Università degli Studi di Trieste

DIPARTIMENTO DI FISICA

Corso di Laurea Magistrale Interateneo in Fisica

Tesi di Laurea Magistrale

Analisi di descrittori meteorologici e di attività solare per lo studio delle relazioni Sole-Terra

Analysis of meteorological and solar activity descriptors to study solar-terrestrial relations

Laureando:
Simone Kodermaz

Relatore:
prof. Mauro Messerotti

Correlatore:
prof. Fulvio Stel

Anno Accademico 2011 - 2012

Sommario

L'irradianza solare, cioè il flusso di energia proveniente dal Sole e diretto verso la Terra, è la principale risorsa di energia del nostro pianeta. Questo flusso di energia varia su diverse scale temporali, da pochi minuti durante i brillamenti solari fino a migliaia di anni, a causa delle variazioni del campo magnetico solare. Tutte queste variazioni hanno effetto sulla Terra, modificandone il clima e il tempo meteorologico. Infatti, anche se queste variazioni sono molto inferiori rispetto al valore totale di irradianza, potrebbero giocare un ruolo fondamentale nel bilancio energetico terrestre, in particolar modo se persistono per lunghi periodi. Oltretutto, la variabilità solare non è uniforme su tutto lo spettro, ma è maggiore per piccole lunghezze d'onda.

In secondo luogo, potrebbero essere presenti meccanismi di *feed-back* che amplificherebbero l'effetto della variabilità solare sul clima terrestre. Uno dei piú studiati in questo senso è la modulazione del flusso di raggi cosmici galattici da parte del vento solare. Il flusso di raggi cosmici è collegato alla ionizzazione dell'aria e alla formazione di nuclei di condensazione, fondamentali per la formazione nuvolosa. La presenza di nuvole influenza il budget energetico dell'atmosfera terrestre modificando l'albedo e rilasciando calore latente durante la condensazione.

L'obiettivo di questo lavoro è lo studio di metodologie per la rilevazione di segnali solari nei parametri meteorologici e climatici, e più precisamente, nelle serie temporali di temperatura e livello colonnare di ozono. Le metodologie utilizzate sono state semplici test di correlazione e analisi di Fourier, analisi mediante trasformate *wavelet*, ed infine analisi mediante *Empirical Mode Decomposition* (EMD) e successiva analisi dello spettro di Hilbert.

I test di correlazione non hanno portato a conclusioni rilevanti. L'analisi di Fourier nel dominio delle frequenze ha segnalato la presenza di componenti in frequenza associabili ad attività solare, ma anche questo risultato non è molto significativo.

L'analisi di Fourier, utile sotto molti punti di vista, possiede però lo svantaggio di perdere l'informazione temporale del segnale analizzato. Allo scopo di risolvere questo inconveniente è stata utilizzata l'analisi wavelet, che

consente di analizzare un segnale nel dominio tempo-frequenza. Questo tipo di analisi consente una migliore identificazione delle componenti di frequenza nel tempo, ma non fornisce dei risultati quantitativi.

L'analisi multiscala del segnale attraverso l'algoritmo dell'*Empirical Mode Decomposition*, consente di scomporre una serie temporale in diversi 'modi intrinseci' (*Intrinsc Mode Functions*), descrivibili come funzioni oscillatorie a diverse scale temporali. La trasformata di Hilbert consente di trasformare gli IMFs in segnali complessi, definendo quindi una fase variabile nel tempo. La derivata di questa quantità rappresenta la frequenza istantanea. Lo spettro di Hilbert diventa quindi la rappresentazione del segnale di partenza nel dominio tempo-frequenza.

L'utilizzo di questa metodologia ha permesso di trovare una componente della serie temporale di temperatura, che potrebbe quantificare il singolo contributo della variabilità solare. Inoltre, l'EMD risulta un metodo alquanto efficace per estrarre un *trend* da una serie temporale, senza dover fare assunzioni arbitrarie, come accade per esempio con una media mobile o un'analisi multiscala tradizionale.

I possibili sviluppi in un futuro lavoro potrebbero essere l'applicazione dell'EMD ad altri parametri meteorologici e climatici su scala locale e globale, o la combinazione dell'EMD con l'analisi wavelet.

Contents

Sommario **1**

Introduction **13**

1 Sun and Earth Climate **15**
Solar Variability 15
The Earth Climate System 18

2 Analysis of Observational and Model-generated Data **28**
2.1 Data Sets 28
2.1.1 Solar Irradiance Platform 28
2.1.2 Solar Irradiance data 29
2.1.3 Meteorological data 37
2.1.4 The R software 37
2.2 Simple correlations and Fourier analysis 41
2.2.1 Correlations 41
2.2.2 Fourier Analysis 43
2.3 Wavelet Analysis 53
2.3.1 Introduction to Wavelet analysis 53
2.3.2 Wavelet analysis of Solar Irradiance data 61
2.3.3 Wavelet analysis of meteorological data 71
2.4 Empirical Mode Decomposition 84
2.4.1 Introduction to EMD 84
2.4.2 Application of EMD to Irradiance data 97
2.4.3 Application of EMD to meteorological data 101

3 Discussion and Conclusions **109**

Acknowledgements **112**

Bibliography **113**

Appendices **117**

A Solar Irradiance modelling by SOLAR2000 **118**

B Application of EMD to Wavelet Analysis **121**

Glossary **126**

List of Figures

1.1 The most important solar activity indicators: Solar irradiance (panel 1), Neutron count rate (panel 2), aa-index (panel 3), aurorae (panel 4), sunspot number (panel 5) and Be-10 measurements (panel 6). Source: [9]. 17
1.2 Globally averaged energy budget of the Earth atmosphere. Source: [14]. 19
1.3 Global annual mean temperature vs. height and atmospheric subdivision. Source: 2010 Pearson Education, Inc. 20
1.4 Schematic of mean meridional tropospheric, stratospheric and mesospheric circulation at the solstice. Source: World Meteorological Organization - WMO, 1985. 22
1.5 Vertical profiles of mixing ratios of the principal constituents of atmosphere. CFC-11 and CFC-12 are chlorofluorocarbons. Source: [6]. 24

2.1 Reference Solar Spectrum. Source: [31]. 29
2.2 Proxies timeline. Source: [31]. 30
2.3 Normalized variation in time of XUV (purple), VUV (red), MUV and NUV (both orange). 33
2.4 Normalized variation in time of XUV (purple), EUV (blue), Ly-α (cyan), FUV , MUV and NUV (all three green). 34
2.5 Normalized variation in time of UV-C (red), UV-B and UV-A (both blue). 35
2.6 Variation in XUV band. The straight black line is the mean value of the band; the yellow and blue line shows the interval of variation due to the change of eccentricity at present time. The red and cyan lines represent the same at Milankovitch maximum. 38

2.7 Variation in VUV band. The straight black line is the mean value of the band; the yellow and blue line shows the interval of variation due to the change of eccentricity at present time. The red and cyan lines represent the same at Milankovitch maximum. 39
2.8 Atmosphere height in hPa and km. 40
2.9 Scatter plot between XUV band and surface temperature anomaly ($°C$) 42
2.10 Value of Student's t vs. time delay. The two series are XUV band irradiance and surface temperature. The time sampling is seasonal (4 data for year). 44
2.11 p-value vs. time delay. The two series are XUV band irradiance and surface temperature. The time sampling is seasonal (4 data per year). 45
2.12 Periodogram showing the spectrum of temperature at surface. The green line indicates annual period, the red line indicates 11-years cycle and the blue line indicates 22-years cycle. 47
2.13 Periodogram showing the spectrum of temperature between 850 and 300 hPa. The green line indicates annual period, the red line indicates 11-years cycle and the blue line indicates 22-years cycle. 48
2.14 Periodogram showing the spectrum of temperature between 300 and 100 hPa. The green line indicates annual period, the red line indicates 11-years cycle and the blue line indicates 22-years cycle. 49
2.15 Periodogram showing the spectrum of temperature between 100 and 50 hPa. The green line indicates annual period, the red line indicates 11-years cycle and the blue line indicates 22-years cycle. 50
2.16 Periodogram of ozone level at 60°N. The green line indicates annual cycle and the red line indicates 11-year cycle. There is no evidence of the latter periodicity. 51
2.17 Periodogram of ozone level at 60°N. The green line indicates annual cycle and the red line indicates 11-year cycle. There is possible evidence of the latter periodicity. 52
2.18 Real (black) and Imaginary (green) part of the Morlet function in the time domain (a) and in the frequency domain (b). . . . 59
2.19 Real (black) and Imaginary (green) part of the Paul function in the time domain (a) and in the frequency domain (b). In this case, $m = 4$ 60

2.20 DOG function in the time domain (a) and in the frequency domain (b). In this case, $m = 4$ 62
2.21 Wavelet spectrum of the Morlet wavelet with m=6, and 24. . . 63
2.22 Wavelet spectrum of the Paul wavelet with m=4 and 40. . . . 63
2.23 Wavelet spectrum of the DOG wavelet with m=2 and 80. . . . 64
2.24 Wavelet spectrum of XUV band with Morlet wavelet of order 6. The green line represents the period of 11 years, the purple line indicates the period of 1 year, and the red line indicates the period of 27 days. 65
2.25 Wavelet spectrum of total solar irradiance with Morlet wavelet of order 6. The green line represents the period of 11 years, the purple line indicates the period of 1 year, and the red line indicates the period of 27 days. 66
2.26 Wavelet coherence spectrum of XUV band and TSI. 67
2.27 Wavelet spectrum of XUV band with a Paul of order 4. The green line represents the period of 11 years, the purple line indicates the period of 1 year, and the red line indicates the period of 27 days. 68
2.28 Wavelet spectrum of XUV band with a DOG of order 2. The green line represents the period of 11 years, the purple line indicates the period of 1 year, and the red line indicates the period of 27 days. 69
2.29 Wavelet spectrum of XUV band with a DOG of order 20. The green line represents the period of 11 years, the purple line indicates the period of 1 year, and the red line indicates the period of 27 days. 70
2.30 Wavelet spectrum of temperature at surface (Morlet of order 6). The green line represents the period of 11 years and the purple line indicates the period of 1 year. 72
2.31 Wavelet spectrum of temperature between 850 and 300 hPa (Morlet of order 6). The green line represents the period of 11 years and the purple line indicates the period of 1 year. 73
2.32 Wavelet spectrum of temperature between 300 and 100 hPa (Morlet of order 6). The green line represents the period of 11 years and the purple line indicates the period of 1 year. 74
2.33 Wavelet spectrum of temperature between 100 and 50 hPa (Morlet of order 6). The green line represents the period of 11 years and the purple line indicates the period of 1 year. 75

2.34 Wavelet spectrum of temperature between 850 and 300 hPa (DOG of order 10). The cyan line indicates the period o f 22 years, the green line represents the period of 11 years and the purple line indicates the period of 1 year. 76
2.35 Wavelet spectrum of temperature between 300 and 100 hPa (DOG of order 10). The cyan line indicates the period o f 22 years, the green line represents the period of 11 years and the purple line indicates the period of 1 year. 77
2.36 Wavelet spectrum of daily global ozone level (Morlet of order 6). The missing data are substituted by the mean value of the time series. 79
2.37 Wavelet spectrum of daily global ozone level (Morlet of order 6). The missing data are substituted by a linear trend (2.37(a)) or by a gaussian white noise (2.37(b)). 80
2.38 Wavelet spectrum of daily ozone level in the latitude interval between $60°N$ and $55°N$ (Morlet of order 6). The missing data are substituted by the mean value of the time series. The purple line indicates the annual period. 81
2.39 Wavelet spectrum of daily ozone level in the latitude interval between $5°N$ and $25°S$ (Morlet of order 6). The missing data are substituted by the mean value of the time series. The purple line indicates the annual period. 82
2.40 Wavelet spectrum of daily ozone level in the latitude interval between $0°N$ and $5°S$ (Morlet of order 6). The missing data are substituted by the mean value of the time series. The purple line indicates the annual period, the blue one a period of about 2.5 years and the orange line indicates a period of about 0.5 years. 83
2.41 Red circles are maxima and blue circles are minima, localized by EMD algorithm. In these examples the signal is a sum of three pure sines: $x(t) = 0.7sin(\pi t) + 0.5sin((2\pi t)/0.7) + 0.3sin((2\pi t)/0.3)$. 86
2.42 Maxima and minima are interpolated by a cubic spline. Red circles are maxima, blue circles are minima. The red line is the upper envelope and the blue line is the lower envelope. . . 87
2.43 The average of the upper and lower envelopes is computed. The purple line represents this average vs. t. 88
2.44 Subtracting the mean from the original signal $x(t)$, the latter (upper plot) is decomposed into a candidate from the first IMF (middle plot) and a residue which contains all the components at lower frequencies (lower plot). 89

2.45 Empirical Mode Decomposition. 90
2.46 Superimposition of the three IMFs and the residue. Blue line: first IMF; red line: second IMF; black line: third IMF; purple line: residue. Straight lines represent the amplitude of the real pure sines used to composed the original signal $x(t)$. 94
2.47 Superimposition of the three orthogonal IMFs and the residue. Blue line: first orthogonal IMF; red line: second orthogonal IMF; black line: third orthogonal IMF; purple line: residue. Straight lines represent the amplitude of the real pure sines used to composed the original signal $x(t)$. 95
2.48 A sinusoidal signal $x(t) = A(t)sin(\omega(t)t)$ with $A(t) = 0.8+\sqrt{t}$ and $\omega(t) = 2\pi(2 + 0.5t)$. 97
2.49 Hilbert Spectrum of signal $x(t)$. The frequency is presented as linear function of time and amplitude is presented here as intensities of black. 98
2.50 Amplitude vs. time obtained by Hilbert analysis (black line). The correct amplitude function is plotted in red and shifted downwards by one unit. 99
2.51 Empirical mode decomposition of XUV solar irradiance. Time sampling: monthly. 100
2.52 Hilbert spectrum of temperature between 850 and 300 hPa. the red line indicates the period at 11 years. The mode that follows the red line indicates that its frequency remains quasi-constant. 102
2.53 XUV irradiance (red line) and 5^{th} IMF of temperature between 850 and 300 hPa (blue line). 103
2.54 4^{th} IMF of XUV irradiance (red line) and 5^{th} IMF of temperature between 850 and 300 hPa (blue line). 104
2.55 Differences in time between solar and temperature extrema: a positive difference indicates a foregoing solar extremum whereas a negative differences indicates a foregoing temperature extremum. 105
2.56 Energy of the components of temperature at 850-300 hPa. . . 106
2.57 XUV irradiance residue (red) and global ozone level residue (green). 107

B.1 (a) A trial signal composed by a pure sine of small amplitude and a linear trend. (b) Wavelet spectrum computed with a Morlet wavelet of order 6 of a trial signal composed by a pure sine of small amplitude and a linear trend. 122

B.2 Wavelet power spectrum of temperature at 850-300 hPa without detrending by EMD. 123
B.3 Wavelet power spectrum of temperature at 850-300 hPa with detrending by EMD. The periodicity around 11 year is more evident. 124

List of Tables

1.1 Chemical composition of Earth atmosphere. Source: NASA . . 23

2.1 Table reporting the main spectral band adopted by the International Standard Organization. Notice that some bands overlap each other. (*): in our work, XUV band refers to a range between 1 and 10 nm. 31
2.2 Summary table with percentual variation for each band. . . . 32
2.3 Mean value of each wavelength-band compared to total solar irradiance in percentage. 36
2.4 Annual and seasonal global temperature deviations for the surface, troposphere, and lower stratosphere derived from a 54-Station Radiosonde Network, from year 1958 to 2010. Source: J. K. Angell (Air Resources Laboratory - National Oceanic and Atmospheric Administration) 40
2.5 Stratospheric Column Ozone. Source: Nimbus 7 TOMS (January 1979 - April 1993) and Earth Probe TOMS (August 1996 - December 2005) satellite measurements. 41
2.6 Three different types of mother wavelet: Morlet, Paul and derivative of gaussian (DOG). 56

A.1 SOLAR2000 modelling setup phases. Source: [31]. 119
A.2 Summary of solar irradiance proxies provided by SOLAR2000. Source: [30]. 120

Introduction

The Sun provides a significant energy input to the Earth atmosphere. Such an energy input varies by different amounts on various time scales according to specific wavelength bands.

The Earth's climate and weather are interested by the solar output, which impacts relevant atmospheric processes. The solar energy flux variations are small compared to the total amount of the incoming energy, but they could play an important role in the global energetic balance and in the terrestrial climate, especially if they persist for several decades. In fact, there are evidences that the Sun had been the major driver of Earth's climate and paleoclimate in the past. However, there is no significant evidence that solar output and variability have contributed in a significant way to the increase in global mean temperature in the past 50 years [10].

A key area of inquiry deals with defining a unified record of solar output from present to pre-industrial era. In order to do this, attention is paid to the study of the links between indices of solar activity and solar irradiance. Increasing our knowledge of these aspects, it would become possible to give a better reconstruction of solar irradiance variations in the past.

Concerning the analysis of the direct mechanisms of solar influence on terrestrial climate, there are two main possibilities. The *top-down* model expects that solar variability affects the climate first perturbing stratosphere and then forcing lower atmospheric levels. In the *bottom-up* view, the interaction between solar energy and terrestrial surface leads to changes in dynamics and thermodynamics from lower to upper altitudes [10].

Furthermore, there are also feed-back mechanisms, that could play a significant role. One of the most studied effects is the modulation of Galactic Cosmic Rays (GCR) flux due to the variations of solar activity. During solar maxima it is found a significant reduction of GCR flux, whereas during solar minima an increase of GCR flux is observed [17]. The modulation of GCR could affect Earth's climate through the ionization of atmosphere, that leads to the production of cloud condensation nuclei (CCN) and the subsequent cloud formation. Clouds are connected with local and global climate, since

their condensation releases latent heat, and their presence changes terrestrial albedo.

Anyway, due to the complex nonlinear coupling of the Sun's emission with the atmosphere system, it is challenging to clearly identify and quantify the solar contribution to climate and weather.

A possible approach consists in a cross analysis of solar activity descriptor and climate descriptor time series, in the attempt to point out the solar influence. To achieve this goal, in this work we used observational and model-generated data sets. In particular, solar irradiance data and meteorological data. Therefore, important topics such as study of the causes of solar variability, paleoclimate indicators and GCR modulation, which are fundamental anyway, have not been taken in account.

The dissertation is organized as follows: A theoretical introduction to Sun variability proxies and to the Earth climate system is illustrated in Chapter 1. The meteorological parameters used in this work and the methodologies used to find solar activity signals are exhibited in Chapter 2. Specifically, Section 2.1 illustrates the sets of data utilized in this dissertation, Section 2.2 presents the tests of correlation and their results, Section 2.3 introduces wavelet analysis and demonstrates the results of its application to the temperature and ozone time series. Section 2.4 introduces Empirical Mode Decomposition (EMD) and reports the results of its application. The final discussions and the conclusions about the main results are exposed in Chapter 3. Some information on the semi-empirical model SOLAR2000 is reported in Appendix A, while a possible application of EMD to wavelet analysis is introduced in Appendix B.

The correlation tests have not led to any significant conclusion, while Fourier analysis has revealed frequency components that could be related with the solar activity. However, doing a Fourier analysis, the temporal information is lost. This limitation has been overcome by wavelet analysis, which has provided some qualitative results of significance, although not quantitative. Finally, the multi-scale analysis performed with Empirical Mode Decomposition has proved to be an efficient methodology to decompose a signal, to achieve a time-frequency analysis through Hilbert Transform, and it has led to some quantitative results of significance, which would prompt some possible future developments.

Chapter 1

Sun and Earth Climate

Solar Variability

The two main carriers through which the Sun transports energy to our planet are, respectively, electromagnetic radiation and energetic particle radiation. Both these carriers influence Earth's system and determine its climate together with other factors, such as volcanoes, human activities and variations of terrestrial orbit dynamics. Moreover, the heliosphere, i.e. the region of space surrounding the solar system dominated by charged particles transported by the solar wind, screens solar system from energetic particles coming from outer space. The variations of these inputs induce changes in the physical state of atmosphere, and the major driver of these solar variations is the solar magnetic field.

Detection techniques and methodologies focused to understand the various aspects of the solar activity, allow us to evaluate the effects of all the solar inputs on climate and hence allow some prediction of the future evolution of Earth's climate.

Indicators of Solar Activity

In order to understand in depth the impact of the Sun on terrestrial climate, it is necessary to recover historical data of solar activity and quantify how our climate reacts to the change in the bulk amount and frequency-composition of the incoming energy from Sun. There are a number of indicators studied in the years by scientists which differ not only in the typology of information provided but also in their characteristic time-scale [9]. The most widely used solar activity indicators analysed in climatology are reported in Figure 1.1. Some of them, like e.g. sunspot number and 10.7 cm radio flux, are indices derived from solar activity observables, whereas others, like e.g. aurorae,

geomagnetic indices, are relevant to planetary perturbations caused by solar activity, i.e., they are indirect indicators, called 'proxies'.

The number of **sunspots** (Figure 1.1, panel 5) has been being recorded accurately for years. It has a quasi periodic cycle of about 9-13 years, known as the 11-year solar cycle. The availability of sunspot number data data covers an interval from the beginning of the 17^{th} century up to now. The data exhibit a period during which no sunspots where observed, that lasted for the second half of the 17^{th} century. It has been called the *Maunder Minimum* and it corresponded to a period of lower than average temperatures recorded in most of the Northern Hemisphere, called *the Little Ice-Age*. There are also two important trends to point out: a low activity period around the beginning of the 19^{th} century and a relative increase of activity in the last 50 years.

A solar activity proxy is the number of **aurorae** (Figure 1.1, panel 4), which are due to the interaction between Earth atmosphere and the charged particles of the solar wind, whose trajectory is deviated by geomagnetic field and focused polewards. Aurorae are also associated with an electric current in the upper atmosphere which causes disturbances in the Earth's magnetic field. This effect is quantified by the so-called **aa-index** [9] (Figure 1.1, panel 3).

The solar wind screens Earth from galactic cosmic rays, and so there is an inverse relation between solar activity and **GCR flux**. The latter is quantified by the number of observed neutrons produced by the impact of a GCR with the atmosphere. These data extend back to around 1950 (Figure 1.1, panel 2).

Another possible indicator of GCR flux, that provides long term data is based on the measurements of **cosmogenic isotopes** such as Beryllium-10 (^{10}Be, Figure 1.1, panel 6) and Carbon-14 (^{14}C). These measurements are deduced from deposits on surface features, for instance ice cores and tree rings. Although these indicators are useful since they cover centuries, their main disadvantage is that they contains information on other chemical and geophysical processes that influence their deposition, then further investigations would be necessary. However, the various processes influencing each nuclide are very different so that by combining them it is possible to separate these effects and thus to reconstruct a reliable solar activity index [4].

The **solar radio flux at 10.7 cm** is commonly used as solar activity indicator. Reliable measurements are available from 1950. Such radio emission is associated with the presence of active regions in the photosphere.

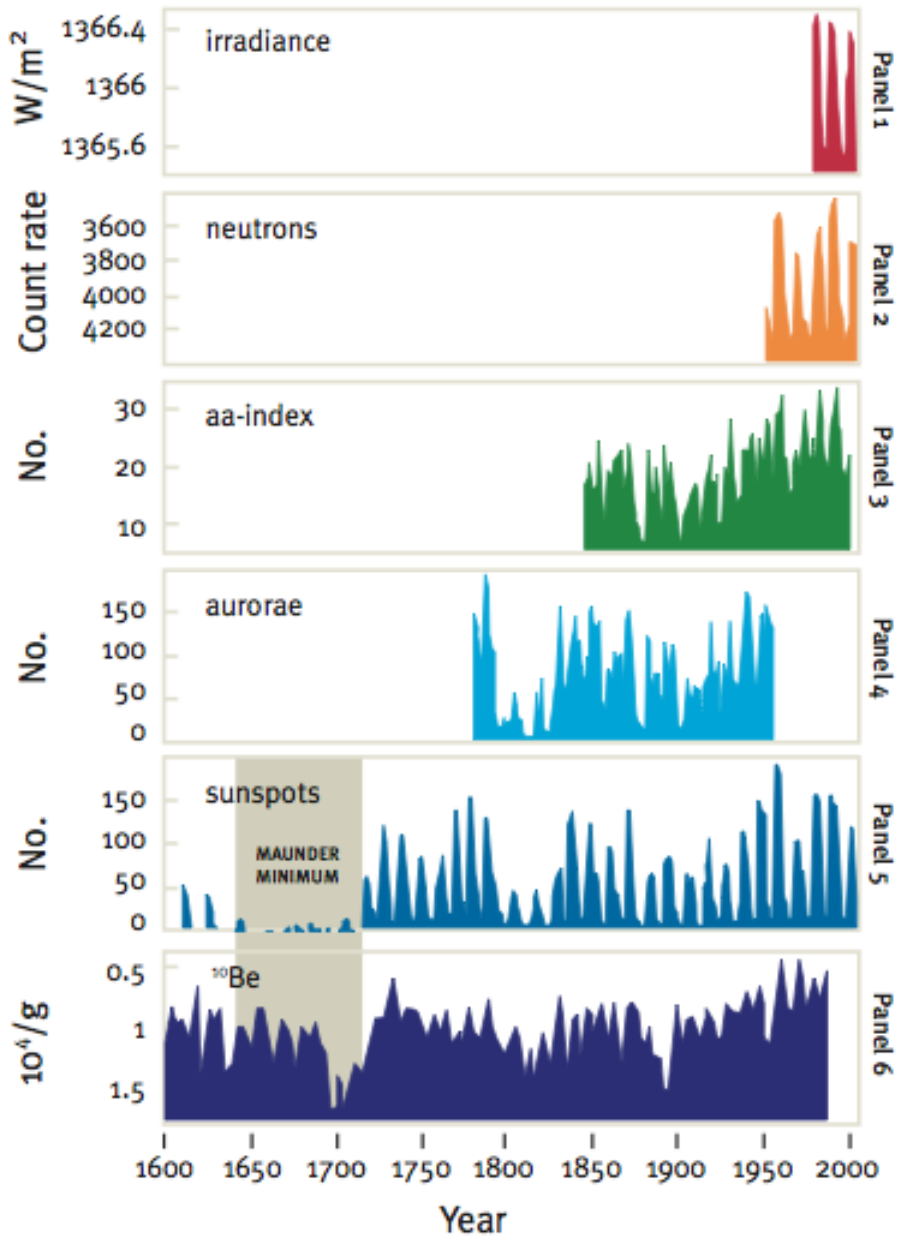


FIGURE 1.1: The most important solar activity indicators: Solar irradiance (panel 1), Neutron count rate (panel 2), aa-index (panel 3), aurorae (panel 4), sunspot number (panel 5) and Be-10 measurements (panel 6). Source: [9].

Solar Radiation

The solar indices and proxies presented in the previous paragraph are useful to describe the evolution of solar activity, since they are associated with different phenomena, but they do not provide any clear indication of the physical processes whereby the Sun influences climate.

The primary physical quantity that could be considered is the total solar irradiance (TSI). However, it does not provide the required accuracy, since the measurements of TSI are subject to fluctuations of atmospheric absorption. Direct measurements from satellites are available since 1978.

The Earth Climate System

The Earth's atmosphere is a layer of gases and water vapour that surrounds the terrestrial surface and is connected with the interplanetary environment. In a first approximation, its physical state is determined by the energy incoming from the Sun, and to a much smaller extent from the heat coming from the interior of the Earth. It is important to point out that the physical state of the terrestrial atmosphere may be perturbed by external phenomena - radiation outbursts, solar wind shocks, cosmic rays - as well as internal ones - such as volcanic eruptions or interactions with the biosphere.

However, it is assumed that the Earth's mean temperature is determined by the absorption of the solar radiation budget, while the latitudinal distribution of absorbed energy is the primary driver for atmospheric circulation. The radiative heating effect is the net result of solar heating and infrared cooling, the latter being related to chemical composition of atmosphere and thermal structure (Figure 1.2).

Atmospheric Structure and Thermal Profile

The atmosphere can be schematically divided into layers according to the vertical profile of temperature, which is determined by physical, chemical and dynamical processes [10] (Figure 1.3).

The lowest layer of the atmosphere is the *troposphere*. Here the temperature decreases with height, and the energy transmission mechanism in this layer is convection. This is due to the air density and to the absorption by atmosphere of infrared radiation coming from the Earth. The upper limit of the troposphere is the *tropopause* at about 12 km of altitude. Actually, the height of the tropopause is not constant and varies with latitude: it spans from about 8 km at the poles to 18 km above the equator. It is important to underline, that most of water vapour lies in the troposphere.

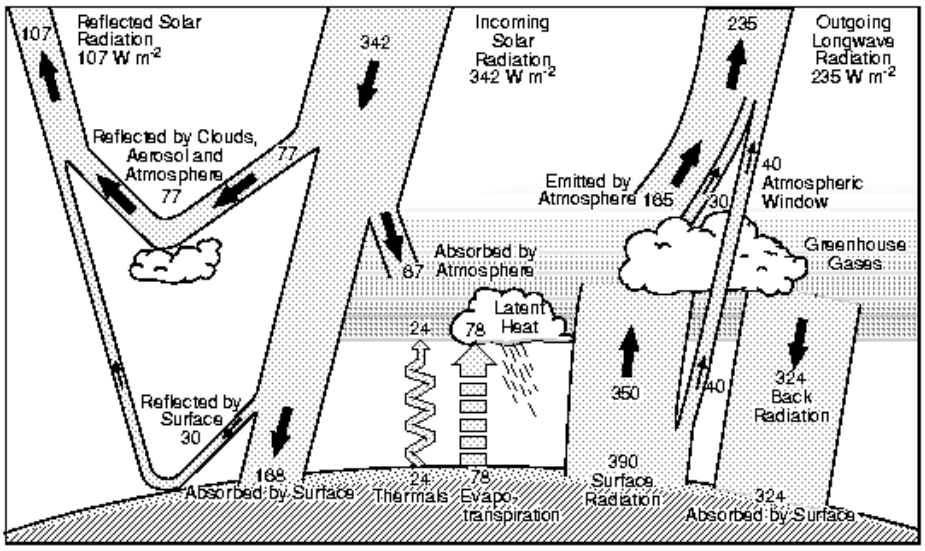


FIGURE 1.2: Globally averaged energy budget of the Earth atmosphere. Source: [14].

Above the tropopause lies the *stratosphere*, which is a layer where the temperature increases with height. Here, there is no vertical mixing and the main way to transport heat are radiative processes. The stratosphere extends up to about 50 km, where the vertical thermal gradient vanishes, and it is bounded by the *stratopause*. Chemically, the stratosphere is characterized by the *ozonosphere*, an ozone layer which determines the vertical profile of the temperature.

From 50 km up to about 90 km lies the *mesosphere* where the temperature decreases again with height. In the mesosphere, the most effective process to transport heat is conduction. Above the mesosphere there is the *mesopause*, which separates the lower layer from the upper one: the *termosphere*, where temperature increases with height again.

The structure shown in Figure 1.3 is subject to latitudinal variations due to the different solar exposition. At the surface the air is warmer near the equator than at the poles, but at the tropopause the reverse is observed. At the stratopause, the temperature decreases monotonically from summer to winter one, and the opposite happens at the mesopause.

Zonal Winds and Mean Meridional Circulation

The planetary mean surface temperature gradient determines the large scale wind circulation [10]. In the troposphere there are two sub-tropical westerly

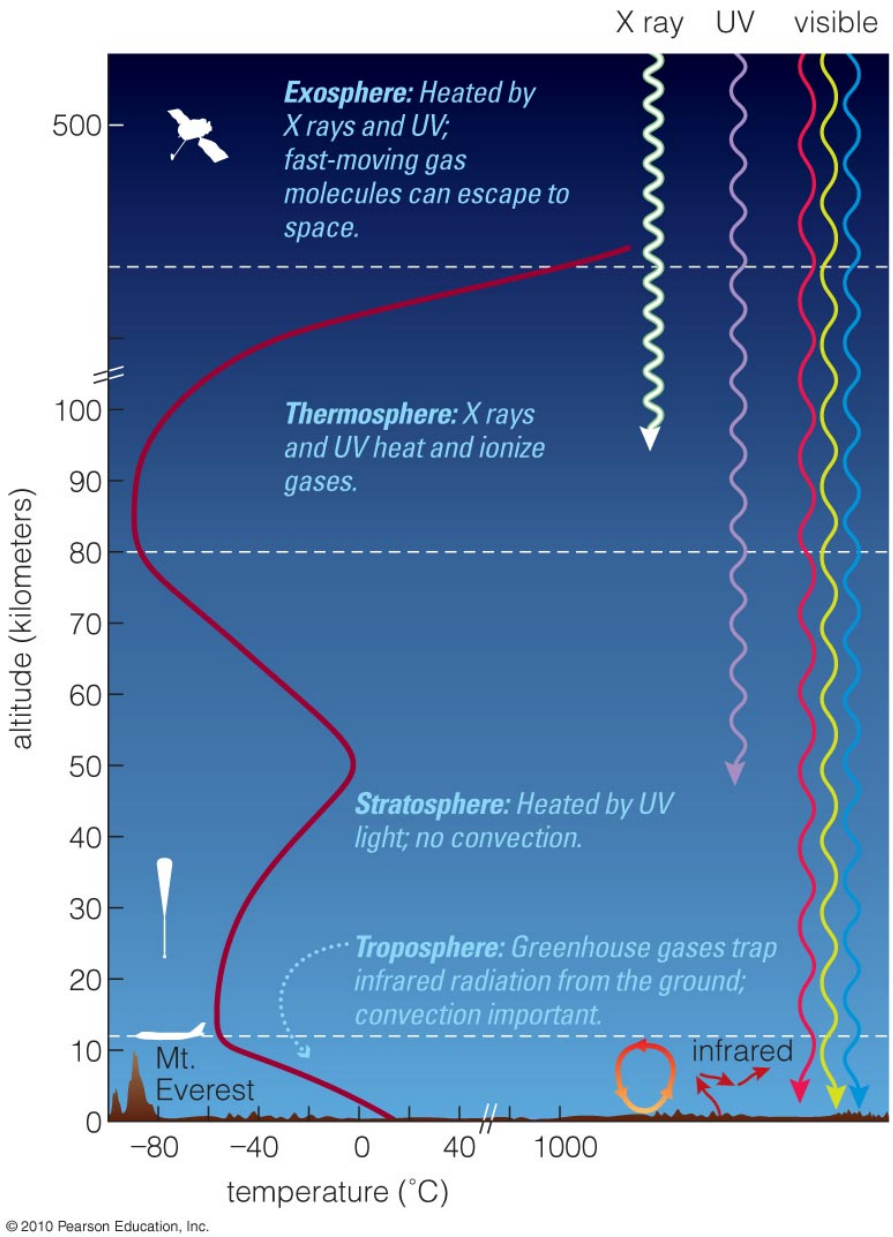


FIGURE 1.3: Global annual mean temperature vs. height and atmospheric subdivision. Source: 2010 Pearson Education, Inc.

jets which are produced by the baroclinic instability, that is the condition when pressure and temperature gradients are not parallel. The mean meridional circulation of the troposphere is dominated by the Hadley cell: air rises from the equator, flows towards the poles, sinks in the sub-tropics and flows back towards equator near the surface with an eastward component due to Coriolis acceleration (trade-winds). In mid-latitudes are active the Ferrel cells which circulates with opposite horizontal vorticity respect to the Hadley ones.

The lower stratosphere is characterized by the Brewer-Dobson circulation, that transports air masses from tropics towards the poles.

In the upper stratosphere - at about 30 hPa - there is a westerly jet in the winter period and an easterly jet in the summer period.

In the mesosphere there is a circulation directed from the hemisphere in the warm season, towards the hemisphere in the cold season. A schematic representation of the mean zonal winds is reported in Figure 1.4.

Chemical Composition

The lower-middle atmosphere is almost well-mixed due to turbulence, that mixes up at shorter timescales than molecular diffusion. In the upper atmosphere (above approximately 100 km) the air is so rarefied that diffusion is much more effective [10]. At these altitudes the free electrons produced by solar ionization have long lifetimes; the *ionosphere* is the layer where free charges are present and it plays a key role in determining the Earth's electric field.

In the atmosphere is also possible to identify the *homosphere* which is the lower atmosphere where the chemical components are well mixed and the concentration of the main constituents is almost homogeneous, and, above it, the *heterosphere* where there is a chemical stratification of the constituents.

The chemical composition of Earth's atmosphere is reported in Table 1.1. Molecular nitrogen exhibits the highest concentration, but it has a marginal role in radiative and chemical processes. The second most relevant constituent is molecular oxygen and it is fundamental in the atmospheric processes, since in the middle atmosphere it is photodissociated into atomic oxygen. Other components such as carbon dioxide and methane are not important in terms of concentration, but they are essential for the greenhouse effect. There are other two important constituents that are not included in Table 1.1, but are crucial for the physical and chemical state of the Earth's atmosphere, i.e. water vapour and ozone. The former is significant since its condensation releases latent heat, while the latter causes the heating of the stratosphere and shields the biosphere from UV-radiation.

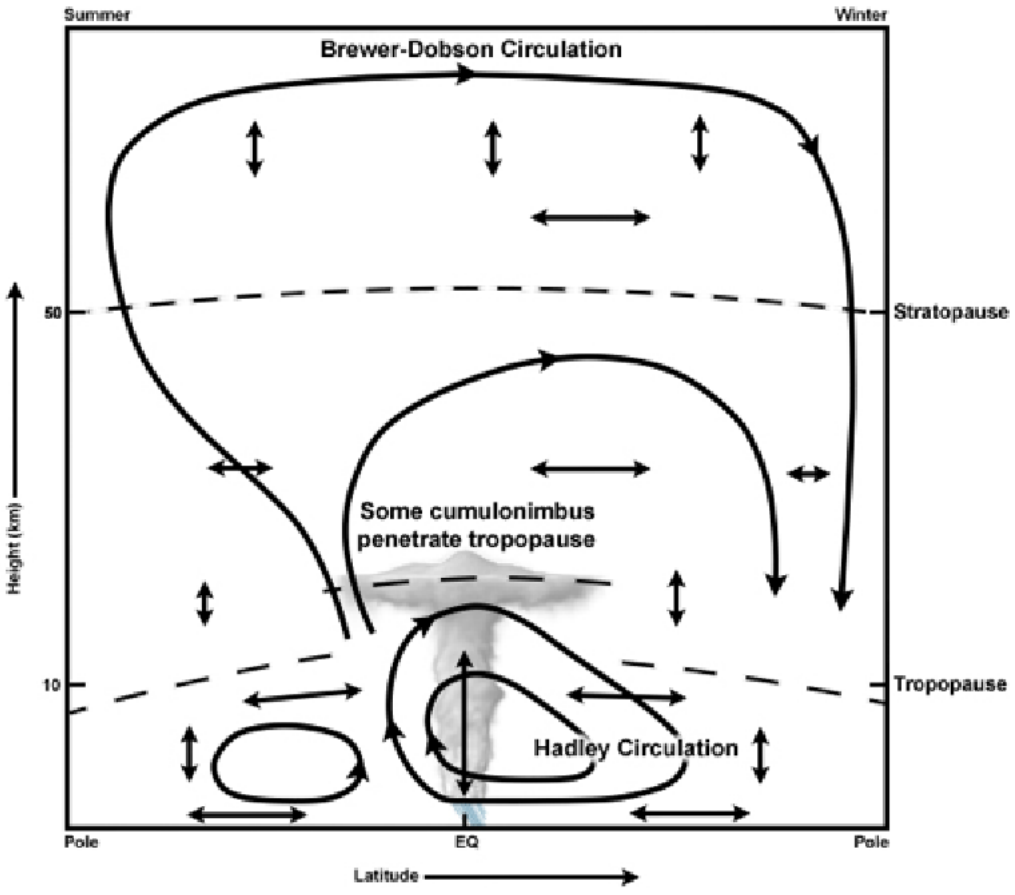


FIGURE 1.4: Schematic of mean meridional tropospheric, stratospheric and mesospheric circulation at the solstice. Source: World Meteorological Organization - WMO, 1985.

Gas	Volume
Nitrogen (N_2)	780,840 ppm (78.08 %)
Oxygen (O_2)	209,460 ppm (20.09 %)
Argon (Ar)	9,340 ppm (0.934 %)
Carbon dioxide (CO_2)	381 ppm
Neon (Ne)	18.18 ppm
Helium (He)	5.24 ppm
Methane (CH_4)	1.745 ppm
Krypton (Kr)	1.14 ppm
Hydrogen (H_2)	0.55 ppm
Water vapour (H_2O)	0 - 40 ppm
Ozone (O_3)	0 - 12 ppm

TABLE 1.1: Chemical composition of Earth atmosphere. Source: NASA

In Figure 1.5 the vertical concentration profiles of the main constituents of atmosphere are reported. In the graph, CFC-11 and CFC-12 are chlorofluorocarbons: these molecules had drawn considerable interest among scientists in the 70s because their presence in the upper troposphere had been upsetting the ozone equilibrium damaging the ozonosphere.

Incidentally, the term *dry atmosphere* refers to the atmosphere composed by all its chemical constituents, except water vapour. It has a peculiar behaviour since Earth's temperature is below the critical temperature of water: water vapour can condense with the subsequent release of latent heat.

Clouds and Precipitations

Clouds play an important role in the radiation, heating and water vapour budgets of the Earth [10]. Cloud distribution is heterogeneous, and generally it is concentrated along the equator and in the mid-latitude belts. They act transporting latent heat from oceans to the atmosphere, reflecting solar radiation back to space and trapping infrared radiation like greenhouse gases do.

A complete physical description of cloud properties and their effect on the Earth's atmosphere would consider:

- relation with Earth radiative budget,
- microphysical and optical properties,
- formation processes,

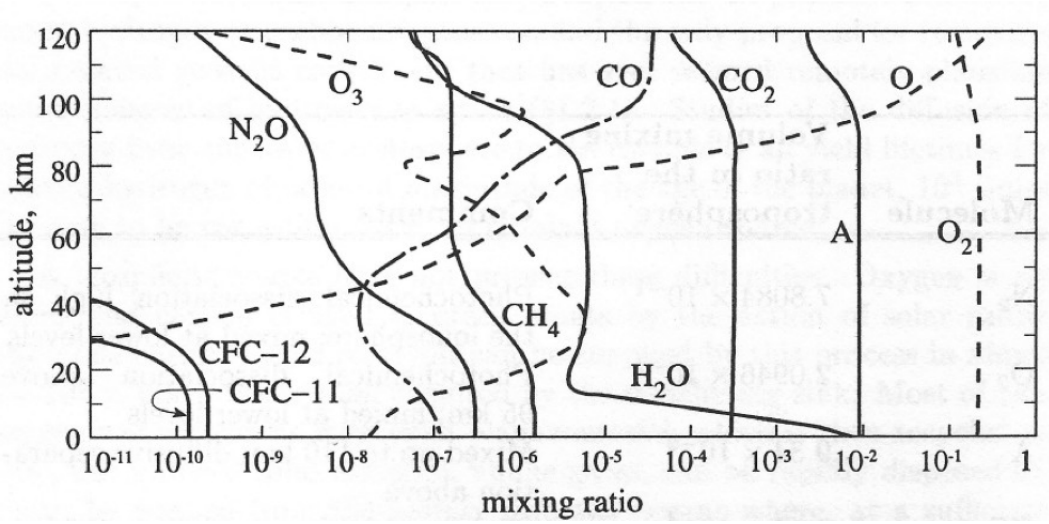


FIGURE 1.5: Vertical profiles of mixing ratios of the principal constituents of atmosphere. CFC-11 and CFC-12 are chlorofluorocarbons. Source: [6].

- connections with the galactic cosmic ray flux.

However, all these important aspects are beyond the scope of this thesis and are not developed in this work.

Weather and Climate

Dealing with the evolution in time of meteorological parameters, such as temperature, pressure or precipitations, it is important to understand the difference between *weather* and *climate*. The weather represents the atmospheric conditions of a particular place at a particular time, while the climate of a geographical region represents the long-term averages of the atmospheric conditions.

Modes of Variability

The Earth's atmosphere exhibits a number of characteristic modes of variability, which are used in describing the climate in geographical regions. In this work, only the most important modes are considered as follows [10].

- **NAO**: *North Atlantic Oscillation* (NAO) is a pattern of atmospheric circulation in the north of the Atlantic Ocean. It is defined as the difference between pressure at the sea-level of Iceland and pressure over the Azores at the sea-level. During positive phases of NAO (NAO+)

the pressure horizontal gradient is stronger and westerly wind are enhanced.

- **ENSO**: *El Niño - Southern Oscillation* is an atmospheric teleconnection that couples ocean and atmosphere. The term *teleconnection* indicates a realtion between climate anomalies at large distance. The oceanic component of ENSO is determined by the heating (*El Niño*) or the cooling (*La Niña*) of the middle-eastern Pacific Ocean. Whereas, its atmospheric component is determined by the *Southern Oscillation*, the periodic change in pressure on the middle-eastern Pacific Ocean. These two components are coupled and ENSO has local effects as well as global impacts.

- **Southern Annular Mode** (SAM) is a variation in the mid- to high-latitude difference in surface pressure is associated with southern polar temperatures and the strength of circumpolar winds [9].

- **QBO**: the *Quasi Biennal Oscillation* is a quasi-periodic oscillation between westerly and easterly winds in the equatorial latitudes in the tropical stratosphere. The period is between 28 and 29 months. Its main effects are the mixing of the stratospheric ozone and the modulation of monsoon precipitation and stratospheric circulation.

The Earth's Electric Field and Ionization of the Atmosphere

Electric currents flowing in the atmosphere are mainly due to charge separation produced by thunderstorms: positive charges flow upwards from surface, which is negatively charged, to ionosphere, that is positively charged [10]. Outside thunderstorms there are no sources of charge separation and a small current flows in the opposite sense. In regions of fair weather there is an electric potential at the ground of roughly $150\ V \cdot m^{-1}$, but during thunderstorms it can reach over $10^5 V \cdot m^{-1}$.

Ionization of the lower atmosphere is produced by natural radioactivity and cosmic rays. The latters are modulated by geomagnetism and are influenced also by the heliospheric magnetic field, that shield our planet from galactic cosmic rays. During solar maxima the flux of galactic cosmic rays is reduced, although the modulation primarily affects lower energy cosmic radiation.

A mechanism proposed by Dickinson (1975) and adopted by Marsh & Svensmark (2000) assumes that modulation of GCR by solar variability affects the formation of cloud condensation nuclei (CCN). Ions produced by a cosmic ray become sites for the nucleation of ultra-fine aerosols, but these

particles have to grow enough to become CCN. Growth may occur through condensation of water vapour or through coagulation among neutral and charged particles.

Chapter 2

Analysis of Observational and Model-generated Data

2.1 Data Sets

In this section a short overview on the semi-empirical model and the set of data used in this work is done. In particular three data sets were used:

1. the solar irradiance data set,
2. the temperature at different altitudes data set,
3. the ozone amount at different latitudes data set.

2.1.1 Solar Irradiance Platform

The SOLAR2000 model was developed in the framework of a collaborative project for the characterization of solar irradiance variability across the spectrum [31]. In order to provide a unified solar spectral variability time series, it is necessary to manage sets of data coming from different and heterogeneous surveys. For example, UV rockets date from 1946, and from that time EUV (10-121 nm band) rocket observations have been made and have provided absolute flux estimates for specific solar conditions. EUV satellite observations exist since 1962. These provide insights into wavelength-dependent daily, solar rotational, active region evolution, and solar cycle variations. Moreover, total solar irradiance (TSI) measurements are essential in this reconstruction. In this purpose, the Earth Radiation Budget Experiment (ERB) carried out successful measurements from November 1978 to early 1993, and ACRIM I - an active cavity radiometer irradiance monitor - carried out measurements from February 1980 to June 1989. Currently, there

are instruments that provide TSI data: ERBS and UARS (=ACRIM II), PMO6-V, DIARAD, SOHO (ESA/NASA) and VIRGO.

The 1-nm reference spectrum in Figure 2.1 is a composite spectrum. Reference spectra have been particularly useful for characterizing the highly variable XUV-EUV spectral range (1-121 nm).

The use of SOLAR2000 consists of five phases: (1) model definition and design; (2) modeling of XUV (1-10 nm) and EUV (10-121 nm) irradiances; (3) modeling of FUV (122-200 nm) irradiances; (4) modeling of UV irradiances, and (5) modeling of visible, infrared and TSI irradiances. The subdivision of the spectrum is shown in Table 2.1.

Further information is given in Appendix A.

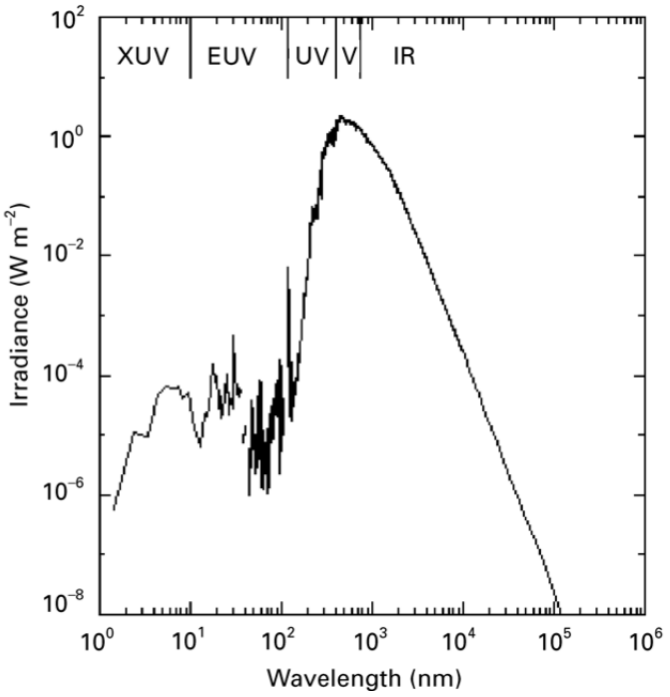


FIGURE 2.1: Reference Solar Spectrum. Source: [31].

2.1.2 Solar Irradiance data

The times series of the solar irradiance data are supplied by the Irradiance Solar Platform (SIP), which contains a semi-empirical model that, starting from real data and solar proxies, provides tables of the solar irradiance level at different wavelengths, in the format and for the duration chosen by the user.

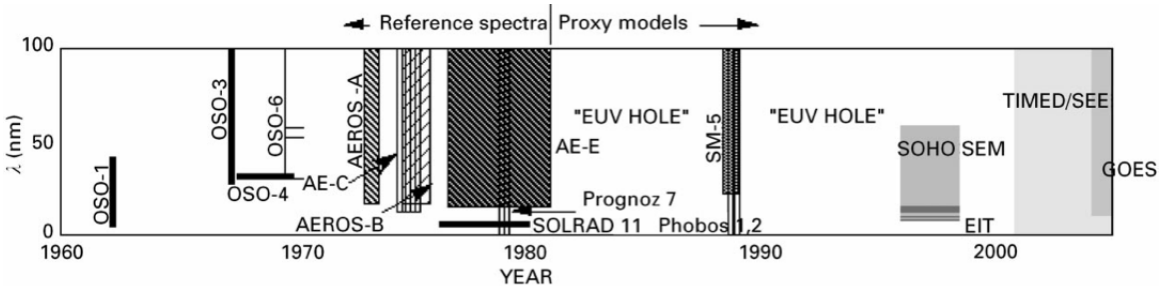


FIGURE 2.2: Proxies timeline. Source: [31].

The maximum time resolution of these time series is 1 day. The maximum available time interval covers a period from the 14^{th} february 1947 until today. The maximum available resolution in wavelength is quite high:

- from 1 to 630 nm: 1 nm
- from 631 to 2500 nm: 2 nm
- from 2500 to 5000 nm: 20 nm
- from 5000 to 10^4 nm: 50 nm
- from 10^4 to $2 \cdot 10^4$ nm: 10^3 nm
- for higher wavelengths there are about fifteen bins with gradually lower wavelength resolution, with increasing wavelength value up to 10^6 nm.
- **total**: 1816 bins

This high resolution, especially at low wavelengths, causes obviously a huge number of bins, and, if one wants to consider every bin at the same time, more than 1800 single time series have to be handled simultaneously. This is quite difficult to perform, and a subdivision of the EM spectrum in a limited number of wavelength-bands has been made. The choice of the subdivision is done according to the International Standard Organization (ISO): a summary is reported in Table 2.1. An important aspect to stress is the fact that some bands overlap each other. For example, Vacuum Ultraviolet (VUV) are comprehensive of Extreme UV (EUV), Lyman-α and Far UV (FUV) ones. Ultraviolet C are comprehensive of part of EUV, Lyman-α, FUV (so the major part of the VUV band) and the more energetic part of Middle UV (MUV). In this work, XUV band refers to photons with wavelength between 1 and 10 nanometers.

Spectral category	Band	Name	Wavelength range (nm)
X-rays	XUV	Soft X-rays	$0.1 \leq \lambda \leq 10$ (*)
UV	UV	Ultraviolet	$100 \leq \lambda \leq 400$
	VUV	Vacuum UV	$10 \leq \lambda \leq 200$
	EUV	Extreme UV	$10 \leq \lambda \leq 121$
	H-Lyman α	H-Lyman α	$121 \leq \lambda \leq 122$
	FUV	Far UV	$122 \leq \lambda \leq 200$
	MUV	Middle UV	$200 \leq \lambda \leq 300$
	NUV	Near UV	$300 \leq \lambda \leq 400$
	UV-C	Ultraviolet C	$100 \leq \lambda \leq 280$
	UV-B	Ultraviolet B	$280 \leq \lambda \leq 315$
	UV-A	Ultraviolet A	$315 \leq \lambda \leq 400$
Visible	VIS	Visible	$380 \leq \lambda \leq 760$
Infrared	IR	Infrared	$760 \leq \lambda \leq 10^6$
	IR-A	Near Infrared	$760 \leq \lambda \leq 1400$
	IR-B	Middle Infrared	$1400 \leq \lambda \leq 3000$
	IR-C	Far Infrared	$3000 \leq \lambda \leq 10^6$

TABLE 2.1: Table reporting the main spectral band adopted by the International Standard Organization. Notice that some bands overlap each other. (*): in our work, XUV band refers to a range between 1 and 10 nm.

After the split of the whole spectrum into bands, an evaluation of the relative time variations of energy is necessary. In order to do this, the time series are firstly normalized at their mean value, and then compared. For clarity, this comparison has been made among three sets of splits of the UV band, since in the splits in Table 2.1 an overlapping between bands exists. The three sets are:

1. XUV, VUV, MUV and NUV;
2. XUV, EUV, Lyman-α, FUV, MUV, NUV;
3. UV-C, UV-B and UV-A.

Figures 2.3, 2.4 and 2.5 depict the time variability of these sets of bands, and Table 2.2 summarises the content of the graphs.

Band	Name	Percentual variation relative to mean value
XUV	Soft X-rays	49% to 360%
VUV	Vacuum UV	97 % to 108%
EUV	Extreme UV	72 %to 213%
Ly-α	Lyman-α	77 %to 143%
FUV	Far UV	$\sim$ constant
MUV	Middle UV	$\sim$ constant
NUV	Near UV	$\sim$ constant
UV-C	Ultraviolet C	99.97 % to 100.05 %
UV-B	Ultraviolet B	$\sim$ constant
UV-A	Ultraviolet A	$\sim$ constant
TSI	Total Solar Irradiance	10^{-3}%

TABLE 2.2: Summary table with percentual variation for each band.

The considerations that can be derived are the following ones:

a. The Total Solar Irradiance (TSI) varies about 1 part over 10^5, so, for various geophysical problems it can be considered as a constant.

b. The XUV band has the most relevant variability, and this variability decreases with increasing wavelength.

c. The maxima of the variation are more pronounced and higher than the minima, due to the peculiar profile of the solar cycle.

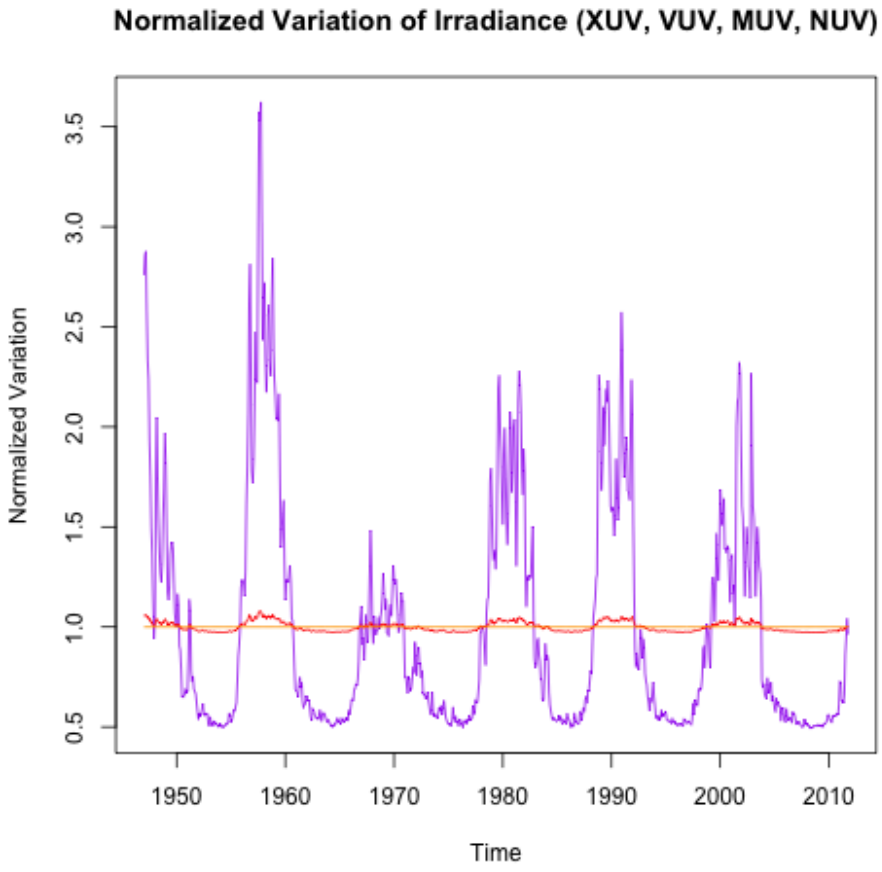


FIGURE 2.3: Normalized variation in time of XUV (purple), VUV (red), MUV and NUV (both orange).

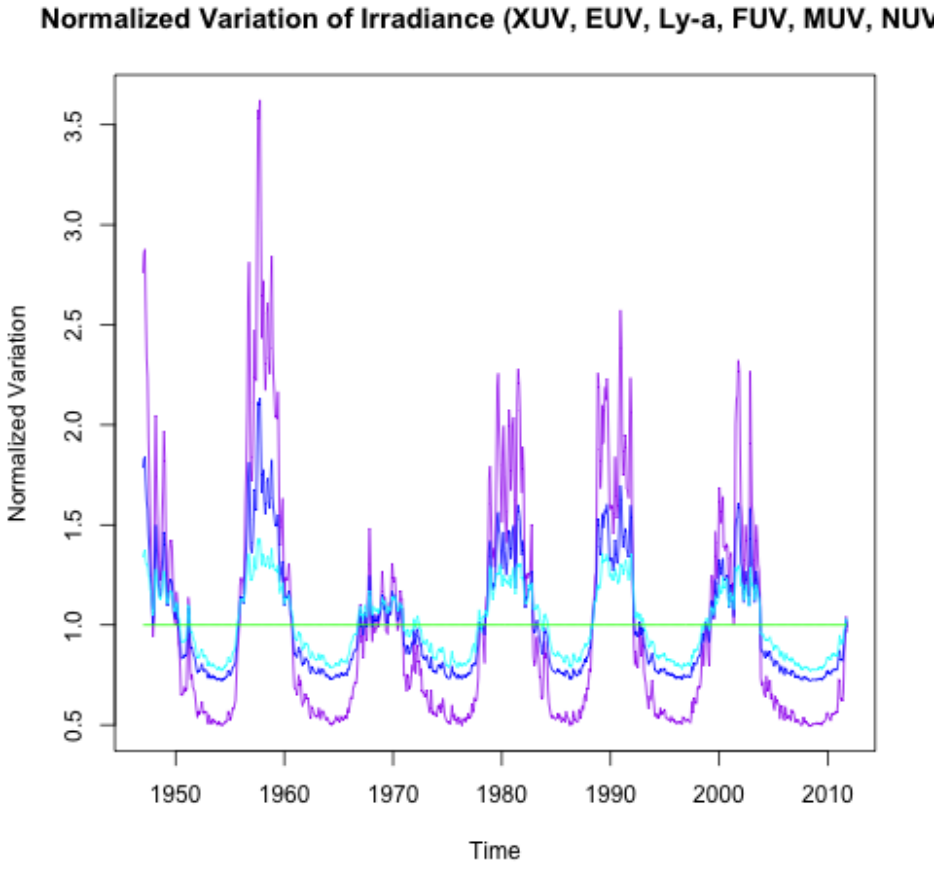


FIGURE 2.4: Normalized variation in time of XUV (purple), EUV (blue), Ly-α (cyan), FUV , MUV and NUV (all three green).

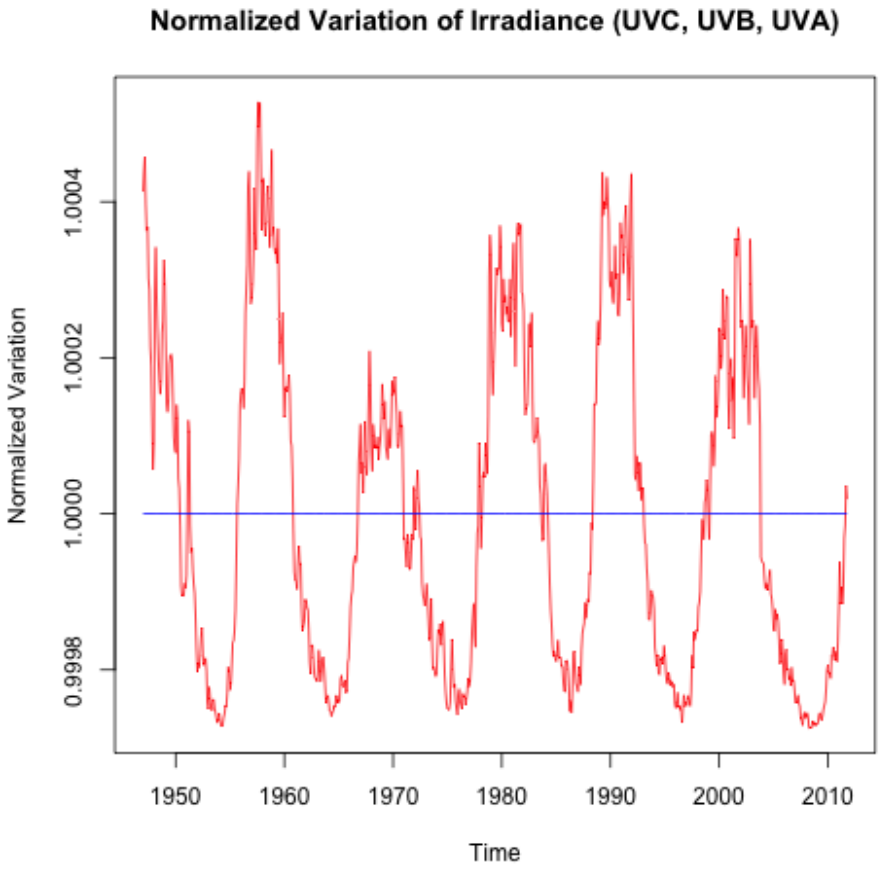


FIGURE 2.5: Normalized variation in time of UV-C (red), UV-B and UV-A (both blue).

d. The bands at wavelengths higher than 122 nm (Lyman-α emission wave-length) could be considered constant. In fact, the variability of VUV is mainly due to EUV and Lyman-α bands, since the variations of FUV are negligible, at this level.

The second important estimation is the irradiance integrated in the wavelength range of each band and then normalized to the total irradiance integrated over the whole spectrum range from XUV to IR. After doing this, one obtains the value relative to the TSI. For clarity, the results are reported in Table 2.3.

Band	Name	Percentual mean value relative to TSI
XUV	Soft X-rays	$3.5 \cdot 10^{-5}$ %
VUV	Vacuum UV	$7.8 \cdot 10^{-3}$%
EUV	Extreme UV	$3.4 \cdot 10^{-4}$%
Ly-α	Lyman-α	$5.4 \cdot 10^{-4}$%
FUV	Far UV	$7 \cdot 10^{-3}$%
MUV	Middle UV	1.1 %
NUV	Near UV	6.7%
UV	Ultraviolet	7.8 %
UV-C	Ultraviolet C	0.5 %
UV-B	Ultraviolet B	1.2%
UV-A	Ultraviolet A	6.1%
VIS	Visible	46%
IR	Infrared	47%

TABLE 2.3: Mean value of each wavelength-band compared to total solar irradiance in percentage.

The main conclusion is that (as already known) the majority of energy comes from the Sun in the visible and IR bands. However, the fact that variations are higher in the UV and XUV bands could have second-order effects to Earth's atmosphere.

A third consideration that can be made, is that on the estimation of the solar variability relative to the Milankovitch cycles, i.e. the change in eccentricity of Earth's orbit during millennia. It could be interesting to know if these solar variations are comparable to the variations due to the long-term orbital variations of Earth or not. The result is in agreement with the estimation of the variation relative to the mean value: XUV, VUV, EUV and Ly-α bands are the most variable, and their fluctuations are of the same order of magnitude of those caused by the Milankovitch cycles. See for instance

Figures 2.6 and 2.7. The fluctuation of the other bands are not comparable to those caused by Milankovitch. To compute the variation of irradiance due to Earth's orbit eccentricity variations, simple geometric considerations are done as follows. Let S_0 be the irradiance in a fixed wavelength band at the mean distance d. If the orbit has an eccentricity value e, then the irradiance value varies, and the upper and the lower extrema are:

$$S_{\max} = \frac{S_0}{(1-e)^2}, \tag{2.1}$$

$$S_{\min} = \frac{S_0}{(1+e)^2}. \tag{2.2}$$

with e varying with respect to time, according to Milankovitch cycles.

2.1.3 Meteorological data

The meteorological data used in this work are a set of temperature anomaly at different altitudes and latitudes, and a set of columnar ozone data at different latitudes. A description of the available data set is reported in Tables 2.4 and 2.5.

Temperature data are annual and seasonal global temperature deviations for the surface, troposphere, and lower stratosphere derived from a 54-Station Radiosonde Network, from year 1958 to 2010. The source is J. K. Angell (Air Resources Laboratory - National Oceanic and Atmospheric Administration).

The ozone data tables were determined from Nimbus 7 TOMS (January 1979 - April 1993) and Earth Probe TOMS (August 1996 - December 2005) satellite measurements.

Moreover, see Figure 2.8 to better understand the scale in altitude of our analysis.

2.1.4 The R software

In this work, the 'R' software has been used to handle data. R is a software environment and a open source programming language, which has several routines for data manipulation, statistical analysis and graphical display. Some of its advantages are [33]:

- effective data handling and storage facility,
- graphical facilities for data analysis and display,
- a simple and effective syntax,

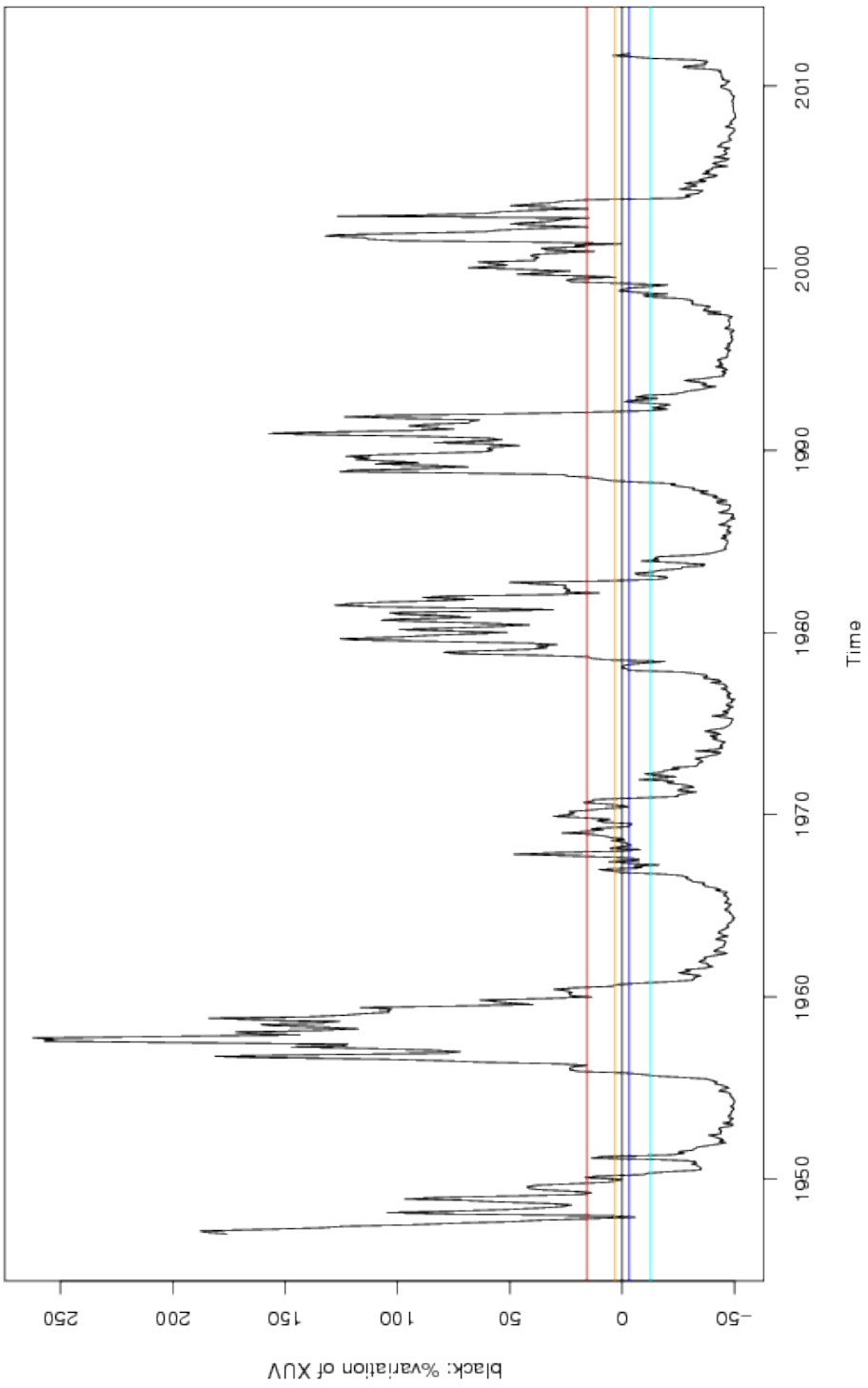


FIGURE 2.6: Variation in XUV band. The straight black line is the mean value of the band; the yellow and blue line shows the interval of variation due to the change of eccentricity at present time. The red and cyan lines represent the same at Milankovitch maximum.

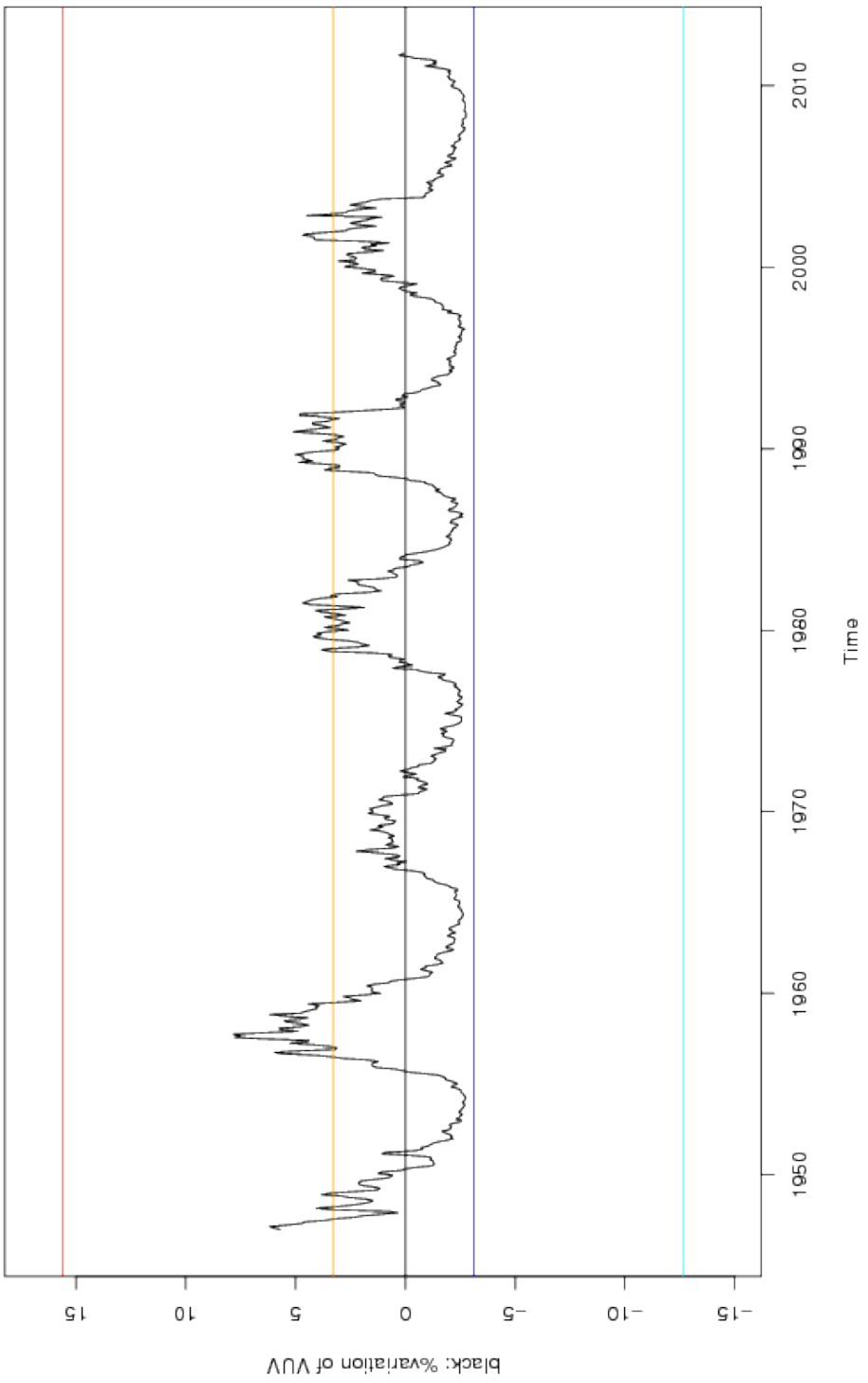


FIGURE 2.7: Variation in VUV band. The straight black line is the mean value of the band; the yellow and blue line shows the interval of variation due to the change of eccentricity at present time. The red and cyan lines represent the same at Milankovitch maximum.

Temperature anomaly ($^{\circ}C$)		
Latitude	Altitude	Time
Global	At surface	Annual
Northern Hemisphere	From 850 hPa to 300 hPa	Seasonal
Southern Hemisphere	From 300 hPa to 100 hPa	
From $90^{\circ}N$ to $60^{\circ}N$	From 100 hPa to 50 hPa	
From $60^{\circ}N$ to $30^{\circ}N$		
From $30^{\circ}N$ to $30^{\circ}S$		
From $30^{\circ}S$ to $60^{\circ}S$		
From $60^{\circ}S$ to $90^{\circ}S$		

TABLE 2.4: Annual and seasonal global temperature deviations for the surface, troposphere, and lower stratosphere derived from a 54-Station Radiosonde Network, from year 1958 to 2010. Source: J. K. Angell (Air Resources Laboratory - National Oceanic and Atmospheric Administration)

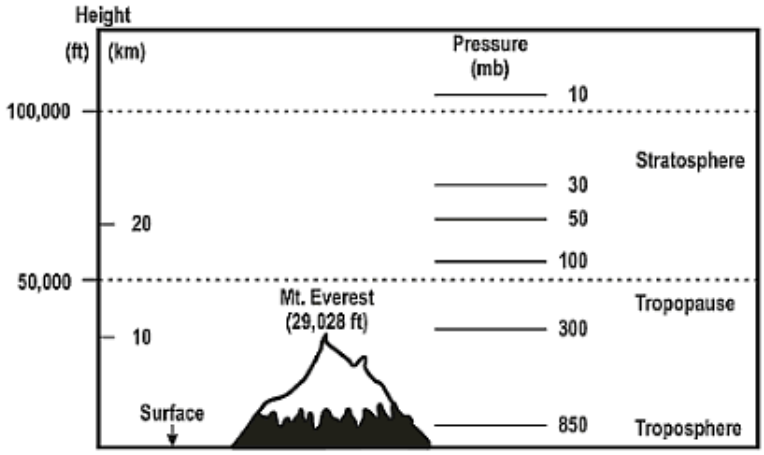


FIGURE 2.8: Atmosphere height in hPa and km.

Ozone level	
Latitude	Time
Global	Annual
From $60°N$ to $45°S$ in 5°-intervals	Monthly
	Daily (only for global data)

TABLE 2.5: Stratospheric Column Ozone. Source: Nimbus 7 TOMS (January 1979 - April 1993) and Earth Probe TOMS (August 1996 - December 2005) satellite measurements.

- a wide collection of open source *packages*, that make data analysis easier and faster.

In particular, the packages used in this work have been `biwavelet` [7] and `EMD` [16] for wavelet and empirical decomposition respectively (see Section 2.3 and 2.4).

2.2 Simple correlations and Fourier analysis

2.2.1 Correlations

The first analysis that can be done is the simple correlation between irradiance and temperature. Since the maximum time resolution of the temperature time series is seasonal, the time resolution of the irradiance data must be adapted to the other time series. This operation has been repeated to every temperature time series, at every altitude and latitude. Since the irradiance band shows the same pattern, the correlations presented in this work involves only the XUV band. As expected, the scatter plots do not show any visible correlation. For instance, see Figure 2.9.

The considerations that one can draw are the following ones.

1. The solar signal could be averaged in space and time; this would impose an averaging in time, since the irradiance could be correlated with temperature at seasonal scale, or annual scale, or at longer scale.

2. The solar signal at t_1 could be correlated with temperature at $t_1 + T$, where T is a unknown response time.

3. The solar signal at t_1 could be correlated with the difference between temperature at t_1 and temperature at $t_0 < t_1$. In other words, it could be correlated with the derivative of temperature.

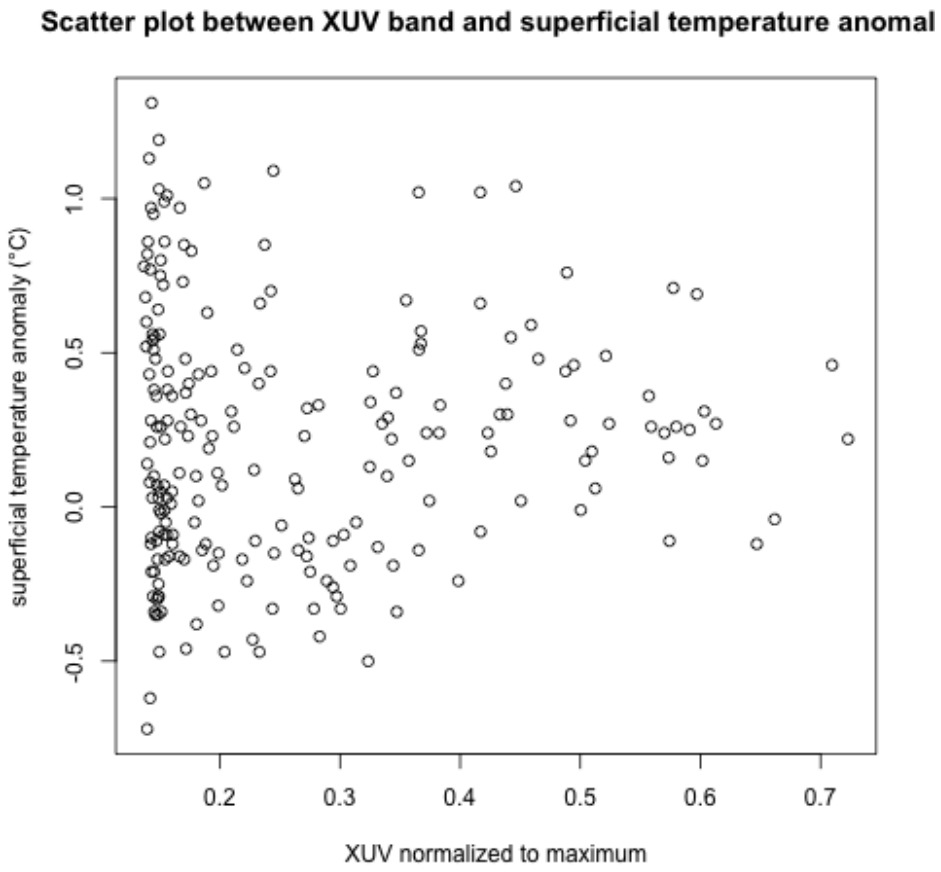


FIGURE 2.9: Scatter plot between XUV band and surface temperature anomaly ($^{\circ}C$)

With regard to point 1., the time series have been averaged in time intervals which could be related to some connections between Sun and Earth, like one year, half a year, 11 years (a solar sunspots cycle), half a solar cycle, 22 years (magnetic solar cycle). Unfortunately, this operation did not reveal any correlation of interest and, anyway, the choice of the temporal window is arbitrary. The typical scatter plot seems like the one presented in Figure 2.9 (eventually with less points) and is not included in this work for brevity.

Point 2. has been developed in the following way. At each scale above mentioned, the correlation coefficient was calculated between the solar irradiance and the temperature time series delayed by a step. The step lag depends obviously on the scale at which the time series are averaged. After having done this, the value of the correlation coefficient was compared with the null-hypothesis that there is no correlation (*Student's t-test*). In Figure 2.10 is graphed the value of t vs. the time lag, while Figure 2.11 reports the p-value. A series of data is considered statistically significant if its p-value is lower than a threshold value (0.05 for 95% of confidence). The two series considered were XUV band irradiance and surface temperature. The time resolution is seasonal (4 data for year). The main result is that the hypothesis that there is correlation between irradiance and temperature cannot be always rejected (at 95% of level of confidence). The fact that the highest delays show the highest correlation is due to the fact that in these cases the number of paired data is very low, and so this is not so significant. Unfortunately, if one plots the solar irradiance and the delayed temperature, the graph does not reveal any clear dependence, since the variance is very high.

The scatter plots between solar irradiance and the derivative of temperature do not reveal anything of interest, and the correlation values are low.

2.2.2 Fourier Analysis

This section illustrates the analysis of time series Fourier spectra, in order to identify some periodicities that can be associated with solar activity.

Temperature

The analysis of Fourier spectra of temperature between 850 and 50 hPa shows the presence of a periodicity around 22 years. This may be connected with the magnetic solar cycle, which has a period of 22 years. Moreover, the temperature between 300 and 100 exhibits, in addition, a period around 11 years, which could be associated with the solar sunspots number cycle of 11 years. The results are reported in Figures 2.12, 2.13, 2.14 and 2.15. It is important

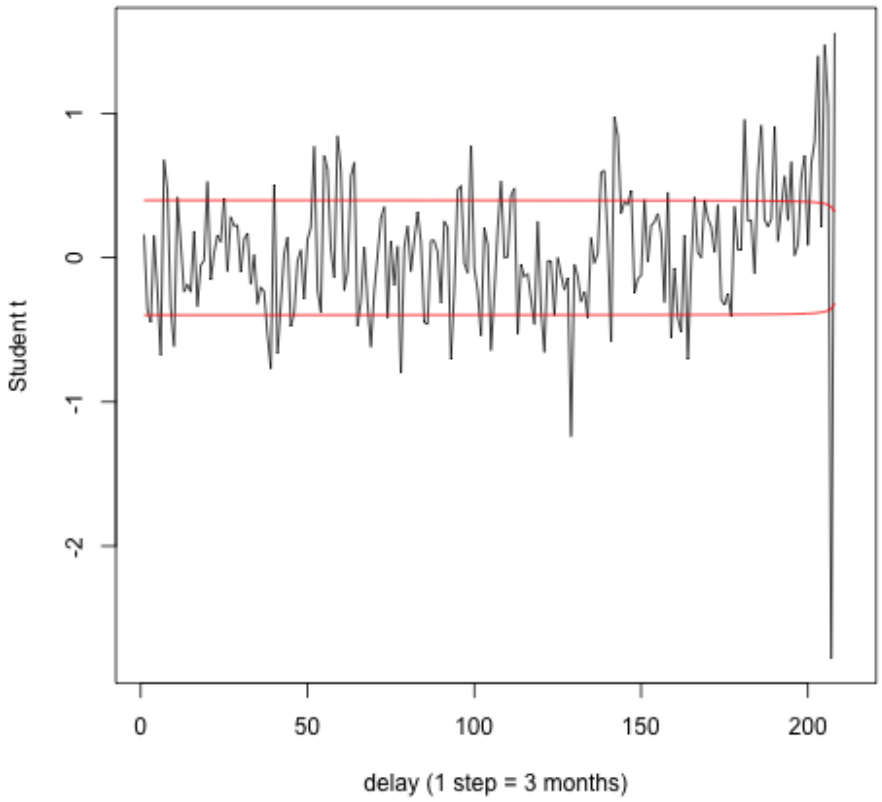


FIGURE 2.10: Value of Student's t vs. time delay. The two series are XUV band irradiance and surface temperature. The time sampling is seasonal (4 data for year).

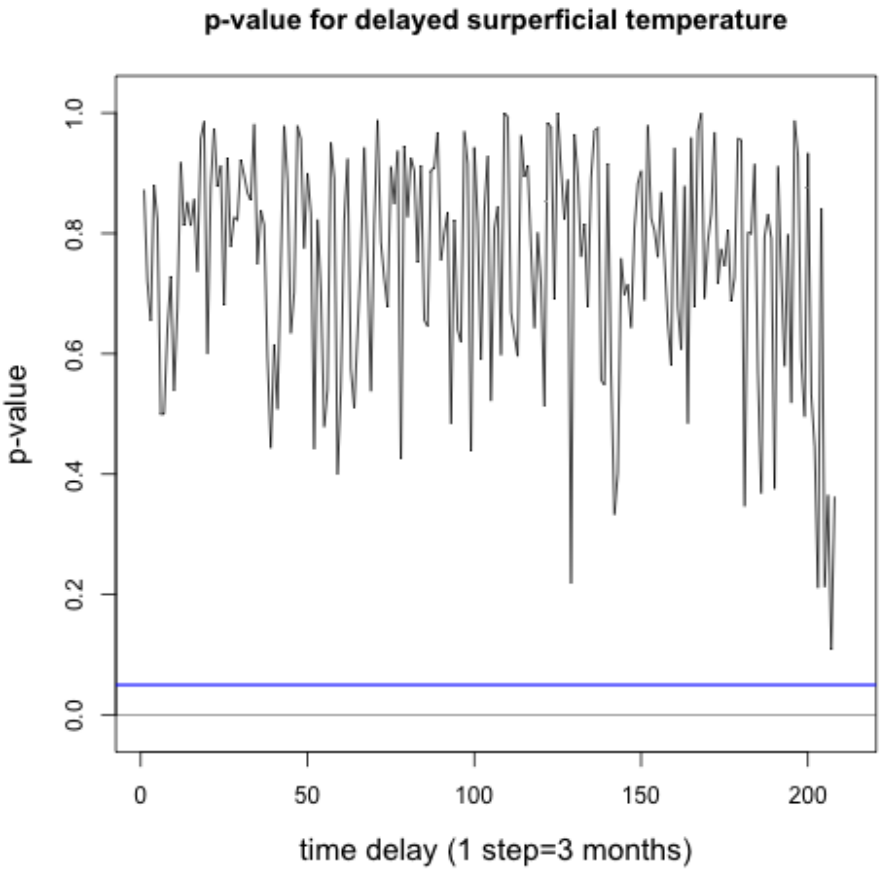


FIGURE 2.11: p-value vs. time delay. The two series are XUV band irradiance and surface temperature. The time sampling is seasonal (4 data per year).

to stress that the presence of periodicity does not necessarily involve necessary a correlation, and does not clearly explain the mechanism(s) by which the Sun affects climate. On the other hand, it is also clear that the existence of a correlation does involve the presence of the same period of variation.

Ozone level

The spectral analysis of ozone level reveals an important aspect of its response to solar irradiance.

The ozone level at higher latitude (above $5°N$ and below $25°S$) shows mainly an annual period, which is a seasonal effect of Earth's revolution around the Sun. Instead, ozone between about $5°N$ and $25°S$, reveals other significant periods of oscillation, including a period of about 11 years, which could be due to solar activity.

For instance, the periodogram of ozone level at 60°N is reported in Figure 2.16. There is no evidence of 11-year periodicity. Conversely, ozone level periodogram at 5°S is reported in Figure 2.17 and there is some possible evidences of 11-year cycle periodicity.

It is possible to draw some results, which are the starting point for the analysis described in the following sections. Variability in XUV and UV band is much higher than VIS and IR. However, there is no evidence of linear correlation between irradiance and temperature. Fourier analysis suggests some evidences of possible periodicities due to solar activity in temperature time series, while the periodograms of columnar ozone level exhibit periodicities around 11 years mainly at low latitudes.

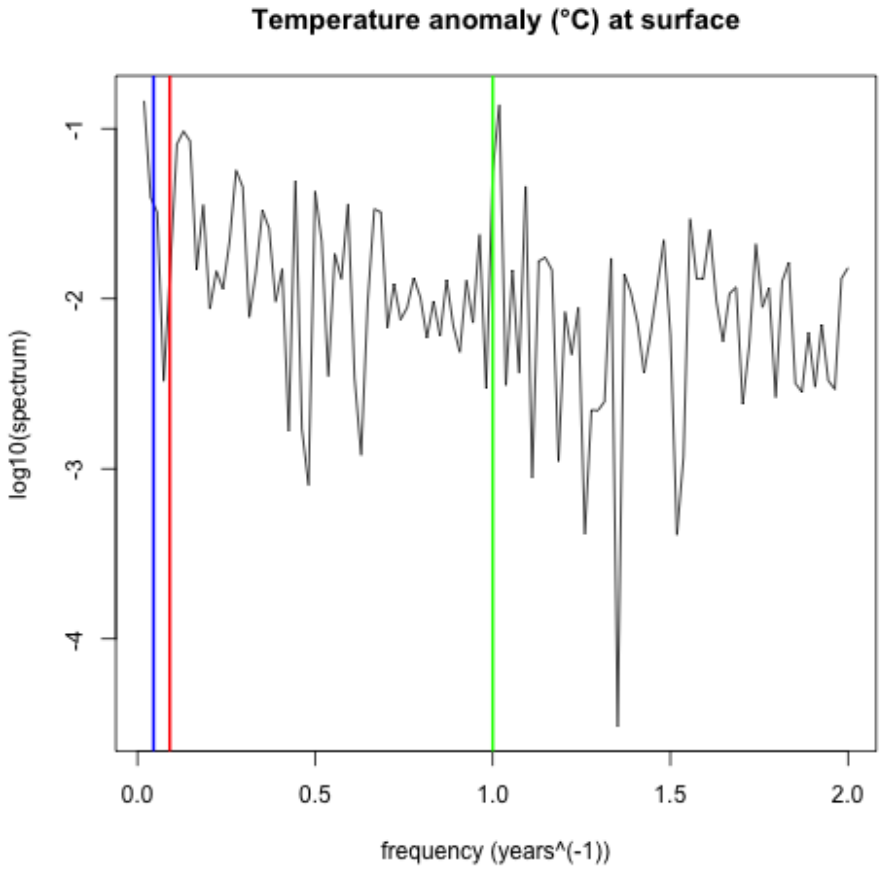


FIGURE 2.12: Periodogram showing the spectrum of temperature at surface. The green line indicates annual period, the red line indicates 11-years cycle and the blue line indicates 22-years cycle.

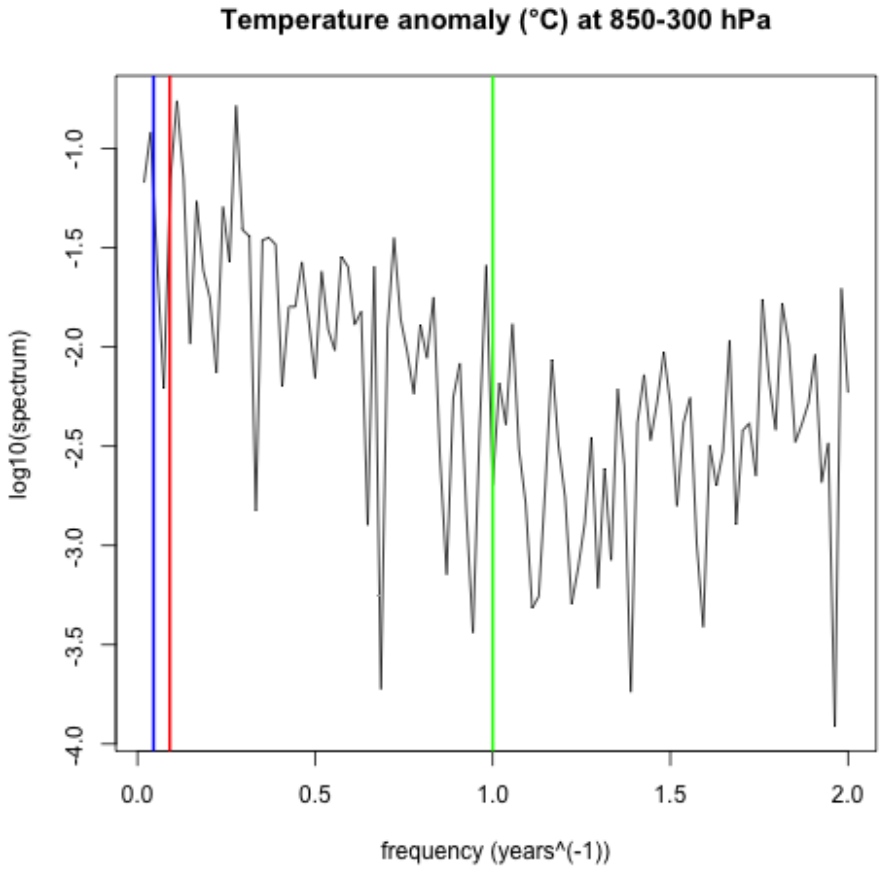


FIGURE 2.13: Periodogram showing the spectrum of temperature between 850 and 300 hPa. The green line indicates annual period, the red line indicates 11-years cycle and the blue line indicates 22-years cycle.

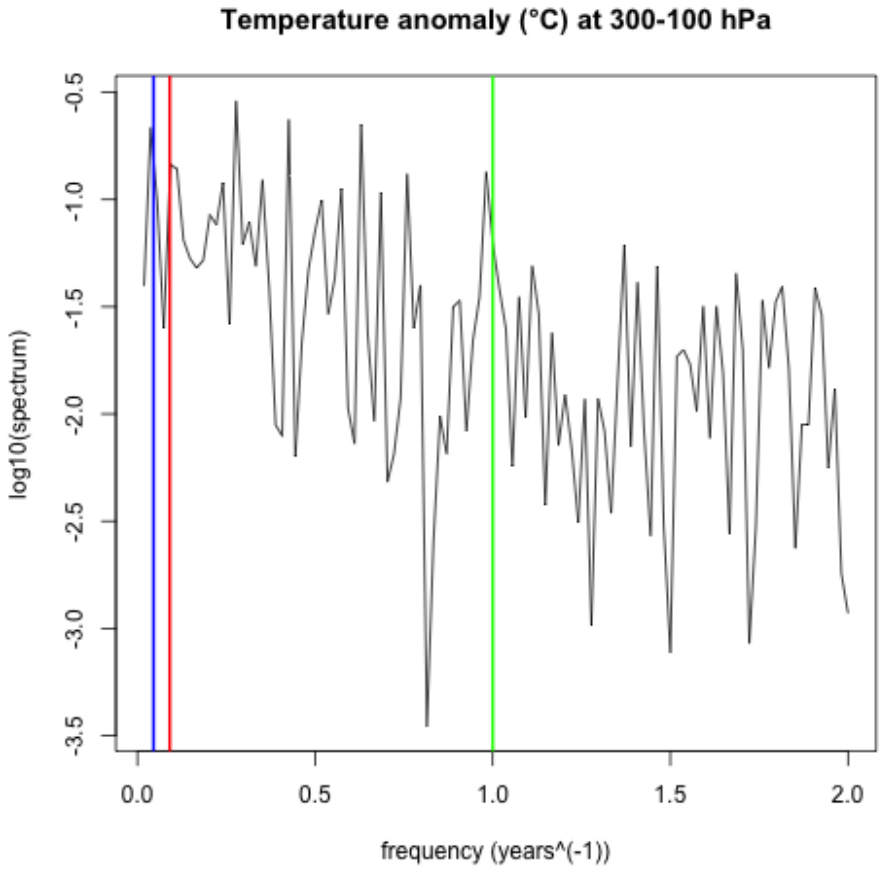


FIGURE 2.14: Periodogram showing the spectrum of temperature between 300 and 100 hPa. The green line indicates annual period, the red line indicates 11-years cycle and the blue line indicates 22-years cycle.

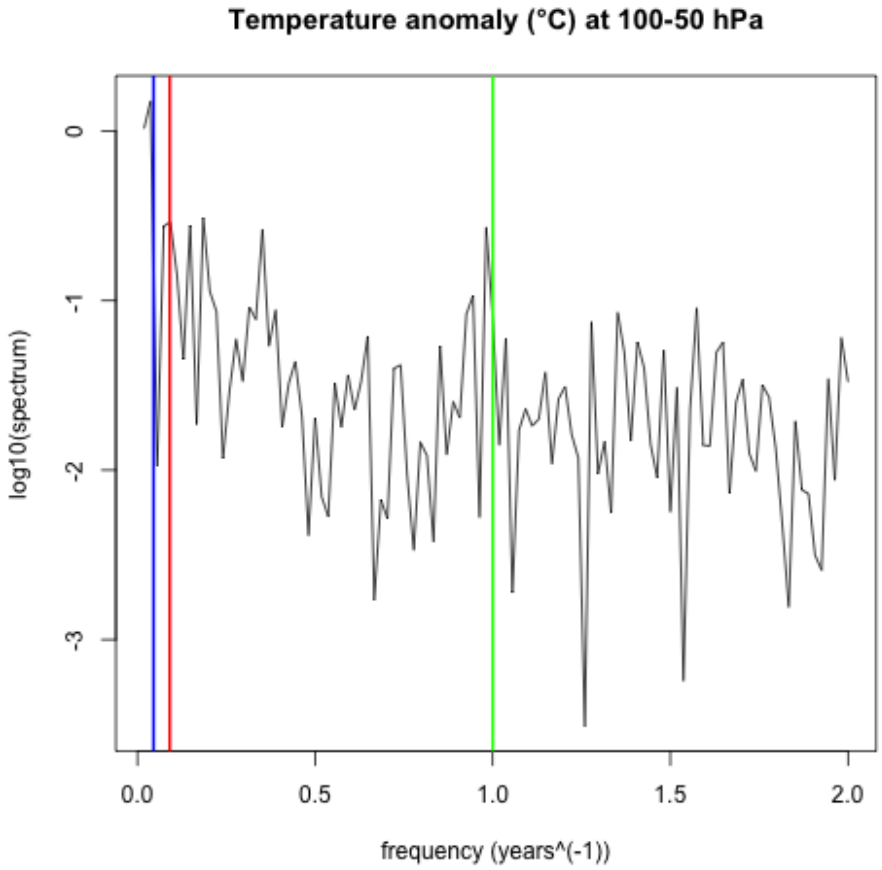


FIGURE 2.15: Periodogram showing the spectrum of temperature between 100 and 50 hPa. The green line indicates annual period, the red line indicates 11-years cycle and the blue line indicates 22-years cycle.

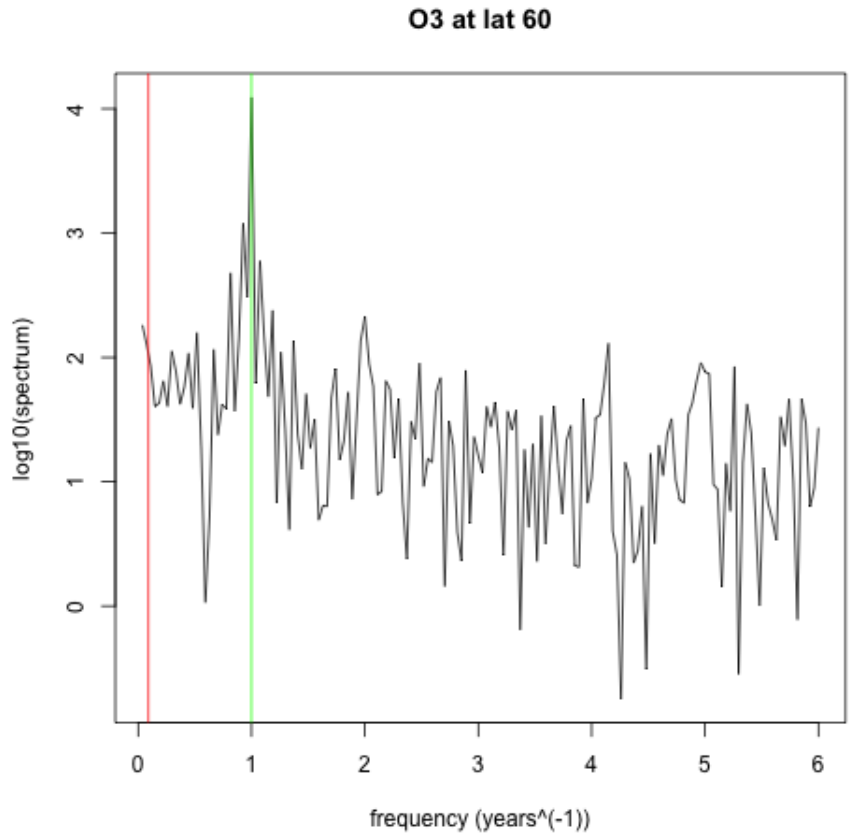


FIGURE 2.16: Periodogram of ozone level at 60°N. The green line indicates annual cycle and the red line indicates 11-year cycle. There is no evidence of the latter periodicity.

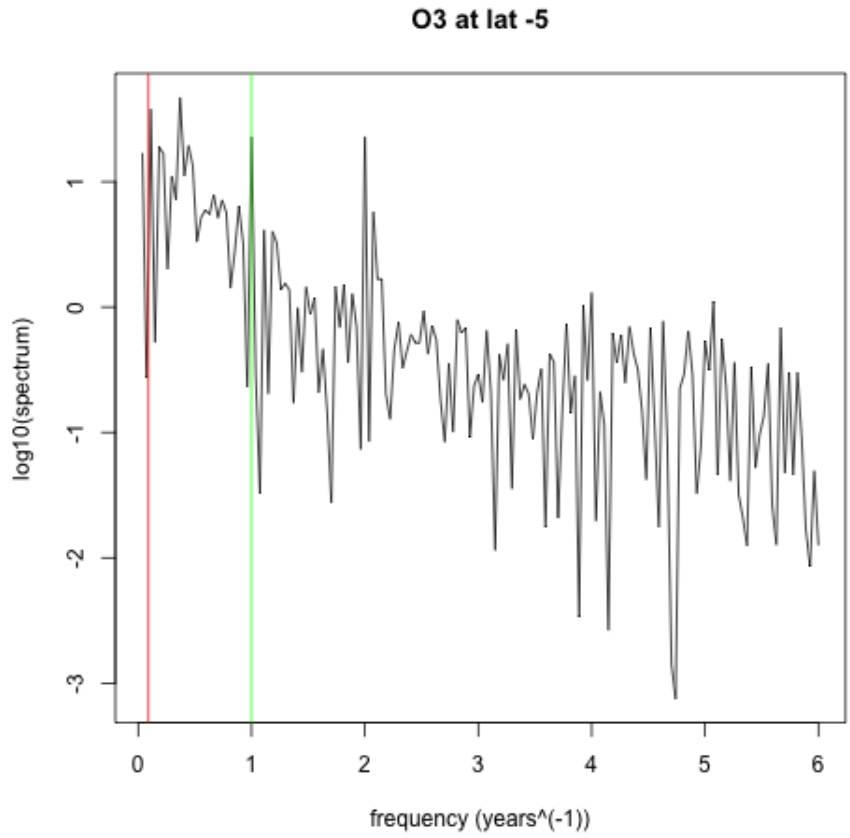


FIGURE 2.17: Periodogram of ozone level at 60°N. The green line indicates annual cycle and the red line indicates 11-year cycle. There is possible evidence of the latter periodicity.

2.3 Wavelet Analysis

2.3.1 Introduction to Wavelet analysis

An intrinsic limit to the Fourier analysis is that, in the transformation from time domain to frequency domain, the temporal information of the time series is lost. For instance, two signals composed by the same frequency components localized in time in two different ways, produce the same Fourier spectrum, and temporal information is lost. On the other hand, the original signal provide no information about the frequency content. In order to overcome this problem, new methodologies have been developed to perform an effective time-frequency analysis.

The first improvement of the simple Fourier Transform is the *Windowed Fourier Transform* (WFT). It consists in performing a Fourier transform on a shifting segment of size T on a time series of time step δt and total length $N\delta t$. The segment is windowed by an arbitrary function (gaussian, boxcar, etc.). The resolution in time-frequency domain is constant and it is given by the product of the time resolution and the frequency resolution. This fact, leads to a sort of 'Indetermination principle', according to which the choice of a high time resolution causes a low frequency resolution (and vice versa). This method represents an inaccurate and inefficient method of time-frequency localization since it imposes an arbitrary time scale T [32].

Wavelet analysis has been developed to study variations of power within a time-series, by decomposing empirical data in time-frequency space. In this work, wavelet analysis has been used to extract some information from time series of meteorological data and solar irradiance data.

First, it is useful to outline the main steps of Wavelet analysis:

1. Choice of a **mother wavelet function**, such that
   - has **zero mean,**
   - is **localized** both in time and in frequency,
   - is **normalized**
2. Choice of a set of scales, which are generally powers of 2 of a initial scale s_0, with time resolution: $s_j = s_0 2^j$, $j = 0, 1, ..., J$, with $J = \log_2(N\delta t/s_0)$.
3. At each scale s, the **convolution** between the wavelet and the signal is performed. The result is the Wavelet Transform as function of scale s and time index n, $W_n(s)$.

4. $W_n(s)$ is, in general, complex-valued, so one can define the Wavelet Power spectrum as $|W_n(s)|^2$.

5. It is also important for the interpretation of results the definition of the so called '**cone of influence**', which is due to the finiteness of the transformed signal. This 'cone' puts a limit on the wavelet analysis at large scale.

6. A **significance test** is defined, with respect to a null hypothesis that the time series has a white-noise (or, possibly, red-noise) Wavelet Spectrum. If a value of the Wavelet Power Spectrum is significantly above that of the white- or red- noise spectrum, then it is considered as a true feature of the time series.

7. Finally, if there are two time series X and Y to be analysed, the Cross-Wavelet Transform is defined, as $W_n^{XY}(s) = W_n^X(s)W_n^{Y*}(s)$, and the **Wavelet Coherence**: $|W_n^{XY}(s)|^2$

The definitions and the kind of wavelet analysis performed in this work are detailed in the following sections.

Definitions

Let us suppose to have a time series x_n with time spacing δt (n = 0,1..., $N-1$), and *wavelet function* $\psi_0(\eta)$ that depends on a non-dimensional time parameter η. This function must have two properties: it must have zero mean and must be localized both in time and frequency. In this type of time-frequency analysis, many different wavelet bases are used and each of them provides different results. A wavelet function can be both *orthogonal* and *non orthogonal*: the first one is usually used to perform a *discrete wavelet transform*, and the second one can be used to perform either the discrete or the *continuous* wavelet transform.

The continuous wavelet transform (CWT) of a discrete sequence x_n is defined as the convolution of x_n with a scaled and shifted version of the wavelet function $\psi_0(\eta)$ ('*' indicates the complex conjugate):

$$W_n(s) = \sum_{i=0}^{N-1} x_i \psi^* \left[\frac{(i-n)\delta t}{s} \right]. \tag{2.3}$$

By varying the wavelet scale s and the time index n one can construct a 2D image in the time index-scale space.

To approximate the CWT, the convolution should be done N times for each time scale. At this point, one can use the convolution theorem to do all

the N convolutions simultaneously in the Fourier space using the Discrete Fourier Transform (DFT) [32]. The DFT of the signal x_n is

$$\hat{x}_k = \frac{1}{N}\sum_{n=0}^{N-1} x_n e^{-2\pi i k n/N}, \tag{2.4}$$

where k is the frequency index. If $\hat{\psi}(s\omega)$ is the Fourier Transform of the function $\psi(t/s)$, then the wavelet transform becomes:

$$W_n(s) = \sum_{k=0}^{N-1} \hat{x}_k \hat{\psi}^*(s\omega_k) e^{i\omega_k n\delta t}, \tag{2.5}$$

where the angular frequency is defined as:

$$\omega_k = \begin{cases} \dfrac{2\pi k}{N\delta t} & : k \leq \frac{N}{2} \\ -\dfrac{2\pi k}{N\delta t} & : k > \frac{N}{2} \end{cases} \tag{2.6}$$

An important factor that has to be considered is the normalization of the wavelet function: in this way, the transforms at each scale are directly comparable to each other.

$$\hat{\psi}(s\omega_k) = \left(\frac{2\pi s}{\delta t}\right)^{1/2} \hat{\psi}_0(s\omega_k), \tag{2.7}$$

$$\psi\left(\frac{(i-n)\delta t}{s}\right) = \left(\frac{\delta t}{s}\right)^{1/2} \psi_0\left(\frac{(i-n)\delta t}{s}\right). \tag{2.8}$$

Generally, the wavelet function $\psi(\eta)$ is complex so also the wavelet transform is complex and it can be divided in real and imaginary part, or amplitude and phase. Finally, one can define the *wavelet power spectrum* as $|W_n(s)|^2$.

In order to compare different wavelet power spectra, it is useful to find a common normalization for the spectrum. The expectation value for $|W_n(s)|^2$ is equal to N times the expectation value for $|\hat{x}_k|^2$. For a white-noise time series, with variance σ^2, the expectation value for $|\hat{x}_k|^2$ is σ^2/N, thus the expectation value for the wavelet power spectrum is $|W_n(s)|^2 = \sigma^2$ at all n and s.

Choice of the Mother Wavelet and the set of scales

A critical point in the wavelet analysis is the arbitrary choice of the wavelet function $\psi_0(\eta)$, and in this important choice there are several factor that should be considered.

A wavelet basis can be **orthogonal** or **nonorthogonal**. In the first case, the wavelet transform produces a spectrum that contains discrete 'blocks' of wavelet power and it is useful for signal processing, since it provides a compact representation of a time series. On the other hand, a nonorthogonal analysis is highly redundant at large scales, since at these scales the wavelet spectrum at adjacent times is highly correlated [32].

A wavelet function can be **real**- or **complex**-valued. A real wavelet is useful to isolate peaks and discontinuities, whereas a complex wavelet is better suited for signals with oscillatory behaviour.

The **resolution** of a wavelet function is determined by the balance between the real space resolution and resolution in the Fourier space.

The **shape** of the wavelet function should reflect the type of signal to be processed. For instance, a signal with sharp jumps or discontinuities will be better analysed by a boxcar-like function, while a smooth signal will be better analysed by a function such as a damped cosine.

In Table 2.6 three different types of wavelet basis function are reported.

	$\psi_0(\eta)$	$\hat{\psi}_0(s\omega)$
Morlet	$\pi^{-1/4}e^{i\omega_0\eta}e^{-\eta^2/2}$	$\pi^{-1/4}H(\omega)e^{-(s\omega-\omega_0)^2/2}$
Paul (order m)	$\frac{2^m i^m m!}{\sqrt{\pi(2m)!}}(1-i\eta)^{-(m+1)}$	$\frac{2^m}{\sqrt{m(2m-1)!}}H(\omega)(s\omega)^m e^{-s\omega}$
DOG (m-th derivative)	$\frac{(-1)^{m+1}}{\sqrt{\Gamma\left(m+\frac{1}{2}\right)}}\frac{d^m}{d\eta^m}\left(e^{-\eta^2/2}\right)$	$\frac{i^m}{\sqrt{\Gamma\left(m+\frac{1}{2}\right)}}(s\omega)^m e^{-(s\omega)^2/2}$

TABLE 2.6: Three different types of mother wavelet: Morlet, Paul and derivative of gaussian (DOG).

Another important point is the choice of scales. It is convenient to write the scales as a fractional power of two:

$$s_j = s_0 2^{j\delta j},\ j = 0, 1, ..., J; \tag{2.9}$$

$$J = \delta j^{-1} log_2(N\delta t/s_0), \tag{2.10}$$

where s_0 is the smallest resolvable scale and J represents the largest scale.

Cone of influence and Significance Test

Since one deals with finite time series, errors will occur at the beginning and the end of the wavelet power spectrum [32]. Furthermore, a common way to make the computation of wavelet power spectrum faster is to pad the end of the time series with zeros before calculating the spectrum, and then to remove them. This process introduces discontinuities at the edges of the signal. The so called *cone of influence* is the region of the wavelet spectrum in which edge effects become important.

There is a relationship between the *equivalent Fourier period* and the wavelet scale s, because the peak in $\hat{\psi}(s\omega)$ does not necessarily occur at frequency s^{-1}. The wavelet scale can be derived analytically for a particular wavelet function by using a pure sine as time series, and then computing the scale s. For instance, for the Morlet wavelet function with $\omega_0 = 6$ the Fourier period λ is $\lambda = 1.03\,s$.

An important aspect of the wavelet power spectrum analysis is the definition of a significance test. The *null* hypothesis is defined as follows. It is assumed that the time series has a mean power spectrum. If a peak of the power spectrum is significantly above the background spectrum, then it can be assumed to be a true feature of the time series with a certain level of confidence. Assuming a mean background spectrum, the distribution of the *Fourier power spectrum* is $N|\hat{x}_k|^2/2\sigma^2$. The corresponding distribution of the *local wavelet power spectrum* is

$$\frac{|W_n(s)|^2}{\sigma^2} \qquad \text{and it is distributed as} \qquad \frac{1}{2}P_k\chi_2^2, \tag{2.11}$$

where P_k is the mean spectrum at the Fourier frequency k, that corresponds to the wavelet scale s.

The *confidence interval* is defined as the probability that the true power at a certain time and scale lies within a certain interval around the estimated wavelet power. From the last Equation 2.11 one can replace the theoretical value $\sigma^2 P_k$ with the 'true' wavelet power $\mathcal{W}_n^2(s)$. The confidence interval is then

$$\frac{2}{\chi_2^2(p/2)}|W_n(s)|^2 \leq \mathcal{W}_n^2(s) \leq \frac{2}{\chi_2^2(1-p/2)}|W_n(s)|^2, \tag{2.12}$$

where p is the desired significance.

The wavelet power spectrum can also be averaged in time or in scale. Averaging in time produces a *time-averaged wavelet spectrum* if the average is computed over a certain period, or a *global wavelet spectrum* if the average is computed over all the local wavelet spectra. Averaging in scale, a new time series is obtained, which indicates the average variance in a certain band [32].

Cross-Spectrum and Wavelet Coherence

Given two time series X and Y, with wavelet transforms $W_n^X(s)$ and $W_n^Y(s)$, the *cross-wavelet spectrum* as $W_n^{XY}(s) = W_n^X(s)W_n^{Y*}(s)$, which is in general complex-valued, and one can define the *cross-wavelet power spectrum* as [32] $|W_n^{XY}(s)|^2$.

The *Wavelet coherence* is defined as the square of the cross spectrum normalized by the individual power spectra, and measures the cross-correlation between two time series as function of time and frequency. Unfortunately, this wavelet coherence is identically one at all times and frequencies; this problem could be avoided by smoothing the cross spectrum before normalization, but in wavelet analysis is not prescribed what sort of smoothing should be performed.

Let us focus on the three types of continuous mother wavelet functions, commonly used in wavelet analysis: Morlet, Paul and derivative of gaussian (hereafter DOG).

Example 1: Morlet Wavelet

The Morlet wavelet is a pure cosine damped by a gaussian function and it is defined as follows in time and frequency domain:

$$\psi_0(\eta) = \pi^{-1/4} e^{im\eta} e^{-\eta^2/2}, \tag{2.13}$$

$$\hat{\psi}_0(s\omega) = \pi^{-1/4} H(\omega) e^{-(s\omega - m)^2/2}, \tag{2.14}$$

where η is the non-dimensional time parameter, m is the wavenumber and H is the Heaviside function. The graphical representation of the Morlet function with parameter $m = 6$ is reported in Figure 2.18.

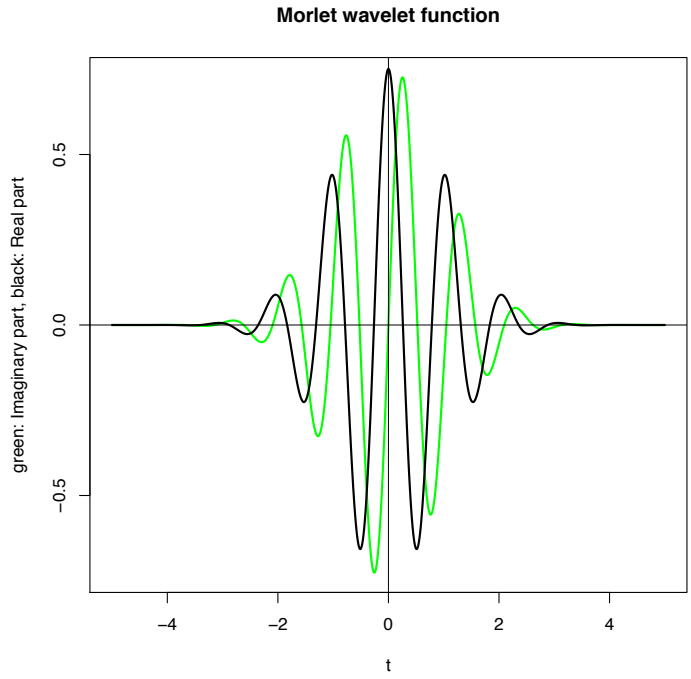
Example 2: Paul Wavelet

The Paul wavelet is another possible choice of basis function, and it decays more quickly than Morlet functions:

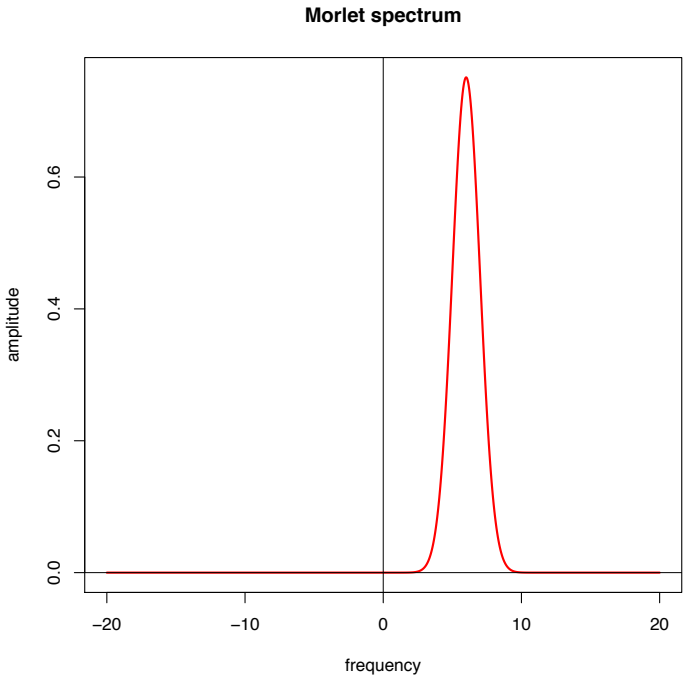
$$\psi_0(\eta) = \frac{2^m i^m m!}{\sqrt{\pi(2m)!}(1 - i\eta)^{-(m+1)}}, \tag{2.15}$$

$$\hat{\psi}_0(s\omega) = \frac{2^m}{\sqrt{m(2m-1)!}} H(\omega)(s\omega)^m e^{-s\omega}. \tag{2.16}$$

In the case of the Paul function, the spectrum in the frequency domain is wider and more asymmetric than the Morlet case. In general, the Paul wavelet in the time domain contains less oscillations than the Morlet wavelet.

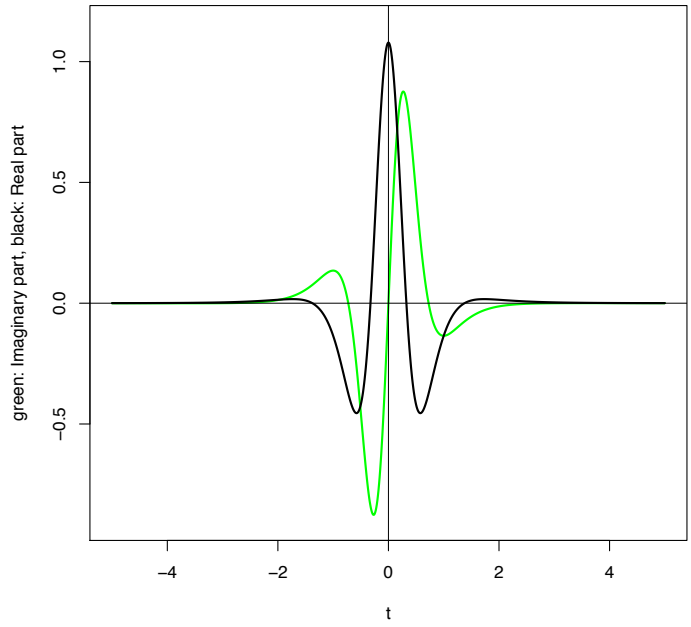


(a) Time domain

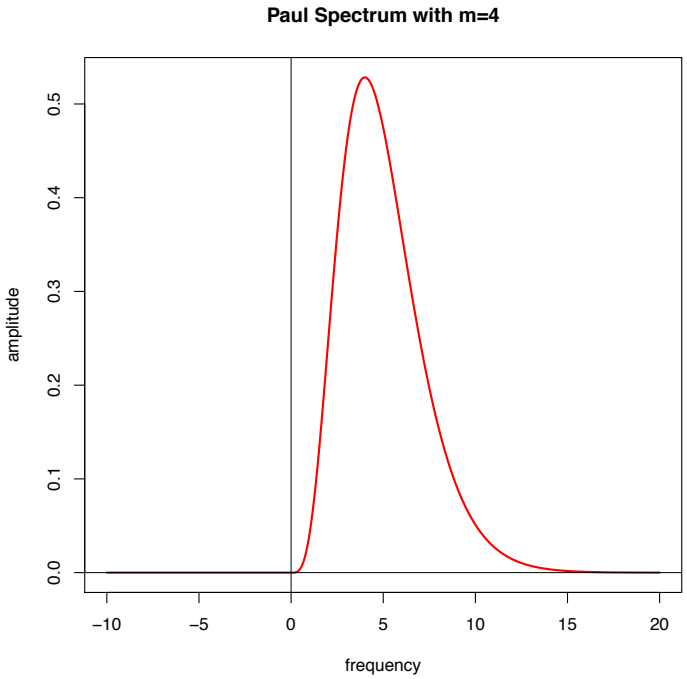


(b) Frequency domain

FIGURE 2.18: Real (black) and Imaginary (green) part of the Morlet function in the time domain (a) and in the frequency domain (b).



(a) Time domain



(b) Frequency domain

FIGURE 2.19: Real (black) and Imaginary (green) part of the Paul function in the time domain (a) and in the frequency domain (b). In this case, $m = 4$

Example 3: DOG Wavelet

The third most commonly used continuous wavelet function is the *Derivative of Gaussian* (DOG), a DOG of order $m = 2$ is also known as '*Mexican hat*'.

$$\psi_0(\eta) = \frac{(-1)^{m+1}}{\sqrt{\Gamma\left(m+\frac{1}{2}\right)}} \frac{d^m}{d\eta^m}\left(e^{-\eta^2/2}\right), \tag{2.17}$$

$$\hat{\psi}_0(s\omega) = -\frac{i^m}{\sqrt{\Gamma\left(m+\frac{1}{2}\right)}} (s\omega)^m e^{-s\omega^2/2}. \tag{2.18}$$

This wavelet decays with the square root of the Gamma function, and, in time localization, its behaviour lies between the Morlet and the Paul wavelet.

Comparison between different mother Wavelets

All the considered three kinds of mother wavelets allow a good time localization, even if the Paul wavelet exhibits the best time localization, the Morlet wavelet the worst one, and the DOG an intermediate localization. In frequency domain, the Morlet wavelet is characterised by the best frequency resolution followed by the DOG and the Paul wavelet, which has the worst frequency resolution [2].

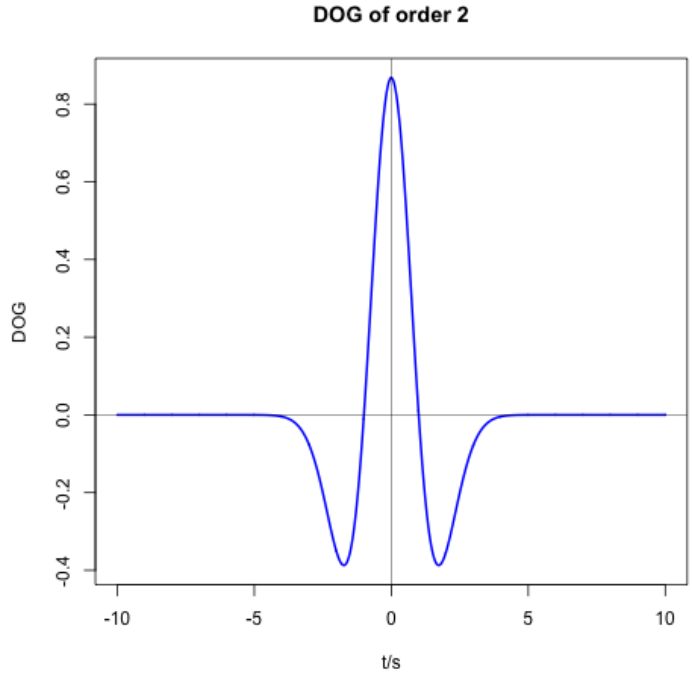
Since the DOG is real-valued, the Wavelet Power Spectrum shows intermittent peaks, where the other two wavelets show an uniform significant peak. This intermittence must not be associated with a true feature of the original signal, because it is an intrinsic feature of a real-valued wavelet.

For example, in Figures 2.21, 2.22 and 2.23 the wavelet spectra of a pure sine of period 0.33 for the Morlet (m=6 and 24), the Paul (m=4 and 40) and the DOG (m=2 and 80) are reported. The best frequency localization with low order m is performed by the Morlet wavelet. The 'cones' centered at the head and tail of the signal contoured by the significance test are produced by the discontinuities at $t = 3.33\,\mathrm{s}$ and $t = 6.66\,\mathrm{s}$, and are the same of those produced by the padding with zeros the time series which has to be analysed.

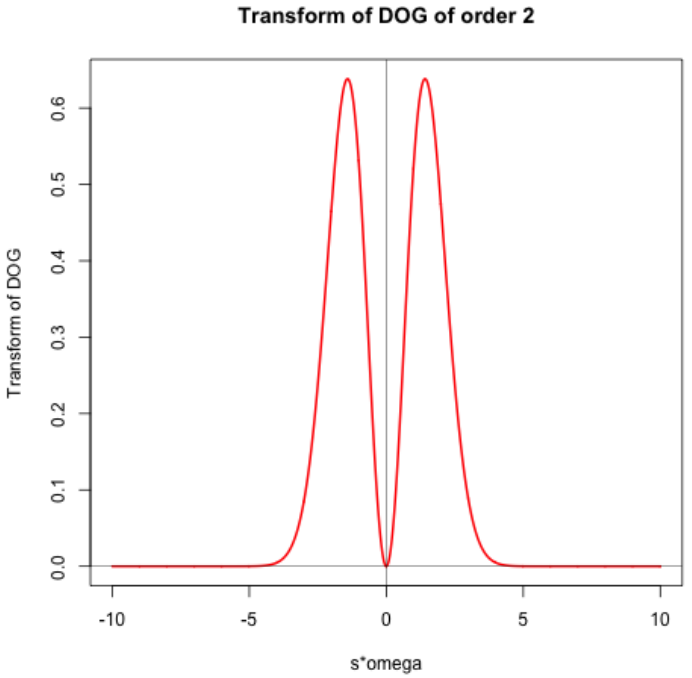
2.3.2 Wavelet analysis of Solar Irradiance data

In this section we describe the Wavelet analysis applied to Solar Irradiance data. In Figure 2.24 the daily data of XUV band (1-10 nm) transformed using a Morlet wavelet of order 6 are depicted. Following the significance contours, three main periods are plotted and localized in time.

The main period is, as expected, that at about 11 years, and could be associated with the solar cycle. The localization in time of this period is not

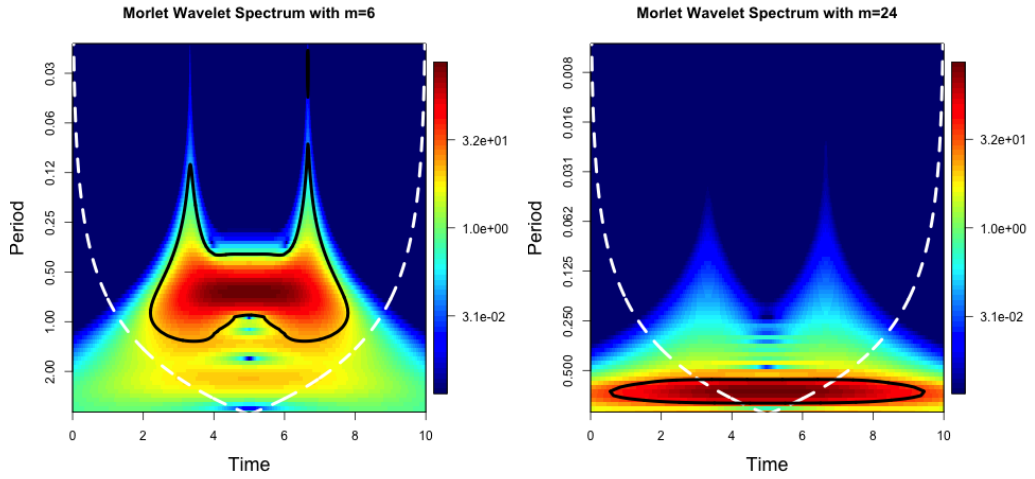


(a) Time domain



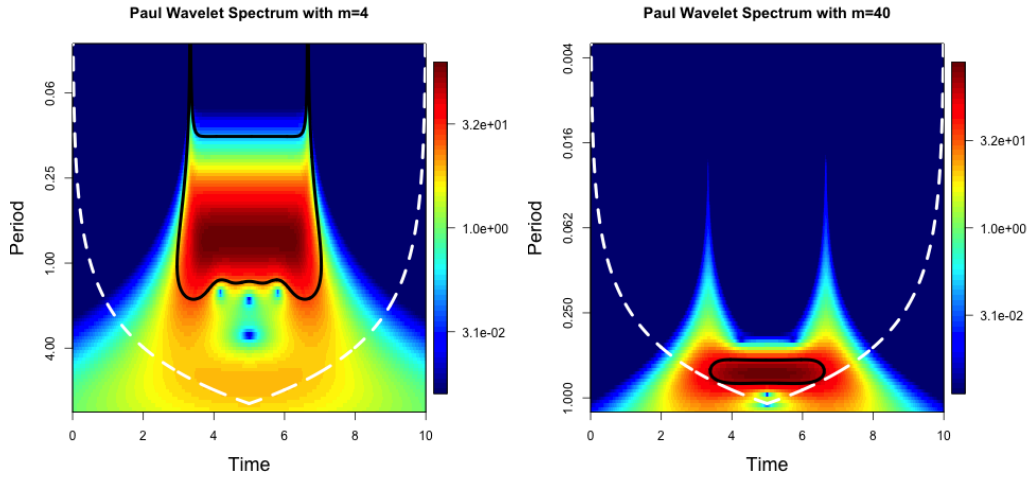
(b) Frequency domain

FIGURE 2.20: DOG function in the time domain (a) and in the frequency domain (b). In this case, $m = 4$



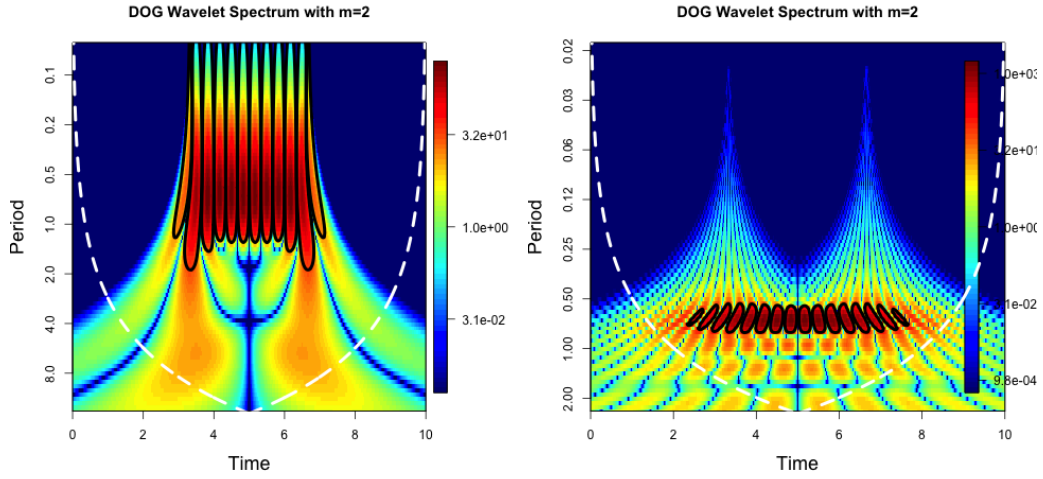
(a) $m = 6$ (b) $m = 24$

FIGURE 2.21: Wavelet spectrum of the Morlet wavelet with m=6, and 24.



(a) $m = 4$ (b) $m = 40$

FIGURE 2.22: Wavelet spectrum of the Paul wavelet with m=4 and 40.



(a) $m = 2$ (b) $m = 80$

FIGURE 2.23: Wavelet spectrum of the DOG wavelet with m=2 and 80.

effective, but this is due to the fact that time localization of low frequencies is worse than that at high frequencies. However, the period of 11 years is a feature appearing in all the time series analyses.

The other two periods pointed out by the wavelet power spectrum are about one year and about one month. The frequency localization is not so effective, but the time localization indicates that the variability at these frequencies are higher during the solar maxima. These two periodicities could be associated with the Earth's orbit (one year period) and the solar rotational period (about 27 days).

The wavelet analysis of other bands and the total solar irradiance (TSI) exhibits a similar behaviour. The wavelet power spectrum of the total solar irradiance is reported in Figure 2.25.

For instance, the wavelet coherence spectrum of XUV and TSI time series is reported in Figure 2.26: it is evident that the coherence is high in the whole time-frequency space.

Figures 2.27 and 2.28 illustrate the wavelet power spectra of the time series of the XUV band calculated with the Paul wavelet (order 4) and the DOG wavelet (order 2, i.e. 'mexican hat'). The periodicities revealed by the spectrum are the same as those shown in the Morlet wavelet spectrum, but are worstly localized in the frequency domain. However, if we adopt a DOG of order 20, the resultant wavelet spectrum is comparable with that obtained with a wavelet of order 6. As a comparison, see Figures 2.24 and 2.29.

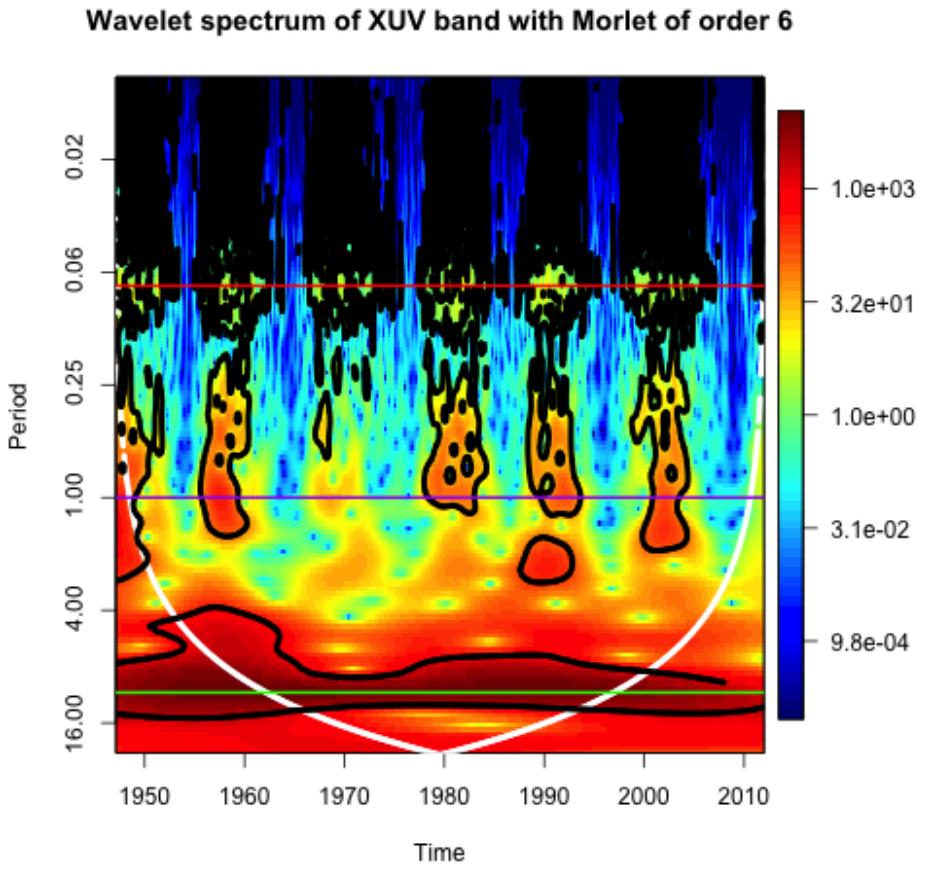


FIGURE 2.24: Wavelet spectrum of XUV band with Morlet wavelet of order 6. The green line represents the period of 11 years, the purple line indicates the period of 1 year, and the red line indicates the period of 27 days.

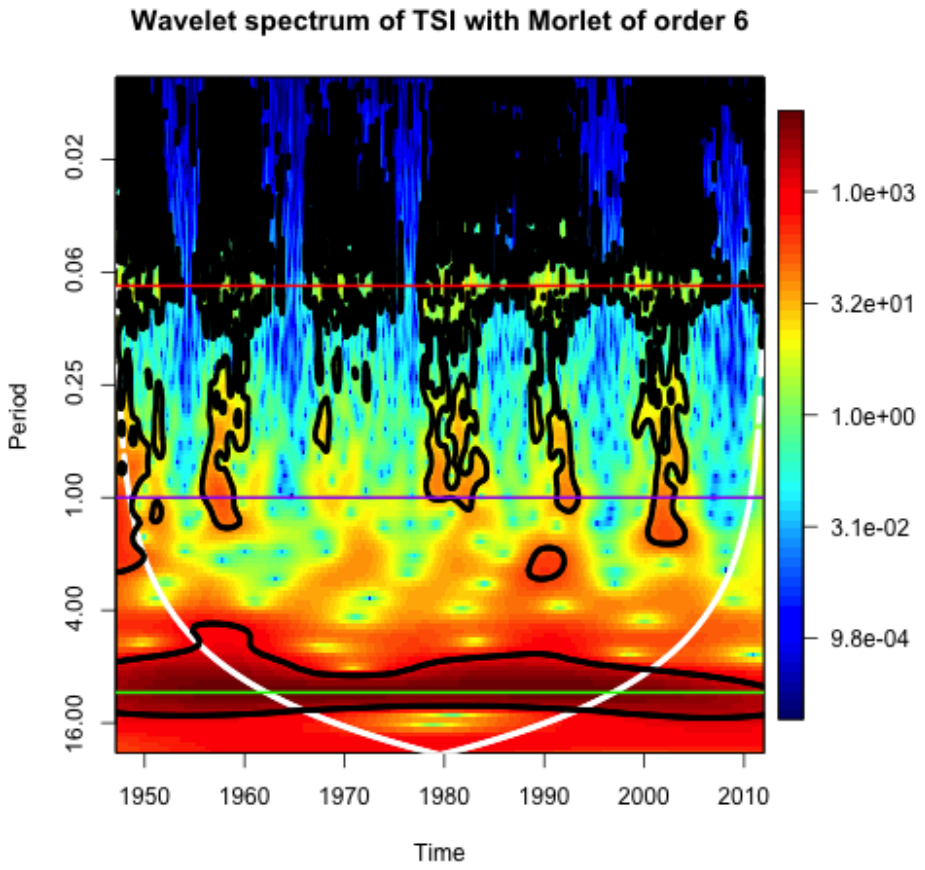


FIGURE 2.25: Wavelet spectrum of total solar irradiance with Morlet wavelet of order 6. The green line represents the period of 11 years, the purple line indicates the period of 1 year, and the red line indicates the period of 27 days.

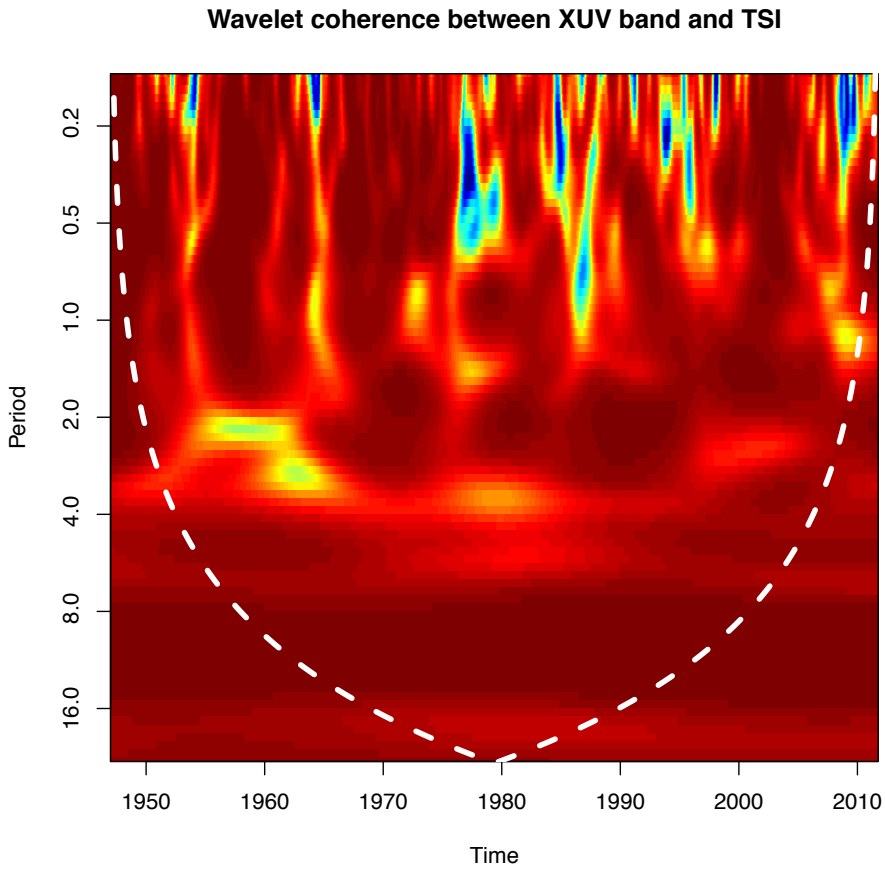


FIGURE 2.26: Wavelet coherence spectrum of XUV band and TSI.

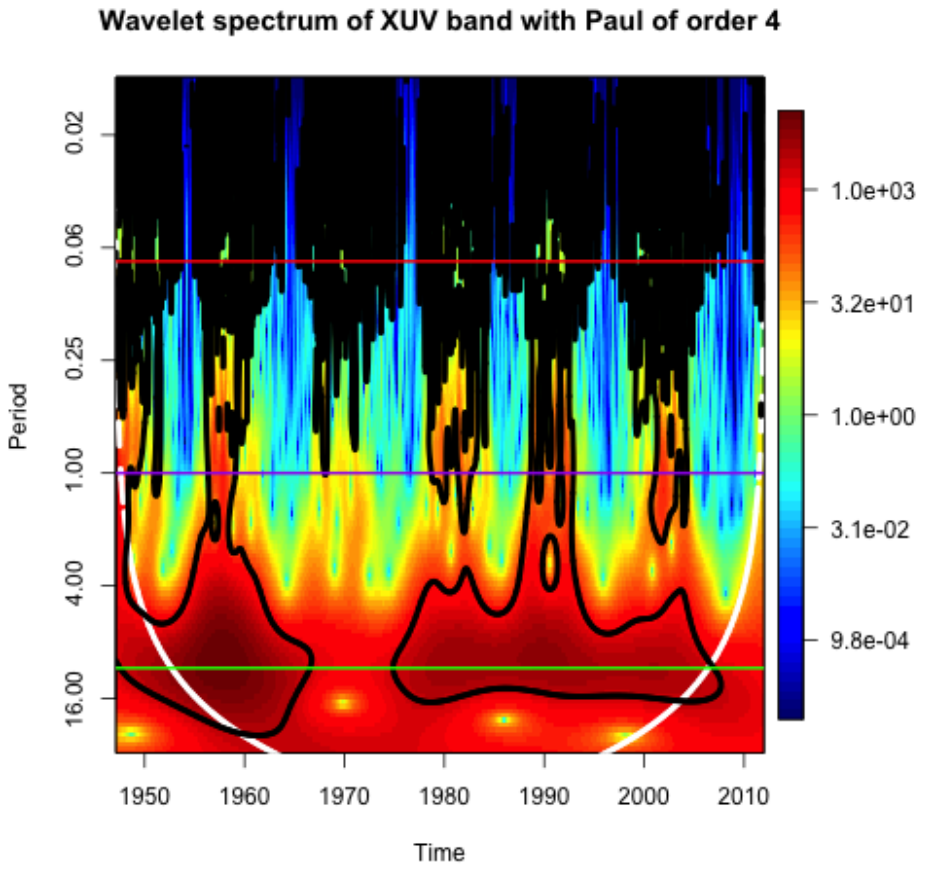


FIGURE 2.27: Wavelet spectrum of XUV band with a Paul of order 4. The green line represents the period of 11 years, the purple line indicates the period of 1 year, and the red line indicates the period of 27 days.

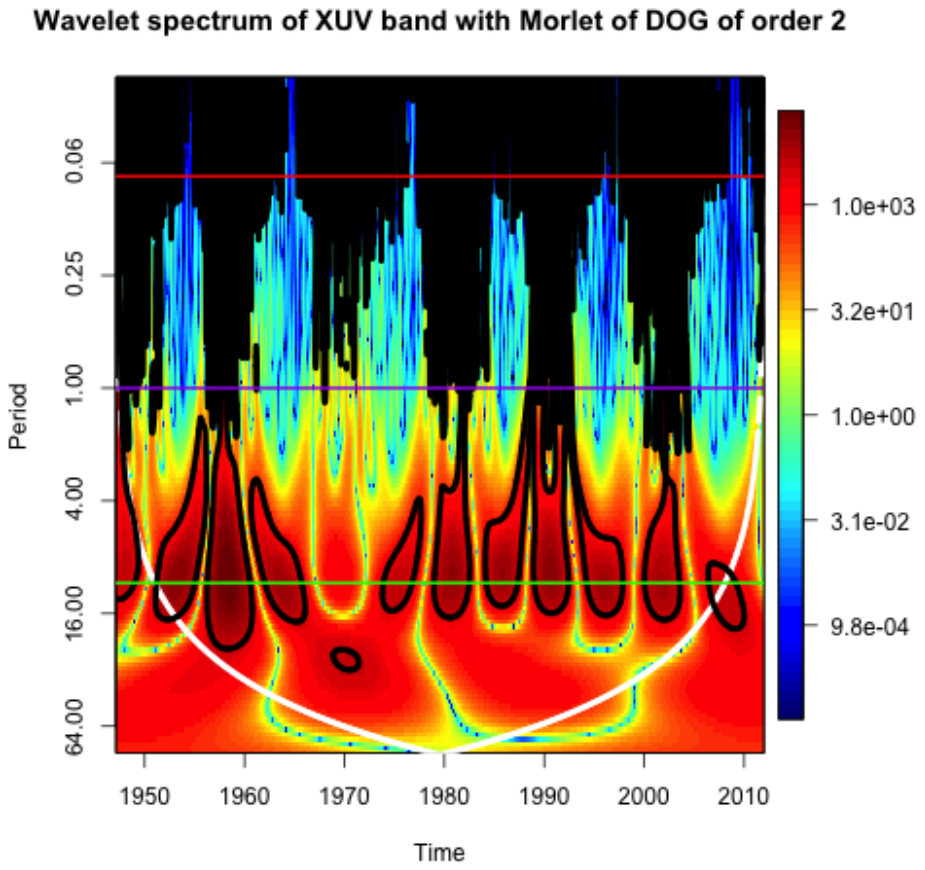


FIGURE 2.28: Wavelet spectrum of XUV band with a DOG of order 2. The green line represents the period of 11 years, the purple line indicates the period of 1 year, and the red line indicates the period of 27 days.

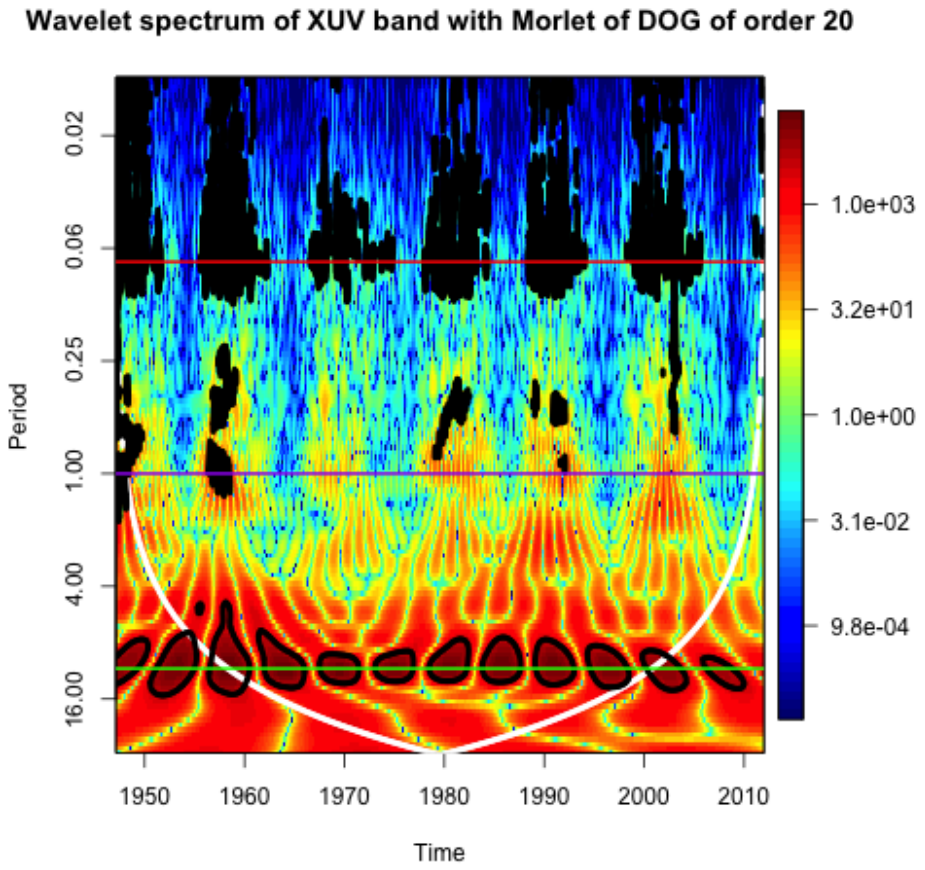


FIGURE 2.29: Wavelet spectrum of XUV band with a DOG of order 20. The green line represents the period of 11 years, the purple line indicates the period of 1 year, and the red line indicates the period of 27 days.

2.3.3 Wavelet analysis of meteorological data

Temperature at different altitudes

In this section the Wavelet analysis of meteorological data is illustrated. First of all, the analysis focuses on the temperature data at four different altitudes: at surface, between 850 and 300 hPa, between 300 and 100 hPa, and between 100 and 50 hPa.

Figures 2.30, 2.31, 2.32 and 2.33 depicted the wavelet spectrum of the four temperature time series. A Morlet wavelet of order 6 was used. In these spectra there are no visible periodicities such as those in the irradiance spectra: the most relevant periods that may be noticed could be a period of about one year in the surface and the 100-50 hPa temperature, and a shorter period than that at 11 year in the 850-300 hPa temperature. Moreover, there is another period around 3-4 years in the temperature at 850-300 and 300-100 hPa, which is not easily associated with solar periodicities, and then it should be linked to other phenomena. Incidentally, the *El Niño - La Niña Southern Oscillation*, which couples atmosphere and ocean, could be associated with this periodicity (see Section 1), since its period of variability is of about 3-4 years. However, the time localization of these frequencies cannot be directly associated with any solar episode, like solar maxima.

Repeating the wavelet analysis using a DOG of order 10 instead of the Morlet one, the same periodicities can be found: temperature at 850-300 hPa exhibits a period slightly shorter than 11 year and a longer period which could reveal the magnetic solar cycle of 22 years. Also the temperature at 300-100 hPa exhibits similar periodicities, and, in addition, an unclear period around 3-4 years. Figures 2.34 and 2.35 illustrate these results. For the sake of simplicity, the spectra of superficial temperature and temperature between 100 and 50 hPa are omitted, since they do not reveal anything of interest.

Ozone at different latitudes

The wavelet analysis was applied to time series of total ozone amount. There are two kind of time series that have been used: the global ozone amount daily data, and a multi-series at 22 different latitudes (from $60°N$ down to $45°S$, with a step of $5°$) with monthly frequency.

The main problem that had to be faced was the fact that these time series are not regular, since several 'gaps' of missing data are present. This fact becomes more important in the wavelet analysis: these discontinuities give rise to 'cones' of artificial signal already seen in Figure 2.18 and following ones.

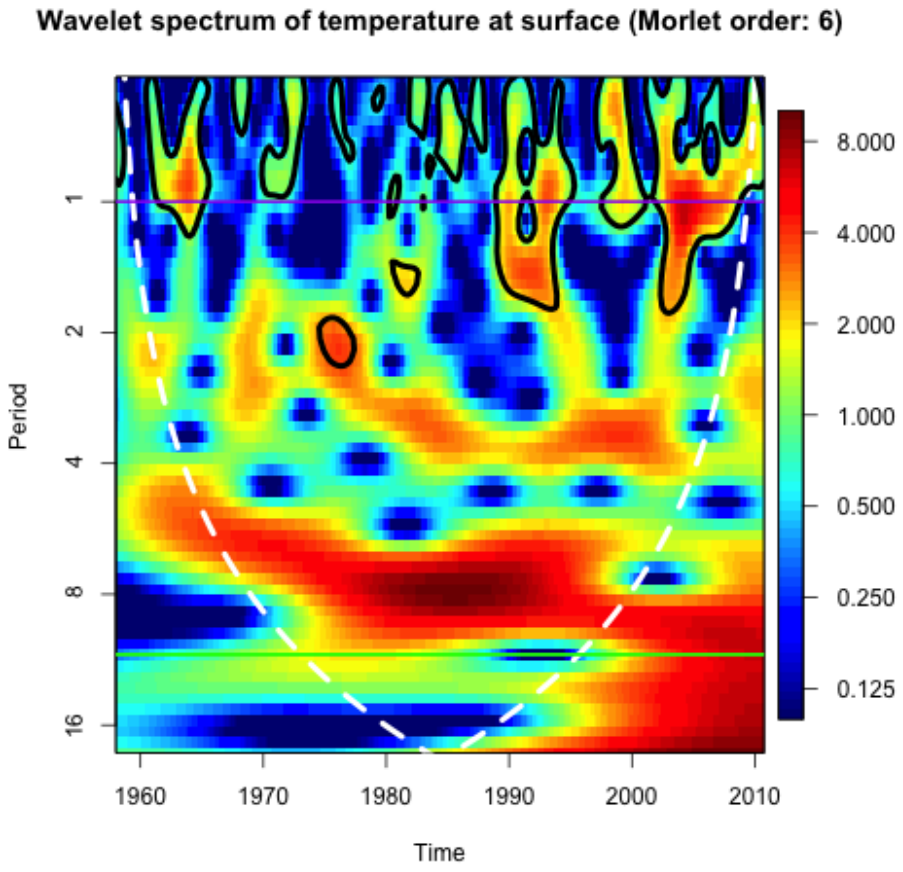


FIGURE 2.30: Wavelet spectrum of temperature at surface (Morlet of order 6). The green line represents the period of 11 years and the purple line indicates the period of 1 year.

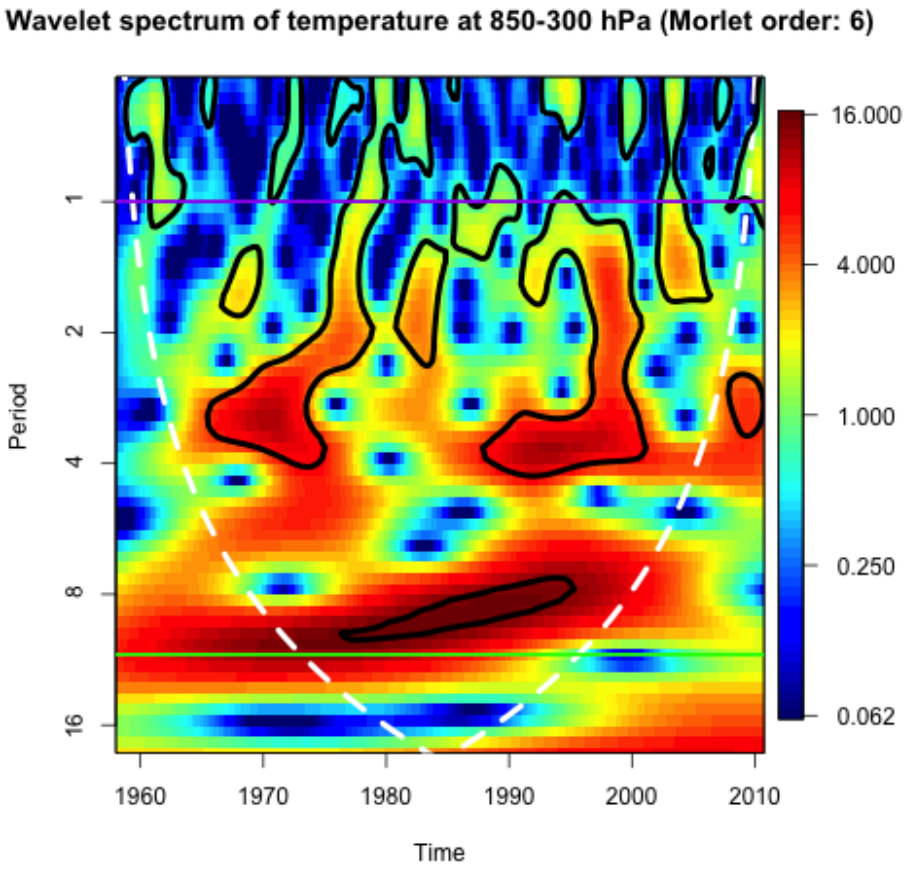


FIGURE 2.31: Wavelet spectrum of temperature between 850 and 300 hPa (Morlet of order 6). The green line represents the period of 11 years and the purple line indicates the period of 1 year.

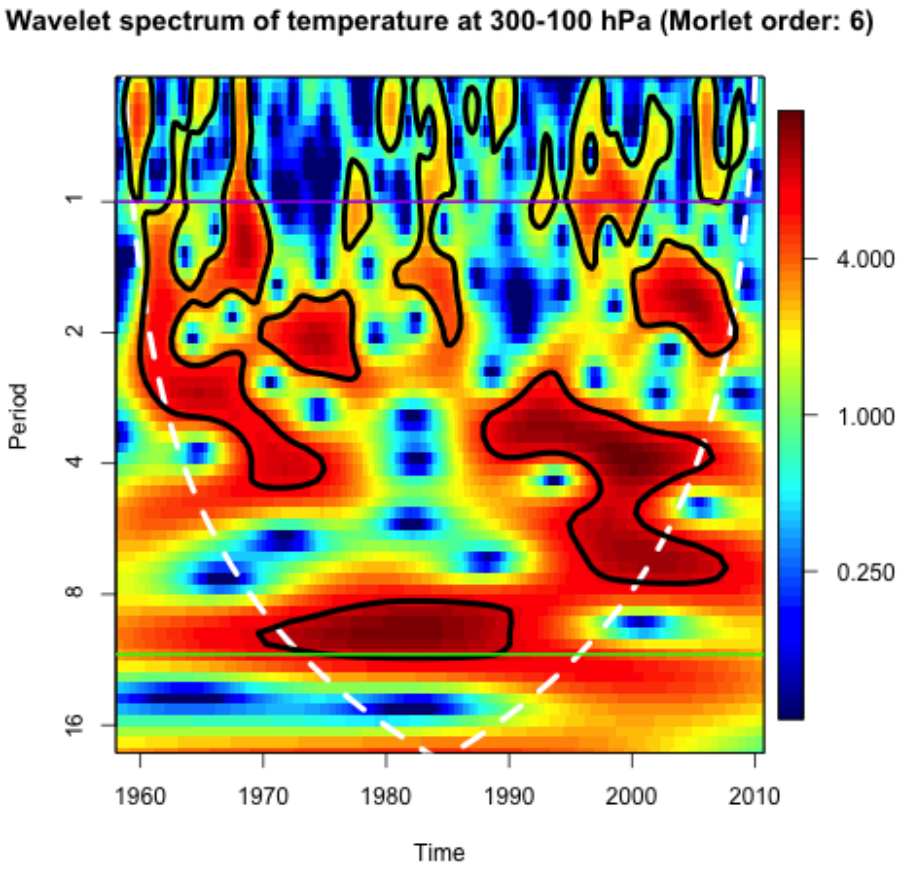


FIGURE 2.32: Wavelet spectrum of temperature between 300 and 100 hPa (Morlet of order 6). The green line represents the period of 11 years and the purple line indicates the period of 1 year.

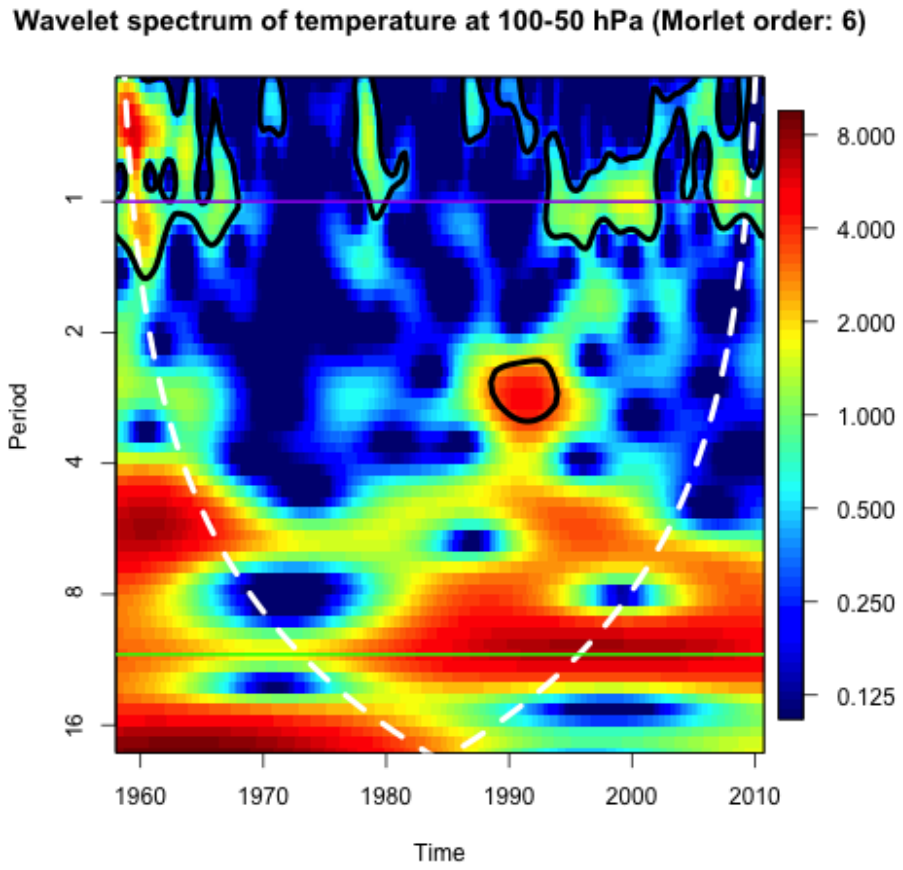


FIGURE 2.33: Wavelet spectrum of temperature between 100 and 50 hPa (Morlet of order 6). The green line represents the period of 11 years and the purple line indicates the period of 1 year.

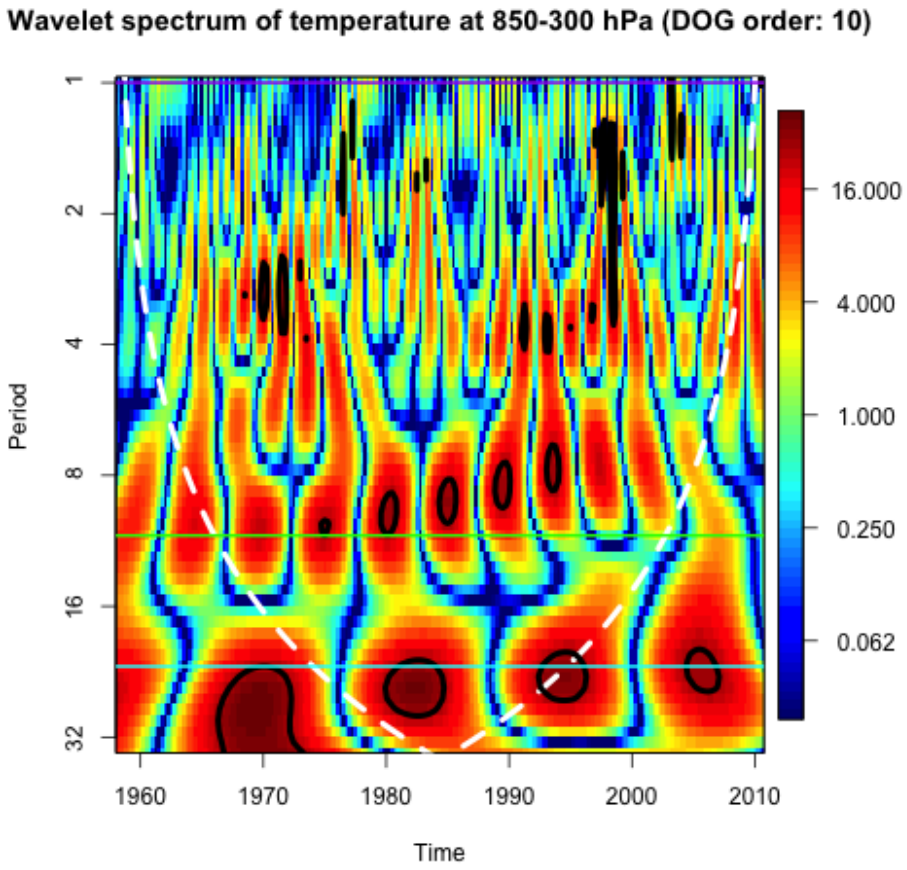


FIGURE 2.34: Wavelet spectrum of temperature between 850 and 300 hPa (DOG of order 10). The cyan line indicates the period o f 22 years, the green line represents the period of 11 years and the purple line indicates the period of 1 year.

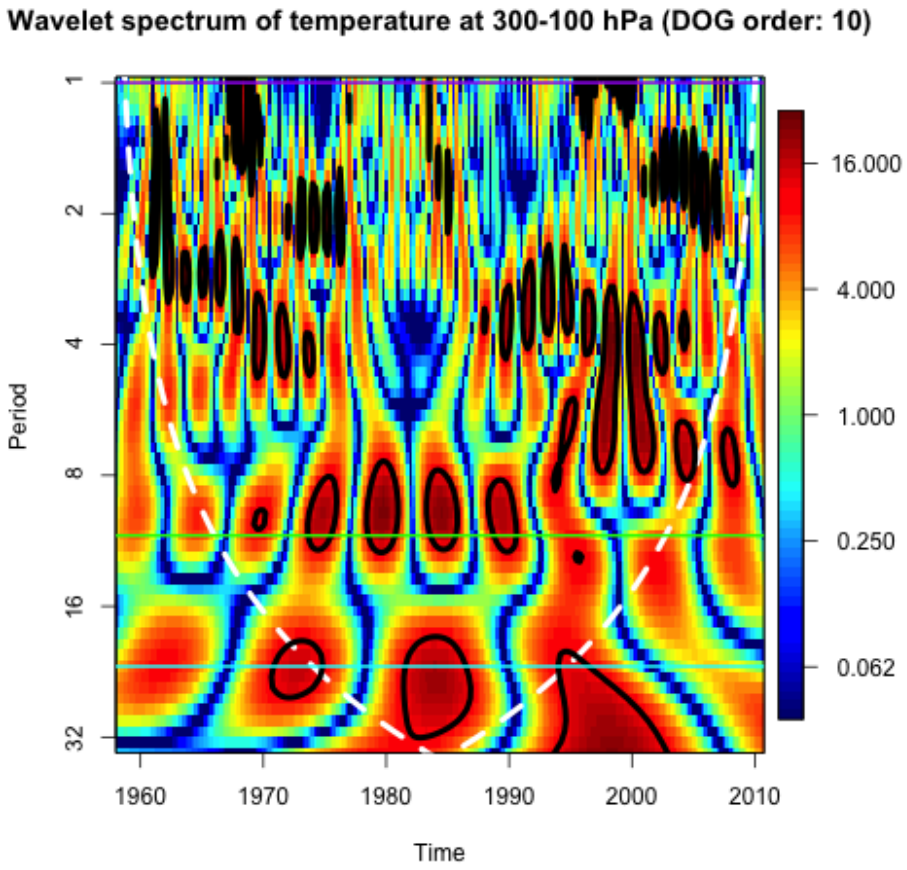


FIGURE 2.35: Wavelet spectrum of temperature between 300 and 100 hPa (DOG of order 10). The cyan line indicates the period o f 22 years, the green line represents the period of 11 years and the purple line indicates the period of 1 year.

Three different workarounds have been adopted:

1. the substitution of a gap of missing data with the constant mean value of the other data of the time series;

2. the substitution of a gap with a linear trend that fits the last datum before the gap with the first datum after the gap;

3. the substitution of a gap with a white-noise signal with mean and variance equal to those of the 'true' data of the time series.

The resultant power spectrum of the daily data is reported in Figure 2.36. The wavelet basis used is a Morlet function of order 6. The annual period is evident and should reflect the Earth's orbit effect. The missing data were here substituted by the mean value of the time series. The other cases show a similar behaviour and, for completeness, are reported in Figure 2.37.

The same analysis has been carried out on the monthly data at different latitudes. Almost the whole northern hemisphere and the southern middle latitudes (from $60°N$ to $5°N$, and from $25°S$ to $45°S$) show a similar spectrum as the previous data. For instance, see Figure 2.38. On the contrary, the data referring to equatorial and south-tropical latitudes reveal a more complex situation. The annual period is not evident and there are several periods of different amplitude localized in certain time intervals, which are not easily associated with any solar periodicity. In Figure 2.39 the spectrum of the ozone value averaged in the latitude interval between $5°N$ and $25°S$ is depicted. Figure 2.40 reports the spectrum of the ozone value in the latitude interval between $0°$ and $5°S$, where other periodicities are evident.

With these findings it is possible to draw some conclusions, to summarise the results obtained in wavelet analysis.

1. The adopted mother wavelets, which proved more suitable have been:
   - Morlet-6 and DOG-20: for, respectively, daily and monthly data.
   - DOG-10: seasonal data

2. The results in terms of temperature anomaly have been that:
   - Long periods are revealed and time localized.
   - Annual frequency events can not be directly associated to solar cycle.

3. Ozone level exhibits a 11-year periodicity, especially at low latitudes, which could be associated with solar irradiance.

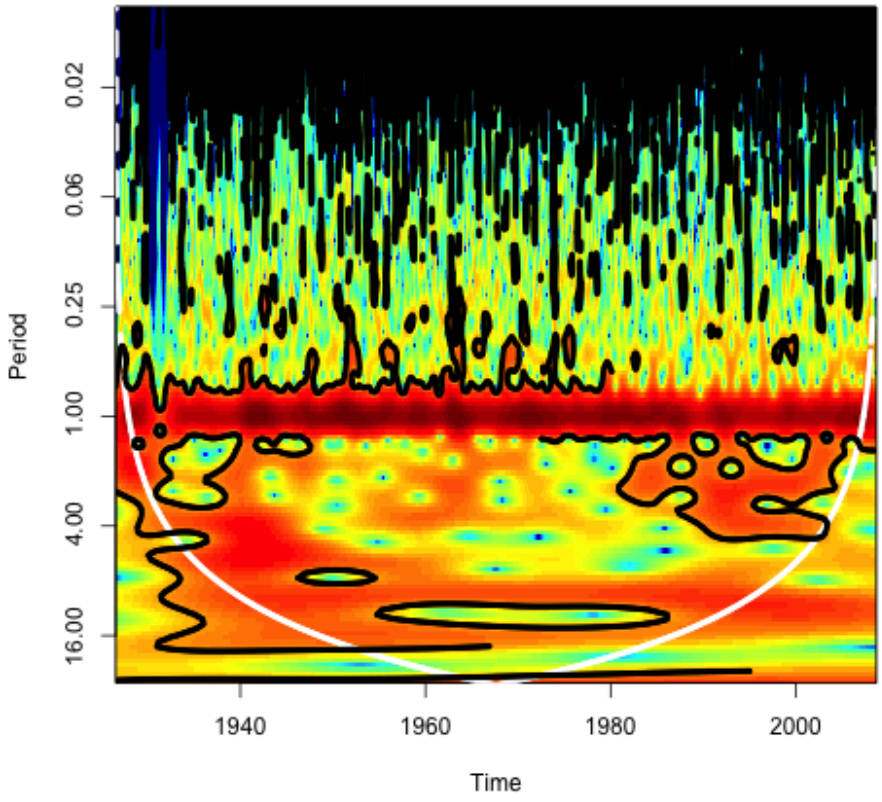
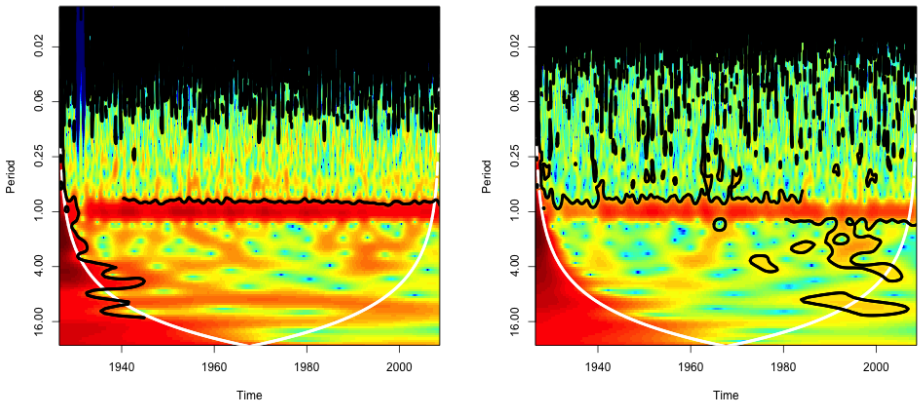


FIGURE 2.36: Wavelet spectrum of daily global ozone level (Morlet of order 6). The missing data are substituted by the mean value of the time series.



(a) linear trend (b) white noise

FIGURE 2.37: Wavelet spectrum of daily global ozone level (Morlet of order 6). The missing data are substituted by a linear trend (2.37(a)) or by a gaussian white noise (2.37(b)).

4. Wavelet analysis is not so efficient to analyse time series with a trend: it could be useful to de-trend data before performing the wavelet transform. See Appendix B.

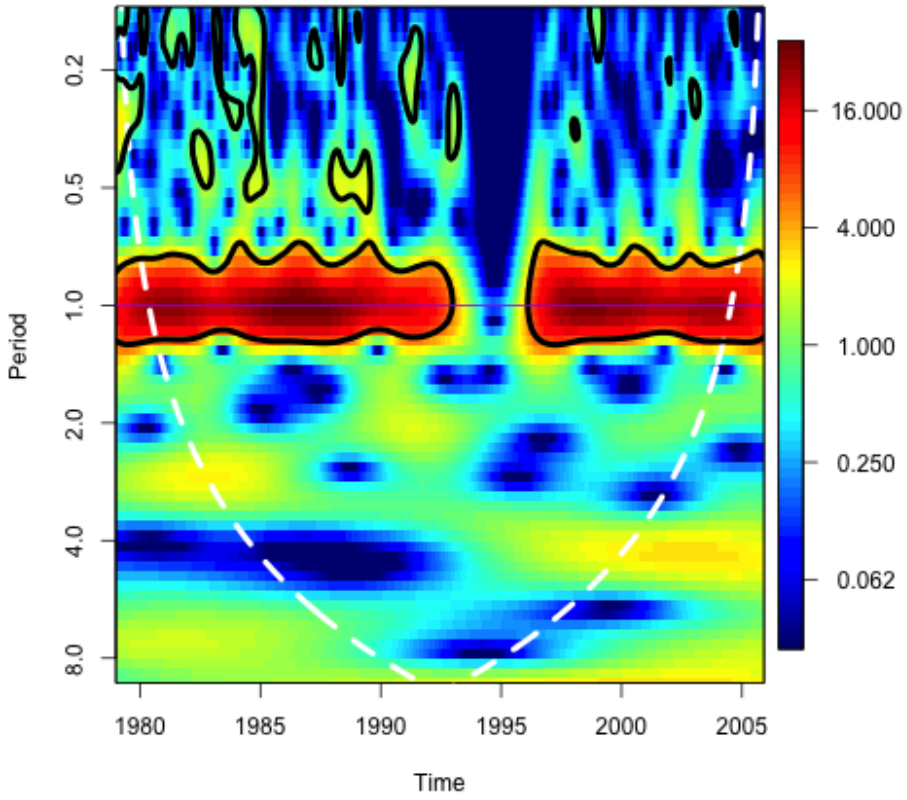


FIGURE 2.38: Wavelet spectrum of daily ozone level in the latitude interval between $60°N$ and $55°N$ (Morlet of order 6). The missing data are substituted by the mean value of the time series. The purple line indicates the annual period.

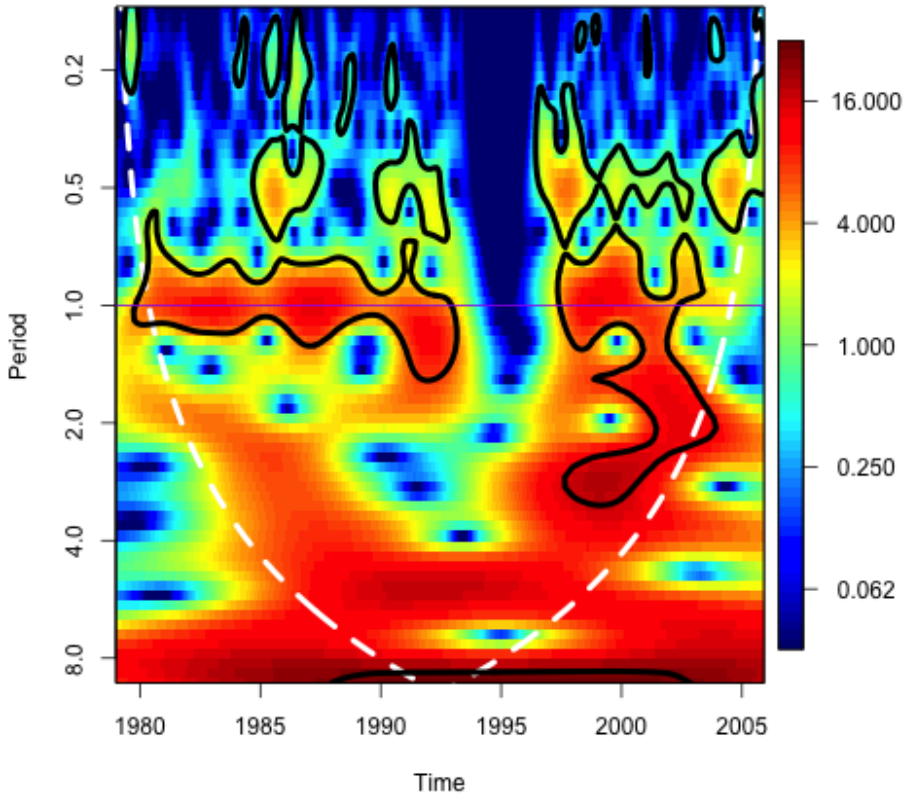


FIGURE 2.39: Wavelet spectrum of daily ozone level in the latitude interval between $5°N$ and $25°S$ (Morlet of order 6). The missing data are substituted by the mean value of the time series. The purple line indicates the annual period.

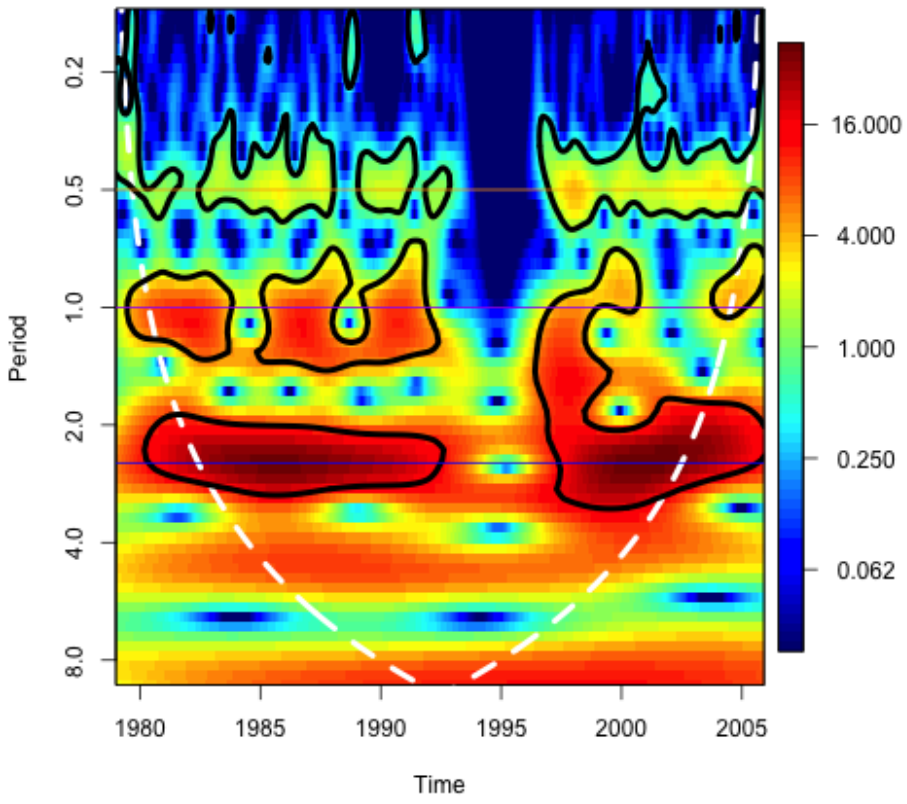


FIGURE 2.40: Wavelet spectrum of daily ozone level in the latitude interval between $0°N$ and $5°S$ (Morlet of order 6). The missing data are substituted by the mean value of the time series. The purple line indicates the annual period, the blue one a period of about 2.5 years and the orange line indicates a period of about 0.5 years.

2.4 Empirical Mode Decomposition

The Empirical Mode Decomposition (EMD) was conceived by Huang et al. (1998) and has been applied in many disciplines of science and engineering. The aim of EMD is to decompose a signal into so-called Intrinsic Mode Functions (IMFs). Furthermore, the Hilbert spectral analysis provides information on frequencies evolving with time and quantifies the variations of energy related to the period of oscillation and the time localization. EMD is also useful to perform a detrending of a signal that does not need arbitrary assumptions, like the choice of a set of filters. This detrending can be applied to wavelet analysis in order to carry out a more effective wavelet spectrum (see Appendix B).

2.4.1 Introduction to EMD

Definition and Derivation of IMFs

The starting point in the definition of an Intrinsic Mode Function (IMF) is the concept of 'frequency', considered as a repetition of an oscillating and periodic pattern ([12], [15], [11], [5]). Focusing on a pure sine function, a single cycle has a local zero mean, and it can be delimited by two successive extrema (maximum or minimum), or by two zeros separated by another zero-crossing. Extending this concept, Huang et al. (1998) defined an IMF as a function with the following features:

1. the number of extrema and zero crossing of a IMF must be equal or differ by one unit

2. the local average of the IMF must always be zero. This implies that the average between the upper and the lower envelope must be zero.

In order to extract the IMFs from the original signal $x(t)$ the following procedure.

1. locate the extrema (Figure 2.41).

2. estimate the envelope of the maxima and the envelope of the minima (Figure 2.42) by an interpolation.

3. the average of the two envelopes $m(t)$ gives the sum of the components at lower frequencies (Figure 2.43).

4. subtract the average $m(t)$ to the original signal $x(t)$, in order to find the candidate for the first IMF $h_{11}(t)$ (Figure 2.45).

5. if the function $h_{11}(t)$ does not correspond to the definition of IMF, apply the routine to $h_{11}(t)$ to find out $h_{12}(t)$, $h_{13}(t)$, ...

This process is called **sifting**. Suppose that, after k repetitions of the routine, the function h_{1k} satisfies the definition of IMF. Therefore, $h_{1k}(t)$ is assumed to be the first IMF $c_1(t)$. The criterions, according to which one can assume h_{1k} as the first IMF, and therefore, stop the sifting process, are the following ones:

1. to fix a maximum number of siftings;
2. the absolute value of the candidate $h_{1k}(t)$ is smaller than a threshold: $|h_{1k}(t)| < \varepsilon, \forall t$;
3. the difference between two consecutive $h_{1,k-1}$, $h_{1,k}$ is smaller than a threshold:
$$\sum_t \left(\frac{h_{1,k} - h_{1,k-1}}{h_{1,k-1}} \right)^2 < \varepsilon. \tag{2.19}$$

The choice of the stopping rule is quite important, because the sifting process causes a loss of amplitude of the sifted function. On the other hand, this is almost the only arbitrary aspect of this kind of analysis.

Once the first IMF $c_1(t)$ is found, it is subtracted from the original signal $x(t)$, and from this first residue the second IMF $c_2(t)$ is extracted with the same sifting process.

At the end, the result is a set of K IMFs and a final K^{th} residue, from which no new IMF can be extracted, since there are no more oscillations but only a **trend**.

$$\begin{aligned} x(t) &= h_1(t) + r_1(t) && (2.20)\\ &= h_1(t) + h_2(t) + r_2(t) && (2.21)\\ &= \sum_{i=1}^{K} h_i(t) + r_K(t) && (2.22) \end{aligned}$$

Let us summarize the main features of Empirical Mode Decomposition[5]:

- **Locality**: the method operates at the scale of *one* oscillation.
- **Adaptivity**: the method is totally *data-driven*
- **Multiresolution**: the method explores the natural scales of the signal with

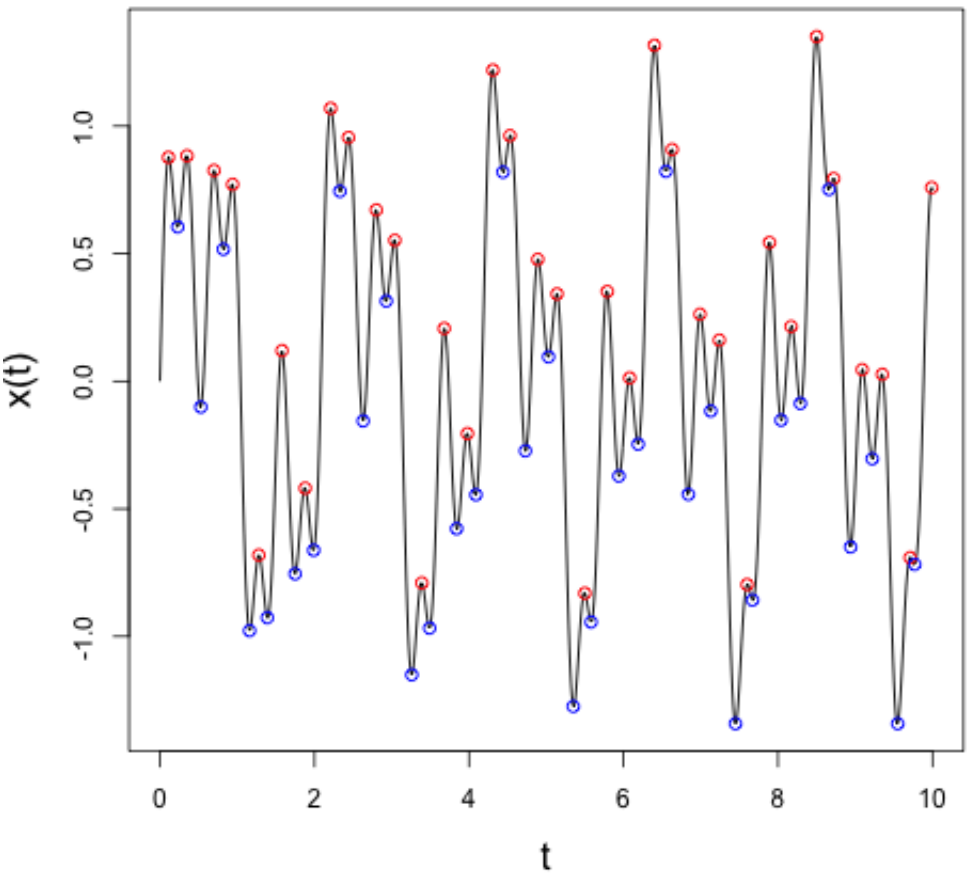


FIGURE 2.41: Red circles are maxima and blue circles are minima, localized by EMD algorithm. In these examples the signal is a sum of three pure sines: $x(t) = 0.7sin(\pi t) + 0.5sin((2\pi t)/0.7) + 0.3sin((2\pi t)/0.3)$.

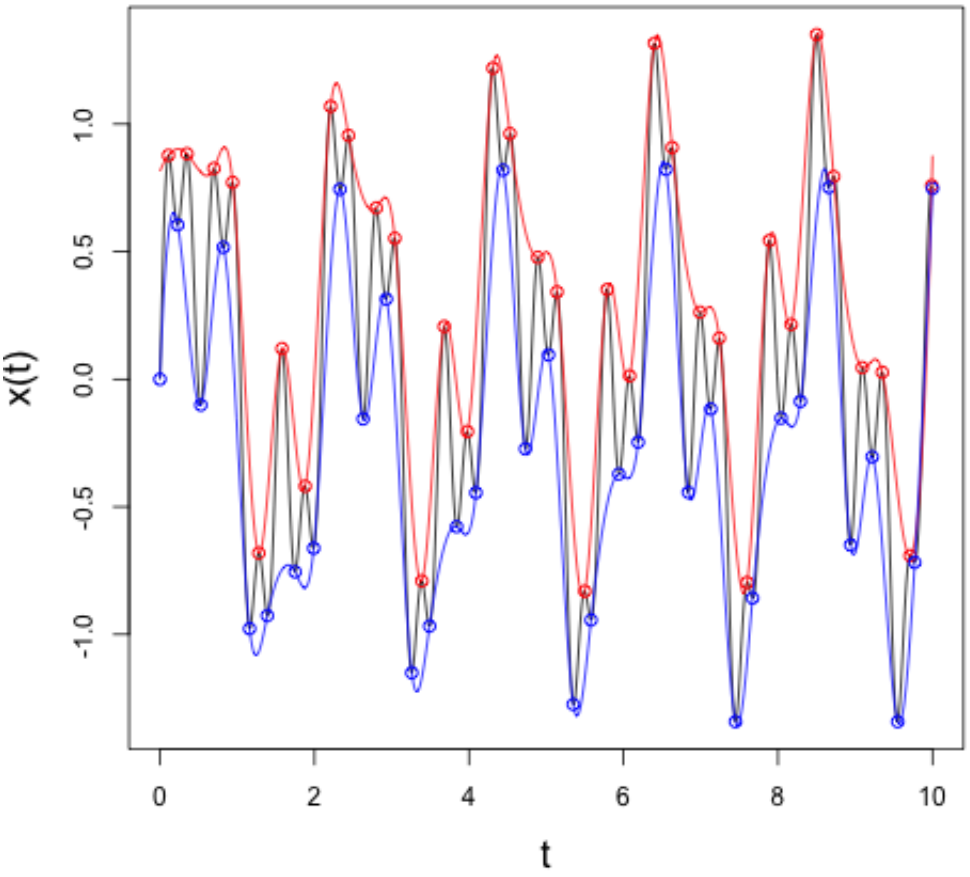


FIGURE 2.42: Maxima and minima are interpolated by a cubic spline. Red circles are maxima, blue circles are minima. The red line is the upper envelope and the blue line is the lower envelope.

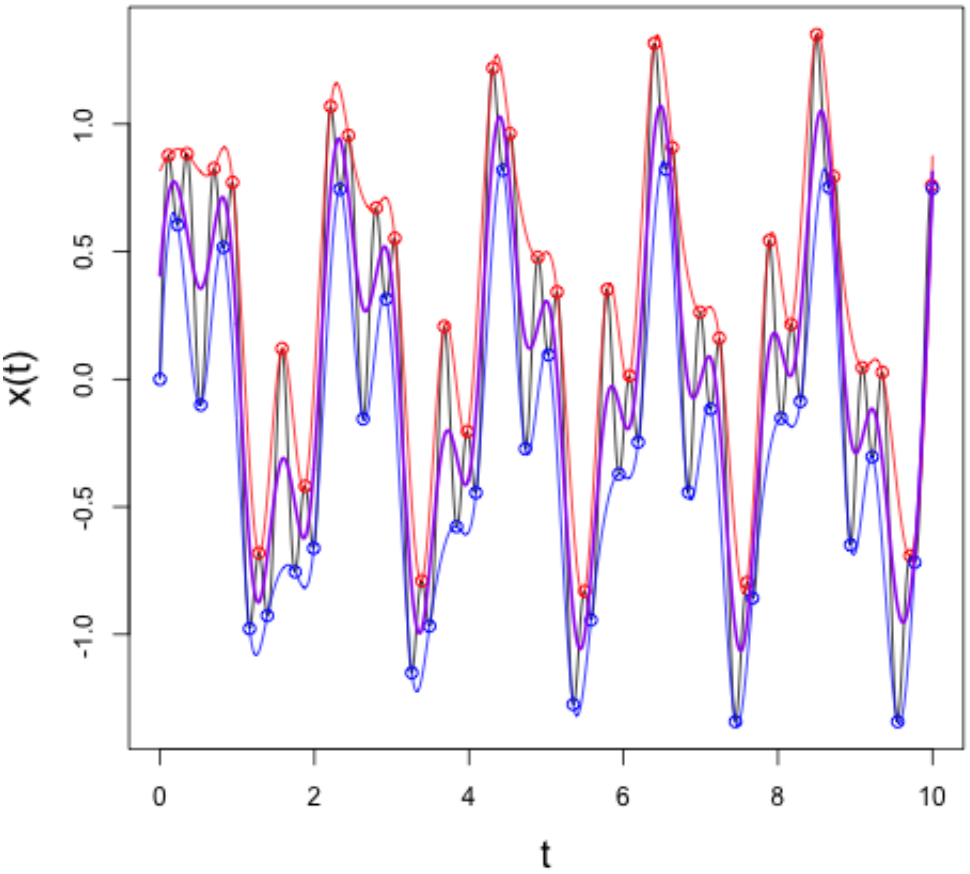


FIGURE 2.43: The average of the upper and lower envelopes is computed. The purple line represents this average vs. t.

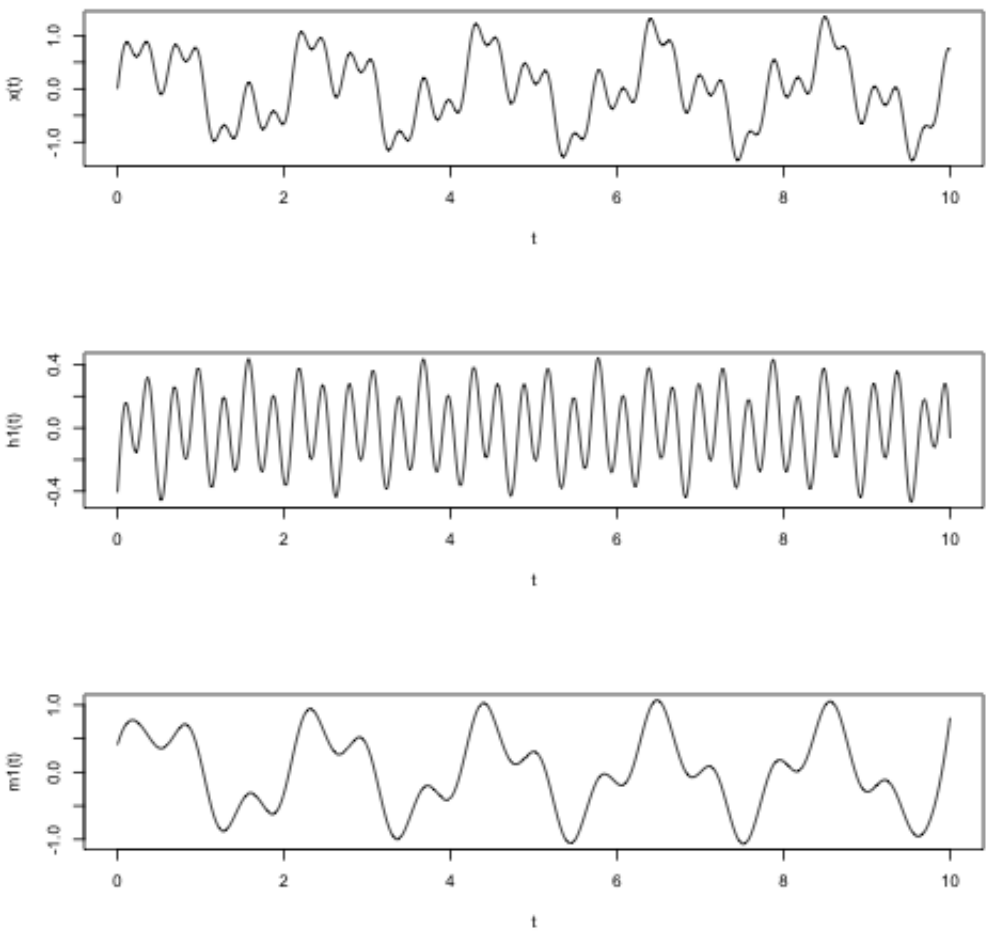


FIGURE 2.44: Subtracting the mean from the original signal $x(t)$, the latter (upper plot) is decomposed into a candidate from the first IMF (middle plot) and a residue which contains all the components at lower frequencies (lower plot).

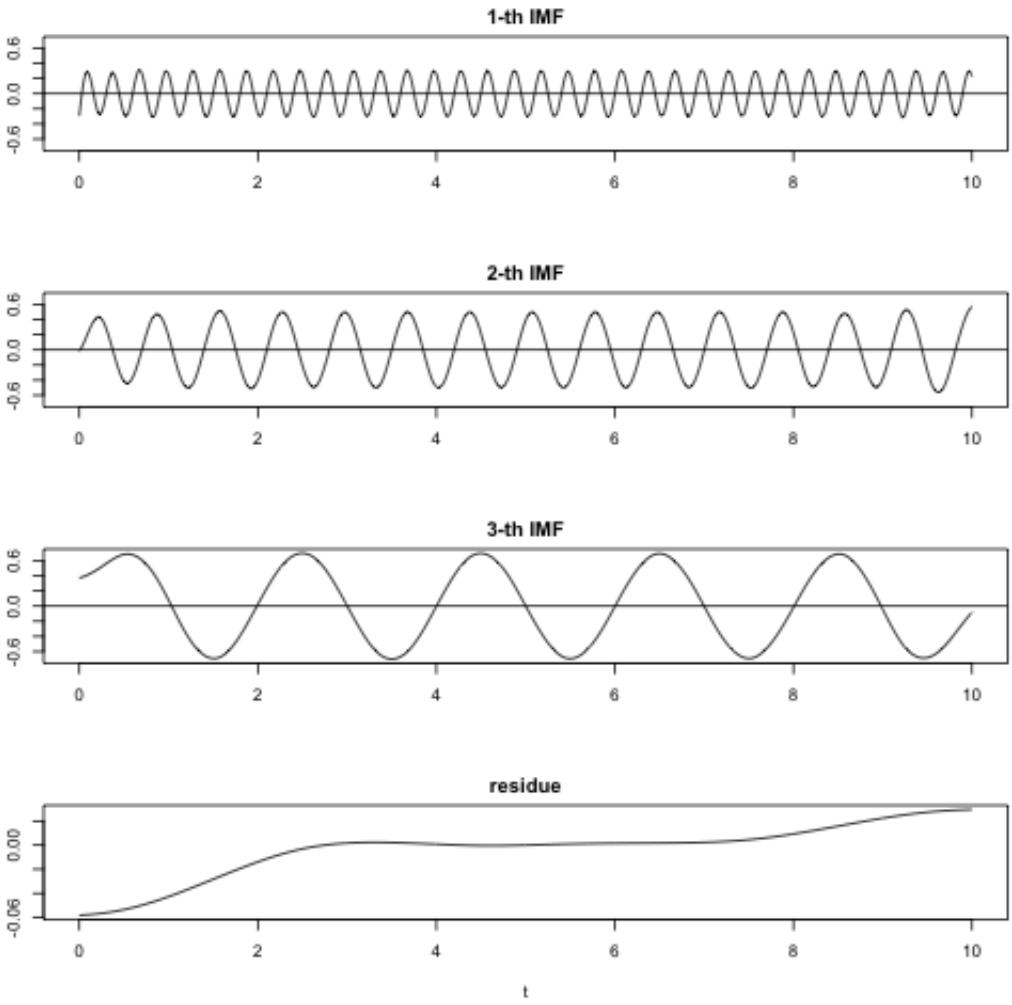


FIGURE 2.45: Empirical Mode Decomposition.

- **no assumptions**: there are no assumptions on the nature of oscillations (e.g. harmonic);
- **no analytic definition**: the decomposition is defined as an output of an algorithm.

A similar decomposition could be performed using a Discrete Wavelet Transform (see Section 2.3), but in this case the scale would be **pre-determined**, instead of adaptive, as in the case of EMD.

Orthogonality and Energy of IMFs

In order to validate the Empirical Mode Decomposition, the set of Intrinsic Mode Functions should be a complete and orthogonal basis. The completeness is ensured, because the sum of the IMFs and the residue reconstructs perfectly the original signal. Concerning orthogonality of the IMFs, Huang et al. (1998) defined an overall index of orthogonality (IO) and a partial index of orthogonality between two modes j and k (IO_{jk}) [20].

Since we are dealing with discrete time series, let us denote the original signal $x(t)$ with x_i ($i = 1, 2, ..., N$). Furthermore, let us denote with $c_j(t)$ the $j-$th IMF (with $j = 1, 2, ..., K$) and with $c_{K+1}(t)$ the final residue $r_K(t)$.

Index of orthogonality:

$$IO = \frac{\sum_{j=1}^{K+1}\sum_{k=1}^{K+1}\int_0^T c_j(t)c_k(t)dt}{\int_0^T x^2(t)dt} \tag{2.23}$$

$$= \frac{\sum_{j=1}^{K+1}\sum_{k=1}^{K+1}\sum_{i=1}^{N} c_{ij}c_{ik}}{\sum_{i=1}^{N} x_i^2}. \tag{2.24}$$

The partial index of orthogonality between modes j and k is:

$$IO_{jk} = \frac{\int_0^T c_j(t)c_k(t)dt}{\left(\int_0^T c_j^2(t)dt + \int_0^T c_k^2(t)dt\right)} \tag{2.25}$$

$$= \frac{\sum_{i=1}^{N} c_{ij}c_{ik}}{\sum_{i=1}^{N}(c_{ij}^2 + c_{ik}^2)}. \tag{2.26}$$

Moreover, it is possible to give a definition of energy of the whole signal, E_x, of each IMF, E_j, and finally, the energy leakage between the $j-$th and the $k-$th IMF, E_{jk}.

$$E_x = \int_0^T x^2(t)dt = \sum_{i=1}^{N} x_i^2, \tag{2.27}$$

$$E_j = \int_0^T c_j^2(t) = \sum_{i=1}^{N} c_{ij}^2, \tag{2.28}$$

$$E_{jk} = \int_0^T c_j(t)c_k(t) = \sum_{i=1}^{N} c_{ij}c_{ik}. \tag{2.29}$$

Theoretically, if the set of functions $c_j(t)$ would be perfectly orthogonal and complete, the indexes of orthogonality should be zero, such as the energy leakage, and the sum of the energy of the functions $c_j(t)$ should be equal to the energy of the starting function $x(t)$:

$$\sum_{j=1}^{K+1} E_j = E_x \tag{2.30}$$

Practically, this never happens and the typical values for the index of orthogonality are about from 10^{-3} to 10^{-2}. This fact implies that the IMFs and the residue are almost orthogonal.

In order to realize a perfect orthogonal set, a new method has been developed by Gram-Schmidt and named 'Orthogonal Empirical Mode Decomposition' [20].

Furthermore, the definition of the IMF energy (Equation 2.28) could be improved, also considering the typical scale of oscillation of a specific Intrinsic Mode Function.

Orthogonal EMD

Aiming at the creation of a set of orthogonal IMFs, the following procedure has been developed by Gram-Schmidt[20]. Let us denote with $c_j(t)$ the IMFs found with the ordinary method (the residue is denoted $c_{K+1}(t)$) and $\hat{c}_j(t)$ the new orthogonal IMFs.

1. The first orthogonal IMF is assumed to be equal to the simple first IMF: $\hat{c}_1(t) = c_1(t)$.

2. Since there is no certainty that $c_2(t)$ is orthogonal to $c_1(t)$, then from $c_2(t)$ partial $\hat{c}_1(t)$ is removed.

   $$\hat{c}_2(t) = c_2(t) - \beta_{21}\hat{c}_1(t), \tag{2.31}$$

   where $\hat{c}_2(t)$ is the second orthogonal IMF and β_{21} is the orthogonality coefficient between $\hat{c}_2(t)$ and $c_1(t)$.

3. Imposing the orthogonality between $\hat{c}_1(t)$ and $\hat{c}_2(t)$, integrating in time:

$$0 = \int_0^T \hat{c}_1(t)\hat{c}_2(t)dt = \int_0^T \hat{c}_1(t)c_2(t)dt - \beta_{21}\int_0^T \hat{c}_1^2(t)dt, \tag{2.32}$$

so we can find the value for β_{21}:

$$\beta_{21} = \frac{\int_0^T \hat{c}_1(t)c_2(t)dt}{\int_0^T \hat{c}_1^2(t)dt} = \frac{\sum_{i=1}^N \hat{c}_{1i}c_{2i}}{\sum_{i=1}^N \hat{c}_{1i}}. \tag{2.33}$$

4. Iterating this procedure one can find all the orthogonalized IMFs and the orthogonality coefficients:

$$\hat{c}_{k+1}(t) = c_{k+1} - \sum_{j=1}^{k} \beta_{k+1,j}\hat{c}_j(t), \tag{2.34}$$

$$\beta_{k+1,j} = \frac{\int_0^T c_{k+1}(t)\hat{c}_j(t)dt}{\int_0^T \hat{c}_j^2(t)} = \frac{\sum_{i=1}^N c_{k+1,i}\hat{c}_{ji}}{\sum_{i=1}^N \hat{c}_{ji}^2}. \tag{2.35}$$

Comparison between EMD and OEMD

The values of orthogonality obtained by OEMD are actually smaller than those obtained by EMD, but the result is not so significant. Moreover, the comparison between the two residuals shows that those obtained by EMD are much more smooth than those obtained by OEMD, which have a content of higher frequencies. This fact is due to the subtraction for each IMF and residue of all the IMFs at higher frequencies.

In order to understand this aspect, see Figures 2.46 and 2.47. In this example the signal is a superimposition of three pure sines, so its local mean is zero-valued. The residue, that in this case should be zero, exhibits a fake trend. However, the amplitude of the latter is about 5% of the original signal. Other trials showed that the extraction of a existing trend is effectively performed by EMD, but these results are not presented in this work.

Another possible approach, to perform the computation of orthogonal intrinsic mode functions, is that to assume as first orthogonal IMF the residue instead of the mode at highest frequency. In this way, the residue of the new set of orthogonal IMFs remains smooth. However, this procedure involves that the IMFs at higher frequency present a trend and their local mean is no longer zero-valued.

In conclusion, the use of OEMD is strictly bound to the kind of problem that has to be studied. For example, OEMD is not so effective to extract the trend of a time series.

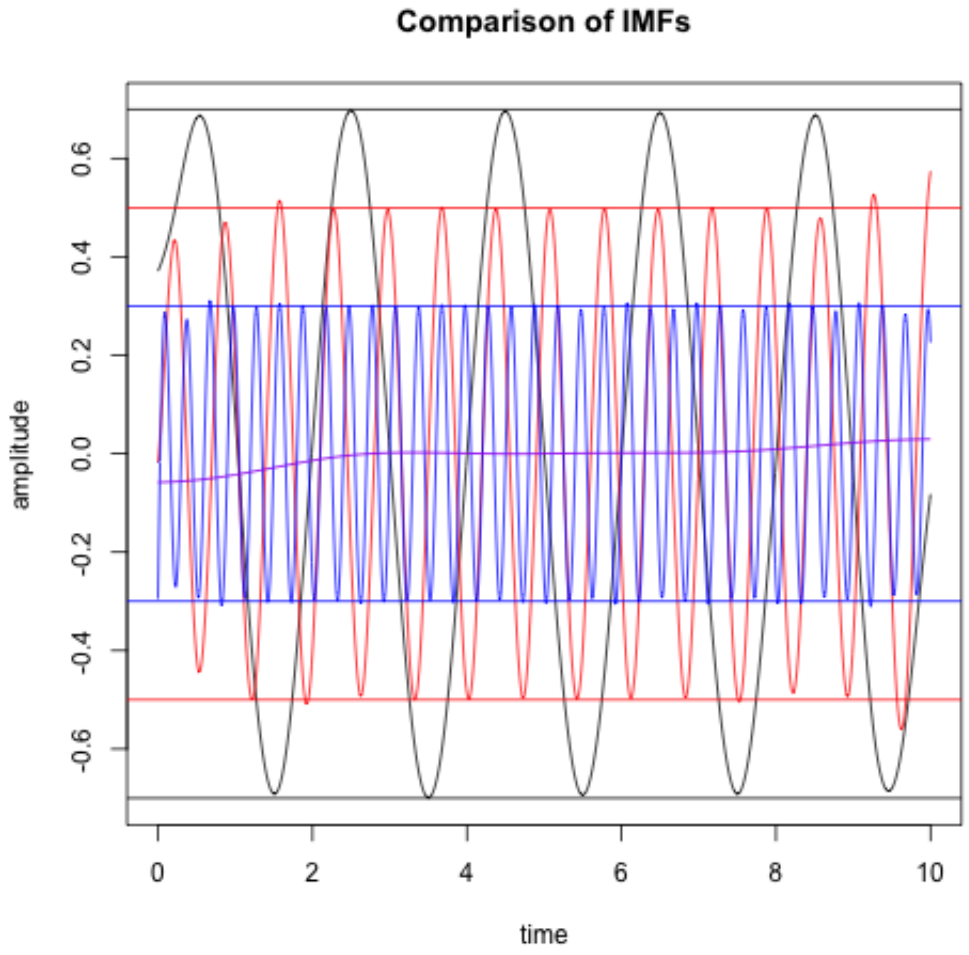


FIGURE 2.46: Superimposition of the three IMFs and the residue. Blue line: first IMF; red line: second IMF; black line: third IMF; purple line: residue. Straight lines represent the amplitude of the real pure sines used to composed the original signal $x(t)$.

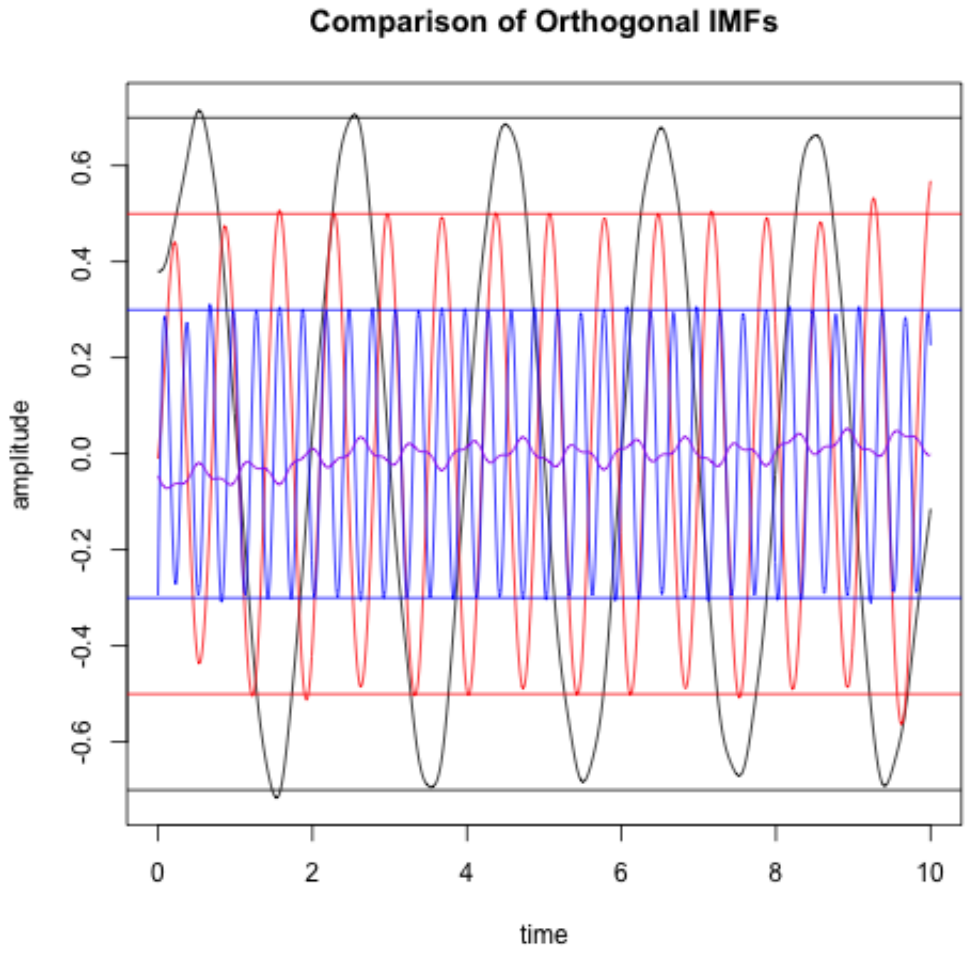


FIGURE 2.47: Superimposition of the three orthogonal IMFs and the residue. Blue line: first orthogonal IMF; red line: second orthogonal IMF; black line: third orthogonal IMF; purple line: residue. Straight lines represent the amplitude of the real pure sines used to composed the original signal $x(t)$.

Hilbert Spectrum

After the Empirical Mode Decomposition, a signal $x(t)$ is decomposed in a set of K Intrinsic Mode Functions at different scales and a residue, which represents the 'trend' of the original signal.

Since there are no assumptions on the nature of the oscillations, the frequency of the oscillations can vary with time. It would be useful at this point, to define a method in order to study the frequency of each mode as function of time. The proposed method utilizes the Hilbert Transform and the Hilbert Spectrum.

Given a real-valued signal $x(t)$, the Hilbert Transform $y(t)$ is defined as[15]:

$$y(t) = \frac{1}{\pi}\, p.v. \int_{-\infty}^{+\infty} \frac{x(s)}{t-s} ds. \tag{2.36}$$

The complex-valued signal $z(t)$ is defined as:

$$z(t) = x(t) + iy(t). \tag{2.37}$$

Writing $z(t)$ in polar coordinates, its amplitude and phase are determined:

$$z(t) = A(t)e^{\theta(t)}. \tag{2.38}$$

where $A(t) = ||z(t)|| = \sqrt{x(t)^2 + y(t)^2}$ and $\theta(t) = arctan\left(\frac{y(t)}{x(t)}\right)$. Since the phase is defined as function of time, the change of phase in a time interval, namely the time derivative of phase, is the instantaneous frequency:

$$\omega(t) = \frac{d\theta(t)}{dt} \tag{2.39}$$

Finally one can plot the amplitude as function of time and instantaneous frequency, in the time-frequency space, to obtain the Hilbert Spectrum. To understand the properties of Hilbert Spectrum, let us define, for instance, a sinusoidal signal $x(t)$ with amplitude and frequency function of time (Equations 2.40 and 2.41, and Figure 2.48):

$$x(t) = A(t)sin(\omega(t)t), \tag{2.40}$$

$$A(t) = 0.8 + \sqrt{t} \qquad \text{and} \qquad \omega(t) = 2\pi(2 + 0.5t). \tag{2.41}$$

Hilbert analysis applied to this particular $x(t)$ provides information regarding frequency and amplitude: points in the time-frequency space represent instantaneous frequency and their intensity represents the value in amplitude at that time.

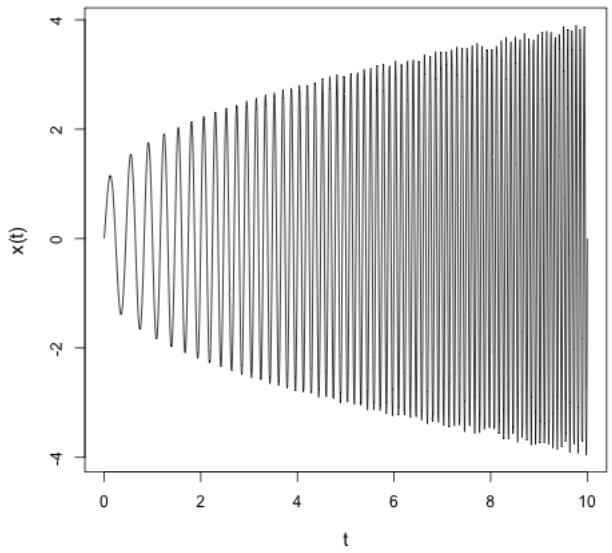


FIGURE 2.48: A sinusoidal signal $x(t) = A(t)sin(\omega(t)t)$ with $A(t) = 0.8 + \sqrt{t}$ and $\omega(t) = 2\pi(2 + 0.5t)$.

In order to evaluate this, see Figure 2.49. Frequency is graphed as a linear function of time, while amplitude increases with time as a square root (Figure 2.50), as initially set.

It is very important to point out that **no information** on the composition of the signal $x(t)$ was used in the algorithm: the trend of amplitude and frequency were extracted directly from the signal $x(t)$ by empirical mode decomposition. Furthermore, the resolution in time-frequency domain, does not vary significantly with frequency, as in the case of wavelet spectrum.

2.4.2 Application of EMD to Irradiance data

The application of Empirical Mode Decomposition to the Irradiance data gives firstly a multiresolution analysis.

From the Intrinsic Mode Functions at lowest scales, i.e. highest frequencies, one can deduce that the fast oscillation are larger during solar maxima (Figure 2.51).

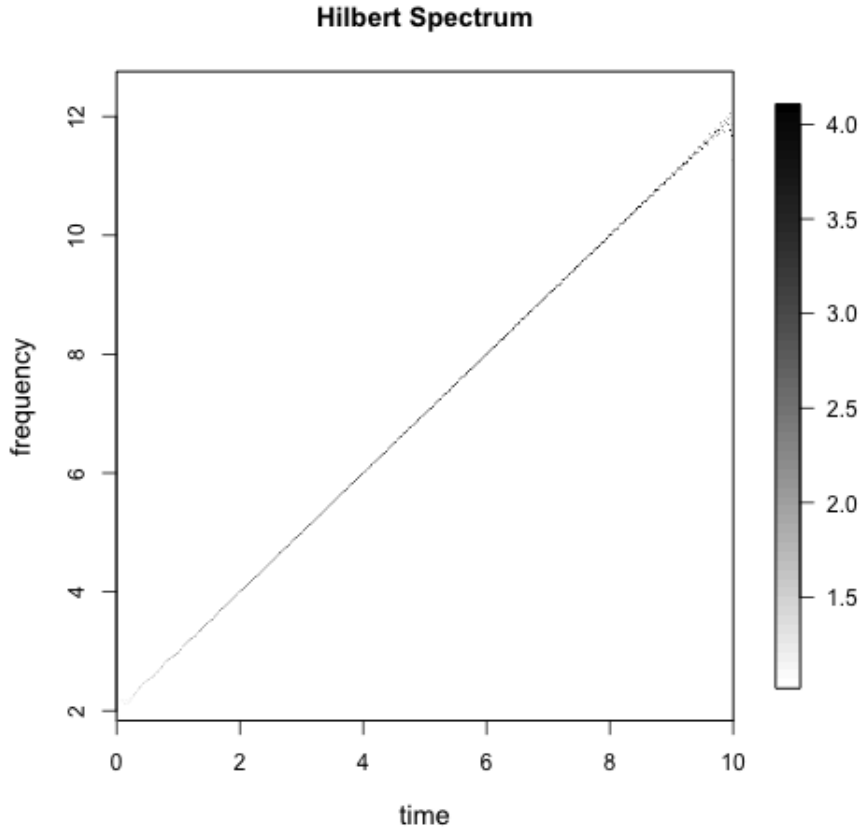


FIGURE 2.49: Hilbert Spectrum of signal $x(t)$. The frequency is presented as linear function of time and amplitude is presented here as intensities of black.

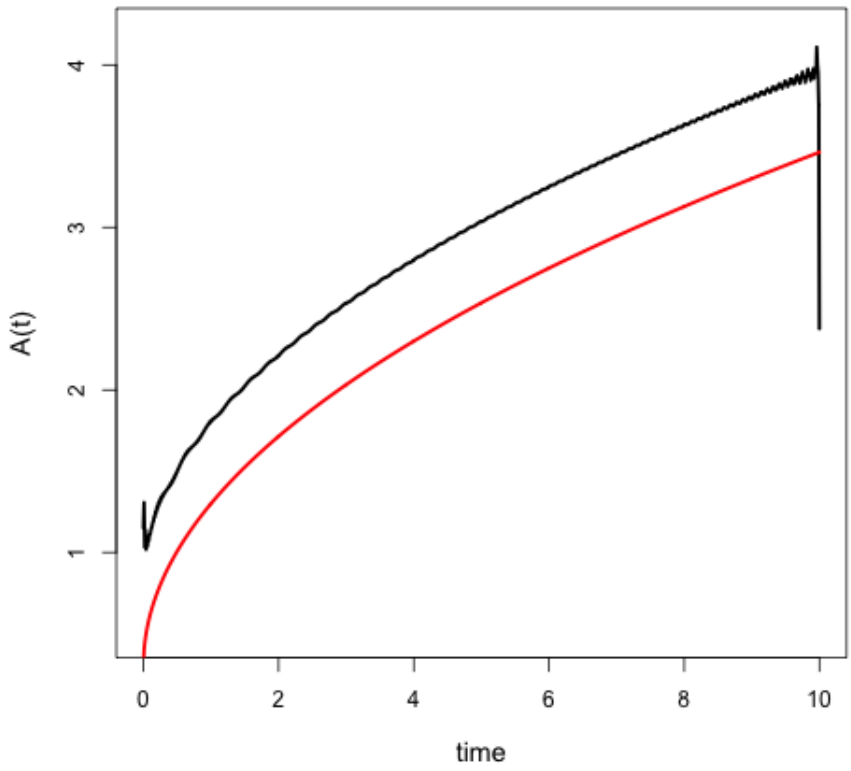


FIGURE 2.50: Amplitude vs. time obtained by Hilbert analysis (black line). The correct amplitude function is plotted in red and shifted downwards by one unit.

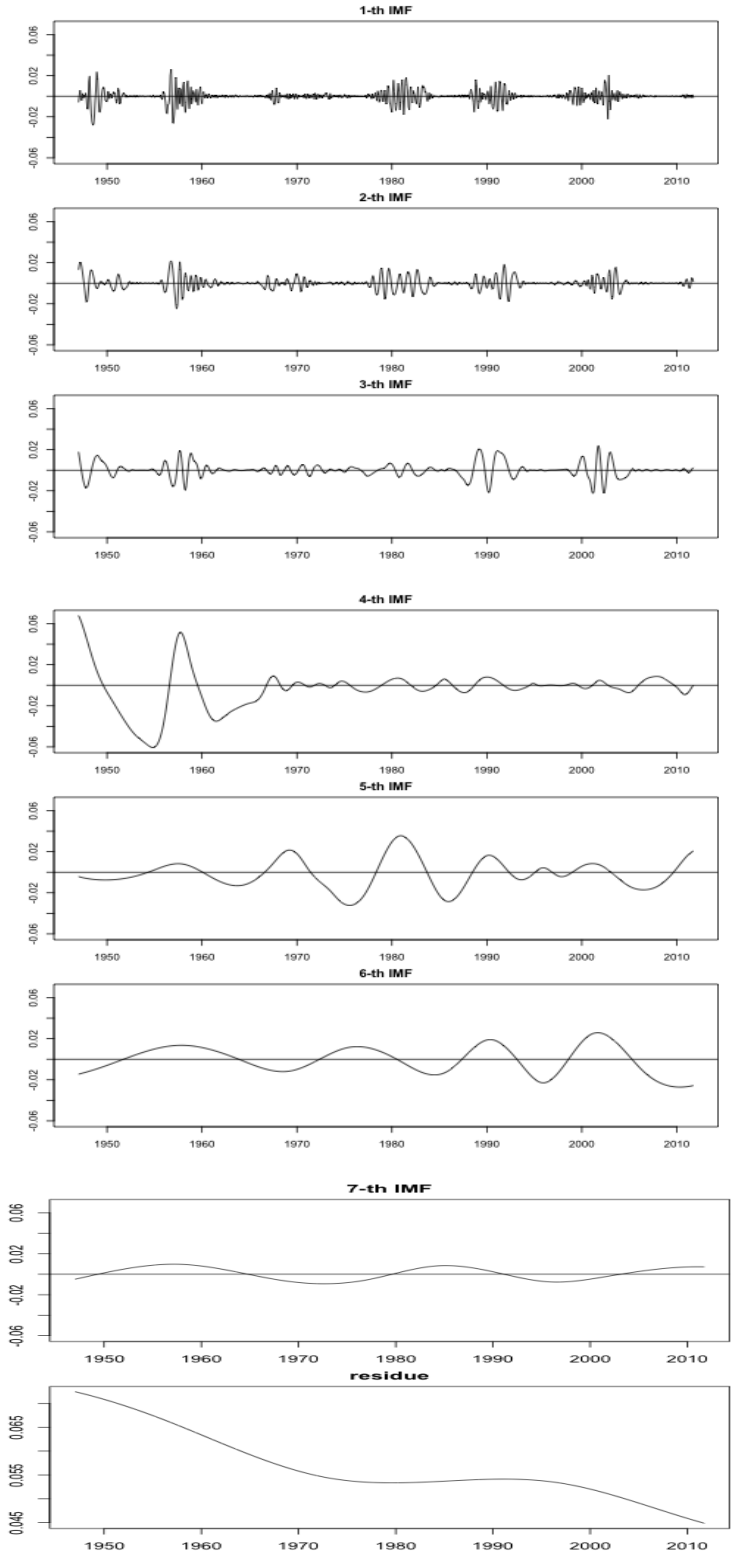


FIGURE 2.51: Empirical mode decomposition of XUV solar irradiance. Time sampling: monthly.

2.4.3 Application of EMD to meteorological data

The application of Empirical Mode Decomposition to the meteorological time series provides a multiresolution decomposition.

The second step consists in computing the correlation between every IMF of the temperature at each latitude and altitude, and every IMF of the solar irradiance.

It is expected that the correlation coefficient should be low for most of the cases, especially for correlation test between time series at very different scales. The hope is that, for IMFs with a comparable scale, the correlation test could reveal some connections of solar activity with the temperature time series.

An interesting result has been obtained, observing the Hilbert spectrum of global seasonal temperature between 850 hPa and 300 hPa (about 1.5 and 9 km in altitude). The Hilbert spectrum is reported in Figure 2.52. The red line indicates the frequency with period equal to 11 years. The fifth IMF of temperature exhibits a quasi-constant periodicity at that frequency for most of the time interval.

Plotting only the fifth IMF versus solar irradiance, they seem to have the same temporal pattern (Figure 2.53).

The evaluation of the correlation values between all the IMF of temperature and irradiance pointed out that there is a discrete correlation (around 0.8) between the fifth IMF of temperature and the fourth IMF of solar irradiance. These are illustrated in Figure 2.54.

The superimposition is not perfect: before around 1985, solar extrema seem to precede those of temperature, whereas, after 1985 the opposite happens. The differences in time between solar and temperature extrema are reported in Figure 2.55. According to this behaviour, an important aspect that could be noticed is that atmospheric temperature before 1980-1985 is nearly constant with a slightly decrease. After 1985 its trend is markedly increasing. This behaviour has been associated with the reduction of sulphur oxide emissions, which had caused acid rain and, for this fact, they had been controlled by environmental policies. However, these emissions had made cloud more reflective, limiting the effects of the CO_2, and so, their reduction caused the decrease of terrestrial albedo and the subsequent increase in temperature.

The evaluation of energy of the IMFs and the residue was obtained according to [20], as explained in Section 2.4.1. The energy of the components of temperature at 850-300 hPa is graphed in Figure 2.56: the energy of the fifth IMF is about the 8.3% of the energy of the whole signal.

The analysis of correlation between IMFs of irradiance and ozone level

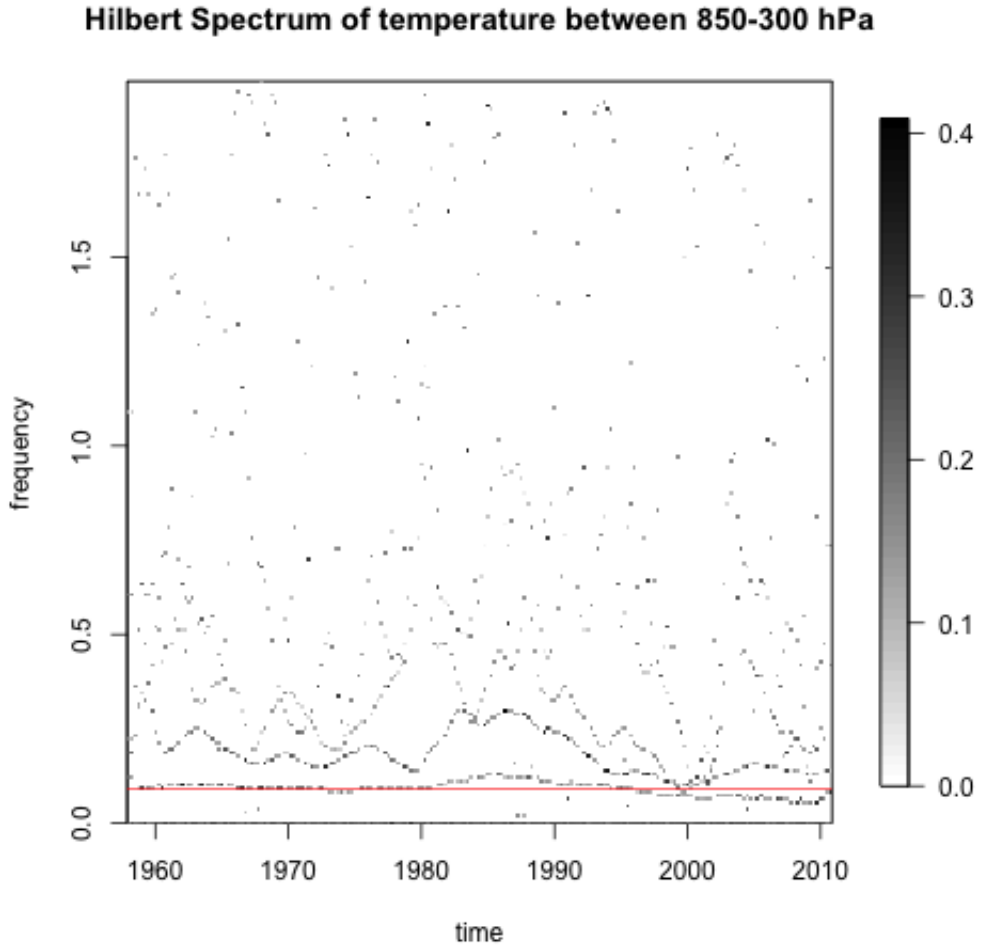


FIGURE 2.52: Hilbert spectrum of temperature between 850 and 300 hPa. the red line indicates the period at 11 years. The mode that follows the red line indicates that its frequency remains quasi-constant.

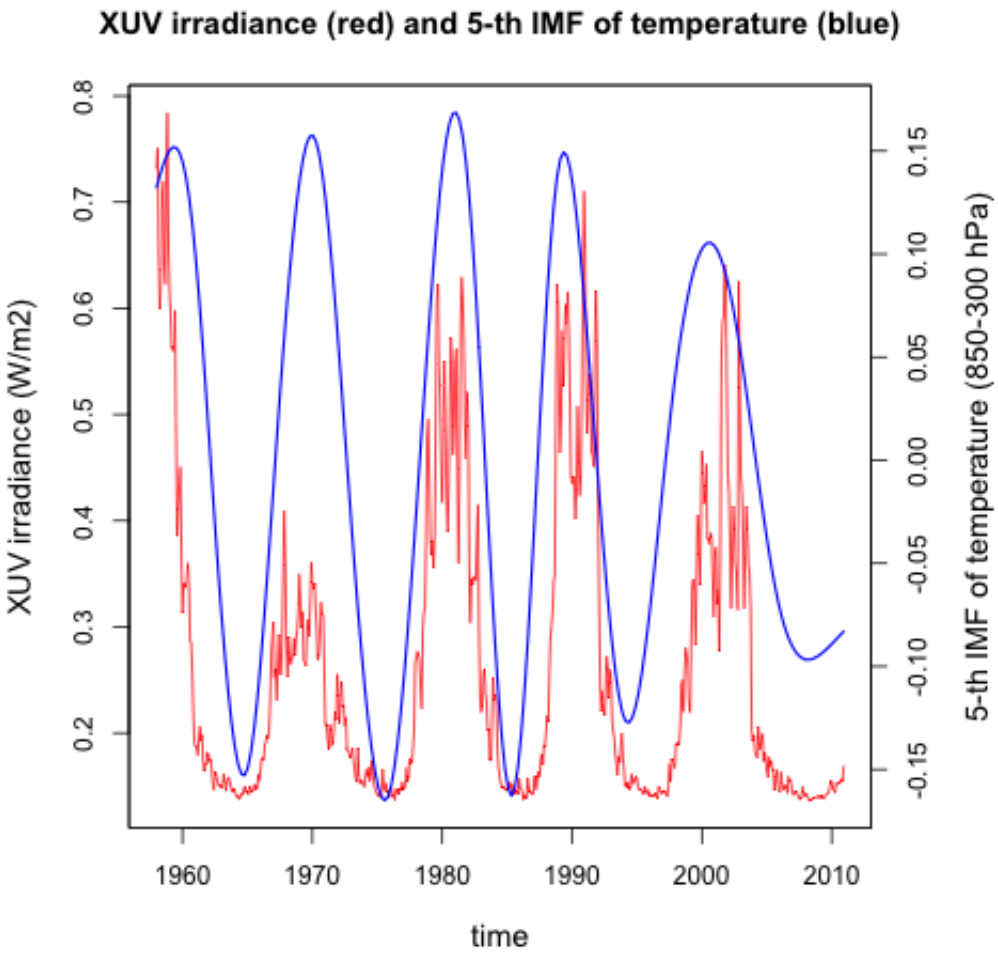


FIGURE 2.53: XUV irradiance (red line) and 5^{th} IMF of temperature between 850 and 300 hPa (blue line).

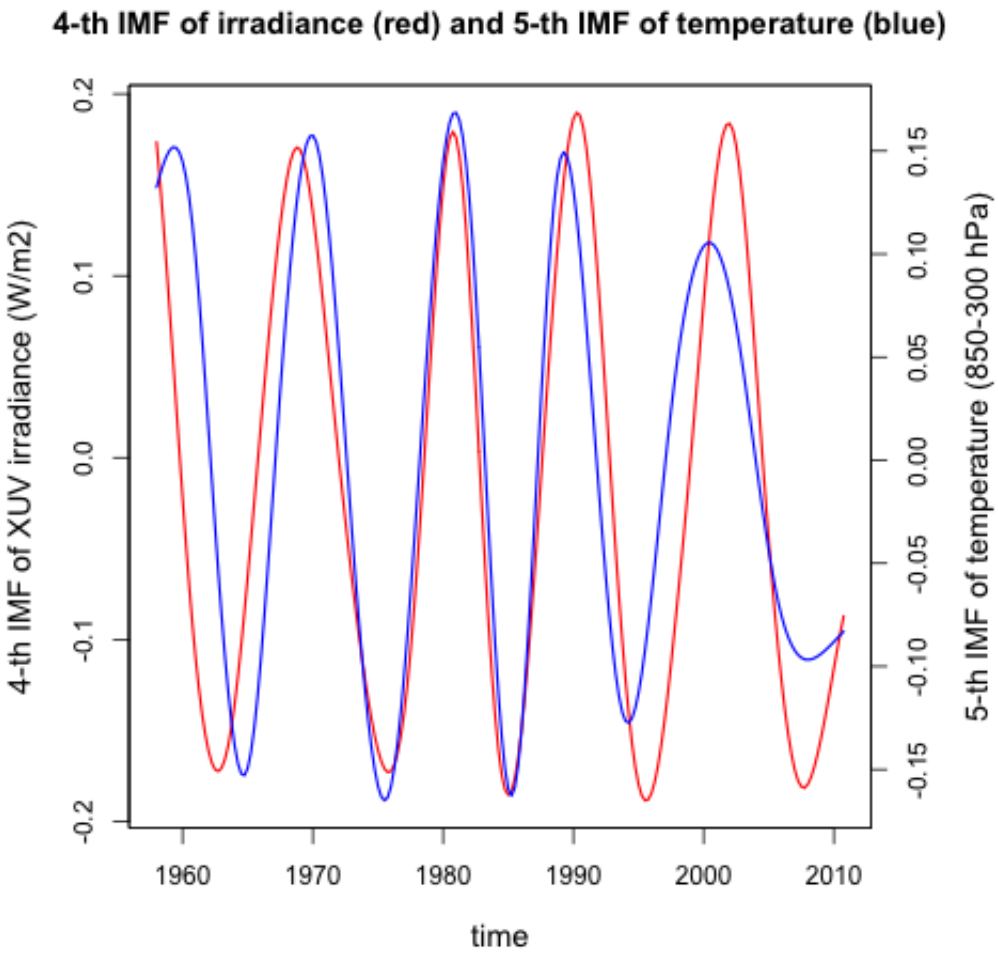


FIGURE 2.54: 4^{th} IMF of XUV irradiance (red line) and 5^{th} IMF of temperature between 850 and 300 hPa (blue line).

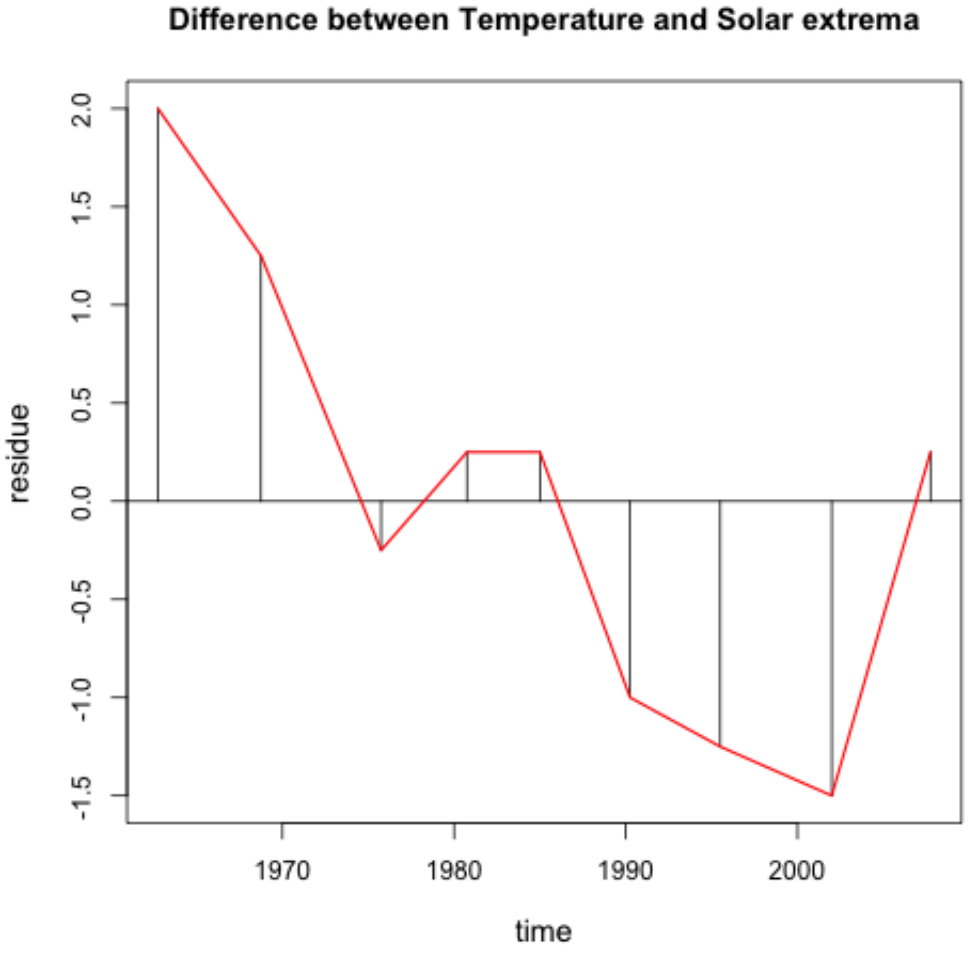


FIGURE 2.55: Differences in time between solar and temperature extrema: a positive difference indicates a foregoing solar extremum whereas a negative differences indicates a foregoing temperature extremum.

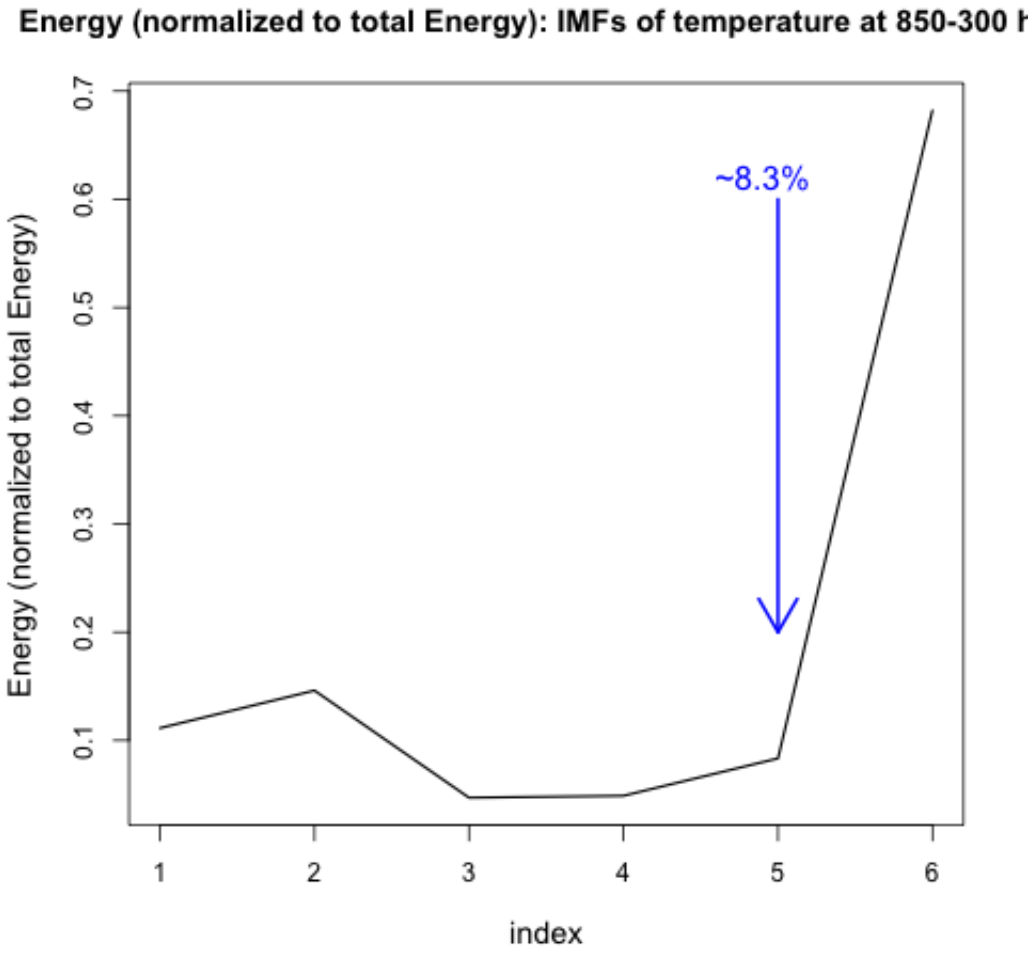


FIGURE 2.56: Energy of the components of temperature at 850-300 hPa.

does not reveal any significant result. Instead, there are some results of interest with regard to residues of irradiance and ozone level. The superimposition of these residuals is illustrated in Figure 2.57. Here, the columnar ozone amount seems to follow the temporal pattern of irradiance, but with a time lag.

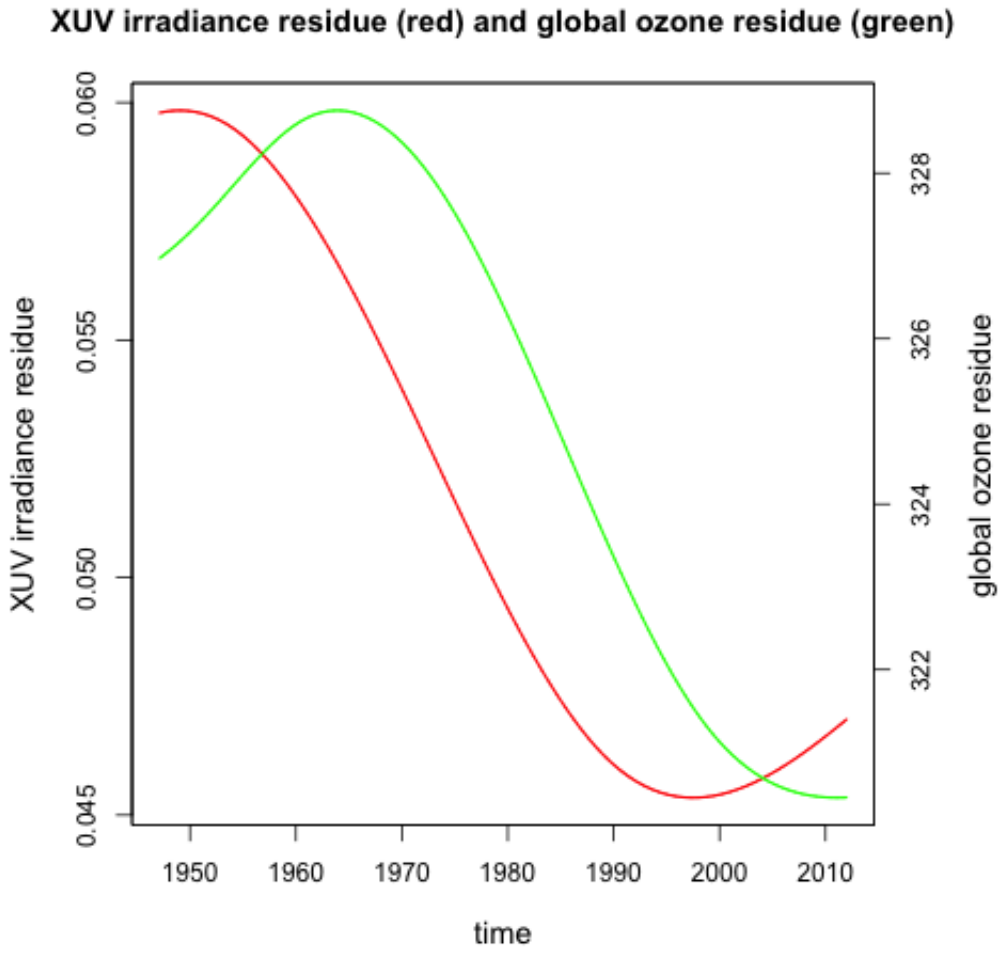


FIGURE 2.57: XUV irradiance residue (red) and global ozone level residue (green).

Chapter 3

Discussion and Conclusions

In this chapter the main conclusions of the analysis of meteorological and solar activity descriptors are drawn.

The main results of the correlation tests and Fourier analysis are the following.

1. Variability in XUV and UV band is much higher than VIS and IR.
2. There is no evidence of linear correlation between irradiance and temperature.
   - neither direct,
   - nor at different scale and time delays,
   - or analyzing the derivative of temperature,
3. Student's t test cannot reject the hypothesis of correlation.
4. There are some evidences of possible periodicities due to solar activity.
5. ozone level show periodicities around 11 years at low latitudes (between $5°N$ and $25°S$)

However, in the Fourier analysis the temporal information is lost, so it is necessary to adopt a more powerful tool to perform a time-frequency analysis.

The wavelet analysis performs a time-frequency analysis computing many convolutions in time domain between the signal and a mother function, which is scaled a number of times in order to explore the frequency domain. This analysis aims at expressing the frequency components as function of time. However, the resolution in the time-frequency domain depends on frequency. The main result of the application of wavelet analysis to irradiance and meteorological time series are the following:

1. The adopted mother wavelets, which resulted more suitable have been:
   - Morlet-6 and DOG-20: for, respectively, daily and monthly data.
   - DOG-10: seasonal data
2. The results in terms of temperature anomaly were that:
   - Long periods are revealed and time localized.
   - Annual frequency events can not be directly associated to solar cycle.
3. Ozone level shows a 11-year periodicity, especially at low latitudes, which could be associated with solar irradiance.

Wavelet analysis has some limitations: the choice of the scales and the mother wavelet are somehow arbitrary, and this fact could affect the validity of the obtained results.

Another methodology, which has been recently used in many scientific fields, is the Empirical Mode Decomposition (EMD) and the Hilbert transform. Its purpose is to perform first a signal decomposition in a set of Intrinsic Mode Functions (IMF) and secondly, a time-frequency analysis. The main advantages of this methodology are[5]:

- **Locality**: the method operates at the scale of *one* oscillation.
- **Adaptivity**: the method is totally *data-driven*.
- **Multiresolution**: the method examinates the natural scales of the signal.
- **No assumptions**: there are no assumptions on the nature of oscillations (e.g. harmonic).
- **No analytic definition**: the decomposition is defined as an output of an algorithm.

The use of the EMD revealed that a component of the temperature between 850 and 300 hPa could follow the temporal pattern of solar irradiance with a periodicity of about 11 year. The quantification of the amplitude of this component with respect to the complete signal was calculated with the sum of the squared amplitudes[20]. According to this approach, the energy of this component (defined as sum of the squared amplitudes) could represent $\sim$8% with respect to the energy of the original time series.

A possible way to further improve this analysis is first of all, to use time series with better resolution and temporal coverage, and secondly, to detrend the time series by EMD, and to perform a Wavelet analysis on detrended data series. Furthermore, it may be suitable to refine the definition of IMF energy considering the scale of the IMF oscillations, in order to improve the quantification of the results obtained by EMD.

It could be useful to consider the equation that describes the total energy of terrestrial atmosphere per unit of mass, in order to understand the different ways where solar irradiance variations could manifest:

$$E = c_V T + gz + \frac{1}{2}|\mathbf{v}|^2 + Lq, \tag{3.1}$$

where c_V is specific heat at constant volume, T is temperature, g is gravitational acceleration, z is altitude, $\mathbf{v}$ is wind velocity, L is latent heat of condensation and q is specific humidity.

Solar irradiance may affect all these term contributing to the total energy of atmosphere, and the analysis of the temperature only reveals just one component of the response of Earth's atmosphere to solar variability. A complete study, should consider also the other contributions, especially those dealing with total wind speed and specific humidity. Therefore, another possible future development could be the application of the methodologies presented in this work to such atmospheric parameters.

Acknowledgements

The completion of this dissertation would not have been possible without the support, the assistance and the encouragement of the following people, to whom I would like to express my sincere gratitude.

First, I owe my deepest gratitude to **Prof. Mauro Messerotti** for having accepted the role of supervisor, following the planning and the realization of this thesis from the beginning. His competence have been fundamental and the accuracy put in all the aspects of the draft of this dissertation has given the confidence in the effectiveness of this work.

Secondly, I wish to thank **Prof. Fulvio Stel**, who raised my attention to the Physics of Atmosphere and advised me in the final moment of the course of my studies. His awareness and his reassurance have been encouraging during the realization of this work and have made possible not to lose sight of our objectives.

I would like to express my gratitude also to **Prof. Fulvio Crisciani** for his collaboration and constructive criticism.

I acknowledge also **Dott. Fabio Raicich** of ISMAR-Trieste for having suggested the application of Empirical Mode Decomposition and **Prof. Edoardo Milotti** for having provided research material on wavelet analysis.

I would also thank **Prof. Silvio Modesti**, **Prof. Andrea Bressan**, **Dott. Silvana Palmiero** and **Dott. Nadia D'Antoni** for all the aspects regarding academic, technical and administrative questions.

Finally, I am thankful to my family and all my friends for having supported me throughout the course of my studies.

Bibliography

[1] P. Addison. The little wave with the big future. *Physics World*, March 2004.

[2] AISN Software Inc. *AutoSignal: Pioneering Automated Signal Analysis and Advanced Filtering*.

[3] L. A. Barford, R. Shane Fazzio, and D. R. Smith. An introduction to wavelets. *Instruments and Photonics Laboratory*, pages HPL–92–124, September 1992.

[4] J. Beer. Long-term indirect indices of solar variability. *Space Science Reviews*, **94**:53–66, 2000.

[5] P. Flandrin. Empirical mode decomposition vs. wavelets for the analysis of scaling processes, 2010.

[6] R. Goody. *Principles of Atmospheric Physics and Chemistry*. Oxford University Press, first edition, 1995.

[7] T. C. Gouhier and A. Grinsted. *Package 'biwavelet': Conduct univariate and bivariate wavelet analyses*, May 2012.

[8] L. J. Gray, J. Beer, M. Geller, J. D. Haigh, Lockwood M., K. Matthes, U. Cubasch, D. Fleitmann, G. Harrison, L. Hood, J. Luterbacher, G. A. Meehl, D. Shindell, B. van Geel, and W. White. Solar influences on climate. *Rev. Geophys.*, **48**, 2010. RG4001, doi:10.1029/2009RG000282.

[9] J. D. Haigh. Solar influences on climate. *Imperial College London, Grantham Institute for Climate Change*, February 2011.

[10] J. D. Haigh, M. Lockwood, and M. S. Giampapa. *The Sun, Solar Analogs and the Climate*. Springer, first edition, 2004.

[11] N. E. Huang, Man-Li C. Wu, S. R. Long, S. S. P. Shen, W. Qu, P. Gloersen, and K. L. Fan. A confidence limit for the empirical mode decomposition and hilbert spectral analysis. *Proc. R. Soc. Lond. A.*, **459**:2317–2345, 2003.

[12] N.E. Huang, Z. Shen, S.R. Long, M.L. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu. The empirical mode decomposition and hilbert spectrum for nonlinear and nonstationary time series analysis. *Proc. Roy. Soc. London A.*, **454**:903–995, 1998.

[13] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, A. Bunde, S. Havlin, and H. E. Stanley. Multifractal detrended fluctuation analysis of nonstationary time series. *arXiv:physics/0202070v1*, February 2002.

[14] J. T. Kiehl and K. E. Trenberth. Earth's annual global mean energy budget. *Bull. Amer. Meteor. Soc.*, **78**:197–208, 1997.

[15] D. Kim and H. S. Oh. EMD: A package for empirical mode decomposition and hilbert spectrum. *The R Journal*, **1/1**, May 2009. ISSN 2073-4859.

[16] D. Kim and H. S. Oh. *Package 'EMD': Empirical Mode Decomposition and Hilbert Spectral Analysis*, January 2013.

[17] J. Kirkby. Cosmic rays and climate. *Surv. Geophys.*, **28**:333–375, 2007.

[18] G. Kopp. Overview and advances in solar radiometry for climate studies, September 2008.

[19] P. Kumar and E. Foufoula-Georgiou. Wavelet analysis for geophysical applications. *Reviews of Geophysics*, **35**(4):385–412, November 1997. Paper number 97RG00427.

[20] Tian li Huang, Wei-Xin Ren, and Meng lin Lou. The orthogonal hilbert-huang transform and its application in earthquake motion recordings analysis. *The* 14^{th} *World Conference on Earthquake Engineering*, October 2008.

[21] J. M. Pap and P. Fox. *Solar Variability and Its Effects on Climate*. AGU Books Board, first edition, 2004.

[22] S. K. Solanki. Solar variability and climate change: Is there a link? *Harold Jeffreys Lecture*, **43**:5.9 – 5.13, October 2002.

[23] Space Environment Technologies. *SOLAR2000 Professional Grade v2.24 Installation Manual and User Guide*, September 2004.

[24] W. K. Tobiska. Status of the solar2000 solar irradiance model. *Phys. Chem. Earth (C)*, **25**(5-6):383–386, 2000.

[25] W. K. Tobiska. Solar2000 irradiances for climate change research, aeronomy and space system engineering. *Advances in Space Research*, **34**:1736–1746, 2004.

[26] W. K. Tobiska. Systems-level space environment specification for satellite and ground system operations. *AIAA-2005-0069*, 2005.

[27] W. K. Tobiska. Solar and geomagnetic space environment specification for operations. *AIAA-2006-0471*, 2006.

[28] W. K. Tobiska. SOLAR2000 v2.30 and SOLARFLARE v1.01: New capabilities for space system operations. *AIAA-2007-0495*, 2007.

[29] W. K. Tobiska. Mitigating orbit planning, satellite operations, and communication surprises from adverse space weather. *AIAA-2008-0453*, 2008.

[30] W. K. Tobiska and S. D. Bouwer. New developments in solar2000 for space research and operations. *Advances in Space Research*, **37**:347–358, 2006.

[31] W. K. Tobiska, T. Woods, F. Eparvier, R. Viereck, L. Floyd, S. D. Bouwer, G. Rottman, and O. R. White. The SOLAR2000 empirical solar irradiance model and forecast tool. *Journal of Atmospheric and Solar-Terrestrial Physics*, **62**:1233–1250, 2000.

[32] C. Torrence and G. P. Compo. A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, October 1997.

[33] W. N. Venables, D. M. Smith, and the R Core Team. *An Introduction to R*, October 2012.

[34] J. R. Ziemke, S. Chandra, and P. K. Bhartia. Two new methods for deriving tropospheric column ozone from toms measurements: The assimilated uars mls/haloe and convective-cloud differential techniques. *J. Geophys. Res*, **103**(22):115–22,127, 1998.

Appendix A

Solar Irradiance modelling by SOLAR2000

In reference at Chapter 1, SOLAR2000 has a five-phase development process: (1) model definition and design; (2) modeling of soft X-rays (XUV) and extreme ultraviolet (EUV) irradiances; (3) modeling of far ultraviolet (FUV) irradiances; (4) modeling of UV irradiances; and (5) representation of visible, infrared, and the total solar irradiance (TSI). These phases of development are summarized in Table A.1.

Furthermore, SOLAR2000 produces also a number of solar irradiance proxies, that can be used by scientific community. A summary of these proxies is provided in Table A.2.

Phase	λ	Source data	Timeframe	Proxies
2	XUV	SOLRAD	1977-1980	*Coronal:*
	1-10 nm	rockets	1976-1998	0.1-0.8 nm
		YOHKOH	1991-	F10.7 (flux at 1.07×10^8nm)
		SNOE	1998-	Coronal holes images
		TIMED	2001-	
2	EUV	SOLRAD	1977-1980	*Coronal:*
	10-121 nm	AE-E	1977-1980	F10.7 (flux at 1.07×10^8nm)
		rockets	1976-1998	*Chromospheric:*
		SOHO	1995-	Lyman-α
		TIMED	2001-	He I 1083.0 nm EW
		ISS	2003-	Mg II c=w
		GOES	2004-	GOES 5 bands
				Ca II K images
3	FUV	AE-E (Ly-α only)	1977-1980	*Chromospheric:*
	122-200 nm	Nimbus 7	1978-1987	Lyman-α
		SME	1981-1989	He I 1083.0 nm EW
		rockets	1976-1998	Mg II c=w
		NOAA 9,11,14	1985-	Ca K 0.1 nm
		ATLAS 1,2	1992,1993	Ca II K images
		UARS	1991-	*Photospheric:*
		GOME	1995-	PSI
		TIMED	2001-	
		EOS	2002-	
		ISS	2003-	
4	UV	Nimbus 7	1978-1987	*Chromospheric:*
	200-400 nm	SME	1981-1989	Lyman-α
		rockets	1976-1998	He I 1083.0 nm EW
		NOAA 9,11,14	1985-	Mg II c=w
		ATLAS 1,2	1992,1993	Ca K 0.1 nm
		EURECA	1992-1993	Ca II K images
		UARS	1991-	*Photospheric:*
		GOME	1995-	PSI
		EOS/TSIM	2002-	
		ISS	2003-	
5	VIS-IR	ATLAS 1,2	1992,1993	*Photospheric:*
	400-1,000,000 nm	*EOS/TSIM*	2002-	PSI
		ISS	2003-	

Table A.1: SOLAR2000 modelling setup phases. Source: [31].

Proxy	Description	Application
$E_{10.7}$	Integrated XUV-EUV energy from 1 to 105 nm reported in 10.7 cm flux units	Empirical thermospheric density and ionospheric models that use 10.7 cm solar flux inputs .
S	Integrated solar spectrum	Research and operations that use a variable TSI.
Q_{EUV}	Total volume thermospheric heating rate	Aeronomy usage for comparison between UV airglow and solar irradiances.
R_{SN}	Derived sunspot number	HF propagation ray-trace algorithms that historically use the International Sunspot Number.
T_{∞}	Exospheric temperature	Comparison with Jacchia exospheric temperatures.
P_{EUV}	XUV-EUV hemispheric power	Aeronomy usage for comparison with Joule and particle precipitation heating.
E_{1_40}	Integrated XUV-EUV energy from 1 to 40 nm	Aeronomy usage for comparison between UV airglow and solar irradiances.
$XE_{10.7}$	Integrated XUV-EUV energy from 1 to 40 nm reported in 10.7 cm solar flux units	Aeronomy usage for comparison between UV airglow and solar irradiances.
X_{b10}	XUV 0.1-0.8 nm 24-h background irradiance index	Operational flare evolution prediction.
X_{hf}	XUV 0.1-0.8 nm hourly flare irradiance index	Operational flare evolution prediction.
$X_{10.7}$	$X_{b10} + X_{hf}$ reported in 10.7 cm solar flux units	Comparison of XUV 0.1-0.8 nm variability with 10.7 cm flux.
E_{SRC}	Integrated Schumann-Runge continuum energy from 125 to 174 nm	Lower thermosphere energy contribution to atmospheric density.
E_{SRB}	Integrated Schumann-Runge band energy from 175 to 204 nm	Mesosphere and stratosphere energy contribution to atmospheric density.

TABLE A.2: Summary of solar irradiance proxies provided by SOLAR2000. Source: [30].

Appendix B

Application of EMD to Wavelet Analysis

One of the weak points of wavelet analysis is that the presence of a trend, or of a component at very low frequency, leads to the misunderstanding of the wavelet power spectra: in such cases a component at high frequency could be not revealed. In order to clarify this concept let us define a trial signal $x(t)$ composed by a pure sine of small amplitude, and a significant graphed linear trend. Such a signal is proposed in figure B.1.a. The wavelet analysis does not emphasize a significant sinusoidal component, and this problem becomes more important when noise is present as well due to secondary processes. A possible wavelet spectrum - adopting a Morlet mother wavelet of order 6 - is presented in Figure B.1.b.

A possible work-around consists in applying Empirical Mode Decomposition to the data series to be analyzed in order to extract first the linear trend (=residue), and secondly, to perform wavelet analysis on the detrended signal. After doing this, the lower frequencies are effectively revealed by the wavelet analysis and significant periodicities. In fact, in the cases under consideration, the 11- and 22-years cycles, are pointed out more easily.

For instance, let us compare the wavelet spectra of temperature at 850-300 hPa obtained with a Morlet mother wavelet of order 6 with and without detrending by EMD method. The two spectra are reported in Figures B.2 and B.3 respectively.

A critical point of the application of the EMD to wavelet analysis is that a nonlinear trend could be considered as a signal at very low frequency. From this point of view, the subtraction of the trend from the whole signal could be considered as a loss of information. Actually, the residue of the EMD is a function, on which the sifting process can not be performed. This means, that there are no oscillations to be singled out, and so no frequencies to be

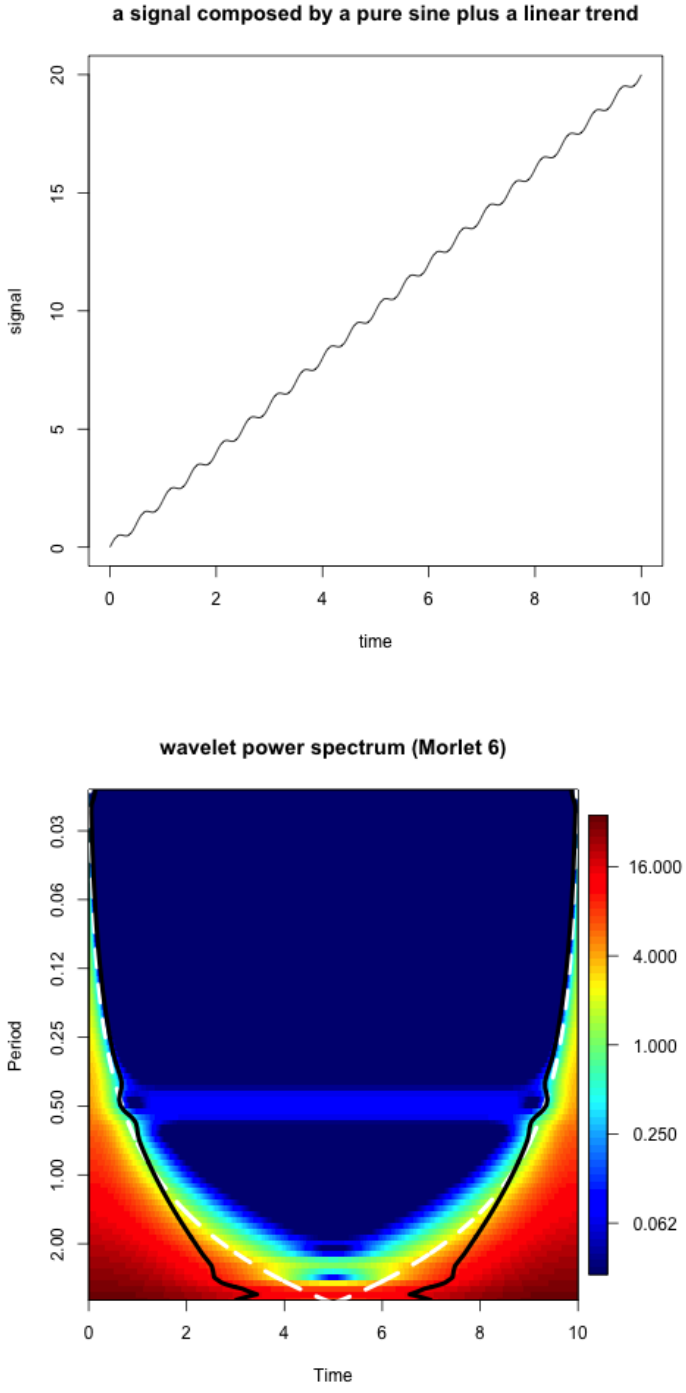


FIGURE B.1: (a) A trial signal composed by a pure sine of small amplitude and a linear trend. (b) Wavelet spectrum computed with a Morlet wavelet of order 6 of a trial signal composed by a pure sine of small amplitude and a linear trend.

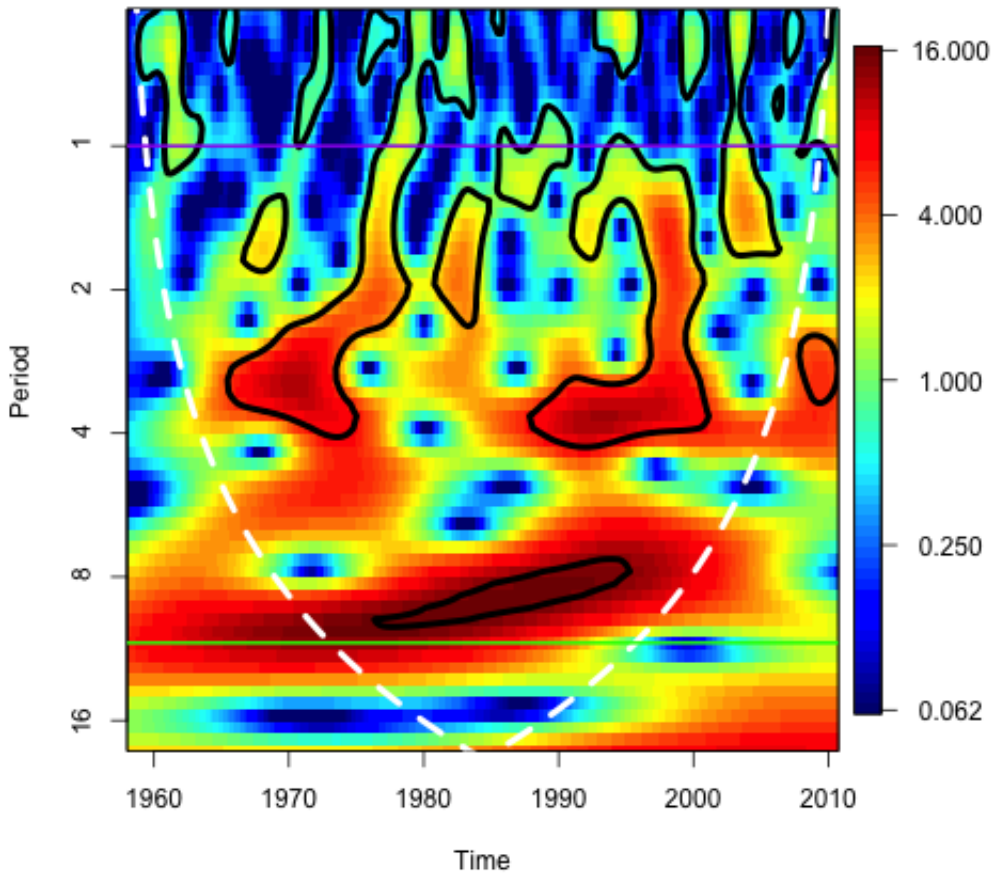


FIGURE B.2: Wavelet power spectrum of temperature at 850-300 hPa without detrending by EMD.

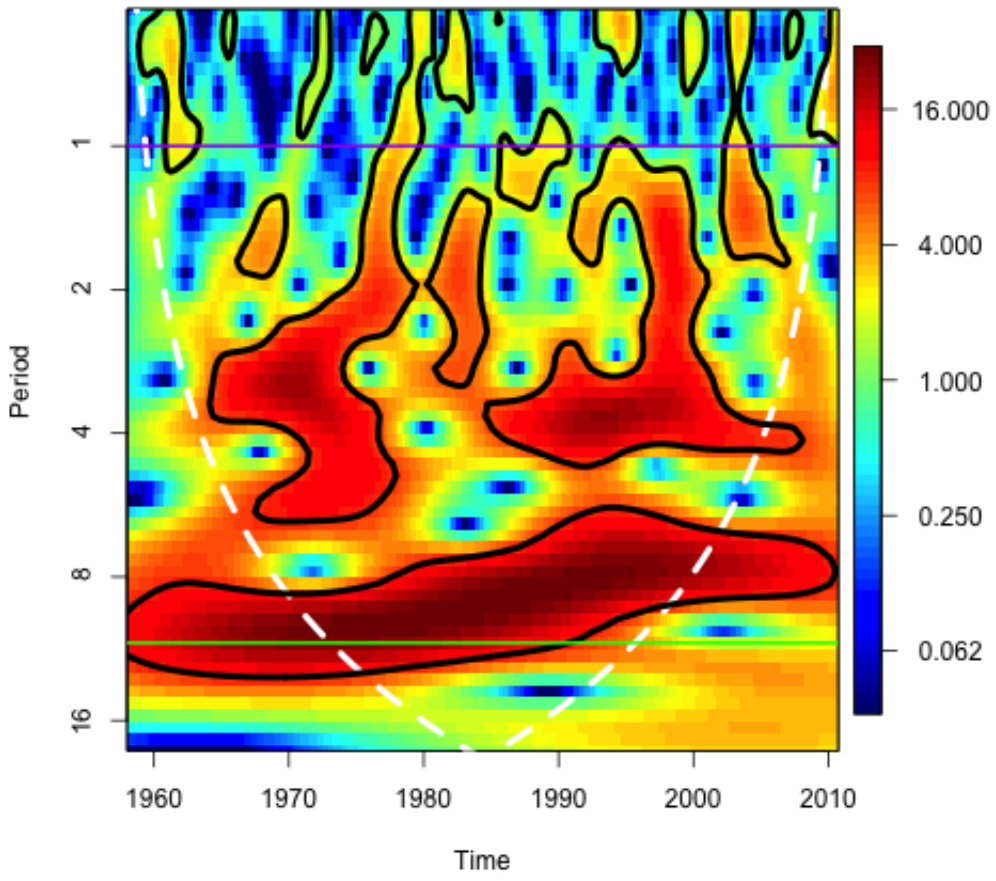


FIGURE B.3: Wavelet power spectrum of temperature at 850-300 hPa with detrending by EMD. The periodicity around 11 year is more evident.

revealed.

The subtraction of the residue of EMD is not useful in the Fourier analysis, since the identification of the frequency components is effective even in presence of a trend.

Glossary

ACRIM: Active Cavity Radiometer Irradiance Monitor Satellite.
CFC: Chlorofluorocarbons.
CCN: Cloud Condensation Nuclei
CWT: Continuous Wavelet Transform.
DFT: Discrete Fourier Transfom.
DOG: Derivative of Gaussian.
DWT: Discrete Wavelet Transform.
EMD: Empirical Mode Decomposition.
EM-Spectrum: Electromagnetic Spectrum.
ENSO: El Niño/La Niña - Southern Oscillation.
EOS/TSIM: Earth Orbit Station.
ESA: European Space Agency.
ERB-Experiment: Earth Radiation Budget Experiment.
ERBS: Earth Radiation Budget Satellite.
EURECA: The European Retrievable Carrier.
EUV: Extreme Ultraviolet.
FUV: Far Ultraviolet.
GCR: Galactic Cosmic Rays.
GOES: Geostationary Operational Environmental Satellite.
GOME: Global Ozone Monitoring Experiment, an instrument on board the ERS-2 satellite.
HS: Hilbert Spectrum.
IMF: Intrinsic Mode Function.
IR: Infrared.
ISO: International Standard Organisation.
ISS: International Space Station.
Ly-α: Lyman α.
MUV: Middle Ultraviolet.
NAO: North-Atlantic Oscillation.
NASA: National Aeronautics and Space Administration.
NOAA: National Oceanic and Atmospheric Administration.

NUV: Near Ultraviolet.
OEMD: Orthogonal Empirical Mode Decomposition.
QBO: Quasi Biennial Oscillation.
SAM:Southern Annular Mode.
SIP: Solar Irradiance Platform.
SME: Solar Mesosphere Explorer.
SOHO: Solar and Heliospheric Observatory.
SOLRAD: experiment carried by the project Galactic Radiation and Background (GRAB).
TIMED: Thermosphere Ionosphere Mesosphere Energetics and Dynamics.
TOMS: Total Ozone Mapping Spectrometer
TSI: Total Solar Irradiance.
UARS: Upper Atmosphere Research Satellite.
UV: Ultraviolet.
VIS: Visible radiation.
VUV: Vacuum Ultraviolet.
WFT: Windowed Fourier Transform.
WMO: World Meteorological Organization.
XUV: soft X-rays.
YOHKOH: solar observatory spacecraft of the Institute of Space and Astronautical Science (Japan).