Università degli Studi di Torino
Scuola di Dottorato

UNIVERSITÀ
DI TORINO

Interaction of cosmic rays with asteroids and terrestrial atmosphere and application of predictive models for the study of solar activity

Ilaria Bizzarri

Università degli Studi di Torino
Scuola di Dottorato

Dottorato in Fisica

Interaction of cosmic rays with asteroids and terrestrial atmosphere and application of predictive models for the study of solar activity

Ilaria Bizzarri

Tutor: Prof.ssa Carla Taricco

Abstract

Knowing solar activity and its variations over time is fundamental for understanding the evolution of the star itself, as well as the consequences that these variations can have on the Earth's climate. Given the scientific interest in these topics, several proxies, mostly linked to cosmic rays, have been identified and analyzed in this thesis. Some of them have been measured at the Monte dei Cappuccini underground Laboratory by the Cosmogeophysics group of the University of Torino, of which I was a member during my Ph.D., while others were studied thanks to collaborations with other research groups.

The collaboration with the Large Volume Detector (LVD) experiment in the Gran Sasso Laboratories (Italy) allowed me to study the longest underground muon flux ever recorded. Atmospheric muons originate from the decay of short-lived mesons, secondary particles produced by the interaction between high-energy cosmic rays and the atmosphere. Their flux carries information about the stratosphere, as the temperature of the latter modulates its intensity. A comprehensive investigation of this 24-year-long series allowed envisaging it as a high-time-resolution integrated proxy of lower-stratospheric temperatures. By using advanced spectral analysis methods, in addition to the well-known annual cycle, it was possible to detect two multiannual variations in the muon flux, including a decadal modulation linked to solar activity. The amplitudes of the multiannual variations are significantly larger than those expected based on the temperature modulations. Such differences may be due to the known difficulties of the adopted temperature dataset (ERA-Interim) in representing long-term variability scales, leading to an artificial attenuation of such long-term modulations.

The collaboration with the Pierre Auger Observatory in Malargüe (Argentina) gave me the opportunity to analyze the so-called scaler data, which represents a low threshold rate of cosmic rays measured on the ground, known to be affected by solar modulation. Several long-term components have been detected in the record, the most powerful of which have periods ranging from a decade down to a few days. In-depth investigations have been carried out to understand the phenomena at the origin of these variabilities. Furthermore, to study the daily variation of the Auger scalers, the short-term components have also been analysed, detecting the presence of periodic variability with periods down to a few hours.

The modulation of cosmic rays due to solar activity directly affects the production of cosmogenic radionuclides in extraterrestrial materials, such as meteorites. Therefore, the measurement of the activity of such radioisotopes represents a valuable proxy for solar activity. With a large-volume HPGe-NaI(Tl) spectrometer located at the Monte dei Cappuccini Underground Laboratory in Torino (Italy), the gamma activity of the Cavezzo meteorite (L5 anomalous chondrite) felt on January 1^{st}, 2020, was measured. Thanks to the high efficiency, selectivity, and low background of the spectrometer, fifteen cosmogenic radioisotopes were detected, with half-lives down to a few days. The HPGe dead layer thickness and detection efficiency were estimated using Geant4-based simulations, allowing to define the activity of each of the revealed radionuclides. The use of a particular

coincidence between the HPGe and NaI(Tl) detectors allowed the detection of the low activity of ^{44}Ti and ^{60}Co. The high ^{22}Na/^{26}Al ratio measured in the Cavezzo meteorite reflects the low activity of the Sun at the beginning of 2020 and is consistent with the expected value for L chondrites.

Thanks to the collaboration with the OSIRIS-Rex mission of NASA, the Monte dei Cappuccini Underground Laboratory will also carry out the activity measurement of the sample taken from the surface of the asteroid Bennu. This gamma-activity measurement will make it possible to detect the cosmogenic radionuclides present in the extraterrestrial material and to determine the cosmic-ray exposure age of the latter. Given the estimated composition and density of the sample, the best geometry of the container that will host the specimen during the measurement has been determined as a function of the expected sample mass. The optimal container shape is the one for which the gamma detection efficiency is maximum. The production rates of ^{26}Al and ^{22}Na have been estimated, and the counting error as a function of the measurement time was determined by Geant4-based simulations.

Finally, during my Ph.D. I developed a Bayesian Long Short-Term Memory model for predicting upcoming solar cycles using the most famous proxy for solar activity: the number of sunspots. The temporal evolution of sunspot activity for Solar Cycle 25 has been forecasted, together with a rigorous estimation of the uncertainty. Because of the underlying complexity of the dynamo mechanism in the solar interior at the origin of the solar cycle phenomenon, the predictions offered by state-of-the-art machine learning algorithms represent valuable tools for understanding the solar cycle progression. In particular, Bayesian deep learning is particularly compelling thanks to recent advances in the field providing improvements in the accuracy and uncertainty quantification compared to classical techniques. The uncertainty of the predicted cycle was obtained by including both the model uncertainty and the error propagation of the sunspot number time series. The forecasted peak amplitude is lower than that of the preceding cycle. Solar Cycle 25 will reach its maximum in the middle of the year 2024, while the next solar minimum is expected in 2030 and will be as deep as the previous one.

Contents

List of Abbreviations

ANN	Artificial Neural Network
AR	AutoRegressive
AU	Astronomic Units
BPTT	BackPropagation Through Time
BR	Branching Ratio
CCD	Charge Coupled Device
CERN	Conseil Européen pour la Recherche Nucléaire
CME	Coronal Mass Ejection
COI	Cone Of Influence
CR	Cosmic Ray
CWT	Continuous Wavelet Transform
DDEP	Decay Data Evaluation Project
DL	Dead Layer
ECMWF	European Center for Medium-range Weather Forecast
ECR	Extragalactic Cosmic Rays
EOF	Empirical Orthogonal Function
FADC	Flash Analog-to-Digital Converter
FD	Fluorescence Detector
FRIPON	Fireball Recovery and InterPlanetary Observation Network
FWHM	Full Width at Half Maximum
GCR	Galactic Cosmic Rays
GPS	Global Positioning System
GRB	Gamma Ray Burst
GWS	Global Wavelet Spectrum
HPGe	High Purity Germanium
IIR	Infinite Impulse Response
INFN	Istituto Nazionale di Fisica Nucleare
ISES	International Space Environmental Services
LINEAR	Lincoln Near-Earth Asteroid Research
LIS	Local Interstellar Spectrum
LNGS	Laboratori Nazionali del Gran Sasso
LSTM	Long Short-Term Memory
LVD	Large Volume Detector
MAE	Mean Absolute Error
MC	Monte Carlo
MC-SSA	Monte Carlo - Singular Spectrum Analysis
MHD	Magneto HydroDynamic
MSE	Mean Square Error
NASA	National Aeronautics and Space Administration
NEA	Near Earth Asteroid
NEO	Near Earth Object
NOAA	National Oceanic and Atmospheric Administration
OSIRIS-REx	Origins, Spectral Interpretation, Resource Identification, Security, Regolith Explorer

PC	Principal Component
PHO	Potentially Hazardous Object
PMT	Photomultiplier
PRISMA	Prima Rete Italiana per la Sorveglianza sistematica di Meteore e Atmosfera
QBO	Quasi-Biennial Oscillation
RC	Reconstructed Component
RNN	Recurrent Neural Network
SCR	Solar Cosmic Rays
SD	Surface Detector
SILSO	Sunspot Index and Long-term Solar Observations
SMM	Solar Maximum Mission
SN	Sunspot Number
SNR	Signal to Noise Ratio
SORCE	SOlar Radiation and Climate Experiment
SSA	Singular Spectrum Analysis
TAGSAM	Touch And Go Sample Acquisition Mechanism
TSI	Total Solar Irradiance
TIM	Total Irradiance Monitor
UT	Universal Time
UTC	Universal Time Coordinated
UTTR	Utah Test and Training Range
WDC	World Data Center
WTC	Wavelet Transform Coherence

Chapter 1

Solar evidence in the muon flux (LVD experiment)

Showers of secondary cosmic rays, composed of photons, electrons, positrons, hadrons, muons, and neutrinos, originate through successive interactions between high-energy cosmic rays and atmospheric atoms and molecules. The hadronic and electromagnetic components extinguish as the atmospheric depth increases, and only muons and neutrinos, the longest-living particles, can reach the ground.

The atmosphere is the vehicle of showers development, and its state rules both the intensity and propagation of the latter. Pressure and temperature are the atmospheric parameters that most influence shower production. Depending on the nature of the shower components, the way these parameters influence their rate and distribution changes, as well as the atmospheric layers that most influence their production. A change in pressure at the ground, corresponding to a change in the overburden matter, substantially modifies the electromagnetic component, minimally affecting the muonic one, which is more penetrating and which content is sensitive to temperature changes. Since muons of higher energy originate from the decay of unstable mesons (namely pions and kaons) in the early stages of the cascade, their production is influenced mainly by the state of the higher levels of the atmosphere (lower stratosphere). The intensity of the muon flux depends on the density of the portion of the atmosphere in which the mesons propagate: the higher the atmospheric temperature, the lower the air density, the lower the probability of mesons interacting before decaying, and the greater the likelihood of producing muons. Therefore, the muon flux can be considered a high-time-resolution integrated proxy of lower-stratospheric temperatures. Thanks to their small cross-section, low energy loss, and long lifetime, high energy (above $\sim$ 1 TeV) muons can penetrate the ground and reach great depths.

This chapter shows the investigation conducted on the 24-year-long series of the muon flux recorded underground at the Gran Sasso Laboratory (Italy) with the Large Volume Detector (LVD). Using advanced spectral methods, namely the singular spectrum analysis (SSA) and the wavelet transform coherence (WTC), the significant components of the LVD muon-flux series and the effective temperature, a parameter commonly used to describe the temperature-driven variations of the atmospheric muons, have been extracted and compared.

The results of such analyses have been published [see Ref. 1].

1.1 Cosmic rays

Cosmic rays (CRs) originate as primary cosmic rays produced in various astrophysical processes outside the Earth's atmosphere. More specifically, they originate from the Sun, from outside of the Solar System in our galaxy, and from distant galaxies, and they move through space at a nearly relativistic speed. About 99% of primary cosmic rays are nuclei of well-known atoms, about 1% are solitary electrons, while a small portion is composed of positrons and antiprotons. Regarding the first ones, about the 90% is represented by simple protons, the 9% are alpha particles (helium nuclei), while about the 1% are heavier elements. Their energy distribution peaks at 300 MeV and CRs with extreme energies are extremely rare.

Depending on their origin, they can be classified into three types:

- solar cosmic rays (SCR), which are high-energy particles (mostly protons), with energies below 10^6 keV, produced in the solar corona and ejected from the Sun by flares and coronal mass ejections (CMEs);
- galactic cosmic rays (GCR), i.e., high-energy particles originating outside our Solar System, with energies comprised between 10^6 and 10^{14} keV;
- extragalactic cosmic rays (ECR), with energies higher than 10^{14} keV, originating from beyond the Milky Way galaxy (usually included in the GCR).

The flux of incoming cosmic rays at the upper atmosphere depends on the Earth's magnetic field, the solar wind, and the particles' energies. Direct measurements of the cosmic ray flux have been possible since the launch of the first satellites in the late 1950s. They can be performed in space by particle detectors installed on satellites or at high altitudes by balloon-borne instruments. Currently, most operating experiments for high-energy CRs are ground-based, allowing for indirect measurement of the primary flux by looking at secondary cosmic rays, the latter produced by the interaction between primary cosmic rays and atmospheric molecules. Figure 1.1 shows a schematic representation of the components of an extensive air shower, composed of photons, electrons, positrons, hadrons, muons, and neutrinos.

Among them, muons make up more than half of the cosmic radiation at sea level. Thanks to their small cross-section, low energy loss, and long lifetime, muons with energy higher than ~ 1 TeV can penetrate deep underground. Underground muons are exploited for a variety of physics analyses, such as for the measurement of the flux, composition, and anisotropies of the GCR [see, e.g., Refs. 2–5], as well as for the validation and calibration of deep detectors. Furthermore, their measurement is of great importance since they are also one of the unavoidable backgrounds in underground laboratories for experiments searching for rare events.

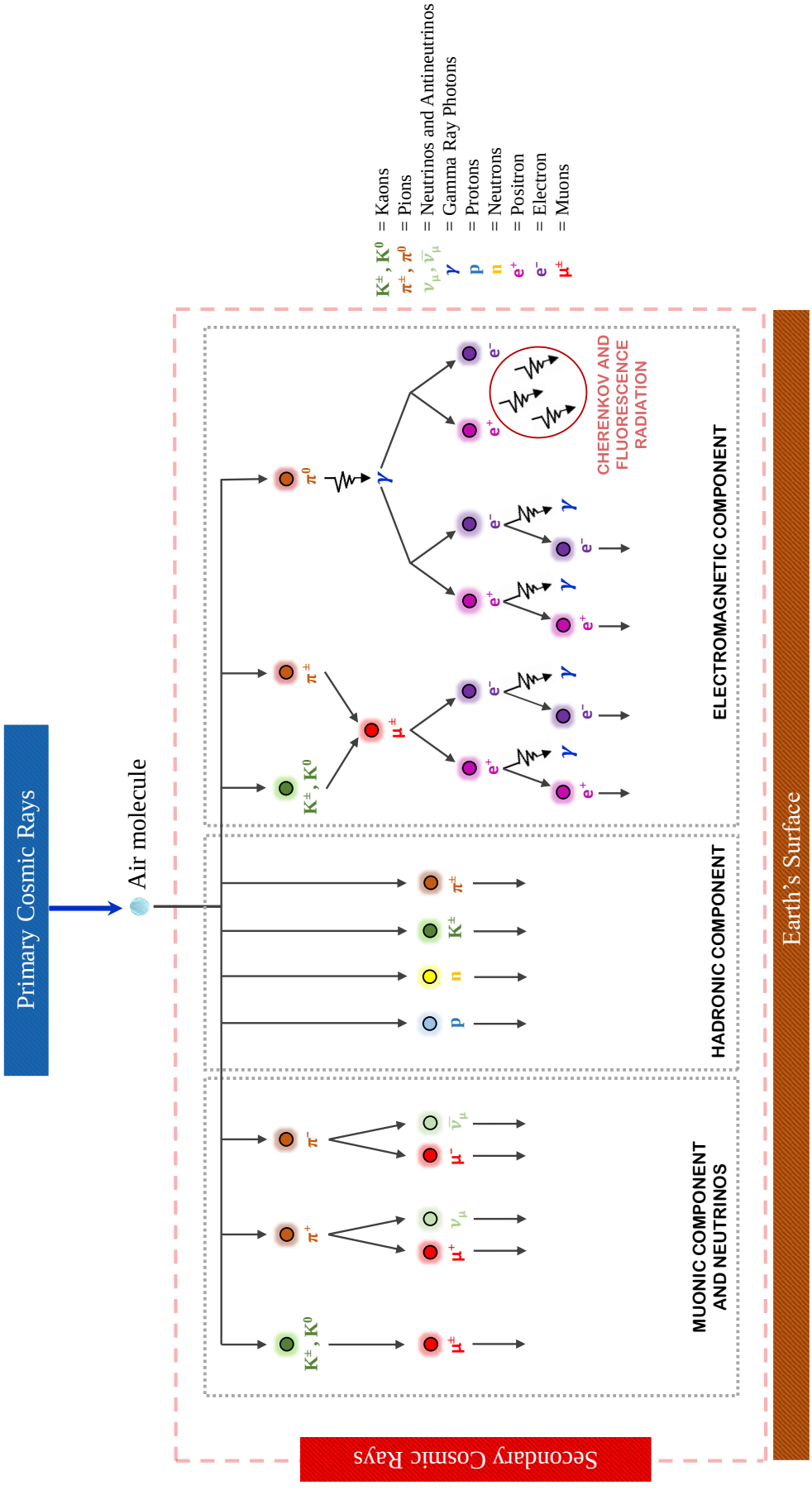

Figure 1.1: Schematic representation of the components of the shower originated through successive interactions between primary cosmic rays and air molecules.

1.2 The Large Volume Detector

The LVD is a neutrino telescope conceived to detect neutrinos from core-collapse supernova explosions in the galaxy [6]. Figure 1.2 shows a frontal view of the detector sited in the INFN Gran Sasso Laboratory (LNGS) (42.4275°N, 13.5333°E) under 1400 m of rock, at a minimal depth of 3100 m water equivalent. It consists of an array of 840 counters, 1.5 m^3 each, which hold, in total, 1000 tons of liquid scintillator. In each counter, the 1.2 tons of scintillator are viewed by three photomultipliers (PMTs) and the start of data collection is determined by the coincidence of the three signals in any counter. When this occurs, the amplitude and time of detection of the PMT signals are acquired, and to make sure to observe the neutrino burst associated with a gravitational collapse in the Milky Way, a dedicated machine-learning software has been developed for the data acquisition to maximize the detector observation time, allowing the number of active counters to vary over time. The energy resolution of the counters, at 10 MeV, is σ/E$\sim$ 20%, while the time of occurrence of each event is measured with a relative accuracy of 12.5 ns and an absolute one of 100 ns.

From the viewpoint of the PMT power supply, trigger, and data acquisition, the array is divided into three sectors called towers. Each sector operates independently from the others and includes 280 counters sharing the same low-voltage power supplies. Moreover, for each tower, counters are organized into 35 modules of eight sharing the same high-voltage divider, the same charge digitizer board, and the same position in the array. This modularity allows LVD to achieve a duty cycle close to 100%, which is essential in the search for the sporadic and unpredictable events of interest. A schematic representation of the LVD counters distribution is shown in Fig. 1.3, with the three towers highlighted in different colors.

The minimum energy, at the surface, for a muon to reach the detector is 1.3 TeV, and its passage through the LVD array can be easily recognised by analysing the time characteristics and amplitude of the induced signal. Only 1 per million cosmic ray muons at sea level have enough energy to penetrate down to the LNGS depth.

The LVD has been continuously taking data for a quarter century, since June 1992, allowing obtaining the longest muon flux time series ever recorded, which is described and shown in the following section.

1.3 Muon data and effective temperature

The LVD muon-flux time series includes 5.48×10^7 muons registered over 8402 days, from January 1994 to December 2017, corresponding to almost 96% of the total live time. The series comprises a few short gaps (of the order of days), as well as longer ones, which happened in 1996 (67 d) and in 2002 (24 d). To obtain a homogeneous series, required for the analysis methods used in the following (see Section 1.4), all the gaps have been filled with a gap-filling procedure based on autoregressive (AR) models, with order defined through the Akaike criterion [7].

Figure 1.4a shows the daily muon flux in terms of relative deviations from the average value, which is $I_\mu^0 = 3.35 \pm 0.0005^{stat} \pm 0.03^{sys} \times 10^{-4}$ m^{-2} s^{-1}. The shaded red bars represent the statistical uncertainties, while the filled data are highlighted in green.

[1] https://web.bo.infn.it/lvd/
[2] https://www.lngs.infn.it/it/lvd

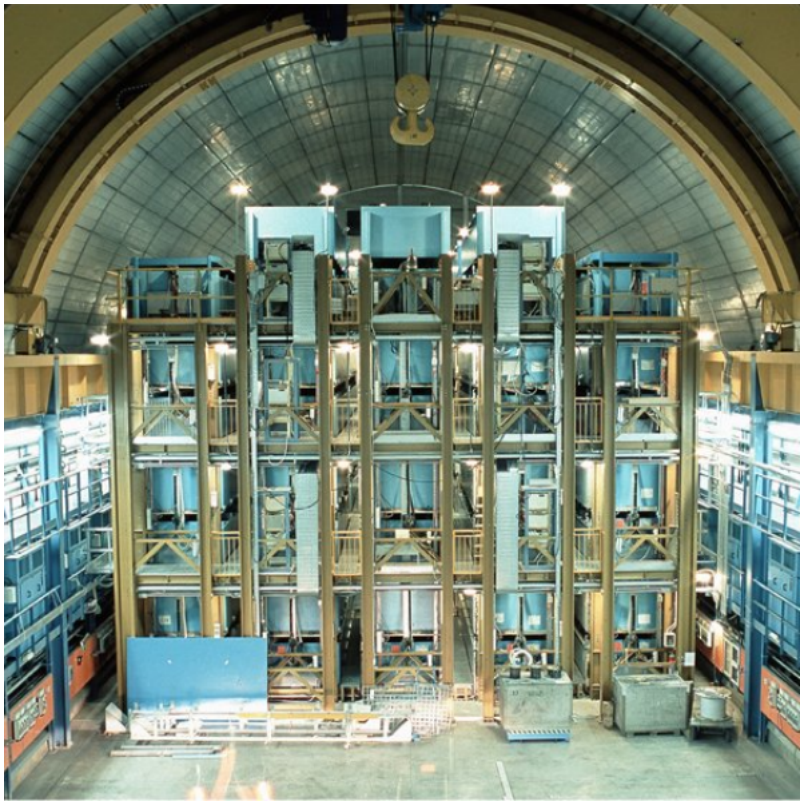

Figure 1.2: Frontal view of the LVD detector located at the LNGS. Picture taken from the LVD website[1].

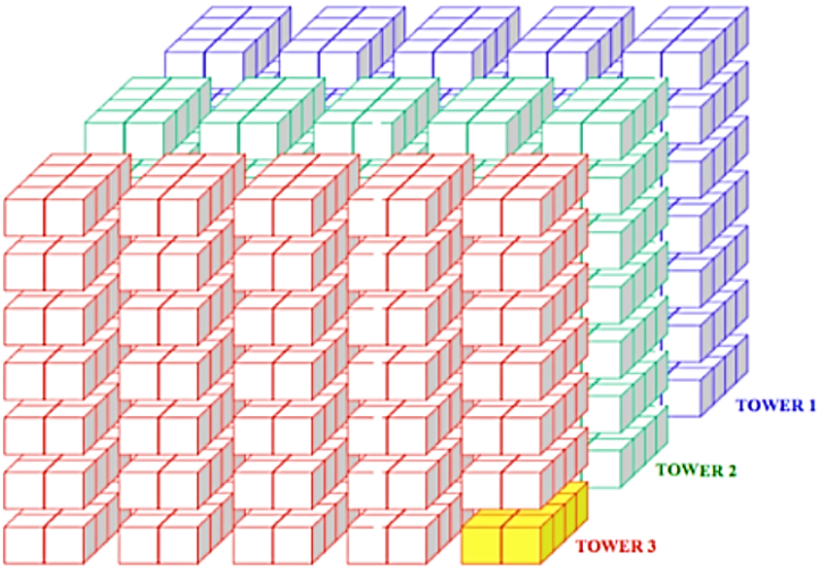

Figure 1.3: Schematic distribution of the LVD counters. Picture taken from LNGS web site[2].

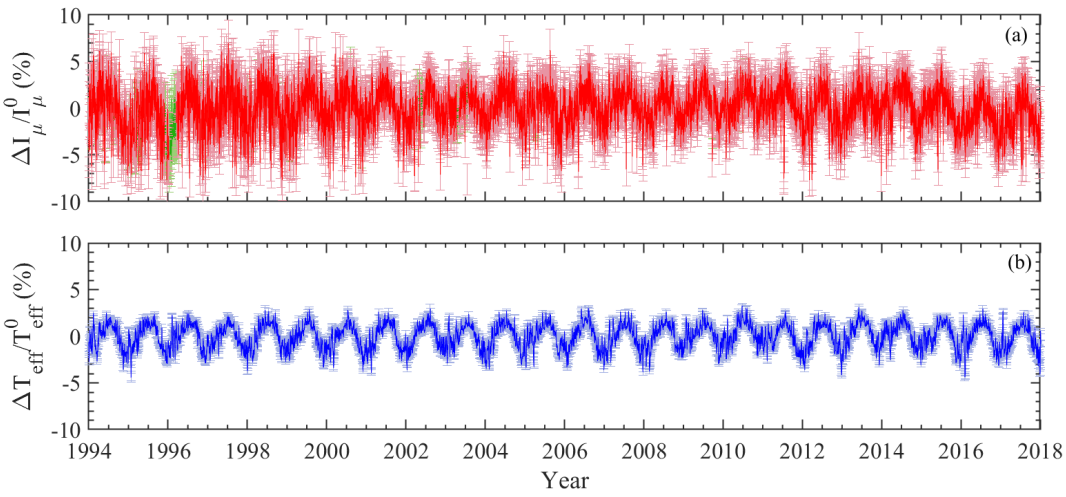

Figure 1.4: Percentage variation of the daily muon flux (panel a) and daily effective temperature (panel b). Shaded-colored bars represent the statistical uncertainties. The green intervals in the muon flux (panel a) represent gaps in the series which have been filled by using autoregressive models with order estimate through the Akaike criterion.

It has been known since the 1950s that the intensity of atmospheric muons is highly correlated with temperature changes in the upper stratosphere [8, 9]. If the temperature gets higher, the air density gets lower, and the probability of interaction between mesons and atmospheric particles is reduced, thus resulting in a higher muon rate.

To account for the fact that muons are not produced over a fixed level in the atmosphere, neither with fixed energy or direction, the so-called effective temperature T_{eff} has been introduced in those kinds of studies. The effective temperature, which can be interpreted as the temperature of an ideal isothermal atmosphere that should reproduce the observed muon intensity, is defined as the weighted average of the temperature over 37 discrete pressure levels in the 1-1000 hPa range. The weights account for the physics of production, propagation, and decay of mesons (pions and kaons), so that larger weights are expected for pressure levels at which the probability of mesons decaying in muons is high. Furthermore, the weights are site-dependent, i.e., they depend on the product between the energy required for the muon to reach the underground detector of interest and its direction of arrival. A detailed description of their estimation for the atmospheric muons detected by the LVD is shown in Ref.[10].

The effective temperature is calculated as:

$$T_{\text{eff}} = \frac{\sum_{n=1}^{N} \Delta X_n T(X_n) W(X_n)}{\sum_{n=1}^{N} \Delta X_n W(X_n)} \tag{1.1}$$

where the index n runs over N = 37 pressure levels. $T(X_n)$, $W(X_n)$, and ΔX_n are the temperature, weight, and thickness of the level X_n, respectively. The temperature data $T(X_n)$ are an interpolation of the ERA-Interim reanalysis dataset to the LVD location from the original $0.7° \times 0.7°$ provided by the European Center for Medium-range Weather Forecast (ECMWF) [11]. The provided dataset consists of one value each 6 h (4 values per day) over 37 discrete pressure levels in the 1-1000 hPa range. The daily series is obtained by averaging the four values per day, and their variance (typically 0.5 K) is used to estimate the uncertainty of the mean value.

Figure 1.4b shows the percentage variation of the daily effective temperature, where $T^0_{\rm eff} = 220.3 \pm 0.5$ K is the mean value.

An effective temperature coefficient α_T is defined as:

$$\frac{\Delta I_\mu}{I^0_\mu} = \int_0^\infty dX_\alpha(X) \frac{\Delta T(X)}{T^0_{\rm eff}} = \alpha_T \frac{\Delta T_{\rm eff}}{T^0_{\rm eff}} \tag{1.2}$$

where $\Delta I_\mu / I^0_\mu$ and $\Delta T_{\rm eff} / T^0_{\rm eff}$ represent the relative deviations from the mean daily muon intensity I^0_μ and mean daily temperature $T^0_{\rm eff}$, respectively. The value of α_T is 0.94 $\pm$ 0.02, obtained with a linear regression.

To better focus on the long-term variability and to avoid redundancy due to high-frequency oscillations, both the muon and effective temperature daily series have been rebinned in 5 days. The coefficient of correlation between daily muon flux and effective temperature is r = 0.56 ($p_{value}<10^{-5}$), while the one between the 5-d rebinned series is, as expected, higher and equal to r = 0.72 ($p_{value}<10^{-5}$).

The 5-d rebinned muon flux and effective temperature time series will be introduced in the next section, together with a detailed description of the methods used to define their spectral content.

1.4 Spectral Analysis

The harmonic analysis of the LVD muon series, as well as that of the effective temperature, has been previously performed [10] by using a classical Fourier-like approach, namely, the Lomb-Scargle method. However, this method is no longer the most appropriate tool in the presence of possible non-sinusoidal variations and for series with highly dynamic power ranges (i.e., including strong and weak spectral components). Therefore, the following advanced spectral methods have been used, which are suitable to efficiently extract and compare the deterministic components of the series under consideration.

1.4.1 Methods

The Wavelet Transform Coherence (WTC) method [12, 13] allows for multivariate analysis of time series. The main advantage of this approach is related to the evolutionary characteristic of the wavelet method, allowing for efficient detection of the amplitude modulation of the different variability components, as well as short-term events. An accurate time resolution at high frequencies and an accurate period resolution at low frequencies characterize these methods (see Appendix A).

The WTC is used to give a general view of the spectral features in common between the muon and the effective temperature time series and to estimate their coherence over a broad range of periods, from days to decades.

To deepen the analysis, the Singular Spectrum Analysis (SSA) [14–16] is applied to both series separately. The SSA is an advanced spectral technique that allows signal extraction from noise. With the use of data-adaptive filters, it is designed to separate a time series into statistically independent components that can be classified as oscillatory patterns or noise. The identified periodic components can be modulated in amplitude and phase.

A recursive procedure based on a Monte Carlo test (MC-SSA) [17] is designed to identify the significant components present in the series. Namely, at a given confidence level (c.l.), a set of null hypotheses are iteratively tested, starting from the simplest one, representing pure noise, up to more complex ones, obtained by superimposing on the noise different combinations of spectral components. To represent the noise, an AR process of order 1 [AR(1)] has been assumed, since it describes well the bias towards low frequencies, present in many time series, and it represents a suitable compromise between the excessive simplicity of white noise [see, e.g., Ref. 15] and higher-order AR processes, which are themselves capable of oscillatory behavior. Considering the time resolution of the series rebinned 5 days, the choice of AR(1) noise is consistent with the findings of the careful analysis of the noise given in Ref.[10].

The rejection of a null hypothesis is achieved by comparing the SSA spectrum of the analysed series with the Monte Carlo band, obtained by applying the SSA to the ensemble of 10000 surrogate series generated from the considered model of the series (null hypothesis). Finally, the model which explains the series, and allows the identification of the significant components at a given confidence level, is the one that cannot be rejected. More details about this technique are given in Appendix B.

1.4.2 Results

The WTC method applied to the 5-d rebinned muon flux and effective temperature time series gives the spectrum illustrated in Fig. 1.5c. Areas with a strong (weak) correlation are represented by the color red (blue). The cone of influence (COI), highlighting regions influenced by edge effects, is represented by the semitransparent area. The black curves delimit the area in which the coherence is significant at 95% c.l., while the black arrows show the relative phase relation: rightward for in-phase, leftward for anti-phase.

The broad red/yellow band surrounded by the black curve, extending over the entire time range between periods of ∼160 to 640 d, indicates the highly significant correlation between the two series on the annual period, as expected. Rightward arrows indicate a full coherence between the two series, hence the superimposed black rightward arrows over the entire annual range indicate that the two series are also fully coherent, i.e., in-phase.

Two other high-correlated bands are present in correspondence with periods of ∼4 y (with larger coherence from about 2008) and ∼10 y. For the latter, the red/yellow band lies outside the COI, thus indicating that the WTC method is not suitable to drive conclusions on such a long-term component. This limitation does not affect the SSA method.

Furthermore, elongated power spots are evident at the 30-d scale along the whole period spanned by the series, corresponding to a coherent behavior of the two series. This feature appears to be localized in winter months, thus suggesting that sudden atmospheric phenomena characterizing the stratosphere [18] may be imprinted in the muon flux.

The SSA is applied separately to the muon flux and effective temperature 5-d rebinned series. The adopted window width is defined as $W = M \cdot \Delta t$, where M is the window length and Δt is the sampling interval. After having tested the robustness of the results for a wide range of M values, the M parameter has been finally set equal to 400. The iterative steps of the Monte Carlo test are shown in Table 1.1, where the null hypothesis and the significant components revealed at 99% c.l., together with their period and variance, are listed for both the muon flux and the effective temperature time series.

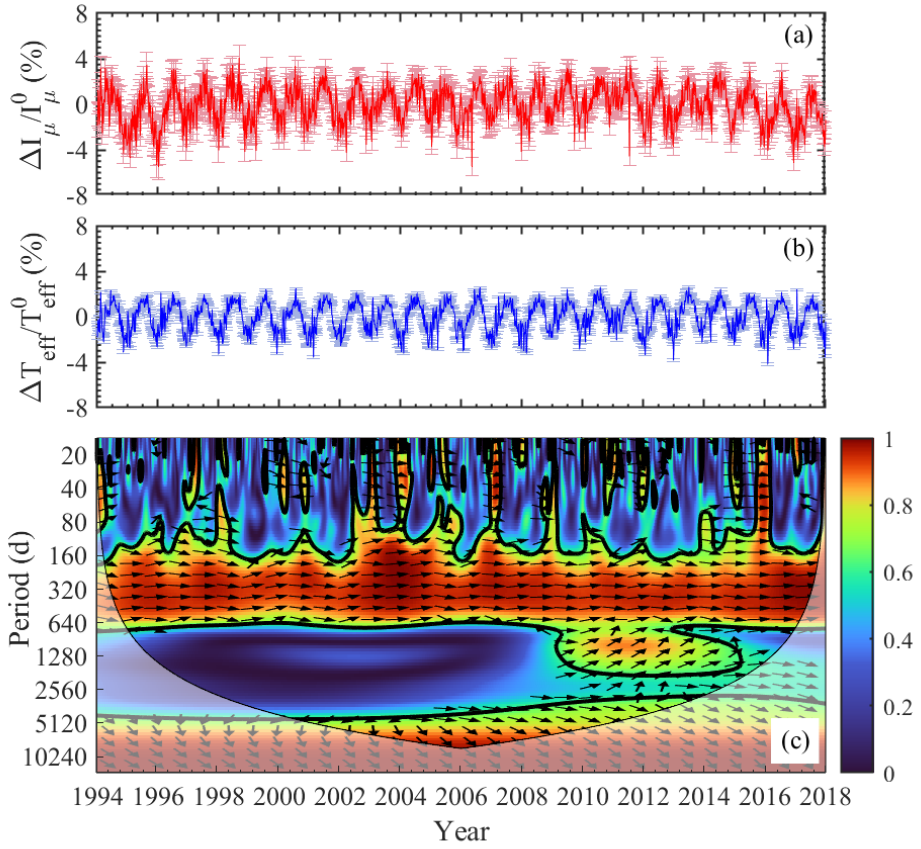

Figure 1.5: Percentage variation of the 5-d rebinned muon flux (panel a) and effective temperature (panel b). Shaded red and blue bars represent the statistical uncertainties of the 5-d rebinned muon and effective temperature time series, respectively. Panel c shows the wavelet transform coherence of the two series. Black curves delimit the area where the coherence is significant at 95% c.l.. Areas with a stronger (weaker) correlation are represented by the red (blue) color and the relative phase relation is shown by the black arrows: rightward for in-phase and leftward for anti-phase. The cone of influence is represented by the semitransparent area at the bottom.

MUON FLUX

Null hypothesis	Significant components	Period (y)	Variance (%)
AR(1)	1,2	1	41.3
AR(1) + RCs(1,2)	3	Trend (∼11 y)	∼3
AR(1) + RCs(1,2,3)	4,5,6	∼4	6.35
AR(1) + RCs(1,2,3,4,5,6)	9,10	∼1	1.4
AR(1) + RCs(1,2,3,4,5,6,9,10)	7,8	0.5	1.6
AR(1) + RCs(1,2,3,4,5,6,7,8,9,10)	11,12	∼1	1.3
AR(1) + RCs(1,2,3,4,5,6,7,8,9,10,11,12)	-	-	-

EFFECTIVE TEMPERATURE

Null hypothesis	Significant components	Period (y)	Variance (%)
AR(1)	1,2	1	67.8
AR(1) + RCs(1,2)	3,4	0.5	3.2
AR(1) + RCs(1,2,3,4)	-	-	-

Table 1.1: Iterative steps of the MC-SSA of the muon flux and effective temperature time series. The null hypothesis and the significant component revealed with a 99 % c.l. at each step are shown, together with the period and variance of the latter. The last null hypothesis for both series is the one that cannot be rejected, for which no more significant components are detected.

The spectra of the muon flux and effective temperature series, corresponding to the non-rejectable null hypothesis, are shown in terms of power vs period in Figs. 1.6a and 1.6b, respectively. The Monte Carlo bands are represented by the grey bars, while the red and blue dots represent the significant components revealed in the two series with a 99% c.l.. The empty dots represent the spectral components that can be parametrized as red noise.

Regarding the muon flux, the significant components revealed with the MC-SSA, which power exceeds the corresponding Monte Carlo bands (red dots in Fig. 1.6a), are a decadal trend and variability modes with 4-, 1-, and 0.5-y periods. The model that describes the muon flux series at 99% c.l. is given by the empirical orthogonal functions (EOFs) 1-12, describing ∼55% of the total signal variance (see Table 1.1), and the AR(1) noise.

In the case of the effective temperature, only two components are significant, with periods of 1 and 0.5 y (blue dots in Fig. 1.6b), which sum describes about 70% of the total signal variance (see Table 1.1). The model describing the effective temperature series at 99% c.l. is given by the EOFs 1-4, superposed to the AR(1) noise.

Figure 1.7 shows the reconstruction of all the significant components of the muon flux (red curves) and effective temperature (blue curve) series revealed by the MC-SSA. The annual and semiannual components are the dominant ones and they are the main cause of the high correlation (r = 0.72) between the two series. In fact, the correlation between the annual/semiannual detrended series is much smaller and equal to r = 0.42 (p_{value}<10^{-5}).

The semiannual component describes the asymmetry of the annual modulation due to the larger temperature variability during winter seasons, plausibly related to the occurrence of sudden warming events. The asymmetry is present for the whole duration of the series, as shown in Fig. 1.8. Panel a shows the muon flux rebinned series (shaded red curve) superimposed by the SSA reconstructed yearly oscillation (red curve) and the latter summed to the half-year component (black curve). Panel c shows the same comparison in a shorter time interval, to better appreciate the asymmetry caused by the semiannual periodicity. Similarly, panel b shows the effective temperature rebinned series (shaded

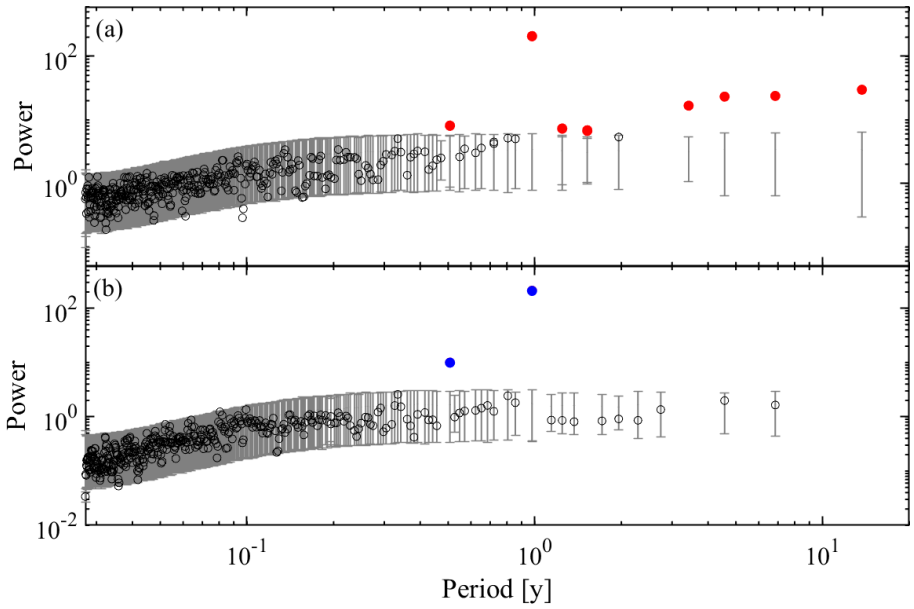

Figure 1.6: (a) MC-SSA spectra of the muon flux and (b) effective temperature series (99% c.l.). The Monte Carlo ensemble size is 10000. The significant spectral components are highlighted in red and blue for the muon and effective temperature, respectively. The empty dots indicate the spectral components that can be parametrized as red noise.

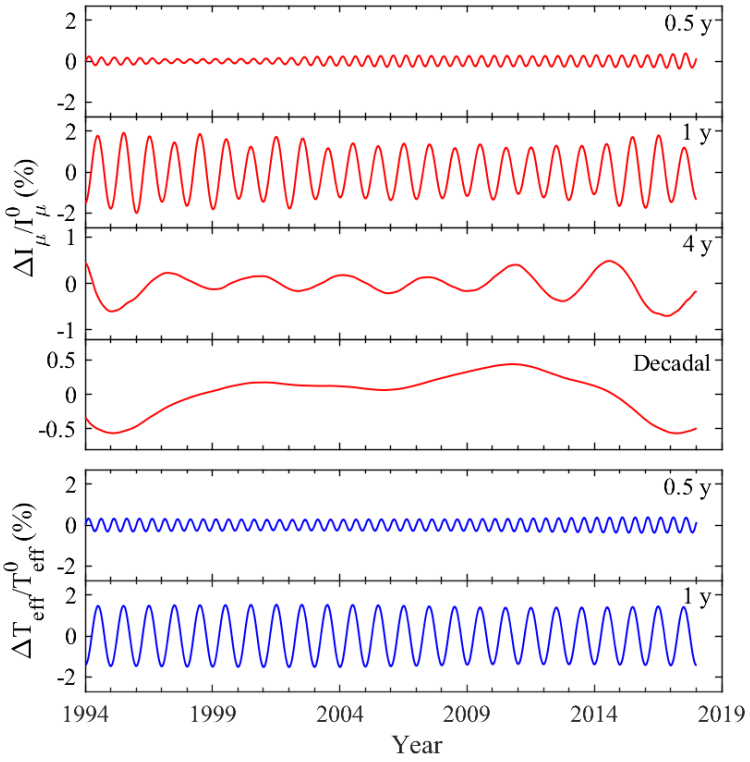

Figure 1.7: The significant MC-SSA components of the muon flux (red curves) and effective temperature (blue curves) series. The decadal component is obtained by using a larger window length (M = 800).

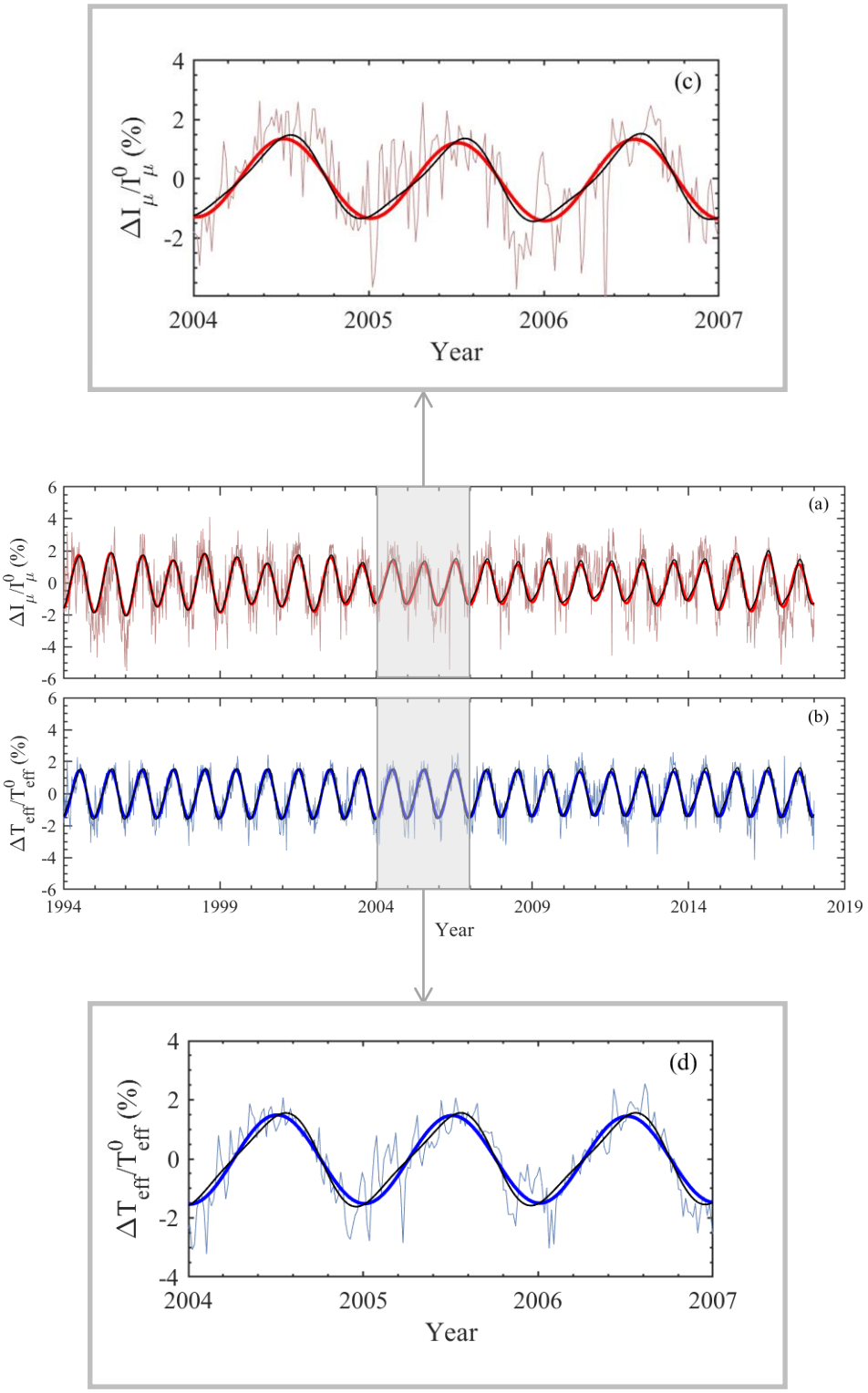

Figure 1.8: Panel a shows the muon flux rebinned series (shaded red curve) superimposed by the annual SSA reconstructed components (red curve) and the sum of the annual and semiannual components (black curve). Panel c shows a section of the series in panel a, where the asymmetry in the yearly oscillation due to the half-year component is clearly visible. Panel b shows the effective temperature rebinned series (shaded blue curve) superimposed by its annual SSA reconstructed components (blue curve) and the sum of the annual and semiannual components (black curve). Panel d shows a section of the series in panel b within the same time range as panel c. Also here, the asymmetry in the yearly oscillation due to the half-year component is clearly visible.

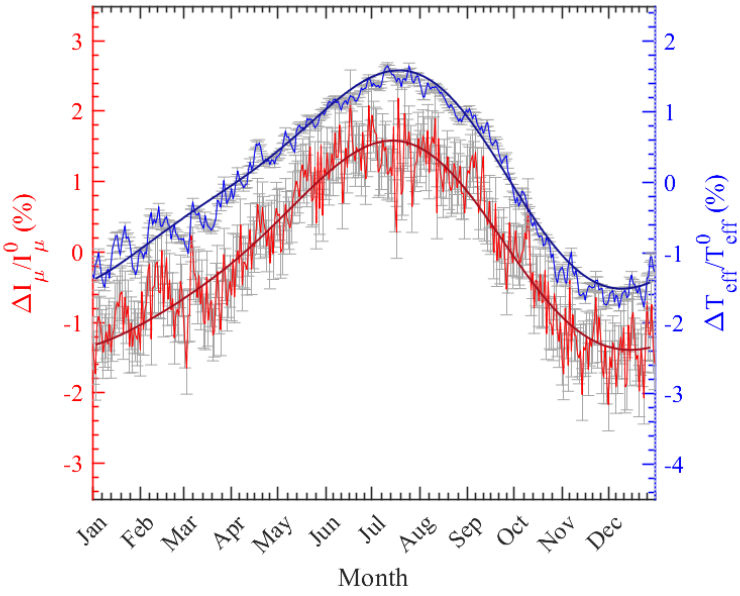

Figure 1.9: Muon flux (red curve) and temperature (blue curve) daily series with the 24-y datasets folded onto a 1-y period. The statistical uncertainties are represented by the light gray bars. The superposed heavy red and heavy blue curves correspond to the singular spectrum analysis (SSA) denoised signal, similarly folded onto a 1-y period, for the muon-flux and effective-temperature series, respectively.

blue curve) superimposed by its SSA reconstructed yearly oscillation (blue curve) and the sum of annual and semiannual components (black curve). Panel d shows a section of the series shown in panel b, within the same time interval as panel c. Also in this case the asymmetry caused by the semiannual variability is clearly visible.

Such asymmetry is clearly visible also in Fig. 1.9, where the muon flux and effective temperature daily series are shown after folding the 24-year-long dataset onto a 1-y period, with statistical uncertainties represented by the light gray bars. The superimposed thicker curves correspond to the yearly oscillations extracted with the SSA similarly folded.

The SSA-based description of the annual oscillation of the two series, more appropriate than the classical sinusoidal one, allowed an accurate determination of the amplitude and position of the maxima and minima on a year-by-year basis, capturing such asymmetry well.

The multiannual components revealed in the muon record do not have counterparts in the effective temperature series. This discrepancy was not expected, since the effective temperature is introduced to represent the muon flux variability. Only a faint indication of the presence of a 4-y component in the effective temperature series is found by applying the MC-SSA to the monthly series and decreasing the confidence level to 95%.

A deep investigation into this discrepancy has been made and will be discussed in the following sections.

1.5 Multiannual variability

As discussed in the previous section, the muon flux's spectral content differs from that of the effective temperature. Multiannual periodic components, in addition to the annual and semiannual dominant cycles, are present in the muon flux series but not in the effective

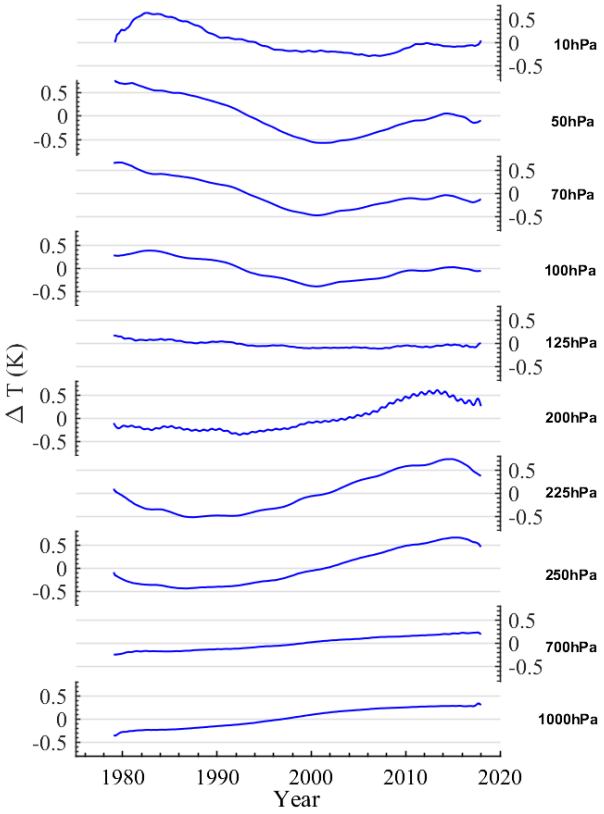

Figure 1.10: Trends revealed by MC-SSA of the monthly atmospheric temperature time series at different pressure levels.

temperature record. This discrepancy is unexpected since the muon flux should reflect temperature variations.

The origin of such variability modes has been deeply investigated by applying the MC-SSA to the monthly raw-temperature time series at several atmospheric levels, ranging from 1 to 1000 hPa, with data extending from 1979 to 2017. Raw temperature data have been taken from the ECMWF dataset. Table 1.2 shows the result of the MC-SSA performed on the raw-temperature time series over 21 of the 37 atmospheric levels. For each level, expressed in pressure (hPa), the percentage variance of the total signal described by each significant component detected using a window of width W=15 y (M=180; Δt=1 month) at 99% c.l. (unless otherwise specified) is shown. The dominant variability mode in each layer is the annual oscillation. A 4-months periodic component is detected mainly in the stratosphere, while a 6-months oscillation is present in all the layers. Several components with periods in the 2-5 years range are also detected. A decadal modulation is observed only between 175 hPa and 300 hPa.

The MC-SSA of the raw-temperature time series also reveals a long-term trend in all layers, which shows the well-known temperature increase in the troposphere and the temperature decrease in the stratosphere. Some of these trends are shown in Fig. 1.10, where the slope change around the tropopause is well visible. The amplitude of the reconstructed trends is in agreement with other findings [see, e.g., Ref. 19]. A decreasing trend is observed up to 125 hPa, no significant trends variations are revealed at 150 and 175 hPa, and an increasing trend is observed from 200 hPa down to the lowest layer (see also the last column of Table 1.2).

Level (hPa)	4 months	6 months	1 year	2-5 years	Decadal	Trend
1	0.96%	4.14%	58.79%	2.18% (∼2.3y, 98% cl)	-	7.3%
5	2%	4.52%	62.37%	-	-	1.43% (95% cl)
10	1.49%	1.69%	81.7%	0.36% (∼5y ,95% cl)	-	0.31% (95% cl)
30	0.56% (98% cl)	2.04%	79.83%	0.65% (∼2.3y, 98% cl) - 0.41% (∼3y, 95% cl) 0.32% (∼4.3y, 95% cl) - 0.18% (∼5.3y, 95% cl)	-	0.93% (98% cl)
50	1.04%	2.73%	68.27%	1.2% (∼2.3y, 95% cl) – 0.73% (∼4y, 90% cl)	-	2.58%
70	4.9%	4%	42.35%	2.48% (∼4y ,95% cl) - 2.12% (∼5y ,95% cl)	-	3.89%
100	5.58%	1.2% (95% cl)	44.76%	4.04% (∼4y, 98% cl)	-	2.57% (98% cl)
125	1.7%	1.32% (95% cl)	56.8%	1.56% (∼3.8y, 95% cl) - 1.18% (∼5.3y , 95% cl)	-	0.53% (95% cl)
150	1.01% (95% cl)	5.83%	45.31%	-	-	-
175	4.22%	11.95%	26.39%	1.98% (∼5y, 95% cl)	1.05% (90% cl)	-
200	5.28%	13.1%	34%	0.46% (∼2.3y, 95% cl) - 1.17% (∼3y , 95% cl) 1.67% (∼5.3y)	0.49% (90% cl)	1.54% (98% cl)
225	2.88%	8.8%	65%	0.68% (∼5y, 95% cl)	0.21% (95% cl)	2.18% (98% cl)
250	0.87%	4.55%	85.62%	0.07% (∼4y, 95% cl) - 0.23% (∼5y, 95% cl)	0.07% (95% cl)	1.31% (98% cl)
300	0.05% (95% cl)	1.78%	92.81%	0.14% (∼2.3y, 95% cl) - 0.12% (∼3y, 95% cl) 0.13% (∼4.3y, 95% cl) - 0.18% (∼5.3y, 95% cl)	0.07% (95% cl)	0.47%
400	-	0.98%	92.4%	0.13% (∼2.3y, 95% cl) - 0.18% (∼3y, 95% cl) 0.17% (∼5.3y, 95% cl)	-	0.3%
500	-	0.74%	92%	0.1% (∼2.3y, 95% cl) - 0.18% (∼3y, 95% cl) 0.1% (∼5.3y, 95% cl)	-	0.32%
600	-	0.67%	91%	0.17% (∼2.3y, 95% cl) - 0.19% (3y, 95% cl) 0.12% (∼5.3y, 95% cl)	-	0.19%
700	-	1%	89.8%	0.11% (∼2.3y, 95% cl) - 0.18% (∼3y, 95% cl) 0.16% (∼5y ,95% cl)	-	0.23%
800	-	1.1%	91.06%	0.17% (∼2.3y, 95% cl) - 0.14% (∼3y, 95% cl) 0.13% (∼4.3y, 95% cl)	-	0.49%
900	0.09% (95% cl)	0.74%	93.59%	0.15% (∼2y, 95% cl) - 0.11% (∼3y, 95% cl) 0.06% (∼4.3y, 95% cl) - 0.05% (∼5.3y, 95% cl)	-	0.33%
1000	0.09% (95% cl)	0.69%	94.34%	0.14% (∼2y, 95% cl) - 0.12% (∼3y, 95% cl) 0.08% (∼5.3y, 95% cl)	-	0.2%

Table 1.2: Variance of the significant (99% c.l., unless otherwise indicated) MC-SSA-components of the atmospheric temperature at different levels.

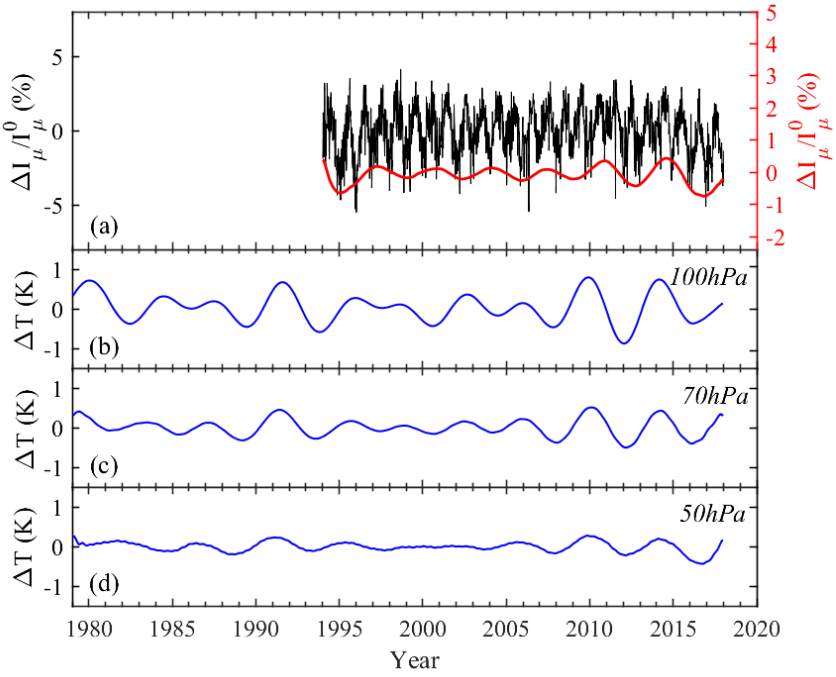

Figure 1.11: The 4-y component revealed by the MC-SSA in the muon flux (red curve, panel a) and in the raw-temperature series at 100 hPa (b), 70 hPa (c), and 50 hPa (d). Panel a also shows the 5-d rebinned muon series (black curve).

1.5.1 The 4-year component

The MC-SSA performed on the raw-temperature time series layer by layer allows to reveal the presence in the low stratosphere of a periodic component similar to the 4-y oscillation observed in the muon flux. At 100 hPa this component reaches its maximum amplitude, which decreases with altitude in the stratosphere and moving towards the troposphere. Figure 1.11 shows the reconstruction of such periodic components at 100 (panel b), 70 (panel c), and 50 hPa (panel d) together with the one found in the muon flux (red curve in panel a). This comparison shows a shape agreement, especially in the last portion of the records.

The same cannot be said about the amplitude. Considering equation 1.2, the $\Delta I_\mu/I^0_\mu \sim$ 1.2% observed between the last maximum and minimum of the 4-y muon flux oscillation should correspond to $\Delta T_{eff}/T^0_{eff}$ ~1.26%, that is to a $\Delta T_{eff} \sim$ 2.8 K. The 50-, 70-, and 100-hPa layers contribute to the effective temperature for 10.5%, 11.24%, and 9.87%, and show, in the same interval, temperature variations of ~ 0.63 K, ~ 0.83 K, and ~ 1.11 K, respectively. Therefore, they contribute to a total variation $\Delta\ T_{eff}^{50-70-100hPa}$ of:

$$\Delta T_{\rm eff}^{50-70-100hPa} \approx (0.63 \cdot 0.105) + (0.83 \cdot 0.1124) + (1.11 \cdot 0.0987) \simeq 0.27K \qquad (1.3)$$

describing only about 1/10 of the expected ΔT_{eff}.

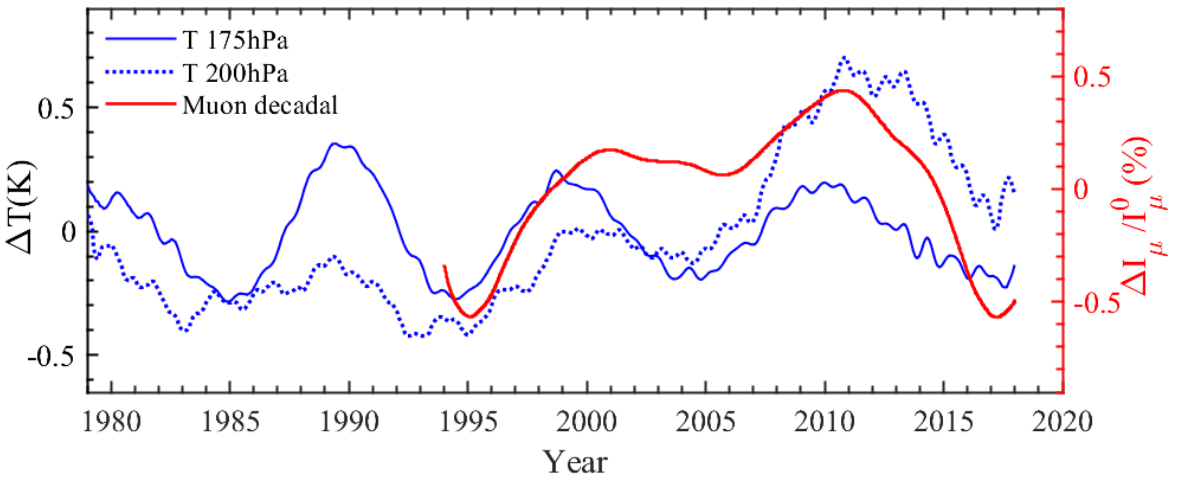

Figure 1.12: The decadal component revealed by MC-SSA in the 5-d muon flux series (red curve) and in the raw-temperature series at 175 (blue curve) and 200 hPa (blue dotted curve).

1.5.2 The decadal component

The raw-temperature time series show a decadal modulation which appears at ∼175 hPa and vanishes downward (see Table 1.2).

The decadal variability at 175 hPa is fully described by the decadal oscillation, while at other layers, such as 200 hPa, an additional trend is needed to describe such a variability mode. It is thus possible that such a trend is also present in the muon series, although it is not detectable due to the limited length of the series.

Moreover, the lack of a muon trend is plausible, since it results from the integration of production at different layers, with higher weights at the tropopause, where the slope change in the temperature trend is observed.

Figure 1.12 shows the comparison between the muon decadal component (red curve) and the temperature decadal variability at 175 hPa (blue curve). The blue dotted curve represents the long-term temperature component at 200 hPa, obtained as the sum of the decadal and trend ones, whose agreement with the decadal modulation of the muon series is improved compared to considering only the decadal variability.

A good shape agreement is visible among the reconstructed decadal oscillations but, similarly to the case of the 4-y component, a discrepancy regarding the amplitude is present. The $\Delta I_\mu/I^0_\mu$ ∼1% of the muon decadal modulation should correspond to a ΔT_{eff} ∼ 2.3 K. The 175- and 200-hPa layers contribute to the effective temperature for 5.1% and 4.2%, respectively. Therefore, the temperature variations of ∼ 0.5 K and ∼ 1.1 K observed in the decadal modulation at 175 and 200 hPa contribute to a total variation $\Delta\ T_{eff}^{175-200hPa}$ of:

$$\Delta T_{\rm eff}^{175-200hPa} \approx (0.5 \cdot 0.051) + (1.1 \cdot 0.042) \simeq 0.07K \tag{1.4}$$

which describes only 1/30 of the expected ΔT_{eff}, a value which is too low considering that at such levels the temperature decadal variation reaches its maximum amplitude.

This discrepancy (together with the one discussed for the 4-y component) may highlight known difficulties of the adopted reanalysis dataset in describing long-term variations

in stratospheric layers [11, 20]. A possible attenuation of long-term wave- and transport-driven modulations may result due to changes in the assimilated global observing system and the presence of time-varying biases in models and observations. A further contribution to the dampening of decadal-scale temperature variations may be due to the assumption of a constant value of solar irradiance in the underlying atmospheric model.

1.5.3 Multiannual components over the Gran Sasso region

The multiannual components discussed in previous sections refer to the Gran Sasso area. To put the findings in a larger-scale perspective, the geographical distribution of temperature over the Northern Hemisphere is considered by analysing temperature maps at different heights in the atmosphere. The pressure levels of interest have been chosen among three atmospheric regions: the stratosphere (50 and 70 hPa), the upper troposphere–lower stratosphere (100, 175, and 250 hPa), and the troposphere (500 hPa). To reveal the presence and the intensity of the 4-y and decadal oscillations at such levels, differences in temperature between maxima and minima for each periodicity have been calculated.

Figure 1.13 shows the geographical distribution of such differences over a large part of the Northern Hemisphere. The periods related to maxima and minima are based on the SSA oscillations shown in Fig. 1.11 and 1.12 for the 4-y and the decadal component, respectively. In the layers suggested by the SSA (100 hPa for the 4-y and 175–200 hPa for the decadal oscillation) local maxima are present, as indicated by the red features in Fig. 1.13, left panel. The 4-y component is also present at 70 and 50 hPa, as shown in Figs. 1.11c and 1.11d, though the amplitude revealed by the SSA in these layers is smaller, possibly due to the dominance of the annual cycle at these heights (see next section). The decadal cycle is only present at 175–200 hPa and rapidly disappears above and below this layer (Fig. 1.13, right panel).

This analysis confirms the results obtained by the SSA performed on the raw temperature time series at different atmospheric levels over the Gran Sasso region and also shows that both cycles are not due to local phenomena, with the 4-y variation extending to several parts of Western Europe and the decadal variation being present at large scales throughout the Northern Hemisphere (Europe, North America, Atlantic Ocean, and high latitudes).

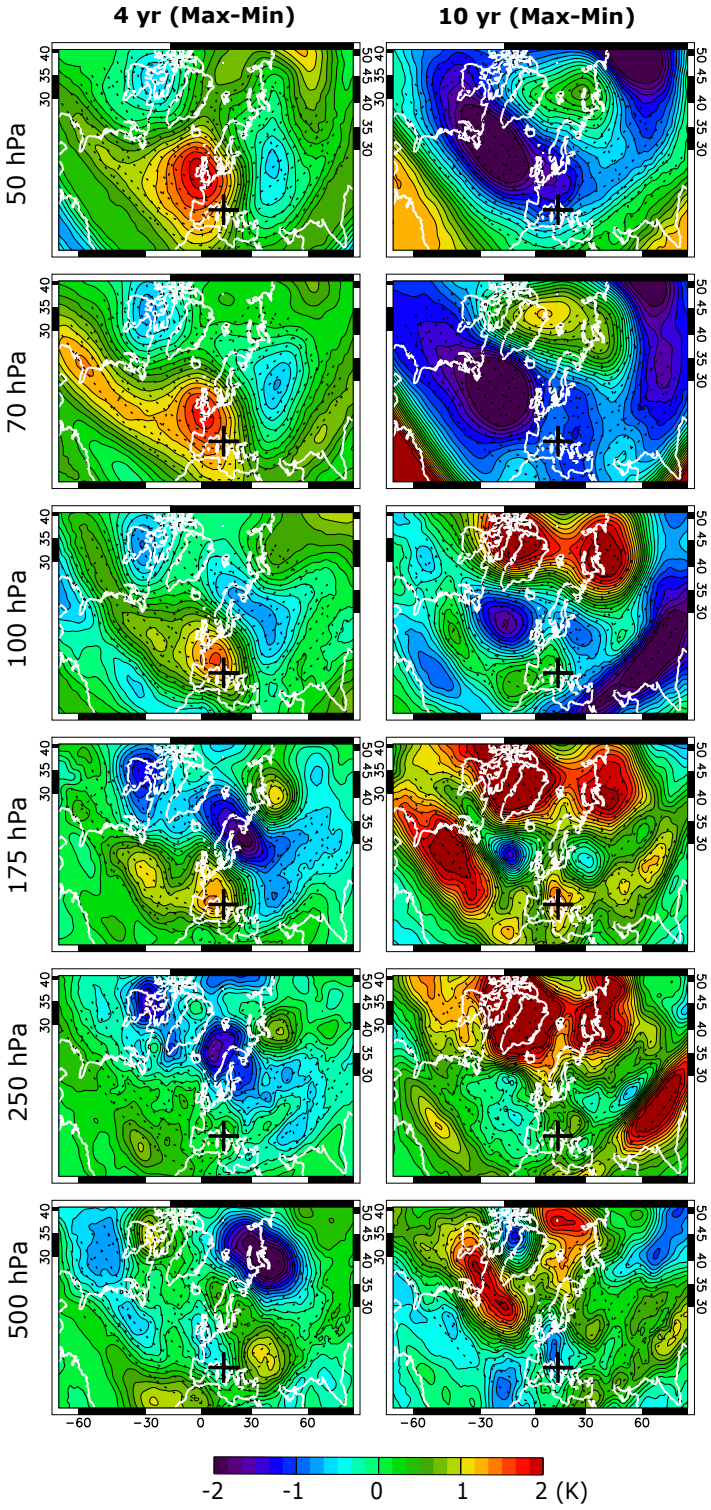

Figure 1.13: Geographical distribution of temperature differences between years of maximum and minimum signal of the 4-y (left) and decadal (right) periodicity. Data are reported for pressure levels ranging from 50 to 500 hPa and for a large fraction of the Northern Hemisphere. The position of the LNGS is indicated with a black cross. The Land's borders are delimited by white curves.

1.5.4 Multiannual components and annual cycle

As previously discussed, the multiannual components revealed in the muon flux are related to temperature variability between 100 and 200 hPa. To explain the reason behind this, it has been hypothesized that the emergence of such multiannual components may be related to a dampening of the dominant annual modulation at these layers.

Figure 1.14a shows the correlation coefficient between the daily series of the muon flux and the raw temperature in the 37 atmospheric levels. Two maxima, centered at $\sim$ 30 and 225 hPa (dashed blue lines), are present, together with a local minimum near the tropopause, between 100 and $\sim$200 hPa (range delimited by the red lines). A similar behaviour was found in Ref.[21].

The observed profile is most likely driven by the dominant annual cycle since its variance gets lower in correspondence with the local minimum of the correlation (see Table 1.2). That is a plausible conjecture, given the dominant role of transport processes within this layer, dampening the annual radiative modulation. To prove this hypothesis, the annual modulation of the temperature at different atmospheric layers, reconstructed with the MC-SSA, was examined. This is shown in Fig. 1.15, where it is clearly visible that the amplitude decreases by a factor $\sim$ 4 from 10 to 100-200 hPa, and then increases again moving towards lower levels.

The correlation between the annual components of the muon flux and the temperature as a function of the pressure level is shown in Fig. 1.14b (blue curve), together with the percentage of the total variance described by the annual component detected in each raw-temperature series (red curve). A correlation minimum is observed around the tropopause, thus demonstrating that the minimum observed in the correlation between the muon flux and the raw temperature series (Fig. 1.14a) is, in fact, related to the annual cycle. However, the high correlation at 175 and 200 hPa, reaching values comparable to the ones at 30 hPa, is in contrast with the behaviour observed in Fig. 1.14a. This can be explained by looking at the variance curve in Fig. 1.14b. Although the correlation at 175 and 200 hPa is good, the variance of the annual temperature component reaches its minimum value at such pressure levels. Moreover, the amplitude of the annual component assumes the minimum value in correspondence with these levels (see Fig. 1.15).

Finally, Fig. 1.14c shows the correlation between the temperature and muon flux time series after detrending them with the corresponding annual cycles reconstructed with the MC-SSA. The fact that, in this case, there is no minimum around the tropopause confirms that the minimum observed in Fig. 1.14a is due to the weakening of the annual temperature cycle.

It is possible to conclude that, thanks to the attenuation of the dominant annual cycle at these levels, the weaker 4- and 10-y components emerge in both the temperature and the muon series.

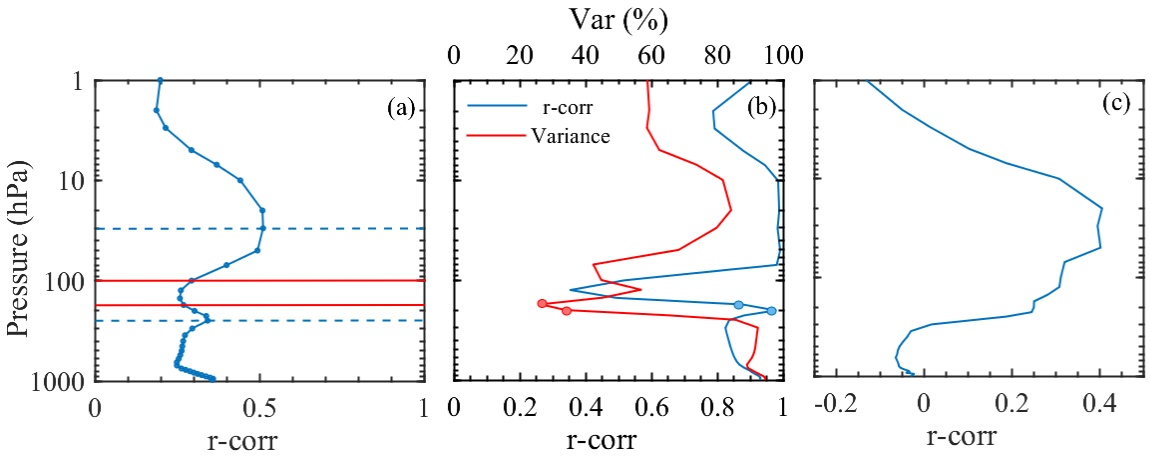

Figure 1.14: (a) Correlation coefficient between muon and temperature daily data from 1 to 1000 hPa. The red lines identify the 100 and 175 hPa levels, while the dashed blue lines mark the two correlation maxima. (b) Correlation coefficient between the annual components of the temperature and muon series (blue curve) and percentage of the total variance of the annual temperature component (dashed red curve) from 1 to 1000 hPa. (c) Correlation coefficient between temperature and muon series after detrending them with the corresponding annual cycle reconstructed with the MC-SSA.

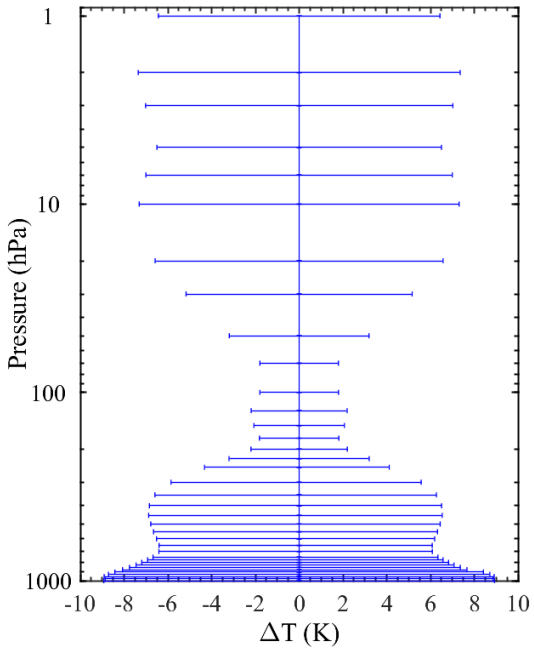

Figure 1.15: Amplitude of the annual cycle reconstructed by the MC-SSA for the raw-temperature time series from 1 to 1000 hPa.

1.5.5 Decadal muon component and solar activity

The decadal oscillation revealed in the muon flux series may suggest a relation with the decadal solar cycle (Schwabe cycle). Cosmic rays with energies below tens of gigaelectronvolt are known to be anticorrelated with the 11-y solar cycle, while the muons generated by them do not have enough energy to reach the LVD underground detector. The geomagnetic cutoff at the location of the latter prevents such low-energy cosmic rays from even reaching the atmosphere and even if low-energy atmospheric neutrinos interacting with the rock surrounding LVD can generate additional muons, their number would be so small to explain the observed modulation. Therefore, such modulation can be related to the decadal stratospheric temperature induced by the solar cycle.

Several complex mechanisms, both direct and indirect, can lead to an impact on the stratospheric temperature due to the Schwabe cycle. Because of their complexity, the decadal temperature modulation imprinted by the 11-y solar cycle can be found only at specific layers and regions in the atmosphere. Several studies have identified an 11-y signal in the Northern Arctic region temperatures at ∼10 hPa due to direct changes in the heating, extending to midlatitudes and then connecting to a region at 100 to 20 hPa at low and midlatitudes due to indirect changes in the global Brewer-Dobson atmospheric circulation [22, 23]. Because of these two regions of influence, the temperature in the Gran Sasso region is expected to show an imprint of the 11-y cycle in the low stratosphere. This is confirmed by the results obtained by applying the MC-SSA to the raw-temperature time series at different layers, which allowed detecting a decadal variation at ∼175-300 hPa (see Table 1.2), with the maximum amplitude reached at 175 hPa.

A recent analysis of 10-y of data from the Borexino experiment [24], also situated in the LNGS, revealed the presence of a decadal modulation of the muon flux. Although the precision in determining a decadal modulation within the 10-y covered by the Borexino data is limited (due to the presence of only one cycle), such a decadal trend is consistent with the one found in the LVD muon data.

In Fig. 1.16, a comparison between the LVD (panel a) and the Borexino (panel b) decadal muon trends is shown. Panel c shows the decadal variation in the raw-temperature series at 175 (blue curve) and 200 hPa (dotted blue curve), which shape is in agreement with the decadal muon modulation.

Both the temperature and muon flux decadal variabilities reproduce the 11-y solar cycle, described by the sunspot number (SN) time series shown in Fig. 1.16e, although a phase misalignment is visible between temperature and sunspots, with temperature maxima and minima anticipating those in the sunspot series.

A better agreement is found by looking at the 11-y solar modulation of geomagnetic activity recorded by the Ap index (see Fig. 1.16d), whose maxima are observed in the descending phases of the 11-y sunspot cycles. Geomagnetic activity can influence stratospheric temperature by regulating the impact of energetic particle precipitation on atmospheric chemistry. Consequent temperature variations can depend directly or indirectly on the considered region [25, 26] and, as a consequence, the indirect correlation between the Ap index and our measurements is not unexpected. Moreover, the quasi-biennial oscillation was found to amplify the effect of the 11-y cycle and geomagnetic activity on the stratosphere, modifying the original solar signal imprinted in the latter.

Those terrestrial effects arise the complexity in the interpretation of the origin of the decadal muon cycle, which will need further and deeper analyses to be understood.

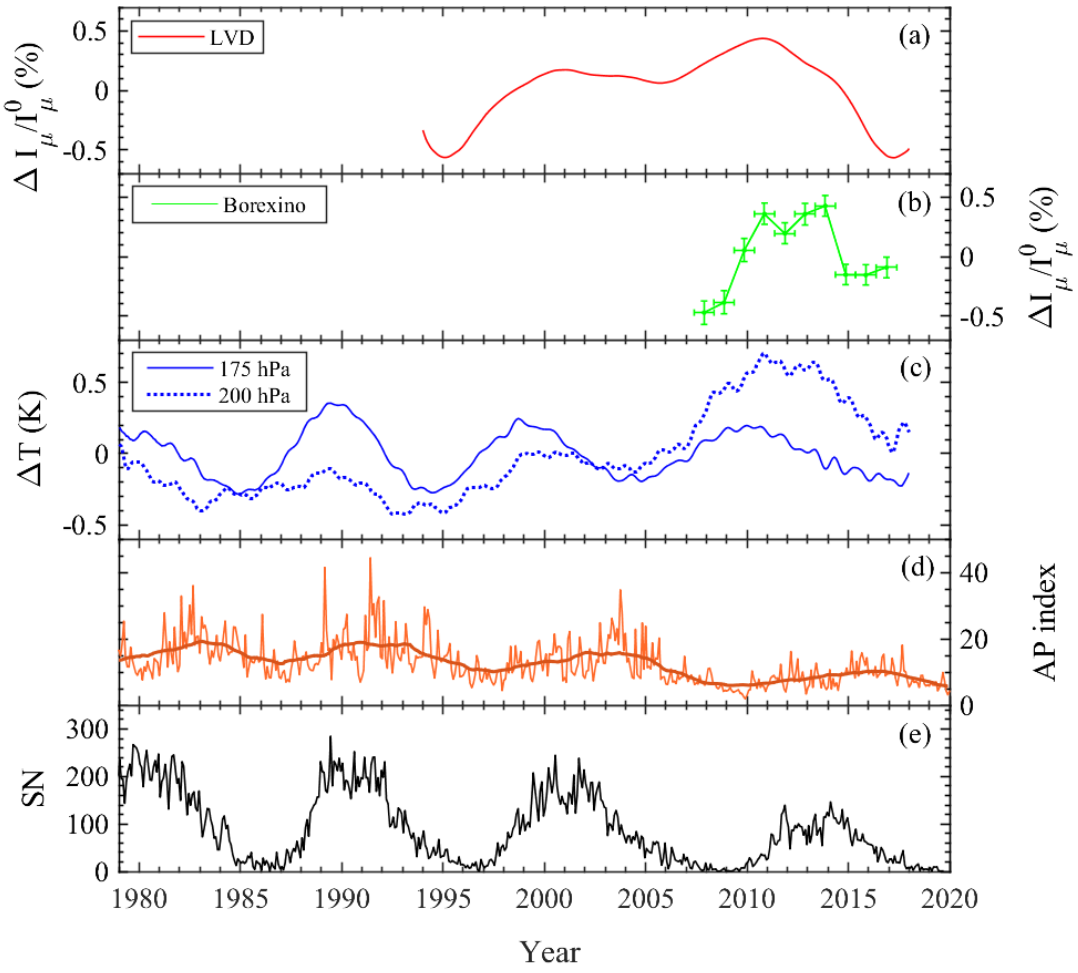

Figure 1.16: (a) MC-SSA decadal muon flux component from the Large Volume Detector (LVD) data, (b) Borexino muon-flux variation data, (c) the MC-SSA decadal components detected in the raw-temperature series at 175 and 200 hPa (decadal plus trend components, dotted curve), (d) monthly mean Ap index (with superimposed running mean), and (e) total sunspot number.

1.6 Conclusions

The spectral analysis of the 24-y muon flux time series observed underground by the LVD in the Gran Sasso Laboratory, performed with the MC-SSA and WTC methods, revealed the presence, in addition to the well-known annual periodicity, of two multiannual components, with periods of $\sim$ 4 and 10 y. Neither of the latter has been detected in the effective temperature time series.

Applying the MC-SSA on the raw-temperature time series from the ERA-Interim reanalysed dataset, a 4-y oscillation was detected near the tropopause, at $\sim$100 hPa, while a decadal modulation was observed at 175-200 hPa. Such weaker oscillations emerge thanks to the dampening of the dominant annual cycle at these levels, which are affected by high-frequency variability related to transport and wave processes.

Such long-scale modulations exhibit a much larger amplitude than would be expected based on temperature variations. These long-term temperature variations make such a small contribution to the effective temperature, due to their small amplitude, that they are not significant in the spectrum of the latter. All these results may represent difficulties of the adopted reanalysis dataset to correctly represent stratospheric temperature variability, due to a possible attenuation of long-term modulations.

In turn, the intensity of muons measured underground is very sensitive to temperature variabilities in the lower stratosphere, thus suggesting that variations in the muon flux may represent an additional unbiased integrated proxy of local stratospheric temperature with a high time resolution.

Chapter 2

Solar evidence in the scaler data (Auger experiment)

Changes in the interplanetary medium related to solar activity variations and transient events modify the magnetic deflections of high-energy particles, thus changing the CR flux in the Earth's atmosphere. The standard detectors for measuring these effects are neutron monitors, but also different detection systems, such as muon telescopes and particle counters from high-energy experiments. Among the latter, there is the Pierre Auger Observatory, located in Malargüe, Argentina. It is composed of an array of 1660 water-Cherenkov surface detector stations, covering 3000 km^2, constantly measuring the rate of particles, as well as of a Fluorescence Detector, composed of 24 telescopes grouped in units of six at four different locations, designed to measure the ultraviolet light emitted by excited nitrogen when an extensive air shower passes through the atmosphere.

The rate of particles in all the surface detectors of the array can be measured using a counting trigger, the so-called scaler mode. The Auger scaler rate corresponds to the number of particles depositing an amount of energy within the range [15-100] MeV and can be used to study solar effects on the cosmic-ray rate on scales from minutes to years. These energy depositions arise from showers of primary cosmic rays with energies from 10 GeV to a few TeV.

The spectral analysis conducted over the scaler data sampled every 48 hours is here presented. Using the SSA method described in the previous chapter, several long-term components were detected and extracted from the noise, the most powerful of which have periods ranging from about 10 y, down to a few days. The presence of these significant components has also been proven by the CWT analysis of the record.

Finally, the short-term components were also analysed by applying the same spectral techniques to shorter intervals of the series with data sampled every fifteen minutes. A daily oscillation, characterized by a double-peaked shape becoming less regular during the solar cycle maxima, has been observed. A high-frequency component has also been detected, the period of which changes from solar minimum to maximum.

2.1 The Pierre Auger Observatory

The Pierre Auger Observatory [27] is located in Malargüe, Argentina (69.3° W, 35.3° S, 1400 m a.s.l.[1]) and covers an area of 3000 km^2. It was conceived to measure the flux, direction-of-arrival distribution, and mass composition of CRs at the highest energies (above 10^{15} keV) with high statistical significance over the whole sky. Due to the extremely low flux of particles at these high energies, a large detection area is required. The Observatory uses two different detectors to combine two detection techniques: the Fluorescence Detector (FD), observing the fluorescence light produced by the secondary particles as they propagate in the atmosphere, and the Surface Detector (SD), directly measuring the particles reaching the ground.

The FD [28] consists of 24 telescopes, grouped in units of six at four different locations, operating only on moonless and clear-weather nights, resulting in an operational average duty cycle of 13% of the total time. It is designed to measure the ultraviolet light emitted by excited nitrogen when an extensive air shower passes through the atmosphere.

The SD [27] is an array of more than 1600 water-Cherenkov detectors in a triangular grid with a 1500 m spacing, each consisting of a polyethylene tank (10 m^2 of area) containing 12 tonnes of high-purity water in a 1.2 m high sealed diffusive-reflective liner, providing a total of about 16000 m^2 of collection area for the full SD array. The Cherenkov light generated by the passage of charged particles through the water volume in each detector is recorded by three photomultipliers (PMTs) symmetrically placed on the top of each tank. For data acquisition, six channels with 10 bit flash analogue-to-digital converters (FADC), with a sampling frequency of 40 MHz, are used, as well as a GPS system for timing.

Figure 2.1, taken from Ref.[29], shows the layout of the Pierre Auger Observatory. The black squares represent the four locations where the FD's telescopes are placed, with gray lines indicating their field of view, while the orange dots mark the position of the Cherenkov detectors.

The detectors are sensitive to charged particles, such as $\mu^{\pm}$ and $e^{\pm}$, which essentially dominate the extensive air shower at ground level, as well as high energy photons, as these convert in e^+e^- pairs in the water volume.

At the operating gain of the surface detectors, a correspondence of 1 FADC $\approx$ 5 MeV is observed between the pulse amplitude in FADC counts and the energy deposited by a particle in the detector.

2.1.1 The Auger scaler data

The recording of low threshold rates (scalers) with all the surface detectors of the array, known as single particle technique [30], is useful for the detection of transient events such as Gamma Ray Bursts (GRBs) [31, 32], expected to produce a coherent increase in the counting rates, as well as to give insight to the rates of low-energy CRs influenced by solar modulation. This data has also been employed in studies of Forbush decreases [29, 33],

[1] above sea level

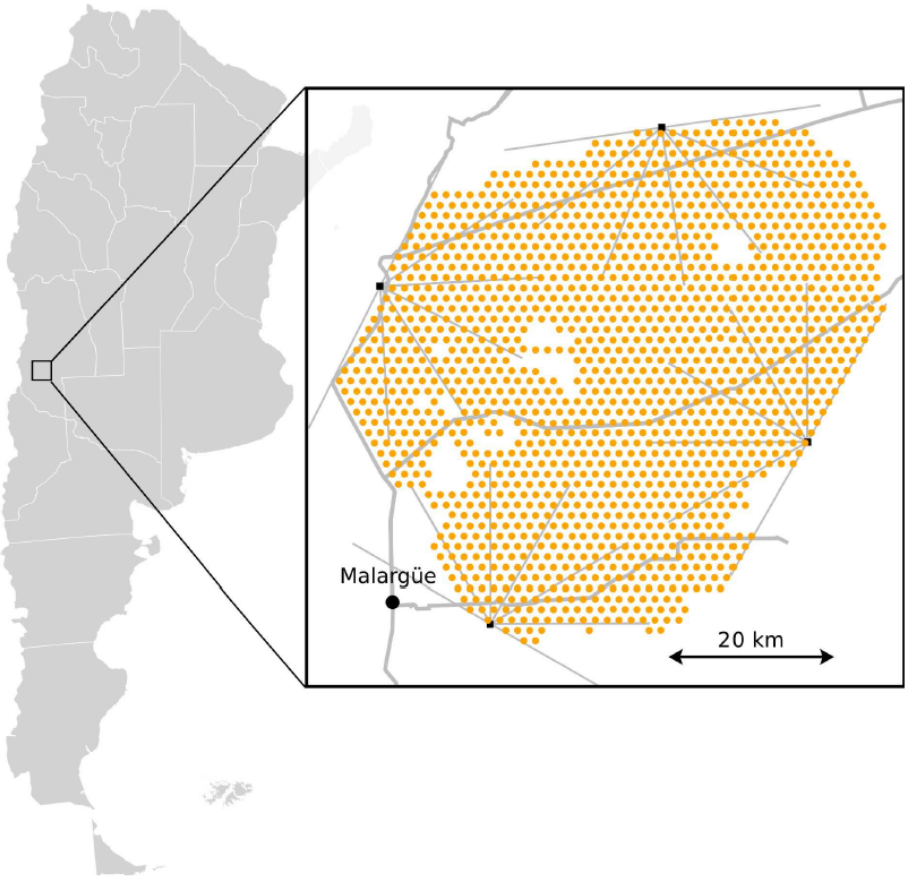

Figure 2.1: Map of the Pierre Auger Observatory at Malargüe, Argentina. The orange dots represent the positions of the surface detectors, with a 1.5 km spacing between nearest neighbors. The black squares indicate the four buildings where the FD's telescopes are placed, with gray lines indicating the field of view of the six telescopes at each location. This picture was taken from Ref.[29].

which are rapid reductions in the observed GCR intensity due to coronal mass ejections (CMEs).

The Auger scaler rates record the counting rates of signals within the range [3-20] FADC counts above baseline, corresponding approximately to the energy range [15-100] MeV, resulting in an average count rate of $\sim 1.8\times10^8$ counts per minute. Data are sent and stored every second from each detector.

Scaler rates are affected by several factors, such as atmospheric conditions, intrinsically non-constant rate of low energy particles, and eventual instrumental instabilities. Therefore, before looking for transient events and studying long-term solar modulations, the Auger scaler rates must be treated and corrected. A detailed description of all the cuts and corrections applied can be found in Ref. [34].

Some of the applied corrections are due to:

- identification of bad stations;
- removing station outliers, corresponding to very high count rates in a single second without other stations being affected as well;
- removing stations without three properly working PMTs;
- pressure correction;
- large excesses due to lightning events.

For each station, the rate measured within a specific time interval is given by the arithmetic mean over all the seconds in the considered interval not removed by one of the correction criteria discussed before. Finally, the scaler rate value associated with such a time interval is obtained by averaging over all the stations.

Because the stations are deployed on different altitudes and have different ages and PMT gains, the scaler rate is not identical between stations. To compensate for these differences, the corrected scaler rate $\Gamma_i^{(c)}$ of a station i is scaled to a known reference value $\langle\Gamma_i\rangle$, obtaining the so-called relative scaler rate $r_i(t)$ as follows:

$$r_i(t) = \frac{\Gamma_i^{(c)}(t)}{\langle\Gamma_i\rangle} \tag{2.1}$$

The chosen reference value is the mean count rate for 2013, which is roughly in the middle of the dataset. The idea behind this scaling is that the relative response of the stations to changes in the physical rate is more uniform than the absolute count rate [see Ref. 34]. Furthermore, it is a way to make the response of the whole SD array independent of the number of currently working stations.

Figure 2.2 shows the time series resulting from the mean of the relative scaler rate $r_i(t)$ over all the stations. In particular, panel a shows the relative scaler rate sampled every 12 hours, where some gaps are present due to the applied corrections. The largest gaps are highlighted in green.

The series has been gap-filled with an AR technique [7], for which the missing data are replaced with estimates extrapolated from forward and reverse AR fits of the remaining samples. The maximum number of samples used in the estimation and the order of the AR model have been chosen as the ones leading to the most accurate reconstruction of some artificial gaps randomly introduced in the series (not corresponding to already present gaps in the record). Finally, the gap-filled points are given by the average of 100 gap-filling estimates, for each of which the optimal parameters for the AR model have been defined, being the set of artificial gaps randomly generated each time.

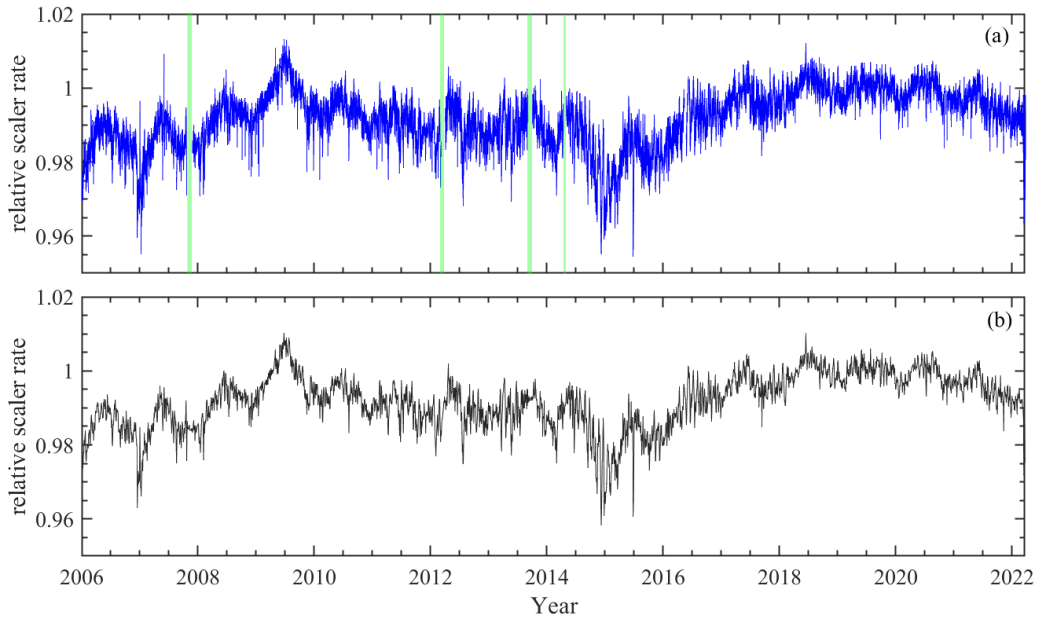

Figure 2.2: Relative scaler rates series, ranging from January 1^{st}, 2006, to March 19^{th}, 2022. The rate contains all corrections from Ref [35]. Panel a shows the series sampled every 12 hours. Some gaps are present in the series, which have been filled thanks to a gap-filling procedure based on an AR model. The largest gaps are highlighted in green. The scaler rate in panel b was obtained by resampling the original series every 48 hours after having gap-filled the series in panel a.

The filled series spans from January 1^{st}, 2006 up to the 19^{th} of March 2022 for a total of 11844 values. A decadal modulation, as well as an annual oscillation, are visible in the series. Two Forbush decreases, one occurring in December 2006 and one in the mid of June 2015, are also visible, both leading, as expected, to an intense reduction of the Auger scaler rate.

2.2 Spectral analysis

To reveal the significant periodic components in the relative scaler rate, an accurate spectral analysis has been performed on the series. The MC-SSA advanced spectral method, already introduced in Section 1.4.1 (see Appendix B for further detail), has been used to define and reconstruct the Auger scaler's significant components.

Furthermore, the spectral content has been confirmed by performing the Continuous Wavelet Transform (CWT) [12, 36]. The CWT method gives an evolutionary analysis of the series in the time-scale plane and allows studying non-stationary features of the signal, such as changes in periodicity, isolated events, trends, and intermittency (see Appendix A).

Both long- and short-term components have been extracted from the series and their origin has been investigated.

2.2.1 Long-term components

To extract the long-term oscillations present in the scaler rate record, the MC-SSA has been applyed to a resampled version of the series to reduce computational time while maintaining a good time resolution. More in detail, starting from the gap-filled version of the 12-hour-sampled series, a 48-hour-sampled one has been obtained, thus reducing the length of the series under analysis to a quarter of the initial one. Such a series is shown in Fig. 2.2b.

The adopted window length was set equal to $M = 450$, and, being $\Delta t = 48$ h, the window width W is about 2.5 years. The initial null hypothesis is a red noise, represented by a first-order autoregressive process AR(1).

The most powerful significant components revealed with the MC-SSA with a 99% c.l. are a decadal trend, an annual oscillation, and variability modes with ∼186-, ∼27- and 13.5-d periods. Figure 2.3 shows the reconstructed significant components (panels a-e), with the percentage of variance described by each (magenta values), and the SN time series sampled every 2 days (panel f). Similar periodicities have been found in neutron monitor data and in several solar indices [37–39].

Decadal variability - The decadal modulation, shown in Fig. 2.3a, is linked to the 11-y solar cycle and describes most of the signal variance, representing ∼65% of the latter. In Fig. 2.4a, this component is directly compared with the decadal one (red curve) revealed by MC-SSA in the SN time series (light red curve) by using the same window length M chosen for the Auger scaler data. Such a series has the same sampling interval as the scaler rates and was obtained by resampling the daily data available from the World Data Center (WDC) - Sunspot Index and Long-term Solar Observations (SILSO)[2], Royal Observatory of Belgium, Brussels [40]. The anticorrelation between the two series is visible.

A delay in the scaler's response to this long-term modulation is not unexpected. The time lag between GCR intensity and solar variability, particularly regarding long-term modulations, has been discussed in several papers [see, e.g., Refs. 38, 41, 42]. Such a delay is related to the polarity state of the heliosphere (positive, $A > 0$, or negative, $A < 0$). A higher delay is usually observed during odd cycles [38, 41, 42] and negative polarity states [42].

Solar cycle 24 showed an unusual polar field reversal, which has been analysed by several authors [43–45]. Following Ref.[43], the southern solar hemisphere reversed polarity in mid-2013, changing from a negative to a positive polarity state. In the northern solar hemisphere, the reversal in the field started as early as June 2012 and was followed by a period of almost zero field strength lasting until the end of 2014, before rising. The total time interval required for the polar field reversal in both hemispheres lasted from June 2012 to November 2014, as shown by the vertical gray bar in Fig. 2.4a.

In the region of positive polarity state ($A > 0$), the maximum observed in the decadal modulation of the scaler rate is almost in phase with the minimum observed in the decadal solar cycle. This is not true for the negative polarity state region, where a lag of about eight months in the Auger scaler's response is observed, lag that decreases in time, reaching a value of about three months for the local scaler's maximum around 2013. The deep minimum observed around 2015 is mainly due to the intense Forbush decrease that occurred in correspondence with such a period (see Fig. 2.2), also causing the apparent higher delay in the GCR response to the solar activity maximum at the beginning of 2014.

[2] http://www.sidc.be/silso/

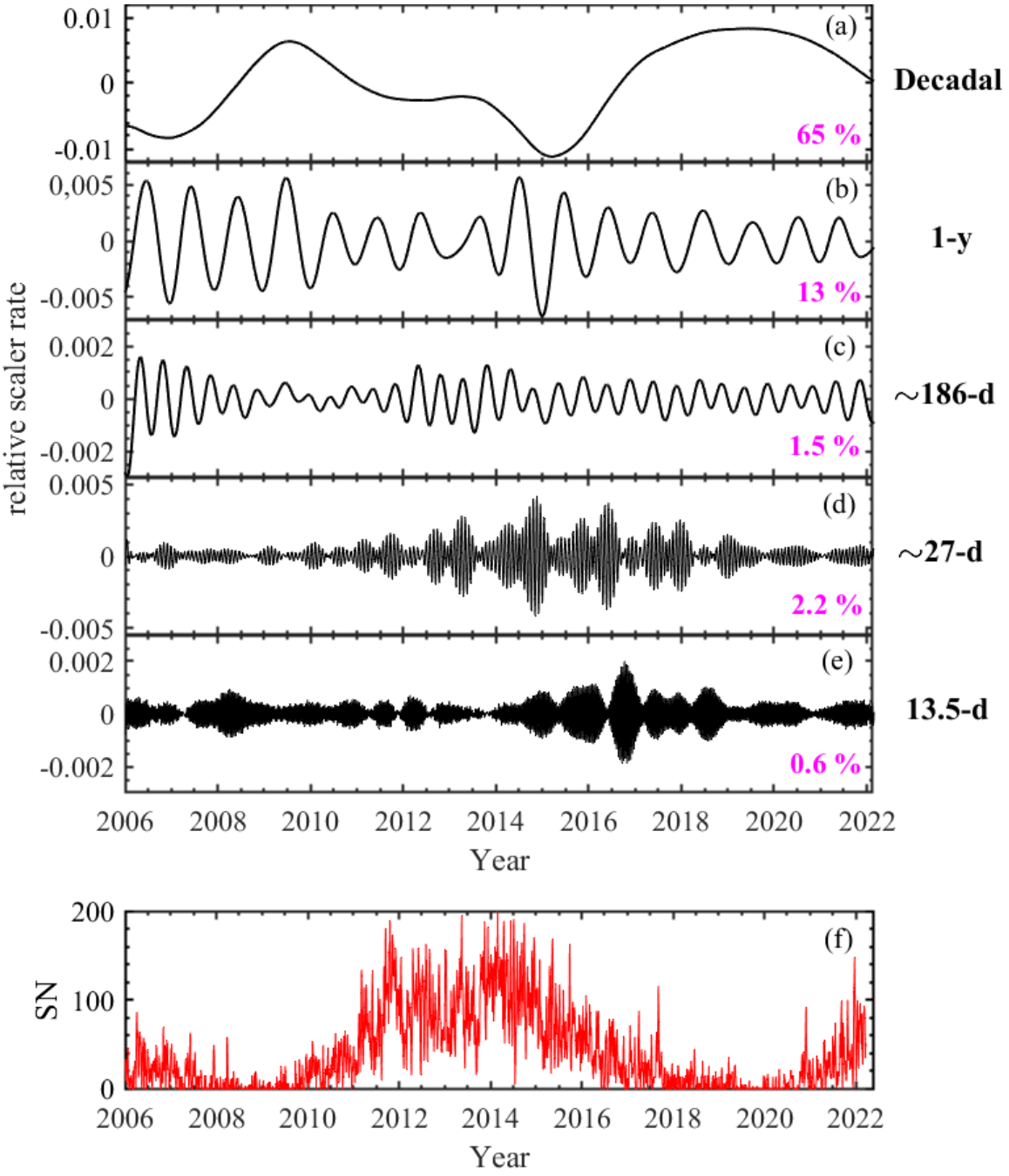

Figure 2.3: The significant MC-SSA components of the relative scaler rate series with decadal (a), annual (b), ~186-d (c), ~27-d (d), and 13.5-d (e) periods. The variance described by each component is also shown (magenta values). The SN time series sampled every 2 days is shown in panel f.

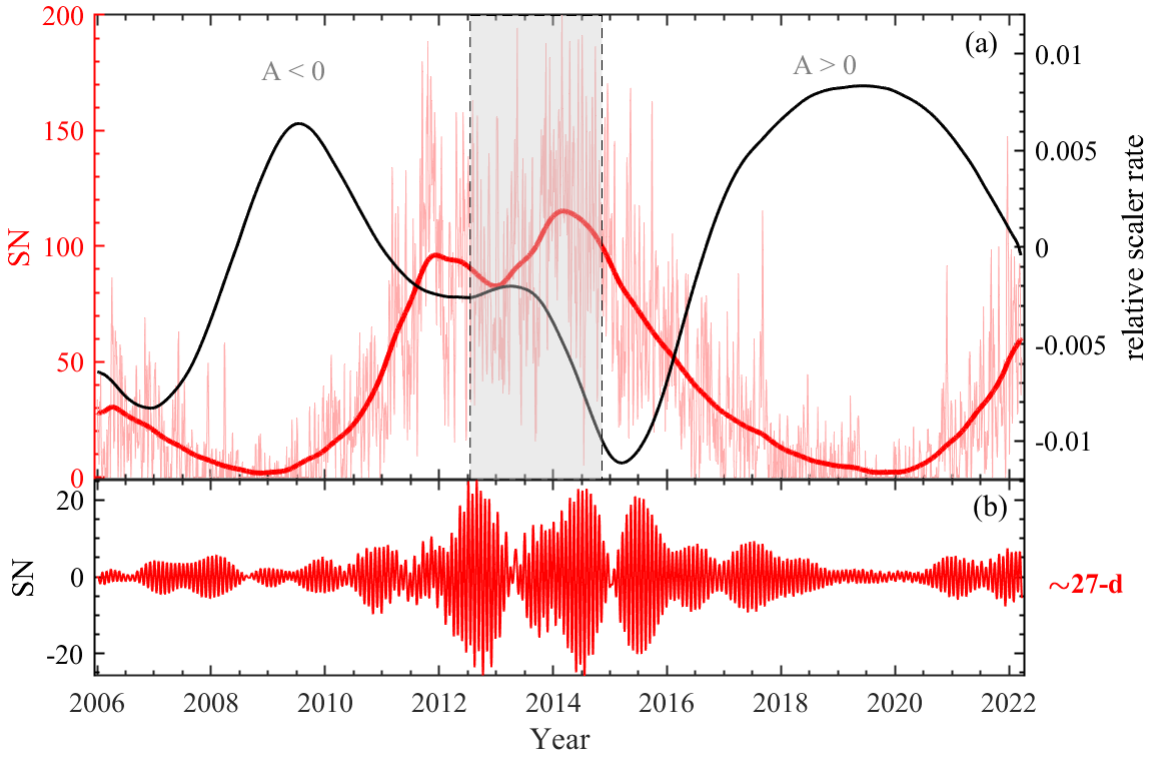

Figure 2.4: (a) Comparison between the decadal trend revealed in the Auger scaler rate (black curve) and the SN series sampled every 2 days (shaded red curve), superimposed by the decadal modulation revealed in the latter by MC-SSA (red curve). An anticorrelation among the decadal trends is clearly visible. The shaded gray bar represents the time interval from the beginning to the end of the polarity state (A) reversal for Solar Cycle 24. Panel b shows the 27-d oscillation revealed in the SN time series.

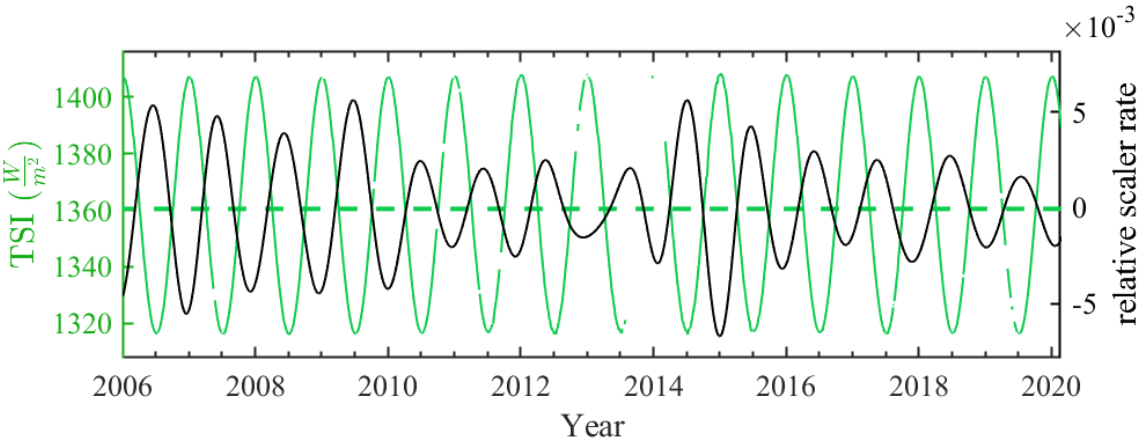

Figure 2.5: Comparison between the Auger scaler's annual modulation (black curve) revealed with MC-SSA and the TSI (green curve) at the top of the atmosphere measured by NASA's SORCE mission. The time interval shown is the one for which data are available for both the scaler rate and the TSI datasets.

The annual variability - The annual oscillation shows minima at the beginning of each year (December-January) and maxima in the middle (June-July). The origin of this component is due to the Earth's revolution around the Sun. When the Earth is at perihelion, the intensity of the GCR flux is reduced, and vice versa at the aphelion. Figure 2.5 shows the comparison between the annual component revealed by the MC-SSA in the Auger scaler rate and the total solar irradiance (TSI) measured by the Total Irradiance Monitor (TIM), launched in January 2003 aboard the NASA SOlar Radiation and Climate Experiment (SORCE), and representing the spatially and spectrally integrated solar radiation incident at the top of Earth's atmosphere[3] [46, 47]. It represents the average daily power per unit area of solar irradiance across a spherical surface surrounding the Sun with a radius equal to the Earth-Sun distance. The latter is a result of NASA's SORCE mission, which provided an invaluable data record about solar irradiance and the impact of the Sun's energy on Earth's weather, climate, and life [48]. The dotted horizontal line in Fig. 2.5 represents the mean TSI value, equal to $\sim$ 1360.5 W/m^2.

Since the physical distance of the Earth from the Sun varies during the year, the TSI at the top of the Earth's atmosphere is $\sim$6.5% higher ($\sim$ 88 W/m^2) in January (when the Earth is at perihelion) than in July, when the Earth is at the farthest point from the Sun. Although few missing data are present in the TSI record, the anti-phase with the annual oscillation of the scaler rate is visible.

The 186-d (Rieger-type) variability - The $\sim$186-d variability shown in Fig. 2.3c is the known Rieger-type periodicity [49], which can be found in several indicators of solar magnetic activity [see, e.g., Refs. 50–54]. The Rieger-type modulation was initially attributed to the 154-days periodicity found by Rieger et al. [49] in the gamma-ray flares observed by the Solar Maximum Mission (SMM) near the maximum of Solar Cycle 21. The analysis of different indicators of solar magnetic activity during the past few cycles, namely, X-ray flares [50, 55–57], sunspot group numbers [51, 52, 58–60], 10.7 cm radio flux and SN [61], occurrence rates of solar flare energetic electrons [62], type II, III, and IV radio bursts [54, 63], as well as microwave [57] and proton [53] flares, confirmed the existence of such a periodicity.

[3] https://disc.gsfc.nasa.gov/datacollection/SOR3TSID_019.html, Accessed 05/03/2022

The physical reason for the occurrence of the Rieger-type periodicity is not completely clear, and several different mechanisms have been suggested to explain the enigmatic features of this component [see, e.g., Refs. 64–69].

The long-term appearance of this periodicity has been studied, revealing that the associated period may vary from about 130 to 190 days [58, 60, 64, 70]. This component was found to be strong in some cycles and weak or lacking in others. Moreover, its period is cycle-dependent. As found in Ref.[70], the solar cycle strength is related to the value of the Rieger-type period: the stronger the solar cycle, the shorter the period.

Monthly variability - The monthly component shown in Fig. 2.3d, with a period of about 27 days, is caused by the combination of solar rotation and an inhomogeneous distribution of long-lived solar active regions, such as sunspots, CMEs, coronal holes, and co-rotating interaction regions [71]. It is linked to the longitudinal asymmetry of the electromagnetic conditions in the heliosphere during one solar rotation.

Due to the solar differential rotation, the rotational period is of about 25 days at the near-equatorial region and reaches values around 38 days near the poles. The rotation period at the equator is known as *sidereal rotation period*, which should not be confused with the *synodic rotation period* of about 26 days, which is the time for a fixed feature on the Sun to rotate to the same apparent position as viewed from Earth. In literature, *Carrington rotation* is often used to define a synodic rotation period of about 27 days, corresponding to a latitude of 26°.

This variability has already been found in several neutron monitor data [see, e.g., Refs. 38, 39, 72, 73] and has also been detected in the SN time series, as shown in Fig. 2.4b.

The amplitude of the 27-d periodic component in the Auger scaler rate reaches its maximum value around solar activity maximum. Furthermore, a higher variability is observed until the beginning of 2018, i.e., for the whole duration of the descending phase of the cycle, most probably because of more robust solar activities in the form of more intense solar flares and CMEs. Ref.[74] shows that as the CMEs go down during the rising phase of solar cycles, the 27-d period modulation too is lower, and this is in agreement with the intensity of the monthly component revealed in Auger scaler data (see Fig. 2.3d).

The 13.5-d variability - The 13.5-d periodic component shown in Fig. 2.3e is associated with both solar active longitudes and tilted dipole structure. It has been found in neutron monitor data and in several solar indices as shown in Ref.[39]. The MC-SSA shows the possible presence of such a component in the SN data, although its power is smaller compared to the other components, and further analyses are needed to confirm its presence.

As discussed in Ref.[74], the 13.5-d period is observed to be more prominent during the declining phase of solar cycles, for the same reasons previously discussed for the 27-d component, and this is in agreement with the variability observed in the reconstructed half-monthly component. As visible in Fig. 2.3e, a higher variability is observed between 2015 and 2019, and around 2008 as well, i.e., in correspondence with solar cycles' declining phases.

All the long-term components revealed by MC-SSA have been confirmed by the CWT spectral analysis. Figure 2.6b shows the continuous wavelet spectrum of the Auger scaler rate (panel a), where areas with a high (low) power are represented by the color red (blue). The black curves enclose regions with a confidence level greater than 90% against the null hypothesis of a red-noise process. The COI, delimiting regions influenced by edge effects, is represented by the white curve. Figure 2.6c shows the Global Wavelet Spectrum (GWS) (black curve), obtained by time-averaging the CWT spectrum over each period, and the corresponding significance levels (red curve).

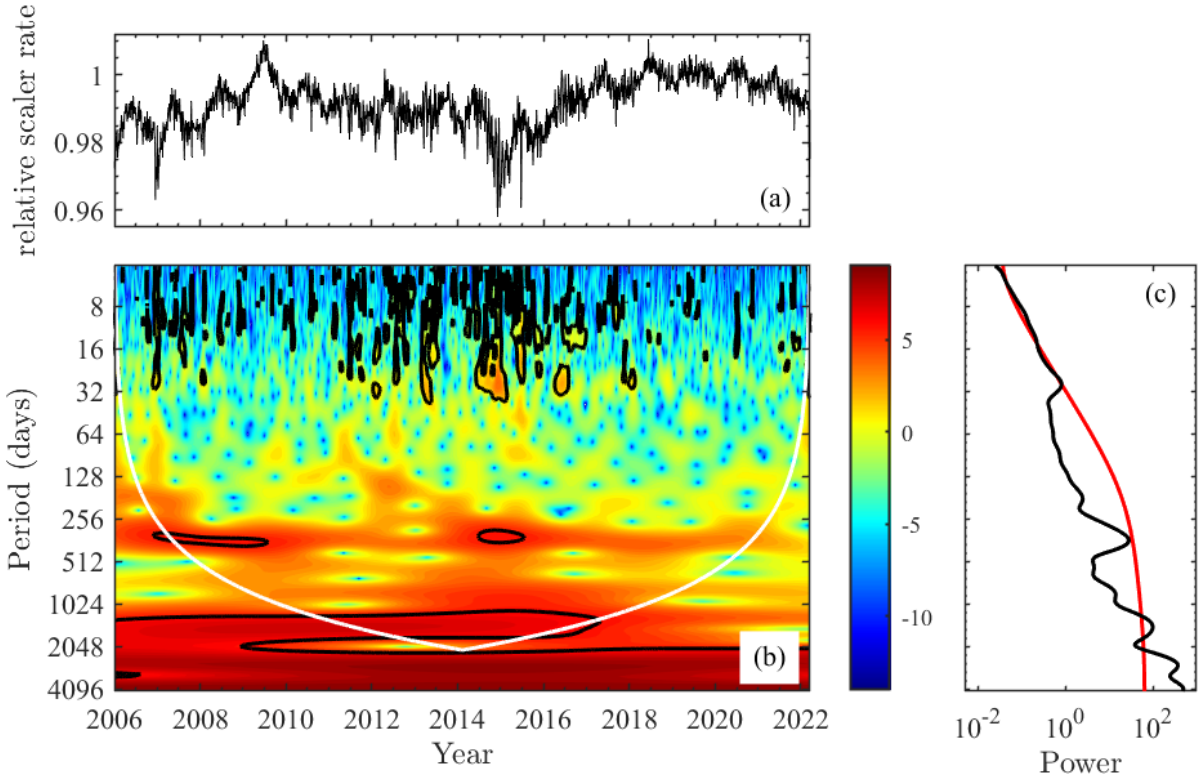

Figure 2.6: Continuous Wavelet power spectrum (panel b) of the Auger scaler rate sampled every 48 hours (panel a) and Global Wavelet Spectrum (panel c). The black curves in panel b enclose regions with a confidence level greater than 90% against a red-noise process null hypothesis, while the COI, indicating regions influenced by edge effects, is represented by the white curve.

The decadal modulation shows the highest power, although it lies outside the COI due to the limited length of the series. The annual periodicity is also detected, showing higher and significant power in correspondence with the time intervals in which its amplitude is larger. The monthly oscillation is highly significant between 2012 and 2017, while it shows a lower power elsewhere. This result is in agreement with that found by the MC-SSA (Fig. 2.3d), i.e., the reconstructed monthly component has higher variability in correspondence with the high-power regions detected in the wavelet spectrum. Furthermore, a peak in the GWS between 8 and 16 days of period is observed, corresponding to the 13.5-d significant component previously discussed. A peak in correspondence with 186-d of period is observed in the GWS (Fig. 2.6c), although it is not significant due to the limitations of the method in detecting peaks showing a large difference in amplitude and close to each other.

2.2.2 Short-term components

To analyse short-term variabilities in Auger scalers, the CWT power spectrum has been estimated for data recorded in the same month (June) of four different years with a sampling interval of 15 minutes. Furthermore, the MC-SSA has been applied on each record to confirm the results obtained with the CWT method and to extract the significant periodic components from the noise. Small gaps in all the series have been filled with the AR model previously described, while large gaps, usually present at the beginning and/or at the end of the records, have not been filled. Two of the series under analysis, those relating to 2010 and 2020, shown in Figs. 2.7a and 2.10a, respectively, correspond to solar minima. The other two, relating to 2012 and 2014, correspond to solar maxima and are shown in Figs. 2.8a and 2.9a, respectively. For a direct comparison of the records, the same range has been used on the y-axis. It is possible to notice that the signals recorded during solar minima show a more regular pattern to those revealed during the solar maximum.

The CWT power spectra of the records show, with a 99% c.l., the presence of daily and half-daily significant components. During solar minimum (2010 and 2020), these components exhibit high power throughout the signals (see Figs. 2.7b and 2.10b). Additionally, an intermittent high-power signal is observed in correspondence with a period of about 8 hours.

Regarding the years 2012 and 2014 (see Figs. 2.8b and 2.9b), the results are different compared to the previous two cases. The 12-h components lose power in several intervals of the records. The 8-h component revealed in the previous two records is not present in these, although high-power sporadic regions are present around 5 hours of period. Moreover, in June 2014, the 24-h modulation also loses power in several intervals of the signal.

With a window length M=200 ($W \simeq 2$ d), the significant components of all four records have been detected with the MC-SSA method with a 99% c.l. against the null hypothesis of a red noise process. Figure 2.11 shows the reconstructed significant components of the four records, together with the variance described by each (magenta values). The first row shows the reconstructed daily components revealed in each series combined with a trend, while the second row shows those with a 12-h period. The half-daily component detected in June 2012 shows a smaller amplitude and describes a lower variance than those extracted in the other three records. The third row in Fig. 2.11 shows the high-frequency significant component detected in the records, also referred to as intra-daily oscillations, describing a lower variance compared to the daily and half-daily ones, and which period shows a variation from one solar minimum to the following. In 2010 and 2020 (solar minima), such a component shows a period of about 8 hours, while, during solar maxima (2012 and 2014), the period is slightly smaller, between 5 and 6 hours, thus confirming the results obtained with the CWT analysis (see Figs. 2.7b, 2.8b, 2.9b, and 2.10b).

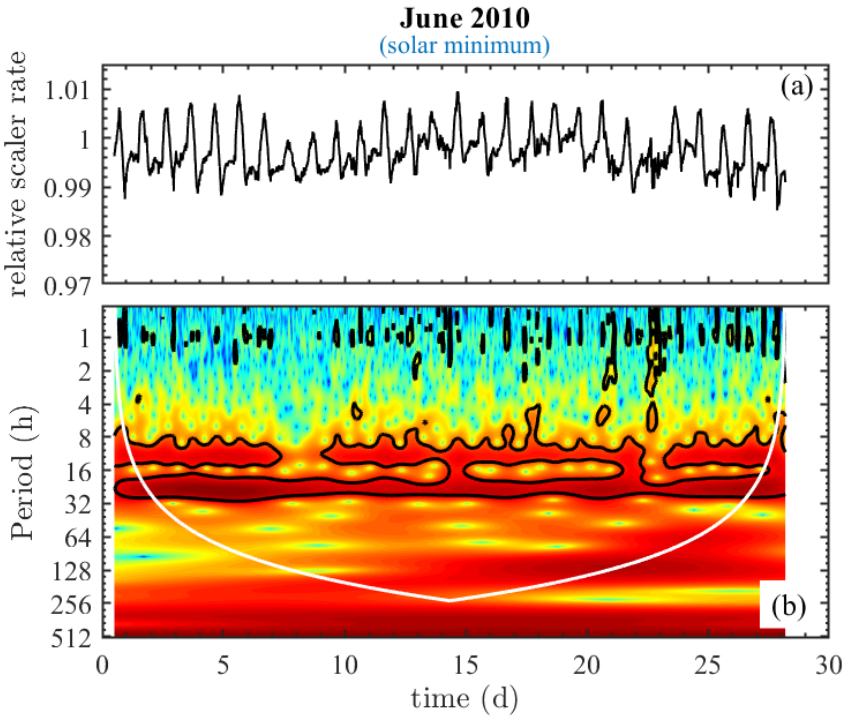

Figure 2.7: (a) Auger scaler rate relative to June 2010, with data sampled every 15 minutes. Panel b shows the wavelet power spectrum of the record. The black curves enclose regions with a confidence level greater than 99% against a red-noise process null hypothesis. The white curve represents the COI, which indicates the area in the time-scale plane influenced by edge effects.

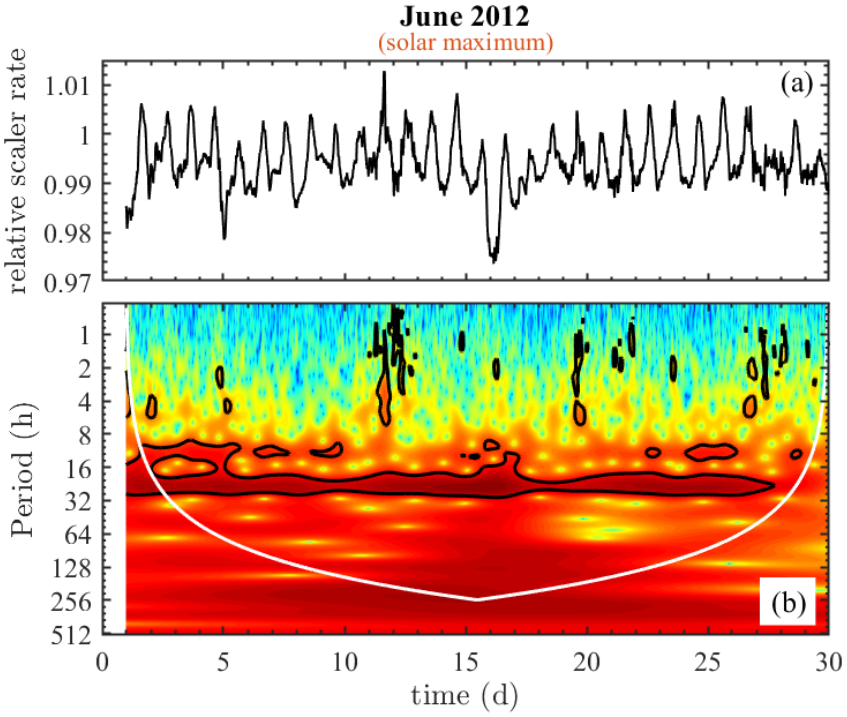

Figure 2.8: (a) Auger scaler rate relative to June 2012, with data sampled every 15 minutes. Panel b shows the wavelet power spectrum of the record. The black curves enclose regions with a confidence level greater than 99% against a red-noise process null hypothesis. The white curve represents the COI, which indicates the area in the time-scale plane influenced by edge effects.

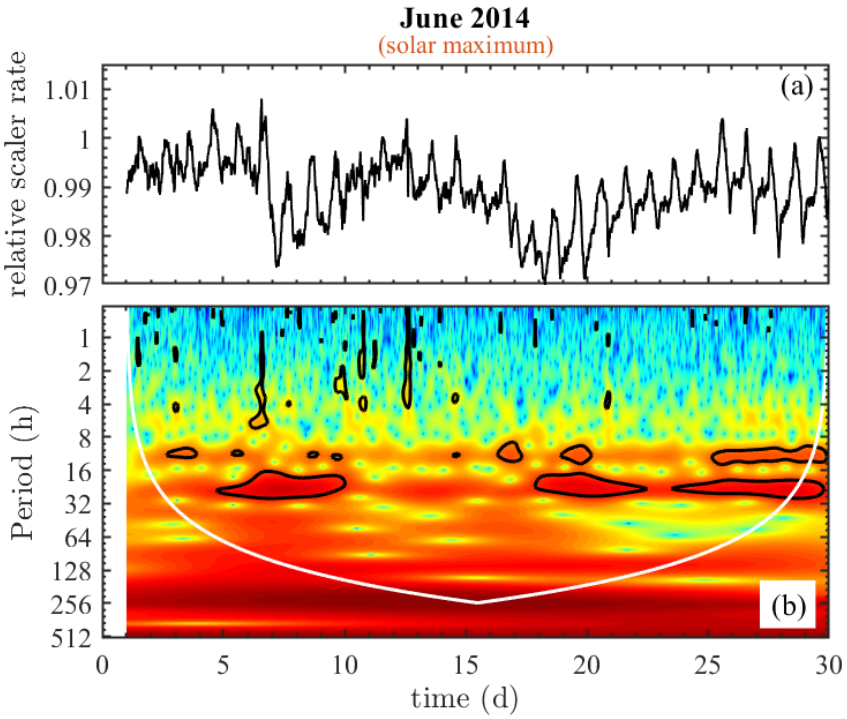

Figure 2.9: (a) Auger scaler rate relative to June 2014, with data sampled every 15 minutes. Panel b shows the wavelet power spectrum of the record. The black curves enclose regions with a confidence level greater than 99% against a red-noise process null hypothesis. The white curve represents the COI, which indicates the area in the time-scale plane influenced by edge effects.

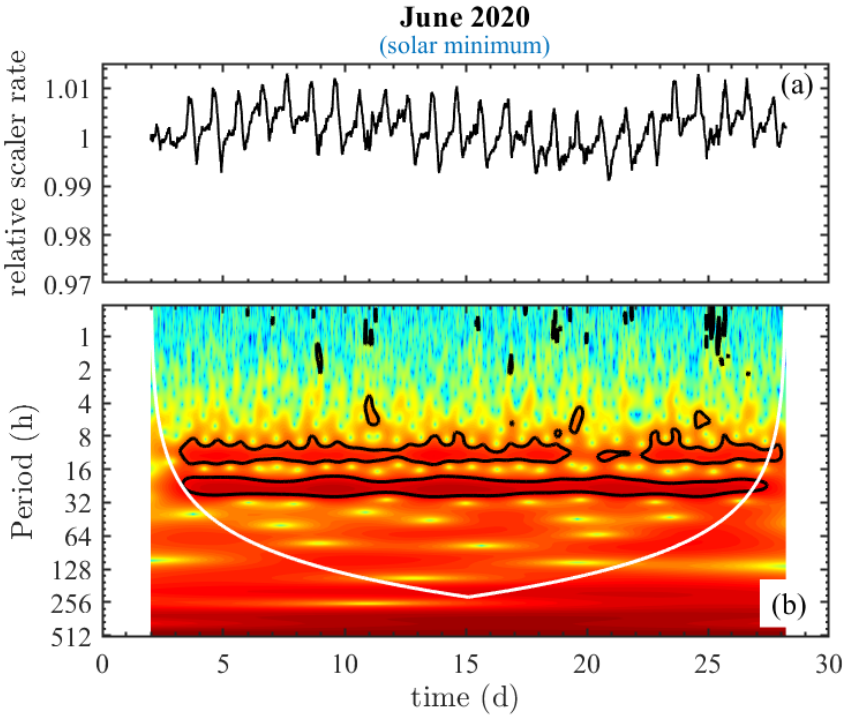

Figure 2.10: (a) Auger scaler rate relative to June 2020, with data sampled every 15 minutes. Panel b shows the wavelet power spectrum of the record. The black curves enclose regions with a confidence level greater than 99% against a red-noise process null hypothesis. The white curve represents the COI, which indicates the area in the time-scale plane influenced by edge effects.

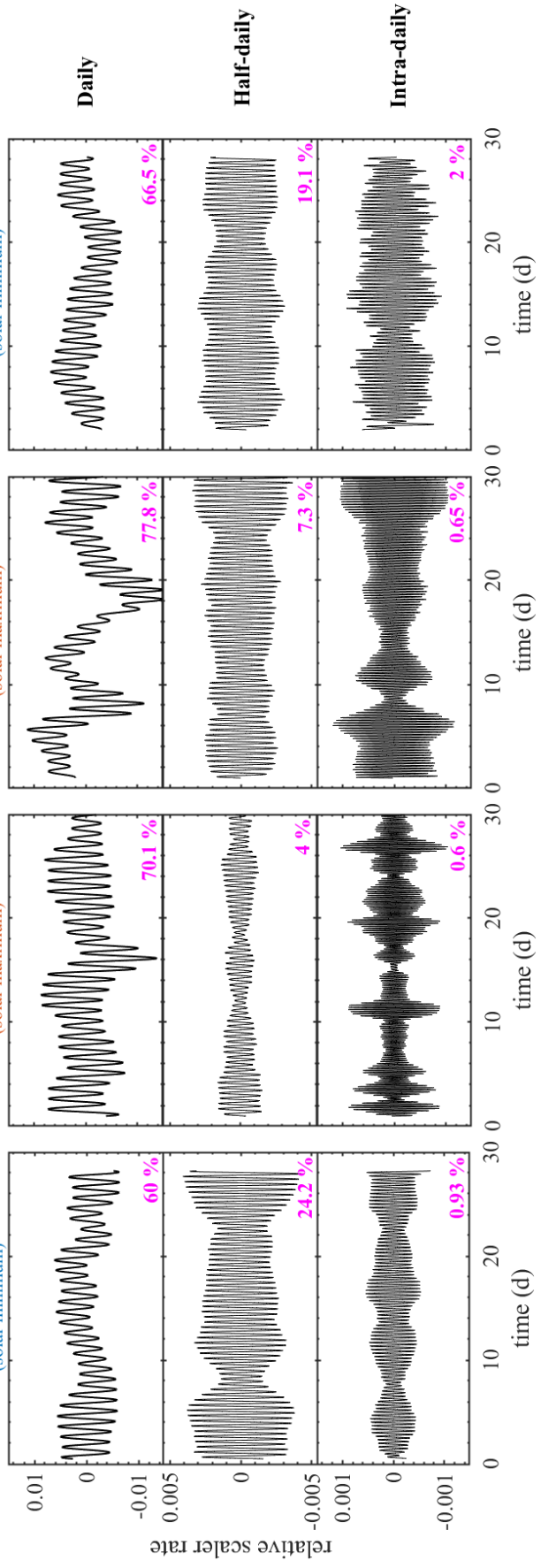

Figure 2.11: Reconstruction of the significant periodic components revealed in the Auger scaler data sampled every 15 minutes relative to the same month (June) of four different years. A daily (first row) and a half-daily (second row) periodic components have been detected. A third high-frequency oscillation is also present (third row), showing a period of about 8-h during 2010 and 2020 (solar minima) and a shorter one, of about 6-h, in 2012 and 2014 (solar maxima). The variance described by each component is also shown (magenta values).

The sum of all the significant components, i.e., the de-noised version of the signals, are shown in Fig. 2.12, together with the total variance (magenta values). The half-daily component is responsible for the double-peaked shape observed in all the records, except in June 2012, where this signal showed a lower amplitude (see Fig. 2.11).

The long-term behaviour of the daily variation has been discussed in Ref.[34] by comparing the daily profiles across the 15-year-long Auger scaler record. The monthly Auger scaler rates are shown in the left panel of Fig. 2.13, taken from Ref.[34], where each month was normalized by rescaling it with the mean scaler rate of the respective month, allowing to emphasize the shape of the daily signal instead of highlighting the long-term modulations. A seasonal variation of the 24-h signal is visible, with the peak's time of occurrence showing a shift of approximately three hours over a year. In January (austral summer), the daily signal maximum is at about 14:00 UTC (11:00 a.m. local time), whereas in austral winter, the maximum occurs at 17:00 UTC. During the solar maximum between 2012 and 2014, the observed seasonal variation was less regular due to a higher rate of transient events.

The right panel of Fig. 2.13 compares the average daily profile of June 2018 (black curve) with the profiles of two individual days of the same month, the latter rescaled with their average rate, to enable the comparison of the shapes themselves. It is visible that the amplitudes and detail of the diurnal variations are distinctively different.

Such results highlight the abilities of the Auger scaler rates in resolving structures on hourly scales with high statistics. The phenomena at the origin of these short-term fluctuations need further investigation. Since the Auger data have already been corrected for pressure [34, 35], a correlation with the latter is not expected, while a possible link with other factors, such as temperature, is currently being studied.

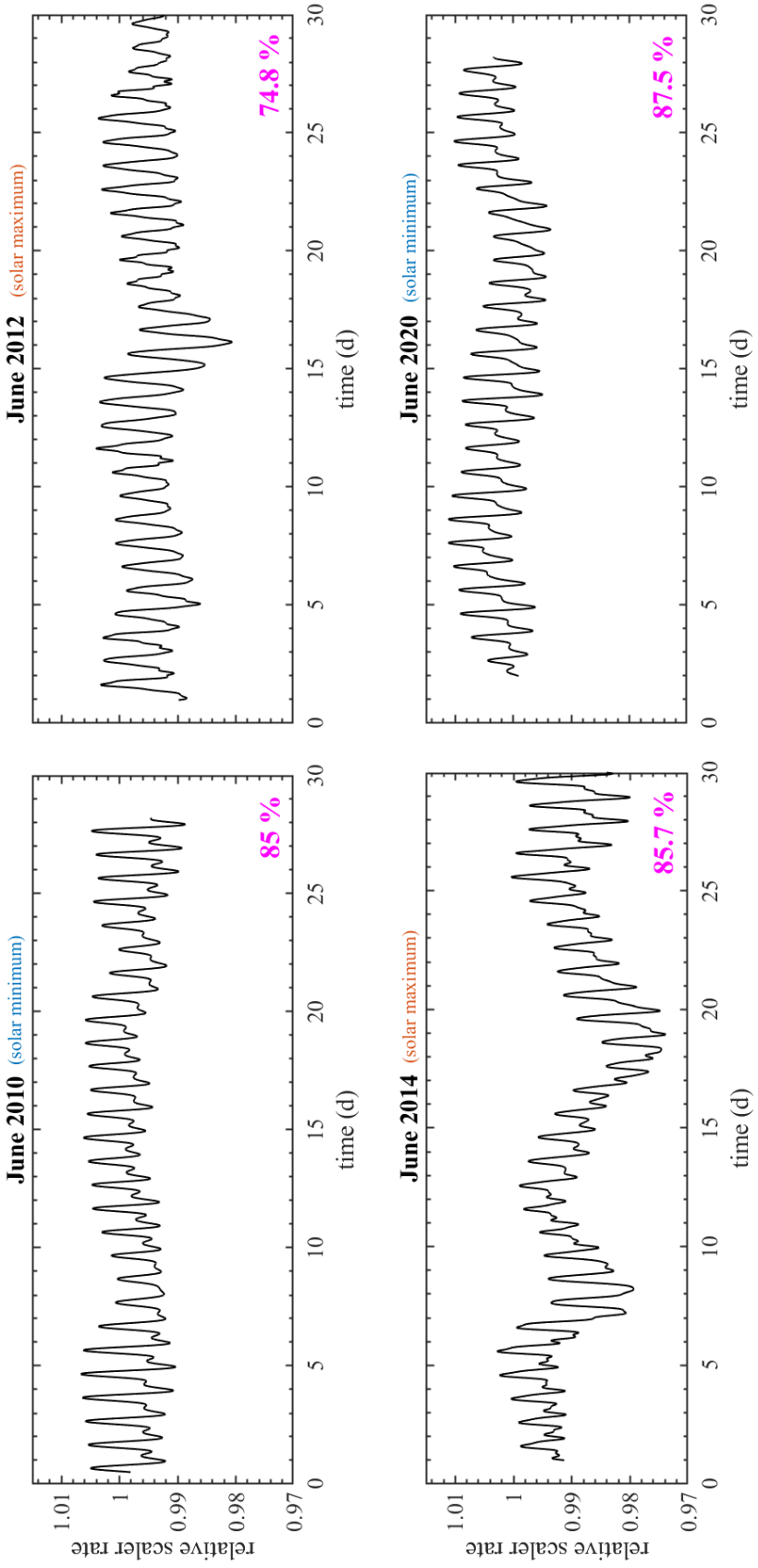

Figure 2.12: Sum of all the significant reconstructed components, representing the de-noised versions of the four signals. The total variance described by each reconstructed signal is also shown (magenta values).

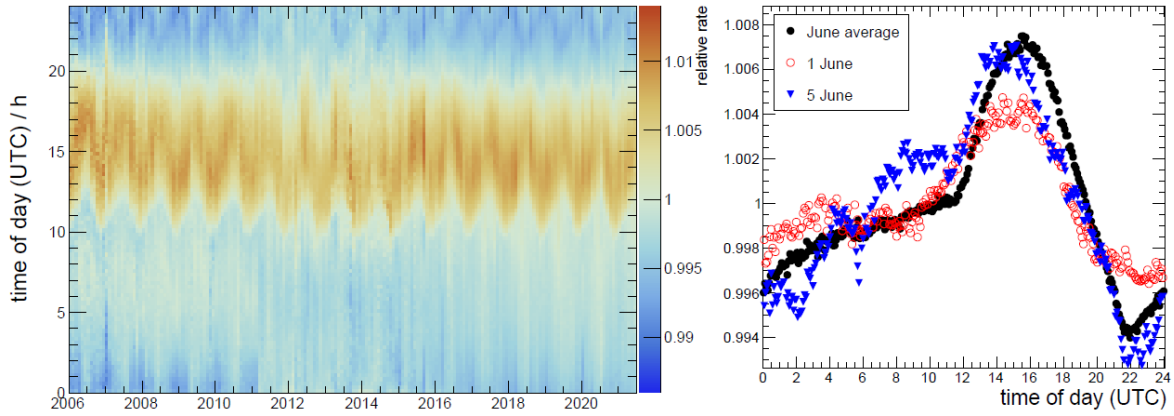

Figure 2.13: Picture taken from Ref.[34]. The left panel shows the monthly average values, where each month was normalized separately to emphasize the shape rather than the long-term decadal trend. On the right, the average daily shape of the observed diurnal pattern of the scaler rate is shown for June 2018 (black curve), together with two single days of the same month as examples of the variations in this daily pattern (red and blue curves).

2.3 Conclusions

By use of the MC-SSA, several long-term components have been detected in the 16-year-long Auger scaler rate record sampled every 48 hours with a 99% c.l. against the null hypothesis of a red noise process. The presence of these periodic components was also confirmed by the CWT analysis.

A dominant decadal modulation, describing most of the signal variance, has been detected, revealing an anticorrelation with the 11-y solar cycle. This component shows a delay of a few months with respect to the decadal solar cycle in correspondence with periods of negative polarity state of the heliosphere.

An annual oscillation, showing maxima in correspondence with austral winters and minima during austral summers, is also present, the origin of which is linked to the Earth-Sun distance variation during the year.

Periodic oscillations with periods of about 186, 27, and 13.5 days have been detected. The first one is the known Rieger-type periodicity, which has been detected in several neutron monitor data, as well as in many solar indices, and which period is inversely proportional to the solar cycle strength. The 27-d modulation is caused by the combination of an inhomogeneous distribution of long-lived solar active regions and solar rotation. The 27-d oscillation in the Auger scaler shows a higher variability in correspondence with both the maximum and the descending phase of Solar Cycle 24. In correspondence with solar cycles' declining phases, more intense solar flares and CMEs usually occur, thus possibly explaining the observed intensity of the monthly component. As the CMEs go down during the rising phase of solar cycles, the 27-d modulation is lower during these periods. This periodicity has also been found in the SN time series. The 13.5-d periodic component is associated with both solar active longitudes and tilted dipole structure, and it has been found in neutron monitor data and several solar indices as well. This component is observed to be more prominent during the declining phase of solar cycles, presenting a higher variability between 2015 and 2019, and around 2008 as well. This is probably due to the same causes discussed for the 27-d component.

Short-term components have been analysed by estimating the wavelet power spectrum of the Auger scaler rate recorded in the same month (June) of four different years sampled every 15 minutes. The significant periodic components have also been detected and reconstructed with the MC-SSA with a 99% c.l.. Two of the analysed series are in correspondence with solar minima (June 2010 and June 2020), while the other two are in correspondence with the maximum activity of Solar Cycle 24 (June 2012 and June 2014). It is possible to notice that the signals recorded during the minimum activity of the Sun show a more regular pattern compared to the others.

Daily and half-daily signals have been detected in all the records. As of June 2014, the 24-h modulation loses power in several intervals of the record. In June 2012, the 12-h oscillation, which describes the observed double-peaked daily shape, is much smaller and explains a lower variance compared to the same component revealed in the other three records.

A high-frequency significant component has also been revealed, showing a period of about 8 hours during solar minima, and a shorter period, comprised between 5 and 6 hours, during Solar Cycle 24 maximum.

The phenomena at the origin of these short-term fluctuations are not clearly known, and a deep investigation is needed to fully understand them. Analysing data recorded in the same month of different years allowed the detection of possible influences due to solar activity. By performing the same analyses on data recorded in different months within the same year, it will be possible to understand the evolution of the daily pattern over time. However, such results highlight the abilities of the Auger scaler rates in resolving structures on hourly scales with high statistics.

Chapter 3

Cosmogenic radioisotopes and meteorites

The concentration of cosmogenic radionuclides in meteorites is a valuable proxy of solar activity [75]. Besides the dependence on the flux of primary galactic cosmic rays (GCR) in space during approximately one half-life of the isotope before the meteorite falls to Earth, where the cosmic ray (CR) irradiation becomes negligible, the production rate of a radioisotope also depends on the composition of the meteorite, the size of the parent meteoroid, and the shielding depth. Solar Cosmic Rays (SCR) produce radioisotopes by spallation reactions only in the outer layers of the meteoroids. Because of the ablation of surface layers, consisting of the loss of mass due to overheating from friction with atmospheric molecules and fragmentation caused by aerodynamic pressure exceeding the crushing strength of the meteoroid, the contribution in isotope production due to SCR is negligible. Therefore, the abundance of cosmogenic radioisotopes in meteorites can serve as a proxy for GCR flux variations, and the measurement of radioisotopes with different half-lives allows these variations to be analysed at different time scales. In particular, the activity of ^{22}Na, having a half-life of about three years, can unveil the 11-year solar cycle. On the other side, the measurement of long-lived cosmogenic radioisotopes provides an independent estimation of the meteoroid size and allows defining the average shielding conditions of the meteorite during its CR exposure. Furthermore, freshly-fallen meteorites are of great interest in planetary science as they allow measuring the activity of cosmogenic radioisotopes with very short half-lives.

In this chapter, the gamma-ray measurement of the Cavezzo L5 anomalous chondrite, recovered on January 4$^{\text{th}}$, 2020, just three days after the fall observed over Northern Italy skies by the PRISMA fireball network, is detailed described. The measurement took place at the Monte dei Cappuccini underground Research Station (Torino, Italy) with a highly selective HPGe-NaI(Tl) spectrometer. Thanks to the high efficiency and low background of the apparatus, fifteen cosmogenic radioisotopes have been detected with half-lives down to a few days. The activity of each radionuclide has been determined after having estimated the dead layer thickness of the HPGe detector (non-active volume of the crystal) and the detection efficiency as a function of the energy with Geant4-based simulations. Moreover, to identify radioisotopes with low gamma activity, a method based on the detection of coincident events recorded over the HPGe detector and the NaI(Tl) scintillator has been optimized. Finally, the ^{22}Na/^{26}Al activity ratio was calculated and compared with the estimated values for L chondrites.

Part of the results discussed in this chapter have been published [see Ref. 76].

3.1 The Cavezzo meteorite: recovery and classification

The PRISMA (Prima Rete Italiana per la Sorveglianza sistematica di Meteore e Atmosfera) network [77, 78], born in 2016 , was conceived to establish a national fireball network in Italy and is a partner of the FRIPON (Fireball Recovery and InterPlanetary Observation Network) collaboration [79]. The network aims to monitor fireballs and bolides by using all-sky cameras distributed all over the country, thus achieving systematic surveillance of the Italian skies. The network has installed almost 70 stations, including operating ones and those in the installation phase, and each station is equipped with a CCD camera (6mm diagonal, 1296 x 966 pixels) operated at 30 Hz, coupled with a short focal lens objective (1.25 mm), to obtain an all-sky field of view. Other similar networks exist, whose primary objectives are bright meteors and the recovery of meteorites, and the birth of the PRISMA project marked the beginning of Italian participation in a network of European observing facilities.

On January 1st, 2020, at 18:26:53 UT, eight PRISMA stations detected a brilliant fireball in the skies of Northern Italy. Figure 3.1 (taken from Ref.[80]) shows the map of the PRISMA stations that detected the fireball (white dots) and the images of the meteor trail seen by each camera. The red line represents the on-ground projection of the reconstructed trajectory.

The analysis of these observations allowed to determine a meteorite residue was most likely to have survived the atmospheric transit. The expected strewn field, computed using a purely ablative model, was confined in an area of about 5 km^2 near the municipality of Cavezzo, Modena [80]. Similar estimates about the atmospheric path of fallen meteorites are available only for 35 meteorites among all the officially classified ones [79, 80]. The analysis of the light-curve profile suggested that the object underwent a fragmentation process during the atmospheric flight at about 30 km height from the ground, so more than one fragment was expected along the final part of its trajectory.

Three days after the bolide was observed, two meteorite specimens were found within the predicted area of fall, at coordinates 44°49′43″.7 N 10°58′19″.5 E. The samples are shown in Fig. 3.2. Specimen F1 weighs 3.12 g, while specimen F2, the largest one, weighs 52.19 g. An evident fusion crust is visible in both samples, covering most of their surface. Sample F2 also presents a less pronounced secondary fusion crust (darker grey coloration on one side), which is compatible with exposure related to the fragmentation event during its atmospheric transit. More details about the fireball data analysis and Cavezzo search and recovery are reported in Ref.[80].

The analysis for classification and characterization of the Cavezzo meteorite has been carried out at the Earth Sciences Department and Natural History Museum of the Firenze University, revealing a discrepancy between the two specimens. The texture, crystal chemistry, and modal mineralogy of specimen F2 are typical of an ordinary L chondrite, while specimen F1 presents anomalous texture, structure, and modal mineralogy compared to the latter. Specimen F1 also shows different lithological and geochemical characteristics, oxygen isotopic composition, and rare element patterns. Because of these discrepancies, the Cavezzo meteorite has been classified as an anomalous L5 chondrite, being the first of this class. Further details about the analysis of the mineralogy, petrography, and geochemistry of the Cavezzo meteorite are reported in Ref.[81].

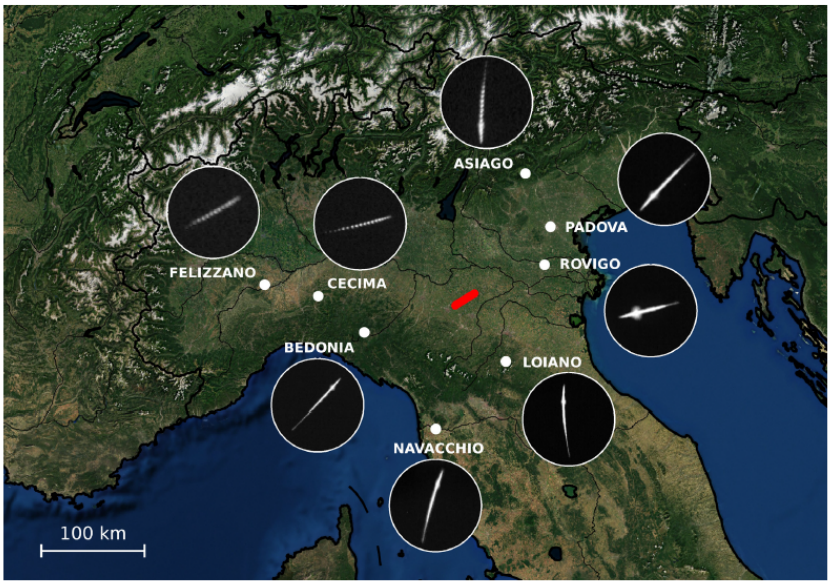

Figure 3.1: Map of the PRISMA stations (white dots) involved in the detection of the brilliant fireball observed on January 1^{st}, 2020. The red line represents the projection on the ground of the fireball bright trajectory. White circles enclose the fireball trail seen by each camera. This picture was taken from Ref.[80].

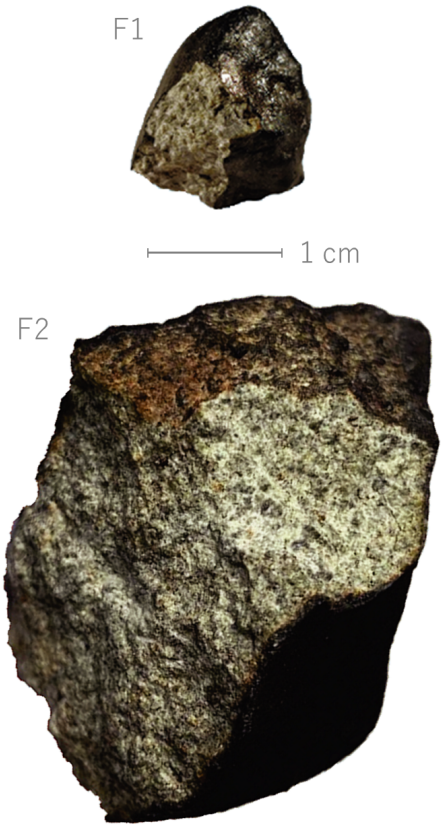

Figure 3.2: The two specimens of the Cavezzo meteorite. An evident fusion crust is visible in both samples, covering most of their surface. The biggest specimen (F2) also presents a less pronounced secondary fusion crust that may be related to fragmentation during its atmospheric transit. A small portion of their surface does not show any fusion features, probably because of the impact on the ground [80].

3.2 Cosmogenic radionuclides in the Cavezzo meteorite

To determine the activities of cosmogenic radionuclides in the Cavezzo meteorite, the F2 sample was measured with a high-efficiency HPGe-NaI(Tl) spectrometer, named GEM90 [82]. The facility is located in the underground Laboratory (70 m.w.e.[1]) of Monte dei Cappuccini in Torino, where several meteorites have already been measured in the past decades [see, e.g., Refs. 82–86]. The detector allows for non-destructive and highly selective measurement of the gamma-activity of the meteorite. The measurement of the main mass of the Cavezzo meteorite indicated the presence of fifteen cosmogenic isotopes with half-lives down to a few days, thus confirming the recent fall of the sample. In the following sections, the experimental setup is described and the list of all the radioisotopes revealed in the F2 sample is shown.

3.2.1 Experimental apparatus: HPGe-NaI(Tl) (GEM90)

The GEM90 spectrometer consists of a p-type closed-ended coaxial hyperpure germanium (HPGe) crystal operating in coincidence with an umbrella of NaI(Tl) scintillator. It can hold samples up to 1 kg of mass. Figure 3.3 shows a schematic representation of the GEM90 detector.

The HPGe detector has a mass of $\sim$ 2 kg, is 79.1 mm in height, 77.4 mm in diameter, and has a volume of 372 cm^3. For the 1332.5 keV ^{60}Co gamma-rays, it has a relative efficiency[2] of 95%, a peak-to-Compton ratio of 88, and a resolution (FWHM[3]) of 1.93 keV. Furthermore, it is surrounded by a 1 mm-thick aluminum crystal.

The NaI(Tl) scintillator, surrounding the Ge detector, has a detecting mass of 55 kg, a total volume of 15000 cm^3, and a resolution at 662 keV (gamma-ray of ^{137}Cs) equal to 8.5%. It is composed of an annular ring and a cylindrical plug sealing the cavity. The annular ring has a mass of 51 kg, is 229 mm high, has a diameter of 305 mm, and a cavity of 12.4 cm hosting the HPGe crystal. The plug has a mass of $\sim$4 kg, is 100 mm high, and has a diameter of 120 mm.

To record signals, the NaI annular ring (or annulus) is coupled with six photomultipliers (PMTs) placed on its upper surface, whereas the cylindrical plug is equipped with only one PMT.

Nitrogen is continuously flushed inside the cavity to minimize the contribution of the ambient radon and its decay products, prevent humidity, and avoid condensation on the cold parts (including electronics).

Although the natural shielding of the rock under which the Laboratory is located provides a cosmic μ-rate 30 times less than at the surface level, a passive Pb-Cd-Cu shield (200-1-40 mm in thickness, respectively) surrounds the spectrometer to further reduce the incoming radiation from the environment, thus considerably reducing the background counts.

[1] meter water equivalent

[2] Photopeak efficiency relative to that of a standard 3 in. $\times$ 3 in. cylindrical NaI(Tl) scintillation crystal, for the 1332.5 keV gamma rays of a ^{60}Co source on-axis at a distance of 25 cm from the detector.

[3] Full Width at Half Maximum

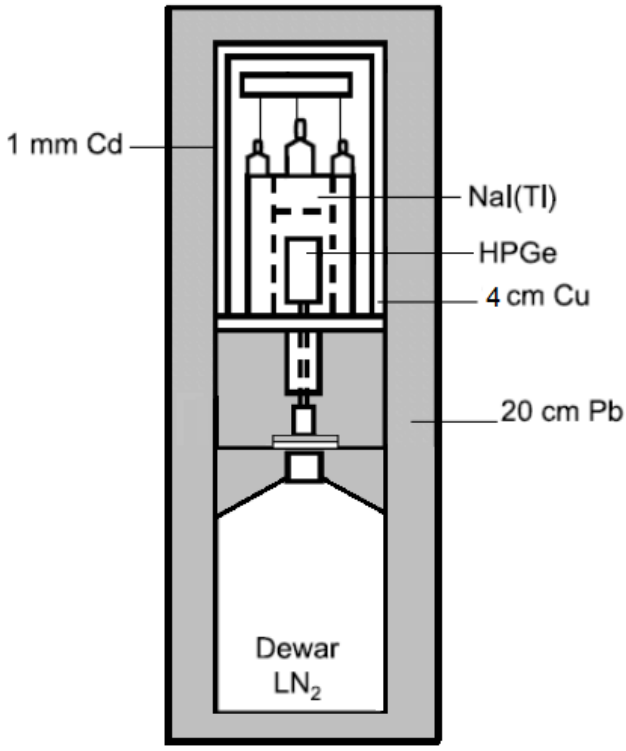

Figure 3.3: Schematic representation of the GEM90 detector.

An isolated dewar containing up to 30 liters of liquid nitrogen (N_2) is held in thermal contact with the detector, keeping the latter at a low temperature. This process inhibits the thermal generation of electron/hole pairs and the resulting leakage current that would interfere with detecting gamma energy deposition.

The GEM90 multi-parametric acquisition system allows for independently recording the signals from HPGe and NaI(Tl), which are stored together with their timestamps [84, 87], allowing to post-process the data and to detect events in coincidence between Ge and NaI (see Section 3.3).

The gamma-activity of the Cavezzo meteorite was determined by measuring the biggest sample (F2) with the GEM90 detector. Figure 3.4 shows the F2 specimen located on the top of the Ge crystal, surrounded by the NaI(Tl) annulus. The six PMTs on the top of the NaI annulus are also visible.

3.2.2 Gamma-spectrometry measurement

The gamma-activity measurement of the F2 specimen lasted $\sim$ 45 days and started about three weeks after the sample was recovered. The counted gamma-ray spectrum in normal mode (HPGe alone) is shown in Fig. 3.5, where some peaks are highlighted and associated with the related cosmogenic radionuclide (red) or naturally occurring radioisotope (black). The ^{238}U and ^{232}Th radioisotopes, together with their daughters along the decay chain in secular equilibrium (e.g., ^{212}Pb, ^{214}Bi, ^{226}Ra, ^{214}Pb, ^{228}Ac, and ^{208}Tl), are just some of the naturally occurring radionuclides which can be present in both the sample and the

Figure 3.4: GEM90 HPGe detector surrounded by the NaI(Tl) annulus. The F2 specimen is placed on top of the germanium crystal. The six PMTs on the top of the NaI annulus are also visible.

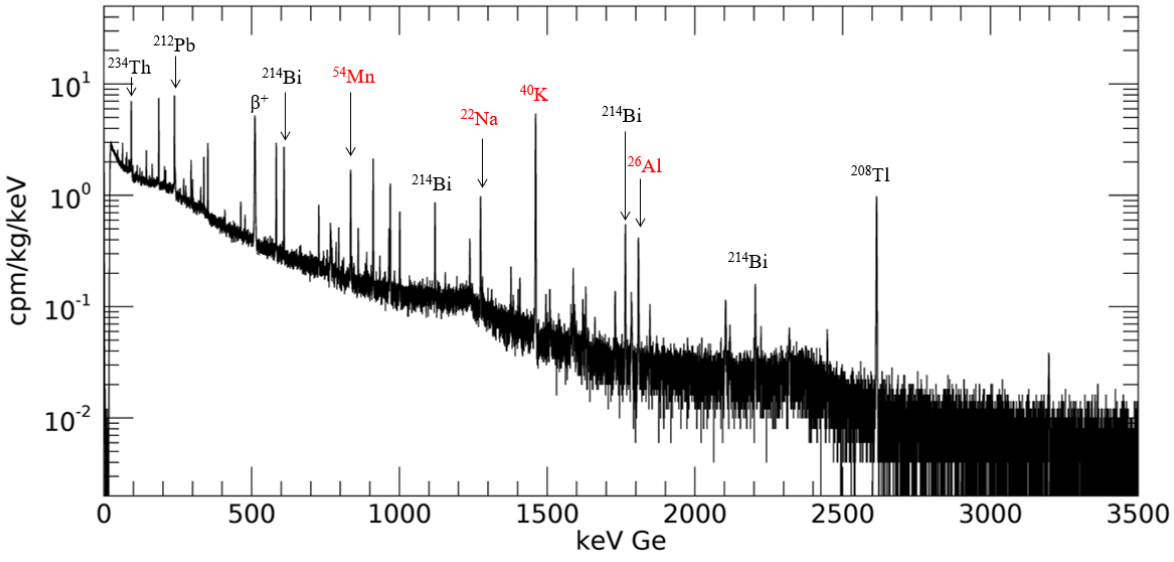

Figure 3.5: The Cavezzo meteorite gamma-ray spectrum in normal mode (HPGe alone, $\sim$ 45 days counting time). Some peaks are highlighted and associated with the related cosmogenic (red) or naturally occurring (black) radionuclide.

Nuclide	λ (s^{-1})	Main target elements
^{48}V	502.24×10^{-9}	Ti, Fe, Ni
^{51}Cr	289.58×10^{-9}	Fe, Ni
^{7}Be	150.74×10^{-9}	C, O, Mg, Al, Si, $S^{(*)}$, $P^{(*)}$, Ca, Ti, Fe, Ni
^{58}Co	113.233×10^{-9}	$Fe^{(\alpha)}$, Ni
^{56}Co	103.87×10^{-9}	Fe, Ni
^{56}Sc	95.749×10^{-9}	Ti, Fe, Ni
^{57}Co	29.515×10^{-9}	Fe, Ni
^{54}Mn	25.698×10^{-9}	Fe, Ni, Mn
^{22}Na	8.439×10^{-9}	Mg, Al, Si, Ca, Ti, Fe, Ni
^{60}Co	4.167×10^{-9}	$Co^{(n)}$, Ni
^{44}Ti	366×10^{-12}	Ti, Fe, Ni
^{26}Al	30.6×10^{-15}	Mg, Al, Si, $S^{(*)}$, $P^{(*)}$. Ca, Ti, Fe, Ni
^{40}K	17.566×10^{-18}	Ca, Ti, Fe, Ni

Table 3.1: Main target elements for the production of several cosmogenic radionuclides produced in meteorites [88]. The decay constant λ of each radioisotope is also shown. If production is not obtained through spallation, the target elements are labeled with (n) for predominance of neutron capture and (α) for reaction induced by alpha particles. The (*) symbol indicates target elements only relevant in iron meteorites or the metal phase of stony-iron.

environment. Thanks to the high selectivity and low background of the system, fifteen cosmogenic radionuclides, with half-lives down to a few days, have been measured. The most intense peak in Fig. 3.5, at the energy of $\sim$ 1461 keV, is due to the decay of ^{40}K. The main target elements of several of the measured radionuclides are listed in table 3.1. The decay constant λ of each radioisotope, expressed in s^{-1}, is also shown.

Table 3.2 lists all the cosmogenic radioisotopes detected in the F2 sample together with:

- decay modes and their probability of occurrence (in percentage);
- half-life;
- gamma-ray energy in keV;
- branching ratio (in percentage);
- measured counts per day (cpd) at the date of fall (corrected by the decay factor).

Nuclide	Decay mode	Half-life	E_γ (keV)	BR (%)	Counts per day
^{47}Ca	β^-(100%)	4.5 d	1297.09	75	137 ± 71
^{52}Mn	ϵ(68.9%) - β^+(31.1%)	5.6 d	1434.06	99.99	50 ± 42
^{48}V	ϵ(50.1%) - β^+(49.9%)	16.0 d	983.52 1312.10	99.85 97.5	51 ± 7 35 ± 8
^{51}Cr	ϵ(100%)	27.7 d	320.08	9.89	25 ± 6
^{7}Be	ϵ(100%)	53.2 d	477.60	10.44	31 ± 3
^{58}Co	ϵ(85.1%) - β^+(14.9%)	70.9 d	810.76	99.44	13 ± 2
^{56}Co	ϵ(80.4%) - β^+(19.6%)	77.2 d	846.76 1238.27	99.94 66.41	15 ± 2 7 ± 2
^{46}Sc	β^-(100%)	83.8 d	889.28 1120.55	99.98 99.97	21 ± 2 8 ± 2
^{57}Co	ϵ(100%)	271.8 d	122.06	85.49	28 ± 3
^{54}Mn	ϵ(100%)	312.2 d	834.85	99.98	195 ± 3
^{22}Na	ϵ(9.6%) - β^+(90.4%)	2.6 y	1274.54	99.94	130 ± 2
^{60}Co*	β^-(100%)	5.3 y	1173.23 1332.49	99.85 99.98	1.52 ± 0.36 1.54 ± 0.44
^{44}Ti*	ϵ(100%)	60 y	1157.02*	99.88	4.4 ± 0.8
^{26}Al	ϵ(18%) - β^+(82%)	717 ky	1129.67 1808.65	2.5 99.76	3 ± 1 64.6 ± 0.9
^{40}K	ϵ(10.7%) - β^-(89.3%)	1250 My	1460.82	10.55	327 ± 4

Table 3.2: Counting rates of the cosmogenic radionuclides measured in the Cavezzo meteorite (F2 sample), reported to the date of fall. (*) Measured with the coincidence technique. ($\star$) Gamma emitted by its short-lived daughter ^{44}Sc.

All the energy levels, transition energies, and branching ratios are retrieved from the Decay Data Evaluation Project (DDEP) through the online application Nucléide-Lara[4].

Short-lived radionuclides with half-lives up to a few days, such as ^{47}Ca, ^{52}Mn, and ^{48}V, have been detected, thus confirming with no doubt the recent fall of the meteorite and its link with the New Year's Eve fireball seen by the PRISMA all-sky cameras.

The measured counting rates of ^{54}Mn, ^{22}Na, ^{26}Al, and ^{40}K radionuclides are determined within an uncertainty of about 1%, while most of the other radionuclides were counted with Due to the short half-lives of ^{47}Ca and ^{52}Mn, their abundances had already halved 4 times when the measurement started, leading to rather high uncertainties in the measured count rates. After a few days of counting, their peaks were totally submerged by the background, and their activities were finally determined on the partial acquisition of 17 days.

Particular attention must be paid to the detection of ^{52}Mn since two other major lines originating from its decay, with a branching ratio (BR) greater than 90% and energies of 744.21 and 935.52 keV, were never detected in any partial spectrum. Although two full-energy peaks are present close to these energies, their counts can be fully explained by the activities of naturally occurring radioisotopes ^{234m}Pa and ^{214}Bi [5]. Furthermore, the ^{52}Mn line at 1434.06 keV suffers from background interference of the ^{234m}Pa gamma at 1434.16 keV (BR = 0.00973%). However, by subtracting the expected counts of the latter from the peak integral, a small positive residue, which could possibly be attributable to the decay of ^{52}Mn, was obtained. Concerning ^{47}Ca, no background peaks were expected in the neighboring channels of the spectrum, and its activity was determined with a relative uncertainty of approximately 50%.

The peaks of ^{44}Ti and ^{60}Co are not visible at all in the normal HPGe spectrum shown in Fig. 3.5, but thanks to the higher selectivity provided by the coincidence between HPGe and NaI(Tl) detectors, they have both been detected.

3.3 Coincidence spectrum

Data from the HPGe and NaI(Tl) scintillator are independently recorded by the digital acquisition chain of the detector [84, 87], allowing obtaining two independent spectra. Figure 3.6 shows the two-dimensional spectrum, that is, the number of events, as a function of their energy, recorded in the HPGe detector in coincidence with events on the NaI scintillator. This is obtained by arranging the time-ordered list of events in a 2D counts histogram, with Ge channels as x-entries and NaI channels as y-entries.

For instance, the sharp vertical lines at $\sim$ 1275 and 1809 keV represent the major gamma-rays of the ^{22}Na and ^{26}Al decays detected on the HPGe in coincidence with the annihilation radiation detected on the NaI scintillator. Two regions of higher counts in correspondence of 511 and 1022 keV are visible, which are highlighted by the white circles in Fig. 3.6 for ^{26}Al. Oblique lines originate from partial energy depositions on both detectors. The most intense one is due to the major line of ^{40}K and starts from an energy of $\sim$ 1461 keV on both axes.

[4] http://www.lnhb.fr/nuclear-data/module-lara

[5] Line of ^{234m}Pa at 742.81 keV (BR = 0.094%) and ^{214}Bi at 934.06 keV (BR = 3.1%)

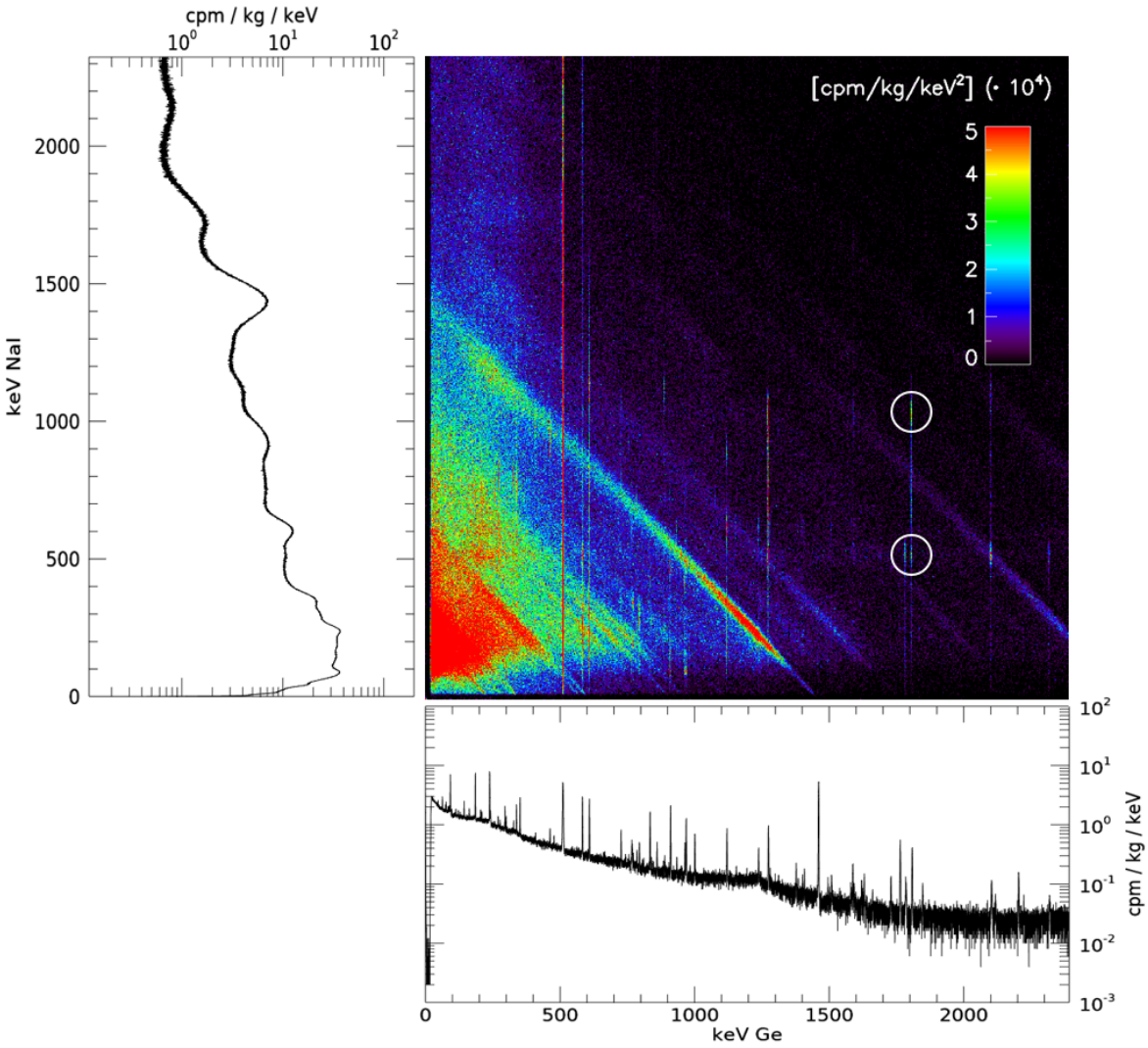

Figure 3.6: Surface representation of the two-dimensional Ge-NaI spectrum, with total Ge and NaI spectra plotted on the corresponding axes. White circles highlight the two regions of higher counts due to the coincidence between the annihilation gamma-rays and the 1809 keV line of ^{26}Al.

Given the two-dimensional coincidence spectrum, the normal one, shown in Fig. 3.5, can be obtained by integrating the two-dimensional spectrum over the whole NaI energy range, i.e., by integrating over the full range of y-entries. On the other hand, coincidence spectra are computed by integrating over limited energy ranges.

The coincidence between HPGe and NaI detectors, providing a substantial reduction of the background level, allows performing event selection and finally measuring ^{44}Ti and ^{60}Co activities.

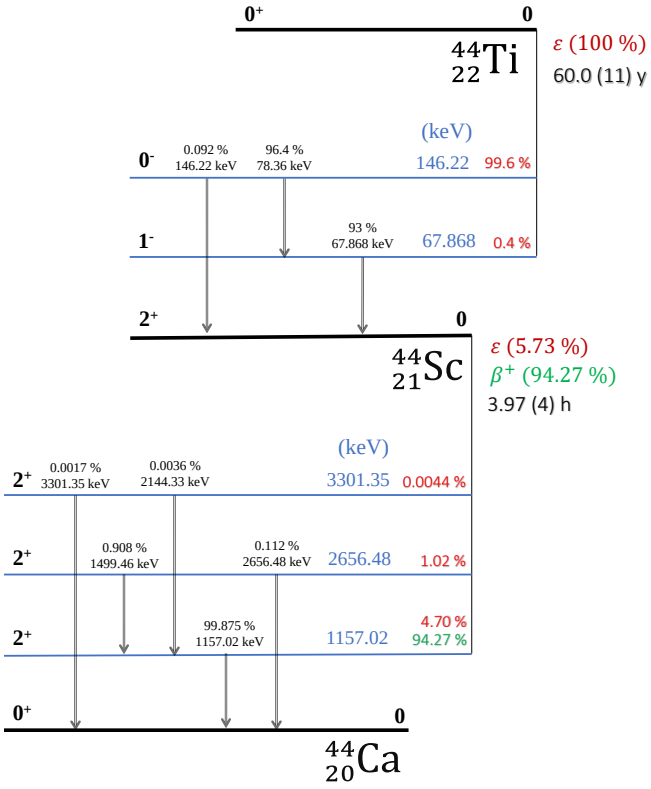

Figure 3.7: ^{44}Ti decay scheme.

3.3.1 Coincidence optimization for ^{44}Ti detection

As visible from the decay scheme in Fig. 3.7, ^{44}Ti decays by electron capture (100%) into ^{44}Sc and has a half-life of $T_{1/2}$= (60.0 ± 1.1) y. Two gamma-rays of ∼ 78 keV (96.4%) and ∼ 68 keV (93%) or a single gamma-ray of ∼ 146 keV energy (0.092%) are emitted in correspondence with ^{44}Ti relaxation to the ground state. These low-energy peaks cannot be accurately revealed since they lie in a spectral region with a high background level. Therefore, the measurement of ^{44}Ti activity can rely on the decay of its short-lived daughter ^{44}Sc ($T_{1/2} \simeq 4$ h), which is in secular equilibrium with its parent, and that turns into ^{44}Ca by β^+ decay (94.27%) emitting a photon of energy 1157.02 keV (99.88%) (see Fig. 3.7).

Unfortunately, the gamma-ray line of ^{44}Sc at 1157.02 keV suffers from the interference of an intense peak at 1155.19 keV originated by the β^- decay of the naturally occurring ^{214}Bi radioisotope, which partial decay scheme is shown in Fig. 3.8.

Figure 3.9a shows the normal Ge spectrum in the ^{44}Sc–^{214}Bi peak region, where it is clear that the ^{214}Bi peak (red curve) totally submerges the signal from ^{44}Sc. Therefore, a confident estimation of ^{44}Ti activity is not achievable on this spectrum, and to overcome this issue, the difference between the decay of ^{44}Sc (β^+ decay) and ^{214}Bi (β^- decay) has been exploited. The annihilation between a positron emitted by the β^+ decay of ^{44}Sc and one electron within the sample produces two collinear photons of 511 keV energy simultaneously to the emission of the 1157.02 keV gamma-ray. The coincidence between the 1157 keV gamma-ray revealed in the Ge detector and the 511 and 1022 keV annihilation photons detected by NaI(Tl) scintillator can be used as an estimate of ^{44}Ti activity [82].

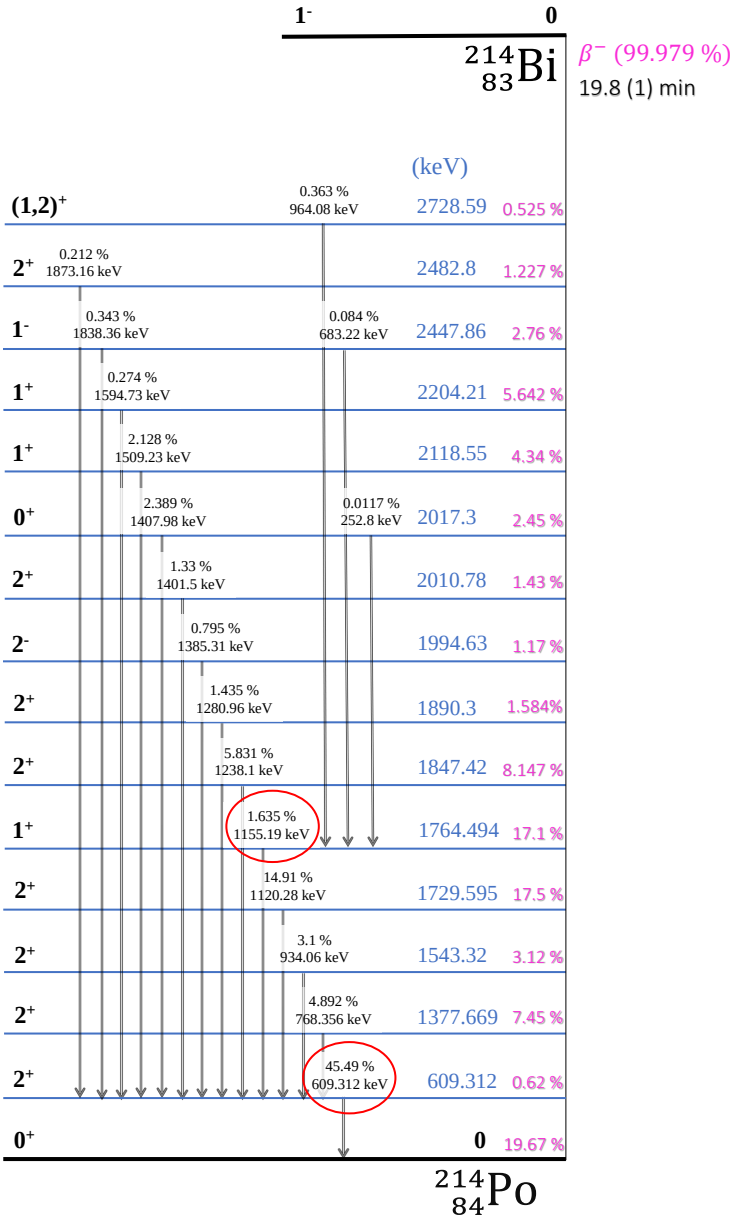

Figure 3.8: ^{214}Bi partial decay scheme.

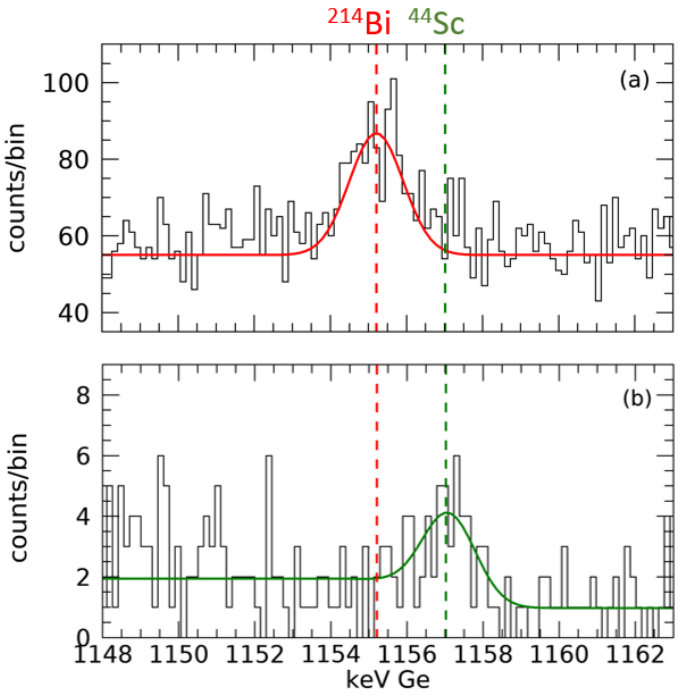

Figure 3.9: Ge total spectrum (panel a) and corresponding Ge coincidence spectrum in the energy range of 1148–1163 keV obtained using both 511 and 1022 keV windows (panel b). Red and green dashed lines represent ^{214}Bi peak at 1155.19 keV and ^{44}Sc peak at 1157.02 keV, respectively. The Levemberg-Marquardt parametric fit for the ^{214}Bi peak in the normal spectrum (red curve) and for the ^{44}Sc peak in the coincidence spectrum (green curve) are also shown.

The optimal energy windows for NaI(Tl) have to be determined to obtain the coincidence spectrum. Windows as wide as several tens of keV are needed due to the poor resolution of the NaI scintillator. Furthermore, the choice of the 511 keV window is particularly tricky due to the 609.3 keV (45.49%) gamma-ray emitted by ^{214}Bi (simultaneously to the 1155.19 keV gamma-ray). A wrong choice of the width of this window would involve the partial inclusion of the peak at 609.3 keV, leading to an incorrect estimate of the ^{44}Ti activity [84, 89].

The energy levels and emitted gammas shown in the partial decay scheme of Fig. 3.8 are the ones most linked to the emission of the interfering peaks of ^{214}Bi (highlighted by the red circles).

The approach introduced in Ref.[90] has been applied to optimize the choice of the coincidence windows for the detection of ^{44}Ti. The whole ensemble S of all possible coincidence windows, parametrized through their centre c and their half-width w, has been considered for both 511 and 1022 keV, and the coincidence Ge spectrum for each (c,w) has been extracted. For each spectrum, the ^{44}Ti and ^{214}Bi counting rates have been computed through a double peak-fit [84, 91], and the final coincidence window was chosen according to the following figure of merit:

$$M = \mathrm{cpm}_{^{44}\mathrm{Ti}} \cdot (1 - \mathrm{cpm}_{^{214}\mathrm{Bi}}) \cdot \mathrm{SNR}_{^{44}\mathrm{Ti}} \tag{3.1}$$

where $\mathrm{SNR}_{^{44}\mathrm{Ti}}$ is the ^{44}Ti signal-to-noise ratio, defined as:

$$\mathrm{SNR}_{^{44}\mathrm{Ti}} = \frac{\mathrm{cpm}_{^{44}\mathrm{Ti}}}{\sigma_{\mathrm{cpm}_{^{44}\mathrm{Ti}}}} \tag{3.2}$$

The final coincidence window (c_0, w_0) is the one that maximizes M, i.e., that simultaneously minimizes the ^{214}Bi interference and maximizes the ^{44}Ti signal-to-noise ratio:

$$(c_0, w_0) = \underset{(c,w)\in S}{argmax}[M(c,w)] \tag{3.3}$$

This optimization algorithm was performed independently for both 511 and 1022 keV windows. Figure 3.9b shows the sum of the coincidence spectra obtained in these two cases, where the background level, and consequently the ^{214}Bi signal, has been reduced by a factor of more than 20 with respect to the normal spectrum (Fig. 3.9a). The peak in the coincidence spectrum (green curve in Fig. 3.9b) can thus be completely attributed to the ^{44}Ti activity.

In order to correct for the count loss due to the application of the coincidence windows, the measured ^{44}Ti activity has been normalized to the ratio between the peak counts in the coincidence (C) and normal (N) spectra of the ^{26}Al peak at 1808.65 keV [82].

Finally, the estimated count rate of ^{44}Ti was found to be (4.4 ± 0.8) cpd.

3.3.2 Coincidence optimization for ^{60}Co detection

A similar approach as the one described in Section 3.3.1 was applied to measure the faint activity of ^{60}Co. As visible in Fig. 3.10, this radionuclide decays through the β^- channel by emitting two gamma-rays at 1173.23 keV (γ_1) and 1332.49 keV (γ_2) from the de-excitation of the daughter ^{60}Ni. The coincidence spectra were built in two cases: the γ_1 peak detected on the HPGe in coincidence with the γ_2 peak detected on the NaI(Tl) scintillator and vice versa.

Figure 3.11 shows the results of the coincidence window optimization for ^{60}Co detection. In this case, the coincidence windows optimized for the detection of γ_1 on the HPGe detector in coincidence with γ_2 on the NaI scintillator, and vice versa, were chosen according to the following figure of merit:

$$Q = \mathrm{cpm} \cdot \mathrm{SNR} \tag{3.4}$$

$$(c_0, w_0) = \underset{(c,w)\in S}{argmax}[Q(c,w)] \tag{3.5}$$

In Figs. 3.11a and 3.11c is visible that the peaks related to γ_1 and γ_2 are not present in the Ge total spectrum. Conversely, in the coincidence spectra shown in Figs. 3.11b and 3.11c, both peaks are revealed thanks to the significant reduction of the background level that allows measuring a counting rate of (1.52 ± 0.36) cpd for the γ_1 peak and (1.54 ± 0.44) cpd for the γ_2 line. For the case of ^{60}Co, both peaks' counts have been normalized to their C/N ratio estimated from a coincidence analysis performed on the measurement

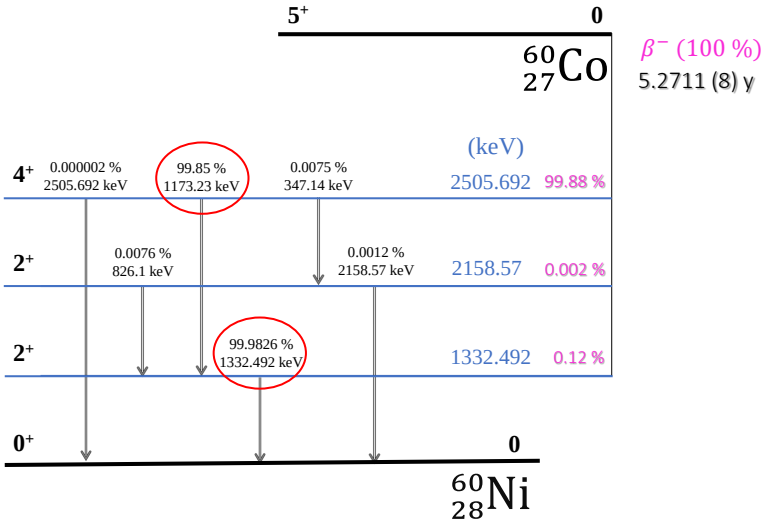

Figure 3.10: ^{60}Co decay scheme.

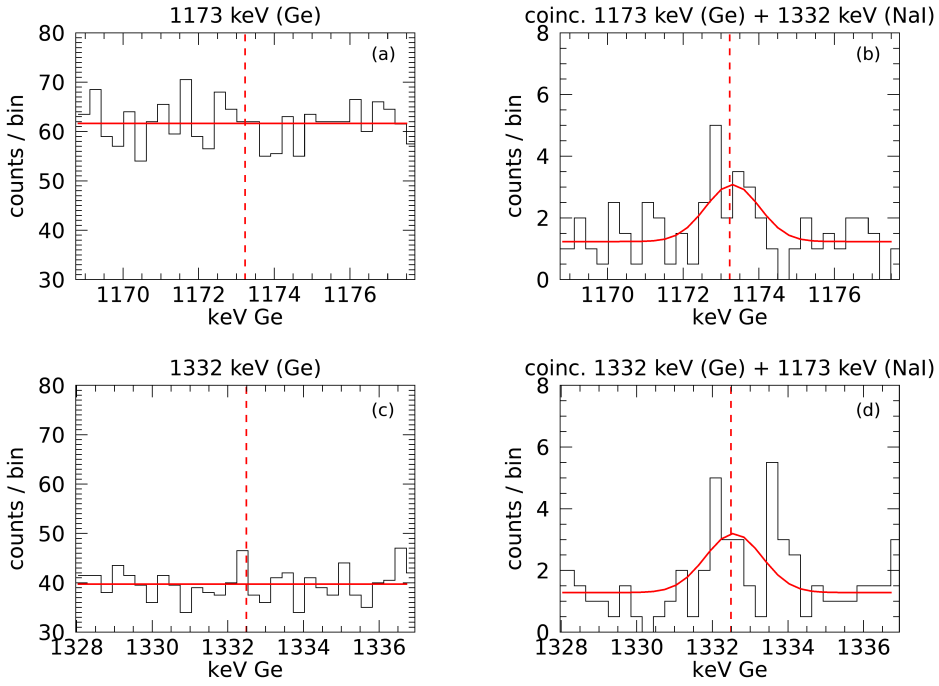

Figure 3.11: Coincidence window optimization for ^{60}Co detection. (a) Ge normal spectra in the energy ranges centered at 1173.23 keV (panel a) and 1332.49 keV (panel c). Panel b shows the coincidence spectrum between the detection of the 1173.23 keV peak on the Ge crystal and the 1332.49 keV peak on the NaI scintillator. Vice versa for panel d. A fit of the peaks is also shown (red curves). Red dashed lines are in correspondence with the 1173.23 and 1332.49 keV peaks of ^{60}Co. A two-channel binning is performed in this case over both spectra.

of the activity of a ^{60}Co standard, which activity was known (see Section 3.4.1). Since the two peaks have very similar BRs (see Table 3.2), and the detector efficiency does not vary significantly over this energy range, similar counts were expected. Therefore, their measured counts are consistent, and this is a countercheck of the accuracy of the coincidence optimization method and count loss normalization adopted.

3.4 Cosmogenic radionuclides activity

To deduce the activity of the cosmogenic radionuclides detected in the Cavezzo meteorite, it is necessary to estimate the detection efficiency η_d as a function of the peak energy E_γ, defined as the ratio between the number of counts (photons) observed by the detector and the number of photons emitted by a source. Given the measured counts, the activity of each radioisotope, e.g., the number of decays per minute (dpm), can thus be estimated as follows:

$$\mathrm{dpm} = \frac{\mathrm{cpm}}{\eta_d} \tag{3.6}$$

The measured counts depend on the activity of the source, the source-to-detector distance, the materials between the emitting source and the detector, and the detector geometry itself. Except for low-energy photons, an incident gamma-ray photon can completely pass through the detector without interacting, that is, before undergoing a significant interaction and losing its energy. Furthermore, many of the gamma-rays may deposit only a fraction of their energy in the detector. The efficiency is determined only by the photons interacting one or more times in the detector and depositing all their energy. Therefore, the efficiency will be substantially less than 100%.

Another effect influencing the detection efficiency is the dead layer (DL) thickness of the HPGe detector, corresponding to the undepleted detector region, close to the outer surface, doped with lithium atoms to get a semiconductor junction. The DL thickness of GEM90 has a nominal value of 0.7 mm but, as a consequence of years of usage and thermal cycles of the apparatus, it can increase due to the lithium atoms continuously diffusing inside the germanium crystal, enlarging the insensitive volume of the latter. Therefore, it is not reliable to use the dead layer thickness provided by the manufacturer and its actual value should be estimated.

In the following sections, the DL thickness estimation and the simulations carried out to determine the detection efficiency values for each of the radionuclides detected, as well as the activity of the latter, are described in detail.

Standard	Energy (keV)	cpm
^{26}Al	1129.67	0.40 ± 0.01
	1808.65	8.86 ± 0.02
	1808.65 - 511*	0.26 ± 0.01
	2938.32	0.051 ± 0.002
^{40}K	1460.82	74.05 ± 0.07
	1460.82 - 511*	0.59 ± 0.03
^{60}Co	1173.23	1314 ± 9
	1332.49	1207 ± 13
	2505.69	66 ± 2

Table 3.3: Results of the measurements of the ^{26}Al, ^{40}K, and ^{60}Co standards. The counts per minute (cpm) of the considered peaks are shown with their uncertainties. The result on single-escape peaks of the main lines of ^{26}Al and ^{40}K (*) are also reported.

3.4.1 Dead layer thickness estimation

As already discussed, to estimate the detection efficiency as a function of the energy, the actual DL thickness of the HPGe detector must be determined. Since the actual thickness of the DL cannot be measured directly, its value has been deduced by comparing the measurements of standards (whose activities are known) with the results of simulations.

Table 3.3 summarizes the results of the measurements of three standards: ^{26}Al, ^{40}K, and ^{60}Co. Since the activity of each sample is known, these measurements have been simulated by increasing the DL thickness from its nominal value (0.7 mm) up to 10 mm, with a step of 0.3 mm. For this purpose, the Monte Carlo simulation toolkit Geant4[6] [92–94] was used. It was developed by the Geant collaboration at CERN and is currently used in several research areas, including high-energy, nuclear and accelerator physics, as well as medical and space science studies. It provides an extensive library of virtual classes to implement for the specific case study, allowing the simulation of the passage of particles through matter.

The geometry and composition of both the apparatus and the sample must be defined. Geant4 already provides a comprehensive library of materials, and specific mixtures of elements, and compounds can also be defined by providing the desired density and the mass/mole fractions.

The left panel of Fig. 3.12 illustrates a schematic representation of the ^{26}Al (panel a), ^{40}K (panel b), and ^{60}Co (panel c) standards, with the darker blue or grey areas representing the active regions. On the right, the snapshots of the Geant4 simulations of the ^{26}Al (panel a), ^{40}K (panel b), and ^{60}Co (panel c) standards activity are shown, with the samples placed over the end cap window of the HPGe detector.

The results of a such procedure are shown in Figs. 3.13, 3.14, and 3.15, for ^{26}Al, ^{40}K, and ^{60}Co, respectively. For each peak, the DL thickness is estimated as the intersection of the simulated values (red curve) with the experimental results (black horizontal line). The red lines plot the 2^{nd}-degree polynomial fit of the difference between measured and simulated counts as a function of the DL thickness, while the red bands mark the 68% confidence intervals of each fit. The vertical dashed black lines indicate the intersection intervals with the experimental result. The DL range shown for ^{60}Co in Fig. 3.15 is shorter compared to the other two standards simply because, for all the peaks, the simulated cpm

[6] Geant4 version 10.2 patch-2 (Geant4.web.cern.ch)

had already reached a value lower than the measured one for a DL value smaller than 5 mm.

Figure 3.16 shows the final estimate of the GEM90 DL thickness (red horizontal line), equal to (3.29 ± 0.06) mm, obtained as the weighted mean of the DL values estimated with the simulation performed over the ^{26}Al (blue), ^{40}K (orange), and ^{60}Co (green) standards.

It should be emphasized that the DL thickness obtained does not correspond exactly to the actual thickness of the outer contact layer of the Ge crystal, but has to be addressed as an equivalent dead layer thickness, also taking into account possible inaccuracies in the modeling of the HPGe geometry and allowing a fine-tuning of the simulation. For example, the size of the tip and bullet radius of the Ge crystal hole are not precisely known (only their nominal values are given), thus affecting the determination of the active volume. Furthermore, the simulation only takes into account the energy deposition within the active volume, while the charge collection process is not modeled.

To validate the estimated DL, the measurement of the Dhajala meteorite, an H3/4 chondrite fallen in 1976 in India [95] has been simulated. This step also verified that the self-absorption of the sample, which is directly proportional to the sample density of the emitting source and mainly affects low-energy gamma-rays, was correctly taken into account in the simulations.

This meteorite specimen was counted at the underground Laboratory of Monte dei Cappuccini in Torino, [82, 83]. It has a volume of 245 cm^3, a mass of 706 g (ρ = 2.88 g cm^{-3}), and a roughly squared form, which was modeled as a parallelepiped of sides 7.0/6.5/5.4 cm. Its composition was obtained from Ref.[96] and normalized to 100%, neglecting elements under 1% of mass fraction.

With the Geant4 toolkit, the Dhajala spectrum was simulated 100 times. Figure 3.17a shows a snapshot of the activity simulation of Dhajala, with the meteorite placed above the end cap window of the HPGe detector. Table 3.4 shows the comparison between measured and simulated counts of the ^{26}Al peaks (including full-energy, single escape, and sum ones) with their uncertainties. The simulated counts and uncertainties are obtained as the mean and standard deviation of the 100 simulated peak counts, respectively. An overall agreement is observed, except for the 1808.65 keV peak where, however, the observed discrepancy is small (3%).

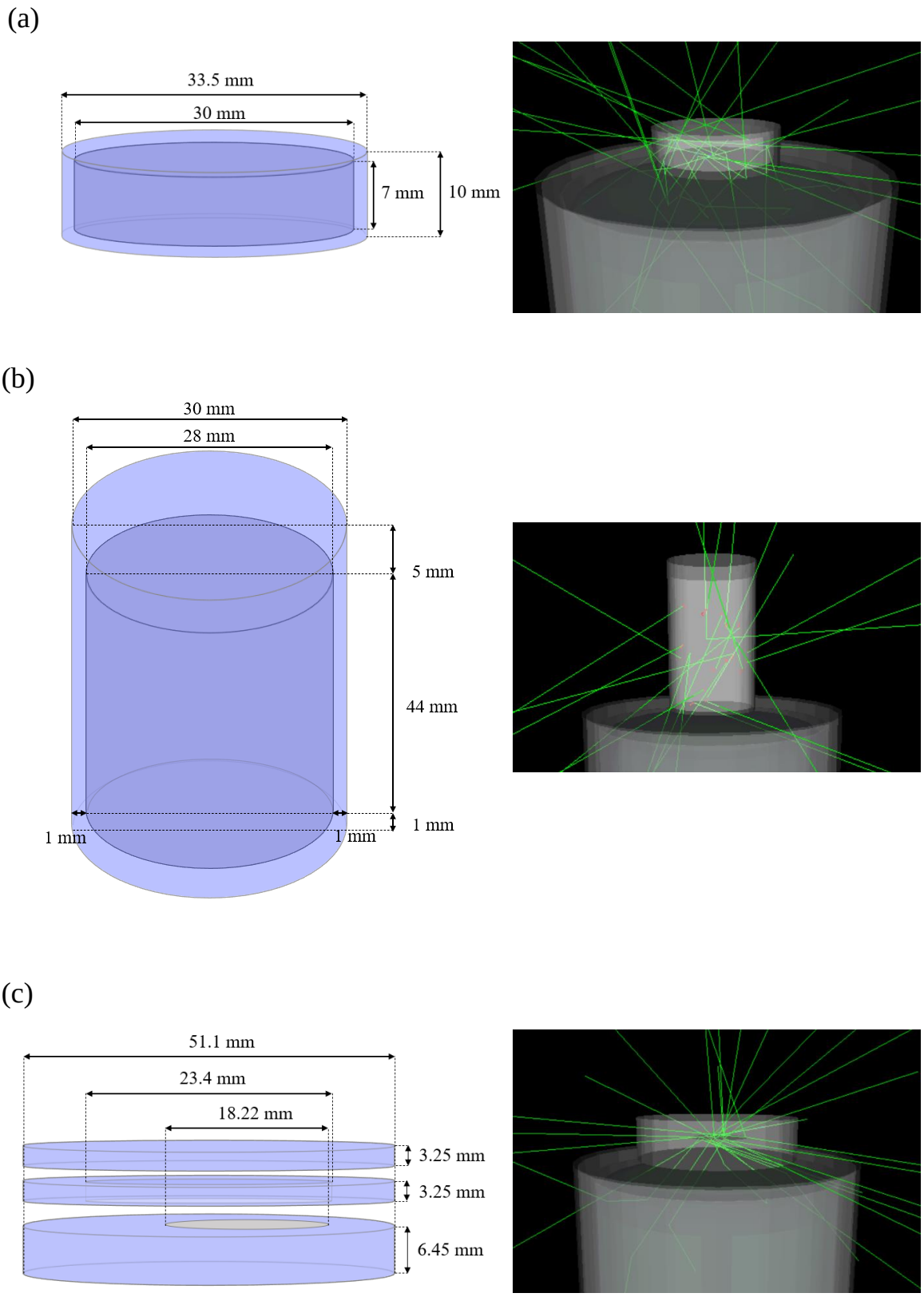

Figure 3.12: (Left column) Schematic representation of the ^{26}Al (panel a), ^{40}K (panel b), and ^{60}Co (panel c) standards, with the darker blue or grey areas representing the active regions. (Right column) Snapshots of the Geant4 simulations of the ^{26}Al (panel a), ^{40}K (panel b), and ^{60}Co (panel c) standards activity. The samples are placed over the end cap window of the HPGe detector.

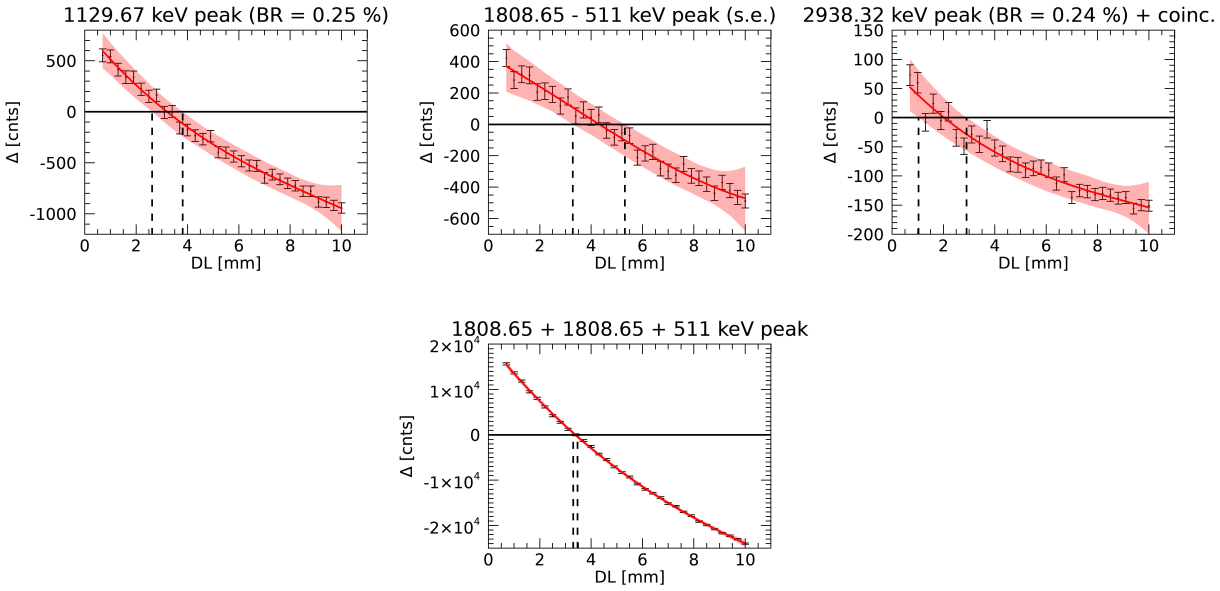

Figure 3.13: Difference (Δ) between ^{26}Al standard simulated and measured counts as a function of the DL thickness of the Ge crystal. The red lines plot the 2$^{\text{nd}}$-degree polynomial fit and the red bands mark the 68% confidence intervals of the fit. The vertical dashed black lines indicate the intersection intervals with $\Delta = 0$, represented by the horizontal black line.

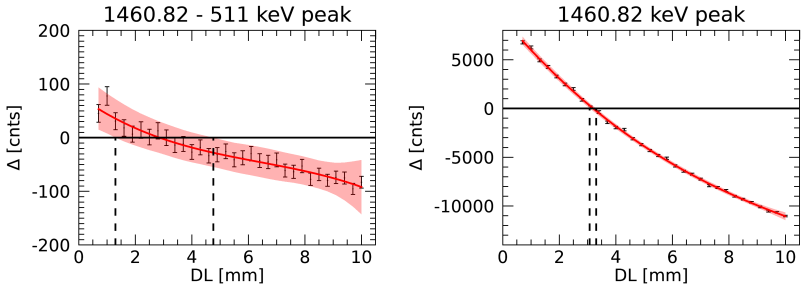

Figure 3.14: Difference (Δ) between ^{40}K standard simulated and measured counts as a function of the DL thickness of the Ge crystal. The red lines plot the 2$^{\text{nd}}$-degree polynomial fit and the red bands mark the 68% confidence intervals of the fit. The vertical dashed black lines indicate the intersection intervals with $\Delta = 0$, represented by the horizontal black line.

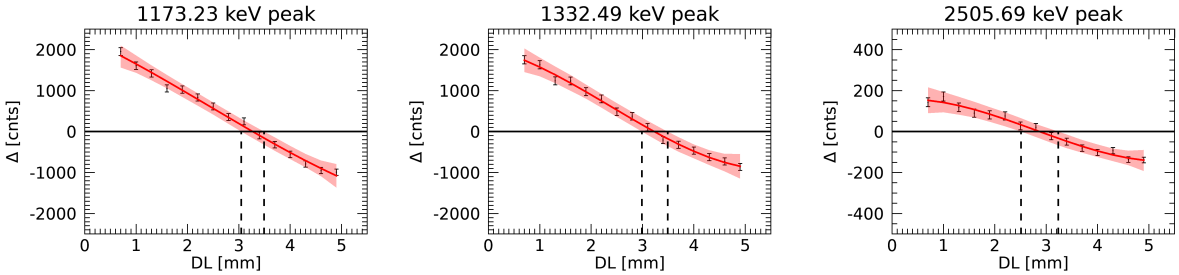

Figure 3.15: Difference (Δ) between ^{60}Co standard simulated and measured counts as a function of the DL thickness of the Ge crystal. The red lines plot the 2$^{\text{nd}}$-degree polynomial fit and the red bands mark the 68% confidence intervals of the fit. The vertical dashed black lines indicate the intersection intervals with $\Delta = 0$, represented by the horizontal black line. The DL range is limited up to 5 mm, since, in all three cases, the polynomial fit has already reached a value lower than $\Delta = 0$.

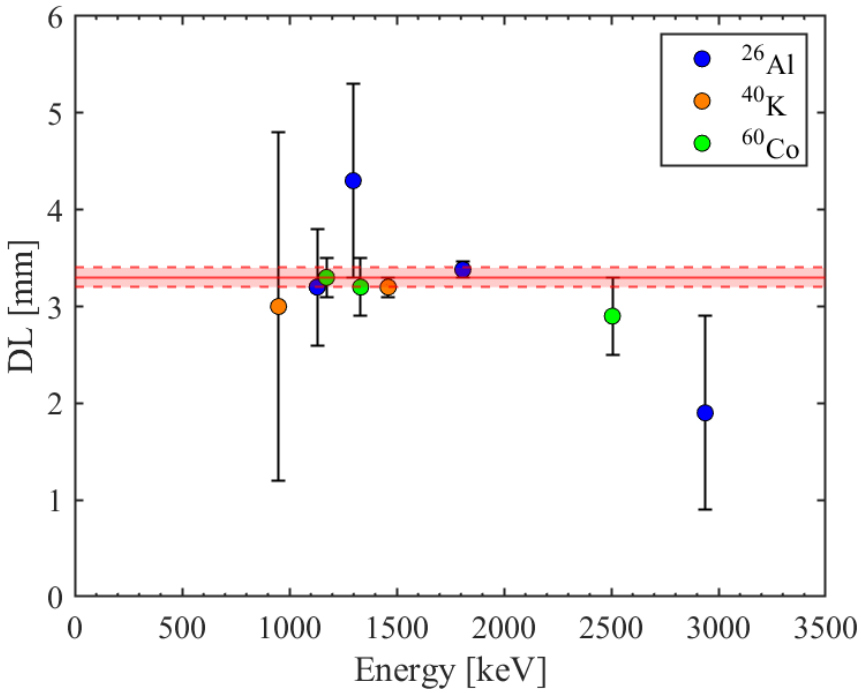

Figure 3.16: Dead layer estimates with associated uncertainties for ^{26}Al, ^{40}K, and ^{60}Co standards in blue, orange, and green, respectively. The red horizontal line is the weighted mean value and the shaded red bar represents its uncertainty.

Energy	measured cpd	simulated cpd
1129.67	15.1 ± 0.7	16.0 ± 0.8
1808.65	434.3 ± 0.6	450 ± 4
1808.65 - 511	11.5 ± 0.6	13.2 ± 0.8
2938.32	1.6 ± 0.1	1.3 ± 0.2

Table 3.4: Comparison between measured and simulated cpd of ^{26}Al peaks with their uncertainties for the Dhajala meteorite.

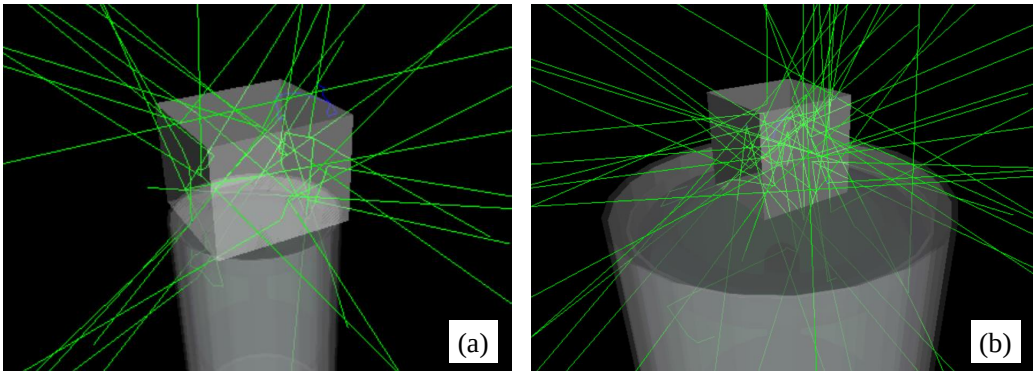

Figure 3.17: Snapshot of the Geant4 simulation of Dhajala (panel a) and Cavezzo (panel b) activity. Both meteorites are placed over the end cap window of the HPGe detector.

3.4.2 Detection efficiency simulations and radioisotopes activity

To deduce the activity of each radionuclide detected in the Cavezzo meteorite, it is necessary to estimate the detection efficiency, which was obtained by simulating the measurement of the Cavezzo meteorite with the Geant4 toolkit. The meteorite has a roughly cubic form, which has been modeled as a cube of 2.5 cm sides with a bulk density of 3.322 g cm^{-3} [80]. Its bulk composition was measured at the Earth Sciences Department and Natural History Museum of the Firenze University, revealing the presence of fourteen elements. Since the density of the Cavezzo meteorite is higher than the Dhajala one, a higher self-absorption is expected for Cavezzo.

Figure 3.17b shows a snapshot of the activity simulation of Cavezzo, with the meteorite placed above the end cap window of the HPGe detector.

One million decays have been simulated 100 times for each of the revealed radioisotopes. Each time, the DL thickness was defined as $(3.29 + x \cdot 0.06)$ mm, where x is a pseudo-random value from a gaussian distribution of null mean and standard deviation equal to one. The net counts have been calculated as the mean of the 100 runs. Finally, the detection efficiency η_d has been estimated following equation 3.6. Examples of such a simulation are shown in Fig. 3.18, where the blue histograms, fitted with gaussian distributions (black curves), represent the statistics of the peak's integral counts. The mean net counts and uncertainties are represented by the vertical red lines and the shaded red bars, respectively, corresponding to $\eta_d = (4.30 \pm 0.02)$ % for ^{57}Co, $\eta_d = (3.19 \pm 0.03)$ % for ^{58}Co, $\eta_d = (1.83 \pm 0.01)$ % for ^{22}Na, and $\eta_d = (1.53 \pm 0.02)$ % for ^{26}Al. The efficiency decreases as the peak energy increases, as expected. High-energy incident gamma-ray photons can pass through the whole detector without interacting. Furthermore, conversely to low-energy photons, they may deposit only a fraction of their energy in the detector, not contributing to the peak efficiency, which incorporates only those gamma-ray photons that deposit all their energy by interacting one or more times in the detector. The low values obtained for the detection efficiency also take into account that not all the gamma-rays emitted by the source are directed toward the detector.

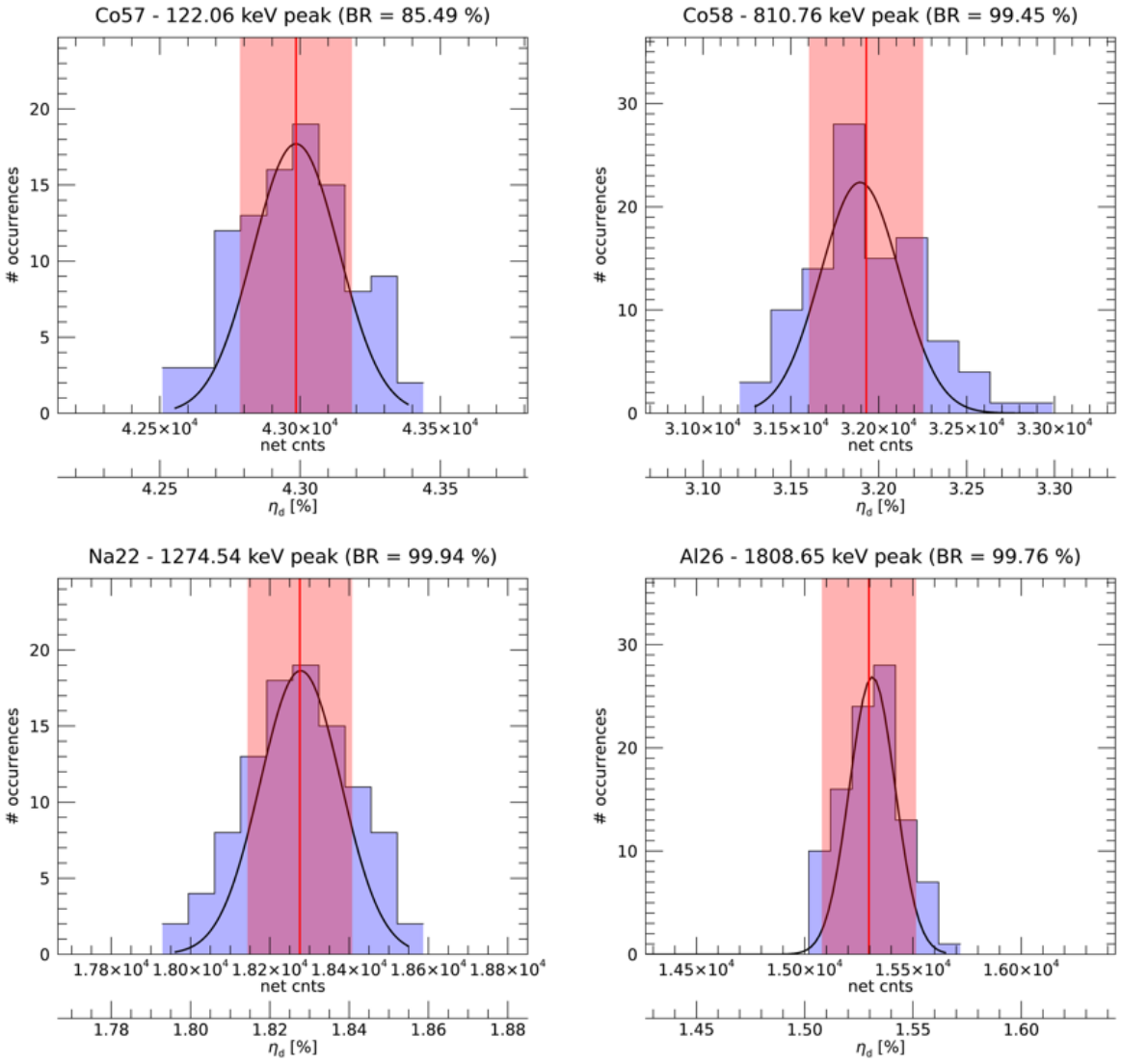

Figure 3.18: Results of the detection efficiency simulation of the ^{57}Co line at 122.06 keV, the ^{58}Co line at 810.76 keV, the ^{22}Na line at 1274.54 keV, and the ^{26}Al line at 1808.65 keV. The blue histograms show the distribution of the efficiency obtained by running 100 times the decay simulation. The mean values are shown by the vertical red line and the shaded red bars represent the associated uncertainties.

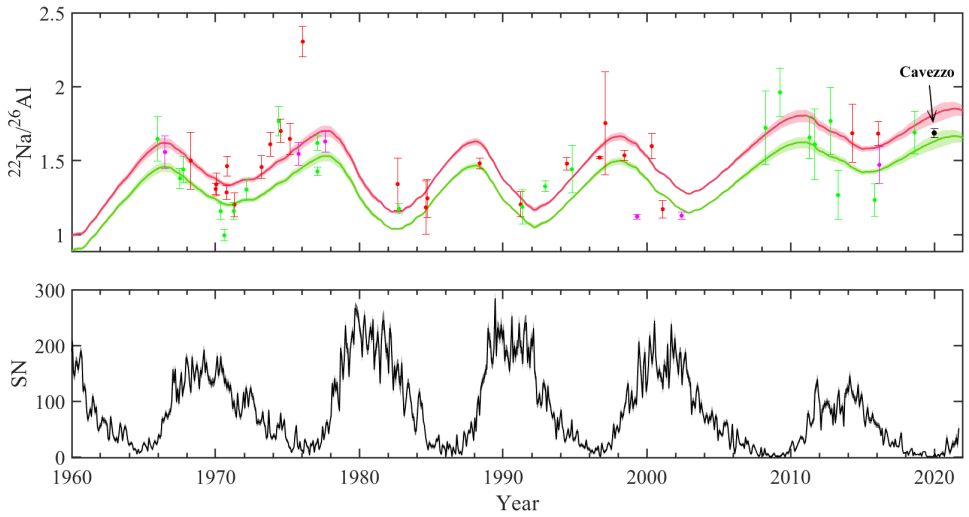

Figure 3.19: Estimated values for the ^{22}Na/^{26}Al activity ratio of H (red curve) and L (green curve) chondrites, with uncertainties represented by the shaded colored areas, and measured values for H (red), L (green), and LL (magenta) chondrites felt within the last five solar cycles [97–126]. The uncertainty associated with each measured value is represented by the vertical bars. The black dot shows the ^{22}Na/^{26}Al activity ratio calculated for the Cavezzo meteorite.

3.5 The ^{22}Na/^{26}Al ratio as a proxy of decadal GCR variations

Due to the half-live of ^{22}Na of about 2.6 y, the activity of such radionuclide reflects the decadal GCR flux variation due to solar activity. In stony meteorites, ^{22}Na and ^{26}Al are produced by the same target elements (see Table 3.1) through spallation. Therefore, the ratio between these GCR-induced activities is independent of the size of the meteoroid and the depth. Moreover, being the ^{26}Al activity insensitive to decadal variations of the GCR flux, due to its half-life of 717 ky, the ^{22}Na/^{26}Al activity ratio can be used as a direct indicator of the latter. By analysing this ratio in stony meteorites felt during the last five Solar Cycles (from 19 to 24), it has been found that GCR intensity decreased by nearly 20% from solar minima to solar maxima.

The expected values for the ^{22}Na/^{26}Al activity ratio can be calculated through the SN series by applying the opposite oscillation of the SN to the mean value of the ratio. Given the *reversed* SN series:

$$SN_{rev}[i] = max(SN) - SN[i] \tag{3.7}$$

where i ranges from 1 to the length of the SN series, the expected ^{22}Na/^{26}Al activity ratio can be estimated as follows:

$$\left(\frac{^{22}Na}{^{26}Al}\right)[n] = \overline{\left(\frac{^{22}Na}{^{26}Al}\right)} \cdot \frac{S[n]}{\overline{S}} \tag{3.8}$$

where:

$$S[n] = \sum_{i=1}^{n} SN_{rev}[i] e^{-\lambda(t_n - t_i)} \tag{3.9}$$

is a support series, which depends on the decay constant λ of ^{22}Na, $\overline{S}$ is the mean value of the latter, and their ratio is multiplied for the mean ^{22}Na/^{26}Al activity ratio expected for the chondrite of interest.

Figure 3.19 shows the ^{22}Na/^{26}Al activity ratio expected for H (red curve) and L (green curve) chondrites, together with measured values of several H (red), L (green), and LL (magenta) chondrites, with relative uncertainties [97–126]. In the bottom panel, the SN [40] time series is shown, indicating a phase difference of about two years between the sunspot minimum and the ^{22}Na/^{26}Al ratio peak. This delay is due to the fact that the activity of a radionuclide is influenced by the GCR flux of the previous 2-3 half-lives.

The high ^{22}Na/^{26}Al ratio measured in the Cavezzo meteorite, equal to 1.69 $\pm$ 0.03 (black dot in Fig. 3.19) reflects the low activity of the Sun around 2018-2020 and is consistent with the expected value for L chondrites.

3.6 Conclusions

Cavezzo is the first meteorite recovered by the Italian PRISMA all-sky camera network, felt on January 1st, 2020, in Northern Italy and classified as an L5 anomalous chondrite. The high efficiency and selectivity of the gamma-ray spectrometer used to measure the gamma activity of the Cavezzo meteorite made it possible to reveal also numerous cosmogenic radionuclides with half-lives down to a few days, thus confirming the recent fall of the sample.

The faint activities of ^{44}Ti and ^{60}Co have been measured thanks to the coincidence optimization between the HPGe and the NaI(Tl) detectors. The optimization of the coincidence window made it possible to almost completely remove the interfering signal due to the naturally occurring ^{214}Bi, thus enabling the estimation of a ^{44}Ti counting rate of (4.4 $\pm$ 0.8) cpd. With a similar approach, the background level has been reduced, thus allowing the detection of the 1173 and 1332 keV peaks of ^{60}Co, which normalized counts turned out to be equal to (1.52 $\pm$ 0.36) cpd and (1.54 $\pm$ 0.44) cpd, respectively. In general, this optimization technique can be used whenever the coincidence between any pair of events is useful to detect the presence of a given radionuclide.

The detection efficiency as a function of the energy was estimated using Monte Carlo simulations performed with the Geant4 toolkit. The equivalent DL thickness was estimated by simulating the activity of three different standards. The comparison between measured and simulated counts, obtained by varying the DL thickness value from 0.7 mm (nominal value) up to 10 mm, made it possible to obtain an estimate of this parameter equal to (3.29 $\pm$ 0.06) mm.

The measurement of the Dhajala meteorite (H3/4 chondrite fell in 1976 in India) was simulated to validate the estimated DL value and to verify the correct modeling of the sample's self-absorption. The overall agreement between the measured and simulated counts for the main peaks of ^{26}Al provided validation of the simulation.

The detection efficiency as a function of the energy has been evaluated by performing Monte Carlo simulations of the Cavezzo measurement with the Geant4 toolkit.

Finally, the ^{22}Na/^{26}Al ratio has been estimated and compared to the expected value. The high value, equal to 1.69 $\pm$ 0.03, reflects the low activity of the Sun during the beginning of 2020 and is consistent with the expected value for L chondrites.

Chapter 4

Cosmogenic radioisotopes in Bennu asteroid (OSIRIS-REx mission)

Asteroids are geologic remnants from the early Solar System that can help answer the most relevant question in planetary science: which are the sources of Earth's organic materials and water, which may have influenced the origin and evolution of life? Understanding whether these materials were already present on Earth during its early planetary formation or whether they were delivered there at a later stage is also a matter of great interest.

Answering these questions is the main purpose of NASA's Origins, Spectral Interpretation, Resource Identification, and Security-Regolith Explorer (OSIRIS-REx) mission. The spacecraft was launched on September 2016 and reached its target asteroid (101955) Bennu two years later. In October 2020, it collected several grams of surface material, which will be returned to Earth in late 2023. Bennu was discovered by the Lincoln Near-Earth Asteroid Research (LINEAR) during a near-Earth asteroids (NEA) survey on September 11^{th}, 1999. It is a B-type asteroid, a carbonaceous asteroid characterized by a blue spectral appearance, probably formed at least 700 million years ago from a much larger, carbon-rich asteroid in the Main Asteroid Belt. It is also classified as a near-Earth object (NEO), specifically an Apollo asteroid, and considered a potentially hazardous object (PHO), as it makes close approaches to Earth every six years. The carbonaceous asteroids are considered to be minimally altered since the formation of the Solar System and to preserve material from the early Solar System rich in volatile compounds and organic molecules. The availability of carbonaceous material, representing the basis of organic molecules necessary for life, was a determining factor that led to the choice of the asteroid Bennu as the target of the OSIRIS-REx mission.

The Monte dei Cappuccini Underground Laboratory in Torino has been chosen by NASA to carry out the radioactivity measurement of the sample, with the aim of detecting the cosmogenic radionuclides present in the extraterrestrial material and determining the cosmic-ray exposure age of the latter by comparing the measured activity of ^{26}Al with the expected saturation trend. With the Cosmogeophysic group of the University of Torino, we estimated the production rates of ^{26}Al and ^{22}Na, expected to be present in the sample, and determined the counting error as a function of the measurement time by Geant4-based simulations. To estimate the counting error, the dead layer thickness of the HPGe detector that will be used for the measurement has been estimated, as well as the gamma detection efficiency. Moreover, a cylindrical container will host the specimen during the measurement, and the best shape of the latter has been determined as a function of the expected sample mass. The ultimate purpose of such preliminary analyses is to define the optimal conditions necessary to minimize the counting error and perform the most precise measurement possible.

4.1 The OSIRIS-REx mission

The primary goal of the OSIRIS-REx[1] mission is the collection of pristine carbonaceous regolith from the asteroid (101955) Bennu. As carbonaceous asteroids like Bennu are believed to preserve material from the early Solar System, in-depth analysis conducted on the sample will help understand what role primitive asteroids played in both the origin of life on Earth and in the formation of planets [127]. In Fig. 4.1, a mosaic image of asteroid Bennu composed of 12 PolyCam images collected by the OSIRIS-REx spacecraft from a range of 15 miles is shown[2].

The OSIRIS-REx spacecraft was launched on September 8^{th}, 2016, on an outbound-cruise trajectory that resulted in a rendezvous with Bennu in November 2018. Several instruments on the spacecraft allowed for measuring the physical, geological, and chemical properties of Bennu, assessing its resource potential, and refining estimates of its probability of impacting the Earth [128]. These data, providing an extensive global data set of Bennu, have been used to find a suitable sampling site.

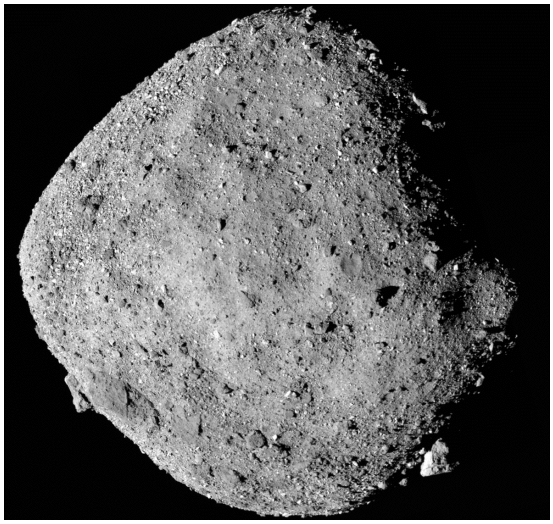

Figure 4.1: Mosaic image of the asteroid Bennu composed of 12 PolyCam images collected by the OSIRIS-REx spacecraft from a range of 15 miles. Picture taken from the NASA website.

On October 20^{th}, 2020, the spacecraft left orbit and descended towards the surface of Bennu for sample collection [129], storing (250 ± 101) grams of material with its Touch And Go Sample Acquisition Mechanism (TAGSAM) [130], consisting of a circular sample collection device connected to the spacecraft by means of an articulated arm that can be extended for several meters. TAGSAM had penetrated 48.8 cm into the subsurface, collecting much more sample mass than the mission's requirement of 60 g [127, 131].

The spacecraft left the asteroid in 2021 and will return the sample to the Utah Test and Training Range (UTTR) on September 2023, whose gamma activity will be measured at the Monte dei Cappuccini Underground Laboratory in Torino.

The following sections discuss the experimental setup that will be used to measure the gamma activity of the sample, as well as the procedure adopted for estimating the precision in determining the activity of some of the radionuclides expected to be present in the extraterrestrial sample.

[1] https://www.nasa.gov/OSIRIS-REx

[2] Picture taken from the NASA website (https://www.nasa.gov/image-feature/goddard/2018/OSIRIS-REx-debuts-new-bennu-images)

4.2 Experimental apparatus: HPGe-NaI(Tl) (GEM150)

The Monte dei Cappuccini underground Laboratory hosts two HPGe-NaI(Tl) spectrometers: the GEM90, described in Section 3.2.1, and the GEM150, which schematic representation is shown in Fig. 4.2. The latter is the most performing one and consists of a p-type closed-ended coaxial HPGe crystal operating in coincidence with an umbrella of NaI(Tl) scintillator. It can hold samples up to $\sim$ 1.5 kg of mass.

The HPGe detector has a mass of $\sim$ 3 kg, a volume of 570 cm^3, a diameter of 91 mm, and is 88.4 mm in height. For the 1332.5 keV ^{60}Co gamma-rays, it has a relative efficiency of $\sim$150%, a resolution (FWHM) of 1.85 keV, and a peak-to-Compton ratio of 104.

The NaI(Tl) scintillator has a detecting mass of 86 kg, a total volume of 24000 cm^3, and a resolution at 662 keV (gamma-ray of ^{137}Cs) equal to 10%. It consists of an annular ring and a cylindrical plug to seal the cavity from above. The ring has a mass of 81 kg, is 300 mm high, has a diameter of 335 mm, and has a cavity of 13.5 cm in which the HPGe is located. The cylindrical plug has a mass of 5 kg, is 100 mm high, and is 125 mm in diameter.

The NaI annulus is coupled with six PMTs placed on its upper surface, while the cylindrical plug is equipped with only one PMT. The system is provided with a nitrogen flushing system to minimize the contribution of the ambient radon and its decay products, as well as to prevent humidity and avoid condensation.

A passive Pb-Cd-Cu shield (200-1-50 mm in thickness, respectively) surrounds the spectrometer to further reduce the incoming radiation from the surrounding environment. The dewar, in thermal contact with the detector, contains up to 60 liters of liquid N_2.

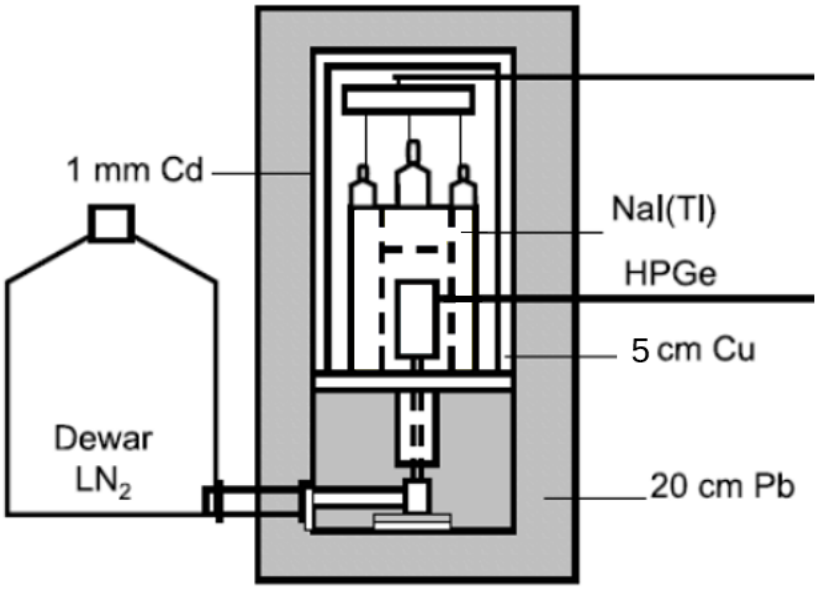

Figure 4.2: Schematic representation of the GEM150 detector.

Standard	Energy [keV]	cpm
^{26}Al	1129.67	0.53 ± 0.01
	1808.65	12.9 ± 0.02
	1808.65 - 511*	0.32 ± 0.01
	2938.32	0.073 ± 0.001
^{60}Co	1173.23	1924 ± 1
	1332.49	1779 ± 1
	2505.69	150.7 ± 0.4

Table 4.1: Results of the measurements of the ^{26}Al and ^{60}Co standards on GEM150. The counts per minute (cpm) of each considered peak are shown with their uncertainties. The results on single-escape peaks of the main lines of ^{26}Al are also reported.

4.2.1 Dead layer thickness estimation

As discussed in Section 3.4.1, since the actual thickness of the DL cannot be directly measured, its value can be inferred by comparing the measurement of standards with simulations. Table 4.1 summarizes the results of the measurements of the ^{26}Al and ^{60}Co standards performed with the GEM150 detector. These measurements have been simulated by increasing the DL thickness from its nominal value (0.7 mm) up to 10 mm for the ^{26}Al standard and 4.9 mm for the ^{60}Co standard, with a step of 0.3 mm.

The results of this procedure are shown in Figs. 4.3 and 4.4, for ^{26}Al and ^{60}Co, respectively. For each peak, the DL thickness was estimated as the intersection of the simulated values (red curve) with the experimental results (black horizontal line). The red lines plot the 2nd-degree polynomial fit of the difference between measured and simulated counts as a function of the DL thickness, while the red bands mark the 68% confidence intervals of each fit. The vertical dashed black lines indicate the intersection intervals with the experimental result, represented by the horizontal black line.

Figure 4.5 shows the final estimate of the GEM150 DL thickness (red horizontal line), equal to (2.24 ± 0.05) mm, obtained as the weighted mean of the DL values estimated with the simulation performed over the ^{26}Al (blue) and ^{60}Co (green) standards. The estimated DL was then used to simulate the measurement of the Almahata Sitta meteorite, an achondrite belonging to the ureilite group with anomalous properties fallen in 2008 in Sudan, which was counted at the Monte dei Cappuccini underground Laboratory [132]. The measured fragment is the largest one and is classified as fragment n°15. It has a mass of 75.536 g, a density of approximately 3.11 g cm^{-3}, and a roughly semi-ellipsoidal shape. Its composition was obtained from Ref.[133].

The Almahata Sitta spectrum was simulated 50 times. Table 4.2 shows the comparison between measured and simulated counts of the ^{26}Al and ^{22}Na main peaks (including full-energy and sum peaks), obtained as the mean value among the 50 simulations, The uncertainty is defined as the standard deviation of the 50 simulated peak counts. A good agreement between measured and simulated cpd is overall observed, thus validating the estimated detector DL thickness.

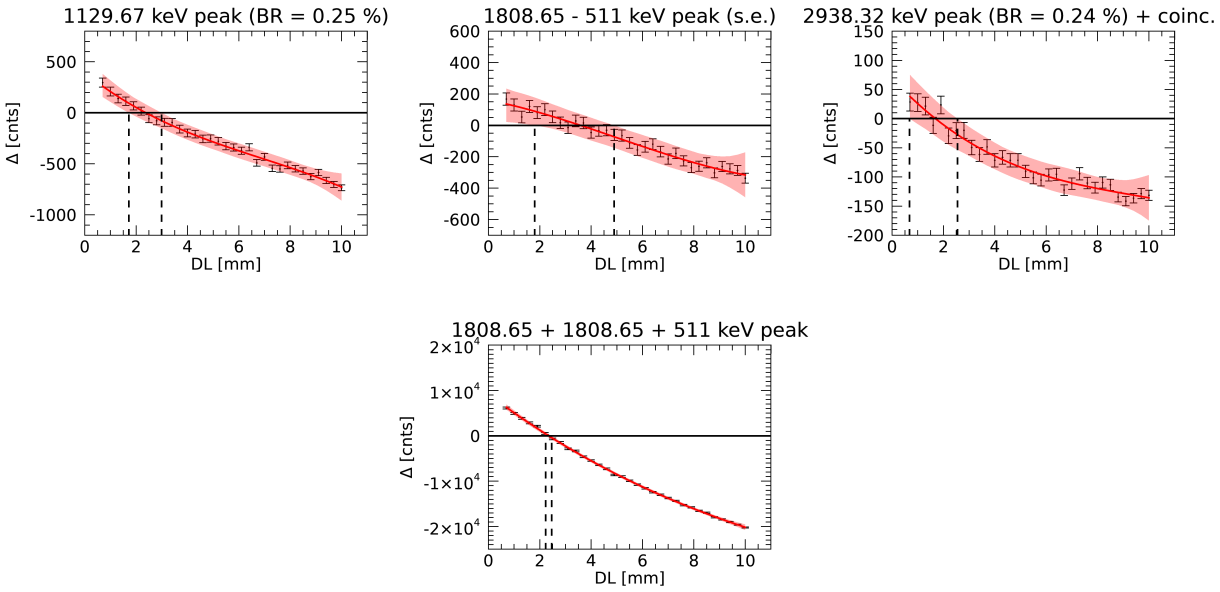

Figure 4.3: Difference (Δ) between ^{26}Al standard simulated and measured counts as a function of the DL thickness of the Ge crystal. The red lines plot the 2$^{\text{nd}}$-degree polynomial fit, and the red bands mark the 68% confidence intervals of the fit. The vertical dashed black lines indicate the intersection intervals with $\Delta = 0$, represented by the horizontal black line.

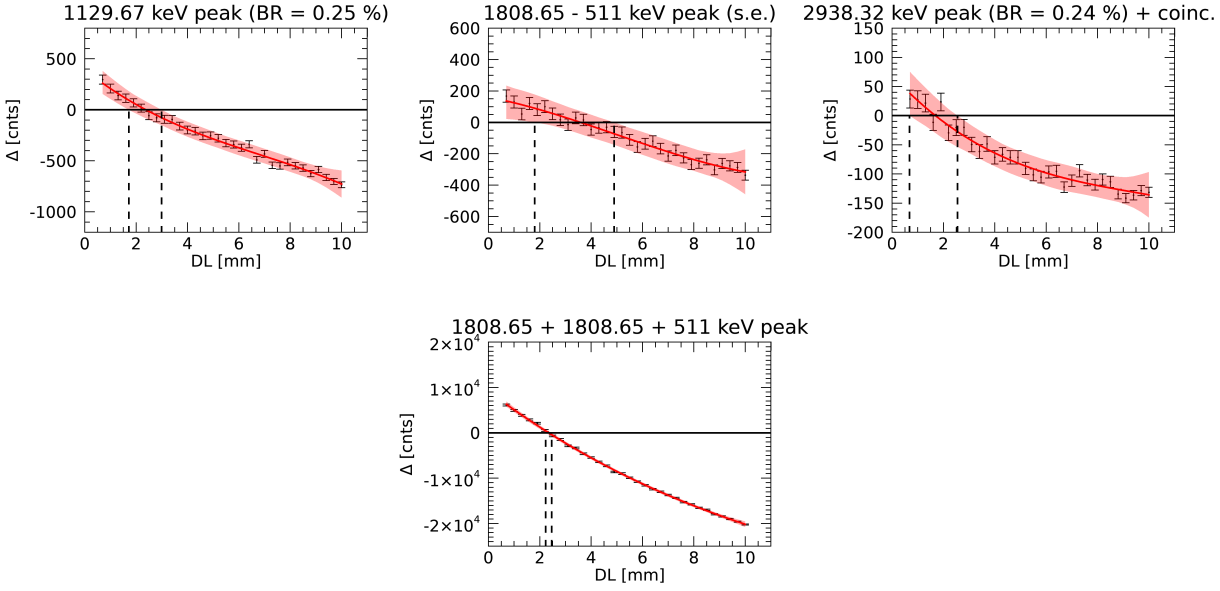

Figure 4.4: Difference (Δ) between ^{60}Co standard simulated and measured counts as a function of the DL thickness of the Ge crystal. The red lines plot the 2$^{\text{nd}}$-degree polynomial fit, and the red bands mark the 68% confidence intervals of the fit. The vertical dashed black lines indicate the intersection intervals with $\Delta = 0$, represented by the horizontal black line.

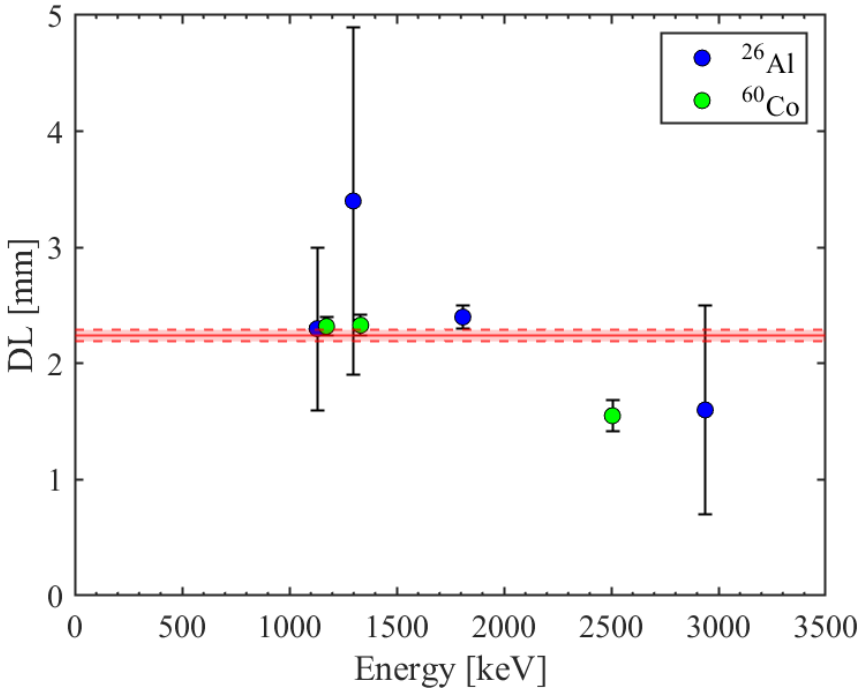

Figure 4.5: Dead layer estimates with associated uncertainties for ^{26}Al and ^{60}Co standards in blue and green, respectively. The red horizontal line is the weighted mean value, and the shaded red bar represents its uncertainty.

Nuclide	Energy	measured cpd	simulated cpd
^{26}Al	1129.67	5.8 ± 1.4	5.7 ± 0.4
	1808.65	144.0 ± 1.4	143.6 ± 1.6
	1808.65 + 511	24.3 ± 0.7	22.9 ± 0.8
^{22}Na	1274.54	273.6 ± 1.4	277 ± 2
	1274.54 + 511	56.2 ± 1.4	53 ± 1

Table 4.2: Comparison between measured and simulated cpd of ^{26}Al and ^{22}Na full-energy and sum peaks with their uncertainties for the Almahata Sitta meteorite.

4.3 Study of the sample container

During its transport from the USA to Italy, the sample will be sealed in a gas-tight nitrogen-filled container with a pressure plate to keep it from moving during handling. An external vessel will also be used to further protect the latter.

Such a container will host the sample for its gamma counting and analyses have been carried out to define its optimal geometry. Due to the symmetry of the HPGe detector, a cylindrical shape represents the best to be measured. Figure 4.6 shows a section of the projected container, where the radioactive sample is shown in grey. The blue and green parts of the figure are made in Teflon and will be in direct contact with the specimen. In particular, the green lid will be inserted to a depth determined by the amount of sample allocated for counting. A system comprised of a spring and an upper block has been designed to provide force to keep the lid in contact with the sample. The height of the spring block will be selected based on the sample volume. Because of the measurement cavity of the HPGe-NaI(Tl) detector, the maximum sample container external diameter allowed is of about 8 cm.

The best dimension of such a container has been estimated through Geant4-based simulations. Figure 4.7 shows a schematic representation of the simulated cylindrical container that will host the sample during the measurement. The darker and lighter blue areas represent the sample and the container, respectively.

The variables in Fig. 4.7 represent:

- the internal diameter D_1;
- the external diameter $D_2 = D_1 + 2$ mm;
- the internal high L_1, that is the maximum high reached by the sample material once it is completely enclosed in the sealed Teflon container, estimated as:

$$L_1 = \frac{m}{\rho\pi(\frac{D_1}{2})^2} \tag{4.1}$$

  where m is the sample mass and ρ the bulk density;

- $L_2 = L_1 + 2$ mm.

In the simulations, a typical composition of a CM chondrite has been used, shown in Table 4.3 [134]. Elements with a weight percentage (wt%) lower than 0.01 have been neglected. The average bulk density ρ of the sample was estimated to be 1.19 kg m^{-3} [135].

Element	H	C	N	O	Na	Mg	Al	Si	P	S	Cl	K	Ca	Ti	Cr	Mn	Fe	Co	Ni	Cu	Zn
wt (%)	1.15	2.32	0.112	42.41	0.4	11.9	1.14	13.2	0.1	3	0.05	0.04	1.2	0.06	0.309	0.17	21.1	0.056	1.25	0.013	0.02

Table 4.3: Expected composition of Bennu (CM chondrite mean composition [134]). The elements are listed in order of increasing atomic number.

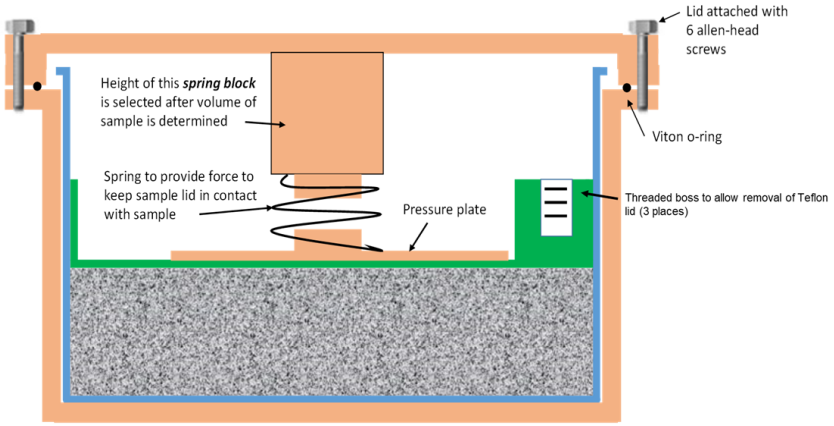

Figure 4.6: Schematic representation of the container that will be used to protect the sample during its transport and host it during the gamma measurement.

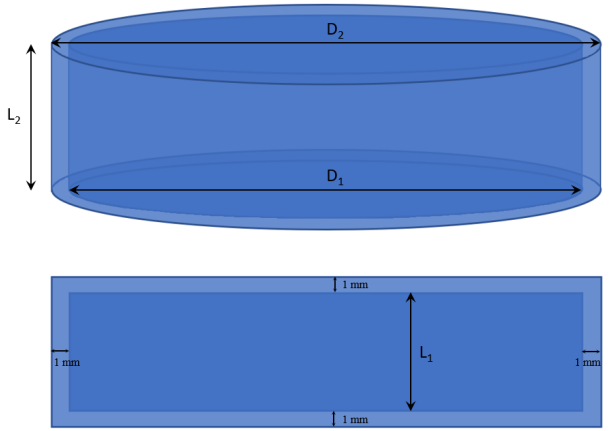

Figure 4.7: Schematic representation of the simulated cylindrical sample container. The scheme on the bottom represents a vertical cutting of the contained. The darker blue area represents Bennu's sample.

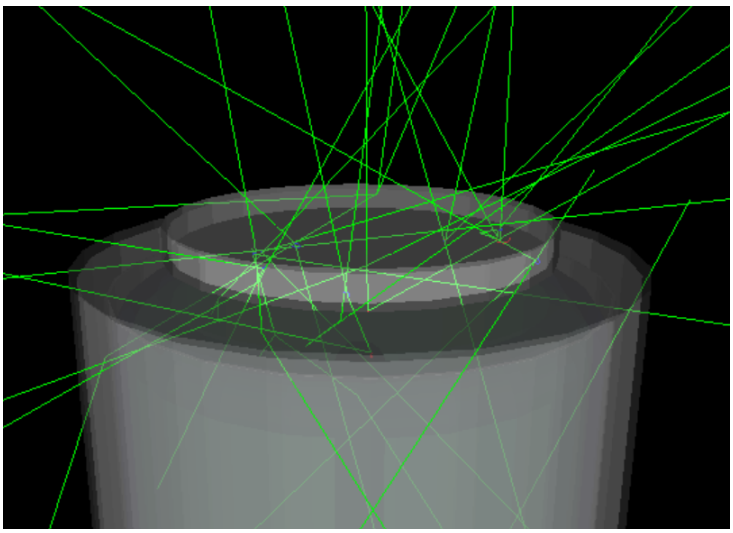

Figure 4.8: Snapshot of the Geant4 simulation of Bennu's sample activity. The cylindrical container is placed over the end cap window of the HPGe detector.

4.4 Detection efficiency estimate

To deduce the activity of ^{26}Al and ^{22}Na in Bennu's sample it is necessary to estimate the detection efficiency η_d with the same approach described in Section 3.4.2. For both radionuclides, $N = 10^6$ decays have been simulated with the Geant4 toolkit. The detection efficiency is finally estimated as the ratio between the measured peak counts and the gamma emitted within the N decays simulated. Figure 4.8 shows a snapshot of the simulation, with the cylindrical sample container placed over the end cap window of the HPGe detector. Because of the high uncertainty regarding the collected mass, several simulations have been performed by varying both the sample mass and the cylindrical contained diameter to define the best configuration (higher detection efficiency) among all possible cases.

Figure 4.9 shows an example of the simulated activity of ^{26}Al (top panel) and ^{22}Na (bottom panel) for a sample mass of 20 g distributed in a cylindrical container with D_1 = 7 cm. Full energy, single escape, double escape, and sum peaks are present.

Figures 4.10 and 4.11 show the estimated detection efficiency for the 1808.65 keV peak of ^{26}Al and the 1274.5 keV peak of ^{22}Na, respectively, as a function of the D_1 diameter and sample mass. The detection efficiency is overall higher for the ^{22}Na peak with respect to the ^{26}Al one, in accordance with the known fact that peaks with lower energies are more efficiently detected. Some mass-diameter combinations are not permitted since the corresponding container high L_2 is not allowed due to the detector cavity dimension. In both cases is possible to notice that the efficiency reaches a maximum value around 70 mm of diameter for each mass. A plateau-like distribution is observed for higher diameter values. Therefore, this value represents the best choice for the diameter D_1 since it leads to the maximum detection efficiency for each possible value of the sample mass within the considered range.

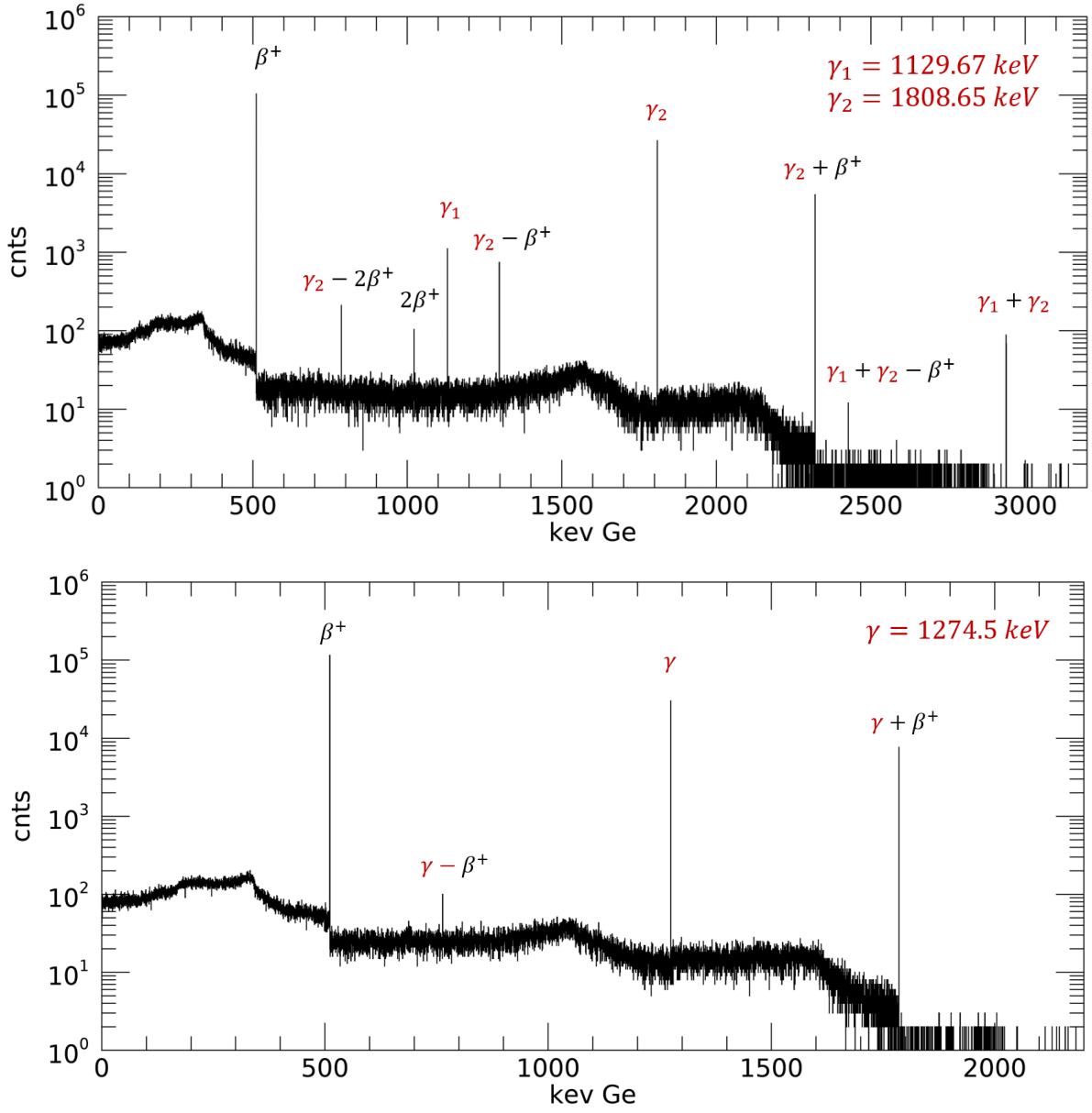

Figure 4.9: Example of the simulation of the ^{26}Al (top panel) and ^{22}Na (bottom panel) activity. Full energy, single escape, double escape, and sum peaks are highlighted.

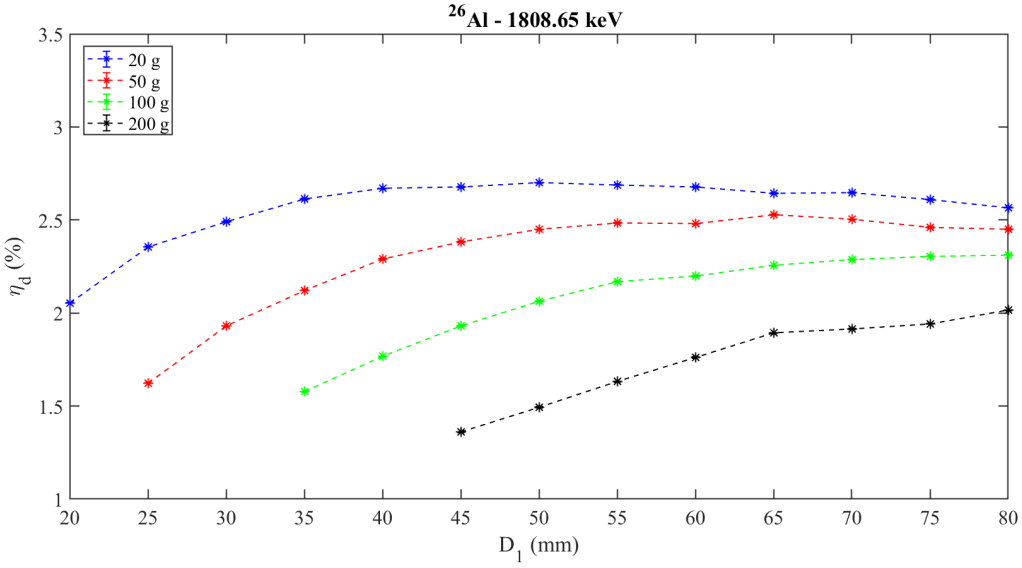

Figure 4.10: Gamma efficiency of the 1808.65 keV peak of ^{26}Al as a function of the container diameter and sample mass. Some mass-diameter combinations are not allowed due to the detector cavity dimensions.

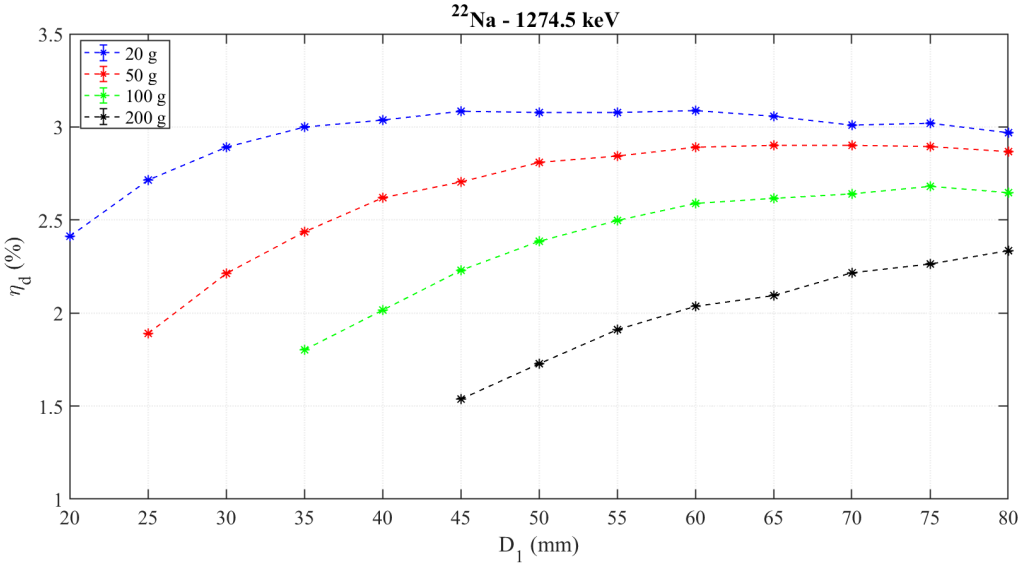

Figure 4.11: Gamma efficiency of the 1274.5 keV peak of ^{22}Na as a function of the container diameter and sample mass. Some mass-diameter combinations are not allowed due to the detector cavity dimensions.

4.5 CR flux parametrization

As already discussed in Chapter 3, the production of cosmogenic nuclides in an asteroid depends on the body size, its bulk chemical composition, and the SCR and GCR flux intensity.

SCR particles are emitted during energetic solar events like flares and CMEs, and consist of 98% protons and 2% alpha particles [136] and a small number of heavier nuclei.

The GCR flux consists of 87% protons, 12% alpha particles, and 1% heavier ions which are injected into the interstellar medium by supernova explosions and reach relativistic speed. The GCR energy spectrum is modulated by the interaction of GCR particles with the solar magnetic field, and this modulation is reflected in the production of cosmogenic radionuclides.

Suitable mathematical parameterizations of both the SCR and GCR spectra exist. For SCR particles, the differential flux density as a function of the rigidity ϕ is usually considered, which for SCR protons is given by:

$$\frac{\delta J_{p,SCR}}{\delta \phi} = J_{0,SCR}(4\pi, E_p > 10MeV)e^{-\frac{\phi}{\phi_0}} \tag{4.2}$$

where $J_{p,SCR}$ is expressed in (protons MV)/(cm^2 s), $J_{0,SCR}(4\pi,E_p > 10\text{MeV})$ is the 4π-integral flux density of protons with energies E_p above 10 MeV, expressed in protons/(cm^2 s), and ϕ_0 is the characteristic rigidity in MV units [88]. The rigidity ϕ of a particle is defined as the relativistic momentum of the latter over its charge.

The energy-differential spectrum of GCR protons at 1 AU ($\sim 1.5\times10^8$ km) can be parametrized as:

$$\frac{\delta J_{GCR}}{\delta E} = J_{LIS}\frac{E(E+2mc^2)}{(E+M)(E+M+2mc^2)} \tag{4.3}$$

where J_{GCR} is expressed in protons/(m^2 s sr), J_{LIS} is a term representing the local interstellar spectrum (LIS), E is the kinetic energy per nucleon, m is the particle mass, and M is the solar modulation parameter expressed in MeV units, described in the following section. Equation 4.3, known as *force-field* approximation, is obtained by assuming azimuthal and spherical symmetry.

There are many expressions for J_{LIS}, mainly differing at low energies. The parameterization used in most of the studies on cosmogenic radionuclides is the one introduced by Castagnoli and Lal [137], which is:

$$J_{LIS,CL} = 9.9 \times 10^8 \cdot (E + 780 \cdot e^{-2.5\times10^{-4}E})^{-2.65} \tag{4.4}$$

where E is expressed in MeV per nucleon while $J_{LIS,CL}$ in protons/(m^2 sr s MeV) per nucleon.

Another LIS parametrization is the one discussed by Usoskin et.al. [138], based on an earlier approach by Burger et.al. [139], equal to:

$$J_{LIS,BG} = \frac{1.9 \times 10^4 \cdot P(E)^{-2.78}}{1 + 0.4866P(E)^{-2.51}} \tag{4.5}$$

where $J_{LIS,BG}$ is expressed in particles/(m^2 sr s GeV) per nucleon, E in GeV per nucleon, and where $P(E)$ is the particle rigidity, given by:

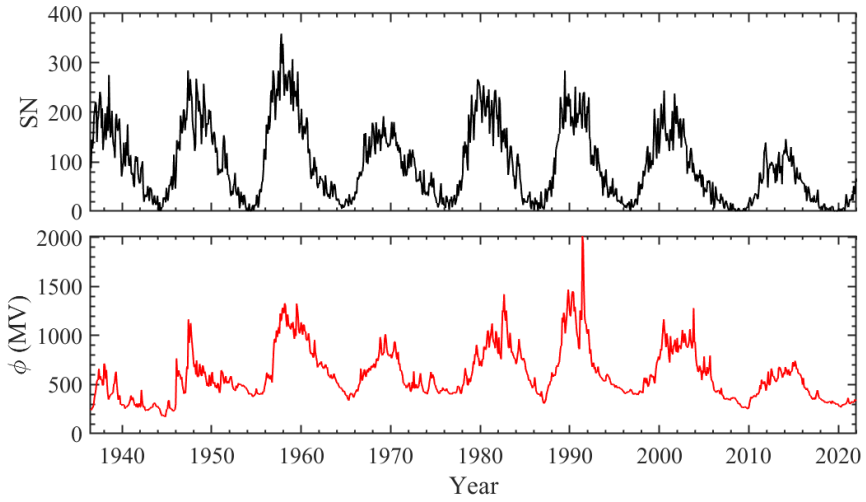

Figure 4.12: Comparison between the sunspot number (top panel) and the rigidity ϕ (bottom panel) monthly series. A direct correlation between the two series is observed.

$$P(E) = \sqrt{E(E + 2mc^2)} \tag{4.6}$$

Primary Galactic alpha particles must be considered separately from primary Galactic protons since they are modulated in different ways. A parametrization for the GCR alpha particles was introduced by Leya et.al.[140]:

$$\frac{\delta J_\alpha}{\delta E} = 5.5 \times 10^7 \cdot \frac{E^k(E + 2m_\alpha c^2)}{(E + 700)(E + 2m_\alpha c^2 + 700)(E + 312500E^{-2.5} + 700)^{(1.65+k)}} \tag{4.7}$$

where J_α is expressed in particles/(m^2 sr s MeV) per nucleon units, E is the energy in Mev per nucleon, and m_α is the mass of the alpha particle (helium nucleus). The adimensional parameter k in equation 4.7 depends on the modulation parameter for alpha particles M_α, and is expressed as:

$$k = 3.572 \times 10^{-3} \cdot M_\alpha - 0.1323 \tag{4.8}$$

The modulation parameter M does not have a physical meaning, but it is an *ad-hoc* parameter used to represent the solar influence on the particle energy spectrum. It can be interpreted as the mean adiabatic energy loss for a galactic particle entering the heliosphere, meaning that such a parameter is anticorrelated with the Galactic particle flux while strongly correlated with solar activity indices. It is calculated as:

$$M = \phi \frac{eZ}{A} \tag{4.9}$$

where ϕ is the rigidity, e is the elementary charge, and Z and A are the atomic number and atomic mass, respectively. As shown in Fig. 4.12, small values of the rigidity ϕ, and therefore of the modulation parameter, are typical of solar minima, whereas larger values refer to solar maxima. The rigidity values shown in Fig. 4.12 have been estimated by Usoskin et.al. [138, 141] using the $J_{LIS,BG}$ parametrization [139].

Figure 4.13 shows the anticorrelation between Castagnoli's parametrization of the proton flux and the modulation parameter. It is important to notice that the flux does not depend on the M parameter when high energies are involved, as high-energy particles are not significantly affected by solar activity.

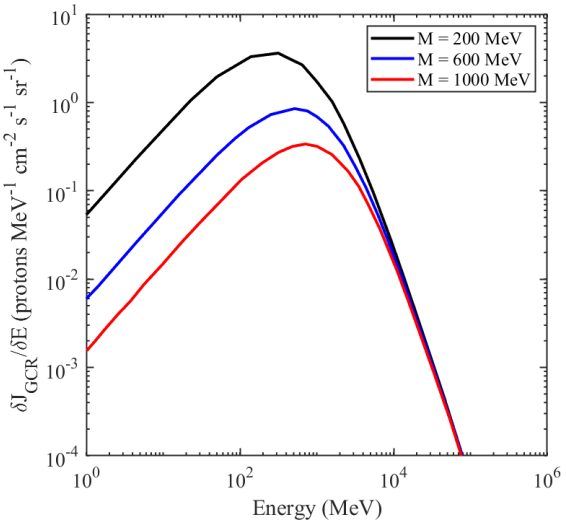

Figure 4.13: Castagnoli's parametrization of the energy-differential spectrum of GCR protons for different values of the modulation parameter M.

4.5.1 Cosmogenic radionuclides production rates: Physical model

As already discussed in Chapter 3, the production of cosmogenic isotopes in celestial bodies is due to their exposure to both SCR and GCR. In physical models of cosmogenic nuclide production, the production rate $P_{i,x}$ of a cosmogenic isotope i at a given depth due to the interaction of the component x of the CR flux with the asteroid, is calculated as the sum of the production rates by SCR and GCR:

$$P_{i,x} = P_{i,SCR,x} + P_{i,GCR,x} \tag{4.10}$$

More in detail, the total production of the isotope i is given by the sum of all the significant components of the flux, including both primary protons and alphas, as well as secondary protons and neutrons produced in the asteroid.

Because of their low energies, the maximum relevant depth for SCR interactions is of the order of magnitude of 10 g cm^{-2}. Therefore, the SCR-related radionuclide production is limited to the outer surface of the irradiated object. The production rate of a cosmogenic nuclide i by irradiation with a SCR proton spectrum can be estimated as follows:

$$P_{i,SCR,p}(R, d, c_s, c_b, J_{0,SCR}, \phi_0) = N_A \sum_j \frac{c_{s,j}}{A_j} \int \sigma_{i,j,p}(E_p) \frac{\delta J_{p,SCR}}{\delta E_p}(E_p, R, d, c_b, J_{0,SCR}, \phi_0) dE_p \tag{4.11}$$

where N_A is the Avogadro's number, R is the asteroid radius, d is the depth, c_s is the chemical abundances vector of target elements, c_b is the bulk chemical composition, A_j is the mass number of the target element j, $\sigma_{i,j,p}$ is the cross-section of the production of isotope i by the interaction between the target element j and the SCR protons with energy E_p, and $J_{0,SCR}$ and ϕ_0 are the parameters already introduced in equation 4.2.

Equation 4.11 is related to the physical model for the production rates discussed in Ref.[142], and it can be considered as the theoretical equation describing the actual physics of the phenomenon. The main complexity in resolving this equation lies in the definition of both the cross-section $\sigma_{i,j,p}$ and the flux dependence on the depth, for which many

empirical and semi-empirical models have been proposed by different authors [see, e.g., Refs. 143, 144].

In analogy to equation 4.11, the production rate of a cosmogenic nuclide i by irradiation with a GCR component x is given by:

$$P_{i,GCR,x}(R,d,c_s,c_b,M) = N_A \sum_j \frac{c_{s,j}}{A_j} \int \sigma_{i,j,x}(E_x) \frac{\delta J_{x,GCR}}{\delta E_x}(E_x,R,d,c_b,M) dE_x \quad (4.12)$$

This method calculates the depth and size-dependent differential spectra $\delta J_{x,GCR}/\delta E_x$ with Monte Carlo codes such as HETC within the HERMES code system [145], or LAHET [146] (both originating from HET [147]) for high-energy protons and neutrons, coupled with MORSE [148] or MCNP [149] for low-energy neutrons.

The GCR production dominates at high depth, since the maximum relevant depth for GCR interactions is of the order of magnitude of 100 g cm^{-2}.

Due to Bennu's CM chondrite-like composition, ^{26}Al and ^{22}Na are two among the radionuclides expected to be present.

The variation with time of the number of radioactive atoms at a given depth in the asteroid can be expressed as:

$$\frac{dN}{dt} = P_i(t) - A(t) = P_i(t) - \lambda N(t) \quad (4.13)$$

where P_i is the radionuclide production rate, obtained as the sum of the $P_{i,x}$ for all the components x of the CR flux, and A is the element's activity given by the product of the decay constant λ and the number of radionuclides.

Due to the long half-life of ^{26}Al, the production rate of this radionuclide can be considered constant. In this way, equation 4.13 can be solved, leading to:

$$A(t) = P_0(1 - e^{-\lambda t}) = P_0(1 - e^{-\frac{t}{\tau}}) \quad (4.14)$$

where P_0 is the production rate at the time t_0=0 when the exposure to CRs started and at which $N(t_0)$=0. By assuming the secular equilibrium, that is, the balance between production and destruction of the radioisotope ($\lambda \cdot t \sim \infty$), the activity of ^{26}Al reaches a value equal to P_0 and, with the model previously described, it is also possible to estimate the activity of ^{22}Na through the ^{22}Na/^{26}Al ratio for the reasons already explained in section 3.5.

For the asteroid Bennu, the mean ^{22}Na/^{26}Al ratio of equation 3.8 was calculated with the model described in Ref.[143], obtaining a value of 1.7. The ^{22}Na/^{26}Al ratio has been reconstructed for the past years, and a value of 2.1 was estimated for October 2020, the period in which the sample acquisition took place.

Since TAGSAM had penetrated 48.8 cm underground for sample collection, the total activity of both radionuclides was calculated down to 50 cm in depth. In Fig. 4.14, the production rate of ^{26}Al (red curve, left panel) and ^{22}Na (red curve, right panel) estimated with the physical model, is shown, obtained as the sum of the GCR (blue curves) and SCR (green curves) productions, as described in equation 4.10.

In summary, with the physical model, it was possible to calculate the ^{26}Al activity by assuming the secular equilibrium, while the ^{22}Na production rate has been indirectly estimated by using the expected ^{22}Na/^{26}Al ratio.

The following section discusses an alternative semi-empirical method for which the production rates of ^{26}Al and ^{22}Na by GCR are independently estimated using Geant4-based simulations. Finally, the total production rates are obtained by combining the results of both the physical and the semi-empirical methods.

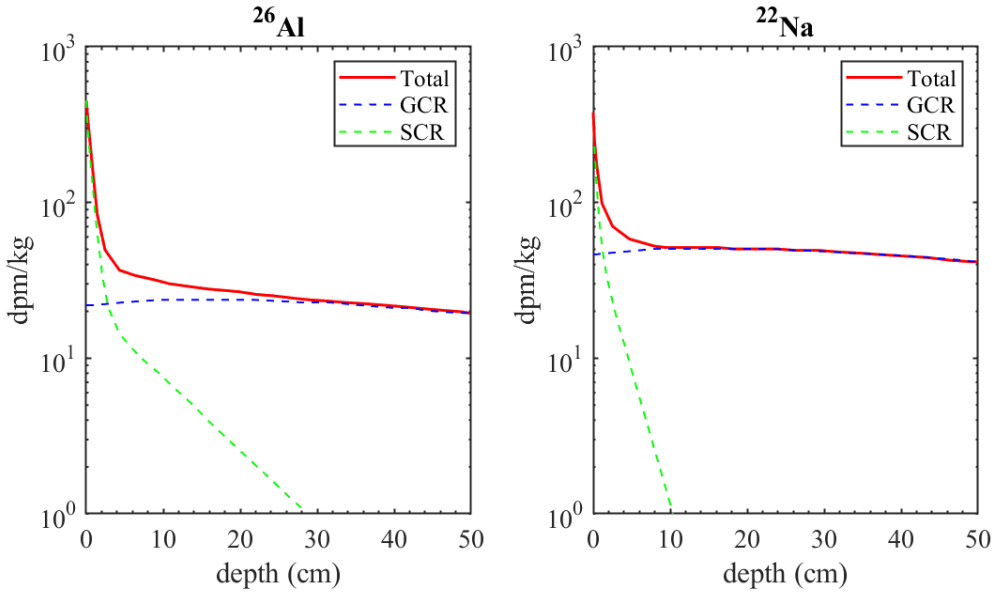

Figure 4.14: Total production rates (red curves) of ^{26}Al (left panel) and ^{22}Na (right panel) obtained as the sum of GCR (blue curves) and SCR (green curves) productions.

4.5.2 Cosmogenic radionuclides production rates: Geant4 model

The Geant4 toolkit [92–94] has been recently successfully employed in the field of cosmogenic radionuclide production [see, e.g., Refs. 140, 150]. Here, the ^{26}Al and ^{22}Na activities have been estimated by simulating the GCR flux irradiating Bennu's surface. Bennu has been modeled as a homogeneous sphere with the typical composition of a CM chondrite described in Table 4.3. An isotropic GCR flux has been implemented, and both protons and alpha particles have been considered thanks to their parametrizations (see equations 4.3, 4.4 and 4.7). The isotropic GCR flux on Bennu's surface was obtained by assuming a cosine-law angular emission and considering the CRs emitting source as a sphere with radius R_{source} (ideally equal to infinity) concentric to the asteroid.

With the hypothesis of secular equilibrium, by simulating a certain amount of flux time, the radionuclide activity is directly determined since the number of radionuclides generated is equal to the number of decays in the same time interval. As discussed in Refs.[151, 152], once there are enough simulated particles such that the number of radionuclides produced is directly proportional to the number of particles generated, the radioactivity A_s obtained within the simulated time interval can be converted in the theoretical activity A_t with the following proportion:

$$A_t = A_s \frac{\Delta T_t}{\Delta T_s} \tag{4.15}$$

where ΔT_s and ΔT_t are related to the number of generated particles and represent the simulated and theoretical time intervals, respectively.

The number of particles per second N_x to simulate was determined by integrating the flux with the cosine-law over the energies, the source surface S, half the solid angle ω, and the angle θ between the radial direction and the actual direction of the emitted particle. Since it is not needful to simulate particles not hitting the asteroid, it is convenient to limit the emission direction by implementing a maximum value of θ. In this way, N_x can be calculated as:

Nuclide	M (MeV)	$\int_0^\infty J_{GCR,p}dE$	N_p	$\int_0^\infty J_{GCR,\alpha}dE$	N_α
^{26}Al	660	1.9×10^3	4.4×10^8	2.3×10^2	5.2×10^8
^{22}Na	309	3.9×10^3	8.9×10^9	3.8×10^3	8.7×10^8

Table 4.4: Number of protons and α particles per second to be simulated, together with the estimated modulation parameter M and the value of the integral of the corresponding GCR-flux parametrizations.

$$N_x = \int_0^\infty \frac{\delta J_{GCR,x}}{\delta E} dE \int_S dS \int_0^{2\pi} d\omega \int_0^{\theta_{max}} cos\theta sin\theta d\theta = 4\pi R_{source}^2 sin^2\theta_{max} \int_0^\infty \frac{\delta J_{GCR,x}}{\delta E} dE \tag{4.16}$$

With the hypothesis of a spherical concentric source, Bennu's radius R_B can be expressed as:

$$R_B = R_{source} \cdot sin\theta_{max} \tag{4.17}$$

Therefore, N_x can be expressed as:

$$N_x = 4\pi R_B^2 \int_0^\infty \frac{\delta J_{GCR,x}}{\delta E} dE \tag{4.18}$$

Since the activity of each radionuclide is affected by the CR flux for at least the past three half-lives, it is incorrect to consider the instantaneous GCR flux while solving equation 4.18. Therefore, an average value of the GCR flux in the past should be considered. For this purpose, a suitable half-life-dependent modulation parameter has been determined for each radionuclide as follows:

$$\overline{M}[n] = \frac{\sum_{i=1}^n M[i] \cdot e^{-\lambda(t_n - t_i)}}{\sum_{i=1}^n e^{-\lambda(t_n - t_i)}} \tag{4.19}$$

where n is the index representing the month of sampling (October 2020). Following equation 4.19, the modulation parameter for ^{22}Na has been estimated, being equal to 309 MeV. Regarding ^{26}Al, because of its very long half-life, the modulation parameter series available (see Fig. 4.12, bottom panel) is too short for the estimate described in equation 4.19, and the value of 660 MeV, estimated in Ref.[140], was considered.

According to Ref.[129], a value of 241 meters has been considered for R_B. Finally, the number of protons N_p and alpha particles N_α per second to be simulated is shown in Table 4.4, together with the estimated value for the integral of the corresponding GCR flux.

By comparing the results of the two approaches adopted for GCR-induced radioactivity, similar results are observed [153] and, since the Geant4-based model also accounts for the secondary particles produced in the extraterrestrial body, the latter has been chosen and summed to the physical-based model estimate of the SCR flux to calculate the total flux. Therefore, by combining the production due to the SCR-physical model with the one obtained from the semi-empirical (simulated) GCR flux, the average activity of ^{26}Al and ^{22}Na up to 50 cm of depth are equal to:

$$^{26}\text{Al activity} = 33.3 \frac{dpm}{kg} \tag{4.20}$$

$$^{22}\text{Na activity} = 54.5 \frac{dpm}{kg} \tag{4.21}$$

For further detail about the adopted procedure, and results regarding other radionuclides expected to be present in Bennu's sample, see Ref.[153].

4.6 Counting error as a function of measurement time

Once the gamma detection efficiency and the expected activity of the radionuclides of interest are estimated, the counting error as a function of the measurement time can be calculated. Given the estimated and measured background activity, the latter obtained by performing gamma activity measurement with no sample on the top of the Ge crystal, the cpm of each of the ^{26}Al and ^{22}Na emitted peaks as a function of the measurement time T_m are calculated by simulations of the expected spectrum. This procedure involves the following steps:

- the background counts are expressed in cpm by scaling them for the measurement time (153.6 days);
- the ^{26}Al (or ^{22}Na) peak of interest is added to the background counts, knowing that the expected cpm is given by:

$$\text{radionuclide activity} \cdot \text{sample mass} \cdot \text{gamma efficiency} \tag{4.22}$$

- the simulated spectrum is obtained by estimating the counts for each channel equal to a random value taken from a Poissonian distribution with a mean value equal to:

$$cpm \cdot T_m \tag{4.23}$$

Examples of the simulated normal spectra after 200 days of counting, and for a sample mass of 20 g placed in a contained 7 cm in diameter, are shown in Fig. 4.15. The panel on top shows the comparison between the measured background (black) and the simulated (red) normal spectra, the latter including the expected 1808.65 keV peak of ^{26}Al. The bottom panel shows the same comparison for the 1274.5 keV peak of ^{22}Na.

Peak counts are finally estimated for each of the 1000 spectrum simulations over a wide measurement time range (from 10 to 200 days). Finally, the counting error, defined as the relative percentage error, is calculated for each simulation as follows:

$$rel_{err}(\%) = \frac{\Delta cnts}{cnts} \cdot 100 \tag{4.24}$$

The final counting error corresponds to the mean value over the 1000 simulations and the associated uncertainty is represented by the standard deviation.

Figure 4.16 shows the counting error as a function of the days of counting and the sample mass for the detection of the 1808.65 keV energy peak emitted by ^{26}Al. Figure 4.17 shows the same results for the 1274.5 keV peak of ^{22}Na. The dotted horizontal grey line in both figures highlights the 1% relative error level. The greater the mass of the sample, the shorter the measurement time required to reach this level.

As expected, the longer the sample is measured, the lowest the error in estimating the radionuclide activity. Thanks to the higher gamma detection efficiency for ^{22}Na, the counting error is overall lower compared to ^{26}Al.

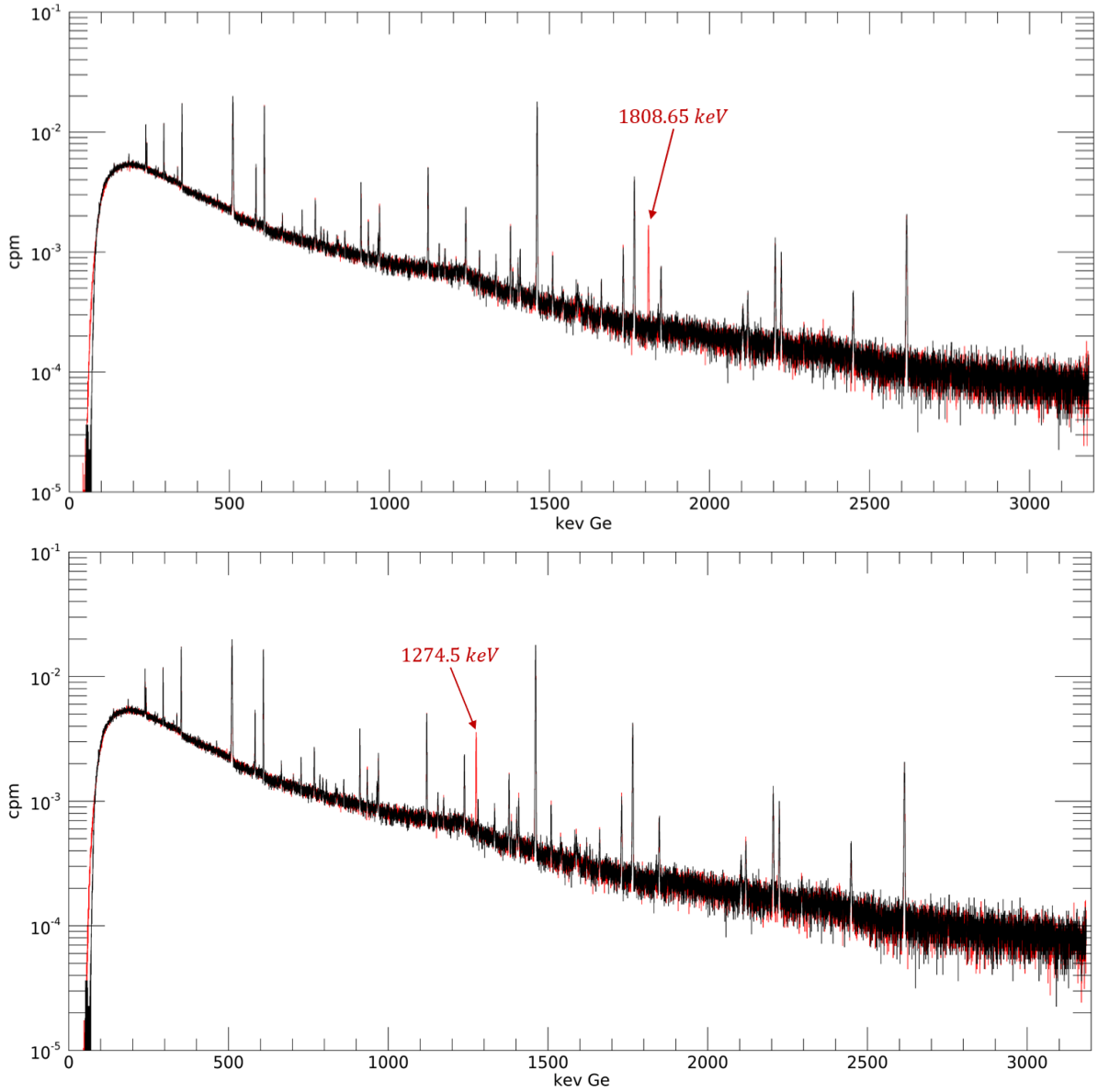

Figure 4.15: Comparison between the measured background spectrum (black) and a simulated one (red), the latter including the expected counts for the 1808.65 keV peak of ^{26}Al (top panel) and the 1274.5 keV peak of ^{22}Na (bottom panel).

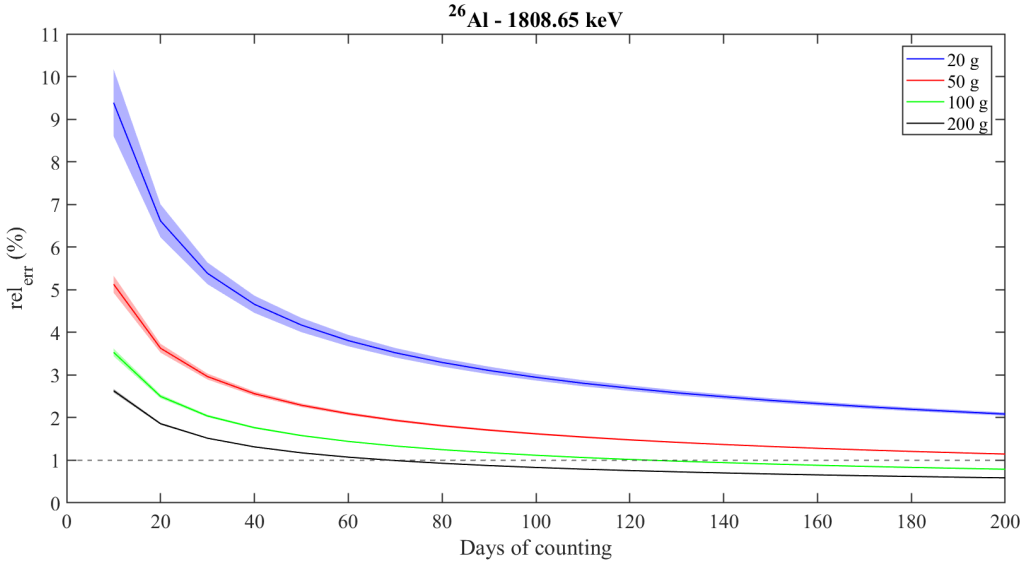

Figure 4.16: Counting error as a function of the days of counting and the sample mass for the detection of the 1808.65 keV peak emitted by ^{26}Al. The horizontal grey dotted line highlights the 1% relative error level.

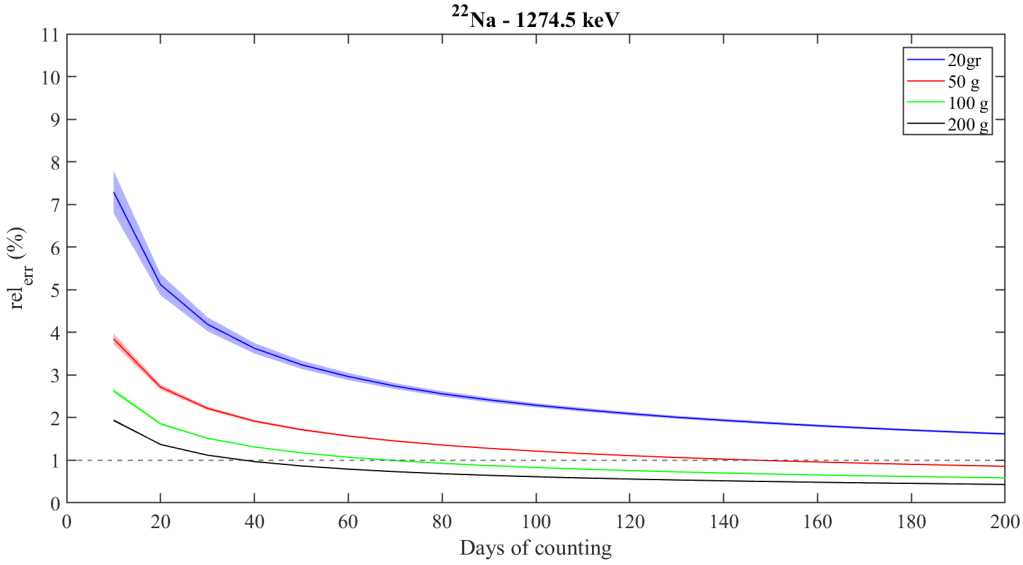

Figure 4.17: Counting error as a function of the days of counting and the sample mass for the detection of the 1274.5 keV peak emitted by ^{22}Na. The horizontal grey dotted line highlights the 1% relative error level.

4.7 Conclusions

The gamma activity of the sample taken from the surface of the asteroid (101955) Bennu will be measured at the Monte dei Cappuccini underground Laboratory in Torino with a high-efficient HPGe-NaI(Tl) spectrometer. This measurement will allow the detection of the cosmogenic radionuclides present in the extraterrestrial material and the determination of the cosmic-ray exposure age of the latter. A cylindrical container will host the specimen during the measurement, the best shape of which was determined as a function of the expected sample mass using Geant4-based simulations.

The spectrometer's DL thickness was determined by simulating the activity of standards, and its estimated value is (2.24 ± 0.05) mm.

The detection efficiencies for the main peaks of ^{26}Al and ^{22}Na were estimated as a function of the sample mass and the internal diameter of the cylindrical container. The optimal internal diameter of the container is equal to 7 cm since for this value the maximum gamma efficiency is obtained for each assumed value of the sample mass.

The production rates of ^{26}Al and ^{22}Na have been estimated by evaluating the total CR flux with a combination of physical and semi-empirical methods.

Finally, the counting error as a function of the sample mass and measurement time was estimated for the 1274.5 and 1808.65 keV peaks of ^{26}Al and ^{22}Na, respectively, by simulations of the expected gamma spectrum. The lower the sample mass, the more counting days are required to achieve a small counting error.

All the analyses described in this chapter have the ultimate goal of defining the optimal conditions necessary to carry out the most precise measurement possible, a measurement that will begin as soon as the sample will arrive at our Laboratory in 2023.

Chapter 5

Solar activity forecast

Forecasting solar activity is crucial nowadays to predict the impact of the Sun on both Earth's magnetosphere and terrestrial climate. Predicting upcoming solar cycles allows anticipating adverse space weather effects on the Earth's environment, as well as power systems, satellites, and manned space missions, produced by coronal transients and traveling interplanetary disturbances.

The mechanism behind the origin of the decadal solar cycle is related to an intrinsic magnetohydrodynamic (MHD) dynamo process sustained by complex interactions between the magnetic field in the solar convection zone and plasma flows [154]. This process implies a decadal cyclic regeneration of the toroidal and poloidal components of the solar magnetic field [155], also known as the Schwabe cycle, which is also responsible for the modulation of the frequency of all eruptive phenomena that may directly impact planetary magnetospheres. As a consequence of such a mechanism, sunspots originate in the solar photosphere due to the buoyant emergence of toroidal flux tubes generated inside the solar interior. Such temporary phenomena provide the most famous proxy of solar activity.

Several predictions have been published in the literature using both physical-based models and neural network approaches [156–178]. Because of the underlying complexity of such a dynamo mechanism, the predictions offered by state-of-the-art machine learning algorithms represent valuable tools for understanding the cycle progression and its forecasting. The latest advances in deep learning techniques provide new prediction models that allow forecasting the evolution of cosmogeophysical time series. Among them, Bayesian deep learning is particularly compelling since it provides an improvement in both uncertainty quantification and accuracy compared to classical techniques.

This chapter describes the forecast of the complete profile of Solar Cycle 25 obtained by employing a deep learning Long Short-Term Memory (LSTM) model. A rigorous uncertainty estimation is obtained by applying a Bayesian approach. Furthermore, two different model validation techniques, namely the Train-Test split and the time series k-fold cross-validation, have been implemented and compared.

The results are published in Ref.[179].

5.1 Sunspots

Sunspots are temporary phenomena occurring on the solar photosphere, appearing as dark spots differing from the surrounding area in terms of temperature and magnetic activity. Although the details of sunspot formation are still under study, it is a well-known fact that these phenomena originate from a distortion of the lines of force of the solar magnetic field due to the Sun's differential rotation. This distortion involves the generation of magnetic flux tubes in the convective zone of the Sun that are projected across the photosphere creating magnetic loops. As a consequence, bipolar sunspots are formed in two belts, one in the northern and one in the southern hemisphere of the Sun, and their number is correlated with the intensity of solar radiation. Spots of the same hemisphere have the same polarity, which is opposite to that of those in the opposite hemisphere.

Sunspots initially appear in the form of tiny pores. Within a few days, they develop, proliferate, widen, and merge together, giving rise to groups of spots, which generally dissolve within about a month. Their characteristic darkening occurs due to the strong magnetic field inhibiting convection in the photosphere. Another factor contributing to this feature is that, although their temperature is between 3000 and 4500 K, their thermal emissivity is much lower than the surrounding regions, which are even brighter due to their highest temperature of the order of 5800 K. Therefore, they appear as dark spots. The central part of a sunspot, corresponding to the darker and colder area, is called *umbra*, while the surrounding area (*penumbra*) is lighted due to its higher temperature of ∼5000 K. The magnetic field reaches its maximum in the middle of the umbra, where the field lines are usually normal to the surface, while moving towards the penumbra the magnetic field's lines are more tilted.

The location of sunspots is not random. At the beginning of each solar cycle, they originate at mid-latitudes, around 30° in both the hemispheres, and, as the cycle proceeds, they start appearing at lower latitudes, almost reaching the equatorial zone.

The sunspot latitude variation, described by the Spörer's experimental law, and well-represented by the butterfly diagram, lasts about 11 years and defines a solar cycle. The maximum activity of the Sun occurs in the middle of each cycle. Moreover, sunspot polarities reverse from one solar cycle to the following.

By far, the SN, originally defined by Wolf in the 19^{th} century [180], is the most commonly high-resolution index employed to evaluate the variable solar activity over the last centuries. It originates from a linear combination of telescope-aided visual counts of the number of sunspots and the number of groups that they form on the visible disk of the Sun.

Sunspots represent a unique proxy since they constitute the longest quantitative and homogeneous global indicator of solar activity, having been recorded since 1610, even though the construction of this series has been subject to periodic revisions since its formulation.

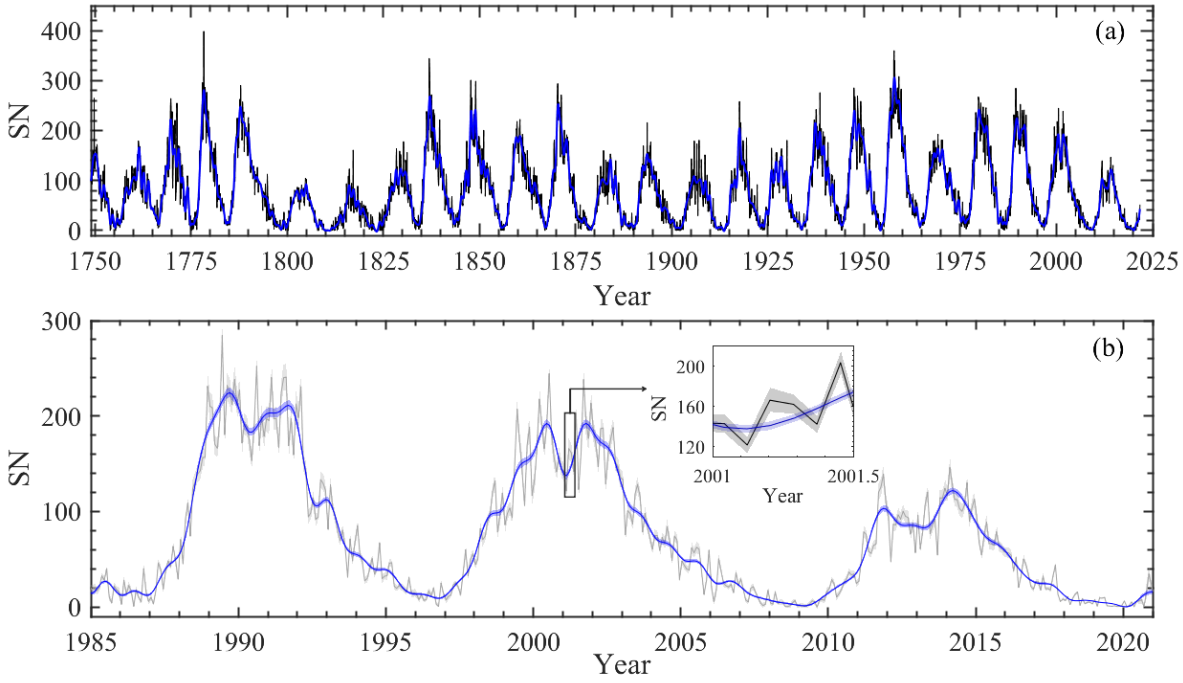

Figure 5.1: Panel a shows the comparison between the original monthly sunspot number series (black curve) and the series once the smoothing has been performed (blue curve). Panel b shows the same comparison over the last three cycles, with relative uncertainties represented as shaded colored bands.

5.2 Data Analysis

5.2.1 Dataset: Sunspot number

The SN dataset is available from the World Data Center (WDC) - Sunspot Index and Long-term Solar Observations (SILSO)[1], Royal Observatory of Belgium, Brussels [40]. This archive is the most authoritative source of SN data, recently recalibrated and revisioned [see Refs. 181, 182], presenting daily values since 1818, monthly means since 1749, and annual means since 1700.

The revised series has been introduced since July 2015 and is now considered version 2.0 of the SN series, although major revisions are expected in the near future since the corrections applied in v2.0 were somewhat crude and the counting of sunspot groups is still a matter of discussion [183].

The SN time series used in this work has been obtained from the SILSO catalogue and consists of 3273 monthly values from January 1749 to September 2021 (black curve in Fig. 5.1a). Besides the dominant decadal Schwabe cycle, the centennial modulation, known as Gleissberg cycle, is also visible, including the Dalton minimum (around 1800) and the Gleissberg one (around 1900). Following Wolf's numbering scheme of solar cycles, the 1755-1766 cycle is traditionally considered as the 1^{st} one, and all the others are numbered in order.

[1] http://www.sidc.be/silso/

5.2.2 Dataset pre-processing

Since the aim of this work is to forecast the next solar cycle and, more in detail, long-term variations of solar activity (periods higher than 1 y), the monthly SN time series has been smoothed using the adaptive filter proposed by Mann [184, 185], that results to be particularly suitable for time series with non-stationary behaviour, among traditional smoothing methods. A Butterworth low-pass filter was applied in association with boundary constraints designed to minimize edge effects, which are:

- *minimum norm*, which tends to minimize the amplitude of the smooth near the boundaries;
- *minimum slope*, which tends to minimize the slope of the smooth near the boundaries;
- *minimum roughness*, which seeks the smoothest behaviour near the time series boundaries by tending to minimize the second derivative of the smooth near the boundaries.

The constraints among these three that minimize the mean square error (MSE), with respect to the raw time series, can be chosen as being the most appropriate for the particular time series and time interval being analysed. The final smoothing S(t) will be a combination of the three smoothing of the time series T(t), obtained by individually applying the three constraints previously introduced:

$$S(t) = w_1 S_1(t) + w_2 S_2(t) + w_3 S_3(t) \tag{5.1}$$

where the coefficients w_i are defined in such a way that the quantity:

$$\sum [S(t) - T(t)] \tag{5.2}$$

is minimized, subject to the constraints:

$$\sum_{i=1}^{3} w_i = 1 \quad , \quad w_i \in [0, 1] \tag{5.3}$$

An open-source Matlab [186] algorithm[2] has been employed for this purpose, which applies a 10-point Butterworth Infinite Impulse Response (IIR) low-pass filter of specified cutoff frequency, that was set equal to f_0 =1/12 months^{-1}. The rough half-width of the smoothing window associated with the filter is Δt=1/f_0 = 12 months. The obtained smoothed time series is shown in Fig. 5.1a (blue curve), and a comparison with the original series over the last three cycles is shown in Fig. 5.1b, together with the relative uncertainty bars.

The uncertainty of the smoothed series has been obtained by propagating the error of the original time series, which is provided with an uncertainty value on each point, by a Gaussian process regression bootstrapping, a Bayesian non-linear regression method [187, 188]. More in detail, an ensemble of 1000 surrogated time series is generated from the SN monthly time series and then smoothed with the procedure previously described, and the standard deviation of this ensemble with respect to the smoothed SN series is used as an estimation of the uncertainty of the latter.

[2] http://www.meteo.psu.edu/holocene/public_html/smoothing08/

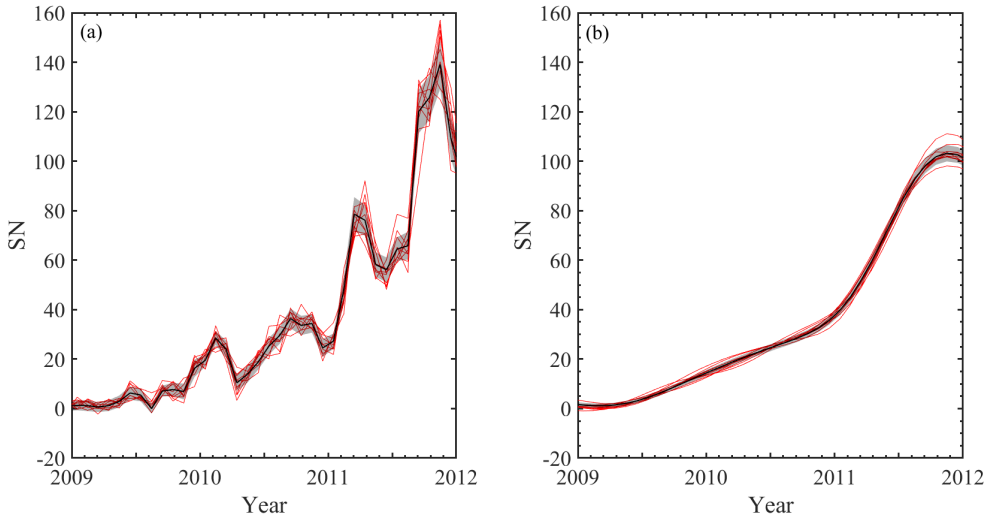

Figure 5.2: (a) Comparison between the original time series (black curve), corresponding to the ascending phase of Solar Cycle 24, and ten of the 1000 surrogated time series generated with the gaussian process regression bootstrapping (red curves). Panel b shows the same comparison after the smoothing process is performed on each series. The black shaded bands represent data uncertainties.

Figure 5.2 shows an example of the surrogated time series generated with such a non-linear method. Panel a shows the comparison between the original monthly series (black curve), corresponding to the ascending phase of Solar Cycle 24, and ten of the 1000 surrogated time series (red curves), while panel b shows the same comparison after performing the smoothing process on each series. The same ensemble will also be used to propagate the uncertainty to the forecast on the next cycle (see Section 5.3.3).

5.2.3 Spectral analysis

Before predicting future values of the smoothed SN time series, a spectral analysis has been performed by Continuous Wavelet Transform (CWT) [12, 36], to detect the deterministic components of the record. As already introduced in Section 2.2, the CWT method gives an evolutionary analysis of the series in the time-scale plane and allows to study non-stationary features of the signal, such as changes in periodicity, isolated events, trends, and intermittency (see Appendix A for more detail).

Figure 5.3 shows the monthly smoothed series (panel a) together with the relative CWT power spectrum (panel b), where areas with a high (low) power are colored red (blue). The black curves enclose regions with a confidence level greater than 99% against the null hypothesis of a red-noise process, while the COI is represented by the white curve. Figure 5.3c shows the Global Wavelet Spectrum (GWS) (black curve), obtained by time-averaging the CWT spectrum over each period, and the corresponding significance levels (red curve).

Most of the signal variance is described by the decadal modulation. The high power of such a component, represented by the orange/red bar extended over the entire time range, shows the well know period modulation, oscillating between about 8 and 12 y, resulting in a large red band centered around 10-y of period.

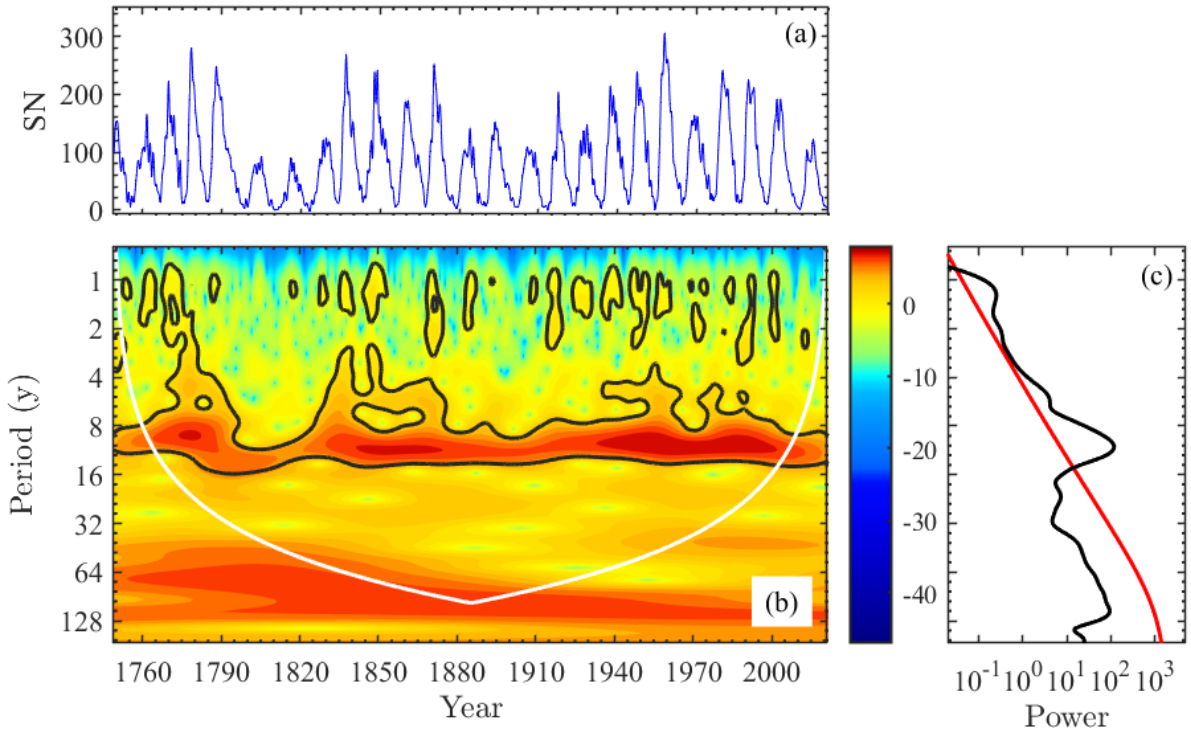

Figure 5.3: Continuous Wavelet power spectrum (panel b) of the SN smoothed time series (panel a) and Global Wavelet Spectrum (panel c). The black curves in panel b enclose regions with a confidence level greater than 99% against a red-noise process null hypothesis, while the COI, indicating regions influenced by edge effects, is represented by the white curve.

A significant amount of spectral power is distributed around 1-y of period, as also visible in Fig. 5.3c. The power of such component is lower around the Dalton and Gleissberg minima and gets higher in correspondence with the Schwabe cycle maxima, as visible in Fig. 5.3b. A less powerful component with a period in the range of 2-3 y is present, although it is not revealed along the whole interval covered by the series. Because of the limited length of the record, the Gleissberg centennial modulation is not detected by the CWT analysis since it lies outside the COI (white curve in Fig. 5.3b).

The results of the CWT analysis are confirmed by the SSA of the series detrended by the decadal cycle, revealing two interannual cycles with periods of about 1.2 y and 2.7 y, significant at a 99% c.l.. This result was obtained by using a null hypothesis of red noise (not shown).

The presence of such oscillations is not unexpected, since the large-scale magnetic field of the Sun is known to undergo considerable variations on timescales much less than a solar cycle but longer than a rotation period. In particular, the Sun's dipole magnetic moment and open magnetic flux are known to exhibit multiple peaks around frequencies of 1 y^{-1}. The generating cause of these fluctuations can be assessed to stochastic processes caused by the periodic emergence of magnetic flux during the solar cycle with a decay time of about 1 y [see, e.g., Ref. 189], which is determined by differential rotation, meridional flow and supergranule diffusion [see also the comprehensive review of Ref. 190].

The last three solar cycles show a double-peak shape, the so-called Gnevyshev gap, believed to be a manifestation of the quasi-biennial oscillation (QBO) [see, e.g., Ref. 191, and references therein], which was found to be a prominent feature during the cycle

maximum phase [192]. Frequencies of a quasi-biennial cycle have been detected in several indicators [see, e.g., Refs. 193–195].

The presence of oscillations with higher frequency than the decadal cycle allowed to identify the limits of the predictive method used in forecasting time series with multiple periodicities.

5.3 Methodology

5.3.1 Long Short-Term Memory Neural Network

An artificial neural network (ANN) is an algorithm which takes inspiration from the functioning principles of the nervous system of evolved organisms, built by interconnecting elementary computational units called nodes, units or neurons. There are several types of ANNs, and they work very well on a large variety of problems. One of their fundamental characteristics is that knowledge about the specific problem of interest is acquired from a learning process and is stored in the network parameters.

The elementary unit of an ANN is a formal neuron. Each neuron is connected at least to another one, and each connection is weighted by a real number describing the degree of importance of the given connection in the neural network. Another parameter that can be introduced in the neural network is the bias, which is an additional weight that does not depend on any input data. This weight is useful to tune the optimal working point of the neuron, by adjusting the output along with the weighted sum of the inputs to the neuron.

Neural network neurons are arranged into layers. The most simple network consists of an input layer and an output one. The network becomes more complex if more layers are added between the two mentioned before. These further layers are usually called *hidden layers*, and ANNs with at least one hidden layer are called deep networks. The hidden layers performs several computation on the features entered through the input layer and transfers the result to the output layer.

Each neuron contains an activation function, which is meant to limit the output of the neuron usually to values in the range [0, 1], or [−1,1], and which is chosen in such a way to introduce non-linearity in the network. If no activation functions are introduced, each neuron will only be performing a linear transformation on the inputs using the weights and biases. All layers will behave the same and the neural network will be just a linear regression model. The bias can also be seen as a parameter used for shifting the activation function toward left or right.

Among all the existing classes of ANNs, recurrent neural networks (RNNs) are particularly suitable to deal with time series forecasting since, differently from traditional neural networks, inputs and outputs are not assumed to be independent. These networks are distinguished from others by their *memory*, which holds information from previous inputs and outputs to influence the current output.

Long Short-Term Memory (LSTM) neural networks are a special kind of RNNs, firstly introduced by Hochreiter and Schmidhuber [196] and later modified and improved by several authors [197, 198].

The architecture of LSTM networks includes layers, each of which is a set of recurrently connected subnets called memory blocks, also known as memory cells, which refer to a single time-steps. An example of three consecutive cells is shown in Fig. 5.4. Each cell contains three multiplicative units called gates, which are:

- the forget gate, which chooses the information to discard from the cell. The output from this gate is:

$$f_t = \sigma(W_{hf} \cdot h_{t-1} + W_{xf} \cdot x_t + b_f) \tag{5.4}$$

  where σ is the sigmoid activation function, b_f is the connection bias, and W_{hf} and W_{xf} are the weight associated to h_{t-1} and x_t, respectively.

- the input gate, which decides which values from the input are useful to update the memory state. The value of i_t is given by:

$$i_t = \sigma(W_{hi} \cdot h_{t-1} + W_{xi} \cdot x_t + b_i) \tag{5.5}$$

- the output gate, which determines what to output based on both the current input and the previous output. The output value from this gate is defined as:

$$o_t = \sigma(W_{ho} \cdot h_{t-1} + W_{xo} \cdot x_t + b_o) \tag{5.6}$$

By operating with these gates, each LSTM cell internally evaluates what to keep in and erase from the memory, given the input of its time-step (x_t) and the output (h_{t-1}) and memory (C_{t-1}) from the previous one, and finally generates its output h_t. In Fig. 5.4, $\tilde{C}_t$ represents a vector of new candidate values, which is estimated as:

$$\tilde{C}_t = tanh(W_{hC} \cdot h_{t-1} + W_{xC} \cdot x_t + b_C) \tag{5.7}$$

where $tanh$ is the hyperbolic tangent activation function. This term, together with the outputs from the forget and input gates, contribute to the new cell state C_t as follows:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{5.8}$$

The internal state C_t, that encodes the network's memory, is the peculiarity that makes LSTM networks well-suited to deal with physical time series, typically exhibiting a certain degree of persistence and/or periodical features.

Finally, the value of h_t is given by:

$$h_t = o_t \cdot tanh(C_t) \tag{5.9}$$

and the output y_t is defined as:

$$y_t = softmax(W_{yh} \cdot h_t + b_y) \tag{5.10}$$

The consistent data flow resulting from this architecture keeps each cell stable (thus, neither exploding nor vanishing [196].

The LSTM network are trained following the Back Propagation Through Time (BPTT) algorithm [199], which is the application of the backpropagation training algorithm to recurrent neural network applied to time series. The goal of the backpropagation training algorithm is to modify the weights of the neural network in order to minimize the error, that is how far the output was compared to its desired value. Conceptually, BPTT works

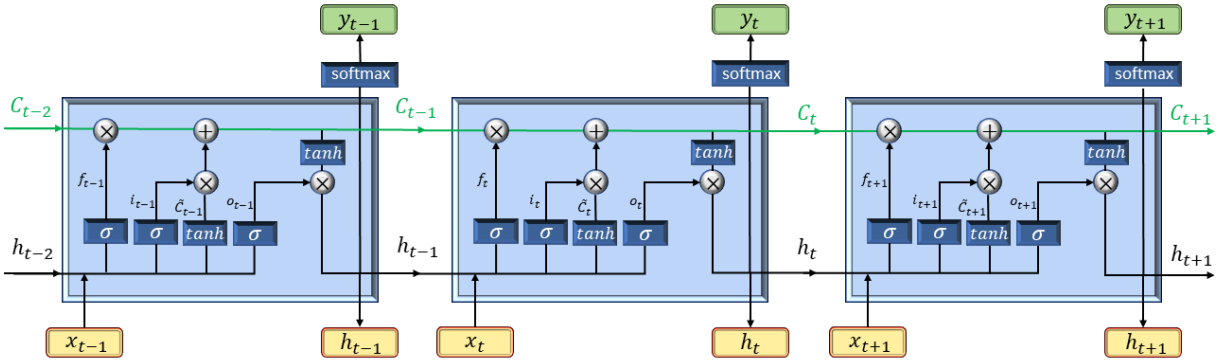

Figure 5.4: Section of an LSTM network chain structure. The horizontal green line running through the top represents the cell's internal state. The gray circles are pointwise operations while the dark blue boxes represent the activation functions. Each LSTM module is related to a given time instant.

by unfolding the network over time. The errors are then calculated and accumulated for each time step and the network is rolled back up. Finally, the weights are updated in order to reduce the error E, also known as *loss function*, as follows:

$$\omega \longleftarrow \omega - \eta \nabla_{\omega} E \tag{5.11}$$

where ω represents a generic weight or bias, η is the learning rate, a hyperparameter that controls how much to change weights and biases to reach a minimum of the error, and $\nabla_{\omega} E$ is the gradient of the loss function with respect to the parameter ω. This process is repeated several times till a minimum of the loss function is reached.

When a lot of training data are available, it is advisable to train the network on a portion of data at a time, usually referred as batch. This training technique is called *batch data processing* and the number of data in each batch is called batch size. Each time the network is fed with a number of samples equal to the batch size, the network weights are updated. The choice of the batch size is not trivial. A batch size which is too large leads to few weights' updates and may not fit into memory, thus taking a long time for the training process to end. Conversely, if the batch size is too small, the number of weights' updates can be too large, leading to overfitting the data. Moreover, it is unlikely that a small number of samples can be a good representation of the entire dataset.

Another important hyperparameter is the number of epochs, which defines the number of times the entire dataset is passed forward and backward through the neural network during the training process. Usually, more than one training epoch is used, allowing the train the network enough until the loss function has been sufficiently minimized.

For a more in-depth description of LSTM neural networks see Refs.[200–202].

5.3.2 Bayesian Neural Network structure and data preparation

Bayesian methods are gaining increasing attention and popularity in the astrophysical context in the last decade [see the recent review of Ref. 203]. In particular, Bayesian inference is central to linking astronomical data to nonlinear astrophysical models in the fields of solar and exoplanetary physics. Their main advantages include the ability to explicitly incorporate into the model prior knowledge about the parameter of interest and

Layer (type)	Activation function	Units	Dropout (%)
Input Layer	-	-	-
LSTM	tanh	[2-200]	-
Dropout	-	-	5
Dense	linear	[2-200]	-
Dropout	-	-	5
Dense	linear	N	-

Table 5.1: Layers of the developed neural network with corresponding activation functions and hidden unit ranges. The dropout percentage between each layer is also shown.

to give a description of the latter in probabilistic terms. On the other hand, the frequentist approach, focused on p-values and null hypothesis significance testing, is subject to common misinterpretation and has been recently criticized, so that likelihood-free methods are now growing in importance. The Bayesian framework offers several alternative indices for the p-value. In particular, the Bayesian method can compute the probability of different effects values given the observed data and some prior expectation, thus resulting in the so-called posterior distribution, which is a distribution of possible values for the parameters.

Following the approach described by Gal and Ghahramani [204], a neural network with dropout applied before every weight layer can be interpreted as a Bayesian approximation of a probabilistic deep Gaussian process. The dropout applied before each layer randomly drops out each hidden unit with a certain probability. Dropout is also a technique for addressing overfitting in neural networks [205].

The structure of the Bayesian neural network developed to forecast the future activity of the Sun is described in Table 5.1. All the network layers are listed together with their activation function, number of units, and dropout percentage. Dense layers are densely-connected neural network layers, generally used to change the dimension of the output of the previous layer. The model was robust enough with a reasonable range of dropout probability, which was set equal to 5%, that is, each time, the 5% of the weights between each pair of layers are randomly dropped out from the network.

Given the dataset D, the model aims to obtain the posterior distribution p(w|D). Gaussian priors for the weight parameters p(w) have been introduced in the LSTM layer and all the Dense layers using the Keras initializer RandomNormal [206].

A supervised learning process has been implemented to train this model, that is examples of input/desired output pairs were presented to the network and the weight coefficients were adjusted to make the predicted and desired outputs as close as possible. The developed network can predict multiple time steps at a time. More specifically, its goal is to predict N values of the series given as input the N previous values. All the algorithms were implemented in Python [207].

Two different model validation techniques have been applied and compared in terms of performance and final forecast:

- Train-Test split technique, that consists in dividing the dataset into three different sets: (i) the training-set, used to train the whole network; (ii) the validation-set, required to diagnose the goodness of the fit during the training process; (iii) the test-set, used to evaluate if the trained network can generalize well if unknown data are given as input.
- Time Series k-fold Cross-Validation technique, which is an adaptation of the classical cross-validation techniques for data where the assumption of independent and identically distributed samples is not valid.

For the sake of simplicity, the Train-Test split and the Time Series k-fold Cross-Validation techniques will be referred to as the TT and CV technique, respectively. Similarly, the models validated with the two validation techniques will be referred to as the TT and CV model, respectively.

A schematic representation of the data-split performed by the TT technique is shown in Fig. 5.5. The training-set length includes about the 88% of the whole series, the validation-set covers almost two cycles, and the test-set length is equal to N, thus including the last solar cycle.

The time series splitting for the CV-technique has been implemented from the machine learning library Scikit-learn [208]. This technique involves splitting the dataset into k consecutive folds, which are arranged to obtain k pairs of training and test data, referred as iterations. The training-set of the first iteration is composed of the first fold, while the test-set is composed of the following one. From one iteration to the next, the training-set gets bigger and bigger, including each time the test-set of the previous iteration, while the new test-set remains of the same size. All surplus data are added to the first training-set.

Figure 5.6 shows a representation of such subdivision, where the training- and test-set of each iteration are shown in blue and red, respectively. The k-fold value was set equal to ten, therefore, the CV-technique involves the evaluation of the neural network over ten different pairs of training- and test-set.

Before starting the network training process, the time series must be transformed into a supervised learning dataset. Since neural networks are less efficient over time series with a large variability scale, the data are re-scaled within $[-1,1]$. In this way, the network's weights can be assigned in a suitable initialization range, which is a crucial requirement in a Bayesian approach [209]. The dataset is then framed to increase the number of samples to be fed to the neural network during the training phase. Each training-, validation- and test-set is converted into a matrix with N columns and a number of rows corresponding to the number of samples using a one-step forward moving window of length N. Dealing with a supervised learning problem, matrices formed by input data and corresponding target data (desired output) are defined for each set. A schematic representation of such framing is shown in Fig. 5.7 for a dataset composed of $n = 10$ values and a moving window of length $N = 3$. Because of the goal of the developed network, giving as input N values of the time series (x_i), the network aims to predict the following N values (y_i), referred to as the desired output. Thanks to the framing process, the input and output data are rearranged into $M \times N$ matrices, where $M = n - 2N + 1$, by moving the windows one step ahead until the end of the dataset is reached, thus allowing to obtain M input/desired output pairs.

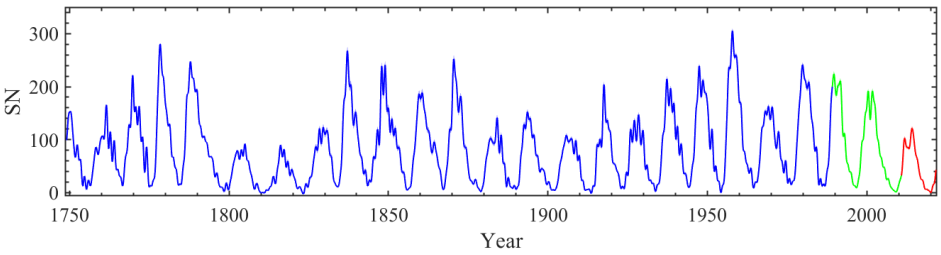

Figure 5.5: The dataset splitting into training-, validation-, and test-set for Train-Test validation technique. The training-set (blue curve) consists of most of the time series ($\sim$ 88%). The validation-set (green curve) covers about the 8% of the data while the test-set length is equal to N ($\sim$ 4% of the whole series, red curve).

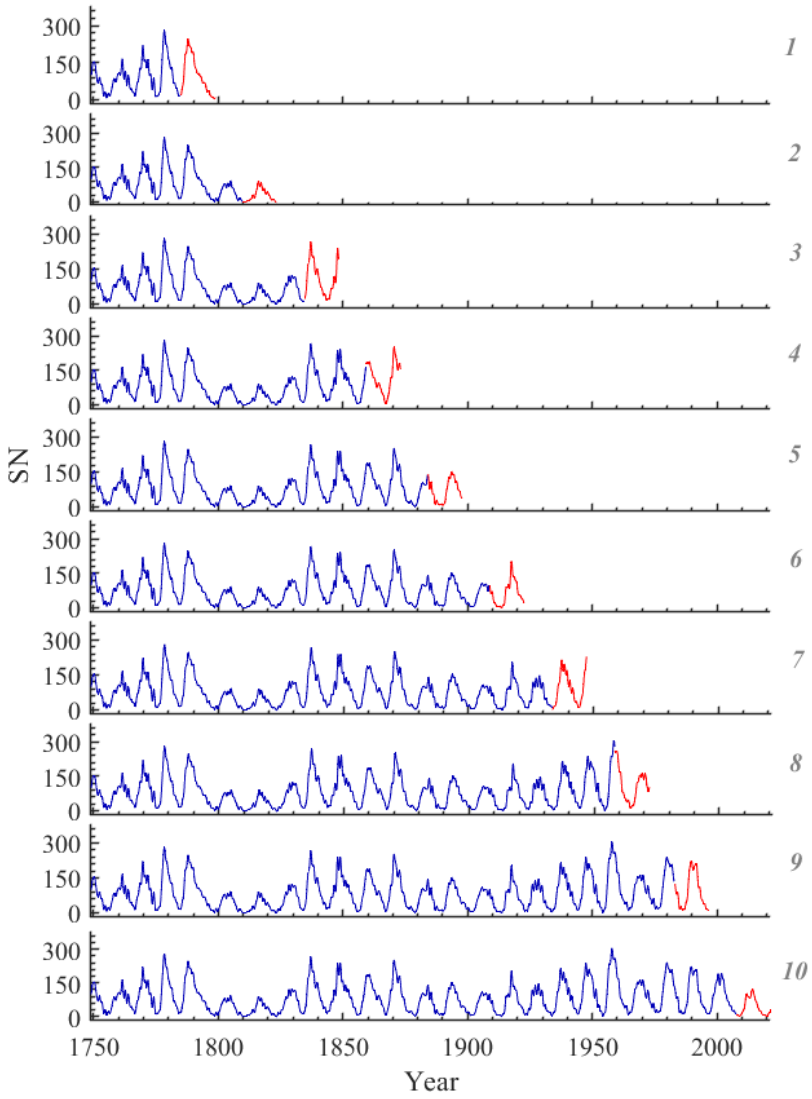

Figure 5.6: Data splitting scheme for the Time Series 10-fold Cross-validation technique. For each of the ten iterations, the training- and test-set are shown in blue and red, respectively. Each test-set is included in the training-set of the next iteration. Each training-set can be interpreted as a superset of those that came before it.

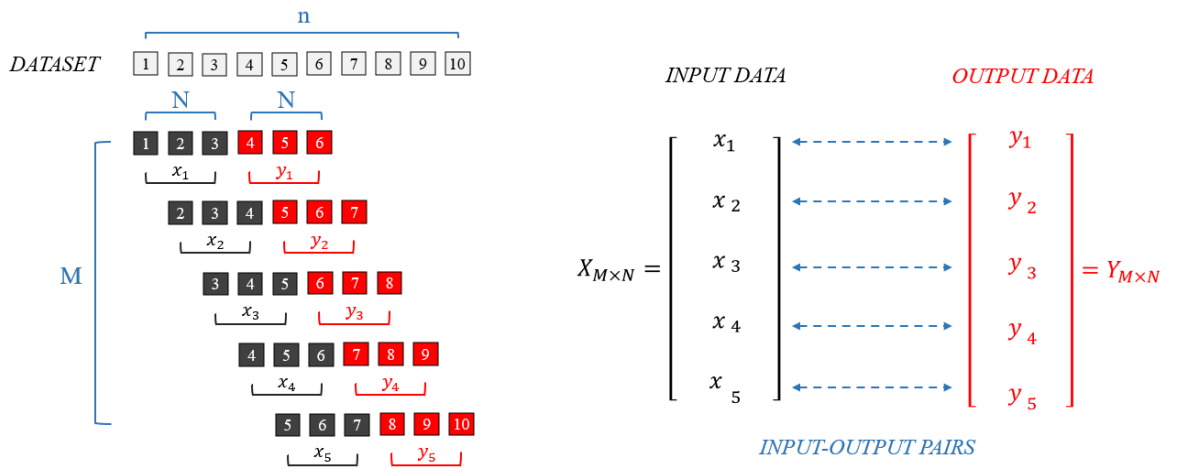

Figure 5.7: Framing of the dataset using a moving window. Giving as input N values of the time series (x_i), the network aims to predict the following N values (y_i). The input and output data are rearranged into M×N matrices, where M = n − 2N + 1, by moving the windows one step ahead until the end of the dataset is reached. This framing allows obtaining M input-output pairs with which the network can be trained.

Regarding the CV model, as a consequence of the network structure, the first N values of the training- and test-set of each iteration are not predicted but only used as input to forecast the following N values. The TT model validation-set is fully predicted, since, to forecast its first N values, the last N values of the training-set are used as input. Similarly, the last N values of the validation-set are used as input to forecast the whole test-set.

The loss function, which has to be minimized to achieve a well-trained network, is the Mean Absolute Error (MAE), defined as follows:

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |y_{i,\mathrm{obs}} - y_{i,\mathrm{pred}}|}{n} \tag{5.12}$$

where $y_{i,\mathrm{obs}}$ is the number of sunspot observed, $y_{i,\mathrm{pred}}$ is the number of sunspots predicted by the neural network and n is the length of the series.

Before training each model, the batch size and the number of hidden nodes (nodes in the hidden layers) have been optimized within the range [2-200] and [5-100], respectively. The chosen parameter combination is the one that gives the lowest MAE after training the network for some epochs. The *Adam* algorithm, an extension based on stochastic gradient descent [210], is chosen to update the weights and optimize the learning rate.

Each model has been trained for fifty epochs. The Keras ModelCheckpoint callback [206] was used to save each time the best model, corresponding to the specific combination of weights which led to the lowest MAE among all the epochs of training. More specifically, the best TT model is the one providing the lowest MAE over the validation-set, while a best CV model is saved for each iteration and corresponds to the one with the lowest MAE over the test-set.

The parameter N was set equal to 130, so the network can predict 130 points in the future (about 11 y), starting from the previous 130. This configuration was chosen according to the following observations:

- the network performance degrades for N > 130, indicating the forecast limit of this approach;
- Solar Cycle 25 started in December 2019 [40] and, given the average duration of solar cycles, 130 months are enough to predict its remaining part together with the next solar minimum.

Before forecasting Solar Cycle 25, the model was trained again using all the available data. Since using the whole series as input does not allow having a validation set available, which is necessary to avoid overfitting, a crucial point concerns the choice of the number of epochs for this further training. For this purpose, both the TT and CV models have been trained on shorter versions of the dataset, and the MAE obtained by predicting known data has been estimated. More in detail, both methods were trained and tested on two different datasets, one obtained by cutting the SN time series at the end of cycle 22 and one with values up to cycle 23. Therefore, predicting the future 130 values of these datasets means predicting cycles 23 and 24, respectively. Future values are forecasted many times by varying the number of epochs, and the final value of this hyperparameter is the one for which the MAE is minimized.

5.3.3 Uncertainty estimation

Deep learning tools for regression do not automatically quantify prediction uncertainty.

In the adopted Bayesian approach, probabilistically removing, or dropping out, inputs to a layer, has the effect of simulating a large number of networks with very different structures and, in turn, making nodes in the network generally more robust to the inputs, thus reducing overfitting and improving model performance.

Following the idea of Gal and Ghahramani [204], the model uncertainty has been estimated using a Monte Carlo dropout (MC dropout): given an input vector x, the output is forecasted with stochastic dropout at each layer K times, thus obtaining y_1, ..., y_K. The final prediction is the sample mean and the relative uncertainty can be approximated by the sample standard deviation. Similarly to Ref.[211], a few hundred predictions have been found to be enough to obtain a stable estimation, and, finally, the K value was set equal to 1000.

In addition to the uncertainty term described above, a further contribution related to the error propagation of the original time series has been introduced. As a first step, this uncertainty was propagated onto the result of the smoothing by a gaussian process regression bootstrapping, as described in Section 5.2.2. The ensemble of 1000 surrogated time series generated for this purpose was fed to the trained neural network, and the same approach has been used once again to propagate the uncertainty to the N forecasted values. In this way, the uncertainty of the SN data was propagated to the final prediction. This error contribution was then summed to the one from the MC dropout that, on the other hand, represents the model uncertainty.

5.4 Results and discussion

The TT model has been trained on the smoothed monthly SN series, with data ranging from January 1749 to March 1989 (about 88% of the dataset). While tuning model hyperparameters, the network skills have been validated over cycles 22 and 23, and its performance has been finally tested over cycle 24. The MAE obtained by the best TT model for the test-set prediction, corresponding to the last N values of the series, is equal to 13.2. About the CV model, the MAEs obtained for test-sets predictions are shown in Table 5.2. Figure 5.8 shows the CV model forecast of the test-set of each iteration (green curve). The grey curve represents the previous data, including both the training-set and the first N values of the test-set (which are never forecasted), while the black curve represents the desired output. In the bottom panel, the forecast of the TT model test-set is also shown (red curve), with the associated uncertainty. It's important to notice that, because of how the CV technique splits the dataset, the test-set has the same length for each iteration, equal to 295 values. Therefore, the CV model test-set forecast consists of $M = 295 - 260 + 1 = 36$ vectors of 130-point length.

The CV model performance improves as the training-set becomes larger and larger, that is from the first to the last iteration.

As already discussed in Section 5.3.2, Solar Cycle 25 was forecasted after training the models using all available data as input for a number of epochs chosen after evaluating the forecasts of known data. Figure 5.9a shows the comparison between the smoothed cycle 23 (black curve) and the TT model forecast of Solar Cycle 23 (red curve) as a function of the epochs the network has been trained again using as input all the values up to cycle 22. In Fig. 5.9c, the same comparison for the CV model forecast (green curve) is shown. Figure 5.9b and 5.9d show the MAE as a function of the number of epochs obtained for the two models, where the coloured stars highlights the most accurate forecasts. Figure 5.10 shows the same results for the forecast of cycle 24, obtained after training the models using as input all the values up to cycle 23.

The best forecasts are shown in Fig. 5.11, in red and green for the TT and CV models, respectively. The associated MAEs are listed in Table 5.3. Forecasts of Solar Cycle 23 of both models (Figs. 5.11a,c) are less accurate than the ones of cycle 24 (Figs. 5.11b,d), mostly because of the exceptional length of the minimum around 2009 [212]. Cycle's length, shape, and maximum value are generally well predicted. This is evident by looking at Table 5.3, where measured and predicted maxima are compared for both cycles. A good agreement between measured and predicted maxima is present both for the amplitude and the time of occurrence.

Looking at the forecasted cycles in Fig. 5.11, is possible to notice that the periodicities of about 1.2 y and 2.7 y (discussed in Section 5.2.3) are not predicted by the two models. While conducting various tests for the best neural network configuration choice, it has been made clear that these components are predicted, together with the decadal cycle, only when overfitting occurs. Furthermore, a good forecast of Solar Cycle 5, which happend to be anomalous with respect to the others, was found to be another sign of overfitting (not shown). Although the developed neural network cannot predict these high-frequency components, solar cycles 23 and 24 are well predicted within the limits of uncertainties (see Table 5.3).

Iteration	MAE
1	71.2
2	49.9
3	31.0
4	42.7
5	21.1
6	23.0
7	19.8
8	18.5
9	19.9
10	15.6

Table 5.2: Mean absolute error obtained for test-set prediction performed by the best model of each CV iteration.

Since the number of epochs for which the lowest MAE was obtained in equal to 10 in almost all the cases, this has been chosen as the number of epochs both models have been trained again using all the available data before forecasting the Solar Cycle 25.

Solar cycle 25 predictions obtained from TT and CV models are shown in Fig. 5.12 in red and green, respectively, together with their uncertainty bars and with Solar Cycles 20 to 24. The forecasts are compatible within the error limits. Both models predict the next solar cycle to be smaller than the previous one, with a peak amplitude some 14% smaller than the peak of the smoothed Solar Cycle 24. The CV model predicts a peak amplitude of 110 $\pm$ 9 on April 2024 ($\pm$ 12 months), while the TT model predicted maximum is of 104 $\pm$ 7 and will occur on July 2024 ($\pm$ 7 months). Both models agree that the next solar minimum will be as deep as the previous one. The minimum predicted by the TT and CV models will occur on June 2030 ($\pm$ 8 months) and on August 2030 ($\pm$ 8 months), respectively. Solar cycle 25 will thus be a relatively normal cycle, with a mean forecast duration of (10.6 $\pm$ 0.7) y.

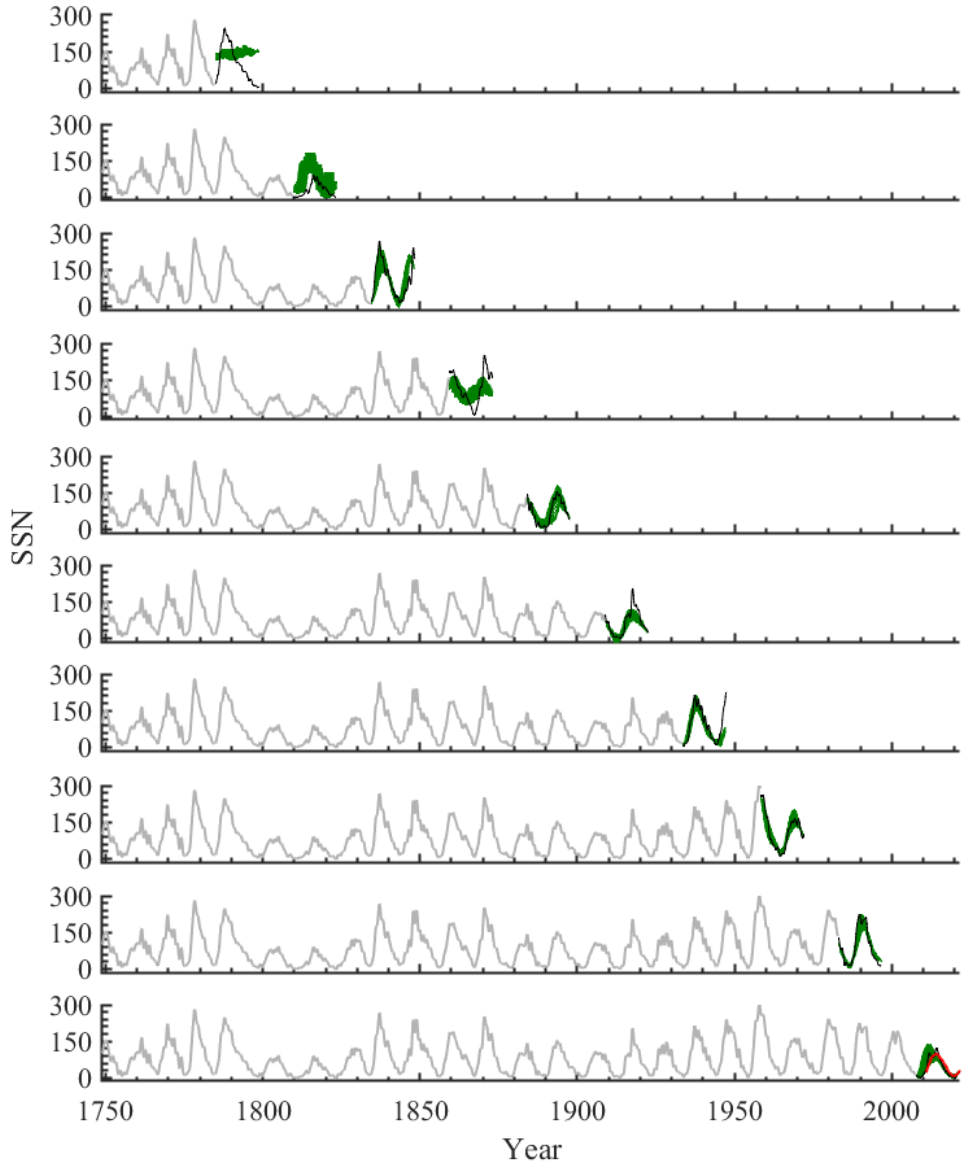

Figure 5.8: Forecast of the test-set of each iteration obtained from the CV model (green curve). The grey curve represents the previous data, including both the training-set and the first N values of the test-set (which are never forecasted), while the black curve represents the desired output. In the bottom panel, the forecast of the TT model test-set is also shown (red curve).

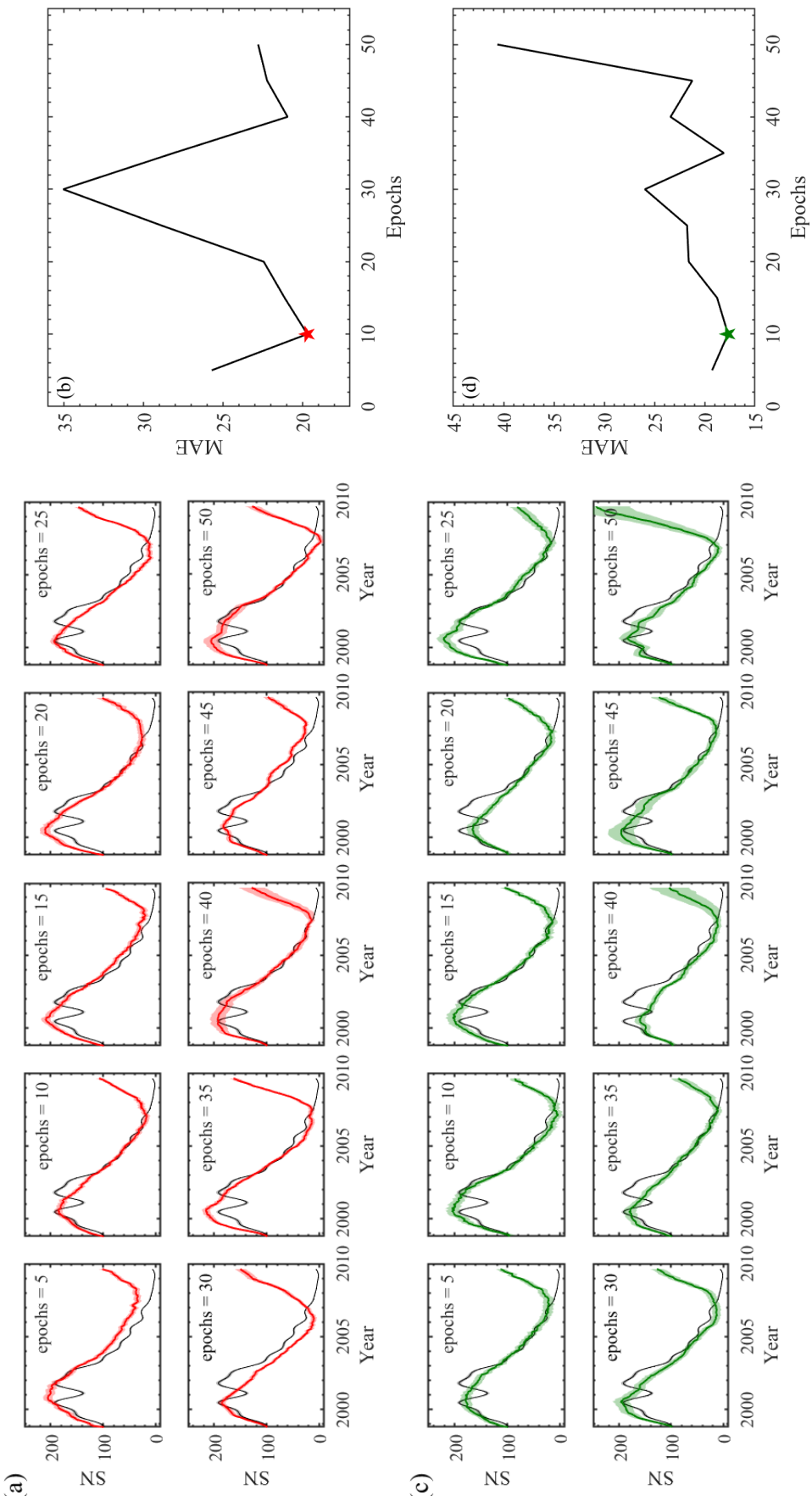

Figure 5.9: Panel a and c show the comparison between the smoothed Solar Cycle 23 (black curve) and the forecasted ones obtained from the TT model (red curve in panel a) and the CV model (green curve in panel c) as a function of the number of epochs the network has been trained again using as input all the values up to cycle 22. The MAEs for the TT and CV models are shown in panel b and d, respectively, as a function of the epochs, where the coloured stars correspond to the lowest MAEs obtained.

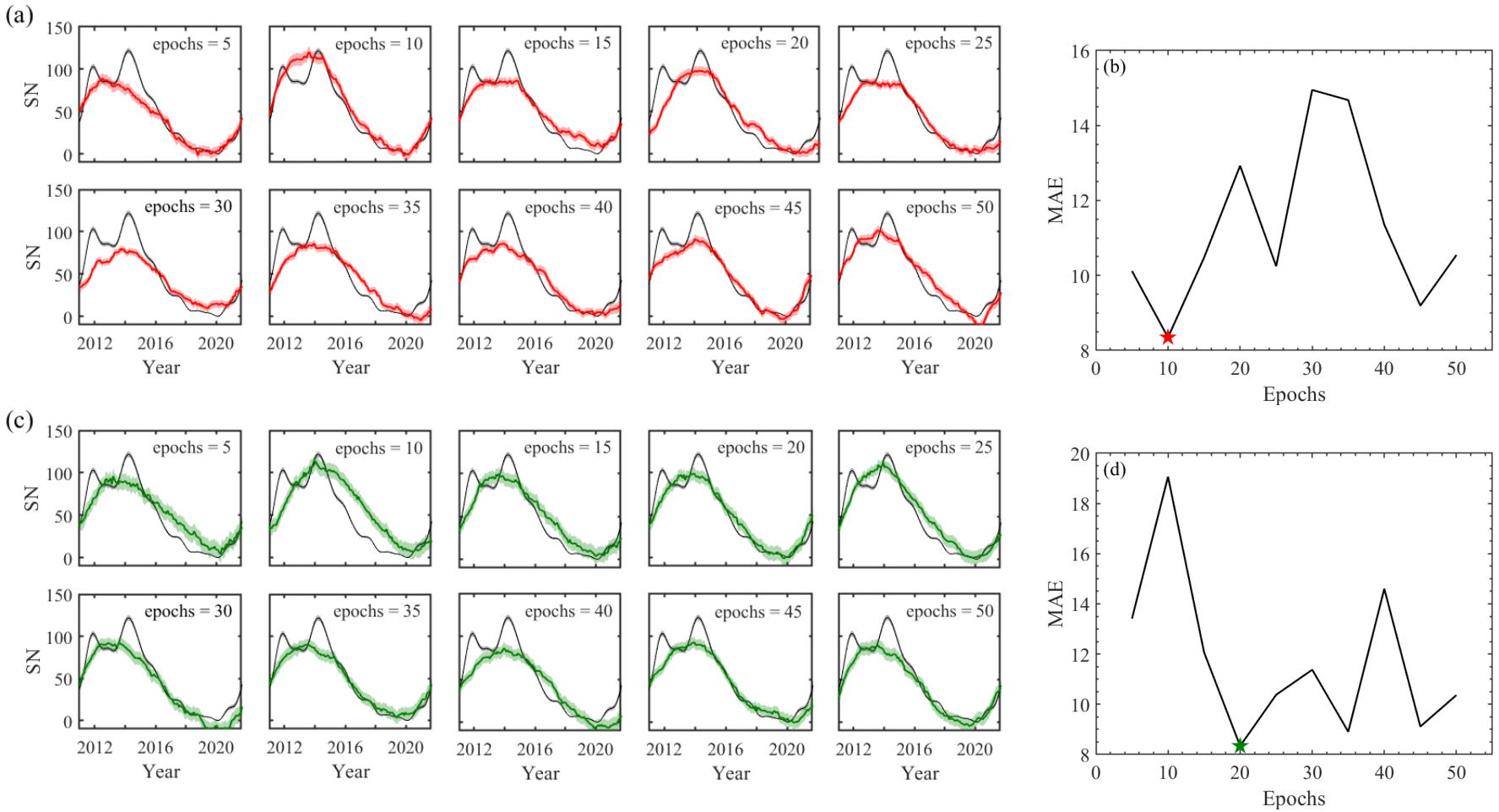

Figure 5.10: Panel a and c show the comparison between the smoothed Solar Cycle 24 (black curve) and the forecasted ones obtained from the TT model (red curve in panel a) and the CV model (green curve in panel c) as a function of the number of epochs the network has been trained again using as input all the values up to cycle 23. The MAEs for the TT and CV models are shown in panel b and d, respectively, as a function of the epochs, where the coloured stars correspond to the lowest MAEs obtained.

MEAN ABSOLUTE ERROR

	Cycle 23	Cycle 24
TT model	19.7	8.4
CV model	17.7	8.3

SOLAR CYCLE MAXIMA

	Cycle 23	Cycle 24
Measured	192 ± 4 June 2000 (± 1 month) 192 ± 5 Oct 2001 (± 1 month)	122 ± 3 Mar 2014 (± 2 months)
TT model	186 ± 7 Aug 2000 (± 5 months)	120 ± 7 Aug 2013 (± 5 months)
CV model	206 ± 12 June 2000 (± 6 months)	99 ± 7 Dec 2013 (± 10 months)

Table 5.3: Performance of the TT and CV models in predicting known cycles. The MAEs obtained for the forecasts of cycles 23 and 24 are shown. Predicted maxima are compared with the measured ones. For Solar Cycle 23, the measured double maxima are both reported.

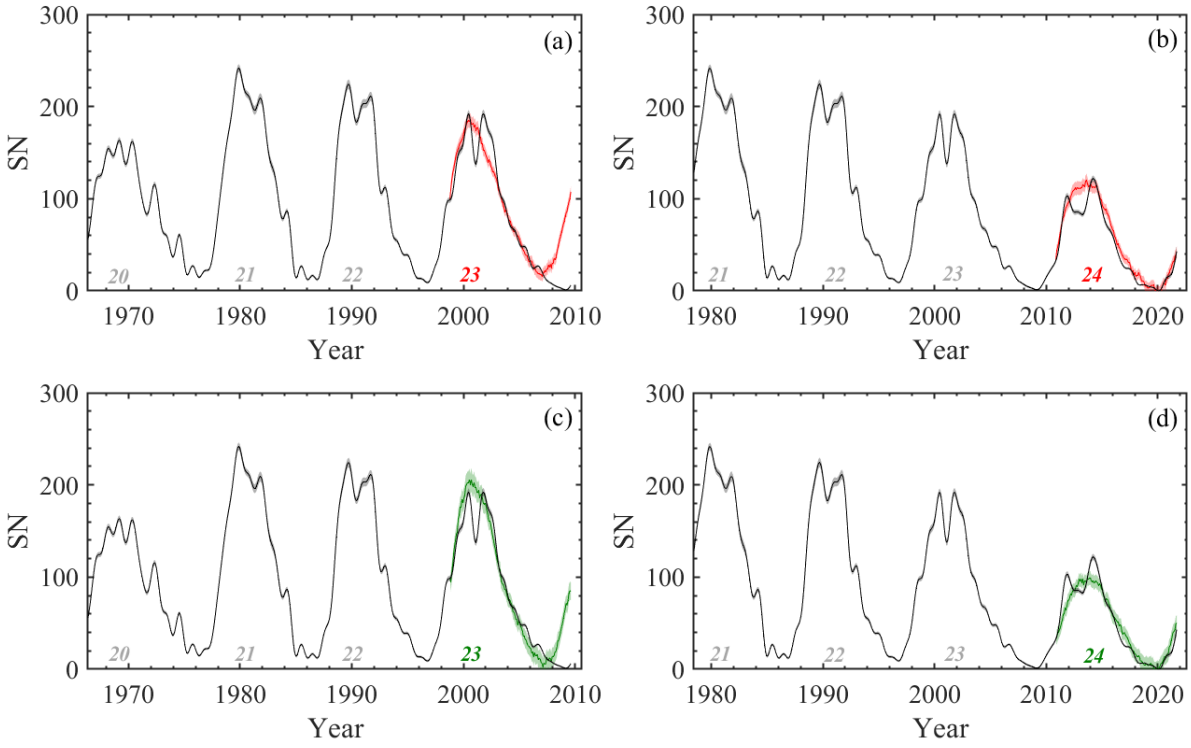

Figure 5.11: Most accurate forecasts of Solar Cycles 23 and 24 performed after training the networks up to cycles 22 and 23, respectively. Red curves represent the TT model forecasts of Solar Cycles 23 (a) and 24 (b). Green curves represent the CV model forecasts of Solar Cycles 23 (c) and 24 (d). All the predictions are shown with the associated uncertainties.

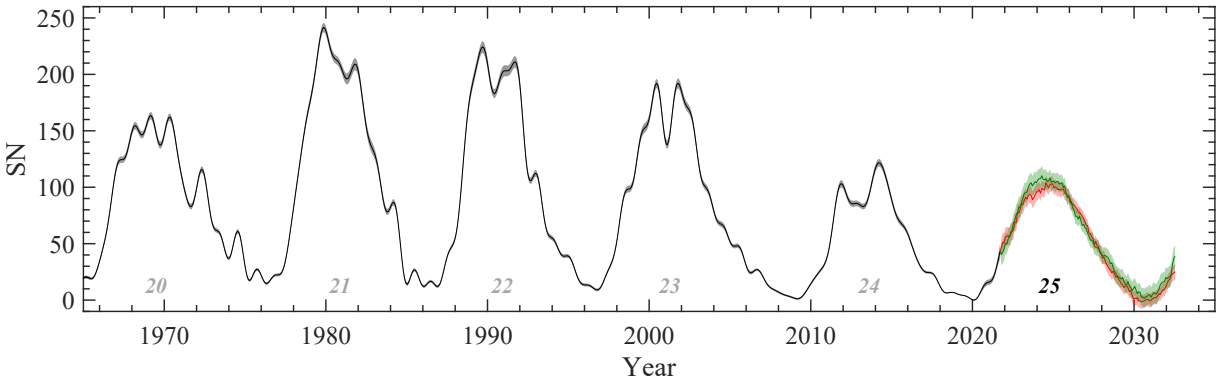

Figure 5.12: Forecasts of Solar Cycle 25 performed by the TT (red curve) and CV models (green curve) together with the smoothed series (black curve) from cycle 20 to 24, with relative uncertainties.

5.4.1 Comparison with other predictions and newly measured data

According to a recent review of Ref.[213], predictions of Solar Cycle 25 based on the solar dynamo framework converge towards a weak or moderate Solar Cycle 25 peaking around 2024-2025. Solar dynamo models driven by observations should represent a standard for forecasting upcoming solar cycles. However, in the above models, it is often implicitly assumed that the dynamical memory of the solar dynamo mechanism is short [see, e.g., Ref. 214], thus allowing for predictions of only the subsequent sunspot cycle. Even worse, numerical simulations with stochastic and deterministic non-linear dynamo models hint at the possibility that it is not even possible to predict the solar cycle using the output from such models [215]. To complicate the problem, although two main proposals exist in the literature for the poloidal field generation - the Babcock-Leighton mechanism [216, 217] and the mean-field α effect [155] - most of the dynamo-based prediction models only consider the former mechanism. As a consequence, the convergence of the physical-based prediction models is probably somewhat biased by this exclusive choice for the underlying dynamo mechanism. Because of the above discussion, it is thus reasonable to consider results obtained from several forecasting techniques not connected with the underlying physics.

Figure 5.13 shows a comparison of various predictions of Solar Cycle 25 obtained by different groups and based on diverse methodologies, taking only into account those works that predicted both the time and the amplitude of the peak. In particular, the comparison is done with physical-based forecasts [156, 167, 172, 173], precursor technique based forecasts [174–176], non-linear model based forecasts [157, 177, 178], statistical forecasts [158–164], spectral methods based forecasts [165, 166], and machine learning/neural network based forecasts [168–171]. It is clear that an agreement between the various forecasts is still lacking, even among predictions made with the same methodology.

Also included is the prediction made by the Solar Cycle 25 Prediction Panel[3], co-chaired by the National Oceanic and Atmospheric Administration (NOAA), the National Aeronautics and Space Administration (NASA), and the International Space Environmental Services (ISES), according to which Solar Cycle 25 will reach a peak in July 2025 ($\pm$ 8 months), with a maximum smoothed SN of 115 ± 10, being thus quite similar to Solar Cycle 24. The amplitudes of Solar Cycle 25 forecasted by the TT and CV models lie within the middle range of the distribution, and their uncertainties are among the smallest.

[3] https://www.swpc.noaa.gov/products/solar-cycle-progression

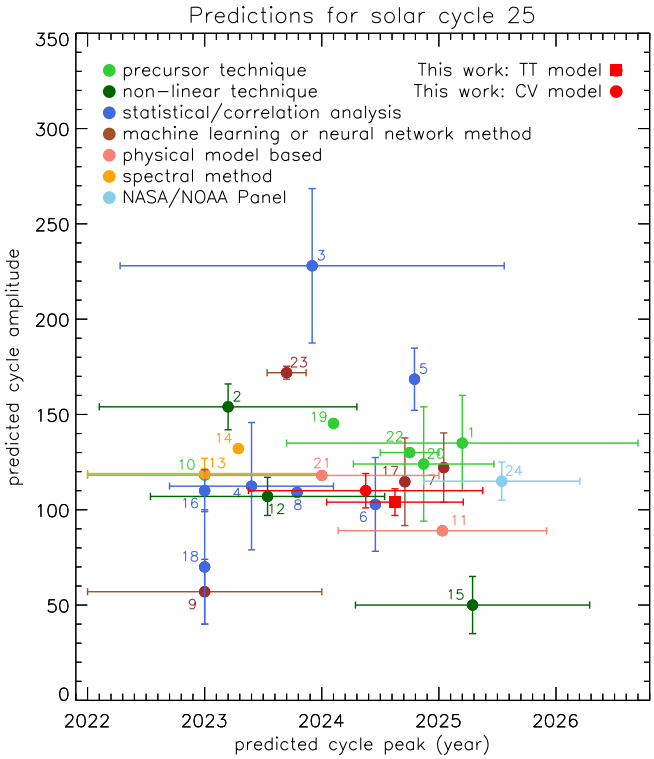

Figure 5.13: Predictions of Solar Cycle 25 by different groups based on diverse methodologies indicated in the legend. References are: [174][1], [177][2], [158][3], [159][4], [162][5], [161][6], [168][7], [160][8], [169][9], [175][10], [167][11], [178][12], [165][13], [166][14], [157][15], [163][16], [170][17], [164][18], [172][19], [173][20], [156][21], [176][22], [171][23], NOAA-NASA-ISES prediction[24].

As discussed in Section 5.2.1, the developed Bayesian networks have been trained using the SN time series with monthly values from the beginning of 1749 up to September 2021. At the current time, new data have been added to the SILSO dataset, therefore is possible to directly compare the obtained forecasts with measured data. Figure 5.14 shows the comparison between the smoothed monthly SN time series (magenta curve), obtained after including new data up to September 2022, and the forecasts of Solar Cycle 25, obtained by the TT model (red curve) and the CV model (green curve), where the shaded colored bars represent the uncertainties. The smoothed series used to train the models is also shown (black curve).

A good agreement is present between forecasted and newly measured data, within the limits of uncertainties. Up to the beginning of 2022, measured data fall within one standard deviation of both model's forecasts. After that period, a better agreement is present between the measured data and the CV model forecast. However, both models fully agree with measured data within three standard deviations (not shown).

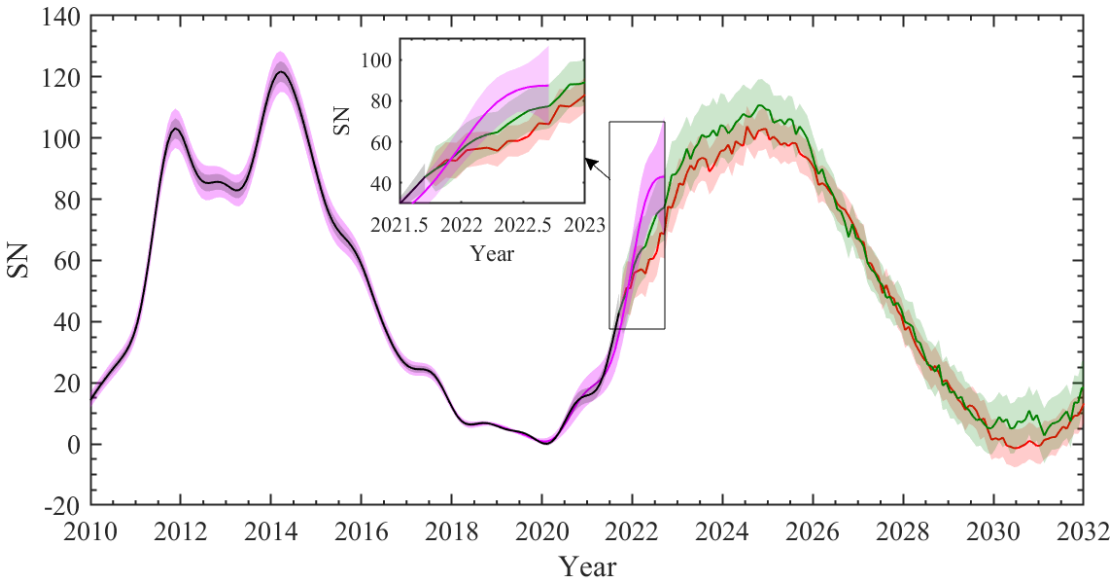

Figure 5.14: Comparison between the forecasts of Solar Cycle 25, obtained by the TT (red curve) and CV models (green curve), and the smoothed monthly SN time series (magenta curve), obtained after including new data up to September 2022. The smoothed series used in this work is also shown (black curve). Data uncertainty, represented by the shaded colored bars, represents ± one standard deviation.

5.4.2 Conclusions

Forecasting solar activity variability is crucial for predicting the impact of the Sun on both Earth's magnetosphere (and ionosphere) and terrestrial climate. Several predictions of the amplitude of Solar Cycle 25 have been recently published, obtained using physical-based models and neural network approaches. The underlying complexity of the dynamo mechanism at the origin of the cyclic nature of the Sun's activity is such that predictions offered by state-of-the-art machine learning algorithms can represent useful tools for cycle progression understanding and forecasting. Among them, Bayesian deep learning provides significant improvements in both accuracy and uncertainty quantification compared to classical techniques. The forecast of the complete profile of Solar Cycle 25 was obtained by applying a deep neural network model on sunspot number data. A refined Bayesian approach has been developed to quantify forecast uncertainty. Two different validation techniques have been implemented, showing fully comparable results. The forecasted Solar Cycle 25 has a peak amplitude of about 14% smaller than cycle 24, and the maximum activity is expected in the middle of the year 2024. The predicted peak amplitude lies within the middle range of the predictions found in the recent literature and is in agreement with the forecast presented by the Solar Cycle Prediction Panel representing NOAA, NASA, and ISES. A good agreement was found by comparing predicted values with newly measured data. The next solar minimum will occur around July 2030 and will be as deep as the previous one, thus suggesting that Solar Cycle 25 will last (10.6 ± 0.7) y. Following the decreasing trend of the last four solar cycles, speculations about a new imminent grand minimum, similar to the Maunder one (1645–1715), have been recently proposed [218, 219].

Bibliography

[1] C. Taricco et al. “Exploration of the stratosphere with cosmic-ray muons detected underground”. In: *Physical Review Research* 4.2 (2022). DOI: https://doi.org/10.1103/PhysRevResearch.4.023226.

[2] T.K. Gaisser, R. Engel, and E. Resconi. *Cosmic rays and particle physics*. Cambridge University Press, 2016. DOI: https://doi.org/10.1017/CBO9781139192194.

[3] P. Lipari. “Lepton spectra in the earth’s atmosphere”. In: *Astroparticle Physics* 1.2 (1993), pp. 195–227. DOI: https://doi.org/10.1016/0927-6505(93)90022-6.

[4] K. Munakata et al. “Large-scale anisotropy of the cosmic-ray muon flux in Kamiokande”. In: *Physical Review D* 56.1 (1997), p. 23. DOI: https://doi.org/10.1103/PhysRevD.56.23.

[5] M. Ambrosio et al. “Search for the sidereal and solar diurnal modulations in the total MACRO muon data set”. In: *Physical Review D* 67.4 (2003). DOI: https://doi.org/10.1103/PhysRevD.67.042002.

[6] N.Y. Agafonova et al. “Implication for the core-collapse supernova rate from 21 years of data of the large volume detector”. In: *The Astrophysical Journal* 802.1 (2015), p. 47. DOI: https://doi.org/10.48550/arXiv.1411.1709.

[7] H. Akaike. “Fitting autoregressive models for prediction”. In: *Annals of the institute of Statistical Mathematics* 21.1 (1969), pp. 243–247. DOI: https://doi.org/10.1007/BF02532251.

[8] P.H. Barrett et al. “Interpretation of cosmic-ray measurements far underground”. In: *Reviews of Modern Physics* 24.3 (1952), p. 133. DOI: https://doi.org/10.1103/RevModPhys.24.133.

[9] P.H. Barrett et al. “Atmospheric temperature effect for mesons far underground”. In: *Physical Review* 95.6 (1954), p. 1573. DOI: https://doi.org/10.1103/PhysRev.95.1573.

[10] N.Y. Agafonova et al. “Characterization of the varying flux of atmospheric muons measured with the Large Volume Detector for 24 years”. In: *Physical Review D* 100.6 (2019). DOI: https://doi.org/10.48550/arXiv.1909.04579.

[11] D.P. Dee et al. “The ERA-Interim reanalysis: Configuration and performance of the data assimilation system”. In: *Quarterly Journal of the royal meteorological society* 137.656 (2011), pp. 553–597. DOI: https://doi.org/10.1002/qj.828.

[12] C. Torrence and G.P. Compo. “A Practical Guide to Wavelet Analysis.” In: *Bulletin of the American Meteorological Society* 79.1 (1998), pp. 61–78. DOI: https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2.

[13] A. Grinsted, J.C. Moore, and S. Jevrejeva. “Application of the cross wavelet transform and wavelet coherence to geophysical time series”. In: *Nonlinear processes in geophysics* 11.5/6 (2004), pp. 561–566. DOI: 10.5194/npg-11-561-2004.

[14] R. Vautard and M. Ghil. "Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series". In: *Physica D: Nonlinear Phenomena* 35.3 (1989), pp. 395–424. DOI: https://doi.org/10.1016/0167-2789(89)90077-8.

[15] R. Vautard, P. Yiou, and M. Ghil. "Singular-spectrum analysis: A toolkit for short, noisy chaotic signals". In: *Physica D: Nonlinear Phenomena* 58.1-4 (1992), pp. 95–126. DOI: https://doi.org/10.1016/0167-2789(92)90103-T.

[16] M. Ghil et al. "Advanced spectral methods for climatic time series". In: *Reviews of geophysics* 40.1 (2002), pp. 3-1–3-41. DOI: https://doi.org/10.1029/2000RG000092.

[17] M.R. Allen and L.A. Smith. "Monte Carlo SSA: Detecting irregular oscillations in the presence of colored noise". In: *Journal of climate* 9.12 (1996), pp. 3373–3404. DOI: 10.1175/1520-0442(1996)009<3373:MCSDIO>2.0.CO;2.

[18] P. Adamson et al. "Observation of muon intensity variations by season with the MINOS far detector". In: *Physical Review D* 81.1 (2010). DOI: https://doi.org/10.48550/arXiv.0909.4012.

[19] A.K. Steiner et al. "Observed temperature changes in the troposphere and stratosphere from 1979 to 2018". In: *Journal of Climate* 33.19 (2020), pp. 8165–8194. DOI: https://doi.org/10.1175/JCLI-D-19-0998.1.

[20] A. Simmons et al. *Global stratospheric temperature bias and other stratospheric aspects of ERA5 and ERA5.1*. 2020. DOI: 10.21957/rcxqfmg0.

[21] V.B. Petkov et al. "Temperature variations of high energy muon flux". In: *International Cosmic Ray Conference*. Vol. 11. 2011, p. 384. DOI: 10.7529/ICRC2011/V11/0685.

[22] L.J. Gray et al. "A lagged response to the 11 year solar cycle in observed winter Atlantic/European weather patterns". In: *Journal of Geophysical Research: Atmospheres* 118.24 (2013), pp. 13–405. DOI: https://doi.org/10.1002/2013JD020062.

[23] T.H.A. Frame and L.J. Gray. "The 11-yr solar cycle in ERA-40 data: An update to 2008". In: *Journal of Climate* 23.8 (2010), pp. 2213–2222. DOI: https://doi.org/10.1175/2009JCLI3150.1.

[24] M. Agostini et al. "Modulations of the cosmic muon signal in ten years of Borexino data". In: *Journal of Cosmology and Astroparticle Physics* 2019.02 (2019), p. 046. DOI: https://doi.org/10.48550/arXiv.1808.04207.

[25] H. Lu et al. "Geomagnetic perturbations on stratospheric circulation in late winter and spring". In: *Journal of Geophysical Research: Atmospheres* 113.D16 (2008). DOI: https://doi.org/10.1029/2007JD008915.

[26] A. Seppälä et al. "Geomagnetic activity signatures in wintertime stratosphere wind, temperature, and wave response". In: *Journal of Geophysical Research: Atmospheres* 118.5 (2013), pp. 2169–2183. DOI: https://doi.org/10.1002/jgrd.50236.

[27] J. Abraham et al. "Properties and performance of the prototype instrument for the Pierre Auger Observatory". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 523.1-2 (2004), pp. 50–95. DOI: https://doi.org/10.1016/j.nima.2003.12.012.

[28] J. Abraham et al. "The fluorescence detector of the Pierre Auger Observatory". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 620.2-3 (2010), pp. 227–251. DOI: https://doi.org/10.48550/arXiv.0907.4282.

[29] Pierre Auger Collaboration. "The Pierre Auger Observatory scaler mode for the study of solar activity modulation of galactic cosmic rays". In: *Journal of Instrumentation* 6.01 (2011). DOI: 10.1088/1748-0221/6/01/P01003.

[30] C. Morello, L. Periale, and G. Navarra. "A search for high-energy cosmic gamma-ray bursts". In: *Il Nuovo Cimento C* 7 (1984), pp. 682–688. DOI: https://doi.org/10.1007/BF02573794.

[31] X. Bertou. "Search for Gamma Ray Bursts using the single particle technique at the Pierre Auger Observatory". In: *Proceedings of the 30th International Cosmic Ray Conference (ICRC 2007)*. 2007. DOI: https://doi.org/10.48550/arXiv.0706.1256.

[32] Pierre Auger collaboration, D. Thomas, et al. "Search for coincidences with astrophysical transients in Pierre Auger Observatory data". In: *Proceedings of the 31st International Cosmic Ray Conference (ICRC 2009)*. 2009.

[33] S. Dasso, H. Asorey, Pierre Auger Collaboration, et al. "The scaler mode in the Pierre Auger Observatory to study heliospheric modulation of cosmic rays". In: *Advances in space research* 49.11 (2012), pp. 1563–1569. DOI: https://doi.org/10.48550/arXiv.1204.6196.

[34] M.L. Schimassek. "Extending the Physics Reach of the Pierre Auger Observatory using Low-Level Trigger Data". PhD thesis. Karlsruher Institut für Technologie (KIT), 2022. DOI: 10.5445/IR/1000149933.

[35] M.L. Schimassek, D. Veberič, and R. Engel. "A New Scaler Analysis Framework". In: *GAP-Note 2019-060* (2019).

[36] P. Kumar and E. Foufoula-Georgiou. "Wavelet analysis for geophysical applications". In: *Reviews of geophysics* 35.4 (1997), pp. 385–412. DOI: https://doi.org/10.1029/97RG00427.

[37] Y.P. Singh. "Study of the solar rotational period and its harmonics in solar activity, interplanetary, geomagnetic, and cosmic ray intensity indicators during solar polarity reversal periods". In: *Solar Physics* 294.3 (2019). DOI: 10.1007/s11207-019-1413-y.

[38] A. López-Comazzi and J.J. Blanco. "Short-and Mid-term Periodicities Observed in Neutron Monitor Counting Rates throughout Solar Cycles 20–24". In: *The Astrophysical Journal* 927.2 (2022), p. 155. DOI: 10.3847/1538-4357/ac4e19.

[39] A. López-Comazzi and J.J. Blanco. "Short-term periodicities observed in neutron monitor counting rates". In: *Solar Physics* 295.6 (2020), p. 81. DOI: 10.1007/s11207-020-01649-5.

[40] SILSO World Data Center. "The International Sunspot Number". In: *International Sunspot Number Monthly Bulletin and online catalogue* (1749-2021). URL: http://www.sidc.be/silso/.

[41] E. Ross and W.J. Chaplin. "The behaviour of galactic cosmic-ray intensity during solar activity cycle 24". In: *Solar physics* 294 (2019), pp. 1–17. DOI: 10.1007/s11207-019-1397-7.

[42] M. Singh, Y.P. Singh, et al. "Solar modulation of galactic cosmic rays during the last five solar cycles". In: *Journal of atmospheric and solar-terrestrial physics* 70.1 (2008), pp. 169–183. DOI: 10.1016/j.jastp.2007.10.001.

[43] P. Janardhan et al. "Solar cycle 24: An unusual polar field reversal". In: *Astronomy and Astrophysics* 618 (2018). DOI: https://doi.org/10.1051/0004-6361/201832981.

[44] N. Karna, S.A. Hess Webber, and W.D. Pesnell. "Using polar coronal hole area measurements to determine the solar polar magnetic field reversal in solar cycle 24". In: *Solar Physics* 289 (2014), pp. 3381–3390. DOI: 10.1007/s11207-014-0541-7.

[45] X. Sun et al. "On polar magnetic field reversal and surface flux transport during solar cycle 24". In: *The Astrophysical Journal* 798.2 (2015), p. 114. DOI: 10.1088/0004-637X/798/2/114.

[46] G. Kopp. "Magnitudes and timescales of total solar irradiance variability". In: *Journal of space weather and space climate* 6 (2016). DOI: https://doi.org/10.1051/swsc/2016025.

[47] G. Kopp. *SORCE Level 3 Total Solar Irradiance Daily Means V019*. Greenbelt, MD, USA, 2019. DOI: 10.5067/7C82ZHS0OPFR.

[48] G. Kopp, G. Lawrence, and G. Rottman. "The total irradiance monitor (TIM): science results". In: *The Solar Radiation and Climate Experiment (SORCE) Mission Description and Early Results* (2005), pp. 129–139.

[49] E. Rieger et al. "A 154-day periodicity in the occurrence of hard solar flares?" In: *Nature* 312.5995 (1984), pp. 623–625. DOI: https://doi.org/10.1038/312623a0.

[50] M. Dimitropoulou, X. Moussas, and D. Strintzi. "Enhanced Rieger-type periodicities' detection in X-ray solar flares and statistical validation of Rossby waves' existence". In: *Monthly Notices of the Royal Astronomical Society* 386.4 (2008), pp. 2278–2284. DOI: 10.1111/j.1365-2966.2008.13203.x.

[51] M. Carbonell and J.L. Ballester. "A short-term periodicity near 155 day in sunspot areas". In: *Astronomy and Astrophysics* 238 (1990), pp. 377–381.

[52] J.L. Ballester, R. Oliver, and F. Baudin. "Discovery of the near 158 day periodicity in group sunspot numbers during the eighteenth century". In: *The Astrophysical Journal* 522.2 (1999), pp. L153–L156. DOI: 10.1086/312232.

[53] T. Bai and E.W. Cliver. *A 154-day periodicity in the occurrence rate of proton flares*. Tech. rep. STANFORD UNIV CA, 1990. DOI: 10.1086/169342.

[54] VV Lobzin, Iver H Cairns, and PA Robinson. "Rieger-type periodicity in the occurrence of solar type III radio bursts". In: *The Astrophysical Journal Letters* 754.2 (2012), pp. L28–L32. DOI: 10.1088/2041-8205/754/2/L28.

[55] B.R. Dennis. "Solar hard X-ray bursts". In: *Solar physics* 100 (1985), pp. 465–490. DOI: 10.1007/BF00158441.

[56] T. Bai and P.A. Sturrock. "The 152-day periodicity of the solar flare occurrence rate". In: *Nature* 327.6123 (1987), pp. 601–604. DOI: https://doi.org/10.1038/327601a0.

[57] J.N. Kile and E.W. Cliver. "A search for the 154 day periodicity in the occurrence rate of solar flares using Ottawa 2.8 GHz burst data, 1955-1990". In: *The Astrophysical Journal* 370 (1991), pp. 442–448. DOI: 10.1086/169831.

[58] J. Lean. "Evolution of the 155 day periodicity in sunspot areas during solar cycles 12 to 21". In: *The Astrophysical Journal* 363 (1990), pp. 718–727. DOI: 10.1086/169378.

[59] M. Carbonell and J.L. Ballester. "The periodic behaviour of solar activity-The near 155-day periodicity in sunspot areas". In: *Astronomy and Astrophysics* 255 (1992), pp. 350–362.

[60] R. Oliver, J.L. Ballester, and F. Baudin. "Emergence of magnetic flux on the Sun as the cause of a 158-day periodicity in sunspot areas". In: *Nature* 394.6693 (1998), pp. 552–553. DOI: https://doi.org/10.1038/29012.

[61] J.L. Lean and G.E. Brueckner. "Intermediate-term solar periodicities-100-500 days". In: *The Astrophysical Journal* 337 (1989), pp. 568–578. DOI: 10.1086/167124.

[62] W. Droege et al. "A 153 day periodicity in the occurrence of solar flares producing energetic interplanetary electrons". In: *The Astrophysical Journal Supplement Series* 73 (1990), pp. 279–283. DOI: 10.1086/191463.

[63] V.K. Verma et al. "Search for a 152-158 days periodicity in the occurrence rate of solar flares inferred from spectral data of radio bursts". In: *Astronomy and Astrophysics Supplement Series* 90 (1991), pp. 83–87.

[64] T.V. Zaqarashvili et al. "Magnetic Rossby waves in the solar tachocline and Rieger-type periodicities". In: *The Astrophysical Journal* 709.2 (2010), p. 749. DOI: 10.1088/0004-637X/709/2/749.

[65] K. Ichimoto et al. "Periodic behaviour of solar flare activity". In: *Nature* 316.6027 (1985), pp. 422–424.

[66] T. Bai and P.A. Sturrock. "The 154-day and related periodicities of solar activity as subharmonics of a fundamental period". In: *Nature* 350.6314 (1991), pp. 141–143. DOI: 10.1038/350141a0.

[67] P.A. Sturrock et al. "Indications of r-mode oscillations in SOHO/MDI solar radius measurements". In: *The Astrophysical Journal* 804.1 (2015), p. 47. DOI: 10.1088/0004-637X/804/1/47.

[68] P.A. Sturrock et al. "An analysis of apparent r-mode oscillations in solar activity, the solar diameter, the solar neutrino flux, and nuclear decay rates, with implications concerning the Sun's internal structure and rotation, and neutrino processes". In: *Astroparticle Physics* 42 (2013), pp. 62–69. DOI: https://doi.org/10.1016/j.astropartphys.2012.11.011.

[69] Y.-Q. Lou. "Rossby-type wave-induced periodicities in flare activities and sunspot areas or groups during solar maxima". In: *The Astrophysical Journal* 540.2 (2000), p. 1102. DOI: 10.1086/309387.

[70] E. Gurgenashvili et al. "Rieger-type periodicity during solar cycles 14–24: estimation of dynamo magnetic field strength in the solar interior". In: *The Astrophysical Journal* 826.1 (2016), p. 55. DOI: 10.3847/0004-637X/826/1/55.

[71] P.K.F. Grieder. *Cosmic rays at Earth*. Elsevier, 2001. DOI: https://doi.org/10.1016/B978-0-444-50710-5.X5000-3.

[72] R. Modzelewska and A. Gil. "Recurrence of galactic cosmic-ray intensity and anisotropy in solar minima 23/24 and 24/25 observed by ACE/CRIS, STEREO, SOHO/EPHIN and neutron monitors-Fourier and wavelet analysis". In: *Astronomy and Astrophysics* 646 (2021). DOI: https://doi.org/10.1051/0004-6361/202039651.

[73] G.A. Bazilevskaya. "Observations of variability in cosmic rays". In: *Space Science Reviews* 94.1-2 (2000), pp. 25–38. DOI: 10.1023/A:1026721912992.

[74] R. Syiemlieh et al. "Analyzing Dominant 13.5 and 27 day Periods of Solar Terrestrial Interaction: A New Insight into Solar Cycle Activities". In: *Research in Astronomy and Astrophysics* 22.8 (2022), p. 085005. DOI: 10.1088/1674-4527/ac7129.

[75] J. Beer, K. McCracken, and R. Steiger. *Cosmogenic radionuclides: theory and applications in the terrestrial and space environments*. Springer Science and Business Media, 2012. DOI: 10.1007/978-3-642-14651-0.

[76] I. Bizzarri et al. "Cosmogenic radionuclides in the Cavezzo meteorite: Gamma-ray measurement and detection efficiency simulations". In: *Applied Radiation and Isotopes* (2023). DOI: https://doi.org/10.1016/j.apradiso.2023.110651.

[77] D. Gardiol, A. Cellino, and M. Di Martino. "PRISMA, Italian network for meteors and atmospheric studies". In: *International Meteor Conference Egmond, the Netherlands, 2-5 June 2016*. Ed. by A. Roggemans and P. Roggemans. 2016, p. 76.

[78] D. Gardiol. "News from the Italian PRISMA fireball network". In: *International Meteor Conference, Pezinok-Modra, Slovakia*. 2019, pp. 81–86.

[79] F. Colas et al. "FRIPON: a worldwide network to track incoming meteoroids". In: *Astronomy and Astrophysics* 644 (2020). DOI: https://doi.org/10.1051/0004-6361/202038649.

[80] D. Gardiol, A. Barghini D. Buzzoni, A. Carbognani, et al. “Cavezzo, the first Italian meteorite recovered by the PRISMA fireball network. Orbit, trajectory, and strewn-field”. In: *Monthly Notices of the Royal Astronomical Society* 501.1 (2021), pp. 1215–1227. DOI: https://doi.org/10.1093/mnras/staa3646.

[81] G. Pratesi et al. “Cavezzo—The double face of a meteorite: Mineralogy, petrography, and geochemistry of a very unusual chondrite”. In: *Meteoritics and Planetary Science* 56.6 (2021), pp. 1125–1150. DOI: https://doi.org/10.1111/maps.13695.

[82] C. Taricco et al. “Galactic cosmic ray flux decline and periodicities in the interplanetary space during the last 3 centuries revealed by 44Ti in meteorites”. In: *Journal of Geophysical Research: Space Physics* 111.A8 (2006). DOI: https://doi.org/10.1029/2005JA01145.

[83] C. Taricco et al. “Mid 19th century minimum of galactic cosmic ray flux inferred from ^{44}Ti in Allegan meteorite”. In: *Advances in Space Research* 41.2 (2008), pp. 275–279. DOI: https://doi.org/10.1016/j.asr.2007.06.050.

[84] P. Colombetti et al. “Experimental set-up for gamma-activity measurements of astromaterials”. In: *2008 IEEE Nuclear Science Symposium Conference Record*. IEEE. 2008, pp. 1802–1805. DOI: 10.1109/NSSMIC.2008.4774742.

[85] P. Colombetti et al. “Low γ activity measurement of meteorites using HPGe–NaI detector system”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 718 (2013), pp. 140–142. DOI: https://doi.org/10.1016/j.nima.2012.07.053.

[86] C. Taricco et al. “Early 18th century cosmic ray flux inferred from ^{44}Ti in Agen meteorite”. In: *Astrophysics and Space Science* 361.10 (2016), pp. 1–5. DOI: https://doi.org/10.1007/s10509-016-2909-7.

[87] P. Colombetti. “Measurement of Cosmogenic Radionuclides in Meteorites by Gamma-ray Spectrometry: Heliospheric Modulation of Cosmic Rays over the Last Three Centuries”. PhD thesis. University of Turin, 2009.

[88] R. Michel and S. Neumann. “Interpretation of cosmogenic nuclides in meteorites on the basis of accelerator experiments and physical model calculations”. In: *Proceedings of the Indian Academy of Sciences-Earth and Planetary Sciences* 107 (1998), pp. 441–457. DOI: https://doi.org/10.1007/BF02841610.

[89] C. Taricco et al. “Experimental set-up and optimization of a gamma-ray spectrometer for measurement of cosmogenic radionuclides in meteorites”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 572.1 (2007), pp. 241–243. DOI: https://doi.org/10.1016/j.nima.2006.10.359.

[90] D. Gardiol et al. “Improvement of the extraction method of faint signals in γ-activity measurements of meteorites”. In: *The European Physical Journal Plus* 132.6 (2017), pp. 1–9. DOI: https://doi.org/10.1140/epjp/i2017-11556-y.

[91] A. Simonits et al. “HyperLab: A new concept in gamma-ray spectrum analysis”. In: *Journal of Radioanalytical and Nuclear Chemistry* 257.3 (2003), pp. 589–595. DOI: https://doi.org/10.1023/A:1025400917620.

[92] S. Agostinelli et al. “GEANT4—a simulation toolkit”. In: *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8.

[93] J. Allison et al. “Geant4 developments and applications”. In: *IEEE Transactions on nuclear science* 53.1 (2006), pp. 270–278. DOI: 10.1109/TNS.2006.869826.

[94] J. Allison et al. "Recent developments in Geant4". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 835 (2016), pp. 186–225. DOI: https://doi.org/10.1016/j.nima.2016.06.125.

[95] A.L. Graham. "Meteoritical Bulletin, No. 55". In: *Meteoritics* 13 (1978).

[96] S.P. Das Gupta et al. "The Dhajala meteorite 1". In: *Mineralogical Magazine* 42.324 (1978), pp. 493–497. DOI: https://doi.org/10.1180/minmag.1978.042.324.12.

[97] P.J. Cressy Jr. *Cosmogenic radionuclides in stone meteorites*. Tech. rep. 1976.

[98] G. Bonino et al. "Solar and galactic cosmic-ray records of the Fermo (H) chondrite regolith breccia". In: *Meteoritics and Planetary Science* 36.6 (2001), pp. 831–839. DOI: https://doi.org/10.1111/j.1945-5100.2001.tb01920.x.

[99] G. Bonino et al. "Cosmogenic radionuclides in four fragments of the Portales Valley meteorite shower: Influence of different element abundances and shielding". In: *Advances in Space Research* 29.5 (2002), pp. 777–782. DOI: https://doi.org/10.1016/S0273-1177(02)00011-X.

[100] J.F. Wacker et al. "The Juancheng and El Paso Superbolides of February 15 and October 9, 1997: Preatmospheric Meteoroid Sizes". In: *Meteoritics and Planetary Science* 33 (1998), p. 160.

[101] H. Neder, M. Laubenstein, and G. Heusser. "Radionuclide concentrations in the freshly fallen meteorite Morávka". In: *Meteoritics and Planetary Science* 36 (2001), p. 146.

[102] S.V.S. Murty et al. "Devgaon (H3) chondrite: Classification and complex cosmic ray exposure history". In: *Meteoritics and Planetary Science* 39.3 (2004), pp. 387–399. DOI: https://doi.org/10.1111/j.1945-5100.2004.tb00100.x.

[103] W.U. Reimold et al. "Thuathe, a new H4/5 chondrite from Lesotho: History of the fall, petrography, and geochemistry". In: *Meteoritics and Planetary Science* 39.8 (2004), pp. 1321–1341. DOI: https://doi.org/10.1111/j.1945-5100.2004.tb00949.x.

[104] T. Kohout et al. "Annama H chondrite — Mineralogy, physical properties, cosmic ray exposure, and parent body history". In: *Meteoritics and Planetary Science* 52.8 (2017), pp. 1525–1541. DOI: https://doi.org/10.1111/maps.12871.

[105] H. Haack et al. "Ejby — A new H5/6 ordinary chondrite fall in Copenhagen, Denmark". In: *Meteoritics and Planetary Science* 54.8 (2019), pp. 1853–1869. DOI: https://doi.org/10.1111/maps.13344.

[106] S. Ghosh et al. "Fall, classification and cosmogenic records of the Sabrum (LL6) chondrite". In: *Meteoritics and Planetary Science* 37.3 (2002), pp. 439–448. DOI: https://doi.org/10.1111/j.1945-5100.2002.tb00826.x.

[107] N. Bhandari et al. "Bhawad LL6 chondrite: Chemistry, petrology, noble gases, nuclear tracks, and cosmogenic radionuclides". In: *Meteoritics and Planetary Science* 40.7 (2005), pp. 1015–1021. DOI: https://doi.org/10.1111/j.1945-5100.2005.tb00170.x.

[108] J.C. Evans et al. "Cosmogenic nuclides in recently fallen meteorites: Evidence for galactic cosmic ray variations during the period 1967–1978". In: *Journal of Geophysical Research: Solid Earth* 87.B7 (1982), pp. 5577–5591. DOI: 10.1029/JB087iB07p05577.

[109] A. Bischoff et al. "The Stubenberg meteorite—An LL 6 chondrite fragmental breccia recovered soon after precise prediction of the strewn field". In: *Meteoritics and Planetary Science* 52.8 (2017), pp. 1683–1703. DOI: https://doi.org/10.1111/maps.12883.

[110] J.C. Barton. “Studies with a low-background germanium detector in the Holborn Underground laboratory”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 354.2-3 (1995), pp. 530–538. DOI: https://doi.org/10.1016/0168-9002(94)01068-4.

[111] W. Osborn et al. “Fall, recovery and description of the Coleman chondrite”. In: *Meteoritics and Planetary Science* 32.6 (1997), pp. 781–790. DOI: https://doi.org/10.1111/j.1945-5100.1997.tb01569.x.

[112] J.M. TRIGO-RODRÍGUEZ et al. “The Berduc L6 chondrite fall: Meteorite characterization, trajectory, and orbital elements”. In: *Meteoritics and Planetary Science* 45.3 (2010), pp. 383–393. DOI: https://doi.org/10.1111/j.1945-5100.2010.01029.x.

[113] A. Bischoff et al. “Jesenice—A new meteorite fall from Slovenia”. In: *Meteoritics and Planetary Science* 46.6 (2011), pp. 793–804. DOI: https://doi.org/10.1111/j.1945-5100.2011.01191.x.

[114] M. Laubenstein et al. “Cosmogenic radionuclides in the Sołtmany (L6) meteorite”. In: *Meteorites* 2 (2012), pp. 45–51. DOI: 10.5277/met120106.

[115] S. Buhl et al. “The meteorite fall near Boumdeid, Mauritania, from September 14, 2011”. In: *Meteorites* 3 (2014), pp. 5–18. DOI: 10.5277/met140101.

[116] P. Jenniskens et al. “Fall, recovery, and characterization of the Novato L6 chondrite breccia”. In: *Meteoritics and Planetary Science* 49.8 (2014), pp. 1388–1425. DOI: https://doi.org/10.1111/maps.12323.

[117] R. Bartoschewitz et al. “The Braunschweig meteorite- a recent L6 chondrite fall in Germany”. In: *Geochemistry* 77.1 (2017), pp. 207–224. DOI: https://doi.org/10.1016/j.chemer.2016.10.004.

[118] D. Ray et al. “Petrography, classification, oxygen isotopes, noble gases, and cosmogenic records of Kamargaon (L6) meteorite: The latest fall in India”. In: *Meteoritics and Planetary Science* 52.8 (2017), pp. 1744–1753. DOI: https://doi.org/10.1111/maps.12875.

[119] J.C. Evans and J.H. Reeves. “Cosmogenic Radionuclides in the WETHERSFIELD-1982 and Tuxtuac Chondrites”. In: *Lunar and Planetary Science Conference*. Vol. 14. 1983, pp. 181–182.

[120] H. Chennaoui Aoudjehane et al. “Gueltat Zemmour: An L4 Ordinary Chondrite Fall in Morocco on August 2018”. In: *82nd Annual Meeting of The Meteoritical Society*. Vol. 82. 2019, p. 6295.

[121] N. Bhandari et al. “The Torino, H6, meteorite shower”. In: *Meteoritics* 24.1 (1989), pp. 29–34. DOI: https://doi.org/10.1111/j.1945-5100.1989.tb00938.x.

[122] N. Bhandari et al. “Exposure history of H5 chondrite Gujargaon”. In: *Meteoritics* 23.2 (1988), pp. 103–106. DOI: https://doi.org/10.1111/j.1945-5100.1988.tb00904.x.

[123] S. Yabuki, M. Shima, and S. Murayama. “Cosmic-Ray-Produced-Radioactive nuclides in freshly fallen chondrites Aomori and Tomiya”. In: *Meteoritics* 20 (1985), p. 789.

[124] A.W.R. Bevan, K.J. McNamara, and J.C. Barton. “The Binningup H5 chondrite: A new fall from Western Australia”. In: *Meteoritics* 23.1 (1988), pp. 29–33. DOI: https://doi.org/10.1111/j.1945-5100.1988.tb00893.x.

[125] M. Shima. “On isotopic analyses of Mihonoseki and other meteorites”. In: *21. Japan conference on radiation and radioisotopes*. 1994.

[126] P. Brown et al. “The fall of the St-Robert meteorite”. In: *Meteoritics and Planetary Science* 31.4 (1996), pp. 502–517. DOI: https://doi.org/10.1111/j.1945-5100.1996.tb02092.x.

[127] D.S. Lauretta et al. "OSIRIS-REx: sample return from asteroid (101955) Bennu". In: *Space Science Reviews* 212.1 (2017), pp. 925–984. DOI: https://doi.org/10.1007/s11214-017-0405-1.

[128] D.S. Lauretta et al. "The unexpected surface of asteroid (101955) Bennu". In: *Nature* 568.7750 (2019), pp. 55–60. DOI: 10.1038/s41586-019-1033-6.

[129] D.S. Lauretta et al. "Spacecraft sample collection and subsurface excavation of asteroid (101955) Bennu". In: *Science* 377.6603 (2022), pp. 285–291. DOI: 10.1126/science.abm1018.

[130] E.B. Bierhaus et al. "The OSIRIS-REx spacecraft and the touch-and-go sample acquisition mechanism (TAGSAM)". In: *Space Science Reviews* 214 (2018), pp. 1–46. DOI: https://doi.org/10.1007/s11214-018-0521-6.

[131] D.S. Lauretta et al. "OSIRIS-REx at Bennu: Overcoming challenges to collect a sample of the early Solar System". In: *Sample return missions*. Elsevier, 2021, pp. 163–194. DOI: https://doi.org/10.1016/B978-0-12-818330-4.00008-2.

[132] C. Taricco et al. "Cosmogenic radioisotopes in the Almahata Sitta ureilite". In: *Meteoritics and Planetary Science* 45.10-11 (2010), pp. 1743–1750. DOI: 10.1111/j.1945-5100.2010.01135.x.

[133] J.M. Friedrich et al. "The elemental composition of Almahata Sitta". In: *Meteoritics and Planetary Science* 45.10-11 (2010), pp. 1718–1727. DOI: https://doi.org/10.1111/j.1945-5100.2010.01105.x.

[134] K. Lodders. "Relative atomic solar system abundances, mass fractions, and atomic masses of the elements and their isotopes, composition of the solar photosphere, and compositions of the major chondritic meteorite groups". In: *Space Science Reviews* 217.3 (2021), pp. 1–33. DOI: https://doi.org/10.1007/s11214-021-00825-8.

[135] D.J. Scheeres et al. "The dynamic geophysical environment of (101955) Bennu based on OSIRIS-REx measurements". In: *Nature Astronomy* 3.4 (2019), pp. 352–361. DOI: https://doi.org/10.1038/s41550-019-0721-3.

[136] J.N. Goswami et al. "Solar flare protons and alpha particles during the last three solar cycles". In: *Journal of Geophysical Research: Space Physics* 93.A7 (1988), pp. 7195–7205. DOI: https://doi.org/10.1029/JA093iA07p07195.

[137] G. Castagnoli and D. Lal. "Solar modulation effects in terrestrial production of carbon-14". In: *Radiocarbon* 22.2 (1980), pp. 133–158. DOI: https://doi.org/10.1017/S0033822200009413.

[138] I.G. Usoskin et al. "Heliospheric modulation of cosmic rays: Monthly reconstruction for 1951–2004". In: *Journal of Geophysical Research: Space Physics* 110.A12 (2005). DOI: https://doi.org/10.1029/2005JA011250.

[139] R.A. Burger, M.S. Potgieter, and B. Heber. "Rigidity dependence of cosmic ray proton latitudinal gradients measured by the Ulysses spacecraft: Implications for the diffusion tensor". In: *Journal of Geophysical Research: Space Physics* 105.A12 (2000), pp. 27447–27455. DOI: https://doi.org/10.1029/2000JA000153.

[140] I. Leya, J. Hirtz, and J.C. David. "Galactic Cosmic Rays, Cosmic-Ray Variations, and Cosmogenic Nuclides in Meteorites". In: *The Astrophysical Journal* 910.2 (2021), p. 136. DOI: 10.3847/1538-4357/abe52f.

[141] I.G. Usoskin, G.A. Bazilevskaya, and G.A. Kovaltsov. "Solar modulation parameter for cosmic rays since 1936 reconstructed from ground-based neutron monitors and ionization chambers". In: *Journal of Geophysical Research: Space Physics* 116.A2 (2011). DOI: https://doi.org/10.1029/2010JA016105.

[142] R. Michel et al. “On the production of cosmogenic nuclides in meteoroids by galactic protons”. In: *Meteoritics* 26.3 (1991), pp. 221–242. DOI: https://doi.org/10.1111/j.1945-5100.1991.tb01043.x.

[143] R.C. Reedy and J.R. Arnold. “Interaction of solar and galactic cosmic-ray particles with the Moon”. In: *Journal of Geophysical Research* 77.4 (1972), pp. 537–555. DOI: https://doi.org/10.1029/JA077i004p00537.

[144] N. Bhandari and M.B. Potdar. “Cosmogenic 21Ne and 22Ne depth profiles in chondrites”. In: *Earth and Planetary Science Letters* 58.1 (1982), pp. 116–128. DOI: https://doi.org/10.1016/0012-821X(82)90107-8.

[145] P. Cloth. “HERMES, High energy radiation Monte Carlo elaborate system”. In: *KFA-IRW-E* 12 (1988).

[146] R.E. Prael and H. Lichtenstein. “User guide to LCS: The LAHET code system”. In: *Group* 10.6 (1989).

[147] T.W. Armstrong and K.C. Chandler. “HETC: a high energy transport code”. In: *Nuclear Science and Engineering* 49.1 (1972), pp. 110–111. DOI: https://doi.org/10.13182/NSE72-A22535.

[148] M.B. Emmett. *MORSE Monte Carlo radiation transport code system*. Tech. rep. Oak Ridge National Lab., TN (USA), 1975. DOI: https://doi.org/10.2172/5627830.

[149] J.F. Briesmeister. “A general Monte Carlo N-particle transport code”. In: *LA-12625-M* (1997).

[150] P.N. Peplowski et al. “Cosmogenic radionuclide production modeling with Geant4: Experimental benchmarking and application to nuclear spectroscopy of asteroid (16) Psyche”. In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 446 (2019), pp. 43–57. DOI: https://doi.org/10.1016/j.nimb.2019.03.023.

[151] S. Peracchi et al. “Modelling of the silicon-on-insulator microdosimeter response within the International Space Station for astronauts’ radiation protection”. In: *Radiation Measurements* 128 (2019), p. 106182. DOI: https://doi.org/10.1016/j.radmeas.2019.106182.

[152] G. Santin et al. “Normalisation modelling sources”. In: *Geant4 [Online]* (2007). URL: http://geant4.in2p3.fr/2007/prog/GiovanniSantin/GSantin_Geant4_Paris07_Normalisation_v07.pdf.

[153] A. Antonione. “Cosmogenic radioactivity estimate and detection in asteroids and meteorites: application to Bennu (OSIRIS-REx mission)”. MA thesis. University of Turin, 2022.

[154] P. Charbonneau. “Dynamo models of the solar cycle”. In: *Living Rev. Sol. Phys.* 17.1, 4 (2020), p. 4. DOI: 10.1007/s41116-020-00025-6.

[155] E.N. Parker. “The Formation of Sunspots from the Solar Toroidal Field.” In: *Astrophys. J.* 121 (1955), p. 491. DOI: 10.1086/146010.

[156] D. Bhowmik P.and Nandy. “Prediction of the strength and timing of sunspot cycle 25 reveal decadal-scale space environmental conditions”. In: *Nat. Commun.* 9, 5209 (2018), p. 5209. DOI: 10.1038/s41467-018-07690-0.

[157] I.N. Kitiashvili. “Application of Synoptic Magnetograms to Global Solar Activity Forecast”. In: *Astrophys. J.* 890.1, 36 (2020), p. 36. DOI: 10.3847/1538-4357/ab64e7.

[158] Y.B. Han and Z.Q. Yin. “A Decline Phase Modeling for the Prediction of Solar Cycle 25”. In: *Sol. Phys.* 294.8, 107 (2019), p. 107. DOI: 10.1007/s11207-019-1494-7.

[159] M.I. Pishkalo. "Preliminary prediction of solar cycles 24 and 25 based on the correlation between cycle parameters". In: *Kinemat. Phys. Celest. Bodies* 24.5 (2008), pp. 242–247. DOI: 10.3103/S0884591308050036.

[160] K.J. Li, W. Feng, and F.Y. Li. "Predicting the maximum amplitude of solar cycle 25 and its timing". In: *J. Atmos. Sol.-Terr. Phys.* 135 (2015), pp. 72–76. DOI: 10.1016/j.jastp.2015.09.010.

[161] A.K. Singh and A. Bhargawa. "An early prediction of 25th solar cycle using Hurst exponent". In: *Astrophys. Space Sci.* 362.11, 199 (2017), p. 199. DOI: 10.1007/s10509-017-3180-2.

[162] F.Y. Li et al. "Solar cycle characteristics and their application in the prediction of cycle 25". In: *J. Atmos. Sol.-Terr. Phys.* 181 (2018), pp. 110–115. DOI: 10.1016/j.jastp.2018.10.014.

[163] K.M. Hiremath. "Prediction of solar cycle 24 and beyond". In: *Astrophys. Space Sci.* 314.1-3 (2008), pp. 45–49. DOI: 10.1007/s10509-007-9728-9.

[164] D.H. Hathaway and R.M. Wilson. "What the Sunspot Record Tells Us About Space Climate". In: *Sol. Phys.* 224.1-2 (2004), pp. 5–19. DOI: 10.1007/s11207-005-3996-8.

[165] R.P. Kane. "Solar Cycle Predictions Based on Extrapolation of Spectral Components: An Update". In: *Sol. Phys.* 246.2 (2007), pp. 487–493. DOI: 10.1007/s11207-007-9059-6.

[166] N.R. Rigozo et al. "Prediction of sunspot number amplitude and solar cycle length for cycles 24 and 25". In: *J. Atmos. Sol.-Terr. Phys.* 73.11-12 (2011), pp. 1294–1299. DOI: 10.1016/j.jastp.2010.09.005.

[167] F.s Labonville, P. Charbonneau, and A. Lemerle. "A Dynamo-based Forecast of Solar Cycle 25". In: *Sol. Phys.* 294.6, 82 (2019), p. 82. DOI: 10.1007/s11207-019-1480-0.

[168] D.I. Okoh et al. "A Hybrid Regression-Neural Network (HR-NN) Method for Forecasting the Solar Activity". In: *Space Weather* 16.9 (2018), pp. 1424–1436. DOI: 10.1029/2018SW001907.

[169] E. Covas, N. Peixinho, and J. Fernandes. "Neural Network Forecast of the Sunspot Butterfly Diagram". In: *Sol. Phys.* 294.3, 24 (2019), p. 24. DOI: 10.1007/s11207-019-1412-z.

[170] T. Dani and S. Sulistiani. "Prediction of maximum amplitude of solar cycle 25 using machine learning". In: *Journal of Physics: Conf. Series*. Vol. 1231. 2019, 012022. DOI: 10.1088/1742-6596/1231/1/012022.

[171] P. Amrita et al. "Prediction of solar cycle 25 using deep learning based long short-term memory forecasting technique". In: *Advances in Space Research* 69.1 (2022), pp. 798–813. DOI: https://doi.org/10.1016/j.asr.2021.10.047.

[172] J.Y. Lu et al. "A Novel Bimodal Forecasting Model for Solar Cycle 25". In: *The Astrophysical Journal* 924.2, 59 (2022), p. 59. DOI: 10.3847/1538-4357/ac3488.

[173] Z.L. Du. "The solar cycle: a modified Gaussian function for fitting the shape of the solar cycle and predicting cycle 25". In: *The Astrophysical Journal* 367.2, 20 (2022), p. 20. DOI: 10.1007/s10509-022-04049-2.

[174] W.D. Pesnell and K.H. Schatten. "An Early Prediction of the Amplitude of Solar Cycle 25". In: *Sol. Phys.* 293.7, 112 (2018), p. 112. DOI: 10.1007/s11207-018-1330-5.

[175] H.R. Helal and A.A. Galal. "An early prediction of the maximum amplitude of the solar cycle 25". In: *J. Adv. Res.* 4.3 (2013), pp. 275–278. DOI: 10.1016/j.jare.2012.10.002.

[176] K. Petrovay et al. "Precursors of an upcoming solar cycle at high latitudes from coronal green line data". In: *J. Atmos. Sol.-Terr. Phys.* 176 (2018), pp. 15–20. DOI: 10.1016/j.jastp.2017.12.011.

[177] V. Sarp et al. “Prediction of solar cycle 25: a non-linear approach”. In: *Mon. Not. R. Astron. Soc.* 481.3 (2018), pp. 2981–2985. DOI: 10.1093/mnras/sty2470.

[178] S. Sello. “Solar cycle activity: an early prediction for cycle #25”. In: *arXiv e-prints* (2019). DOI: https://doi.org/10.48550/arXiv.1902.05294.

[179] I. Bizzarri et al. “Forecasting the solar cycle 25 using a multistep Bayesian neural network”. In: *Monthly Notices of the Royal Astronomical Society* 515.4 (2022), pp. 5062–5070. DOI: https://doi.org/10.1093/mnras/stac2013.

[180] R. Wolf. “Mittheilungen über die Sonnenflecken I”. In: *Astronomische Mitteilungen Der Eidgenossischen Sternwarte Zurich* 1 (1850), pp. 3–13. URL: https://ui.adsabs.harvard.edu/abs/1850MiZur...1....3W.

[181] F. Clette et al. “Revisiting the Sunspot Number. A 400-Year Perspective on the Solar Cycle”. In: *Space Sci. Rev.* 186.1-4 (2014), pp. 35–103. DOI: 10.1007/s11214-014-0074-2.

[182] F. Clette and L. Lefèvre. “The New Sunspot Number: Assembling All Corrections”. In: *Sol. Phys.* 291.9-10 (2016), pp. 2629–2651. DOI: 10.1007/s11207-016-1014-y.

[183] K. Petrovay. “Solar cycle prediction”. In: *Living Rev. Sol. Phys.* 17.1, 2 (2020), p. 2. DOI: 10.1007/s41116-020-0022-z.

[184] M.E. Mann. “On smoothing potentially non-stationary climate time series”. In: *Geophysical Research Letters* 31.7, L07214 (2004). DOI: 10.1029/2004GL019569.

[185] M.E. Mann. “Smoothing of climate time series revisited”. In: *Geophysical Research Letters* 35.16, L16708 (2008). DOI: 10.1029/2008GL034716.

[186] MATLAB. *version 9.7.0.1296695 (R2019b)*. Natick, Massachusetts: The MathWorks Inc., 2019.

[187] B. Efron and R.J. Tibshirani. *An introduction to the bootstrap*. CRC press, 1994. DOI: https://doi.org/10.1201/9780429246593.

[188] P.D.W Kirk and M.P.H. Stumpf. “Gaussian process regression bootstrapping: exploring the effects of uncertainty in time course data”. In: *Bioinformatics* 25.10 (2009), pp. 1300–1306. DOI: https://doi.org/10.1093/bioinformatics/btp139.

[189] Y.M. Wang and Jr. Sheeley N.R. “On the Fluctuating Component of the Sun’s Large-Scale Magnetic Field”. In: *The Astrophysical Journal* 590.2 (2003), pp. 1111–1120. DOI: 10.1086/375026.

[190] D.H. Hathaway. “The Solar Cycle”. In: *Living Rev. Sol. Phys.* 12.1, 4 (2015), p. 4. DOI: 10.1007/lrsp-2015-4.

[191] S. Mancuso et al. “Spatio-Temporal Evolution and North-South Asymmetry of Quasi-Biennial Oscillations in the Coronal Fe xiv Emission”. In: *Sol. Phys.* 293.8, 124 (2018), p. 124. DOI: 10.1007/s11207-018-1346-x.

[192] M.N. Gnevyshev. “On the 11-Years Cycle of Solar Activity”. In: *Sol. Phys.* 1.1 (1967), pp. 107–120. DOI: 10.1007/BF00150306.

[193] A.C. Cadavid et al. “Independent global modes of solar magnetic field fluctuations”. In: *Solar Physics* 226.2 (2005), pp. 359–376. DOI: 10.1007/s11207-005-8187-0.

[194] K.J. Li et al. “The periodicity of high-latitude solar activity”. In: *Solar Physics* 239.1 (2006), pp. 493–501. DOI: https://doi.org/10.1007/s11207-006-0008-6.

[195] A. Vecchio and V. Carbone. “Spatio-temporal analysis of solar activity: main periodicities and period length variations”. In: *Astronomy and Astrophysics* 502.3 (2009), pp. 981–987. DOI: https://doi.org/10.1051/0004-6361/200811024.

[196] S. Hochreiter and J. Schmidhuber. "Long Short-Term Memory". In: *Neural Comput.* 9.8 (1997), pp. 1735–1780. DOI: https://doi.org/10.1162/neco.1997.9.8.1735.

[197] F. Gers, J. Schmidhuber, and F. Cummins. "Learning to forget: continual prediction with LSTM". In: *Neural Comput.* 12.1 (2000), pp. 2451–2471. DOI: 10.1162/089976600300015015.

[198] A. Graves and J. Schmidhuber. "Framewise phoneme classification with bidirectional LSTM and other neural network architectures". In: *Neural Netw.* 18.5 (2005). IJCNN 2005, pp. 602–610. DOI: https://doi.org/10.1016/j.neunet.2005.06.042.

[199] P.J. Werbos. "Backpropagation through time: what it does and how to do it". In: *Proceedings of the IEEE* 78.10 (1990), pp. 1550–1560. DOI: 10.1109/5.58337.

[200] A. Graves. "Generating Sequences With Recurrent Neural Networks". In: *arXiv e-prints* (2013). DOI: https://doi.org/10.48550/arXiv.1308.0850.

[201] C. Olah. *Understanding LSTM Networks*. 2015. URL: http://colah.github.io/posts/2015-08-Understanding-LSTMs/.

[202] F. Chollet. *Deep learning with Python*. Simon and Schuster, 2021. ISBN: 9781617294433.

[203] E.D. Feigelson et al. "21st Century Statistical and Computational Challenges in Astrophysics". In: *Annu. Rev. Stat. Appl.* 8 (2021), pp. 493–517. DOI: 10.1146/annurev-statistics-042720-112045.

[204] Y. Gal and Z. Ghahramani. "Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning". In: *Proceedings of The 33rd International Conference on Machine Learning*. Ed. by Maria Florina Balcan and Kilian Q. Weinberger. Vol. 48. Proceedings of Machine Learning Research. PMLR, 2016, pp. 1050–1059.

[205] N. Srivastava et al. "Dropout: a simple way to prevent neural networks from overfitting". In: *The journal of machine learning research* 15.1 (2014), pp. 1929–1958. DOI: 10.5555/2627435.2670313.

[206] F. Chollet et al. *Keras*. https://keras.io. 2015.

[207] G. Van Rossum and F.L. Drake Jr. *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam, 1995.

[208] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830. URL: http://scikit-learn.sourceforge.net.

[209] C.M. Bishop. *Neural networks for pattern recognition*. Oxford university press, 1995. ISBN: 9780198538646.

[210] D.P. Kingma and J. Ba. "Adam: A Method for Stochastic Optimization". In: *arXiv e-prints* (2014). DOI: https://doi.org/10.48550/arXiv.1412.6980.

[211] L. Zhu and N. Laptev. "Deep and Confident Prediction for Time Series at Uber". In: *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*. 2017, pp. 103–110. DOI: 10.1109/ICDMW.2017.19.

[212] C.T. Russell, J.G. Luhmann, and L.K. Jian. "How unprecedented a solar minimum?" In: *Reviews of Geophysics* 48.2 (2010). DOI: 10.1029/2009RG000316.

[213] D. Nandy. "Progress in Solar Cycle Predictions: Sunspot Cycles 24-25 in Perspective". In: *Sol. Phys.* 296.3, 54 (2021). DOI: 10.1007/s11207-021-01797-2.

[214] A.R. Yeates, D. Nandy, and D.H. Mackay. "Exploring the Physical Basis of Solar Cycle Predictions: Flux Transport Dynamics and Persistence of Memory in Advection- versus Diffusion-dominated Solar Convection Zones". In: *Astrophys. J.* 673.1 (2008), pp. 544–556. DOI: 10.1086/524352.

[215] P.J. Bushby and S.M. Tobias. "On Predicting the Solar Cycle Using Mean-Field Models". In: *The Astrophysical Journal* 661.2 (2007), pp. 1289–1296. DOI: 10.1086/516628.

[216] H.W. Babcock. "The Topology of the Sun's Magnetic Field and the 22-YEAR Cycle." In: *Astrophys. J.* 133 (1961), p. 572. DOI: 10.1086/147060.

[217] R.B. Leighton. "A Magneto-Kinematic Model of the Solar Cycle". In: *Astrophys. J.* 156 (1969), p. 1. DOI: 10.1086/149943.

[218] N.V. Zolotova and D.I. Ponyavin. "Is the new Grand minimum in progress?" In: *J. Geophys. Res. (Space Phys.)* 119.5 (2014), pp. 3281–3285. DOI: 10.1002/2013JA019751.

[219] V. Zharkova et al. "Upcoming modern grand minimum and solar activity prediction backwards five millennia". In: *EGU General Assembly Conference Abstracts*. EGU General Assembly Conference Abstracts. 2018, p. 8066.

[220] D.B. Percival and A.T. Walden. *Wavelet methods for time series analysis*. Vol. 4. Cambridge university press, 2000. DOI: https://doi.org/10.1017/CBO9780511841040.

[221] C. Torrence and P.J. Webster. "Interdecadal changes in the ENSO–monsoon system". In: *Journal of climate* 12.8 (1999), pp. 2679–2690. DOI: https://doi.org/10.1175/1520-0442(1999)012<2679:ICITEM>2.0.CO;2.

Appendix A

Continuous Wavelet Transform (CWT) and Wavelet Transform Coherence (WTC)

The Continuous Wavelet Transform (CWT) allows for evolutionary spectral analysis of a series in the timescale plane [12, 36, 220]. The concept of scale, typical of this method, represents a time duration that can be properly translated into a Fourier period and hence into a frequency.
Conceptually, the CWT is a sliding cross-correlation between a finite-length sequence $X(n)$ and a family of finite-energy, zero-mean waveforms, known as *wavelets*, defined as:

$$W_n^X(s) \sim \frac{1}{\sqrt{s}} \sum_{n'=0}^{N-1} X(n') \Psi_0^*(\frac{n'-n}{s}) \tag{A.1}$$

where Ψ_0 is the reference wavelet, commonly known as *mother wavelet*, s is the scale factor, and n is the lag. The wavelet power spectrum or scalogram, defined as the absolute value squared of the wavelet transform, is constructed by varying the wavelet scale s and translating along the localized time index n. It can be considered as an estimate of the similarity between the signal and the wavelet functions, giving a measure of the time series variance at each scale (period) and at each time.
Errors at the beginning and end of the wavelet power spectrum are to be expected when dealing with time series of finite length. This issue is usually solved by zero-padding the signal, that is, by adding zeroes at the end of the time series before applying the wavelet transform. These added zeros constitute a discontinuity in the signal causing a distortion of the CWT values on both edges from their hypothetical value, i.e., the one hypothetically expected if the signal had an infinite amount of data.
The cone of influence (COI) is the area in the scalogram potentially affected by these edge effect artifacts, where the wavelet power spectra are distorted due to the finite length of the series. The larger the scale, the wider the time intervals adjacent to the edges where the CWT is disturbed. Within the COI region, the information provided by the scalogram is an accurate time-frequency representation of the data, while outside of it, information should be treated as suspect. In practice, the COI is defined as the area in which the wavelet power caused by a discontinuity at the edge has dropped to e^2 of the edge value.

To study the relative phase of the common oscillations of two signals, the wavelet transform coherence (WTC) represents a suitable method. Following Ref.[221], the WTC of two time series $X(n)$ and $Y(n)$ is defined as:

$$R_n^2(s) = \frac{|S(s^{-1}W_n^{XY}(s))|^2}{S(s^{-1}|W_n^X(s)|^2)S(s^{-1}|W_n^Y(s)|^2)} \tag{A.2}$$

where S is a smoothing operator, W^x is the wavelet transform of the time series $X(n)$, and $W^{XY} = W^X W^{Y*}$ is the cross wavelet transform of the two time series. This definition closely resembles that of a traditional correlation coefficient, and it is useful to think of wavelet coherence as a localized correlation coefficient in the time-frequency space.
The WTC spectrum in Section 1.4.2 was obtained using the Paul mother wavelet, which is narrower in time than the generally applied Morlet one, thus producing a better temporal localization of transient events. The spectrum was calculated using the wavelet coherence MATLAB package from Ref.[13].

Appendix B

Singular Spectrum Analysis (SSA) and Monte Carlo test (MC-SSA)

The Singular Spectrum Analysis (SSA) is a non-parametric spectral method that allows to efficiently extract the deterministic components of a time series from noise [14–16]. It uses data-adaptive filters to separate the time series into statistically independent components that can be classified as oscillatory patterns, modulated in amplitude and phase.
The SSA methodology applied on a time series $x(n)$, consists of three basic steps:

- embedding the time series of length N in a vector space of proper dimension M;
- computing the $M \times M$ lag-covariance matrix C_D of the data;
- diagonalizing the matrix C_D to extract its eigenvectors E_k and corresponding eigenvalues λ_k, with $k = 1, M$.

The lag-covariance matrix C_D can be estimated as:

$$C_D = \frac{1}{N'} \cdot D^T \cdot D \tag{B.1}$$

where $N^{'}$ = N - M + 1 is the embedding dimension, and D is the trajectory matrix, defined as follows:

$$D = \begin{pmatrix} x(1) & x(2) & \dots & x(M) \\ x(2) & x(3) & \dots & x(M+1) \\ \vdots & \vdots & \ddots & \vdots \\ x(N') & x(N^{'}+1) & \dots & x(N) \end{pmatrix} \tag{B.2}$$

The diagonalization of the matrix C_D yields the diagonal matrix $\Lambda_D = E_D^T C_D E_D$, where $\Lambda_D = diag(\lambda_1, \lambda_2, \lambda_3, ..., \lambda_M)$, with $\lambda_1 > \lambda_2 > \lambda_3 > ... > \lambda_M > 0$, and E_D is the $M \times M$ matrix having the corresponding eigenvectors $E_k, k = 1, M$, as its columns.
For each E_k, also known as Empirical Orthogonal Functions (EOFs), the k^{th} principal component (PC) is constructed, representing a time series of length $N^{'}$ obtained by projecting the original time series on the eigenvector E_k:

$$P_k(t^{'}) = \sum_{j=1}^{M} x(t^{'}+j) E_k(j) \tag{B.3}$$

The corresponding eigenvalue λ_k describes its variance, which can be interpreted as the percentage of the time series $x(n)$ described by the k^{th} component.
Given a subset of eigenvalues, it is possible to extract time series of length N by combining the corresponding PCs. These time series, called Reconstructed Components (RCs), capture the variability associated with the eigenvalues of interest, and are estimated as follows:

$$R_k(t) = \frac{1}{M_t} \sum_{k \in K} \sum_{j=L_t}^{U_t} P_k(t-j) E_k(j) \tag{B.4}$$

The values of the normalization factor M_t and the lower and upper bound of summation L_t and U_t differ between the central part of the time series and its end points [15, 16]:

$$(M_t, L_t, U_t) = \begin{cases} (\frac{1}{t}, 1, t) & \text{if} \quad 1 \leq t \leq M-1 \\ (\frac{1}{M}, 1, M) & \text{if} \quad M \leq t \leq N' \\ (\frac{1}{N-t+1}, t-N+M, M) & \text{if} \quad N'+1 \leq t \leq N \end{cases} \tag{B.5}$$

To reliably identify the oscillatory components in a series, that is to assess the statistical significance of the variability modes detected inside the records, a Monte Carlo (MC) recursive method is used [17]. At a given confidence level, a set of models for the analyzed time series (null hypotheses) are iteratively tested, starting from the simplest one, representing pure noise, up to more complex ones, which are obtained by including different combinations of spectral components superimposed on the noise. A Monte Carlo ensemble of surrogate time series is generated from the assumed model, and the SSA is applied to each surrogate series. Then, data are projected onto the EOFs of the null hypothesis to test whether it is possible to distinguish the series from the ensemble. The rejection of a null hypothesis is achieved by comparing the SSA spectrum of the analysed series with the Monte Carlo band, obtained by applying the SSA to the ensemble of surrogate series generated from the considered model of the series. Finally, the model explaining the series, and allowing the identification of the significant components at a given confidence level, is the one that cannot be rejected.
The AR(1) noise assumption is usually used since a large class of physical processes generates series with larger power at lower frequencies. This is done to avoid overestimating the system predictability by underestimating the amplitude of the stochastic component of the time series [15].
The MC-SSA analyses were performed using the SSA-MTM Toolkit[1].

[1]Freeware SSA-MTM Toolkit at http://www.atmos.ucla.edu/tcd/ssa.

Ringraziamenti

Il mio percorso di dottorato, come quello di molti altri, è stato decisamente "anomalo". Il primo anno è stato rallentato da una pandemia globale. Il secondo, di conseguenza, è stato carico di lavoro arretrato, ma ha portato con sé i primi soddisfacenti frutti. Il terzo, infine, è stato ricco di traguardi e pieno di consapevolezza delle proprie capacità e delle tante competenze acquisite.

Desidero innanzitutto ringraziare la mia tutor, la Prof.ssa Carla Taricco, per avermi guidata nel mio percorso di ricerca, per le conoscenze trasmesse, per la fiducia riposta in me e per avermi sempre dato preziosi consigli.

Ringrazio il coodinatore del dottorato, il Prof. Paolo Olivero, per la sua disponibilità e professionalità nel gestire ogni situazione.

Ringrazio i miei genitori per il loro immenso sostegno. Sin dal mio primo giorno da studentessa universitaria, nell'ormai lontano 2012, la loro infinita fiducia nelle mie capacità mi ha sempre spinta a dare il meglio fino a raggiungere quest'ultimo importante traguardo.

Ringrazio con tutto il cuore Dario: compagno di studi, di ricerca, di conferenze e di pranzi al sushi. Un punto di riferimento in questi 3 anni, sempre disponibile e pronto a dare una mano nei momenti di difficoltà.

Un ringraziamento al gruppo di Cosmogeofisica dell'Università di Torino, del quale ho fatto parte in questi anni, e a tutti i gruppi di ricerca con i quali ho avuto la possibilità di collaborare. Grazie perchè, con tutti voi, ho avuto modo di portare avanti ricerce profondamente interessanti.

Ringrazio il mio compagno Alvin, sempre al mio fianco, sempre pronto a dare un consiglio valido e sempre pronto a sostenermi nelle mie scelte.

E infine, perché no, ringrazio me stessa, per aver creduto nelle mie capacità, per essermi spinta oltre gli ostacoli e per aver sempre cercato di dare il meglio in tutto ciò che ho fatto.

E questa tesi la dedico proprio a me!

Grazie a tutti!