

11

A Brief Review of the Sun–Climate Connection, with a New Insight Concerning Water Vapour

Dr Willie Soon & Dr Sallie Baliunas

Summary

This chapter gives a brief historical overview of the scientific study of the Sun–climate relationship. We then provide evidence for the existence of physical links between total solar irradiance (TSI) and Northern Hemisphere surface air temperature records from Arctic, US, Ireland and China, as well as the composite four-region record, as recently reported in Soon, Connolly and Connolly (2015). The results for TSI and conterminous US and Ireland temperature records is new. We use quality-controlled datasets for both TSI and surface temperature records avoiding issues that can arise from the contamination of temperature records by non-climatic factors. This is achieved by either adopting rural-only temperature datasets or urbanisation-corrected records. In this study, we show how TSI may be correlated with atmospheric water vapour. Left for future investigation is the development of a comprehensive theory of the climate and Sun–climate connection.

Early history

Early observations and speculations in classical Greece some 2300 to 2400 years ago, can be traced to Theophrastus (371–287 BCE) who suggested a connection between sunspots, and rain and wind (Soon and Yaskell 2003).

The Jesuit astronomer Giovanni Battista Riccioli (1598–1671) also wondered about a connection between sunspot activity and weather. Riccioli, in *Almagestum novum* published in 1651, noted the warm and dry weather in Italy during September of 1632 when sunspots were absent, in contrast to the cold June of 1642 when sunspots were prevalent. Galindo and Saladino (2008) recently pointed out a remarkable, early insight from the Mexican astronomer and meteorologist, Jose Antonio Alzate (1737–1799), who commented in 1784 that:

The light conveyed by the Sun, efficiently influences vegetables and live-stock markets: this assertion should be present in the minds of those savants dealing with meteorological observations (none has implemented it). Pliny tells us in his ‘Natural History’ that, in times of Augustus, the Sun was dim and epidemics and food scarcity struck; modern physicists ascribe the mentioned dimness to the numerous sunspots that masked the solar disc at that time; from the year [17]69 I have been observing the Sun most days till the year [17]83, always numerous sunspots. In relation to the Eclipse of August 15, [17]84, I surveyed the Solar body using a telescope with much magnification and great clarity and confirmed that it was completely spotless. I continued observing daily up to 29 October, and no single spot appeared. Would the seasonal alterations experienced in Europe and here, depend partly on that cause? Only experience will tell.

With the benefits of modern measurements and equipment, the specific suggestion by Alzate of more dark sunspots leading to a dimmer Sun can now be rejected (see, for example, Scafetta and Willson 2014). For a more knowledgeable understanding of the Sun, fast-forward another 125 years, when Annie Maunder and Walter Maunder (1908), with their long experience of observing the Sun, remarked in *The Heavens and their Story* that:

A [‘great’] spot like that of February, 1892 is enormous by itself, but it is a small object compared to the sun; and spots of such size do not occur frequently, and last but a very short time. We have no right to expect,

A BRIEF REVIEW OF THE SUN–CLIMATE CONNECTION

therefore, that a time of many sunspots should mean any appreciable falling off in the light and heat we have from the sun. Indeed, since the surface around the spots is generally bright beyond ordinary, it may well be that a time of many spots means no falling off, but rather the reverse.

This acute observational fact alone tells us that a study of TSI variability, applied to a physical study of Sun–climate relations, must not be blindly equated to the study of sunspot count statistics *per se* (an important issue, which we will discuss again further, below).

In the English-speaking world, the idea and proposal concerning sunspots, weather and climate by William Herschel (1738–1822), in his twin papers published in *Philosophical Transactions of the Royal Society*, is popularly known and cited. These pioneering observations and remarks by Herschel (1801a, 1801b) have been considered both a positive support for and, further, deep scepticism of the physical reality of the Sun–weather–climate relationship. In his May 1801 paper, which is possibly less read and cited, he wrote:

But we need not in the future be at a loss how to come at the truth of the current temperature of this climate, as the thermometrical observations, which are now regularly published in the *Philosophical Transactions*, can furnish us with a proper standard, with which the solar phenomena may be compared. ... [A]lthough I have, in my first Paper, sufficiently noticed the want of a proper criterion for ascertaining the temperature of the early periods where the sun has been recorded to have been without spots, and have also referred to future observations for shewing whether a due distribution of dry and wet weather, with other circumstances which are known to favour the vegetation of corn, do or do not require a certain regular emission of the solar beams ... For, if the thermometer, which will be our future criterion, should establish the symptom we have assigned, of a defective or copious emission of the solar rays, or even help us to fix on different ones, as more likely to point out the end we have in view, we may leave it entirely to others, to determine the use to which a fore-knowledge of the probable

temperature of an approaching summer, or winter, or perhaps or both may be applied; but still it may be hoped that some advantage may be derived, even in agricultural economy, from an improved knowledge of the nature of the sun, and of the causes, or symptoms, of its emitting light and heat more or less copiously.

We have quoted from both Alzate and Herschel in order to explain the limitations that existed to the study of the Sun–climate connection during their times. As well, it raises awareness about similar limitations in our own era; physical quantities are needed for a more quantitative and secure study, and to claim any relationship between the Sun, weather and climate. In the late eighteenth and early nineteenth centuries, thermometer records were neither widely available nor mature in their precision. As a result, Alzate and Herschel were forced to adopt the price of wheat or food as an indirect measure of meteorological and climatic fluctuations.

Thomas Jefferson (1743–1826), while concurring with Herschel’s remark, was also calling for a firmer physical basis to study our climatic history. In marginal notes from his weather diary made at Monticello from 1 January 1810 to 31 December 1816, Jefferson says:

It is a common opinion that the climates of the several states of our union have undergone a sensible change since the dates of their first settlements; that the degrees of both cold and heat are moderated. The same opinion prevails as to Europe; if facts gleaned from history give reasons to believe that, since the times of Augustus Caesar, the climate of Italy, for example, has changed regularly at the rate of 1° of Fahrenheit’s thermometer for every century. May we not hope that the methods invented in latter times for measuring with accuracy the degrees of heat and cold, and the observations which have been and will be made and preserved, will at length ascertain this curious fact in physical history?

Measurements remain a problem because, as we shall discuss and amplify later this chapter, even with the given thermometer records, the

hardest tasks of any serious scientific study and analysis is to discern, and therefore remove, any contamination of the datasets by non-climatic factors. This problem is explained in detail by Connolly and Connolly (2014a; 2014b; 2014c). In addition, any Sun–climate study that relies strictly on sunspot-number records for quantitative assessment of solar activity level or strength may miss the correct physical and dynamic ranges of solar variability. It has long been warned, for example in Menzel (1959), that the ‘zero’ in sunspot numbers has an ambiguous meaning and value.

Before we consider these issues in more detail, let us briefly summarise findings from other contemporary studies of the Sun–climate connection.

Contemporary studies

In several previous publications, we have also shown and discussed the physical plausibility and relationship between TSI variations and various key climatic indicators and paleoproxies. Examples of these are the North Atlantic Meridional Overturning Circulation Index; Tropical Atlantic sea-surface temperatures; glacier mass and movements; zones and areal extent of tropical rain belts; and even waveguide teleconnection patterns, which have originated meteorologically, and where certain apparent correlations with zero-time lag as well as delayed responses of up to 50 to 100 years are all documented (Hormes et al. 2006; Soon 2009; Soon et al. 2011; Asmerom et al. 2013; Cruz-Rico et al. 2015; Yan et al. 2015). A more comprehensive discussion on the nomenclature of solar and climatic oscillations is given in Soon et al. (2014).

Recently, since the pioneering works discussed above, the scientific pursuit of the Sun–climate connection has been expanded, both in content and context. At the same time, it is openly admitted and discussed that recent studies are also not without controversies, with some having both technical strengths and weaknesses. Courtillot, Le Mouél and colleagues, in a series of papers (Le Mouél et al. 2009;

Courtillot et al. 2010; Kossobokov et al. 2010), introduced more advanced statistical techniques and various transformed climatic variables for addressing the non-stationary and non-linear nature of the solar forcing and climatic response relations. Van Loon et al. (2012) examined the multidecadal patterns of sea-level pressure and the associated quasi-stationary wave in the North Atlantic region, and suggested that this pattern could be related to the 50- to 100-year so-called Gleissberg–Yoshimura cycle (Solheim et al. 2012); while examining the relations between sunspot-cycle length and surface temperatures from Norway and the North Atlantic region offered the prediction of a significant temperature decrease during sunspot cycle 24 (that is, from about 2008 to 2018–2019).

Zhao and Feng (2014), from the Space Weather community, emphasised the solar activity and Earth–climate connection in terms of the multidecadal-centennial to bicentennial modulation. Many others are studying solar activity and weather and climate relations through charged-particle effects, including through incoming galactic cosmic rays and the global atmospheric electric circuit that are both in turn modulated by the solar wind (Svensmark & Friis-Christensen 1997; Soon et al. 2000; Yu 2002; Shaviv 2005; Harrison & Stephenson 2006; Usoskin et al. 2006; Miyahara et al. 2008; Harrison & Usoskin 2010; Kirkby et al. 2011; Tinsley 2012; Voiculescu & Usoskin 2012; Svensmark et al. 2013; Nicoll & Harrison 2014; Veretenenko & Ogurtsov 2014; Yu & Luo 2014; Zhou et al. 2014; Lam & Tinsley 2016; Prikyl et al. 2016).

The latest progress report in studying the cosmic-rays–cloud hypothesis by the CERN CLOUD experiment (Kirkby et al. 2016) has confirmed the importance of cosmic rays in elevating the natural background – in the absence of sulphuric acid that is known to be mostly an anthropogenic factor – of the biogenic emission of aerosol particles

of 1.7 nanometers, or larger. However, there is still a long way to go to find a complete answer to the mystery of cloud formation in pristine and polluted environments.

Finally, the quasi-review paper by van Geel and Ziegler (2013) offered some perspective and context for solar activity and climatic covariations from the point of view of paleoclimate proxies. They concluded that the United Nations' (UN) Intergovernmental Panel on Climate Change (IPCC) reports have indeed underestimated the Sun's role in climate change. This conclusion challenges the community consensus-driven review on the Sun–climate connection recently published by Gray et al. (2010). In addition, Soon (2014) also independently reviewed the consensus Sun–climate discussion, as reported in the latest IPCC AR5 Working Group I report, and found serious misrepresentation, missing physical mechanisms, and political biases.

The difficulty of measuring TSI

The first issue that must be addressed in any solar TSI-climate study will concern the choice of measurement for TSI. All measurements are approximations of the total solar energy arriving at the Earth's atmospheric column and/or its surface, but there is no universally agreed measure. This issue is not widely understood, even among climate scientists.

It is shown in Soon, Connolly and Connolly (2015) that most of the often-quoted TSI records are based on the work of Wang, Lean and Sheeley (2005). These records have been subsequently used in the IPCC reports. They are very closely correlated to the sunspot-number records and, hence, are likely unable to capture our hypothesised, longer-term changes and modulation of solar irradiance rooted in various magnetic structures and complexes. This fact is consistent with the pioneering observations by Maunder and Maunder (1908) quoted earlier.

Indirect evidence supporting longer-term modulation of TSI, beyond the eleven-year cycle, such as sunspots variations, can be found by studying the changes in the fluid dynamics associated with the three-dimensional structures of magnetic dark spots, and related changes in solar convective zone properties (Hoyt 1979; Hathaway 2013; Bludova et al. 2014; Livingston & Watson 2015).

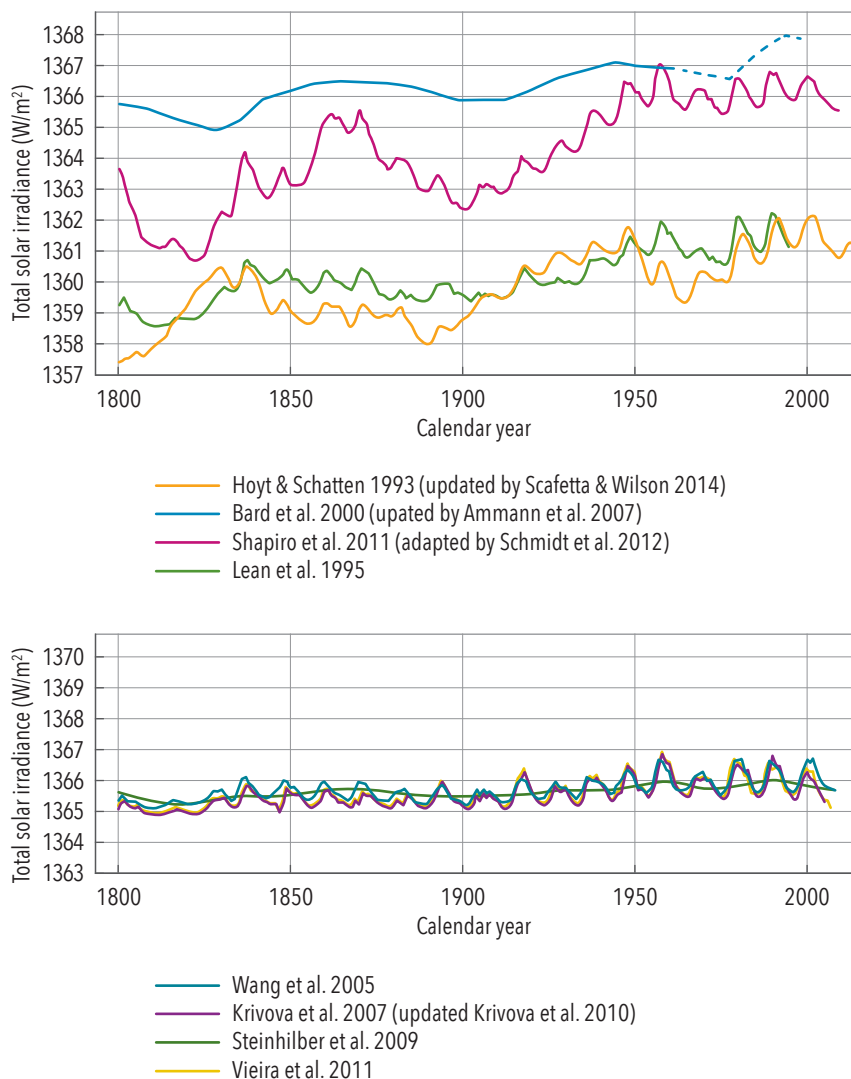
Figure 11.1 compiles seven proposed TSI reconstructions and the TSI index originally proposed by Hoyt and Schatten (1993), with updates (orange line top chart).

The diverse ranges of reconstructed TSI histories and values are explained mainly from the wide range of unknowns and uncertainties involved in the study of TSI beyond the direct observations by satellite-borne radiometers available only from 1979 to the present. The differences have been studied and discussed in detail in Soon, Connolly and Connolly (2015). In addition, as discussed above, TSI reconstructions that are based mainly on sunspot records alone may fail to fully capture the true amplitude and variation of the TSI index.

We chose the TSI reconstruction from Hoyt and Schatten (1993) because their work involves the most diverse types and ranges of proxy values for solar irradiance estimation – sunspot-cycle amplitude, sunspot-cycle length, solar equatorial-rotation rate, fraction of penumbral spots, and the decay rate of the approximate eleven-year sunspot cycle.

Their assumption is that each of those slightly different proxies will most likely capture some part of the underlying mechanism responsible for modulating the solar magneto-convection-induced processes that affect TSI. In an *à priori* sense, we note that magneto-fluid dynamical processes on the Sun need not strictly produce eleven-year-like cycles of high and near-zero sunspot numbers, as specified artificially by the paleoclimate modelling community (see Schmidt et al. 2011, 2012). In that regard, it is important to note that in the most recent work on

Figure 11.1 Different measures of total solar irradiance



The top chart shows the results from four studies suggesting relatively high solar variability since 1800. (The dashed blue line represents the work of Ammann et al. 2007, which was used to update Bard et al. 2000.) This contrasts with the bottom chart showing the results from low solar variability studies used as 'natural forcings' for the CMIP5 Global Climate hindcasts in the IPCC's Fifth Assessment Report (5AR).

Source: Adapted by permission from Elsevier – Soon, W, Connolly, R & Connolly, M, 'Re-evaluating the role of solar variability on Northern Hemisphere temperature trends since the 19th century', *Earth-Science Reviews*, vol. 150, pp. 409-452, copyright 2015.

physical modelling of TSI variability by Fontenla et al. (2011), up to nine solar magnetic features and structures were considered:

1. dark quiet-Sun inter-network
2. quiet-Sun inter-network
3. quiet-Sun network lane
4. enhanced network
5. plage (that is, not facula)
6. facula (that is, very bright plage)
7. sunspot umbra
8. sunspot penumbra
9. hot facula.

Despite this important progress, results are still unsatisfactory. The reason for this is that it has been found, and known, that even the ‘quiet’ part of the Sun may consist of ‘unresolved’ small-scale magnetic fields that vary, perhaps more than likely in a non-linear fashion, with the solar cycle in both mean strength and spatial distribution (Schuhle et al. 2000; Trujillo Bueno et al. 2004; Schnerr & Spruit 2011; Stenflo 2012; Stenflo & Kosovichev 2012).

Correlating TSI with surface temperatures

Many researchers have failed to find a relationship between TSI and changes in the Earth’s surface temperatures because they used the wrong TSI proxy measure, and then chose a corrupted surface temperature record.

It is important when aiming towards a physical study of the Sun–climate connection to search for multidecadal covariations as a physical measure of the Sun’s manifest activity. One relationship that has been explored is the TSI index in rural and urbanisation-corrected temperature datasets from the Northern Hemisphere. When high-quality surface temperature records are used with the Hoyt and Schatten TSI data, intriguing correlations can be identified (1993). Specifically, Soon

(2005) and Soon (2009) found a strong apparent correlation for the Arctic; Soon et al. (2011) found a similar correlation with China; Soon and Legates (2013) also found a correlation between TSI and trends in the ‘Equator-to-Pole (Arctic) Temperature Gradient’.

How can one be sure, however, that those temperature records have not been contaminated with non-climatic effects and factors? If they have been contaminated, how does one quantify the non-climatic components, and therefore, remove those factors from the raw or unmodified datasets?

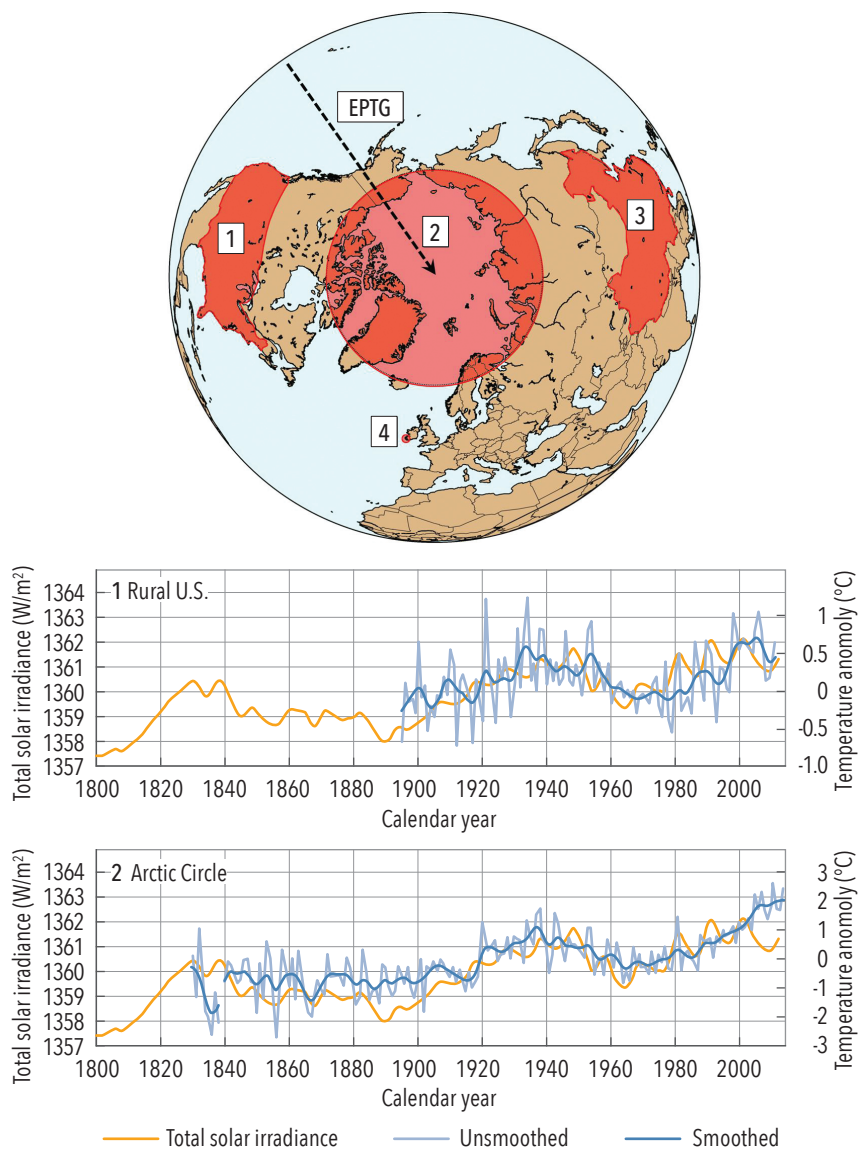
Figure 11.2 shows the temperature anomalies for key regions after these questions have been answered in detail, with the workings described in Soon, Connolly and Connolly (2015). The statistical correlations shown in Figure 11.2, considering how diverse the geographical areas and climatic zones that are covered, are very impressive.

Correlating TSI with water vapour concentrations

A so far much neglected key variable that is important in understanding ocean–land–atmosphere climate systems is the total atmospheric water vapour. However, a direct measurement of this metric is not available for a period of 100 years, and so the record must be constructed by using indirect deduction from the so-called climate reanalyses study (Hersbach et al. 2015; Kobayashi et al. 2015; Poli et al. 2016).

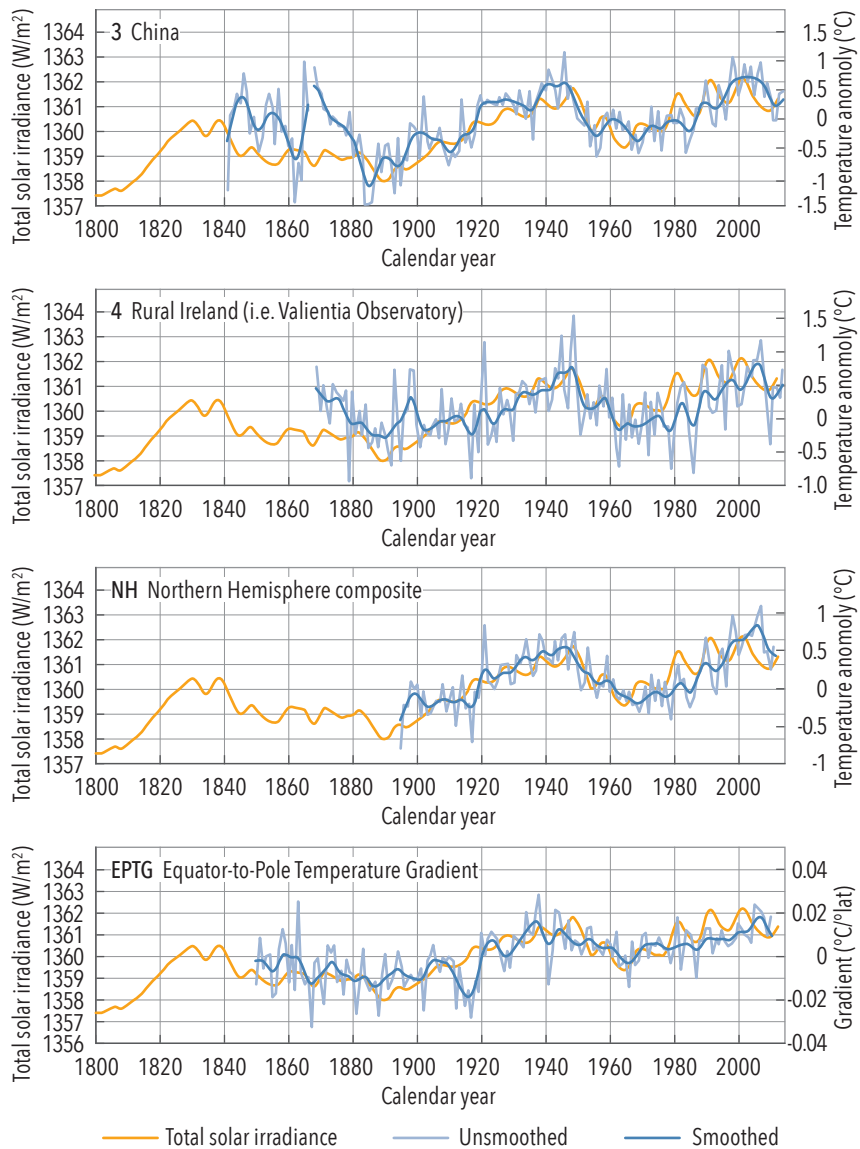
Figure 11.3 shows the correlation between the TSI index of Hoyt and Schatten and the total column water vapour content, as deduced from the European Centre for Medium-Range Weather Forecasts’ model outputs. The relationship is close and meaningful, especially for the multidecadal variation and modulation. However, we note that the inter-annual changes are more closely related to internal oscillating components of the climate system, such as the El Niño–Southern Oscillation (ENSO) or the North Atlantic Oscillation (NAO) factors. The close correlation, perhaps, reflects a real physical relationship between

Figure 11.2 Comparing TSI and Northern Hemisphere temperatures

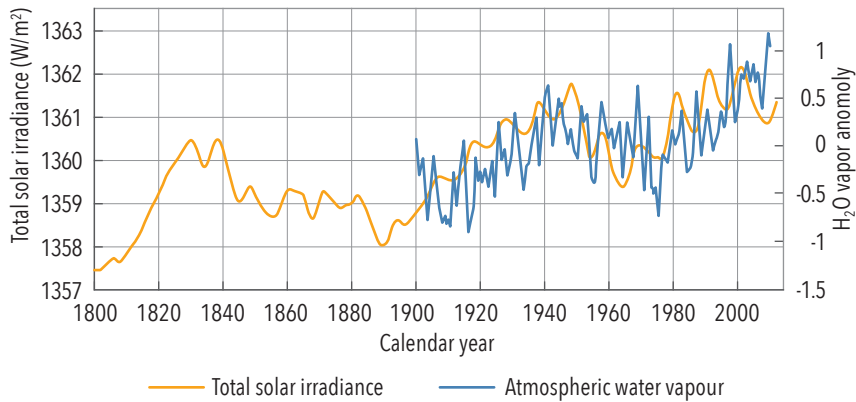


Comparison between the Hoyt and Schatten TSI trends (orange line) and the various temperature indices (blue lines) from the Northern Hemisphere regions (1. Rural US, 2. Arctic Circle, 3. China, 4. Rural Ireland, and the Northern Hemisphere composite record), as well as the Northern Hemispheric Equator-to-Pole temperature gradients (EPTG).

A BRIEF REVIEW OF THE SUN-CLIMATE CONNECTION



Source: Adapted from Soon, Connolly and Connolly 2015.

Figure 11.3 Comparing TSI and water vapour trends

Comparison between the Hoyt and Schatten TSI trends (orange line) and the total atmospheric water vapour content anomaly (over ice-free region of the Earth's oceans) based on the twentieth-century climate reanalyses outputs from the climate model of the European Centre for Medium-Range Forecasts.

Source: Data courtesy of Drs A Simmons and H Hersbach.

sea-surface temperatures and the atmospheric water vapour, as convincingly emphasised in Hersbach et al. (2015).

If the results from Figure 11.2 and Figure 11.3, both in terms of physical correlations and mechanisms, can be fully confirmed, we will then have very clear empirical evidence for supporting a Sun–climate connection on multidecadal timescales. This is in the sense that it is the TSI-induced changes in modulating the direct thermal impact on the surface heat budget, as well as the amplification factor from the covarying atmospheric water vapour that best explained all the available observed climatic records.

In contrast, this fact will imply that the role of rising atmospheric CO₂ may be, at most, minimal when explaining the surface temperature and the atmospheric water vapour records. Indeed, in Soon, Connolly and Connolly (2015), we went a step further – by using the time series in combination with the atmospheric CO₂ forcing – and found that the

quantitative constraints for the Earth's 'climate sensitivity' to the doubling of atmospheric CO₂ was anywhere between 0.4 to 2.5 °C. The high value is obtained by assuming, without any strong supporting arguments, that the Northern Hemisphere derived composite temperature is driven foremost by rising atmospheric CO₂ content alone. The lower value of 0.4 °C for CO₂-doubling climate sensitivity is obtained by assuming the observed empirical correlation shown in Figure 11.2 is valid, and then fitting only a CO₂ forcing curve to the residual temperature trend after accounting for the TSI factor.

Conclusion

It is hoped that the study of Sun–climate relations will remain true to science and the scientific method, and be able to be better interpreted using high-quality data.

We are, perhaps, moving slowly towards a unifying theory of climate (Lions et al. 1993; Monin & Shishkov 2000; Li & Wang 2008; Essex 2011, 2013; Palmer et al. 2014). This will be further strengthened when there is a better appreciation of the socio-political factors, which currently obscure a better understanding of the Sun–climate connection. This would be helped by a proper definition of 'climatology' (for example, Landsberg 1972; Gutmann 1989; Arguez & Vose 2011).

In Soon, Connolly and Connolly (2015) the importance of using uncontaminated surface temperature datasets is explained, along with reasons for choosing the Hoyt and Schatten TSI-record. The question about the veracity of TSI being adopted for such a study of the Sun–climate connection must necessarily remain open and be answered in time by solar physics. Even more importantly is the fact that several other promising mechanisms, especially those involving the charged particles and electric-circuit phenomena outlined above, must all be fully examined and integrated into a more comprehensive framework, which should, ultimately, give us the correct superposition rules for the Sun–Earth climatic connection in multiscale spatial and temporal domains.