Received: 1 August 2021 | Revised: 28 November 2021 | Accepted: 23 December 2021

DOI: 10.1002/joc.7662

RESEARCH ARTICLE

International Journal of Climatology RMetS

Near-surface soil thermal regime and land–air temperature coupling: A case study over Spain

Camilo Melo-Aguilar[1,2] | Fidel González-Rouco[1,2] | Norman J. Steinert[1,2] | Hugo Beltrami[3] | Francisco José Cuesta-Valero[4,5] | Almudena García-García[4,5] | Felix García-Pereira[1,2] | Elena García-Bustamante[6] | Pedro José Roldán-Gómez[1,2] | Thomas Schmid[6] | Jorge Navarro[6]

[1]Universidad Complutense de Madrid and Instituto de Geociencias (UCM-CSIC), Madrid, Spain

[2]Facultad de Ciencias Físicas, Universidad Complutense de Madrid, Madrid, Spain

[3]Environmental Science Program, Memorial University of Newfoundland, St. Jhons's, Newfoundland, Canada

[4]Climate and Atmospheric Science Institute, St Francis Xavier University, Antigonish, Nova Scotia, Canada

[5]Department of Remote Sensing, Helmholtz Centre for Environmental Research – UFZ, Leipzig, Germany

[6]Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain

Correspondence
Camilo Melo-Aguilar, Universidad Complutense de Madrid and Instituto de Geociencias (UCM-CSIC), Madrid 28040, Spain.
Email: camelo@ucm.es

Funding information
Ministerio de Ciencia, Innovación y Universidades, Grant/Award Number: RTI2018-102305-B-C21; Ministerio de Economía, Industria y Competitividad, Gobierno de España, Grant/Award Number: BES-2015-075019

Abstract

Understanding the near-surface soil thermal regime and its connection to the atmospheric state is important for the assessment of several climate-related processes. However, the lack of in situ soil temperatures measurements limits the analysis of such processes. In this study, we have developed a quality-controlled soil temperature database for Spain that consists of 39 sites spanning from 1987 to 2018. We have used this database to assess the near-surface soil thermal regime. Likewise, we evaluate at seasonal to multidecadal timescales the land–air temperature coupling over Spain by analysing the structure of the surface air temperature (SAT) and the ground surface temperature (GST) covariance and also their long-term evolution. In addition, we have employed the ERA5-Land reanalysis to test the consistence between observations and reanalysis. The results show that the near-surface soil thermal structure is dominated by conduction despite some influence of hydrology-related processes. Regarding the land–air temperature coupling, we have found a strong connection between SAT and GST. However, in the summer months there is an offset in SAT–GST at some sites due to limited evaporation and enhanced sensible heat fluxes. Furthermore, multidecadal SAT–GST decoupling may exist over some sites as a response to decreasing precipitation. The ERA5-Land represents the observations' climatology well, but it underestimates the summer soil temperature observations and the long-term trends at some sites.

KEYWORDS

land–air temperature coupling, observations, soil temperature, soil thermal diffusivity

1 | INTRODUCTION

Soil temperature is a key environmental variable that is controlled by the energy and water exchanges at the land surface and the thermal properties of the soil. The surface energy balance establishes a link between the atmosphere and the ground. This generally translates into a strong connection between the state of the

atmosphere and the land (Bonan, 2015) that result in a coupling between the surface air temperature (SAT) and the ground surface temperature (GST). In turn, the GST affects the amount of heat flowing into/out-of the soil, thus shaping the distribution of temperatures below the surface.

The study of the subsurface thermal structure and its connection to the atmosphere thermal state has been an important field of research in climate science because of the influence of the near-surface soil thermal regime on water, heat and carbon exchanges (Steinert *et al.*, 2021) and biogeochemical processes like soil carbon stability (Zhang *et al.*, 2016). Likewise, a branch of paleoclimate studies assume that subsurface temperatures record a signal of the past surface temperature variations as a result of the air–ground thermal connection and the conductive propagation of surface temperature changes into the subsurface (Cuesta-Valero *et al.*, 2021). Understanding the connection between the atmosphere and subsurface thermal state is also important for future climate change assessments. It is expected that soil temperature will increase over the 21st century as a response to atmospheric warming at global (Soong *et al.*, 2020) and regional (Araghi *et al.*, 2019) scales. This may potentially enhance the carbon release from soils (Crowther *et al.*, 2016) and related climate feedbacks, or produce drier soil conditions due to accelerated evaporation rates with implications in the occurrence of extreme heatwaves and drought events (Guerreiro *et al.*, 2018; Chen *et al.*, 2021).

In situ soil temperature observations allow exploring the air–ground thermal connection. For instance, the downward propagation of surface temperature changes have been explored with the use of soil temperature observations. Smerdon *et al.* (2004) showed that the downward propagation of surface temperature variations is effectively explained by the laws of heat conduction that entail an exponential decay and linear phase shift of surface temperature changes as a function of depth (Carslaw and Jaeger, 1959). This is supported by other works that yield similar evidence of the prevalence of thermal conduction on controlling the distribution of temperature in the subsurface (e.g., Putnam and Chapman, 1996; Schmidt *et al.*, 2001; Baker and Baker, 2002; An *et al.*, 2016; Tong *et al.*, 2017).

Similarly, land–atmosphere interactions can be analysed with the use of soil and air temperature observations. This is important because the bond between SAT and GST may be altered by land surface changes altering the ground surface energy balance. These changes occur at different timescales, daily, seasonal cycles, decadal, centennial and longer-term variations as a response to external forcings (Melo-Aguilar *et al.*, 2018). In such cases, the temperature signal that propagates into the subsurface may be decoupled from the temperature variations in the overlying atmosphere. Smerdon *et al.* (2003) showed that in areas with large amount of snow cover, the coupling between the air and soil temperatures at seasonal scales is interrupted because of the insulating effect of snow cover. SAT and GST decoupling may also develop at interannual to decadal timescales in response to changes in snow cover seasonality (Beltrami and Kellman, 2003; Bartlett *et al.*, 2004; García-García *et al.*, 2019; Soong *et al.*, 2020). Other factors can also lead to short-term SAT–GST decoupling. For example, the type of vegetation cover may produce large differences between SAT and GST with the GST generally warmer than the SAT on daily and annual cycles (Cermak *et al.*, 2017).

The use of soil temperature data has provided insights into the heat transfer mechanisms within the subsurface and the role of surface processes on the SAT–GST coupling at short timescales. However, most of these works have been carried out at a few locations in North America because of the lack in ST observations in other parts of the world.

In the recent decades, soil temperature databases with local or regional coverage around the globe have been developed. For instance, Hu and Feng (2003) developed a soil temperature database for the U.S. including five soil layers down to a depth of 1 m. Likewise, the U.S. Climate Reference Network (USCRN; Bell *et al.*, 2013) and the Soil Climate Analysis Network (SCAN; Schaefer and Paetzold, 2000) includes soil-temperature records since the late 20th century. Some regional soil temperature observational datasets have also been developed in other parts of the world such as Canada (Qian *et al.*, 2011), China (Zhang *et al.*, 2016) and Russia (Zhang and Barry, 2000).

Despite the current developments in soil temperature datasets around the globe, there are still limitations in their spatial and temporal availability (Hao *et al.*, 2014; Huang *et al.*, 2020) as ST has received less attention than other meteorological variables (e.g., air temperature and precipitation; Qian *et al.*, 2011). Therefore, expanding the spatial and temporal coverage of soil temperature databases is of great interest to improve our understanding of processes that influence the exchange of energy between the atmosphere and the subsurface thermal state, the soil thermodynamics and the response to atmospheric warming. This is relevant for several climate studies, including agricultural applications, CO_2 release from soils to the atmosphere (Hicks Pries *et al.*, 2017) and even paleoclimate reconstructions.

In this work, we have developed a database of Soil Temperature Observations for Spain (SoTOS) that includes records for five soil layers down to a depth of 1 m and surface air temperature (2 m above ground) at 39 observatories spanning from 1987 to 2018. We have performed a data quality control (QC) in order to eliminate erroneous measurements and homogenize the

TABLE 1 List of AEMET weather stations included in SoTOS

Site	Code	Altitude (m.a.s.l.)	Longitude (°)	Latitude (°)	Location
1	4642E	19	353.0903	37.2800	Huelva (Ronda este)
2	7178I	62	358.8306	38.0028	Murcia
3	4452	185	353.1708	38.8833	Badajoz (Talavera, base aérea)
4	4116I	622	356.2597	38.9508	Almagro (Famet)
5	4121	627	356.0803	38.9894	Ciudad Real
6	8178D	674	358.1392	39.0069	Albacete (Observatorio)
7	3469A	405	353.6606	39.4722	Cáceres (Carretera Trujillo)
8	B691	40	3.0197	39.7464	La Puebla
9	3260B	516	355.9506	39.8847	Toledo (Buenavista)
10	8096	956	357.8619	40.0667	Cuenca
11	8368U	900	358.8772	40.3517	Teruel
12	2444	1130	355.3214	40.6500	Ávila (Observatorio)
13	3168C	635	356.8258	40.6611	Guadalajara (El Serranillo)
14	2462	1890	355.9897	40.7806	Navacerrada (Puerto)
15	2465	1005	355.8728	40.9478	Segovia (Observatorio)
16	2867	790	354.5039	40.9456	Salamanca (Matacán)
17	0200E	412	2.1253	41.4194	Barcelona (Fabra)
18	2614	656	354.2664	41.5167	Zamora (Observatorio)
19	0222	168	2.1689	41.6139	Caldes de Montbui
20	9771C	192	0.5950	41.6258	Lleida (Observatorio)
21	2422	735	355.2333	41.6500	Valladolid (Observatorio)
22	9434	247	358.9919	41.6619	Zaragoza (Aeropuerto)
23	2539	846	355.1500	41.7000	Valladolid (Villanubla)
24	0149D	280	1.8392	41.7211	Manresa (La Culla)
25	2030	1082	357.5333	41.7667	Soria (Observatorio)
26	0341	611	2.2278	41.8478	Tona (Escola)
27	1495	255	351.3680	42.2236	Vigo (Peinador)
28	1690A	143	352.1397	42.3278	Ourense (Granxa Deputación)
29	2331	890	356.3675	42.3561	Burgos (Villafría)
30	1484C	107	351.3836	42.4400	Pontevedra (Mourente)
31	1549	534	353.4000	42.5639	Ponferrada
32	2661	916	354.3506	42.5889	León (Virgen del Camino)
33	9091O	508	357.2772	42.8839	Vitoria
34	1428	364	351.5731	42.8994	Santiago de Compostela
35	1505	444	352.5439	43.1153	Rozas (Aeródromo)
36	1024E	252	357.9605	43.3075	San Sebastián (Igueldo)
37	1249I	336	354.1267	43.3536	Oviedo (El Cristo)
38	1387	58	351.5806	43.3672	A Coruña
39	1111	52	356.2003	43.4917	Santander

dataset. We use SoTOS to explore the shallow subsurface thermal regime via the evaluation of the downward conductive heat propagation (i.e., soil surface temperature changes characteristics). Subsequently, we address the coupling between SAT and GST at seasonal to multidecadal timescales.

Additionally, we compare SoTOS and the ERA5-Land reanalysis (ERA5-L; Muñoz Sabater, 2019). The comparison

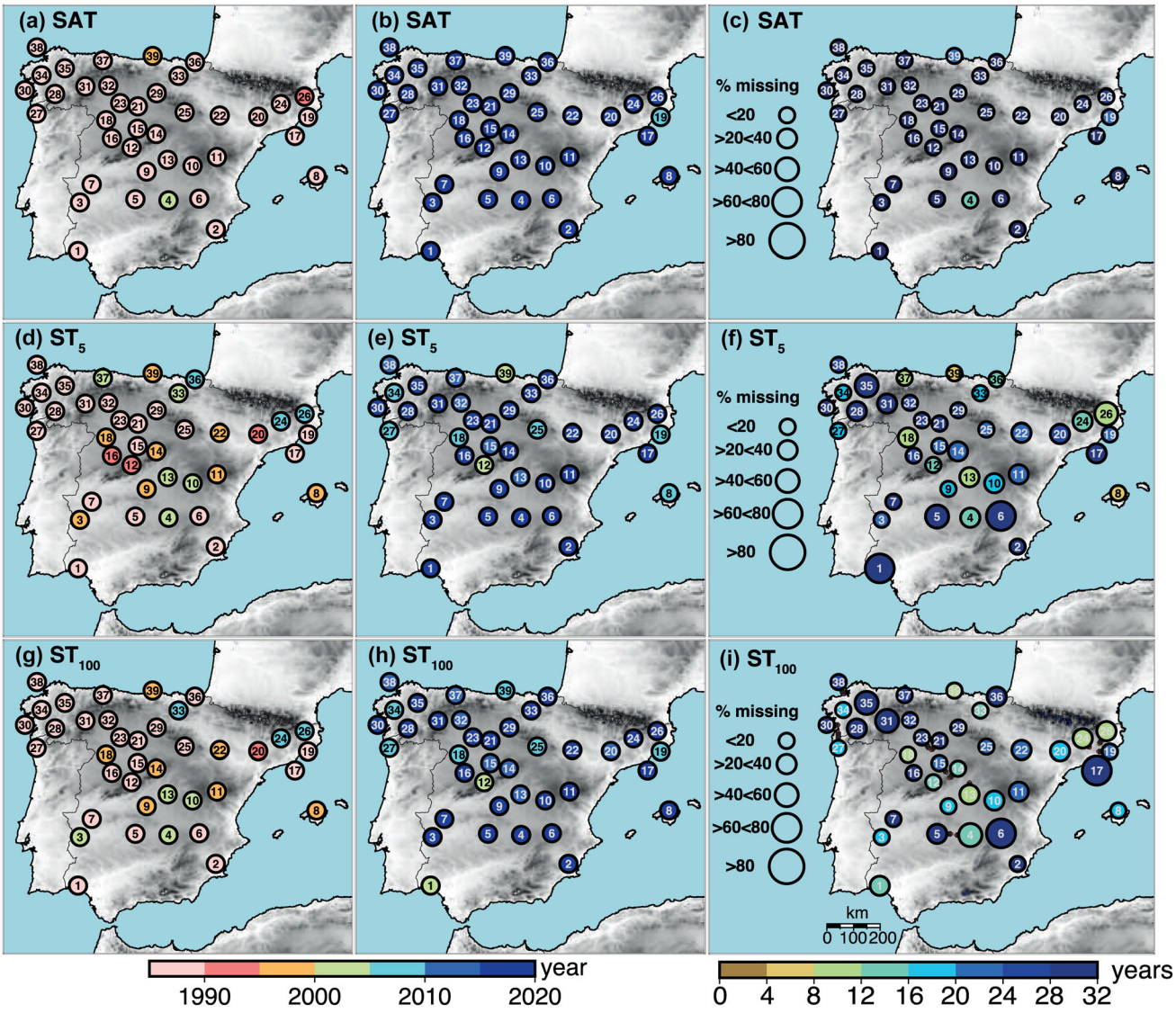

FIGURE 1 Initial (left) and final (middle) date of records for each of the 39 sites considered in SoTOS. Percentage of missing values relative to the time interval between the initial final dates of observations (right). Surface air temperature (SAT) and soil temperature (ST) at 5 and 100 cm are shown as examples (top to bottom). The arrows at some sites (e.g., 4, 5) point to the actual location while the symbol has been moved to avoid overlap [Colour figure can be viewed at wileyonlinelibrary.com]

provides information regarding the consistency of ERA5-L for representing the subsurface thermal state and the land–air interactions from local to regional scales in Spain. This analysis is relevant as ERA5-L is employed in climate research, in operational weather prediction, drought, flood risk management and soil moisture feedbacks on climate (Gallego-Elvira *et al.*, 2016).

2 | DATA

2.1 | Soil temperature observational dataset for Spain

A set of soil temperature (ST) and surface air temperature (2 m above ground; SAT) observations were collected from the Spanish Meteorological Agency (Agencia Estatal de Meteorología; AEMET) for a total of 39 sites over the Iberian Peninsula and the Balearic Islands spanning from 1987 to 2018 (Table 1). The ST data are measured at 5, 10, 20, 50 and 100 cm depth (ST_5, ST_{10}, ST_{20}, ST_{50} and ST_{100}, hereafter). The SAT and ST_5, ST_{10}, ST_{20} data are provided on intradaily resolution with measurements available at 0000, 0700, 1300 and 1800 GMT; for SAT, daily maximum and minimum values are also available. For the ST_{50} and ST_{100}, only one observation per day is taken at 1800 GMT.

The SAT data present a relatively homogeneous distribution of initial and final dates of available records in 1985 and 2018, respectively (Figure 1). The length of available ST data is highly variable among sites and for the different soil layers due to difference in the initial

and ending dates for ST records (Figure 1d,e,g,h). Thus, the ST data do not cover the full 1987–2018 period at all sites. In addition, missing data also affect the availability of records (Figure 1c,f,i), while data are significantly reduced for both ST_5 and ST_{100}. In a few cases, the missing data can reach up to 60–80%.

For the analyses developed in this study, daily averages are estimated from the available intradaily records. For the SAT, the daily mean is obtained as the average between the daily maximum and minimum values since both of these records are available. This is a standard procedure to obtain daily average in meteorological temperature data (WMO, 2018). For ST_5, ST_{10} and ST_{20} daily maximum and minimum values are not available. Therefore, the daily mean is calculated using the minimum and maximum values among the four available daily measures. The minimum temperature value is regularly captured at 0700 GMT for soil layers close to the surface, whereas the maximum value is recorded between the 1300 and 1800 depending on both depth and season.

2.1.1 | Quality control

We applied a QC to the original intradaily data and the SAT daily maximum and minimum values before estimating the daily averages. The QC consisted of a two-step process that identifies suspicious data that are flagged as measurement errors. The erroneous measurements were corrected when possible or discarded otherwise.

The first step detects extreme values that are measurement errors with a high level of confidence. To do so, we established two criteria: (a) unrealistically large or small temperature values based on absolute values above 50°C and below −50°C (Gonzalez-Hidalgo *et al.*, 2015) and (b) spikes in temperature data. The spikes are identified by taking the differences between three consecutive measurements, centered on each of the data points. The smallest absolute difference is considered and compared with a threshold value. If the difference exceeds this threshold, the data point is flagged (spike) and deleted (Reek *et al.*, 1992). Imposing a correct limit to determine whether a certain data point is an error is crucial to avoid eliminating extreme values that are due to meteorological extremes. We imposed a threshold of 10°C for maximum differences representing a conservative approach. Figure 2a illustrates results of the internal consistency check for site-16. The erroneous data are flagged (red points) and deleted.

In the second step, other situations that affect the quality of temperature data, such as changes in the average or trends over some periods of the time series, that is, inhomogeneities (González-Rouco *et al.*, 2001), are identified. Hu and Feng (2003) showed that, in some cases,

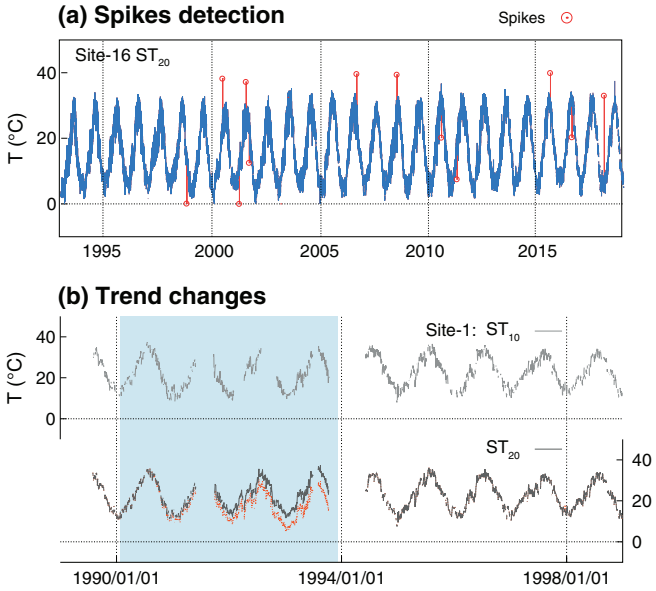

FIGURE 2 Example of the spikes detection for the ST20 in site 16 (a) and trend changes in time series for site 1 (b) as an example. Red circles in (a) stand for erroneous flagged and corrected data. The shaded area in (b) highlights the period 1989–1994 in which the ST20 presents a change in the trend that is not evident in the ST at the other available layers. In this case, ST10 is used as a reference for the correction of the change in the ST20 [Colour figure can be viewed at wileyonlinelibrary.com]

changes in the average/trend in some of the ST layers can be corrected. This is possible if such a behaviour is present over short periods of time and if there is a reference ST from a different soil layer that is not affected by this situation. As it is assumed that soil temperature at different soil layers is a function of the downward propagation of the surface temperature signal through mostly conductive heat transport, the ST at different soil layers contains a similar long-term behaviour (trend) provided that shallow depths are considered. Thus, the reference depth trend can be used to estimate the corrected data for the segment in the affected one (Hu and Feng, 2003),

$$ST_e = ST_c - (a_c + b_c t) + (a_r + b_r t) - \Delta ST_{rc}, \quad (1)$$

where ST_e and ST_c represent the estimated and the original ST in the affected segment, respectively, t is the time in days, and $a_c + b_c t$ and $a_r + b_r t$ represent the linear fit to the data for the affected ST and the reference ST in the affected segment, with a and b as the regression coefficients. ΔST_{rc} is the difference in the temperature average between the affected and reference ST, excluding the affected segment. ΔST_{rc} is included to account for the temperature variations with depth.

Equation (1) is employed to estimate the corrected data for the segment of the ST_{20} (shaded region; 1990–

TABLE 2 Summary of the corrected ST data according to the changes in trend or amplitude over some segments of the time series, indicating the site number (see Table 1), the corrected and reference ST layers, the time interval of corrected records and the percentage of corrected records

	Site	Corrected ST layer	Reference ST layer	Corrected period	% of data
Changes in trend	1	ST_{20}	ST_5	Aug 31, 1990–Feb 28, 1994	20.8
	1	ST_{100}	ST_{50}	Aug 4, 1992–Oct 31, 1997	55.4
	5	ST_{100}	ST_{50}	Jan 1, 1987–Dec 31, 1988	8.8
	8	ST_{100}	ST_{50}	Jul 1, 1998–Feb 28, 2003	30.9
	23	ST_{100}	ST_{50}	Jan 1, 2005–Dec 31, 2007	9.3
	27	ST_{100}	ST_{50}	Jan 1, 1989–Dec 31, 1995	38.7
	36	ST_{20}	ST_{10}	Jan 1, 1995–Mar 31, 1999	13.2
Changes in amplitude	29	ST_{20}–ST_{100}		Jan 1, 2002–Jan 31, 2008	29

1994) in the example depicted in Figure 2b. In this case, ST_{10} is used as the reference temperature. A summary of the corrected segments and the percentage of fixed records relative to the total length of the series is shown in Table 2.

Similar to the changes in the average of time series, there are cases with changes in the amplitude over segments of time series that also impact the quality of the data. These types of errors can be detected by the cumulative sum of squares (Peterson *et al.*, 1998). In some cases, it is possible to obtain corrected ST data for the segments that exhibit such type of error (Hu and Feng, 2003). To do so, the segment of the time series with the correct amplitude is used to rescale the amplitude of the segment that depicts this type of error. This procedure was only applied for site-29, in which the original data are affected from 2002 to 2008 at all soil layers (Table 2).

2.2 | Reanalysis data

This study also considers the ERA5-L reanalysis that provides simulated fields of land surface states. The ERA5-L runs the stand-alone H-TESSEL land surface model (IFS, 2018), that is, not coupled to the atmospheric or oceanic components. It is forced by atmospheric variables from the ERA5 atmospheric reanalysis (Hersbach *et al.*, 2020). ERA5-L has a spatial resolution of 9 km, with hourly temporal frequency output covering the period from 1981 to present. The H-TESSEL is discretized into four layers (Table 3) down to a depth of 2.89 m. The soil temperatures are defined at full layers while the heat fluxes are defined at the interface between layer i and $i+1$. The model soil layer depths do not exactly match those of the observations, however, the vertical distribution of soil levels in the H-TESSEL land surface model is comparable to those in SoTOS. To allow for comparison, the ERA5-L soil layers are linearly interpolated to the SoTOS layers.

TABLE 3 Soil layers and node depths in the H-TESSEL land surface model. The node depth at which the soil temperatures are defined corresponds to the middle of each layer

Layer	Layer depth (m)	Node depth (m)
L1	0–0.07	0.035
L2	0.07–0.28	0.175
L3	0.28–1	0.64
L4	1–2.89	1.945

We use soil temperature data from ERA5-L over the region between 35°–45°N and 10°W–5°E, for the 1985–2018 period at monthly resolution. Other variables such as volumetric soil water content and 2 m air temperature, are also considered for some analyses developed herein.

3 | METHODS

3.1 | Subsurface thermal regime

To assess the shallow subsurface thermal regime, we evaluate the propagation of surface temperature variations into the subsurface. This process is controlled by the one-dimensional time-dependent heat conduction equation (Carslaw and Jaeger, 1959),

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial z^2}, \tag{2}$$

where κ is the thermal diffusivity of the medium, and z and t stand for depth and time, respectively. Assuming that the temperature variations at the surface can be represented by simple harmonic variations and considering

a homogeneous subsurface with no internal heat production, the solution of Equation (2) as a function of time and depth takes the form,

$$T(z,t)=Ae^{\left(-z\sqrt{\pi/\tau\kappa}\right)}\cos\left(\frac{2\pi}{\tau}t-z\sqrt{\pi/\tau\kappa}\right), \quad (3)$$

where A and τ represent the amplitude of the temperature change and the period of oscillation, respectively. Equation (3) shows that the amplitude of the temperature variations is attenuated exponentially with depth. Likewise, there is a linear phase shift ($\phi=z\sqrt{\pi/\tau\kappa}$) with depth relative to the initial phase of the oscillation at the surface.

We evaluated these characteristics in the 39 SoTOS sites. The amplitude and phase of the annual signal from the daily soil temperature measurements at 5, 10, 20, 50 and 100 cm depth are extracted using linear least-squares fitting of the data to a sine-wave model of the form,

$$T(t)=A\sin\left(\frac{2\pi}{\tau}t-\phi\right). \quad (4)$$

The annual signal is employed as a reference because the seasonal cycle is closely approximated to a simple-harmonic wave. Likewise, the annual signal is traceable within the upper meter of the soil, whereas higher-frequency oscillations (e.g., diurnal cycle) are completely attenuated within the first meter (Putnam and Chapman, 1996). The first soil layer (ST_5) is used as a reference to estimate the phase shift of the subsequent layers (i.e., ST_{10}, ST_{20}, ST_{50} and ST_{100}). Once the amplitude and phase shift of the annual signal have been obtained for all of the available soil layers, a linear regression of both the natural logarithm of the amplitude and phase shift versus depth is performed. In either case, the regression slope allows estimating an apparent thermal diffusivity (κ) as the propagation of the signal depends on κ and the period of oscillation (see Equation (3)). The apparent thermal diffusivity can be thought of as an estimation of the conductive propagation of the surface temperature variations in the shallow subsurface. SoTOS allows for evaluating the subsurface conductive heat transport assumption for various sites with different climatological and soil thermal properties, thus expanding on the analysis of previous works (e.g., Smerdon *et al.*, 2004).

For the implementation of this procedure, it is convenient that temperature records for all soil layers at individual locations have the same time span. All SoTOS sites contain enough information to ensure that soil temperature records are available during a common time period at all soil layers. However, the specific period of data availability may differ from site to site due to the variability in the temporal coverage of records for the different sites (Figure 1). We have imposed a condition of a minimum of 5 years of continuous data for all soil layers at each of the sites so that the time span of available data is sufficient to provide reliable results.

3.2 | SAT–GST coupling

To explore the relationship between SAT and GST at interannual to decadal timescales, we employ the SAT and the ST_5 variables. The latter is considered as the GST since this is the closest available level to the surface. First, the Pearson's correlation coefficient between SAT and GST is estimated over the whole available period using monthly anomalies (i.e., annual cycle removed by subtracting the corresponding monthly long-term average from each monthly value). Also, the differences between the SAT and the GST mean values are evaluated as this provides information about the energy exchange across the air–ground interface (García-García *et al.*, 2019). Second, the temporal evolution of the SAT–GST relationship is explored by estimating the SAT and GST linear trends over the 1988–2018 period. This is done for selected sites that include SAT and GST information for the complete period. Such an assessment allows identifying potential SAT–GST decoupling processes on interannual to decadal timescales. The SAT–GST relationship is addressed on seasonal and annual averages to explore the variability in the energy exchanges at the surface across the year.

All of the analyses described for SoTOS are also implemented using the ERA5-L dataset. We describe the consistency between observations and reanalysis throughout the article.

4 | RESULTS AND DISCUSSION

4.1 | SAT and ST climatologies

Within the 1987–2018 period, the mean annual SAT is above 10°C for the bulk of the SoTOS sites (Figure 3). The highest values of approximately 20°C are recorded over the southernmost part of the Iberian Peninsula (e.g., sites 1 and 2) with a gradual decrease towards northern Spain. However, there are some areas over northeast Spain, the Ebro Valley and the Balearic Islands in which the annual mean temperature can be as high as in the southern half of the Iberian Peninsula (>15°C; sites 17, 19, 20 and 22). The northern Inner Plateau

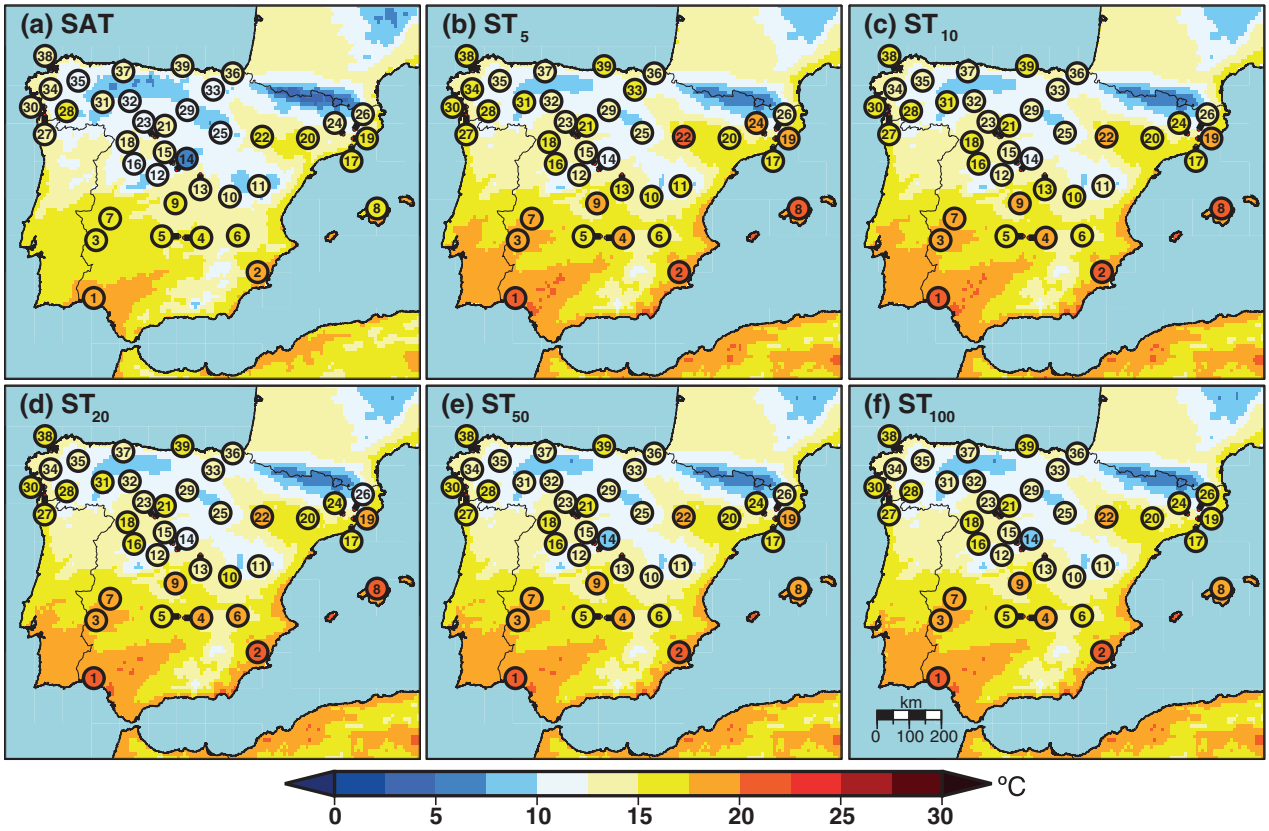

FIGURE 3 Mean annual temperature for SAT, ST5, ST10, ST20, ST50 and ST100 (from a to f, respectively). The period employed to obtain the mean values may differ from site to site because of the differences in the available data for both sites and variables. The mean temperature from the ERA5-L reanalysis is shown in the background for the 1985–2018 period. The soil layers of ERA5-L have been linearly interpolated to the SoTOS depths [Colour figure can be viewed at wileyonlinelibrary.com]

exhibits, on average, the coldest annual SAT with temperatures around 11°C. The minimum annual SAT (6.2°C) is observed at high elevation in the Sierra de Guadarrama (site 14). The ERA5-L represents the mean SAT of the SoTOS sites. Likewise, it shows a consistent spatial distribution of temperatures in agreement with the observations.

STs at all depths show a similar spatial distribution as SAT in the observations and the reanalysis. Interestingly, mean STs are warmer than SAT on average, which is more noticeable in the SoTOS sites. Indeed, the SoTOS sites show slightly higher STs than ERA5-L.

To provide a metric regarding the consistency between ERA5-L and the observations, Taylor diagrams (Taylor, 2001) are developed for the annual averages and for the Northern Hemisphere winter (DJF) and summer (JJA) seasons (Figure 4). This evaluation is carried out by selecting the nearest grid points (co-located) in the ERA5-L grid to the observational sites. For SAT, ERA5-L reproduces the observations with a correlation above 0.9 ($p < 0.05$) and standard deviation ratios distributed around 1 for the bulk of the sites both in the annual case and the JJA and DJF seasons (Figure 4a,d). ERA5-L shows lower skill reproducing the ST_5 observations as the correlation coefficients decrease for the annual average in comparison to SAT, although most of them still have high correlation (>0.7). Similarly, the standard deviation ratios are systematically below 1 revealing that ERA5-L underestimates the variability of observations. The JJA and DJF cases suggest that most of the annual average variability for ST_5 is influenced by the JJA season as the data in this season are spread towards lower correlation coefficients and standard deviation ratios below 1. Whereas in DJF, correlation is higher and standard deviation ratios cluster around 1 in most of the cases. A similar picture to the ST_5 is depicted for the ST_{100}. The correlation coefficients range from 0.4 to 0.9 in the annual data, and the standard deviation ratios are distributed below 1. The JJA and DJF seasons show a wider range of correlation values ranging from 0.2 to 0.95 and standard deviation ratios below the 0.5 arc for the JJA season. The reduction in the ability of the ERA5-L to

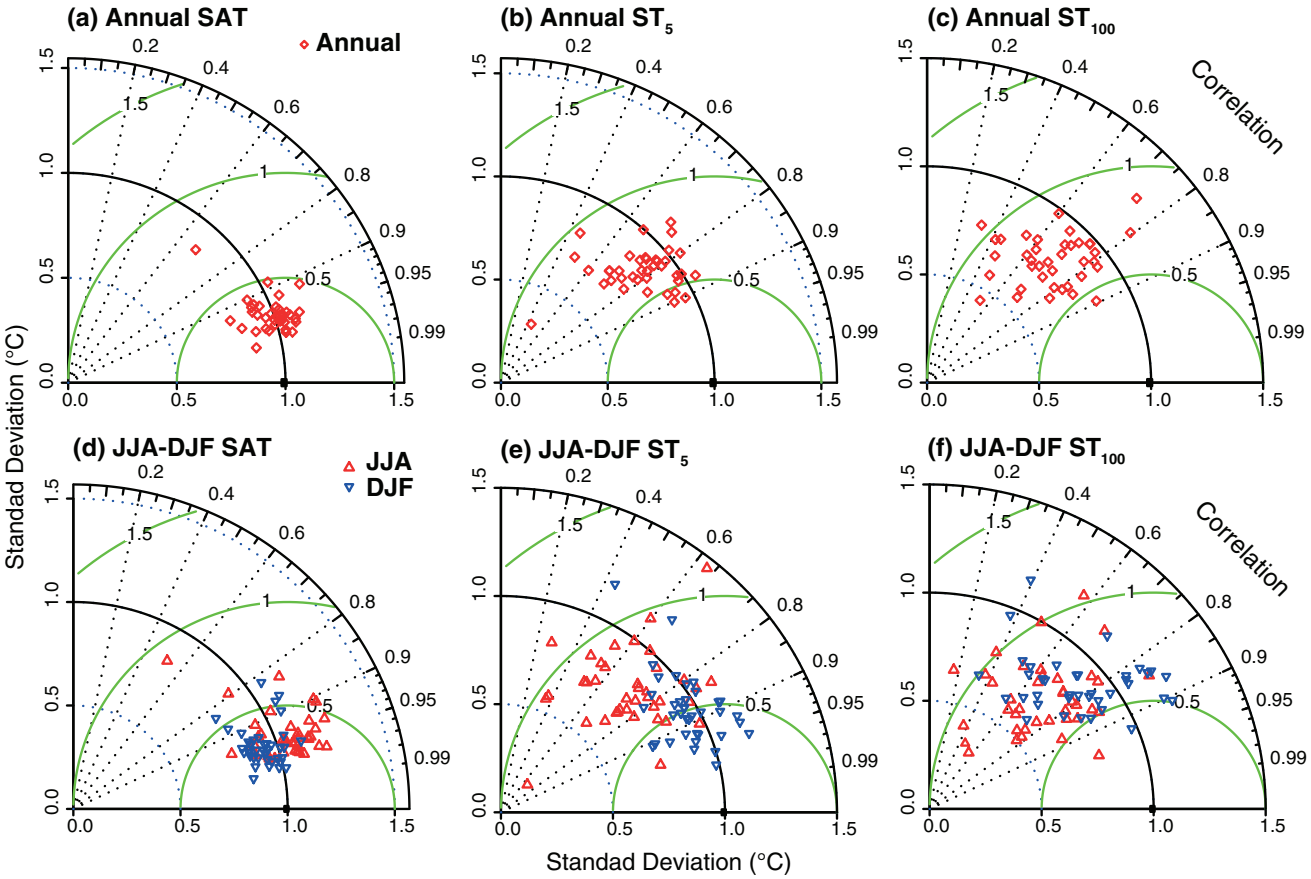

FIGURE 4 Taylor diagrams of the monthly anomalies at each SoTOS site for the annual SAT (a), ST5 (b), ST100 (c) as well as JJA and DJF for SAT (d), ST5 (e) and ST100 (f). Taylor diagrams show the correlation coefficient on a clockwise angle scale. The radial coordinate gives the standard deviation ratio: high correlation coefficients and standard deviation ratio around 1 indicate low RMSE values (concentric green circles) [Colour figure can be viewed at wileyonlinelibrary.com]

reproduce the soil thermal state over the SoTOS sites may be related to issues such as the model spatial resolution and the interpolation process (Yang and Zhang, 2018). It may also be influenced by assimilating only surface data (Balsamo *et al.*, 2015) or to the use of a shallow bottom boundary condition placement in the model (Smerdon and Stieglitz, 2006). The lower fit in JJA may be related to turbulent heat fluxes due to interactions with vegetation (Johannsen *et al.*, 2019).

4.2 | Subsurface thermal regime

4.2.1 | Conductive heat transport

Figure 5a illustrates the time series of daily SAT and ST data for site 2. Note that there may be gaps in the ST data. Despite such gaps, there are some periods in which data are available at all depths (e.g., 1997–2015). These periods of corresponding coverage of ST data permit the extraction of both the amplitude and phase of the different soil layers in a homogeneous frame, so that the analysis of the conductive heat transport can be performed. All the SoTOS sites contain ST information at all depths for corresponding periods of at least five consecutive years (not shown).

The amplitude and phase of the annual signal for SAT and ST at all depths are extracted using a linear least-squares fitting as described in section 3. This information is employed to represent the annual cycle (Figure 5b). Interestingly, the SAT annual signal shows a lower amplitude relative to the ST at the upper layers. This feature is common to the majority of the SoTOS sites. The nature of such a response is discussed in section 4.3.1. A simple visual inspection reveals the amplitude attenuation and phase shift with depth of the downward propagating wave, visible, for example, in ST_{100} relative to the upper layers. This, qualitatively illustrates the characteristic of the conductive heat transport of the surface temperature signal into the soil.

To provide a quantitative measure on the conductive-dominated shallow soil thermal regime, a linear

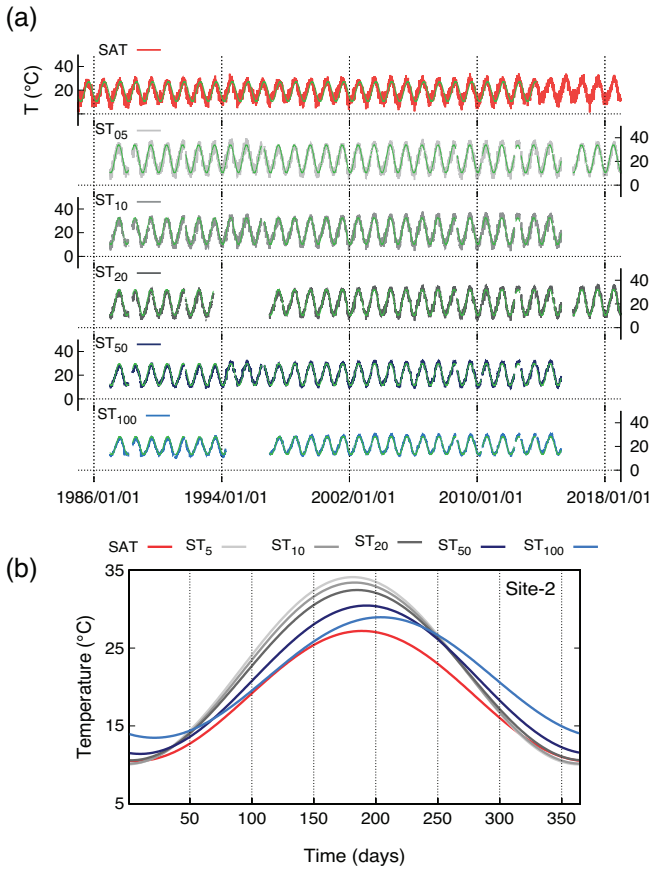

FIGURE 5 Final quality controlled daily SAT and ST5, ST10, ST20, ST50 and ST100 for site 2 as an example. The annual cycle for SAT and ST at all depths is shown in (b) [Colour figure can be viewed at wileyonlinelibrary.com]

regression model is applied to the natural logarithm of the amplitude (Ln-A, hereafter) and the phase shift (P-S, hereafter) versus depth (Figure 6). The regression lines show the linear dependence of both Ln-A and P-S with depth. In all SoTOS cases, the r^2 of either the Ln-A or P-S is above 0.95. The resulting slope is used to estimate an apparent thermal diffusivity of the shallow subsurface, as an average over the entire year across the subsurface's upper meter (indicated in each of the panels), and to obtain a metric of conductive heat transport (see section 3).

The resulting thermal diffusivity obtained from the Ln-A and P-S can often be different for the SoTOS sites. In some cases, such difference is relatively small within the range of a few tenths of $1\times10^{-7}\mathrm{m}^2\cdot\mathrm{s}^{-1}$ (e.g., site 2), while some other cases show larger differences (e.g., site 21). In previous works that have used a similar approach to estimate the subsurface apparent thermal diffusivity, the values obtained from either the Ln-A or the P-S closely coincide (e.g., Smerdon *et al.*, 2003; 2004). However, SoTOS differs from those case studies by including ST observations only for the subsurface upper meter, whereas in other works, the soil temperature profiles go down to at least 3 m depth and reach >10 m. In such cases, the bottom part of the soil profiles is less affected by soil hydrology-related processes (e.g., evapotranspiration) that may have a considerable influence on altering the pure conductive regime (see section 4.2.2). Gao *et al.* (2008) found that, for the diurnal cycle, estimating the near-surface soil thermal diffusivity from the amplitude attenuation or the phase shift with depth may yield slightly different results if only thermal conduction is considered. Likewise, Tong *et al.* (2017) showed that the estimation of the amplitude and phase of the daily cycle at different depths is highly sensitive to small changes in liquid water flux. Thus, neglecting nonconductive subsurface processes when estimating apparent thermal diffusivity might produce the differences in the apparent thermal diffusivity.

4.2.2 | Soil moisture influence on the apparent thermal diffusivity

Even though we have not considered the effect of soil hydrology in our estimation of the apparent thermal diffusivity, the influence of soil water content can be observed in the spatial distribution of estimated apparent thermal diffusivities (Figure 7). The results of the ERA5-L, shown in the background, are based on interpolated ST layers to the SoTOS depths. Thus, in both cases, the thermal diffusivity represents the mean value over the upper 1 m depth. The ERA5-L apparent thermal diffusivity yields lower values than those obtained from the observations using either the Ln-A or P-S. Therefore, the results have been scaled by a factor of 2 to improve spatial patterns visualization. There is a similar pattern in the observations and the reanalysis, with larger values over the northern coastal edge of the Bay of Biscay (e.g., sites 36, 37 and 39) and some other sites dispersed over the Peninsula (e.g., sites 3, 21 and 25). The ERA5-L soil moisture content (Figure 7c) reveals a strong influence of soil water content on the apparent thermal diffusivity. In general, larger thermal diffusivities coincide with larger soil water content. This relationship is especially evident in the Ln-A map (Figure 7a), but such a signal can also be detected in the P-S map (Figure 7b). Although other factors such as the soil texture and mineralogy, bulk density, salt concentration and organic matter content also influence soil thermal diffusivity (Abu-Hamdeh and Reeder, 2000), the soil water content accounts for a large part of the spatial variability in the apparent thermal diffusivity for the study area as soil water content affects the soil thermal conductivity (Patil *et al.*, 2011) and the volumetric heat capacity (McCumber

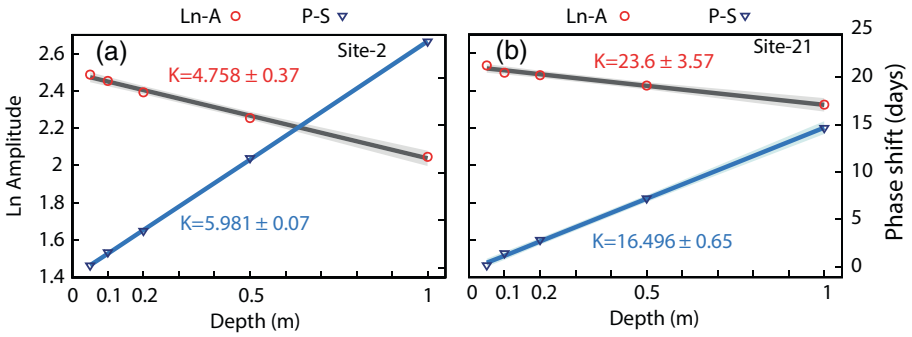

FIGURE 6 Linear regression of the natural logarithm of the amplitude (red circles) and phase shift of the annual signal of STs at the levels with available data (blue triangles) relative to the ST5 as a function of depth for sites 2 (a) and 21 (b) as an example. The shaded region indicates the 95% confidence interval. The estimated apparent thermal diffusivity from both the amplitude attenuation (grey) and phase shift (blue) are shown in each panel. The units for thermal diffusivity are given as ×10−7 $m^2 \cdot s^{-1}$ [Colour figure can be viewed at wileyonlinelibrary.com]

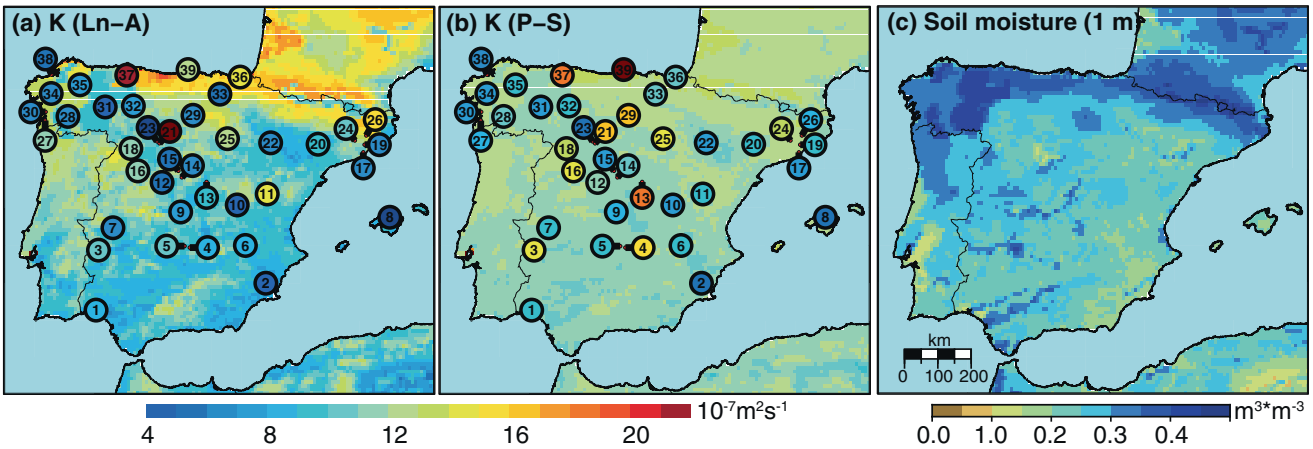

FIGURE 7 Apparent thermal diffusivity for the upper meter of the subsurface calculated from the slope of the linear regression of the amplitude attenuation (a) and phase shift (b) as a function of depth for the 39 SoTOS sites. The results from ERA5-L are shown in the background and scaled by a factor of 2 to enhance visibility. (c) ERA5-L average volumetric water content for the upper subsurface meter obtained as a weighted average over the layer thickness [Colour figure can be viewed at wileyonlinelibrary.com]

and Pielke, 1981). In the case of the ERA5-L data, there are also differences in the estimation of the apparent thermal diffusivity form the Ln-A and the P-S. In this case, besides the thermal effects related to phase changes, which are included in the H-TESSEL model, the errors in the representation of the amplitude attenuation and phase shift with depth due to having a shallow boundary condition at the bottom may also have an effect. It can be around 20% in amplitude and 5% in phase shift for the propagating of signals between 1 day and 1 year (Viterbo and Beljaars, 1995).

Another way to address the influence of soil water content on the shallow subsurface thermal regime, is to compare the evolution of the apparent thermal diffusivity and its relationship to soil moisture. Although no measurements of soil moisture are available for the SoTOS sites, the ERA5-L data can be used as an alternative. Likewise, observed precipitation may be used as a proxy for soil moisture as the spatio-temporal variability of soil moisture depends on precipitation (Sehler *et al.*, 2019). To estimate the variations with time in the apparent thermal diffusivity, we have selected the sites with the longest consecutive ST records at all depths. The thermal diffusivity is estimated for chunks of three consecutive years using a 1-year moving window. The same approach is applied to ERA5-L using the co-located grid points to the observational sites in SoTOS over the 1981–2018 period.

Figure 8 illustrates the evolution of the apparent thermal diffusivity from SoTOS and ERA5-L, the ERA5-L co-located grid point soil water content as the weighted average over the layer thickness for the upper meter of the soil and the observed precipitation for sites 28. The

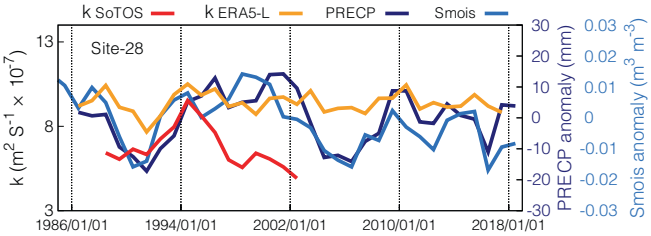

FIGURE 8 Evolution of thermal diffusivity in SoTOS (ERA5-L) estimated for chunks of three consecutive years within the 1994–2014 (1981–2018) period using a 1-year moving window. The thermal diffusivity is presented only for the ln-A analysis as described in Figure 6. The entire soil mode profile is employed to estimate the thermal diffusivity of ERA5-L. The reference depths to evaluate the relationship between amplitude attenuation and phase shift with depth are the midpoint of each model soil layer (i.e., 0.035, 0.175, 0.64 and 1.945 m depth). The estimated thermal diffusivity from ERA5-L is scaled by a factor of 2 for visibility. Precipitation (volumetric water content) is also shown over the full 1981–2018 period as an average of three consecutive years to match the thermal diffusivity estimations. The ERA5-L volumetric water content is presented as a weighted average over the entire soil profile (i.e., L1, L2, L3 and L4). Results are shown for site 28 as an example [Colour figure can be viewed at wileyonlinelibrary.com]

SoTOS thermal diffusivity spans from 1987 to 2003. Only the results from the Ln-A are depicted because in the reanalysis, the results from the P-S show less variability with time. Note that there is good agreement between ERA5-L soil moisture and precipitation with a Pearson correlation coefficient of 0.67 ($p < 0.05$). Similarly, the apparent thermal diffusivity from SoTOS and ERA5-L show similar variability ($r = 0.44$) and also to the variations in soil moisture and precipitation. The temporal variability in the apparent thermal diffusivity and soil moisture/precipitation suggests that the variation in soil water content influences the thermal diffusivity variability with time. The effect of soil water content on thermal conductivity is shown in a study carried out in the Czech Republic where the high-est thermal conductivities were measured in soils with relatively high soil water content (Kodešová *et al*., 2013).

Despite the differences between the Ln-A and P-S thermal diffusivity estimates, the resulting values, in either case, agree with in situ measurements in different parts of the world. For instance, measured thermal diffusivity of soil, in the area of Ogun state (Southwestern Nigeria) for the upper 1.5 m soil depth ranges from 3.46 to 7.5 × 10−7 $m^2 \cdot s^{-1}$ for sandy soils with relatively low soil moisture content (Oladunjoye and Sanuade, 2012). Similarly, Sáez Blázquez *et al*. (2017) found that soil thermal conductivity for the Avila region (Spain) ranges from ∼1.44 to 2.5 $W \cdot m^{-1} \cdot K^{-1}$ depending on the soil type and moisture conditions. Assuming typical values of density (ρ) and specific heat (c_p) for soil mineral and soil organic matter around 1,800 $kg \cdot m^{-3}$ and 1,400 $J \cdot kg^{-1} \cdot K^{-1}$ (Kluitenberg, 2002), these conductivity values would roughly correspond to thermal diffusivities of 6–10 × 10^{-7} $m^2 \cdot s^{-1}$ ($\kappa = \alpha / c_p \rho$). This range of variations is similar to the estimated thermal diffusivity from the analysis developed herein.

4.3 | Land–atmosphere interactions

4.3.1 | SAT–GST coupling

The SAT–GST coupling is usually assessed by analysing the differences between SAT and GST mean values as well as their covariance structure over a reference time period since this provides information on the interactions across the air–ground interface (Bartlett *et al*., 2004; González-Rouco *et al*., 2009; Melo-Aguilar *et al*., 2018). Figure 9 shows the SAT–GST differences (top) and the correlation coefficients (bottom) for DJF (left) and JJA (centre) seasons separately and for the annual average (right). For a few SoTOS sites (i.e., sites 6, 14, 26, 38 and 29), the existence of missing data in GST, either in DJF or JJA over relatively large periods, prevents calculating the correlation coefficients. Therefore, these sites are excluded from the analysis.

For DJF, there is a strong bond between SAT and GST as low nonsignificant differences and high correlation coefficients (>0.9) are calculated for most of the SoTOS sites. The ERA5-L data agree in representing the SAT-GST differences and correlations in DJF. The reanalysis correlations show a general distribution of values above 0.95 except in the mountain areas. Particularly, the low correlation coefficients over the Pyrenees are worth noting. A similar pattern with high negative SAT–GST differences is observed over this region, which is likely the response to the snow cover insulating effect that prevents the contact of the cold air with the ground surface (Melo-Aguilar *et al*., 2018).

The JJA maps show different behaviour compared to DJF since the majority of SoTOS sites show negative SAT–GST differences with some large values distributed over both the northern and the southern plateau and the Mediterranean and Atlantic coasts that can reach >7°C

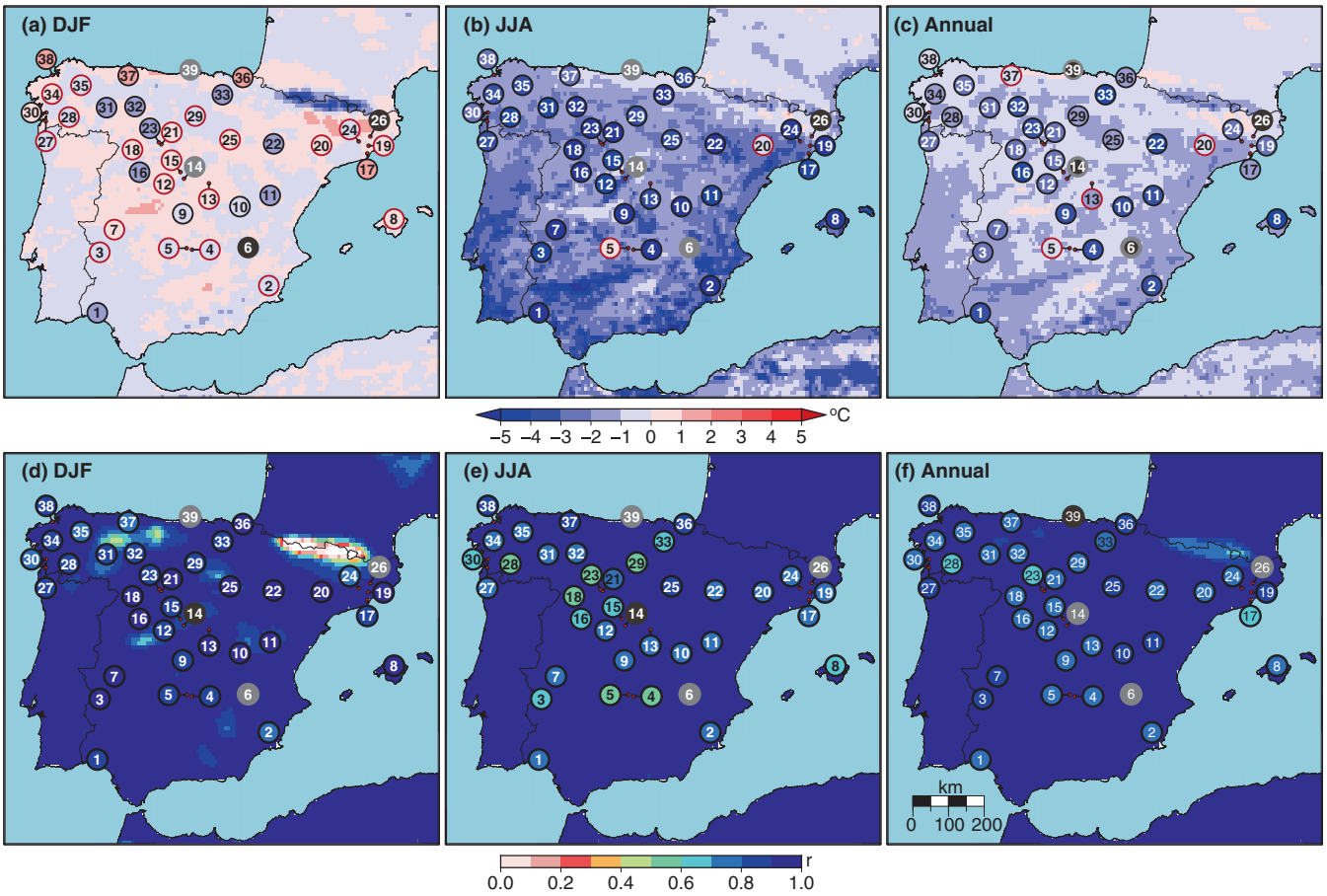

FIGURE 9 Mean SAT-GST differences at the 39 sites for (a) DJF, (b) JJA and (c) annual. Pearson correlation coefficients between SAT and GST for (d) DJF, (e) JJA and (f) annual. The period of reference to estimate the differences and correlations may differ from site to site, depending on the available data for each of the sites. Black (red) contour circles show statistical significant (nonsignificant) values ($p < 0.05$). To evaluate the significance, a *t* test has been applied and temporal autocorrelation has been considered. The grey-coloured sites represent those locations with no sufficient data to estimate the correlations. The SAT–GST differences are calculated from absolute temperature data using the original daily values while for the estimation of the correlation coefficients, monthly anomalies are employed. The results from the ERA5-L dataset are shown in the background [Colour figure can be viewed at wileyonlinelibrary.com]

in absolute value. Likewise, the correlation coefficients decrease for some of the SoTOS sites, although they are still high (>0.5) in all cases. ERA5-L shows good agreement with the SAT–GST differences in the summer months with negative differences that coincide with those of SoTOS. However, the SAT–GST differences are of smaller magnitude compared to SoTOS. Likewise, ERA5-L does not represent the lower correlation depicted by some SoTOS sites in this season. Other works have also found that ERA reanalysis products (e.g., ERA5-L, ERA5 and ERA-Interim/land; Balsamo *et al.*, 2015) tend to underestimate GST observations. Johannsen *et al.* (2019) showed that these products underestimate summertime daily maximum GST over the Iberian Peninsula. Similarly, a cold bias has also been reported over China, especially during summer (Yang and Zhang, 2018; Li *et al.*, 2020). Johannsen *et al.* (2019) found that the treatment of vegetation cover in the H-TESSEL may partially explain the underestimation of the GST since ERA5 overestimates the total vegetation cover over the Iberian Peninsula. The latter may lead to a higher surface–atmosphere coupling via turbulent exchanges dissipating more energy from the ground surface back to the atmosphere, especially in semi-arid regions. Indeed, the location of most of the SoTOS sites corresponds to bare soil, short grass and shrub surfaces (not shown). These land cover types have been shown to enhance the land surface radiative heating. A dry bare surface would warm more than a vegetated covered area and consequently have a larger SAT–GST offset (Cermak and Bodri, 2018).

The annual average indicates a larger influence from JJA since there is a general distribution of warmer GST relative to SAT for both SoTOS and ERA5-L with negative differences that range from 0 to −5°C. The lower

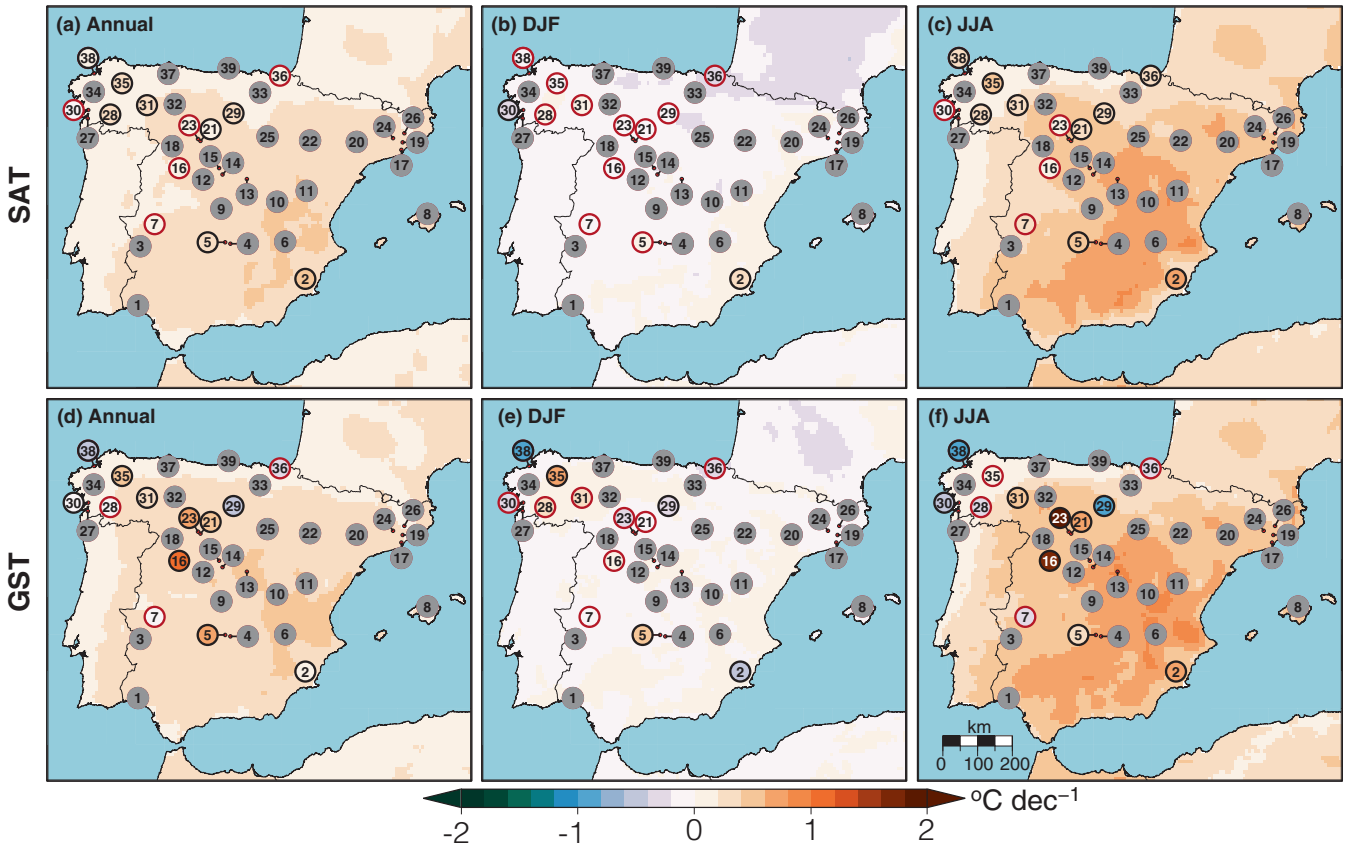

FIGURE 10 Temperature trends for the 1988–2018 period for SAT (top) and GST (right). Trends are calculated from monthly anomalies and indicated in °C·decade^{-1}. Black (red) contour circles indicate statistical significant (nonsignificant) values ($p < 0.05$). To evaluate the significance of trends, a *t* test has been applied. Temporal autocorrelation has been considered for the estimation of both the standard deviations of the regression residuals and the reduced degrees of freedom using a lag-1 autoregressive statistical model (Santer *et al*., 2000; Hartmann *et al*., 2013). Note that only selected SoTOS sites are coloured as these are the sites including GST information or the complete 1987–2018 period. The grey-coloured sites represent the locations that have no sufficient data to estimate the trends. The results from the ERA5-L dataset are shown in the background [Colour figure can be viewed at wileyonlinelibrary.com]

correlations over the Pyrenees depicted in DJF in the ERA5-L data are also observed in the annual average, thus, highlighting the influence of the winter season on the annual average over this area.

The combined analysis of SAT–GST differences and correlations indicates a stronger SAT–GST coupling in DJF compared to JJA over Spain. Larger SAT–GST offset in JJA relative to DJF has been reported by Putnam and Chapman (1996), who explored the SAT and GST relationship using observational data from one location in an arid region of northwest Utah. They showed that radiative heating of the ground surface during the summer months leads to large SAT–GST offset from March through October. They found that the annual cycle of GST minus SAT differences closely tracks the annual solar cycle. In the winter months, as the solar radiation decreases, the differences between SAT and GST average to zero. A similar physical response may explain the SAT–GST relationship over some sites included in SoTOS, especially those located over the Inner Plateau, the Ebro Valley and the coastal areas of the Mediterranean sea and the Atlantic Ocean. The latter is due to the high solar irradiance in the summer months (Sancho-Avila *et al*., 2012) and the low precipitation rates over the mentioned areas (Serrano-Notivoli *et al*., 2017). The soil moisture content from the ERA5-L provides additional support since the areas with low (high) water content in the first model layer closely coincide with the areas of high (low) SAT–GST differences during the summer months (not shown). As less water is available to evaporate, the energy that reaches the land surface is mostly balanced by sensible heat, thus reinforcing the radiative heating of the ground surface. Indeed, the JJA SAT–GST differences and volumetric water content for the first model layer have a spatial pattern correlation of 0.67.

4.3.2 | Interannual to decadal SAT–GST relationship

At this point, the question of seasonal effects on the SAT–GST relationship at longer timescales is remaining. If the seasonal SAT–GST offset does not change from

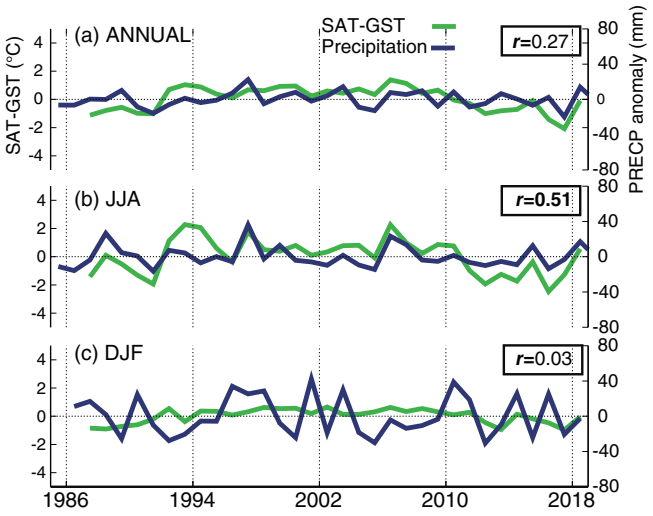

FIGURE 11 Evolution of SAT–GST and total precipitation anomaly over the 1985–2018 period for (a) annual, (b) JJA and (c) DJF mean for site 21. The Pearson correlation coefficient is indicated in each panel. Bold indicates significant correlations ($p < 0.05$) based on a t test [Colour figure can be viewed at wileyonlinelibrary.com]

year to year, the effect would be a constant bias with no consequences on the SAT–GST tracking at longer than annual timescales (Bartlett, 2005). However, if the relative influence of seasonal processes changes with time, the long-term SAT–GST coupling may change. In the present study, SoTOS allows for evaluating the SAT–GST relationship over a three-decade period for the sites with the most extensive length in observations for both SAT and GST. For those sites, the SAT and GST linear trends for the annual data and DJF and JJA seasons are estimated to assess their evolution at interannual to multidecadal timescales. In general, positive trends are observed in the SAT for the annual averages, with some SoTOS sites delivering statistically significant values (Figure 10). The GST trend maps also show a predominant picture of positive values. However, there are some differences relative to SAT for some of the SoTOS sites. For instance, a pattern of larger GST increase relative to SAT emerges over the Inner Plateau. Likewise, sites 29 and 30 show significant negative trends. The ERA5-L data also show widespread GST warming over Spain. However, it underestimates the larger GST warming compared to SAT shown by some of the SoTOS sites (e.g., 5, 16, 21 and 23).

The trend analysis for DJF and JJA provides insights into the mechanisms that control annual data response. While DJF shows negligible positive trends for both SAT and GST, the JJA months show widespread warming over the Peninsula, denoting a larger contribution to the annual averages. Likewise, in JJA, the same pattern as the annual case of larger GST warming relative to the SAT for some SoTOS sites is observed, particularly over the Inner Plateau. The latter is not fully captured by ERA5-L. The trend analysis based on a Mann–Kendall test (Mann, 1945; Kendall, 1975) shows consistent results (not shown) with the linear trends, thus providing confidence on the analysis.

A comparison between SAT and GST differences to precipitation reveals a strong influence of water availability on the SAT–GST relationship (Figure 11). Variations in summer precipitation, and also in SON season (not shown), highly correlate to SAT–GST differences. Further, decreasing precipitation over the last few decades controls the warmer GST response relative to SAT at this site on interannual to multidecadal timescales. This effect dominates the annual average response, thus leading to SAT–GST decoupling at these timescales. In DJF, there is a more stable evolution of SAT-GST differences. In this season, the low incoming energy limits evaporative processes, thus, variations in winter precipitation do not impact the SAT–GST relationship. Gallego-Elvira *et al.* (2016) evaluated the soil moisture control on land surface energy balance using satellite observations and reanalysis data. They found that the land surface warms faster than the overlying atmosphere under water-stressed conditions due to limitations on evaporation and enhanced sensible heat fluxes. Additionally, they showed that this effect is larger over short vegetation and bare soil areas than forested regions due to differences in aerodynamic resistance and hydrological dry spell behaviour. These are indeed the dominant land surface conditions of most of the SoTOS sites (not shown), thus, contributing to the larger GST increase compared to SAT.

The existence of changes in the SAT–GST relationship with time suggests some SAT–GST decoupling at interannual to multidecadal timescales. Nevertheless, these results should be interpreted with caution as trends based on short records are very sensitive to the beginning and end dates due to the natural variability and sometimes do not indicate an actual long-term trend (Hartmann *et al.*, 2013). Despite this, SoTOS shows some consistent patterns over the Inner Plateau that point to multidecadal SAT–GST decoupling.

5 | CONCLUSIONS

Here, a soil temperature observational dataset for Spain was developed including observations at five different soil layers (5, 10, 20, 50 and 100 cm depth) at 39 meteorological stations where SAT records also exist. We performed a quality control procedure which resulted in the deletion of 6,193 erroneous records (0.19%)

from a total of 5×10^6 measurements. Likewise, 23,817 (0.47%) records were corrected either from changes in the trend or in variance over segments of the series. The quality control procedure improved data quality but is not necessarily error free. Future work may incorporate additional corrections. The final dataset is presented at a daily temporal resolution spanning from 1987 to 2018.

The development of SoTOS offers a unique opportunity to evaluate the performance of regional-climate model simulations on representing the temporal and spatial variability of soil temperatures. This is relevant because soil temperature is a key variable that influences the representation of many physical processes in the atmosphere and below the surface in numerical weather prediction, regional modelling and land-reanalysis approaches. A more realistic simulation of subsurface temperatures (Soong *et al*., 2020) can ultimately influences relevant climate and socio-economic aspects such as food production or extreme heatwaves and droughts events (Ukkola *et al*., 2018), with relevance for Sustainable Development Goals objectives (United Nations, 2015).

We showed that the ERA5-L adequately represents the SAT and ST climatology over the SoTOS sites. Nevertheless, ERA5-L tends to underestimate the summer ST observations at some SoTOS sites. Our result agrees with previous work on the consistency of ERA reanalysis products over the Iberian Peninsula and China (Yang and Zhang, 2018; Johannsen *et al*., 2019; Li *et al*., 2020).

We used SoTOS to evaluate the shallow subsurface thermal regime and the coupling between SAT and GST on seasonal to multidecadal timescales.

The results show that the conductive model (i.e., amplitude decay and phase shift with depth of surface temperature changes) explains most of the thermal regime within the shallow subsurface. Nonetheless, there are some deviations from the purely conductive propagation of the annual surface temperature signal both in the observations and the ERA5-L. Such deviations yield different results to the apparent thermal diffusivity of the shallow subsurface from the amplitude decay or the phase shift variations as a function of depth for the bulk of the SoTOS sites. We have shown that this may be the result of soil hydrology processes such as latent heat release/uptake that alter the conductive regime. However, in the ERA5-L case, besides the influence of nonconductive processes, other factors associated to the shallow bottom boundary condition placement in the H-TESSEL land surface model may distort the heat propagation regime.

The influence of soil water content on the deviations from the purely conductive propagation is evident in the spatial distribution of thermal diffusivity and soil water content. On a regional scale, the spatial pattern of apparent thermal diffusivity roughly coincides between the estimates from observations and reanalysis. The ERA5-L data show that areas with high (low) soil water content coincide to areas with high (low) apparent thermal diffusivity. Likewise, in the temporal domain, soil water content variations are correlated to changes in the apparent thermal diffusivity. Despite the influence of soil water content, the estimated apparent thermal diffusivity, assuming purely conductive regime, is within the range of measured thermal diffusivity at different locations, including some areas in the Iberian Peninsula.

Our results show a strong connection between SAT and GST over Spain. Nevertheless, during the summer months, SAT–GST decoupling may exist due to radiative heating of the ground surface, leading to warmer GST relative to the SAT. This physical response is mainly driven by incoming solar radiation in JJA and the low precipitation rates on summer in the region. As water becomes a limiting factor, the incoming energy heats the ground surface faster than the air above. This is a general pattern in the centre and the southern parts of Spain. Also, the predominant location over bare soil and short grass surface characteristics of the sites included in SoTOS enhance the ground surface's radiative heating during the summer months. Some consistent variations in the SAT–GST relationship on interannual to multidecadal timescales over some SoTOS sites were found. The latter is represented by a higher GST increase relative to SAT over the last three decades as a response to decrease in precipitation (and soil moisture) increasing the radiative heating of the ground surface. The ERA5-L, does not show the same SAT–GST decoupling as it underestimates the decadal soil temperature trends over the 1988–2018 period.

Even though the results suggest that some SAT–GST decoupling may exist at decadal timescales, this does not necessarily imply that the SAT–GST coupling is corrupted at longer timescales. The use of soil temperature observational datasets with longer temporal coverage would be desirable to further explore this issue. However, the availability of long-term ST measurements is limited. The use of Earth System Models output or long-term reanalyses covering the full 20th century (e.g., Laloyaux *et al*., 2016; Poli *et al*., 2016) or even extending back to the mid-19th century (Slivinski *et al*., 2019) emerges as an opportunity to continue this line of research.

AUTHOR CONTRIBUTIONS

Camilo Melo-Aguilar: Conceptualization; data curation; formal analysis; investigation; methodology; visualization; writing – original draft. **Fidel González-Rouco:** Conceptualization; funding acquisition; investigation; methodology; project administration; resources; writing – review

and editing. **Norman J. Steinert:** Methodology; writing – review and editing. **Hugo Beltrami:** Writing – review and editing. **Francisco José Cuesta-Valero:** Writing – review and editing. **Almudena García-García:** Writing – review and editing. **Felix García-Pereira:** Writing – review and editing. **Elena García-Bustamante:** Writing – review and editing. **Pedro José Roldán-Gómez:** Writing – review and editing. **Thomas Schmid:** Writing – review and editing. **Jorge Navarro:** Writing – review and editing.

ACKNOWLEDGEMENTS

We gratefully acknowledge the IlModels (CGL2014-59644-R) and GreatModelS (RTI2018-102305-B-C21) projects. We also thank the Spanish Meteorological Agency (AEMET) for the data. The ERA5-Land reanalysis used in this work is funded under the Copernicus Regulation and operated by ECMWF under the ECMWF Agreement. Francisco José Cuesta-Valero is funded by the Alexander von Humboldt Foundation.

FUNDING INFORMATION

Ministerio de Economía, Industria y Competitividad, Grant/Award Number: BES-2015-075019; Ministerio de Ciencia, Innovación y Universidades, RTI2018-102305-B-C21.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

ORCID

Camilo Melo-Aguilar https://orcid.org/0000-0003-0060-1853
Fidel González-Rouco https://orcid.org/0000-0001-7090-6797
Norman J. Steinert https://orcid.org/0000-0002-2154-5857
Hugo Beltrami https://orcid.org/0000-0001-9576-8933
Francisco José Cuesta-Valero https://orcid.org/0000-0003-1577-671X
Almudena García-García https://orcid.org/0000-0003-1333-4774
Felix García-Pereira https://orcid.org/0000-0001-8491-1175
Elena García-Bustamante https://orcid.org/0000-0002-2677-0252
Pedro José Roldán-Gómez https://orcid.org/0000-0002-6413-9314
Thomas Schmid https://orcid.org/0000-0002-2849-4730
Jorge Navarro https://orcid.org/0000-0003-4476-8612

REFERENCES

Abu-Hamdeh, N.H. and Reeder, R.C. (2000) Soil thermal conductivity effects of density, moisture, salt concentration, and organic matter. *Soil Science Society of America Journal*, 64, 1285–1290.

An, K., Wang, W., Zhao, Y., Huang, W., Chen, L., Zhang, Z., Wang, Q. and Li, W. (2016) Estimation from soil temperature of soil thermal diffusivity and heat flux in sub-surface layers. *Boundary-Layer Meteorology*, 158, 473–488.

Araghi, A., Adamowski, J., Martinez, C.J. and Olesen, J.E. (2019) Projections of future soil temperature in northeast Iran. *Geoderma*, 349, 11–24.

Baker, J.M. and Baker, D.G. (2002) Long-term ground heat flux and heat storage at a mid-latitude site. *Climatic Change*, 54, 295–303.

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T. and Vitart, F. (2015) ERA-Interim/land: a global land surface reanalysis data set. *Hydrology and Earth System Sciences*, 19, 389–407.

Bartlett, M.G. (2005) Snow effect on north american ground temperatures, 1950–2002. *Journal of Geophysical Research*, 110, F03008.

Bartlett, M.G., Chapman, D.S. and Harris, R.N. (2004) Snow and the ground temperature record of climate change. *Journal of Geophysical Research*, 109, F04008.

Bell, J.E., Palecki, M.A., Baker, C.B., Collins, W.G., Lawrimore, J. H., Leeper, R.D., Hall, M.E., Kochendorfer, J., Meyers, T.P., Wilson, T. and Diamond, H.J. (2013) U.S. climate reference network soil moisture and temperature observations. *Journal of Hydrometeorology*, 14, 977–988.

Beltrami, H. and Kellman, L. (2003) An examination of short- and long-term air–ground temperature coupling. *Global and Planetary Change*, 38, 291–303.

Bonan, G. (2015) Surface energy fluxes. In: *Ecological Climatology: Concepts and Applications*, 3rd edition. Cambridge: Cambridge University Press, pp. 193–208.

Carslaw, H.S. and Jaeger, J.C. (1959) *Conduction of Heat in Solids*, 2nd edition. New York, NY: Oxford University Press.

Cermak, V. and Bodri, L. (2018) Attribution of precipitation changes on ground–air temperature offset: granger causality analysis. *International Journal of Earth Sciences*, 107, 145–152.

Cermak, V., Bodri, L., Kresl, M., Dedecek, P. and Safanda, J. (2017) Eleven years of ground–air temperature tracking over different land cover types. *International Journal of Climatology*, 37, 1084–1099.

Chen, L., Wang, G., Miao, L., Gnyawali, K.R., Li, S., Amankwah, S. O.Y., Huang, J., Lu, J. and Zhan, M. (2021) Future drought in CMIP6 projections and the socioeconomic impacts in China. *International Journal of Climatology*, 41, 4151–4170.

Crowther, T.W., Todd-Brown, K.E.O., Rowe, C.W., Wieder, W.R., Carey, J.C., Machmuller, M.B., Snoek, B.L., Fang, S., Zhou, G., Allison, S.D., Blair, J.M., Bridgham, S.D., Burton, A.J., Carrillo, Y., Reich, P.B., Clark, J.S., Classen, A.T., Dijkstra, F. A., Elberling, B., Emmett, B.A., Estiarte, M., Frey, S.D., Guo, J., Harte, J., Jiang, L., Johnson, B.R., Kröel-Dulay, G., Larsen, K. S., Laudon, H., Lavallee, J.M., Luo, Y., Lupascu, M., Ma, L.N., Marhan, S., Michelsen, A., Mohan, J., Niu, S., Pendall, E., Peñuelas, J., Pfeifer-Meister, L., Poll, C., Reinsch, S., Reynolds, L.L., Schmidt, I.K., Sistla, S., Sokol, N.W., Templer, P.H., Treseder, K.K., Welker, J.M. and Bradford, M.A. (2016) Quantifying global soil carbon losses in response to warming. *Nature*, 540, 104–108.

Cuesta-Valero, F.J., García-García, A., Beltrami, H., González-Rouco, J.F. and García-Bustamante, E. (2021) Long-term global

ground heat flux and continental heat storage from geothermal data. *Climate of the Past*, 17, 451–468.

Gallego-Elvira, B., Taylor, C.M., Harris, P.P., Ghent, D., Veal, K.L. and Folwell, S.S. (2016) Global observational diagnosis of soil moisture control on the land surface energy balance. *Geophysical Research Letters*, 43, 2623–2631.

Gao, Z., Lenschow, D.H., Horton, R., Zhou, M., Wang, L. and Wen, J. (2008) Comparison of two soil temperature algorithms for a bare ground site on the loess plateau in China. *Journal of Geophysical Research: Atmospheres*, 113, D1815. https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2008JD010285

García-García, A., Cuesta-Valero, F.J., Beltrami, H. and Smerdon, J.E. (2019) Characterization of air and ground temperature relationships within the CMIP5 historical and future climate simulations. *Journal of Geophysical Research: Atmospheres*, 124, 3903–3929.

Gonzalez-Hidalgo, J.C., Peña Angulo, D., Brunetti, M. and Cortesi, N. (2015) Motedas: a new monthly temperature database for mainland Spain and the trend in temperature (1951–2010). *International Journal of Climatology*, 35, 4444–4463.

González-Rouco, J.F., Beltrami, H., Zorita, E. and Stevens, B. (2009) Borehole climatology: a discussion based on contributions from climate modelling. *Climate of the Past*, 5, 99–127.

González-Rouco, J.F., Jiménez, J.L., Quesada, V. and Valero, F. (2001) Quality control and homogeneity of precipitation data in the southwest of europe. *Journal of Climate*, 14, 964–978.

Guerreiro, S.B., Dawson, R.J., Kilsby, C., Lewis, E. and Ford, A. (2018) Future heat-waves, droughts and floods in 571 european cities. *Environmental Research Letters*, 13, 034009.

Hao, G., Zhuang, Q., Pan, J., Jin, Z., Zhu, X. and Liu, S. (2014) Soil thermal dynamics of terrestrial ecosystems of the conterminous United States from 1948 to 2008: an analysis with a process-based soil physical model and ameriflux data. *Climatic Change*, 126, 135–150.

Hartmann, D., Klein Tank, A., Rusticucci, M., Alexander, L., Brönnimann, S., Charabi, Y., Dentener, F., Dlugokencky, E., Easterling, D., Kaplan, A., Soden, B., Thorne, P., Wild, M. and Zhai, P. (2013) *Observations: Atmosphere and Surface*. Cambridge and New York, NY: Cambridge University Press, pp. 159–254.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S. and Thépaut, J.-N. (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146, 1999–2049.

Hicks Pries, C.E., Castanha, C., Porras, R.C. and Torn, M.S. (2017) The whole-soil carbon flux in response to warming. *Science*, 355, 1420–1423.

Hu, Q. and Feng, S. (2003) A daily soil temperature dataset and soil temperature climatology of the contiguous United States. *Journal of Applied Meteorology*, 42, 1139–1156.

Huang, R., Xi Huan, J., Zhang, C., Yuan Ma, H., Zhuo, W., Yi Chen, Y., Hai Zhu, D., Wu, Q. and Mansaray, L.R. (2020) Soil temperature estimation at different depths, using remotely-sensed data. *Journal of Integrative Agriculture*, 19, 277–290.

IFS. (2018) *ECMWF Integrated Forecast System Documentation. Part IV: Physical processes. No. 4 in IFS Documentation*. European Centre for Medium-Range Weather Forecasts (ECMWF). Available at: https://www.ecmwf.int/en/elibrary/18714-ifs-documentation-cy45r1-part-iv-physical-processes [Accessed 30th August 2020].

Johannsen, F., Ermida, S., Martins, J.P.A., Trigo, I.F., Nogueira, M. and Dutra, E. (2019) Cold bias of ERA5 summertime daily maximum land surface temperature over Iberian Peninsula. *Remote Sensing*, 11, 2570.

Kendall, M. (1975) *Rank Correlation Methods*, 4th edition. London: Charles Griffin.

Kluitenberg, G. (2002) Heat capacity and specific heat. In: *Methods of Soil Analysis: Part 4 Physical Methods*. Madison, Wisconsisn: Soil Science Society of America, Inc. Vol. 5, pp. 1201–1208.

Kodešová, R., Vlasáková, M., Fér, M., Teplá, D., Jakšík, O., Neuberger, P. and Adamovský, R. (2013) Thermal properties of representative soils of the Czech Republic. *Soil and Water Research*, 8, 141–150.

Laloyaux, P., Balmaseda, M., Dee, D., Mogensen, K. and Janssen, P. (2016) A coupled data assimilation system for climate reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 142, 65–78.

Li, M., Wu, P. and Ma, Z. (2020) A comprehensive evaluation of soil moisture and soil temperature from third-generation atmospheric and land reanalysis data sets. *International Journal of Climatology*, 40, 1–23.

Mann, H.B. (1945) Nonparametric tests against trend. *Econometrica*, 13, 245–259.

McCumber, M.C. and Pielke, R.A. (1981) Simulation of the effects of surface fluxes of heat and moisture in a mesoscale numerical model: 1. Soil layer. *Journal of Geophysical Research: Oceans*, 86, 9929–9938.

Melo-Aguilar, C., González-Rouco, J.F., García-Bustamante, E., Navarro-Montesinos, J. and Steinert, N. (2018) Influence of radiative forcing factors on ground–air temperature coupling during the last millennium: implications for borehole climatology. *Climate of the Past*, 14, 1583–1606.

Muñoz Sabater, J. (2019) *ERA5-land monthly averaged data from 1981 to present*. Copernicus Climate Change Service (C3S) Climate Data Store. https://doi.org/10.24381/cds.68d2bb30 [Accessed 16th September 2020].

Oladunjoye, M.A. and Sanuade, O.A. (2012) Thermal diffusivity, thermal effusivity and specific heat of soils in olorunsogo powerplant, southwestern nigeria. *IJRRAS*, 13, 502–521.

Patil, M., Waghmare, R., Halder, S. and Dharmaraj, T. (2011) Performance of noah land surface model over the tropical semi-arid conditions in western India. *Atmospheric Research*, 99, 85–96.

Peterson, T.C., Vose, R., Schmoyer, R. and Razuvaëv, V. (1998) Global historical climatology network (GHCN) quality control of monthly temperature data. *International Journal of Climatology*, 18, 1169–1179.

Poli, P., Hersbach, H., Dee, D.P., Berrisford, P., Simmons, A.J., Vitart, F., Laloyaux, P., Tan, D.G.H., Peubey, C., Thépaut, J.-N., Trémolet, Y., Hólm, E.V., Bonavita, M., Isaksen, L. and Fisher, M. (2016) ERA-20C: an atmospheric reanalysis of the twentieth century. *Journal of Climate*, 29, 4083–4097.

Putnam, S.N. and Chapman, D.S. (1996) A geothermal climate change observatory: first year results from emigrant pass in Northwest Utah. *Journal of Geophysical Research: Solid Earth*, 101, 21877–21890.

Qian, B., Gregorich, E.G., Gameda, S., Hopkins, D.W. and Wang, X. L. (2011) Observed soil temperature trends associated with climate change in Canada. *Journal of Geophysical Research: Atmospheres*, 116, D02106.

Reek, T., Doty, S.R. and Owen, T.W. (1992) A deterministic approach to the validation of historical daily temperature and precipitation data from the cooperative network. *Bulletin of the American Meteorological Society*, 73, 753–765.

Sáez Blázquez, C., Farfń Martín, A., Martín Nieto, I., Carrasco García, P., Sánchez Pérez, L.S. and González Aguilera, D. (2017) Thermal conductivity map of the avila region (Spain) based on thermal conductivity measurements of different rock and soil samples. *Geothermics*, 65, 60–71.

Sancho-Avila, J.M., Riesco-Martin, J., Jiménez-Alonso, C., Sánchez-de Cos-Escuin, M.C., Montero-Caldoso, J. and López-Bartolome, M. (2012) *Atlas de radiación solar en España utilizando datos del saf de clima de eumesat*. Madrid, Spain: Agencia Estatal de Meteorología. Technical report.

Santer, B.D., Wigley, T.M.L., Boyle, J.S., Gaffen, D.J., Hnilo, J.J., Nychka, D., Parker, D.E. and Taylor, K.E. (2000) Statistical significance of trends and trend differences in layer-average atmospheric temperature time series. *Journal of Geophysical Research: Atmospheres*, 105, 7337–7356.

Schaefer, G. L. and Paetzold, R. F. (2000) *SNOTEL (SNOwpack TELemetry) and SCAN (Soil Climate Analysis Network)*. Presented at the Automated Weather Station (AWS) Workshop, March 6–10, Lincoln, Nebr.

Schmidt, W.L., Gosnold, W.D. and Enz, J.W. (2001) A decade of air–ground temperature exchange from Fargo, North Dakota. *Global and Planetary Change*, 29, 311–325.

Sehler, R., Li, J., Reager, J. and Ye, H. (2019) Investigating relationship between soil moisture and precipitation globally using remote sensing observations. *Journal of Contemporary Water Research & Education*, 168, 106–118.

Serrano-Notivoli, R., Beguería, S., Saz, M.A., Longares, L.A. and de Luis, M. (2017) Spread: a high-resolution daily gridded precipitation dataset for Spain—an extreme events frequency and intensity overview. *Earth System Science Data*, 9, 721–738.

Slivinski, L.C., Compo, G.P., Whitaker, J.S., Sardeshmukh, P.D., Giese, B.S., McColl, C., Allan, R., Yin, X., Vose, R., Titchner, H., Kennedy, J., Spencer, L.J., Ashcroft, L., Brönnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram, T. A., Crouthamel, R., Domínguez-Castro, F., Freeman, J.E., Gergis, J., Hawkins, E., Jones, P.D., Jourdain, S., Kaplan, A., Kubota, H., Blancq, F.L., Lee, T.-C., Lorrey, A., Luterbacher, J., Maugeri, M., Mock, C.J., Moore, G.K., Przybylak, R., Pudmenzky, C., Reason, C., Slonosky, V.C., Smith, C.A., Tinz, B., Trewin, B., Valente, M.A., Wang, X.L., Wilkinson, C., Wood, K. and Wyszyński, P. (2019) Towards a more reliable historical reanalysis: improvements for version 3 of the twentieth century reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, 145, 2876–2908.

Smerdon, J.E., Pollack, H.N., Cermak, V., Enz, J.W., Kresl, M., Safanda, J. and Wehmiller, J.F. (2004) Air–ground temperature coupling and subsurface propagation of annual temperature signals. *Journal of Geophysical Research*, 109, D21107.

Smerdon, J.E., Pollack, H.N., Enz, J.W. and Lewis, M.J. (2003) Conduction-dominated heat transport of the annual temperature signal in soil. *Journal of Geophysical Research*, 108, 2431.

Smerdon, J.E. and Stieglitz, M. (2006) Simulating heat transport of harmonic temperature signals in the earth's shallow subsurface: lower-boundary sensitivities. *Geophysical Research Letters*, 33, L14402.

Soong, J.L., Phillips, C.L., Ledna, C., Koven, C.D. and Torn, M.S. (2020) CMIP5 models predict rapid and deep soil warming over the 21st century. *Journal of Geophysical Research: Biogeosciences*, 125, e2019JG005266.

Steinert, N.J., González-Rouco, J.F., de Vrese, P., García-Bustamante, E., Hagemann, S., Melo-Aguilar, C., Jungclaus, J. H. and Lorenz, S.J. (2021) Increasing the depth of a land surface model. Part II: temperature sensitivity to improved subsurface thermodynamics and associated permafrost response. *Journal of Hydrometeorology*, 22, 3231–3254.

Taylor, K.E. (2001) Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106, 7183–7192.

Tong, B., Gao, Z., Horton, R. and Wang, L. (2017) Soil apparent thermal diffusivity estimated by conduction and by conduction–convection heat transfer models. *Journal of Hydrometeorology*, 18, 109–118.

Ukkola, A.M., Pitman, A.J., Donat, M.G., De Kauwe, M.G. and Angélil, O. (2018) Evaluating the contribution of land–atmosphere coupling to heat extremes in CMIP5 models. *Geophysical Research Letters*, 45, 9003–9012.

United Nations. (2015) *Transforming our world: the 2030 agenda for sustainable development. resolution adopted by the general assembly on September 25, 2015 (a/res/70/1)*. Available at: https://www.unfpa.org/sites/default/files/resource-pdf [Accessed 21st November 2021].

Viterbo, P. and Beljaars, A. (1995) *An improved land surface parametrization scheme in the ECMWF model and its validation*. Reading: ECMWF. Report number: 52. Available at: https://www.ecmwf.int/node/12951 [Accessed 12th August 2020].

WMO. (2018) *Guide to Climatological Practices*. Geneva: World Meteorological Organization. Technical report.

Yang, K. and Zhang, J. (2018) Evaluation of reanalysis datasets against observational soil temperature data over China. *Climate Dynamics*, 50, 317–337.

Zhang, H., Wang, E., Zhou, D., Luo, Z. and Zhang, Z. (2016) Rising soil temperature in China and its potential ecological impact. *Scientific Reports*, 16, 35530.

Zhang, T. and Barry, R.G.G.D. (2000) *Russian historical soil temperature data*. Arctic Data Center. Available at: https://data.eol.ucar.edu/dataset/106.ARCSS078 [Accessed 12th September 2020].

How to cite this article: Melo-Aguilar, C., González-Rouco, F., Steinert, N. J., Beltrami, H., Cuesta-Valero, F. J., García-García, A., García-Pereira, F., García-Bustamante, E., Roldán-Gómez, P. J., Schmid, T., & Navarro, J. (2022). Near-surface soil thermal regime and land–air temperature coupling: A case study over Spain. *International Journal of Climatology*, 1–19. https://doi.org/10.1002/joc.7662