Atmospheric Research 284 (2023) 106585

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmosres

The effect of the Madden–Julian Oscillation on the global electric circuit

Alexander V. Kozlov [a,b], Nikolay N. Slyunyaev [a,b,*], Nikolay V. Ilin [a,b], Fedor G. Sarafanov [a,b], Alexander V. Frank-Kamenetsky [c]

[a] Lobachevsky State University of Nizhny Novgorod, 23 Gagarina Ave., 603950 Nizhny Novgorod, Russia
[b] Institute of Applied Physics, Russian Academy of Sciences, 46 Ulyanova St., 603950 Nizhny Novgorod, Russia
[c] Arctic and Antarctic Research Institute, 38 Beringa St., 199397 St. Petersburg, Russia

ARTICLE INFO

Keywords:
Global electric circuit
Ionospheric potential
Fair-weather electric field
Madden–Julian Oscillation
Convection
Modelling

ABSTRACT

The Madden–Julian Oscillation (MJO) is the most dominant component of the climate variability in the tropics on the timescale of tens of days. Occurring irregularly, the MJO consists of an eastward propagation of certain large-scale coupled structures in atmospheric circulation and deep convection. Here we investigate the effect of the MJO on the direct current global electric circuit (GEC), using both numerical simulations and the results of electric field measurements. We have reproduced the atmospheric dynamics for 1980–2020 with the help of the Weather Research and Forecasting model and meteorological reanalysis data; this allowed us to simulate 41 years of the GEC variation by parameterising contributions to the ionospheric potential (IP) in terms of precipitation and convection. Besides, we have analysed the results of surface potential gradient (PG) measurements at the Vostok station in Antarctica during 2006–2020. When averaged over many days, both the simulated IP and measured fair-weather PG show sinewave-like variations with the MJO phase, which manifest a statistically significant correlation with the MJO cycle. A more elaborate investigation involving an application of empirical orthogonal functions shows that the dynamics of contributions to the IP follows the patterns of variability of deep convection typical for the MJO. The variation of the IP on the MJO scale can largely be represented as a superposition of two basic oscillating patterns of convection, one of which is located over the Maritime Continent and its surrounding seas and the other is located over the Indian Ocean and part of South-East Asia. These oscillating patterns together describe the eastward progression of perturbations and can be naturally associated with the two components of the Real-time Multivariate MJO index, widely used to describe the MJO. Together with earlier results about links between the GEC and the El Niño—Southern Oscillation, the results concerning the MJO obtained in this study indicate that the GEC is an important part of the Earth system, which reflects climate variability on different timescales.

1. Introduction

The Madden–Julian Oscillation (MJO) is the most dominant component of the climate variability in the tropics on the timescale of tens of days. MJO events occur irregularly and typically last for 30–90 days. An important aspect of the MJO is the coupling between the large-scale atmospheric circulation and deep convection; during each MJO event large-scale coupled structures slowly propagate eastwards at an average speed of 5–10 m s^{-1}. This phenomenon affects all longitudes with the most pronounced effect over the Indian Ocean, the Maritime Continent and the Pacific Ocean.

The MJO was first discovered by Madden and Julian (1971, 1972). Over the following 50 years it has been widely studied from both observational and theoretical perspectives, with hundreds of papers dedicated to analysing different aspects of this oscillation (for a review see, e.g., Madden and Julian, 1994; Zhang, 2005; Zhang et al., 2020). Such interest in the MJO is caused by the fact that, being the main component of the intraseasonal variability in the tropical atmosphere (Lau and Waliser, 2012), it influences global rainfall patterns and the development of tropical cyclones, monsoons and even the El Niño—Southern Oscillation (ENSO).

Despite the numerous papers devoted to studying the MJO in the context of meteorology and climate, it appears that there have been only few studies addressing its impact on atmospheric electricity. Using satellite data and the results of Schumann resonance measurements in Antarctica, Anyamba et al. (2000) have shown that intraseasonal

* Corresponding author at: Lobachevsky State University of Nizhny Novgorod, 23 Gagarina Ave., 603950 Nizhny Novgorod, Russia.
E-mail address: slyunyaev.n@ipfran.ru (N.N. Slyunyaev).

https://doi.org/10.1016/j.atmosres.2022.106585
Received 31 May 2022; Received in revised form 15 December 2022; Accepted 16 December 2022
Available online 20 December 2022

oscillations in deep convection are reflected in the variation of the Schumann resonance intensity. Schumann resonances are excited in the Earth–ionosphere cavity by the global distribution of lightning discharges; therefore it is not surprising that global changes in tropical deep convection (often accompanied by lightning) may influence their characteristics. Stolz et al. (2017), who studied anomalies in lightning activity over the Indian Ocean on the basis of satellite data, have found some patterns associated with different phases (i.e. stages) of the MJO cycle. Recently Beggan and Musur (2019) searched for MJO patterns in Schumann resonance data from Eskdalemuir observatory in the UK; they have found that the intensity and frequency of Schumann resonances correlate with the indices characterising the MJO during the cold phase of ENSO (La Niña) but not during its warm phase (El Niño).

Notwithstanding that lightning is indeed associated with deep convection and the Schumann resonances are known to reflect modes of climate variability such as the ENSO (e.g., Sátori et al., 2009; Williams et al., 2021), strictly speaking, lightning flash rates are related to convection only indirectly. In this regard, a more natural concept to consider is the direct current global electric circuit (GEC), which consists of the distribution of quasi-stationary electric fields and currents in the atmosphere, maintained by charge separation inside thunderstorms and electrified shower clouds (ESCs) all over the Earth (see, e.g., Rycroft et al., 2008; Williams, 2009; Williams and Mareev, 2014). The direct current GEC is often contrasted with the Schumann resonances, which are sometimes referred to as the alternating current GEC owing to the global nature of their origin (Rycroft et al., 2012).

The direct current GEC is essentially different from the alternating current GEC in that the latter is maintained by lightning-producing clouds, whereas the generators of the former include not only thunderstorms but also ESCs (i.e. electrified clouds without lightning). It is therefore slightly more natural to search for the effect of changes in deep convection (in particular, of those associated with the MJO) in the intensity of the direct current GEC rather than in the Schumann resonance intensity, since the former should keep more information about convection than the latter. In addition, ESCs are especially abundant in the vicinity of the Maritime Continent (e.g., Liu et al., 2010), which is a key region for the MJO.

Two important quantitative characteristics of the direct current GEC are the ionospheric potential (IP) and the fair-weather potential gradient (PG). Being the sum of contributions from electrified clouds all over the globe, the IP is arguably the most fundamental measure of the GEC intensity. The IP is convenient for theoretical discussion of the GEC and its numerical simulation but is rather difficult to measure. On the other hand, the fair-weather PG is affected by various local factors at the site concerned but can be measured on a regular basis. Tacza et al. (2022), who performed spectral analysis of the PG recorded at several locations, have shown that the PG variability manifests, among others, the periodicity of about 45 days, probably associated with the MJO.

Recently Slyunyaev et al. (2021b,c) have employed simulations to show that changes in deep convection during ENSO events (El Niños and La Niñas) influence the IP (which characterises the direct current GEC intensity) and its diurnal variation. These findings turned out to agree with the results following from PG measurements at the Earth's surface (Harrison et al., 2011; Lavigne et al., 2017; Slyunyaev et al., 2021a).

Here we use the same approach in order to investigate the effect of the MJO on the direct current GEC, using both numerical simulations with the Weather Research and Forecasting (WRF) model and the results of electric field measurements in Antarctica. In Section 2 we describe the employed methods and data, in Section 3 we present the main results, which are further analysed and explained in Section 4. Sections 5 and 6 are devoted to further discussion (including an evaluation of the possible influence of solar activity on the results) and concluding remarks.

2. Methods and data

2.1. Indices characterising the MJO

It is customary to describe climate variations with the help of special indices, which simplify the consideration of these phenomena. There are many such indices characterising the MJO, the most widespread of which is the Real-time Multivariate MJO index (RMM), introduced by Wheeler and Hendon (2004). We shall now give a brief description of this index.

In climatology it is quite common to calculate the so-called empirical orthogonal functions (EOFs) and principal components (PCs) for different physical parameters in order to identify the MJO (e.g., Knutson and Weickmann, 1987; Slingo et al., 1999; Lo and Hendon, 2000; Matthews, 2000; Kessler, 2001). EOFs of a physical process are mutually orthogonal spatial patterns determined by the data, whereas PCs are time-dependent coefficients of these patterns. The first EOF is selected such that it describes the largest possible portion of the data variance; similarly, the second EOF describes the largest part of the remaining variance, and so on (e.g., Zhang and Moore, 2015, Ch. 6; Hannachi, 2021, Ch. 3). The calculation of EOFs and PCs makes it possible to decompose a complicated process into a sum of relatively simple components, the first few of which are often sufficient to describe the main part of the observed variation.

To introduce the RMM index, Wheeler and Hendon (2004) applied EOF analysis to three variables, namely zonal wind velocities at 850 hPa and 200 hPa, inferred from meteorological reanalysis data, and the outgoing long-wave radiation (OLR) flux, observed from satellites. A combination of these three variables allows one to identify patterns typical for the MJO in both atmospheric circulation (reflected by zonal wind velocities) and deep convection (reflected by OLR).

The MJO always occurs simultaneously with modes of climate variability with different temporal and spatial scales. Therefore all of the three variables forming the basis of the RMM are first processed to remove the majority of the non-MJO variations. After that, concentrating on the equatorial region bounded by the circles of latitude at 15° S and 15° N and averaging everything over 2.5°-wide meridional strips in this region (corresponding to different longitudes), one obtains for each day three data sets of length 144. These three data sets are then combined together to form 432-dimensional vectors, from which after some normalisation the EOFs are extracted. The PCs corresponding to the first two EOFs obtained in this manner form the two-dimensional RMM index (these PCs are referred to as RMM1 and RMM2). Simply put, RMM1 and RMM2 are dimensionless coefficients (depending on the day) of the two most prominent basic spatial patterns for the data vectors combining zonal wind velocities at two different heights with OLR.

From a geographical perspective, RMM1 describes an oscillation of convection over the Maritime Continent together with the surrounding seas, while RMM2 primarily corresponds to its variation over the Indian Ocean (see the curves for OLR in Wheeler and Hendon, 2004, Fig. 1). The superposition of these components yields the eastward propagation of a certain convective structure along the equator on the MJO timescale. It is convenient to illustrate the state of the MJO with a point in the (RMM1, RMM2) plane, which is shown schematically in Fig. 1. This plane is divided into eight regions, according to which eight phases of the MJO are formally distinguished. These phases represent different stages of a typical MJO cycle, each of which is characterised by its own patterns of convection and atmospheric circulation; two-dimensional indices, such as the RMM, allow us to define these stages in a more formal manner. Roughly speaking, during an average MJO event, a point in the (RMM1, RMM2) plane moves around the origin, passing through different phases; however, actual trajectories during specific events may be much more complicated (see Wheeler and Hendon, 2004, Fig. 7).

In this study the values of RMM1 and RMM2 for 1980–2020 have been taken from a time series maintained by the Australian Bureau of Meteorology, available online at http://www.bom.gov.au/climate/mjo.

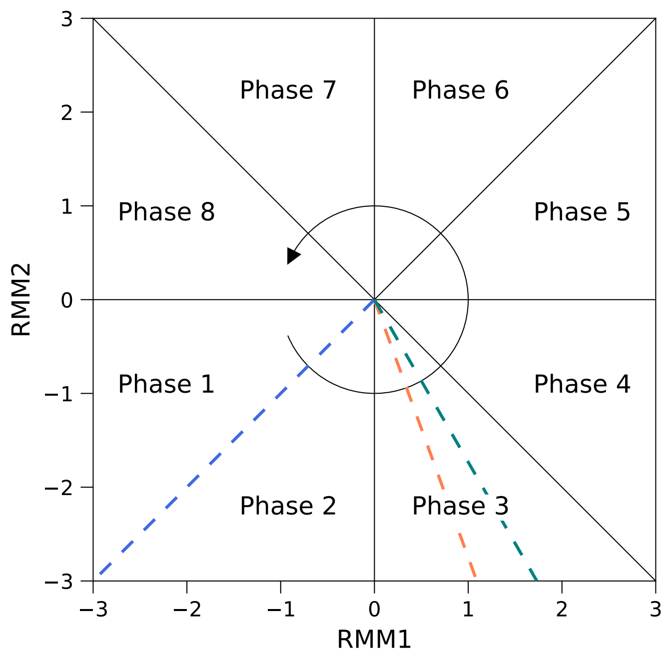

Fig. 1. The division of the (RMM1, RMM2) plane into eight regions which determine eight phases of the MJO. The circular arrow schematically indicates a typical ('average') trajectory of a point during an MJO event (actual trajectories during specific events are often much more complicated; see, e.g., Wheeler and Hendon, 2004, Fig. 7). Dashed lines indicate the directions in the (RMM1, RMM2) plane providing maximum correlation with the simulated ionospheric potential (orange line), the negative of OLR (teal line) and the fair-weather potential gradient at Vostok (blue line); for more details, see Sections 3 and 5.

2.2. Modelling the atmospheric dynamics

To simulate the GEC variation for this study, we used the approach which was earlier employed in Slyunyaev et al. (2021b,c) and described in detail by Ilin et al. (2020). First, we simulated the atmospheric dynamics for 1980–2020 with the help of the WRF model (Powers et al., 2017), namely version 4.3 of the Advanced Research WRF (ARW) solver (Skamarock et al., 2019), running it globally on a $1^\circ \times 1^\circ$ latitude-longitude grid with 51 altitude levels from the Earth's surface to the 10-mbar level and with time steps of 150 s. The following set of physics schemes for the WRF was chosen: mp_physics = 2, ra_lw_physics = 4, ra_sw_physics = 4, radt = 30, sf_sfclay_physics = 1, sf_surface_physics = 2, bl_pbl_physics = 1 and cu_physics = 2. Precipitation was calculated on the basis of the parameterisation of convection, which is the only possible option on a large-scale grid; the hybrid sigma-pressure vertical coordinate, the hydrostatic approximation and polar filtering for latitudes greater than 60° were used in simulations.

To initialise the WRF, we used the ERA5 data on pressure levels (Hersbach et al., 2018). ERA5 stands for the fifth generation ECMWF (European Centre for Medium-Range Weather Forecasts) reanalysis for the global climate and weather; it is a consistent data set combining model data with global observations. Setting the ERA5 data as initial conditions, we performed 48-h free runs of the WRF starting at 0:00 UTC and 12:00 UTC on every third day from 31 December 1979 to 28 December 2020. In each model run we restricted our attention to the 24-h interval between 18 and 42 h since the model initiation, as the model needs some time to achieve the correct reproduction of the data absent in initial conditions (see Ilin et al., 2020). In this way we obtained two 24-h data sets for every third day during 1980–2020, thus reproducing the long-term atmospheric dynamics.

2.3. Simulation of the GEC variation

Next, we simulated the long-term variation of the IP using the parameterisation suggested by Slyunyaev et al. (2019) on the basis of earlier ideas of Mareev and Volodin (2014). This parameterisation expresses contributions to the IP from different grid cells in terms of climate variables such that the total IP V is given by the equation

$$V = \sum_i \frac{j_0 H S_i}{\sigma_0 S_\mathrm{E}} \frac{P_i}{W_i} \times \left(\exp\left(-\frac{z_i^\mathrm{l}}{H} \right) - \exp\left(-\frac{z_i^\mathrm{u}}{H} \right) \right) \times \begin{cases} 0, & \varepsilon_i < \varepsilon_0, \\ 1, & \varepsilon_i \geq \varepsilon_0, \end{cases} \quad (1)$$

where j_0 is the characteristic value of the charge separation current density inside electrified clouds, σ_0 and H are the surface value and scale height for the air conductivity (in calculations we assume $H = 6$ km), S_E is the total area of the Earth's surface, S_i is the area covered by the ith model grid column; $P_i, W_i, \varepsilon_i, z_i^\mathrm{l}$ and z_i^u are the amount of precipitation totalled over a symmetric two-hour interval, the total amount of precipitable water, the maximum convective available potential energy (CAPE) and the lower and upper boundaries of the mixed phase region (which we approximate by the heights of the 0 °C and −38 °C isotherms) in the ith column respectively and ε_0 is a threshold CAPE value, assumed to be 1 kJ kg^{-1} (this value gives the best agreement of the mean diurnal variation of the IP with the results of measurements). The sum in (1) is taken over all grid columns, but can be restricted to specific cells if one is looking for a contribution to the IP from some specific region. For the derivation and discussion of formula (1), see Ilin et al. (2020). Here we restrict ourselves to noting that two basic ideas behind this parameterisation are (i) elimination of the grid columns with weak convective activity on the basis of CAPE and (ii) estimating the area covered by electrified clouds (which maintain the direct current GEC) in each column from the amount of precipitation.

To simulate the long-term variation of the IP during 1980–2020, we calculated its average value for every third day on the basis of the 24-h data sets described above with the help of the aforementioned parameterisation. Using two model runs with different initial moments for each day allows us to reduce the effect of systematic uncertainties in the parameterisation of climate variables. Finally, we adjusted the value of j_0 in (1) such that the total IP averaged over 1980–2020 became equal to 240 kV, which is a rather typical value for the IP (Markson, 2007).

2.4. Electric field measurements at Vostok station

Apart from simulations with the WRF model, in this study we use the results of continual PG measurements at the Russian Vostok station in the Antarctic (78.5° S, 106.9° E, 3488 m above sea level) during 2006–2020. Collected in a clean location, these PG values make up a unique long-term data set, which can be used to study the GEC variation (e.g., Burns et al., 2012, 2017).

The rotating-dipole field mill, presently operating at Vostok station, was installed there at the end of 2005 under an agreement between the Russian Antarctic Expedition and the Australian Antarctic Division (Burns et al., 2017). The field mill is placed about 3 m above the snow surface, upwind of the main station buildings. The PG data are collected in the form of mean values over consecutive 10-s periods. The data have a major gap during the second half of 2017 and several minor gaps; also, for technical reasons, only 5-min averages are available for some periods of time (in particular, for the entire years 2017 and 2020).

To simplify the analysis, we transformed the field mill data into average values over consecutive UTC hours; if both 10-s data and 5-min data were available for a particular day, we took the former. Considering that sometimes there are small gaps in the data, when averaging, we retained only those hours for which at least 80% of the 10-s or 5-min PG values are known. All measured PG values were divided by a form factor equal to 3 in order to eliminate the disturbance caused by the metal pole supporting the field mill.

To select the 'fair-weather' days at Vostok (more precisely, the least disturbed days, for we do not consider actual weather), we employed the approach used earlier by Slyunyaev et al. (2021a). When dealing with PG measurements in the Antarctic, one can avoid using actual meteorological data to eliminate the days affected by bad weather, and criteria based on the PG values themselves usually work well (e.g., Burns et al., 2012, 2017). The following formal procedure was applied:

1. We excluded the days with incomplete/absent hourly data.
2. We excluded the days with negative/zero hourly PG values.
3. We excluded the days with hourly PG values exceeding 300 V m^{-1}.
4. Among the remaining days we retained only those for which the diurnal peak-to-peak amplitude does not exceed 150% of the diurnal mean.

2.5. Other data

We use daily 1°×1° OLR data downloaded from the NOAA (National Oceanic and Atmospheric Administration) PSL (Physical Sciences Laboratory) website at https://psl.noaa.gov/mddb2/showDataset.html?datasetID=37. These data are derived from the radiance observations by the HIRS (High Resolution Infrared Radiation Sounder) instruments on board NOAA TIROS-N (Television and Infra-Red Observation Satellite) series and EUMETSAT (European Organisation for the Exploitation of Meteorological Satellites) MetOp (Meteorological Operational satellite) satellites; the processing algorithm is described by Lee (2014).

We also employ data on sunspot numbers from the World Data Center SILSO (Sunspot Index and Long-term Solar Observations), Royal Observatory of Belgium, Brussels, available at https://www.sidc.be/silso/datafiles.

3. The effect of the climatological MJO on the GEC

3.1. Simulations

We shall first search for patterns related to the MJO in the simulated IP variation for 1980–2020 (see Section 2.3). To this end, we will average the values of regional contributions to the IP over the days corresponding to each of the eight MJO phases. These phases are determined according to the polar angle of a point in the (RMM1, RMM2) plane (see Fig. 1); on average, during an MJO event (i.e. one cycle) a point in this plane moves around the origin, passing through consecutive phases. To detect the occurrence of MJO events, one usually considers not only the phase but also the distance from the origin (i.e. the amplitude), which makes it possible to identify strong MJO activity (e.g., Wheeler and Hendon, 2004). However, in this study we will not distinguish between weak and strong MJO and will not consider the RMM amplitude at all; such an approach allows us to use more days for our analysis, and we have not found any dependence of the observed trends on the RMM amplitude (although, of course, their magnitude becomes greater if only strong MJO events are considered).

One of the most important aspects of the MJO is the eastward propagation of a certain large-scale convective structure in the tropics. According to our parameterisation of the IP, expressed by Eq. (1), contributions to the total IP from individual grid cells are to a large extent determined by CAPE and precipitation, which links them with deep convection. Therefore it is reasonable to expect that MJO patterns will be distinguishable in contributions to the IP.

Considering that the MJO is a very irregular phenomenon, here we shall not focus on individual MJO events, which may be substantially different from each other. Instead, we average the diurnal mean values of contributions to the IP over the days corresponding to each phase of the MJO. Then, subtracting the long-term mean value of each contribution from the average values over specific phases, we pass to anomalies of contributions, which are easier to interpret: a positive anomaly means that the contribution to the IP is generally higher than usual during the MJO phase in question, whereas a negative anomaly means that it is generally lower than usual.

Fig. 2 demonstrates how such anomalies in contributions to the IP change with the MJO phase. We observe that positive and negative anomalies propagate towards the east one after another as the MJO phase increases. This behaviour reflects the eastward propagation of regions with enhanced and reduced convective activity during MJO events (see, e.g., Madden and Julian, 1972, Fig. 16; Wheeler and Hendon, 2004, Figs. 8 and 9). It is also interesting that Fig. 2 resembles, to some extent, satellite-based maps of lightning flash rates over the Indian Ocean during different MJO phases published by Stolz et al. (2017, Fig. 8), with an especially good agreement near the coast of Sumatra. At the same time we should note that generally speaking, the distribution of lightning flash rates need not be similar to those of convection and contributions to the IP (in particular, owing to the important contribution of ESCs).

Thus on average, simulated contributions to the IP vary in accordance with the dynamics of convection on the MJO timescale, which should not be surprising in view of the remarks made above. It is much more interesting to look at the behaviour of global GEC parameters such as the IP or fair-weather PG, which can be actually measured. Fig. 3a shows the average values of the total IP for different MJO phases. We observe a sine-like variation with a maximum in phase 3 and a minimum in phase 7. The peak amplitude of this variation is about 12 kV.

Fig. 3b demonstrates the average values of RMM1 and RMM2 for different MJO phases (on the basis of the data for 1980–2020). These variations are well approximated by sine curves, as one might have expected from the definition of MJO phases in terms of the polar angle in the (RMM1, RMM2) plane (Fig. 1). The variations of RMM1 and RMM2 have similar amplitudes, but the former is ahead of the latter by one quarter of the period (i.e. by two MJO phases), which simply reflects the fact that in the (RMM1, RMM2) plane the RMM1 and RMM2 axes are at right angles to each other.

Comparing Fig. 3a with Fig. 3b, we conclude that the dependence of the IP on the MJO phase is to a large extent similar to that of RMM2 multiplied by -1. More precisely, there is a clear negative correlation between the IP and RMM2 with the correlation coefficient $r = -0.93$. At the same time the correlation coefficient between the IP and RMM1 is only $r = 0.33$.

Since we consider $N = 8$ MJO phases, we can estimate the significance of the observed correlations by using a two-tailed test on the basis of Student's t-distribution with $N - 2 = 6$ degrees of freedom. If two samples of size N correspond to uncorrelated variables and r is the sample correlation coefficient, the quantity

$$q = \sqrt{\frac{(N-2)r^2}{(1-r^2)}}$$

obeys precisely this distribution (for a more precise formulation of this result, see, e.g., Kendall and Stuart, 1961, Ch. 26, 31). At the 1% significance level the null hypothesis that the two variables are uncorrelated is rejected if $q \geqslant 3.71$, or $|r| \geqslant 0.83$. Using this criterion, we deduce that the negative correlation between the IP and RMM2, for which $r = -0.93$, is statistically significant at the 1% level, while the correlation between the IP and RMM1 with $r = 0.33$ is not significant.

More generally, given that the RMM index is essentially two-dimensional, instead of RMM1 or RMM2 we can consider its projection on an arbitrary direction in the (RMM1, RMM2) plane, that is, the quantity

$$\mathrm{RMM1}\cdot\cos\phi + \mathrm{RMM2}\cdot\sin\phi, \quad (2)$$

where ϕ is the polar angle of the direction in question (measured from the RMM1 axis). The correlation coefficient r between (2) and the IP varies depending on ϕ. A computation shows that r attains its maximum value $r = 0.99$ when $\phi = 290°$; in Fig. 1 this direction is indicated by an orange dashed line. Thus we infer that IP values, when averaged over

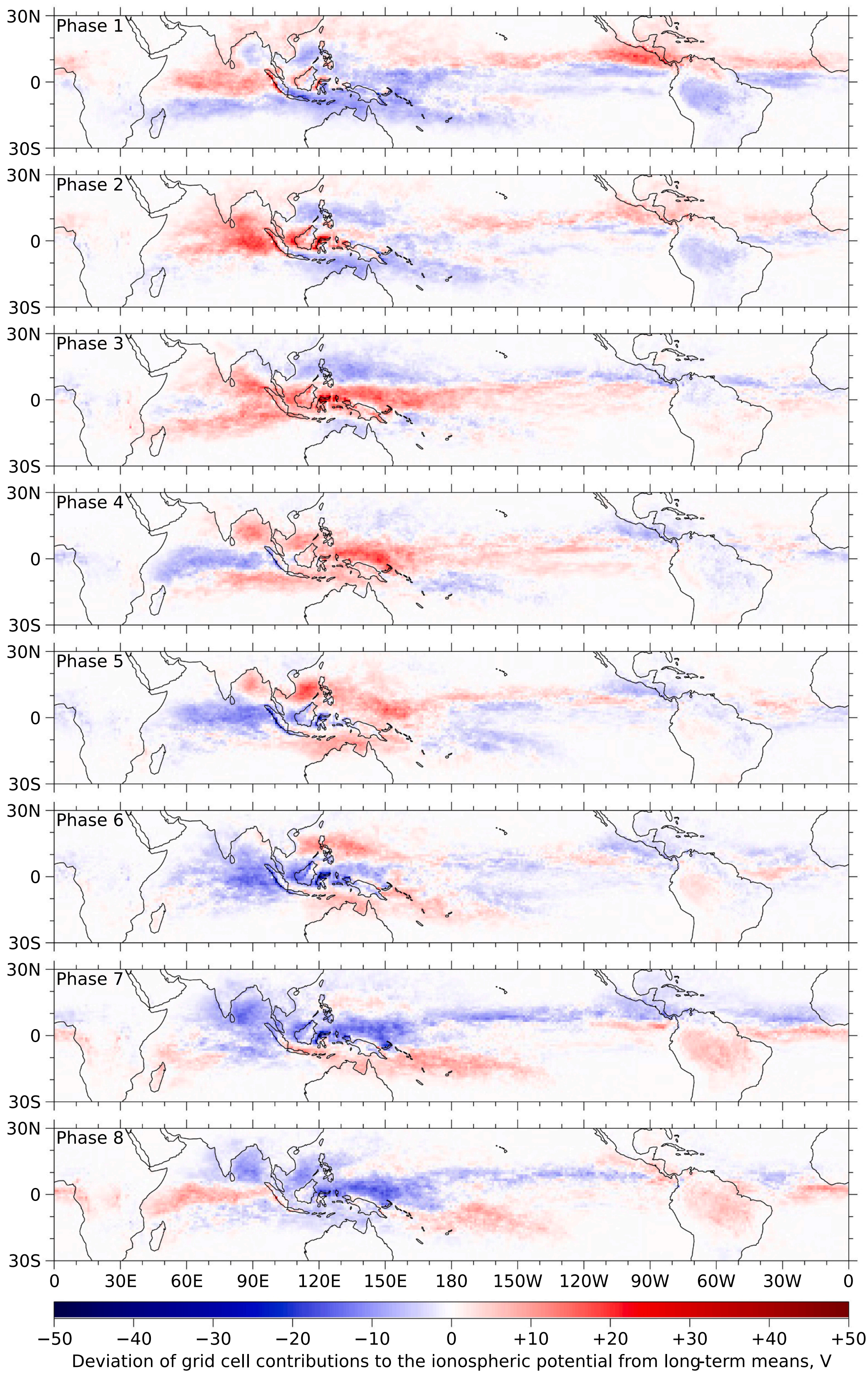

Fig. 2. Anomalies in contributions to the ionospheric potential from individual grid columns during each MJO phase (in other words, for the contribution of each grid column the long-term mean value is subtracted from average values corresponding to different MJO phases).

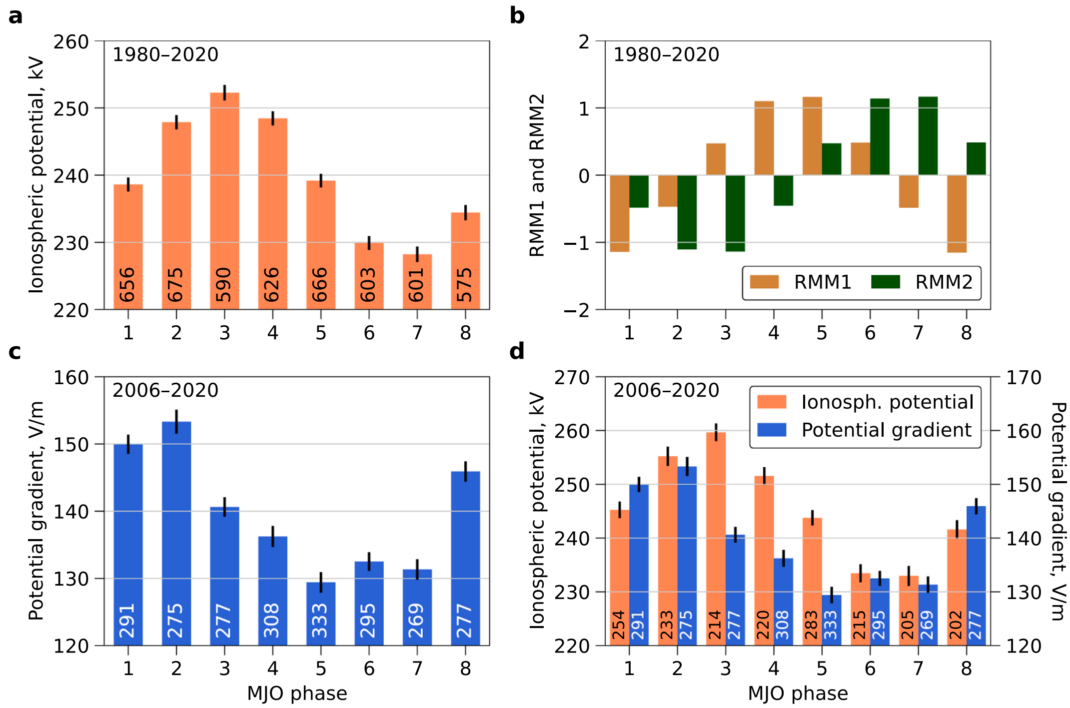

Fig. 3. (a): Average values of the simulated ionospheric potential for each MJO phase (on the basis of simulations for 1980–2020). (b): Average values of the RMM index components for each MJO phase (on the basis of the data for 1980–2020). (c): Average values of the fair-weather potential gradient at Vostok for each MJO phase (on the basis of measurements during 2006–2020). (d): Comparison between the potential gradient values shown in panel (c) and the values of the simulated ionospheric potential for the same time period (2006–2020). Annotations in panels (a), (c) and (d) indicate the number of days in the source data corresponding to each phase of the MJO. Error bars in panels (a), (c) and (d) show the values plus and minus one standard error.

MJO phases, are very tightly correlated with the MJO cycle. Note that the negative of RMM2 corresponds to $\phi = 270^\circ$, which is close to the optimum direction $\phi = 290^\circ$, whereas RMM1 corresponds to $\phi = 0^\circ$, which is nearly perpendicular to this direction. This agrees with our earlier observation that the IP is correlated with RMM2 but not with RMM1.

3.2. Measurements

We have just seen that the simulated IP manifests a sine-like variation on the MJO timescale. Now we shall look for a similar effect in the results of fair-weather PG measurements at the Antarctic Vostok station during 2006–2020 (see Section 2.4).

The mean values of the fair-weather PG at Vostok corresponding to different MJO phases are shown in Fig. 3c. Once again we see a sine-like variation, albeit less smooth than that of the IP and ahead in phase; this variation has a maximum in phase 2 and a minimum around phases 5–7. Strictly speaking, the PG exhibits two local minima, one in phase 5 and another one in phase 7, separated by a small local maximum in phase 6; however, considering that the PG values for phases 5–7 are close to each other, we do not regard the values at phase 5 and phase 7 as two separate minima.

Comparing the dynamics of the simulated IP for 1980–2020 (Fig. 3a) and the fair-weather PG at Vostok for 2006–2020 (Fig. 3c) on the MJO scale, we observe that the two variations are somewhat different from one another. To make such a comparison more substantiated, we should compare the PG values with the IP values averaged over the same 2006–2020 period; this is done in Fig. 3d. The correlation coefficient between the two variations is $r = 0.50$ (which is insufficient for the statistical significance at the 1% level), but there is still a phase shift of about one eighth of the period (i.e. one MJO phase) between them. If we shift the PG variation by one phase, thus eliminating this phase difference with the IP, the correlation coefficient increases, reaching the value $r = 0.90$. This gives us a statistically significant correlation at the 1% level.

It is also interesting to directly compare the fair-weather PG with components of the RMM index. The correlation coefficient between the PG and RMM1 is equal to $r = -0.68$; its counterpart for RMM2 is almost the same, namely $r = -0.67$. At the 1% level this does not yield statistical significance, for which we need $|r| \geqslant 0.83$; note, however, that we would have a lower threshold if we divided the (RMM1, RMM2) plane into more than eight phases.

If we consider arbitrary linear combinations of RMM1 and RMM2 (i.e. arbitrary directions in the (RMM1, RMM2) plane), the correlation coefficient between (2) and the fair-weather PG reaches its maximum value $r = 0.95$ at $\phi = 225^\circ$; this direction is shown in Fig. 1 by a blue dashed line. Note that it bisects the right angle between the two axes, which is consistent with our earlier remark that the PG values have almost the same degree of correlation with RMM1 and RMM2. We also observe that the optimum direction for the PG differs from that obtained earlier for the IP ($\phi = 290^\circ$); this difference suggests a phase shift of 65° (about one and a half MJO phases) between the two quantities, which agrees with Fig. 3d.

To summarise, both IP values and PG values (averaged over MJO phases) are tightly correlated with the MJO cycle but their variations on the MJO scale appear to have a conspicuous phase difference. Error bars in Fig. 3 denote the values of the IP and PG plus and minus one standard error (which is equal to the standard deviation of the mean). By looking at Fig. 3d, we conclude that the disagreement between the shapes of two variations cannot be attributed to statistical errors alone. Other possible reasons behind it include simulation uncertainties, parameterisation flaws, local effects at Vostok station and solar activity. We shall continue the discussion of these issues in Section 5.3.

4. Further analysis and explanation of the physical mechanism

Since we have inferred the relationship between the MJO and the IP from theoretical simulations with the WRF model, it is possible to analyse this relationship in more detail by extracting more information from contributions of individual grid columns. In this section we will try to decompose the IP variation into simpler oscillations with a view to understanding the physical mechanism behind the observed effect.

We have already mentioned that in studies of the MJO it is quite common to perform EOF analysis, decomposing the complicated variation into a sum of simpler components. EOFs are certain eigenvectors of the data which represent the basic spatial patterns, and the time-dependent coefficients of these patterns are called PCs. In Section 2.1 we have described how Wheeler and Hendon (2004) employed EOF analysis (combining zonal wind velocities and OLR into a single data set) to introduce the RMM index, which we use to formally distinguish the eight MJO phases. Here we shall apply EOF analysis to contributions to the IP; this will allow us to reduce the complicated IP variability to several simple oscillations, which are easier to interpret in terms of the MJO.

4.1. Preliminary processing of the IP data

Following Wheeler and Hendon (2004), we restrict our attention to the equatorial region bounded by the circles of latitude at 15° S and 15° N. According to our simulations, on average, this region contributes about 86% of the total IP, which makes it the key to understanding the MJO–GEC relationship. Summing simulated contributions to the IP over each of the 360 $1^{\circ}\times 30^{\circ}$ meridional strips near the equator (15° S–15° N in latitude and 0°–1° E, 1° E–2° E and so on up to 1° W–0° in longitude), we obtain one data set of length 360 for each day. Such reduction simplifies the complicated nature of the evolution of contributions to the IP during the MJO cycle by removing its latitudinal structure, which can be clearly seen in Fig. 2. Considering that the main aspect of the MJO is the eastward propagation of convection, in the first approximation it suffices to concentrate only on the longitudinal structure and its dynamics.

Now that the MJO always occurs simultaneously with other modes of climate variability, we will try to identify in the IP data patterns corresponding to non-MJO variations and remove them. A similar approach has been applied by Wheeler and Hendon (2004) prior to the EOF analysis employed to define the RMM index.

First, we remove the effect of the ENSO on the GEC intensity (e.g., Harrison et al., 2011; Slyunyaev et al., 2021c). To this end, we search for linear regression relationships between the diurnal mean contributions to the IP from $1^{\circ}\times 30^{\circ}$ equatorial strips at different longitudes and the sea surface temperature (SST) averaged over the Niño 3.4 region (5° S–5° N in latitude and 120° W–170° W in longitude), which quantitatively characterises the ENSO. Then for each $1^{\circ}\times 30^{\circ}$ strip we use the corresponding linear fit to remove ENSO variability from the time series of diurnal contributions while keeping their long-term mean value the same (more precisely, from contributions to the IP we subtract the values predicted by the linear fit on the basis of the Niño 3.4 SST on the respective days, adjusting the constant term of this linear fit such that the long-term mean does not change).

Next, we remove seasonal variability of the GEC (Adlerman and Williams, 1996). To do so, we apply the discrete Fourier transform to the time series of contributions to the IP from each $1^{\circ}\times 30^{\circ}$ strip, remove the first four harmonics of the seasonal cycle (those corresponding to $T = 365.25$ days, $T/2, T/3$ and $T/4$) and then compute a new time series by applying the inverse transform. Note that it is important to perform this step after the removal of ENSO variability, inasmuch as SST has its own seasonal harmonics and the subtraction of a time series proportional to the Niño 3.4 SST enhances seasonal harmonics of contributions to the IP. Also, note that removing particular Fourier harmonics does not change long-term mean values of contributions to the IP, since the mean value of each harmonic is zero.

Fig. 4a shows how the variation of the total IP with the MJO phase changes after restricting to contributions from the equatorial region (15° S–15° N) and removing non-MJO variability according to the procedure described above. To make the comparison easier, we pass to anomalies, subtracting the long-term mean values (240 kV for the total IP and about 207 kV for the equatorial contribution). We observe that the two variations are similar to each other, which supports our hypothesis that the observed variation of the total IP is primarily caused by the dynamics of the equatorial contribution.

4.2. Calculation of the EOFs and PCs for the IP data

Having removed ENSO variability and seasonal variability from the IP data, we obtain a time series of 360-dimensional vectors (one vector for each simulated day) containing contributions to the IP from consecutive $1^{\circ}\times 30^{\circ}$ equatorial meridional strips corresponding to different longitudes. Let us denote such vector for day d by

$$\mathbf{V}(d) = (V_1(d), \dots, V_{360}(d)).$$

In order to proceed with EOF analysis, we need to form a vector whose components have zero mean values, namely

$$\mathbf{U}(d) = (U_1(d), \dots, U_{360}(d)),$$

where $U_j(d) = V_j(d) - \langle V_j(d)\rangle$, the angle brackets denoting the long-term average over d.

Now we shall briefly describe the concept of EOF analysis (for more details, see, e.g., Zhang and Moore, 2015, Ch. 6) and apply it to $\mathbf{U}(d)$. We can interpret $U_1(d)$, …, $U_{360}(d)$ as the coordinates of the 360-dimensional vector $\mathbf{U}(d)$ with respect to the standard basis of $\mathbb{R}^{360}$ $\mathbf{e}^{(1)}$, …, $\mathbf{e}^{(360)}$. The main idea of EOF analysis is to choose another orthonormal

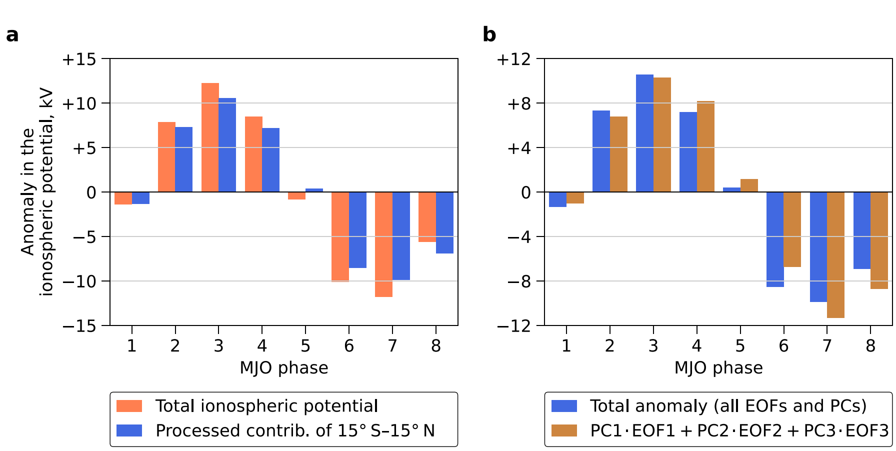

Fig. 4. (a): Anomalies in the total simulated ionospheric potential and in the contribution to the ionospheric potential of the equatorial region (15° S–15° N) corresponding to each MJO phase. Anomalies are computed relative to the long-term means. In the case of the equatorial contribution the variability related to the ENSO and seasonal variation have been removed from the data before plotting. (b): Anomalies in the contribution of the equatorial region to the ionospheric potential (the same as in panel (a)) and the part of its variability determined by the first three empirical orthogonal functions (EOFs).

basis $\mathbf{f}^{(1)}, \dots, \mathbf{f}^{(360)}$ such that the first few components (corresponding to $\mathbf{f}^{(1)}, \mathbf{f}^{(2)}$, etc.), in a sense, describe the main part of the variability of $\mathbf{U}(d)$. Thus we shall have

$$\mathbf{U}(d) = \sum_{j=1}^{360} U_j(d)\mathbf{e}^{(j)} = \sum_{j=1}^{360} C_j(d)\mathbf{f}^{(j)}; \tag{3}$$

the elements of the new basis $\mathbf{f}^{(1)}, \dots, \mathbf{f}^{(360)}$ will be our EOFs and the coefficients $C_1(d), \dots, C_{360}(d)$ will be our PCs. Note that the jth EOF $\mathbf{f}^{(j)}$ is actually a 360-dimensional vector, written as

$$\mathbf{f}^{(j)} = \left(f_1^{(j)}, \dots, f_{360}^{(j)}\right) \tag{4}$$

in the basis $\mathbf{e}^{(1)}, \dots, \mathbf{e}^{(360)}$, i.e. a function of longitude.

More precisely, the first EOF $\mathbf{f}^{(1)}$ is chosen so as to maximise the variance of $C_1(d)$, then the second EOF $\mathbf{f}^{(2)}$ is chosen among the remaining dimensions so as to maximise the variance of $C_2(d)$, etc. This means that the series

$$\mathbf{U}(d) = \sum_{j=1}^{360} C_j(d)\mathbf{f}^{(j)}, \tag{5}$$

or, in the original basis $\mathbf{e}^{(1)}, \dots, \mathbf{e}^{(360)}$,

$$(U_1(d), \dots, U_{360}(d)) = \sum_{j=1}^{360} C_j(d)\left(f_1^{(j)}, \dots, f_{360}^{(j)}\right),$$

decomposes the variation of $\mathbf{U}(d)$ into a sum of mutually orthogonal (in $\mathbb{R}^{360}$) components $C_1(d)\mathbf{f}^{(1)}, \dots, C_{360}(d)\mathbf{f}^{(360)}$, the first few of which account for a large part of the data variance. For more precise definitions and a brief description of the procedures used to compute the EOFs and PCs, see Appendix A.

According to (3), the anomaly of the total equatorial contribution to the IP (with non-MJO variability removed) on day d is given by

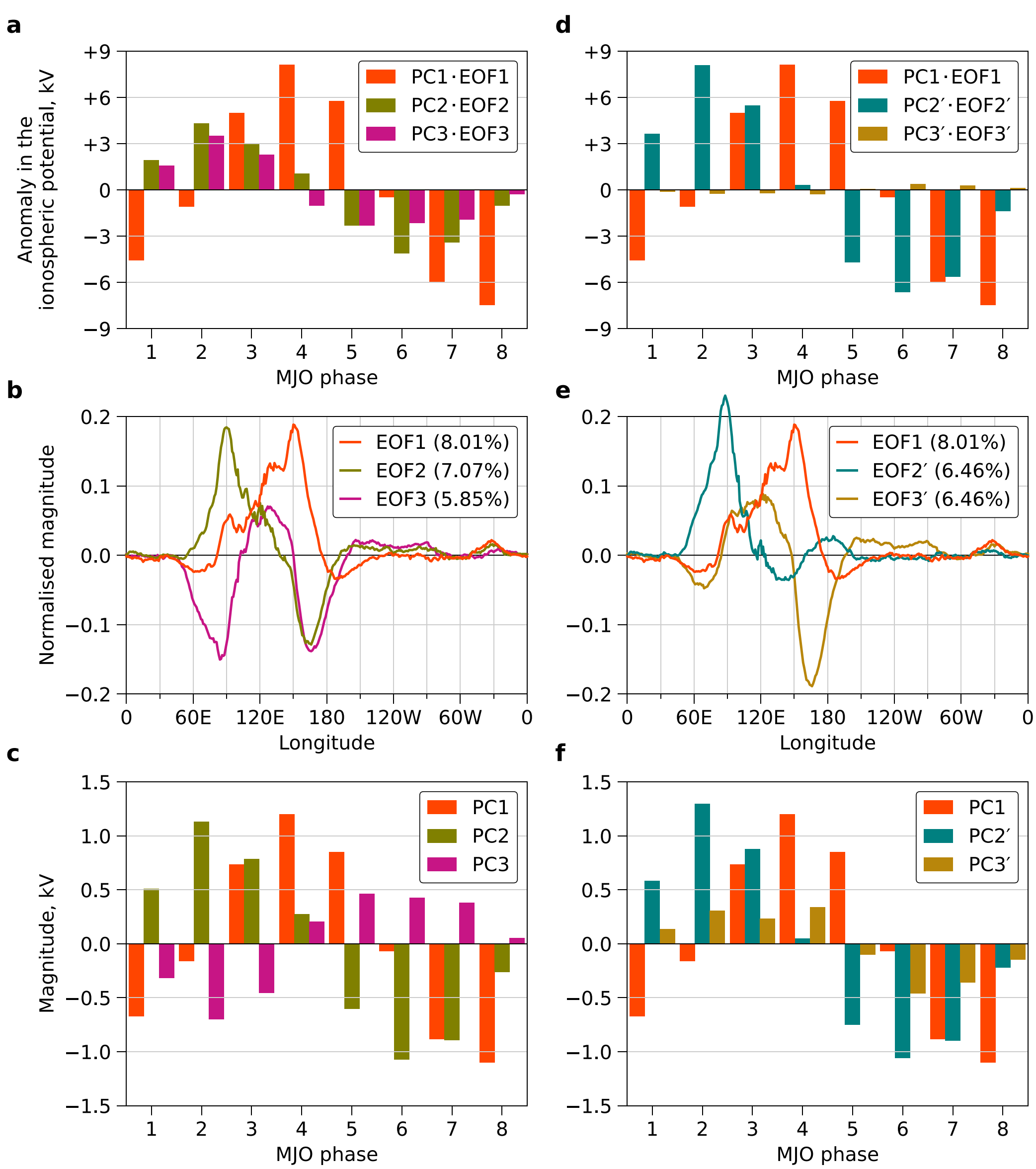

Fig. 5. (a): The variability of the contribution from the equatorial region (15° S–15° N) to the ionospheric potential on the MJO scale corresponding to each of the first three empirical orthogonal functions (EOFs) taken separately. (b): The spatial structure of the first three EOFs. (c): The principal components (PCs) corresponding to the first three EOFs averaged over different MJO phases. (d)–(f): Similar to (a)–(c) for the adjusted (rotated) EOFs. The numbers in parentheses in panels (b) and (e) indicate portions of the total variance explained by specific EOFs.

$$U_{\mathrm{tot}}(d)=\sum_{j=1}^{360}U_j(d)=\sum_{i,j=1}^{360}C_i(d)f_j^{(i)}. \tag{6}$$

If instead of the full sum in (5) we consider its main part

$$\widetilde{\mathbf{U}}(d)=C_1(d)\mathbf{f}^{(1)}+C_2(d)\mathbf{f}^{(2)}+C_3(d)\mathbf{f}^{(3)}, \tag{7}$$

where only the first three EOFs are retained, the corresponding total anomaly will be

$$\widetilde{U}_{\mathrm{tot}}(d)=\sum_{j=1}^{360}C_1(d)f_j^{(1)}+\sum_{j=1}^{360}C_2(d)f_j^{(2)}+\sum_{j=1}^{360}C_3(d)f_j^{(3)}. \tag{8}$$

Fig. 4b compares the average values of U_{tot} and $\widetilde{U}_{\mathrm{tot}}$ for particular MJO phases. We notice that the first three EOFs already produce the sine-like variation observed in the full sum, which suggests that it suffices to consider (7) and (8) instead of (5) and (6) in order to explore the physical mechanism behind the effect.

Let us now decompose the variation of $\widetilde{U}_{\mathrm{tot}}$ on the MJO scale, shown in Fig. 4b, into separate components corresponding to the first three EOFs. Fig. 5a demonstrates each of the three summands in (8) averaged over different MJO phases. The contribution corresponding to EOF1, namely

$$U_{\mathrm{tot}}^{(1)}(d)=C_1(d)\sum_{j=1}^{360}f_j^{(1)}, \tag{9}$$

has a sine-like variation with a maximum in phase 4 and a minimum in phase 8, whereas those corresponding to the EOF2 and EOF3, that is,

$$U_{\mathrm{tot}}^{(2)}(d)=C_2(d)\sum_{j=1}^{360}f_j^{(2)}, \tag{10}$$

$$U_{\mathrm{tot}}^{(3)}(d)=C_3(d)\sum_{j=1}^{360}f_j^{(3)}, \tag{11}$$

have similar sine-like variations with maxima in phase 2 and minima around phase 6.

Fig. 5b shows the first three EOFs (EOF1 $\mathbf{f}^{(1)}$, EOF2 $\mathbf{f}^{(2)}$ and EOF3 $\mathbf{f}^{(3)}$) as functions of longitude (in accordance with (4)). EOF1 describes a pronounced anomaly in contributions to the IP located between 80° E and 180° with a maximum at about 150° E. EOF2 and EOF3 are most pronounced between 50° E and 170° W; they both have a prominent minimum between 160° E and 170° E, whereas around 90° E they have extrema of opposite signs.

To make the situation clearer, in Fig. 5c we demonstrate average values of the PCs corresponding to the first three EOFs (i.e. PC1 C_1, PC2 C_2 and PC3 C_3) for each MJO phase. We notice that average PCs are proportional to average anomalies in total contributions from different EOFs, shown in Fig. 5a; this is actually an easy consequence of Eqs. (9)–(11) and the fact that EOFs only describe the spatial structure and do not change with time.

4.3. Adjustment of the EOF basis

Fig. 5c indicates that PC2 and PC3 vary in antiphase on the MJO scale; at the same time the corresponding EOFs, EOF2 and EOF3, as functions of longitude show similar behaviour to the east of about 150° E but vary oppositely around 90° E (see Fig. 5b). Therefore we shall obtain somewhat simpler spatial patterns if in our basis of EOFs we change EOF2 $\mathbf{f}^{(2)}$ and EOF3 $\mathbf{f}^{(3)}$ to

$$\mathbf{f}^{(2')}=\frac{\mathbf{f}^{(2)}-\mathbf{f}^{(3)}}{\sqrt{2}},\quad \mathbf{f}^{(3')}=\frac{\mathbf{f}^{(2)}+\mathbf{f}^{(3)}}{\sqrt{2}},$$

also denoted below by EOF2′ and EOF3′ respectively. Accordingly, instead of PC2 $C_2(d)$ and PC3 $C_3(d)$ we shall have

$$C_{2'}(d)=\frac{C_2(d)-C_3(d)}{\sqrt{2}},\quad C_{3'}(d)=\frac{C_2(d)+C_3(d)}{\sqrt{2}}$$

as the new PCs PC2′ and PC3′. This technique is called rotation of EOFs (see, e.g. Zhang and Moore, 2015, Ch. 6; Hannachi, 2021, Ch. 4); roughly speaking, it allows one to adjust the automatically selected basis to a specific problem (note that we computed our original EOFs without making any assumptions about the MJO, its timescale and its phases).

We can now repeat the previous analysis using the adjusted EOFs $\mathbf{f}^{(1)}$, $\mathbf{f}^{(2')}$ and $\mathbf{f}^{(3')}$ instead of the original basis $\mathbf{f}^{(1)},\mathbf{f}^{(2)},\mathbf{f}^{(3)}$. Eqs. (7) and (8) turn into

$$\widetilde{\mathbf{U}}(d)=C_1(d)\mathbf{f}^{(1)}+C_{2'}(d)\mathbf{f}^{(2')}+C_{3'}(d)\mathbf{f}^{(3')}, \tag{12}$$

$$\widetilde{U}_{\mathrm{tot}}(d)=\sum_{j=1}^{360}C_1(d)f_j^{(1)}+\sum_{j=1}^{360}C_{2'}(d)f_j^{(2')}+\sum_{j=1}^{360}C_{3'}(d)f_j^{(3')}, \tag{13}$$

Eq. (9) remains the same, and instead of (10) and (11) we now have

$$U_{\mathrm{tot}}^{(2')}(d)=C_{2'}(d)\sum_{j=1}^{360}f_j^{(2')}, \tag{14}$$

$$U_{\mathrm{tot}}^{(3')}(d)=C_{3'}(d)\sum_{j=1}^{360}f_j^{(3')}. \tag{15}$$

Fig. 5e and f is the counterpart of Fig. 5b and c for the adjusted EOFs EOF1, EOF2′, EOF3′ (i.e. $\mathbf{f}^{(1)},\mathbf{f}^{(2')},\mathbf{f}^{(3')}$) and the new PCs PC1, PC2′, PC3′ (i.e. $C_1(d),C_{2'}(d),C_{3'}(d)$). Comparing Fig. 5b with Fig. 5e, we ascertain that the spatial structure of the new basic patterns is indeed simpler than that of the original ones. In particular, EOF2′ describes a pronounced anomaly in contributions to the IP located between 50° E and 120° E with a single prominent maximum near 90° E. As for EOF3′, its spatial structure is somewhat more complicated, with the main extremum near 170° E. However, by looking at the corresponding PCs during different MJO phases (Fig. 5f) we infer that on average, the absolute value of PC3′ (which is the coefficient of EOF3′ in (12)) is considerably smaller than those of PC1 and PC2′, which means that EOF3′ will be less important to us.

Next, we pass to anomalies of the total equatorial contribution to the IP. Fig. 5d demonstrates the three summands in the new decomposition (13), namely (9), (14) and (15), averaged over different MJO phases (similarly to Fig. 5a showing the three summands in (8)). It is easy to see that the summand (15), corresponding to EOF3′, becomes negligible. This is not surprising: not only is PC3′ small in comparison with the other two coefficients but also positive and negative terms in the sum of 360 longitudinal components of EOF3′ offset each other (see the plot of EOF3′ in Fig. 5e). The other two summands (9) and (14), corresponding to EOF1 and EOF2′, manifest clear sine-like oscillations of similar magnitudes, one of which is ahead of the other by one quarter of the period (i.e. by two MJO phases). The variation of the anomaly of the total equatorial contribution to the IP, shown in Fig. 4, can be to a large extent approximated by the sum of these basic oscillations.

4.4. Longitudinal structure of basic oscillations

Thus the adjustment of the original EOF basis has made it possible to reduce the observed IP variation with the MJO phase (Fig. 3a) to a superposition of two basic oscillations, obtained by averaging (9) and (14) over specific phases (Fig. 5d). In terms of spatial patterns, the variations (9) and (14) come from summing the components

$$\mathbf{U}^{(1)}(d)=C_1(d)\mathbf{f}^{(1)} \tag{16}$$

and

$$\mathbf{U}^{(2')}(d) = C_{2'}(d)\mathbf{f}^{(2')} \tag{17}$$

over longitudes. We will now analyse these components in more detail, averaging them over MJO phases without longitudinal summation (i.e., keeping the spatial structure).

The left column of Fig. 6 shows the average longitudinal structure of the anomalies in equatorial contributions to the IP (i.e. of $\mathbf{U}(d)$ given by (5)) for different MJO phases. This is a one-dimensional representation of Fig. 2: anomalies in pixel contributions to the IP, shown in Fig. 2, add up to anomalies in contributions of $1^\circ \times 30^\circ$ meridional strips, shown in the left column of Fig. 6; that is why red and blue regions in Fig. 2, indicating positive and negative anomalies respectively, generally correspond to the intervals of positive and negative values in the left column of Fig. 6.

The middle column of Fig. 6 demonstrates how the left column changes if instead of the whole sum (5), which can be rewritten in terms of the adjusted EOFs as

$$\mathbf{U}(d) = C_1(d)\mathbf{f}^{(1)} + C_{2'}(d)\mathbf{f}^{(2')} + C_{3'}(d)\mathbf{f}^{(3')} + \sum_{j=4}^{360} C_j(d)\mathbf{f}^{(j)},$$

we consider only its main part given by the sum of (16) and (17). Comparing the two columns of the figure, we deduce that the first two adjusted EOFs, EOF1 and EOF2′, are sufficient to capture the main patterns in the dynamics of contributions to the IP on the MJO scale.

Finally, the right column of Fig. 6 decomposes the middle column into two components, showing separately the average anomalies corresponding to each of the first two (adjusted) EOFs, given by (16) and (17). We observe that the complicated variability on the MJO scale shown in the middle column is just a superposition of two simple oscillations, whose longitudinal structures never change and are determined by EOF1 and EOF2′, plotted in Fig. 5e, and whose magnitudes

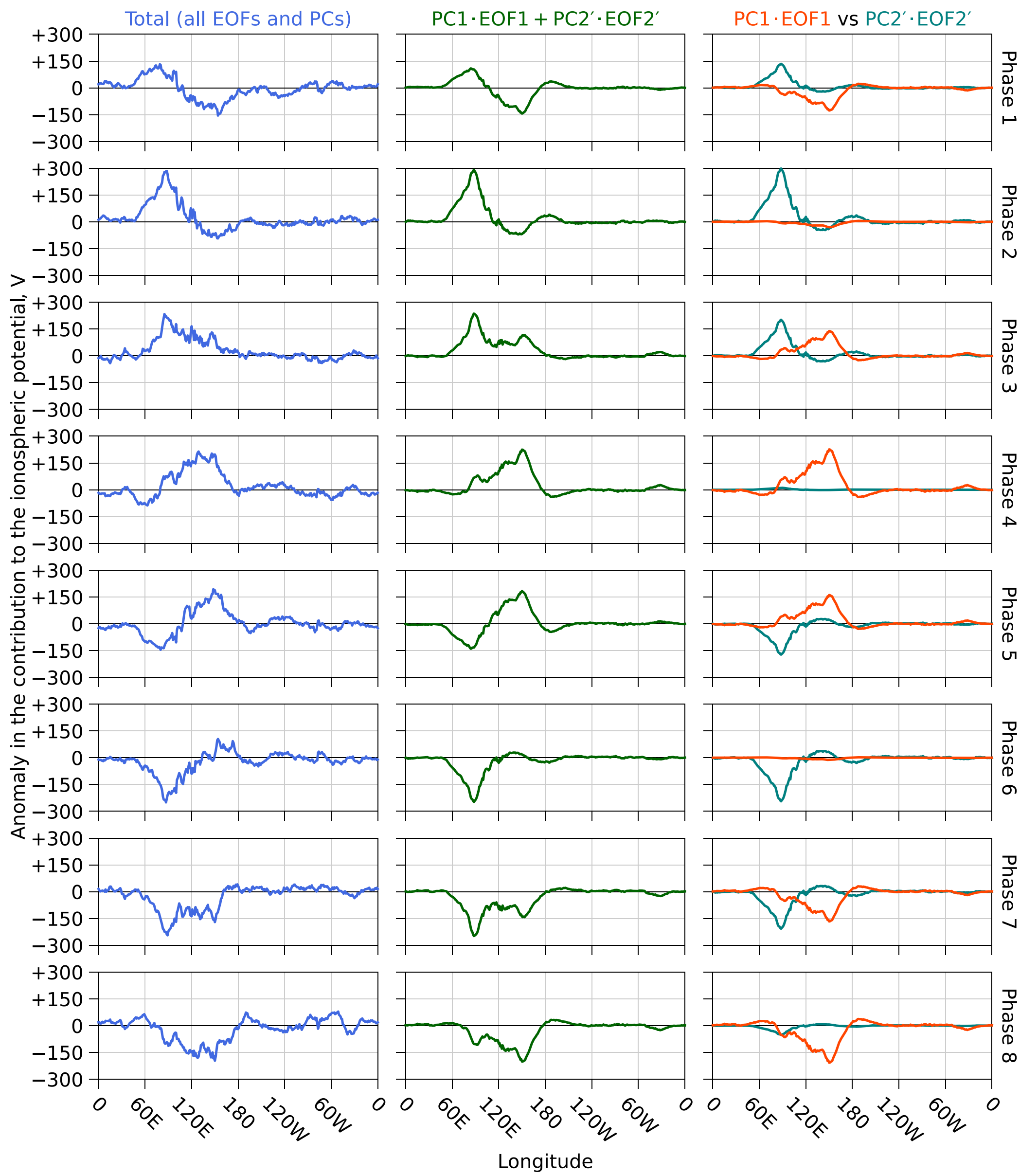

Fig. 6. Left column: Anomalies in the contributions to the ionospheric potential of $1^\circ \times 30^\circ$ meridional strips (15° S to 15° N) corresponding to grid cells along the equator during each MJO phase. Anomalies are computed relative to the long-term means. The variability related to the ENSO and seasonal variation has been removed from the data before plotting. Middle column: The same with only the variability determined by the first two adjusted (rotated) empirical orthogonal functions (EOFs) retained. Right column: The variability corresponding to the first two adjusted EOFs shown separately.

exhibit sine-like variations determined by the PC1 and PC2′, plotted in Fig. 5f.

Thus for the component (16), corresponding to EOF1, the main part of the spatial structure is located between 80° E and 180° with a pronounced maximum at about 150° E, and this structure on average exhibits a periodic variation on the MJO scale with a maximum in phase 4 and a minimum in phase 8. Similarly, for the component (17), corresponding to EOF2′, the spatial structure is most conspicuous between 50° E and 120° E and has a prominent maximum at about 90° E; its average variation with the MJO phase has a maximum in phase 2 and a minimum in phase 6.

4.5. The link between the IP and the MJO

To summarise the above, we have identified two basic sinusoidal oscillations of contributions to the IP with the phase of the MJO (see the right column of Fig. 6), one of which is primarily located over the Maritime Continent and its surrounding seas (around 80° E–180°) and the other is mainly located over the Indian Ocean and part of South-East Asia (around 50° E–120° E). The latter oscillation is ahead of the former by one quarter of the period, that is, by two MJO phases; when contributions corresponding to different longitudes are added together, the two oscillations reach similar magnitudes (see Fig. 5d) and in combination produce a sine-like variation with a maximum in phase 3 and a minimum in phase 7. This accounts for the majority of the variability of the total IP on the MJO scale (see Fig. 4) and thus reduces the problem of explaining the variation shown in Fig. 3a to understanding the aforementioned pair of basic oscillations.

Let us now recall that the two-dimensional RMM index, which we used to formally distinguish eight MJO phases, has been introduced by Wheeler and Hendon (2004) on the basis of EOF analysis applied to a combined data set of zonal wind velocities and OLR (which serves as a proxy for deep convection; see Section 2.1). If we look at their EOFs, we observe that the first one corresponds to a pronounced negative anomaly in OLR between 70° E and 180° with a flat maximum between 120° E and 140° E, while the second one describes a prominent positive anomaly between 40° E and 110° E with a maximum near 80° E (see the curves for OLR in Wheeler and Hendon, 2004, Fig. 1). Considering that contributions to the IP estimated using (1) also reflect the geographical distribution of deep convection, it is natural to search for a correspondence between our EOFs for the IP and the EOFs of Wheeler and Hendon (2004) for OLR. Comparing the two sets of EOFs, we conclude that our EOF1 and EOF2′ describe more or less the same convection patterns as the first two EOFs of Wheeler and Hendon (2004). More precisely, EOF1 and EOF2 for OLR approximately match our EOF1 and the negative of our EOF2′ for the IP (i.e. $\mathbf{f}^{(1)}$ and $-\mathbf{f}^{(2')}$) respectively; note that the greater is OLR, the smaller are the corresponding contributions to the IP.

The first two PCs obtained by Wheeler and Hendon (2004) are precisely RMM1 and RMM2. Under our correspondence RMM1 and RMM2 should approximately match (up to sign) our PC1 and PC2′; more precisely, RMM1 should match our PC1 (C_1) and RMM2 should match the negative of our PC2′ ($-C_{2'}$). In accordance with the definition of MJO phases (see also Fig. 3b), RMM1 exhibits a nearly sinusoidal variation with a maximum between phase 4 and phase 5 while RMM2 manifests a similar variation with a maximum between phase 6 and phase 7. At the same time, looking once again at Fig. 5f, we recall that C_1 and $-C_{2'}$ attain their maxima in phase 4 and in phase 6 respectively. This means that on the MJO scale our PC1 and PC2′ indeed describe more or less the same oscillations of convection as RMM1 and RMM2, with only a slight phase shift. This small phase shift can be attributed to the difference between OLR and contributions to the IP, as well as to the fact that here we consider only the two main components of each variable.

Thus we see that the two main oscillating patterns in anomalies of contributions to the IP nearly match the oscillating OLR patterns that were used by Wheeler and Hendon (2004) to define the RMM index. More generally, our EOFs for contributions to the IP look similar to typical sets of basic OLR patterns found in the literature on the MJO. For example, the first three EOFs for OLR obtained by Kessler (2001, Fig. 2) appear to match our original EOF2, EOF1 and EOF3, plotted in Fig. 5b, respectively; similarly, two-dimensional EOFs for OLR obtained by Lo and Hendon (2000, Fig. 2) and Matthews (2000, Fig. 2) seem to describe the same patterns as our original EOF1 and EOF2 (in the latter case we also need to change their order). In this connection one can say that the EOFs obtained by Wheeler and Hendon (2004) are different from other sets of basic patterns, as they more directly correspond to our adjusted functions EOF1 and EOF2′, plotted in Fig. 5e.

We conclude that the two basic oscillations in anomalies of contributions to the IP, which account for the majority of the total IP variability with the MJO phase, approximately correspond to basic oscillating convection patterns typical for the MJO, in particular to those used to define the RMM index. In view of the fact that on average, both components of the RMM manifest sinusoidal variations with the MJO phase (Fig. 3b), it is not surprising that we obtain a sine-like variation for the IP as well. In a sense, contributions to the IP provide another quantitative characteristic of deep convection; that is why the eastward propagation of convective structures typical for the MJO is reflected in the IP and in the GEC in general.

5. Further discussion

5.1. The observed relationship and its reproducibility

The results presented in this study clearly indicate that the MJO influences the intensity of the direct current GEC. We have established that on average, both the simulated IP and the fair-weather PG measured in the Antarctic manifest clear sine-like variations with the MJO phase and exhibit a statistically significant correlation with indices characterising the MJO.

It is important to emphasise that the behaviour of GEC parameters on the MJO scale shown in Fig. 3a and c is not a mere coincidence. Similarly to Fig. 3a, Fig. 7a–d demonstrates the dependence of the IP on the MJO phase, but now instead of the entire period of 1980–2020, IP values are averaged over the four consecutive decades from 1981–1990 to 2011–2020. Although typical values of the simulated IP are not the same for different decades, the sine-like shape of the variation is clearly visible in each case.

Likewise, Fig. 7e and f shows the counterparts of Fig. 3c for two overlapping 10-year periods, 2006–2015 and 2011–2020 (we could have split the period of 2006–2020, for which we have the PG data, into non-overlapping intervals, but we do not do so since we want these intervals to be large enough in order for average trends to be reliable). Comparing Fig. 7e and f with Fig. 3c, we see that the shape of the PG variation with the MJO phase is more or less the same regardless of the time period used for averaging (although in the case of 2011–2020 the maximum shifts from phase 2 to phase 1, the PG values corresponding to these phases are within one standard error of each other).

Thus we have identified in the GEC variation some patterns clearly linked to the MJO. Despite the fact that connections between the MJO and Schumann resonances have been investigated before (Anyamba et al., 2000; Beggan and Musur, 2019), in the literature there appear to be no reports of the influence of the MJO on the direct current GEC, except that the periodicity of about 45 days has been recently identified in the PG spectrum by Tacza et al. (2022). Together with the results concerning links between the GEC and the ENSO (Harrison et al., 2011; Lavigne et al., 2017; Slyunyaev et al., 2021a–c), our findings indicate that the GEC is an important part of the Earth system, which reflects climate variability on different timescales. Note, however, that for both the ENSO and the MJO one needs to average everything over a large number of days in order to identify clear trends in the GEC behaviour.

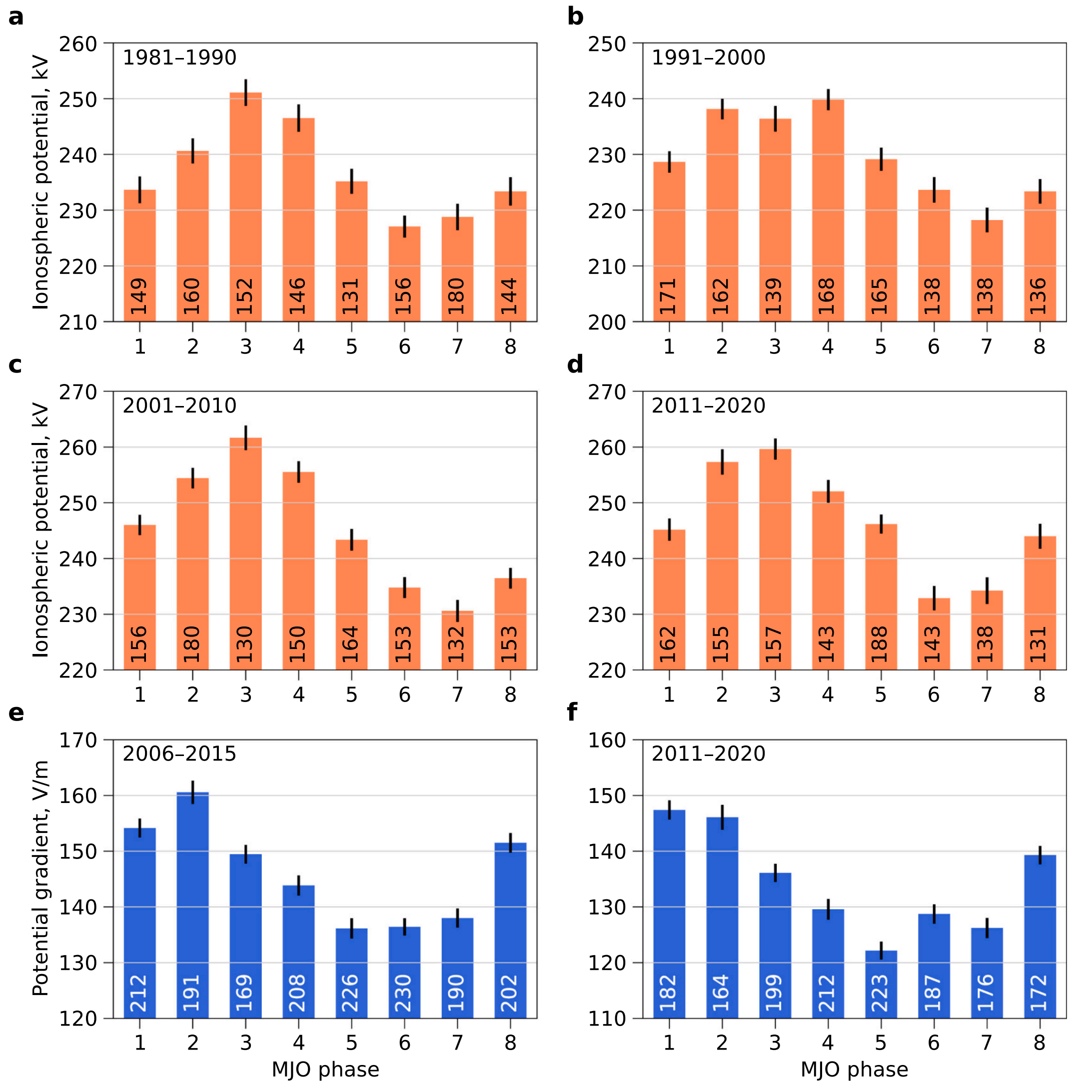

Fig. 7. (a)–(d): Average values of the simulated ionospheric potential for each MJO phase on the basis of simulations for the four decades during 1981–2020 considered separately. (e), (f): Average values of the fair-weather potential gradient at Vostok for each MJO phase on the basis of measurements during 2006–2015 and 2011–2020. Annotations in each panel indicate the number of days in the source data corresponding to each phase of the MJO; error bars show the values plus and minus one standard error.

5.2. Further remarks concerning EOF analysis

It is important to point out that simulations with the WRF model allowed us not only to identify the effect of the MJO on the GEC but also to reveal the mechanism behind this effect. To this end, we applied EOF analysis, commonly used to study the MJO, to anomalies in contributions to the IP in the equatorial region averaged along the meridians. The EOFs provide a basis of spatial patterns selected such that the coefficients (PCs) of the first few of them describe as much of the total variance as possible; this allowed us to reduce the complicated original process to a superposition of a few simple oscillating patterns.

The numbers in parentheses in Fig. 5b and e indicate which portion of the total variance is explained by each EOF. Although the first three EOFs together explain only about 21% of the total variance, one should understand that the IP data contain variability on all timescales and the larger part of this variability does not matter for the MJO. We can illustrate this with Fig. 5d, which demonstrates that the contribution of EOF3′ to the variation on the MJO scale is negligible in comparison with those of EOF1 and EOF2′; at the same time EOF3′ explains about 6% of the total variance, which is close to the explained variances of EOF1 (about 8%) and EOF2′ (about 6%).

As demonstrated in Section 4.5, the first few EOFs for contributions to the IP actually describe more or less the same patterns as the first few EOFs for OLR in the literature on the MJO. The variation of the total simulated IP on the MJO scale (Fig. 3a) is to a large extent produced by a sum of two basic oscillating patterns EOF1 and EOF2′, which roughly correspond to the basic patterns of convection for the RMM index. Accordingly, the corresponding PCs PC1 and PC2′ can be naturally associated with RMM1 and RMM2. Thus, from the perspective of the MJO, contributions to the IP generally reproduce the dynamics of deep convection, which explains the variation of GEC parameters with the phase of the MJO. In a sense, simulated contributions to the IP provide another means of describing the propagation of convection during the MJO cycle. One could even introduce a new index describing the MJO on the basis of contributions to the IP (in place of commonly used characteristics of convection such as OLR).

5.3. Phase shift between the IP and the PG

As shown in Section 3, both the simulated IP and the fair-weather PG at Vostok manifest a clearly visible and statistically significant correlation with the MJO cycle. At the same time we have noticed a phase shift between the variations of the two parameters: the IP reaches its maximum in phase 3 and shows the highest correlation with (2) when $\phi = 290^\circ$, whereas the PG attains its maximum in phase 2 and exhibits the highest correlation with (2) when $\phi = 225^\circ$ (this is actually the boundary between phase 1 and phase 2).

From a theoretical perspective, the fair-weather PG at the Earth's surface is proportional to the total IP. This is owing to the global nature of the GEC: the currents driven by charge separation in thunderstorms and ESCs all over the Earth flow upwards, together maintaining the total IP, then spread out in the ionosphere and eventually return to the Earth's surface in fair-weather regions. Despite the fact that local conductivity perturbations may substantially affect surface electric field measurements, in clean locations after averaging over many fair-weather days the results of such measurements are supposed to reflect main trends in the IP dynamics. Therefore it is rather unusual that both the IP and the PG are tightly correlated with the MJO cycle, but specific characteristics of the MJO which provide the most prominent correlation (chosen among linear combinations of RMM1 and RMM2) are different for the two quantities.

5.3.1. Possible shortcomings of the IP parameterisation and simulation

It is natural to suppose that the observed discrepancy between the IP variation and the PG variation may be caused by shortcomings of simulations, for it is clear that both large-scale modelling of the atmospheric dynamics with the WRF and our parameterisation of the IP in terms of climate variables are not precise. It is likely that such simulations to a certain extent overestimate or underestimate contributions to the IP from different regions of the Earth's surface.

Considering that our parameterisation generally predicts large contributions to the IP from the Maritime Continent and Eastern Pacific Ocean (see, e.g., Ilin et al., 2020), it is possible that these contributions are somewhat overestimated in simulations. In view of the fact that our EOF1 characterises a prominent anomaly in contributions to the IP across this region (Fig. 5e), such an overestimation could also have increased the magnitude of the part of the IP variability on the MJO scale which corresponds to EOF1. Note that this is just a hypothesis, since the magnitude of the anomaly of any variable need not be proportional to the mean value of the variable itself. If this is the case, however, then the contribution corresponding to EOF1 in Fig. 5d should actually have a lower amplitude than that corresponding to EOF2′. This would make the variation of the total IP, largely determined by the sum of these contributions, dominated by the component corresponding to EOF2′, which would shift the maximum of the sine-like variation shown in Fig. 3a to the left (here and below we imagine certain interpolations of eight-phase bar plots, where maxima and minima can occur between integer phase numbers).

Thus decreasing the contributions of the Maritime Continent and Eastern Pacific Ocean in simulations might reduce the phase difference between the variations of the IP and PG. Note, however, that a shift of the position of the maximum in Fig. 3a cannot be large: even if we set to zero the magnitude of the anomaly corresponding to EOF1, the maximum of the sum over the first two EOFs would shift by only one MJO phase, as one sees immediately from Fig. 5d. In addition, the analogy between our first two PCs and the components of the RMM index (see Section 4.5) predicts similar magnitudes of PC1 and PC2′ (see Fig. 3b) and, by inference, similar magnitudes of the components of the IP variability shown in Fig. 5d, which contradicts the hypothesis of overestimation. Moreover, the same analogy suggests that the maximum of our PC1 (roughly corresponding to RMM1) should be attained close to phases 4–5, where RMM1 reaches its maximum; similarly, the maximum of our PC2′ (roughly corresponding to the negative of RMM2) should be close to the minimum of RMM2 in phases 2–3. These considerations support the variations shown in Fig. 5d and f and indicate that it is unlikely for the maximum of the total IP (determined by the combination of the components shown in Fig. 5d) to be reached in phase 2.

To make the correspondence between the IP and deep convection easier to comprehend, we also consider the variation with the MJO phase of the OLR flux averaged over the equatorial region lying between 15° S and 15° N. This variation is shown in Fig. 8a. We observe that the mean equatorial OLR precisely mirrors the variation of the simulated IP (Fig. 3a) with a clear minimum in phase 3 and a flat maximum in phase 7. If we consider correlations with arbitrary linear combinations of RMM1 and RMM2, as we did in Section 3 for the IP and the PG, we obtain a maximum of the correlation coefficient between (2) and the negative of OLR at $\phi = 300^\circ$, with the value $r = 0.98$. This direction is shown in Fig. 1 by a teal dashed line, and it is definitely closer to the corresponding direction for the IP ($\phi = 290^\circ$) than to its counterpart for the PG ($\phi = 225^\circ$).

Considering that the OLR data that we use are based on satellite observations (see Section 2.5), which are independent of our simulations, this implies that our parameterisation is more or less reliable in reproducing the variation of equatorial convection on the scale of the MJO—at least the parameterisation is not that wrong in determining the phase of the variation. At the same time it is quite possible that it slightly underestimates the amplitude, considering that the simulated IP variation shown in Fig. 3a has a smaller amplitude than that of the PG at Vostok shown in Fig. 3c and we know that for the diurnal variation of the GEC large-scale simulations also predict a smaller amplitude than the value inferred from measurements (Ilin et al., 2020).

Thus the IP variation predicted by our simulations is consistent with the dynamics of convection in the equatorial region lying between 15° S and 15° N inferred from OLR data. This is not unexpected, as we have mentioned earlier that in our simulations the region between 15° S and 15° N contributes about 86% of the total IP and thereby determines its variation with the MJO phase (see Fig. 4a). However, it is still possible that our parameterisation somewhat underestimates the contributions of higher latitudes, which may respond differently to the MJO. This idea is supported by Fig. 8c and d, which shows the variation of the simulated CAPE with the MJO phase, when averaged over 15° S–15° N and over 30° S–30° N respectively. As is the case with our simulated IP (Fig. 3a) and the negative of the equatorial OLR (Fig. 8a), the average CAPE for 15° S–15° N (shown in Fig. 8c) reaches its main maximum in phase 3. At the same time if we average CAPE over a wider region between 30° S and 30° N (see Fig. 8d), this maximum shifts to phase 2, thus yielding a variation similar to that of the PG at Vostok (Fig. 3c).

These observations suggest that higher latitudes (beyond 15° S–15° N) may actually determine a larger part of the total IP than the 14% predicted by our parameterisation. Note, however, that CAPE itself does not provide a good proxy for contributions to the IP; for example, on the diurnal timescale a parameterisation based on CAPE alone (without precipitation) notably underestimates the magnitude of the variation and incorrectly predicts the hour of its maximum (see Ilin et al., 2020, Fig. 3b). On the MJO scale the situation is similar. Even though the variation of CAPE shown in Fig. 8d agrees in phase with that of the PG, its peak-to-peak amplitude is much smaller. Also, according to Fig. 8c, the equatorial CAPE has a secondary maximum in phase 8, which is not observed in the variations of the IP, PG or equatorial OLR. Therefore, although the variation of CAPE implies that our parameterisation of the IP may overestimate the equatorial contribution relative to higher latitudes, its improvement requires further research.

In this connection it is also interesting to note that the MJO is supposed to be linked to the active and break phases of the Indian summer monsoon (Madden and Julian, 1994), which should enhance convection over Southern Asia precisely during phase 2 of the MJO. If our parameterisation of the IP indeed underestimates contributions of non-equatorial regions (including Southern Asia), this may partially explain the observed phase difference between the PG and the simulated IP (with the maximum of the former in phase 2 and the maximum of the latter shifted to phase 3). Note, however, that this factor can be

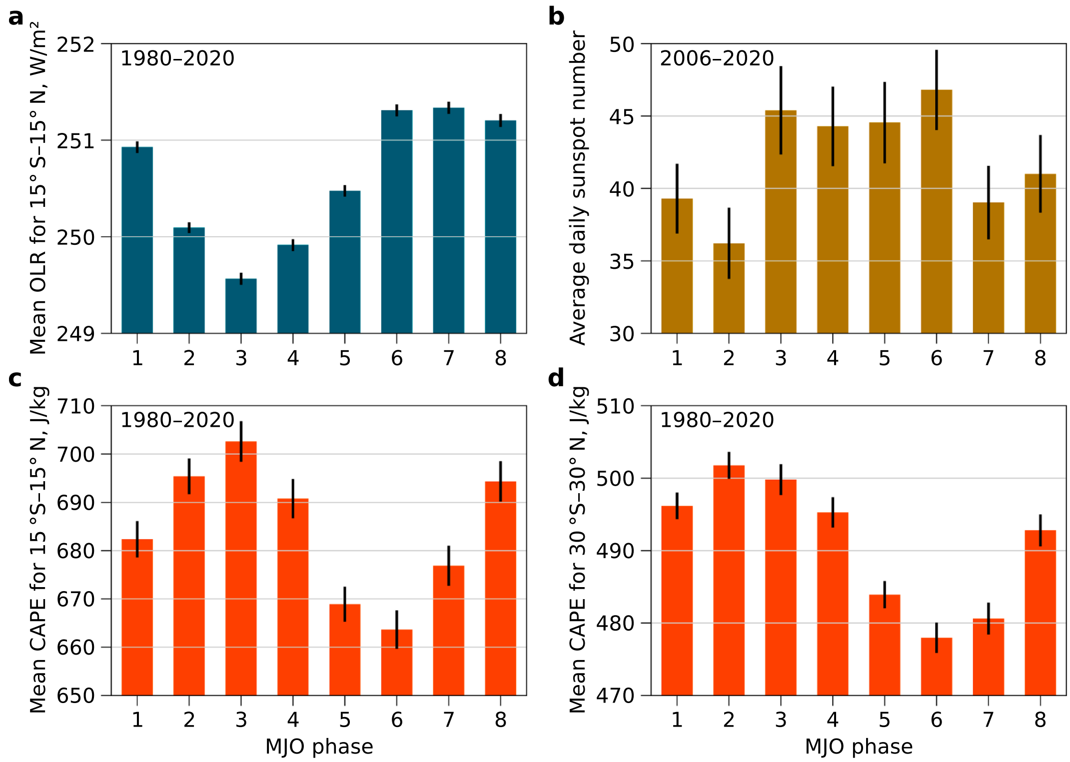

Fig. 8. (a): Average values of the OLR flux in the equatorial region (15° S–15° N) for each MJO phase (on the basis of the data for 1980–2020). (b): Values of the daily sunspot number for each MJO phase, averaged over the same days that were used to plot the potential gradient variation in Fig. 3c and d. (c): Average values of the simulated CAPE in the equatorial region (15° S–15° N) for each MJO phase; more precisely, maximum CAPE values in different grid columns are averaged over the equatorial region (on the basis of the data for 1980–2020). (d): The same as (c) for a wider region (30° S–30° N). Error bars show the values plus and minus one standard error.

important only during Northern Hemisphere summers.

5.3.2. Possible influence of solar activity

Another possible reason for the observed disagreement between the IP and PG is related to various factors which influence PG measurements. One such factor which should be taken into account when one interprets the results of measurements in the Antarctic is solar activity (Burns et al., 2017). Solar activity influences atmospheric ionisation both directly (via solar cosmic rays) and indirectly (by modulating the flux of galactic cosmic rays). This affects the atmospheric conductivity on many timescales, from several hours (e.g., after solar flares) to many years (during 11-year solar cycles).

Conductivity perturbations due to solar activity influence both the IP and the PG. Since the IP is a global index uniquely determined by the distributions of the conductivity and source currents, it is affected by conductivity changes in the entire atmosphere. Regarding the PG, it follows from Ohm's law that this quantity at the Earth's surface in fair-weather regions (where the current is nearly vertical) can be expressed as

$$\frac{d\phi}{dz}(0) = \frac{V}{\sigma(0)}\left(\int \frac{dz}{\sigma(z)}\right)^{-1},$$

where V is the IP, z is the vertical coordinate (with $z = 0$ corresponding to the Earth's surface) and $\phi(z)$ and $\sigma(z)$ denote the local profiles of the potential and conductivity respectively. Therefore the PG is especially affected by the variability of the conductivity in the air column over the site where measurements are performed. This may be particularly important for Antarctic measurements in view of the greater ionisation and more pronounced influence of solar activity at high latitudes.

In this study we did not take into account conductivity perturbations due to solar activity when simulating the IP (Eq. (1) assumes a universal conductivity profile $\sigma_0\exp(z/H)$ all over the Earth). We did not try to eliminate the effects related to solar activity from the results of PG measurements at Vostok either. Therefore it is indeed natural to expect that this effect will be able to partially explain the disagreement between the variations of the IP and PG.

Since a typical timescale for MJO events is about 30–90 days, it is reasonable to suppose that the quasi-periodic 27-day solar variability, associated with the solar rotation period and related phenomena (e.g., corotating interaction regions), will be particularly important in our problem. Note that Füllekrug and Fraser-Smith (1996), who have identified a variation with a period of 20–30 days in Schumann resonance data, attributed this variation precisely to the solar rotation period. However, Anyamba et al. (2000) have shown later that the variability of the Schumann resonance intensity on this timescale may be related to the MJO as well.

Despite all these considerations, it is unlikely that 27-day solar variability could result in a systematic difference between the simulated IP and the PG at Vostok. Indeed, since MJO events occur irregularly, any periodic or quasi-periodic variability is smoothed out after averaging over MJO phases (provided that the data set covers a large number of MJO events). The same argument applies, e.g., to 11-year solar cycles and, more generally, justifies the neglect of solar activity in simulations.

As a further illustration, Fig. 8b shows how solar activity varied over the 'fair-weather' days at Vostok selected according to our criteria listed in Section 2.4; this figure presents the value of the daily sunspot number averaged over the selected days within each phase of the MJO. Although solar activity does not manifest any stable trend on the MJO scale (calculations show that other selections of 'fair-weather' days and time spans yield different results), the variation shown in Fig. 8b is still substantial and may have influenced the results of measurements. In particular, comparing Fig. 3c with Fig. 8b, we see that phase 2, where the measured PG attains its maximum, is characterised by the lowest solar activity among all phases (with our selection of 'fair-weather' days). However, if we try to correct PG values according to the variation of the sunspot number, we observe that a linear correction cannot yield a good agreement with the IP.

Thus solar variability must have influenced PG values at Vostok, but it is unlikely that it is the main reason for the observed discrepancy between the PG and the simulated IP. In addition, note that although equations of the GEC predict that the IP should decrease with increasing ionisation and conductivity (see, e.g., our formula (1)), there are some studies reporting the opposite (e.g., Markson, 1981). This suggests that cosmic radiation may influence not only the conductivity but also source currents inside electrified clouds (which we assume constant in our simulations). However, there presently does not exist a consistent theory describing such effects.

5.3.3. *Possible shortcomings of PG measurements*

Finally, it is possible that the observed phase difference between the variations of the IP and PG is the result of other local factors at Vostok. PG data are generally more variable than the results of simulations (compare Fig. 7a–d with Fig. 7e and f), and probably we need more fair-weather days to fully understand whether the discrepancy between the two variations is a systematic effect and not a coincidence. PG values may be influenced by many varying local factors; this usually makes the long-term variation of the diurnal mean PG more difficult to interpret than, say, the diurnal variation of the PG, where one can eliminate the day-to-day variability by passing to relative values. Further research is needed to completely understand all these issues.

6. Conclusions

Let us now summarise the main findings of this study.

Both numerical simulations of the IP dynamics with the WRF and PG measurements in the Antarctic imply that the MJO systematically influences the direct current GEC. On average, both the simulated IP and measured fair-weather PG exhibit sine-like variations with the MJO phase; these variations show a statistically significant correlation with appropriately chosen projections of the two-dimensional RMM index, which characterises the MJO.

A more detailed investigation involving an application of EOF analysis to simulated contributions to the IP shows that the dynamics of these contributions reflects the dynamics of deep convection typical for the MJO. The majority of the IP variability on the MJO scale is produced by a superposition of two basic oscillating patterns of convection, one of which is located over the Maritime Continent and its surrounding seas and the other is located over the Indian Ocean and part of South-East Asia. These oscillating patterns can be naturally associated with the two components of the RMM index.

Despite a small phase difference between the variations of the simulated IP and measured PG (which is probably related to the shortcomings of simulations, but may also have been influenced by solar variability and local effects at the site of measurements), the results obtained in this study clearly indicate that the variation of the GEC contains some patterns related to the MJO. Together with earlier results about links between the GEC and the ENSO, this means that the GEC is an important part of the Earth system, which reflects climate variability on different timescales.

CRediT authorship contribution statement

Alexander V. Kozlov: Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Nikolay N. Slyunyaev:** Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Nikolay V. Ilin:** Conceptualization, Methodology, Software, Writing - review & editing. **Fedor G. Sarafanov:** Methodology, Software, Writing – review & editing, Validation. **Alexander V. Frank-Kamenetsky:** Data curation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Links to the data and code can be found in the Acknowledgements section of the paper.

Acknowledgements

The authors thank Evgeny Mareev for a helpful discussion of the material presented in this paper. The authors also thank Earle Williams and an anonymous reviewer of this paper for valuable comments and recommendations. This work was supported by the University of Nizhny Novgorod development programme 'Priority-2030' within the strategic project 'Comfortable environment' and by a grant from the Government of the Russian Federation (contract No. 075–15-2019–1892). The authors are grateful to the Australian Antarctic Division and to Gary Burns, the chief investigator of the project 'Atmospheric electrical circuits, weather and climate research', for their contribution to the development of instruments and software for collecting the electric field data at Vostok station. Continuous registration and collection of the electric field data at Vostok station is carried out by the specialists of the Russian Antarctic Expedition within the framework of the project 'Integrated oceanological, climatic, glaciological, and geophysical studies of Antarctica and the Southern Ocean'. PG data from Vostok, source codes and scripts used to analyse the results and plot the images are available at https://git260.ipfran.ru/eee/mjo-gec-2022. The WRF-ARW model can be downloaded from https://www2.mmm.ucar.edu/wrf/users. Simulations with the WRF model used the ERA5 data (Hersbach et al., 2018), downloaded from the Copernicus Climate Change Service (C3S) Climate Data Store, so that their results contain modified C3S information 2021; neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains. A time series of RMM1 and RMM2 values by the Australian Bureau of Meteorology is available online at http://www.bom.gov.au/climate/mjo. The NOAA's OLR—Daily CDR (Climate Data Record): PSL Interpolated Version data set (Lee, 2014) is provided by the NOAA PSL, Boulder, Colorado, USA, and can be found at https://psl.noaa.gov/mddb2/showDataset.html?datasetID=37. The online sunspot number catalogue of the World Data Center SILSO, Royal Observatory of Belgium, Brussels, is available at https://www.sidc.be/silso/datafiles.

Appendix A. Definition and calculation of the EOFs and PCs for the IP

Here we provide more precise definitions of EOFs and PCs (for more details, see, e.g., Zhang and Moore, 2015, Ch. 6). We will write all equations for the specific data set of anomalies in latitudinally averaged contributions to the IP, which we consider in Section 4.2.

Let

$$(\mathbf{a}\cdot\mathbf{b}) = \sum_{j=1}^{360} a_j b_j \text{ for } \mathbf{a} = \sum_{j=1}^{360} a_j \mathbf{e}^{(j)},\ \mathbf{b} = \sum_{j=1}^{360} b_j \mathbf{e}^{(j)}$$

and

$$\|a\| = \sqrt{\sum_{j=1}^{360} a_j^2} \text{ for } \mathbf{a} = \sum_{j=1}^{360} a_j \mathbf{e}^{(j)}$$

be the standard inner product and the standard Euclidean norm in $\mathbb{R}^{360}$. The first EOF $\mathbf{f}^{(1)}$ is defined as a unit vector ($\|\mathbf{f}^{(1)}\| = 1$) which minimises

$$\epsilon_1 = \left\langle \left\|\mathbf{U}(d) - (\mathbf{U}(d)\cdot\mathbf{f}^{(1)})\mathbf{f}^{(1)}\right\|^2 \right\rangle.$$

Likewise, the second EOF $\mathbf{f}^{(2)}$ is defined as a unit vector orthogonal to $\mathbf{f}^{(1)}$ ($\|\mathbf{f}^{(2)}\| = 1, (\mathbf{f}^{(1)}\cdot\mathbf{f}^{(2)}) = 0$) which minimises

$$\epsilon_2 = \left\langle \left\|\mathbf{U}(d) - (\mathbf{U}(d)\cdot\mathbf{f}^{(1)})\mathbf{f}^{(1)} - (\mathbf{U}(d)\cdot\mathbf{f}^{(2)})\mathbf{f}^{(2)}\right\|^2 \right\rangle.$$

In a similar way one defines the third EOF $\mathbf{f}^{(3)}$, and so on. The PCs are just coordinates of $\mathbf{U}(d)$ with respect to the new basis, with the jth PC given by the formula

$$C_j(d) = (\mathbf{U}(d)\cdot\mathbf{f}^{(j)}). \tag{A.1}$$

Using the fact that $\langle U_j(d)\rangle = 0$ for all j, it is not hard to show that

$$\epsilon_1 = \left\langle \|\mathbf{U}(d)\|^2 \right\rangle - \left\langle (\mathbf{U}(d)\cdot\mathbf{f}^{(1)})^2 \right\rangle = \sum_{j=1}^{360} \text{Var } U_j(d) - \text{Var } C_1(d),$$

where Var stands for the estimate of the variance. Similarly,

$$\epsilon_2 = \left\langle \|\mathbf{U}(d)\|^2 \right\rangle - \left\langle (\mathbf{U}(d)\cdot\mathbf{f}^{(1)})^2 \right\rangle - \left\langle (\mathbf{U}(d)\cdot\mathbf{f}^{(2)})^2 \right\rangle = \sum_{j=1}^{360} \text{Var } U_j(d) - \text{Var } C_1(d) - \text{Var } C_2(d),$$

and so on. Thus the first EOF $\mathbf{f}^{(1)}$ maximises Var $C_1(d)$, the second EOF $\mathbf{f}^{(2)}$ is chosen among the remaining dimensions and maximises Var $C_2(d)$, etc.

The actual computation of EOFs is based on finding the eigenvalues and eigenvectors of the covariance matrix Σ with elements $\Sigma_{jk} = \text{Cov}(U_j(d), U_k(d))$. It is not hard to show that this matrix is symmetric and positive semi-definite (Zhang and Moore, 2015, Proposition 6.1), hence its eigenvalues are real and non-negative. It can be shown that the first EOF $\mathbf{f}^{(1)}$ is an eigenvector of Σ corresponding to the largest eigenvalue, the second EOF $\mathbf{f}^{(2)}$ is an eigenvector corresponding to the second largest eigenvalue, and so on (Zhang and Moore, 2015, Theorems 6.1 and 6.3). The PC time series are then inferred from Eq. (A.1). The explained variance of the jth EOF $\mathbf{f}^{(j)}$ is equal to Var $C_j(d)$; it can be determined as the jth eigenvalue of the covariance matrix Σ (Zhang and Moore, 2015, Proposition 6.2).

References

Adlerman, E.J., Williams, E.R., 1996. Seasonal variation of the global electrical circuit. J. Geophys. Res. 101, 29679–29688. https://doi.org/10.1029/96JD01547.

Anyamba, E., Williams, E., Susskind, J., Fraser-Smith, A., Fullekrug, M., 2000. The manifestation of the Madden–Julian oscillation in global deep convection and in the Schumann resonance intensity. J. Atmos. Sci. 57, 1029–1044. https://doi.org/10.1175/1520-0469(2000)057<1029:TMOTMJ>2.0.CO;2.

Beggan, C.D., Musur, M.A., 2019. Is the Madden–Julian Oscillation reliably detectable in Schumann Resonances? J. Atmos. Sol. Terr. Phys. 190, 108–116. https://doi.org/10.1016/j.jastp.2019.05.009.

Burns, G.B., Frank-Kamenetsky, A.V., Tinsley, B.A., French, W.J., Grigioni, P., Camporeale, G., Bering, E.A., 2017. Atmospheric global circuit variations from Vostok and Concordia electric field measurements. J. Atmos. Sci. 74, 783–800. https://doi.org/10.1175/JAS-D-16-0159.1.

Burns, G.B., Tinsley, B.A., Frank-Kamenetsky, A.V., Troshichev, O.A., French, W.J.R., Klekociuk, A.R., 2012. Monthly diurnal global atmospheric circuit estimates derived from Vostok electric field measurements adjusted for local meteorological and solar wind influences. J. Atmos. Sci. 69, 2061–2082. https://doi.org/10.1175/JAS-D-11-0212.1.

Füllekrug, M., Fraser-Smith, A.C., 1996. Further evidence for a global correlation of the Earth-ionosphere cavity resonances. Geophys. Res. Lett. 23, 2773–2776. https://doi.org/10.1029/96GL02612.

Hannachi, A., 2021. Patterns Identification and Data Mining in Weather and Climate. Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-030-67073-3.

Harrison, R.G., Joshi, M., Pascoe, K., 2011. Inferring convective responses to El Niño with atmospheric electricity measurements at Shetland. Environ. Res. Lett. 6, 044028 https://doi.org/10.1088/1748-9326/6/4/044028.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J.N., 2018. ERA5 hourly data on pressure levels from 1979 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). doi:10.24381/cds.bd0915c6. Updated daily.

Ilin, N.V., Slyunyaev, N.N., Mareev, E.A., 2020. Toward a realistic representation of global electric circuit generators in models of atmospheric dynamics. J. Geophys. Res. Atmos. 125, e2019JD032130 https://doi.org/10.1029/2019JD032130.

Kendall, M.G., Stuart, A., 1961. The Advanced Theory of Statistics. Vol. 2. Inference and Relationship. Charles Griffin & Co. Ltd., London.

Kessler, W.S., 2001. EOF representations of the Madden–Julian oscillation and its connection with ENSO. J. Climate 14, 3055–3061. https://doi.org/10.1175/1520-0442(2001)014<3055:EROTMJ>2.0.CO;2.

Knutson, T.R., Weickmann, K.M., 1987. 30–60 day atmospheric oscillations: Composite life cycles of convection and circulation anomalies. Mon. Wea. Rev. 115, 1407–1436. https://doi.org/10.1175/1520-0493(1987)115<1407:DAOCLC>2.0.CO;2.

Lau, W.K.M., Waliser, D.E. (Eds.), 2012. Intraseasonal Variability in the Atmosphere-Ocean Climate System, second ed. Springer, Berlin, Germany. https://doi.org/10.1007/978-3-642-13914-7.

Lavigne, T., Liu, C., Deierling, W., Mach, D., 2017. Relationship between the global electric circuit and electrified cloud parameters at diurnal, seasonal, and interannual timescales. J. Geophys. Res. Atmos. 122, 8525–8542. https://doi.org/10.1002/2016JD026442.

Lee, H.T., 2014. Climate Algorithm Theoretical Basis Document (C-ATBD): Outgoing Longwave Radiation (OLR)—Daily. NOAA's Climate Data Record (CDR) Program, CDRP-ATBD-0526. URL:https://www.ncei.noaa.gov/pub/data/sds/cdr/CDRs/Outgoing%20Longwave%20Radiation%20-%20Daily/AlgorithmDescription_01B-21.pdf.

Liu, C., Williams, E.R., Zipser, E.J., Burns, G., 2010. Diurnal variations of global thunderstorms and electrified shower clouds and their contribution to the global electrical circuit. J. Atmos. Sci. 67, 309–323. https://doi.org/10.1175/2009JAS3248.1.

Lo, F., Hendon, H.H., 2000. Empirical extended-range prediction of the Madden–Julian oscillation. Mon. Wea. Rev. 128, 2528–2543. https://doi.org/10.1175/1520-0493(2000)128<2528:EERPOT>2.0.CO;2.

Madden, R.A., Julian, P.R., 1971. Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. J. Atmos. Sci. 28, 702–708. https://doi.org/10.1175/1520-0469(1971)028<0702:DOADOI>2.0.CO;2.

Madden, R.A., Julian, P.R., 1972. Description of global-scale circulation cells in the tropics with a 40–50 day period. J. Atmos. Sci. 29, 1109–1123. https://doi.org/10.1175/1520-0469(1972)029<1109:DOGSCC>2.0.CO;2.

Madden, R.A., Julian, P.R., 1994. Observations of the 40–50-day tropical oscillation—A review. Mon. Wea. Rev. 122, 814–837. https://doi.org/10.1175/1520-0493(1994)122<0814:OOTDTO>2.0.CO;2.

Mareev, E.A., Volodin, E.M., 2014. Variation of the global electric circuit and ionospheric potential in a general circulation model. Geophys. Res. Lett. 41, 9009–9016. https://doi.org/10.1002/2014GL062352.

Markson, R., 1981. Modulation of the Earth's electric field by cosmic radiation. Nature 291, 304–308. https://doi.org/10.1038/291304a0.

Markson, R., 2007. The global circuit intensity: Its measurement and variation over the last 50 years. Bull. Amer. Meteor. Soc. 88, 223–241. https://doi.org/10.1175/BAMS-88-2-223.

Matthews, A.J., 2000. Propagation mechanisms for the Madden-Julian Oscillation. Quart. J. Roy. Meteor. Soc. 126, 2637–2651. https://doi.org/10.1002/qj.49712656902.

Powers, J.G., Klemp, J.B., Skamarock, W.C., Davis, C.A., Dudhia, J., Gill, D.O., Coen, J.L., Gochis, D.J., Ahmadov, R., Peckham, S.E., Grell, G.A., Michalakes, J., Trahan, S., Benjamin, S.G., Alexander, C.R., Dimego, G.J., Wang, W., Schwartz, C.S., Romine, G. S., Liu, Z., Snyder, C., Chen, F., Barlage, M.J., Yu, W., Duda, M.G., 2017. The Weather Research and Forecasting model: Overview, system efforts, and future directions. Bull. Amer. Meteor. Soc. 98, 1717–1737. https://doi.org/10.1175/BAMS-D-15-00308.1.

Rycroft, M.J., Harrison, R.G., Nicoll, K.A., Mareev, E.A., 2008. An overview of Earth's global electric circuit and atmospheric conductivity. Space Sci. Rev. 137, 83–105. https://doi.org/10.1007/s11214-008-9368-6.

Rycroft, M.J., Nicoll, K.A., Aplin, K.L., Harrison, R.G., 2012. Recent advances in global electric circuit coupling between the space environment and the troposphere. J. Atmos. Sol. Terr. Phys. 90–91, 198–211. https://doi.org/10.1016/j.jastp.2012.03.015.

Sátori, G., Williams, E., Lemperger, I., 2009. Variability of global lightning activity on the ENSO time scale. Atmos. Res. 91, 500–507. https://doi.org/10.1016/j.atmosres.2008.06.014.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Liu, Z., Berner, J., Wang, W., Powers, J.G., Duda, M.G., Barker, D.M., Huang, X.Y., 2019. A Description of the Advanced Research WRF Version 4. NCAR Technical Note NCAR/TN-556+STR. https://doi.org/10.5065/1dfh-6p97.

Slingo, J.M., Rowell, D.P., Sperber, K.R., Nortley, F., 1999. On the predictability of the interannual behaviour of the Madden–Julian Oscillation and its relationship with El Niño. Quart. J. Roy. Meteor. Soc. 125, 583–609. https://doi.org/10.1002/qj.49712555411.

Slyunyaev, N.N., Frank-Kamenetsky, A.V., Ilin, N.V., Sarafanov, F.G., Shatalina, M.V., Mareev, E.A., Price, C.G., 2021a. Electric field measurements in the Antarctic reveal patterns related to the El Niño—Southern Oscillation. Geophys. Res. Lett. 48, e2021GL095389 https://doi.org/10.1029/2021GL095389.

Slyunyaev, N.N., Ilin, N.V., Mareev, E.A., 2019. Modeling contributions of continents and oceans to the diurnal variation of the global electric circuit. Geophys. Res. Lett. 46, 5516–5525. https://doi.org/10.1029/2019GL083166.

Slyunyaev, N.N., Ilin, N.V., Mareev, E.A., Price, C.G., 2021b. The global electric circuit land–ocean response to the El Niño—Southern Oscillation. Atmos. Res. 260, 105626. https://doi.org/10.1016/j.atmosres.2021.105626.

Slyunyaev, N.N., Ilin, N.V., Mareev, E.A., Price, C.G., 2021c. A new link between El Niño—Southern Oscillation and atmospheric electricity. Environ. Res. Lett. 16, 044025 https://doi.org/10.1088/1748-9326/abe908.

Stolz, D.C., Rutledge, S.A., Xu, W., Pierce, J.R., 2017. Interactions between the MJO, aerosols, and convection over the central Indian Ocean. J. Atmos. Sci. 74, 353–374. https://doi.org/10.1175/JAS-D-16-0054.1.

Tacza, J., Nicoll, K.A., Macotela, E., 2022. Periodicities in fair weather potential gradient data from multiple stations at different latitudes. Atmos. Res. 276, 106250 https://doi.org/10.1016/j.atmosres.2022.106250.

Wheeler, M.C., Hendon, H.H., 2004. An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. Mon. Wea. Rev. 132, 1917–1932. https://doi.org/10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2.

Williams, E., Bozóki, T., Sátori, G., Price, C., Steinbach, P., Guha, A., Liu, Y., Beggan, C. D., Neska, M., Boldi, R., Atkinson, M., 2021. Evolution of global lightning in the transition from cold to warm phase preceding two super El Niño events. J. Geophys. Res. Atmos. 126, e2020JD033526 https://doi.org/10.1029/2020JD033526.

Williams, E., Mareev, E., 2014. Recent progress on the global electrical circuit. Atmos. Res. 135–136, 208–227. https://doi.org/10.1016/j.atmosres.2013.05.015.

Williams, E.R., 2009. The global electrical circuit: A review. Atmos. Res. 91, 140–152. https://doi.org/10.1016/j.atmosres.2008.05.018.

Zhang, C., 2005. Madden-Julian Oscillation. Rev. Geophys. 43, RG2003 https://doi.org/10.1029/2004RG000158.

Zhang, C., Adames, Á.F., Khouider, B., Wang, B., Yang, D., 2020. Four theories of the Madden-Julian Oscillation. Rev. Geophys. 58, e2019RG000685 https://doi.org/10.1029/2019RG000685.

Zhang, Z., Moore, J.C., 2015. Mathematical and Physical Fundamentals of Climate Change. Elsevier, Amsterdam, Netherlands. https://doi.org/10.1016/C2013-0-14403-0.