

Correlation between Global Lightning and Whistlers observed at Tihany, Hungary

A. B. Collier,^{1,2} B. Delpont,² A. R. W. Hughes,² J. Lichtenberger,³ P. Steinbach,^{3,4}

J. Öster^{2,5,6} and C. J. Rodger⁷

Hermanus Magnetic Observatory, Hermanus, 7200, South Africa.

School of Physics, University of KwaZulu-Natal, Durban, 4041, South Africa.

5 Space Research Group, Eötvös University, Budapest, Pf 32 H-1518, Hungary.

Research Group for Geology, Geophysics and Space Sciences, MTA-ELTE,
Eötvös University, Budapest, Pf 32 H-1518, Hungary.

Space and Plasma Physics, School of Electrical Engineering, Royal Institute of
Technology, SE-100 44 Stockholm, Sweden.

10 Department of Astronomy and Space Physics, Uppsala University, SE-751 20
Uppsala, Sweden.

Department of Physics, University of Otago, Dunedin 9054, New Zealand.

Abstract. Although the generation and propagation mechanisms for whistlers are fairly well understood, the location and extent of the lightning source region for the whistlers observed at a given station are currently unknown.

The correlation of whistler observations against global lightning data allows an estimate of the size and position of the source region. For whistlers detected at Tihany, Hungary, an area of positive correlation with radius ~ 1000 km was found to be centred on the conjugate point. Although the maximal sample correlation coefficient was relatively low, $r = 0.065$, it has a high statistical significance, indicating that it is extremely improbable that the whistlers and lightning in this region are actually uncorrelated. Other smaller areas of positive correlation were found further afield in South America and the Maritime Continent. Lightning in the northern hemisphere displayed a negative correlation with whistlers at Tihany.

1. Introduction

1.1. Whistler Mechanism

Whistlers are dispersed Very Low Frequency (VLF) emissions observed both on the ground and in space. The mechanism expounded in the pioneering work of *Storey* [1953] is generally acknowledged to account for their production. In this classical model, a whistler is initiated by a lightning stroke, which produces an intense pulse of electromagnetic radiation with broad spectral content, brief duration and peak power ~ 20 GW, an appreciable fraction of which lies in the VLF range. Within the Earth-ionosphere waveguide this pulse, or spheric, is not significantly dispersed and can travel considerable distances with low attenuation ~ 1 dB/Mm [*Bernstein et al.*, 1974]. However, some portion of the initial energy may penetrate upwards through the ionosphere and enter the magnetosphere. In the inhomogeneous and anisotropic magnetospheric plasma the waves propagate along magnetic field lines in the whistler mode. Since the whistler mode is dispersive, the propagation delay varies with frequency, transforming the initial pulse into a complex tone with a unique frequency-time signature determined by the magnetic field strength and plasma density along the path traversed through the magnetosphere. An example of a whistler spectrogram is presented in Figure 1.

When a whistler is received on the ground it has completed a journey with at least three major components: (i) sub-ionospheric propagation from the source to the footpoint of the duct; (ii) field-aligned propagation to the opposite hemisphere; and (iii) sub-ionospheric propagation from the ionospheric duct exit point at the conjugate footpoint to the receiver.

The morphology of whistlers varies appreciably with latitude. Whereas at $L \lesssim 2.4$ whistlers are generally observed as isolated traces, corresponding to propagation along a single path through the magnetosphere, at higher latitudes whistlers are commonly multi-path, where a sin-

gle lightning stroke produces a signal which propagates along multiple magnetospheric paths. The whistlers considered in this study conform to the former category.

1.2. Ducting

Due to anisotropy in the whistler mode dispersion relation, the passage of the waves through
 50 the magnetosphere is roughly aligned with the magnetic field. If, in addition, the waves are trapped within a field-aligned plasma density irregularity, or duct, the wave normal direction is more strongly confined to the magnetic field. While in the duct the signal should also be amplified by wave-particle interactions [*Brice*, 1960; *Liemohn*, 1967; *Pasmanik et al.*, 2002]. There may also be unducted portions of the propagation path through the ionosphere and mag-
 55 netosphere. It is thought that all whistlers observed on the ground are ducted.

Because the refractive index of the neutral atmosphere, $\mu \sim 1$, is much lower than that in the plasma medium, $\mu \gg 1$, waves entering the ionosphere from below are refracted into a transmission cone defined by wave normal directions which lie close to the vertical. In order for a whistler to become trapped in a duct, the refracted wave normal must lie within the trap-
 60 ping cone, which is symmetric around the magnetic field direction [*Helliwell*, 1965]. At low latitudes there is little or no overlap between the transmission and trapping cones. As a consequence, whistlers are most commonly observed at middle to high geomagnetic latitudes, where the magnetic field has a large inclination, the field lines are approximately vertical and the transmission and trapping cones intersect. The intensity of the magnetic field regulates the width of
 65 the transmission cone and therefore also has an influence on the degree of overlap. Closer to the poles, where the magnetic field strength at ionospheric altitudes is higher, the whistler mode refractive index in the ionosphere is lower and the transmission cone is broader. Regions of depleted magnetic field strength are thus not favourable locations for the footpoint of a duct.

Only those waves incident upon the ionosphere from above with small wave normal angle
 70 with respect to the vertical are able to refract from the plasma medium to the neutral atmosphere.
 All other waves are reflected back into the magnetosphere. Thus, the most favourable conditions
 for the whistlers to penetrate to the ground also occur at higher latitudes.

Ducted whistlers retain the smallest wave normal angles and it is thus generally believed
 that almost all whistlers which enter into the waveguide at the conjugate point are ducted. In
 75 the absence of a duct the waves are magnetospherically reflected and may undergo numerous
 reflections, possibly forming a uniform band of wave energy [Sonwalkar and Inan, 1989]. Since
 whistler mode waves interact with the ambient magnetospheric plasma, causing electrons to be
 scattered into the loss cone, these magnetospherically reflected whistlers play a significant role
 in regulating the population of the radiation belts [Lauben *et al.*, 1999; Johnson *et al.*, 1999;
 80 Rodger *et al.*, 2004].

Access to a duct need not necessarily occur at its base, as it is feasible for a whistler to be
 trapped after leaking in through the side of the duct [Strangeways and Rycroft, 1980]. The
 termination altitude of the duct strongly influences the proportion of wave energy which is
 trapped through either its side or base, where the side is favoured for ducts extending to lower
 85 altitudes [Strangeways, 1981]. Similarly, waves may leak out of a duct when their wavelength
 becomes comparable to the width of the duct, and those which exit on the lower latitude side of
 the duct may still have near-perpendicular incidence on the ionosphere [Strangeways, 1986].

Although the majority of lightning is confined to the tropics and sub-tropics [Christian *et al.*,
 2003], these regions do not play a major role in generating whistlers. Whistler observations near
 90 the geomagnetic equator are rare and thought to arise due to propagation in the waveguide from
 an ionospheric exit point at higher latitudes [Storey, 1955; Rao *et al.*, 1974; Helliwell, 1965].

There is evidence to suggest that a low-latitude cutoff exists for whistlers at around 16° geomagnetic latitude, which may be due to the paucity of ducts at low latitudes [Rao *et al.*, 1974], but might also be due to matching conditions at the ionospheric boundary. Thorne and Horne
 95 [1994] concluded that VLF signals launched at an invariant latitude below 15° (equivalent to $L \sim 1.1$) remained trapped in the ionosphere.

1.3. Source Region

The model outlined above is broadly accepted and little subsequent work has been done to verify it in the light of improved equipment and analysis techniques. However, despite the relatively long history of investigations into the whistler phenomenon, some of the details of the
 100 chain between the initial lightning discharge and the reception of its dispersed electromagnetic signature are still unclear. The efficiency of transionospheric leakage and transformation to the whistler mode, known to depend on ionospheric conditions and magnetic field inclination, have been examined and it has been established that the majority of lightning strokes generate an upgoing, incipient whistler detectable on Low Earth Orbit (LEO) satellites [for example
 105 Hughes, 1981; Li *et al.*, 1991; Hughes and Rice, 1997; Holzworth *et al.*, 1999]. These whistlers are predominantly unducted and the proportion which are ducted to the conjugate hemisphere is currently unknown. There is also rather scarce and inconclusive experimental verification of the typical characteristic of the lightning source region.

The size and location of the effective source region is a source of contention. It is well known
 110 that sferics can travel enormous distances in the Earth-ionosphere waveguide with minimal attenuation. Whistlers have been associated with lightning strokes occurring more than 2000 km from the duct footprint [Weidman *et al.*, 1983; Carpenter and Orville, 1989; Li *et al.*, 1991; Clilverd *et al.*, 1992; Holzworth *et al.*, 1999]. The lightning source region for whistlers detected

at a given location may thus be rather large. Furthermore, it is assumed that this source region
 115 is centred on the magnetic conjugate point. Yet this assumption has not been validated. *Yoshino*
 [1976] observed that the majority of whistlers observed in Sugadaira, Japan, occurred when
 there was thick cloud cover within ~ 500 km southwest of the conjugate point. In contrast,
 more recently *Chum et al.* [2006], in an analysis based on isolated pairs of lightning strokes and
 fractional hop whistlers detected on LEO satellites, found that the point at which the sferic pulse
 120 penetrates the ionosphere is $\lesssim 1500$ km from discharge. Their data included both ducted and
 non-ducted whistlers, although it is probable that the majority of the whistlers were unducted
 [*Hughes and Rice*, 1997]. *Collier et al.* [2006] employed a source region with radius 600 km
 (selected for rather pragmatic reasons: their lightning data did not extend more than ~ 600 km
 south of the conjugate point) in a case study which suggested that whistlers at Tihany, Hungary,
 125 are more likely to arise from strokes to the southeast of the conjugate point. Both *Yoshino*
 [1976] and *Collier et al.* [2006] found that the most effective source region was displaced from
 the conjugate point towards the magnetic pole.

Since propagation in the waveguide may occur both before and after the signal's passage
 through the magnetosphere, it is in principle possible for both the causative discharge and the
 130 VLF receiver to be significantly distant from the footpoints of the guiding magnetic field line.
 Indeed the path through the magnetosphere need not have a footpoint at either the source or the
 receiver, but may be displaced with respect to both [*Clilverd et al.*, 1992]. *Allcock and McNeill*
 [1966] found that transmission loss was minimised for paths aligned along either the magnetic
 meridian of the source or the receiver. This stands to reason since these configurations result
 135 in the minimum loss of power due to the cylindrically symmetric expansion of the wavefront
 in the waveguide. A path at the magnetic longitude of the source, however, appears to be the

most favourable [*Shimakura et al.*, 1987; *Ladwig and Hughes*, 1989]. Furthermore, based on the overlap of transmission and trapping cones, strokes at L higher than the duct footpoint are most likely to produce whistlers [*Helliwell*, 1965, Fig. 3–23].

140 Since the power of a signal radiating symmetrically in the Earth-ionosphere waveguide is inversely proportional to the distance from the source, it is not unreasonable to suppose that the likelihood of duct excitation by a given lightning discharge is highest in close proximity to the stroke. Consequently one would suppose that a region of limited extent around the conjugate point represents the area in which a whistler's causative lightning stroke is most likely
145 to have been located. Certainly this is a reasonable assumption if there is an appreciable level of lightning activity in the vicinity of the conjugate point. For regions of regular thunderstorm activity, the number of lightning strokes within a reasonably large area around the footpoint of a field line far exceeds the number of whistlers recorded near the conjugate point, indicating that lightning is a necessary but not sufficient condition for the generation of a ducted whistler.
150 If, however, the conjugate point is located in a region of infrequent lightning then it is probable that a realistic source region should have considerably greater range. Proximity is thus not the only consideration and there must be other factors involved: the total power radiated and the discharge type (cloud-to-cloud (CC) or cloud-to-ground (CG)) may also play a role.

In some instances it is possible to identify the causative sferic for 2-hop whistlers [*Carpenter*,
155 1959; *Carpenter and Orville*, 1989]. The location of the causative sferic is traditionally consistent with propagation times calculated using the DE–1 diffusive equilibrium model [*Park*, 1972] for the distribution of electron density along the magnetic field lines and the simplified Appleton-Hartree dispersion relation. It has, however, been demonstrated, using a more realistic electron density distribution and an improved dispersion relation, that the calculated nose time

may differ by up to 500 ms [*Lichtenberger et al.*, 2008b]. There is also uncertainty regarding the distance that the signal propagated in the waveguide before and after passing through the magnetosphere. Furthermore, the global flash rate of $44 \pm 5 \text{ s}^{-1}$ [*Christian et al.*, 2003] gives a typical interval of ~ 20 ms between sferics, leading to significant ambiguity in the selection of the correct sferic amongst a number of potential candidates. In addition, due to sub-ionospheric propagation conditions, the actual causative sferic may or may not be visible on a spectrogram. It is thus possible that the causative sferic identified using traditional techniques may not be valid, but simply occurs with serendipitous timing. In either case, in the absence of further information it is almost impossible to uniquely determine the location of the discharge associated with a given whistler.

This paper is an attempt to obtain a better understanding of the positions of causative lightning strokes. The application of whistlers as a remote-sensing tool [for example, *Carpenter*, 1963] would be greatly enhanced by more accurate knowledge of the initial wave source and the parameters of the path taken through the magnetosphere.

2. Data

The analysis presented here is a statistical comparison of whistler data from a single terrestrial VLF receiver and global lightning data. The lightning and whistler data are both time sequences. The distinction between a time series and a time sequence is that, whereas the former consists of measurements of a given parameter at various times, the latter is a sequence of times at which a particular event or phenomenon occurred. The whistler data for Tihany consist of a sequence indicating the observation time of each whistler. The World Wide Lightning Location Network (WWLLN) data, in addition to the epoch of each lightning stroke identified, also reflect the corresponding geographical location.

2.1. Whistler Data

The VLF data is from Tihany, Hungary (46.89° N 17.89° E, $L = 1.80$), where the local time is UTC + 1 h. The point magnetically conjugate to Tihany is located at 33.08° S 28.08° E, just off the southeast coast of South Africa. The magnetic field at the conjugate point has declination and inclination of -26.75° and -64.72° respectively. The horizontal component is thus directed towards the northwest.

The whistler times were acquired using the Automatic Whistler Detector (AWD) developed at Eötvös University [*Lichtenberger et al.*, 2008a]. The AWD algorithm is based on two-dimensional image correlation: dynamic spectra are computed for 16 s overlapping data windows and correlated with an idealised template whistler. The latter is derived from a model which utilises *Bernard's* [1973] approximation and covers the dispersion range 20 to $80 \text{ s}^{1/2}$. The choice of this range is based on the average dispersion of $\sim 45 \text{ s}^{1/2}$ for whistlers recorded at Tihany. It is also possible that two-hop whistlers may have a dispersion which falls within the upper portion of this range, however, the seasonal distribution of the whistlers identified by the system indicates a dearth of whistlers during the Boreal summer, which suggests that multi-hop whistlers are only seldom, or perhaps never, detected. The detector identifies the whistler epoch as the time at which the whistler traces passes through $\sim 7 \text{ kHz}$. Prior to 10 June 2002 timing information was derived from the DCF radio time signal originating in Germany. Since the AWD receiver is near the limits of the range of this network, the system time during this interval was not particularly reliable. Global Positioning System (GPS) timing was subsequently employed. Although the detection system has been operational for a number of years, it is still considered to be in a experimental state, and this should be taken into account when considering the data. The system has false-positive and false-negative rates of 20% to 50% and 10% respectively

[*Lichtenberger et al.*, 2008a], meaning that it is unlikely to miss a whistler if there is indeed one
 205 present in the data, but more often mistakenly indicates a whistler when there is not actually
 one present in the data.

Between 27 February 2002 and 18 May 2005, 681107 whistlers were identified at Tihany, an
 average of 578 per day. A clear seasonal variation is apparent with a daily average of 1371 in
 summer and 99 in winter. The diurnal distribution of whistler activity is plotted in Figure 2. The
 210 majority of whistlers are observed between dusk and dawn. This pattern persists throughout the
 year but is most apparent during the Austral summer months when whistlers are more preva-
 lent. Naturally, lightning and propagation conditions must both play a role in the incidence
 of whistlers. Thunderstorm activity around the conjugate point has a diurnal peak in the late
 afternoon and a seasonal maximum during the Austral summer [*Collier et al.*, 2006].

215 The nocturnal nature of whistlers is well known and may be attributed to the elevated trans-
 parency of the ionosphere at night, while during the day it is effectively opaque to VLF due to
 absorption in the D region [*Helliwell*, 1965]. From satellite measurements it is known that sig-
 nals from VLF transmitters are 10 dB stronger at night than during the day. This is consistently
 observed in the enhanced coupling of VLF transmitter signals into the plasmasphere on the
 220 night side [*Green et al.*, 2005; *Gamble et al.*, 2008]. The attenuation of a plane wave normally
 incident on the ionosphere increases with frequency [*Helliwell*, 1965, Fig. 3–31] but decreases
 with geomagnetic latitude and, at mid-latitudes is ~ 10 dB lower at night than during the day for
 frequencies around 2 kHz [*Helliwell*, 1965, Fig. 3–35]. For whistlers observed on the ground
 this implies an attenuation which is 20 dB greater during the day, making them undetectable by
 225 many receivers.

Figure 3 displays the distribution of L obtained by applying the fitting procedure described by *Tarcsai* [1975] to whistlers observed at Tihany. Of the 148724 whistler traces to which the procedure was applied, only the results for 125066 were retained since the remainder were poor fits. The range of dispersion was from 14 to $181 \text{ s}^{1/2}$, with an average of $\langle D \rangle = 53 \text{ s}^{1/2}$. The resulting range of L was 1.75 to 3.42, with a mean of 2.33. These results may be contrasted with those of *Tarcsai et al.* [1988] who, using 985 whistlers detected at Tihany between December 1970 and May 1975, found that L extended from 1.4 to 3.2. It is apparent that most of the whistlers observed at Tihany travel along ducts at L higher than that of the station. This marked asymmetry may be indicative of a low latitude cutoff [*Rao et al.*, 1974].

2.2. Lightning Data

At any instant up to 2000 thunderstorms are active over the surface of the Earth. Most of this activity is concentrated in the tropics over the land masses in South America, Africa and the Maritime Continent. There is a well defined diurnal variation in global lightning activity which is dictated by the local diurnal cycle over each of these regions [*Price*, 1993].

The acquisition of global lightning statistics has been facilitated by improved satellite data and terrestrial networks with more extensive coverage. The WWLLN operates in the VLF range, where attenuation is low, and is thus able to detect global lightning activity with only a limited number of receivers [*Dowden et al.*, 2002, 2008; *Lay et al.*, 2004]. The number of receivers in the network has increased from 11 in March 2003 to more than 30 as of June 2008. This has been accompanied by a commensurate increase in the number of lightning strokes reported by the network [*Rodger et al.*, 2008]. While WWLLN does not group strokes together into flashes, it primarily responds to lightning with large currents [*Rodger et al.*, 2006], and as such will typically report the first stroke in a multi-stroke flash, as the subsequent strokes generally

have peak currents which are somewhat less than half that of the first stroke [Uman, 1987, pp. 123–125]. Each stroke is assigned a time and location, where the temporal and spatial accuracy
 250 are $\sim 30 \mu\text{s}$ and $\sim 10 \text{ km}$ respectively [Rodger *et al.*, 2005].

WWLLN records both CG and CC strokes, but does not discriminate between the two classes. This distinction may be significant for the generation of whistlers. The network also does not measure the current moment or determine the polarity of detected strokes. WWLLN seldom identifies strokes with peak current less than 25 kA [Dowden *et al.*, 2008]. Lay *et al.* [2004],
 255 using data from a local lightning detection network in Brazil, found the mean peak current for WWLLN events was between 70 and 80 kA, which is appreciably above the average for global lightning activity. The bias towards lightning strokes carrying larger currents arises because a given event is only registered by the network if it is detected by at least five sites. The data therefore include a disproportionate number of intense lightning events.

260 Comparison with other lightning detection systems has established that WWLLN still provides an accurate and representative reflection of global lightning activity. The global lightning distribution obtained from an analysis of WWLLN data [Rodger *et al.*, 2006] is consistent with that obtained by Christian *et al.* [2003]. The high threshold current results in a global detection efficiency of only around 2% for all strokes and 5% for CG strokes [Rodger *et al.*, 2005, 2006;
 265 Jacobson *et al.*, 2006]. Due to the low sensitivity of WWLLN, the absolute rates are, depending on the longitude sector, roughly an order of magnitude too low. The detection efficiency of WWLLN varies both spatially and temporally both as a result of the non-uniform distribution of receivers and the fact that not all of the receivers are operational at any given time. There is also a spurious enhancement over the Maritime Continent, which arises due to the high density
 270 of WWLLN nodes in this region [Rodger *et al.*, 2006].

Since its inception both the triggering techniques used to identify sferics at each of the receivers and the algorithm used to assimilate the sferic data have been improved several times. Prior to August 2003, WWLLN used the trigger time rather than the Time of Group Arrival (TOGA) [Dowden *et al.*, 2002], with the consequence that locations could be in error by more than 100 km [R. H. Holzworth, private communication]. However, after this date locations are typically accurate to between 10 and 20 km [Rodger *et al.*, 2005]. The raw TOGA data, which have been archived since August 2004, were recently subjected to an updated algorithm which has improved the accuracy of the lightning stroke locations and removed some artefacts from the data.

2.3. Joint Data Availability

Although neither the whistler nor lightning detection systems are infallible, for the purpose of this study all events registered by either system are regarded as equally authentic. It is, however, only meaningful to make comparisons between these data sets for periods during which both detection systems were functioning properly. To this end, intervals when either system was not operating were identified.

Figure 4 reflects the status of the AWD for the period from 24 February 2002 until 18 May 2005. The colour assigned to a particular day indicates whether the system was recording data and whether or not any whistlers were detected during the day. For the first few months the results from the detector were erratic and there were numerous days on which no whistlers were identified. However, after this initial period the system stabilised and has since operated reliably. No data were recorded during the period from the end of August 2004 until the beginning of November 2004 because of the failure of the antenna pre-amplifier caused by a nearby lightning stroke.

WWLLN operates on data from numerous stations distributed around the globe. At any given time a subset of these stations may not be operational. Although missing nodes do not render the network inoperable, the reliability of the data is compromised. Since the conjugate point for Tihany is located close to the coast of South Africa, it is anticipated that the majority of the whistlers detected at Tihany should arise from lightning strokes located somewhere in this vicinity. Although the operation of the two stations in South Africa (at Durban and Hermanus) is critical to the coverage of the network over Africa, the relative detection efficiency of a given station actually declines for strokes occurring within a few hundred km of that station [Rodger *et al.*, 2006, Fig. 7], so that observations from the rest of the network contribute most to locating strokes close to the conjugate point. Although regular records are maintained of the operation of the stations in the network, it is difficult to use this information to resolve the absolute efficiency of the network for a particular region of the globe. In addition, empirical identification of those days which should be excluded on the basis of poor lightning data is problematic. One approach might be to exclude periods during which WWLLN did not record strokes over South Africa for two or more successive days. Certainly there are numerous periods during which this criterion applies. However, such a solution is somewhat arbitrary and exclusion on this basis might bias the results in favour of locations around South Africa, so it is not invoked.

In light of the fact that the operation of the WWLLN is hard to quantify, it was assumed to be functional (although not necessarily at optimal efficiency) throughout the period of the analysis. However, data was only retained for days on which the Tihany logs reflect that the AWD was in operation. On this basis, data for 96 days were excluded.

3. Results

The efficacy of a given region of the globe in producing whistlers is assessed by performing
 315 a correlation between whistler incidence at Tihany and lightning occurrence within that region.
 The objective of this analysis is to identify those portions of the globe for which this correlation
 is high.

3.1. Correlation

The lightning data was first projected onto a 3° by 3° spatial grid. It is possible that the limited
 location accuracy of WWLLN may influence the statistical analysis of the lightning-whistler
 320 correlations, but the use of a relatively coarse grid ensures that the effects of spatial uncertainty
 are minimised. The whistler and lightning data sets were then prepared for the correlation
 analysis by dividing the time period between 1 January 2003 and 19 May 2005 into intervals of
 length $\Delta t = 1$ min. The number of events during each interval was then determined. Finally,
 the event counts were reduced to Boolean values, simply indicating the presence or absence of
 325 activity, but not its absolute intensity.

The latitudinal variation in the area of the cells in the lightning grid was not taken into account:
 cells at higher latitudes, which represent smaller areas, are treated in the same way as equatorial
 cells. For example, the cells at the conjugate latitude are 16% smaller than those at the equator.
 This implies that, all else being equal, there would be an increase in lightning counts in cells
 330 closer to the equator and a reduced count for higher latitude cells. However, the fact that this
 analysis is based on the presence or absence of lightning, rather than the absolute lightning
 count, implies that this does not have a significant influence on the results and the small bias is
 largely negated by the Boolean treatment of the data.

The selection of an appropriate value for Δt was based on a few considerations: (i) the delay
 335 between the causative lightning stroke and the reception of a whistler at Tihany is typically ~ 1 s;
 and (ii) computational efficiency (reducing Δt results in more time consuming calculations).
 The choice of Δt should be sufficiently long to minimise the risk of a lightning stroke and the
 associated whistler being allocated to different time intervals. Yet, Δt should be sufficiently
 short to reduce the risk of chance coincidence. The hazard of a temporal mismatch might
 340 be reduced by applying an offset, equal to the typical whistler propagation time, to either the
 whistler or lightning time sequence. However, the fact that Δt is substantially longer than the
 typical whistler propagation time suggests that the probability of a lightning stroke and the
 resulting whistler being assigned to separate intervals is very small indeed, so that there would
 be negligible gain from such a transformation.

345 To illustrate the typical form of the data, a contingency table for the cell centred at 34.5° S
 28.5° E, which contains the conjugate point, is presented in Table 1. It is apparent that there
 is appreciable asymmetry between each of the marginal totals. Applying the Fisher exact test
 for count data (appropriate for an unbalanced contingency table) to the data in Table 1 yields
 a vanishingly small p -value, indicating that the null hypothesis (odds ratio equal to unity) can
 350 be readily rejected: there is a significant relationship between lightning and whistler activity,
 although in the overwhelming majority of intervals there was an absence of either phenomenon.

Locations which are more likely to be the source of causative lightning strokes are those
 for which one finds a lightning stroke during the same interval as a whistler or an absence of
 lightning when no whistlers were observed. Conversely, locations which have lightning when
 355 there are no whistlers, or vice versa, are probably not within the source region. There are
 thus four possible cases: a positive result if there is both lightning and whistler activity or

neither; a negative result if there is either lightning or whistler activity but not both. This logic is equivalent to the condition

$$(W \cap S) \cup (\overline{W} \cap \overline{S}) \equiv \overline{W \oplus S} \quad (1)$$

where W and S are Boolean sequences for whistlers and lightning respectively, the length of each sequence corresponding to the number of time intervals of length Δt spanning the period of the analysis, and \oplus denotes the exclusive-OR operator. It might be supposed that $W \cap S$ would be a sufficient indicator, effectively counting intervals in which there were both whistlers and lightning, however regions of perennially intense lightning activity are continuously active and therefore introduce an unfair bias towards the tropics. The inclusion of $\overline{W} \cap \overline{S}$ restores the balance by accounting for those intervals in which there were neither whistlers nor lightning. Whereas $W \cap S$ emphasise tropical regions of high lightning activity, $\overline{W} \cap \overline{S}$ favours higher latitudes where lightning is scarce.

Figure 5 illustrates the above scheme, comparing the number of observed whistlers to lightning strokes within 600 km of the conjugate point for a 2 h period on 26 February 2004. The histograms reflect counts accumulated in 1 min bins, indicating both total counts (empty bars) as well as a Boolean count (shaded bars). Below the histograms are rug plots indicating the precise epoch of each of the whistler or lightning events. Of the 120 intervals, 48% had whistlers and 78% had lightning. Intervals for which condition (1) was satisfied are indicated by blue diamonds. The condition is true for 56% of the intervals.

It is interesting to note that although the frequency of lightning strokes within the chosen region is roughly constant (a few strokes per 1 min interval) for the full period illustrated in Figure 5, there are sparse whistlers observed during the first hour but many during the second

hour. This effect is due to the passage of the day-night terminator, where the latter portion of the data corresponds to night conditions.

380 Although conceptually revealing, an analysis based on (1) was problematic because no means was readily available to evaluate uncertainties. The relationship between whistler and lighting activity was thus assessed using the Pearson product-moment correlation coefficient, ρ , being the ratio of the covariance of the two sequences to the product of their respective standard deviations. The value of ρ is an indication of the extent to which variations in two variables
 385 occur in concert. The correlation coefficient is defined only if each of the standard deviations is both finite and non-zero. The sample correlation coefficient, r , is an unbiased estimator of the population correlation coefficient, ρ . The value of r lies in the range $[-1, +1]$, where the limits correspond to strong correlation (+1) or anti-correlation (-1). If the variables are independent then the correlation is zero. A vanishingly small value for r does not, however, necessarily imply
 390 that two quantities are unrelated. Two standardised variables are uncorrelated if the expected value of their product equals the product of their expected values. They are independent if their joint probability distribution function is the product of their individual probability distribution functions.

In general the correlation between two sequences may be calculated at various lags, where
 395 the sequences are shifted relative to each other. However, since this analysis is concerned with the corresponding time intervals in the two sequences, only zero lag is considered. If, however, the interval Δt were made so short that whistlers and the causative lightning strokes frequently occurred in different intervals, then non-zero lags would have to be accounted for.

The correlation between the Boolean sequences illustrated in Figure 5 is 0.18. This statistic,
 400 however, applies to only a limited time interval and an extended region around the conjugate

point. Figure 6 displays the correlation between the whistler sequence at Tihany and the lightning sequence in each of the grid cells spanning the globe. Data are plotted only for cells in which the correlation is significant at the $\alpha = 0.01$ level. It is apparent that there are regions of non-vanishing correlation, both positive and negative, but that the majority of the cells have values close to zero.

4. Discussion

A preliminary examination of the data in Figure 6 reveals that the correlation is:

1. generally positive (negative) in the southern (northern) hemisphere;
2. enhanced in a compact region around the conjugate point;
3. relatively high over regions in South America, southern Africa and the Maritime Continent.

There is a distinct transition in the sign of the correlation coefficient across the geographic equator: whereas in the southern hemisphere correlation is predominantly positive, it is principally negative north of the equator. There is no corresponding reversal across the geomagnetic equator. This can be explained by the seasonal difference in lightning activity in the two hemispheres: whistlers in Tihany are most prevalent during the southern hemisphere summer [Collier *et al.*, 2006], which is a time of profuse lightning activity south of the equator, but little lightning in the northern hemisphere. The situation is reversed during the southern hemisphere winter. The occurrence of Tihany whistlers is thus in phase with seasonal lightning activity in the southern hemisphere, but in antiphase with the northern hemisphere lightning season.

The spatially coherent region of positive correlation is centred on the conjugate point and extends to a radius of approximately 1000 km. This indicates that the majority of whistlers are well correlated with lightning activity within a few hundred km of the conjugate point. This

region should encompass most of the causative lightning discharges and is thus interpreted as the principal source region for whistlers at Tihany. This finding concurs with the mechanism outlined by *Storey* [1953], but places a reasonable upper bound on the extent of the principal
 425 source region.

The extended regions of relatively large positive correlation in tropical South America, southern Africa and over the Maritime Continent might initially appear to be artefacts of the analysis procedure, since it seems somewhat unlikely that whistlers are triggered by lightning strokes occurring at such great distances from the footpoint of the magnetic field line. However, sferics
 430 may travel enormous distances in the waveguide before they become too attenuated to generate detectable whistlers [*Li et al.*, 1991; *Holzworth et al.*, 1999; *Meredith et al.*, 2006; *Chum et al.*, 2006]. The strokes identified by WWLLN, and hence those considered in this analysis, are those with an above average intensity, so it is not implausible that they radiate sufficient energy to trigger a whistler at distances ~ 12000 km. This suggests that sferics generated by lightning
 435 over South America and the Maritime continent are able to survive propagation across the Atlantic and Indian Oceans. The waveguide propagation time for this distance is ~ 30 ms, much shorter than the typical magnetospheric propagation time for whistlers received at Tihany, which is ~ 500 ms. Furthermore, there is persuasive evidence that, for locations which have a dearth of lightning activity in the vicinity of their conjugate point, the majority of whistlers originate
 440 from distant lightning strokes [*Collier et al.*, 2009]. The extreme negative correlations achieved over the Maritime Continent north of the equator may be attributed to the non-uniform spatial distribution of the WWLLN detection efficiency, which is enhanced in this vicinity [*Rodger et al.*, 2006, Fig. 12].

A comparison of Figure 6 with global lightning activity [*Christian et al.*, 2003; *Rodger et al.*,
 445 2006] reveals the fact that regions of high lightning activity do not necessarily engender high
 correlation. The area of maximal lightning activity in equatorial Africa is uncorrelated with the
 whistlers. Maximal correlation occurs off the southeast coast of South Africa, an area of only
 moderate lightning activity. The active areas in tropical South America and the Maritime conti-
 450 nent are associated with a lesser degree of positive correlation. The area of positive correlation
 over South America is more extensive than that over the Maritime continent possibly due to the
 smaller land-to-ocean ratio in the latter area, where greater lightning activity generally occurs
 over land.

Although the spatial distribution of the correlation coefficients in Figure 6 is suggestive of a
 pattern which is in accord with expectation, the range is rather low, from a minimum of -0.037
 455 to a maximum of 0.065 . Under conventional conditions a correlation of 0.065 would be deemed
 inconclusive. Two topics should therefore be addressed: why are the correlations so low and
 might one still ascribe any significance to them?

The limited range of sample correlation coefficients may be attributed to a variety of sources:
 the appreciable mismatch in occurrence frequency between lightning and whistlers; the division
 460 of the lightning data amongst a number of cells; the imperfect efficiency of both the whistler
 and lightning detection systems; failure to isolate periods during which either system was not
 fully functional; or the inability of WWLLN to distinguish between CG and CC discharges.

Of the reasons mooted above, the first two are the most compelling. In general the number of
 lightning strokes exceeds the number of whistlers during a given interval [*Chum et al.*, 2006].
 465 This analysis proceeds on the premise that every whistler is generated by a lightning stroke but
 not every lightning stroke results in a whistler. The relatively low correlation even around the

conjugate point arises principally due to the fact that a given lightning stroke rarely results in a whistler. The mean flash rate around the conjugate point is roughly $10 \text{ km}^{-2} \text{ year}^{-1}$ [*Christian et al.*, 2003; *Collier et al.*, 2006], so that a region of radius 1000 km centred on the conjugate point has on average ~ 3600 flashes per hour. Values for the average flash multiplicity (number of strokes per flash) vary appreciably according to the observation technique. If one assumes a conservative mean multiplicity of 2.5 [*Rakov and Huffines*, 2003], then this translates to ~ 9000 strokes per hour. The average whistler detection rate at Tihany is ~ 24 per hour. Neglecting imperfections in the detection systems, it is possible to conclude that only approximately one in every 375 lightning strokes produces a whistler detectable on the ground. It is difficult to achieve a substantial correlation between two phenomena whose frequencies of occurrence differ by at least two orders of magnitude!

The distribution of the causative lightning strokes over a number of spatial cells also has a deleterious effect on the correlation. This is illustrated schematically in Figure 7. The horizontal grids represent a sequence of time intervals. Shaded cells indicate intervals which contain activity, while empty cells are inactive. Above the dashed line is a hypothetical sequence of whistlers, while below the line are two scenarios for gridding the lightning data. In the first scenario all of the causative strokes are located within a single spatial cell and the correlation with the whistler sequence is perfect. However, in the second scenario the original cell is divided into four smaller cells. Now the lightning strokes are divided into four separate sequences according to their location within each of the smaller cells. The correlations between each of the four new lightning sequences and the original whistler sequence are now significantly less than unity. The analogous effect applies to the data presented in Figure 6, where the causative strokes are distributed over numerous spatial cells.

It would indeed be of interest to conduct a similar analysis using a data set which differentiates between CG and CC discharges, since this might resolve the controversy regarding which of CG or CC strokes is more likely to yield a whistler. Although whistlers are associated with the most energetic lightning strokes [*Helliwell*, 1965, p. 121], which are most likely to be CG strokes, there is conflicting evidence [*Ferencz et al.*, 2007] which suggests that the primary source of whistlers are CC strokes, although this may be partly due to the higher incidence of CC lightning.

There is a subtle northwest-southeast orientation of the region of enhanced correlation, suggesting that strokes which are displaced parallel to the horizontal projection of the magnetic field (the local declination is -27°) relative to the conjugate point are preferred. This result concurs with *Yoshino* [1976], who found that strokes located between the conjugate point and the magnetic pole were most likely to generate whistlers. Furthermore, although *Strangeways* [1981] found that strokes at lower L were more likely to excite a duct, *Helliwell* [1965] indicates that matching between the neutral atmosphere and the ionosphere is best when waves approach the entry point from higher magnetic latitudes. Our results are not able to confirm or refute either of these hypotheses.

4.1. Equatorial Correlation

The regions of high lightning activity in the tropics have flash rates in excess of $10 \text{ km}^{-2} \text{ year}^{-1}$ [*Christian et al.*, 2003]. A 3° by 3° cell therefore experiences around 10^6 flashes per year or, on average, one flash every 30 s. Therefore, every time interval which contains a whistler at Tihany is very likely to also have a lightning discharge associated with it in these high activity cells. So, one might be tempted to attribute the elevated correlations over South America and the Maritime Continent to coincidence due to the profusion of lightning in

these areas. However, two factors oppose this line of reasoning. Firstly, in keeping with the logic of (1), the intervals which do not have whistlers, yet do have lightning activity, should make a negative contribution towards the correlation. Secondly, the diurnal variation in the correlation over these regions mirrors that around the conjugate point. Maximal correlations in these remote areas are not contemporaneous with peak local lightning activity.

The region of maximal correlation near the conjugate point is one in which there is only relatively low lightning activity. It is thus evident that the positive correlation in this area is not simply due to an overwhelming frequency of lightning fortuitously timed with respect to the whistlers. This fact may be further appreciated with reference to Figure 8, which represents the relationship between the correlation coefficient and the level of lightning activity. It is immediately evident that regions with high lightning density do not necessarily produce above average correlation. The grid cells with the highest correlation actually have very low stroke density. The 0.95 quantile of stroke density occurs at $0.243 \text{ km}^{-2} \text{ year}^{-1}$. The mean correlation coefficient for cells with stroke densities greater than the 0.95 quantile is 0.0032, which indicates that the regions of greatest lightning activity are only marginally correlated with whistlers at Tihany.

4.2. Diurnal Variation

The fact that whistlers are observed at Tihany with much greater probability during the hours of darkness [Collier *et al.*, 2006], negatively impacts on the quality of a correlation based on data from all local times, since during daylight any lightning stroke, no matter how favourably situated, is unlikely to generate a whistler. The data were thus decomposed into 8 units corresponding to consecutive 3 h intervals. The correlation analysis was then repeated, yielding the data displayed in Figure 9.

During the period 06:00 to 15:00 UTC, when whistlers are scarce at Tihany, there is little significant correlation anywhere. However, between the mid-afternoon and predawn hours the pattern of relatively high correlation emerges. This regimen applies not only around the conjugate point but also over South America and the Maritime Continent. The fact that the variation of the correlation over these remote regions follows the same diurnal pattern as that near the conjugate point suggests that the effectiveness of lightning strokes over these regions is heightened when the ionosphere over Tihany and its conjugate point are in darkness and that the impulses from the distant strokes are travelling a significant distance in the waveguide before entering the magnetosphere. Furthermore, it is important to note that the peaks in lightning activity over the Maritime Continent and South America occur at roughly 08:00 UTC and 20:00 UTC respectively [Price, 1993]. In the case of the former, the period of enhanced correlation does not correspond to the peak in lightning activity, indicating again that it is ionospheric transparency over the source and receiver that is operative.

Various observations of whistlers indicate that paths at magnetic longitudes nearest to the source or receiver are most favourable [Allcock and McNeill, 1966; Shimakura *et al.*, 1987; Ladwig and Hughes, 1989; Clilverd *et al.*, 1992]. If one were to apply this, for example, to causative lightning strokes over South America then two options exist: either the whistlers are ducted to the northern hemisphere from around the lightning discharge location and then propagate in the waveguide to Tihany, or the spheric travels in the waveguide and enters a duct in the vicinity of Tihany's conjugate point. One might fairly readily discard the first option in light of the low-latitude cutoff for whistlers. This then leaves the latter option, which is also compatible with the fact that the illumination of the ionosphere over Tihany and its conjugate point appears to control the correlation over South America.

4.3. Correlation and Causation

The foregoing analysis indicates that the region of highest correlation between lightning activity and Tihany whistlers is located within a region centred on the conjugate point. One is thus inclined to infer a deterministic relationship between the two phenomena. However, even
 560 a strong correlation between two quantities does not automatically imply that there is a causal relationship between them. Certainly it is the case that statistically independent quantities are always uncorrelated, but the converse is not necessarily true. However, a positive correlation does indeed *suggest* a cause-and-effect relationship and may be taken as non-conclusive evidence of such.

565 Furthermore, although correlation does not imply causation, a correlation that supports an existing hypothesis is more compelling than one obtained on an investigative basis. By analogy, it is more impressive for a golfer to hit a hole-in-one immediately after predicting the feat than simply doing so by chance. In this study the positive correlation surrounding the conjugate point supports the existing theory advanced by *Storey* [1953].

4.4. Statistical Significance

570 In assessing the plausibility of these results the statistical significance of the calculated correlation coefficients becomes relevant. Figure 6 reflects only those cells which have $p < 0.01$, where the p -value represents the probability of achieving a result at least as extreme as that obtained assuming that the null hypothesis (no correlation) is, in fact, true.

One in one hundred random data sets might be expected to yield correlations exceeding a
 575 given value at the $\alpha = 0.01$ significance level on the basis of chance alone. The likelihood of this occurring might be enhanced by the selection of an area of the globe for which there exists a qualitative similarity between lightning incidence and whistler occurrence. However, the fact

that elevated correlations are observed in a spatially coherent region surrounding the conjugate point suggests that this is not a chance occurrence and supports the existing hypothesis regarding the formation of whistlers.

Various caveats apply to the interpretation of the correlation coefficient associated with time sequences. The statistical significance of r depends on the sample size or, more specifically, the number of independent observations. The standard technique for calculating the correlation p -value assumes that the populations from which the samples are drawn are normally distributed and that the samples within each of the sequences are independent. Although the two sequences used for calculating the correlation coefficient consist of Boolean samples, it can be shown that the resulting statistic r has a normal distribution. The individual samples are not, however, drawn from a normal distribution. This issue will be addressed in a forthcoming publication. The latter requirement, that each of the series used to calculate the correlation coefficient represents a sequence of statistically independent samples, is rarely satisfied for time series, which generally possess an appreciable degree of autocorrelation. It is quite clear that this is not true of the lightning data, since the presence of lightning in a given time interval certainly increases the likelihood of lightning in the subsequent intervals. The resulting uncertainties may be underestimated as a result. Autocorrelation does not, however, have an effect on the calculated correlation coefficient itself.

The situation may be remedied by employing an effective sample size, \hat{N} , to estimate the number of degrees of freedom, thereby making allowance for autocorrelation within the data. The effective sample size is a function of the nominal sample size and the first order autocorrelation of both of the variables [Dawdy and Matalas, 1964]. Application of an effective sample size should only be considered if the autocorrelation estimates are statistically significant. The

distribution of autocorrelations for sequences of N random samples is approximately normal with zero mean and variance $1/N$ [Panofsky and Brier, 1958; Chatfield, 1975]. Thus, with a probability of roughly 0.95, the autocorrelation of a random sequence lies within $\pm 2/\sqrt{N}$ of zero. A value lying outside this range is significant at the $\alpha = 0.05$ level. The autocorrelation of a random sequence of length 1251360 is thus likely to be very small indeed. Although it is possible to obtain spurious autocorrelation values for short random sequences, for sequences of reasonable length, this contingency does not arise.

The nominal sample size, based on the number of time intervals spanning the period of interest, was $N = 1251360$. The resulting values of the effective sample size varied from 502691 to 1251948, with the lowest values corresponding to those areas with the highest lightning activity. A reduction in the effective sample size results in a larger sample standard deviation, which, in turn, broadens the confidence intervals. The average effective sample size was $\langle \hat{N} \rangle = 1116992 \pm 1972$, which does not differ significantly from the nominal sample size. The width of the confidence intervals is adjusted by the factor

$$\sqrt{(N - 2)/(\hat{N} - 2)}. \quad (2)$$

Since the ratio N/\hat{N} is at most ~ 2 , the factor in (2) is ~ 1.4 or less, and on average is very close to unity. Therefore the effect of autocorrelation for the data considered here does not have a dramatic effect on the width of the confidence intervals.

In principle the most robust technique for determining the significance of r would be to use a permutation test. This would obviate the need for computing an effective sample size. For the large number of cells (120×60) and the long sequences for each cell (1251360 elements) the computations would be both formidable and arduous. However, to illustrate that the use of a parametric test in this case is not inappropriate, we compared the confidence intervals

obtained for the correlation coefficient between two synthetic sequences. The 99.0% confidence interval obtained under the assumption that the correlation coefficient has a normal distribution is $[0.599, 0.601]$, centred on the estimate of 0.600. Several options exist for constructing the bootstrap confidence interval: a range of 2.576 standard deviations either side of the mean yields $[0.598, 0.602]$, while the empirical 0.5% and 99.5% percentiles are $[0.598, 0.602]$, both of which are in good agreement with the parametric estimate.

Once the effective sample size is known, determination of the p -value associated with the correlation coefficient is straightforward [Fisher, 1941, p. 186]. Figure 10(a) plots the relationship between the sample correlation coefficient and the associated p -value for all cells on the geographic grid. The dashed horizontal line indicates the 0.01 significance level. The null hypothesis may be rejected at the 0.01 level for all points lying below this line. It is thus apparent that correlations with absolute magnitude greater than 0.002 are statistically significant and contradict the null hypothesis. Locations which conform to this criterion for statistical significance occupy 49% of the area of the Earth.

Confidence intervals may be readily computed provided that the populations are normally distributed and pairs of observations are selected at random. Figure 10(b) relates the 99.0% confidence intervals to the calculated values of r . The confidence interval is ± 0.00230 . It is clear that only a small range of correlation coefficients (delimited by the vertical dotted lines) have a confidence interval which includes the null hypothesis.

5. Conclusions

A correlation analysis may suggest the possibility of a deterministic relationship between two phenomena. Motivated by such an indication, a valid theory may be constructed on the basis of other factors beyond mere correlation. If, however, a correlation analysis corroborates some

aspect of an existing theory, then this is very compelling evidence in favour of the veracity of the theory.

640 Whistler data from Tihany has been correlated against the corresponding global lightning data from WWLLN. The objective of this analysis was to examine the co-occurrence of lightning and whistler activity, not to obtain a one-to-one relationship between individual whistlers and the causative lightning discharges. The investigation was thus not focused on the absolute number of whistlers or strokes in a given interval but whether or not there was such activity. The
645 whistler and lightning counts were therefore reduced to Boolean sequences, simply indicating the presence or absence of activity in a given interval. The use of intervals of reasonable duration and Boolean values partially ameliorated the imperfections of the lightning and whistler detection systems.

The region of highest correlation is centred on the conjugate point with a radial extent of
650 ~ 1000 km. This suggests that the majority of causative strokes are located close to the conjugate point, and is in accord with the generally accepted theory of whistler generation. This can thus be regarded as the principal source region for whistlers recorded at Tihany. There are also, however, areas of positive correlation further afield, indicating that sferics may travel great distances in the waveguide before entering a duct. The fact that the diurnal variation in
655 the correlation coefficient in these remote regions is congruent with that around the conjugate point indicates that the characteristics of the ionosphere in the vicinity of receiver's meridian are operative in determining the generation of a whistler by a remote lightning stroke.

The principal source region is also not circular, but exhibits an elongation along an axis in the northwest-southeast direction. The declination at the conjugate point is -27° , so that the axis
660 is roughly aligned with the local magnetic meridian. Furthermore, the region of positive cor-

relation appears to extend to a greater distance in the southeasterly direction, which is towards the magnetic pole, consistent with the fact that the majority of whistlers at Tihany propagated at L higher than that of the station. These facts support the model for coupling between the atmosphere and ducts presented by *Helliwell* [1965, Fig. 3–23], which suggests that at low
 665 and medium latitudes, the coupling improves for waves originating from the poleward side of the footpoint of the duct. This should be contrasted with *Strangeways* [1981], who contended that the transmission of whistlers through the ionosphere is most effective when the lightning discharge is equatorward of the ionospheric footprint of the duct.

Although the correlation between whistlers and lightning is strongly significant over 49%
 670 of the area of the Earth, the correlation coefficients are small, with the implication that only a minute fraction of the variance in either of the variables can be explained by linear regression on the other variable. However, this was not the objective of this study, which was simply intended to identify those regions of the globe with the highest correlation.

A possible source of concern with the local significance tests on the correlation coefficients
 675 for points which are likely to exhibit spatial correlation is whether or not the null hypothesis is rejected for a significant fraction of the globe. To assess this issue, global significance might be determined using Monte Carlo resampling as described by *Livezey and Chen* [1983]. However, it is readily apparent from Figures 6 and 10(a) that the regions which violate the null hypothesis constitute only a small fraction of the globe. Furthermore, the dominant region is centred on
 680 that portion of the globe which is magnetically conjugate to the receiver, which is the area from which whistlers were *a priori* assumed to originate.

The technique outlined here promises to provide a convincing indication of the most likely source of whistlers detected at a given location. The present analysis should be regarded as

preliminary. Since the efficiency of WWLLN has improved significantly over the period for
 685 which whistler data is available, the quality of the correlations is not consistent over the duration.
 Because the quality and reliability of the WWLLN data is not always well known, it may be
 worthwhile repeating this analysis using lightning data from a different source such as ZEUS
 (<http://sifnos.engr.uconn.edu/system.htm>).

Acknowledgments. We are grateful to Sherry Bremner for her contribution to the data analy-
 690 sis. We acknowledge the numerous sites hosting WWLLN nodes and their efforts in maintaining
 the consistent global coverage of the network.

References

- Allcock, G. M., and F. A. McNeill (1966), The reception of whistler mode signals at a point
 remote from the transmitter's magnetic conjugate point, *J. Geophys. Res.*, *71*(9), 2285–2294.
- Bernard, L. C. (1973), A new nose extension method for whistlers, *J. Atmos. Terr. Phys.*, *35*,
 695 871–880.
- Bernstein, S. L., M. L. Burrows, J. E. Evans, A. S. Griffiths, D. A. McNeil, C. W. Niessen,
 I. Richer, D. P. White, and D. K. Willim (1974), Long-range communications at extremely
 low frequencies, *Proc. IEEE*, *62*(3), 292–312.
- Brice, N. M. (1960), Traveling wave amplification of whistlers, *J. Geophys. Res.*, *65*(11), 3840–
 700 3842.
- Carpenter, D. L. (1959), Identification of whistler sources on visual records and a method of
 routine whistler analysis, *Tech. rep.*, Radioscience Laboratory, Stanford Electronics Labora-
 tories, Stanford University.

Carpenter, D. L. (1963), Whistler evidence of a ‘knee’ in the magnetospheric ionization density profile, *J. Geophys. Res.*, 68(6), 1675–1682.

Carpenter, D. L., and R. E. Orville (1989), The excitation of active whistler mode signal paths in the magnetosphere by lightning; two case studies, *J. Geophys. Res.*, 94(A7), 8886–8894.

Chatfield, C. (1975), *The Analysis of Time Series: Theory and Practice*, Chapman and Hall.

Chow, V. T. (Ed.) (1964), *Handbook of Applied Hydrology, A Compendium of Water-Resources Technology*, McGraw-Hill Book Company, New York.

Christian, H. J., et al. (2003), Global frequency and distribution of lightning as observed from space by the optical transient detector, *J. Geophys. Res.*, 108(D1), doi:10.1029/2002JD002347.

Chum, J., F. Jiricek, O. Santolik, M. Parrot, G. Diendorfer, and J. Fiser (2006), Assigning the causative lightning to the whistlers observed on satellites, *Annales Geopys.*, 24(11), 2921–2929.

Clilverd, M. A., N. R. Thomson, and A. J. Smith (1992), Observation of two preferred propagation paths for whistler mode vlf signals received at a non-conjugate location, *J. Atmos. Terr. Phys.*, 54(7–8), 1075–1079, doi:10.1016/0021-9169(92)90073-T.

Collier, A. B., A. R. W. Hughes, J. Lichtenberger, and P. Steinbach (2006), Seasonal and diurnal variation of lightning activity over southern africa and correlation with european whistler observations, *Annales Geopys.*, 24(2), 529–542.

Collier, A. B., S. Bremner, A. R. W. Hughes, J. Lichtenberger, P. Steinbach, C. J. Rodger, and G. McDowell (2009), Spatial distribution of causative lightning discharges for whistlers observed at Dunedin, New Zealand, *Geophys. Res. Lett.*, in preparation.

Dawdy, D. R., and N. C. Matalas (1964), *Statistical and probability analysis of hydrologic data, Part III. Analysis of variance, covariance, and time series*, chap. 8–III, pp. 8.68–8.90, in *Chow* [1964].

Dowden, R. L., J. B. Brundell, and C. J. Rodger (2002), VLF lightning location by time of group arrival (TOGA) at multiple sites, *J. Atmos. Solar-Terr. Phys.*, *64*(7), 817–830, doi:10.1016/S1364-6826(02)00085-8.

Dowden, R. L., et al. (2008), World-wide lightning location using VLF propagation in the earth-ionosphere waveguide, *IEEE Antennas Propag. Mag.*, *50*(5), 40–60, doi:10.1109/MAP.2008.4674710.

Ferencz, O. E., C. Ferencz, P. Steinbach, J. Lichtenberger, D. Hamar, M. Parrot, F. Lefeuvre, and J.-J. Berthelier (2007), The effect of subionospheric propagation on whistlers recorded by the DEMETER satellite – observation and modelling, *Annales Geophys.*, *25*, 1103–1112.

Fisher, R. A. (1941), *Statistical Methods for Research Workers*, eighth ed., Oliver and Boyd.

Gamble, R. J., C. J. Rodger, M. A. Clilverd, J.-A. Sauvaud, N. R. Thomson, S. L. Stewart, R. J. McCormick, M. Parrot, and J.-J. Berthelier (2008), Radiation belt electron precipitation by man-made vlf transmissions, *J. Geophys. Res.*, *113*, A10211, doi:10.1029/2008JA013369.

Green, J. L., S. Boardsen, L. Garcia, W. W. L. Taylor, S. F. Fung, and B. W. Reinisch (2005), On the origin of whistler mode radiation in the plasmasphere, *J. Geophys. Res.*, *110*(A3), A03201, doi:10.1029/2004JA010495.

Helliwell, R. A. (1965), *Whistlers and Related Ionospheric Phenomena*, Stanford University Press.

Holzworth, R. H., R. M. Winglee, B. H. Barnum, Y. Q. Li, and M. C. Kelley (1999), Lightning whistler waves in the high-latitude magnetosphere, *J. Geophys. Res.*, *104*(A8), 17,369–

17,378.

750 Hughes, A. R. W. (1981), Satellite measurements of whistler dispersion at low latitudes, *Adv. Space Res.*, *1*(1), 377–380, doi:10.1016/0273-1177(81)90138-1.

Hughes, A. R. W., and W. K. Rice (1997), A satellite study of low latitude electron and proton whistlers, *J. Atmos. Solar-Terr. Phys.*, *59*(10), 1217–1222, doi:10.1016/S1364-6826(96)00111-3.

755 Jacobson, A. R., R. H. Holzworth, J. Harlin, R. L. Dowden, and E. H. Lay (2006), Performance assessment of the world wide lightning location network (wwlln), using the los alamos sferic array (lasa) array as ground-truth, *J. Atmos. Ocean. Tech.*, *23*, 1082–92.

Johnson, M. P., U. S. Inan, and D. S. Lauben (1999), Subionospheric vlf signatures of oblique (nonducted) whistler-induced precipitation, *Geophys. Res. Lett.*, *26*(23), 3569.

760 Ladwig, J. M., and A. R. W. Hughes (1989), An asymmetry in the direction of arrival of whistlers at Sanae, Antarctica, *J. Atmos. Terr. Phys.*, *51*(2), 61–65, doi:10.1016/0021-9169(89)90105-0.

Lauben, D. S., U. S. Inan, and T. F. Bell (1999), Poleward-displaced electron precipitation from lightning-generated oblique whistlers, *Geophys. Res. Lett.*, *26*(16), 2633–2636.

765 Lay, E. H., R. H. Holzworth, C. J. Rodger, J. N. Thomas, O. Pinto, Jr, and R. L. Dowden (2004), Wwll global lightning detection system: Regional validation study in brazil, *Geophys. Res. Lett.*, *31*, L03102, doi:10.1029/2003GL018882.

Li, Y. Q., R. H. Holzworth, H. Hu, M. McCarthy, R. D. Massey, P. M. Kintner, J. V. Rodrigues, U. S. Inan, and W. C. Armstrong (1991), Anomalous optical events detected by rocket-borne
770 sensor in the WIPP campaign, *J. Geophys. Res.*, *96*(A2), 1315–1326.

Lichtenberger, J., C. Ferencz, L. Bodnár, D. Hamar, and P. Steinbach (2008a), Automatic whistler detector and analyzer system: Automatic whistler detector, *J. Geophys. Res.*, *113*, A12201, doi:10.1029/2008JA013467.

Lichtenberger, J., C. Ferencz, O. E. Ferencz, D. Hamar, P. Steinbach, and C. Rodger (2008b),
 775 Obtaining plasmaspheric electron density profiles: A new whistler inversion model, poster presented at the XXIX URSI General Assembly.

Liemohn, H. B. (1967), Cyclotron-resonance amplification of VLF and ULF whistlers, *J. Geophys. Res.*, *72*(1), 39–55.

Livezey, R. E., and W. Chen (1983), Statistical field significance and its determination by
 780 monte carlo techniques, *Mon. Weather Rev.*, *111*(1), 46–59, doi:10.1175/1520-0493(1983)111<0046:SFSaid>2.0.CO;2.

Meredith, N. P., R. B. Horne, M. A. Clilverd, D. Horsfall, R. M. Thorne, and R. R. Anderson (2006), Origins of plasmaspheric hiss, *J. Geophys. Res.*, *111*, A09217, doi:10.1029/2006JA011707.

785 Panofsky, H. A., and G. W. Brier (1958), *Some Applications of Statistics to Meteorology*, Pennsylvania State University Press, University Park, Pennsylvania.

Park, C. G. (1972), Methods of determining electron concentrations in the magnetosphere from nose whistlers, *Tech. Rep. 3454–1*, Radioscience Laboratory, Stanford Electronics Laboratories, Stanford University.

790 Pasmanik, D. L., A. G. Demekhov, D. Nunn, V. Y. Trakhtengerts, and M. J. Rycroft (2002), Cyclotron amplification of whistler-mode waves: A parametric study relevant to discrete vlf emissions in the earth's magnetosphere, *J. Geophys. Res.*, *107*(A8), 1162, doi:10.1029/2001JA000256.

Price, C. (1993), Global surface temperatures and the atmospheric electrical circuit, *Geophys. Res. Lett.*, 20(13), 1363–1366.

Rakov, V. A., and G. R. Huffines (2003), Return-stroke multiplicity of negative cloud-to-ground lightning flashes, *J. Appl. Meteorol.*, 42(10), 1455–1462, doi:10.1175/1520-0450(2003)042<1455:RMONCL>2.0.CO;2.

Rao, M., V. V. Somayajulu, S. K. Dikshit, and B. A. P. Tantry (1974), Low-latitude cutoff for whistlers observed on the ground, *J. Geophys. Res.*, 79(25), 3867–3869.

Rodger, C. J., R. J. McCormick, and M. A. Clilverd (2004), Testing the importance of precipitation loss mechanisms in the inner radiation belt, *Geophys. Res. Lett.*, 31, L10803, doi:10.1029/2004GL019501.

Rodger, C. J., J. B. Brundell, and R. L. Dowden (2005), Location accuracy of VLF world-wide lightning location (WWLL) network: Post-algorithm upgrade, *Annales Geophys.*, 23(2), 277–290.

Rodger, C. J., S. Werner, J. B. Brundell, E. H. Lay, N. R. Thomson, R. H. Holzworth, and R. L. Dowden (2006), Detection efficiency of the vlf world-wide lightning location network (wwlln): initial case study, *Annales Geophys.*, 24, 3197–3214.

Rodger, C. J., J. B. Brundell, R. H. Holzworth, and E. H. Lay (2008), Growing detection efficiency of the world wide lightning location network, in *Proceedings of the Workshop on Coupling of Thunderstorms and Lightning Discharges to Near-Earth Space*, in review.

Ruhnke, L. H., and J. Latham (Eds.) (1983), *Proceedings in Atmospheric Electricity*, A. Deepak Publishing.

Shimakura, S., A. Tsubaki, and M. Hayakawa (1987), Very unusual low latitude whistlers with additional traces of the earth-ionosphere waveguide propagation effect, *J. Atmos. Terr. Phys.*,

49(11–12), 1081, doi:10.1016/0021-9169(87)90091-2.

Sonwalkar, V. S., and U. S. Inan (1989), Lightning as an embryonic source of vlf hiss, *J. Geophys. Res.*, *94*(6), 6986–6994.

820 Storey, J. R. K. . L. R. O. (1955), An attempt to observe whistling atmospherics near the magnetic equator, *Nature*, *175*, 36–37, doi:10.1038/175036a0.

Storey, L. R. O. (1953), An investigation of whistling atmospherics, *Phil. Trans. Roy. Soc.*, *A-246*, 113–141.

Strangeways, H. J. (1981), Trapping of whistler-mode waves in ducts with tapered ends, *J.*
825 *Atmos. Terr. Phys.*, *43*(10), 1071–1079, doi:10.1016/0021-9169(81)90022-2.

Strangeways, H. J. (1986), Whistler leakage from narrow ducts, *J. Atmos. Terr. Phys.*, *48*(5), 455–462, doi:10.1016/0021-9169(86)90122-4.

Strangeways, H. J., and M. J. Rycroft (1980), Trapping of whistler-waves through the side of ducts, *J. Atmos. Terr. Phys.*, *42*(11–12), 983–994, doi:10.1016/0021-9169(80)90112-9.

830 Tarcsai, G. (1975), Routine whistler analysis by means of accurate curve fitting, *J. Atmos. Terr. Phys.*, *37*, 1447–1457.

Tarcsai, G., P. Szemerédy, and L. Hegymegi (1988), Average electron density profiles in the plasmasphere between $L = 1.4$ and 3.2 deduced from whistlers, *J. Atmos. Terr. Phys.*, *50*(7), 607–611, doi:10.1016/0021-9169(88)90058-X.

835 Thorne, R. M., and R. B. Horne (1994), Landau damping of magnetospherically reflected whistlers, *J. Geophys. Res.*, *99*(A9), 17,249–17,258.

Uman, M. A. (1987), *The Lightning Discharge*, Academic Press, San Diego, California.

Weidman, C. D., E. P. Krider, C. G. Park, and B. Blockley (1983), Correlated measurements of lightning radiation fields and whistlers, in *Ruhnke and Latham* [1983], pp. 334–337.

- ⁸⁴⁰ Yoshino, T. (1976), Low-latitude whistlers and cloud distributions in the conjugate area, *J. Geophys. Res.*, *81*(A25), 4793–4796.

Figure 1. Whistler recorded at Tihany, Hungary, at 20:45 UTC on 6 March 2002.

Figure 2. Average diurnal variation of whistler count rate at Tihany.

Figure 3. Distribution of L obtained from fitting whistlers observed at Tihany.

	# whistlers	recording
green	= 0	no
red	= 0	yes
grey	> 0	yes

Figure 4. Status of the AWD operating at Tihany, Hungary, indicating whether the system was recording data and whether or not any whistlers were identified.

Table 1. Contingency table reflecting the proportion of time intervals for which whistlers were observed at Tihany and lightning was detected in the cell centred at 34.5° S 28.5° E. Counts are normalised relative to the total number of time intervals, 1251360.

	no lightning	lightning	total
no whistlers	0.938	0.002	0.940
whistlers	0.060	0.001	0.060
total	0.997	0.003	1

Figure 5. Occurrence histograms for lightning strokes and whistlers during a 2 h interval around dusk on 26 February 2004. Histogram bins have 1 min width. The red bars represent the resulting Boolean histogram. The blue diamonds indicate those bins for which the condition (1) is satisfied.

Figure 6. Correlation between whistler observations at Tihany, Hungary, and global lightning strokes for $\Delta t = 1$ min. Data are plotted only in cells for which the correlation is statistically significant. The geomagnetic equator is indicated by the dashed curve. The location of Tihany is reflected by a pentagram. The conjugate point is surrounded by circles at intervals of 200 km up to a distance of 1000 km. The magnetic meridian linking Tihany to the conjugate point is indicated by a dotted curve.

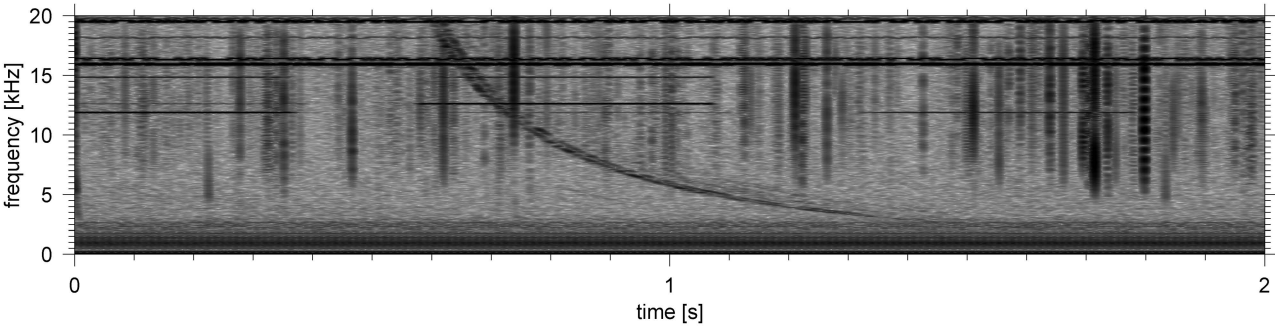
Figure 7. Schematic illustration of the effect of dividing causative lightning strokes over a number of spatial cells.

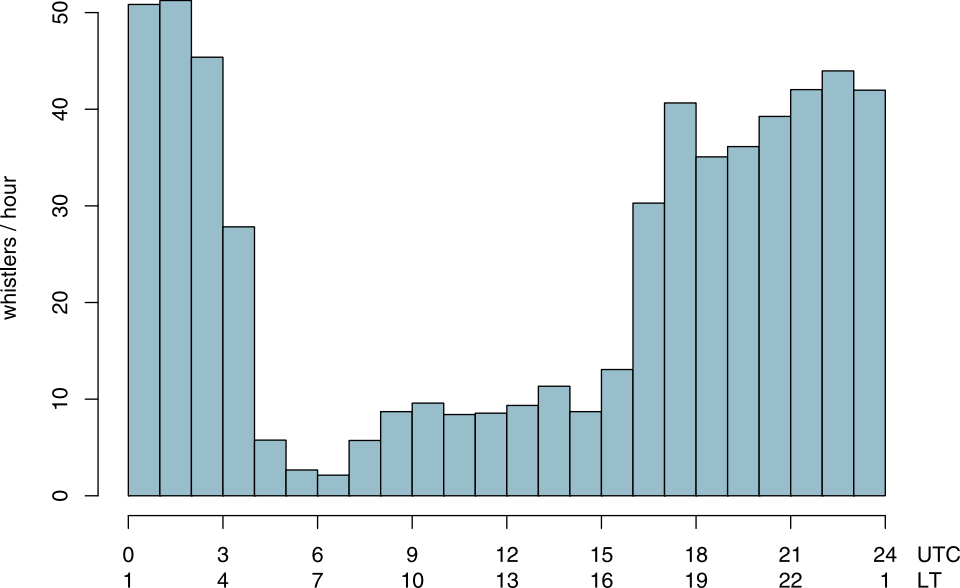
Figure 8. Relationship between correlation coefficient and lightning stroke rate for all geographic cells. The red points correspond to cells which lie within 1000 km of the conjugate point. The vertical dashed line corresponds to the 0.95 quantile in stroke density.

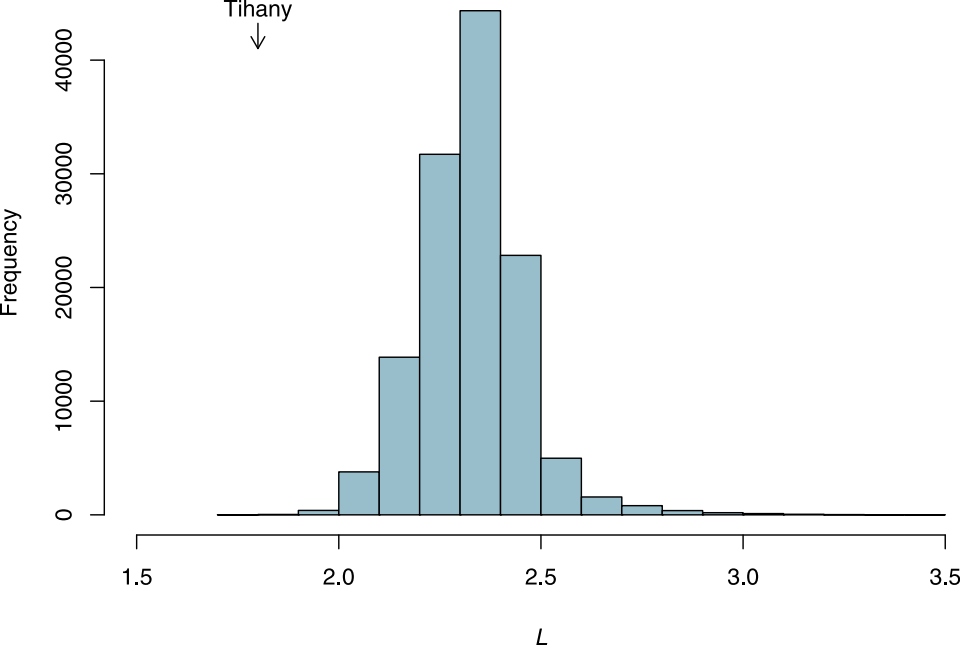
Figure 9. Correlation coefficients derived from data decomposed into 3 h periods. The UTC for each period is indicated to the right of each panel.

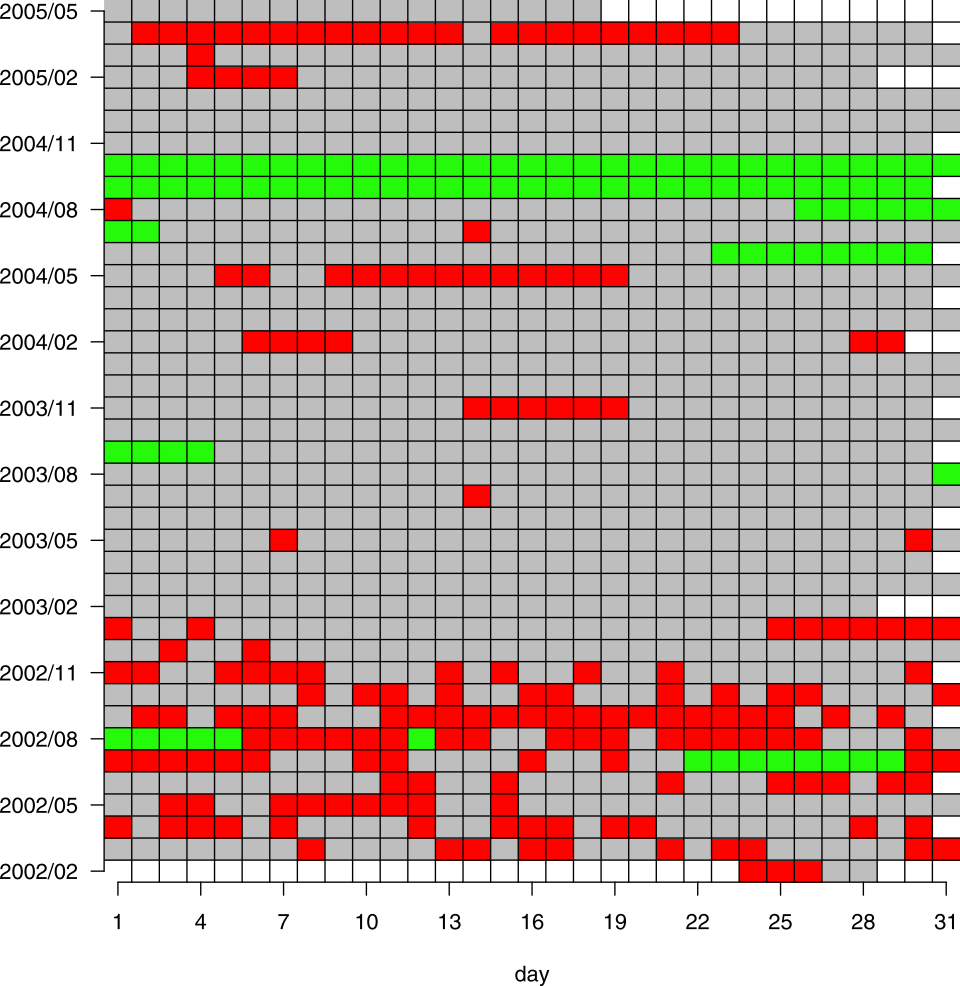
(a)(b)
 p -con-
 value
 dence
 in-
 ter-
 vals

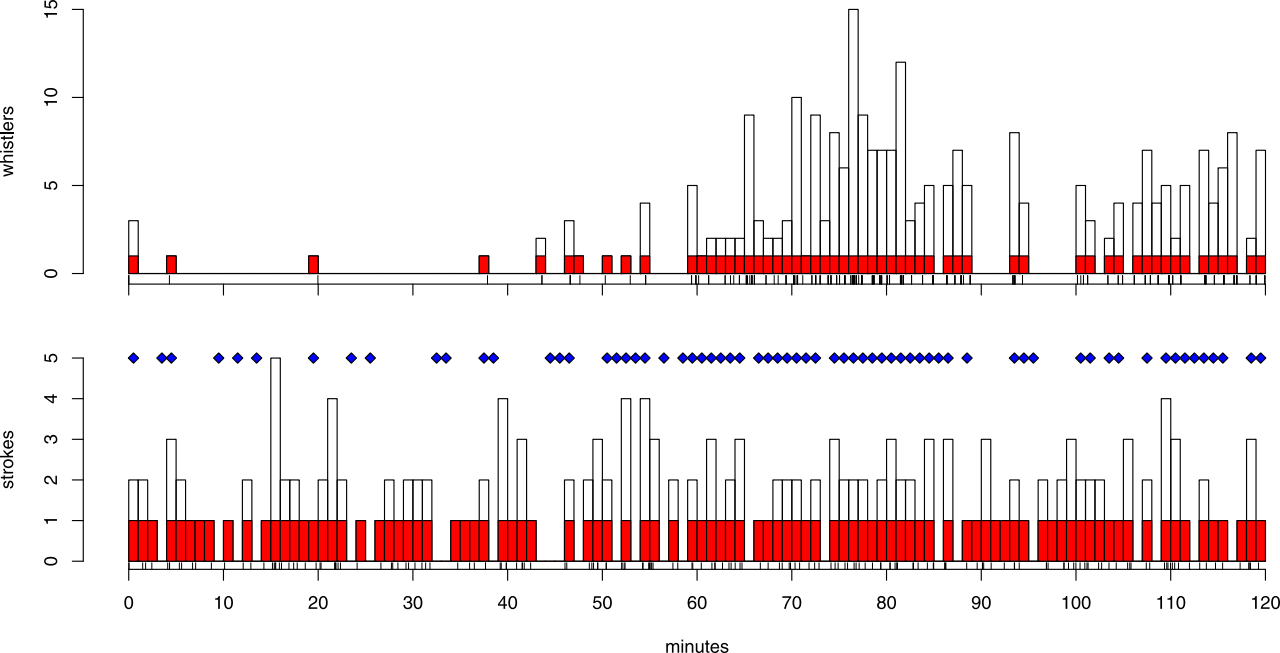
Figure 10. Statistical significance of the calculated correlation coefficients: (a) logarithm of the probability that the null hypothesis (zero correlation) is valid, and (b) 99.0% confidence intervals as a function of calculated correlation coefficient.

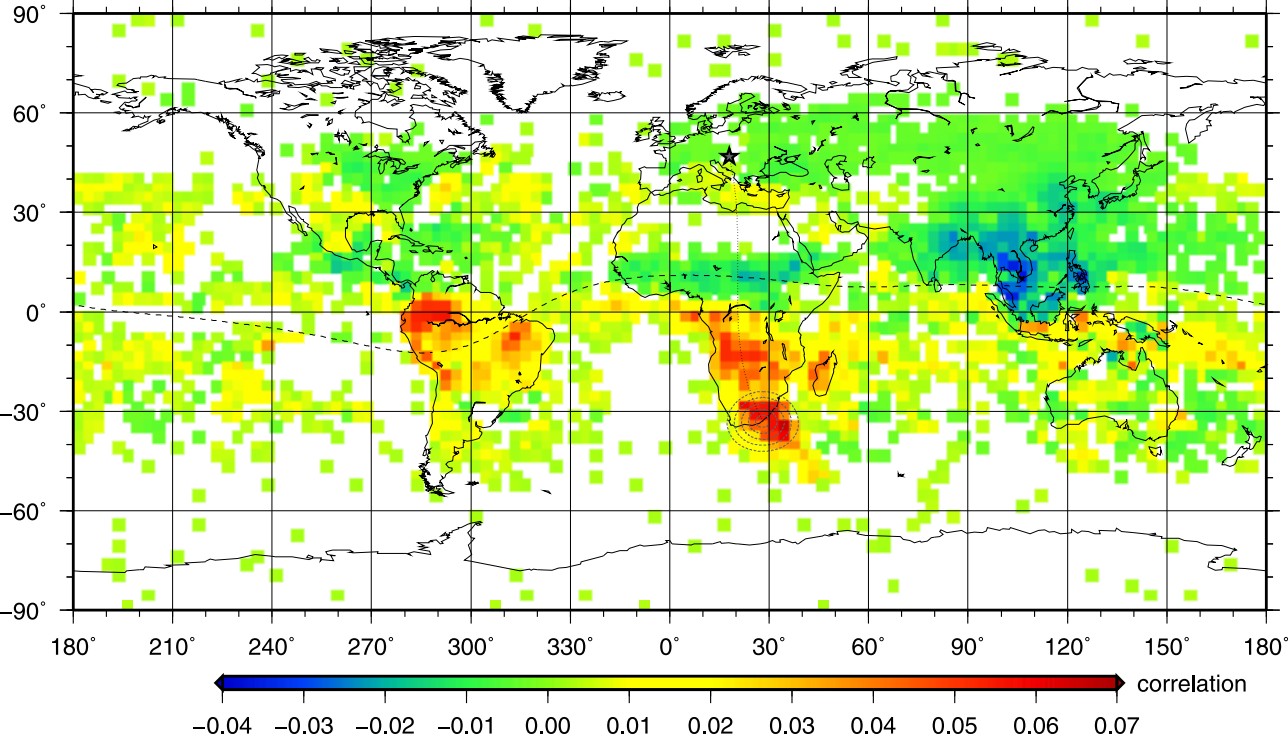


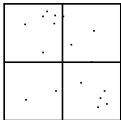
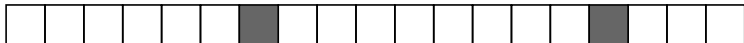
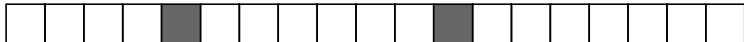


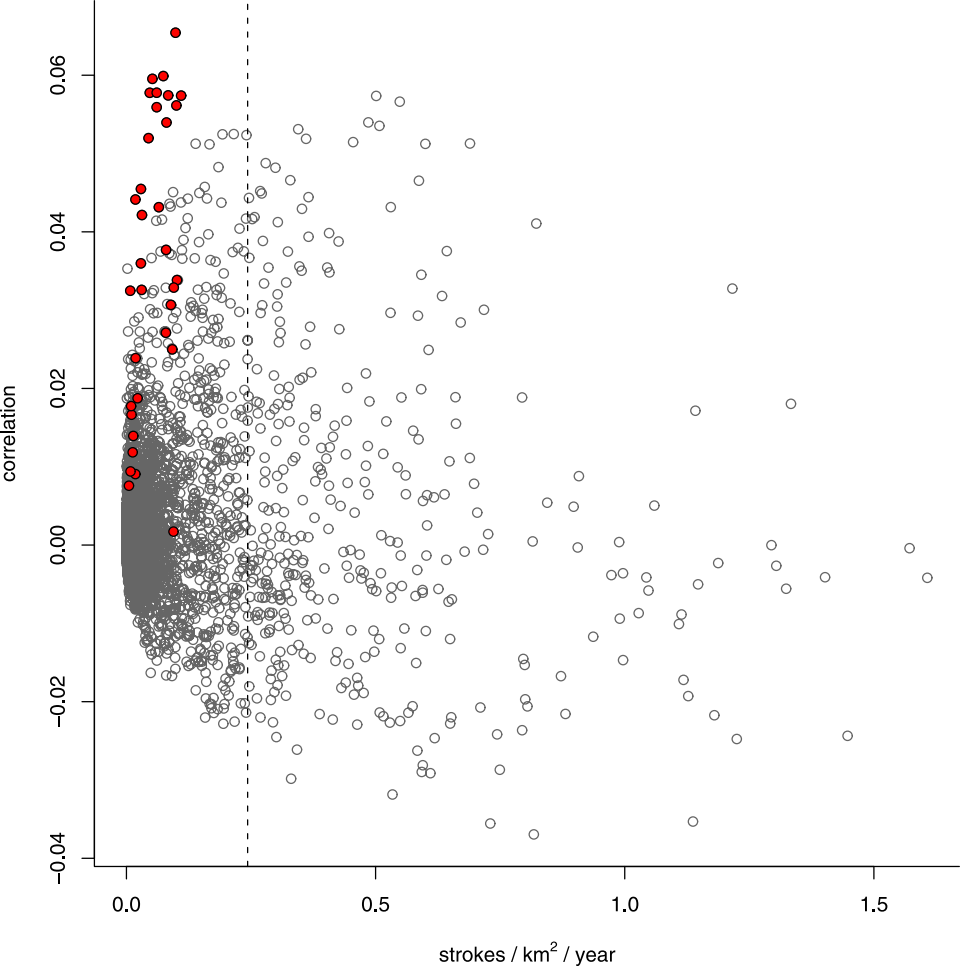


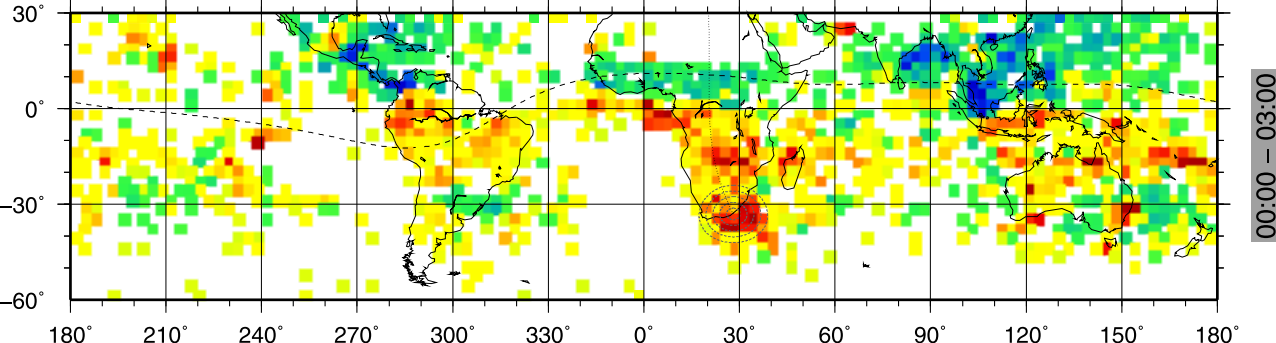


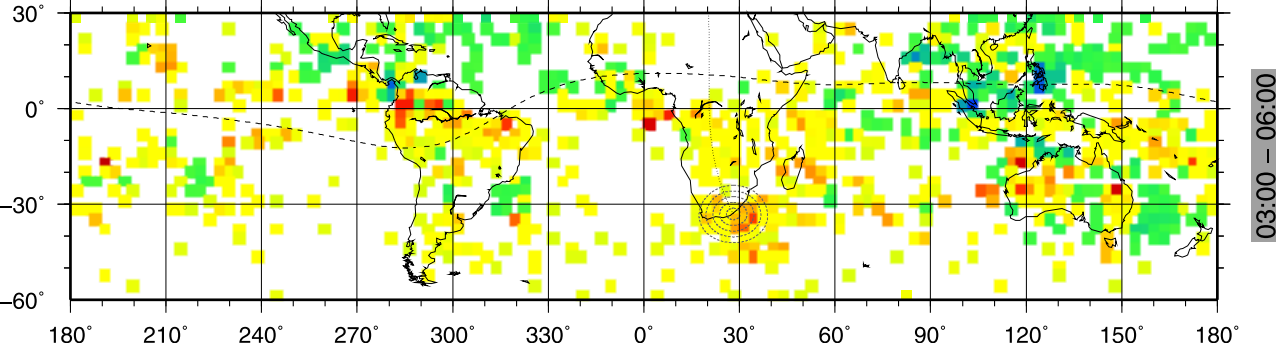


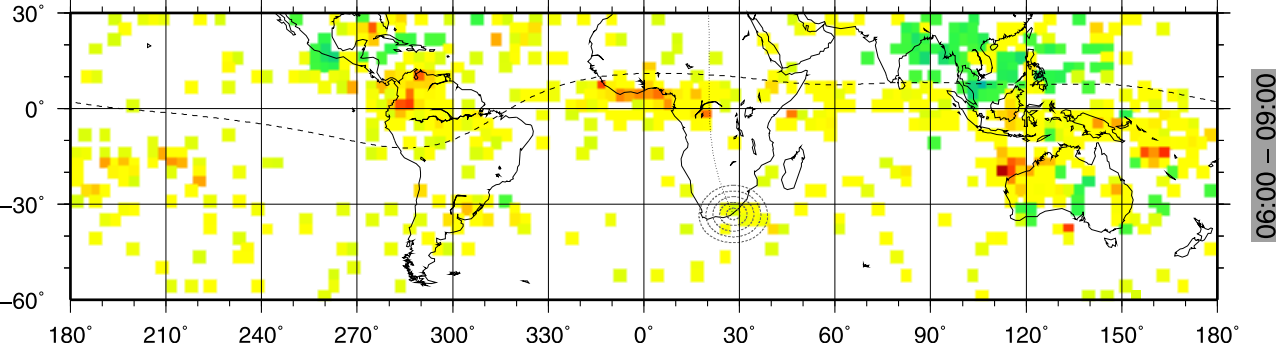


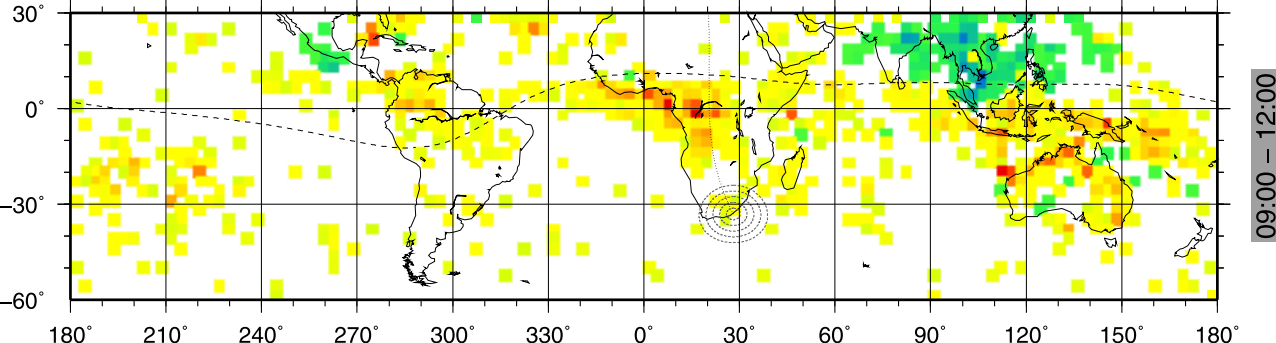


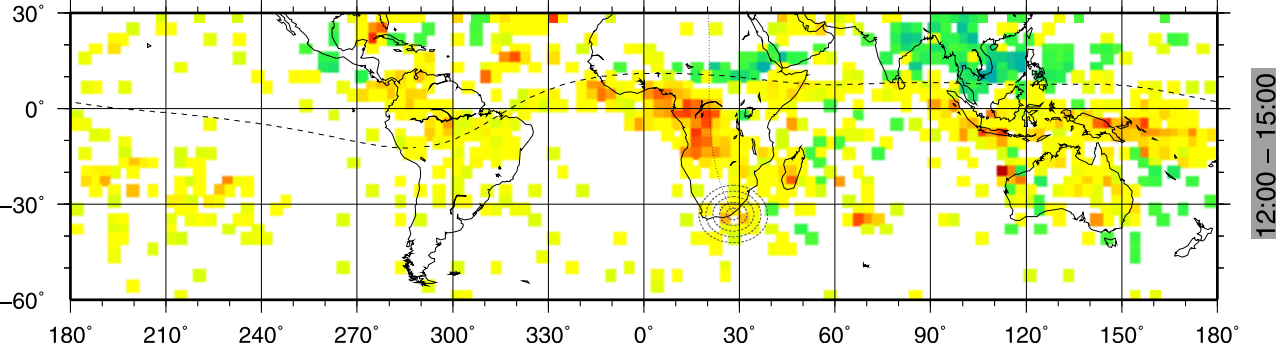


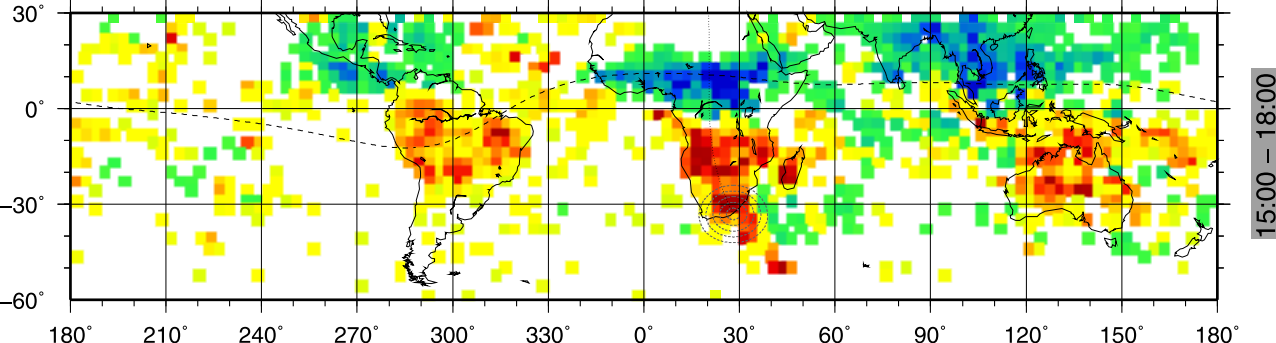


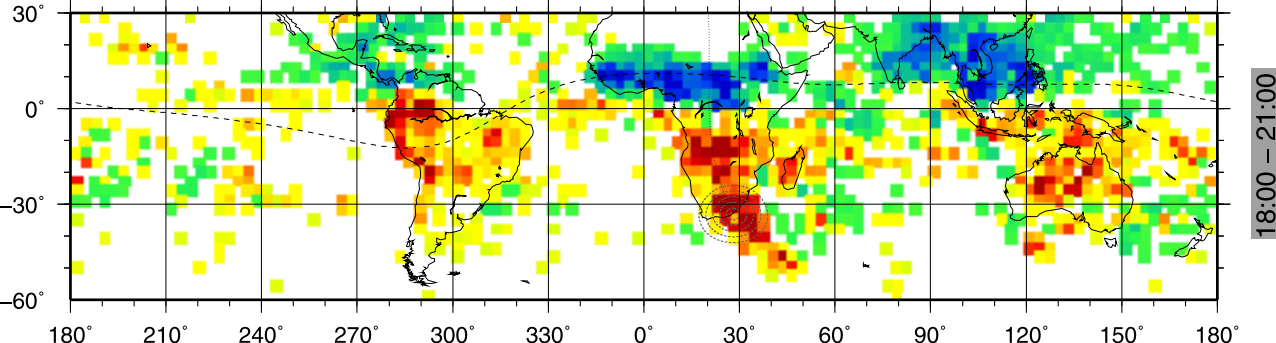


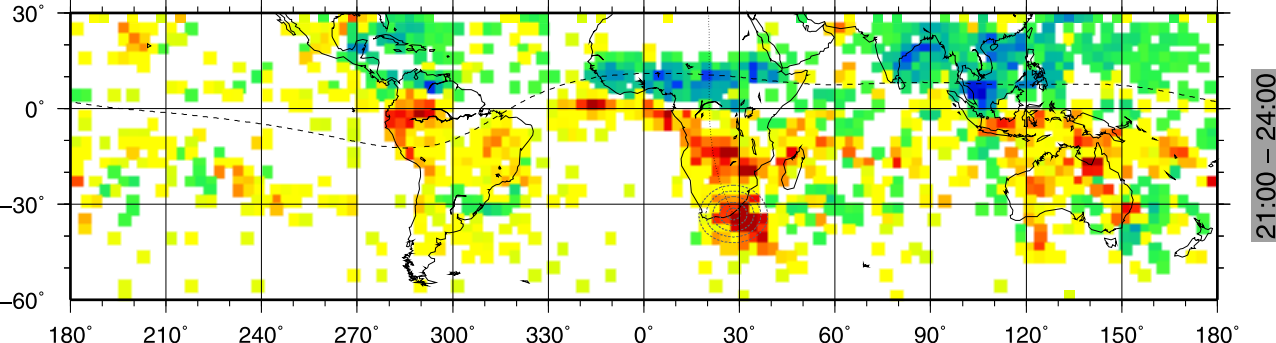












21:00 - 24:00

