Korte, M., Constable, C. (2006): On the use of calibrated relative paleointensity records to improve millennial-scale geomagnetic field models. - Geochemistry Geophysics Geosystems (G3), Vol. 7, Q09004

DOI: 10.1029/2006GC001368.

GEOCHEM. GEOPHYS. GEOSYS., VOL. ???, XXXX, DOI:10.1029/,

On the use of calibrated relative paleointensity records to improve millennial-scale geomagnetic field models

M. Korte[1], C. G. Constable[2]

[1]GeoForschungsZentrum Potsdam, Germany

[2]Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California at San Diego, USA

Abstract. Current millennial-scale time-varying global geomagnetic field models suffer from a lack of intensity data compared to directional data, because only thermoremanently magnetized material can provide absolute information about the past field strength. The number of archeomagnetic artefacts that can provide such data diminishes rapidly prior to 3000 BC. Sediment cores provide time series of declination and inclination and of variations of magnetization: the latter can reflect relative geomagnetic field variations if suitably normalized. We propose a calibration technique based on predictions from global models and use the CALS7K.2 model to calibrate relative paleointensity records from 22 globally distributed locations and assess whether they reflect actual field variations. All except a few contain useful information for 0 to 7 ka and could be used to improve the existing models. Using synthetic data from a numerical dynamo simulation we show that with the existing directional data the distribution of intensity data has an important influence on model quality. Intensity data from a broad range of latitudes seem particularly important. This study opens the possibility of extending global time-varying geomagnetic field models further back in time than the current 7 kyr interval.

1. Introduction

Our knowledge of long-term geomagnetic field evolution is mainly obtained from indirect measurements using the remanent magnetization of archaeological material, lava flows or sediments. While direct observations of the magnetic field directions declination and inclination go back several centuries, a method to determine the absolute intensity was only developed in 1832 by Gauss (see *Jonkers et al.* [2003] for an overview).

Similarly, but for different reasons, it is easier to obtain directional information from remanently magnetized material than absolute intensities. Starting from the initial *Thellier* [1941] technique several methods have been developed to determine absolute paleointensities by comparison between thermal demagnetization of the natural remanent magnetization (NRM) and acquisition of thermoremanent magnetization in a known laboratory field (see *Valet* [2003] for a review). Obviously these methods work only for material carrying a thermoremanent magnetization, i.e. archeomagnetic materials and lavas, but not sediments. The remanent magnetization carried by detrital grains is generally supposed to be a linear function of the ambient field but is influenced by lithological factors such as grain size and concentration of the magnetic material in the sediment, as well as by properties of the non-magnetic matrix [*Levi and Banerjee*, 1976; *King et al.*, 1983; *Tauxe*, 1993; *Tauxe et al.*, 2006]. The intensity of magnetization must be appropriately normalized to mitigate these other influences. The resulting record can re-

flect the field variations, but not the absolute values of past geomagnetic field intensity.

Attempts have been made to calibrate sedimentary intensity records by using absolute archeomagnetic intensities intensities, including efforts involving nearby paleointensity observations to global virtual axial dipole moment results [e.g. *King et al.*, 1983; *Constable*, 1985; *Peck et al.*, 1996; *Brachfeld and Banerjee*, 2000]. The global distribution of currently available absolute paleointensity results, however, is very uneven and extremely sparse for the southern hemisphere [*Korte et al.*, 2005], and the influence of field morphology means that the suitability of comparisons decreases with increasing distance.

Recently a global field model spanning the past 7000 years has been developed [*Korte and Constable*, 2005], based on archeomagnetic and lake sediment directional data and archeomagnetic absolute intensities [*Korte et al.*, 2005] and using the same modeling technique as applied to recent direct field observation data [see *Bloxham and Jackson*, 1992; *Korte and Constable*, 2005]. This model, named CALS7K.2 (Continuous model from Archeomagnetic and Lake Sediment data of the past 7 k years), has limited spatial and temporal resolution compared with models covering only recent decades or centuries but can provide predictions of general field evolution from 5000 BC to 1950 AD for any location. However, the accuracy of specific predictions varies due to the distribution and quality of the underlying data. More than 13000 declination and inclination data respectively have been used for

CALS7K.2, but only 3092 absolute intensity data were available, with a large fraction coming from Europe and no more than 165 from the Southern hemisphere.

We investigate the possibility of using CALS7K.2 as a calibration tool for relative paleointensity data and determine whether a constant scaling factor at each location is sufficient for calibration to absolute intensity. We assess the quality of the sediment intensity records as a representation of magnetic field changes and check for potential improvements to millennial scale geomagnetic field models by including the calibrated intensity data from 22 locations. In section 6 we test the importance of good intensity data distributions, using synthetic data generated by a numerical dynamo simulation.

2. Relative paleointensity data

Several authors made their relative sedimentary intensity results available for this study, providing us with a reasonable world-wide data distribution. Table 1 lists locations, references and abbreviation codes, which we will use in the following to address individual data series. Figure 1 displays the locations on a global map.

We briefly recall some significant limitations of paleomagnetic data in general and sedimentary intensity records in particular. Inferring paleointensity from sediments is not straightforward. The magnetization of the sediment in general tends to reflect lithology and volume of magnetic material rather than variations in the geomagnetic field. *King et al.* [1983] and *Tauxe* [1993]

established minimal criteria they considered necessary before a sediment paleointensity record can be supposed to reflect actual field variation. These criteria are uniform magnetic mineralogy with homogeneous concentration and grain sizes, uncomplicated behaviour during demagnetisation to yield good directional records, no coherence of paleointensity with bulk magnetic parameters and agreement among estimates of paleointensity using different normalizations. The data series used in this study are quite heterogeneous, as they are drawn from work by numerous different authors and sites with a broad range of lithologies. In some but not all of the studies the limitations are well documented, and in general we can expect that these minimal criteria are met to different degrees by the individual records. It is also important to determine whether these criteria provide a genuinely useful guide in selecting appropriate material for sedimentary paleointensity studies [*Frank et al.*, 2003]. The normalization for magnetic material concentration variations is commonly done by normalizing the NRM with ARM, but occasionally IRM or SIRM are used, see Table 1. Ideally the parameter used for normalization should excite the same spectrum of magnetic particles as participate in the NRM, and this may not correspond to just one of ARM, IRM, or SIRM, because of their different sensitivities to particular grain sizes. Some authors [*Brachfeld and Banerjee*, 2000; *Frank et al.*, 2003] have proposed a secondary normalization as a possible means to compensate magnetic grain size variations with time. Other factors associated with the non-magnetic matrix have also been shown to be important in determining the quality of

paleointensity records in some circumstances [e.g. *Barton et al.*, 1980; *Katari et al.*, 2000; *Katari and Bloxham*, 2001; *Tauxe*, 1993; *Tauxe et al.*, 2006].

A second problem is determining the age of sediments, and when the magnetization was acquired. The most commonly used method is radiocarbon (^{14}C) dating. Radiocarbon ages can be quite uncertain due to carbon reservoir effects, and have to be calibrated to calendar ages. Fourteen of the sediment records we used are radiocarbon dated, and ten of these came with uncalibrated ages. We calibrated those (BAI, BAR, EAC, LEB, ON1 to ON4, PAD, PEP) using the CALIB program by *Stuiver and Reimer* [1993]; *Stuiver et al.* [1998] with the same parameter settings used for the sedimentary records providing directional data for CALS7K.2, see *Korte et al.* [2005] for details. For Lake Pepin (PEP), this calibrated age scale differs from that given by *Brachfeld and Banerjee* [2000]. Their calibrated ages are tuned to the ages determined for another North American lake by comparing features of inclination, a method known as paleomagnetic dating. Paleomagnetically dated records are not suitable for global field modeling because the dating is not independent from secular variation and field morphology and consequently we preferred to use an independent (although not necessarily better) age model. For this calibration study, however, independent age control is not a stringent requirement and we also included two geomagnetically dated records (BYE, LAR) for which no reasonable independent age models were available. The Bjorn Drift and Gardar Drift records (BJO,GAR) were dated by correlating the oxygen isotopic record (δ^{18}O) to a reference signal given in

uncalibrated ^{14}C ages. Consequently we had to calibrate the ages given for BJO and GAR like the other radiocarbon ages. Finally, four of the records (FRG, NAU, POH, SAR) had been dated by varve-counting by their authors, a method that we supposed to be more accurate than radiocarbon dating. Information on the dating method is included in Table 1.

We aim to study variations on the centennial to millennial scale, which means that high temporal resolution of the paleomagnetic record from the sediments is required. Although post-depositional effects may also play a role this is mainly determined by the sedimentation rate. Information about approximate sedimentation rates is included in Table 1, where a single number means a relatively constant sedimentation rate and a range of values hints at changes of sedimentation rate with time. These rates are approximate, because this information was sometimes difficult to extract from the publications, and in some cases they are based on uncalibrated and in others on calibrated ages. The difference is unlikely to exceed 10% for these data, which is not significant in the context of evaluating the expected overall resolution of the records. However, the variability of sedimentation rate down core is often unknown. Dating by varve counting can provide a good record of variations of accumulation rates over time, but with radiocarbon dating the sedimentation rates are usually interpolated between a number of age tie points. The number is included in Table 1 to give an indication of the reliability of inferred ranges in sedimentation rates. When there is only one age tie point over the time interval of interest a constant sedimentation rate

is assumed, rather than confirmed. In cases with several tie points several methods are used to construct an interpolated age model. Linear or low degree polynomial interpolation (the latter often over longer time spans than studied here) are most commonly used for these data sets. In lake sediments covering the two most recent centuries and having good sedimentation rate control, this rate often rises significantly at about the 18^{th} or early 19^{th} century.

The last column of Table 1 contains brief information about the method used in determining paleomagnetic results along the sediment core. When u-channels are passed through a magnetometer, this can provided a high spatial density of measurements, but very dense measurements are not independent because the response functions of the magnetometer coils have an effective width of several centimeters. Results based on u-channels are thus smoothed compared with those based on individual samples and edge effects must be considered at breaks in the core and near the ends. As we will discuss below, the PEP data are affected by such edge effects.

3. Calibration of relative intensities

We adopt a straightforward approach to calibrating the individual time series at each location:

1. Compare relative intensities F_{dat} to model predictions F_{mod} at the same locations and times and calculate the ratio F_{mod}/F_{dat} for each individual datum.

2. Determine the median ratio F_{mod}/F_{dat} for the whole time series and multiply the relative intensity data by that scaling factor.

3. Compare this calibrated time series with the model predictions.

The success of this approach will of course depend on the quality of both the relative intensity time series and the model predictions and it is important to make a careful assessment of the results. For an ideal relative intensity record and completely accurate model we expect the ratio F_{mod}/F_{dat} to be constant over the whole time span, but the two examples shown in Fig. 3 clearly indicate that this is not the case: for the ESC record, which is the most extreme example, the ratios varies by more than a factor of two. We need some guidelines for evaluating whether these deviations from a constant factor are compatible with the particular record providing useful information about paleointensity information.

From a geomagnetic point of view we should expect axial dipole (g_1^0) variations to be the same at all locations (apart from the scaling factor), and we might expect that these variations will dominate the long period signal in all the records (note that this does not exclude the possibility that they also contribute to the record at shorter periods where their coherent variations may be masked by non-axial-dipole contributions). We can test this by looking at intensity predictions from CALS7K.2 for all of our sediment data locations. Figure 2 shows that the long term trend is broadly similar on several thousand year time scales, producing about a factor of two variation over 7 kyr , but at shorter periods the predicted signal is quite variable

from one region to another. Thus the short term (1 kyr or less) deviations in F_{mod}/F_{dat} could be a reflection of inadequacies in the model especially in regions with few or no data. Large variations on an intermediate time scale must be a cause for some concern, but we cannot rule out a priori the possibility that they too might arise from deficiencies in CALS7K.2. Other factors that might affect F_{mod}/F_{dat} were mostly discussed in the previous section, and are (a) large but essentially uncorrelated errors in relative paleointensity estimates – this is what our modeling assumes, and some outliers in individual records look like this (e.g. EAC); (b) mismatches in age scales, including general uncertainties, temporal changes in sedimentation rate, or hiatus in sedimentation – these generally give rise to temporally correlated uncertainties, and could cause model mismatches on timescales of interest for field modeling; (c) inadequate normalization and/or change in grain size and/or mineralogy on time-scales which could well be the same as those of interest for the field – these errors are also temporally correlated.

Visual inspection did not reveal very long-term trends in the ratios F_{mod}/F_{dat} for any of the time series, with the possible exceptions of the earliest 3kyr of ON1 to ON4, all lying within a few hundred km of one another: these data have very low temporal resolution, making it difficult to determine whether they are internally consistent. We plotted histograms of the scaling ratios for all the data series and found that they all are nearly normally distributed: the remainder were slightly long-tailed towards larger ratios. Our representative examples of the estimated ratios through time in

Fig. 3 show a large dynamic range, but the absence of long term trends is reassuring. The short term deviations up to a few hundred years might be expected from the quality of both the data and the model and we will see later that when included directly in the modeling the misfits obtained for them are comparable to those for archeointensity data used in CALS7K.2.

Table 2 lists the median scaling factor and its standard deviation for each site. The large range in scaling factors, reflecting very large differences between the relative intensity values from different sites, is to be expected from differences in sedimentary environments and normalization strategies. Each lake can be expected to have a different scaling factor because of its unique lithology. The same is true for the marine sediments considered here which are sufficiently distant from one another to have distinct sedimentary environments. Thus the differences in factors among sites ON1 to ON4 as well as BJO and GAR are unremarkable.

The standard deviation in the scaling within each record mainly ranges from 14 to 37 %, with the four exceptions of LAR (54 %), ON2 (55 %), ON4 (63 %) and EAC (75 %). LAR and ON1-4 have poor age control, and for EAC the reason is some strong outliers in the record. We did not remove those a priori, because data rejection purely on the basis of scattered data could be quite subjective. The median should not be affected by them and in the modeling described below the iterative rejection of outliers will be able to deal with them properly. Fig. 4 displays the comparison between

model predictions and the paleointensity sediment time series calibrated by following the three steps described above.

Long-term trends in data and model predictions agree in general. The agreement at shorter periods is more variable, but although there are significant deviations, they mostly are of the same order as those observed between the archeomagnetic intensity data used in constructing the model and CALS7K.2 predictions. The root mean square (rms) misfit between all scaled sedimentary intensity values and CALS7K.2 is 14.3 μT. This is slightly larger than both our averaged estimated uncertainties for absolute archeo-intensities (11 μT) and the rms misfit between CALS7K.2 and all archeo-intensity data (10.8 μT), but several of the rms misfits for individual locations (included in Table 2) are below those values. Figure 5 shows that the distribution of residuals has the same general characteristics for both archeomagnetic and scaled sedimentary intensities with many very closely fit values and a large number of outliers compared to a normal distribution. Strong outliers from a normal distribution occur mainly on the positive side for the calibrated intensity data, but the residuals are only slightly biased, the average is 1.8 μT, significantly less than for the archeomagnetic data (3.6 μT).

As with the absolute intensity data used in CALS7K.2 there are some times and places where the model disagrees with the observations, which will be discussed in detail in section 5.

4. Including calibrated intensity values in global modeling

In using CALS7K.2 to calibrate the relative paleointensity data we must also keep in mind that any disagreement between model predictions and calibrated intensity data might be caused by shortcomings of CALS7K.2. The results presented in Fig. 4 led us to explore the possibility of using the calibrated relative intensity data in future updates of CALS7K.2. We included the calibrated sedimentary intensity data in the modeling process and studied the effect on achievable fit to both intensity and directional data and differences between the models. We have directional data from most of the sedimentary records, but for CALS7K.2 only the directional data from BAI, BAR, EAC, LEB, PAD and POH were used. In order to study purely the influence of the calibrated intensity records we did not include the additional directional data in the following. The resulting numbers of data are given in Table 3 where the calibrated sedimentary intensities account for 5848 of the intensity data.

For the modeling process we needed to assign error estimates in accordance with those used previously in constructing CALS7K.2 (see *Korte et al.* [2005]; *Korte and Constable* [2005] for details). Following that scheme for age errors, we assigned 5 years to intensities dated by varve counting and the uncertainties obtained from the CALIB calibration process or 25 years to calibrated radiocarbon dated intensities. The geomagnetically dated intensities, which should not be used in improving the model, were assigned age uncertainties of 51 years. Note, that these dating error estimates are mostly

not reasonable uncertainty estimates for individual data (which might be much larger), but used in the same way as for the original CALS7K model to obtain consistent weighting of the data according to their dating uncertainties. The data uncertainties themselves are hard to assess. Nine of the relative intensity records (BAI, BIR, ESC, ON1 to ON4, PEP and STL) came with uncertainty estimates. When calibrated like the data, their averages range from 1.7 to 8.1μT. For this study, we chose to use the average, 4.6μT, as uncertainty estimate for all our calibrated intensity data. Our current modeling method does not allow us to take age uncertainties into account directly, instead we estimate their effect with respect to uncertainty in the field value based on secular variation studies. That way, the age uncertainties combined with the above data uncertainties result in an average uncertainty estimate of 6.0μT for sediment intensity data, clearly lower than our previous uncertainty estimate for archeomagnetic intensity of 11.0μT. This might not be reasonable for improving the model, but for the current purpose of testing the influence of the scaled intensity an increased weight for those data is perhaps not unreasonable.

The model consists of regularised spherical harmonic coefficients with the temporal continuity described through cubic B-splines. We followed our previously used modeling scheme of iterative data rejection and applied the same criteria for choice of model parameters, see *Korte and Constable* [2005] for details. Spatial and temporal regularization factors and the resulting

norms are given in Table 4, which compares directly to Table 4 in *Korte and Constable* [2005].

Figure 6 shows the fit of the modified model to the calibrated sediment intensity data. The rms misfit to those 5848 data is now reduced from 14.3 to 9.7μT, while the overall fit to all data used in constructing the model has hardly changed. The fact that the normalized rms misfit of all intensity data was smaller than 1.0 in CALS7K.2 and is larger now reflects the significantly lower uncertainty estimates for intensity in the large number of sediment intensity data.

The calibrated sedimentary intensity records more than double the number of intensity data available for global modeling. Changes to the resulting model mainly occur where previously there were no observations or where the new intensity data provide conflicting information.

The comparison of Fig. 4 and Fig. 6 shows how much we can improve the fit of the model to the new intensity time series by including them in the modeling. While some data are fit significantly better (e.g. FRG, PAD), some strong variations in others are not (e.g. LEB, BJO). This improved fit does not significantly change the fit to the previously used data, confirming the internal consistency of the whole data set. Figure 7 compares the fit of CALS7K.2 and the new model to some representative examples from the original data set. There is hardly any visible difference in most cases, and these will be disussed in more detail in the following section.

Repeating the calibration of the original relative sedimentary data with the new model shows that the scatter in individual scaling factors could in general be lowered by a few percent (see Table 2). The reduction in scatter, however, is quite variable and it is not systematic in a way that for example the strongest reduction can be achieved for the data with the largest scatter. Indeed, the strongest reduction in the scatter arises for regions where little or no (intensity) data had been available for CALS7K.2.

5. Data quality and model sensitivity

To gain better understanding of the quality of sedimentary intensity data and the sensitivity of the spherical harmonic models we look in detail at the individual records and how they are fit by the models in this section. We do so roughly by geographic region, as the records are ordered in the panels in Figs. 4 and 6. It starts with North America (PEP to STL) and the northern Atlantic (BJO, GAR) at the top of the left column, Scandinavia (SAR to BYE) and below southern Europe (MEZ). The right column from top to bottom covers the Near East (BIR) and Siberia (BAI), and then the southern hemisphere with the Indonesian - Australian region (ON1 to EAC), southern Argentina (ESC) and Antarctica (LAR, PAD).

5.1. North America to Northern Europe

Steep drop-offs and gaps in the PEP record occur at breaks in the core due to an edge effect of the pass-through magnetometer measurements, but the agreement between data and model predictions for the individual pieces

is otherwise quite remarkable for most of the time interval. This is a region where directional, but no archeomagnetic intensity data had been used in CALS7K.2, yet much of the multi-centennial scale variations seen in PEP and LEB are already roughly fit by CALS7K.2. Clear improvements are made by the modified model, which is reflected both in fit to the data and reduction in scatter of the calibration factor. The two records agree well, and the high intensities between 800 BC and 0 AD are fit better by the modified model, although it cannot fit the full amplitude of this structure seen in the LEB data. As these two locations are close together and the fit to the third northern American record, STL, also is quite good, this might suggest that the normalisation for the LEB data did not work adequately at that time. The sedimentation rates are comparable for these three lakes and the age models seem to work well, even for PEP which is only confirmed by one age tie point prior to 1830.

The BJO and GAR records, which are rather close together, show little agreement in their clear variations and hardly any of these variations are fit by CALS7K.2. Although a reduction in misfit and scatter of scaling factor can be achieved by the modified model for both records, the visual comparison of data and model in Fig. 6 reveals that the fits are still unsatisfying. No archeomagnetic intensity and rather few directional data from Iceland have been used in the models, and it is impossible to decide whether either of these sediment records reflects actual intensity variations reliably. Comparisons among the two records and the model predictions suggest that the

disagreement is not simply a problem of dating, even though this would be a likely explanation as neither record has good Holocene age control.

Four of the five Scandinavian lake sediments are varved, and they agree very well both with respect to time and the rather weak intensity variations seen on the sub-millennial scale in that region. Misfit and scatter of scaling factor are already low when calibrated with CALS7K.2, so there is only slight improvement with the modified model. It is noteworthy that for three of these sites we simply combined all the results from two separate cores each, which obviously did not reduce the quality of the records. The geomagnetic dating used for BYE is supported by this good fit. The modified model tries to fit the low intensities seen in POH and BYE between 0 and 800 AD better, but interestingly at the cost of slightly worsening the fit to the POH inclination data in that interval (see Fig. 7). Together with the fact that this structure is not seen clearly in SAR, FRG or NAU this might suggest a problem with normalisation of the POH intensity data, which, however, leaves open the question of why a similar structure is seen in the BYE data.

5.2. Southern Europe, Near East, Siberia

MEZ and BIR are the only sediment intensity records that lie near regions with notable amounts of archeomagnetic intensity data. A reasonable fit is given for MEZ both by CALS7K.2 and the modified model except for the 0 to 1000 AD time interval, but no improvements are gained by the modified model in misfit or scatter of scaling factor. For BIR, on the other hand, the visual fit to both models is quite unsatisfactory, also with hardly any

change in the high misfit and scatter of scaling factor with the modified model. *Frank et al.* [2003] state that the standard criteria for relative intensity determination from sediments are not met by the BIR record and they applied a second normalisation to remove grain size effects in the sediment record. Despite this, the modified model is unable to improve the fit to the high intensities between 4500 BC and 3200 BC, indicating that at least in that interval the BIR intensity data are incompatible with other data from the same general region. With respect to the shorter period variations, e.g. about 0 AD, we have to note that the variability of intensity (and directions) in the European-Near Eastern region seems to be fit insufficiently by the models, as seen in the Bulgarian data in Fig. 7.

With BAI, the modified model gives a slight improvement to misfit and scatter of scaling factor, but some visual disagreements seen in CALS7K.2 remain even though no other intensity data exist from that region. In this case it looks like some shifts (to account for dating errors) or compressions/expansions (to account for varying sedimentation rate) to the time scale might improve the fit between data and models. We drew a similar conclusion from the directional data at that site [*Korte and Constable*, 2005], but it is not obvious whether the modifications to the age model suggested from the comparison of the different components to the model predictions agree. A more detailed investigation will be necessary to determine if an adjustment to the age model of BAI is warranted based on the global models.

5.3. Equatorial Pacific

With only one result every 500 years the data from ON1 to ON4 are clearly at the limit of what is useful for models trying to describe centennial variations. However, it would be useful to calibrate intensity data with sedimentation rates as low as these, which generally go further back in time and thus could offer the possibility to extend the current millennial scale models to even longer time scales. The four rather closely adjacent records show significant differences in scatter in scaling factor and agreement with CALS7K.2. Although regional differences in the geomagnetic field which are not resolved by the model might play a role, it rather seems that in this case it is possible to judge the quality of the records from the comparison to the model. The modified model clearly reduces misfit and scatter of scaling factor for ON2, but the same cannot be achieved for ON4. Slight improvements are obtained for ON1 and ON3, which already agreed better with CALS7K.2. The inclusion of records with sedimentations rates as low as these cannot improve the current models significantly, but neither does it impair the models.

5.4. Southern hemisphere

The high values of scatter in scaling factor and misfit for EAC are caused by several outliers in the data. The visual comparison reveals a reasonable fit between CALS7K.2 and both the EAC and BAR data, with a few multi-centennial variations already described in that model although virtually no southern hemisphere absolute intensity data had been available for the con-

struction of that model. Directional data from these two sites have been used in CALS7K.2 and little change is seen here in the modified model.

The data situation is similar in South America, where three records of directional lake sediments from almost the same location as ESC had been used. However, here the variations in ESC are not fit by CALS7K.2 but the fit improves slightly in the modified model, interestingly going along with slight improvements of fit to the directional data (see MNT declination at about 4400 BC and inclination at about 2500 BC in Fig. 7). Once again some slight disagreement in time between similar variations in data and model suggests some incoherency in time scales, although the ESC data seem comparatively well-constrained with 5 age tie points.

The intensity variations seen in the model are clearly similar at the location of PAD further south, and there is very good agreement in particular between the modified model and these data. This is all the more remarkable as PAD shows significant changes in sedimentation rate based on 8 age tie points over the past 7 kyrs, and the sediment is not magnetically uniform so that *Brachfeld et al.* [2000] themselves stated that they are hesitant to interpret the NRM/SIRM record as a relative paleointensity record. However, the strong maximum of intensity in PAD at 1000 AD (Fig. 6) which also is fit remarkably well, is probably a spurious effect, caused by inadequate normalization for varying sediment properties. The fact that the model shows too high a maximum at the site of LAR (and probably also ESC, where no data are available) in the vicinity at this time and that the fit to inclina-

tion at site PAD itself is worse (Fig 7) supports that suspicion. The fact that this probably spurious structure is strongly represented in the model most likely results from the scarcity of data (particularly intensity) in the Southern Hemisphere.

The record of LAR, near to PAD, however, does not agree with CALS7K.2. Although a reduction of scatter in scaling factor is achieved by the modified model the misfit increases. However, the disagreement might again be a dating problem. One could imagine that with the data from around 3000 BC shifted to older ages and the the variations between 1000 BC and the top of the core compressed somewhat there could be reasonable agreement between data and model. As for BAI, a detailed comparison of the directional data could help to decide whether this is a problem of dating or sediment properties, but this is beyond the scope of the current work.

5.5. Summary

We can draw some general conclusions from all the individual observations described above: Data which are coherent over comparatively large areas are most easily satisfied with existing strategies for global millennial scale modeling, which seek to recover robust large scale field structure. Thus global millennial scale magnetic field models can give useful hints about the quality of sedimentary relative intensity records, by facilitating the identification of records which are either partially or totally incoherent with other regional data. This can be true even when no absolute intensity data or even not much data at all had been available from the region in question before. Par-

ticularly when the directional records can also be taken into account, it can be possible to distinguish between problems with the time scale or sediment properties unsuitable for recovering actual magnetic field variations. However, as most of the other data used in constructing the global models are also affected by large uncertainties, both in measurements and dating, the models cannot always provide a decisive answer on the quality of individual data sets. The inclusion of sedimentary intensity data can improve global models but the suitability of the records should ideally be established beforehand. Although the data used here are based on different methods, cover a range of sedimentation rates and differently constrained age models, most of them proved to provide useful paleointensity information for the time period of interest.

6. Influence of the intensity data distribution

Early global modeling efforts were based only on directional data, with a constraint on the axial dipole evolution to compensate for the missing absolute field strength information [*Constable et al.*, 2000; *Korte and Constable*, 2003]. This approach was based on the fact that for a field with two and only two poles and an infinitely dense global coverage of accurate directional data the field can be reconstructed accurately except for a scaling factor [*Hulot et al.*, 1997]. With CALS7K.2 directional coverage, our results suggested that in many areas intensity data mainly serve as a scaling factor. This raised the question of how important it is to have good global coverage with intensity data. One might suppose that good temporal coverage is most

important for intensity, that the spatial distribution of data is less critical, or that data from even a small region would be adequate. We investigate this in a simple test of using synthetic data from a magnetic field model produced by a geodynamo simulation.

6.1. Modeling of synthetic data

We use a numerical model developed by *Wicht* [2002]. Dynamo I, a simulation showing reversals, is driven by an imposed temperature difference across the outer core and its parameters are: Rayleigh number $Ra = 810$, Ekman number $E = 3 \times 10^{-4}$, Prandtl number $Pr = 1$ and magnetic Prandtl number $Pr_m = 3$. Although numerical models in general still differ in a number of ways from data-inferred geomagnetic field behavior, we suppose that the characteristics are similar enough for this test. One time step of the model can be considered roughly equivalent to 10 years, so we used 696 time steps from a period of stable polarity to represent the 6950 years from CALS7K. A comparison of the average spectral distribution of Dynamo I from within this time interval to current geomagnetic models reveals that there is less power in the lower spherical harmonic degrees seen at the Earth's surface. By multiplying the axial dipole coefficient of Dynamo I by a factor of 2 the spectrum becomes quite comparable to recent ones (see Fig. 8), suggesting that there will be a comparable amount of spatial structure, which is the important criterion for our test.

Initially, six synthetic datasets were compiled from the model predictions of Dynamo I with this boosted axial dipole. All contain the distribution of

directional data which was available for CALS7K.2, i.e. a realistic, inhomogeneous spatial and temporal distribution. Various intensity distributions, labeled D1–D6 were investigated, and these are described in Table 5. All have approximately the same number of data as used in CALS7K.2 except for D6 where complete vector information is included for all CALS7K.2 locations. Several modifications of datasets D3 and D4 centered on different longitudes or latitudes, respectively, were used to check the sensitivity of the results. These modifications of D3 and D4 allowed us to assess the influence of hemispheric and zonal asymmetries in the numerical simulation results.

All datasets were modeled with the technique used for CALS7K.2. Spatial and temporal regularization parameters were chosen in each iteration to minimize the root mean square (rms) misfit to the data and according to the criteria used for CALS7K.2. The synthetic data do not contain uncertainties, so we are able to fit them very well.

6.2. Results

The differences to the models caused by the intensity data distribution is studied by comparing how well the original spatial power spectrum can be reproduced. As a general guide, Fig. 8(a) and (b) show how the surface and core spectra of the time-averaged CALS7K.2 spectrum compare to the full resolution attainable with the modern POMME3.0 satellite model of *Maus et al.* [2006]. Temporal averaging reduces the overall power (generating a systematic offset), but the important differences for our purposes arise from the systematic deviations in trend above degree 4 that reflect lack of spatial

resolution in CALS7K.2. Figure 8(c) and (d) show the time-averaged spectra of models from datasets D1 to D6 compared to the spectrum of the original numerical model. The overall CALS7K.2 data distribution synthesized in dataset D1 performs much better than the CALS7K.2 model because there is no noise in the synthetic data. As expected, the better distributions of intensity data (D5 and D6) resolve the most structure, and the very localized intensity data in D2 performs worst. The poor performance of D2 is more or less independent of where the patch of intensity data is located.

We were initially surprised to find a strong contrast in the results from D3 and D4. D3, with intensity data from all latitudes but a single longitude, reproduces the original model almost as well as with good global intensity distributions (D5 and D6), while D4 with a single latitude and complete range of longitudes performs as poorly as D2.

Although the good results with D3 might be rationalized in terms of latitudinally distributed intensity data resolving the dominantly dipolar field structure we found that the initial results in Figure 8 were somewhat fortuitously good. Fig. 9 shows the results of similar experiments along different meridians, and these show a substantial amount of variation, indicating that 7kyr is not long enough to average to a zonal field. In this sense the structure of the numerical model is not dissimilar to the Earth's magnetic field. However, Fig. 10 confirms, for a variety of small circles, that D4-type intensity data from a single latitude gives poor agreement with the original spectrum, and that the performance in this case is generally worse than when data like

D3 are available. We could not obtain a more satisfying reconstruction than from D4 with intensity data coming only from any other longitudinal circle. The spectra from two of the other examples shown in Fig. 10 agree closely with that of D4, while two others significantly overestimate the spatial structure. In these two, the intensity data come from the 45°S (circles) and additionally the latitudes of the directional data have been reversed (diamonds). The overestimation must be a consequence of a more complex structure of the numerical model field in the southern than the northern hemisphere, again structure that can be regarded as similar to the real geomagnetic field. In this simulation with no data uncertainties our criteria for damping lead to some spurious spatial variation in areas with less field structure and less data in the case of dense sampling of areas with complex structure. Similar overestimations do occur with simulated data from all latitudes (see circles in Fig. 9), but they are by far less strong.

We conclude that a good latitudinal distribution of intensity data is important to improve current millennial scale global models. However, even in the best case these results also show that without a significant improvement in the global distribution of directional data we will not be able to obtain full spatial resolution beyond spherical harmonic degree 6.

7. Conclusions

We have shown that the global geomagnetic field model CALS7K.2 can be a useful tool for calibrating relative paleointensity data. There is generally reasonable agreement between model predictions and sedimentary intensity

data, especially for long-term variations. Short term variations show occasional systematic departures from the model, but are nevertheless fit at about the same level as archeointensity and directional data in the original model. This lends credence to both the quality of the model and the reliability of relative paleointensity results, but as always the quality of individual data sets is best evaluated by their consistency with other data from the same general region. Problematic sediment intensity results can best be identified when even a model including those data cannot achieve an improved fit to them. Of the 22 records used in this study most reflect field intensity variations rather well, with the possible exceptions of BJO, GAR, BIR, LAR, and the top part of PAD. It is encouraging to note that these could also have been identified a priori as potentially suspect based on their poor age control (BJO, GAR, LAR) or variability in sediment properties (BIR, PAD). The rms misfits between model predictions and the calibrated relative intensity data series studied here are similar to the average error estimates for the archeomagnetic intensity data used to construct the global model. Inclusion of calibrated sediment intensity records in the modeling improves their fit to the model, while mostly not affecting the agreement with the previously used data. The most obvious changes to the model occur in regions that previously lacked adequate data. It is clear that where no reasonable constraint is provided by directional or other intensity information the time variations represented in the calibrated intensity data do more than just supply an overall scaling factor for the model. Slight improvements to intensity

even in reasonably well-constrained regions, however, indicate that with the currently available distribution and quality of directional data the intensity information in general cannot be regarded as just a scaling factor.

Constant calibration factors were applied to all relative paleointensity time series used in this study, and indicated that suitably normalized records generally reflect actual field variations and are not dominated by lithological effects. In a few cases variations in the calibrated intensity records remain incompatible with the updated global model of directional and intensity data. These disagreements are no worse than misfits due to regional incompatibility of data in the directional or archeointensity records. Using current millennial scale models it is generally not possible to decide whether they are due to lithological variations in the sediments, or features of the field that cannot be fit either due to their local extent and the limited resolution of the model or due to errors in any of the data. The relative paleointensity records we used had very different sediment properties and were distributed worldwide, and we conclude that with a suitable normalization the use of a constant calibration factor is usually sufficient for time series of a few millennia. Although secondary calibrations have occasionally been used it may be difficult to determine whether they can be justified on the basis of improved fit to existing models.

Our results indicate that there are good prospects for improving the CALS7K.2 global millennial scale model by including relative paleointensity data. Re-calibration of the data by the new model shows that the simple

approach of iteratively improving the model by including sedimentary intensities scaled by the initial model is adequate. However, the reliability of sedimentary paleointensity as a reflection of actual field behaviour should be established both by their internal consistency (using rock magnetic tests) and by using a test model to establish whether they are consistent with other observations for the same geographic region. Questionable data can be rejected or down-weighted using the same general strategy as employed for directional and archeointensity data in CALS7K. The test model approach might also be taken to carefully guide revision of the time scales of some sedimentary records (including records providing only directional data). However, we note again that it is important to avoid circularity in the modeling, and that geomagnetically dated records are unsuitable for global field modeling.

The use of sedimentary relative intensity records may eventually make it possible to balance the ratio of directional and intensity data underlying millennial scale models which could be an important factor in improving the model's resolution. Even more intriguing is the possibility of extending global continuous models further into the past; at the moment their time span is limited by the diminishing number of archeomagnetic data with increasing age. Longer sediment records could be calibrated by the factor determined from their overlap with CALS7K.2, yielding the absolute intensity information which is lacking for older epochs. Through a study on predictions from a numerical model we have also shown that with the currently available distribution of directional data it is important to include a broad distribution

of intensity information rather than limiting it to a scaling factor from one region. Good intensity coverage from all latitudes is particularly important for resolving the smaller-scale structure of the field.

Acknowledgements

We thank all the colleagues who made their data available for our study, either personally or by publishing the complete data sets along with their results. The constructive comments of two anonymous reviewers and editor John Tarduno are gratefully acknowledged. The cooperation of the authors was greatly facilitated by support from Alexander von Humboldt-Foundation under the follow-up sponsorship program. CC acknowledges support from NSF grant EAR0337712 and EAR0537986.

References

Barton, C., M. McElhinny, and D. Edwards (1980), Laboratory studies of depositional DRM, *Geophys. J.R. astr. Soc.*, *61*, 355–377.

Bloxham, J., and A. Jackson (1992), Time-dependent mapping of the magnetic field at the core-mantle boundary, *J. Geophys. Res.*, *97*, 19,537–19,563.

Brachfeld, S., G. D. Acton, Y. Guyodo, and S. K. Banerjee (2000), High-resolution paleomagnetic records from Holocene sediments from the Palmer Deep, Western Antarctic Peninsula, *Earth Planet. Sci. Let.*, *181*, 429–441.

Brachfeld, S., E. Domack, C. Kissel, C. Laj, A. Leventer, S. I. R. Gilbert, A. Camerlenghi, and L. B. Eglinton (2003), Holocene history of the Larsen Ice Shelf constrained by geomagnetic paleointensity dating, *Geology*, *31*, 749–752.

Brachfeld, S. A., and S. K. Banerjee (2000), A new high-resolution geomagnetic relative paleointensity record for the North American Holocene: A comparison of sedimentary and absolute intensity data, *J. Geophys. Res.*, *105*, 821–834.

Channell, J. (1999), Geomagnetic paleointensity and directional secular variation at Ocean Drilling Program (ODP) site 984 (Bjorn Drift) since 500 ka: Comparison with ODP site 983 (Gardar Drift), *J. Geophys. Res.*, *104*, 22,937–22,951.

Channell, J. E. T., D. A. Hodell, and B. Lehman (1997), Relative geomagnetic paleointensity and δ^{18}o at ODP Site 983 (Gardar Drift, North Atlantic) since 350 ka, *Earth Planet. Sci. Let.*, *153*, 103–118.

Constable, C. G. (1985), Eastern Australian geomagnetic field intensity over the past 14000 yr, *Geophys. J. Roy. Astr. Soc.*, *81*, 121–130.

Constable, C. G., and L. Tauxe (1987), Palaeointensity in the pelagic realm: marine sediment data compared with archaeomagnetic and lake sediment records, *Geophys. J. Roy. Astr. Soc.*, *90*, 43–59.

Constable, C. G., C. L. Johnson, and S. P. Lund (2000), Global geomagnetic field models for the past 3000 years: transient or permanent flux lobes?, *Phil. Trans. R. Soc. Lond. A*, *358*, 991–1008.

Frank, U., M. Schwab, and J. Negendank (2003), Results of rock magnetic investigations and relative paleointensity determinations on lacustrine sediments from Birkat Ram, Golan Heights (Israel), *J. Geophys. Res.*, *108(B8)*, doi:10.1029/2002JB002,049.

Frank, U., E. Schnepp, N. Nowaczyk, A. Ramrath, and J. F. Negendank (2005), Magnetostratigraphic investigations of Late Quaternary sediments from Lago di Mezzano, Central Italy: rock magnetism and relative paleointensity variations, in preparation.

Gogorza, C. S. G., J. M. Lirio, H. N. nez, M. Chaparro, H. R. Bertorello, and A. M. Sinito (2004), Paleointensity studies on Holocene-Pleistocene sediments from Lake Escondido, Argentina, *Phys. Earth Planet. Int.*, *145*, 219–238.

Hulot, G., A. Khokhlov, and J. L. LeMouël (1997), Uniqueness of mainly dipolar magnetic fields recovered from directional data, *Geophys. J. Int.*, *129*, 347–354.

Jonkers, A. R., A. Jackson, and A. Murray (2003), Four centuries of geomagnetic data from historical records, *Rev. Geophys.*, *41(2)*, doi:10.1029/2002RG000,115.

Katari, K., and J. Bloxham (2001), Effects of sediment aggregate size on DRM intensity: a new theory, *Earth Planet. Sci. Lett.*, *186*, 113–122.

Katari, K., L. Tauxe, and J. King (2000), A reassessment of postdepositional remanent magnetism: preliminary experiments with natural sediments, *Earth Planet. Sci. Lett.*, *183*, 147–160.

King, J. W., S. K. Banerjee, and J. Marvin (1983), A new rock-magnetic approach to selecting sediments for geomagnetic paleointensity studies: Application to paleointensity for the last 4000 years, *J. Geophys. Res.*, *88*, 5911–5921.

Korte, M., and C. G. Constable (2003), Continuous global geomagnetic field models for the past 3000 years, *Phys. Earth Planet. Interiors*, *140*, 73–89.

Korte, M., and C. G. Constable (2005), Continuous geomagnetic field models for the past 7 millennia: 2. CALS7K, *Geochem., Geophys., Geosys.*, *6, Q02H16*, doi:10.1029/2004GC000,801.

Korte, M., A. Genevey, C. G. Constable, U. Frank, and E. Schnepp (2005), Continuous geomagnetic field models for the past 7 millennia: 1. a new global data compilation., *Geochem., Geophys., Geosys.*, *6, Q02H15*, doi:10.1029/2004GC000,800.

Levi, S., and S. K. Banerjee (1976), On the p ossibility of obtaining relative paleointensities from lake sediments, *Earth Planet. Sci. Lett.*, *29*, 219–226.

Maus, S., M. Rother, C. Stolle, W. Mai, S. Choi, H. Lühr, D. Cooke, and C. Roth (2006), Third generation of the Potsdam Magnetic Model of the Earth (POMME), *J. Geophys. Res.*, p. submitted.

Ojala, A. E. K., and T. Saarinen (2002), Palaeosecular variation of the Earth's magnetic field during the last 10000 years based on on the annually laminated sediment of Lake Nautajärvi, central Finland, *The Holocene*, *12*, 391–400.

Peck, J. A., J. W. King, S. M. Colman, and V. A. Kravchinsky (1996), An 84-kyr paleomagnetic record from the sediments of Lake Baikal, Siberia, *J. Geophys. Res.*, *101*, 11,365–11,385.

Saarinen, T. (1998), High-resolution paleosecular variation in northern Europe during the last 3200 years, *Phys. Earth Planet. Int.*, *106*, 299–309.

Schwab, M., F. Neumann, T. Litt, J. Negendank, and M. Stein (2004), Holocene palaeoecology of the Golan Heights (Near East): investigation of lacustrine sediments from Birkat Ram crater lake, *Quat. Sci. Rev.*, *23*, 1723–1731.

Snowball, I., and P. Sandgren (2002), Geomagnetic field variations in northern Sweden during the Holocene quantified from varved lake sediments and their implications for cosmogenic nuclide production rates, *The Holocene*, *12*, 517–530.

Snowball, I., and P. Sandgren (2004), Geomagnetic field intensity changes in Sweden between 9000 and 450 cal BP: extending the record of "archaeomagnetic jerks" by means of lake sediments and the pseudo-Thellier technique, *Earth Planet. Sci. Let.*, *in press*, in press.

St-Onge, G., J. Stoner, and C. Hillaire-Marcel (2003), Holocene paleomagnetic records from the St. Lawrence Estuary, eastern Canada: centennial- to millennial-scale geomagnetic modulation of cosmogenic isotopes, *Earth Planet. Sci. Let.*, *209*, 113–130.

Stuiver, M., and P. J. Reimer (1993), Extended ^{14}C data base and revised CALIB 3.0 ^{14}C age calibration program, *Radiocarbon*, *35*, 215–230.

Stuiver, M., P. J. Reimer, E. Bard, J. W. Beck, G. S. Burr, K. A. Hughen, B. Kromer, G. McCormac, J. van der Plicht, and M. Spurk (1998), INTCAL98 radiocarbon age calibration, 24,000-0 cal BP, *Radiocarbon*, *40*, 1041–1083.

Tauxe, L. (1993), Sedimentary records of relative paleointensity of the geomagnetic field: Theory and practice, *Rev. Geophys.*, *31*, 319–354.

Tauxe, L., J. Steindorf, and A. Harris (2006), Depositional remanent magnetization: toward an improved theoretical and experimental foundation, *Earth Planet. Sci. Let.*, p. in press.

Thellier, E. (1941), Sur la vérification d'une méthode permettant de déterminer l'intensité du champ terrestre dans le passé, *C. R. Hebd. Seances Acad. Sci.*, *212*, 281–283.

Valet, J.-P. (2003), Time variations in geomagnetic intensity, *Rev. Geophys.*, *41(1)*, doi:10.1029/2001RG000,104.

Wicht, J. (2002), Inner core conductivity in numerical dynamo simulations, *Phys. Earth Planet. Inter.*, *132*, 281–320.

Code	Location	Lat. [deg]	Long. [deg]	Reference	Normalization at [mT]	Dating	Age model tie points	Sedim. rate [cm/kyr]	Method
BAI	Lake Baikal	52.2	106.5	*Peck et al.* [1996][3]	NRM/ARM 10	^{14}C	4	14	samp.
BAR	Lake Barrine	-17.2	145.6	*Constable* [1985]	NRM/ARM 10	^{14}C	11	74	samp.
BIR	Birkat Ram	33.3	35.7	*Frank et al.* [2003]	NRM/ARM[5] 20	^{14}C[4]	2	∼160	samp.
BJO	Bjorn Drift ODP 984	61.4	-24.1	*Channell* [1999]	NRM/IRM 35-45	δ^{18}O/^{14}C	<1[7]	22	u-chan.
BYE	Byestadsjön	57.4	15.3	*Snowball and Sandgren* [2004][1]	NRM/ARM 20	geomagnetic	na	∼43	samp.
EAC	Lake Eacham	-17.3	145.6	*Constable* [1985]	NRM/ARM 10	^{14}C	11	110	samp.
ESC	Lake Escondido	-41.0	-71.3	*Gogorza et al.* [2004]	NRM/ARM 20	^{14}C	5	12-66	samp.
FRG	Frangsjön	64.0	19.7	*Snowball and Sandgren* [2002][1]	NRM/ARM 40	varves	na	40	samp.
GAR	Gardar Drift ODP 983	60.4	-23.6	*Channell et al.* [1997]	NRM/IRM 25-60	δ^{18}O/^{14}C	<1	∼27	u-chan.
LAR	Larsen Ice Shelf	-64.8	-60.4	*Brachfeld et al.* [2003]	NRM/ARM 20	geomagnetic	na	10-23	u-chan.
LEB	Lake LeBoeuf	42.0	-79.9	*King et al.* [1983][1]	DRM/ARM 20	^{14}C	5	∼200	samp.
MEZ	Lago di Mezzano	42.6	11.9	*Frank et al.* [2005]	NRM/ARM 20	^{14}C	5	45-270	samp.
NAU	Nautajärvi	61.8	24.7	*Ojala and Saarinen* [2002]	NRM/ARM 20	varves	na	66	samp.
ON1	Ontong-Java Pl. ERDC 83Bx	-1.4	157.3	*Constable and Tauxe* [1987]	NRM/ARM	^{14}C	<1	2.2	samp.
ON2	Ontong-Java Pl. ERDC 92Bx	-2.2	157.0	*Constable and Tauxe* [1987]	NRM/ARM	^{14}C	<1	2.5	samp.
ON3	Ontong-Java Pl. ERDC 102Bx	-3.6	161.3	*Constable and Tauxe* [1987]	NRM/ARM	^{14}C	<1	1.7	samp.
ON4	Ontong-Java Pl. ERDC 120Bx	0.0	158.7	*Constable and Tauxe* [1987]	NRM/ARM	^{14}C	<1	2.5	samp.
PAD	Palmer Deep	-64.9	-64.2	*Brachfeld et al.* [2000]	NRM/SIRM 20-50	^{14}C	8	90-600	u-chan.
PEP	Lake Pepin	44.4	-92.1	*Brachfeld and Banerjee* [2000]	NRM/ARM 10-40	^{14}C	1	150[6]	u-chan.
POH	Pohjajärvi	62.8	28.0	*Saarinen* [1998]	NRM/ARM 20	varves	na	100	samp.
SAR	Sarsjön	64.0	19.6	*Snowball and Sandgren* [2002][1]	NRM/ARM 40	varves	na	∼40	samp.
STL	St. Lawrence Est.	48.6	-68.6	*St-Onge et al.* [2003][2]	NRM/ARM 20-40	^{14}C	5	120-180	u-chan.

Table 1. Relative paleointensity records used in this study.

[1] both published core results combined; [2] core 2220 only; [3] data not directly displayed in paper, by pers. comm. from author; [4] updated dating from *Schwab et al.* [2004] with shift by 600 yrs to correct for hardwater effect [*pers. comm. U. Frank*]; [5] additional normalisation for grain size correction used; [6] strong increase in sedimentation rate from 1830 to present; [7] <1 means that the age tie point is older than 7ka;

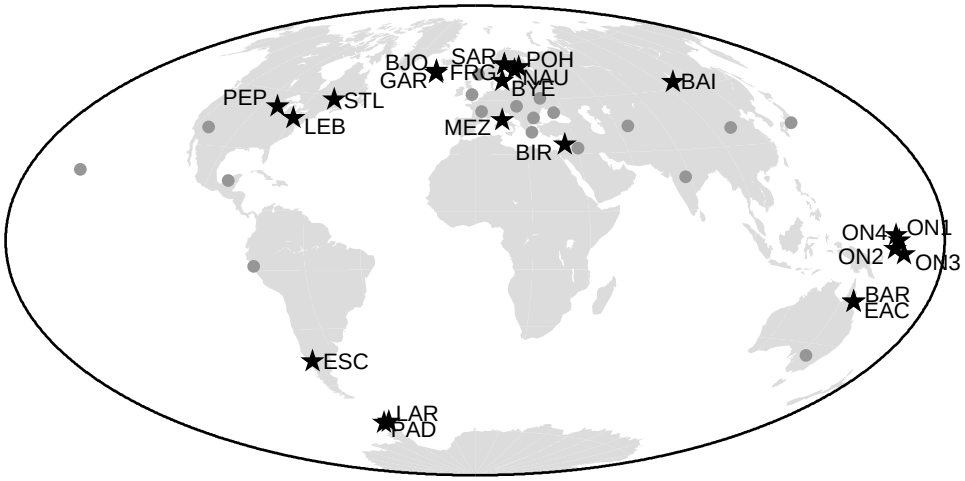

Figure 1. Locations of relative paleointensity records used in this study. See Table 1 for references. The gray dots are average locations of archeointensity data used in constructing the global model CALS7K.2.

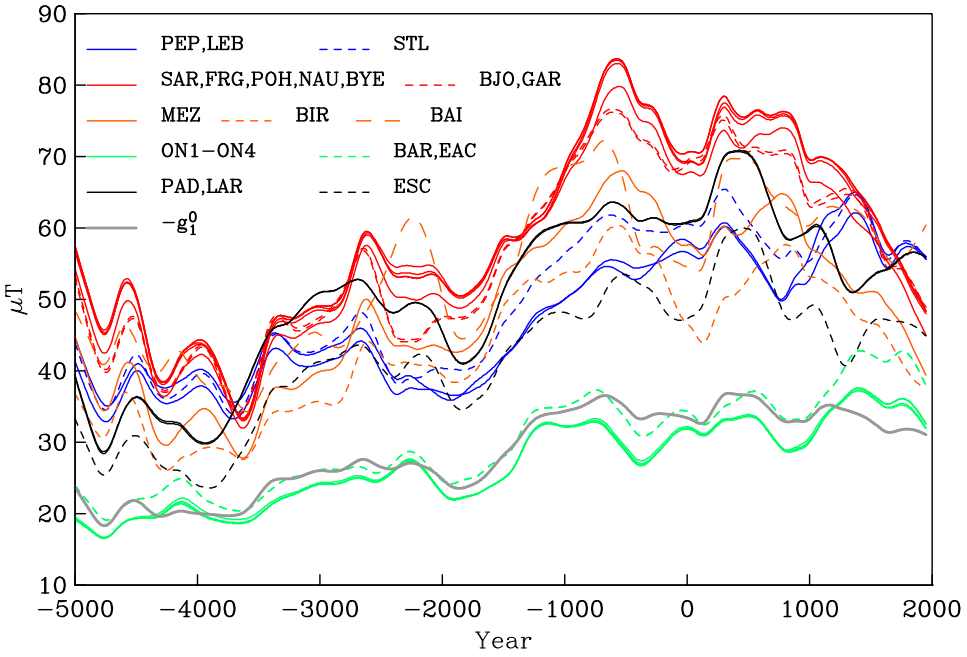

Figure 2. Predictions from CALS7K.2 of field intensity variations with time for all the relative intensity locations listed in Table 1. Colors indicate approximate regional groupings, except for MEZ, BIR, and BAI which have no nearby records. Gray line gives the variation in axial dipole coefficient g_1^0. Note that some records mirror the trend in g_1^0 while in others it is masked by substantial non-axial-dipole contributions.

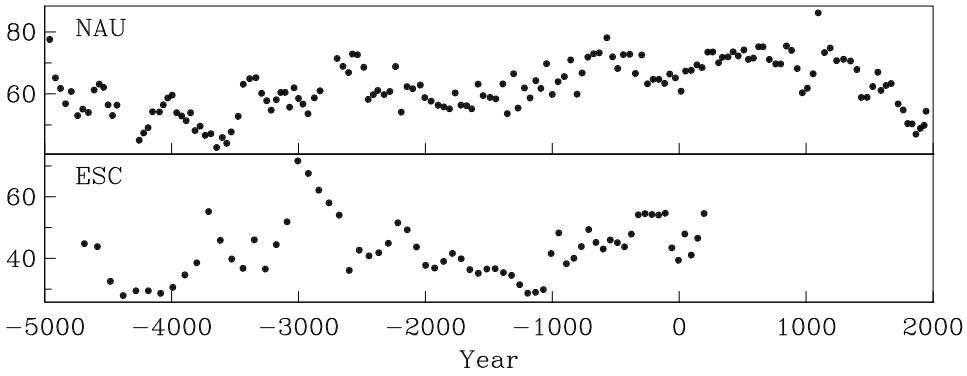

Figure 3. Ratio of model predictions to data values (F_{mod}/F_{dat}) for time series of NAU and ESC in μT.

Table 2. Calibration factors with standard deviation σ and root mean square (rms) misfit of CALS7K.2 to sediment intensity data. The last two columns give standard deviation in scaling factor and the misfit for the new model described in section 4.

Site	Median Factor	σ	σ in %	rms	σ_{new} in %	rms_{new}
BAI	54.53	10.73	20	9.15	16	7.40
BAR	43.37	10.81	25	9.39	24	10.16
BIR	49.57	15.70	32	18.7	28	17.00
BJO	18721	6957	37	16.90	29	12.97
BYE	59.10	12.40	21	11.08	16	8.27
EAC	53.94	40.58	75	16.11	72	15.75
ESC	42.67	9.31	22	8.73	19	7.13
FRG	53.95	13.55	25	9.31	22	6.92
GAR	13198	3626	27	13.24	21	10.19
LAR	45.75	24.84	54	17.00	37	21.30
LEB	139.30	36.90	26	13.25	19	10.00
MEZ	56.61	11.86	21	10.51	23	11.35
NAU	61.76	8.25	13	7.62	9	5.13
ON1	114.26	36.07	31	7.17	27	6.20
ON2	206.42	113.95	55	8.43	36	6.19
ON3	90.11	28.90	32	7.34	29	6.52
ON4	287.50	181.73	63	10.08	69	10.04
PAD	47.11	13.64	29	20.21	18	9.51
PEP	154.90	44.15	29	10.24	21	7.05
POH	68.77	9.73	14	8.5	12	6.98
SAR	70.87	10.15	14	7.63	11	6.72
STL	44.34	6.40	14	7.33	9	4.41

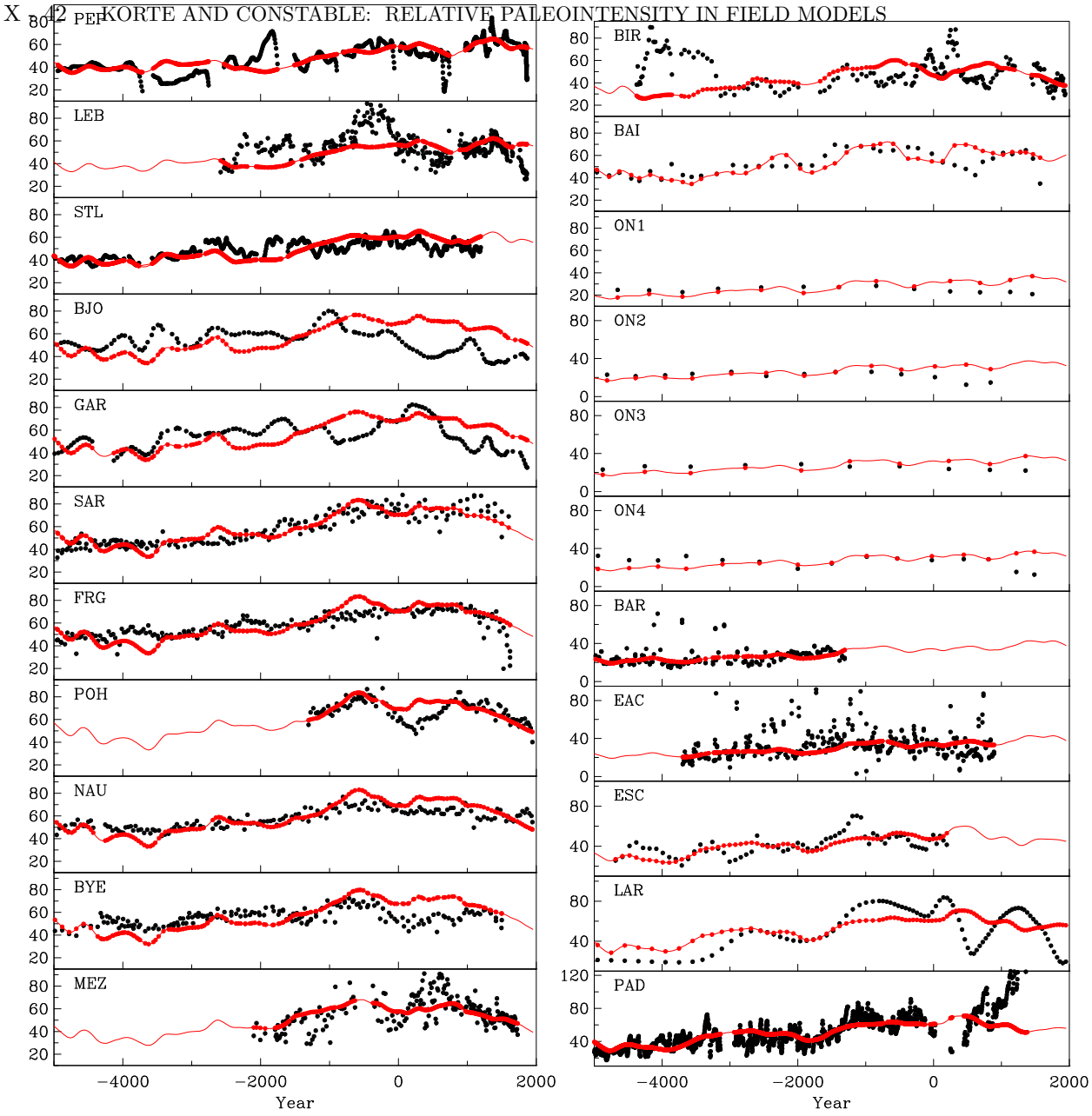

Figure 4. Comparison of CALS7K.2 model predictions (red) and calibrated sedimentary paleointensity time series (black) in μT.

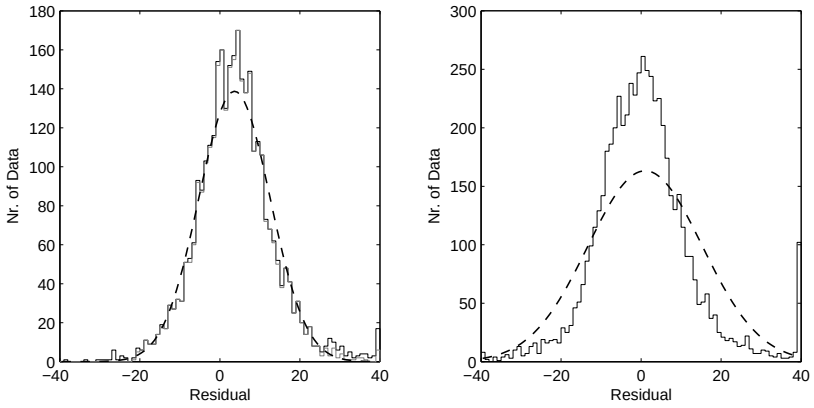

Figure 5. Histograms of residuals between intensity data and CALS7K.2 predictions. Left: absolute archeomagnetic data, all data before (black) and data after (gray) iterative data rejection used for CALS7K.2. Right: scaled sedimentary intensity data. The dashed curves are normal distributions with same mean and standard deviation.

Table 3. Number of data and misfit. N is the number of data, and rms is the root mean square misfit to a constant axial dipole of 30 μT before and after data rejection, denoted by subscripts i and r, respectively. R, % is the number of rejected data in percent of the initial data, and rms_f is the misfit to the final modified model.

Component		N_i	rms_i	R, %	N_r	rms_r	rms_f
Total		38201	2.41	14.6	32614	1.71	1.01
Inclination		16085	1.89	15.1	13663	1.40	0.94
Declination		13080	2.93	22.0	10206	1.64	1.00
Intensity	all	9036	2.41	3.2	8745	2.17	1.14
	absolute	3188	1.54	4.0	3062	1.46	0.96
	calibrated	5848	2.78	2.8	5683	2.48	1.23

Table 4. Parameters and norms for modified model. The annotations *b.r.* and *a.r.* mean before and after data rejection.

Iteration	rms	λ (nT^{-2})	Spatial norm (nT^2)	τ ($nT^{-2}yr^4$)	Temporal norm (nT^2yr^{-4})	Intensity factor
$1b.r.$	1.93	10^{-7}	10×10^{10}	10^{-1}	44×10^3	2
$2b.r.$	1.90	10^{-7}	82×10^9	10^{-1}	66×10^3	2
$3b.r.$	1.89	10^{-7}	91×10^9	10^{-1}	92×10^3	2
$1a.r.$	1.13	5×10^{-7}	35×10^9	1×10^{-1}	12×10^3	2
$2a.r.$	1.05	5×10^{-7}	53×10^9	1×10^{-1}	28×10^3	2
$3a.r.$	1.03	5×10^{-7}	54×10^9	1×10^{-1}	42×10^3	2
$4a.r.$	1.01	5×10^{-8}	67×10^9	1×10^{-1}	57×10^3	2

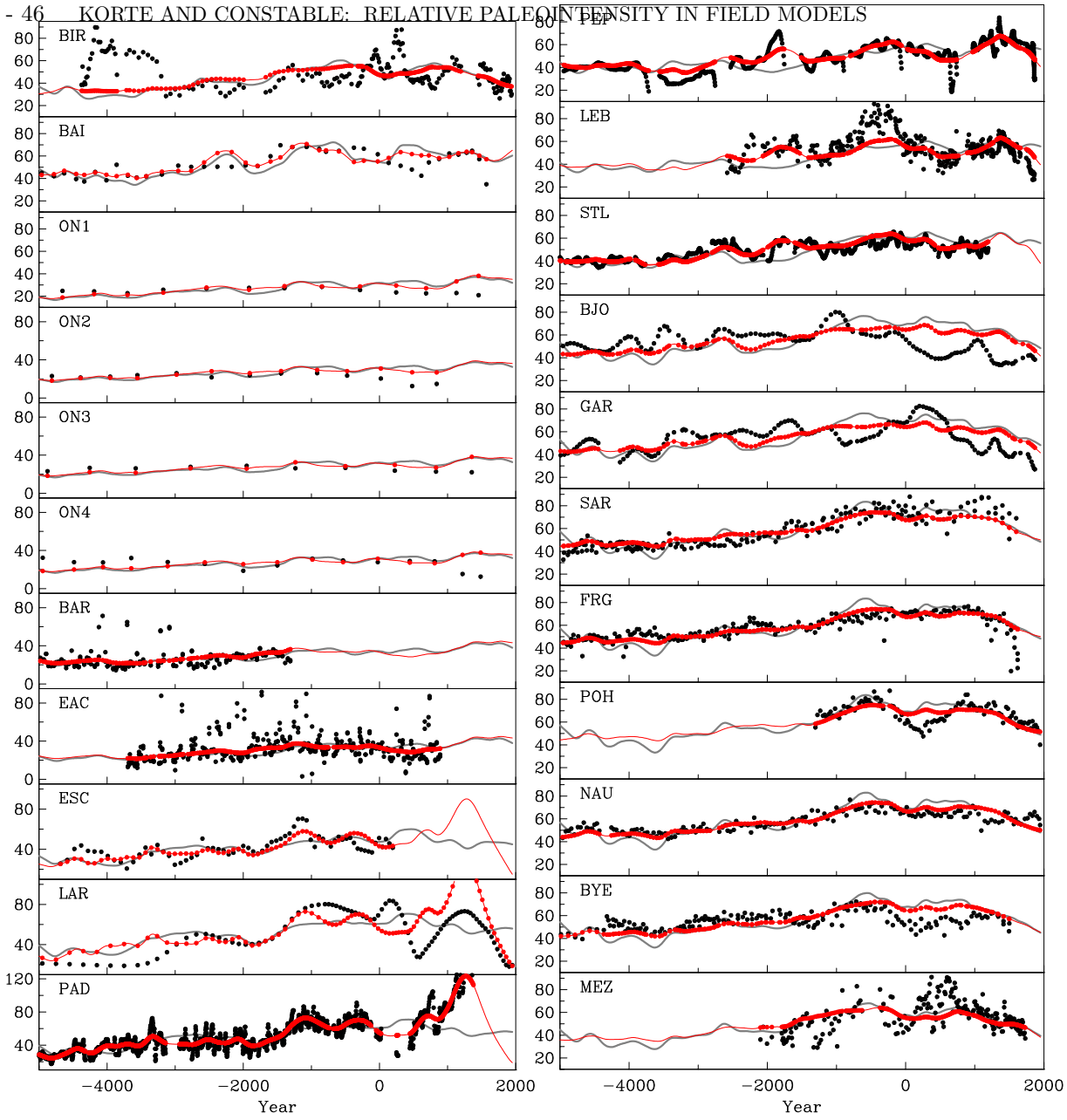

Figure 6. Comparison of calibrated sediment intensity data (black) and predictions from modified model (red) in μT. Gray line is CALS7K.2 prediction.

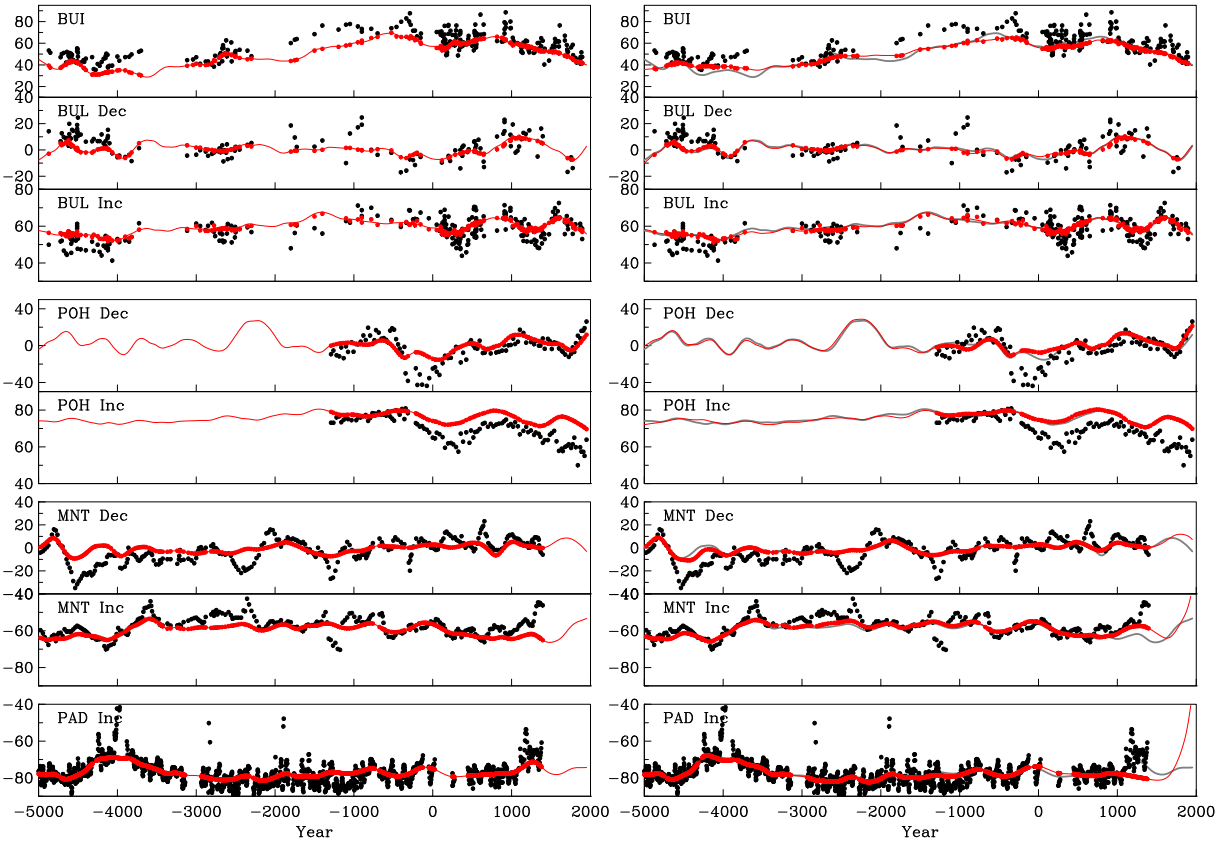

Figure 7. Comparison of fit of CALS7K.2 (red, left) and the model from this study (red, right) to some of the data (black): archeomagnetic data from the region of Bulgaria (BUI Intensity, BUL directions), sediment directional data from Scandinavia (POH), Argentina (MNT) and Antarctica (PAD). Gray line in the left panels is the prediction of CALS7K.2 again.

Table 5. Distribution of intensity data F in synthetic datasets with overall number of data.

Dataset	Latitude	Longitude	F distribution	Nr. of Data
D1	scattered	scattered	as in CALS7K.2	27061
D2	45 to 50	-5 to 5	localized F data	27101
D3	-80 to 80	0	all lat., one long.	27101
D4	45	0 to 360	all long., one lat.	27101
D5	-80 to 80	0 to 360	well-distributed	27101
D6	scattered	scattered	all I loc. of CALS7K.2	41316

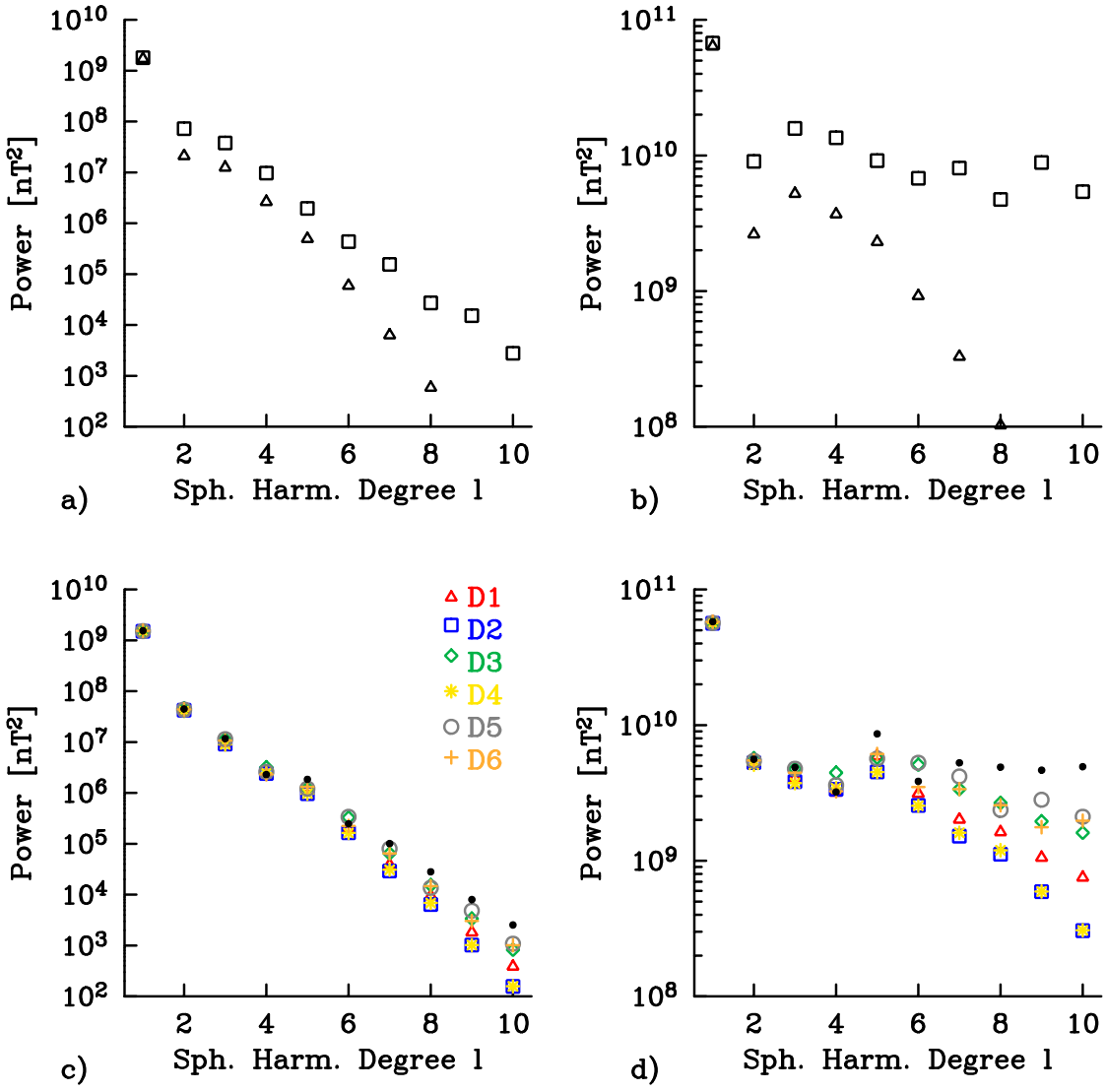

Figure 8. Time-averaged power spectra of CALS7K.2 (triangles) compared to a current high-resolution field model (squares, POMME3.0 *Maus et al.* [2006]) at a) the Earth's surface and b) the core-mantle boundary. Power spectra of final models from the six different synthetic datasets compared to the original numerical model (black dots) which was used to predict these synthetic data. All spectra are averaged spectra over the whole time interval, a) at the Earth's surface and b) at the core-mantle boundary, where the effects of diminished spatial resolution are enhanced.

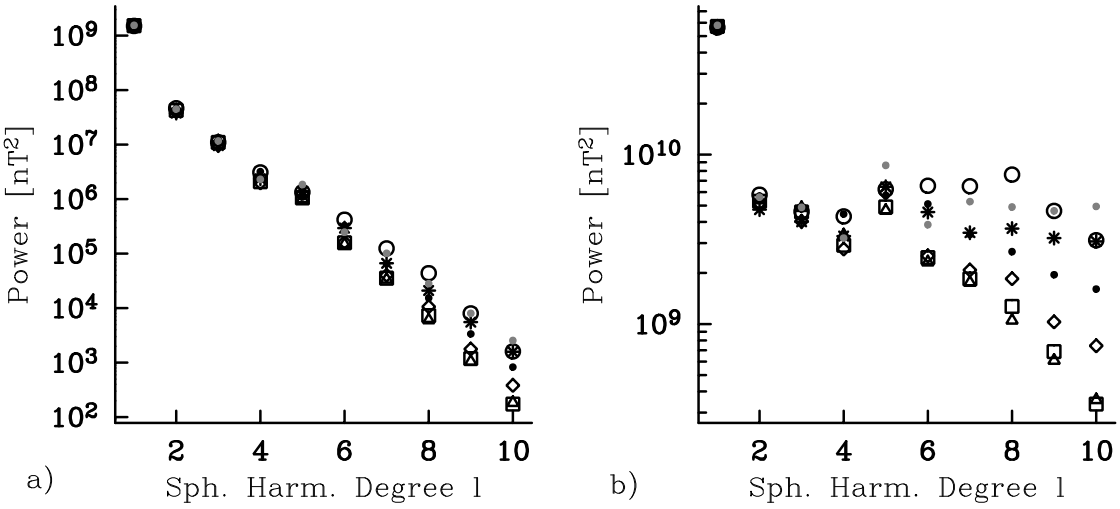

Figure 9. Time averaged power spectra of models based on dataset D3 (black dots) and modifications thereof with intensity data coming from several latitudinal profiles: at 60°E (triangles), 90°E (squares), 180°E (diamonds) and 160°W (asterisks), respectively, and at 60°E with all directional data positions also shifted east by 60° (circles). Gray dots are from the original numerical model, a) at the Earth's surface and b) at the core-mantle boundary.

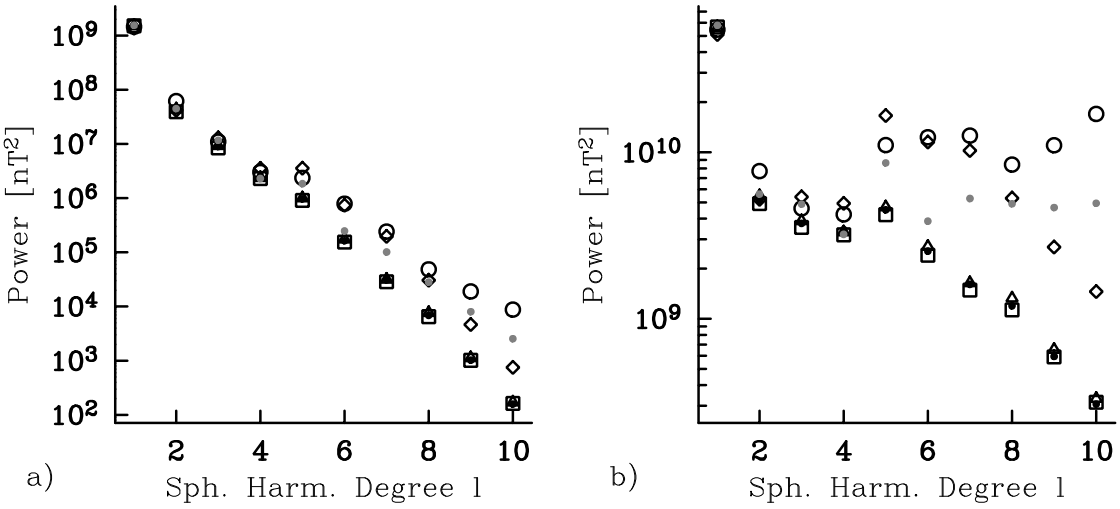

Figure 10. Time averaged power spectra of models based on dataset D4 (black dots) and modifications thereof with intensity data coming from different longitudinal circles: 30°N (triangles), 60°N (squares) and 45°S, respectively, and 45°S with latitudes of directional data positions reversed so that most data come from the southern hemisphere (diamonds). Gray dots are from the original numerical model, a) at the Earth's surface and b) at the core-mantle boundary.