A Comparison of Methods for Estimating the Geoelectric Field

R.S. Weigel

June 30, 2016

Abstract

The geoelectric field is the primary input used for estimation of geomagnetically induced currents (GICs) in conducting systems. We compare three methods for estimating the geoelectric field given the measured geomagnetic field at four locations in the U.S. The methods include using: (1) a pre–existing 1-D conductivity model, (2) a conventional 3-D frequency domain method, and (3) a robust and bias–corrected remote reference 3-D frequency domain method. The quality of the estimations are determined using the power spectrum (in the period range 9.1 to 18,725 seconds) of estimation errors along with the prediction efficiency summary statistic. It is shown that with respect to these quality metrics, Method 1 produces average electric field estimation errors that can be equal to or significantly above that measured (due to under- or over-estimation, respectively) and Method 3 produces in general reliable but lower quality estimates than Method 2 for the time intervals and locations considered.

1 Introduction

Historically, estimation of geoelectric field magnitudes used in GIC studies have often been made with 1-D conductivity models as these were the only models available over large phys-

iographic regions in the U.S. or because their historical use made them useful for comparison (e.g., *Pulkkinen et al.*, 2012; *Wei et al.*, 2013; *Viljanen et al.*, 2014; *Boteler*, 2015; *NERC*, 2015). These 1-D conductivity models are developed based on local geology and magnetotelluluric and seismic surveys and are intended as a first-order approximation of the 3-D conductivity structures that may exist (*Fernberg*, 2012; *Boteler*, 2015).

Over the past decade, the EMScope (*Schultz* [2009]) component of the EarthScope project (*Meltzer*, 2003) has developed transfer functions over a large span of the U.S. and made the transfer functions publicly available (*Schultz et al.* [2016]). These transfer functions were developed to model conductivity structures but can be used for the estimation of the geoelectric field given the geomagnetic field on Earth's surface (*Bedrosian and Love*, 2015).

There is a large body of literature in the magnetotelluluric (MT) community on computing surface impedance tensors, $\mathcal{Z}(\omega)$, that connect the geomagnetic field, $\mathbf{B}$ to the geoelectric field $\mathbf{E}$ on Earth's surface using $\mathbf{E}(\omega) = \mathcal{Z}(\omega)\mathbf{B}(\omega)$, where $\mathbf{B}(\omega) = [B_x(\omega), B_y(\omega)]^T$ and $\mathbf{E}(\omega) = [E_x(\omega), E_y(\omega)]^T$ (*Chave and Jones* [2012] and references therein). The primary objective of such estimates is for an impedance tensor that can be used to compute a 3-D model of Earth's conductivity. These impedance tensors are derived using statistical methods that are customized to reduce bias and increase robustness for such conductivity model estimates. The quantity minimized in developing these models is a weighted residual, where the weights and data intervals used are based on an iterative process and the residual used is either a L_2 or L_1 norm that depends on the magnitude of the residual (*Simpson and Bahr* [2005] chapter 4; *Chave and Jones* [2012] chapter 5). In general, the quality of the computed transfer function is assessed by their visual characteristics, error bars, and consistency when different data segments are used in their computation (e.g., *Jones et al.* [1989]; *Fujii et al.* [2015]).

In contrast, for estimating GICs, which are computed using the estimated $\mathbf{E}(t)$ (*Lehtinen and Pirjola*, 1985; *Pulkkinen et al.*, 2010; *Viljanen et al.*, 2012; *NERC*, 2015), the primary objective is to estimate $\mathbf{E}(t)$ given $\mathbf{B}(t)$ with a high degree of accuracy. The primary as-

sessment of the quality of the estimation is generally based on the overall match between the predicted and measured time series. The most common metric for assessing quality is either a visual inspection of the predicted $\mathbf{E}(t)$ or GIC(t) computed using $\mathbf{E}(t)$ versus the measured $\mathbf{E}(t)$ or GIC(t), the histogram of error, and/or a sum-of-squares error-based statistic (*McKay*, 2003; *Pulkkinen et al.*, 2010; *Love and Swidinsky*, 2014). In this work, we also consider the frequency dependence of the error in order to identify situations where, for example, one method may better estimate high frequency fluctuations than low frequency variations.

This difference in assessment of the quality of estimates between the MT and GIC community motivated the use of a conventional method (*Sims et al.*, 1971) for the estimation of $\mathcal{Z}$. MT researchers often cite the results of *Jones et al.* [1989] which showed that complex robust and remote reference methods were superior to conventional spectral processing methods in estimating impedance tensors. However, to date, no comparison has been made to determine the influence of the additional layers of computation involved for robust and remote reference processing on the quality of the electric field estimates from the perspective of the GIC community.

2 Methods

The three methods considered for estimating the surface geoelectric field given measurements of the surface geomagnetic field are given below. Method 1 is referred to as a 1-D method because the impedance tensor depends only with depth. Methods 2 and 3 are referred to as 3-D methods because their impedance tensors depend on depth and horizontal directions.

2.1 Method 1

A surface impedance is computed from a pre-existing 1-D model of conductivity, σ, versus depth, d, using

$$Z_n(\omega) = F\big(\sigma(d), d, \omega\big) \tag{1}$$

where the function F provides the surface impedance from the use of Wait's recursion formula (Wait [1954]; see also formula 2.33 of Simpson and Bahr 2005). $E_x(\omega)$ and $E_y(\omega)$ are computed using

$$\begin{aligned} E_x(\omega) &= Z_n(\omega) B_y(\omega) \\ E_y(\omega) &= -Z_n(\omega) B_x(\omega) \end{aligned} \tag{2}$$

and then $E_x(t)$ and $E_y(t)$ are computed from the inverse fourier transforms of $E_x(\omega)$ and $E_y(\omega)$, respectively.

2.2 Method 2

$\mathbf{E}(t)$ and $\mathbf{B}(t)$ measurements are used to solve for $\mathcal{Z}$ in

$$\mathbf{E}(\omega) = \mathcal{Z}(\omega)\mathbf{B}(\omega) \tag{3}$$

where

$$\mathcal{Z} = \begin{bmatrix} Z_{xx}(\omega) & Z_{xy}(\omega) \\ Z_{yx}(\omega) & Z_{yy}(\omega) \end{bmatrix} \tag{4}$$

using a linear least squares method (*Sims et al.*, 1971; *Simpson and Bahr*, 2005). In this work, the evaluation frequencies were selected to be logarithmically spaced and the spectral and cross-spectral values required for computing $\mathcal{Z}$ at each evaluation frequency are determined using a Parzen averaging window. The largest evaluation base frequency was set at 0.25 seconds and the ratio of consecutive frequencies is $\sqrt{2}$; the actual evaluation frequencies were chosen to be the frequency from the fast fourier transformed measurements nearest to the evaluation base frequency; the actual ratio of evaluation frequencies varied

between 1.25 and 1.5. $E_x(t)$ and $E_y(t)$ are computed from the inverse fourier transforms of $E_x(\omega)$ and $E_y(\omega)$, respectively, after linear interpolation of the components of $\mathcal{Z}$ on to a uniform frequency grid with frequency spacing of $1/N$ Hz, where N is the length of the 1-second-cadence prediction segment.

We have considered using linearly spaced evaluation frequencies and a rectangular window of various widths along with a Bartlett averaging window. The most important factor was the use of logarithmically spaced evaluation frequencies. With this, the use of a Parzen averaging window provided slight improvements over that for a rectangular or Bartlett averaging window. Linearly spaced evaluation frequencies with any window resulted in higher errors at periods above 10^3 s but similar errors below.

2.3 Method 3

Method 3 is similar to Method 2 except the estimate of $\mathcal{Z}$ is made using a robust regression method and auxiliary remote reference measurements are used (*Chave and Jones* [2012]). We have not implemented this algorithm but rather have used pre-computed impedance tensors from the MTScope (*Schultz et al.*, 2016) and used them to compute estimates of $E_x(t)$ and $E_y(t)$ in the same way as Method 2. The provided impedance tensors have frequencies that are approximately logarithmically spaced in the period range of 9.1-18,725 seconds, with ratios of evaluation periods in the range of 1.25–1.64. To compute a predicted electric field, linear interpolation was used to obtain impedances on a uniform frequency grid. All of the transfer functions used in this work had the highest provider-assessed quality score (5 on a scale of 0-5).

3 Data

The four stations listed in Table 1 were selected because they fell into one of the physiographic regions for which the 1-D conductivity models of (*Fernberg* [2012]) is available and also had

ID	Location	1-D model
UTP17	The Cove, UT	CL-1
GAA54	Gator Slide, GA	CP-2
ORF03	Jewell, OR	PB-2
RET54	Buffalo Cove, NC	PT-1

Table 1: Site locations and applicable 1-D conductivity models used in this work

four-day time intervals with few spikes and no baseline offsets; the first four-day interval that had these characteristics was selected. The raw instrument count measurements were used after conversion to physical units with a constant scale factor. Data spikes were manually removed and the **E** measurements were filtered by zeroing frequencies outside of the range of 9.1-18,725 seconds, corresponding to the range of available impedances for Method 2. For purposes of estimation, this zeroing should be avoided because we have found that the zeroed components of the signal are still predictable in the sense that the spectrum of the prediction error for Method 3 is less than the spectrum of the predictand (as discussed in the following section).

The four-day intervals of 1-second cadence measurements were split into two-day segments. To determine out-of-sample estimation errors for Method 2, the first two-day interval was used for computing the model and the second interval was used for testing. We have also computed results for when the second interval was used for training and the first interval was used as the test set and the overall trends and the results are similar; for brevity, these results are not presented in this paper.

Because the exact intervals used for determining the models for Method 3 are not known, all results should be considered as in-sample. However, because the model has a small number of free parameters (approximately twice the number of evaluation frequencies), overfitting is not expected to be a concern for any of the models.

4 Results

The summary statistic of the prediction efficiency was used as an overall measure of estimation quality and the spectrum of prediction errors was also considered. The prediction efficiency, $PE = 1 - ARV$, where the average relative variance $ARV = \langle (p-t)^2 \rangle / \sigma_t$, where p is the prediction and t is the target time series; a prediction efficiency of 1 corresponds to a perfect prediction, and a prediction efficiency of 0 corresponds to a prediction that is no better than using the average of t as a predictor. The advantage of the prediction efficiency is in this interpretation and due to the fact that high correlations that occur when the prediction signal is a scaled version of the measured signal will result in low prediction efficiencies.

Table 2 shows the results for these summary statistics for both in-sample and out-of-sample segments. The primary feature is the ordering of the out-of-sample prediction efficiencies. In all cases, $PE(\text{Method 2}) > PE(\text{Method 3}) > PE(\text{Method 1})$ and the separation between Method 3 and Method 1 is greater than that for Method 2 and Method 3. We note that the difference between Method 2 and Method 3 can be dependent on the zeroing of periods outside of the range of $9.1 - 18,725$ seconds, with the separation sometimes becoming larger when this constraint is removed. As an example, when this constraint is removed for RET54, the training/testing prediction efficiencies for E_y for Method 2 slightly increase from 0.96/0.93 to 0.97/0.94 and for Method 3 they decrease from 0.87/0.88 to 0.81/0.78. This is explained by the fact that variations outside of the range of $9.1 - 18,725$ seconds are predictable for this site.

The smoothed error spectra are shown in Figure 1 and the data used for their computation are shown in Figures 2-5. All of the time series displayed in Figures 2-5 were filtered to have zero spectral amplitudes outside of the range of $9.1 - 18,725$ seconds, and the first and last 18,725 seconds were omitted in the computation of correlations and prediction efficiencies. The error spectra and time series for E_x are not shown, but the results and conclusions are the similar to that for E_y. The smoothed error spectra were computed using the same

approach for the spectral components of the transfer function for Method 2; logarithmically spaced evaluation frequencies were used along averages weighted with a Parzen window.

Consistent with the prediction efficiency results, in Figure 1 in most cases the error spectra is lower at all periods for Method 2, and Method 3 has amplitudes that are larger than Method 1.

The spectral errors are not flat, indicating that the residuals are frequency dependent. As noted by *Egbert and Booker* [1986], many of the assumptions required for estimating error bars and bias in $\mathcal{Z}$ for least-squares-based methods are not applicable and in robust estimation of the impedance tensors for Method 3, the contribution of the error for each $\mathcal{Z}_{ij}$ is approximately scaled by this error curve.

In Figures 2, 3, and 5, the intermittent spikes in the error time series are due to spikes in the measured magnetic field that were not obvious or removed when the four-day time series of $E_x(t)$ and $E_y(t)$ was visually inspected.

5 Summary and Conclusions

We have shown that Method 1 produces results that are inferior to Method 2 and Method 3. The primary reasons are that (1) the transfer functions cover a very large geographic region over which the transfer function can change – the transfer functions computed for Methods 2 and 3 show that within the physiographic regions defined by *Fernberg* [2012], significant differences in the transfer function exist; and (2) the assumption that $\mathcal{Z}_{xy} = -\mathcal{Z}_{yx}$ – for all of the sites considered, the ratio of these impedances range from ∼0.2 to 5.

It was shown ~~have shown~~ that for data sets without many defects (spikes and baseline jumps), a straightforward algorithm (Method 2) for computing a transfer function yields near equal or better estimates estimates than a method that uses a remote reference and attempts to reduce bias (Method 3) in the estimate of the transfer function.

The MT literature most often uses the frequency domain method and many works advo-

cate the use of robust methods along with remote reference measurements. These methods have been argued to be important when making unbiased estimates of the characteristics of transfer functions (*Chave and Jones* [2012]). However, in practice, remote reference data may not be available, and the most straightforward statistical method should be used to simplify interpretation. We have shown that a conventional least squares frequency domain method that can give reliable and accurate out-of-sample estimates of the geoelectric field. We note that it is an open question as to whether Method 2 produces reliable estimates of the actual transfer function; although the computed transfer functions from Method 2 differ from those of Method 3, without a ground truth for comparison, it is not possible to make firm conclusions in this regard.

From the GIC perspective, the method to use for estimating the geoelectric field given geomagnetic field measurements depends on a number of factors and the results indicate that when possible both Methods 2 and 3 are viable options. It is an open question as to how much revised estimates of historical geoelectric field estimates made with Method 1 (e.g., *Pulkkinen et al.*, 2012; *Wei et al.*, 2013) will change when Methods 2 or 3 are used. For the sites considered, Method 1 produced both over- and under-estimates and variations that had the correct scale (but poor detailed resolution).

6 Acknowledgments

We acknowledge Anna Kelbert, Gary Egbert, and Adam Schultz for assistance clarifications on the calibration of measurements from the USArray MT measurements.

The USArray MT TA project was led by PI Adam Schultz and Gary Egbert. They would like to thank the Oregon State University MT team and their contractors, lab and field personnel over the years for assistance with data collection, quality control, processing and archiving. They also thank numerous districts of the U.S. Forest Service, Bureau of Land Management, the U.S. National Parks, the collected State land offices, and the many private

landowners who permitted access to acquire the MT TA data. USArray TA was funded through NSF grants EAR-0323311, IRIS Subaward 478 and 489 under NSF Cooperative Agreement EAR-0350030 and EAR-0323309, IRIS Subaward 75-MT under NSF Cooperative Agreement EAR-0733069 under CFDA No. 47.050, and IRIS Subaward 05-OSU-SAGE under NSF Cooperative Agreement EAR-1261681.

References

Bedrosian, P. A., and J. J. Love, Mapping geoelectric fields during magnetic storms: Synthetic analysis of empirical United States impedances, *Geophys. Res. Lett.*, *42*, 10, 2015.

Boteler, D. H., The evolution of Quebec earth models used to model geomagnetically induced currents, *IEEE Transactions on Power Delivery*, *30*, 2171–2178, 2015.

Chave, A. D., and A. G. Jones, *The Magnetotelluric Method: Theory and Practice*, Cambridge University Press, 2012.

Egbert, G. D., and J. R. Booker, Robust estimation of geomagnetic transfer functions, *Geophysical Journal International*, *87*, 173–194, 1986.

Fernberg, P., One-dimensional earth resistivity models for selected areas of the continental United States and Alaska, 2012.

Fujii, I., T. Ookawa, S. Nagamachi, and T. Owada, The characteristics of geoelectric fields at Kakioka, Kanoya, and Memambetsu inferred from voltage measurements during 2000 to 2011, *Earth, Planets and Space*, *67*, 2015.

Jones, A. G., A. D. Chave, G. Egbert, D. Auld, and K. Bahr, A comparison of techniques for magnetotelluric response function estimation, *J. Geophys. Res.*, *94*, 14,201–14,213, 1989.

Lehtinen, M., and R. Pirjola, Currents produced in earthed conductor networks by geomagnetically-induced electric fields, *Annales Geophysicae*, *3*, 479–484, 1985.

Love, J. J., and A. Swidinsky, Time causal operational estimation of electric fields induced in the earth's lithosphere during magnetic storms, *Geophys. Res. Lett.*, *41*, 2266–2274, 2014.

McKay, A. J., Geoelectric fields and geomagnetically induced currents in the United Kingdom, 2003.

Meltzer, A., Earthscope: Opportunities and challenges for earth-science research and education, *The Leading Edge*, *22*, 2003.

NERC, NERC Project 2013-03 - benchmark geomagnetic disturbance event description, 2015.

Pulkkinen, A., R. Kataoka, S. Watari, and M. Ichiki, Modeling geomagnetically induced currents in hokkaido, japan, *Advances in Space Research*, *46*, 1087 – 1093, 2010.

Pulkkinen, A., E. Bernabeu, J. Eichner, C. Beggan, and A. W. P. Thomson, Generation of 100-year geomagnetically induced current scenarios, *Space Weather*, *10*, 2012.

Schultz, A., EMScope: A continental scale magnetotelluric observatory and data discovery resource, *Data Science Journal*, *8*, 2009.

Schultz, A., G. D. Egbert, A. Kelbert, T. Peery, V. Clote, B. Fry, S. Erofeeva, and staff of the National Geoelectromagnetic Facility and their contractors (2006-2018), *doi:10.17611/DP/EMTF/USARRAY/TA*, 2016.

Simpson, F., and K. Bahr, *Practical Magnetotellurics*, Cambridge University Press, 2005.

Sims, W. E., F. X. Bostick, and H. W. Smith, The estimation of magnetotelluric impedance tensor elements from measured data, *Geophysics*, *36*, 938–942, 1971.

Viljanen, A., R. Pirjola, M. Wik, A. Ádám, E. Prácser, Y. Sakharov, and J. Katkalov, Continental scale modelling of geomagnetically induced currents, *Journal of Space Weather and Space Climate*, *2*, A17, 2012.

Viljanen, A., R. Pirjola, E. Prácser, J. Katkalov, and M. Wik, Geomagnetically induced currents in Europe. Modelled occurrence in a continent-wide power grid, *Journal of Space Weather and Space Climate*, *4*, A09, 2014.

Wei, L. H., N. Homeier, and J. L. Gannon, Surface electric fields for north america during historical geomagnetic storms, *Space Weather*, *11*, 451–462, 2013.

Training set

	cc_x	PE_x	cc_y	PE_y
Method 1	0.92	-0.91	0.75	0.12
Method 2	0.99	0.97	0.99	0.98
Method 3	0.97	0.88	0.97	0.93

Testing set

	cc_x	PE_x	cc_y	PE_y
Method 1	0.92	-1.67	0.71	-0.49
Method 2	0.99	0.97	0.99	0.98
Method 3	0.98	0.92	0.97	0.93

(a) UTP17 summary statistics based on data displayed in Figure 2.

Training set

	cc_x	PE_x	cc_y	PE_y
Method 1	0.44	0.19	0.44	-0.08
Method 2	0.94	0.89	0.89	0.79
Method 3	0.93	0.86	0.84	0.68

Testing set

	cc_x	PE_x	cc_y	PE_y
Method 1	0.46	0.21	0.56	-0.11
Method 2	0.92	0.84	0.94	0.87
Method 3	0.91	0.83	0.93	0.82

(b) GAA54 summary statistics based on data displayed in Figure 3.

Training set

	cc_x	PE_x	cc_y	PE_y
Method 1	0.85	0.72	0.68	0.37
Method 2	0.96	0.92	0.92	0.84
Method 3	0.91	0.81	0.85	0.61

Testing set

	cc_x	PE_x	cc_y	PE_y
Method 1	0.69	0.45	0.67	-0.02
Method 2	0.92	0.84	0.89	0.79
Method 3	0.80	0.59	0.92	0.67

(c) ORF03 summary statistics based on data displayed in Figure 4.

Training set

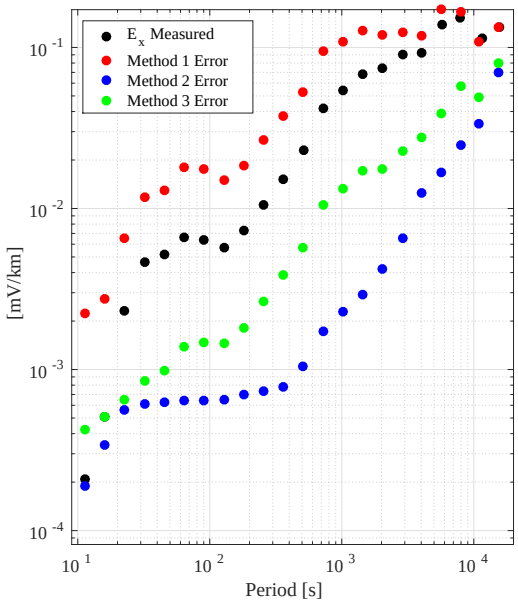
	cc_x	PE_x	cc_y	PE_y
Method 1	0.91	0.79	0.62	0.26
Method 2	0.98	0.97	0.98	0.96
Method 3	0.96	0.85	0.95	0.87

Testing set

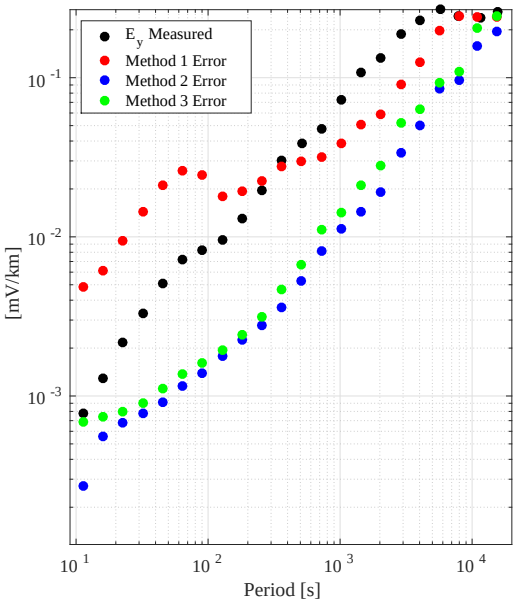
	cc_x	PE_x	cc_y	PE_y
Method 1	0.91	0.81	0.58	0.25
Method 2	0.98	0.96	0.97	0.93
Method 3	0.98	0.90	0.96	0.88

(d) RET54 summary statistics based on data displayed in Figure 5.

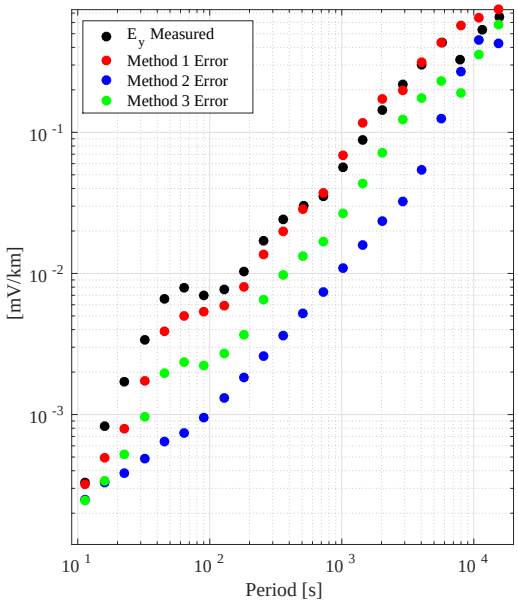
Table 2: Summary error statistics based on time series shown in Figures 2–5. The first and last 18,725 seconds of the four-day segments were omitted from the calculations.



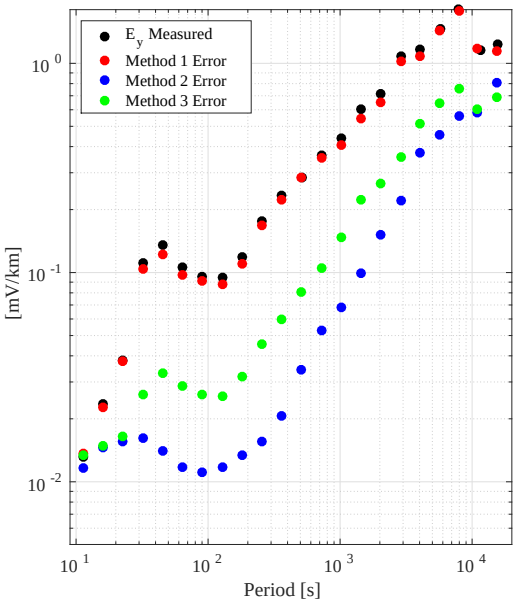
(a) UTP17 error spectrum based on data displayed in Figure 2.



(b) GAA54 error spectrum based on data displayed in Figure 3.



(c) ORF03 error spectrum based on data displayed in Figure 4.



(d) RET54 error spectrum based on data displayed in Figure 5.

Figure 1: Spectra of test interval errors.

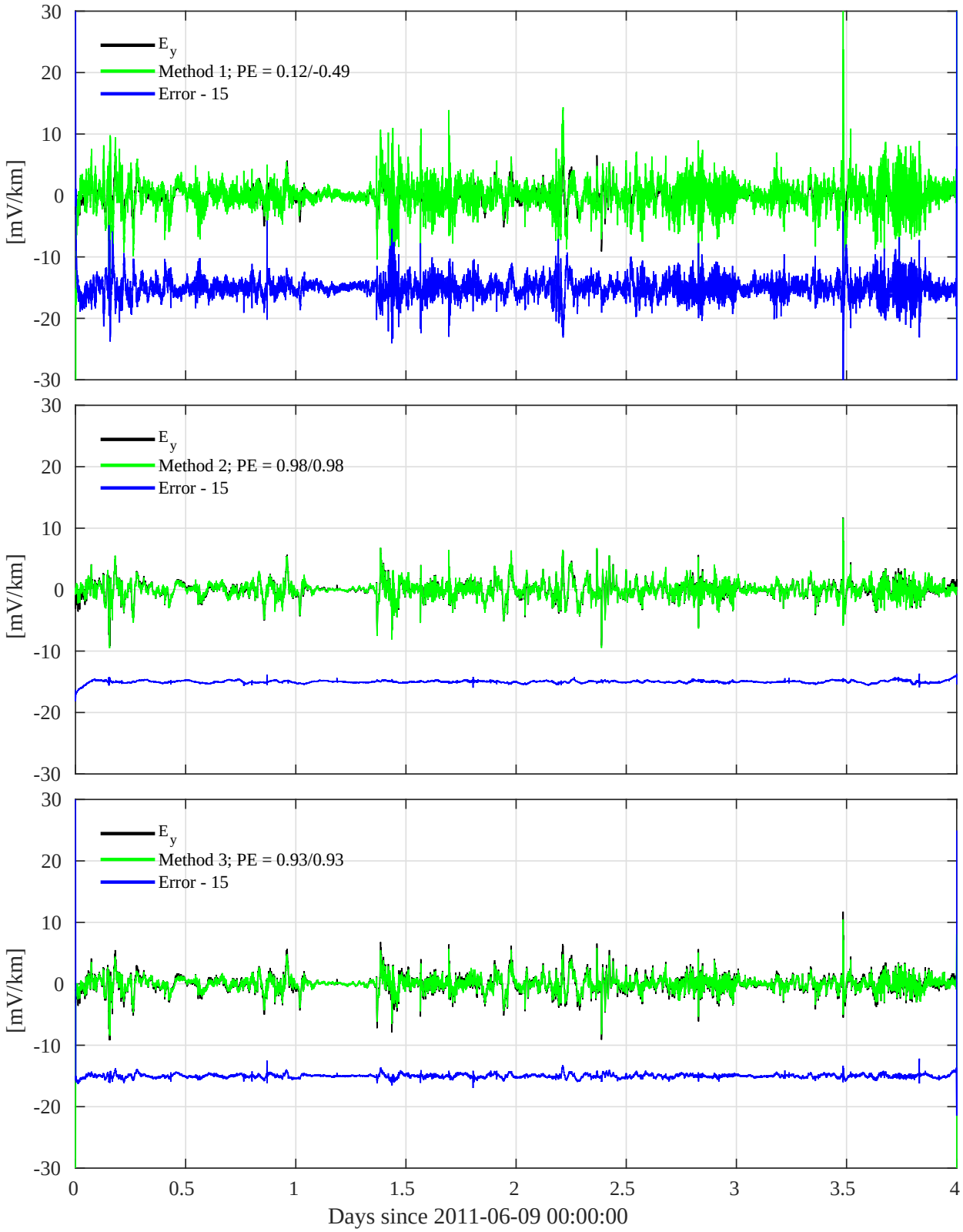


Figure 2: Measured, predicted, and error time series for E_y at UTP17. The two prediction efficiencies are for the training/testing intervals, which correspond to the first and last half of the four-day interval.

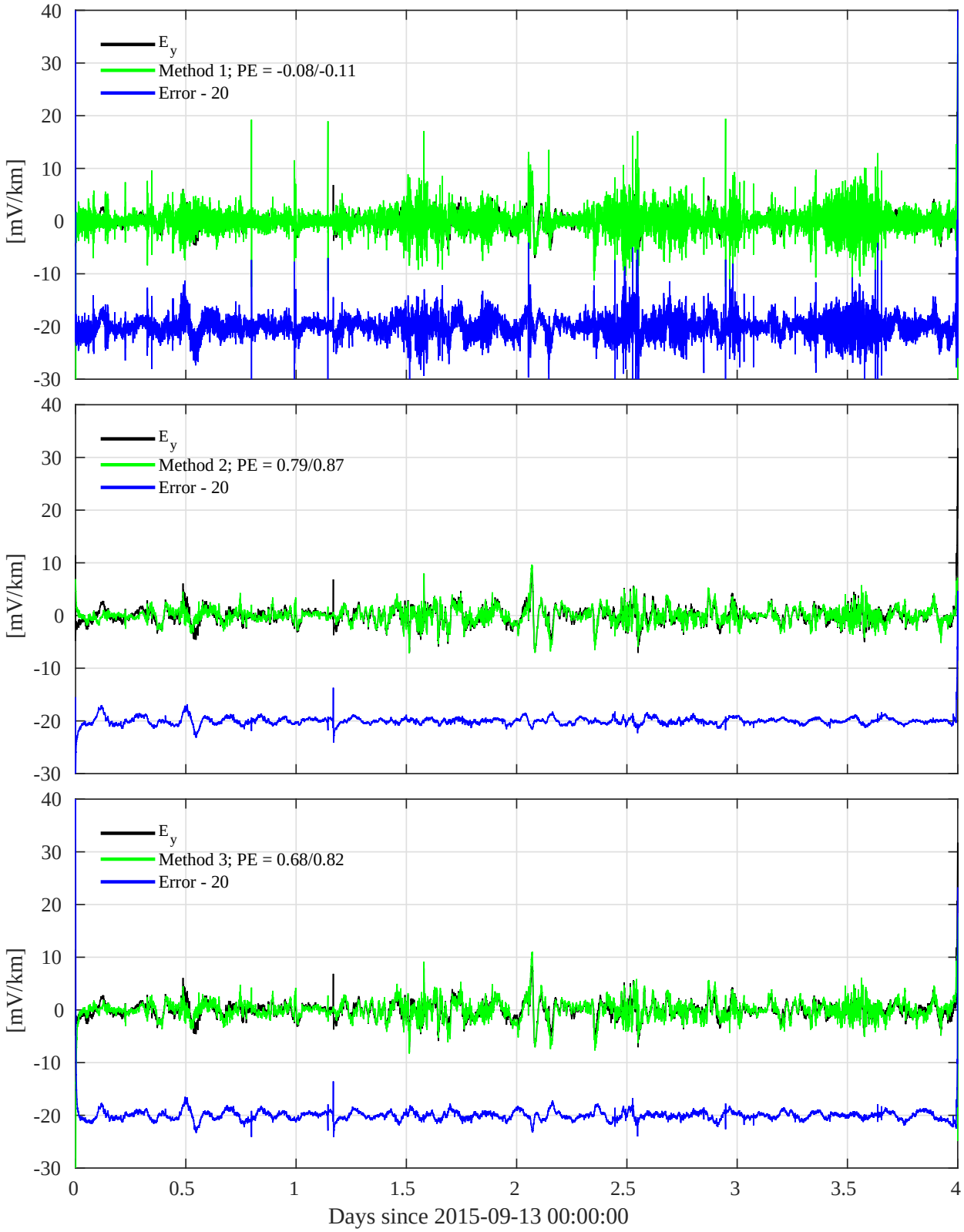


Figure 3: Measured, predicted, and error time series for E_y at GAA54.

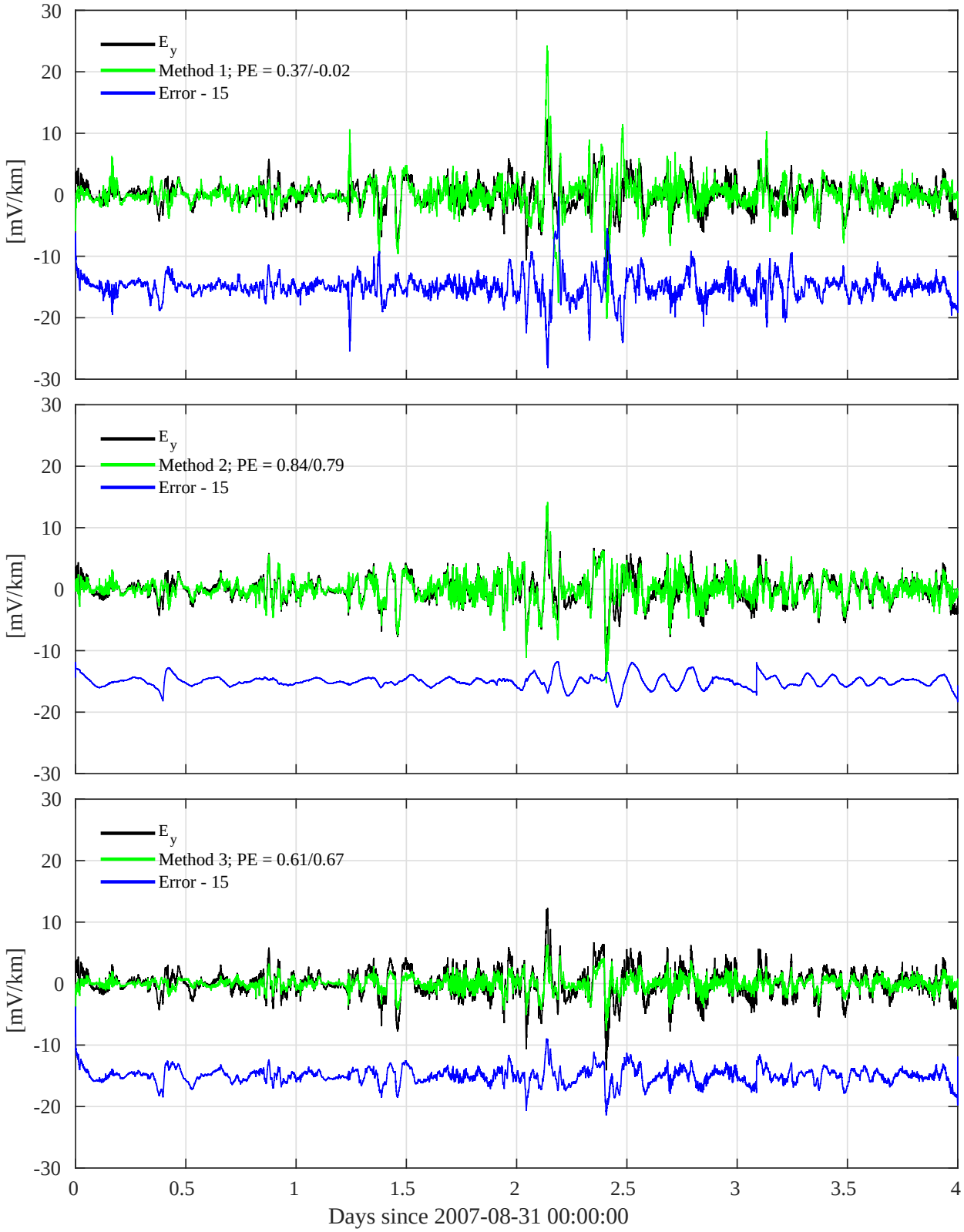


Figure 4: Measured, predicted, and error time series for E_y at ORF03.

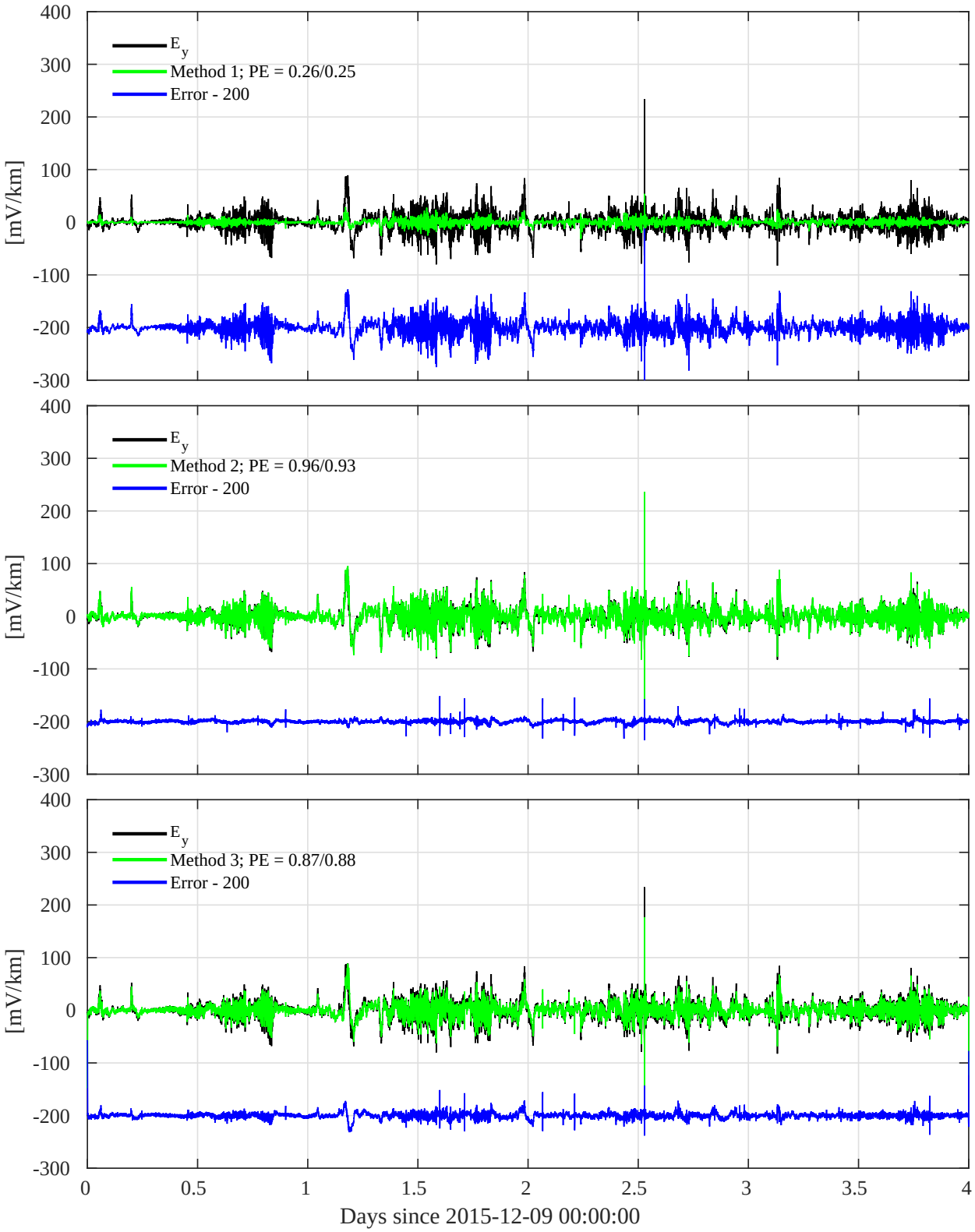


Figure 5: Measured, predicted, and error time series for E_y at RET54.