Supplementary Material

Transfer Entropy between South Atlantic Anomaly and Global Sea Level for the last 300 years

S. A. Campuzano, A. De Santis, F. J. Pavón-Carrasco, M. L. Osete, E. Qamili

Correspondence: S. A. Campuzano: sacampuzano@ucm.es

1 Supplementary Figures and Tables

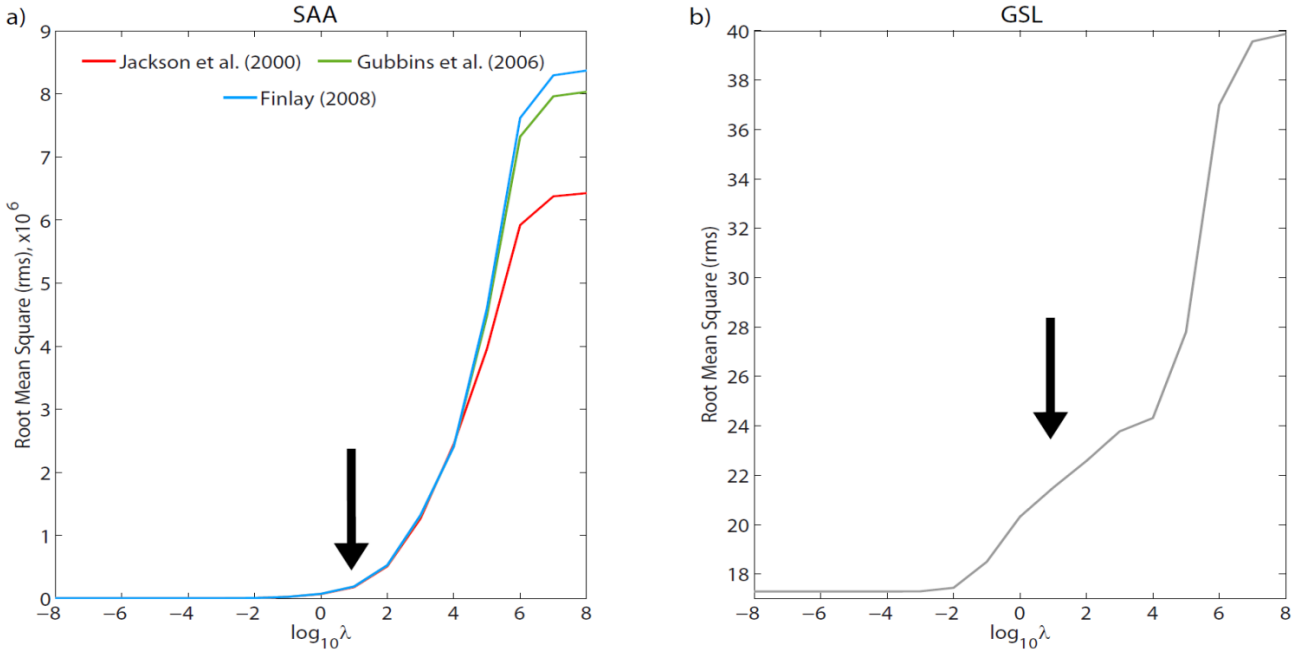
The supporting information included here are figures and tables that complement the analysis carried out along the paper. Figure S1 represents the root mean square (rms) error in function on different damping parameters (λ), used to determinate its optimal value according to the trade-off curves. Figure S2 is a schematic representation to better understand how works the transfer entropy from a system J (SAA anomalies) to other system I (GSL anomalies). Figure S3 represents the results given by the mutual information and the autocorrelation function to obtain the optimal embedding parameters (k_{SAA} and k_{GSL}). Table S1 shows different sets of optimal parameters (number of bins, S, and embedding parameters, k). Table S2 contains the results of TE according to the sets of parameters given in Table S1. More details about the results from the Table S2 are given below.

Table S2 reflects the results of the tests carried out changing both number of bins (S) and embedding parameters (k or l). We consider different cases for each information flow sense. From SAA to GSL (the other sense, from GSL to SAA, is analogous) we have checked 3 different values for the memory of the SAA time series (l) using the optimal embedding parameter for the GSL series (k) as follows: a) l = 1. b) l = k. c) l equal to the optimal embedding parameter for the SAA series. In order to see how the number of bins affects these results, we have repeated the analysis using the different S values contained in Table S1.

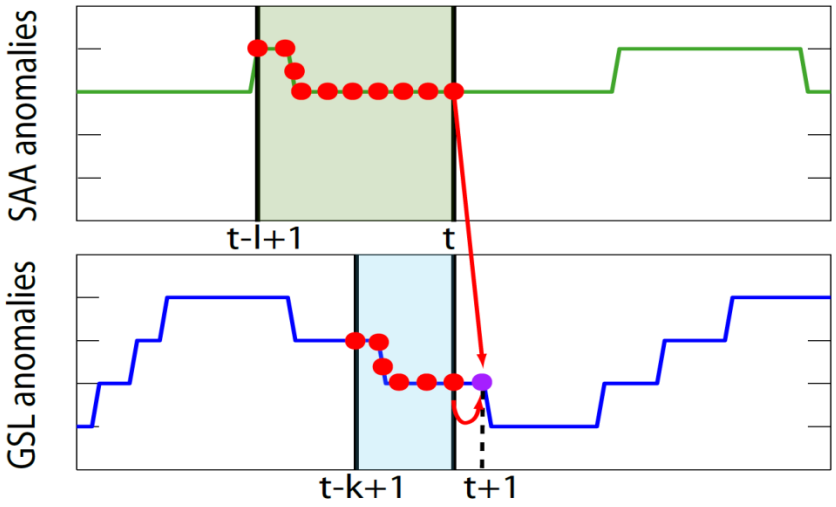
The results show that the TE increases when more information is available in the system. However, the statistical significance decreases due to the finite sample effect [e.g. Marschinski and Kantz, 2002]. In any case, the best results, according to the obtained significance, are provided by the optimal value of S = 4 and l = 1. Moreover, regardless of the selection of the number of bins and the embedding dimension, the predominant sense of the TE is from SAA to GSL anomalies. This reinforces the result reached.

Finally we would like to point out what happens when an incorrect number of bins is chosen (see Table S2d). In this case, the optimal value of the GSL anomalies series is badly conditioned (see Figure S3b) registering spurious local maxima. If we consider the highest maximum (S=11) as the optimal S value, we are overestimating the degrees of freedom of the time series, increasing the amount of information available in GSL anomalies and enhancing the TE from GSL to SAA anomalies. However, the statistical significance is much lesser than in the rest of reported results reflecting this bad conditioning.

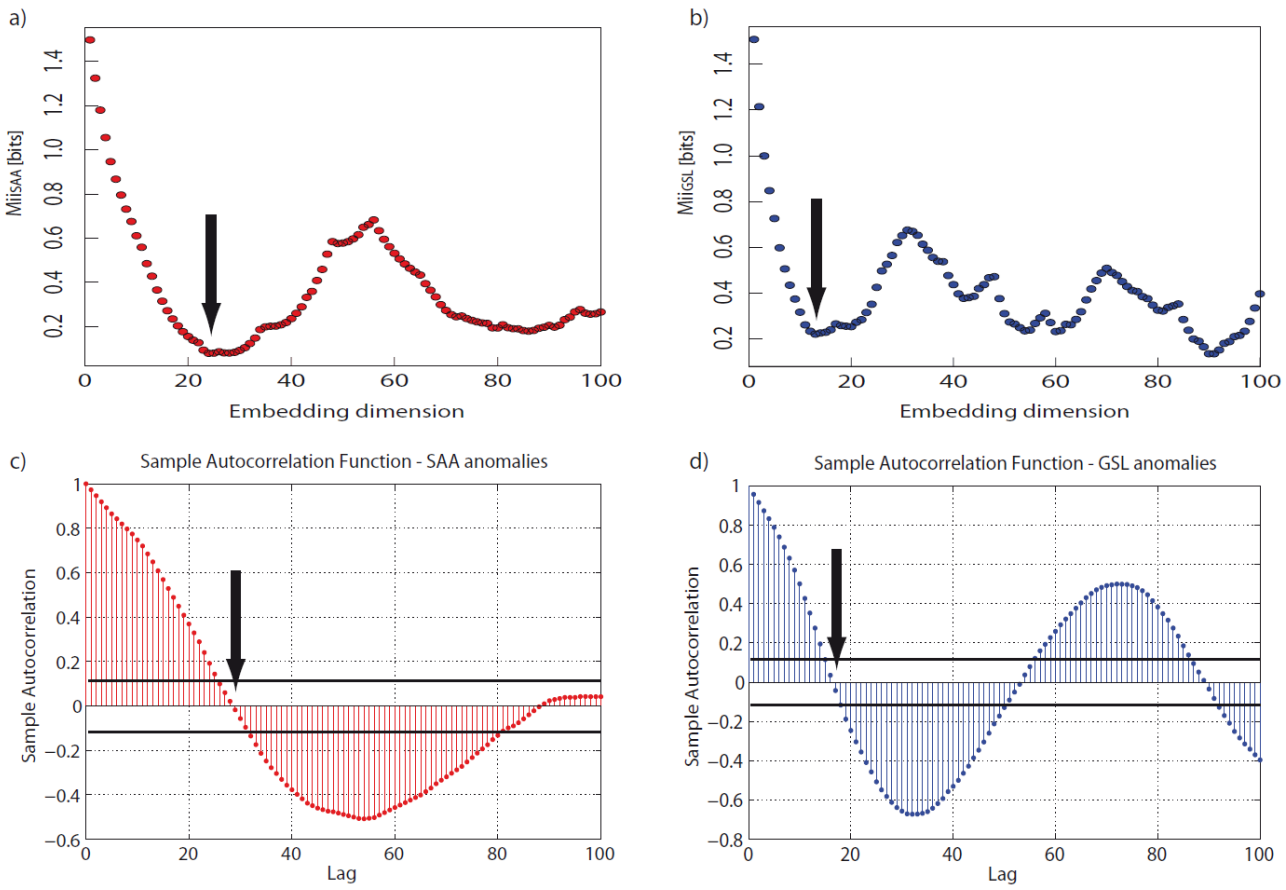
1.1 Supplementary Figures



Supplementary Figure 1. Root mean square (rms) error in function on different damping parameters (λ). The arrows indicate the points where this parameter is optimal, according to the trade-off curves.



Supplementary Figure 2. Transfer entropy (red arrow) between a source J (SAA anomalies) and a target process I (GSL anomalies). Green and blue boxes indicate past states (memory) of both processes, the purple circle indicates the future value of I (GSL anomalies).



Supplementary Figure 3. Mutual information a) for SAA anomalies computed from Jackson et al. [2000] and b) for GSL anomalies. The arrows indicate the first minimum, which represents the optimal embedding dimension k. The sample autocorrelation function is also plotted in c) for SAA anomalies from Jackson et al. [2000] and d) for GSL anomalies. Again, the arrows mark the optimal embedding dimension given by the first zero of the autocorrelation function.

1.2 Supplementary Tables

Supplementary Table 1. Different sets of optimal parameters: number of bins S and embedding dimension k, for SAA and GSL anomalies. As it can be observed, the highest values of embedding dimension are reported when S=3. This means that the memory of system is longer, which can be due to there is less information when the number of possible states of the system is small.

a)

OPTIMIZATION PARAMETERS				
	SAA surface extent			GSL
	Jackson et al. [2000]	*Gubbins et al.* [2006]	*Finlay* [2008]	
S	5	5	5	5
k	23	22	21	15

b)

OPTIMIZATION PARAMETERS				
	SAA surface extent			GSL
	Jackson et al. [2000]	*Gubbins et al.* [2006]	*Finlay* [2008]	
S	4	4	4	4
k	24	26	26	13

c)

OPTIMIZATION PARAMETERS				
	SAA surface extent			GSL
	Jackson et al. [2000]	*Gubbins et al.* [2006]	*Finlay* [2008]	
S	3	3	3	3
k	35	30	30	16

d)

OPTIMIZATION PARAMETERS				
	SAA surface extent			GSL
	Jackson et al. [2000]	*Gubbins et al.* [2006]	*Finlay* [2008]	
S	5	9	9	11
k	23	26	23	13

Supplementary Table 2. Transfer entropy and statistical significance (in brackets) from SAA to GSL anomalies and from GSL to SAA anomalies. The different tables represent the values of TE with different sets of optimal number of bins S and optimization embedding parameters (see Table 1S). In bold the TE calculated with l=1 and optimal k is shown, in bold and italics when k=l and only in italics when k and l are the embedding dimension optimized for each time series (k=kSAA and l=kGSL for TEGSL->SAA; k=kGSL and l=kSAA for TESAA->GSL).

a)	*Jackson et al.* [2000]	*Gubbins et al.* [2006]	*Finlay* [2008]
$TE_{SAA->GSL}$ [bits]	**0.076 (13%)**	**0.090 (53%)**	**0.083 (36%)**
	0.13 (26%)	***0.12 (3%)***	***0.12 (3%)***
	0.14 (11%)	*0.14 (12%)*	*0.13 (4%)*
$TE_{GSL->SAA}$ [bits]	**0.045 (90%)**	**0.049 (94%)**	**0.056 (98%)**
	0.082 (4%)	***0.11 (13%)***	***0.12 (21%)***
	0.082 (27%)	*0.11 (50%)*	*0.11 (36%)*

With parameters given in Table 1Sa

b)	*Jackson et al.* [2000]	*Gubbins et al.* [2006]	*Finlay* [2008]
$TE_{SAA->GSL}$ [bits]	**0.091 (85%)**	**0.10 (98%)**	**0.11 (99%)**
	0.15 (6%)	***0.18 (39%)***	***0.18 (43%)***
	0.20 (15%)	*0.21 (21%)*	*0.20 (9%)*
$TE_{GSL->SAA}$ [bits]	**0.039 (67%)**	**0.027 (48%)**	**0.027 (48%)**
	0.13 (24%)	***0.12 (13%)***	***0.12 (15%)***
	0.12 (77%)	*0.093 (31%)*	*0.091 (26%)*

With parameters given in Table 1Sb

c)	*Jackson et al.* [2000]	*Gubbins et al.* [2006]	*Finlay* [2008]
$TE_{SAA->GSL}$ [bits]	**0.038 (35%)**	**0.043 (68%)**	**0.042 (63%)**
	0.084 (1%)	***0.12 (19%)***	***0.12 (20%)***
	0.13 (0%)	*0.14 (1%)*	*0.14 (1%)*
$TE_{GSL->SAA}$ [bits]	**0.019 (69%)**	**0.016 (48%)**	**0.018 (55%)**
	0.076 (3%)	***0.097 (24%)***	***0.098 (20%)***
	0.068 (58%)	*0.075 (41%)*	*0.075 (39%)*

With parameters given in Table 1Sc

d)	*Jackson et al.* [2000]	*Gubbins et al.* [2006]	*Finlay* [2008]
$TE_{SAA->GSL}$ [bits]	**0.021 (46%)**	**0.021 (27%)**	**0.021 (30%)**
	0.021 (1%)	***0.030 (1%)***	***0.025 (2%)***
	0.032 (3%)	0.044 (2%)	0.043 (6%)
$TE_{GSL->SAA}$ [bits]	**0.045 (59%)**	**0.026 (32%)**	**0.063 (83%)**
	0.082 (0%)	***0.061 (0%)***	***0.092 (0%)***
	0.082 (7%)	*0.061 (4%)*	*0.092 (4%)*

With parameters given in Table 1Sd

References

Finlay, C.C. (2008), Historical variation of the geomagnetic axial dipole, Phys. Earth Planet. Interiors. 170, 1-14.

Gubbins, D., Jones, A.L., Finlay, C.C. (2006), Fall in Earth's Magnetic Field is erratic, Science, 312 (5775), 900-902.

Jackson, A., Jonkers, A.R.T., Walker, M.R. (2000), Four centuries of geomagnetic secular variation from historical records, Philos. Trans. R. Soc. Lond. A 358 (1768), 957-990.

Marschinski, R. and Kantz, H. (2002), Analysing the information flow between financial time series: An improved estimator for transfer entropy, Eur. Phys. J. B., 30, 275–281, doi: 10.1140/epjb/e2002-00379-2.