Discrete Chi-square Method can model and forecast El Niño between 1870 and 2024

Lauri Jetsu[1,*]

[1]Department of Physics, P.O. Box 64, FI-00014 University of Helsinki, Finland
[*]lauri.jetsu@helsinki.fi

ABSTRACT

Here, we show that our Discrete Chi-square Method (DCM) can model and forecast the El Niño phenomenon between the years 1870 and 2024. This result may be the "holy grail" of climatology. The solar forcing most probably causes the detected 5.850 ± 0.085, 12.82 ± 0.40 and 19.30 ± 0.83 year signals. These strict periodicities indicate that the solar cycle is caused by the planets. Our results confirm that the DCM is an ideal time series analysis method for detecting an unknown number of signals superimposed on an unkown trend.

Keywords El Niño, Solar cycle, Time series analysis, Discrete Chi-square Method (DCM)

The El Niño phenomenon is considered unpredictable.[1–8] Our analysed data are the yearly means in the The Seasonal Mean Nino 4 HadISST1.1 (NOAA PSL) sample.[1] There were no values for the last four months of 2025. The full data contains $n = 155$ yearly means between 1870 and 2024.

Method

The Discrete Chi-Square Method (DCM) was formulated in Paper I[9]. This method can detect an unknown number of signals superimposed on an unknown trend. We have applied the DCM to Algol's eclipse epochs (Paper II[10]) and the sunspot data (Paper III[11]). Here, we apply the most recently published DCM version (Paper IV[12]).

The data are $y_i = y(t_i) \pm \sigma_i$, where t_i are the observing times and σ_i are the errors $(i = 1, 2, ..., n)$. The data window is $\Delta T = t_n - t_1$. The mid point is $t_{\rm mid} = t_1 + \Delta T/2$.

The DCM model is

$$g(t) = g_{K_1,K_2,K_3}(t) = h(t) + p(t). \tag{1}$$

The periodic function is

$$h(t) = h(t, K_1, K_2) = \begin{cases} 0, & \text{if } K_1 = 0 \\ \sum_{i=1}^{K_1} h_i(t), & \text{if } K_1 \geq 1 \end{cases} \tag{2}$$

$$h_i(t) = \sum_{j=1}^{K_2} B_{i,j} \cos(2\pi j f_i t) + C_{i,j} \sin(2\pi j f_i t). \tag{3}$$

It is a sum of K_1 harmonic $h_i(t)$ signals having frequencies f_i. The aperiodic trend function is

$$p(t) = p(t, K_3) = \begin{cases} 0, & \text{if } K_3 = -1 \\ \sum_{k=0}^{K_3} p_k(t), & \text{if } K_3 = 0, 1, 2, ... \end{cases} \tag{4}$$

where

$$p_k(t) = M_k \left[\frac{2(t - t_{\rm mid})}{\Delta T} \right]^k. \tag{5}$$

The $h_i(t)$ signals are superimposed on the aperiodic K_3 order polynomial trend $p(t)$. Function $h(t)$ repeats itself in time, while the function $p(t)$ does not. The details of DCM formalism are explained in Paper IV, and are therefore not repeated here. We search for periods between 3 and 100 years.

[1]The data from https://psl.noaa.gov/data/timeseries/month/data/nino4.long.anom.data were retrieved on January 29th, 2026.

Results

Trend

We detect significant signals only from the weighted data. Therefore, the trend and the forecasts are solved from the weighted data. The weighted DCM analysis results for all data are given in Table S1. For one, two, three and four signals, the best trend alternative is the linear trend $K_1 = 1$. We use this $K_1 = 1$ trend in our DCM analysis of El Niño data. For the $\mathscr{M}$=3 model, the positive linear trend is $2M_1 = 0.56 \pm 0.20$ C^{o}during $\Delta T = 154$ years (Equation 5), like in all $\mathscr{M}$=1-4 models.

Forecasting data

We use the first half of all data $(n = 78)$ to forecast the second half $(n = 77)$. The results of weighted DCM analysis are given in Table S2. The periods of the three strongest signals are 5.850 ± 0.085, 12.82 ± 0.40 and 19.30 ± 0.83 years (Model $\mathscr{M}$=3). The forecasts of the two and three signals models, open circles, are shown in Figures S1 and S2. For the second half of the data, the forecast χ^2 values for the one, two, three and four signals models are 421, 336, 370 and 428, respectively. Hence, the two signal $\mathscr{M}$=2 model gives the best forecast.

All data

The DCM analysis results for all weighted data are given in Table S3. The three signals $\mathscr{M}$=3 model is the best. The four signal model is unstable (UM). The three and four signal models are shown in Figures S3 and S4. For both alternatives, the predicted values, open circles, after 2024 are given in Table S4. Both alternatives give essentially the same forecast.

For the sake of consistency, we also give the results of non-weighted analysis of all data (Table S5). We detect the same signals as from non-weighted data, but the signal significances Q_F are lower. This confirms that the data error information is crucial.

Discussion

Our DCM can model and forecast the El Niño phenomenon. The solar forcing is the only possible explanation for these results. Within their error limits, the detected three El Niño signals could represent the Hale cycle[13], the solar cycle[14,15] and half the solar cycle. These strict El Niño periodicities indicate that the planets could cause a stationary and deterministic solar cycle.[11] A stochastic and non-stationary solar cycle[16,17] could not cause the detected regular El Niño periodicities.

References

1. Thirumalai, K. *et al.* Future increase in extreme el niño supported by past glacial changes. *Nature* **634**, 374–380 (2024).
2. Lu, Z. *et al.* Increased frequency of multi-year el niño–southern oscillation events across the holocene. *Nat. Geosci.* 1–7 (2025).
3. Cai, W. *et al.* Increasing frequency of extreme el niño events due to greenhouse warming. *Nat. climate change* **4**, 111–116 (2014).
4. Hu, R. *et al.* Predicting the 2023/24 el niño from a multi-scale and global perspective. *Commun. Earth & Environ.* **5**, 675 (2024).
5. Liu, Y. *et al.* Enhanced multi-year predictability after el niño and la niña events. *Nat. communications* **14**, 6387 (2023).
6. Ludescher, J. *et al.* Improved el niño forecasting by cooperativity detection. *Proc. Natl. Acad. Sci.* **110**, 11742–11745 (2013).
7. Timmermann, A. *et al.* El niño–southern oscillation complexity. *Nature* **559**, 535–545 (2018).
8. Liang, X. S. *et al.* El niño modoki can be mostly predicted more than 10 years ahead of time. *Sci. Reports* **11**, 17860 (2021).
9. Jetsu, L. Discrete Chi-square Method for Detecting Many Signals. *The Open J. Astrophys.* **3**, 4, DOI: 10.21105/astro.2002.03890 (2020). 2002.03890.
10. Jetsu, L. Say Hello to Algol's New Companion Candidates. *Astrophys. J.* **920**, 137, DOI: 10.3847/1538-4357/ac1351 (2021). 2005.13360.
11. Jetsu, L. Do the planets cause the sunspot cycle? *arXiv e-prints* arXiv:2311.08317, DOI: 10.48550/arXiv.2311.08317 (2025). Submitted to Scientific Reports, August 13th, 2025, 2311.08317.
12. Jetsu, L. Discrete Chi-Square Method can model and forecast complex time series. *arXiv e-prints* arXiv:2509.01540, DOI: 10.48550/arXiv.2509.01540 (2026). Submitted to Journal of Statistical Theory and Practice, January 31st, 2026, 2509.01540.

13. Hale, G. E., Ellerman, F., Nicholson, S. B. & Joy, A. H. The Magnetic Polarity of Sun-Spots. *Astrophys. J.* **49**, 153, DOI: 10.1086/142452 (1919).

14. Schwabe, M. Sonnenbeobachtungen im Jahre 1843. Von Herrn Hofrath Schwabe in Dessau. *Astron. Nachrichten* **21**, 233 (1844).

15. Wolf, R. Bericht über neue Untersuchungen über die Periode der Sonnenflecken und ihrer Bedeutung von Herrn Prof. Wolf. *Astron. Nachrichten* **35**, 369, DOI: 10.1002/asna.18530352504 (1852).

16. Charbonneau, P. Dynamo models of the solar cycle. *Living Rev. Sol. Phys.* **17**, 1–104, DOI: 10.1007/s41116-020-00025-6 (2020).

17. Usoskin, I. G. A history of solar activity over millennia. *Living Rev. Sol. Phys.* **20**, 2, DOI: 10.1007/s41116-023-00036-z (2023).

Acknowledgements

This work has made use of NASA's Astrophysics Data System (ADS) services.

Author contributions statement

L.J. peformed this research and wrote the manuscript.

Consent for publication

The author grants permission for publication in Nature

Additional information

Supplementary information The online version contains supplementary material available.

Data availability

All data files, DCM Python code, DCM control files and Python code are stored to the https://zenodo.org/records/18595333. Our Supplementary material contains instructions for repeating the whole DCM analysis.

Declarations

Competing interests

The author declares no competing interests.

Correspondence and requests for materials should be addressed to L.J.

Funding No funding was received.

Ethical approval and consent to participate Not applicable.

Supplementary material
Discrete Chi-square Method can model and forecast El Niño between 1870 and 2024

L. Jetsu
Department of Physics, P.O. Box 64, FI-00014 University of Helsinki, Finland

The $\mathscr{M} = 1$ analysis in Table S1 can be repeated using these two commands

```
cp longC14K110.dat dcm.dat
python dcm.py
```

The whole analysis can be repeated in similar fashion.

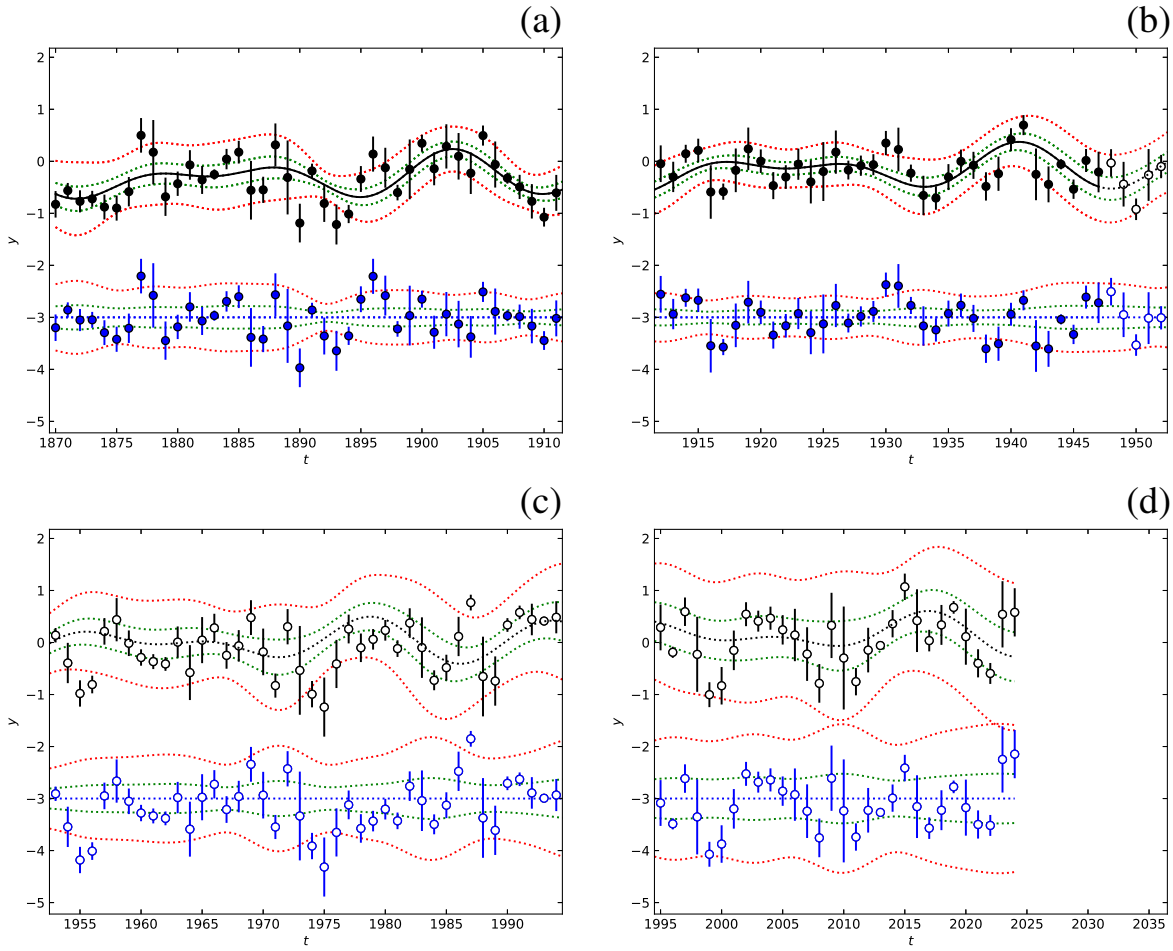

Figure S1. First half forecast (Table S2: $\mathscr{M}$=2): Forecasting data (closed black circles) and forecasted data (open black circles). Green and red dottted lines denote one and three sigma model $g(t)$ errors. Residuals (closed and open blue circles) are offset to level of -3.

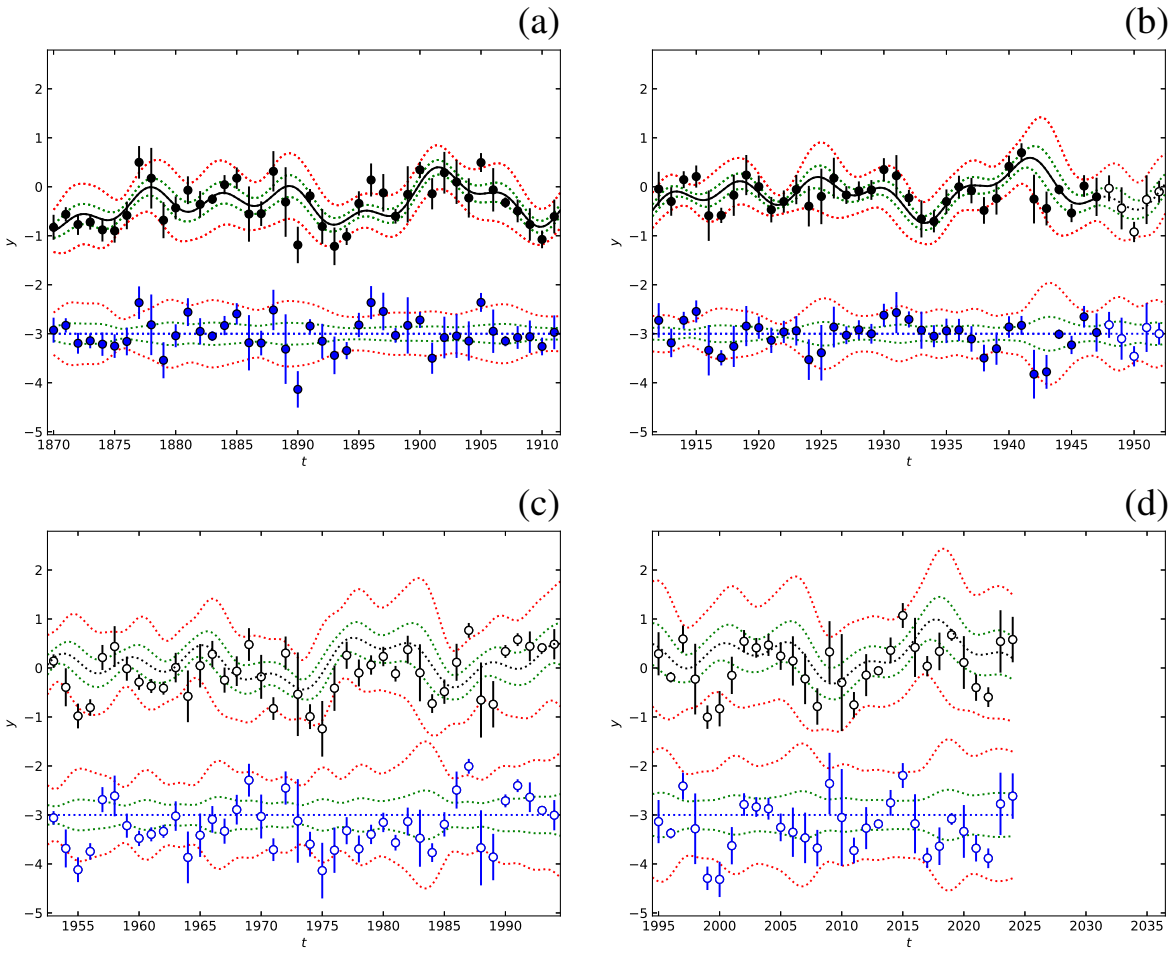

Figure S2. First half forecast (Table S2: $\mathscr{M}$=3). Otherwise as in Figure S1.

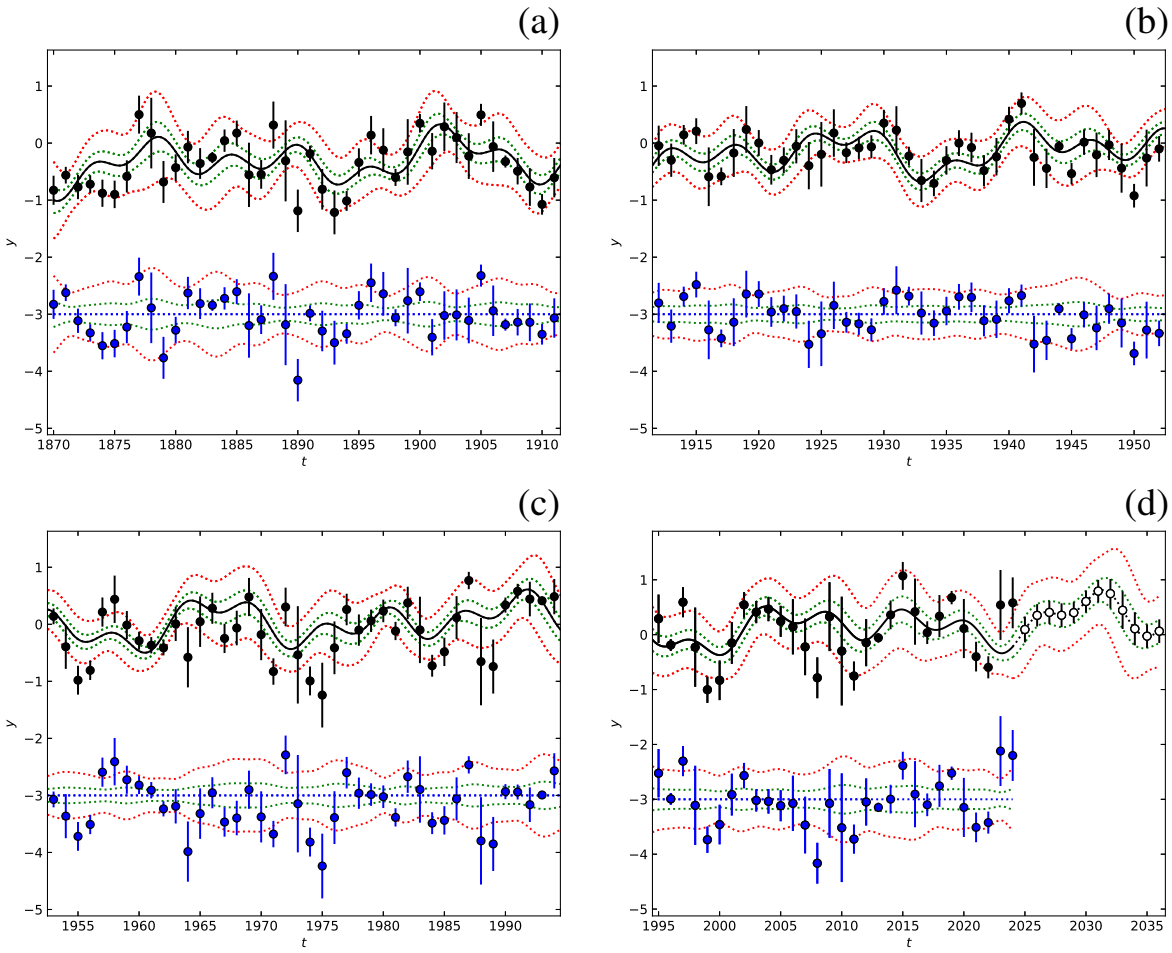

Figure S3. All data model $\mathscr{M}$=3 in Table S3. Otherwise as in Figure S1.

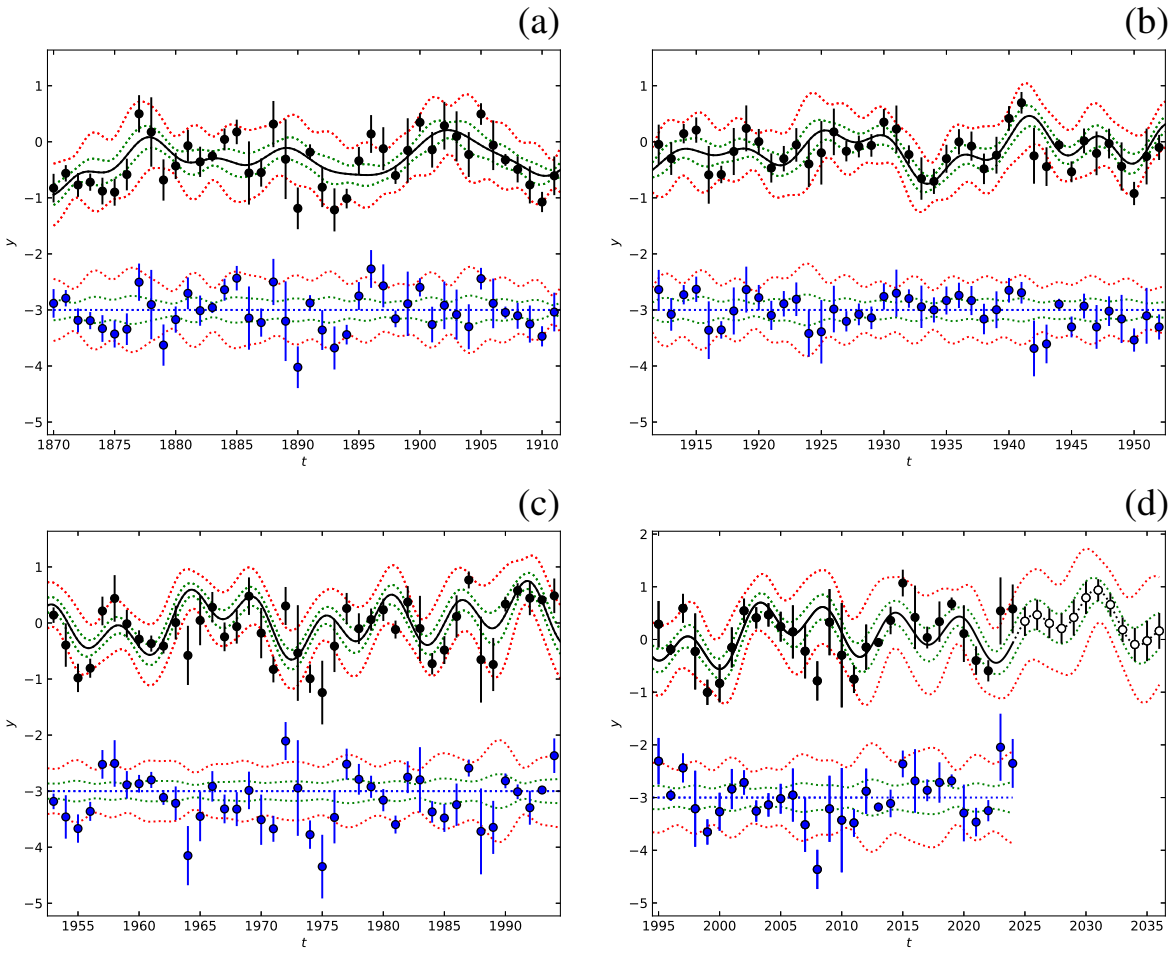

Figure S4. All data model $\mathscr{M}$=4 in Table S3. Otherwise as in Figure S1.

Table S1. Trend in all data. Weighted DCM analysis performed between $P_{\rm min} = 3$ and $P_{\rm max} = 100$. (1) $\mathscr{M}$ model, (2) η free parameters and χ^2 of model, (3) First signal: P_1 period, A_1 peak to peak amplitude and $t_{\rm min,1}$ primary minimum epoch. (4-6) Next signals. (7-8) Fisher-test: F_χ test statistic and $Q_{\rm F}$ critical level. (9) Control file. Arrows point to best model in Fisher-test comparison.

		Period analysis				Fisher-test		
		Data: Original weighted data ($n = 155, \Delta T = 154$: `Clong.dat`)						
$\mathscr{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathscr{M}$=2	$\mathscr{M}$=3	
	η (-)	A_1 (C^o)	A_2 (C^o)	A_3 (C^o)	A_4 (C^o)	F_χ (-)	F_χ (-)	Control file
	χ^2 (-)	$t_{\rm min,1}$ (y)	$t_{\rm min,1}$ (y)	$t_{\rm min,1}$ (y)	$t_{\rm min,1}$ (y)	Q_F (-)	Q_F (-)	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
One signal								
$\mathscr{M}$=1		12.67±0.33	-	-	-	↑	↑	
$\mathscr{M}_{1,1,0,\chi^2}$	4	0.62±0.15	-	-	-	26.9	14.0	`longC14K111.dat`
	516	1871.2±1.3	-	-	-	6.7×10^{7}	2.7×10^{-6}	
$\mathscr{M}$=2		12.65±0.20	-	-	-	-	←	
$\mathscr{M}_{1,1,1,\chi^2}$	5	0.51±0.14	-	-	-	-	1.02	`longC14K111.dat`
	437	1871.3±1.3	-	-	-	-	0.31	
$\mathscr{M}$=3		12.65±0.43			-	-	-	
$\mathscr{M}_{1,1,2,\chi^2}$	6	0.52±0.12			-	-	-	`longC41K112.dat`
	434	1871.4±1.1			-	-	-	
Two signals								
						$\mathscr{M}$=5	$\mathscr{M}$=6	
$\mathscr{M}$=4		12.70±0.17	26.8±2.0	-	-	↑	↑	
$\mathscr{M}_{2,1,0,\chi^2}$	7	0.62±0.12	0.39±0.16	-	-	27.4	13.6	`longC14K210.dat`
	456	1870.8±1.0	1870.3±3.4	-	-	5.6×10^{-7}	3.8×10^{-6}	
$\mathscr{M}$=5		5.677±0.089	12.64±0.31	-	-	-	←	
$\mathscr{M}_{2,1,1,\chi^2}$	8	0.393±0.086	0.47±0.13	-	-	-	0	`longC14K211.dat`
	384	1875.65±.56	1871.3±1.2	-	-	-	1	
$\mathscr{M}$=6		5.677±0.077	12.64±0.42		-	-	-	
$\mathscr{M}_{2,1,2,\chi^2}$	9	0.388±0.093	0.47±0.12		-	-	-	`longC41K212.dat`
	384	1875.66±0.76	1871.4±1.2		-	-	-	
Three signals								
						$\mathscr{M}$=8	$\mathscr{M}$=9	
$\mathscr{M}$=7		5.619±0.091	12.74±0.48	21.83±0.81	-	↑	↑	
$\mathscr{M}_{3,1,0,\chi^2}$	10	0.41±0.13	0.52±0.11	0.40±0.12	-	29.2	15.0	`longC14K310.dat`
	407	1871.04±0.77	1870.7±1.7	1870.0±3.2	-	2.6×10^{-7}	1.2×10^{-6}	
$\mathscr{M}$=8		5.662±0.077	12.77±0.12	21.3±1.5	-	-	←	
$\mathscr{M}_{3,1,1,\chi^2}$	11	0.40±0.13	0.47±0.10	0.37±0.12	-	-	0.82	`longC14K311.dat`
	338	1870.3±0.62	1870.4±0.87	1870.8±3.2	-	-	0.36	
$\mathscr{M}$=9		5.661±0.073	12.76±0.22	21.2±1.2	-	-	-	
$\mathscr{M}_{3,1,2,\chi^2}$	12	0.40±0.11	0.49±0.10	0.39±0.12	-	-	-	`longC41K312.dat`
	336	1870.3±0.74	1870.6±1.2	1871.1±2.8	-	-	-	
Four signals								
						$\mathscr{M}$=11	$\mathscr{M}$=12	
$\mathscr{M}$=10		5.597±0.066	12.77±0.19	19.4±1.6 IF	21.5±2.1 IF	↑	↑	
$\mathscr{M}_{4,1,0,\chi^2}$	13	0.44±0.10	0.57±0.11	0.4±1.3 AD	0.4±1.4 *AD*	28.1	14.2	`longC14K410.dat`
UM	362	1871.56±0.57	1870.3±1.2	1872.8±3.7	1871.8±3.3	4.5×10^{-7}	2.5×10^{-6}	
$\mathscr{M}$=11		5.496±0.093 IF	5.66±0.13 IF	12.76±0.15	21.1±1.5	-	←	
$\mathscr{M}_{4,1,1,\chi^2}$	14	0.3±1.8 AD	0.4±1.8 AD	0.49±0.13	0.39±0.11	-	0.45	`longC14K411.dat`
UM	300	1873.8±1.2	1870.2±1.2	1870.60±0.97	1871.3±2.1	-	0.50	
$\mathscr{M}$=12		5.497±0.087 IF	0.566±0.062 IF	12.75±0.17	21.0±1.0	-	-	
$\mathscr{M}_{4,1,2,\chi^2}$	15	0.34±0.94 AD	0.42±0.90 AD	0.50±0.11	0.398±0.084	-	-	`longC41K412.dat`
UM	299	1873.8±1.0	1870.18±0.99	1870.6±0.97	1871.4±2.2	-	-	

Table S2. Periods in first half data. Otherwise as in Table S1.

		Period analysis				Fisher-test			
		Data: Original weighted data ($n = 78, \Delta T = 77$: `Chalf.dat`)							
$\mathscr{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathscr{M}$=2	$\mathscr{M}$=3	$\mathscr{M}$=4	
	η (-)	A_1 (C^o)	A_2 (C^o)	A_3 (C^o)	A_4 (C^o)	F_χ (-)	F_χ (-)	F_χ(-)	Control file
	χ^2 (-)	$t_{\mathrm{min},1}$ (y)	$t_{\mathrm{min},1}$ (y)	$t_{\mathrm{min},1}$ (y)	$t_{\mathrm{min},1}$ (y)	Q_F (-)	Q_F (-)	Q_F (-)	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
One signal									
$\mathscr{M}$=1		19.1±1.2	-	-	-	$\leftarrow$	$\uparrow$	$\uparrow$	
$\mathscr{M}_{1,1,1,\chi^2}$	5	0.44±0.12	-	-	-	5.62	5.95	5.78	`halfC14K111.dat`
	168	1873.6±2.0	-	-	-	0.0016	5.0×10^{-5}	7.5×10^{-6}	
Two signals									
$\mathscr{M}$=2		12.82±0.54	18.78±0.65	-	-	-	$\leftarrow$	$\uparrow$	
$\mathscr{M}_{2,1,1,\chi^2}$	8	0.44±0.12	0.53±0.13	-	-	-	5.3	4.91	`halfC14K211.dat`
	135	1870.4±1.4	1874.3±1.9	-	-	-	0.0025	0.00034	
Three signals									
$\mathscr{M}$=3		5.580±0.085	12.82±0.40	19.32±0.83	-	-	-	$\leftarrow$	
$\mathscr{M}_{3,1,1,\chi^2}$	11	0.42±0.14	0.488±0.087	0.499±0.096	-	-	-	3.88	`halfC14K311.dat`
	109	1875.01±0.51	1870.2±1.4	1873.5±2.3	-	-	-	0.012	
Four signals									
$\mathscr{M}$=4		3.437±0.052 IF	5.70±0.13IF	12.71±x0.17	21.2±1.2	-	-	-	
$\mathscr{M}_{4,1,1,\chi^2}$	14	0.30±0.12 AD	0.40±0.31 AD	0.46±0.13	0.440±0.096	-	-	-	`halfC14K411.dat`
	92	1872.51±0.66	1875.6±1.2	1870.8±1.1	1870.9±2.9	-	-	-	

Table S3. Periods in all weighted data. Otherwise as in Table S1.

		Period analysis				Fisher-test			
		Data: Original weighted data ($n = 155, \Delta T = 154$: `Clong.dat`)							
$\mathscr{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathscr{M}$=2	$\mathscr{M}$=3	$\mathscr{M}$=4	
	η (-)	A_1 (C^o)	A_2 (C^o)	A_3 (C^o)	A_4 (C^o)	F_χ (-)	F_χ (-)	F_χ(-)	Control file
	χ^2 (-)	$t_{\mathrm{min},1}$ (y)	$t_{\mathrm{min},1}$ (y)	$t_{\mathrm{min},1}$ (y)	$t_{\mathrm{min},1}$ (y)	Q_F (-)	Q_F (-)	Q_F (-)	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
One signal									
$\mathscr{M}$=1		12.65±0.20	-	-	-	$\uparrow$	$\uparrow$	$\uparrow$	
$\mathscr{M}_{1,1,1,\chi^2}$	5	0.51±0.14	-	-	-	12.4	6.98	7.10	`longC14K111.dat`
	437	1871.3±1.3	-	-	-	2.7×10^{-7}	1.6×10^{-6}	1.9×10^{-8}	
Two signals									
$\mathscr{M}$=2		5.677±0.089	12.64±0.31	-	-	-	$\uparrow$	$\uparrow$	
$\mathscr{M}_{2,1,1,\chi^2}$	8	0.393±0.086	0.47±0.13	-	-	-	6.48	6,53	`longC14K211.dat`
	384	1875.65±0.56	1871.3±1.2	-	-	-	0.00038	4.1×10^{-6}	
Three signals									
$\mathscr{M}$=3		5.662±0.077	12.78±0.12	21.3±1.5	-	-	-	$\uparrow$	
$\mathscr{M}_{3,1,1,\chi^2}$	11	0.40±0.13	0.47±0.10	0.37±0.12	-	-	-	5.91	`longC14K311.dat`
	338	1870.27±0.62	1870.36±0.88	1870.8±3.3	-	-	-	0.00079	
Four signals									
$\mathscr{M}$=4		5.496±0.093 IF	5.66±0.13 IF	12.76±0.15	21.09±1.5	-	-	-	
$\mathscr{M}_{4,1,1,\chi^2}$	14	0.3±1.8 AD	0.4±1.8 AD	0.491±0.13	0.39±0.11	-	-	-	`longC14K411.dat`
UM	300	1873.8±1.2	1870.2±1.1	1870.6±0.97	1871.3±2.	-	-	-	

Table S4. Forecasts (Figures S3 and S4: open circles).

	Figure S3		Figure S4	
t	y	σ	y	σ
(y)	(C^o)	(C^o)	(C^o)	(C^o)
2025	0.09	0.24	0.35	0.28
2026	0.35	0.20	0.47	0.29
2027	0.41	0.21	0.31	0.28
2028	0.35	0.22	0.21	0.28
2029	0.40	0.19	0.42	0.34
2030	0.60	0.20	0.79	0.31
2031	0.79	0.20	0.94	0.21
2032	0.75	0.27	0.66	0.21
2033	0.45	0.34	0.18	0.23
2034	0.11	0.30	-0.10	0.31
2035	-0.03	0.21	-0.03	0.38
2036	0.06	0.22	0.16	0.34

Table S5. Periods in all non-weighted data. Otherwise as in Table S1.

		Period analysis				Fisher-test			
		Data: Original non-weighted data ($n = 155, \Delta T = 154$: `Rlong.dat`)							
$\mathscr{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathscr{M}$=2	$\mathscr{M}$=3	$\mathscr{M}$=4	
	η (-)	A_1 (-)	A_2 (-)	A_3 (-)	A_4 (-)	F_R (-)	F_R (-)	F_R(-)	Control file
	R (-)	$t_{\mathrm{min},1}$ (y)	$t_{\mathrm{min},1}$ (y)	$t_{\mathrm{min},1}$ (y)	$t_{\mathrm{min},1}$ (y)	Q_F (-)	Q_F (-)	Q_F (-)	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
One signal									
$\mathscr{M}$=1		12.71±0.12	-	-	-	←	←	←	
$\mathscr{M}_{1,1,1,R}$	5	0.389±0.074	-	-	-	3.67	3.77	1.11	`longR14K110.dat`
	27.1	1871.19±0.86	-	-	-	0.013	0.0016	0.36	
Two signals									
$\mathscr{M}$=2		5.651±0.030	12.69±0.15	-	-	-	←	←	
$\mathscr{M}_{2,1,1,R}$	8	0.319±0.064	0.384±0.088	-	-	-	3.77	-0.092	`longR14K211.dat`
	25.2	1870.26±0.41	1871.3±1.1	-	-	-	0.0016	1	
Three signals									
$\mathscr{M}$=3		5.650±0.032	9.12±0.12	12.69±0.27	-	-	-	←	
$\mathscr{M}_{3,1,1,R}$	11	0.327±0.084	0.299±0.087	0.381±0.079	-	-	-	-3.50	`longR14K311.dat`
	23.4	1870.28±0.46	1872.5±1.2	1871.4±1.3	-	-	-	1	
Four signals									
$\mathscr{M}$=4		5.496±0.093 IF	5.662±1.3 IF	12.76±0.16	21.1±1.5	-			
$\mathscr{M}_{4,1,1,R}$	14	0.3±1.9 AD	0.4±1.8 AD	0.491±0.49±0.13	0.39±0.11	-			`longR14K411.dat`
UM	25.3	1873.8±1.2	1870.2±1.2	1870.6±0.97	1871.3±2.1	-			