Draft version 2026-05-26
Typeset using LaTeX **twocolumn** style in AASTeX63

Discrete Chi-Square Method discovers solar forcing signals in El Niño time series

L. Jetsu [1]

[1] *Department of Physics, P.O. Box 64, FI-00014 University of Helsinki, Finland*

(Received 2026-05-26 20:27:58+03:00)

ABSTRACT

Discrete Chi-square Method (DCM) can detect multiple signals superimposed on an arbitrary trend. DCM's backbone is Gauß-Markov theorem that Least Squares (LS) is the best unbiased estimator for linear regression models. DCM is robust because it computes a massive number of linear model LS fits. Discrete Fourier Transform (DFT) and other frequency-domain methods have numerous application limitations. None of those limitations constrains DCM. Fisher-test provides signal significance estimates and identifies the best DCM model, which is validated by Forecast-test. Simulations verify the groundbreaking Window Dimension Effect (WD-effect): *"For any sample window ΔT, DCM inevitably detects the correct $p(t)$ trend and $h(t)$ signal(-s) when sample size n and/or data accuracy σ increase."* WD-effect "sees through time". DCM's model analytical solution is ill-posed. We present a computational well-posed solution. Mainstream considers El Niño phenomenon chaotic. Usual forecasts are probabilistic, not deterministic. We use El Niño time series to stress-test DCM. It detects the multi-periodic "Big Wave" superimposed on linear global warming trend. This gives accurate El Niño forecasts. Our real-time forecast outperforms those of official agencies. Only solar forcing, not chaotic ocean-atmosphere coupling, can cause the "Big Wave". DCM has now detected multi-periodicity in the Pacific Ocean and sunspot record. Only a strictly multi-periodic solar dynamo, not the mainstream stochastic dynamo, can cause this. Planetary clockwork tidal forces probably drive solar dynamo. Future El Niño models must integrate astrophysical cycles with chaotic climatological fluid dynamics. Validating our analysis now can save trillions (USD) in El Niño damages.

Keywords: Sun: solar-terrestrial relations, activity, magnetic fields, Methods: statistical, data analysis, Celestial mechanics: tides

1. INTRODUCTION

Jetsu (2020) formulated DCM. This current follow-up paper aims to demonstrate that DCM outperforms all other parametric frequency-domain time series analysis methods. We validate this claim in our stress test, where DCM is applied to the notorious El Niño time series. The mainstream El Niño dynamical models rely on supercomputers to simulate the chaotic and nonlinear coupling of the ocean and atmosphere (Timmermann et al. 2018; Hu et al. 2024). The current short-term El Niño forecasts are usually probabilistic and typically fail beyond 18 months (Chen et al. 2004; Lenssen et al. 2024). In contrast, our stress test will show that DCM can provide reliable, deterministic, long-term El Niño forecasts.

Cheng et al. (2015) stated, "Forecasting the evolution of complex systems is noted as one of the 10 grand challenges of modern science." We will show that DCM can detect the signal(-s) and the trend in complex time series, and can also forecast the time series evolution. Our DCM is based on the LS method, which was originally formulated by Legendre (1805). Gauß (1809) connected this method to the principles of probability and the normal distribution. A few years later, he showed that the LS method gives the best unbiased estimates for the free parameters of linear models, if the zero mean data errors are equal, normally distributed and uncorrelated (Gauß 1821). The extended Gauß-Markov theorem states that the LS method gives the best estimates for the free parameters of linear models, even if the data errors do not pass all the above-mentioned criteria (Wooldridge

lauri.jetsu@helsinki.fi

2010). The Gauß-Markov theorem is the backbone of our DCM. This time series analysis method performs a massive number of LS fits to find the best model for the data.

Hadamard (1902) defined the three conditions for a well-posed problem. The solution for the problem determines these conditions.

C_1 Existence: A solution exists.

C_2 Uniqueness: The solution is unique.

C_3 Stability: Small changes in the input data cause small changes in the solution.

Stability means that the solution behaves predictably as the input varies. In other words, there exists a continuous mapping from the input space to the solution space. If any of these conditions are not satisfied, the problem is classified as ill-posed.

Ill-posed problems arise in many fields of science (Rudin 1976; Tikhonov & Arsenin 1977; Lavrentiev et al. 1986). For example, Piskunov et al. (1990) solved the stellar surface imaging inverse problem using the Tikhonov (1963) regularisation technique. Other typical examples are partial differential equation solutions (Hadamard 1923), non-linear parameter estimation (Bard 1974) and regularisation of inverse problems (Engl et al. 1996; Vogel 2002). Ill-posed problems are encountered in time series analysis (Timmer & König 1995; Bailer-Jones 2012; Berger 2013; Box et al. 2015) because the models for the data are often non-linear (Tong 1990; Tsay & Chen 2018). The unknown free parameters of the model, namely the frequencies, are in the arguments of trigonometric functions representing the signals. This causes non-linearity because the model partial derivatives contain free parameters of the model.

Our DCM is a frequency-domain parametric time series analysis method. Here, we discuss shortly such methods in the order of increasing model complexity. Numerous methods can be applied if the time series is stationary (constant mean and variance). The Fast Fourier Transform (FFT) for evenly spaced data (Cooley & Tukey 1965) and DFT for unevenly spaced data (Lomb 1976; Scargle 1982) are the most widely-used methods. Astronomers have published several extensions of these methods (Zechmeister & Kürster 2009; VanderPlas 2018). FFT and DFT assume that the correct model for the time series is a pure sine. The spectral estimating techniques fail if the data are non-stationary. Therefore, trends changing the sample mean and/or variance must be removed before applying spectral estimation (Nerlove 1964). There are numerous methods for detecting one signal superimposed on a trend, especially for evenly spaced data (Cleveland et al. 1990; Shumway & Stoffer 2006; Brockwell & Davis 2009). Kay & Marple (1981) compared many spectral estimating techniques that can detect more than one signal from evenly spaced data. They discussed how the sample window (ΔT) limits the frequency resolution for detecting many signals and causes leakage that shifts power from one spectral peak to another. Exactly correct frequency values may not be detected because the leakage can shift the periodogram peaks. Ghaderpour et al. (2021) presented different techniques for reducing spectral leakage. They also discussed techniques for detecting many signals in unevenly spaced data, if these signals are pure sines. Astronomers apply DFT pre-whitening technique to detect many signals from unevenly spaced data (Reinhold et al. 2013; Zhu & Jia 2018). The data must be detrended before applying this technique, which detects one frequency at a time, until it detects no new frequency. There are more complex modelling problems. For example, the signals are not necessarily pure sines or the sample window ΔT is shorter than the signal period(-s).

Reliable forecasting is the ultimate test for any time series analysis method (Cheng et al. 2015; Jetsu 2025). Forecasting can work for linear and stationary processes, but it is challenging for non-linear or non-stationary ones. Other forecasting challenges are overfitting and forecast error estimation (Makridakis et al. 1987; Lefrancois 1989; Grushka-Cockayne et al. 2017; Petropoulos et al. 2022).

Our DCM is a frequency-domain parametric time series analysis method. Such methods have their own particular limitations that constrain their applications. Here is our list of those application limitations (AL).

1. Data errors (level of noise) are unknown.
2. Data error information is not utilised.
3. Data must be evenly spaced.
4. Model parameter errors are unknown.
5. Model and forecast errors are unknown.
6. Sample window is shorter than signal period(s).
7. Presence and shape of trend are unknown.
8. Sample window causes leakage.
9. Leakage weakens frequency resolution.
10. Signal shapes are not pure sines.
11. Number of signals is unknown.

12. Correct model alternative is unknown.
13. Signal significances are unknown.
14. Model solution is ill-posed.
15. Complex non-linear model forecasts fail.

We will show that these ALs do not constrain the DCM.

The modelling and forecasting of El Niño is considered extremely challenging (Thirumalai et al. 2024; Lu et al. 2025; Cai et al. 2014; Hu et al. 2024; Liu et al. 2023b; Ludescher et al. 2013; Timmermann et al. 2018; Liang et al. 2021). In the mainstream climatological models, El Niño emerges from the non-linear dynamical coupling between the Pacific Ocean and atmosphere (Bjerknes 1969; Zebiak & Cane 1987). Supercomputers must be used to solve millions of complex fluid dynamic equations of this chaotic interaction (Cane et al. 1986; Barnston et al. 2012). The "spring predictability barrier" reduces forecast accuracy from March to May due to high ocean-atmosphere instability (Webster & Yang 1992; Lau & Yang 1996). Even the best El Niño forecasts fail in 1.5-years (Chen et al. 2004; Zhao et al. 2024).

The annual cost of El Niño damage is roughly one trillion dollars (Callahan & Mankin 2023; Liu et al. 2023a). An accurate El Niño prediction for just one year ahead would save the global economy trillions of dollars by mitigating long-term productivity losses (Hsiang et al. 2011; Xu et al. 2026). If decadal reliable El Niño forecasts were possible, they would allow us to prepare for these climatological disasters well in advance. Such forecasts would stabilise global agriculture and economy be replacing reactive guesswork with proactive planning (Solaraju-Murali et al. 2022).

The total solar irradiance varies by about 0.1% during the solar cycle (Willson & Hudson 1991; Fröhlich & Lean 1998). Some studies have proposed a connection between solar cycle and climate temperature (Friis-Christensen & Lassen 1991; Connolly et al. 2021). Recent research has confirmed that solar irradiance fluctuations have a negligible impact on global warming (Gray et al. 2010; Kopp & Lean 2011; Lockwood 2012). In the mainstream solar dynamo models, the stochastic and non-stationary solar cycle cannot cause any strictly periodic solar forcing (Usoskin 2017; Charbonneau 2020). The alternative planetary tidal dynamo models could provide a precise solar cycle "clocking" mechanism (Klevs et al. 2023; Mouël et al. 2025; Stefani et al. 2025; Jetsu 2025). Our DCM has already detected extremely significant, strictly periodic signals in the sunspot record (Jetsu 2025). Reliable decedal El Niño forecasts may succeed, if the solar forcing is discovered to cause strictly periodic changes in the Niño 4 index data. This would challenge to the mainstream climate and solar dynamo theories.

This study proceeds through following stages. We present DCM formulation (Section 2.1). The WD-effect is discussed (Section 2.2). Complex time series are simulated using DCM model (Sect. 2.3). DFT and its ALs are presented (Section 2.4). We compare our DCM to the renowned DFT. The time series analysis of simulated data sets shows that the ALs of DFT do not constrain DCM (Sections 3.1-3.8). Then, we present the identification of the best DCM model for the data (Section 3.9), DCM significance estimates (Section 3.10), the well-posed computational DCM model solution (Section 3.11) and DCM forecasts (Section 3.12). We summarise why DCM outperforms other frequency-domain parametric time series analysis methods (Section 4). In our stress test, we show that DCM can model and forecast the El Niño time series (Section 5). Many tables and figures are published only in Supplementary material.

2. METHODS

The observations are $y_i = y(t_i) \pm \sigma(t_i)$, where t_i are the observing times and σ_i are the errors ($i = 1, 2, ..., n$). The sample window is $\Delta T = t_n - t_1$. The mid point is $t_{\rm mid} = t_1 + \Delta T/2$. The mean and the standard deviation of all y_i values are denoted with m and s.

2.1. *Discrete Chi-square method (DCM)*

Jetsu (2020) introduced DCM which is a frequency-domain parametric time series analysis method. The primary objective of this current paper is to show that DCM outperforms all other similar methods.

DCM model is

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

where the integer values K_1, K_2 and K_3 are called the model orders. The notation $g_{K_1,K_2,K_3}(t)$ is used to specify these orders. DCM model is a sum of two functions. These functions are the periodic function

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

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

and the aperiodic function

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

where

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

For $k \geq 1$, the $p_k(t)$ function full range is

$$\begin{cases} 2|M_k|, & \text{if } k \text{ is odd} \\ |M_k|, & \text{if } k \text{ is even.} \end{cases} \quad (6)$$

The free parameters of $g(t)$ model are

$$\begin{aligned} \beta &= [\beta_1, \beta_2, ..., \beta_\eta] \\ &= [B_{1,1}, C_{1,1}, f_1, ..., B_{K_1,K_2}, C_{K_1,K_2}, f_{K_1}, \\ &\quad M_0, ..., M_{K_3}]. \end{aligned} \quad (7)$$

The number of free parameters is

$$\eta = K_1 \times (2K_2 + 1) + K_3 + 1. \quad (8)$$

We divide the free parameters β into two groups β_I and β_{II}. The first group are the frequencies

$$_I = [f_1, ..., f_{K_1}]. \quad (9)$$

Due to this group, all free parameters are not eliminated from all partial derivatives $\partial g/\partial\beta_i$. This makes the $g(t)$ model *non-linear*. If the β_I frequencies have constant values, the multipliers $2\pi j f_i$ in Equation 3 become constants. In this case, the model becomes *linear* because the partial derivatives $\partial g/\partial\beta_i$ no longer contain any free parameters. The LS solution for the second group of remaining free parameters

$$\begin{aligned} \beta_{II} &= [B_{1,1}C_{1,1}, ..., B_{K_1,K_2}, C_{K_1,K_2}, \\ &\quad M_0, ..., M_{K_3}] \end{aligned} \quad (10)$$

becomes *unique*. This solution passes the C_1, C_2 and C_3 conditions of a well-posed problem.

Let us assume that we search for periods between $f_{\rm min}$ and $f_{\rm max}$. The non-linear $g(t)$ model becomes linear, if the tested β_I frequencies are fixed to any constant values. The sum $h(t)$ of signals $h_1(t, f_1)$, $h_2(t, f_2)$ ... and $h_{K_1}(t, f_{K_1})$ does not depend on the order in which these signals are added. For example, the two signal $K_1 = 2$ model symmetry is

$$\begin{aligned} h(t) &= h_1(t, f_1) + h_2(t, f_2) \\ &= h_1(t, f_2) + h_2(t, f_1) \end{aligned} \quad (11)$$

Both (f_1, f_2) and (f_2, f_1) combinations give the same value for the $g(t)$ model. Since this symmetry applies to any K_1 number of signals, we compute the linear $g(t)$ models only for all tested frequency combinations

$$f_{\rm max} \geq f_1 > f_2 > ... > f_{K_1} \geq f_{\rm min}. \quad (12)$$

This frequency space symmetry idea reduces CPU consumption dramatically. For example, were this symmetry not exploited, the four signal $K_1 = 4$ search would give $4! = 24$ exactly the same solutions from four-dimensional tested grid in frequency space. This would cause unnecessary use of CPU in the preliminary testing all possible frequency combinations inside a tesseract (a four-dimensional cube). The search for the best frequency combination from $4! = 24$ different non-linear iterations would be a pointless exercise (see Equation 22).

DCM model residuals

$$\epsilon_i = y(t_i) - g(t_i) = y_i - g_i. \quad (13)$$

give the sum of squared residuals

$$R = \sum_{i=1}^{n} \epsilon_i^2, \quad (14)$$

and the Chi-square

$$\chi^2 = \sum_{i=1}^{n} \frac{\epsilon_i^2}{\sigma_i^2}. \quad (15)$$

For every tested $\beta_I = [f_1, f_2, ..., f_{K_1}]$ frequency combination, the LS fit gives DCM test statistic

$$z = z(f_1, f_2, ..., f_{K_1}) = \sqrt{R/n}, \text{ unknown } \sigma_i \quad (16)$$

$$z = z(f_1, f_2, ..., f_{K_1}) = \sqrt{\chi^2/n}, \text{known } \sigma_i \quad (17)$$

computed for a linear model. The value of z is unique. The errors can be unknown in Equation 16 (AL1). For known errors, DCM uses this information in Equation 17 (AL2). In the preliminary long search, we test an evenly spaced grid of $n_{\rm L}$ frequencies between $f_{\rm min}$ and $f_{\rm max}$. This search gives the best frequency candidates $f_{1,\rm mid}, ...f_{K_1,\rm mid}$.

In the final short search, we test a denser grid of $n_{\rm S}$ frequencies within an interval

$$[f_{\rm i,mid} - a, f_{\rm i,mid} + a], \quad (18)$$

where $i = 1, .., K_1$, $a = c(f_{\rm min} - f_{\rm max})/2$ and $c = 0.05$.

The total number of all tested long and short search frequency combinations is

$$\binom{n_f}{K_1} = \frac{n_f!}{K_1!(n_f - K_1)!}, \quad (19)$$

where $n_f = n_L$ and $n_f = n_S$, respectively. The even or uneven data spacing is irrelevant because the LS fit result for every tested frequency combination does not depend on this spacing (AL3).

The global periodogram minimum

$$z_{\min} = z(f_{1,\text{best}}, f_{2,\text{best}}, ..., f_{K_1,\text{best}}) \tag{20}$$

is at the tested frequencies $f_{1,\text{best}}$, $f_{2,\text{best}}, ...,$ $f_{K_1,\text{best}}$. This tested frequency combination gives the best linear model for the data. The periodogam value z is a scalar, which is computed from K_1 frequency values. It is possible to plot the $K_1 = 2$ two signal periodogram $z(f_1, f_2)$ as a map, where f_1 and f_2 are the coordinates, and $z = z(f_1, f_2)$ is the height. For more than two signals, there is no direct graphical z plot because that requires more than three dimensions. Our solution for this dimensional problem is simple. We plot only the following one-dimensional slices of the full periodogram

$$\begin{aligned}
z_1(f_1) &= z(f_1, f_{2,\text{best}}, ..., f_{K_1,\text{best}}) \\
z_2(f_2) &= z(f_{1,\text{best}}, f_2, f_{3,\text{best}}, ..., f_{K_1,\text{best}}) \\
z_3(f_3) &= z(f_{1,\text{best}}, f_{2,\text{best}}, f_3, f_{4,\text{best}}, ..., f_{K_1,\text{best}}) \\
z_4(f_4) &= z(f_{1,\text{best}}, f_{2,\text{best}}, f_{3,\text{best}}, f_4, f_{5,\text{best}}, f_{K_1,\text{best}}) \\
z_5(f_5) &= z(f_{1,\text{best}}, f_{2,\text{best}}, f_{3,\text{best}}, f_{4,\text{best}}, f_5, f_{K_1,\text{best}}) \\
z_6(f_6) &= z(f_{1,\text{best}}, f_{2,\text{best}}, f_{3,\text{best}}, f_{4,\text{best}}, f_{5,\text{best}}, f_6).
\end{aligned} \tag{21}$$

In the above-mentioned $K_1 = 2$ map, the slice $z_1(f_1)$ would represent the height z at the location $(f_1, f_{2,\text{best}})$ when moving along the straight constant line $f_2 = f_{2,\text{best}}$ that crosses the global minimum $z_{\min}$ (Equation 20) at the coordinate point $(f_{1,\text{best}}, f_{2,\text{best}})$. These one-dimensional periodogram slices (Equation 21) allow us "to see inside" the multi-dimensional structure of DCM test statistic z (Equations 16 and 17). This kind of visualisation is important in time series analysis (Su & Wu 2024).

The short search gives the best $f_{1,\text{best}}, ...,$ $f_{K_1,\text{best}}$ frequencies for the data. These frequencies are the *unique* initial values for the first group of free parameters $\beta_{\text{I,initial}} = [f_{1,\text{best}}, ..., f_{K_1,\text{best}}]$ (Equation 9). The linear model for these constant $[f_{1,\text{best}}, ..., f_{K_1,\text{best}}]$ frequencies gives the *unique* initial values for the second group $\beta_{\text{II,initial}}$ of free parameters (Equation 10). The non-linear iteration

$$\beta_{\text{initial}} = [\beta_{\text{I,initial}}, \beta_{\text{II,initial}}] \rightarrow \beta_{\text{final}} \tag{22}$$

gives the final free parameter values β_{final}.

Furlan & Mortarino (2020) emphasise that the analytical error estimates for the non-linear model free parameters can be tricky. They compare different analytical free parameter error estimating methods by using the computational statistical bootstrap technique (Efron & Tibshirani 1986). They conclude that the analytical error estimates become less reliable if the number of free parameters increases. For our non-linear DCM model, we determine the i:th signal parameters

$$P_i = 1/f_i = \text{Period} \tag{23}$$
$$A_i = \text{Peak to peak amplitude} \tag{24}$$
$$t_{\text{i,min,1}} = \text{Deeper primary minimum} \tag{25}$$
$$t_{\text{i,min,2}} = \text{Secondary minimum (if present)} \tag{26}$$
$$t_{\text{i,max,1}} = \text{Higher primary maximum} \tag{27}$$
$$t_{\text{i,max,2}} = \text{Secondary maximum (if present)} \tag{28}$$

and the trend parameters

$$M_k = \text{Polynomial coefficients.} \tag{29}$$

Of these model parameters, all $P_i = 1/f_i$ and all M_k estimates are among the free parameters β. However, the A_i, $t_{\text{i,min,1}}$, $t_{\text{i,min,2}}$, $t_{\text{i,max,1}}$ and $t_{\text{i,max,2}}$ model parameters depend on β values. The analytical solutions for these model parameters are simple for pure sines ($K_2 = 1$), but become quite complicated for double waves ($K_2 = 2$). Clearly, the analytical solution for the errors of DCM model parameters would be tedious. The analytical solution for the model error

$$g(t) \pm \sigma_{g(t)} \tag{30}$$

would be practically impossible because the number of DCM model free parameters is large, especially if the data contains many signals.

We use the computational statistical bootstrap technique (Efron & Tibshirani 1986) to solve the above-mentioned analytical problems. In our bootstrap, the tested frequencies are the same as in the short search. We select a random sample ϵ^* from the residuals ϵ of DCM model (Equation 13). Any ϵ_i value can be chosen to the ϵ^* sample as many times as the random selection happens to favour it. We create $J = 1, 2, ..., n_{\text{B}}$ residual random samples ϵ^*_J. Every ϵ^*_J sample gives one *artificial* bootstrap data set

$$y^*_J = g + \epsilon^*_J. \tag{31}$$

Each y^*_J sample gives *one* free parameter estimate β_J. This bootstrap procedure gives n_{B} free parameter estimates β_J. The standard deviations for all n_{B} estimates for $P_i = 1/f_i$ and all M_k give the errors of these model parameters. We use each β_J to compute the model for a dense grid of time points. This gives us n_{B} *numerical* A_i $t_{\text{i,min,1}}$, $t_{\text{i,min,2}}$, $t_{\text{i,max,1}}$ and $t_{\text{i,max,2}}$ estimates. The standard deviation of those n_{B} estimates is the error of these model parameters. The β_J values give n_{B} estimates for $g(t)$ at any time t. The standard deviation of these $g(t)$ estimates gives the error $\sigma_{g(t)}$ of Equation 30.

Note that these β_J values can also be used to compute the errors for $h(t), h_i(t), p(t)$ and $p_k(t)$. Our computational statistical bootstrap approach gives not only the errors estimates for DCM model parameters (AL4), but also the model error inside ΔT and the forecast error outside ΔT (AL5: Equation 30).

There are, of course, totally wrong DCM models for the data. For example, DCM can be forced search for too few or too many K_1 signals, or the selected $p(t)$ trend order K_3 can be wrong, as shown in Figures 5-10 by Jetsu (2020). Such DCM models are unstable and we denote them with "UM", like in Jetsu (2025). These unstable models have three signatures

Intersecting frequencies (IF)

Dispersing amplitudes (AD)

Leaking periods (LP)

Intersecting frequencies occur when the signal frequencies in the data are very close to each other. We give the following example of how this instability can arise in the two signal model. If the frequency f_1 approaches the frequency f_2, both $h_1(t)$ and $h_2(t)$ signals become essentially one and the same signal. The LS fit fails because it makes no sense to model the same signal twice.

Dispersing amplitudes instability can occur, if the two signal frequencies are too close to each other. The LS fit finds a model, where two high amplitude signals nearly cancel out each other. The low amplitude signal, the sum of these two high amplitude signals, fits to the data.

There are DCM models where the detected frequency f is outside the tested frequency interval between $f_{\min}$ and $f_{\max}$. This leaking periods instability may indicate that the chosen tested period range is wrong.

DCM model (Equation 1) is more sophisticated than the models of our former time series analysis methods, the Three Stage Period Search (Jetsu & Pelt 1999, TSPA, Equation 1) and the Continuous Period Search (Lehtinen et al. 2011, CPS, Equation 3). TSPA and CPS can detect only one signal ($K_1 = 1$) from stationary time series ($K_3 = 0$). TSPA can detect pure sine and double wave signals ($K_2 = 1$ or 2). The extension of TSPA, the CPS method, tests three alternatives: $K_2 = 1$ or 2, or no signal at all. DCM model is more sophisticated because it can have any arbitrary K_1, K_2 and K_3 combination. DCM sum $g(t) = h(t) + p(t)$ of arbitrary periodic and aperiodic functions represents a universal model because the innumerable K_1, K_2 and K_3 combinations allow unlimited complexity. Even more complex signal shape (K_2) combinations could be used. For example, these shapes could be $K_{2,i=1} = 3$ (3rd harmonic), $K_{2,i=2} = 1$ (pure sine) and $K_{2,i=3} = 2$ (double wave), where if $K_{2,i}$ denotes the shape of i:th signal.

2.2. *WD-effect: Spearhead of DCM*

In this section, we discuss what causes the WD-effect defined in our abstract:

> *"For any sample window ΔT, DCM inevitably detects the correct $p(t)$ trend and $h(t)$ signal(-s) when sample size n and/or data accuracy σ increase."*

The consequences of this effect are also discussed.

The t_i, y_i and σ_i data are inside the rectangle $\Delta T \times \Delta y = (t_n - t_1) \times [\max(y_i + \sigma_i) - \min(y_i - \sigma_i)]$. The LS fit results do not depend on ΔT. If the measuring time intervals δt_i for each observation y_i fulfil $\delta t_i \ll \Delta T$, the R and χ^2 values obtained from *all* LS (Equation 19) depend *only* on y_i changes in ΔY direction, but not on the t_i changes in ΔT direction. For fixed y_i and σ_i, the residuals ϵ_i determine the R and χ^2 values. These residuals measure *only* the Δy direction. Hence, the periodogram z values (Equation 16 or 17) do not depend on ΔT. If ΔT decreases, the R and χ^2 estimates can measure the g_i model details inside the $\Delta T \times \Delta y$ rectangle only if n increases and/or σ decreases. Better data reveal these model details.

Plenty of concrete examples will confirm that the WD-effect is real (Sections 3.1-3.7). We will show that DCM can detect signals(-s) when the sample window ΔT is shorter than the period P value(-s). DCM surpasses AL6. The data window ΔT length is irrelevant. This means that DCM can model an infinitesimal time series, as well as forecast its future and past, if the sample size (n) and/or the data accuracy (σ) are sufficient. This revolutionary achievement allows DCM to "see through time".

2.3. *Simulated DCM model time series samples*

We use seven different $g(t)$ models to simulate 21 artificial complex time series (Sections 3.1-3.7). The sample window of all simulated time series is $\Delta T = 1$. The $n^\star$ simulated time points $t_i^\star$ are drawn from a random uniform distribution $U(0, \Delta T, n^\star)$. The first and last time point values are then modified to $t_1^\star = 0$ and $t_n^\star = \Delta T$. Hence, the distance between independent frequencies (Loumos & Deeming 1978; Kay & Marple 1981; Press et al. 1992) is always

$$f_0 = 1/\Delta T = 1. \tag{32}$$

The $n^\star$ residuals $\epsilon^\star(t_i^\star)$ of simulated model are drawn from a random normal distribution $N(0, \sigma, n^\star)$, where σ is the accuracy of simulated data. The simulated data

Table 1. Model 1: DCM analysis between $P_{\min} = 0.63$ and $P_{\max} = 5.70$. (1) Simulated P_1, A_1, $t_{1,\min,1}$, $t_{1,\max,1}$ and M_0 values. (2-4) Detected values for different n and SN combinations. Two lowest lines specify electronic data files and DCM analysis control files.

(1)	(2)	(3)	(4)
	$n = 50$	$n = 50$	$n = 100$
Model 1	SN = 10	SN = 50	SN = 100
$P_1 = 1.9$	1.58 ± 0.21	1.822 ± 0.042	1.863 ± 0.017
$A_1 = 2.0$	1.64 ± 0.33	1.894 ± 0.072	1.941 ± 0.025
$t_{1,\min,1} = 1.35$	1.20 ± 0.11	1.312 ± 0.021	1.3321 ± 0.0088
$t_{1,\max,1} = 0.40$	0.4098 ± 0.0058	0.4007 ± 0.0013	0.40056 ± 0.00078
$M_0 = 1.0$	1.20 ± 0.18	1.061 ± 0.036	1.030 ± 0.014
Data file	`Model1n50SN10.dat`	`Model1n50SN50.dat`	`Model1n100SN100.dat`
Control file	`dcmModel1n50SN10.dat`	`dcmModel1n50SN50.dat`	`dcmModel1n100SN100.dat`

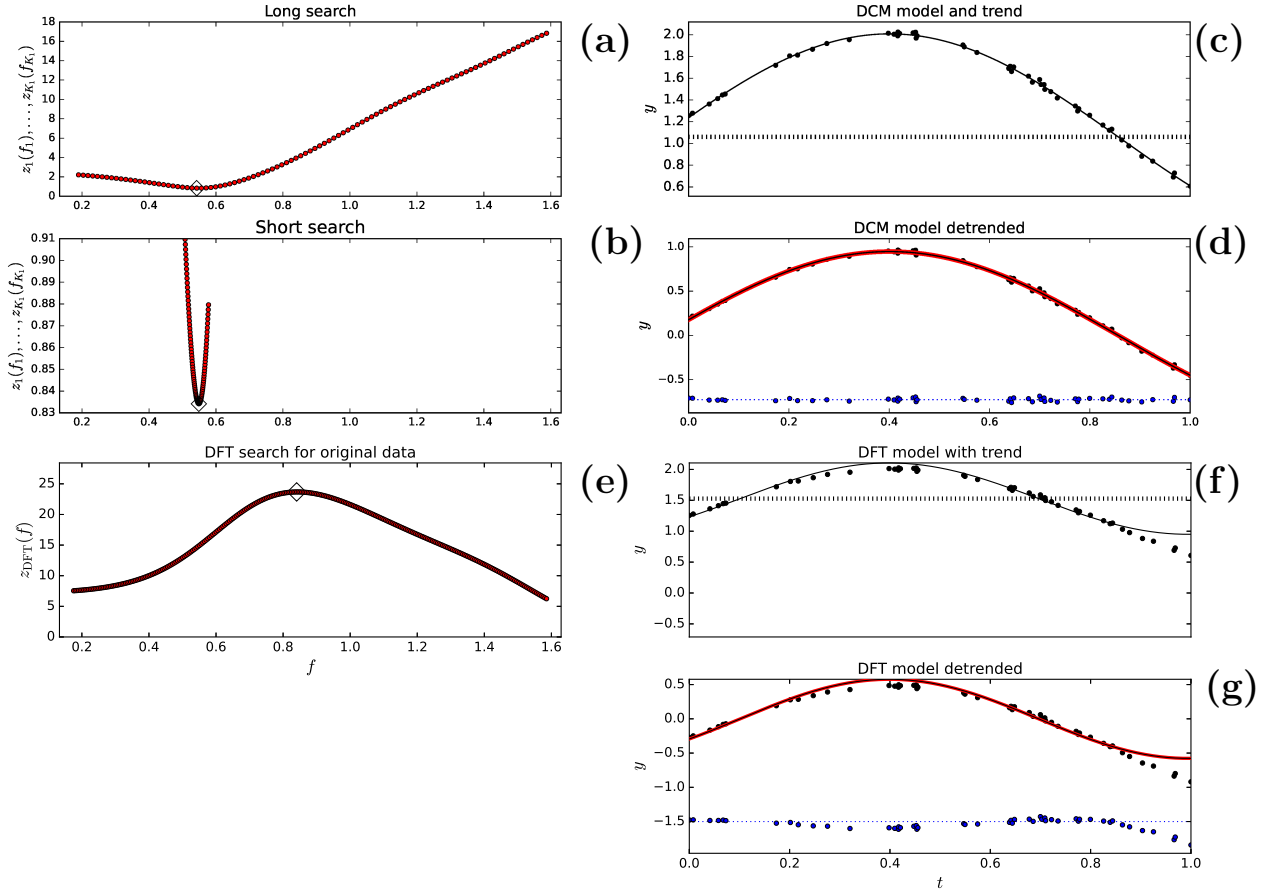

Figure 1. Model 1 (Table 1: $n = 50$, SN = 50 simulation). (a) DCM long search periodogram $z_1(f_1)$ gives best period at 1.843 (diamond). (b) DCM short search periodogram $z_1(f_1)$ gives best period at 1.822 (diamond). (c) DCM model $g(t)$ (black continuous line), DCM trend $p(t)$ (black dashed line) and data y_i (black dots). (d) DCM model detrended $g(t) - p(t)$ (black continuous line), DCM signal $h_1(t)$ (red thick continuous line), detrended data $y(t_i) - p(t_i)$ (black dots) and DCM model residuals $y(t_i) - g(t_1)$ (blue dots) offset to -0.65 level (blue dotted line). (e) DFT periodogram $z_{\rm DFT}(f)$ gives best period at 1.190 (Diamond). (f) DFT model $g_{\rm DFT}(t)$ (black continuous line), DFT trend $p_{\rm DFT}(t)$ (black dashed line) and data y_i (black dots). (g) DFT model detrended $g_{\rm DFT}(t) - p_{\rm DFT}(t)$ (black continuous line), DFT pure sine $s_{\rm DFT}(t)$ (red thick continuous line), detrended data $y(t_i) - p_{\rm DFT}(t_i)$ (black dots) and DFT model residuals (blue dots) offset to -1.5 level (blue dotted line).

are

$$y^\star(t_i^\star) = g(t_i^\star) + \epsilon^\star(t_i^\star) \\ = h(t_i^\star) + p(t_i^\star) + \epsilon^\star(t_i^\star).$$

The peak to peak amplitudes of all simulated signals is $A = 2$. Our definition for the signal to noise ratio is

$$\mathrm{SN} = (A/2)/\sigma. \quad (33)$$

2.4. *Discrete Fourier Transform (DFT)*

DFT is one of the most prevalent frequency-domain parametric time series analysis methods for unevenly spaced time series. It searches for the best pure sine model for the data. The equivalent DCM model has the orders $K_1 = 1, K_2 = 1$ and $K_3 = 0$. Any time series analysis method must be remarkable if it performs better than the distinguished DFT. Therefore, we will compare how DCM and DFT perform in the analyses of simulated time series (Sections 3.1-3.7). We search for signals in these simulated time series by using the frequently applied DFT version formulated by Horne & Baliunas (1986)[1]. Our notation for DFT test statistic is $z_{\mathrm{PDF}}(f)$ (Horne & Baliunas 1986, Equation 1). The notations for DFT model are

$$g_{\mathrm{DFT}}(t) = s_{\mathrm{DFT}}(t) + p_{\mathrm{DFT}}(t), \quad (34)$$

where $s_{\mathrm{DFT}}(t)$ is the sum of pure sine signals and $p_{\mathrm{DFT}}(t)$ is the trend. The pure sine signals for the y_i data and the ϵ_i residuals are denoted with $s_{y,\mathrm{DFT}}(t)$ and $s_{\epsilon,\mathrm{DFT}}(t)$, respectively. Our DFT analyses of simulated time series may fail due to AL6-AL10.

The time series is "too short" (Kay & Marple 1981; Scargle 1982, 1989, AL6)

$$P > \Delta T. \quad (35)$$

The time series is non-stationary because it contains "a trend" (Nerlove 1964; Chianca et al. 2005; Kim et al. 2009, AL7)

$$p_{\mathrm{DFT}}(t) \neq 0. \quad (36)$$

Due to the leakage caused by the sample window (AL8), the signal frequencies are "too close" (Loumos & Deeming 1978; Kay & Marple 1981; Martinez & Kurtz 1990, AL9).

$$|f_1 - f_2| < f_0 = \Delta T^{-1}. \quad (37)$$

The signals are not "pure sines" (Bretthorst 1988; Koen 1995; Baluev 2009, AL10)

$$K_2 \neq 1. \quad (38)$$

[1] This paper had over 2600 citations in December 2025

Many parametric frequency-domain time series analysis methods, like DFT, can be applied only to stationary data. Trends changing the time series mean or variance must be removed before applying these methods (Nerlove 1964). DFT can detect only one period at the time. Such frequency-domain parametric time series analysis methods are hereafter called "one-dimensional". Our DFT analysis of simulated data proceeds through two stages. First, the $p_i = p_{\mathrm{DFT}}(t_i)$ trend is removed from the simulated data $y_i = y(t_i)$. Then, the iterative pre-whitening technique (Reinhold et al. 2013; Zhu & Jia 2018; Ghaderpour et al. 2021) is applied to search for the pure sine signals. We search for the first signal by applying DFT to the detrended $y_i - p_i$ simulated data. In the second signal search, the DTF is applied to the original model residuals $\epsilon_i = (y_i - p_i) - g_{\mathrm{y,DFT}}(t_i)$. At any stage, this combination of detrending and iterative pre-whitening analysis may fail and corrupt the results obtained at the next stages. For example, all DFT analyses of simulated time series will fail already at the detrending stage (Sections 3.1-3.7). Our DCM has no corrupting separate stages because the trend and the signal(-s) are detected (modelled) simultaneously.

3. RESULTS

The validity of the WD-effect is critical to the credibility of our work. We confirm this validity thoroughly by applying DCM and DFT to seven different simulated time series (Sections 3.1 - 3.7). As we proceed, these time series become increasingly complex. DCM analysis succeeds for all time series. DFT analysis fails for every time series.

3.1. *Model 1*

Our first time series simulation model is the one signal model

$$g(t) = (A_1/2)\cos\left[\frac{2\pi(t - t_{1,\max,1})}{P_1}\right] + M_0. \quad (39)$$

We give the P_1, A_1, $t_{1,\max,1}$ and M_0 values in Table 1 (Column 1). This sample is "too short" (Equation 35) because $P_1 = 1.9\Delta T$. The constant mean level M_0 is unknown (Equation 36). We perform DCM and DFT time series analysis between $P_{\min} = P_1/3 = 0.63$ and $P_{\max} = 3P_1 = 5.70$.

Model 1 is a DCM model, where $K_1 = 1$, $K_2 = 1$, $K_3 = 0$, $B_{1,1} = (A_1/2)\cos(2\pi f_1 t_{1,\max,1})$ and $C_{1,1} = (A_1/2)\sin(2\pi f_1 t_{1,\max,1})$. We give DCM analysis results for three samples having different n and SN combinations (Table 1: Columns 2-4). For each sample, this table specifies the electronic data file and the electronic

DCM control file.[2] The detected P_1, A_1, $t_{1,\max,1}$ and M_0 values are correct and accurate even for the combination $n = 50$ and SN = 50. Regardless of $\Delta T < P_1$, these model parameter values become more accurate and converge to the correct simulated values when n and SN increase. This confirms the WD-effect.

A graphical presentation of DCM analysis results is shown for Model 1 simulated time series, where $n = 50$ and SN = 50 (Figures 1a-d). DCM long search $z_1(f_1)$ periodogram minimum is at $P_1 = 1.843$ (Figure 1a: diamond). DCM short search periodogram $z_1(f_1)$ gives the best period at $P_1 = 1.822$ (Figure 1b: diamond). The continuous black line denoting DCM model $g(t)$ fits perfectly to the black dots denoting the data y_i (Figure 1c). The mean $p(t) = M_0 = 1.061 \pm 0.036$ is correct (Figure 1c: dashed black line). The detrended model $g(t) - p(t)$ (black continuous line), the detrended data $y(t_i) - p(t_i)$ (black dots) and the pure sine signal $h_1(t)$ (red thick continuous line) are shown in Figure 1d. Note that the thick continuous red line stays under the thin continuous black line because $h_1(t) = g(t) - p(t)$. DCM residuals (blue dots) are offset to the level of -0.65 (blue dotted line). These residuals are stable and display no trends.

DFT detects the wrong period $P_1 = 1.190$ (Figure 1e: Diamond). DFT mean level estimate $M_0 = 1.527$ is also wrong (Figure 1f: Dashed black line). The black dots denoting the data y_i deviate from the continuous black line denoting DFT model $g_{\mathrm{DFT}}(t)$. The detrended model $g_{\mathrm{DFT}}(t) - p_{\mathrm{DFT}}(t)$ (continuous black line), the detrended data $y(t_i) - p_{\mathrm{DFT}}(t_i)$ (black dots) and the signal $s_{\mathrm{y,DFT}}(t)$ (continuous thick red line) are shown in Figure 1g. The thin black line covers the thick red line because $s_{\mathrm{y,DFT}}(t) = g_{\mathrm{DFT}}(t) - p_{\mathrm{DFT}}(t)$. DFT residuals (blue dots) offset to the level of -1.5 (blue dotted line) display obvious trends, especially at the end of analysed sample.

For the simulated time series of Model 1, DCM analysis succeeds, but DFT analysis fails.

3.2. *Model 2*

Our next one signal time series simulation model is

$$g(t) = (A_1/2)\cos\left[\frac{2\pi(t - t_{1,\max,1})}{P_1}\right] + M_0 + M_1 T + M_2 T^2 \qquad (40)$$

where $T = [2(t - t_{\mathrm{mid}})]/\Delta T$. We give the P_1, A_1, $t_{1,\max,1}$, M_0, M_1 and M_2 values in Table 2. As a DCM model, the orders of Model 2 are $K_1 = 1$, $K_2 = 1$ and $K_3 = 2$. The simulated time series is "too short" because the period P_1 is $1.9 \times \Delta T$ (Equation 35). The parabolic trend $p(t)$ is unknown (Equation 36). Again, we use DCM and DFT time series analysis methods to search for periods between $P_{\min} = P_1/3 = 0.63$ and $P_{\max} = 3P_1 = 4.70$.

DCM analysis results are given in Table 2. These results are not very accurate for the $n = 100$ and SN = 100 combination, but they definitely improve for larger n and SN values. The WD-effect ensures that the detected values converge to the correct simulated model parameter values. The short sample window, $\Delta T < P_1$, does not mislead DCM analysis.

For Model 2, DCM analysis results are illustrated for the $n = 10\ 000$ and SN = 100 combination (Figures 2a-d). DCM long search $z_1(f_1)$ periodogram minimum is at $P_1 = 1.843$ (Figure 2a: diamond). DCM short search gives the value $P_1 = 1.852$ (Figure 2b: diamond). DCM model $g(t)$ is so good that it's continuous black line is totally covered by the black dots representing the y_i data (Figure 2c). Therefore, the colour of this $g(t)$ line has been changed from black to white. The results for the parabolic trend coefficients $M_0 = 1.079 \pm 0.096$, $M_1 = 0.229 \pm 0.026$ and $M_2 = 0.451 \pm 0,064$ of the dashed black $p(t)$ line are correct. In Figure 1d, the white continuous line shows the detrended model $g(t) - p(t)$. The black dots show the detrended data $y(t_i) - p(t_i)$ and the red thick continuous line shows the pure sine signal $h_1(t)$. Note that the red thick line is under the white thin line because $h_1(t) = g(t) - p(t)$. DCM residuals (blue dots) are offset to the level of -0.65. The colour of dotted line, which denotes this offset level, has been changed from blue to white. The distribution of these DCM model residuals is stable, as expected for a random normal distribution.

The wrong period $P_1 = 0.697$ is detected by DFT (Figure 2e: Diamond). DFT estimates for the trend $p(t)$ coefficients, $M_0 = 1.92$, $M_1 = -0.15$ and $M_2 = -0.56$ are also wrong (Figure 2f: dashed black line). The data y_i (black dots) deviate from DFT model $g_{\mathrm{DFT}}(t)$ (black continuous line), especially in the end of the sample (Figure 2g). For the detrended DFT model, the thin black line covers the thick red line because $s_{\mathrm{y,DFT}}(t) = g_{\mathrm{DFT}}(t) - p_{\mathrm{DFT}}(t)$ (Figure 2g). The $s_{\mathrm{y,DFT}}(t)$ sine curve peak to peak amplitude is far below the correct $A_1 = 2.0$ value. DFT residuals (blue dots) are offset to the level of -1.5 (blue dotted line). The trends of these residuals confirm that DFT analysis fails.

Only DCM (not DFT) succeeds in the analysis of Model 2 simulated time series.

[2] See declaration "Code and data availability".

3.3. *Model 3*

Table 2. Model 2. DCM analysis between $P_{\rm min} = 0.63$ and $P_{\rm max} = 4.70$. Notations as in Table 1.

(1)	(2)	(3)	(4)
	$n = 1\ 000$	$n = 10\ 000$	$n = 10\ 000$
Model 2	SN = 100	SN = 100	SN = 200
$P_1 = 1.9$	1.98 ± 0.34	1.852 ± 0.055	1.933 ± 0.040
$A_1 = 2.0$	2.4 ± 2.8	1.84 ± 0.20	2.13 ± 0.16
$t_{1,\rm min,1} = 1.35$	1.39 ± 0.17	1.326 ± 0.027	1.367 ± 0.020
$t_{1,\rm max,1} = 0.40$	0.4015 ± 0.0020	0.40007 ± 0.00072	0.40017 ± 0.00075
$M_0 = 1.0$	0.8 ± 1.4	1.079 ± 0.096	0.937 ± 0.080
$M_1 = 0.25$	0.29 ± 0.21	0.229 ± 0.026	0.266 ± 0.020
$M_2 = 0.50$	0.62 ± 0.52	0.451 ± 0.064	0.540 ± 0.049
Data file	`Model2n1000SN100.dat`	`Model2n10000SN100.dat`	`Model2n10000SN200.dat`
Control file	`dcmModel2n1000SN100.dat`	`dcmModel2n10000SN100.dat`	`dcmModel2n10000SN200.dat`

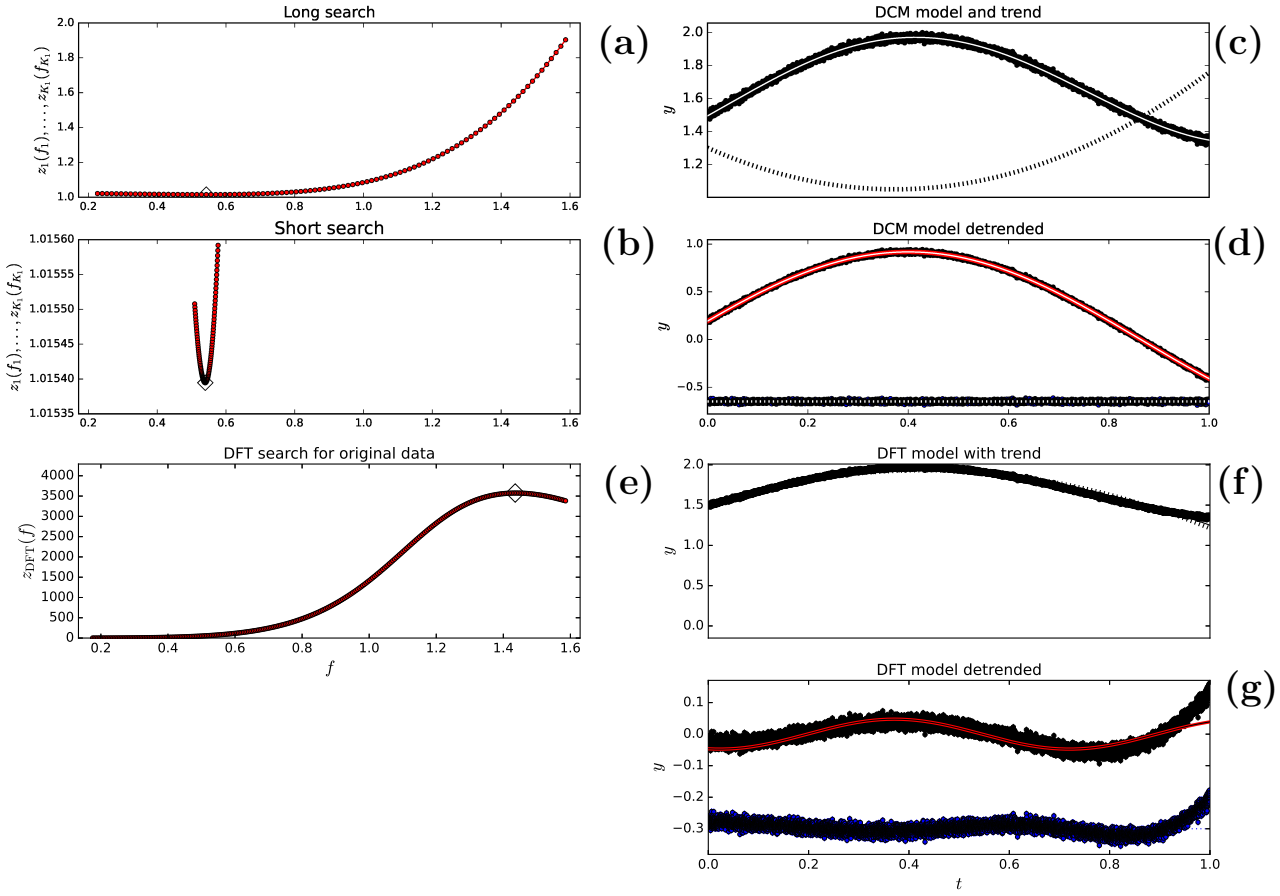

Figure 2. Model 2 (Table 2: $n = 10\ 000$, SN = 100 simulation). (c) Colour of $g(t)$ line has been changed from black to white. (d) Colour of $g(t)$ line has been changed from black to white. Colour of offset level -0.3 dotted line has been changed from blue to white. Locations of best periods (diamonds) are explained in text (Section 3.2). Otherwise, notations are as in Figure 1.

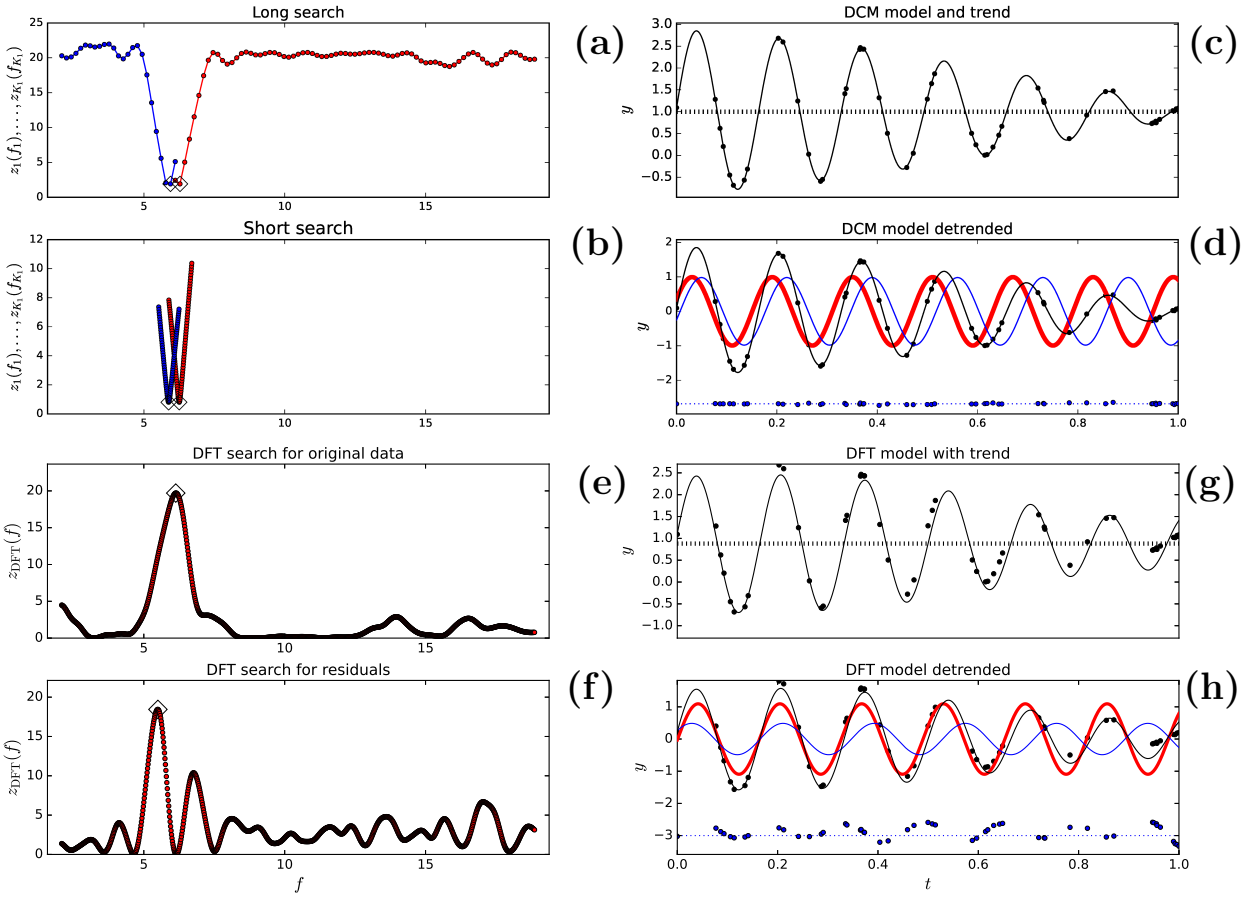

Figure 3. Model 3 (Table 3: $n = 50$, SN $= 50$ simulation). (a) DCM long search periodograms $z_1(f_1)$ (red) and $z_2(f)$ (blue) give best periods at 0.160 and 0.168 (diamonds). (b) DCM short search periodograms $z_1(f_1)$ (red) and $z_2(f)$ (blue) give best periods at 0.160 and 0.170 (diamonds). (c) DCM model $g(t)$ (black continuous line), DCM trend $p(t)$ (black dashed line) and data y_i (black dots). (d) DCM model detrended $g(t) - p(t)$ (black continuous line), DCM signal $h_1(t)$ (red thick continuous line), DCM signal $h_2(t)$ (blue continuous thin line), detrended data $y(t_i) - p(t_i)$ (black dots) and DCM model residuals $y(t_i) - g(t_1)$ (blue dots) offset to -3.0 level (blue dotted line) (e) DFT periodogram $z_{\rm DFT}(f)$ for the original data gives best period at 0.163 (diamond). (f) DFT periodogram $z_{\rm DFT}(f)$ for the sine model residuals gives best period at 0.182 (diamond). (g) DFT model $g_{\rm DFT}(t)$ (black continuous line), DFT trend $p_{\rm DFT}(t)$ (black dashed line) and data y_i (black dots). (h) DFT model detrended $g_{\rm DFT}(t) - p_{\rm DFT}(t)$ (black continuous line), DFT pure sine model for original data $s_{y,{\rm DFT}}(t)$ (red thick continuous line), DFT pure sine model for first residuals $s_{\epsilon,{\rm DFT}}(t)$ (blue continuous thin line), detrended data $y(t_i) - p_{\rm DFT}(t_i)$ (black dots) and DFT model residuals (blue dots) offset to -3.0 level (blue dotted line).

Table 3. Model 3. DCM analysis between $P_{\rm min} = 0.053$ and $P_{\rm max} = 0.480$. Notations as in Table 1.

(1)	(2)	(3)	(4)
	$n = 50$	$n = 50$	$n = 100$
Model 3	SN = 10	SN = 50	SN = 100
$P_1 = 0.16$	0.15898 ± 0.00068	0.15989 ± 0.00014	0.159950 ± 0.000077
$A_1 = 2.0$	1.85 ± 0.10	1.987 ± 0.036	1.994 ± 0.015
$t_{1,\rm min,1} = 0.11$	0.1122 ± 0.0020	0.11080 ± 0.00088	0.11004 ± 0.00030
$t_{1,\rm max,1} = 0.03$	0.0328 ± 0.0022	0.03086 ± 0.00094	0.03007 ± 0.00033
$P_2 = 0.17$	0.17025 ± 0.00066	0.17018 ± 0.00024	0.16996 ± 0.00080
$A_2 = 2.0$	2.001 ± 0.090	1.962 ± 0.046	2.002 ± 0.012
$t_{2,\rm min,1} = 0.135$	0.1312 ± 0.0022	0.13461 ± 0.00087	0.13495 ± 0.00031
$t_{2,\rm max,1} = 0.05$	0.0461 ± 0.0024	0.04952 ± 0.00099	0.04997 ± 0.000038
$M_0 = 1.0$	0.996 ± 0.011	0.9985 ± 0.0032	0.9995 ± 0.0012
Data file	`Model3n50SN10.dat`	`Model3n50SN50.dat`	`Model3n100SN100.dat`
Control file	`dcmModel3n50SN10.dat`	`dcmModel3n50SN50.dat`	`dcmModel3n100SN100.dat`

Our third time series simulation model is the two signal model

$$g(t) = (A_1/2)\cos\left[\frac{2\pi(t - t_{1,\rm max,1})}{P_1}\right] + (A_2/2)\cos\left[\frac{2\pi(t - t_{2,\rm max,1})}{P_2}\right] + M_0. \quad (41)$$

We give the P_1, P_2, A_1,A_2, $t_{1,\rm max,1}$, $t_{2,\rm max,1}$ and M_0 values in Table 3 (Column 1). DCM orders of Model 3 are $K_1 = 2$, $K_2 = 1$ and $K_3 = 0$. The simulated time series is not "too short" (Equation 35) because $P_1 < P_2 < \Delta T$. The $p(t)$ trend mean level M_0 is unknown (Equation 36). The two frequencies are "too close" because $\Delta f = f_1 - f_2 = f_0/2.72$ (Equation 37). We apply DCM and DFT time series analyses to search for periods between $P_{\rm min} = P_1/3 = 0.053$ and $P_{\rm max} = 3P_1 = 0.480$.

DCM analysis results for different n and SN combinations are given in Table 3 (Columns 2-4). The results are surprisingly accurate even for the small $n = 50$ sample having a low SN = 10. Due to the WD-effect, the detected model parameter values converge to the correct simulated values when n and SN increase.

DCM results for Model 3 are illustrated in Figures 1a-d ($n = 50$ and SN = 50 combination). DCM long search $z_1(f_1)$ periodogram (red) and $z_2(f)$ periodogram (blue) minima are at the best periods $P_1 = 0.159$ and $P_2 = 0.168$ (Figure 3a: diamonds). DCM short search gives $P_1 = 0.160$ and $P_2 = 0.170$ (Figure 3b: diamonds). The black continuous DCM model $g(t)$ curve crosses trough the black dots of y_i data (Figure 1c). The result $M_0 = 0.9985 \pm 0.0032$ for the dashed black $p(t)$ trend line is correct. Our Figure 3d shows the detrended DCM model $g(t) - p(t)$ (black continuous line), the detrended data $y(t_i) - p(t_i)$ (black dots), the pure sine signal $h_1(t)$ (red thick continuous line) and the pure sine signal $h_2(t)$ (blue thin continuous line). DCM residuals (blue dots) are offset to the level of -3.0 (blue dotted line). These residuals are small and their level is stable.

DFT detects the wrong period $P_1 = 1.163$ for the original data (Figure 3e: diamond). This is an expected result because the detected period should be close to $(P_1+P_2)/2$ when the peak to peak amplitudes of the simulated data, $A_1 = A_2$, are equal (Jetsu 2025). The two DFT periodogram $z_{\rm DFT}(f)$ peaks at frequencies $1/P_1$ and $1/P_2$ overlap and merge into one peak. The period $P_2 = 0.182$ detected for the residuals is also wrong (Figure 3f: diamond). DFT trend $p_{\rm PDF}(t)$ estimate $M_0 = 0.880$ fails (Figure 3f: dashed black line). The black dots y_i show minor deviations from the continuous black DFT model $g_{\rm DFT}(t)$ line (Figure 3g). The detrended model $g_{\rm DFT}(t) - p_{\rm DFT}(t)$ (continuous black line), the detrended data $y(t_i) - p_{\rm DFT}(t_i)$ (black dots), the pure sine signal $s_{\rm y,DFT}(t)$ for the data (continuous thick red line) and the pure sine signal $s_{\epsilon,\rm DFT}(t)$ for the residuals (continuous blue thin line) are shown in Figure 3h. Note that the $s_{\rm y,DFT}(t)$ and $s_{\epsilon,\rm DFT}(t)$ signal amplitudes are far from equal. DFT residuals (blue dots) offset to the level of -3.0 (blue dotted line) are not stable.

DCM analysis of Model 3 simulated time series succeeds, but DFT analysis does not.

3.4. *Model 4*

The next time series simulation model is

$$g(t) = (A_1/2)\cos\left[\frac{2\pi(t - t_{1,\rm max,1})}{P_1}\right] + (A_2/2)\cos\left[\frac{2\pi(t - t_{2,\rm max,1})}{P_2}\right] + M_0 + M_1 T + M_2 T^2, \quad (42)$$

where $T = [2(t - t_{\rm mid})]/\Delta T$. In this model, two signals are superimposed on an unknown parabolic trend. The P_1, P_2, A_1,A_2, $t_{1,\rm max,1}$, $t_{2,\rm max,1}$, M_0, M_1 and M_2

Table 4. Model 4. DCM analysis between $P_{\min} = 0.053$ and $P_{\max} = 0.480$. Notations as in Table 1.

(1)	(2)	(3)	(4)
	$n = 50$	$n = 50$	$n = 100$
Model 4	SN = 10	SN = 50	SN = 100
$P_1 = 0.16$	0.1606 ± 0.0084	0.15982 ± 0.00023	0.16001 ± 0.000058
$A_1 = 2.0$	1.97 ± 0.15	1.958 ± 0.051	2.012 ± 0.014
$t_{1,\min,1} = 0.11$	0.1134 ± 0.0036	0.11059 ± 0.00088	0.10977 ± 0.00033
$t_{1,\max,1} = 0.03$	0.0331 ± 0.0040	0.03068 ± 0.00097	0.02976 ± 0.00035
$P_2 = 0.17$	0.1744 ± 0.0012	0.17009 ± 0.00026	0.169958 ± 0.000092
$A_2 = 2.0$	1.74 ± 0.17	1.982 ± 0.044	2.018 ± 0.017
$t_{2,\min,1} = 0.135$	0.1321 ± 0.0044	0.1342 ± 0.0011	0.13528 ± 0.00042
$t_{2,\max,1} = 0.05$	0.0460 ± 0.0048	0.0492 ± 0.0012	0.05031 ± 0.00034
$M_0 = 1.0$	1.015 ± 0.024	0.9997 ± 0.0030	1.0012 ± 0.0016
$M_1 = 0.25$	0.219 ± 0.023	0.2565 ± 0.0066	0.2495 ± 0.0019
$M_2 = 0.5$	0.485 ± 0.050	0.511 ± 0.011	0.4983 ± 0.0033
Data file	`Model4n50SN10.dat`	`Model4n50SN50.dat`	`Model4n100SN100.dat`
Control file	`dcmModel4n50SN10.dat`	`dcmModel4n50SN50.dat`	`dcmModel4n100SN100.dat`

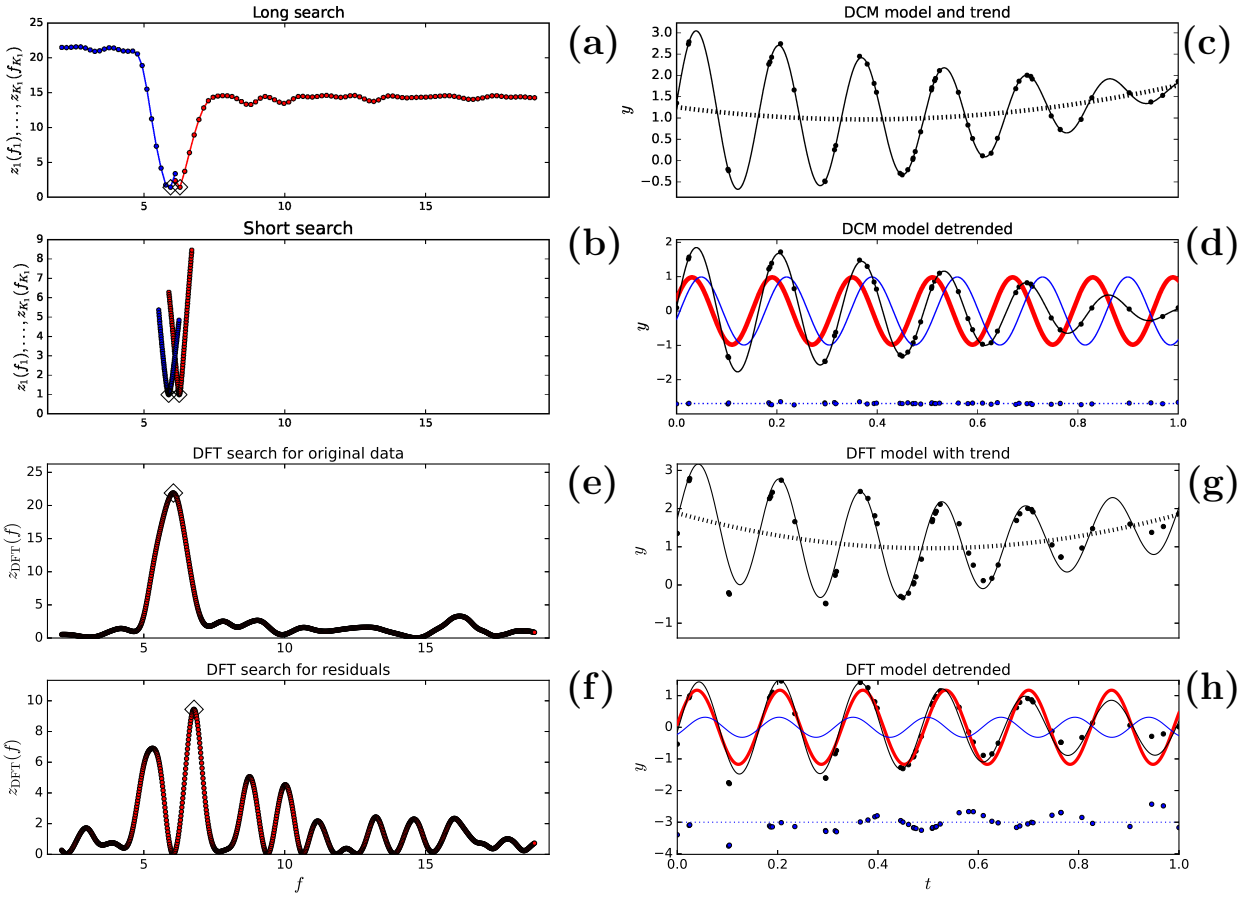

Figure 4. Model 4 (Table 4: $n = 50$, SN = 50 simulation). Notations as in Figure 3. Best periods (diamonds) are explained in Section 3.4.

values are given in Table 4. This simulated sample is not "too short" (Equation 35). The polynomial trend $p(t)$ is unknown (Equation 36). The $\Delta f = f_1 - f_2 = f_0/2.72$ difference means that the frequencies are "too close" (Equation 37). DCM and DFT time series analysis methods are used to search for periods between $P_{\rm min} = P_1/3 = 0.053$ and $P_{\rm max} = 3P_1 = 0.480$.

Model 4 is a DCM model having orders $K_1 = 2$, $K_2 = 1$ and $K_3 = 2$. Our DCM analysis results for different n and SN combinations are given in Table 4 (Columns 2-4). DCM detects the correct P_1, P_2, A_1, A_2, $t_{\rm max,1}$, $t_{\rm max,2}$, M_0, M_1 and M_2 values even for the lowest $n = 50$ and SN = 10 combination. As the simulated data n and SN values increase, the WD-effect ensures that all detected values converge to the correct simulated model parameter values.

Our Figures 4a-d illustrate DCM analysis results for one sample of simulated time series (Model 4: $n = 50$ and SN = 50). DCM long search best periods are at $P_1 = 0.159$ and $P_2 = 0.168$ (Figure 4a: diamonds). DCM short search values are $P_1 = 0.160$ and $P_2 = 0.170$ (Figure 4b: diamonds). The continuous DCM model $g(t)$ black line crosses through all black dots representing the y_i data (Figure 4c). DCM detects the correct polynomial trend $p(t)$ coefficients $M_0 = 0.9997 \pm 0.0030$, $M_1 = 0.2565 \pm 0.0066$ and $M_2 = 0.511 \pm 0.011$ (Figure 4c: dashed black line). The detrended DCM model $g(t) - p(t)$ (black continuous line), the detrended data $y(t_i) - p(t_i)$ (black dots), the pure sine signal $h_1(t)$ (red thick continuous line) and the pure sine signal $h_2(t)$ (blue thin continuous line) are shown in Figure 4d. DCM residuals (blue dots) offset to the level of -3.0 show no trends and are extremely stable.

Since the peak to peak amplitudes of the simulated data signals are equal, $A_1 = A_2$, the expected result for DFT analysis of original data is $(P_1 + P_2)/2$, where P_1 and P_2 are the simulated signal periods (Jetsu 2025). DFT detects this expected wrong period $P_1 = 1.165$ for the original data (Figure 4e: diamond). A wrong period $P_2 = 0.147$ is also detected for the residuals (Figure 4f: diamond). DFT estimate $M_0 = 0.962$ for the $p(t)$ trend is close to the correct value $M_0 = 1$, but the $M_1 = -0.017$ and $M_2 = 0.903$ estimates are wrong (Figure 4g: dashed black line). The continuous black DFT model $g_{\rm DFT}(t)$ line deviates from the black dots of data y_i, especially in the beginning and end of the sample. In our Figure 4h, the black dots are the detrended data $y(t_i) - p_{\rm DFT}(t_i)$ and the continuous black line is the detrended DFT model $g_{\rm DFT}(t) - p_{\rm DFT}(t)$. The continuous thick red line is the pure sine signal $s_{\rm y,DFT}(t)$ for the original data and the continuous thinner blue line is the pure sine signal $s_{\epsilon,\rm DFT}(t)$ for the residuals. The $s_{\epsilon,\rm DFT}(t)$ signal amplitude is far below the simulated correct value $A_2 = 2.0$. The blue dots representing DFT model residuals are offset to the level of -3.0 and show clear trends.

Our Model 4 simulated time series analysis succeeds for DCM and fails for DFT.

3.5. *Model 5*

The mathematical time series equation

$$\begin{aligned} g(t) = {} & (A_1/2)\cos\left[\frac{2\pi(t - t_{1,\rm max,1})}{P_1}\right] \\ & + (A_2/2)\cos\left[\frac{2\pi(t - t_{2,\rm max,1})}{P_2}\right] \\ & + M_0 + M_1 T + M_2 T^2 \end{aligned} \tag{43}$$

for our fifth model is the same as for Model 4 (Equation 42). However, this new Model 5 differs from the earlier Model 4 because we use totally different P_1, P_2, A_1,A_2, $t_{1,\rm max,1}$, $t_{2,\rm max,1}$, M_0, M_1 and M_2 values (Table 5, Column 1). The simulated time series is "too short" for both P_1 and P_2 periods (Equation 35). There is "a trend" $p(t)$ (Equation 36). The signal frequencies f_1 and f_2 are "too close" (Equation 37). The three main reasons that can cause the failure of DFT analysis are present. We perform DCM and DFT time series analysis between $P_{\rm min} = P_1/3 = 0.47$ and $P_{\rm max} = 3P_1 = 4.20$.

Model 5 has DCM orders $K_1 = 2$, $K_2 = 1$ and $K_3 = 2$. We give DCM analysis results for different n and SN combinations in Table 5 (Columns 2-4). DCM fails to detect the correct P_1, P_2, ..., M_1 and M_2 values for the the lowest $n = 10\ 000$ and SN = 1 000 combination (Table 5, Column 2). This shows that DCM can fail, just like any other time series analysis method, if the quality of data is too low. Due to the WD-effect, DCM results for higher n and SN combinations are correct (Table 5, Columns 3-4).

We show DCM analysis results for one sample of Model 5 simulated time series (Figures 5a-d: $n = 10\ 000$ and SN = 10 000). The long and short DCM searches give $P_1 = 1.407$ and $P_2 = 1.917$ (Figure 5a: diamonds), and $P_1 = 1.393$ and $P_2 = 2.024$ (Figure 5b: Diamonds). The continuous line denoting the model $g(t)$ is totally covered by the black dots of y_i data (Figure 5c). Therefore, we use white colour to highlight this DCM model $g(t)$ line. The detected polynomial trend $p(t)$ coefficients $M_0 = 0.81 \pm 0.16$, $M_1 = 0.230 \pm 0.013$ and $M_2 = 0.596 \pm 0.084$ are correct (Figure 5c: dashed black line). We show the detrended DCM model $g(t) - p(t)$ (white continuous line), the detrended data $y(t_i) - p(t_i)$ (black dots), the signal $h_1(t)$ (red thick continuous line) and the signal $h_2(t)$ (blue thin continuous line) in Figure 5d. DCM residuals (blue dots) are offset to the level of -1.5. These blue dots appear black because there

Table 5. Model 5. DCM analysis between $P_{\min} = 0.47$ and $P_{\max} = 4.20$. Notations as in Table 1.

(1)	(2)	(3)	(4)
	$n = 10\ 000$	$n = 10\ 000$	$n = 100\ 000$
Model 5	SN = 1 000	SN = 10 000	SN = 10 000
$P_1 = 1.4$	1.404 ± 0.046	1.393 ± 0.014	1.4019 ± 0.0035
$A_1 = 2.0$	2.27 ± 0.45	1.97 ± 0.13	1.994 ± 0.029
$t_{1,\min,1} = 1.1$	1.129 ± 0.023	1.1030 ± 0.0028	1.09766 ± 0.00082
$t_{1,\max,1} = 0.4$	0.4267 ± 0.0012	0.4066 ± 0.0059	0.3967 ± 0.0020
$P_2 = 1.9$	5.24 ± 0.83	2.023 ± 0.094	1.854 ± 0.023
$A_2 = 2.0$	5 ± 24	2.35 ± 0.34	1.918 ± 0.074
$t_{2,\min,1} = 1.55$	3.23 ± 0.50	1.599 ± 0.030	1.5292 ± 0.0066
$t_{2,\max,1} = 0.6$	0.61 ± 0.34	0.587 ± 0.030	0.6024 ± 0.0064
$M_0 = 1.0$	-2 ± 12	0.81 ± 0.16	1.055 ± 0.035
$M_1 = 0.25$	-1.27 ± 0.53	0.230 ± 0.013	0.2590 ± 0.0029
$M_2 = 0.5$	3.8 ± 1.8	0.596 ± 0.084	0.470 ± 0.020
Data file	`Model5n10000SN1000.dat`	`Model5n10000SN10000.dat`	`Model5n100000SN10000.dat`
Control file	`dcmModel5n10000SN1000.dat`	`dcmModel5n10000SN10000.dat`	`dcmModel5n100000SN10000.dat`

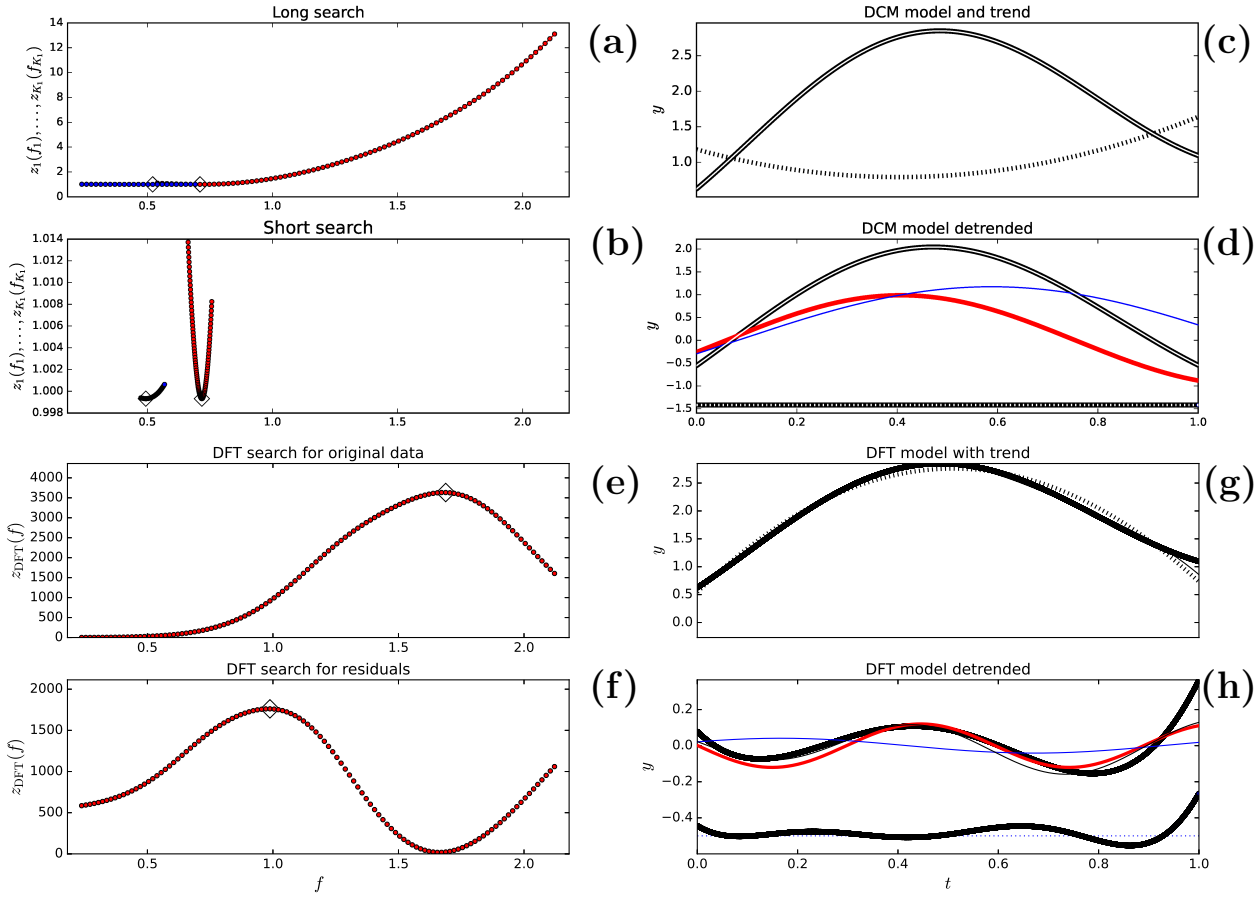

Figure 5. Model 5 (Table 5: $n = 10\ 000$, SN = 10 000 simulation). Notations as in Figure 3. Best periods (diamonds) are explained in Section 3.5.

Table 6. Model 6. DCM analysis between $P_{\rm min} = 0.053$ and $P_{\rm max} = 0.480$. Notations as in Table 1.

(1)	(2)	(3)	(4)
	$n = 50$	$n = 50$	$n = 100$
Model 6	SN = 10	SN = 50	SN = 100
$P_1 = 0.16$	0.15966 ± 0.00020	0.160022 ± 0.000021	0.160024 ± 0.000019
$A_1 = 2.0$	$1.983 \pm \pm 0.041$	2.0031 ± 0.0086	2.0013 ± 0.0032
$t_{1,\min,1} = 0.0979$	0.09868 ± 0.00057	0.097933 ± 0.000096	0.097934 ± 0.000079
$t_{1,\min,2} = 0.0225$	0.02283 ± 0.00070	0.022563 ± 0.000077	0.022403 ± 0.000071
$t_{1,\max,1} = 0.1421$	0.14241 ± 0.00064	0.142099 ± 0.000079	0.142100 ± 0.000077
$t_{1,\max,2} = 0.0575$	0.05827 ± 0.00057	0.05761 ± 0.00011	0.057448 ± 0.000074
$M_0 = 0$	0.003 ± 0.012	-0.0055 ± 0.0026	0.00176 ± 0.00088
Data file	`Model6n50SN10.dat`	`Model6n50SN50.dat`	`Model6n100SN100.dat`
Control file	`dcmModel6n50SN10.dat`	`dcmModel6n50SN50.dat`	`dcmModel6n100SN100.dat`

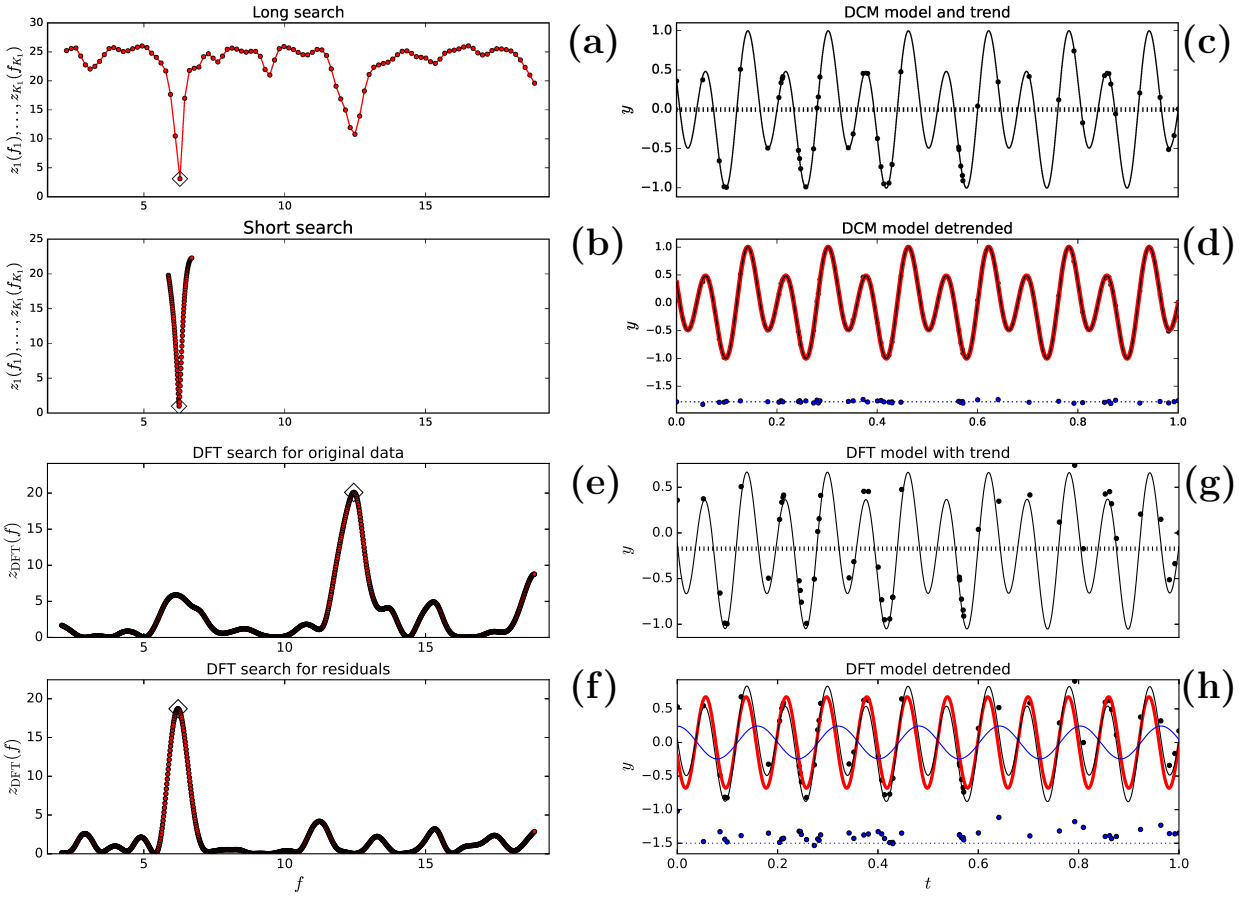

Figure 6. Model 6 (Table 6: $n = 50$, SN = 50 simulation). Notations as in Figure 3. Best periods (diamonds) are explained in Section 3.6.

Table 7. Model 7. DCM analysis between $P_{\min} = 0.4$ and $P_{\max} = 3.6$. Notations as in Table 1.

(1)	(2)	(3)	(4)
	$n = 100$	$n = 1\ 000$	$n = 10\ 000$
Model 7	SN = 5 000 000	SN = 1 000 000	SN = 1 000 000
$P_1 = 1.2$	1.19917 ± 0.00070	1.20000 ± 0.00063	1.20000 ± 0.00024
$A_1 = 2.0$	2.20 ± 0.20	1.94 ± 0.14	1.990 ± 0.057
$t_{1,\min,1} = 0.6892$	0.6835 ± 0.0050	0.6900 ± 0.0043	0.6888 ± 0.0019
$t_{1,\min,2} = 0.1134$	0.1235 ± 0.0098	0.1160 ± 0.0067	0.1128 ± 0.0026
$t_{1,\max,1} = 1.0195$	1.0253 ± 0.0052	1.0176 ± 0.0039	1.0188 ± 0.0018
$t_{1,\max,2} = 0.3779$	0.3705 ± 0.0082	0.3804 ± 0.0050	0.3780 ± 0.0020
$P_2 = 1.4$	1.4050 ± 0.0048	1.3986 ± 0.00034	1.3997 ± 0.0014
$A_2 = 2.0$	1.871 ± 0.098	2.049 ± 0.095	2.009 ± 0.040
$t_{2,\min,1} = 0.9262$	0.938 ± 0.012	0.9231 ± 0.0090	0.9252 ± 0.0036
$t_{2,\min,2} = 0.2766$	0.260 ± 0.016	0.281 ± 0.012	0.2771 ± 0.0046
$t_{2,\max,1} = 1.3109$	1.3165 ± 0.0049	1.3105 ± 0.0039	1.3101 ± 0.0014
$t_{2,\max,2} = 0.5864$	0.5817 ± 0.0037	0.5860 ± 0.0042	0.5865 ± 0.0015
$M_0 = 1.0$	0.950 ± 0.050	1.011 ± 0.035	1.002 ± 0.013
$M_1 = 0.25$	0.2447 ± 0.0067	0.2548 ± 0.0039	0.2509 ± 0.0018
$M_2 = 0.5$	0.551 ± 0.050	0.488 ± 0.035	0.498 ± 0.014
Data file	`Model7n100SN5000000.dat`	`Model7n1000SN1000000.dat`	`Model7n10000SN1000000.dat`
Control file	`dcmModel7n100SN5000000.dat`	`dcmModel7n1000SN1000000.dat`	`dcmModel7n10000SN1000000.dat`

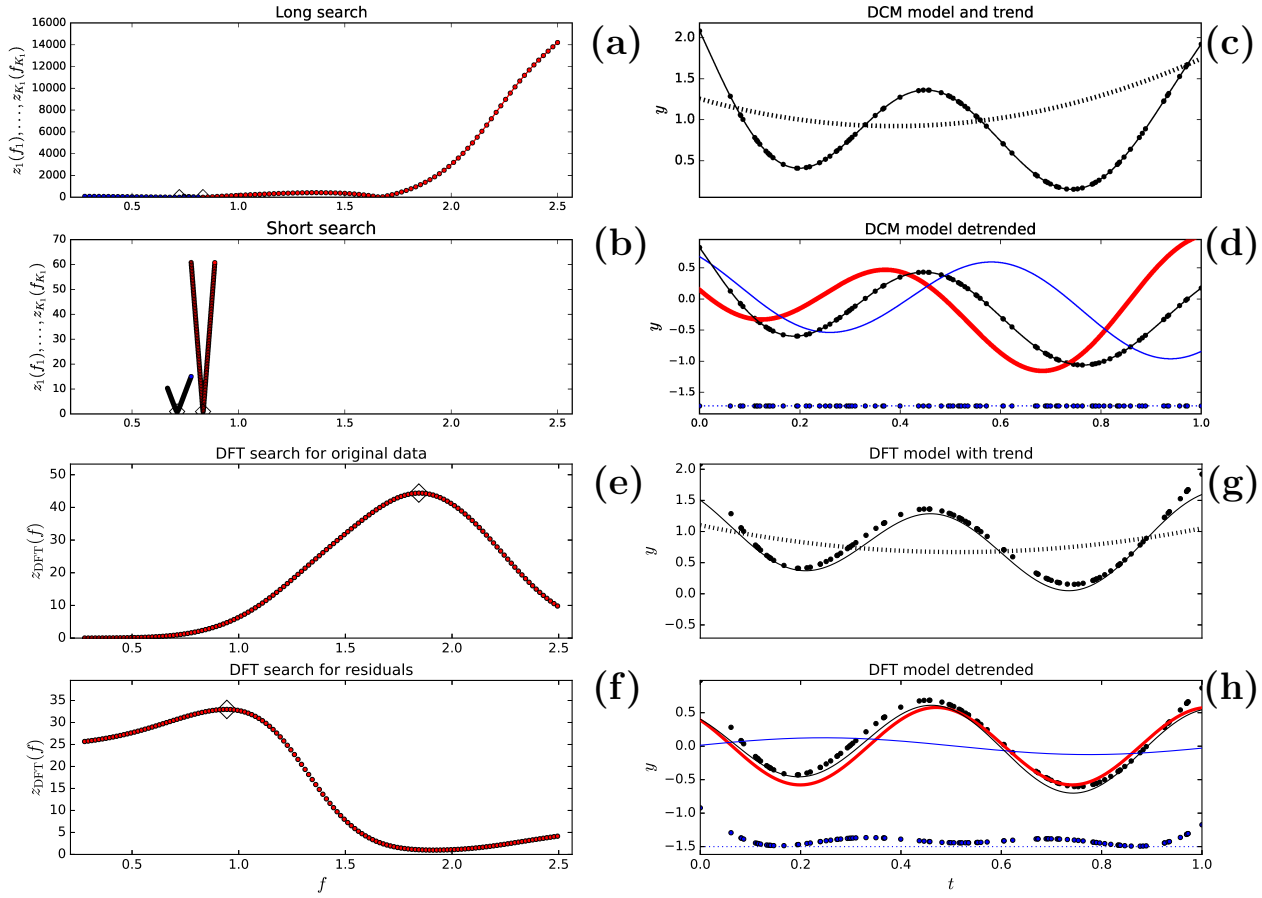

Figure 7. Model 7 (Table 7: $n = 100$, SN = 5 000 000 simulation). Notations as in Figure 3. Best periods (diamonds) are explained in Section 3.7.

are 10 000 of them. For obvious reasons, the -1.5 level of these residuals is highlighted by a white dotted line. DCM model residuals are stable and show no trends.

DFT detects the wrong periods for the original data (Figure 5e: diamond at $P_1 = 0.592$) and for the residuals (Figure 5f: diamond at $P_2 = 1.012$). DFT estimates for the $p(t)$ trend, $M_0 = 2.758$, $M_1 = 0.091$ and $M_2 = -2.120$, are also wrong (Figure 5g: dashed black line). DFT model $g_{\rm DFT}(t)$ black continuous line deviates from the black dots of data y_i, especially in the end of the simulated data sample. The black dots denoting the detrended data $y(t_i) - p_{\rm DFT}(t_i)$ and the continuous black line denoting the detrended DFT model $g_{\rm DFT}(t) - p_{\rm DFT}(t)$ are shown in Figure 5h. DFT model $g_{\rm DFT}(t)$ gives very low amplitudes for the pure sine signal $s_{\rm y,DFT}(t)$ (continuous thick red line) and pure sine signal $s_{\epsilon,\rm DFT}(t)$ (continuous thinner blue line). These amplitudes are far below the correct simulated values $A_1 = A_2 = 2$. The offset level for DFT model residuals (blue dots) is -0.5. These residuals show strong trends.

DCM analysis succeeds for Model 5 simulated time series, but DFT analysis fails.

3.6. *Model 6*

Our sixth time series simulation model is

$$g(t) = c_1 \cos\left[\frac{2\pi t}{P_1}\right] + c_2 \cos\left[\frac{4\pi(t-c_3)}{P_1}\right] + M_0, \quad (44)$$

where $P = 0.16$ and $M_0 = 0$. The coefficients $c_1 = 0.3655$, $c_2 = 0.7310$ and $c_3 = 0.3000$ determine the A_1, $t_{1,\rm min,1}$, $t_{1,\rm min,2}$, $t_{1,\rm max,1}$ and $t_{1,\rm max,2}$ values given in Table 6 (Column 1). This simulated time series is not "too short" because $P_1 < \Delta T$ (Equation 35). There is no trend because $p(t) = M_0 = 0$ (Equation 36). This simulated time series contains only one signal (Equation 37). However, this simulation Model 6 is not a $K_2 = 1$ pure sine model (Equation 38). This double wave simulation model has two unequal minima and maxima. Its DCM orders are $K_1 = 1$, $K_2 = 2$ and $K_3 = 0$. We perform DCM and DFT time series analysis between $P_{\rm min} = P_1/3 = 0.053$ and $P_{\rm max} = 3P_1 = 0.480$.

DCM time series analysis results for different n and SN combinations are given in Table 6 (Columns 2-4). DCM detects the correct P_1, A_1, $t_{1,\rm min,1}$ $t_{1,\rm min,2}$ $t_{1,\rm max,1}$ $t_{1,\rm max,2}$ and M_0 values even for the lowest $n = 50$ and SN = 10 combination. For increasing n and SN, the results for the model parameters converge to correct values (WD-effect).

We demonstrate DCM analysis results for simulated time series having $n = 50$ and SN = 50 (Figures 6a-d). The long and short searches give $P_1 = 0.159$ (Figure 6a: diamond) and $P_1 = 0.160$ (Figure 6b: diamond). DCM model $g(t)$ black line covers the black dots denoting the data y_i (Figure 6c). The "trend" at $p(t) = M_0 = -0.0055 \pm 0.0026$ is correct. The detrended model $g(t) - p(t)$ (black continuous line), the pure sine signal $h_1(t)$ (red thick continuous line) and the detrended data $y(t_i) - p(t_i)$ (black dots) are displayed in Figure 6d. The thick continuous red line stays under the thin continuous black line because $h_1(t) = g(t) - p(t)$. DCM residuals (blue dots) are offset to the level of -1.8 (blue dotted line). These residuals show no trends and their level is stable.

DFT detects the wrong period $P_1 = 0.080$ (Figure 6e: Diamond). This results is exactly half of the correct simulated value $P_1 = 0.160$. The reason for this "detection" is that the double wave dominates because $c_2 = 2c_1$ in Model 6 (Equation 44). DFT mean level estimate $M_0 = -0.172$ fails (Figure 6f: Dashed black line). DFT analysis of the residuals $\epsilon_i = y(t_i) - [s_{\rm y,DFT}(t_i) + p_{\rm DFT}(t_i)]$ gives $P_2 = 0.161$, which is nearly equal to the correct simulated $P_1 = 0.160$ value. The black dots denoting the data y_i show minor deviations from continuous black line denoting DFT model $g_{\rm DFT}(t)$ (Figure 6g). The detrended model $g_{\rm DFT}(t) - p_{\rm DFT}(t)$ (continuous black line), the detrended data $y(t_i) - p_{\rm DFT}(t_i)$ (black dots) and the signal $s_{\rm y,DFT}(t)$ (continuous thick red line) are shown in Figure 6h. DFT analysis residuals (blue dots) are offset to the level of -1.5 (blue dotted line). These residuals display trends.

We conclude that DFT "detects" the $P_1/2$ and P_1 periods, where P_1 is the correct simulated period value. However, DFT two signal model is not the correct model for these Model 6 simulated time series, which contains only one signal. If the correct period is P and the correct model is a double wave ($K_2 = 2$), DCM pure sine model ($K_1 = 1$) analysis can also give the values $P/2$ and P.

DCM analysis succeeds for Model 6 simulated time series. DFT analysis fails.

3.7. *Model 7*

In this section, we analyse our most complex time series. The seventh simulation model is

$$g(t) = c_1 \cos\left[\frac{2\pi t}{P_1}\right] + c_2 \cos\left[\frac{4\pi(t-c_3)}{P_1}\right] + c_4 \cos\left[\frac{2\pi t}{P_2}\right] + c_5 \cos\left[\frac{4\pi(t-c_6)}{P_2}\right] + M_0 + M_1 T + M_2 T^2 \quad (45)$$

where $T = [2(t - t_{\rm mid})]/\Delta T$, $P_1 = 1.2$, $P_2 = 1.4$, $M_0 = 1$ $M_1 = 0.25$ and $M_2 = 0.5$. The coefficients $c_1 = 0.3687$, $c_2 = 0.7374$, $c_3 = 0.4000$, $c_4 = 0.3708$, $c_5 = 0.7416$ and

$c_6 = 0.6000$ determine the A_1, $t_{1,\min,1}$, $t_{1,\min,2}$, $t_{1,\max,1}$, $t_{1,\max,2}$, A_2, $t_{2,\min,1}$, $t_{2,\min,2}$, $t_{2,\max,1}$ and $t_{2,\max,2}$ values given in Table 7 (Column 1). This simulated time series is "too short" because $\Delta T < P_1 < P_2$ (Equation 35). The parabolic $p(t)$ represents "a trend" (Equation 36). The signal frequencies f_1 and f_2 are "too close" because $\Delta f = f_1 - f_2 = 0.12 < f_0 = 1$ (Equation 37). The two $h_1(t)$ and $h_2(t)$ signals are not "pure sines" (Equation 38). These signals are double waves having two unequal minima and maxima. All four main reasons that can cause the failure of DFT analysis are present (Equations 35-38). Therefore, this simulated time series is the most complex one of all analysed seven time series. Our DCM and DFT time series analysis of Model 7 simulated time series is performed between $P_{\min} = P_1/3 = 0.4$ and $P_{\max} = 3P_1 = 3.6$.

DCM orders of Model 7 are $K_1 = 2$, $K_2 = 2$ and $K_3 = 2$. This model has $\eta = 13$ free parameters (Equation 8). We give DCM analysis results for different n and SN combinations in Table 7 (Columns 2-4). DCM analysis results are displayed for one sample of Model 7 simulated time series (Figures 7a-d: $n = 100$ and SN = 5 000 000). Since there are $\eta = 13$ free model parameters, this $n = 100$ time series is quite small for time series analysis. The long and short DCM searches give $P_1 = 1.200$ and $P_2 = 1.385$ (Figure 7a: diamonds), and $P_1 = 1.199$ and $P_2 = 1.405$ (Figure 7b: Diamonds). The black $g(t)$ model line goes through all black y_i data dots (Figure 7c). DCM detects the correct polynomial trend $p(t)$ coefficients $M_0 = 0.950 \pm 0.050$, $M_1 = 0.2447 \pm 0.0067$ and $M_2 = 0.551 \pm 0.050$ (Figure 7c: dashed black line). The detrended DCM model $g(t) - p(t)$ (white continuous line), the detrended data $y(t_i) - p(t_i)$ (black dots), the signal $h_1(t)$ (red thick continuous line) and the signal $h_2(t)$ (blue thin continuous line) are displayed in Figure 7d. The $n = 100$ residuals (blue dots) are offset to the level of -1.8 (dotted blue line). These DCM model residuals are small and their level is stable. These results confirm that DCM time series analysis method can detect complex non-linear models ($\eta = 13$) from very small time series ($n = 100$), if the data are extremely accurate (SN = 5 000 000). These results demonstrate the power of WD-effect because both periods P_1 and P_2 are shorter than the sample window ΔT, the trend is a parabola and the signals are not pure sines.

DFT time series analysis gives the wrong periods for the original data (Figure 7e: diamond at $P_1 = 0.542$ and for the residuals (Figure 7f: diamond at $P_2 = 1.059$. DFT also gives wrong $p(t)$ trend coefficients $M_0 = 0.670$, $M_1 = -0.025$ and $M_2 = 0.406$ (Figure 7g: dashed black line). DFT model $g_{\mathrm{DFT}}(t)$ (black continuous line) deviates from the data y_i (black dots). This deviation is largest at the beginning and the end of the simulated time series. In our Figure 7h, the black dots denote the detrended data $y(t_i) - p_{\mathrm{DFT}}(t_i)$, the continuous black line denotes the detrended DFT model $g_{\mathrm{DFT}}(t) - p_{\mathrm{DFT}}(t)$ and the continuous thick red line denotes pure sine signal $s_{\mathrm{y,DFT}}(t)$ detected from the original data. DFT model $g_{\mathrm{DFT}}(t)$ gives very low amplitude for the pure sine signal $s_{\epsilon,\mathrm{DFT}}(t)$ detected from the residuals (continuous thinner blue line). The correct simulated peak to peak amplitude for this second signal is much larger, $A_2 = 2$. DFT model residuals (blue dots) are offset to -1.5 level and show strong trends.

DCM analysis succeeds for Model 7 simulated time series. DFT analysis fails, as predicted by Equations 35-38.

3.8. *Summary of simulations*

For all seven simulated time series, all DCM analyses succeed and all DFT analyses fail (Sections 3.1 - 3.7).

DCM detects exactly the correct signal(-s) and trend. Clearly, the AL6-AL10 do not constrain our DCM. We do admit that the simulated n and SN values of Models 2, 5 and 7 are extreme and unrealistic for most cases of real data. However, those model simulations are necessary for demonstrating the WD-effect. We could keep on adding complexity by simulating a larger number of signals (K_1), more complex signal shapes (K_2) and/or higher order polynomial trends (K_3). The WD-effect ensures that DCM would detect those complex models when the sample size (n) and/or the signal to noise (SN) increase. The other way round, the correct model can be found *if* the correct DCM model combination K_1, K_2 and K_3 is tested. Furthermore, totally wrong K_1, K_2 and K_3 combinations give unstable DCM models ("UM").

The analyses of seven simulated time series expose DFT's weaknesses. Every analysis fails! The AL6-AL10 are not the only causes for these failures. In general, the solutions for one-dimensional time series analysis models are ill-posed, if the data contain an unknown trend and/or more than one signal. For example, the detrending of simulated data misleads DFT even when the correct K_3 value is known (Sections 3.1-3.7). This correct K_3 value can be unknown for real data. The combination of detrending and iterative pre-whitening can fail at any stage. Even if the detrending succeeds, the leakage can cause the detection of wrong first signal frequency. This corrupts the model residuals and the whole analysis. For any one-dimensional time series analysis method, like DFT, the search for many signals

superimposed on an unknown trend can fail when the combination of detrending and pre-whitening is applied.

Our DCM outperforms the esteemed DTF. DCM detects the correct signal(s) and trend, but DFT does not. For the sunspot data, DCM detects signals superimposed on a constant trend (Jetsu 2025). The simulations presented here support the conclusion that those solar signals and the trend are correct. This would explain why DCM is the first time series analysis method that detects Jupiter's exact $11.^{y}86$ period in the sunspot record. The one-dimensional time series analysis methods, like DFF, have failed to find exactly correct periods.

3.9. *Best model*

The K_1, K_2 and K_3 orders of the best model are not necessarily known when some time series analysis method is applied to the real time series. We know a priori the best DCM and DFT model orders for the simulated time series of Models 1-7 (Sections 3.1-3.7). It could be argued that our DCM analysis succeeds only for this reason.

All alternative K_1, K_2 and K_3 order models are nested. For example, the *simple* one signal $K_1 = 1$ model $g_1(t)$ is a special case of the *complex* two signal $K_1 = 2$ model $g_2(t)$ having $A_2 = 0$. We use the Fisher-test to compare any pair of simple $g_1(t)$ and complex $g_2(t)$ models. The model parameters (Equation 14: R_1, R_2), (Equation 15: χ_1,χ_2) and (Equation 8: $\eta_1 < \eta_2$) give the test statistic of Fisher-test

$$F_R = \left(\frac{R_1}{R_2} - 1\right)\left(\frac{n - \eta_2 - 1}{\eta_2 - \eta_1}\right) \tag{46}$$

$$F_\chi = \left(\frac{\chi_1^2}{\chi_2^2} - 1\right)\left(\frac{n - \eta_2 - 1}{\eta_2 - \eta_1}\right). \tag{47}$$

The Fisher-test is based on the null hypothesis

> H_0: *"The complex model $g_2(t)$ does not provide a significantly better fit to the data than the simple model $g_1(t)$."*

Under this hypothesis, the F_R and F_χ test statistic parameters have an F distribution with $\nu_1 = \eta_2 - \eta_1$ and $\nu_2 = n - \eta_2$ degrees of freedom (Draper & Smith 1998). The critical level $Q_F = P(F_R \geq F)$ or $Q_F = P(F_\chi \geq F)$ is the probability that F_R or F_χ exceeds the numerical value F. If

$$Q_F < \gamma_F = 0.001, \tag{48}$$

we reject the H_0 hypothesis, which means that the complex $g_2(t)$ model is better than the simple $g_1(t)$ model. The pre-assigned significance level $\gamma_F = 0.001$ represents the probability that we falsely reject the H_0 hypothesis when it is in fact true.

Larger F_R or F_χ values have smaller Q_F critical levels. Hence, the probability for the H_0 hypothesis rejection increases when F_R or F_χ increases. If the number of complex model free parameters η_2 increases, the R_2 or χ_2^2 values decrease. This increases the F_R or F_χ values because the terms $(R_1/R_2 - 1)$ or $(\chi_1^2/\chi_2^2 - 1)$ increase (Equations 46 and 47: first terms). At the same time, the $(n - \eta_2 - 1)/(\eta_2 - \eta_1)$ penalty term decreases (Equations 46 and 47: second terms), and this decreases the F_R or F_χ values. This penalty term prevents the use of too high η_2 values (too complex models).

Here, we illustrate how the Fisher-test finds the best model from a group of numerous alternative nested DCM models. The Fisher-test is used to find the best model for the simulated time series of Model 1 combination $n = 100$ and SN = 100 (Table 1: column 4). In other words, we assume that the correct DCM model orders K_1, K_2 and K_3 are unknown, which can be the case for real time series. The eight tested models contain one or two signals, and no trend or a constant trend or a linear trend or a parabolic trend. We compare all these eight models $\mathcal{M}$=1-8 against each other (Table 8). The special model number notation"$\mathcal{M}$" is used because the notations "M_0, ..., M_{K_3}" have already been reserved for the $p(t)$ trend.

Model $\mathcal{M}$=2 has the known correct Model 1 orders $K_1 = 1$, $K_2 = 1$ and $K_3 = 0$. For example, the Fisher-test between the simple model $\mathcal{M}$=1 ($\eta = 3$, $\chi^2 = 738$) and the complex model $\mathcal{M}$=2 ($\eta = 4$, $\chi^2 = 84.7422$) gives $F = 733$ (Equation 47). The critical level[3] for this very large F value is extremely significant, $Q_F < 10^{-16}$. This means that the H_0 hypothesis must be rejected, and the complex $\mathcal{M}$=2 model is certainly better than the simple $\mathcal{M}$=1 model (Equation 48). The upward arrow "↑" in Table 8 indicates that $\mathcal{M}$=2 model is a better model than $\mathcal{M}$=1 model. A closer look at Table 8 reveals that $\mathcal{M}$=2 model is better than all other models because the "↑" and "←" arrows of all other models point toward $\mathcal{M}$=2 model. There is no need to test models having more than two signals because all two signal $\mathcal{M}$=5-8 models are already unstable (Table 8:"UM"). The Fisher-test finds the correct DCM model for Model 1 simulated time series.

In the above example, the Fisher-test finds the correct number of signals (K_1) and the correct trend (K_3) for the pure sine signal alternative ($K_2 = 1$). We do not test the double wave signal alternative ($K_2 = 2$) against the pure sine signal alternative ($K_1 = 1$) because the

[3] The highest achievable accuracy for the computational f.cdf subroutine in scipy.optimize python library is 10^{-16}.

Table 8. Fisher tests. DCM analysis of Model 1 simulated time series ($n = 100, \mathrm{SN} = 100$, data file `Model1n100SN100.dat`). (1) Simple model: Model number $\mathcal{M}$, model g_{K_1,K_2,K_3}, number of free parameters η, chi-square χ^2 and control file name. (2-8) Complex model: Hypothesis H_0 rejected ↑ (Complex model better), Hypothesis H_0 not rejected ← (Simple model better), test statistic F and critical level Q_F. Note that no Fisher-test is done between models $\mathcal{M}$=4 and $\mathcal{M}$=5 because $\eta_1 = \eta_2 = 6$.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Fisher-test ($\gamma = 0.001$)			
				Complex model			
Simple model	$\mathcal{M}$=2	$\mathcal{M}$=3	$\mathcal{M}$=4	$\mathcal{M}$=5	$\mathcal{M}$=6	$\mathcal{M}$=7	$\mathcal{M}$=8
$\mathcal{M}$=1	↑	↑	↑	↑	↑	↑	↑
$g_{1,,1,-1}$	$F = 733$	$F = 367$	$F = 242$	$F = 242$	$F = 180$	$F = 142$	$F = 121$
$\eta = 3, \chi^2 = 738$	$Q_\mathrm{F} < 10^{-16}$	$Q_\mathrm{F} < 10^{-16}$	$Q_\mathrm{F} < 10^{-16}$	$Q_\mathrm{F} < 10^{-16}$	$Q_\mathrm{F} < 10^{-16}$	$Q_\mathrm{F} < 10^{-16}$	$Q_\mathrm{F} < 10^{-16}$
`dcmModel1K11-1.dat`							
$\mathcal{M}$=2	-	←	←	←	←	←	←
$g_{1,1,0}$	-	$F = 1.0451$	$F = 0.5404$	$F = 0.5159$	$F = 0.3553$	$F = 0.2784$	$F = 0.7686$
$\eta = 4, \chi^2 = 84.7422$	-	$Q_\mathrm{F} = 0.309$	$Q_\mathrm{F} = 0.584$	$Q_\mathrm{F} = 0.599$	$Q_\mathrm{F} = 0.785$	$Q_\mathrm{F} = 0.891$	$Q_\mathrm{F} = 0.575$
`dcmModel1K110.dat`							
$\mathcal{M}$=3	-	-	←	←	←	←	←
$g_{1,1,,1}$	-	-	$F = 0.0464$	$F = -0.0021$	$F = 0.0213$	$F = 0.0336$	$F = 0.7027$
$\eta = 5, \chi^2 = 83.8104$	-	-	$Q_\mathrm{F} = 0.830$	$Q_\mathrm{F} = 1$	$Q_\mathrm{F} = 0.979$	$Q_\mathrm{F} = 0.992$	$Q_\mathrm{F} = 0.592$
`dcmModel1K111.dat`							
$\mathcal{M}$=4	-	-	-	←	←	←	←
$g_{1,1,2}$	-	-	-	No Test	$F = -0.0033$	$F = 0.0277$	$F = 0.9215$
$\eta = 6, \chi^2 = 83.7686$	-	-	-		$Q_\mathrm{F} = 1$	$Q_\mathrm{F} = 0.973$	$Q_\mathrm{F} = 0.434$
`dcmModel1K112.dat`							
$\mathcal{M}$=5 UM : AD	-	-	-	-	←	←	←
$g_{2,1,-1}$	-	-	-	-	$F = 0.0447$	$F = 0.0515$	$F = 0.9377$
$\eta = 6, \chi^2 = 83.8123$	-	-	-	-	$Q_\mathrm{F} = 0.833$	$Q_\mathrm{F} = 0.950$	$Q_\mathrm{F} = 0.426$
`dcmModel1K21-1.dat`							
$\mathcal{M}$=6 UM : IF, AD	-	-	-	-	-	←	←
$g_{2,1,0}$	-	-	-	-	-	$F = 0.0587$	$F = 1.3840$
$\eta = 7, \chi^2 = 83.7716$	-	-	-	-	-	$Q_\mathrm{F} = 0.809$	$Q_\mathrm{F} = 0.256$
`dcmModel1K210.dat`							
M=7 UM : IF, AD	-	-	-	-	-	-	←
$g_{2,1,1}$	-	-	-	-	-	-	$F = 2.7081$
$\eta = 8, \chi^2 = 83.7176$	-	-	-	-	-	-	$Q_\mathrm{F} = 0.103$
`dcmModel1K211.dat`						-	
$\mathcal{M}$=8 UM : IF, AD	-	-	-	-	-	-	-
$g_{2,1,2}$	-	-	-	-	-	-	-
$\eta = 9, \chi^2 = 81.2721$	-	-	-	-	-	-	-
`dcmModel1K212.dat`						-	

number of tested models would increase from 8 to 16, and Table 8 would become excessively large. One example of testing the $K_2 = 2$ signal models against each other can be found in Jetsu (2025, Table S7).

We conclude that the best model for the real time series can be found by applying the Fisher-test to any arbitrary number of different nested DCM or DFT models.

3.10. *Significance estimates*

Jetsu (2020) or Jetsu (2025) gave no signal significance estimate for the first detected period. Here, we use the Fisher-test for this purpose. The one signal model is the complex model. The logical simple model alternative is $g(t) = p(t) = m \equiv$ white noise having standard deviation s. However, white noise is not the only possible alternative simple model. The other possible nested no signal DCM polynomial models are

$$g(t) = p(t, K_3) \tag{49}$$

where $h(t) = 0$, $K_3 = 0, 1, 2...$ The $K_3 = 0$ polynomial $g(t) = p(t) = M_0$ represents white noise. The Fisher-test gives the critical level Q_F for rejecting the H_0 hypothesis when this hypothesis is in fact true (Equation 48). Therefore, this Q_F value represents the probability of false signal detection. In cases $Q_\mathrm{F} < 10^{-16}$, the signal detection is absolutely certain.

We use the Model 1 simulated time series combination $n = 100$ and $\mathrm{SN} = 100$ to demonstrate the significance estimation for the first detected period. The correct model for the first detected signal is $g_{1,1,0}(t)$. The Fisher-test is used to compare this correct $g_{1,1,0}(t)$ model to different polynomial models $g(t) = p(t, K_3)$ having $K_3 = 0, 1, .., 7$ (Table 9: Equation 49). The

$\mathcal{M}$=1, 2 and 3 models have less free parameters than the correct $g_{1,1,0}(t)$ model, but the χ^2 values of these three polynomial models are so large that the Fisher-test is merely a formality. The $\mathcal{M}$=4 model has the same number of free parameters as the correct $g_{1,1,0}(t)$ model, but the comparison of χ^2 values reveals that this third order $p(t)$ polynomial is not the correct model for the time series. The next model $\mathcal{M}$=5 has more free parameters than the correct $g_{1,1,0}(t)$ model. This fourth order $p(t)$ polynomial $\mathcal{M}$=5 model must be rejected because it has a larger χ^2 value than the correct $g_{1,1,0}(t)$ model. The critical levels $Q_{\rm F} \gg \gamma = 0.001$ for the remaining $\mathcal{M}$=6, 7 and 8 polynomial models are so large that these models must also be rejected. The results in Table 9 confirm that the one signal model is better than any polynomial model in Equation 49. Hence, the analysed Model 1 simulated time series must contain at least the $h_1(t)$ pure sine signal. For the constant, linear or parabolic $p(t)$ trend alternatives, the significance for this $h_1(t)$ signal is $Q_{\rm F} < 10^{-16}$.

After the detection of the first strongest signal, the Fisher-test critical level $Q_{\rm F}$ values increase for the next detected weaker signals. In other words, the signal significances decrease. Typical examples can be found in Jetsu (2025, Tables S5-S16). No new signals are detected when $Q_{\rm F} > \gamma = 0.001$ because the critical level exceeds the pre-assigned significance level and the H_0 hypothesis is no longer rejected. For the Model 1 simulated $n = 100$ and SN = 100 combination, Table 8 shows that this time series contains only one signal.

For this Model 1 time series, there is no need to discuss DFT significance estimates (Horne & Baliunas 1986, their Equation 22) because this method fails to detect the correct period.

We conclude that the Fisher-test identifies the correct DCM model (Section 3.9), as well as gives the signal significance $Q_{\rm F}$ estimates (Section 3.10).

3.11. *Ill-posed problem*

The solution for the non-linear DCM model is an ill-posed problem (Equations 1-5). We present a computational statistical solution that fulfils the C_1, C_2 and C_3 conditions of a well-posed problem. DCM model is just one special case of a non-linear model. Our technique can be applied to solve other non-linear models: Use the free parameters that make the model non-linear (Equation 9: β_I) for solving the remaining other free parameters (Equation 10: β_{II}).

3.11.1. *Existence* (C_1)

Fourier (1822) transformed the original function into the frequency domain. The modern DFT time series analysis method transforms the original time series into the frequency domain and gives the best frequency for a pure sine model. Gauß (1821) presented the LS method, which minimises the differences between the data and the linear model. Our DCM does the same by testing a large number of linear models. The data spacing, even or uneven, is irrelevant for these models. For every chosen DCM model, the total number of tested linear models is

$$n_{\rm Lin} = \binom{n_{\rm L}}{K_1} + (1 + n_{\rm B}) \times \binom{n_{\rm S}}{K_1}, \qquad (50)$$

where the number of tested long and short search frequencies is $n_{\rm L}$ and $n_{\rm S}$, respectively. The number of bootstrap samples is $n_{\rm B}$. In other words, we solve this ill-posed problem by using brute computational force. If

Table 9. Fisher tests for Model 1 $n = 100, {\rm SN} = 100$ simulated time series (electronic data file `Model1n100SN100.dat`). (1) Polynomial model (Equation 49). (2) DCM model $g_{1,1,0}$. Note that $\mathcal{M}$=1-3 polynomials represent simple models, and $\mathcal{M}$=5-8 polynomials represent complex models. Otherwise as in Table 8.

(1)	(2) Fisher-test ($\gamma = 0.001$) $g_{1,1,0}$ $\eta = 4, \chi^2 = 84.7422$
Polynomial	
$\mathcal{M}$=1	↑
$g(t) = p(t, K_3 = 0)$	$F = 63218$
$\eta = 1$, $\chi^2 = 169260$	$Q_{\rm F} < 10^{-16}$
$\mathcal{M}$=2	↑
$g(t) = p(t, K_3 = 1)$	$F = 54366$
$\eta = 2$, $\chi^2 = 97076$	$Q_{\rm F} < 10^{-16}$
$\mathcal{M}$=3	↑
$g(t) = p(t, K_3 = 2)$	$F = 1945$
$\eta = 3$, $\chi^2 = 1820$	$Q_{\rm F} < 10^{-16}$
$\mathcal{M}$=4	↑
$g(t) = p(t, K_3 = 3)$	No test
$\eta = 4$, $\chi^2 = 469$	
$\mathcal{M}$=5	↑
$g(t) = p(t, K_3 = 4)$	$F = -0.3718$
$\eta = 5$, $\chi^2 = 85.0787$	$Q_{\rm F} = 1$
$\mathcal{M}$=6	↑
$g(t) = p(t, K_3 = 5)$	$F = 0.3403$
$\eta = 6$, $\chi^2 = 84.1265$	$Q_{\rm F} = 0.712$
$\mathcal{M}$=7	↑
$g(t) = p(t, K_3 = 6)$	$F = 0.3524$
$\eta = 7$, $\chi^2 = 83.7794$	$Q_{\rm F} = 0.787$
$\mathcal{M}$=8	↑
$g(t) = p(t, K_3 = 7)$	$F = 0.9394$
$\eta = 8$, $\chi^2 = 81.3819$	$Q_{\rm F} = 0.445$

the time series contains only zero mean white noise, the Gauß-Markov theorem ensures that a LS fit solution *exists* for every tested frequency combination.

3.11.2. *Uniqueness* (C_2)

We make the most of the Gauß-Markov theorem. When the numerical values of the tested frequencies β_I (Equation 9) are fixed, the model becomes linear and the solution for the other free parameters β_{II} is *unique* (Equation 10). All possible β_I frequency combinations are tested (Equation 12). For every tested frequency combination β_I, the linear model gives a *unique* value for the test statistic z (Equations 16 and 17). From all tested frequency combinations, we select the best frequency combination $\beta_{\rm I,best}$ which minimises z. The linear model for this best frequency combination $\beta_{\rm I,best}$ gives *unique* values for the remaining other free parameters $\beta_{\rm II,best}$. The only goal for the massive DCM search (Equation 50) is to find these *unique* initial free parameter values $\beta_{\rm initial} = [\beta_{\rm I,best}, \beta_{\rm II,best}]$ for the non-linear iteration that gives the *unique* final free parameter values $\beta_{\rm final}$ (Equation 22).

We use the Fisher-test to compare many different non-linear DCM models against each other. The selection criterion for the best model is *unique* (Equation 48). The best DCM model is not necessarily the correct model, if this correct model is not among the compared models. The correct model must be able to forecast the future and past data. We formulate the Forecast-test for alternative DCM models (Equation 53). The order of Fisher- and Forecast-tests can be reversed. However, the former uses all data, while the latter uses a subset of all data, the forecasting data. In ideal cases, both tests identify the same best and correct model, like in Section 3.12.

3.11.3. *Stability* (C_3)

The artificial bootstrap data sets (Equation 31) represent "small changes in the input data", while the bootstrap results for the model parameters represent "small chances in the solution" (Section 1: C_3 condition formulation). We routinely check the *stability* of these solutions (Jetsu 2025, Figure S5). The unstable models, where the model parameter changes are large, are rejected (Section 2.1: "UM" models).

There are additional signatures of *stability*. The z periodogram solution is unique for every tested β_I frequency combination. If these periodograms are continuous and their changes are not irregular (e.g., like in Figs. 1a-b), DCM model solution is *stable* because it does not change by increasing the number of tested frequencies $n_{\rm L}$ and $n_{\rm S}$. Furthermore, the solutions for all seven simulated time series converge when the n and SN values increase (Tables 1 - 7). The different n and SN combinations give the same *stable* DCM model solution.

Table 10. DCM time series analysis between $P_{\rm min} = 0.53$ and $P_{\rm max} = 4.80$ for forecasting data (Model = 3 combination $n = 50$ and SN = 10 simulated data: first forty observations). (1) Model. (2) Forecasting data test statistic z (Equation 17). (3) Forecasted data test statistic $z_{\rm Fore}$ (Equation 53). (4) Data file. (5) Control file. (6) Figure where forecast is shown.

(1) Model	(2) z	(3) $z_{\rm Fore}$	(4) Data file	(5) Control file	(6) Fig
$g_{1,1,0}$	2.27	8.61	`Model3n40SN10.dat`	`dcmModel3K110.dat`	8a
$g_{2,1,0}$	0.76	0.88	`Model3n40SN10.dat`	`dcmModel3K210.dat`	8b
$g_{3,1,0}$	0.70	0.91	`Model3n40SN10.dat`	`dcmModel3K310.dat`	8c

We conclude that our computational statistical DCM model solution fulfils the C_1, C_2 and C_3 conditions of the solution for a well-posed problem (Sections 3.11.1 - 3.11.3).

3.12. *Forecast*

There are numerous techniques for forecasting a time series (e.g., Hamilton 1994; Hastie et al. 2001; Kazemi & Rodrigues 2025). DCM model $g(t_i, \beta)$ can be used to forecast. We divide all data into the forecasting data and the forecasted data. The time points, the observations and the errors of these samples are

n forecasting data values t_i, y_i and σ_i

n' forecasted data values t'_i, y'_i and σ'_i

DCM gives the best forecasting data model

$$g_i = g(t_i, \beta_{\rm Fore}), \tag{51}$$

where $\beta_{\rm Fore}$ are the free parameter values. The $t_{\rm mid,Fore} = (t_n + t_1)/2$ and $\Delta T_{\rm Fore} = t_n - t_1$ values are computed from the forecasting data time points t_i. The forecasting model values at any arbitrary time t can be obtained from the $\beta_{\rm Fore}$, $t_{\rm mid,Fore}$ and $\Delta T_{\rm Fore}$ values.

The forecasted data model values are

$$g'_i = g(t'_i, \beta_{\rm Fore}), \tag{52}$$

where $t_{\rm mid} = t_{\rm mid,Fore}$ and $\Delta T = \Delta T_{\rm Fore}$. We do not compute "new" $t_{\rm mid}$ and ΔT values from the forecasted data time points t'_i because the correct g'_i values are obtained only from the $\beta_{\rm Fore}$, $t_{\rm mid,Fore}$ and $\Delta T_{\rm Fore}$ combination of the forecasting model (Equation 52). The n' forecasted data model residuals

$$\epsilon'_i = y'_i - g'_i$$

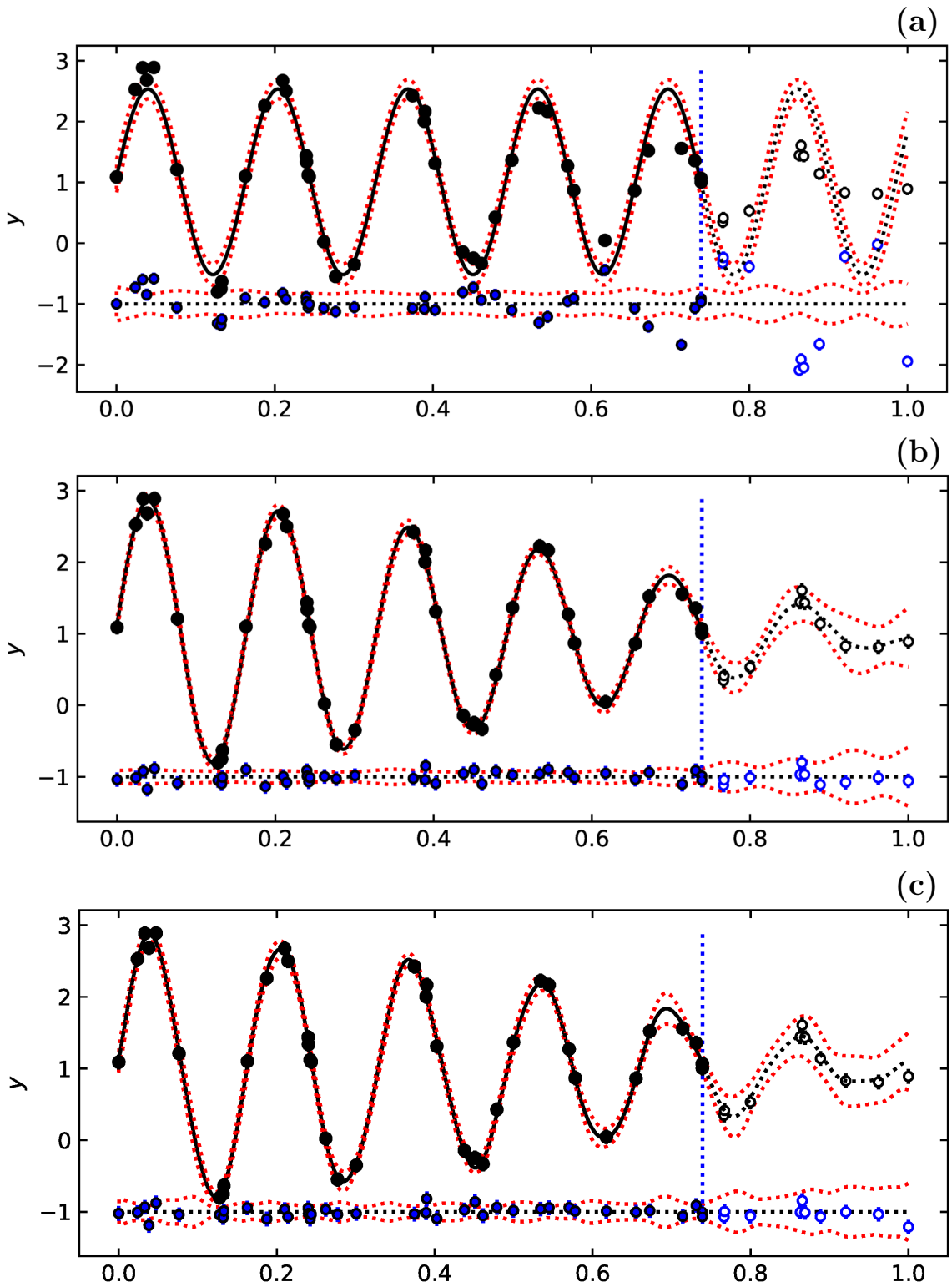

Figure 8. DCM model forecast (Model 3 simulated time series for combination $n = 50$ and SN = 10: first forty observations are forecasting data). (a) One signal model $g_{1,1,0}$ results. Forecasting data are t_i, y_i and σ_i ($n = 40$ black dots) and forecasted data are ten last observations t'_i, y'_i and σ'_i ($n = 10$ open black dots). Continuous black line of forecasting data model $g(t, \beta_{\mathrm{Fore}})$ ends to vertical dotted blue line, where dotted black line of forecasted model $g'(t, \beta_{\mathrm{Fore}})$ begins. Dotted red line denotes $\pm 3\sigma_{g(t)}$ errors of both models (Equation 30). Residuals of forecasting data model $\epsilon_i = y_i - g_i$ (blue dots), residuals of forecasted data model $\epsilon_i =' y'_i - g'_i$ (open blue dots) and $\pm 3\sigma$ errors of both models (red dotted line) are offset to level -1 (blue dotted line). (b) Two signal model $g_{2,1,0}$ results. Otherwise as in "a". (c) Three signal model $g_{3,1,0}$ results. Otherwise as in "a".

give the *forecasted data test statistic*

$$z_{\rm Fore} = \text{z test statistic for forecasted data } t'_i,\ y'_i \text{ and } \sigma'_i \text{ (Equation 16 or 17).} \quad (53)$$

This parameter $z_{\rm Fore}$ measures how well the forecast (Equation 52) obtained from the forecasting data works for the forecasted data. If DCM detects a new signal from the forecasting data, there are two alternatives:

> If this new signal is real, the $z_{\rm Fore}$ value of forecasted data decreases.
>
> If this new signal is unreal, the $z_{\rm Fore}$ value of forecasted data increases.

This "Forecast-test" technique revealed at least five real signals in the sunspot record (Jetsu 2025)

We compute the $z_{\rm Fore}$ parameter value from the *known* forecasted data t'_i, y'_i and σ'_i values (Equation 53). Forecasts are possible even if all t'_i, y'_i or σ'_i forecasted data values are *unknown*. In this case, the t_i, y_i and σ_i values of all data can be used as forecasting data. The best DCM model $g(t_i, \beta_{\rm Fore})$ for all data determines the correct $\beta_{\rm Fore}$, $t_{\rm mid,Fore} = (t_n + t_1)/2$ and $\Delta T_{\rm Fore} = t_n - t_1$ combination of the forecasting model. The n' forecasted data time points t'_i can be created for any arbitrary chosen sample window $\Delta T' = t'_n - t'_i$. The n' forecasted $g'_i = g'_i(t'_i, \beta_{\rm Fore})$ values are obtained from Equation 52. These g'_i values can be used, for example, to compute the *forecasted mean level*

$$m_{\rm Fore} = \frac{1}{n'} \sum_{i=1}^{n'} g'_i \quad (54)$$

during the chosen sample window $\Delta T' = t'_n - t'_1$. Jetsu (2025) used this $m_{\rm Fore}$ parameter to postdict the known $\Delta T'$ time intervals of prolonged solar activity minima, like the Maunder minimum between the years 1640 and 1720 (Usoskin et al. 2007). Since the known mean level of all sunspot data was m, Jetsu (2025) used these three criteria for correct postdictions of past prolonged activity minima $\Delta T'$ time intervals:

> If DCM detects a real new signal in all data, the $m_{\rm Fore}$ value decreases.
>
> If DCM detects many real signals in all data, the $m_{\rm Fore}$ value falls below m.
>
> If DCM detects an unreal new signal in all data, the $m_{\rm Fore}$ value increases.

We use Model 3 combination $n = 50$ and SN = 10 simulated data (Table 3, column 2) to illustrate DCM forecasting technique. The black dots in Figures 8a-c are the first $n = 40$ forecasting data values t_i, y_i and σ_i. The open black dots denote the $n' = 10$ forecasted data values t'_i, y'_i and σ'_i. The continuous black line is the forecasting model $g(t)$ and the dotted black line is the forecasted model $g'(t)$. The red dotted line shows the $\pm 3\sigma$ errors of both models. The blue dots are the forecasting data $\epsilon_i = y_i - g_i$ residuals and the open blue dots are the forecasted data $\epsilon'_i = y'_i - g'_i$ residuals.

Model 3 is the sum of two $P_1 = 0.16$ and $P_2 = 0.17$ pure sine signals ($K_1 = 2, K_2 = 1$) superimposed on the constant mean level $M_0 = 1$ ($K_3 = 0$). We compute the z and $z_{\rm Fore}$ values for the $g_{1,1,0}$, $g_{2,1,0}$ and $g_{3,1,0}$ models (Table 10). These three models have the same $K_2 = 1$ and $K_3 = 0$ orders as Model 3, but their signal numbers $K_1 = 1, 2$ or 3 are different, the $g_{2,1,0}$ model being the correct simulation Model 3.

The correct $g_{2,1,0}$ model gives the smallest $z_{\rm Fore} = 0.88$ value (Table 10). Therefore, it is a better forecasting model than the $g_{1,1,0}$ and $g_{3,1,0}$ models. The one signal $g_{1,1,0}$ model forecast fails because the blue open circles denoting the forecasted data residuals $\epsilon'_i = y'_i - g'_i$ show large deviations from to the blue dotted line offset level of $\epsilon'_i = -1$ (Figure 8a). The two signal $g_{2,1,0}$ model forecast succeeds because all open blue dots denoting the forecasted data residuals stay close to the blue dotted line offset level $\epsilon'_i = -1$, as well as inside the red dotted $\pm 3\sigma$ model error limits (Figure 8b).

The three signal model $g_{3,1,0}$ periods are $P_1 = 0.0643 \pm 0.0018$, $P_2 = 0.1592 \pm 0.0012$ and $P_3 = 0.1716 \pm 0.0012$. The amplitudes of these pure sine signals are $A_1 = 0.092 \pm 0.025$, $A_2 = 1.879 \pm 0.16$ and $A_3 = 1.82 \pm 0.15$. The periods and the amplitudes of the two strongest P_2 and P_3 signals are correct because they are the same as in Model 3 simulation (Table 3, column 1). Therefore, the $g_{3,1,0}$ forecast appears nearly as good as the $g_{2,1,0}$ forecast because the residuals ϵ'_i are close to the offset level of $\epsilon'_i = -1$ (blue open dots), and these residuals also stay inside the red dotted $\pm 3\sigma$ model error limits (Figure 8c). The amplitude $A_1 = 0.092$ of the third $P_1 = 0.0643$ signal is very low. Due this weak "unreal" P_1 signal, the $g_{3,1,0}$ model has a larger $z_{\rm Fore} = 0.91$ value than the correct $g_{2,1,0}$ model (Table 10). Finally, we note that the three signal model $g_{3,1,0}$ is not unstable ("UM") although the simulated time series contains only two signals.

For the forecasting data, the extreme Fisher-test critical levels $Q_{\rm F} < 10^{-13}$ confirm that the two signal $g_{2,1,0}$ ($\chi^2 = 22.946$) and the three signal $g_{3,1,0}$ ($\chi^2 = 19.790$) models are certainly better than the one signal $g_{1,1,0}$ ($\chi^2 = 206.3$) model. The forecasting data parameters $n = 40$, $\eta_1 = 7$, $\eta_2 = 10$, $\chi_1 = 22.946$ and $\chi_2 = 19.790$

for the simple $g_{2,1,0}$ model and the complex $g_{3,1,0}$ model give the Fisher-test critical level $Q_F = 0.22 > \gamma_F = 0.001$. The $g_{2,1,0}$ model beats the $g_{3,1,0}$ model because the H_0 hypothesis is not rejected. Hence, the Fisher-test confirms that the $g_{2,1,0}$ model is the best model for the forecasting data. The best DCM model (Fisher-test, $g_{2,1,0}$) is also the correct DCM model (Forecast-test, $g_{2,1,0}$). The double-check works for this particular Model 3 simulated time series!

DFT cannot detect the correct frequencies for the forecasting data because the simulated frequencies are "too close" (Equation 37). Therefore, DFT forecast cannot succeed.

The analyses of all seven complex time series indicate that the relative accuracy of amplitude and period estimates are lower than the relative accuracy of signal minimum and maximum epoch estimates (Tables 1 - 7). This statistical effect is the same when DCM is applied to the time series of any arbitrary phenomenon. This effect would, for example, explain why our solar cycle amplitude forecasts are less accurate than our solar cycle minimum and maximum epoch forecasts (Jetsu 2025). DCM detects the correct simulated period values, but DFT detects less accurate period values, which are not always correct. The leakage of DFT spectral power (Kay & Marple 1981; Ghaderpour et al. 2021) would explain why correct signal periods have not been detected earlier from the sunspot data.

DCM analysis results can be double-checked. We use the Fisher-test to identify the best DCM model from all tested DCM models (Equation 48). The correct model may not be among the tested models. Hence, the best model is not necessarily the correct one. This best model must be correct if it passes the Forecast-test (Equation 53). This Fisher- and Forecast-test combination double-checks DCM analysis results. The perfect result is that the best model is the correct model. Therefore, the Forecast-test also prevents overfitting.

4. DISCUSSION

DCM proceeds through two stages. It computes the R (Equation 14) or χ^2 (Equation 15) values for a massive number of LS fits (Equation 50). The Gauß-Markov theorem ensures that DCM model having the lowest R or χ^2 is inevitably *always* found, and the result for the non-linear iteration is unique (Equation 22). This first stage devours CPU. The second fast stage, the Fisher-test, then reveals the best DCM model alternative (Equation 48). Practical time series analysis applications, like digital signal processing, require fast computational algorithms (Kay & Marple 1981). Long computation time is the main AL constraining DCM. For example, Jetsu (2025) computed about 4.5 million LS fits to search for four signals in the monthly sunspot data ($n = 3287, K_1 = 4, K_2 = 1, K_3 = 0, n_L = 100, n_S = 30, n_B = 20$). This required several months of CPU. The parallel Python code computations took a few days. Nevertheless, the CPU used is beside the point if a well-posed computational solution can be found to any challenging scientific ill-posed problem.

DCM computation time AL constraint is amply compensated in real observations because there is no need to wait for the repetition of the signal(-s). The WD-effect is the spearhead of DCM because the sample window (ΔT) can be infinitesimally short. Fast accurate observations are the best approach. If the noise ($\sigma \equiv$ SN) cannot be eliminated it is always possible to increase the sample size (n).

It is time to summarise why the fifteen ALs (Section 1) of other frequency-domain parametric time series analysis methods do not constrain DCM.

1. Data errors (level of noise) are unknown.

   DCM can solve both alternatives: errors σ_i known or unknown (Equations 16, 17, 46 and 47). All DCM model parameter solutions in Tables 1-7 converge to the correct values when the simulated data n and/or SN increase (σ_i decrease). Due to this WD-effect, sufficient noise reduction always leads to the correct model detection.

2. Data error information is not utilised.

   DCM utilises this information. If the errors σ_i are known, DCM performs weighted LS fits which minimise the model χ^2 for every tested frequency combination (Equation 15). This gives DCM test statistic z (Equation 17). The Fisher-test utilises the χ^2 values of all alternative tested DCM models to identify the best DCM model (Equation 47).

3. Data must be evenly spaced.

   DCM performance is independent of data spacing. Even or uneven spacing is irrelevant for all n_{Lin} LS fits (Equation 50). However, we do admit that long gaps can mislead even these LS fits.

4. Model parameter errors are unknown.

   DCM gives error estimates for the model parameters of Equations 23-29. For non-linear

models, such as DCM model, the analytical solution for the model parameter errors is a highly complex effort (Furlan & Mortarino 2020). DCM solves these error estimates using the computational statistical bootstrap technique (Equation 31).

5. Model and forecast errors are unknown.

   The computational bootstrap technique (Equation 31) gives n_B estimates for β_J, $t_{\mathrm{mid,J}}$ and ΔT_J, where $J = 1, 2, ...n_B$. These estimates give n_B values for $g(t)$ at any time t inside and outside the sample window ΔT. The standard deviation of these n_B values gives the error limits $\sigma_{g(t)}$ in Equation 30. The n_B estimates for β_J, $t_{\mathrm{mid,J}}$ and ΔT_J also give the error limits for the functions $h(t), h_i(t)$ and $p(t)$ in DCM model $g(t)$.

6. Sample window is shorter than signal period(s).

   DCM solves this problem. The sample window is shorter than the period(s) in Model 1, 2, 5 and 7 time series simulations. Regardless of this, all model parameter estimates converge to correct values when n and/or SN increase (Tables 1, 2, 5 and 7). Due to this WD-effect, the sample window ΔT value is irrelevant for DCM. It can perceive the future and the past much "earlier" (from shorter ΔT) than has been previously thought, like in Equations 35 and 37.

7. Presence and shape of trend are unknown.

   DCM can test any K_1, K_2 and K_3 model combination. The Fisher-test reveals the combination of the best DCM model. The Forecast-test can double-check this result. These tests give the correct K_3 trend order. Due to the WD-effect, the polynomial trend coefficients M_k converge to correct values in all simulated time series (Sections 3.1-3.7). The trend is absent (stationary time series) if $K_3 = -1 \equiv p(t) = 0$ or $K_3 = 0 \equiv p(t) = M_0 =$ constant (Equations 4 and 5).

8. Sample window causes leakage.

   Leakage does not constrain DCM because the sample window ΔT has no effect on the LS fits (Equation 50: n_{Lin}). All frequencies converge to exactly correct values when the n and/or SN values of simulated time series increase (Tables 1-7). This confirms the absence of leakage. Due to the WD-effect, the performance of DCM does not depend on the sample window ΔT.

9. Leakage weakens frequency resolution.

   There is no leakage because the sample window is irrelevant (AL8). DCM frequency resolution is limited only by the sample size n and the data accuracy σ. The WD-effect ensures that all frequency detections for the simulated time series converge to exactly correct values (Tables 1-7).

10. Signal shapes are not pure sines.

    DCM always finds the correct K_1, K_2 and K_3 values (AL7). The K_2 value determines the signal shape. The $K_2 = 1$ pure sine shape is the simplest. Higher K_2 values allow DCM modelling of more complex shapes. We demonstrate the $K_2 = 2$ double wave signal detections in Sections 3.6 and 3.7.

11. Number of signals is unknown.

    The correct K_1, K_2 and K_3 values can always be found (AL7). The K_1 value determines the number of signals. For all simulated time series, the detected signal frequencies and amplitudes converge to exactly correct values (Sections 3.1-3.7). Furthermore, Jetsu (2020) showed that DCM models having too few or too many signals are often unstable ("UM").

12. Correct model alternative is unknown.

    The values of K_1, K_2 and K_3 determine the correct model (AL7). The $(n - \eta_2 - 1)/(\eta_2 - \eta_1)$ penalty terms prevent overfitting when the Fisher-test is used to identify the best DCM model from all tested DCM models (Equations 46 and 47). This best model is also the correct model if it passes the Forecast-test (Equation 53). The WD-effect ensures that better data can inevitably reveal the correct model.

13. Signal significances are unknown.

    The Fisher-test gives the Q_{F} critical levels (signal significances) for all detected signals

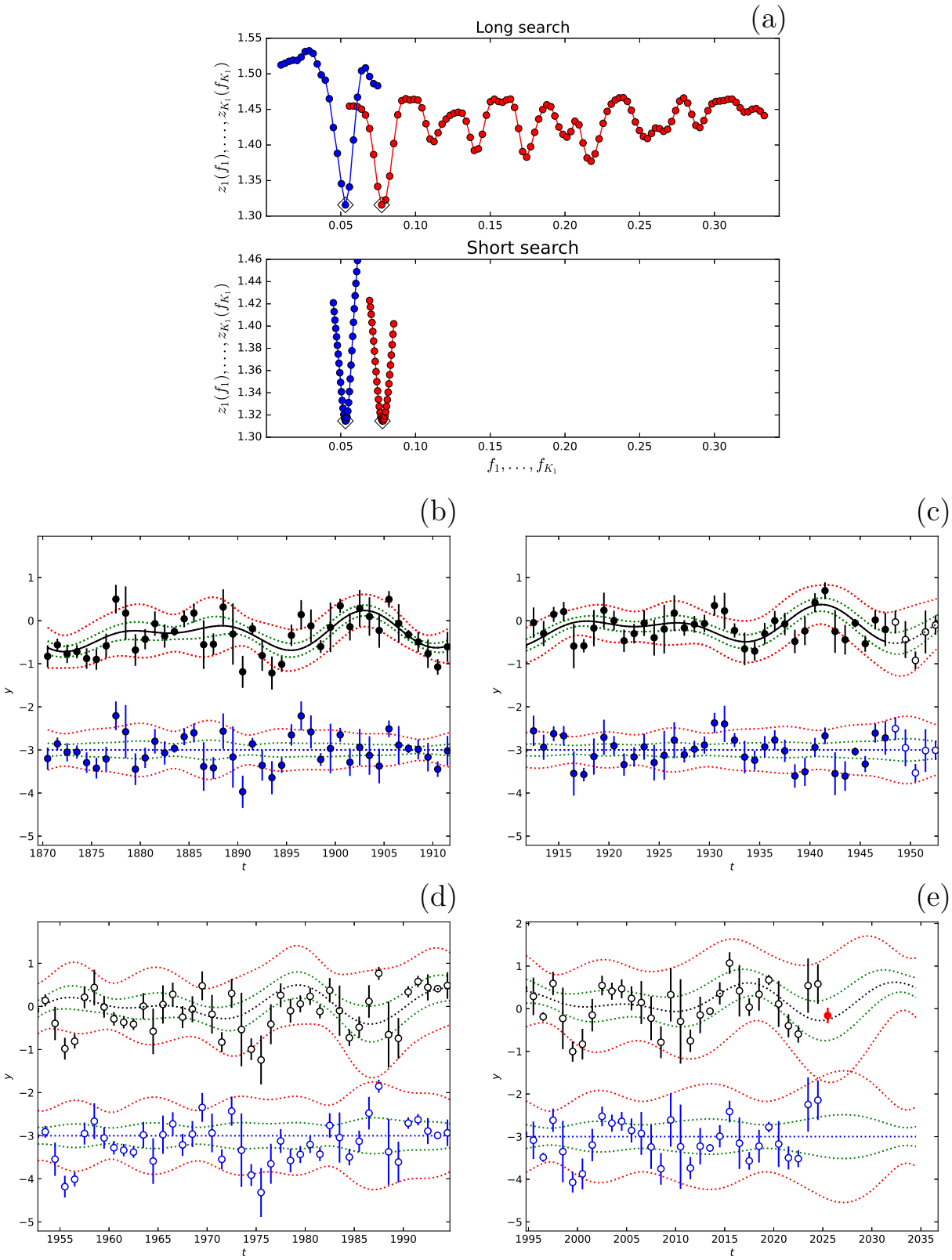

Figure 9. First half forecast (Table S4: $\mathcal{M}$=2): (a) Periodogram (Equation 21). (b-e) Forecasting data (closed black circles) and forecasted data (open black circles). Red closed circle denotes 2025 yearly mean which is discussed later in Section 5.2.3. Green and red dottted lines denote one and three sigma model $g(t)$ errors. Residuals (closed and open blue circles) are offset to level of -3. Units are x-axis: $[t_i]$=y, y-axis: $[y_i] = {}^{\circ}$C and error bars = $[1 \times \sigma_i] = {}^{\circ}$C.

(Sections 3.9 and 3.10). This Q_F is the probability of falsely rejecting the H_0 hypothesis when it is in fact true. Thus, the Q_F value represents the probability of false signal detection. This detection is absolutely certain in the $Q_F < 10^{-16}$ cases. Our preassigned significance level for signal detection is $\gamma = 0.001$ (Equation 48).

14. Model solution is ill-posed.

    The *analytical* solution for the non-linear DCM model is ill-posed (Equation 1). We show that there exists a well-posed *computational* solution (Section 3.11). The WD-effect ensures that this solution can always be found by increasing the sample size (n) and/or the signal to noise ratio (SN), regardless of the time series complexity (K_1, K_2 and K_3 combination). The main DCM constraint is that the solutions for more complex models require more computation time when the number of signals increases (Equation 50).

15. Complex non-linear model forecasts fail.

    Our computational solution for the non-linear DCM model is well-posed (Section 3.11). The forecast of this well-posed solution can be double-checked (Section 3.12). First, the Fisher-test identifies the best DCM model from all tested DCM models (Section 3.9). Then, the Forecast-test reveals if this best DCM model is also the correct one (Equation 53). Again, the WD-effect ensures that the correct model and forecast can be found for any complex time series (K_1, K_2 and K_3 combination) when n and/or SN increase. We conclude that DCM rises to meet the challenge of "forecasting the evolution of complex systems" (Cheng et al. 2015).

DCM turns things upside down. There would be no need for time series analysis, if the correct frequencies were already *known*. The tested frequencies are already *known*. DCM does not search for *unknown* frequencies, it just tests *known* frequencies. DCM tests all possible frequency combinations. The LS fit for every frequency combination gives unique R and χ^2 values. The best frequency combination minimises R or χ^2. This gives the unique initial free parameter β values for the non-linear iteration. The frequencies are never *unknown* in this process. The Gauss-Markov theorem, the well-posed computational model solution and the revolutionary WD-effect make absolutely sure that DCM succeeds for any number of signals (K_1), all signal shapes (K_2) and every polynomial trend (K_3). The Fisher-test and the Forecast-test can double-check that DCM model solution is correct. DCM analysis of intensive, large and accurate time series can see through time: a glimpse of the future and the past.

Our DCM is a remarkable method because it outperforms DFT. Gauß (1777–1855) developed an early form of the Fast Fourier Transform (FFT) algorithm for astronomical purposes, but did he not publish it. Perhaps our DCM should be re-named the Slow Gauß Transform (SGT).

5. VALIDATION ON COMPLEX TERRESTRIAL TIME SERIES: EL NIÑO-SOUTHERN OSCILLATION

We subject DCM to a rigorous stress test to evaluate its performance limits. DCM is applied to the Niño 4 index time series which measures the sea surface temperature anomalies in the central-western equatorial Pacific Ocean. An El Niño event threshold is that the Niño 4 index exceeds 0.5 °C (Hanley et al. 2003; Bunge & Clarke 2009; Ren & Jin 2011). The mainstream climatological models treat El Niño as a non-linear, chaotic phenomenon. The forecasting of El Niño events is challenging and even the best forecasts fail after 18 months (Ludescher et al. 2013; Cai et al. 2014; Timmermann et al. 2018; Liang et al. 2021; Liu et al. 2023b; Hu et al. 2024; Thirumalai et al. 2024; Lu et al. 2025).

5.1. *Data*

We apply DCM to the Niño 4 index from The Seasonal Mean Nino 4 HadISST1.1 (NOAA PSL) sample. These NOAA[4] data were retrieved on January 29th, 2026. For the year 2025, four monthly mean values from September to December were missing. Therefore, we analyse $n = 155$ yearly mean values and $n = 1860$ monthly mean values between January, 1870 and December, 2024. All analysed data samples are summarised in Table S1.

5.2. *DCM analysis results*

5.2.1. *Trend in all yearly mean data*

For all yearly mean Niño 4 index data, we detect significant signals only from the weighted values (Table S1: `Clong.dat`). Therefore, the correct K_3 order for the polynomial trend $p(t)$ is solved from these weighted all yearly mean data. The weighted DCM analysis results for all yearly mean data are given in Table S2. For one,

[4] https://psl.noaa.gov/data/timeseries/month/data/nino4.long.anom.dat

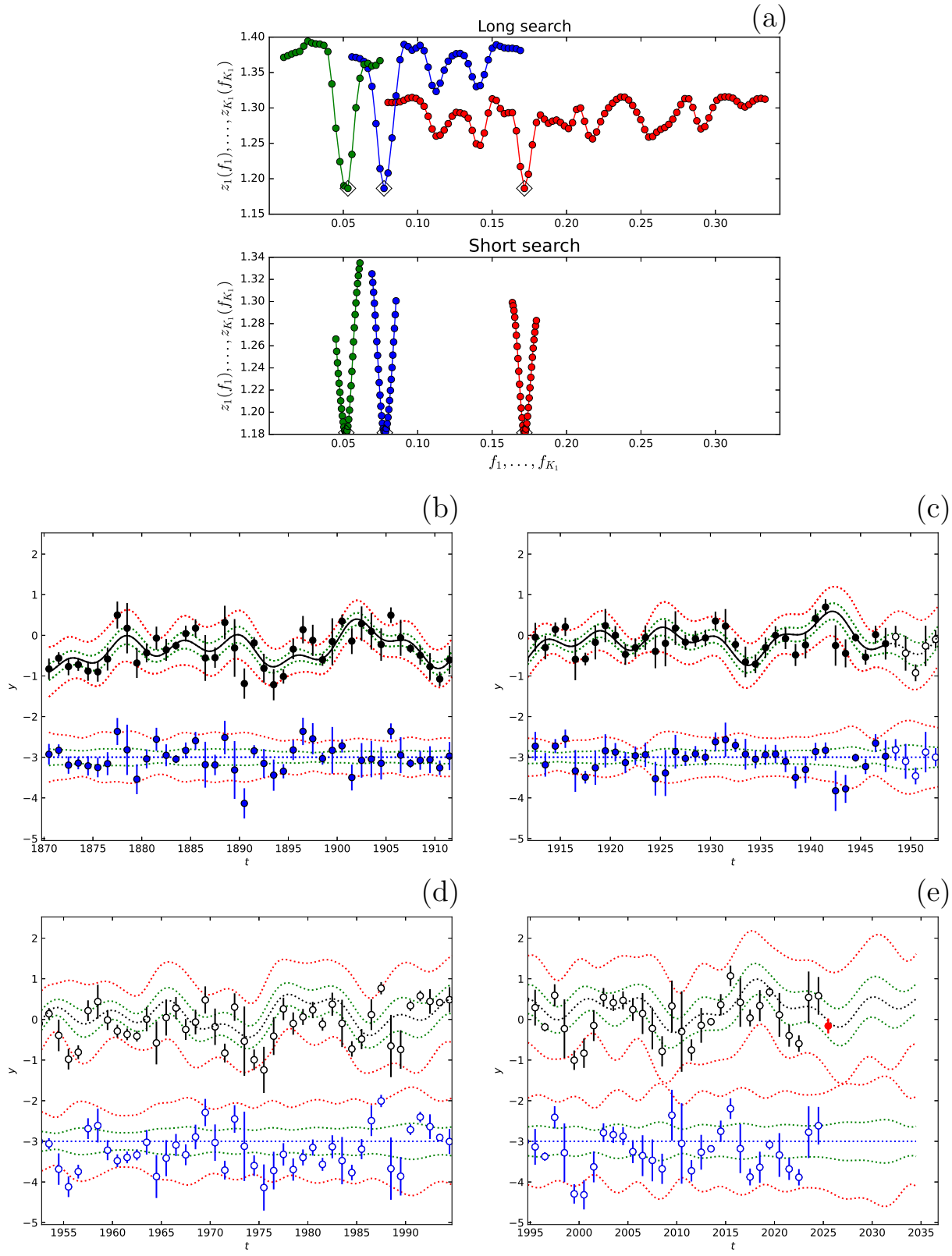

Figure 10. First half forecast (Table S4: $\mathcal{M}$=3). Otherwise as in Figure 9.

two, three, and four signal models, the best alternative is always the linear trend $K_3 = 1$. We use this $K_3 = 1$ trend in all our subsequent DCM analysis of Niño 4 index data. For all yearly mean data, the positive linear temperature increase trend for the best three signal model $\mathcal{M}$=3 (Table S2) is $2M_1 = 0.56 \pm 0.20$ °C during $\Delta T = 154$ years (Equation 5). This trend is approximately the same in all models (Table S2: $\mathcal{M}$=1-4). This positive linear trend pattern confirms the climate change-induced global warming.

5.2.2. *Forecast of first half of all yearly mean data*

We use the first half of all yearly mean Niño 4 index data ($n = 78$) to forecast the second half ($n = 77$). The weighted DCM analysis results for the first half are given in Table S4. DCM detects many signals. The sum of these interfering signals is hereafter called the "Big Wave". The periods of the three strongest "Big Wave" signals are $P_1 = 5.580 \pm 0.085$, $P_2 = 12.82 \pm 0.40$ and $P_3 = 19.30 \pm 0.83$ years (Table S4, $\mathcal{M}$=3). For the pure white noise simple model alternative, the first signal significance is $Q_F = 6.2 \times 10^{-5}$ (Table S3, $\mathcal{M}$=1). The one, two and three signal significances are $Q_F = 0.0016$, 0.0025 and 0.013 (Table S4, $\mathcal{M}$=1-3). For the pre-assigned significance level $\gamma = 0.005$, the $\mathcal{M}$=2 model is the best. The $\mathcal{M}$=2 and $\mathcal{M}$=3 model periodograms show clearly defined, very sharp minima (Figures 9a and 10a). The simulations of Models 1-7 (Sections 3.1-3.7) have already confirmed that only real signals cause such clear periodogram minima. The two and three signal model forecasts, dotted black lines, are shown in Figures 9b-e and 10b-e. Both forecasts can reproduce the changes of forecasted yearly Niño 4 index means, the open black circles, during the second half. For the second half, the first half forecast χ^2 values for the one, two, three and four signal models are 421, 336, 370 and 428, respectively. Hence, the two signal $\mathcal{M}$=2 model gives the best forecast for the second half of weighted yearly mean data. The red closed circles denoting the observed 2025 yearly mean (Figures 9e and 10e) are discussed in greater detail in the next Section 5.2.3.

5.2.3. *All weighted yearly mean data*

DCM analysis results for all weighted yearly mean Niño 4 index data are given in Table S5. We detect the same "Big Wave" periods that are already detected from the first half of weighted yearly mean data. The three signal $\mathcal{M}$=3 model is the best. The $\mathcal{M}$=3 model DCM periodogram shows three clearly defined, sharp minima (Figure 11a). The "Big Wave" signal periods are $P_1 = 5.662 \pm 0.077$, $P_2 = 12.78 \pm 0.12$ and $P_3 = 21.3 \pm 1.5$ years. The peak to peak amplitudes are $A_1 = 0.40 \pm 0.13$ °C, $A_2 = 0.47 \pm 0.10$ °C and $A_3 = 0.37 \pm 0.12$ °C. We show this $\mathcal{M}$=3 model in Figures 11b-e. The Fisher-test comparison between the simple pure white noise model and the complex one signal model gives signal significance $Q_F = 3.9 \times 10^{-13}$ (Table S6, $\mathcal{M}$=1). This Q_F value represents the probability that the strongest signal, the first detected 12.65 years signal of model $\mathcal{M}$=1, is not real. All other pure polynomial model $g(t) = p(t)$ alternatives are also rejected (Table S6, $\mathcal{M} \geq 2$). The second and third signal significances are $Q_F = 2.7 \times 10^{-7}$ and $Q_F = 0.00038$ (Table S5: $\mathcal{M}$=2-3). We conclude that the detections of all these three "Big Wave" signals are significant.

The results for the unstable four signal model (Table S5, $\mathcal{M}$=4, UM) are shown in Figures 12a-e. The $\mathcal{M}$=3 and $\mathcal{M}$=4 model alternatives for all yearly mean weighted data give essentially the same "Big Wave" forecast values between 2025 and 2036 (Table 11).

Table 11. "Big Wave" yearly mean forecasts. (1) Year. (2-3) Figure 11e: open circles. (4-5) Figure 12e: open circles.

	Figure 11e		Figure 12e	
(1)	(2)	(3)	(4)	(5)
t	y	$\pm 1\sigma$	y	$\pm 1\sigma$
(y)	(°C)	(°C)	(°C)	(°C)
2025.5	0.05	0.25	0.30	0.22
2026.5	0.30	0.21	0.42	0.21
2027.5	0.36	0.20	0.26	0.23
2028.5	0.31	0.17	0.16	0.27
2029.5	0.35	0.17	0.37	0.28
2030.5	0.56	0.20	0.74	0.28
2031.5	0.74	0.21	0.89	0.31
2032.5	0.70	0.31	0.61	0.30
2033.5	0.40	0.37	0.14	0.31
2034.5	0.06	0.32	-0.14	0.33
2035.5	-0.08	0.25	-0.08	0.34
2036.5	0.01	0.20	0.11	0.35

The *old* data from NOAA database were retrieved on January 29th, 2026. The four last monthly mean values were missing for the year 2025. Today, May 2nd, 2026, we have retrieved those missing values. The *old* and *new* data after December 2024 are given in Table S8.

The data gap after December 2024 is a fortunate coincidence for us. The *new* monthly mean values in Table S8 (column 3) give the *new* 2025 yearly mean value $y_i \pm 1\sigma_i = -0.16 \pm 0.18$ °C. The red closed circle now highlights this *observed new* 2025 yearly value in Figures 9e, 10e, 11e and 12e. All four figures confirm that the DCM forecast for this *new* 2025 yearly mean value succeeds! For first half of data between 1870 and 1947, both DCM models $\mathcal{M}$=2 and $\mathcal{M}$=3 can, after a gap of 78 years, forecast the year 2025 *observed new* value

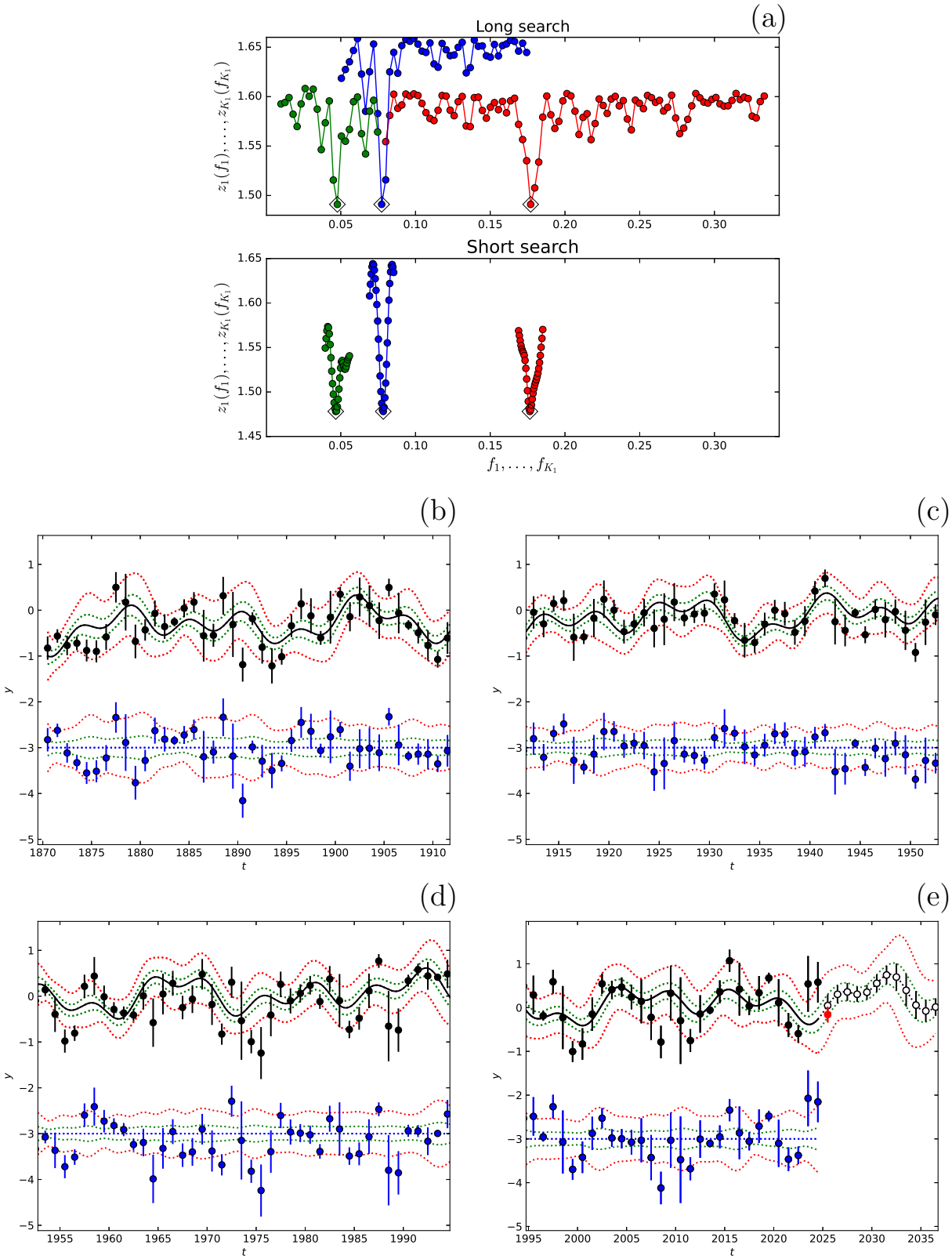

Figure 11. "Big Wave" for all weighted yearly mean data model $\mathcal{M}$=3 in Table S5. Red circle is added *new observed* 2025 year value. Otherwise as in Figure 9.

within $\pm 1\sigma$ (Figures 9e and 10e). For the best $\mathcal{M}$=3 model of all yearly means, the "Big Wave" forecast for the year 2025 (black open circle) and the *observed new* 2025 value (red closed circle) agree within $\pm 1\sigma$ (Figure 11e). Even the unstable $\mathcal{M}$=4 model for all yearly means gives an accurate 2025 forecast (Figure 12e). In short, the *observed new* 2025 value not only confirms both all data forecasts, but it also confirms both first half data forecasts. All these forecasts can, of course, be dismissed just as lucky guesses. Who wants to take the risk that our severe "Big Wave" forecasts for the years 2030, 2031 and 2032 are also "lucky" (Table 11)?

Both DCM models $\mathcal{M}$=2 and $\mathcal{M}$=3 for the first half of yearly mean $n = 78$ sample give excellent forecasts for the second half (Figures 9b-e and 10b-e). For the first six years, the observed second half data $\pm 1\sigma_i$ errors match the first half forecast $\pm 1\sigma_{g(t)}$ errors. Therefore, our forecasts for 2025 - 2036 should be even more accurate because DCM can utilise a two times larger $n = 155$ sample of all weighted yearly means (Table 11, Figures 11e and 12e).

The three signal model is the best one for all weighted yearly mean data (Table S5, $\mathcal{M}$=3). This best forecasting model is computed for the $n = 155$ yearly means (Figure 11e, black closed circles). We can check our real-time yearly mean forecast against only one new yearly mean value, the year 2025 (Figure 11e, red closed circle). However, every monthly mean t_i and y_i value after December 2024 can be used to compute the yearly sliding mean values $t_{12,i}$, $y_{12,1}$ and $\sigma_{12,i}$ for past 12 months. To be precise, the yearly sliding means are computed from each monthly mean t_i and y_i, and the eleven earlier monthly means. The observed yearly sliding means are given in Table 12 and displayed in Figure 13 (open red small circles). These yearly sliding mean values confirm that the $\mathcal{M}$=3 model (Table S5) real-time forecast after December 2024 has become true.

While there is no fixed numerical Niño 4 index limit for El Niño events, the general academic standard is +0.5 °C (Hanley et al. 2003; Bunge & Clarke 2009; Ren & Jin 2011). The Niño 4 index variance is much lower than the Niño 3.4 index variance. Consequently, a lower Niño 4 index anomaly value is statistically more significant. We show the $\mathcal{M}$=3 model (Table S5) "Big Wave" forecast for the next decade in Figure 13. The forecasted $\pm 1\sigma$ yearly mean values between the years 2026 and 2029 rise close to the El Niño limit +0.5 °C (continuous horizontal black line). We forecast a prolonged three year El Niño period, where the yearly mean Niño 4 index values will exceed this +0.5 °C limit (Table 11, 2030-2032). A four year cooling trend should begin in the year 2033.

For the sake of consistency, we also give DCM results for all non-weighted yearly Niño 4 index data (Table S7). We detect the signatures of the same "Big Wave" signals, but the signal significances $Q_{\rm F}$ are low. This confirms that the yearly data error information is crucial for weighted DCM analysis.

5.2.4. *All monthly means*

Here, we analyse $n = 1860$ monthly mean Niño 4 index values between January, 1870 and December, 2024. The computation of twelve month means may have prevented the detection of shorter periods from yearly mean data. We allow shorter period detections by lowering $P_{\rm min} = 3$ years to 2 years. The value $P_{\rm max} = 100$ years remains the same. The four signal DCM model is the best one (Table S9, $\mathcal{M}$=4). The signal periods are $P_1 = 3.6436 \pm 0.0046$, $P_2 = 4.7477 \pm 0.0080$, $P_3 = 5.647 \pm 0.011$, and $P_4 = 12.704 \pm 0.063$ years. The 12.7 years signal has the highest peak to peak amplitude $A_4 = 0.392 \pm 0.038$ °C.

Four prominent, sharp periodogram minima (Figure 14a) indicate the presence of true periodic signals. The first signal significance is $Q_{\rm F} < 10^{-16}$ (Table S10, $\mathcal{M}$=1-4). The second, third and fourth signal significances are also $Q_{\rm F} < 10^{-16}$ (Table S9). All four signal detections are absolutely certain. DCM search for the fifth signal from the $\mathcal{M}$=4 model residuals gives $R = 529$. This confirms that these monthly mean data contain *only four* signals because the $\mathcal{M}$=4 model has $R = 475$.

Table 12. Yearly sliding means after December 2024 (Figure 13: small red open circles). (1-2) Monthly t_i and y_i. (3-5) Sliding means for past twelve months $t_{12,i}$, $y_{12,i}$ and $\sigma_{12,i}$.

(1) t_i (y)	(2) y_i (°C)	(3) $t_{12,i}$ (y)	(4) $y_{12,i}$ (°C)	(5) $\sigma_{12,i}$ (°C)
2025.04	-0.50	2024.58	0.42	0.48
2025.12	-0.36	2024.66	0.29	0.46
2025.21	-0.13	2024.75	0.20	0.42
2025.29	-0.09	2024.83	0.12	0.38
2025.38	0.02	2024.91	0.06	0.34
2025.42	0.14	2025.00	0.03	0.31
2025.54	0.05	2025.08	-0.01	0.27
2025.62	0.06	2025.16	-0.04	0.24
2025.71	-0.12	2025.25	-0.07	0.22
2025.79	-0.32	2025.33	-0.11	0.22
2025.88	-0.37	2025.41	-0.16	0.20
2025.96	-0.19	2025.50	-0.15	0.19
2026.04	-0.15	2025.58	-0.12	0.16
2026.12	0.12	2025.66	-0.08	0.16
2026.21	0.35	2025.75	-0.04	0.20
2026.29	0.30	2025.83	-0.01	0.22

We show the best DCM model for all monthly mean data in Figures 14b-e. The consecutive model residuals (blue closed circles) show strong residual correlation which is typical for climatological time series (Zwiers &

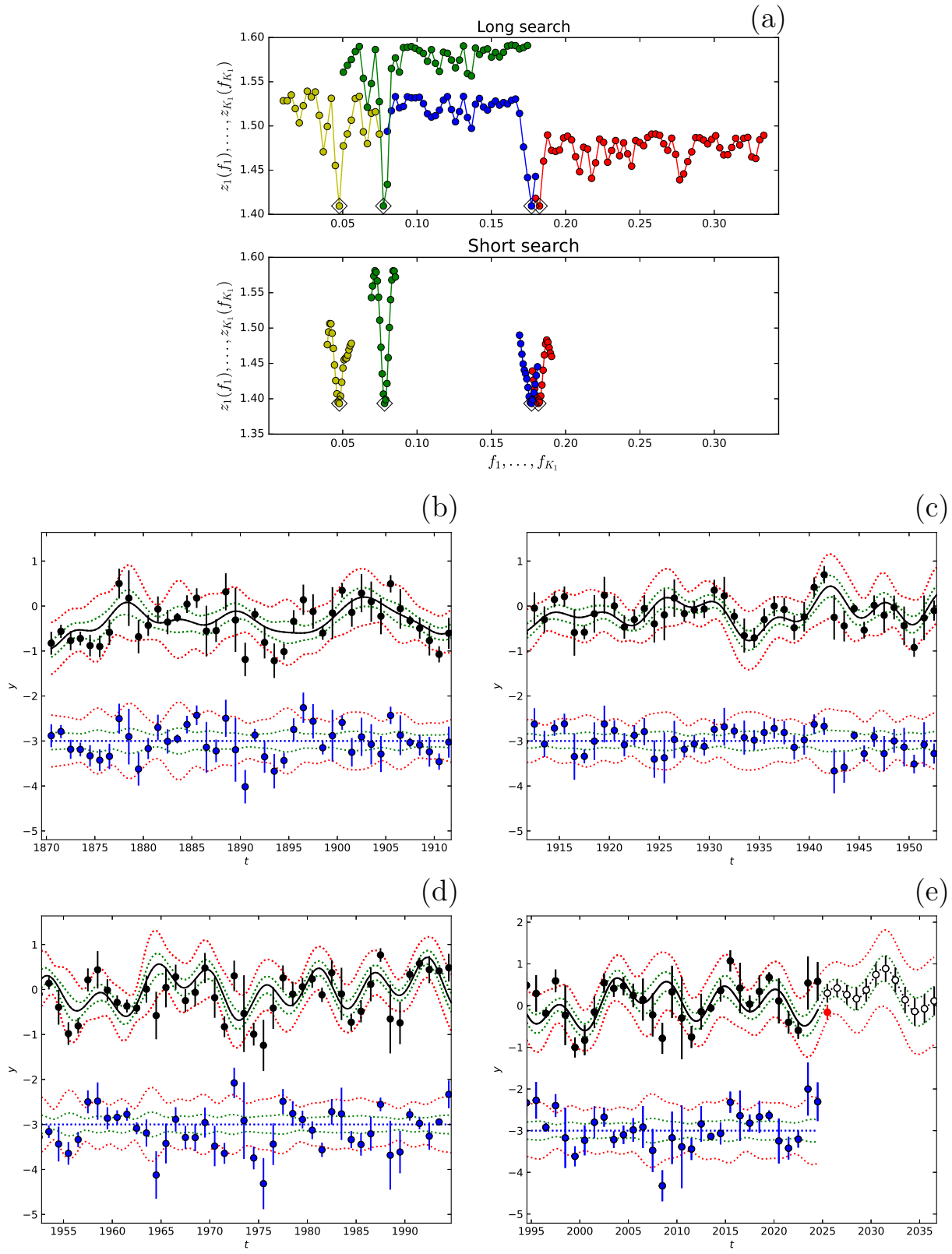

Figure 12. "Big Wave" for all weighted yearly mean data model $\mathcal{M}$=4 in Table S5. Otherwise as in Figure 10.

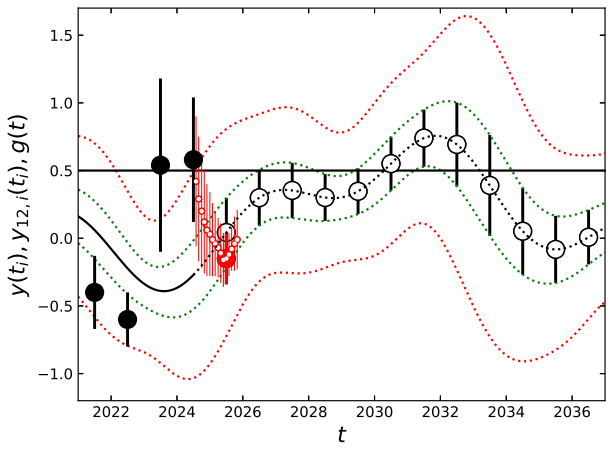

Figure 13. Real-time forecast validation for $\mathcal{M}$=3 model (Table S5). Model curves, observed large closed black circles and forecasted large open white circles are as in Figure 11e. Small open red circles denote $y_{i,12}$ sliding means (Table 12). Observed 2025 yearly mean (red large closed circle) falls within $\pm 1\sigma$ confidence interval of deterministic forecast (large open black circle) generated from data before December 2024. Continuous horizontal black line shows El Niño +0.5 °C limit. Note that our real-time model also successfully captured neutral period in early 2026 and subsequent warming at spring 2026 (Table S8: January-April, 2026). This demonstrates that real-time Niño 4 index forecast is currently synchronised with "Big Wave" periodicities. All error bars are $\pm 1\sigma$. Units are [y] (x-axis) and [°C] (y-axis)

Von Storch 1995; Guemas et al. 2014). Due to this residual correlation, the bootstrap procedure gives too small model error $\sigma_{g(t)}$ estimates (Equation 30). For this reason, we give only the next decade 2025-2036 monthly mean forecast values, but not their error estimates (Table S11). The monthly mean values after December 2024 are not analysed, but they are displayed in Figure 14e (red closed circles).

The correlating monthly mean model residuals (blue closed dots) consistently show how rapidly and strongly the $y(t_i)$ values react to the $g(t_i)$ peaks and valleys (Figures 14b-e). The extrema of $y(t_i)$ and $g(t_i)$ coincide. This is confirmed by the extreme significance $p = 9.3 \times 10^{-109}$ for their linear correlation (Figure 15). However, the scatter of monthly $y(t_i)$ values at any $g(t_i)$ value is about ±1 °C. This means that DCM cannot forecast the monthly mean Niño 4 index values.

5.2.5. *Separate months*

We also analyse separately all months from January to December (Tables S12-S23). The one year gap between consecutive values can weaken the residual correlation problem. We compare the observed and forecasted monthly values after December 2024 (Figures S1-S12: red closed circles and black open circles). All red closed circles are within the forecast $\pm 3\sigma_{g(t)}$ error limits denoted by the red dotted lines. We give the forecasted separate monthly mean values for the next decade (Table S24). Note that Table S24 also gives the forecast error estimates because the residual correlation problem for separate months is weak or totally absent.

5.3. *Solar forcing hypothesis*

DCM analysis of Niño 4 index data reveals that some strictly periodic mysterious "engine" warms and cools the Pacific Ocean. It is highly improbable that some *internal* oceanic "engine" functions with such mechanical regularity. The primordial "engine" *outside* the ocean must be the Sun. We apply DCM to the observed anomalies of the sea surface temperatures $T_{\rm Sea}(t)$. If the "engine" is the Sun, the $T_{\rm Sea}(t)$ changes

$$y(t) = T_{\rm Sea}(t) = T_{\rm Sol}(t) + T_{\rm Glo}(t), \tag{55}$$

are the sum of solar forcing $T_{\rm Sol}(t)$ and global warming $T_{\rm Glo}(t)$, where the units are °C.

Our model $g(t) = h(t) + p(t)$ is the sum of a periodic $h(t)$ function and an aperiodic $p(t)$ function. If the observed linear polynomial trend $p(t)$ represents the global warming

$$T_{\rm Glo}(t) = p(t), \tag{56}$$

we get

$$y(t) - p(t) = T_{\rm Sea}(t) - p(t) = T_{\rm Sol}(t). \tag{57}$$

In this case,

$$\begin{aligned} T_{\rm Sol}(t) &= y(t) - p(t) \\ &\approx g(t) - p(t) = h(t) \propto F_{\rm Sun}(t) \end{aligned} \tag{58}$$

where $F_{\rm Sun}(t)$ is the sum of solar output flux signals having the units Wm^{-2}. The maximum $h(t)$ solar forcing effect is $A_1 + A_2 + A_3 = 1.2 \pm 0.2$ °C for yearly means (Table S5) and $A_1 + A_2 + A_3 + A_4 = 1.41 \pm 0.06$ °C for monthly means (Table S9). The $\pm 1\sigma$ limits of these two estimates overlap.

We use the relation

$$F_{\rm Sun}(t) \xrightarrow{\rm Cli} T_{\rm Sol}(t) \tag{59}$$

to describe how $F_{\rm Sun}(t)$ triggers the $T_{\rm Sol}(t)$ changes. The symbol "$\xrightarrow{\rm Cli}$" contains all unknown climatological processes that transform Wm^{-2} to °C. For the monthly mean values, the linear correlation between

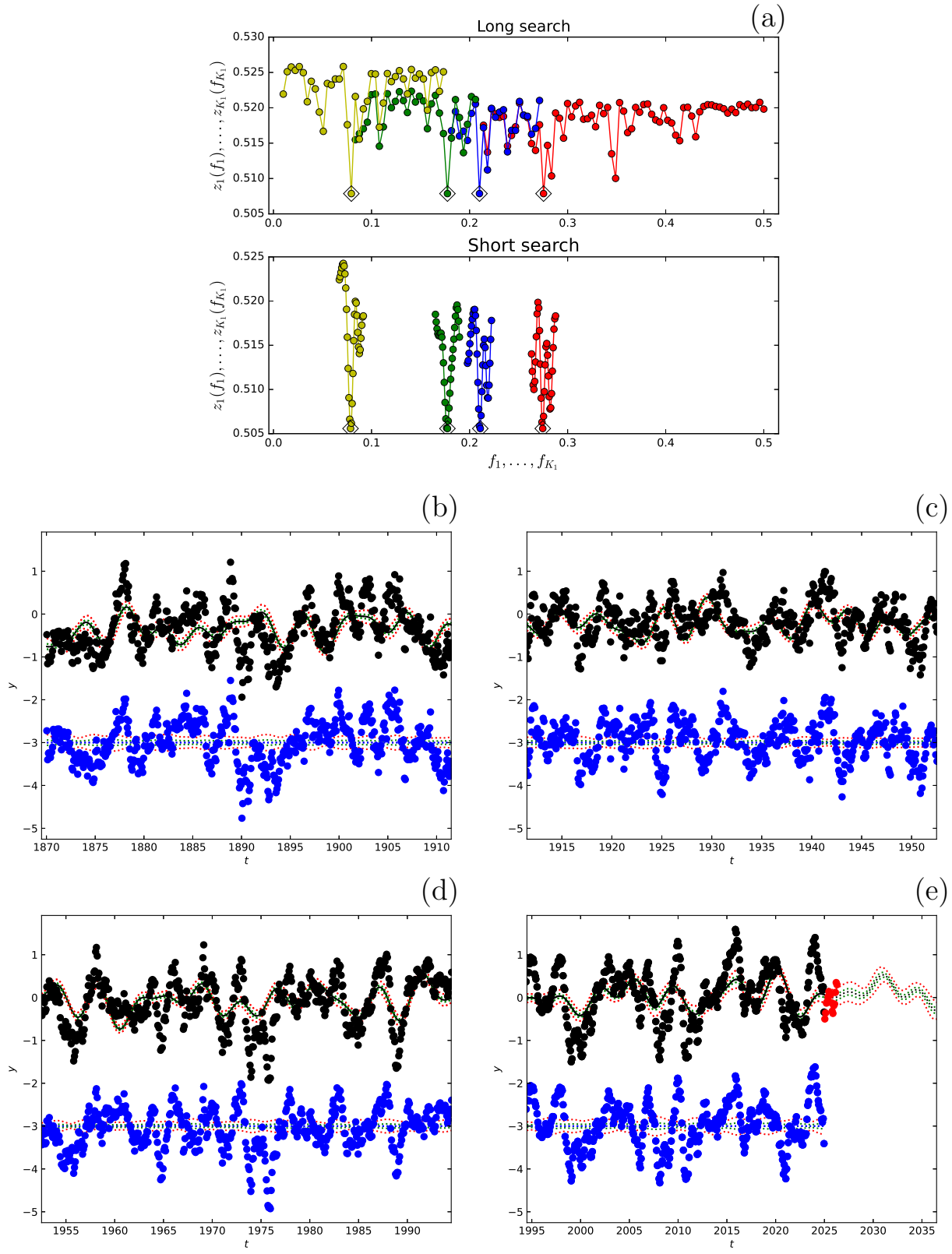

Figure 14. All monthly mean data model $\mathcal{M}$=4 in Table S9. Red dots are new updated monthly mean values from Table S8. Otherwise as in Figure 10.

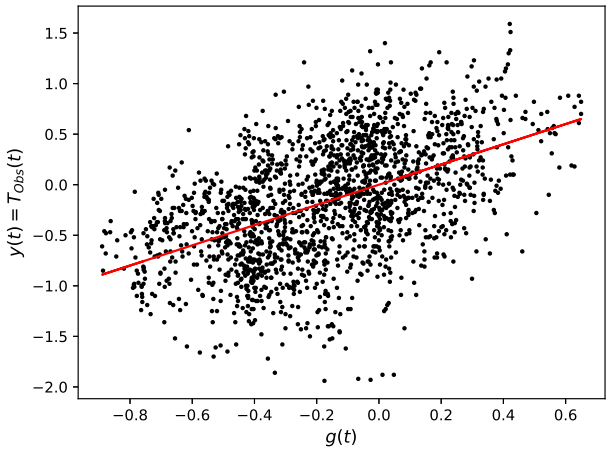

Figure 15. Correlation between monthly mean $g(t_i)$ and $y(t_i)$ values. Linear correlation coefficient $r = 0.501$ significance is $p = 9.3 \times 10^{-119}$. Red line denotes $y(t) = a[g(t) + b$, where $a = 1.001 \pm 0.040$ and $b = 0.000 \pm 0.013$. Units are [°C] (x- and y-axis).

$g(t) - p(t) \propto F_{\rm Sun}(t)$ and $y(t) - p(t) \approx T_{\rm Sol}(t)$ is absolutely certain (Figure 16: $p = 5.7 \times 10^{-92}$), but the ± 1 °C scatter of monthly $y(t_i) - p(t_i)$ values prevents accurate forecasting. For the yearly means, however, forecasting succeeds because the average of twelve month means eliminates the chaotic "$\xrightarrow{\text{Cli}}$" effect in Equation 59.

5.4. *Holy grail: El Niño, the Sun and the Planets*

DCM can model and forecast the "Big Wave" in Niño 4 index data. We detect three signals in all yearly mean data and four signals in all monthly mean data. The two strongest 12.8 and 5.7 year signals are the same. The shorter 3.6 and 4.7 year signals are detected only from the monthly mean data because the computation of twelve month averages eliminates them from yearly mean data. Between January 1870 and December 2024, the first half of the yearly means effectively forecasts the second half (Figures 9b-e and 10b-e: open circles). Both first half DCM models for yearly Niño 4 index means before the year 1948 can accurately forecast the *observed new* 2025 yearly mean (Figures 9e and 10e: closed red circle). The "Big Wave" signals must be very stable because DCM forecasts come true in the year 2025 after a gap of 78 years. In a chaotic system, a 78-year forecast should be mathematically impossible. For the two times larger complete yearly mean sample, only time will reveal if our ten year "Big Wave" forecast after December 2024 succeeds (Table 11, Figures 11e and 12e: open black circles). Our real-time yearly sliding mean DCM forecast after December 2024 has come true (Figure 13). Due to the "$\xrightarrow{\text{Cli}}$" effect (Equation 59), the monthly mean data are too chaotic for accurate forecasting (Figures 15 and 16).

Here, we discuss the hypothesis that *external* solar forcing, rather than *internal* climate dynamics, governs the general pattern of Niño 4 index changes. Some studies have found correlation between the solar cycle and the climate temperature (Friis-Christensen & Lassen 1991; Connolly et al. 2021). The mainstream scientific consensus is that solar activity plays a minor role in global warming (Gray et al. 2010; Kopp & Lean 2011; Lockwood 2012). For the period 1870 - 2024, our DCM estimates for all yearly means give + 0.6 °C for linear global warming trend and ± 0.6 °C for solar forcing superimposed on this trend.

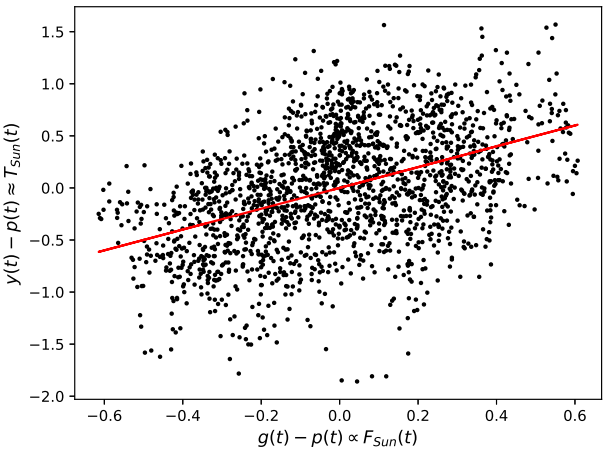

Figure 16. Correlation between monthly $g(t_i) - p(t_i)$ and $y(t_i) - p(t_i)$ values. Linear correlation coefficient $r = 0.447$ significance is $p = 5.7 \times 10^{-92}$. Red line denotes $y(t) - p(t) = a[g(t) - p(t)] + b$, where $a = 0.999 \pm 0.046$ and $b = 0.000 \pm 0.012$. Units are [°C] (x- and y-axis).

Our arguments are transparent.

Argument 1: If the "Big Wave" is a real phenomenon, solar forcing is the only logical cause.

Argument 2: If the solar forcing causes the "Big Wave", this forcing must be strictly periodic.

Argument 3: If the solar forcing is strictly periodic, the solar dynamo must be strictly periodic.

Argument 4: If the solar dynamo is strictly periodic, the planetary tides are the only logical cause.

The main counterarguments are

Counterargument 1: El Niño is an internal climate phenomenon, where the ocean and the atmosphere are dynamically coupled (Bjerknes 1969; Philander 1989).

Counterargument 2: The solar dynamo is stochastic and non-stationary (Usoskin 2017; Charbonneau 2020).

Counterargument 3: Planetary gravitational fields are too weak to sustain a strictly periodic solar dynamo (Kudryavtsev et al. 2022; Charbonneau 2022).

5.4.1. *Argument 1*

DCM detects the same strictly periodic signals from half of the yearly Niño 4 index means, all yearly Niño 4 index means and all monthly Niño 4 index means. The signal signicances are extreme, like $Q_{\rm F} < 10^{-16}$ for all four signals detected from all monthly mean Niño 4 index values. DCM models give accurate forecasts. All these results confirm that the "Big Wave" must be a real phenomenon. If Counterargument 1 were true, such a chaotic system would not cause the "Big Wave". If the ocean or atmosphere cannot cause the "Big Wave", the only remaining alternative is the Sun.

5.4.2. *Argument 2*

If the solar forcing were not strictly periodic, what other alternative phenomenon could cause the observed regular "Big Wave"? If the solar forcing were irregular, the global warming and climate response noise "$\xrightarrow{\rm Cli}$" (Equation 59) could prevent DCM detection of "Big Wave". For example, this forcing must be strictly periodic and stationary because the yearly means between the years 1870 and 1947 can accurately forecast the 2025 yearly mean (Figures 9e and 10e: closed red circle).

5.4.3. *Argument 3*

If the solar dynamo is stochastic and non-stationary (Counterargument 2), the solar forcing can not be strictly periodic. There are dynamo models, where the synchronised response to planetary tides drives a multi-periodic solar dynamo (Klevs et al. 2023; Mouël et al. 2025; Stefani et al. 2025). The solar dynamo may be multi-periodic and stationary because DCM detects the extremely significant, strictly periodic S_{11^y}, S_{10^y} and $S_{11.^y86}$ signals from the sunspot record (Jetsu 2025). The above subscripts refer to the signal periods in years.

5.4.4. *Argument 4*

Counterargument 3 is not true, if the synchronised planetary tides can provide the "clocking" mechanism for the solar dynamo (Klevs et al. 2023; Mouël et al. 2025; Stefani et al. 2025). The unending interference of planetary tides could explain the regular "Big Wave" in the Niño 4 index time series. Except for planets, there is nothing else in the solar system that can cause the regular signals in the period range of the "Big Wave". The planetary tides could also explain the strictly periodic S_{11^y}, S_{10^y} and $S_{11.^y86}$ signals in the sunspot data (Jetsu 2025). It may not be a coincidence that the $S_{11.^y86}$ signal period detected from the sunspot record is equal to Jupiter's orbital period of 11.86 years. The tidally driven solar dynamo could explain both the "Big Wave" in the Pacific Ocean and the strictly periodic signals in the sunspot record (Jetsu 2025). If the solar dynamo were indeed stochastic and non-stationary, our DCM would never detect extremely significant, strictly periodic signals in these two different physical systems.

5.4.5. *Damping and amplifying in solar forcing*

DCM detected the S_{10^y}, S_{11^y} and $S_{11.^y86}$ signals from the sunspot record (Jetsu 2025). These solar forcing signals would amplify or dampen each other in the Pacific Ocean. Our notation $S_{12.^y78}$ refers to the strongest 12.78 ± 0.12 years signal in the yearly mean Niño 4 index data (Table S5, $\mathcal{M}$=3).

The beat frequency of two arbitrary frequencies $f_1 = 1/p_1$ and $f_2 = 1/p_2$ is

$$f_{1,2} = |f_1 - f_2|. \tag{60}$$

One connection between the sunspot signals S_{10^y} ($p_1 = 1/f_1$), S_{11^y} ($p_2 = /f_2$) and $S_{11.^y86}$ ($p_3 = 1/f_3$), and the El Niño signal $S_{12.^y78}$ ($p_4 = 1/f_4$), is outlined in Table S25. If we take the "beat" $f_{2,3}$ between f_2 and f_3, and subtract this beat from the frequency f_3, we get

$$\begin{aligned} 12.87 \equiv 0.0777 &= f_3 - f_{2,3} \\ \approx f_4 = 0.0782 &\equiv 12.78 = p_4. \end{aligned} \tag{61}$$

This relation is well within the $\pm 1\sigma = 0.12$ error limits. The f_1, f_2, f_3 and f_4 frequencies are nearly equidistant (Table S25). The chain is $10 \xrightarrow{-0.009} 11 \xrightarrow{-0.007} 11.86 \xrightarrow{-0.006} 12.78$. The 12.78 period may represent the low frequency caused by interference between the 11 and 11.86 periods. For example, Stefani et al. (2019) discuss this $S_{12.^y78}$ signal in connection to their planetary tidal synchronisation dynamo model.

We show that the strongest $S_{12.^y78}$ signal in Niño 4 index data can be connected to the strongest S_{11^y} signal and the third strongest $S_{11.^y86}$ signal which DCM has detected in the sunspot data (Jetsu 2025). Therefore, it is unnecessary to begin a long discussion about all possible connections between the signals detected in the

Pacific Ocean and the Sun. The main point is that DCM signal detections in these two different physical systems can be connected.

5.4.6. *General comments*

Our *mathematical* model detects regularities that the *climatological* models fail to detect, let alone forecast. Mathematics detects signals that, according to Climatology, should not be there. The counterargument is that simple mathematics cannot explain the complex coupling between the ocean and atmosphere. Our simple physical model is that the solar forcing causes the "Big Wave". The Pacific Ocean is the Earth's largest, and therefore the most robust, thermal reservoir for measuring this forcing in real-time. If the sum of many signals moderates the solar forcing, it is no surprise that we have so far failed to discover this "Big Wave" interference in the sea surface temperature anomalies. Furthermore, the "Big Wave" is buried under the global warming trend and the strong noise "$\xrightarrow{\text{Cli}}$" caused by the climate response (Equation 59). El Niño seems to work like a "clock" because our DCM can model and forecast the "Big Wave". This result may be the "holy grail" of climatology. Our "Big Wave" discovery does not fully solve the El Niño riddle, like the chaotic Niño 4 index monthly means. However, the "Big Wave" does give a good starting point for modelling the climatological "$\xrightarrow{\text{Cli}}$" effects (Equation 59).

The mainstream dynamical models simulate ocean-atmosphere coupling using supercomputers. The "spring predictability barrier" prevents long-term forecasting. Even the best dynamical model El Niño forecasts fail in 1.5 years (Zhao et al. 2024). These dynamical models may be over-complicated because our DCM analysis confirms that short-term weather does not affect long-term trends. Our mathematical DCM yearly mean forecast can connect seven "Big Wave" decades to the next seven decades (Figures 9 and 10). This should be impossible if the climatological chaos takes over within 1.5 years, not to mention the obstacle of breaking through seventy "spring predictability barriers". A reliable El Niño forecast even one year ahead would save trillions of dollars (Hsiang et al. 2011; Callahan & Mankin 2023; Liu et al. 2023a; Xu et al. 2026). Given these massive financial stakes, we urge the scientific community to verify our findings immediately.[5] An ordinary PC can detect the first two "Big Wave" signals in about one minute. The third signal detection takes a few minutes. William of Ockham (1287–1347) stated that if two competing hypotheses make the same predictions, the simplest hypothesis is usually the best one (Occam's razor). This comes true when an ordinary PC gives better El Niño forecasts than multi-million-dollar supercomputers.

Since December 2024, DCM has achieved a critical milestone by forecasting in real-time the 2025 yearly mean within a $\pm 1\sigma$ margin (Figure 13). Our real-time validation also accurately captured the neutral conditions in early 2026 and the current warming trend at the end of spring 2026. All these validity checks confirm DCM's predictive power of strictly periodic signals. These results confirm that DCM can provide deterministic forecasts for complex climate transitions. The next months will show if the Pacific Ocean continues to dance the "Big Wave" tune. That would also finally confirm that the solar cycle is multi-periodic and stationary (Jetsu 2025). It is fortunate for us that the Niño 4 index forecasts can be checked much faster the sunspot cycle forecasts.

The "notorious" Niño 4 index time series is actually considerably simpler than our simulated Model 7 time series discussed in Section 3.7. That time series is pathological because both signal periods are shorter than the data window. Both signals are not pure sines and they are superimposed on a parabolic polynomial trend. This pathology does not hinder the success of the Model 7 time series DCM analysis. Because of this, there is no rational reason to doubt the results of our Niño 4 index DCM analysis. As shown in Figures 9a, 10a, 11a, and 14a, the sharp periodogram minima unmistakably indicate a successful analysis. Official El Niño forecasts from NOAA and similar agencies often cite probabilities in the 10% – 90% range. DCM probability that the four strongest monthly mean data "Big wave" signals are real is larger than $1 - 10^{-16} = 0.9\ 999\ 999\ 999\ 999\ 999$ (Table S9). If our DCM can decode the Earth's most complex non-linear phenomenon (El Niño) with this significance, it is powerful enough to handle complex astrophysical time series. Our El Niño stress test confirms that DCM is a robust, cross-disciplinary tool for detecting non-stationary signals in the presence of extreme variance.

6. CONCLUSIONS

The frequency-domain parametric time series analysis methods methods suffer from many application limitations (Section 1: AL1 - AL15). We demonstrate thoroughly that none of those limitations apply to DCM (Sections 3 - 4). By dividing the data into separate forecasting and forecasted samples, DCM is able to "see through time". The Fisher-test and the Forecast-test can confirm the correct model for the data. Due to the

[5] See declaration "Code and data availability".

WD-effect, the time span of data window is irrelevant because the performance of DCM depends only on the quantity and quality of data.

The mainstream considers the El Niño phenomenon so chaotic and non-linear that even multi-million-dollar supercomputer long-term forecasts are doomed to fail. In our stress test, we apply DCM to the Niño 4 index time series (Section 5). DCM can model and forecast El Niño. For the yearly means, the DCM model for the first seven decades can forecast the next seven decades. Despite the 78-year gap, both DCM models for the 1780 - 1947 yearly means can accurately predict the *observed new* 2025 yearly mean (Figures 9e and 10e: red closed circle). As for all yearly means before December 2024, we perform a real-time DCM forecast validation. This DCM real-time forecast for the new data after December 2024 has come true (Figure 13). The rapid warming of sea surface temperatures observed now, in May 2026, provides the latest striking real-time validation.

We detect the extremely significant and strictly periodic "Big Wave" from the yearly and monthly Niño 4 index data. No such precise "clock" can exist *inside* the Pacific Ocean. Strictly periodic solar forcing is the only possible *external* "clock". The mainstream stochastic and non-stationary solar dynamo (Charbonneau 2020; Usoskin 2023) cannot cause strictly periodic solar cycles. Planetary tides could provide such a precise multi-periodic solar dynamo "clocking" mechanism (Klevs et al. 2023; Mouël et al. 2025; Stefani et al. 2025; Jetsu 2025). By definition, a system cannot be characterized as chaotic (e.g., El Niño) or stochastic (e.g., solar dynamo) if it produces strictly periodic, extremely significant signals that exhibit phase-stability for decades (e.g., red closed circles in Figures 9e and 10e).

At the moment, our solar dynamo and El Niño claims are off-mainstream. However, that scientific narrative is unimportant, for example, to commodity producers who deserve reliable real-time El Niño forecasts. The simple connection between the Sun, the Planets and the "Big Wave" may seem like the holy grail of Climatology. However, these three elements are merely physical phenomena. The mathematical DCM is the real scientific holy grail. DCM has now discovered extremely significant, strictly periodic signals in the sunspot and El Niño data. Both discoveries require that the solar dynamo is stationary and multi-periodic. These claims are now under an objective, real-time "review" by the Pacific Ocean itself.

ACKNOWLEDGEMENTS

This work has made use of NASA's Astrophysics Data System (ADS) services. We thank Dr. Jouni Räisänen for his comments of the manuscript.

CODE AND DATA AVAILABILITY

All manuscript data files, control files and Python codes are stored to
https://zenodo.org/uploads/19854621
The manual for repeating all our analysis is also published there.

REFERENCES

Bailer-Jones, C. A. L. 2012, A&A, 546, A89, doi: 10.1051/0004-6361/201220109

Baluev, R. V. 2009, Monthly Notices of the Royal Astronomical Society, 395, 1541, doi: 10.1111/j.1365-2966.2009.14634.x

Bard, Y. 1974, Nonlinear Parameter Estimation (New York, NY: Academic Press)

Barnston, A. G., Tippett, M. K., L'Heureux, M. L., Li, S., & DeWitt, D. G. 2012, Bulletin of the American Meteorological Society, 93, 631

Berger, J. O. 2013, Statistical decision theory and Bayesian analysis (London: Springer Science & Business Media)

Bjerknes, J. 1969, Monthly weather review, 97, 163

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. 2015, Time series analysis: forecasting and control (Hoboken, New Jersey: John Wiley & Sons)

Bretthorst, G. L. 1988, in Maximum-Entropy and Bayesian Methods in Science and Engineering: Foundations (London: Springer), 75–145

Brockwell, P. J., & Davis, R. A. 2009, Time series: theory and methods (New York, NY: Springer science & Business media)

Bunge, L., & Clarke, A. J. 2009, Journal of Climate, 22, 3979

Cai, W., Borlace, S., Lengaigne, M., et al. 2014, Nature climate change, 4, 111

Callahan, C. W., & Mankin, J. S. 2023, Science, 380, 1064

Cane, M. A., Zebiak, S. E., & Dolan, S. C. 1986, Nature, 321, 827

Charbonneau, P. 2020, Living Reviews in Solar Physics, 17, 1, doi: 10.1007/s41116-020-00025-6

—. 2022, Frontiers in Astronomy and Space Sciences, 9, 853676

Chen, D., Cane, M. A., Kaplan, A., Zebiak, S. E., & Huang, D. 2004, Nature, 428, 733
Cheng, C., Sa-Ngasoongsong, A., Beyca, O., et al. 2015, Iie Transactions, 47, 1053
Chianca, C., Ticona, A., & Penna, T. 2005, Physica A: Statistical Mechanics and its Applications, 357, 447, doi: https://doi.org/10.1016/j.physa.2005.03.047
Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. 1990, Journal of Official Statistics, 6, 3
Connolly, R., Soon, W., Connolly, M., et al. 2021, Research in Astronomy and Astrophysics, 21, 131
Cooley, J. W., & Tukey, J. W. 1965, Mathematics of computation, 19, 297
Draper, N. R., & Smith, H. 1998, Applied Regression Analysis (Hoboken, New Jersey: John Wiley & Sons, Inc.), doi: 10.1002/9781118625590
Efron, B., & Tibshirani, R. 1986, Statistical Science, 1, 54
Engl, H., Hanke, M., & Neubauer, A. 1996, Regularization of Inverse Problems, Mathematics and Its Applications (Dordrecht: Kluwer Academic Publishers Group)
Fourier, J. B. J. 1822, Théorie analytique de la chaleur (Firmin Didot)
Friis-Christensen, E., & Lassen, K. 1991, Science, 254, 698
Fröhlich, C., & Lean, J. 1998, Geophysical Research Letters, 25, 4377
Furlan, C., & Mortarino, C. 2020, Computational Statistics, 35, 1951
Gauß, C. F. 1821, Commentationes Societatis Regiae Scientiarum Gottingensis Recentiores, 1, 1
Gauß, K. F. 1809, Theoria motvs corporvm coelestivm in sectionibvs conicis solem ambientivm. (Sumtibus F. Perthes et I.H. Besser)
Ghaderpour, E., Pagiatakis, S. D., & Hassan, Q. K. 2021, Applied Sciences, 11, doi: 10.3390/app11136141
Gray, L. J., Beer, J., Geller, M., et al. 2010, Reviews of Geophysics, 48
Grushka-Cockayne, Y., Jose, V. R. R., & Lichtendahl Jr, K. C. 2017, Management Science, 63, 1110
Guemas, V., Auger, L., & Doblas-Reyes, F. J. 2014, Journal of Applied Meteorology and Climatology, 53, 637
Hadamard, J. 1902, Princeton University Bulletin, 13, 49
—. 1923, Lectures on Cauchy's Problem in Linear Partial Differential Equations, Mrs. Hepsa Ely Silliman memorial lectures (New Haven, Connecticut: Yale University Press)
Hamilton, J. D. 1994, Time Series Analysis (Princeton: Princeton University Press), doi: 10.1515/9780691218632
Hanley, D. E., Bourassa, M. A., O'Brien, J. J., Smith, S. R., & Spade, E. R. 2003, Journal of Climate, 16, 1249, doi: 10.1175/1520-0442(2003)16⟨1249:AQEOEI⟩2.0.CO;2
Hastie, T., Tibshirani, R., & Friedman, J. 2001, The Elements of Statistical Learning, Springer Series in Statistics (New York, NY, USA: Springer New York Inc.)
Horne, J. H., & Baliunas, S. L. 1986, ApJ, 302, 757, doi: 10.1086/164037
Hsiang, S. M., Meng, K. C., & Cane, M. A. 2011, Nature, 476, 438
Hu, R., Lian, T., Liu, T., et al. 2024, Communications Earth & Environment, 5, 675
Jetsu, L. 2020, The Open Journal of Astrophysics, 3, 4, doi: 10.21105/astro.2002.03890
—. 2025, arXiv e-prints, arXiv:2311.08317, doi: 10.48550/arXiv.2311.08317
Jetsu, L., & Pelt, J. 1999, A&AS, 139, 629, doi: 10.1051/aas:1999411
Kay, S., & Marple, S. 1981, Proceedings of the IEEE, 69, 1380, doi: 10.1109/PROC.1981.12184
Kazemi, M., & Rodrigues, P. C. 2025, Computational Statistics, 40
Kim, D.-W., Protopapas, P., Alcock, C., Byun, Y.-I., & Bianco, F. B. 2009, Monthly Notices of the Royal Astronomical Society, 397, 558, doi: 10.1111/j.1365-2966.2009.14967.x
Klevs, M., Stefani, F., & Jouve, L. 2023, Solar Physics, 298, 90
Koen, C. 1995, Ap&SS, 230, 307, doi: 10.1007/BF00658188
Kopp, G., & Lean, J. L. 2011, Geophysical Research Letters, 38
Kudryavtsev, S. M., Cionco, R. G., & Soon, W. W.-H. 2022, Proceedings of the International Astronomical Union, 18, 139
Lau, K.-M., & Yang, S. 1996, Quarterly Journal of the Royal Meteorological Society, 122, 945
Lavrentiev, M. M., Romanov, V. G., & Shishatskii, S. P. 1986, Translations of Mathematical Monographs, Vol. 64, Ill-posed problems of mathematical physics and analysis (Providence, R.I.: American Mathematical Society)
Lefrancois, P. 1989, International Journal of Forecasting, 5, 553
Legendre, A.-M. 1805, Journal de l'École Polytechnique
Lehtinen, J., Jetsu, L., Hackman, T., Kajatkari, P., & Henry, G. W. 2011, A&A, 527, A136, doi: 10.1051/0004-6361/201015454
Lenssen, N., DiNezio, P., Goddard, L., et al. 2024, Geophysical Research Letters, 51, e2023GL106988, doi: https://doi.org/10.1029/2023GL106988
Liang, X. S., Xu, F., Rong, Y., et al. 2021, Scientific Reports, 11, 17860
Liu, Y., Cai, W., Lin, X., Li, Z., & Zhang, Y. 2023a, Nature Communications, 14, 5887

Liu, Y., Donat, M. G., England, M. H., et al. 2023b, Nature communications, 14, 6387
Lockwood, M. 2012, Surveys in Geophysics, 33, 503
Lomb, N. R. 1976, Ap&SS, 39, 447, doi: 10.1007/BF00648343
Loumos, G. L., & Deeming, T. J. 1978, Ap&SS, 56, 285, doi: 10.1007/BF01879560
Lu, Z., Schultze, A., Carré, M., et al. 2025, Nature Geoscience, 1
Ludescher, J., Gozolchiani, A., Bogachev, M. I., et al. 2013, Proceedings of the National Academy of Sciences, 110, 11742
Makridakis, S., Hibon, M., Lusk, E., & Belhadjali, M. 1987, International Journal of Forecasting, 3, 489
Martinez, P., & Kurtz, D. W. 1990, MNRAS, 242, 636, doi: 10.1093/mnras/242.4.636
Mouël, J.-L. L., Courtillot, V., Kossobokov, V., et al. 2025, arXiv preprint arXiv:2511.18939
Nerlove, M. 1964, Econometrica: Journal of the Econometric Society, 241
Petropoulos, F., Apiletti, D., Assimakopoulos, V., et al. 2022, International Journal of forecasting, 38, 705
Philander, S. G. 1989, International geophysics series, 46, X
Piskunov, N. E., Tuominen, I., & Vilhu, O. 1990, A&A, 230, 363
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Cambridge University Press
Reinhold, T., Reiners, A., & Basri, G. 2013, A&A, 560, A4, doi: 10.1051/0004-6361/201321970
Ren, H.-L., & Jin, F.-F. 2011, Geophysical Research Letters, 38
Rudin, W. 1976, Principles of Mathematical Analysis, International series in pure and applied mathematics (New York, NY: McGraw-Hill)
Scargle, J. D. 1982, ApJ, 263, 835, doi: 10.1086/160554
—. 1989, ApJ, 343, 874, doi: 10.1086/167757
Shumway, R. H., & Stoffer, D. S. 2006, Time series analysis and its applications: with R examples (New York, NY: Springer-Verlag)
Solaraju-Murali, B., Bojovic, D., Gonzalez-Reviriego, N., et al. 2022, Climate services, 27, 100303
Stefani, F., Giesecke, A., Seilmayer, M., Stepanov, R., & Weier, T. 2019, arXiv preprint arXiv:1910.10383
Stefani, F., Horstmann, G., Mamatsashvili, G., & Weier, T. 2025, Solar Physics, 300, 1
Su, E. C.-Y., & Wu, H.-M. 2024, Computational Statistics, 39, 1937
Thirumalai, K., DiNezio, P. N., Partin, J. W., et al. 2024, Nature, 634, 374
Tikhonov, A. N. 1963, Dokl. Akad. Nauk SSSR, 151, 501
Tikhonov, A. N., & Arsenin, V. Y. 1977, Solutions of ill-posed problems (Washington, D.C.: John Wiley & Sons, New York: V. H. Winston & Sons), xiii+258
Timmer, J., & König, M. 1995, A&A, 300, 707
Timmermann, A., An, S.-I., Kug, J.-S., et al. 2018, Nature, 559, 535
Tong, H. 1990, Non Linear Time Series: A Dynamical System Approach (Oxford, England: Clarendon Press)
Tsay, R. S., & Chen, R. 2018, Nonlinear time series analysis, Wiley Series in Probability and Statistics (Hoboken, New Jersey: John Wiley & Sons)
Usoskin, I. G. 2017, Living Reviews in Solar Physics, 14, 3
—. 2023, Living Reviews in Solar Physics, 20, 2, doi: 10.1007/s41116-023-00036-z
Usoskin, I. G., Solanki, S. K., & Kovaltsov, G. A. 2007, A&A, 471, 301, doi: 10.1051/0004-6361:20077704
VanderPlas, J. T. 2018, ApJS, 236, 16, doi: 10.3847/1538-4365/aab766
Vogel, C. R. 2002, SIAM Classics in Applied Mathematics, Vol. 29, Computational Methods for Inverse Problems (Philadelphia, Pennsylvania: SIAM)
Webster, P. J., & Yang, S. 1992, Quarterly Journal of the Royal Meteorological Society, 118, 877
Willson, R. C., & Hudson, H. S. 1991, Nature, 351, 42
Wooldridge, J. M. 2010, Econometric Analysis of Cross Section and Panel Data, 2nd edn. (Cambridge, MA, USA: The MIT Press)
Xu, Y., Zhu, W., Samanta, D., & Horton, B. P. 2026, Nature Climate Change, 1
Zebiak, S. E., & Cane, M. A. 1987, Monthly Weather Review, 115, 2262
Zechmeister, M., & Kürster, M. 2009, A&A, 496, 577, doi: 10.1051/0004-6361:200811296
Zhao, S., Jin, F.-F., Stuecker, M. F., et al. 2024, Nature, 630, 891
Zhu, F. R., & Jia, H. Y. 2018, Ap&SS, 363, 138, doi: 10.1007/s10509-018-3332-z
Zwiers, F. W., & Von Storch, H. 1995, Journal of Climate, 8, 336

Supplementary material

Table S1. Samples. (1) Description. (2-3) Size and time span. (4) Errors available. (5) Electronic data file.

(1) Description (-)	(2) n (-)	(3) ΔT (y)	(4) σ_i (y)	(5) Data file (-)
All yearly	155	154	yes	`Clong.dat`
First half all yearly	78	77	yes	`Chalf.dat`
All yearly	155	154	no	`Rlong.dat`
All monthly	1860	154.9	no	`mAll.dat`
All January	155	154	no	`Jan.dat`
All February	155	154	no	`Feb.dat`
All March	155	154	no	`Mar.dat`
All April	155	154	no	`Apr.dat`
All May	155	154	no	`May.dat`
All June	155	154	no	`Jun.dat`
All July	155	154	no	`Jul.dat`
All August	155	154	no	`Aug.dat`
All September	155	154	no	`Sep.dat`
All October	155	154	no	`Oct.dat`
All November	155	154	no	`Nov.dat`
All December	155	154	no	`Dec.dat`

Table S2. Trend in all yearly mean weighted data. DCM analysis between $P_{\min} = 3$ and $P_{\max} = 100$ years. We use $\gamma = 0.001$ (Equation 48) for yearly data. Otherwise as in Table 8

		Period analysis				Fisher-test ($\gamma = 0.001$)		
		Data: Original weighted data ($n = 155$, $\Delta T = 154$: `Clong.dat`)						
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\mathcal{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathcal{M}$=2	$\mathcal{M}$=3	
	η (-)	A_1 (^{o}C)	A_2 (^{o}C)	A_3 (^{o}C)	A_4 (^{o}C)	F_χ (-)	F_χ (-)	Control file
	χ^2 (-)	$t_{\min,1}$ (y)	$t_{\min,1}$ (y)	$t_{\min,1}$ (y)	$t_{\min,1}$ (y)	Q_F (-)	Q_F (-)	
One signal								
$\mathcal{M}$=1		12.67 ± 0.33	-	-	-	↑	↑	
$\mathcal{M}_{1,1,0,\chi^2}$	4	0.62 ± 0.15	-	-	-	26.9	14.0	`longC14K110.dat`
	516	1871.7 ± 1.3	-	-	-	6.7×10^{7}	2.7×10^{-6}	
$\mathcal{M}$=2		12.65 ± 0.20	-	-	-	-	←	
$\mathcal{M}_{1,1,1,\chi^2}$	5	0.51 ± 0.14	-	-	-	-	1.02	`longC14K111.dat`
	437	1871.8 ± 1.3	-	-	-	-	0.31	
$\mathcal{M}$=3		12.65 ± 0.43			-	-	-	
$\mathcal{M}_{1,1,2,\chi^2}$	6	0.52 ± 0.12			-	-	-	`longC14K112.dat`
	434	1871.9 ± 1.1			-	-	-	
Two signals								
						$\mathcal{M}$=5	$\mathcal{M}$=6	
$\mathcal{M}$=4		12.70 ± 0.17	26.8 ± 2.0	-	-	↑	↑	
$\mathcal{M}_{2,1,0,\chi^2}$	7	0.62 ± 0.12	0.39 ± 0.16	-	-	27.4	13.6	`longC14K210.dat`
	456	1871.3 ± 1.0	1870.8 ± 3.4	-	-	5.6×10^{-7}	3.8×10^{-6}	
$\mathcal{M}$=5		5.677 ± 0.089	12.64 ± 0.31	-	-	-	←	
$\mathcal{M}_{2,1,1,\chi^2}$	8	0.393 ± 0.086	0.47 ± 0.13	-	-	-	0	`longC14K211.dat`
	384	1876.15 ± .56	1871.8 ± 1.2	-	-	-	1	
$\mathcal{M}$=6		5.677 ± 0.077	12.64 ± 0.42		-	-	-	
$\mathcal{M}_{2,1,2,\chi^2}$	9	0.388 ± 0.093	0.47 ± 0.12		-	-	-	`longC14K212.dat`
	384	1876.06 ± 0.76	1871.9 ± 1.2		-	-	-	
Three signals								
						$\mathcal{M}$=8	$\mathcal{M}$=9	
$\mathcal{M}$=7		5.619 ± 0.091	12.74 ± 0.48	21.83 ± 0.81	-	↑	↑	
$\mathcal{M}_{3,1,0,\chi^2}$	10	0.41 ± 0.13	0.52 ± 0.11	0.40 ± 0.12	-	29.2	15.0	`longC14K310.dat`
	407	1871.54 ± 0.77	1871.2 ± 1.7	1870.5 ± 3.2	-	2.6×10^{-7}	1.2×10^{-6}	
$\mathcal{M}$=8		5.662 ± 0.077	12.78 ± 0.12	21.3 ± 1.5	-	-	←	
$\mathcal{M}_{3,1,1,\chi^2}$	11	0.40 ± 0.13	0.47 ± 0.10	0.37 ± 0.12	-	-	0.82	`longC14K311.dat`
	338	1870.77 ± 0.62	1870.86 ± 0.88	1871.3 ± 3.3	-	-	0.36	
$\mathcal{M}$=9		5.661 ± 0.073	12.76 ± 0.22	21.2 ± 1.2	-	-	-	
$\mathcal{M}_{3,1,2,\chi^2}$	12	0.40 ± 0.11	0.49 ± 0.10	0.39 ± 0.12	-	-	-	`longC14K312.dat`
	336	1870.8 ± 0.74	1871.1 ± 1.2	1871.6 ± 2.8	-	-	-	
Four signals								
						$\mathcal{M}$=11	$\mathcal{M}$=12	
$\mathcal{M}$=10		5.597 ± 0.066	12.77 ± 0.19	19.4 ± 1.6 IF	21.5 ± 2.1 IF	↑	↑	
$\mathcal{M}_{4,1,0,\chi^2}$	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	1872.06 ± 0.57	1870.8 ± 1.2	1873.3 ± 3.7	1872.3 ± 3.3	4.5×10^{-7}	2.5×10^{-6}	
$\mathcal{M}$=11		5.496 ± 0.093 IF	5.66 ± 0.13 IF	12.76 ± 0.15	21.1 ± 1.5	-	←	
$\mathcal{M}_{4,1,1,\chi^2}$	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	1874.3 ± 1.2	1870.7 ± 1.2	1871.10 ± 0.97	1871.8 ± 2.1	-	0.50	
$\mathcal{M}$=12		5.497 ± 0.087 IF	0.566 ± 0.062 IF	12.75 ± 0.17	21.0 ± 1.0	-	-	
$\mathcal{M}_{4,1,2,\chi^2}$	15	0.34 ± 0.94 AD	0.42 ± 0.90 AD	0.50 ± 0.11	0.398 ± 0.084	-	-	`longC14K412.dat`
UM	299	1874.3 ± 1.0	1870.68 ± 0.99	1871.1 ± 0.97	1871.9 ± 2.2	-	-	

Table S3. Significance of one signal model $\mathcal{M}$=1 in Table S4. Note that we use $\gamma = 0.005$ for this smallest sample. Otherwise as in Table 9.

(1)	(2)
	Fisher-test ($\gamma = 0.005$)
	$g_{1,1,1}, n = 78$
Polynomial	$\eta = 5, \chi^2 = 168$
$\mathcal{M}$=1	↑
$g(t) = p(t, K_3 = 0)$	$F = 7.2$
$\eta = 1$, $\chi^2 = 235$	$Q_{\rm F} = 6.2 \times 10^{-5}$
$\mathcal{M}$=2	↑
$g(t) = p(t, K_3 = 1)$	5.14
$\eta = 2$, $\chi^2 = 204$	$Q_{\rm F} = 0.0028$
$\mathcal{M}$=3	↑
$g(t) = p(t, K_3 = 2)$	$F = 7.3$
$\eta = 3$, $\chi^2 = 202$	$Q_F = 0.0013$
$\mathcal{M}$=4	↑
$g(t) = p(t, K_3 = 3)$	$F = 12.86$
$\eta = 4$, $\chi^2 = 198$	$Q_{\rm F} = 0.00060$
$\mathcal{M}$=5	↑
$g(t) = p(t, K_3 = 4)$	No test
$\eta = 5$, $\chi^2 = 192$	$Q_{\rm F} = 1$
$\mathcal{M}$=6	↑
$g(t) = p(t, K_3 = 5)$	$F = -8.9$
$\eta = 6$, $\chi^2 = 192$	$Q_{\rm F} = 1$
$\mathcal{M}$=7	↑
$g(t) = p(t, K_3 = 6)$	$F = -4.2$
$\eta = 7$, $\chi^2 = 191$	$Q_{\rm F} = 1$
$\mathcal{M}$=8	↑
$g(t) = p(t, K_3 = 7)$	$F = -2.4$
$\eta = 8$, $\chi^2 = 188$	$Q_{\rm F} = 1$

Table S4. Periods in first half yearly mean data. Otherwise as in Table S2.

		Period analysis				Fisher-test ($\gamma = 0.005$)			
		Data: Original weighted data ($n = 78$, $\Delta T = 77$: `Chalf.dat`)							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathcal{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathcal{M}$=2	$\mathcal{M}$=3	$\mathcal{M}$=4	
	η (-)	A_1 ($^{\rm o}$C)	A_2 ($^{\rm o}$C)	A_3 ($^{\rm o}$C)	A_4 ($^{\rm o}$C)	F_χ (-)	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 (-)	Q_F (-)	
One signal									
$\mathcal{M}$=1		19.05 ± 0.95	-	-	-	↑	↑	↑	
$\mathcal{M}_{1,1,1,\chi^2}$	5	0.44 ± 0.16	-	-	-	5.62	5.95	5.78	`halfC14K111.dat`
	168	1874.2 ± 2.1	-	-	-	0.0016	5.0×10^{-5}	7.5×10^{-6}	
Two signals									
$\mathcal{M}$=2		12.81 ± 0.44	18.8 ± 1.	-	-	-	↑	↑	
$\mathcal{M}_{2,1,1,\chi^2}$	8	0.44 ± 0.10	0.53 ± 0.18	-	-	-	5.3	4.91	`halfC14K211.dat`
	135	1870.9 ± 1.45	1874.8 ± 2.5	-	-	-	0.0025	0.00034	
Three signals									
$\mathcal{M}$=3		5.82 ± 0.14	12.82 ± 0.43	19.33 ± 0.74	-	-	-	←	
$\mathcal{M}_{3,1,1,\chi^2}$	11	0.42 ± 0.13	0.49 ± 0.10	0.50 ± 0.12	-	-	-	3.88	`halfC14K311.dat`
	109	1875.51 ± 0.94	1874.0 ± 1.3	1874.0 ± 1.8	-	-	-	0.013	
Four signals									
$\mathcal{M}$=4		3.437 ± 0.052	5.830 ± 0.047	12.82 ± 0.45	19.57 ± 0.59	-	-	-	
$\mathcal{M}_{4,1,1,\chi^2}$	14	0.30 ± 0.12	0.48 ± 0.13	0.498 ± 0.097	0.54 ± 0.12	-	-	-	`halfC14K411.dat`
	92	1873.01 ± 0.66	1875.42 ± 0.39	1870.7 ± 1.8	1873.6 ± 2.2	-	-	-	

Table S5. Periods in all yearly mean weighted data. Otherwise as in Table S2.

		Period analysis				Fisher-test ($\gamma = 0.001$)			
		Data: Original weighted data ($n = 155$, $\Delta T = 154$: `Clong.dat`)							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathcal{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathcal{M}$=2	$\mathcal{M}$=3	$\mathcal{M}$=4	
	η (-)	A_1 ($^{\rm o}$C)	A_2 ($^{\rm o}$C)	A_3 ($^{\rm o}$C)	A_4 ($^{\rm o}$C)	F_χ (-)	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 (-)	Q_F (-)	
One signal									
$\mathcal{M}$=1		12.65 ± 0.20	-	-	-	↑	↑	↑	
$\mathcal{M}_{1,1,1,\chi^2}$	5	0.51 ± 0.14	-	-	-	12.4	6.98	7.10	`longC14K111.dat`
	437	1871.8 ± 1.3	-	-	-	2.7×10^{-7}	1.6×10^{-6}	1.9×10^{-8}	
Two signals									
$\mathcal{M}$=2		5.677 ± 0.089	12.64 ± 0.31	-	-	-	↑	↑	
$\mathcal{M}_{2,1,1,\chi^2}$	8	0.393 ± 0.086	0.47 ± 0.13	-	-	-	6.48	6,53	`longC14K211.dat`
	384	1876.15 ± 0.56	1871.8 ± 1.2	-	-	-	0.00038	4.1×10^{-6}	
Three signals									
$\mathcal{M}$=3		5.662 ± 0.077	12.78 ± 0.12	21.3 ± 1.5	-	-	-	↑	
$\mathcal{M}_{3,1,1,\chi^2}$	11	0.40 ± 0.13	0.47 ± 0.10	0.37 ± 0.12	-	-	-	5.91	`longC14K311.dat`
	338	1870.77 ± 0.62	1870.86 ± 0.88	1871.3 ± 3.3	-	-	-	0.00079	
Four signals									
$\mathcal{M}$=4		5.496 ± 0.093 IF	5.66 ± 0.13 IF	12.76 ± 0.15	21.09 ± 1.5	-	-	-	
$\mathcal{M}_{4,1,1,\chi^2}$	14	0.3 ± 1.8 AD	0.4 ± 1.8 AD	0.49 ± 0.13	0.39 ± 0.11	-	-	-	`longC14K411.dat`
UM	300	1874.3 ± 1.2	1870.7 ± 1.1	1871.1 ± 0.97	1871.9 ± 22.	-	-	-	

Table S6. Significance of one signal model $\mathcal{M}$=1 in Table S5. Otherwise as in Table 9.

(1)	(2)
	Fisher-test ($\gamma = 0.001$)
	$g_{1,1,1}, n = 155$
Polynomial	$\eta = 5, \chi^2 = 437$
$\mathcal{M}$=1	↑
$g(t) = p(t, K_3 = 0)$	$F = 19.9$
$\eta = 1$, $\chi^2 = 670$	$Q_{\rm F} = 3.9 \times 10^{-13}$
$\mathcal{M}$=2	↑
$g(t) = p(t, K_3 = 1)$	11.4
$\eta = 2$, $\chi^2 = 537$	$Q_{\rm F} = 9.2 \times 10^{-7}$
$\mathcal{M}$=3	↑
$g(t) = p(t, K_3 = 2)$	$F = 17.0$
$\eta = 3$, $\chi^2 = 537$	$Q_{\rm F} = 2.1 \times 10^{-7}$
$\mathcal{M}$=4	↑
$g(t) = p(t, K_3 = 3)$	$F = 33.8$
$\eta = 4$, $\chi^2 = 536$	$Q_{\rm F} = 3.6 \times 10^{-8}$
$\mathcal{M}$=5	↑
$g(t) = p(t, K_3 = 4)$	No test
$\eta = 5$, $\chi^2 = 514$	$Q_{\rm F} = 1$
$\mathcal{M}$=6	↑
$g(t) = p(t, K_3 = 5)$	$F = -22$
$\eta = 6$, $\chi^2 = 514$	$Q_{\rm F} = 1$
$\mathcal{M}$=7	↑
$g(t) = p(t, K_3 = 6)$	$F = -8.2$
$\eta = 7$, $\chi^2 = 492$	$Q_{\rm F} = 1$
$\mathcal{M}$=8	↑
$g(t) = p(t, K_3 = 7)$	$F = -5.4$
$\eta = 8$, $\chi^2 = 492$	$Q_{\rm F} = 1$

Table S7. Periods in all non-weighted yearly mean data. Otherwise as in Table S2.

		Period analysis				Fisher-test ($\gamma = 0.001$)			
		Data: Original non-weighted data ($n = 155$, $\Delta T = 154$: `Rlong.dat`)							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathcal{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathcal{M}$=2	$\mathcal{M}$=3	$\mathcal{M}$=4	
	η (-)	A_1 (-)	A_2 (-)	A_3 (-)	A_4 (-)	F_R (-)	F_R (-)	F_R(-)	Control file
	R (-)	$t_{\rm min,1}$ (y)	$t_{\rm min,1}$ (y)	$t_{\rm min,1}$ (y)	$t_{\rm min,1}$ (y)	Q_F (-)	Q_F (-)	Q_F (-)	
One signal									
$\mathcal{M}$=1		12.71 ± 0.12	-	-	-	←	←	←	
$\mathcal{M}_{1,1,1,R}$	5	0.389 ± 0.074	-	-	-	3.67	3.77	1.11	`longR14K110.dat`
	27.1	1871.69 ± 0.86	-	-	-	0.013	0.0016	0.36	
Two signals									
$\mathcal{M}$=2		5.651 ± 0.030	12.69 ± 0.15	-	-	-	←	←	
$\mathcal{M}_{2,1,1,R}$	8	0.319 ± 0.064	0.384 ± 0.088	-	-	-	3.77	-0.092	`longR14K211.dat`
	25.2	1870.76 ± 0.41	1871.8 ± 1.1	-	-	-	0.0016	1	
Three signals									
$\mathcal{M}$=3		5.650 ± 0.032	9.12 ± 0.12	12.69 ± 0.27	-	-	-	←	
$\mathcal{M}_{3,1,1,R}$	11	0.327 ± 0.084	0.299 ± 0.087	0.381 ± 0.079	-	-	-	-3.50	`longR14K311.dat`
	23.4	1870.78 ± 0.46	1873.0 ± 1.2	1871.9 ± 1.3	-	-	-	1	
Four signals									
$\mathcal{M}$=4		5.496 ± 0.093 IF	5.662 ± 1.3 IF	12.76 ± 0.16	21.1 ± 1.5	-			
$\mathcal{M}_{4,1,1,R}$	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	1874.3 ± 1.2	1870.7 ± 1.2	1871.1 ± 0.97	1871.8 ± 2.1	-			

Table S8. Monthly Niño 4 index means after December 2024. (1) Month. (2) Available on 29.01.2026. (3) Available on 02.05.2026. Monthly values for 2026 are not yet *finalised* in NOAA.

(1)	(2)	(3)
Month	y	y
(-)	°C	°C
Jan-25	-0.50	-0.50
Feb-25	-0.36	-0.36
Mar-25	-0.13	-0.13
Apr-25	-0.09	-0.09
May-25	0.02	0.02
Jun-25	0.14	0.14
Jul-25	0.05	0.05
Aug-25	0.06	0.06
Sep-25	-	-0.12
Oct-25	-	-0.32
Nov-25	-	-0.37
Dec-25	-	-0.19
Jan-26	-	-0.15
Feb-26	-	0.12
Mar-26	-	0.35
Apr-26	-	0.30

Table S9. Periods in non-weighted all months data. Otherwise as in Table S2.

		Period analysis				Fisher-test ($\gamma = 0.001$)			
		Data: Original non-weighted data ($n = 1860$, $\Delta T = 154.9$: `mAll.dat`)							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathcal{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathcal{M}$=2	$\mathcal{M}$=3	$\mathcal{M}$=4	
	η (-)	A_1 (°C)	A_2 (°C)	A_3 (°C)	A_4 (°C)	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 (-)	
One signal									
$\mathcal{M}$=1		12.716 ± 0.047	-	-	-	↑	↑	↑	
$\mathcal{M}_{1,1,1,R}$	5	0.396 ± 0.029	-	-	-	32.9	33.2	35.3	`mAllR14K111.dat`
	556.8	1871.62 ± 0.37	-	-	-	$Q_{\mathrm{F}} < 10^{-16}$	$Q_{\mathrm{F}} < 10^{-16}$	$Q_{\mathrm{F}} < 10^{-16}$	
Two signals									
$\mathcal{M}$=2		5.651 ± 0.012	12.700 ± 0.042	-	-	-	↑	↑	
$\mathcal{M}_{2,1,1,R}$	8	0.348 ± 0.029	0.389 ± 0.028	-	-	-	31.8	34.7	`mAllR14K211.dat`
	528.6	1870.75 ± 0.19	1871.76 ± 0.32	-	-	-	$Q_{\mathrm{F}} < 10^{-16}$	$Q_{\mathrm{F}} < 10^{-16}$	
Three signals									
$\mathcal{M}$=3		4.7459 ± 0.0082	5.640 ± 0.012	12.694 ± 0.043	-	-	-	↑	
$\mathcal{M}_{3,1,1,R}$	11	0.335 ± 0.023	0.340 ± 0.034	0.388 ± 0.030	-	-	-	35.7	`mAllR14K311.dat`
	502.6	1870.60 ± 0.13	1870.93 ± 0.15	1871.83 ± 0.29	-	-	-	$Q_{\mathrm{F}} < 10^{-16}$	
Four signals									
$\mathcal{M}$=4		3.6436 ± 0.0046	4.7477 ± 0.0080	5.647 ± 0.011	12.704 ± 0.063	-	-	-	
$\mathcal{M}_{4,1,1,R}$	14	0.345 ± 0.030	0.332 ± 0.031	0.342 ± 0.025	0.392 ± 0.038	-	-	-	`mAllR14K411.dat`
	475.0	1872.68 ± 0.11	1870.55 ± 0.14	1870.82 ± 0.19	1871.74 ± 0.39.	-	-	-	

Table S10. Significance of one signal model $\mathcal{M}$=1 in Table S9. Otherwise as in Table 9.

(1)	(2)
	Fisher-test ($\gamma = 0.001$)
	$g_{1,1,1}, n = 1860$
Polynomial	$\eta = 5, R = 557.1$
$\mathcal{M}$=1	↑
$g(t) = p(t, K_3 = 0)$	$F = 63.9$
$\eta = 1,\ R = 634$	$Q_{\rm F} < 10^{-16}$
$\mathcal{M}$=2	↑
$g(t) = p(t, K_3 = 1)$	$F = 39.8$
$\eta = 2,\ R = 593$	$Q_{\rm F} < 10^{-16}$
$\mathcal{M}$=3	↑
$g(t) = p(t, K_3 = 2)$	$F = 59.7$
$\eta = 3,\ R = 593$	$Q_{\rm F} < 10^{-16}$
$\mathcal{M}$=4	↑
$g(t) = p(t, K_3 = 3)$	$F = 92.8$
$\eta = 4,\ R = 585$	$Q_{\rm F} < 10^{-16}$
$\mathcal{M}$=5	↑
$g(t) = p(t, K_3 = 4)$	No test
$\eta = 5,\ R = 584$	$Q_{\rm F} = 1$
$\mathcal{M}$=6	↑
$g(t) = p(t, K_3 = 5)$	$F = -34$
$\eta = 6,\ R = 584$	$Q_{\rm F} = 1$
$\mathcal{M}$=7	↑
$g(t) = p(t, K_3 = 6)$	$F = -33.5$
$\eta = 7,\ R = 578$	$Q_{\rm F} = 1$
$\mathcal{M}$=8	↑
$g(t) = p(t, K_3 = 7)$	$F = -22.3$
$\eta = 8,\ R = 578$	$Q_{\rm F} = 1$

Table S11. All months "Big Wave" forecast (Figure 14). (1) Year. (2-13) Month. Symbols ↑ and ↓ tell that preceding value is smaller or bigger.

(1)	(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)		(12)		(13)	
Year	Jan		Feb		Mar		Apr		May		Jun		Jul		Aug		Sep		Oct		Nov		Dec	
(y)	(°C)		(°C)		(°C)		(°C)		(°C)		(°C)		(°C)		(°C)		(°C)		(°C)		(°C)		(°C)	
2025	0.06	↑	0.05	↓	0.04	↓	0.03	↓	0.02	↓	0.02	↓	0.01	↓	0.00	↓	-0.00	↓	-0.01	↓	-0.01	↓	-0.01	↑
2026	-0.00	↑	0.00	↑	0.01	↑	0.01	↑	0.02	↑	0.03	↑	0.04	↑	0.05	↑	0.06	↑	0.08	↑	0.09	↑	0.10	↑
2027	0.11	↑	0.12	↑	0.12	↑	0.13	↑	0.13	↑	0.13	↑	0.13	↑	0.13	↓	0.13	↓	0.12	↓	0.12	↓	0.11	↓
2028	0.10	↓	0.09	↓	0.09	↓	0.08	↓	0.07	↓	0.06	↓	0.05	↓	0.05	↓	0.04	↓	0.04	↓	0.04	↑	0.04	↑
2029	0.05	↑	0.06	↑	0.07	↑	0.08	↑	0.10	↑	0.12	↑	0.14	↑	0.17	↑	0.19	↑	0.22	↑	0.25	↑	0.28	↑
2030	0.31	↑	0.34	↑	0.37	↑	0.40	↑	0.42	↑	0.45	↑	0.47	↑	0.49	↑	0.50	↑	0.52	↑	0.52	↑	0.53	↑
2031	0.53	↑	0.52	↓	0.52	↓	0.50	↓	0.49	↓	0.47	↓	0.45	↓	0.42	↓	0.39	↓	0.36	↓	0.33	↓	0.30	↓
2032	0.27	↓	0.23	↓	0.20	↓	0.17	↓	0.14	↓	0.12	↓	0.09	↓	0.07	↓	0.05	↓	0.03	↓	0.02	↓	0.01	↓
2033	0.01	↓	0.01	↓	0.01	↑	0.02	↑	0.02	↑	0.03	↑	0.05	↑	0.06	↑	0.08	↑	0.09	↑	0.11	↑	0.13	↑
2034	0.14	↑	0.15	↑	0.17	↑	0.17	↑	0.18	↑	0.19	↑	0.19	↑	0.18	↓	0.18	↓	0.16	↓	0.15	↓	0.13	↓
2035	0.11	↓	0.09	↓	0.06	↓	0.03	↓	0.00	↓	-0.03	↓	-0.06	↓	-0.10	↓	-0.13	↓	-0.16	↓	-0.19	↓	-0.22	↓

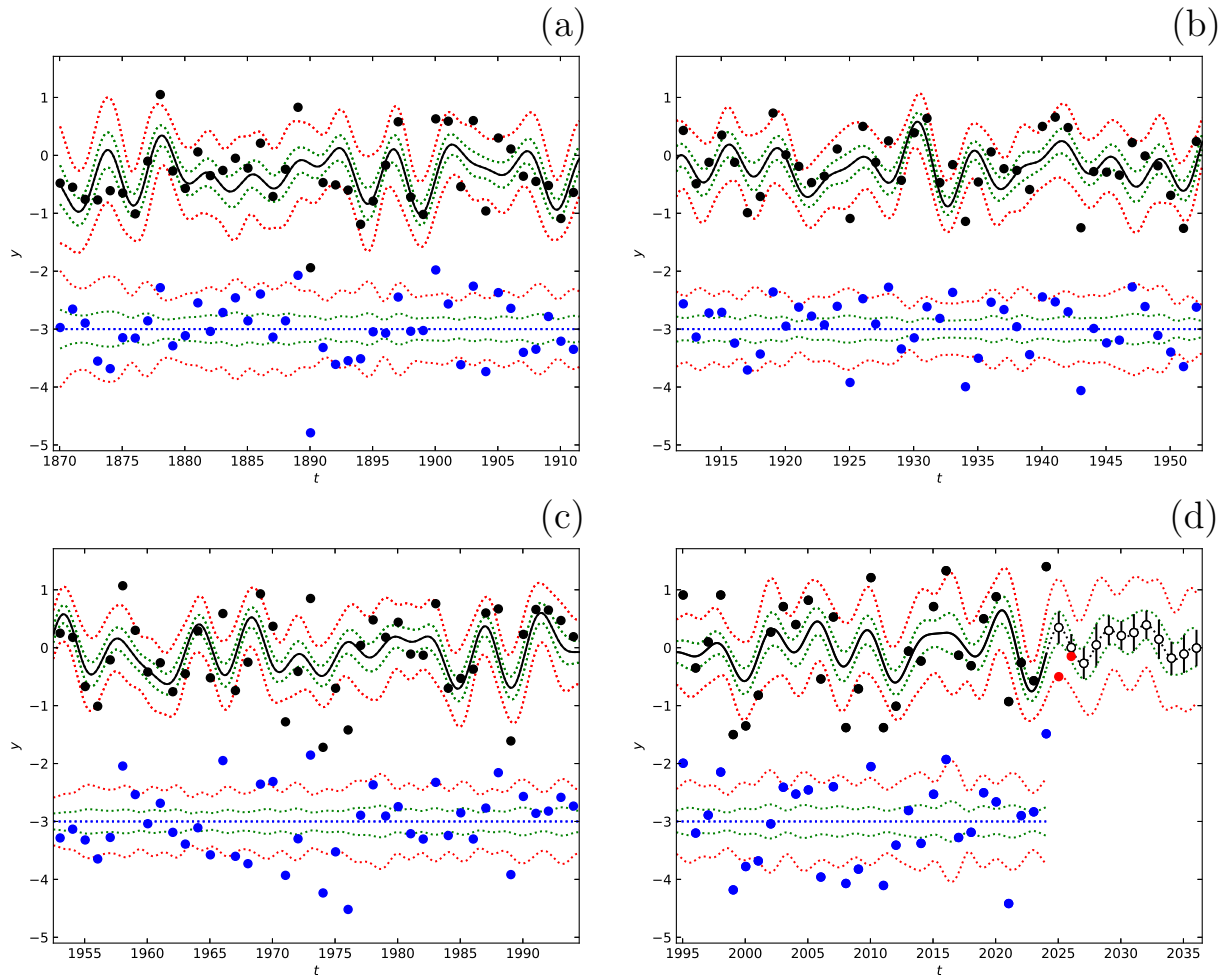

Figure S1. January data model $\mathcal{M}$=4 in Table S12. Red dots from Table S8 (Column 3). Otherwise as in Figure 10.

Table S12. Periods in January data. Note that we use the pre-assigned significance level $\gamma = 0.05$ (Equation 48) because data of separate months data are so noisy. Otherwise as in Table S2.

		Period analysis				Fisher-test ($\gamma = 0.05$)			
		Data: Original non-weighted data ($n = 155$, $\Delta T = 154$: `Jan.dat`)							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathcal{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathcal{M}$=2	$\mathcal{M}$=3	$\mathcal{M}$=4	
	η (-)	A_1 ($^{\rm o}$C)	A_2 ($^{\rm o}$C)	A_3 ($^{\rm o}$C)	A_4 ($^{\rm o}$C)	F_R (-)	F_R (-)	F_R(-)	Control file
	R (-)	$t_{\rm min,1}$ (y)	$t_{\rm min,1}$ (y)	$t_{\rm min,1}$ (y)	$t_{\rm min,1}$ (y)	Q_F (-)	Q_F (-)	Q_F (-)	
One signal									
$\mathcal{M}$=1		3.766 ± 0.030	-	-	-	↑	↑	↑	
$\mathcal{M}_{1,1,1,R}$	5	0.46 ± 0.12	-	-	-	3.33	3.40	3.26	`JanR14K111.dat`
	64.0	1872.18 ± 0.46	-	-	-	0.021	0.0035	0.0012	
Two signals									
$\mathcal{M}$=2		3.767 ± 0.040	4.734 ± 0.046	-	-	-	↑	↑	
$\mathcal{M}_{2,1,1,R}$	8	0.45 ± 0.11	0.46 ± 0.13	-	-	-	3.14	3.09	`JanR14K211.dat`
	59.9	1872.18 ± 0.44	1871.02 ± 0.59	-	-	-	0.027	0.0071	
Three signals									
$\mathcal{M}$=3		3.766 ± 0.045	4.734 ± 0.026	12.72 ± 0.18	-	-	-	↑	
$\mathcal{M}_{3,1,1,R}$	11	0.45 ± 0.10	0.46 ± 0.11	0.44 ± 0.11	-	-	-	2.91	`JanR14K311.dat`
	56.2	1872.19 ± 0.44	1871.02 ± 0.57	1871.7 ± 1.4	-	-	-	0.037	
Four signals									
$\mathcal{M}$=4		3.768 ± 0.018	4.737 ± 0.048	5.643 ± 0.15	12.68 ± 0.30	-	-	-	
$\mathcal{M}_{4,1,1,R}$	14	0.45 ± 0.12	0.44 ± 0.12	0.41 ± 0.10	0.43 ± 0.11	-	-	-	`JanR14K411.dat`
	52.9	1872.21 ± 0.44	1870.94 ± 0.62	1870.84 ± 0.73	1871.9 ± 1.4.	-	-	-	

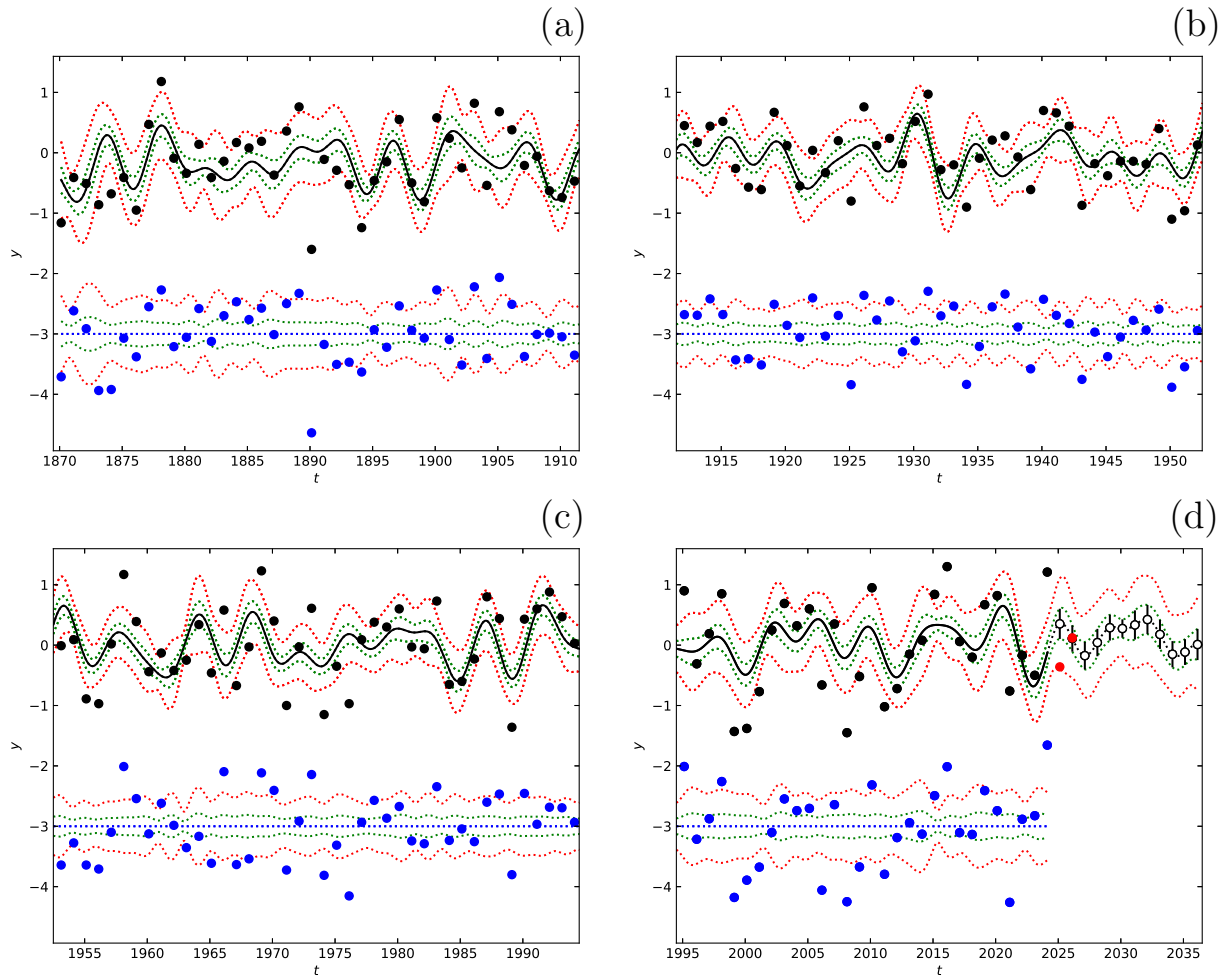

Figure S2. February data model $\mathcal{M}$=4 in Table S13. Otherwise as in Figure S1.

Table S13. Periods in February data. Otherwise as in Table S12.

		Period analysis				Fisher-test ($\gamma = 0.05$)			
		Data: Original non-weighted data ($n = 155, \Delta T = 154$: `Feb.dat`)							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
M		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathcal{M}$=2	$\mathcal{M}$=3	$\mathcal{M}$=4	
	η (-)	A_1 (°C)	A_2 (°C)	A_3 (°C)	A_4 (°C)	F_R (-)	F_R (-)	F_R(-)	Control file
	R (-)	$t_{\min,1}$ (y)	$t_{\min,1}$ (y)	$t_{\min,1}$ (y)	$t_{\min,1}$ (y)	Q_F (-)	Q_F (-)	Q_F (-)	
One signal									
$\mathcal{M}$=1		5.652 ± 0.088	-	-	-	↑	↑	↑	
$\mathcal{M}_{1,1,1,R}$	5	0.48 ± 0.13	-	-	-	3.55	3.36	3.23	`FebR14K111.dat`
	55.8	1870.73 ± 0.46	-	-	-	0.016	0.0039	0.0013	
Two signals									
$\mathcal{M}$=2		5.656 ± 0.025	12.74 ± 0.22	-	-	-	↑	↑	
$\mathcal{M}_{2,1,1,R}$	8	0.47 ± 0.12	0.441 ± 0.086	-	-	-	3.02	2.92	`FebR14K211.dat`
	52.0	1870.68 ± 0.75	1871.3 ± 1.6	-	-	-	0.032	0.010	
Three signals									
$\mathcal{M}$=3		4.741 ± 0.052	5.644 ± 0.089	12.73 ± 0.69	-	-	-	↑	
$\mathcal{M}_{3,1,1,R}$	11	0.40 ± 0.11	0.45 ± 0.11	0.43 ± 0.101	-	-	-	2.72	`FebR14K311.dat`
	48.9	1870.960 ± 0.60	1870.88 ± 0.48	1871.37 ± 1.8	-	-	-	0.046	
Four signals									
$\mathcal{M}$=4		3.770 ± 0.034	4.742 ± 0.046	5.647 ± 0.031	12.73 ± 0.22	-	-	-	
$\mathcal{M}_{4,1,1,R}$	14	0.37 ± 0.11	0.398 ± 0.097	0.45 ± 0.13	0.436 ± 0.094	-	-	-	`FebR14K411.dat`
	46.2	1872.12 ± 0.44	1870.93 ± 0.67	1870.80 ± 0.47	1871.31 ± 1.5	-	-	-	

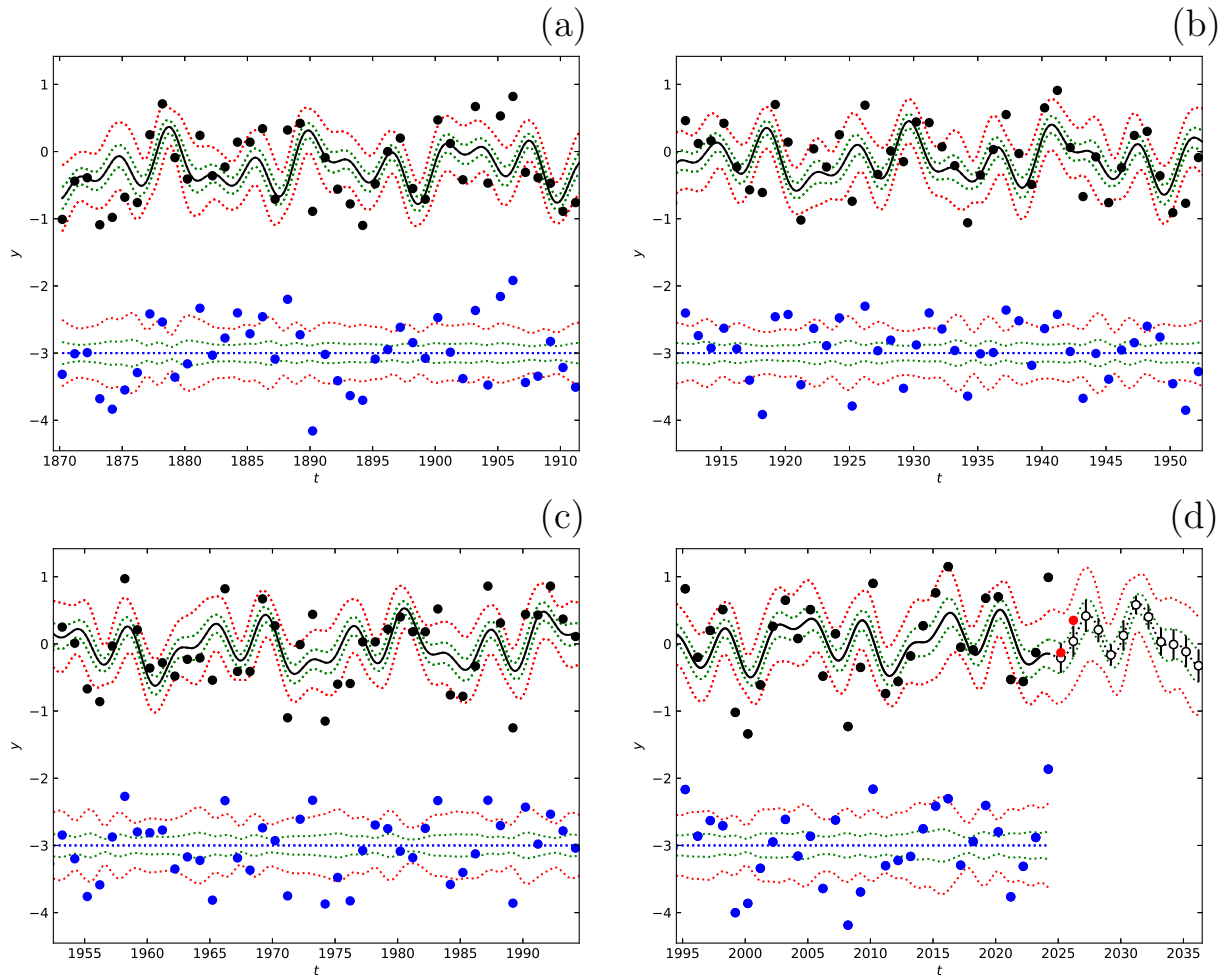

Figure S3. March data model $\mathcal{M}$=3 in Table S14. Otherwise as in Figure S1.

Table S14. Periods in March data. Otherwise as in Table S12.

		Period analysis				Fisher-test ($\gamma = 0.05$)			
		Data: Original non-weighted data ($n = 155, \Delta T = 154$: `Mar.dat`)							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathcal{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathcal{M}$=2	$\mathcal{M}$=3	$\mathcal{M}$=4	
	η (-)	A_1 ($^{\circ}$C)	A_2 ($^{\circ}$C)	A_3 ($^{\circ}$C)	A_4 ($^{\circ}$C)	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 (-)	
One signal									
$\mathcal{M}$=1		5.665 ± 0.029	-	-	-	↑	↑	↑	
$\mathcal{M}_{1,1,1,R}$	5	0.45 ± 0.13	-	-	-	3.94	4.27	3.71	`MarR14K111.dat`
	44.1	1870.67 ± 0.64	-	-	-	0.0097	0.00054	0.00033	
Two signals									
$\mathcal{M}$=2		5.669 ± 0.042	12.68 ± 0.56	-	-	-	↑	↑	
$\mathcal{M}_{2,1,1,R}$	8	0.444 ± 0.092	0.412 ± 0.085	-	-	-	4.33	3.41	`MarR14K211.dat`
	40.8	1870.61 ± 0.53	1871.9 ± 1.1	-	-	-	0.0059	0.0036	
Three signals									
$\mathcal{M}$=3		3.633 ± 0.019	5.671 ± 0.0289	12.68 ± 0.13	-	-	-	←	
$\mathcal{M}_{3,1,1,R}$	11	0.42 ± 0.12	0.45 ± 0.10	0.418 ± 0.090	-	-	-	2.36	`MarR14K311.dat`
	37.4	1873.16 ± 0.26	1870.58 ± 0.50	1871.85 ± 0.77	-	-	-	0.074	
Four signals									
$\mathcal{M}$=4		3.633 ± 0.027	4.585 ± 0.079	5.671 ± 0.034	12.68 ± 0.15	-	-	-	
$\mathcal{M}_{4,1,1,R}$	14	0.419 ± 0.093	0.311 ± 0.084	0.45 ± 0.10	0.418 ± 0.080	-	-	-	`MarR14K411.dat`
	35.6	1873.14 ± 0.17	1870.51 ± 0.48	1870.63 ± 0.65	1871.9 ± 1.5	-	-	-	

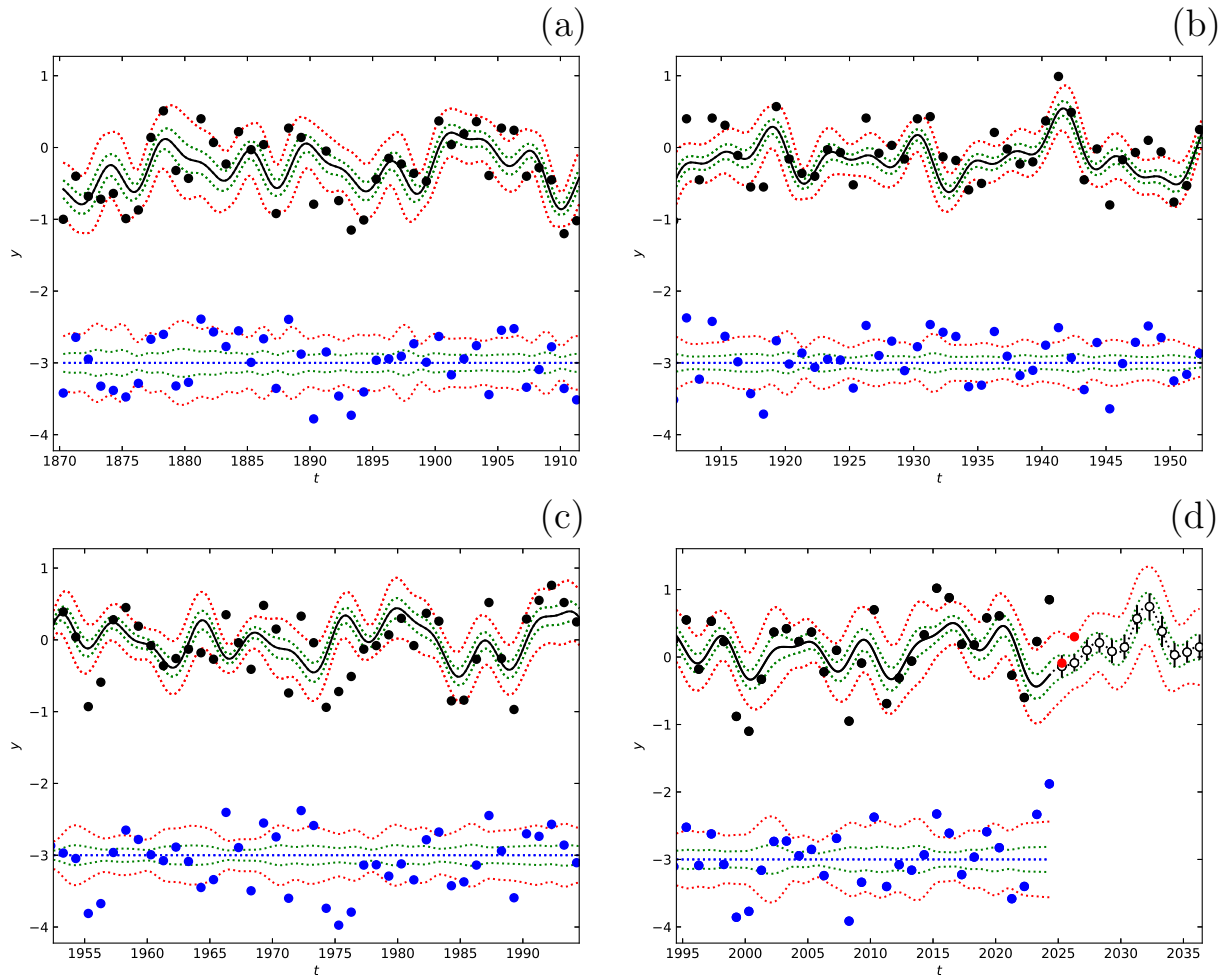

Figure S4. April data model $\mathcal{M}$=4 in Table S15. Otherwise as in Figure S1.

Table S15. Periods in April data. Otherwise as in Table S12.

		Period analysis				Fisher-test ($\gamma = 0.05$)			
		Data: Original non-weighted data ($n = 155, \Delta T = 154$: `Apr.dat`)							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathcal{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathcal{M}$=2	$\mathcal{M}$=3	$\mathcal{M}$=4	
	η (-)	A_1 (^{o}C)	A_2 (^{o}C)	A_3 (^{o}C)	A_4 (^{o}C)	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 (-)	
One signal									
$\mathcal{M}$=1		12.69 ± 0.13	-	-	-	↑	↑	↑	
$\mathcal{M}_{1,1,1,R}$	5	0.42 ± 0.10	-	-	-	5.06	4.29	3.92	`AprR14K111.dat`
	30.8	1871.90 ± 0.92	-	-	-	0.0023	0.00052	0.00018	
Two signals									
$\mathcal{M}$=2		5.663 ± 0.052	12.66 ± 0.14	-	-	-	↑	↑	
$\mathcal{M}_{2,1,1,R}$	8	0.388 ± 0.073	0.416 ± 0.068	-	-	-	3.29	3.13	`AprR14K211.dat`
	27.9	1870.73 ± 0.44	1872.1 ± 1.3	-	-	-	0.022	0.0065	
Three signals									
$\mathcal{M}$=3		5.665 ± 0.017	12.73 ± 0.13	18.83 ± 0.84	-	-	-	↑	
$\mathcal{M}_{3,1,1,R}$	11	0.382 ± 0.097	0.416 ± 0.068	0.31 ± 0.11	-	-	-	2.82	`AprR14K311.dat`
	26.1	1870.71 ± 0.48	1871.65 ± 0.89	1874.0 ± 1.6	-	-	-	0.040	
Four signals									
$\mathcal{M}$=4		3.760 ± 0.039	5.669 ± 0.069	12.74 ± 0.12	18.85 ± 0.37	-	-	-	
$\mathcal{M}_{4,1,1,R}$	14	0.276 ± 0.074	0.382 ± 0.096	0.42 ± 0.13	0.310 ± 0.090	-	-	-	`AprR14K411.dat`
	24.6	1872.27 ± 0.38	1870.65 ± 0.56	1871.59 ± 0.89	1873.9 ± 1.6	-	-	-	

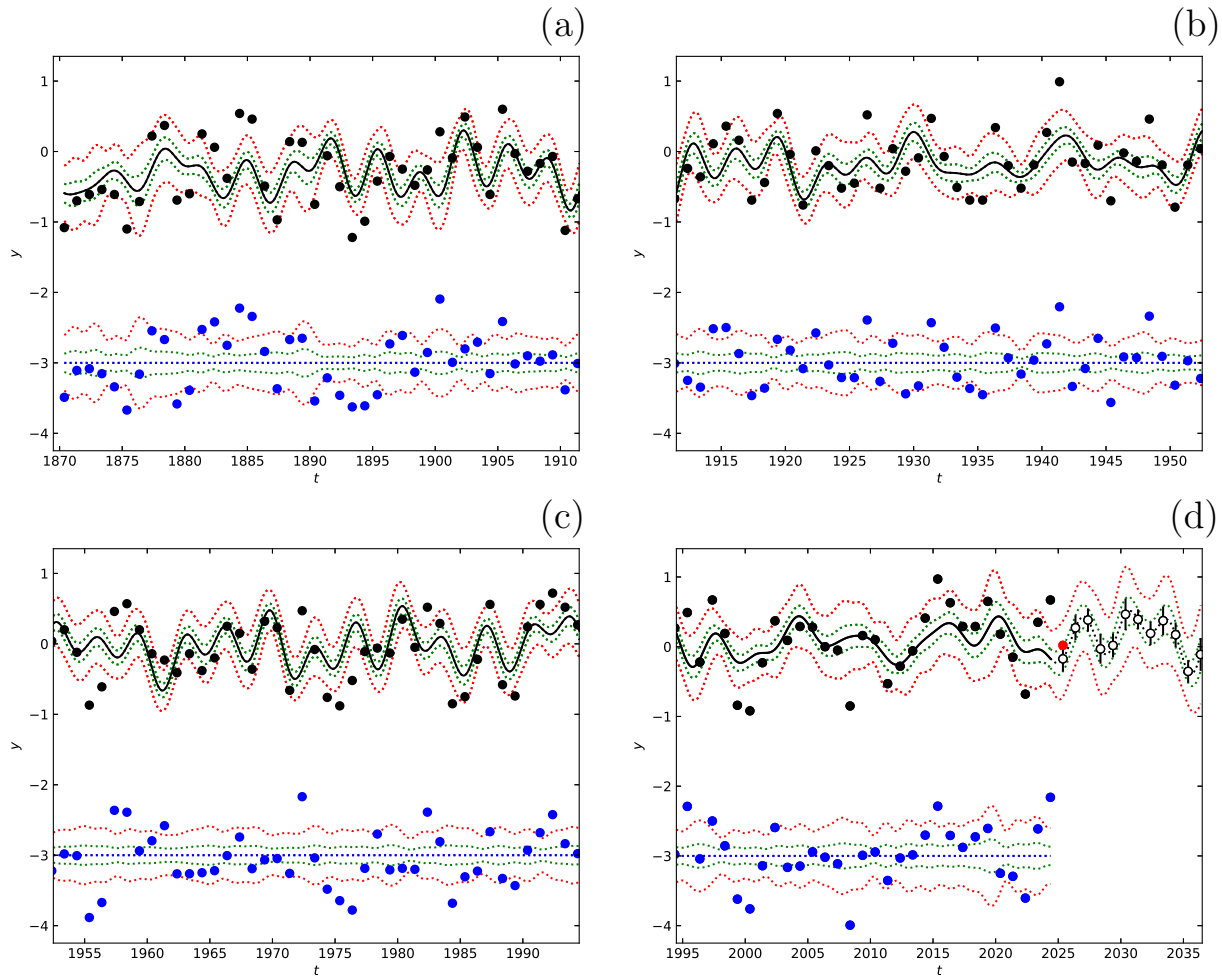

Figure S5. May data model $\mathcal{M}$=4 in Table S16. Otherwise as in Figure S1.

Table S16. Periods in May data. Otherwise as in Table S12.

		Period analysis				Fisher-test ($\gamma = 0.05$)			
		Data: Original non-weighted data ($n = 155, \Delta T = 154$: `May.dat`)							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathcal{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathcal{M}$=2	$\mathcal{M}$=3	$\mathcal{M}$=4	
	η (-)	A_1 (°C)	A_2 (°C)	A_3 (°C)	A_4 (°C)	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 (-)	
One signal									
$\mathcal{M}$=1		12.67 ± 0.24	-	-	-	↑	↑	↑	
$\mathcal{M}_{1,1,1,R}$	5	0.358 ± 0.070	-	-	-	5.67	3.39	3.38	`MayR14K111.dat`
	29.7	1872.4 ± 1.2	-	-	-	0.0010	0.0037	0.00088	
Two signals									
$\mathcal{M}$=2		5.664 ± 0.045	12.65 ± 0.15	-	-	-	↑	↑	
$\mathcal{M}_{2,1,1,R}$	8	0.329 ± 0.073	0.353 ± 0.075	-	-	-	2.93	3.37	`MayR14K211.dat`
	27.6	1870.76 ± 0.56	1872.51 ± 0.96	-	-	-	0.036	0.0060	
Three signals									
$\mathcal{M}$=3		3.388 ± 0.025	5.666 ± 0.062	12.65 ± 0.30	-	-	-	↑	
$\mathcal{M}_{3,1,1,R}$	11	0.288 ± 0.090	0.330 ± 0.057	0.356 ± 0.080	-	-	-	3.26	`MayR14K311.dat`
	26.0	1873.42 ± 0.40	1870.76 ± 0.66	1872.55 ± 0.39	-	-	-	0.023	
Four signals									
$\mathcal{M}$=4		3.383 ± 0.024	3.551 ± 0.042	5.666 ± 0027	12.65 ± 0.12	-	-	-	
$\mathcal{M}_{4,1,1,R}$	14	0.304 ± 0.098	0.30 ± 0.10	0.332 ± 0.067	0.356 ± 0.081	-	-	-	`MayR14K411.dat`
	24.3	1873.53 ± 0.33	1872.03 ± 0.37	1870.7 ± 0.47	1872.5 ± 1.4	-	-	-	

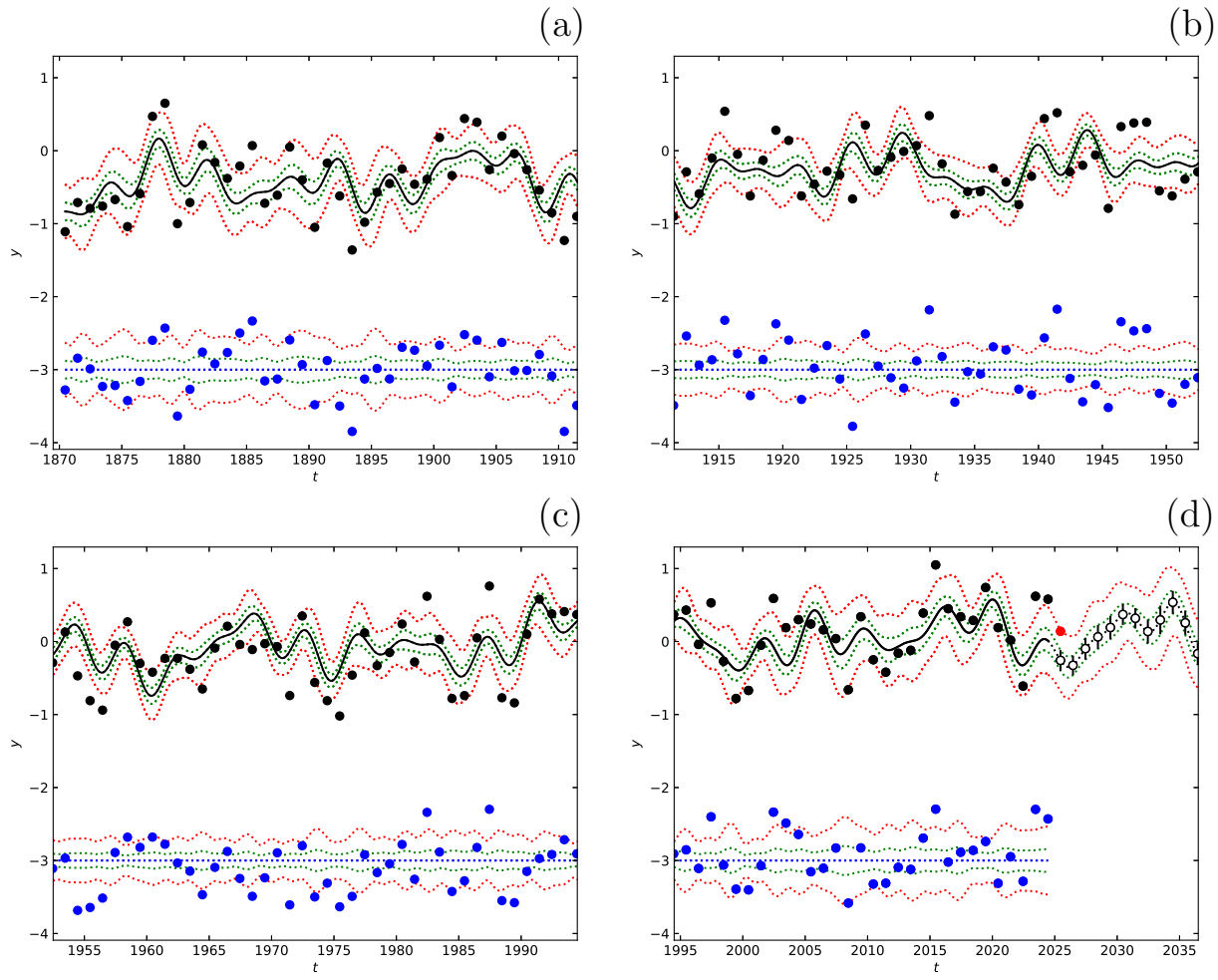

Figure S6. June data model $\mathcal{M}$=4 in Table S17. Otherwise as in Figure S1.

Table S17. Periods in June data. Otherwise as in Table S12.

		Period analysis				Fisher-test ($\gamma = 0.05$)			
		Data: Original non-weighted data ($n = 155, \Delta T = 154$: `Jun.dat`)							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathcal{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathcal{M}$=2	$\mathcal{M}$=3	$\mathcal{M}$=4	
	η (-)	A_1 (^{o}C)	A_2 (^{o}C)	A_3 (^{o}C)	A_4 (^{o}C)	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 (-)	
One signal									
$\mathcal{M}$=1		12.72 ± 0.15	-	-	-	↑	↑	↑	
$\mathcal{M}_{1,1,1,R}$	5	0.42 ± 0.10	-	-	-	4.67	4.07	4.00	`JunR14K111.dat`
	27.4	1872.00 ± 0.86	-	-	-	0.0038	0.00084	0.00014	
Two signals									
$\mathcal{M}$=2		4.740 ± 0.022	12.72 ± 0.14	-	-	-	↑	↑	
$\mathcal{M}_{2,1,1,R}$	8	0.34 ± 0.10	0.42 ± 0.11	-	-	-	3.26	3.42	`JunR14K211.dat`
	25.0	1870.74 ± 0.33	1872.05 ± 0.79	-	-	-	0.023	0.0035	
Three signals									
$\mathcal{M}$=3		4.740 ± 0.018	12.75 ± 0.15	22.46 ± 0.55	-	-	-	↑	
$\mathcal{M}_{3,1,1,R}$	11	0.342 ± 0.072	0.411 ± 0.072	0.297 ± 0.068	-	-	-	3.42	`JunR14K311.dat`
	23.4	1870.73 ± 0.47	1871.9 ± 1.0	1891.5 ± 1.6	-	-	-	0.019	
Four signals									
$\mathcal{M}$=4		3.647 ± 0.012	4.741 ± 0.037	12.76 ± 0.12	22.44 ± 1.7	-	-	-	
$\mathcal{M}_{4,1,1,R}$	14	0.289 ± 0.069	0.338 ± 0.09	0.415 ± 0.077	0.302 ± 0.086	-	-	-	`JunR14K411.dat`
	21.8	1872.53 ± 0.35	1870.8 ± 0.36	1871.81 ± 0.99	1891.5 ± 1.8	-	-	-	

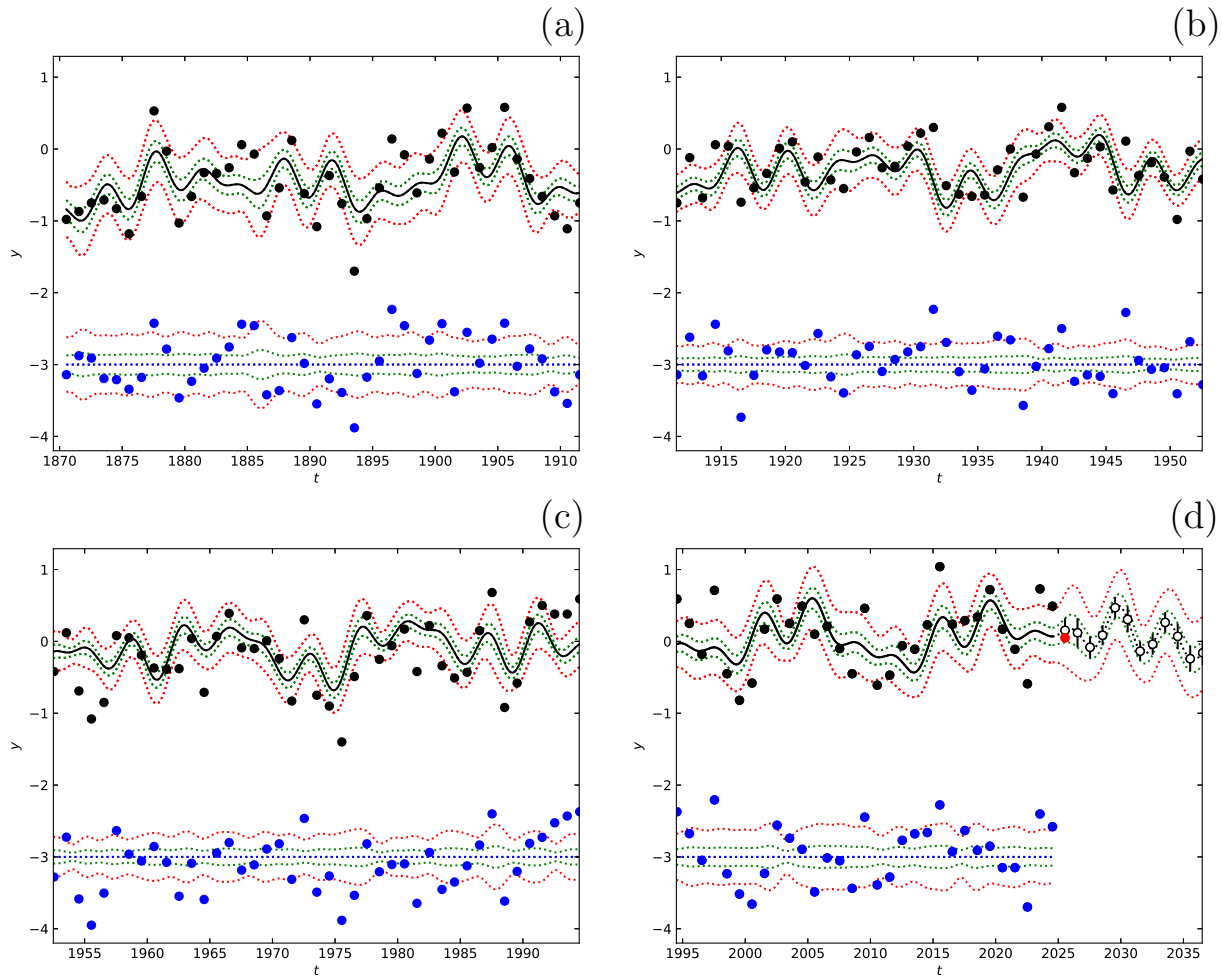

Figure S7. July data model $\mathcal{M}$=4 in Table S18. Otherwise as in Figure S1.

Table S18. Periods in July data. Otherwise as in Table S12.

		Period analysis				Fisher-test ($\gamma = 0.05$)			
		Data: Original non-weighted data ($n = 155, \Delta T = 154$: `Jul.dat`)							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathcal{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathcal{M}$=2	$\mathcal{M}$=3	$\mathcal{M}$=4	
	η (-)	A_1 (°C)	A_2 (°C)	A_3 (°C)	A_4 (°C)	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 (-)	
One signal									
$\mathcal{M}$=1		12.76 ± 0.17	-	-	-	↑	↑	↑	
$\mathcal{M}_{1,1,1,R}$	5	0.372 ± 0.049	-	-	-	4.07	4.23	4.24	`JulR14K111.dat`
	28.5	1871.3 ± 1.4	-	-	-	0.0082	0.00059	7.1×10^{-5}	
Two signals									
$\mathcal{M}$=2		12.77 ± 0.18	20.0 ± 1.0	-	-	-	↑	↑	
$\mathcal{M}_{2,1,1,R}$	8	0.372 ± 0.093	0.330 ± 0.081	-	-	-	4.14	4.06	`JulR14K211.dat`
	26.3	1871.1 ± 1.3	1873.0 ± 2.1	-	-	-	0.0076	0.00087	
Three signals									
$\mathcal{M}$=3		4.738 ± 0.048	12.77 ± 0.16	19.90 ± 0.71	-	-	-	↑	
$\mathcal{M}_{3,1,1,R}$	11	0.331 ± 0.085	0.369 ± 0.078	0.327 ± 0.069	-	-	-	3.75	`JulR14K311.dat`
	24.2	1870.78 ± 0.40	1871.2 ± 1.3	1873.2 ± 1.6	-	-	-	0.012	
Four signals									
$\mathcal{M}$=4		3.545 ± 0.010	4.740 ± 0.043	12.77 ± 0.12	19.97 ± 0.98	-	-	-	
$\mathcal{M}_{4,1,1,R}$	14	0.302 ± 0.070	0.333 ± 0.08	0.370 ± 0.081	0.329 ± 0.047	-	-	-	`JulR14K411.dat`
	22.4	1872.25 ± 0.34	1870.73 ± 0.42	1871.2 ± 1.2	1873.1 ± 2.1	-	-	-	

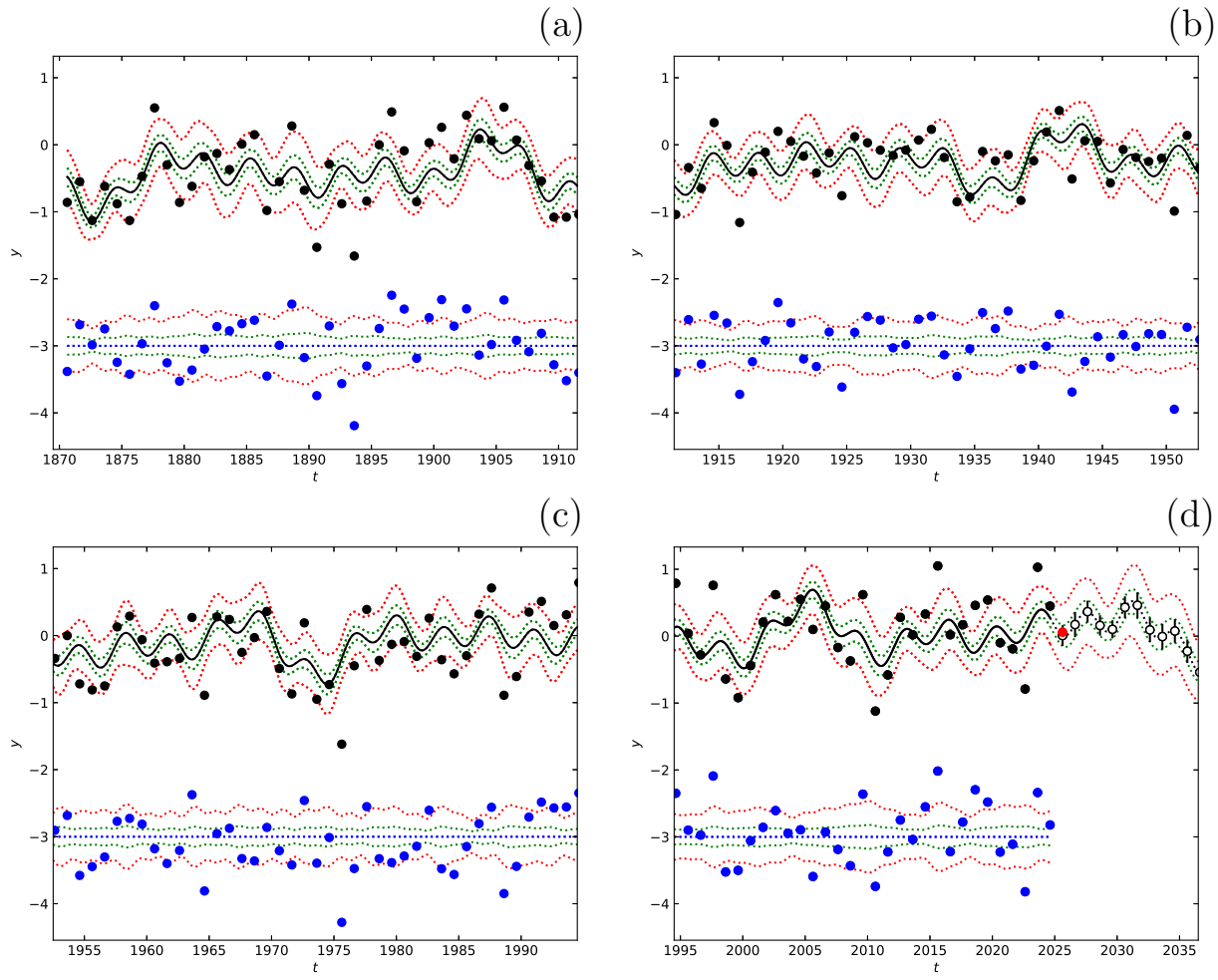

Figure S8. August data model $\mathcal{M}$=4 in Table S19. Red dot is -0.30 for 2025. Otherwise as in Figure 11

Table S19. Periods in August data. Otherwise as in Table S12.

		Period analysis				Fisher-test ($\gamma = 0.05$)			
		Data: Original non-weighted data ($n = 155$, $\Delta T = 154$: `Aug.dat`)							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathcal{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathcal{M}$=2	$\mathcal{M}$=3	$\mathcal{M}$=4	
	η (-)	A_1 (°C)	A_2 (°C)	A_3 (°C)	A_4 (°C)	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 (-)	
One signal									
$\mathcal{M}$=1		3.643 ± 0.036	-	-	-	↑	↑	↑	
$\mathcal{M}_{1,1,1,R}$	5	0.38 ± 0.14	-	-	-	3.86	37.4	3.56	`AugR14K111.dat`
	35.4	1872.50 ± 0.36	-	-	-	0.011	0.0017	0.00051	
Two signals									
$\mathcal{M}$=2		3.647 ± 0.011	20.58 ± 0.89	-	-	-	↑	↑	
$\mathcal{M}_{2,1,1,R}$	8	0.38 ± 0.10	0.368 ± 0.071	-	-	-	3.43	3.26	`Aug14K211.dat`
	32.8	1872.51 ± 0.23	1871.8 ± 2.7	-	-	-	0.019	0.0050	
Three signals									
$\mathcal{M}$=3		3.647 ± 0.017	12.56 ± 0.42	20.58 ± 0.44	-	-	-	↑	
$\mathcal{M}_{3,1,1,R}$	11	0.38 ± 0.10	0.331 ± 0.081	0.367 ± 0.071	-	-	-	2.92	`AugR14K311.dat`
	30.6	1872.51 ± 0.32	1873.0 ± 1.5	1871.9 ± 2.0	-	-	-	0.036	
Four signals									
$\mathcal{M}$=4		3.648 ± 0.010	9.13 ± 0.19	12.51 ± 0.31	20.6 ± 1.2	-	-	-	
$\mathcal{M}_{4,1,1,R}$	14	0.382 ± 0.084	0.311 ± 0.064	0.334 ± 0.084	0.353 ± 0.092	-	-	-	`AugR14K411.dat`
	28.8	1872.51 ± 0.36	1872.94 ± 0.64	1873.6 ± 1.1	1871.9 ± 1.9	-	-	-	

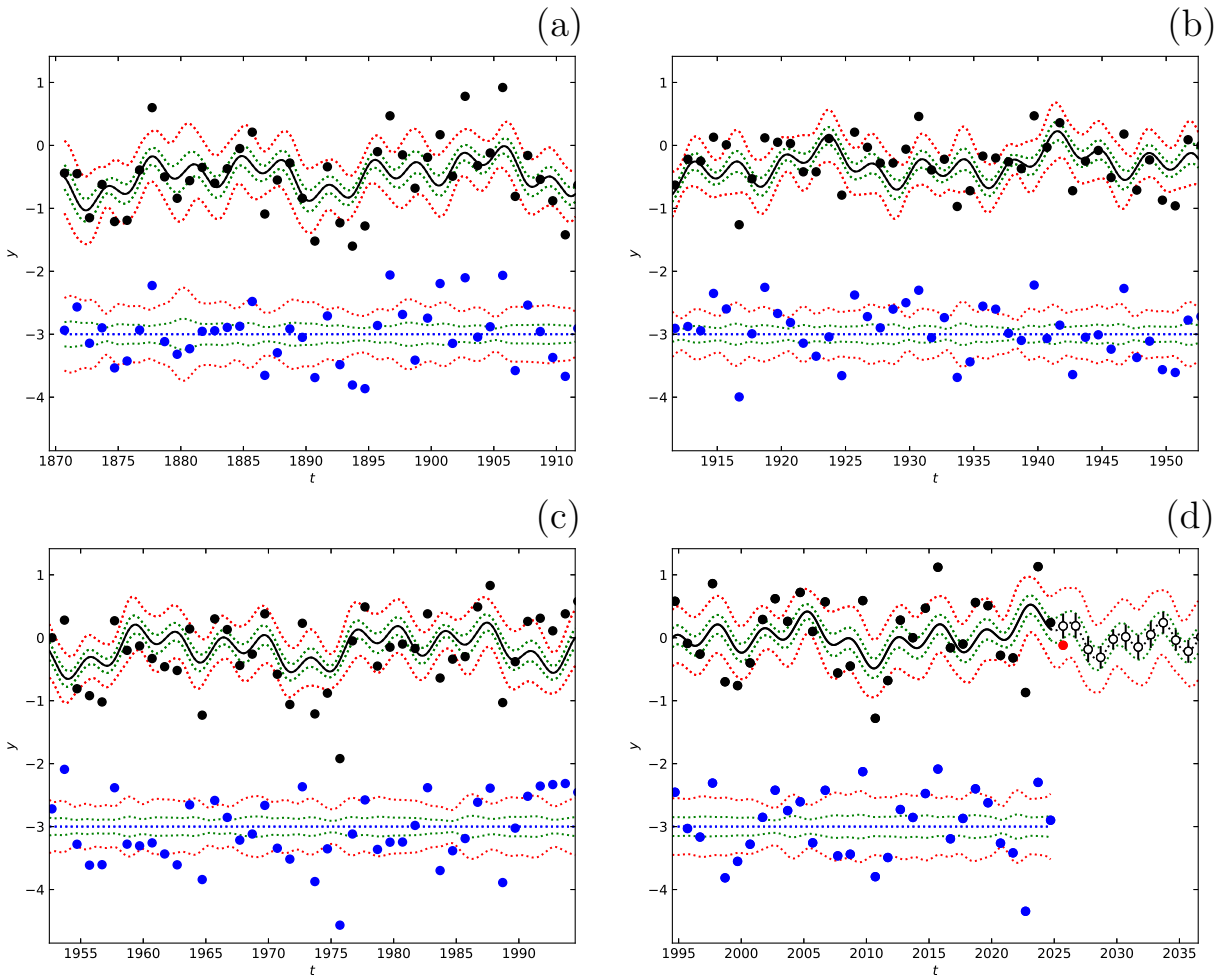

Figure S9. September data model $\mathcal{M}$=3 in Table S20. Otherwise as in Figure S1.

Table S20. Periods in September data. Otherwise as in Table S12.

		Period analysis				Fisher-test ($\gamma = 0.05$)			
		Data: Original non-weighted data ($n = 155$, $\Delta T = 154$: `Sep.dat`)							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathcal{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathcal{M}$=2	$\mathcal{M}$=3	$\mathcal{M}$=4	
	η (-)	A_1 (°C)	A_2 (°C)	A_3 (°C)	A_4 (°C)	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 (-)	
One signal									
$\mathcal{M}$=1		9.12 ± 0.11	-	-	-	↑	↑	↑	
$\mathcal{M}_{1,1,1,R}$	5	0.350 ± 0.078	-	-	-	2.84	2.81	2.78	`SepR14K111.dat`
	43.5	1873.23 ± 0.99	-	-	-	0.040	0.013	0.0050	
Two signals									
$\mathcal{M}$=2		3.546 ± 0.022	20.1 ± 1.3	-	-	-	↑	↑	
$\mathcal{M}_{2,1,1,R}$	8	0.35 ± 0.13	0.350 ± 0.076	-	-	-	2.70	2.66	`Sep14K211.dat`
	41.1	1872.32 ± 0.34	1872.3 ± 2.9	-	-	-	0.048	0.018	
Three signals									
$\mathcal{M}$=3		3.546 ± 0.018	9.13 ± 0.21	19.99 ± 0.89	-	-	-	←	
$\mathcal{M}_{3,1,1,R}$	11	0.35 ± 0.11	0.340 ± 0.094	0.341 ± 0.083	-	-	-	2.53	`SepR14K311.dat`
	38.9	1872.31 ± 0.47	1873.18 ± 0.81	1872.6 ± 2.3	-	-	-	0.060	
Four signals									
$\mathcal{M}$=4		3.545 ± 0.025	4.749 ± 0.071	9.12 ± 0.11	19.92 ± 0.93	-	-	-	
$\mathcal{M}_{4,1,1,R}$	14	0.35 ± 0.12	0.32 ± 0.11	0.343 ± 0.079	0.340 ± 0.082	-	-	-	`SepR14K411.dat`
	36.9	1872.35 ± 0.27	1875.18 ± 0.74	1873.2PM1.4	1872.7 ± 2.3	-	-	-	

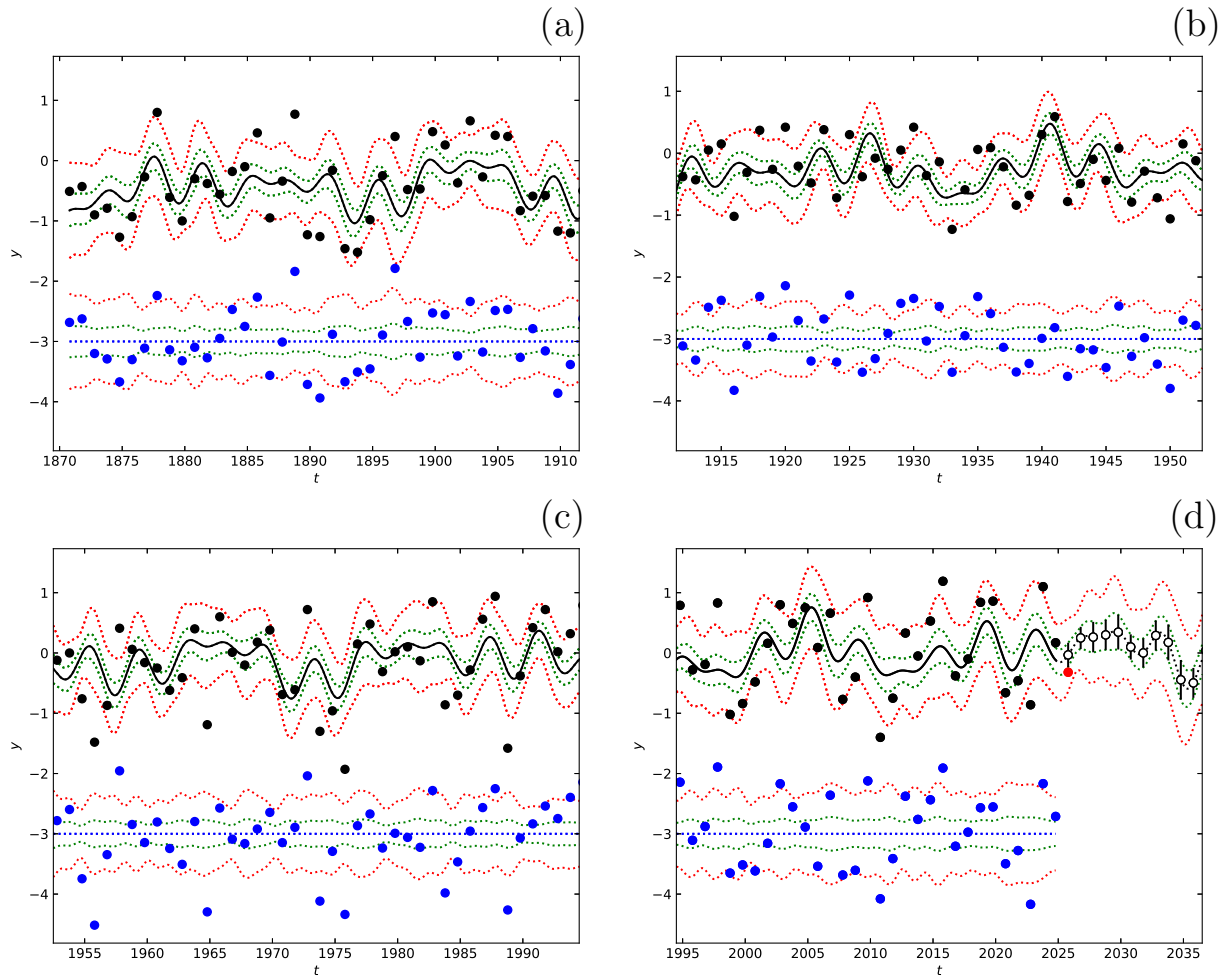

Figure S10. October data model $\mathcal{M}$=4 in Table S21. Otherwise as in Figure S1.

Table S21. Periods in October data. Otherwise as in Table S12.

		Period analysis				Fisher-test ($\gamma = 0.05$)			
		Data: Original non-weighted data ($n = 155$, $\Delta T = 154$: `Oct.dat`)							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathcal{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathcal{M}$=2	$\mathcal{M}$=3	$\mathcal{M}$=4	
	η (-)	A_1 (°C)	A_2 (°C)	A_3 (°C)	A_4 (°C)	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 (-)	
One signal									
$\mathcal{M}$=1		3.354 ± 0.049	-	-	-	↑	↑	↑	
$\mathcal{M}_{1,1,1,R}$	5	0.38 ± 0.13	-	-	-	2.51	2.70	2.78	`OctR14K111.dat`
	56.0	1872.40 ± 0.54	-	-	-	0.039	0.016	0.0049	
Two signals									
$\mathcal{M}$=2		3.542 ± 0.022	4.576 ± 0.045	-	-	-	←	↑	
$\mathcal{M}_{2,1,1,R}$	8	0.38 ± 0.11	0.40 ± 0.12	-	-	-	2.46	2.65	`Oct14K211.dat`
	52.9	1871.62 ± 0.41	1874.92 ± 0.88	-	-	-	0.065	0.018	
Three signals									
$\mathcal{M}$=3		3.542 ± 0.022	4.578 ± 0.047	12.82 ± 0.47	-	-	-	↑	
$\mathcal{M}_{3,1,1,R}$	11	0.381 ± 0.093	0.419 ± 0.11	0.367 ± 0.093	-	-	-	2.75	`OctR14K311.dat`
	50.3	1872.42 ± 0.40	1874.88 ± 0.48	1883.4 ± 2.0	-	-	-	0.045	
Four signals									
$\mathcal{M}$=4		3.542 ± 0.035	5.579 ± 0.058	12.84 ± 0.39	20.4 ± 1.0	-	-	-	
$\mathcal{M}_{4,1,1,R}$	14	0.38 ± 0.10	0.40 ± 0.15	0.38 ± 0.13	0.38 ± 0.11	-	-	-	`OctR14K411.dat`
	47.5	1872.44 ± 0.46	1874.86 ± 0.47	1883.0 ± 1.4	1872.0 ± 2.7	-	-	-	

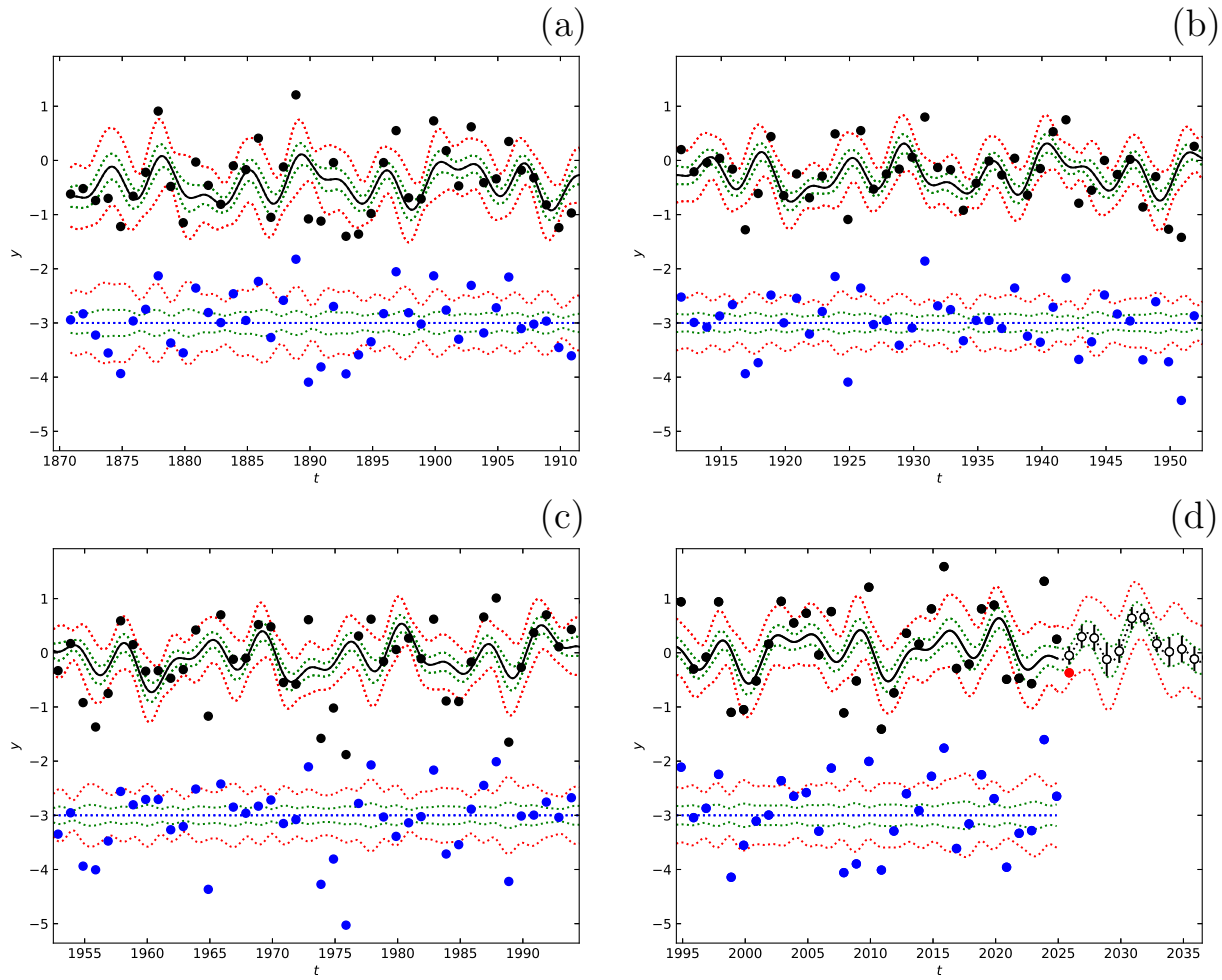

Figure S11. November data model $\mathcal{M}$=3 in Table S22. Otherwise as in Figure S1.

Table S22. Periods in November data. Otherwise as in Table S12.

		Period analysis				Fisher-test ($\gamma = 0.05$)			
		Data: Original non-weighted data ($n = 155$, $\Delta T = 154$: `Nov.dat`)							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathcal{M}$		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathcal{M}$=2	$\mathcal{M}$=3	$\mathcal{M}$=4	
	η (-)	A_1 (°C)	A_2 (°C)	A_3 (°C)	A_4 (°C)	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 (-)	
One signal									
$\mathcal{M}$=1		5.665 ± 0.052	-	-	-	↑	↑	↑	
$\mathcal{M}_{1,1,1,R}$	5	0.47 ± 0.10	-	-	-	2.86	3.02	3.08	`NovR14K111.dat`
	64.8	1876.05 ± 0.40	-	-	-	0.039	0.0081	0.0021	
Two signals									
$\mathcal{M}$=2		3.647 ± 0.019	5.668 ± 0.078	-	-	-	↑	↑	
$\mathcal{M}_{2,1,1,R}$	8	0.43 ± 0.12	0.47 ± 0.14	-	-	-	3.07	3.06	`Nov14K211.dat`
	61.2	1872.54 ± 0.48	1876.01 ± 0.57	-	-	-	0.039	0.0076	
Three signals									
$\mathcal{M}$=3		3.647 ± 0.041	5.670 ± 0.049	12.86 ± 0.63	-	-	-	←	
$\mathcal{M}_{3,1,1,R}$	11	0.44 ± 0.14	0.46 ± 0.12	0.43 ± 0.13	-	-	-	2.67	`NovR14K311.dat`
	57.5	1872.56 ± 0.44	187596 ± 0.59	1883.1 ± 1.2	-	-	-	0.050	
Four signals									
$\mathcal{M}$=4		3.647 ± 0.029	5.667 ± 0.072	9.095 ± 0.090	12.90 ± 0.31	-	-	-	
$\mathcal{M}_{4,1,1,R}$	14	0.43 ± 0.13	0.47 ± 0.15	0.42 ± 0.12	0.432 ± 0.096	-	-	-	`NovR14K411.dat`
	54.1	1872.55 ± 0.41	1876.01 ± 0.67	1873.54 ± 0.78	1882.8 ± 1.7	-	-	-	

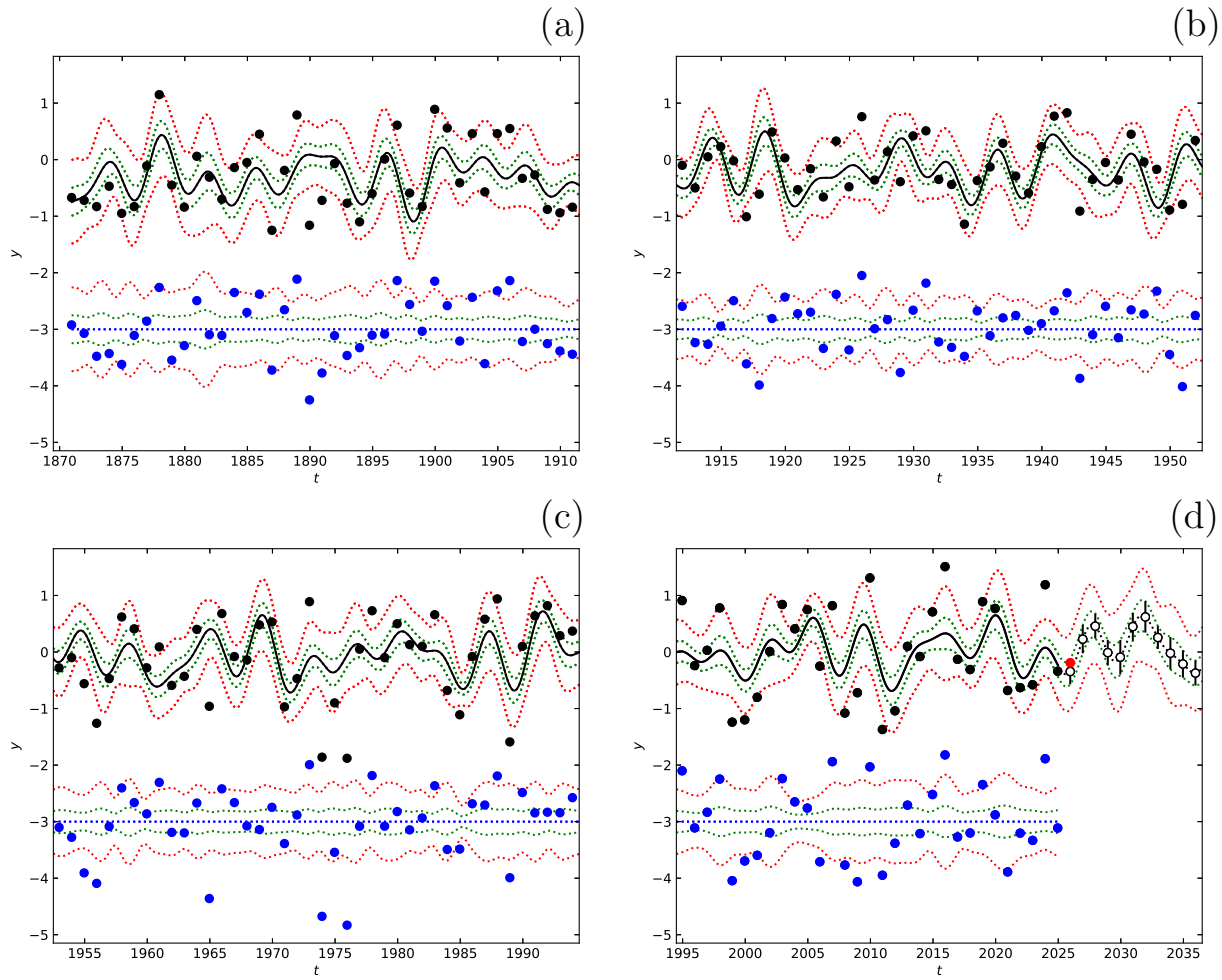

Figure S12. December data model $\mathcal{M}$=4 in Table S23. Otherwise as in Figure S1.

Table S23. Periods in December data. Otherwise as in Table S12.

		Period analysis				Fisher-test ($\gamma = 0.05$)			
		Data: Original non-weighted data ($n = 155$, $\Delta T = 154$: `Dec.dat`)							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
M		P_1 (y)	P_2 (y)	P_3 (y)	P_4 (y)	$\mathcal{M}$=2	$\mathcal{M}$=3	$\mathcal{M}$=4	
	η (-)	A_1 (°C)	A_2 (°C)	A_3 (°C)	A_4 (°C)	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 (-)	
One signal									
$\mathcal{M}$=1		3.645 ± 0.019	-	-	-	↑	↑	↑	
$\mathcal{M}_{1,1,1,R}$	5	0.48 ± 0.12	-	-	-	4.97	4.20	3.94	`DecR14K111.dat`
	64.8	1872.66 ± 0.45	-	-	-	0.0026	0.00064	0.00017	
Two signals									
$\mathcal{M}$=2		3.645 ± 0.016	12.75 ± 0.38	-	-	-	↑	↑	
$\mathcal{M}_{2,1,1,R}$	8	0.49 ± 0.15	0.50 ± 0.16	-	-	-	3.20	3.20	`Dec14K211.dat`
	58.8	1872.67 ± 0.36	1871.29 ± 0.95	-	-	-	0.025	0.0056	
Three signals									
$\mathcal{M}$=3		3.643 ± 0.029	5.664 ± 0.090	12.74 ± 0.18	-	-	-	↑	
$\mathcal{M}_{3,1,1,R}$	11	0.49 ± 0.10	0.43 ± 0.10	0.50 ± 0.11	-	-	-	3.06	`DecR14K311.dat`
	55.1	1872.69 ± 0.38	1876.19 ± 0.61	1871.35 ± 0.98	-	-	-	0.030	
Four signals									
$\mathcal{M}$=4		3.644 ± 0.040	4.566 ± 0.020	5.664 ± 0.058	12.74 ± 0.32	-	-	-	
$\mathcal{M}_{4,1,1,R}$	14	0.49 ± 0.10	0.42 ± 0.10	0.428 ± 0.092	0.50 ± 0.12	-	-	-	`DecR14K411.dat`
	51.7	1872.71 ± 0.38	1875.35 ± 0.59	1876.22 ± 0.80	1871.4 ± 1.2	-	-	-	

Table S24. Separate months. "Big Wave" forecast (Figures S1-S12: open circles). Otherwise as in Table S11.

(1) Year (y)	(2) Jan (°C)	(3) Feb (°C)	(4) Mar (°C)	(5) Apr (°C)	(6) May (°C)	(6) Jun (°C)	(8) Jul (°C)	(9) Aug (°C)	(10) Sep (°C)	(11) Oct (°C)	(12) Nov (°C)	(13) Dec (°C)
2025	$0.3 \pm 0.3 \uparrow$	$0.4 \pm 0.3 \uparrow$	$-0.2 \pm 0.2 \downarrow$	$-0.1 \pm 0.2 \uparrow$	$-0.2 \pm 0.2 \downarrow$	$-0.3 \pm 0.1 \downarrow$	$0.2 \pm 0.2 \uparrow$	$0.0 \pm 0.2 \downarrow$	$0.2 \pm 0.2 \uparrow$	$-0.0 \pm 0.2 \downarrow$	$-0.0 \pm 0.2 \downarrow$	$-0.3 \pm 0.2 \downarrow$
2026	$0.0 \pm 0.2 \uparrow$	$0.1 \pm 0.2 \uparrow$	$0.0 \pm 0.2 \downarrow$	$-0.1 \pm 0.1 \downarrow$	$0.3 \pm 0.2 \uparrow$	$-0.3 \pm 0.1 \downarrow$	$0.1 \pm 0.2 \uparrow$	$0.2 \pm 0.2 \uparrow$	$0.2 \pm 0.2 \uparrow$	$0.2 \pm 0.2 \uparrow$	$0.3 \pm 0.2 \uparrow$	$0.2 \pm 0.2 \downarrow$
2027	$-0.3 \pm 0.3 \downarrow$	$-0.2 \pm 0.2 \uparrow$	$0.4 \pm 0.2 \uparrow$	$0.1 \pm 0.2 \downarrow$	$0.4 \pm 0.2 \uparrow$	$-0.1 \pm 0.1 \downarrow$	$-0.1 \pm 0.2 \uparrow$	$0.4 \pm 0.2 \uparrow$	$-0.2 \pm 0.2 \downarrow$	$0.3 \pm 0.3 \uparrow$	$0.3 \pm 0.2 \uparrow$	$0.5 \pm 0.2 \uparrow$
2028	$0.0 \pm 0.4 \downarrow$	$0.0 \pm 0.2 \downarrow$	$0.2 \pm 0.2 \uparrow$	$0.2 \pm 0.1 \uparrow$	$-0.0 \pm 0.2 \downarrow$	$0.1 \pm 0.2 \uparrow$	$0.1 \pm 0.1 \uparrow$	$0.2 \pm 0.2 \uparrow$	$-0.3 \pm 0.2 \downarrow$	$0.3 \pm 0.2 \uparrow$	$-0.1 \pm 0.3 \downarrow$	$-0.0 \pm 0.2 \uparrow$
2029	$0.3 \pm 0.3 \uparrow$	$0.3 \pm 0.2 \downarrow$	$-0.2 \pm 0.2 \downarrow$	$0.1 \pm 0.2 \uparrow$	$0.0 \pm 0.2 \downarrow$	$0.2 \pm 0.2 \uparrow$	$0.5 \pm 0.1 \uparrow$	$0.1 \pm 0.1 \downarrow$	$-0.0 \pm 0.2 \downarrow$	$0.3 \pm 0.3 \uparrow$	$0.0 \pm 0.2 \downarrow$	$-0.1 \pm 0.3 \downarrow$
2030	$0.2 \pm 0.3 \uparrow$	$0.3 \pm 0.2 \uparrow$	$0.1 \pm 0.2 \downarrow$	$0.1 \pm 0.2 \uparrow$	$0.5 \pm 0.2 \uparrow$	$0.4 \pm 0.1 \downarrow$	$0.3 \pm 0.2 \downarrow$	$0.4 \pm 0.2 \uparrow$	$0.0 \pm 0.2 \downarrow$	$0.1 \pm 0.2 \uparrow$	$0.6 \pm 0.2 \uparrow$	$0.4 \pm 0.2 \downarrow$
2031	$0.3 \pm 0.3 \downarrow$	$0.3 \pm 0.2 \uparrow$	$0.6 \pm 0.1 \uparrow$	$0.6 \pm 0.2 \downarrow$	$0.4 \pm 0.1 \downarrow$	$0.3 \pm 0.1 \downarrow$	$-0.1 \pm 0.1 \downarrow$	$0.5 \pm 0.2 \uparrow$	$-0.1 \pm 0.2 \downarrow$	$0.0 \pm 0.2 \uparrow$	$0.7 \pm 0.1 \uparrow$	$0.6 \pm 0.3 \downarrow$
2032	$0.4 \pm 0.2 \downarrow$	$0.4 \pm 0.2 \uparrow$	$0.4 \pm 0.2 \downarrow$	$0.7 \pm 0.2 \uparrow$	$0.2 \pm 0.2 \downarrow$	$0.1 \pm 0.2 \downarrow$	$-0.0 \pm 0.1 \downarrow$	$0.1 \pm 0.2 \uparrow$	$0.0 \pm 0.2 \downarrow$	$0.3 \pm 0.2 \uparrow$	$0.2 \pm 0.2 \downarrow$	$0.3 \pm 0.2 \uparrow$
2033	$0.1 \pm 0.3 \downarrow$	$0.2 \pm 0.2 \uparrow$	$0.0 \pm 0.2 \downarrow$	$0.4 \pm 0.2 \uparrow$	$0.4 \pm 0.2 \downarrow$	$0.3 \pm 0.2 \downarrow$	$0.3 \pm 0.1 \downarrow$	$-0.0 \pm 0.2 \downarrow$	$0.2 \pm 0.2 \uparrow$	$0.2 \pm 0.3 \downarrow$	$0.0 \pm 0.3 \downarrow$	$-0.0 \pm 0.3 \downarrow$
2034	$-0.2 \pm 0.3 \downarrow$	$-0.2 \pm 0.2 \uparrow$	$-0.0 \pm 0.2 \uparrow$	$0.0 \pm 0.2 \uparrow$	$0.2 \pm 0.2 \uparrow$	$0.5 \pm 0.2 \uparrow$	$0.1 \pm 0.2 \downarrow$	$0.1 \pm 0.2 \uparrow$	$-0.0 \pm 0.2 \downarrow$	$-0.4 \pm 0.3 \downarrow$	$0.1 \pm 0.2 \uparrow$	$-0.2 \pm 0.2 \downarrow$
2035	$-0.1 \pm 0.3 \uparrow$	$-0.1 \pm 0.2 \downarrow$	$-0.1 \pm 0.2 \downarrow$	$0.1 \pm 0.2 \uparrow$	$-0.4 \pm 0.2 \downarrow$	$0.3 \pm 0.2 \uparrow$	$-0.2 \pm 0.2 \downarrow$	$-0.2 \pm 0.2 \uparrow$	$-0.2 \pm 0.2 \uparrow$	$-0.5 \pm 0.2 \downarrow$	$-0.1 \pm 0.2 \uparrow$	$-0.4 \pm 0.2 \downarrow$
2036	$-0.0 \pm 0.3 \uparrow$	$0.0 \pm 0.3 \uparrow$	$-0.3 \pm 0.2 \downarrow$	$0.1 \pm 0.2 \uparrow$	$-0.1 \pm 0.2 \downarrow$	$-0.2 \pm 0.2 \downarrow$	$-0.2 \pm 0.2 \uparrow$	$-0.5 \pm 0.1 \downarrow$	$0.0 \pm 0.2 \uparrow$	$0.0 \pm 0.3 \uparrow$	$-0.1 \pm 0.2 \downarrow$	$-0.1 \pm 0.2 \uparrow$

Table S25. Signal S_{11^y}, S_{10^y}, $S_{11.^y86}$ and $S_{12.^y78}$ connections.

S_{10^y}	S_{11^y}	$S_{11.^y86}$	$S_{12.^y78}$
p_1	p_2	p_3	p_4
10 y	11 y	11.86 y	12.78 y
f_1	f_2	f_3	f_4
0.1000 y^{-1}	0.0909 y^{-1}	0.0843 y^{-1}	0.0782 y^{-1}
	$f_1 - f_2$	$f_2 - f_3$	$f_3 - f_4$
	0.0091 y^{-1}	0.0066 y^{-1}	0.0061 y^{-1}