

ТЕСТИРОВАНИЕ ДЕТЕРМИНИРОВАННОЙ КОМПОНЕНТЫ ВРЕМЕННЫХ РЯДОВ ЧИСЕЛ ВОЛЬФА МЕТОДАМИ КОМБИНАТОРНОЙ СЛОЖНОСТИ

Макаренко Н.Г., Волобуев Д.М., Князева И.С.

Главная (Пулковская) астрономическая обсерватория РАН, Санкт-Петербург, Россия

TESTING THE DETERMINISTIC COMPONENT OF TIME SERIES OF WOLF NUMBERS BY COMBINATORIAL COMPLEXITY METHODS

Makarenko N.G., Volobuev D.M., Knyazeva I.S.

Pulkovo Observatory, St Petersburg, Russia

<https://doi.org/10.31725/0552-5829-2023-211-214>

Wolf number series are usually considered usually as deterministically generated observed solar activity dynamics, in the sense of Takens. Therefore, in solar-terrestrial coupling problems, these series were considered as a deterministic signal with a wide range of periods, in the range (2–200) years, which provided any climatic applications. However, the validity of this set of Fourier harmonics is weakly conditioned: recent revisions have reduced it to only a single, apparently chaotic 11-year mode. Alternative tests based on embedology methods are limited by the small length of the instrumental series. In this work, we use combinatorial methods of algebraic complexity and Topological Data Analysis (TDA) to test the reality of a deterministic component in the Wolf number series. We find no confirmation of the deterministic nature of Wolf cycles. Our estimates are consistent with the hypothesis that the 11-year cycles are a manifestation of dynamical noise observed on a timescale much shorter than the lifetime of the star.

Введение

Сложность моделей распределенной динамики, согласованных с наблюдаемыми временными рядами, обусловлена тем, что последние не содержат явной информации о фазовой геометрии системы. Кроме того, доступные наблюдениям масштабы намного меньше, чем характерные времена системы; данные нелинейные, нестационарные и содержат шумы неизвестной природы. Для динамики Солнечной Активности (СА) примером являются ряды чисел Вольфа. В задачах солнечно-земных связей их обычно рассматривают как *детерминированно-порожденная наблюдаемая*, в смысле Такенса [1]. Поэтому присутствие в них детерминированной компоненты считается неизбежным. Так, считается, что Фурье спектр этих рядов содержит богатый набор периодов, от квази-двухлеток до 200-летних, которые и обеспечивают любые климатические приложения. Однако реальность этого набора обусловлена слабо: недавние ревизии редуцировали его лишь до одной, хаотической 11-летней моды [2]. Тесты, основанные на методах эмбедологии [1], ограничены длиной инструментального ряда, которая недостаточна для корректного анализа. Напомним,

что необходимая для реконструкции m -мерного аттрактора длина ряда N оценивается как $N > 42^m$ [3].

Использование топологических шаблонов, основанных на идее сохранить симметрию второго порядка, связанную с инверсией магнитного поля, было реализовано в [4] с помощью перехода от запаздывающих координат (u_1, u_2) в R^2 к модифицированным (X, Y) :

$$\begin{aligned} u_1 &= X^2 - Y^2 \\ u_2 &= 2XY \end{aligned},$$

инвариантным относительно изменения знака. Фазовая реконструкция по числам Вольфа стала похожа, в новых координатах, на геометрию аттрактора Реслера. К сожалению, правдоподобность этой эвристики нельзя доказать на основе короткого инструментального ряда.

В этой работе мы используем методы алгоритмической сложности [5, 6]. Основой является понятие случайной последовательности вещественных чисел $A = \{a_i\}_{i=1}^N$, измеренной в дискретные моменты времени $i = 1, 2, \dots, N$. Ее сложность можно оценить как длину программы $P(A)$ универсальной машины Тьюринга в виде минимальной бинарной последовательности команд. Длина программы $|P(A)|$ имеет нижнюю грань $K(A) \geq |P(A)|$ при $N \rightarrow \infty$. Случайная последовательность *несжимаема*, в том смысле, что информация, необходимая для ее репродуцирования, не меньше информации, необходимой для ее посимвольного кодирования, и имеет максимальное значение $K(A)$. Сложность неслучайной, «сжимаемой» последовательности вычисляется длиной минимальной программы $|P(A)|$, которая сама является случайной, в указанном выше смысле.

Величину $K(A)$ можно использовать в качестве параметрической меры сложности для диагностики динамических режимов системы, полагая, что A является типичной проекцией наблюдаемой фазовой точки, в смысле Такенса [1]. Для этого используют следующие соображения символической динамики. Пусть в фазовом пространстве M , $\dim(M) = m$ задана динамическая система в форме группы диффеоморфизмов $f^t : M \rightarrow M$ или каскада, для дискретного времени и Липшиц-непрерывная наблюдаемая $\varphi(x) \equiv \varphi(f^i(x_0)) = a_i \in A$ для некоторой начальной типичной точки $x_0 \in M$. Каждая траектория $t \rightarrow x(t)$ посредством отображения $\varphi(f^t x)$ генерирует символическую последовательность (временной ряд) $A = \{a_i\}_{i=1}^N$ как сдвиг $A \rightarrow f(A)$. Идентифицируем траекторию в M разбиением его на ε -боксы в

дискретные моменты времени $\tau: x_i = x(i\tau), i = 1, 2, \dots, N$. Тогда пучок полутраекторий, выходящих из бокса с центром x_0 , ε -неразрешим. Однако в момент $t = k\tau$ мы можем обнаружить, что число занятых боксов $\mathcal{N}(\varepsilon, t)$ из-за расхождения траекторий. Двойной предел (порядок существенен):

$$h_T = \lim_{\varepsilon \rightarrow 0} \lim_{t \rightarrow \infty} \log(\mathcal{N}(\varepsilon, t)/t)$$

называют топологической энтропией системы [7]. Ее можно интерпретировать как скорость, с которой мы получаем информацию Хартли о системе: $I = \log(\mathcal{N}(\varepsilon, t))$, наблюдая *a priori* равновероятные последовательности $\{A\}$. Энтропия h_T зависит от эволюции всего пучка траекторий, но можно предположить, что справедливы условия эргодичности. Тогда топологическая энтропия связана со сложностью отдельной траектории $x(t)$, которую можно оценить алгоритмической сложностью ее типичной проекции $K(\varphi(x))$. Диффеоморфизмы $f^t: M \rightarrow M$ и сдвиги $A \rightarrow f(A)$ топологически не эквивалентны, но для широко класса систем справедливо неравенство [7]: $K(\varphi(x)) \leq h_T$. Это мотивирует важность оценок алгоритмической сложности наблюдаемой (временного ряда). Мы используем эти методы для тестирования детерминированной компоненты СА. Мы пытаемся проверить гипотезу о существовании детерминированного хаоса, режим которого мы наблюдаем в форме 11-летней циклической моды. Такая гипотеза подробно обсуждалась в [4, 8]. В принципе, существует несколько возможных вариантов. Один из них, оценка алгоритмической сложности ряда Вольфа на основе алгоритма Лемпеля-Зива [6], по словарю численного решения для аттрактора Ресслера. Второй основан на реконструкции топологии из временного ряда [9] с последующей оценкой чисел Бетти по фильтрации полученного комплекса методами Топологического Анализа Данных (TDA). Мы приводим здесь результаты наиболее простого теста, основанного на перемешивании ряда и вычисления перестановочной энтропии [10].

Методика

Ради краткости поясним ее на примере. Рассмотрим последовательность, или «слово» из семи «букв», которыми служат целые числа: $x = (4, 7, 9, 10, 6, 11, 3)$. Можно думать, что это начальный фрагмент временного ряда. Подсчитаем, сколько пар отсчетов удовлетворяют отношениям порядка: $\pi: x_t < x_{t+1}$ and $\pi: x_t > x_{t+1}$. Первому варианту будет соответствовать 4 из 6 случаев: $\{4 < 7; 7 < 9; 9 < 10; 6 < 11\}$, а второму два случая $\{10 > 6; 11 > 3\}$. Соответствующая этим двум отношениям энтропия Шеннона будет равна: $H(2) = -(4/6)\log(4/6) - (2/6)\log(2/6) \approx 0.918$.

Аналогичным образом можно получить $H(3)$ для отношений порядка троек отсчетов. В общем случае, для каждого типа π таких отношений, нетрудно оценить вероятности:

$$p(\pi) = \frac{\#\{t | t \leq T - n, (x_{t+1}, \dots, x_{t+n}) \text{ has type } \pi\}}{T - n + 1}$$

и соответствующие энтропии [10]. Полученные оценки можно сравнить с энтропиями случайной, т.е. хорошо темперированной последовательности [5]. Мы использовали модифицированный вариант оценок перестановочной энтропии [11] для сравнения сложности среднемесячных инструментальных значений чисел Вольфа (SSN2) с суррогатным рядом, полученным случайным перемешиванием ряда с сохранением только парных корреляций. Мы не обнаружили существенных отличий перестановочных энтропий оригинала и его суррогатного аналога. Этот результат согласуется с гипотезой о хаотической природе 11-летней моды. Разумеется, достоверность вывода требует дальнейших исследований.

Литература

1. *Sauer T., et al.* Embedology // J. Stat. Phys. 1991. V. 65. P. 579-616.
2. *Frick P.F., et al.* Spectral Properties of Low-order Dynamo Systems // Russian Journal of Nonlinear Dynamics. 2022. V. 18. P. 289-296.
3. *Smith L.A.* Intrinsic limits on dimension calculations // Phys. Lett. A. 1988. V. 133. P. 283-288.
4. *Letellier C. et al.* Evidence for low dimensional chaos in sunspot cycles // Astronomy & Astrophysics. 2006. V. 449. P. 379-387.
5. *Chaitin G.J.* Randomness and mathematical proof // Scient. Amer. 1975. V. 232. P. 47-52.
6. *Kaspar F., Schuster H.G.* Easily calculable measure for the complexity of spatiotemporal patterns // Phys. Rev. A. 1987. V. 36. №. 2. P. 842-848.
7. *Брудно Ф.Ф.* Топологическая энтропия // УМН. 1974. Т. 29. вып.6(180). С. 157-158
8. *Kadtke J., Kremliovsky M.* Classifying complex deterministic signals // Kravtsov Y.A., Kadtke J.B. (ed.). Predictability of complex dynamical systems. Springer, 1996. P.79.
9. *Muldoon M.R. et al.* Topology from time series // Physica D: Nonlinear Phenomena. 1993. V. 65. №. 1-2. P. 1-16.
10. *Bandt C., Pompe B.* Permutation entropy: a natural complexity measure for time series // Phys. Rev. Lett. 2002. V. 88. P. 174102.
11. *Azami H., Escudero J.* Improved multiscale permutation entropy for biomedical signal analysis: Interpretation and application to electroencephalogram recordings // Biomedical Signal Processing and Control. 2016. V. 23. P. 28-41.