



УДК: 530.182, 51-73

MSC 2010: 37M10

Оптимальный подбор параметров прогностических моделей в методе нелинейной причинности по Грейнджеру в приложении к сигналам, характеризваемым хорошо выраженными временными масштабами

М. В. Корнилов, И. В. Сысоев, Б. П. Безручко

Задача определения наличия и направления связи между различными системами по их экспериментальным наблюдаемым временным рядам актуальна во многих областях знания. Одним из подходов к ее решению является метод нелинейной причинности по Грейнджеру, основанный на построении прогностических моделей. Эффективность этого подхода зависит от подбора параметров модели. Для сигналов с характерным временным масштабом колебаний важнейшими параметрами являются временной лаг, используемый при реконструкции вектора состояния, и дальность прогноза.

Получено 21 апреля 2014 года

После доработки 8 августа 2014 года

Работа выполнена при поддержке Российского фонда фундаментальных исследований, грант № 14-02-00492, и гранта Президента Российской Федерации для государственной поддержки ведущих научных школ Российской Федерации, НШ-1726.2014.2.

Корнилов Максим Вячеславович

kornilovmv@gmail.com

Сысоев Илья Вячеславович

ivssci@gmail.com

Саратовский государственный университет им. Н. Г. Чернышевского
410012, Россия, г. Саратов, ул. Астраханская, д. 83

Безручко Борис Петрович

bezruchkobp@gmail.com

Саратовский государственный университет им. Н. Г. Чернышевского
410012, Россия, г. Саратов, ул. Астраханская, д. 83

Саратовский филиал Института радиотехники и электроники им. В. А. Котельникова РАН
410019, Россия, г. Саратов, ул. Зеленая, д. 38

В данной работе предложены два критерия оценки работоспособности метода нелинейной причинности по Грейнджеру, которые позволяют выбрать значения лага и дальности прогноза, при которых достигаются лучшие чувствительность и специфичность; чувствительность определяется тем, насколько слабую связь может детектировать метод, а под специфичностью понимается способность избегать ложноположительных результатов. В результате применения предложенных критериев на примере нескольких однонаправленно связанных эталонных систем были получены практические рекомендации по подбору параметров модели: лага и дальности прогноза.

Ключевые слова: поиск связанности, причинность по Грейнджеру, моделирование по временным рядам

1. Введение

Определение наличия и направления связи между подсистемами сложных систем (задача поиска связанности) по их экспериментальным наблюдаемым (записям их колебаний) временным рядам актуально во многих областях знания. Для решения этой задачи разработано множество методов, в том числе: вычисление функции кросс-корреляции, когерентности, индекса фазовой синхронизации, функции взаимной информации, а также подходы, основанные на построении прогностических моделей, к которым относится и рассматриваемый метод причинности по Грейнджеру [1]. Автор разработал его применительно к экономике, но сейчас он успешно применяется, например, в физиологии для выявления связей между различными отделами головного мозга [2–4], в климатологии — для предсказания поведения муссонов [5]. Несмотря на прозрачность идеи метода, его работоспособность зависит от деталей процедуры реализации.

В данной работе мы предлагаем новый подход к реализации метода для случая, когда временные реализации имеют ярко выраженные собственные характерные временные масштабы (один или несколько), которым соответствуют максимумы в спектре частот. Когда спектр узкополосный, его характерный масштаб удобно определять по автокорреляционной функции. Такие системы достаточно распространены в природе: например, ритмическое изменение солнечной активности (число пятен на Солнце), колебательные изменения потенциалов головного мозга (например, пик-волновые комплексы при абсансной эпилепсии или θ -ритм во время сна), сигналы сердца (основной сердечный ритм в кардиограмме), и др. Идея подхода состоит в том, что параметры модели, имеющие смысл времени (временной лаг l , используемый при вложении для реконструкции вектора состояния [6–8]) и дальность прогноза τ следует связать с характерным масштабом наблюдаемых колебаний. Некоторые исследования в данном направлении уже проводились: в работе [9] показано, что недостаточная частота выборки и, как следствие, слишком большие значения τ и l ведут к систематическим неустраняемым ошибкам в определении направления связанности, в первую очередь — к появлению ложноположительных выводов о наличии связи. В работе [10] было показано, что к подобным ошибкам могут вести и слишком малые значения τ , и было предложено использовать значение τ , равное четверти характерного периода наблюдаемых колебаний.

Однако в целом вопрос об оптимальном выборе параметров τ и l до сих пор далек от решения даже для отдельных узких классов сигналов, в том числе вследствие того, что использованные ранее численные критерии работоспособности метода причинности по Грейнджеру носили слишком частный характер — учитывали только один из параметров метода:

либо дальность прогноза, как в [10], либо шаг выборки, как в [9], либо вид используемых аппроксимирующих функций, как в [11]. Поэтому в данной работе предложены и апробированы на ряде эталонных связанных нелинейных систем два новых численных критерия, характеризующих работоспособность метода в зависимости от τ и l при различных величинах силы связи. На основе применения этих критериев сделаны общие выводы о наилучших и наихудших значениях τ и l для моделей, используемых в методе причинности по Грейнджеру для определения связанности систем по сигналам, имеющим выраженный временной масштаб (период колебаний). Показана степень зависимости работоспособности метода от уровня нелинейности исходных данных, который определялся по старшему ляпуновскому показателю.

В оригинальной работе [1] использовались линейные авторегрессионные модели, но в настоящее время в качестве аппроксимирующих функций часто используют степенные полиномы [11], а также радиальные базисные функции [12]. В данном исследовании рассматривались только модели с полиномиальной нелинейностью общего вида; такие модели достаточно универсальны и в то же время наиболее часто используются на практике вследствие своей простоты. При этом брались небольшие значения размерности модели и порядка полинома, что обусловлено в первую очередь желанием сделать результаты работы полезными с практической точки зрения, когда объем данных сильно ограничен вследствие специфики эксперимента, нестационарности или желания получить приемлемое временное разрешение при построении моделей во временном окне, как это делалось, например, в работе [13]. Такие модели содержат не очень много коэффициентов, которые можно надежно оценить по ряду небольшой длины.

2. Причинность по Грейнджеру

Напомним суть предложенного Грейнджером метода. Пусть имеются записи колебаний двух систем — ряд $\{x_n\}_{n=1}^N$ от системы X и ряд $\{y_n\}_{n=1}^N$ от системы Y , где n ($n = 1, 2, \dots, N$) — дискретное время, N — длина рядов. На основе анализа реализаций $\{x_n\}_{n=1}^N$ и $\{y_n\}_{n=1}^N$, которые в общем случае содержат и шумы, требуется определить, влияет система Y на систему X или нет. На первом шаге строится индивидуальная модель (динамическая система):

$$x'_{n+\tau} = f(x_n, x_{n-l}, \dots, x_{n-(D_s-1)l}, \mathbf{c}^s), \quad (2.1)$$

где x'_n — предсказанное значение в момент времени n и может отличаться от измеренного значения x_n , f — аппроксимирующая функция (если она нелинейна, метод называется нелинейной причинностью по Грейнджеру), l — лаг модели (количество шагов дискретного времени между точками ряда, образующими D_s -мерный вектор состояний модели $\mathbf{x}_n(x_n, x_{n-1}, \dots, x_{n-(D_s-1)l})$, τ — дальность прогноза (расстояние во времени между предсказываемой точкой и наиболее близкой точкой вектора состояния), D_s — размерность индивидуальной модели (число точек временного ряда, составляющих вектор состояния, который реконструируется методом временных задержек [7, 8]), \mathbf{c}^s — неизвестный вектор коэффициентов, который подбирается методом наименьших квадратов так, чтобы минимизировать среднеквадратичную ошибку аппроксимации (2.1):

$$\varepsilon_s^2 = \frac{1}{N} \sum_{n=\tau+(D_s-1)l+1}^N (x'_n - x_n)^2. \quad (2.2)$$

Следующим шагом строится совместная модель, в которой кроме данных из ряда $\{x_n\}_{n=1}^N$ используется D_a членов из ряда $\{y_n\}_{n=1}^N$:

$$x''_{n+\tau} = g(x_n, x_{n-l}, \dots, x_{n-(D_s-1)l}, y_n, y_{n-l}, \dots, y_{n-(D_a-1)l}, \mathbf{c}^j), \quad (2.3)$$

где x''_n есть предсказанное моделью значение, \mathbf{c}^j — коэффициенты совместной модели. Минимизированная среднеквадратичная ошибка прогноза совместной модели, аналогично (2.2), имеет вид

$$\varepsilon_j^2 = \frac{1}{N} \sum_{n=\tau+(\max(D_s, D_a)-1)l+1}^N (x''_n - x_n)^2. \quad (2.4)$$

При $\varepsilon_j^2 < \varepsilon_S^2$ говорят, что Y действует на X (системы связаны).

В качестве меры связанности, как правило, используется показатель *улучшения прогноза*:

$$PI = 1 - \frac{\varepsilon_j^2}{\varepsilon_S^2}. \quad (2.5)$$

Если $PI = 0$ (учет сигнала Y не помог в предсказании X), то считают, что Y не воздействует на X . Если же $PI \rightarrow 1$ (учет сигнала Y существенно улучшил предсказание X), считают, что Y воздействует на X .

Практика показывает, что выбор параметров описанной процедуры (лага l , дальности прогноза τ , размерностей D_s и D_a , вида нелинейных функций f и g) существенным образом определяет работоспособность метода. Например, использование слишком малых или слишком больших значений τ может приводить к большому числу ошибок — положительных выводов о наличии связи, когда она в действительности отсутствует [9, 10]. Пренебрежение нелинейностью при моделировании часто приводит к тому, что реально существующие связи не детектируются [11, 14]. Работоспособности метода можно добиться, разрабатывая специализированные технологии его реализации для выбранного достаточно узкого класса систем, исходя из дополнительных соображений, например, априорной информации о свойствах исследуемой системы, или базируясь на результатах аналитического и численного анализа базовых (эталонных для этого класса) моделей. Как уже упоминалось, мы делаем это для систем, в движении которых присутствуют характерные временные реализации.

3. Методика исследования работоспособности метода

Работоспособность методов детектирования связанности двух систем принято характеризовать их чувствительностью и специфичностью. Чувствительность определяется тем, насколько слабую связь может детектировать метод. Различные факторы (ограниченность наблюдаемых временных рядов, шумы измерений, использование слишком простых базисных функций, недостаточная точность расчетов и т. п.) не позволяют детектировать связь при сколь угодно малых ее уровнях. Под специфичностью понимается способность метода избегать ложноположительных результатов. Хорошая специфичность означает относительно малое число ложных связей (найденных там, где их на самом деле нет), а плохая — большое их число. В типичном случае метод причинности по Грейнджеру (рис. 1) дает ненулевые значения PI в обе стороны даже для однонаправленно связанных систем [15]. Такого рода ошибки часто обусловлены конечностью объема экспериментальных данных

или неудачным выбором параметров метода. Поэтому при расчете PI часто проводят тестирование *значимости* полученных значений — вероятности того, что ненулевые значения показателя улучшения прогноза получены не случайно. Значимость можно оценивать различными способами, например, исходя из соображений о теоретических свойствах остатков модели [9] или на основе генерации различных суррогатов [10, 13, 15].

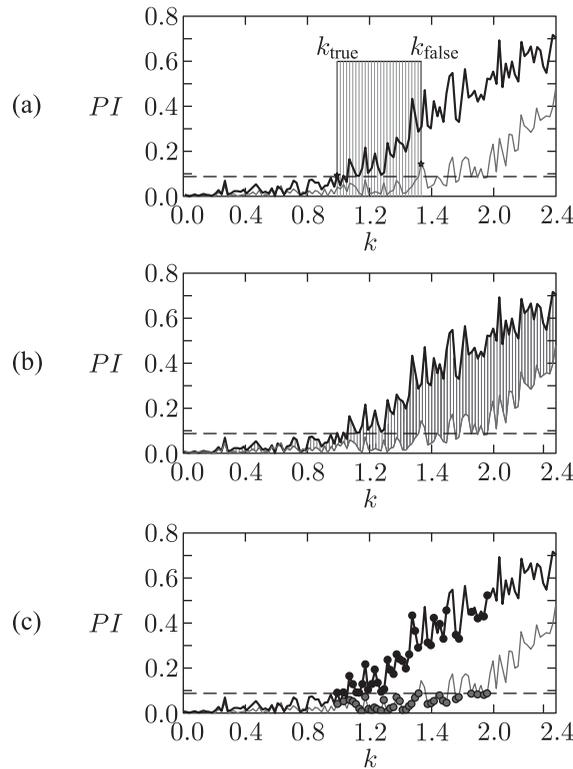


Рис. 1. Зависимость показателя улучшения прогноза PI от коэффициента связи k , когда только одна (возбуждающая) система действует на вторую (ведомую): (а) типичный вид зависимости PI от k , где k_{true} — минимальное значение, при котором связь значимо обнаруживается методом в заведомо верном направлении, а k_{false} — минимальное значение k , при котором связь значимо обнаруживается в неверном направлении (серая линия); (б) проиллюстрирован критерий (3.1) оценки эффективности метода нелинейной причинности по Грейнджеру; (с) проиллюстрирован критерий (3.2), точками отмечены значения, при которых $\delta = 1$.

В работе [10] был предложен подход для оценки эффективности метода причинности по Грейнджеру на системах связанных уравнений, состоящий в определении критического значения коэффициента связанности \bar{k} , при котором значение PI становится значимым. Обозначим как $\bar{k} = k_{true}$ минимальное значение, при котором связь значимо обнаруживается методом в заведомо верном направлении, а $\bar{k} = k_{false}$ — минимальное значение, при котором связь значимо обнаруживается в заведомо ложном направлении. Поскольку зависимость $PI(k)$ обычно носит изрезанный характер (см. рис. 1а) вследствие наличия шумов, различных начальных условий и конечности длины ряда (неравномерный обход аттрактора), часто необходимо усреднение по реализациям.

Однако для заявленной нами цели — исследования зависимости результатов от параметров l и τ — полученные кривые $PI(k)$ нуждаются в дополнительной обработке, поскольку фактически приходится рассматривать зависимость $PI(k, l, \tau)$. Такую зависимость

невозможно представить на графике в интуитивно понятной форме. Поэтому необходимо разработать дополнительные, интегральные критерии. При практическом применении метода величина силы связи неизвестна, а значит, в целях исследования следует рассмотреть максимально широкий диапазон величин k . Нижняя граница такого диапазона соответствует отсутствию связи $k = 0$, верхняя — синхронизации ведущей и ведомой систем, поскольку в случае синхронного поведения выявить направление связи уже невозможно из-за того, что значения двух измеренных временных рядов соотносятся взаимно однозначно, а следовательно, информация о направлении связи потеряна. Для оценки верхней границы значения коэффициента связи $k = k_{\max}$ по методике, предложенной в работе [16], оценивался коэффициент фазовой синхронизации. Это обосновано тем, что в системах с характерным выраженным временным масштабом фазовая синхронизация, как правило, может быть хорошо диагностирована и при этом предшествует другим видам синхронизации [17].

Практика показывает, что при выборе параметров модели приходится соблюдать оптимальный баланс между высокой чувствительностью и высокой специфичностью, поскольку улучшение одного из этих свойств модели, как правило, ведет к ухудшению другого. Для выбора оптимальных значений l и τ были введены два численных критерия, в некотором смысле подобных критерию Шварца [18], позволяющему найти компромисс между (1) точностью описания ряда моделью и (2) размером модели и надежностью оценки коэффициентов.

1. Первый критерий учитывает, насколько в среднем для всех рассмотренных значений k различаются величины PI в заведомо верную (PI_{true}) и в заведомо ложную (PI_{false}) стороны (см. рис. 1b), и определяется формулой

$$S_1 = \langle PI_{\text{true}}(k) - PI_{\text{false}}(k) \rangle_k. \quad (3.1)$$

Величина первого критерия визуально определяется средней шириной заштрихованной области на рисунке 1b. Поскольку согласно (2.5) PI есть нормированная величина и изменяется от 0 до 1, величина S_1 может изменяться на отрезке от -1 до 1. Если для всех значений k $PI_{\text{true}} = 1$ (связи в верную сторону определяются идеально при любой силе связи), а $PI_{\text{false}} = 0$ (связи в ложную сторону всегда отсутствуют), то согласно (3.1) $S_1 = 1$. Если оказывается, что $\langle PI_{\text{true}} \rangle_k = \langle PI_{\text{false}} \rangle_k$ (улучшение прогноза в верную и ложную стороны в среднем одинаково, то есть метод не способен определить направление связи), $S_1 = 0$. Если же $PI_{\text{false}} = 1$ и $PI_{\text{true}} = 0$ (в ложную сторону связи находятся всегда, а в истинную — никогда), $S_1 = -1$.

2. Второй критерий (см. рис. 1c) учитывает значимость полученных значений PI , но не учитывает их абсолютные величины: вводится величина $\delta(k)$; $\delta(k) = 1$, если связь в верную сторону определена как значимая, а связь в неверную — как незначимая, иначе $\delta(k) = 0$. Далее значения $\delta(k)$ усредняются по всем k , итоговая величина S_2 лежит в диапазоне от 0 (нет таких k , при которых связи в верную сторону определяются как значимые, а в неверную — как незначимые) до 1 (для всех k связи в верную сторону определяются как значимые, а в неверную — как незначимые). Таким образом,

$$S_2 = \frac{1}{Z_k} \sum \delta(k), \quad (3.2)$$

$$\delta(k) = \begin{cases} 1, & PI_{\text{true}} > PI_{\text{sign}} \wedge PI_{\text{false}} < PI_{\text{sign}}, \\ 0, & PI_{\text{true}} \leq PI_{\text{sign}} \vee PI_{\text{false}} \geq PI_{\text{sign}}, \end{cases}$$

где Z_k — число рассмотренных значений коэффициента связи, PI_{sign} — это 95 %-ный уровень значимости, полученный при использовании суррогатных временных рядов, полученных от тех же систем, что и исследуемые на связанность ряды, при тех же значениях коэффициентов, но без связи (аналогично [15]).

Критерии (3.1) и (3.2) дополняют друг друга. Если критерий S_1 достигает значения, близкого к 1, можно утверждать, что метод работает хорошо при данных параметрах, а если близок к 0 или отрицателен — метод работает плохо. При этом проверку значимости можно не проводить. Если же критерий S_1 принимает некоторые промежуточные между 0 и 1 значения, то тестирование на значимость существенно, и работоспособность метода определяется по критерию S_2 , согласно чему выбор параметров l и τ тем более оптимален, чем больше S_2 .

4. Численный эксперимент

4.1. Выявление зависимости значений представленных критериев от l и τ при различных значениях старшего ляпуновского показателя

Применение предложенных в разделе 3 критериев (3.1) и (3.2) проиллюстрируем на связанных однонаправленной связью вида $k \cdot y_2$ (значения коэффициента связи k были ограничены 0.5) системах Рёсслера [19]:

$$\begin{cases} \dot{x}_1 = -x_2 - x_3, \\ \dot{x}_2 = x_1 + a_1 x_2 + k y_2, \\ \dot{x}_3 = b_1 - (c_1 - x_1) x_3, \\ \dot{y}_1 = -y_2 - y_3, \\ \dot{y}_2 = y_1 + a_2 y_2, \\ \dot{y}_3 = b_2 - (c_2 - y_1) y_3. \end{cases} \quad (4.1)$$

Здесь и далее:

- параметры эталонных осцилляторов были выбраны таким образом, чтобы получить хаотические режимы для каждой подсистемы как при отсутствии, так и при наличии связи;
- уравнения интегрировались методом Рунге–Кутты 4-го порядка, с шагом $\Delta t = 0.01$;
- отсекался переходной процесс длиной в 1000 единиц безразмерного времени;
- в численном эксперименте использовались ряды длиной 20 480 значений.

В качестве наблюдаемых использовались временные ряды координаты x_2 каждой из подсистем.

Использовано 4 набора параметров¹ для системы Рёсслера, при которых старшие ляпуновские показатели ведомой системы принимали значения 0.6, 0.8, 0.1 и 0.12 при нулевом

¹ $a_1 = 0.146, b_1 = 0.3, c_1 = 10, a_2 = 0.16, b_2 = 0.2, c_2 = 16; a_1 = 0.176, b_1 = 0.19, c_1 = 11, a_2 = 0.172, b_2 = 0.2, c_2 = 10; a_1 = 0.2, b_1 = 0.3, c_1 = 10, a_2 = 0.2, b_2 = 0.15, c_2 = 7.5; a_1 = 0.205, b_1 = 0.05, c_1 = 20, a_2 = 0.20, b_2 = 0.05, c_2 = 11.$

значении коэффициента связи k . Системы Рёсслера при выбранных параметрах имеют более или менее размытый (в зависимости от величины старшего ляпуновского показателя) пик в спектре на частоте ~ 0.18 (см. рис. 2а–д), по которому сложно было определить характерный период T ведомой системы. Однако его можно вычислить по автокорреляционной функции (см. рис. 2е–h) как положение первого локального максимума; для систем с $\Lambda_1 = 0.06$, $\Lambda_1 = 0.08$ и $\Lambda_1 = 0.10$ он равен приблизительно 620 отсчетам (6.2 единиц безразмерного времени), для системы с $\Lambda_1 = 0.12$ — 655 отсчетам. Видно, что с увеличением Λ_1 степень регулярности колебаний в целом уменьшается, что соответствует более быстрому спадаанию корреляции за характерный период.

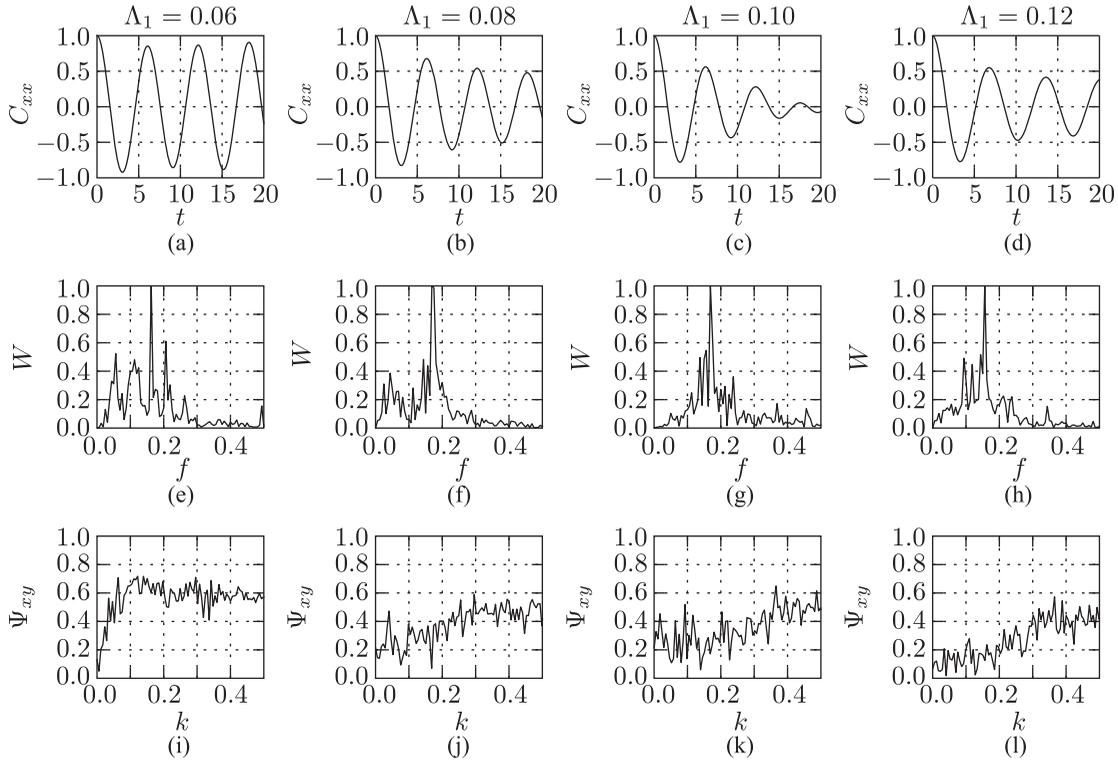


Рис. 2. Автокорреляционные функции (а) и спектры мощности (б) ведомой системы Рёсслера при различных параметрах и величинах старшего ляпуновского показателя Λ_1 .

Зависимости значений (3.1), (3.2) от l при различных значениях τ и Λ_1 представлены на рисунке 3, откуда видно, что при $l = T/2$ наблюдается четкий минимум графика функции для большинства значений τ . Аналогичный минимум наблюдается при $l = T$, хотя он менее выражен. Согласно критерию (3.1), наилучшие результаты метод нелинейной причинности по Грейнджеру демонстрирует при значениях l , лежащих в интервалах от $T/4$ до $T/2$ и от $T/2$ до $3T/4$ при всех τ . По критерию (3.2), учитывающему значимость, такие значения l также являются удачными, но лучшие результаты достигаются при малом (но не единичном) l и τ , близком к половине характерного периода. Те же выводы можно сделать, рассмотрев двумерные зависимости $S_1(\tau, l)$ и $S_2(\tau, l)$, построенные для случая $\Lambda_1 = 0.1$ (рис. 4). При этом заметно, что оба критерия указывают на очень низкую эффективность метода при дальности прогноза $\tau = 1$, особенно при $l < T/4$, что согласуется с выводами, полученными в работе [10]. Низкая работоспособность метода при величинах l ,

кратных $T/2$, скорее всего может объясняться большой корреляцией компонент вектора состояния, вследствие чего они несут недостаточно информации о рассматриваемом сигнале. Эта гипотеза подтверждается тем, что для наборов параметров, соответствующих быстрее спадающей за период колебаний автокорреляционной функции, значения l , кратные $T/2$, оказываются не столь плохими, как и в случае более медленного спада автокорреляции.

Графики, представленные на рисунке 3, позволяют сделать вывод о том, что для всех значений τ наиболее сильно зависимость критериев (3.1) и (3.2) от значения l выражена при малых значениях старшего ляпуновского показателя, то есть чем меньше старший ляпуновский показатель, тем важнее для эффективной работы метода нелинейной причинности по Грейнджеру правильный выбор величин l и τ .

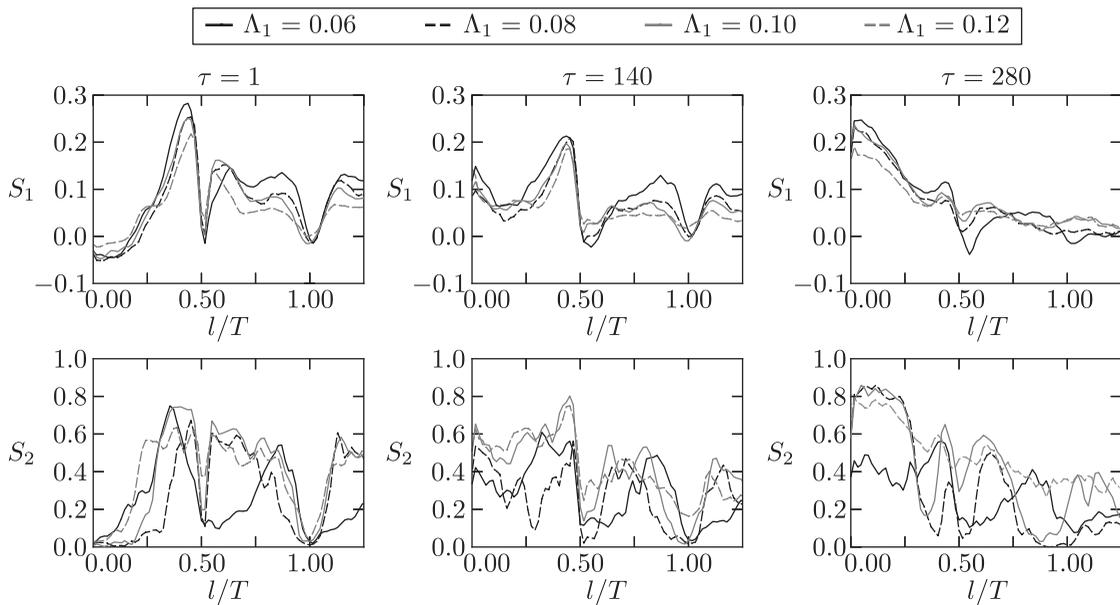


Рис. 3. Зависимости величины критериев (3.1) и (3.2) от лага модели (измеряется в характерных периодах) при различных дальностях прогноза для двух однонаправленно связанных систем Рёсслера (4.1), построенные при использовании моделей с размерностями $D_s = 2$ и $D_a = 1$ и при аппроксимации функций f и g полиномами общего вида.

Следует отметить, что зависимость $S_2(l)$ при $\Lambda_1 = 0.06$ не совсем похожа на аналогичные зависимости при других значениях Λ_1 и лежит существенно ниже для ряда значений l . Подробный анализ зависимостей $PI(k)$ для этого случая показывает, что это может быть обусловлено высокой степенью синхронизации ведомой и ведущей подсистем (коэффициент фазовой синхронизации порядка 0.7 для большинства значений k), вследствие чего используемые для определения значимости суррогаты, построенные на основе несвязанных систем, не позволяют определить направление связи: связь в обе стороны детектируется как значимая (то есть метод причинности по Грейнджеру не проходит тест на специфичность), хотя величина PI в верную сторону больше, чем в неверную.

Аналогичные представленным на рисунке 3 результаты были получены при других сочетаниях размерности индивидуальной модели D_s и степени полинома ν , в частности, при $D_s = 3, \nu = 3$ и $D_s = 2, \nu = 2$ и при более детальном переборе значений τ (с шагом 10 отсчетов, что соответствует примерно $0.016T$).

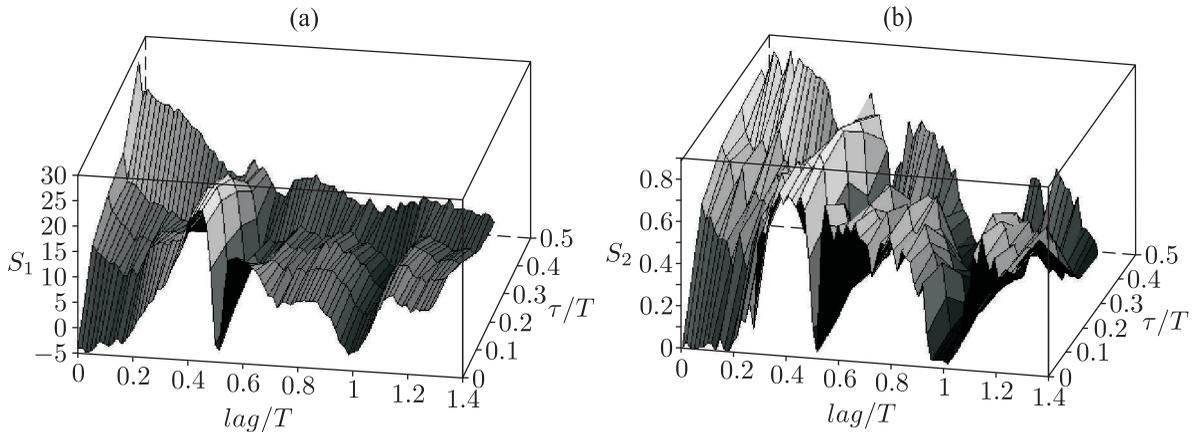


Рис. 4. Зависимости величины критериев (3.1) и (3.2) от значений дальности прогноза τ и лага l (измеряются в характерных периодах) для двух однонаправленно связанных систем Рёсслера (4.1), построенные при использовании моделей с размерностями $D_s = 2$ и $D_a = 1$.

4.2. Выявление зависимости значений представленных критериев от l и τ для различных эталонных систем

Общность сделанных выводов подтверждает численный эксперимент с однонаправленно связанными уравнениями двух генераторов Кияшко–Пиковского–Рабиновича [20].

$$\begin{cases} \dot{x}_1 = 0.15x_1 + x_2 - 0.93x_3 + ky_1, \\ \dot{x}_2 = -x_1, \\ 0.2\dot{x}_3 = x_1 - f(x_3), \\ \dot{y}_1 = 0.23y_1 + y_2 - 0.755y_3, \\ \dot{y}_2 = -y_1, \\ 0.2\dot{y}_3 = y_1 - f(y_3), \end{cases} \quad (4.2)$$

$$f(p) = 8.592p - 22p^2 + 14.408p^3.$$

Характерный период ведомой системы равен примерно 611 точкам (6.11 единиц безразмерного времени). Значения коэффициента связи были ограничены 0.038. Старший ляпуновский показатель для ведомой подсистемы в отсутствие связи составляет $\Lambda_1 = 0.059$, что меньше, чем для системы Рёсслера.

Из рисунка 5 видно, что ранее сформулированные выводы подтверждаются. Во-первых, заметны четкие минимумы зависимостей $S_1(l)$ для l , близких к $T/2$ и T . Во-вторых, для $l < T/4$ работоспособность метода при $\tau = 1$ существенно ниже, чем при других значениях τ . В-третьих, кривые $S_1(l)$, построенные для $\tau \neq 1$, пересекаются примерно в одной точке при значениях l , немного более $T/4$. Значения критерия (3.1) малы в сравнении с 1, поэтому важным также является анализ результатов критерия (3.2), учитывающего значимость. Кривые на графике $S_2(l)$ имеют менее выраженные минимумы при $l = T/2$ и $l = T$, чем в случае предыдущей системы, при этом заметно, что метод в целом неплохо работает для любых l и τ .

Аналогичные результаты были получены с использованием модели с размерностью $D_s = 3$ и полиномом степени $\nu = 2$. Полученные результаты подтверждают сделанный ранее

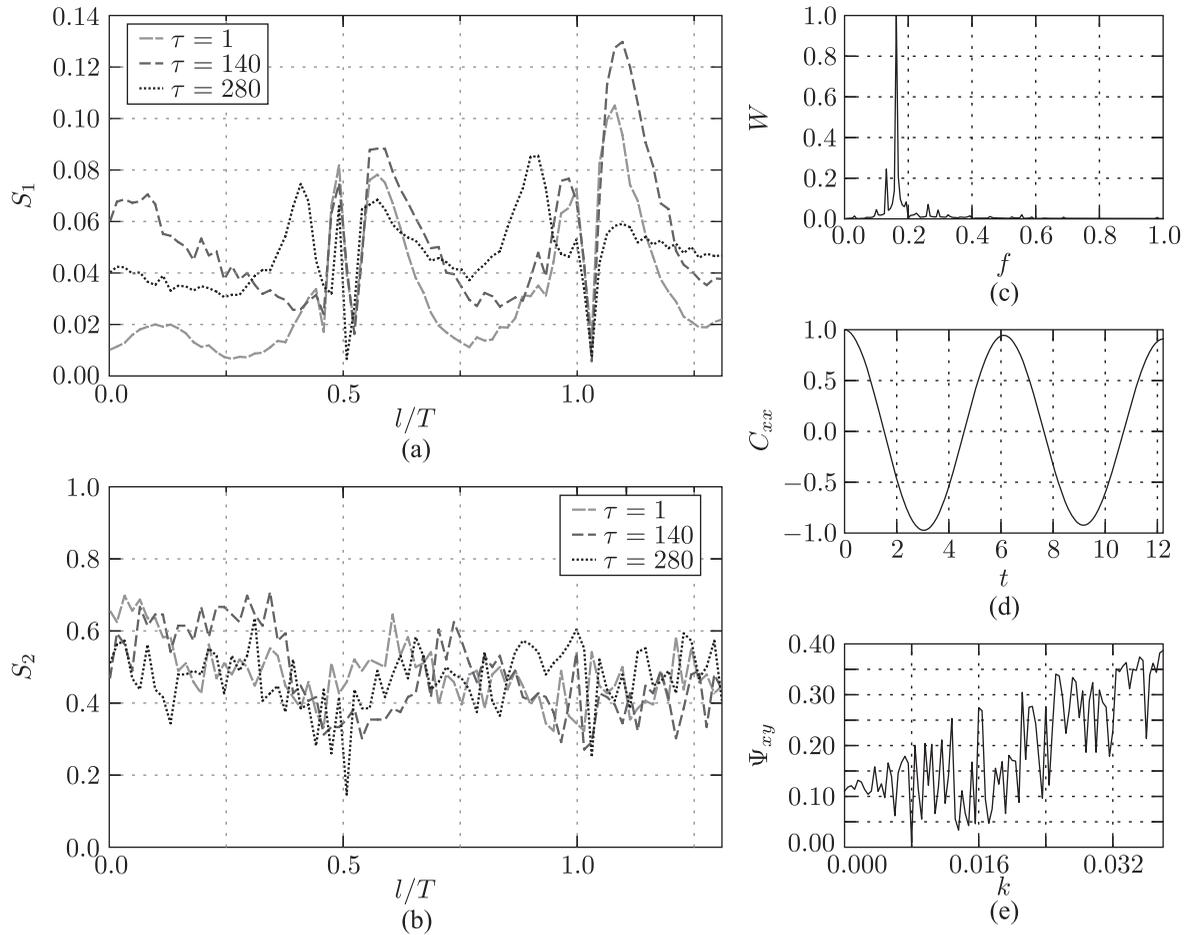


Рис. 5. Зависимости величины критериев (3.1) и (3.2) от лага модели (измеряется в характерных периодах) при различных дальностях прогноза для однонаправленно связанных уравнений генераторов Кияшко – Пиковского – Рабиновича (4.2), построенные при использовании моделей с размерностями $D_s = 2$ и $D_a = 1$ и при аппроксимации функций f и g полиномами общего вида (а) и (б) соответственно, (с) — спектр мощности ведомой системы, (д) — автокорреляционная функция для ведомой системы, (е) — зависимость коэффициента фазовой синхронизации ведущей и ведомой систем Ψ_{xy} от коэффициента связи k .

вывод, что при наличии одного выраженного характерного масштаба значения l , близкие к $T/2$ и T , не следует использовать при определении связанности методом причинности по Грейнджеру.

Перейдем к рассмотрению систем, обладающих большим уровнем нелинейности (старший ляпуновский показатель которых находится в диапазоне от 0.1 до 0.2). Одним из примеров таких систем являются связанные линейной однонаправленной связью уравнения генераторов Анищенко – Астахова [21]:

$$\begin{cases} \dot{x}_1 = 1.106x_1 + x_2 - x_1x_3 + ky_1, \\ \dot{x}_2 = -x_1, \\ \epsilon \dot{x}_3 = -0.68x_3 + \theta(x_1)x_1^2, \end{cases}$$

$$\begin{cases} \dot{y}_1 = 1.111y_1 + y_2 - y_1y_3, \\ \dot{y}_2 = -y_1, \\ \dot{y}_3 = -0.7y_3 + \theta(y_1)y_1^2, \\ \theta(x) = \begin{cases} 0, & \text{при } x \leq 0, \\ 1, & \text{при } x > 0. \end{cases} \end{cases} \quad (4.3)$$

Значения коэффициента связи были ограничены 0.15. Старший ляпуновский показатель ведомой подсистемы $\Lambda_1 = 0.136$.

На рисунке 6с, d представлены спектр мощности и автокорреляционная функция ведомой подсистемы, из графика которой можно видеть, что она имеет один характерный масштаб колебаний, равный примерно 683 отсчетам (6.83 единицы безразмерного времени), несмотря на два пика в спектре мощности. При этом автокорреляционная функция явно асимметрична, в отличие от функций систем Рёсслера и уравнений генератора Кияшко–Пиковского–Рабиновича, которые визуалью мало отличались от затухающей синусоиды (см. рис. 2а–d и рис. 5d).

Результаты применения обоих критериев представлены на графиках (рис. 6а–b). На графике зависимости $S_1(l)$ можно видеть два минимума аналогично тому, что наблюдалось для ранее рассмотренных систем, однако расположены не на величинах $l = T/2$ и $l = T$, а на меньших значениях, положение минимумов различно для различных τ . Также минимумы менее выражены. Это, скорее всего, является следствием наличия второго пика (в спектре на большей частоте), соответствующего меньшему периоду. Эти же выводы подтверждаются зависимостью $S_2(l)$, но для нее минимумы расположены близко к значениям $l = T/2$ и $l = T$, определенным по автокорреляционной функции. Лучшие результаты работы метода причинности по Грейнджеру получаются в диапазоне $T/4 < l < T/3$, что в целом согласуется с результатами, полученными в разделе 4.1. При этом значение дальности прогноза $\tau = 1$ вновь оказывается менее эффективно при малых l , чем большие значения.

Рассмотрим еще одну эталонную систему, которая, как и система (4.3), демонстрирует сложную существенно нелинейную колебательную динамику, — уравнения автогенераторов с полутора степенями свободы [22]:

$$\begin{cases} \dot{x}_1 = (F_1(x_3) - x_1)/3 + ky_3, \\ \dot{x}_2 = x_1 - x_3, \\ \dot{x}_3 = x_2 - 0.21x_3, \end{cases} \quad (4.4)$$

$$\begin{cases} \dot{y}_1 = (F_2(y_3) - y_1)/3, \\ \dot{y}_2 = y_1 - y_3, \\ \dot{y}_3 = y_2 - 0.22y_3, \end{cases} \quad (4.5)$$

$$\begin{aligned} F_1(z) &= 26z \exp(-z^2), \\ F_2(z) &= 28z \exp(-z^2). \end{aligned} \quad (4.6)$$

Значения коэффициента связи k были ограничены 0.2. Старший ляпуновский показателем системы $\Lambda_1 = 0.148$.

В качестве наблюдаемых использовались координаты x_1 и y_1 , как и ранее, но уравнения интегрировались с шагом 0.03, чтобы на один характерный период приходилось примерно столько же измерений, сколько для систем, рассмотренных ранее. По спектру

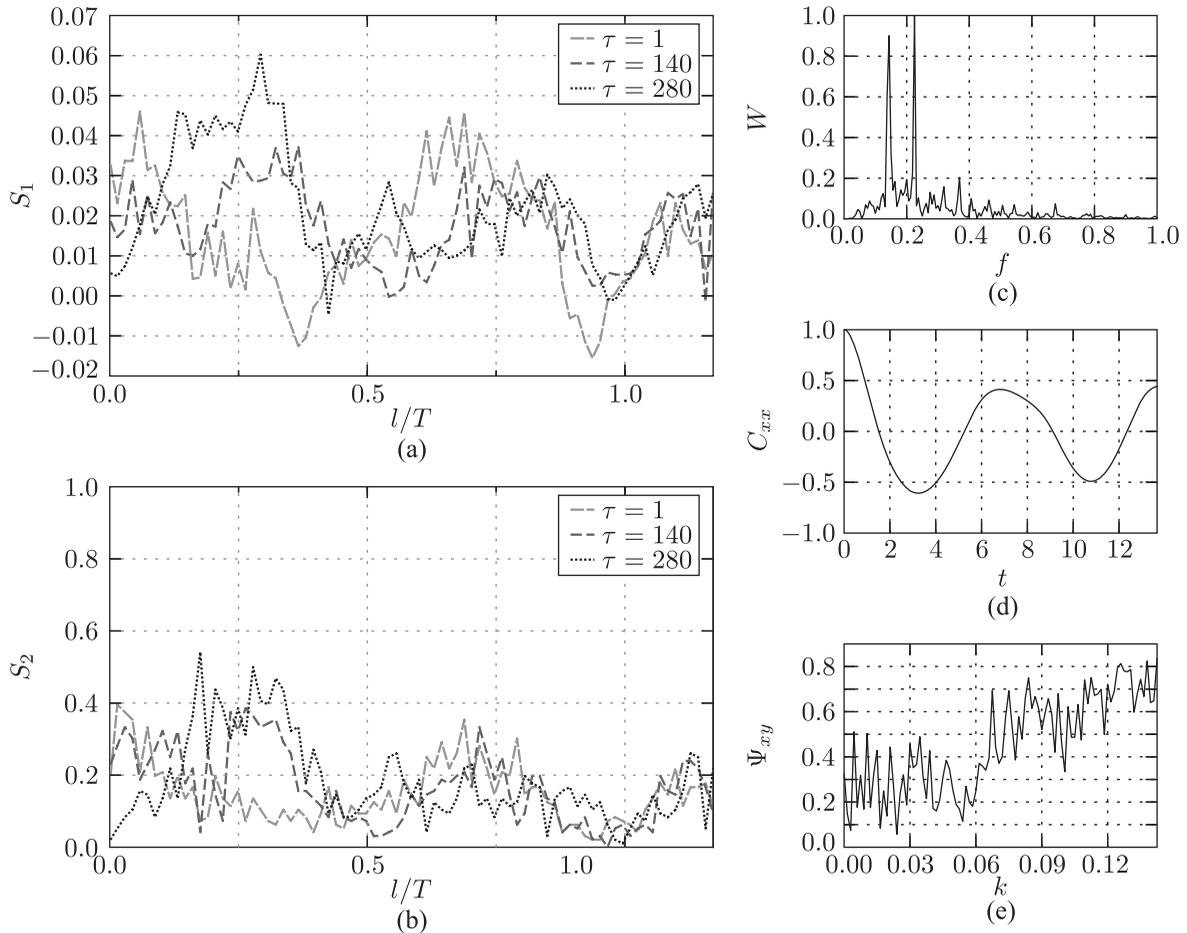


Рис. 6. Зависимости величины критериев (3.1) и (3.2) от лага модели (измеряется в характерных периодах) при различных дальностях прогноза для однонаправленно связанных уравнений генераторов Анищенко – Астахова (4.3), построенные при использовании моделей с размерностями $D_s = 2$ и $D_a = 1$ и при аппроксимации функций f и g полиномами общего вида (а) и (б) соответственно, (с) – спектр мощности ведомой системы, (д) – автокорреляционная функция для ведомой системы, (е) – зависимость коэффициента фазовой синхронизации ведущей и ведомой систем Ψ_{xy} от коэффициента связи k .

мощности (рис. 7с) оценить характерный временной масштаб невозможно, поскольку он имеет несколько пиков. Основной период, определяемый из графика автокорреляционной функции, равен 20.2 единицы безразмерного времени (или примерно 673 точки для перевыбранного ряда).

Низкая автокорреляция непосредственно отражается на зависимости $S_1(l)$, минимумы в которой на величинах порядка $T/2$ и T выражены слабее, чем у рассмотренных ранее систем, при этом только при $\tau = 140$ (немного менее четверти характерного периода) значения критерия $S_1(l)$ стабильно выше 0.06 при всех рассмотренных значениях l , в то время как кривые $S_1(l)$ при $\tau = 1$ и $\tau = 280$ колеблются возле нуля. В целом низкие значения критерия (3.1) дают повод усомниться в работоспособности метода для данной системы при выбранных размерностях и степени аппроксимирующего полинома, но график $S_2(l)$ показывает, что с учетом проверки значимости можно достоверно определить наличие и направленность воздействия более чем в 70 % случаев, если использовать удачное сочетание τ и l .

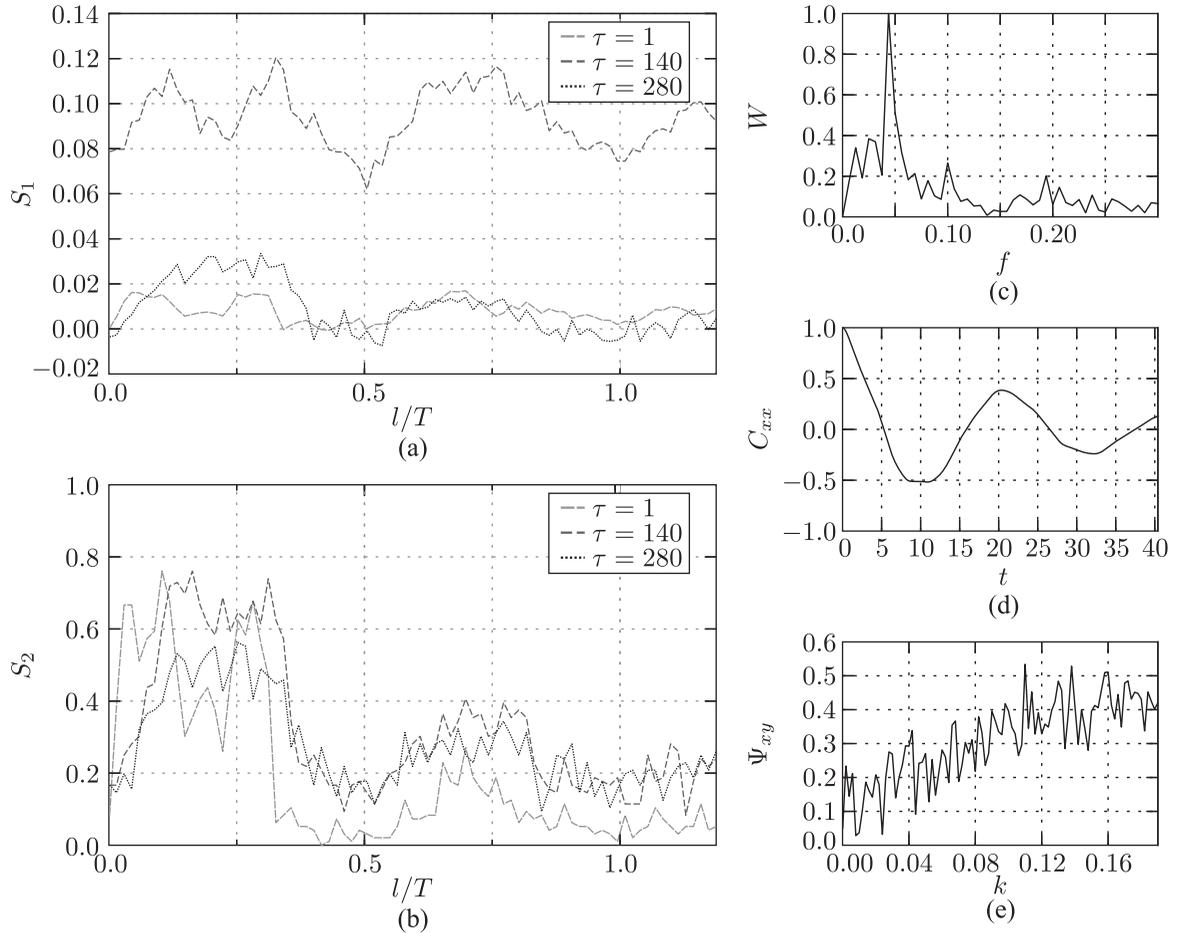


Рис. 7. Зависимости величины критериев (3.1) и (3.2) от лага модели (измеряется в характерных периодах) при различных дальностях прогноза для однонаправленно связанных уравнений генераторов с полутора степенями свободы (4.4), построенные при использовании моделей с размерностями $D_s = 2$ и $D_a = 1$ и при аппроксимации функций f и g полиномами общего вида, шаг выборки составлял 0.03 единицы безразмерного времени, наблюдались координаты x_1 и y_1 (а) и (б) соответственно, (с) — спектр мощности ведомой системы, (д) — автокорреляционная функция для ведомой системы, (е) — зависимость коэффициента фазовой синхронизации ведущей и ведомой систем Ψ_{xy} от коэффициента связи k .

Кроме того, при использовании $\tau \approx T/4$ вероятность верных выводов достигает не менее 60% для всех значений l в интервале от $T/12$ до $T/3$, что в очередной раз подтверждает результаты работы [10]. Аналогичные результаты можно получить и при использовании $\tau = 1$, но в более узком диапазоне значений l .

5. Заключение

Активное использование методов оценки связанности, основанных на построении эмпирических прогностических моделей, при обработке экспериментальных данных напрямую ставит вопрос о надежности таких подходов и зависимости получаемых результатов от параметров этих методов. Для метода причинности по Грейнджеру важными параметрами

являются лаг l , используемый при реконструкции вектора состояния методом временных задержек, и дальность прогноза τ , в особенности для сигналов, имеющих характерный временной масштаб колебаний. В данной работе для такого типа сигналов на основе рассмотрения различных тестовых примеров (эталонных систем нелинейной динамики и радиофизики) сформулирован ряд рекомендаций по выбору l и τ .

Для решения поставленной задачи нами разработаны два интегральных критерия эффективности метода причинности по Грейнджеру, позволяющие на эталонных примерах определять лучшие значения τ и l для широкого диапазона сил связи для однонаправленно связанных систем. Один из критериев базируется на средней разнице между улучшением прогноза в заведомо верную и заведомо ложную сторону и является нормированным, изменяясь от -1 до 1 . Второй критерий учитывает также уровень значимости выводов о связанности. В данной работе значимость рассчитывалась с помощью суррогатных временных рядов, взятых от тех же эталонных систем, но при отсутствии связи на уровне 0.05 (возможен любой другой способ оценки значимости). Выбор диапазона сил связи, при котором будут проводиться исследования, является важным: на практике уровень связи, как и ее наличие, не известны и, зачастую, именно их требуется найти, поэтому при рассмотрении эталонных систем мы использовали весь диапазон связей, при котором не наступает синхронизация, препятствующая определению направления связанности.

При проведении исследования мы ограничили перебор других параметров метода причинности по Грейнджеру, иначе оно оказалось бы слишком объемным. Для повышения надежности в ряде примеров аналогичные расчеты были проведены и для систем с большими размерностями и степенями полинома, при этом получены полностью аналогичные результаты.

По результатам проведенного анализа для класса систем с выраженным временным масштабом колебаний можно сделать следующие выводы:

1. Выбор значения l , равного или кратного половине характерного масштаба колебаний, является неудачным. Это обусловлено высокой корреляцией (большой линейной зависимостью) между выбранными таким образом компонентами вектора состояния, вследствие чего они вместе несут слишком мало (недостаточно) информации об изучаемом временном ряде. Этот вывод является надежным, поскольку продемонстрирован на всех рассмотренных эталонных системах.
2. Зависимость работоспособности метода причинности по Грейнджеру от выбора значений параметров l и τ выражена сильнее (минимумы в зависимости предложенных критериев от l и τ глубже, а максимумы выше) при меньших величинах старшего ляпуновского показателя ведомой подсистемы. Этот вывод сделан на основе рассмотрения как одной и той же системы с различными ляпуновскими показателями (при разных значениях параметров), так и при сопоставлении результатов, полученных для различных эталонных систем.
3. Можно предложить две стратегии выбора оптимальных значений l и τ . В первом случае используется значение лага, меньшее половины характерного периода, например $T/4$ или $T/3$, поскольку для большинства систем в этой области имеется глобальный максимум зависимости предложенных критериев от величины l для $\tau < T/3$. Во втором случае следует ориентироваться на сочетание большой дальности прогноза (порядка $T/2$) и малого лага (например, $l = T/10$). Какая стратегия предпочтительнее — в значительной степени зависит от специфики рассматриваемого объекта.

Представленные рекомендации имеют определенную область применения, связанную с тем, насколько хорошо рассматриваемая система и наблюдаемый от нее временной ряд вписывается в выбранный класс. Наилучшим образом они работают для слабо нелинейных систем — систем с явно выраженным пиком в спектре, у которых автокорреляционная функция спадает медленно (не более, чем вдвое за период) и старший ляпуновский показатель мал (скажем, не превышает 0.1). Поскольку все эти характеристики можно оценить по экспериментальному временному ряду, предложенные критерии несложно проверить на практике. Следует отметить, что и для систем, не полностью подходящих под данное определение (например, при наличии нескольких соизмеримых пиков в спектре или немного большего ляпуновского показателя), выводы в целом остаются справедливыми.

Список литературы

- [1] Granger C. W. J. Investigating causal relations by econometric models and cross-spectral methods // *Econometrica*, 1969, vol. 37, no. 3, pp. 424–438.
- [2] Baccalá L. A., Sameshima K., Ballester G., Do Valle A. C., Timo-Laria C. Studying the interaction between brain structures via directed coherence and Granger causality // *Appl. Signal Process.*, 1998, vol. 5, no. 1, pp. 40–48.
- [3] Tass P., Smirnov D., Karavaev A., Barnikol U., Barnikol T., Adamchic I., Hauptmann C., Pawelczyk N., Maarouf M., Sturm V., Freund H.-J., Bezruchko B. The causal relationship between subcortical local field potential oscillations and Parkinsonian resting tremor // *J. Neural Eng.*, 2010, vol. 7, no. 1, 016009.
- [4] Gourévitch B., Le Bouquin-Jeannès R., Faucon G. Linear and nonlinear causality between signals: Methods, examples and neurophysiological applications // *Biol. Cybern.*, 2006, vol. 95, no. 4, pp. 349–369.
- [5] Мохов И. И., Смирнов Д. А. Эмпирические оценки воздействия естественных и антропогенных факторов на глобальную приповерхностную температуру // *Докл. РАН*, 2009, т. 426, № 5, с. 679–684.
- [6] Яхно Ю. В., Мольков Я. И., Мухин Д. Н., Лоскутов Е. М., Фейгин А. М. Реконструкция оператора эволюции как способ анализа электрической активности мозга при эпилепсии // *Изв. вузов. ПНД*, 2011, т. 19, № 6, с. 156–172.
- [7] Packard N. H., Crutchfield J. P., Farmer J. D., Shaw R. S. Geometry from a time series // *Phys. Rev. Lett.*, 1980, vol. 45, no. 9, pp. 712–716.
- [8] Kougioumtzis D. State space reconstruction parameters in the analysis of chaotic time series: The role of the time window length // *Phys. D*, 1996, vol. 95, no. 1, pp. 13–28.
- [9] Smirnov D., Bezruchko B. Spurious causalities due to low temporal resolution: Towards detection of bidirectional coupling from time series // *Europhys. Lett.*, 2012, vol. 100, no. 1, 10005, 6 pp.
- [10] Сысоева М. В., Диканев Т. В., Сысоев И. В. Выбор временных масштабов при построении эмпирической модели // *Изв. вузов. ПНД*, 2012, т. 20, № 2, с. 54–62.
- [11] Chen Y., Rangarajan G., Feng J., Ding M. Analyzing multiple nonlinear time series with extended Granger causality // *Phys. Lett. A*, 2004, vol. 324, no. 1, pp. 26–35.
- [12] Marinazzo D., Pellicoro M., Stramaglia S. Nonlinear parametric model for Granger causality of time series // *Phys. Rev. E*, 2006, vol. 73, no. 6, 066216, 6 pp.
- [13] Sitnikova E., Dikanav T., Smirnov D., Bezruchko B., van Luijtelaar G. Granger causality: Corticothalamic interdependencies during absence seizures in WAG/Rij rats // *J. Neurosci. Meth.*, 2008, vol. 170, no. 2, pp. 245–254.
- [14] Сысоев И. В., Караваев А. С., Наконечный П. И. Роль нелинейности модели в диагностике связей при патологическом треморе методом грейнджеровской причинности // *Изв. вузов. ПНД*, 2010, т. 18, № 4, с. 81–86.

- [15] Корнилов М. В., Сысоев И. В. Влияние выбора структуры модели на работоспособность метода нелинейной причинности по Грейнджеру // Изв. вузов. ПНД, 2013, т. 21, № 2, с. 74–88.
- [16] Allefeld C., Kurths J. Testing for phase synchronization // Internat. J. Bifur. Chaos Appl. Sci. Engrg., 2004, vol. 14, no. 2, pp. 405–416.
- [17] Пиковский А., Розенблюм М., Куртс Ю. Синхронизация: Фундаментальное нелинейное явление. Москва: Техносфера, 2003. 494 с.
- [18] Schwarz G. Estimating the dimension of a model // Ann. Statist., 1978, vol. 6, no. 2, pp. 461–464.
- [19] Rössler O. E. An equation for continuous chaos // Phys. Lett. A, 1976, vol. 57, no. 5, pp. 397–398.
- [20] Кияшко С. В., Пиковский А. С., Рабинович М. И. Автогенератор радиодиапазона со стохастическим поведением // Радиотехника и электроника, 1980, т. 25, № 2, с. 336–343.
- [21] Анищенко В. С., Астахов В. В. Экспериментальное исследование механизма возникновения и структуры странного аттрактора в генераторе с инерционной нелинейностью // Радиотехника и электроника, 1983, т. 28, № 6, с. 1109–1125.
- [22] Дмитриев А. С., Кислов В. Я. Стохастические колебания в автогенераторе с инерционным запаздыванием первого порядка // Радиотехника и электроника, 1984, т. 29, № 12, с. 2389–2398.

Optimal selection of parameters of the forecasting models used for the nonlinear Granger causality method in application to the signals with a main time scales

Maxim V. Kornilov¹, Ilya V. Sysoev², Boris P. Bezrychko³

^{1,2,3}Saratov State University named after N. G. Chernyshevsky
Astrahanskaya st. 83, Saratov, 410012, Russia

³Saratov Branch, Kotelnikov Institute of Radio-Engineering and Electronics, Russian Academy of Sciences
Zelenaya 38, Saratov, 410019, Russia

¹kornilovmv@gmail.com, ²ivssci@gmail.com, ³bezruchkobp@gmail.com

The detection of coupling presence and direction between various systems using their time series is a common task in many areas of knowledge. One of the approaches used to solve it is nonlinear Granger causality method. It is based on the construction of forecasting models, so its efficiency depends on selection of model parameters. Two parameters are important for modeling signals with a main time scales: lag that is used for state vector reconstruction and prediction length.

In this paper, we propose two criteria for evaluating performance of the method of nonlinear Granger causality. These criteria allow to select lag and prediction length, that provide the best sensitivity and specificity. Sensitivity determines the weakest coupling method can detect, and specificity refers to the ability to avoid false positive results. As a result of the criteria application to several etalon unidirectionally coupled systems, practical recommendations for the selection of the model parameters (lag and prediction length) were formulated.

MSC 2010: 37M10

Keywords: search for coupling, Granger causality, modeling from time series

Received April 21, 2014, accepted August 8, 2014

Citation: *Rus. J. Nonlin. Dyn.*, 2014, vol. 10, no. 3, pp. 279–295 (Russian)