На правах рукописи

ИСТОМИН Илья Александрович

Новые подходы к исследованию временных рядов

специальность: 01.04.02 — теоретическая физика

Автореферат
диссертации на соискание ученой степени
кандидата физико-математических наук

Москва – 2006

Работа выполнена на физическом факультете Московского государственного университета им. М.В.Ломоносова

Научный руководитель
доктор физико-математических наук, профессор
Александр Юьевич Лоскутов

Официальные оппоненты:
доктор физико-математических наук, профессор
Генри Эдгарович Норман
доктор физико-математических наук, профессор
Андрей Геннадьевич Попов

Ведущая организация
Институт космических исследований РАН

Защита состоится ".......".............2006 года в .... часов на заседании диссертационного совета К.501.001.17 физического факультета МГУ им. М.В.Ломоносова по адресу: 119992, Москва, Ленинские Горы, МГУ, физический факультет, ауд. ....

С диссертацией можно ознакомиться в библиотеке физического факультета МГУ им. М.В.Ломоносова.

Автореферат разослан ".......".....................2006 года.

Ученый секретарь
диссертационного совета К.501.001.17
доктор физико-математических наук,
профессор П.А.Поляков

Общая характеристика работы

Актуальность темы

Считается, что основной задачей естественных наук является описание некоторого явления на основе наблюдений. Но такой подход не является единственным. Уже в средние века, например, главной задачей исследователей было не описание настоящего, а предсказание будущего. Считалось, что настоящие может видеть каждый и понять настоящее могут многие; другое дело будущее ... В последующем стали считать, что задача прогнозирования не может быть решена без построения точной модели. Тем не менее, это не так.

В настоящее время задача прогноза и анализа сложных явлений вновь выходит на первый план. Это связано в том числе с быстрым развитием как новых математических методов, так и вычислительной техники. Как можно представить явление для прогнозирования и анализа? Вполне достаточным здесь оказывается использование понятия временного ряда. Временным рядом можно назвать любую последовательность чисел, полученную в результате измерения через определенные промежутки времени какой-либо величины, характеризующей рассматриваемый процесс.

Задачи анализа и прогнозирования какого-либо явления могут быть сведены, в рамках нелинейной динамики, к исследованию и продолжению временного ряда на основе уже имеющейся его части. Отличительной особенностью методов, основанных на теории динамических систем, является то, что эти методы не требуют явным образом строить модель системы, породившей временной ряд. Тем не менее, задача построения модели по временному ряду — одна из основных, которая имеет непосредственную связь с проблемой прогноза.

Конечно же, задача прогнозирования в своей классической постановке уже давно и достаточно детально рассмотрена в рамках математической статистики и не только. Давно известны алгоритмы типа авторегрессии, которые с успехом используются как для целей различных исследований так и в практическом плане, например в финансовом анализе или в метерологии. Однако, во-первых, полученные там результаты относятся чаще к линейным моделям, а в настоящее же время основной интерес для исследования представляют модели сложных нелинейных процессов. Во-вторых, и авторегрессионные методы прогнозирования, и развившиеся в последнее время нейросетевые способы не имели достаточно строгого теоретического обоснования до появления теории динамических систем. Но самое главное заключается не в обосновании, а в том, что нелинейная динамика указала пути совершенствования старых и развития новых способов прогнозирования временных рядов, наложив определенные, хотя и минимальные, ограничения на функцию связи прогнозируемого и предыдущих значений ряда. В рамках теории динамических систем, в качестве таких моделей-кандидатов используется наиболее широкий класс всевозможных дифференцируемых динамических систем, так что свойства будущей модели ограничены минимально.

Таким образом, в конце 80-х годах возникло новое направление в анализе временных рядов, связанное с использованием идей нелинейной динамики. Эти подходы применялись с тех пор и по настоящее время к широкому спектру проблем, однако в большом числе случаев результаты их использования были неоднозначны. Поэтому в последние годы стали появляться работы, в которых отмечались ограничения методов прогнозирования. То есть, возникло противоречие между сравнительно простыми, ясными и привлекательными идеями, которые лежат в основе подхода нелинейной динамики к прогнозированию временных рядов, и трудностями, связанными с получением конкретных численных результатов прогноза, особенно для систем естественного происхождения. Иными словами, осталось большое количество открытых вопросов. Решить некоторых из них и явилось целью данной диссертационной работы.

В диссертации рассматриваются два принципиально различных метода анализа и прогнозирования временных рядов, объединенных общими принципами нелинейной динамики: метод сингулярного спектрального анализа (ССА) и группа методов под названием локальная аппроксимация (ЛА). Большинство *удовлетворительных* результатов применения методов нелинейной динамики, к сожалению, относится к модельным временным рядам, порожденными системами небольшой размерности. Поэтому в качестве объекта исследования были выбраны временные ряды естественного происхождения, характеризующие солнечную активность и ряд данных по температуре у поверхности Земли. Применение методов нелинейной динамики к этим временным рядам позволило, помимо прогноза, подтвердить некоторые известные идеи и выдвинуть новые гипотезы.

Другой важной задачей стала систематизация и обобщение большого количества вариантов метода локальной аппроксимации, для чего была построена обобщенная теория данного метода прогнозирования. На основании построенной обобщающей теоретической модели были сделаны некоторые выводы о предпочтительности использования того или иного варианта метода локальной аппроксимации.

Цели работы

- Применение методов нелинейной динамики для анализа сложных систем естественного происхождения.

- Построение общего решения задачи прогноза в рамках теории динамических систем.
- Разработка метода, позволяющего находить как дальние корреляции, так и скрытые закономерности во временных рядах естественного происхождения.

Научная новизна

1. В рамках теории динамических систем разработаны новые методы обработки временных рядов естественного происхождения.
2. Предложено новое обоснование гипотезы о существовании 80-летнего цикла активности Солнца.
3. Обоснован новый способ, позволяющий выявить скрытые взаимосвязи между различными системами естественного происхождения.
4. Предложено новое обоснование гипотезы о связи глобального потепления с активностью Солнца.
5. Построена обобщенная теория, позволяющая найти аналитическое решение задачи прогноза для различных вариантов метода локальной аппроксимации.
6. На основании предложенной обобщенной теории даны оценки качества прогноза.

Практическая ценность

1. Предложен новый подход к исследованию возможности прогноза временных рядов естественного происхождения.
2. Разработаны алгоритмы, позволяющие выбирать оптимальный метод прогноза в зависимости от характера и особенностей исследуемого временного ряда.

3. Предложен новый способ выявления дальних корреляций в системах различного происхождения.

Публикации

По результатам диссертационной работы опубликовано 7 работ, в том числе 4 статьи в рецензируемых журналах. Перечень публикаций приведен в заключительной части автореферата. Работа докладывалась на многочисленных научных семинарах и конференциях, среди которых Европейский астрономический конгресс JENAM (Москва, 26 мая – 3 июня 2000г.), международный конгресс по приложениям нелинейных явлений (Shanghai, China, 2005) и второй Европейский форум по нелинейности в естественных науках (Heraklion, Crete, Greece, 2005).

Содержание работы

Диссертация (123 страницы) включает введение, четыре главы, заключение и список литературы (90 ссылок).

Во введении рассмотрена история вопроса, обоснована актуальность проведенных исследований, формулируются цели, дана краткая характеристика основных полученных результатов и их научная новизна. Также приведено краткое содержание диссертации.

Первая глава представляет собой обзор литературы. Здесь рассматриваются базовые понятия математической и реальной динамических систем, временного ряда как среза такой системы и т.п. Здесь же приведен так называемый *метод полной реконструкции.* Данный подход является своего рода классическим при прогнозировании временных рядов в рамках нелинейной динамики. Метод позволяет во многих случаях предсказать дальнейшее поведение системы методом реконструкции соответствующих уравнений.

Задача прогнозирования временных рядов может быть решена методом полной реконструкции. Но при ее решении применительно к временным рядам естественного происхождения почти всегда возникают непреодолимые трудности, вызванные чрезвычайной сложностью процессов, порождающих эти ряды. Таким образом, к сожалению, успешное применение методов прогнозирования, связанных с полным восстановлением динамики системы, на реальных временных рядах является редким исключением.

Тем не менее, в силу своей практической значимости, задача прогноза временных рядов давно решалась с переменным успехом методами, далекими как от реконструкции динамической системы, так и от нелинейной динамики вообще. Все такие методы прогнозирования временных рядов основаны на выражении предсказываемого значения ряда через предыдущие его значения. Иными словами вначале интуитивным образом предлагается некоторая (часто линейная) модель с набором параметров. При помощи этой модели пытаются связать очередное значение ряда с некоторым количеством предыдущих значений. Потом прогоняют эту модель через имеющийся ряд данных и оценивают разницу между предсказываемыми моделью и реальными значениями ряда. Затем минимизируют обобщенное отклонение модели от реального поведения ряда и получают параметры модели. Далее, имея модель с уточненными параметрами, можно уже попытаться спрогнозировать неизвестное значение ряда.

К вышеупомянутым методам в первую очередь следует отнести набор приемов, связанных с *авторегрессией* и нейросетевыми способами прогнозирования. Обсуждение этих способов прогнозирования также находит свое отражение во второй части первой главы.

Здесь же описываются базовые идеи прогнозирования временных рядов в рамках нелинейной динамики. Эти методы предусматривают переход от анализа требующего прогнозирования имеющегося одномерного временного ряда

$$x_t = x_1, x_2, \ldots$$

к анализу и прогнозированию многомерного ряда векторов

$$\overline{X}_t = (x_t, x_{t+1}, x_{t+2}, \ldots, x_{t+\tau-1}).$$

В таком ряде каждый вектор $\overline{X}_t$ образуется из определенного числа τ последовательных (т.е. взятых с единичным шагом) значений исходного одномерного временного ряда.

Далее приводится сопоставление авторегрессионных способов прогнозирования со способами прогнозирования в рамках нелинейной динамики.

Заканчивается первая глава представлением двух методов прогнозирования: ССА и ЛА. Первый является глобальным методом, опосредовано использующим в выражении прогноза все имеющиеся в распоряжении значения исходного временного ряда, а второй основан на поиске схожих локальных подобластей в исходном ряде.

Вторая глава подробно описывает используемые далее методы анализа и прогноза временных рядов: метод сингулярного спектрального анализа и метод локальной аппроксимации. Эти описания и введенная в них терминология необходимы для изложения материала в последующих главах.

Главный элемент всех алгоритмов прогнозирования и анализа временных рядов в рамках нелинейной динамики — метод запаздываний — подразумевает переход от рассмотрения исходного одномерного временного ряда к анализу многомерного ряда (ряда векторов), в котором каждый вектор образуется из определенного числа τ последовательных значений исходного одномерного временного ряда. Полученная матрица представляет собой последовательность частей-векторов исходного ряда: каждая такая часть (вектор) представляет собой столбец высотой τ, а первый элемент (координата) каждого последующего вектора на один шаг сдвинут относительно первого элемента (координаты) предыдущего вектора:

$$X = \begin{pmatrix} x_1 & x_2 & x_3 & \dots & x_\tau & \dots & x_n \\ x_2 & x_3 & x_4 & \dots & x_{\tau+1} & \dots & x_{n+1} \\ x_3 & x_4 & x_5 & \dots & x_{\tau+2} & \dots & x_{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_\tau & x_{\tau+1} & x_{\tau+2} & \dots & x_{2\tau-1} & \dots & x_N \end{pmatrix}$$

Здесь $x_1, x_2, \dots$ — значения ряда в моменты времени t=1, 2 и т.д. Каждый столбец — вектор в τ-мерном пространстве запаздываний; последовательность таких векторов задает матрицу наблюдений $X_{\tau\times(N-\tau+1)}$, где N — число элементов исходного ряда. Эта матрица, в каждом столбце которой стоят "куски"одного и того же ряда, сдвинутые друг относительно друга, и будет многомерным представлением исходного скалярного ряда в пространстве запаздываний.

Для метода ССА следующим шагом будет обработка матрицы X по методу главных компонент.

Целью применения этого метода является снижение размерности имеющегося пространства запаздываний и переход к новым, информативно более обоснованным переменным. Такие новые переменные называют главными компонентами (ГК). Отличительной особенностью применения метода главных компонент (МГК) в алгоритме метода ССА является то, что он здесь используется для одновременной обработки всей матрицы X.

В МГК переход к главным компонентам осуществляется через ортогональное линейное преобразование: производится разложение многомерного ряда X размерности τ по ортогональному базису такой же размерности, где каждый следующий базисный вектор строится вдоль (оставшегося) направления максимальной дисперсии. При последующем отборе ГК мерой информативности считается величина дисперсии.

Для перехода к главным компонентам для матрицы X строится корреляционная матрица, которая затем раскладывается на собственные значения и собственные векторы:

$$C = \frac{1}{n} X X^T = V \Lambda V^T.$$

Здесь Λ — диагональная матрица собственных чисел, а $V = \left(V^1, V^2, \ldots, V^\tau\right)$ — ортогональная матрица собственных векторов, которую представляют как матрицу перехода к главным компонентам

$$Y = V^T X = (Y_1, Y_2, \ldots, Y_\tau)$$

исходного ряда. При этом собственные значения $\lambda_1, \lambda_2, \ldots, \lambda_\tau$ можно рассматривать как вклад главных компонент Y_1, Y_2, ..., Y_τ в общее информационное содержание исходного временного ряда. По полученным главным компонентам можно полностью восстановить исходную матрицу X: $X = \sum_{i=1}^{\tau} V^i Y_i$, а по матрице X, в свою очередь, можно восстановить исходный временной ряд.

Для восстановления временного ряда в алгоритме ССА используются не все главные компоненты $Y_1, Y_2, \ldots, Y_\tau$, а лишь их часть $r < \tau$, существенная с точки зрения информационной содержательности.

Прогнозирование по методу основывается на том, что прогнозирование ряда на один шаг по времени вперед эквивалентно построению нового вектора в пространстве запаздываний. У такого вектора будет единственная неизвестная координата, а остальные будут известны из построения. Эта единственная неизвестная координата находится из условия минимизации проекции этого нового вектора на r-мерную гиперплоскость. Таким образом, задача прогнозирования сводится к задаче минимизации по одной переменной. Эта задача решается аналитически, что и даёт искомое следующее значение ряда.

Для метода ЛА следующим шагом после построения матрицы X будет выбор локального представление: конкретного вида функции, связывающей следующее значения ряда с предыдущими. Для наиболее распространенного варианта — линейной аппроксимации — вид локального представления будет таким:

$$x_{t+1} = a_0 + \mathbf{x_t^T a}.$$

Здесь $\mathbf{x}$ — столбец матрицы X, а параметры задаются через a_0 и вектор a. Для аппроксимации нулевого порядка $x_{t+1} = a_0$.

Далее производится отбор соседей, под которыми подразумеваются векторы $\mathbf{x}$, ближайшие в евклидовой метрике к последнему известному вектору в пространстве запаздываний. Этот последний вектор, после которого начнется прогноз, называют стартовым.

После того как соседи отобраны, по ним производится оценка параметров выбранного представления, после чего уже можно построить прогноз следующего значения ряда.

Для прогноза на несколько шагов обычно используются один из следующих способов способов: итеративный, итеративный с пересчетом и прямой прогноз.

При *итеративном* способе после прогноза на один шаг вперед полученное значение добавляется к исходному ряду, а затем выбранная модель локального представления применяется ещё раз с параметрами $(a_0, \mathbf{a})$, найденными ещё на первом шаге. Затем процедура рекуррентно повторяется. При *итеративном способа с пересчетом* параметры представления на каждом шаге оцениваются заново. При *прямом прогнозе* и стартовый вектор, и все его соседи, в отличие от двух предыдущих методов, остаются неизменными, зато параметры представления на каждом шаге находятся заново за счет анализа эволюции каждого соседа на такое количество шагов, на которое нацелен данный шаг прогноза.

Преимуществом локального метода ЛА можно считать более достоверный краткосрочный прогноз, в то время как ССА позволяет работать с нестационарными рядами как в плане относительно долгосрочного оценочного прогнозирования, так и с точки зрения выявления определенных скрытых периодичностей. Возможность выделения и анализа информационно обусловленных составляющих (не периодических) также рассмотрена в этой главе.

Третья глава посвящена применению методов прогнозирования к системам естественного происхождения. Вначале главы выявлены закономерности и различные периодичности в динамике солнечного излучения, а также дан прогноз солнечной активности на ближайшие годы. В качестве методов исследования используется методы ССА и ЛА. Как показано в диссертации, первый метод даёт высокую достоверность предсказания амплитуды 11-летнего солнечного цикла и пригоден для выявления более продолжительных циклов, а второй метод позволяет точно прогнозировать поведение солнечной активности на краткосрочную перспективу. Во второй части главы рассмотрен нестандартный подход к использованию глобальных методов прогнозирования на примере метода ССА и проблемы корреляции между глобальным потеплением на Земле и активностью Солнца.

Достаточно давно было замечено, что активность Солнца зависит от количества пятен, видимых на его диске. Это количество сильно меняется в течение 11-и лет, называемых солнечным циклом. Сопутствующее этому изменение в структуре магнитных полей Солнца опосредованно влияет на климат Земли, имеет вероятную связь с природными катастрофами, а также часто сказывается на других сферах жизни.

Использовать количество солнечных пятен в качестве меры солнечной активности было предложено Вольфом в 1848 году. Для этой цели он предложил рассмотреть сумму из общего количества видимых на солнечном диске пятен и десятикратного числа регионов, в которых эти пятна располагались. Последнее слагаемое призвано согласовать результаты измерений, проведенных при различных условиях. Позже были реконструированы среднемесячные значения чисел Вольфа с 1749 года и среднегодовые значения, начиная с 1700 года (эти ряды и используется в настоящей работе).

В диссертации рассматривается применение метода ССА к ряду среднемесячной солнечной активности. При этом используются не характерные высокие размерности: например размерность исследуемого многомерного ряда составила $\tau = 500$ составляющих, а размерность, оставленная для реконструкции, составила $r \approx 200$. В результате были получены как удовлетворительные прогнозы солнечной активности, так и выделены различные компоненты, которые отвечают за различные квазипериодические компоненты ряда.

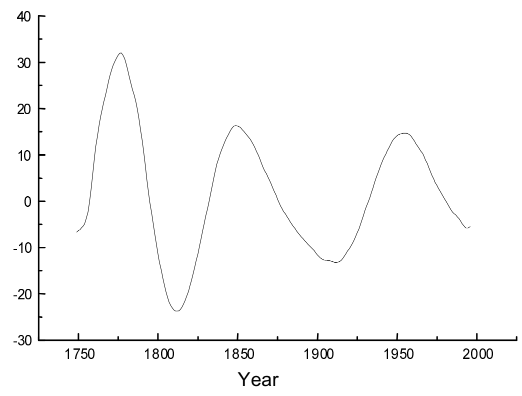

Рис. 1: Реконструкция предполагаемого 80-летнего цикла

Далее в работе рассматривается ряд среднегодовых чисел Вольфа. Помимо очевидного 11-летнего цикла, в процессе анализа были выделены компоненты, отвечающие и за предполагаемый 80-летний цикл активности Солнца. Ряд чисел Вольфа, восстановленный только по этим компонентам, показан на рис.1.

Затем в работе рассмотрена возможность использования метода ССА для прогнозирования ряда из среднегодовых чисел Вольфа. Для этого был взят ряд среднегодовых значений чисел Вольфа по 2006 год. Данный среднегодовой ряд заканчивается на минимуме активности Солнца, и представляется интересным описать два его последующих цикла. Продолжим среднегодовую последовательность на 24 точки вплоть до 2030 года. Ряд раскладывался на 33 компоненты и для прогнозирования отбирались первые 11 из них. Эти два значение были отобраны тотальным перебором в поисках лучшего качества прогноза по известной части ряда.

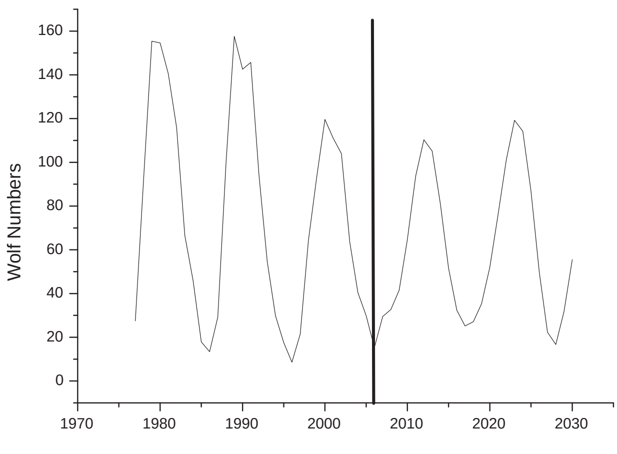

Рис. 2: Прогноз активности Солнца с 2007 по 2030 годы

Результат прогноза с 2007 до 2030 года показан на рис.2. Вертикальной чертой отделены реальные настоящие данные от прогноза на будущее. Из этого рисунка можно видеть, что в ближайшем будущем Солнце будет пребывать в относительно спокойном (по сравнению с двумя предыдущими периодами) состоянии, а уровень очередного максимума в районе 2012 года будет сравнительно низким.

Во втором разделе третьей главы рассматривается упоминавшийся выше нестандартный подход к корреляции между различными системами на примере активности Солнца и ряда данных по температуре воздуха у поверхности Земли.

Исходным материалом для составления ряда данных по среднегодовой температуре Земли являются данные о среднемесячной температуре, полученные от более чем трех тысяч метеостанций на континентах. Покрытие земной поверхности этими метеостанциями не является равномерным как по географии, так и по времени. Так что трудности, связанные с формированием ряда данных по температуре, требуют отдельного рассмотрения, чему и посвящена соответствующая часть третьей главы. Более подробно эти вопросы рассмотрены в приведенных в диссертации ссылках.

Итак, рассматриваются два ряда: ряд со среднегодовыми значениями числа Вольфа и ряд данных по среднегодовой температуре у поверхности земли. Данные наблюдений за глобальной среднегодовой температурой имеются с 1856 года. Для значений солнечной активности аналогичные данные имеются с 1700 года. Первый ряд содержит 148 точки, а второй — 304. Вначале в работе приводится простейший корреляционный анализ, показывающий неоднозначность определения корреляции между этими рядами. Далее проводится перебор параметров прогнозирования методом ССА (размерность и количество оставленных для реконструкции главных компонент). В результате анализа обнаруживается, что корреляция между качеством прогноза и стартовой точкой для ряда среднегодовых температур была аналогична ряду активности Солнца: после локального минимума и на подъеме качество прогноза лучше, чем на максимуме или на спаде. Также было обнаружено, что наилучший прогноз для обоих рядов получается при одинаковых значениях как размерности вложения τ, так и количества оставленных для реконструкции главных компонент r.

Таким образ, можно говорить о практической идентичности параметров прогнозирования исследованных рядов методом ССА. Идентичность параметра τ говорит о близости размерности вложения аттракторов соответствующих динамических систем. Тот факт, что при этом для наилучшего прогнозирования следует оставлять одинаковое количество главных компонент ряда, может говорить об одинаковой размерности самих аттракторов. Следовательно, в силу близости параметров динамики систем можно допустить, что имеет место возможная корреляция между процессами, происходящими на Солнце, и глобальным потеплением на Земле.

Четвертая глава посвящена построению и анализу общей модели метода локальной аппроксимации. Данная модель дает аналитическое решение задачи прогноза методом ЛА для любого варианта этого метода. Анализ этого решения позволяет сделать ряд качественных выводов и рекомендаций о предпочтительности того или иного варианта метода ЛА. Кроме того, предложенная модель может быть полезна сама по себе для дальнейшего аналитического исследования метода ЛА.

Таким образом, в четвертой главе показано, что локальное представление модели ЛА любого порядка может быть записано в виде:

$$x_{t+1} = a_0 + \mathbf{x}_t^T\,\mathbf{a},$$

где $\mathbf{a}$ представляет собой обобщенный вектор параметров, а $\mathbf{x}$ представляет собой обобщенный вектор в пространстве запаздываний.

$$\mathbf{a}^T = \left(\, a_1 \;\; a_2 \;\; \ldots \;\; a_\tau \;\; a_{11} \;\; a_{12} \;\; \ldots \;\; a_{111} \;\; a_{112} \;\; \ldots \, \right),$$

$$\mathbf{x}_t^T = \left(\, x_t \;\; x_{t-1} \;\; \ldots \;\; x_{t-\tau+1} \;\; x_t^2 \;\; x_t \cdot x_{t-1} \;\; \ldots \;\; x_t^3 \;\; x_t^2 \cdot x_{t-1} \;\; \ldots \, \right).$$

Введение обобщенных векторов позволяет учесть все порядки аппроксимации метода ЛА.

Параметры, входящие в состав обобщенного вектора $\mathbf{a}$, находятся из решения задачи минимизации системы уравнений $\mathbf{Y} = \mathbf{I}\, a_0 + \mathbf{X}\, \mathbf{a}$:

$$\Phi = (\mathbf{Y} - \mathbf{I}\, a_0 - \mathbf{X}\, \mathbf{a})^T \; (\mathbf{Y} - \mathbf{I}\, a_0 - \mathbf{X}\, \mathbf{a}) \to \underset{\{a_0,\, \mathbf{a}\}}{min} \, .$$

Здесь обобщенная матрица соседей обозначена как $\mathbf{X}$, искомые коэффициенты как $(a_0, \mathbf{a})$, $\mathbf{I}$ — единичный вектор, у которого количество компонент равно числу соседей (под соседями подразумеваются отобранные по критерию близости в евклидовой метрике части многомерного ряда запаздываний), а $\mathbf{Y}$ — вектор значений ряда, в которые переходят соседи стартового вектора за определенное количество шагов. Это количество шагов прогнозирования далее обозначим через T, а последний вектор многомерного ряда, после которого начинается прогноз, обозначим индексом L. После того, как параметры модели будут найдены, общий вид решения задачи прогноза по методу ЛА запишется так:

$$\hat{x}_{L+T} = \bar{Y} + \left(x_L^T - \bar{\mathbf{X}} \right) \, \hat{\mathbf{a}}.$$

Здесь $\bar{Y}$ — среднее значение для $\mathbf{Y}$ или, что то же самое, прогноз по ЛА0, и $\bar{\mathbf{X}}$ — вектор средних значений координат столбцов в матрице соседей $\mathbf{X}$, который мы далее будем называть *усредненным* соседом. Последнее выражение показывает, что прогноз полученный методом ЛА, есть сумма прогноза метода ЛА нулевого порядка и линейной комбинации отклонения стартового вектора от *усредненного* соседа.

Далее в четвертой главе на основании полученного общего аналитического решения строятся модельные уравнения прогноза для различных способов прогнозирования на несколько шагов вперед: для итеративного прогноза, для итеративного способа с пересчетом параметров и для прямого прогноза.

В результате показано, что итеративный вариант нулевого порядка дает аппроксимацию нулевого порядка, в то время как при прямом прогнозе нулевого порядка уже получается аппроксимация первого порядка, а в частном случае равномерного распределении векторов-соседей вокруг стартового вектора можно даже получить второй порядок аппроксимации.

Далее в работе исследуется асимптотические свойства прогноза при итеративном и прямом способах прогноза на несколько шагов вперед (итеративный с пересчетом параметров не рассматривался в силу невозможности его аналитического выражения). В результате была показана предпочтительность прямого варианта.

Наконец, в последнем разделе четвертой главы исследуется влияние точности вычислений на итеративный и прямой прогноз, результатом чего опять может служить вывод о предпочтительности именно прямого прогноза.

Таким образом, исходя из построенной общей модели локальной аппроксимации, можно заключить, что универсальным вариантом для прогнозирования временных рядов методом локальной аппроксимации является прямой прогноз малого порядка ЛА0 или ЛА1. При этом, если исходный ряд длинный и стационарный, то лучше использовать ЛА0, а если найденные соседи неравномерно распределены вокруг стартового вектора в пространстве запаздываний, то можно попробовать использовать метод ЛА1 в прямом варианте.

Защищаемые положения

1. Обоснована возможность применения теории динамических систем и Такенса–Мане для анализа временных рядов естественного происхождения с использованием высокой размерности представления.

2. Используя предложенный подход, показана возможность выявления скрытых закономерностей в реальных временных рядах: долгопериодических и короткопериодических составляющих. В частности, подтверждена гипотеза о существовании т.н. цикла Гляйсберга в динамике солнечной активности продолжительностью около 80 лет.

3. Используя глобальные методы прогнозирования, обоснован новый способ выявления корреляций между системами естественного происхождения посредством анализа порождаемых ими временных рядов.

4. Предложено новое объяснение связи глобального потепления на Земле с активностью Солнца.

5. На основе общего подхода, использующего теорию динамических систем и метод локальной аппроксимации, найдено общее аналитическое решение задачи прогноза.

6. В рамках теории динамических систем даны оценки качества прогноза и предложены критерии выбора того или иного варианта прогнозирования в зависимости от природы временного ряда.

Публикации

1. A.Loskutov, I.Istomin, K.Kuzanyan. Prediction of Wolf number dynamics using singular spectrum analysis method.— *Abstracts of the Joint European and National Astronomical Meeting* JENAM 2000, Moscow, Russia, 26 May – 3 June 2000, p.129.

2. А.Ю.Лоскутов, И.А.Истомин, О.Л.Котляров, К.М.Кузанян. Исследование закономерностей магнитной активности Солнца методом сингулярного спектрального анализа.— *Письма в Астрон. журнал*, 2001, т.27, No11, с.867–876.

3. A.Loskutov, I.A.Istomin, K.M.Kuzanyan and O.L.Kotlyarov. Testing and forecasting the time series of the solar activity by singular spectrum analysis.— *Nonlin. Phenomena in Complex Syst.*, 2001, v.4, No1, p.47–57.

4. И.А.Истомин, О.Л.Котляров, А.Ю.Лоскутов. К проблеме обработки временных рядов: расширение возможностей метода локальной аппроксимации посредством сингулярного спектрального анализа.— *Теор. и матем. физика*, 2005, т.142, No1, с.148–159.

5. А.Ю.Лоскутов, О.Л.Котляров, И.А.Истомин, Д.И.Журавлев. Проблемы нелинейной динамики. III. Локальные методы прогнозирования временных рядов.— *Вестн. Моск. ун-та, сер. Физ.-астр.*, 2002, No6, с.3–21.

6. A.Loskutov, I.A.Istomin, K.M.Kuzanyan and O.L.Kotlyarov. Testing and forecasting time-series of the Solar activity by singular spectrum analysis.— http://xxx.lanl.gov/ps/nlin/-/0010027.

7. A.Loskutov, I.Istomin and O.Kotlyarov. Data analysis: generalizations of the local approximation method by singular spectrum analysis.— http://xxx.lanl.gov/abs/nlin.cd/0109022.