

УДК 004.932.2+004.932.72'1

М.И. Денисенко, В.Н. Беловодский

Донецкий национальный технический университет, г. Донецк
кафедра компьютерных систем мониторинга

ОПРЕДЕЛЕНИЕ РАЗМЕРНОСТИ ПРИ РЕКОНСТРУКЦИИ МОДЕЛИ: ПОДХОДЫ, МЕТОДЫ, СРАВНИТЕЛЬНЫЙ АНАЛИЗ

Аннотация

Денисенко М.И., Беловодский В.Н. Определение размерности при реконструкции модели: подходы, методы, сравнительный анализ. Рассмотрены основные этапы реконструкции обыкновенных дифференциальных уравнений на базе временных рядов. Выполнено исследование и сравнительный анализ методов определения размерности модели.

Ключевые слова: реконструкция уравнений, размерность модели, временные ряды.

Цель статьи – провести исследование этапов реконструкции обыкновенных дифференциальных уравнений на базе временных рядов.

Постановка задачи исследования. Исследовать основные методы определения размерности модели при реконструкции, провести их сравнительный анализ и на его основе выбрать метод для дальнейшего исследования в области прогнозирования.

Актуальность

Имея информацию о погодных условиях, метеорологи, как правило, могут предсказывать вероятность наступления и продолжительность тех или иных природных явлений: наличие или отсутствие осадков, облачность, направление и интенсивность ветра. Благодаря применению компьютерной техники в последние десятилетия возросла и точность прогнозов погоды. Но, в ряде случаев до сих пор случаются несовпадения прогнозируемых и реальных метеоусловий. Поэтому задача анализа и прогнозирования метеопараметров по временным рядам остается актуальной и сейчас. В первом приближении она заключается в формировании математической модели, конечной или дифференциальной, и последующего построения прогноза на ее основе.

Обзор предметной области исследования

В данной работе предполагается, что формирование математической модели будет проводиться на базе временных рядов, получаемых с метеостанции Vantage Pro 2, установленной на факультете КНТ ДонНТУ. Данная метеостанция проводит замеры ряда показателей (температуру,

влажность воздуха, давление, скорость ветра) и передачу их на серверы КСМ с помощью системы «АКИАМ». Интервал замеров составляет 10 минут. Таким образом, в процессе эксплуатации постоянно формируются и постепенно накапливаются временные ряды. Наличие этой информации делает вполне реальной задачу разработки прогностического комплекса.

Процедуру получения модельных уравнений по временным рядам, которая достаточно часто обсуждается в научной литературе последнего десятилетия, называют *реконструкцией* уравнений.

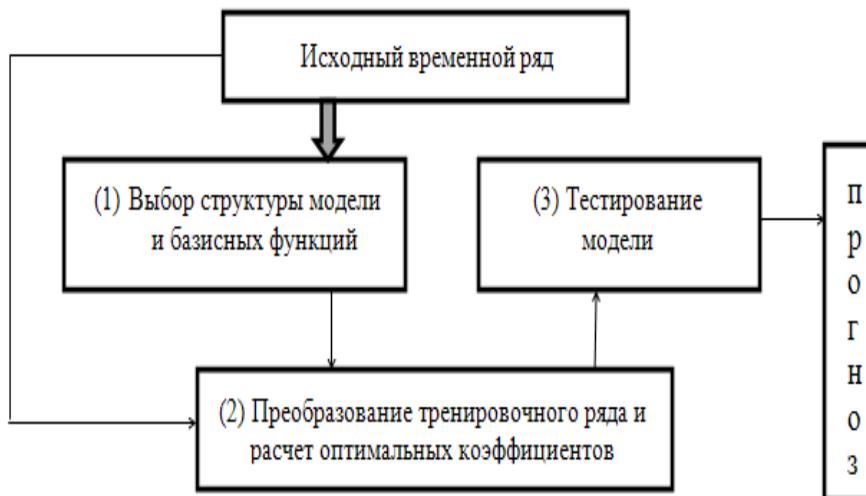


Рис.1. Схема процесса моделирования по временному ряду.

Реконструкция дифференциальных уравнений проводится по схеме (рис. 1). На первом этапе (1) на основании априорных данных ослабляются шумы, с необходимой частотой выбираются точки, восстанавливаются дополнительные переменные и т.п. Далее, искомую функцию F представляют с помощью выбранных на первом этапе *базисных* функций с неизвестными коэффициентами и находят значения коэффициентов, которые обеспечивают наилучшее соответствие модели и экспериментальных данных. На третьем этапе результаты предсказания с помощью полученной модели сравниваются с данными тестовой части ряда. При этом оценивают качество получившейся модели, для чего используются критерии, определяемые целью моделирования [1].

Анализ последних исследований. На кафедре КСМ в течение последних четырех лет выполнен ряд магистерских работ в этом направлении –

выпускниками Гриценко А.В и Сивяковым А.С. была разработана программная система Fcomplex, предназначенная для краткосрочного прогнозирования метеопараметров по временным рядам с использованием конечных математических моделей. Прогнозирование в ней осуществляется на базе временных рядов, которые берутся с метеостанции Vantage Pro 2. Fcomplex – это программная система для составления краткосрочных прогнозов значений метеопараметров "Температура", "Давление", "Влажность" и "Скорость ветра". Для определения размерности модели в ней используются метод главных компонент и ложных соседей, а для построения аппроксимирующих зависимостей – искусственные нейронные сети и метод Эглайса [5].

Обоснование и пути решения выделенных задач. В данной работе расчет прогноза предполагается проводить по традиционной схеме путем выполнения следующих этапов:

1. получение временного ряда из базы данных (БД), содержащей данные снимаемые с метеостанции Vantage Pro 2;
2. анализ данных;
3. определение размерности модели;
4. прогнозирование;
5. проверка эффективности модели.

В данной работе хотелось бы остановиться на третьем из перечисленных этапов, так как он является одним из важнейших при реконструкции уравнений.

Выбор размерности модели можно осуществлять на основе предварительного анализа временного ряда. В литературе описаны различные подходы к решению этой задачи. Рассмотрим некоторые из них, которые можно отнести к числу основных. Это метод Брумхеда-Кинга, оценка корреляционной размерности методом Грассбергера-Прокаччия и метод ложных ближайших соседей.

Идея метода Брумхеда-Кинга состоит в следующем. Сначала по скалярному ряду восстанавливаются методом задержек векторов в пространстве вложения достаточно большой размерности. Затем в этом пространстве производится преобразование поворота, причем направления новых осей координат выбираются в соответствии с теми направлениями, вдоль которых наиболее интенсивно развивается движение. Оценкой сверху для размерности служит количество направлений, вдоль которых имеется существенная протяженность траектории. Составляющие векторов вдоль тех направлений, которые слабо «осваиваются траекторией» списываются на влияние шума и ими пренебрегают [1].

Метод ложных соседей основан на проверке свойства, что фазовая траектория, восстановленная в пространстве достаточной размерности не должна иметь самопересечений. При пробной размерности D (исследуемая

модель в виде временного ряда) для каждого восстановленного вектора x_k отыскивают одного (самого близкого) соседа; увеличив D на 1, определяют, какие из соседей оказались ложными (сильно разошлись), а какие – истинными. Подсчитывают отношение числа ложных соседей к общему числу восстановленных векторов. Если при увеличении D это число становится малым при некотором значении D^* , то последнее и есть оценка размерности пространства, в котором достигается вложение траектории моделируемого движения [2].

Приведём один из алгоритмов вычисления относительного количества ложных ближайших соседей в восстановленном аттракторе. В лаговом пространстве выбираются две соседние точки, принадлежащие восстановленному аттрактору – x_i и x_j , вычисляется расстояние между ними, а также расстояние между следующими двумя точками – x_{i+1} и x_{j+1} . Далее, определяется соотношение между этими расстояниями:

$$R = \frac{|x_{i+1} - x_{j+1}|}{|x_i - x_j|}$$

Поскольку мы имеем дело, как правило, с гладкими процессами, то после одной итерации точки сильно не разбегутся даже при больших значениях показателя Ляпунова λ . Поэтому, если полученная величина превышает некоторое пороговое значение (обычно 10), то точки x_i и x_j можно считать ложными соседями. Таким образом, в идеальном случае, минимально необходимой будет такая размерность лагового пространства, при которой число ложных соседей равно нулю [4].

Метод Грассбергера–Прокаччия заключается в восстановлении аттрактора, «похожего» на исходный, последовательным сдвигом на величину τ . Для оценки размерности вложения последовательно получают новые размерности и измеряют некоторую характеристику получившегося многомерного ряда. После некоторого значения эта величина обычно перестает увеличиваться, что говорит о достижении размерности вложения. В качестве такой характеристики в нашем случае используется корреляционный интеграл $C(\epsilon)$. Корреляционный интеграл (корреляционный показатель) – это вероятность того, что временной ряд содержит пару точек, расстояние между которыми не превышает ϵ .

Метод заключается в построении зависимости корреляционного интеграла от характеристического расстояния:

$$C(l) = \lim_{n \rightarrow \infty} \frac{1}{N^2} \sum_{ij} \theta[l - |X_i - X_j|], \quad (1),$$

где l – характеристическое расстояние, θ – функция Хэвисайда, X_i и X_j точки d -мерного фазового пространства, N – количество экспериментально полученных точек фазовой траектории. Зависимость (1) может быть выражена как

$$C(l) \propto l^{D_c}, \quad (2),$$

где D_c – корреляционная размерность исследуемого аттрактора.

Как известно, для простых объектов, таких как линия или плоскость, корреляционная размерность будет равна размерности объекта. При увеличении размерности фазового пространства корреляционная размерность растет, выходя при определенном значении в насыщение. Уровень насыщения является фрактальной размерностью, а значение размерности фазового пространства – минимальной размерностью вложения аттрактора [3].

Чаще всего для определения размерности модели при реконструкции уравнений используется метод Грассбергера-Прокаччиа, который является одним из наиболее распространенных для определения корреляционной размерности. Его использование позволяет анализировать исходный ряд, поместив его в n -мерное пространство, что дает основания для получения более точных данных. Кроме того, программная система Fcomplex составлялась на основе методов главных компонент и метода ложных соседей, поэтому для дальнейших исследований будет использован метод Грассбергера-Прокаччиа.

Выводы. В работе описаны основные этапы реконструкции дифференциальных уравнений на базе временных рядов и методы определения размерности формируемой модели. В настоящее время результаты этого анализа уже используются при проектировании прогностической системы.

Список литературы

1. Безручко Б.П., Смирнов Д.А. Реконструкция обыкновенных дифференциальных уравнений по временным рядам. Учебно-методическое пособие, – Саратов: ГосУНЦ «Колледж», 2000 – 46 с.
2. Безручко Б.П., Смирнов Д.А. Математическое моделирование и хаотические временные ряды. Саратов: ГосУНЦ «Колледж», 2005. – 320 с.
3. Исследователь. [Electronic resource] / Интернет-ресурс. – Режим доступа: <http://ligis.ru/effects/stat/modules/sttimser.html> – Анализ временных рядов.
4. M. V. Kennel, R. Brown, and H. D. I. Abarbanel. Determining embedding dimension for phasespace reconstruction using a geometrical construction// Phys. Rev. A, 45, 3403, 1992.
5. Гриценко А.В. Реконструкция уравнений и прогнозирование метеопараметров по их временным рядам. [Текст] – Донецк, ДонНТУ, 2010. – 149 с.