

Федеральное государственное автономное образовательное учреждение  
высшего образования «Уральский федеральный университет имени первого  
Президента России Б.Н. Ельцина»

На правах рукописи



**Тимошенкова Юлия Сергеевна**

**Разработка методики интеграции формальных методов  
прогнозирования временных рядов и метода ассимиляции  
данных**

2.3.1. Системный анализ, управление и обработка информации,  
статистика

**Автореферат**  
диссертации на соискание учёной степени  
кандидата технических наук

Работа выполнена в учебно-научном центре «Информационная безопасность» Института радиоэлектроники и информационных технологий – РТФ ФГАОУ ВО «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина».

Научный руководитель: доктор технических наук, профессор,  
**Поршнев Сергей Владимирович**

Официальные оппоненты: **Авдеевко Татьяна Владимировна**, доктор технических наук, профессор, ФГБОУ ВО «Новосибирский государственный технический университет», г. Новосибирск, профессор кафедры теоретической и прикладной информатики;

**Зобнин Борис Борисович**, доктор технических наук, профессор, ФГБОУ ВО «Уральский государственный горный университет», г. Екатеринбург, профессор кафедры информатики;

**Никольцев Евгений Витальевич**, доктор технических наук, профессор, ФГБОУ ВО «МИРЭА – Российский технологический университет», г. Москва, профессор кафедры КБ-14 «Цифровых технологий обработки данных» Института кибербезопасности и цифровых технологий.

Защита состоится 06 декабря 2022 г. в 11:00 часов на заседании диссертационного совета УрФУ 2.3.12.13 по адресу: 620002, г. Екатеринбург, ул. Мира, 19, ауд. И-420 (зал Ученого совета).

С диссертацией можно ознакомиться в библиотеке и на сайте ФГАОУ ВО «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина»,

<https://dissovet2.urfu.ru/mod/data/view.php?d=12&rid=3914>

Автореферат разослан «\_\_\_\_\_» \_\_\_\_\_ 2022 года.

Ученый секретарь  
диссертационного совета



Сафиуллин Николай Тахирович

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

### Актуальность темы исследования и степень ее разработанности.

Под термином «прогноз», традиционно, понимают научно обоснованное суждение о возможных состояниях выбранного объекта (динамической системы (ДС)) в будущем и/или его (ее) траекториях в пространстве состояний, а также сроках изменения данных состояний. Методы прогнозирования, помимо общеизвестного ежедневного прогноза погоды, применяется, например: в автоматических системах управления производственными процессами (H.S. Hippert); для прогнозирования геофизических и метеорологических явлений на Земле (N. Rescher, L.I. Fen et al.); в геологии и геофизике (J. Benes, G. Herrera et al.); в медицине (J.M. Wu, A.F. Hundley, R.G. Fulton, E.R. Myers, A.E.E. Hridoy et al.); в политологии (S.I. Bretschneider, W.L. Gorr, G. Grizzle, E. Klay); в макроэкономике и государственном управлении (L.R. Weatherford, S.E. Kimes, R.D. Lee, L.R. Carter, J.A. Fox, A.R. Piquero); в экономике (L. Gyorf, G. Lugosi, F. Udina); в управлении финансами (V. Kodogiannis, A. Lolis).

Таким образом, задача прогнозирования будущих состояний ДС на основе информации о ее предыдущих состояниях является актуальной в различных областях человеческой деятельности (G. Box, H. Jenkins, G. Reinsel, G. Ljung).

Во многих случаях задачу прогнозирования изменения состояния ДС удается свести к задаче прогнозирования упорядоченных по времени последовательностей, составленных из значений параметров процессов  $X(t)$ , порожденных данной ДС, измеренных в узлах равномерной временной сетки  $t_i = i\Delta t$ ,  $i = \overline{0, N-1}$ , где  $\Delta t = \text{const}$  – шаг временной сетки. Данные последовательности называют временным рядом (ВР). При этом различают следующие типы ДС:

- ДС, описываемые известными математическими моделями, которые построены на основе физических законов, определяющих ее эволюцию во времени;
- ДС, описываемые с помощью феноменологических моделей порожденных ими ВР – математических моделей ВР, построенных на основе их анализа, так как до конца неизучены физические законы, описывающие их динамику.

Прогнозирование ВР, порожденных ДС первого типа, осуществляется на основе использования соответствующей математической модели ДС, прогнозирование ВР второго типа – на основе использования феноменологической (формальной) модели ВР, выбираемой исходя из анализа свойств прогнозируемого ВР. Соответственно, методы прогнозирования ВР разделяют на методы прогнозирования ВР, основанные на использовании математической модели ДС, породившей данный ВР, и методы прогнозирования, основанные на использовании формальных моделей ВР – формальные методы прогнозирования ВР.

Также на практике приходится решать задачу прогнозирования ВР, у которых известны как его ранее спрогнозированные в выбранные моменты времени значения, так и соответствующие им наблюдаемые значения ВР (например,

ВР, составленный из значений оптовых цен на электроэнергию, ВР, составленный из значений обменных курсов валют и т.д.). Наличие данной информации, потенциально, дает возможность проводить сравнение соответствующего спрогнозированного и наблюдаемого значений ВР и на его основе корректировать значения последующих спрогнозированных значений ВР, улучшая, тем самым, точность прогноза. Такой подход известен как метод ассимиляции данных (Data Assimilation – DA) и используется, например, при прогнозировании погоды. Необходимым условием его использования является наличие математической модели ДС, породившей прогнозируемый ВР. Однако, на практике также требуется прогнозировать ВР, порожденные ДС, описывающими феноменологическими моделями. При этом понятно, что спрогнозированные на основе феноменологических моделей значения ВР, потенциально, также можно корректировать на основе их сравнения с соответствующими наблюдаемыми значения ВР. Между тем, для ВР данного типа, методов, аналогичных методу DA, на момент начала диссертационного исследования не существовало.

Начало разработки формальных методов прогнозирования ВР датируется 1910 г., когда было предложено использовать для прогнозирования ВР модель скользящего среднего (MA). В 1920 г. в работах G. Yule и J.C. Walker впервые была предложена авторегрессионная модель (AR) ВР. Далее в 1970 г. G. Voh и H. Jenkins объединили AR- и MA-модели ВР в единую модель, получившей название авторегрессии-скользящего среднего (ARMA). Затем в 1976 г. на основе ARMA-модели G. Voh и H. Jenkins была разработана модель авторегрессии-проинтегрированного скользящего среднего (ARIMA).

В 1989 г. R. Vautard, M. Ghil предложили метод сингулярного спектрального анализа (SSA), относящийся к классу непараметрических методов анализа и прогнозирования ВР, основанный на декомпозиции исследуемого ВР на некоторый аддитивный набор более простых ВР, называемых главными компонентами (ГК), которые могут быть интерпретированы с точки зрения представлений о динамике ДС, породившей данный ВР (например, тренд, гармонические составляющие, шум и т.д.). В основании метода SSA лежит преобразование одномерного ВР в многомерный ВР с помощью однопараметрической сдвиговой процедуры (отсюда в русскоязычной литературе данный метод получил название «Гусеница») и далее исследование полученной многомерной траектории с помощью сингулярного разложения и восстановления (аппроксимации) ВР по выбранным ГК. Отметим, что данный метод не требует стационарности анализируемого ВР, в отличие, например, от Фурье-анализа и параметрических методов спектрального оценивания. Исследованием различных теоретических и прикладных аспектов метода SSA также занимались Н.Э. Голяндина, А.А. Жиглевский, С.В. Поршнев и др.

В 1968 г. А.Г. Ивахненко был предложен формальный метод прогнозирования ВР, получивший название метод группового учета аргументов (GMDH), основанный на гипотезе о том, что значение  $x_N$  ВР  $x_i$  зависит от  $B$  его предыдущих значений:

$$x_N = F(x_{N-1}, x_{N-2}, \dots, x_{N-B}). \quad (1)$$

Здесь в качестве функции  $F(\dots)$  он предложил использовать полином Колмогорова-Габора:

$$x_N = a_0 + \sum_{i=1}^m a_i x_{N-i} + \sum_{i=1}^m \sum_{j=1}^m a_{i,j} x_{N-i} x_{N-j} + \\ \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m a_{i,j,k} x_{N-i} x_{N-j} x_{N-k},$$

$m \leq B$ , коэффициенты которого вычисляются с использованием метода наименьших квадратов (МНК), который, однако, оказывается весьма затратным с вычислительной точки зрения.

Далее в 1982 г. им же был предложен самоорганизующийся GMDH, в котором используются полиномы Колмогорова-Габора второго порядка, аргументами которых являются все возможные сочетания пар входных переменных  $\{x_{N-1}, x_{N-2}, \dots, x_{N-B}\}$ . В результате строится множество регрессионных моделей ВР первого уровня. Из них выбираются те регрессионные модели, дисперсия ошибок остатков которых оказывается меньше выбранного значения. Затем строятся значения, в которых в качестве входных переменных полиномов используются регрессионные модели второго уровня полиномов первого уровня т.д. Описанная выше процедура построения полиномов продолжается до тех пор пока погрешность полиномов данного уровня не оказывается меньше погрешности полинома предыдущего уровня. Данный алгоритм оказывается менее затратным с вычислительной точки зрения, нежели классический GMDH.

Один из активно используемых в настоящее время формальных методов прогнозирования ВР основан на использовании искусственных нейронных сетей (1943 г. W. McCulloch и W. Pitts), выполняющих роль функции  $F(\dots)$  в (1). При этом оказывается, что наилучшее качество прогнозов, в рассматриваемой задаче, обеспечивают рекуррентные нейронные сети (RNN) с долгой краткосрочной памятью (LSTM) (1997 г. J. Hochreiter и J. Schmidhuber).

В тоже время, несмотря на популярность формальных методов прогнозирования ВР, в них отсутствует возможность коррекции спрогнозированных значений ВР на основе сравнения соответствующих спрогнозированного и наблюдаемого значений ВР.

Свободным от отмеченного выше недостатка формальных методов прогнозирования ВР оказывается метод ассимиляции данных (Data Assimilation, DA), предложенный в работах М. Talagrand и J. Courier. Однако, обязательным условием использования DA является знание математической модели ДС, породившей данный ВР, которое, как оказывается, выполняется на практике далеко не всегда.

В этой связи **цель исследования**, состоящая в разработке научно-обоснованной методики интеграции формальных методов прогнозирования ВР и метода DA является актуальной.

Для достижения поставленной цели потребовалось решить следующие задачи:

1. Анализ существующих методов прогнозирования и коррекции ВР.
2. Разработка методики интеграции формальных методов прогнозирования ВР в DA.
3. Разработка программной реализации методики интеграции формальных методов прогнозирования ВР в DA.
4. Экспериментальная апробация разработанной методики интеграции формальных методов прогнозирования ВР в DA на примере ВР, порожденных ДС, описываемыми феноменологическими моделями.

**Объект исследования:** методы прогнозирования и коррекции спрогнозированных значений ВР с целью увеличения их точности.

**Предмет исследования:** методика интеграции формальных методов прогнозирования ВР, порожденных ДС, описываемыми феноменологическими моделями, и метода DA.

**Научная новизна:** диссертационного исследования заключается в:

1. Обосновании возможности интеграции формальных методов прогнозирования ВР, порожденных ДС, описываемыми феноменологическими моделями, и метода DA.
2. Разработке методики интеграции формальных методов прогнозирования ВР в метод DA и ее программной реализации.
3. Обосновании выбора метрик и набора показателей, обеспечивающих количественную оценку качества прогнозирования ВР.

**Теоретическая значимость** диссертационного исследования заключается в разработке методики интеграции формальных методов прогнозирования ВР, порожденных ДС, описываемыми феноменологическими моделями, и метода DA.

**Практическая значимость** диссертационного исследования заключается в разработке программного комплекса, обеспечивающего использование разработанной методики для прогнозирования реальных ВР, в котором реализованы:

1. Формальные методы прогнозирования ВР: ARIMA, SSA, GMDH, LSTM-net.
2. Классический метод DA.
3. Методика интеграции формальных методов прогнозирования ВР и метода Data Assimilation.

Работа выполнена в рамках проекта № 20-37-90006/20 от 21.08.2020, поддержанного грантом РФФИ «Аспиранты» 2020 г.

**Методология и методы исследования.** В работе использованы методы анализа и прогнозирования ВР, методы ассимиляции данных, методы математической статистики, вычислительной математики и системного анализа.

### **Основные положения, выносимые на защиту:**

1. Разработана научно обоснованная методика, обеспечивающая интеграцию формальных методов прогнозирования ВР и метода DA.

2. Выбранные метрики и набор показателей для оценки качества прогнозирования ВР обеспечивают возможность количественного сравнения результатов прогнозирования ВР с помощью разработанной методики интеграции формальных методов прогнозирования ВР и метода DA.

3. Результаты прогнозирования ВР, представляющих собой ВР, порожденные ДС, описываемыми феноменологическими моделями, подтверждают работоспособность разработанной методики интеграции формальных методов прогнозирования ВР и метода DA и ее программной реализации, использованы: ВР «Air Passengers», составленный из данных о числе пассажиров, перевезенных авиакомпаниями США в период с января 1949 по декабрь 1960; ВР, составленного из среднемесячных значений чисел Вольфа (числа наблюдаемых солнечных пятен), сглаженных по 13 месяцам, являющихся количественными показателями активности Солнца с августа 1913 по сентябрь 2021 гг.; ВР, составленного из значений переводных курсов доллар/рубль, евро/рубль в период с ноября 2021 г. по апрель 2022 г.; представляющие собой ВР, порожденные ДС, описываемыми феноменологическими моделями, подтверждают работоспособность разработанной методики интеграции формальных методов прогнозирования ВР и метода DA и ее программной реализации.

**Достоверность** полученных результатов подтверждается обоснованным применением методов прогнозирования ВР, результатами тестовых испытаний разработанного программного комплекса «Прогнозирование временных рядов», а также сходимостью результатов прогнозирования ВР с помощью разработанной методики интеграции с соответствующими результатами их наблюдения.

**Внедрение результатов диссертационного исследования.** Результаты диссертационного исследования внедрены в ООО «Эйрбэйс» (акт об использовании результатов от 05.03.2022), а также используются в федеральном государственном автономном образовательном учреждении высшего образования «Уральский федеральный университет им. первого президента России Б.Н. Ельцина» при подготовке бакалавров и специалистов по укрупненной группе специальностей 10.00.00 Информационная безопасность (акт об использовании результатов от 21.07.2022).

**Апробация работы.** Результаты работы докладывались и обсуждались на научно-практических конференциях: Международная конференция студентов, аспирантов и молодых ученых «Информационные технологии, телекоммуникации и системы управления», 2017, Екатеринбург, Россия; XI International scientific and technical conference «Applied Mechanics and Dynamics Systems», 14–16 ноября 2017, Омск, Россия; 1st International Conference on Physics, Mathematics and Statistics (ICPMS 2018), 12–14 May 2018, Shanghai, China; 2018 International Conference on Applied Mathematics and Computational Science (ICAMCS.NET 2018), 6-8 october 2018, Budapest, Hungary; International Conference on Numerical

Analysis and Applied Mathematics 2018 (ICNAAM 2018), 13-18 september 2018, Rhodes, Greece; 4th International Workshop on Radio Electronics and Information Technologies (REIT-Autumn 2018), Екатеринбург, Россия; Международная конференция 2019 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT), 25-26 апреля 2019, Екатеринбург, Россия; Международная конференция 2020 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT), 14–15 мая 2020, Екатеринбург, Россия; Международная конференция 2021 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT), 13–14 мая 2021, Екатеринбург, Россия.

**Личный вклад.** Автор обосновал возможность интеграции формальных методов прогнозирования ВР и метода DA, разработал методику интеграции формальных методов прогнозирования ВР и метода DA и ее программную реализацию.

**Публикации.** По теме диссертации опубликовано 14 работ, из них 9 статей опубликованы в рецензируемых научных изданиях, определенных ВАК РФ и Аттестационным советом УрФУ, включая 8 статей в изданиях, индексируемых в международных цитатно-аналитических базах Scopus и Web of Science; имеется 3 свидетельства о государственной регистрации программы для ЭВМ.

**Объем и структура работы.** Диссертация состоит из введения, 4 глав, заключения и 2 приложений. Полный объем диссертации составляет 162 страницы, включая 77 рисунков и 6 таблиц. Список литературы содержит 106 наименований.

## Содержание работы

Во **введении** обоснована актуальность выбранной темы исследований, проводимых в рамках данной диссертационной работы, сформулированы цели и задачи работы. Представлена актуальность работы, основные результаты. Приводятся положения, выносимые на защиту и научная новизна.

В **первой главе** введены основные термины и определения, используемые в диссертации; описаны алгоритмы формальных методов прогнозирования ВР, в том числе: ARIMA, SSA, GMDH, LSTM, а также метода DA, используемого для прогнозирования и коррекции спрогнозированных значений ВР, порожденных ДС, описываемыми с помощью математических моделей; сформулированы цель и задачи исследования.

**Вторая глава** посвящена разработке методики интеграции формальных методов прогнозирования ВР и метода DA, в котором полагается, что эволюция ДС, породившей прогнозируемый ВР, описывается в дискретном времени следующей системой уравнений:



$$\begin{cases} \vec{x}_{k+1} = M(\vec{x}_k, t_k) + \vec{w}_k, \\ \vec{y}_k = H(\vec{x}_k, t_k) + \vec{v}_k, \end{cases} \quad (2)$$

где  $\vec{x}$  – вектор состояния динамической системы,  $M$  – оператор или функция перехода, определяющая эволюцию системы во времени,  $\vec{x}_k$  – вектор состояния системы в момент  $t_k$ ,  $x_{k+1}$  – вектор состояния системы в момент времени  $t_{k+1}$ ,  $\vec{y}_k$  – вектор наблюдений,  $H(\vec{x}_k, t_k)$  – оператор наблюдения, связывающий многомерное состояние системы с одномерным вектором наблюдений, в предположении о том, что одновременно наблюдаются значения всех переменных системы, оказывается единичной матрицей,  $\vec{w}_k$ ,  $\vec{v}_k$  – ошибки модели и наблюдения, соответственно. Данное уравнение может быть использовано для прогнозирования одномерных ВР без векторного представления.

Далее было проведено сравнение выражения, используемого для решения задачи прогнозирования ВР, порожденных ДС, описываемыми феноменологическими моделями:

$$x_{k+1} = F(x_k, x_{k-1}, \dots, x_{k-p}, \alpha_0, \alpha_1, \dots, \alpha_p) + w_k \quad (3)$$

и первого уравнения в (2), которое позволило сделать вывод о том, что с точностью до аддитивного члена  $w_k$  (ошибки прогноза), данные уравнения имеют одинаковую структуру. Следовательно, функция  $F$ , по своей сути, аналогична оператору  $M$  в (2), то есть использование (3) эквивалентно переходу от описания ДС в пространстве состояний к ее описанию в  $p$ -мерном координатном пространстве. На основе использования данной аналогии, предложена следующая система уравнений, объединяющая метод DA и формальные методы прогнозирования ВР:

$$\begin{cases} x_{k+1} = F(x_k, t_k) + w_k, \\ y_k = x_k + v_k, \end{cases} \quad (4)$$

где  $F$  – функция, определяемая выбранным формальным методом прогнозирования ВР;  $y_k$  – наблюдаемое значение состояния ДС, где второе уравнение связывает  $x_k$  с наблюдаемым значением  $y_k$ , а  $w_k$ ,  $v_k$  – ошибки модели и наблюдения, соответственно.

Схема вычислительного алгоритма согласно (4) представлена на рисунке 1, из которого видно, что данный алгоритм реализуется выполнением следующей последовательности действий:

1. Выбор формального метода прогнозирования.
2. Задать параметры метода DA и формального метода прогноза.
3. Получить ансамбль прогнозов размером  $N$  в точке с помощью формального метода и добавления шума (для фильтра Калмана  $N = 1$ , для The Ensemble Kalman Filter (EnKF-фильтра)  $N \geq 1$ ).
4. Получить данные наблюдений системы в прогнозируемый момент времени.

5. Рассчитать матрицы ковариации ошибок прогноза.
  6. Выполнить коррекцию прогноза путем применения матриц ковариации и передаточного коэффициента фильтра Калмана  $K$  для каждой точки ансамбля размером  $N$ .
  7. Обновить каждый член ансамбля прогноза и матрицы ковариаций с использованием соответствующих значений.
  8. Выполнить пункты 3–7 для каждой прогнозируемой точки.
  9. Оценить точность полученного прогноза.
- Работоспособность предложенной методики интеграции формальных методов прогнозирования ВР и метода ДА подтверждают результаты прогнозирования координат аттрактора Лоренца:

$$\begin{aligned}\frac{\partial X}{\partial t} &= \sigma(Y - X), \\ \frac{\partial Y}{\partial t} &= -X \cdot Z + r \cdot X - Y, \\ \frac{\partial Z}{\partial t} &= X \cdot Y - Z \cdot Z,\end{aligned}$$

где  $\sigma = 10$ ,  $r = 28$ ,  $b = 8/3$ , представленный на рисунке 2, визуальный анализ которого позволяет сделать вывод о том, что погрешность прогнозирования предложенной методикой интеграции формальных моделей прогнозирования ВР и метода ДА оказывается сравнимой с аналогичной величиной классического метода ДА. Данный результат позволил сделать обоснованный вывод о работоспособности предложенной методики интеграции.

Отметим, что для каждой координаты системы дифференциальных уравнений (СДУ) Лоренца прогнозирование формальным методом выполнялось отдельно, то есть база прогноза и параметры метода у каждой координаты свои и независимые, поэтому в (4) уравнения прогнозов для СДУ Лоренца с использованием АР-модели можно записать в виде:

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \\ z_{k+1} \end{bmatrix} = \begin{bmatrix} F_x(x_k, \dots, x_{k-10}, \alpha_{x0}, \dots, \alpha_{x10}) + w_{xk} \\ F_y(y_k, \dots, y_{k-10}, \alpha_{y0}, \dots, \alpha_{y10}) + w_{yk} \\ F_z(z_k, \dots, z_{k-10}, \alpha_{z0}, \dots, \alpha_{z10}) + w_{zk} \end{bmatrix}.$$

Выполнен сравнительный анализ прогнозирования с помощью классического метода ДА и разработанной методики интеграции на примере координат аттрактора Лоренца с параметрами, проводящими систему в хаотический режим, который показал, что разработанная методика дает результат не хуже и может применяться наравне с классическим методом ДА при применении к ВР, порождённым ДС, описываемыми известными математическими моделями.

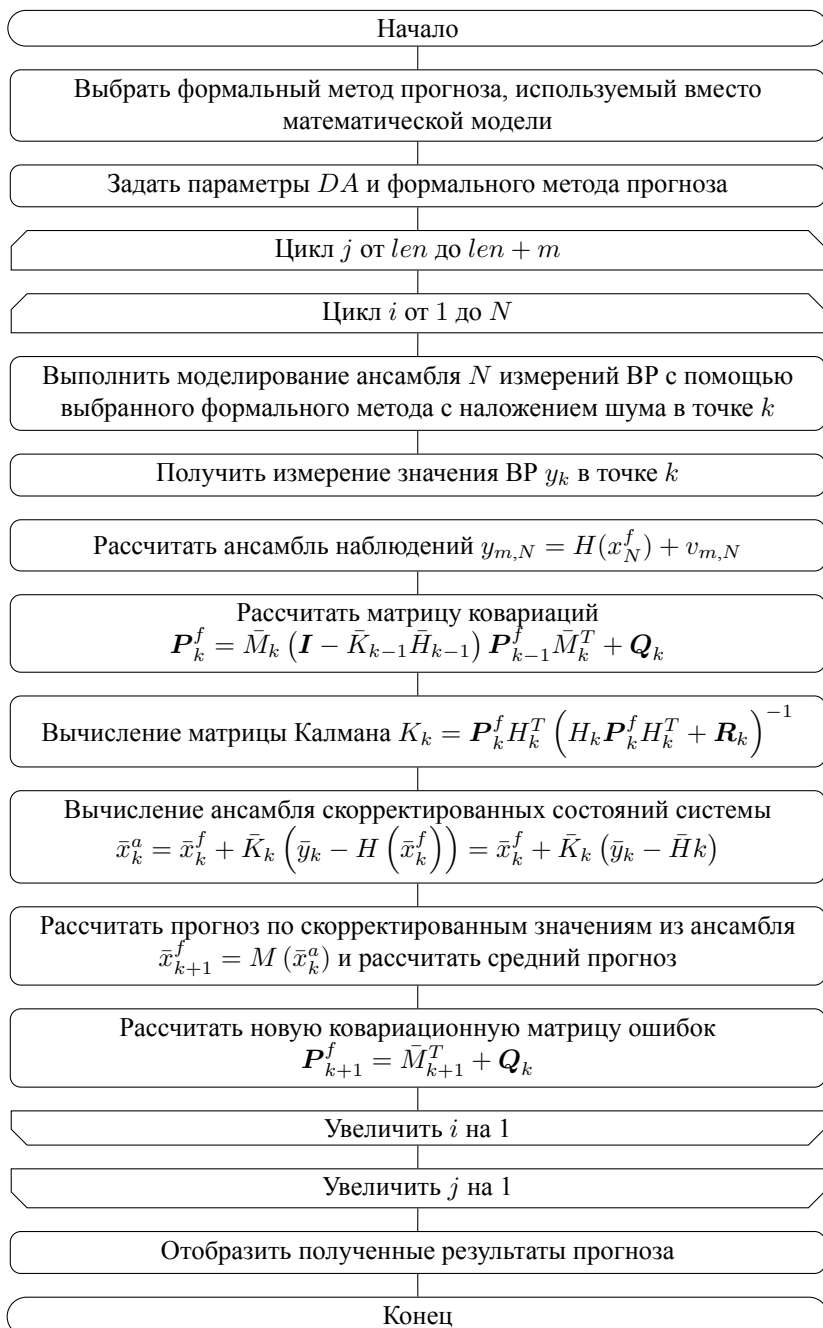


Рисунок 1 — Схема алгоритма интеграции формальных методов прогнозирования ВР и метода DA

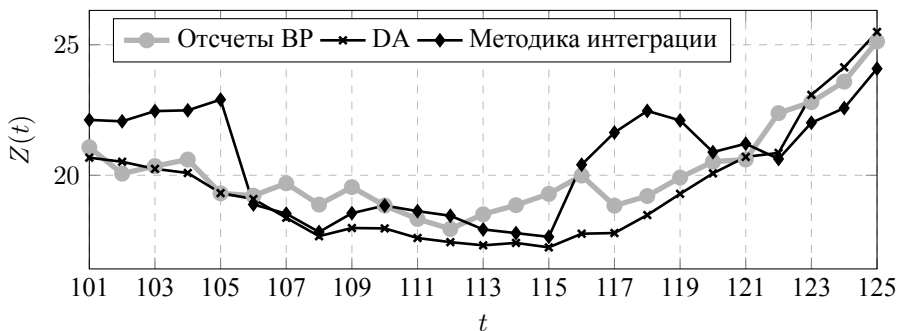


Рисунок 2 — Типичные результаты прогнозирования решений координаты  $Z$  для СДУ Лоренца

В **третьей главе** описан разработанный программный комплекс «Прогнозирование временных рядов» (ПК ПБР) на языке Python 3 и его пользовательский интерфейс, предназначенный для прогнозирования ВР. ПК ПБР совместим с популярными операционными системами (Microsoft Windows выше 7 версии, Linux выше 16.04, macOS выше 10.12.6), минимальные системные требования 2 GB RAM, для поддержки графического процессора требуется графический процессор с поддержкой CUDA.

ПК ПБР состоит из трех программных библиотек (ПБ), обеспечивающих прогнозирование ВР, и модулей, обеспечивающих взаимодействие пользователя с комплексом:

- ПБ № 1 «Формальные методы прогнозирования временных рядов», содержащий функции, реализующие следующие формальные методы прогнозирования ВР: ARIMA, SSA, GMDH;
- ПБ № 2 «Прогнозирование временных рядов с помощью нейронной сети LSTM», содержащий функции, обеспечивающие прогнозирование ВР на основе нейронной сети LSTM;
- ПБ № 3 «Прогнозирование временных рядов на основе интеграции формальных методов и метода ассимиляции данных», содержащий функции, обеспечивающие прогнозирование ВР на основе интеграции формальных методов прогнозирования ВР и метода DA;
- модуля *Estimator*, содержащего функции генерации модельных ВР или загрузки данных из файла;
- модуля *Error*, содержащего функции, используемые для расчета точности прогнозирования ВР;
- модуля *Visualisation*, содержащего функции для визуализации результатов прогноза в графическом и текстовом виде.

В **четвертой главе** проведен анализ результатов применения разработанной методики интеграции для прогнозирования реальных ВР, порожденных ДС, описывающихся с помощью феноменологических моделей. С целью оценки

качества прогнозирования с помощью разработанной методики интеграции формальных методов прогнозирования ВР и метода ДА, для получения статистически устойчивых оценок точности прогнозирования выбранных ВР  $x_i, i = \overline{1, N}$ , формировались ансамбли, состоящие из 50-ти статистически независимых значений прогнозов значений ВР  $\{x_k^f\}_m, k = \overline{1, K}, K$  – число прогнозируемых точек,  $m = \overline{1, 50}$ ,  $m$  – номер независимого испытания. Прогноз выполнялся ретроспективно, что обеспечило возможность вычисления ансамблей погрешностей спрогнозированных значений ВР:

$$\{\Delta^{<f>} x_k\}_m = \{x_k^{<f>}\}_m - x_{N+k}^{obs},$$

где  $x_{N+k}^{obs}$  –  $(N + k)$ -ое наблюдаемое значение прогнозируемого ВР,  $f = AR, SSA, GMDH, LSTM$  – множество реализованных в ПК ПВР формальных методов интегрированных в разработанную методику интеграции формальных методов прогнозирования ВР и метода ДА.

Кроме точечных оценок погрешности прогнозов для каждого  $k$ -го ансамбля погрешностей прогнозов вычислялись их статистические характеристики, в том числе средние по номеру точки прогнозирования (что аналогично, по времени) значения смещения прогнозов:

$$\bar{\Delta}_{\text{mean}}^{(f)}(i_{2\text{ st}}, i_{2\text{ fin}}) = \frac{1}{i_{2\text{ fin}} - i_{2\text{ st}}} \sum_{i_{2\text{ st}}=i_{2\text{ st}}}^{i_{2\text{ fin}}} \text{mean}(\Delta \tilde{x}_{i_2}^{(f)}), \quad (5)$$

и среднее значение среднеквадратического отклонения:

$$\bar{\Delta}_{\text{std}}^{(f)}(i_{2\text{ st}}, i_{2\text{ fin}}) = \frac{1}{i_{2\text{ fin}} - i_{2\text{ st}}} \sum_{i_{2\text{ st}}=i_{2\text{ st}}}^{i_{2\text{ fin}}} \text{std}(\Delta \tilde{x}_{i_2}^{(f)}), \quad (6)$$

где  $i_{2\text{ st}}$  – номер начальной точки прогноза, используемой для оценки выбранных показателей точности прогноза,  $i_{2\text{ fin}}$  – номер конечной точки прогноза, используемой для оценки выбранных показателей точности прогноза.

**ВР «Air Passengers», составленный из данных о числе пассажиров, перевезенных авиакомпаниями США в период с января 1949 по декабрь 1960.** (пример прогноза на рис. 3)

Анализ значений полученных показателей (таб. 1) позволили сделать следующие выводы:

1. При увеличении дальности прогнозирования погрешность прогнозов, вычисленных с помощью формальных методов прогнозирования ВР, увеличивается, в то время как при использовании методики интеграции, наоборот, при увеличении дальности прогнозирования погрешность прогнозирования, как минимум, не увеличивается. (Данная особенность, с нашей точки зрения, связана с тем, что при инициализации параметров матриц ковариации ошибок  $P$  и  $Q$  взяты нулевые начальные условия, а далее эти значения элементов данных матриц обновляются согласно полученным оценкам ошибок прогноза и наблюдений, что и обеспечивает возможность долгосрочного прогнозирования.)

Таблица 1 — Значения усредненных погрешностей прогнозов ВР «Air Passengers» для  $i_{2\text{ ст}} = 6$  и  $i_{2\text{ фин}} = 12$

| Метод   | mean (5) | std (6) |
|---------|----------|---------|
| DA&AR   | 1,964    | 0,1288  |
| DA&SSA  | 4,192    | 0,0102  |
| DA&GMDH | 11,408   | 0,0444  |
| DA&LSTM | 7,417    | 0,6180  |

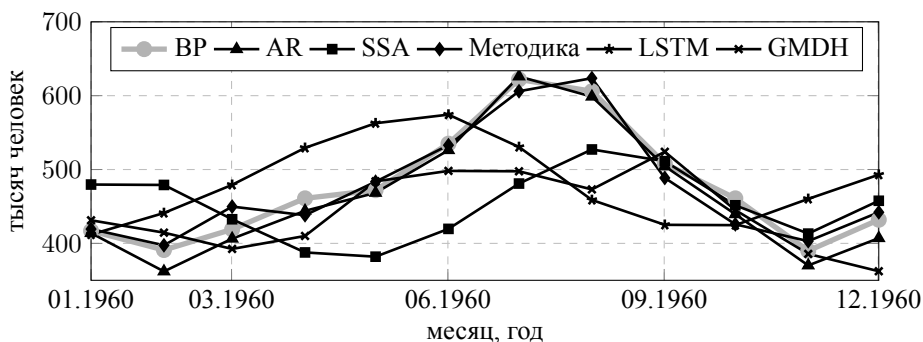


Рисунок 3 — Результаты прогнозирования ВР «Air Passengers» с помощью формальных методов прогнозирования ВР (AR, SSA, GMDH, LSTM) и разработанной методики интеграции формальных методов прогнозирования ВР и метода DA на основе AR(12)

2. Продемонстрировано, что локальная погрешность прогнозов, вычисленных с использованием методики интеграции, и размер области рассеяния возможных значений прогнозов, является случайными величинами, зависящими от локальных характеристик прогнозируемого ВР.

3. Предложены количественные характеристики, обеспечивающие сравнение погрешности результатов прогнозирования ВР с помощью методики интеграции при использовании различных методов прогнозирования ВР.

4. При использовании формальных методов в методике наблюдается тенденция уменьшения математического ожидания с увеличением дальности прогноза.

**ВР, составленный из среднемесячных значений чисел Вольфа, с августа 1913 по сентябрь 2021 гг. (пример прогноза на рис. 4)**

Значения выбранных показателей погрешности прогноза ВР (таб. 2) позволили сделать выводы:

1. При увеличении дальности прогнозирования погрешность прогнозов, вычисленных с помощью формальных методов прогнозирования ВР, увеличивается, в то время как при использовании методики интеграции, наоборот, при увеличении дальности прогнозирования погрешность прогнозирования, как минимум, не увеличивается.

Таблица 2 — Значения усредненных погрешностей прогнозов чисел Вольфа для  $i_{2\text{ st}} = 115$  и  $i_{2\text{ fin}} = 152$

| Метод   | mean (5) | std (6) |
|---------|----------|---------|
| DA&AR   | 3,373    | 0,094   |
| DA&SSA  | 3,241    | 0,542   |
| DA&GMDH | 8,126    | 1,792   |
| DA&LSTM | 5,428    | 1,368   |

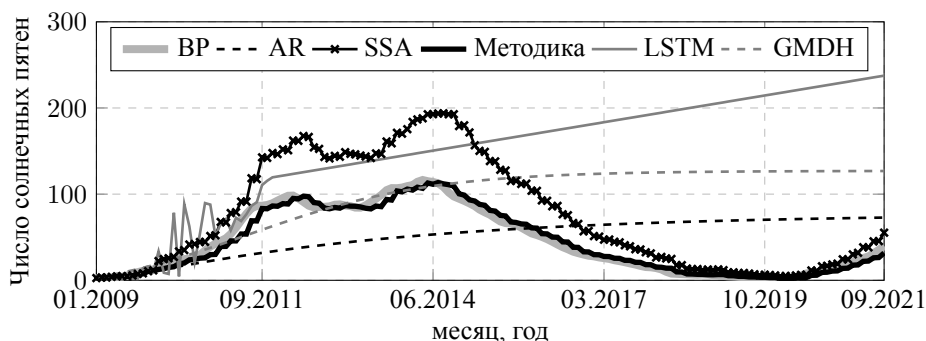


Рисунок 4 — Результаты прогнозирования BP чисел Вольфа с помощью формальных методов прогнозирования BP (AR, SSA, GMDH, LSTM) и разработанной методики интеграции формальных методов прогнозирования BP в и метода DA на основе AR(2)

2. Наименьшая погрешность прогноза получена при использовании метода AR в методике интеграции. При этом порядок  $p$  метода AR при использовании в методике интеграции можно брать невысоким, что позволяет сделать выбор параметров проще, а расчет коэффициентов быстрее и эффективнее.

3. Предложены количественные характеристики, обеспечивающие сравнение погрешности результатов прогнозирования BP с помощью методики интеграции при использовании различных методов прогнозирования BP.

4. При использовании формальных методов в методике интегрирования наблюдается тенденция уменьшения математического ожидания ошибки с дальностью прогноза.

**BP, составленный из значений переводных курсов доллар/рубль, евро/рубль в период с ноября 2021 г. по апрель 2022 г.** (пример прогноза на рис. 5–6)

Значения выбранных показателей погрешности прогноза BP (таб. 3–4) позволили сделать выводы:

1. Усредненная погрешность прогнозирования, реализованного с помощью разработанной методики интеграции, в сравнении с аналогичной величиной у соответствующих формальных методов прогнозирования BP оказывается

Таблица 3 — Значения усредненных погрешностей прогнозов курса доллар/рубль для  $i_{2\text{ st}} = 10$  и  $i_{2\text{ fin}} = 22$

| Метод   | mean (5) | std (6) |
|---------|----------|---------|
| DA&AR   | 0,923    | 0,191   |
| DA&SSA  | 1,570    | 0,035   |
| DA&GMDH | 1,138    | 0,560   |
| DA&LSTM | 1,489    | 0,042   |

Таблица 4 — Значения усредненных погрешностей прогнозов курс евро/рубль для  $i_{2\text{ st}} = 10$  и  $i_{2\text{ fin}} = 22$

| Метод   | mean (5) | std (6)               |
|---------|----------|-----------------------|
| DA&AR   | 1,961    | $1,703 \cdot 10^{-1}$ |
| DA&SSA  | 2,947    | $1,721 \cdot 10^{-1}$ |
| DA&GMDH | 2,565    | 1,061                 |
| DA&LSTM | 2,130    | 1,572                 |

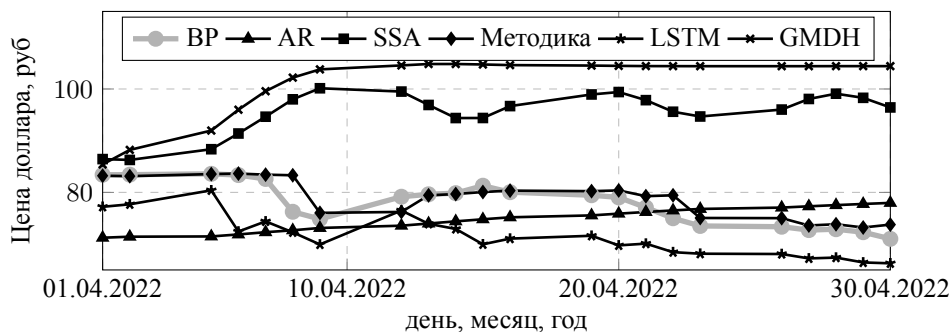


Рисунок 5 — Результаты прогнозирования BP курса доллар/рубль с помощью формальных методов прогнозирования BP (AR, SSA, GMDH, LSTM) и разработанной методики интеграции формальных методов прогнозирования BP и метода DA на основе AR(4)

более чем на порядок выше у последних (например, средняя погрешность прогнозирования курса доллар/рубль на основе AR-модели составляет 10,275, в то время как AR-модель, интегрированная в DA – 0,923).

2. Среди формальных методов прогнозирования, интегрированных в DA, наименьшую усредненную погрешность прогнозирования имеют AR-модели (0,923 и 1,961 для курсов доллар/рубль и евро/рубль соответственно). В тоже время усредненные погрешности прогнозов, вычисленные при использовании методов SSA, GMDH, LSTM, интегрированных в DA, не превосходят 3, что свидетельствует о возможности их использования на практике для прогнозирования курсов доллар/рубль, евро/рубль.



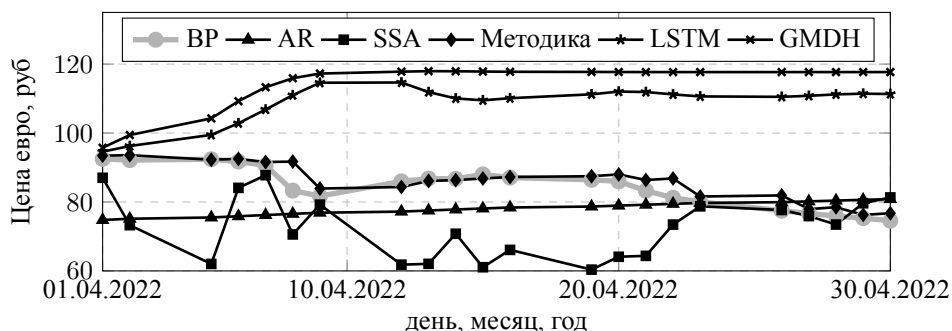


Рисунок 6 — Результаты прогнозирования BP курса евро/рубль с помощью формальных методов прогнозирования BP (AR, SSA, GMDH, LSTM) и разработанной методики интеграции формальных методов прогнозирования BP и метода DA на основе AR(2)

В **заключении** подведены основные итоги исследования:

1. Проведен анализ современного состояния методов прогнозирования и коррекции BP, который показал:

а) существует два типа ДС, порождающих прогнозируемые BP:

– ДС, описываемые известными математическими моделями, которые строятся на основе физических законов, определяющих эволюцию данного типа ДС во времени;

– ДС, описываемые феноменологическими моделями (точнее феноменологическими моделями, порожденных ими BP), которые строятся на основе анализа прогнозируемого BP, а потому оказываются зависящими от используемого метода анализа BP.

б) методы прогнозирования BP в зависимости от типа ДС разделяются на два независимых класса:

– методы прогнозирования BP, основанные на использовании математических моделей BP;

– формальные методы прогнозирования BP, основанные на использовании феноменологических моделей BP (авторегрессионная модель BP; модели, основанные на представлении прогнозируемых BP в виде линейных комбинации более простых BP (декомпозиции); самоорганизующиеся регрессионные модели; нейросетевые модели).

в) на практике актуальной является задача прогнозирования BP, у которых оказываются известными как ранее спрогнозированные, так и соответствующие наблюдаемые значения BP, порождённые обоими типами ДС, что потенциально, обеспечивает возможность уточнения прогнозов значений BP;

г) для уточнения прогнозируемых значений BP, порожденных ДС, описываемых математическими моделями, по мере поступления наблюдаемых его значений используется метод ассимиляции данных (Data Assimilation), однако,

отсутствует методика позволяющая использовать в методе Data Assimilation феноменологические модели ВР.

2. Разработана научно обоснованная методика интеграции формальных моделей ВР и метода Data Assimilation.

3. Разработан на языке Python 3 программный комплекс «Прогнозирование временных рядов», обеспечивающая практическое использование методики интеграции формальных методов прогнозирования ВР и метода Data Assimilation.

4. Подтверждена работоспособность методики интеграции формальных методов прогнозирования ВР и метода ассимиляции данных результатами прогнозирования ВР, порожденных ДС обоих типов, и получены оценки качества прогнозирования ВР разработанной методики интеграции формальных методов прогнозирования и метода DA.

**Перспективы дальнейшей разработки темы исследования** заключаются в изучении возможности разработки методики автоматизированного подбора параметров формальных методов прогнозирования.

## **Публикации по теме диссертации**

**Статьи, опубликованные в рецензируемых научных журналах и изданиях, определенных ВАК РФ и Аттестационным советом УрФУ:**

1. **Тимошенкова Ю.С.** Методика интеграции формальных методов прогнозирования временных рядов в Data Assimilation / **Тимошенкова Ю. С.**, Поршнев С. В., Сафиуллин Н. Т. // International Journal of Open Information Technologies. – 2022. – № 4 (10). – С. 15–23. (0,73 п.л. / 0,24 п.л.)

2. **Timoshenkova Y.** About Influence of Weight Initialization Algorithms on Accuracy of the Forecast with LSTM-net for Harmonic Signals / **Timoshenkova Y.**, Porshnev S., Safiullin N. // 2021 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBREIT). – 2021. – No. 9455014. – pp. 330-333. (0,47 п.л. / 0,21 п.л.) (Scopus)

3. **Timoshenkova Y.** The Dependence of the Sunspot Forecast Accuracy Using LSTM Networks from Number of Cycles in the Training Set / **Timoshenkova Y.**, Safiullin N. // 2020 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBREIT). – 2020. – No. 9117641. – pp. 452-455. (0,45 п.л. / 0,22 п.л.) (Scopus)

4. **Timoshenkova Y.** Application of data assimilation methods based on Kalman ensemble filters for the Lorenz attractor / **Timoshenkova Y.**, Porshnev S., Safiullin N. // International Conference on Numerical Analysis and Applied Mathematics (ICNAAM 2018). AIP Conference Proceedings. – 2019. – Vol. 2116. – No. 390017. – pp. 1-4. (0,33 п.л. / 0,11 п.л.) (Scopus, Web of Science)

5. **Timoshenkova Y.** The Required Number of Sunspot Cycles in the Training Set for a Better Accuracy of the Forecast with Artificial Neural Network / **Timoshenkova Y.**, Safiullin N. // 2019 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBREIT). – 2019. – No. 8736555. – pp. 248-251. (0,4 п.л. / 0,2 п.л.) (Scopus, Web of Science)

6. **Timoshenkova Y.** Application of data assimilation algorithms based on Kalman ensemble filters for the Lorenz attractor / **Timoshenkova Y.**, Porshnev S., Safiullin N. // CEUR Workshop Proceedings. – 2018. – Vol. 2274. – pp. 82-87. (0,3 п.л. / 0,1 п.л.) (Scopus)

7. **Timoshenkova Y.** On the Possibility of the Forecast Correction Lorenz Systeme Use Particle Filter / **Timoshenkova Y.**, Porshnev S., Safiullin N., Ponomareva O. // 2018 International Conference on Applied Mathematics & Computational Science (ICAMCS.NET). – 2018. – No. 8955771. – pp. 27-30. (0,32 п.л. / 0,08 п.л.) (Scopus, Web of Science)

8. **Timoshenkova Y.** On the possibility of correction of the forecasting of the Lorenz attractor dynamic characteristics using experimental data and data assimilation / **Timoshenkova Y.**, Porshnev S., Safiullin N. // Journal of Physics: Conference Series. – 2018. – Vol. 1053(1). – No. 012004. – pp. 1-7. (0,35 п.л. / 0,12 п.л.) (Scopus)

9. **Timoshenkova Y.** On the Possibility of the Forecast Correction for Inaccurate Observations Based on Data Assimilation / **Timoshenkova Y.**, Safiullin N. // 2017 Dynamics of Systems, Mechanisms and Machines (Dynamics). – 2017. – No. 8239519. – pp. 1-4. (0,33 п.л. / 0,16 п.л.) (Scopus, Web of Science)

#### **Патенты и свидетельства о регистрации программы:**

10. Свидетельство о государственной регистрации программы для ЭВМ № 2021681386 RU. Программная библиотека «Прогнозирование временных рядов на основе интеграции формальных методов и метода ассимиляции данных» / **Тимошенкова Ю.С.**, Поршнев С.В. – № 2021680432; дата поступления 07.12.2021; дата государственной регистрации в реестре программ для ЭВМ 21.12.2021 (Российская Федерация).

11. Свидетельство о государственной регистрации программы для ЭВМ № 2021618161 Программная библиотека «Прогнозирование временных рядов с помощью нейронной сети LSTM» / **Тимошенкова Ю.С.**, Поршнев С.В. – № 2021617124; дата поступления 07.05.2021; дата государственной регистрации в реестре программ для ЭВМ 24.05.2021 (Российская Федерация).

12. Свидетельство о государственной регистрации программы для ЭВМ № 2021615541 Программная библиотека «Формальные методы прогнозирования временных рядов» / **Тимошенкова Ю.С.**, Поршнев С.В. – № 2021614038; дата поступления 07.03.2021; дата государственной регистрации в реестре программ для ЭВМ 09.04.2021 (Российская Федерация).

#### **Другие публикации:**

13. **Тимошенкова Ю.С.** О возможности коррекции прогноза с помощью Data Assimilation / **Тимошенкова Ю.С.**, Сафиуллин Н.Т. // Актуальные проблемы развития технических наук: сборник статей. Москва: ООО "Эдитус". – 2018. – С. 120-125. (0,22 п.л. / 0,11 п.л.)

14. **Тимошенкова Ю.С.** О возможности коррекции прогноза по неточным наблюдениям на основе Data Assimilation / **Тимошенкова Ю.С.**, Сафиуллин Н.Т. // Наука и Мир. – 2017. – №5(45). – Т. 1. – С. 55-63. (0,44 п.л. / 0,22 п.л.)