

Санкт-Петербургский государственный университет

Прикладная математика и информатика

Производственная практика 2 (научно-исследовательская работа)

ИСПОЛЬЗОВАНИЕ МЕТОДА SSA В МАШИННОМ ОБУЧЕНИИ ДЛЯ  
ПРОГНОЗА ВРЕМЕННЫХ РЯДОВ

Выполнил:

Ежов Федор Валерьевич

группа 20.М03-мм



Научный руководитель:

к.физ.-мат.н., доцент

Голяндина Нина Эдуардовна

Кафедра Статистического Моделирования

Санкт-Петербург

2022

# Оглавление

|  |           |
|--|-----------|
| <b>Введение . . . . .</b>  | <b>3</b>  |
| <b>Глава 1. Алгоритм Singular Spectrum Analysis . . . . .</b>    | <b>4</b>  |
| 1.1. Этап 1. Построение траекторной матрицы (Вложение) . . . . . | 4         |
| 1.2. Этап 2. Singular Value Decomposition (SVD) . . . . .        | 5         |
| 1.3. Этап 3. Группировка первых $r$ собственных троек . . . . .  | 5         |
| 1.4. Этап 4. Диагональное усреднение . . . . .                   | 5         |
| <b>Глава 2. Архитектуры нейронных сетей . . . . .</b>            | <b>7</b>  |
| 2.1. Artificial Neural Network (ANN) . . . . .                   | 7         |
| 2.2. Recurrent neural network (RNN) . . . . .                    | 8         |
| 2.3. Long short-term memory(LSTM) . . . . .                      | 8         |
| 2.4. Gated Recurrent Unit (GRU) . . . . .                        | 9         |
| <b>Глава 3. Использование SSA в машинном обучении . . . . .</b>  | <b>10</b> |
| 3.1. Задача . . . . .  | 10        |
| 3.2. Подготовка данных . . . . .                                 | 10        |
| 3.3. Обучение нейронной сети . . . . .                           | 12        |
| 3.4. Прогнозирование . . . . .                                   | 13        |
| 3.5. Метрики . . . . .   | 13        |
| 3.6. Методика применения и сравнения методов . . . . .           | 13        |
| <b>Глава 4. Модельные данные . . . . .</b>                       | <b>16</b> |
| 4.1. Сумма двух синусов . . . . .                                | 16        |
| <b>Глава 5. Реальные данные . . . . .</b>                        | <b>43</b> |
| 5.1. Среднемесячные осадки в Индии . . . . .                     | 43        |
| 5.2. Данные Earth Orientation Parameters (EOP) . . . . .         | 51        |
| 5.3. Погода . . . . .  | 59        |
| <b>Заключение . . . . .</b>                                      | <b>66</b> |
| <b>Список литературы . . . . .</b>                               | <b>67</b> |

## Введение

В статье [1] было проведено сравнение обычного метода ANN и гибридного метода SSA-ANN на реальных данных. Итоги сравнения показали, что гибридные методы портят результаты прогноза из-за эффекта «data leaking». В предыдущих семестрах, было продемонстрирован обратный эффект на тех же данных, а также модельных. Количество методов пополнилось рекуррентными нейронными сетями RNN [2], GRU [3], LSTM [4], а также их гибридными аналогами SSA-RNN, SSA-GRU, SSA-LSTM. В предыдущем семестре была рассмотрена методика исследования сравнения методов, которая показалась весьма успешной, особенно на реальных данных среднемесячных осадков в Индии. Анализ реальных данных и анализ модельного ряда приводил к схожим выводам, в связи с чем возникла идея при анализе реальных данных сначала строить их модель и потом исследовать методы на модели. Модель удобна тем, что в ней можно менять параметры, можно проверять устойчивости результатов, и пр.

В этом семестре была поставлена задача: улучшить методику, применить к новым данным, добавить прогноз по SSA. Методику отработать на модельных примерах. Использовать модельные примеры для объяснения результатов сравнения.

Были рассмотрены данные ЕОР, данные характеристики погоды в Санкт-Петербурге, а также модельные данные. В методику добавился прогноз по методу SSA, отображение результатов прогнозов, отображение выделения сигнала с помощью метода SSA, проверка устойчивости результатов.

Структура работы следующая. В главе 1 описан алгоритм препроцессинга Singular Spectrum Analysis (SSA). В главе 2 подробно описаны все методы прогнозирования, используемые в работе. В главе 3 описана основные методы и схемы, применяемые в экспериментах в области ML, описанных в главах 4 и 5. В главе 4 описаны эксперименты на модельных данных. В главе 5 описаны эксперименты на реальных данных.

## Глава 1

### Алгоритм Singular Spectrum Analysis

Метод SSA используется для разложение исходного ряда в сумму рядов, которые легко интерпретировать и понять их поведение. Обычно исходный ряд раскладывается в сумму трех рядов: тренд — медленно меняющаяся компонента, сезонность — циклическая компонента с фиксированным периодом и шум. Информацию про базовый алгоритм SSA и связанные с методом фундаментальные понятия можно найти в книге «Analysis of time series structure: SSA and related techniques» [5].

Алгоритм SSA состоит из четырех этапов:

1. Построение траекторной матрицы (Вложение).
2. SVD.
3. Группировка первых  $r$  собственных троек.
4. Диагональное усреднение.

Рассмотрим каждый этап подробнее.

Пусть  $\mathbf{X}_N = (x_1, \dots, x_N)$  — временной ряд, где  $N > 3$ . Также будем предполагать, что найдется хоть одно  $x_i \neq 0$ , то есть ряд не нулевой. Обычно считается, что  $x_i = f(i\Delta)$  для некоторой функции  $f(t)$ , где  $t$  — время, а  $\Delta$  — некоторый временной интервал.

#### 1.1. Этап 1. Построение траекторной матрицы (Вложение)

Выберем целое  $L$  — длина окна, такое что  $1 < L < N$ . Тогда  $K = N - L - 1$ . Построим вектора  $X_i = (x_i, \dots, x_{i+L-1})^T$ , для  $1 \leq i \leq K$ . Составим из векторов  $X_i$  траекторную матрицу:

$$\mathbf{X} = [X_1 : \dots : X_K] = \begin{pmatrix} x_1 & x_2 & x_3 & \cdots & x_K \\ x_2 & x_3 & x_4 & \cdots & x_{K+1} \\ x_3 & x_4 & x_5 & \cdots & x_{K+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & x_{L+2} & \cdots & x_N \end{pmatrix}.$$

Получили матрицу  $\mathbf{X}$  размерностью  $L \times K$ , составленную из пересекающихся частей исходного временного ряда. Можно заметить, что на побочных диагоналях стоят одинаковые числа, такая матрица называется ганкелевой. Существует взаимно-однозначное соответствие между ганкелевыми матрицами  $L \times K$  и рядами длиной  $N = L + K - 1$ .

Операцию получения из ряда  $\mathbf{X}_N$  траекторную матрицу  $\mathbf{X}$  обозначим:

$$\mathbf{X} = \mathcal{T}_L(\mathbf{X}_N),$$

обратную операцию обозначим, как  $\mathcal{T}^{-1}$  соответственно.

## 1.2. Этап 2. Singular Value Decomposition (SVD)

На данном этапе применяется метод SVD к траекторной матрице  $\mathbf{X}$ . Пусть  $\mathbf{S} = \mathbf{X}\mathbf{X}^T$  и  $\lambda_1 > \dots > \lambda_L$  — собственные числа матрицы  $\mathbf{S}$ ,  $U_1, \dots, U_L$  — ортонормированная система базисных векторов, соответствующих собственным числам. Обозначим  $V_i = \frac{\mathbf{X}^T U_i}{\sqrt{\lambda_i}}$  и  $d = \max\{i : \lambda_i > 0\}$ . Тогда сингулярное разложение матрицы  $\mathbf{X}$  запишется следующим образом:

$$\mathbf{X} = \mathbf{X}_1 + \dots + \mathbf{X}_d, \text{ где } \mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T,$$

Набор  $(\sqrt{\lambda_i}, U_i, V_i^T)$  будем называть  $i$ -й собственной тройкой.

## 1.3. Этап 3. Группировка первых $r$ собственных троек

На этапе группировки из всех значений  $\{1 \dots d\}$  берутся первые  $r$ . Пусть,  $I = \{1, \dots, r\}$ , тогда результирующая матрица соответствующая группе  $I$  имеет вид:  $\mathbf{X}_I = \mathbf{X}_1 + \dots + \mathbf{X}_r$ .

## 1.4. Этап 4. Диагональное усреднение

Пусть  $\mathbf{Y}$  — матрица  $L \times K$  и  $L < K$ .  $y_{ij}$  — элементы матрицы, где  $1 \leq i \leq L$ ,  $1 \leq j \leq K$ ,  $N = L + K - 1$ . Диагональное усреднение преобразует матрицу  $\mathbf{Y}$  в ряд  $g_0, \dots, g_{N-1}$  по формуле:

$$g_k = \begin{cases} \frac{1}{k+1} \sum_{m=1}^{k+1} y_{m,k-m+2} & , \text{ для } 0 \leq k < L-1, \\ \frac{1}{L} \sum_{m=1}^L y_{m,k-m+2} & , \text{ для } L-1 \leq k < K, \\ \frac{1}{N-k} \sum_{m=k-K+2}^{N-K+1} y_{m,k-m+2} & , \text{ для } K \leq k < N. \end{cases}$$

Применяя диагональное усреднение к результирующей матрице группы  $I$ , получаем ряд  $\widehat{\mathbf{F}} = (f_1, \dots, f_{N-1})$ . Ряд  $\widehat{\mathbf{F}}$  назовем оценкой сигнала, полученной с помощью SSA. Процедуру выделения сигнала с помощью SSA обозначим как:

$$\widehat{\mathbf{F}} = \text{SSA}_{L,r}(\mathbf{F}),$$

где  $L$  — длина окна в SSA,  $r$  — количество первых собственных троек, участвующие в построении  $\widehat{\mathbf{F}}$ .

## Глава 2

# Архитектуры нейронных сетей

В данной главе рассматриваются нейронные сети используемые в работе. Всего рассмотрено четыре архитектуры: одна линейная (называемая ANN), и три рекуррентных нейронных сети — RNN, LSTM, GRU.

Каждая нейронная сеть имеет параметры и гиперпараметры. Архитектура или модель нейронной сети определяется с точностью до гиперпараметров. Термин модель означает конкретный нейросетевой метод (например ANN или LSTM), не стоит путать его с моделью временного ряда. Нейронные сети оптимизируют (подгоняют) параметры для решения задачи прогнозирования сигнала ряда во время процесса «обучения». Процессом обучения называется оптимизация параметров нейронной сети с помощью градиентного метода «Back Propagation». Нейронные сети как и все методы машинного обучения «обучаются с учителем», то есть на парах «признаки — предсказываемые значения».

### 2.1. Artificial Neural Network (ANN)

Artificial Neural Network (ANN) включает в себя входной слой, ряд скрытых слоев и выходной слой, каждый слой содержит несколько узлов. Считается, что ANN с одним скрытым слоем обеспечивает достаточную сложность для моделирования нелинейных взаимосвязей данных. ANN в данной работе формализуется следующим образом:

Входные данные, на которых модель учится делать предсказания:

$$X = [x_1, \dots, x_T].$$

Предсказания модели:

$$\hat{Y} = [\hat{y}_1, \dots, \hat{y}_R].$$

Формула, описывающая модель:

$$y_k = \phi_2 \left( \sum_{j=1}^h w_{jk}^{(2)} \phi_1 \left( \sum_{i=1}^T w_{ij}^{(1)} x_i + \theta_j^{(1)} \right) + \theta_k^{(2)} \right), k = [1, \dots, R],$$

где  $T$  — размер входного вектора, на котором выполняется прогноз.  $h$  — размер скрытого слоя.  $w$  и  $\theta$  — веса модели.  $\phi$  — функции активации.  $R$  — размер выходного вектора-прогноза.

Список некоторых функций активации:

1. Линейная функция активации:  $\phi(x) = x$ .

2. Сигмоида:  $\phi(x) = \frac{1}{1 + e^{-x}}$ .

3. ReLU:  $\phi(x) = \begin{cases} 0, & x < 0, \\ x, & x \geq 0. \end{cases}$

## 2.2. Recurrent neural network (RNN)

Модель recurrent neural network (RNN), использует рекурсию в своих архитектурах для решения задач, где данные содержатся в некоторой последовательности (например, временные ряды, текстовые задачи и др.). В RNN присутствует вектор скрытого слоя, служащий для «накопления» информации.

$X = [x_1, \dots, x_T]$  — вектор входных данных.

$J_t = [j_1, \dots, j_h]$  — вектор скрытого слоя в момент  $t$ .

$\hat{Y}_T = [\hat{y}_1, \dots, \hat{y}_R]$  — вектор выходных данных в момент  $T$ .

Следующие формулы описывают модель:

$$J_t = f_j(Ux_t + \theta_1 + WJ_{t-1} + \theta_2),$$

$$\hat{Y}_T = f_y(VJ_T + \theta_3),$$

где  $t = [1, \dots, T]$ ,  $U, V, W, \theta_i$  — веса модели, вычисляющиеся в процессе обучения,  $f_h, f_y$  — функции активации (в большинстве случаев сигмоиды).

## 2.3. Long short-term memory(LSTM)

Long short-term memory (LSTM) — разновидность рекуррентных моделей, с добавлением второго скрытого слоя, используемого для «долгосрочной» памяти.

$X = [x_1, \dots, x_T]$  — вектор входных данных.

$J_t = [j_1, \dots, j_h]$  — вектор скрытого слоя в момент  $t$ .

$C_t = [c_1, \dots, c_J]$  — вектор скрытого слоя в момент  $t$ .

$\hat{Y}_T = [\hat{y}_1, \dots, \hat{y}_R]$  — вектор выходных данных в момент  $T$ .

Следующие формулы описывают модель:

$$\begin{aligned} f_t &= \sigma(W_f \cdot [J_{t-1}, x_t] + b_f), \quad i_t = \sigma(W_i \cdot [J_{t-1}, x_t] + b_i), \\ \tilde{C}_t &= \tanh(W_c \cdot [J_{t-1}, x_t] + b_c), \quad C_t = f_t * C_{t-1} + i_t * \tilde{C}_t, \\ o_t &= \sigma(W_o \cdot [J_{t-1}, x_t] + b_o), \quad J_t = o_t * \tanh(C_t), \\ \hat{Y}_T &= f(VJ_T + b_0), \end{aligned}$$

где  $t = [1, \dots, T]$ ,  $W, b$  — веса модели,  $\sigma(x)$ ,  $\tanh(x)$ ,  $f(x)$  — функции активации.

Оператор  $*$  — производит поэлементное умножение.

## 2.4. Gated Recurrent Unit (GRU)

Gated Recurrent Unit (GRU) — модель похожая на LSTM, но без дополнительного скрытого слоя.

$X = [x_1, \dots, x_T]$  — вектор входных данных.

$J_t = [j_1, \dots, j_h]$  — вектор скрытого слоя в момент  $t$ .

$\hat{Y}_T = [\hat{y}_1, \dots, \hat{y}_R]$  — вектор выходных данных в момент  $T$ .

Следующие формулы описывают модель:

$$\begin{aligned} z_t &= \sigma(W_z \cdot [J_{t-1}, x_t] + b_z), \quad r_t = \sigma(W_r \cdot [J_{t-1}, x_t] + b_r), \\ \tilde{J}_t &= \tanh(W \cdot [r_t * J_{t-1}, x_t] + b), \quad J_t = (1 - z_t) * J_{t-1} + z_t * \tilde{J}_t, \\ Y_T &= f(VJ_T + b_0), \end{aligned}$$

где  $t = [1, \dots, T]$ ,  $W, b$  — веса модели,  $\sigma(x)$ ,  $\tanh(x)$ ,  $f(x)$  — функции активации.

Оператор  $*$  — производит поэлементное умножение.

## Глава 3

# Использование SSA в машинном обучении

### 3.1. Задача

Рассмотрим  $Z_N$  — временной ряд длины  $N$  и задачу: с помощью модели некоторой нейронной сети  $f$  на основе  $T$  последовательных точек ряда  $Z_N$ , предсказать следующие  $R$  точек ряда.

$$[\hat{z}_{i+T+1}, \dots, \hat{z}_{i+T+R}] = f([z_{i+1}, \dots, z_{i+T}]).$$

Считаем, что  $Z_N = S_N + \xi_N$ , где  $S_N$  — сигнал,  $\xi_N$  — шум, случайный процесс с нулевым мат. ожиданием. Тогда возникает идея, подавать на вход методу  $f$  не сам ряд, а оценку сигнала  $\hat{S}_N$ , полученную с помощью метода SSA. Методы  $f$ , которым на вход подается  $\hat{S}_N$  называем гибридными. Разница между обычными и гибридными методами заключается только в данных, поступающих на вход. Таким образом решение поставленной задачи, можно разбить на несколько частей: подготовка данных, обучение методов, прогнозирование.

### 3.2. Подготовка данных

$Z_N$  — изначальный временной ряд длиной  $N$ . Можем представить ряд в виде траекторной матрицы для длины окна  $T + R$ :

$$\mathbf{Z} = \mathcal{T}_{T+R}(Z_N) = \left( \begin{array}{cccc|cccc} z_1 & z_2 & \cdots & z_T & z_{T+1} & \cdots & z_{T+R-1} & z_{T+R} \\ z_2 & z_3 & \cdots & z_{T+1} & z_{T+2} & \cdots & z_{T+R} & z_{T+R+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ z_{N-T-R+1} & z_{N-T-R+2} & \cdots & z_{N-R} & z_{N-R+1} & \cdots & z_{N-1} & z_N \end{array} \right).$$

Матрица  $\mathbf{Z}$  имеет размерность  $(N - T - R + 1) \times (T + R)$ . Левую часть матрицы  $\mathbf{Z}$  обозначим  $\mathbf{Z}^x$ , правую —  $\mathbf{Z}^y$ . Разобьем матрицу по строчкам на три части: train, val, test. Пусть  $\tau$ ,  $v$  и  $t$  номера последних строчек в каждой соответствующей части. Обозначим часть матрицы  $\mathbf{Z}$  с  $a$  по  $b$  строчку и с  $c$  по  $d$  столбец как  $\mathbf{Z}_{a,b}^{(c,d)}$ . Тогда train,

val, test записываются как:  $\mathbf{Z}_{train} = \mathbf{Z}_{1,\tau}^{(1,T+R)}$ ,  $\mathbf{Z}_{val} = \mathbf{Z}_{\tau+T+R,v}^{(1,T+R)}$ ,  $\mathbf{Z}_{test} = \mathbf{Z}_{v+T+R,t}^{(1,T+R)}$ . В этих же обозначениях  $\mathbf{Z}^x = \mathbf{Z}_{1,t}^{(1,T)}$ ,  $\mathbf{Z}^y = \mathbf{Z}_{1,t}^{(T+1,T+R)}$ .

### SSA-preprocessing

В этом разделе опишем алгоритм предобработки SSA для тренировочной выборки. Пусть  $L, r$  — гиперпараметры, описанные в разделе «1.1. Алгоритм SSA».

1. Преобразуем train часть матрицы  $\mathbf{Z}$  во временной ряд  $\tilde{\mathbf{Z}} = \mathcal{T}^{-1}(\mathbf{Z}_{train})$ .
2. Получим ряд  $\hat{\mathbf{Z}} = SSA_{L,r}(\tilde{\mathbf{Z}})$ .
3. Получаем траекторную матрицу  $\hat{\mathbf{Z}} = \mathcal{T}_{T+R}(\hat{\mathbf{Z}})$ . Матрица  $\hat{\mathbf{Z}}$  и будет результатом работы предобработки SSA для тренировочной выборки.

Предобработка SSA для валидационной или тестовой выборок отличается от предыдущей, ввиду разных предназначений выборок. В отличие от тренировочной выборки, о которой все известно, считается, что о валидационной и тестовой выборках нет информации. В этих случаях SSA-обработку следует применять так, чтобы предыдущие значения ряда не получали информации от будущих («заглядывание в будущее»).

Пусть  $Z_{b,e} = [z_b, z_{b+1}, \dots, z_e]$  подряд ряда  $Z$ , где  $b$  — начальный индекс,  $e$  — конечный индекс. Пусть  $p$  — тоже индекс ряда, такой что  $b < p < e$ . Следующий алгоритм описывает процедуру получения ряда  $Z_{p+1,e}$ , обработанного с помощью SSA без «заглядывание в будущее»:

1. Пусть есть ряд  $Z_{b,e}$  и задано  $p$ . Тогда  $Q = e - p$  — размер ряда  $Z_{p+1,e}$ . Пусть  $\hat{\mathbf{Z}}_Q = (\hat{z}_1, \dots, \hat{z}_Q)$  — ряд размера  $Q$ .
2. Для каждого  $i = [1, \dots, Q]$  получим

$$\hat{\mathbf{Z}}'_{b+i-1,p+i} = SSA_{L,r}(Z_{b+i-1,p+i}).$$

Присвоим значение последнего элемента полученного ряда  $\hat{z}'_{p+i}$  значению ряда  $\hat{\mathbf{Z}}_Q$  с соответствующим индексом,  $\hat{z}_i = \hat{z}'_{p+i}$ .

3. Получили ряд  $\hat{\mathbf{Z}}_Q$  размера  $Q$ , значения которого являются значениями ряда  $Z_{p+1,e}$ , обработанные с помощью SSA без «заглядывания в будущее».

Обозначим процедуру получения  $\widehat{\mathbf{Z}}_Q = \text{SSA}^{(p)}(\mathbf{Z}_{b,e})$ . Заметим, что для предобработки валидационной и тестовой выборок логично взять  $Q = \tau + T$  ( $\tau + T$  — длина ряда  $\mathcal{T}^{-1}(\mathbf{Z}_{train}^x)$ , который является тренировочной выборкой представленным в виде ряда). Алгоритм предобработки для валидационной выборки запишется следующим образом:

1. Запишем  $\mathbf{Z}_{1,v}^{(1,T+1)}$  как  $\mathbf{Z}_{1,v+T+1}$ .
2. Выберем  $p = \tau + T + 1$ .
3. Получим  $\widehat{\mathbf{Z}}_Q = \text{SSA}^{(p)}(\mathbf{Z}_{1,v+T+1})$ .
4. Перейдем обратно к траекторной матрице  $\widehat{\mathbf{Z}}_{val} = \mathcal{T}_{T+1}(\widehat{\mathbf{Z}}_Q)$ , которая будет результатом предобработки SSA для валидационной выборки.

Размерность  $\widehat{\mathbf{Z}}_{val}$  будет совпадать с размерностью  $\mathbf{Z}_{val}$ . Аналогичным образом строится тестовая выборка.

### 3.3. Обучение нейронной сети

Оптимизация параметров нейросетевой модели проводится с помощью процедуры обратного распространения ошибки на тренировочной выборке. Модель учится по  $\mathbf{Z}_{train}^x$  предсказывать  $\mathbf{Z}_{train}^y$ . Эпоха — цикл прохода всех строчек из тренировочной выборки матрицы  $\mathbf{Z}_{train}$  в процессе обучения. Валидационная выборка используется для оценки точности модели и оптимизации параметров. Переобучение предотвращается с помощью механизма ранней остановки. Данный алгоритм останавливает обучение, если ошибка на валидационной выборке растет на протяжение некоторого количества эпох.

Перед началом обучения нужно выбрать гиперпараметры модели и количество эпох. Так как в обучение используется алгоритм ранней остановки, возьмем заведомо большое количество эпох.

Алгоритм обучения модели после выбора архитектуры:

1. Инициализация модели со случайными весами.
2. На тренировочной выборке  $\mathbf{Z}_{train}$  оптимизируются веса  $w, \theta$  с заданным количеством эпох. Модель учится по данным строчкам  $\mathbf{Z}_{train}^x$  предсказывать соответствующие строчки  $\mathbf{Z}_{train}^y$ . Для каждой  $i$ -ой эпохи считается  $\varepsilon_i$  — ошибка на валидационной выборке. Для валидационной выборки  $\mathbf{Z}_{val}^x$  строится прогноз  $\widehat{\mathbf{Z}}_{val}^y$ . Ошибка  $\varepsilon_i$  получается сравнением  $\widehat{\mathbf{Z}}_{val}^y$  с  $\mathbf{Z}_{val}^y$  по какой-нибудь метрике (например MSE).

3. Срабатывает механизм ранней остановки. Присваиваем параметрам модели те, которые были получены при минимальной ошибке  $\min(\varepsilon_i)$ .

### 3.4. Прогнозирование

После того как модель обучена, можно перейти к прогнозированию точек ряда.

1. Возьмем  $\mathbf{Z}_{test}^x$  и  $\mathbf{Z}_{test}^y$ .
2. Представим  $\mathbf{Z}_{test}^x = [Z_{test}^{x,1} : \dots : Z_{test}^{x,Q}]^T$ , где  $Q$  — количество строчек в тестовой матрицы  $\mathbf{Z}_{test}$ .
3. Для каждой строчки матрицы  $\mathbf{Z}_{test}^x$  получаем прогноз с помощью обученной модели. Запишем результат прогноза как матрицу  $\widehat{\mathbf{Z}}^y = [\widehat{Z}^{y,1} : \dots : \widehat{Z}^{y,Q}]^T$ .
4. Далее можно сравнить  $\widehat{\mathbf{Z}}^y$  с  $\mathbf{Z}_{test}^y$  по какой-нибудь метрике.

### 3.5. Метрики

С помощью метрик MSE и RMSE можно мерить размер ошибки полученного прогноза.

$$\text{MSE}(\mathbf{Z}_{test}^y, \widehat{\mathbf{Z}}^y) = \frac{1}{Q} \text{diag} \left( (\mathbf{Z}_{test}^y - \widehat{\mathbf{Z}}^y)(\mathbf{Z}_{test}^y - \widehat{\mathbf{Z}}^y)^T \right),$$

$$\text{RMSE}(\mathbf{Z}_{test}^y, \widehat{\mathbf{Z}}^y) = \sqrt{\text{MSE}(\mathbf{Z}_{test}^y, \widehat{\mathbf{Z}}^y)}.$$

### 3.6. Методика применения и сравнения методов

В этом разделе описана методика применения и сравнения обычных и гибридных прогнозов, а также прогнозов полученных с помощью SSA.

**Прогноз по SSA** Подберем для метода SSA параметры  $L$  и  $r$ , которые дают наилучшие результаты на валидационной выборке. Для параметров для  $L$  и  $r$  зададим сетку, по которой будем перебирать комбинации параметров, оценивая ошибку прогноза на один шаг вперед с помощью метрики RMSE на валидационной выборке. Для пяти лучших пар параметров построим прогноз на тестовой выборке и получим оценку ошибки. Каждую ошибку отобразим на графике с помощью горизонтальной прямой.

**Выбор параметров SSA в гибридных моделях** Прежде чем сравнивать обычные и гибридные модели, нужно выбрать параметры SSA для гибридных моделей. Сделать это можно двумя способами:

1. Подобрать параметры на основе SSA-анализа тренировочной части ряда или исходя из других источников информации.
2. Перебрать параметры для метода SSA и выбрать лучшие параметры, основываясь на точности прогноза методом.

**Обычные и гибридные методы** Для гибридных моделей фиксируем гиперпараметры SSA  $L$  и  $r$ , выбранные заранее. Будем сравнивать методы по сетке гиперпараметров  $T$  — размер входного вектора и  $h$  — размер скрытого слоя нейронной сети (модели нейронных сетей подробно описаны в главе 2). Каждую пару  $T$  и  $h$  можно представить как ячейку в таблице 3.1. Для каждой ячейки получаем ошибку по метрике RMSE на тестовой выборке. Ошибку можно усреднить по столбцам или по строкам, и построить по графику зависимости ошибки от гиперпараметров  $T$  или  $h$  для каждого метода соответственно. Методы можно будет сравнить на графиках.

Таблица 3.1. Сетка гиперпараметров.

|          | $T_1$        | $T_2$        | $\dots$  | $T_n$        |
|----------|--------------|--------------|----------|--------------|
| $h_1$    | $(T_1, h_1)$ | $(T_2, h_1)$ | $\dots$  | $(T_n, h_1)$ |
| $h_2$    | $(T_1, h_2)$ | $(T_2, h_2)$ | $\dots$  | $(T_n, h_2)$ |
| $\vdots$ | $\vdots$     | $\vdots$     | $\ddots$ | $\vdots$     |
| $h_m$    | $(T_1, h_m)$ | $(T_2, h_m)$ | $\dots$  | $(T_n, h_m)$ |

**Проверка устойчивости результатов** Для того чтобы оценить устойчивость результатов, фиксируется небольшая сетка параметров  $T$  и строится таблица аналогичная 3.1. Проводится  $n$  реализаций для каждой ячейки из таблицы. Так как параметры в нейронных сетях инициализируется случайно, то каждый раз получается разный результат. Для каждого  $T$  полученные результаты заносятся на отдельный график зависимости ошибки от параметра  $h$  в виде точек. По этим графикам можно оценить устойчивость результатов.

**Отображение результатов** Чтобы объяснить успешность или неуспешность прогноза и в целом посмотреть на картину прогнозирования будем строить график прогноза на фоне правильных ответов.

**Восстановление ряда с помощью SSA** Для оценки корректности использования предобработки SSA будем строить график восстановленного сигнала с помощью SSA на фоне временного ряда.

## Глава 4

### Модельные данные

В данном разделе на модельных рядах проведены эксперименты с целью, показать влияние различных особенностей временных рядов и способов выбора параметров методов на результаты точности прогнозов. В разделе рассмотрено:

1. Влияние длины временного ряда.
2. Влияние выбора параметра  $r$ .
3. Влияние наличия шума во временном ряде.
4. Ряд с плохо отделимым сигналом.

#### 4.1. Сумма двух синусов

Рассмотрим следующий ряд (рис. 4.1), состоящий из элементов:

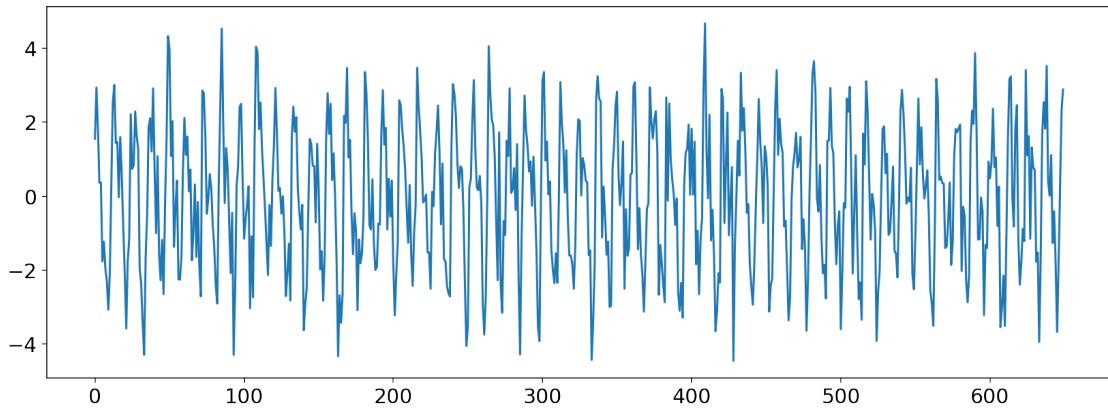
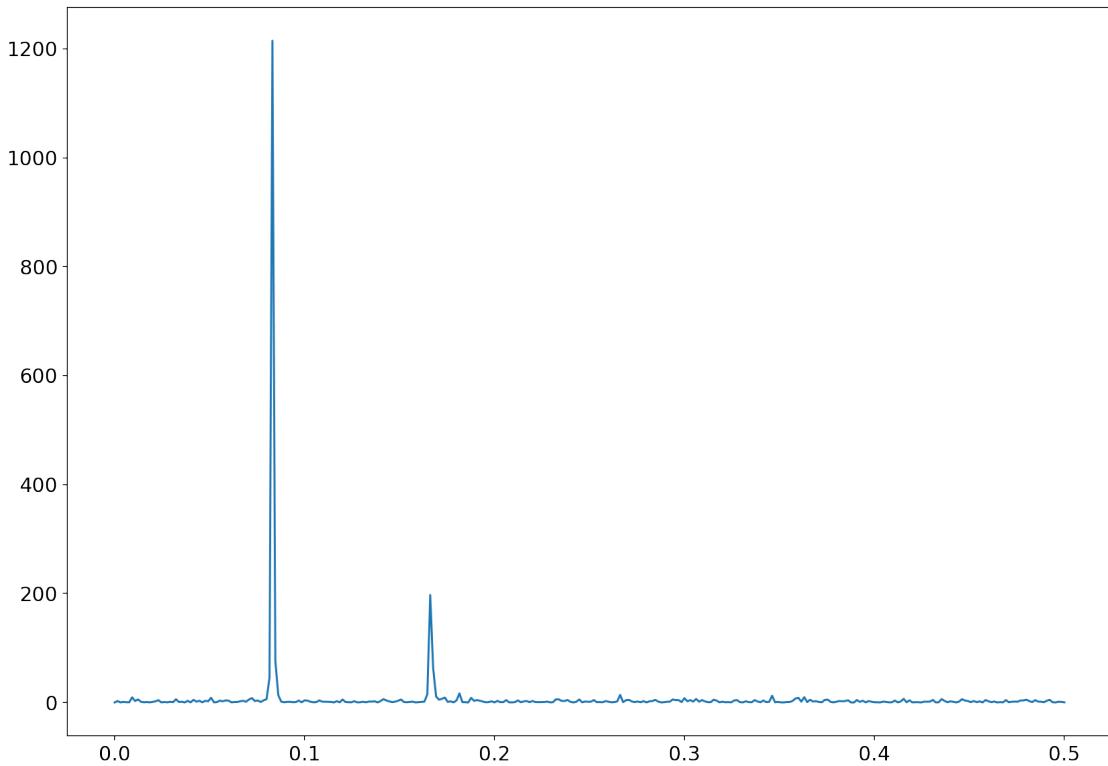
$$z_i = \left( \sin\left(2\pi \frac{i}{6}\right) + 2 \cdot \sin\left(2\pi \frac{i}{12}\right) \right) + \kappa \varepsilon_i,$$

где  $\varepsilon_i \in N(0, 1)$ ,  $\kappa$  задает размер шума в ряде.

Обозначим  $X_N$  ряд с  $\kappa = 0.3$ , а  $Z_N$  с  $\kappa = 1.5$ . Далее в этой главе рассматриваем ряды:  $X_{650}$ ,  $Z_{650}$ .

В данном ряде легко оценить сигнал, это можно увидеть на периодограмме (рис. 4.2). Шум в ряде не очень большой, хотя видно что он присутствует. Из формулы видно, что ряд стоит из двух синусов с периодами 12 и 6, общий период ряда 12. Далее в экспериментах будем перебирать параметры  $T$  и  $L$  по сетке с шагом кратным 12. Будем считать аналитически верными параметрами для этого ряда  $L = 175$ ,  $r = 4$ . Так как ранг ряда равен 4, а  $L = 175$  удовлетворяет асимптотической разделимости.

В экспериментах будем разбивать ряд  $Z_{650}$  на тренировочную, валидационную, тестовую выборки по 350, 150, 150 точек соответственно.

Рис. 4.1. Ряд  $Z_{650}$ .Рис. 4.2. Периодограмма ряда  $Z_{650}$ .

#### 4.1.1. Влияние выбора параметра $r$

Поставим задачу сравнить обычные и гибридные методы в случае разным выбранным параметром  $r$  в гибридных методах. Будем проводить сравнение на ряде  $Z_{650}$ . Разобьем наше исследование на три части: правильно выбранное  $r = 4$ , недостаточно большое  $r = 2$ , слишком большое  $r = 6$ . Каждую часть будем исследовать с помощью методологии, описанной в главе 3.6.

## Прогноз по SSA

Сравним точность прогнозирования методом SSA при разных параметрах  $r$ . Зададим следующую сетку параметров  $L = \{12, 24, \dots, 168\}$ ,  $r = \{2, 4, 6\}$ . Посмотрим на результаты на рисунке 4.3. На графике видно, что наилучшие результаты достигаются при  $r = 4$ , худшие результаты достигаются при  $r = 6$ .

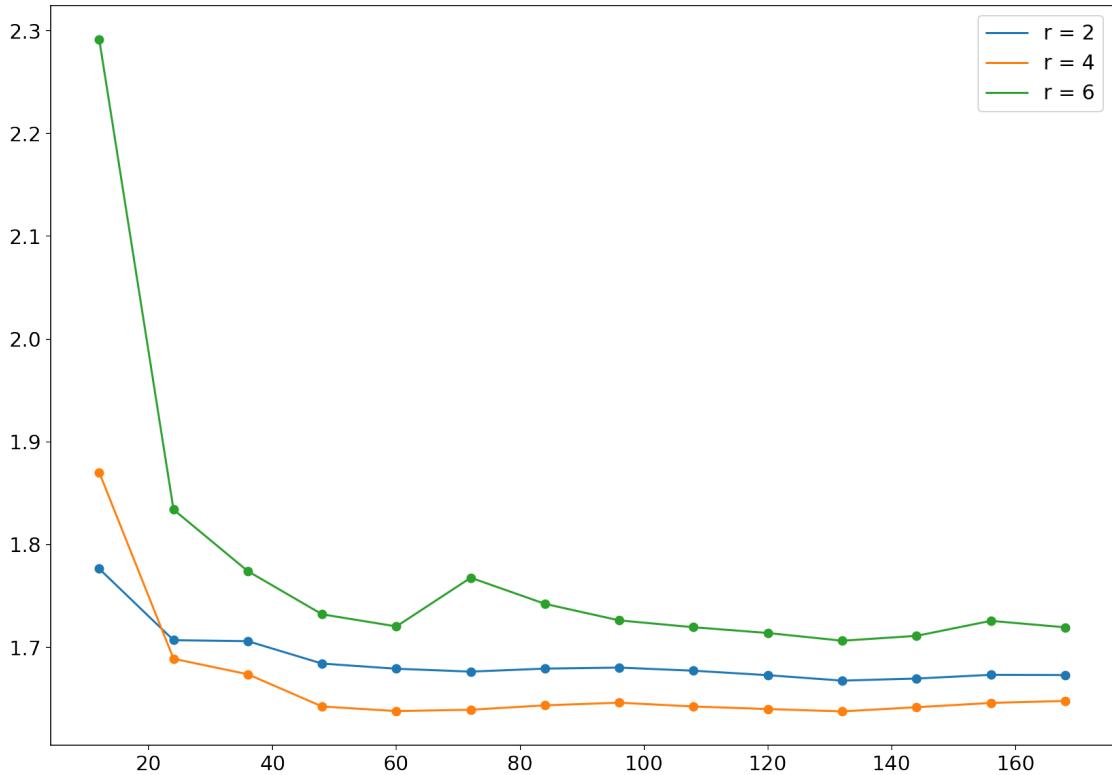


Рис. 4.3. Прогнозирование с помощью метода SSA.

Выделим лучшее  $L$  для каждого  $r$ . Исходя из графика для всех  $r$  это будет  $L = 132$ . Далее будем рассматривать лучшие комбинации гиперпараметров SSA вместе с  $L = 175$  и  $r = \{2, 4, 6\}$ .

## Восстановление SSA

Посмотрим, как метод SSA восстанавливает тренировочную выборку для выбранных пар на графиках ниже. На графике 4.4 видно, что метод не восстанавливает ряд полностью. У оценки сигналы маленькая амплитуда. На графике 4.5, где  $r = 4$ , метод SSA очень хорошо аппроксимирует сигнал. А для  $r = 6$  на графике 4.6 видно, что в оценку сигнала просочился шум, который мешает точность решать задачу прогнозирования. На графиках видно, что разницы между восстановленными рядами нету, далее

будем рассматривать только параметры  $r = \{2, 4, 6\}$  и  $L = 175$ .

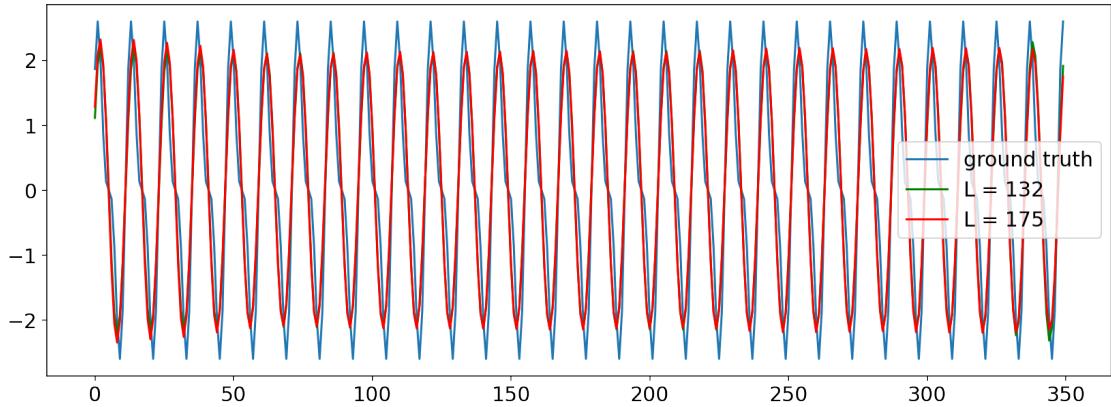


Рис. 4.4. Восстановление тренировочной выборки с помощью метода SSA.  $r = 2$

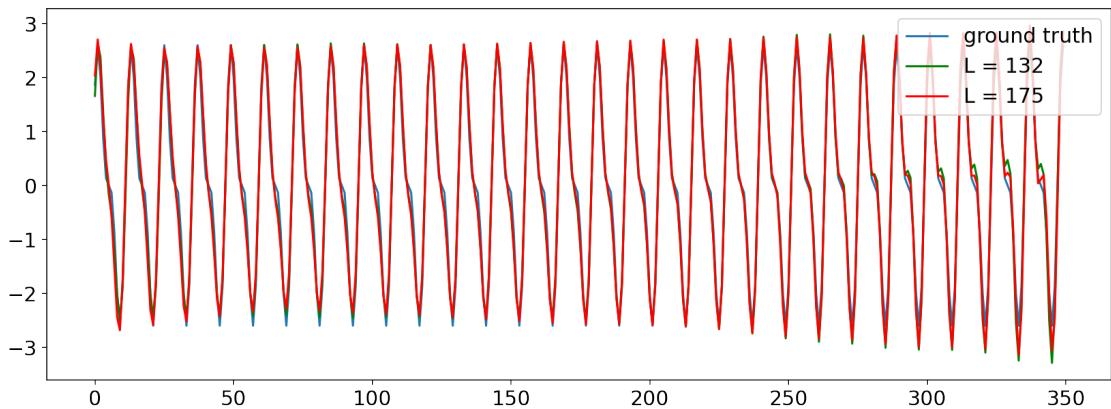


Рис. 4.5. Восстановление тренировочной выборки с помощью метода SSA.  $r = 4$

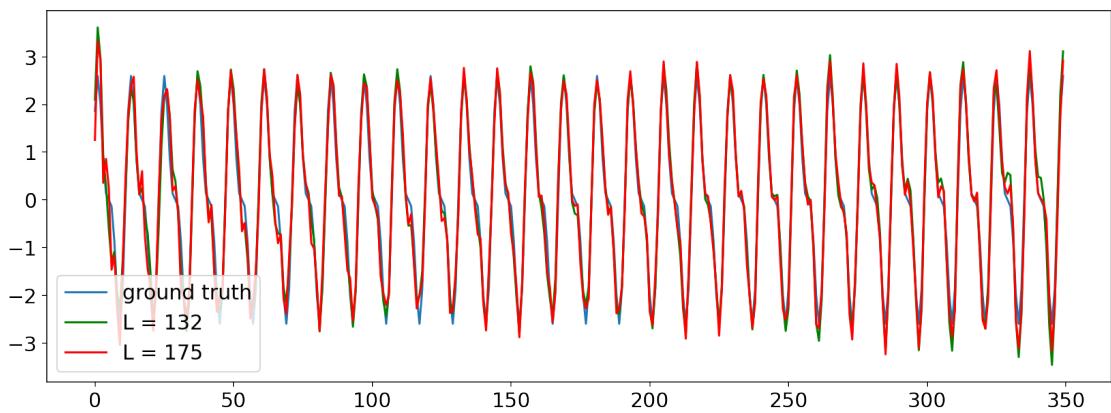


Рис. 4.6. Восстановление тренировочной выборки с помощью метода SSA.  $r = 6$

## Сравнение обычных и гибридных методов

Пусть задана следующая сетка параметров:  $T = \{12, 24, \dots, 144\}$ ,  $h = \{10, 25, \dots, 100\}$ .

Для метода SSA в гибридных моделях возьмем пары параметров, выбранные заранее.

На графиках 4.7, 4.8, 4.9, 4.10, 4.11, 4.12 представлены результаты сравнения по сетке параметров, заданной выше. Можно заметить, что для  $r = 2$  и  $r = 4$  гибридные методы явно лучше, чем обычные. Для  $r = 6$  результаты начинают смешиваться. В целом для  $r = 2$  и  $r = 4$  ситуация выглядит похоже, какие-то гибридные методы показывают результаты лучше для  $r = 2$ , другие для  $r = 4$ . Также для  $r = 4$  кривые выглядят более горизонтально, что уменьшает зависимость ошибки от выбора гиперпараметров нейронной сети. Для точности прогнозов методом SSA все осталось также, как и в предыдущей части эксперимента.

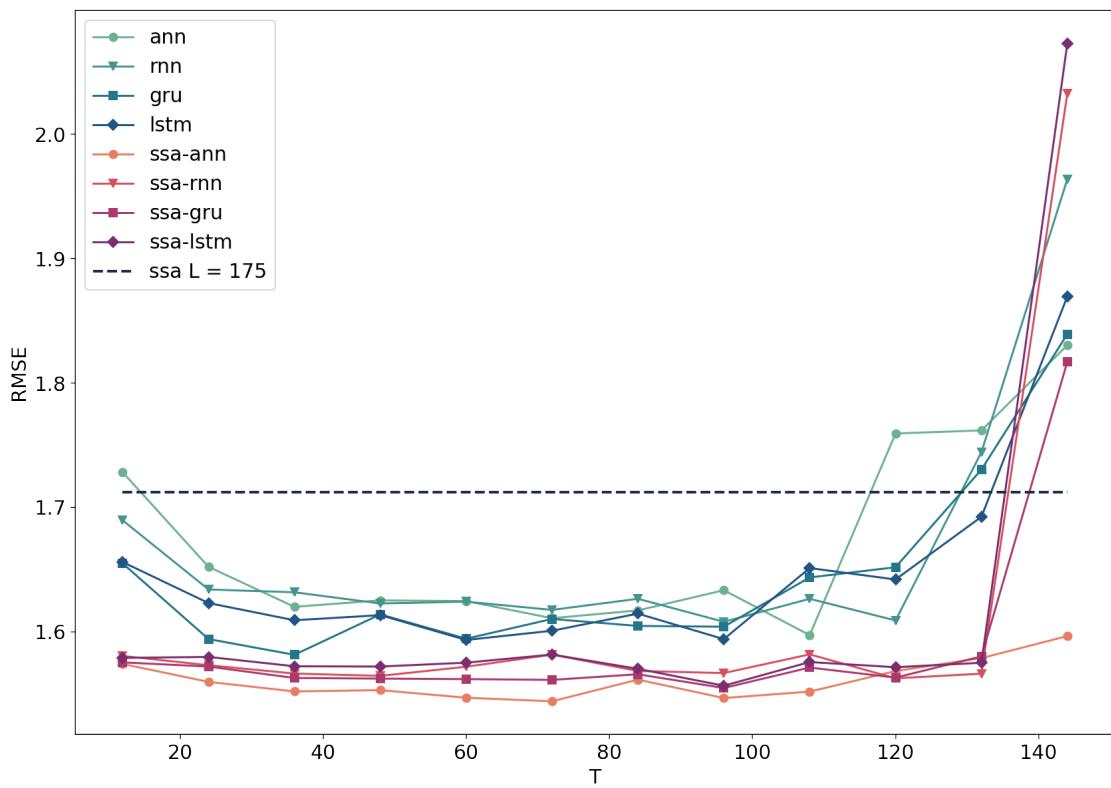
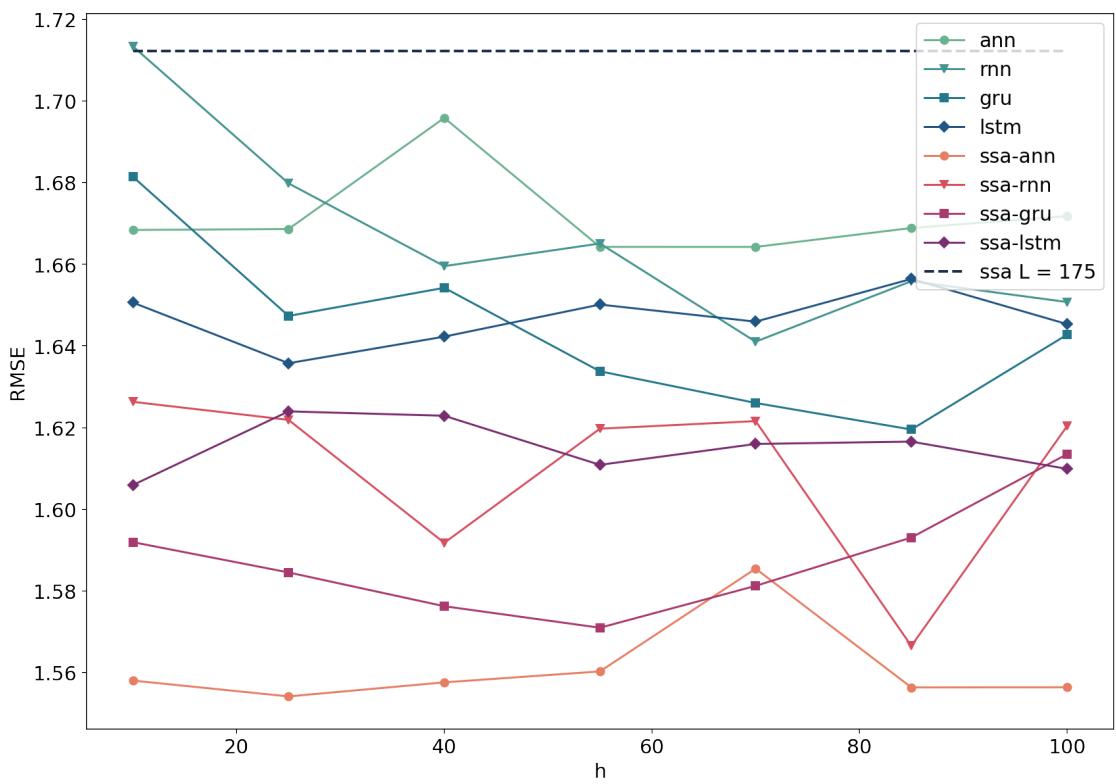
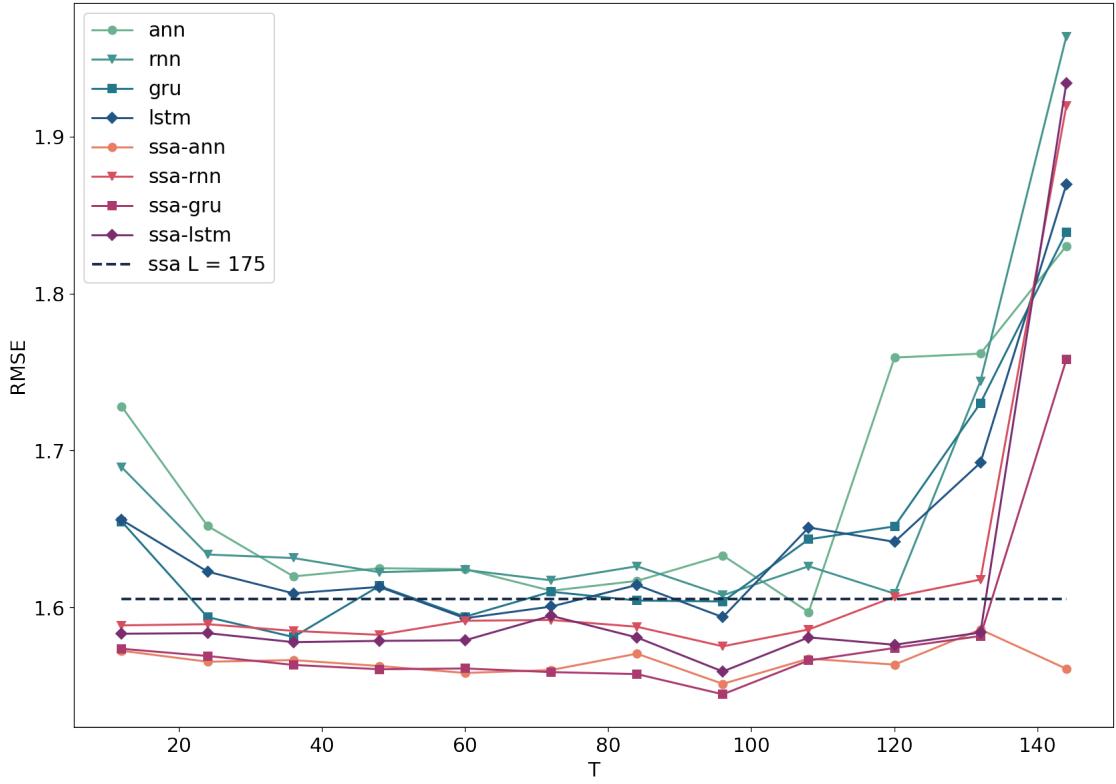
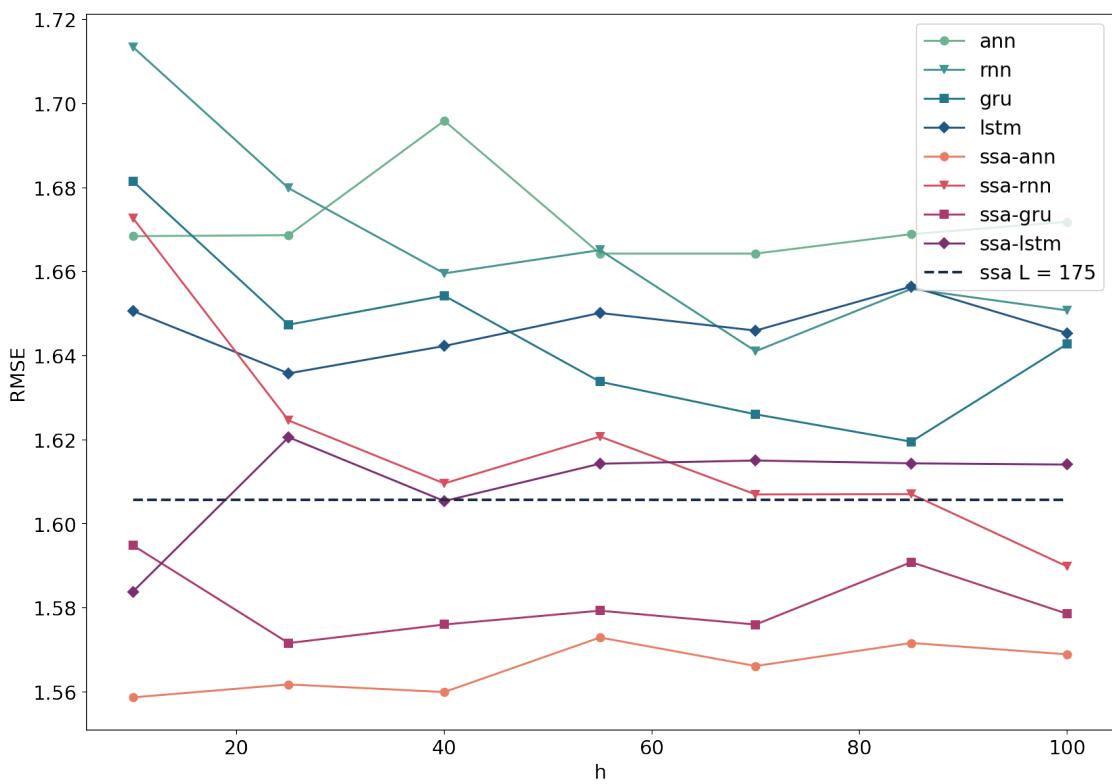
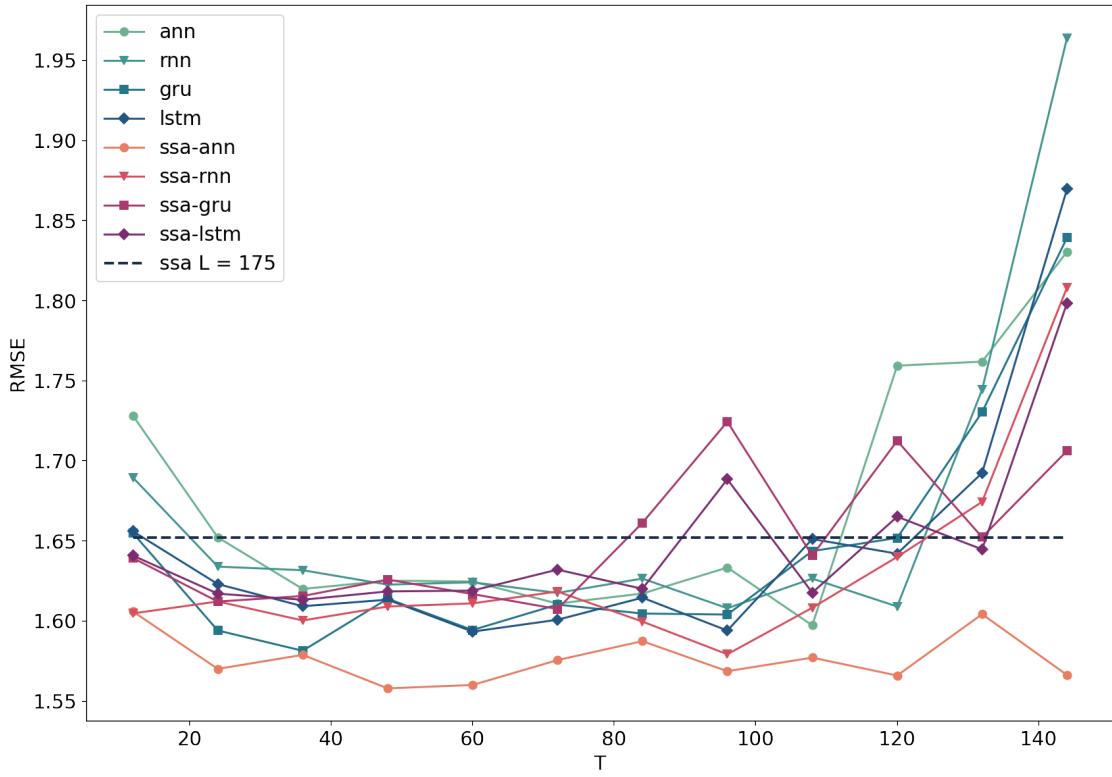
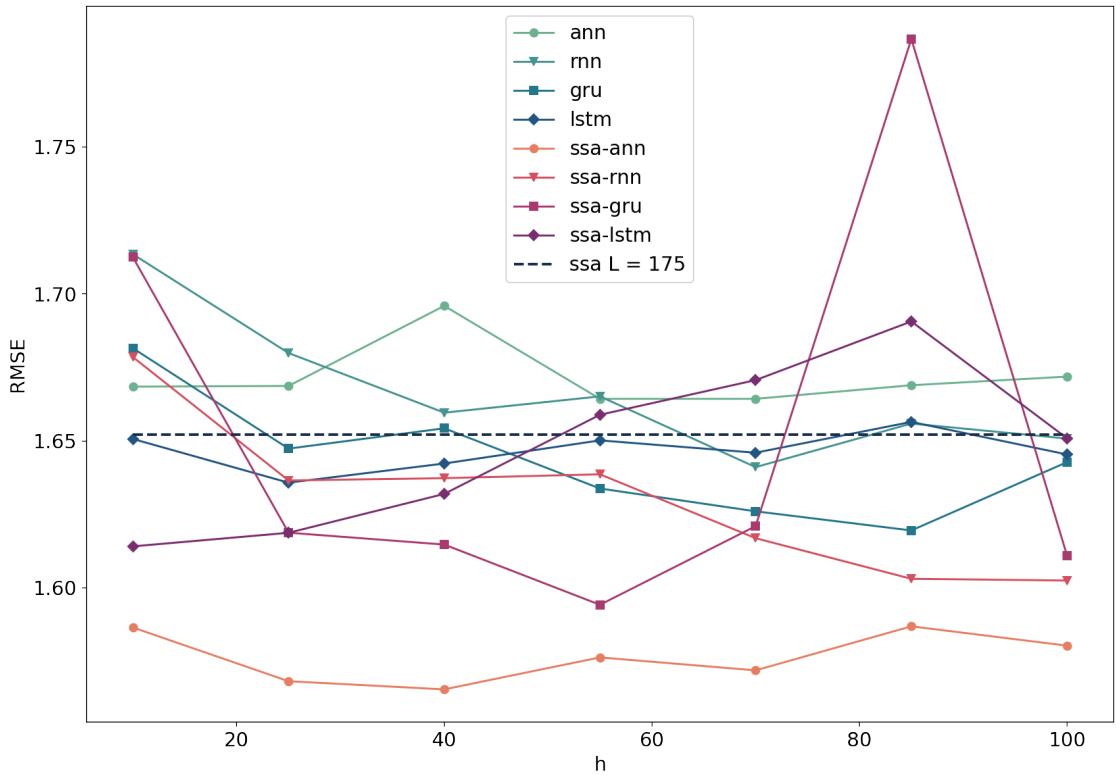


Рис. 4.7. Сравнение методов.  $L = 175$ ,  $r = 2$ .

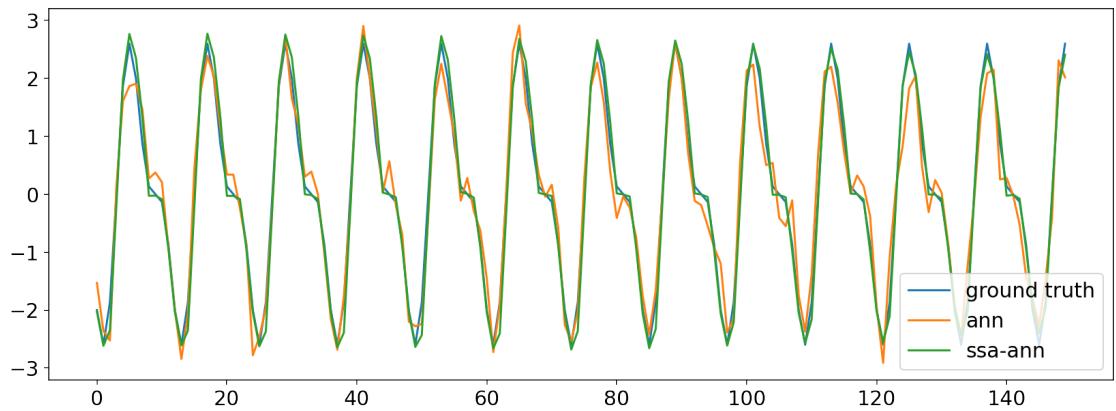
Рис. 4.8. Сравнение методов.  $L = 175$ ,  $r = 2$ .Рис. 4.9. Сравнение методов.  $L = 175$ ,  $r = 4$ .

Рис. 4.10. Сравнение методов.  $L = 175$ ,  $r = 4$ .Рис. 4.11. Сравнение методов.  $L = 175$ ,  $r = 6$ .

Рис. 4.12. Сравнение методов.  $L = 175$ ,  $r = 6$ .

### Отображение результатов

Ниже на графиках представлены результаты прогнозирования методами. На графиках можно подтвердить выводы полученные ранее. Так на графиках для  $r = 6$  видно, что в прогноз просочился шум. Прогнозы для  $r = 2$  и  $r = 4$  выглядят похоже.

Рис. 4.13. Отображение результатов для ANN и SSA-ANN.  $r = 2$

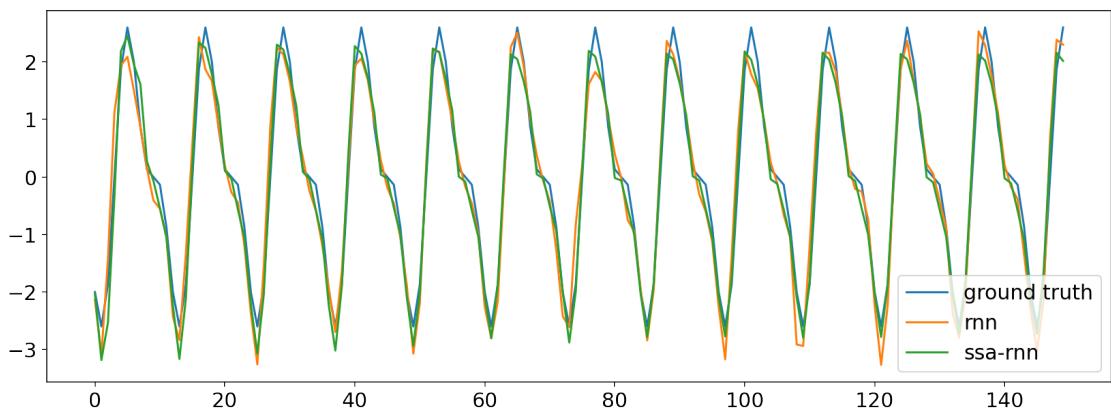


Рис. 4.14. Отображение результатов для RNN и SSA-RNN.  $r = 2$

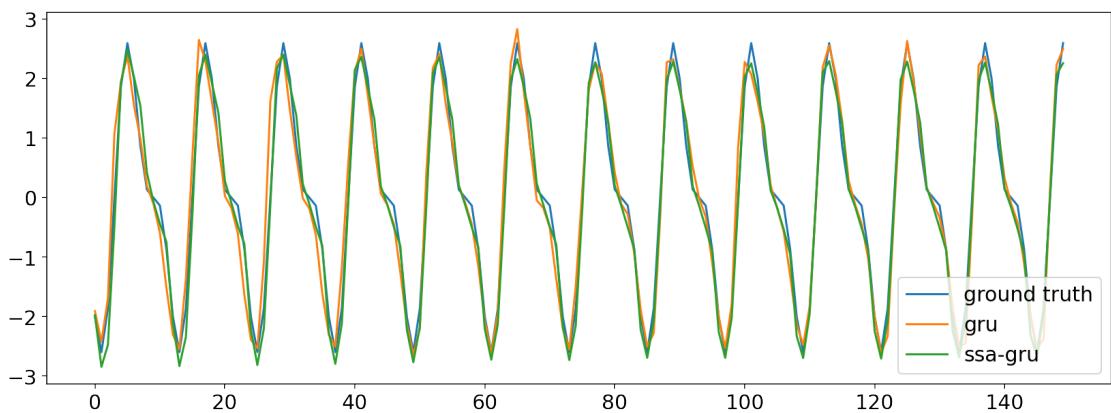


Рис. 4.15. Отображение результатов для GRU и SSA-GRU.  $r = 2$

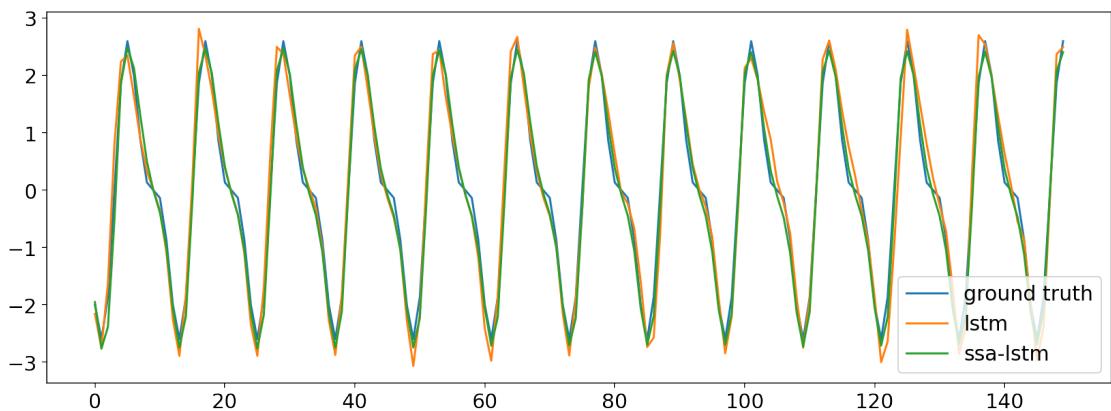


Рис. 4.16. Отображение результатов для LSTM и SSA-LSTM.  $r = 2$

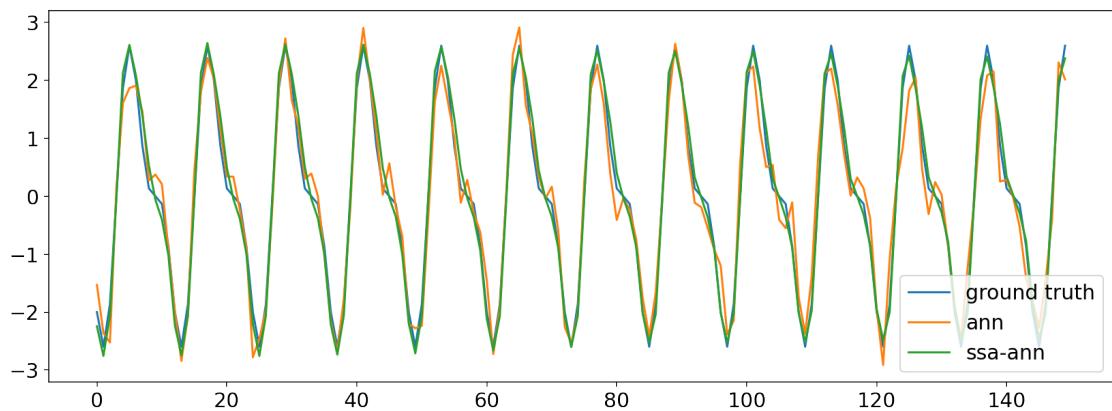


Рис. 4.17. Отображение результатов для ANN и SSA-ANN.  $r = 4$

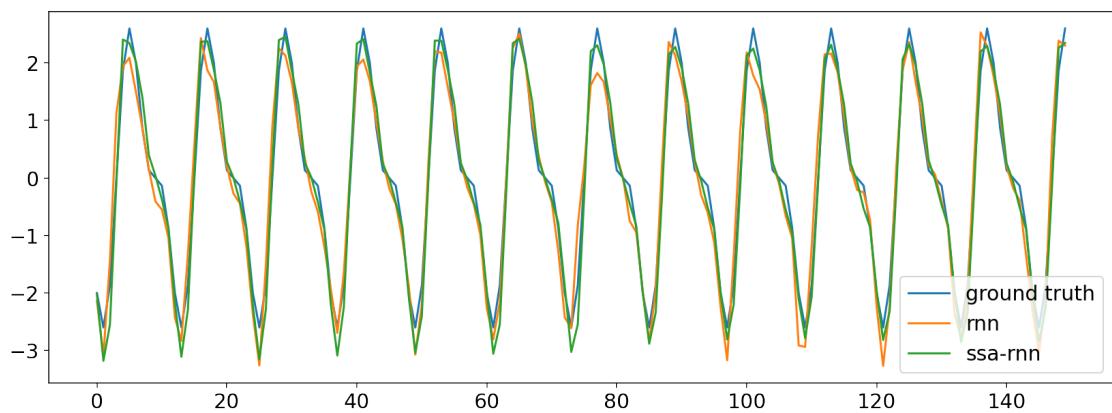


Рис. 4.18. Отображение результатов для RNN и SSA-RNN.  $r = 4$

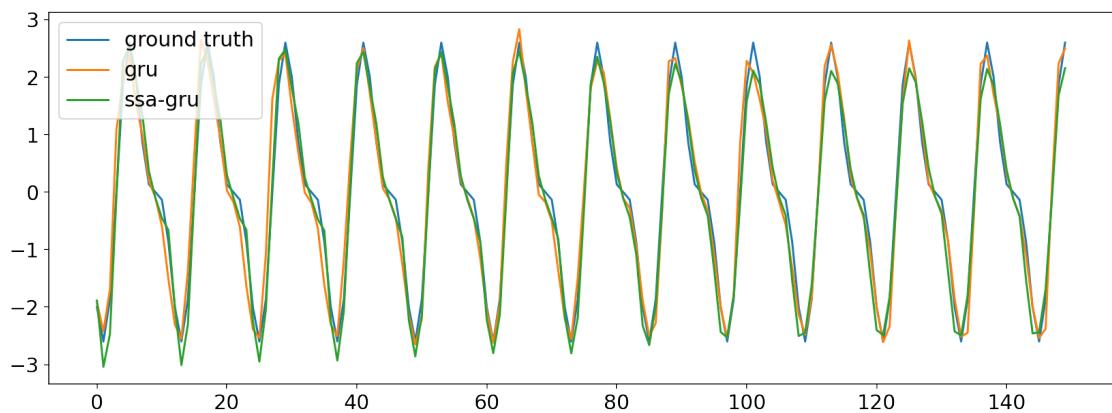


Рис. 4.19. Отображение результатов для GRU и SSA-GRU.  $r = 4$

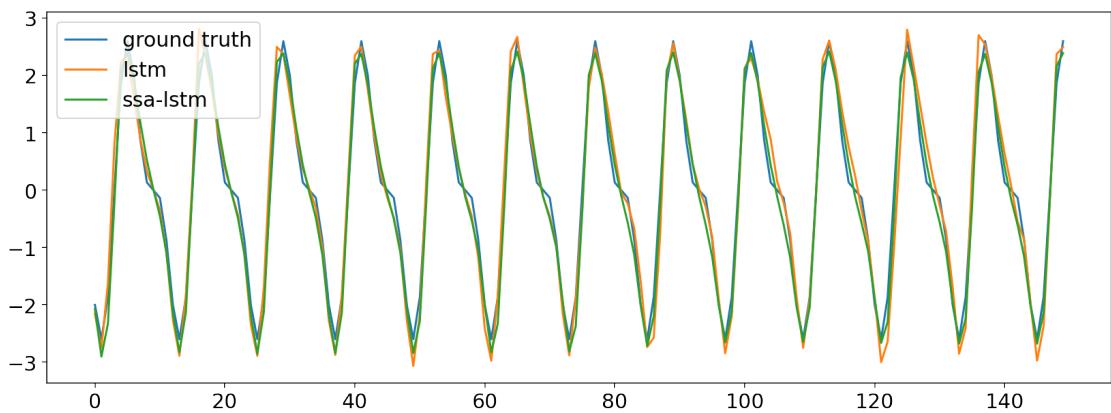


Рис. 4.20. Отображение результатов для LSTM и SSA-LSTM.  $r = 4$

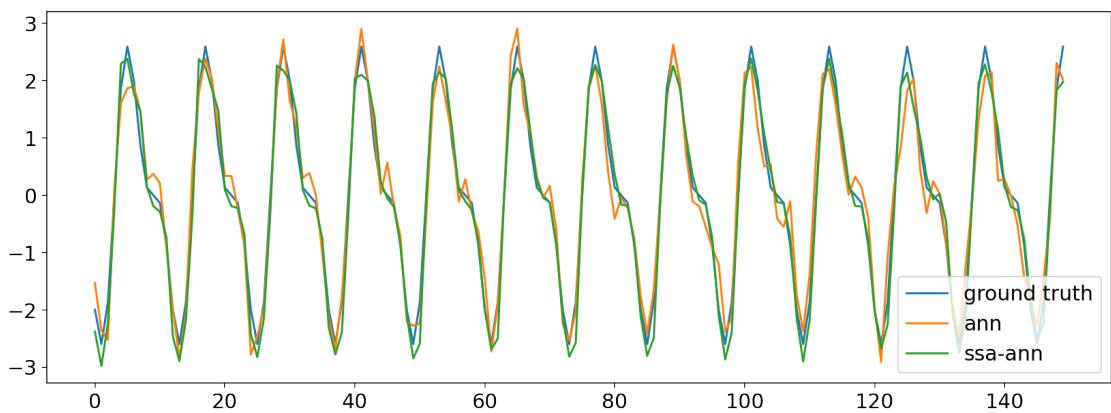


Рис. 4.21. Отображение результатов для ANN и SSA-ANN.  $r = 6$

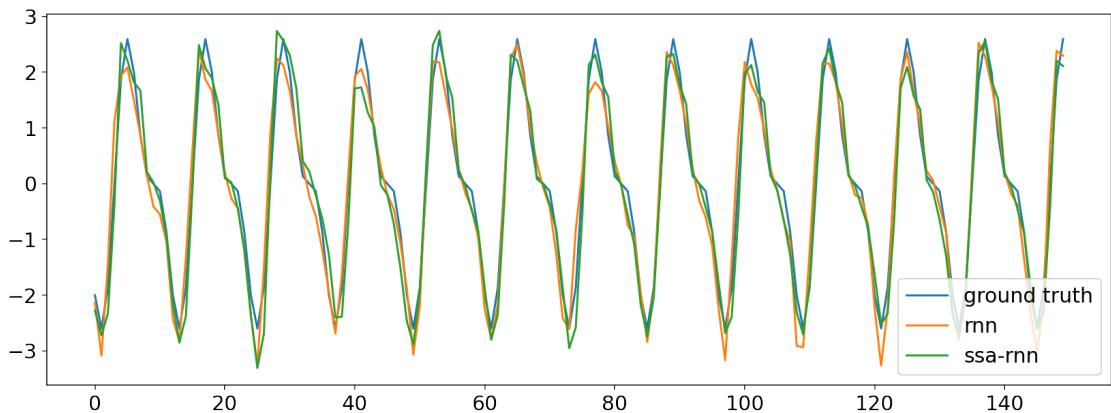


Рис. 4.22. Отображение результатов для RNN и SSA-RNN.  $r = 6$

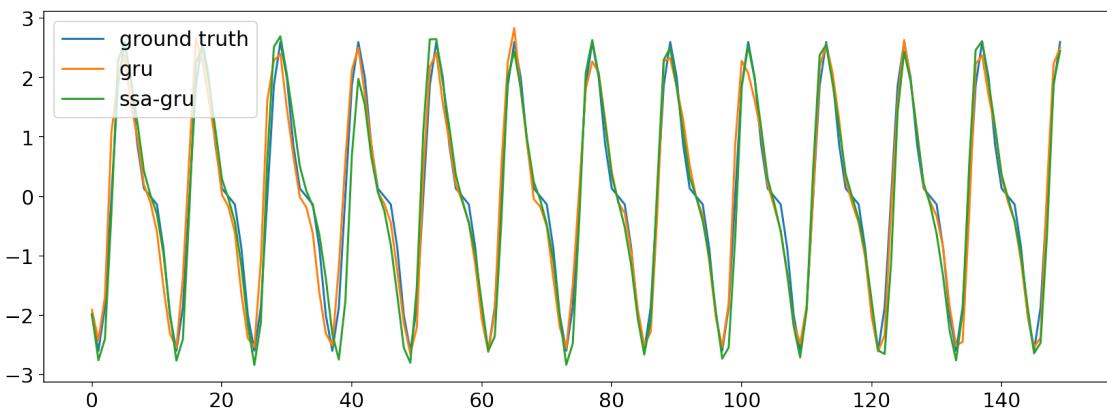


Рис. 4.23. Отображение результатов для GRU и SSA-GRU.  $r = 6$

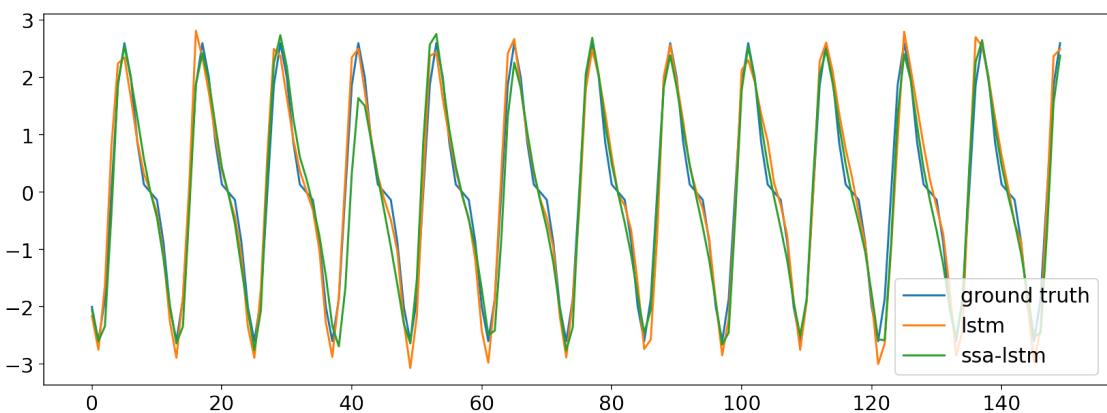


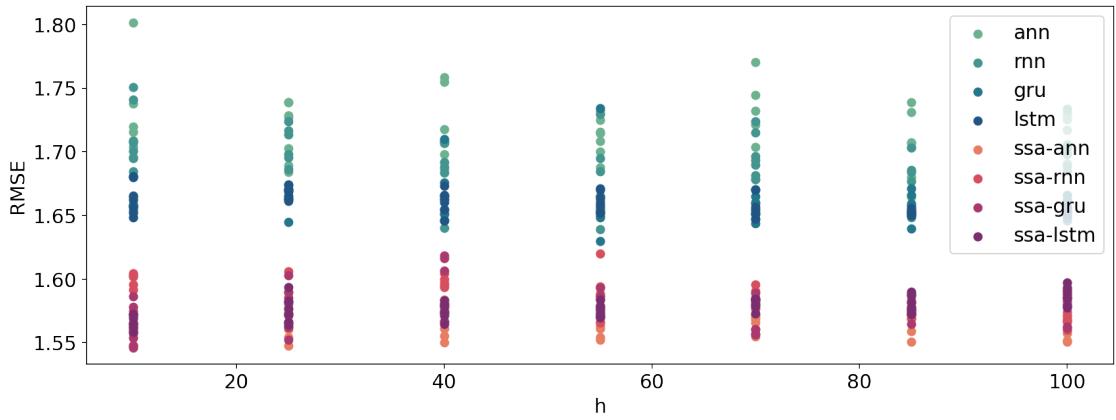
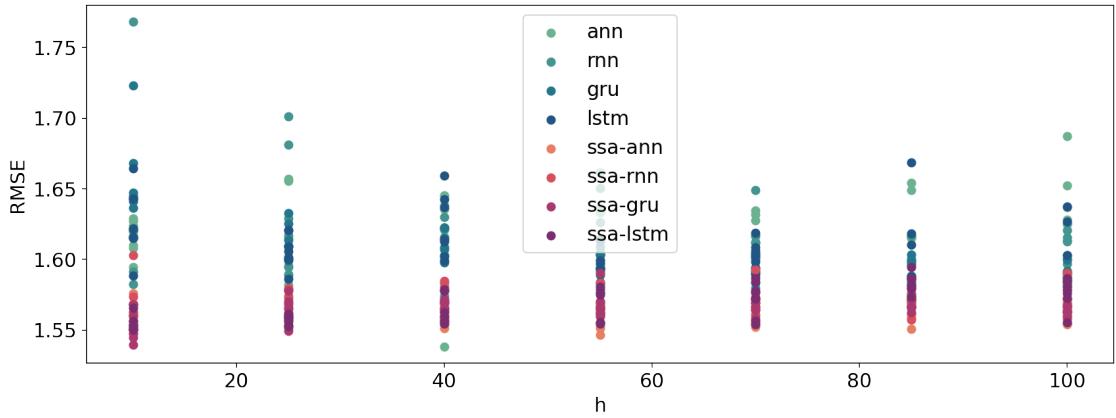
Рис. 4.24. Отображение результатов для LSTM и SSA-LSTM.  $r = 6$

## Проверка устойчивости

Чтобы исключить случайность в полученных результатах, проведем сравнение для разных начальных весов методов. Зафиксируем новую сетку для параметра  $T = \{12, 84\}$ . Сетка для параметр  $h$  останется прежней. Будем получать каждый результат по 7 раз, инициализируя метод с новыми весами. Полученные результаты отображены на рисунках ниже. На них подтверждается, выводы сделанные ранее. Заключаем, что полученные результаты устойчивые.

## Выводы

На ряде  $Z_{650}$  было продемонстрировано сравнение обычных и гибридных методов, и метода SSA. Из полученных результатов можем сделать выводы, что для ряда с несложным выделяемым сигналом выбор аналитически верных параметров является

Рис. 4.25. Проверка устойчивости.  $r = 2$ ,  $L = 12$ .Рис. 4.26. Проверка устойчивости.  $r = 2$ ,  $L = 84$ .

оптимальным.

#### 4.1.2. Влияние выбора параметра $r$ для ряда с небольшим шумом

Поставим задачу сравнить обычные и гибридные методы в случае разным выбранным параметром  $r$  в гибридных методах на временном ряде с небольшим шумом. Эксперимент аналогичен другому, описанному в главе 4.1.1. Будем проводить сравнение на ряде  $X_{650}$ .

#### Прогноз по SSA

Сравним точность прогнозирования методом SSA при разных параметрах  $r$ . Зададим следующую сетку параметров  $L = \{12, 24, \dots, 168\}$ ,  $r = \{2, 4, 6\}$ . Посмотрим на результаты на рисунке 4.31. На графике видно, что наилучшие результаты достигаются при  $r = 4$ , худшие результаты достигаются при  $r = 2$ .

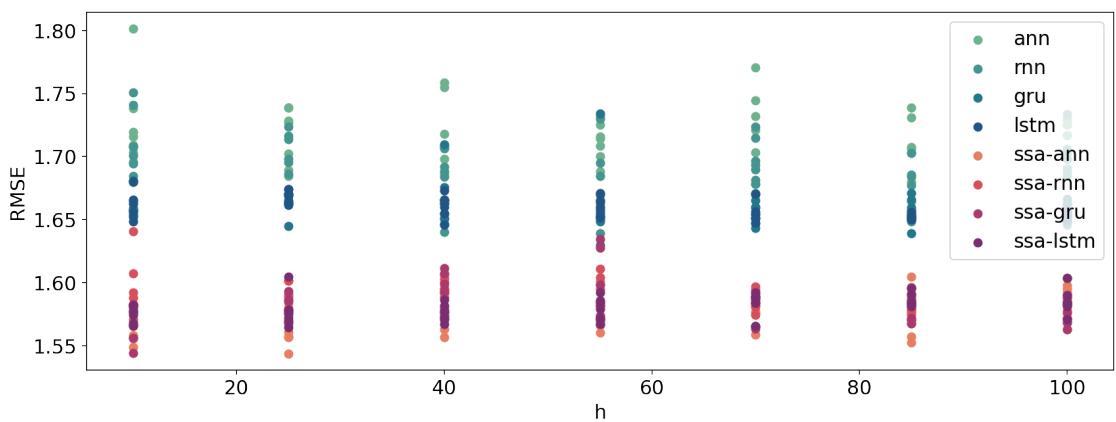
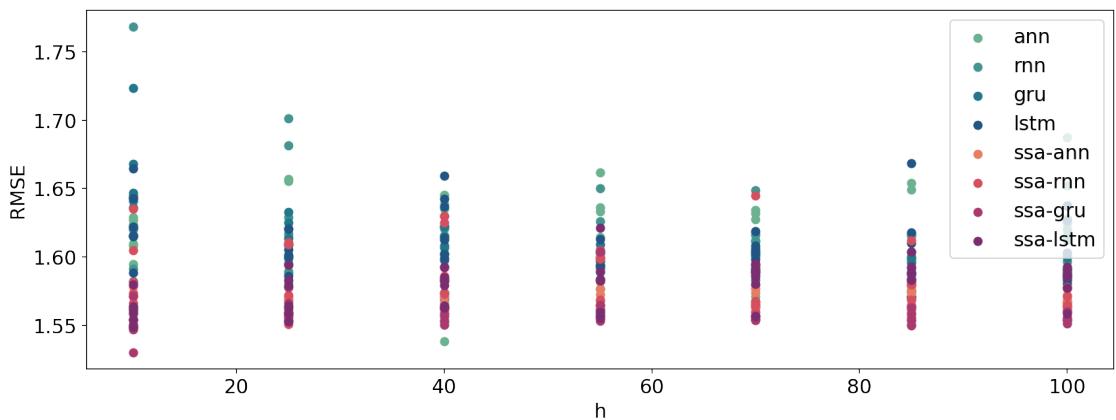
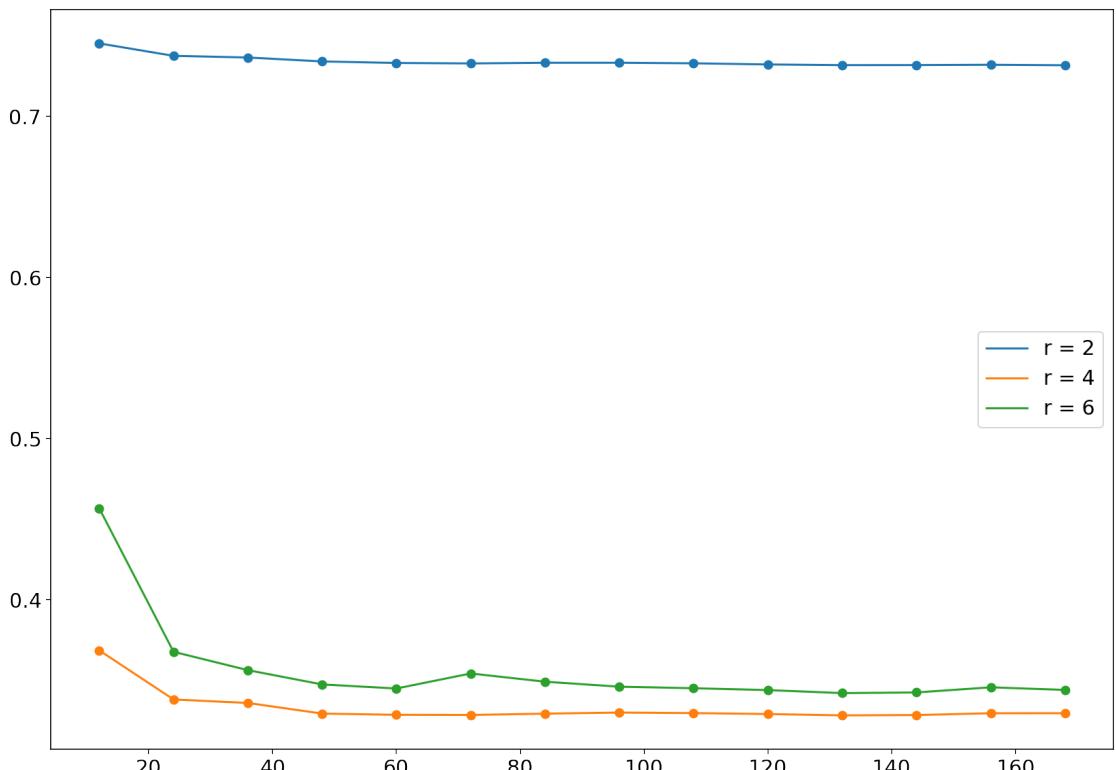
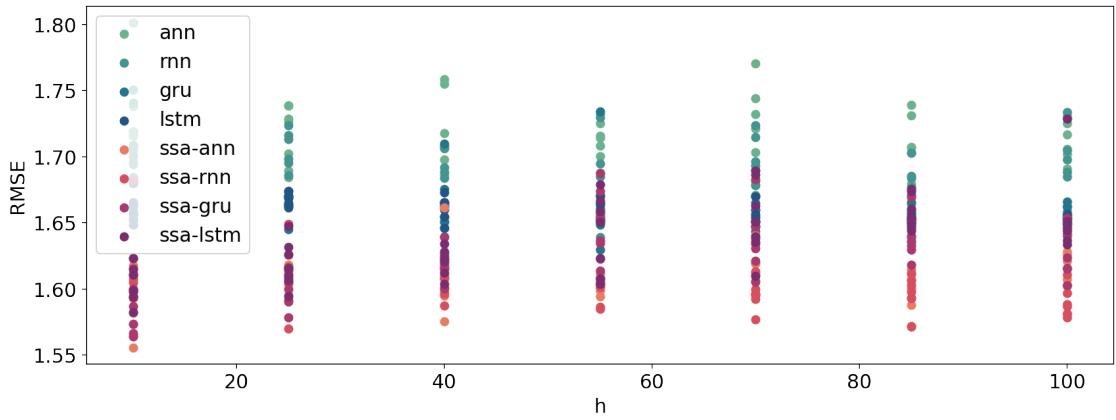
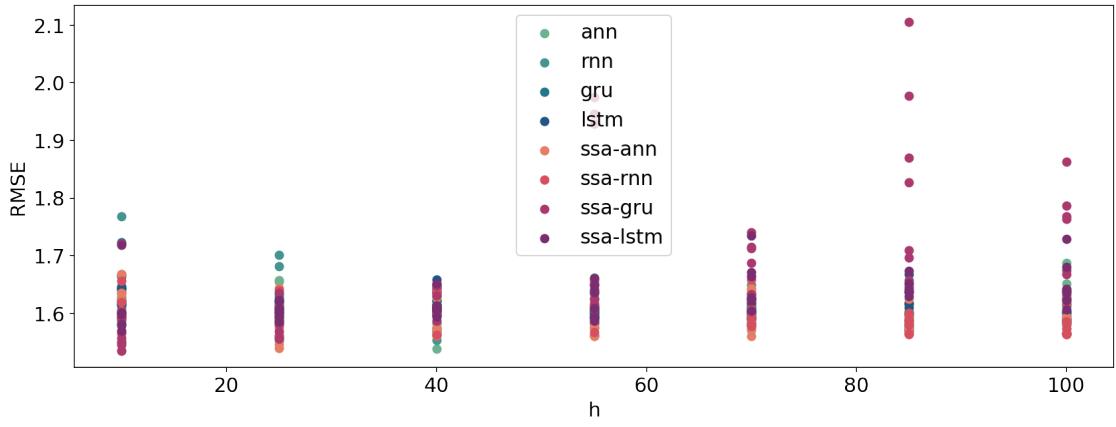
Рис. 4.27. Проверка устойчивости.  $r = 4$ ,  $L = 12$ .Рис. 4.28. Проверка устойчивости.  $r = 4$ ,  $L = 84$ .

Рис. 4.31. Прогнозирование с помощью метода SSA.

Рис. 4.29. Проверка устойчивости.  $r = 6$ ,  $L = 12$ .Рис. 4.30. Проверка устойчивости.  $r = 6$ ,  $L = 84$ .

Выделим лучшее  $L$  для каждого  $r$ . Исходя из графика для  $r = \{4, 6\}$  это будет  $L = 132$ , для  $r = 2$  будет  $L = 168$ . Далее будем рассматривать лучшие комбинации гиперпараметров SSA вместе с  $L = 175$  и  $r = \{2, 4, 6\}$ .

## Восстановление SSA

Посмотрим, как метод SSA восстанавливает тренировочную выборку для выбранных пар на графиках ниже. На графике 4.32 видно, что метод не восстанавливает ряд полностью. Результат очень похожи на те, что в главе 4.1.1. На графиках 4.33, 4.34, видно, что результаты идентичны. В главе 4.1.1 для  $r = 6$  в сигнал попадал шум, так как в этом примере шум небольшой, это сильно не портит оценку сигнала. На графиках видно, что разницы между восстановленными рядами нету, далее будем рассматривать только параметры  $r = \{2, 4, 6\}$  и  $L = 175$ .

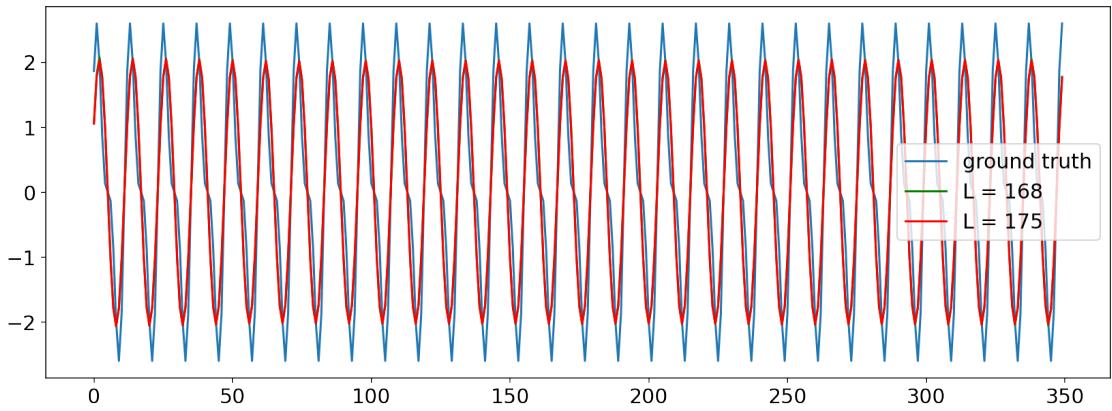


Рис. 4.32. Восстановление тренировочной выборки с помощью метода SSA.  $r = 2$

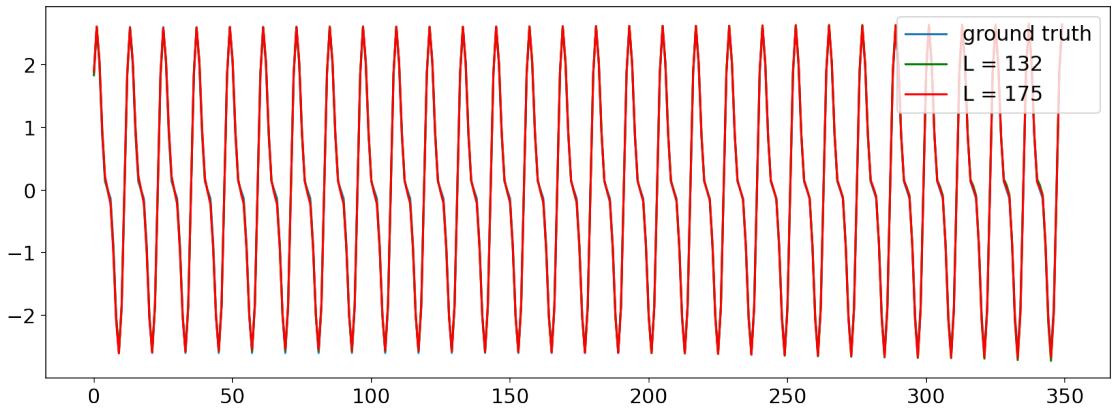


Рис. 4.33. Восстановление тренировочной выборки с помощью метода SSA.  $r = 4$

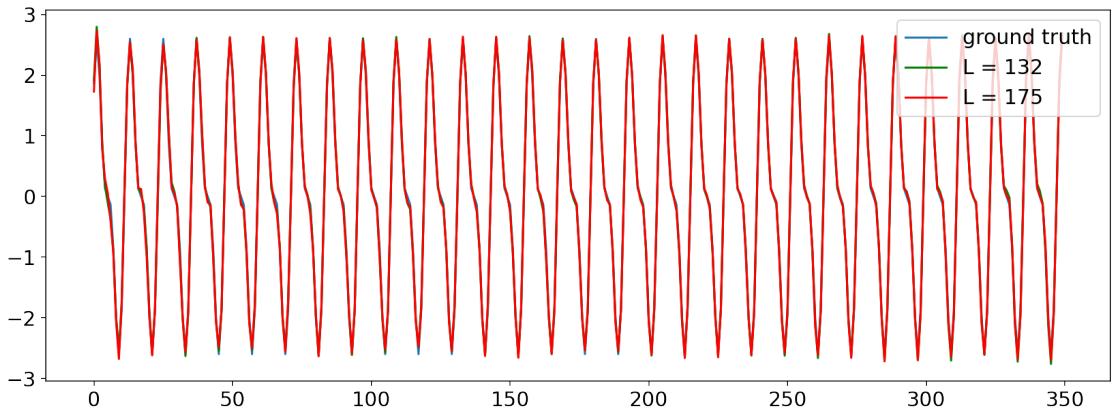


Рис. 4.34. Восстановление тренировочной выборки с помощью метода SSA.  $r = 6$

### Сравнение обычных и гибридных методов

Пусть задана следующая сетка параметров:  $T = \{12, 24, \dots, 144\}$ ,  $h = \{10, 25, \dots, 100\}$ .

Для метода SSA в гибридных моделях возьмем пары параметров, выбранные заранее.

На графиках 4.35, 4.36, 4.37, 4.38, 4.39, 4.40 представлены результаты сравнения по сетке параметров, заданной выше. Можно заметить, что гибридные методы явно лучше, чем обычные. Для всех  $r$  результаты очень похожие. Можно заметить, что для  $r = 4$  ошибка меньше, чем для других  $r$ , а также разрыв в точности между обычными и гибридными методами виден больше. Разрывы на графиках небольшие, так как в ряде очень слабый шум. Для точности прогнозов методом SSA все осталось также, как и в предыдущей части эксперимента.

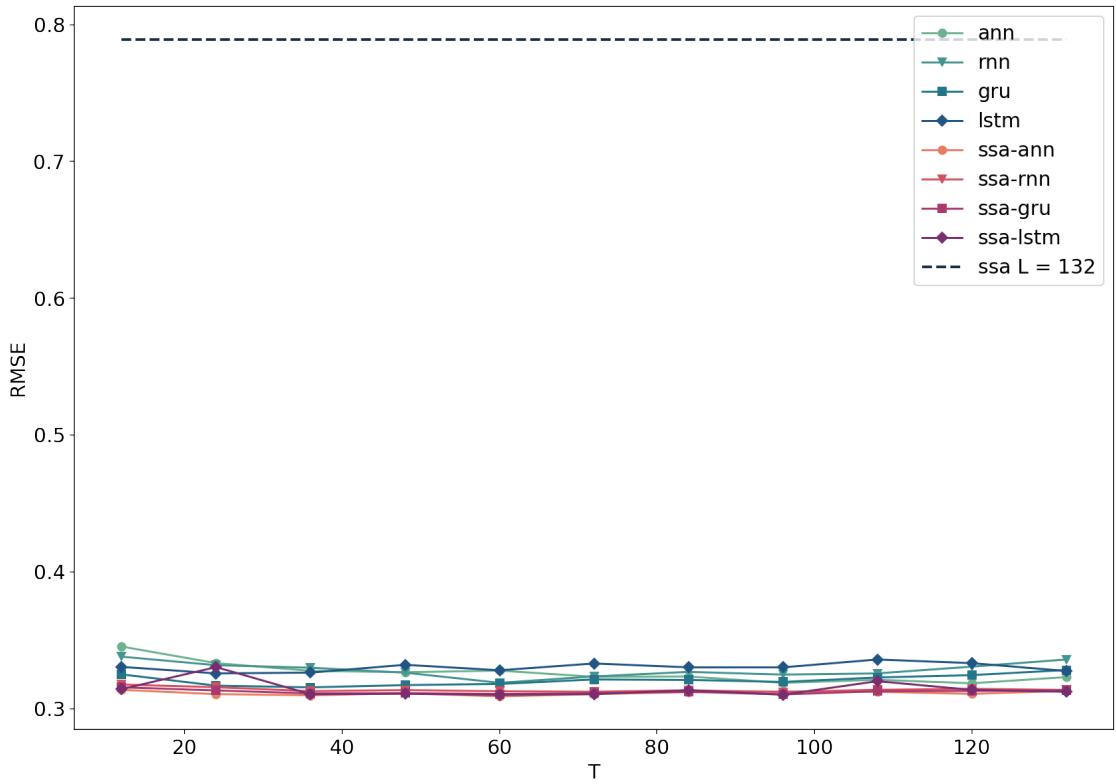
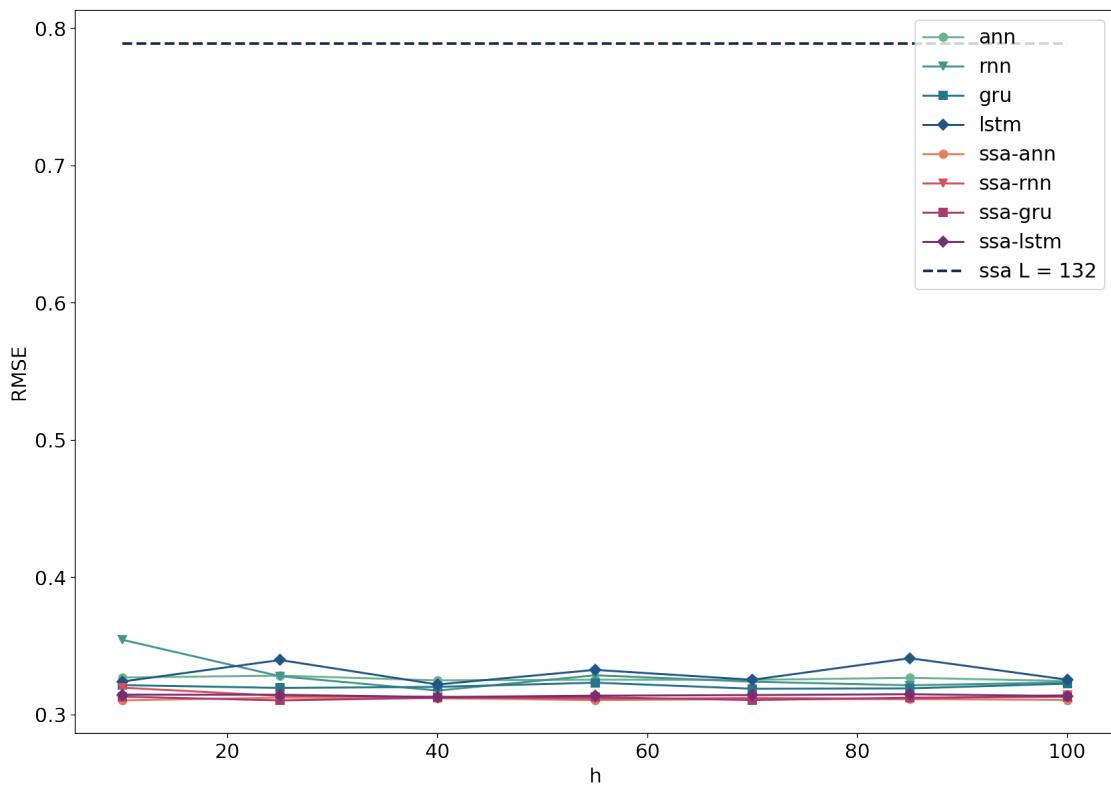
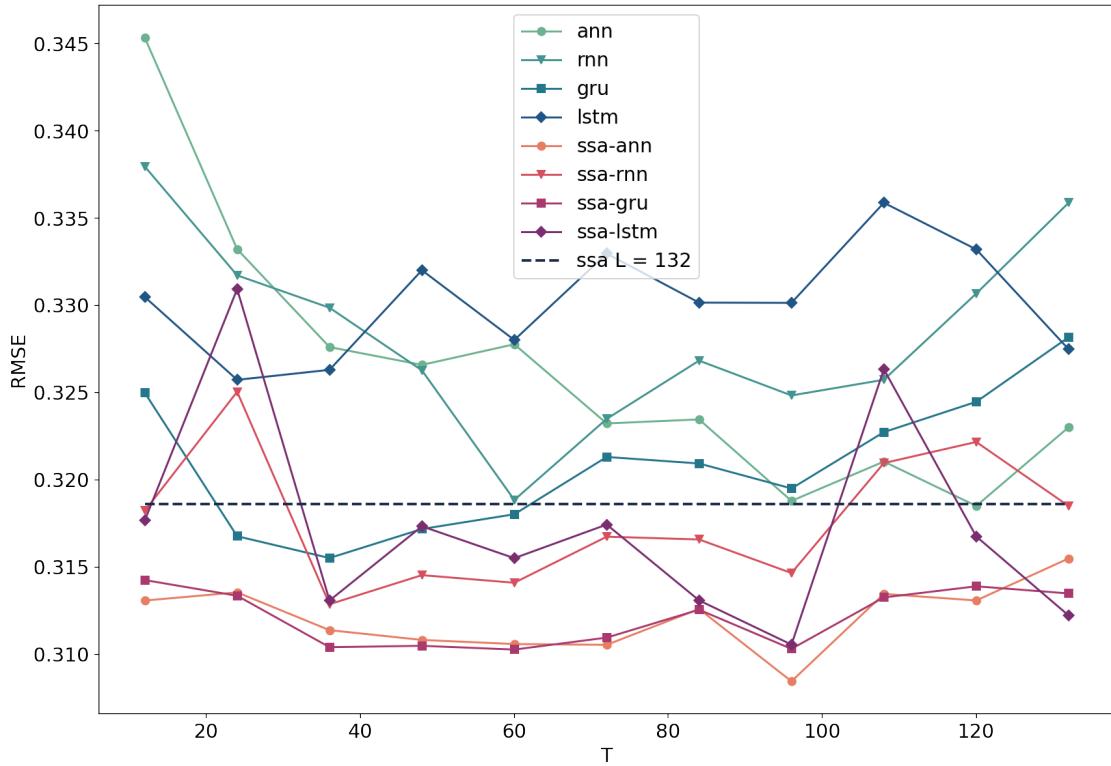
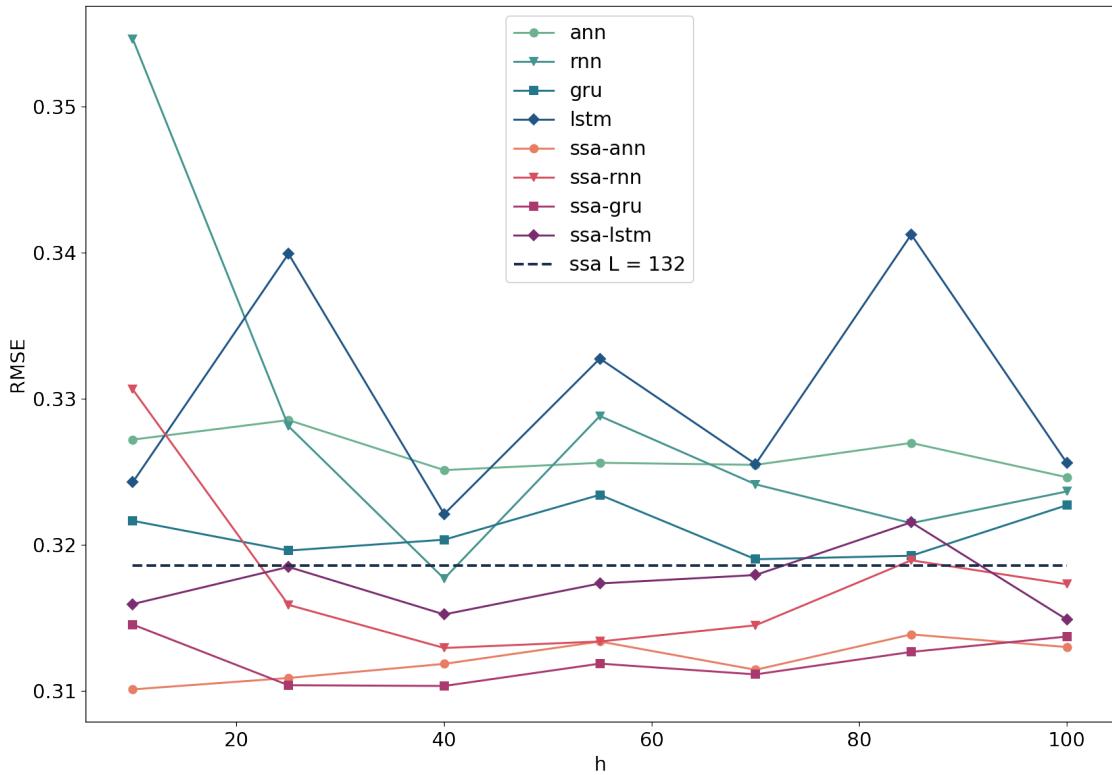
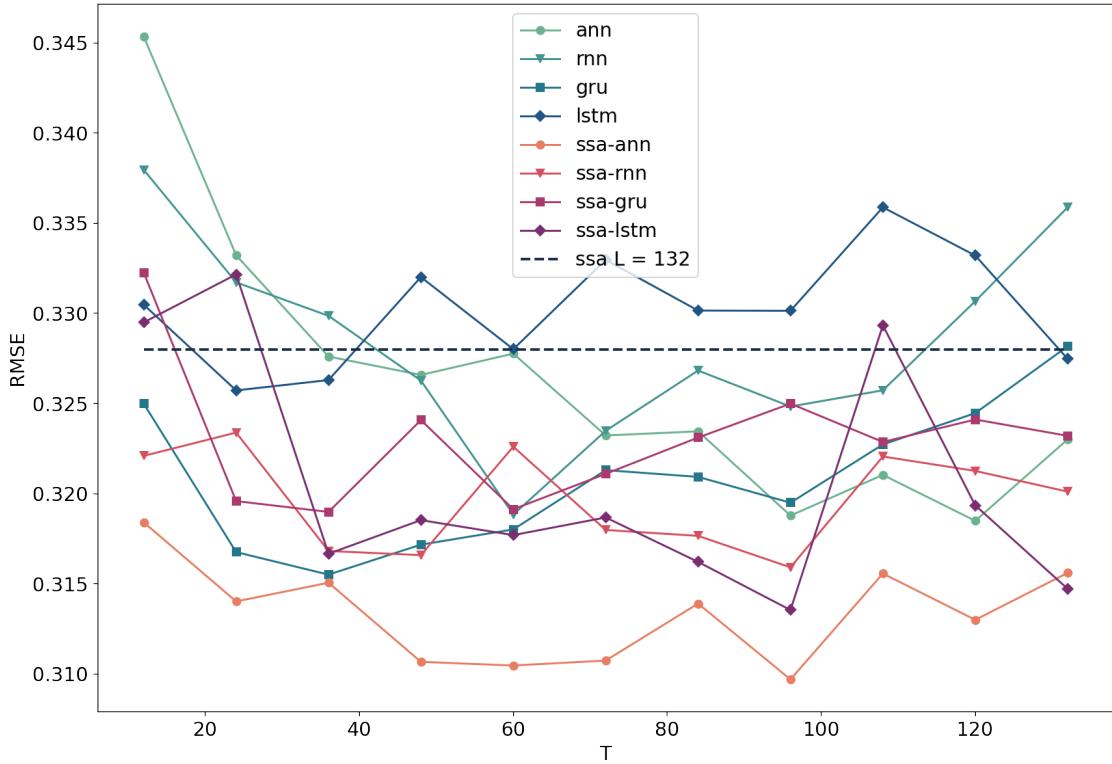
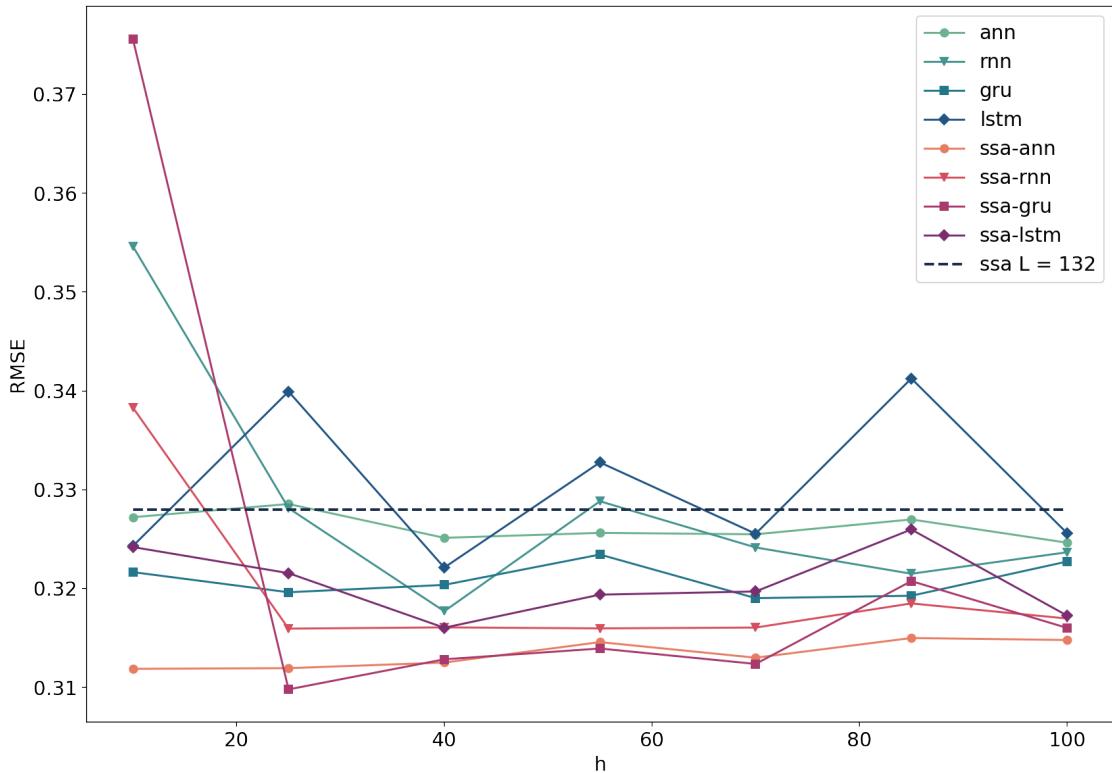


Рис. 4.35. Сравнение методов.  $L = 175$ ,  $r = 2$ .

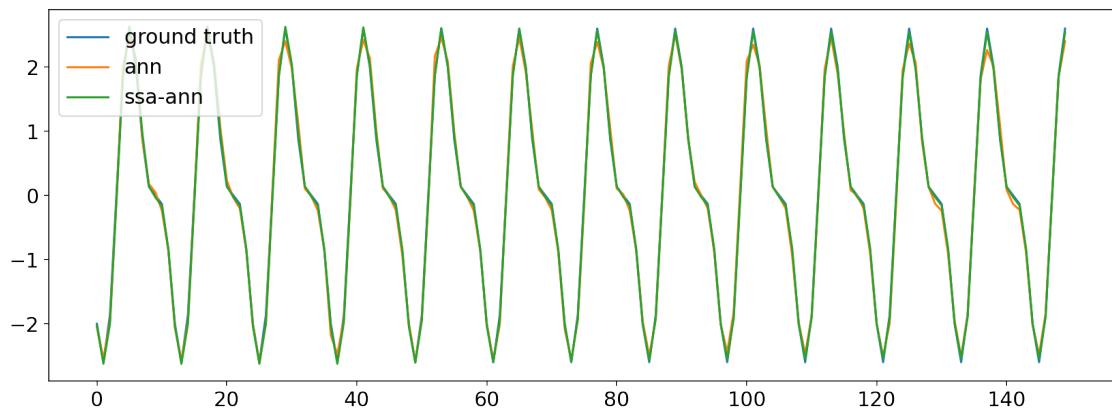
Рис. 4.36. Сравнение методов.  $L = 175$ ,  $r = 2$ .Рис. 4.37. Сравнение методов.  $L = 175$ ,  $r = 4$ .

Рис. 4.38. Сравнение методов.  $L = 175$ ,  $r = 4$ .Рис. 4.39. Сравнение методов.  $L = 175$ ,  $r = 6$ .

Рис. 4.40. Сравнение методов.  $L = 175$ ,  $r = 6$ .

### Отображение результатов

На графиках ниже представлены результаты прогнозирования методами. Из-за маленького шума все прогнозы похожи друг на друга.

Рис. 4.41. Отображение результатов для ANN и SSA-ANN.  $r = 2$

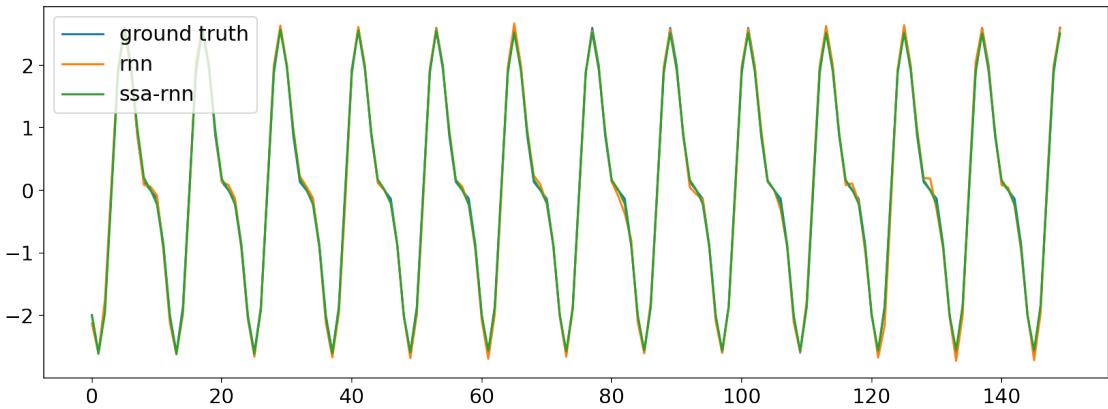


Рис. 4.42. Отображение результатов для RNN и SSA-RNN.  $r = 2$

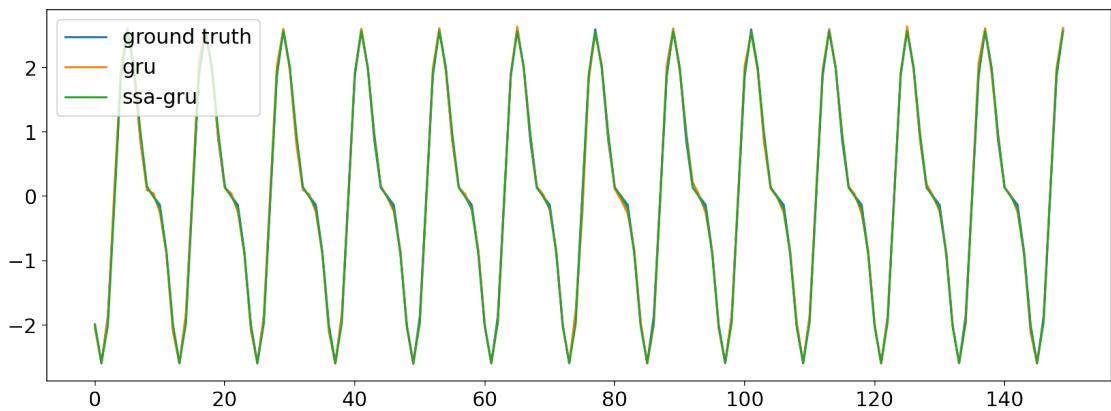


Рис. 4.43. Отображение результатов для GRU и SSA-GRU.  $r = 2$

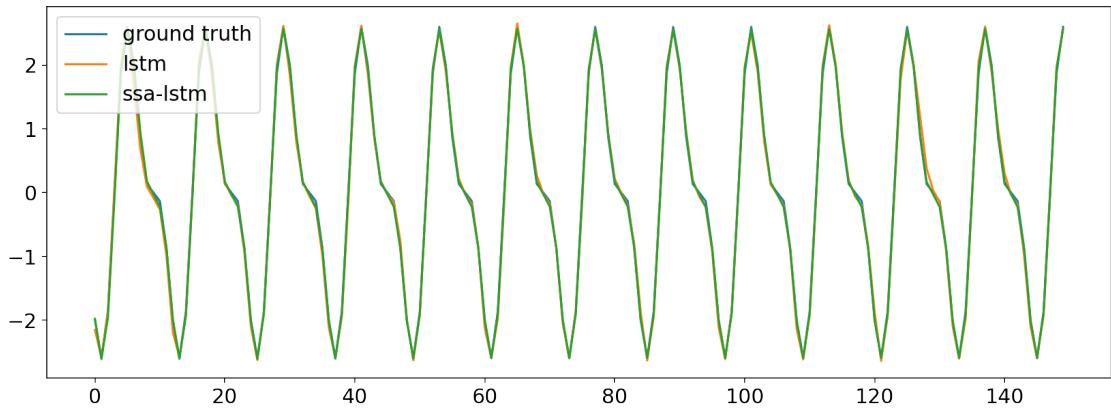


Рис. 4.44. Отображение результатов для LSTM и SSA-LSTM.  $r = 2$

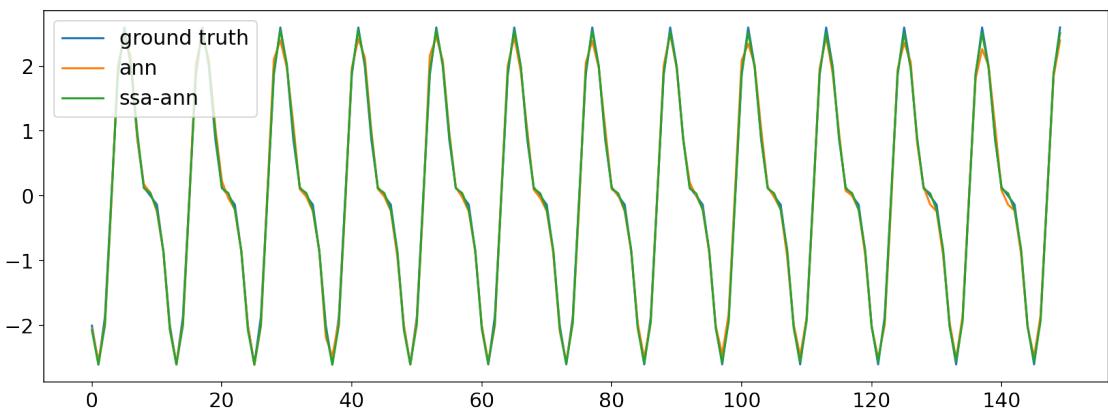


Рис. 4.45. Отображение результатов для ANN и SSA-ANN.  $r = 4$

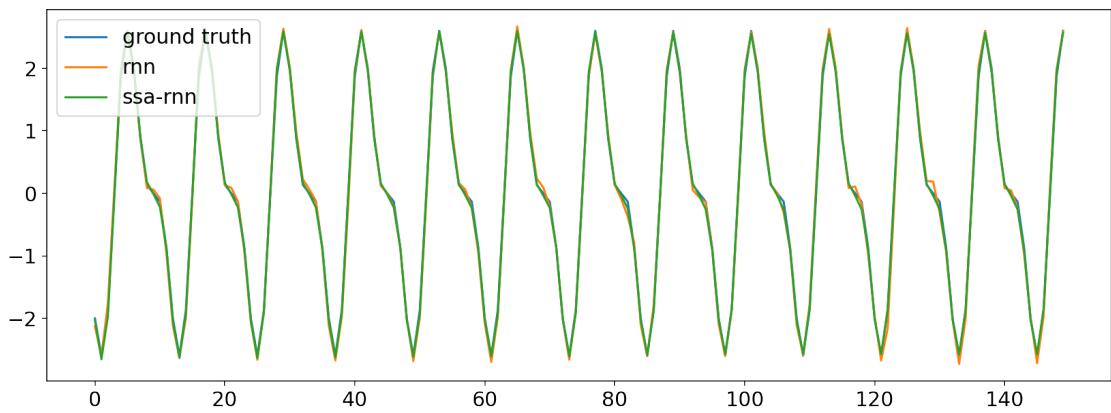


Рис. 4.46. Отображение результатов для RNN и SSA-RNN.  $r = 4$

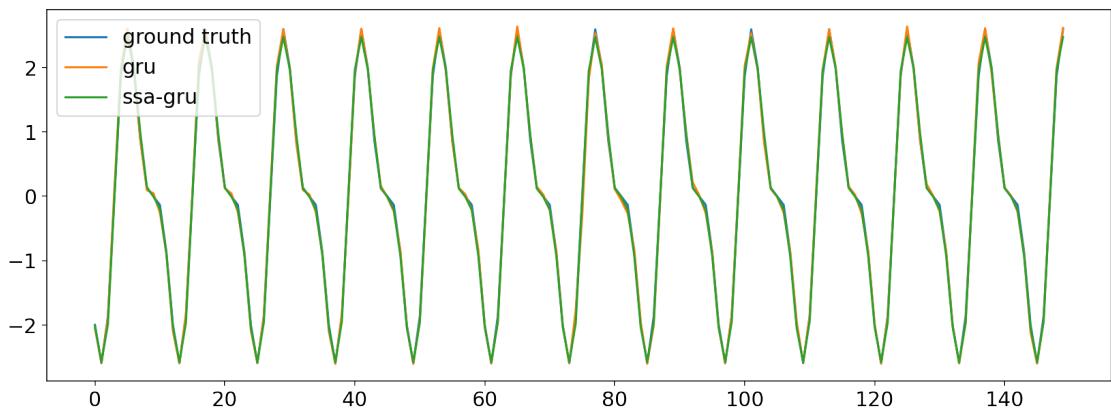


Рис. 4.47. Отображение результатов для GRU и SSA-GRU.  $r = 4$

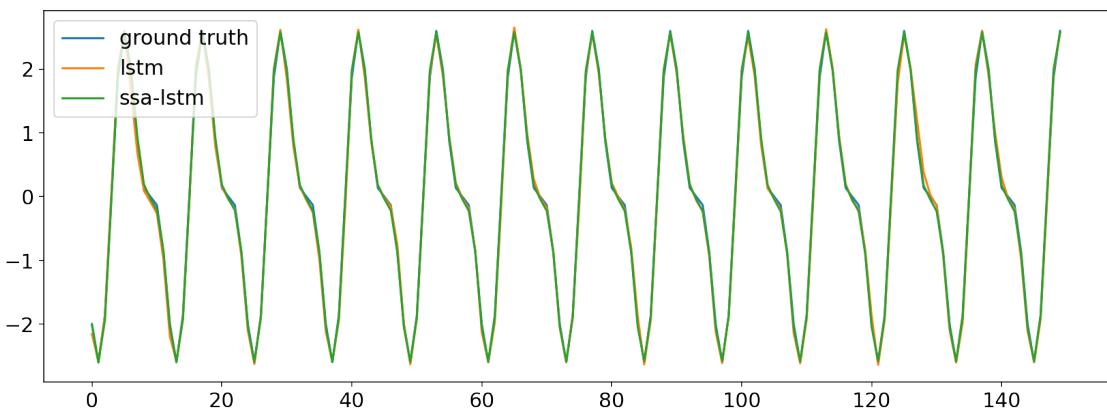


Рис. 4.48. Отображение результатов для LSTM и SSA-LSTM.  $r = 4$

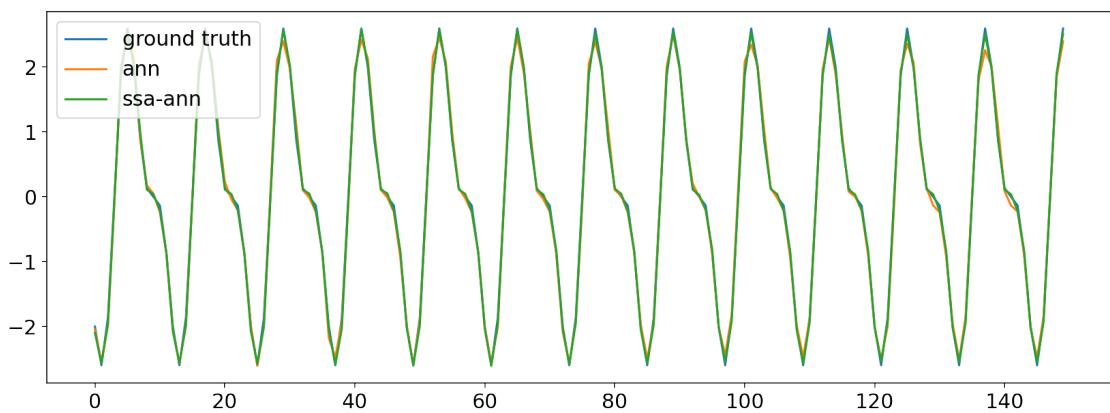


Рис. 4.49. Отображение результатов для ANN и SSA-ANN.  $r = 6$

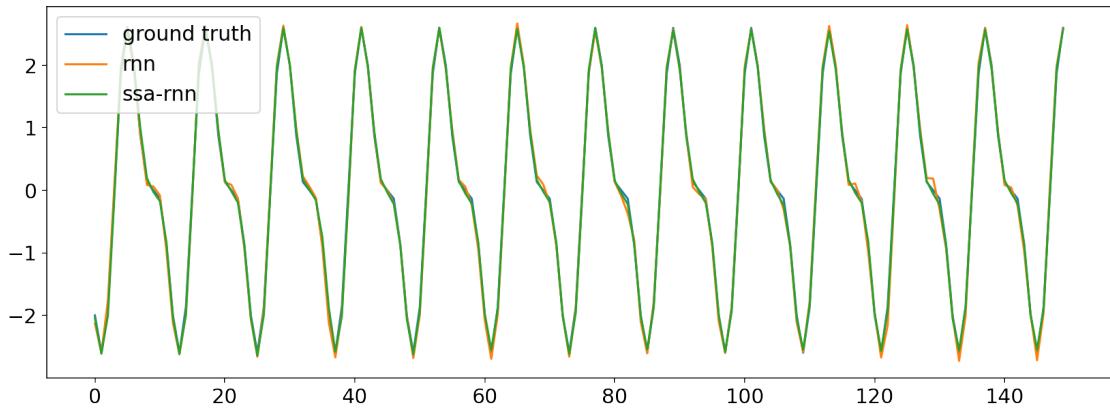


Рис. 4.50. Отображение результатов для RNN и SSA-RNN.  $r = 6$

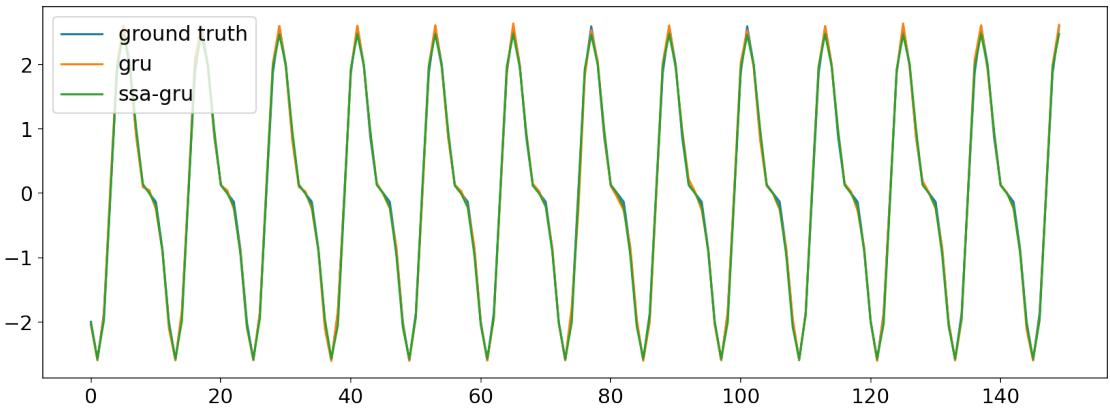


Рис. 4.51. Отображение результатов для GRU и SSA-GRU.  $r = 6$

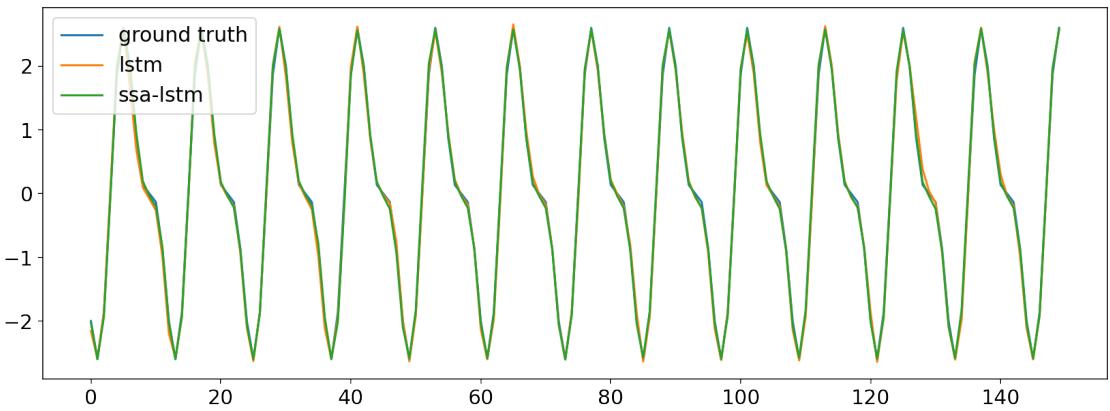


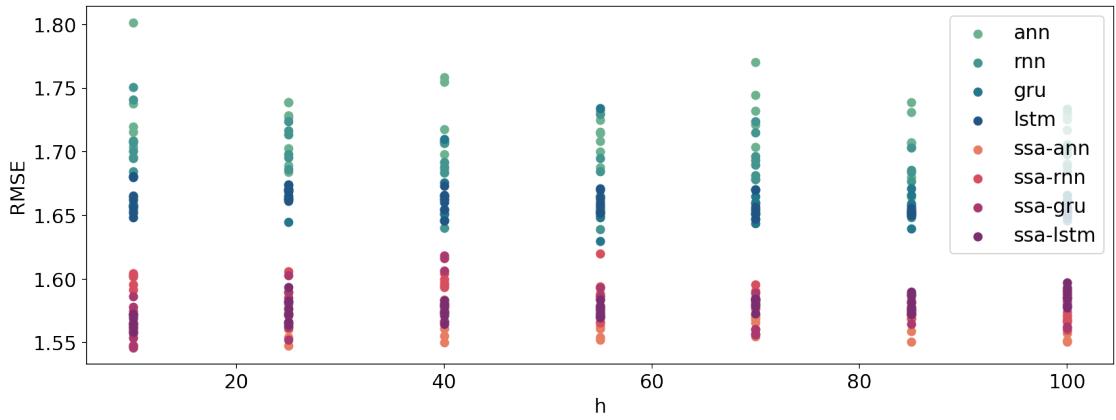
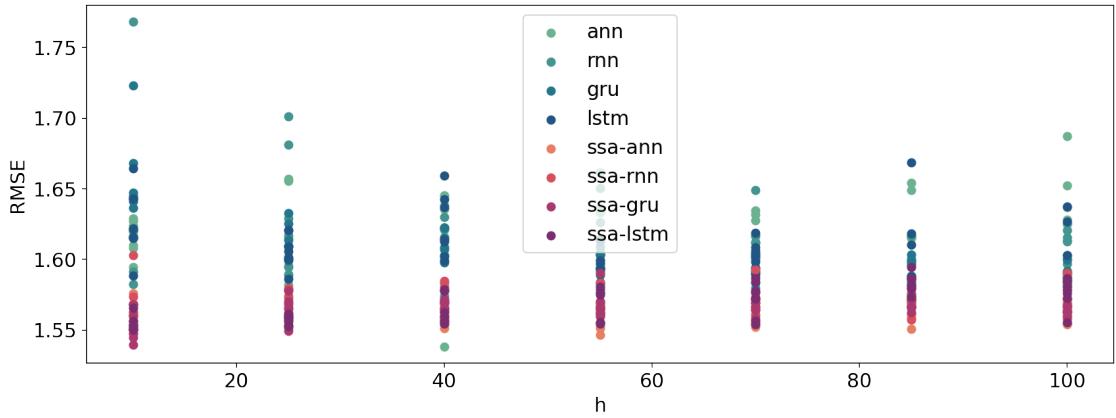
Рис. 4.52. Отображение результатов для LSTM и SSA-LSTM.  $r = 6$

## Проверка устойчивости

Чтобы исключить случайность в полученных результатах, проведем сравнение для разных начальных весов методов. Зафиксируем новую сетку для параметра  $T = \{12, 84\}$ . Сетка для параметр  $h$  останется прежней. Будем получать каждый результат по 7 раз, инициализируя метод с новыми весами. Полученные результаты отображены на рисунках ниже. На них подтверждается, выводы сделанные ранее. Заключаем, что полученные результаты устойчивые.

## Выводы

На ряде  $X_{650}$  было продемонстрировано сравнение обычных и гибридных методов, и метода SSA. Из полученных результатов можем сделать выводы, что для ряда с несложным выделяемым сигналом выбор аналитически верных параметров является оп-

Рис. 4.53. Проверка устойчивости.  $r = 2$ ,  $L = 12$ .Рис. 4.54. Проверка устойчивости.  $r = 2$ ,  $L = 84$ .

тимальным и для маленького шума. Также в эксперименте было продемонстрирована меньшая разница в точности между обычными и гибридными методами. Было показано, что выбор параметра  $r$  меньше, чем нужно, ведет к большему ухудшению точности, в отличие от выбора в большую сторону.

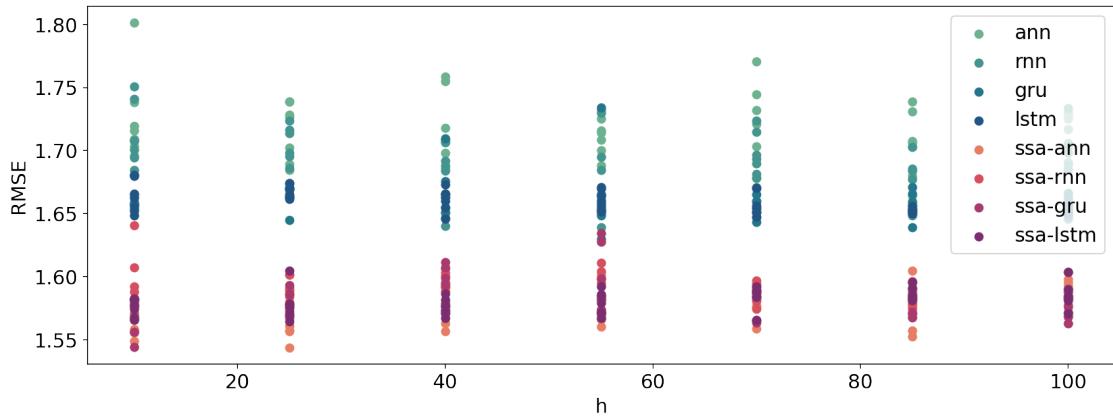


Рис. 4.55. Проверка устойчивости.  $r = 4$ ,  $L = 12$ .

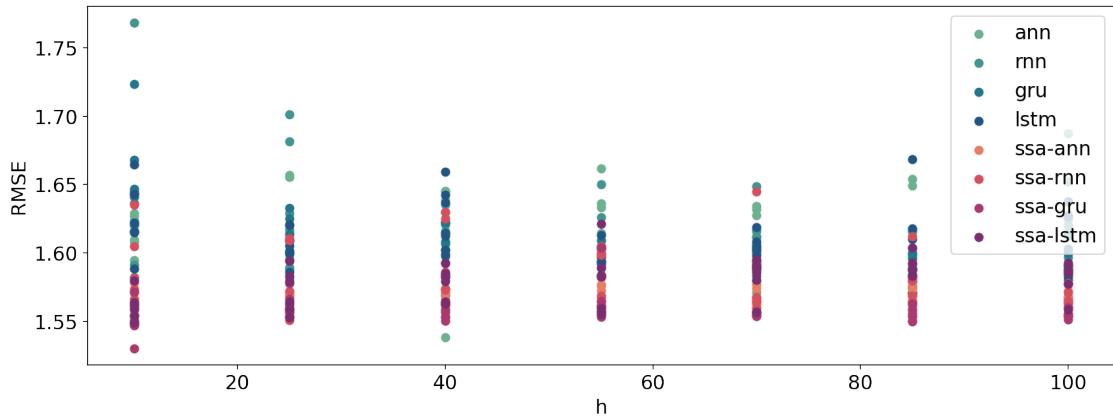


Рис. 4.56. Проверка устойчивости.  $r = 4$ ,  $L = 84$ .

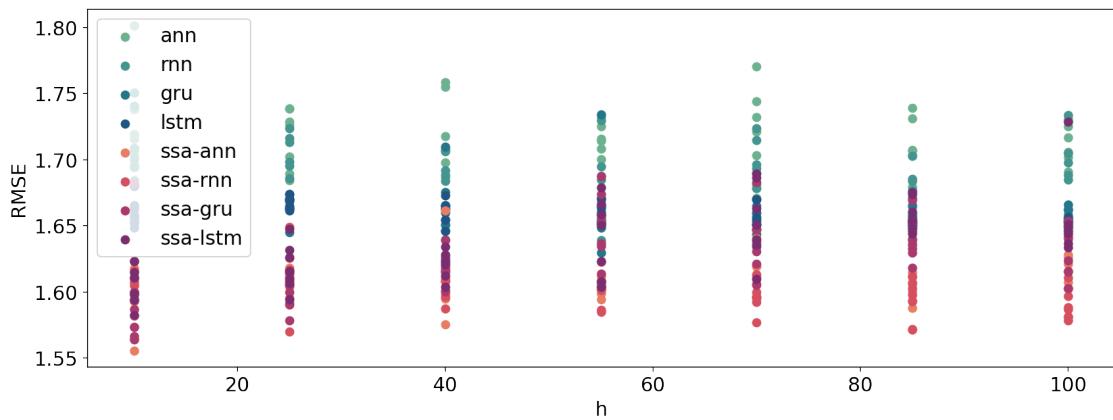


Рис. 4.57. Проверка устойчивости.  $r = 6$ ,  $L = 12$ .

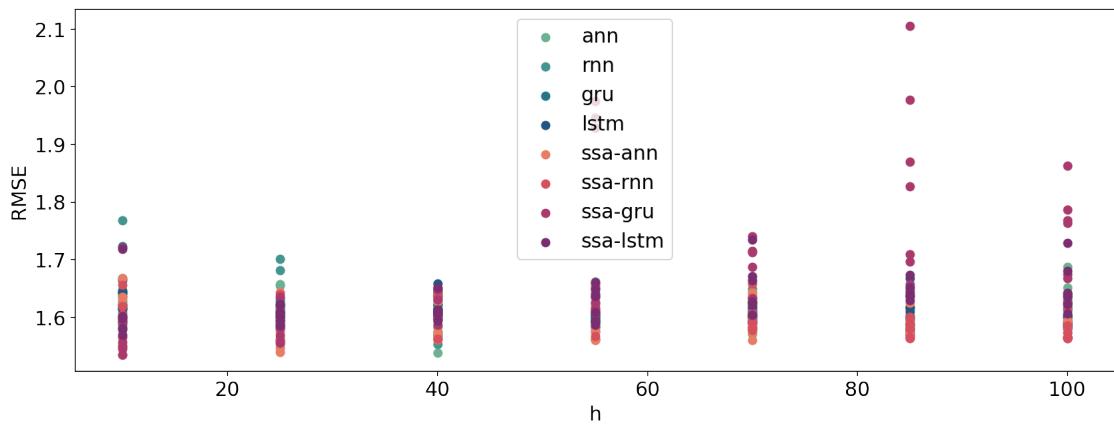


Рис. 4.58. Проверка устойчивости.  $r = 6$ ,  $L = 84$ .

## Глава 5

### Реальные данные

#### 5.1. Среднемесячные осадки в Индии

Рассмотрим следующий ряд  $Z_{1500}$  (рис. 5.1) взятый из статьи [1]. Данные «Indian Rain» показывают среднемесячное количество осадков в Индии.

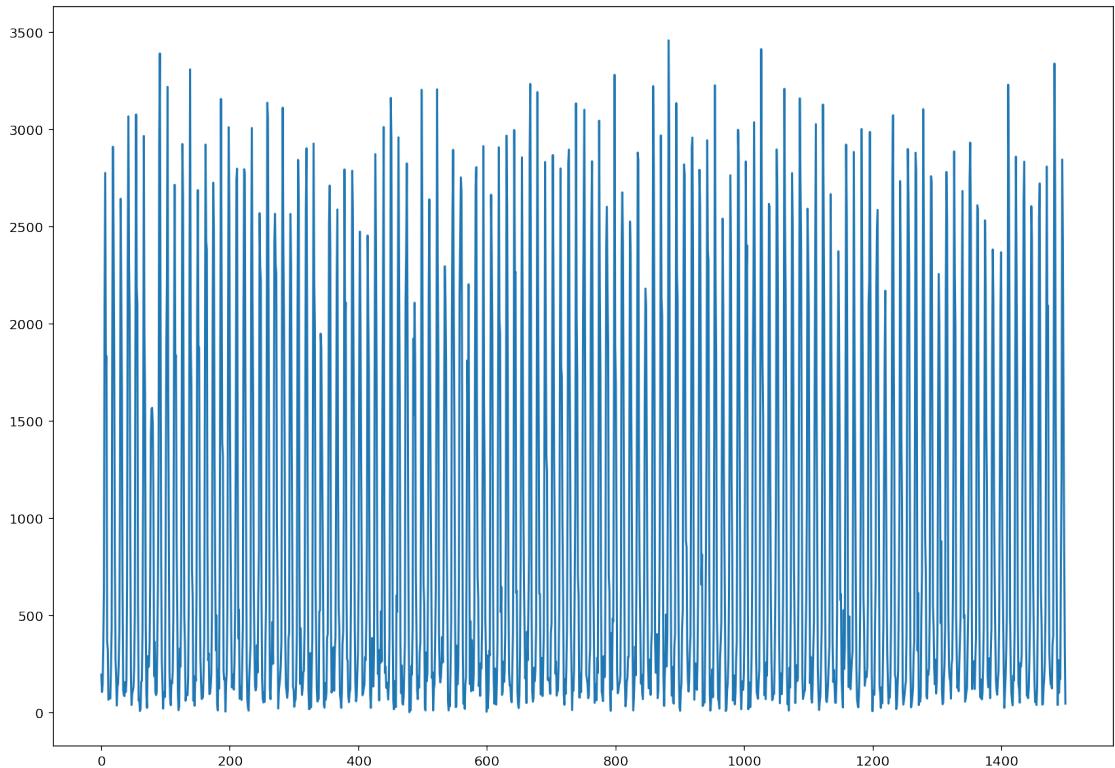


Рис. 5.1. Данные «Indian Rain».

В экспериментах будем разбивать ряд  $Z_{1500}$  на тренировочную, валидационную, тестовую выборки по 750, 500, 250 точек соответственно.

На рисунке 5.2 можно увидеть периодограмму ряда. Видно, что ряд имеет три периодики и трендовую составляющую. Исходя из рис. 5.1 это константный тренд. Ввиду этого, будем считать параметры  $r = 7$  и  $L = 375$  аналитически верными для метода SSA и гибридных методов. Так как ранг ряда скорее всего равен 7, а  $L = 375$  удовлетворяет асимптотической разделимости. Также, так как это данные по месяцам, то период ряда кратен 12. Далее в экспериментах будем перебирать параметры  $T$  и  $L$  по сетке с шагом кратным 12.

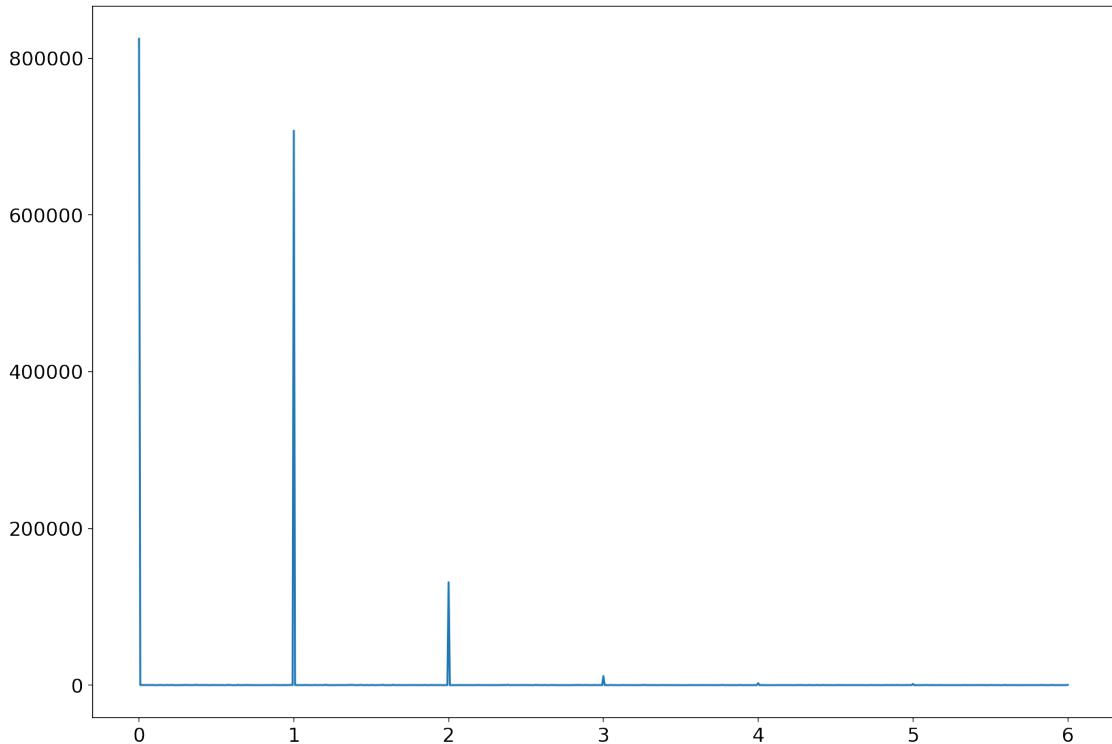


Рис. 5.2. Периодограмма ряда Z<sub>1500</sub>.

### 5.1.1. Сравнение прогнозов, полученных с помощью метода SSA, обычных и гибридных методов

Сравним метод SSA, обычные и гибридные методы по способу, описанным в главе 3.6.

#### Прогноз по SSA

Сравним точность прогнозирования методом SSA при разных параметрах. Зададим следующую сетку параметров  $L = \{12, 24, \dots, 375\}$ ,  $r = \{5, 7, 9, 11\}$ . Посмотрим на результаты на рисунке 5.3. На графике видно, что наилучшие результаты достигаются при  $r = 11$ . Разницы в параметрах  $L$  нет. Далее посмотрим на две пары параметров  $r = 11$ ,  $L = 375$  и  $r = 7$ ,  $L = 375$ .

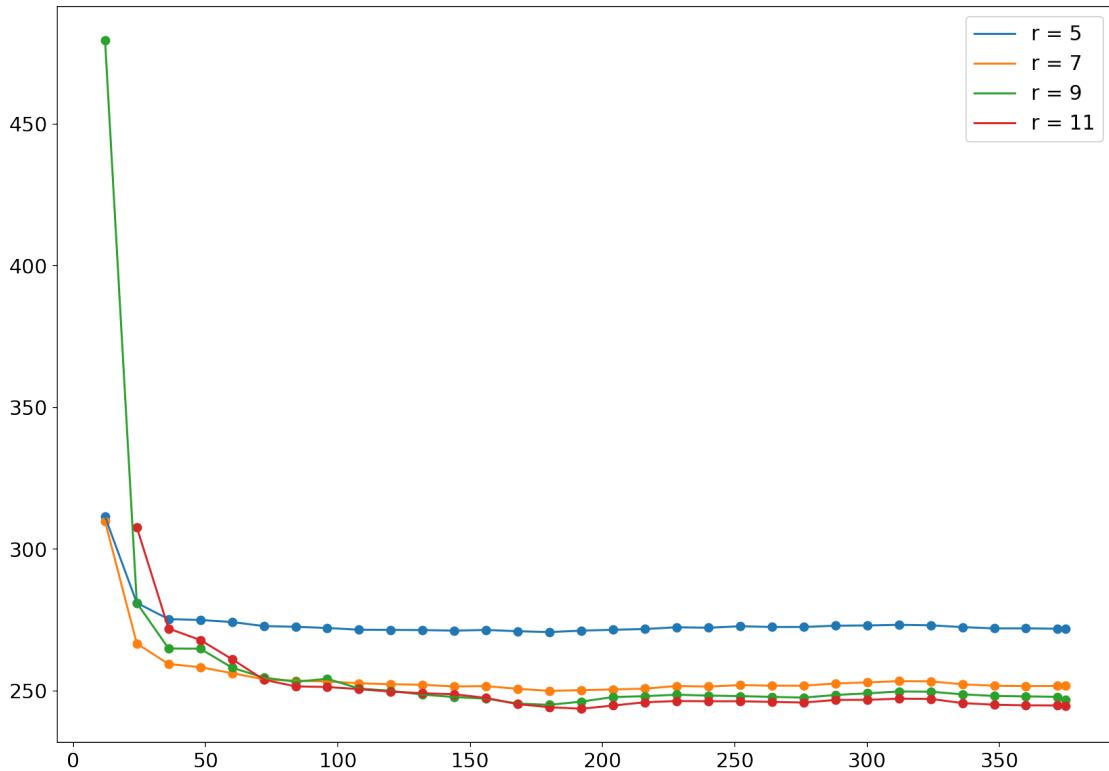


Рис. 5.3. Прогнозирование с помощью метода SSA.

### Восстановление SSA

Посмотрим, как метод SSA восстанавливает тренировочную выборку для выбранных пар на рис. 5.4, 5.5, 5.6. На графиках видно, что метод весьма хорошо выделил сигнал. Для  $r = 7$  пики в оценке одинаковые. Для  $r = 11$  пики у оценки немного скачут, возможно это эффект шума, попавшего в ряд. Далее будем использовать параметр  $r = 7$  в методе SSA и гибридных параметрах, параметр  $L = 375$  в гибридных моделях.

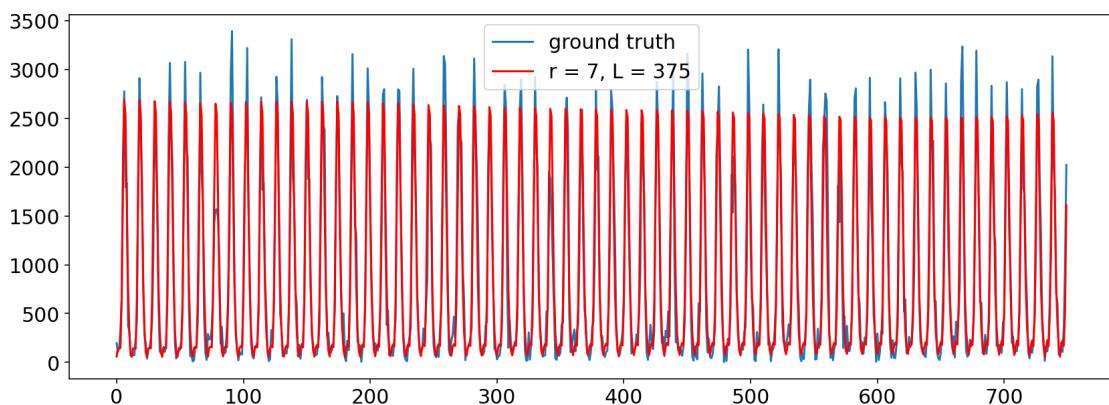


Рис. 5.4. Восстановление тренировочной выборки с помощью метода SSA.  $r = 7$ ,  $L = 375$ .

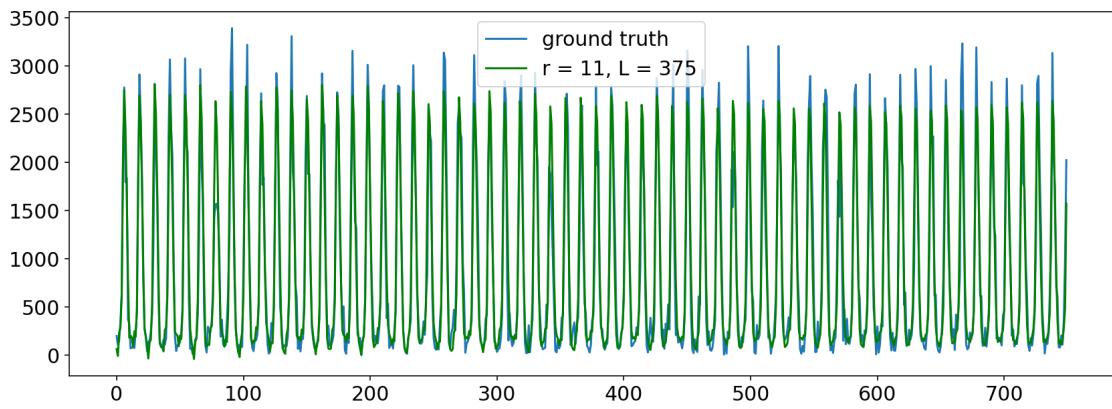


Рис. 5.5. Восстановление тренировочной выборки с помощью метода SSA.  $r = 11$ ,  $L = 375$ .

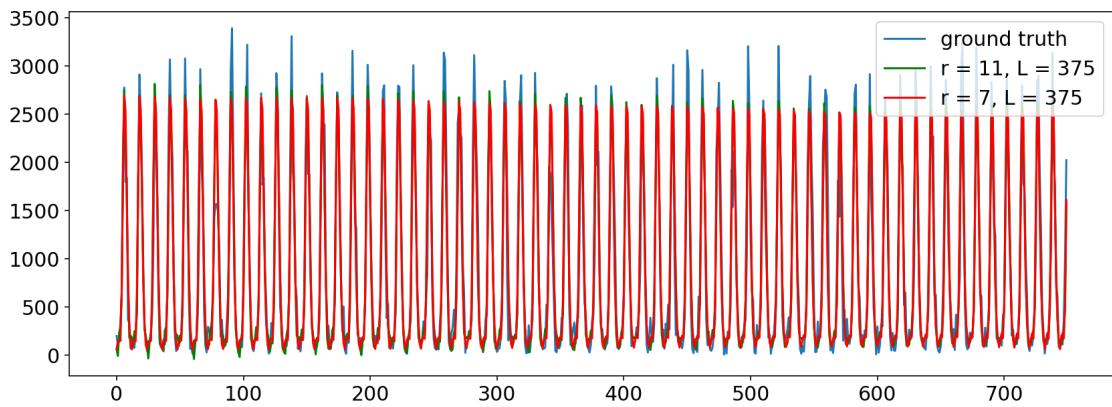


Рис. 5.6. Восстановление тренировочной выборки с помощью метода SSA. Обе пары.

## Сравнение методов

Пусть задана следующая сетка параметров:  $T = \{12, 48, \dots, 480\}$ ,  $h = \{10, 25, \dots, 100\}$ .

Для метода SSA зафиксируем параметр  $r = 7$ , а параметр  $L$  будем перебирать по сетке  $\{12, 24, \dots, 480\}$ .

На рис. 5.7, 5.8 видно, что гибридные методы показывают наилучшие результаты. Прогноз с помощью метода SSA находится посередине. На графике 5.7 особенно виден отрыв для  $T = 12$  (левый край графика). Также видно, что у гибридных методов почти нет зависимости ошибки от выбора параметра  $T$ .

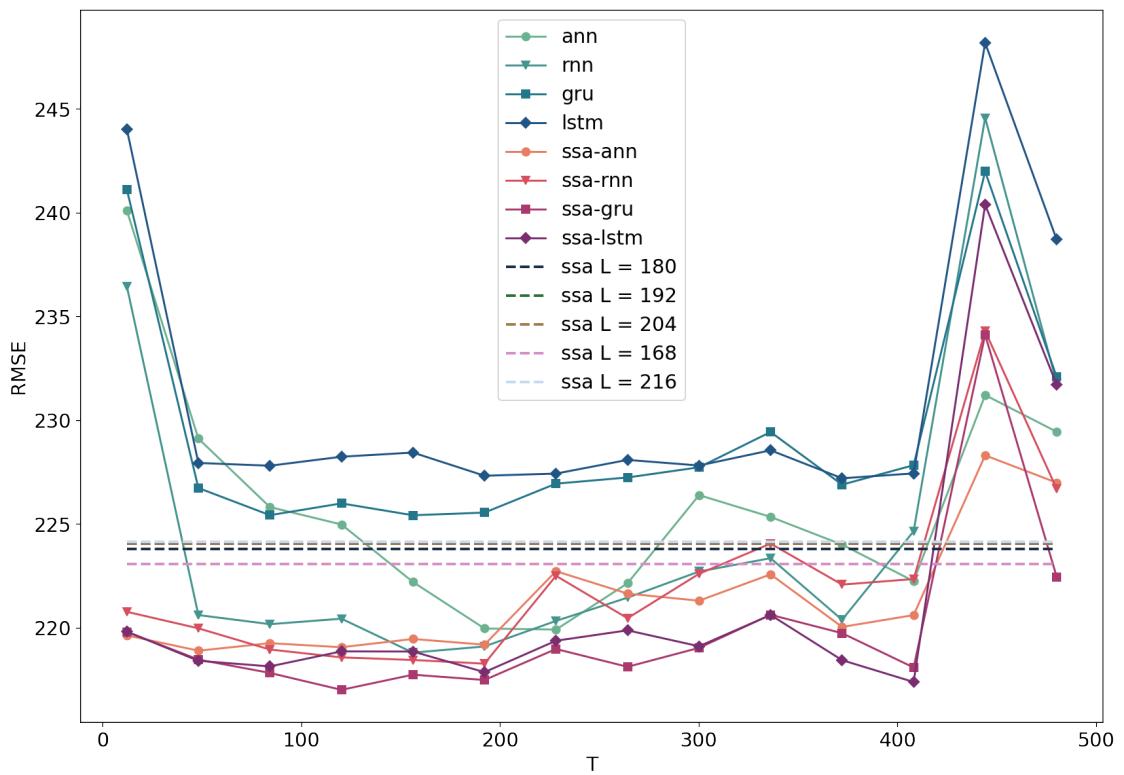


Рис. 5.7. Сравнение методов на данных «Indian Rain».

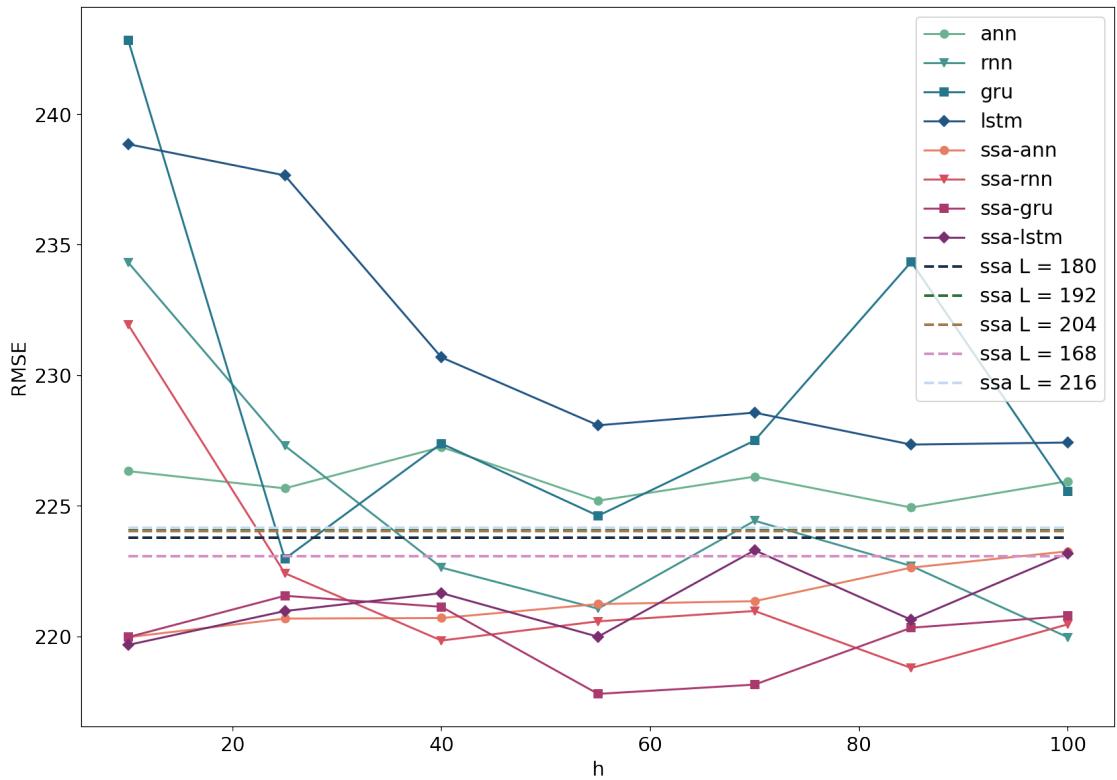


Рис. 5.8. Сравнение методов на данных «Indian Rain».

## Отображение результатов

На графиках ниже видно, что прогнозирование обычными и гибридными методами практически одинаковое.

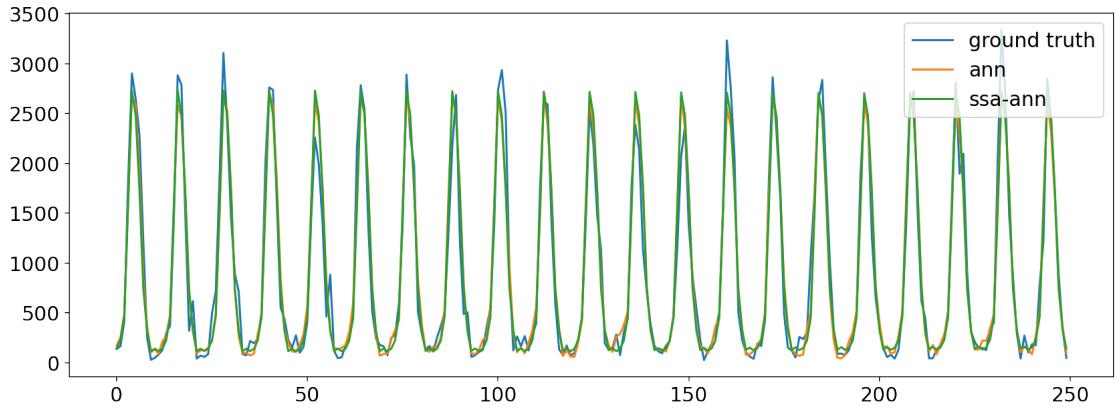


Рис. 5.9. Отображение результатов для ANN и SSA-ANN.

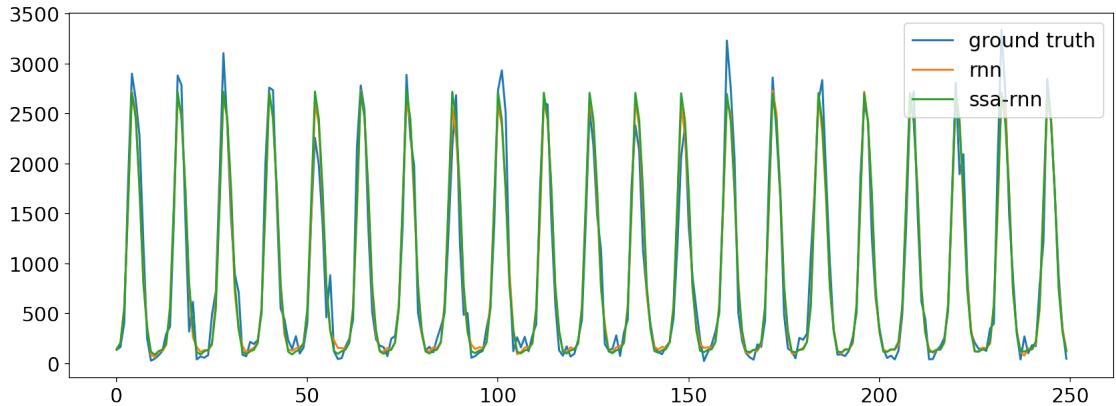


Рис. 5.10. Отображение результатов для RNN и SSA-RNN.

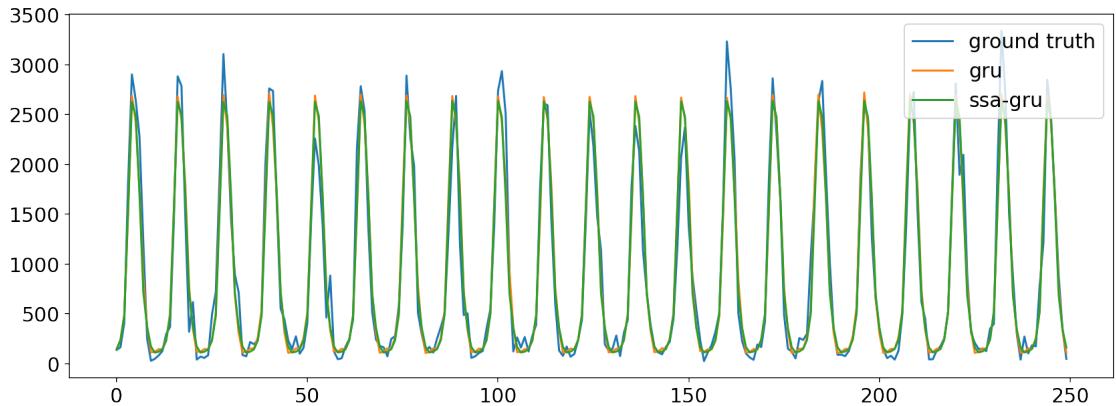


Рис. 5.11. Отображение результатов для GRU и SSA-GRU.

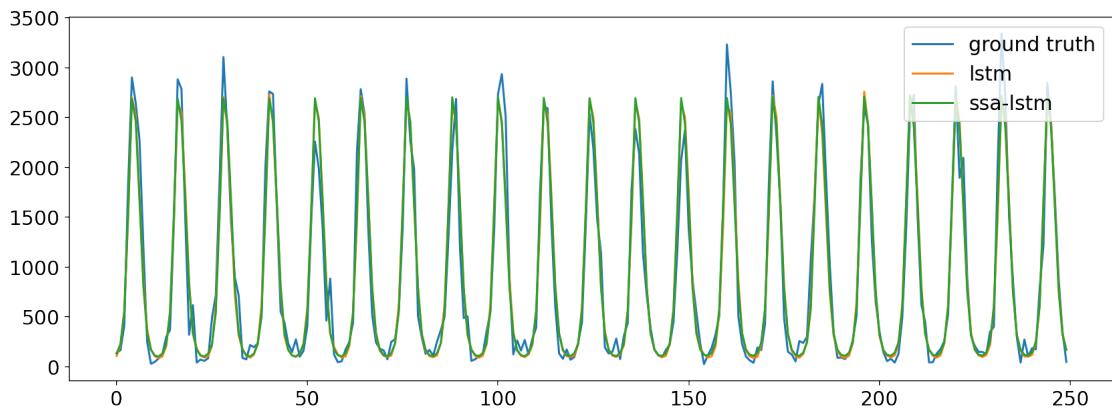


Рис. 5.12. Отображение результатов для LSTM и SSA-LSTM.

### Проверка устойчивости

Чтобы исключить случайность в полученных результатах, проведем сравнение для разных начальных весов методов. Зафиксируем новую сетку для параметра  $T = \{12, 156\}$ . Сетка для параметр  $h$  останется прежней. Будем получать каждый результат по 7 раз, инициализируя метод с новыми весами. Полученные результаты отображены на рис. 5.13, 5.14.

На рисунках видно, что полученные ранее результаты не случайны. Гибридные модели показывают лучшие результаты, особенно хорошо это видно для  $L = 12$ .

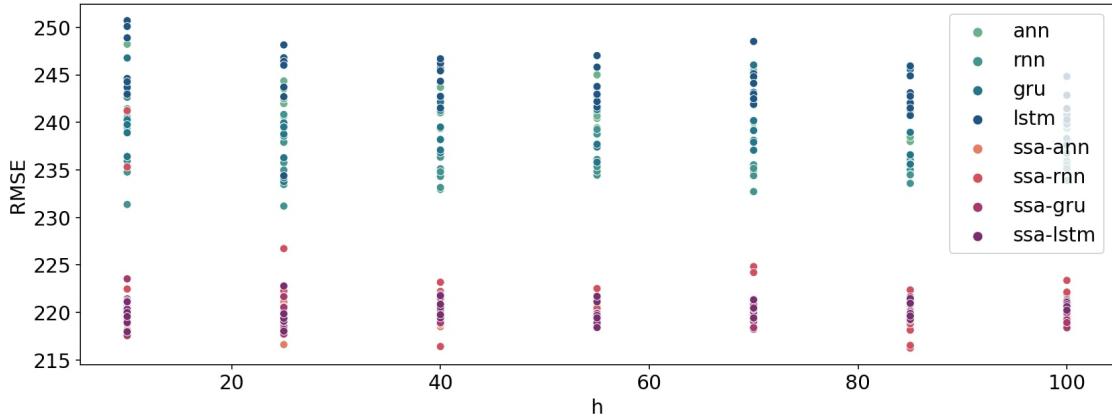


Рис. 5.13. Проверка устойчивости.  $L = 12$ .

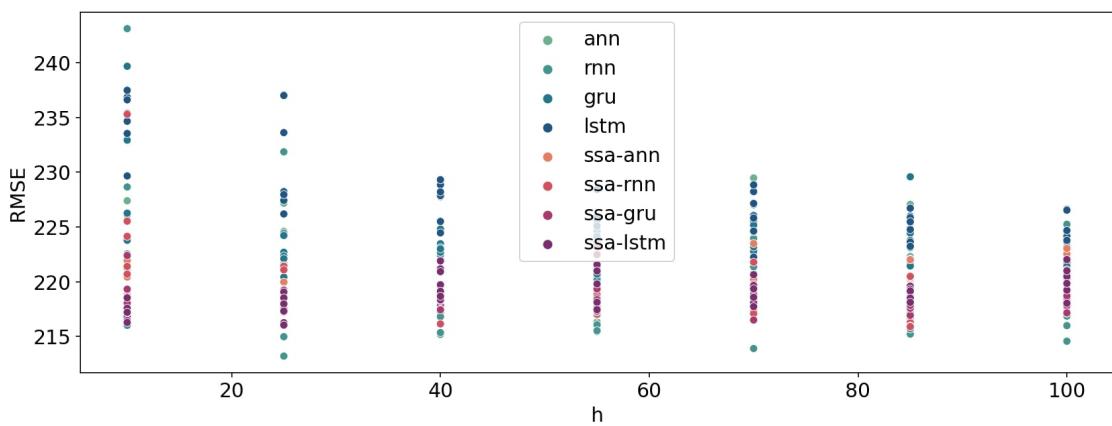


Рис. 5.14. Проверка устойчивости.  $L = 156$ .

## Выводы

Эксперимент показал, что данных «Indian rain» использование гибридных методов приводит к хорошему приросту в точности. Также использование гибридных методов снижает зависимость ошибки от выбора параметров модели, что позволяет выбрать менее сложную модель, а также увеличивает количество пар «признаки – предсказываемые значения».

## 5.2. Данные Earth Orientation Parameters (EOP)

Рассмотрим следующий ряд  $Z_{717}$  (рис. 5.15). Временной ряд отображает координату по оси абсцисс земного полюса.

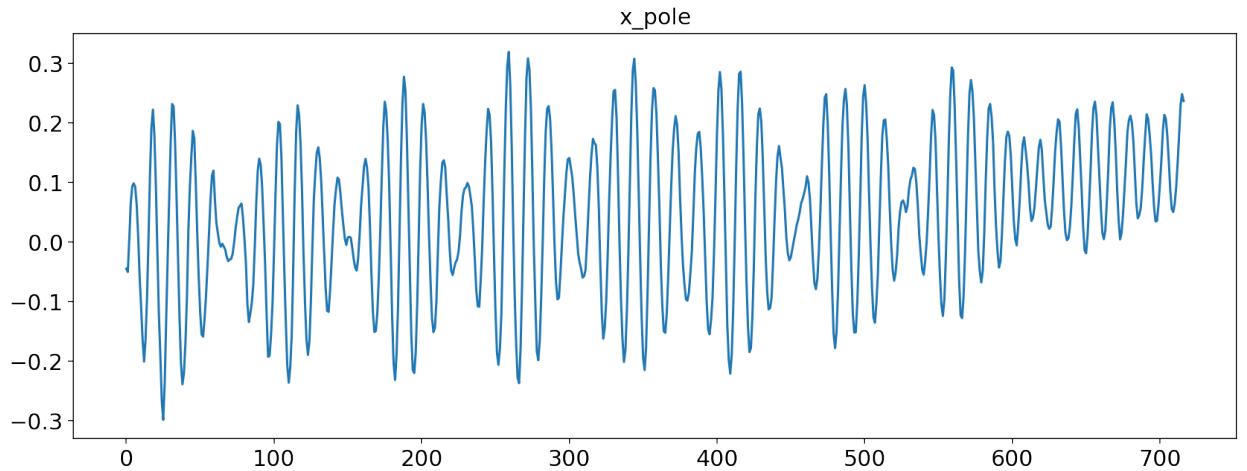


Рис. 5.15. Ряд  $Z_{717}$ .

Вычтем тренд из  $Z_{717}$  (рис. 5.16). Будем считать, что правая часть ряда не похожа на остальной временной ряд, поэтому удалим ее, чтобы избежать искаженных результатов. Эта операция сократит размер ряда  $Z_{717}$  до 620 точек. Обозначим ряд слева от вертикальной черты  $Z_{620}$ . Далее все эксперименты в разделе 5.2 проводятся на  $Z_{620}$ . Также, так как это данные по месяцам, то период ряда кратен 12. Но исходя из периодограммы, видно, что с пиком в 1 есть еще пик, который вдвое больше. Считаем, что это пик в 14, тогда общий период ряда равен 13. Далее в экспериментах будем перебирать параметры  $T$  и  $L$  по сетке с шагом кратным 13.

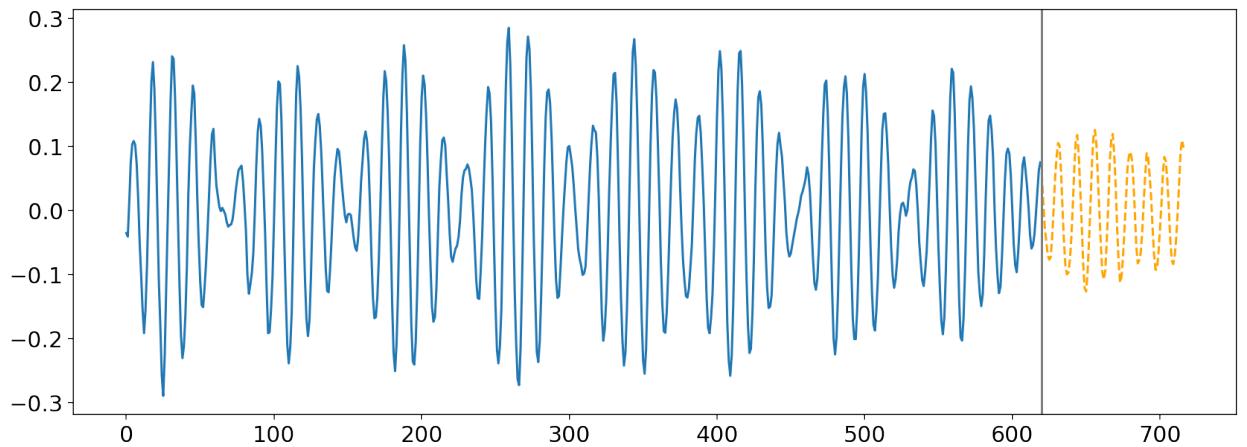


Рис. 5.16. Ряд  $Z_{717}$  без тренда.

В экспериментах будем разбивать ряд  $Z_{620}$  на тренировочную, валидационную, тестовую выборки по 320, 150, 150 точек соответственно.

Посмотрим на периодограмму ряда  $Z_{620}$  на рис. 5.17. На графике видно две близкие периодики, которые смешались. Это говорит, что у ряда сложный сигнал, который будет трудно выделить корректно.

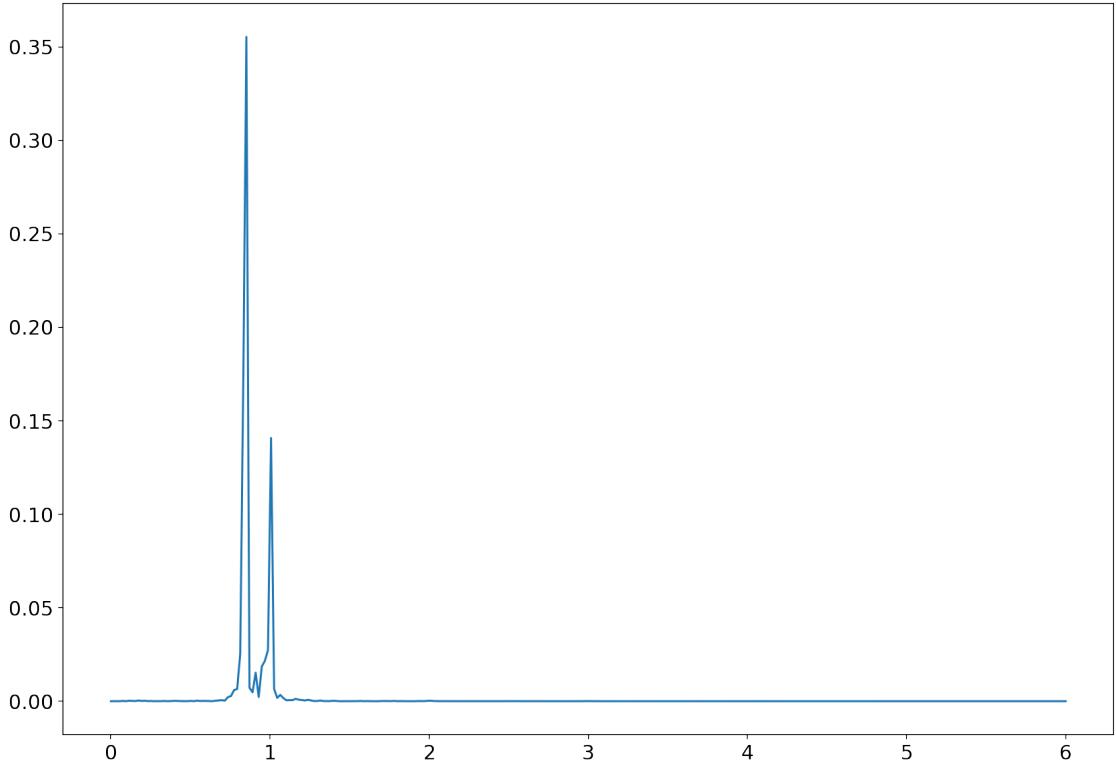


Рис. 5.17. Периодограмма ряда  $Z_{620}$ .

### 5.2.1. Сравнение прогнозов, полученных с помощью метода SSA, обычных и гибридных методов

Сравним метод SSA, обычные и гибридные методы по способу, описанным в главе 3.6. В разделе «Прогноз по SSA» подберем оптимальные параметры для метода SSA и гибридных моделей.

#### Прогноз по SSA

Сравним точность прогнозирования методом SSA при разных параметрах. Зададим следующую сетку параметров  $L = \{13, 26, \dots, 160\}$ ,  $r = \{6, 8, 12, 16, 18\}$ .

Посмотрим на результаты на рисунке 5.18. На графике видно, что с ростом  $L$  растет и ошибка. Наилучшая точность достигается при  $r = \{12, 16, 18\}$  и маленьком

$T$ . Такие результаты могут свидетельствовать о том, что сигнал ряда  $Z_{620}$  не конечно-го ранга. Использование маленького  $T$  и большого  $r$  в препроцессинге SSA приведет к сильной аппроксимации оценкой сигнала ряда, что делает использование SSA неце-лесообразным. Возьмем пару параметров по середине:  $L = 78$ ,  $r = 18$ , далее будем использовать эту пару в гибридных моделях.

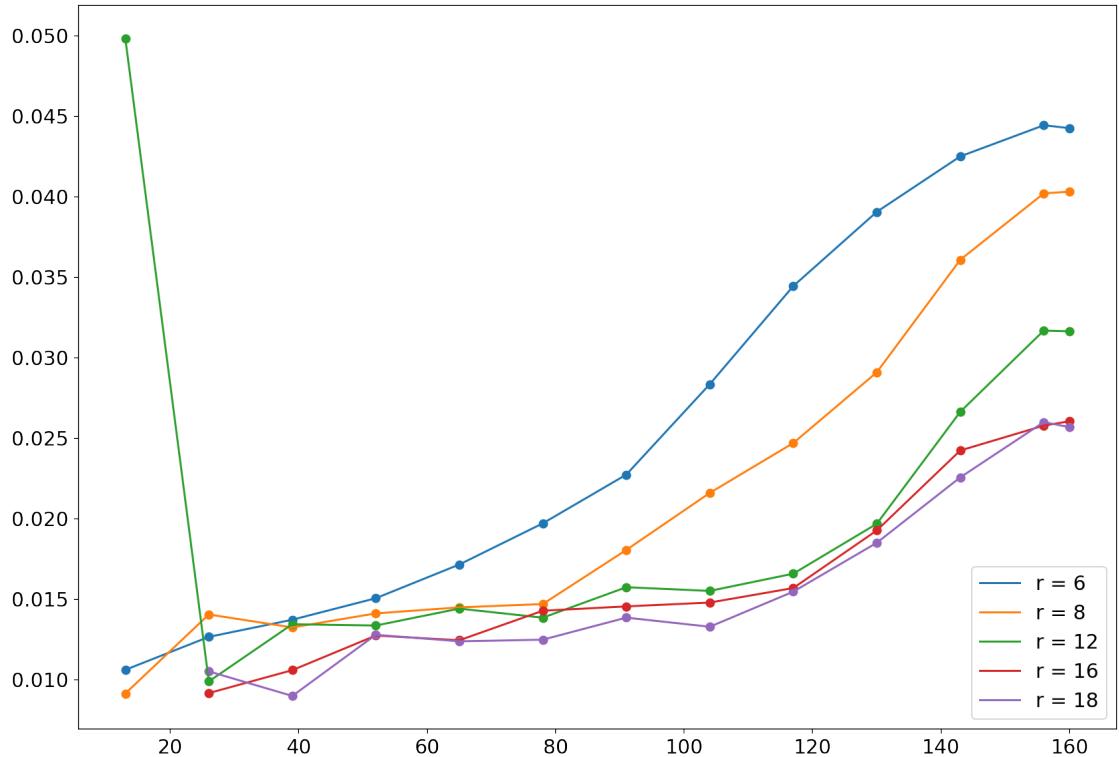


Рис. 5.18. Прогнозирование с помощью метода SSA.

## Восстановление SSA

Посмотрим, как метод SSA восстанавливает тренировочную выборку для выбранных пар на рис. 5.19. На графике видно, что метод неплохо выделил сигнал. Видно, что оценка сигнала сильно аппроксимирует временной ряд, кроме нескольких мест.

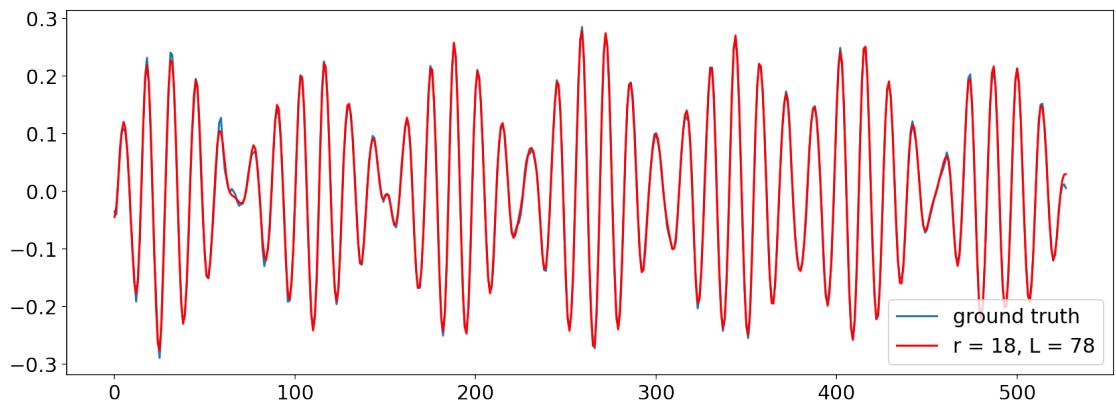


Рис. 5.19. Восстановление тренировочной выборки с помощью метода SSA.  $r = 18$ ,  $L = 78$

## Сравнение методов

Пусть задана следующая сетка параметров:  $T = \{13, 42, \dots, 143\}$ ,  $h = \{10, 25, \dots, 100\}$ .

Для метода SSA зафиксируем параметр  $r = 12$ , а параметр  $L$  будем перебирать по сетке  $\{13, 26, \dots, 130\}$ .

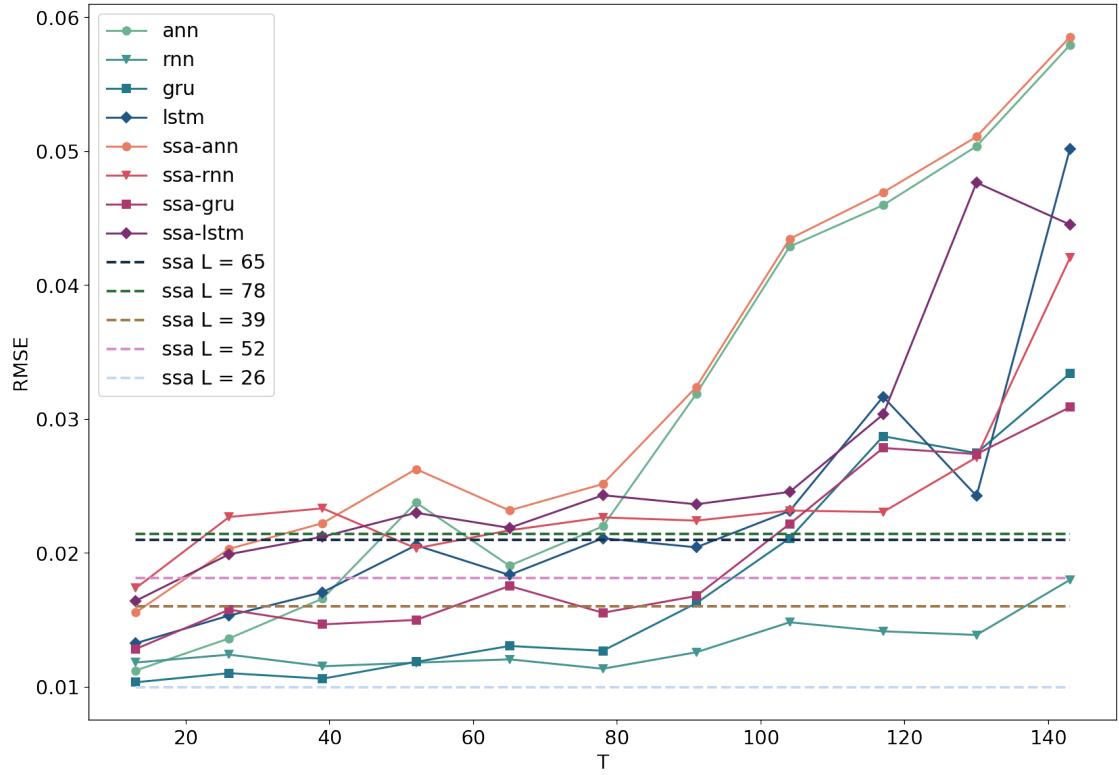


Рис. 5.20. Сравнение методов на данных ЕОР.

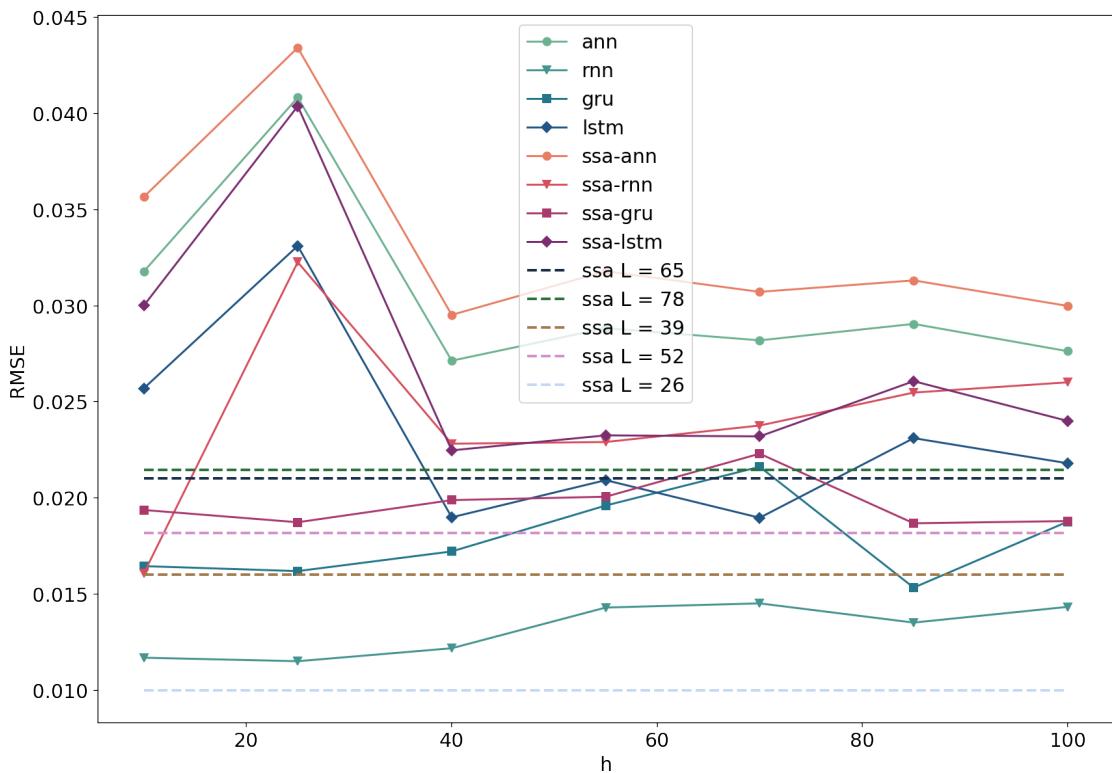


Рис. 5.21. Сравнение методов на данных ЕОР.

Посмотрим на результаты на рис. 5.20. На графике можно видеть, что наилучший результат показывает метод SSA с параметром  $L = 26$ . Видно гибридные методы сильно уступают в точности обычным методам. Методы GRU и RNN показывают хорошую точность.

### Отображение результатов

На графиках ниже видно, что прогнозирование сильно расходится в месте так называемого «перехода». Ближе всего в этом месте прогнозирует метод RNN, что видно на графике 5.23. Также есть сильное расхождение с сигналом ряда в конце прогноза (правая часть графиков). Для всех гибридных методов ошибка в этих местах больше, чем для обычных. В остальных местах значительной разницы между прогнозами не замечено.

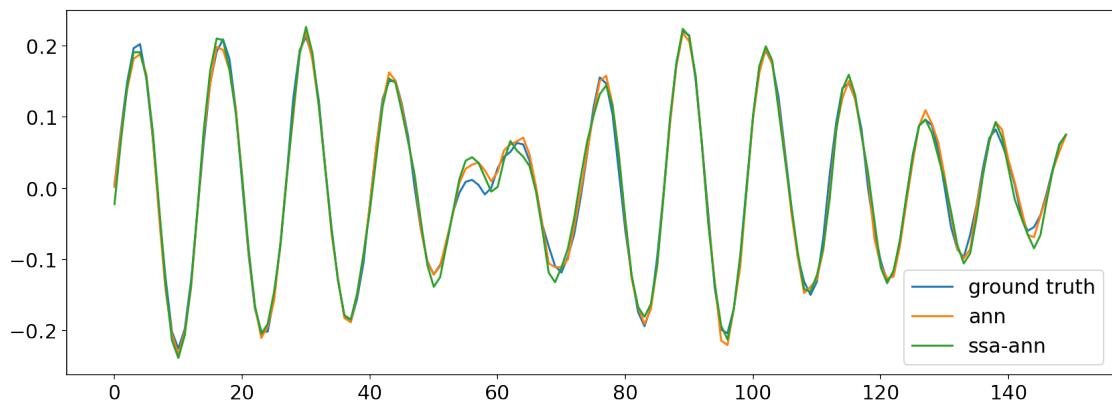


Рис. 5.22. Отображение результатов для ANN и SSA-ANN.

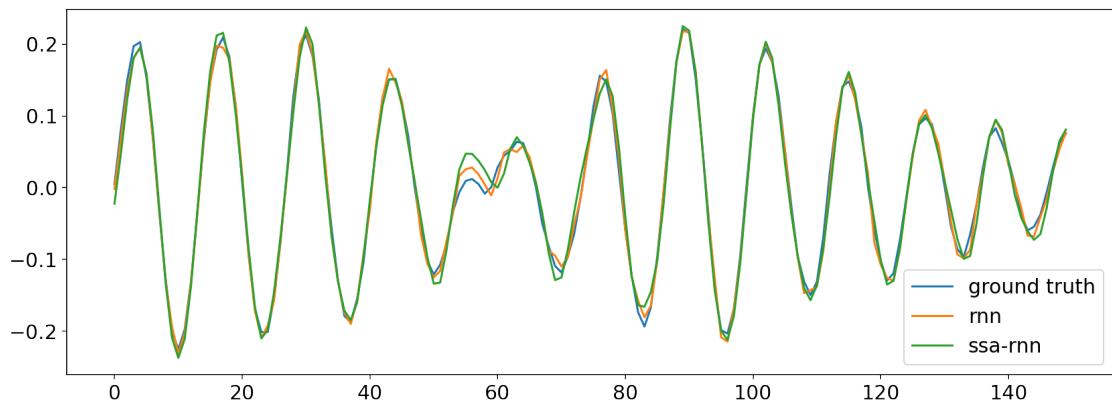


Рис. 5.23. Отображение результатов для RNN и SSA-RNN.

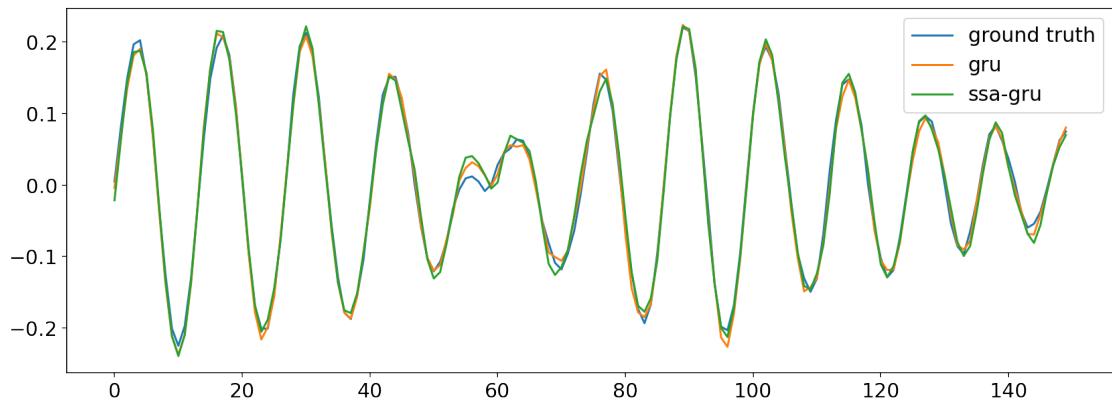


Рис. 5.24. Отображение результатов для GRU и SSA-GRU.

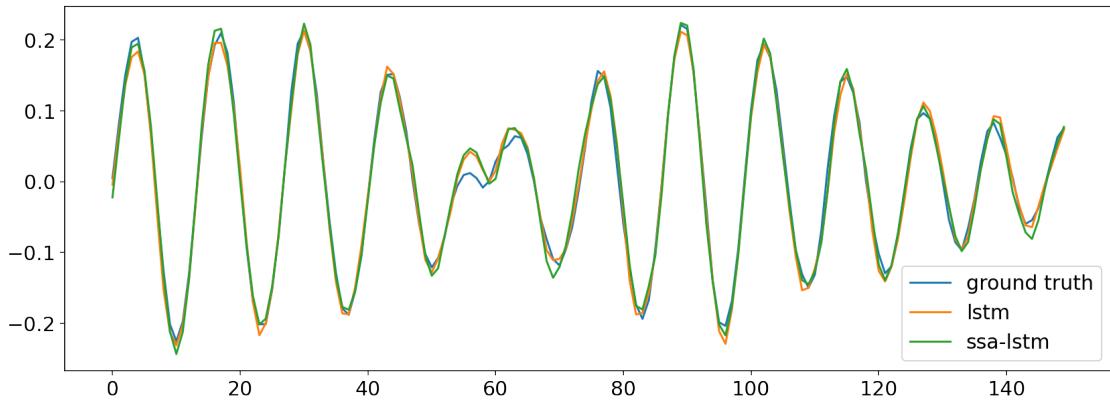


Рис. 5.25. Отображение результатов для LSTM и SSA-LSTM.

### Проверка устойчивости

Чтобы исключить случайность в полученных результатах, проведем сравнение для разных начальных весов методов. Зафиксируем новую сетку для параметра  $T = \{13, 91\}$ . Сетка для параметр  $h$  останется прежней. Будем получать каждый результат по 7 раз, инициализируя метод с новыми весами. Полученные результаты отображены на рис. 5.26, 5.27. На рисунках видно, что полученные ранее результаты не случайны.

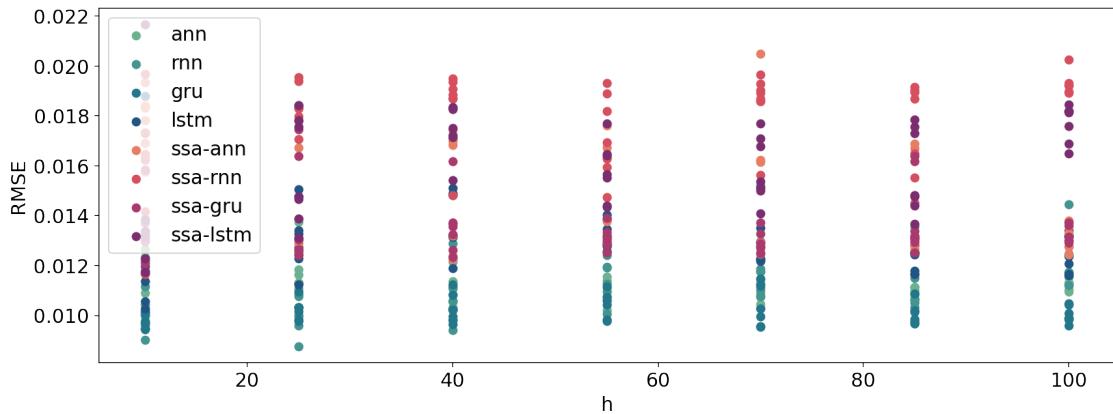
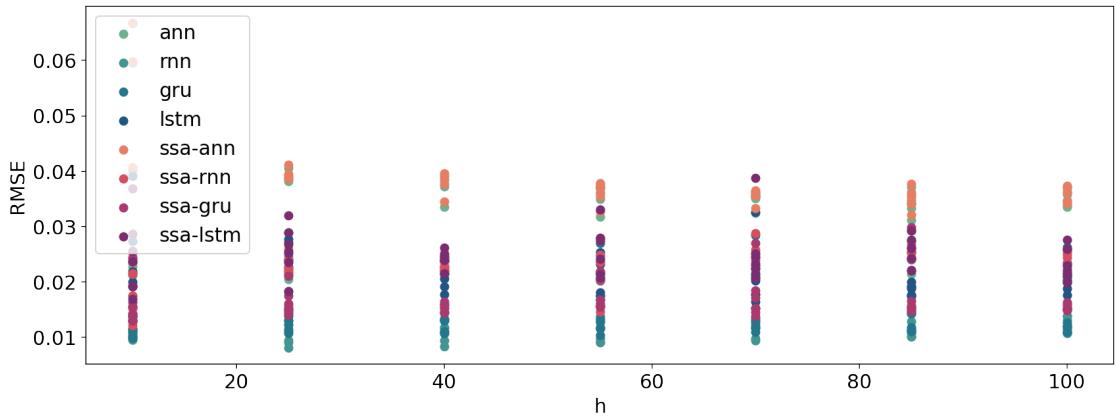


Рис. 5.26. Проверка устойчивости.  $L = 13$ .

Рис. 5.27. Проверка устойчивости.  $L = 91$ .

## Выводы

В ходе сравнение обычных, гибридных методов и метода SSA наилучшие результаты показал метод SSA с параметрами  $L = 26, r = 12$ . Но использование таких параметров в гибридных моделях нецелесообразно, так как оценка сигнала слишком сильно аппроксимирует ряд. Гибридные методы показали результаты хуже, чем обычные методы. Причина таких результатов может быть нехватка данных или сложный вид сигнала, который метод SSA не может выделить. Также в данном ряде не наблюдается наличие шума, что также может ставить использование метода SSA в качестве препроцессинга под сомнение.

### 5.3. Погода

Рассмотрим следующий ряд  $Z_{828}$  (рис. 5.28). Данный ряд отображает одну из характеристик погоды в городе Санкт-Петербурге.

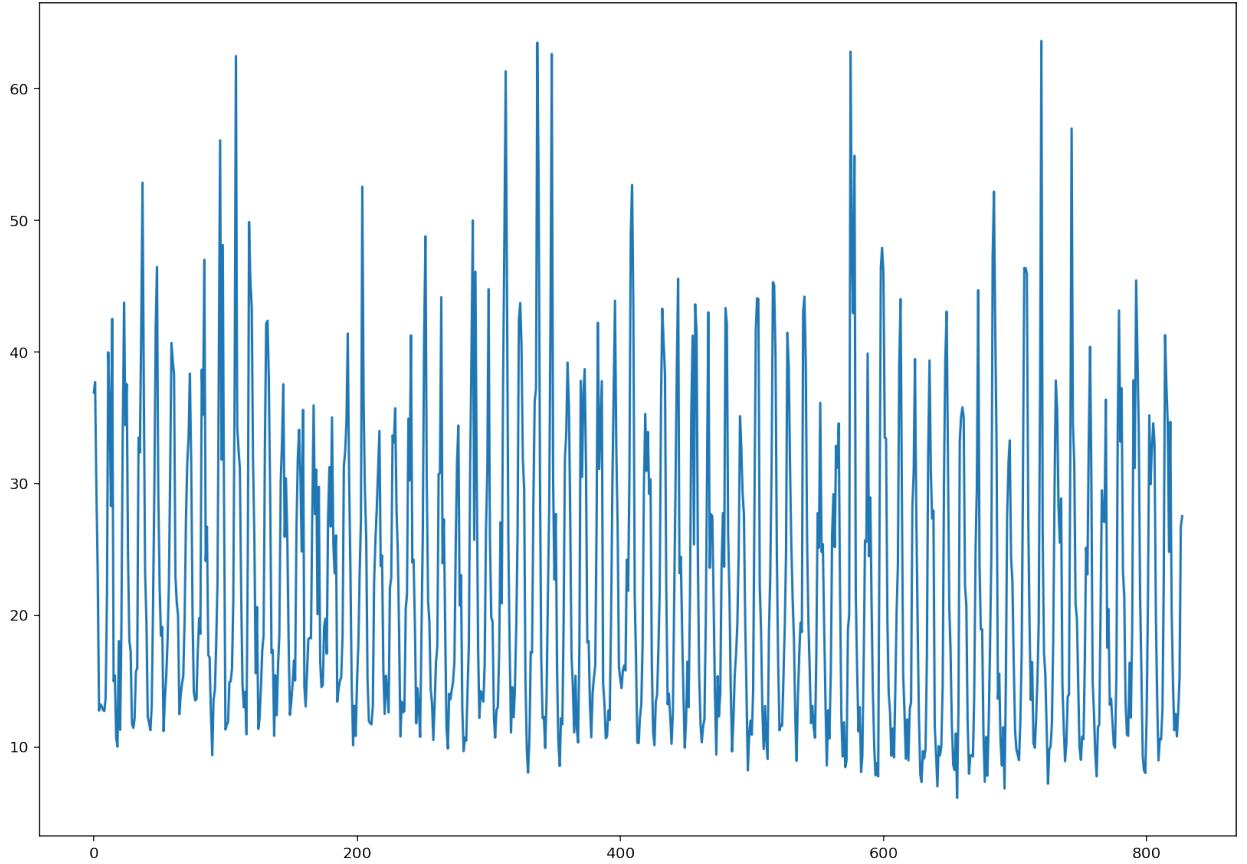


Рис. 5.28. Ряд  $Z_{828}$ .

В экспериментах будем разбивать ряд  $Z_{828}$  на тренировочную, валидационную, тестовую выборки по 528, 150, 150 точек соответственно.

На рис. 5.29 представлена периодограмма ряда  $Z_{828}$ . Из нее видно, что у ряда есть тренд и две периодики. Одна периодика слабо выраженная и может смешаться с шумом. Ввиду этого, будем считать параметры  $r = 5$  и  $L = 264$  аналитически верными для метода SSA и гибридных методов. Так как ранг ряда скорее всего равен 5, а  $L = 264$  удовлетворяет асимптотической разделимости. Также, так как это данные по месяцам, то период ряда кратен 12. Далее в экспериментах будем перебирать параметры  $T$  и  $L$  по сетке с шагом кратным 12.

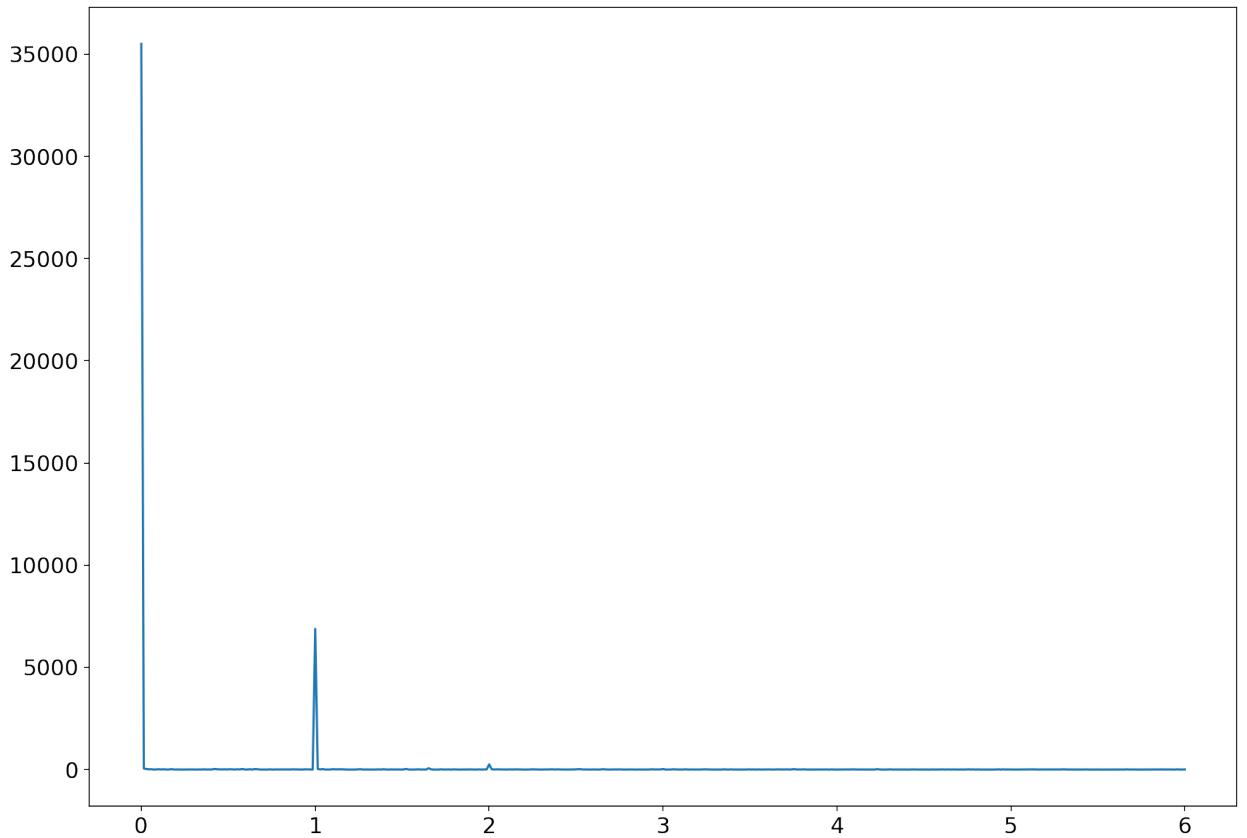


Рис. 5.29. Периодограмма ряда Z<sub>828</sub>.

### 5.3.1. Сравнение прогнозов, полученных с помощью метода SSA, обычных и гибридных методов

Сравним метод SSA, обычные и гибридные методы по способу, описанным в главе 3.6.

#### Прогноз по SSA

Сравним точность прогнозирования методом SSA при разных параметрах. Зададим следующую сетку параметров  $L = \{12, 24, \dots, 264\}$ ,  $r = \{3, 5, 7, 9\}$ . Посмотрим на результаты на рисунке 5.30. На графике видно, что наилучшие результаты достигаются при  $r = 5$ . Нету сильной разницы в точности при  $r = 5$ , поэтому дальше будем рассматривать пару  $r = 5$ ,  $L = 264$ .

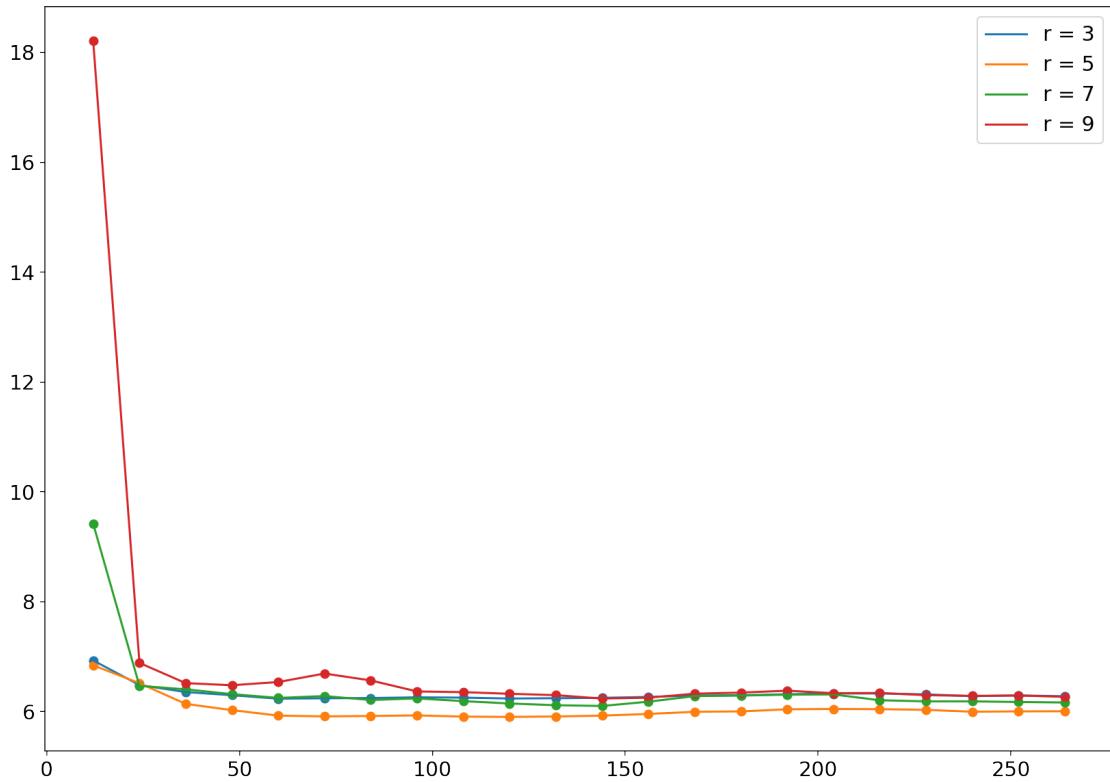


Рис. 5.30. Прогнозирование с помощью метода SSA.

### Восстановление SSA

Посмотрим, как метод SSA восстанавливает тренировочную выборку для выбранных пар на рис. 5.31. На графике видно, что метод неплохо выделил сигнал и оценка сигнала сильно не аппроксимирует временной ряд.

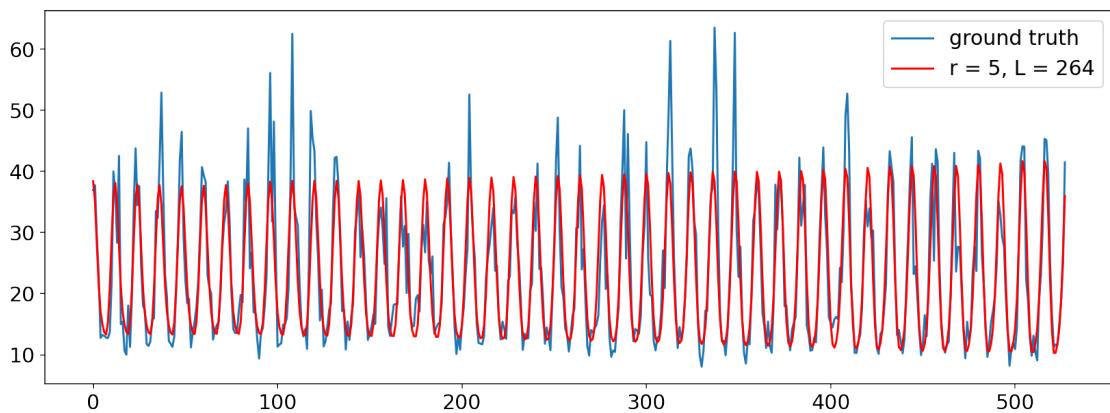


Рис. 5.31. Восстановление тренировочной выборки с помощью метода SSA.  $r = 5$ ,  $L = 264$

## Сравнение методов

Пусть задана следующая сетка параметров:  $T = \{12, 48, \dots, 120\}$ ,  $h = \{10, 25, \dots, 100\}$ .

Параметры для SSA в гибридных методах выберем  $L = 264$ ,  $r = 5$ . Для метода SSA зафиксируем параметр  $r = 5$ , а параметр  $L$  будем перебирать по сетке  $\{12, 24, \dots, 264\}$ .

Посмотрим на результаты на рис. 5.32, 5.33, видно, что лучшие результаты показывает метод SSA. Сложно оценить, какие методы показали себя лучше. На графике все результаты перемешаны, но кривые гибридных методов лежат немного ниже, чем обычных. Такие результаты могли получиться из-за маленького количества данных. А также возможно малого размера валидационной выборки, что ограничивает выбор параметра  $T$ .

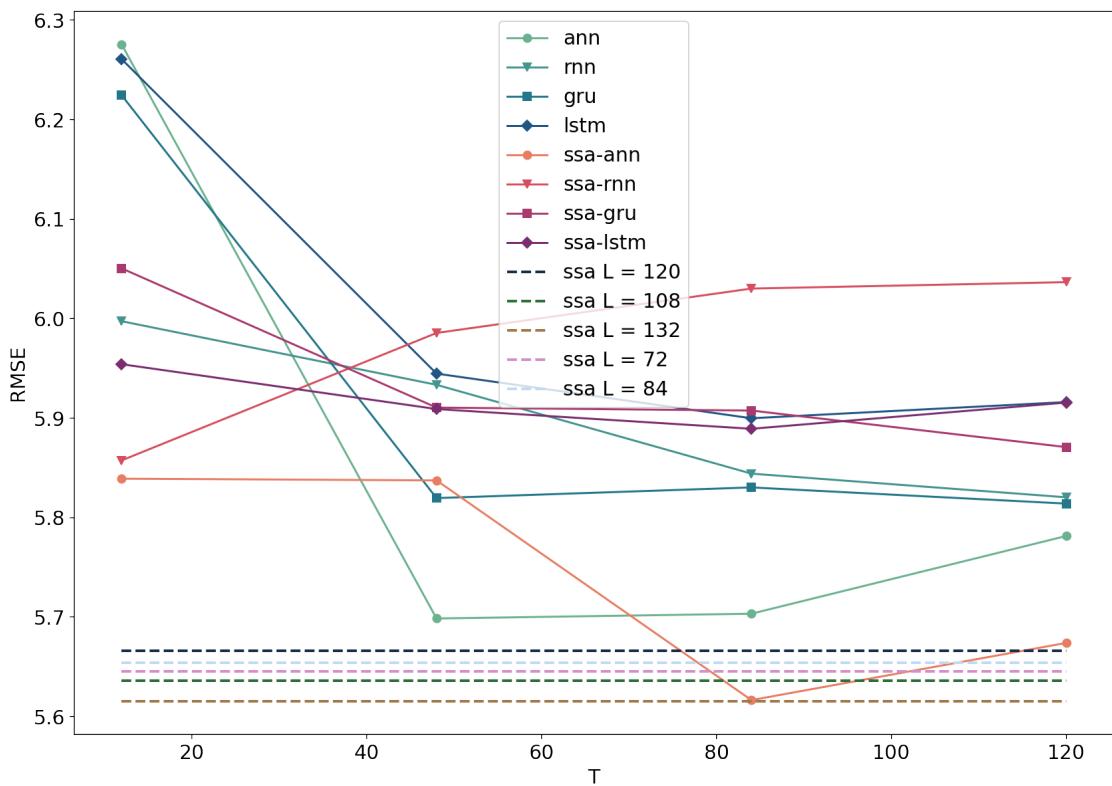


Рис. 5.32. Сравнение методов на данных погоды.

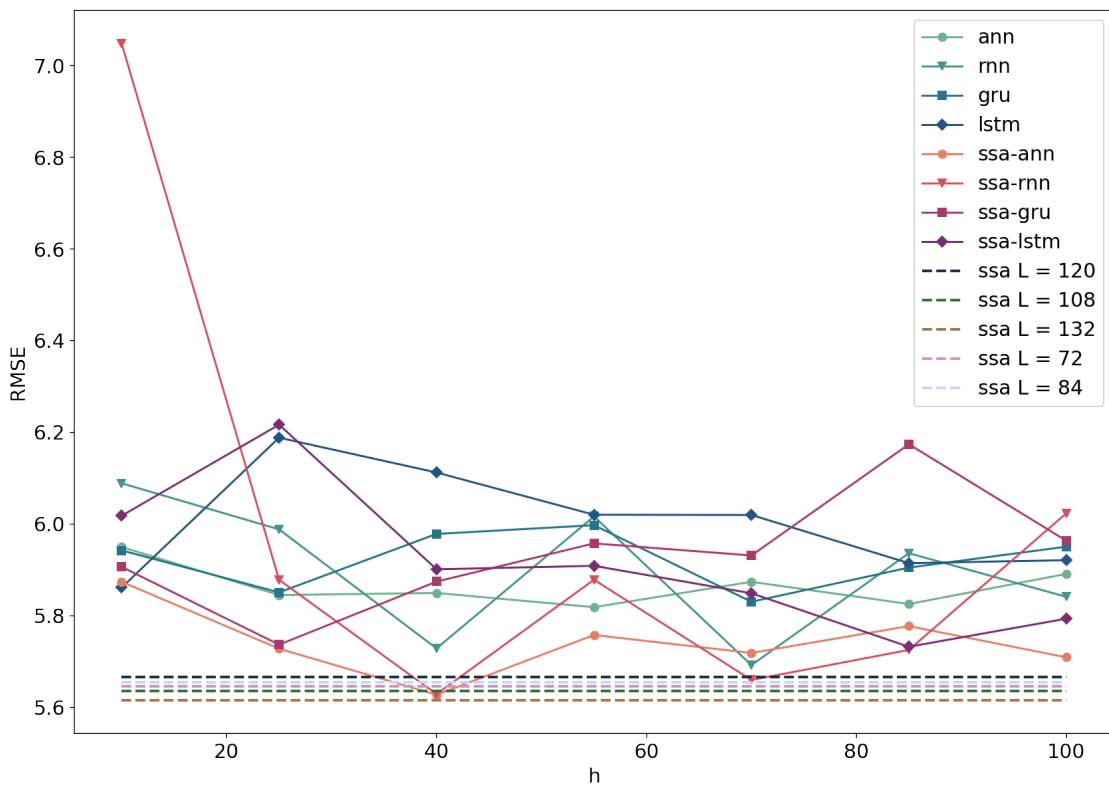


Рис. 5.33. Сравнение методов на данных погоды.

### Отображение результатов

На графиках ниже видно, что прогнозирование обычными и гибридными методами похожее. Но заметно, что прогноз гибридных методов более стабилен. Эффект этого объясняется препроцессингом SSA. Видно, что прогнозы обычных методов, не использующие препроцессинг, более подвержены влиянию шума в ряде.

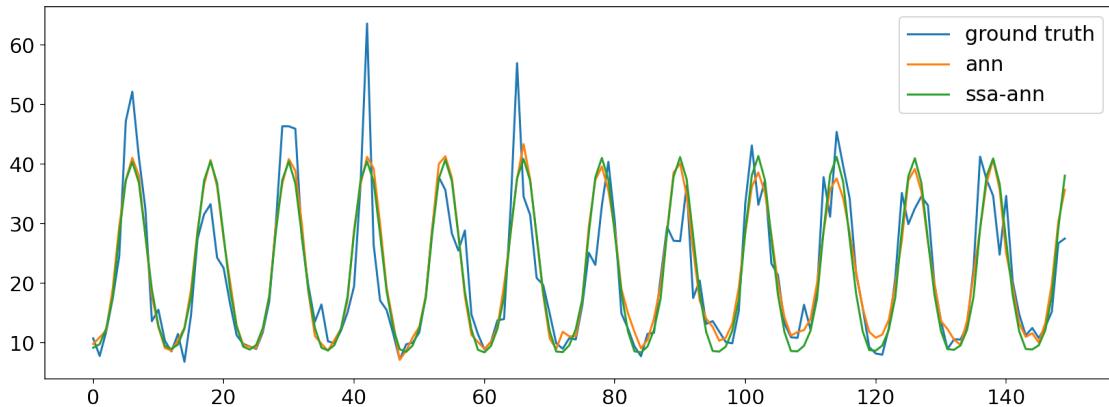


Рис. 5.34. Отображение результатов для ANN и SSA-ANN.

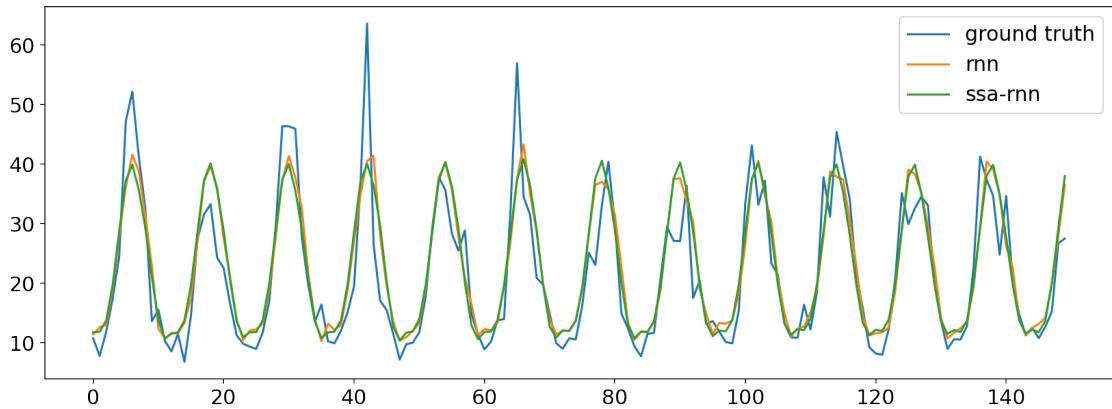


Рис. 5.35. Отображение результатов для RNN и SSA-RNN.

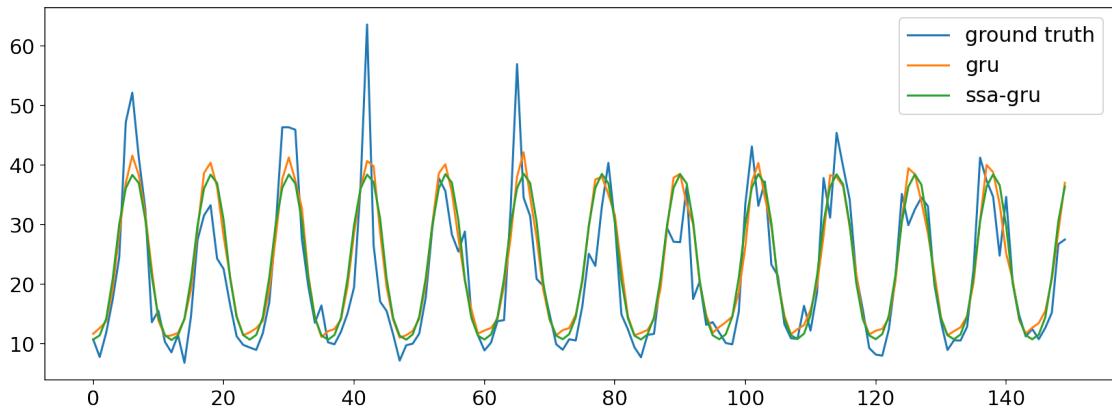


Рис. 5.36. Отображение результатов для GRU и SSA-GRU.

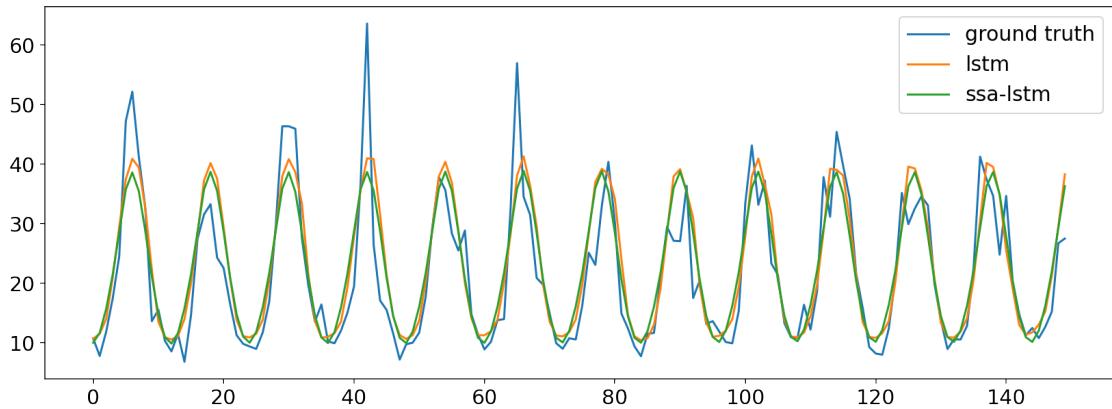


Рис. 5.37. Отображение результатов для LSTM и SSA-LSTM.

## Проверка устойчивости

Чтобы исключить случайность в полученных результатах, проведем сравнение для разных начальных весов методов. Зафиксируем новую сетку для параметра  $T =$

{12, 84}. Сетка для параметр  $h$  останется прежней. Будем получать каждый результат по 7 раз, инициализируя метод с новыми весами. Полученные результаты отображены на рисунках ниже. На них подтверждается, выводы сделанные ранее. Заключаем, что полученные результаты устойчивые.

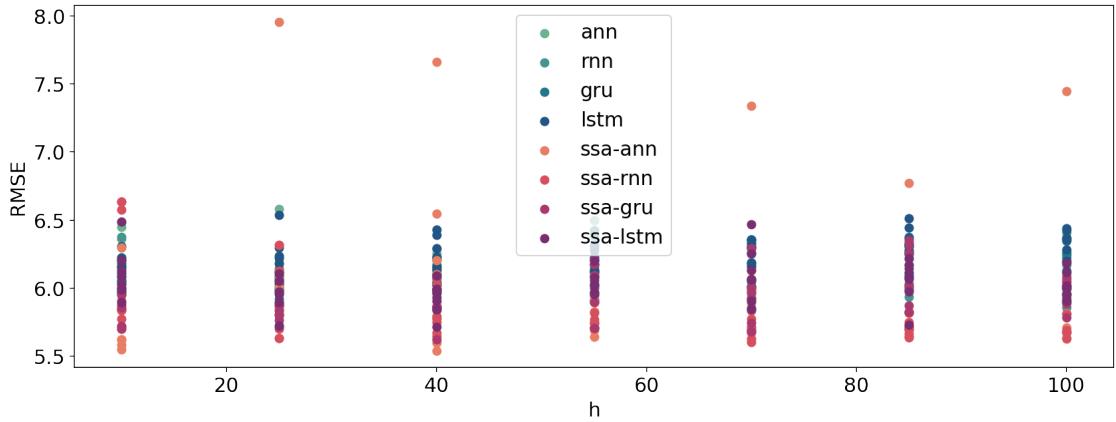


Рис. 5.38. Проверка устойчивости.  $L = 12$ .

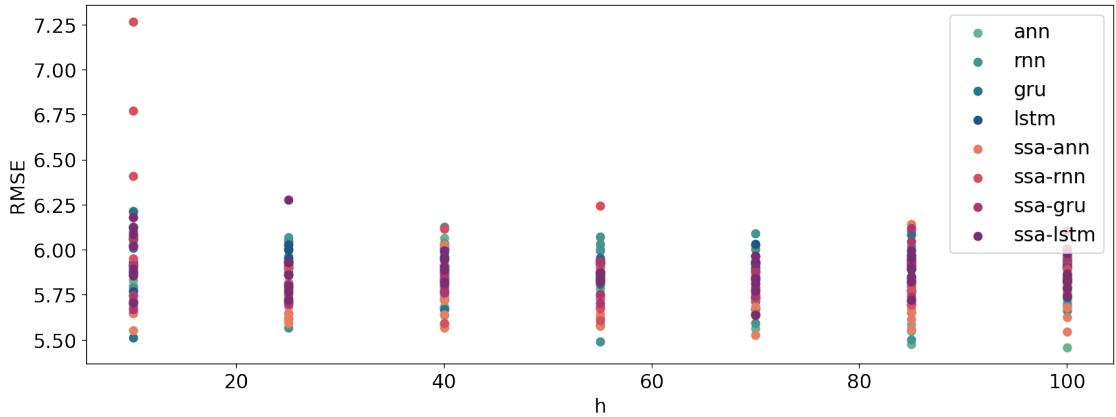


Рис. 5.39. Проверка устойчивости.  $L = 84$ .

## Выводы

Сравнение показало, что гибридные методы не дают прироста в точности на данных погоды. Такие результаты возможны из-за малого количества данных, нейронные сети не могут уловить зависимость в данных и получить общее представление о данных, хоть и метод SSA корректно выделил сигнал.

## Заключение

В работе был рассмотрена методика исследования сравнения методов, которая показалась весьма успешной. На данных Indian Rain удалось продемонстрировать успешное применение гибридных методов. Результаты на данных ЕОР и погоды в Санкт-Петербурге показали, что использование гибридных методов не всегда приводит к улучшению результата. Полученные результаты показали, что при маленьком шуме неразумно применять препроцессинг SSA, так как главная цель препроцессинга, отделить сигнал от шума. Также на модельных данных были проведены эксперименты, которые показали к каким результатам приводит использование гибридные методы при разных параметрах SSA.

## Список литературы

1. Du K., Zhao Y., Lei J. The wrong usage of singular spectral analysis and discrete wavelet transform in hybrid models to predict hydrological time series // Journal of Hydrology. — 2017. — 06. — Vol. 552.
2. Elman J. L. Finding Structure in Time // Cogn. Sci. — 1990. — Vol. 14. — P. 179–211.
3. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation / Cho K., van Merriënboer B., Gülcühre Ç., Bougares F., Schwenk H., and Bengio Y. // CoRR. — 2014. — Vol. abs/1406.1078. — arXiv : 1406.1078.
4. Hochreiter S., Schmidhuber J. Long Short-term Memory // Neural computation. — 1997. — 12. — Vol. 9. — P. 1735–80.
5. Golyandina N., Nekrutkin V., Zhigljavsky A. A. Analysis of Time Series Structure - SSA and Related Techniques // Monographs on statistics and applied probability. — 2001.