

**Ежов Федор Валерьевич**

**Выпускная квалификационная работа**

**ИСПОЛЬЗОВАНИЕ МЕТОДА SSA В МАШИННОМ ОБУЧЕНИИ ДЛЯ  
ПРОГНОЗА ВРЕМЕННЫХ РЯДОВ**

Уровень образования: магистратура

Направление 01.04.02 «Прикладная математика и информатика»

Основная образовательная программа ВМ.5751.2020 «Математическое моделирование,  
программирование и искусственный интеллект»

Научный руководитель:

Доцент, кафедра статистического

моделирования

к.ф.-м.н., доцент<sup>a</sup> Н. Э. Голяндина

Рецензент:

Программист, Майкрософт Израиль

А. Ю. Шлемов

---

<sup>a</sup> Где рецензент на обоих листах?

Saint Petersburg State University  
Applied Mathematics and Computer Science  
Statistical Modelling

**EZHOV Fedor Valerievich**

**Graduation Project**

**ON USING THE SSA METHOD IN MACHINE LEARNING TO PREDICT  
TIME SERIES**

Scientific Supervisor:  
Associate Professor, Department of  
Statistical Modelling N. E. Golyandina

Reviewer:  
Software developer, Microsoft R&D, Israel  
A. Y. Shlemov

Saint Petersburg  
2022

# Оглавление

<b>Введение . . . . .</b>	<b>5</b>
<b>Глава 1. Алгоритм Singular Spectrum Analysis . . . . .</b>	<b>6</b>
1.1. Этап 1. Построение траекторной матрицы (Вложение) . . . . .	6
1.2. Этап 2. Singular Value Decomposition (SVD) . . . . .	7
1.3. Этап 3. Группировка первых $r$ собственных троек . . . . .	7
1.4. Этап 4. Диагональное усреднение . . . . .	7
<b>Глава 2. Архитектуры нейронных сетей . . . . .</b>	<b>9</b>
2.1. Artificial Neural Network (ANN) . . . . .	9
2.2. Recurrent neural network (RNN) . . . . .	10
2.3. Long short-term memory (LSTM) . . . . .	10
2.4. Gated Recurrent Unit (GRU) . . . . .	11
<b>Глава 3. Использование SSA в машинном обучении . . . . .</b>	<b>12</b>
3.1. Задача . . . . .	12
3.2. Подготовка данных . . . . .	12
3.3. Обучение нейронной сети . . . . .	14
3.4. Прогнозирование . . . . .	15
3.4.1. Прогнозирование с помощью нейронных сетей . . . . .	15
3.4.2. Прогнозирование SSA . . . . .	15
3.5. Метрики качества прогноза . . . . .	15
3.6. Методика применения и сравнения методов . . . . .	15
<b>Глава 4. Модельные данные . . . . .</b>	<b>19</b>
4.1. Сумма двух синусов с белым шумом . . . . .	19
4.1.1. Влияние выбора параметра $r$ . . . . .	20
4.1.2. Влияние выбора параметра $r$ для ряда с небольшим шумом . .	28
4.1.3. Влияние длины ряда . . . . .	36
4.2. Сумма двух синусов с красным шумом. Ряд с трудно отделимым сигналом	44
4.2.1. Сравнение прогнозов, полученных с помощью метода SSA, обыч- ных и гибридных методов . . . . .	45

4.3. Суммарные результаты по модельный рядам . . . . .	62
<b>Глава 5. Реальные данные . . . . .</b>	<b>64</b>
5.1. Среднемесячные осадки в Индии . . . . .	64
5.1.1. Сравнение прогнозов, полученных с помощью метода SSA, обычных и гибридных методов . . . . .	65
5.2. Earth Orientation Parameters (EOP) . . . . .	71
5.2.1. Сравнение прогнозов, полученных с помощью метода SSA, обычных и гибридных методов . . . . .	73
5.3. Погода . . . . .	81
5.3.1. Сравнение прогнозов, полученных с помощью метода SSA, обычных и гибридных методов . . . . .	82
<b>Заключение . . . . .</b>	<b>87</b>
<b>Список литературы . . . . .</b>	<b>89</b>
<b>Приложение А. Модельные данные . . . . .</b>	<b>90</b>
A.1. Сумма синусов с белым шумом . . . . .	90
A.1.1. Влияние $r$ . . . . .	90
A.1.2. Влияние $r$ при маленьком шуме . . . . .	98
A.1.3. Влияние длины . . . . .	108
A.2. Красный шум . . . . .	109
<b>Приложение Б. Реальные данные . . . . .</b>	<b>131</b>
B.1. Среднемесячные осадки в Индии . . . . .	131
B.2. Погода . . . . .	132

## Введение

<sup>1</sup> 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.

---

<sup>1</sup> TODO

## 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 «обратного распространения ошибки». В данной работе нейронные сети «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 ряд  $Z$ , 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 гиперпараметров. 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 обучении 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

### 3.4.1. Прогнозирование с помощью нейронных сетей

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.

### 3.4.2. Прогнозирование SSA

## 3.5. Метрики качества прогноза

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  $L$  и  $r$  зададим сетку, по которой будем перебирать комбинации параметров, оценивая ошибку прогноза на один шаг вперед с помощью метрики RMSE на валидационной выборке. Построим график с

кривыми зависимости ошибки от параметра  $L$ . Каждая кривая будет соответствовать одному параметру  $r$  из соответствующей сетки.

Далее для пяти лучших пар параметров построим прогноз на тестовой выборке и получим оценку ошибки. Каждую ошибку отобразим на графике с помощью горизонтальной прямой вместе с прогнозами обычных и гибридных методов.

**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.
2. Перебрать параметры для метода SSA и выбрать лучшие параметры, основываясь на точности прогноза методом, полученной на предыдущем этапе.

**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.

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

Заметим, что для модельных данных, где сигнал известен, можно также использовать данные сигнала для анализа ошибок методов.

**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 инициализируются 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 прогнозов** Чтобы объяснить успешность или неуспешность прогноза и в целом посмотреть на картину прогнозирования, будем строить график прогноза на один шаг на тестовых данных 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 (его тренировочной части).

**Таблица с результатами** Для каждого модельного примера вместе с графиками сравнения будет приводиться таблица с числовыми результатами, которые строятся на основе таблицы с суммарными результатами 4.5. Таблица 4.5 объединяет результаты, полученные для всех модельных рядов из главы 4. Опишем сначала общую таблицу. В столбце «ssa-params» отображены параметры SSA в гибридных моделях, если параметры не указаны, значит в строчке записаны результаты для негибридных моделей. Префикс «b» значит, что значение в таблице отображает ошибку лучшей нейронной сети (например, лучшая из ANN, или лучшая из гибридной SSA-ANN с указанными параметрами SSA). Префикс «m» означает, что значение отображает среднее значение ошибок для соответствующих негибридных, обычных методов или метода SSA. При этом для методов с использованием нейронных сетей при вычислении общего среднего по методу откидывались наименьшее и наибольшее среднее по  $T$  в таблице 3.1 из средних при разных  $h$ . Для метода SSA показана наименьшее значение ошибки для префикса «b», для префикса «m» показано среднее значение по пяти лучшим прогнозом, при этом для удобства сравнения во всех строчках ошибки SSA-прогноза повторены.<sup>1</sup>

В отдельных таблицах для каждого примера префикс «b» значит, что значение

---

<sup>1</sup> А для реальных рядов такая таблица и подтаблицы будут?

в таблице отображает ошибку лучшей нейронной сети из своих аналогов (например, лучшая из негибридных, или лучшая из гибридных с указанными параметрами SSA). Префикс «m» означает, что значение отображает среднее значение ошибок для негибридных, обычных методов или метода SSA, где среднее считается по значениям из таблицы 4.5. Для метода SSA значения повторены из таблицы 4.5. <sup>2</sup>

---

<sup>2</sup> Проверьте, правильно ли я переписала этот кусочек.

## Глава 4

### Модельные данные

В данном разделе на модельном ряде проведены эксперименты с целью показать влияние различных особенностей временных рядов и способов выбора параметров методов на результаты точности прогнозов и на результаты сравнения методов прогноза. В разделе рассмотрено:

1. Влияние выбора параметра  $r$ .
2. Влияние величины шума во временном ряде.
3. Влияние красного шума во временном ряде.

#### 4.1. Сумма двух синусов с белым шумом

Рассмотрим следующий ряд, 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. Так как сигнал ряда состоит из двух синусов, его ранг равен 4. Обозначим  $X_N$  ряд с  $\kappa = 0.3$ , а  $Z_N$  с  $\kappa = 1.5$ . Далее в этом разделе рассматриваем ряды:  $X_{650}$ ,  $Z_{650}$ ,  $Z_{1500}$ . На рис. 4.1 представлен ряд  $Z_{650}$ .

В экспериментах будем разбивать ряды длиной 650 на тренировочную, валидационную, тестовую выборки по 350, 150, 150 точек соответственно. А ряды длиной 1500 — по 750, 500, 250 точек соответственно.

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. В данном примере сигнал 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  $L = 175$  (половина длины тренировочной выборки) и  $r = 4$ , так как ранг ряда равен 4, а  $L = 175$  удовлетворяет условию асимптотической разделимости.

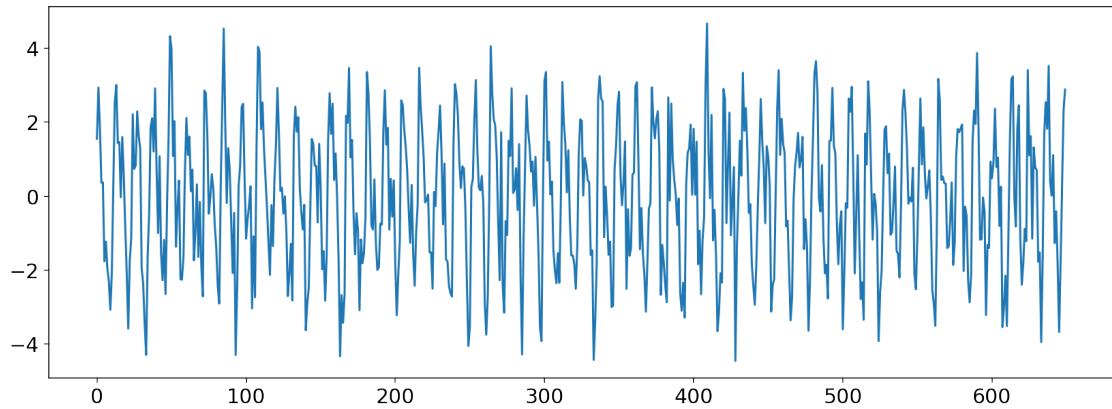


Рис. 4.1. Ряд «сумма синусов с белым шумом»,  $Z_{650}$ .

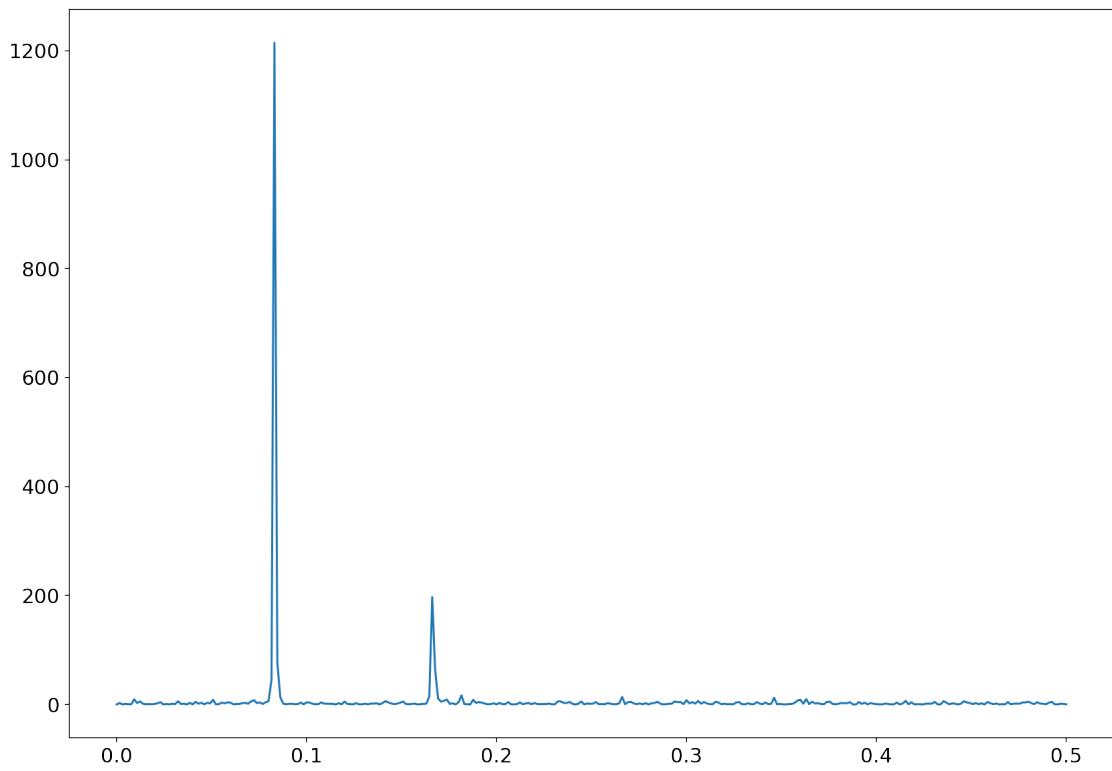


Рис. 4.2. Периодограмма ряда «сумма синусов с белым шумом»,  $Z_{650}$ .

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

Поставим задачу сравнить обычные и гибридные методы при разных выбранных значениях параметра  $r$  в гибридных методах. Будем проводить сравнение на ряде  $Z_{650}$ . Далее будем работать с тремя различными параметрами: правильно выбранное  $r = 4$ , недостаточно большое  $r = 2$ , слишком большое  $r = 6$ . Каждый вариант будем исследовать с помощью методологии, описанной в разделе 3.6.

В тексте работы будем приводить только часть рисунков, а полный анализ можно найти в приложении к работе.

## Прогноз по SSA

Сравним точность прогнозирования методом SSA при разных параметрах  $r$  на валидационном периоде. Зададим следующую сетку параметров  $L = \{12, 24, \dots, 175\}$ ,  $r = \{2, 4, 6\}$ . Посмотрим на результаты на рисунке 4.3. На графике видно, что наилучшие результаты достигаются при  $r = 4$ , худшие результаты достигаются при  $r = 6$ .

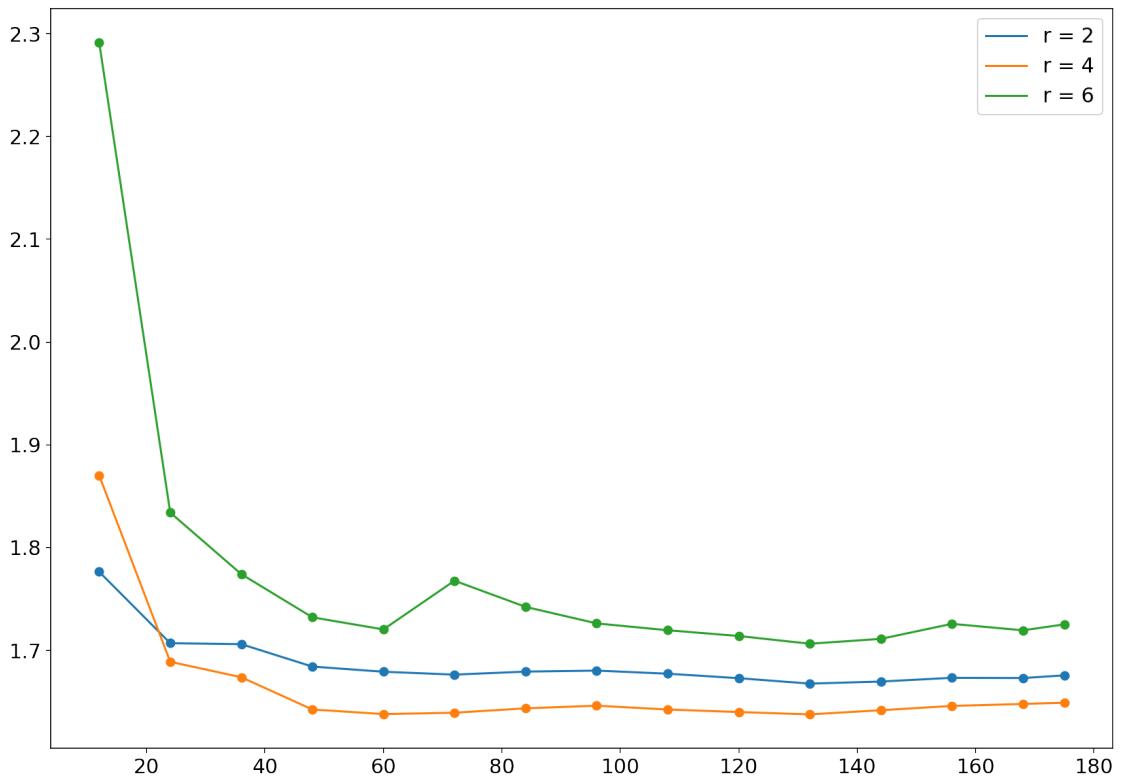


Рис. 4.3. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . RMSE прогноз на валидационной части.

Выделим лучшее  $L$  для каждого  $r$ . Исходя из графика, для всех  $r$  можно положить  $L = 132$ . Также добавим  $L = 175$  к лучшим параметрам и будем рассматривать комбинации параметров  $L = \{132, 175\}$  и  $r = \{2, 4, 6\}$ .

## Восстановление SSA

Посмотрим на графиках 4.4–4.6, как метод SSA восстанавливает тренировочную выборку для выбранных пар параметров. На графике 4.4 видно, что при  $r = 2$  метод не восстанавливает ряд полностью, что естественно, так как не учтен синус с периодом 6.

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 разница между параметрами  $L = 132$ ,  $L = 175$  маленькая, далее будем рассматривать только параметры  $r = \{2, 4, 6\}$  и  $L = 175$ .

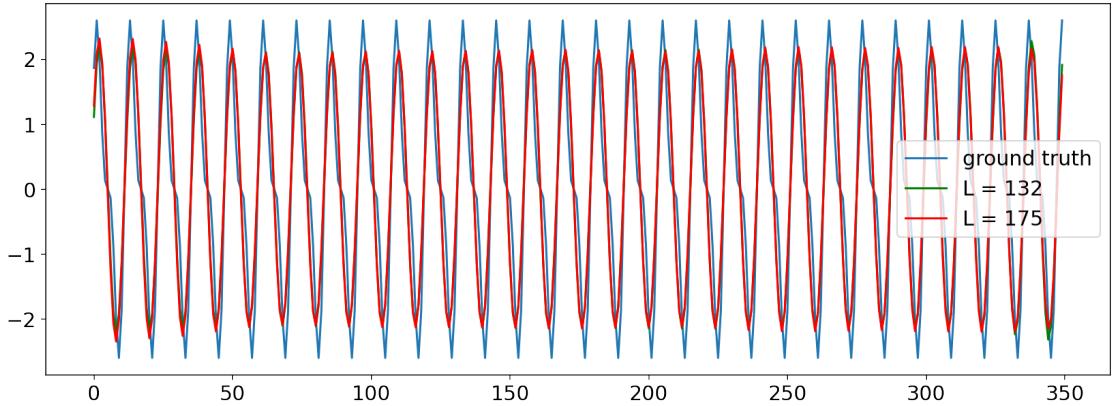


Рис. 4.4. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Восстановление тренировочной выборки с помощью метода SSA.  $r = 2$

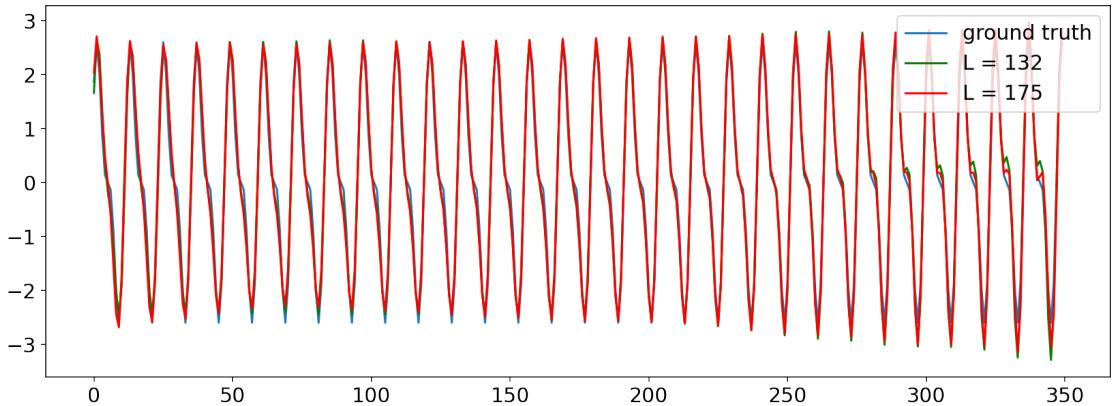


Рис. 4.5. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Восстановление тренировочной выборки с помощью метода SSA.  $r = 4$

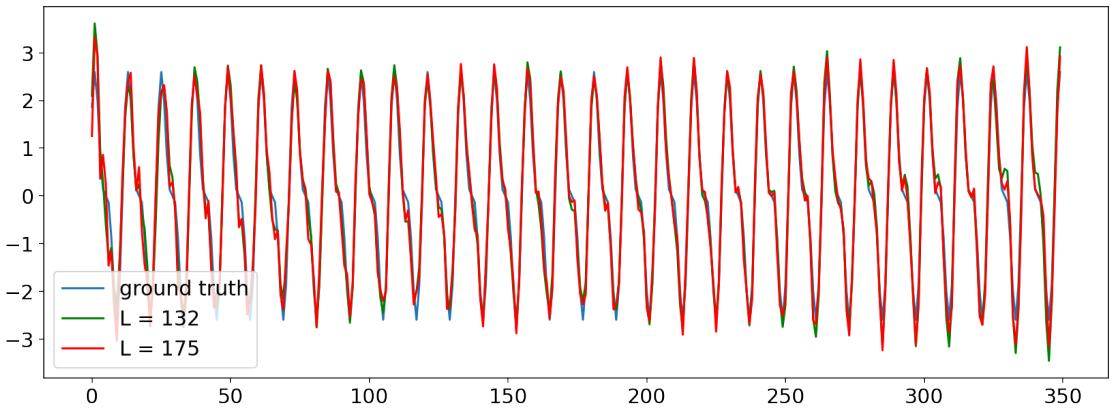


Рис. 4.6. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Восстановление тренировочной выборки с помощью метода SSA.  $r = 6$

### Сравнение обычных и гибридных методов

Для нейронных сетей зададим следующую сетку параметров:  $T = \{12, 24, \dots, 132\}$ ,  $h = \{10, 25, \dots, 100\}$ . Для метода SSA в гибридных моделях возьмем пары параметров, выбранные заранее.

На графиках 4.7—4.9 (больше графиков в приложении А.1.1), 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.

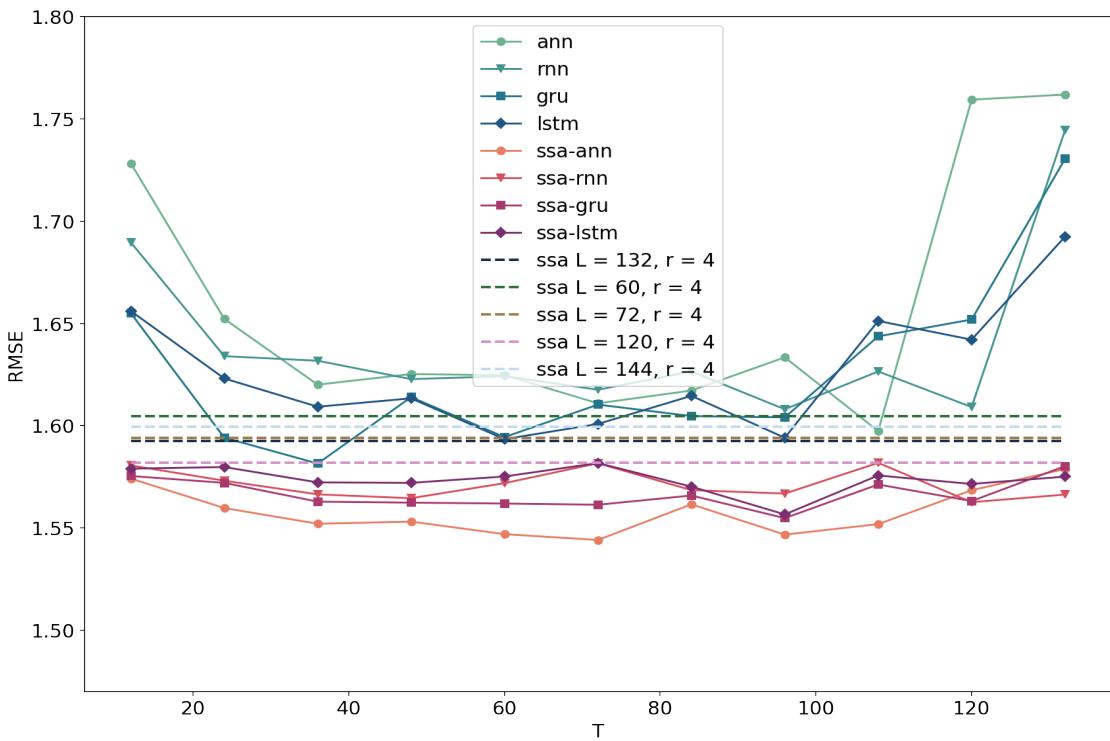


Рис. 4.7. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Ошибки прогноза в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 2$ .

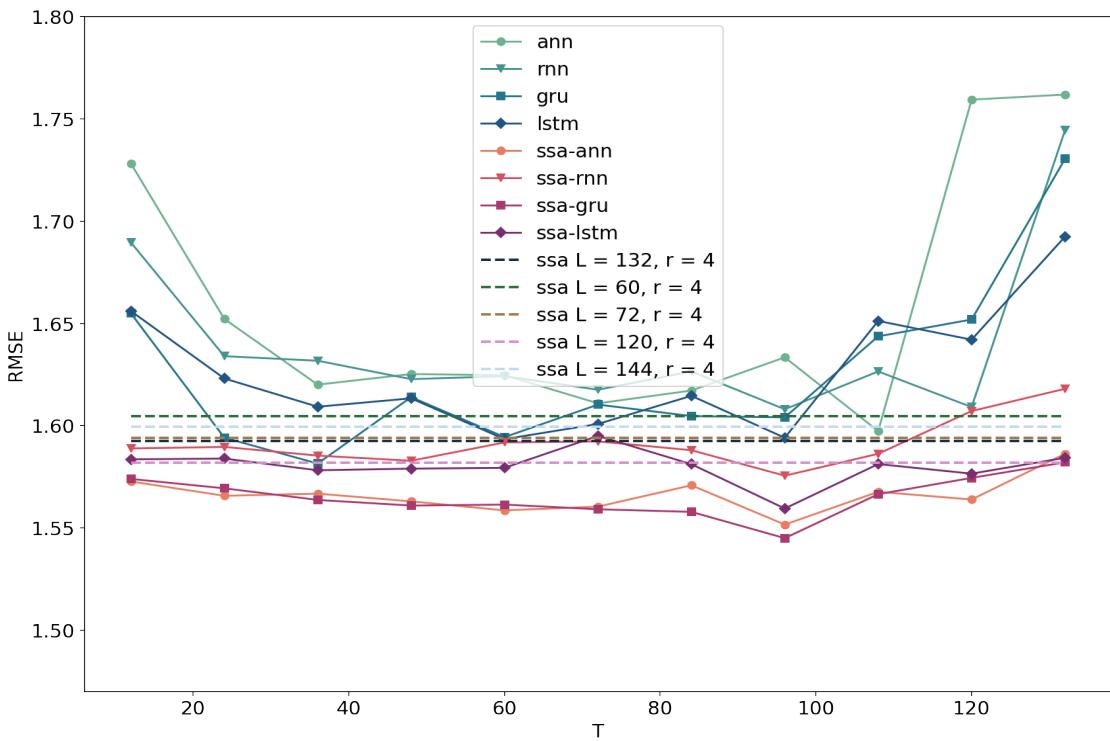


Рис. 4.8. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Ошибки прогноза в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 4$ .

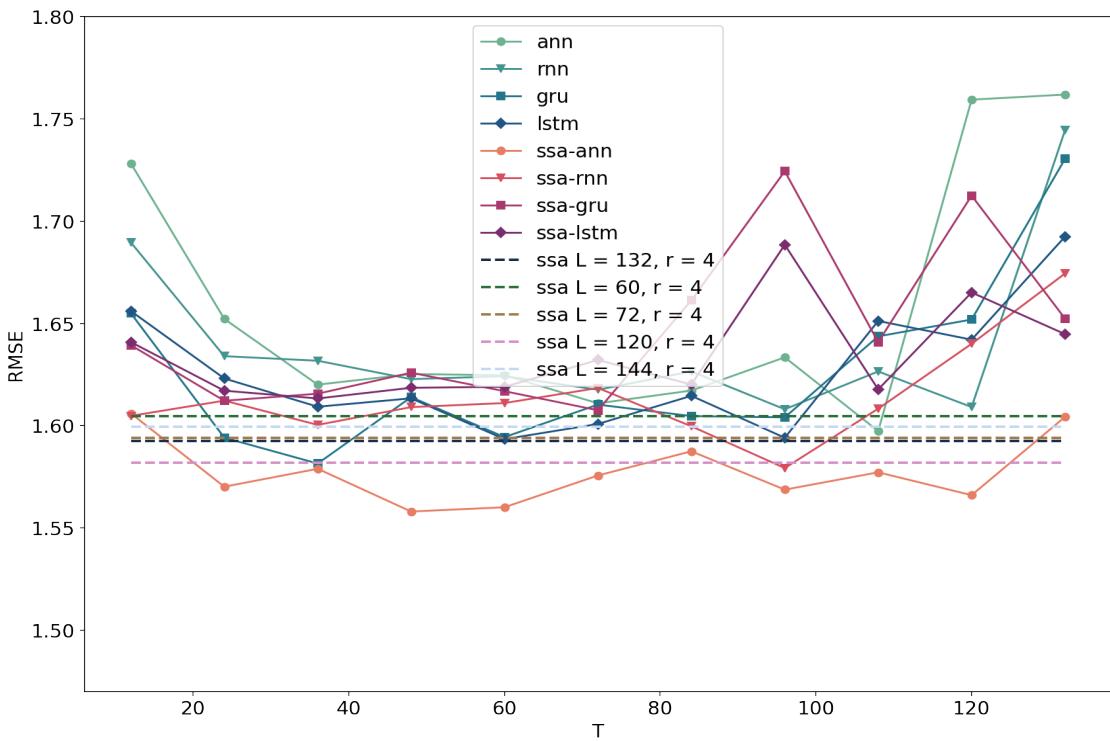


Рис. 4.9. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Ошибки прогноза в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 6$ .

### Отображение прогнозов

На графиках 4.10–4.12 (больше графиков в приложении А.1.1) представлены результаты прогнозирования методам ANN. 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, но в случае  $r = 2$  форма прогноза больше похожа на форму сигнала. Прогноз для  $r = 4$  выглядит более гладким.

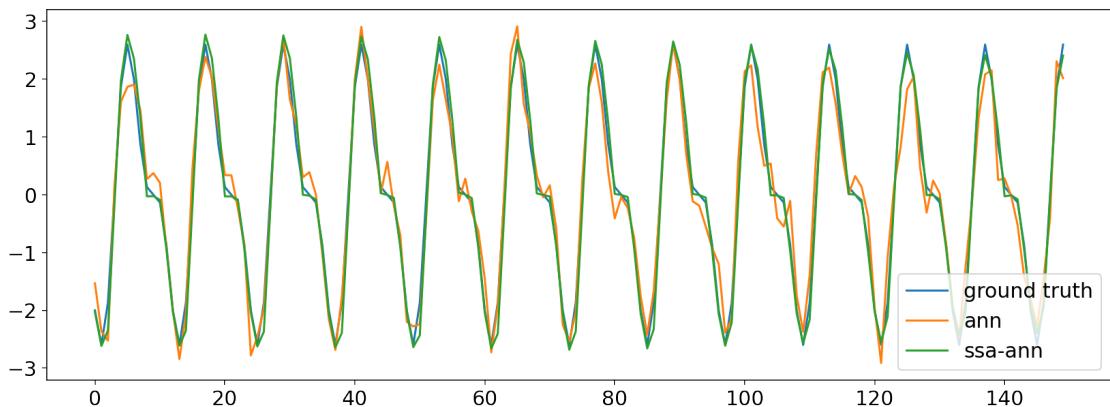


Рис. 4.10. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Прогноз для ANN и SSA-ANN.  $r = 2$

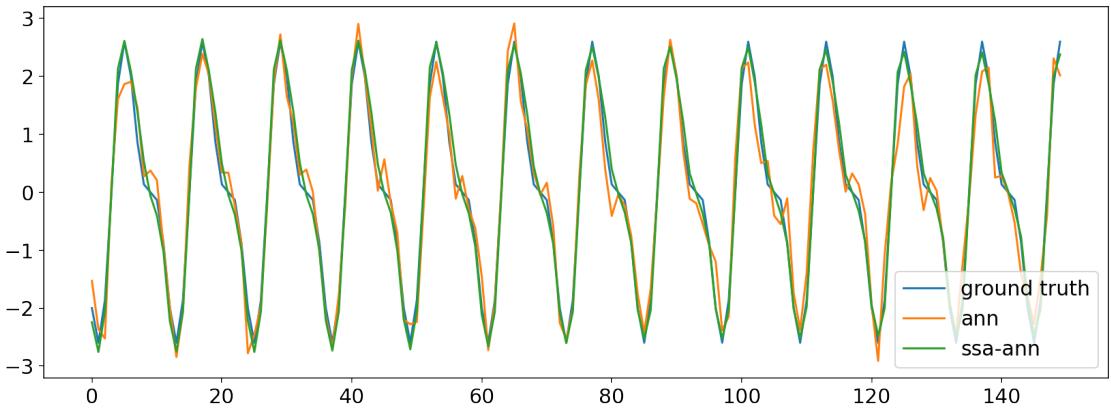


Рис. 4.11. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Прогноз для ANN и SSA-ANN.  $r = 4$

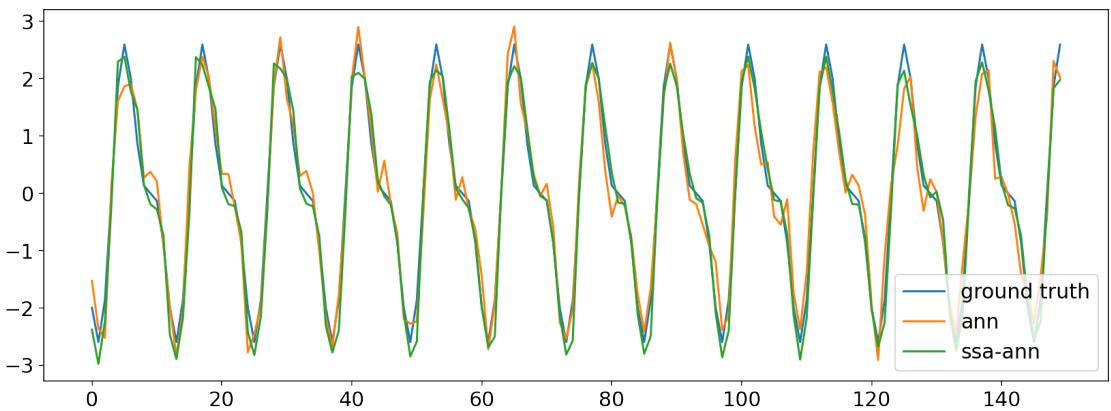


Рис. 4.12. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Прогноз для ANN и SSA-ANN.  $r = 6$

## Проверка устойчивости

Чтобы исключить случайность в полученных результатах, проведем сравнение для разных начальных весов методов. Зафиксируем уменьшенную сетку для параметра  $T = \{12, 84\}$ . Сетка для параметра  $h$  останется прежней. Будем получать каждый результат по 7 раз, инициализируя метод с новыми весами. Полученные результаты отображены на рисунках 4.13—4.14 (больше графиков в приложении А.1.1). 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. В таблице 4.1 продемонстрированы величины ошибок по метрике RMSE для методов.

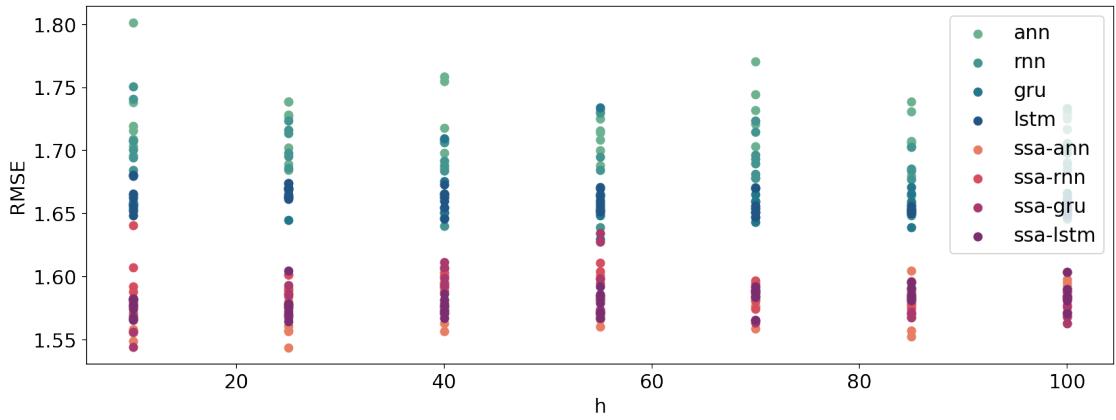


Рис. 4.13. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$r = 4, L = 175, T = 12.$

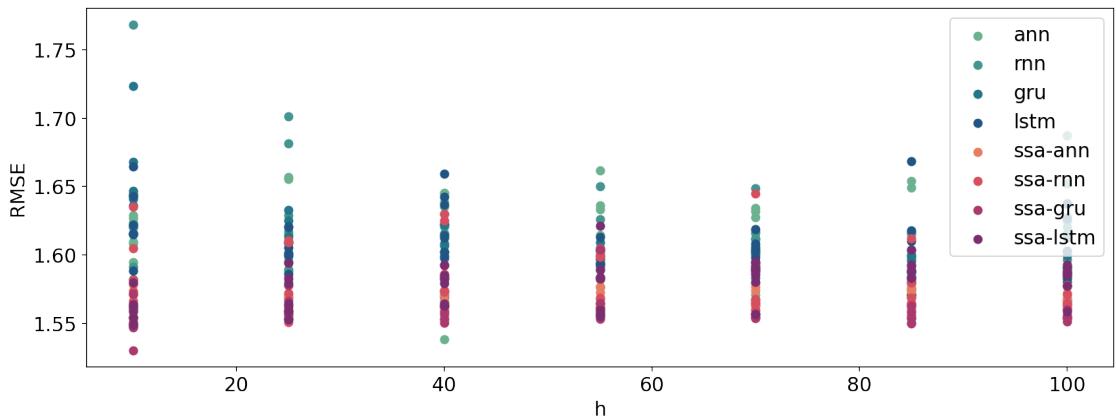


Рис. 4.14. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$r = 4, L = 175, T = 84.$

Из значений в таблицы 4.1 и полученных ранее результатов можем сделать выводы, что для ряда с несложным выделяемым сигналом выбор аналитически верных параметров приводит к одному из лучших результатов. Что удивительно, выбор ранга  $r = 2$  меньше ранга сигнала привел к улучшению, хоть и небольшому, точности прогноза. При этом остаётся непонятным, каким образом нейронные сети делают прогноз с хорошей точностью при том, что на вход подаётся неправильное восстановление сигнала в результате предобработки SSA.

Гибридные методы показали наилучшую точность. Наибольшую ошибку показал метод SSA. Как и говорилось, наименьшая ошибка достигается для  $r = 2$  в гибридных моделях. Средняя ошибка для  $r = 2$  и  $r = 4$  не сильно отличается. Для  $r = 6$  средняя ошибка гибридных методов выше.

Таблица 4.5 показывает, что из негибридных методов немного лучше работает LSTM, однако в паре с SSA лучший результат у ANN.

1

Таблица 4.1. Усредненные и лучшие результаты прогнозов по RMSE.

ssa-params	b-nn	m-nn	b-ssa	m-ssa
-	1.547	1.635	1.581	1.586
r2	1.528	1.567	1.581	1.586
r4	1.532	1.575	1.581	1.586
r6	1.533	1.613	1.581	1.586

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

Поставим задачу сравнить обычные и гибридные методы в случае разных выбранных параметрах  $r$  в гибридных методах на временном ряде с небольшим шумом. Эксперимент аналогичен описанному в разделе 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  $L = \{12, 24, \dots, 175\}$ ,  $r = \{2, 4, 6\}$ . Посмотрим на результаты на рисунке 4.15. На графике видно, что наилучшие результаты достигаются при  $r = 4$ , худшие результаты достигаются при  $r = 2$ , что отличает этот пример от предыдущего с большим шумом.

<sup>1</sup> Везде нужно исправить надписи к таблицам, причем в том же стиле - с названием ряда и пр.

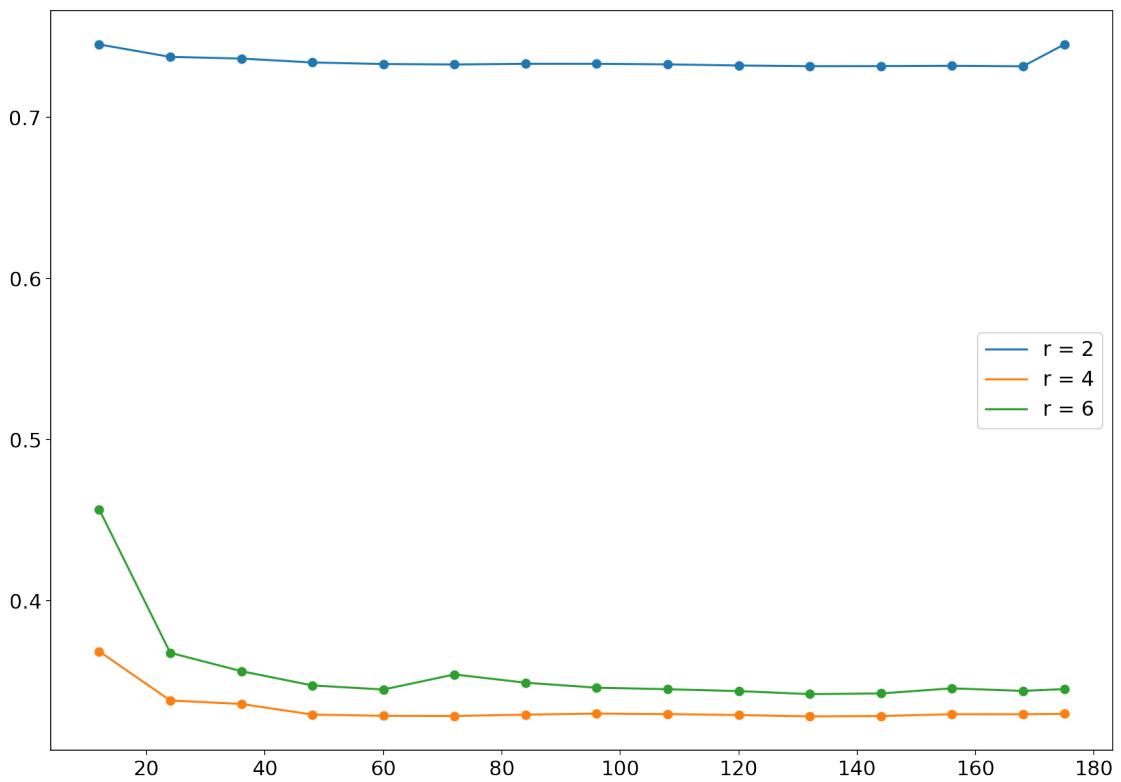


Рис. 4.15. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . RMSE прогноз на валидационной части.

Выделим лучшее  $L$  для каждого  $r$ . Исходя из графика для  $r = \{4, 6\}$  это будет  $L = 132$ , для  $r = 2$  будет  $L = 168$ . Также добавим  $L = 175$  к лучшим параметрам и будем рассматривать комбинации параметров  $L = \{132, 168, 175\}$  и  $r = \{2, 4, 6\}$ .

## Восстановление SSA

Посмотрим, как метод SSA восстанавливает тренировочную выборку для выбранных пар на графиках ниже. На графике 4.16 видно, что метод не восстанавливает ряд полностью. Результат очень похожи на те, что в главе 4.1.1. На графиках 4.17, 4.18, видно, что восстановление очень хорошее. На графике 4.18 можно заметить влияние шума, видно, как в оценках сигнала дрожат пики. На графиках видно, что разницы между параметрами  $L$  нет, далее будем рассматривать только параметры  $r = \{2, 4, 6\}$  и  $L = 175$ .

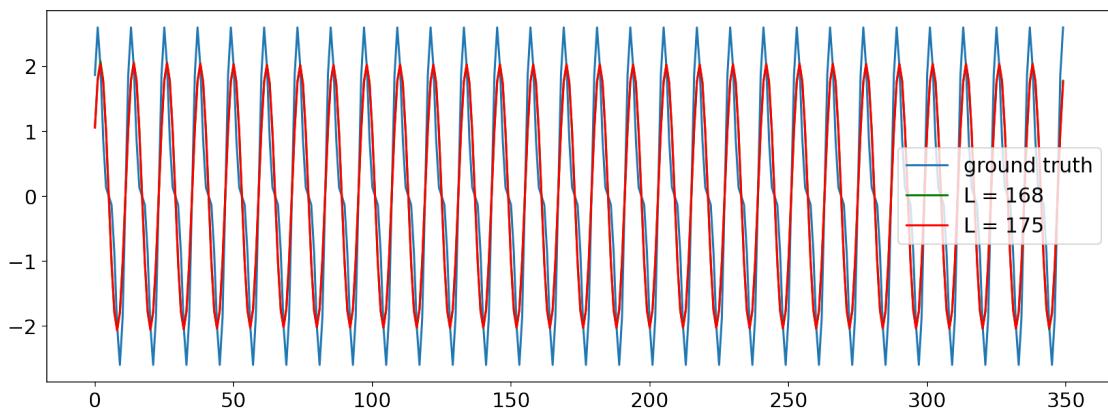


Рис. 4.16. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Восстановление тренировочной выборки с помощью метода SSA.  $r = 2$

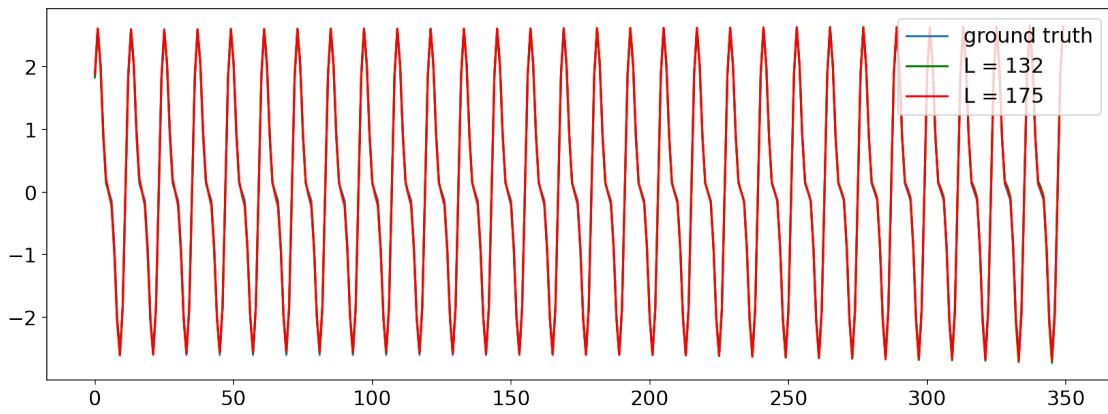


Рис. 4.17. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Восстановление тренировочной выборки с помощью метода SSA.  $r = 4$

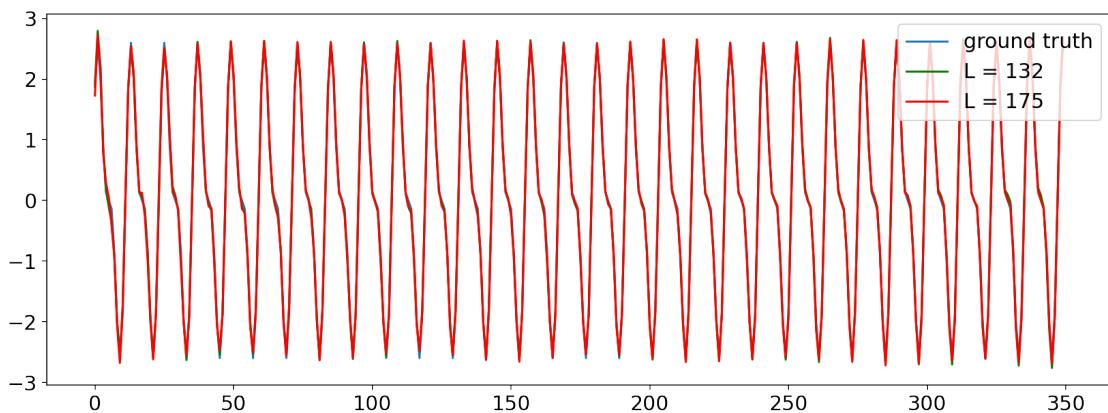


Рис. 4.18. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Восстановление тренировочной выборки с помощью метода SSA.  $r = 6$

## Сравнение обычных и гибридных методов

Пусть задана следующая сетка параметров:  $T = \{12, 24, \dots, 132\}$ ,  $h = \{10, 25, \dots, 100\}$ .

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

На графиках 4.19—4.21 (больше графиков в приложении А.1.2) представлены результаты сравнения по сетке параметров, заданной выше. Можно заметить, что гибридные методы явно лучше, чем обычные. В целом, результаты аналогичны результатам в разделе 4.1.1.

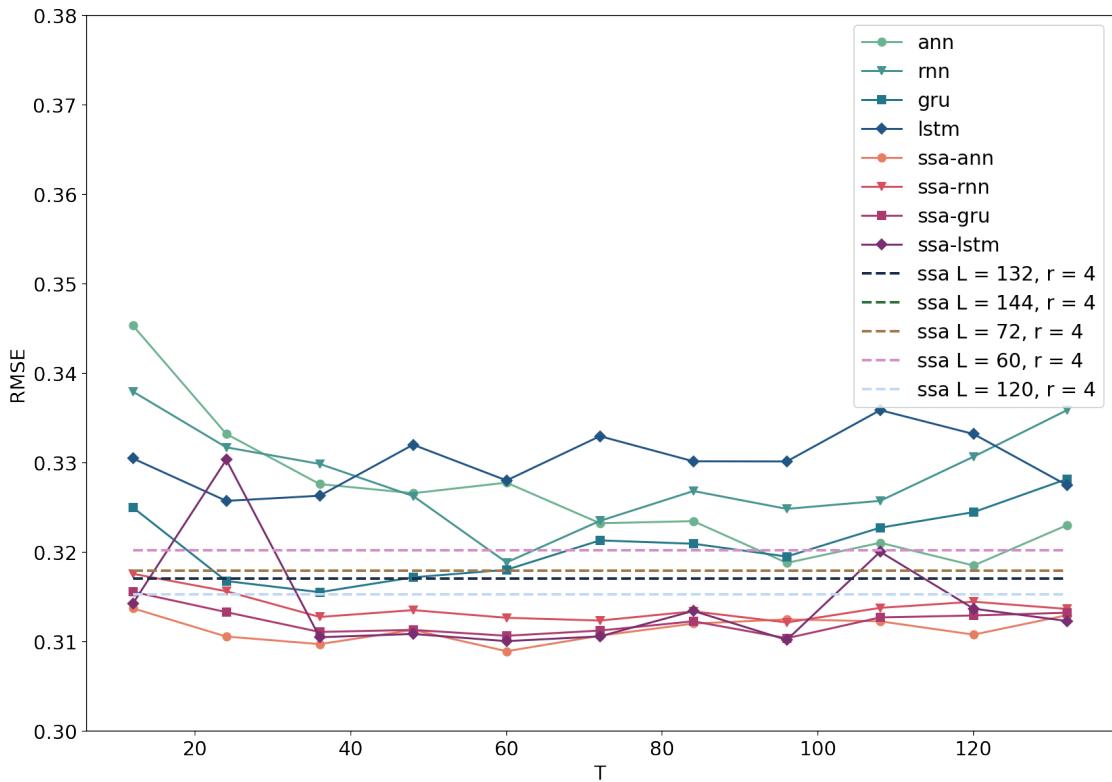


Рис. 4.19. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Ошибки прогноза в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 2$ .

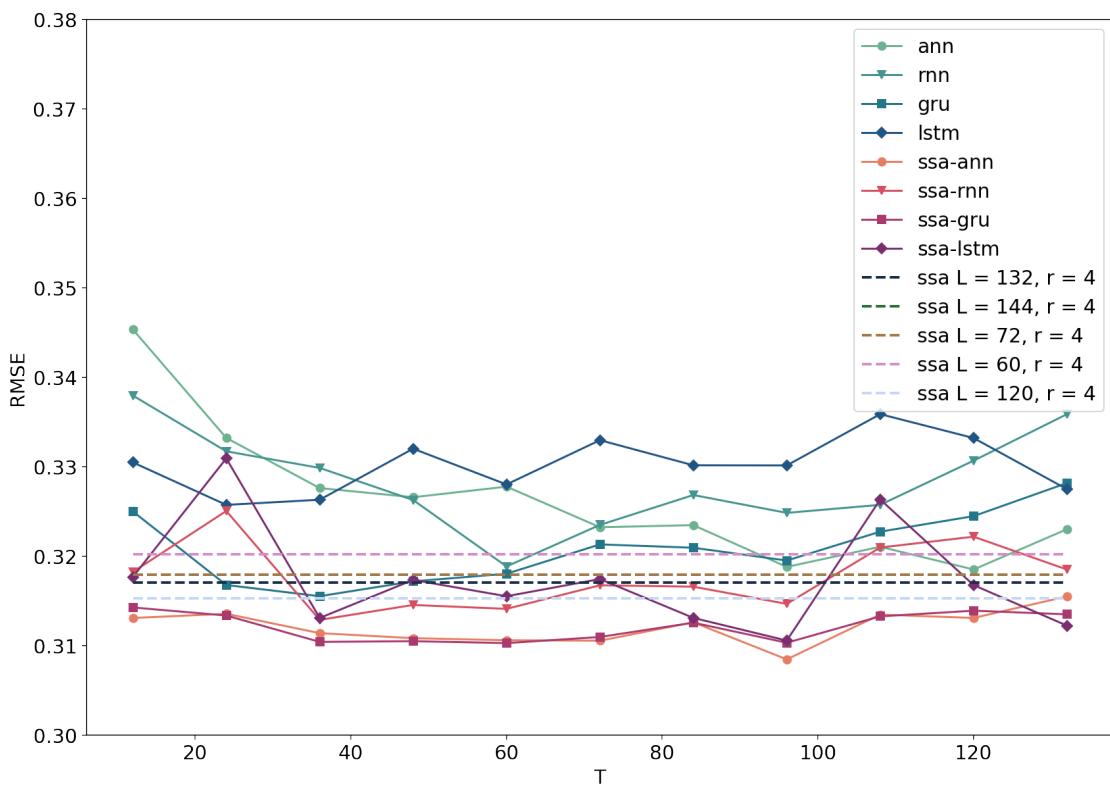


Рис. 4.20. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Ошибки прогноза в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 4$ .

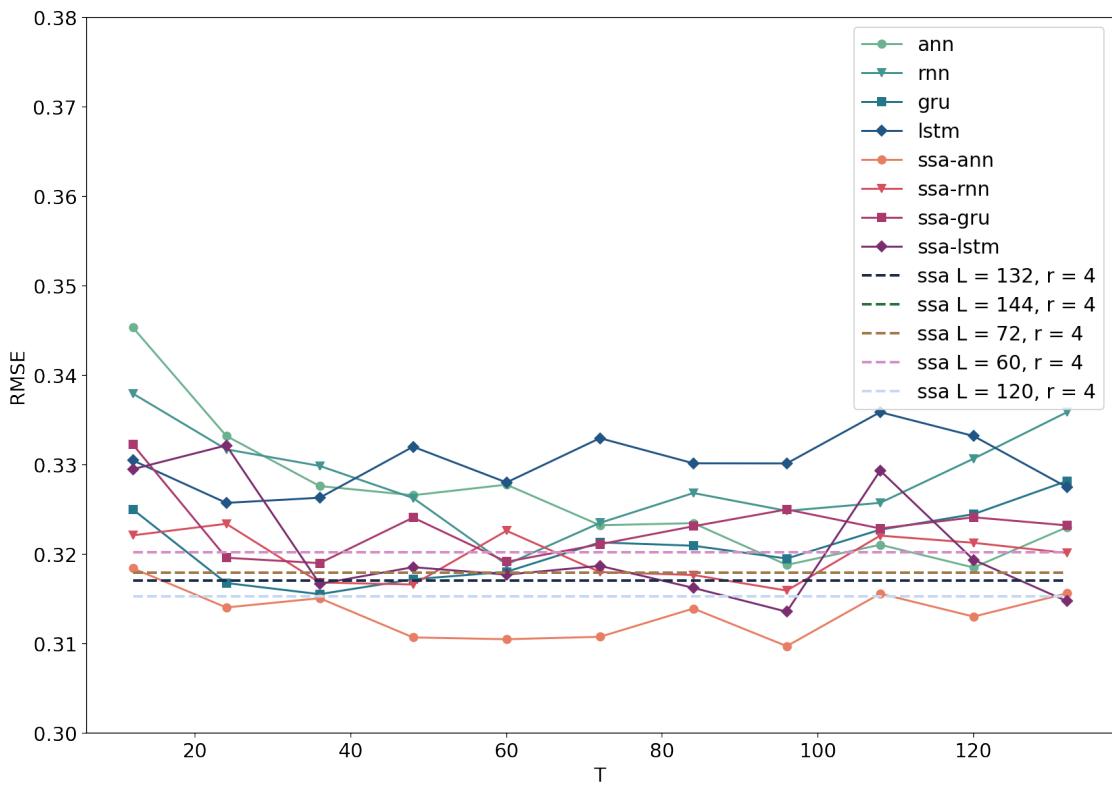


Рис. 4.21. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Ошибки прогноза в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 6$ .

### Отображение прогнозов

На графиках 4.22–4.24 (больше графиков в приложении А.1.2) представлены результаты прогнозирования методами. Из-за маленького шума все прогнозы похожи друг на друга.

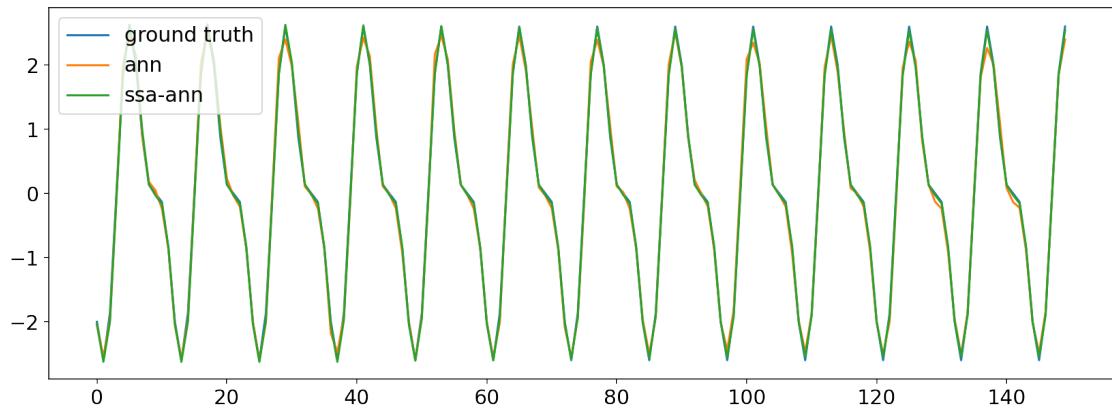


Рис. 4.22. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Прогноз результатов для ANN и SSA-ANN.  $r = 2$

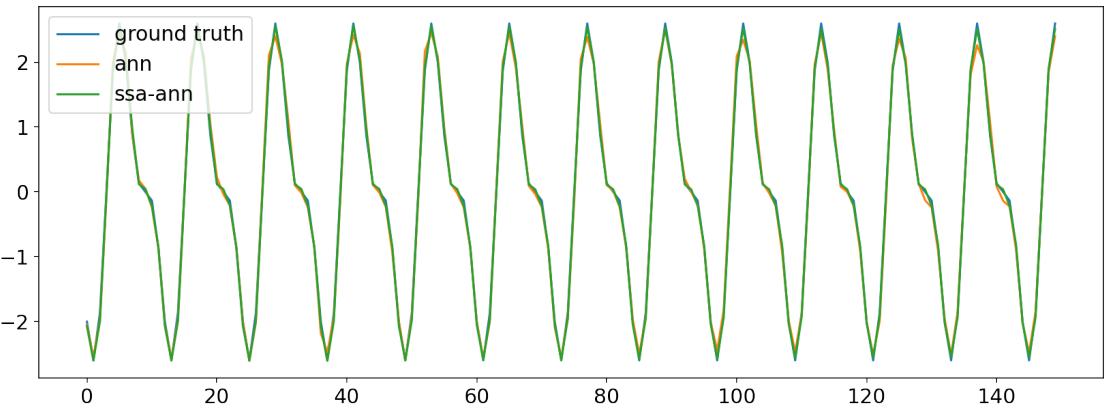


Рис. 4.23. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Прогноз результатов для ANN и SSA-ANN.  $r = 4$

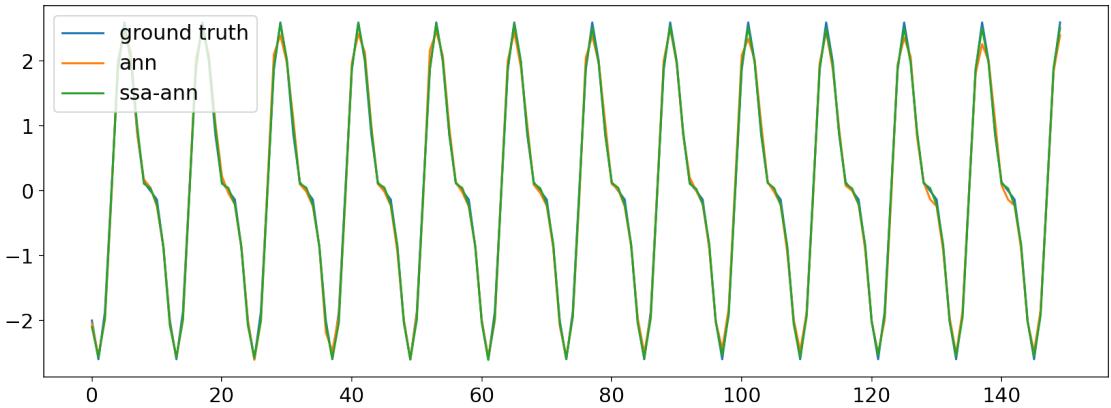


Рис. 4.24. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Прогноз результатов для ANN и SSA-ANN.  $r = 6$

### 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 параметра  $h$  останется прежней. Будем получать каждый результат по 7 раз, инициализируя метод с новыми весами. Полученные результаты отображены на рисунках 4.25—4.26 (больше графиков в приложении А.1.2). На них подтверждается, выводы сделанные ранее. Заключаем, что полученные результаты устойчивые.

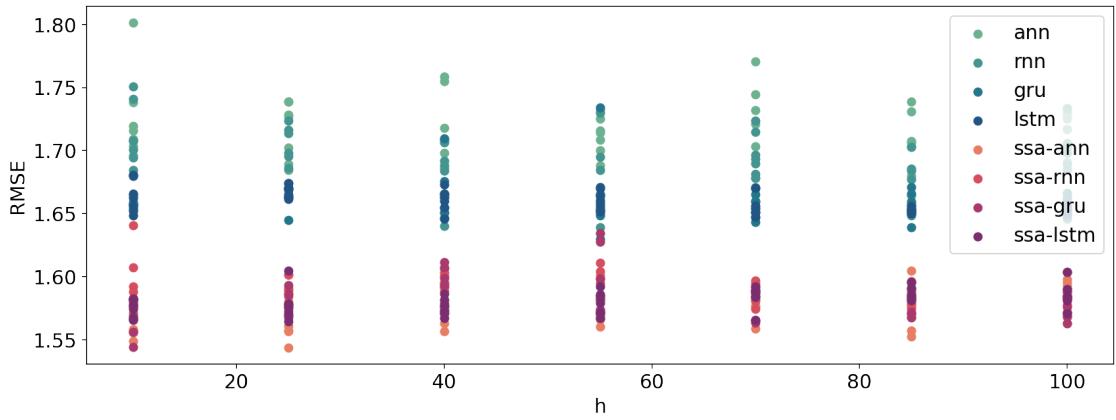


Рис. 4.25. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Проверка устойчивости.

$r = 4, L = 175, T = 12.$

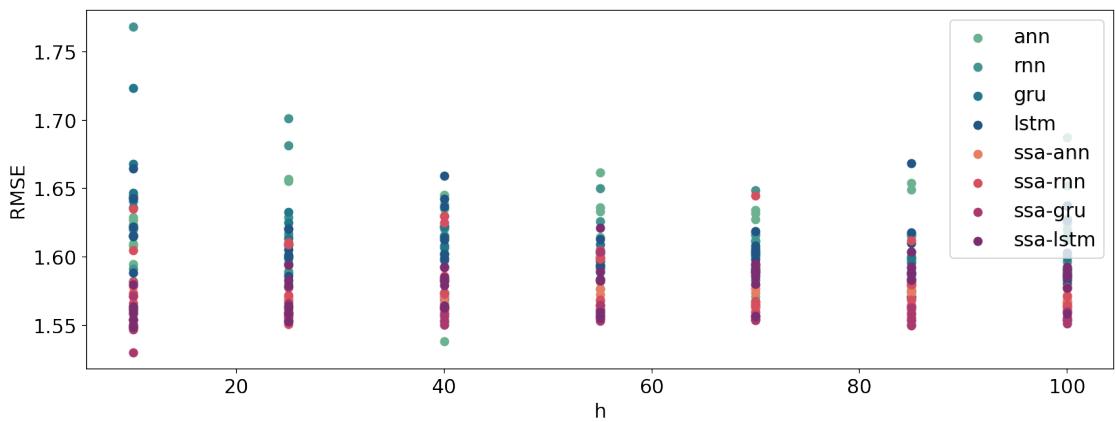


Рис. 4.26. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Проверка устойчивости.

$r = 4, L = 175, T = 84.$

## Выводы

Таблица 4.2. Усредненные лучшие результаты прогнозов методов.

	b-nn	m-nn	b-ssa	m-ssa
<b>ssa-params</b>				
-	0.308	0.325	0.315	0.316
r2	0.307	0.313	0.315	0.316
r4	0.305	0.314	0.315	0.316
r6	0.304	0.316	0.315	0.316

На ряде  $X_{650}$  было продемонстрировано сравнение обычных и гибридных методов, и метода SSA. Из полученных результатов и таблицы 4.2 (таблица является аналогичной таблице 4.1) можем сделать выводы, что полученные результаты аналогичны результатам, полученным в разделе 4.1.1. Только теперь наилучшие результаты достигаются для  $r = 6$ . Но из таблицы видно, что разница между параметрами  $r$  не велика. Негибридные методы проигрывают гибридным в точности, но не сильно. Если сравнивать средние значения, то видно, что ошибка прогноза по SSA примерно равна ошибке гибридных методов и они вместе выигрывают у негибридных нейронных сетей.

Таблица 4.5 показывает, что ... <sup>2</sup>

#### 4.1.3. Влияние длины ряда

Поставим задачу сравнить обычные и гибридные методы на временных рядах разной длины. Будем проводить сравнение на рядах  $Z_{650}$ ,  $Z_{1500}$ . В ходе эксперимента хотим выяснить, как длина ряда влияет на точность предсказаний.

#### Прогноз по SSA

Сравним точность прогнозирования методом SSA при разных параметрах  $r$ . Для ряда  $Z_{650}$  воспользуемся полученными результатами из раздела 4.1.1. Для ряда  $Z_{1500}$  зададим следующую сетку параметров  $L = \{12, 48, \dots, 375\}$ ,  $r = \{2, 4, 6\}$ . Посмотрим на результаты на рисунке 4.27. На графике видно, что наилучшие результаты достигаются при  $r = 4$ , худшие результаты достигаются при  $r = 2$ .  $r = 6$  дает неплохие результаты. Также заметим, что ошибка сильно упала по сравнению с рядом  $Z_{650}$  (рис. 4.3). Также для ряда  $Z_{650}$  наихудшие результаты показывал параметр  $r = 6$ , для  $Z_{1500}$  наоборот. Это объяснимо тем, что при росте длины ряда, влияние шума, становится меньше. Алгоритму SSA становится проще выделить сигнал.

---

<sup>2</sup> Наверно, нужно везде написать что-то про сравнение сетей между собой, примерно как сделала в пред примере. А в разделе с суммарной таблице можно написать, какая сеть чаще оказывалась победителем.

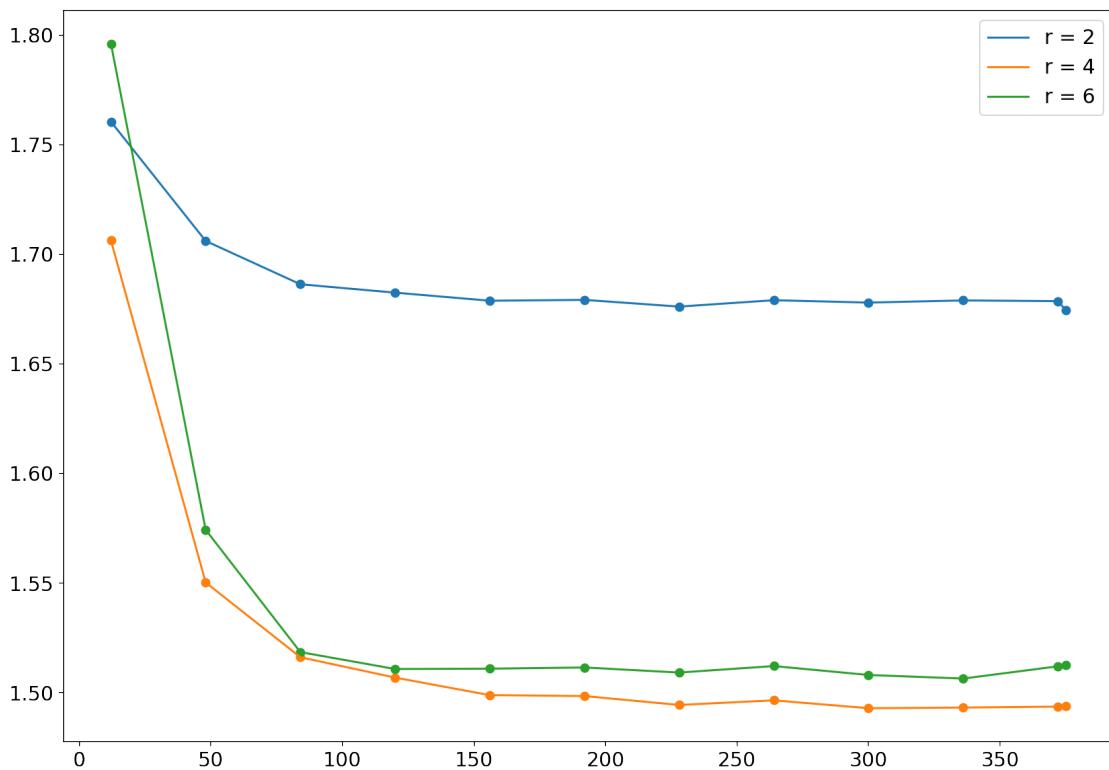


Рис. 4.27. «Сумма синусов с белым шумом». Ряд  $Z_{1500}$ . RMSE прогноз на валидационной части.

На основе полученных результатов зафиксируем наилучшие параметры SSA. Для ряда  $Z_{650}$  параметры  $L = 175$  и  $r = 4$ , для ряда  $Z_{1500}$  параметры  $L = 375$  и  $r = 4$ . Далее в гибридных моделях и прогнозе SSA будем использовать эти параметры.

### Восстановление SSA

Посмотрим на то, как метод SSA восстановил тренировочную выборку ряда  $Z_{1500}$  с выбранными ранее параметрами (рис. 4.28). На графике видно, что метод хорошо выделил сигнал, но в самом конце амплитуда начала уменьшаться. Результаты восстановления сигнала для  $Z_{650}$  были представлены ранее (рис. 4.5).

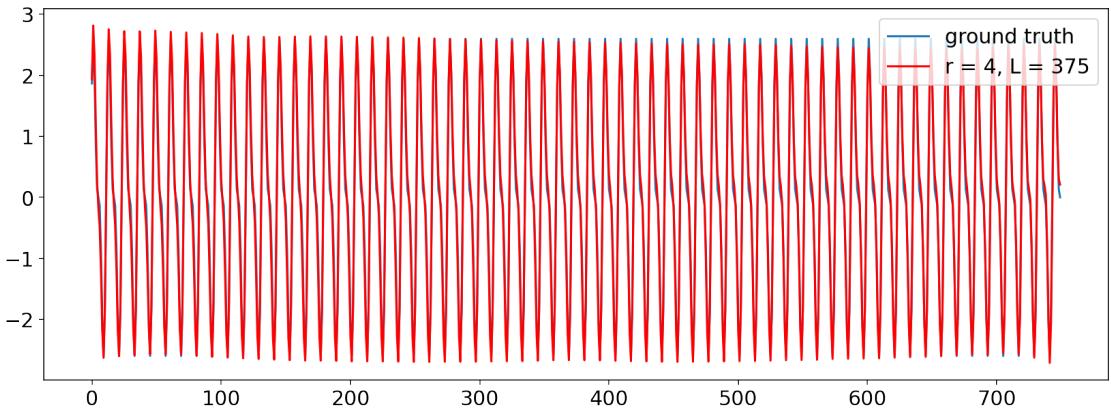


Рис. 4.28. «Сумма синусов с белым шумом». Ряд  $Z_{1500}$ . Восстановление тренировочной выборки с помощью метода SSA.  $r = 4$

### Сравнение обычных и гибридных методов

Для нейронных сетей зададим следующую сетку параметров:  $T = \{12, 48, \dots, 444\}$ ,  $h = \{10, 25, \dots, 100\}$ . Для метода SSA в гибридных моделях возьмем пары параметров, выбранные выше.

На графиках 4.29 и 4.30, представлены результаты сравнения для ряда  $Z_{1500}$ . Результаты для ряда  $Z_{650}$  были продемонстрированы ранее (рис. 4.8, А.5). На графиках ниже, мы можем наблюдать превосходство гибридных методов. Но такое же превосходство мы наблюдали на графиках 4.8, А.5 для ряда  $Z_{650}$  ранее. Также в видно что точность для ряда  $Z_{1500}$  намного больше, чем для  $Z_{650}$ . Это видно по двум пунктирным линиям соответствующим прогнозам SSA (линия  $L = 375$ ,  $r = 4$  соответствует такой же линии на рисунках 4.8, А.5).

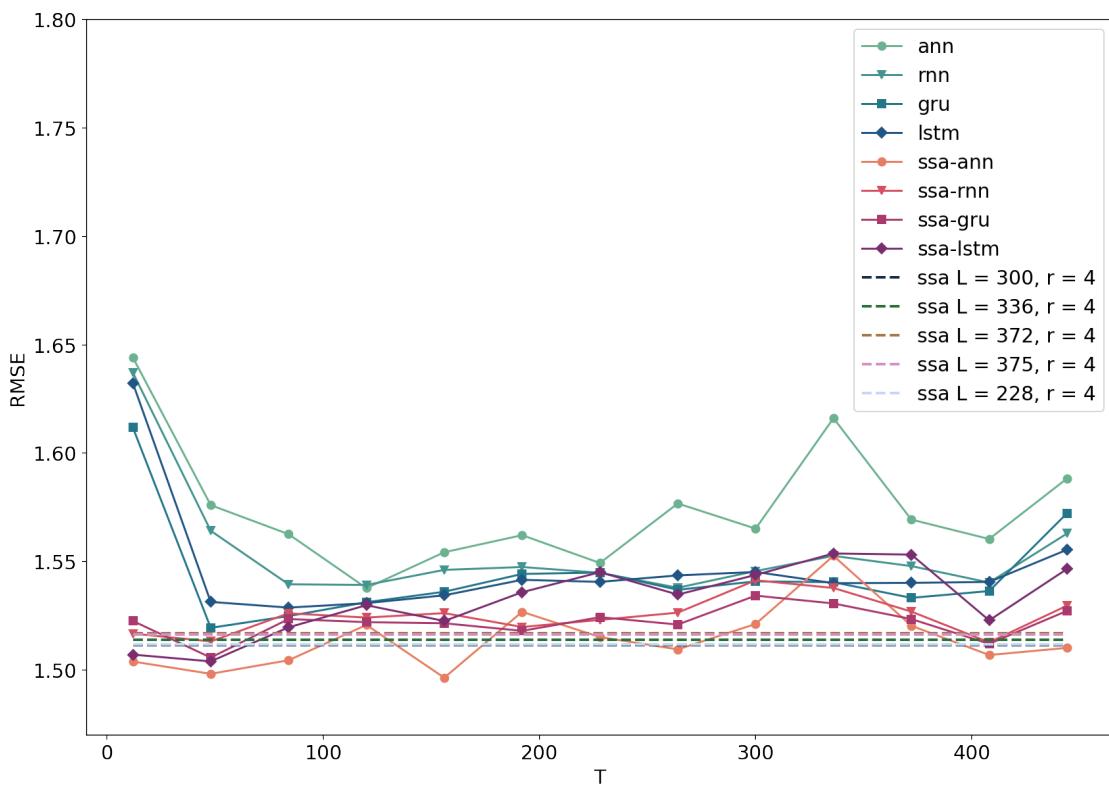


Рис. 4.29. «Сумма синусов с белым шумом». Ряд  $Z_{1500}$ . Ошибки прогноза в зависимости от параметра  $T$ .  $L = 375$ ,  $r = 4$ .

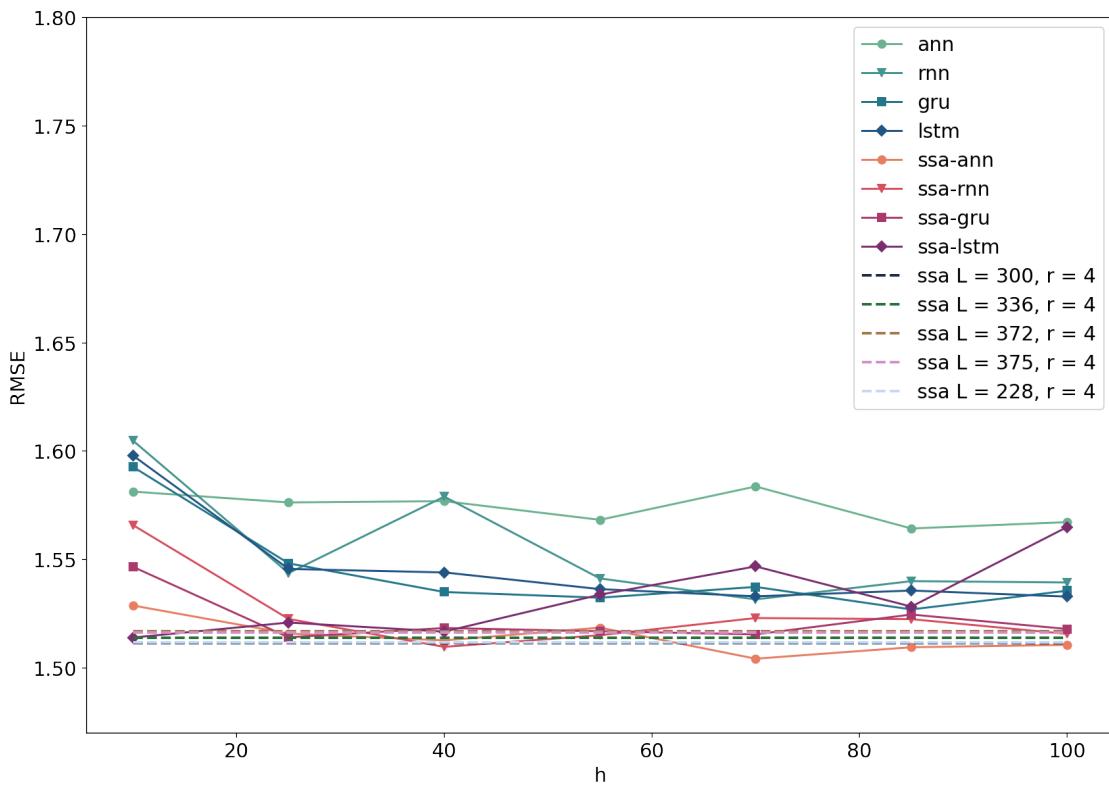


Рис. 4.30. «Сумма синусов с белым шумом». Ряд  $Z_{1500}$ . Ошибки прогноза в зависимости от параметра  $h$ .  $L = 375$ ,  $r = 4$ .

### Отображение прогнозов

Ниже на графиках представлены результаты прогнозирования методами для ряда  $Z_{1500}$ . На графиках видно, что предсказания для гибридных методов больше напоминают сигнал ряда. Видно, что в прогноз обычных методов проникает шум.

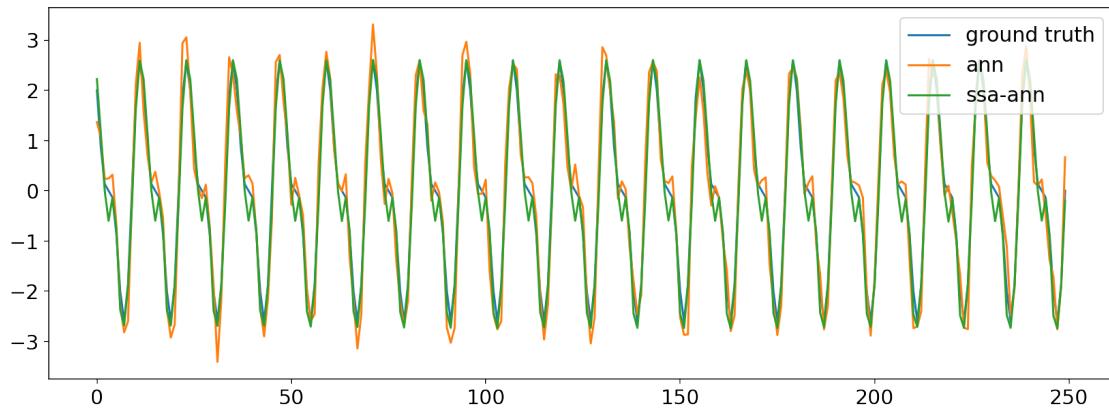


Рис. 4.31. «Сумма синусов с белым шумом». Ряд  $Z_{1500}$ . Прогноз для ANN и SSA-ANN.  $r = 4$

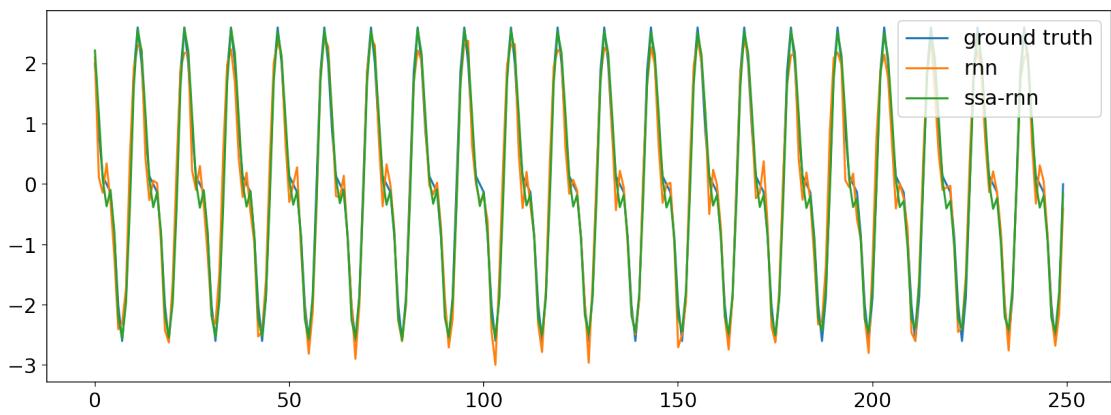


Рис. 4.32. «Сумма синусов с белым шумом». Ряд  $Z_{1500}$ . Прогноз для RNN и SSA-RNN.  $r = 4$

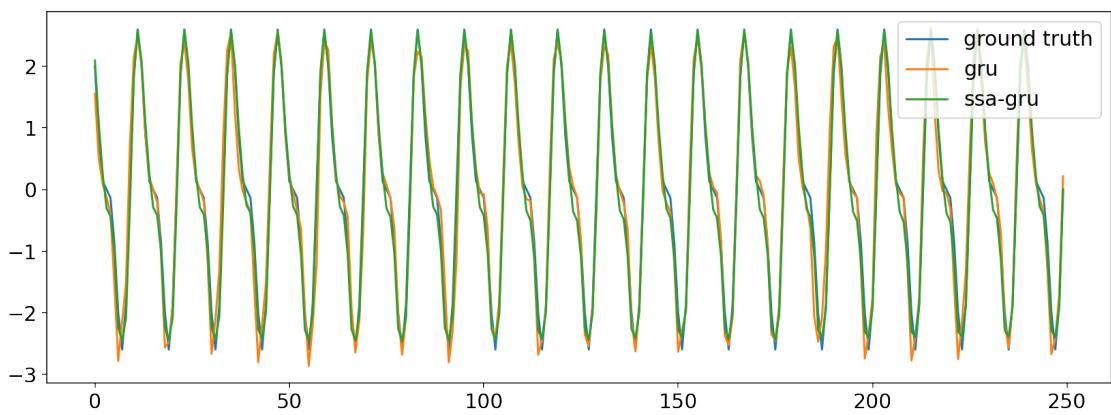


Рис. 4.33. «Сумма синусов с белым шумом». Ряд  $Z_{1500}$ . Прогноз для GRU и SSA-GRU.  $r = 4$

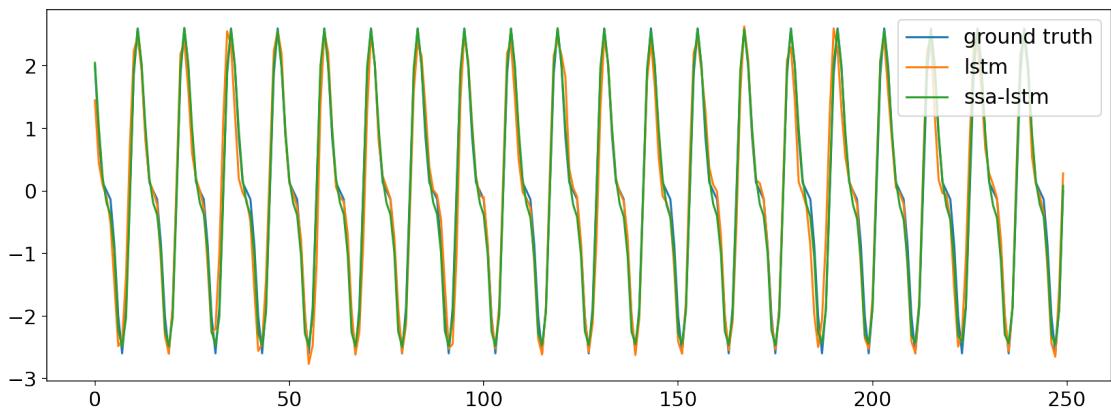


Рис. 4.34. «Сумма синусов с белым шумом». Ряд  $Z_{1500}$ . Прогноз для LSTM и SSA-LSTM.

$$r = 4$$

## 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 параметра  $h$  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.

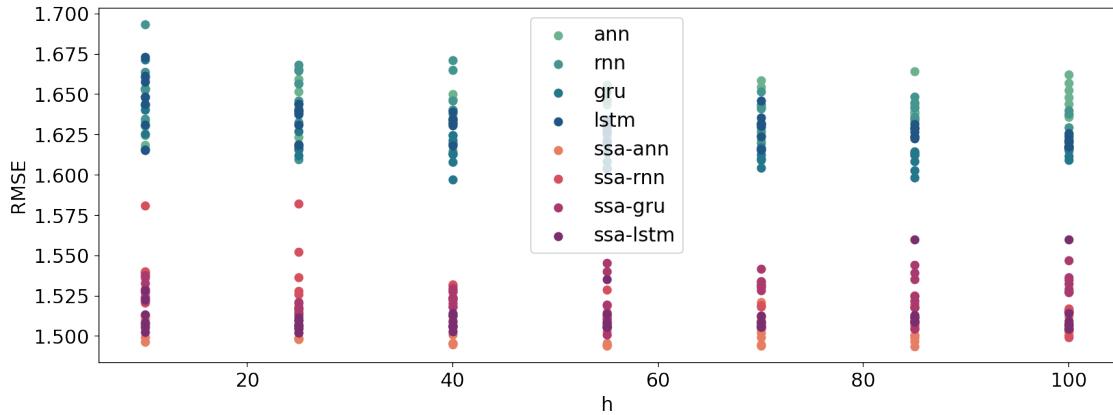


Рис. 4.35. «Сумма синусов с белым шумом». Ряд  $Z_{1500}$ . Проверка устойчивости.

$$r = 4, L = 175, T = 12.$$

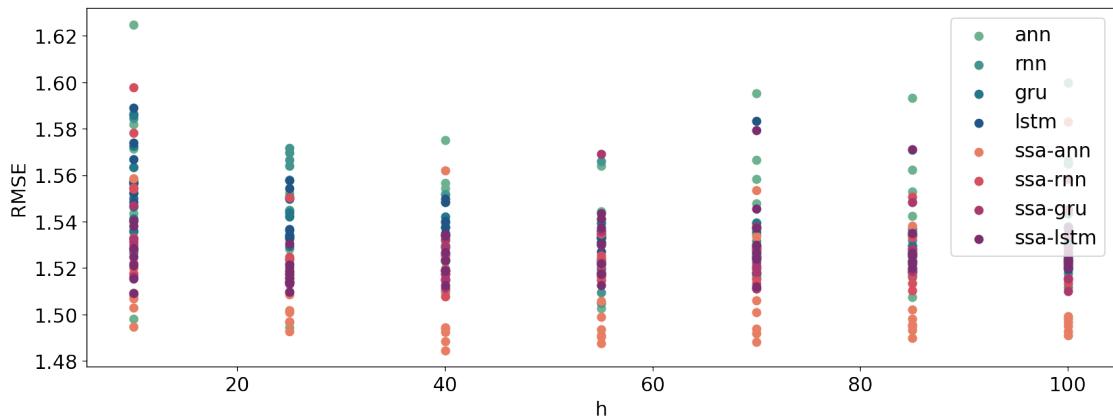


Рис. 4.36. «Сумма синусов с белым шумом». Ряд  $Z_{1500}$ . Проверка устойчивости.

$$r = 4, L = 375, T = 156.$$

## Выводы

Из полученных результатов и таблицы 4.3 (таблица является аналогичной таблице 4.1) можно сделать выводы, что увеличение длины ряда позволяет улучшить прогноз

нейронными сетями, что ведет также к улучшению точности прогнозирования гибридными методами. Также увеличение длины ряда позволяет лучше выделить сигнал, что также приводит к улучшению точности гибридных методов. Из полученных результатов можем заключить, что увеличение длины ряда положительно сказывается на точности прогнозирования обычных и гибридных методов, особенно для метода SSA. Из таблицы видно, что средняя ошибка прогноза методом SSA, меньше, чем средняя ошибка для методов с нейронными сетями.

Таблица 4.3. Усредненные и лучшие результаты прогнозов.

		b-nn	m-nn	b-ssa	m-ssa
ts	ssa-params				
$Z_{650}$	-	1.547	1.635	1.581	1.586
	$L = 175, r = 4$	1.532	1.575	1.581	1.586
$Z_{1500}$	-	1.504	1.550	1.511	1.512
	$L = 375, r = 4$	1.484	1.520	1.511	1.512

## 4.2. Сумма двух синусов с красным шумом. Ряд с трудно отделимым сигналом

Рассмотрим следующий ряд  $Z_{650}$  (рис. 4.37), состоящий из элементов:

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

где  $\xi_i = \xi_{i-1} + \sigma \varepsilon_i$ ,  $\sigma = 1.2$ ,  $\varepsilon_i \sim N(0, 1)$ .

Сигнал данного ряда идентичен сигналу ряда из раздела 4.1. Отличие состоит в шуме в этой главе он красный, что делает выделение сигнала сложной задачей, так как он сильно смешивается с шумом. Это можно увидеть на периодограмме (рис. 4.38).

Другим важным отличием этого примера является то, что красный шум в отличие от белого шума допускает осмысленный прогноз сам по себе. Поэтому можно ставить две разные задачи — прогноз сигнала и прогноз ряда целиком (для белого шума эти задачи не различаются).

Из формулы видно, что сигнал стоит из двух синусов с периодами 12 и 6, общий период ряда 12. В дальнейшем будем выбирать сетки для параметров  $T$  и  $L$  кратные 12. Исходя из раздела 4.1, можно зафиксировать аналитически верными параметры  $L = 175$ ,  $r = 4$  для прогноза сигнала, но для так как из-за красного шума сигнал сложно выделить, то возможно данная пара будет не оптимальной даже для прогноза сигнала. Далее в разделе сравним разные параметры для метода SSA, выберем те, которые дают хорошую точность, а также не слишком хорошо аппроксимируют ряд оценкой сигнала, иначе гибридные методы будут не отличаться от негибридных. Выбирая лучшие пары, будем ориентироваться на прогноз всего ряда.

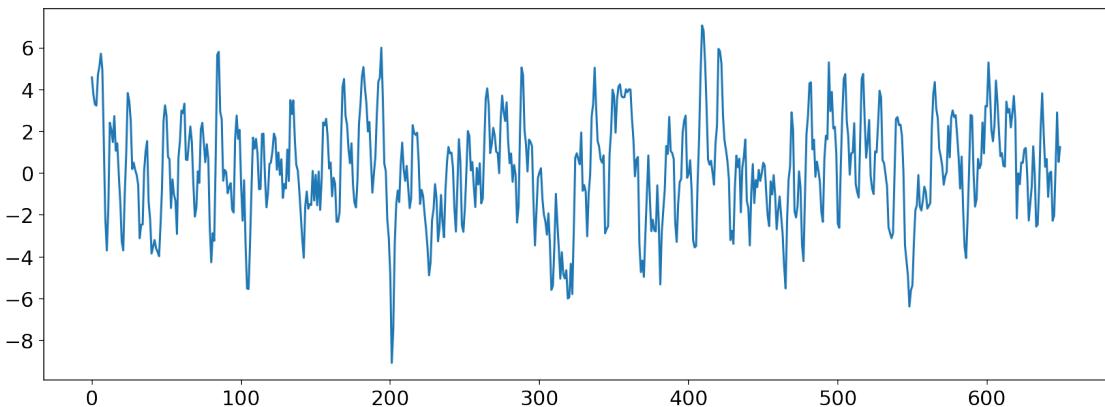


Рис. 4.37. Ряд суммы синусов с красным шумом  $Z_{650}$ .

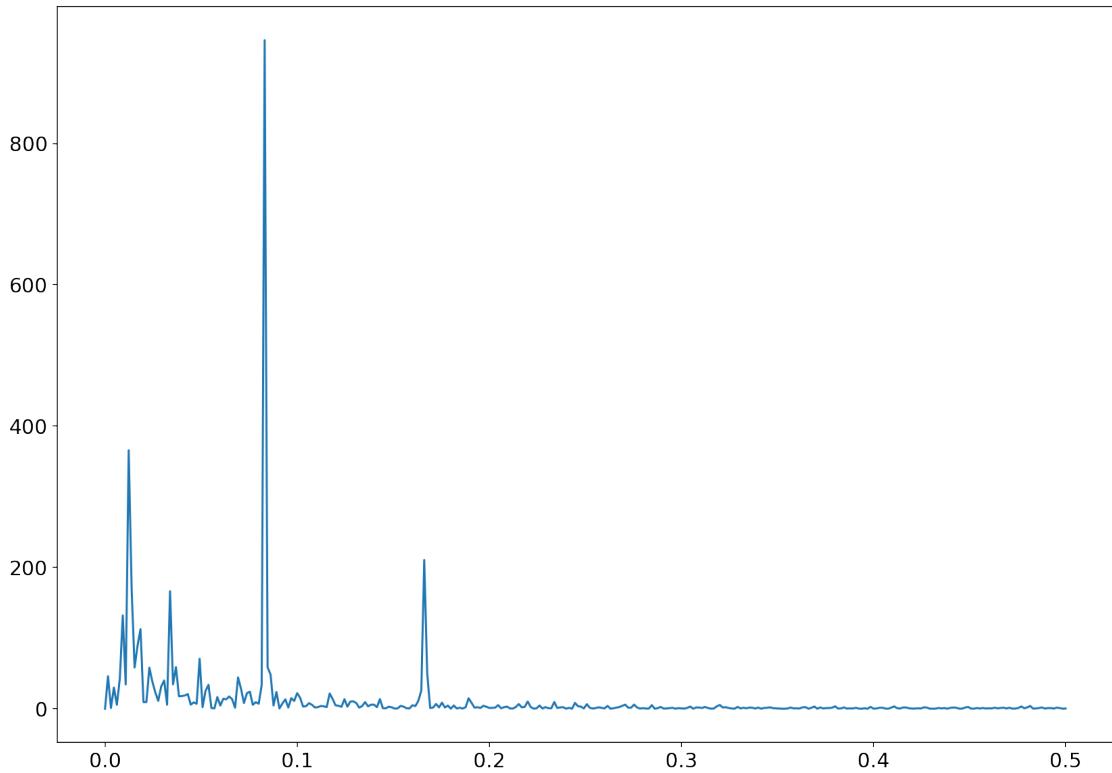


Рис. 4.38. Периодограмма ряда Z<sub>650</sub>.

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

Поставим задачу сравнить обычные, гибридные методы и метод SSA на ряде Z<sub>650</sub>. Проводить сравнение будем по методу, описанному в главе 3.6. В ходе эксперимента хотим посмотреть как поведут себя гибридные методы, столкнувшись с рядом, в котором сложно корректно выделить сигнал. Будем проводить сравнение для аналитически верной пары  $L = 175$ ,  $r = 4$ , а также для оптимальной пары, которую выделим на этапе сравнения параметров SSA прогноза.

#### Прогноз по SSA

Зададим следующую сетку параметров  $L = \{12, 24, \dots, 175\}$ ,  $r = \{2, 4, 6, 8, 10, 12, 16\}$ .

Посмотрим на результаты на рисунке 4.39. На графике видно, что чем меньше  $r$ , тем выше ошибка. Для  $r \geq 8$  разница между ошибками не такая большая, и результаты начинают смешиваться. Из графика видно, что  $r = 2$  дает результаты сильно хуже, чем остальные варианты. Также заметим, что ошибка для аналитически верной пары растет с ростом  $L$ . Пока что сложно выделить, лучшую пару параметров. Прежде чем

сделать это, посмотрим как каждая пара параметров выделяет сигнал на тренировочной выборки.

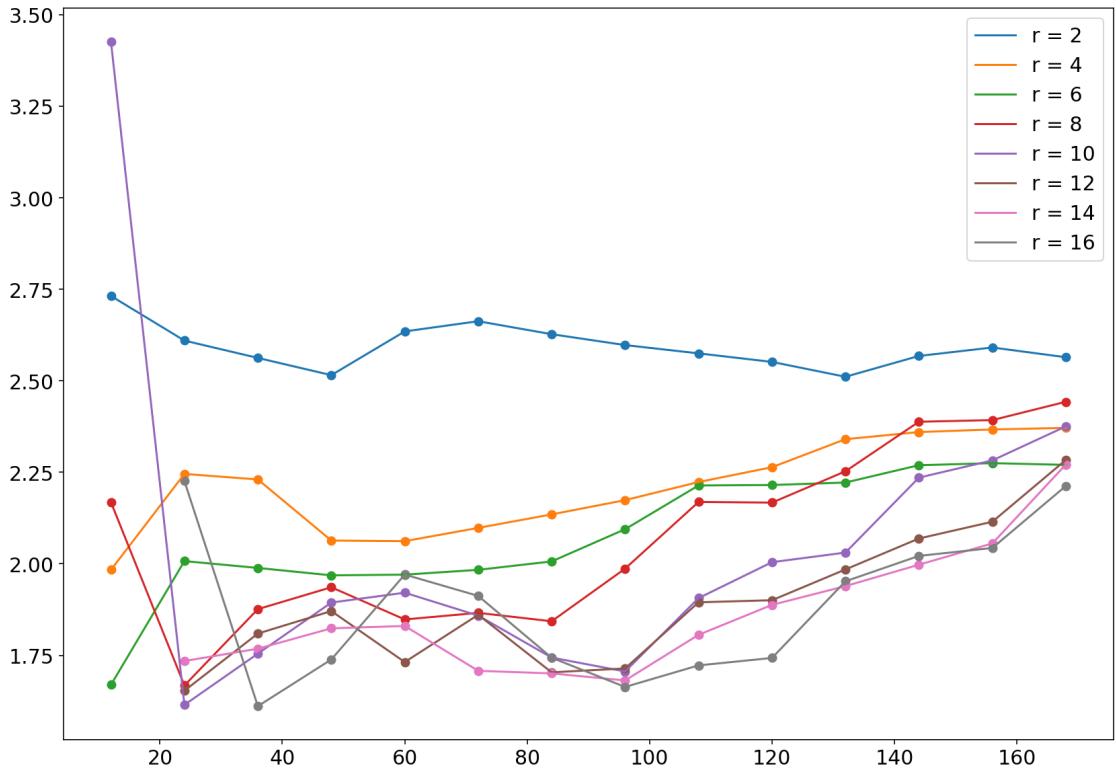


Рис. 4.39. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . RMSE прогноз на валидационной части.

## Восстановление SSA

Посмотрим как SSA восстанавливает тренировочную выборку. На графике показана зависимость ошибки восстановления тренировочной выборки от параметра  $L$ . Сетку для параметров оставим такую же. На графике 4.40 представлены результаты. Видно, что наилучшие результаты показывает пара  $r = 2, L = 175$ . Для других  $r$  ситуация выглядит похоже. Для сравнения возьмем пару  $r = 6, L = 175$  и аналитически верную пару  $r = 4, L = 175$ . Также возьмем две пары  $r = 16, L = 36$  и  $r = 14, L = 84$ , как показывающие низкую ошибку на графике 4.39. Для выбранных пар покажем как метод SSA восстанавливает сигнал.

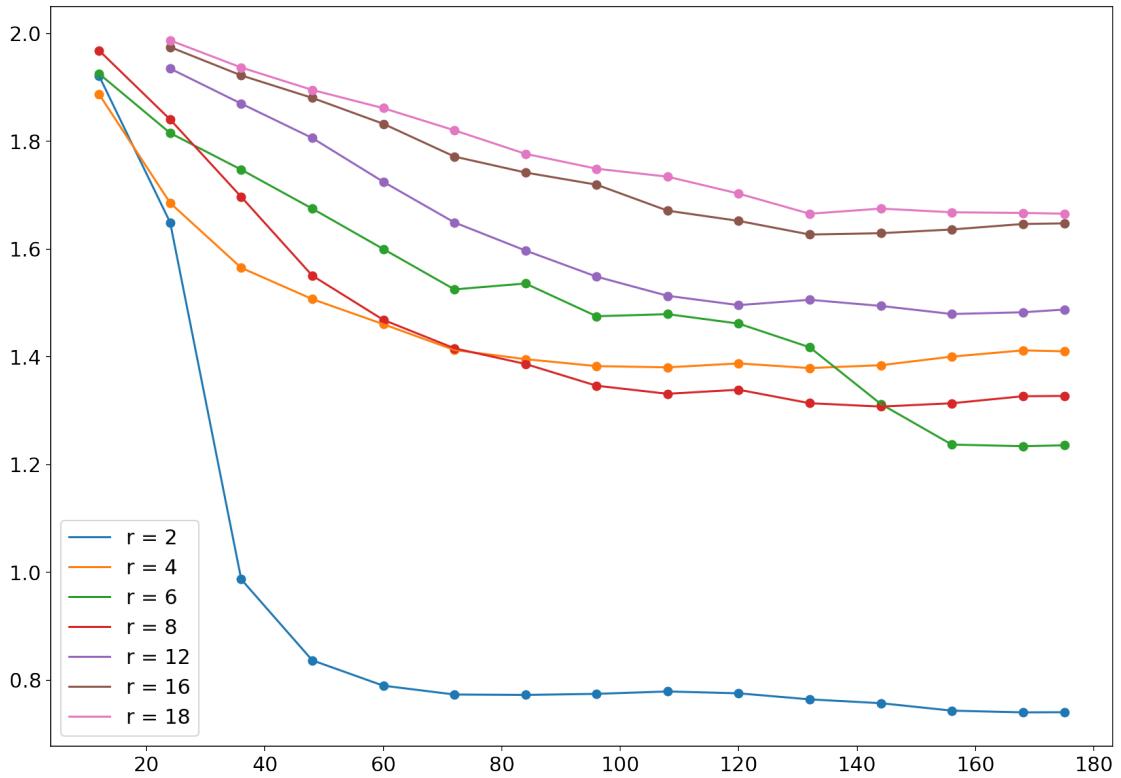


Рис. 4.40. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибка восстановление тренировочной части в зависимости от параметра  $L$ .

Ниже на графиках представлены результаты. Видно, что лучше всего метод SSA восстанавливает сигнал с параметрами  $L = 175$ ,  $r = 2$  (рис. 4.41). Для остальных пар наблюдается влияние шума. На рисунках 4.42, 4.43, 4.44, 4.45 можно наблюдать, что для параметров  $L = 175$  и  $r = \{4, 6\}$  влияние шума меньше, чем для пар  $L = 36$ ,  $r = 16$ ,  $L = 84$ ,  $r = 14$ . На рисунке 4.46 можно заметить, что разницы в восстановленных сигналах для пар  $L = 175$   $r = 4$  и  $L = 175$   $r = 6$  не очень большая. Аналогичные выводы можно сделать для пар  $L = 36$ ,  $r = 16$ ,  $L = 84$ ,  $r = 14$  на рисунке 4.47.

Далее в эксперименте рассмотрим три пары  $L = 175$ ,  $r = 2$ ,  $L = 175$ ,  $r = 4$  и  $L = 84$ ,  $r = 14$ . Пара  $L = 175$ ,  $r = 2$  дает наилучшее восстановление ряда. Выбирая из пар  $L = 36$ ,  $r = 16$ ,  $L = 84$ ,  $r = 14$ , взяли вторую, так как в теории она дает меньшую аппроксимацию оценкой сигнала ряда. Пару  $L = 175$ ,  $r = 4$  выбрали как аналитически верную.

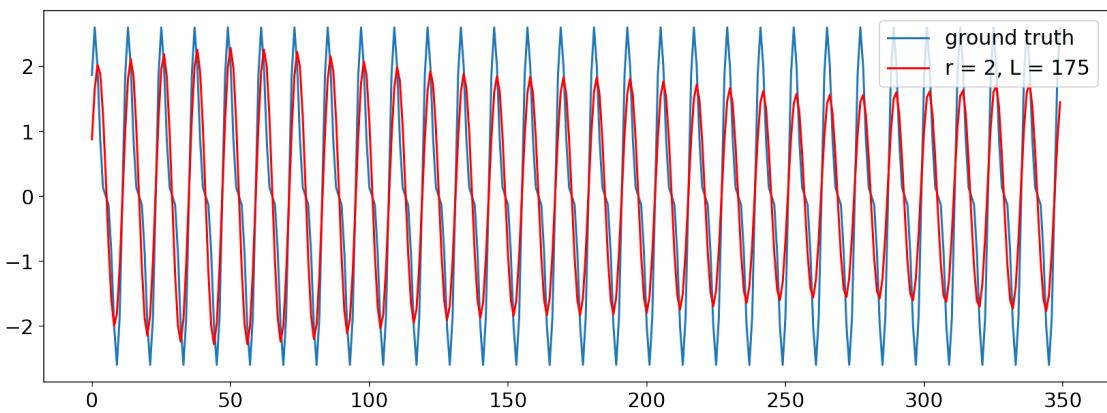


Рис. 4.41. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Восстановление тренировочной выборки с помощью метода SSA.  $L = 175$ ,  $r = 2$ .

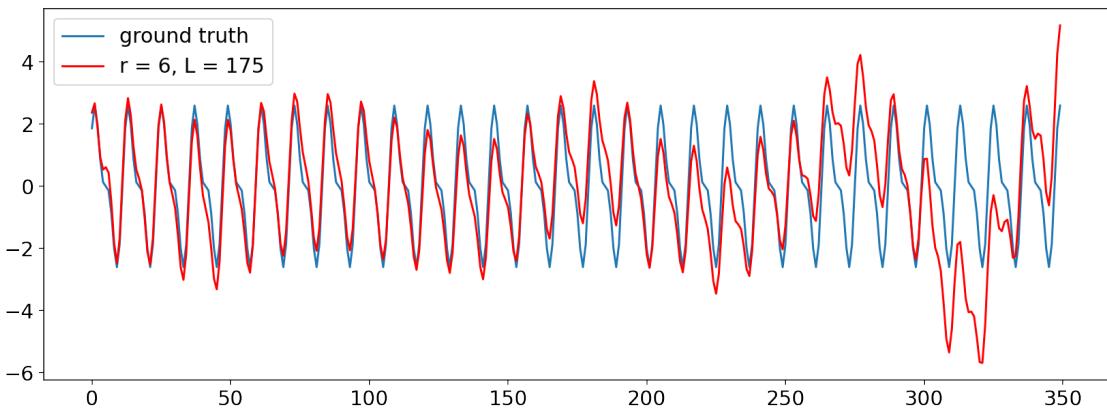


Рис. 4.42. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Восстановление тренировочной выборки с помощью метода SSA.  $L = 175$ ,  $r = 6$ .

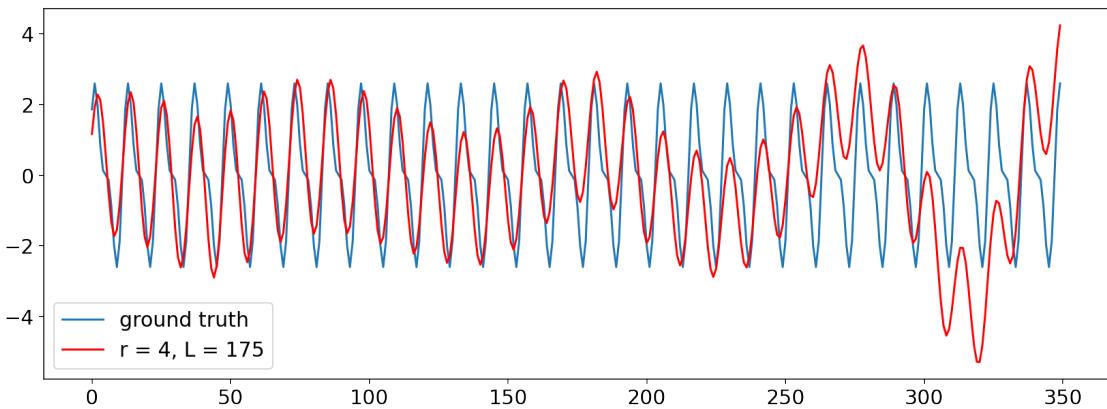


Рис. 4.43. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Восстановление тренировочной выборки с помощью метода SSA.  $L = 175$ ,  $r = 4$ .

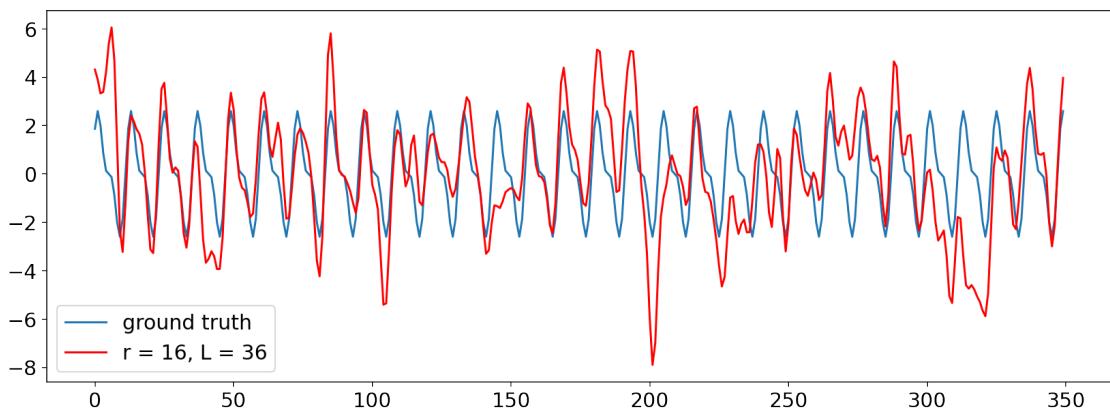


Рис. 4.44. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Восстановление тренировочной выборки с помощью метода SSA.  $L = 36$ ,  $r = 16$ .

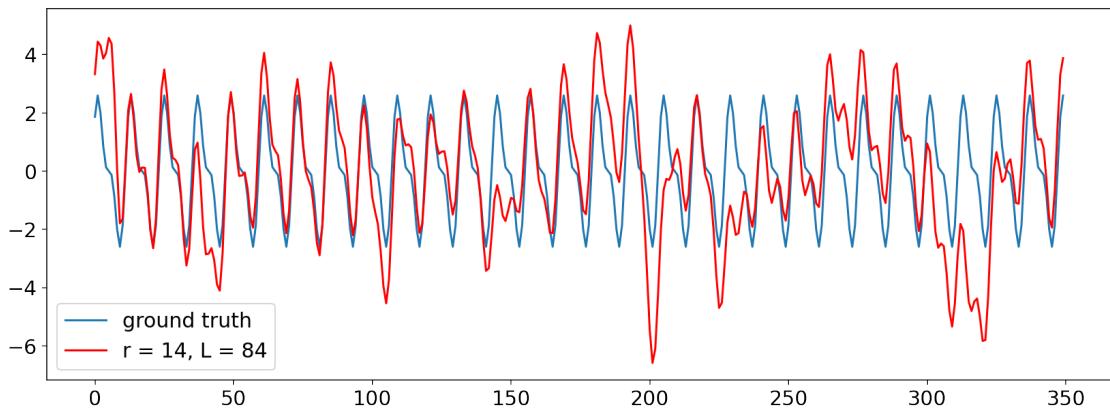


Рис. 4.45. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Восстановление тренировочной выборки с помощью метода SSA.  $L = 84$ ,  $r = 14$ .

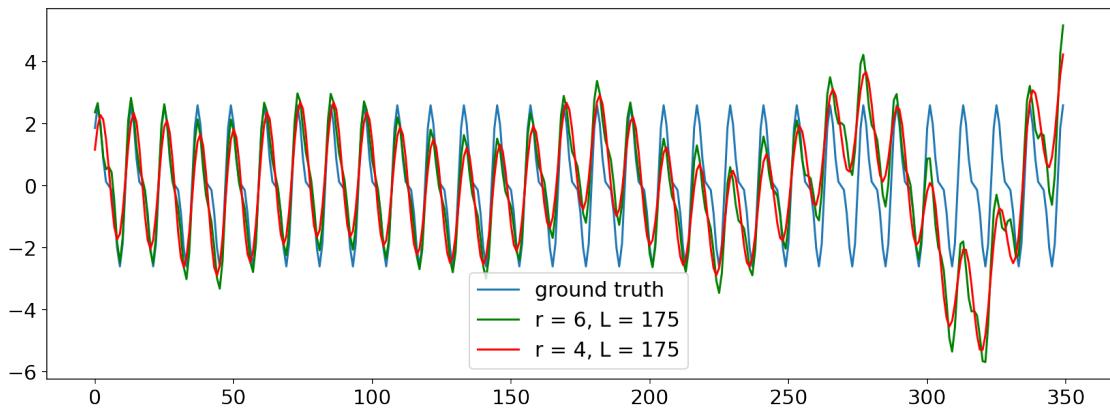


Рис. 4.46. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Восстановление тренировочной выборки с помощью метода SSA.  $L = 175$ ,  $r = 4$  и  $L = 175$ ,  $r = 6$ .

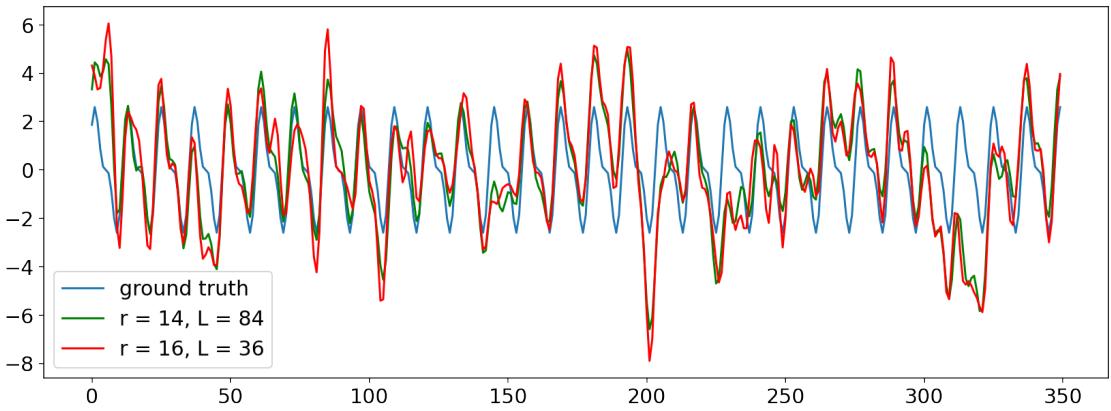


Рис. 4.47. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Восстановление тренировочной выборки с помощью метода SSA.  $L = 84$ ,  $r = 14$  и  $L = 36$ ,  $r = 16$ .

### Сравнение обычных и гибридных методов

Для нейронных сетей зададим следующую сетку параметров:  $T = \{12, 24, \dots, 132\}$ ,  $h = \{10, 25, \dots, 100\}$ . Для метода SSA в гибридных моделях возьмем пары параметров, выбранные выше.

Для прогнозу по методу SSA зададим сетку, что была ранее:  $L = \{12, 24, \dots, 175\}$ ,  $r = \{2, 4, 6, 8, 10, 12, 16\}$ .

Будем сравнивать получившиеся прогнозы относительно сигнала и временного ряда. На графиках 4.48—4.50 (больше графиков в приложении А.2) представлены результаты сравнения относительно временного ряда, диапазон ошибки на оси ординат лежит от 1.3 до 2.2. На графиках 4.51—4.53 результаты сравнения относительно сигнала ряда, диапазон ошибки на оси ординат лежит от 0.25 до 2.2.

На графиках, где показано отклонение от временного ряда, можно заметить, что гибридные методы сильно проигрывают обычным методам. На графике 4.48 можно видеть среднюю ошибку 2.09 для гибридных моделей с параметрами  $L = 175$ ,  $r = 2$ . На графике 4.49 средняя ошибка достигает 2.07 для гибридных моделей с параметрами  $L = 175$ ,  $r = 4$ . На графике 4.50 средняя ошибка падает до 1.95 для гибридных моделей с параметрами  $L = 84$ ,  $r = 14$ . Для обычных методов средняя ошибка составляет примерно 1.4.

На графиках, где показано отклонение от сигнала ряда, можно наблюдать совершенно противоположную ситуацию. Ошибка прогноза гибридных методов сильно упала, а вот ошибка прогноза обычных методов выросла до 1.6. На графике 4.51 можно

видеть среднюю ошибку 0.37 для гибридных моделей с параметрами  $L = 175$ ,  $r = 2$ . На графике 4.52 средняя ошибка достигает 0.87 для гибридных моделей с параметрами  $L = 175$ ,  $r = 4$ . На графике 4.53 средняя ошибка 1.49 для гибридных моделей с параметрами  $L = 84$ ,  $r = 14$ .

Можно заметить, что средняя ошибка для обычных нейронных сетей составляет, как уже говорилось, 1.4, если считать отклонение от ряда. В случае, если считать отклонение от сигнала, средняя ошибка увеличивается до значение 1.6. Эффект легко заметить на графиках ниже. Можно сделать вывод, что обычные нейронные сети пытаются прогнозировать ряд вместе с шумом.

Также видна тенденция, что чем больше параметр  $r$  в гибридных моделях, тем меньше ошибка, посчитанная относительно отклонения от временного ряда. А чем меньше параметр  $r$ , тем больше ошибка, посчитанная относительно отклонения от сигнала ряда. Это говорит о том, что обычные и гибридные методы решают две разные задачи в случае красного шума. Обычные методы пытаются прогнозировать временной ряд вместе с шумом. Гибридные методы прогнозируют сигнал ряда, при условии что в оценку ряда не попал существенное количество шумовых компонент.

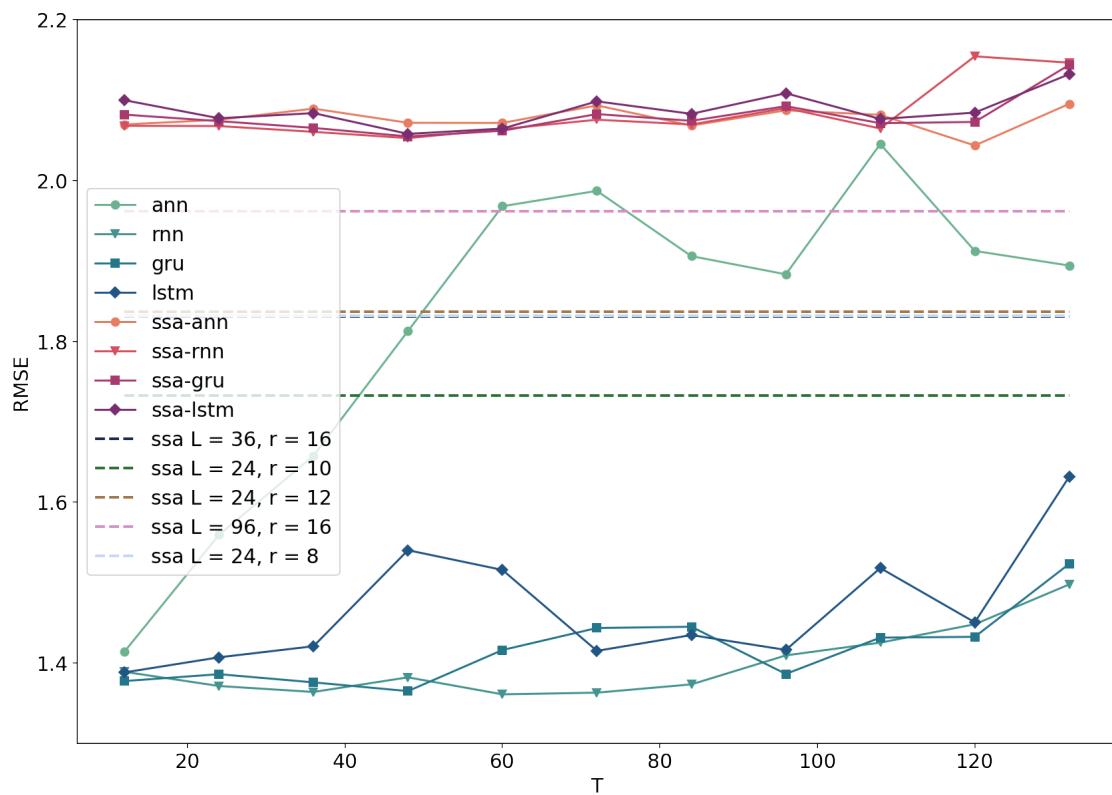


Рис. 4.48. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно ряда в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 2$ .

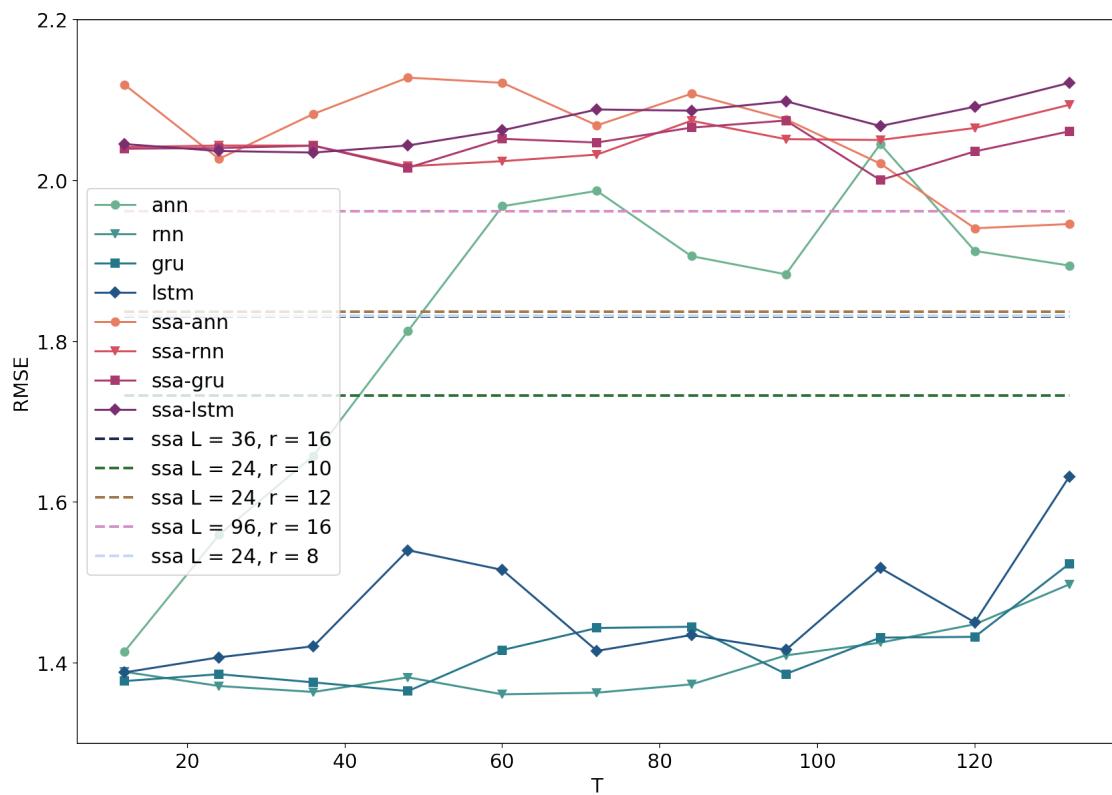


Рис. 4.49. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно ряда в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 4$ .

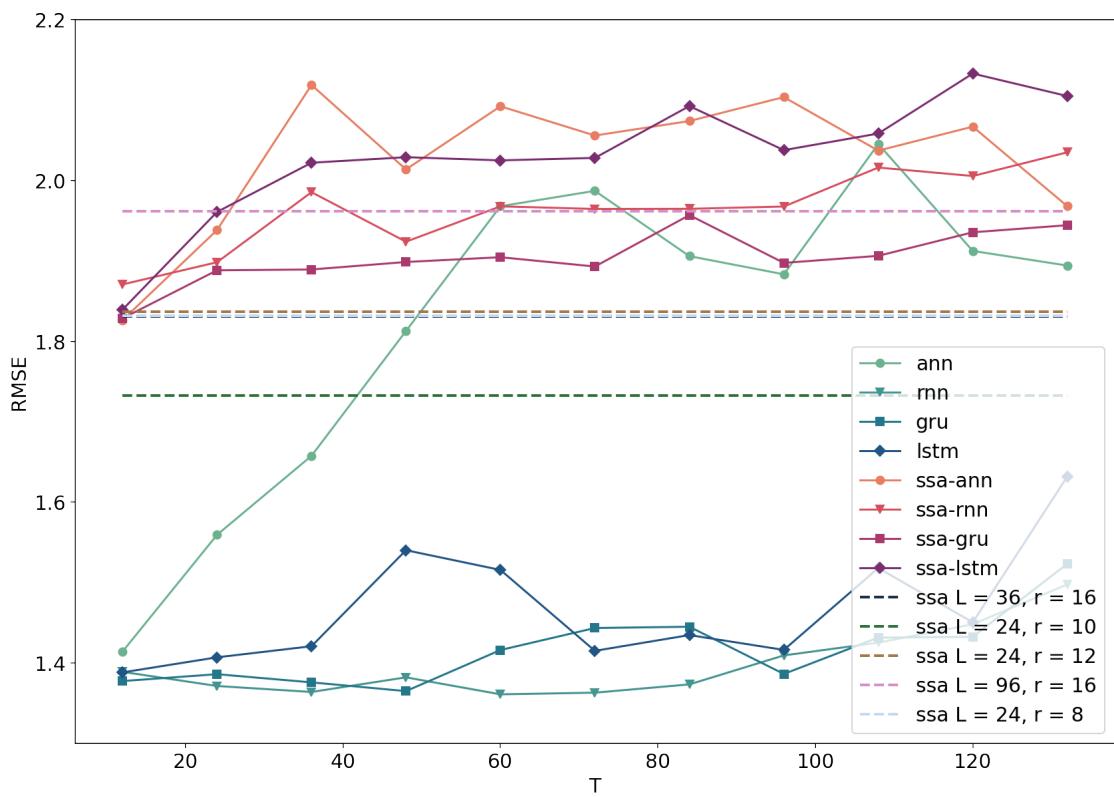


Рис. 4.50. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно ряда в зависимости от параметра  $T$ .  $L = 84$ ,  $r = 14$ .

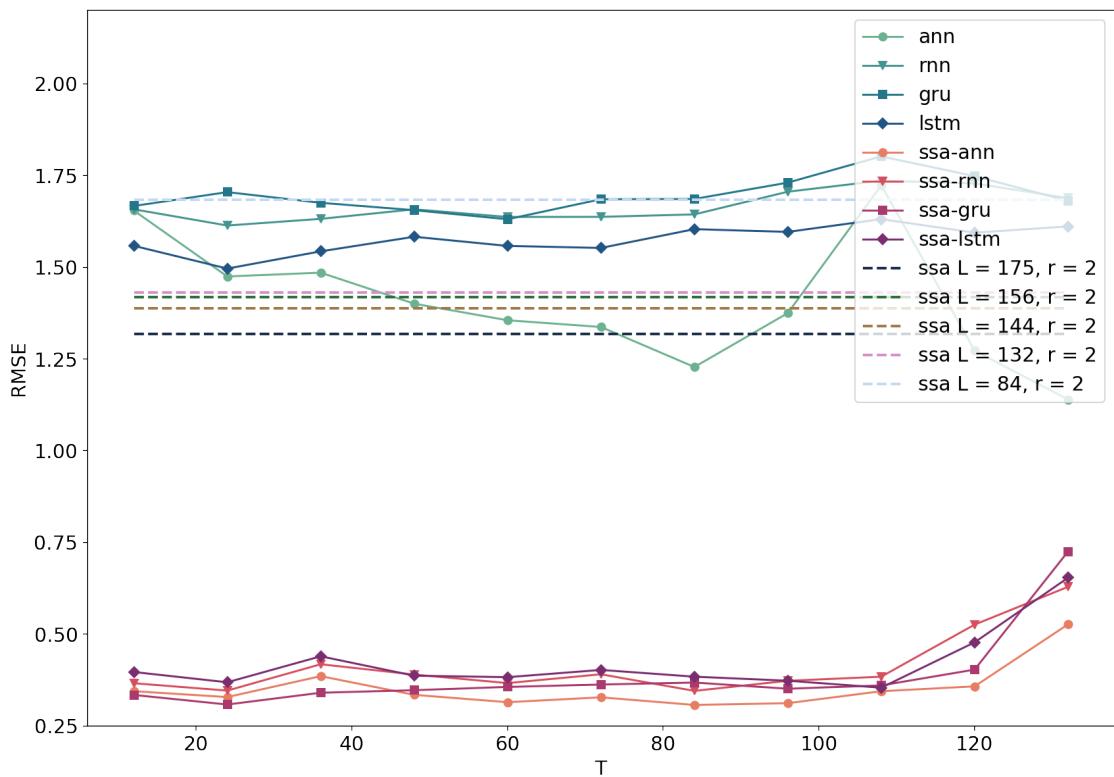


Рис. 4.51. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно сигнала в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 2$ .

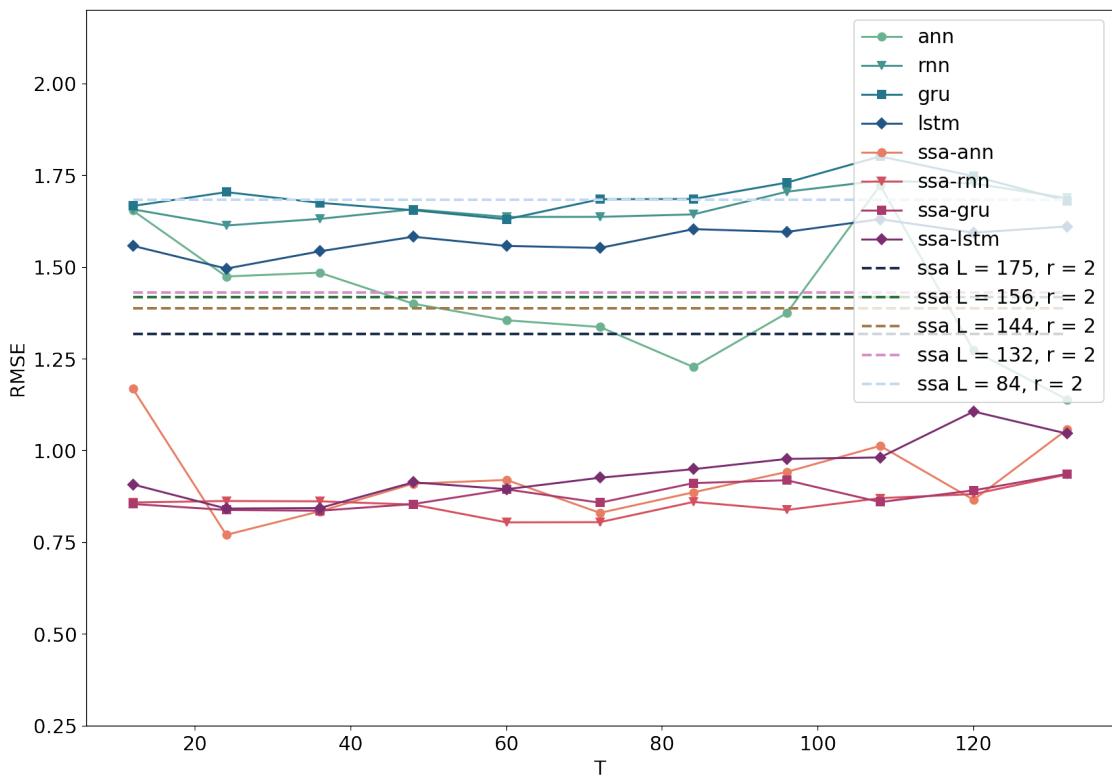


Рис. 4.52. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно сигнала в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 4$ .

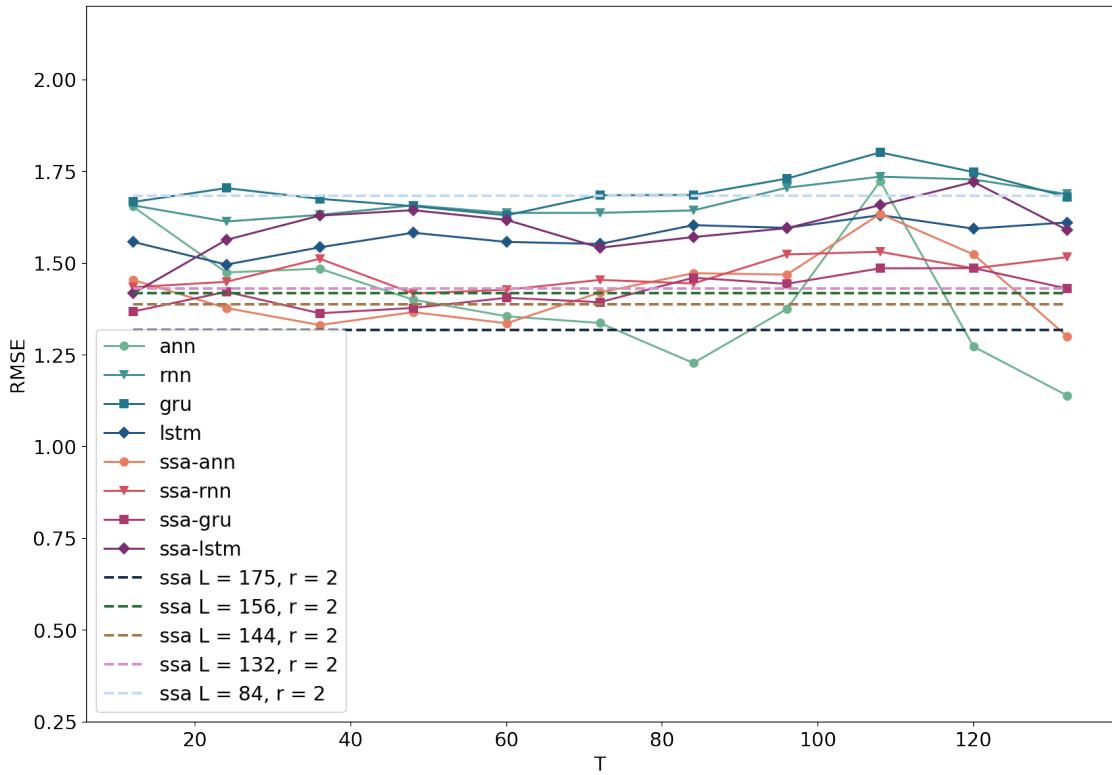


Рис. 4.53. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно сигнала в зависимости от параметра  $T$ .  $L = 84$ ,  $r = 14$ .

### Отображение прогнозов

На графиках 4.54—4.59 (больше графиков в приложении А.2) представлены результаты прогнозирования методами (на графиках показано отклонение от сигнала ряда). Можем заметить, как сильно обычные методы пытаются предсказывать шум. Для пары  $L = 84$ ,  $r = 14$  у гибридных методов предсказание сигнала получилось не очень хорошее. Это объяснимо тем, что в оценке сигнала попало слишком много шумовых компонент. Для пары  $L = 175$ ,  $r = 4$  получилось хорошее предсказание сигнала методов SSA-RNN и SSA-GRU. Для пары  $L = 175$ ,  $r = 2$  хорошее предсказание получилось для трех методов SSA-RNN, SSA-GRU и SSA-LSTM.

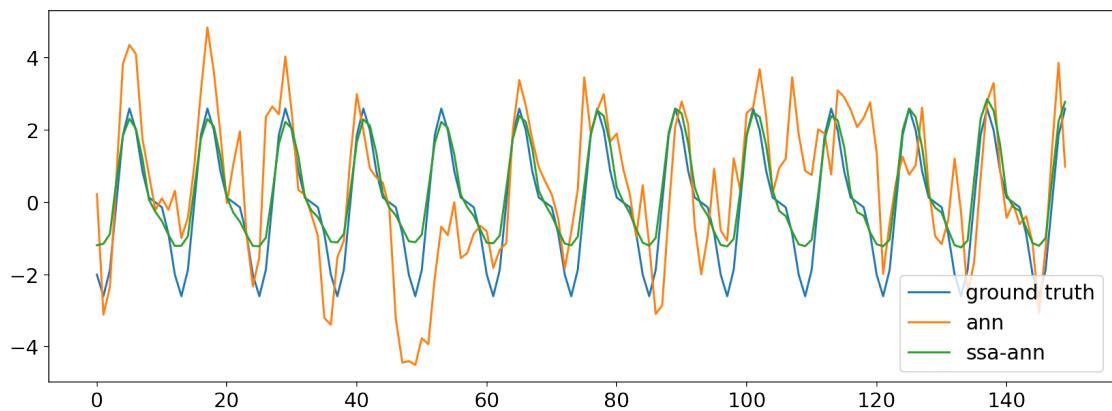


Рис. 4.54. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Прогноз результатов для ANN и SSA-ANN.  $L = 175$ ,  $r = 2$

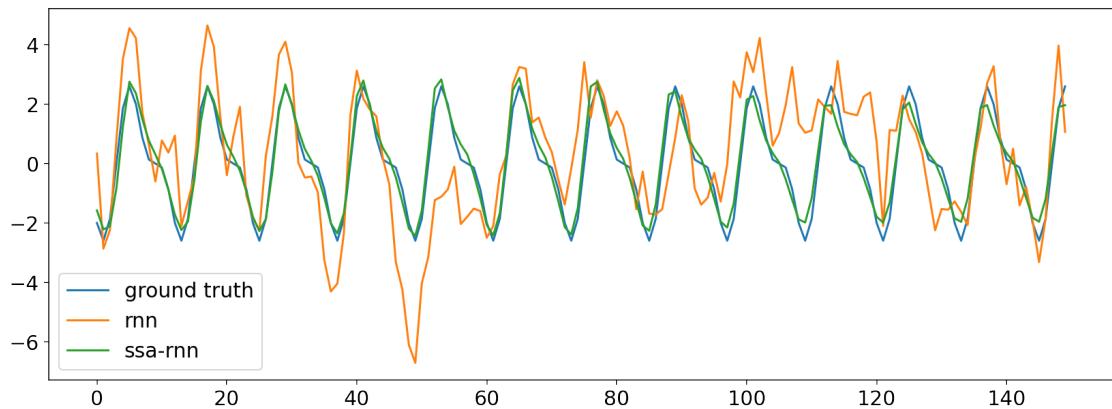


Рис. 4.55. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Прогноз результатов для RNN и SSA-RNN.  $L = 175$ ,  $r = 2$

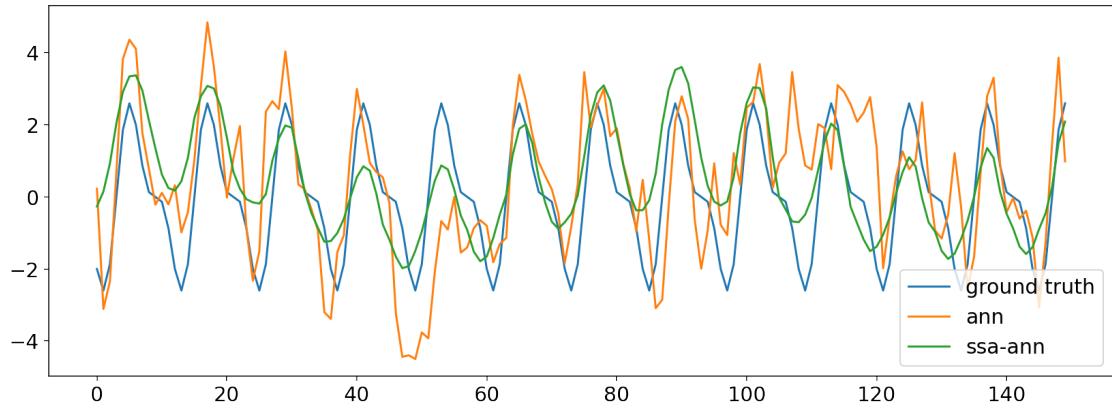


Рис. 4.56. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Прогноз результатов для ANN и SSA-ANN.  $L = 175$ ,  $r = 4$

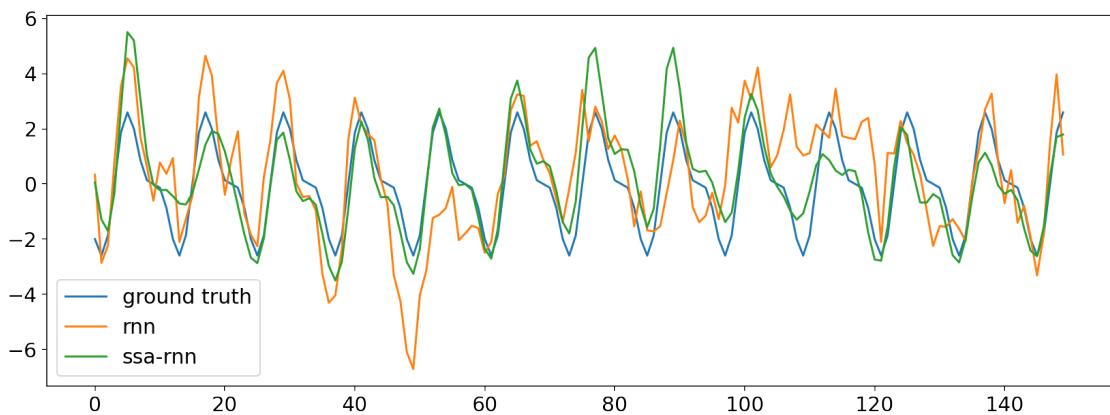


Рис. 4.57. «Сумма синусов с красным шумом». Ряд Z<sub>650</sub>. Прогноз результатов для RNN и SSA-RNN.  $L = 175$ ,  $r = 4$

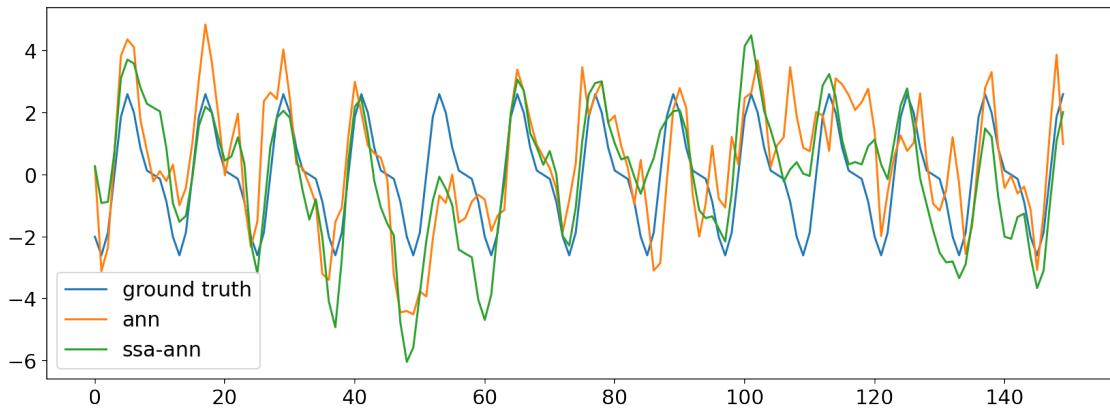


Рис. 4.58. «Сумма синусов с красным шумом». Ряд Z<sub>650</sub>. Прогноз результатов для ANN и SSA-ANN.  $L = 84$ ,  $r = 14$

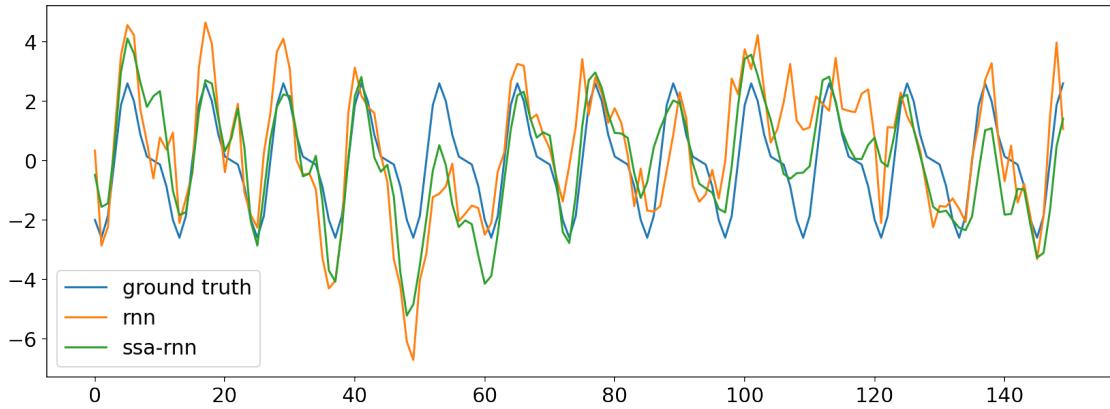


Рис. 4.59. «Сумма синусов с красным шумом». Ряд Z<sub>650</sub>. Прогноз результатов для RNN и SSA-RNN.  $L = 84$ ,  $r = 14$

## 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 параметра  $h$  останется прежней. Будем получать каждый результат по 7 раз, инициализируя метод с новыми весами. Полученные результаты отображены на рисунках 4.60—4.63 (больше графиков в приложении А.2). На них подтверждается, выводы сделанные ранее. Заключаем, что полученные результаты устойчивые.

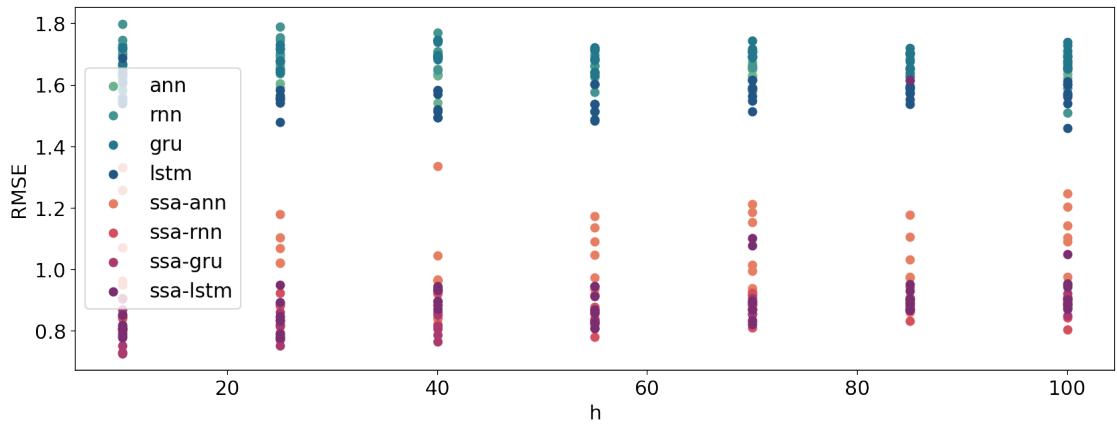


Рис. 4.60. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$$r = 4, L = 175, T = 12$$

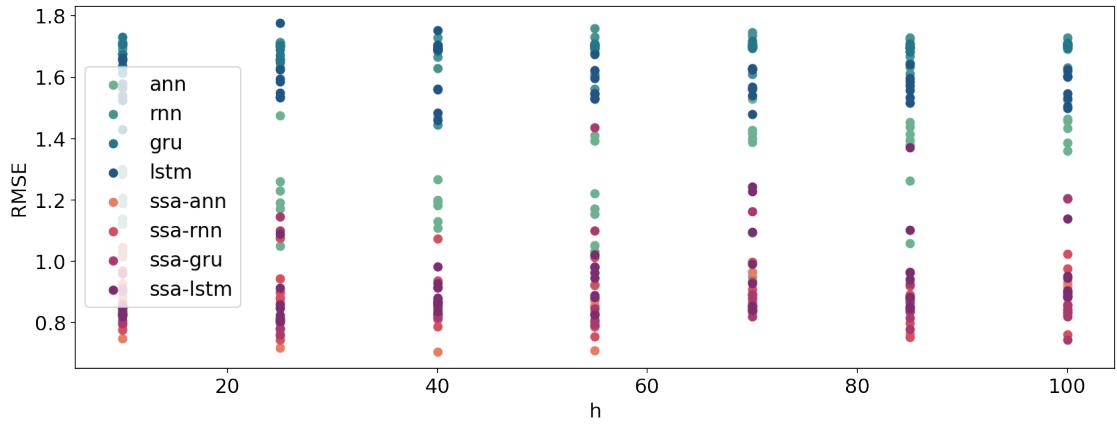


Рис. 4.61. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$$r = 4, L = 175, T = 84.$$

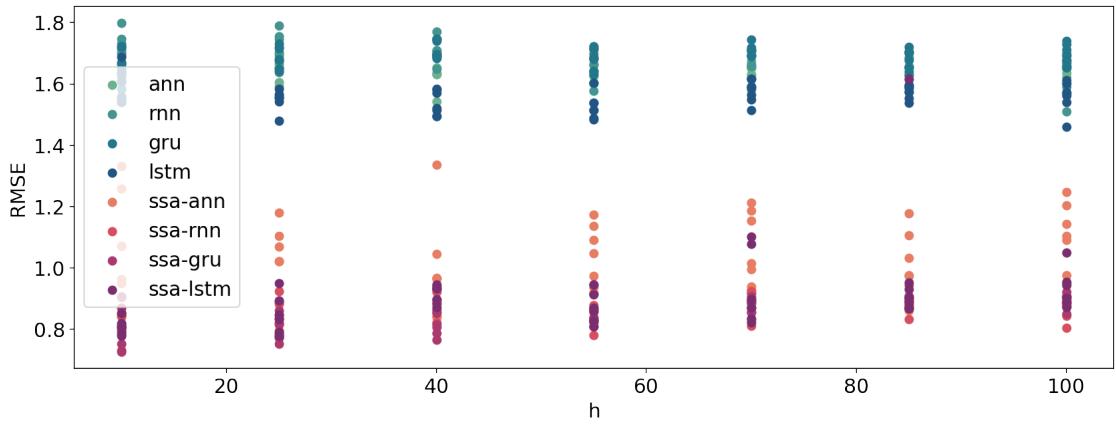


Рис. 4.62. «Сумма синусов с красным шумом». Ряд Z<sub>650</sub>. Проверка устойчивости.

$$r = 4, L = 175, T = 12$$

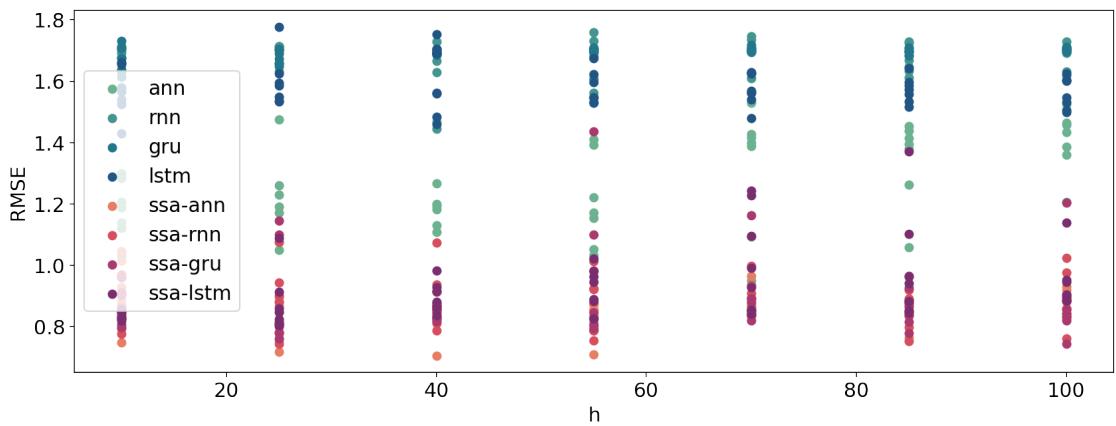


Рис. 4.63. «Сумма синусов с красным шумом». Ряд Z<sub>650</sub>. Проверка устойчивости.

$$r = 4, L = 175, T = 84.$$

## Выводы

Из полученных результатов и таблицы 4.4 (таблица является аналогичной таблице 4.1) видно, что при красном шуме обычные нейронные сети прогнозируют временной ряд вместе с шумом и выполняют эту задачу лучше, чем гибридные методы, что логично, так как в гибридных методах на вход сети подается сигнал без шума. Метод SSA находится между ними при ошибке относительно всего ряда.

Если считать ошибку относительно сигнала ряда, метод SSA показывает наименьшие значения ошибки. Гибридные методы в этом случае прогнозируют сигнал ряда точнее, чем негибридные, при условии, что во время препроцессинга в оценку ряда не попали шумовые компоненты. Гибридные методы имеют меньшую ошибку относительно сигнала ряда, чем обычные методы.

Эксперимент также показал, что при прогнозе сигнала желательно выбирать параметр  $r$  ниже оптимального, если есть вероятность смешивания шума с сигналом ряда. Так даже, если сигнал не был полностью восстановлен, то нейронная в гибридном методе сеть сможет корректно спрогнозировать сигнал ряда. Это было продемонстрировано для пары параметров SSA  $r = 2$ ,  $L = 175$ , гибридные методы получили среднюю ошибку 0.37, в сравнении с 1.6 для обычных методов.

Таблица 4.4. Усредненные лучшие результаты прогнозов методов.

ssa-params	b-nn	m-nn	b-ssa	m-ssa
-	1.294	1.523	1.733	1.803
$L = 175, r = 2$	1.939	2.083	1.733	1.803
$L = 175, r = 4$	1.850	2.056	1.733	1.803
$L = 84, r = 14$	1.736	1.989	1.733	1.803
<hr/>				
-	0.999	1.589	1.319	1.398
$L = 175, r = 2$	0.164	0.385	1.319	1.398
$L = 175, r = 4$	0.694	0.898	1.319	1.398
$L = 84, r = 14$	0.998	1.489	1.319	1.398

### 4.3. Суммарные результаты по модельный рядам

В таблице 4.5 собраны результаты по всем модельным примерам.

Таблица 4.5. RMSE для модельных примеров.

		b-ann		m-ann		b-rnn		m-rnn		b-gru		m-gru		b-lstm		m-lstm		b-ssa		m-ssa	
experiment	ssa-params																				
edsine	-		1.577	1.657		1.547	1.635		1.563	1.623		1.566	1.625		1.581	1.586					
	r2		1.528	1.558		1.554	1.571		1.545	1.566		1.545	1.574		1.581	1.586					
	r4		1.542	1.567		1.551	1.587		1.532	1.566		1.551	1.581		1.581	1.586					
	r6		1.533	1.578		1.557	1.614		1.563	1.629		1.592	1.632		1.581	1.586					
edsine1500	-		1.506	1.574		1.506	1.549		1.504	1.538		1.511	1.539		1.511	1.512					
	r4		1.484	1.513		1.494	1.520		1.495	1.519		1.496	1.529		1.511	1.512					
edsineN0.3	-		0.312	0.326		0.308	0.325		0.311	0.321		0.314	0.330		0.315	0.316					
	r2		0.307	0.311		0.310	0.313		0.308	0.312		0.309	0.314		0.315	0.316					
	r4		0.306	0.312		0.309	0.316		0.305	0.312		0.307	0.317		0.315	0.316					
	r6		0.304	0.313		0.310	0.317		0.304	0.315		0.307	0.320		0.315	0.316					
ser	-		1.381	1.819		1.316	1.396		1.294	1.415		1.296	1.461		1.733	1.803					
	r2LH		1.939	2.078		2.014	2.088		2.025	2.078		2.021	2.088		1.733	1.803					
	r4LH		1.850	2.057		1.982	2.051		1.954	2.045		1.949	2.071		1.733	1.803					
	r14L84		1.781	2.031		1.801	1.968		1.736	1.914		1.800	2.043		1.733	1.803					
serws	-		0.999	1.404		1.451	1.679		1.527	1.697		1.340	1.577		1.319	1.398					
	r2LH		0.164	0.349		0.287	0.404		0.275	0.367		0.286	0.421		1.319	1.398					
	r4LH		0.694	0.919		0.725	0.856		0.729	0.876		0.792	0.943		1.319	1.398					
	r14L84		0.998	1.443		1.238	1.475		1.221	1.439		1.339	1.597		1.319	1.398					

## 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.

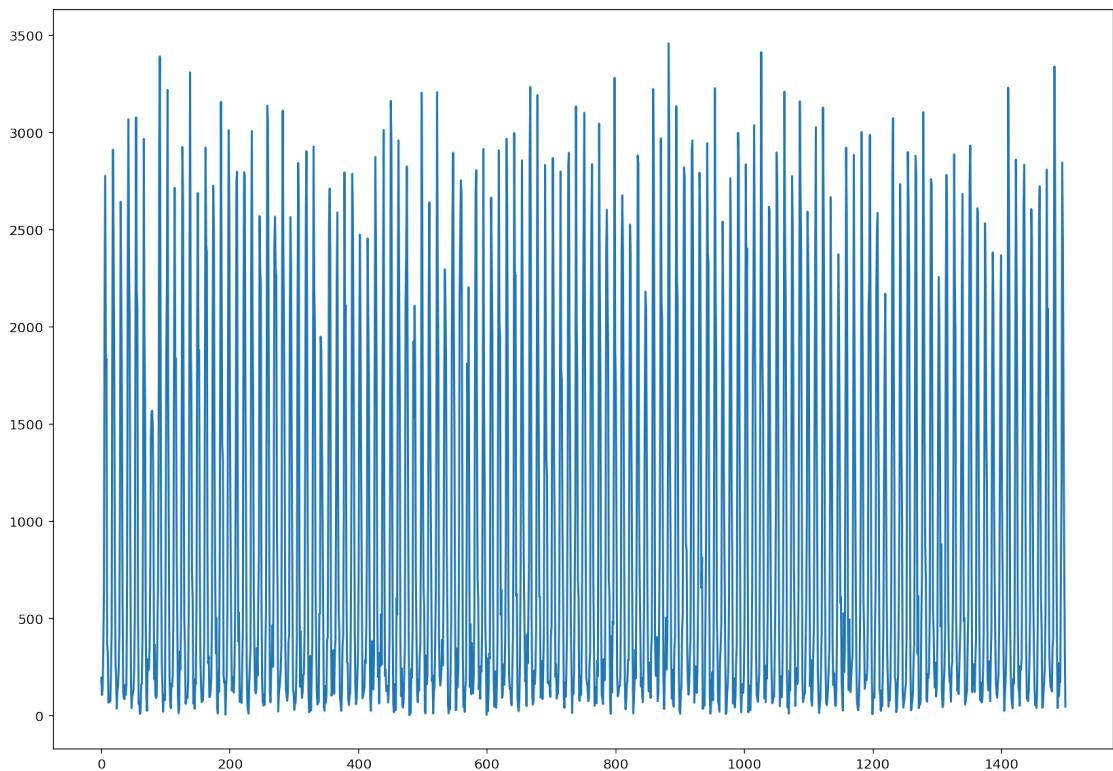


Рис. 5.1. Данные «Indian Rain».

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 в ряде равен 12. Далее в экспериментах будем перебирать параметры  $T$  и  $L$  по сетке с шагом кратным 12.

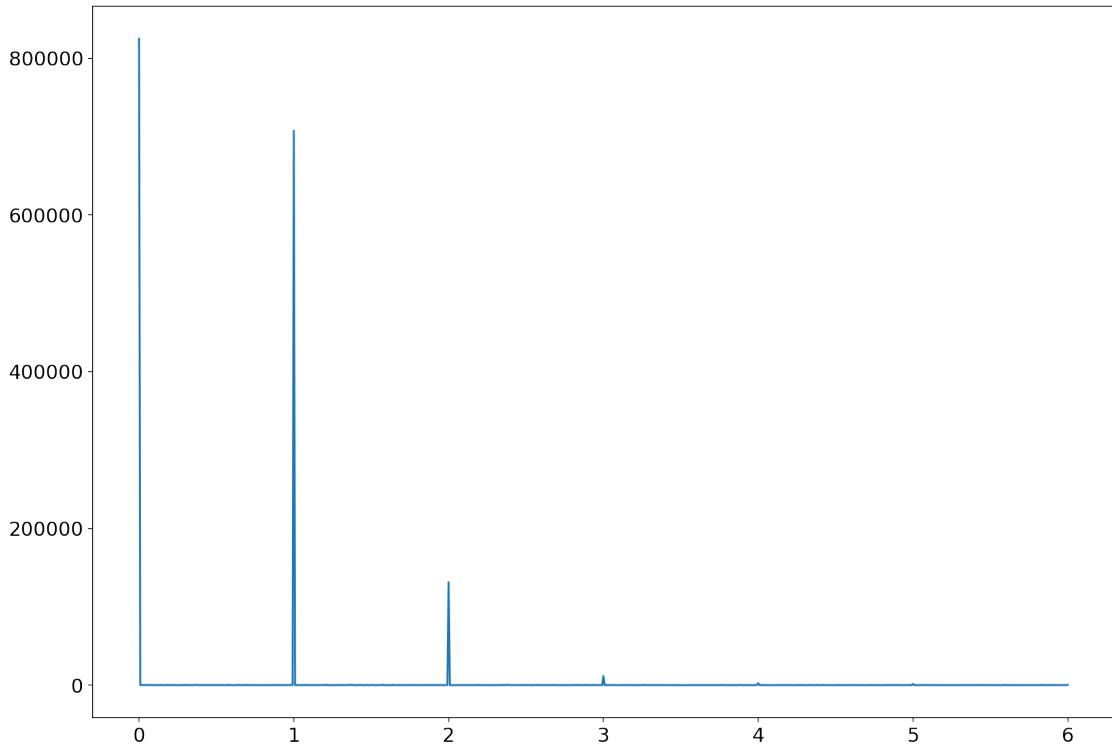


Рис. 5.2. Периодограмма ряда «Indian Rain».

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

Сравним метод SSA, обычные и гибридные методы по методике, описанной в разделе 3.6. <sup>1</sup>

#### 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  $r = 7$ ,  $L = 375$ .

---

<sup>1</sup> Вы часто пишете глава вместо раздел. Проверьте везде.

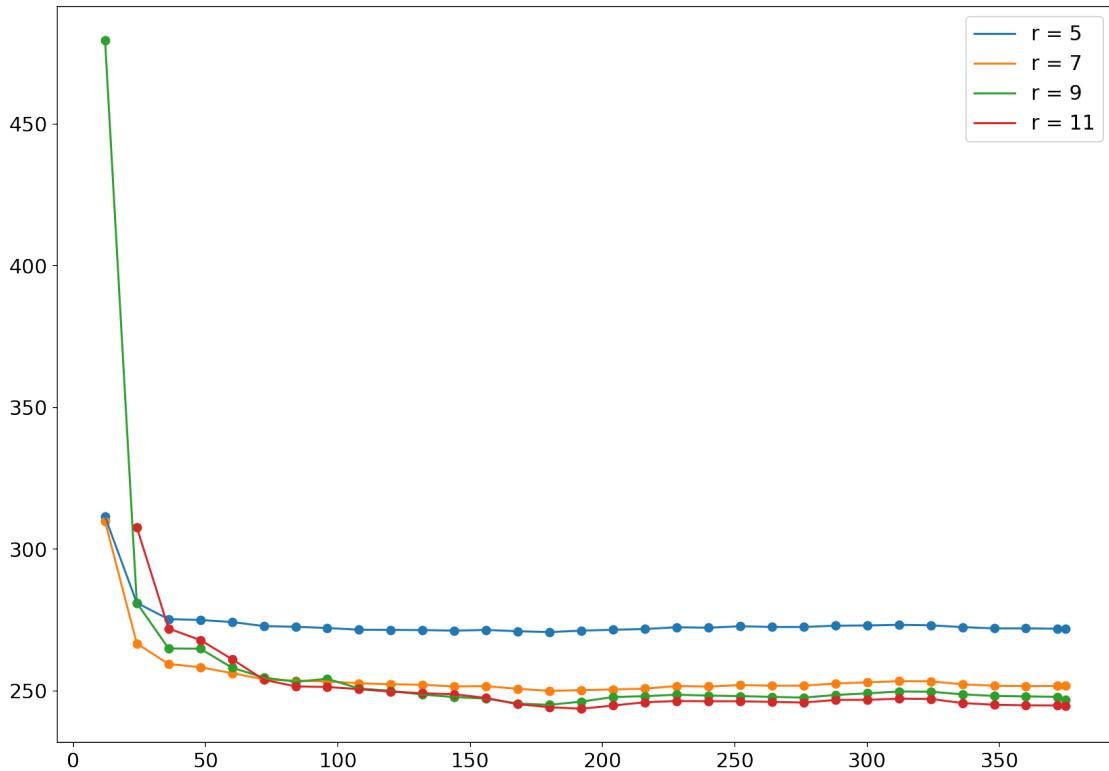


Рис. 5.3. Данные «Indian Rain». Прогнозирование с помощью метода SSA.

### 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.

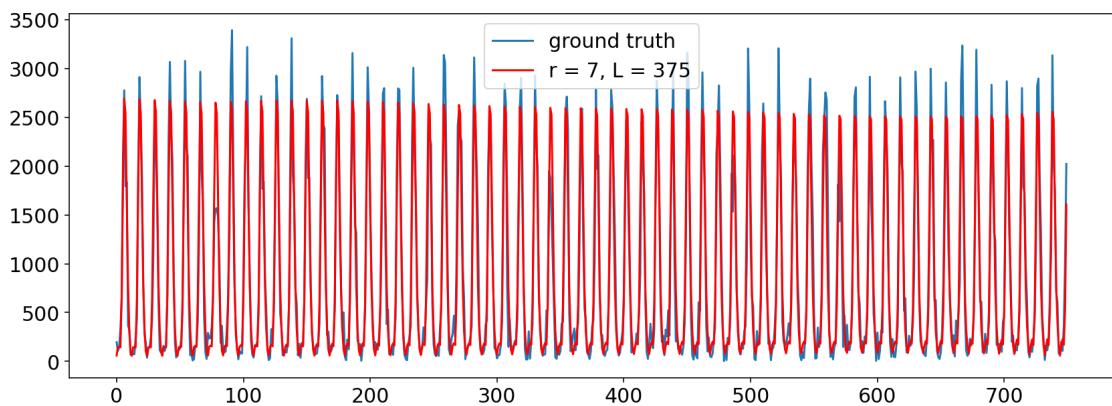


Рис. 5.4. Данные «Indian Rain». Восстановление тренировочной выборки с помощью метода SSA.  $r = 7$ ,  $L = 375$ .

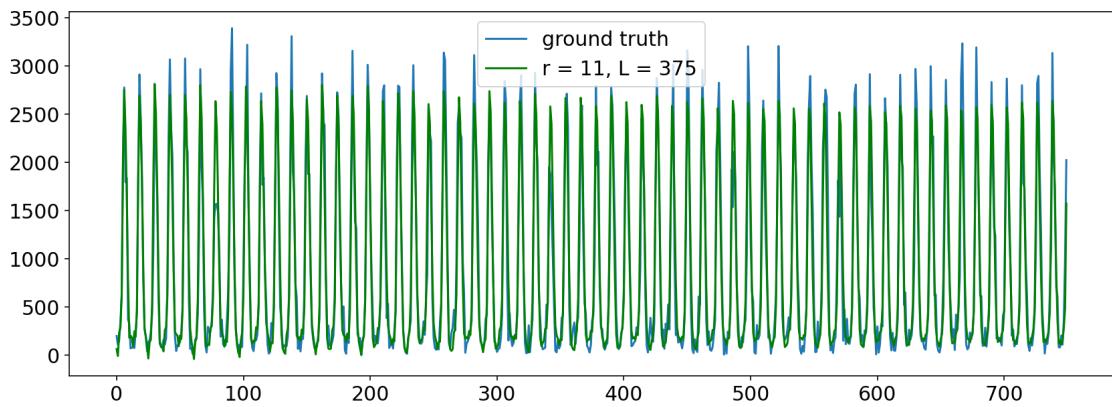


Рис. 5.5. Данные «Indian Rain». Восстановление тренировочной выборки с помощью метода SSA.  $r = 11$ ,  $L = 375$ .

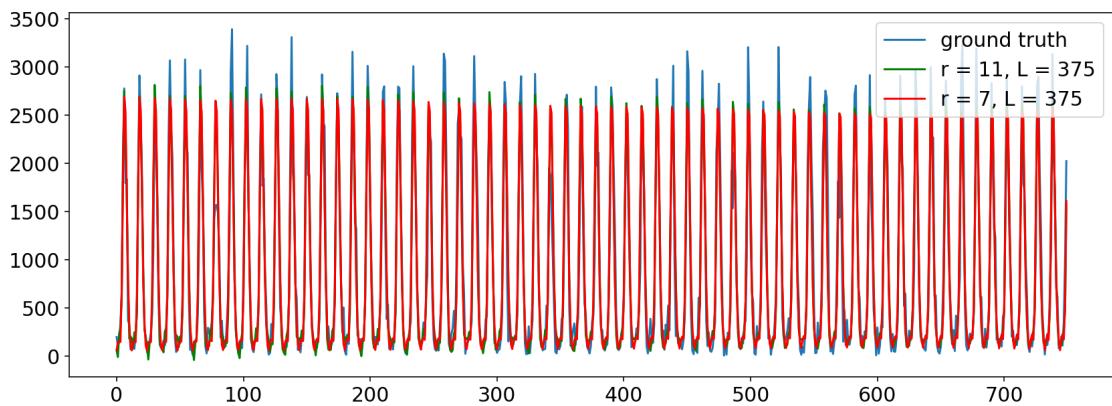


Рис. 5.6. Данные «Indian Rain». Восстановление тренировочной выборки с помощью метода SSA. Обе пары.

## Сравнение методов

Для нейронных сетей зададим следующую сетку 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  $T$ .

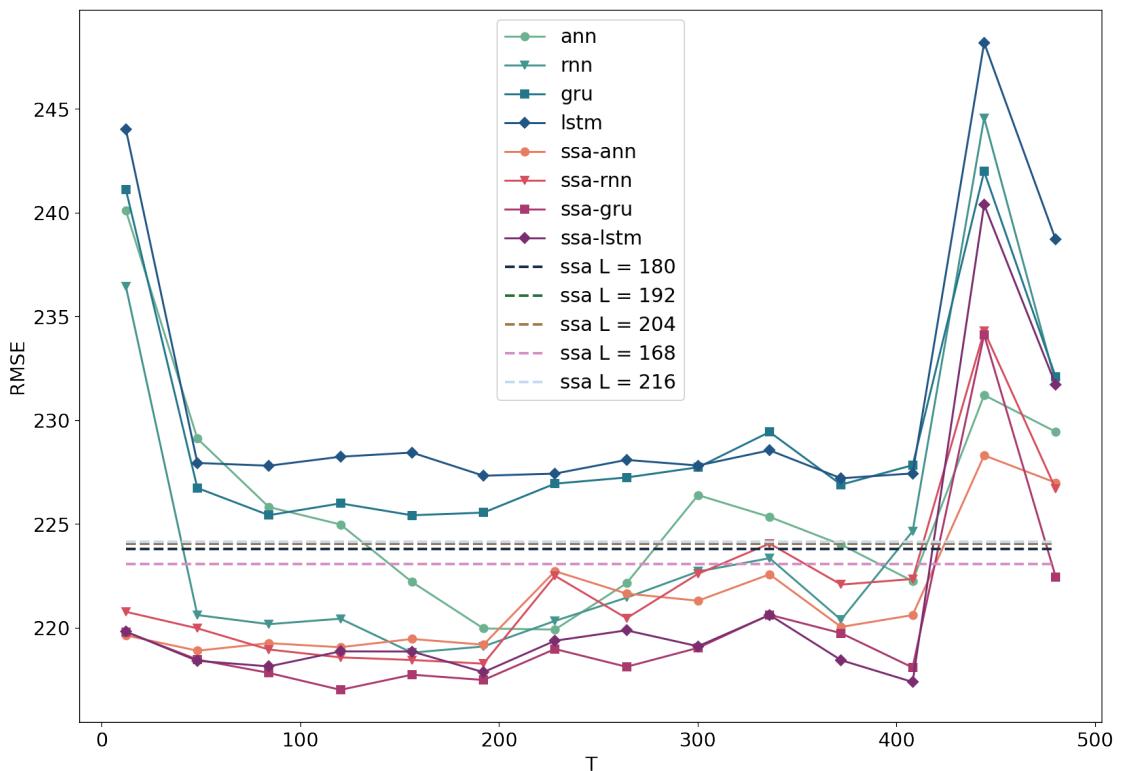


Рис. 5.7. Данные «Indian Rain». Ошибки прогноза в зависимости от параметра  $T$ .

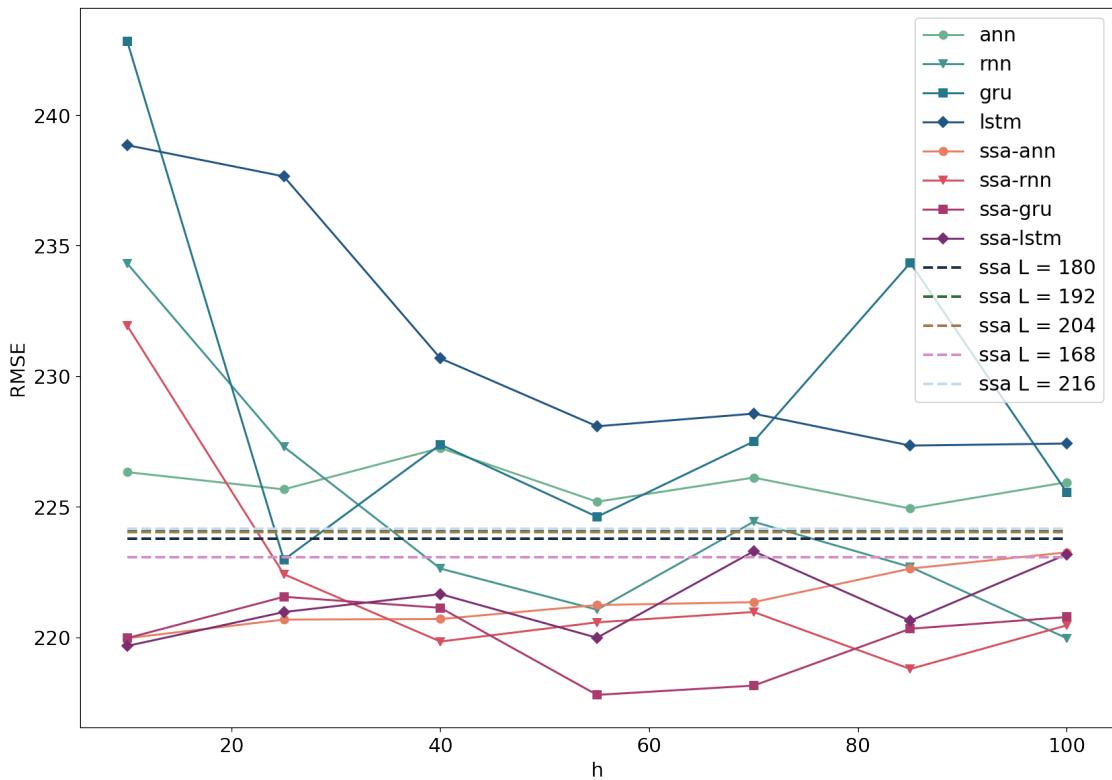


Рис. 5.8. Данные «Indian Rain». Ошибки прогноза в зависимости от параметра  $h$ .

## Отображение прогнозов

На графиках Б.1–Б.4 видно, что прогнозирование обычными и гибридными методами практически одинаковое.

## 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 параметра  $h$  останется прежней. Будем получать каждый результат по 7 раз, инициализируя метод с новыми весами. Полученные результаты отображены на рис. 5.9, 5.10.

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

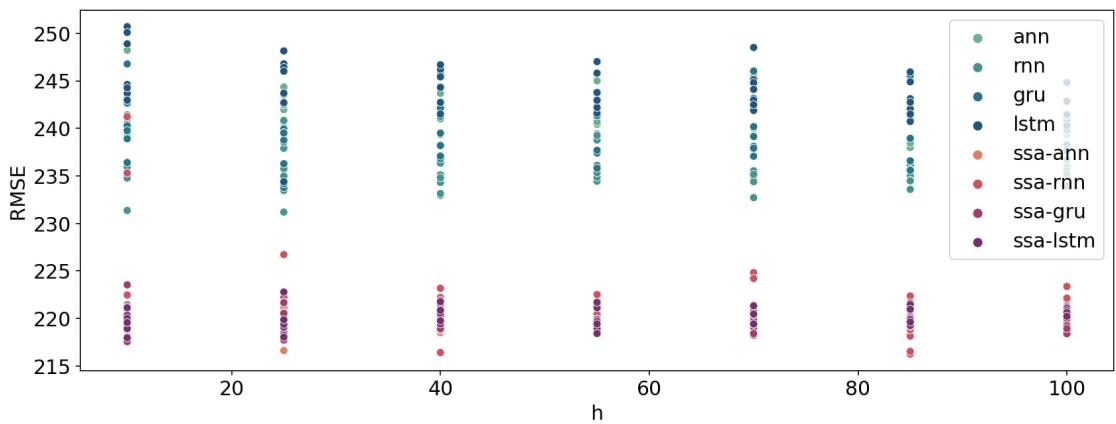


Рис. 5.9. Данные «Indian Rain». Проверка устойчивости.  $T = 12$ .

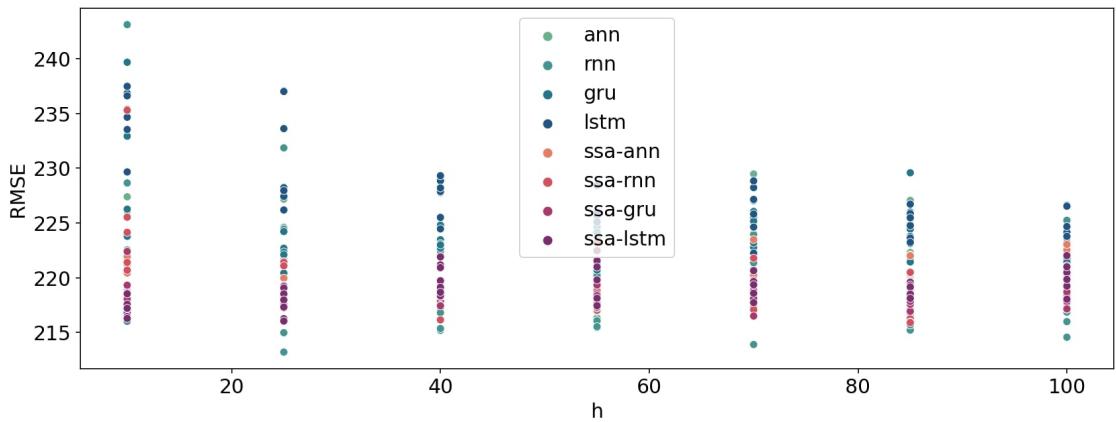


Рис. 5.10. Данные «Indian Rain». Проверка устойчивости.  $T = 156$ .

## Выводы

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

Успех гибридных методов объясним тем, что ряд «Indian Rain» имеет простой сигнал конечного ранга. Ряд имеет хорошую длину, что обеспечивает лучшее выделение сигнала и большее количество пар «признаки – предсказываемые значения», это положительно сказывается на точности предсказаний гибридных моделей. Также в ряде есть шум, что делает использование препроцессинга SSA логичным.

## 5.2. Earth Orientation Parameters (EOP)

Рассмотрим следующий ряд  $Z_{717}$  (рис. 5.11). Временной ряд отображает координату по оси абсцисс земного полюса.<sup>2</sup>

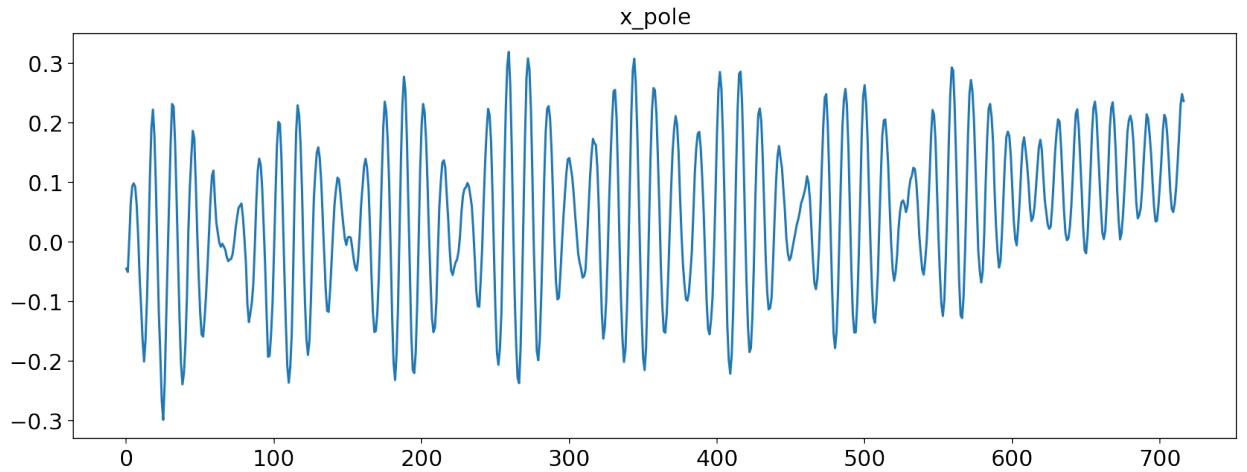


Рис. 5.11. Временной ряд « $x$  pole».

Вычтем тренд из  $Z_{717}$  (рис. 5.12). Будем считать, что правая часть ряда не похожа на остальную временнной 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.

---

<sup>2</sup> Везде нужны ссылки на данные. Для Погоды не нашла, откуда конкретно скачала, поэтому дайте ссылка на kaggle.

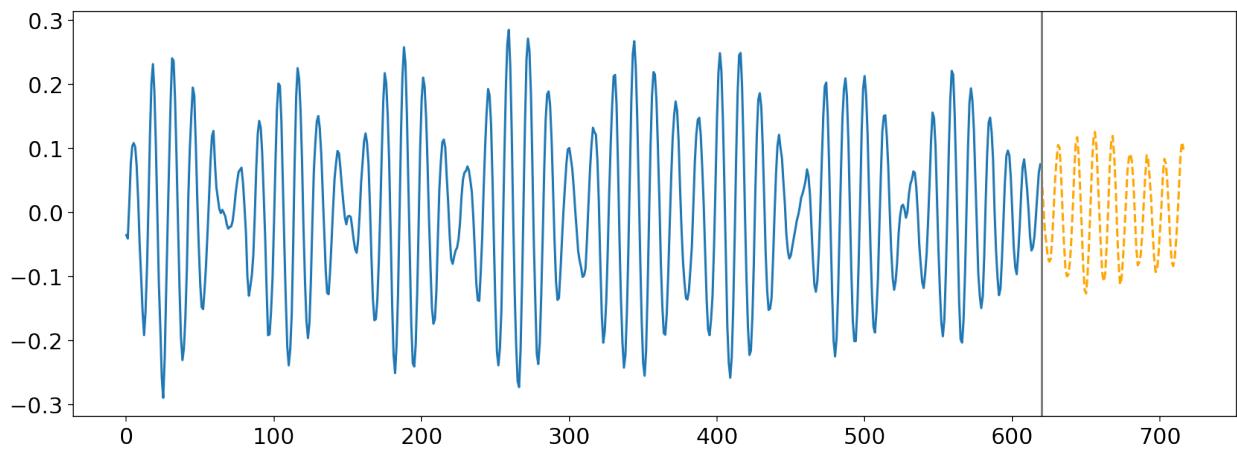


Рис. 5.12. Временной ряд «x pole» без тренда.

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. 5.13. На графике видно две близкие периодики, которые смешались. Это говорит, что у ряда сложный сигнал, который будет трудно выделить корректно.

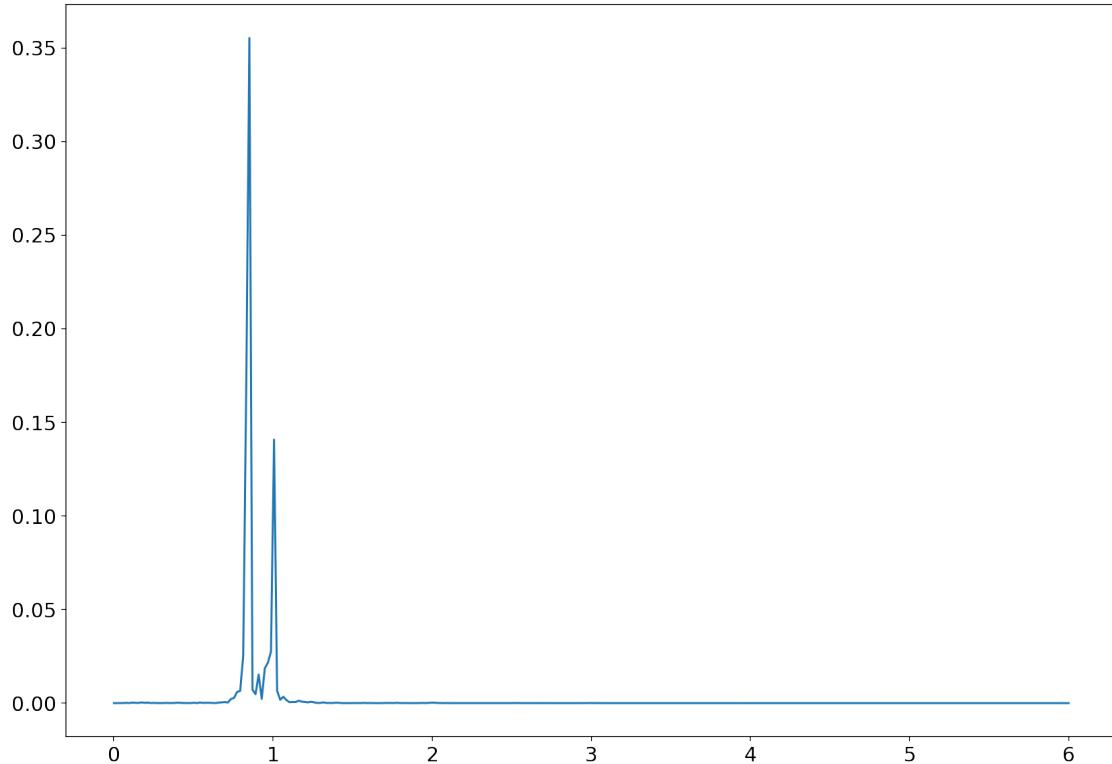


Рис. 5.13. Периодограмма ряда «x pole» без тренда.

### 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  $L = \{13, 26, \dots, 160\}$ , так как основная частота около  $1/13$ , и  $r = \{6, 8, 12, 16, 18\}$ .

Посмотрим на результаты на рисунке 5.14. 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 нене-лесообразным. Возьмем пару параметров по середине:  $L = 78$ ,  $r = 18$ , далее будем использовать эту пару в гибридных моделях.

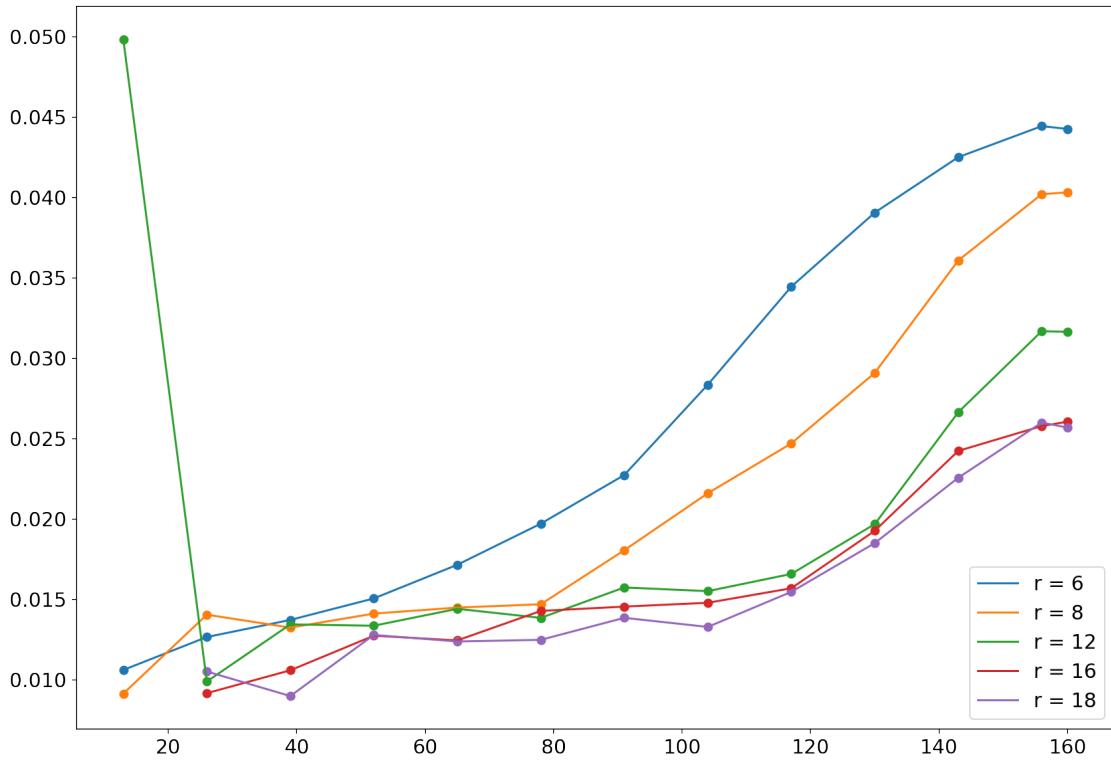


Рис. 5.14. Ряд «x pole» без тренда. Прогнозирование с помощью метода SSA.

## Восстановление SSA

Посмотрим, как метод SSA восстанавливает тренировочную выборку для выбранных пар на рис. 5.15. На графике видно, что метод неплохо выделил сигнал. Видно, что оценка сигнала хорошо аппроксимирует временной ряд, кроме нескольких мест.

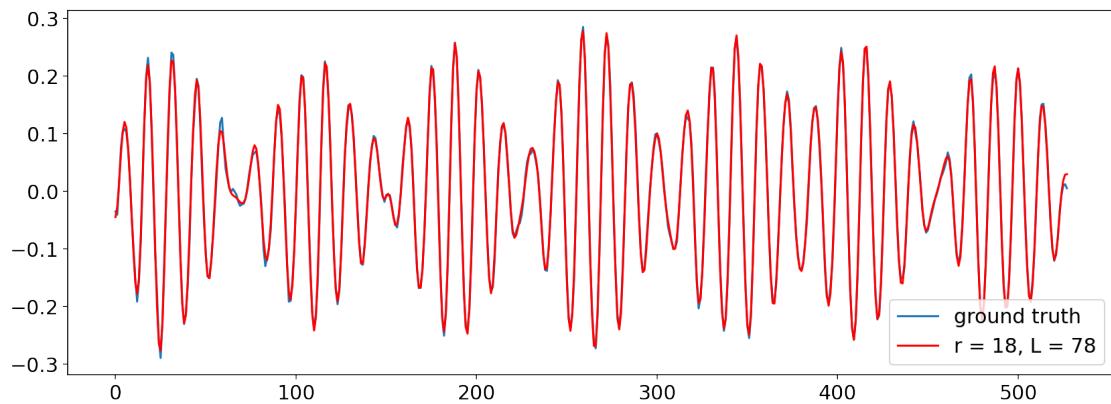


Рис. 5.15. Ряд «x pole» без тренда. Восстановление тренировочной выборки с помощью метода SSA.  $r = 18$ ,  $L = 78$

## 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  $\{13, 26, \dots, 130\}$ .

3

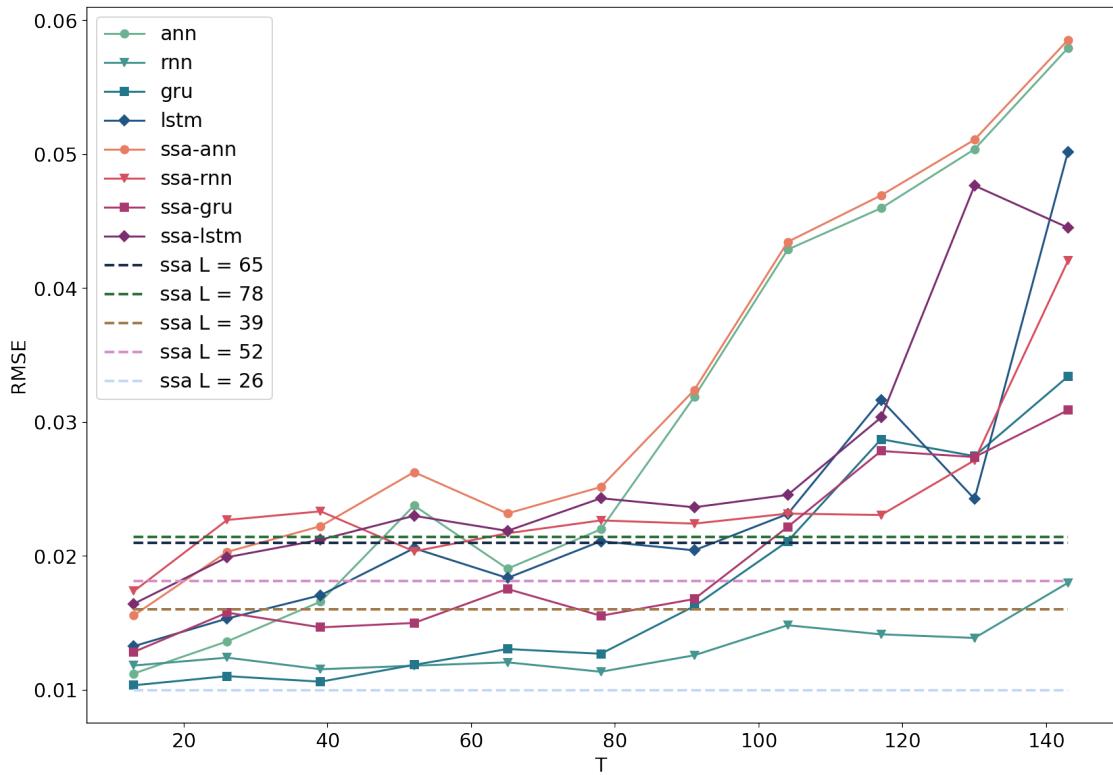


Рис. 5.16. Ряд «x pole» без тренда. Ошибки прогноза в зависимости от параметра  $T$ .

<sup>3</sup> В реальных примерах явная проблема в том, как усреднять для общей таблице, так как всё сильно зависит от  $T$ . Тут вообще усреднения по  $T$  и по  $h$  вызывают вопросы... Усреднять по  $T$  от 13 до 65 и по  $h$  от 40 до 100 и сделать пару соотв. картинок?

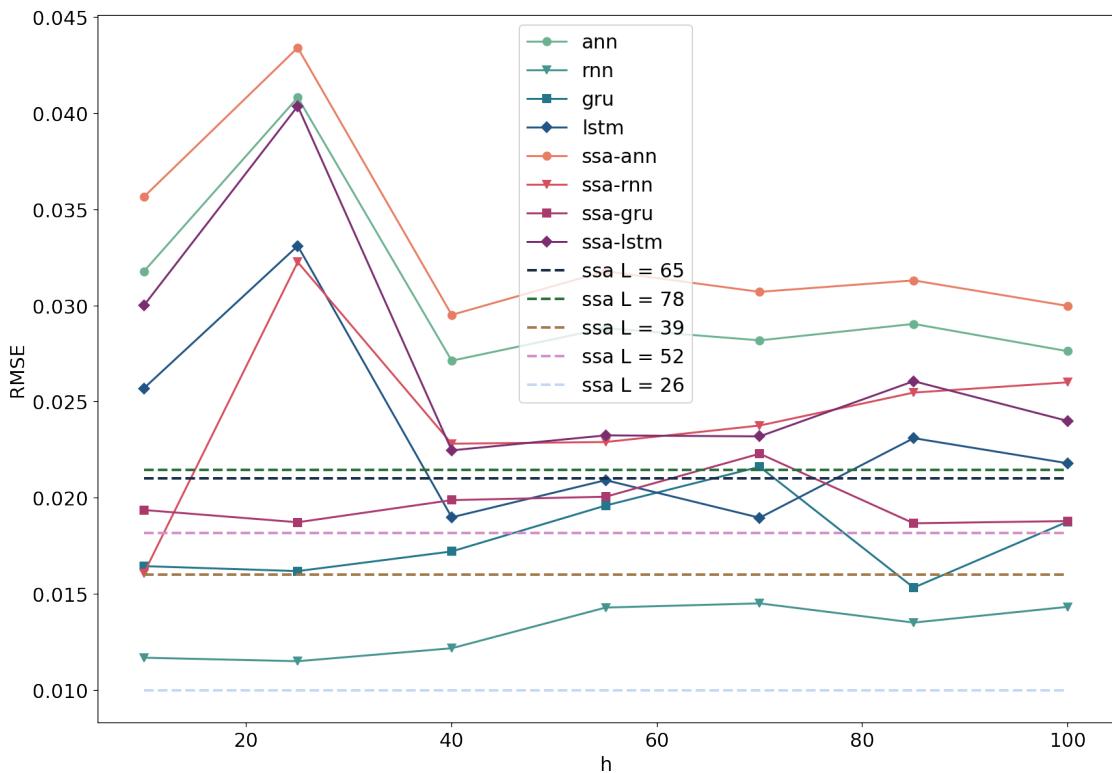


Рис. 5.17. Ряд «x pole» без тренда. Ошибки прогноза в зависимости от параметра  $h$ .

Посмотрим на результаты на рис. 5.16. На графике можно видеть, что наилучший результат показывает метод SSA с параметром  $L = 26$ . Видно, что гибридные методы сильно уступают в точности обычным методам. Методы GRU и RNN показывают хорошую точность.

### Отображение прогнозов

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 5.19. 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.

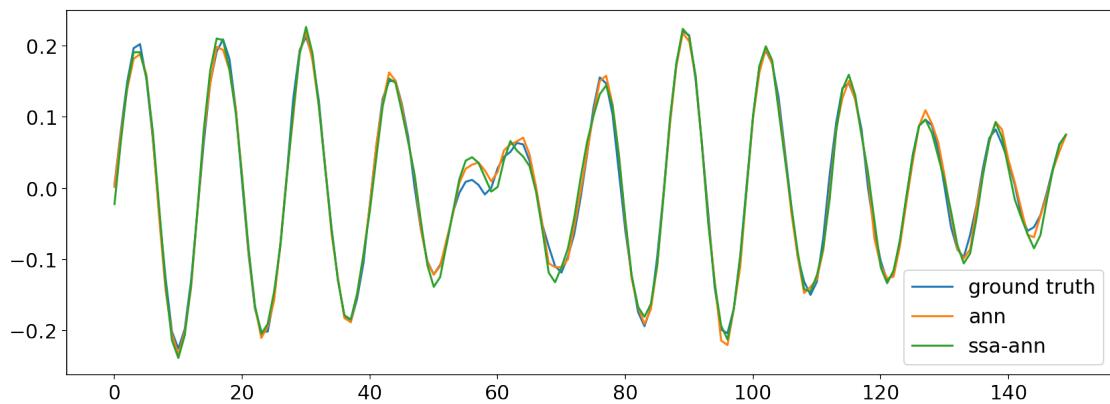


Рис. 5.18. Ряд «x pole» без тренда. Прогноз для ANN и SSA-ANN.

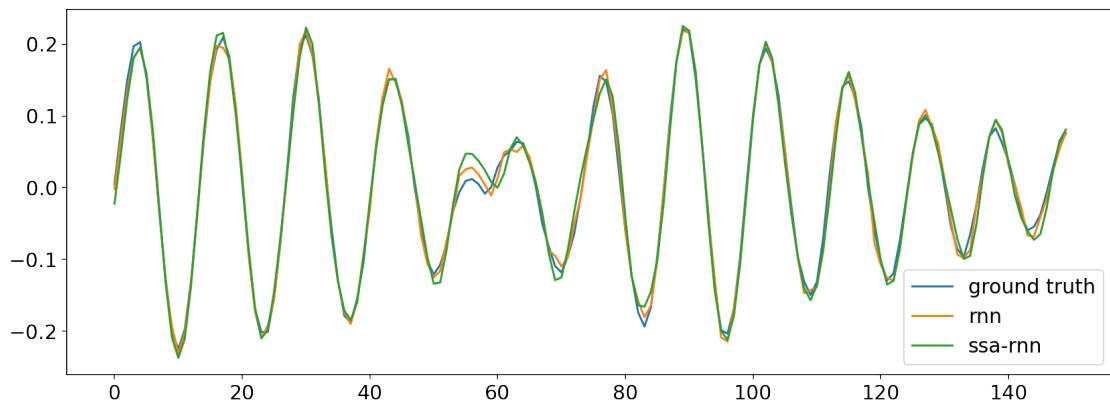


Рис. 5.19. Ряд «x pole» без тренда. Прогноз для RNN и SSA-RNN.

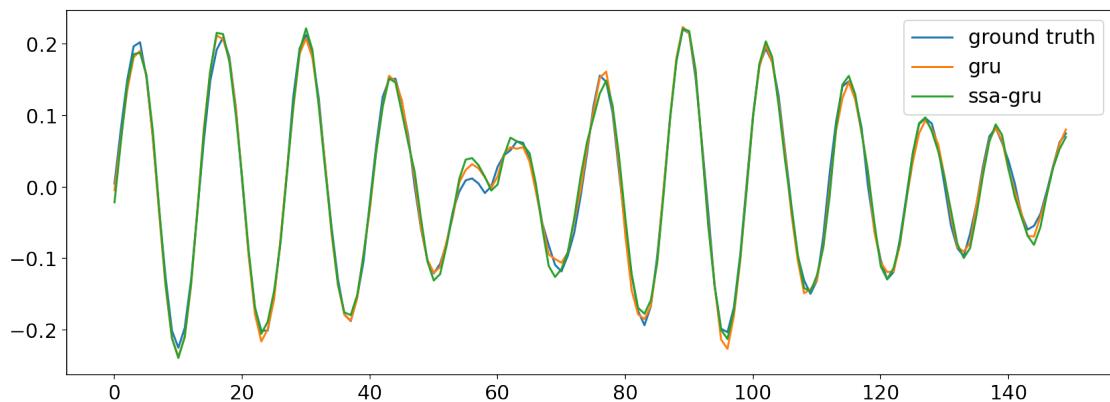


Рис. 5.20. Ряд «x pole» без тренда. Прогноз для GRU и SSA-GRU.

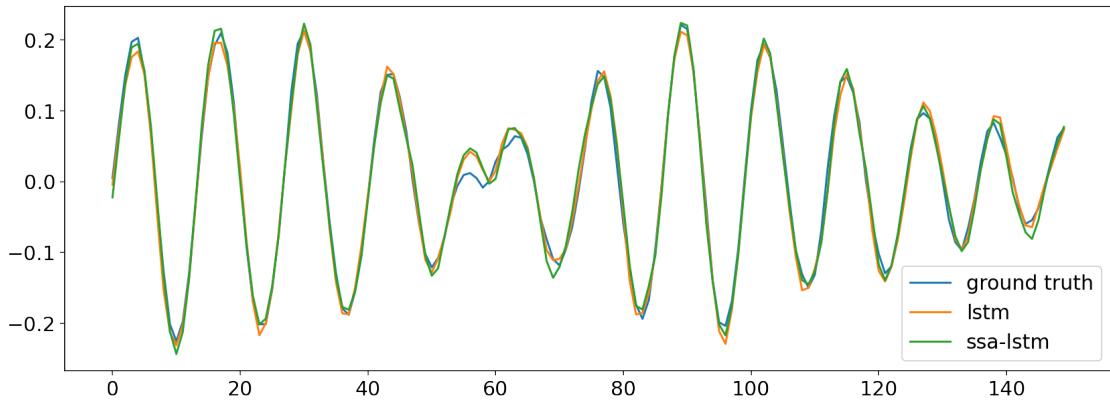


Рис. 5.21. Ряд «x pole» без тренда. Прогноз для LSTM и SSA-LSTM.

### Проверка устойчивости

4

Чтобы исключить случайность в полученных результатах, проведем сравнение для разных начальных весов методов. Зафиксируем новую сетку для параметра  $T = \{13, 91\}$ . Сетка для параметра  $h$  останется прежней. Будем получать каждый результат по 7 раз, инициализируя метод с новыми весами. Полученные результаты отображены на рис. 5.22, 5.23. На рисунках видно, что полученные ранее результаты не случайны.

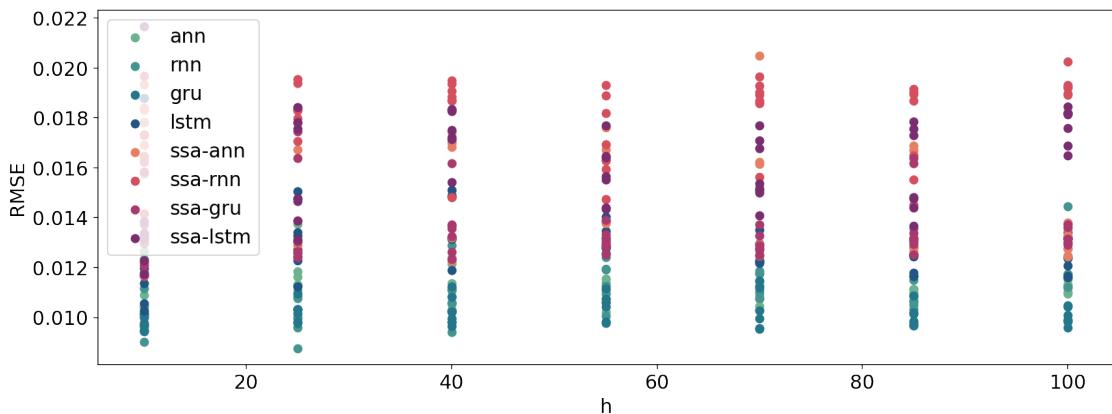


Рис. 5.22. Ряд «x pole» без тренда. Проверка устойчивости.  $T = 13$ .

<sup>4</sup> Может, проверку устойчивости везде кроме первого модельного ряда и дождей перенести в Приложение?

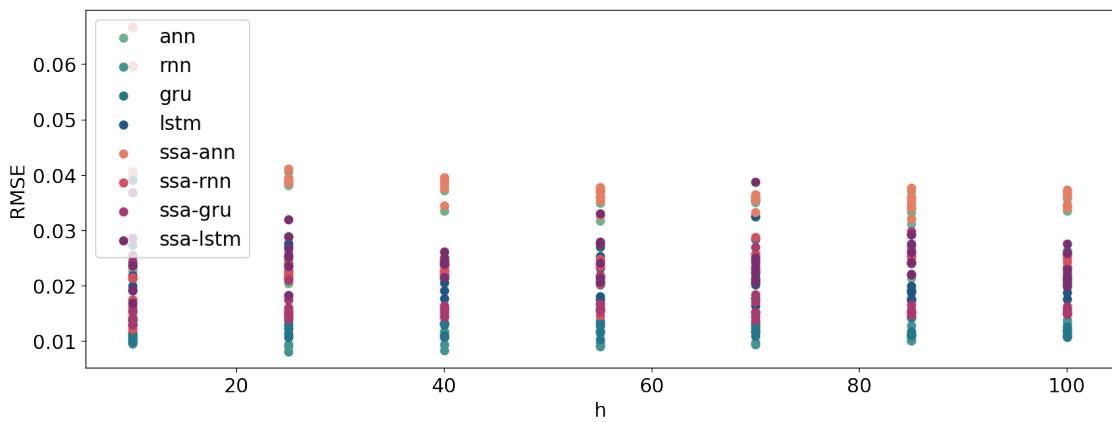


Рис. 5.23. Ряд «x pole» без тренда. Проверка устойчивости.  $T = 91$ .

## Выводы

В ходе сравнение обычных, гибридных методов и метода SSA наилучшие результаты показал метод SSA с параметрами  $L = 26, r = 12$ , что равносильно использование линейной рекуррентной формулы порядка 25 для прогноза. Использование таких параметров в гибридных моделях нецелесообразно, так как оценка сигнала слишком сильно аппроксимирует ряд. Гибридные методы показали результаты хуже, чем обычные методы. Когда там пишем, то имеется в виду, что, конечно, гибридные методы выбором большого  $r$  можно сделать практически эквивалентными негибридным, поэтому слово «хуже» подразумевает использование небольшого  $r$ .

Причина таких результатов может быть в том, что сигнал ряда скорее всего неконечного ранга, что делает корректное выделение сигнала затруднительным. Ошибки на картинках с прогнозами регулярно происходят в местах переходов между биениями, что говорит о том, что SSA некорректно выделяет сигнал, а это, возможно, приводит гибридные методы к плохим результатам. Также проблема может возникнуть в небольшой длине ряда, что может плохо сказаться на выделении сигнала методом SSA и обучении нейронных сетей (дефицит обучающих пар). Шум ряда маленький, что ставит под сомнение использование препроцессинга SSA. Как было показано в эксперименте 4.1.2 в случае маленького шума, логично выбрать параметр  $r$  соответствующим рангу сигнала или больше. В этом случае мы не можем выделить оптимальный параметр аналитический, что заставляет искать параметр  $r$  перебором, что в свою очередь сподвигает выбрать параметр  $r$  побольше. Недостаточно большой параметр  $r$  с большой вероятностью приведет к некорректному выделению сигнала, так как все компоненты

сложного сигнала могут не попасть в оценку. Большой параметр  $r$  должен захватить весь сигнал, но ввиду того что сигнал ряда неконечного ранга, параметр  $L$  должен быть выбран относительно малым. Большой параметр  $r$  и малый параметр  $L$  приводят нас к сильной аппроксимации ряда оценкой сигнала, что делает препроцессинг SSA бессмысленным.

### 5.3. Погода

Рассмотрим следующий ряд  $Z_{828}$  (рис. 5.24). Данный ряд отображает одну из характеристик погоды в городе Санкт-Петербурге.

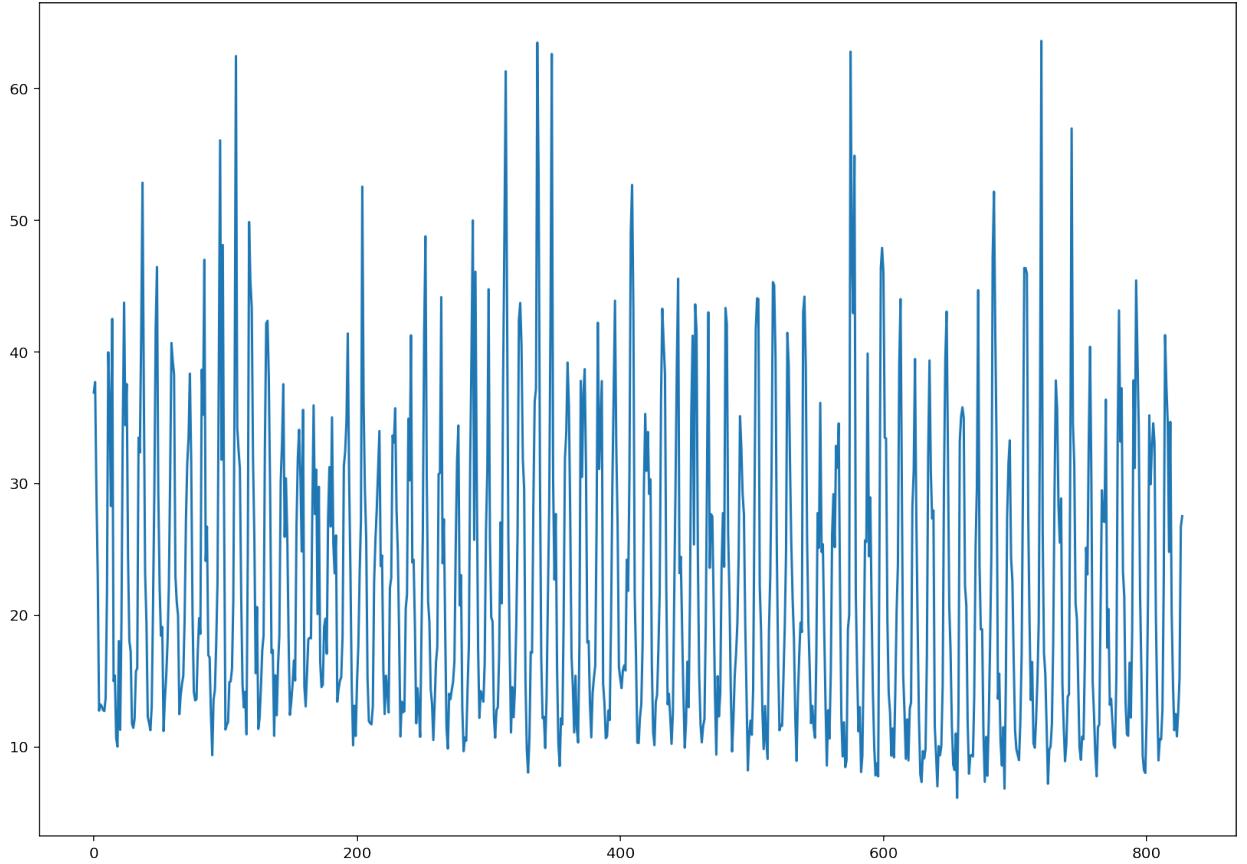


Рис. 5.24. Временной ряд характеристики погоды в Санкт-Петербурге.

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

На рис. 5.25 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 равен 12. Далее в экспериментах будем перебирать параметры  $T$  и  $L$  по сетке с шагом кратным 12.

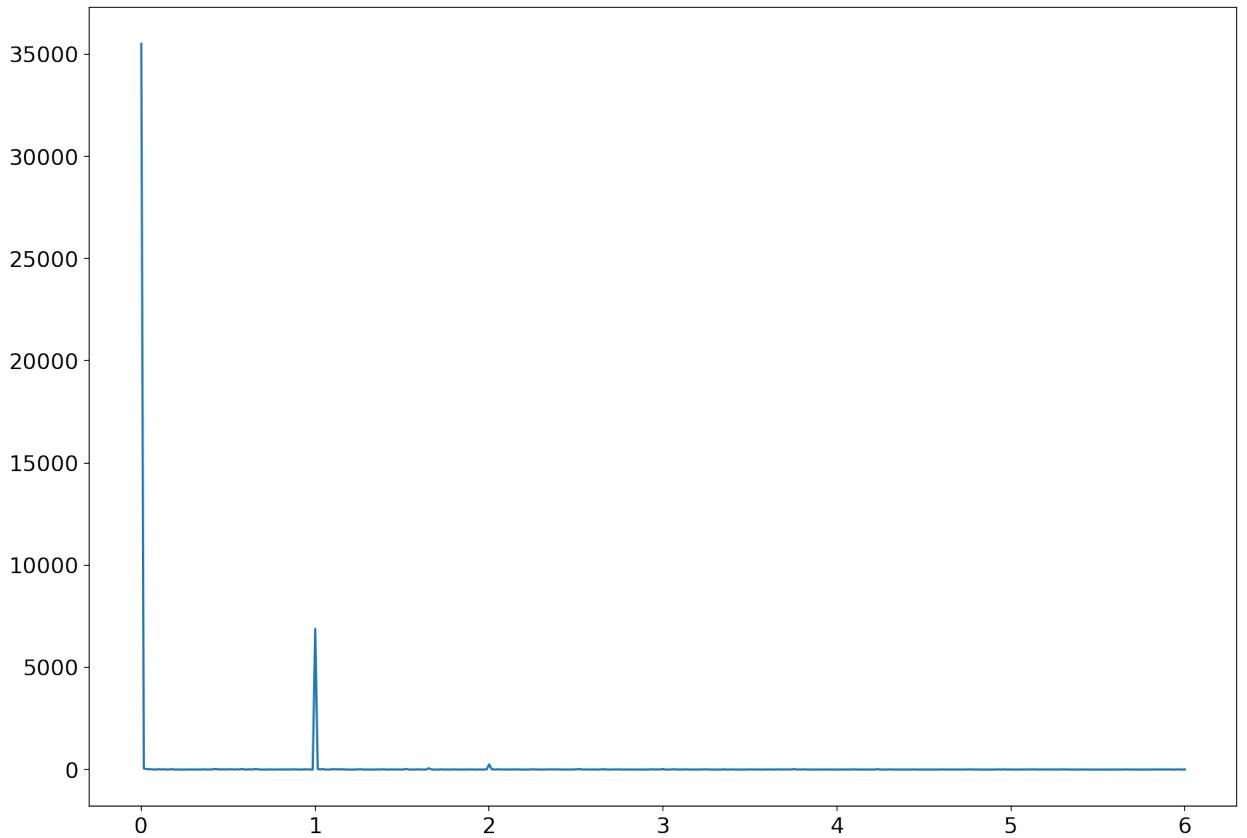


Рис. 5.25. Периодограмма ряда характеристики погоды.

### 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 5.26. На графике видно, что наилучшие результаты достигаются при  $r = 5$ . Нет сильной разницы в точности при  $r = 5$ , поэтому дальше будем рассматривать пару  $r = 5$ ,  $L = 264$ .

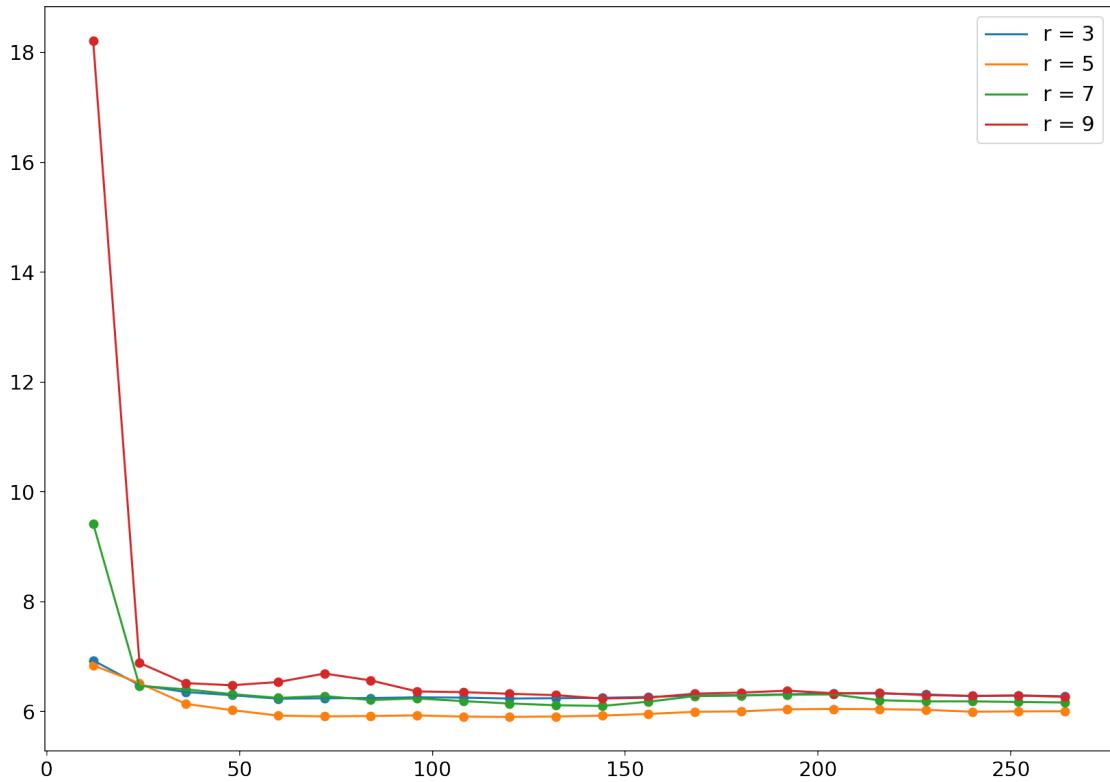


Рис. 5.26. Погода в Санкт-Петербурге. Прогнозирование с помощью метода SSA.

### Восстановление SSA

Посмотрим, как метод SSA восстанавливает тренировочную выборку для выбранных пар на рис. 5.27. На графике видно, что метод неплохо выделил сигнал и оценка сигнала не слишком сильно аппроксимирует временной ряд.

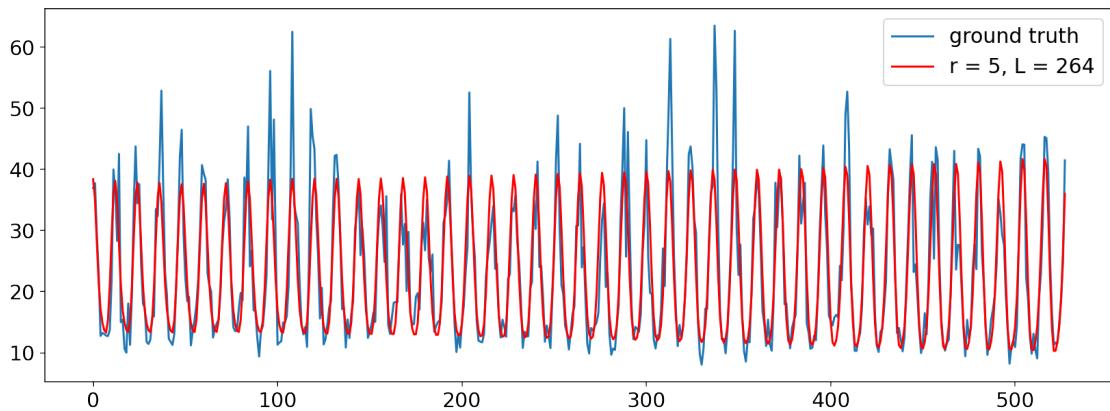


Рис. 5.27. Погода в Санкт-Петербурге. Восстановление тренировочной выборки с помощью метода SSA.  $r = 5$ ,  $L = 264$

## Сравнение методов

Для нейронных сетей зададим следующую сетку 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  $\{12, 24, \dots, 264\}$ .<sup>5</sup>

Посмотрим на результаты на рис. 5.28, 5.29, 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. Тем не менее, в некоторых частных случаях гибридные методы показывают хорошую точность. Например SSA-ANN на рис. 5.28 и SSA-ANN, SSA-RNN на рис. 5.29. Такие результаты могли получиться из-за маленького количества данных. А также возможно малого размера валидационной выборки, что ограничивает выбор параметра  $T$ .

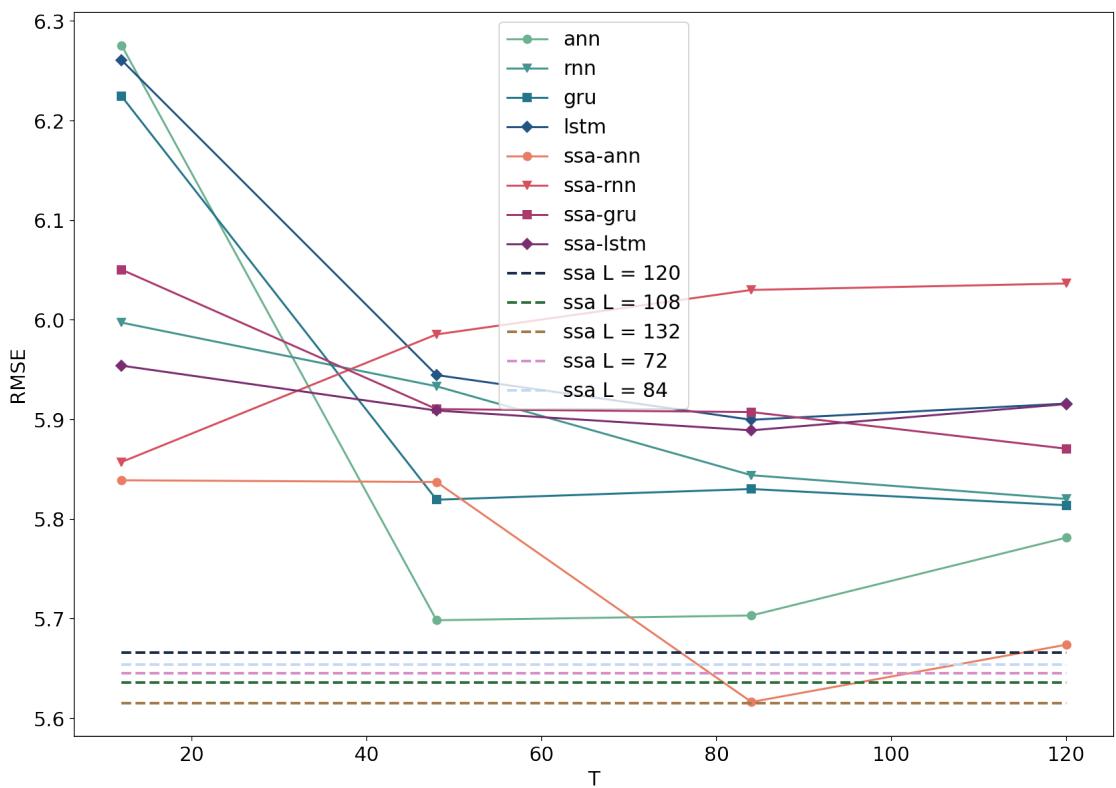


Рис. 5.28. Погода в Санкт-Петербурге. Ошибки прогноза в зависимости от параметра  $T$ .

<sup>5</sup> А какая логика в том, что для SSA вы фиксируете  $r$ ?

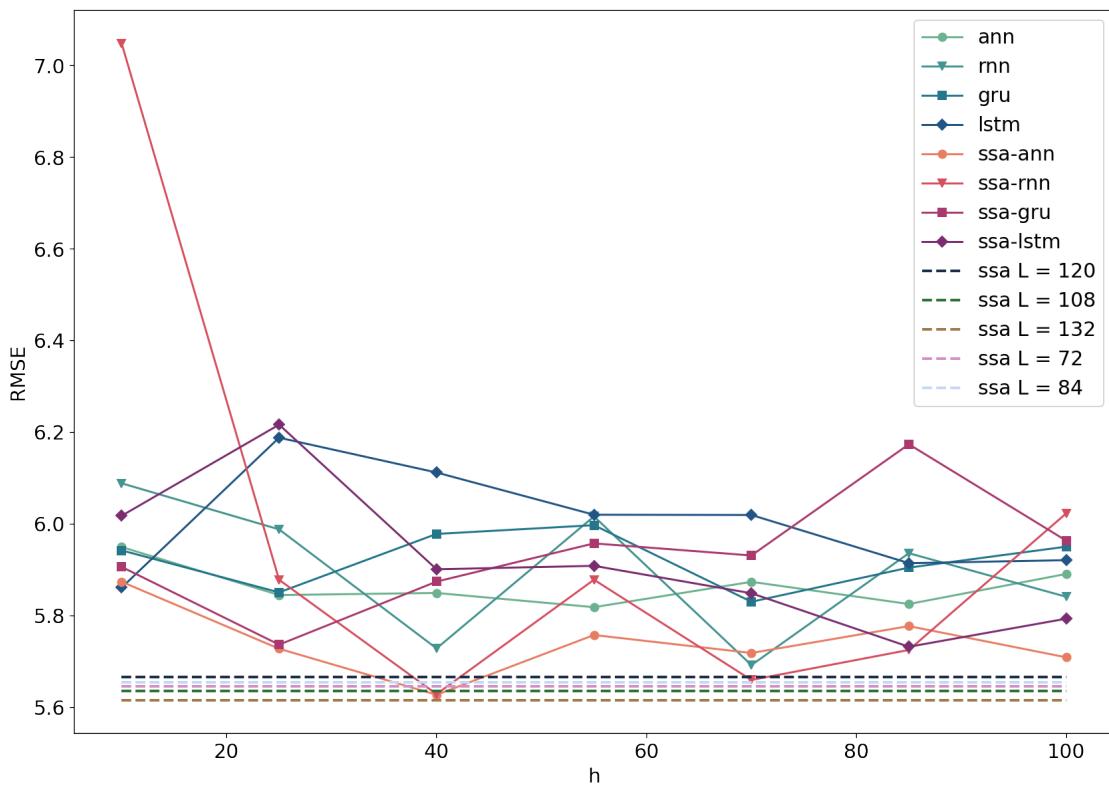


Рис. 5.29. Погода в Санкт-Петербурге. Ошибки прогноза в зависимости от параметра  $h$ .

### Отображение прогнозов

На графиках Б.5—Б.8 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.<sup>6</sup>

### 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 параметра  $h$  останется прежней. Будем получать каждый результат по 7 раз, инициализируя метод с новыми весами. Полученные результаты отражены на рисунках 5.30, 5.31. На них подтверждается, выводы сделанные ранее. Заключаем, что полученные результаты устойчивые.

<sup>6</sup> Вот это «видно» и «заметно» совсем не видно и не заметно, т.е. непонятно, что имеется в виду не субъективное

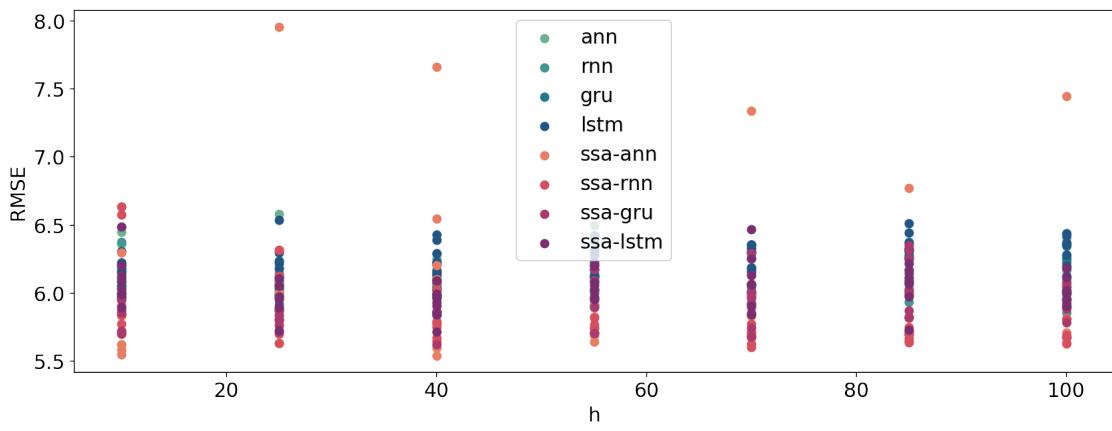


Рис. 5.30. Погода в Санкт-Петербурге. Проверка устойчивости.  $T = 12$ .

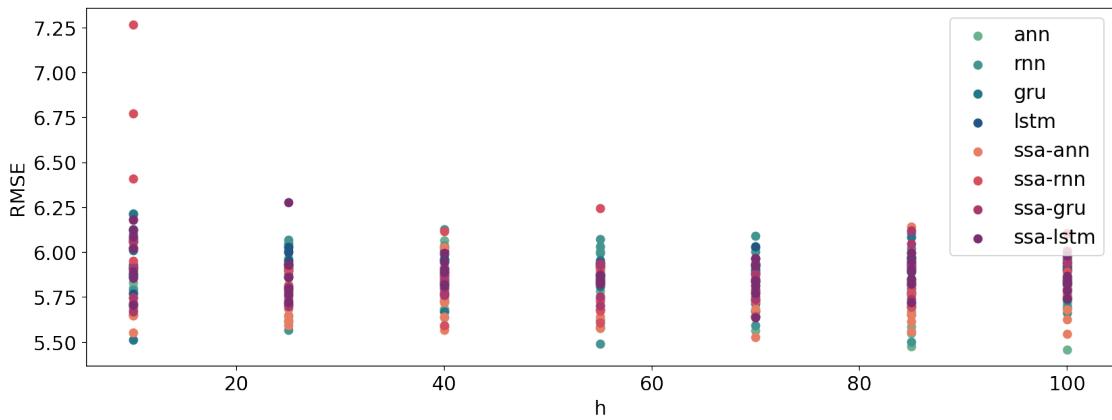


Рис. 5.31. Погода в Санкт-Петербурге. Проверка устойчивости.  $T = 84$ .

## Выводы

Сравнение показало, что гибридные методы не дают сильной прироста в точности на данных погоды. Но все же в некоторых частных случаях получается достичь преимущества над обычными методами.

Данный ряд похож на ряд «Indian Rain», но с сигналом немного сложнее. Сигнал ряда погоды конечного ранга, в ряде присутствует шум. Аналитические выбранные параметры для метода SSA совпадают с выбранными перебором. Ряды погоды и «Indian Rain» очень похожи, но получить успех «Indian Rain» не получилось. Единственное их различие это длина ряда. Это приводит к выводу, что 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 для использования препроцессинга SSA важно наличие трех факторов:

1. Наличие в ряде шума.
2. Длина ряда.
3. Метод SSA может корректно выделить сигнал (получить оценку сигнала).
4. Какова цель прогноза, прогноз сигнала или прогноз всего ряда.

По первому пункту, цель препроцессинга SSA в задачи прогнозирования временных рядов заключается в очистке ряда от шума, так если первый пункт не выполнен, не понятно зачем использовать препроцессинг SSA. Если шум в ряде маленький, то препроцессинг SSA все еще можно использовать, если SSA может корректно получить оценку сигнала. В этом случае лучше выбрать параметр  $r$  оптимальный или больше, чтобы точно захватить сигнал. Даже если шум попадет в оценку, если он небольшой, то его влияние будет незначительным. В случае значительного шума, стоит выбрать параметр  $r$  оптимальный или меньше, чтобы шум не искажал оценку сигнала.

Немаловажную роль играет и длина ряда, так на достаточно длинных рядах удается лучше выделить сигнал методом SSA. Также параметры нейронных сетей лучше оптимизируются во время процесса обучения, что положительно сказывается на точности предсказаний.

По третьему пункту, если метод SSA не может корректно выделить сигнал, это приведет использование гибридных методов к некачественным прогнозам. В этом случае стоит отказаться от использования препроцессинга SSA.

Так на основе изложенных выше информации, можно сказать, что использование препроцессинга на данных ЕОР является бессмысленным. Данные «Indian Rain»

в свою очередь, являются идеальным экземпляром для использования гибридных методов. Данные погоды в Санкт-Петербурге можно рассматривать, как данные «Indian Rain», но меньшей длины, что подчеркивает важность длины временных рядов.

## 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. Analysis of Time Series Structure - SSA and Related Techniques. — Boca Raton, FL : Chapman and Hall/CRC, 2001.

## Приложение А

### Модельные данные

#### A.1. Сумма синусов с белым шумом

##### A.1.1. Влияние $r$

В этом разделе приложены графики имеющие отношения к эксперименту из главы 4.1.1.

#### Сравнение обычных и гибридных методов

На графиках ниже показаны результаты сравнение обычных и гибридных методов, а также метода SSA. Графики являются приложением к разделу 4.1.1.

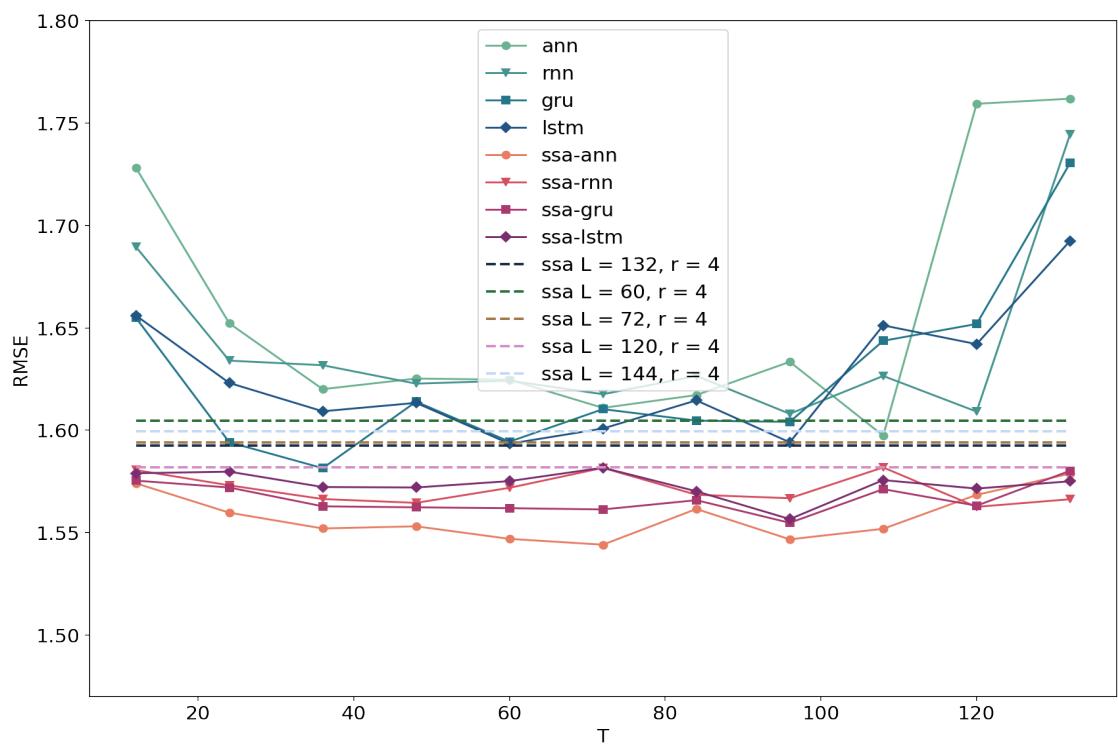


Рис. A.1. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Ошибки прогноза в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 2$ .

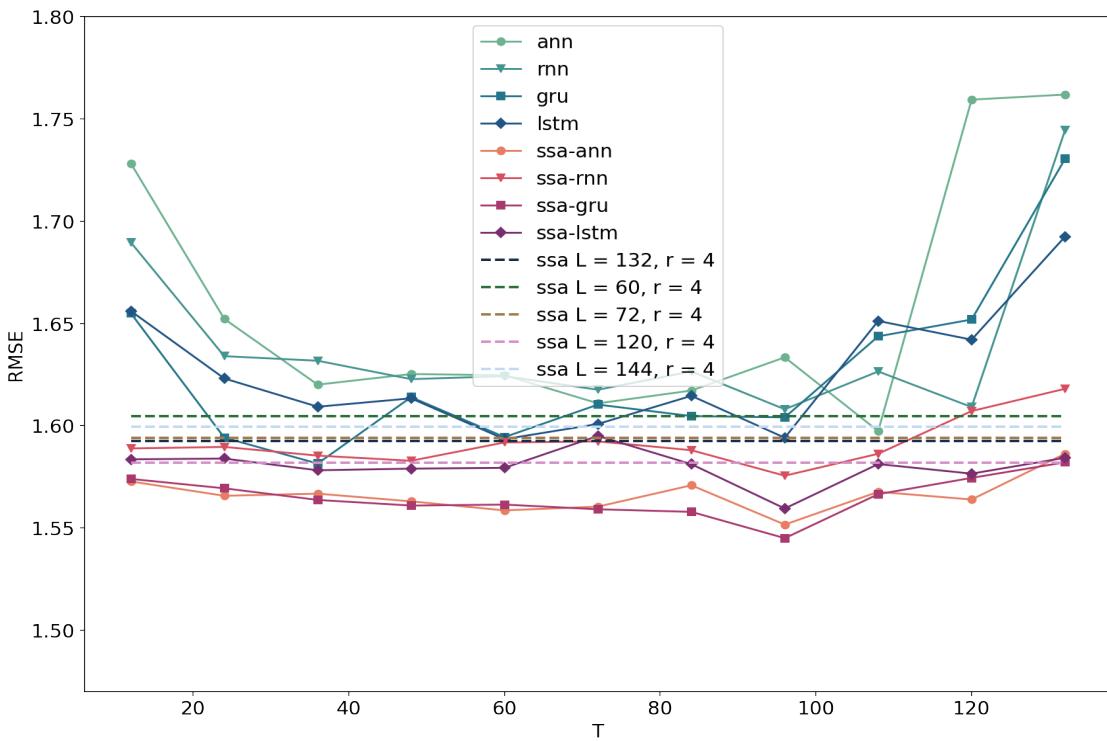


Рис. А.2. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Ошибки прогноза в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 4$ .

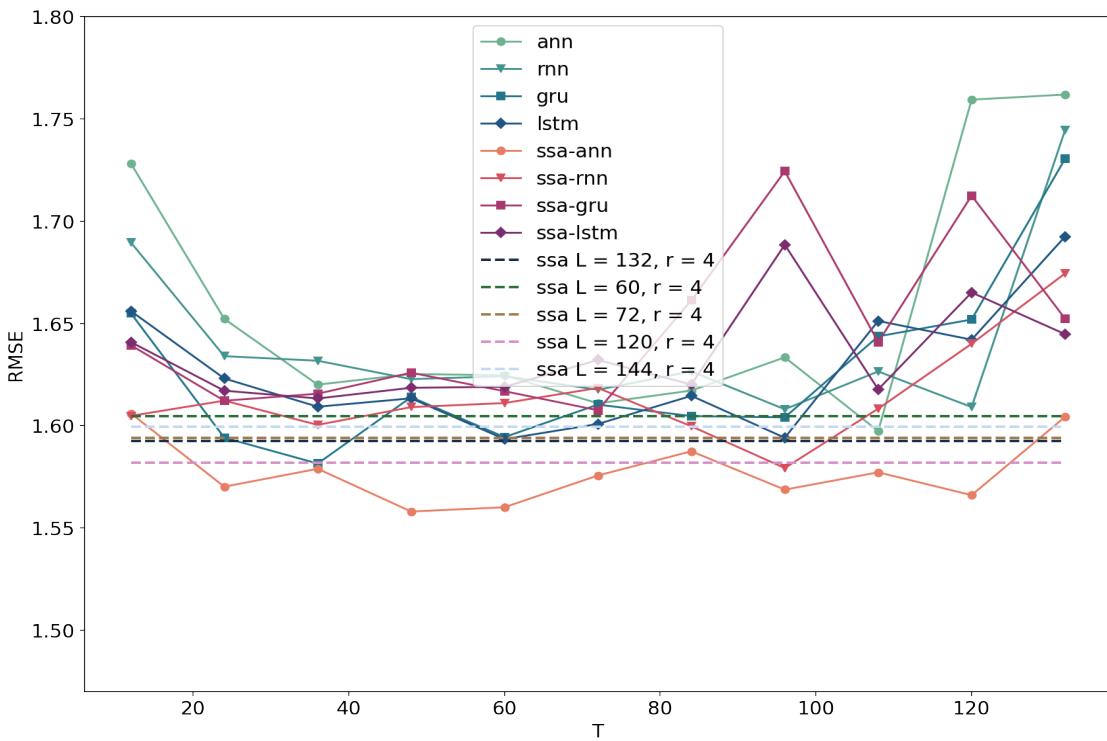


Рис. А.3. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Ошибки прогноза в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 6$ .

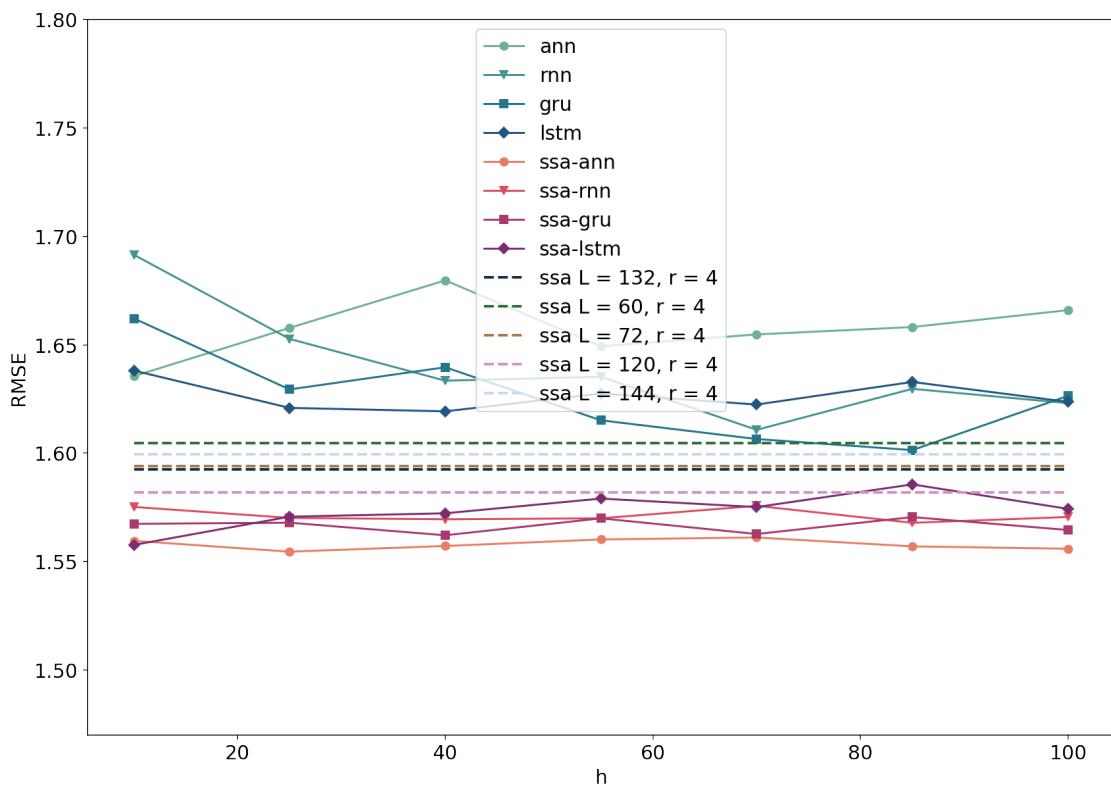


Рис. A.4. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Ошибки прогноза в зависимости от параметра  $h$ .  $L = 175$ ,  $r = 2$ .

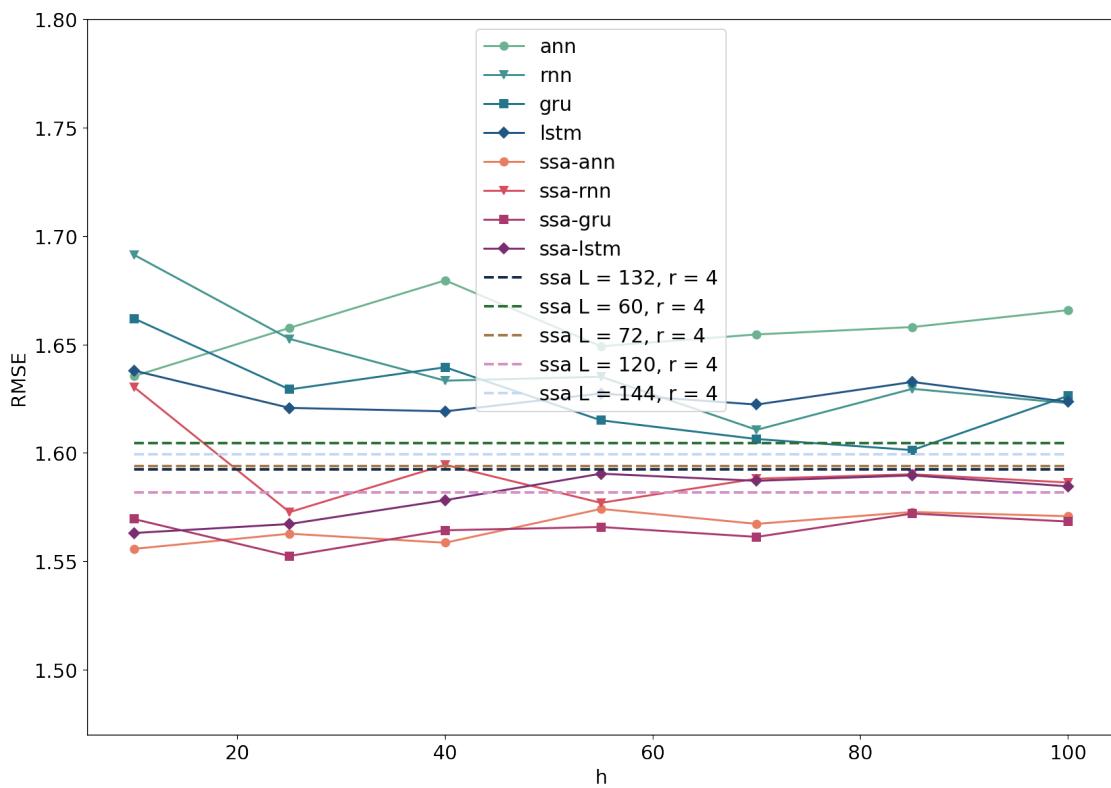


Рис. A.5. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Ошибки прогноза в зависимости от параметра  $h$ .  $L = 175$ ,  $r = 4$ .

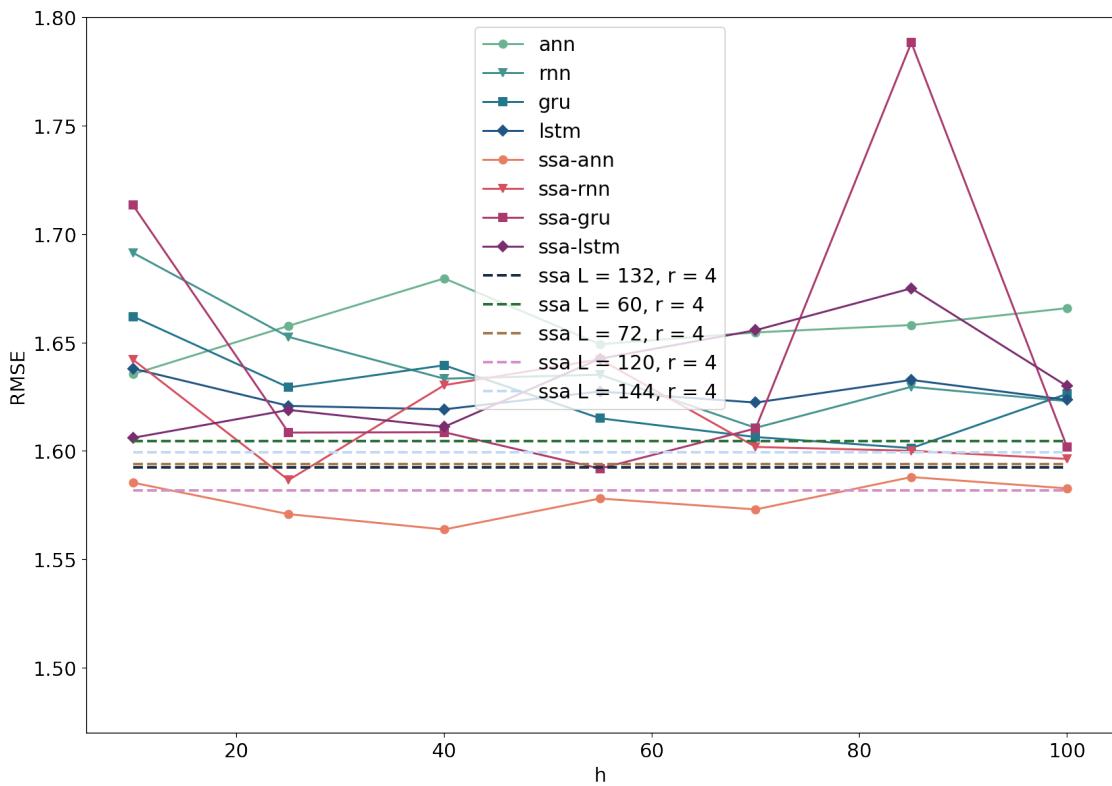


Рис. A.6. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Ошибки прогноза в зависимости от параметра  $h$ .  $L = 175$ ,  $r = 6$ .

### Отображение прогнозов

На графиках ниже показаны результаты прогнозирования тестовой выборки обычных и гибридных методов. Графики являются приложением к разделу 4.1.1.

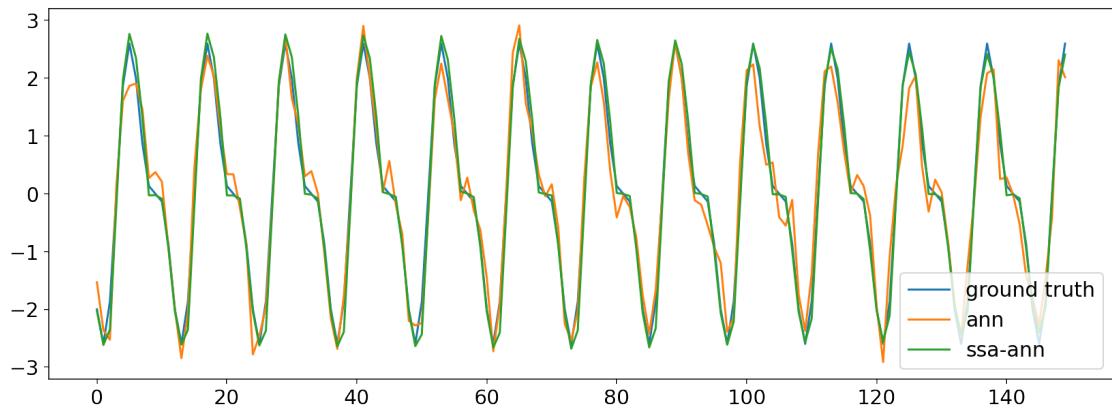


Рис. A.7. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Прогноз для ANN и SSA-ANN.  $r = 2$

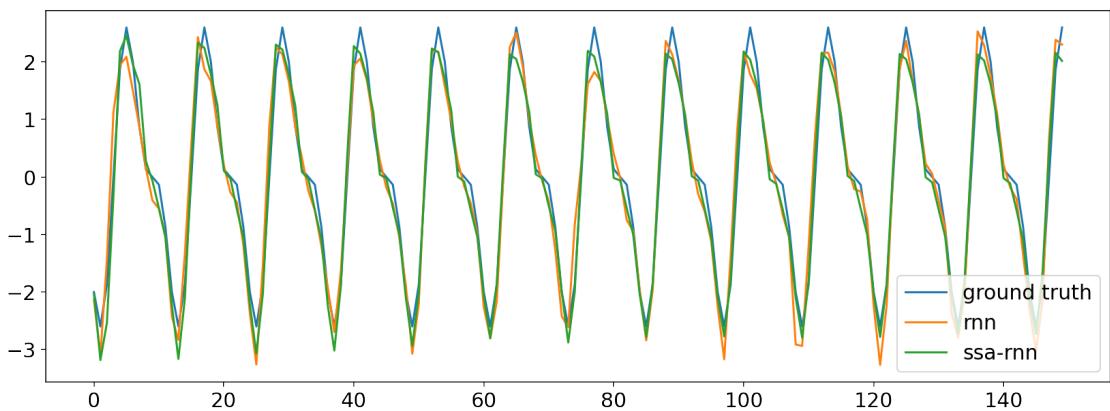


Рис. А.8. «Сумма синусов с белым шумом». Ряд Z<sub>650</sub>. Прогноз для RNN и SSA-RNN.  $r = 2$

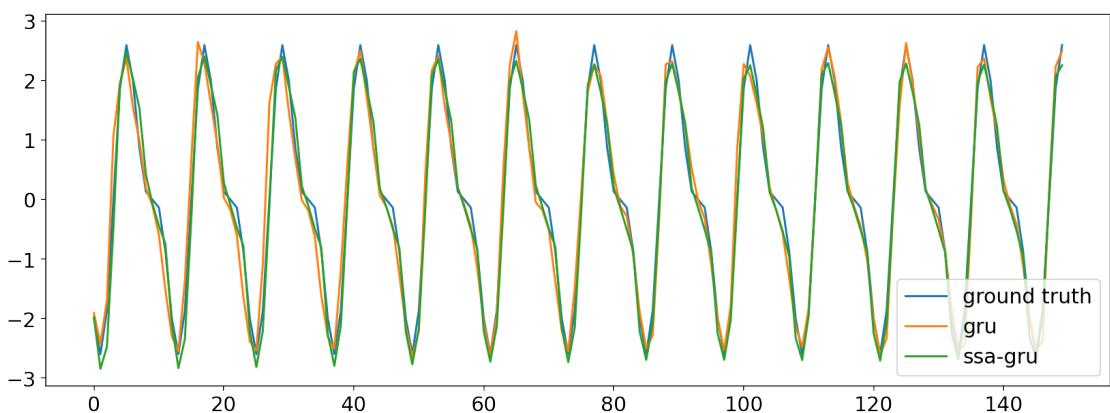


Рис. А.9. «Сумма синусов с белым шумом». Ряд Z<sub>650</sub>. Прогноз для GRU и SSA-GRU.  $r = 2$

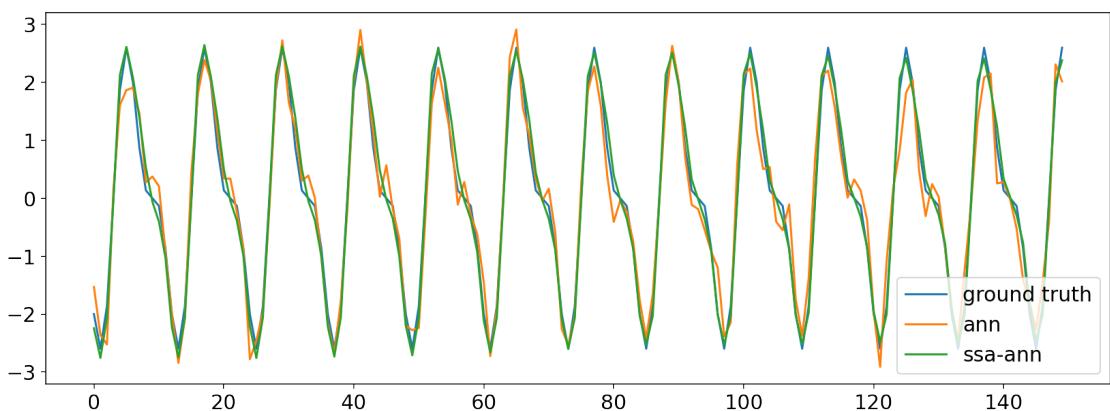


Рис. А.10. «Сумма синусов с белым шумом». Ряд Z<sub>650</sub>. Прогноз для ANN и SSA-ANN.  $r = 4$

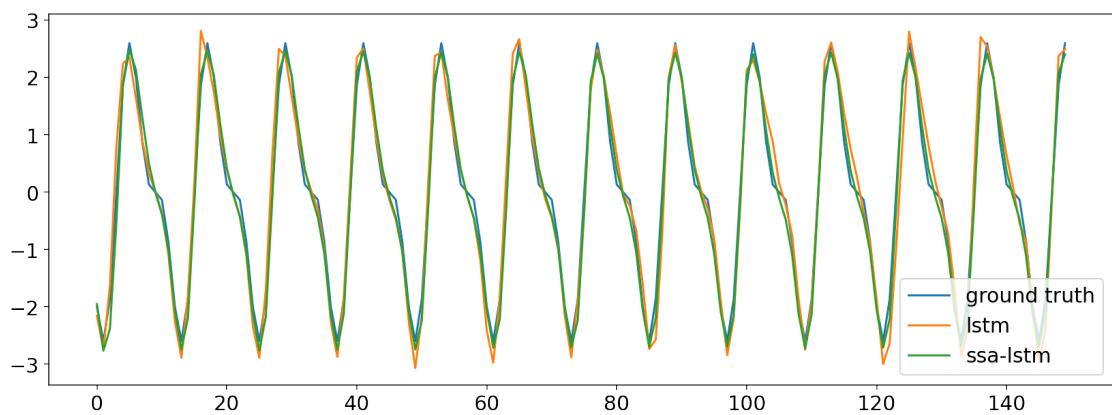


Рис. А.11. «Сумма синусов с белым шумом». Ряд Z<sub>650</sub>. Прогноз для LSTM и SSA-LSTM.  $r = 2$

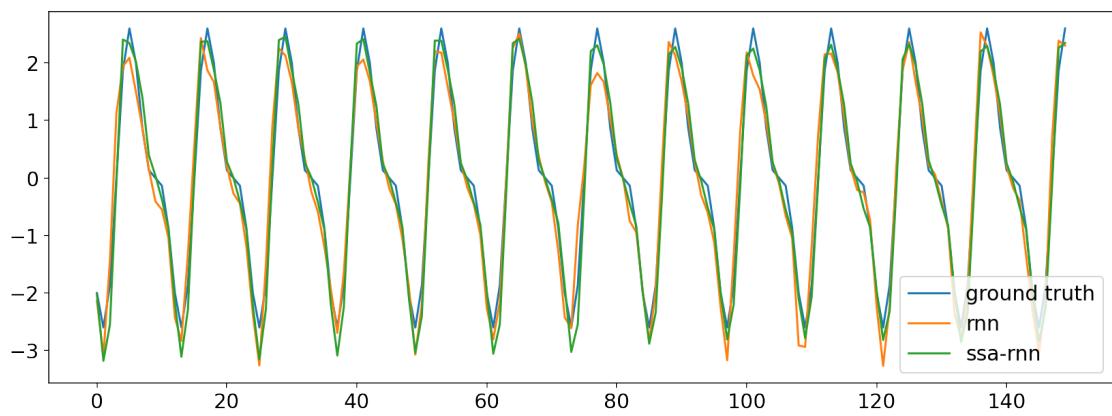


Рис. А.12. «Сумма синусов с белым шумом». Ряд Z<sub>650</sub>. Прогноз для RNN и SSA-RNN.  $r = 4$

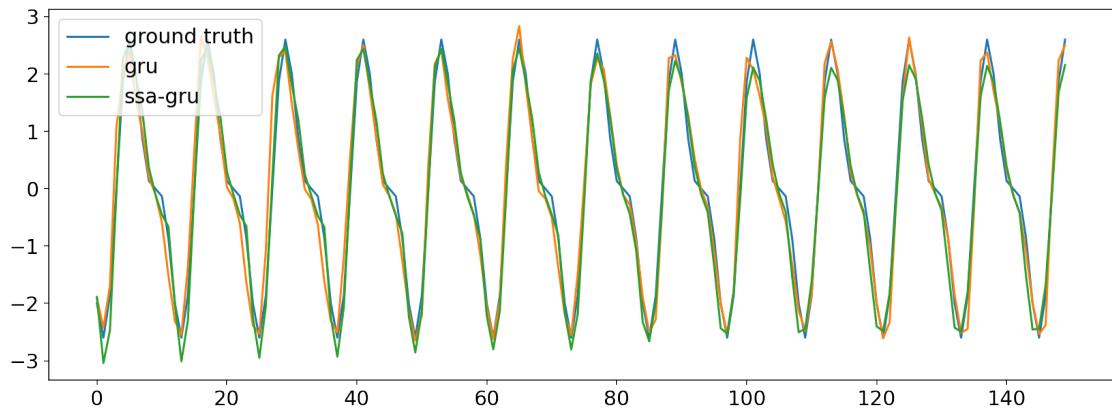


Рис. А.13. «Сумма синусов с белым шумом». Ряд Z<sub>650</sub>. Прогноз для GRU и SSA-GRU.  $r = 4$

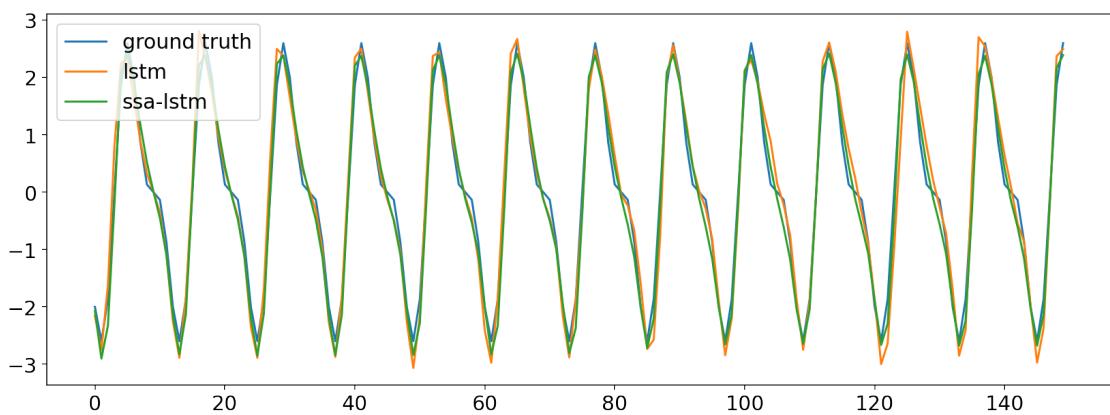


Рис. А.14. «Сумма синусов с белым шумом». Ряд Z<sub>650</sub>. Прогноз для LSTM и SSA-LSTM.  $r = 4$

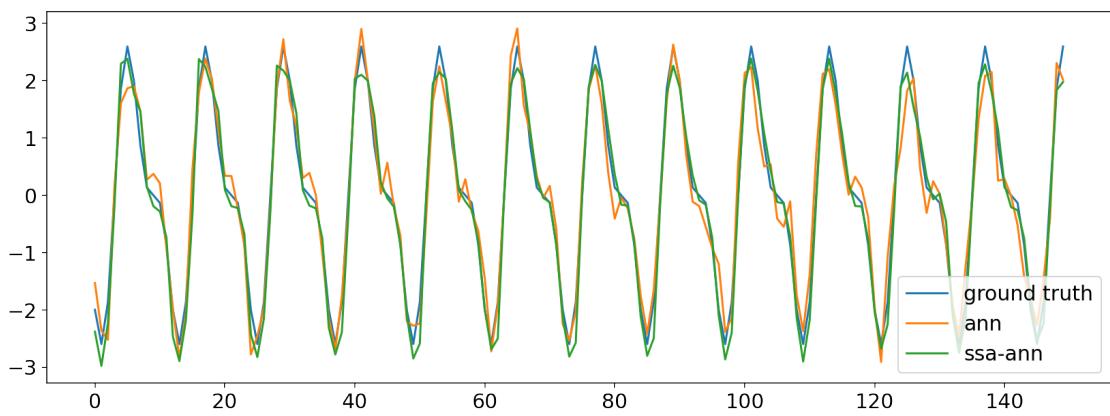


Рис. А.15. «Сумма синусов с белым шумом». Ряд Z<sub>650</sub>. Прогноз для ANN и SSA-ANN.  $r = 6$

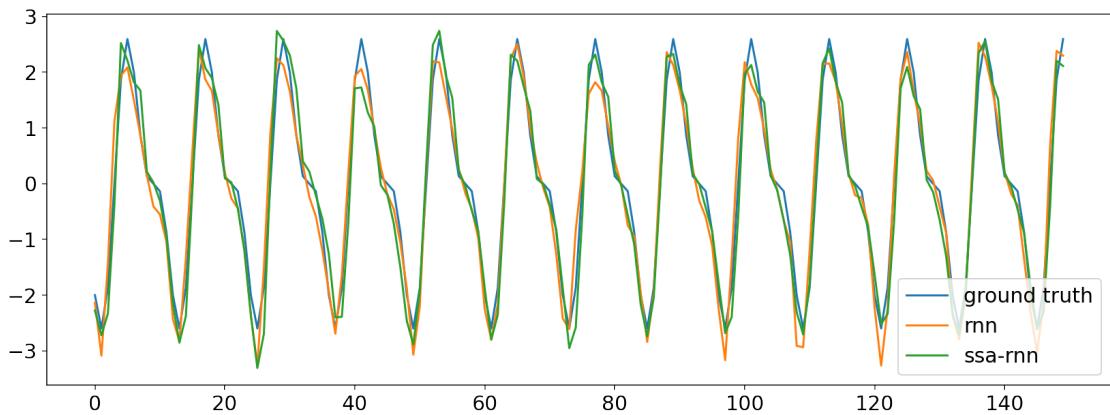


Рис. А.16. «Сумма синусов с белым шумом». Ряд Z<sub>650</sub>. Прогноз для RNN и SSA-RNN.  $r = 6$

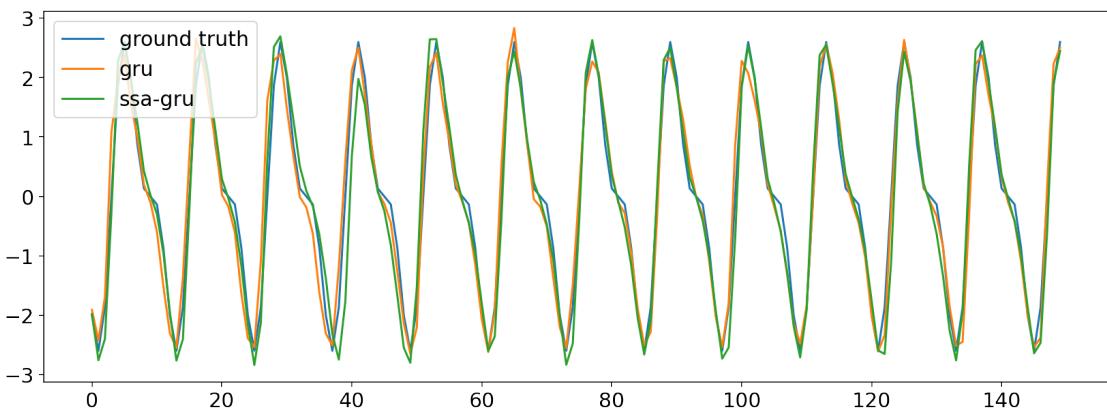


Рис. A.17. «Сумма синусов с белым шумом». Ряд Z<sub>650</sub>. Прогноз для GRU и SSA-GRU. r = 6

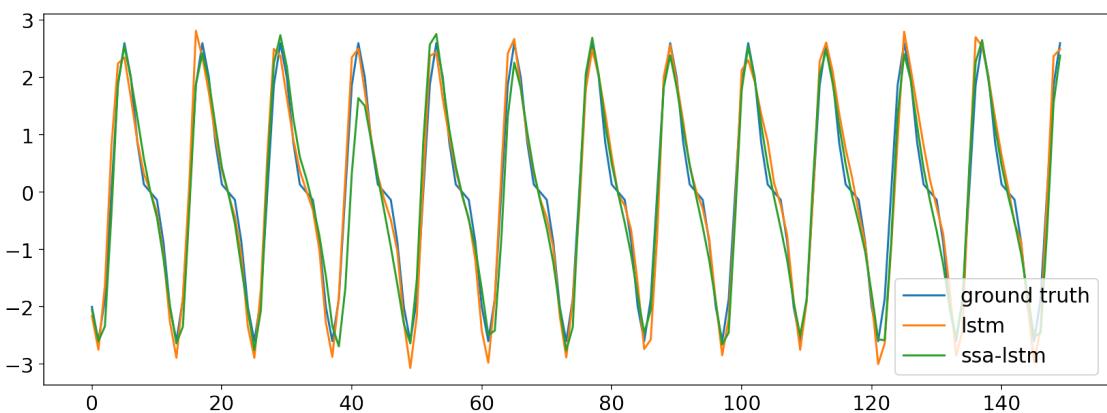


Рис. A.18. «Сумма синусов с белым шумом». Ряд Z<sub>650</sub>. Прогноз для LSTM и SSA-LSTM. r = 6

## Проверка устойчивости

На графиках проверяются устойчивость результатов по методике из раздела 3.6. Графики являются приложением к разделу 4.1.1.

### A.1.2. Влияние r при маленьком шуме

В этом разделе приложены графики имеющие отношения к эксперименту из главы 4.1.2.

## Сравнение обычных и гибридных методов

На графиках ниже показаны результаты сравнение обычных и гибридных методов, а также метода SSA. Графики являются приложением к разделу 4.1.2.

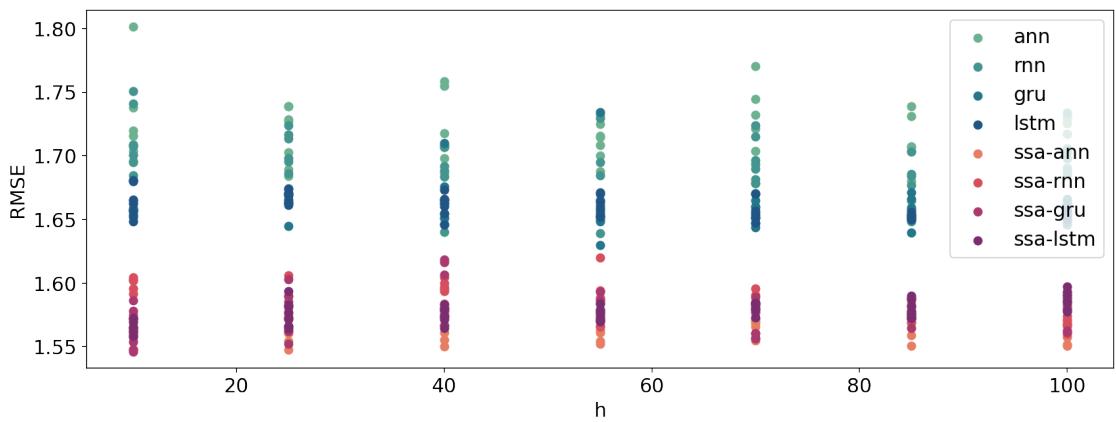


Рис. A.19. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$r = 2, L = 175, T = 12$ .

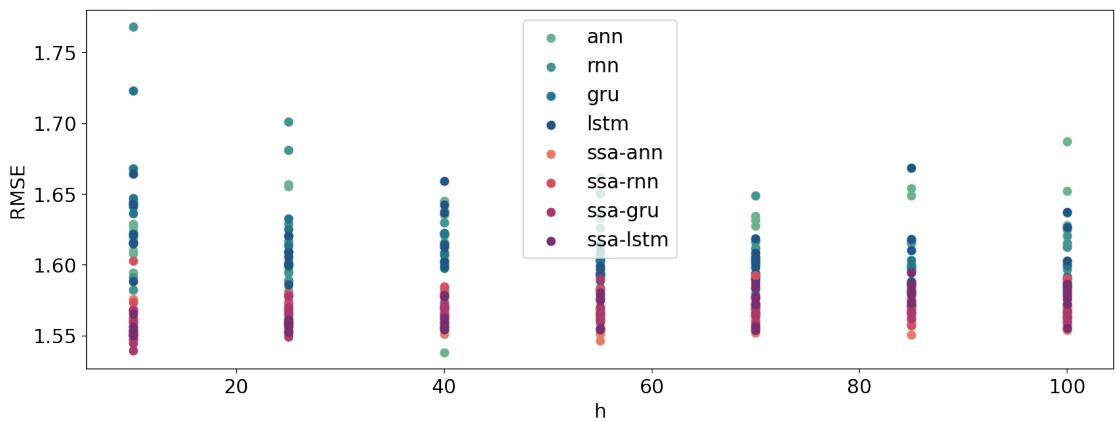


Рис. A.20. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$r = 2, L = 175, T = 84$ .

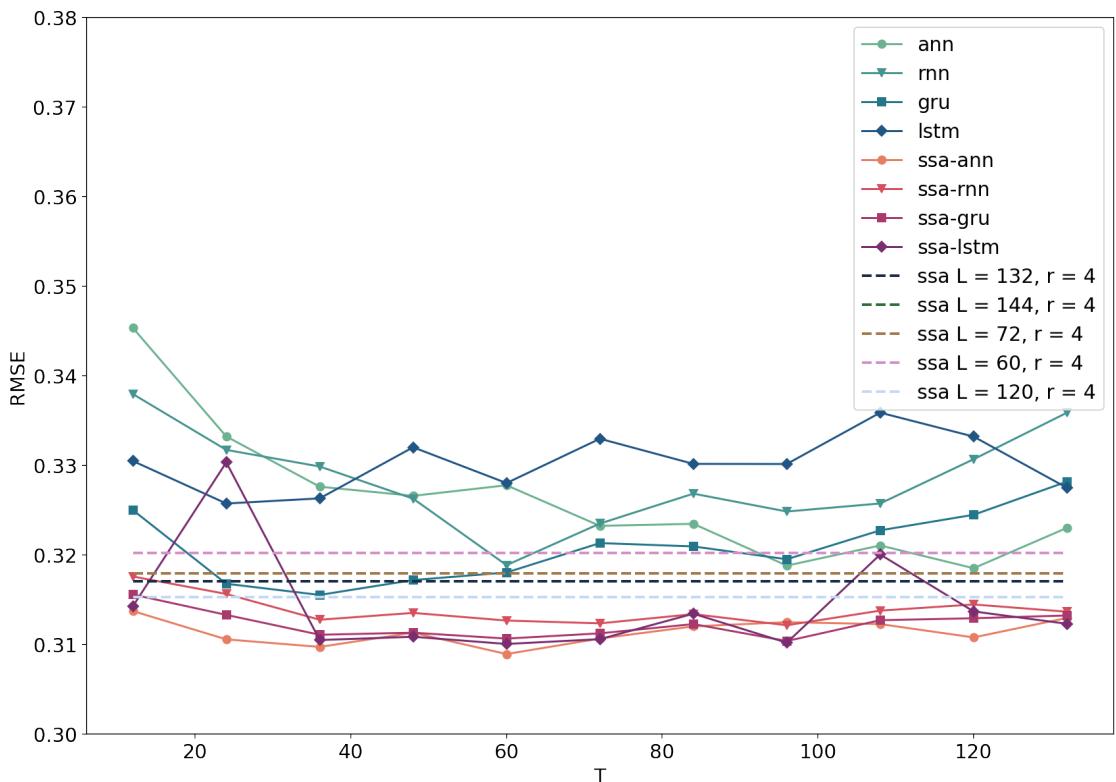
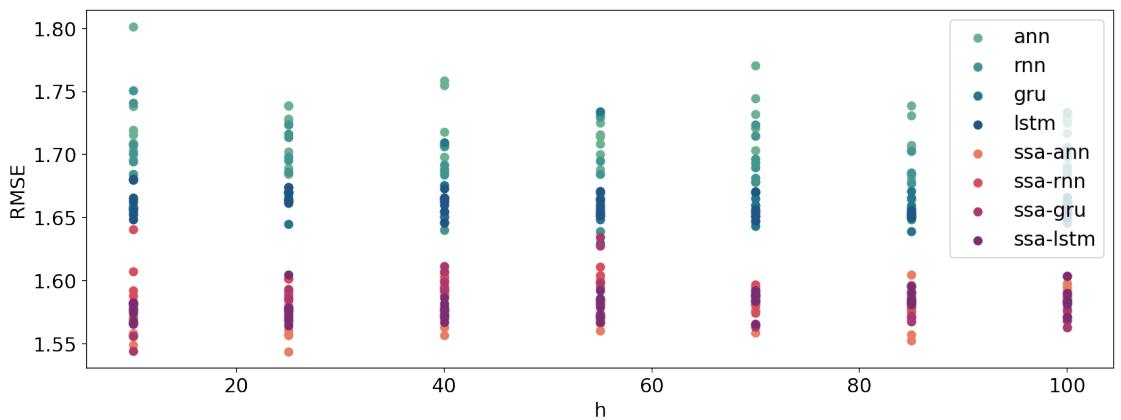
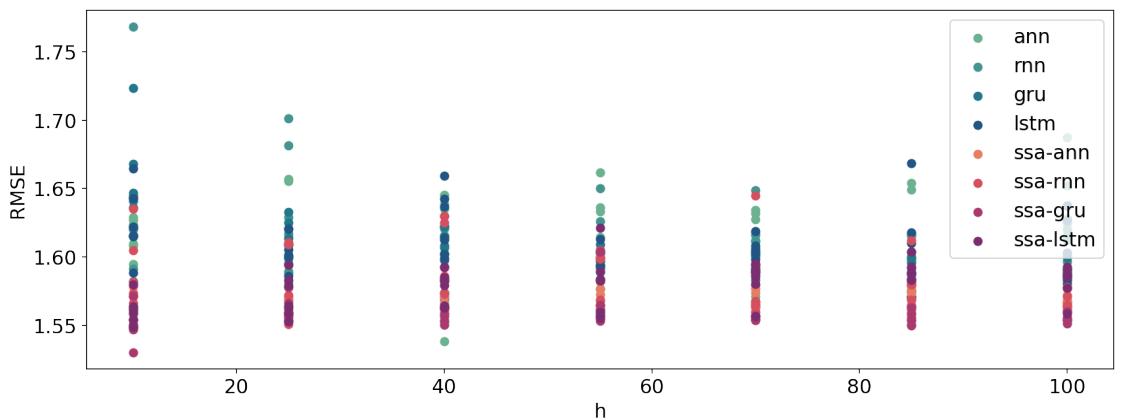
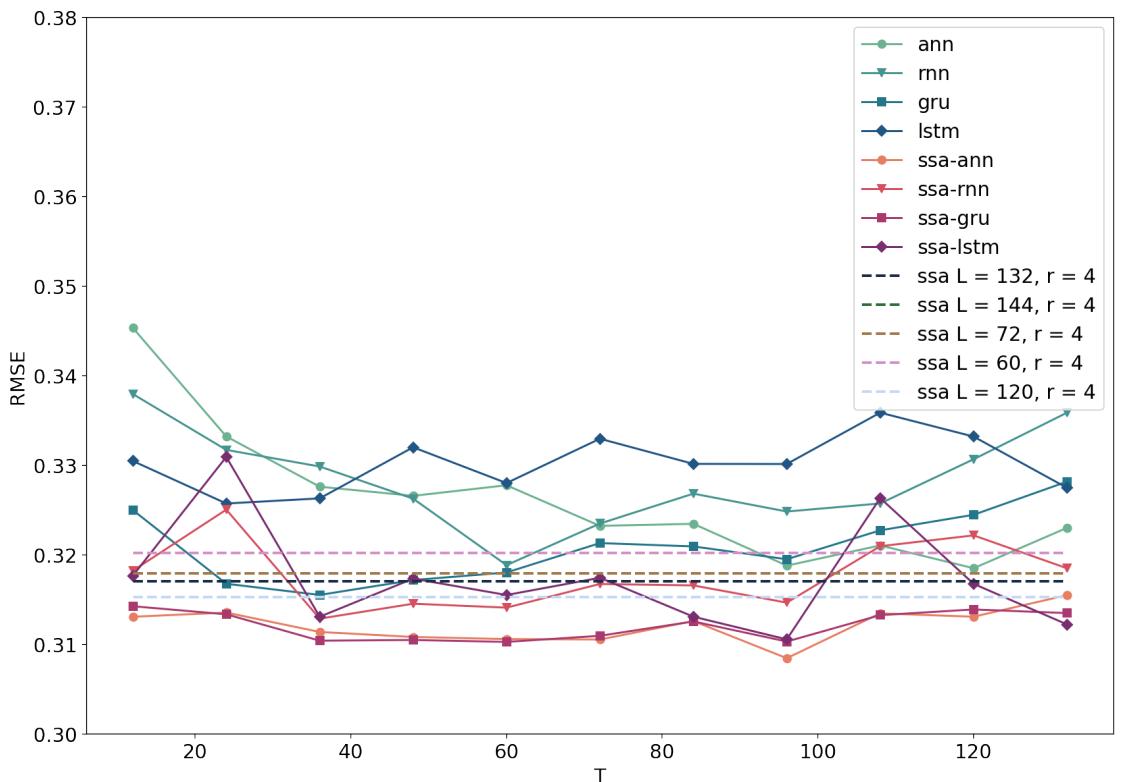
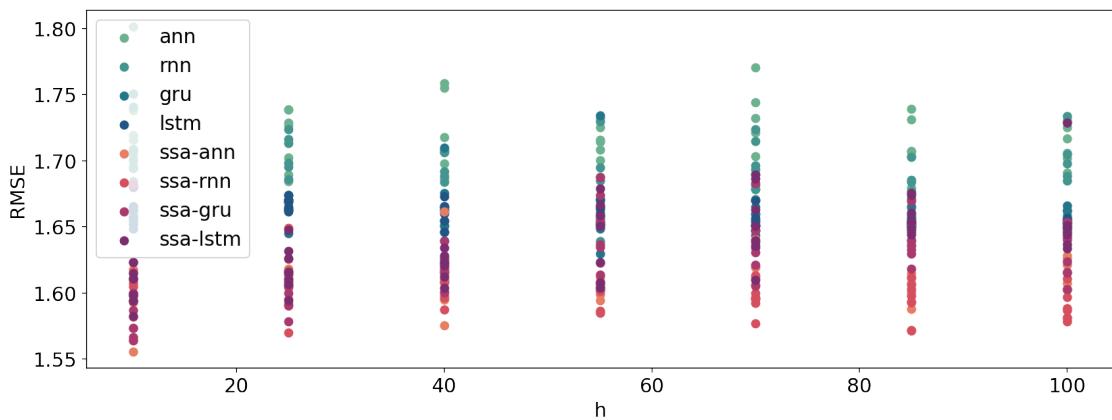
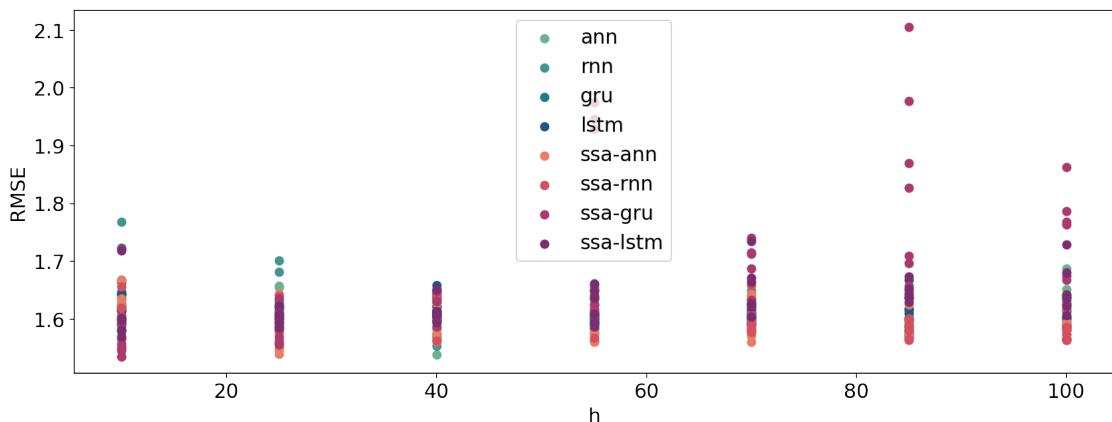
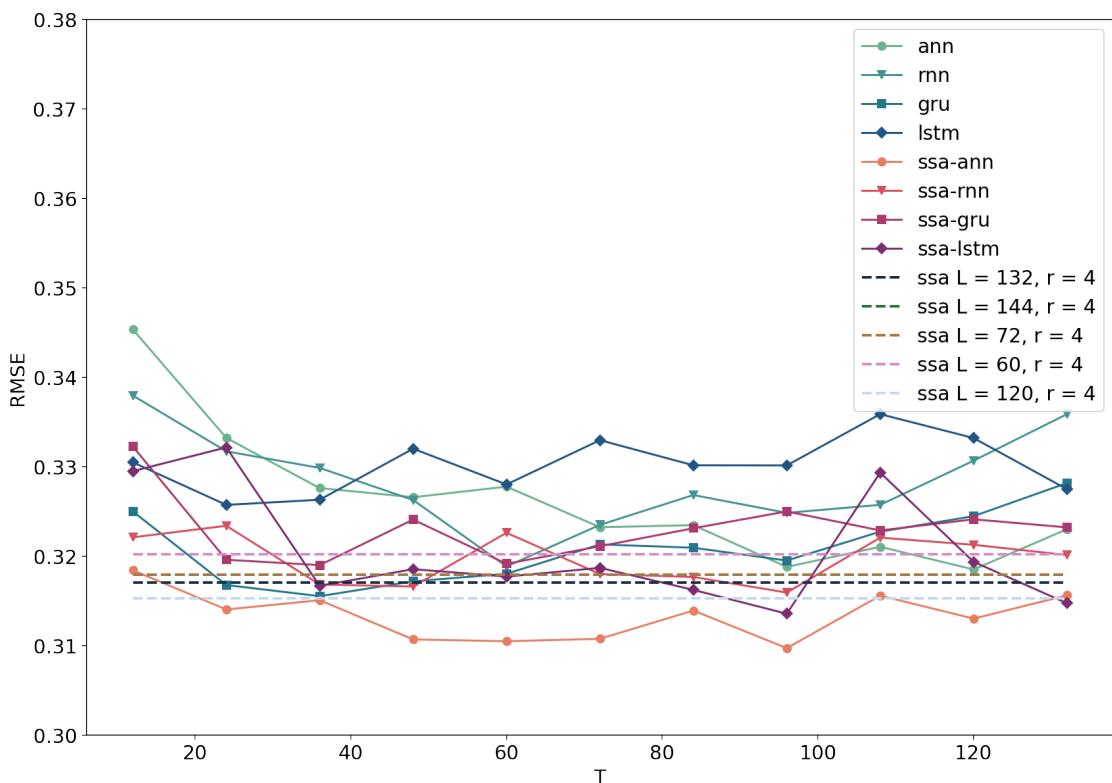


Рис. A.25. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Ошибки прогноза в

зависимости от параметра  $T, L = 175, r = 2$

Рис. A.21. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Проверка устойчивости. $r = 4, L = 175, T = 12$ .Рис. A.22. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Проверка устойчивости. $r = 4, L = 175, T = 84$ .Рис. A.26. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Ошибки прогноза взависимости от параметра  $T, L = 175, r = 4$

Рис. А.23. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Проверка устойчивости. $r = 6, L = 175, T = 12$ .Рис. А.24. «Сумма синусов с белым шумом». Ряд  $Z_{650}$ . Проверка устойчивости. $r = 6, L = 175, T = 84$ .Рис. А.27. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Ошибки прогноза в

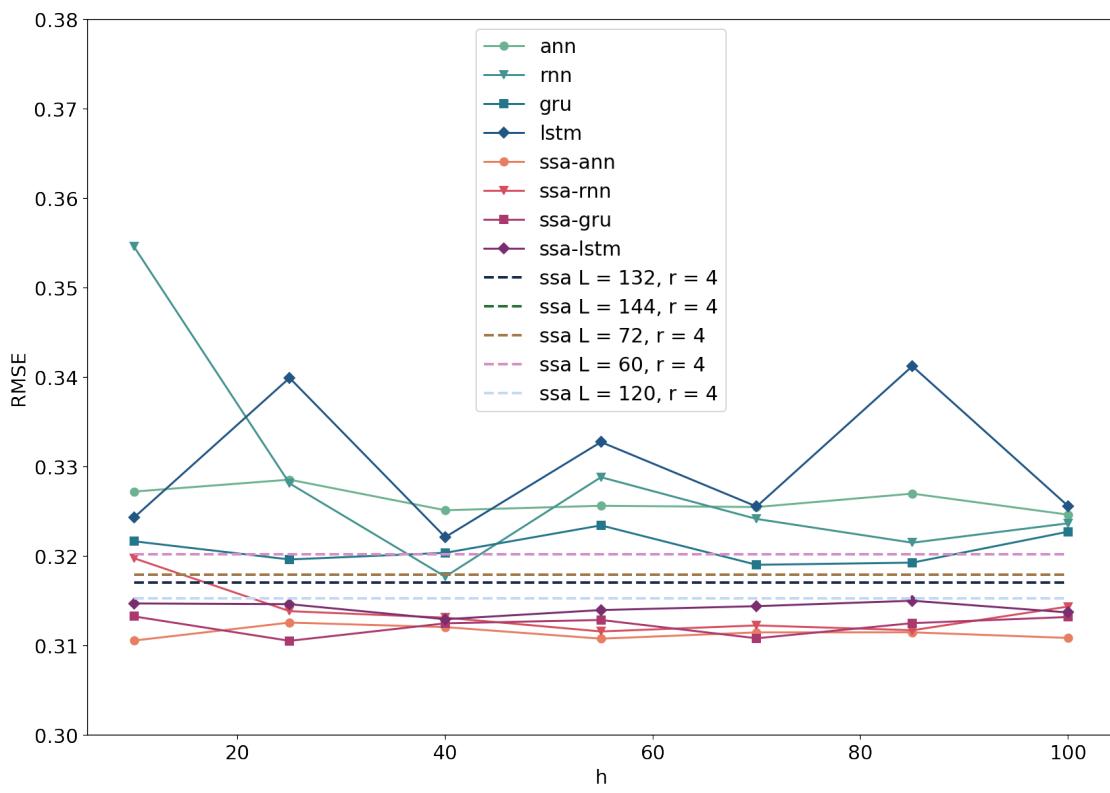


Рис. A.28. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Ошибки прогноза в зависимости от параметра  $h$ .  $L = 175$ ,  $r = 2$ .

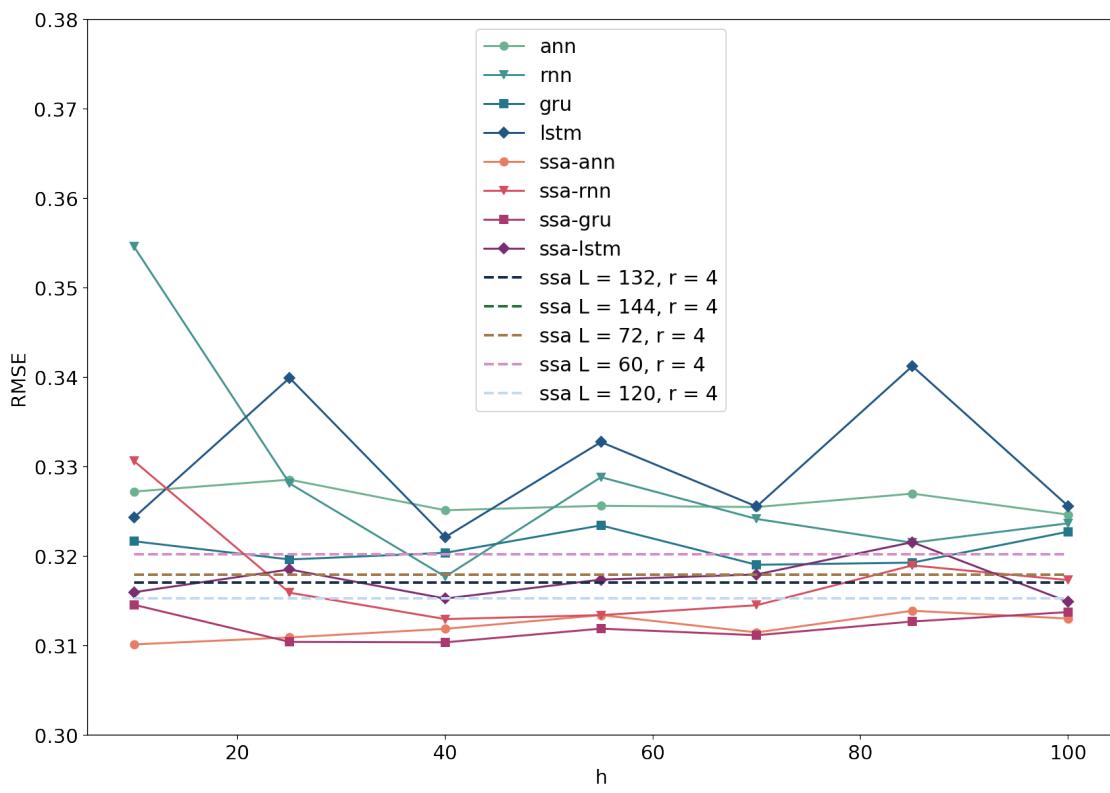


Рис. A.29. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Ошибки прогноза в зависимости от параметра  $h$ .  $L = 175$ ,  $r = 4$ .

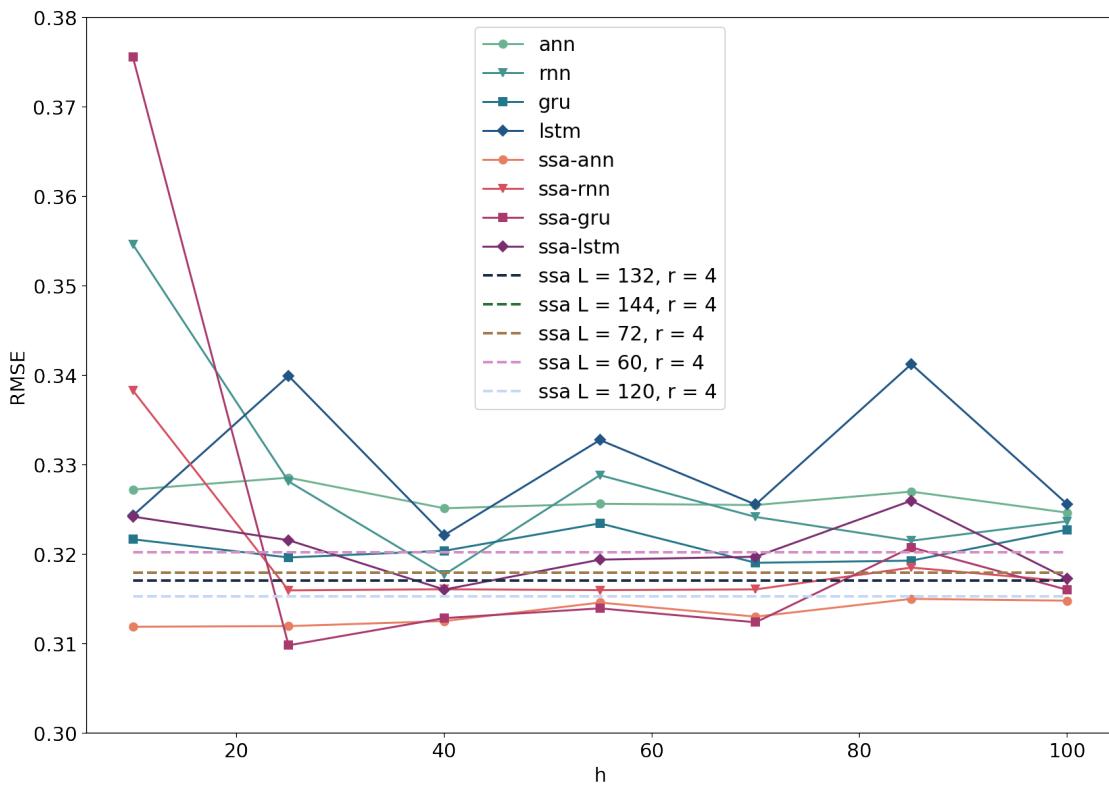


Рис. А.30. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Ошибки прогноза в зависимости от параметра  $h$ .  $L = 175$ ,  $r = 6$ .

### Отображение прогнозов

На графиках ниже показаны результаты прогнозирования тестовой выборки обычных и гибридных методов. Графики являются приложением к разделу 4.1.2.

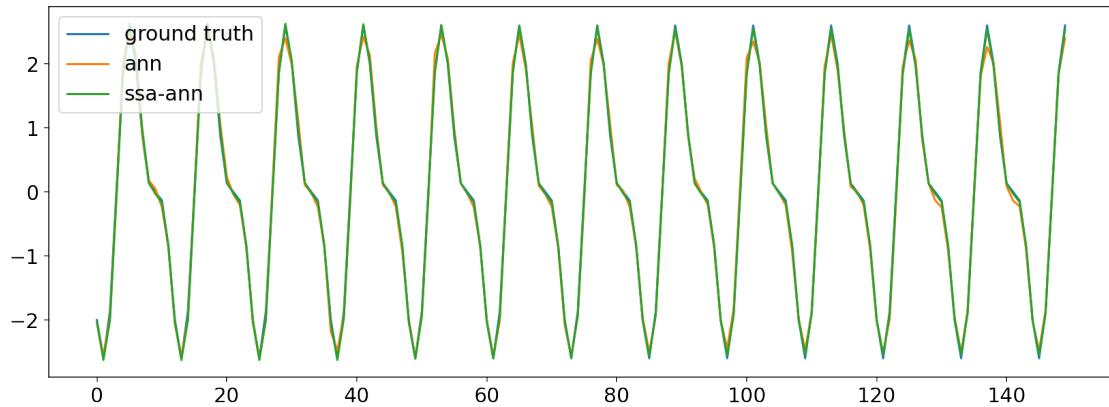


Рис. А.31. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Прогноз результатов для ANN и SSA-ANN.  $r = 2$

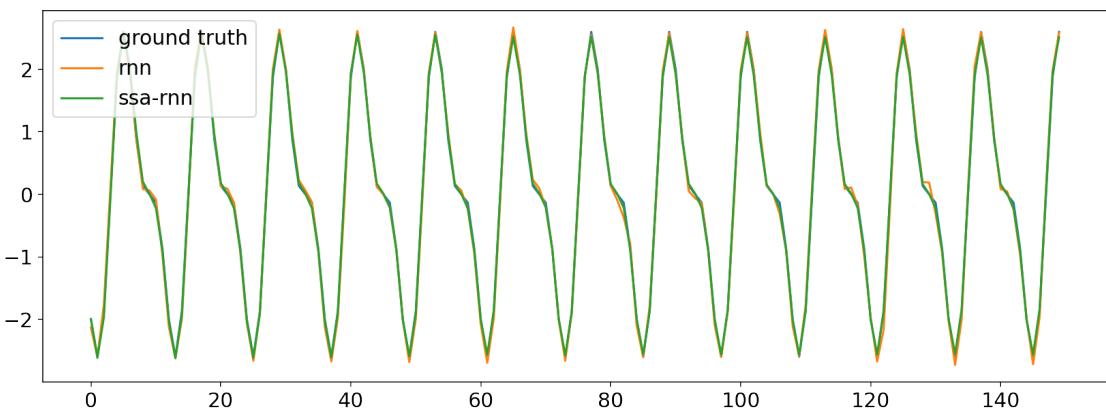


Рис. А.32. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Прогноз результатов для RNN и SSA-RNN.  $r = 2$

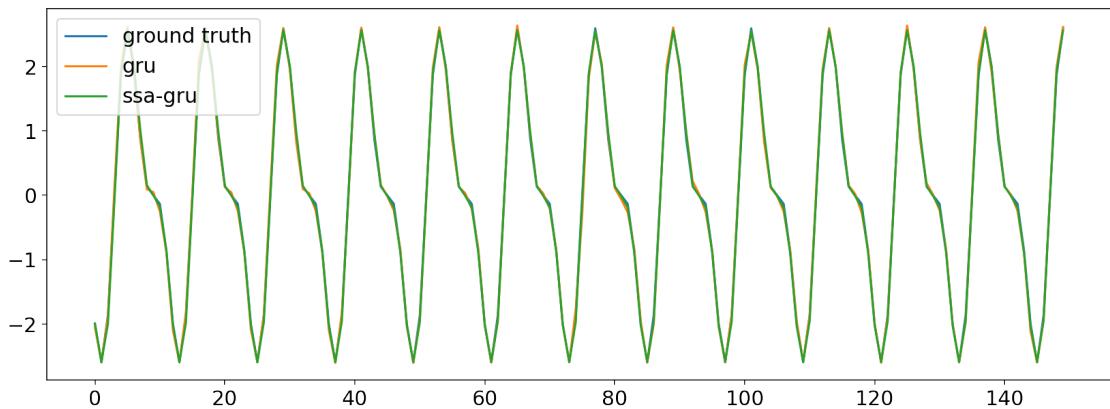


Рис. А.33. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Прогноз результатов для GRU и SSA-GRU.  $r = 2$

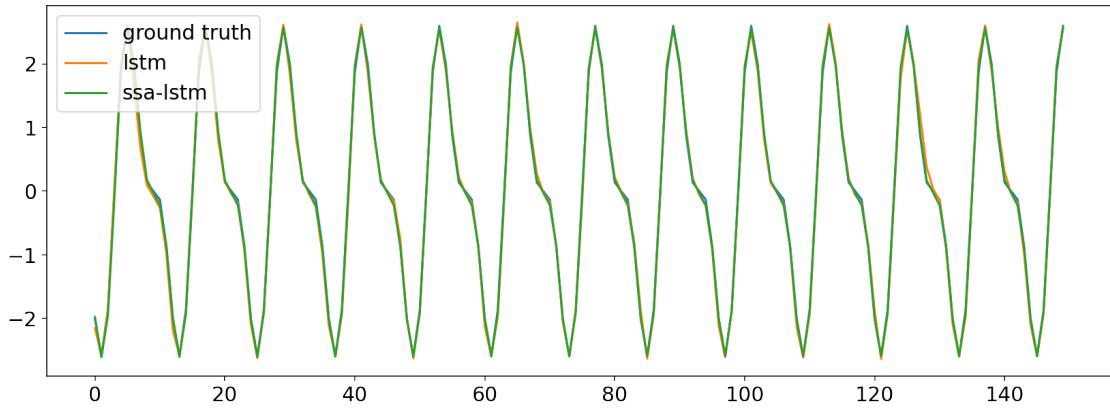


Рис. А.34. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Прогноз результатов для LSTM и SSA-LSTM.  $r = 2$

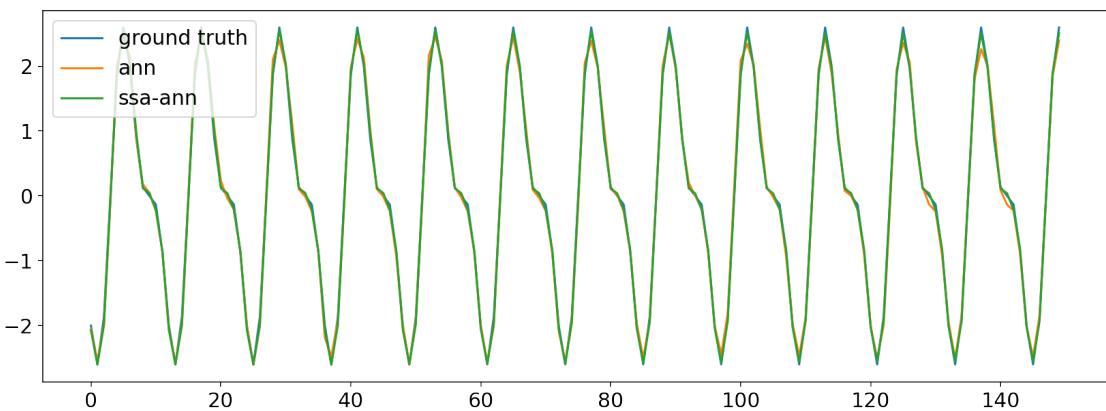


Рис. А.35. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Прогноз результатов для ANN и SSA-ANN.  $r = 4$

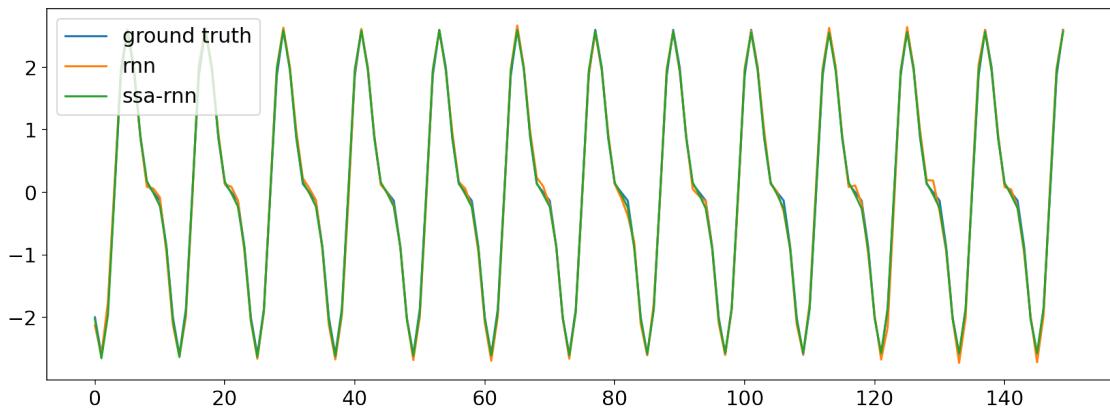


Рис. А.36. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Прогноз результатов для RNN и SSA-RNN.  $r = 4$

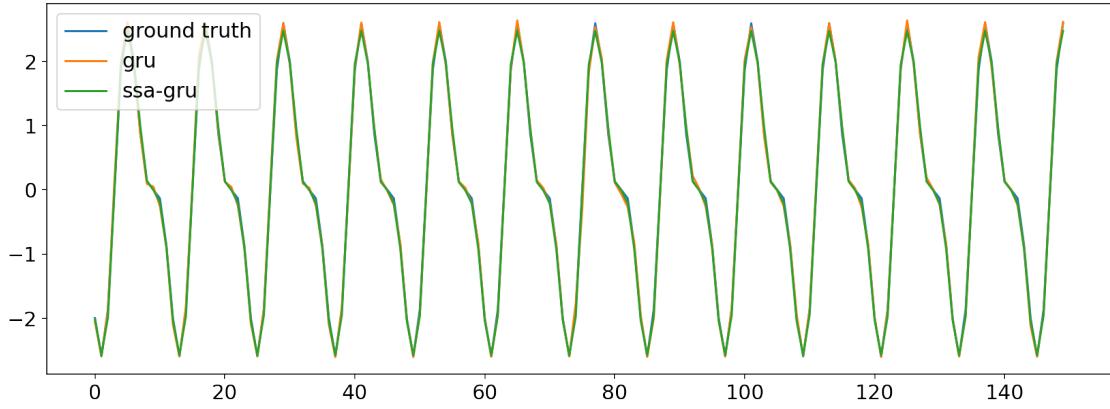


Рис. А.37. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Прогноз результатов для GRU и SSA-GRU.  $r = 4$

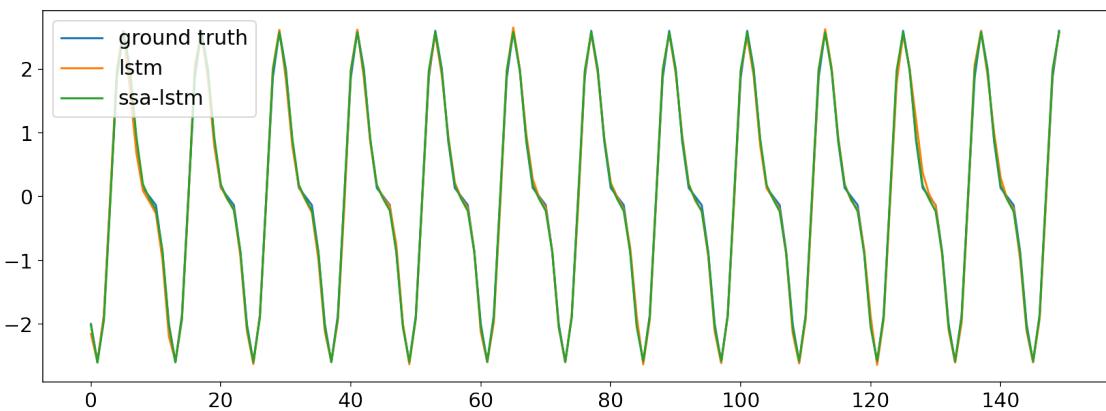


Рис. А.38. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Прогноз результатов для LSTM и SSA-LSTM.  $r = 4$

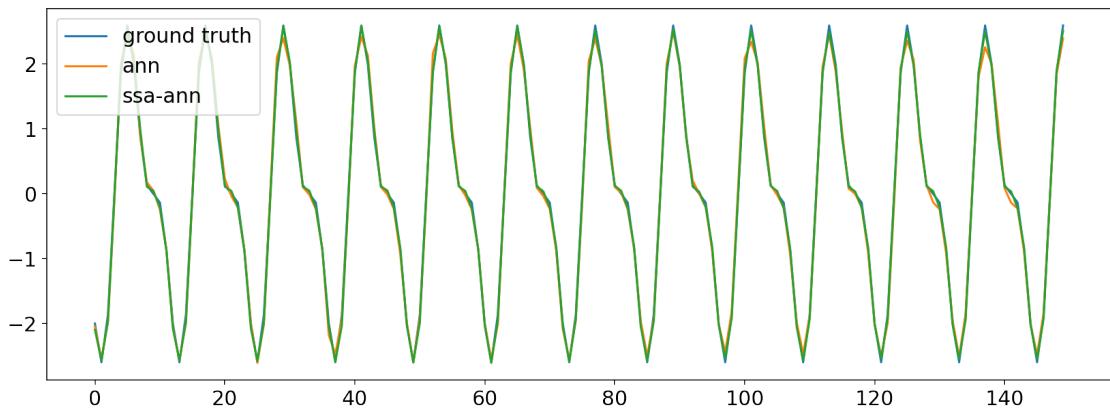


Рис. А.39. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Прогноз результатов для ANN и SSA-ANN.  $r = 6$

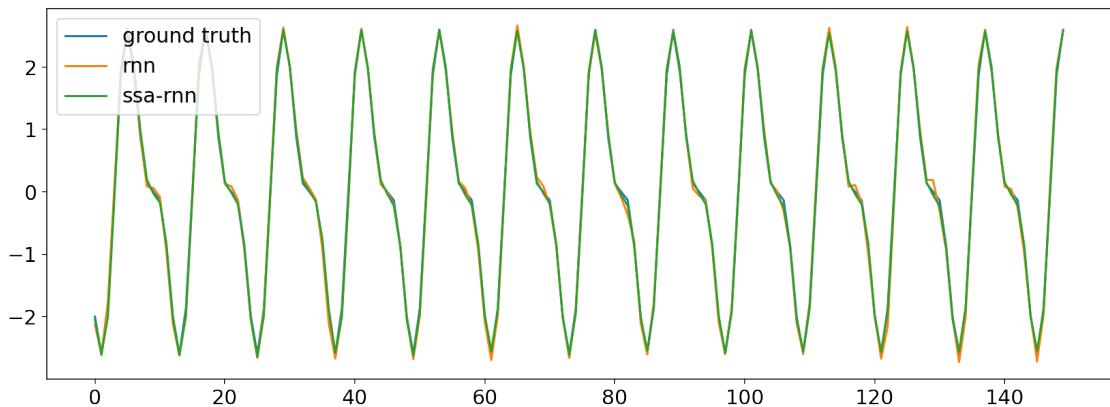


Рис. А.40. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Прогноз результатов для RNN и SSA-RNN.  $r = 6$

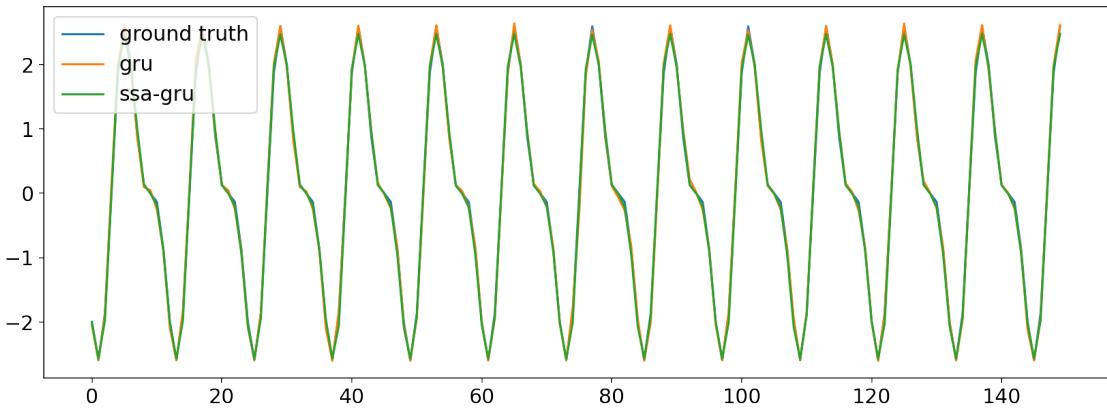


Рис. A.41. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Прогноз результатов для GRU и SSA-GRU.  $r = 6$

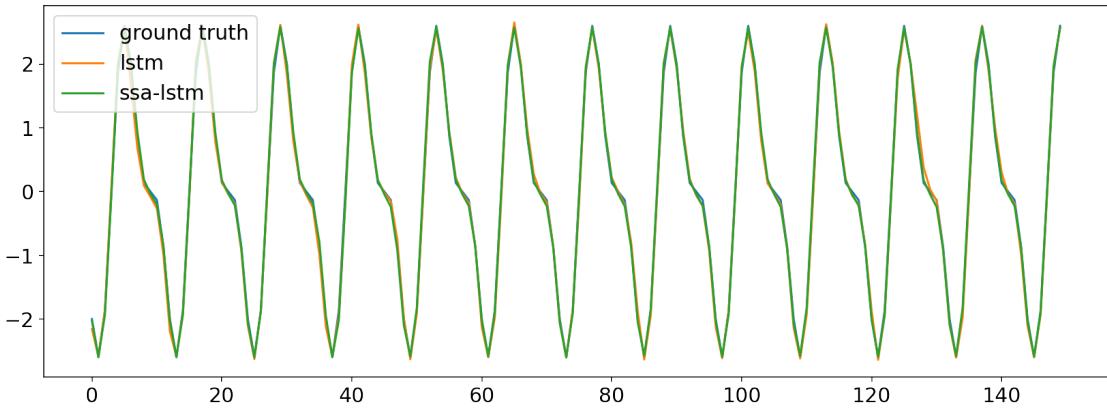


Рис. A.42. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Прогноз результатов для LSTM и SSA-LSTM.  $r = 6$

### Проверка устойчивости

На графиках проверяются устойчивость результатов по методике из раздела 3.6. Графики являются приложением к разделу 4.1.2.

#### A.1.3. Влияние длины

##### Отображение прогнозов

1

---

<sup>1</sup> Пусто?

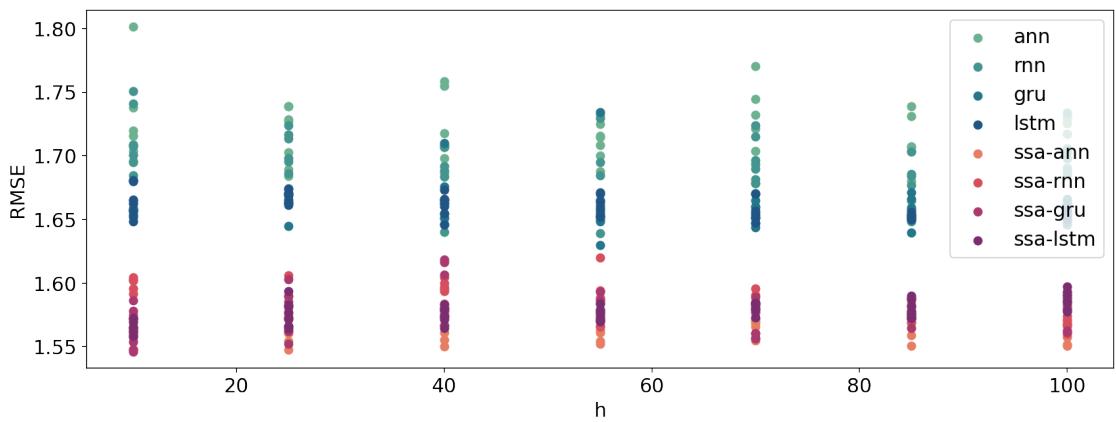


Рис. А.43. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Проверка устойчивости.

$r = 2, L = 175, T = 12.$

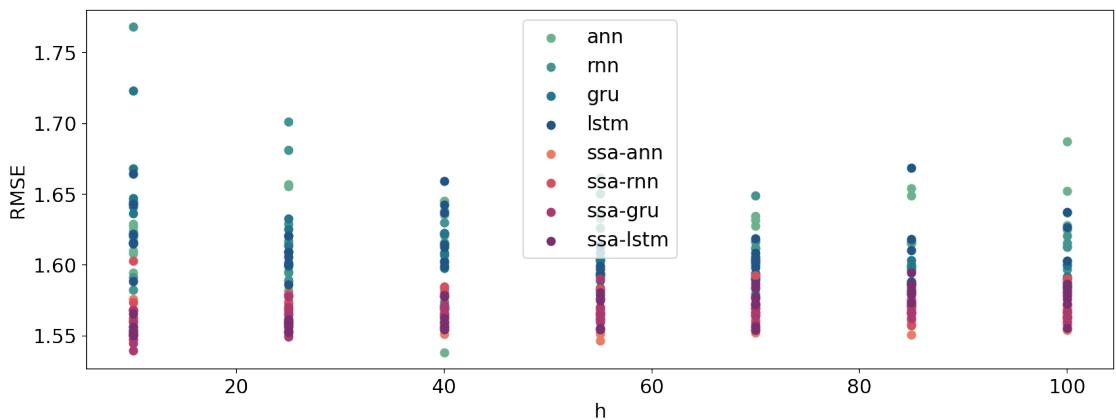


Рис. А.44. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Проверка устойчивости.

$r = 2, L = 175, T = 84.$

## Проверка устойчивости

2

## A.2. Красный шум

В этом разделе приложены графики имеющие отношения к эксперименту из главы 4.2.1.

---

<sup>2</sup> Пусто?

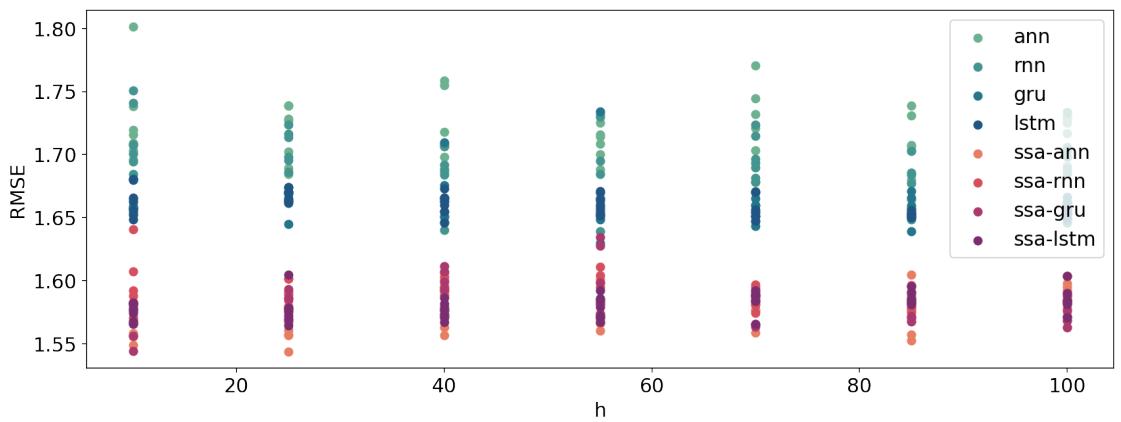


Рис. А.45. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Проверка устойчивости.

$$r = 4, \quad L = 175. \quad T = 12.$$

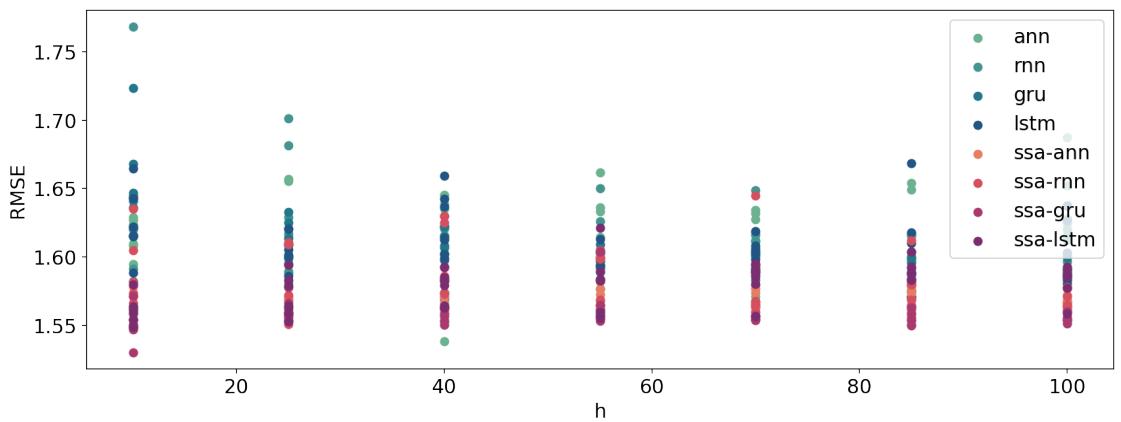


Рис. А.46. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Проверка устойчивости.

$$r = 4, \quad L = 175. \quad T = 84.$$

### Сравнение обычных и гибридных методов

На графиках ниже показаны результаты сравнение обычных и гибридных методов, а также метода SSA. Графики являются приложением к разделу 4.2.1.

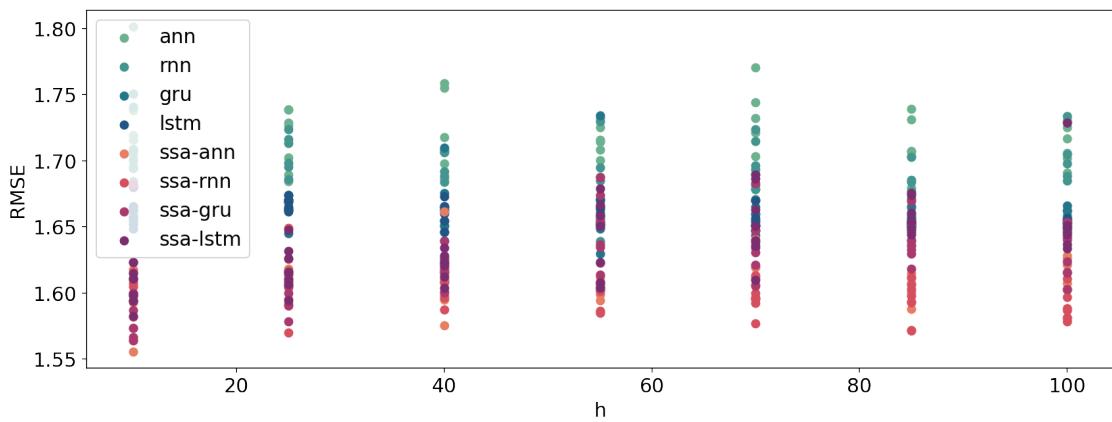


Рис. А.47. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Проверка устойчивости.

$r = 6, L = 175, T = 12$ .

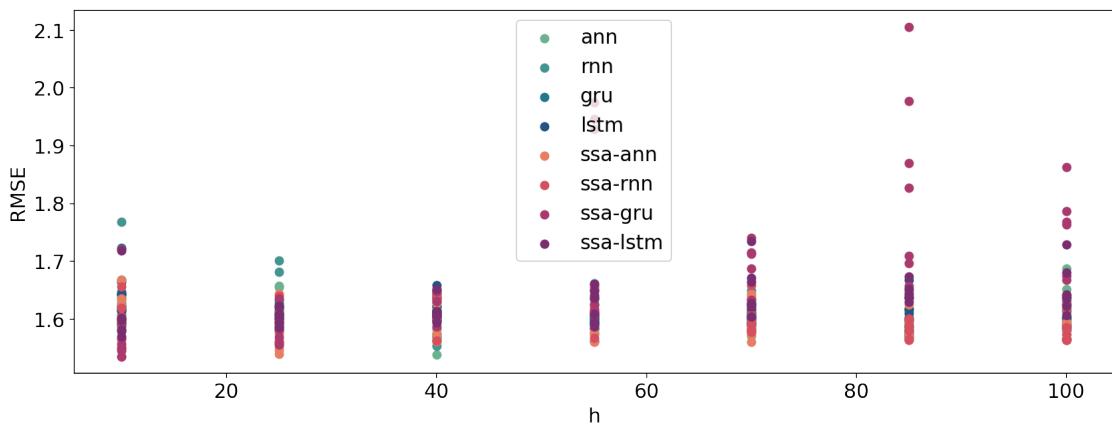


Рис. А.48. «Сумма синусов с небольшим белым шумом». Ряд  $X_{650}$ . Проверка устойчивости.

$r = 6, L = 175, T = 84$ .

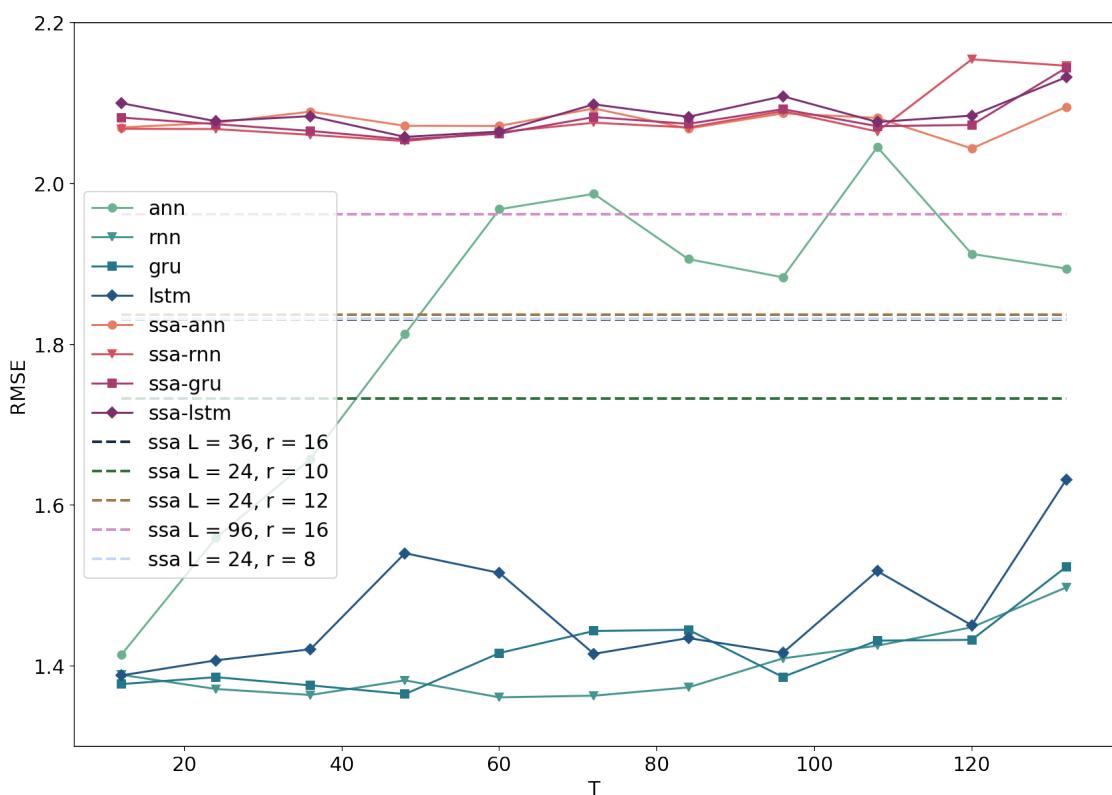


Рис. А.49. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно

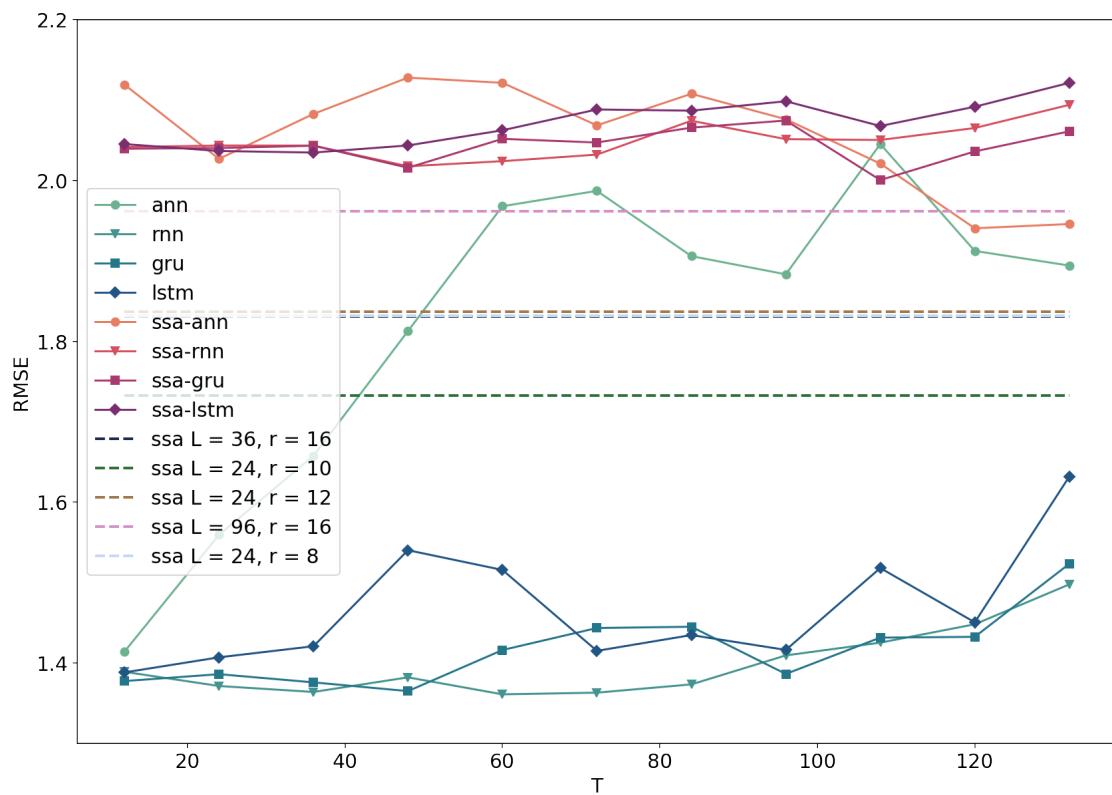


Рис. А.50. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно ряда в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 4$ .

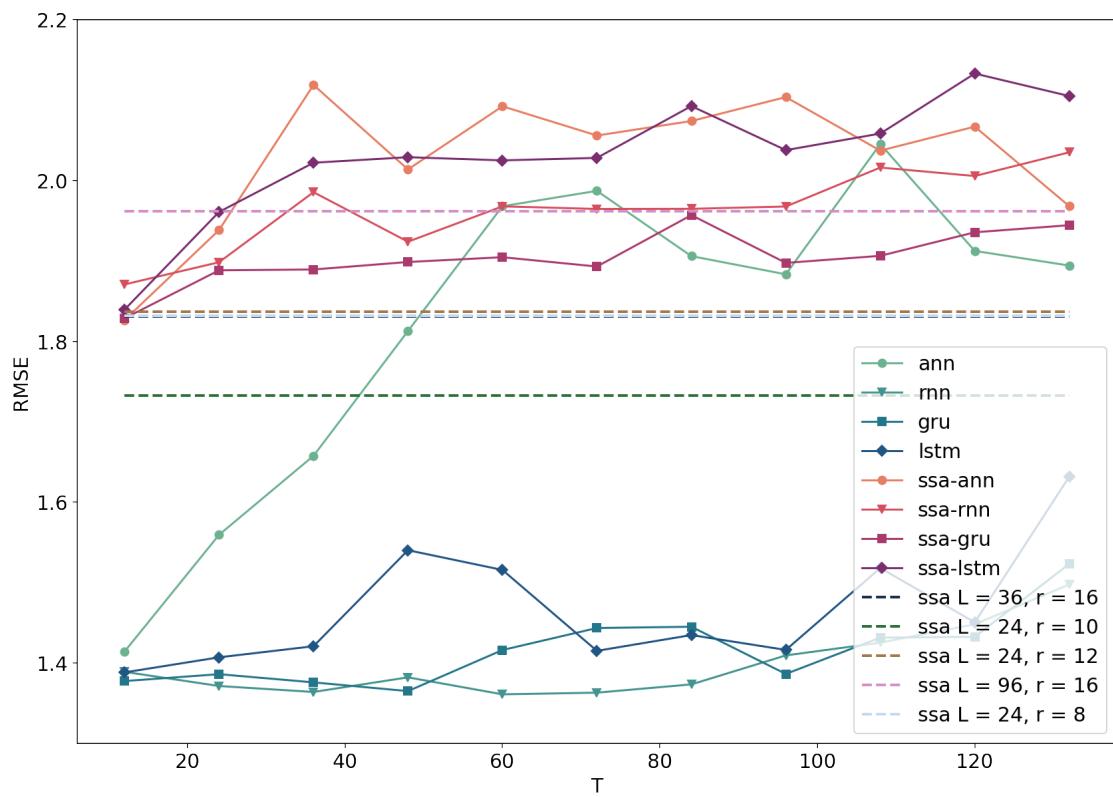


Рис. А.51. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно ряда в зависимости от параметра  $T$ .  $L = 84$ ,  $r = 14$ .

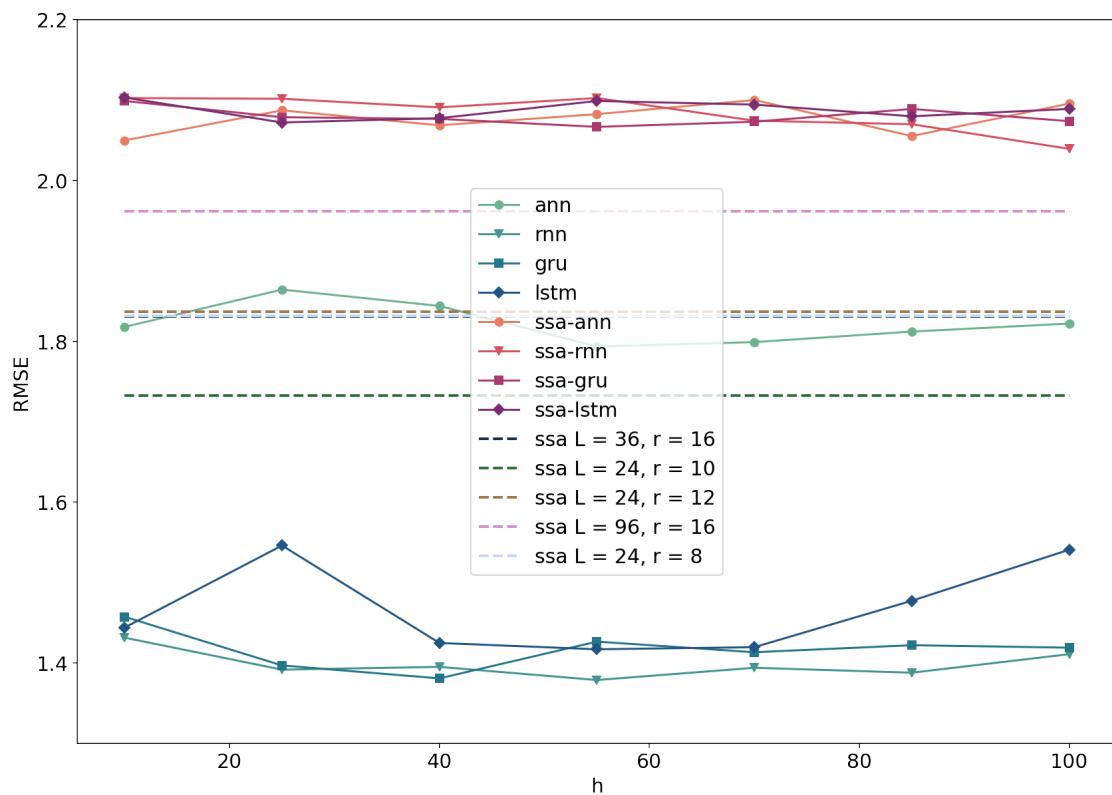


Рис. A.52. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно ряда в зависимости от параметра  $h$ .  $L = 175$ ,  $r = 2$ .

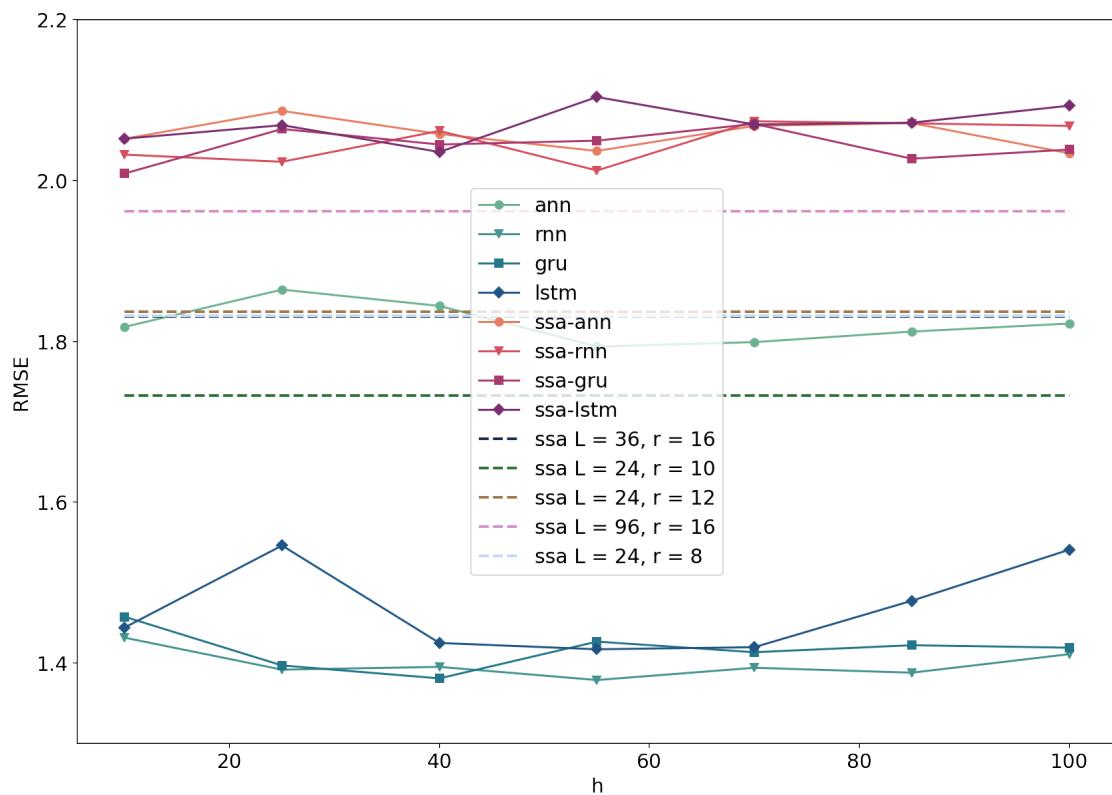


Рис. A.53. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно ряда в зависимости от параметра  $h$ .  $L = 175$ ,  $r = 4$ .

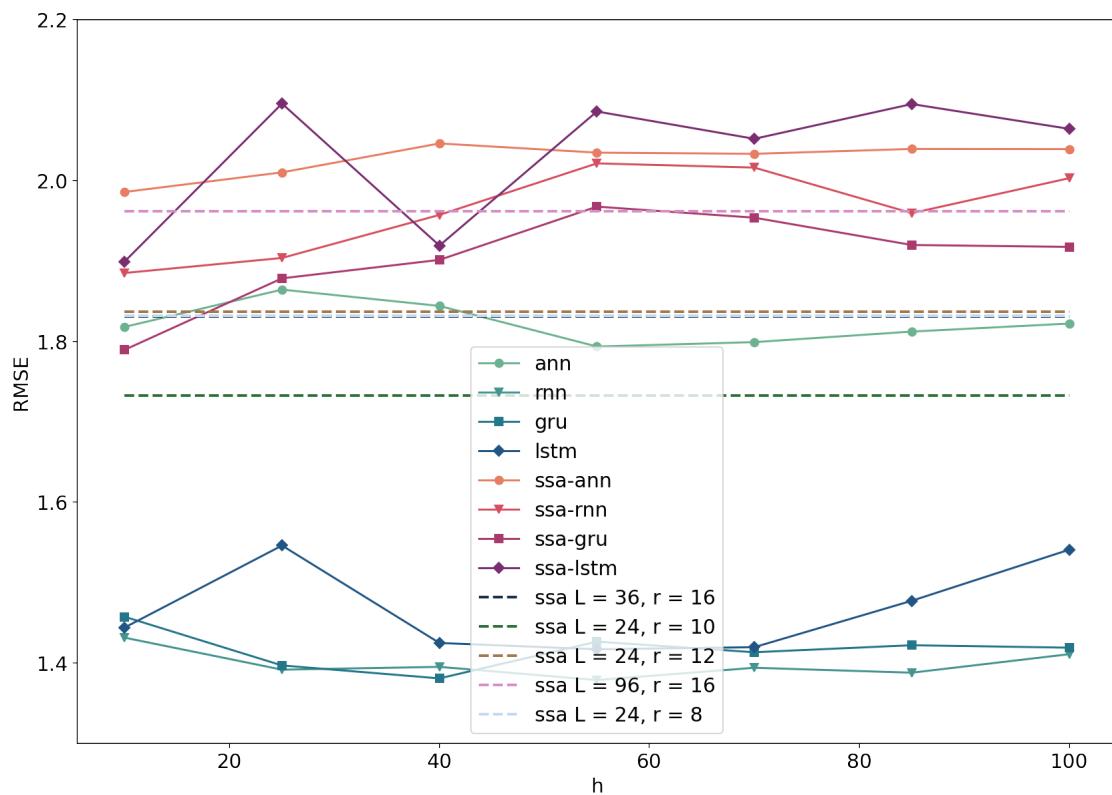


Рис. A.54. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно ряда в зависимости от параметра  $h$ .  $L = 84$ ,  $r = 14$ .

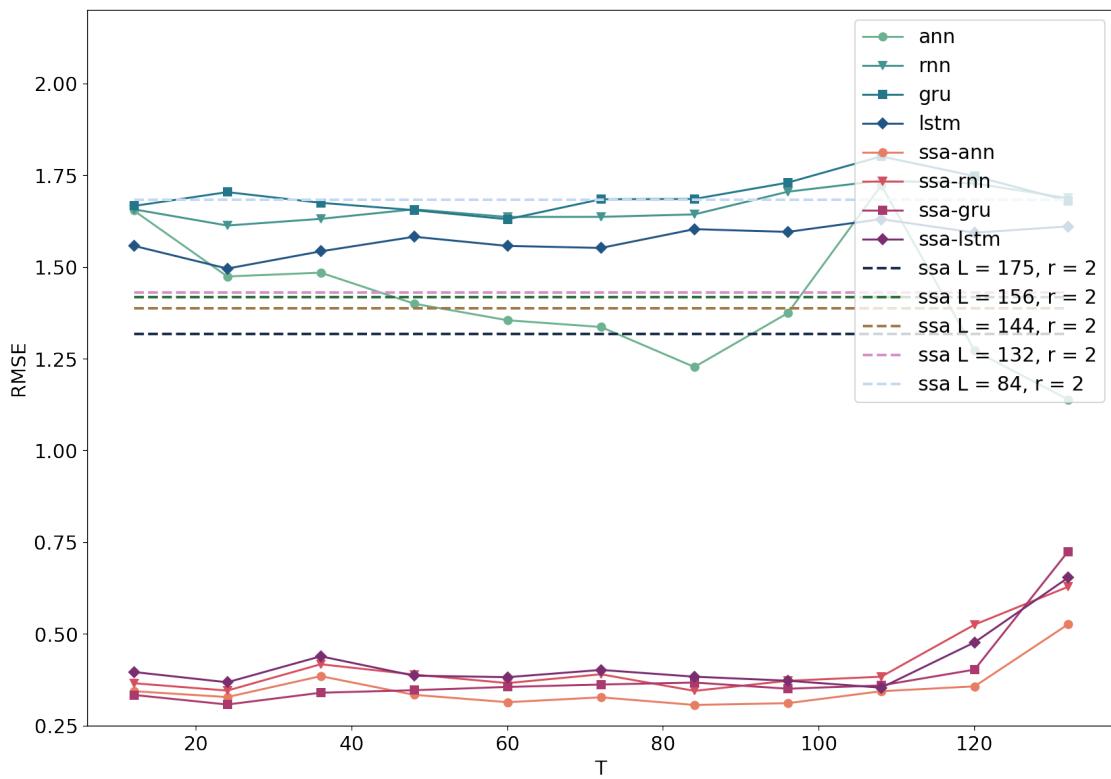


Рис. А.55. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно сигнала в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 2$ .

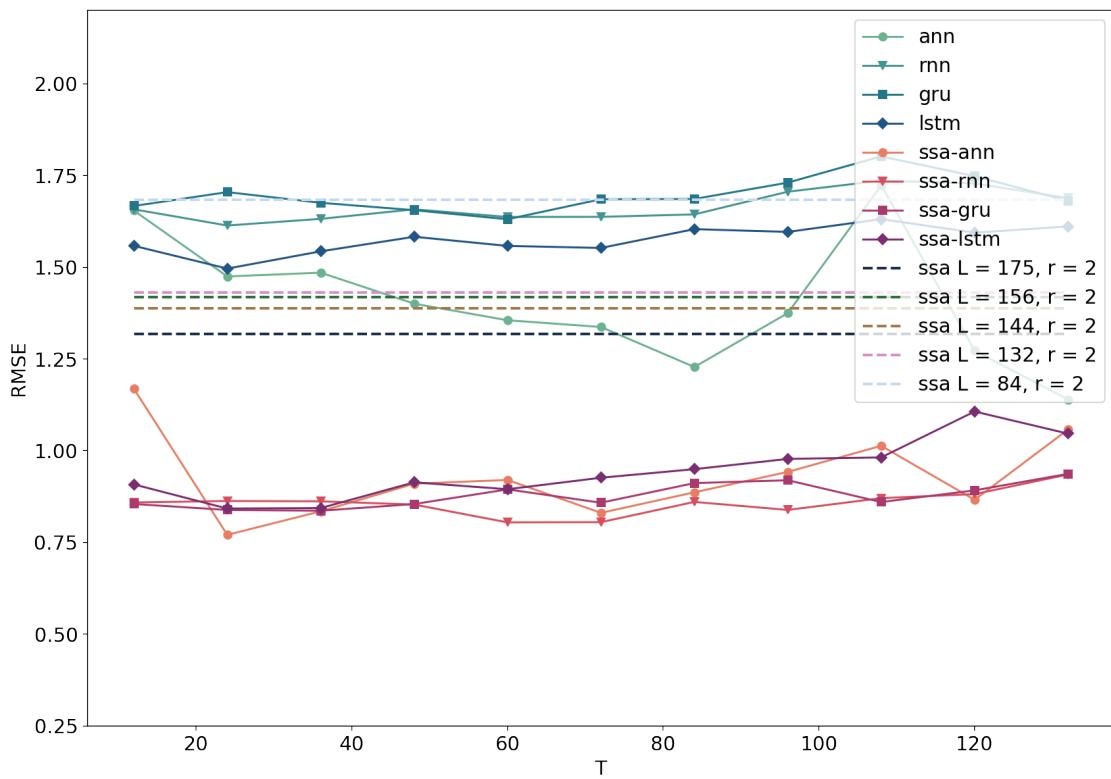


Рис. А.56. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно сигнала в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 4$ .

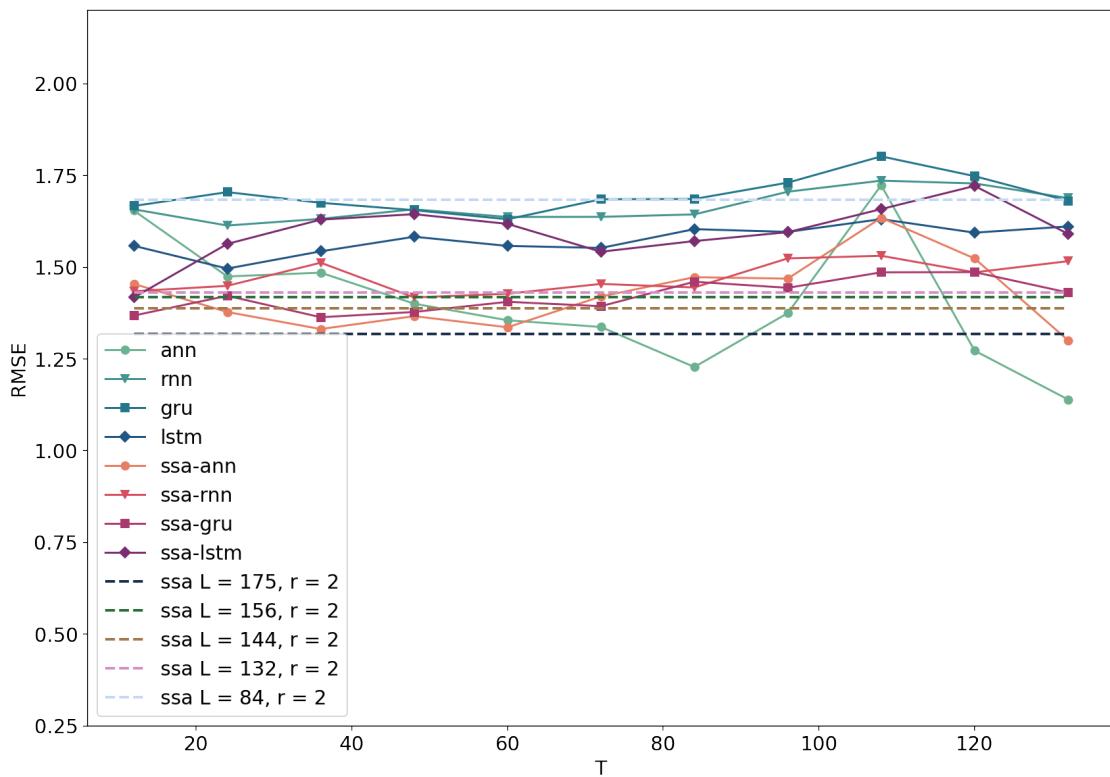


Рис. А.57. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно сигнала в зависимости от параметра  $T$ .  $L = 84$ ,  $r = 14$ .

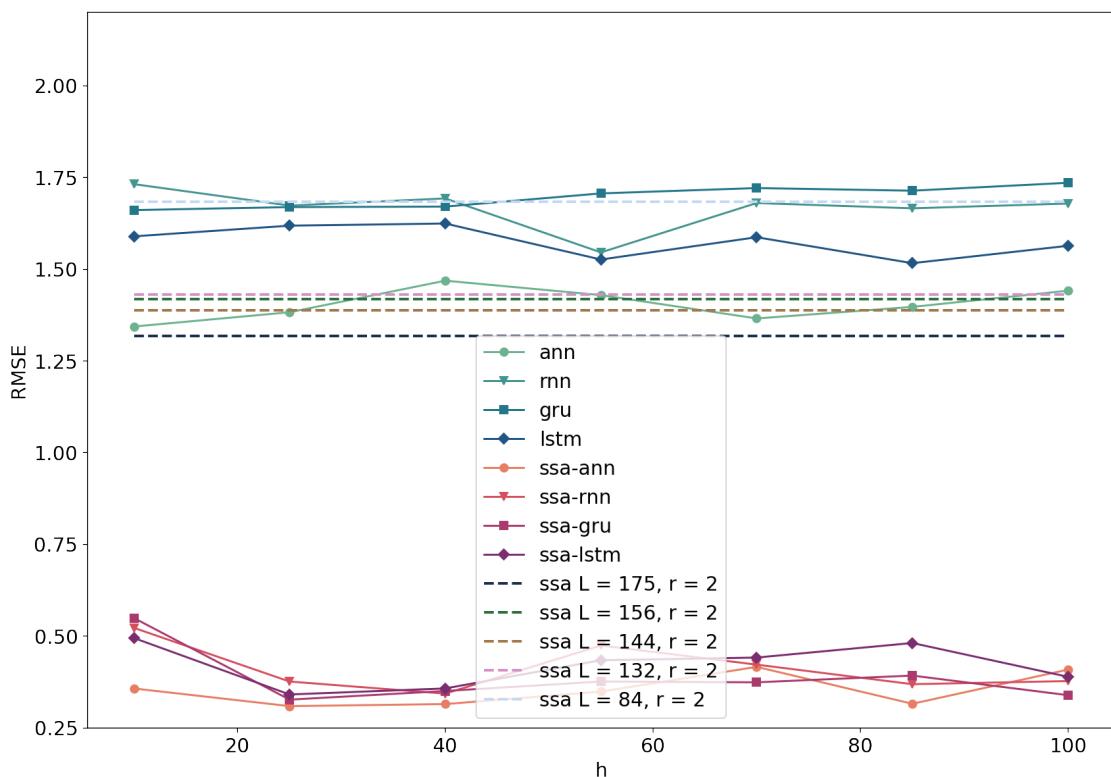


Рис. A.58. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно сигнала в зависимости от параметра  $h$ .  $L = 175, r = 2$ .

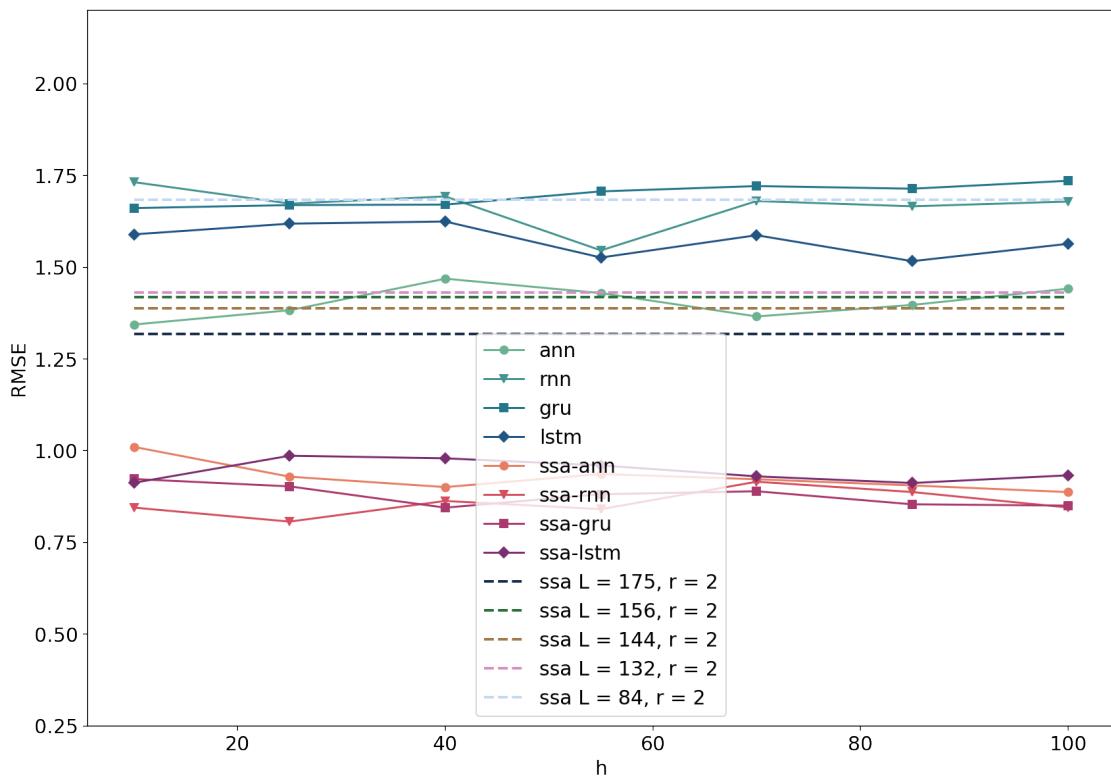


Рис. А.59. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно сигнала в зависимости от параметра  $h$ .  $L = 175$ ,  $r = 4$ .

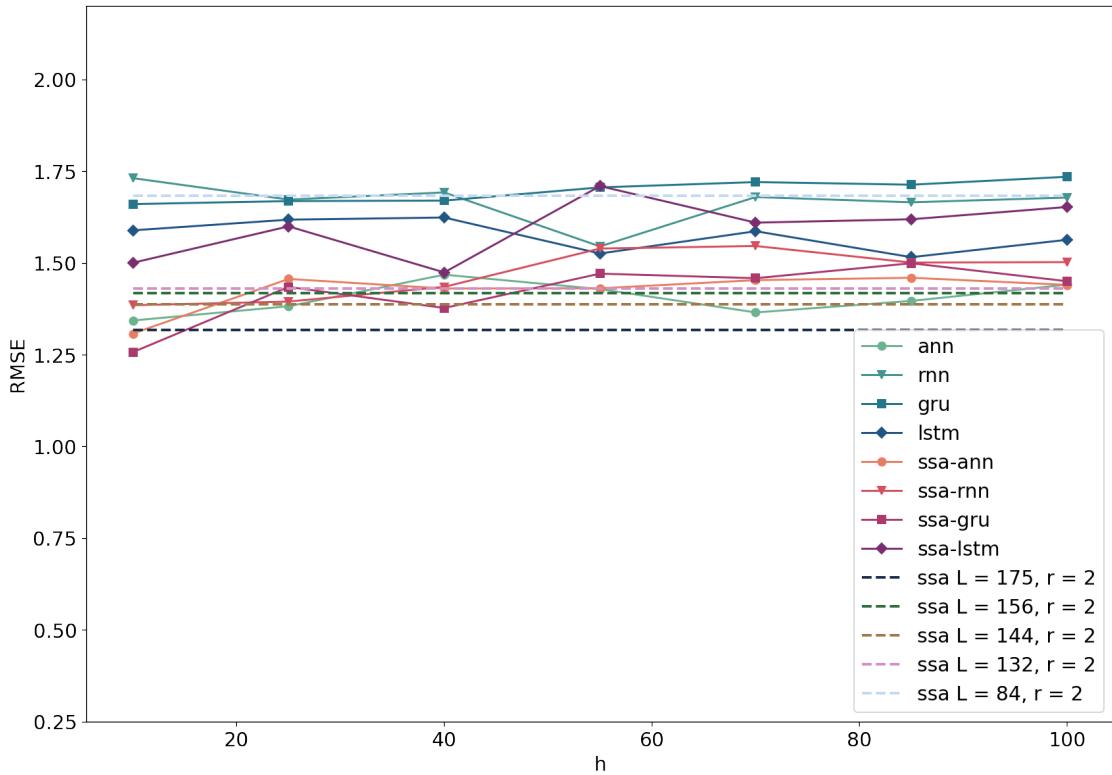


Рис. А.60. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Ошибки прогноза относительно сигнала в зависимости от параметра  $h$ .  $L = 84$ ,  $r = 14$ .

### Отображение прогнозов

На графиках ниже показаны результаты прогнозирования тестовой выборки обычных и гибридных методов. Графики являются приложением к разделу 4.2.1.

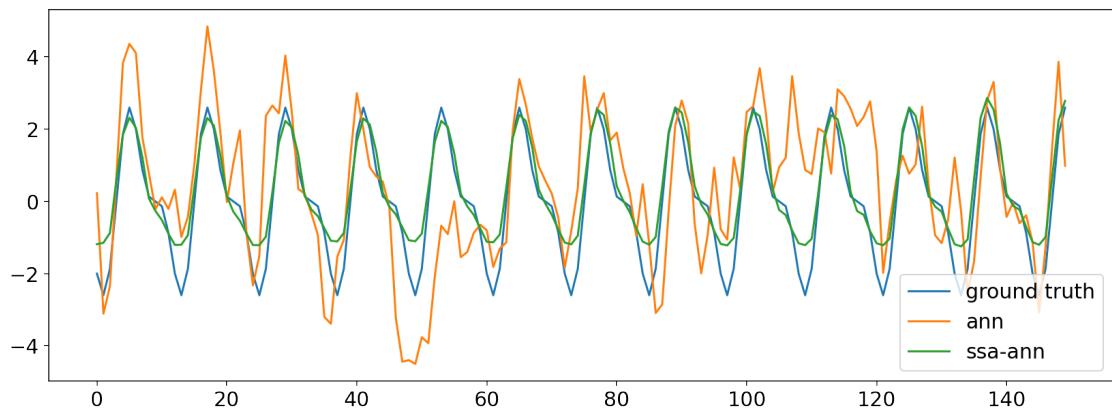


Рис. А.61. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Прогноз результатов для ANN и SSA-ANN.  $L = 175$ ,  $r = 2$

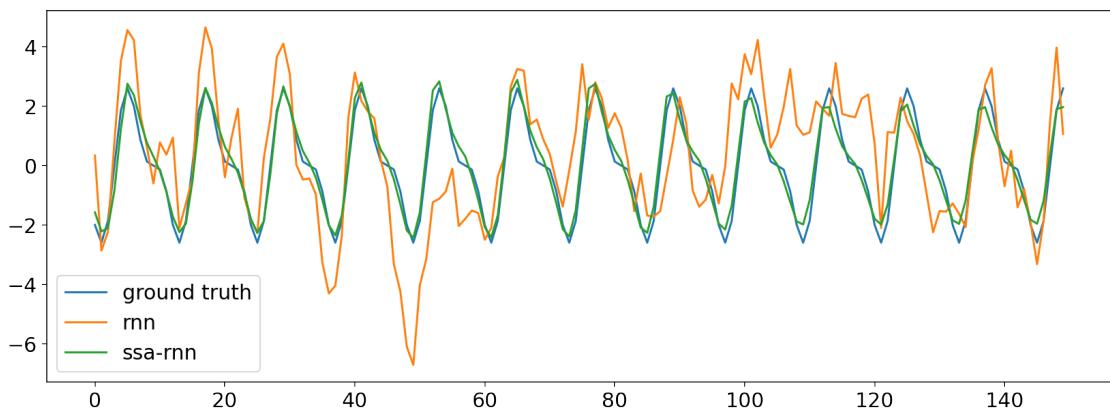


Рис. А.62. С«Сумма синусов с красным шумом». Ряд Z<sub>650</sub>. Прогноз результатов для RNN и SSA-RNN.  $L = 175$ ,  $r = 2$

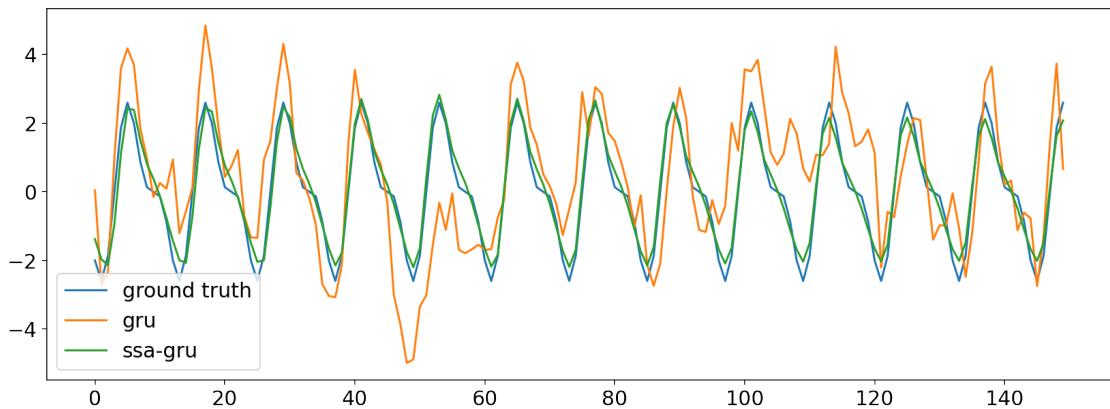


Рис. А.63. «Сумма синусов с красным шумом». Ряд Z<sub>650</sub>. Прогноз результатов для GRU и SSA-GRU.  $L = 175$ ,  $r = 2$

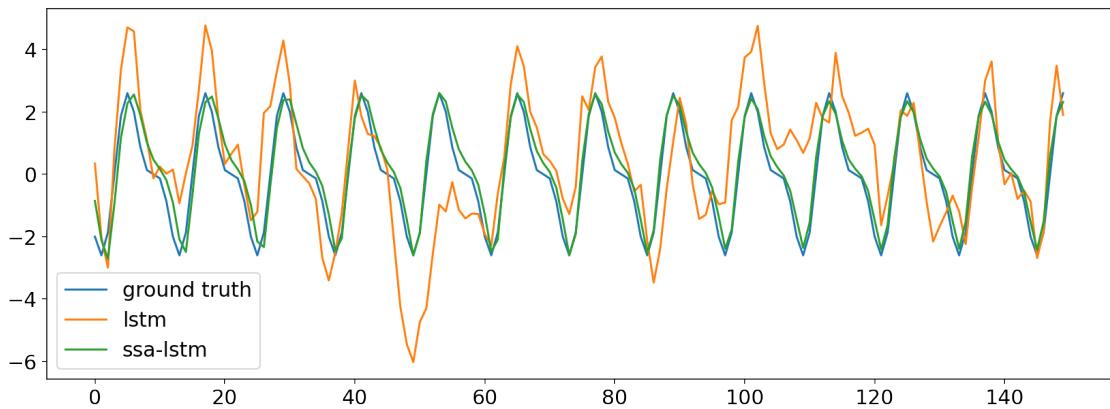


Рис. А.64. «Сумма синусов с красным шумом». Ряд Z<sub>650</sub>. Прогноз результатов для LSTM и SSA-LSTM.  $L = 175$ ,  $r = 2$

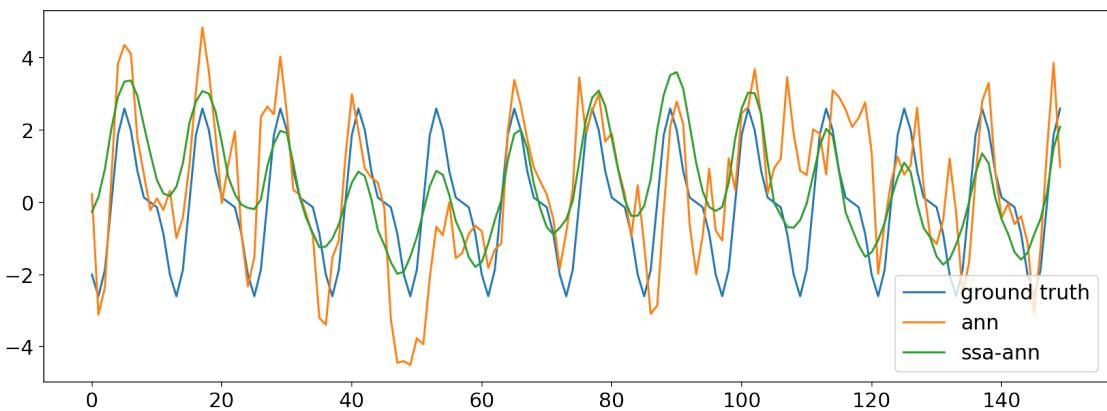


Рис. А.65. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Прогноз результатов для ANN и SSA-ANN.  $L = 175$ ,  $r = 4$

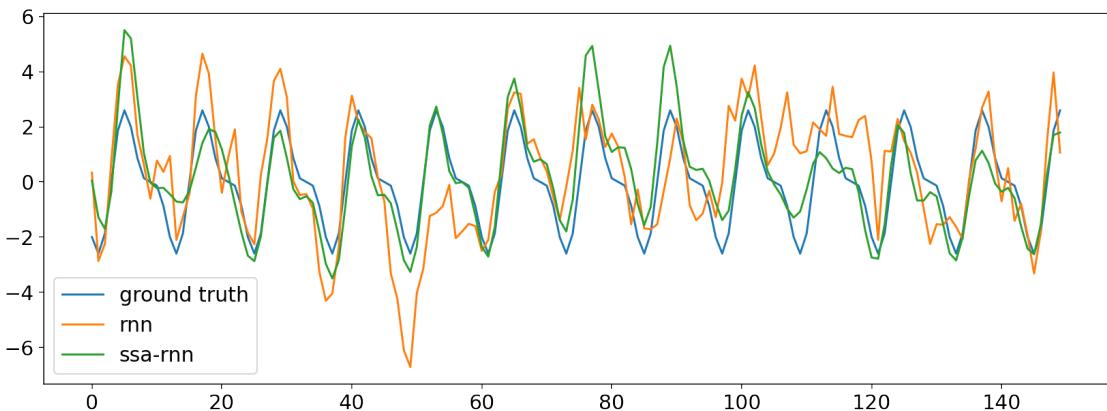


Рис. А.66. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Прогноз результатов для RNN и SSA-RNN.  $L = 175$ ,  $r = 4$

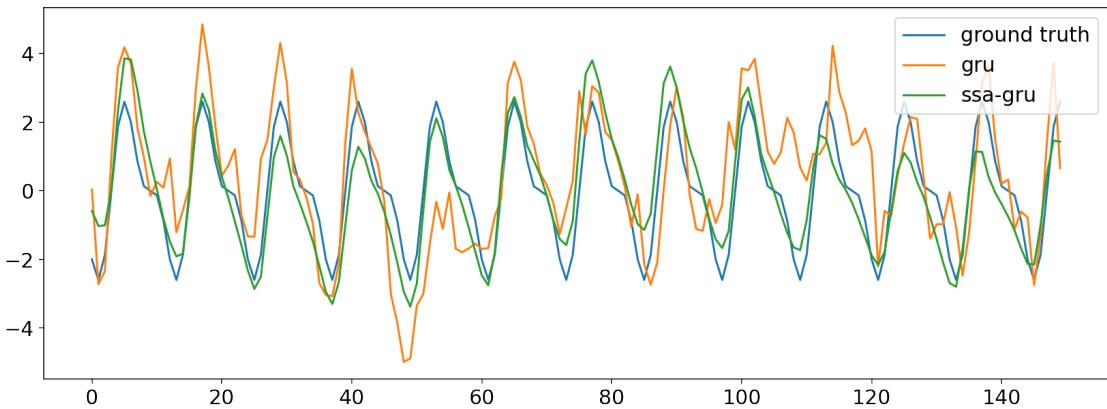


Рис. А.67. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Прогноз результатов для GRU и SSA-GRU.  $L = 175$ ,  $r = 4$

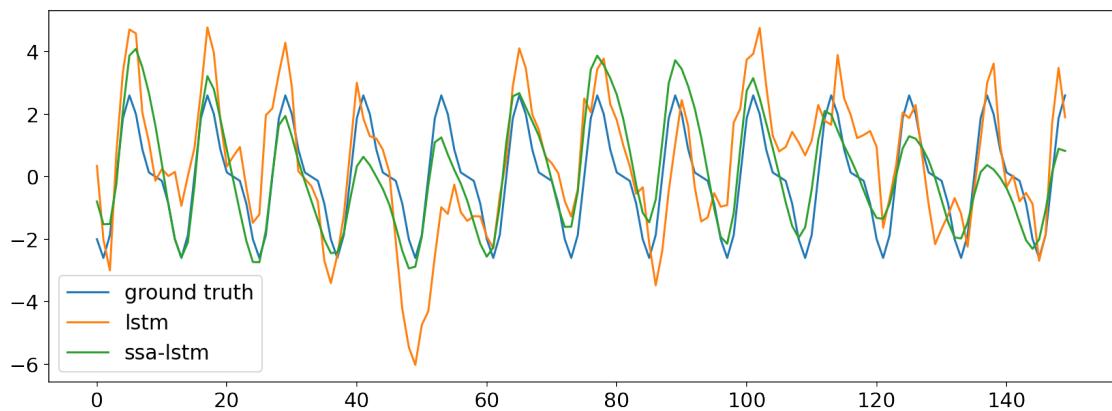


Рис. А.68. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Прогноз результатов для LSTM и SSA-LSTM.  $L = 175$ ,  $r = 4$

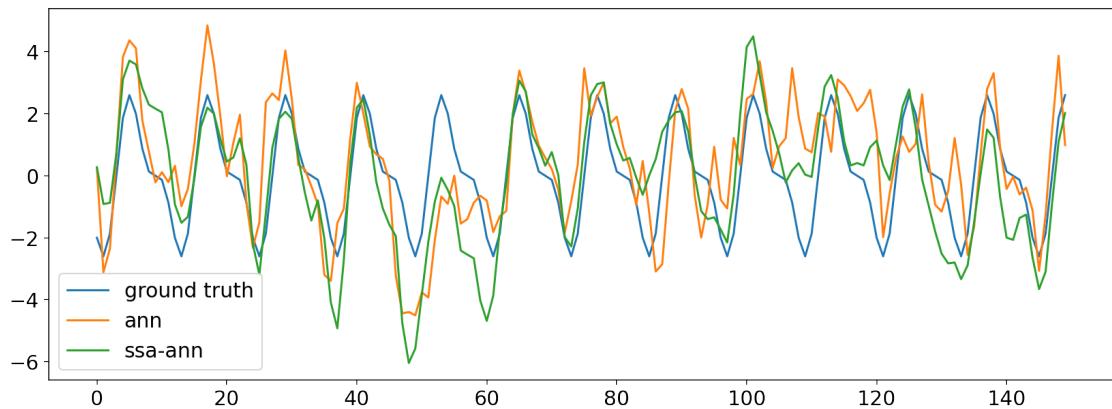


Рис. А.69. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Прогноз результатов для ANN и SSA-ANN.  $L = 84$ ,  $r = 14$

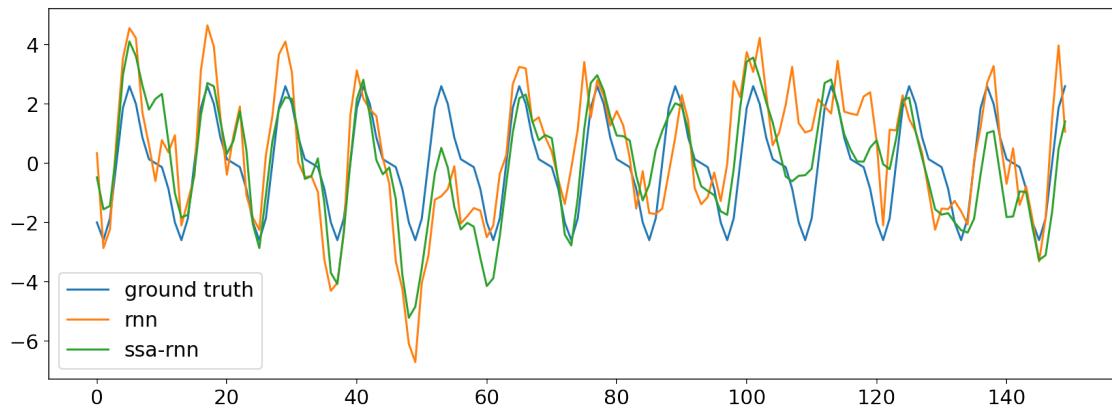


Рис. А.70. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Прогноз результатов для RNN и SSA-RNN.  $L = 84$ ,  $r = 14$

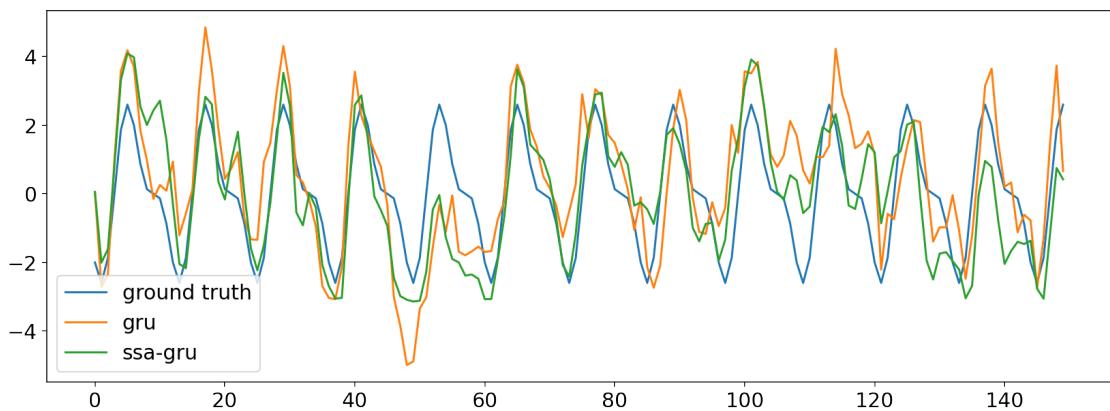


Рис. А.71. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Прогноз результатов для GRU и SSA-GRU.  $L = 84$ ,  $r = 14$

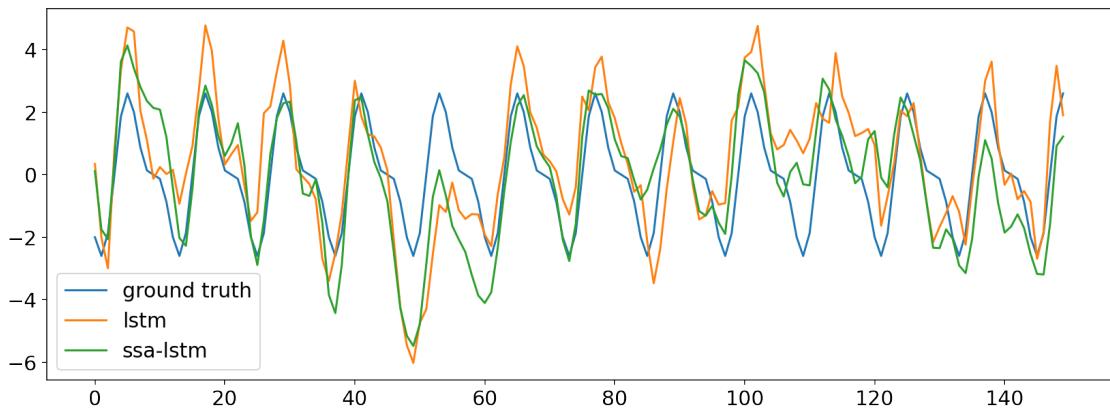


Рис. А.72. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Прогноз результатов для LSTM и SSA-LSTM.  $L = 84$ ,  $r = 14$

### Проверка устойчивости

На графиках проверяются устойчивость результатов по методике из раздела 3.6. Графики являются приложением к разделу 4.2.1.

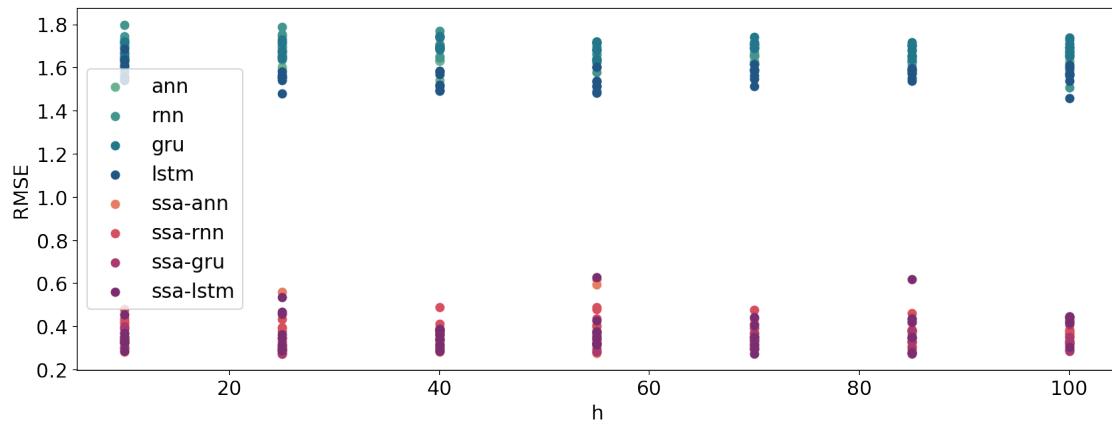


Рис. А.73. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$r = 2, L = 175, T = 12$ .

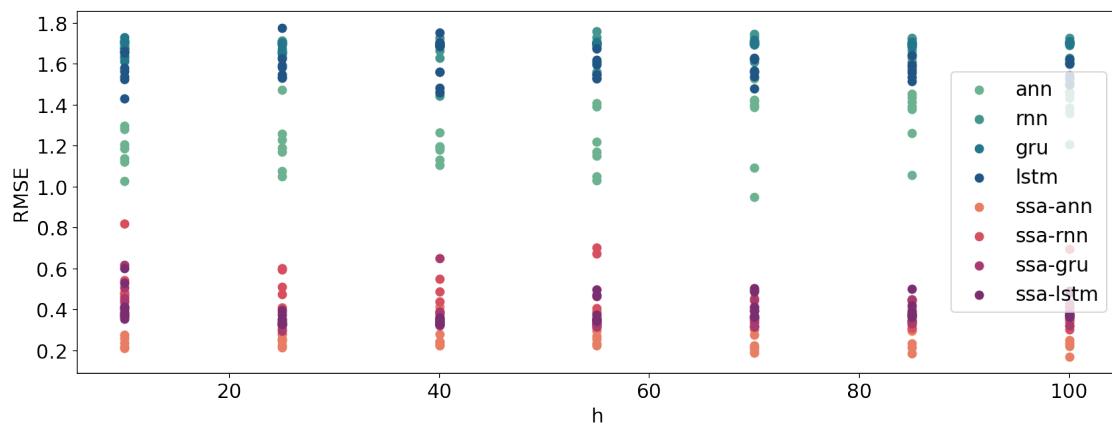


Рис. А.74. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$r = 2, L = 175, T = 84$ .

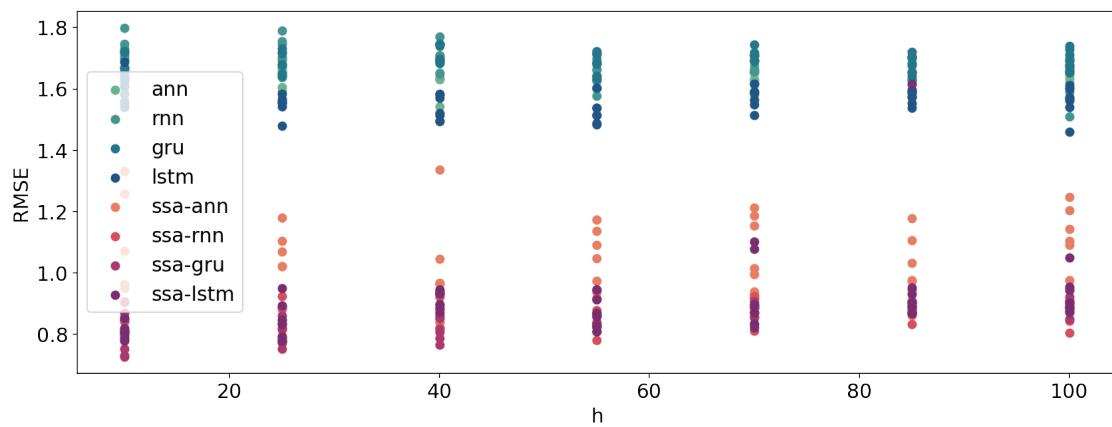


Рис. А.75. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$r = 4, L = 175, T = 12$

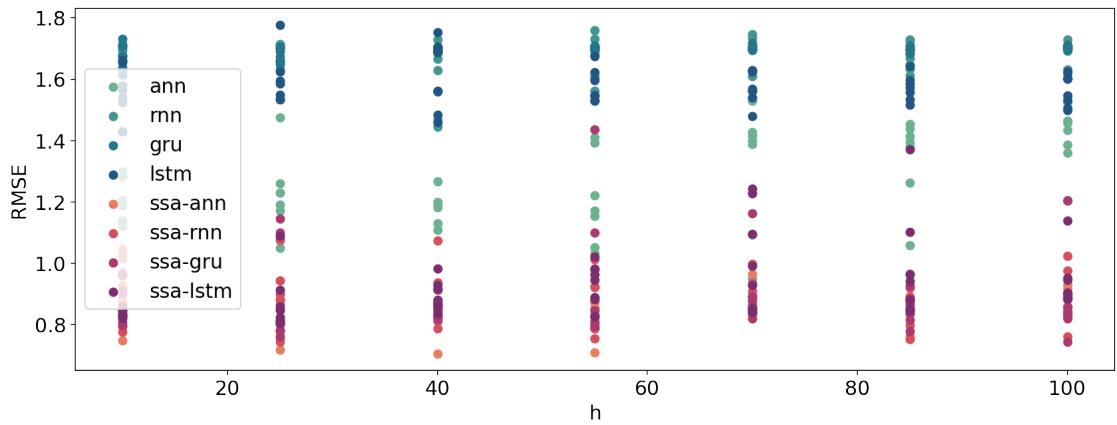


Рис. А.76. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$r = 4, L = 175, T = 84$ .

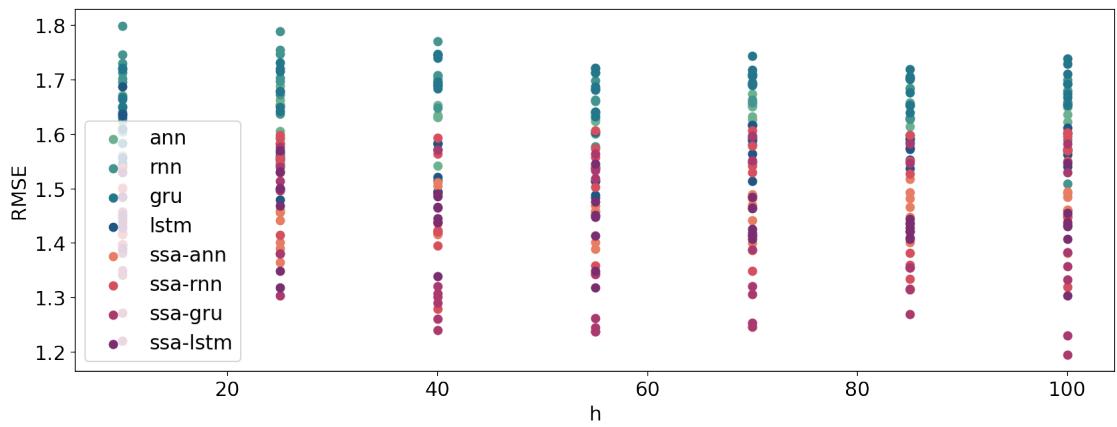


Рис. А.77. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$r = 14, L = 84, T = 12$

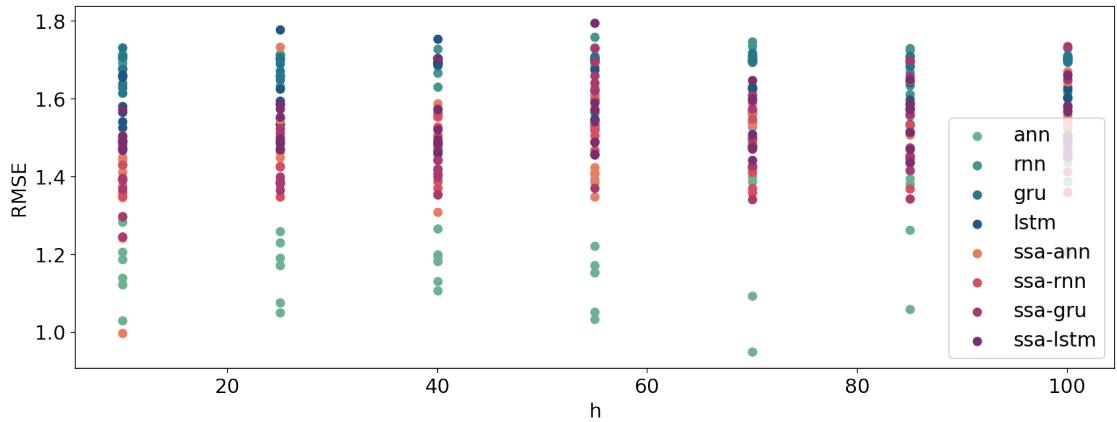


Рис. А.78. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$r = 14, L = 84, T = 84$ .

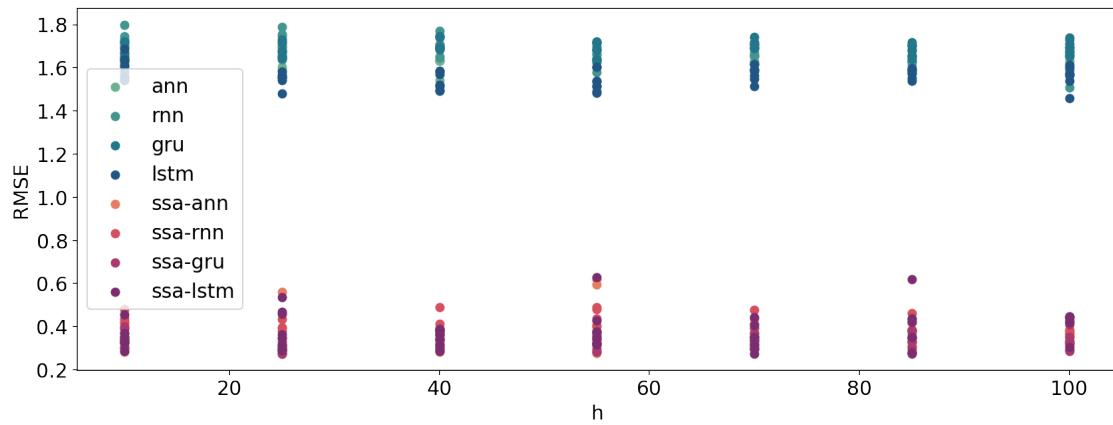


Рис. А.79. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$r = 2, L = 175, T = 12$ .

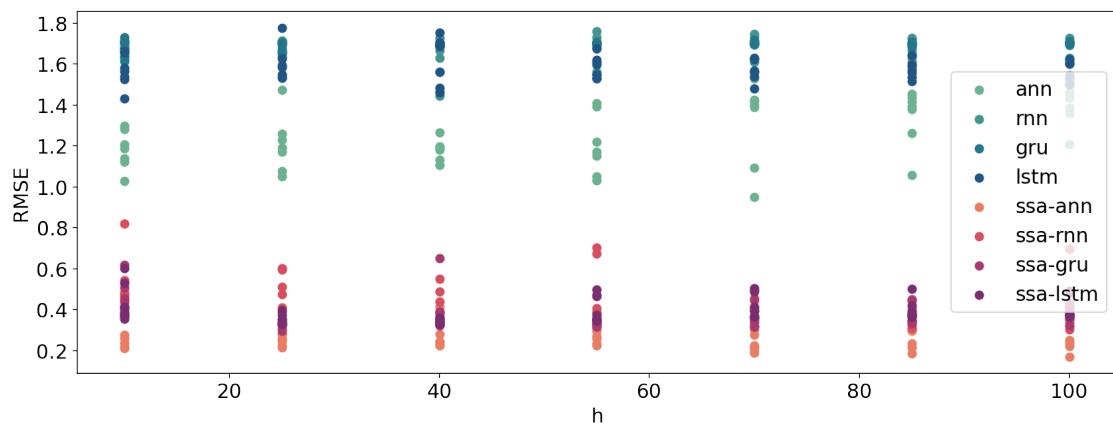


Рис. А.80. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$r = 2, L = 175, T = 84$ .

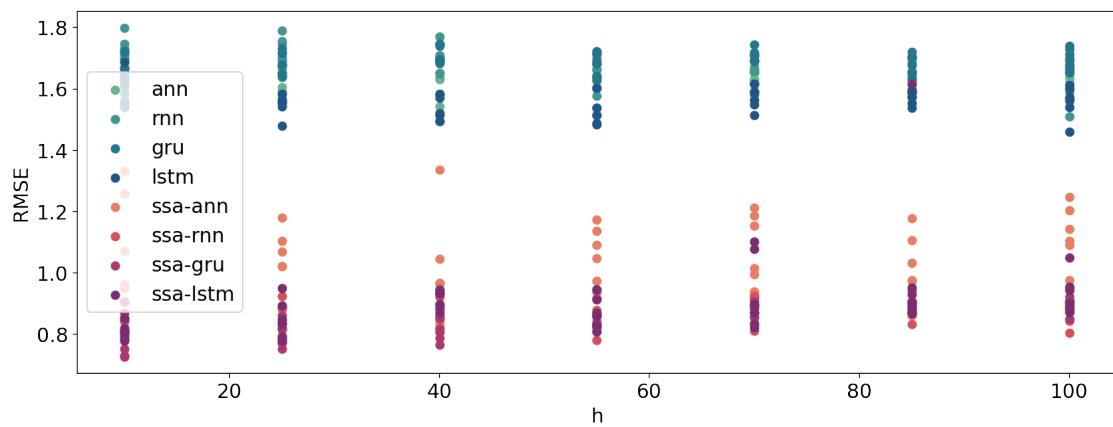


Рис. А.81. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$r = 4, L = 175, T = 12$

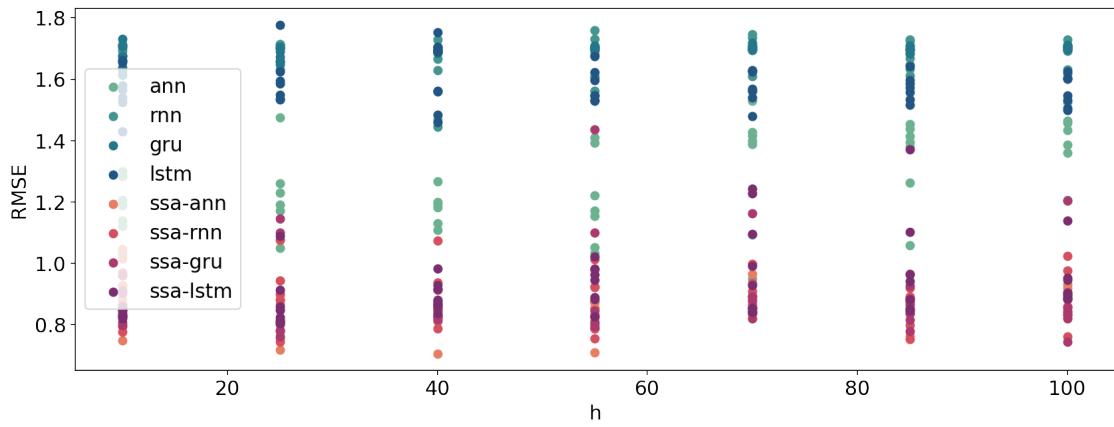


Рис. A.82. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$r = 4, L = 175, T = 84$ .

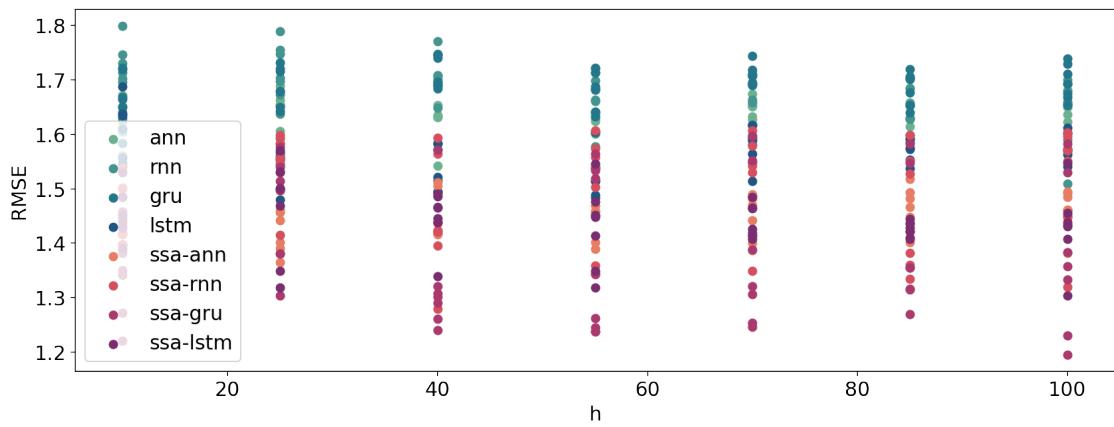


Рис. A.83. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$r = 14, L = 84, T = 12$

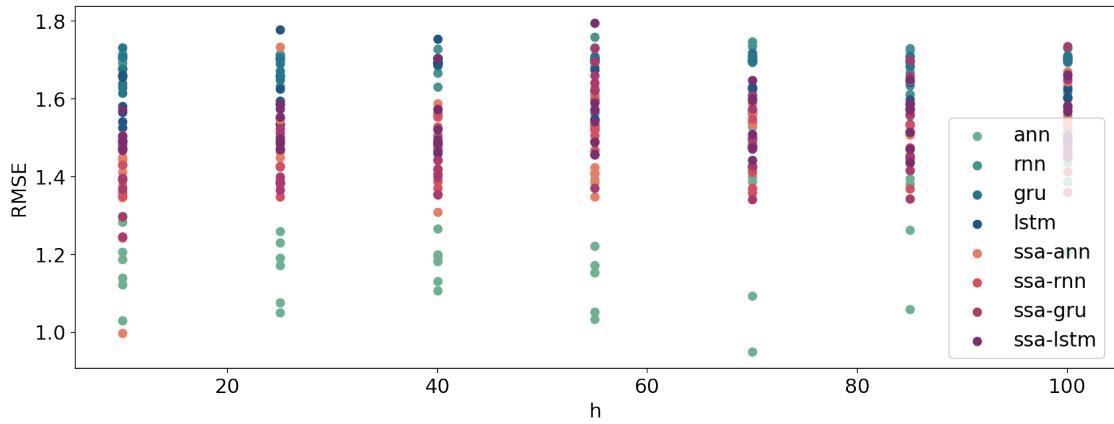


Рис. A.84. «Сумма синусов с красным шумом». Ряд  $Z_{650}$ . Проверка устойчивости.

$r = 14, L = 84, T = 84$ .

## Приложение Б

### Реальные данные

#### Б.1. Среднемесячные осадки в Индии

##### Отображение прогнозов

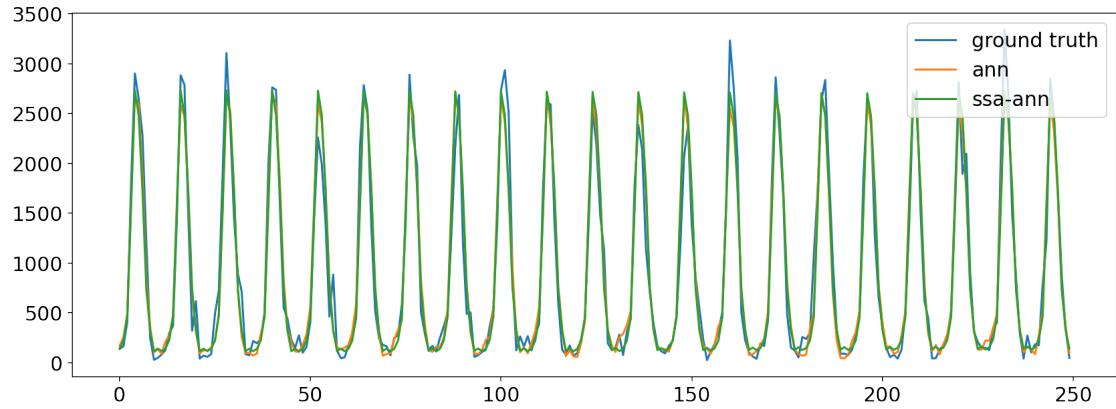


Рис. Б.1. Данные «Indian Rain». Прогноз для ANN и SSA-ANN.

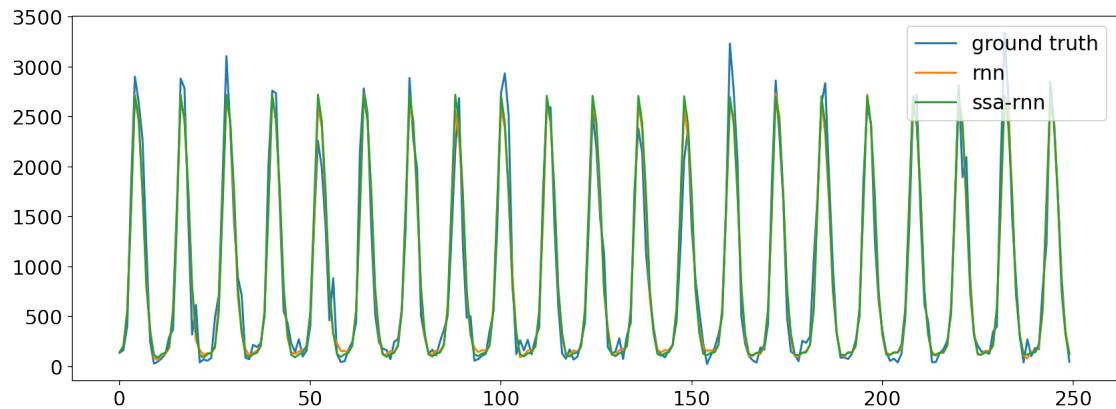


Рис. Б.2. Данные «Indian Rain». Прогноз для RNN и SSA-RNN.

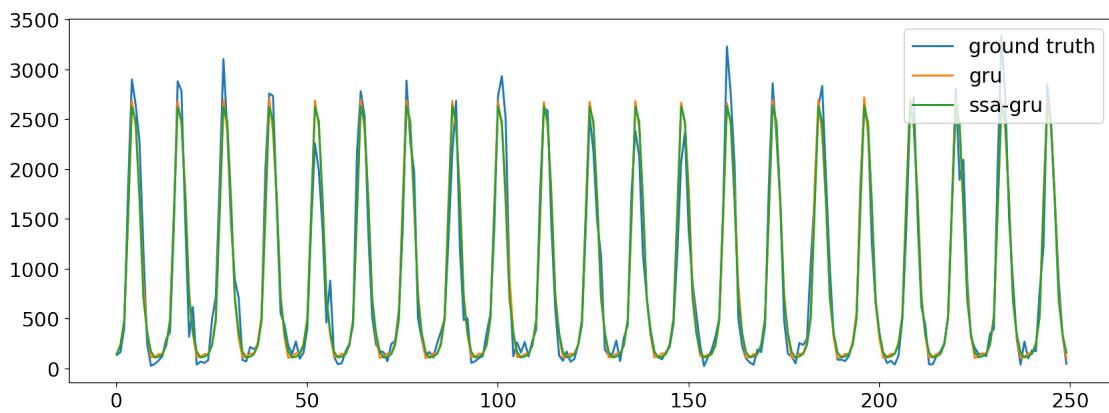


Рис. Б.3. Данные «Indian Rain». Прогноз для GRU и SSA-GRU.

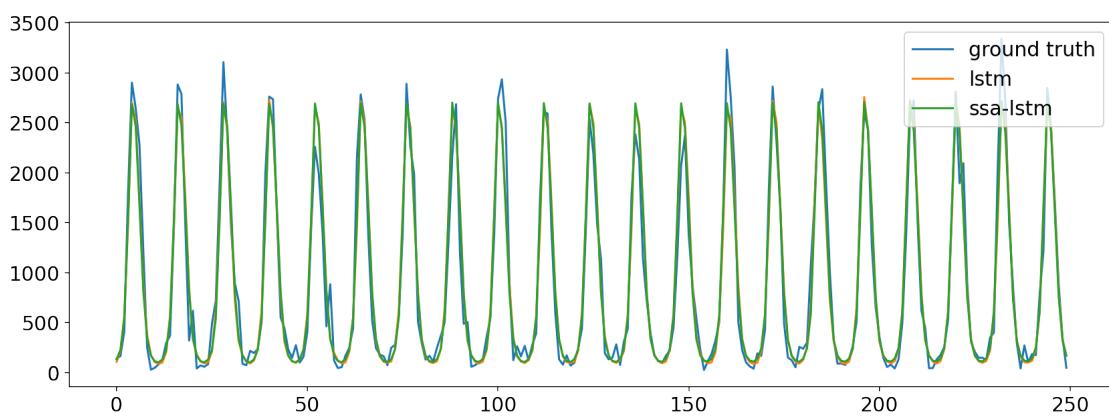


Рис. Б.4. Данные «Indian Rain». Прогноз для LSTM и SSA-LSTM.

## Б.2. Погода

### Отображение прогнозов

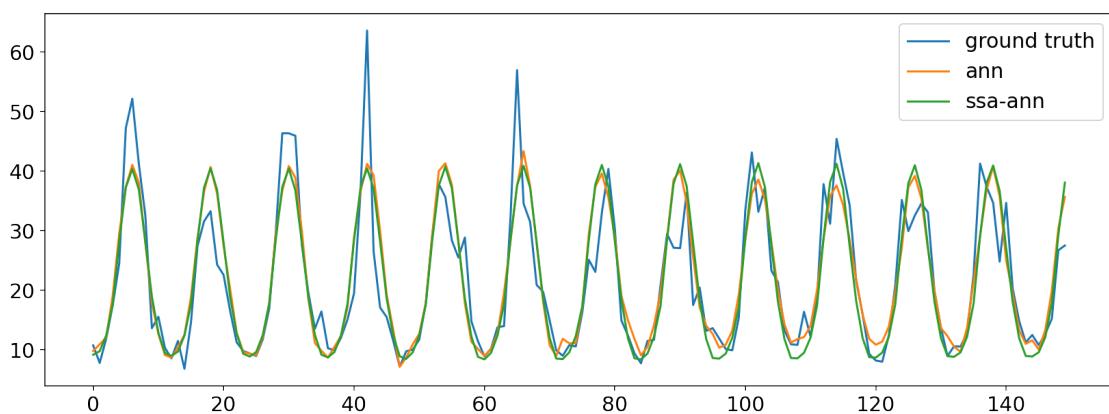


Рис. Б.5. Погода в Санкт-Петербурге. Прогноз для ANN и SSA-ANN.

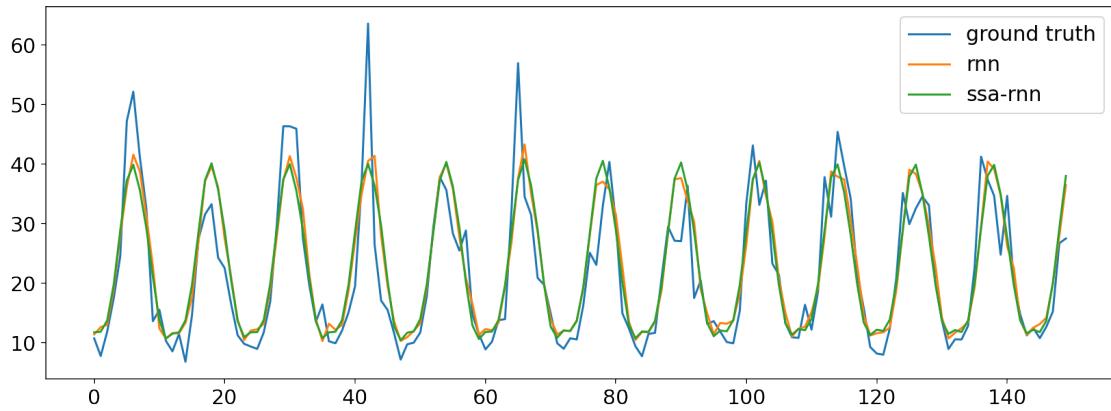


Рис. Б.6. Погода в Санкт-Петербурге. Прогноз для RNN и SSA-RNN.

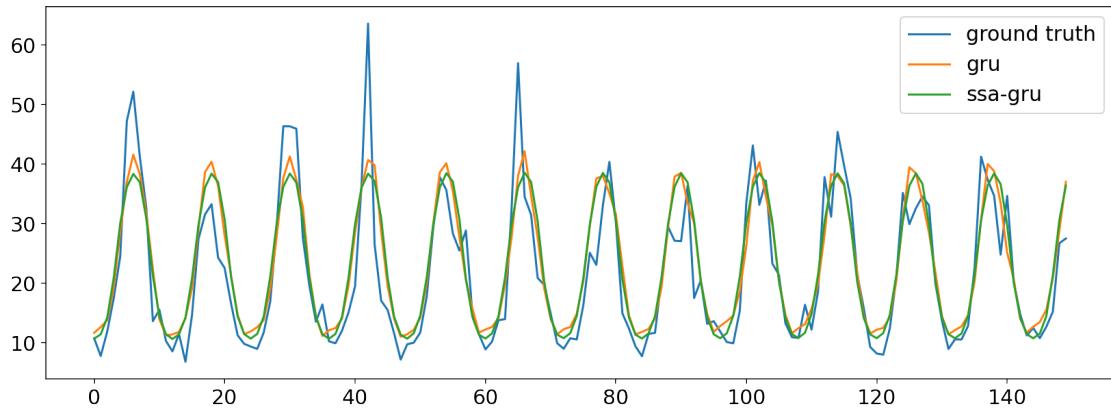


Рис. Б.7. Погода в Санкт-Петербурге. Прогноз для GRU и SSA-GRU.

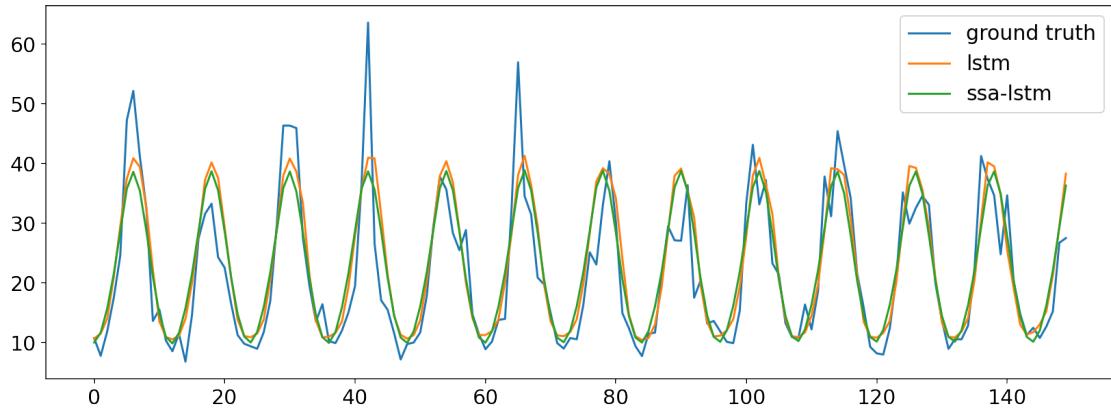


Рис. Б.8. Погода в Санкт-Петербурге. Прогноз для LSTM и SSA-LSTM.