

Санкт-Петербургский государственный университет

Прикладная математика и информатика

Производственная практика 2 (научно-исследовательская работа)

ИСПОЛЬЗОВАНИЕ МЕТОДА SSA В МАШИННОМ ОБУЧЕНИИ ДЛЯ  
ПРОГНОЗА ВРЕМЕННЫХ РЯДОВ

Выполнил:

Ежов Федор Валерьевич  
группа 20.М03-мм

Научный руководитель:

к.физ.-мат.н., доцент

Голяндина Нина Эдуардовна

Кафедра Статистического Моделирования

Санкт-Петербург

2022

# Оглавление

<b>Введение . . . . .</b>	4
<b>Глава 1. Алгоритм Singular Spectrum Analysis . . . . .</b>	5
1.1. Этап 1. Построение траекторной матрицы (Вложение) . . . . .	5
1.2. Этап 2. Singular Value Decomposition (SVD) . . . . .	6
1.3. Этап 3. Группировка первых $r$ собственных троек . . . . .	6
1.4. Этап 4. Диагональное усреднение . . . . .	6
<b>Глава 2. Архитектуры нейронных сетей . . . . .</b>	8
2.1. Artificial Neural Network (ANN) . . . . .	8
2.2. Recurrent neural network (RNN) . . . . .	9
2.3. Long short-term memory(LSTM) . . . . .	9
2.4. Gated Recurrent Unit (GRU) . . . . .	10
<b>Глава 3. Использование SSA в машинном обучении . . . . .</b>	11
3.1. Задача . . . . .	11
3.2. Подготовка данных . . . . .	11
3.3. Обучение нейронной сети . . . . .	13
3.4. Прогнозирование . . . . .	14
3.5. Метрики . . . . .	14
3.6. Методика применения и сравнения методов . . . . .	14
<b>Глава 4. Модельные данные . . . . .</b>	17
4.1. Сумма двух синусов с белым шумом . . . . .	17
4.2. Сумма двух синусов с красным шумом. Ряд с трудно отделимым сигналом	48
<b>Глава 5. Реальные данные . . . . .</b>	68
5.1. Среднемесячные осадки в Индии . . . . .	68
5.2. Данные Earth Orientation Parameters (EOP) . . . . .	77
5.3. Погода . . . . .	86
<b>Заключение . . . . .</b>	94

<b>Список литературы . . . . .</b>	95
------------------------------------	----

## Введение

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

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

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

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

## Глава 1

### Алгоритм Singular Spectrum Analysis

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

Алгоритм SSA состоит из четырех этапов:

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

Рассмотрим каждый этап подробнее.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

## 1.4. Этап 4. Диагональное усреднение

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

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

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

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

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

## Глава 2

# Архитектуры нейронных сетей

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

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

### 2.1. Artificial Neural Network (ANN)

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

Входные данные, на которых модель учится делать предсказания:

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

Предсказания модели:

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

Формула, описывающая модель:

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

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

Список некоторых функций активации:

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

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

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

## 2.2. Recurrent neural network (RNN)

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

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

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

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

Следующие формулы описывают модель:

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

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

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

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

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

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

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

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

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

Следующие формулы описывают модель:

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

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

Оператор  $*$  — производит поэлементное умножение.

## 2.4. Gated Recurrent Unit (GRU)

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

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

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

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

Следующие формулы описывают модель:

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

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

Оператор  $*$  — производит поэлементное умножение.

## Глава 3

# Использование SSA в машинном обучении

### 3.1. Задача

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

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

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

### 3.2. Подготовка данных

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

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

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

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

### SSA-preprocessing

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

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

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

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

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

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

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

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

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

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

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

### 3.3. Обучение нейронной сети

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

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

Алгоритм обучения модели после выбора архитектуры:

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

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

### 3.4. Прогнозирование

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

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

### 3.5. Метрики

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

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

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

### 3.6. Методика применения и сравнения методов

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

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

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

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

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

Таблица 3.1. Сетка гиперпараметров.

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

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

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

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

## Глава 4

### Модельные данные

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

1. Влияние выбора параметра  $r$ .
2. Влияние наличия шума во временном ряде.

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

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

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

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

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

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

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

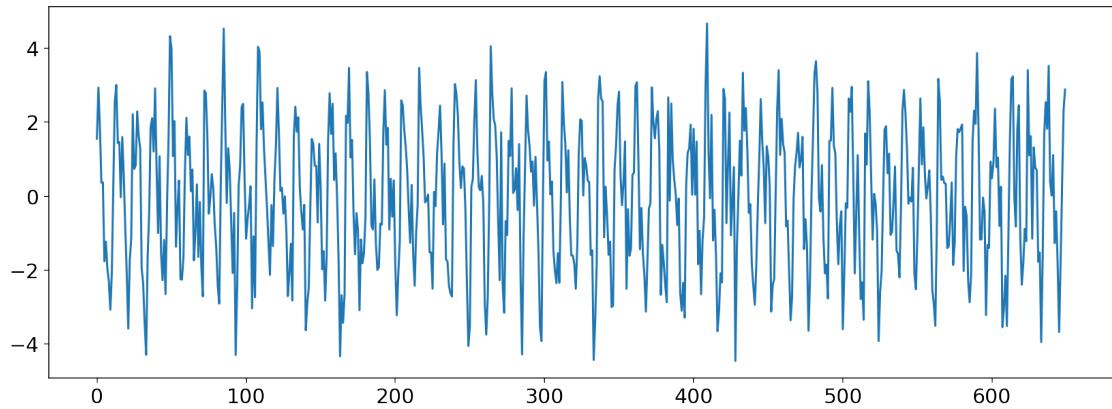


Рис. 4.1. Ряд суммы синусов с белым шумом. Ряд  $Z_{650}$ .

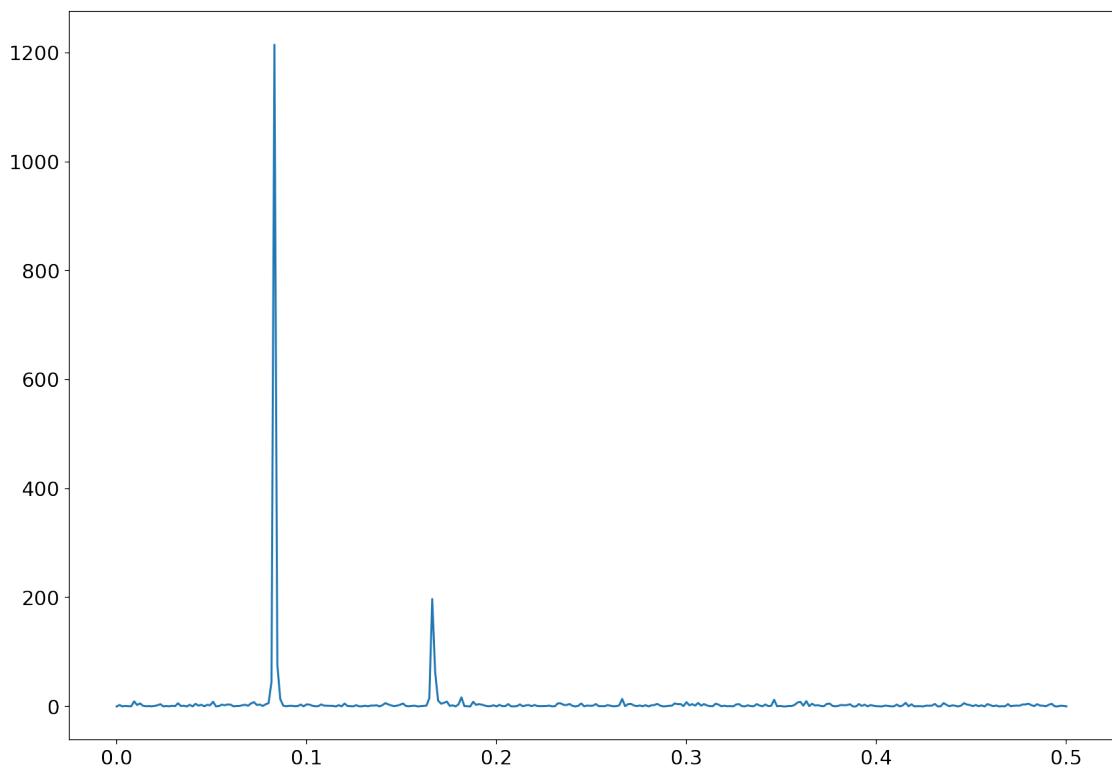


Рис. 4.2. Периодограмма ряда суммы синусов с белым шумом. Ряд  $Z_{650}$ .

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

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

## Прогноз по SSA

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

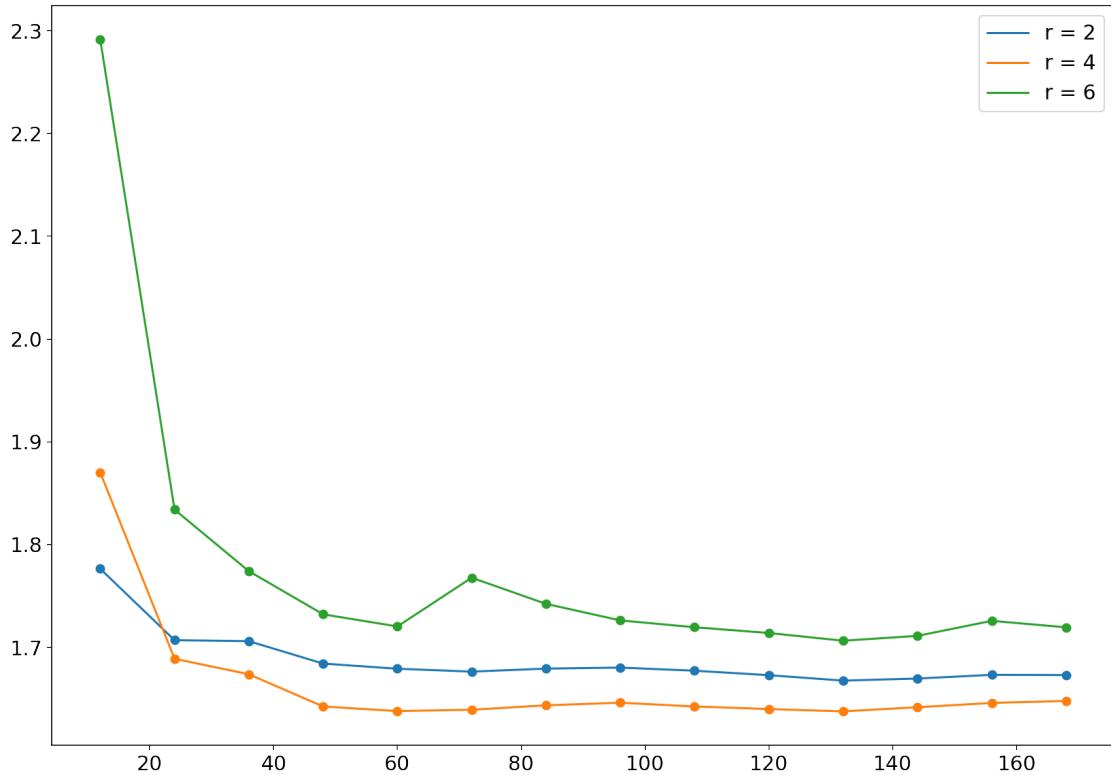


Рис. 4.3. Сумма синусов с белым шумом. Прогнозирование с помощью метода SSA.

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

## Восстановление SSA

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

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

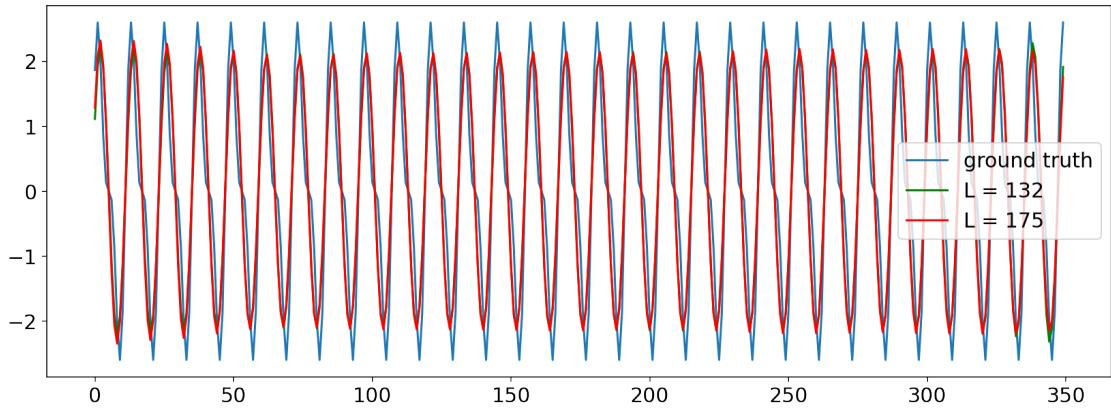


Рис. 4.4. Сумма синусов с белым шумом. Восстановление тренировочной выборки с помощью метода SSA.  $r = 2$

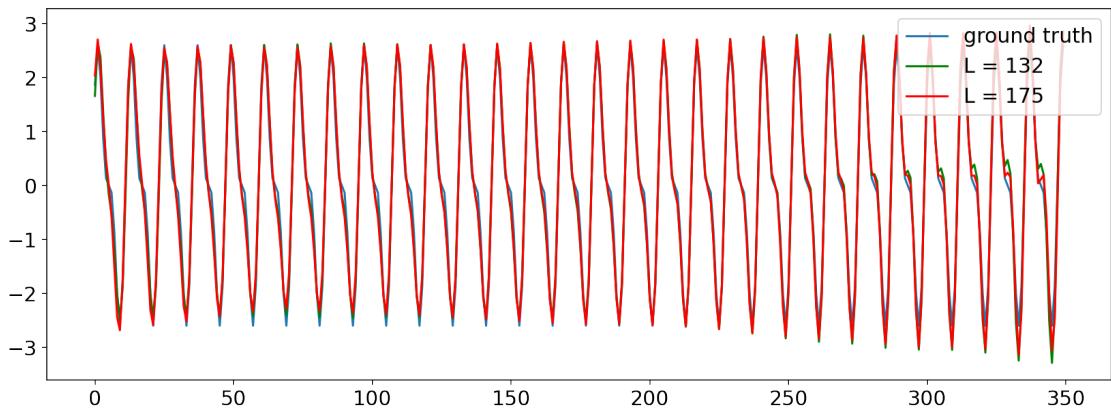


Рис. 4.5. Сумма синусов с белым шумом. Восстановление тренировочной выборки с помощью метода SSA.  $r = 4$

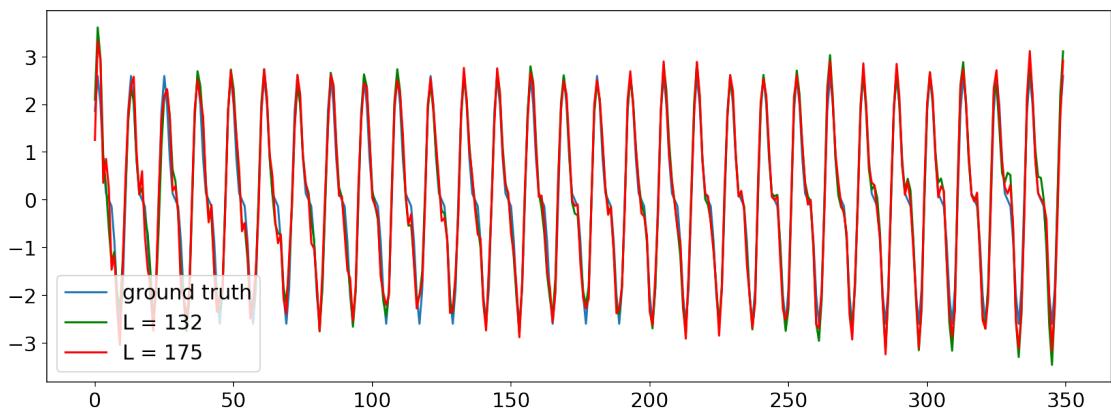


Рис. 4.6. Сумма синусов с белым шумом. Восстановление тренировочной выборки с помощью метода SSA.  $r = 6$

## Сравнение обычных и гибридных методов

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

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

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

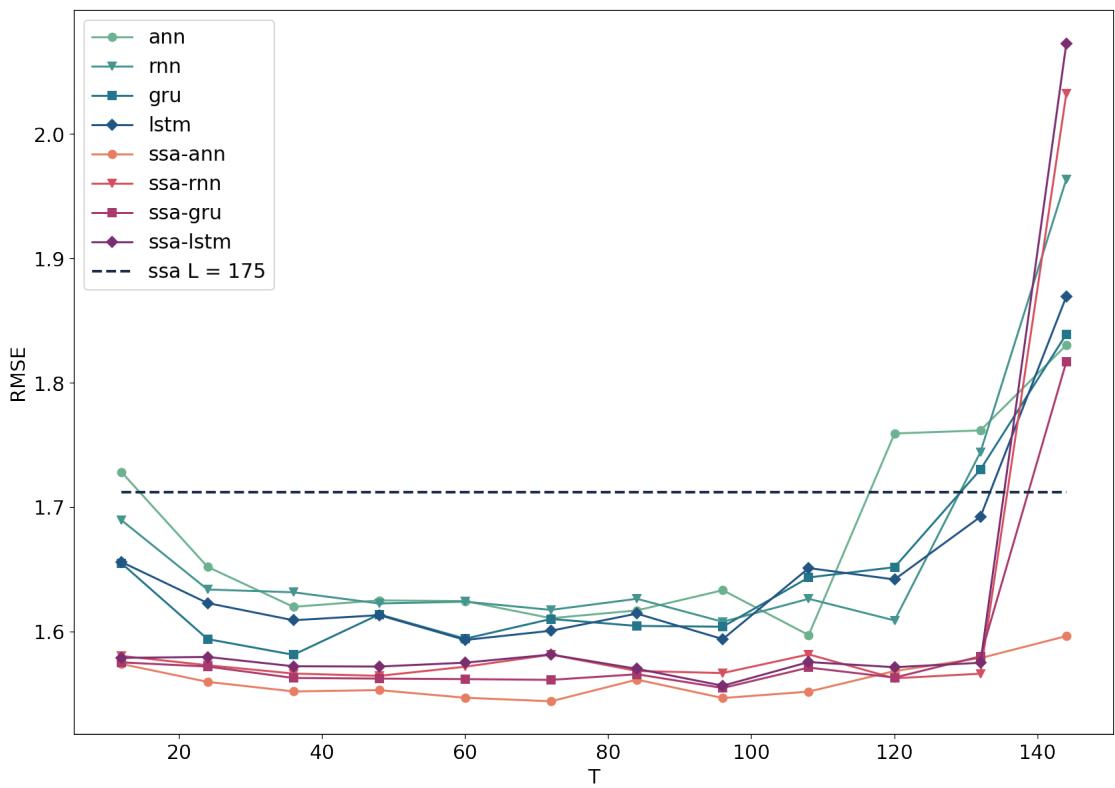


Рис. 4.7. Сумма синусов с белым шумом. Ошибки прогноза в зависимости от параметра  $T$ .

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

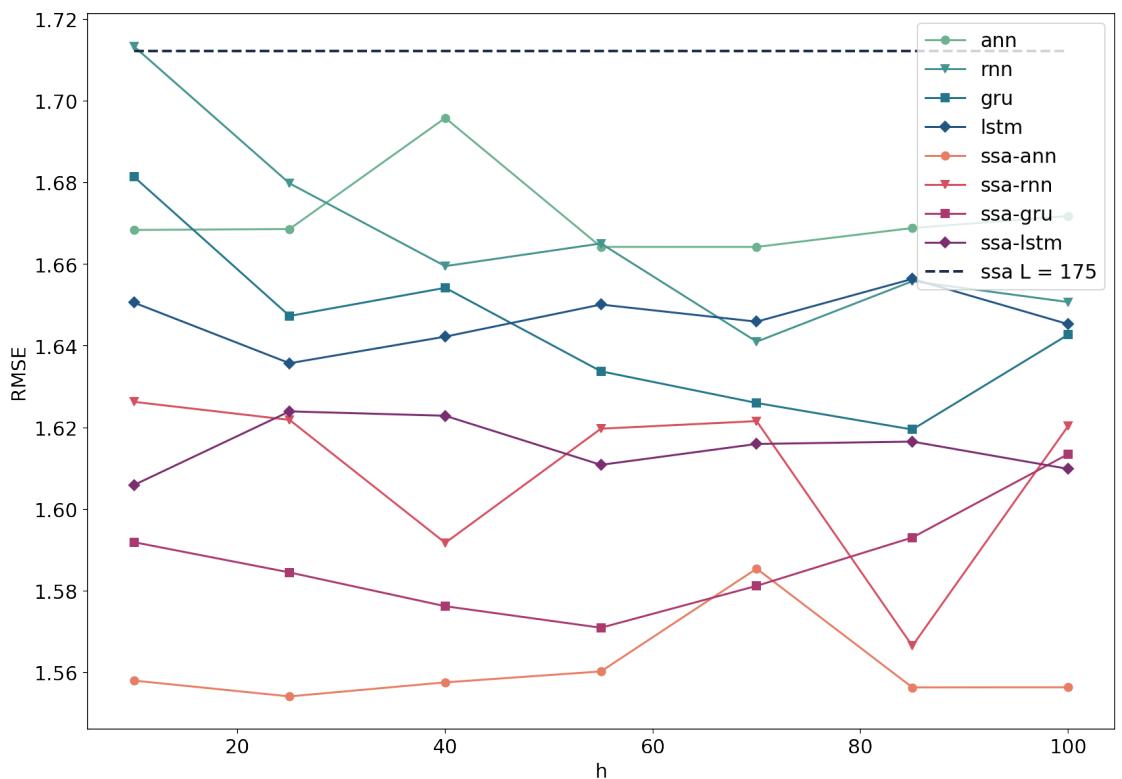


Рис. 4.8. Сумма синусов с белым шумом. Ошибки прогноза в зависимости от параметра  $h$ .

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

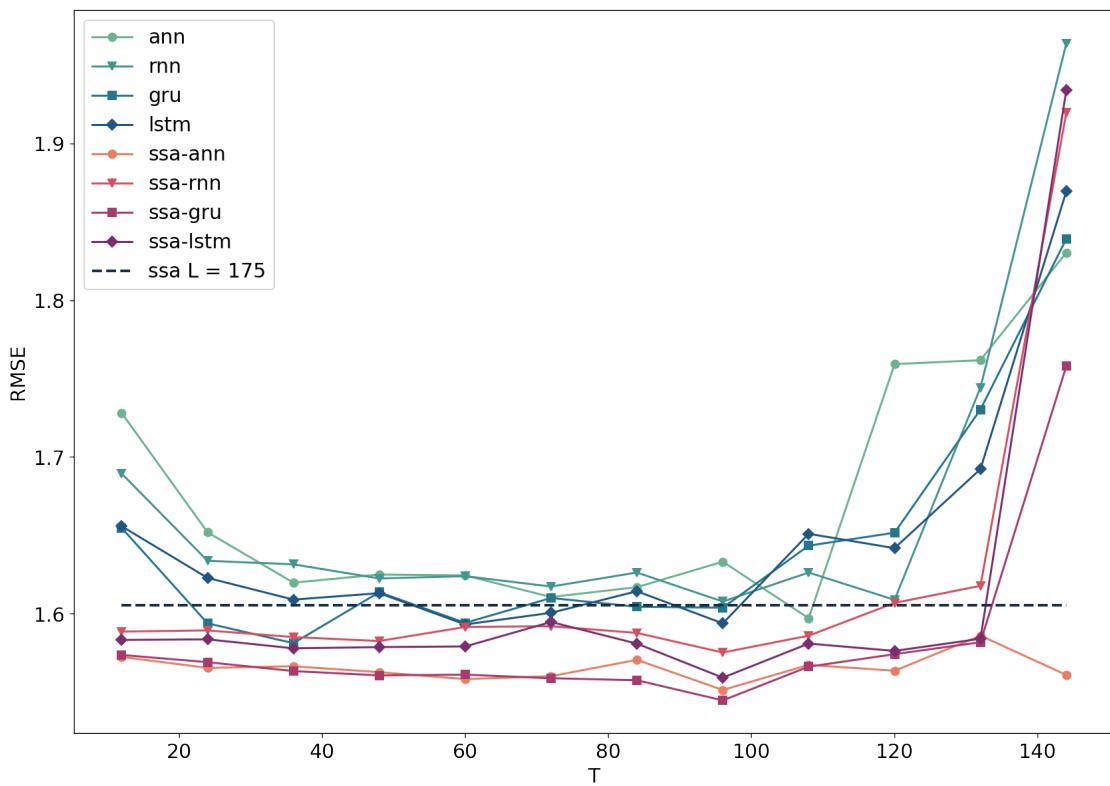


Рис. 4.9. Сумма синусов с белым шумом. Ошибки прогноза в зависимости от параметра  $T$ .

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

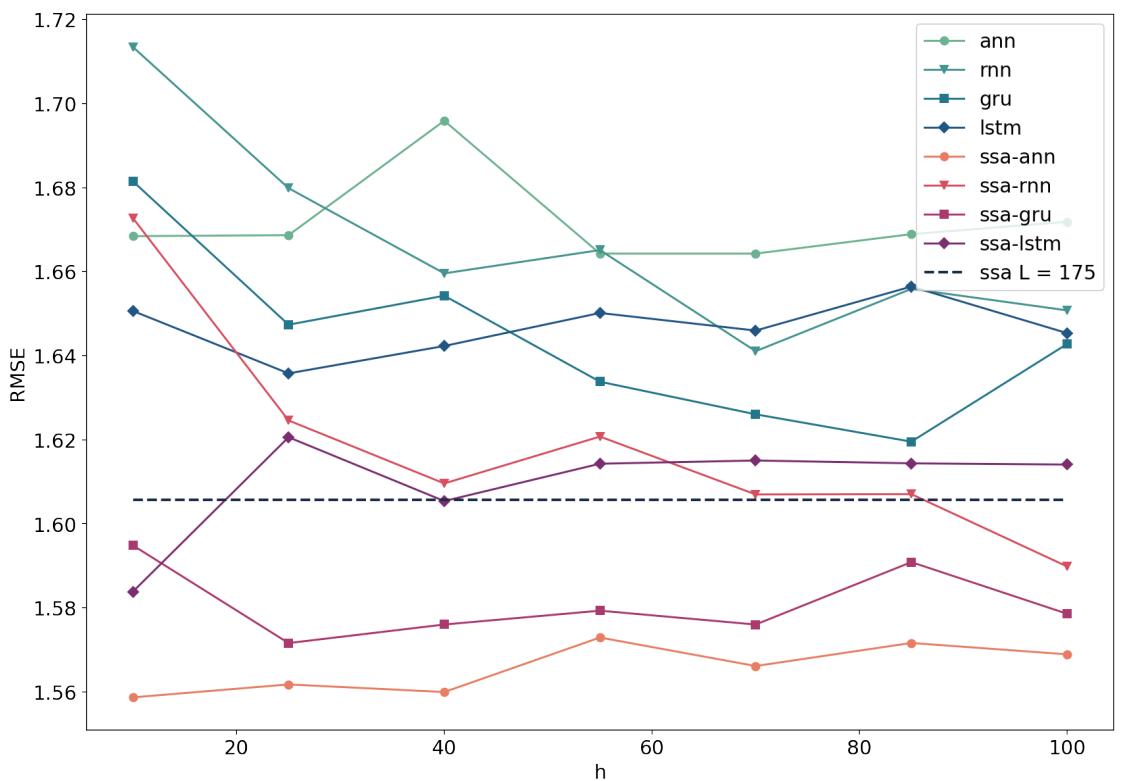


Рис. 4.10. Сумма синусов с белым шумом. Ошибки прогноза в зависимости от параметра  $h$ .

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

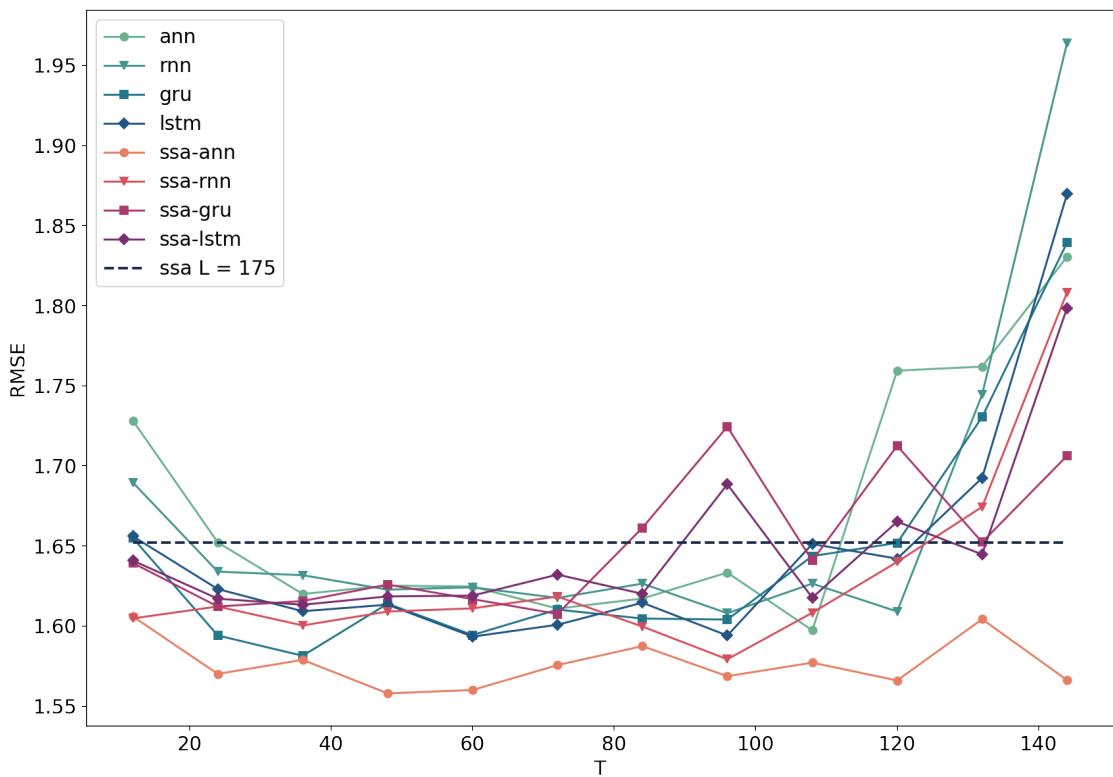


Рис. 4.11. Сумма синусов с белым шумом. Ошибки прогноза в зависимости от параметра  $T$ .

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

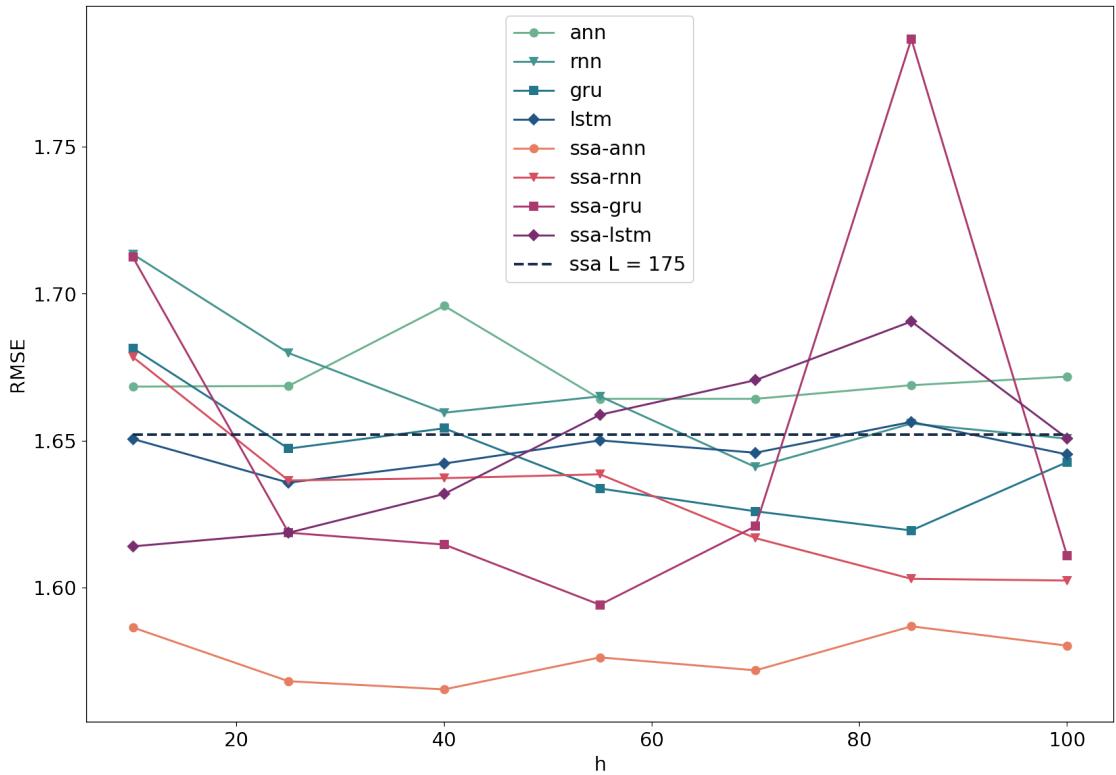


Рис. 4.12. Сумма синусов с белым шумом. Ошибки прогноза в зависимости от параметра  $h$ .

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

### Отображение прогнозов

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

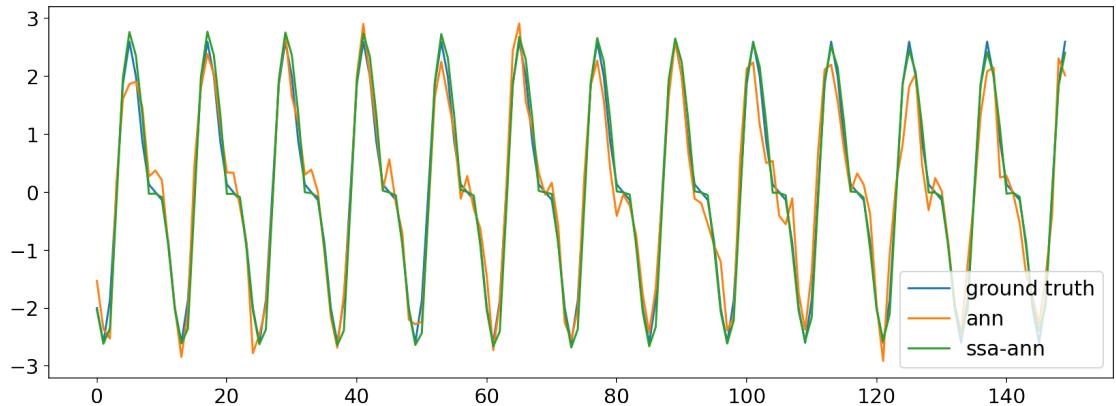


Рис. 4.13. Сумма синусов с белым шумом. Прогноз для ANN и SSA-ANN.  $r = 2$

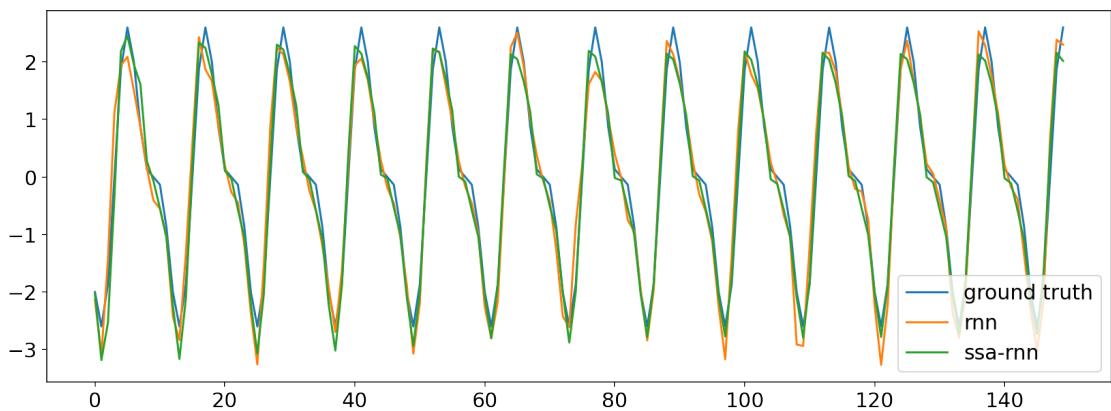


Рис. 4.14. Сумма синусов с белым шумом. Прогноз для RNN и SSA-RNN.  $r = 2$

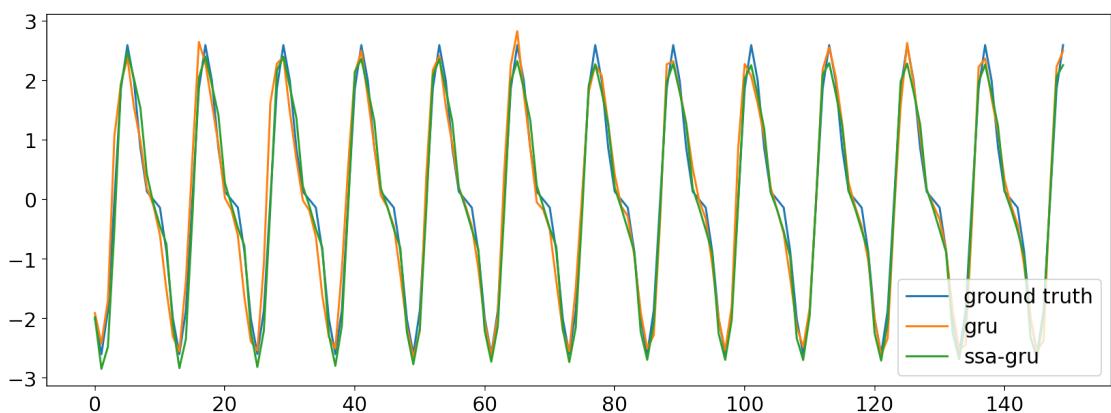


Рис. 4.15. Сумма синусов с белым шумом. Прогноз для GRU и SSA-GRU.  $r = 2$

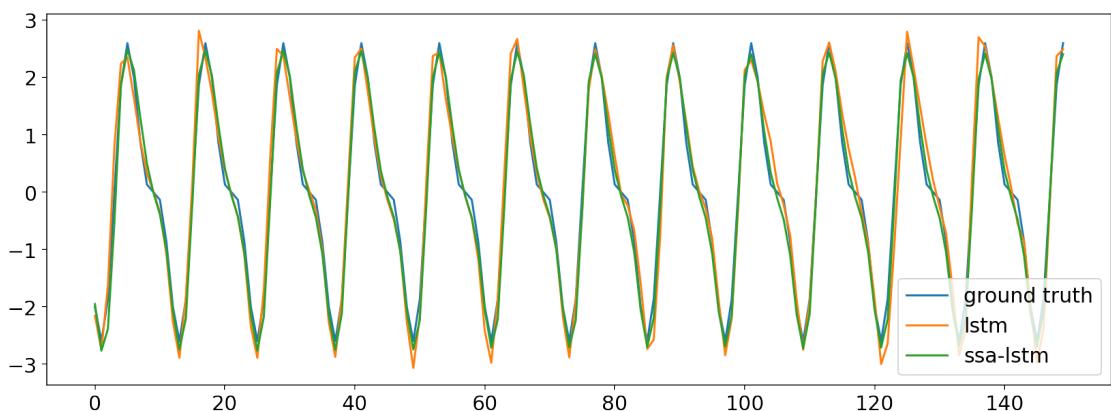


Рис. 4.16. Сумма синусов с белым шумом. Прогноз для LSTM и SSA-LSTM.  $r = 2$

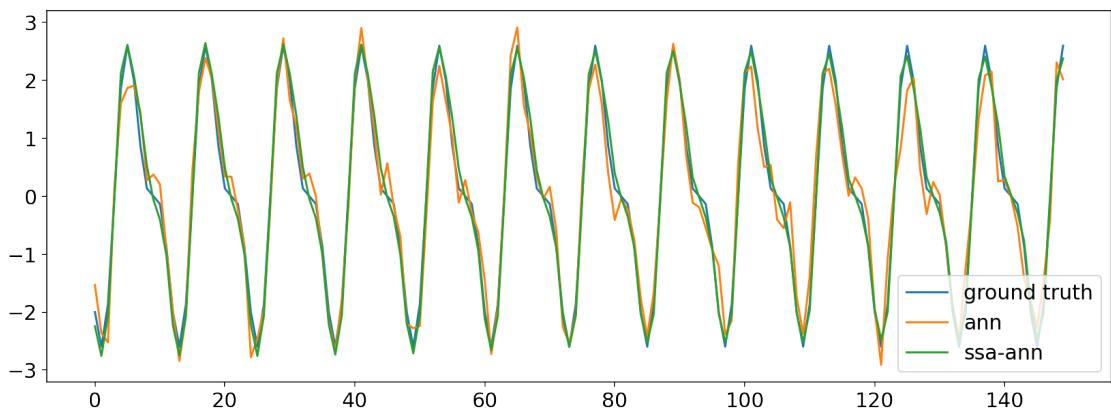


Рис. 4.17. Сумма синусов с белым шумом. Прогноз для ANN и SSA-ANN.  $r = 4$

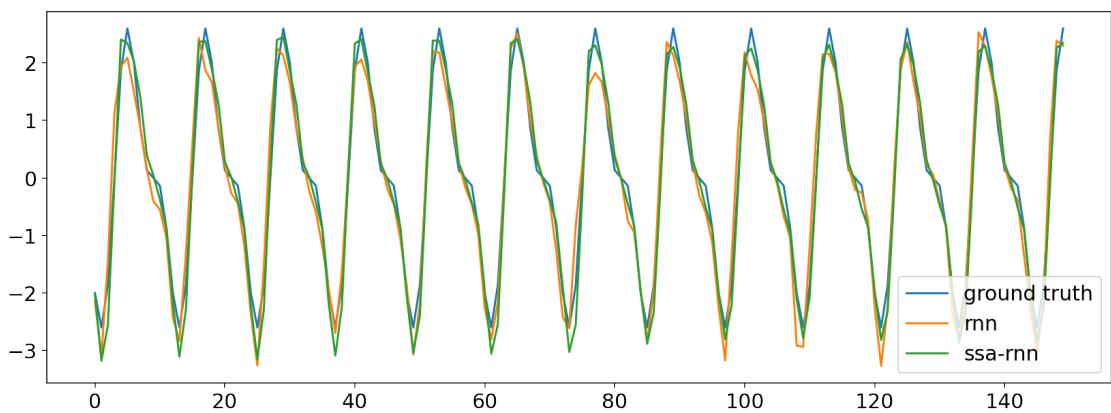


Рис. 4.18. Сумма синусов с белым шумом. Прогноз для RNN и SSA-RNN.  $r = 4$

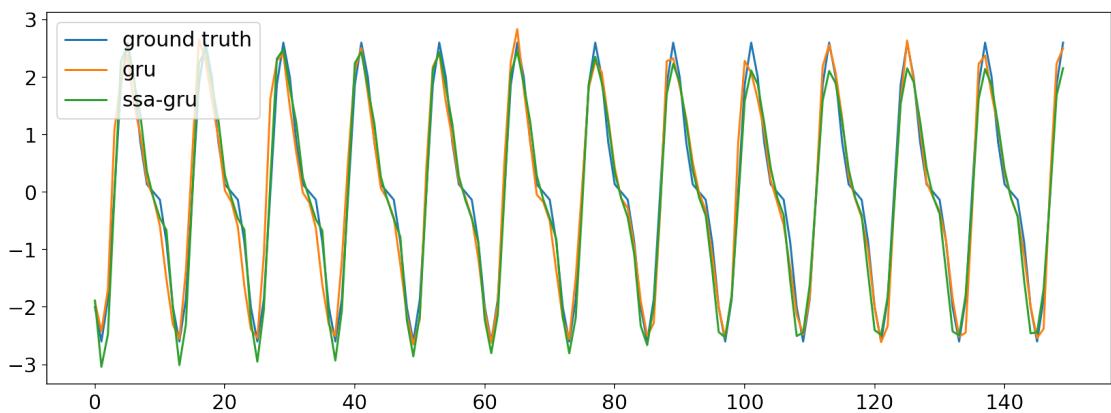


Рис. 4.19. Сумма синусов с белым шумом. Прогноз для GRU и SSA-GRU.  $r = 4$

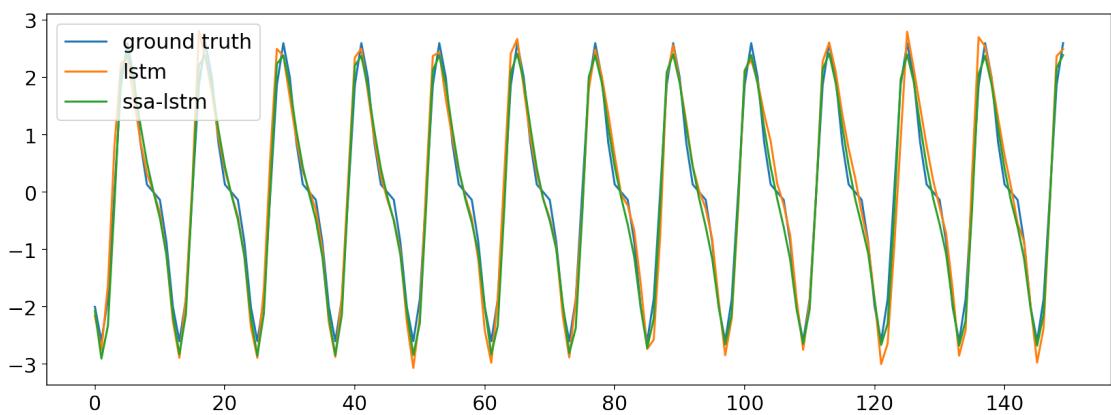


Рис. 4.20. Сумма синусов с белым шумом. Прогноз для LSTM и SSA-LSTM.  $r = 4$

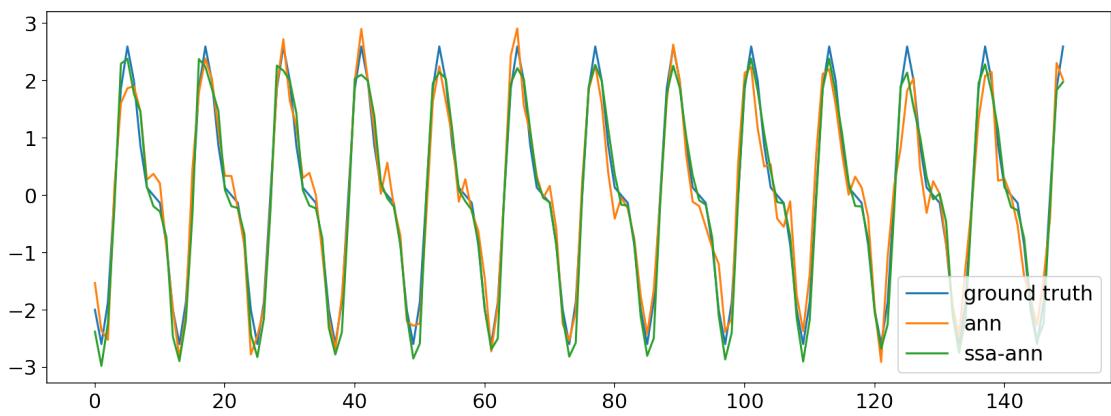


Рис. 4.21. Сумма синусов с белым шумом. Прогноз для ANN и SSA-ANN.  $r = 6$

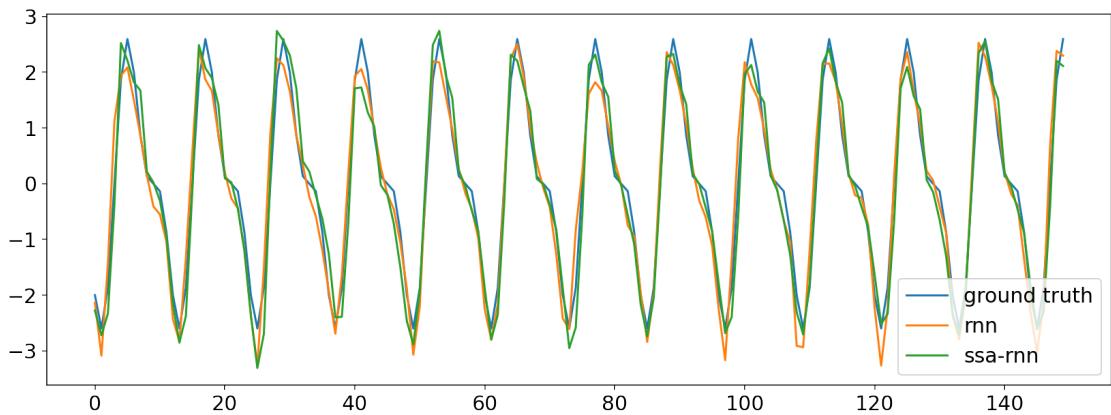


Рис. 4.22. Сумма синусов с белым шумом. Прогноз для RNN и SSA-RNN.  $r = 6$

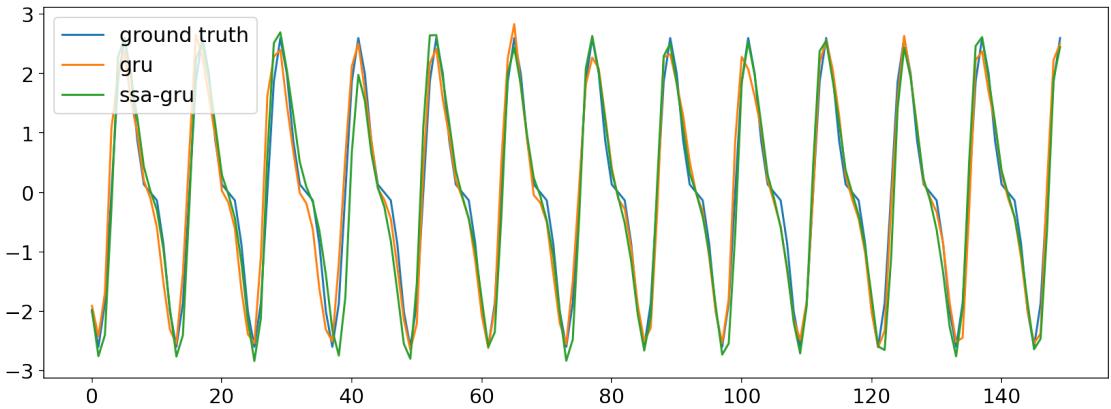


Рис. 4.23. Сумма синусов с белым шумом. Прогноз для GRU и SSA-GRU.  $r = 6$

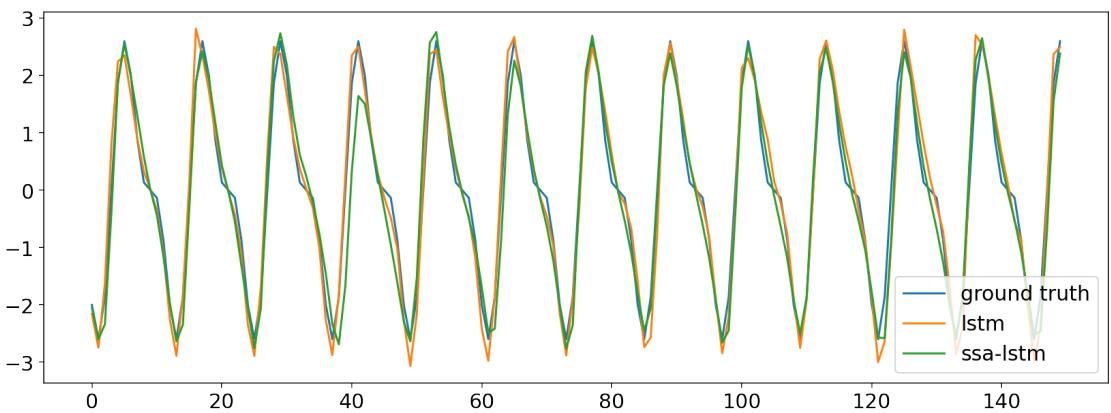


Рис. 4.24. Сумма синусов с белым шумом. Прогноз для LSTM и SSA-LSTM.  $r = 6$

## Проверка устойчивости

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

## Выводы

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

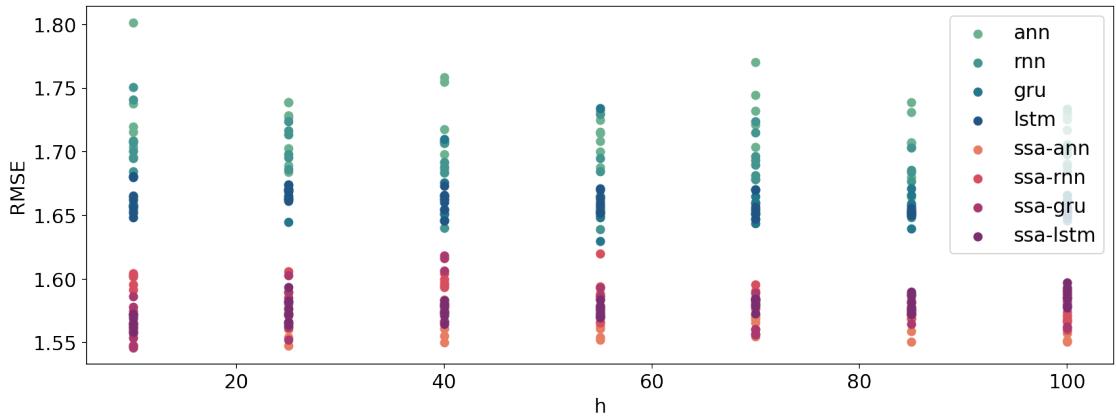


Рис. 4.25. Сумма синусов с белым шумом. Проверка устойчивости.  $r = 2$ ,  $L = 175$ .  $T = 12$ .

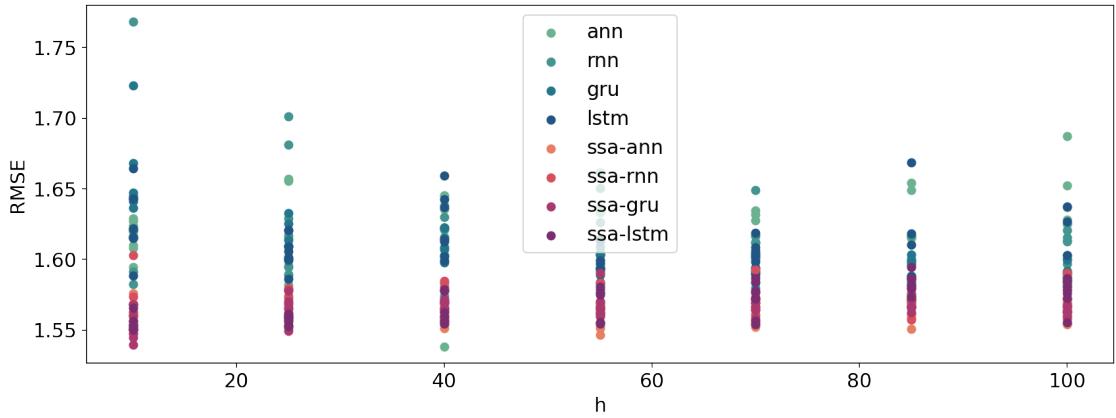


Рис. 4.26. Сумма синусов с белым шумом. Проверка устойчивости.  $r = 2$ ,  $L = 175$ .  $T = 84$ .

оптимальным.

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

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

#### Прогноз по SSA

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

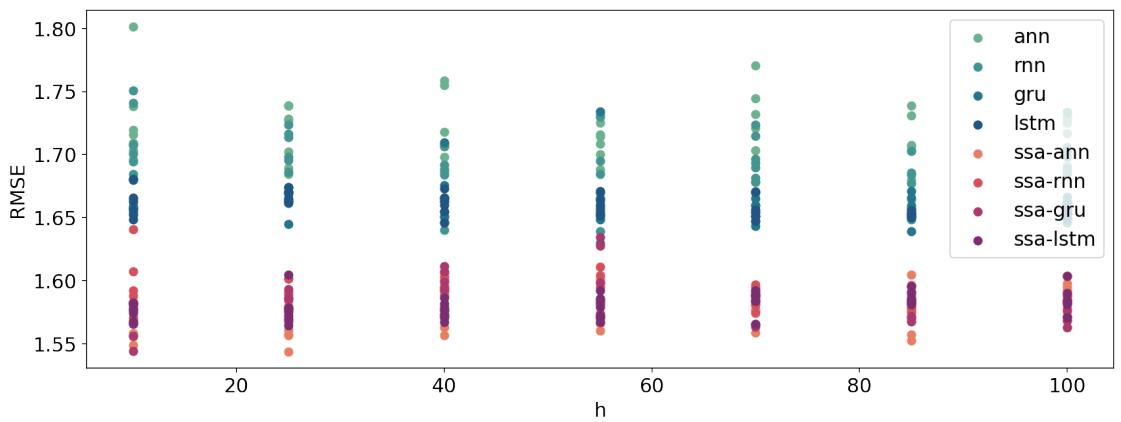


Рис. 4.27. Сумма синусов с белым шумом. Проверка устойчивости.  $r = 4$ ,  $L = 175$ .  $T = 12$ .

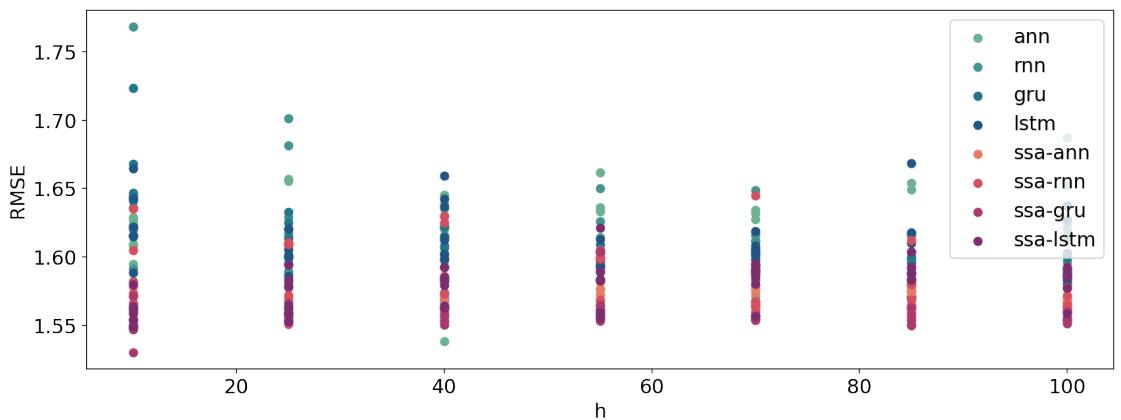


Рис. 4.28. Сумма синусов с белым шумом. Проверка устойчивости.  $r = 4$ ,  $L = 175$ .  $T = 84$ .

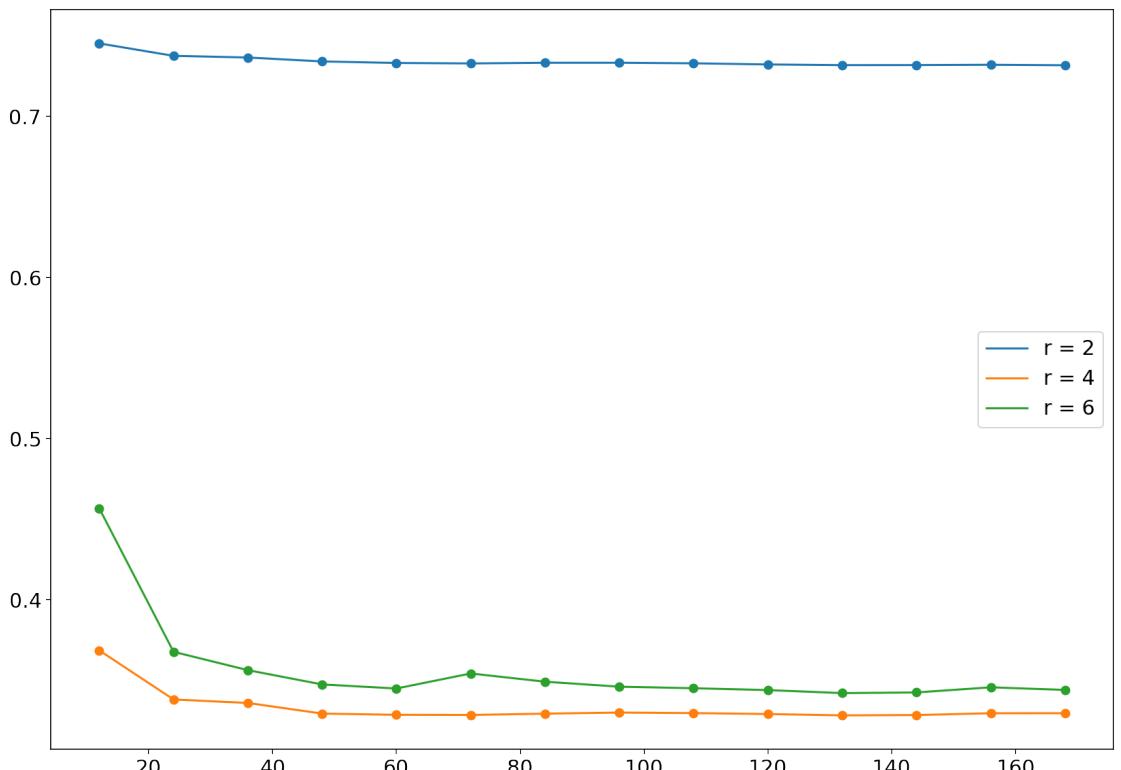


Рис. 4.31. Сумма синусов с небольшим белым шумом. Прогнозирование с помощью метода SSA.

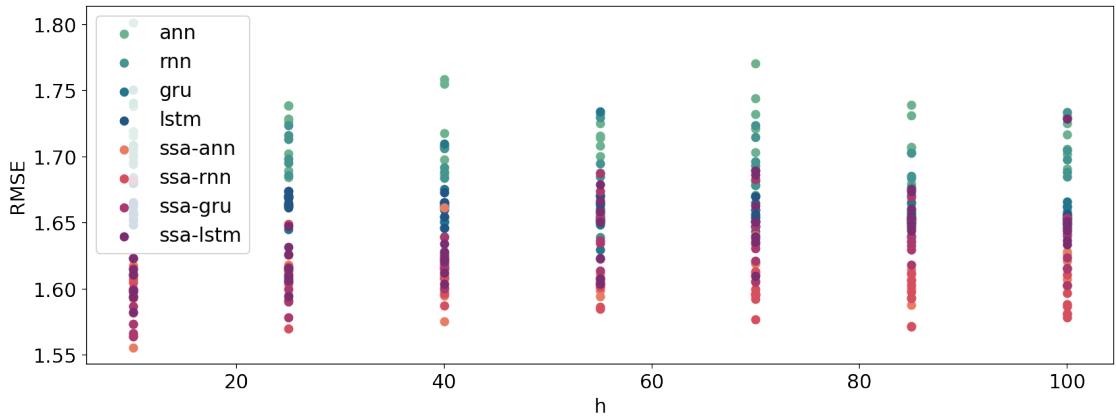


Рис. 4.29. Сумма синусов с белым шумом. Проверка устойчивости.  $r = 6$ ,  $L = 175$ .  $T = 12$ .

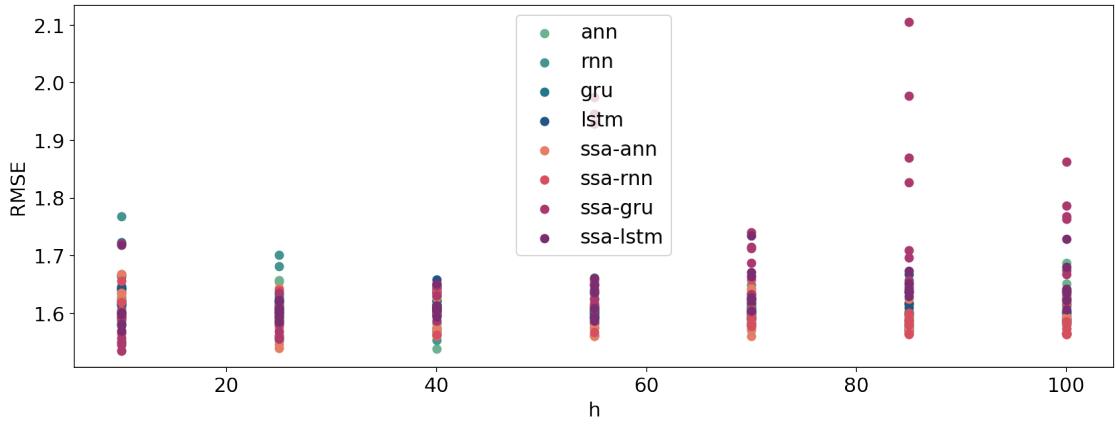


Рис. 4.30. Сумма синусов с белым шумом. Проверка устойчивости.  $r = 6$ ,  $L = 175$ .  $T = 84$ .

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

## Восстановление SSA

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

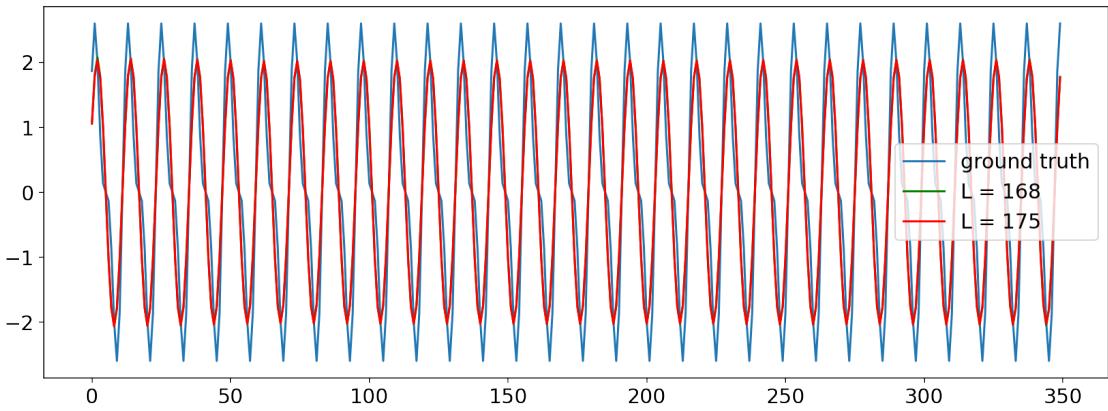


Рис. 4.32. Сумма синусов с небольшим белым шумом. Восстановление тренировочной выборки с помощью метода SSA.  $r = 2$

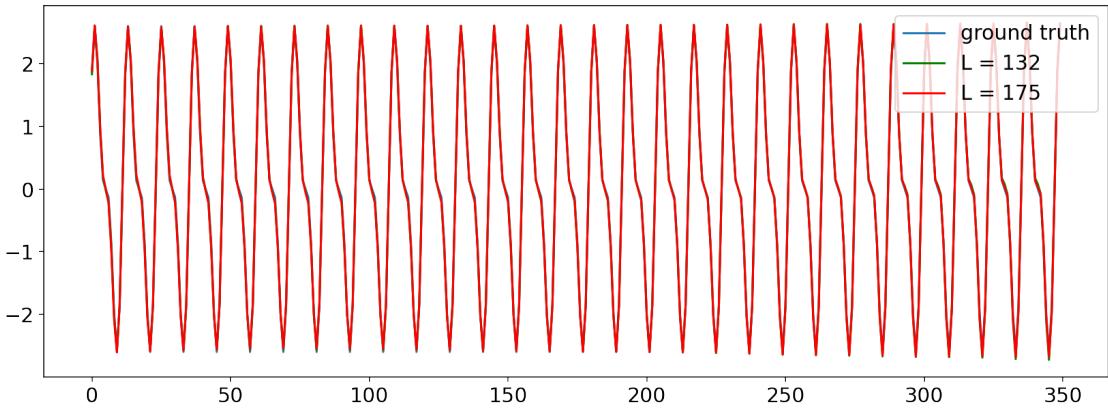


Рис. 4.33. Сумма синусов с небольшим белым шумом. Восстановление тренировочной выборки с помощью метода SSA.  $r = 4$

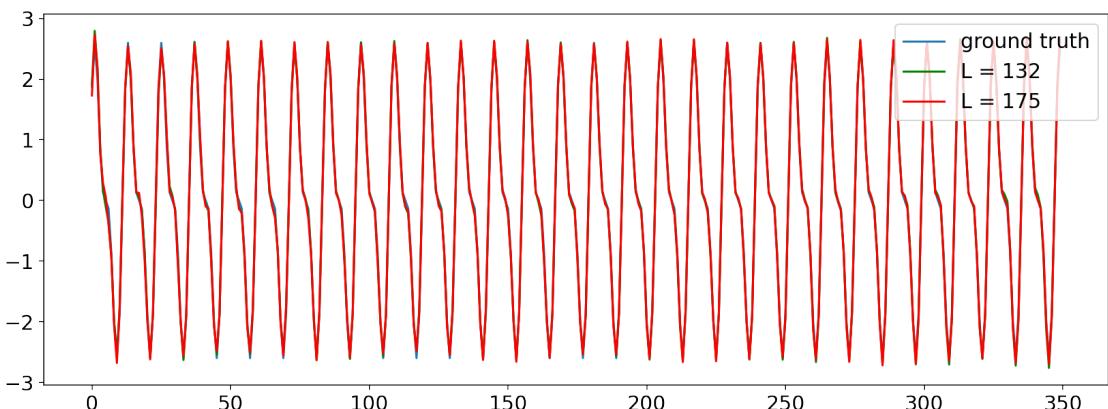


Рис. 4.34. Сумма синусов с небольшим белым шумом. Восстановление тренировочной выборки с помощью метода SSA.  $r = 6$

## Сравнение обычных и гибридных методов

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

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

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

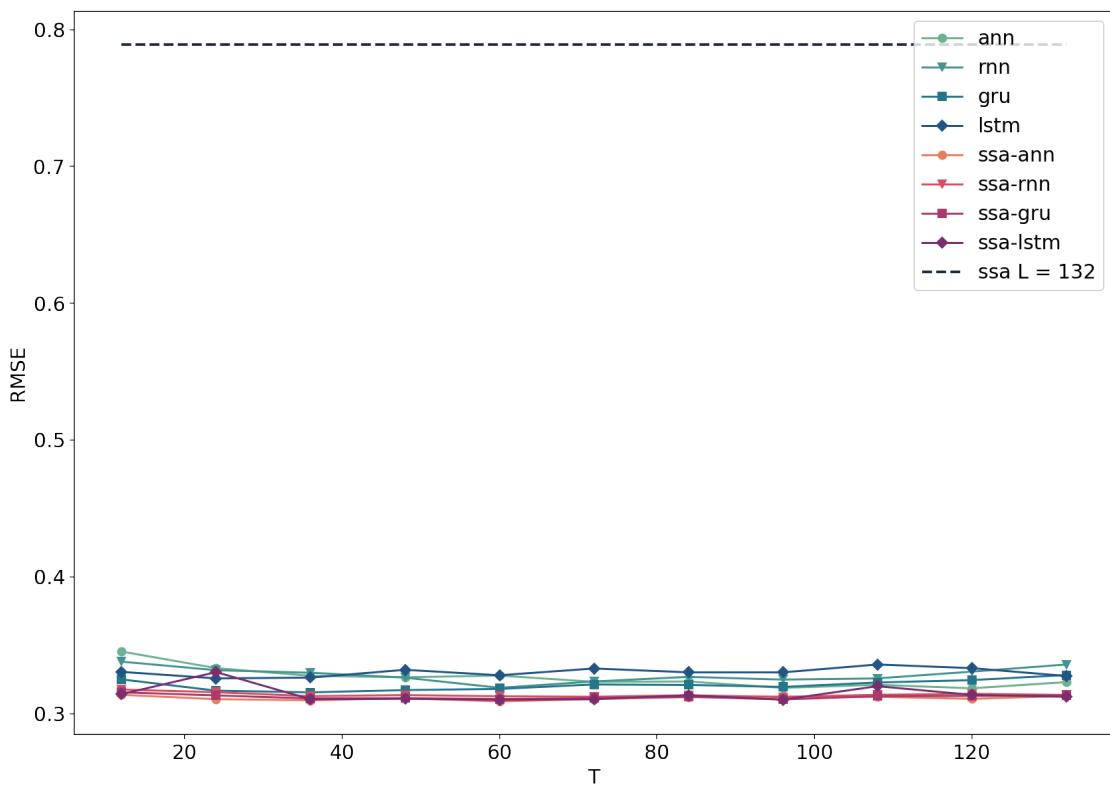


Рис. 4.35. Сумма синусов с небольшим белым шумом. Ошибки прогноза в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 2$ .

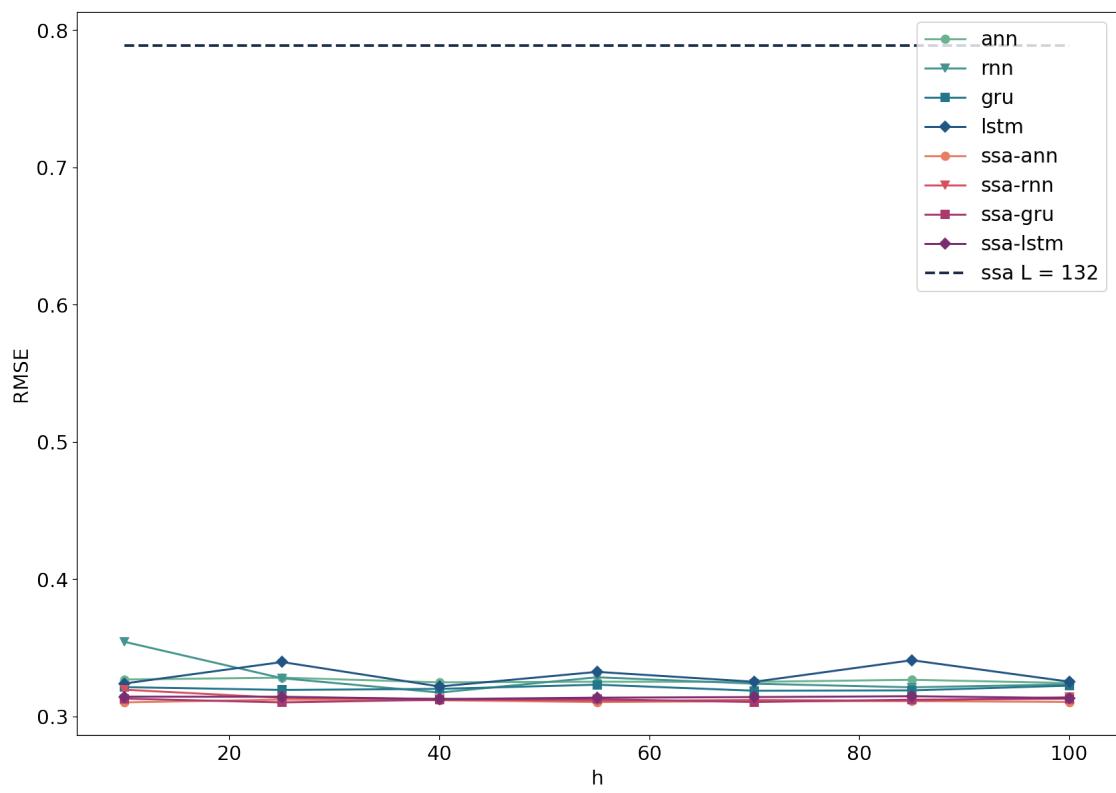


Рис. 4.36. Сумма синусов с небольшим белым шумом. Ошибки прогноза в зависимости от параметра  $h$ .  $L = 175$ ,  $r = 2$ .

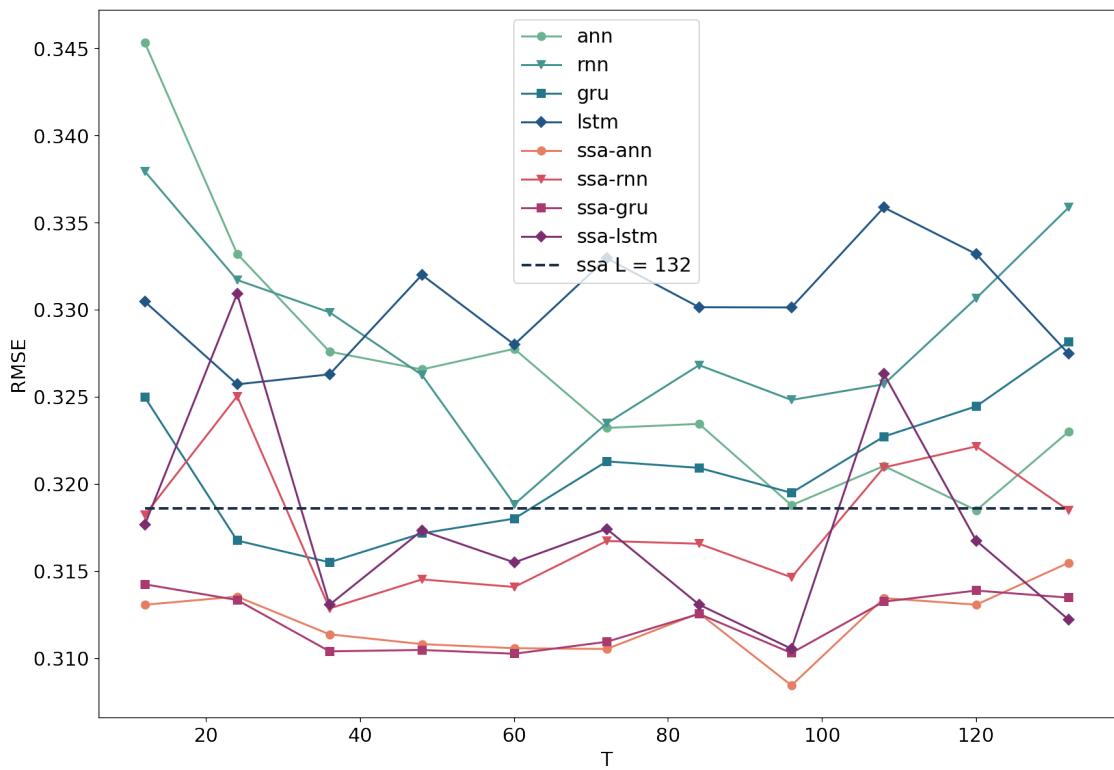


Рис. 4.37. Сумма синусов с небольшим белым шумом. Ошибки прогноза в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 4$ .

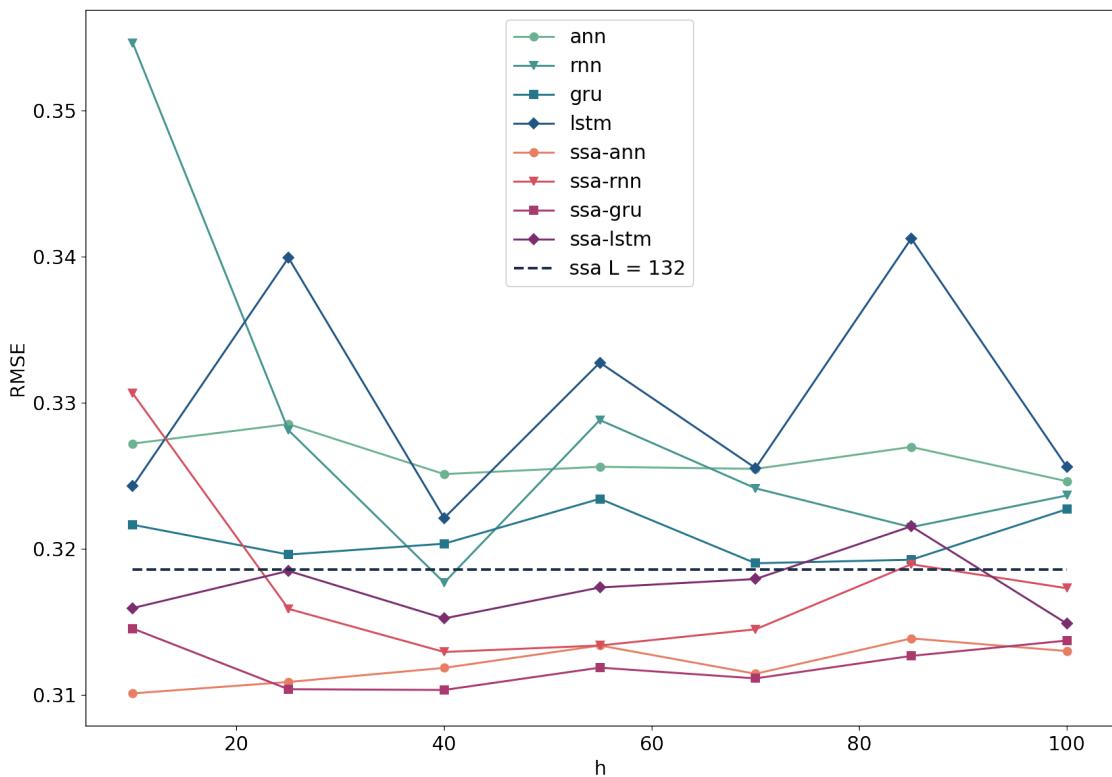


Рис. 4.38. Сумма синусов с небольшим белым шумом. Ошибки прогноза в зависимости от параметра  $h$ .  $L = 175$ ,  $r = 4$ .

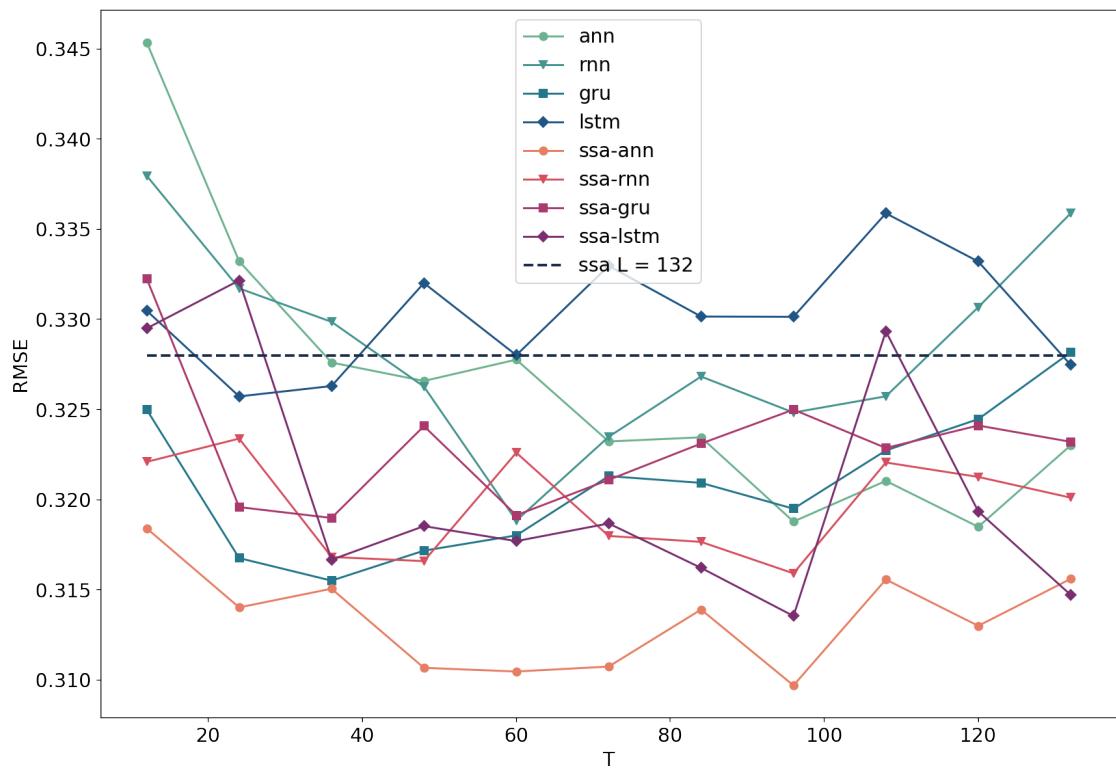


Рис. 4.39. Сумма синусов с небольшим белым шумом. Ошибки прогноза в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 6$ .

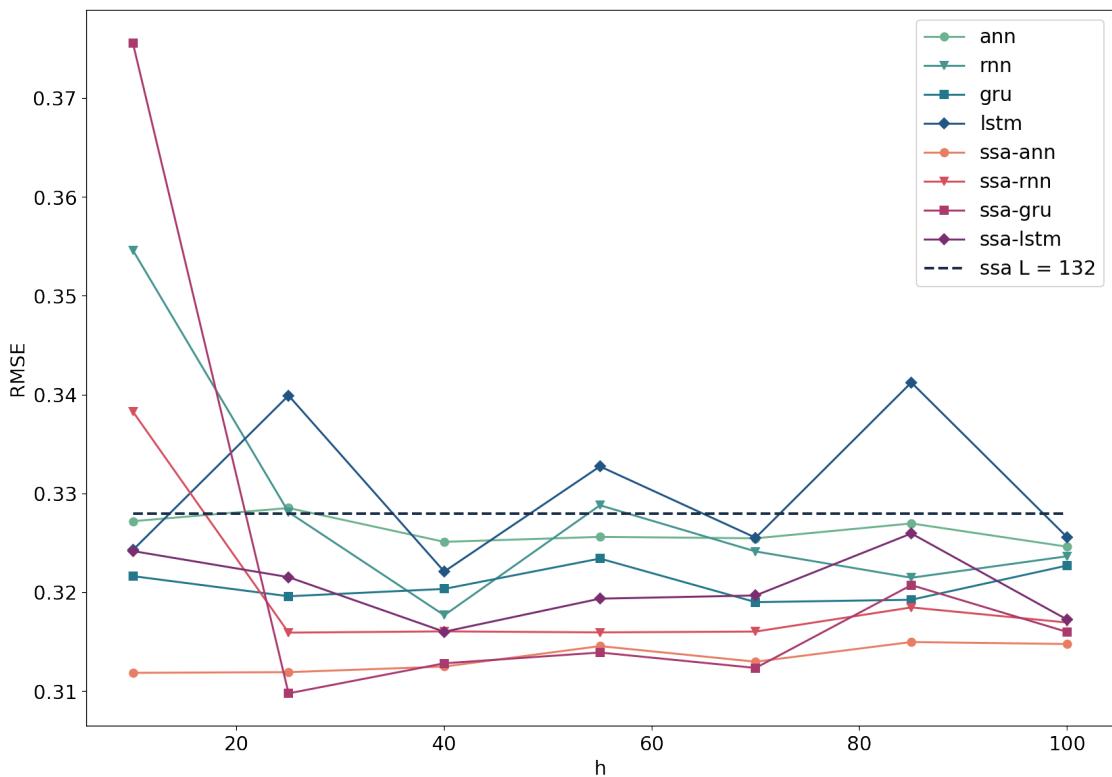


Рис. 4.40. Сумма синусов с небольшим белым шумом. Ошибки прогноза в зависимости от параметра  $h$ .  $L = 175$ ,  $r = 6$ .

### Отображение прогнозов

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

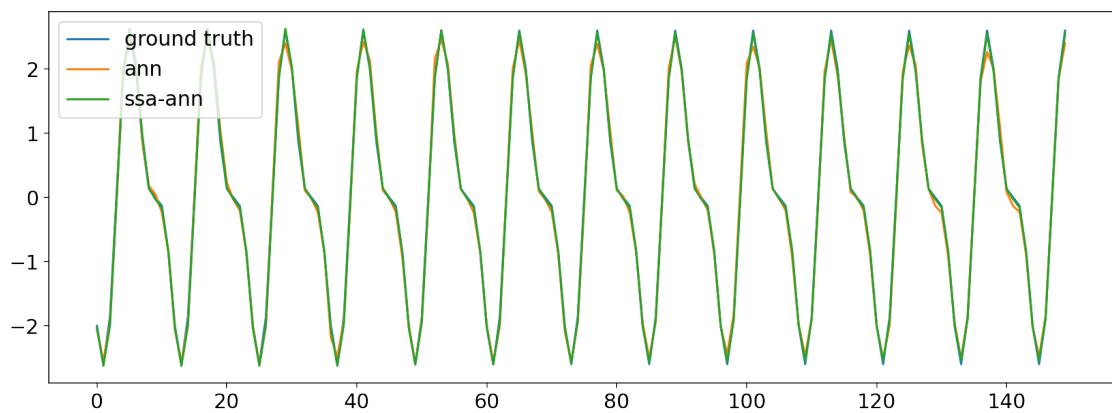


Рис. 4.41. Сумма синусов с небольшим белым шумом. Прогноз результатов для ANN и SSA-ANN.  $r = 2$

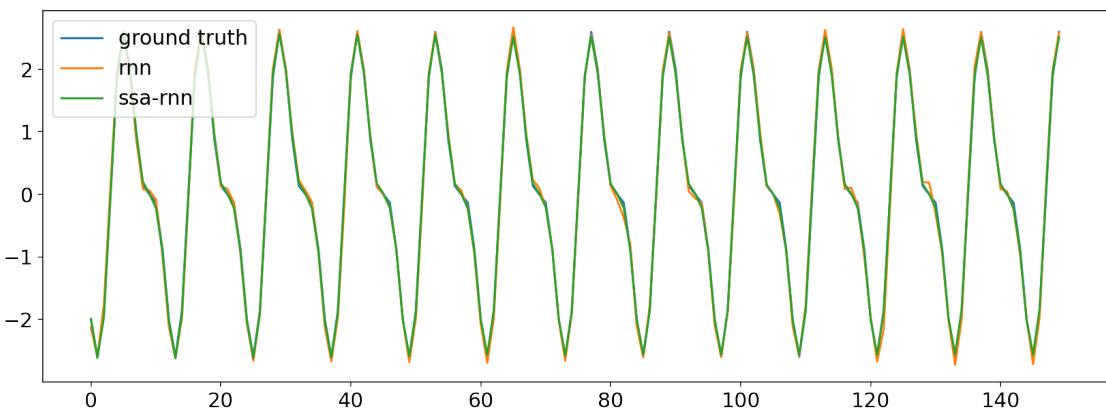


Рис. 4.42. Сумма синусов с небольшим белым шумом. Прогноз результатов для RNN и SSA-RNN.  $r = 2$

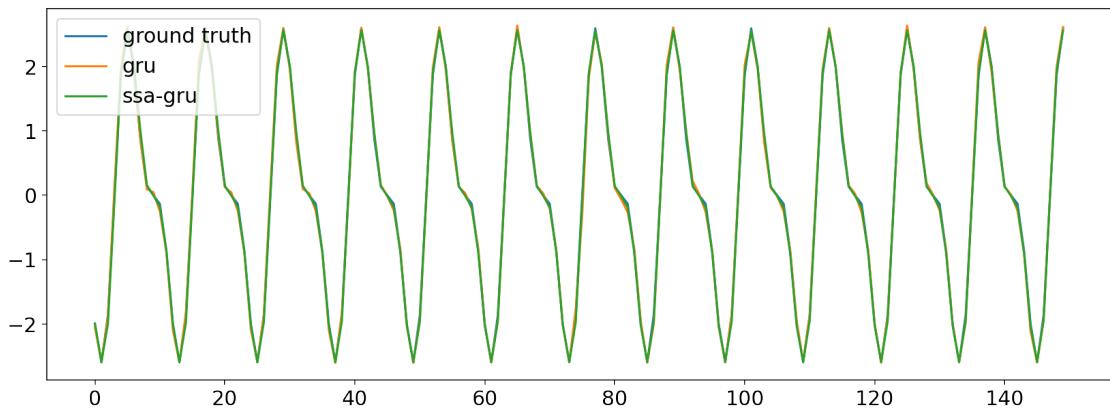


Рис. 4.43. Сумма синусов с небольшим белым шумом. Прогноз результатов для GRU и SSA-GRU.  $r = 2$

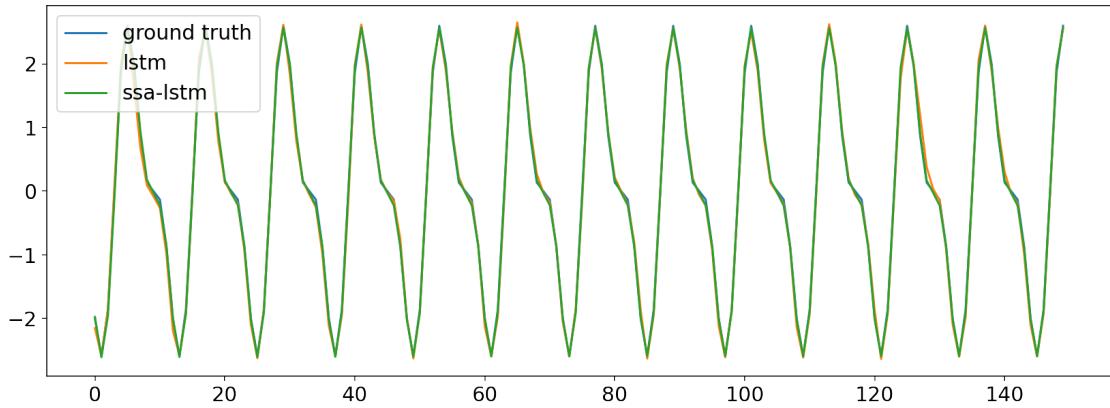


Рис. 4.44. Сумма синусов с небольшим белым шумом. Прогноз результатов для LSTM и SSA-LSTM.  $r = 2$

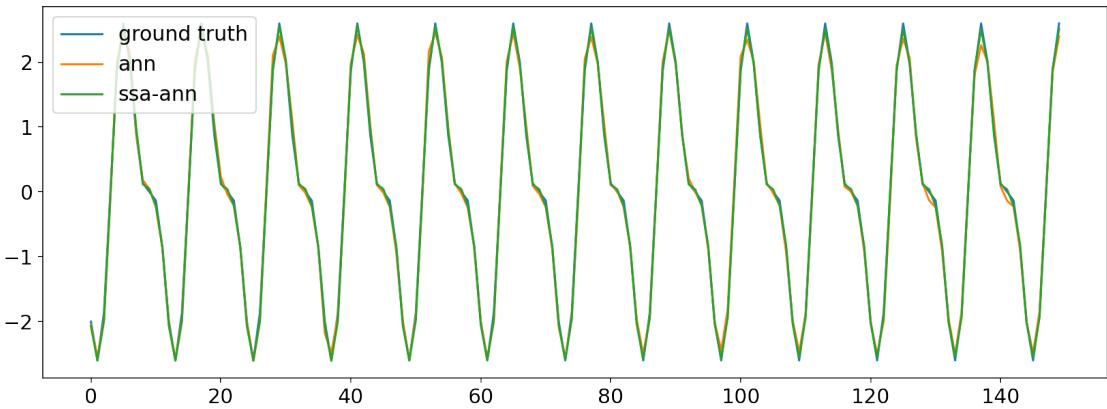


Рис. 4.45. Сумма синусов с небольшим белым шумом. Прогноз результатов для ANN и SSA-ANN.  $r = 4$

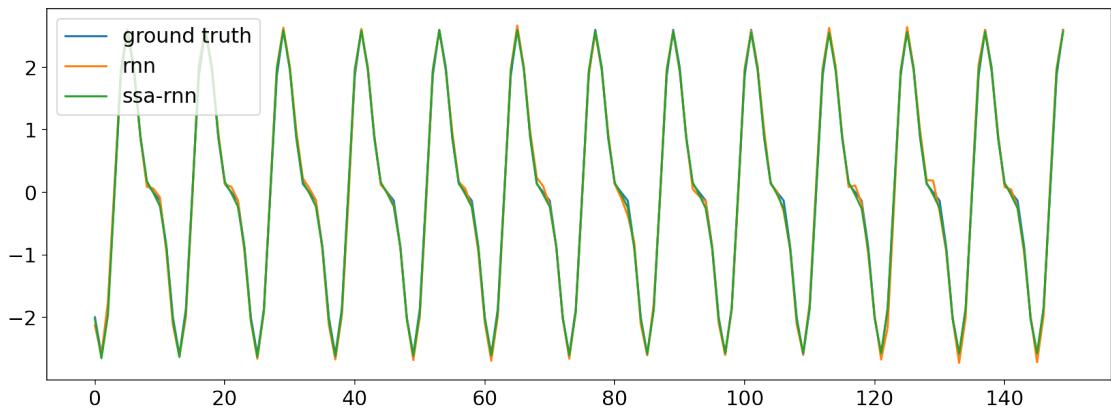


Рис. 4.46. Сумма синусов с небольшим белым шумом. Прогноз результатов для RNN и SSA-RNN.  $r = 4$

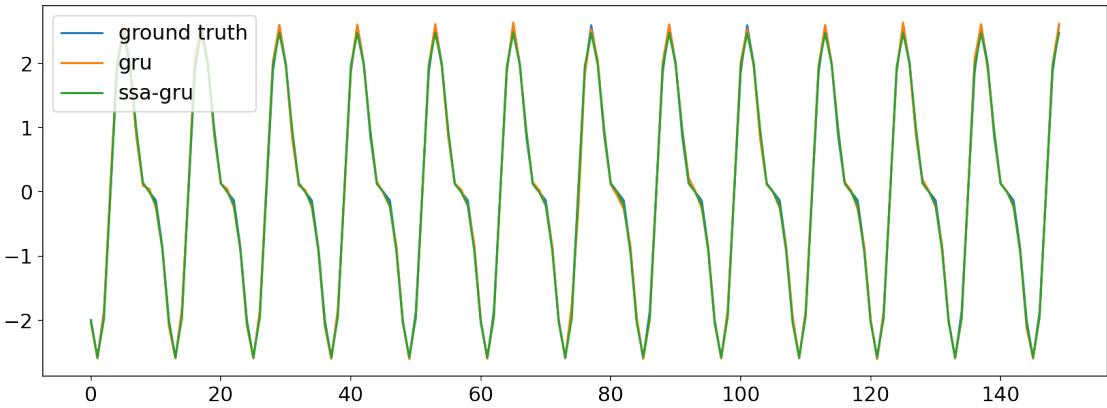


Рис. 4.47. Сумма синусов с небольшим белым шумом. Прогноз результатов для GRU и SSA-GRU.  $r = 4$

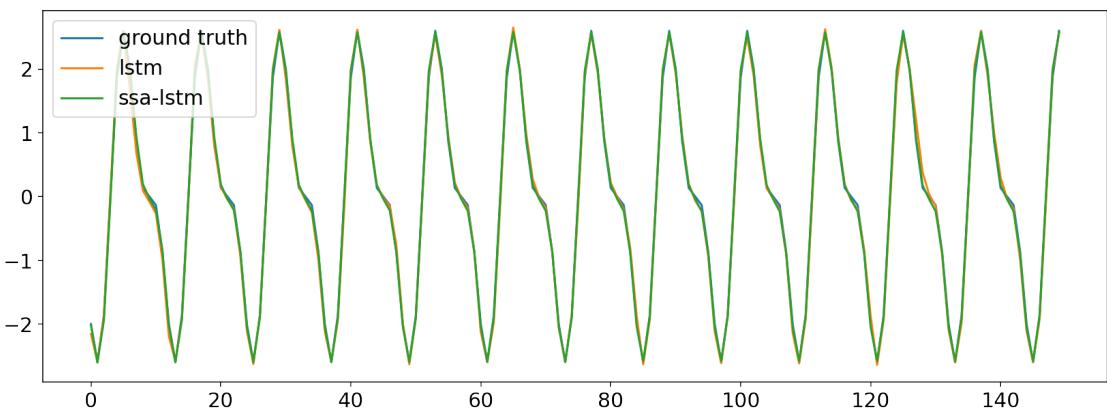


Рис. 4.48. Сумма синусов с небольшим белым шумом. Прогноз результатов для LSTM и SSA-LSTM.  $r = 4$

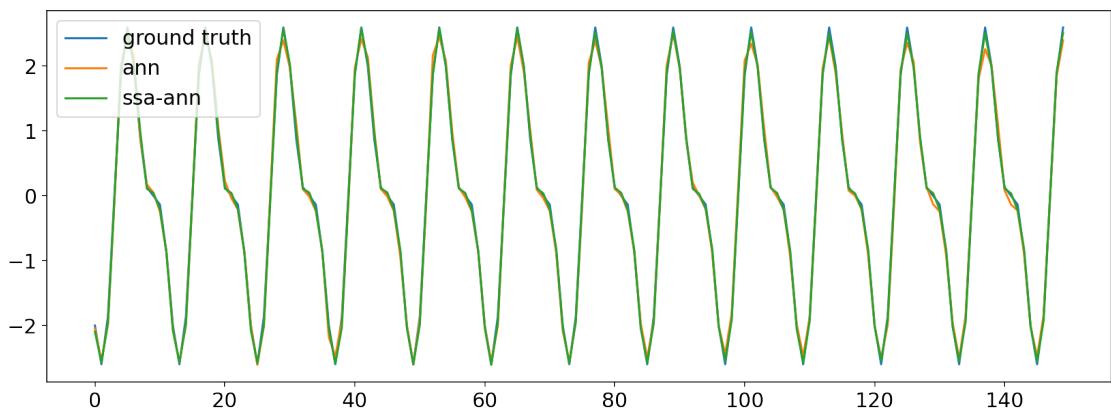


Рис. 4.49. Сумма синусов с небольшим белым шумом. Прогноз результатов для ANN и SSA-ANN.  $r = 6$

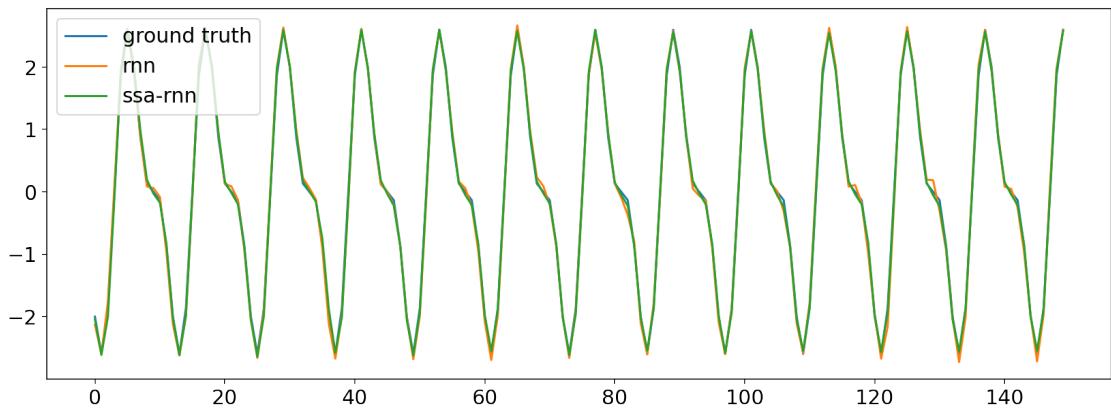


Рис. 4.50. Сумма синусов с небольшим белым шумом. Прогноз результатов для RNN и SSA-RNN.  $r = 6$

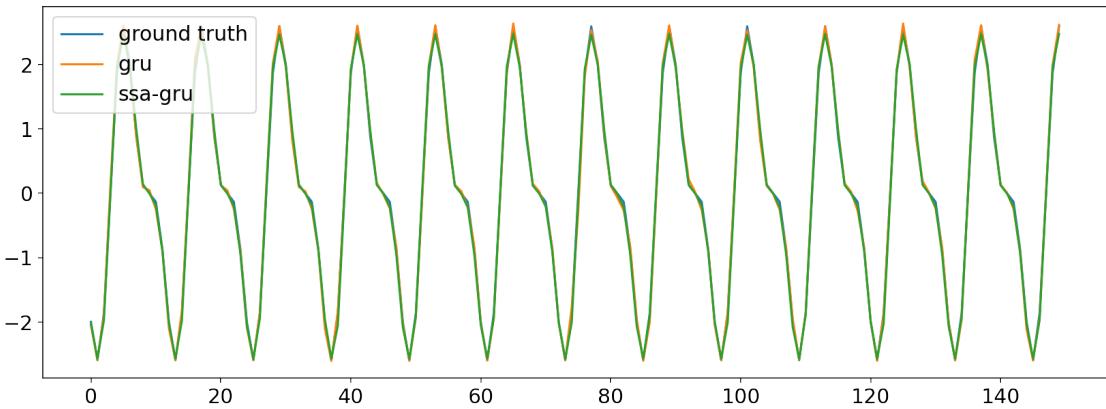


Рис. 4.51. Сумма синусов с небольшим белым шумом. Прогноз результатов для GRU и SSA-GRU.  $r = 6$

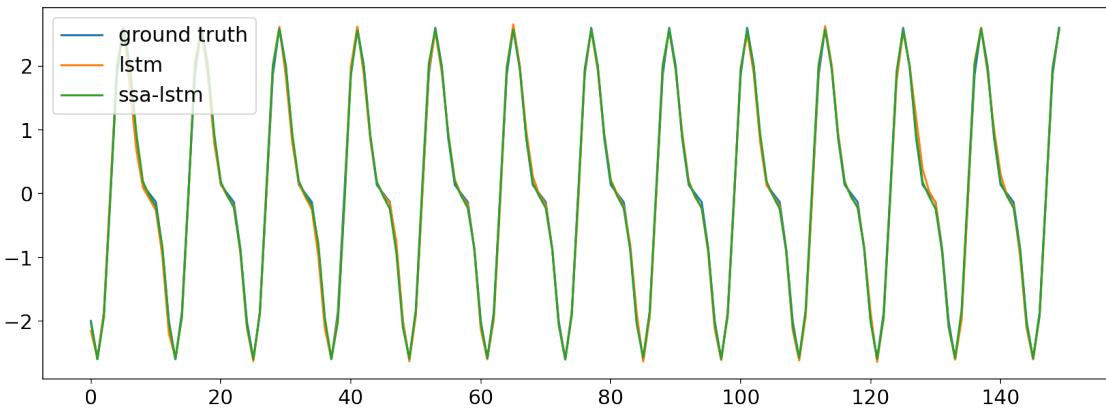


Рис. 4.52. Сумма синусов с небольшим белым шумом. Прогноз результатов для LSTM и SSA-LSTM.  $r = 6$

### Проверка устойчивости

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

### Выводы

На ряде  $X_{650}$  было продемонстрировано сравнение обычных и гибридных методов, и метода SSA. Из полученных результатов можем сделать выводы, что для ряда с

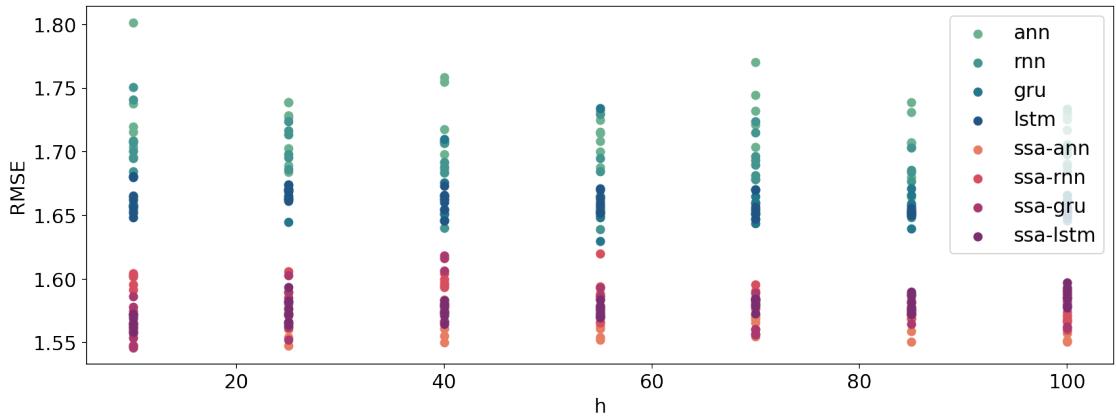


Рис. 4.53. Сумма синусов с небольшим белым шумом. Проверка устойчивости.  $r = 2$ ,  $L = 175$ .  
 $T = 12$ .

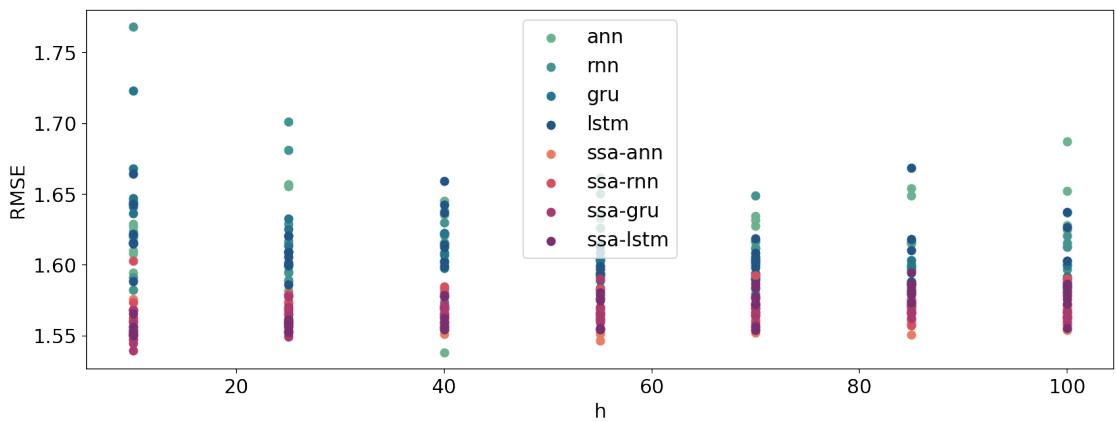


Рис. 4.54. Сумма синусов с небольшим белым шумом. Проверка устойчивости.  $r = 2$ ,  $L = 175$ .  
 $T = 84$ .

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

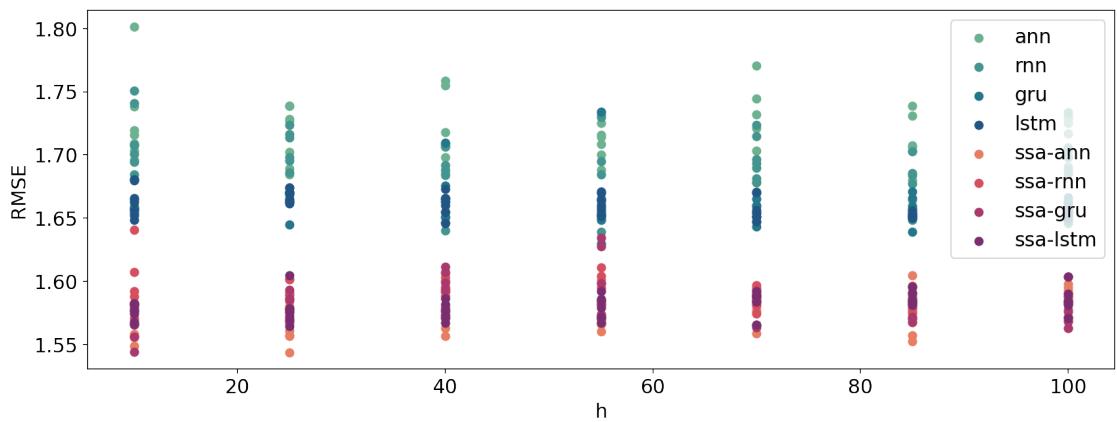


Рис. 4.55. Сумма синусов с небольшим белым шумом. Проверка устойчивости.  $r = 4$ ,  $L = 175$ .  
 $T = 12$ .

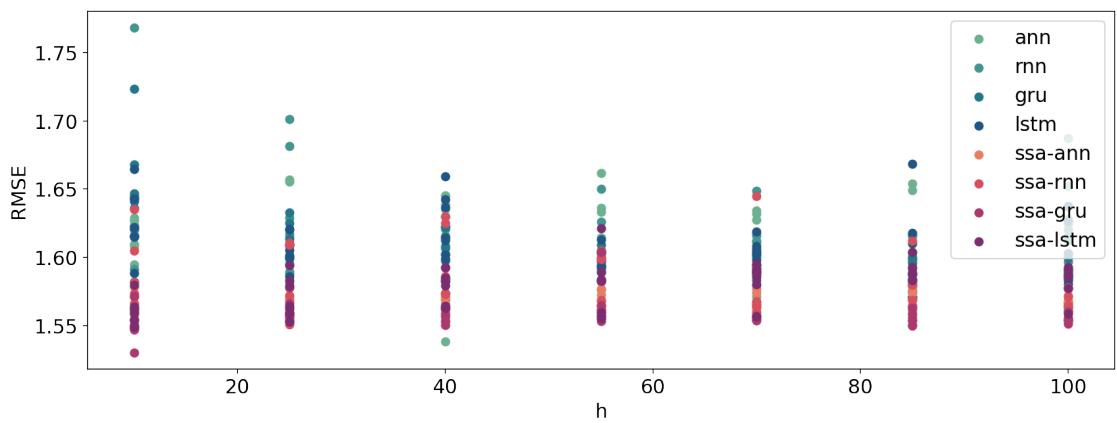


Рис. 4.56. Сумма синусов с небольшим белым шумом. Проверка устойчивости.  $r = 4$ ,  $L = 175$ .  
 $T = 84$ .

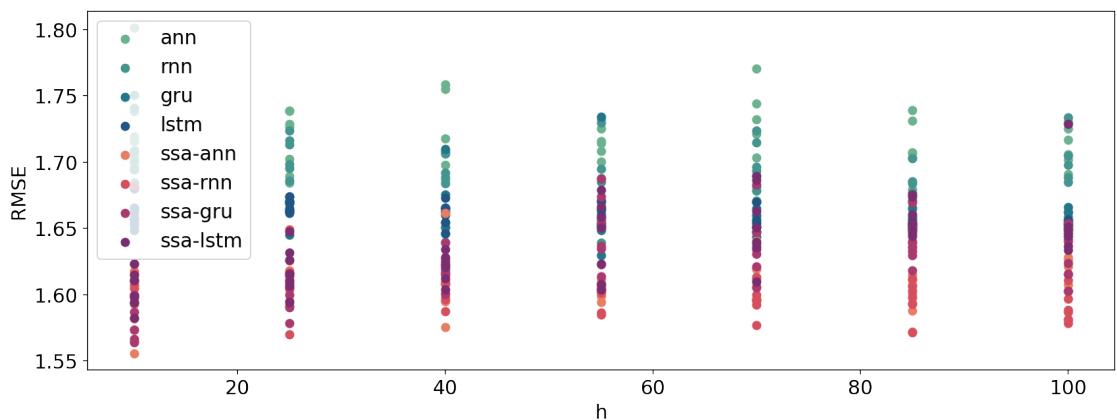


Рис. 4.57. Сумма синусов с небольшим белым шумом. Проверка устойчивости.  $r = 6$ ,  $L = 175$ .  
 $T = 12$ .

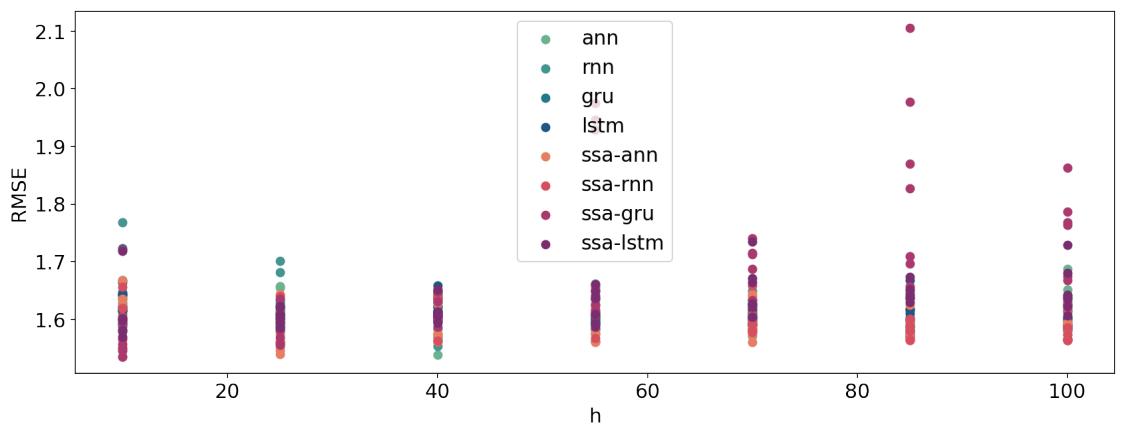


Рис. 4.58. Сумма синусов с небольшим белым шумом. Проверка устойчивости.  $r = 6$ ,  $L = 175$ .  
 $T = 84$ .

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

Рассмотрим следующий ряд  $Z_{650}$  (рис. 4.59), 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  $\xi_i = \xi_{i-1} + \sigma \varepsilon_{i-1}$ ,  $\sigma = 1.2$ ,  $\varepsilon_i \in N(0, 1)$ .

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. 4.60). Из формулы видно, что ряд 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  $L = 175$ ,  $r = 4$ , но так как из-за красного шума сигнал сложно выделить, то не возможно данная пара будет не оптимальной. Далее в главе сравним разные параметры для метода SSA, выберем те которые дают хорошую точность, а также не сильную аппроксимацию ряд оценкой сигнала.

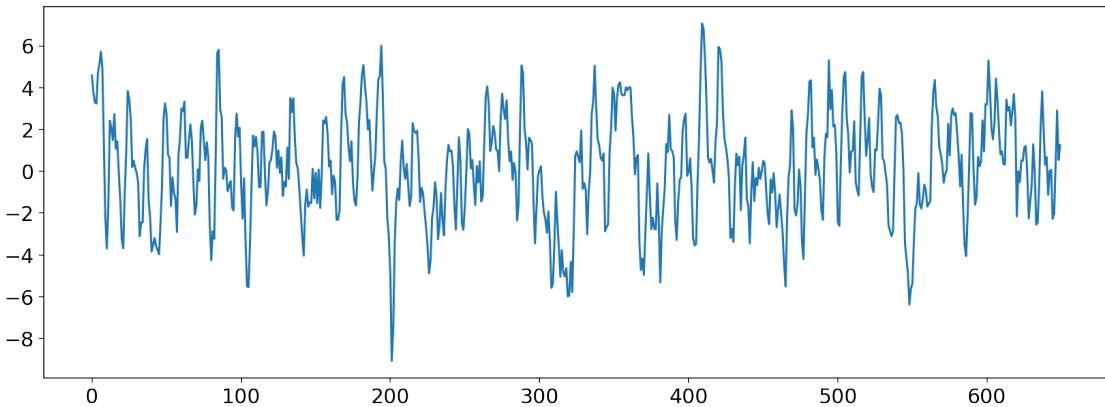


Рис. 4.59. Ряд суммы синусов с красным шумом  $Z_{650}$ .

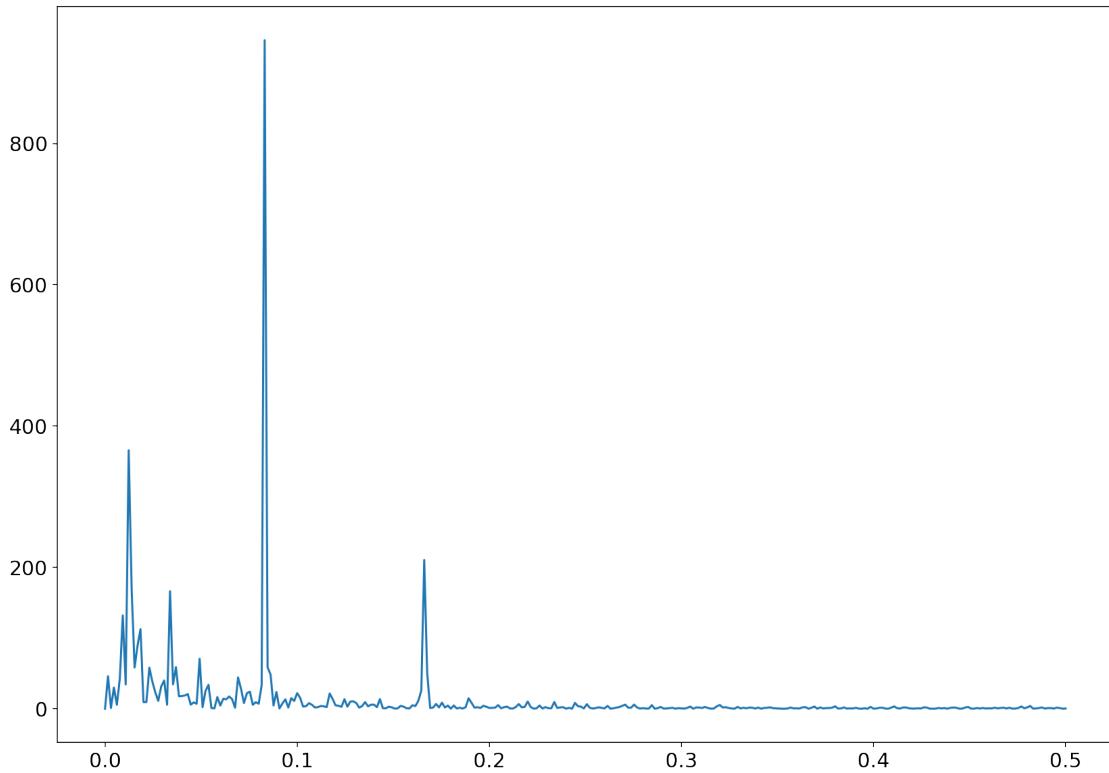


Рис. 4.60. Периодограмма ряда Z<sub>650</sub>.

#### 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, описанным 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 в методе SSA.

#### 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 4.61. На графике видно, что чем меньше  $r$ , тем выше ошибка. Для  $r \geq 8$  разница между ошибками не такая большая, и результаты начинают смещиваться. Из графика видно, что  $r = 2$  дает результаты сильно хуже, что говорит о сильном влиянии шума. Также заметим, что ошибка для аналитически верной пары растет с ростом  $L$ .

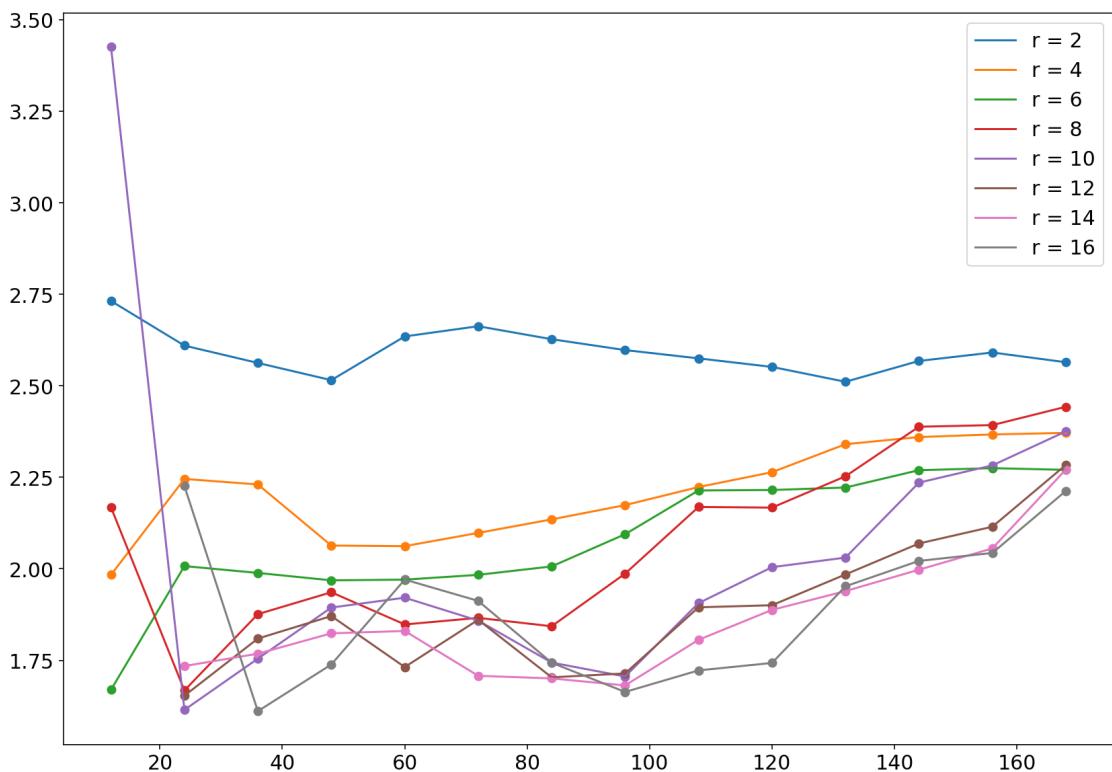


Рис. 4.61. Сумма синусов с красным шумом. Прогнозирование с помощью метода SSA.

### Восстановление SSA

Посмотрим как SSA восстанавливает тренировочную выборку. Посмотрим на график зависимости ошибки восстановления тренировочной выборки от параметра  $L$ . Сетку для параметров оставим такую же. На графике 4.62 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 4.61. Для выбранных пар покажем как метод SSA восстанавливает сигнал.

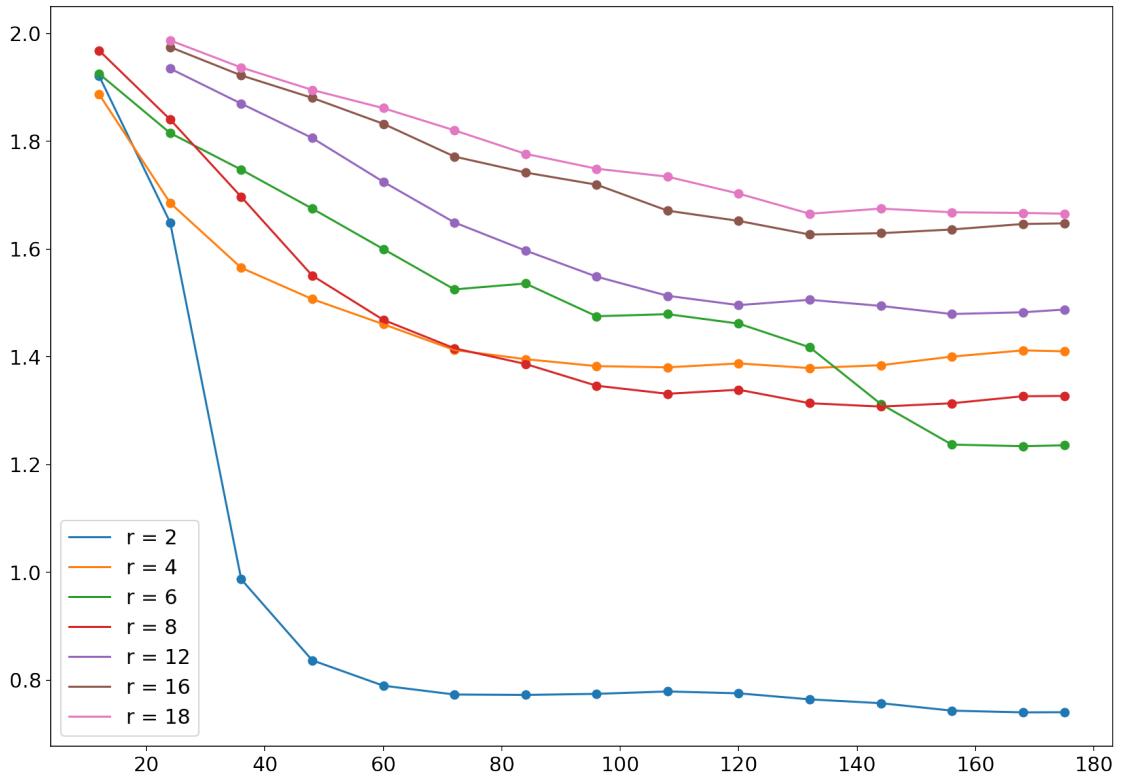


Рис. 4.62. Сумма синусов с красным шумом. Зависимость ошибки восстановления от параметра  $T$ .

Ниже на графиках представлены результаты. Видно, что лучше всего метод SSA восстанавливает сигнал с параметрами  $L = 175, r = 2$  (рис. 4.63). Для остальных пар наблюдается влияние шума. На рисунках 4.64, 4.65, 4.66, 4.67 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 4.68 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 4.69.

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.

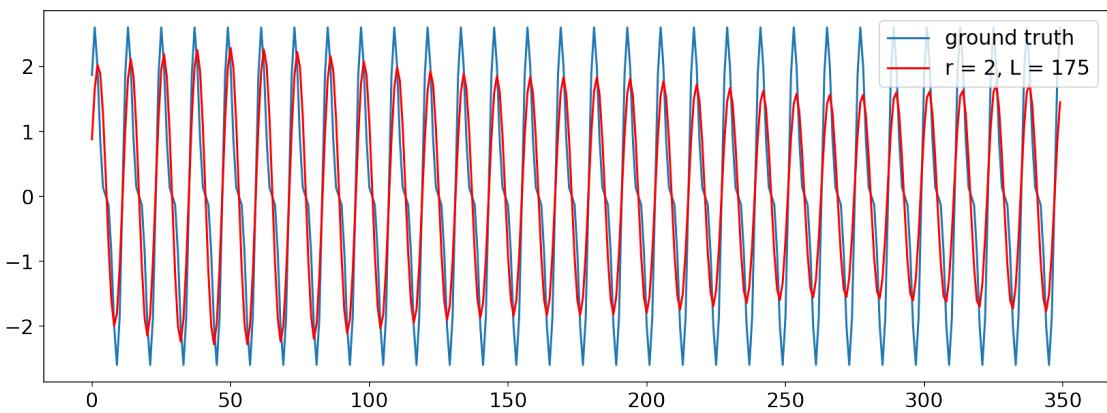


Рис. 4.63. Сумма синусов с красным шумом. Восстановление тренировочной выборки с помощью метода SSA.  $L = 175$ ,  $r = 2$ .

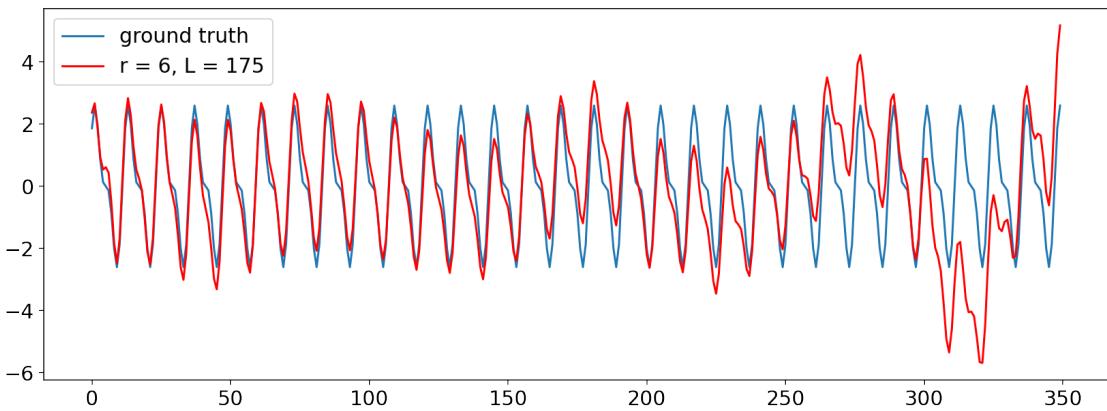


Рис. 4.64. Сумма синусов с красным шумом. Восстановление тренировочной выборки с помощью метода SSA.  $L = 175$ ,  $r = 6$ .

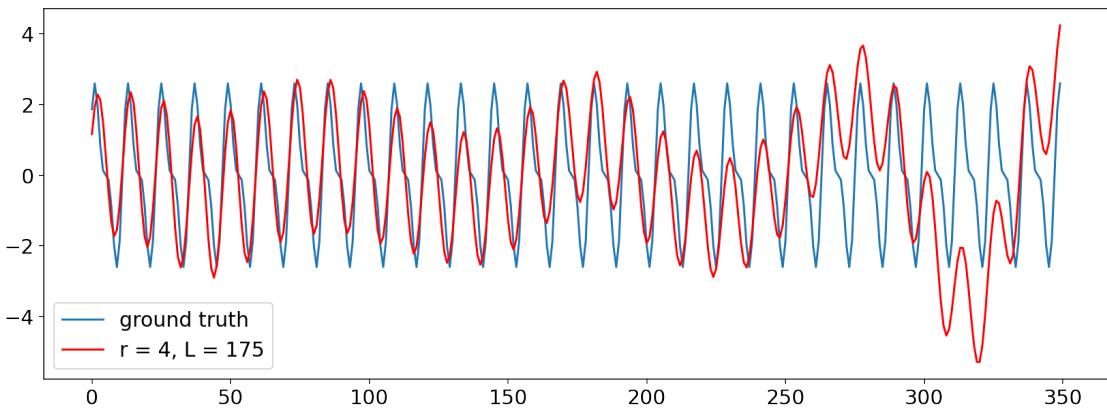


Рис. 4.65. Сумма синусов с красным шумом. Восстановление тренировочной выборки с помощью метода SSA.  $L = 175$ ,  $r = 4$ .

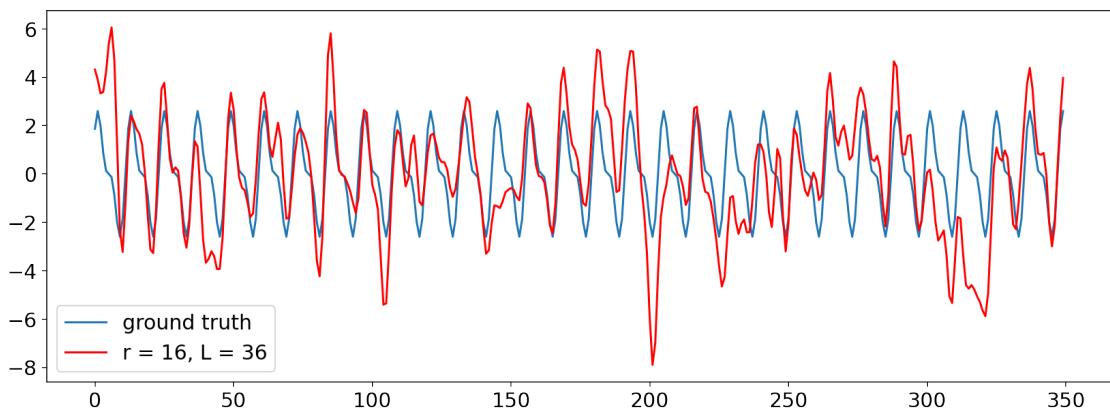


Рис. 4.66. Сумма синусов с красным шумом. Восстановление тренировочной выборки с помощью метода SSA.  $L = 36$ ,  $r = 16$ .

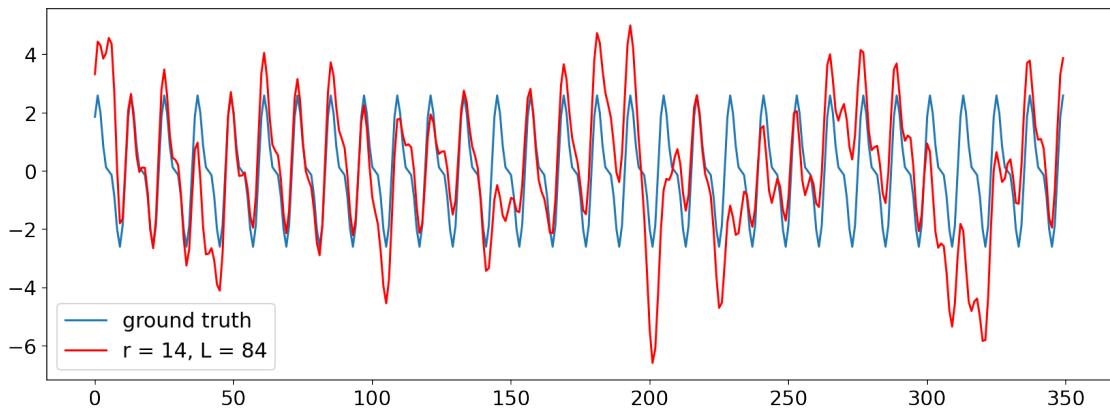


Рис. 4.67. Сумма синусов с красным шумом. Восстановление тренировочной выборки с помощью метода SSA.  $L = 84$ ,  $r = 14$ .

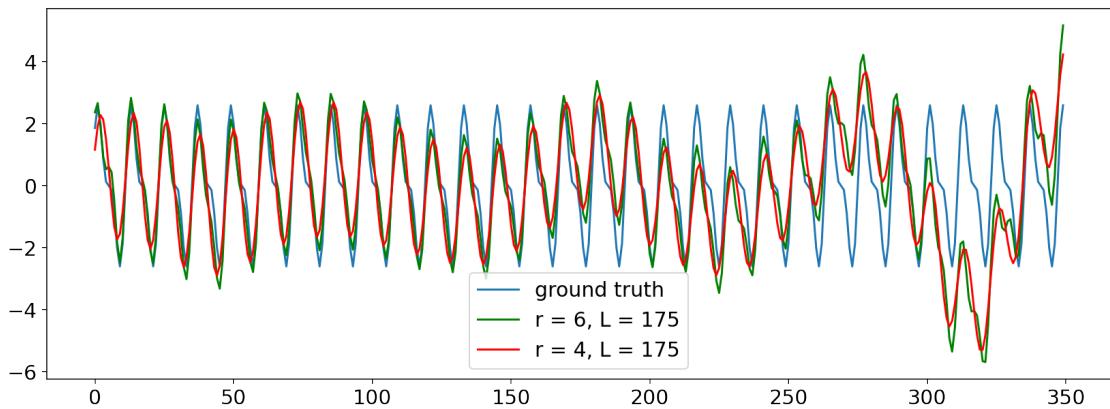


Рис. 4.68. Сумма синусов с красным шумом. Восстановление тренировочной выборки с помощью метода SSA.  $L = 175$ ,  $r = 4$  и  $L = 175$ ,  $r = 6$ .

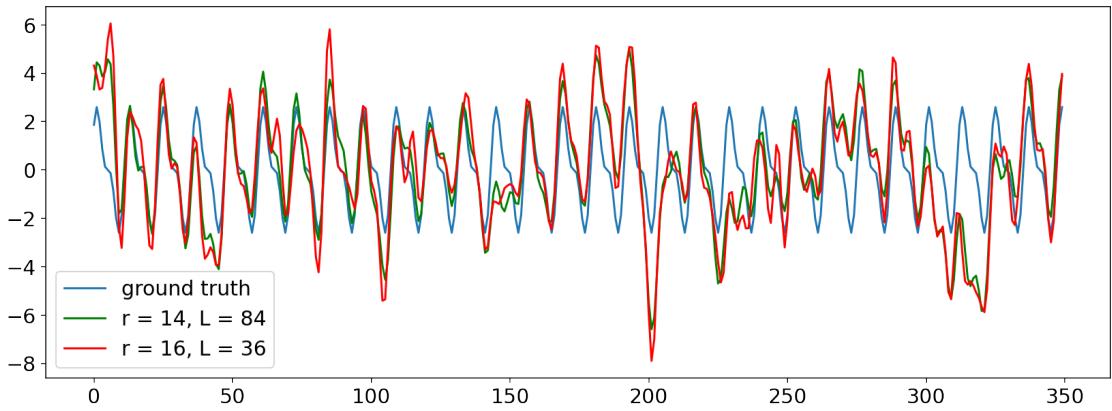


Рис. 4.69. Сумма синусов с красным шумом. Восстановление тренировочной выборки с помощью метода SSA.  $L = 84$ ,  $r = 14$  и  $L = 36$ ,  $r = 16$ .

### Сравнение обычных и гибридных методов

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

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

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

На графиках 4.70, 4.71, 4.72, 4.73, 4.74, 4.75 представлены результаты сравнения по сетке параметров, заданной выше. На графиках видно, что метод SSA и гибридные методы сильно проигрывают обычных нейронным сетям. Это объяснимо тем, что ошибка считается относительно всего ряда. Возможно, гибридные методы предсказывают сигнал ряда, а он в свою очередь сильно отличен от временного ряда с шумом. Чтобы проверить это посмотрим на примеры предсказаний в следующем разделе.

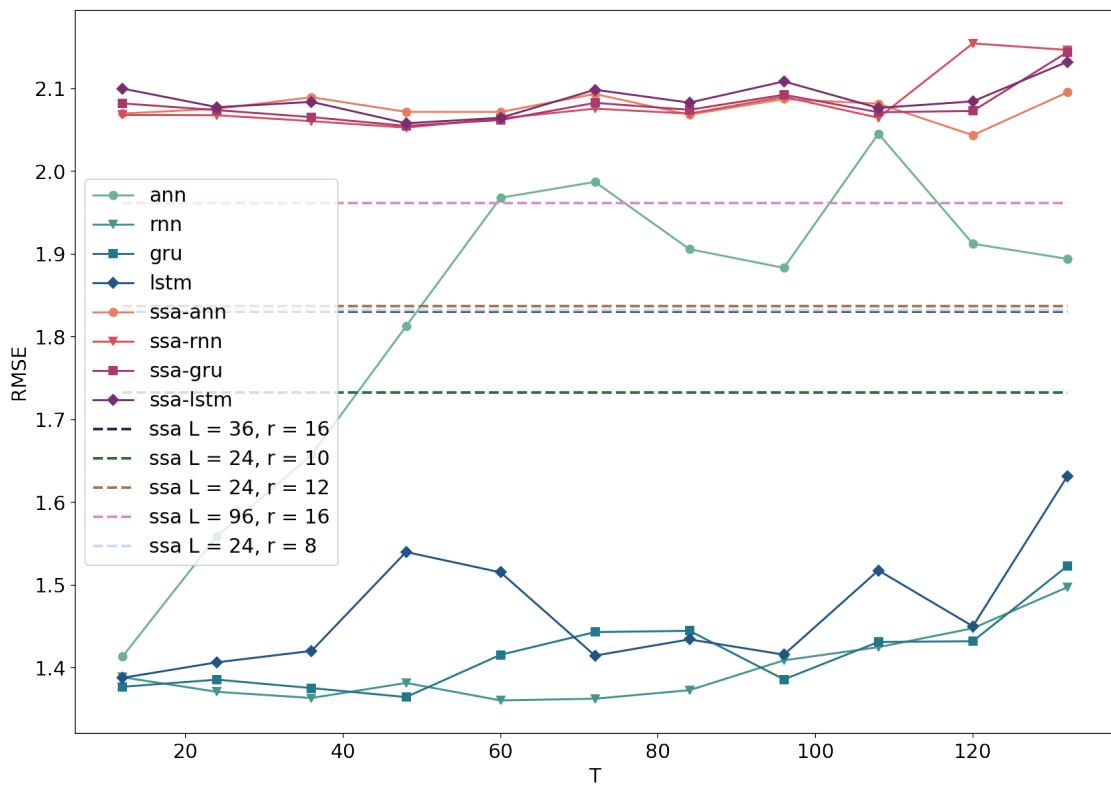


Рис. 4.70. Сумма синусов с красным шумом. Ошибки прогноза в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 2$ .

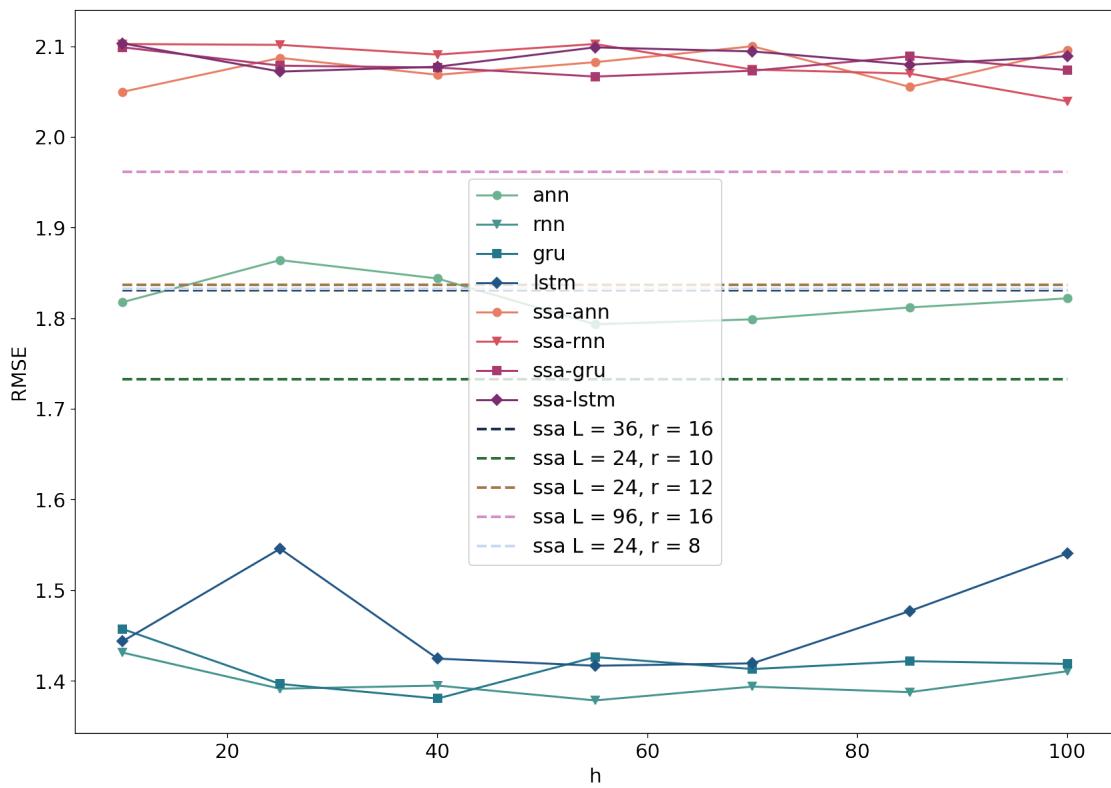


Рис. 4.71. Сумма синусов с красным шумом. Ошибки прогноза в зависимости от параметра  $h$ .  $L = 175$ ,  $r = 2$ .

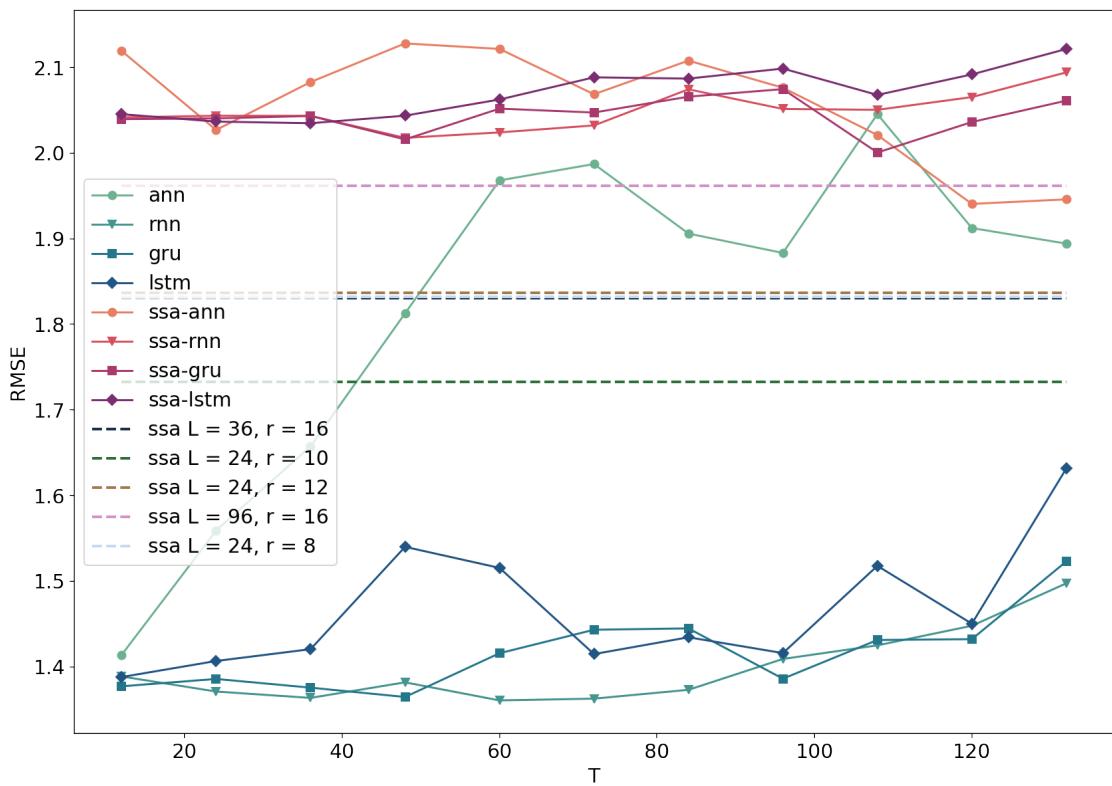


Рис. 4.72. Сумма синусов с красным шумом. Ошибки прогноза в зависимости от параметра  $T$ .  $L = 175$ ,  $r = 4$ .

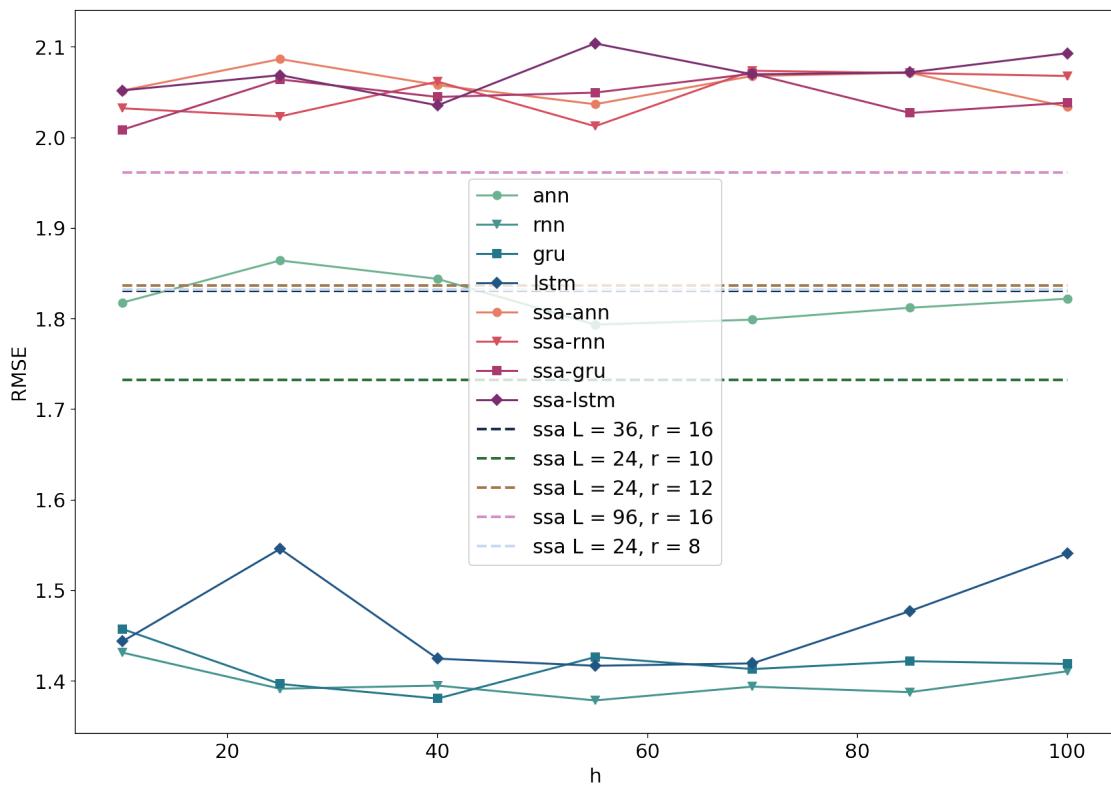


Рис. 4.73. Сумма синусов с красным шумом. Ошибки прогноза в зависимости от параметра  $h$ .  $L = 175$ ,  $r = 4$ .

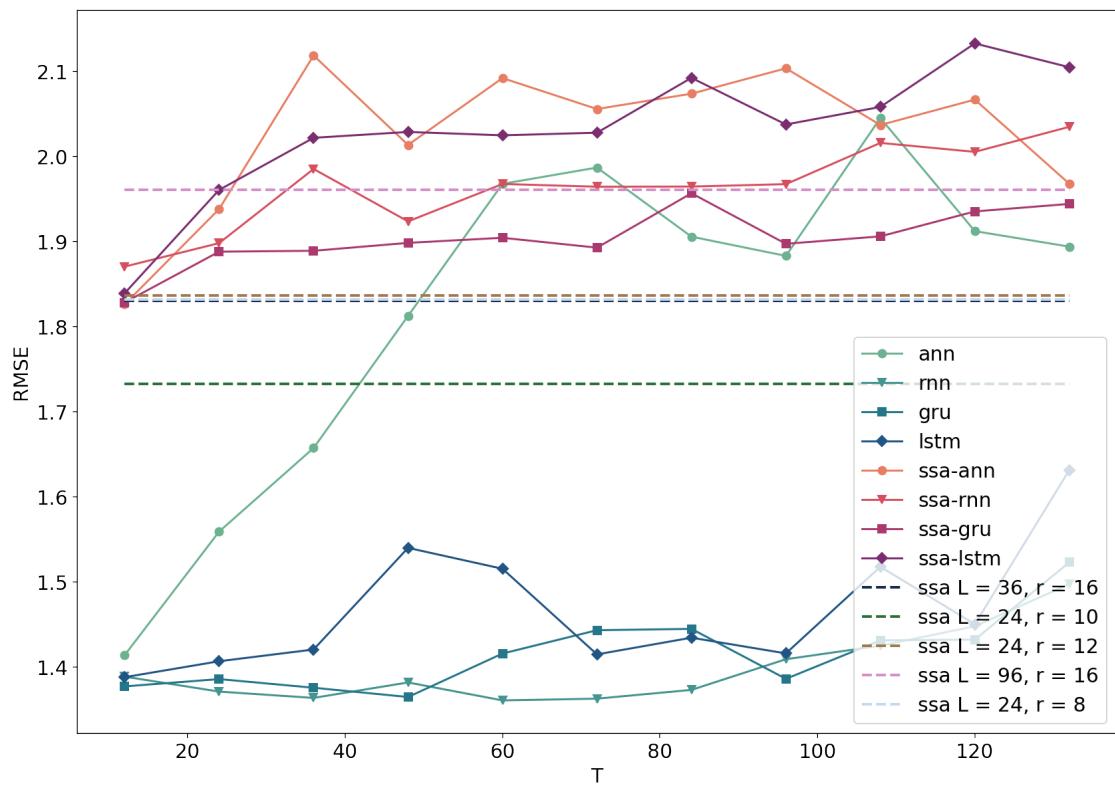


Рис. 4.74. Сумма синусов с красным шумом. Ошибки прогноза в зависимости от параметра  $T$ .  $L = 84$ ,  $r = 14$ .

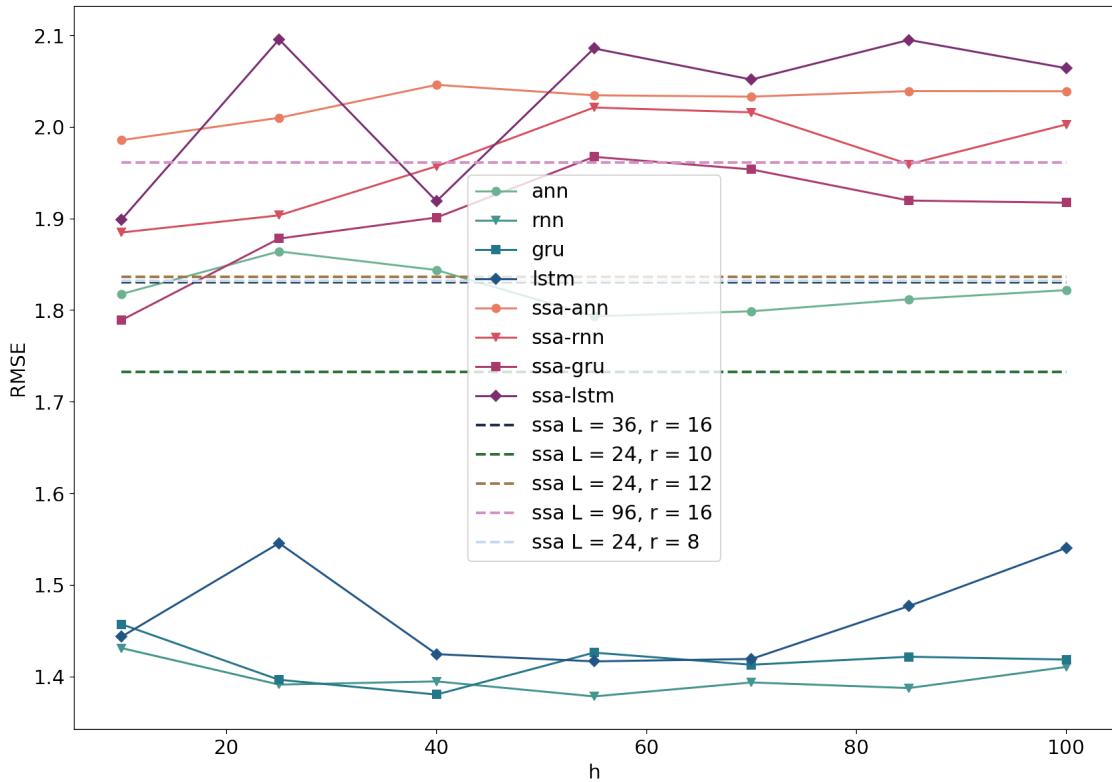


Рис. 4.75. Сумма синусов с красным шумом. Ошибки прогноза в зависимости от параметра  $h$ .  $L = 84$ ,  $r = 14$ .

### Отображение прогнозов

На графиках ниже представлены результаты прогнозирования методами (на графиках показано отклонение от сигнала ряда). Можем заметить как сильно обычные методы пытаются предсказывать шум. Для пары  $L = 84$ ,  $r = 14$  предсказание сигнала получилось не очень хорошее. 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.

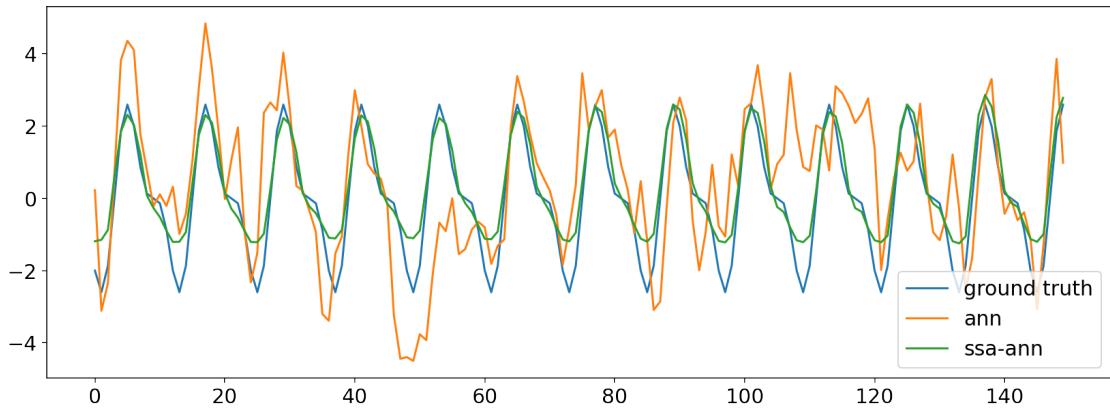


Рис. 4.76. Сумма синусов с красным шумом. Прогноз результатов для ANN и SSA-ANN.

$$L = 175, r = 2$$

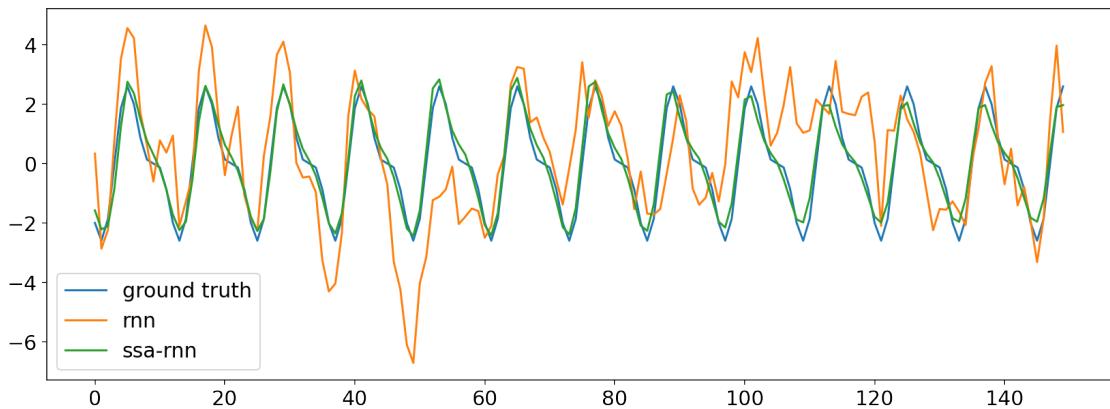


Рис. 4.77. Сумма синусов с красным шумом. Прогноз результатов для RNN и SSA-RNN.

$$L = 175, r = 2$$

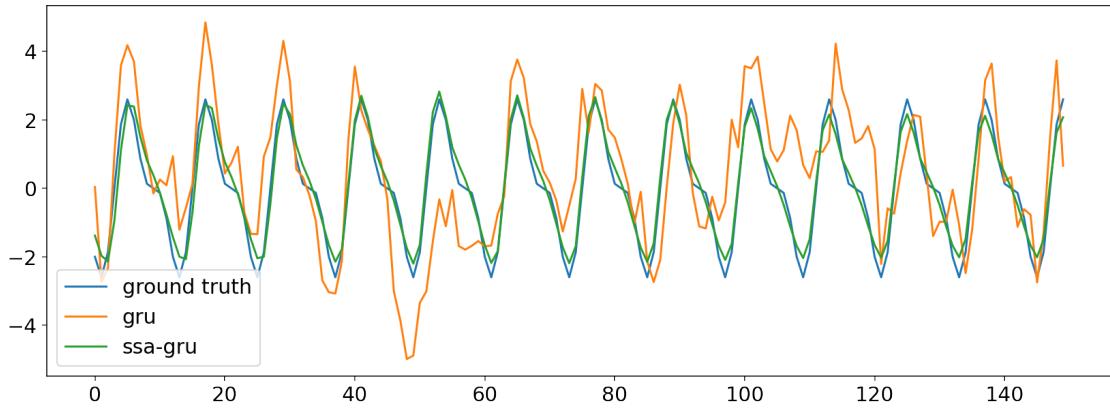


Рис. 4.78. Сумма синусов с красным шумом. Прогноз результатов для GRU и SSA-GRU.

$$L = 175, r = 2$$

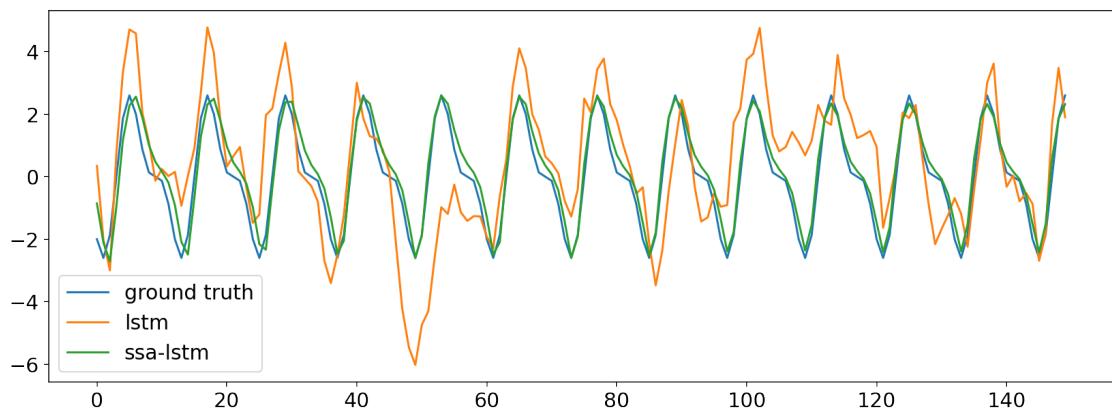


Рис. 4.79. Сумма синусов с красным шумом. Прогноз результатов для LSTM и SSA-LSTM.

$$L = 175, r = 2$$

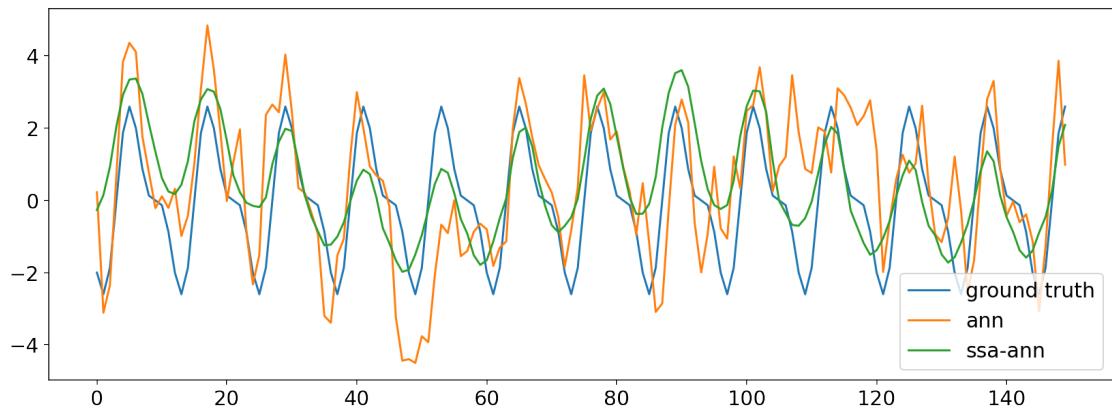


Рис. 4.80. Сумма синусов с красным шумом. Прогноз результатов для ANN и SSA-ANN.

$$L = 175, r = 4$$

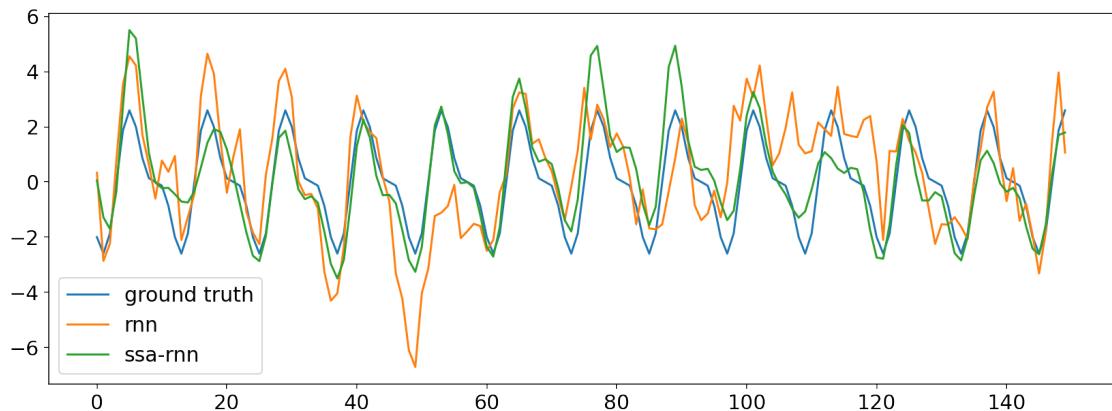


Рис. 4.81. Сумма синусов с красным шумом. Прогноз результатов для RNN и SSA-RNN.

$$L = 175, r = 4$$

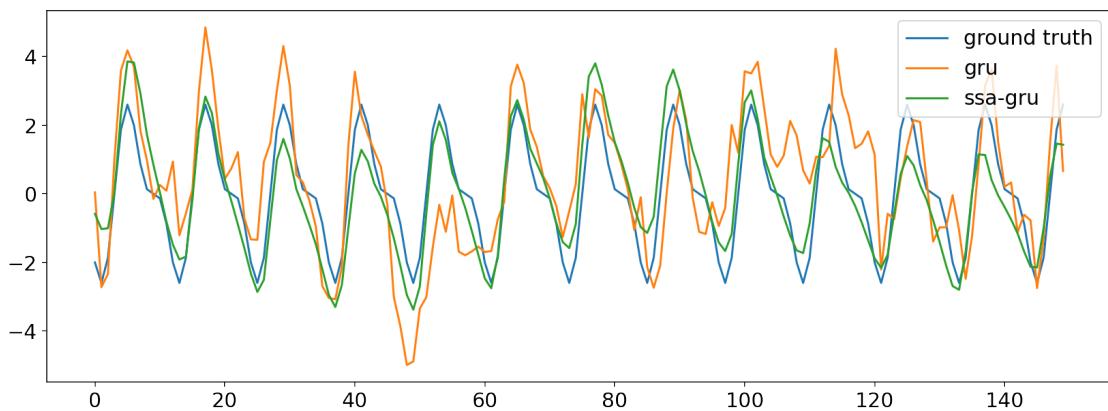


Рис. 4.82. Сумма синусов с красным шумом. Прогноз результатов для GRU и SSA-GRU.

$$L = 175, r = 4$$

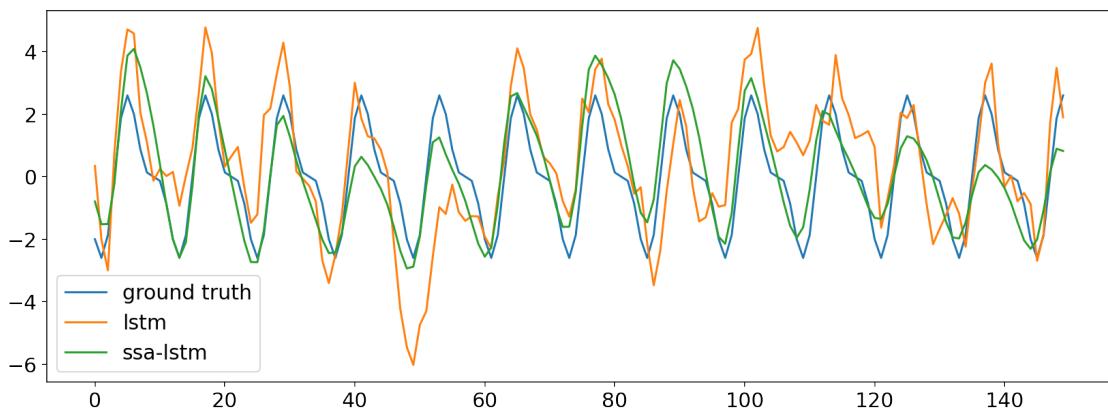


Рис. 4.83. Сумма синусов с красным шумом. Прогноз результатов для LSTM и SSA-LSTM.

$$L = 175, r = 4$$

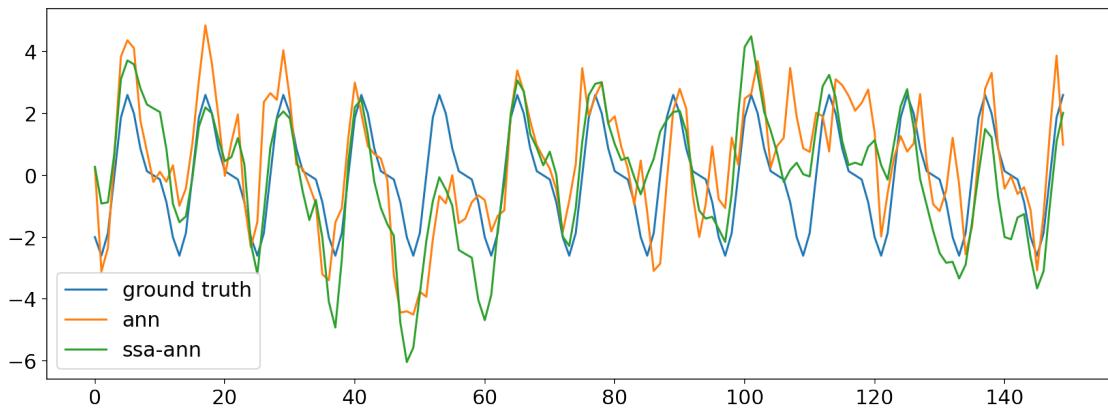


Рис. 4.84. Сумма синусов с красным шумом. Прогноз результатов для ANN и SSA-ANN.

$$L = 84, r = 14$$

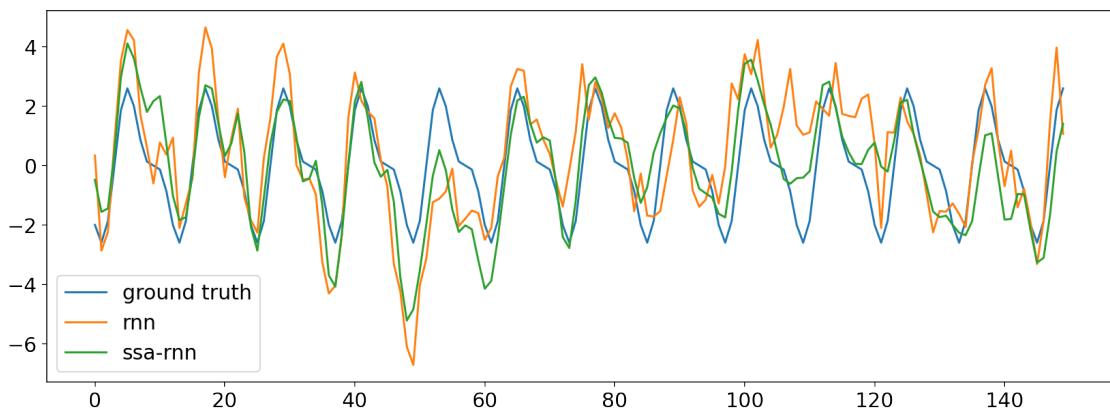


Рис. 4.85. Сумма синусов с красным шумом. Прогноз результатов для RNN и SSA-RNN.

$$L = 84, r = 14$$

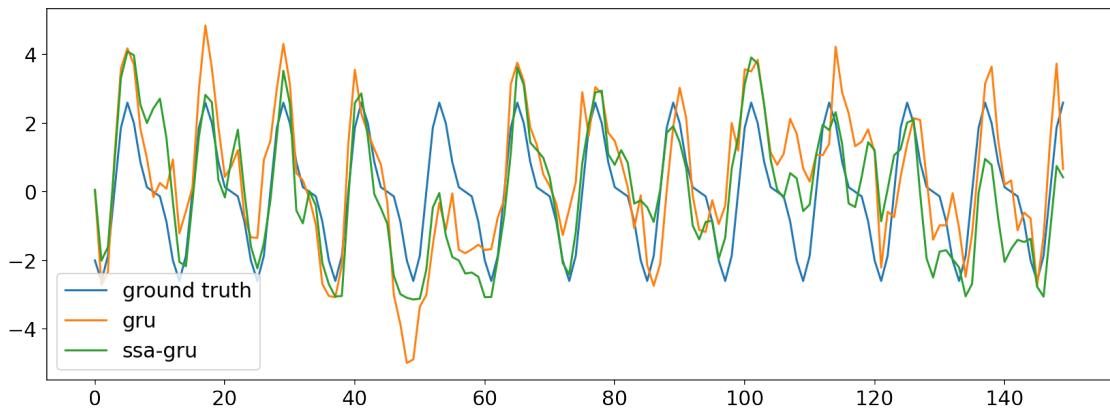


Рис. 4.86. Сумма синусов с красным шумом. Прогноз результатов для GRU и SSA-GRU.

$$L = 84, r = 14$$

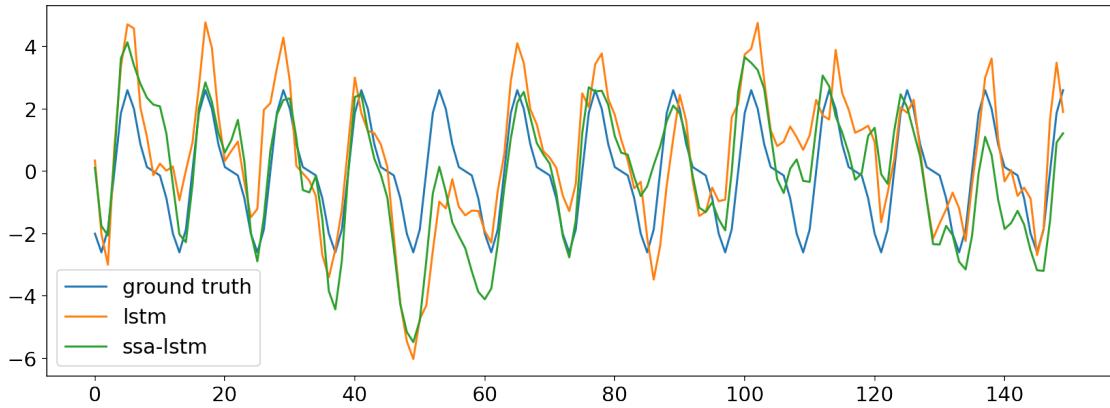


Рис. 4.87. Сумма синусов с красным шумом. Прогноз результатов для LSTM и SSA-LSTM.

$$L = 84, r = 14$$

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

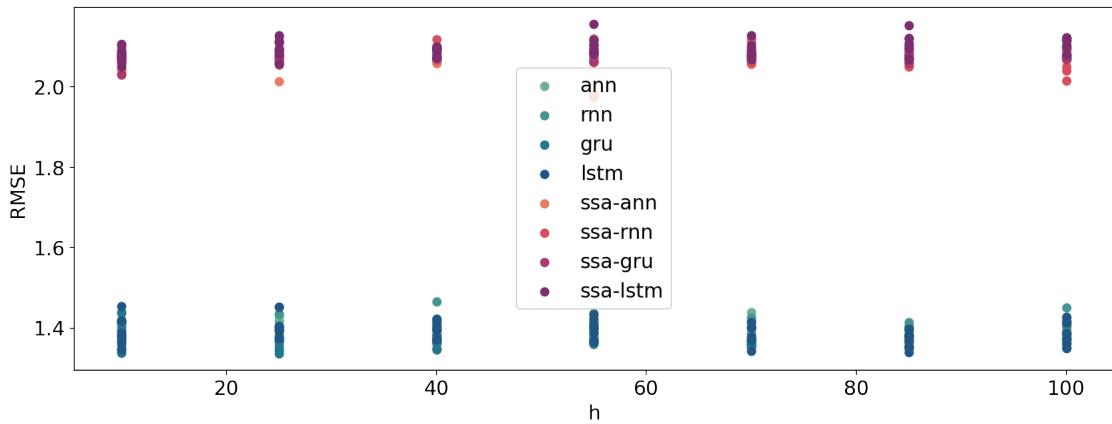


Рис. 4.88. Сумма синусов с красным шумом. Проверка устойчивости.  $r = 2$ ,  $L = 175$ .  $T = 12$ .

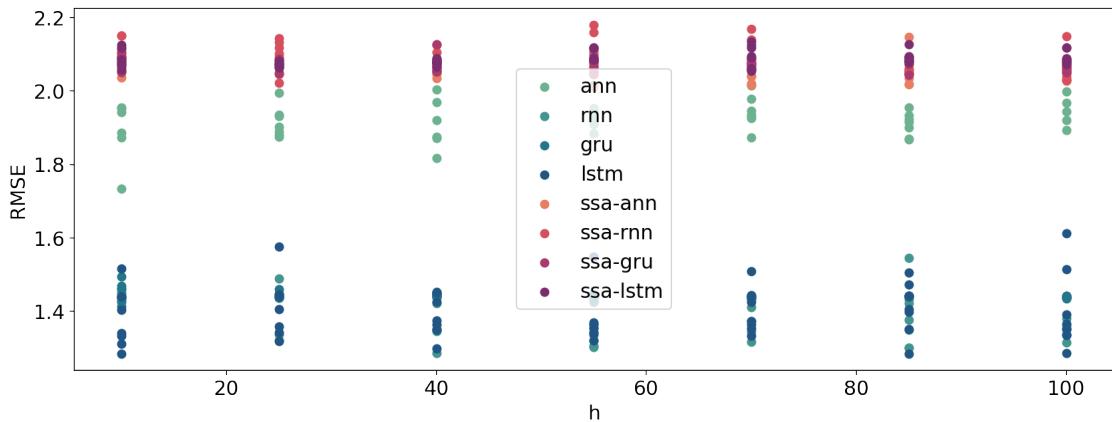


Рис. 4.89. Сумма синусов с красным шумом. Проверка устойчивости.  $r = 2$ ,  $L = 175$ .  $T = 84$ .

## Выводы

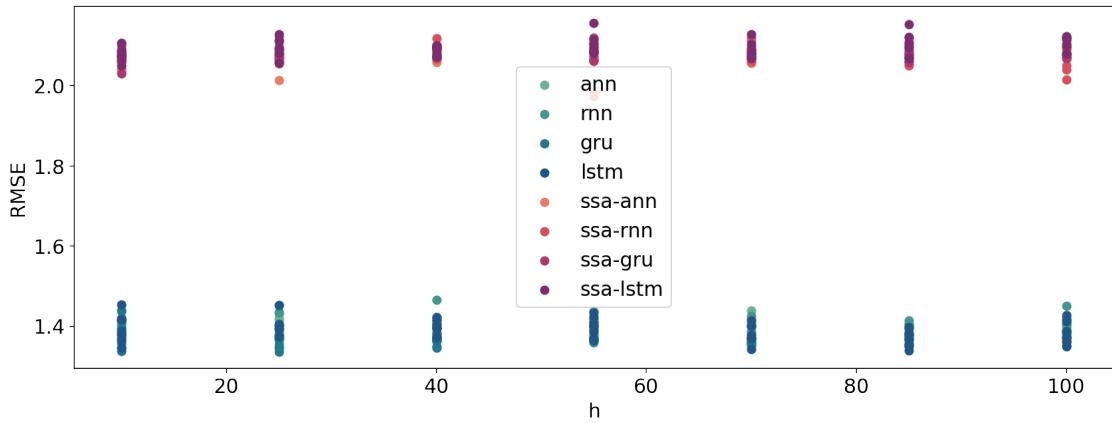


Рис. 4.90. Сумма синусов с красным шумом. Проверка устойчивости.  $r = 4$ ,  $L = 175$ .  $T = 12$

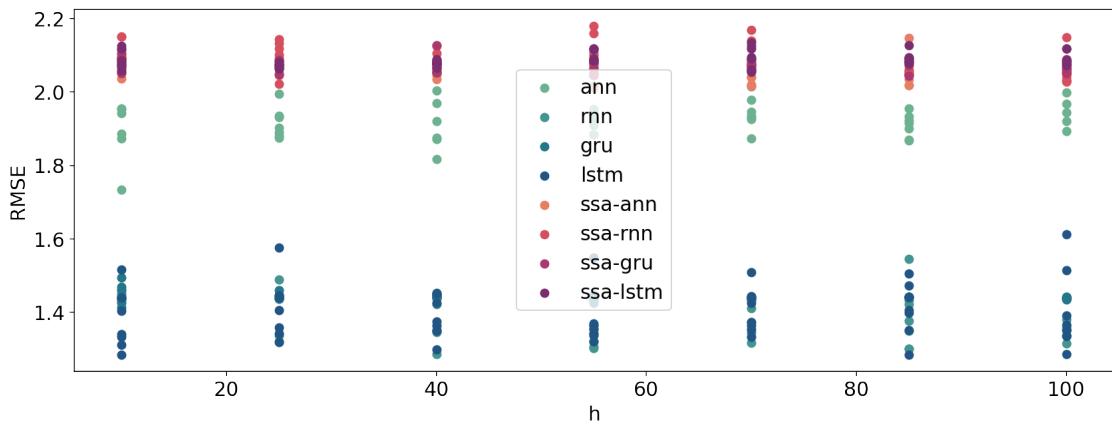


Рис. 4.91. Сумма синусов с красным шумом. Проверка устойчивости.  $r = 4$ ,  $L = 175$ .  $T = 84$ .

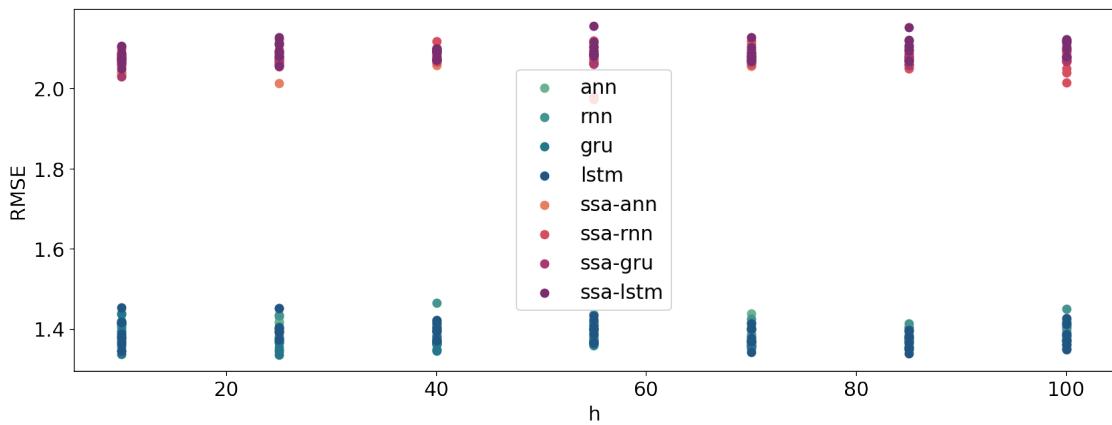


Рис. 4.92. Сумма синусов с красным шумом. Проверка устойчивости.  $r = 14$ ,  $L = 84$ .  $T = 12$

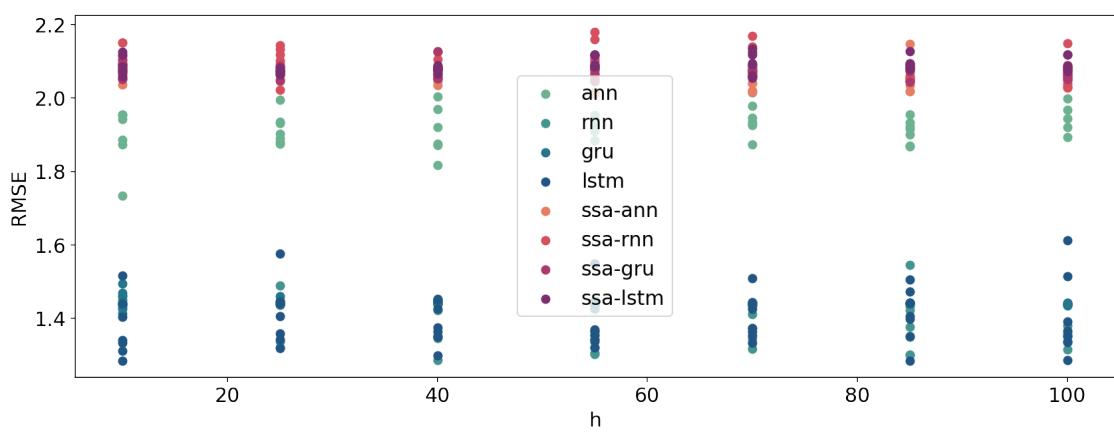


Рис. 4.93. Сумма синусов с красным шумом. Проверка устойчивости.  $r = 14$ ,  $L = 84$ .  $T = 84$ .

## Глава 5

### Реальные данные

#### 5.1. Среднемесячные осадки в Индии

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

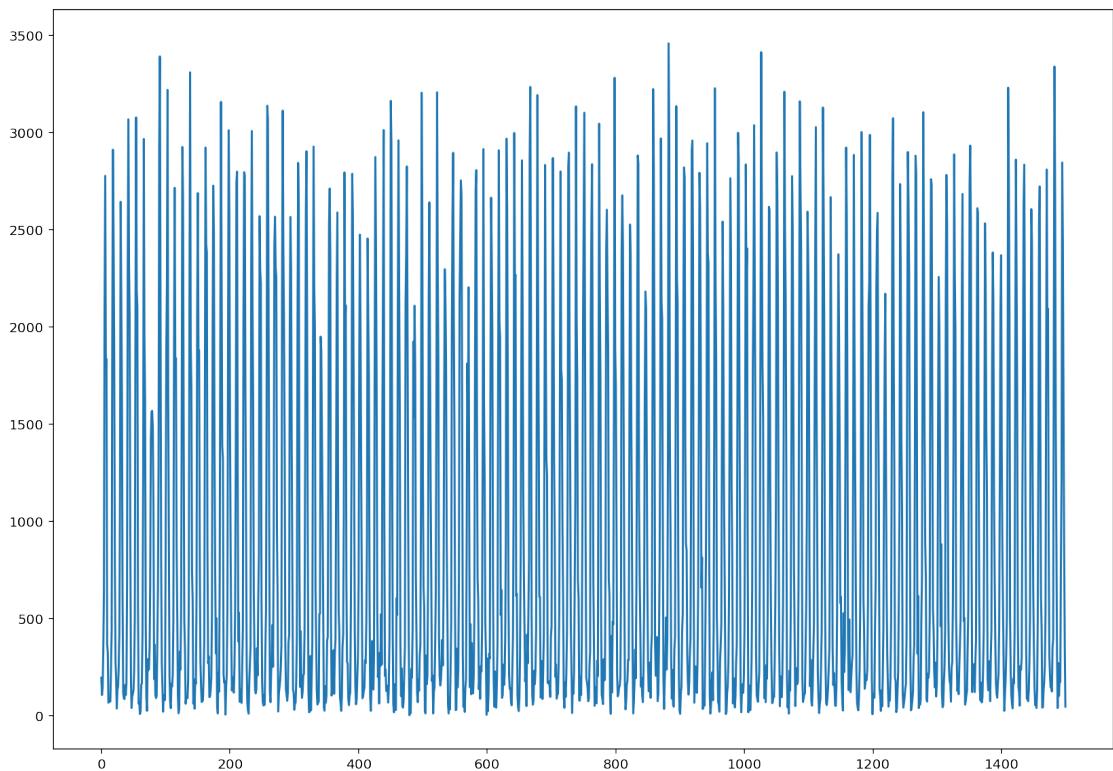


Рис. 5.1. Данные «Indian Rain».

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

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

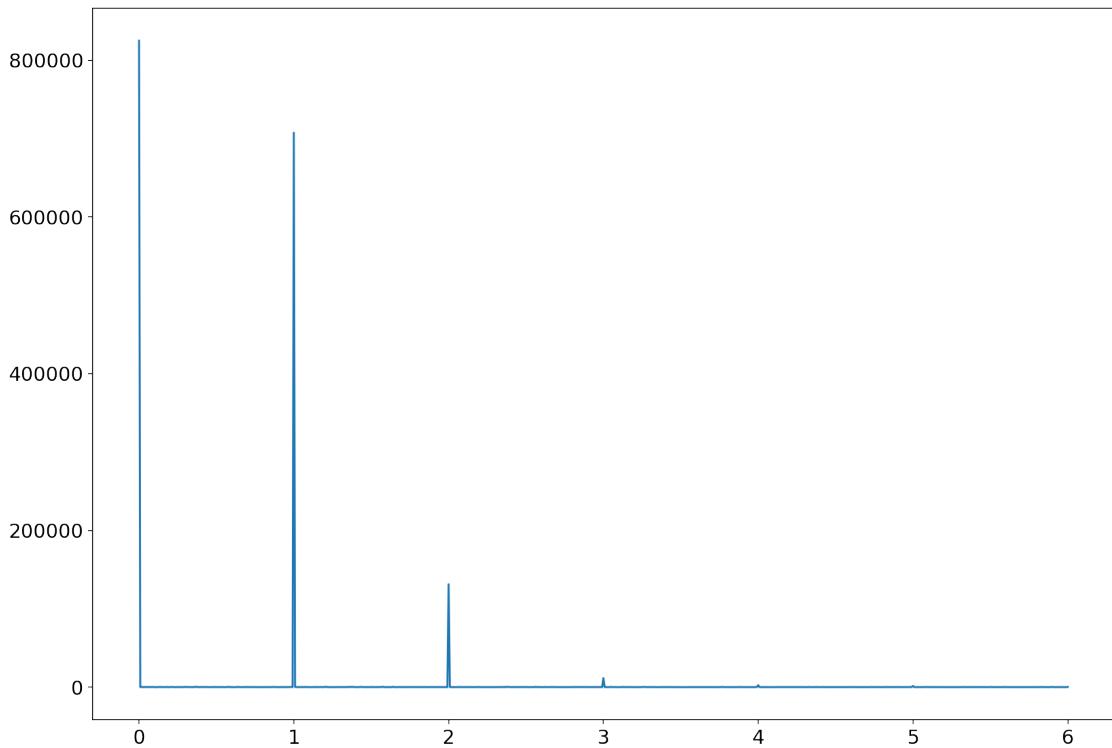


Рис. 5.2. Периодограмма ряда «Indian Rain».

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

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

#### Прогноз по SSA

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

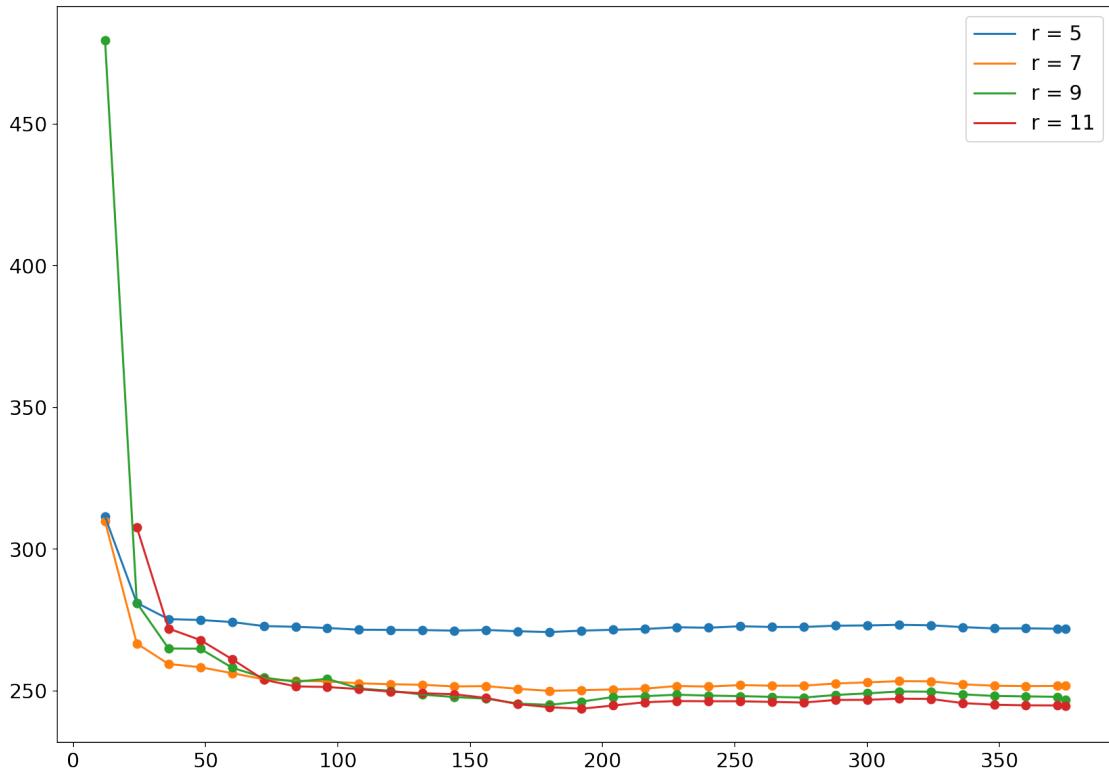


Рис. 5.3. Данные «Indian Rain». Прогнозирование с помощью метода SSA.

### 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.

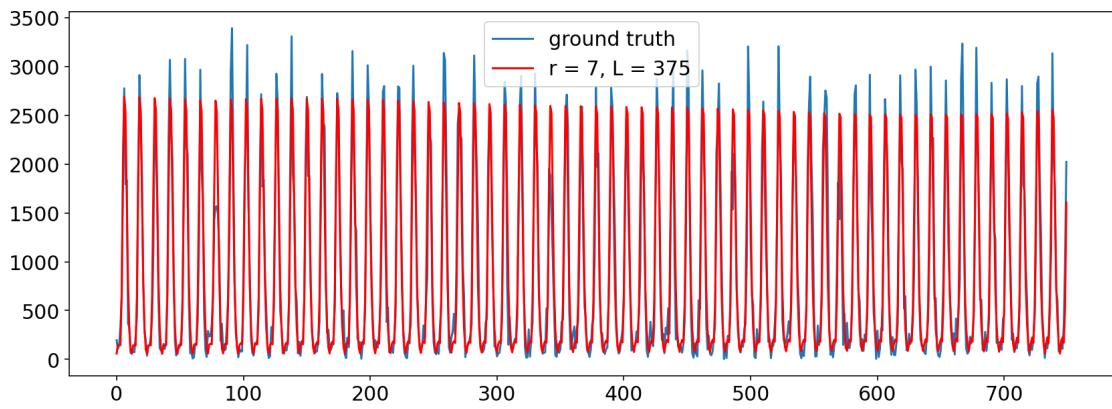


Рис. 5.4. Данные «Indian Rain». Восстановление тренировочной выборки с помощью метода SSA.  $r = 7$ ,  $L = 375$ .

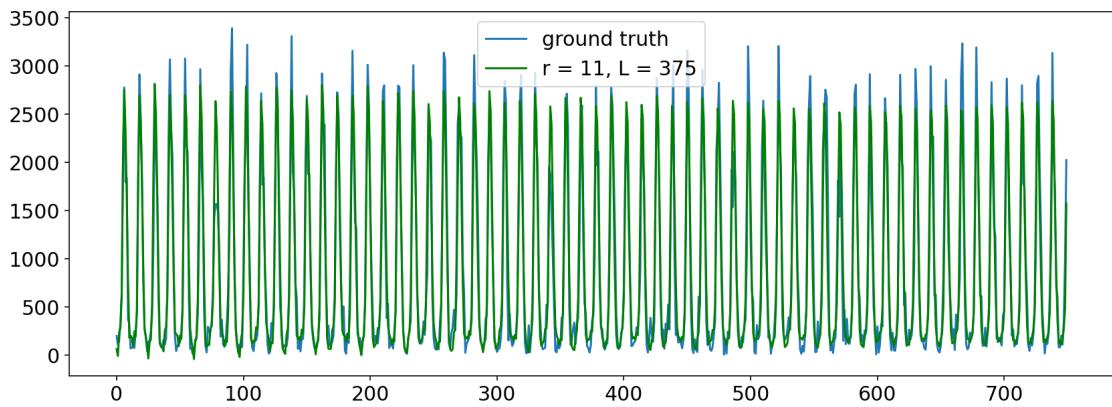


Рис. 5.5. Данные «Indian Rain». Восстановление тренировочной выборки с помощью метода SSA.  $r = 11$ ,  $L = 375$ .

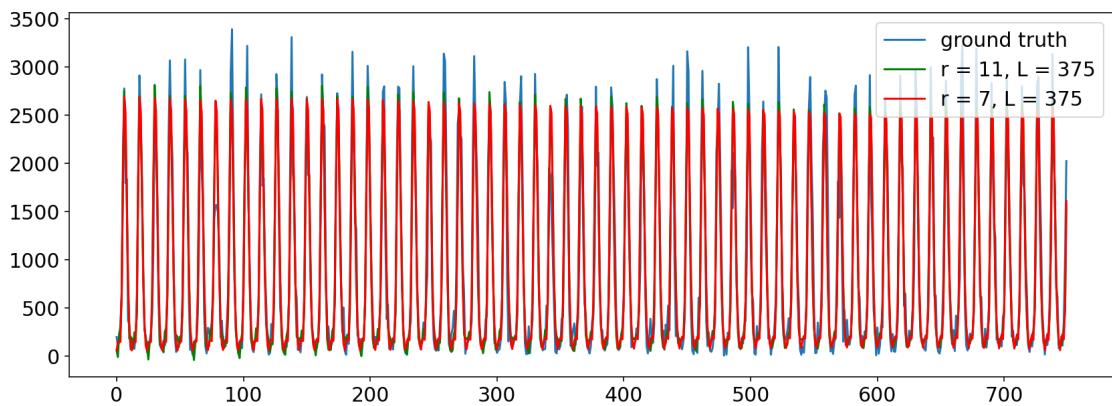


Рис. 5.6. Данные «Indian Rain». Восстановление тренировочной выборки с помощью метода SSA. Обе пары.

## Сравнение методов

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

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

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

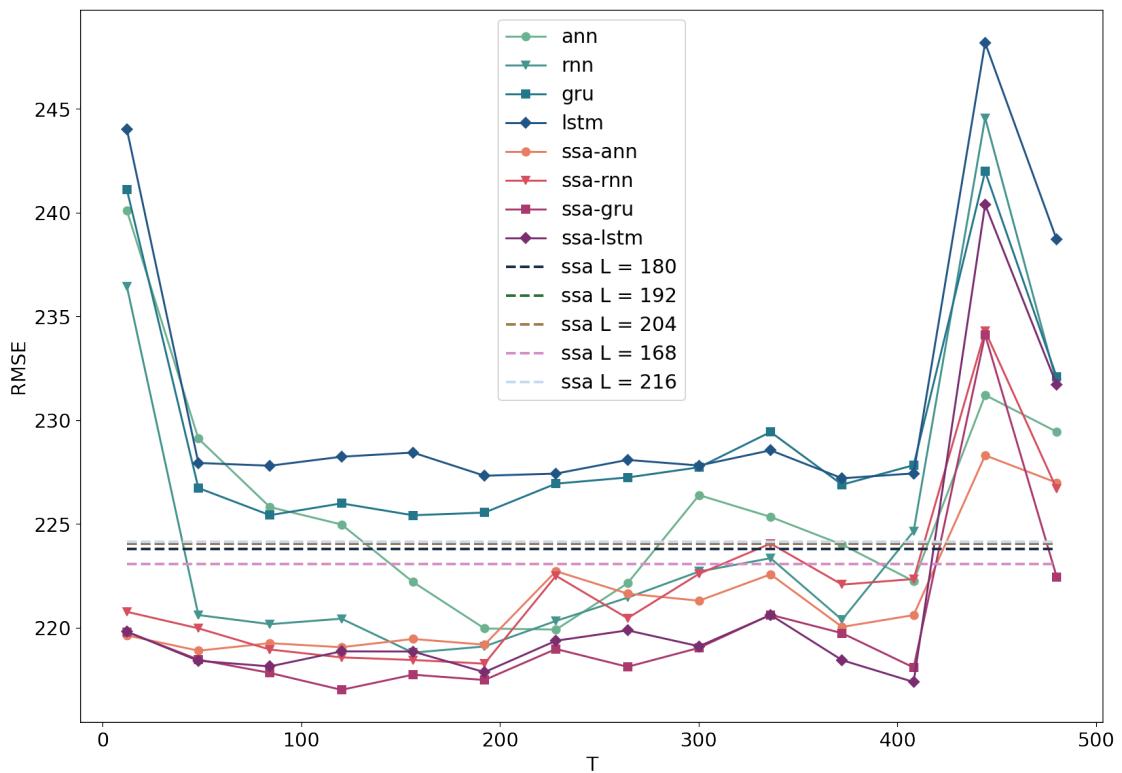


Рис. 5.7. Данные «Indian Rain». Ошибки прогноза в зависимости от параметра  $T$ .

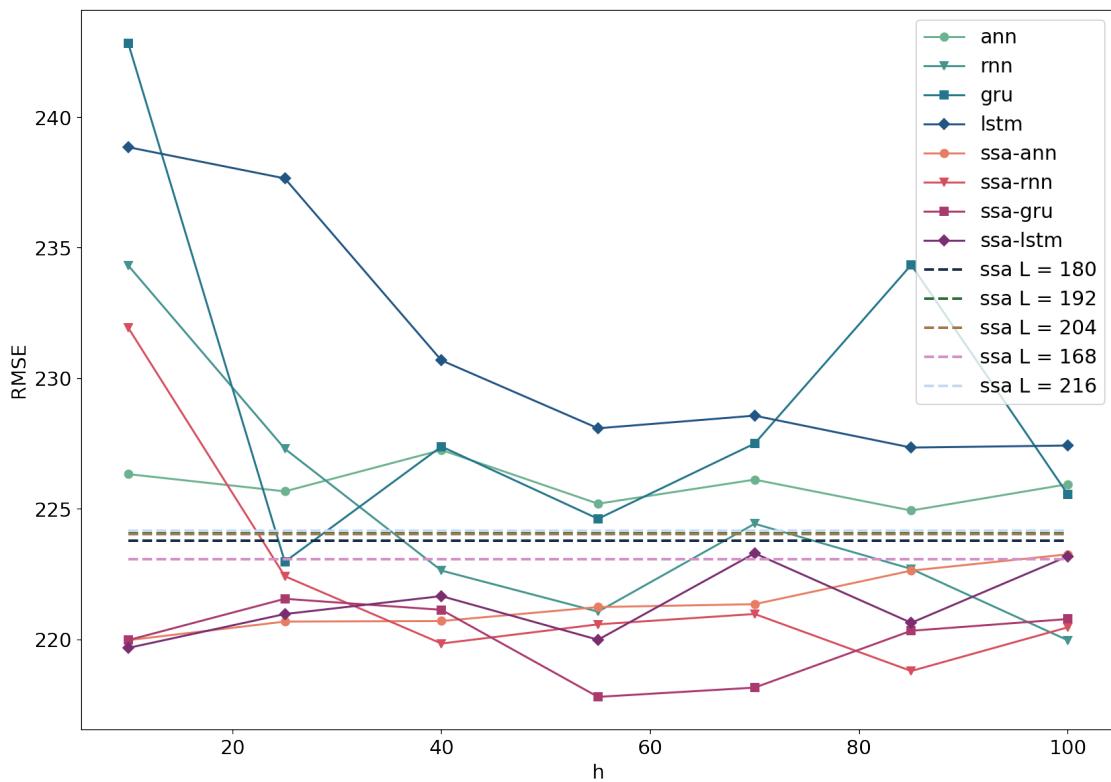


Рис. 5.8. Данные «Indian Rain». Ошибки прогноза в зависимости от параметра  $h$ .

### Отображение прогнозов

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

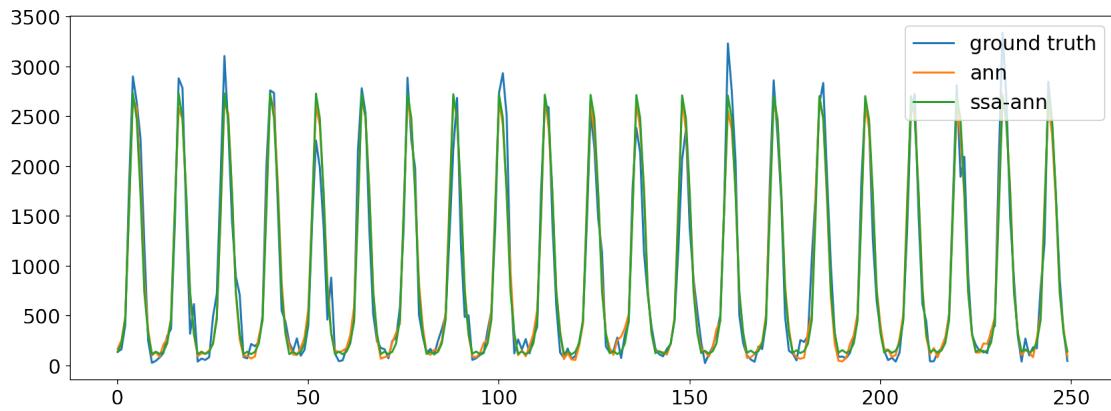


Рис. 5.9. Данные «Indian Rain». Прогноз для ANN и SSA-ANN.

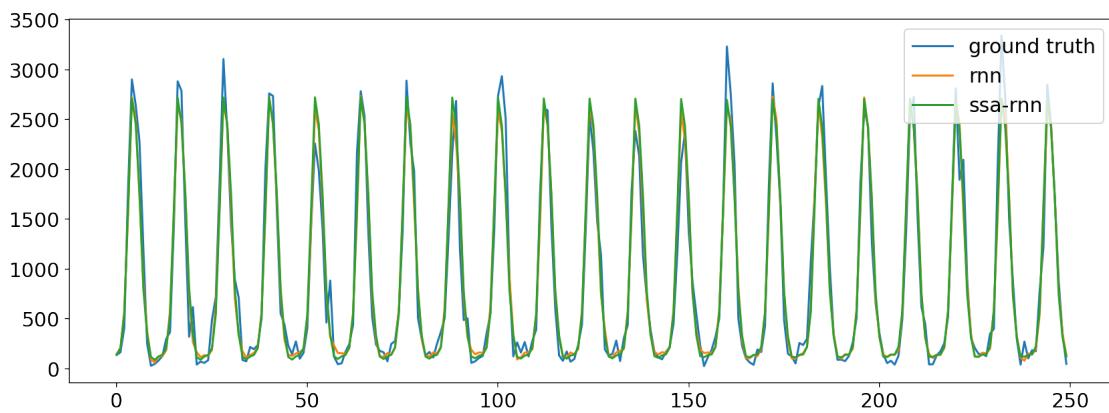


Рис. 5.10. Данные «Indian Rain». Прогноз для RNN и SSA-RNN.

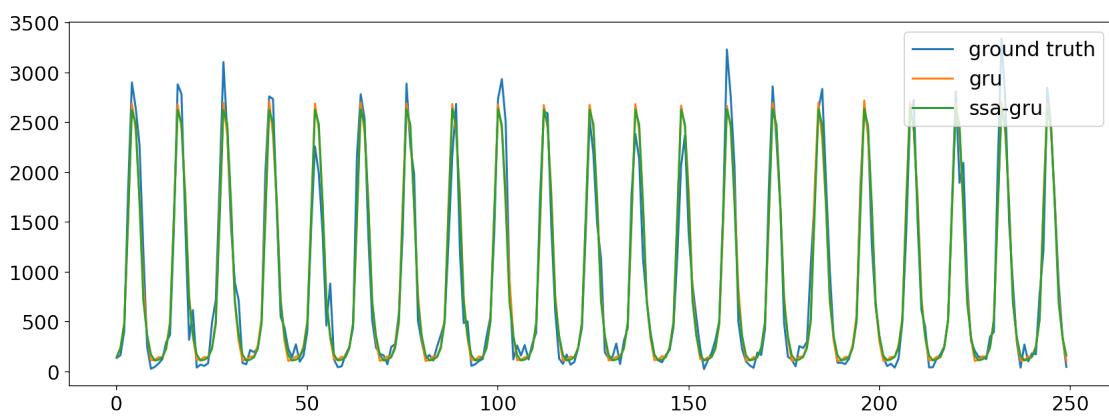


Рис. 5.11. Данные «Indian Rain». Прогноз для GRU и SSA-GRU.

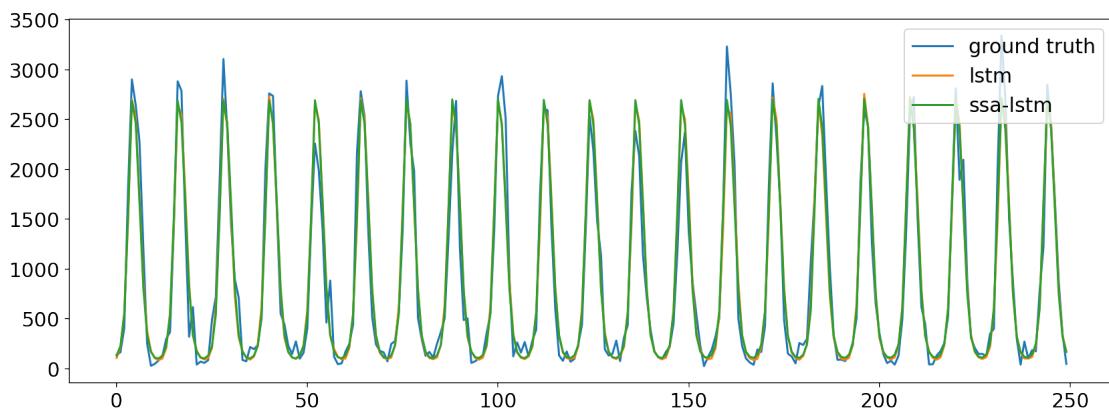


Рис. 5.12. Данные «Indian Rain». Прогноз для LSTM и SSA-LSTM.

## Проверка устойчивости

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

{12, 156}. Сетка для параметр  $h$  останется прежней. Будем получать каждый результат по 7 раз, инициализируя метод с новыми весами. Полученные результаты отображены на рис. 5.13, 5.14.

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

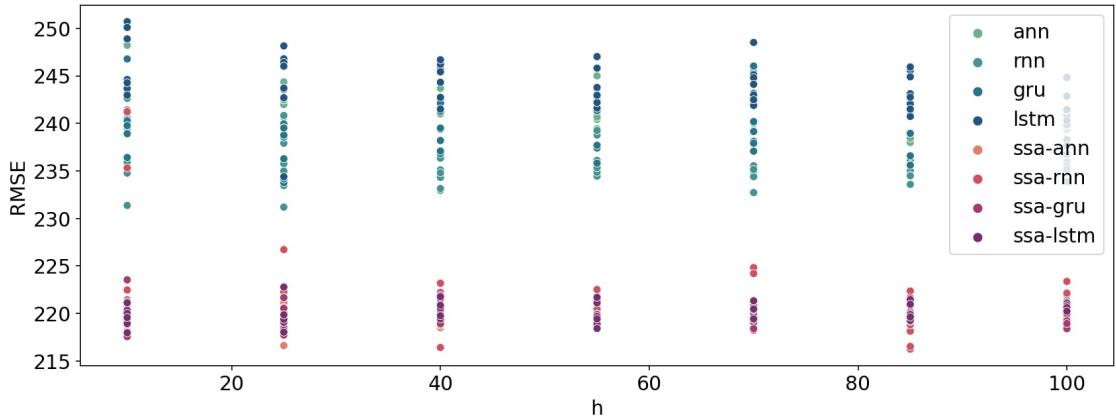


Рис. 5.13. Данные «Indian Rain». Проверка устойчивости.  $T = 12$ .

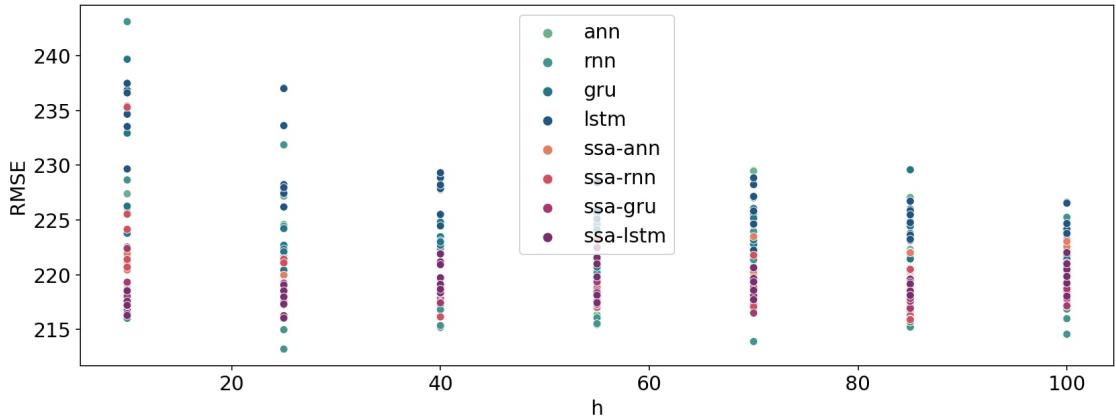


Рис. 5.14. Данные «Indian Rain». Проверка устойчивости.  $T = 156$ .

## 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. Данные Earth Orientation Parameters (EOP)

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

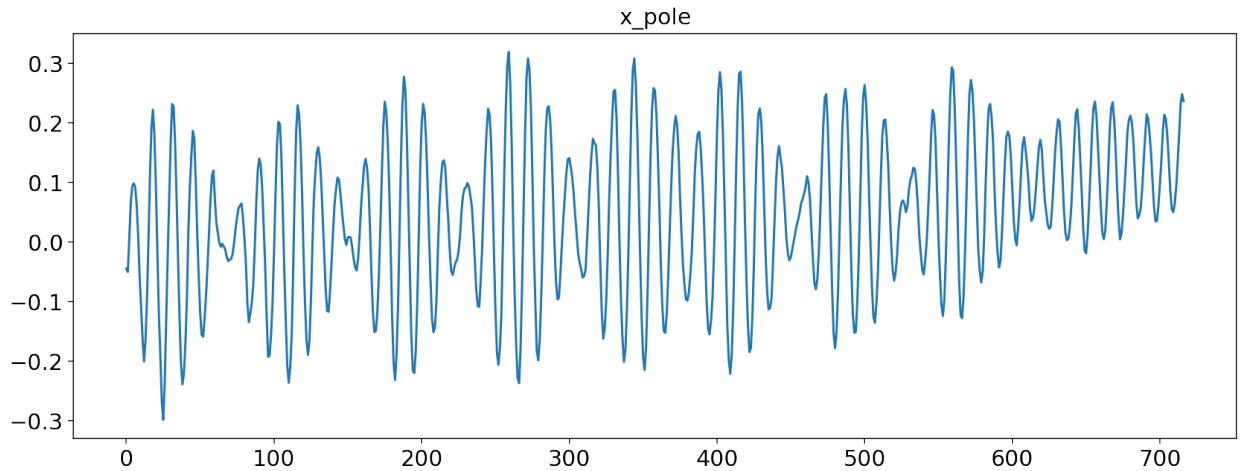


Рис. 5.15. Временной ряд « $x$  pole».

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

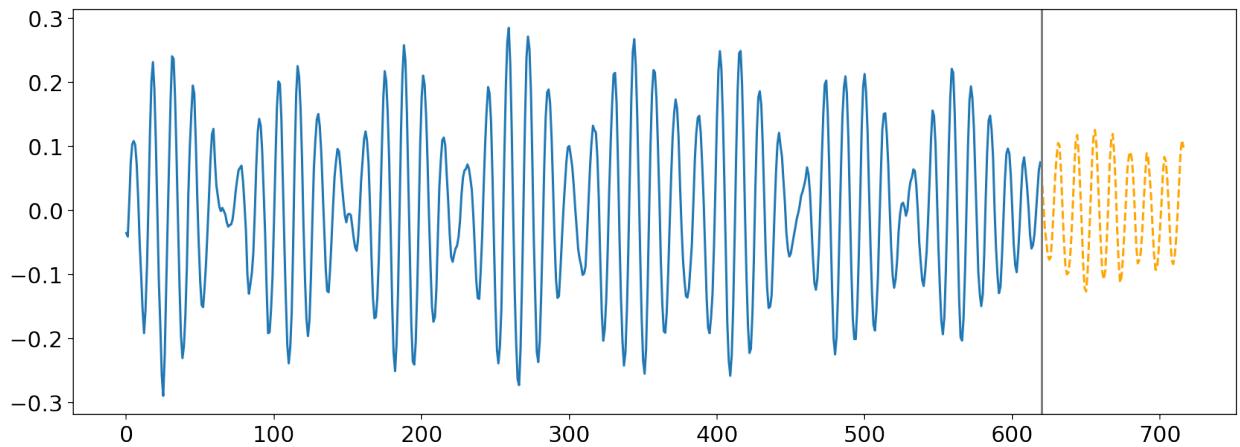


Рис. 5.16. Временной ряд « $x$  pole» без тренда.

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

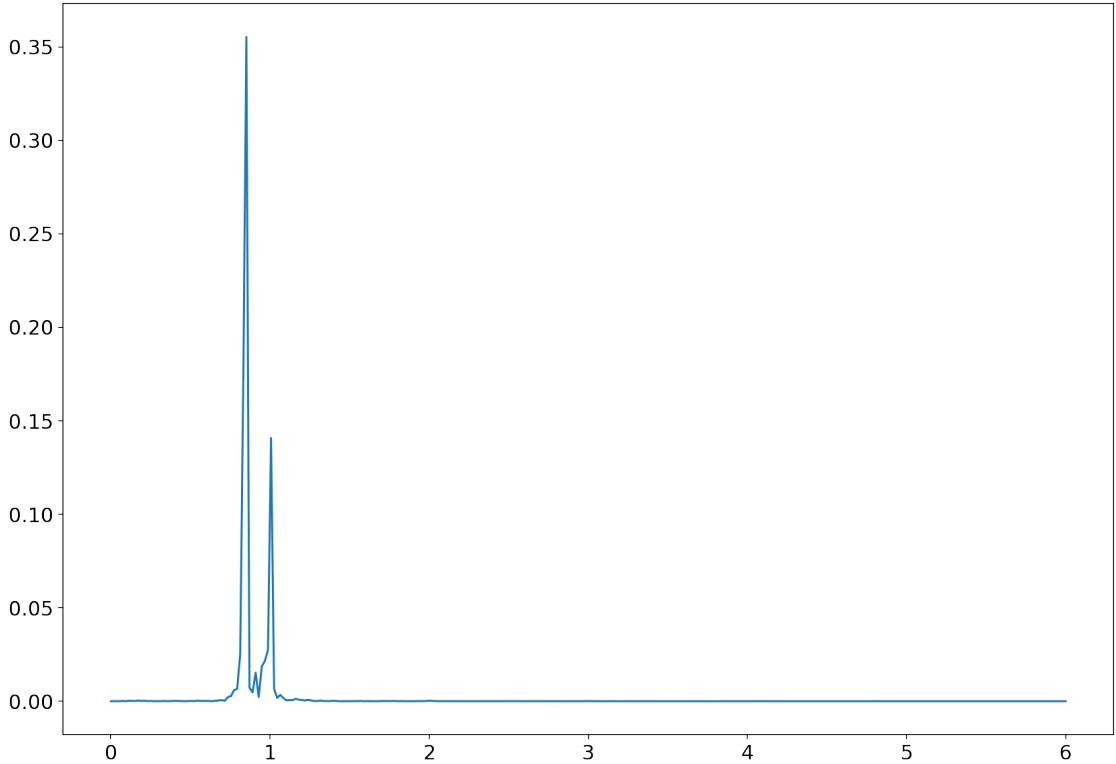


Рис. 5.17. Периодограмма ряда «x pole» без тренда.

### 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  $L = \{13, 26, \dots, 160\}$ ,  $r = \{6, 8, 12, 16, 18\}$ .

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

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

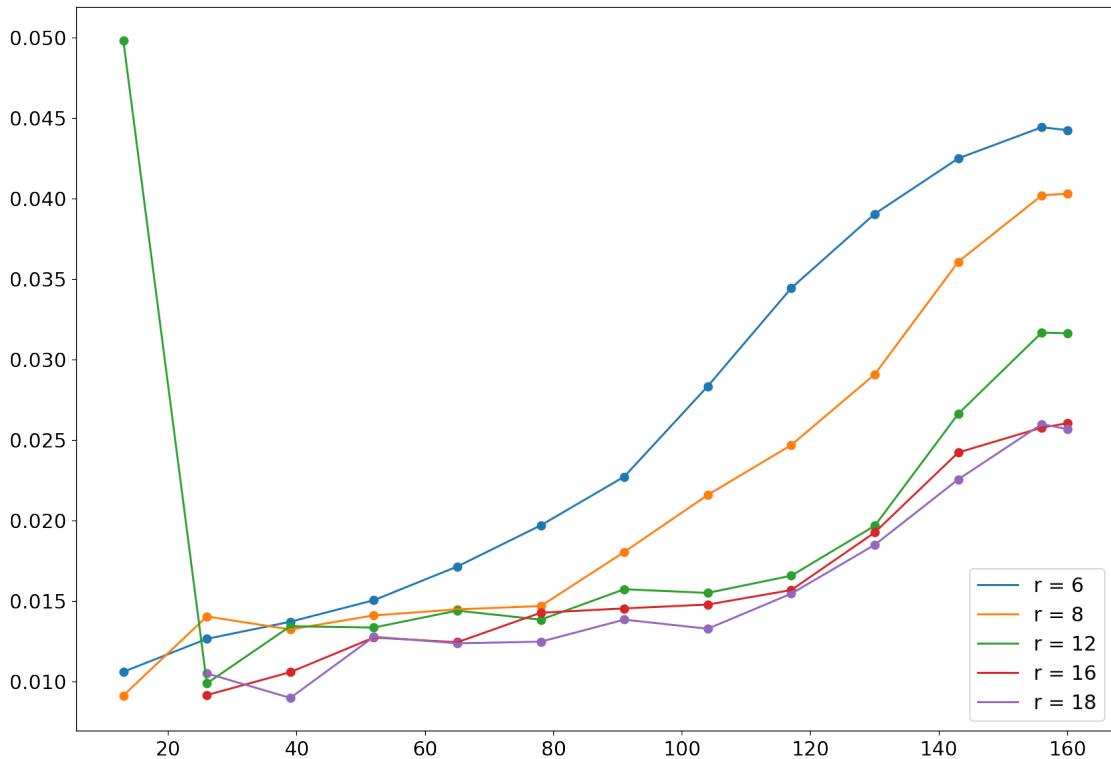


Рис. 5.18. Ряд «x pole» без тренда. Прогнозирование с помощью метода SSA.

## Восстановление SSA

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

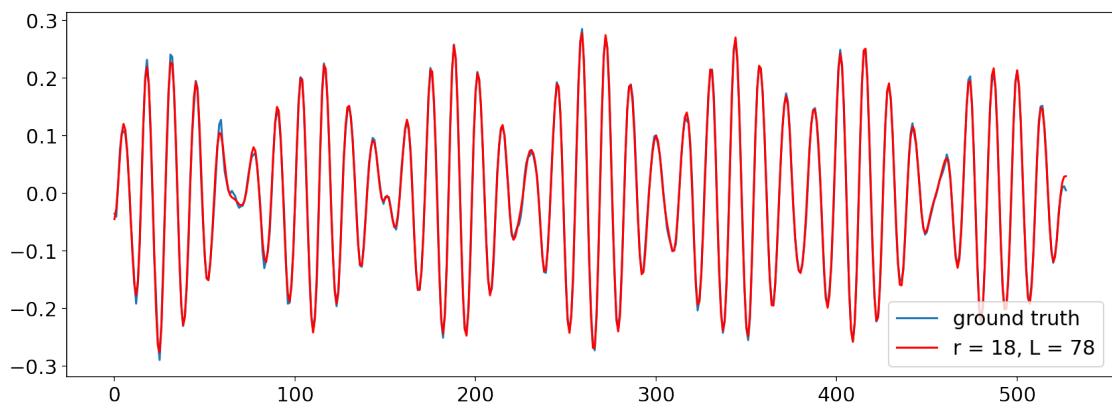


Рис. 5.19. Ряд «x pole» без тренда. Восстановление тренировочной выборки с помощью метода SSA.  $r = 18$ ,  $L = 78$

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

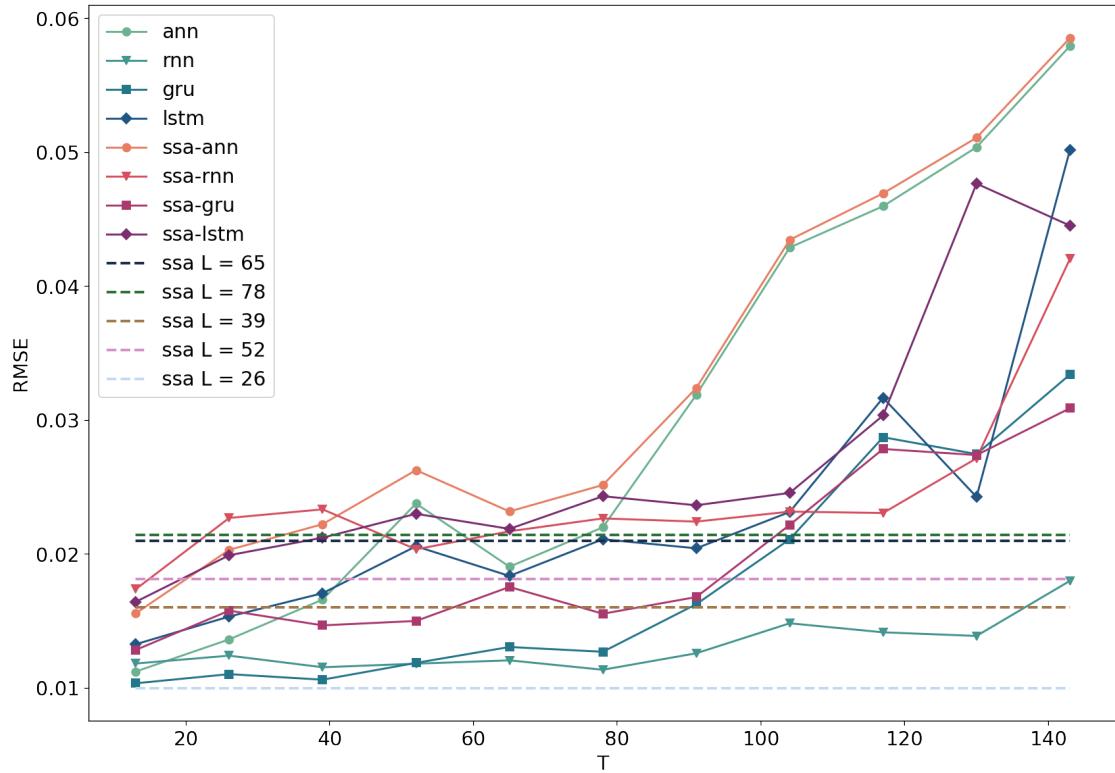


Рис. 5.20. Ряд «x pole» без тренда. Ошибки прогноза в зависимости от параметра  $T$ .

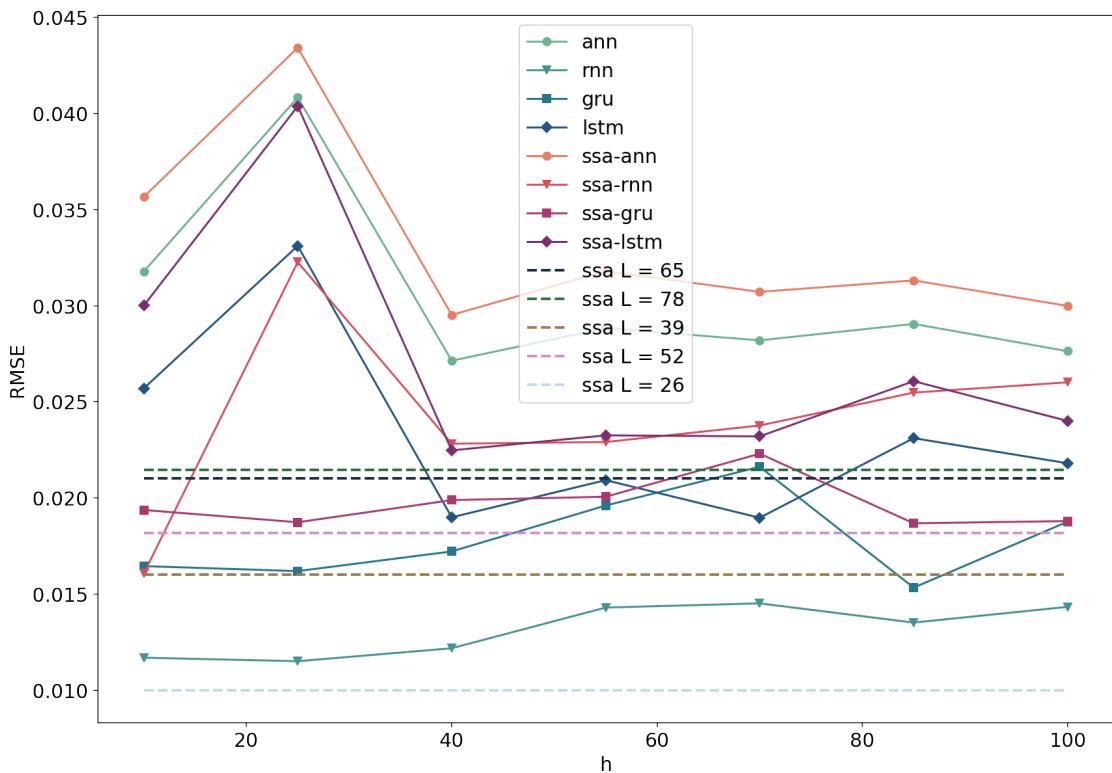


Рис. 5.21. Ряд «x pole» без тренда. Ошибки прогноза в зависимости от параметра  $h$ .

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

### Отображение 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 5.23. Также есть сильное расхождение с сигналом ряда в конце прогноза (правая часть графиков). Для всех гибридных методов ошибка в этих местах больше, чем для обычных. В остальных местах значительной разницы между прогнозами не замечено.

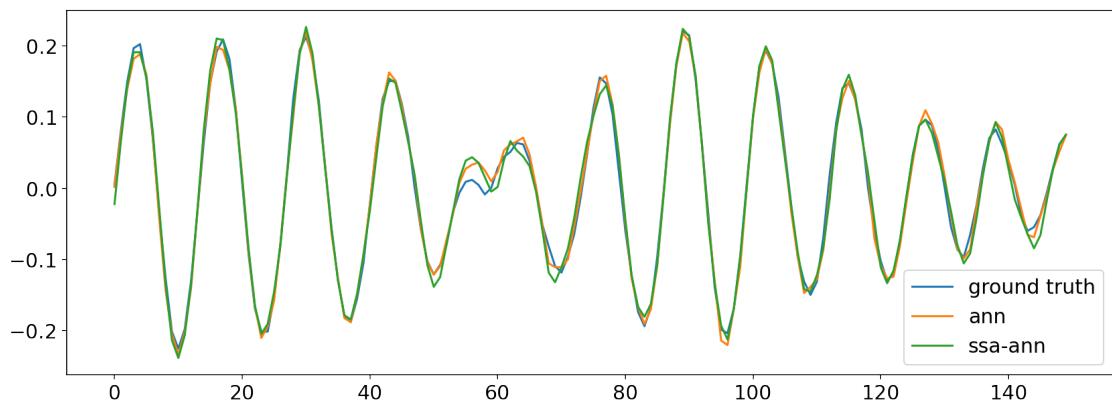


Рис. 5.22. Ряд «x pole» без тренда. Прогноз для ANN и SSA-ANN.

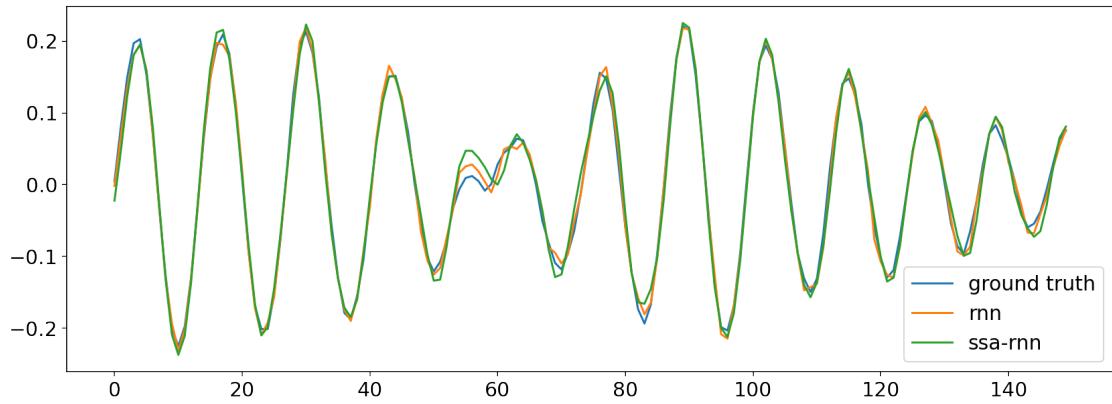


Рис. 5.23. Ряд «x pole» без тренда. Прогноз для RNN и SSA-RNN.

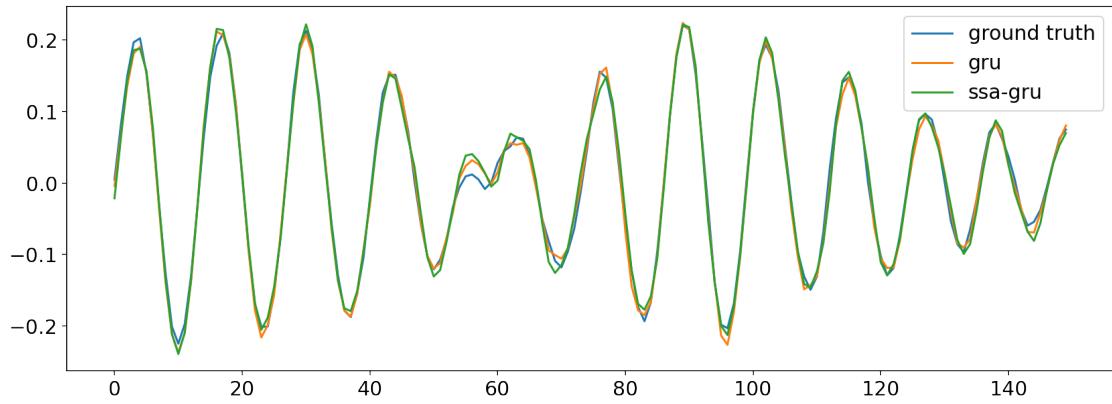


Рис. 5.24. Ряд «x pole» без тренда. Прогноз для GRU и SSA-GRU.

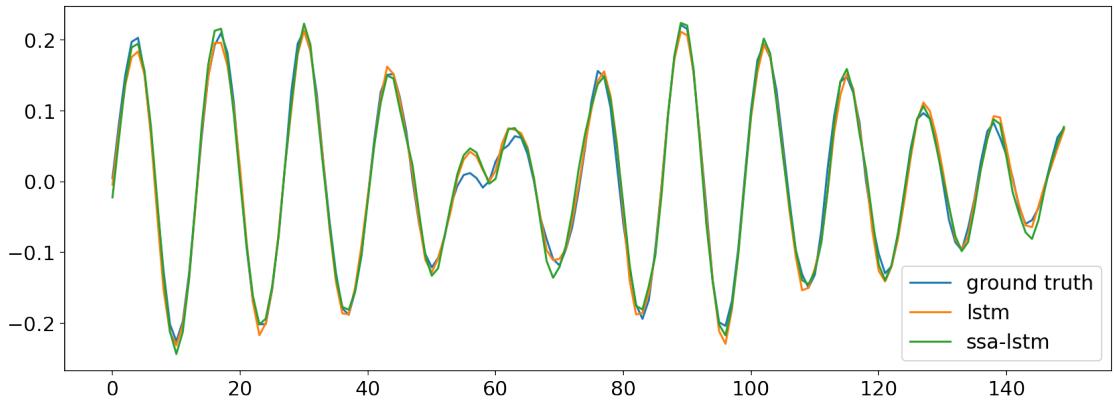


Рис. 5.25. Ряд «x pole» без тренда. Прогноз для LSTM и SSA-LSTM.

### Проверка устойчивости

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

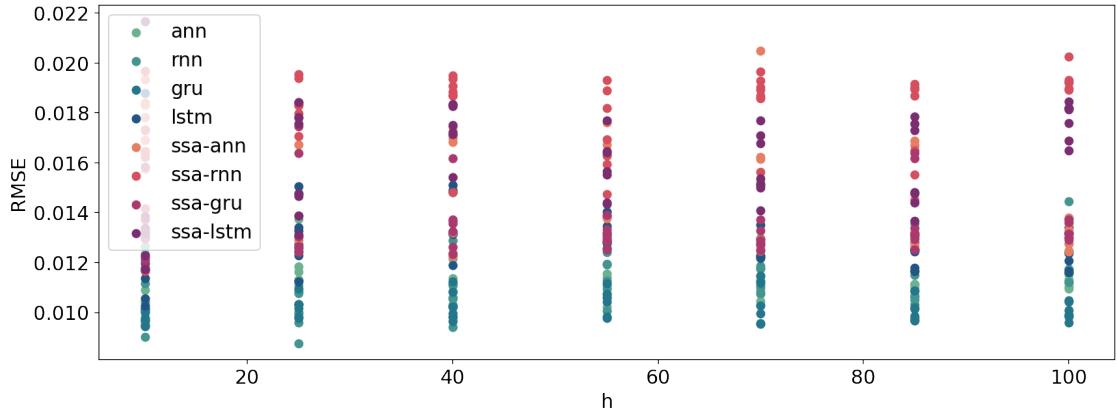


Рис. 5.26. Ряд «x pole» без тренда. Проверка устойчивости.  $T = 13$ .

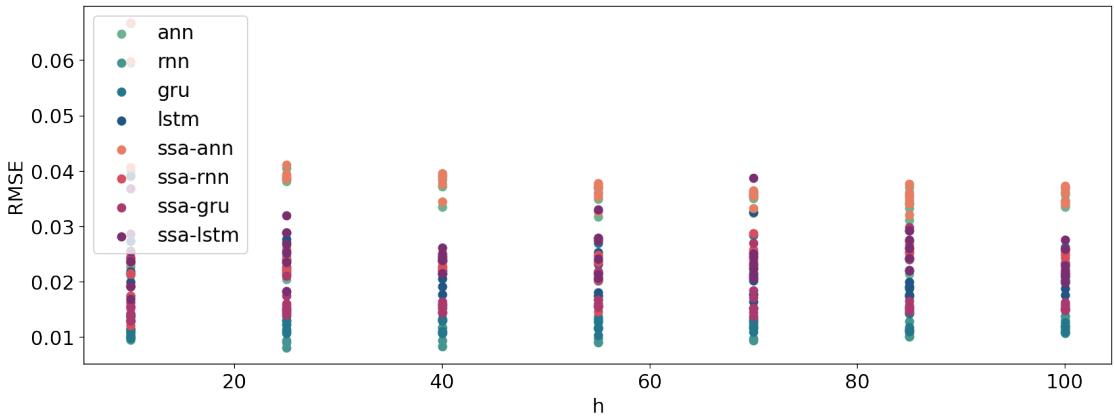


Рис. 5.27. Ряд «x pole» без тренда. Проверка устойчивости.  $T = 91$ .

## 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 для обычного ряда. Использование таких параметров в гибридных моделях неподходящим, так как оценка сигнала слишком сильно аппроксимирует ряд. Гибридные методы показали результаты хуже, чем обычные методы.

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 2 местах, что говорит о том, что SSA некорректно выделяет сигнал, а это 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 пре-процессинга SSA. Как было показано в эксперименте 4.1.2 в случае маленького шума, логично выбрать параметр  $r$  оптимальный 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 оценкой сигнала ряда, что

делает препроцессинг SSA бессмысленным.

### 5.3. Погода

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

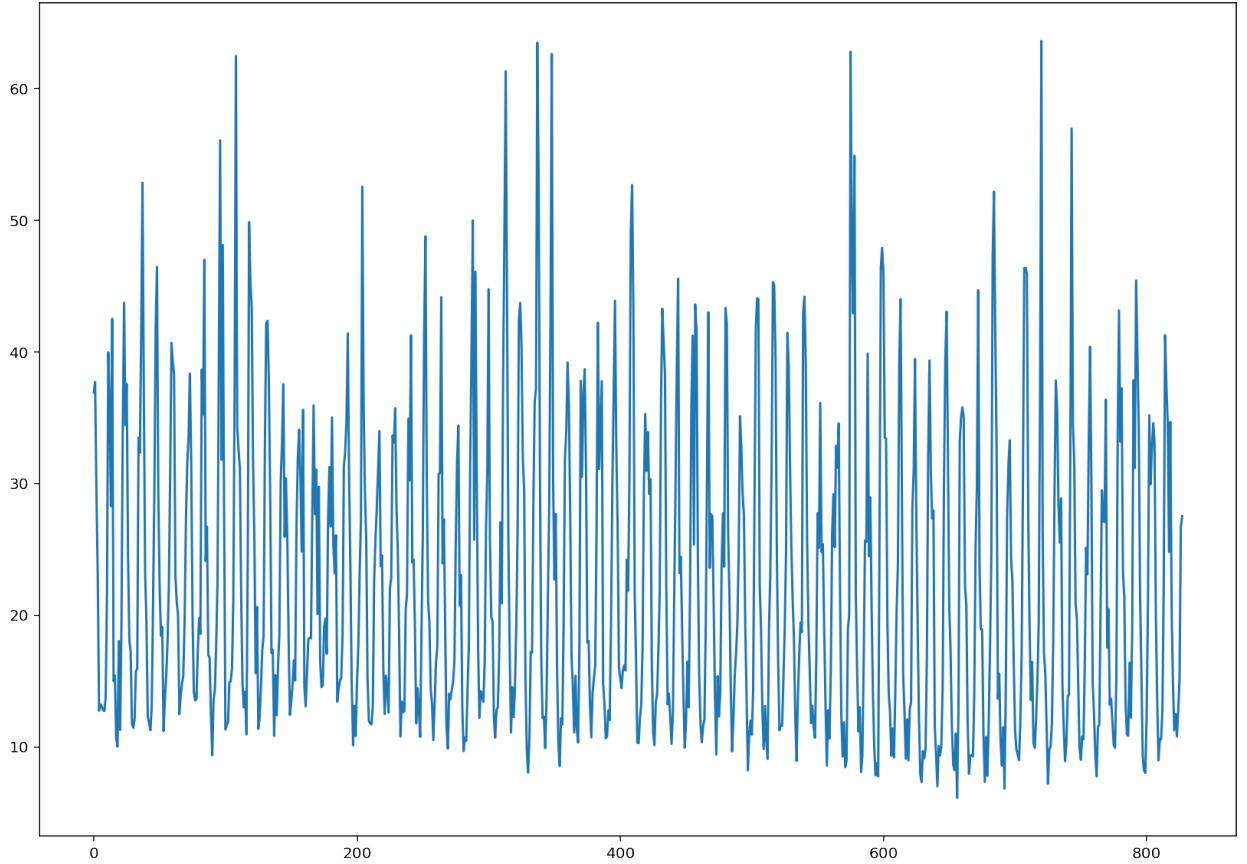


Рис. 5.28. Временной ряд характеристики погоды в Санкт-Петербурге.

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

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

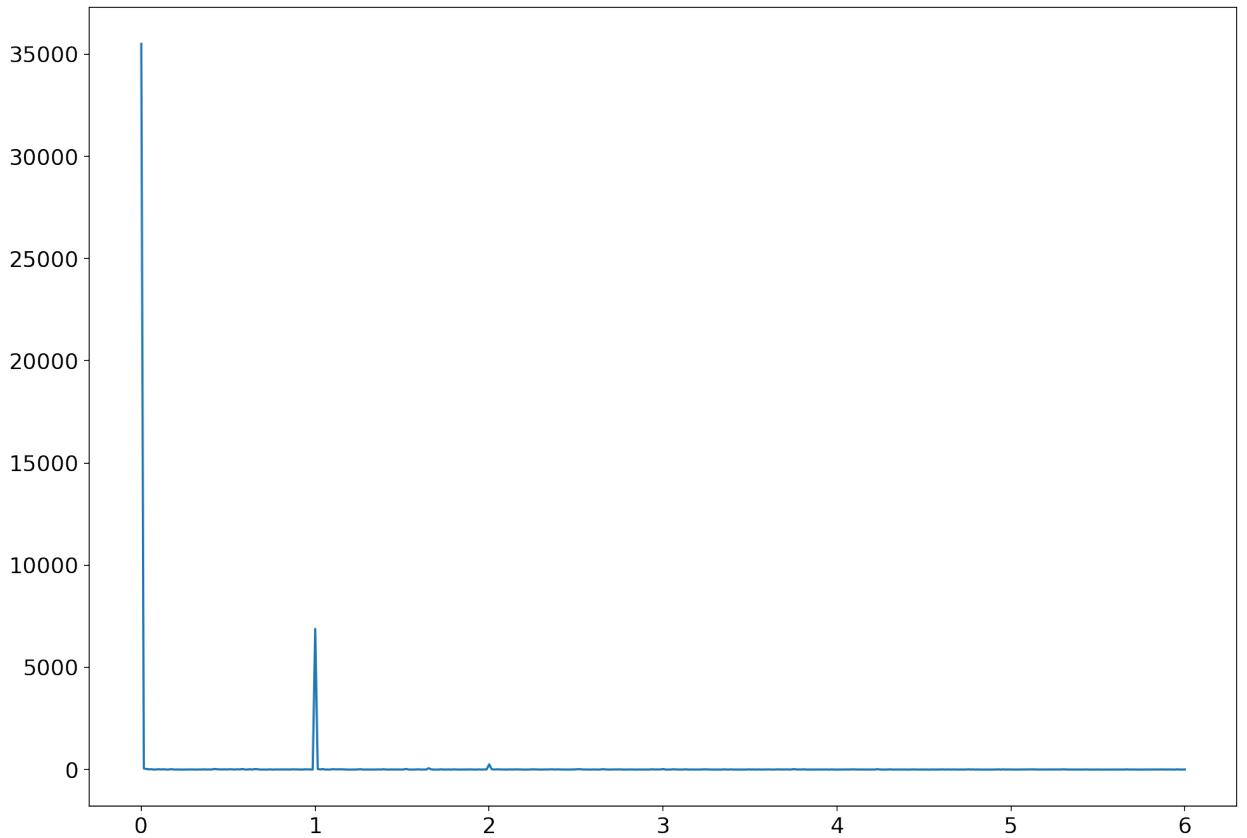


Рис. 5.29. Периодограмма ряда характеристики погоды.

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

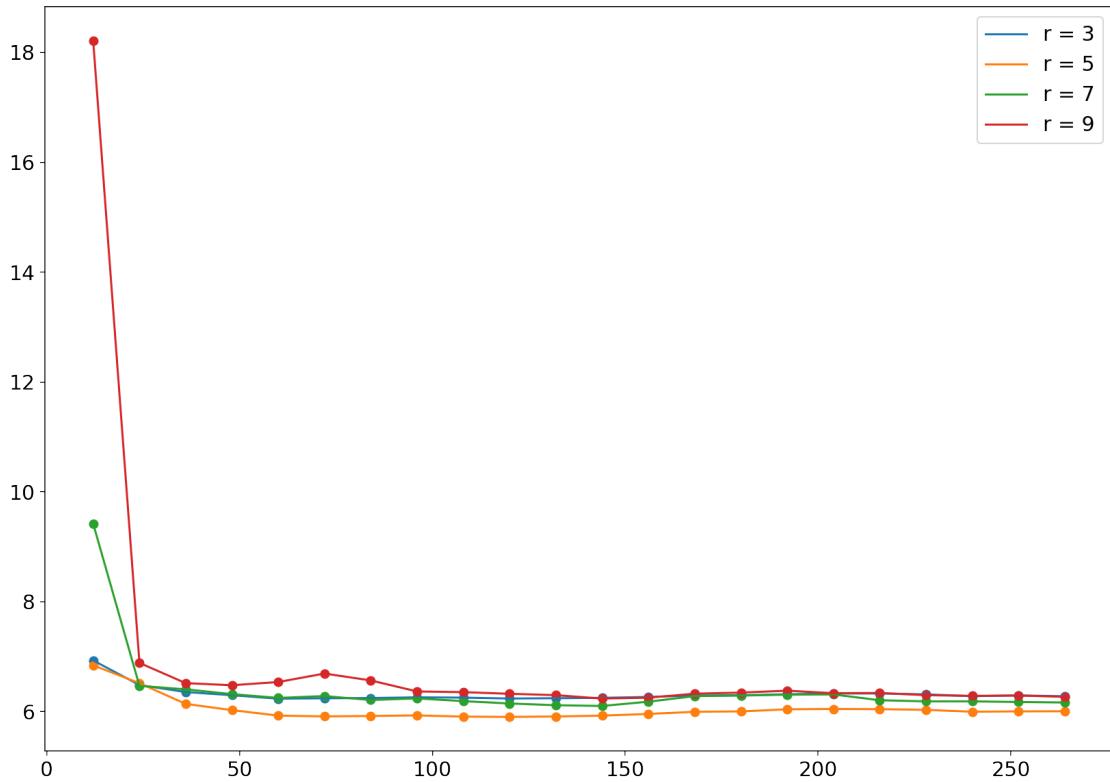


Рис. 5.30. Погода в Санкт-Петербурге. Прогнозирование с помощью метода SSA.

### Восстановление SSA

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

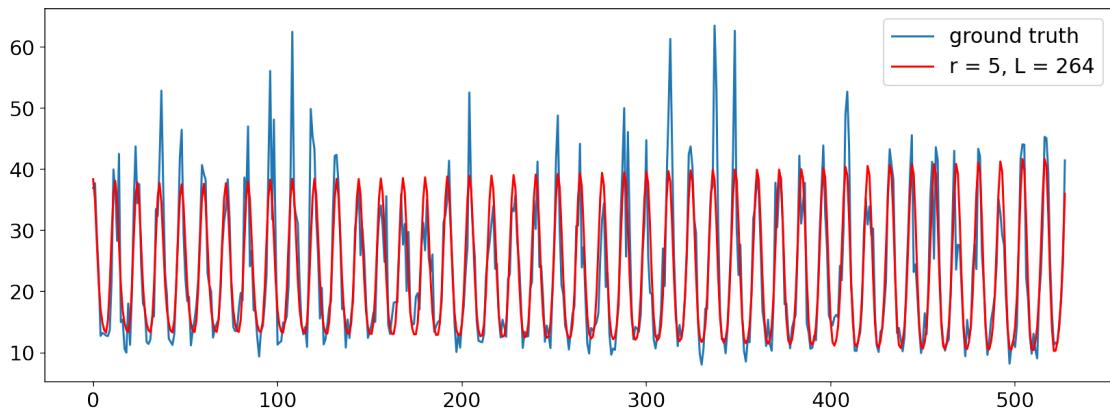


Рис. 5.31. Погода в Санкт-Петербурге. Восстановление тренировочной выборки с помощью метода SSA.  $r = 5$ ,  $L = 264$

## Сравнение методов

Пусть задана следующая сетка параметров:  $T = \{12, 48, \dots, 120\}$ ,  $h = \{10, 25, \dots, 100\}$ . Параметры для SSA в гибридных методах выберем  $L = 264$ ,  $r = 5$ . Для метода SSA зафиксируем параметр  $r = 5$ , а параметр  $L$  будем перебирать по сетке  $\{12, 24, \dots, 264\}$ .

Посмотрим на результаты на рис. 5.32, 5.33, видно, что лучшие результаты показывает метод SSA. Сложно оценить, какие методы показали себя лучше. На графике все результаты перемешаны, но кривые гибридных методов лежат немного ниже, чем обычных. Тем не менее, в некоторых частных случаях гибридные методы показывают хорошую точность. Например SSA-ANN на рис. 5.32 и SSA-ANN, SSA-RNN на рис. 5.33. Такие результаты могли получиться из-за маленького количества данных. А также возможно малого размера валидационной выборки, что ограничивает выбор параметра  $T$ .

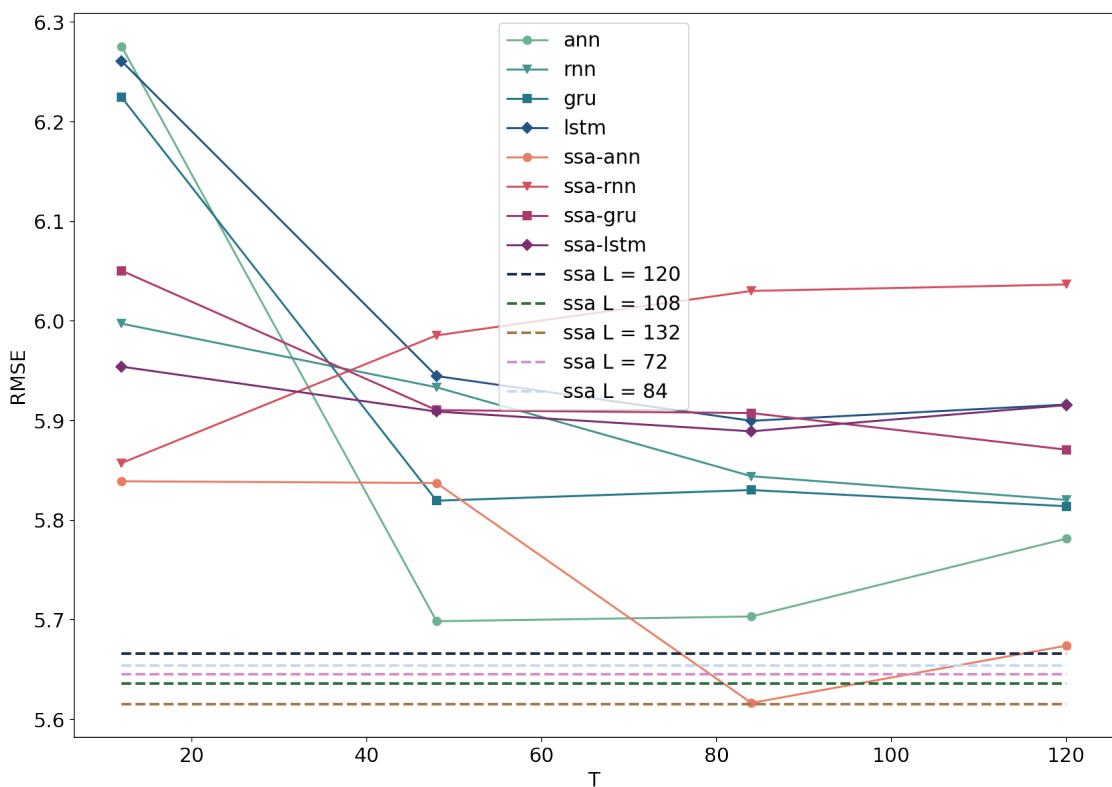


Рис. 5.32. Погода в Санкт-Петербурге. Ошибки прогноза в зависимости от параметра  $T$ .

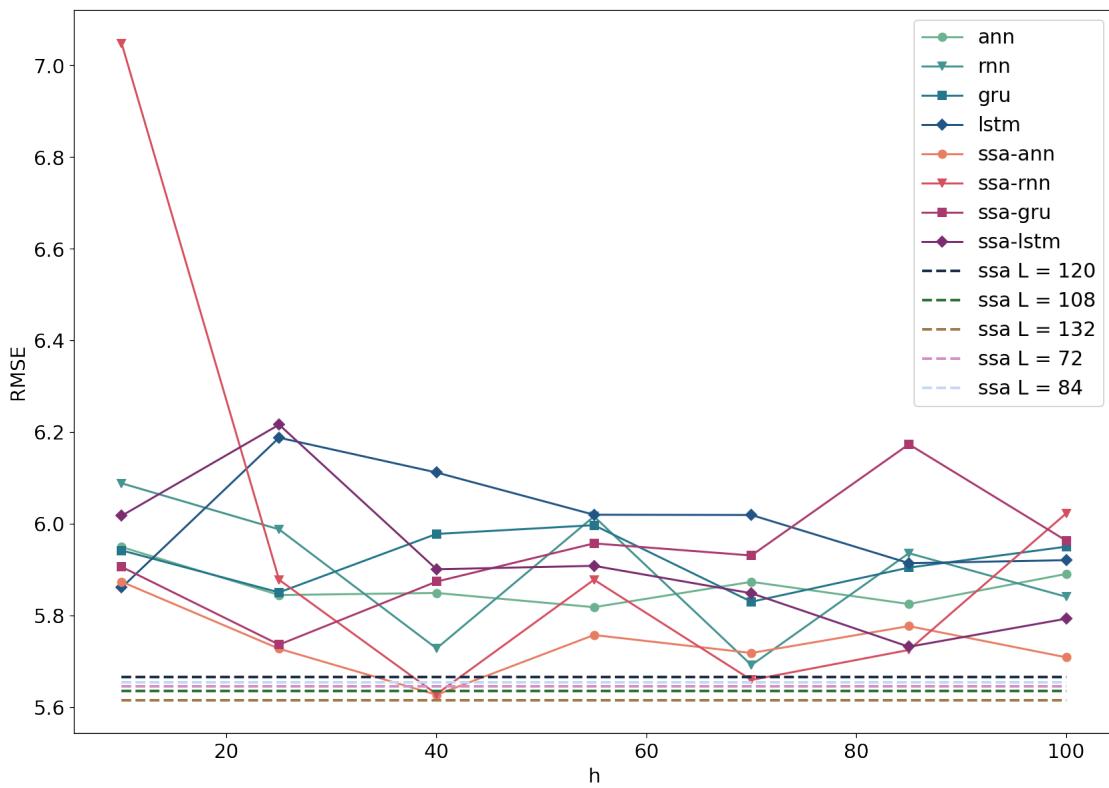


Рис. 5.33. Погода в Санкт-Петербурге. Ошибки прогноза в зависимости от параметра  $h$ .

### Отображение прогнозов

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

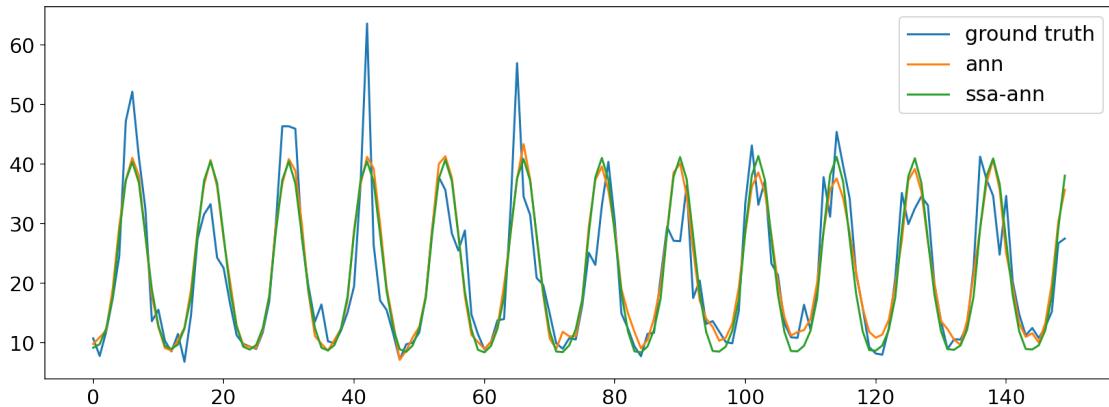


Рис. 5.34. Погода в Санкт-Петербурге. Прогноз для ANN и SSA-ANN.

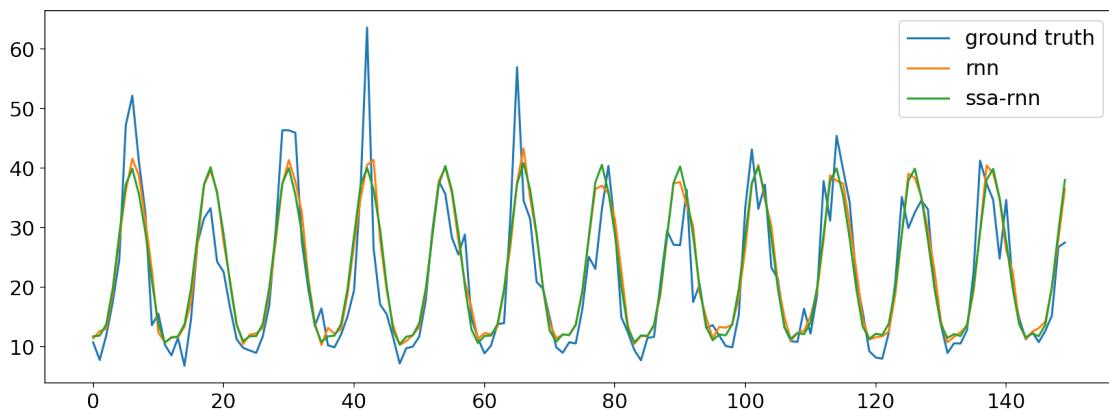


Рис. 5.35. Погода в Санкт-Петербурге. Прогноз для RNN и SSA-RNN.

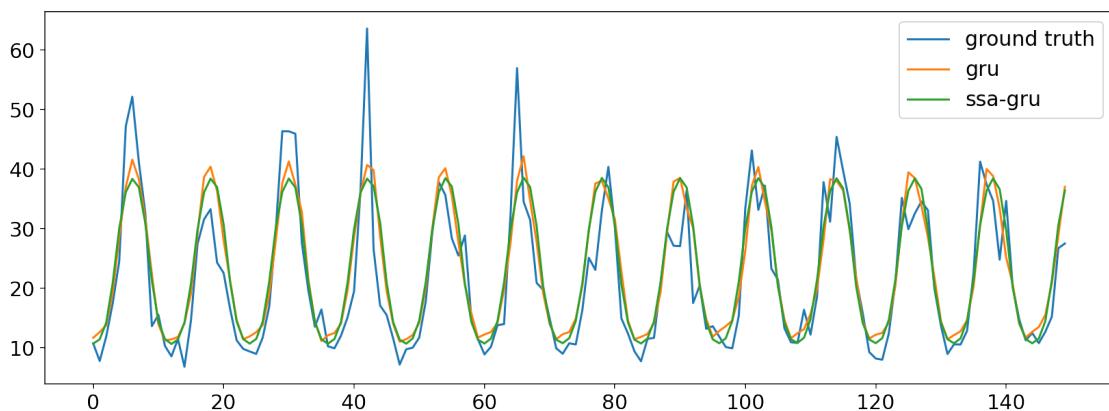


Рис. 5.36. Погода в Санкт-Петербурге. Прогноз для GRU и SSA-GRU.

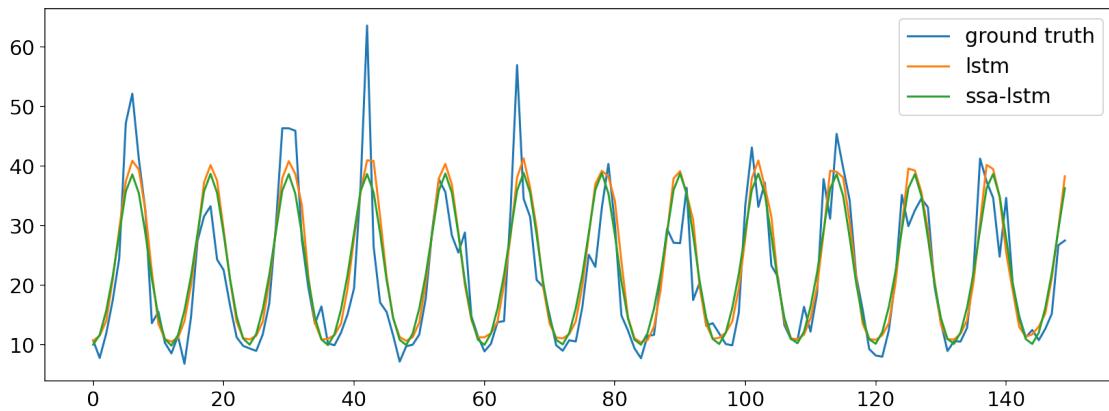


Рис. 5.37. Погода в Санкт-Петербурге. Прогноз для LSTM и SSA-LSTM.

## Проверка устойчивости

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

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

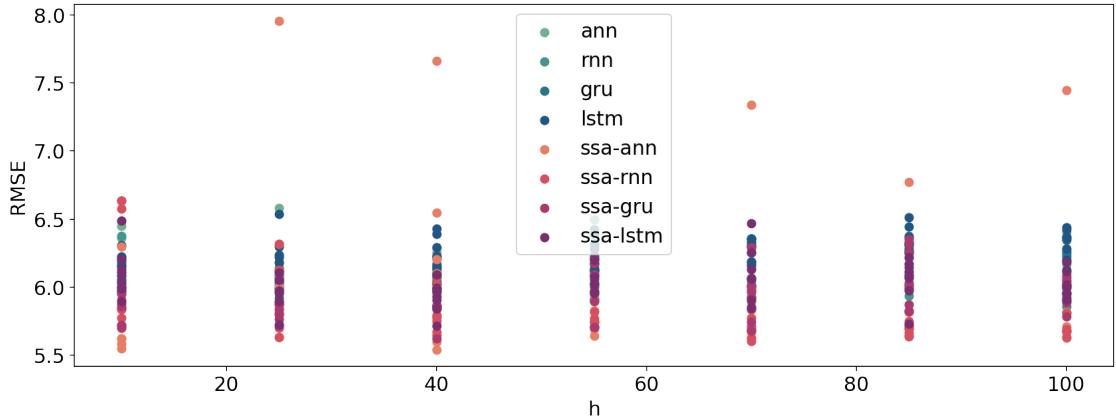


Рис. 5.38. Погода в Санкт-Петербурге. Проверка устойчивости.  $T = 12$ .

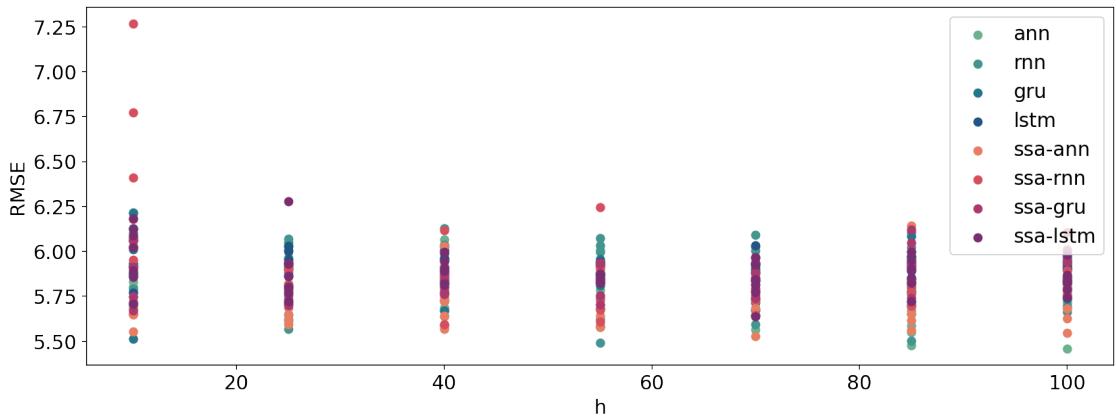


Рис. 5.39. Погода в Санкт-Петербурге. Проверка устойчивости.  $T = 84$ .

## 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. Что 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. Метод SSA может корректно выделить сигнал (получить оценку сигнала).
3. Длина 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 неискажал оценку сигнала.

По второму 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.

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.

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 ме-

тодов. 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.