

Підвищення ефективності якості прогнозування часових рядів інструментів фінансових ринків за рахунок агрегації різних довжин тайм-фреймів

Дипломна робота на здобуття кваліфікації
бакалавра

системного аналізу

Декрет Маргарита Володимирівна, КА-25

Керівник к.т.н., ст. викл. кафедри ММСА

Мурга Микола Олексійович

Вступ

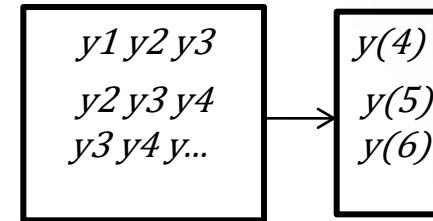
Класичний підхід прогнозування використовує аналіз часового ряду на одному тайм-фреймі. Однак очевидною є можливість збільшення ефективності якості прогнозування, завдяки агрегації оцінок одного і того ж прогнозованого значення на різних довжинах тайм-фреймів.

Новизна роботи полягає в розробці моделі за допомогою нестандартного підходу поділу даних та проведення порівняльного аналізу методів Back Propagation, авторегресії та методу головних компонент за допомогою критеріїв RMSE, MAE, MAPE та maxAPE.

Об'єкт дослідження:	часові ряди інструментів фінансових ринків
Предмет дослідження:	моделі, методи та технології прогнозування часових рядів інструментів фінансових ринків
Мета дослідження:	визначити оптимальні набори тайм-фреймів для прогнозування значень цін фінансових інструментів
Завдання:	❖ Реалізувати програмний додаток, що дозволяє проводити дослідження у відповідності до мети роботи. ❖ Визначити величини тайм-фреймів, на яких можливо виконувати прогнозування значення цін фінансових інструментів. ❖ Побудувати модель на оптимальних наборах тайм-фреймів для ефективного прогнозування за допомогою методів: гусениці-SSA, на основі моделі АРКС та нейронної мережі Back Propagation ❖ Провести порівняльний аналіз ефективності методів прогнозування значень цін інструментів фінансових ринків.

Побудова моделі та критерії оцінювання

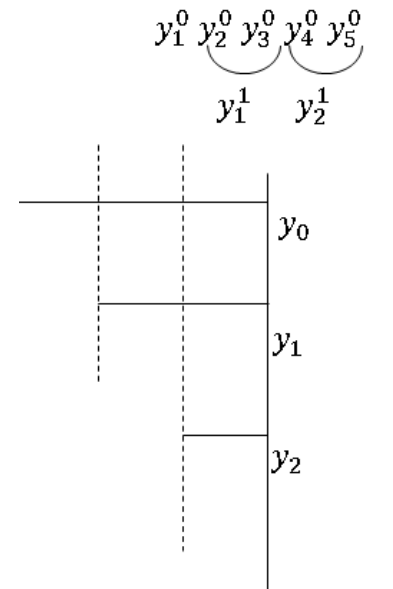
- $MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i}$



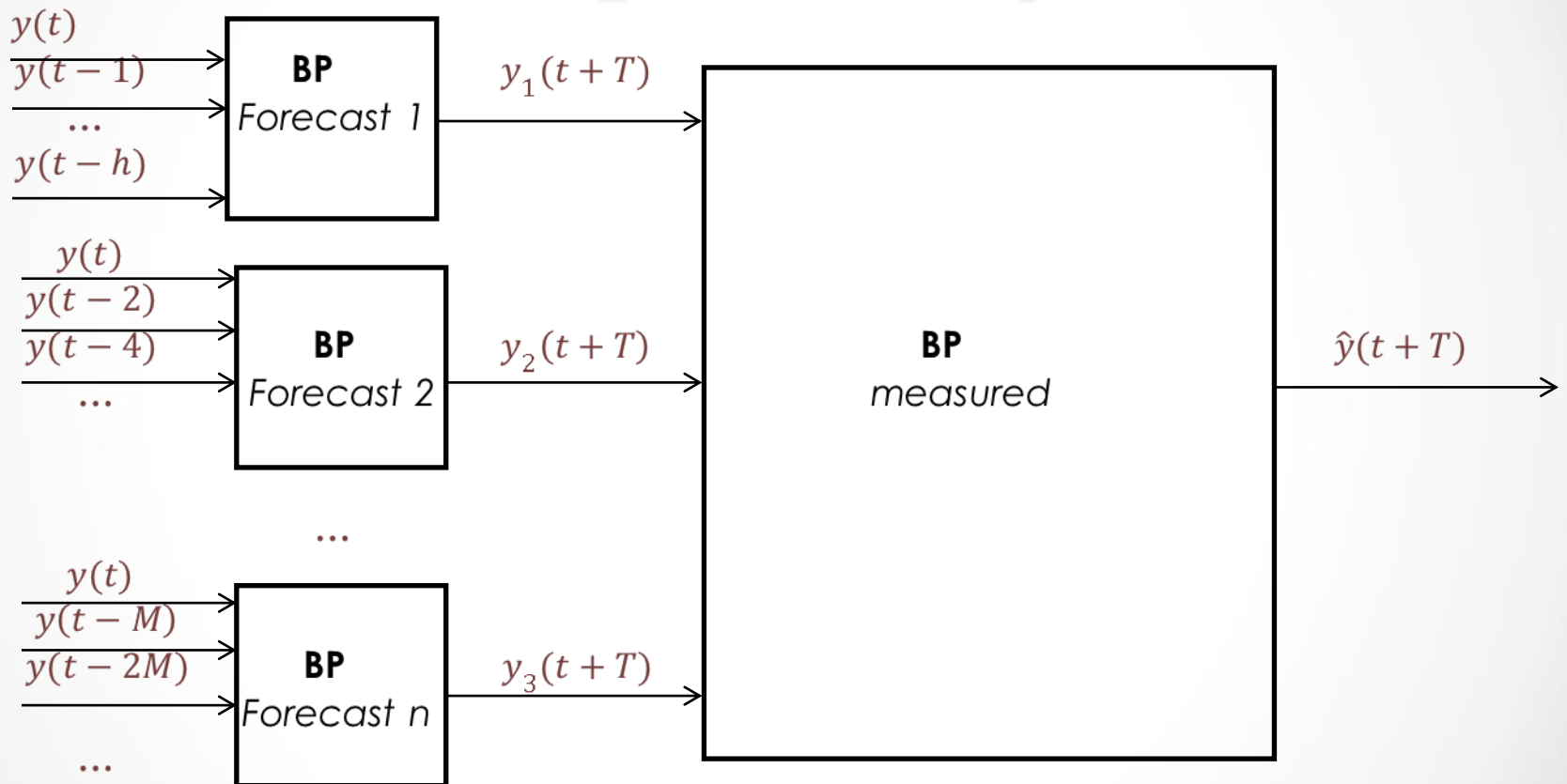
- $MaxAPE = \max \frac{|y_i - \hat{y}_i|}{y_i}$

- $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$

- $MAE = \max_i |y_i - \hat{y}_i|$



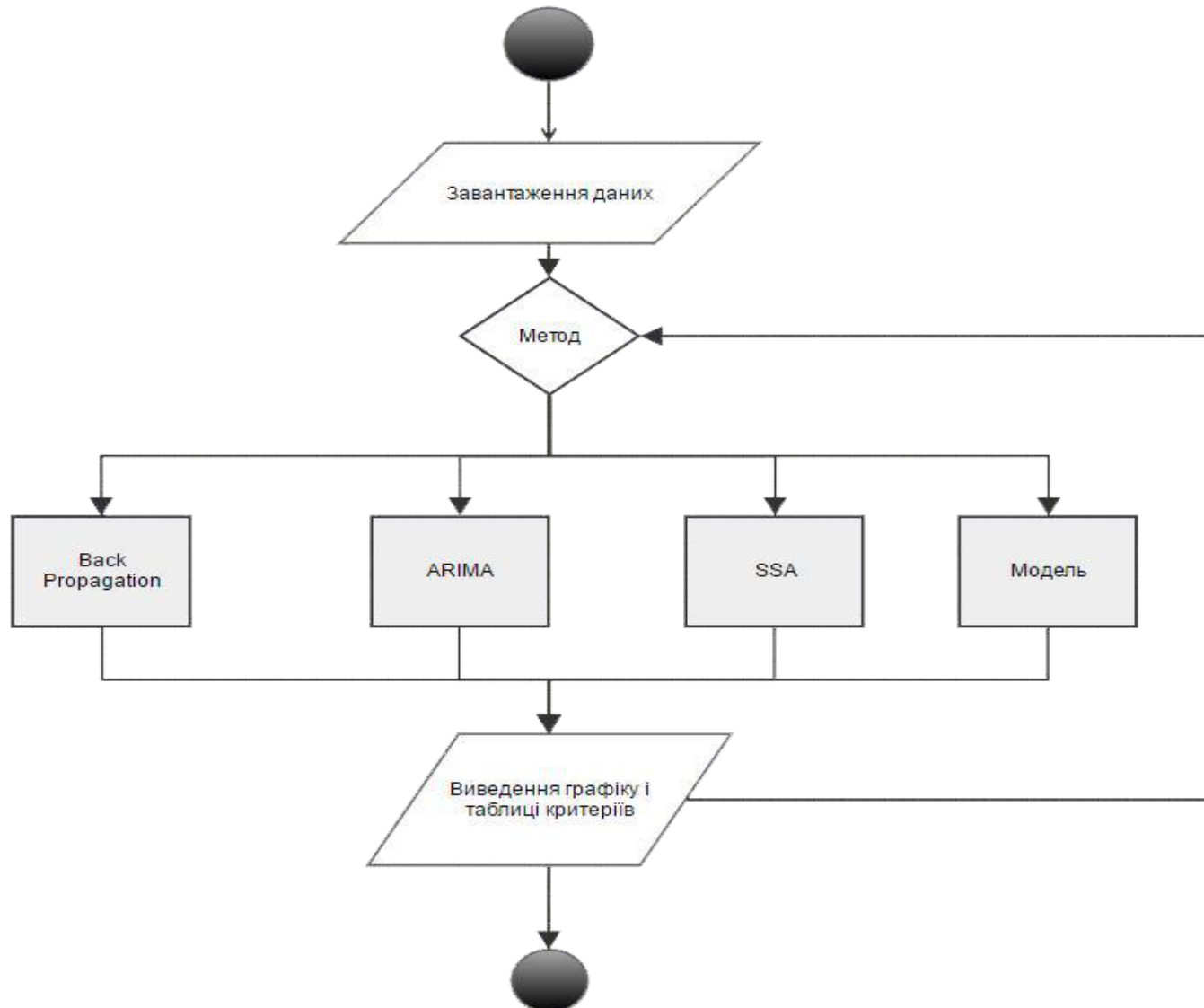
Модель прогнозування



$$y_1(t+T) = ax + c, \text{ а } y_2(t+T) = cx + d, \dots y_n(t+T) = ex + f, \text{ то}$$

$$\hat{y}(t+T) = ay_1(t+T) + by_2(t+T) + \dots + cy_n(t+T) + d.$$

Реалізація програмного продукту

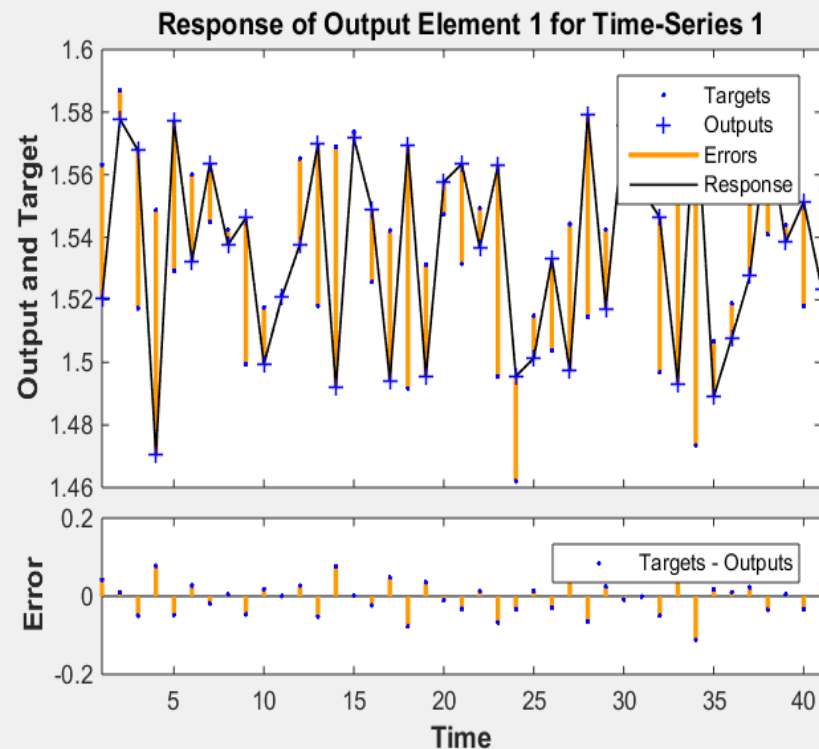
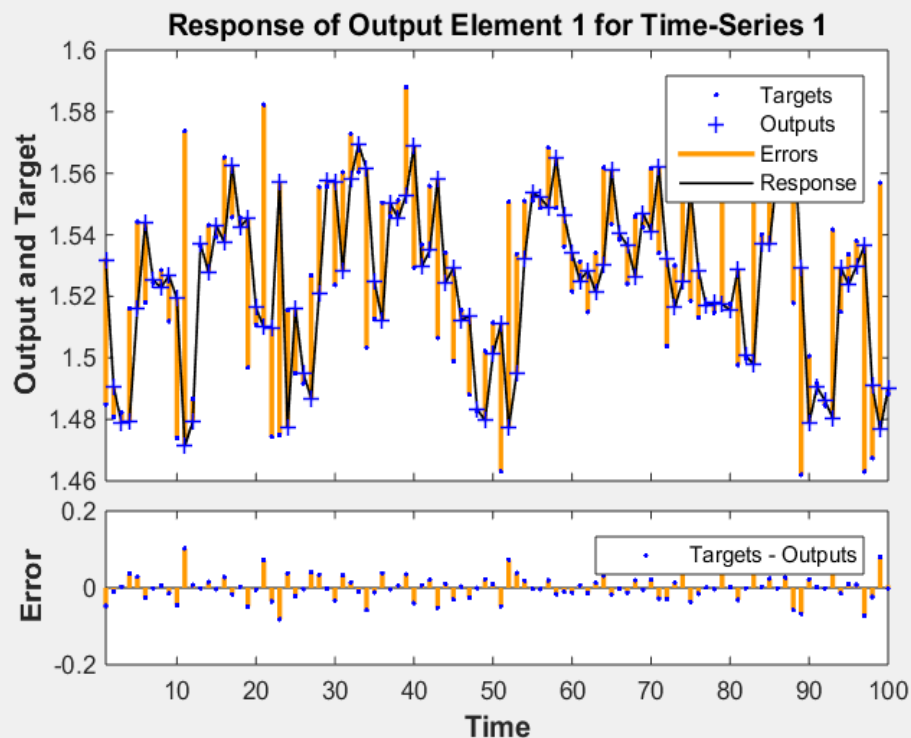
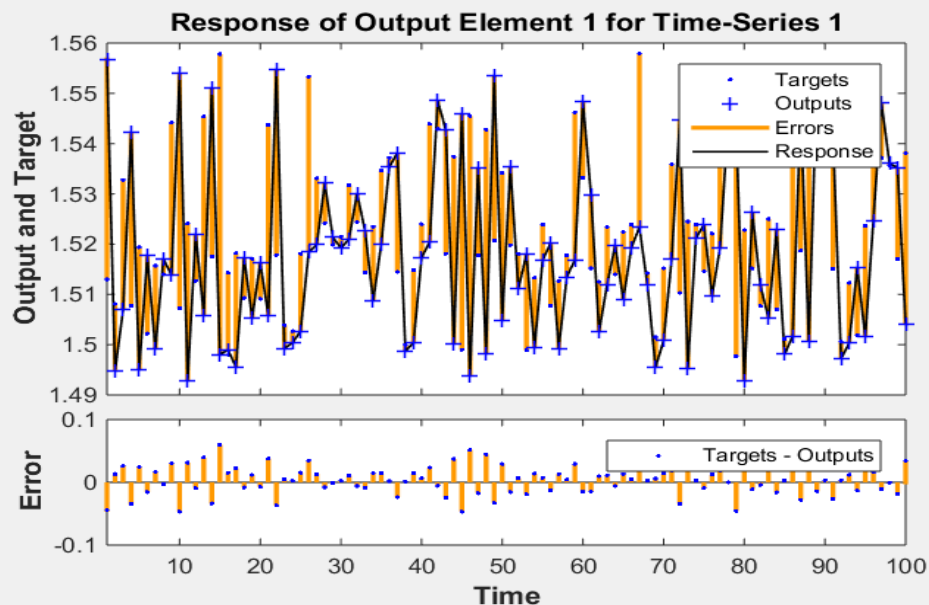


Реалізація за допомогою методу Back Propagation

В ході експерименту розглянуто пару фунт/долар за січень-жовтень 2015 рік. Виконані прогнози по годинам, дням та тижням по самим тайм-фреймам та по їх приростах окремо в двох таблицях відповідно:

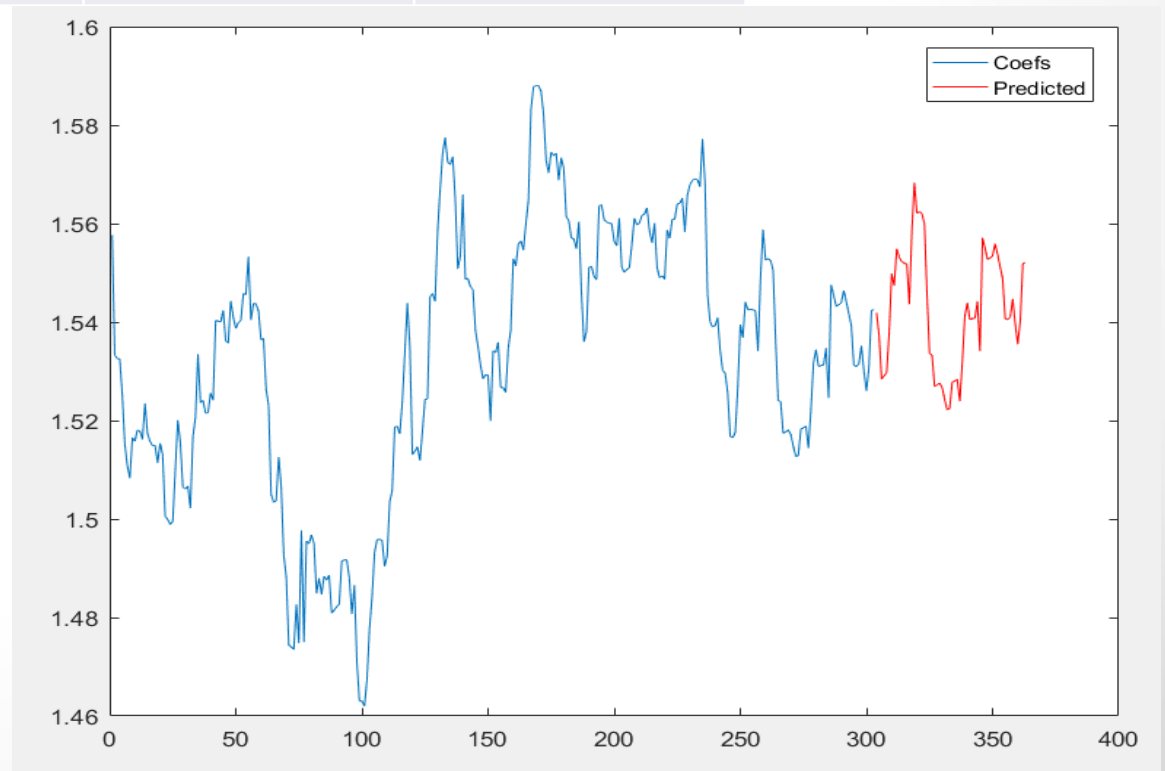
	Години	Дні	Тижні
MAPE	1.4802e-06	1.1415e-05	5.9807e-04
RMSE	0.9995	0.9984	0.9879
maxAPE	0.5511	1.4128	1.6285
MAE	0.8444	2.1865	2.5147

Графіки прогнозу по дням для вхідної вибірки та її приростів



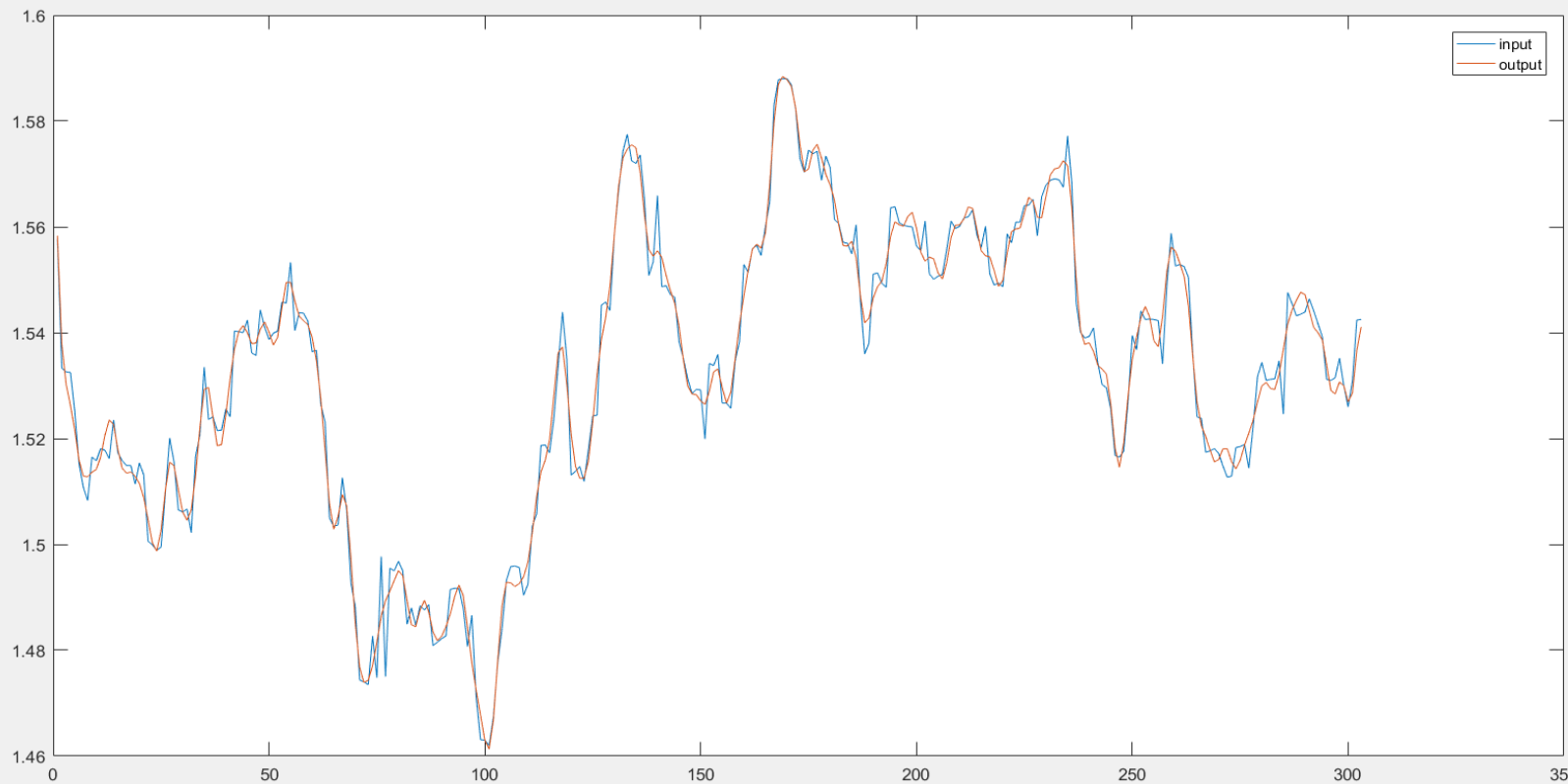
Прогнозування за допомогою авторегресії

	Години	Дні	Тижні
MAPE	7.6200e-04	0.0045	0.0348
RMSE	0.9989	0.9970	0.9773
maxAPE	2.1079	4.7616	5.0196
MAE	3.277197	6.9614	7.3386



Реалізація методом головних компонент SSA

	Години	Дні	Тижні
MAPE	6.1345e-05	5.6615e-04	0.0094
RMSE	0.9990	0.9967	0.9757
maxAPE	0.3082	0.9742	1.4707
MAE	0.4703	1.4369	2.2720

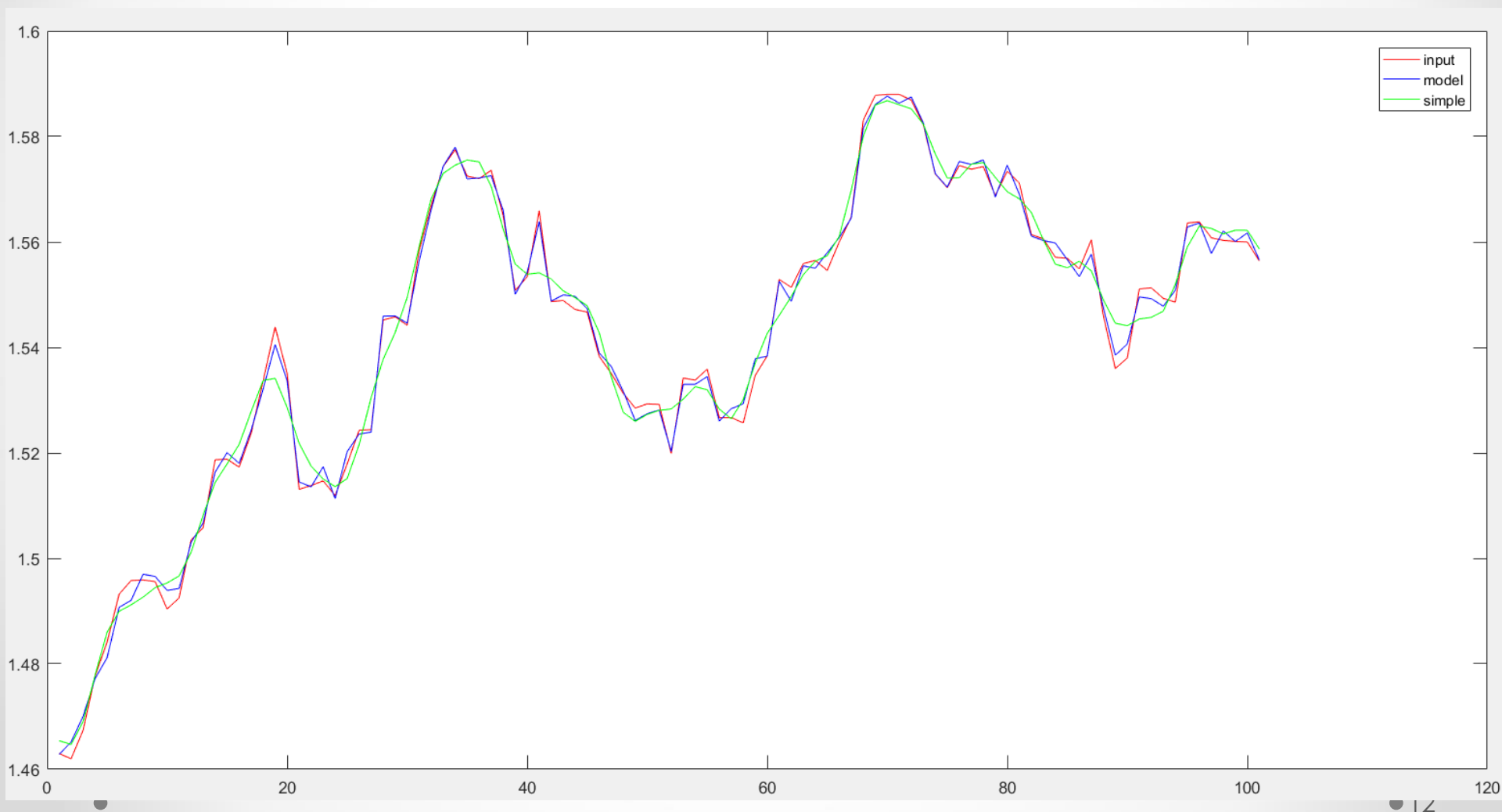


Дослідження ефективності моделі

HM	n	Модель	SSA	Звичайний прогноз	Модель
MAPE	1.1415e-05	0.0779			
RMSE	0.9984	0.6340	MAPE	6.3799e-04	3.3175e-04
maxAPE	1.4128	0.9320	RMSE	0.9966	0.9966
MAE	2.1865	1.4365	maxAPE	1.0553	0.5329
			MAE	1.5565	0.8091

Порівняння	Нейронна мережа	Метод головних компонент
MAPE	0.0779	3.3175e-04
RMSE	0.6340	0.9966
maxAPE	0.9320	0.5329
MAE	1.4365	0.8091

Графік порівняння звичайного методу і моделі



Апробація результатів роботи

...

Результати роботи апробовано на 18 міжнародній
конференції SALT-2016

Висновки

- Робота присвячена актуальній тематиці застосування сучасних інтелектуальних технологій для прогнозування значень цін фінансових інструментів. В ході її виконання були отримані наступні науково-практичні результати:
 - Для прогнозу за допомогою алгоритму Back Propagation значення критеріїв MAPE та MAE зростає зі збільшенням довжини тайм-фрейму в межах одного порядку від 0,0011 до 0,0872, що можна пояснити більшими амплітудами коливань на більших тайм-фреймах, а максимальне відхилення покращується з ущільненням інтервалу. Оцінка RMSE ж навпаки збільшується від тижнів до годин. Це вказує на те, що тайм-фрейми, прогнозовані по годинам, ймовірніше потрапляють в очікувані значення.
 - Для побудованої моделі MAPE та RMSE мають гірші значення за рахунок величини вибірки, але відхилення $\max \text{APE}$ зменшилось з 1.4128 до 0.9320, що свідчить про можливість ефективного використання моделі для нестабільних значень.

- Порівнявши методи отримали, що нейронна мережа дає найточніший результат ($MAPE=1.1415e-05$), $MAPE$ для методу головних компонент становить $5.6615e-04$, а для авторегресії 0.0045 , решта критеріїв порядком не відрізняються, але $RMSE$ найкращий теж в нейронній мережі і складає 0.9984 , а значення $maxAPE$ і MAE менші у метода головних компонент (0.9742 і 1.4369 відповідно).
- Порівнюючи результати для створеної моделі нейронної мережі та методу головних компонент видно, що результат по всіх чотирьох критеріях має меншу похибку, $RMSE$ 0.8963 для нейронної мережі і 0.9966 , внаслідок того, що в методі гусениці можна оптимізувати кожен математичний крок послідовно, в той час як в нейронній мережі змінювалась лише вхідна вибірка, тому модель методу головних компонент адаптована під нашу ціль краще.
- Експериментальна модель, отримана за допомогою методу головних компонент показала практично точне повторення вхідної вибірки, така модель буде дуже доцільною на фінансвоих ринках, де потрібно передбачити різку зміну відношення валют.

Дякую за увагу!