ERASMUS UNIVERSITEIT ROTTERDAM

Erasmus School of Economics

Policy Economics

Predicting short-term exchange rates with a Neural Network and structural variables

Authors:
I.M. Ruige, 448542

Instructor:
dr. F. P. L. Ward
Second Assessor:
dr. L. C. G. Pozzi

1st May 2023

Abstract

Predicting short-term exchange rates is challenging due to complex, non-linear relationships between macroeconomic variables and exchange rates. Creating a model that beats the Random Walk is difficult (Meese and Rogoff, 1983). This study addresses this challenge using a Neural Network model incorporating structural macroeconomic variables for exchange rate prediction. The model was trained using Bayesian optimization and evaluated against a Random Walk benchmark. Results show that the Neural Network model outperforms the Random Walk model regarding prediction accuracy. Additionally, the SHapley Addititative exPlanations (SHAP) method is used to interpret the model's decisions.

This study contributes to the literature by demonstrating the potential of neural networks for short-term exchange rate prediction. Further research exploring alternative network architectures and variables could enhance predictive capabilities.

Contents

1 Introduction

In recent years, exchange rates have shown significant influence on global economies and investments, as shown by events such as the fall of the pound sterling due to an unreliable financial plan of the United Kingdom, which ultimately led to the prime minister's resignation. Such incidents highlight the profound interactions between exchange rates, politics, and societies. Exchange rates not only have the power to trigger political discord but also affect import and export decisions in less extreme situations. Folk wisdom dictates that a lower exchange rate promotes exports while a higher rate stimulates imports. This folk wisdom can be directly derived from the Purchasing Power Parity theory. Furthermore, exchange rates influence companies' investment strategies in different countries, as demonstrated by models created by Froot and Stein (1991) and Dewenter (1995). Empirical research, such as those conducted by Cushman (1985) and Froot and Stein (1991), substantiates the claim that investments in countries are partly dependent on exchange rates. Overall, the effects of exchange rates on policy and investment are well-researched and significant (Dash (2018)).

Since both investments and policy decisions are future-oriented, accurate predictions of future exchange rates can be highly beneficial. Predicting exchange rates has been a well-studied problem, but the exchange rate is notoriously difficult to forecast (Ince and Trafalis (2006)). Exchange rates are highly influenced by a plethora of external factors, including the economic, social, political, and psychological behaviors of individual traders (Dash (2018)). Achieving predictions with a higher degree of accuracy than a Random Walk (RW) poses a difficult challenge (Meese and Rogoff, 1983). Numerous papers in the literature attempt to predict exchange rates using linear models; however, the assumption of linear correlation with predictors in foreign exchange rates are rarely fulfilled and is often non-linear. Neural Networks have the ability to fit any functional form in datasets (Hornik et al., 1989), making them a promising approach for predicting foreign exchange rates. Many studies, such as those employing a Multi-Layer Perceptron Network (MLP) (Galeshchuk (2016)) or a hybrid neural network with trading indicators from Ni and Yin (2009), consider using neural networks as a promising model for predicting exchange rates.

Kilian and Taylor (2003) found strong predictability for periods of 2 or 3 years in the future; however, they did not find evidence for shorter horizons. This paper aims to explore whether short-term exchange rate predictions are possible. Unlike most previous studies, which train neural networks solely on time series data of exchange rates, this paper incorporates structural macroeconomic variables. Following the promising results of Ait-Sahalia et al. (2022), who demonstrated that Neural Networks (NN) models could predict stock price volatility 14% better than a random walk using extremely short timeframes, this paper will predict exchange rates through an NN model with both structural data of economies and time series data of exchange rates. The exchange rate pairs USD/CAD, USD/CHF, and CHF/CAD will be investigated.

This paper is organized as follows: Section 2 describes the extensive dataset for exchange rates, along with the fundamental and explanatory variables employed in the analysis. Section 3 discusses the methodology, covering the hypothesis of no Triangular Arbitrage, the Random Walk (RW) model, the setup and notation of the Neural Network models, and the benchmarks and criteria for assessing the predictive quality of the models, based on Mean Squared Error(MSE), Mean Absolute Error, and Symmetric Absolute Mean Percentage Error. Further subsections in this Methodology cover the activation function based on the REctified Linear Unit. The optimization of the hyperparameters is based on Bayesian Optimization. Additionally, the overfitting issue is addressed through the Dropout method. Also, the necessary Batch Normalization of the data for the Neural Network is explained. Finally, the Forecast Interpretation based on the SHapley Addititative exPlanations method is explained. In Section 4, the paper results are presented, with the Neural Network based on the MSE loss function being the best-performing model. Additionally to the model prediction, the No Triangular Arbitrage is rejected, and the interpretations of the important variables are given. Finally, in Section 5, the findings are summarized, namely, the most explanatory variables are given, and the rejection of the No Triangular Arbitrage hypotheses is stated. We finally find that the optimized Neural Network outperforms the RW model on all used metrics.

2 Data

The dataset encompasses a vast range of variables from January 2012 to January 2022, classified into seven categories: Commodities, Exchange Rates, Metals, Macro-Economic variables, Interest Rates, and Swap Rates. A large dataset is necessary to train a neural network; thus, daily granularity is used where possible. The only exceptions are the macroeconomic variables, as these are published either on a monthly or quarterly basis. When the less granular information is published, it will be set as the last publicized value for the higher granularity until the new publication.

First and foremost, the exchange rates considered are the United States Dollar (USD), Swiss Franc (CHF), and Canadian Dollar (CAD), along with all their possible combinations. Excluding these combinations could create a possibility for triangular arbitrage.

Secondly, commodities such as metals, crude oil, gas, rice, and grain, among others, are included. These commodities serve as a proxy for the inflation rate, which is not published daily. Most countries predict the inflation rate based on a basket of goods; these goods will be used as a daily proxy alongside the inflation rate. Chen et al. (2014) demonstrates that these commodities possess explanatory power for inflation rates.

Thirdly, interest rates are also considered in the model. Interest rates play a crucial role in determining exchange rates, as they influence investment flows between countries and affect the relative attractiveness of financial assets (Dornbusch (1976), Frankel (1979)). Hence, the interest rates are added.

Also, the swap rates represent a fixed rate paid or received in exchange for a floating rate in an interest rate swap and give important information about the perceived risk with financial transactions, which might, in turn, affect exchange rates.

Lastly, a set of macroeconomic variables, such as trade deficit, money supply, trade deficit, and market returns, among others, will be used to predict the exchange rate. These macroeconomic structural variables have a proven track record of influencing exchange rates (Meese and Rogoff, 1983). Therefore, incorporating them into the model may be beneficial.

3 Methodology

In this Section, the methodology is explained. First, the No Triangular Arbitrage Hypothesis is presented, which will be tested. Then, the models used are specified. Various Machine Learning (ML) approaches, such as Random Forest or Support Vector Machines, can be explored to predict exchange rates. However, the literature suggests that Deep Learning approaches, such as Neural Networks, perform best for exchange rate predictions (Wang and Chen, 2021 & Galeshchuk and Mukherjee, 2017). In Section 3.2, the employed models are specified, which include a Random Walk model for comparison and the Neural Network models. Subsequently, the Activation Function is explained, followed by the Loss Functions and their optimization. The Dropout method is then discussed to prevent overfitting. Next, an explanation of the training and test data splits is provided. The Hyperparameter tuning in combination with Network Architecture is then considered. A description of the performance measures follows this. Finally, the SHapley Additive exPlanations (SHAP) method is described for interpreting the variable importance in the NN.

3.1 No Triangular Arbitrage Hypothesis

Triangular arbitrage is a risk-free trading strategy that exploits discrepancies in foreign exchange rate pricing to generate profit. Triangular arbitrage operates by converting one currency to another, then to a third currency, and finally back to the original currency, resulting in a profit. According to Fama (1970), there should be no triangular arbitrage opportunities in an efficient market. This is expressed by:

$$E_{AC} = E_{AB} \times E_{BC} \tag{1}$$

Here, E_{AC} represents the exchange rate from currency A to C, E_{AB} from currency A to B, and E_{BC} from currency B to C. Fenn et al. (2009) demonstrates that small and short-lived arbitrage opportunities sometimes exist in exchange rates.

This paper will empirically test the hypothesis of no triangular arbitrage in the foreign exchange rates used in the dataset. If the hypothesis is confirmed, the model will incorporate a penalization for triangular arbitrage opportunities.

3.2 Model Specifications

Let $R_{1:T} = (R_1, R_2, ..., R_T)$ be the $(3 \times T)$ exchange rate matrix, where $R_t = (r_{1,t}, r_{2,t}, r_{3,t})'$ is the vector of the three exchange rates used. Additionally, let $\mathbf{f}_{1:T} = (\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_T)$ represent the predictor data matrix $(K \times T)$, where the vector $\mathbf{f}_t = (f_{1,t}, f_{2,t}, ..., f_{k,t})'$ contains the predictor information at time t.

3.2.1 Random Walk

The Random Walk (RW) model is used as a baseline model. The RW model is specified as follows:

$$R_t = \mu + R_{t-1} + \varepsilon_t \tag{2}$$

In this equation, R_t denotes the exchange price at time t, and μ is the intercept. R_{t-1} represents the exchange rate at time $t-1$, and ε refers to the error term. The error term has an expected value of $E(\varepsilon_t) = 0$. Its variance, $Var(\varepsilon_t) = \sigma_t^2$, is determined by the standard deviation in the training set. Additionally, ε_t is assumed to be normally distributed.

This model will be used as a benchmark. Meese and Rogoff (1983) demonstrate that the Random Walk model is a crucial model to outperform if the goal is to have a model with additional explanatory power instead of random guessing.

3.2.2 Artificial Neural Network

I use an Artificial Neural Network (ANN) to model potential nonlinear specifications within the exchange rate parameters. ANN models consist of a network architecture of nodes that can pass signals through (multiple) hidden layers. This results in one of the more powerful machine learning techniques due to the ability of the NN to approximately fit any functional form in any dataset (Hornik et al., 1989).

The ANN models provide an approximation of the exchange rate prediction function. Through a layered structure, as shown in Figure 1. In this Section, most of the notation is based on Athey and Imbens (2019).

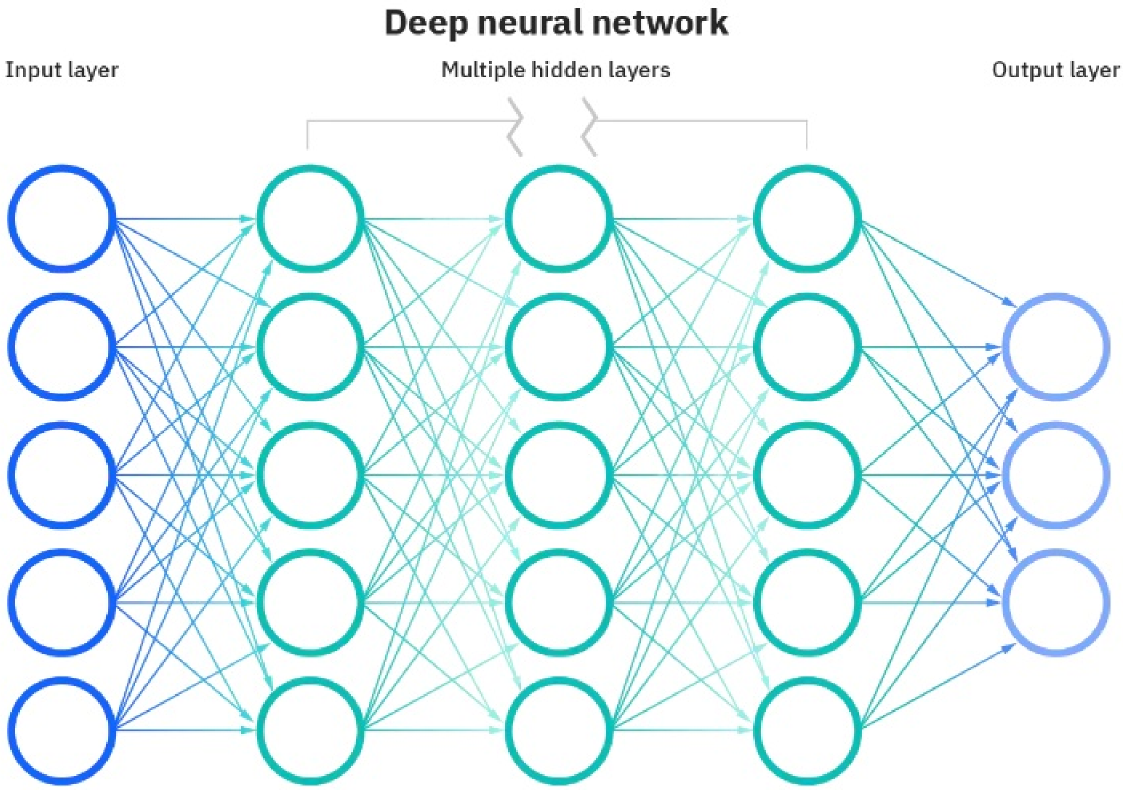

Figure 1: An example of a Deep Neural Network with multiple layers

The NN model consists of an input layer, multiple hidden layers, and an output layer. The network layers each consist of nodes denoted by $n_m^{(i)}$, where m depicts the node in the layer i. Each layer has a set of M_i nodes, which can differ per layer. For example, the input layer consists of K nodes and, since it is the first layer, the layer results in $\{n_1^{(1)}, n_2^{(1)}, \ldots, n_K^{(1)}\}$. Each node receives the input of the prediction data of time t: $\{f_{1,t}, f_{2,t}, \ldots, f_{K,t}\}$ this results in an activation signal, a linear combination of the input:

$$Z_{k,t}^{(1)} = \sum_{j=1}^{K} \beta_{k,j}^{(1)} f_{j,t}, \text{ for } k = 1, \ldots, K_1 \tag{3}$$

In this instance, $\beta_{k,j}^{(1)}$ is the scaling coefficient of input parameter j to the activation node k. The $Z_{k,t}^{(1)}$ represents the node value of each node in the first layer. These nodes' values go through an activation function and then become the input for the next layer. In general form, this is represented by:

$$Z_{k,t}^{(i+1)} = \sum_{j=1}^{K_i} \beta_{kj}^{(i+1)} h(Z_{tj}^{(i)}), \text{ for } k = 1, \ldots, K_{i+1} \tag{4}$$

Here $\beta_{kj}^{(i+1)}$ is the scaling coefficient for the next connected layer node, and $h(\cdot)$ equals the activation function, which will be explained later in Section 3.3. The final layer and prediction for the exchange rate is represented by:

$$R_t = \sum_{j=1}^{K_{l-1}} \beta_k^{(l)} h(Z_{tk}^{(l-1)}) + \varepsilon_t \tag{5}$$

Where R_t is the predicted exchange rate, l represents the number of the last layer, and ε_t represents the error term.

3.3 Activation Function

As an activation Function the *Rectified Linear Unit* (ReLU) is used and is given by (Nair and Hinton, 2010):

$$h_{ReLU}(Z) = max(0, Z) \tag{6}$$

In this equation, Z represents the node value. The ReLU function is one of the most popular used activation functions (Nwankpa et al., 2018). Hara et al. (2015) states that the ReLU's popularity is due to the fast convergence because of the nonzero derivative of the ReLU function. This results in a fast, trainable, and relatively precise NN compared to other activation functions and hence is used in most in the state-of-the-art models (Nwankpa et al., 2018).

3.4 Batch Normalization

Batch normalization is a technique designed to improve the training of Deep Neural Networks by rescaling the values of the activation functions during the training of the NN. This leads to better generalization and performance (Ioffe and Szegedy, 2015). Unscaled data might lead to a faster training process and difficulty in choosing hyperparameters. Section 3.8 will explore hyperparameters.

During training, batch normalization works by normalizing the input to a layer by adjusting its mean to center around 0 and variance for each mini-batch, a small subset of data, resulting in a normally distributed signal. Effectively the process is at each hidden layer:

$$\mu = \frac{1}{b}\sum_i Z^{(i)} \tag{7}$$

μ is the batch mean, b is the size of the batch, $Z^{(i)}$ are all the nodes in layer i.

$$\sigma^2 = \frac{1}{b}\sum_i (Z^{(i} - \mu)^2 \tag{8}$$

σ^2 is the batch variance.

$$Z^{(i)}_{norm} = \frac{Z^{(i)} - \mu}{\sqrt{\sigma^2 - \varepsilon}} \tag{9}$$

Where ε is a constant used for numerical stability.

$$Z_{scaled} = \gamma * Z^{(i)}_{norm} + \beta \tag{10}$$

Where β is a constant for bias compensation, and γ is used for the scalability of the variance. Both variables are optimized for each node in the layer during the training process.

The effective smoothing of the training process creates a better model according to Santurkar et al. (2018); hence, it is almost standard to implement this process in a Neural Network.

3.5 Loss Functions

For NN, the loss functions play a pivotal role; a loss function quantifies the difference between the predicted and the true values. The training of a NN effectively minimizes this loss function, thereby improving the model's predictive accuracy. Various loss functions exist, each suitable for different purposes. In this Section, two loss functions are proposed, a loss function based on the Mean Squared Error (MSE) and a personalized loss function with penalization for triangular arbitrage.

3.5.1 Mean Squared Error Loss

Since the prediction of exchange rates is a regression problem, the Mean Squared Error (MSE) is a natural starting point. The MSE is given by:

$$MSE = \frac{1}{N}\sum_{t=1}^{N}(R_t - \hat{R}_t)^2 \tag{11}$$

Where N is the number of observations, R_t is the true exchange rate at time t, and $\hat{R}_t$ is the predicted exchange rate at time t.

3.5.2 No Triangular Arbitrage Loss

Alongside the exchange rate being a regression problem. In this paper, the hypothesis of No Triangular Arbitrage is a factor. Combining the hypothesis in the loss function results in the following:

$$loss = \frac{1}{N}\sum_{t=1}^{N}(R_t - \hat{R}_t)^2 - E_{AC} + E_{AB} \times E_{BC} \tag{12}$$

3.5.3 Loss Function Optimization

The MSE and the loss function based on Equation (12) are tried to optimize the NN model. The optimization of the loss function is done through AdaGrad (Adaptive Gradient Algorithm), an off-the-shelf loss function optimizer (Duchi et al., 2011). It is a popular loss function due to its fast convergence and robustness for different data scales. The latter property is particularly attractive in the current setting since the data has a wide variety of scales.

3.6 Preventing Overfitting

This paper uses 122 variables on 7830 exchange rate observations, which may result in overfitting of the model. Overfitting is the issue where the model specializes in fitting the training data too closely, leading to poor generalization and, thus, poor performance for unseen data prediction. As explained in Section 3.9, the only measures for prediction are out-of-sample; hence, it is important to address this issue. The prevention of overfitting is done through regularization techniques. Liang and Liu (2015) state that the most popular and effective method to do this is the Dropout method. In this subsection, I will discuss the process of the Dropout method.

The Dropout is a regularization technique introduced by Srivastava et al. (2014), which involves randomly "dropping out" or deactivating a subset of neurons during the training of the model. The process is done by initiating a dropout rate p. Then, each node in the network is assigned a random value between 0 and 1. Each node with a value below p is deactivated by setting the activation function of the node to 0. The result of this process is shown in Figure 2.

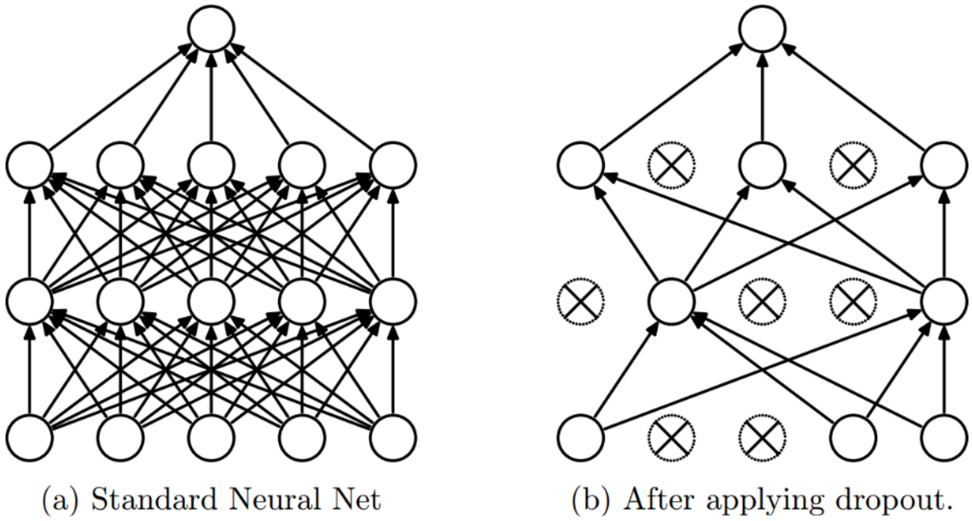

Figure 2: The process of the Dropout method shown in an example of a Neural Net. The figure is from Srivastava et al. (2014).

According to Srivastava et al. (2014), this process introduces randomness and noise, forcing the model to create multiple neuron combinations, thus creating a more robust model.

3.7 Data splitting

During the Neural Network training process, the data is split into two sets: the training set and the test set. The training set contains the first 80% of the observations, and the test set includes the last 20%. This is done to maintain the time series structure of the data. Preserving the temporal order of the time series helps avoid look-ahead bias in the model training. It resembles a "real-world" problem more closely, where future values are predicted with only historical data, not training on possible future observations. Aside from more closely resembling "real-world" problems, Falessi et al. (2018) suggest that preserving the time series increases model performance.

3.8 Hyperparameter Optimization

Hyperparameter optimization is essential for the behavior of the NN since it determines the shape and size of the NN. Hyperparameters are parameters set before a model's training begins, and they control the behavior of the training algorithm. In this application, the tuned hyperparameters include the dropout rate, the number of nodes in each layer, and the learning rate. The fixed hyperparameter is the number of layers.

3.8.1 Tuned Hyperparameters

The hyperparameters are tuned using Bayesian Hyperparameter Optimization (Snoek et al., 2012). A grid of possible outcomes is created for each hyperparameter to tune the hyperparameters. The used grid for each hyperparameter is:

- The number of nodes in each layer [3,122], with the size of the grid chosen based on the rule of thumb that the optimal number of nodes falls between the size of the input layer (currently 122) and the size of the output layer (3) (Zoph and Le, 2016). With a step size of 7, this creates 17 possible outcomes.
- The dropout rate p [0,1) is a natural choice since it represents the entire probability space of the dropout. With a step size of 0.01, this creates 100 possibilities. The Dropout is further explained in Section 3.6.
- The learning rate has four options [0.01,0.001,0.0001,0]. The learning rate determines the size of the update of the weights of each node in optimizing the gradient descent of the loss function of the Neural Network (Smith, 2017).

The literature presents multiple optimization methods, ranging from straightforward brute force grid searches to more sophisticated models such as Bayesian Hyperparameter Optimization (Snoek et al., 2012). Wu et al. (2019) and Shahriari et al. (2015) show that the latter is one of the better ways to find the correct value for hyperparameters for a given problem. Instead of using brute force, this method employs a probabilistic model based on a Gaussian Process, resulting in a faster and more efficient training process for the model. Consequently, it is one of the most widely used off-the-shelf methods.

3.8.2 Set Hyperparameters

A Neural Network can grow automatically in a more advanced setting, requiring a much larger dataset than the current research (Krizhevsky et al., 2017). Real et al. (2017) propose a method for automating Neural Network architecture. However, a more Brute-Force method is employed due to data limitations in the training set for the current setting. Layers are added until the NN no longer shows improvement.

3.9 Forecasting Performance Measures

Three popular error statistics for regression ML models are used to assess the quality of the predictions and to determine whether the NN model outperforms the RW model. These include the Mean Absolute Error (MAE), the Mean Squared Error (MSE), and the Symmetric Mean Absolute Percentage Error (SMAPE). Out-of-sample predictions are used for all of the performance measures since in-sample predictions are not very informative due to the adaptive nature of NN methods. The out-of-sample data consists of the last 20% of the observations to preserve the time series structure.

3.9.1 Mean Absolute Error

MAE is a popular metric for measuring prediction accuracy. It calculates the average absolute difference between the predicted and actual values. A lower MAE indicates better model prediction. The MAE formula is given by:

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|R_t - \hat{R}_t| \tag{13}$$

Where N is the number of observations, R_t is the true exchange rate at time t, and $\hat{R}_t$ is the predicted exchange rate at time t.

3.9.2 Mean Squared Error

MSE is another widely used metric for evaluating forecast accuracy. MSE calculates the average of the squared differences between the predicted and actual values. A lower MSE indicates a better forecast. The MSE formula is given by:

$$MSE = \frac{1}{N}\sum_{t=1}^{N}(R_t - \hat{R}_t)^2 \tag{14}$$

Where N is the number of observations. R_t is the true exchange rate at time t, and $\hat{R}_t$ is the predicted exchange rate at time t.

3.9.3 Symmetric Mean Absolute Percentage Error

Symmetric Mean Absolute Percentage Error (SMAPE) is a metric that measures forecast accuracy as a percentage of the actual values. It calculates the average of the absolute differences between the predicted and actual values, divided by the average of their sum. The result is multiplied by 100 to express it as a percentage. SMAPE ranges from 0 to 200, with 0 indicating perfect predictions and higher values indicating less accurate predictions. In the context of exchange rate forecasting, a low SMAPE means that the model's predicted exchange rates are relatively accurate compared to the actual exchange rates in percentage terms.

The Symmetric Mean Absolute Percentage Error (SMAPE) is given by:

$$SMAPE = \frac{1}{N}\sum_{t=1}^{N}\frac{|R_t - \hat{R}_t|}{(R_t + \hat{R}_t)/2} * 100 \tag{15}$$

Where R_t and $\hat{R}_t$ are, respectively, the actual and the predicted exchange rate.

3.10 Forecast Interpretation

In this subsection, I present the interpretation technique to better understand the influence of each variable on the forecast. The interpretability of ML, as explained by Doshi-Velez and Kim (2017), is the ability to explain the model in understandable terms for a human. A NN is much harder to interpret than a linear regression. The attractive feature of the NN is its flexibility. To benefit from the flexibility and provide an interpretation of the NN, two questions must be asked:

- Which predictor variables contribute to the forecast?
- What are the partial relations between the predictor variables and the forecasts?

To answer these questions, I use the method introduced by Lundberg and Lee (2017) called SHapley Addititative exPlanations (SHAP). These SHAP values will construct the partial dependence and the variable importance measure.

The SHAP values originate in the Shapley Values from Shapley (1953). Shapley values quantify the contribution of each player in a payout function. In the ML setting, each predictor is a "player," and the contribution to the model is the "payout ."According to Lundberg and Lee (2017), the SHAP method is the only method proven to be efficient in the sense that the sum contribution of all the predictors adds up to the forecast itself. Aside from its efficiency, the SHAP method is also proven to be symmetric, implying that if the Shapley value of two predictors is the same, the marginal contribution to the forecast of the predictor is the same. These two properties make SHAP a very attractive method for interpreting NN.

Let Φ be the matrix of matrices: $\Phi_{r,k,t} = (\phi_{1,k,t}, \phi_{2,k,t}, \phi_{3,k,t})'$. Then the exchange rate forecast $\hat{R}_t$ at time t in the test sample has the SHAP explanation model given by:

$$\hat{R}_t = \Phi_{r,0} + \sum_{k=1}^{K} \Phi_{r,k,t} \tag{16}$$

In the model, $\Phi_{r,k,t}$ is the attribution of feature k to the exchange rate at time t. $\Phi_{r,0}$ denotes the expected forecast overall test sample return predictions hence a (1x3) vector in the current setup implying that $\Phi_{r,0} = E(\hat{R}_t)$. Equation 16 can be rewritten as:

$$\sum_{k=1}^{K} \Phi_{r,k,t} = \hat{R}_t - E(\hat{R}_t) \tag{17}$$

By using the additive model from Equation (17), an explanation of the importance of each predictor to the forecast can be given. I illustrate this process with an example of four predictor variables in Figure 3.

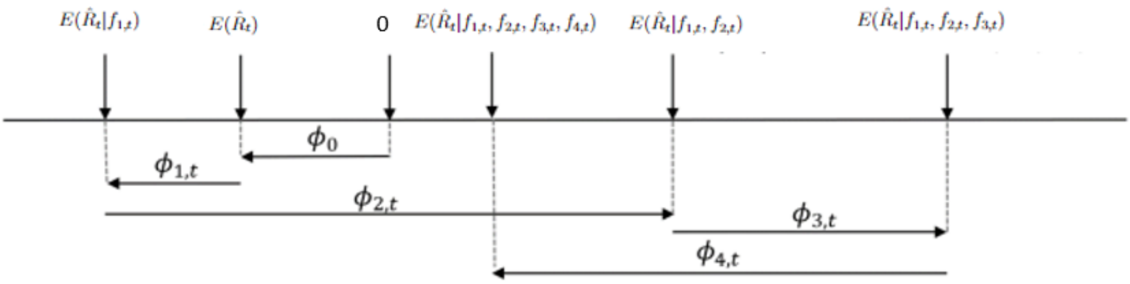

Figure 3: Illustration of the SHAP process: This figure displays the individual contribution of the predictor variable denoted by $f_{k,t}$ to the final forecast $\hat{R}_t$. In this example, K = 4 is used as the number of predictor variables. $E(\hat{R}_t)$ is the so-called base value.

Let F be the set that contains all the predictor variables. A Shapley value $\Phi_{r,t,k}$ is estimated by averaging the forecast, including and excluding the k^{th} predictor variable, in all subsets S of F. For each subset S, the model is trained, both using the set with predictor variable $S \bigcup \{k\}$ and the model without the predictor variable S. Then the difference between the two is the Shapley value for Φ_r, k, t resulting in:

$$\Phi_{r,k,t} = \sum_{S \subseteq F \setminus \{k\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [(\hat{E}_t|S \cup \{k\}) - (\hat{E}_t|S)] \tag{18}$$

Estimating models on all subsets S would require estimating $2^{|F|}$ models, which is practically infeasible. Lundberg and Lee (2017) propose an efficient model to approach this: the Deep SHAP model[1]. They restrict training certain model types and, based on sampling, have found a way to develop faster model-specific approximations.

Shapley values are estimated for all predictor variables. In this research, the number of samples is limited to 100 training examples for the model due to the heavy calculation required to compute the values. In the programming example of Lundberg and Lee (2017), a sample of 100 is also used and suggested to be satisfactory for explanatory purposes. The average Shapley values from this analysis will help to explain the first sub-question of this Section. A scatterplot is also created from these 100 observations, which helps answer the partial dependence question.

[1]https://github.com/slundberg/shap

4 Results

In this Section, the results of the research are discussed and presented. First, the No Triangular Arbitrage is addressed. Then the Prediction Performance is discussed. After which, the interpretation of the NN utilizing the SHAP method is presented.

4.1 No Triangular Arbitrage

First, the results for the hypothesis in Equation (1) are presented in Table 1. We see that depending on the number of decimals, the triangular arbitrage hypothesis is violated more often than not for more than three decimals. For three decimals, it is violated about half of the time. Hence, I reject the hypothesis of no triangular arbitrage and the loss function based on the MSE Equation (11) will yield better results than the loss function based on Equation (12). The findings are in line with Aiba et al. (2002) who find triangular arbitrage opportunities in foreign exchange markets. Also, Fenn et al. (2009) find these triangular arbitrage opportunities, but they state that these are short-lived and of small magnitude.

Cutoff Decimal	Number of Violations	Percentage of Violations
1×10^{-3}	1222	46.82%
1×10^{-4}	2434	93.26%
1×10^{-5}	2594	99.39%
1×10^{-6}	2609	99.96%

Table 1: Cutoff on each decimal on 2610 of Observations of the No Triangular Arbitrage hypothesis

4.2 Prediction Performance

In this Section, we compare the results of the three used models. The metrics in all of the models are based on 522 out-of-sample predictions for each exchange rate. Making the average prediction of 1566 observations out-of-sample.

4.2.1 Random Walk Results

In this Subsection, we present the performance of the Random Walk model (2) for predicting exchange rates. The results are assessed using three metrics: Mean Absolute Error (MAE), Mean Squared Error (MSE), and Symmetric Mean Absolute Percentage Error (SMAPE). The performance of the models is reported in Table 2 for three currency pairs: USD/CHF, USD/CAD, and CAD/CHF, as well as the average for the three exchange rates.

	USD CHF	USD CAD	CAD CHF	Average
MAE	0.0335	0.0783	0.0589	0.0576
MSE	0.0017	0.0098	0.005	0.0057
SMAPE	3.1120	10.2430	8.2491	7.2014

Table 2: Comparison of prediction of exchange rates based on the three metrics of MAE, MSE, and SMAPE, predicted through the Random Walk model (2)

In Table 2, we see that for the average exchange rate, we mispredict by 0.0576 due to MAE being this. The MSE is the smallest for the USD/CHF pair and the largest for the USD/CAD pair, indicating the largest misprediction for the USD/CAD pair. We also see this result for the SMAPE metric. On average, we are 10.24% off on the prediction for the USD/CAD pair. The average forecast is off by about 7.2%.

4.2.2 Neural Network Results

In this Section, the results of both Neural Network model predictions are shown.

	USD CHF	USD CAD	CAD CHF	Average
MAE	0.20198	0.43747	0.09185	0.07662
MSE	0.08787	0.23637	0.01322	0.01402
SMAPE	32.31866	12.08401	11.38957	18.59741

Table 3: Performance Metrics for the Neural Network with Loss Function Based on no arbitrage in predicting the different exchange rates. With Hyperparameters of 64 nodes in the first layer, 122 in the second, 122 third layer, a Dropout rate of 0.3, and a learning rate of 0.01

In Table 3, we see the results of the Neural Network based on the No Triangular Arbitrage Loss function (12). We already saw in Section 4.1 that the No Triangular Arbitrage Hypotheses was oftentimes violated, hence expected a worse performance than for an MSE-based model. We see this result the average MAE is 0.07662, which is the highest of all models. A MSE of 0.01402 for the average prediction, which is also the highest of all the models, and the same holds for the average SMAPE of 18.59741. This model also performs worse on all metrics for all exchange rates tried.

In the model in Table 4, the Neural Network is trained on the MSE loss function (11).

This model has the lowest average MAE (0.04418), MSE (0.00474), and SMAPE (5.37573) of all the models, as can be seen in Table 4. Indicating that this model is the best-performing model in all fields.

Both Zhang and Hu (1998) and Galeshchuk (2016) outperform the RW as well.

	USD/CHF	USD/CAD	CHF/CAD	Average
MAE	0.03567	0.03123	0.06565	0.04418
MSE	0.00424	0.00430	0.00595	0.00474
SMAPE	3.34673	3.97756	8.80292	5.37573

Table 4: Performance Metrics for the Neural Network with Loss Function Based on no MSE in predicting the different exchange rate. With Hyperparameters, 32 nodes in the first layer, 122 in the second, 122 third layer, a Dropout rate of 0.4, and a learning rate of 0.01.

4.3 SHAP Interpretation

In this Section, the SHAP interpretation of the Neural Network model as used in Table 4 is presented. Only the SHAP values of the best working Neural Network model are calculated. Since these calculations are very computationally intensive, for each of the predicted exchange rates, the model has a different feature importance. Note that only the 20 most important variables are shown.

4.3.1 USD/CHF SHAP

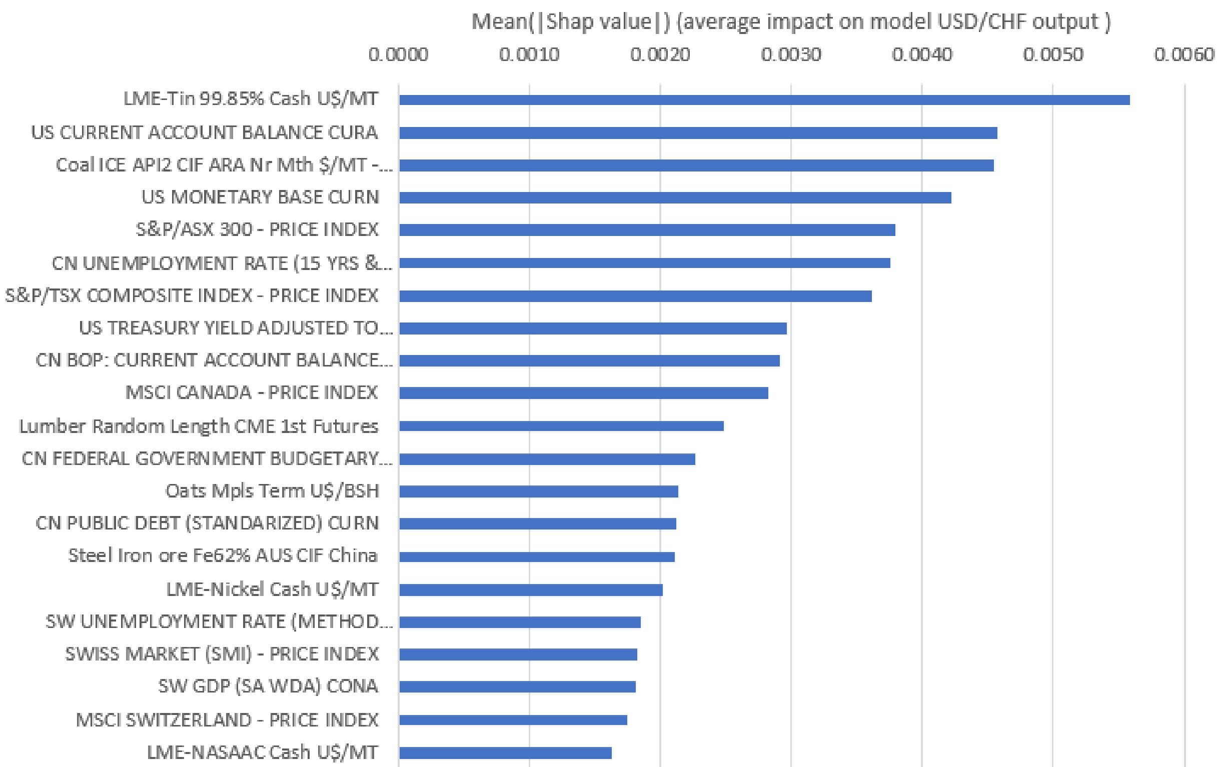

Figure 4: The absolute mean SHAP impact of the 20 most important variables on the USD/CHF exchange rate, the mean is taken because SHAP is defined as the deviations from the base value

In Figure 4, we can see that the most important predictor is Tin for the exchange rate between USD and CHF. In the same Figure, we also see macroeconomic structural influential predictors. For example, the US current account balance, the Canadian unemployment rate, and the US treasury yield. Interestingly enough, we see that the current model suggests that Canadian macroeconomic structural variables influence the exchange rate between the USD and CHF more than the ones of Switzerland do. This is something that would require more research, this could be a NN quirk, or one suggestion might be that Canada's economy and that of the US are interwoven. According to the US government[2] 718 billion USD was traded in goods in 2019, which is a very significant chunk of the total American economy. Hence, for a definite conclusion, this would require more research, but that falls outside the scope of this research.

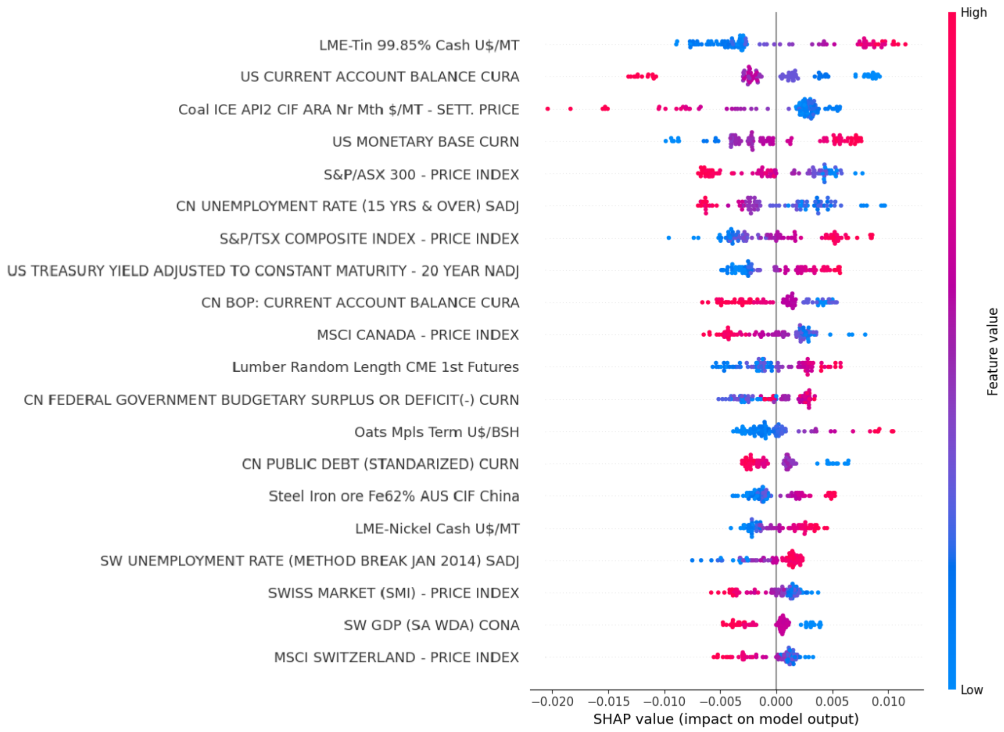

Figure 5: SHAP USD/CHF the higher the feature relative to its own range, the redder the dot, and the lower, the bluer the dot. The placement of the dot indicates the importance of the observation.

In Figure 5, the SHAP relations of the most important predictors with respect to the USD/CHF exchange rate are shown. For all of the variables, we see that the SHAP importance depends on the relative feature value. If the color scheme is from blue left to right red, the variable has a positive relationship with the exchange rate. Meaning that a higher value of that value will result in numerically a higher exchange rate. The other way around also holds true. The purple dots indicate the average value of the feature

[2]https://ustr.gov/countries-regions/americas/canada

value. Note that an average value of the feature does not necessarily mean the average impact since the power of the Neural Network is Non-Linear relationships.

Some interesting results of this are that the price of Tin is positively correlated with the height of the USD/CHF, suggesting that the price of Tin leads to a relative appreciation of the USD compared to the CHF.

We find that in the current setup, the US current account balance is negatively correlated with the USD/CHF rate. Suggesting that a high current account for the US correlates with a low appreciation of the USD compared to the CHF. This is in line with monetary theory from Pilbeam (2018).

Another finding is that a high yield on the US 20-year bond correlates with a higher appreciation of the USD compared to the CHF. This is in line with Clarida and Waldman (2008).

Furthermore, we see that high unemployment in swiss results in a higher appreciation of the dollar, which is in line with expectations about productivity (Chen and Rogoff, 2003). In Chen and Rogoff (2003), an indirect relationship between the unemployment rate and exchange through productivity is found.

4.3.2 USD/CAD SHAP

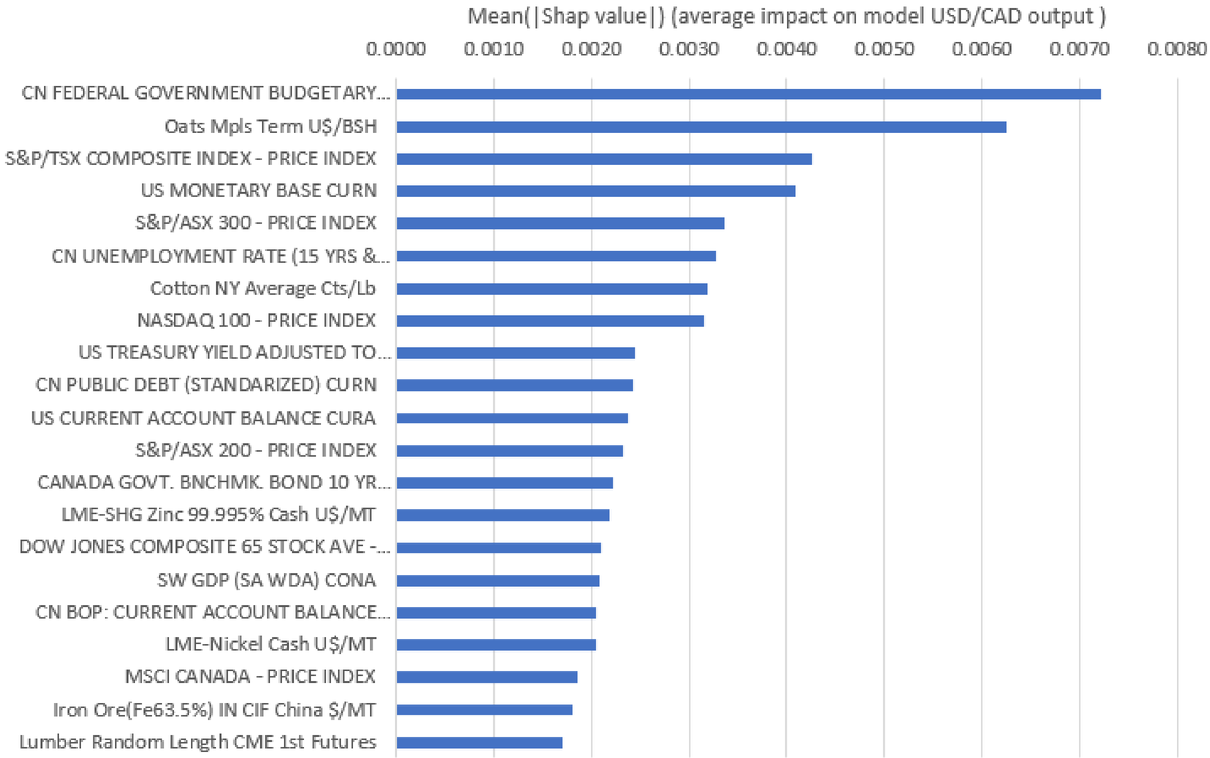

Figure 6: The absolute mean SHAP impact of the 20 most important variables on the USD/CAD exchange rate, the mean is taken because SHAP is defined as the deviations from the base value

In Figure 7, we can see the 20 most important variables for the prediction of the USD/CAD. We see that of the most important variables, many are macroeconomic structural variables. The Canadian Government Budgetary Surplus or Deficit CURN is the most important predictor, with an average absolute value of 0.0072. Figure 7 suggests that a higher Canadian budget surplus correlates with a lower appreciation of the CAD relative to the USD. Pilbeam (2018) also finds this result.

We also see that the S&P/TSX Price index, which is S&P relative to the Canadian stock market, plays an important role, with a higher price on the Canadian stock market increasing the price of the CAD relative to the USD. A similar result was found for the CHF in Section 4.3.1. The MSCI Canadian price index also shows this.

Another striking result is that a higher government bond yield of the Canadian government correlates with an increased price of CAD relative to the USD. The other way around also happens for the USD 20-year treasury yield and an appreciation for the USD with a higher yield. This result is in line with the findings of Pilbeam (2018).

Also, we see once more that a higher current account balance of the US correlates with a low exchange rate of the USD relative to the CAD.

A higher Canadian Unemployment leads to a depreciation of the CAD. We also saw this result for the CHF and Swiss unemployment.

A high Dow Jones price correlates with a high US dollar appreciation.

Figure 7: SHAP USD/CHF the higher the feature relative to its own range, the redder the dot, and the lower, the bluer the dot. The placement of the dot indicates the importance of the observation.

4.3.3 CHF/CAD SHAP

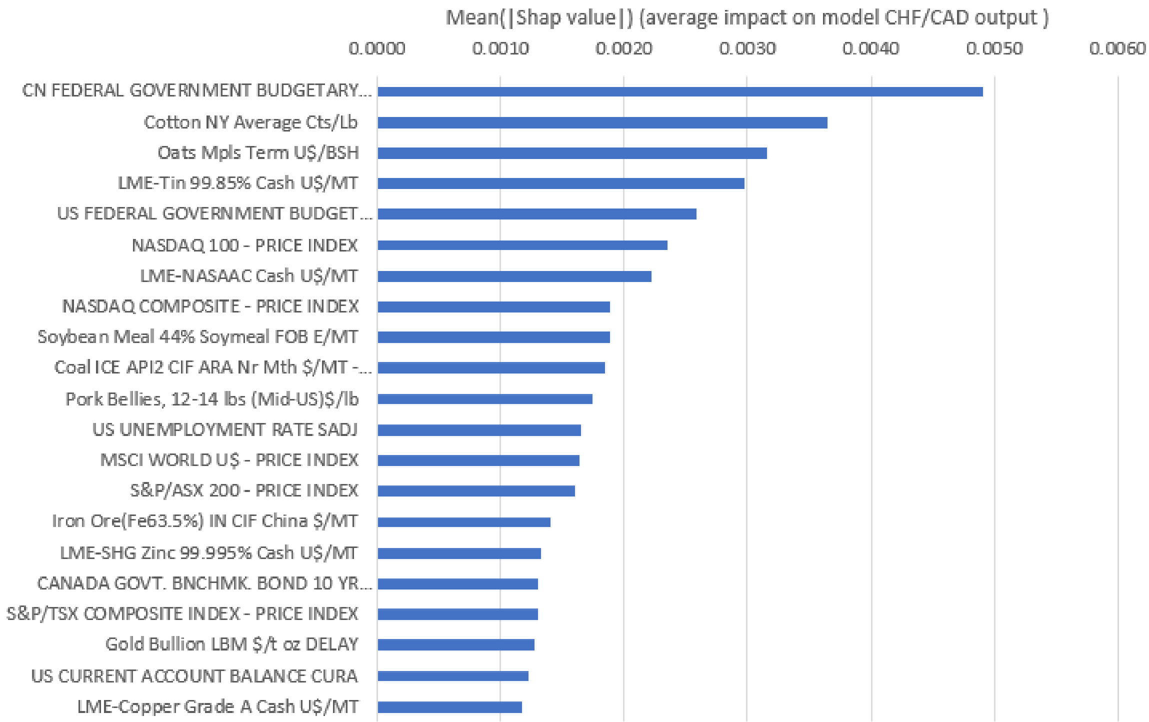

Figure 8: The absolute mean SHAP impact of the 20 most important variables on the CHF/CAD exchange rate, the mean is taken because SHAP is defined as the deviations from the base value

We see once more in Figure 8 that the Canadian Government Budgetary Surplus or Deficit CURN with an average absolute value of 0.00494 is the most important predictor. The relationships between CHF/CAD are shown in Figure 9. In Figure 9, we see, for example, that the Canadian Government Budgetary Surplus or Deficit CURN has the relationship that a higher surplus indicates a relatively low appreciation of the Canadian dollar. Meaning that the influence of the lower supply of bonds and hence higher prices so lower yields outweigh the benefit of a strong economy and stability. For the Comparison between Canada and Swiss according to Pilbeam (2018).

We see the same occurrence of the relative importance of a third economy in between the exchange rate of two other economies as we saw in Section 4.3.1. The reason as to why would be interesting for further research. However, one can argue that the predictive power is more important than the workings of a Neural Network.

We see once more that a high yield in the bond of a county correlates with a higher appreciation of the currency.

Interestingly for all the metals and the LME, a high price correlates with a high appreciation of the CHF compared to the CAD, except for gold and Tin.

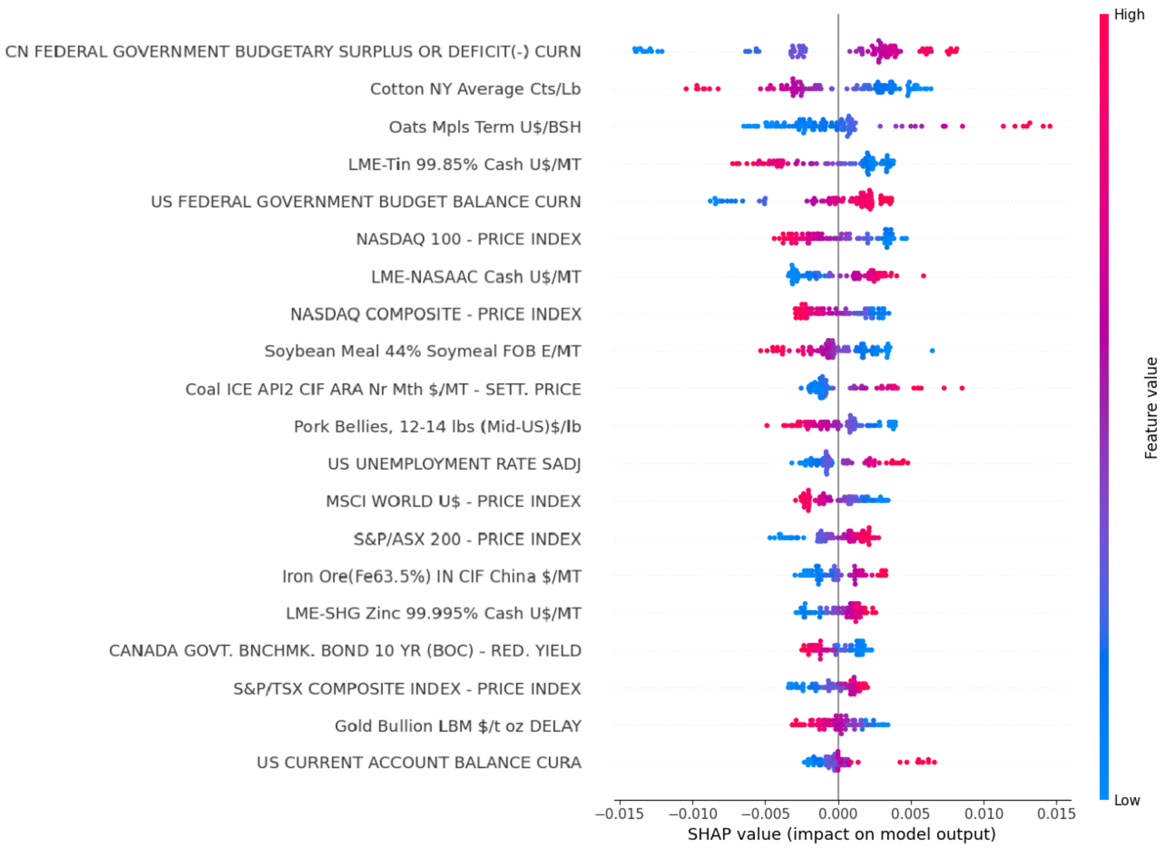

Figure 9: SHAP USD/CHF the higher the feature relative to its own range, the redder the dot, and the lower, the bluer the dot. The placement of the dot indicates the importance of the observation.

5 Conclusion

In conclusion, our study investigated the performance of various models for predicting exchange rates, focusing on using structural variables as input for NN to determine whether this approach would yield better results than a Random Walk. The no triangular arbitrage hypothesis is frequently violated, suggesting that a Neural Network based on a Mean Squared Error (MSE) loss function is more suitable for predictions.

To find the best-performing NN model for the current problem, we optimized various hyperparameters using Bayesian optimization. This allowed us to identify the optimal combination of hyperparameters that led to the best model performance for the current set-up.

Our results show that the NN model employing an MSE-based loss function and optimized hyperparameters outperforms the other models and even the RW model. With the lowest Mean Absolute Error (MAE), Mean Squared Error (MSE), and Symmetric Mean Absolute Percentage Error (SMAPE) values. In contrast, the model utilizing a no arbitrage-based loss function performed the worst across all metrics. Suggesting that exploring alternative loss functions and network architectures could be fruitful for future research in increasing the predictive capabilities of NN models in financial markets. Alternative architecture suggestions could, for example, include Recurrent Neural Networks (RNNs) or Long Short-Term Memory Networks (LSTMs), which might be better suited for capturing temporal patterns in exchange rate data.

In the current research, we have also employed the SHapley Additive exPlanations (SHAP) method to gain a deeper understanding of the effects of the input features on prediction outcomes. We found that various structural variables greatly impacted the model prediction. We now have looked mostly at structural variables, but incorporating political and legal factors as additional features may lead to a more robust prediction of exchange rates. Cross-market analysis might also be another expansion suited for future research. We saw that the Canadian market significantly affected the USD/CHF exchange rate. Hence, adding more markets to the features might increase explanatory power of the model.

Another promising future addition would be implementing a rolling window framework for forecasts, which dynamically adjusts the training and testing data over time; this, however, might be very computationally intensive.

References

Aiba, Y., Hatano, N., Takayasu, H., Marumo, K. & Shimizu, T. (2002). Triangular arbitrage as an interaction among foreign exchange rates. *Physica A: Statistical Mechanics and its Applications*, *310*(3-4), 467–479.

Ait-Sahalia, Y., Fan, J., Xue, L. & Zhou, Y. (2022). How and when are high-frequency stock returns predictable? *Available at SSRN 4095405.*

Athey, S. & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, *11*, 685–725.

Chen, Y.-c. & Rogoff, K. (2003). Commodity currencies. *Journal of international Economics*, *60*(1), 133–160.

Chen, Y.-c., Turnovsky, S. J. & Zivot, E. (2014). Forecasting inflation using commodity price aggregates. *Journal of Econometrics*, *183*(1), 117–134.

Clarida, R. H. & Waldman, D. (2008). Is bad news about inflation good news for the exchange rate? and, if so, can that tell us anything about the conduct of monetary policy? *Asset prices and monetary policy* (pp. 371–396). University of Chicago Press.

Cushman, D. O. (1985). Real exchange rate risk, expectations, and the level of direct investment. *The Review of Economics and Statistics*, 297–308.

Dash, R. (2018). Performance analysis of a higher order neural network with an improved shuffled frog leaping algorithm for currency exchange rate prediction. *Applied Soft Computing*, *67*, 215–231. https://doi.org/https://doi.org/10.1016/j.asoc.2018.02.043

Dewenter, K. L. (1995). Do exchange rate changes drive foreign direct investment? *Journal of Business*, 405–433.

Dornbusch, R. (1976). Expectations and exchange rate dynamics. *Journal of political Economy*, *84*(6), 1161–1176.

Doshi-Velez, F. & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608.*

Duchi, J., Hazan, E. & Singer, Y. (2011). Adaptive subgradient methods for online learning and stochastic optimization. *Journal of machine learning research*, *12*(7).

Falessi, D., Narayana, L., Thai, J. F. & Turhan, B. (2018). Preserving order of data when validating defect prediction models. *California Polytechnic State University.*

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*(2), 383–417.

Fenn, D. J., Howison, S. D., McDonald, M., Williams, S. & Johnson, N. F. (2009). The mirage of triangular arbitrage in the spot foreign exchange market. *International Journal of Theoretical and Applied Finance*, *12*(08), 1105–1123.

Frankel, J. A. (1979). On the mark: A theory of floating exchange rates based on real interest differentials. *The American economic review*, *69*(4), 610–622.

Froot, K. A. & Stein, J. C. (1991). Exchange rates and foreign direct investment: An imperfect capital markets approach. *The quarterly journal of economics*, *106*(4), 1191–1217.

Galeshchuk, S. (2016). Neural networks performance in exchange rate prediction. *Neurocomputing*, *172*, 446–452. https://doi.org/https://doi.org/10.1016/j.neucom.2015.03.100

Galeshchuk, S. & Mukherjee, S. (2017). Deep networks for predicting direction of change in foreign exchange rates. *Intelligent Systems in Accounting, Finance and Management*, *24*(4), 100–110.

Hara, K., Saito, D. & Shouno, H. (2015). Analysis of function of rectified linear unit used in deep learning. *2015 international joint conference on neural networks (IJCNN)*, 1–8.

Hornik, K., Stinchcombe, M. & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural networks*, *2*(5), 359–366.

Ince, H. & Trafalis, T. B. (2006). A hybrid model for exchange rate prediction. *Decision Support Systems*, *42*(2), 1054–1062.

Ioffe, S. & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *International conference on machine learning*, 448–456.

Kilian, L. & Taylor, M. P. (2003). Why is it so difficult to beat the random walk forecast of exchange rates? *Journal of International Economics*, *60*(1), 85–107.

Krizhevsky, A., Sutskever, I. & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90.

Liang, J. & Liu, R. (2015). Stacked denoising autoencoder and dropout together to prevent overfitting in deep neural network. *2015 8th International Congress on Image and Signal Processing (CISP)*, 697–701. https://doi.org/10.1109/CISP.2015.7407967

Lundberg, S. M. & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, *30*.

Meese, R. A. & Rogoff, K. (1983). Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of international economics*, *14*(1-2), 3–24.

Nair, V. & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. *Proceedings of the 27th international conference on machine learning (ICML-10)*, 807–814.

Ni, H. & Yin, H. (2009). Exchange rate prediction using hybrid neural networks and trading indicators. *Neurocomputing*, *72*(13-15), 2815–2823.

Nwankpa, C., Ijomah, W., Gachagan, A. & Marshall, S. (2018). Activation functions: Comparison of trends in practice and research for deep learning. *arXiv preprint arXiv:1811.03378.*

Pilbeam, K. (2018). *Finance and financial markets*. Bloomsbury Publishing.

Real, E., Moore, S., Selle, A., Saxena, S., Suematsu, Y. L., Tan, J., Le, Q. V. & Kurakin, A. (2017). Large-scale evolution of image classifiers. *International Conference on Machine Learning*, 2902–2911.

Santurkar, S., Tsipras, D., Ilyas, A. & Madry, A. (2018). How does batch normalization help optimization? *Advances in neural information processing systems*, *31*.

Shahriari, B., Swersky, K., Wang, Z., Adams, R. P. & De Freitas, N. (2015). Taking the human out of the loop: A review of bayesian optimization. *Proceedings of the IEEE*, *104*(1), 148–175.

Shapley, L. S. (1953). Stochastic games. *Proceedings of the national academy of sciences*, *39*(10), 1095–1100.

Smith, L. N. (2017). Cyclical learning rates for training neural networks. *2017 IEEE winter conference on applications of computer vision (WACV)*, 464–472.

Snoek, J., Larochelle, H. & Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems*, *25*.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *The journal of machine learning research*, *15*(1), 1929–1958.

Wang, J. & Chen, Y. (2021). Adaboost-based integration framework coupled two-stage feature extraction with deep learning for multivariate exchange rate prediction. *Neural Processing Letters*, *53*(6), 4613–4637.

Wu, J., Chen, X.-Y., Zhang, H., Xiong, L.-D., Lei, H. & Deng, S.-H. (2019). Hyperparameter optimization for machine learning models based on bayesian optimizationb. *Journal of Electronic Science and Technology*, *17*(1), 26–40. https://doi.org/https://doi.org/10.11989/JEST.1674-862X.80904120

Zhang, G. & Hu, M. Y. (1998). Neural network forecasting of the british pound/us dollar exchange rate. *Omega*, *26*(4), 495–506. https://doi.org/https://doi.org/10.1016/S0305-0483(98)00003-6

Zoph, B. & Le, Q. V. (2016). Neural architecture search with reinforcement learning. *arXiv preprint arXiv:1611.01578.*

A. Appendix

	2 Layers NTA	1 Layer NTA	2 Layers MSE	1 Layer MSE
MSE	0.0553	1.3532	0.0058	0.2166
MAE	0.1916	0.8740	0.0688	0.3007
SMAPE	25.8163	98.9333	7.8547	33.3600

Table 5: All the models Neural Network with number of layers tried, NTA is the loss function based on NTA, MSE loss function based on MSE

B. Appendix

A list of all the used variables

- Crude Oil-WTI Spot Cushing U$/BBL
- Cotton NY Average Cts/Lb
- Lumber Random Length CME 1st Futures
- Cocoa-ICCO Daily Price US$/MT
- Coffee-ICO Composite Daily ICA c/lb
- NYCE-(FCOJ-A) ORANGE JUICE N-C. LG. - OPEN INTEREST
- Raw Sugar-ISA Daily Price c/lb
- Corn No.2 Yellow U$/Bushel
- Wheat No.2,Soft Red U$/Bu
- Oats Mpls Term U$/BSH
- CBT-ROUGH RICE NON-REP. LONG - OPEN INTEREST
- Soybean Meal 44% Soymeal FOB E/MT
- Yellow Soybn US NO.1 Sth Dvprt U$/Bsh
- CME-FEEDER CATTLE NON-COML. LONG - OPEN INTEREST
- Lean HOG 53-54% US 3 AREA Del U$/Cwt
- CME-LIVE CATTLE NON-COML. LONG - OPEN INTEREST
- Pork Bellies, 12-14 lbs (Mid-US)$/lb
- Steel Iron ore Fe62% AUS CIF China
- Iron Ore(Fe63.5%) IN CIF China $/MT

- Coal ICE API2 CIF ARA Nr Mth $/MT - SETT. PRICE
- Gold Bullion LBM $/t oz DELAY
- LME-Copper Grade A Cash U$/MT
- Palladium U$/Troy Ounce
- London Platinum Free Market $/Troy oz
- Silver, Handy&Harman (NY) U$/Troy OZ
- LME-NASAAC Cash U$/MT
- LME-SHG Zinc 99.995% Cash U$/MT
- LME-Nickel Cash U$/MT
- LME-Lead Cash U$/MT
- LME-Tin 99.85% Cash U$/MT
- US DOLLAR SW DEPOSIT (FT/RFV) - MIDDLE RATE
- US DOLLAR 6M DEPOSIT (FT/RFV) - MIDDLE RATE
- US DOLLAR 1Y DEPOSIT (FT/RFV) - MIDDLE RATE
- US REPO RATE 3M - RFV COMPOSITE - OFFERED RATE
- US DOLLAR 1M DEPOSIT (FT/RFV) - MIDDLE RATE
- US DOLLAR 3M DEPOSIT (FT/RFV) - MIDDLE RATE
- SWISS FRANC 3M DEPOSIT (FT/RFV) - MIDDLE RATE
- SWISS FRANC 1M DEPOSIT (FT/RFV) - MIDDLE RATE
- SWISS FRANC S/T DEPOSIT (FT/RFV) - MIDDLE RATE
- SWISS FRANC 1Y DEPOSIT (FT/RFV) - MIDDLE RATE
- US DOLLAR S/T DEPOSIT (FT/RFV) - MIDDLE RATE
- SWISS FRANC 6M DEPOSIT (FT/RFV) - MIDDLE RATE
- SWISS FRANC SW DEPOSIT (FT/RFV) - MIDDLE RATE
- RF SWITZERLAND GVT BMK BID YLD 10Y - RED. YIELD
- RF SWITZERLAND GVT BMK BID YLD 1Y - RED. YIELD
- RF SWITZERLAND GVT BMK BID YLD 2Y - RED. YIELD
- RF SWITZERLAND GVT BMK BID YLD 3M - RED. YIELD

- RF SWITZERLAND GVT BMK BID YLD 1M - RED. YIELD
- RF SWITZERLAND GVT BMK BID YLD 4Y - RED. YIELD
- RF SWITZERLAND GVT BMK BID YLD 5Y - RED. YIELD
- RF SWITZERLAND GVT BMK BID YLD 6M - RED. YIELD
- RF US GVT BMK BID YLD 10Y - RED. YIELD
- RF US GVT BMK BID YLD 2Y - RED. YIELD
- RF US GVT BMK BID YLD 3M - RED. YIELD
- RF US GVT BMK BID YLD 5Y - RED. YIELD
- RF US GVT BMK BID YLD 30Y - RED. YIELD
- RF US GVT BMK BID YLD 1Y - RED. YIELD
- RF US GVT BMK BID YLD 1M - RED. YIELD
- RF US CORP BMK BBB 10Y - RED. YIELD
- CANADIAN DOLLAR 3M DEPOSIT (FT/RFV) - MIDDLE RATE
- CANADIAN DOLLAR 1M DEPOSIT (FT/RFV) - MIDDLE RATE
- CANADIAN DOLLAR 1Y DEPOSIT (FT/RFV) - MIDDLE RATE
- CANADIAN DOLLAR 6M DEPOSIT (FT/RFV) - MIDDLE RATE
- CANADIAN DOLLAR SW DEPOSIT (FT/RFV) - MIDDLE RATE
- CANADA GOVT. BNCHMK. BOND 10 YR (BOC) - RED. YIELD
- CANADA TREASURY BILL 3 MONTH - MIDDLE RATE
- CANADA GOVERNMENT BOND 5-10 YEAR - RED. YIELD
- CANADA TREASURY BILL 1 YEAR (BOC) - MIDDLE RATE
- RF CANADA GVT BMK BID YLD 10Y - RED. YIELD
- CANADA CHARTERED BK. CONV. MTGE. 1 YR - MIDDLE RATE
- SWISS SBI DMS.GVT. 3-7Y (SF)(BID) - PRICE INDEX
- SWISS SBI DMS.GVT. 7+Y (SF)(BID) - PRICE INDEX
- SWISS SBI DMS.GVT. 1-3Y (SF) - PRICE INDEX
- US BENCHMARK 10 YEAR DS GOVT. INDEX - PRICE INDEX
- MSCI WORLD U$ - PRICE INDEX

- S&P/ASX 200 - PRICE INDEX
- S&P/TSX COMPOSITE INDEX - PRICE INDEX
- S&P/TSX 60 INDEX - PRICE INDEX
- MSCI CANADA - PRICE INDEX
- SWISS MARKET (SMI) - PRICE INDEX
- MSCI SWITZERLAND - PRICE INDEX
- SWISS MARKET INDEX - AGGREGATE PRICE
- S&P 500 COMPOSITE - PRICE INDEX
- NASDAQ COMPOSITE - PRICE INDEX
- S&P 500 TR (1970) - TOT RETURN IND
- NASDAQ 100 - PRICE INDEX
- S&P/ASX 300 - PRICE INDEX
- DOW JONES COMPOSITE 65 STOCK AVE - PRICE INDEX
- USD TO AUD 1M FWD OR (WMR) - EXCHANGE RATE
- US NOMINAL ADVANCED FOREIGN ECONOMIES DOLLAR INDEX NADJ
- US TRADE WEIGHTED EXCHANGE RATE, NOMINAL (STANDARDIZED) NADJ
- US GDP (AR) CONA
- CN GDP (AR) CONA
- SW GDP (SA WDA) CONA
- SW PRODUCTIVITY - GDP PER EMPLOYED PERSON SADJ
- US OUTPUT PER HOUR OF ALL PERSONS - BUSINESS SECTOR VOLA
- CN LABOUR PRODUCTIVITY - BUSINESS SECTOR VOLA
- US PUBLIC DEBT OUTSTANDING CURN
- SW EXTERNAL DEBT - PUBLIC SECTOR CURN
- CN PUBLIC DEBT (STANDARIZED) CURN
- US CURRENT ACCOUNT BALANCE CURA
- SW BOP: CURRENT ACCOUNT BALANCE CURN

- CN BOP: CURRENT ACCOUNT BALANCE CURA
- US CPI - ALL URBAN: ALL ITEMS SADJ
- CN CPI NADJ
- SW CPI SADJ
- US GOODS TRADE BALANCE ON A BALANCE OF PAYMENTS BASIS CURA
- CN VISIBLE TRADE BALANCE (BALANCE OF PAYMENTS BASIS) CONA
- SW VISIBLE TRADE BALANCE (TOTAL 2) CURN
- US UNEMPLOYMENT RATE SADJ
- SW UNEMPLOYMENT RATE (METHOD BREAK JAN 2014) SADJ
- CN UNEMPLOYMENT RATE (15 YRS & OVER) SADJ
- US TREASURY YIELD ADJUSTED TO CONSTANT MATURITY - 20 YEAR NADJ
- CN GOVERNMENT BOND YIELD - OVER 10 YEARS (EP) NADJ
- SW YIELD 10-YEAR GOVERNMENT BONDS NADJ
- CN OVERNIGHT MONEY MARKET FINANCING RATE (EP) NADJ
- US INTERBANK RATE - 3 MONTH (LONDON) (MONTH AVG) NADJ
- US MONETARY BASE CURN
- SW MONETARY AGGREGATE M3 CURN
- CN CREDIT LIABS.OF NONFINL,NON-MORTGAGE LOANS-CHART
- ERED BANKS
- US FEDERAL GOVERNMENT BUDGET BALANCE CURN
- CN FEDERAL GOVERNMENT BUDGETARY SURPLUS OR DEFICIT(-) CURN