A Hybrid Approach of Wavelet Transform, Convolutional Neural Networks and Gated Recurrent Units for Stock Liquidity Forecasting

Mohamed Ben Houad[1], Mohammed Mestari[2], Khalid Bentaleb[3], Adnane El Mansouri[4] and Salma El Aidouni[5]
Laboratory of 2IACS, ENSET, Hassan II University, PO Box 159 Bd Hassan II,
28830, Mohammedia, Morocco

***Abstract*—Stock liquidity forecasting is critical for investors, issuers, and financial market regulators. The object of this study is to propose a method capable of accurately predicting the liquidity of stocks. The few studies on stock liquidity forecasting have focused on single models such as Seasonal Auto-Regressive Integrated Moving Average with eXogenous factors, the nonlinear autoregressive network with exogenous input, and Deep Learning. A new trend in forecasting which attempts to combine several approaches is emerging at the moment. Inspired by this new trend, we propose a hybrid approach of Wavelet Transform, Convolutional Neural Networks, and Gated Recurrent Units to predict stock liquidity. Our model is tested on daily data of companies listed on the Casablanca Stock Exchange from 2000 to 2021. Its forecasting performances are evaluated based on the Mean Absolute Error, the Root Mean Square Error, the Mean Absolute Percentage Error, Theil's U statistic, and the correlation coefficient. Finally, the outperformance of the proposed model is confirmed by comparison with other reference forecasting models. This study contributes to the enrichment of the field of prediction of financial risks and can constitute a framework of analysis allowing to help the stakeholders of the financial markets to forecast the liquidity of the actions.**

***Keywords*—Stock liquidity; wavelet transform; convolutional neural networks; GRU cell; Casablanca stock exchange**

I. INTRODUCTION

"A stock is considered liquid if large transactions can be made rapidly without significantly impacting the stock price and without incurring substantial losses, and if any price variation caused by a random shock is quickly adjusted" [1]. Before the financial crisis of 2007-2008, stock liquidity risk was largely underestimated by investors, financial market regulators, and researchers. Yet, it has negative financial and economic consequences. In fact, It increases equity market risk [2], [3], [4], and reduces bank liquidity [5], [6]. Stock liquidity affects financial stability [7], [8]. It also impacts the financial structure and cost of capital [9], [10], the dividend distribution policy [11], and the risk of corporate failures [12], [13].

Forecasting stock liquidity is crucial for investors, issuers, and financial market regulators. It allows investors to forecast the illiquidity premium to be charged to compensate for lower returns [14]. Stock liquidity prediction helps issuers to choose the right time to go public, increase their capital or carry out financial packages such as takeover bids, sales, or the outs. Financial market regulators are also concerned with liquidity predictions as it allows them to act a priori to safeguard financial stability.

Only a few studies have attempted to address forecasting stock liquidity. We are aware of only two research articles in this area. In a comparative study, [15], concludes that the Nonlinear Autoregressive Network with eXogenous inputs (NARX) has better predictive performance than the Seasonal Auto-Regressive Integrated Moving Average with eXogenous factors (SARIMAX). These authors found that SARIMAX method is inaccurate because stock liquidity is irregular, noisy, and nonlinear time series. However, this study has some shortcomings.

It is based on only 108 observations, a number that we consider to be low for the learning processes of an algorithm capable of making effective predictions. The training of this type of algorithm on the basis of the large dataset can generate the problems of vanishing and exploding of the gradient. Since the number of hidden layers is low, NARX neural networks cannot capture hidden functional relationships in the historical stock liquidity dataset. As a result, their predictive performance is very limited.

The author in [16] compared linear regression model, multilayer perceptron, and Long Short Term Memory (LSTM). Based on daily data of companies listed on the Ho Chi Min and Hanoi Stock Exchanges in Vietnam from January 2011 to December 2019, the authors conclude that the LSTM model has the lowest Mean Square Error (MSE). This result seems logical to us because LSTM neural networks have more advantages than the linear regression model and the multilayer perceptron. Compared to linear regression, the LSTM model capture the characteristics of even nonlinear data. The multilayer perceptron assumes that inputs and outputs are independent of each other. On the other hand, the LSTM model takes into account the temporal dependencies while avoiding the problem of vanishing gradient. However, despite the advantages of the LSTM model, the results of this study are less convincing. First, the predictive performance is only evaluated on the basis of a single criterion (MSE). Second, despite their power, LSTM neural networks alone cannot capture all abrupt and dynamic changes in financial time series [17].Stock liquidity is noisy, volatile, and non-linear. It requires pre-processing before forecasting.

The Wavelet Transform (WT) is an effective tool for denoising the most complex time series. By decomposing a signal into different scales, the WT can capture the hidden features of the time series. Therefore, to further improve the efficiency and accuracy of forecasting, researchers have started to develop hybrid models that combine WT and deep learning algorithms. These hybrid models are used to predict time series, such as wind [18], solar energy [19], water quality [20], nickel futures price [21], gold returns [22], and stock prices [23]. Exploring this approach, we propose a hybrid model that combines a Wavelet Transform (WT), a Convolutional Neural Network (CNN), and a Recurrent Neural Network (RNN) with a Gated Recurrent Unit (GRU) layer. The purpose of this study is to find out if stock liquidity denoising by WT can improve the predictive performance of deep learning algorithms. The proposed WT-CNN-GRU model is tested on daily data of companies listed on the Casablanca Stock Exchange from 2000 to 2021. Our database contains 5478 trading days.

This study brings three novelties. First, the originality of our approach is the denoising of stock liquidity data by the WT before proceeding to the forecasting by deep learning algorithms. Second, unlike previous studies that used unique signal decomposition methods to denoise the data, we use adaptive approaches consistent with the characteristics of different measures of stock liquidity. Third, the proposed model showed better performance in accurately predicting the strong disruptions in stock liquidity caused by the COVID-19 pandemic.

While the predictive performance of the previously presented models is measured by a single criterion, the proposed model in this study is evaluated by a multitude of parameters such as Mean Absolute Error, the Root Mean Square Error, the Mean Absolute Percentage Error, Theil's U statistic and the correlation coefficient. By showing superiority over previous studies, the proposed model is considered a step forward in improving the prediction of stock liquidity.

The rest of the paper proceeds as follows. The next section describes the adopted methodology. Section 3 presents the empirical process, results, and comparative analysis of stock liquidity forecasting. Section 4 presents the conclusions and offers some suggestions and perspectives.

II. Proposed Methodology

To forecast stock liquidity, the proposed methodology is a hybrid approach between a WT, CNN, and GRU. Fig. 1 shows the general procedure of our model.

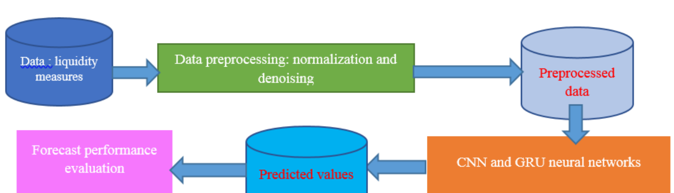

Fig. 1. Flowchart of the WT–CNN–GRU Model.

The normalized and denoised stock liquidity measures are inputs to the mixed CNN-GRU model. The detailed steps for processing the pre-processed data in the CNN-GRU model are shown in Fig. 2.

This section discusses the main steps of the hybrid WT-CNN-GRU model: (1) data preprocessing (2) a hybrid CNN-

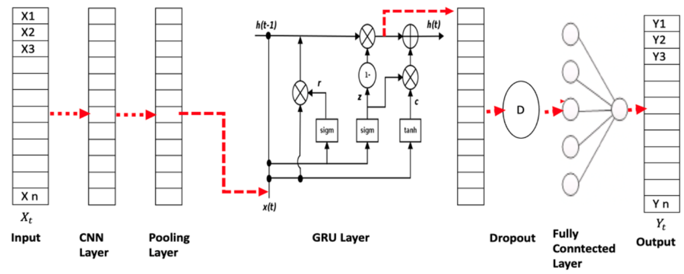

Fig. 2. Flow Chart of the CNN-GRU Model.

GRU neural network and (3) forecasting performance evaluation.

A. Data Pre-Processing

1) Data Normalization: The normalized data is calculated using the Z-Score method. Z-scores measure the distance between a data point and the means in terms of standard deviation. The standardized data set has a mean of 0 and a standard deviation of 1, and retains the shape properties of the original data set (same skewness and kurtosis). The standardized data are obtained by (1).

$$x^* = \frac{(x - \bar{x})}{\sigma} \tag{1}$$

$\bar{x}$ is the mean of the original data and σ is the standard deviation of the original data. Normalization supports machine learning algorithms in measuring the distance between the standard deviation and the mean of processed data samples. Conversely, The original data can be derived as follows:

$$x = \sigma x^* + \bar{x} \tag{2}$$

2) Wavelet Transform: Financial series are noisy, volatile, nonlinear, and non-stationary. As a consequence, they require pre-processing. The Discrete Wavelet Transform (DWT) is a mathematical tool that decomposes the input signal into several physically significant components, invisible in the raw data. These components can be frequencies, trends, edges, or breaks. This facilitates the analysis of each component in isolation and the reconstruction of the original signal into the desired components to be extracted. This facilitates the denoising of the input signal. The purpose of this step in our model is to eliminate the noise that can characterize stock liquidity.

There are several signal decomposition techniques such as Maximum Overlap Discrete Wavelet Transform (MODWT), Empirical Mode Decomposition (EMD), Empirical Wavelet Transform (EWT), Tunable Q-factor Wavelet Transform (TQWT), and Variational Mode Decomposition (VMD). The choice between these techniques depends on the characteristics of the input signal [1]. MODWT is adapted for signals containing oscillations with trends or transitions. TQWT and VMD are most suitable for signals containing high or low-frequency oscillations. EWT is intended for extracting low-frequency oscillations. EMD is used when the input signal contains trends.

However, before proceeding with the decomposition of input signals, it is necessary to analyze them in time frequency to choose the most appropriate decomposition method. The Continuous Wavelet Transform (CWT) is more efficient for performing time-frequency analysis of a signal than the DWT because the scales are discretized more finely in CWT. In this study, we prefer Morse Wavelets because it has the advantage of varying amplitude and frequency over time [24].

B. A Hybrid CNN-GRU Model

1) Convolutional Neural Network: A Convolutional Neural Network (CNN) is a network architecture for deep learning that learns directly from the data, eliminating the need for manual feature extraction. Fig. 3 shows a simple CNN architecture.

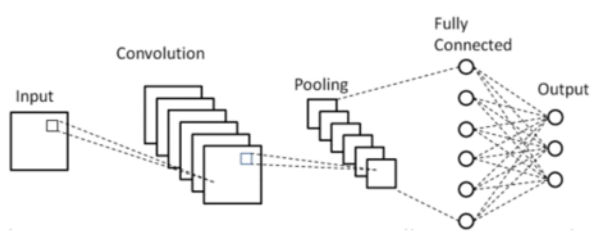

Fig. 3. Basic Architecture of the Convolutional Neural Network.

In addition to the input and output layers, three different layers are normally present in CNNs, such as the convolution layer, the pooling layer, and the fully connected layer. The convolution layer is a set of filters whose purpose is to extract local features from the input layer. This ensures that the network focuses on low-level features in the first hidden layer, then it assembles them into higher-level features in the next hidden layer, etc. Convolution layers are used in our study to extract chaotic, irregular, and fluctuating features from liquidity measurements. Pooling layers are used to retain only the most relevant features of the liquidity measures and to deepen them. Pooling can be of two types, maximum pooling, and average pooling. In this study, we retain maximum pooling because pooling by the mean is an outlier. The fully connected layer is similar to a Feedforward Neural Network whose goal is to extract the global feature of the inputs. Each neuron in these layers is connected to all hidden neurons in the previous layer.

2) Recurrent Neural Network: The Closed Recurrent Unit Layer: In contrast to deep Multilayer Perceptron (MLP), Recurrent Neural Networks (RNN) have interdependent input and output layers. RNNs are suitable for modeling sequential data and their associated temporal dynamics with greater accuracy. However, simple RNNs are generally characterized by the vanishing gradient problem, where, depending on the activation function, information "vanish with time," and the term nonlinearity is often inadequate for longer-term memory. To overcome this problem, Long Short-Term Memory networks (LSTMs) have been developed. LSTMs help to preserve errors that can be back-propagated across time and layers. By maintaining these errors, LSTMs allow RNNs to continue learning more efficiently across many time steps.

Compared to LSTM networks, GRUs [25], have only two gates; a reset gate and an update gate. The update gate behaves similarly to the forget gate in LSTM by deciding which information to keep and which new information to add, while the reset gate is another mechanism to determine the amount of past temporal information to delete. We retain GRUs in our model because they are less easy to construct than LSTMs, due to fewer tensor operations.

As mentioned earlier, this study proposes a hybrid model of CNN and GRU, with corresponding parameters summarized in Table I.

TABLE I. PARAMETER OF THE CNN-GRU MODEL

Hybrid model	Parameter	Values
Network CNN	Input layer	1
	Dimension convolution layer	1 D
	Number of convolution layers	2
	Number of Pooling layers	2
	Number of filters	1
	Width of filters	1
	Pooling method	Max Pooling
GRU network	Number of GRU layers	1
	Number of masked units	300
	Activation function to update the masked state	Tanh
	Activation function to be applied to gates	Sigmoid
Dropout layer Options	dropout rate	0.4
Options Learning Network CNN-GRU	Epochs	300
	optimizer	Adam
	Initial learning rate	0.005
	Loss function	Mean Square Error

3) Evaluation of the Forecasting Performance: To evaluate the prediction performance of the proposed model, five statistical evaluation indicators are used to compare the performance of the associated models, including MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error), RMSE (Root Mean Square Error), Theil's U statistic and the correlation coefficient (R^2) which can be calculated as follows:

$$MAE(X,\vec{X}) = \frac{\sum_{t=1}^{T} \|X_t - \vec{X}_t\|}{T} \tag{3}$$

$$MAPE(X,\vec{X}) = \left(\sum_{t=1}^{T} \frac{\|X_t - \vec{X}_t\|}{X_t} \right) \Big/ T \tag{4}$$

$$RMSE(X,\vec{X}) = \sqrt{\frac{\sum_{t=1}^{T} (X_t - \vec{X}_t)^2}{T}} \tag{5}$$

$$\begin{aligned} U(X,\vec{X}) = &\left(\sqrt{\frac{\sum_{t=1}^{T} (X_t - \vec{X}_t)^2}{T}} \right) \\ &\Big/ \left(\sqrt{\frac{\sum_{t=1}^{T} (\vec{X}_t)^2}{T}} + \sqrt{\frac{\sum_{t=1}^{T} (X_t)^2}{T}} \right) \end{aligned} \tag{6}$$

$$R^2(X.\vec{X}) = 1 - \left(\frac{\sum_{t=1}^{T} (X_t - \vec{X}_t)^2}{\sum_{t=1}^{T} (X_t - \bar{X})^2} \right) \quad (7)$$

Where T is the number of observations, X is the actual value, $\vec{X}$ is the forecasted value.

III. Empirical Results and Discussion

A. Sample, Liquidity Measures, and Data Analysis

1) Sample and Liquidity Measures: To test our model, we collect the best bid and ask price, closing price, and daily trading volume of 75 companies listed on the C.S.E from 04/01/2000 to 12/31/2021. These data come from the CDG Capital Bourse database. Stock liquidity is evaluated using three indicators to account for its multidimensional nature, including the displayed range (Qs), the Amihud illiquidity measure (Amh) and the zero return (Zr). The formulas for calculating these indicators are as follows:

$$Qs_{i,t} = \frac{(P_{i,t}^A - P_{i,t}^B)}{P_{i,t}^M} \quad (8)$$

$$Amh_{i,t} = \left(\sum_{t=1}^{D} \frac{\|r_{i,t}\|}{Vol_{i,t}} \right) \Big/ D_{i,t} \quad (9)$$

$$Zr_{i,t} = \begin{cases} 1, & r_{i,d} = 0 \\ 0, & r_{i,d} \neq 0 \end{cases} \quad (10)$$

With $P_{i,t}^A$, $P_{i,t}^B$, $P_{i,t}^M$, $r_{i,t}$ and $Vol_{i,t}$ are the best ask price, best bid price, closing price, daily return, and daily volume of stock i , respectively. The displayed range $Qs_{i,t}$ measures the depth of the market. The larger the spread between the best ask price and the best bid price, relative to the market price, the lower the liquidity of the stock [5]. Amihud's illiquidity ratio ($Amh_{i,t}$) describes the change in daily price for a given trading volume; a low trading volume generating a higher return is synonymous with illiquidity. The zero return ($Zr_{i,t}$) measures the number of days when the return is zero. It takes the value 1 if the return is zero and 0 otherwise. A high number of days of zero return is synonymous with low liquidity.

Each indicator is calculated daily per company; the aggregate daily indicator is an average of all companies. In sum, we have 5268 observations for Qs, 5469 observations for Amh and 5475 for Zr.

2) Descriptive Statistics: Fig. 4 shows that the measures of stock liquidity are chaotic, erratic, asymmetric, and non-linear to time. These characteristics are evidenced by the descriptive statistics presented in Table II. The liquidity measures are volatile as indicated by the coefficients of variation that are close to 0.5, and exceed 1 for the Amihud illiquidity ratio (Amh). Respectively, the Skewness and Kurtosis coefficients, indicate that the displayed range (Qs) and Amihud illiquidity ratio (Amh) are left skewed and pointed, while the zero return (Zr) is right skewed and flattened.

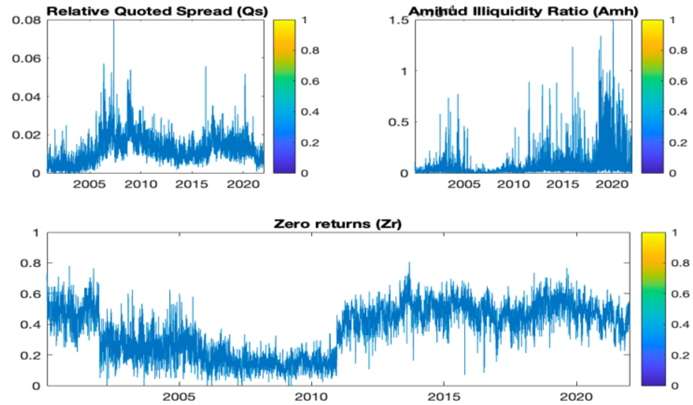

Fig. 4. Evolution of the Liquidity of Shares in the Casablanca Stock Exchange.

TABLE II. Descriptive Statistics

	Qs	Amh	Zr
Number of observations	5268	5475	5475
Average	0,012	6,06E-06	0,362
Median	0,012	2,33E-06	0,397
Std	0,007	1,16E-05	0,167
Coefficient of variation	0,607	1,914	0,461
Kurtosis	6,512	34,004	1,874
Skewness	1,116	4,805	-0,196

B. Data Pre-Processing

The raw liquidity measures are first normalized by the Z-Score method; then denoised by the WT. However, since the choice of the decomposition method for the liquidity measures depends on the characteristics of the data, we first perform a time-frequency analysis of the liquidity measures. From Fig. 5, we can observe that Qs, Amh, and Zr experience low and medium frequency oscillations.

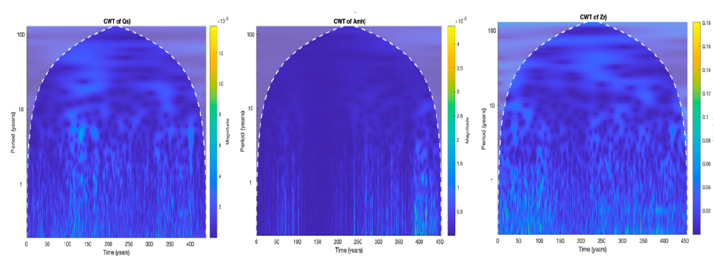

Fig. 5. Magnitudes Scalograms of Liquidity Measures.

Based on the time-frequency analysis while following the approach of [26], we present in III the methods used to decompose the stock liquidity measures.

TABLE III. Multiresolution Analysis Techniques

Input signal	Signal characteristics					Decomposition technique technique
	Low frequency	Average frequency	Increase in frequency	Breaking	Trend	
Qs	Yes	Yes	No	No	No	VMD
Amh	Yes	Yes	No	No	No	VMD
Zr	Yes	Yes	No	Yes	No	MODWT

Since Qs and Amh experience low and medium frequency oscillations and peaks but no trend or break, the most appropriate method to decompose them is VMD. The latter decomposes the original signal into K intrinsic mode function (IMF) components. Fig. 6(a) shows the decomposition of Qs into five MFIs. The first IMFs (IMF1 to IMF3) locate low frequency oscillations, IMF4 locates mid-frequency oscillations and IMF 5 locates peaks. The Qs signal has experienced a few spikes between the year 2000 and 2021; the first ones are related to the effect of the 2007-2008 financial crisis while the last ones are due to COVID-19. Qs is noisy with midrange frequencies and spikes. Therefore, to reconstruct the noisy Qs signal, we neutralize IMF4 and IMF5. Fig. 6(a) highlights the original signal and the reconstructed signal.

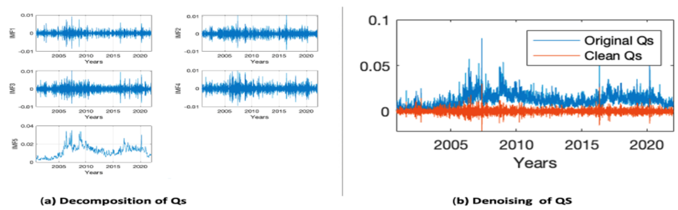

Fig. 6. Decomposition and Denoising of Qs.

Fig. 7(a) clarifies the decomposition of the Amh signal into five MFIs. This decomposition clearly shows that under the effect of COVID-19, the Amh signal experienced mid-frequency oscillations (MFI4) and spikes (MFI5) during 2020 and 2021. This has caused Amh to become very noisy. Thus, in order to denoise it, we reconstruct the signal while neglecting IMF1, IMF2, and IMF3. This reconstruction is shown in Fig. 7(b).

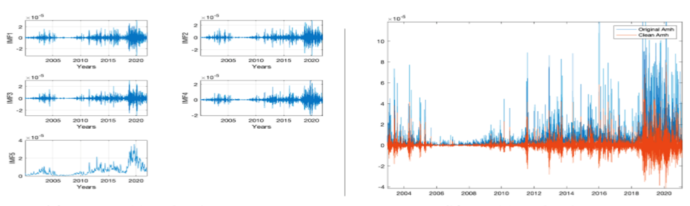

Fig. 7. Decomposition and Denoising of Amh.

Zr is decomposed by the MODWT method, as it experiences low and medium frequency oscillations and a jump in 2011. The MODWT method decomposes the original signal into wavelet coefficients and scaling coefficient. The wavelet coefficients identify high-frequency oscillations, while the scaling coefficients capture trends and jumps in a time series. Fig. 8(a) shows the MODWT algorithm's decomposition of the Zr signal into five levels using the orthogonal Daubechies wavelet, with a level 1. We can observe from Fig. 8(a) that the Zr signal is noisy by medium frequency oscillations. Fig. 8(b) shows the reconstruction of the Zr signal after removing the noise.

1) Experimentation of the WT-CNN-GRU Model:

a) Presentation of the Empirical Results: In subsection III-B1, Qs and Amh were decomposed by the VMD method while Zr was decomposed by the MODWT method.

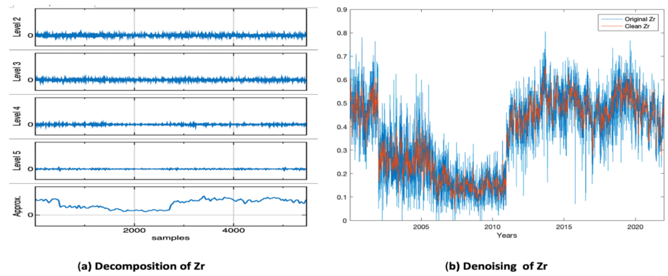

Fig. 8. Decomposition and Denoising of Zr.

After the decomposition and denoising operations, we have 5469 MFIs for Amh, 5268 MFIs for Qs and 5475 approximation coefficients for Zr. The selected MFIs and approximation coefficients constitute inputs for the CNN-GRU model. The data for these inputs are further splitted into 80% training data and 20% test data. Table IV details this split.

TABLE IV. Training and Test Data

Liquidity measurement	Training data		Test data	
	Number of observations	Period	Number of observations	Period
Qs	4215	November 01, 2000 to October 03, 2017	1053	October 04, 2017 to December 31, 2021
Amh	4370	January 04, 2000 to July 19, 2017	1099	July 20, 2017 to December 31, 2021
Zr	4380	January 04, 2000 to July 31, 2017	1095	August 01, 2017 to December 31, 2021

The CNN-GRU model is trained according to the parameters described in Table I. A comparative study between the forecasted and actual values of the three denoised liquidity measures is shown in Fig. 9 to Fig. 11. We can observe that the forecasted values are almost equal to the actual values and that our model is able to make accurate forecasting even under sudden shocks, such as the case of COVID-19 in 2020 and 2021. From Fig. 9 to Fig. 11, we can also notice that the number of outliers in the test errors of the proposed model is very small.

Table V lists the metrics for evaluating the predictive performance of test data. From Table V, we can notice that the proposed model gives very low MAE, MAPE, and RMSE, Theil's U statistics less than 1, and (R^2) close to 1. Our model demonstrates excellent forecasting performance compared to models proposed in previous studies.

The author in [5] proposed a NARX neural network to predict the liquidity of stocks listed on the Casablanca Stock Exchange. Their model is evaluated based on MSE which indicates a value of 0.0083 for Qs and 0.0023 for Zr. The author in [6] estimates that the LSTM model is more efficient than the linear regression method and the multilayer perceptron. The results of this study indicate that the LSTM model exhibits better MSEs. The MSE of the Amihud ratio indicates a value of 0.0169 and 0.0252 on the Ho Chi Min and Hanoi Stock Exchange respectively.

Statistically, our model far exceeds the performance of models postulated by previous studies. Moreover, these models are only tested on a small number of observations. The authors in [5] and [6] only tested their model on the basis of 12 and 110 observations respectively, whereas our model is tested on 1053 observations for Qs, 1099 for Amh, and 1095 for Zr. As a result, the MSEs of earlier models are less reliable. The commonality of previous studies is the use of MSE as an endpoint. However, the latter is more sensitive to outliers. On the contrary, our model is evaluated on the basis of multiple indicators.

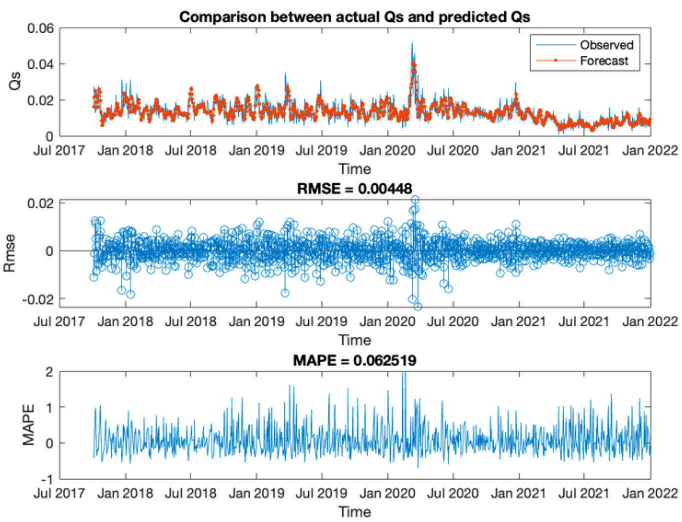

Fig. 9. Comparison between Actual Qs and Qs Predicted by the WT-CNN-GRU Model.

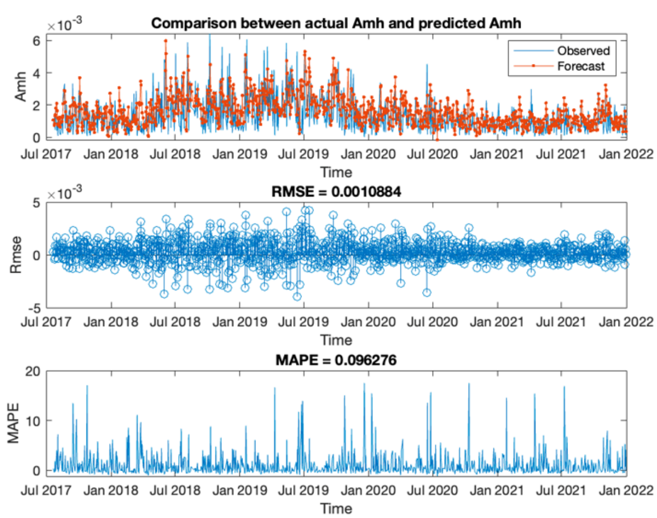

Fig. 10. Comparison between Actual Amh and Amh Predicted by the WT-CNN-GRU Model.

b) Comparison with Similar and Alternative Models: To prove the effectiveness of the proposed model, we compare its performance to similar models, such as WT-CNN-LSTM and WT-CNN-BILSTM. In addition, since the focus of our study is whether the introduction of WT improves the predictive ability of deep neuron networks, the proposed model is compared to alternative models such as CNN-GRU, CNN- LSTM and CNN-BILSTM. While the inputs of the similar models (WT-CNN-LSTM and WT-CNN-BILSTM) are denoised data, those of the alternative models (CNN-LSTM, CNN, BILSTM, and CNN-GRU,) are normalized raw data.

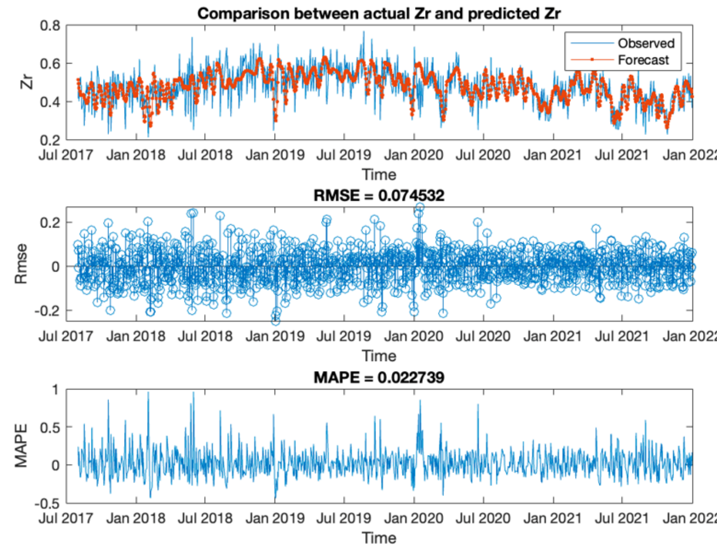

Fig. 11. Comparison between Actual Zr and Zr Predicted by the WT-CNN-GRU Model.

TABLE V. Indicators for Evaluating the Predictive Performance of the CNN-GRU Model

Liquidity measurement	MAE	MAPE	RMSE	Theil's U	R^2
Qs	3.1e-05	0.062	0.004	0.072	0.7379
Amh	0.0002	0.096	0.001	0.037	0.6325
Zr	-0.0025	0.022	0.074	0.079	0,7880

As showcased in Table VI, if we compare the proposed WT-CNN-GRU model with the CNN-GRU model, we can notice that the denoising of the data by WT improved the forecasting results of stock liquidity. At the level of Qs, the WT was able to reduce MAPE and Theil's U by 33% and 57% respectively, and increase R^2 by 48%. In addition, denoising the Amh data contributed to the improvement of MAE, MAPE, RMSE, Theil's U and R^2 by 154%, 84%, 21%, 94% and 40% respectively. Finally, the denoising of Zr data by WT strongly improved MAE, MAPE, RMSE, Theil's U and R^2 by 14%, 72%, 24%, 9% and 97% respectively. These improved results are obtained not only from the denoising of the data, but also from the appropriate choice of the decomposition methods.

In Table VI, we can also notice that the proposed model WT-CNN-GRU also outperforms similar models, such as WT-CNN-LSTM and WT-CNN-BILSTM. The GRU closed cell is therefore the most effective hybrid model when it is compared to the LSTM and BILSTM cells. GRU contains several hidden layers, which can efficiently identify the fluctuation characteristics of liquidity measures. GRU also optimizes the network structure and reduces information redundancy.

TABLE VI. EVALUATION CRITERIA VALUES FOR DIFFERENT FORECASTING MODELS

	Models	MAE	MAPE	RMSE	Theil's U	R^2
Qs	WT-CNN-LSTM	6.71e-05	0.09	0.008	0.1619	0.7174
	WT-CNN-BILSTM	0.0001	0.117	0.004	0.1834	0.4953
	WT-CNN-GRU	3.1e-05	0.062	0.004	0.0723	0.7379
	CNN-LSTM	0.0004	0.07	0.005	0.2143	0.1063
	CNN-BILSTM	8.6e-04	0.095	0.005	0.2056	0.0674
	CNN-GRU	2.3e-05	0.0933	0.004	0.1683	0.4961
Amh	WT-CNN-LSTM	0.0024	0.1479	0.0011	0.3694	0.3639
	WT-CNN-BILSTM	2.7e-05	0.3631	0.0011	0.3840	0.0223
	WT-CNN-GRU	0.0002	0.096	0.0011	0.0371	0.6325
	CNN-LSTM	0.0001	0.2558	0.0015	0.5295	0.0092
	CNN-BILSTM	0.0001	0.2937	0.0010	0.3704	0.0842
	CNN-GRU	-3.7e-04	0.5907	0.0014	0.6524	-0.4526
Zr	WT-CNN-LSTM	-0.0025	0.0226	0.07082	0.0741	0.6104
	WT-CNN-BILSTM	0.0040	0.0190	0.07482	0.0778	0.33970
	WT-CNN-GRU	-0.0025	0.022	0.074	0.079	0,7880
	CNN-LSTM	-0.028	0.0851	0.0849	0.0913	0.2285
	CNN-BILSTM	-0.0074	0.0473	0.0861	0.0906	0.1020
	CNN-GRU	0.0022	0.0793	0.0976	0.0865	0.3992

IV. CONCLUSION

Forecasting equity liquidity is crucial for investors, issuers, and financial market regulators. As a financial series, stock liquidity is non-stationary, non-linear, chaotic, and noisy. Therefore, it is very difficult to accurately forecast inventory liquidity. The purpose of this study is to propose a model capable of effectively predicting inventory liquidity. Inspired by the hybrid research stream in financial time series forecasting, we proposed a WT-CNN-GRU model.

By testing it on all stocks listed on the B.V.C, our model showed excellent forecasting performance compared to models from previous studies and other similar or alternative models, such as WT-CNN-LSTM, WT-CNN -BILSTM, CNN-GRU, CNN-LSTM, and CNN-BILSTM. These improved performances are jointly explained by the neural networks WT, CNN, and the closed cell GRU. Choosing the right method for data decomposition and denoising was key to improving the results. The CNN was used to capture features that the WT did not capture, and they worked together to denoise the data. The GRU cell captured the time dependencies of stock liquidity, which has an advantage over the LSTM or BILSTM cell due to its ability to quickly catch time dependencies while avoiding information redundancy.

This study has three contributions. First, it is considered to be one of the few studies that have addressed the issues of forecasting inventory liquidity. Second, unlike previous studies that used a single data denoising method, we opted for methods consistent with our data. Third, the proposed model can predict stock liquidity even in the face of adverse shocks, such as the COVID-19 pandemic.

This study contributes to the enrichment of the forecast field of the financial series. It can be a useful analytical framework capable of helping investors, issuers, and financial market regulators predict stock liquidity. The model proposed in this study is also extended to predict other financial risks.

However, to verify its robustness, we suggest that future researchers test it in other emerging and developed markets. We also believe that our results can be improved by integrating into the proposed model global and specific exogenous variables for companies listed on stock exchanges.

REFERENCES

[1] T. Chordia, L. Shivakumar, et A. Subrahmanyam, "Liquidity dynamics across small and large firms", Econ. Notes, vol. 33, no 1, p. 111-143, 2004.

[2] R. Ma, H. D. Anderson, et B. R. Marshall, "Market volatility, liquidity shocks, and stock returns: Worldwide evidence", Pac.-Basin Finance J., vol. 49, p. 164-199, 2018.

[3] N. Cakici et A. Zaremba, "Liquidity and the cross-section of international stock returns", J. Bank. Finance, vol. 127, p. 106123, 2021.

[4] T. Zhang et S. H. Lence, "Liquidity and asset pricing: Evidence from the Chinese stock markets", North Am. J. Econ. Finance, vol. 59, p. 101557, 2022.

[5] A. K. Mishra, B. Parikh, et R. W. Spahr, "Stock market liquidity, funding liquidity, financial crises and quantitative easing", Int. Rev. Econ. Finance, vol. 70, p. 456-478, 2020.

[6] A. K. Mishra, B. Parikh, et R. W. Spahr, "Contemporaneous linkages: Funding liquidity and stock market spirals", Int. J. Finance Econ., vol. 26, no 4, p. 5912-5929, 2021.

[7] T. Geithner, "Liquidity and financial markets", 2007.

[8] C. G. Rösch et C. Kaserer, "Reprint of: Market liquidity in the financial crisis: The role of liquidity commonality and flight-to-quality", J. Bank. Finance, vol. 45, p. 152-170, 2014.

[9] J. A. Skjeltorp et B. A. Ødegaard, "When do listed firms pay for market making in their own stock?", Financ. Manag., vol. 44, no 2, p. 241-266, 2015.

[10] U. Shahzad, J. Liu, et F. Luo, "Stock liquidity and corporate trade credit strategies: evidence from China", J. Bus. Econ. Manag., vol. 23, no 1, p. 40-59-40-59, 2022.

[11] S. Stereńczak et J. Kubiak, "Dividend policy and stock liquidity: Lessons from Central and Eastern Europe", Res. Int. Bus. Finance, vol. 62, p. 101727, 2022.

[12] J. Brogaard, D. Li, et Y. Xia, "Stock liquidity and default risk", J. Financ. Econ., vol. 124, no 3, p. 486-502, 2017.

[13] H. H. Trinh, C. P. Nguyen, W. Hao, et U. Wongchoti, "Does stock liquidity affect bankruptcy risk? DID analysis from Vietnam", Pac.-Basin Finance J., vol. 69, p. 101634, 2021.

[14] Y. Amihud, A. Hameed, W. Kang, et H. Zhang, "Stock liquidity and the cost of equity capital in global markets", J. Appl. Corp. Finance, vol. 27, no 4, p. 68-74, 2015.

[15] M. B. Houad et Y. Oubouali, "Forecasts of the liquidity of shares listed on the Casablanca Stock Exchange. Comparison between ARIMA modeling and NARX neural networks"
"Prévisions de la liquidité des actions cotées à la bourse des valeurs de Casablanca. Comparaison entre la modélisation ARIMA et les réseaux de neurones NARX", Rev. Gest. Organ., vol. 10, no 2, p. 83-99, 2018.

[16] P. Q. Khang et al., "Machine learning for liquidity prediction on Vietnamese stock market", Procedia Comput. Sci., vol. 192, p. 3590-3597, 2021.

[17] M. Vogl, P. G. Rötzel, et S. Homes, "Forecasting performance of wavelet neural networks and other neural network topologies: A comparative study based on financial market data sets", Mach. Learn. Appl., vol. 8, p. 100302, 2022.

[18] H. Malik, A. K. Yadav, F. P. G. Márquez, et J. M. Pinar-Pérez, "Novel application of Relief Algorithm in cascaded artificial neural network to predict wind speed for wind power resource assessment in India", Energy Strategy Rev., vol. 41, p. 100864, 2022.

[19] F. Rodríguez, I. Azcárate, J. Vadillo, et A. Galarza, "Forecasting intra-hour solar photovoltaic energy by assembling wavelet based time-frequency analysis with deep learning neural networks", Int. J. Electr. Power Energy Syst., vol. 137, p. 107777, 2022.

[20] C. Song, L. Yao, C. Hua, et Q. Ni, "A novel hybrid model for water quality prediction based on synchrosqueezed wavelet transform technique and improved long short-term memory", J. Hydrol., vol. 603, p. 126879, 2021.

[21] Q. Gu, Y. Chang, N. Xiong, et L. Chen, "Forecasting Nickel futures price based on the empirical wavelet transform and gradient boosting decision trees", Appl. Soft Comput., vol. 109, p. 107472, 2021.

[22] M. Risse, "Combining wavelet decomposition with machine learning to forecast gold returns", Int. J. Forecast., vol. 35, no 2, p. 601-615, 2019.

[23] X. Liu, H. Liu, Q. Guo, et C. Zhang, "Adaptive wavelet transform model for time series data prediction", Soft Comput., vol. 24, no 8, p. 5877-5884, 2020.

[24] S. C. Olhede et A. T. Walden, "Generalized morse wavelets", IEEE Trans. Signal Process., vol. 50, no 11, p. 2661-2670, 2002.

[25] J. Chung, C. Gulcehre, K. Cho, et Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling", ArXiv Prepr. ArXiv14123555, 2014.

[26] J. Gilles, "Empirical wavelet transform", IEEE Trans. Signal Process., vol. 61, no 16, p. 3999-4010, 2013.