

High-Frequency Equity Index Futures Trading Using Recurrent Reinforcement Learning with Candlesticks

Patrick Gabrielsson and Ulf Johansson

Department of Information Technology

University of Borås

Sweden

Email : {patrick.gabrielsson, ulf.johansson}@hb.se

Abstract—In 1997, Moody and Wu presented recurrent reinforcement learning (RRL) as a viable machine learning method within algorithmic trading. Subsequent research has shown a degree of controversy with regards to the benefits of incorporating technical indicators in the recurrent reinforcement learning framework. In 1991, Nison introduced Japanese candlesticks to the global research community as an alternative to employing traditional indicators within the technical analysis of financial time series. The research, accumulated over the past two and a half decades, contains conflicting results with regards to the utility of using Japanese candlestick patterns in exploiting inefficiencies in financial time series. In this paper, we combine features based on Japanese candlesticks with recurrent reinforcement learning to produce a high-frequency algorithmic trading system for the E-mini S&P 500 index futures market. Our empirical study shows a statistically significant increase in both return and Sharpe ratio compared to relevant benchmarks, suggesting the existence of exploitable spatio-temporal structure in Japanese candlestick patterns and the ability of recurrent reinforcement learning to detect and take advantage of this structure in a high-frequency equity index futures trading environment.

I. INTRODUCTION

The computational power offered by contemporary computer systems, coupled with the abundance of available financial data, has led to a proliferation of research based on data-driven machine learning methods within the algorithmic trading literature. Since the inception of electronic exchanges over three decades ago, a variety of machine learning paradigms have shown their benefits in forecasting financial time series, in reducing trade execution costs and in creating profitable, risk-adjusted trading strategies.

In 1997, Moody and Wu introduced a policy-based recurrent reinforcement learning (RRL) model to create risk-adjusted, cost-sensitive trading strategies, using a differential Sharpe ratio as an objective function for optimization purposes, and a set of lagged returns together with a recurrent trade position signal as the basic inputs to the trading system [1]. Since then, numerous studies have contributed to an increased understanding of RRL-based trading systems, but with controversial results with respect to the utility of incorporating technical indicators as additional inputs. Dempster and Leemans found no increase in performance when adding additional technical indicators as inputs to the RRL models as compared to using only lagged returns [2].

Japanese Candlesticks, a technical analysis technique, were invented in 18th century Japan, but were first introduced to the global research community in 1991 when Nison published a book in English on the subject [3]. Research into the utility of using Japanese Candlesticks in predicting market moves shows conflicting results, where numerous studies find no useful structure in Japanese Candlestick patterns [4-9].

The datasets used in RRL and Japanese Candlestick research are mostly based on daily market prices. Furthermore, existing intraday studies in RRL use Forex market data [2,10,11]. In this paper, we supply the basic RRL model with additional inputs, in the form of derived Japanese Candlestick features, based on high-frequency data from the E-mini S&P 500 equity index futures market. Our empirical results show a statistically significant increase in both return and Sharpe ratio, in a frictionless setting, when benchmarked against the basic RRL model, a random trading strategy and a simple buy-and-hold trading strategy.

This paper is structured as follows. Section II introduces RRL and Japanese Candlesticks, with a review of related work in Section III. Section IV presents the methodology adopted in this paper, followed by the experimental approach used in Section V. The results from the experiments are presented in Section VI and conclusive remarks are given in Section VII.

II. BACKGROUND

A. Recurrent Reinforcement Learning

The basic recurrent reinforcement learning (RRL) algorithm was introduced by Moody and Wu in 1997 [1], with more detailed presentations in later publications [11,12]. In its most basic form, the RRL model is equivalent to a single-layer, single-unit recurrent neural network with a hyperbolic tangent function as its activation function

$$F_t = \tanh(\sum_{i=0}^m w_i r_{t-i} + w_{m+1} F_{t-1} + w_{m+2} v) \quad (1)$$

where F_t is the output [-1,+1] of the network at time t , F_{t-1} is the output of the network at the previous time step $t-1$, v is the threshold bias and $w_0 \dots w_{m+2}$ are the network weights. The terms $r_{t,i}$ refer to the non-recurrent inputs to the network in the form of lagged returns, where the return r_t at time t is defined in terms of the market prices p_t and p_{t-1} at time t and $t-1$ respectively to create either a simple market return (2) or a relative market return (3)

$$r_t = p_t - p_{t-1} \quad (2)$$

$$r_t = \frac{p_t - p_{t-1}}{p_{t-1}} \quad (3)$$

The only recurrent input to the network is the network's previous output F_{t-1} , corresponding to the trading decision made at time $t-1$. In the literature, this signal is usually post-processed to create a binary $\{+1, -1\}$ long/short trading strategy or a ternary $\{+1, 0, -1\}$ long/neutral/short trading strategy. The excess trading return R_t at time t for a binary trading strategy can be expressed as

$$R_t = r_t^f + \text{sgn}(F_{t-1})(r_t - r_t^f) - \delta |\text{sgn}(F_t) - \text{sgn}(F_{t-1})| \quad (4)$$

where $\text{sgn}(x)$ is the *signum* function, returning +1 for $x \geq 0$ and -1 otherwise, r_t is the market return at time t , r_t^f is the return of a risk-free asset at time t and δ is the transaction cost for a trade. If the r_t^f term is dropped from (4) and if multiple shares μ are traded, a general expression for a single-asset return is obtained as

$$R_t = \mu (\text{sgn}(F_{t-1})r_t - \delta |\text{sgn}(F_t) - \text{sgn}(F_{t-1})|) \quad (5)$$

Supervised learning using a neural network is usually accomplished by optimizing an appropriate objective function, such as minimizing the sum of squared errors between the network's actual- and target output. This scheme requires datasets labeled with the true target values in order to train the network. To avoid a time-consuming labeling procedure, a reinforcement learning scheme can be used instead, where the concept of a temporal difference is employed. In its simplest form, temporal difference learning simply uses the network's output at time t as the target value for the network at time $t-1$. Moody and Wu used the Sharpe ratio (SR) as their objective function

$$SR = \frac{\text{mean}(R_t)}{\text{std}(R_t)} \quad (7)$$

i.e. the ratio between the average trading return and its standard deviation. An online version can be expressed using an exponential moving average of the SR

$$EMASR_t = \frac{A_t}{\kappa_\eta \sqrt{B_t - A_t^2}} \quad (8)$$

where

$$A_t = A_{t-1} + \eta(R_t - A_{t-1})$$

$$B_t = B_{t-1} + \eta(R_t^2 - B_{t-1})$$

$$\kappa_\eta = \sqrt{\frac{1-\eta}{1-\eta^2}}$$

and η is the adaption rate. The differential Sharpe ratio (DSR) is then obtained through a first-order Taylor expansion in the adaption rate η in (8), yielding the expression

$$DSR_t = \left. \frac{dEMASR_t}{d\eta} \right|_{\eta=0} = \frac{B_{t-1}(R_t - A_{t-1}) - \frac{1}{2}A_{t-1}(R_t^2 - B_{t-1})}{(B_{t-1} - A_{t-1}^2)^{\frac{3}{2}}} \quad (9)$$

In order to maximize the DSR, gradient ascent (10) can be used to update a specific network weight incrementally

$$w_t = w_{t-1} + \rho \Delta w_t \quad (10)$$

where Δw_t is the weight delta and hyper-parameter ρ is the network's learning rate. The weight delta is obtained by using the chain rule to express the derivative of the DSR with respect to a specific network weight

$$\Delta w_t = \frac{dD_t}{dw_t} = \frac{dD_t}{dR_t} \left(\frac{dR_t}{dF_t} \frac{dF_t}{dw_t} + \frac{dR_t}{dF_{t-1}} \frac{dF_{t-1}}{dw_{t-1}} \right) \quad (11)$$

where

$$\frac{dD_t}{dR_t} = \frac{B_{t-1} - A_{t-1}R_t}{(B_{t-1} - A_{t-1}^2)^{\frac{3}{2}}}$$

$$\frac{dR_t}{dF_t} = -\mu \cdot \sigma \cdot \text{sgn}(F_t - F_{t-1})$$

$$\frac{dR_t}{dF_{t-1}} = \mu \cdot (r_t + \sigma \cdot \text{sgn}(F_t - F_{t-1}))$$

This is equivalent to the standard backpropagation algorithm for updating the weights in a non-recurrent network. In order to calculate the derivative of the recurrent trading signal F_t with respect to the network weights at time t , the exact solution requires all derivatives from time 0 to t to be computed. Although, using a stochastic update procedure similar to backpropagation through time (BPTT) [13], the derivative for the recurrent trading signal with respect to the network weights at time t can be approximated by using the value of the derivative at the previous time step

$$\frac{dF_t}{dw_t} \approx \frac{\partial F_t}{\partial w_t} + \frac{\partial F_t}{\partial F_{t-1}} \frac{dF_{t-1}}{dw_{t-1}} \quad (12)$$

All terms in (11) and (12) can be calculated recursively in terms of their values in the previous time step. The machine learning model presented above constitutes the basic RRL model presented by Moody and Wu [1,11-12]. The profit, i.e. the cumulative trading return, can be calculated as an absolute profit (13) when using the simple return (2) or as a percentage (14) when using the relative return (3)

$$P_T = \sum_{t=0}^T R_t \quad (13)$$

$$P_T = \prod_{t=0}^T (1 + R_t) - 1 \quad (14)$$

B. Japanese Candlesticks

Japanese Candlesticks were invented in Japan in the 18th century in order to analyze the contemporary rice market, but were only introduced to the global research community in 1991 [3]. A candlestick is made up of four prices; the open-, high-, low- and closing price (often abbreviated as OHLC) during a specific time interval. The open price corresponds to the price level at the beginning of the time interval, whereas the closing price corresponds to the price level at the end of the time interval. The high- and low prices are the highest and the lowest price levels that occur during the time interval.

Candlesticks are often visualized on technical charts according to Fig. 1 and resemble the box and whiskers in a box-and-whiskers plot. The box, or body, of a candlestick is bounded above and below by the opening- and closing price, whereas the peak of the upper whisker constitutes the high price and the peak of the lower whisker constitutes the low price during the time interval. The range between the opening and closing price is called the candlestick's real body. If the closing price is higher than the opening price (a bullish interval), the real body is visualized using a transparent- or white color (left candlestick in Fig. 1). Conversely, if the closing price is lower than the opening price (a bearish interval), the real body is visualized using an opaque- or black color (right candlestick in Fig. 1). The candlestick's upper whisker is called the upper shadow and the lower whisker is called the lower shadow.

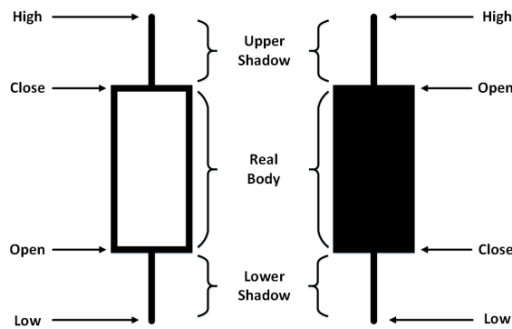


Fig. 1. Japanese candlesticks; bullish (left), bearish (right).

Although various information can be extracted from a single candlestick, patterns consisting of two or three consecutive candlesticks are commonly used in the chart analysis literature [3]. By analyzing candlestick patterns, information about trends and reversals can be extracted. Fig. 2 shows examples of three candlestick patterns; the single candlestick pattern called a Doji (left), the two-candlestick pattern called a Bullish Harami (middle) and the three-candlestick pattern called an Evening Star (right). As an example, an Evening Star pattern following an uptrend is indicative of a trend-reversal.

Candlesticks are commonly used in conjunction with trend indicators, such as a 5-minute simple moving average.

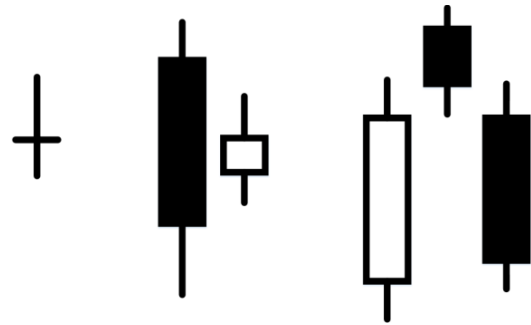


Fig. 2. Candlestick patterns; Doji (left), Bullish Harami (middle), Evening Star (right).

III. RELATED WORK

Following the introduction of the basic RRL algorithm [1], Moody and Wu experimented with creating single-asset trading strategies, asset allocation strategies and rebalancing portfolios in the intraday Forex market (USD/GBP), daily equity index (S&P 500) market and the US treasury bills market, yielding positive returns and Sharpe ratios in the presence of transaction costs [11,12]. They also experimented with various objective functions, besides the basic differential Sharpe ratio, where a modified Sterling ratio was shown to be beneficial. The basic RRL model was also benchmarked with a trading strategy based on the off-policy, online Q-learning reinforcement learning algorithm, where RRL outperformed its benchmark.

Gold created single- and multilayer versions of the basic RRL model using intraday data for 25 Forex markets, with profitable outcomes in all but a few of the markets [10]. One important conclusion from this study was that single-layer RRL models outperformed multi-layer RRL models.

Dempster and Leemans created an automated trading system [2] consisting of three layers; the RRL machine learning layer, a risk management layer and an optimization layer. Intraday data for the EUR/USD currency pair were used, yielding superior results for their three-layer models as compared to the basic RRL model. One important result from this study was that no increase in performance was observed when adding additional technical indicators as inputs.

Bertoluzzo and Corazza used RRL to create ternary trading strategies (long, short, neutral) for 9 daily equity indices. Two major differences from previous studies included the use of a stop-loss risk management strategy and the use of *the reciprocal of the returns weighted direction symmetry index* as their objective function [14]. Profitable results were reported for all but one equity index.

Gorse compared RRL with genetic programming (GP), using daily, weekly and monthly opening prices of the S&P 500 stock index [15]. The study showed that RRL worked better on higher frequency data, whereas GP performed better on monthly data. Once again, experiments with RRL based on multi-layer networks did not improve performance.

Maringer and Ramtohl introduced threshold-based regime-switching models to RRL in 2010, using a GARCH-based volatility measure as the transition variable [16]. Their

experiments using data from 4 equity index markets showed an increase in performance as compared to the basic RRL model. In subsequent studies, experiments were conducted with different transition variables [17] and smooth regime-switching RRL models for creating portfolio- and single-asset trading strategies [18], using daily data for various equity indices. The studies showed promising results for volume-based transition variables and smooth regime-switching RRL models.

Maringer and Zhang further investigated the transition variable selection problem by comparing the performance of different regime-switching RRL models using four different transition variables [19]. Daily closing prices and trading volume for 238 S&P 500 index constituent stocks were used and showed that the volume-based transition variable yielded the highest Sharpe ratios. Following this, an automated regime-switching transition function model was introduced, which outperformed earlier models on daily data for 20 SPI stocks, where the price-to-earnings ratio transition variable produced the highest Sharpe ratios [20]. Experiments with two different ensemble-based schemes for updating the parameters of the RRL models were also conducted using daily prices for all 500 constituents of the S&P 500 index, resulting in higher Sharpe ratios [21].

Zhang and Maringer also conducted experiments using the genetic algorithm (GA) to select an optimal subset of input parameters, consisting of lagged returns, technical indicators, fundamental indicators and econometric indicators, to the RRL model. Daily data for 238 [22] and 180 [23] constituent stocks of the S&P 500 index were used. Their results showed that a GA-optimized model could outperform the basic RRL model (using only lagged returns of closing prices as inputs), suggesting that the joint information found in a mixture of technical-, fundamental- and econometric indicators is beneficial for RRL model performance.

Japanese Candlestick patterns have successfully been used in determining trends and predicting trend reversals [24-37], but there also exists numerous studies that do not find any evidence of usable structure in Japanese Candlestick patterns, using intraday- and daily data for equity and equity index markets [4-9]. Proponents of Japanese Candlesticks have conducted studies using data for the US equity and equity index markets [24,27,32,36], European equity index markets [33,37], Korean equity markets [25], Taiwanese equity markets [26,28-30,35], Philippine equity markets [31] and Forex markets [34-35]. Three studies used intraday data [32,34,37] and two studies incorporated market volume in addition to the open-, high- low- and closing prices [28,32]. All studies resulted in a positive outcome with respect to the utility of Japanese Candlesticks. Most studies conduct a statistical evaluation of Japanese Candlestick patterns, but in [25] an expert system was developed and in [27] a tapped-delay, feed-forward neural network was used to create a trading model. In [31], candlestick features were used in conjunction with fuzzy decision trees and in [32,34] rule-based trading systems were developed. More recently, computer vision techniques have been used to create a content-based image retrieval system (CBIR) based on 7 wavelet patterns derived from candlestick charts [36] and to decrease trade execution costs [37].

IV. METHOD

The related work reflects the controversy with regards to useable structure in Japanese Candlestick patterns and in the ability of the RRL model to benefit from additional inputs besides lagged returns, based on closing prices. Furthermore, most studies found in the literature regarding RRL and Japanese Candlesticks are based on daily equity and equity index market prices. The focus of this study is therefore in determining if there exists useful structure in Japanese Candlestick patterns for intraday equity index futures markets and if RRL can benefit from additional lagged inputs of features derived from Japanese Candlesticks.

A. Data Acquisition and Feature Extraction

The E-mini S&P 500 equity index futures contract (ES), traded on the Chicago Mercantile Exchange, was chosen for the research work. The contract has a quarterly expiry schedule with a minimum tick size of 0.25 index points, where each tick is worth \$12.50 per tick and contract. Two months worth (5th July – 2nd September 2011) of tick data was downloaded from Slickcharts [38]. The data was aggregated into 1-minute bars, each including the open-, high-, low- and close prices. Missing data points, i.e. missing tick data for one or more minutes, was handled by using the same price levels as the previously existing aggregated data point. Four non-parametric features, similar to [31], were extracted from a single candlestick as shown in Fig 3.

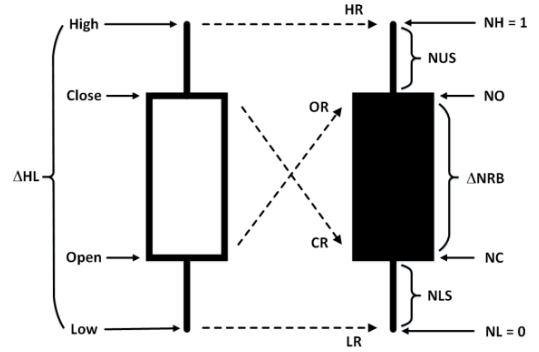


Fig. 3. Derived candlesticks features.

The first feature ΔHL , uses the high and low price at time t , and is the candlestick's normalized range

$$\Delta HL_t = \frac{High_t - Low_t}{Low_t} \quad (15)$$

The second feature NUS , is the normalized upper shadow (16) and is computed by subtracting the maximum of the normalized open- (17) and closing price (18) from 1.0

$$NUS_t = 1 - \max(NO_t, NC_t) \quad (16)$$

$$NO_t = \frac{Open_t - Low_t}{High_t - Low_t} \quad (17)$$

$$NC_t = \frac{Close_t - Low_t}{High_t - Low_t} \quad (18)$$

The third feature NLS , is the normalized lower shadow (19) and is computed by choosing the minimum of the normalized open- (17) and closing price (18)

$$NLS_t = \min(NC_t, NO_t) \quad (19)$$

The fourth feature ΔNRB , is the normalized real body (20) and is computed by subtracting the normalized open price (17) from the normalized closing price (18)

$$\Delta NRB_t = NC_t - NO_t \quad (20)$$

In addition to these three single candlestick features, three relative candlestick features were derived; the relative closing price return (21), the relative high price return (22) and the relative low price return (23)

$$CR_t = \frac{Close_t - Close_{t-1}}{Close_{t-1}} \quad (21)$$

$$HR_t = \frac{High_t - High_{t-1}}{High_{t-1}} \quad (22)$$

$$LR_t = \frac{Low_t - Low_{t-1}}{Low_{t-1}} \quad (23)$$

The relative closing price return (21) is identical to (3). These 7 features were lagged 1-4 time steps to create a total of 35 inputs in addition to the recurrent trading signal.

B. Dataset Partitioning

The dataset was split into training-, validation- and test datasets using a rolling window scheme (see Fig. 4). The initial 14000 data points (10 trading days) were used to train 100 RRL models using the derived candlestick features as lagged inputs. During the training procedure, a validation set consisting of the 1400 data points (1 trading day) immediately following the training dataset was used together with an early stopping criterion to prevent overfitting the RRL models. The early stopping criterion halted training when the Sharpe ratio on the validation set started to decrease. Finally, the trained RRL model was used to trade the test set consisting of the 1400 data points (1 trading day) immediately following the validation set. All three datasets were then rolled forward by 1400 data points (1 trading day) and the procedure above was repeated. By rolling the three windows over the entire dataset, 31 consecutive test datasets were obtained (the rightmost 31 columns in Fig. 4).

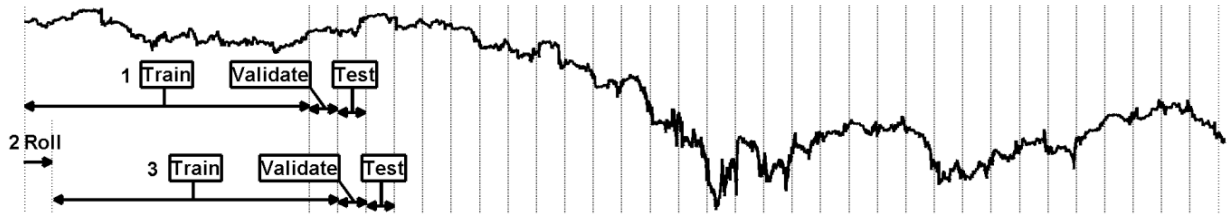


Fig. 4. Dataset partitioning; 10-day training window, 1-day validation window, 1-day test window (with 1-day rolling window scheme).

C. Objective Function and Performance Measures

The differential Sharpe ratio (9) was used as the RRL algorithm's objective function. As performance measures, the Sharpe ratio (7) and the return percentage (14) were used.

D. Benchmarks

In order to validate the performance of our Candlestick RRL (CR) model, we benchmarked its performance against the basic RRL (BR) model, which only uses lagged closing price returns as its non-recurrent inputs (21), a random "zero intelligence" trading strategy (ZI) and a buy-and-hold (BH) trading strategy. All models produced binary $\{+1, -1\}$ long/short trading signals. The random strategy randomly selected between a long position (+1) or a short (-1) position during each time step, whereas the buy-and-hold strategy simply held a long position (+1) during the entire trading period.

V. EXPERIMENTS

The entire dataset was used according to the rolling widow scheme. Instead of selecting the best performing model from the validation set, we evaluated the performance of all models on the test dataset. With the rolling window scheme, we ended up testing the performance of our models on 31 trading days. All models were created with the following hyper-parameter settings; the number of contracts traded $\mu = 1$, trading costs $\delta = 0$ in (5), the learning rate $\rho = 0.01$ in (10) and the adaption rate parameter $\eta = 0.05$ in (8).

We also repeated the above experiments using trading costs. Interest fees (i.e. over-night carry costs) are not applicable in high-frequency trading since all positions are closed-out before the end of the trading day. Furthermore, taxes are not accounted for in this study. The complex effects of slippage were relaxed, using the simplified assumption that when trading one-lots of the front-month E-mini S&P 500 futures contract, there would be enough liquidity in the market for a market order of a single contract to cross immediately without moving the market. Furthermore, the market spread was assumed to be 1 tick (0.25 index points). Commission fees and transaction fees are highly dependent on which brokerage is used, exchange membership and other circumstantial factors. We made the simple assumption that these fees were negligible compared to the market spread cost associated with a market order. Therefore, we included the market spread (0.25 index points) as our only cost per contract and round trip, i.e. when switching a position between long and short, we included a transaction cost of $2 \times 0.25 = 0.50$ index points.

VI. RESULTS

The performance measures for the four models over the 31 test datasets are tabulated in Table I (Sharpe ratio) and Table II (return) respectively. The Shape ratios, calculated using (7), are displayed as 10^{-2} and the returns, calculated using (14), are given as a percentage. Each value in both tables are medians for 100 models, where values not surrounded by parentheses are for the frictionless case (no transaction costs) and values within parentheses reflect the performance with transaction costs included. The four different types of models are abbreviated as BH (the buy-and-hold trader), ZI (the "zero intelligence" random trader), BR (the basic RRL trader using only lagged closing price returns as non-recurrent inputs) and CR (the RRL trader using lagged candlestick features as non-recurrent inputs).

TABLE I
Sharpe Ratios (10^{-2})

#	BH	ZI	BR	CR
1	3.04 (3.00)	0.03 (-48.10)	9.85 (-0.96)	14.22 (2.61)
2	0.12 (0.07)	-0.17 (-61.43)	14.58 (0.01)	18.82 (-0.18)
3	-0.70 (-0.73)	0.13 (-39.73)	4.47 (-0.63)	8.79 (-0.72)
4	-0.76 (-0.81)	-0.53 (-55.64)	9.85 (-0.81)	15.50 (0.08)
5	-5.19 (-5.24)	-0.88 (-56.81)	9.05 (-5.24)	15.42 (4.94)
6	-0.16 (-0.20)	-0.13 (-50.85)	8.20 (-0.28)	15.74 (-0.20)
7	-1.19 (-1.23)	-0.27 (-40.78)	4.16 (-1.29)	8.56 (-1.23)
8	-1.67 (-1.69)	0.14 (-36.69)	5.35 (-0.74)	13.39 (-1.69)
9	-2.86 (-2.90)	0.22 (-44.82)	10.11 (2.07)	14.29 (2.00)
10	-1.30 (-1.33)	-0.18 (-39.43)	7.78 (-1.26)	10.45 (0.80)
11	-3.16 (-3.19)	0.13 (-33.10)	7.06 (1.81)	9.81 (-3.19)
12	-0.73 (-0.74)	-0.21 (-18.99)	3.60 (-0.76)	7.90 (-0.74)
13	-3.34 (-3.35)	0.05 (-17.12)	4.99 (-2.93)	6.60 (3.15)
14	0.15 (0.14)	-0.21 (-18.49)	0.30 (-1.25)	1.40 (-0.16)
15	1.20 (1.89)	0.29 (-18.65)	3.96 (-1.61)	6.63 (-1.21)
16	0.90 (0.89)	0.55 (-20.59)	4.92 (-1.59)	9.03 (0.17)
17	1.72 (1.71)	-0.26 (-27.82)	5.61 (-2.53)	6.61 (-1.69)
18	1.86 (1.83)	-0.32 (-40.47)	0.50 (-1.89)	6.68 (-1.89)
19	-1.38 (-1.41)	-0.13 (-39.46)	5.60 (-0.57)	6.01 (0.68)
20	0.92 (0.89)	-0.07 (-42.15)	4.00 (-0.96)	6.55 (0.56)
21	-4.67 (-4.69)	0.23 (-32.79)	6.96 (1.50)	9.58 (-4.69)
22	-1.02 (-1.03)	-0.35 (-24.35)	4.73 (-0.52)	6.07 (0.76)
23	2.11 (2.09)	-0.28 (-36.11)	5.58 (-2.14)	6.32 (1.39)
24	1.89 (1.86)	-0.10 (-35.42)	5.72 (-1.75)	7.47 (-1.92)
25	-0.18 (-0.21)	0.02 (-34.57)	4.24 (-0.95)	8.91 (-0.08)
26	1.74 (1.72)	0.13 (-31.20)	6.54 (-0.79)	15.53 (-1.76)
27	4.16 (4.12)	0.24 (-45.68)	9.06 (-3.86)	16.65 (-0.63)
28	0.06 (0.02)	0.18 (-48.36)	9.35 (-0.16)	13.65 (-0.10)
29	3.05 (3.01)	0.21 (-47.10)	11.94 (2.11)	13.14 (2.74)
30	-0.59 (-0.62)	-0.59 (-36.77)	9.00 (-0.23)	15.23 (-0.28)
31	-3.50 (-3.53)	0.19 (-34.55)	9.06 (0.89)	11.68 (3.16)

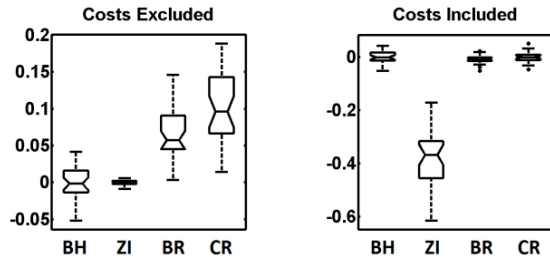


Fig. 5. Overall Sharpe ratios for 100 models and 31 trading days.

Fig. 5 and Fig. 6 below the two tables contain box plots for the 100 models of each model type over all 31 test datasets, i.e. each box plot is for 100 models run over 31 trading days. The box plots were created using Matlab's *boxplot* command with the "notch" feature turned on, i.e. two medians are significantly different at the 5% level if their notch intervals do not overlap. In Fig. 5, the left box plot shows the Sharpe ratios for the frictionless case (no trading costs) and the right box plot depicts the Sharpe ratios with transaction costs included. Likewise, in Fig.6, the left box plot displays the returns with trading costs excluded and the right box plot visualizes the returns with transaction costs included. The box plots convey the general performance between the four different model types across the 31 trading days and the tables compares their performance for each trading day individually.

TABLE II
Returns (%)

#	BH	ZI	BR	CR
1	1.45 (1.44)	0.01 (-23.08)	4.78 (-0.45)	6.94 (1.25)
2	0.04 (0.02)	-0.06 (-23.08)	5.00 (0.00)	6.45 (-0.06)
3	-0.43 (-0.45)	0.06 (-22.79)	2.68 (-0.39)	5.34 (-0.44)
4	-0.30 (-0.32)	-0.21 (-23.05)	3.87 (-0.32)	6.12 (0.03)
5	-1.99 (-2.01)	-0.34 (-23.36)	3.54 (-2.01)	6.07 (1.92)
6	-0.08 (-0.10)	-0.06 (-23.31)	2.43 (-0.13)	6.26 (0.04)
7	-0.75 (-0.77)	-0.20 (-24.05)	0.57 (-0.80)	3.37 (-0.77)
8	-1.14 (-1.16)	0.08 (-23.53)	1.37 (-0.52)	6.88 (0.83)
9	-1.58 (-1.60)	0.11 (-24.29)	3.93 (1.15)	6.47 (1.57)
10	-0.85 (-0.87)	-0.13 (-24.27)	2.94 (-0.83)	4.69 (0.81)
11	-2.58 (-2.60)	0.08 (-24.97)	2.93 (1.47)	4.87 (2.58)
12	-1.17 (-1.19)	-0.40 (-25.30)	1.50 (-1.22)	4.33 (0.92)
13	-5.61 (-5.63)	-0.03 (-25.61)	1.71 (-4.94)	3.70 (5.70)
14	0.16 (0.13)	-0.45 (-27.28)	0.16 (-2.15)	0.76 (-0.38)
15	1.84 (1.82)	0.38 (-26.18)	0.01 (-2.63)	3.83 (-2.05)
16	1.27 (1.24)	0.73 (-26.84)	1.49 (-2.40)	4.50 (-1.43)
17	1.75 (1.72)	-0.30 (-25.74)	1.97 (-2.56)	2.24 (-1.81)
18	1.23 (1.21)	-0.23 (-25.66)	0.23 (-1.27)	2.05 (-1.56)
19	-0.96 (-0.98)	-0.10 (-25.55)	1.55 (-0.41)	2.08 (0.92)
20	0.57 (0.55)	-0.06 (-25.51)	0.47 (-0.62)	2.79 (-0.62)
21	-3.97 (-4.00)	0.18 (-25.84)	2.96 (1.26)	3.94 (4.06)
22	-1.29 (-1.31)	-0.48 (-26.44)	1.13 (-0.68)	1.93 (1.18)
23	1.69 (1.66)	-0.25 (-26.75)	1.67 (-1.73)	1.97 (-1.73)
24	1.51 (1.49)	-0.10 (-26.35)	1.80 (-1.42)	2.90 (-1.76)
25	-0.17 (-0.19)	-0.01 (-25.93)	0.60 (-0.79)	4.09 (0.10)
26	1.58 (1.56)	0.09 (-26.06)	2.71 (-0.75)	10.74 (1.56)
27	2.41 (2.38)	0.13 (-25.60)	4.02 (-2.20)	8.06 (-2.48)
28	0.02 (0.00)	0.08 (-25.00)	3.82 (-0.09)	5.80 (-0.06)
29	1.66 (1.64)	0.10 (-25.20)	4.89 (1.16)	5.75 (1.39)
30	-0.45 (-0.47)	-0.45 (-25.11)	4.91 (-0.19)	8.70 (-0.47)
31	-2.77 (-2.79)	0.13 (-25.41)	5.33 (0.69)	6.33 (2.36)

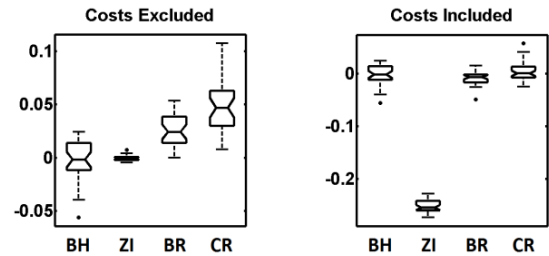


Fig. 6. Overall returns for 100 models and 31 trading days.

Table I shows the candlestick-based RRL model (CR) outperforming the basic RRL model (BR) during all 31 trading days, with positive Sharpe ratios during each trading day, in the case where no transaction costs are included. In fact, this result is statistically significant at the 5% level with regards to the median Sharpe ratios during each trading day. Equivalently, by inspecting Table II, one can observe that the candlestick-based RRL model (CR) has a higher return than the basic RRL model (BR) during all 31 trading days, with positive returns each trading day, in the case where no transaction costs are included. This result is also significant each trading day with regards to the median returns. This suggests that there exists exploitable structure in features derived from Japanese candlesticks, using intraday data for the E-mini S&P 500 futures market, and that the RRL algorithm is able to detect and take advantage of this spatio-temporal information. Both the RRL-based models are also significantly better than the zero intelligence (ZI) model and the buy-and-hold (BH) model, where, in general, the buy-and-hold model outperforms the zero intelligence model in a bullish market and the zero intelligence model outperforms the buy-and-hold model in a bearish market.

The left box plots in Fig. 5 and Fig. 6 show the overall performance of all 100 models over all trading days in the frictionless setting. The buy-and-hold (BH) model's median is centered around the zero line in both plots. The reason for this can be inferred by examining the return curves for each trading day, where the buy-and-hold model is profitable in a rising (bullish) market and incurs a loss in a declining (bearish) market. The zero intelligence trader's (ZI) medians are also centered around zero, but have a tighter inter-quartile range compared to the buy-and-hold model. The box plots show the basic RRL trader (BR) significantly outperforming both the buy-and-hold (BH) trader and the zero intelligence trader (ZI) over all 100 models and all 31 trading days. Furthermore, the candlestick-based RRL (CR) trader significantly outperforms all other three models.

To get a general idea of the trading profiles of the four models, Fig. 7 shows the average equity curves (returns) of the 100 models for the first 5 trading days in the frictionless case (top) and with trading costs included (bottom). The general pattern in the top row, shows rising equity curves with little drawdown for the two RRL-based traders, fluctuations around the zero line for the zero intelligence trader and essentially the market return for the buy-and-hold trader.

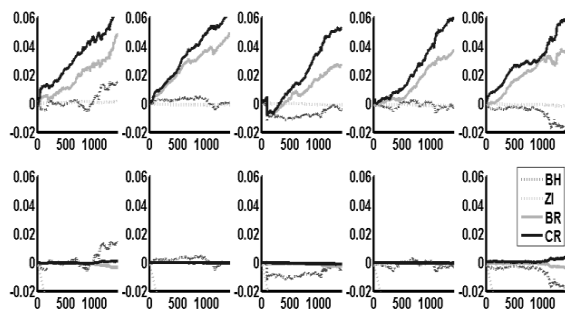


Fig. 7. Returns for days 1-5; without costs (top), with costs (bottom).

When transaction costs are accounted for, the picture changes dramatically. For the buy-and-hold (BH) trader, the values in Table I and Table II are essentially unchanged, since only one round trip is made per trading day. On the other hand, the zero intelligence (ZI) trader fails miserably with large negative Sharpe ratios and returns, which is reflected in the right box plots in Fig. 5 and Fig 6, and by a barely discernible sharply decreasing curve in the bottom row in Fig. 7. For the two RRL-based traders, the Sharpe ratios and returns are mostly negative, with similar inter-quartile ranges as the buy-and-hold trader, as seen in Fig. 5 and Fig. 6. Furthermore, the candlestick-based RRL (CR) trader performs slightly better than the basic RRL (BR) trader, although the difference in median Sharpe ratios and returns between the two RRL-based traders and the buy-and-hold trader are not significant. Any profits have essentially been wiped-out by the market spread cost when trading market orders.

VII. CONCLUSIONS

In this paper, we have reviewed prior research on recurrent reinforcement learning (RRL) in algorithmic trading and found conflicting results with regards to the benefit of incorporating technical indicators as additional inputs to the RRL algorithm. We also reviewed existing research regarding the utility of Japanese candlestick patterns in financial time series analysis, where a divide between proponents and opponents exists.

We hypothesized that lagged features, derived from Japanese candlesticks, using intraday data for the E-mini S&P 500 futures market, contain useful spatio-temporal structure and that the RRL algorithm can be used to detect and take advantage of this information in creating a binary trading strategy. To test our hypothesis, we trained 100 models using the differential Sharpe ratio as our objective function and tested the models' performance on 31 trading days worth of data, using a rolling window approach. Our models were benchmarked against a buy-and-hold model, a zero intelligence model and a basic RRL model, in both a frictionless environment (no trading costs) and with trading costs included.

Our results showed significantly higher median Sharpe ratios and returns for the candlestick-based RRL model as compared to the three benchmark models, in the frictionless setting, suggesting that there indeed does exist exploitable structure in Japanese candlestick patterns for intraday equity index futures trading and that the RRL algorithm can take advantage of this information. Although, the current candlestick-based model is too simplistic when transaction costs are included, where the cost of the market spread wiped-out all profits in our empirical study when submitting market orders.

In a future study, it would be interesting to add additional features, based on technical indicators, as inputs to the RRL algorithm for an intraday trading scenario (similar to Zhang and Maringer [22-23] for daily data). It would also be interesting to investigate multi-layered architectures further and compare the results with Gold [10] and Gorse [15]. Since the objective function directly affects the performance of the trading model, it would also be interesting to investigate alternative objective functions, other than the differential Sharpe ratio.

REFERENCES

- [1] J. Moody and L.Wu, "Optimization of Trading Systems and Portfolios", in *Proceedings of the IEEE/IAFE 1997, Computational Intelligence for Financial Engineering (CIFER)*, pp. 300–307, 1997.
- [2] M. A. H. Dempster and V. Leemans, "An Automated FX Trading System Using Adaptive Reinforcement Learning", *Expert Systems with Applications*, vol. 30, pp. 543–552, 2006.
- [3] S. Nison, "Japanese Candlestick Charting Techniques", New York Institute of Finance, 1991.
- [4] J. H. Fock, C. Klein and B. Zwergel, "Performance of Candlestick Analysis on Intraday Futures Data", *Journal of Futures Markets*, vol. 13, pp. 28–40, 2005.
- [5] B. R. Marshall, M. R. Young and L.C Rose, "Candlestick technical trading strategies: Can they create value for investors", *Journal of Banking and Finance*, vol. 30, pp. 2303–2323, 2006.
- [6] B. R. Marshall, M. R. Young and R. Cahan, "Are Candlestick Technical Trading Strategies Profitable in the Japanese Equity Market?", *Review of Quantitative Finance and Accounting*, vol. 31, no. 2, pp. 191–207, 2008.
- [7] M. J. Horton, "Stars, crows, and doji: The use of candlesticks in stock selection", *The Quarterly Review of Economics and Finance*, vol. 49, pp. 283–94, 2009.
- [8] M. Duvinage, P. Mazza and M. Petitjean, "The Intraday Performance of Market Timing Strategies and Trading Systems Based on Japanese Candlesticks", *Quantitative Finance*, vol. 13, no. 7, pp. 1059–1070, 2013.
- [9] Prado et al., "On the effectiveness of candlestick chart analysis for the Brazilian stock market", 17th International Conference in Knowledge Based and Intelligent Information and Engineering Systems (KES2013), vol. 22, pp. 1136–1145, 2013.
- [10] C. Gold, "FX Trading via Recurrent Reinforcement Learning", in *Proceedings of the 2003 IEEE International Conference on Computational Intelligence for Financial Engineering*, pp. 363–370, 2003.
- [11] J. Moody and M. Shaffell, "Learning to Trade via Direct Reinforcement", *IEEE Transactions on Neural Networks*, vol. 12, no. 4, pp. 875–889, 2001.
- [12] J. Moody, L. Wu, Y. Liao and M. Shaffell, "Performance Functions and Reinforcement Learning for Trading Systems and Portfolios", *Journal of Forecasting*, vol. 17, pp. 441–470, 1998.
- [13] P. Werbos, "Backpropagation Through Time: What it Does and How to Do it", *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560, 1990.
- [14] F. Bertoluzzo and M. Corazza, "Making Financial Trading by Recurrent Reinforcement Learning", in *Lecture Notes in Computer Science*, B. Apolloni, R. Howlett and L. Jain, Eds., Springer, vol. 4693, pp. 619–626, 2007.
- [15] D. Gorse, "Application of Stochastic Recurrent Reinforcement Learning to Index Trading", in *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, pp. 123–128, 2011.
- [16] D. Maringer and T. Ramtohl, "Threshold Recurrent Reinforcement Learning Model for Automated Trading", in *EvoApplications 2010, Lecture Notes in Computer Science*, C. Di Chio, A. Brabazon, G. A. Di Caro, M. Ebner, M. Farooq, A. Fink, J. Grah, G. Greenfield, P. Machado, M. O'Neill, E. Tarantino, N. Urquhart, Eds., Springer, vol. 6025, pp. 212–221, 2010.
- [17] D. Maringer and T. Ramtohl, "Regime-Switching Recurrent Reinforcement Learning in Automated Trading", in *Natural Computing in Computational Finance, Studies in Computational Intelligence*, A. Brabazon, M. O'Neill, D. Maringer, Eds., Springer, vol. 4, pp. 93–121, 2012.
- [18] D. Maringer and T. Ramtohl, "Regime-Switching Recurrent Reinforcement Learning for Investment Decision Making", *Computational Management Science*, vol. 9, pp. 89–107, 2012.
- [19] D. Maringer and J. Zhang, "Transition Variable Selection for Regime Switching Recurrent Reinforcement Learning", in *Proceedings of the 2014 IEEE Conference on Computational Intelligence for Financial Engineering and Economics*, pp. 407–412, 2014.
- [20] J. Zhang, "Automating Transition Functions: A Way to Improve Trading Profits with Recurrent Reinforcement Learning", in *Artificial Intelligence Applications and Innovations, IFIP Advances in Information and Communication Technology*, L. Iliadis, I. Maglogiannis, H. Papadopoulos, Eds., Springer, vol. 436, pp. 39–49, 2014.
- [21] J. Zhang and D. Maringer, "Two Parameter Update Schemes for Recurrent Reinforcement Learning", in *2014 IEEE Congress on Evolutionary Computation (CEC)*, pp. 1449–1453, 2014.
- [22] J. Zhang and D. Maringer, "Indicator Selection for Daily Equity Trading with Recurrent Reinforcement Learning", in *GECCO 2013*, pp. 1757–1758, 2013.
- [23] J. Zhang and D. Maringer, "Using a Genetic Algorithm to Improve Recurrent Reinforcement Learning for Equity Trading", *Computational Economics*, pp. 1–17, 2015.
- [24] G. Caginalp and H. Laurent, "The Predictive Power of Price Patterns", *Applied Mathematical Finance*, vol. 5, pp. 181–205, 1998.
- [25] K. H. Lee and G.S. Jo, "Expert system for predicting stock market timing using a candlestick chart", *Expert Systems with Applications*, vol. 16, no. 4, pp. 357–364, 1999.
- [26] Y. Goo, D. Chen and Y. Chang, "The Application of Japanese Candlestick Trading Strategies in Taiwan", *Investment Management and Financial Innovations*, vol. 4, pp. 49–71, 2007.
- [27] M. Jasemi, A. M. Kimiagari and A. Memariani, "A modern neural network model to do stock market timing on the basis of the ancient investment technique of Japanese Candlestick", *Expert Systems with Applications*, vol. 38, no. 4, pp. 3884–3890, 2011.
- [28] Y. Shiu and T. Lu, "Pinpoint and synergistic trading strategies of candlesticks", *International Journal of Economics and Finance*, vol. 3, pp. 234–244, 2011.
- [29] T. Lu and Y. Shiu, "Tests for two-day candlestick patterns in the emerging equity market of Taiwan", *Emerging Markets in Finance and Trade*, vol. 48, pp. 41–57, 2012.
- [30] T. Lu, Y. Shiu and T. Liu, "Profitable candlestick trading strategies: the evidence from a new perspective", *Review of Financial Economics*, vol. 21, pp. 63–68, 2012.
- [31] C. N. Ochotorena et al., "Robust stock trading using fuzzy decision trees", 2012 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFER), pp. 1–8, 2012.
- [32] N. Goumatianos N, I. T. Christou and P. Lindgren, "Stock Selection System: Building Long/Short Portfolios Using Intraday Patterns", *International Conference On Applied Economics (ICOAE) 2013*, vol. 5, pp. 298–307, 2013.
- [33] T. Lu and J. Chen, "Candlestick charting in European stock markets", *JASSA*, no. 2, pp. 20–25, 2013.
- [34] N. Goumatianos N, I. T. Christou and P. Lindgren, "Intraday Business Model Strategies on Forex Markets: Comparing the Performance of Price Pattern Recognition Methods", *Journal of Multi Business Model Innovation and Technology*, vol. 2, no. 1, pp. 1–34, 2013.
- [35] T. Lu, "The profitability of candlestick charting in the Taiwan stock market", *Pacific-Basin Finance Journal*, vol. 26, pp. 65–78, 2014.
- [36] C. Tsai and Z. Quan, "Stock Prediction by Searching for Similarities in Candlestick Charts", *ACM Transactions on Management Information Systems (TMIS)*, vol 5, no. 2, pp. 9:1–9:21, 2014.
- [37] B. Detollenaere and P. Mazza, "Do Japanese Candlesticks Help Solve the Trader's Dilemma?", *Journal of Banking & Finance*, vol. 48, pp. 386–395, 2013.
- [38] Slickcharts, 2011. [Online]. Available: <http://www.slickcharts.com>.