# Modelling and Simulation of Directional Financial Time Series

**Mahayaudin M. Mansor, David A. Green and Andrew V. Metcalfe**

*School of Mathematical Sciences, University of Adelaide, Adelaide 5005 SA, Australia*
*Email: mohdmahayaudin.mansor@adelaide.edu.au*

**Abstract:** A stationary time series model is directional if it has properties that are not symmetric with respect to time, for example if, after the expected value is adjusted to 0, the expected value of the product of the variable squared and the following variable differs from the expected value of the product of the variable and the following variable squared. Marked directionality in time series is apparent from a comparison of time series plots in chronological order and in reverse order (time-to-go), but formal quantitative tests are needed to determine the strength of evidence against an assumption of reversibility (no directionality). Linear time series models with Gaussian noise (LGN models) are reversible whereas linear models with other noise distributions and non-linear models are directional. Evidence of directionality in time series can be found in many disciplines including financial time series. Examples of directional financial time series are presented, together with values of directionality statistics. A method of calculating the p-values of the directional statistics, corresponding to a null hypothesis of reversibility, using simulation is presented. The time series are then modelled by threshold autoregressive models, with noise distributions, which can be non-Gaussian, fitted to match the empirical distribution of the residuals. The fitting criterion is a weighted sum of the error sum of squares and a measure of discrepancy between values of the directionality statistics in the time series and long simulations from the potential model. Strategies for determining suitable weights for the criterion will be investigated. The directional time series models will be assessed in terms of the distribution of extreme values of the time series, which is a feature that was not used in the fitting process. The assessment will be made by comparing the distribution of extremes from: the observed time series; a long simulation from the fittted directional model; and a long simulation from a LGN model fitted to the time series on the basis of the Akaike information criterion (AIC).

*Keywords:* *Directional time series, reversibility, financial time series, randomization test, threshold autoregressive models*

M.M. Mansor, D.A. Green and A.V. Metcalfe, Modelling and simulation of directional financial time series.

## 1   INTRODUCTION

Many financial time series show non-stationarity inasmuch as there is a trend, or there are seasonal effects. For example, sales of summer clothing are seasonal and may be increasing in line with an increasing population. In contrast there is no reason to suppose that there is any seasonality or long term trend in exchange rates. In any case, if a trend or seasonality are identified in a time series they can be estimated and removed to give a deseasonalized and detrended series that can be considered a realisation of a stationary random process.

In this investigation we assume that the time series is a realisation of a stationary random process. A stationary random process is reversible if it is not possible to distinguish between a realisation in its time order and the same realisation in reverse time order (time-to-go). If a stationary random process is not reversible it is directional. Linear causal models with independent Gaussian errors are reversible. Linear causal models with non-Gaussian errors are directional, as are non-linear models. For some time series it is possible to discern directionality by comparing a plot against time with a plot against time-to-go. For example, the monthly exchange rate of the British pound sterling against the U.S. dollar (GBPUSD XRT), shown in the upper frame of Figure 1, does show three marked plummets and relatively slow recoveries. The reverse time plot (time-to-go) is clearly different in these respects, in fact we see rapid increases, followed by slow decreases.
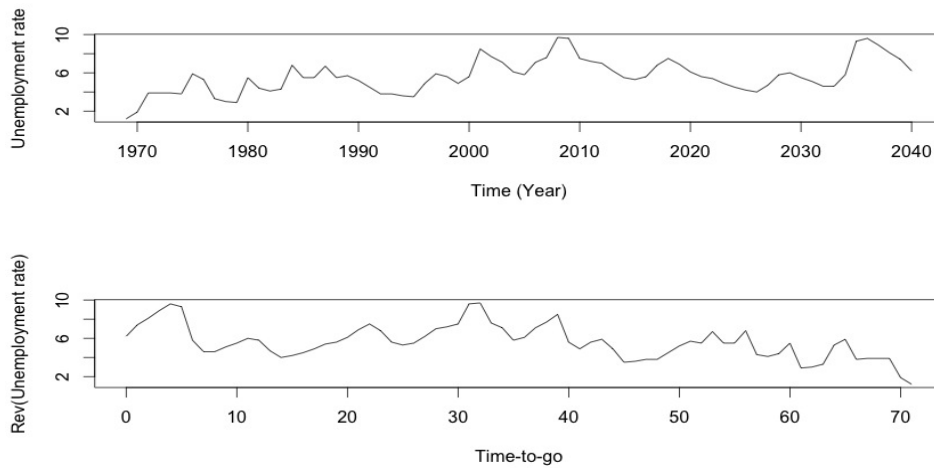
## 2   FINANCIAL DATA
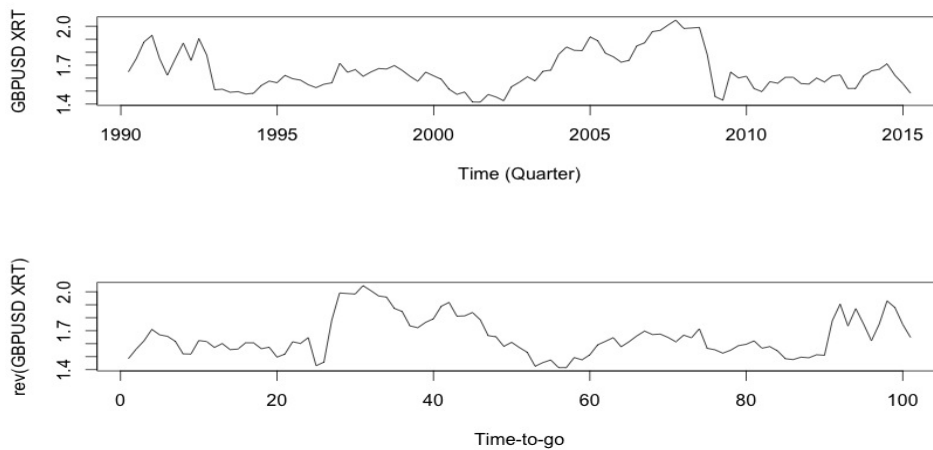


**Figure 1**. Quarterly trading records of GBPUSD XRT from March 1990 to March 2015
[Data courtesy of Bloomberg]
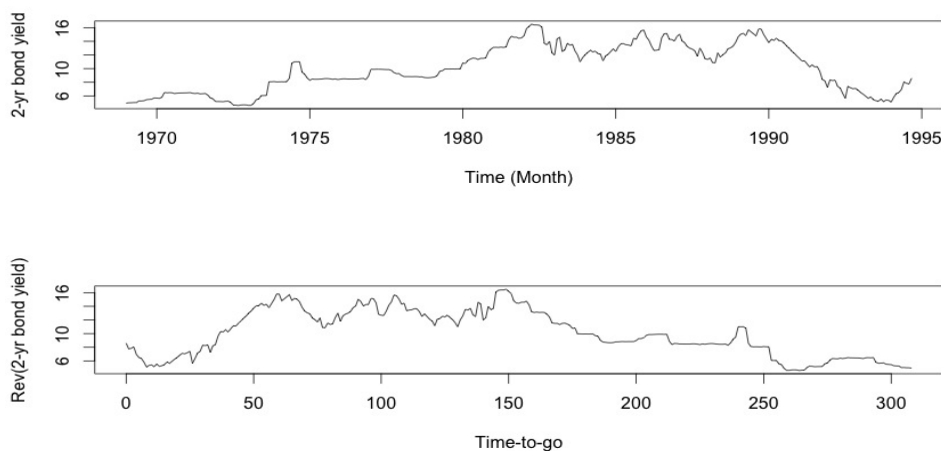


**Figure 2**. Interest rates (per annum) for Australian Government Bond Yield 2-yr securities, monthly from Jan 1969 to Sep 1994 [Data courtesy of Reserve Bank of Australia]

**Figure 3**. Yearly U.S. unemployment rate from 1969 to 2014 [Data courtesy of the U.S. Bureau of Labor Statistics]

Graphical inspection for the financial series show some degree of directionality in its time order (above) and in its reverse time order (below). The exchange rates (Figure 1) show a few well defined plummets. In contrast there is little qualitative difference between the bond interest rate (Figure 2) in time order and time-to-go order. In the case of the unemployment (Figure 3) increases generally appear to be steeper than the decreases.

## 3 TESTING FOR DIRECTIONALITY

There are many statistics that have been designed to detect directionality in time series (e.g. Lawrance, 1991, Beare and Seo, 2012 and Wild et al., 2014). The methods can be categorised into methods based on original series and methods based on first differences. A general, and intuitive, indicator is the skewness of the first order lag one differences. A time series modelled by random variables is given by

$$\{X_t\} \quad \text{for} \quad t = 1, 2, ..., n \tag{1}$$

and the first order lag one differences are

$$Y_t = X_t - X_{t-1} \quad \text{for} \quad t = 2, 3, ..., n. \tag{2}$$

The mean of the differences is $(x_n - x_1)/n$ which is approximately 0 for any long time series. If there is a tendency for rapid increases to be followed by more gradual recessions there will be rather more negative differences which will on average be less than the average of the positive differences. However there will be fewer positive differences. It follows that the distribution of differences will be positively skewed: there will be some outlying positive differences associated with the rapid increases; and the median will be negative and to the left of the mean which is near 0. In contrast, if there is a tendency for rapid decreases to be followed by relatively slow recoveries the distribution of differences will have negative skewness. We use the product moment measure of skewness, and refer to the skewness of the differences, $\gamma_Y$ as the directionality.

The skewness in $Y_t$ can be estimated by

$$\hat{\gamma}_Y = \frac{\sum_{t=1}^n (y_t - \bar{y})^3}{[\sum_{t=1}^n (y_t - \bar{y})^2]^{3/2}}. \tag{3}$$

There is evidence of directionality if the estimate $\hat{\gamma}_Y$ of $\gamma_Y$ is statistically significantly different from 0.

**Table 1**. Statistics of time series

| Statistic | GBPUSD XRT | 2-yr Bond Yeild | Unemployment Rate |
|---|---|---|---|
| Time interval | quarterly | monthly | yearly |
| Length, $n$ | 101 | 309 | 72 |
| Median, $M$ | 1.62 | 9.90 | 5.50 |
| Mean, $\bar{x}$ | 1.65 | 10.06 | 5.63 |
| Standard dev., $s$ | 0.1525 | 3.3542 | 1.7596 |
| Lag 1 ACF | 0.8490 | 0.9840 | 0.7650 |
| Directionality, $\hat{\gamma}$ | -0.7566 | 0.3039 | 1.0246 |
| Two-sided P-value | ( 0.004 ) | ( 0.028 ) | ( 0.00 ) |
| Ratio | [ 3.18 ] | [ 2.18 ] | [ 3.89 ] |

The directionality of the exchange rates (Figure 1), which does show clear rapid decreases, is -0.75. The directionality of the bond yields (Figure 2) and the unemployment rate (Figure 3) is 0.30 and 1.02 respectively. Inspection of time series plot and time-to-go plot of the unemployment rates (Figure 3) does show steeper ascents than descents in the former, and steeper descents than escents in the latter.

All three directionalities shown to be statistically different from 0 (two-sided P-values in Table 1) using a simulation procedure (Mansor et al., 2015). In this simulation procedure, we fit an AR($p$) model to the time series of length $n$ using the AIC criterion to select $p$, which is given in Table 2. The null hypothesis is that the time series is reversible. We simulate 1000 time series of length $n$ from the fitted AR($p$) model, using a Gaussian distribution with mean 0 and variance equal to the variance of the residuals for the errors. For each of the simulated time series, we calculate $\hat{\gamma}_Y$. We determine the proportion of the simulations for which the test statistic is absolutely more extreme than the calculated $\hat{\gamma}_Y$ from the original time series. We return this proportion from the simulation as the two-sided P-value. The ratio in Table 1 is the test statistic in absolute terms to its standard deviation from the simulation.

## 4  MODELLING DIRECTIONALITY

For all three time series the residuals from the AR(1) model have a much smaller standard deviation ($\sigma_{est}$), given in Table 2 (row 2, Order 1), than that of the original time series ($s$), given in Table 2 (row 1, Order 0 i.e. the original time series). In contrast, the further reductions in $\sigma_{est}$ when AR($p$) models, as selected by the AIC, are fitted, given in Table 2 (row 3, AIC) is slight.

**Table 2**. Estimated standard deviation of errors for AR models

| Order | GBPUSD XRT | 2-yr Bond Yeild | Unemployment Rate |
|---|---|---|---|
| 0 | $s = 0.1525$ | $s = 3.3542$ | $s = 1.7596$ |
| 1 | AR(1), $\sigma_{est} = 0.0811$ | AR(1), $\sigma_{est} = 0.6029$ | AR(1), $\sigma_{est} = 1.1409$ |
| AIC | AR(3), $\sigma_{est} = 0.0798$ | AR(10), $\sigma_{est} = 0.5934$ | AR(2), $\sigma_{est} = 1.1214$ |

These results suggest that the simple AR(1) model is a reasonable first approximation for all three time series. However, in all cases, the directionality of the realisation of the fitted AR(1) model, using random permutation of the residuals as errors, does not match the directionality calculated from the observed time series (Table 3). The objective is to improve this aspect of the modeling by introducing a non-linear TAR(1) model. The TAR(1) model is given by

$$X_t = \begin{cases} \alpha_U X_{t-1} + \epsilon_t & \text{if } X_{t-1} > T_U \\ \alpha_L X_{t-1} + \epsilon_t & \text{if } X_{t-1} < T_L \end{cases} \tag{4}$$

We fit the TAR(1) model by finding the values of $\alpha_L$ and $\alpha_U$ that minimize the discrepancy,

$$\omega = \sum_{t=2}^{n} r_t^2 + \phi(\hat{\gamma}_Y - \tilde{\gamma}_Y)^2, \tag{5}$$

where $\{r_t\}$ are the residuals, $\tilde{\gamma}_Y$ is an estimate of the directionality of the TAR(1) model, and $\phi$ is the weight given to agreement between the $\tilde{\gamma}_Y$ and the target directionality value $\hat{\gamma}_Y$. The estimate of directionality of the TAR(1) model is calculated from a long simulation (length $N = 10^5$) of the TAR(1) model using random sampling, with replacement from the residuals. We use $\hat{\alpha}$ from the AR(1) model as the inital value for $\alpha_L$ and $\alpha_U$ when fitting the TAR(1) model. The constraints for stability are that $-1 < \alpha_L, \alpha_U < 1$. The optimization is carried out with the *optim()* function in *R* using the Nelder-Mead option (e.g. Nash, 2014). We consider an TAR(1) with 80% threshold for each series.

**Table 3**. Directionality measure modelled by AR(1) and TAR(1)_80% using resampled residuals

| Financial Series | Target $\hat{\gamma}$ | AR(1) | TAR(1)_80% | $\phi$ |
|---|---|---|---|---|
| Quarterly exchange rates of GBPUSD | -0.7566 | $\tilde{\gamma} = -0.4031$ $\alpha = 0.8487$ $\sigma_{est} = 0.0811$ | $\tilde{\gamma} = -0.7273$ $\alpha_L = 0.9981, \alpha_U = 0.9361$ $\sigma_{est} = 0.0805$ | $10^1$ |
| Monthly interest rates of 2-yr Government bond yield | 0.3039 | $\tilde{\gamma} = 0.4164$ $\alpha = 0.9838$ $\sigma_{est} = 0.6029$ | $\tilde{\gamma} = 0.3040$ $\alpha_L = 0.7587, \alpha_U = 0.7310$ $\sigma_{est} = 0.9467$ | $10^8$ |
| Yearly United States unemployment rates | 1.0246 | $\tilde{\gamma} = 0.9932$ $\alpha = 0.7652$ $\sigma_{est} = 1.1409$ | $\tilde{\gamma} = 1.0231$ $\alpha_L = 0.6170, \alpha_U = 0.8349$ $\sigma_{est} = 0.9959$ | $10^7$ |

• In the case of the exchange rates, the TAR(1) gives a close agreement with the observed directionality, and a slightly lower estimate of the standard deviation of errors.
• In the case of the bond rates, which has a value of $\alpha$ close to 1 for the AR(1) model, the closer agreement of the TAR(1) model with the observed directionality is at the expense of an increase in estimated standard deviation of the errors.
• In the case of the U.S. unemployment rates, with the lowest value of $\alpha$ in the AR(1) model, both AR(1) and TAR(1) give a close fit to the observed directionality. The TAR(1) has a slightly lower estimated standard deviation of the errors.
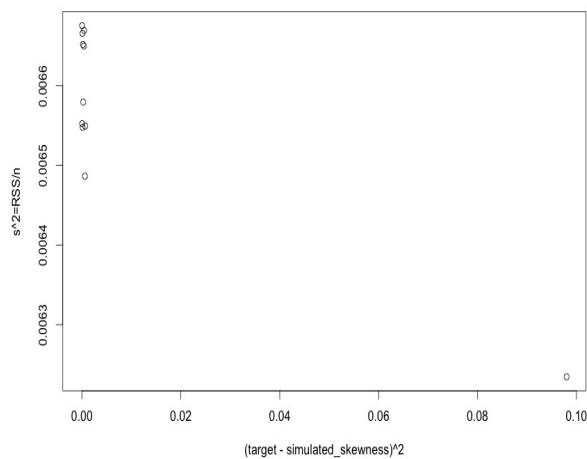


**Figure 4**. Trade-off between minimising the sum of squared residuals and minimising the discrepancy in directionality by increasing $\phi$ for the GBPUSD exchange rates time series

We illustrate in Figure 4 the balance achieved between the sum of squared residuals and the discrepancy between the target directionality and the simulated directionality of the quarterly GBPUSD exchange rates for $N=10^5$.

## 5 SIMULATION OF DIRECTIONAL TIME SERIES

We simulated 1000 time series of length 72 years, the length of the original time series for the U.S. unemployment rates using an AR(1) and a TAR (1; 80% threshold, 2 parameter) models. The upper and lower coefficients for TAR (1) are the optimised parameters of $\alpha_L$=0.6197 and $\alpha_U$=0.8359 (Table 3). For, both models two different error distributions were fitted to the residuals: an Extreme Value Type 1 (Gumbel) distribution of minima (GD); and back-to-back Weibull distribution, that is one Weibull distribution was fitted to the positive residuals and another to the absolute values of the negative residuals, (WD). The maxima, over 72 years, from the 1000 simulations, with GD and with WD, were stored. Their boxplots are shown in Figure 5, and the quantile-quantile (Q-Q) plots are shown in Figure 6.
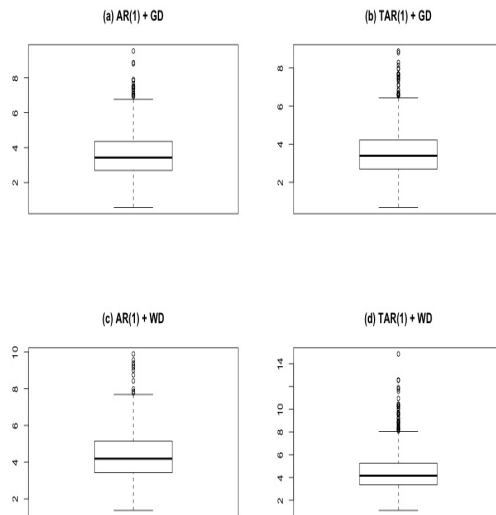


**Figure 5**. Boxplot of the simulated maxima in the U.S. unemployment rate for each model
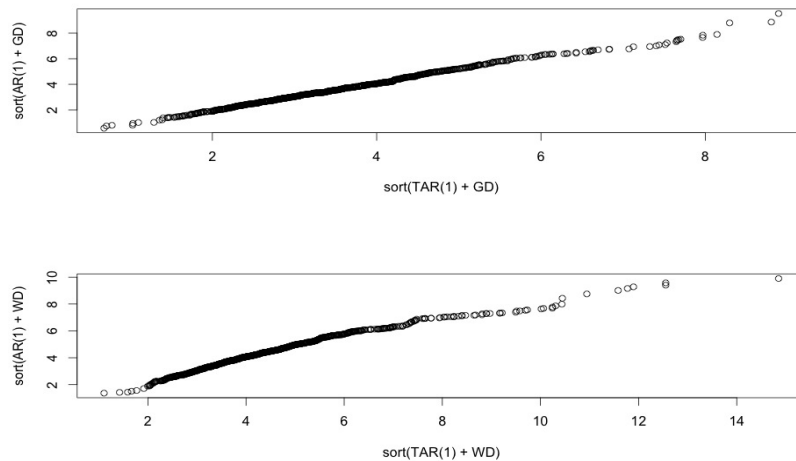


**Figure 6**. Q-Q plot of the simulated maxima in the U.S. unemployment rate, between TAR(1) versus AR(1)

There is no discernible difference in the boxplots for the distributions of maxima with the GD distribution (Figure 5 upper). In both cases the distribution of the maxima is positively skewed as is typical for extreme values. However, there is a noticeable difference between TAR(1) and AR(1) with WD errors, such that the maxima from the TAR(1) are more skewed and more extreme. The Q-Q plots (Figure 6 lower) show that the TAR(1) with WD has generally higher values than AR(1) with WD. Moreover, the Q-Q plots (Figure 6 upper) also show that TAR(1) has slightly higher values than AR(1) with GD.

M.M. Mansor, D.A. Green and A.V. Metcalfe, Modelling and simulation of directional financial time series.

## 6 CONCLUSIONS

We focus on directionality which we have defined in terms of skewness of the first differences of the time series. The financial time series of the quarterly GBPUSD exchange rates, the monthly Australian bond yield and the yearly U.S. unemployment rates are shown to be directional and the estimated values of the lag 1 autocorrelation are 0.8490, 0.9840 and 0.7650 respectively. All three series are well modelled as AR(1) if a least squares error criterion is adopted.

Many economic time series are well modelled as an AR(1). If high frequency sampling has the consequence that the parameter of the AR(1) model ($\alpha$) is close to 1, then the estimate of directionality is very close to the skewnnes of the errors. We choose to model the time series of U.S. unemployment in more detail as it had the lowest value of $\alpha$ (0.76). In contrast, if an AR(1) model has $\alpha$ close to 0, then the skewness of errors has little effect on directionality.

We have demonstrated that in the intermediate range, in particular for GBPUSD exchange rates and U.S. unemployment rates TAR(1) model, fitted with discrepancy strategy, can improve the fit to observed directionality and decrease the estimate standard deviation of the errors ($\sigma_{est}$) (Table 3). We have also shown that the TAR(1) model combined with a back-to-back Weibull model (WD) for the residuals gives substantially higher extreme values than the AR(1) model combined with WD for the residuals, in the case of the U.S. unemployment series. A link between extreme values and directionality has been discussed by Soubeyrand et al. (2014) in the context of rainfall, and Mansor et al. (2015) in the context of sunspots.

### REFERENCES

Beare, B. K. and J. Seo (2012). Time irreversible copula-based markov models. *Econometric Theory*, 1–38.

Lawrance, A. (1991). Directionality and reversibility in time series. *International Statistical Review/Revue Internationale de Statistique*, 67–79.

Mansor, M. M., M. E. Glonek, D. A. Green, and A. V. Metcalfe (2015). Modelling directionality in stationary geophysical time series. *Proceedings of the international work-conference on Time Series (ITISE 2015)*, 755–766.

Mansor, M. M., D. A. Green, and A. V. Metcalfe (2015). Detecting directionality in time series. *in preparation*.

Nash, J. (2014). On best practice optimization methods in r. *Journal of Statistical Software 60*(2), 1–14.

Soubeyrand, S., C. E. Morris, and E. K. Bigg (2014). Analysis of fragmented time directionality in time series to elucidate feedbacks in climate data. *Environmental Modelling & Software 61*, 78–86.

Wild, P., J. Foster, and M. Hinich (2014). Testing for non-linear and time irreversible probabilistic structure in high frequency financial time series data. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 177*(3), 643–659.