

# Volatility Forecasts in Financial Time Series with HMM-GARCH Models

Xiong-Fei Zhuang and Lai-Wan Chan

Computer Science and Engineering Department  
The Chinese University of Hong Kong, Shatin, Hong Kong  
{xfzhuang, lwchan}@cse.cuhk.edu.hk

**Abstract.** Nowadays many researchers use GARCH models to generate volatility forecasts. However, it is well known that volatility persistence, as indicated by the sum of the two parameters  $G_1$  and  $A_1$  [1], in GARCH models is usually too high. Since volatility forecasts in GARCH models are based on these two parameters, this may lead to poor volatility forecasts. It has long been argued that this high persistence is due to the structure changes (e.g. shift of volatility levels) in the volatility processes, which GARCH models cannot capture. To solve this problem, we introduce our GARCH model based on Hidden Markov Models (HMMs), called HMM-GARCH model. By using the concept of hidden states, HMMs allow for periods with different volatility levels characterized by the hidden states. Within each state, local GARCH models can be applied to model conditional volatility. Empirical analysis demonstrates that our model takes care of the structure changes and hence yields better volatility forecasts.

## 1 Introduction

Volatility analysis of financial time series is an important aspect of many financial decisions. For example, fund managers, option traders and etc., are all interested in volatility forecasts in order to either construct less risky portfolios or obtain higher profits. Hence, there is always a need for good analysis and forecasting of volatility.

In the last few decades, many volatility models have been put forward. The most popular and successful models among these are the autoregressive conditional heteroskedasticity (ARCH) models by Engle [2] and extended to generalized ARCH (GARCH) by Bollerslev [3]. Their success stems from their ability to capture some stylized facts of financial time series, such as time-varying volatility and volatility clustering. Other volatility models include the stochastic volatility (SV) models and etc..

Although standard GARCH models improve the in-sample fit a lot compared with constant variance models, numerous studies find that GARCH models give unsatisfactory forecasting performances, see [4]. Andersen and Bollerslev [5] pointed out that GARCH models give good volatility forecasting by increasing the data sampling frequency, such as intra-day data. Increasing the sampling

frequency may also lessen the effect of the structure changes. Our research work, however, focus on the commonly used inter-day data. Hence the effect of structure changes is unavoidable.

In our paper we argue that the usually overstated volatility persistence in GARCH models may be the cause of poor forecasting performances. And many researchers show that this well-known high persistence may originate from the structure changes in the volatility processes, which GARCH models cannot capture. For example, Lamoureux [1] demonstrated that any shift in the structure of financial time series (e.g. the shift of unconditional variance) is likely to lead to misestimation of the GARCH parameters in such a way that they imply too high a volatility persistence.

One approach for modelling volatility structure changes is to use a Hamilton type regime switching (RS) model [6]. The earlier RS applications [7] tend to be rigid, where conditional variance is a constant within each regime. Recent extensions by Cai [8] apply an ARCH specification into the RS model to allow conditional variance to be time dependent.

Our goal is to solve the problem of excessive persistence in original GARCH models by introducing Hidden Markov Models to allow for different volatility states (periods with different volatility levels) in time series. And also, within each state, we allow GARCH models to model the conditional variance. As a result, our model is more parsimonious than Cai [8] because they use an ARCH class specification in each regime. The resulting HMM-GARCH model indeed yields better volatility forecast compared to original GARCH models for both artificial data and real financial data, in-sample as well as out-of-sample.

The overview of this paper is as follows. In section 2, we will formally define our HMM-GARCH model. In section 3, we will give a detailed description of the data, methodology and the empirical results of our volatility forecasting experiments. Finally, our conclusion is in section 4.

## 2 HMM-GARCH Model

### 2.1 GARCH Models

First, let us consider a time series  $Y$ , and observation  $y_t$  ( $t = 1, 2, \dots, T$ ) is the value of  $Y$  at time step  $t$ . Here we consider the GARCH(1,1) models, which are adequate for modelling volatilities in most financial time series (see [9]). The GARCH(1,1) model is as follows:

$$y_t = C + \epsilon_t \quad (1)$$

$$\sigma_t^2 = K + G_1 \sigma_{t-1}^2 + A_1 \epsilon_{t-1}^2 \quad (2)$$

In the conditional mean model Eq.(1), each observation  $y_t$  consists of a conditional mean  $C$ , plus an uncorrelated, white noise innovation  $\epsilon_t$ . In the conditional variance model Eq.(2), the conditional variance  $\sigma_t^2$  consists of a constant  $K$ , plus a weighted average of last period's forecast,  $\sigma_{t-1}^2$ , and the last period's squared

innovation  $\epsilon_{t-1}^2$ . Also we need  $K > 0$ ,  $G_1 > 0$  and  $A_1 > 0$  to ensure a positive conditional variance  $\sigma_t^2$ . A convenient way to express the innovation is  $\epsilon_t = Z_t \sigma_t$ , where  $Z_t$  is an i.i.d. process with zero mean and unit variance. For the parameter estimates of the GARCH(1,1) model, the likelihood functions have been maximized by using the BFGS optimization algorithm in the MATLAB optimization routines.

## 2.2 Hidden Markov Models

The basic concept of a Hidden Markov model is a doubly embedded stochastic process with an underlying stochastic process (the state sequence) that is not observable or hidden. The state sequence is a Markov process and it is called hidden because it can only be observed through another stochastic process (observation sequence), where each observation is a probabilistic function of the corresponding state.

We now define the notation of an HMM [10] which will be used later. Given the time series  $Y$ , an HMM is characterized by the following: 1)  $N$ , the number of states in the model. In our model the states refer to different variance levels. We denote the state set as  $S = \{S_1, S_2, \dots, S_N\}$ , and the state at time  $t$  as  $q_t$ ,  $q_t \in S$ . 2) The state transition probability distribution  $A = \{a_{ij}\}$  where  $a_{ij} = P[q_t = S_j | q_{t-1} = S_i]$ ,  $1 \leq i, j \leq N$ . 3) The observation probability distribution  $B$ . 4) The initial state distribution  $\pi$ . For convenience, we used the compact notation  $\lambda = (A, B, \pi)$  to indicate the complete parameter set of the model.

Given the form of the HMMs, the goal is to find the best model for a given time series through optimally adjusting model parameters ( $\lambda = (A, B, \pi)$ ).

## 2.3 HMM-GARCH Model

Our model is a hybrid model of the original GARCH models and HMMs.

First, we use HMMs to divide the entire time series into regimes with different volatility levels. The return of the time series is assumed to be modelled by a mixture of probability densities and each density function corresponds to a hidden state with its mean and variance. Viterbi algorithm [10] in HMMs is used to find the state sequence in the time series. Then we get the subsets of original time series corresponding to different states (volatility levels).

Second, within each regimes, we allow GARCH models with different parameter sets to model the local conditional variance as:

$$y_t = C^i + \epsilon_t \quad (3)$$

$$\sigma_t^2 = K^i + G_1^i \sigma_{t-1}^2 + A_1^i \epsilon_{t-1}^2 \quad (4)$$

where  $i$  denotes the state of the time series at time  $t$ .  $K^i$ ,  $G_1^i$  and  $A_1^i$  are the parameter sets of local GARCH models related to state  $i$ .

Third, for the volatility forecast  $\sigma_t^2$  of the global model, we need to predict the state  $i$  of time series at time  $t + 1$  (next state). To make the prediction of

the next state, we define  $\alpha_t(i) = P(y_1, y_2, \dots, y_t, q_t = i | \lambda)$ , we can then estimate the probability of next state in terms of the transition probabilities  $a_{ij}$  as:

$$P(q_{t+1} = j | y_1, y_2, \dots, y_t, \lambda) = \frac{P(y_1, y_2, \dots, y_t, q_{t+1} = j | \lambda)}{P(y_1, y_2, \dots, y_t | \lambda)} \quad (5)$$

$$= \frac{\sum_{i=1}^N \alpha_t(i) a_{ij}}{\sum_{j=1}^N (\sum_{i=1}^N \alpha_t(i) a_{ij})} \quad (6)$$

where  $\alpha_t(i)$  can be estimated from the forward-backward algorithm [10].

After the next state  $i$  at time  $t + 1$  has been determined as above, we can choose the corresponding local GARCH model with parameter sets  $K^i$ ,  $G_1^i$  and  $A_1^i$  to make volatility forecast.

### 3 Volatility Forecast Evaluation and Comparison

#### 3.1 Data and Methodology

We used both artificial data sets and real financial data sets in our volatility forecast experiments. We considered both the in-sample forecasting performances and the out-of-sample forecasting performances.

First, we used artificial data because we know the exact regime switching processes in the time series in order to testify if our model solve the problems of excessive persistence in original GARCH models. We generated an artificial data set of total length 550 that switches between two GARCH processes. The diagonal elements  $a_{ii}$  of the transition matrix  $A$  are  $a_{11} = 0.98$  and  $a_{22} = 0.96$ .

Second, to test if our model is useful in practice, we used real financial data sets (stock return time series) in our experiments. All 5 stocks are chosen from the Hang Seng Index (HSI) components with stock ID: 1, 3, 11, 17, 179 (from 2001/02/01 to 2003/05/01), a total of 550 days.

For all data sets, we use the rolling window in the experiments, 500 observations (about 2 years) were employed for training and in-sample evaluation purposes. The next observation was used for out-of-sample evaluation purposes.

#### 3.2 Empirical Results and Evaluation

Since the actual volatility at time  $t$  is not observable, we need some measures of volatility to evaluate the forecasting performance. In this paper we use the standard approach suggested by Pagan and Schwert [7]. A proxy for the actual volatility  $\hat{\sigma}_t^2$  is given by

$$\hat{\sigma}_t^2 = (y_t - \bar{y})^2 \quad (7)$$

where  $\bar{y}$  is the mean of the time series over the sample period.

Mean Squared Error (MSE), a commonly used method in the literature, is used in this paper:

$$MSE = T^{-1} \sum_{t=1}^T (\hat{\sigma}_t^2 - \sigma_t^2)^2 \quad (8)$$

Table 1 and 2 below show the evaluation results. We use a two-state HMM-GARCH model in our experiments. In both tables, t-v represents true value, HG stands for HMM-GARCH model and o-G stands for original GARCH model.  $s_1$  and  $s_2$  indicate the two states with low and high volatility levels, respectively.  $MSE_1$  is the in-sample MSE while  $MSE_2$  is the out-of-sample MSE. The asterisks (\*) means the results is significant at the 10% level.

**Table 1.** MSE for the artificial data set and the true parameter sets compared with those obtained from HMM-GARCH model and original GARCH models.

models	$C$	$K$	$G_1$	$A_1$	$G_1+A_1$	$MSE_1$	$MSE_2$
t-v ( $s_1$ )	$1 \cdot 10^{-3}$	$1 \cdot 10^{-5}$	0.20	0.10	0.30	/	/
( $s_2$ )	$1 \cdot 10^{-3}$	$5 \cdot 10^{-5}$	0.60	0.30	0.90		
HG ( $s_1$ )	$1 \cdot 10^{-3}$	$1 \cdot 10^{-5}$	0.35	0.16	0.51	$\cdot 10^{-5}$	$\cdot 10^{-7} *$
( $s_2$ )	$1 \cdot 10^{-3}$	$7 \cdot 10^{-5}$	0.65	0.32	0.97		
o-G	$1 \cdot 10^{-3}$	$2 \cdot 10^{-5}$	0.58	0.41	0.99	$\cdot 10^{-5}$	$\cdot 10^{-7} *$

**Table 2.** MSE for the stock return data sets and the parameter sets obtained from HMM-GARCH model and original GARCH models. (We ignore  $C, K$  here because we only care about the volatility persistence for real financial data.)

stock	models	$G_1$	$A_1$	$G_1+A_1$	$MSE_1$	$MSE_2$
001	HG ( $s_1$ )	0.88	0.00	0.88	$0.13 \cdot 10^{-5}$	$0.03 \cdot 10^{-5} *$
	( $s_2$ )	0.06	0.00	0.06		
	o-G	0.80	0.12	0.92	$0.14 \cdot 10^{-5}$	$0.04 \cdot 10^{-5} *$
003	HG ( $s_1$ )	0.55	0.11	0.66	$0.49 \cdot 10^{-7} *$	$0.16 \cdot 10^{-7} *$
	( $s_2$ )	0.47	0.00	0.47		
	o-G	0.60	0.13	0.73	$0.51 \cdot 10^{-7} *$	$0.18 \cdot 10^{-7} *$
011	HG ( $s_1$ )	0.88	0.00	0.88	$0.11 \cdot 10^{-6}$	$0.01 \cdot 10^{-6} *$
	( $s_2$ )	0.01	0.00	0.01		
	o-G	0.90	0.06	0.96	$0.13 \cdot 10^{-6}$	$0.03 \cdot 10^{-6} *$
017	HG ( $s_1$ )	0.93	0.02	0.95	$0.16 \cdot 10^{-4}$	$0.02 \cdot 10^{-4} *$
	( $s_2$ )	0.03	0.00	0.03		
	o-G	0.67	0.29	0.96	$0.20 \cdot 10^{-4}$	$0.03 \cdot 10^{-4} *$
179	HG ( $s_1$ )	0.57	0.00	0.57	$0.74 \cdot 10^{-5} *$	$0.07 \cdot 10^{-5} *$
	( $s_2$ )	0.07	0.00	0.07		
	o-G	0.95	0.02	0.97	$0.82 \cdot 10^{-5} *$	$0.12 \cdot 10^{-5} *$

The results above show that HMM-GARCH model recovers the switching processes between two different volatility regimes with different volatility persistence ( $G_1 + A_1$ ). However, the original GARCH models can not capture such volatility structure changes and always show a very high volatility persistence. As a result, we can see that HMM-GARCH model gives better volatility forecasts because the MSE of HMM-GARCH model is significantly smaller than the original GARCH models for most of the time.

## 4 Conclusion

This paper is based on the well-known fact that the volatility persistence of widely-used GARCH models is usually too high so that original GARCH models give poor volatility forecasts. And one possible reason for this excessive persistence is the structure changes (e.g. shift of volatility levels) in the volatility processes, which GARCH models cannot capture.

Therefore, we developed our HMM-GARCH model to allow for both different volatility states in time series and state specific GARCH models within each state. Our model shares the basic regime-switching concept of other recent RS applications (see [8], and [11]), and is more parsimonious in each state by allowing GARCH type heteroskedasticity.

The empirical results for both artificial data and real financial data show that the excessive persistence problems disappear in our model. And as a result, the forecasting performance of our model outperforms original GARCH models for both in sample and out-of-sample evaluation. These results suggest that it is promising to study volatility persistence in more detail, including the hidden regime-switching mechanisms, to improve volatility forecasts in future research.

## Acknowledgement

The authors would like to thank The Research Grants Council, HK for support.

## References

1. Lamoureux, C., and etc., Persistence in Variance, Structural Change, and the GARCH Model. *J. Bus Econ. Statist.*, 8:2, 225-234, 1990.
2. Engle, Robert F., Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50:4, 987-1007, 1982.
3. Bollerslev, T. Generalized Autoregressive Conditional Heteroskedasticity. *J. Econometrics*, 31, 307-327, 1986.
4. Figlewski, S. Forecasting Volatility. *Financial Markets, Institutions and Instruments*, 6, 1-88, 1997.
5. Andersen, T.G. and T. Bollerslev Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts. *Inter. Econ. Review*, 39, 885-905, 1998.
6. Hamilton, J.D. Time series analysis. Princeton University Press, 1994.
7. Pagan, A. R., G. W. Schwert, "Alternative Models for Conditional Stock Volatility. *J. Econometrics*, 45, 267-290, 1990.
8. Cai, J. A Markov Model of Unconditional Variance in ARCH. *J. Bus Econ. Statist.*, 12, 309-316, 1994.
9. Bollerslev, and etc., ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. *J. Econometrics*, 52, 5-59, 1992.
10. Rabiner, L.R., A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. of the IEEE*, 77:2, 1989.
11. Ser-Huang Poon and Clive W.J. Granger Forecasting Volatility in Financial Markets: A Review. *J. Economic Literature*, XLI, 478-539, 2003.