

# 基于 Expectile 和 Realized GARCH 模型的波动率预测

高雷阜, 李伟梅  
(辽宁工程技术大学 运筹与优化研究院, 辽宁 阜新 123000)

**摘 要:**Realized GARCH 模型是预测波动率的经典模型之一,最小化非对称二次损失函数的 Expectile 对收益率尾部分布更加敏感,我们在 Realized GARCH 模型的基础上引入 Expectile 提出 Expectile-Realized GARCH 模型。以沪深 300 指数的高频收益率为例建模分析,对比不同模型下的波动率预测效果,发现 Expectile-Realized GARCH 模型较 Realized GARCH 模型对波动率预测能力更好。其中,当风险水平为 95% 时,对应的 Expectile-Realized GARCH 波动率预测能力最好。  
**关键词:**波动率预测;Expectile;Realized GARCH 模型;高频数据  
**中图分类号:**0213.2 **文章标识码:**A **文章编号:**1007-3221(2022)02-0099-05 **doi:**10.12005/orms.2022.0049

Volatility Prediction Based on Expectile and Realized GARCH Model  
GAO Lei-fu , LI Wei-mei  
(Institute for Optimization and Decision Analytics, Liaoning Technical University, Fuxin 123000, China)

**Abstract:**The Realized GARCH model is one of the classical models to predict the volatility, and the expectile based on minimizing asymmetric quadratic loss function is more sensitive to the tail distribution of return rate. We establish Expectile-Realized GARCHmodel based on Realized GARCH and expectile. We model and analyze the high frequency return rate of Shanghai and Shenzhen 300 index, and compare the volatility prediction results under different models. It is found that the Expectile-Realized GARCHmodel is better than Realized GARCH-model in forecasting volatility. When the risk level is 95%, the predictive volatility is the best.  
**Key words:**volatility prediction; Realized GARCH model; expectile; high frequency data

## 0 引言

收集、分析风险相关信息,进行风险评估,然后制定、实施相应的对策,使风险被控制在所能接受的范围内,这是金融风险管理的主要内容。其中,风险评估在风险管理中扮演着重要的角色,而风险预测对风险评估的准确性有重大影响。因此,在当代经济复杂的环境下,如何精确预测风险是风险管理中不可或缺的研究内容。

波动率是评估风险的核心方法之一。在金融时间序列早期的相关研究中,研究者发现风险预测误差与金融市场的波动相关。1982 年 Engle<sup>[1]</sup>提出的条件异方差模型(ARCH 模型)刻画了这种相关性,但 ARCH 模型待估计的参数繁多不利于快

速计算。为了修正 ARCH 模型,Bollerslev<sup>[2]</sup>提出的 GARCH 模型被称为广义 ARCH 模型,克服了 ARCH 模型的不足。自此之后,GARCH 模型被广大研究者推广并提出很多改进模型,如 TGARCH 模型、EGARCH 模型等,但这些模型多建立在低频数据上,预测的波动率精确度有限。随着科技的发展,高频数据逐渐走进风险预测研究领域。Andersen<sup>[3]</sup>等人基于高频数据,提出已实现方差 RV,证实了 RV 更加接近实际风险波动。2012 年 Hasen<sup>[4]</sup>将 RV 与 GARCH 模型结合提出了 Realized GARCH 模型(下文简称 RGARCH),进一步提高了波动率预测效能。与此同时,其他研究者相继将已实现波动率用于各种 GARCH 衍生模型。其中,王天一和黄卓<sup>[5]</sup>研究了高频数据下波动率的建模及应用。GARCH 及 GARCH 衍生模型对波动

率的预测都需要假设模型分布,但是满足模型条件的分布很难找到。分位数回归的 GARCH 模型不需要假定模型分布,将分位数回归引入波动率预测模型<sup>[6,7]</sup>,可缓解分布假设造成的模型误差,但分位数无法准确捕捉分布尾部信息。Expectile(期望分位数)由 Newey Powell<sup>[8]</sup>提出,和分位数具有一一对应的关系,Expectile 比分位数对分布的尾部信息更加敏感<sup>[9,10,11]</sup>。将 Expectile 和高频数据引进 GARCH 模型是否会碰撞出更精彩的波动率预测结果目前还未有研究。

本文基于分位数回归 GARCH 模型的思想及 Expectile、高频数据的良好性质,提出 Expectile-RealizedGARCH(简称 E-RGARCH)模型,旨在提高波动率预测能力,使得风险预测更加准确。

# 1 模型

## 1.1 RGARCH(1,1)模型<sup>[12]</sup>

$$\begin{cases} r_t = \mu_t + \sqrt{h_t} \varepsilon_t \\ h_t = \omega + \alpha z_{t-1} + \beta h_{t-1} \\ z_t = \lambda + \delta h_t + \tau(\varepsilon_t) + u_t \end{cases} \quad (1)$$

其中, $r_t$ 为资产收益率, $\mu_t$ 为 $r_t$ 的均值, $h_t = E(r_t^2 | F_{t-1})$ 是条件波动率, $F_{t-1}$ 是 $t-1$ 时刻的信息集, $\varepsilon_t$ 为标准化残差,满足 $\varepsilon_t^{i.i.d.} \sim std$ 。 $z_t$ 是已实现波动率测度, $\tau(\varepsilon_t) = \eta_1 \varepsilon_t + \eta_2 (\varepsilon_t^2 - 1)$ , $(\eta_1 < 0, \eta_2 \geq 0)$ 是杠杆函数,用来描述正负收益率对波动率的不对称影响,且 $E[\tau(\varepsilon_t)] = 0$ ,残差项 $\mu_t^{i.i.d.} \sim N(0, \sigma_{u_t}^2)$ , $\varepsilon_t$ 和 $u_t$ 相互独立。

## 1.2 E-RGARCH(1,1)模型

对于指定的风险水平 $\hat{\alpha}$ 和给定的随机变量 $x \in X$ ,在满足一阶平稳的条件下, $x$ 的第 $\tau$ 个 Expectile 值为<sup>[9]</sup>

$$\hat{\mu}_\tau = \min_u E\{\rho_{\hat{\alpha}}(x-u) - \rho_{\hat{\alpha}}(x)\} \quad (2)$$

其中,

$$\rho_{\hat{\alpha}}(x) = \begin{cases} \hat{\alpha} x^2 & x \geq 0 \\ (1 - \hat{\alpha}) x^2 & x < 0 \end{cases} \quad (3)$$

为非对称平方损失函数。 $\tau$ 为<sup>[11]</sup>

$$\tau = \frac{\hat{\alpha} \cdot q_x(\hat{\alpha}) - \int_{-\infty}^{q_x(\hat{\alpha})} x dF_X(x)}{2E(x) - 2 \int_{-\infty}^{q_x(\hat{\alpha})} x dF_X(x) - (1 - 2\hat{\alpha}) q_x(\hat{\alpha})} \quad (4)$$

$q_x(\hat{\alpha})$ 为 $x$ 的 $\hat{\alpha}$ 分位数, $F_X(x)$ 为 $x$ 的密度函数。Expectile 和分位数之间具有一一对应的关系,Expectile与自变量尾部分布和自变量值都有关系,对尾部分布更加敏感。Expectile 回归可以估计出数据在不同分位点的信息。

在 RGARCH(1,1)模型的基础上引入Expectile 得到 E-RGARCH(1,1)模型

$$\begin{cases} r_t = \mu_t + \sqrt{h_t} \varepsilon_t \\ h_t = \omega_\tau + \alpha_\tau z_{t-1} + \beta_\tau h_{t-1} \\ z_t = \lambda_\tau + \delta_\tau h_t + \tau(\varepsilon_t) + u_t \end{cases} \quad (5)$$

其中, $\omega_\tau, \alpha_\tau, \beta_\tau, \lambda_\tau, \delta_\tau$ 分别为 $\omega, \alpha, \beta, \lambda, \delta$ 的第 $\tau$ 个 Expectile 值。

## 2 模型估计

设 $\varepsilon_t \sim^{i.i.d.} std$ ,结合公式(1)中的 RGARCH 模型,令 $\theta = (\omega, \alpha, \beta, \eta_1, \eta_2, \delta, \lambda) std$ 为分布下 RGARCH(1,1)模型的参数向量。基于极大似然估计法<sup>[12]</sup>,可得风险水平 $\hat{\alpha}$ 下的最优参数 $\hat{\theta}$ 。根据 $\varepsilon_t$ 的标准误差,应用期望分位数回归<sup>[7]</sup>得到模型在不同分位点 $\tau$ 下的最优参数,公式如下

$$\hat{\theta}_{\hat{\alpha}} = \arg \min_{\theta} E[\rho_{\hat{\alpha}} \frac{r_t - \mu_t}{\sqrt{h_t}}] \quad (6)$$

## 3 条件波动率预测与评价

### 3.1 样本外动态波动率滚动预测

采用固定时间窗口宽度的样本外 1 步滚动预测法<sup>[13]</sup>对不同模型的条件波动率进行预测。具体方法为:将样本数据分为 $M$ 个样本内和 $N$ 个样本外,样本内用于参数估计,样本外用于预测。当一个新的观测值增加到样本中,删除第一个观测值并重新估计模型,得到一个新的波动率预测值,重复进行 $N$ 次,得到 $N$ 个样本外波动率预测结果。

根据均方差(MSE)<sup>[14]</sup>

$$MSE = \frac{1}{N} \sum_{t=1}^N (h_t - RV_t)^2 \quad (7)$$

评价模型预测能力。其中, $h_t$ 为第 $t$ 个条件波动率预测值, $RV_t$ 为第 $t$ 个已实现波动率。MSE 值越小,代表预测能力越好。

## 4 实证研究

选取沪深 300 指数 2017 年 10 月 13 日至 2020 年 7 月 1 日共 659 个交易日的高频 5min 历史收盘价为数据样本,数据来源于 wind,仿真模拟过程借助 R 语言 3.5.2 中的 rugarch 等包完成。基于高频数据,根据 $RV_t = \sum_{d=1}^{48} r_{t,d}^2$ 得到已实现波动率 $RV$ ,其中 $r_{t,d} = \ln p_{t,d+1} - \ln p_{t,d}$ 为对数收益率。 $p_{t,d}$ 表示第 $t$ 个

交易日的第  $d$  个 5min 收盘价<sup>[15,16]</sup>。基于高频数据的日收益率记为  $R_t$ ,  $R_t$  和  $RV$  的时间序列趋势图如图 1、图 2 所示。

图 1 中日收益率  $R_t$  表现出尖峰厚尾特征,图 2 中已实现波动率  $RV$  具有波动聚集性。

4.1 数据预处理

分析日收益率  $R_t$  的数据特征,判断能否应用 RGARCH 模型对其进行预测。

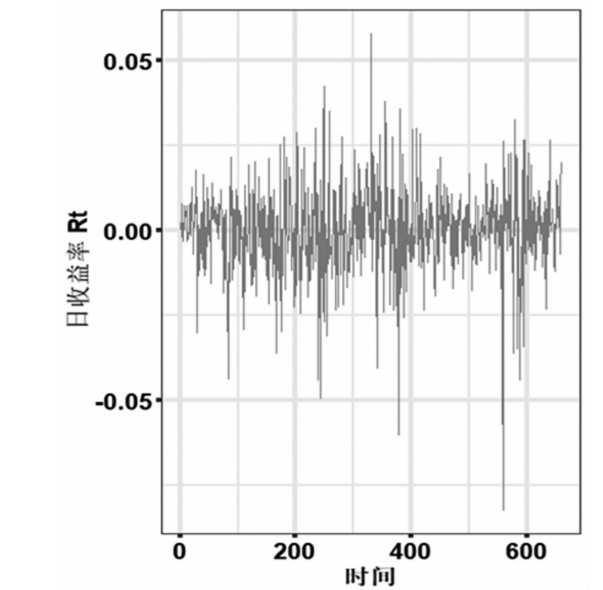
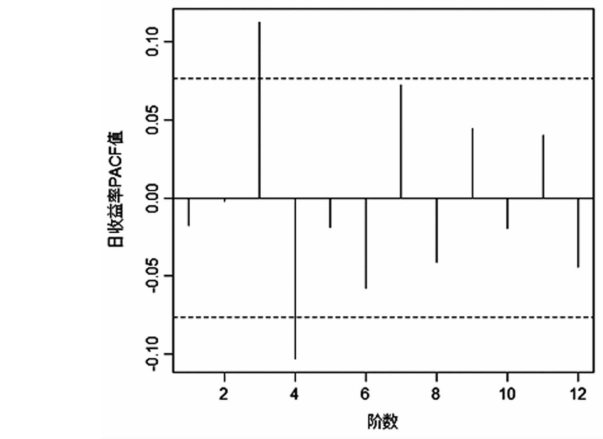


图 1 日收益率  $R_t$  时间序列图

表 1 中日收益率具有向 0 集中的趋势,收益率分布不分散。偏度系数结果表示分布偏左。样本峰度说明分布具有高峰厚尾。JB 检验统计量  $P$  值小于 0.05,说明不服从正态分布。ADF 检验结果



4.1.1 描述性统计量

分析日收益率的描述性统计特征,结果如表 1 所示

表 1 日收益率描述性统计量					
均值	标准差	偏度	峰度	J-B	ADF
0.0001	0.0131	-0.5590	4.0027	0	0.010

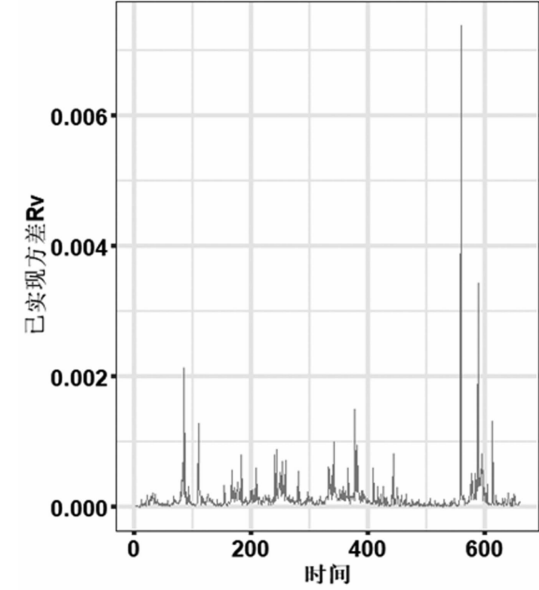


图 2 已实现波动率  $RV$  时间序列图

说明  $R_t$  为平稳序列。

4.1.2 自相关性检验

应用自相关 ACF 和偏自相关 PACF 可对日收益率  $R_t$  进行相关性分析,结果见图 3。

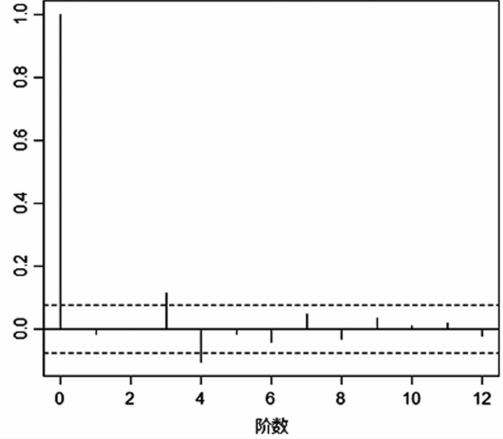


图 3 日收益率  $R_t$  的 ACF 和 PACF 检验

根据图 3, ACF 和 PACF 函数值基本在置信区间内(图中虚线区域内),说明日收益率  $R_t$  自相关性较弱。

Box-Ljung 检验<sup>[16,17]</sup>是对时间序列是否存在滞后相关的一种统计检验,检验日收益率  $R_t$  滞后相

关性,若不相关,则为白噪声序列。

表 2 日收益率 $R_t$ 的 Box-Ljung 检验	
Box-Ljung $R_t$ 检验	
Data: $R_t$	
X-squared = 16.152, df = 5, p - value = 0.006424	

根据表 2,滞后阶为 12 时  $P$  值大于 5%,日收益率  $R_t$  为白噪声序列。

结合图 3 和表 2,我们认为沪深 300 指数的日收益率  $R_t$  不存在显著高阶自相关性,可以应用 RGARCH(1,1)模型对其分析研究。

4.2 模型建立

4.2.1 均值模型的建立

由于日收益率  $R_t$  序列不存在显著的高阶自相关性,设定均值方程的形式为: $r_t = \mu + \sqrt{h_t} \varepsilon_t$ ,  $\mu$  为常数,此处设  $\mu = 0$ 。RGARCH(1,1)模型为

$$\begin{cases} r_t = \sqrt{h_t} \varepsilon_t \\ h_t = \omega + \alpha z_{t-1} + \beta h_{t-1} \\ z_t = \lambda + \delta h_t + \tau(\varepsilon_t) + u_t \end{cases} \quad (8)$$

在的 RGARCH(1,1)模型引入 Expectile,得到相应的 E-RGARCH(1,1)模型为

$$\begin{cases} r_t = \sqrt{h_t} \varepsilon_t \\ h_t = \omega_\tau + \alpha_\tau z_{t-1} + \beta_\tau h_{t-1} \\ z_t = \lambda_\tau + \delta_\tau h_t + \tau(\varepsilon_t) + u_t \end{cases} \quad (9)$$

4.2.2 模型参数估计

选取前 559 个交易日的日收益率  $R_t$  作为样本内数据估计模型参数,得到的结果如表 3 和表 4 所示。

表 3 RGARCH(1,1)模型参数估计结果

参数	RGARCH(1,1)模型
$\omega$	1.8542 <sup>-</sup>
$\alpha$	0.3580 <sup>-</sup>
$\beta$	0.6134 <sup>-</sup>
$\eta_1$	-0.0586 <sup>--</sup>
$\eta_2$	0.1803 <sup>-</sup>
$\delta$	0.8652 <sup>-</sup>
$\lambda$	0.5701 <sup>-</sup>

表 4 E-RGARCH(1,1)模型参数估计结果

分位点 $\tau$	0.29	0.57	0.61
置信水平 $\hat{\alpha}$	0.99	0.95	0.90
$\omega$	1.8542 <sup>-</sup>	1.8596 <sup>-</sup>	1.8546 <sup>-</sup>
$\alpha$	0.3580 <sup>-</sup>	0.3592 <sup>-</sup>	0.3582 <sup>-</sup>
$\beta$	0.6134 <sup>-</sup>	0.6135 <sup>-</sup>	0.6133 <sup>-</sup>
$\eta_1$	-0.0586 <sup>--</sup>	-0.0587 <sup>--</sup>	-0.0582 <sup>--</sup>
$\eta_2$	-0.1803 <sup>-</sup>	0.1806 <sup>-</sup>	0.1802 <sup>-</sup>
$\delta$	0.8652 <sup>-</sup>	0.8622 <sup>-</sup>	0.8649 <sup>-</sup>
$\lambda$	0.5701 <sup>-</sup>	0.5708 <sup>-</sup>	0.5697 <sup>-</sup>

注:参数估计结果中<sup>-</sup>表示1%显著水平,<sup>--</sup>表示5%显著水平。

4.2.3 条件波动率预测和评价

在以上参数估计基础上,将 2017 年 10 月 13 日至 2020 年 2 月 3 日的 559 个数据视作样本内数据,2020 年 2 月 4 日至 2020 年 7 月 1 日的 100 个数据视作样本外数据,对 RGARCH(1,1)模型和 E-RGARCH(1,1)模型的条件波动率进行预测,其预测结果与样本外已实现方差的对比如图 4 所示。

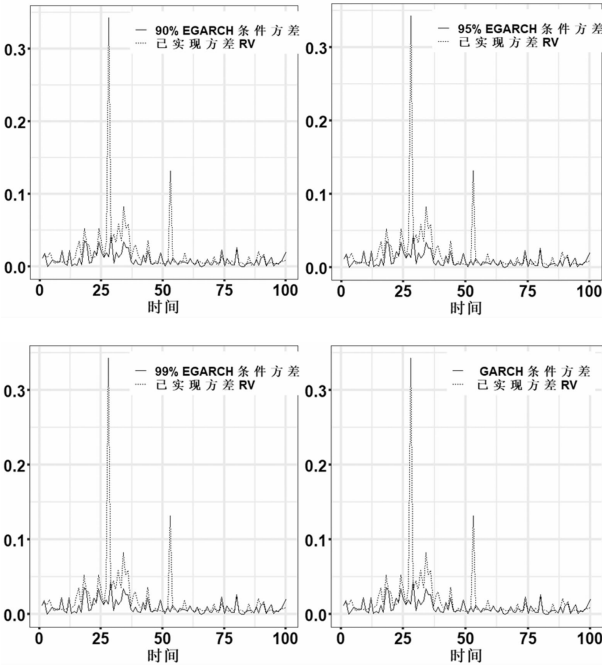


图 4 RGARCH(1,1)模型和不同风险水平下的 E-RGARCH(1,1)模型条件波动率预测图

图 4 中 RGARCH(1,1)模型和 E-RGARCH(1,1)模型对条件波动率预测都有较好的效果,以下应用均方差(MSE)对预测效能作进一步评价,结果如表 5 所示。

表 5 不同模型的 MSE 值

模型	$\hat{\alpha}$	$\tau$	MSE
RGARCH(1,1)	-	-	1.796807
E-RGARCH(1,1)	0.90	0.61	1.6364
E-RGARCH(1,1)	0.95	0.57	1.567238
E-RGARCH(1,1)	0.99	0.29	1.576889

根据表 5 结果,E-RGARCH(1,1)模型波动率预测能力均比 RGARCH(1,1)模型好,当风险水平为 0.95 时,E-RGARCH(1,1)的波动率预测能力最好。

5 结论

基于高频数据的 E-RGARCH 模型较传统的 RGARCH 模型对波动率预测更加准确的,这对金融市场风险管理有一定的指导意义。本文采用极

大似然法估计模型参数,并没有考虑参数的约束信息,而贝叶斯估计法假定了参数的先验分布,使得参数估计效果更佳。因此,基于贝叶斯估计的 E-RGARCH 模型可作为未来研究方向。

## 参考文献:

- [1] Engle R F. Autoregressive conditional heteroscedasticity with estimates of the variance of U. K. Inflation [J]. *Econometrica: Journal of the Econometric Society*, 1982, 4(50): 988-1003.
- [2] Bollerslev T. Generalized autoregressive conditional heteroskedasticity [J]. *Eeri Research Paper*, 1986, 31(3): 307-327.
- [3] Andersen T G, Bollerslev T. Answering the skeptics: yes, standard volatility models do provide accurate forecasts [J]. *International Economic Review*, 1998, 39(4): 885-905.
- [4] Hansen P R, Huang Z, Shek H H. Realized GARCH: a joint model for returns and realized measures of volatility [J]. *Journal of applied econometrics*, 2012, 27(6): 877-906.
- [5] 王天一,黄卓. 基于高频数据的波动率建模及应用研究评述 [J]. *经济学动态*, 2012, (3): 142-145.
- [6] 解其昌. 分位数回归方法及其在金融市场风险价值预测中的应用 [D]. 四川:西南财金大学学位论文, 2012: 60-112.
- [7] 林琦,欧思歆. 基于 GARCH 族类模型的开放式基金收益率风险分析 [J]. *数学的实践与认识*, 2019, 49(24): 109-113.
- [8] Newey W K, Powell J L. Asymmetric least squares estimation and testing [J]. *Econometrica*, 1987, 55(4): 819-847.
- [9] 赵军. Expectile 回归和最优资产组合中的变量选择问题 [D]. 浙江:浙江大学学位论文, 2017: 23-26.
- [10] 许启发,丁晓涵. 基于 Expectile 回归的均值—ES 组合投资决策 [J]. *中国管理科学*, 2018, 26(10): 21-26.
- [11] Jones M C. Expectiles and m-quantiles are quantiles [J]. *Statistics & Probability Letters*, 1994, 20(2): 149-153.
- [12] 黄友珀,唐振鹏,周熙雯. 基于偏 t 分布 realized GARCH 模型的尾部风险估计 [J]. *系统工程理论与实践*, 2015, 35(9): 2201-2206.
- [13] 陈王,马锋. 高频视角下中国股市动态 VAR 模型研究 [J]. *运筹与管理*, 2020, 29(2): 85-190.
- [14] 杨科,陈浪楠. 股市波动率的短期预测模型和预测精度评价 [J]. *管理科学学报*, 2012, 15(5): 25-27.
- [15] 李洪成,尚秀芬,郝瑞丽. 金融数据分析导论 [M]. 北京:机械工业出版社, 2013.
- [16] 吴鑫育,王莉莉. 基于高频数据的中国股市波动率预测研究 [J]. *东北农业大学学报(社会科学版)*, 2019, 17(3): 4-9.
- [17] 樊鹏英,兰勇. 高频数据下基于 PGARCH 模型的 VAR 估计方法及应用 [J]. *系统工程理论与实践*, 2017, 37(8): 2053-2058.