Bayesian Workflow

Quan Tran

2024-06-02

1 Bayesian Workflow

1.1 Loading Data and Preprocessing

In this section, we fetched the data from Yahoo Finance, calculated necessary quantities such as squared returns and daily realized volatility.

2 Week 1: Exploratory Data Analysis and Choosing a Research Question

2.1 Goal

After this week, you should have:

- Setting up your project, for example, using the provided templates
- Formulating a research question & finding a dataset
- Visualising and getting familiar with characteristics of your data (e.g., range, data types)
- Adding your first notes and visualisations to the workflow diary
- Picking an initial model & documenting your reasoning and the strategies you used to choose it
- Obtaining posterior samples using your initial model with default priors
- Documenting what you observe and any issues you encounter in the workflow diary

3 Introduction

3.1 Volatility Explained

Volatility is a statistical measure used to evaluate the dispersion of returns for a given security or market index. Usually, riskier securities have higher corresponding volatility.

Volatility also often refers to the amount of uncertainty or risk related to the size of changes in a security's value. A higher volatility means that a security's value can potentially be spread out over a larger range of values. This means that the price of the security can change dramatically over a short time period in either direction. A lower volatility means that a security's value does not fluctuate dramatically, and tends to be more steady.

3.2 Volatility types

There are two types of volatility: Implied volatility and Realized volatility. Implied volatility (IV), also known as projected volatility, is one of the most important metrics for options traders. On the other hand, Realized volatility measures what actually happened in the past.

3.3 Research question

The research question this project attempts to answer is whether we can model the volatility of gold price using on the daily returns. ARCH models will be utilized to achieve this goal since they are commonly employed in modeling financial time series that exhibit time-varying volatility and volatility clustering, i.e. periods of swings interspersed with periods of relative calm.

3.4 ARCH Model

Autoregressive conditional heteroskedasticity (ARCH) model is a statistical model for time series data that describes the variance of the current error term or innovation as a function of the actual sizes of the previous time periods' error terms; often the variance is related to the squares of the previous innovations.

Let r_t the return on day t, we have

$$r_t = \mu + \epsilon_t$$

where

- μ is some unknown constant
- $\epsilon(t)$ denote the error terms (return residuals, with respect to a mean process), i.e. the series terms.These $a(t)$ are split into a stochastic piece $\eta(t)$ and a time-dependent standard deviation σ_t characterizing the typical size of the terms so that

$$\epsilon_t = \sigma_t \eta_t$$

where

- $\sigma(t)$ is the time-dependent noise
- $\eta(t)$ stochastic noise, e.g., $\eta(t) \sim \mathcal{N}(0, 1)$

The random variable $\eta(t)$ is a strong white noise process. The series σ_t is modeled by

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, $i > 0$.

In this project, we only explore the ARCH[1] model due to its simplicity. Therefore, the likelihood of the model would be

$$r_i \sim \mathcal{N}\left(\mu, \sqrt{\alpha_0 + \alpha_1 * (r_{i-1} - \mu)^2}\right) \quad \text{for} \quad i \in \{1, ..., t\}$$

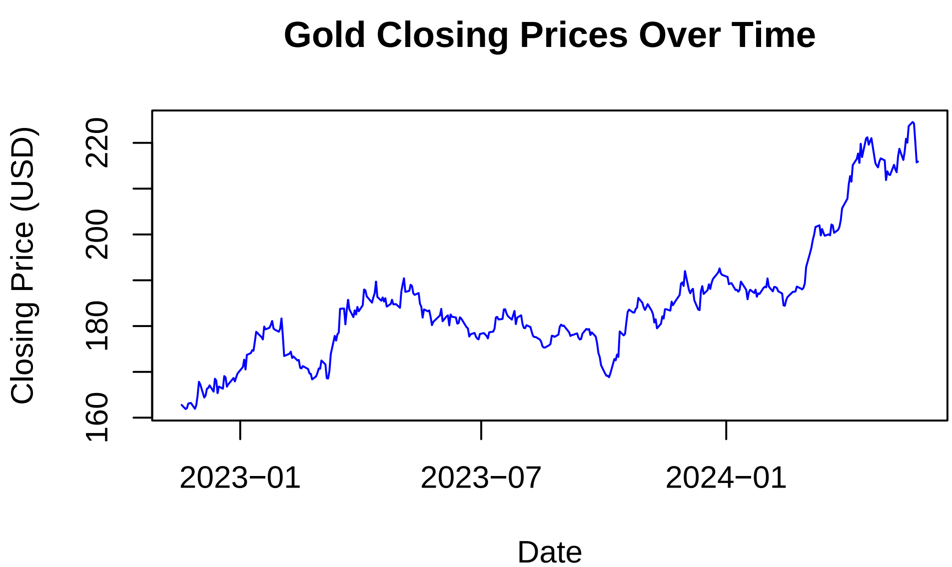
where

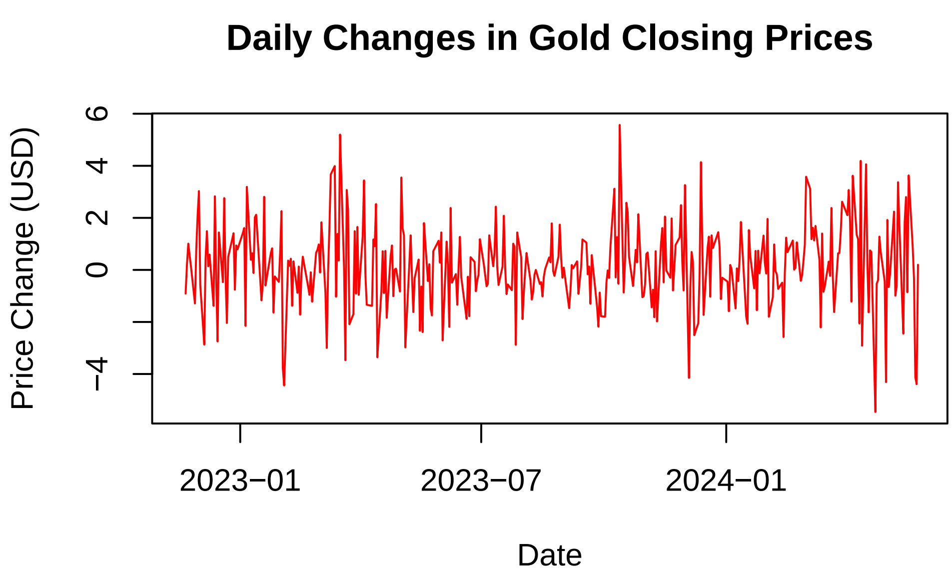
r_i is the return in percentage on day i and μ is the average return.

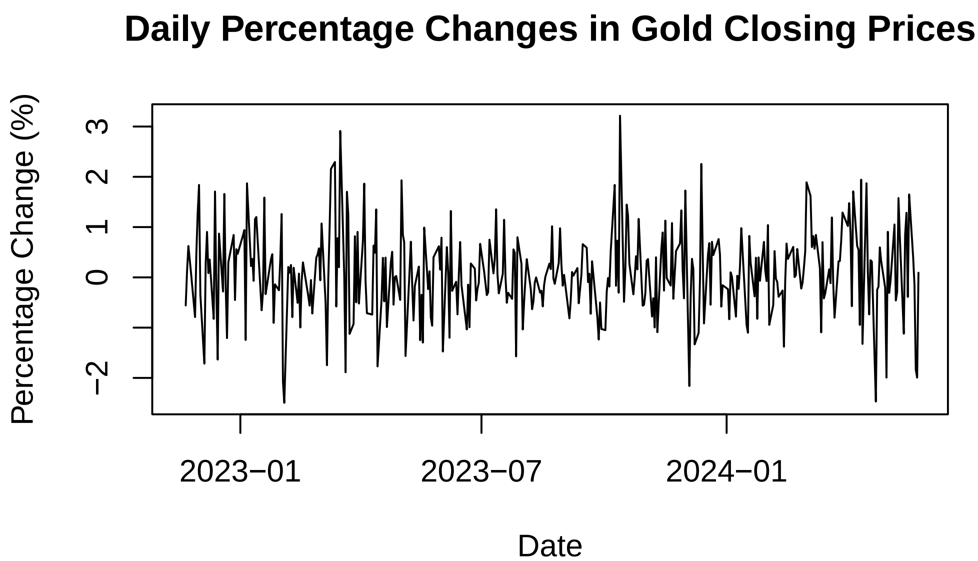
3.5 Exploratory Data Analysis

The dataset is then split into two, training set (304 rows) and testing set (76 rows).

Then we plot some exploratory visualization to gain some overview understanding of how the gold price has been behaving for the few periods in the training set.







As you may notice, there are many big swings in both direction, i.e., there are rises and falls of more than 1% over a sustained period of time, suggesting a volatile market.

4 Week 2: Prior Choice

4.1 Goal

After this week, you should have:

- Proposed priors for each parameter in your model, with justification
- Performed a prior predictive check to ensure that your priors are reasonable

4.2 Prior choices and justification

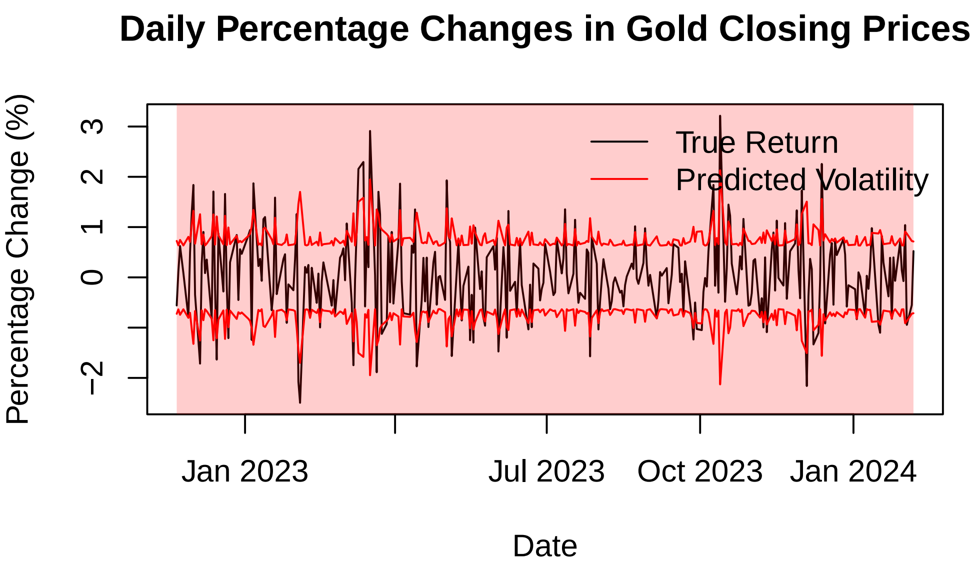
To ensure stationarity in the model, we constrain the values of the parameters α's to lie within the interval [0, 1], with a preference for values closer to 0. Accordingly, we have chosen to use Beta distributions with parameters ($\alpha = 2$) and ($\beta = 3$), effectively skewing the distribution towards 0 and accommodating the desired parameter behavior while not making the distribution too narrow.

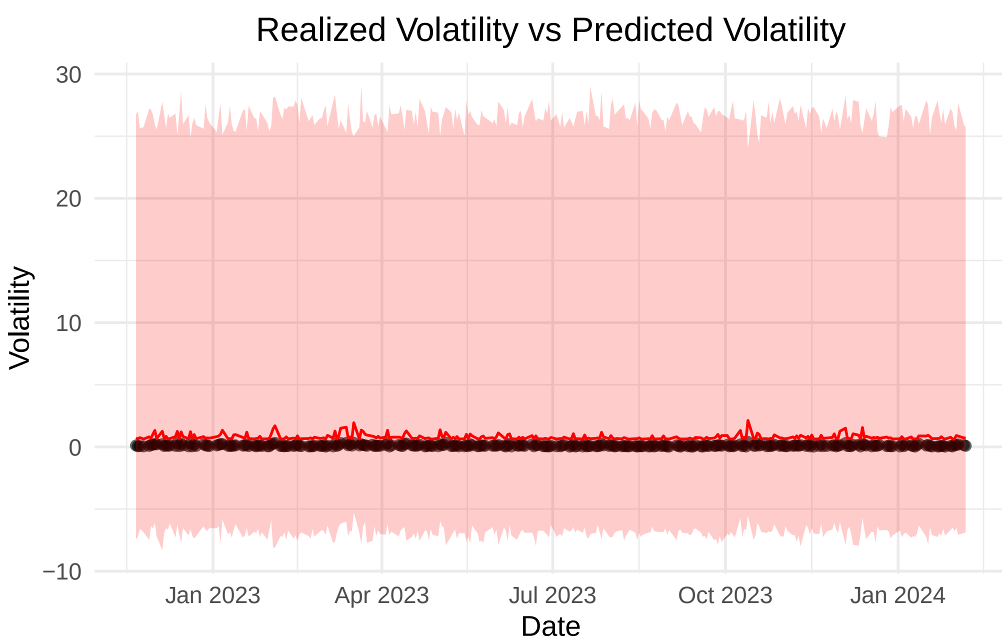
For the average return, we assume a weakly informative Normal prior.

$$\mu \sim \mathcal{N}(0, 10)$$

$$\alpha_i \sim Beta(2, 3)$$

Subsequently, we attempt to carry out the prior predictive checks to inspect whether the priors cover the feasible values - the realized volatility.





From the plots, we can see that the priors chosen covers the feasible range well, suggesting that we can incorporate these priors in the model.

5 Week 3: Model Fitting and Checking

5.1 Goal

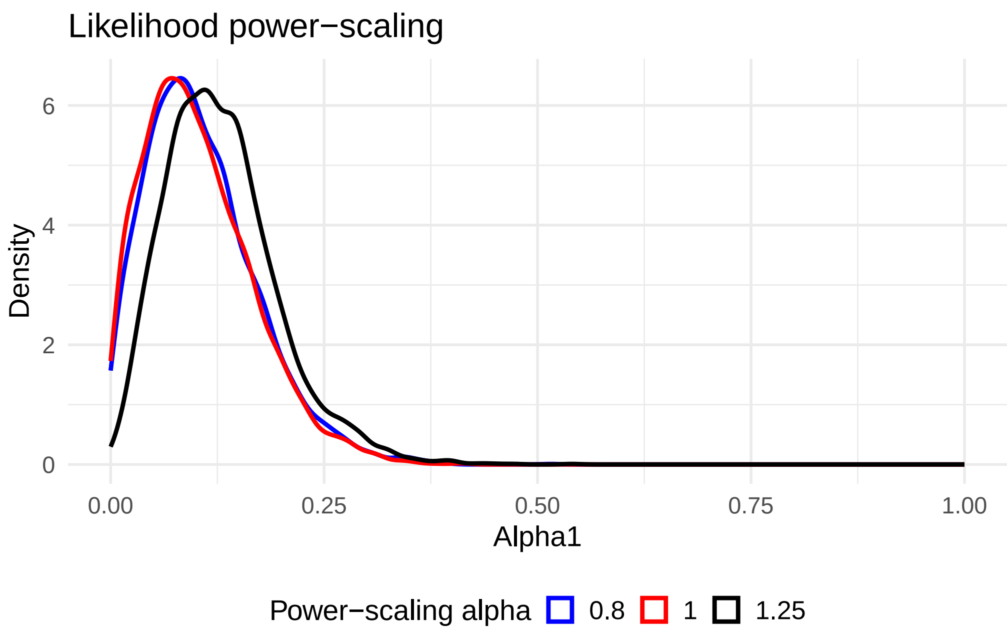
After this week, you should have:

- Fitted your model with chosen priors to your data
- Performed diagnostic checks for quality/stability of fitting
- Performed prior sensitivity assessment
- Performed predictive performance assessment

5.2 ARCH model fitting and checking

5.3 Likelihood Sensitivity Analysis

Due to limited time and computational resources, we only carry out the likelihood power-scaling analysis, leaving out the prior power-scaling one.



The posterior distribution of α_1 is somewhat sensitive to the choice of scaling factor. This sensitivity is evident in the slight shifts in the peaks and the changes in the spread of the distributions.

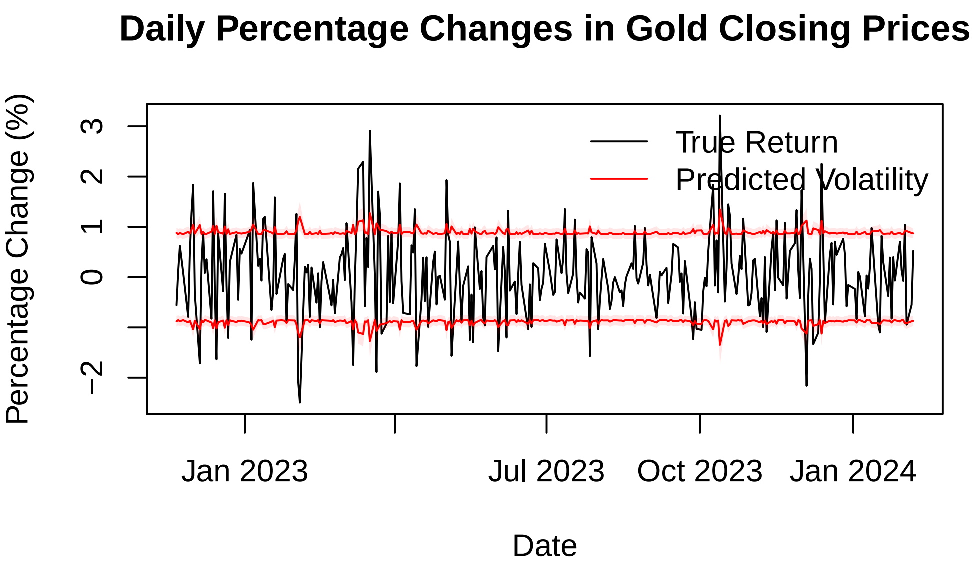
As the scaling factor increases (from 0.8 to 1.25), the variance of the posterior distribution appears to decrease slightly. This means the model becomes more confident in its estimate of α_1 when the likelihood is given more weight.

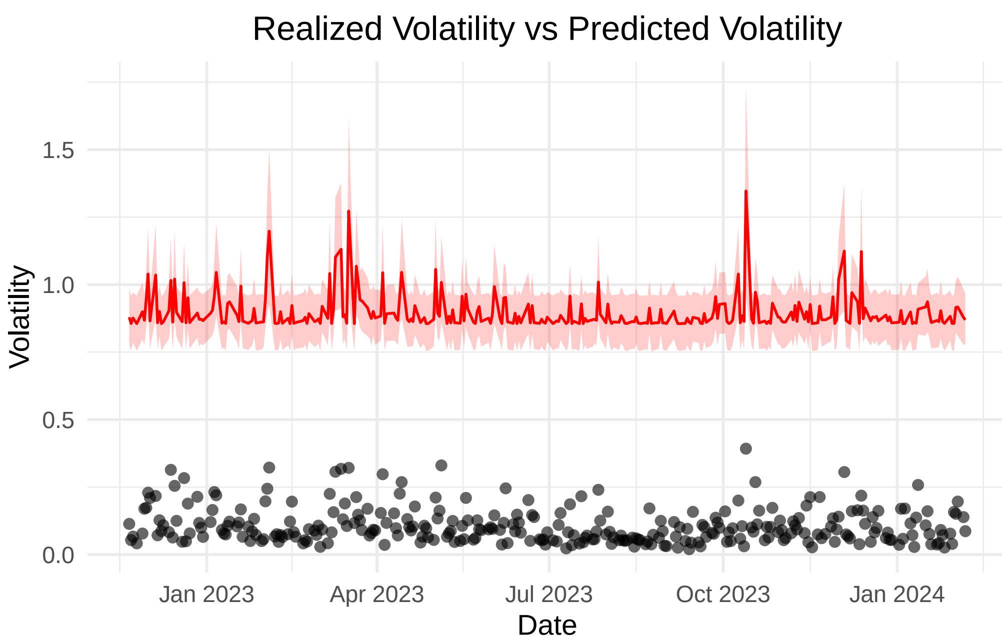
In conclusion, this suggests that the prior distribution may be playing a non-negligible role in shaping the posterior.

5.4 Posterior Predictive Checks

After fitting the model, we move on to evaluate how the model fit the data by plotting the predicted volatility over variability of the return.

We see that the model is a good fit, and the uncertainty is tiny. Please kindly note that they do not fit exactly because the black line and the red line are not showing exactly the same quantity. However, it is quite flat, which is common with ARCH models, and susceptible to bursts - spikes of volatility, which is not desirable.

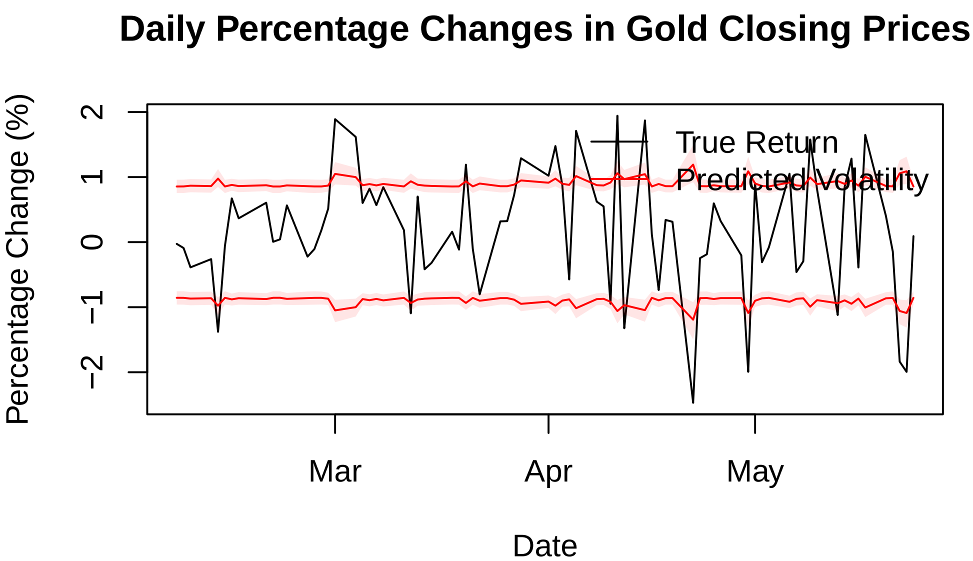


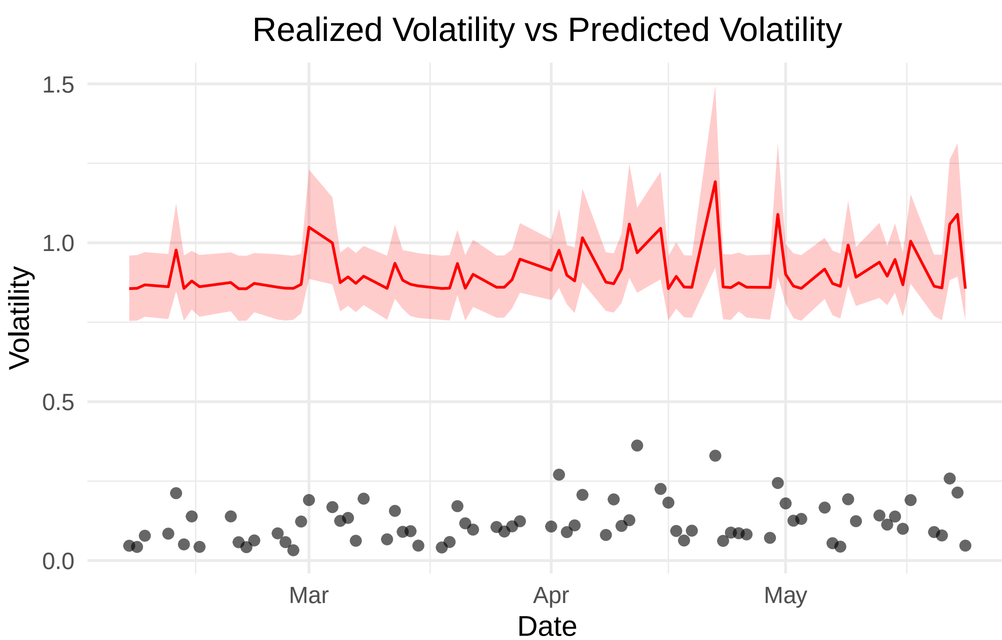


The predicted volatility and realized volatility show clear correlation. However, the gap between them is quite far and thus undesirable. This can be improved by incorporating previous volatility in the model, which will be introduced later on in the extended model.

5.5 Predictive Performance Visual Assessment

Subsequently, we let the model predict the volatility on unseen testing data to evaluate its predictive performance.





The predicted volatility and realized volatility exhibit some correlation, indicating that the model is a good fit. The big gap between the predicted and realized volatility can be explained that the realized volatility is also just a proxy of true volatility - intrinsic, underlying volatility of an asset that is not directly observable.

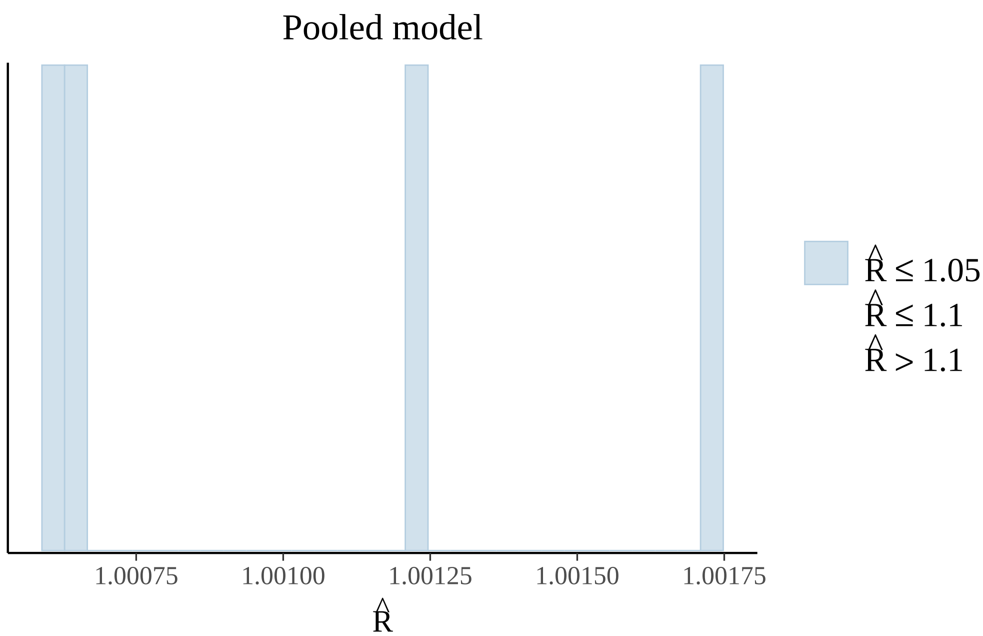
The uncertainty is also wider, reflecting the fact that the model is forecasting on unseen testing set.

5.6 Convergence Diagnostics

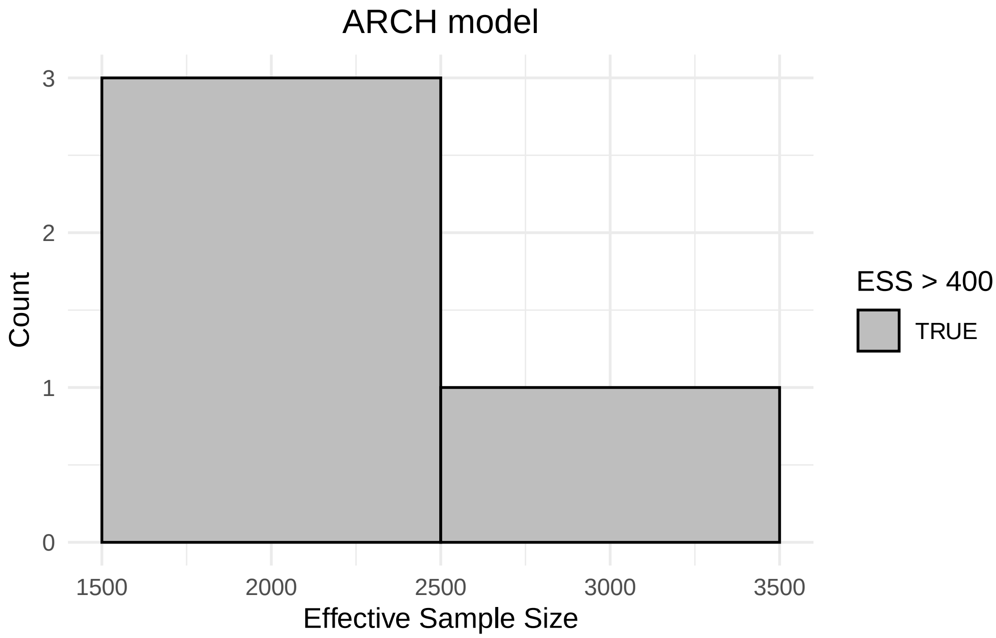
The convergence diagnostics ($\hat{R}$, ESS, divergence) for the ARCH model are all good.

```
Processing csv files: C:/Users/trana/AppData/Local/Temp/RtmpcDQb7A/arch-202406022141-1-6c4891.csv, C:/
```

```
Checking sampler transitions treedepth.
Treedepth satisfactory for all transitions.

Checking sampler transitions for divergences.
1 of 4000 (0.03%) transitions ended with a divergence.
These divergent transitions indicate that HMC is not fully able to explore the posterior distribution.
Try increasing adapt delta closer to 1.
If this doesn't remove all divergences, try to reparameterize the model.

Checking E-BFMI - sampler transitions HMC potential energy.
E-BFMI satisfactory.

Effective sample size satisfactory.

Split R-hat values satisfactory all parameters.

Processing complete.
```




```
Number of divergent transitions: 1
```



6 Week 4: Extending Models and Model Selection

6.1 Goal

After this week, you should have:

- Decided on whether a model expansion or selection approach is relevant for your research question, with justification
- Proposed a second model (or an expansion to the first), building on the issues/diagnostics/concepts from previous weeks

6.2 GARCH Model

A better way to model the heterogeneity is to use a generalized autoregressive conditional heteroskedasticity (GARCH) model. It is similar to the ARCH model, but it not only takes into account the previous data values, but also the previous volatility values. Therefore, this makes it less susceptible to bursts as the volatility is propagated over time.

The series σ_t is modeled by

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$, $i > 0$.

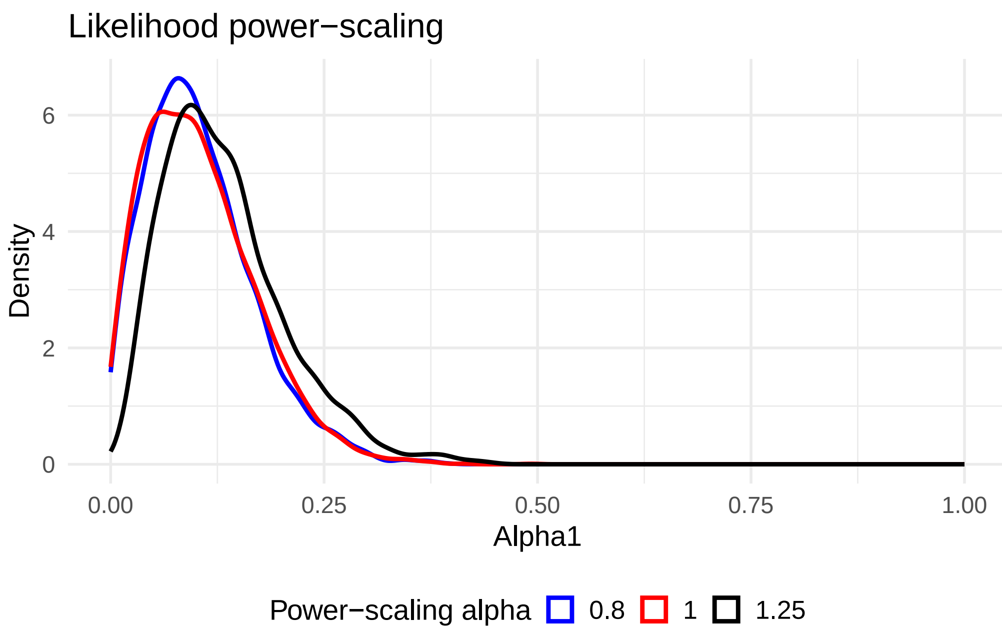
We only explore the GARCH[1] model in this project due to its simplicity. Therefore, the likelihood of the model would be

$$r_i \sim \mathcal{N}\left(\mu, \sqrt{\alpha_0 + \alpha_1 * (r_{i-1} - \mu)^2 + \beta_1 * \sigma_{i-1}^2}\right) \quad \text{for} \quad i \in \{1, ..., t\}$$

6.3 Prior choices and justification

The priors used and justification is similar to those in the ARCH Model.

6.4 Likelihood Sensitivity Analysis

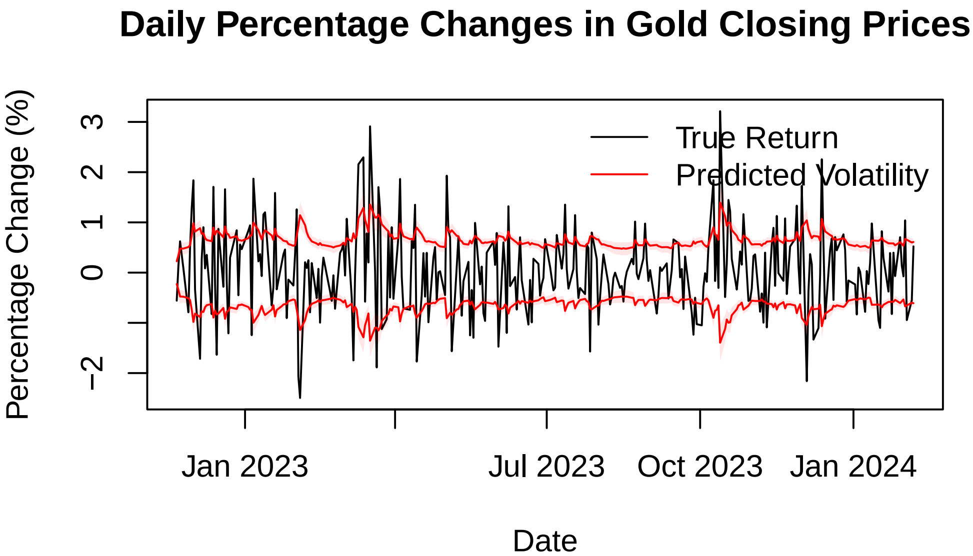


The posterior distribution of α_1 shows a moderate degree of sensitivity to the likelihood power-scaling factor. This can be seen in the shifts in the peak density and the changes in the spread of the distributions.

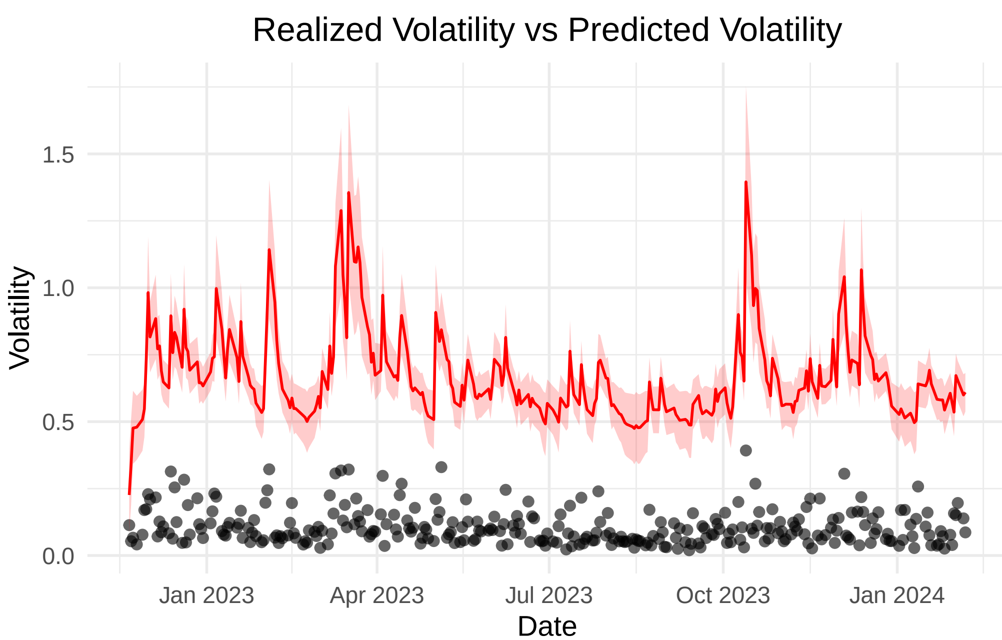
As the scaling factor increases, the variance of the posterior distribution decreases slightly. This means that the model becomes more confident in its estimate of α_1 when the likelihood is given more weight. Additionally, the peak (mode) of the distribution shifts slightly towards higher values as the scaling factor increases, suggesting that the data supports a slightly larger value of α_1 than the prior might have initially suggested.

6.5 Posterior Predictive Checks

After fitting the model, we move on to evaluate how the model fit the data by plotting the predicted volatility over variability of the return.

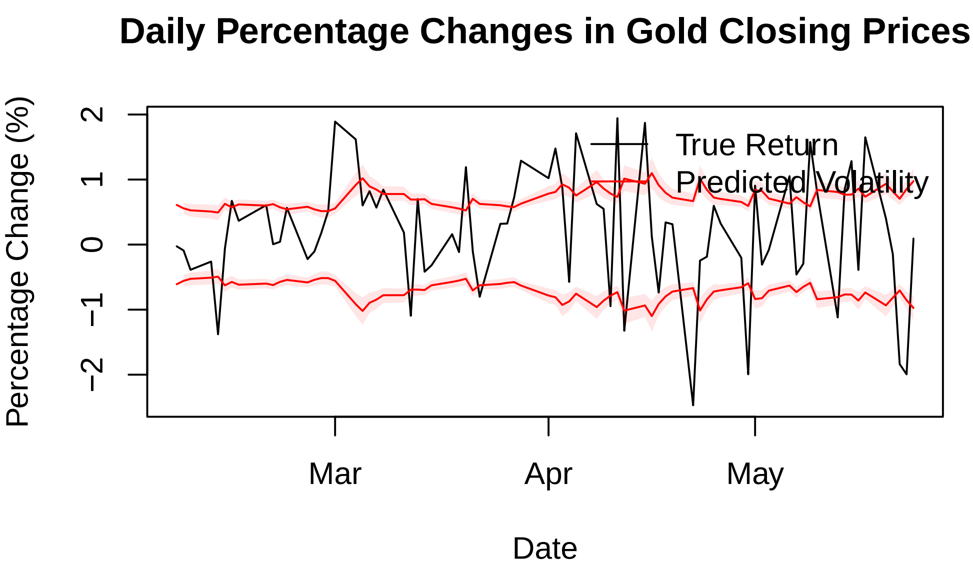


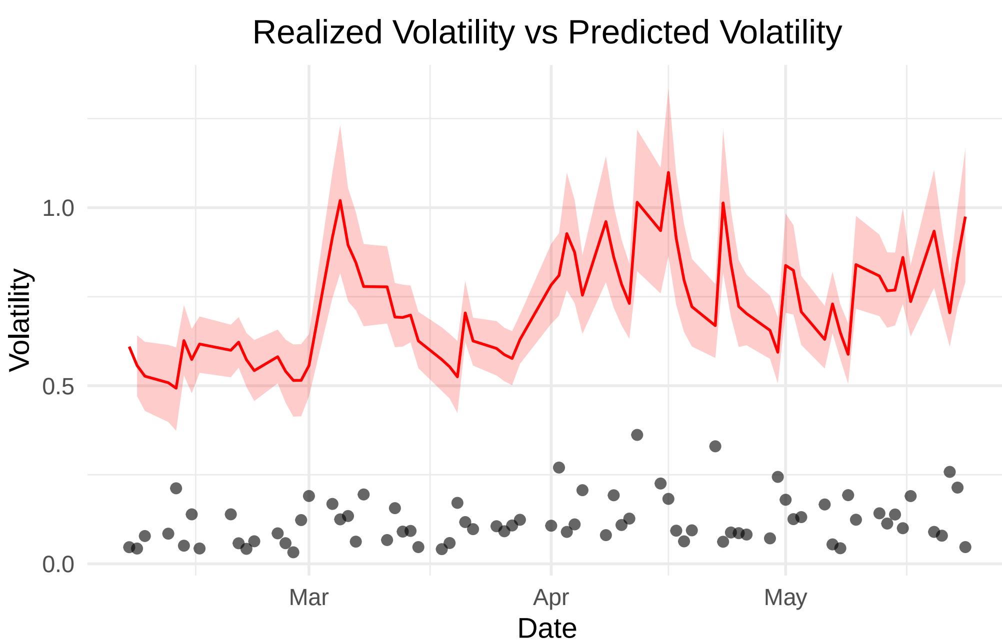
As depicted, we get a really nice correlation with the percentage changes. Compared with the ARCH model, the replicated volatility is now much smoother and closer-matching to the data. The uncertainty is also very tiny.



As you may notice, the gap between the replicated volatility of GARCH model and the realized volatility is much shorter than that of ARCH model.

6.6 Predictive Performance Visual Assessments





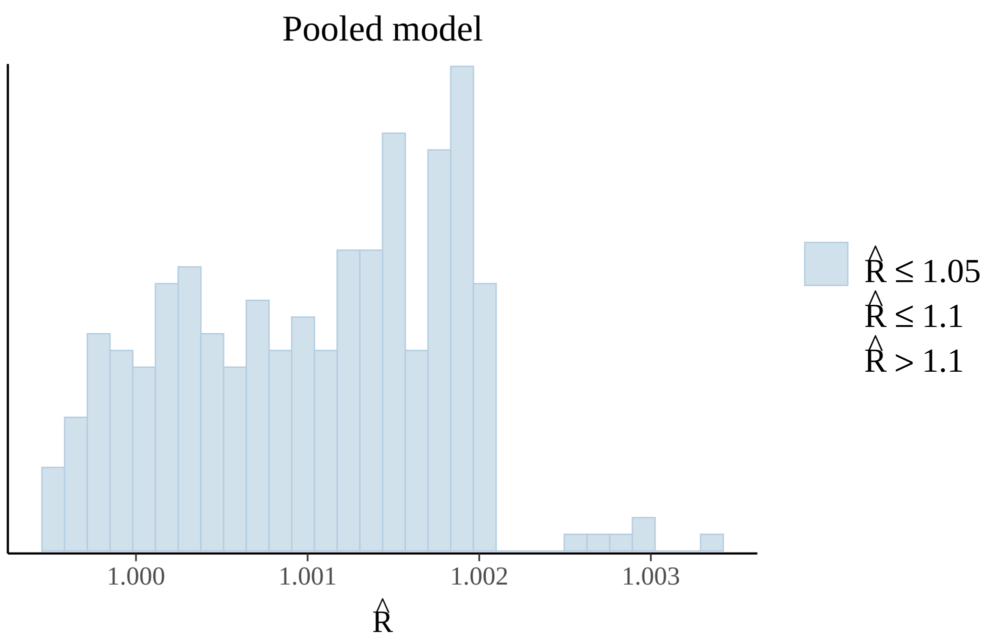
The uncertainty of the predictions on testing set is now wider than that on training set, reflecting the fact that there should be more uncertainty on unseen data. The gap between on prediction volatility and realized volatility is also much shorter than that in ARCH Model.

6.7 Convergence Diagnostics

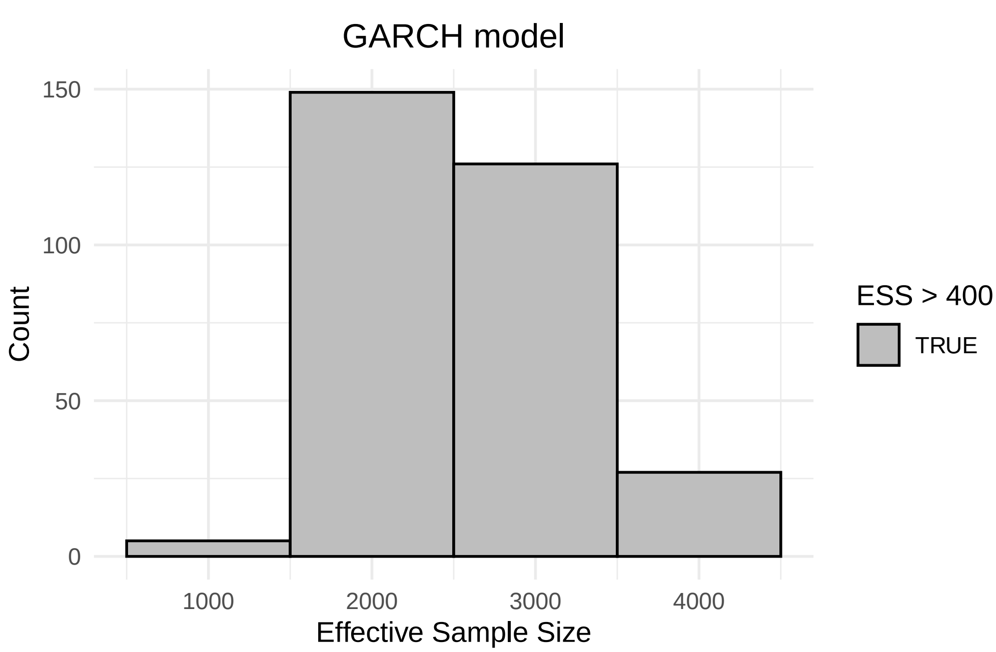
The convergence diagnostics ($\hat{R}$, ESS, divergence) for the ARCH model are all good.

```
Processing csv files: C:/Users/trana/AppData/Local/Temp/RtmpcDQb7A/garch-202406022141-1-77b650.csv, C:

Checking sampler transitions treedepth.
Treedepth satisfactory for all transitions.
```

```
Checking sampler transitions for divergences.
No divergent transitions found.

Checking E-BFMI - sampler transitions HMC potential energy.
E-BFMI satisfactory.

Effective sample size satisfactory.

Split R-hat values satisfactory all parameters.

Processing complete, no problems detected.
```




```
Number of divergent transitions: 0
```



7 Week 5: Interpreting and Presenting Model Results

7.1 Goal

After this week, you should have:

- Prepared a concise summary of your results and how they answer your research question
- Prepared a visualisation of your results that is suitable for presentation to a non-technical audience

7.2 Model Selection

We will decide which model out of the two is in favor based on their predictive performances.

7.3 Forecast Evaluatoin

Our measure of predictive accuracy is based on the average forecast loss achieved by a model/strategy/proxy triplet. A model that provides a smaller average loss is more accurate and therefore preferred. Choices for loss functions are extensive, and their properties vary Volatility forecast comparison can be tricky because forecasted values must be compared against an ex post proxy of volatility, rather than its true, latent value. Patton (2009) identifies a class of loss functions that is attractively robust in the sense that they asymptotically generate the same ranking of models regardless of the proxy being used. The Patton class is comprised of a continuum of loss functions indexed by a param- eter on the real line. It rules out all but two losses traditionally used in the volatility forecasting literature:

$$\begin{aligned}
\text{QL:} \quad & L(\hat{\sigma}_t, h_{t|t-k}) = \frac{\sigma_t^2}{h_{t|t-k}} - \log \frac{\sigma_t^2}{h_{t|t-k}} - 1 \\
\text{MSE:} \quad & L(\hat{\sigma}_t, h_{t|t-k}) = \left(\sigma_t^2 - h_{t|t-k}\right)^2
\end{aligned}$$

where σ_t^2 is an unbiased ex post proxy of conditional variance (such as realized volatility or squared returns) and $h_{t|t-k}$ is a volatility forecast based on t - k information (k > 0). The quasi-likelihood (QL) loss, named for its close relation to the Gaussian likelihood, depends only on the multiplicative forecast error, $\frac{\sigma_t^2}{h_{t|t-k}}$. The mean squared error (MSE) loss depends solely on the additive forecast error, $\sigma_t^2 - h_{t|t-k}$.

7.4 Metrics Result

```
             ARCH     GARCH
QL Loss 1.2870516 1.0855145
MSE     0.6188636 0.3863066
```

Since the GARCH model is performing better in terms of QL Loss and MSE error, we would select the GARCH model as the main model for practical uses and further development.

7.5 Conclusion

In conclusion, after evaluate the performances of the two models, we are in favor of GARCH model is forecasting volatility better in terms of numerical metrics and visual assessments.

However, this model is not perfect and can be further developed by incorporating more lag terms in the model. We also can introduce a hierarchical structure on time periods as periods may have different random noises.