Alder Capital
Investment Management

Sense and Volatility

Mark Caslin *points out the flaws in existing volatility forecasting methods, and proves why alternative routes are worth looking into*

The best-known toolkit for measuring risk has been provided and serviced by JP Morgan since 1994. At the time it seemed strange that they were providing this for free. But as the RiskMetrics home page says:

"RiskMetrics is based on, but differs significantly from, the market risk management systems developed by JP Morgan for its own use." I am not surprised that it differs significantly, since it is not very good, and probably too dangerous for a serious institution to rely on. What I intend to do in this article is draw attention to some of the areas for improving both RiskMetrics and its big forecasting rival, the Garch method.

The 'garbage in garbage out' rule

Good forecasting relies on the quality of the information available. Ironically, in forecasting future volatility, the problem is to find what volatility was in the past. Volatility can not be measured directly like a spot price; all we have are samples of the spot moves generated from that volatility. From these we must infer the level of volatility that was present.

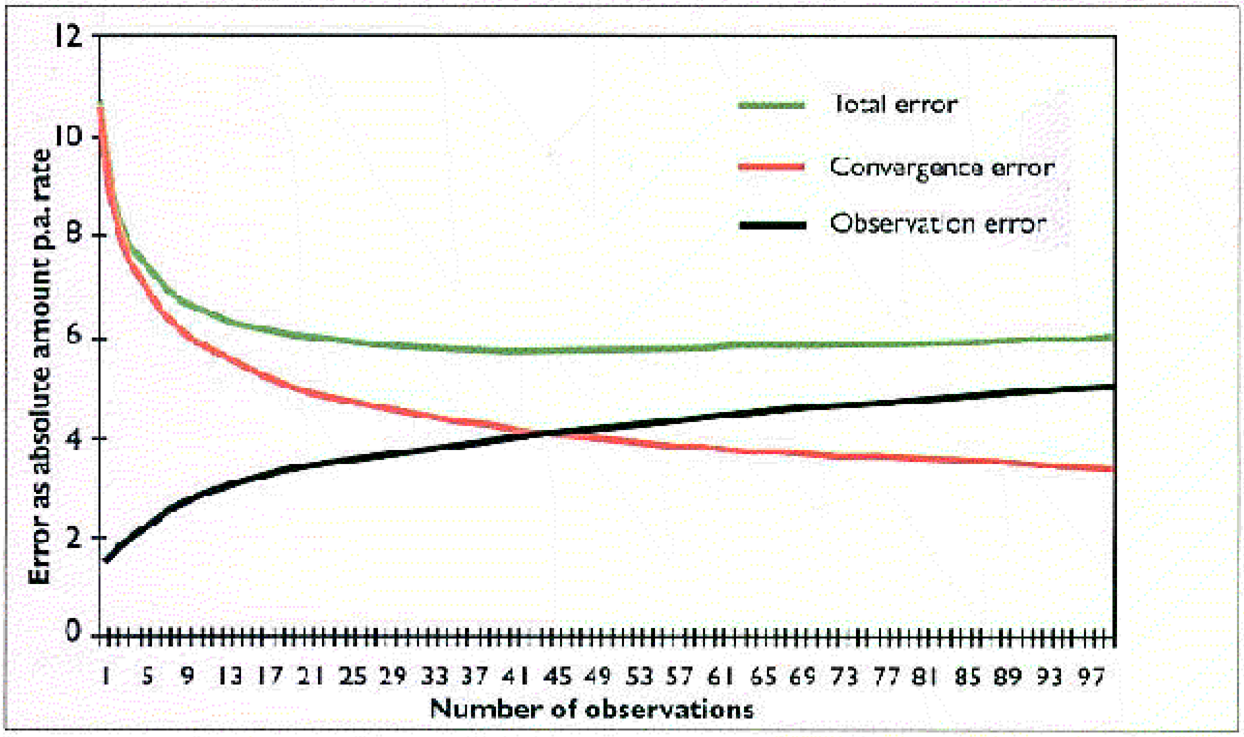
Both RiskMetrics and Garch use one value per day to estimate what volatility was. Consider 7 January 1998: at 8.00am US$/¾ was 1.3314, and at 8.00am the next day it was 1.3310. Both of the above methods would estimate volatility at virtually 0%. But as everybody trading that day knows, US$/¾ traded as low as 1.3080. Would you be comfortable relying on a risk management system that could miss a move of this magnitude?

Clearly, we need to observe the price more than once a day. (Start talking to your database manager now, and don't leave until there is a deadline and someone to blame.) Of course, with each extra snapshot, you introduce more possibilities for observation error. Let's take a simple example and examine the trade-off.

Assume X is the log of the change in the US/JPY price from one day to the next and assume that we can't exactly measure the real change or true X value. We can only read the observed price from a screen display, with each observation having some added but hidden error, because the market price is not exactly the same as the screen price.

Assume we take n observations per day, and that all the errors and price changes are independent of each other. The result is represented graphically in Figure 1 (left). As you take more observations, the sampling error due to the underlying volatility decreases, but the observation error increases as you take more observations. The best compromise in this example is at or near 48 observations, and it is stable. Any number from 20 to 60 works equally well.

Figure 1:



In this simple scenario, the size of the error involved in taking only one observation per day is twice as big as it need be. Those who have tried estimating the parameters for one value per day Garch models have come up against the problem of flat likelihood functions; not surprising really. And if they use it to forecast unseen data (the future for example), they will have even bigger problems.

Don't bother with the heavy maths if it is all based on

<u>one</u> observation per day

I agree that this is a simple scenario, but the potential for improving on the one observation per day is enormous. Going from one-point estimation to multi-point estimation will take a lot of research - the main problems being stale prices, observation error and serial correlation. But it must be done if you require a serious forecasting system. Note that I have also ignored the bias introduced by the observation errors and assumed that this can be factored out. Start today and give it the time it deserves.

Figure 1 shows a typical compromise between big convergence error and big observation error. This illustration is based on a very simple example, too simple to believe. However, it does question the conventional norm of one observation per day. Come on, guys, don't bother with the heavy maths if it is all based on one observation per day.

Exponential decay or exponential weights?

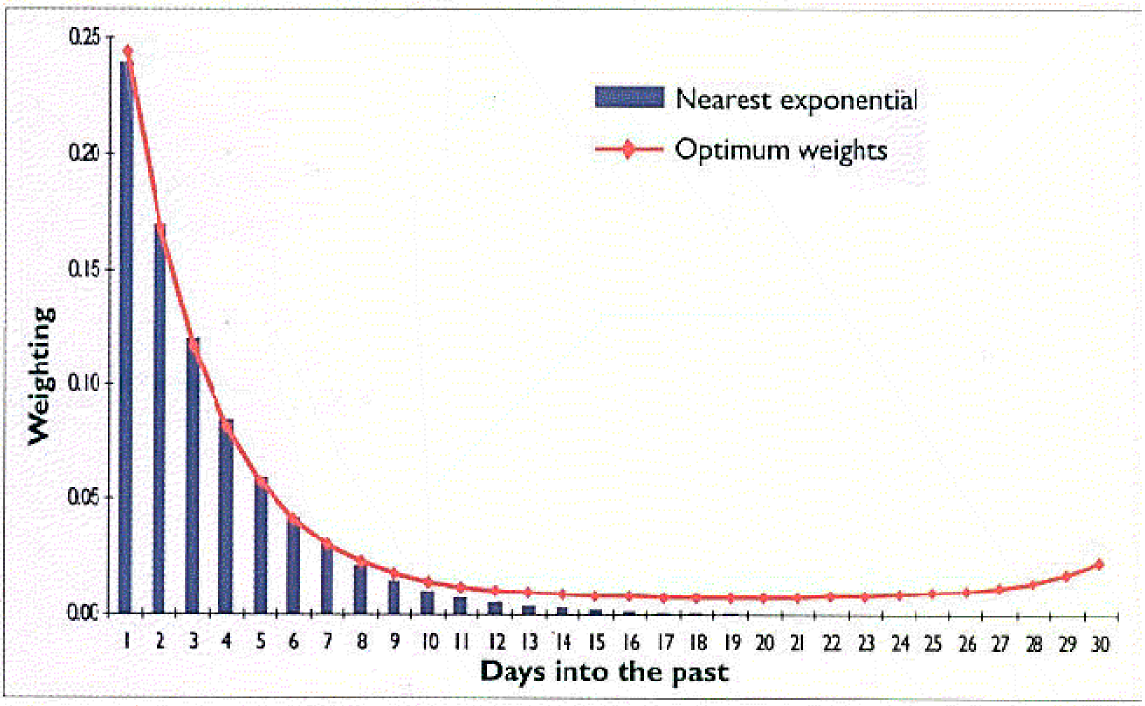
Most systems, including RiskMetrics and Garch, talk of volatility decaying exponentially. This sounds reasonable, but what does it mean exactly? Clearly it deserves closer examination.

For now, let's accept the commonly used assumption that there is a mean or long run volatility level. This allows us the concept of correlation, and we could now assume that the correlation of the underlying variance decays exponentially.

Where will that lead us? To the RiskMetrics-type exponential smoothing or to Garch? Well the bad news is that it leads to neither, but the good news is that it leads to better forecasting. Let tomorrow's forecast be a weighted average of the past observations. The desired objective is to choose the weights that minimise the difference between the forecast and the actual.

Figure 2 (below) illustrates the best weights to use for some sample parameters. (Thirty days is too short for real life, but it makes for a better graph!) Also included on the graph is the best fit for an exponentially weighted curve. It can be seen that no exponential curve has the desired properties. Any exponential curve that approximates the front end closely will decay too rapidly, and any exponential curve that fits the back end closely, will underweight the front end. Exponential decay clearly does not reconcile with exponential weights. At first this is a surprise, but with some thought it makes perfect sense.

Figure 2: Optimum weighting for exponential decay



A good rule of thumb in statistics is that the more independence you can get the better the estimator. Now consider, for example, day 15. Day 15 is highly correlated with days 12, 13 and 14, and also with days 16, 17 and 18, so if you like it has many close cousins. Day 1, by comparison, only has days 2, 3 and 4 as close cousins. Obviously, no days into the future are known. Therefore, it makes sense to decrease the weighting for days like day 15 and increase it for days like day 1. Indeed, you will see that the weightings increase slightly towards the end for the same reason. Remember that you are trying to estimate the mean as well as capturing the correlation effect.

I could have developed a model that incorporated the long run mean more explicitly, but then I would need to estimate it. Of course, in doing that I would have to use the same data that I am applying the weights to, so why not let the weights do all the work?

The 'optimum weights' chosen, clearly allow the volatility tracking system to react quickly, but without the short memory of an exponential system. Decay slows down as time increases.

The EMS crisis of late '92/early '93 demanded a system that could react quickly. But it is now 1998, and have you forgotten what can happen? It is said that volatility jumps up and trickles down; let me add another little rule: 'React quickly but don't forget'.

I believe exponential weightings are unsuitable for volatility forecasting. They fail to react quickly while still maintaining the long memory. Of course, I have again made some simplifying assumptions to illustrate the point, the main one being the assumption of a stable mean volatility level.

Is there a stable mean volatility level? I think not! Now, excuse me while I make some coffee for the database manager.