SOPA PIRANHA

Sopa Piranha*

"Catch a Wavelet"

October 16, 2006 Issue #13

WOLF INTERNATIONAL

Quarterly Commentary:
This quarter's commentary is focused on risk measurement. Ten years ago I published a piece on risk management noting that "**good risk management is not just risk measurement"**. Now I am turning the subject the other way around noting that "good risk management cannot be done without proper risk measurement". Specifically this piece targets risk measurement for the hedge fund industry. As the industry has matured and data has become more available, quantitative studies have become more viable. The environment for writing this seems right as well since **record numbers of hedge funds are going out of business.** We seem to be setting a new high water mark for frauds in the industry. And while we are on the subject of water, lately some fund of funds investors are asking how they can prevent a tsunami hitting while they were expecting a series of 3 foot waves from their multi-strategy fund.

I approach the topic by first covering the basic tools of risk measurement including comments about why hedge funds present unique problems. Next, I **outline the more advanced factor tools** being used by hedge funds and fund of funds. Last, I move to what I believe is the **future of risk measurement** describing the wavelet models being developed at Wolf to solve the unique challenges in hedge funds and fund of funds. **For those of us who avoid formulas**, I have written this in a way that the methods can be clearly understood while skipping right over the math.

I have set out to show that proper risk measurement is a risk management opportunity. If risk measurement, monitoring, and management are done extremely well, **either more risk can be taken or more return will be generated for the same unit of risk**. Either way great risk measurement is the first step to better results and more revenue.

<u>Section One: Why don't basic risk measurement tools work very well for hedge funds?</u> Many clients ask us to measure certain risks in our portfolios that allow them to consolidate risk with traditional asset classes. These factors often **fail to capture the essence of risk** in our portfolios of hedge funds. Items such as the **sector exposure, the size of market cap, the geographic location, and gross exposure can be misleading**. Hedge fund investing adds risk dimensions to a portfolio such as correlation to **poor liquidity in markets, use of leverage,**

high turnover, heavy use of derivatives, correlation of unrelated assets, and poor transparency to investors. Below I describe the challenges of accurate measurement of risks in hedge funds.

I. **A Summary of the basics:** Risk measurement is quite simple for long only liquid portfolios. Off the shelf software can be used to measure correlation to risk factors such as equity beta, fixed income duration, investment concentration, and volatility exposure. Actual positions can be put into the software. The result is a quite accurate report that shows the **value at risk** (VaR), or the amount of money that can be lost each day in 95% of all market environments. For simple portfolios, **imposing tight daily loss limits using VaR covers most of what is needed**. Therefore, the measurement problem of a basic portfolio comes down to the 5% of market environments that are not normal.

   Stress testing is used to cover the 5% stress periods when a portfolio's behavior is unpredictable. **This is usually the result of poor overall market liquidity.** For most simple portfolios, one can do an analysis of the return results in the last ten highest periods of stress. This will produce enough information to predict the worst case scenario. Ninety percent of all hedge funds rely on these simple measures of VaR and stress testing. They are enough for long only portfolios and **are not enough for hedge funds because they fail to measure first time events**. Therefore only about 1000 hedge funds have risk measurement tools that extend far enough.

II. **Leverage is the main culprit**. Leverage reduces the margin of error by magnifying mistakes. This makes first time extreme events a potential disaster and requires better technology for measuring the possible impact. It requires that **measurement error be reduced by the amount of the leverage**. So a portfolio at five times leverage needs to reduce its measurement error by five times. One of the ways to do this is to respond more quickly by measuring risk in real time. Another way is to improve the tools. Measurement tools that can improve accuracy are addressed in section two.

III. **Correlation exposure is the second culprit**. Hedge funds rely on the spread relationships that are supposed to operate within a predictable range. Correlation on all types of spreads is highly unstable over time. As we have recently seen in natural gas, even a calendar spread can blow out beyond anyone's expectation. **For purposes of our work, correlation refers to the relationship between two risk factors in our portfolio**. We want to know how reliable the correlation is. Most hedge fund trades have two sides and so this two side correlation analysis is critical to how much an individual trade can be leveraged. In section three, we deal with this through wavelet analysis of correlation.

IV. **Coherence can be deadly**. Hedge funds also rely on diversification benefits from multiple uncorrelated asset factor risks. For purposes of our work this refers to the matrix of correlations in a very large portfolio or group of funds. This is something termed **coherence.** It was covered in a prior Sopa publication #7 available on our website. Coherence specifically is a problem when all asset risks have unusually high correlation. It is well known that this **coherence tends to coincide with periods of stress due to poor liquidity**. This marks the extreme left tail event that keeps us all from taking more risk. Again wavelet analysis of the factor matrix can help measure this risk.

V. **Non-stationarity is a fourth culprit**. A good hedge fund by definition is moving its risk to the opportunities and as the strategies change it therefore has variable risk. Said another way, if a track record of a fund is evaluated over a 5 year period, one is evaluating a **moving target of risks**. This can be an example of good portfolio management or bad style drift. Related to this is the problem of comparing funds over **different time frames** that imply different opportunity sets. Section three specifically uses wavelet theory to address these problems.

VI. **Last, the problem of poor transparency** makes doing all of the above impossible for many investors. The combination of sections two and three provides a solution to the dilemma. **The tools used in combination with manager meetings, reports and good market knowledge are an improvement on basic tools of risk measurement.**

I have left out many important items to reduce the length of this already very long Sopa. I have intentionally not discussed problems and solutions involved in dealing with non normal distributions, data integrity, volume of trade analysis, persistence analysis, and default analysis to name a few. I have focused on those items unique to hedge funds that I feel are **not addressed in standard models or literature**. This is in keeping with my task of removing left tail risk so more returns can be generated.

<u>Section Two: How can factor analysis help reduce risk measurement error?</u>

I said in section one that leverage in hedge funds created the need to **reduce measurement error to near zero**. This means that identifying changing risks needs to be done frequently and with great accuracy. Factor analysis can fill the remaining gap by identifying more risk factors that may be unknown. It can also point out the rate of change of risk factors.

I. **What is factor analysis?** Factor methods of risk measurement determine the aggregate factors that explain a series of returns. A series of returns can be highly correlated to a stock index like the S and P 500 but this correlation does not show that the same stock is also correlated to many other factors. Perhaps some of those factors show even higher correlation. **Factor analysis is designed to locate those factors that best explain the risks**. If those risks change, the analysis shows new risks that should be monitored. For example, if a stock of a Company suddenly takes over another Company in an unrelated business, the factor study will show how the new stock should perform as the aggregate of the risks each had.

Factor analysis is done either as forward modeling or as inverse modeling. **Forward modeling** uses pre selected factors that are assumed to contribute to risks. For example, if I am selecting a factor in order to analyze a fixed income manager, I will first select a fixed income security such as the US ten year treasuries. The most popular **inverse approach**, Principal Component Analysis, derives statistical factors from a universe of probable contributing time series. Both are described below in greater detail for those readers not yet familiar.

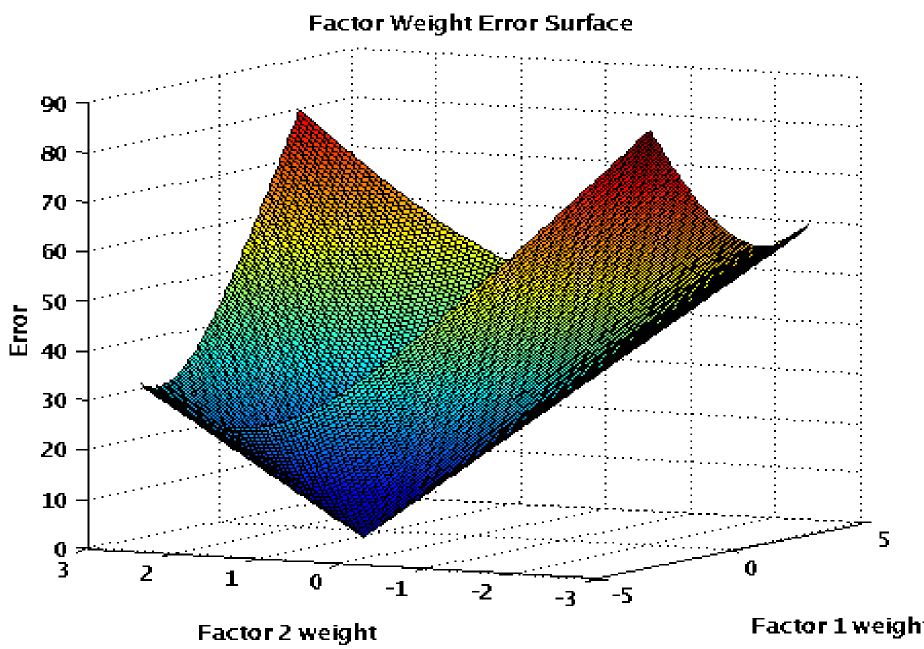
II. **Static forward modeling.** SFM is a “find the factors” game. Given a set of returns, the model searches for the combination of factors that nearly matches the return stream. **By definition it is a replication strategy if those factors are futures contracts or assets that can be traded**. The model simply correlates the return stream to a set of factors and then selects the factor that explains more of the data than the others. Then it deducts this data and again correlates the remaining return stream to the remaining possible factors to find the factor that next best explains the data. And so on until it runs out of chosen factors.

In practice, if a portfolio is being examined for risk, and a new factor emerges, **one can know what risk is being taken without any position information**. For example, if I am a risk manager at a large multi-strategy fund and see a new factor emerge in the form of wheat futures but I know that no one here has any limits to trade wheat, I am immediately warned to go look for the person who might be trading something without permission.

A factor may emerge showing a large short risk position in the VIX volatility index. Perhaps the trades of four people when combined are creating that risk. Each of them individually may be small and within tolerances. **In aggregate they may be too large**. This is different from simple correlation analysis or VaR models in that it is **more likely to locate unintended factors**. It has obvious benefits to determine if the history of returns is coming from the factors the manager says he is good at versus a set of factors that are an accident. Large quantities of funds can be quickly analyzed by comparing what the manager says to what the factors say. If they are not the same, walk away.

In our industry the factor approach described above is often used to separate market beta from alpha. If the explanatory factor is determined to be the S and P 500 for example, it is assumed to be showing beta in the portfolio. And, if it cannot be explained, it is often assumed to be alpha, or risk free return. We find that more often the remaining unexplained data is due to a lack of factors being searched.

To show how this works, I use a two security risk below, such as a large cap stock long and an S and P index short. We can pre select the factors since they are easy. Our model then performs an optimization by calculating the minimum error for different combinations of weights of the explanatory factors. For the 2-position universe, we can visualize the error calculation as a surface, where the minimum error is a point at which the **elevation is zero as in the blue area below**. It best explains the combination of factors with their respective weights. Later in this paper I will ask that you understand error maps so please check the chart below carefully.



A 2-position factor analysis is performed by determining the minimum elevation on an error surface. The elevation is zero at (-1,1) indicating short 1x factor 1 and long 1x factor 2.

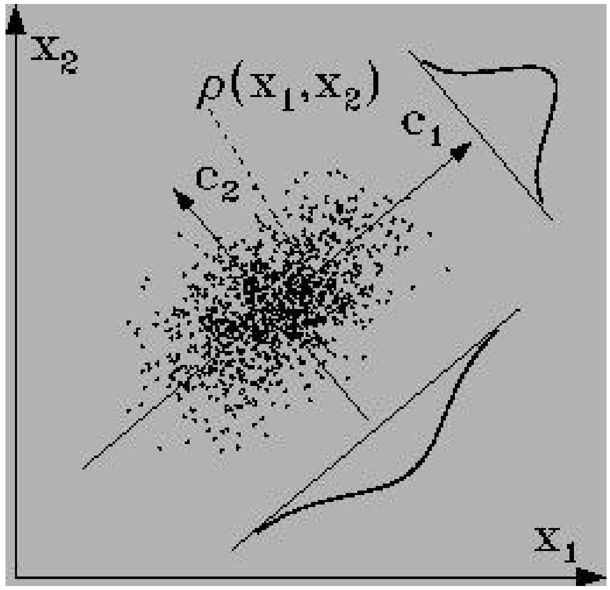
Clearly in the absence of transparency we would only know a manager's returns, not the factors likely to explain the portfolio. Herein lies the challenge of factor modeling; we must extract not just weightings of factors known a priori, **but we must determine the**

factors themselves. The game of locating the factors is the most common and most valuable use of the tool to help remove risk measurement error.

III. **Inverse modeling via principal component analysis:** PCA is a predictive tool that is a “find the statistics” game. Its incremental value over forward modeling is simply that the process is designed to **find all the return stream explanatory data** (not just some factors) in a computationally easy way. It is more complete.

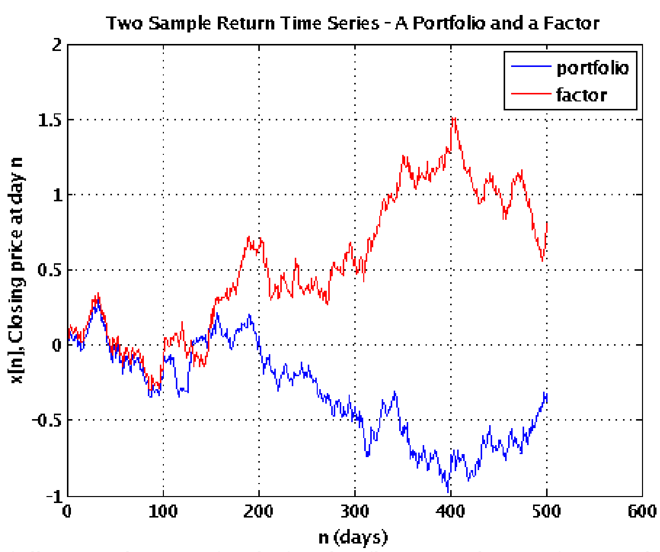
PCA takes all the time data points, for example all the months of a manager’s returns, and averages them with all the other months essentially compressing the data. The data is then put into a covariance matrix pairing the months. **This is done so that explanatory information (called eigenvectors) can be extracted using statistical methods**. The eigenvector with the maximum variance from the main data set has the most explanatory power. This is not a factor like the price of oil; it is a statistical factor with no explicit meaning. Once you have that statistical factor, you can then locate a market factor or set of factors that match that explanatory statistical data and use them. But it is important to note that the original output is a statistical data set and not a factor to be traded. **The statistical factors must be further interpreted. The value is in that further interpretation.**

One expects these highly explanatory statistical factors to closely match powerful market predictors such as beta or interest rates. The PCA process output is shown visually below where the scatter plot has points far from the center with higher variance. These **coincide with the tails of the distribution curves** shown at top right and bottom right. One need only understand that this is a process of locating highly explanatory data that tend to coincide with those times when we are either **making a lot of money or losing a lot**. Since these left and right tails can make or break our performance, studying them is more useful than studying the muddle in the middle.



Principal Component Analysis creates factors that are vectors of maximum variance. These so-called eigenvectors are statistical factors that must be correlated with market factors.

IV. **Dynamic factor analysis.** One of the major challenges in hedge fund factor modeling stems from the fact that – unlike more stationary mutual fund counterparts – hedge fund portfolios are *dynamic and are changing frequently*. A multi-strategy fund, in particular, which draws from a diverse investment pool, may reallocate sector or style weights frequently. Dynamic factor modeling, which permits either the factor weights or a combination of the factors and weights to change with time, helps measure **the non-stationarity of hedge fund portfolios**. The need for recognizing correlation across different time horizons is illustrated in the figure below.



The factor and portfolio correlate perfectly for the first 100 days; afterwards they diverge. This figure illustrates the need for multi-scale correlation in factor analysis.

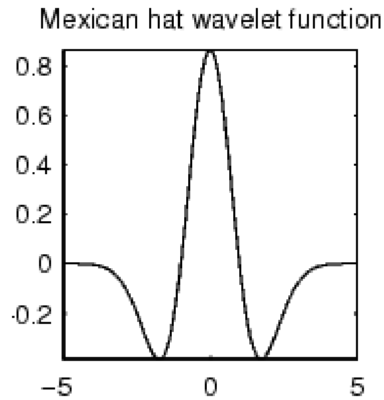
The portfolio returns and the factor correlate perfectly (ρ=1) during the first 100 days, but not during days 101-500. Using a single correlation value across the entire 500-day horizon gives a correlation value of ρ= -0.76, which **completely obscures the highly-probable attribution of returns to an unlevered long position** in the factor during days 1-100. To make matters more confusing, it's equally as possible to select too short a time horizon for correlation measurement, thus diluting recognition of longer time horizon attributions. The unfortunate result of opposed objectives between recognizing long and short time horizons suggests we need a **method of multi-scale correlation**.

Section two has focused on factor techniques that all hedge funds should employ to better understand the risks in their own portfolios. Fund of funds who complain about lack of knowledge or transparency can go a long way toward understanding their manager's risks by employing these techniques. In section three, I move onto new horizons that deal with the **big problem of non-stationarity** in hedge fund strategies.

Section Three: How can wavelets be used to reduce risk measurement problems unique to hedge funds? In section one I expose the problems of **correlation, non-stationarity and poor transparency** in risk measurement. We have developed a factor model approach that **simultaneously addresses these problems through use of wavelets**. **These systems monitor manager factor exposure and market conditions**. Monitoring of manager factor exposure is essentially a measurement of time-varying, or **non-stationary, correlation** between returns and market factors. We can measure how a factor risk correlates in many time frames and thus expose more information about the sub periods easily. Monitoring of market conditions is done through a measurement of **coherence** among the market factors themselves.

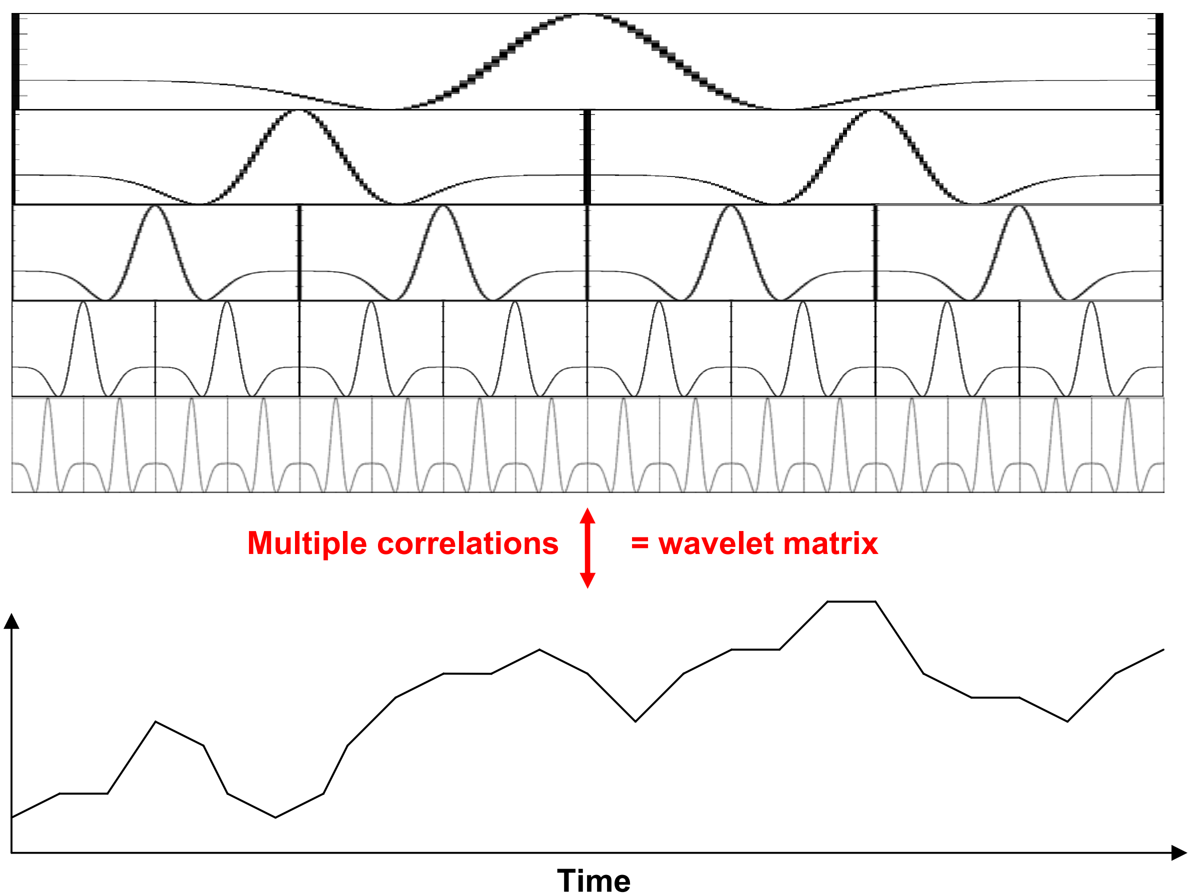
Our risk monitoring system relies heavily on **wavelet technology to augment existing industry-standard risk measurement tools**. Wavelets are a relatively new (about 20 years) and thriving realm of signal processing research. Wavelets are widely used in audio and image analysis and compression as well as pattern recognition. We find that they are a viable tool for hedge funds.

I. **What is a wavelet?** There are many different wavelet functions, but for the sake of easy discussion I'm choosing as an example my favorite wavelet, the (New) Mexican hat, which is shown below. You can think of it as a small scale visual example of a price movement that has symmetry. **The end game is to find patterns in your price data** that match your wavelet so that these sub periods can be analyzed or added up in a more meaningful way.



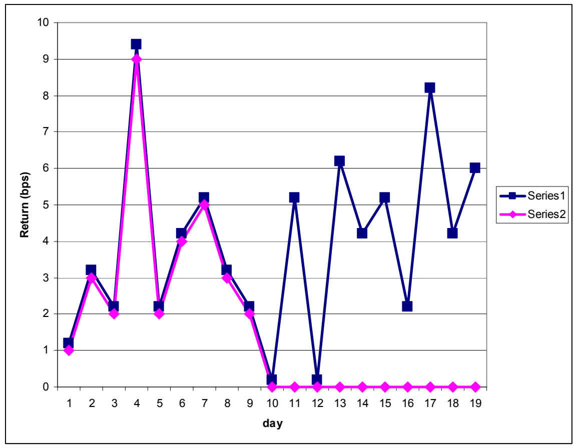
A Mexican hat wavelet is one of many wavelet functions. Think of a wavelet as a price movement that has symmetry.

A wavelet is "wave-like" in appearance and has finite width. Note below the different size scales of the same hat wavelet below. Price comparisons with these multiple scaled wavelets produce different correlations. These are combined to form a wavelet coefficient matrix. **Wavelet analysis is a correlation of a time series to these shifted, scaled wavelet functions. The math is simply a way to do this fast.**



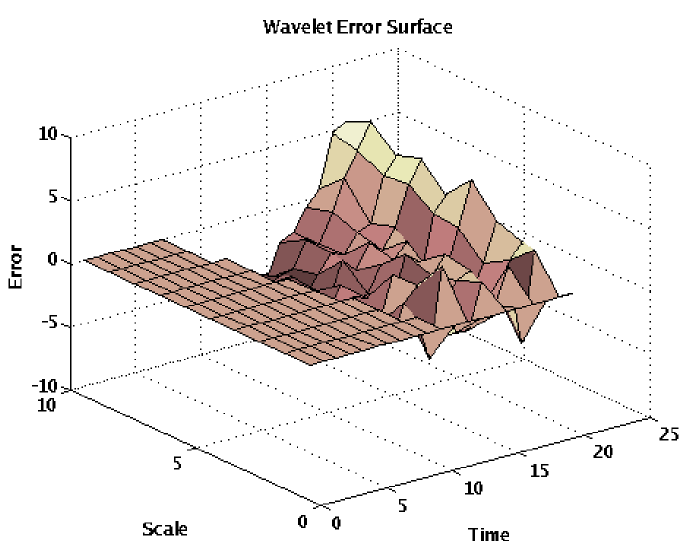
A wavelet matrix is created by correlating stretched and shifted wavelets with the time series.

II. **How can I use wavelets for factor analysis?** We use wavelets in a forward factor modeling algorithm. The factor universe consists of indices, commodity prices, interest rates, etc. We determine a wavelet coefficient matrix for each factor and for the price series or returns being analyzed, and then we compare the returns with each coefficient matrix through an **error map**. Consider the two time series of price data below.



The two time series of price data are identical for the first 10 time points, after which the second time series is at price zero. Wavelets are ideally suited for detecting abrupt changes in correlation – or nonstationarities – such as that occurring at time 10.

fter which the
let coefficient
ɔws a perfect
n the second



Subtracting one wavelet coefficient matrix from another for the above time series gives a wavelet error map. The zero-error at all scales at the beginning indicates perfect correlation.

Where the error map is zero, the factor being analyzed perfectly describes the returns. Its correlation is one. In reality the error map is never perfectly zero, nor does a single factor describe returns. Accurate factor analysis with wavelets requires an optimization procedure to determine the multiple factors and their corresponding weights that describe returns. I discuss the optimization procedure in section V. I know this is about the point where you are counting how many more pages of this are coming. Hang in there, the hard part is in an appendix you can skip.

A problem that confronts us in using error maps to perform attribution (or factor analysis) lies in the fact that **there is no unique solution of factor weights producing a zero-error map; there are infinite solutions.** Consider on a given day that a portfolio is up 100 bps. I can arbitrarily choose *any* symbol in the universe as the factor de jour. If on the same day that the portfolio is up 100 bps, Texas Instruments stock is up 20 bps, I can attribute my portfolio gain to a 5x long position in Texas Instruments. On the following day if portfolio returns are down 50 bps and Microsoft is up 25 bps, I can attribute my loss to a 2x short position in Microsoft. I could select any factor – or combination of factors – each day to get a wildly variable daily attribution of returns. **Given that the notion of persistence is itself embedded in the error map, however, we tend to accept solutions in which the error map shows low error at higher scales to reflect more persistent factor attribution.** I detail the mechanism for selecting more persistent descriptions in Section VI.

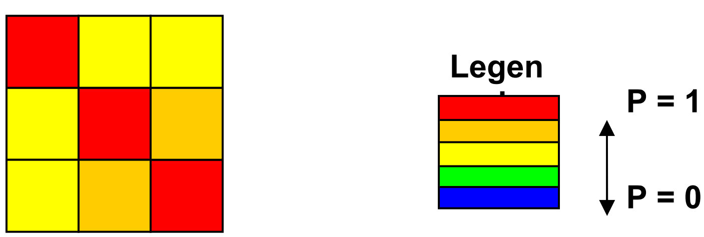
III. **How do I overcome non-stationarity through wavelets?** Recall that non-stationarity shows up in hedge funds because their strategies change as market opportunities change. Our use of wavelets takes a unique approach to the factor analysis problem by exploring attribution of returns to market factors at **all possible time horizons**. The above error map example illustrates the utility of wavelets in describing correlation between two time series, but more importantly it reveals the ability of wavelets to detect **changes in correlation across time, or non-stationarities.** As the error map becomes non-zero at approximately time 12, the correlation has collapsed. **This collapse in attribution indicates a style shift and warns me to call the manager to determine why he has changed strategy.** This does not need to be a full scale change in strategy to be noticed. In fact it will take some time before the divergence is apparent. The subtle changes to market factors can be shown in real time with live data feeds. The key is to pay attention.

IV. **How do I measure factor coherence with wavelets?** Exploiting the same benefits of wavelets that helped us to overcome the problem of non-stationarity in attribution calculations, we can perform multi-scale correlation **among factors** to recognize **a collapse in factor dispersion (or high coherence), which represents an extremely risky environment** as I discussed in a prior Sopa. Recall that a red hot heat map shows that everything correlates and coherence is high. An all blue heat map shows low coherence. Consider a 3-factor portfolio with the following factor correlation heat map as measured over the past 20 trading days:

The above 3-factor heat map shows correlation among all possible factor pairs. The top-left to bottom-right diagonal elements in a heat map are always hot. Other elements in this heat map are relatively cool.

We create heat maps such as the one above from an error map for each pair of factors. Conventional correlation calculations work directly between factors for a single time window, but our use of wavelet maps permits correlation calculations for multiple time horizons simultaneously. Directly from these error maps, with no additional calculations, we can create wavelet error maps between each pair of factors in the universe **for any time scale**. Moreover, **we can track changes in factor universe coherence by monitoring the coherence temperature over time.**

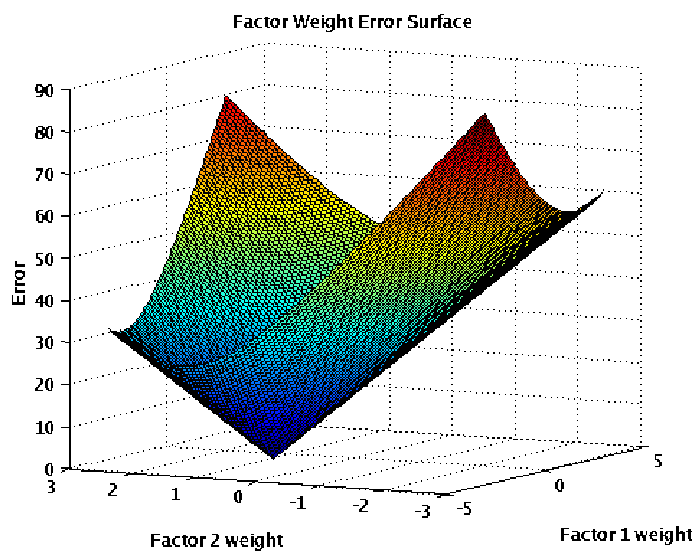
For example, suppose from a 20-day scale wavelet error map we created the above heat map and the next 20 days show a heat map as follows. We can see visually a **reduction in dispersion that serves as a warning sign of potentially impending collapse in dispersion.** These time frames coincide with periods of stress and negative returns. Warnings in advance can be used by hedge funds to lower risk or by fund of funds to reduce risk concentrations.



The heat map now shows warmer conditions among all factor pairs (when compared with the previous heat map). The increase in total heat represents an increase in coherence, or reduction in dispersion.

V. How do I perform forwarding modeling calculations? Our factor analysis consists of two steps. In the first step, we select the factor universe and **calculate wavelet matrices for each factor and the portfolio returns**. In the second step, we perform an optimization or fitting procedure **to determine the factor weights**, assessing the fitness of a given set of factor weights using error maps. Specifically, an optimization considers a set of factor weights to build a composite wavelet matrix as the weighted sum of factors, and this wavelet matrix is compared with the actual returns wavelet matrix to create an error map. In the absence of simplifying assumptions or a methodical approach, the optimization procedure can be complicated. In this section I discuss several possible approaches for calculating factor weights**.**

As a starting point, consider the previously-discussed contrived example in which the optimization takes the form of finding the minimum error on the below 2-factor error surface.



The factor error surface above worked nicely for a 2-factor example, but in reality error surfaces are never this smooth. Optimization therefore requires more sophisticated techniques than finding the minimum of a smooth error map.

In this example the error surface is quite simple in that it is contains no local minima. **In reality error surfaces are never this smooth**. The multi-dimensional optimization surfaces such as those encountered in our wavelet-based factor analysis are highly nonlinear, which means simple algorithms fail to converge to the appropriate solution. A brute-force search through multidimensional space quickly becomes intractable even with a modest number of factors. Consider, for example, parsing each factor into a range of allowable leverages between -5 and 5, with 0.25 increments **Assuming a computer could**

perform 1 billion calculations per second, exhaustion of a 15-factor space would require about 50 million *years* of computation.

Given the exponential computational time required relative to the number of factors, selection of as few factors as possible is requisite for realization of solutions. Additionally constraining the range of weights to tighter leverage bounds, when such leverage limits are known, makes the computation more feasible. Even with substantial effort in constraining the factor dimensionality and weight range, an exhaustive search of the space is, more often than not, impossible. **The model implementor should therefore seek to relax the allowable error to acceptable levels that will permit the model to terminate at an 'acceptable minimum' error rather than searching impossibly for the global minimum.**

In the appendix I put all of the above sections into a tidy package, detailing the integration of wavelets and optimization to show exactly how our factor model works. The discussion is necessarily mathematical, which is why I've left it as an appendix for those who want to dig deeper or dump this document onto their quants.

For those of you who are wondering if wavelets are really additive and worth doing, I can say that they will reveal more sub period data and force a hard look at the source of returns in sub periods. **They will also allow a mathematically simple way of doing this over very large data** sets, such as a whole data base of hedge fund managers. Like any risk measurement tool, they are as good as their implementation and use in risk management. For ourselves, we are convinced and work is underway.

Conclusions: Good risk management cannot be done without proper risk measurement, and I have been especially focused these past 6 months on R&D that equip Wolf with the best measurement and management tools in the industry. Wolf's Rogue Wavelet risk monitoring system, used as a supplement to other traditional measurements, overcomes the challenges of measuring **non-stationarity, asset correlation, and market coherence**, some of the main culprits of risk miss-measurement. As we finalize development and testing of our system we will use it **to determine when 3-foot market waves are likely to give rise to a tsunami**. We also use it **to determine when we can take more risk to make more money.**

Good luck in your investing and we hope you enjoyed your piranha soup…
Mari Kooi
Wolf International
October 16, 2006

Appendix: Putting It All Together: Wolf's Rogue Wavelet Risk Monitoring System

Our factor analysis system performs forward modeling for attribution of tangible factors to explain manager returns. The system integrates wavelet decomposition, pattern-matching and optimization for solution identification as follows.

1. Select *n* time series factors, denoted $f_i[m]$, or simply f_i, (where m is a discrete-time index) for inclusion in the analysis. $F = \{f_1, f_2, \ldots f_n\}$
2. Denote the manager return time series as R[m].
3. Choose ŝ, the typical symbol turnover time horizon. Note that selection of ŝ far too small will lead to a failure to capture the true persistency of an investment style, while choice of ŝ far too large will lead to longer computation times than necessary.
4. Perform wavelet decomposition at scale 150% of ŝ on time series F and R. Denote the wavelet decomposition of F as $W_F = \{W_{f1,} W_{f2,} \ldots W_{fn}\}$ and the decomposition of R as W_R.
5. Assume

$$\check{R}[m] \approx \sum_{<i>} w_i[m] \cdot F[m]$$

   Note that the validity of the assumption rests on good selection of factor set $F = \{f_1, f_2, \ldots f_n\}$. This is true for all forward modeling techniques.

6. Search for a solution of factor weights w_i by optimizing on m=1, the first wavelet matrix cross-section (or column), which shows the wavelet coefficients at all scales 1 through 1.5·ŝ at time/index m=1. Denote the m=1 cross-section of wavelet decomposition of R as $W_{R1,}$ Ř as $W_{\check{R}1}$, and that for the set F as W_{F1}. The error map cross-section at m=1 for a given solution vector $\mathbf{w_i[1]}$ takes the form $W_{E1,i} = W_{\check{R}1,i} - W_{R1,I}$ (scale i). Determine how close solution vector $\mathbf{w_i}$ is (in the wavelet domain) to tracking the real returns by calculating the normalized total error of the solution as

$$NTE = \sum_{<i>} |W_{E,i}| \div \sum_{<i>} |W_{R,i}|$$

7. Determine a 'persistence coefficient' as

$$PC = \sum_{<i=1>}^{\hat{s}} |W_{E,i}| \div \sum_{<i>} |W_{E,i}|$$

The persistence coefficient, which ranges between 0 and 1, serves the purpose of quantifying how much the solution vector $\mathbf{w_i[1]}$ distributes wavelet error energy into short-term, non-persistent attribution. A higher PC means the solution is more desirable because it is explained by more persistent factors.

8. Terminate the optimization (see #6 above) when NTE ≤ ñ and PC ≥ ρ, where ñ is a user-selected value determining how tightly tracking should be matched and ρ is a user-selected value determining how much of the returns must be explained by factors with trade time horizons of ŝ and above. Choice of ñ too small and/or ρ too large may result in the optimizer failing to find a solution.
9. The optimization in #6 becomes computationally infeasible for even a relatively modest number of factors. **Therefore this optimization uses simulated annealing or least-squares fitting.**
10. After finding a solution vector $\mathbf{w_i[1]}$ to the m=1 cross section, determine solutions sequentially for cross-sections m=2, 3,…ᴍ (where ᴍ is the last sequence value). The optimization for m=2,3,… ᴍ should *not* be done as stand-alone optimizations. Rather, using the previous cross-section solution, the current solution should be calculated by descending the k-dimensional gradient to find the local minimum error.

An incremental descent through $\mathbf{w_i}$ space in the direction of $\boldsymbol{\varphi} = \partial W_E/\partial w_1 + \partial W_E/\partial w_2 + \ldots + \partial W_E/\partial w_i$ leads to the local minimum, or equivalently the optimum solution. Upon reaching the local minimum the gradient-descent solution is accepted if NTE ≤ ñ and PC ≥ ρ, as in #9 above; otherwise a non-stationarity requires that the optimization be done as for m=1. **If NTE and PC differ substantially for consecutive time increments, a style shift has likely occurred.**

A Note On Optimization:

The fact that brute-force searches of even a constrained space are too computationally time-intensive motivates a more methodical search of the solution space. **Adaptive simulated annealing (ASA) is an optimization tool well-suited to locating global minima on highly nonlinear surfaces of any dimension.** It combines a random sampling of solution space with a time decay mechanism that performs a gradient-descent around the randomly-chosen space. In the words of Lester Ingber (ASA expert):

> *Consider a mountain range, with two "parameters," e.g., along the North-South and East-West directions. We wish to find the lowest valley in this*

terrain. ASA approaches this problem similar to using a bouncing ball that can bounce over mountains from valley to valley. We start at a high "temperature," where the temperature is an ASA parameter that mimics the effect of a fast-moving particle in a hot object like a hot molten metal, thereby permitting the ball to make very high bounces and being able to bounce over any mountain to access any valley, given enough bounces. As the temperature is made relatively colder, the ball cannot bounce so high, and it also can settle to become trapped in relatively smaller ranges of valleys.

Another optimization approach, which is likely much faster still than simulated annealing, involves solution of simultaneous equations of an over-sampled time series to create a least-squares fit of wavelet coefficients. The need for oversampling to create more equations is motivated by the fact that the system is otherwise under-determined. In the interest of keeping an already-confusing section as simple as possible, I'll omit details of exactly how over-sampling and least-squares fitting works. **Suffice it to say for now that we're still comparing how well it works against simulated annealing and other approaches.**