Risk Management of Hedge Funds using Fuzzy Neural- and Genetic Algorithms

Clemens H. Glaffig
Panathea Capital Partners

clemens.glaffig@panathea.de

August 20, 2004

Abstract

The article investigates the use of adaptive learning algorithms in constructing dynamic portfolios replicating the return characteristics of a given hedge fund. The emphasis is on out of sample conditional predictive capabilities as necessary to serve as a valuable risk management tool, rather than simply explaining hedge fund behaviour over an in sample period. The algorithms learn dynamic trading rules and strategies along with which factors to base those on within an integrated learning mechanism. It thus generelaizes previous approaches by exploring a wide class of nonlinear and dynamic trading strategies to participate in explaining and predicting hedge fund behaviour. The conditional predictive capabilites of the algorithms can specifically be employed to quantify future fund behaviour. It will be useful in constructing quantitative risk measures for individual hedge funds. The article also provides some empirical data for the out of sample behaviour of this method.

1. Introduction

Supported by the recent meltdown of the major equity markets, alternative assets and trading styles have moved further into focus of institutional investors. Specifically hedge funds now form an integral part of the asset allocation of most major institutional investors. However, the hedge fund crisis of 1998 has left its marks in the form of increased demands to understand inherent risks and install mechanisms to quantify and manage those. Given the sparse data, their dynamically changing exposure to markets and lack of full transparancy of a lot of hedge funds, classical risk measures, like empirical VAR, will be difficult to derive and often not properly account for potential hedge fund risk. In some cases the use of classical risk measures even turns out counterbeneficial. More advanced methods to analyse and quantify hedge fund risk, both on a stand alone and marginal basis, are necessary.

1.1 Previous work

Considerable research has been conducted on the analysis of Asset Based Style Factors ("ABSF"), the main driving factors for managed portfolio returns. Sharpe (1992) used the analysis of ABSF to explain the behaviour of mutual funds. The extension for hedge funds included behavioural factors (like the ability to go long and short) and factors, that are not directly tradable (like volatility), see e.g. Schneeweis & Spurgin (1998, 1999) and Fung & Hsieh (1997). These appoaches provided a further understanding, which factors besides the underlying markets drive and influence a hedge fund's behaviour. In contrast to mutual funds, which often depend on a few factors mostly in a linear and stationary way, Hedge Fund return characteristics exhibit strong nonlinear behaviour, trading strategies employed are often substantially more dynamic. To account for the dynamically changing and nonlinear exposure to market factors, Fung & Hsieh (1999, 2002, 2003) used certain options as additional asset based factors to better track and understand trend following strategies, fixed income- and equity funds. Agarwal & Naik (2002) apply this approach to a broad range of hedge fund styles, including the returns of specific simple trading strategies as style factors. Their analysis proves that nonlinearities are part of the true and dominant economical risk- and style factors. Gupta et al. (2003) allow for quadratic terms in their factor analysis to investigate, e.g., market timing abilities of managers. Further approaches by Abernathy & Weisman (2000) also included the return of fixed trading strategies as factors.

In most of the previous approaches of modeling hedge fund return characteristics, the focus was on categorizing and grouping hedge funds, describing common underlying risk factors like option type nonlinearities or on methods to evaluate performance against more appropriate benchmarks obtained from the ABSF. To serve as a risk management tool and to help construct valuable quantitative risk measures, an analysis of hedge fund return characteristics has to move from answering the question "Are we in trouble?" to being able to provide answers to the question "By how much are we in trouble?". It should be able to provide conditional predictive capabilities. While this was not the main purpose of most previous models, providing the necessary conditional predictions on an individual fund level will not always be achieved by buy and hold strategies in underlying market factors plus a preselect set of nonlinear tradable instruments and/or trading strategies.

1.2 Goals of the study

Methods of soft computing have been amply applied in the past to forecast market behaviour, often with very mixed success. While trying to establish a functional relationship between a set of market factors and the market behaviour itself may be a

daunting task, using methods of soft computing to map market factors to conditional predictions of a fund's behaviour is footed on firmer grounds: Any trading strategy, be it based on technical, fundamental or other rules applied within a fund, is in itself a mapping from market factors to a fund's conditional behaviour. In addition, trading strategies applied by managers are substantially more stationary (in the mathematical sense) than markets and thus lend themselves to be studied by methods of soft computing. The degree of stationarity can be further increased, if trading rules are "fuzzified".
The current study proposes a Fuzzy Strategy Mapping ("FSM"), derived via integrated neuro fuzzy- and genetic algorithms, mapping historic hedge fund and market data into a dynamic replicating portfolio. The FSM aims at providing a conditional predictive capability to serve as a basis for quantitative risk measures for single hedge funds or a portfolio including hedge funds. It will have to go further than previous approaches by not only analysing main - including nonlinear - factors, but also to include dynamic strategies explaining hedge fund return characteristics. While the tendency of previous research has been to search for a linear functional relationship between hedge fund returns and market factors, using options as factors to account for nonlinearities and a small set of preselected trading strategies to account for dynamic trading behaviour, the FSM let's the data determine which trading rules, strategies and market factors best explain the fund's return characteristics. In spirit, the method is similiar to option pricing using soft computing, where the model learns the price of an option as well as the exposures (the deltas) for an option replicating portfolio from historic market data (see Hutchinson et al.).

1.3 General Features of the FSM

The FSM will distinguish between two types of market factors: Position factors, the assets of the replicating portfolio and input factors which will determine the dynamic exposure to the position factors. The input factors will encode the trading rules of the replicating portfolio, but no positions will be taken in them. This distinction corresponds to the way most money managers operate: The decision to enter a market or trade an asset is based on information of a number of factors beyond the pure value of that market or price of that asset. Technical trading rules use functions of the time series of assets, balance sheet ratios drive the trading decisions of fundamentally oriented equity funds and intermarket analysis mixes the information of several markets to take positions. All of these examples give reason to distinguish a trading decision from the pure position in the analysis and to allow within the model enough degrees of freedom for the trading rules. In the FSM, the exposures can be determined using fundamental factors, technical factors or simple rules, related to the history of asset prices. With the flexible character of trading rules and strategies admitted to explain hedge fund behaviour, the FSM can synthetically account for any option and any dynamic portfolio insurance strategy. As a result, the FSM is able to provide the necessary conditional predictive capabilities to serve as a basis for quantitative risk measures.

The paper is organised as follows: The first parts of Section 2 will describe the FSM and the construction of the replicating portfolio. The next part of Section 2 will describe the learning mechanism. Section 3 briefly sketches an extension of the FSM to a general nonlinear mapping, in which a nonlinear and fuzzy relationship between market factors and the fund's behaviour is established. Section 4 describes how risk measures based on the FSM may be constructed. Section 5 provides some empirical results, analysing out of sample capabilities of the FSM.

2. Description of the FSM

2.1 Features of the FSM

The FSM will construct a replicating portfolio depending on: (i) The position factors, which will constitute the assets of the portfolio, the positions, the portfolio will invest in. (ii) The input factors, which will determine what the exposure to the individual position factors will be, in which no positions will be taken.

For each position factor, a set of trading rules will be considered, each trading rule being a function of some of the input factors and resulting in an exposure to that position factor. The overall exposure to a position factor will be the sum of the exposures of all trading rules for that position factor. The exposure produced by an individual rule will be the product of two "weights": A validity weight w, measuring how valid a rule is and a leverage/direction weight c, giving a rule direction (long/short in the position factor) and potential leverage.

The input- and position factors will be selected by employing a genetic search algorithm, the trading rules will be derived using a simple neuro fuzzy approach. Both, the selection of factors and the derivation of trading rules will be integrated into one optimization mechanism. This has the advantage, that the true optimal factors for the given setting will be found – factors will not be selected by their individual correlation to the fund, but by their role and contribution in the optimisation process.

The FSM will find the optimal combination of trading rules and market factors, optimising a given objective function. This objective function can be the standard mean square deviation of hedge fund returns to FSM returns, potentially weighted to better capture downside risk. Here, the objective function also includes a term penalising high complexity.

2.2 The Replicating Portfolio

Consider a given pool of input- and position factors. For each position factor B_i, i=1 to I, there will be R rules, each depending on the set of input factors A_k, k=1 to K. The number of rules R will be kept constant for all position factors. Each rule r_j will be of multiplicity M(j), combining M(j) different conditions on the input factors via an "and" link into rule r_j and will quantify the statement:

> If a function $F_{i,j,n}$ of the input factors at time t is close to $Z_{i,j,n}$ for all n = 1 ,...,M(j) i.e., $F_{i,j,1}$ is close to $Z_{i,j,1}$ "and" $F_{i,j,2}$ is close to $Z_{i,j,2}$ "and"..... $F_{i,j,M(j)}$ is close to $Z_{i,j,M(j)}$ then the exposure at time t to the i-th position factor is c(i,j) w(i,j,t).

The Z's and c's are parameters, which will be learned. The validity weights w contain further parameters, the σ's (see below). The F's are pre processing functions of the input factor time series and will here be chosen in advance, before the learning phase of the model. For rule r_j, the validity weight w relative to the position factor B_i at time t is given by:

$$w(i,j,t) = \exp(-\tfrac{1}{2}\Sigma_n (F_{ijn}(A_1,..., A_K, t) - Z_{ijn})^2/\sigma_{ijn}^2), \qquad (1)$$

where $A_i = (A_i(s), s \leq t)$, j runs from 1 to R, i runs from 1 to I, the sum over n goes from 1 to M(j).

Assuming observation of values and prices of the input factors occurs only at integer times, the replicating portfolio P at time s ε (t,t+1] will then be:

$$P(s) = \Sigma_{i,j}\, c(i,j)\, w(i,j,t)\, B_i(s) \qquad (2)$$

The complete set of parameters to be learned consists of the Z's, σ's and the c's.

Since the functional form of the validity weights w limits them to vary between 0 and 1, they are only a measure of how valid a rule is, making no statement about direction or leverage. This is done by the c's, which will determine with what size and what direction a rule will enter.

Typically the preprocessing functions F in the validity weights w will be uni- or bivariate functions of one or two input factors only, smoothing the time series of the input factors like moving averages over small periodes, functions of moving averages or "wavelet-smoothed" market data. If-then rules are inplemented by choosing the F's as Heavyside functions. A specific example is given in section 5.

The exposures will be kept constant during a period from t to t+1, so that the value of the replicating portfolio at time t+1, just before the next adjustment of the validity weight is:

$$P(t+1) = \Sigma_{i,j}\ c(i,j)\ w(i,j,t)\ B_i(t+1) \tag{3a}$$

$$= \Sigma_i \{ \Sigma_j\ c(i,j)\ w(i,j,t)\}\ B_i(t+1)$$

$$= \Sigma_i\ w^*(i,t)\ B_i(t+1) \tag{3b}$$

The portfolio return $R_P(t+1)$ over period [t, t+1] can be expressed as

$$R_P(t+1) = \{P(t+1) - P(t+)\}/P(t+)$$

$$= \{\Sigma_i\ w^*(i,t)\ [B_i(t+1) - B_i(t)]\}/\Sigma_i\ w^*(i,t)\ B_i(t)$$

$$= \Sigma_i\ v^*(i,t)\ AR_i(t+1) \tag{3c}$$

where P(t+) reflects the value of the portfolio at time t, immediately after adjusting the validity weight, $AR_i(t+1)$ is the absolute return or the price increment of factor B_i over period [t, t+1] and

$$v^*(i,t) = w^*(i,t)/\Sigma_i\ w^*(i,t)\ B_i(t) \tag{3d}$$

Notes:

- To account for staleness in hedge fund values, time shifts of the factors can be exploited.
- By setting one of the position factors to constantly one, it is possible to obtain market dependent alphas, i.e. it is possible to partition the fund's alpha according to different phases of markets.
- The replicating portfolio is constructed using the asset prices of the position factors rather than the returns to account for changing proportions due to market changes.
- The preprocessing functions F can also be subjected to an optimisation procedure.

To determine the parameters, the following error function will be minimised:

$$E_\delta (T(t_0)) = \tfrac{1}{2} \sum_t [R_H (t) - R_P (t)]^2 + \delta(\sum_{i,j} c(i,j)^2) \qquad (4)$$

Where t runs over the training set $T(t_0)$ of historical data prior to the set time t_0, $R_H(t+1)$ is the hedge fund's return over period [t, t+1] and δ is a monotone increasing function in the sum of the squared leverage weights, to favor less complex portfolios.

Alternative error functions, like weighted squared errors, that penalise a misspecification of negative returns more heavily, might be more appropriate for risk management purposes and are easily implemented.

2.3 Parameter Determination – the Learning of the Model

The learning of the model consists of two parts: (i) For a given set of factors, the optimal portfolio minimising the error (4), i.e., the optimal paramters (the Z's, σ's, c's) will have to be determined and (ii) the optimal set of factors (input and position) will have to be chosen out of a pool of admissible factors. The learning mechanism employed will integrate (i) and (ii) into one process, in which for a fixed set of factors, the parameters will be determined by minimising (4) over a set of training data and then select that set of factors, for which the optimal portfolio is "best". The best set of factors will be chosen by employing a genetic search, attaching a "quality" to each set of factors and searching for the set of factors with the highest qualtiy. The quality for a set of factors will reflect the error of its optimal portfolio, a measure of its ability to generalise over a set of validation data different from the training data and a penalty function, that will lower the quality for large factor sets. The additional features of the quality function beyond the pure error on the training set are necessary to avoid overfitting and to further reduce complexity, as it is clear that a pure minimisation of (4) over the training set can be bettered by simply adding more "independent" factors, without improving the prediction capabilities. The quality of a given set of factors will then be given by:

$$Q = - E_\delta - V + \varphi(I+K) \qquad (5)$$

where E_δ and V are the errors over the training data according to (4) and validation data respectively, φ is a monotone decreasing function in the number I+K of input- and position factors used.

Fixed at the outset is the set of admissable factors ADM = $\{A_1, ..., A_L, B_1, .. ,B_N\}$ with input factors A_k ,k=1 to L; and position factors B_i ,i= 1 to N. Fixed is also the number of generations, G, for the genetic search. Furthermore, a time t is set, beyond which the FSM will produce conditional predictions, based on learning the data prior to t. The combined mechanism of factor selection and parameter optimisation will be as follows (a summary follows below):

Step 1:
Let T(t) be the set of training dates $(t_1 , t_2 , ... t_T)$; $t_i \leq t$ for all i, used at time t, over which the parameters will be determined. V(t) will denote a set validation dates $(s_1 , s_2 , ... s_V)$; $s_j \leq t$ for all j with T(t) and V(t) being disjoint. Typically one will choose $\max(t_i ;i = 1 \text{ to } T) \leq \min(s_j ; j = 1 \text{ to } V)$, so that the validation period will follow subsequent to the training period. V(t) will be used to see how well a portfolio would

have done in predicting the fund's returns out of sample before time t and will be used in the selection of the factors.

Step 2:
A population, Pop(g), of subsets of input- and position factors for generation g=1, with $|Pop(g)| = p$, is generated by identifying a subset of factors, $S_n(g)$, for n= 1 to p, with an L+N dimensional binary string $[S_n(g)] \varepsilon \{0,1\}^{L+N}$, such that $S_n \sim [S_n]$ according to:

$$[S_n(g)](i) = \begin{cases} 1 & \text{if } A_i \text{ is a chosen input factor of the n-th member } S_n(g) \text{ of Pop(g) for i= 1,...,L} \\ 1 & \text{if } B_{i-L} \text{ is a chosen position factor of the n-th member } S_n(g) \text{ of Pop(g) for i= L+1,...,L+N} \\ 0 & \text{otherwise} \end{cases} \qquad (6)$$

i.e. each individual $S_n(g)$ of the population Pop(g) is identified with that binary string $[S_n(g)]$, which has ones at the places representing the factors to be used and zeros at the places representing factors not to be used.
$[\bullet] : Pop(g) \rightarrow \{0,1\}^{L+N}$, denotes the natural injective embedding of Pop(g) into the set of binary strings according to the definition above.
It will be convenient to write $S_n(g)$ and $[S_n(g)]$ as the concatenations $(S_n(g,A) ,S_n(g,B))$ and $([S_n(g,A)] ,[S_n(g,B)])$ respectively with $S_n(g,A)$ representing the set of input factors used by the n-th member of Pop(g) and $[S_n(g,A)]$ the L-dimensional binary substring representing $S_n(g,A)$. Respectively for $S_n(g,B)$ and $[S_n(g,B)]$.

Step 3:
For each set of factors $S_n(g)$ of the current population Pop(g) , the optimal portfolio $P^*_n(g)$ over T(t) will be constructed by determining the parameters (the Z's, σ's, c's), that minimize the error function E_δ (T(t)) given $S_n(g)$:

$$P^*_n(g) = \{P_n(g; Z, \sigma, c) \text{ s.th. } (Z, \sigma, c) = \operatorname{argmin}\{ E_\delta (T(t), Z, \sigma, c \mid S_n(g))\} \qquad (7)$$

where $P_n(g; Z, \sigma, c)$ is the portfolio as in (2), based on the factors $S_n(g)$.

$P^*_n(g)$ will constitute the best – relative to the error function (4) - the factors $S_n(g)$ can do in replicating the fund returns over the training period T(t).
To determine the optimal portfolio $P^*_n(g)$ for each factor set $S_n(g)$, the return of $P^*_n(g)$ will be viewed as the output of a neural network, with input neuron vectors at time $t_k \varepsilon$ T(t):

$$\begin{aligned} IN1(t_k) &= \{ S_n(g,A)(s); s \varepsilon \Lambda(t_k)\} \\ IN2(t_k) &= \{ S_n(g,B)(s); s \varepsilon \{ t_k , t_k\text{-}1\}\} \end{aligned} \qquad (8)$$

where $S_n(g,A)(s)$ and $S_n(g,B)(s)$ denotes the set of input- and position factors of the n-th member of Pop(g), all evaluated at time s. $\Lambda(t_k)$ is the collection of all dates $s \leq t_k$ as required by all preprocessing functions F_{ijn} (the F's are functions of the time series of the input factors and will typically depend on the history path of a factor at time t_k). IN1 & IN2 are both vectors, whose dimension depend on the number of factors of $S_n(g)$ and $|\Lambda|$.

The hidden neurons $H(i, t_k)$ of the network can be taken to be:

$$H(i, t_k) = w^* (i, t_k) \quad (9)$$

Where the w^* are as in (3b).

The output neuron will be

$$O(t_k) = R(P^*_n(g)) \text{ as in (3c)} \quad (10)$$

The input neurons IN1 are processed in the hidden layer, i.e. they are variables of the hidden neurons. The input neurons IN2 are directly passed to the output neuron, as can be easily seen from the definition of the neurons and (3b) & (3c).
With this setting, the error function (4) will be minimised over the training set T(t), using standard network learning methods (e.g. gradient descent backpropagation, see Jang et al (1997)). That will produce for each factor set $S_n(g)$ the optimal replicating portfolios $P^*_n(g)$, n= 1,... p over the training period T(t). For each factor set $S_n(g)$ a quality Q(n) is attached as defined in (5):

$$Q(n) = - E_\delta (n) – V(n) + \varphi(I_n+K_n) \quad (11)$$

where E_δ (n) is the value of the error function (4) for the n-th optimal portfolio $P^*_n(g)$, V(n) is the validation error (squared deviations of returns) on the validation set for $P^*_n(g)$ and $\varphi(I_n+K_n)$ is the penalty term on the number of input- and position factors of $S_n(g)$.

Step 4
Out of the current popoulation Pop(g) of factor sets, choose the individual $S_{top}(g)$, with highest quality attached:

$$S_{top}(g) = S_n(g) \mid n= \text{argmax}\{Q(i); i= 1 \text{ to } p\} \quad (12)$$

This will be the set of factors within the current population, which produces an optimal portfolio, having the best mixture of fit to hedge fund returns over the training set, predictive power on the validation set and least complexity, as expressed by (4) & (5). $S_{top}(g)$ will be used to build the population of the next generation: Set $S_1(g+1) = S_{top}(g)$. Discard all other individuals $S_n(g) \neq S_{top}(g)$ of the current population of factor subsets. Generate the next generation of individuals in the population by mutating the binary string $[S_{top}(g)]$, representing the factors of the top performer, i.e. by randomly selecting sites in the binary string $[S_{top}(g)]$ and flipping those. The new population Pop(g+1) of factors subsets will then be represented by the following set of binary strings:

$$[Pop(g+1)] = \{ [S_{top}(g)],\ M_i ([S_{top}(g)]); i= 1 \text{ to } p-1\} \quad (13)$$

where the M_i represent the mutation operators applied to the top performing string.

Step 5:
Go back to Step 3 and repeat the cycle for generations g= 2 to G. At the end of it, the overall optimal portfolio (the optimal set of factors along with the optimal Z's, σ's, and c's) is obtained, i.e. one has established the optimal portfolio P^* for the training set T(t):

$$P^* = P^*_{top}(G) \quad (14)$$

Summary

The learning mechanism thus works in two intertwined phases, it selects the factors to be used by identifying a set of factors with a binary string or equivalently with a "corner" of the unit cube in L+N dimensional space. For each corner of that cube, the optimal replicating portfolio will be learned, by optimising the parameters. In this way, an optimal replicating portfolio can be associated to each corner of the cube. The genetic algorithm then searches throughout the corners of the cube for the best overall portfolio with highest quality. The corner of the cube with the portfolio producing the highest quality will then correspond to the optimal set of factors, such that its optimal portfolio as in (2) will produce the lowest error on the combined training and validation set, with limited complexity.

Remarks

- Given the sparse data for a lot of hedge funds, to further stabilize the performance of the network, ensemble learning can be employed, that is, the parameter optimisation will be run a number of times on different training sets, generating an ensemble of replicating portfolios, see e.g. Breimann (1996). The individual ensemble portfolios will be generated by bootstrapping the training data set to produce "different" training sets. The optimisation on the different training sets will produce varying sets of factors, parameters and therefore varying replicating portfolios for the various ensemble members. The final replicating portfolio will then be the simple arithmetic average over the ensemble portfolios.
- To limit the genetic search, one can heuristically fix a number of reasonable and obvious factors - like various equity/sector indices for long/short equity funds – and will run the search on a smaller universe of 5-15 additional factors. In this case, even an exhaustive search through all of the factor combinations makes sense.
- In the set of training data, extreme or tail events for most of the market factors are naturally included, given the market behaviour during the last several years. In the future, it might be necessary to include a mechanism to ensure that tail events for the important factors are included in the training set to capture stop loss behaviour or other portfolio insurance type behaviour of the fund. Also, to speed up adaption, it is sometimes advantageous to apply a "short memory" effect to the training data, weighing recent data more heavily.

3. Nonlinear "Portfolios"

The FSM as a tradable portfolio is linear in the position factors and nonlinear in the input factors. Having a replicating portfolio has the advantage of providing explanations not only as to what factors drive the fund but also what tradable strategies could be underlying the fund – the model has explanatory capabilites.

If explanations beyond which factors drive the fund's behaviour are not important, it is also possible to derive a nonlinear fuzzy functional relationship between all factors and the behaviour of the fund. For funds using mainly nonlinear instruments, like an actively managed option portfolio, the nonlinear approach may sometimes provide good results with less complexity than linear portfolios. This can be achieved by a simple alteration of the previous model:

Instead of constructing a portfolio as in (2), one can output the following nonlinear function of the factors, whereby no distinction between input and position factors will be taken. The preprocessing functions of the factor values will be normalized to lie in [0,1]:

$$O(t) = 1/(1+\exp(- \Sigma_j c(j)\, w(j,t))) \tag{15}$$

with

$$w(j,t) = \exp(-½\Sigma_n (F_{jn}(B_1,..., B_K, t) - Z_{jn})^2/\sigma_{jn}^2), \tag{16}$$

O(t) can then be trained against the sequential, normalized returns of the fund as before, viewing the B_K as input neurons and the w(j,t) as hidden neurons to a neural network with output O(t).

This model can be interpreted as a conditioned perceptron, in which for a given (fuzzy) condition of the market a certain nonlinear functional relation between factors and the funds behaviour exists.

4. Risk Measures based on the FSM

Adequate quantitative risk measures for hedge funds are difficult to derive. The standard mean-variance approach is inadaquate, as the true return distribution will deviate substantially from the normal distribution. VAR or CVAR caculations could be based on empirical distributions, but those are often meaningless given the short history of a lot of (open) funds – for a fund with a 4 year monthly history, the 98% and 99% confidence monthly VAR will produce the same number: The historic maximal monthly drawdown. It is therefore natural to employ ABSFs or the position factors of the FSM, to construct a more meaningful risk measure. However, to produce a meaningful risk measure, the factor model used must have certain minimal conditional predictive power, as provided by the FSM. For the FSM, one can be reasonably confident, that a replicating portfolio like in (2) will closely resemble the fund's behaviour based on the outcome of the position factors. Specifically, as (3c) will closely resemble the fund return over a period [t, t+1], the conditional distribution of the fund's return, R_H (t+1), can be approximated by the distribution of

$$R_P (t+1) = \Sigma_i\ v^* (i,t)\ AR_i(t+1) \tag{17}$$

and hence, can be constructed by the joint conditional distribution of the $AR_i(t+1)$, which can be obtained by

- Using theoretical marginal distributions for the $AR_i(t+1)$ and using an appropriate copula function to account for the correlation of the factors. This correlation can be time dependent (see Nelson (1998) and Cherubini et al. (2004)).
- Using the empirical joint distribution of the $AR_i(t+1)$ or using the empirical marginal distributions and coupling them again with an appropriate copula function.

These approximations of the distribution of fund's returns provides a reasonable basis for distribution based risk measures like VAR, CVAR etc.
Another approach to obtain a risk measure, is of course to use a family of scenario assumptions on the factors and use the maximal drawdown of the FSM over all these scenarios. This would produce a coherent risk measure, see Artzner et al. (1998).

5. Empirical Results

This section describes the out of sample behaviour of the FSM applied to two cases: The CSFB Tremont Composite Index and a long/short equity fund. For the index, the period over which the out of sample behaviour of the FSM is analysed ranges from Juli 1994 to June 2004 and includes the "difficult" period from summer 1998 to summer 1999. The single fund was chosen to illustrate the performance of the FSM under a major style drift in the fund: The manager of the fund changed during the considered period and with it, the investment style changed rather dramatically, as a further correlation/comovement analysis with factors and hedge fund styles would demonstrate. The analysed period ranges from April 1999 to April 2004. For the CSFB Tremont Composite Index and the single fund, out of sample predictions for the monthly returns over the respective periods were generated and analysed.

The FSM used in this analysis was kept simple. The pool of admissible factors, out of which the FSM selected, included: The S&P 500, Indices for small cap growth and value stocks, indices for large cap growth and value stocks, Nikkei 225, Euro Stoxx 50, VIX, BBB Credit Spreads, the Merrill Lynch High Yield Index, 10 yr US Government yield, Euro, JPY and the CRB Index. For all position factors only rules depending on that position factor as input factor were considered. For each factor, at most six rules were allowed, each having a multiplicity, M(j), of one. The factor preprocessing functions for all rules was:

$$F_j(B_i(t)) = B_i\,(t)/MAV_n(t,B_i) \tag{18}$$

where $MAV_n(t,B_i)$ is the n-period moving average of B_i , with n = 3 to 6.

Further improvements can be achieved by adjusting the pool of factors and/or the preprocessing functions to allow for greater flexibility and to account for specifics of a given fund.

Note: If – like in this case – only one preprocessing function is used, care has to be applied initializing the weights for the neural net learning. Specifically the Z's should be "spaced", else they might converge to the same value during learning, resulting in an unefficient double up of rules.

The FSM was initially trained to predict the first monthly return following immediately the training period plus the validation period. The validation period was taken to be one month only, immediately following the training period and preceding the month to be predicted. Monthly adaption to new data was used to generate the entire sequence of monthly returns. The linear correlation between the actual returns and the out of sample predictions was 85% for the index over the period Juli 1994 to June 2004 and 72% for the single fund over the period April 1999 to April 2004. The Spearman Rank Correlation was used in conjunction with the linear correlation to account for more general, nonlinear comovements of the two return series and turned out to be 82% for the index and 63% for the single fund over the respective periods. The Kolmogoroff-Smirnov Test was employed to test for the null hypothsis of equality of the empirical distribution functions. The test for the index prevailed up to α=0.41, for the single fund up to 0.14, strongly emphasising the similiarity of the two distribution functions and thus highlighting the conditional predictive capability, even of a simple FSM. Exhibit 1 shows the results of the actual monthly returns of the Index versus the out of sample returns of the FSM. Exhibit 2 shows the empirical distribution function of the monthly index returns vs the empirical distribution function of the out of sample FSM predicted returns. Exhibit 3 shows the results of the actual monhtly returns of the single fund vs. the out of sample monthly FSM returns.

6. Conclusions

While for a description of qualitative hedge fund behaviour simpler methods may suffice, to be able to quantify future hedge fund behaviour a further sharpening of previous methods is required to provide the needed predictive power. Letting the data determine the instruments, the trading rules and the degree of nonstationarity, the FSM constructs a replicating portfolio that depends in a nonlinear and dynamic way on market factors. It does so by distinguishing between factors in which positions will be taken and factors purely serving the decision process for trading, allowing for a flexible formulation of trading rules that determine the exposures to assets. By learning and adapting to both, the dynamics and the optimal factor set, in an integrated way, the FSM is able to capture complex trading behaviour and adjusts quickly to shifts in trading paradigms. In addition, constructing a replicating portfolio, the FSM has also substantial explanatory power. The predictive capabilites of the FSM might therfore prove useful in building efficient and reliable risk measures.

References:

Abernathy, J. and Weisman, A. *"The Dangers of Historical Hedge Fund Data",* Risk Budgeting (London: Risk Books), (2000).

Aggrawal,V. And Naik, N. *"Risk and Portfolio Decisions Involving Hedge Funds",* The Review of Financial Studies, (2002).

Artzner, P.; Delbaen, J.; Eber, J. and Heath, D. *„Coherent Risk Measures"*, Mathematical Finance 9, pp 203-228 (1999)

Breiman, L *" Bagging Predictors",* Machine Learning, 24, pp 123 – 140, (1996)

Cherubini, U., Luciano, E. And Vecchiato, W. "Copula Methods in Finance", John Wiley Series, forthcoming ,(2004)

Fung, W. And Hsieh, D.A. *"Empirical Charcteristics of Dynamic Trading Stratgeis: The Case of Hedge Fundss",* Review of Financial Studies, 10, pp 275 - 302 (1997).

Fung, W. And Hsieh, D.A. *"The Risk of Hedge Fund Strategies: Theory and Eviodence from Trend Followers",* Review of Financial Studies, 41 (2001), pp.313-341.

Fung, W. And Hsieh, D.A. *"The Risk in Fixed-Income Hedge Fund Styles",* Journal of Fixed Income, 58, 16 – 27, (2002).

Fung, W. And Hsieh, D.A. *"The Risk in Equity Hedge Fund Styles",* Working Paper, Duke University (2003).

Gupta,B.; Cerrahoglu, B. and Daglioglu, A *"Evaluating Hedge Fund Performance: Traditional versus Conditional Approaches",* Journal of Alternative Investments, Winter 2003.

Hutchinson, J.; Lo,A. And Poggio, T. "*A Nonparametric Approach to the Pricing and Hedging of Derivative Securities via Learning Networks"*, Journal of Finance 49 (1994).

Jang, J.-S. , Sun, C.-T. And Mizutani, E. “Neuro Fuzzy and Soft Computing”, Prentice Hall (1997)

Nelson, R. “Introduction to Copulas”, Springer Verlag, (1998)

Schneeweis, T. and Spurgin, R.*“Multifactor Analysis of Hedge Fund, Managed Futures and Mutual Fund Return and Risk Characteristics”,* Journal of Alternative Investments, Fall 1998.

Schneeweis, T. and Spurgin, R.*“Qunatitative Analysis of Hedge Funds and Managed Futures and Risk Characteristics”,* In R. Lake, ed., Evaluating and Implementing Hedge Fund Strategies, (1999).

Sharpe, W. *“Asset allocation: Mangement Style and Performance Measurement”,* Journal of Portfolio Management, Vol 18, No 2 (1992).

Appendix 1

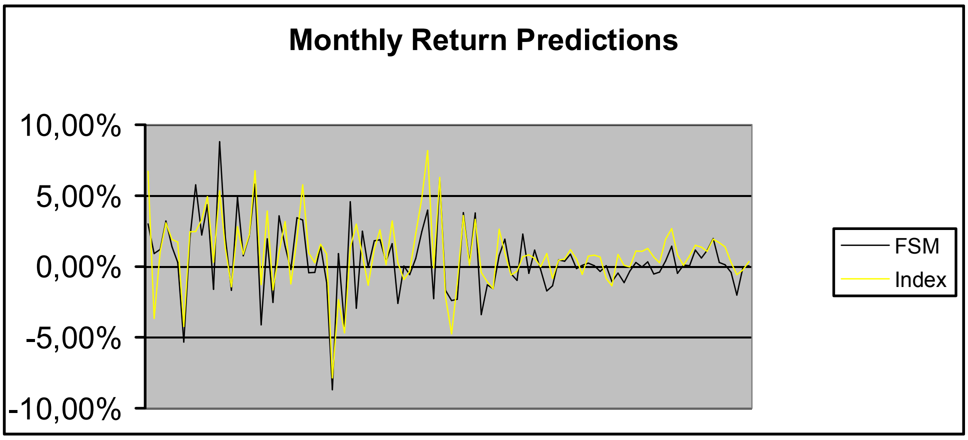

Exhibit 1: *Actual monthly returns of the CSFB Tremont Composite Index vs. monthly, out of sample FSM returns during the period July 1994 – June 2004.*

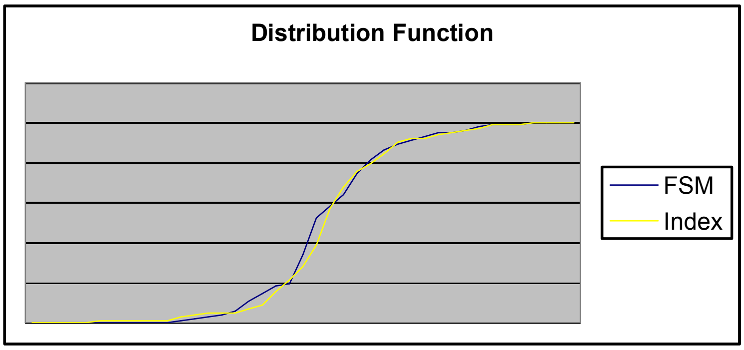

Exhibit 2: *Distribution function of the actual monthly returnsof the CSFB Tremont Composite Index vs. Distribution function of the monthly out of sample FSM returns during the period July 1994 – June 2004.*

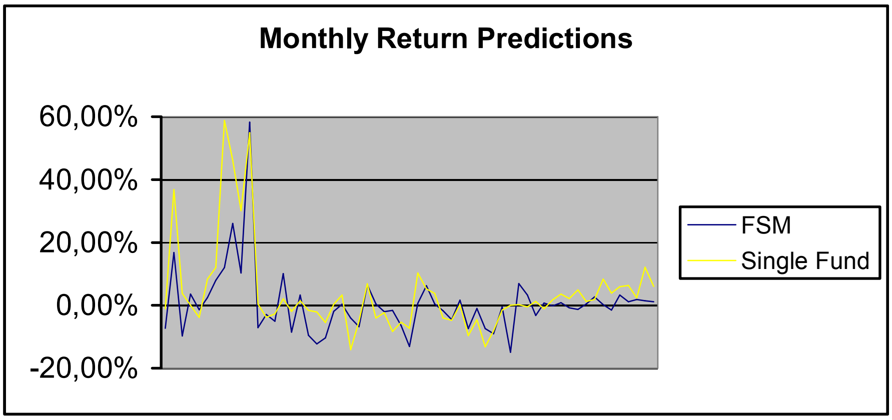

Exhibit 3: *Actual monthly returns of the single long/short equity fund vs. monthly, out of sample FSM returns during the period April 1999 – April 2004.*