ON BIOLOGICALLY INSPIRED PREDICTIONS OF THE GLOBAL FINANCIAL CRISIS

Peter Sarlin
Turku Centre for Computer Science - TUCS
Department of Information Technologies
Åbo Akademi University
FIN-20520, Turku, Finland
E-mail: psarlin@abo.fi

KEYWORDS
financial crises, early warning system, neural networks, genetic algorithms, neuro-genetic model.

ABSTRACT

This paper evaluates the performance of biologically inspired early warning systems (EWS) for systemic financial crises. We create three EWSs: a logit model, a standard back-propagation neural network (NN) and a neuro-genetic (NG) model that uses a genetic algorithm for choosing the optimal NN configuration. The performance of the NN-based models are compared with the benchmark logit in terms of utility for policymakers. For creating the NN-based EWSs, we use two training schemes for parsimonious and generalized models and advocate adopting to the EWS literature the scheme using validation sets for better generalization of data-driven models. The performance evaluation shows that NN-based models, in general, outperform the logit model. The key finding is, however, that NG models not only provide largest utility for policymakers as an EWS, but also in form of decreased expertise and labor needed for, and uncertainty caused by, manual calibration of a NN.

INTRODUCTION

The first wave of the ongoing global financial crisis, while being a self-evident truth in ex post predictions, has demonstrated the importance of ex ante identification of elevated risks that may lead to a systemic financial crisis. Early warning systems (EWSs) have generally treated the task of giving signals of a crisis as a binary-choice problem; however, the modeling techniques have evolved over the years. In the very early days of financial stability surveillance, data were analyzed by hand (e.g. Ramser and Foster, 1931). To the turn of the century, most commonly applied methods have been conventional statistical techniques: discriminant analysis (DA) (e.g. Frank and Cline, 1971), signaling approach (e.g. Kaminsky et al., 1998) and different forms of generalized linear models (GLM), such as logit and probit analysis (e.g. Demirgüc-Kunt and Detragiache, 2000). These methods, however, suffer from assumptions violated more often than not, such as normality of the indicators, distributional assumptions on the relationship between the indicators and the response, and the absence of interactions between indicators. For instance, Lo Duca and Peltonen (2011) show that crisis probabilities increase non-linearly with the increase of vulnerabilities. In past years, some non-parametric, oftentimes distribution-free and non-linear, techniques have been introduced to financial stability surveillance. The key methods in non-parametric EWSs have so far been based upon biologically inspired computing in general and artificial neural networks (NNs) in particular.[1] NNs are effective data-driven non-linear function approximators, but alas they are no panacea for binary-choice classification. To fully benefit from capabilities of NNs, they need to be provided with their computational demands (i.e. large samples and computing power) and specific training schemes for generalization.

Recently, circumstances for financial stability surveillance have changed. While the degree of financial integration and connectivity, and its interplay with the stability of the system, has increased, the soar in data availability and computing power has been remarkable. The former factor advocates further research on early warning modeling whereas the latter enables data-driven modeling of complex relationships. Even though non-parametric methods have produced some promising results in crisis prediction, they have seldom been implemented for financial stability surveillance in practice.[2] One reason might be the complexity of setting up a non-parametric model, not the least a NN. While there exist a large number of "rules-of-thumbs" for parameterization of NNs, manual calibration is still an expertise-demanding and laborious task as well as increases the risk of being trapped in a suboptimum. To this end, Sarlin and Marghescu (2012)

[1] NNs have been implemented as EWSs in, for instance, Nag and Mitra (1999), Franck and Schmied (2004), Peltonen (2006), Aminian *et al*. (2006), Fioramanti (2008), Sarlin and Marghescu (2011) and Sarlin and Peltonen (2011).

[2] For instance, recent work at the European Central Bank, such as Alessi and Detken (2011), Lo Duca and Peltonen (2011) and Hollo *et al*. (2012), includes little computational intelligence.

Proceedings 26th European Conference on Modelling and Simulation ©ECMS Klaus G. Troitzsch, Michael Möhring, Ulf Lotzmann (Editors)
ISBN: 978-0-9564944-4-3 / ISBN: 978-0-9564944-5-0 (CD)

used genetic algorithms (GAs) for finding an optimum model configuration for a NN in currency crisis prediction. Another obstacle for financial stability surveillance might be the problem of parsimonious and generalized modeling; the analyst should know how to avoid overfitting. While NN-based EWSs have generally applied early stopping at the level of performance of a benchmark statistical model (e.g. Fioramanti, 2008; Sarlin and Marghescu, 2012), there is still room for improving accuracy with a balance between complexity and generalization power. What is missing from the EWS literature is using a third dataset, a validation set, for optimizing generalization of a model, something common already in the early NN literature (e.g. Ripley, 1994; Hagan et al., 1996).

This paper tests the usefulness of NN-based models for predicting systemic financial crises. First, we test a wide range of manually calibrated configurations for standard NNs. Second, we apply a standard GA for finding the optimal configuration of the NN – and coin it a neuro-genetic (NG) model. The rationale for this is to test whether NN-based models are better than logit models and the effect of automated calibration of the NG model. For parsimonious models, we implement two training schemes. First, we test whether NN-based models that perform equally well as a logit model on in-sample data perform better on out-of-sample data. Second, we train NN-based models by optimizing utility for policymakers on a third dataset, a validation set. This indicates how much better, if at all, the NN-based models performs when attempting optimal generalization.

METHODOLOGY

This section introduces biologically inspired computing, the underlying data and the evaluation framework. Biologically inspired computing involves processing data in ways that imitate how nature works. While NNs attempt to mimic the functioning of biological neurons in human brains, GAs imitate natural biological evolution dynamics. The dataset consists of two parts: macro-financial indicators and dates of financial crises. And the evaluation framework attempts to measure utility for policymakers by incorporating preferences between type I and type II errors.

Neural networks and genetic algorithms

NNs are data-driven non-linear function approximators composed of a system of units, which are interconnected by weighted links. Units of a NN are ordered into three layers: input, hidden and output layers. Through learning, the weights between units are set to find the best possible approximation of a function f that maps input data x to output values y, i.e. $y=f(x)$. The learning of the back-propagation algorithm can be thought of as a two-step process. First, it evaluates the derivative of the error function with respect to the weights. Then, the obtained derivatives are used for optimal adjustment of the weights.

The NN used in this paper is the very basic multi-layer perceptron with back-propagation learning.[3] The outputs of units U_j are transformed to an output of the NN Y_j as follows:

$$U_j = \sum (X_i * w_{ij}) \quad (1)$$

$$Y_j = f_{th}(U_j + t_j) \quad (2)$$

where U_j equals the sum of inputs X_i multiplied by weights w_{ij} and output Y_j is the sum of the internal values U_j and a bias t_j transformed by an activation function f_{th} into [0,1] and by a threshold *th* into {0,1}. In this paper, we test the following functions f_{th} for a mapping into [0,1]: sigmoidal, hyperbolic, gaussian and linear.

The back-propagation algorithm functions in two steps. It starts at the output layer by computing first the error term e_j and then the weight w_{ij}:

$$e_j = Y_j * (1 - Y_j) * \left(d_j - Y_j\right) \quad (3)$$

$$w_{ij} = w'_{ij} + (1 - M) * \alpha * e_j * (X_i + \varepsilon) + M * (w'_{ij} - w''_{ij}) \quad (4)$$

where the error term e_j equals the product of the output Y_j, its complement $1 - Y_j$, and the difference between the desired output d_j and the output Y_j. Then, weight w_{ij} is determined by the previous weight w'_{ij}; the product of the complement of the momentum factor $1 - M$, learning rate α, the error term e_j, and the input X_i and its input noise ε; and the momentum factor multiplied by the difference between the previous and next to previous weights w'_{ij} and w''_{ij}. On the next layer, while the adjustment proceeds as in Eq. (4), e_j is determined by a modified version of Eq. (3):

$$e_j = Y_j * (1 - Y_j) * \sum (e_k * w'_{jk}) \quad (5)$$

where the sum of the error terms e_k for each neuron k in the succeeding layer multiplied by the respective

[3] The model specification follows, but extends, that in Sarlin and Marghescu (2012). For more detailed presentations of the algorithms, see for example Hagan *et al.* (1996), Haykin (1998) and Bishop (2007).

weights from the previous epoch w'_{ij} replaces the difference between the desired and actual output in Eq. (3). Eq. (5) is applied to each hidden layer. This procedure iterates until fulfilment of a specified stopping criterion or after a specified number of epochs.

An obvious feature of this types of standard NN frameworks is that several parameters need to be specified. These include the learning rate α, momentum factor M, input noise ε, and the number of hidden layers and hidden units per layer. Learning rate sets the extent of adjustment to the old weight, while momentum allows a change to the weights to persist for a number of epochs. Input noise is a small random variation to each input that prevents learning of exact input values and increases thus generalization. In addition, the input-output mapping $f_{th} \in [0,1]$ performed by the activation function can take various forms. Another delicate question for NN modeling is the number and optimum combination of input indicators included in a model. Supervision of an expert analyst is required for obtaining a final, but still subjective, model. User-specified parameterization is not only time-consuming, but also has the drawback of using personal, possibly biased, experience in setting parameters that might overlook optimal solutions. As manual repetition obviously does not ensure even local convergence, it increases the risk of the NN being trapped in a suboptimum.

The choice of indicators, NN parameters and NN architecture can be resolved by a GA (Holland, 1975). GAs can be thought of as stochastic search methods for solving an optimization problem (for instance, maximize utility for a policymaker of an EWS over different NN configurations):

$$\max\{g(s)|s \in \Omega\} \quad (6)$$

where s is a candidate binary string solution to function g that belongs to the solution set Ω. In biological terms this translates to 'individuals' (candidate solutions) represented by 'structures' of genetic information (genotypes) and the 'population' (solution set Ω). The population evolves by genetic operators, such as the weakest structures replacing, by some 'reproduction', the best structures, and by random variation. 'Crossovers' determine the intermingling of features on the same string (or reproduction), where a one-point crossover creates interchanged offspring of the two parent solutions at a randomly chosen point, and 'mutations' determine a random change in a structure. On the candidate solutions, a crossover is performed on the best solutions and a mutation on the rest. Hence, specifying a parameter value for the mutation rate also sets the crossover rate. We specify each candidate solution to have a structure composed of three strings: one for input indicators, one for the NN architecture and one for the NN parameters. This gives the option to only optimize parts of the NN configuration. Let us exemplify the functioning with a simple example. For five inputs, three solutions could be, for instance: s_1={1,1,1,0,1}, s_2={0,0,1,1,1} and s_3={1,0,1,1,1}, where s_1 includes inputs 1-3 and 5, s_2 only inputs 3-5 and s_3 inputs 1 and 3-5, and the performance of $s_1 > s_3 > s_2$. Then, a mutation rate of 0.4 would introduce a random mutation to s_2, such that s_2={0,0,0,1,1}, for instance. If s_1 is read first and crossed over with s_3 after the second element, the offspring would be {1,1,1,1,1}, i.e. inclusion of all inputs 1-5.

Data

The dataset is an extended version of that in Lo Duca and Peltonen (2011) consisting of macro-financial indicators and the occurrence of a systemic financial crisis. The sample includes 28 countries, 10 advanced and 18 emerging economies, from 1990:1–2011:2. The macro-financial indicators consist of 14 quarterly country-specific vulnerabilities that proxy asset price developments and valuations, credit developments and leverage as well as traditional macroeconomic measures. The indicators, and their properties, are shown in Table 1. The statistical properties show that they are skewed and non-mesokurtic, and hence not exhibiting normal distributions. Each indicator is normalized into historic country-specific percentiles to control for cross-country differences, as is common in the EWS literature.

Table 1: The indicatiors and their statistical properties

No.	Variable	Abbreviation	Mean	SD	Min.	Max.	Skew.	Kurt.	KSL	AD
1	Inflation[a]	Inflation	0.89	5.17	-10.15	42.53	4.80	26.72	0.29	263.90
2	Real GDP[b]	Real GDP growth	3.73	3.76	-17.54	14.13	-0.86	3.16	0.06	11.34
3	Real equity prices[b]	Real equity growth	5.93	33.01	-84.40	257.04	0.99	4.31	0.05	7.28
4	Credit to private sector to GDP[a]	Leverage	3.48	51.64	-62.78	1673.04	22.76	673.35	0.29	Inf
5	Stock market capitalisation to GDP[a]	Equity valuation	3.90	28.32	-62.79	201.55	0.77	2.41	0.03	3.86
6	Government deficit to GDP[c]	Government deficit	0.01	0.05	-0.19	0.22	-1.09	3.46	0.09	35.90
7	Global Inflation[a]	Global inflation	0.03	0.64	-1.33	2.29	0.71	1.28	0.08	12.12
8	Global real equity prices[b]	Global real equity growth	2.31	19.08	-40.62	37.77	-0.57	-0.68	0.15	41.90
9	Global stock market capitalisation to GDP[a]	Global equity valuation	0.89	17.41	-40.54	27.46	-0.50	-0.43	0.09	19.11
10	Real credit to private sector to GDP[b]	Real credit growth	234.07	4724.00	-69.42	101870.34	20.76	429.59	0.51	Inf
11	Current account deficit to GDP[c]	CA deficit	-0.02	0.07	-0.27	0.10	-0.98	0.73	0.09	33.12
12	Global real GDP[b]	Global real GDP growth	1.84	1.59	-6.34	4.09	-3.02	11.74	0.20	122.16
13	Global real credit to private sector to GDP[b]	Global real credit growth	3.87	1.68	-0.23	7.20	-0.21	-0.31	0.07	8.82
14	Global credit to private sector to GDP[a]	Global leverage	1.15	2.79	-2.79	11.21	1.84	3.40	0.22	105.26

Note: Transformations: [a], deviation from trend; [b], annual change; [c], level. KSL: Lilliefors' adaption of the Kolmogorov-Smirnov normality test. AD: the standard Anderson-Darling normality test. All normality tests are significant on a 1% level.

Using the financial stress index (FSI) in Lo Duca and Peltonen (2011), we have an objective criterion to date financial crises.[4] The FSI is transformed to an ideal leading indicator that perfectly signals a systemic financial crisis before the event (i.e. the predicted variable) by first defining crisis periods as those when the FSI moves above the 90th percentile of a country-specific distribution, and then by setting the pre-crisis dummy to '1' in the 18 months preceding the systemic financial crisis, and to '0' in all other periods. While the NN and GA methodology is flexible for different forms of input data, the rationale behind using a dataset previously introduced in the literature is to test performance on objective data rather than those that fit the model (i.e. data dredging). To objectively test model performance on the global financial crisis that started in 2007, the sample is split into two sets: the in-sample set (1990:4–2005:1) and out-of-sample set (2005:2–2009:2).

Performance evaluation

In this paper, performance of EWSs is evaluated by computing a goodness-of-fit measure that incorporates policymakers preferences. As is common in the literature, the EWSs issue signals when the estimated probability of a crisis is above a specified threshold. The outcomes are classified into a contingency matrix, including true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). We follow the approach employed in Demirgüc-Kunt and Detragiache (2000) and Alessi and Detken (2011) by calibrating an optimal model and threshold for policy action that maximizes the utility for a policymaker with specific preferences about giving false alarms (type I errors) and missing crises (type II errors). The policymaker has the following loss function:

$$L(\mu) = \mu(FN/(FN+TP)) + (1-\mu)(FP/(FP+TN)) \quad (7)$$

where the parameter μ represents the relative preference of the policymaker between FNs and FPs. Given differences in class size, the policymaker is equally concerned of missing crises and issuing false signals for a value of $\mu = 0.5$, less concerned of issuing false signals when $\mu > 0.5$, and less concerned of missing crises when $\mu < 0.5$. Hence, the utility of one specific threshold is given by subtracting from the expected value of a best-guess with those preferences:

$$U = Min(\mu, 1-\mu) - L(\mu) \quad (8)$$

The performance of a model is solicited by choosing a threshold on the probability of a crisis such that utility U is maximized.

BIOLOGICALLY INSPIRED PREDICTIONS OF THE GLOBAL FINANCIAL CRISIS

To evaluate the NN and NG models, it is necessary to compare them with a benchmark, such as conventional statistical techniques from previous EWS literature. These methods (e.g. DA and logit/probit analysis) can in fact be related to very simple NNs; in particular, NNs with no hidden layer, single-layer perceptrons, with a threshold activation function and logistic activation function (Ripley, 1994; Sarle, 1994). The obvious problem with most statistical methods is that all assumptions on data properties are seldom met. As methods from the GLM family have less restrictive assumptions (e.g. normality of the indicators), we turn to logit analysis. It is preferred over probit analysis as its more fat-tailed nature corresponds better to the rarity of crises.

[4] The FSI consists of five components: the spread of the 3-month interbank rate over the rate of the 3-month government bill; quarterly equity returns; realized volatility of a main equity index; realized volatility of the exchange rate; and realized volatility of the yield on the 3-month government bill.

Even very simple NNs have been shown to be universal approximators by following any continuous function to any desired accuracy (Hornik et al., 1989). This said, the focus of data-driven NN applications in real-world settings should rather be on parsimonity and generalization than on fitting models to all non-linearities and complexities in data. Another common concern is the extent of data dredging when conducting data-driven analysis. To this end, we build NN-based EWSs using two objective training, or early stopping, schemes:

- Scheme 1: Training is performed until in-sample performance of a conventional benchmark model has been reached.
- Scheme 2: Data is divided into three datasets: train, validation and test sets. Models are trained on the train set and the one with optimal validation performance is chosen.

The in-sample dataset is used for estimating a logit model. The estimates of the model are then used to solicit the probability of a crisis and further the utility for policy action. In training scheme 1, the in-sample utility of the logit model for policy action, U=0.25, is used as a benchmark (stopping criterion) when training the NNs. The rationale behind this is twofold: it attempts to prevent overfitting and enables testing whether a NN that is equally good on the in-sample performs better on out-of-sample data. We manually test the performance of NN configurations over a wide set of possibilities as well as of the automated NG model. In training scheme 2, the in-sample dataset is randomly split as follows: 80% train set and 20% validation set. This gives us three datasets: train (in-sample, 80%), validation (in-sample, 20%) and test (out-of-sample) sets. Then we train NN and NG models by optimizing utility for policymakers on the validation set. In practice, we train models for 200 epochs, evaluate them at each epoch and choose the one that maximizes utility on the validation set. This allows testing how much better, if at all, the NN and NG models perform when attempting an optimal model. As NNs are sensitive to initial conditions of the weights, the training of NN and NG models is repeated ten times with randomized starting weights and biases, and then the one with the fastest convergence (least epochs) is chosen. Robustness of GA parameters is tested by varying generation count, population size and mutation rate. The number of crossovers and the fitness criterion are kept fixed because of the shortness of the strings and the aim of the GA training (i.e. utility for policymakers), respectively.

We assume a benchmark policymaker to be equally concerned of missing crises and giving false alarms (μ=0.5), but also show model performance for those with other preferences (μ={0.4,0.6}). For all models using training scheme 1 (in-sample U=0.25), the final NN elements and GA parameters, as well as their out-of-sample utility for policymakers, are shown in Table 2.[5] The results clearly depict differences in model performance (best models per μ are bold). The NN-based models, while having similar in-sample performance by definition, show consequently better out-of-sample performance than the benchmark logit model. Best overall performance is shown by the NG model and the best NN model (NN1) follows parameterization practices common in the literature (see e.g. Peltonen, 2006; Fioramanti, 2008). Most notably, the optimal GA configuration uses only 9 indicators (bold in Table 2). When examining differences in performance for different policymakers' preferences, one can observe that the logit model fails for those emphasizing the share of correctly called crises (U=0), the NN results are somewhat mixed and the NG model brings utility for all types of policymakers (U={0.10,0.19,0.08} for μ={0.4,0.5,0.6}). While NNs are heuristic in nature, the consistency in the slight superiority is likely to be a result of the parsimonious training scheme.

Training scheme 2 attempts better generalization by being less restrictive in terms of parsimonity but still preventing overfitting. Table 3 summarizes in-sample (includes both train and validation sets), validation and out-of-sample performance of the logit model, the best-performing NN model (NN1) and the NG model using training scheme 2. The table shows that, in principle, when allowing for longer training, and thus also a better fit to data, model performance improves on all datasets (best performing model per μ and dataset is bold). This is obvious when it comes to in-sample data, but the validation as well as out-of-sample data still need sufficient parsimonity for decent performance. Compared to training scheme 1, while there is only a small increase in out-of-sample utility of NN1 (from 0.18 to 0.19), the performance of the NG model increases significantly (from 0.19 to 0.22). Out-of-sample utility of the NG model is 70% better than that of the logit model and 17% better than that of the model from training scheme 1.

[5] Model robustness has been tested with respect to different forecast horizons. In general, for longer horizons (e.g. 24 months) there is a slight increase in utility and for shorter (e.g. 6 and 12 months) a slight decrease. However, the relative performance is similar to the benchmark results with a horizon of 18 months.

Table 2: The Configuration of the Models as well as Their Utility for Policymakers using Training Scheme 1

	Logit	*NN1*	*NN2*	*NN3*	*NN4*	*NN5*	*NN6*	*NN7*	*NN8*	*NN9*	*NG*
NN configuration											
No. of nodes in the input layer:	-	14	14	14	14	14	14	14	14	14	9
No. of nodes in the output layer:	-	1	1	1	1	1	1	1	1	1	1
Number of hidden layers:	-	1	1	1	2	2	1	1	1	2	1
No. of nodes in layer 1:	-	2	2	8	2	8	2	2	2	20	4
No. of nodes in layer 2:	-	0	0	0	2	8	0	0	0	20	0
Activation function:	-	Sigmoid	Sigmoid	Sigmoid	Sigmoid	Sigmoid	Gaussian	Hyperbolic	Linear	Sigmoid	Sigmoid
No. of training epochs:	-	15	11	9	22	11	21	13	12	5	6
Learning rate α:	-	0.9	0.8	0.95	0.9	0.9	0.9	0.9	0.9	0.9	0.97
Momentum m:	-	0.9	0.2	0.8	0.9	0.9	0.9	0.9	0.9	0.9	0.15
Input noise ε:	-	0.01	0	0	0.01	0.01	0.01	0.01	0.01	0.01	0.01
GA configuration											
Generation count:	-	-	-	-	-	-	-	-	-	-	10
Population size Ω:	-	-	-	-	-	-	-	-	-	-	3
No. of crossovers:	-	-	-	-	-	-	-	-	-	-	1
Mutation rate:	-	-	-	-	-	-	-	-	-	-	0.10
Fitness criterion:	-	-	-	-	-	-	-	-	-	-	train error
Out-of-sample utility											
μ=0.4	0.07	0.08	0.06	0.07	0.06	0.08	0.06	0.07	0.08	0.08	**0.10**
μ=0.5	0.13	0.18	0.15	0.16	0.16	0.16	0.14	0.14	0.17	0.16	**0.19**
μ=0.6	0.00	**0.08**	0.04	0.05	0.06	0.04	0.02	0.01	0.07	0.05	**0.08**

Table 3: The Utility for Policymakers of Models on In-sample, Validation and Out-of-sample Data using Training Scheme 2

Model	**Epochs**	**In-sample**			**Validation set**			**Out-of-sample**		
		μ=0.4	*μ=0.5*	*μ=0.6*	*μ=0.4*	*μ=0.5*	*μ=0.6*	*μ=0.4*	*μ=0.5*	*μ=0.6*
Logit	-	0.16	0.25	0.14	-	-	-	0.07	0.13	0.00
NN1	36	0.19	0.29	0.19	0.09	0.19	0.10	0.09	0.19	0.09
NG	17	**0.20**	**0.30**	**0.21**	**0.13**	**0.22**	**0.12**	**0.13**	**0.22**	**0.11**

CONCLUSIONS

This paper compared the utility for policymakers of a NN and NG model with that of a conventional logit model. Given similar in-sample performance, we have shown that the NN-based, in particular NG, models outperform the conventional statistical model. The introduction of an early stopping method using a third dataset, a validation set, further increases the difference in utility. Hence, the main finding is that NG models not only provide more utility for policymakers in terms of predictive capabilities, but also in form of decreased expertise and labor needed for, and uncertainty caused by, manual calibration of a NN. The paper also motivates adoption of validation sets for increased generalization of data-driven EWSs.

ACKNOWLEDGMENTS

The author wants to thank Barbro Back, Marco Lo Duca and Tuomas Peltonen, as well as seminar participants at the Data Mining and Knowledge Management Laboratory at Åbo Akademi University, for useful comments and discussions.

REFERENCES

Alessi, L. and C. Detken. 2011. "Quasi real time early warning indicators for costly asset price boom/bust cycles: a role for global liquidity." *European Journal of Political Economy* 27, No.3, 520–533.

Aminian, F.; E. Suarez; M. Aminian; and D. Walz. 2006. "Forecasting Economic Data with Neural Networks." *Computational Economics* 28, 71–88.

Bishop, C. 2007. *Pattern Recognition and Machine Learning*. Springer, Germany, Heidelberg.

Demirgüc-Kunt, A. and E. Detragiache. 2000. "Monitoring Banking Sector Fragility: A Multivariate Logit Approach." *World Bank Economic Review* 14, 287–307.

Fioramanti, M. 2008. "Predicting sovereign debt crises using artificial neural networks: a comparative approach." *Journal of Financial Stability* 4, No.2, 149–164.

Franck, R. and A. Schmied. 2003. "Predicting currency crisis contagion from East Asia to Russia and Brazil: an

artificial neural network approach." AMCB Working Paper No 2.
Frank, C. and W. Cline. 1971. "Measurement of Debt Servicing Capacity: An Application of Discriminant Analysis." *Journal of International Economics* 1, 327–344.
Hagan, M.T.; H.B. Demuth; and M. Beale. 1996. *Neural Networks Design*. PWS Publishing, Boston, MA.
Haykin, S. 1998. *Neural Networks: A Comprehensive Foundation*. Second Edition. Prentice Hall, New Jersey.
Holland, J. 1975. *Adaptation In Natural and Artificial Systems*. The University of Michigan Press, Ann Arbour.
Hollo, D.; M. Kremer; and M. Lo Duca. 2012. "CISS – A composite indicator of systemic stress in the financial system." ECB Working Paper, forthcoming.
Hornik, K.; M. Stinchcombe; and H. White. 1989. "Multilayer feedforward networks are universal approximators." *Neural Networks* 2, 359–366.
Kaminsky, G.; S. Lizondo; and C. Reinhart. 1998. "Leading indicators of currency crises." *IMF Staff Papers* 45, 1–48.
Lo Duca, M. and T. Peltonen. 2011. "Macro-Financial Vulnerabilities and Future Financial Stress – Assessing Systemic Risks and Predicting Systemic Events." ECB Working Paper, No. 1311.
Nag, A. and A. Mitra. 1999. "Neural networks and early warning indicators of currency crisis." Reserve Bank of India, Occasional Papers 20(2), 183–222.
Peltonen, T.A., 2006. "Are emerging market currency crises predictable? A test." ECB Working Paper, No. 571.
Ramser, J. and L. Foster. 1931. "A demonstration of ratio analysis." Bureau of Business Research, University of Illinois,Urbana, IL, Bulletin 40.
Ripley, B. 1994. "Neural Networks and Related Methods for Classification." *Journal of the Royal Statistical Society* 56, 409–456.
Sarle, W. 1994. "Neural Networks and Statistical Models". In *Proceedings of the Nineteenth Annual SAS Users Group International Conference*, Cary, NC: SAS Institute, 1538–1550.
Sarlin, P. and D. Marghescu. 2011. "Visual Predictions of Currency Crises using Self-Organizing Maps." *Intelligent Systems in Accounting, Finance and Management* 18(1), 15–38.
Sarlin, P. and D. Marghescu. 2012. "Neuro-Genetic Predictions of Currency Crises." *Intelligent Systems in Accounting, Finance and Management*, forthcoming.
Sarlin, P. and T.A. Peltonen. 2011. "Mapping the state of financial stability." ECB Working Paper No. 1382.

AUTHOR BIOGRAPHY

PETER SARLIN was born in Turku, Finland and obtained his MSc (Econ.) degree in 2009 from Åbo Akademi University. At the same university, he is currently pursuing a PhD degree in information systems with an emphasis on financial stability surveillance with computational intelligence. Peter has regularly also visited central banks, such as the European Central Bank and the Bank of Finland. His e-mail address is: psarlin@abo.fi and his Web-page can be found at research.it.abo.fi/personnel/psarlin.