SHORT-TERM LOAD FORECASTING USING ARTIFICIAL NEURAL NETWORKS

Francisco Sandoval

Departamento de Tecnología Electrónica
Universidad de Málaga, Espagne

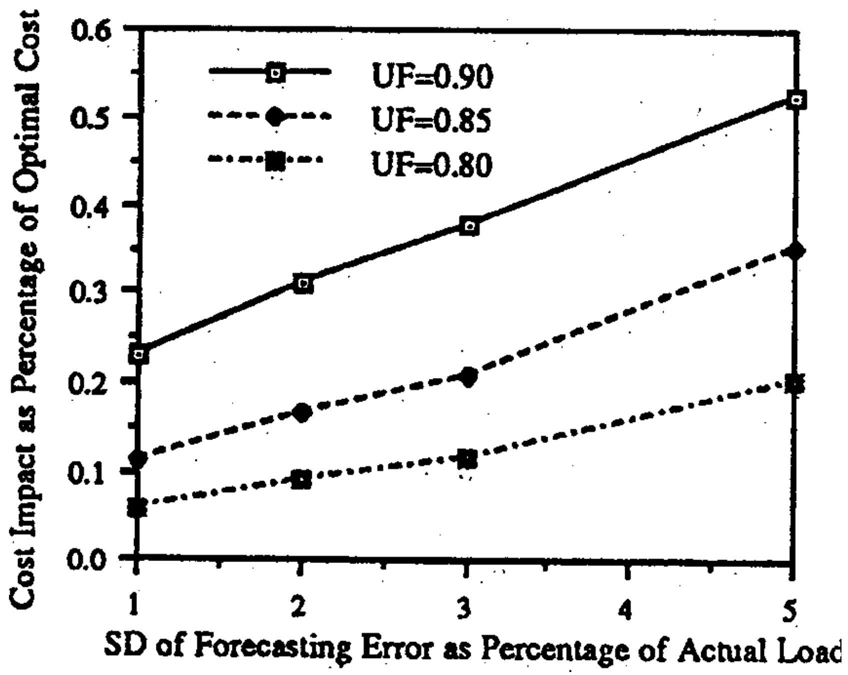
CONTENTS

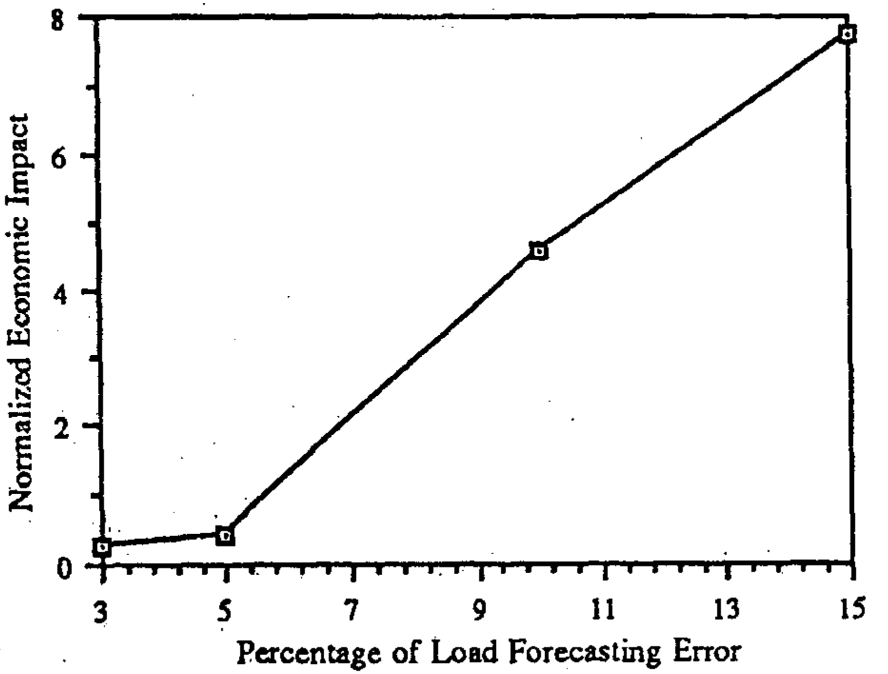
INTRODUCTION

- Electric industry needs to predict load consumption in the short, medium and long term.
- Short term is important for: economic scheduling of generation capacity, security analysis, final price in deregulated markets, etc.

- Load forecasting is a difficult task:
  - ⇨ Load series is complex and exhibits several level of seasonality.
  - ⇨ Many exogenous variables, specially weather-related variables.

INTRODUCTION





Non-linear cost function

INTRODUCTION

- Conventional statistical forecasting models belong to:

    ➪ Time series (univariate) models, where the load is modeled as a function of its past observed values: multiplicative autoregressive models, dynamic linear and non linear models, method based on Kalman filtering, etc.

    ➪ Causal models: the load is modeled as a function of some exogenous factors, specially weather and social variables: Box and Jenkins transfer function, ARMAX models, non parametric regression, etc.

INTRODUCTION

- Artificial intelligence techniques applied to the load forecasting:
    - Expert systems.
    - Fuzzy inference.
    - Artificial neural networks.
    - Hybrids of the above techniques.

- ANNs have not entirely convinced to researches and utilities. Why?

INTRODUCTION

- ANNs are considered as a "black box", and it is difficult to manipulate and to know what happens inside.
- Proposed ANN architectures that seem to be too large for the data samples available to model.
- Models are not systematically tested, and the results of the test not always are satisfactory presented.
- It is difficult to establish benchmarks where to compare procedures and results.

- What could we do to properly design ANN-based forecasting systems?

CONTENTS

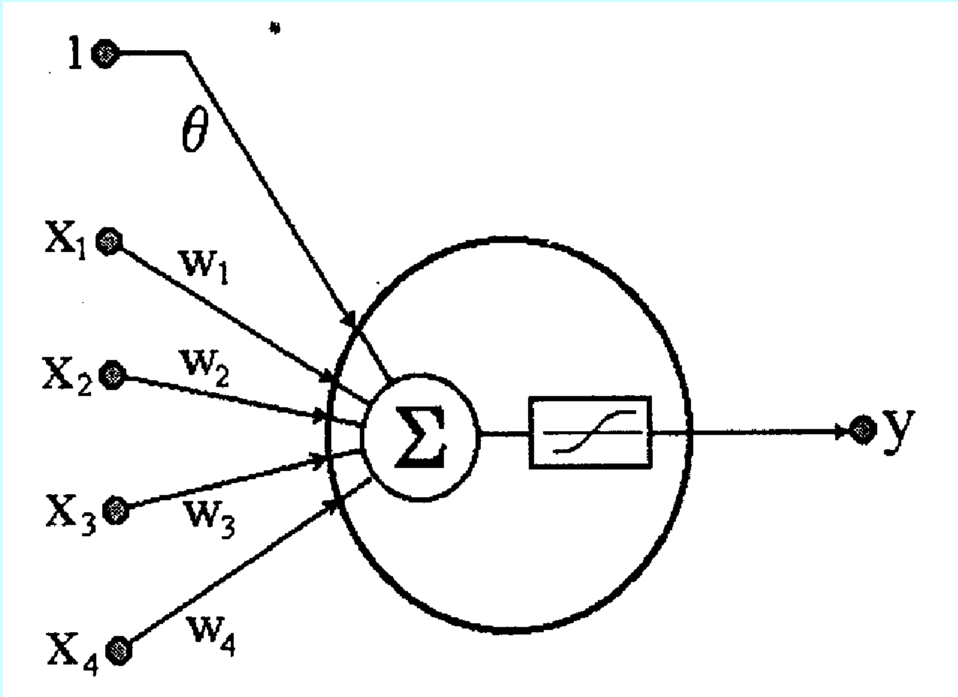
ON ARTIFICIAL NEURAL NETWORKS

- ANNs are a biologically-inspired attractive paradigm of computation for many applications: pattern recognition, system identification, cognitive modeling, etc.

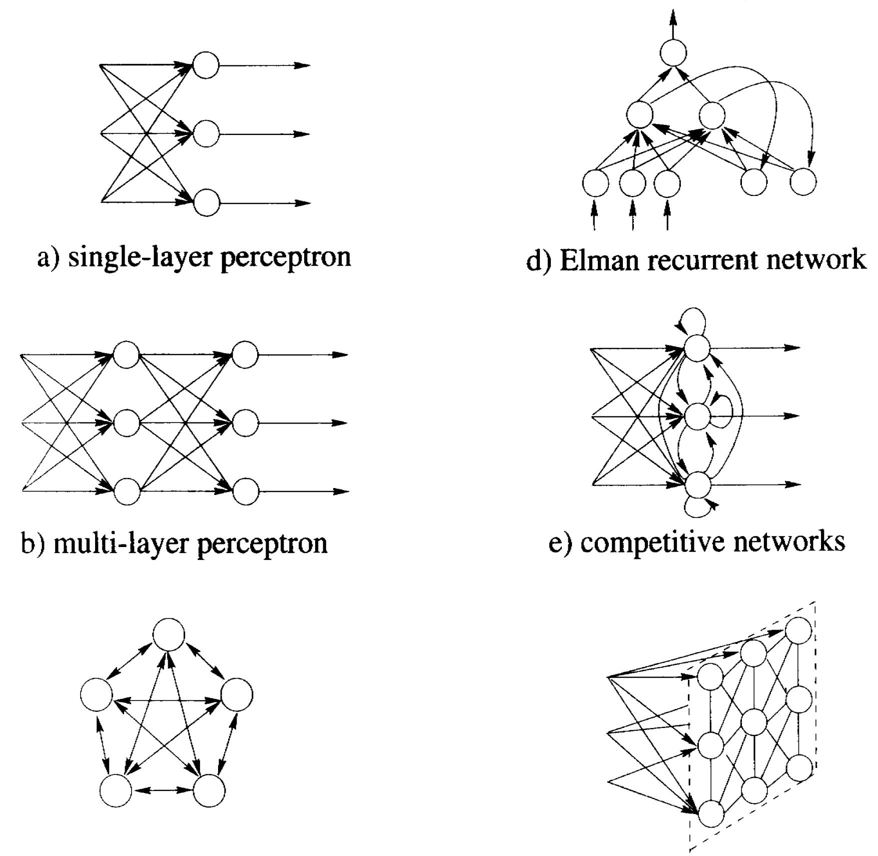
- Properties of ANNs are:
  - ➪ Capability of “learning” and “self-organizing” to carry out a given task: ill-defined and input/output mapping.
  - ➪ Potential for massively parallel computation.
  - ➪ Robustness in the presence of noise.
  - ➪ Resilience to the failure of components.

ON ARTIFICIAL NEURAL NETWORKS

- The basic unit is the neuron: inputs, weights, bias, and ouput.
- Many topologies.



• Topologies of ANN:



a) single-layer perceptron

d) Elman recurrent network

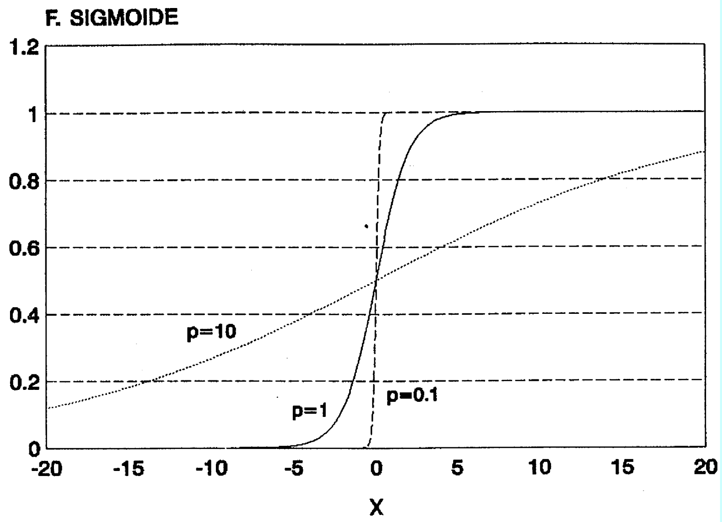
b) multi-layer perceptron

e) competitive networks

c) Hopfield network

f) self-organizing maps

Sigmoid activation function



ON ARTIFICIAL NEURAL NETWORKS

- Neural learning:

  ➪ Supervised learning: perceptron learning algorithm; Least Mean Square (LMS) rule and delta rule; and backpropagation algorithm and its derivatives.

  ➪ Unsupervised learning: Hebbian learning; competitive learning (WTA, vector quantization, and learning vector quantization); self-organization features maps (SOMs).

ON ARTIFICIAL NEURAL NETWORKS

- Suitability of ANN for load forecasting:

  ➪ ANN are able to approximate numerically any continuous function to the desired accuracy. ANNs could be seen as multivariate, nonlinear and nonparametric methods.

  ➪ ANNs are data-driven method, i.e., it is not necessary to postulate tentative models and then to estimate their parameters, so

  ➪ Given a sample of input and output vectors, ANNs are able to automatically map the relationship between them.

ON ARTIFICIAL NEURAL NETWORKS

ANNs and STATISTICS

Neural Networks	Statistics
Learning	Model Estimation
Supervised Learning	Non-linear regression
Unsupervised Learning	Cluster Analysis
Weights	Parameters
Inputs	Independent Variables
Outputs	Dependent Variables

ON ARTIFICIAL NEURAL NETWORKS

- What do ANNs offer with respect statistical methods?
- The answer is not entirely clear, however:
  - ANNs are more attractive for many non-specialists.
  - Statisticians are mainly concentrated on linear models and a small number of parameters.
  - ANNs are easy to implement and can be easily tuned to particular problems.
  - ANNs can be implemented in hardware.
  - ANN can be used as modules in hybrid systems.

- ANNs and statistics are not competitive techniques.

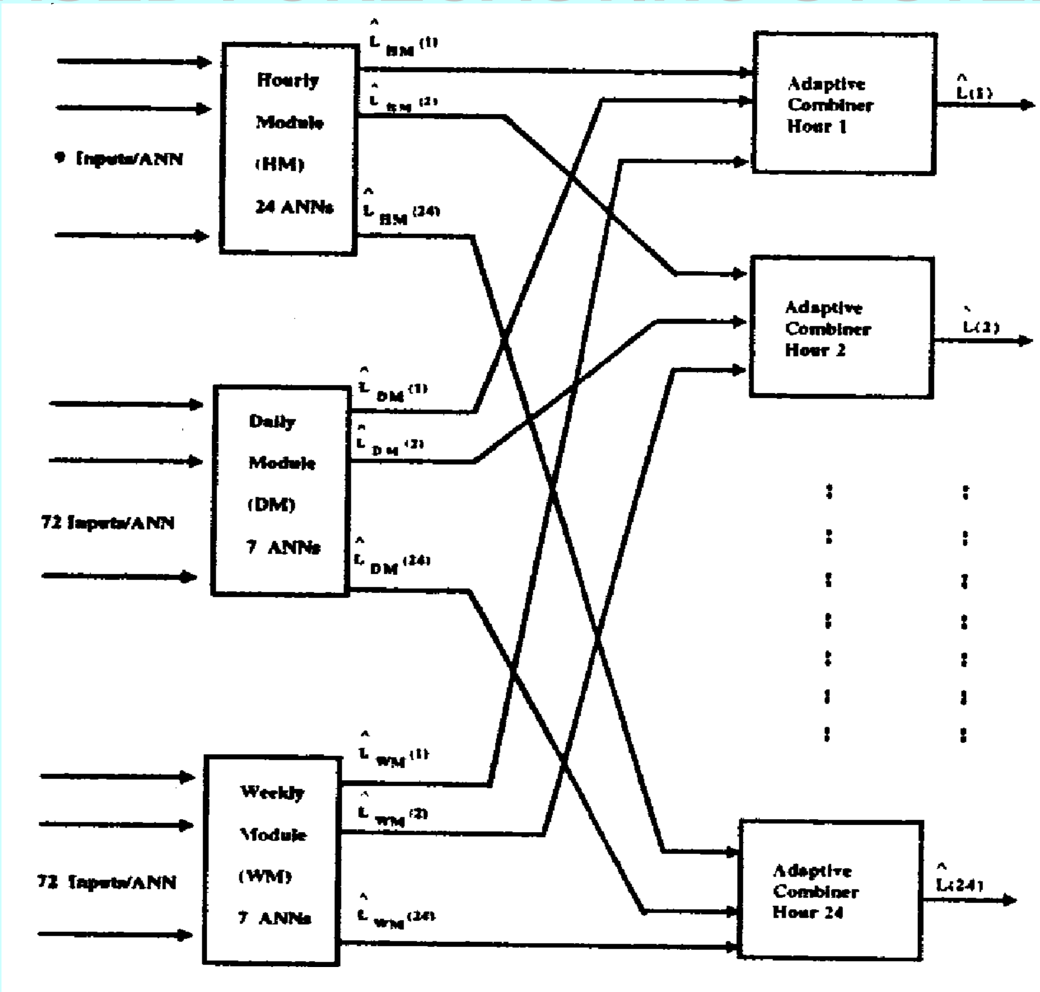
CONTENTS

ANN-BASED FORECASTING SYSTEMS

- Two main architectures: one output node or several output nodes.
- One output node:
  - ⇨ Next hour's load; next day's peak; next day's integral load.
  - ⇨ Forecast load profile: repeatedly forecasting one hourly load at a time; or using a system with 24 ANNs in parallel, one for each hour of the day.
- Several output nodes: 24 nodes to forecast the load profile.
- Other situations

ANN-BASED FORECASTING SYSTEMS

- 153 inputs
- 38 ANNs



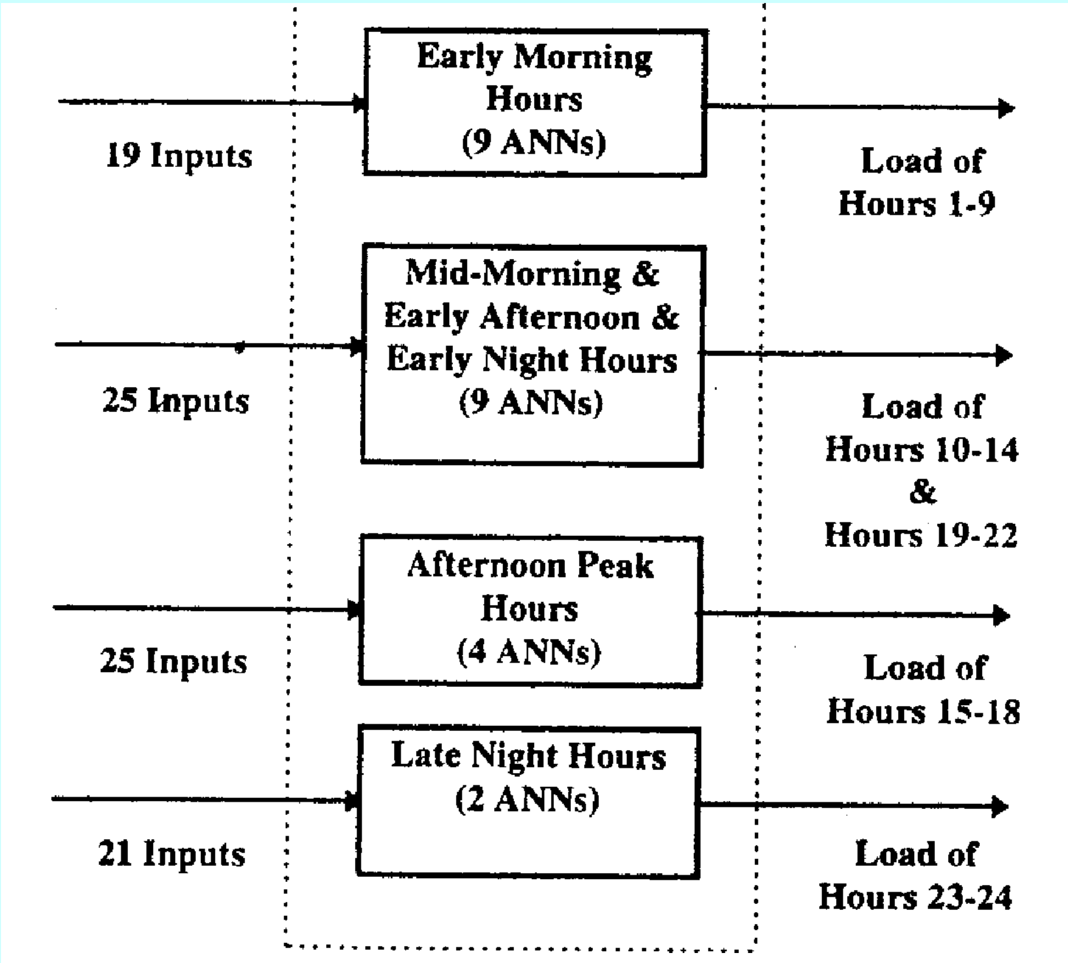
- 24 outputs

Khotanzad et Al, IEEE Trans. Power Systems, 13(4), 1998

Dpto. Tecnología Electrónica
Universidad de Málaga

ANN-BASED FORECASTING SYSTEMS

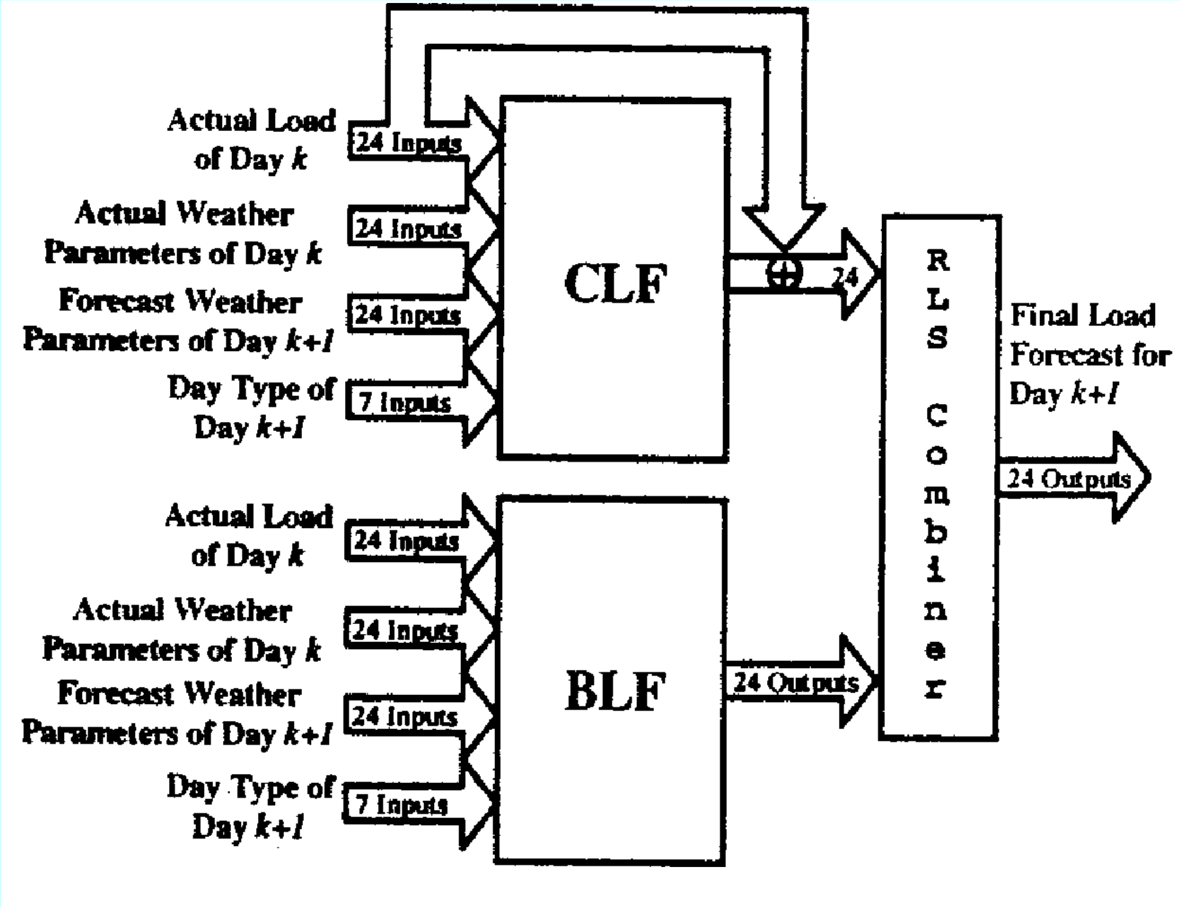
- 90 inputs
- 24 ANNs



- 24 outputs

ANN-BASED FORECASTING SYSTEMS

- 2x79 inputs
- 2 ANNs



24 outputs

ANN-BASED FORECASTING SYSTEMS

		Days-Ahead						
Utility	ANNSTLF Generation	1	2	3	4	5	6	7
1	Three	1.91	2.29	2.53	2.71	2.87	3.03	3.15
	Two	2.11	2.61	2.96	3.26	3.52	3.76	4.25
2	Three	2.72	3.44	3.63	3.77	3.79	3.83	3.80
	Two	2.94	3.57	3.78	3.89	3.99	4.07	4.33
3	Three	1.89	2.25	2.38	2.45	2.53	2.58	2.65
	Two	1.95	2.45	2.66	2.76	2.93	3.04	3.38
4	Three	2.02	2.37	2.51	2.58	2.61	2.65	2.69
	Two	2.86	3.76	4.10	4.32	4.46	4.59	4.94
5	Three	1.97	2.38	2.61	2.66	2.65	2.65	2.74
	Two	2.06	2.38	2.58	2.73	2.84	2.96	3.36
6	Three	1.57	1.86	1.99	2.08	2.14	2.17	2.18
	Two	2.04	2.57	2.79	2.97	3.10	3.23	3.61
7	Three	2.29	2.79	2.90	3.00	3.05	3.10	3.18
	Two	2.39	2.98	3.22	3.38	3.51	3.61	4.00
8	Three	2.22	2.91	3.15	3.28	3.39	3.45	3.50
	Two	2.04	2.57	2.79	2.97	3.10	3.23	3.61
9	Three	1.63	2.04	2.20	2.32	2.40	2.41	2.50
	Two	1.82	2.28	2.50	2.69	2.87	3.06	3.48
10	Three	2.32	2.97	3.25	3.38	3.44	3.52	3.56
	Two	2.40	3.09	3.38	3.60	3.77	3.92	4.25
AVERAGE	Three	**2.05**	**2.53**	**2.72**	**2.82**	**2.89**	**2.94**	**2.99**
	Two	**2.26**	**2.83**	**3.08**	**3.26**	**3.41**	**3.55**	**3.92**
Accuracy Improvement (%)		9	11	12	14	15	17	24

		Utilities					
	Generation	1	2	3	4	5	Avg.
All Hours	Three	4.40	5.42	5.29	4.62	9.68	**5.88**
	Two	5.62	6.89	5.21	5.57	10.07	**6.67**
Peak Load	Three	6.33	5.59	6.27	7.75	9.17	**7.02**
	Two	8.85	10.18	6.84	9.50	10.15	**9.10**

Holiday forecasts

ANN-BASED FORECASTING SYSTEMS

- Fuzzy logic:
  - Front-end fuzzy processor
  - Fuzzy engine after ANN

CONTENTS

PROBLEMS IN DESIGNING ANN-BASED FORECASTING SYSTEMS

- Designing an ANN is not a simple task:
  - Data preprocessing
  - ANN designing itself
  - ANN implementation
  - Validation

CONTENTS

PROBLEMS IN DESIGNING ANN-BASED FORECASTING SYSTEMS

- Data preprocessing: to make the forecasting problem more manageable
- To reduce the dimension of the input vectors.
- To 'clean' the data (removing outliers, missing values or any irregularities, ...)
- Partitioning the input space: classification of the input data (past load profiles or weather data).
  - ⇨ Holidays and special days pose a mayor problem.
  - ⇨ Not many classes to have enough data in each profile.
- Input data covering all the input space.
- Analog and digital input data problem.
- Normalization

CONTENTS

PROBLEMS IN DESIGNING ANN-BASED FORECASTING SYSTEMS

- ANN design:
  - Selecting architecture: mainly the multilayer perceptron (MLP) (feedforward or recurrent networks).
  - Number of hidden layers, neurons per layer, type of activation function, number of output neurons
  - Automatic design using evolutionary techniques (Genetic algorithms, evolution strategies and evolutionary programming).

PROBLEMS IN DESIGNING ANN-BASED FORECASTING SYSTEMS

- ANN design: Number of output neurons

- One output ANN to produce one-step-ahead forecasts.
- Forecasting profiles:
  - ➪ Iterative forecasting. Forecasts of the last hours will be based on the forecast of the earlier ones.
  - ➪ Multi-model forecasting. One model for each hour of the day.
  - ➪ Single-model multivariate forecasting. A multivariate method to forecast all the loads at once, so each profile is represented by a 24-dimensional vector.

PROBLEMS IN DESIGNING ANN-BASED FORECASTING SYSTEMS

- ANN design: Number of input nodes

- The load itself, as the load series is strongly autocorrelated
  - ➪ Forecasting profiles as 24-dimension vectors: data from one or two past days
  - ➪ Forecasting hourly loads: select lagged load values by the analysis of autocorrelation functions (risk of discarding lagged variables with strong nonlinear correlation)
- Exogenous variables. Temperature, which is non-linearly related with the load.
- Other weather variables (cloudiness, humidity, wind, ..) are usually unavailable.

PROBLEMS IN DESIGNING ANN-BASED FORECASTING SYSTEMS

- ANN design: Number of hidden neurons

- There is little theoretical basis for the decision and very few successful heuristics:
    - ⇨ If there are too few neurons, the model will not be able to model the data well.
    - ⇨ If there are too many, the model will overfit the data.

- In most cases, the selection is performed by trial and error.

- In very few cases using evolutionary techniques.

PROBLEMS IN DESIGNING ANN-BASED FORECASTING SYSTEMS

- ANN design: Overfitting and Overparameterization

- Overfitting means estimating a model that fits the data so well that it ends including the noise and then produces poor forecasting. This comes about because the model was over-trained or it was too complex.
- To avoid overtraining:
  - ➪ Cross-validation: the sample set is split into a training set and a validation set.
  - ➪ Regularization techniques: modifying the cost function to be minimized by adding a term that penalizes for the complexity of the model

Dpto. Tecnología Electrónica
Universidad de Málaga

PROBLEMS IN DESIGNING ANN-BASED FORECASTING SYSTEMS

- ANN design: Overfitting and Overparameterization

  - Overparameterization: it produces overfitting due to the complexity of the model.
  - The user adds to the ANN a large number of variables and neurons, without taken into account the number of parameters to be estimated.
  - The adequate rate between the number of sample points required for training and the number of weights in the network has not yet clearly identified.
  - It is difficult to establish how many parameters are too many for a given sample size.

CONTENTS

Dpto. Tecnología Electrónica
Universidad de Málaga

PROBLEMS IN DESIGNING ANN-BASED FORECASTING SYSTEMS

- ANN implementation: Once the ANN has been designed, it must be trained to estimate its parameters.

- To select the training algorithm: backpropagation, quick-propagation, Levenberg-Marquard, ...
- When training must stop?
  - ➪ After a fixed number of iterations
  - ➪ After the error reaches some specific value

- The above criteria are not adequate since they may lead to overfitting of the model.
- Training samples must be appropriately selected.

PROBLEMS IN DESIGNING ANN-BASED FORECASTING SYSTEMS

- ANN implementation: Evaluating the effectiveness of implementation.

- Was the ANN properly trained and tested, so that its performance was the best it could achieve?
  - The ANN was well fitted to the data: the errors in the training sample must be reported.
  - The ANN performances in the training and in the test samples were comparable.
  - The ANN performances across different test samples were coherent

CONTENTS

Dpto. Tecnología Electrónica
Universidad de Málaga

PROBLEMS IN DESIGNING ANN-BASED FORECASTING SYSTEMS

- ANN validation: To examine the ANN errors with samples other than those used for parameter estimation (out of sample errors, as opposed to insample errors).
- Was the performance method fairly compared to that of some well-accepted method? Four requirements are needed:

a) its performance should be compared to that of well accepted methods

  ⇨ Compared to some “naïve” method, which provides a benchmark.

  ⇨ Compared to that of a good standard method (?): ARMAX, regression model, other ANN, fuzzy engines, ..

PROBLEMS IN DESIGNING ANN-BASED FORECASTING SYSTEMS

- ANN validation:

b) the comparison must be based on the performance on test samples.
c) the size of the test samples must be adequate, so that some inference might be drawn.
d) to examine the error using standard techniques:

- ➪ Mean Absolute Percent Errors (MAPE). Relationship between MAPE and cost function.
- ➪ Standard deviation (STD) of the errors.
- ➪ Error distribution.

CONTENTS

A SPECIFIC PROPOSAL

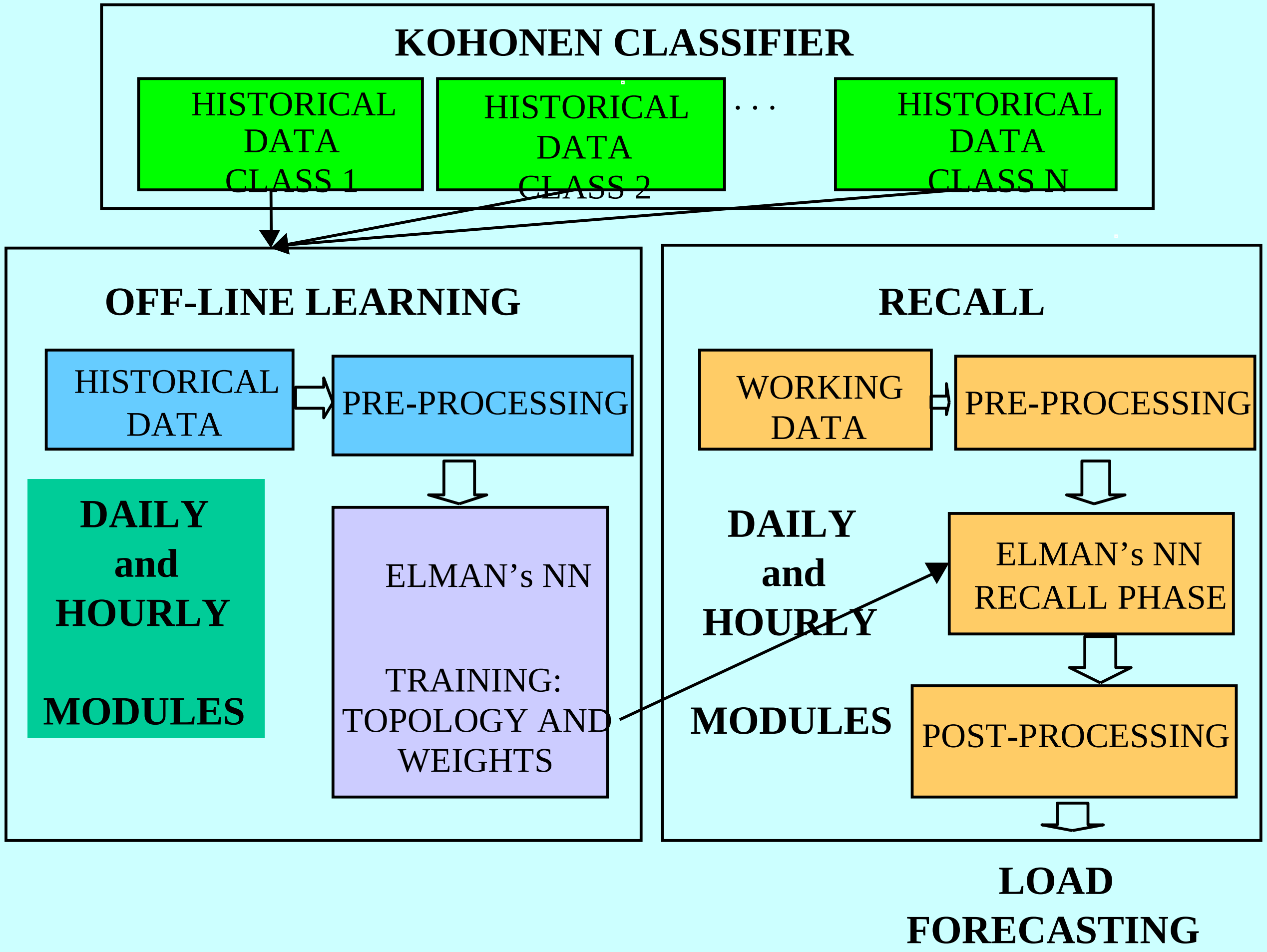
- The forecasting process is developed in three phases:

  1) Using historical data, days are classified according to their load profile (Kohonen's SOM).

  2) For each class, an ANN is built and trained.

  3) Recall phase, where prediction is carried out.

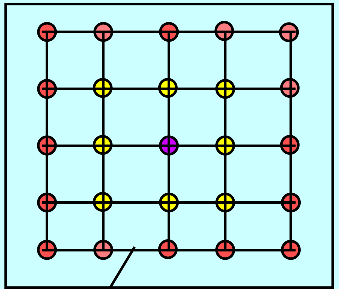
A SPECIFIC PROPOSAL



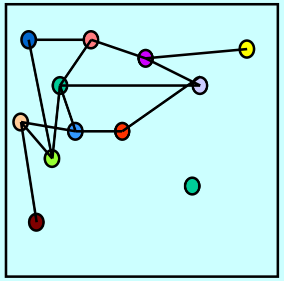
A SPECIFIC PROPOSAL

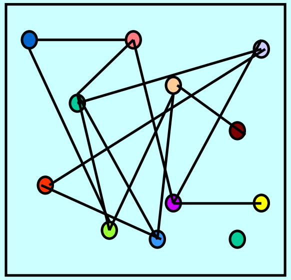
- Kohonen's algorithm (SOM)

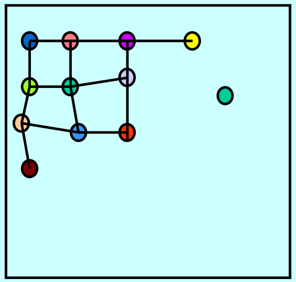
Topologie



$$w_j = (w_{j1}, w_{j2}, \ldots, w_{jn})$$







L'algorithme

- Choix des valeurs initiales de a et V
- Initialisation aléatorire des poids .
- Normalisation
- Répéter jusqu'à convergence
    - Présentation d'un patron d'entrée

$$x_k = (x_{k1}, x_{k2}, \ldots, x_{kn})$$

    - Détermination du neurone le plus actif

$$max_j(|w_j \cdot x_k|)$$

    - Mise à jour des poids dans le voisinage V du neurone le plus actif

$$w_j(t+1) = w_j(t) + a(x_k - w_j)$$

    - Normalisation de *w*
    - Mise à jour de a y V

- -

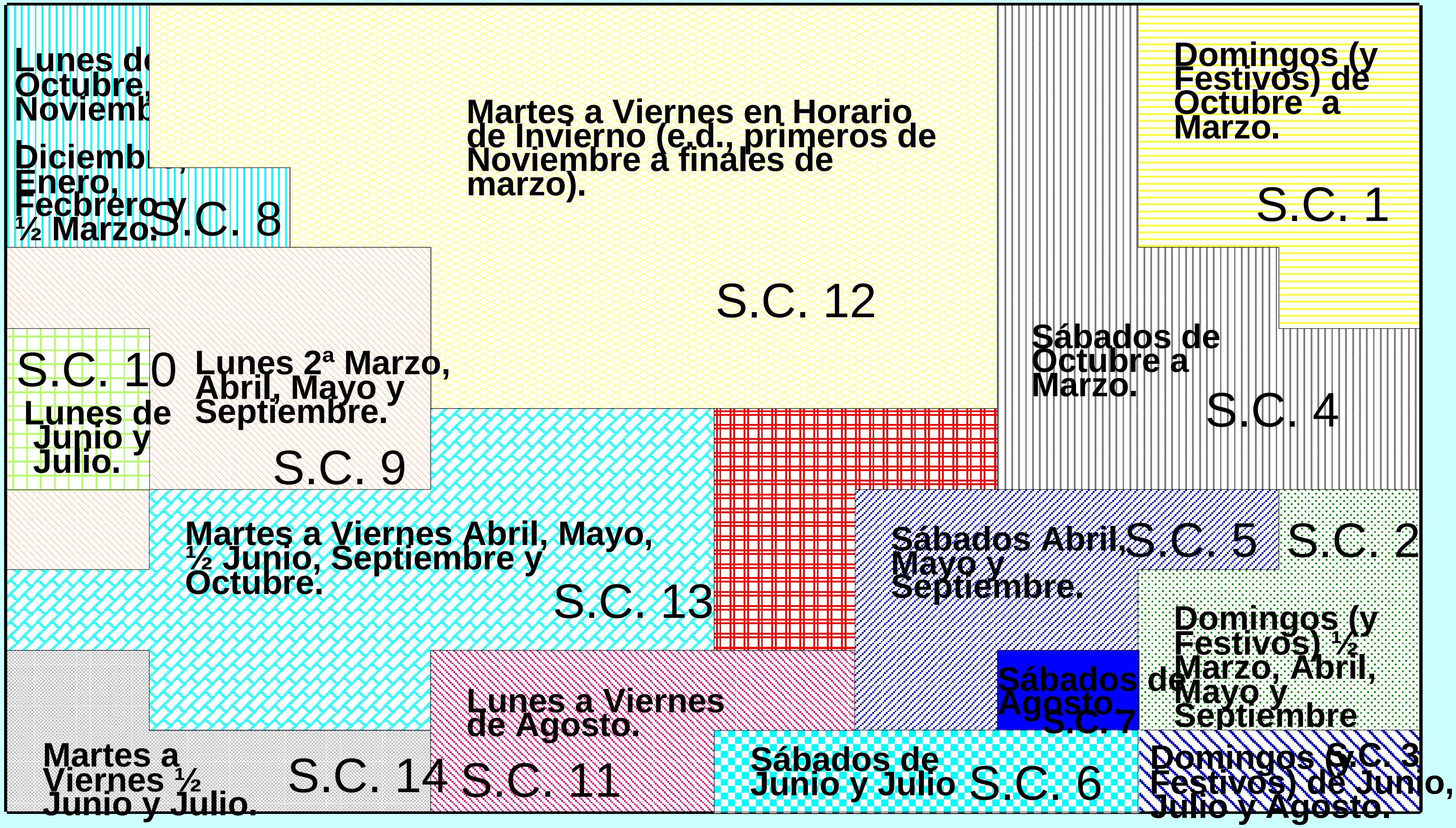
A SPECIFIC PROPOSAL

- Clasification

Class 1	Sunday from October to March
Class 2	Sundays from April, May and September
Class 3	Sundays in June, July and August
Class 4	Saturdays from October to March
Class 5	Saturdays in April, May and September
Class 6	Saturdays in June and July
Class 7	Saturdays in August
Class 8	Mondays from October to March
Class 9	Mondays in April, May and September
Class 10	Mondays in June, July and August
Class 11	Tuesday to Friday of August
Class 12	Tuesday to Friday from November to March
Class 13	Tuesday to Friday of April, May, first fortnight of June, September and October
Class 14	Tuesday to Friday of second fortnight of June and July
Class 15 (Special Class)	Easter Week

A SPECIFIC PROPOSAL

- Two dimensional view:



- Easter class does not appear

Dpto. Tecnología Electrónica
Universidad de Málaga

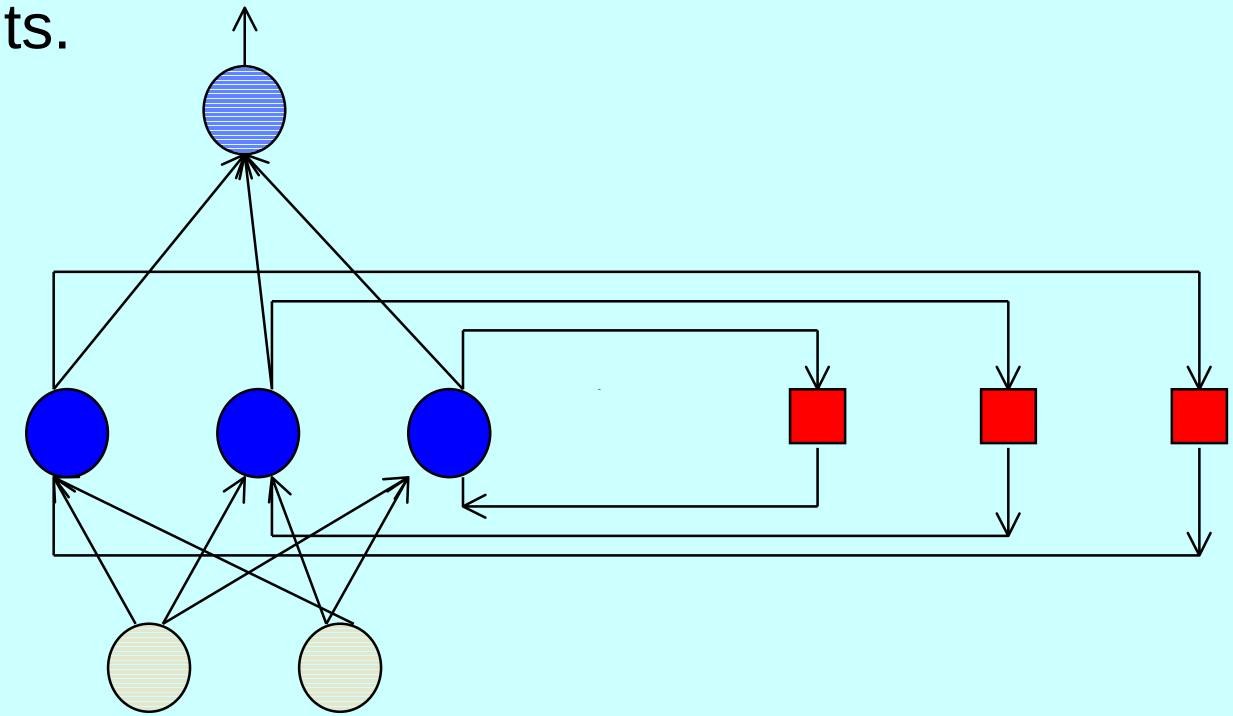
A SPECIFIC PROPOSAL

- Learning phase: a different ANN for each class.
- Input data: the hourly demand, the integral demand and the daily temperatures at 6:00 and 18:00, from January 1989 till February 1999.
- Training: 1989-1995; cross-validation: 1996; testing: 1997-February 1999.

- Correlation between series to predict and other variables is performed.
- As inputs to the ANN, we use the previous daily load of its class (24 inputs) and the predicted integral demand. Only analog inputs are used. The output is the predicted load profile (24 outputs).

A SPECIFIC PROPOSAL

- Elman recurrent neural networks has been used: feedforward nets with their hidden layer recycled back as inputs.



- Selection of the number of hidden neurons by experimentation.

A SPECIFIC PROPOSAL

- Normalization of the input data:

$$Y_S(h) = \frac{I_{MAX} - I_{MIN}}{Y_{MAX} - Y_{MIN}} Y_U(h) + \frac{Y_{MAX} I_{MIN} - Y_{MIN} I_{MAX}}{Y_{MAX} - Y_{MIN}}$$

- Re-training: demand patterns are changing due to long term trends (life-span). In our system the re-learning process is periodic (every year), automatically performed and completely transparent to the user.

A SPECIFIC PROPOSAL

- Results: MAPE and STD

$$MAPE = \frac{1}{N} \sum_{h=1}^{N} \frac{\left| P_F(h) - P_A(h) \right|}{P_A(h)}$$

A SPECIFIC PROPOSAL

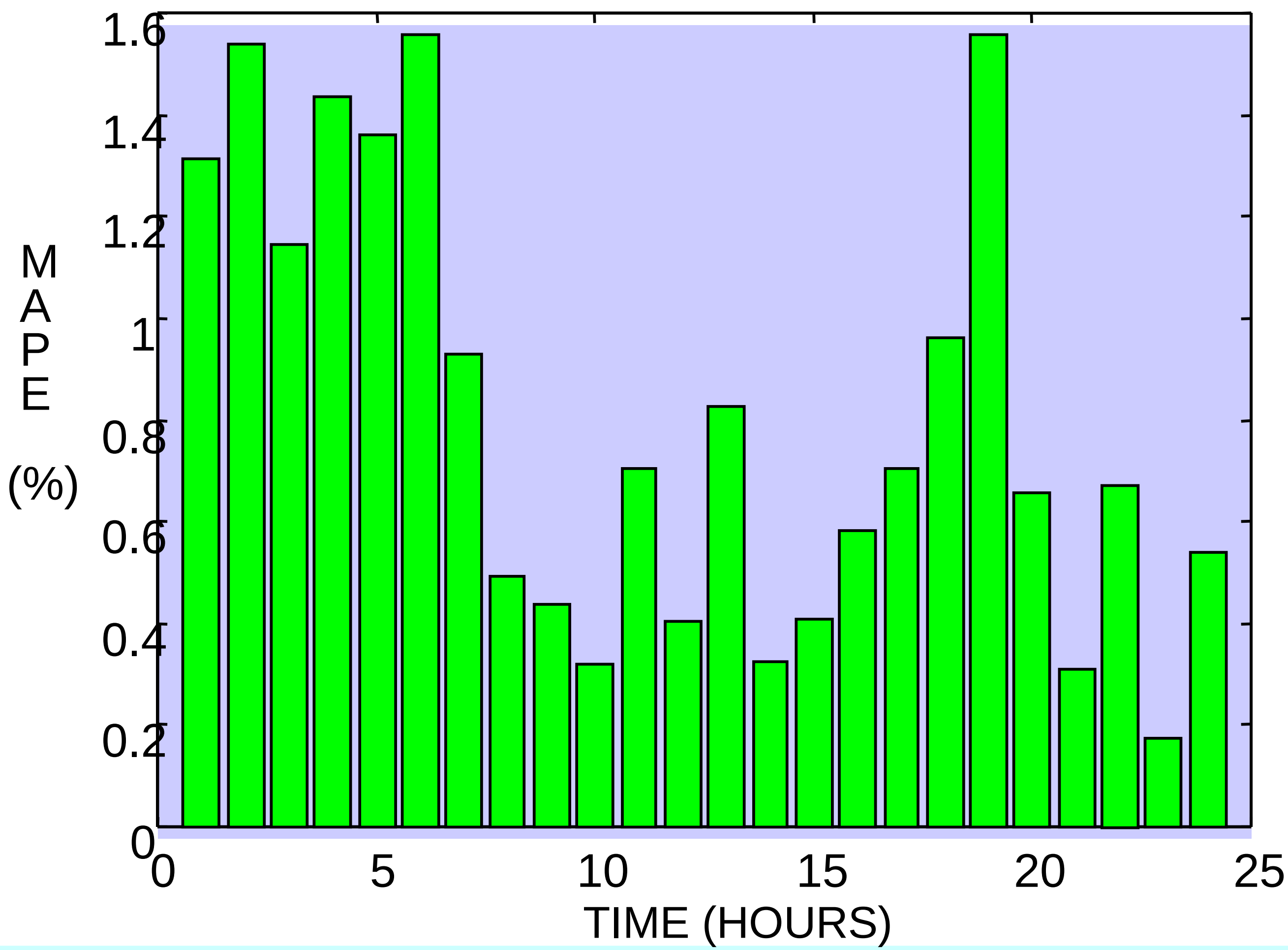
	TRAINING (1989-1996)		TESTING					
			1997		1998		1999 (February, 17th)	
	MAPE(%)	STD	MAPE(%)	STD	MAPE(%)	STD	MAPE(%)	STD
Class 1	1.39	0.76	1.55	1.33	1.74	1.42	1.42	0.91
Class 2	1.65	1.03	1.74	1.10	1.87	1.22		
Class 3	1.20	0.69	1.26	0.87	1.54	1.02		
Class 4	1.37	1.06	1.52	1.19	1.66	1.28	1.61	1.23
Class 5	1.32	0.94	1.43	1.12	1.39	0.97		
Class 6	1.24	0.78	1.36	0.84	1.40	0.89		
Class 7	1.09	1.02	1.32	0.68	0.64	0.42		
Class 8	1.71	1.31	1.70	1.38	1.90	1.53	1.17	0.93
Class 9	1.50	1.11	1.62	1.23	1.63	1.11		
Class 10	1.30	0.90	1.38	1.26	1.34	0.99		
Class 11	1.49	1.07	1.52	1.19	1.55	1.15		
Class 12	1.29	0.96	1.34	0.97	1.30	0.98	1.15	0.84
Class 13	1.40	1.05	1.47	1.14	1.67	1.20		
Class 14	1.03	0.75	1.30	0.94	1.48	1.09		
Class 15 Easter Week	1.70	1.14	1.82	1.27	1.91	1.31		

A SPECIFIC PROPOSAL

MAPE for class 12, February 1998

MAPE=0.8%

STD=0.62



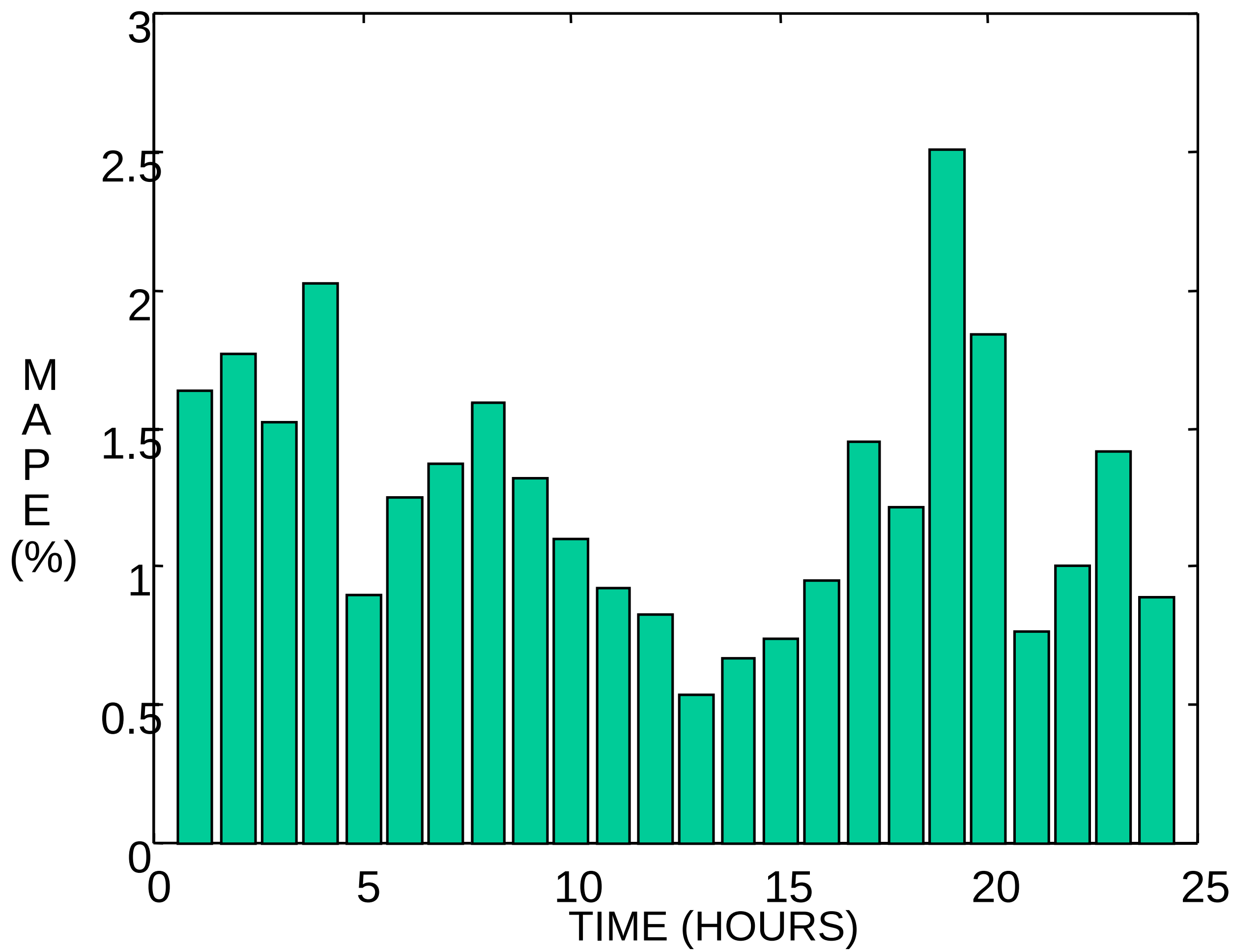
TE
Dpto. Tecnología Electrónica
Universidad de Málaga

A SPECIFIC PROPOSAL

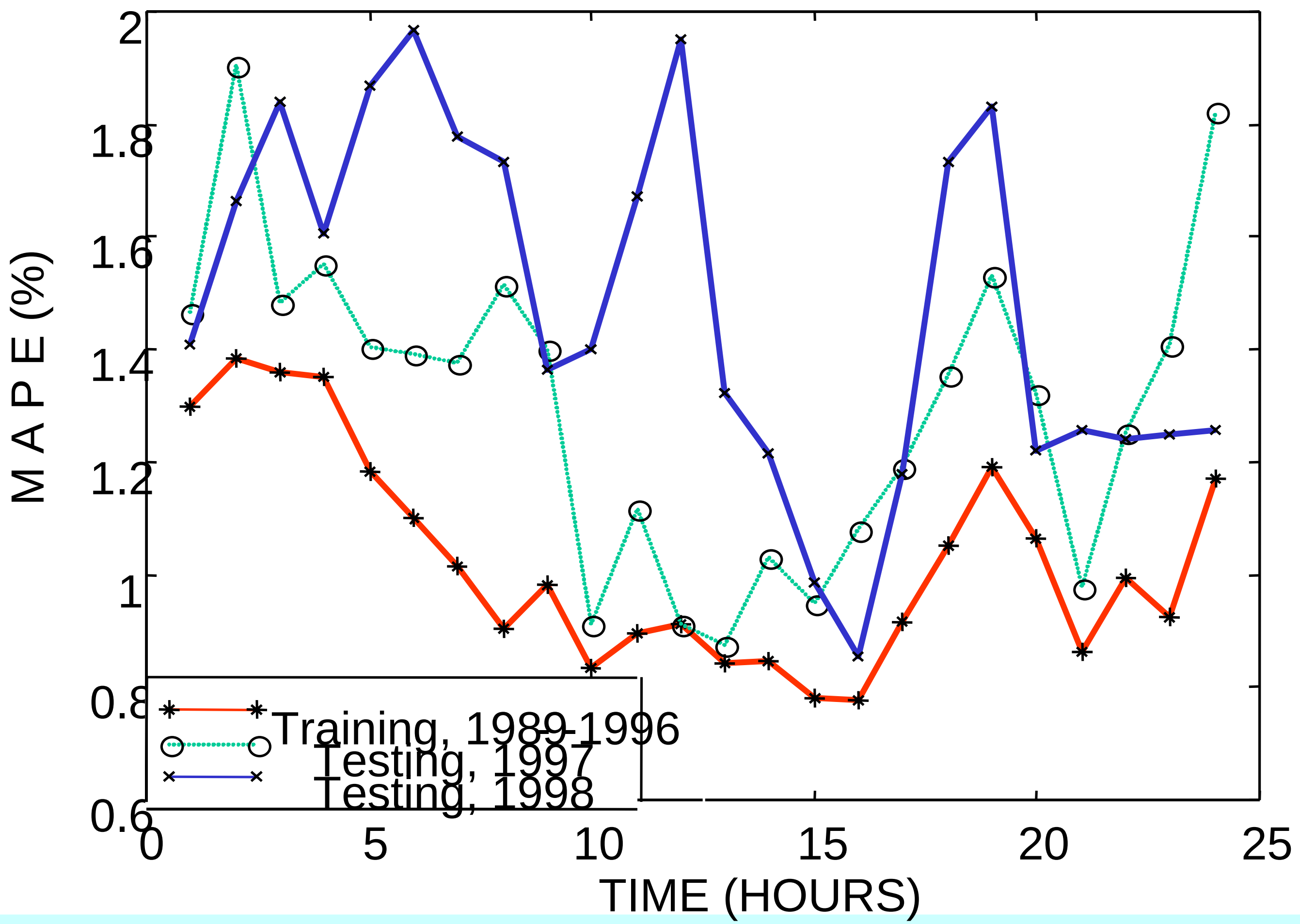
MAPE for 2nd fortnight October 1998, class 13

MAPE=1.20%

STD= 0.86



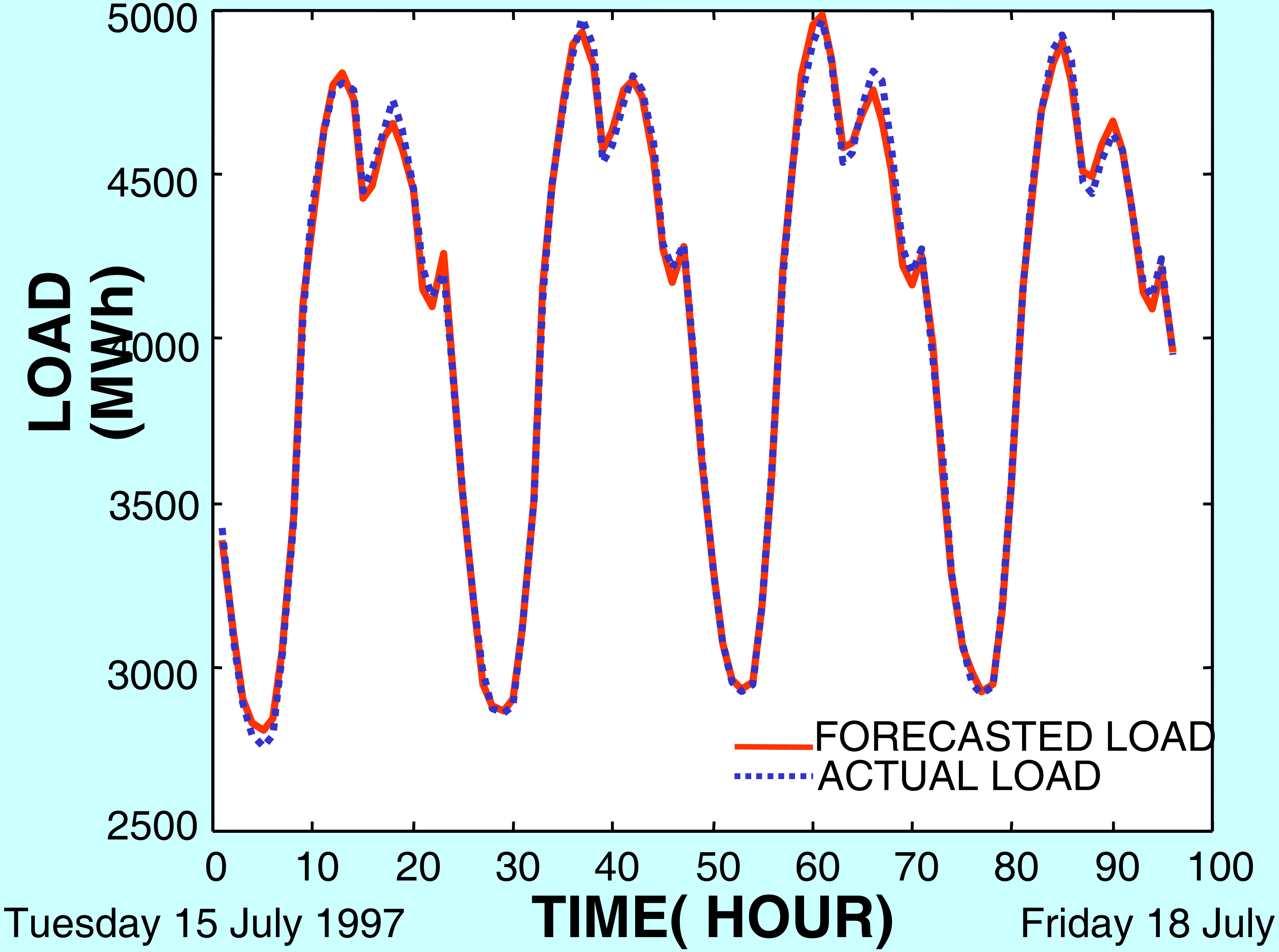
A SPECIFIC PROPOSAL



Training and testing error for class 14

TE
Dpto. Tecnología Electrónica
Universidad de Málaga

A SPECIFIC PROPOSAL



MAPE=0.61%

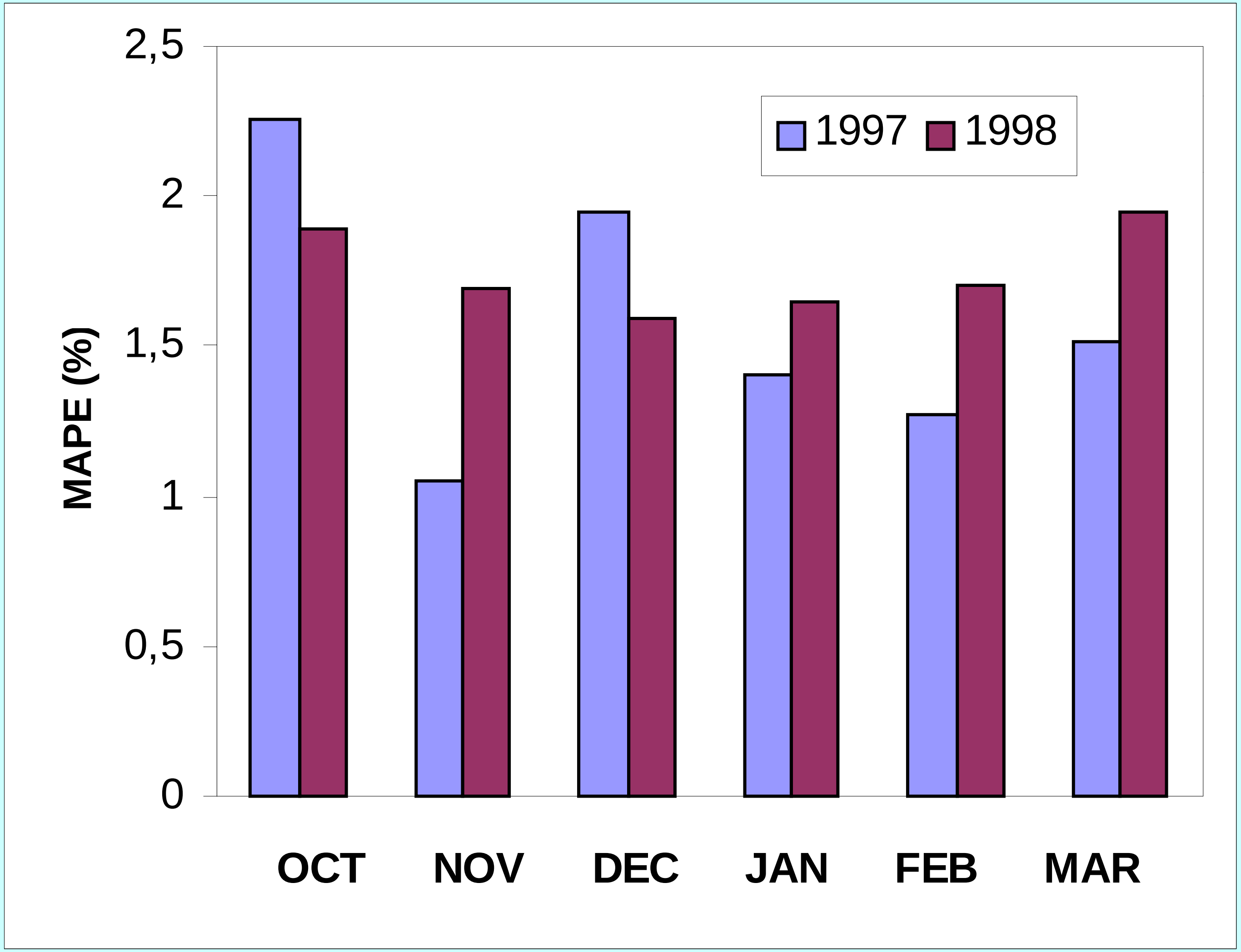
Tuesday 15 July 1997

Friday 18 July 1997

TE
Dpto. Tecnología Electrónica
Universidad de Málaga

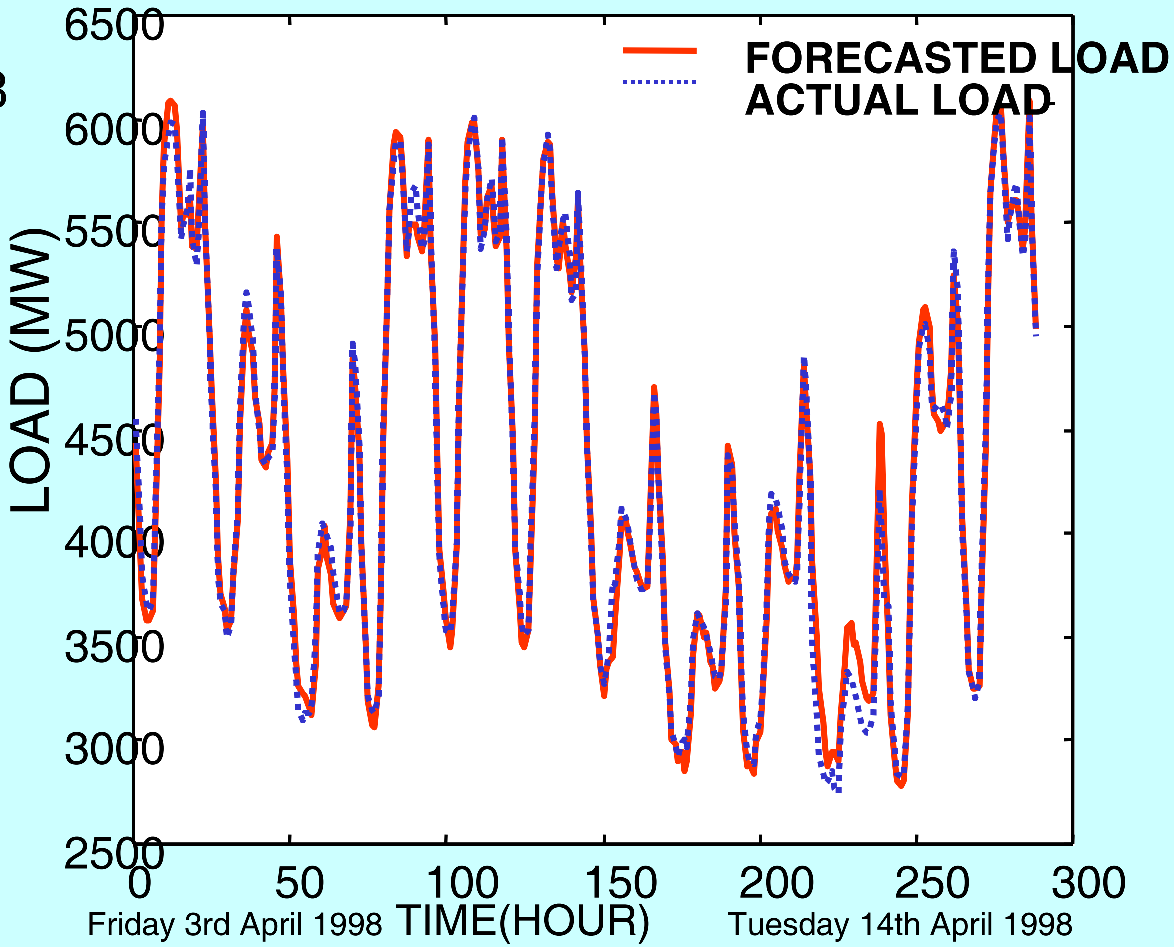
A SPECIFIC PROPOSAL

Comparisons of load forecasting errors for months of class 1, years 1997 and 1998



A SPECIFIC PROPOSAL

Easter week 1998



Dpto. Tecnología Electrónica
Universidad de Málaga

CONTENTS

Dpto. Tecnología Electrónica
Universidad de Málaga

CONCLUSIONS

- Input pattern should be as homogeneous as possible.

- ANN design must avoid overfitting and overparameterization.

- Error of ANNs shoul be performed using standard techniques

- More research on the behavior of large ANN is needed.

- More rigorous standards should be adopted in the reporting of experiments.

SHORT-TERM LOAD FORECASTING USING ARTIFICIAL NEURAL NETWORKS

Francisco Sandoval

Departamento de Tecnología Electrónica
Universidad de Málaga, Espagne