This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Eseye, Abinet Tesfaye; Lehtonen, Matti; Tukia, Toni; Uimonen, Semen; Millar, Robert
Efficient Feature Selection Strategy for Accurate Electricity Demand Forecasting

Published in:
Proceedings of the IEEE PES Europe Conference on Innovative Smart Grid Technologies, ISGT-Europe 2019

DOI:
[10.1109/ISGTEurope.2019.8905713](https://doi.org/10.1109/ISGTEurope.2019.8905713)

Published: 01/01/2019

Document Version
Peer reviewed version

Please cite the original version:
Eseye, A. T., Lehtonen, M., Tukia, T., Uimonen, S., & Millar, R. (2019). Efficient Feature Selection Strategy for Accurate Electricity Demand Forecasting. In *Proceedings of the IEEE PES Europe Conference on Innovative Smart Grid Technologies, ISGT-Europe 2019* (IEEE PES Innovative Smart Grid Technologies Conference Europe). IEEE. <https://doi.org/10.1109/ISGTEurope.2019.8905713>

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

© 2019 IEEE. This is the author's version of an article that has been published by IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Efficient Feature Selection Strategy for Accurate Electricity Demand Forecasting

Abinet Tesfaye Eseye, Matti Lehtonen, Toni Tukia, Semen Uimonen, and John Millar

Department of Electrical Engineering and Automation

Aalto University

Espoo, Finland

Email: abinet.eseye, matti.lehtonen, toni.tukia, semen.uimonen, john.millar@aalto.fi

Abstract—Accurate electric load prediction offers an important input information for various smart decisions in energy systems. This paper focuses on short-term forecasting of the electric load of small-scale decentralized energy systems (buildings, energy communities, microgrids, virtual power plants, local energy internets, etc.), which are newly evolving energy system models. Quite few researchers have done feature selection before training forecast models, which is an important preprocessing task of data mining and broadly practiced for knowledge exploration in expert and intelligent systems. This paper proposes a feature selection strategy to find the most significant and non-repetitive variables for accurate short-term electric load prediction in distributed energy systems in general and buildings in particular. In the devised strategy, Binary Genetic Algorithm (BGA) is employed for the variable selection task and Gaussian Process Regression (GPR) is applied for quantifying the fitness score of the variables. The proposed feature selection strategy is implemented and validated using actual electricity consumption data of various buildings located in Otaniemi area of Espoo, Finland. The results are compared with those obtained by other predictor selection methods and show outperformed performances.

Keywords—BGA, electricity demand forecasting, feature selection, fitness evaluation measure, GPR, smart grid.

I. INTRODUCTION

Numerous challenges, ranging from inadequate electricity supply to increasing demand, confronts distributed energy systems. The electricity consumption curves of decentralized energy systems are dissimilar from the typical electricity consumption curves that denote big-scale demands at the global, national or regional levels. This lets the traditional techniques unsuitable for their direct application in distributed energy systems since the demand is much smaller and highly volatile. This indicates the recent development of distributed controls for local energy systems needs more effective feature selection (FS) methods and prediction techniques for optimal and efficient usage of the electricity.

FS is a process of finding a cluster of the most significant variables for a forecasting or classification task. It is essential to lower computing time and storage memory size, increase generalization and interpretability, simply complexity, and reduce overfitting.

The central logic when developing a FS tool is that the initial dataset contains some features that are either repetitive or unimportant and can thus be removed without causing significant destruction of information. A number of researches have shown that repetitive and unimportant variables decrease the performance and generalization competence of forecast systems. That is why, recently, FS researches are becoming quite prominent.

The focus of this study is to devise and develop a feature selection tool for accurate electric load prediction in distributed energy systems in general and buildings in particular.

Forecasting accuracy is the central goal of almost all prediction researches. As thoroughly shown in [1] and [2], the accuracy of forecasting approaches relies on the feature scope that is formed via the initial feature sets and FS method. FS is typically used in machine learning (ML) applications as one of the preprocessing tasks, where a feature subset is established by eradicating variables with inferior or insignificant value and highly repetitive [3]. Nevertheless, only quite few prediction approaches have performed FS ahead of fitting forecasting models.

Numerous metaheuristic optimization techniques have been employed as search methods for FS. For example, Particle Swarm Optimization (PSO) [4], Ant Colony Optimization (ACO) [5] and Genetic Algorithm (GA) [6]. GA has been widely used due to its higher suitability and effective searching capability. It belongs to the class of the artificial intelligent (AI) searching algorithms and has been successfully applied for solving numerous optimization problems [7].

We propose a hybrid ML-based FS strategy using the combination of Binary Genetic Algorithm (BGA) and Gaussian Process Regression (GPR) for accurate short-term prediction of electricity demand in buildings. BGA is a type of GA that functions by first encoding the feature space (candidate solutions) in binary bitstrings. This makes the BGA even more suitable for FS tasks than the traditional GA. In the proposed hybrid BGA-GPR FS strategy, the BGA is used to select the most significant and non-repetitive variables, while the GPR is employed to measure the fitness of the variables for the BGA execution.

This study mainly aims to (1) reveal the importance of a robust FS for accurate electricity demand forecasting, (2) present efficient and effective ML-based hybrid FS method for electricity demand prediction, and (3) enhance forecasting accuracy by making use of FS ahead of training forecasting models.

The rest parts of the study are outlined below. Section II presents the relevant prior works on FS. Section III provides the data and states the FS problem. Sections IV and V describe the working principle and mathematical modeling the BGA and GPR, respectively. Section VI discusses the devised BGA-GPR-based FS strategy. The experimental results and validations are provided in Section VII. Section VIII concludes the paper.

II. PRIOR WORKS

FS techniques can be categorized as filter, wrapper and embedded methods [1]. Filter methods do not rely on any forecast model and they order variables based on statistical behavior. They use a correlation value to grade a predictor subset. The Filter FS

method comprises correlation-based [8], mutual information-based [9], and principal component analysis-based methods [10]. Wrapper methods assess variable subsets depending on their value to a particular predictor or classifier. They assume the FS as a searching problem that handles numerous combinations of variables, measured, and compared with other combinations. Compared to the filter methods, wrapper methods show better performances because several predictor subsets are measured by the forecast model at every step [11]. Embedded techniques combine the variable selection task into the forecast model learning task. For example, the regularization technique [1] is one type of an embedded FS.

Table I provides recent FS works on demand forecasting in the power and energy sectors.

TABLE I. PRIOR FS WORKS

Application	FS Type	References
Power systems	Filter	[8], [12], [13]
	Wrapper	[14], [15]
Energy systems	Filter	[16]
	Wrapper	[17]

Reference [18] devised a hybrid of GA and ACO for feature selection in the electric load prediction problem. Reference [19] proposed the combination of GA and support vector regression (SVR) for feature selection to forecast hotel-room booking. Reference [20] applied a GA-based variable selection to forecast the demand of Chinese retail industries. Reference [21] devised an improved GA for FS in the demand forecasting in the health sector.

In the FS process, the variables are sorted and chosen based on their evaluated values of the fitness function. The subset of variables that gives the best value of the fitness function is selected. This paper employs the residual (error) of GPR model as the fitness function of the BGA.

As the literature survey in this study shows, the traditional GA with the standard framework is employed for FS by most researches [15]. The traditional GA functions with the real-valued variables to reduce the fitness function. This decreases the efficiency of the FS and causes computation burden. This problem is addressed in this paper by substituting the traditional GA by the BGA and combining it with the effective fitness evaluation measure (GPR residual).

III. FEATURE SPACE AND FS PROBLEM

The candidate variables of the original feature space for this FS work includes historical electric load, seasonal parameters, weather parameters, occupancy, and economic factor (electricity price). The variables f_i , $i = 1, 2, \dots, 24$, in Table II represent the candidate variables.

Thus, the feature space is a matrix of order $m \times n$, where $m = 8760$ is the number of observations, which is a one-year (2017) hourly sample of the features and $n = 24$ is the number of candidate features.

TABLE II. FEATURE SPACE OF THE FS PROBLEM

Feature Index	Feature (f_i)	Unit/Scale	Data Category
1	Hour of the day	1-24	Seasonality/calendar
2	Day of the week	1-7	
3	Month of the year	1-12	
4	Season of the year	1-4	
5	Period of the day	1-3	Occupancy (number of people)
6	Holiday/Weekend indicator1	0-1	
7	Holiday/weekend indicator2	0-2	Weather
8	Ambient air temp.	°C	
9	Dew-point temp.	°C	
10	Relative humidity	%	
11	Precipitation	mm/h	

12	Snow depth	cm	
13	Ambient air pressure	hPa	
14	Horizontal visibility	m	
15	Wind direction	deg	
16	Wind speed	m/s	
17	Gust speed	m/s	
18	Cloud cover	0-8	
19	Global solar radiation	Watt/m ²	
20	Sunshine duration	s	Economy
21	Electricity price	¢/kWh	
22	Previous 24h average load	kW	
23	24h lagged load	kW	
24	168h lagged load	kW	Electric load

In this study, the optimization problem formulated in (1) is solved in order to find the most significant and non-repetitive feature subset from the set of the candidate variables (feature space) given in Table II.

FS Problem:

Given that:

$$f_r \in \mathbb{Z}^+, 1 \leq f_r \leq 24, \text{ and } \beta \in \mathbb{R}^+ \ni 0 \leq \beta \leq 100 \quad (1)$$

where, f_r is the number of features in the reduced feature subset and β is the forecasting error (in %). Find a variable subset of f_i from Table II such that the objectives β and f_r are minimized.

IV. BINARY GENETIC ALGORITHM (BGA)

GA is a population-based metaheuristic optimization technique that was motivated by the Charles Darwin theory of human evolution and genetics theory [22]. The GA operates on the chromosomes (candidate solutions) to create a new population (offsprings) through its three genetic operators - selection, crossover and mutation. The fitness of the chromosomes is evaluated using a fitness function. The fitness function gives numerical values that is used for ranking the chromosomes. BGA is a type of GA that operates by first encoding the chromosomes as bitstrings to minimize or maximize the objective function. It is more efficient and stable than the standard real-valued GA. It also reduces computation burden and time. The flowchart of the BGA is shown in Figure 1.

Figure 1. Flowchart of BGA

V. GAUSSIAN PROCESS REGRESSION (GPR)

GPR is a mathematical model to fit nonlinear relationships between variables using probabilistic distributions over functions. A Gaussian process (GP) describes distributions over functions in such a way that, if we pick up any two or more points in the functions, samples of the targets at these points trace a joint (multivariate) Gaussian distribution (GD). That is, a GP is expressed as a gathering

of arbitrary variables, any specified amount of which have a joint GD.

The GPR output y of a function f at input x is given by:

$$y = f(x) + \varepsilon \quad (2)$$

where, $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ is normal distribution with zero mean and σ standard deviation.

The GP distribution of the function $f(x)$ is defined as:

$$f(x) = GP(m(x), k(x, x')) \quad (3)$$

As shown in (3), the GP distribution is expressed by the mean (m) and covariance (k) functions.

The mean function $m(x)$ designates the estimated value of the function at x :

$$m(x) = E[f(x)] \quad (4)$$

The covariance function $k(x, x')$ denotes the relation between the values of the function for different inputs x and x' :

$$k(x, x') = E[(f(x) - m(x))(f(x') - m(x')))] \quad (5)$$

The function k is called the GP kernel [23]. The GPR model in this study uses the squared exponential kernel expressed beneath:

$$k(x, x') = \sigma_f^2 \exp\left(-\frac{1}{2} \sum_{r=1}^d \frac{(x_r - x'_r)^2}{\sigma_r^2}\right) \quad (6)$$

where, σ_r is the scale-length of the predictor r , $r = 1, 2, \dots, d$ and σ_f is the standard deviation of the dataset. The parameter θ is given by:

$$\theta_r = \log \sigma_r, \text{ for } r = 1, 2, \dots, d \quad (7)$$

$$\theta_{d+1} = \log \sigma_f \quad (8)$$

The parameters σ_r and σ_f are controlled to increase or decrease the correlation between the points and therefore the distribution of the function. After the mean and kernel functions are obtained, the GP is applied to obtain the prior and future values of the function using the historical observation of the dataset.

VI. PROPOSED FEATURE SELECTION STRATEGY

As described in the above sections, the proposed FS methodology for electricity demand prediction is based on the combination of the BGA and GPR (BGA-GPR-based FS). Figure 2 shows the flowchart for the detail operating mechanism of the proposed BGA-GPR-based FS. There are five crucial sub-tasks in the BGA: chromosome encoding, fitness evaluation, selection technique, genetic operators, and termination condition. An initial population is created and evaluated employing the fitness function. A gene value of '1' designates the particular variable pointed by the position of the '1' is selected for the fitness assessment. While a value of '0' indicates the specific variable is not selected.

The chromosomes are sorted based on their fitness values. The top n fittest offsprings (Elitism of size n) are selected to continue with the following generation. The remaining offsprings in the population genetically move via the crossover and mutation operators to generate crossover and mutation offsprings, respectively. The selection, crossover and mutation offsprings are then form the new population (generation) [24], [25].

In this paper, the fitness of the chromosomes is evaluated using the GPR model. The BGA fitness function is formulated by the mean squared error (MSE) of the GPR model predictive residuals. That is, the MSE of the actual target and the GPR model output is calculated for every variable subset given in Table II. The fitness function is expressed as follows:

$$fit = \frac{1}{n} \sum_{i=1}^n (T_i - y_i)^2 \quad (9)$$

where T is a vector of target values (electric load) and n is the number of samples.

The goal of the BGA is to minimize the fitness function (9) by selecting a feature subset with the best fitness (lowest MSE) over iterations.

Figure 2. Flowchart of BGA-GPR-based FS

The BGA parameters employed in this study are given in Table III.

TABLE III. PARAMETERS OF BGA

Parameter	Value
Population size	24
Genomelength	24
Population type	Bitstring
Fitness function	GPR residual
Number of generations	100
Stall Generation Limit	50
Selection mechanism	Tournament selection
Tournament size	2
Mutation function	Uniform mutation
Mutation rate	0.1
Crossover function	arithmetic crossover (logical XOR)
Crossover fraction	0.8
Elite count	2

The BGA stops when it reaches at a desired optimal point. The optimal point associates with the desired predictor subset in question. Following the BGA convergence, the chromosome that achieved the best fitness value is selected and decoded to obtain the desired feature subset as illustrated in Figure 3.

Figure 3. Final feature subset decoding

VII. RESULTS AND DISCUSSIONS

In this study, the BGA-GPR-based FS strategy is implemented and validated based on four electric load datasets (actual measurement records) collected from four different buildings in the Otaniemi area of Espoo, Finland. The buildings are Building A (residential building), Building B (educational building), Building C (office building) and Building D (mixed use building). The buildings have a peak (in 2017) total electric load of 236kW, 626kW, 21kW and 78kW, respectively. A one-year (2017) hourly observation, 8760 values, of both the candidate variables in the predictor space and target variable are employed for FS task. The FS is executed for each of the buildings. The experimental results obtained are given in Table IV.

TABLE IV. FS RESULTS

Customer (Dataset) Type	Selected Features	Best Fitness	Without FS
Building A - Residential	1, 2, 3, 4, 5, 7, 8, 9, 10, 13, 15, 16, 17, 19, 20, 21, 22, 23, 24	41.21	72.16
Building B - Educational	1, 2, 3, 4, 5, 7, 8, 9, 10, 13, 15, 17, 19, 20, 21, 23, 24	9.48	268.50
Building C - Office	1, 2, 3, 4, 5, 7, 8, 13, 17, 19, 20, 21, 22, 23, 24	0.38	11.44
Building D - Mixed-use	1, 2, 3, 4, 5, 7, 8, 9, 13, 16, 17, 21, 22, 23, 24	0.12	71.87
All (Union)	1, 2, 3, 4, 5, 7, 8, 9, 10, 13, 15, 16, 17, 19, 20, 21, 22, 23, 24		

As shown in Table IV, the number of predictors selected by the devised FS strategy is much smaller than the size of the original predictor space. Thus, we can say this is due to the existence of insignificant and repetitive information by most of the features in the original predictor space. The BGA at the end chooses the variable subset that holds the most significant and non-repetitive features. For the residential building, the improvement in fitness value (MSE) using the BGA-GPR FS chosen predictors to fit the electric load by the GPR model is 42.9% over the initial feature set (without FS). Likewise, the chosen variable subsets for the educational, office and mixed-use buildings have achieved improvements of 96.5%, 96.7% and 99% respectively over the initial to variable set, with respect to the GPR residual (MSE) performance measure.

Features 1, 2, 3, 4, 5, 7, 8, 9, 10, 13, 15, 16, 17, 19, 20, 21, 22, 23, and 24, which designate the hour of the day, day of the week, month of the year, season of the year, period of the day, holiday/weekend indicator, air temperature, dew-point, relative humidity, air pressure, wind direction, wind speed, gust speed, solar radiation, sunshine duration, electricity price, previous 24h average electric load, 24h lagged electric load, and 168h lagged electric load, respectively, are selected at least for one of the building types. For the sake of consistency, the feature subset consisting of these 19 predictors is

selected to establish the training input for the accurate short-term prediction of the electricity demands of the buildings.

In order to verify the performance of the proposed BGA-GPR-based FS further, the obtained FS results are compared with FS results using other two conventional FS methods, namely: Correlation-based FS (C FS) and Neighborhood Component Analysis Regression-based FS (NCA FS). The Correlation-based FS first computes the Pearson and Spearman correlations of each variable with the target, and it then takes the maximum of the two correlation coefficients. A variable with correlation value greater than a specified threshold (0.5 in this study) is chosen as the significant variable and involved in the final predictor subset. The NCA FS works based on the neighborhood component analysis (NCA) regression modeled over the variable subsets versus the target. The NCA FS finds the variable weights utilizing a diagonal adaptation of the NCA regression model. The model realizes the FS by regularizing the weights of the variables.

Table V gives the comparison of the FS results by the evaluated FS techniques.

TABLE V. COMPARISON OF FS RESULTS

Customer (Dataset) Type	Selected Features			Best Fitness		
	BGA-GPR FS	NCA FS	C FS	BGA-GPR FS	NCA FS	C FS
Building A	1, 2, 3, 4, 5, 7, 8, 9, 10, 13, 15, 16, 17, 19, 20, 21, 22, 23, 24	1, 2, 3, 4, 6, 7, 8, 9, 10, 13, 15, 16, 17, 19, 20, 21, 22, 23, 24	1, 3, 4, 5, 21, 22, 23, 24	41.21	50.47	54.96
Building B	1, 2, 3, 4, 5, 7, 8, 9, 10, 13, 15, 17, 19, 20, 21, 23, 24	1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 13, 15, 17, 19, 20, 21, 23, 24	2, 3, 4, 5, 6, 7, 10, 13, 19, 20, 21, 22, 23, 24	9.48	14.35	18.67
Building C	1, 2, 3, 4, 5, 7, 8, 13, 17, 19, 20, 21, 22, 23, 24	1, 2, 3, 4, 7, 8, 13, 17, 19, 20, 21, 22, 23, 24	1, 2, 3, 4, 5, 6, 7, 10, 13, 19, 20, 21, 22, 23, 24	0.38	0.50	0.57
Building D	1, 2, 3, 4, 5, 7, 8, 9, 13, 16, 17, 21, 22, 23, 24	1, 2, 3, 4, 5, 7, 9, 11, 12, 13, 16, 17, 18, 21, 22, 23, 24	1, 2, 3, 4, 5, 6, 7, 13, 19, 21, 22, 23, 24	0.12	0.19	0.43

As revealed in Table V, the devised BGA-GPR-based FS resulted in the variable subset with the best fitness value (lowest MSE). Therefore, the feature subset chosen by the devised FS strategy holds more important and non-repetitive variables than the other techniques. That is, an electric load prediction model whose training input is constituted by the variable subset obtained by the devised FS strategy can give accurate forecasting.

VIII. CONCLUSION

In this paper, a BGA-FS-based FS method for short-term electric load prediction is devised and implemented. The method used the GPR fitness function to select the combination of variables from the specified initial feature space. The devised FS has obtained a feature subset that achieved a better fitness value than the initial dataset with all the candidate variables. The FS results obtained by the devised FS

method are also compared with FS results obtained by other two conventional FS methods. The FS results by the proposed method outperformed the other FS results obtained by the other methods with respect to the MSE criterion formulated with the GPR model. In addition, the devised FS method has been applied for four different building electricity demands (residential, educational, office and mixed-use building buildings). It obtained the best training input variable subsets for improved short-term prediction of the electric loads of all the building types.

REFERENCES

- Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Com-puters & Elec. Eng.*, 40, 16–28.
- Han Li, A.T.Eseye, Jianhua Zhang, and Dehua Zheng, “A Double-stage Hierarchical Hybrid PSO-ANFIS Model for Short-term Wind Power Forecasting”, *IEEE GreenTech*, Denver, CO., USA, March 29–31, 2017.
- Vieira, S. M., Sousa, J. M. C. , & Runkler, T. A. (2010). Two cooperative ant colonies for feature selection using fuzzy models. *Expert Sys. with Apps.*, 37, 2714–2723.
- Gunasundari, S., Janakiraman, S., & Meenambal, S. (2016). Velocity bounded Boolean particle swarm optimization for improved feature selection in liver and kidney disease diagnosis. *Expert Sys. with Apps.*, 56, 28–47.
- Ali Jan Ghasab, M., Khamis, S., Mohammad, F., & Jahani Fariman, H. (2015). Feature decision-making ant colony optimization system for an automated recognition of plant species. *Expert Sys. with Apps.*, 42, 2361–2370.
- Ghareb, A. S., Bakar, A. A. , & Hamdan, A. R. (2016). Hybrid feature selection based on enhanced genetic algorithm for text categorization. *Expert Sys. with Apps.*, 49, 31–47.
- Gen, M., & Cheng, R. (2000). *Genetic algorithms and engineering optimization*: Vol. 7. John Wiley & Sons.
- Koprinska, I. , Rana, M. , & Agelidis, V. G. (2015). Correlation and instance based feature selection for electricity load forecasting. *Knowledge-Based Sys.*, 82 , 29–40 .
- Doquire, G., & Verleysen, M. (2013). Mutual information-based feature selection for multi label classification. *Neurocomput.*, 122, 148–155.
- Uguz, H. (2011). A two-stage feature selection method for text categorization by using information gain, principal component analysis and genetic algorithm. *Knowledge-Based Sys.*, 24, 1024–1032.
- Sikora, R., & Piramuthu, S. (2007). Framework for efficient feature selection in ge- netic algorithm based data mining. *European Journal of Oper. Research.*, 180, 723–737.
- Voronin, S., & Partanen, J. (2014). Forecasting electricity price and demand using a hybrid approach based on wavelet transform, ARIMA and neural networks. *Int. Journal of Energy Research*, 38, 626–637.
- Jurado, S., Nebot, A., Mugica, F., & Avellana, N. (2015). Hybrid methodologies for electricity load forecasting: Entropy-based feature selection with machine learning and soft computing techniques. *Energy*, 86, 276–291.
- Hu, Z. Y., Bao, Y. K., & Xiong, T. (2014). Comprehensive learning particle swarm optimization based memetic algorithm for model selection in short-term load fore- casting using support vector regression. *App. Soft Comput.*, 25, 15–25.
- Kazemi, S. M. R., Hoseini, M. M. S., Abbasian-Naghneh, S., & Rahmati, S. H. A. (2014). An evolutionary-based adaptive neuro-fuzzy inference system for intelligent short-term load forecasting. *Int. Trans. in Oper. Research*, 21, 311–326.
- Tonkovic, Z., Zekic-Susac, M., & Somolanci, M. (2009). Predicting natural gas consumption by neural networks. *Tehni `cki Vjesnik*, 16, 51–61.
- Salcedo-Sanz, S. , Munoz-Bulnes, J. , Portilla-Figueras, J. A. , & Del Ser, J. (2015). One-year-ahead energy demand estimation from macroeconomic variables using computational intelligence algorithms. *Energy Conv. & Man.*, 99, 62–71.
- Sheikhan, M., & Mohammadi, N. (2012). Neural-based electricity load forecasting using hybrid of GA and ACO for feature selection. *Neural Comp. & Apps.*, 21, 1961–1970.
- Urraca, R., SanzGarcia, A., Fernandez-Ceniceros, J., Sodupe-Ortega, E., & Martinez-dePison, F. J. (2015). Improving hotel room demand forecasting with a hybrid GA-SVR methodology based on skewed data transformation, feature selection and parsimony tuning. In E. Onieva, I. Santos, E. Osaba, H. Quintian, & E. Corchado (Eds.). In *Hybrid AI systems: Vol. 9121* (pp. 632–643).
- Liu, Y., Yin, Y. F., Gao, J. J., & Tan, C. L. (2008). Wrapper feature selection optimized SVM model for demand forecasting. In G. J. Wang, J. Chen, M. R. Fellows, & H. D. Ma (Eds.), *Proc. of the 9th int. conf. for young comp. scientists: Vol. 1–5* (pp. 953–958) .
- S.Jiang, K-S.Chin, L.Wang, G.Qu and K. L. Tsui, “Modified genetic algorithm-based feature selection combined with pre-trained deep neural network for demand forecasting in outpatient department,” *Expert Sys. with Apps.*, 82 (2017) 216–230.
- Holland John. *Adaptation in Natural and Artificial Systems*. Cambridge, MA: MIT Press. ISBN 978-0262581110; 1992.
- Jäkel, F., Schölkopf, B., & Wichmann, F. A. (2007). A tutorial on kernel methods for categorization. *Journal of Math. Psy.*, 51(6), 343–358.
- Mathworks T. *Statistics Toolbox User’s Guide* The MathWorks, Inc. 3 Apple Hill Drive Natick, MA 01760-2098, 2013.
- Civicioglu P. Transforming Geocentric Cartesian Coordinates to Geodetic Coordinates by Using Differential Search Algorithm. *Comps. & Geos.* 46: 2012; 229247.