Energy 164 (2018) 183–193

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

Automatic generation of models for energy demand estimation using Grammatical Evolution

J.M. Colmenar [a], J.I. Hidalgo [b], S. Salcedo-Sanz [c,*]

[a] Department of Computer Science, Universidad Rey Juan Carlos, Móstoles, Madrid, Spain
[b] Department of Computer Arquitecture, Universidad Complutense de Madrid, Madrid, Spain
[c] Department of Signal Processing and Communications, Universidad de Alcalá, Alcalá de Henares, Madrid, Spain

ARTICLE INFO

Article history:
Received 2 March 2018
Received in revised form
16 August 2018
Accepted 27 August 2018
Available online 30 August 2018

Keywords:
Energy demand estimation
Macro-economic variables
Grammatical evolution
Meta-heuristics

ABSTRACT

The estimation of total energy demand in a country from macro-economic variables is an important problem useful to evaluate the robustness of the country's economy. Since the first years of this century, meta-heuristics approaches have been successfully applied to this problem, for different countries and problem's parameterizations. Many of these works optimize prediction models which are based on classical polynomial or simple exponential relationships, which may not be the best option for an accurate energy demand estimation prediction. In this paper the use of Grammatical Evolution is proposed to generate new models for total energy demand estimation at country level. Grammatical Evolution is a class of Genetic Programming algorithm, which is able to automatically generate new models from input variables. In this case, Grammatical Evolution considers macro-economic variables from which it is able to generate new models for total energy demand estimation of a country, with a temporal prediction horizon of one year. The models generated by the Grammatical Evolution are further optimized in order to adjust their parameters to the energy demand estimation. This process is carried out by means of a Differential Evolution approach, which is run for every model generated by the Grammatical Evolution. Thus, the algorithmic proposal consists of a hybrid method, involving Grammatical Evolution for model generation and a Differential Evolution meta-heuristic for the models' parameter tuning. The performance of the proposed approach has been evaluated in two different problems of total energy demand estimation in Spain and France, with excellent results in terms of prediction accuracy.

1. Introduction

Energy demand is one of the most important indicators of a country's level of development [1]. Energy demand had been increasing at the beginning of the century in what seemed an unstoppable race, pushed up by an aggressive industrialization of developed and developing countries, the rapid population growth and the process of economy globalization produced by the massive development of Information and Communication Technologies [2]. This scenario changed around 2008, year from which the recent worldwide crisis produced a sudden reduction in the demand of total energy in almost all developed countries in the World [3]. Moreover, the increasing of the energy demand in the last three years has been interpreted as an indicator of economic recovery in many countries. Thus, it can be stated that, as the economy grows, the energy demand increases exponentially. This produces important environmental issues, contributing to global warming and climate change [4], which may compromise the future of next generations. Currently, over 80% of the energy demand in the world is covered by non-renewable energy sources such as coal or petroleum. The dependence of these non-renewable resources is huge in developing countries, whereas renewable resources are mainly deployed in developed ones. This seems to indicate that countries with a growing industrial activity happen to be more energy demanding from inefficient non-renewable sources than others with economies based on alternative sectors. In this context, predicting medium and long-term energy demand of a country from macro-economic indicators is a key problem faced by policy makers, with impact in countries' economies, nations' development and also with important environmental implications [5].

Different previous approaches have tackled this important problem of energy demand estimation, from different perspectives.

* Corresponding author.
E-mail address: sancho.salcedo@uah.es (S. Salcedo-Sanz).

https://doi.org/10.1016/j.energy.2018.08.199
0360-5442/

Acronyms and nomenclature

Acronyms

ARIMA	Auto-Regressive Integrated Moving Average model
ACO	Ants Colony Optimization
BVNS	Basic Variable Neighborhood Search
CRO	Coral Reefs Optimization algorithm
DE	Differential Evolution
ELM	Extreme Learning Machine
GA	Genetic Algorithm
GE	Grammatical Evolution
GDP	Gross Domestic Product
GP	Genetic Programming
PSO	Particle Swarm Optimization
TOE	Tons Oil Equivalent
VNS	Variable Neighborhood Search

Nomenclature

CxProb	Crossover probability (GE)
E_{n+1}	Actual energy demand at time $n+1$
$\widehat{E}_{n+1}$	Energy demand estimation at time $n+1$
G	Number of generations
MtProb	Mutation probability (GE)
MF	Mutation factor (DE)
N	Population size
P	Parent population
P_m	Population of models
P_o	Offspring population
RF	Recombination factor (DE)

Interestingly, many of the previous approaches have been focussed on estimating energy demand in emerging economies such as Turkey. The first approach to this problem was discussed in Ref. [6], where a genetic algorithm was used to optimize the parameters of an exponential model for energy demand estimation in Turkey. From this first contribution, other researchers have focused on energy demand estimation in Turkey, such as [7] where ARIMA methods were applied to estimate energy demand related to fuel [8], where an ACO algorithm was applied to energy demand estimation [9], where a PSO algorithm was applied to the problem [10], where the problem of energy demand estimation was tackled with artificial neural networks and [11] where PSO was also applied to energy demand in Turkey. The energy demand estimation problem problem has also been tackled in other countries, such as Korea [12] where artificial neural networks were applied; China, with works involving different algorithms such as [13] where a hybrid PSO-GA was discussed, or [14], where a hybrid Bat algorithm with Simulated Annealing was applied; Iran [15], where the energy demand estimation in the country was tackled by using Evolutionary Algorithms, or Spain [16], where Harmony Search and ELMs was discussed and [17], where an approach based on VNS was successfully applied to this problem. All these approaches have used social-economic predictive variables as inputs for a large variety of algorithms. There are different versions of the problem focused on alternative prediction variables (instead of total energy), such as electricity demand [18], industrial energy [19], transportation energy [20] or automotive fuels [21].

As can be seen, many of these previous works are based on different prediction models optimized by means of meta-heuristic algorithms. Note that the prediction models used were proposed in the first works devoted to this problem. Exponential function models and linear approaches were first proposed in Ref. [6], and massively used since then in the majority of works dealing with meta-heuristics for optimizing these models. In Ref. [15] several new models based on logarithmic and alternative exponential functions were proposed. All of them were optimized by a real-encoding GA for energy demand estimation in Iran. Since then, there is not, to our knowledge, a work dealing with suggesting alternative prediction models for energy demand estimation in the literature.

This paper tries to fill this lack of new energy prediction models by means of using Grammatical Evolution (GE) [22] for generating novel and robust models for energy demand estimation. GE is a subclass of the more general Genetic Programming (GP) [23], which is able to evolve very different models (or programs) using in this case an encoding in the form of chromosomes. The performance of this proposal is analyzed for generating new energy demand models in two real problems of one-year prediction of energy demand in Spain and France.

The structure of the rest of the paper is as follows: next section summarizes the most important characteristics of GE for models extraction. Section 3 details the specific grammars considered in this paper to be evolved. Section 4 shows how to hybridize GE with a DE algorithm in order to obtain the best set of parameters of the energy prediction models. Section 5 presents the experiments carried out and the results obtained in two real problems of energy demand estimation in Spain and France. A comparative study among all models generated and alternative existing models for energy demand is carried out for the case of Spain. Finally, Section 6 closes the paper by giving some final conclusions and remarks on future extension of this work.

2. Grammatical Evolution in models extraction

Grammatical Evolution is an implementation of GP where the individuals are represented by chromosomes instead of trees [22]. In order to generate the phenotype of an individual, GE performs a decoding process which is based on a context free grammar. This way, the grammar defines the search space of the algorithm and lessens the problem of the high number of non-effective individuals that is common in GP. Besides, given that the individuals are represented by chromosomes, the classical genetic operators like the crossover and mutation can be applied with a much lower complexity than the operators that work on the GP trees.

Many problems have been tackled with GE in very different areas such as Medicine [24], Computer Science [25] or Structures Optimization [26], among others. One of the key features of GE is the ability of the grammar to direct the exploration of the optimization algorithm. This way, it is possible to include specific information related to the problem at hand into the grammar, making the search more effective. Therefore, the design of the grammar is a key issue in GE.

The GE algorithm is similar to the GA scheme, and it is shown in Algorithm 1. The input parameters of the algorithm correspond to the size of the population (N), the number of generations (G), the probability of crossover (*CxProb*) and the probability of mutation (*MtProb*). As can be seen in the pseudo-code, the initial population is generated and evaluated, and a loop begins where each iteration corresponds to a different generation. The genetic operators are applied to the individuals from the population (steps 5 and 6), and