Renewable Energy 75 (2015) 93–101

Contents lists available at ScienceDirect

Renewable Energy

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

A Coral Reefs Optimization algorithm with Harmony Search operators for accurate wind speed prediction

S. Salcedo-Sanz [a,*], A. Pastor-Sánchez [a], J. Del Ser [b], L. Prieto [c], Z.W. Geem [d]

[a] Department of Signal Processing and Communications, Universidad de Alcalá, Alcalá de Henares, Madrid 28871, Spain
[b] Optima Unit, Tecnalia Research & Innovation, Zamudio 48170, Bizkaia, Spain
[c] Department of Energy Resource, Iberdrola, Madrid, Spain
[d] Department of Energy IT, Gachon University, Seongnam 461-701, South Korea

ARTICLE INFO

Article history:
Received 5 May 2014
Accepted 18 September 2014
Available online

Keywords:
Short term wind speed prediction
Feature selection
Coral Reefs Optimization
Harmony Search
Extreme learning machines

ABSTRACT

This paper introduces a new hybrid bio-inspired solver which combines elements from the recently proposed Coral Reefs Optimization (CRO) algorithm with operators from the Harmony Search (HS) approach, which gives rise to the coined CRO-HS optimization technique. Specifically, this novel bio-inspired optimizer is utilized in the context of short-term wind speed prediction as a means to obtain the best set of meteorological variables to be input to a neural Extreme Learning Machine (ELM) network. The paper elaborates on the main characteristics of the proposed scheme and discusses its performance when predicting the wind speed based on the measures of two meteorological towers located in USA and Spain. The good results obtained in these experiments when compared to naïve versions of the CRO and HS algorithms are promising and pave the way towards the utilization of the derived hybrid solver in other optimization problems arising from diverse disciplines.

1. Introduction

The importance of renewable sources is currently underpinned by a plethora of initiatives and reports, which have indeed crystallized in a set of recommendations and directives on wind energy published by relevant institutions and the European Commission itself. As to mention the recently posted Strategic Research Agenda of the European Wind Energy Association (EWEA) evinces that in 2013 the wind energy sector reached 117 GW of installed capacity in Europe, with a potential wind electricity yield of 257 TWh that amounts to 8% of the overall European yearly electricity consumption [1]. This report also sets a goal recommendation of 20% share of wind energy in the total European electricity consumption by 2020. These figures elucidate the essential necessity to develop regulatory and technology sharply aimed at improving dramatically the efficiency of turbines, substations and ancillary equipment when delivering wind energy to consumers. This challenge is closely related to the deployment, operational and management costs of wind farms, from which short-term wind speed prediction arise as one of the most relevant exogenous factors impacting on the efficiency and suitability of the wind park at hand.

Indeed, the stochastic nature of the wind resource makes it of utmost importance to infer an accurate prediction of its dynamics in order to estimate the wind farm production or yield in advance, as well as to avoid power transport problems – due to e.g. overcapacity – through the transmission network. To this end, modern short-term wind speed prediction models are mainly based on the combination of physical and statistical algorithms [2]. The physical (meteorological) models can be either global, meso-scale or even local [3], whereas statistical models are usually included in the prediction systems to improve the initial meteorological estimation of the wind dynamics. In the last few years, many different statistical approaches have been applied to wind speed prediction, including linear schemes [4], classical Box–Jenkins methodologies such as auto-regressive models [5] and other time-series analytical procedures such as the Mycielski algorithm [6], different clustering algorithms [7] and computational approaches such as neural networks [8–10], ensembles of neural learners [11], Bayesian methods [12], Support Vector Machines [13–15], or hybrid methods [16–18], among others.

The successful experiments and satisfactory results achieved when applying computational intelligence algorithms to wind speed prediction have motivated the extensive use of these algorithms in combination with meteorological models: in the majority

* Corresponding author. Tel.: +34 91 885 6731; fax: +34 91 885 6699.
E-mail addresses: sancho.salcedo@uah.es (S. Salcedo-Sanz), javidelser@gmail.com (J. Del Ser).

http://dx.doi.org/10.1016/j.renene.2014.09.027
0960-1481/

of cases, the meteorological models provide the set of variables that will be used by the computational intelligence approach in order to produce a prediction. However, the number of meteorological variables that can be input to computational intelligence algorithms is huge, ranging from wind speed to wind direction, temperature, pressure, humidity and other physical phenomena related to the magnitude to be predicted. Their cardinality increases even more if a grid of nodes is considered, usually set at different heights in which these variables can be measured or calculated. If this is the case, it is essential to impose a variable selection criteria to reduce the problem to a reasonable size for subsequent learning stages to perform adequately. In most cases, variable selection is carried out by randomly choosing a few meteorological variables under experimental criteria [8,9], but it is intuitive to note that a comprehensive analysis to discern the best set of variables to be input to a computational intelligence algorithm would potentially improve the quality of the results obtained. This process and the criteria on which it is based is well known in computational intelligence as Feature Selection Problem (FSP). In this context, specific works on FSP have recently appeared in the literature related to wind speed prediction. In Refs. [19] and [20] Particle Swarm Optimization (PSO) and Differential Evolution (DE) algorithms are proposed together with a K-nearest neighbors approach in order to select the best set of variables in a wind speed prediction problem. A classical multi-layer perceptron network is used as computational intelligence algorithm to improve the outcome of a physical model. More recently, in Ref. [21] a genetic algorithm is applied for inferring the best set of features to feed a neural network in a wind speed prediction problem.

This paper joins this line of research by presenting a novel hybrid bio-inspired algorithm for dealing with a hard FSP in a wind speed prediction problem. Specifically, the Coral Reefs Optimization algorithm is hybridized with Harmony Search operators for isolating the best meteorological variables that serve as inputs to a neural network, which acts as a regression approach to provide the final wind speed prediction. The Coral Reefs Optimization (CRO) algorithm is a novel bio-inspired approach from the family of evolutionary solvers. In its naïve version presented in Refs. [28], the novelty of this optimization algorithm was shown to reside in the selection process, which resembles the procedure by which corals fight for space in marine reefs. In this paper the CRO is modified to include Harmony Search operators in the reproduction part of the algorithm (broadcast spawning and brooding), which gives rise to a novel optimization scheme – hereafter coined as CRO-HS – with improved exploration properties with respect to the initially proposed CRO algorithm. This new solver utilizes a fast training neural network (Extreme Learning Machine, ELM) to evaluate the quality of the sets of input variables or *corals* iteratively refined by the CRO-HS. To shed light on the performance of the proposed algorithm two real problems are tackled, from data in Oregon (USA) and Zaragoza (Spain). In these problems we discuss the effect of the HS operators when included within the CRO search thread, and compare the obtained results with those of the naïve CRO and HS approaches.

The rest of this article is structured as follows: Section 2 defines the FSP problem under consideration, including the fitness function and the description of the ELM employed for its calculation. Section 3 delves into the proposed CRO-HS algorithm by sequentially describing the HS, CRO and the hybrid approach proposed in this manuscript. Section 4 embodies the experimental part of the paper, where the performance of the proposed approach is evaluated and benchmarked to original versions of its compounding meta-heuristics. Finally, Section 5 ends the paper with some concluding remarks.

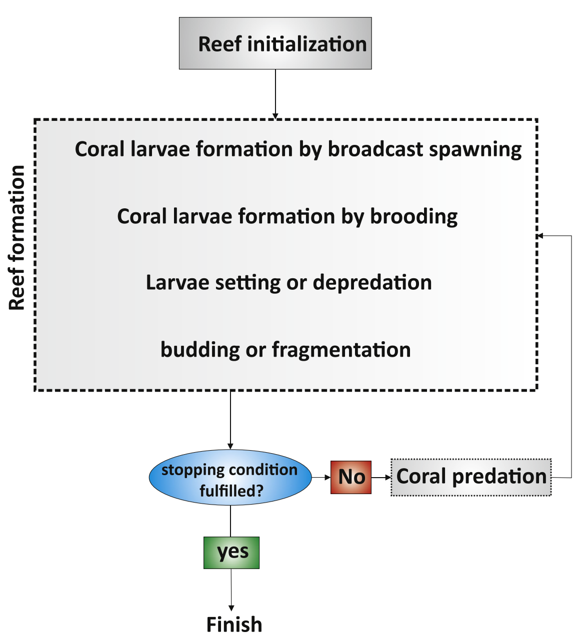

Fig. 1. Outline of the CRO algorithm.

2. Problem formulation

Let us consider an area of study (usually a wind farm or a prospection site) where a measuring tower $\mathfrak{M}$ measures wind speed values at a given rate. On the other hand, we consider a grid Ω surrounding $\mathfrak{M}$, formed by $N\times N$ nodes, where a time series of M meteorological variables are collected at each node of the grid, usually obtained from a given physics-based prediction model. The wind speed series in $\mathfrak{M}$, and the meteorological series in the points of the grid are assumed to be synchronized in time. The problem consists of predicting the wind speed values in $\mathfrak{M}$ by using the predictive meteorological variables of the grid points. Fig. 2 shows an example of a grid Ω and a measuring tower $\mathfrak{M}$. Note that we will consider a reduced number m of final meteorological variables out of the total $n = M \cdot N^2$ available for the wind speed prediction.

Bearing this in mind, the encoding for a given solution to this problem is as follows: each meteorological variable included in the prediction system needs a total of four parameters to be identified: (x,y,id,ma), where x stands for the x-coordinate in the grid, y stands for the y-coordinate in the grid, id stands for the variable identifier and ma is a binary indicator such that value 1 means that a moving-average of the series of the variable is being considered, and value 0 indicates that such a moving-average is not considered in the estimation. The final encoding of a solution can be hence displayed as a 4 m-length vector.

$$\Xi = [x_1, y_1, id_1, ma_1, \ldots, x_\mathrm{m}, y_\mathrm{m}, id_\mathrm{m}, ma_\mathrm{m}], \tag{1}$$

i.e. as the concatenation of the aforementioned 4 parameters for each of the feature variables considered for the wind speed prediction.

The objective or fitness function used to evaluate the goodness of a given solution Ξ to the FSP in the context of wind speed prediction is the root mean square error between the real wind speed v and the prediction given by a model $\widehat{v}$, which is computed as,

$$e(\Xi) \triangleq \frac{1}{K}\sqrt{\sum_{j=1}^{K}\left(v_j - \widehat{v}_j\right)^2}, \tag{2}$$

over a given validation set with K samples.