WIND ENERGY POTENTIAL ASSESSMENT BASED ON WIND DIRECTION MODELLING AND MACHINE LEARNING

P. Krömer*, S. Mišák†, J. Stuchlý†, J. Platoš*

Abstract: Precise wind energy potential assessment is vital for wind energy generation and planning and development of new wind power plants. This work proposes and evaluates a novel two-stage method for location-specific wind energy potential assessment. It combines accurate statistical modelling of annual wind direction distribution in a given location with supervised machine learning of efficient estimators that can approximate energy efficiency coefficients from the parameters of optimized statistical wind direction models. The statistical models are optimized using differential evolution and energy efficiency is approximated by evolutionary fuzzy rules.

1. Introduction

The knowledge of wind speed and direction is very important for wind energy generation, integration, and management. Predicting the amount of generated wind energy is essential for the safe and effective operation of stochastic renewable energy sources such as wind turbines and wind farms. The estimation of wind energy potential [28] is essential for the growth of green, renewable energy applications.

In the recent years, the construction of large wind power plants has stagnated, partly because of the exceeded transmission capacity of overhead power lines. However, small power plants with nominal power in the order of tens to hundreds of kilowatts are getting into the forefront of interest. This interest is given by the development of new technologies for off-grid systems. The Off-Grid systems are understood as systems independent of the supply of electric or heat energy from an external power grid and as systems using mainly renewable source [36]. Although

---
*Pavel Krömer, Jan Platoš – Corresponding author, Department of Computer Science, Faculty of Electrical Engineering and Computer Science, VŠB – Technical University of Ostrava, 17. listopadu 15, 708 33 Ostrava – Poruba, Czech Republic, E-mail: jan.platos@vsb.cz, pavel.kromer@vsb.cz

†Stanislav Mišák, Jindřich Stuchlý, Department of Electrical Energy, Faculty of Electrical Engineering and Computer Science, VŠB – Technical University of Ostrava, 17. listopadu 15, 708 33 Ostrava – Poruba, Czech Republic, E-mail: stanislav.misak@vsb.cz, jindrich.stuchly@vsb.cz

photovoltaic systems are used as dominant renewable sources of electric energy in most cases, they are not sufficient for energy production for the whole year due to its stochastic character. Hence, power plants are used as additional sources of electric or heat energy in these Off-Grids. The cooperation of wind power plants and photovoltaic systems is very advantageous. The energy production from wind power plants is dominant in autumn and winter seasons and energy from photovoltaic systems is dominant in spring and summer seasons [49]. The main problem with the construction of small power plants in Off-Grids lies in finding an acceptable plant location with good meteorological conditions including high wind velocity as well as fixed wind directionality [11, 33].

Current wind power plant energy production forecasting models are mostly based on wind velocity predictions. These models have a relatively good accuracy for large wind power plants that have wind turbine hub placed in an altitude of about 100 meters where the wind direction is stable. However, off-grid systems with nominal power of hundreds of kilowatts are usually composed of small wind power plants that are installed on top of towers with a maximum height of tens of meters and variable meteorological conditions. Another issue associated with small and mid-size wind power plants is their frequent placement in urban environment. In such locations, meteorological and in particular wind conditions are thanks to the presence of high-rise buildings significantly different from conditions found in the altitude of 100m. [1]. For example, wind velocity is smaller and wind directionality has a high variance. For these cases, prediction models based solely on wind velocity forecasts are less accurate [27,38]. Therefore, the development of tools and methods to estimate the volume of electric and heat energy produced by wind power plants with respect to both, wind velocity and wind direction forecasts, is very desirable.

This work introduces a novel two-stage method that can be used to estimate location-specific wind energy efficiency coefficient. In the first phase, an accurate statistical model of annual wind direction distribution [9,10] is created and carefully optimized [23, 24]. The differential evolution algorithm is in this work used for accurate optimization of model parameters. In the second stage, a recent multi-paradigm machine learning method called evolutionary fuzzy rules [30, 31, 41] is applied as an estimator of the annual wind energy efficiency coefficient. Machine-learning-based estimation of the annual wind energy efficiency coefficient is the major original contribution of this work.

The rest of this paper is organized in the following way: Section 2 summarizes relevant recent methods for wind energy potential assessment and wind direction modelling. It also provides a high–level outline of the proposed approach. Section 3 provides a brief introduction into the fundamentals of the employed methods, i.e. statistical wind direction modelling, differential evolution, and evolutionary fuzzy rules, and their application in the area of wind energy potential assessment. Computational experiments illustrating and evaluating the proposed approach are presented in Section 4 and conclusions are drawn in Section 5.

2.   Background and proposed approach

With regard to the results of the available research in the literature, it should be noted that most of the research work deals with the issue of the wind potential assessment of the wind power plant in different point of view.

Study presented in [35] focuses on onshore wind power on the African continent and strives to provide estimates of theoretical, geographical and technical potential using a GIS analysis, while the analysis is based on currently available wind power technology, wind power plant location or related socio-economic and geographic constraints. Methodology described in [44] is based on GIS as well, but adapted to the island/regional requirements and it has been even applied to a practical case in the Canary Islands.

Different techniques such as mesoscale with GIS implementation are used for wind energy potential assessment for Off-Shore systems [51]. Next, a good result in Wind energy yield assessment has been achieved in portfolio theory application in [34].

As an extremely useful can be considered the new suggested approach to assess an index of the local wind energy potential presented in [42]. Of course, the numerical methods for determining of wind energy potential are also present in currently available literature [4, 15, 43]. Closely related with the wind potential assessment is the forecasting of the wind power plant power output in short-term [32] mid-term [14] and long-term [48] time series.

2.1 Wind direction modelling

Wind direction modelling is an important step in optimization of wind energy potential assessment, wind farm layout [19], wind turbine operations and control [26], and wind energy integration. It is well-understood that accurate location-specific modelling of wind power output is essential for accurate management of such systems [21, 25, 26]. Statistical modelling can be used to create various types of accurate descriptive models of wind activity [25], that can be further optimized by advanced methods such as artificial neural networks [28, 45], support vector machines [32, 45], genetic algorithms [24], particle swarm optimization [23], and hybrid soft-computing approaches such as genetic fuzzy systems [16]. Simulations represent another possible way of quantifying different aspects of wind activity including wind direction and wind speed [21].

2.2 Proposed approach

In this work, we propose and evaluate a novel two–stage approach for wind energy efficiency estimation. It is based on statistical modelling of annual wind direction distribution, optimization of model parameters by differential evolution, and nature-inspired learning of an energy efficiency estimator based on the soft computing concept of evolutionary fuzzy rules.

The proposed approach works in the following way. First, a set of wind direction measurements from certain location is used to establish a statistical model of annual wind direction distribution in this place. Then, differential evolution (DE) is used to optimize model parameters. In the second stage, the optimized model is used as an input of an energy efficiency estimator in the form of artificially evolved fuzzy rule which approximates annual wind energy efficiency from its parameters.

The development of accurate energy efficiency estimator is a crucial part of the proposed method. In the proposed framework, a fuzzy rule evolved by genetic pro-
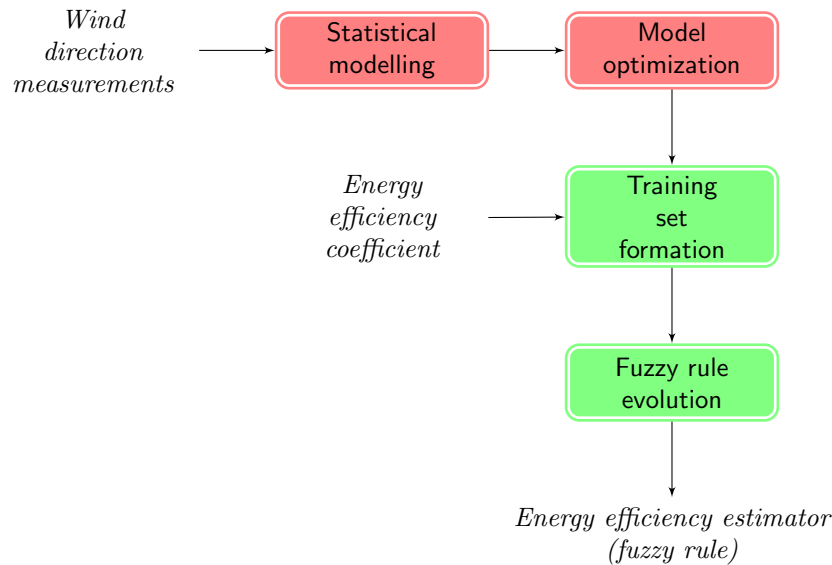
Fig. 1 *Supervised learning of energy efficiency estimator.*

gramming is used as an energy efficiency estimator. The evolution of the estimator, described in Fig. 1, is a supervised machine learning process that requires a training set composed of pairs of optimized wind direction distribution models and known annual wind energy production and known wind energy efficiency coefficients. The basics of all employed methods, i.e. statistical wind direction distribution modelling, differential evolution, wind direction distribution model optimization, and fuzzy rule evolution, are described in the following section.

3. Wind energy potential assessment

The proposed framework for wind energy potential assessment is based on three building blocks. It employs statistical modelling of annual wind direction distribution by a finite mixture of simple von Mises distributions and an optimization of distribution model parameters by differential evolution. Then, an estimation of wind energy efficiency coefficient by an evolutionary fuzzy rule machine-learned from a training data set is performed.

3.1 Statistical modelling of wind direction

Finite Mixture of simple von Mises distributions (MvM) can be used to represent multimodal directional data [23, 24]. Simple von Mises distribution (SvM), defined for random circular variable θ, has two parameters: μ represents the prevailing wind direction, and κ (concentration parameter) indicates the variance around the mean. The probability density function of this distribution (SvM-*pdf*) is given by

$$f(\theta; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu)}, \tag{1}$$

where $\kappa \geq 0$, $0 \leq \mu \leq 2\pi$, $0 \leq \theta \leq 2\pi$ and $I_0(\kappa)$ is the modified Bessel function of the first kind and order zero. The SvM-*pdf* is symmetrical and unimodal, and therefore can only approximate directional data with a single prevailing direction. For $\kappa = 0$, SvM distribution becomes uniform around the circle with all directions equally probable. When the modelled data contains more than one prevailing direction, it is necessary to use a mixture of such distributions. Finite mixture model of simple von Mises distributions is defined by probability density function (MvM-*pdf*)

$$\phi(\theta; \boldsymbol{\nu}) = \sum_{j=1}^{k} \omega_j \cdot f_j(\theta; \mu_j, \kappa_j), \tag{2}$$

where j is the index of particular SvM-*pdf* with parameters μ_j and κ_j, θ is an angular variable ($0 \leq \theta \leq 2\pi$), and $\boldsymbol{\nu}$ is a vector parameter

$$\boldsymbol{\nu} = (\boldsymbol{\mu}, \boldsymbol{\kappa}, \boldsymbol{\omega}) =$$
$$= (\mu_1, \ldots, \mu_k, \kappa_1, \ldots, \kappa_k, \omega_1, \ldots, \omega_k). \tag{3}$$

The weight of each mixture member has to be nonnegative and satisfy the following conditions

$$0 \leq \omega_j \leq 1 \quad \forall j \in \{1, \ldots, k\}, \qquad \sum_{j=1}^{k} \omega_j = 1. \tag{4}$$

Maximum likelihood estimates of MvM distribution vector parameter $\boldsymbol{\nu}$ lead to a system of nonlinear equations that must be solved numerically [20]. The parameter estimates for the MvM distribution can be evaluated using an expectation maximization algorithm fully described in [6]. In this work, the differential evolution is adopted in place of commonly used numerical methods. It is used to optimize an initial analytical approximation of distribution parameters obtained from data.

3.2 Differential evolution

The DE is a versatile and easy to use stochastic evolutionary optimization algorithm [40]. It is a population-based optimizer that evolves a population of real encoded vectors representing the solutions to given problem. The DE was introduced by Storn and Price in 1995 and it quickly became a popular alternative to the more traditional types of evolutionary algorithms. It evolves a population of candidate solutions by iterative modification of candidate solutions by the application of the differential mutation and crossover [40]. In each iteration, so called trial vectors are created from current population by the differential mutation and further modified by various types of crossover operator. In the end, the trial vectors compete with existing candidate solutions for survival in the population.

The DE starts with an initial population of N real-valued vectors. The vectors are initialized with real values either randomly or so, that they are evenly

spread over the problem space. The latter initialization leads to better results of the optimization [40]. During the optimization, the DE generates new vectors that are scaled perturbations of existing population vectors. The algorithm perturbs selected base vectors with the scaled difference of two (or more) other population vectors in order to produce the trial vectors. The trial vectors compete with members of the current population with the same index called the target vectors. If a trial vector represents a better solution than the corresponding target vector, it takes its place in the population [40].

The two most significant parameters of the DE are scaling factor and mutation probability [40]. The scaling factor $F \in [0, \infty]$ controls the rate at which the population evolves and the crossover probability $C \in [0, 1]$ determines the ratio of elements that are transferred to the trial vector from its opponent. The size of the population and the choice of operators are other important parameters of the optimization process.

The basic operations of the classic DE can be summarized using the following formulae [40]: the random initialization of the ith vector with N parameters is defined by

$$x_j^i = \text{rand}(b_j^{\text{L}}, b_j^{\text{U}}), \quad j \in \{1, \dots, N\}, \tag{5}$$

where b_j^{L} is the lower bound of j-th parameter, b_j^{U} is the upper bound of j-th parameter, and $\text{rand}(a, b)$ is a function generating a random number from the range $[a, b]$. A simple form of the standard differential mutation is given by

$$\mathbf{v}^i = \mathbf{v}^{r1} + F(\mathbf{v}^{r2} - \mathbf{v}^{r3}), \tag{6}$$

where F is the scaling factor and $\mathbf{v}^{r1}$, $\mathbf{v}^{r2}$, and $\mathbf{v}^{r3}$ are three random vectors from the population. The vector $\mathbf{v}^{r1}$ is the base vector, $\mathbf{v}^{r2}$ and $\mathbf{v}^{r3}$ are the difference vectors, and $\mathbf{v}^i$ is the trial vector. It is required that $i \neq r1 \neq r2 \neq r3$.

The uniform (binomial) crossover that combines the target vector, $\mathbf{x}^i$, with the trial vector, $\mathbf{v}^i$, is given by

$$v_j^i = \begin{cases} v_j^i & \text{if } (\text{rand}(0,1) < C) \text{ or } j = j_{\text{rand}} \\ x_j^i, & \text{otherwise} \end{cases} \tag{7}$$

for each $j \in \{1, \dots, N\}$. The random index j_{rand} is in the above selected randomly as $j_{\text{rand}} = \text{rand}(1, N)$. The uniform crossover replaces the parameters in $\mathbf{v}^i$ by the parameters from the target vector $\mathbf{x}^i$ with probability $1 - C$. The outline of the classic DE according to [18,40] is summarized in Algorithm 1. The DE is a successful evolutionary algorithm designed for continuous parameter optimization driven by the idea of scaled vector differentials. That makes it an interesting alternative to the wide spread genetic algorithms that are designed to work primarily with discrete encoding of the candidate solutions. As well as GA, the DE represents a highly parallel population based stochastic search meta-heuristic. In contrast to the GA, the differential evolution uses the real encoding of candidate solutions and different operations to evolve the population. It results in different search strategies and different directions found by the DE during crawling of a fitness landscape of the problem domain.

Algorithm 1 A summary of classic Differential Evolution.

Initialize the population P consisting of M vectors using Eq. (5)
Evaluate an objective function ranking the vectors in the population
while Termination criteria not satisfied **do**
    Let G = number of current generation
    for $i \in \{1, \dots, M\}$ **do**
        Differential mutation: Create trial vector $\mathbf{v}^i$ according to Eq. (6)

        Validate the range of coordinates of $\mathbf{v}^i$. Optionally adjust coordinates of $\mathbf{v}^i$ so, that it is
        valid solution to given problem

        Perform uniform crossover. Select randomly one parameter j_{rand} in $\mathbf{v}^i$ and modify the
        trial vector using Eq. (7)

        Evaluate the trial vector.
        if trial vector $\mathbf{v}^i$ represent a better solution than target vector $\mathbf{x}^i$ **then**
            add $\mathbf{v}^i$ to P^{G+1}
        else
            add $\mathbf{x}^i$ to P^{G+1}
        end if
    end for
end while

3.2.1   Optimizing a finite mixture of von Mises distributions by the DE

To optimize the parameters of a finite mixture of von Mises distributions, one needs to define a real-valued representation of candidate solutions, a fitness function, and a population initialization algorithm.

In this work, we use a parameter optimization strategy introduced by Heckenbergerova, et al. [23]. The authors have used particle swarm optimization to search for good parameters of the MvM model. The DE, however, has similar real-valued nature and identical problem representation can be used. In this work, the mixture of k von Mises distributions is represented by a candidate vector $\mathbf{v} = (v_1, \dots, v_n), v_i \in [0, 1]$ with three parts encoding the vector parameter $\boldsymbol{\nu} = (\boldsymbol{\mu}, \boldsymbol{\kappa}, \boldsymbol{\omega})$ respectively

$$\mathbf{v} = (\overbrace{v_1, \dots, v_k}^{\boldsymbol{\mu}}, \underbrace{v_{k+1}, \dots, v_{2k}}_{\boldsymbol{\kappa}}, \overbrace{v_{2k+1}, \dots, v_n}^{\boldsymbol{\omega}}). \tag{8}$$

The decoding of $\boldsymbol{\mu}$ involves scaling of v_i, $i \in \{1, \dots, k\}$ to $[0, 2\pi]$, and the decoding of $\boldsymbol{\kappa}$ requires scaling of v_i, $i \in \{k+1, \dots, 2k\}$ to $[0, 700]$. The upper bound of κ_j has been chosen with respect to precision of the numerical algorithm used to implement Bessel function $I_0(\kappa_j)$.

To satisfy the constraint (4) for the weights of mixture members, ω_j, the following decoding rule has been devised

$$\omega_j = \begin{cases} v_i, & \text{if } j = 1 \\ v_i \cdot \left(1 - \sum_{l=1}^{j-1} w_l\right), & \text{otherwise.} \end{cases} \tag{9}$$

This is the simplest decoding that guarantees that the sum of all weights ω_j is equal to 1 and does not impose any additional constraints on candidate vector handling.

All vectors created during the optimization process are valid candidate solutions representing a mixture of k von Mises distributions.

In the proposed approach, the initial population of particles is formed from randomly perturbed analytical estimates of $\boldsymbol{\mu}$, $\boldsymbol{\kappa}$, and $\boldsymbol{\omega}$ created according to the principles outlined in [23]. The fitness of each candidate solution is evaluated using root mean-squared error (RMSE), defined as

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{T}^{i=1} (O_i - np_i)^2}, \tag{10}$$

where T is the number of frequency classes, O_i is the observed frequency of i-th class, n is the sum of all observed frequencies, and p_i is the theoretical (modelled) probability of i-th frequency class. RMSE is a common measure often used to evaluate the differences between predicted and observed values. It combines intuitive interpretation with good mathematical properties [50]. Low value of RMSE suggests good fit of observed and theoretical probabilities while large values of RMSE are associated with loose correspondence between observed and modeled phenomena.

3.3 Evolutionary fuzzy rules

Fuzzy classification is an umbrella term for different methods capable of efficient soft classification of data. In contrast to its crisp counterpart, fuzzy classification provides a more sensitive tools for data analysis [7]. Fuzzy decision trees and fuzzy if-then rules are prime examples of efficient, transparent, and intelligible fuzzy classifiers and value estimators [7, 46].

The need for interpretable and linguistically comprehensible classification and regression tools is widely recognized [12, 47]. It is also well-known that bio-inspired and evolutionary methods possess the ability to learn and optimize various types of fuzzy systems [12, 13] and data mining models [3, 5]. Evolutionary fuzzy rules (FR) [30, 31, 41] are simple yet powerful classification and regression instruments based on the merger of fuzzy information retrieval (IR) and genetic programming.

Fuzzy information retrieval uses extended Boolean queries that consist of search terms, operators, and weights, and evaluates them against an internal representation (index) of a collection of documents. It is based on the fuzzy set theory and fuzzy logic that facilitate flexible and accurate search [39]. Evolutionary fuzzy rules use similar basic concepts, data structures, and operations, and apply them to general data processing tasks such as classification, prediction, and so forth. Here, the concepts of information retrieval are employed to interpret data and to define the classification or regression models. Symbolic rules of such models are then evolved using genetic programming [2], a generic, problem-independent meta-heuristic machine learning algorithm.

The data processed by a fuzzy rule is a real valued matrix. Each row of the matrix corresponds to a single data record interpreted as a fuzzy set of features. Such a general, real valued matrix $\boldsymbol{D}$ with m rows (data records) and n columns (data attributes, features) can be mapped to an IR index that describes a collection of objects.

526

A fuzzy rule takes the form of a weighted symbolic expression roughly corresponding to an extended Boolean query in the fuzzy IR analogy. The rule consists of weighted feature names and weighted aggregation operators. The evaluation of such an expression assigns to each data record a real value from the range $[0, 1]$. Such valuation can be interpreted as an ordering, labeling, or a fuzzy set induced on the data records.

The fuzzy rule is a symbolic expression that can be parsed into a tree structure consisting of nodes and leafs. There are three types of leafs (a.k.a. terminal nodes)

a) *Feature* node which represents the name of a feature (a search term in the IR analogy). It specifies a requirement for a particular feature in the currently processed data record.

b) *Past feature* node which defines a requirement on certain feature in a previous data record. The index of the previous data record (current - 1, current - 2, etc.) is a parameter of the node.

c) *Past output* node which puts the requirement on a previous output of the predictor. The index of the previous output (current - 1, current - 2, etc.) is a parameter of the node.

A fuzzy rule can be expressed using a simple infix notation

feature1:0.5 and:0.4 (feature2[1]:0.3 or:0.1 ([1]:0.1 and:0.2 [2]:0.3)),

where *feature1:0.5* is a feature node, *feature2[1]:0.3* is a past feature node, and *[1]:0.5* is a past output node. Different node types can be used when dealing with different data sets. For example, the past feature node and past output node are useful for the analysis of time series and data sets where the ordering of the records matters. But their usage is pointless for the analysis of regular data sets. The feature node is the basic building block of classifiers and predictors developed for arbitrary data.

The evaluation of a node, *f:a*, with value $f \in [0, 1]$ and weight $a \in [0, 1]$, is performed using a threshold interpretation of the retrieval status value (RSV) concept known from fuzzy information retrieval [8]

$$g(f, a) = \begin{cases} P(a)\frac{f}{a}, & f < a \\ P(a) + Q(a)\frac{f-a}{1-a}, & f \geq a \end{cases},$$ (11)

where $P(a) = \frac{1+a}{2}$ and $Q(a) = \frac{1-a^2}{4}$ are coefficients used to fine-tune the threshold curve.

Typical evolutionary fuzzy rules support *and*, *or*, *not*, *prod*, and *sum* operator nodes. However, more general or domain specific operators can be used as well. Both nodes and leafs are weighted to soften the criteria they represent.

The operators *and*, *or*, *not*, *prod*, and *sum* are evaluated using fuzzy set operations. In this study, the standard t-norm and s-norm [52] are used to implement *and* and *or* operators, respectively

$$t(x, y) = \min(x, y),$$
$$s(x, y) = \max(x, y),$$

operator *not* is evaluated using the standard fuzzy complement

$$c(x) = 1 - x,$$

and *prod* and *sum* operators, respectively, using the product t-norm and its dual s-norm, bounded sum

$$t_{\text{prod}}(x, y) = xy,$$
$$s_{\text{sum}}(x, y) = a + b - ab.$$

However, other classes of complement, intersection, and union models [29, 37] can be used as well.

A fuzzy rule is a simple version of a general fuzzy rule-based system that consists of a single expression describing soft requirements on data records in terms of their features. In evolutionary fuzzy rules, this expression is evolved using genetic programming [2]. The tree structures, corresponding to the parsed fuzzy rules, are developed by an iterative application of crossover, mutation, and selection operators in order to find an accurate model of the training data.

The general process of rule evolution is used for data-driven search for custom classifiers or predictors. Different data sets may by characterized by different properties and different hidden structure, and the adaptability of genetic programming is essential for the evolution of problem-specific fuzzy rules. On the other hand, the stochastic nature of genetic programming introduces probabilistic elements into the process of rule evolution.

The evolutionary fuzzy rules, although machine-generated, retain the understandable structure and ease of interpretation inherited from the extended Boolean search expressions, and allow a soft classification/regression without the complexity and computational costs of full-featured fuzzy rule-based systems [30]. For more information on fuzzy rules, their structure, weighting, evaluation, evolution, and applications, see e.g. [30, 31, 41].

In this work, FRs are used as a machine learning tool to discover annual wind energy coefficients from parameters of optimized statistical models (MvM) of wind direction distributions.

4. Experiments

All building blocks of the proposed energy potential assessment framework were implemented from scratch in C++ and a series of computational experiments was conducted. They used a unique real-world data set combining the information about long-term wind direction measurements and the volumes of wind energy produced by particular wind power plants. The aim of the experiments was to confirm the viability of the proposed approach to accurately learn models of wind energy potential from real–world data, collected from several wind turbines operating in the Czech Republic. This section discusses the data, the experiments and their parameters, and obtained results.

4.1 Description of sites and data

This sub-section describes the wind turbine properties as well as data properties. Wind turbines properties are summarized in following Tab. I.

Site	Manufacturer	Type	Installed power [kW]	Year of installation
Drahany	Vestas	V90	2000	2006
Kamen	Vestas	V90	2000	2008
Lipna	Vestas	V90	2000	2008
Maletin	Vestas	V90	2000	2008
Veseli 26367	Vestas	V90	2000	2008
Maletin	Vestas	V90	2000	2008
Veseli 26368	Vestas	V90	2000	2008

Tab. I *Wind turbine properties.*

As can be seen from Tab. I, all the wind turbines are from the same manufacturer, are the same type with the same installed power. The only differences are in the place of installation, the altitude and also in the surrounding environment.

All the data are provided by the wind turbine operator and are obtained from the installed monitoring system which is integral of each wind turbine. Moreover, the monitoring system is the same in each wind turbine, because they have the same properties. The experimental data were collected for year 2011, where there were no missing data in the dataset. All the data were stored into database in 10-min average values.

4.2 Optimizing wind direction distribution parameters by DE

A /DE/1/rand differential evolution with population size 100, scaling factor, $F = 0.9$, crossover probability, $C = 0.9$, and maximum number of generations, $G_{\max} = 50,000$, was used to optimize the parameters of a finite mixture of 6 simple von Mises distributions for each site. The fixed parameters were set on the basis of best practices [40], initial trials, and past experience of the authors. The number of simple von Mises distributions corresponds to the number of distributions suggested in the literature [9, 23].

The optimization algorithm used problem representation and fitness function introduced in Section 3.2.1. Because of the stochastic nature of the algorithm, each experiment was performed 30 times, independently. The results of the optimization in terms of RMSE are summarized in Tab. III. For comparison purposes, the table also contains for each location the RMSE of analytical wind distribution model (i.e. model without optimization), created according to the principles summarized e.g. in [23]. The parameters of the best models found for each location are listed in Tab. II and visually illustrated in Figs. 2–7.

Tab. III and Figs. 2–7 clearly illustrate that the DE-optimized models fit the wind direction data for each location significantly better than the models with

Site	μ	κ	ω
Drahany	(0.86, 3.55, 5.53, 2.99, 1.16, 4.91)	(189.98, 13.08, 32.13, 25.17, 5.94, 0.7)	(0.02, 0.02, 0.05, 0.07, 0.07, 0.77)
Kamen	(5.76, 2.97, 6.2, 2.93, 4.77, 5.37)	(34.95, 157.66, 9.05, 12.73, 5.72, 0)	(0.06, 0.08, 0.08, 0.17, 0.25, 0.37)
Lipna	(5.17, 1.91, 6.28, 0.13, 1.95, 5.24)	(43.97, 42.48, 0.45, 14.2, 3.56, 2.96)	(0.05, 0.08, 0.1, 0.12, 0.25, 0.38)
Maletin	(1.6, 5.05, 5.48, 6.28, 2.22, 3.77)	(0, 34.02, 34.59, 1.9, 4.72, 1.31)	(0, 0.08, 0.19, 0.2, 0.24, 0.29)
Veseli WEA26367	(5.28, 1.55, 0.02, 3.02, 2.26, 5.06)	(700, 18.32, 0.65, 3.93, 34.7, 6.28)	(0.01, 0.1, 0.11, 0.22, 0.25, 0.31)
Veseli WEA26368	(0.39, 2.75, 0.65, 4.14, 3.43, 0)	(700, 18.45, 0.46, 4.08, 43.41, 5.65)	(0.01, 0.09, 0.12, 0.23, 0.25, 0.31)

Tab. II *Parameters of best found wind direction models.*

Site	RMSE [-]			
	DE-optimized model			Analytical
	min	avg (σ)	max	model
Drahany	78.97	91.90 (12.74)	123.36	258.20
Kamen	64.77	86.85 (21.31)	136.27	565.70
Lipna	49.16	61.33 (10.33)	89.27	498.71
Maletin	83.07	96.37 (8.91)	107.10	452.94
Veseli 26367	62.89	69.06 (8.78)	83.67	618.46
Veseli 26368	41.82	43.63 (2.09)	48.59	609.10

Tab. III *Minimum, average, and maximum RMSE of optimized wind distribution models found for each location by the DE. Standard deviation, σ, is provided as well.*

parameters set by the analytical method. This is in line with previous results on optimization of wind direction distribution parameters by nature-inspired methods presented in [23,24]. In this work, the optimized wind direction distribution models, describing accurately annual wind direction distribution in given locations, are further utilized to approximate wind energy potential of the corresponding sites.

(a) Wind rose diagram.

(b) MvM distribution components.
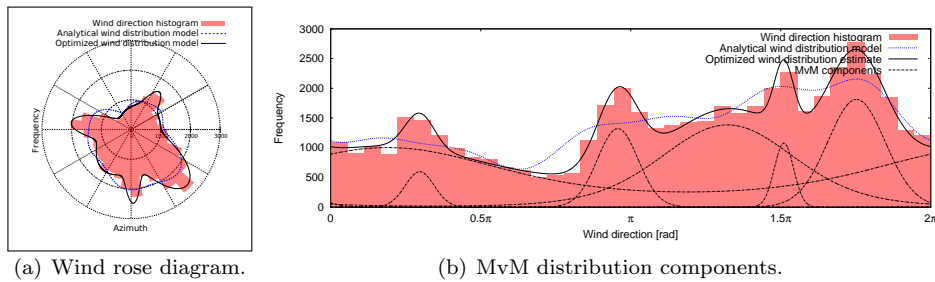
Fig. 2 *Optimized wind direction model for Drahany site.*

(a) Wind rose diagram.

(b) MvM distribution components.
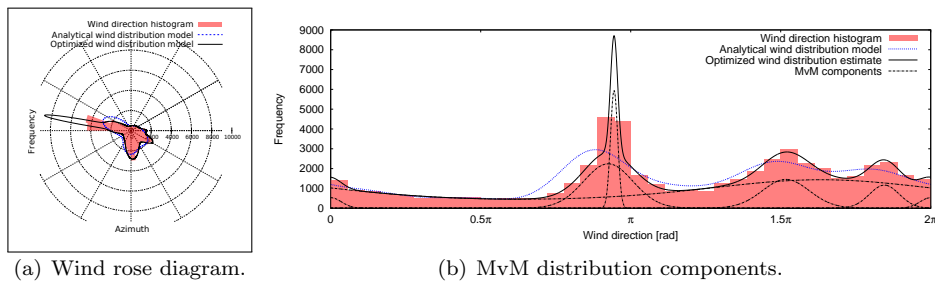
Fig. 3 *Optimized wind direction model for Kamen site.*

4.3   Learning annual wind energy efficiency coefficient by FR

Wind direction distribution models, optimized by the DE, and annual wind energy efficiency coefficients, based on actual production of wind power plants located at corresponding sites, were used to evolve energy efficiency estimator based on evolutionary fuzzy rules. Genetic programming [2] was used to evolve a fuzzy rule mapping wind distribution model parameters onto annual energy efficiency coefficients. The coefficients of all wind power plants considered in this study are summarized in Tab. IV.

Site	Annual wind energy efficiency coefficient [%]
Drahany	28.738
Kamen	29.969
Lipna	27.864
Maletin	19.930
Veseli 26367	25.090
Veseli 26368	22.929

Tab. IV *Annual wind energy efficiency coefficients.*

The GP employed 100 candidate fuzzy rules, mutation probability $m = 0.2$, crossover probability $c = 0.9$, a limit of 50,000 generations, and the F-score measure known from the area of IR as a fitness function [30]. Also in this case, the parameters were set with respect to best practices and initial trial-and-error runs.

(a) Wind rose diagram.   (b) MvM distribution components.
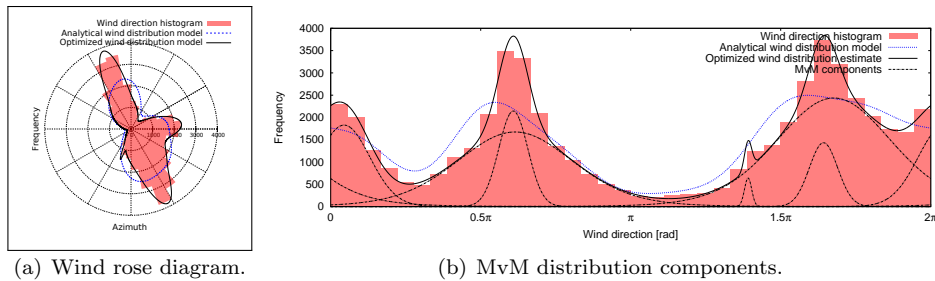
Fig. 4 *Optimized wind direction model for Lipna site.*

Because of the small number of input – output (i.e. optimized wind distribution model – wind energy efficiency coefficient) pairs, a leave-one-out cross-validation [22] was used to evaluate the energy efficiency estimator. In the leave-one-out cross-validation, $n - 1$ data records are used to evolve a model (i.e. they form a training data set) and the last remaining record is employed to asses the generalization ability of the model (i.e. assumes the role of a test data set). Every possible combination of training and test data sets is used and average accuracy of the model is reported.

(a) Wind rose diagram. (b) MvM distribution components.
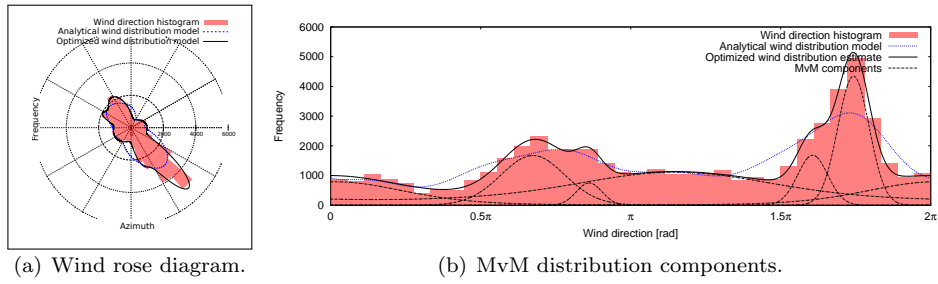
Fig. 5 *Optimized wind direction model for Maletin site.*

(a) Wind rose diagram. (b) MvM distribution components.
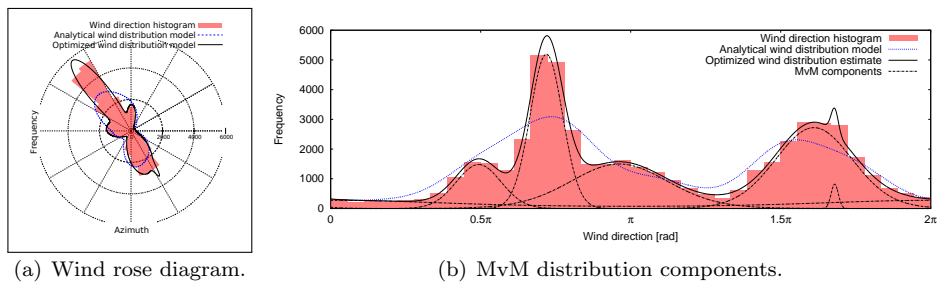
Fig. 6 *Optimized wind direction model for Veseli 26367 site.*

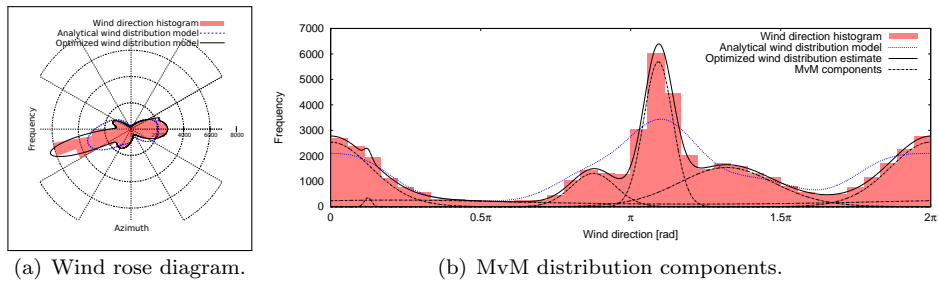(a) Wind rose diagram. (b) MvM distribution components.

Fig. 7 *Optimized wind direction model for the Veseli 26368 site.*

Also the GP is a stochastic procedure and so the cross-validation was performed 30 times independently. Results of annual wind energy efficiency coefficient learning are summarized in Tab. V. The table shows that the best evolved estimators were able to approximate the annual energy efficiency coefficient perfectly, with 0 RMSE, average estimators achieved RMSE smaller than 0.5 %, and even the worst found estimators were able to approximate the annual wind energy efficiency coefficient with error smaller than 5%. An example of a perfect estimator (i.e. an estimator with RMSE = 0 for both, training and test data) is graphically illustrated in Fig. 8. The figure illustrates the combination of wind direction distribution model parameters that are employed to approximate the annual efficiency coefficient. Labels of

533

	RMSE [%]		
Data	min	avg (σ)	max
Training	0	0.43 (0.49)	2.68
Test	0	0.37 (0.85)	4.90

Tab. V *Results of annual wind energy efficiency coefficient learning.*
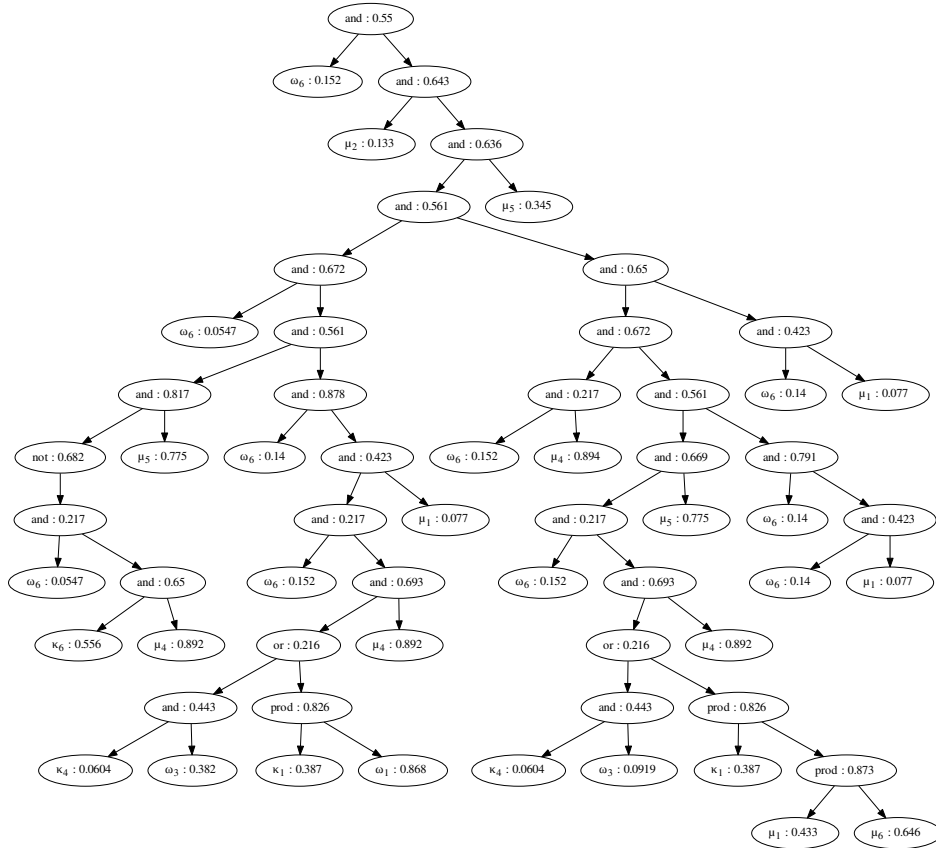
Fig. 8 *A sample annual wind energy efficiency coefficient estimator.*

the terminal nodes show which model parameters are utilized by the estimator and labels of the non-terminal nodes illustrate which operators are used to aggregate the values of terminals and branches.

5. Conclusions

This study proposed a two-stage method for accurate location-specific wind energy potential assessment. The method uses accurate statistical modelling of wind direction distribution and machine learning of annual efficiency coefficient estimators

for wind energy. In this framework, an annual wind direction distribution model in the form of a finite mixture of von Mises distributions is developed from long-term wind direction observations. Its parameters are then optimized by differential evolution to provide an accurate representation of the conditions on the investigated site.

The parameters of such optimized model are in the second phase of the proposed procedure used to approximate the annual wind energy efficiency coefficient for this location by a machine-learned estimator. The estimator was in this study based on evolutionary fuzzy rules, a hybrid soft-computing classification and regression method based on fuzzy information retrieval and genetic programming. Computational experiments have shown that evolutionary fuzzy rules can find good energy efficiency estimators. Although only a limited number of samples was used for estimator evolution and evaluation due to the uniqueness of this type of data, the performed leave-one-out cross-validation analysis suggests that the evolved estimators have a good ability to generalize and can be used to approximate the energy efficiency coefficient of unknown sites with high accuracy.

Extending of this work may be done in many ways. Different methods for optimization of the wind direction distribution parameters evaluated within the context of energy efficiency approximation and different regression methods used in place of evolutionary fuzzy rules are two of them. Moreover, the optimized wind direction distribution models, together with e.g. rejection sampling [17], can be used to simulate the wind direction on specific sites in-silico with a high degree of precision. Such realistic models used as a part of large-scale wind energy generation, transmission, and consumption simulations that can contribute to a new level of truly 'smart' energy grid operations.

Acknowledgement

References

[1] ABOHELA I., HAMZA N., DUDEK S. Effect of roof shape, wind direction, building height and urban configuration on the energy yield and positioning of roof mounted wind turbines. *Renewable Energy*. 2013, 50, pp. 1106–1118, doi: 10.1016/j.renene.2012.08.068.

[2] AFFENZELLER M., WINKLER S., WAGNER S., BEHAM A. *Genetic Algorithms and Genetic Programming: Modern Concepts and Practical Applications*. Chapman & Hall/CRC, 2009, doi: 10.1201/9781420011326.

[3] AFZAL W., TORKAR R. On the application of genetic programming for software engineering predictive modeling: A systematic review. *Expert Systems with Applications*. 2011, 38 (9), pp. 11984–11997, doi: 10.1016/j.eswa.2011.03.041.

[4] ARSLAN T., BULUT Y.M., YAVUZ A.A. Comparative study of numerical methods for determining weibull parameters for wind energy potential. *Renewable and Sustainable Energy Reviews*. 2014, 40, pp. 820–825, doi: 10.1016/j.rser.2014.08.009.

[5] BACARDIT J., LLORÀ X. Large-scale data mining using genetics-based machine learning. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery.* 2013, 3 (1), pp. 37–61, doi: 10.1002/widm.1078.

[6] BANERJEE A., DHILLON I.S., GHOSH J., SRA S. Clustering on the unit hypersphere using von mises-fisher distributions. *J. Mach. Learn. Res.* 2005, 6, pp. 1345–1382.

[7] BEZDEK J.C., KELLER J., KRISNAPURAM R., PAL N.R. *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing (The Handbooks of Fuzzy Sets).* Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005, doi: 10.1007/b106267.

[8] BORDOGNA G., PASI G. Modeling vagueness in information retrieval. In *Lectures on information retrieval.* Springer-Verlag New York, Inc., New York, NY, USA, 2001, pp. 207–241, doi: 10.1007/3-540-45368-7_10.

[9] CARTA J.A., BUENO C., RAMÍREZ P. Statistical modelling of directional wind speeds using mixtures of von mises distributions: Case study. *Energy Conversion and Management.* 2008, 49(5), pp. 897–907, doi: 10.1016/j.enconman.2007.10.017.

[10] CARTA J.A., RAMÍREZ P., BUENO C. A joint probability density function of wind speed and direction for wind energy analysis. *Energy Conversion and Management.* 2008, 49(6), pp. 1309–1320, doi: 10.1016/j.enconman.2008.01.010.

[11] CHEN N., LI Y., XIANG H. A new simulation algorithm of multivariate short-term stochastic wind velocity field based on inverse fast fourier transform. *Engineering Structures.* 2014, 80, pp. 251–259, doi: 10.1016/j.engstruct.2014.09.012.

[12] CORDÓN O. A historical review of evolutionary learning methods for mamdani-type fuzzy rule-based systems: Designing interpretable genetic fuzzy systems. *International Journal of Approximate Reasoning.* 2011, 52(6), pp. 894–913, doi: 10.1016/j.ijar.2011.03.004.

[13] CORDÓN O., GOMIDE F., HERRERA F., HOFFMANN F., MAGDALENA L. Ten years of genetic fuzzy systems: current framework and new trends. *Fuzzy Sets and Systems.* 2004, 141(1), pp. 5–31, doi: 10.1016/S0165-0114(03)00111-8.

[14] CROONENBROECK C., DAHL C.M. Accurate medium-term wind power forecasting in a censored classification framework. *Energy.* 2014, 73, pp. 221–232, doi: 10.1016/j.energy.2014.06.013.

[15] DE ANDRADE C.F., AO MAIA NETO H.F., ROCHA P.A.C., DA SILVA M. E.V. An efficiency comparison of numerical methods for determining weibull parameters for wind energy applications: A new approach applied to the northeast region of brazil. *Energy Conversion and Management.* 2014, 86, pp. 801–808, doi: 10.1016/j.enconman.2014.06.046.

[16] DE LA ROSA J.J.G., PÉREZ A.A., SALAS J.C.P., LEO J.G.R., NOZ A.M.M. A novel inference method for local wind conditions using genetic fuzzy systems. *Renewable Energy.* 2011, 36(6), pp. 1747–1753, doi: 10.1016/j.renene.2010.12.017.

[17] DEVROYE L. *Non-Uniform Random Variate Generation.* Springer-Verlag, New York, NY, 1986, doi: 10.1007/978-1-4613-8643-8.

[18] ENGELBRECHT A. *Computational Intelligence: An Introduction, 2nd Edition.* Wiley, New York, NY, USA, 2007, doi: 10.1002/9780470512517.

[19] FENG J., SHEN W.Z. Modelling wind for wind farm layout optimization using joint distribution of wind speed and wind direction. *Energies.* 2015, 8(4), pp. 3075–3092, doi: 10.3390/en8043075.

[20] FISHER N.I. *Statistical analysis of circular data.* Cambridge University Press, 1995, doi: 10.1017/cbo9780511564345.

[21] GAUMOND M., RÉTHORÉ P.-E., OTT S., PEÑA A., BECHMANN A., HANSEN K.S. Evaluation of the wind direction uncertainty and its impact on wake modeling at the horns rev offshore wind farm. *Wind Energy.* 2014, 17(8), pp. 1169–1178, doi: 10.1002/we.1625.

[22] HASTIE T., TIBSHIRANI R., FRIEDMAN J. *The Elements of Statistical Learning.* Springer Series in Statistics. Springer New York Inc., New York, NY, USA, 2001, doi: 10.1007/978-0-387-21606-5.

[23] HECKENBERGEROVÁ J., MUSÍLEK P., KRÖMER P. Optimization of wind direction distribution parameters using particle swarm optimization. In *Afro-European Conference for Industrial Advancement*, A. Abraham, P. Krömer, and V. Snasel, Eds., vol. 334 of *Advances in Intelligent Systems and Computing*. Springer International Publishing, 2015, pp. 15–26, doi: 10.1007/978-3-319-13572-4_2.

[24] HECKENBERGEROVÁ J., MUSÍLEK P., MEJZNAR J., VANČURA M. Estimation of wind direction distribution with genetic algorithms. In *26th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*. 2013, pp. 1–4, doi: 10.1109/CCECE.2013. 6567681.

[25] HIRAT, Y., MANDIC D.P., SUZUKI H., AIHARA K. Wind direction modelling using multiple observation points. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. 2008, 366(1865), pp. 591–607, doi: 10. 1098/rsta.2007.2112.

[26] HIRATA Y., SUZUKI H., AIHARA K. Wind modelling and its possible application to control of wind farms. In *Signal Processing Techniques for Knowledge Extraction and Information Fusion*, D. Mandic, M. Golz, A. Kuh, D. Obradovic, and T. Tanaka, Eds. Springer US, 2008, pp. 23–36, doi: 10.1007/978-0-387-74367-7_2.

[27] JUNG J., BROADWATER R.P. Current status and future advances for wind speed and power forecasting. *Renewable and Sustainable Energy Reviews*. 2014, 31(0), pp. 762–777, doi: 10.1016/j.rser.2013.12.054.

[28] JUNG S., KWON S.-D. Weighted error functions in artificial neural networks for improved wind energy potential estimation. *Applied Energy*. 2013, 111(0), pp. 778–790, doi: 10.1016/ j.apenergy.2013.05.060.

[29] KLIR G.J., YUAN B. *Fuzzy Sets and Fuzzy Logic; Theory and Applications*. Prentice Hall, Upper Saddle River, N. Y., 1995, doi: 10.1142/2895.

[30] KRÖMER P., OWAIS S.S.J., PLATOŠ J., SNÁŠEL V. Towards new directions of data mining by evolutionary fuzzy rules and symbolic regression. *Computers & Mathematics with Applications*. 2013, 66(2), pp. 190–200, doi: 10.1016/j.camwa.2013.02.017.

[31] KRÖMER P., PLATOŠ J., SNÁŠEL V., ABRAHAM A. Fuzzy classification by evolutionary algorithms. In *Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on* (2011), pp. 313–318, doi: 10.1109/ICSMC.2011.6083684.

[32] LAHOUAR A., BEN HADJ SLAMA J. Wind speed and direction prediction for wind farms using support vector regression. In *Renewable Energy Congress (IREC), 2014 5th International* (March 2014), pp. 1–6, doi: 10.1109/IREC.2014.6826932.

[33] MASSERAN N. Markov chain model for the stochastic behaviors of wind-direction data. *Energy Conversion and Management*. 2015, 92(0), pp. 266–274, doi: 10.1016/j.enconman. 2014.12.045.

[34] MEDIMOREC D., TOMŠIČ Ž. Portfolio theory application in wind potential assessment. *Renewable Energy*. 2015, 76(0), pp. 494–502, doi: 10.1016/j.renene.2014.11.033.

[35] MENTIS D., HERMANN S., HOWELLS M., WELSCH M., SIYAL S.H. Assessing the technical wind energy potential in africa a gis-based approach. *Renewable Energy*. 2015, 83(0), pp. 110–125, doi: 10.1016/j.renene.2015.03.072.

[36] MIŠÁK S., STUCHLÝ J., PLATOŠ J., KRÖMER P. A heuristic approach to active demand side management in off-grid systems operated in a smart-grid environment. *Energy and Buildings*. 2015, 96, pp. 272–284, doi: http://dx.doi.org/10.1016/j.enbuild.2015.03.033.

[37] MUSÍLEK P., GUANLAO R., BARREIRO G. Genetic programming of fuzzy aggregation operations. *Journal of Intelligent and Fuzzy Systems*. 2005, 16(2), pp. 107–118.

[38] OLAOFE Z.O. A 5-day wind speed & power forecasts using a layer recurrent neural network (lrnn). *Sustainable Energy Technologies and Assessments*. 2014, 6, pp. 1–24, doi: 10.1016/ j.seta.2013.12.001.

[39] PASI G. Fuzzy sets in information retrieval: State of the art and research trends. In: *Fuzzy Sets and Their Extensions: Representation, Aggregation and Models*, H. Bustince, F. Herrera, and J. Montero, Eds., 220, *Studies in Fuzziness and Soft Computing*. Springer Berlin/Heidelberg, 2008, pp. 517–535, doi: 10.1007/978-3-540-73723-0_26.

[40] PRICE K.V., STORN R.M., LAMPINEN J.A. *Differential Evolution A Practical Approach to Global Optimization.* Natural Computing Series. Springer-Verlag, Berlin, Germany, 2005.

[41] PROKOP L., MIŠÁK S., NOVOSÁD T., KRÖMER P., PLATOŠ J., SNÁŠEL V. Artificially evolved soft computing models for photovoltaic power plant output estimation. In: *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, SMC 2012, Seoul, Korea (South), October 14-17, 2012* (2012), pp. 1011–1016, doi: `10.1109/ICSMC.2012.6377861`.

[42] RITTER M., SHEN Z., CABRERA B.L., ODENING M., DECKERT L. Designing an index for assessing wind energy potential. *Renewable Energy.* 2015, 83, pp. 416–424, doi: `10.1016/j.renene.2015.04.038`.

[43] ROCHA P.A.C., DE SOUSA R.C., DE ANDRADE C.F., DA SILVA M. E.V. Comparison of seven numerical methods for determining weibull parameters for wind energy generation in the northeast region of brazil. *Applied Energy.* 2012, 89(1), pp. 395–400, doi: `10.1016/j.apenergy.2011.08.003`.

[44] SCHALLENBERG-RODRÍGUEZ J., DEL PINO J.N. Evaluation of on-shore wind techno-economical potential in regions and islands. *Applied Energy.* 2014, 124, pp. 117–129, doi: `10.1016/j.apenergy.2014.02.050`.

[45] TAGLIAFERRI F., VIOLA I., FLAY R. Wind direction forecasting with artificial neural networks and support vector machines. *Ocean Engineering.* 2015, 97, pp. 65–73, doi: `10.1016/j.oceaneng.2014.12.026`.

[46] VERIKAS A., GUZAITIS J., GELZINIS A., BACAUSKIENE M. A general framework for designing a fuzzy rule-based classifier. *Knowledge and Information Systems.* 2011, 29(1), pp. 203–221, doi: `10.1007/s10115-010-0340-x`.

[47] WANG L.-X., MENDEL J. Generating fuzzy rules by learning from examples. *Systems, Man and Cybernetics, IEEE Transactions on.* 1992, 22(6), pp. 1414–1427, doi: `10.1109/21.199466`.

[48] WEEKES S., TOMLIN A., VOSPER S., SKEA A., GALLANI M., STANDEN J. Long-term wind resource assessment for small and medium-scale turbines using operational forecast data and measure–correlate–predict. *Renewable Energy.* 2015, 81, pp. 760–769, doi: `10.1016/j.renene.2015.03.066`.

[49] WIDÉN J., CARPMAN N., CASTELLUCCI V., LINGFORS D., OLAUSON J., REMOUIT F., BERGKVIST M., GRABBE M., WATERS R. Variability assessment and forecasting of renewables: A review for solar, wind, wave and tidal resources. *Renewable and Sustainable Energy Reviews.* 2015, 44, pp. 356–375, doi: `10.1016/j.rser.2014.12.019`.

[50] WITTEN I., FRANK E., HALL M. *Data Mining: Practical Machine Learning Tools and Techniques: Practical Machine Learning Tools and Techniques.* The Morgan Kaufmann Series in Data Management Systems. Elsevier Science, 2011, doi: `10.1016/c2009-0-19715-5`.

[51] YAMAGUCHI A., ISHIHARA T. Assessment of offshore wind energy potential using mesoscale model and geographic information system. *Renewable Energy.* 2014, 69, pp. 506–515, doi: `10.1016/j.renene.2014.02.024`.

[52] ZADEH L.A. Fuzzy sets. *Information and Control.* 1965, 8, pp. 338–353, doi: `10.1016/s0019-9958(65)90241-x`.