ARTICLE IN PRESS

Renewable Energy xxx (2015) 1–11

Contents lists available at ScienceDirect

Renewable Energy

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

Automatic identification of wind turbine models using evolutionary multiobjective optimization

William La Cava [a, *], Kourosh Danai [a], Lee Spector [b], Paul Fleming [c], Alan Wright [c], Matthew Lackner [a]

[a] *Dept. of Mechanical and Industrial Engineering, University of Massachusetts Amherst, Amherst, MA 01003, USA*
[b] *School of Cognitive Science, Hampshire College, Amherst, MA, USA*
[c] *National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, CO, USA*

ARTICLE INFO

Article history:
Received 8 April 2015
Received in revised form
25 September 2015
Accepted 28 September 2015
Available online xxx

Keywords:
Wind energy
System identification
Genetic programming
Multiobjective optimization

ABSTRACT

Modern industrial-scale wind turbines are nonlinear systems that operate in turbulent environments. As such, it is difficult to characterize their behavior accurately across a wide range of operating conditions using physically meaningful models. Customarily, the models derived from wind turbine data are in 'black box' format, lacking in both conciseness and intelligibility. To address these deficiencies, we use a recently developed symbolic regression method to identify models of a modern horizontal-axis wind turbine in symbolic form. The method uses evolutionary multiobjective optimization to produce succinct dynamic models from operational data while making minimal assumptions about the physical properties of the system. We compare the models produced by this method to models derived by other methods according to their estimation capacity and evaluate the trade-off between model intelligibility and accuracy. Several succinct models are found that predict wind turbine behavior as well as or better than more complex alternatives derived by other methods. We interpret the new models to show that they often contain intelligible estimates of real process physics.

1. Introduction

As wind energy grows across the globe and new offshore wind turbine installations encounter new operating environments, the models that inform the design and control of these multimillion-dollar machines become increasingly important. Typical multi-megawatt wind turbines exhibit nonlinear behavior and are subject to wind (and sometimes wave) disturbances that are often hard to estimate. These properties make the simulation of their dynamics not only challenging but also site-dependent, because of the influence of wind, wave, and foundation characteristics. Accordingly, the first-principles models of wind turbines, such as the one embedded in the aero-hydro-elastic simulation tool FAST [1], are prone to cumulative discrepancies between prediction and reality. These models are also computationally expensive to run because of their fairly comprehensive representation of wind turbine dynamics. Although the use of engineering models is fundamental to the structural design and loads analysis process, model-based controllers preferably rely on a customized model of the real system in the field, rather than a first-principles model that may miss key elements present in the real system [2].

As an alternative to potentially inaccurate and computationally expensive first-principles models, empirical models of wind turbines are obtained from experimental data to provide a customized representation of the wind turbine. These models are usually in the form of auto-regressive moving-average (ARMAX) models [2–5], neural networks [6], or fuzzy logic models [7], among others, to provide the structural flexibility for adapting the model according to the measured observations. Although these empirical models provide an effective means of estimation/prediction, they have the major drawback of lacking transparency about the physics of the process [8]. This lack of transparency obscures the knowledge of the process that is gained through their development. Ideally, the model should not only be accurate, but intelligible so that the user acquires the insight attained through the model's development. A well-formed model serves two purposes: (i) it improves knowledge of the underlying dynamics of the system; and (ii) it improves the ability of the wind turbine controller to extract power and minimize loads on the turbine.

* Corresponding author.
E-mail address: wlacava@umass.edu (W. La Cava).

http://dx.doi.org/10.1016/j.renene.2015.09.068
0960-1481/

Please cite this article in press as: W. La Cava, et al., Automatic identification of wind turbine models using evolutionary multiobjective optimization, Renewable Energy (2015), http://dx.doi.org/10.1016/j.renene.2015.09.068

In order to improve the intelligibility of adapted models, empirical models in the form of symbolic equations can be formulated by symbolic regression [9,10]. In symbolic regression, the process variables, inputs, and parameters (constants) are treated as symbols and integrated as blocks to form candidate model structures. Free of restrictions from the form (structure), the search is typically conducted by genetic programming (GP) for candidate models having the best-fit outputs to the measured observations [9]. However, in the absence of a presumed model structure and guided only by the prediction error (i.e., the difference between the modeled and measured outputs), symbolic regression often yields illegible, albeit accurate, models that do not convey any of the physics of the process. The method proposed for modeling here safeguards against this potential shortcoming by two innovations. First, it uses a novel GP method known as epigenetic linear genetic programming (ELGP) that combines the flexibility of stack-based GP representations with an epigenetic encoding to allow for topological search of the candidate model structures, leading to less complex and more accurate results than traditional GP [11,12]. Second, it uses an evolutionary multi-objective optimization (EMO) framework [13] that includes the complexity of the model as an objective in order to yield accurate models that are as intelligible as possible.

In this paper we evaluate the applicability of the proposed ELGP method in identifying wind turbine models based on experimental data collected in normal closed-loop operation from the three-bladed Controls and Advanced Research Turbine (CART3), a turbine maintained by the National Renewable Energy Laboratory (NREL). The paper is organized as follows. First, we present a brief overview of wind turbine mechanics. We then review previous system identification work. Next, the problem formulation as sought by multiobjective optimization is presented, followed by a description of the proposed ELGP method. We then detail the wind turbine identification procedure and analyze results pertaining to local and global models of the wind turbine. The paper concludes with a discussion of the intelligibility of the identified models as they inform the physics of the process.

2. Wind turbine mechanics

Identification of wind turbine models is a difficult undertaking because of the many layers of nonlinearity governing their behavior. Moreover, modern horizontal-axis wind turbines (HAWTs) are controlled using variable-speed and variable-blade pitch operation, further complicating the dynamics. Consider for instance the steady-state aerodynamic rotor torque (Q_R) and thrust (T_R) generated by the rotor operating in freestream wind speed V, defined by:

$$Q_R = \frac{1}{2}\rho\pi R^3 C_q(\lambda,\beta)V^2 \tag{1}$$

$$T_R = \frac{1}{2}\rho\pi R^2 C_T(\lambda,\beta)V^2 \tag{2}$$

where the tip speed ratio $\lambda = \Omega R/V$ relates the rotor speed Ω to the wind speed V, ρ is the air density, R is the rotor radius, β is the pitch angle of the blades (assumed pitching collectively). C_q and C_T are the torque and thrust coefficients, respectively, defining the corresponding generated lift as functions of λ and β. The overall C_q is a function of local aerofoil drag and lift coefficients C_d and C_l, local incidence angle with the wind, ϕ, and local tip speed ratio λ_r, defined by the strip theory calculation of C_q, as:

$$C_q = \left(8\big/\lambda^3\right)\int_{\lambda_h}^{\lambda} \sin^2\phi(\cos\phi - \lambda_r\sin\phi)(\sin\phi + \lambda_r\cos\phi)\left[1 - \frac{C_d}{C_l}\cot\phi\right]\lambda_r^2 d\lambda_r \tag{3}$$

Because it is difficult to obtain the lift and drag coefficients at each position along the blade due to small inconsistencies in fabrication and local shape deflections, they are often estimated empirically [14]. The inaccuracy of estimated nonlinear coefficient surfaces C_q and C_T, compounded with the measurement uncertainty and stochasticity of V, impedes prediction of the aerodynamic torque and thrust response of the system.

Control actions are limited to actuating the collective pitch β, the generator torque T_G, and the yaw angle ψ. Because of the highly nonlinear nature of the wind turbine behavior, a pitch action of the same magnitude may result in very different aerodynamic forces depending on the instantaneous wind speed and rotor speed, requiring the employment of gain scheduling for pitch control [15]. In addition to aerodynamic nonlinearities, the turbine has low-frequency periodic excitations induced by the rotating blades at once-per-revolution (1P) and thrice-per-revolution (3P) that are normally within the same frequency range as the fore-aft (FA) and side–side (SS) natural frequencies of the tower, requiring the added provision of avoiding dynamic coupling between these excitations and that of the pitch control that affects Ω. Similarly, the first mode of the wind turbine drivetrain can be excited by the generator torque commands, so the generator control must account for this fundamental design objective as well. From the above anecdotes it follows that an accurate model of the wind turbine is essential for designing a reliable controller. This need for model accuracy motivates data-based modeling approaches that can account for turbine-specific observations and provide confident estimates of wind turbine behavior.

3. Related work

Most system identification attempts at modeling wind turbines have focused on producing linear time-invariant (LTI) models via ARMAX models [2,4] or modified forms of closed-loop subspace identification (SSID) [3,5]. Although LTI models seem to be effective in characterizing simulated wind turbine behavior at specific operating wind speeds [2,4], they provide only localized representation. As a remedy, SSID methods have been extended to account for the time-varying, nonlinear dynamics of the wind turbines to form global models. For example, Van der Veen [5] showed that Wiener and Hammerstein systems could be used to identify global wind turbine dynamics by providing the model with the nonlinear aerodynamic torque and thrust relations (Eqs. (1) and (2)), as well as the surface functions for (C_p) and (C_T) that vary with the tip speed ratio λ and pitch angle β. This approach, however, requires good knowledge of these two surface functions, which rely on first principles. Another approach to global modeling associates the nonlinearities with the azimuth angle of the rotor and uses a linear parameter-varying (LPV) model to conduct closed-loop identification of the wind turbine dynamics [3]. In this case, the dynamics of the turbine are assumed to vary periodically, so the matrices of the state space model are defined in terms of the azimuth position of the rotor. This approach provides good predictions of the hub moments at the rotor and tower top motion.

The above approaches, albeit in 'black-box' form, are attractive because of their incorporation of expert knowledge in modeling some of the nonlinearities and for their accommodation of control