



Jul 12th, 9:50 AM - 10:10 AM

## Pre-processing SCADA data from wind turbines as a previous stage to the application of prognosis algorithms

Alejandro Blanco

*Data and Signal Processing Research Group, U Science Tech, University of Vic - Central University of Catalonia, Smartive-ITESTIT SL, Parc Audiovisual de Terrassa, pere.marti@uvic.cat*

Juan José Cárdenas

*Smartive-ITESTIT SL, Parc Audiovisual de Terrassa*

Isaac Justicia

*Smartive-ITESTIT SL, Parc Audiovisual de Terrassa*

Jordi Solé-Casals

*Data and Signal Processing Research Group, U Science Tech, University of Vic - Central University of Catalonia, jordi.sole@uvic.cat*

Pere Marti-Puig

*University Data and Signal Processing Research Group, U Science Tech, University of Vic - Central University of Catalonia, pere.marti@uvic.cat*

Follow this and additional works at: <https://scholarsarchive.byu.edu/iemssconference>



Part of the [Civil Engineering Commons](#), [Data Storage Systems Commons](#), [Environmental Engineering Commons](#), [Hydraulic Engineering Commons](#), and the [Other Civil and Environmental Engineering Commons](#)

---

Blanco, Alejandro; Cárdenas, Juan José; Justicia, Isaac; Solé-Casals, Jordi; and Marti-Puig, Pere, "Pre-processing SCADA data from wind turbines as a previous stage to the application of prognosis algorithms" (2016). *International Congress on Environmental Modelling and Software*. 4.  
<https://scholarsarchive.byu.edu/iemssconference/2016/Stream-C/4>

This Event is brought to you for free and open access by the Civil and Environmental Engineering at BYU ScholarsArchive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU ScholarsArchive. For more information, please contact [scholarsarchive@byu.edu](mailto:scholarsarchive@byu.edu), [ellen\\_amatangelo@byu.edu](mailto:ellen_amatangelo@byu.edu).

# Pre-processing SCADA data from wind turbines as a previous stage to the application of prognosis algorithms

Alejandro Blanco<sup>1,2</sup>, Juan José Cárdenas<sup>2</sup>, Isaac Justicia<sup>2</sup>, Jordi Solé-Casals<sup>1</sup>, Pere Martí-Puig<sup>1</sup>  
[pere.marti@uvic.cat](mailto:pere.marti@uvic.cat), [jordi.sole@uvic.cat](mailto:jordi.sole@uvic.cat)

<sup>1</sup>Data and Signal Processing Research Group, U Science Tech, University of Vic - Central University of Catalonia (UVic-UCC), Perot Rocaguinarda 17, 08500 Vic, Spain

<sup>2</sup>Smartive-ITESTIT SL, Parc Audiovisual de Terrassa, Carretera BV-1274, Km1, 08226 Terrassa, Barcelona, Spain

**Abstract:** Modern wind turbines in operation today record more than 200 analogous variables at intervals of 5 to 10 minutes by their SCADA (Supervisory Control and Data Acquisition) system. The SCADA data often comprise temperature values from a variety of measurement positions in the turbine, pressure data, electrical quantities such as line currents and voltages or pitch-motor currents, tower vibration, etc. The Wind Turbine's data pre-processing is a task that faces diverse difficulties. The main are the atypical values produced by a miss configuration of the SCADA, the information lost about the alarms states (activations/deactivations) and the information loss when the connection is lost between SCADA and our data capture server in some update period. This document covers the communication with an OPC (Open Platform Communication) server that is a widely used in Wind Turbine SCADAs, which implements a publisher-subscriber paradigm. Because of this paradigm we are applying two corrections over the data, the first is online and the second offline. The first correction is taking into account each wind turbine data notification frequency so the OPC must notify data variables with a timestamp difference of about 10 minutes, if not, then a correction is made applying a linear regression once the next update is received. This will create a calculated value for a missing field each time that it is lost. The second correction is offline, over the already received data, which could have a second and not less important problem that covers the wind turbine's alarm. This step is offline because the data variable values are required, so this has to be made once the empty data variable blocks are filled.

**Keywords:** Wind farms; SCADA data; Logically inconsistent data management; Data cleaning.