Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 78 (2015) 447 – 452

Energy Procedia

6th International Building Physics Conference, IBPC 2015

# Data-driven performance evaluation of ventilated photovoltaic double-skin facades in the built environment

Leon Gaillard[a,*], Guillaume Ruedin[b], Stéphanie Giroux-Julien[b], Marc Plantevit[c], Mehdi Kaytoue[d], Syamimi Saadon[a], Christophe Ménézo[a], Jean-François Boulicaut[d]

[a]*INSA-Lyon, CETHIL UMR 5008, INSA-EDF Chair "Habitats and Energy Innovations", Villeurbanne 69621, France*
[b]*University of Lyon 1, CETHIL UMR 5008, 69621 Villeurbanne, France*
[c]*University of Lyon, CNRS, University of Lyon 1, LIRIS, UMR5205, 69621 Villeurbanne, France*
[d]*University of Lyon, CNRS, INSA-Lyon, LIRIS, UMR5205, 69621 Villeurbanne, France*

**Abstract**

In this paper we present a collaborative data science project bringing together experts from the fields of physical and information sciences to tackle to challenges and opportunities of analysing large and rich datasets obtained from the monitoring of building integrated PV systems operating in the built environment. We present data mining analysis techniques to classify data according to environmental conditions and system performance, to distinguish between nominal and anomalous behaviour, and to identify instrumentation faults. These methods were implemented using data from the RESSOURCES project to construct evaluate the performance of a PV envelope, and to validate a simplified physical model to predict thermal and aerodynamic behaviour.

Peer-review under responsibility of the CENTRO CONGRESSI INTERNAZIONALE SRL

*Keywords:* Full-scale; PV double-skin façade; Monitoring; Data mining

## 1. Introduction

Europe strives for a decrease in green house gas emissions by a factor of 4 by 2050 relative to 1990 levels, and a reduction of 20% of energy consumption by 2020 whilst raising the contribution of renewable energy to 20%. The

* Corresponding author. Tel.: +33-4-7243-8813; fax: +33-4-72438811.
*E-mail address:* leon.gaillard@insa-lyon.fr


Peer-review under responsibility of the CENTRO CONGRESSI INTERNAZIONALE SRL
doi:10.1016/j.egypro.2015.11.694

building sector accounts for a large portion of emissions and energy consumption, and this is reflected in national regulations such as in France where by 2020 all new buildings should be net energy positive (BEPOS). Building integrated photovoltaic (BIPV) systems represent an interesting solution, which may include active or passive cooling techniques to control PV cell temperature whilst improving building thermal performance. For the large-scale deployment of PV envelopes, several technological hurdles must be overcome. Such components are often multifunction in nature, and must meet criteria including electrical and thermal production, but also daylighting, sound insulation and aesthetics. The complexity of the urban environment, heterogeneous and fluctuating in nature, also presents a major challenge to BIPV architectures [1]. Full scale experimentation under real conditions is essential in order to validate the concepts and predict the performance over the lifetime of the system [2].

As part of the RESSOURCES project, realistic prototype PV envelopes were designed and constructed on real buildings in France to study their behavior under natural ventilation operating conditions: two for individual houses and one for an office building. Each facade comprised tinted double glass PV modules of variable cell configurations plus fully transparent tinted panels. This paper makes use of results from the HBS Technal prototype, presented in figure 1. The prototype, of vertical pleated geometry, was installed at Toulouse on the W.N.W wall of a three-storey, open-plan office building. Measuring 7.4 m in height, 4.0 m in width, and with an airgap of 60-80cm, the prototype covers the first two floors of the building. The PV components are divided into a stack of three separate arrays, hereafter referred to as ‘blocs’, connected to a constant load, with a total rated power of 1.2 kW. The data used in this paper span a three month period, 21/06/2012 – 20/09/2012. Previous results [3] have shown that despite the level of geometrical complexity of the HBS prototype and the variability of the environment, the behaviour of the system follows regular and periodic trends over timescales of one day and one year.

The data retrieved from this and other PV installations contain a wealth of information that could be further exploited to improve the understanding of the technology, but novel methods are needed to analyze sizeable datasets comprising a large number of parameters, a multiplicity of spatial and temporal scales, and a level of instrumentation that is typically too coarse to reveal the detailed interaction of phenomena such as wind and shading effects. To this end, the CETHIL and LIRIS laboratories initiated with the INSA-Lyon/EDF research chair a collaborative data science project in 2012, initially within the framework of the AMADOUER project (CNRS MASTODONS programme), and subsequently by the INSA-Lyon BQR project SOLSTICE. To date this work has focused upon well-established data mining techniques and the creation of data exploration tools.

**Nomenclature**

| | |
|---|---|
| $G_{dir}$ | direct solar radiation (pyrheliometer), ($Wm^{-2}$) |
| $G_i$ | in-plane total solar radiation (pyranometer), ($Wm^{-2}$) |
| $G_{iT}$ | total horizontal radiation (rooftop), ($Wm^{-2}$) |
| $G_{ref}$ | radiation intensity at reference conditions, 1000 ($Wm^{-2}$) |
| k | number of clusters, (-) |
| $P_c$ | rated electrical power, (W) |
| $P_{dc}$ | instantaneous produced electrical power, (W) |
| $\eta$ | instantaneous performance ratio (-) |

## 2. Data mining methods

The analysis was developed with the aid of the open source software Knime [4]. A workflow was first assembled to manage the interface with the MySQL database repository, and handle the tasks of spatial and temporal aggregation, and filtering, that were previously undertaken using Scilab. Several data mining techniques were then appended to this analysis chain, and in particular the k-means clustering algorithm. Rarely used in scientific research of solar systems in urban environments, these methods have been demonstrated effective in other domains such as life sciences research (gene expression studies) for many years [5]. The k-means method is one of the oldest and most widely used clustering routines [6, 7]. The algorithm regroups data into k partitions or clusters in n dimensions, which for the current application correspond to indicators of system behavior or environmental factors. The only