

Article

# Hygrothermal Dynamic and Mould Growth Risk Predictions for Concrete Tiles by Using Least Squares Support Vector Machines

**Roberto Zanetti Freire** <sup>1,\*</sup> , **Gerson Henrique dos Santos** <sup>2</sup> and **Leandro dos Santos Coelho** <sup>1,3</sup>

<sup>1</sup> Industrial and Systems Engineering Graduate Program—PPGEPS, Polytechnic School—EP, Pontifical Catholic University of Paraná—PUCPR, Rua Imaculada Conceição, 1155, Curitiba 80215-901, Brazil; leandro.coelho@pucpr.br

<sup>2</sup> Department of Mechanical Engineering, Federal Technological University of Paraná—UTFPR, Av. Monteiro Lobato, Km 04, Ponta Grossa 84016-210, Brazil; gsantos@utfpr.edu.br

<sup>3</sup> Department of Electrical Engineering, Electrical Engineering Graduate Program—PPGEE, Federal University of Paraná—UFPR, Av. Cel. Francisco H. dos Santos, 210, Curitiba 81531-970, Brazil

\* Correspondence: roberto.freire@pucpr.br; Tel.: +55-41-3271-1330

Received: 3 July 2017; Accepted: 24 July 2017; Published: 26 July 2017

**Abstract:** The hygrothermal analysis of roofs is relevant due to the large areas exposed to a wide range of weather conditions, these directly affecting the energy performance and thermal comfort of buildings. However, after a long life service, the solar absorptivity coatings of roofs can be altered by mould accumulation. Based on two well established mathematical models, one that adopts driving potentials to calculate temperature, moist air pressure and water vapor pressure gradients, and the other to estimate the mould growth risk on surfaces, this research introduces an approach to predict mould growth considering a reduced computational effort and simulation time. By adopting multiple MISO (Multiple-Input, Single-Output) Nonlinear AutoRegressive with eXogenous inputs (NARX) models, a machine learning technique known as Least Squares Support Vector Machines (LS-SVM), a maximum margin model based on structural risk minimization, was used to predict vapor flux, sensible heat flux, latent heat flux, and mould growth risk on roof surfaces. The proposed model was validated in terms of the Multiple Correlation Coefficient ( $R^2$ ), Mean Square Error (MSE) and Mean Absolute Error (MAE) performance indices considering as input the weather file from Curitiba city—Brazil, showing consistent precision when compared to the results of a validated numerical model.

**Keywords:** support vector machines; machine learning; system identification; concrete tiles; hygrothermal performance; mould growth

## 1. Introduction

Most recent studies in building physics focus on energy savings and/or thermal comfort. Taking into account that a considerable amount of energy attributed to buildings is used to provide thermal comfort, and that in modern societies people spend over 90% of their time indoors [1], buildings became responsible for a considerable amount of energy demand worldwide [2,3]. In Brazil, residential and commercial buildings are responsible for almost 45% of the country's energy demand [4], which progressively motivates energy conservation studies for promoting building energy efficiency.

The design of energy efficient buildings depends on a significant amount of variables, the required indoor climate conditions, internal gains, the outdoor prevailing climatic conditions, and the choice of building construction materials and insulation among other variables. Accurate methods to predict hygrothermal performance of building envelopes became necessary, and software

capable of performing whole-building energy simulation and analysis became available [5–9]. However, the complexity of considering moisture presence in this type of simulation requires both time-consuming and complex computational codes [10–13]. However, an alternative way to evaluate the influences of coupled heat and moisture transfer in building can be performed by adopting computational intelligence and machine learning techniques [14]. Moreover, this type of technology can be also used in the analysis of building energy demand and energy savings [15–17].

In warm climates, during clear sky conditions, up to about  $1 \text{ kW/m}^2$  of solar radiation can be incident on a roof surface, and between 20% and 95% of this radiation is typically absorbed. The roof color that is apparent from the reflected visible part of the solar radiation usually gives an indication of the value of solar absorption [18]. In cases of high radiation incidence, the use of proper insulating or higher roof solar reflectance can reduce the solar energy absorbed by the roof, providing economy in the usage of air conditioning in warmer climate countries. Additionally, in cold weather, it can avoid heat losses, by increasing the energy efficiency of the whole building. Due to these facts, the application of different types of thermal insulation and special building materials significantly increased in recent years [19] and became a valuable strategy for making buildings more sustainable [20–22].

In this context, reflective roof or insulation coatings have been utilized for increasing the energy savings potential of building envelopes, and several studies can be found in the literature, such as [18,19,23–27]. Nevertheless, due to the low cost, concrete or ceramic tiles are widely used in roofs in Brazil. Those tiles enable mould or algae to grow and normally no paint layer or impermeable films (Figure 1a) are provided. In asphalt roofing shingles found widely in European and North American buildings, this phenomenon is commonly observed (Figure 1b). The mould growth decreases the tile durability, worsens the aesthetic appearance of buildings and increases the solar absorptivity.



**Figure 1.** (a) Mould growth on concrete tiles; (b) Mould growth on asphalt roofing shingles.

Taking into account that the cost of both heating and cooling in buildings is directly affected by roof performances, this article presents an approach combining a numerical computational code and an artificial intelligence method in order to predict the hygrothermal behavior of building roofs, focusing on the evaluation of mould growth risk on these structures. The main idea is to perform a nonlinear system identification by using data obtained from the results of the numerical model. In this manner, the main objectives are to reduce computational costs and to provide consistent approximation when compared to an already validated numerical model.

SVMs (Support Vector Machines) have already proven to be a promising approach in nonlinear identification and modelling. This technique was developed based on statistical learning (details in [28]) and was originally created to solve classification problems. SVM refers to a kernel-based method, similar to artificial neural network (ANN) models, which constitute an approximate implementation of the structural risk minimization principle [29]. Considering structures called nuclei (kernels), SVMs go beyond the hyperplanes generated initially, been widely applied in classification [30–32] and nonlinear regression [33–35] areas, mapping the input data in a space with characteristics of high dimensionality.

A variation of SVMs, known as LS-SVMs (Least Squares Support Vector Machines), adopted in this work, was proposed, evaluated, and compared to the classical version of SVMs [36,37] for a regression/identification task. It involves the equality constraints only, where the solution is obtained by solving a system of linear equations. In the application presented in this work, both sensible and latent heat flows, vapor flow and mould growth risk on concrete tiles, which are constantly adopted in Brazilian buildings, were predicted using the LS-SVM approach considering Multiple-Input, Single-Output (MISO) models. The technique was selected by considering the learning capability of the SVM, especially related to nonlinearities available in the system caused by moisture presence.

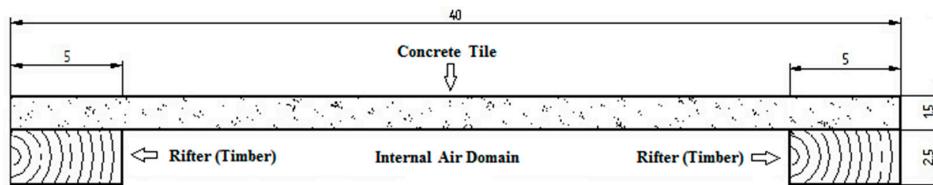
The next section of this article presents the data acquisition procedures, where a validated mathematical model was adopted to generate data for the system identification procedures. The data set analysis is presented in Section 3, followed by a detailed description of the LS-SVM technique in Section 4. Section 5 presents and discusses both training and validation results found by the LS-SVM. Finally, Section 6 addresses the conclusions and future work of this type of research.

## 2. Data Acquisition Procedures

This section presents the mathematical model adopted for data acquisition. In order to obtain a consistent data set, simulation procedures indicated in this section were used for the system identification procedures presented on Section 4 of the present study.

### 2.1. Mathematical Model

The model for the porous media domain (Figure 2) has been elaborated considering the differential governing equations for moisture, air and energy balances [38]. The transient terms of each governing equation have been written in terms of the driving potentials to take more advantage of the Multitridiagonal-Matrix Algorithm (MTDMA) solution algorithm [39].



**Figure 2.** Physical domain and dimensions (cm) of the roof.

### 2.2. Porous Element Domain

The model is based on averages taken over a representative elementary volume (REV). The moisture transport has been divided into liquid and vapor flows as shown in Equation (1):

$$\mathbf{j} = \mathbf{j}_l + \mathbf{j}_v, \quad (1)$$

where  $\mathbf{j}$  is the density of moisture flow rate ( $\text{kg}/\text{m}^2 \cdot \text{s}$ ),  $\mathbf{j}_l$ , the density of liquid flow rate ( $\text{kg}/\text{m}^2 \cdot \text{s}$ ) and,  $\mathbf{j}_v$ , the density of vapor flow rate ( $\text{kg}/\text{m}^2 \cdot \text{s}$ ). The liquid transport calculation is based on the Darcy equation:

$$\mathbf{j}_l = K(\nabla P_{suc} - \rho_l \mathbf{g}), \quad (2)$$

where  $K$  is the liquid water permeability ( $\text{s}$ ),  $P_{suc}$ , the suction pressure ( $\text{Pa}$ ),  $\rho_l$ , the liquid water density ( $\text{kg}/\text{m}^3$ ) and  $\mathbf{g}$  the gravity ( $\text{m}/\text{s}^2$ ).

The capillary suction pressure can be written as a function of temperature and moisture content in the following form:

$$\nabla P_{suc} = \frac{\partial P_{suc}}{\partial T} \nabla T + \frac{\partial P_{suc}}{\partial P_v} \nabla P_v. \quad (3)$$

Similar to the liquid flow, the vapor flow is calculated from the Fick equation considering effects of both vapor pressure and air pressure driving potentials:

$$\mathbf{j}_v = \delta_v \nabla P_v - \rho_v \frac{kk_{rg}}{\mu_g} \nabla P_g, \quad (4)$$

where  $\delta_v$  is the vapor diffusive permeability ( $\text{s}$ ),  $P_v$ , the partial vapor pressure ( $\text{Pa}$ ),  $\rho_v$ , the vapor density ( $\text{kg}/\text{m}^3$ ),  $k$ , the absolute permeability ( $\text{m}^2$ ),  $k_{rg}$ , the gas relative permeability,  $\mu_g$ , the dynamic viscosity ( $\text{Pa.s}$ ) and,  $P_g$ , the gas pressure. The first term after the equality represents vapor diffusion, and the second one vapor advection. The water mass conservation equation can be described as:

$$\frac{\partial w_m}{\partial t} \nabla \cdot \mathbf{j}, \quad (5)$$

where  $w_m$  is the moisture content ( $\text{kg}/\text{m}^3$ ).

This moisture content conservation equation—Equation (5)—can be written in terms of the three driving potentials as:

$$\frac{\partial w_m}{\partial \phi} \frac{\partial \phi}{\partial P_v} \frac{\partial P_v}{\partial t} + \frac{\partial w_m}{\partial \phi} \frac{\partial \phi}{\partial T} \frac{\partial T}{\partial t} = \nabla \cdot \left[ -K \frac{\partial P_{suc}}{\partial T} \nabla T - \left( K \frac{\partial P_{suc}}{\partial P_v} - \delta_v \right) \nabla P_v + \rho_v \frac{kk_{rg}}{\mu_g} \nabla P_g + K \rho_l \mathbf{g} \right]. \quad (6)$$

In the proposal model, the air transport is individually considered through the dry-air mass balance. In this way, the dry-air conservation equation can be expressed as:

$$\frac{\partial \rho_a}{\partial t} = \nabla \cdot \mathbf{j}_a, \quad (7)$$

with the air flow calculated by the following expression:

$$\mathbf{j}_a = \delta_v \nabla P_v - \rho_a \frac{kk_{rg}}{\mu_g} \nabla P_g, \quad (8)$$

where  $\rho_a$  is the dry-air density ( $\text{kg}/\text{m}^3$ ),  $\mathbf{j}_a$ , the dry-air flow rate density ( $\text{kg}/\text{m}^2 \cdot \text{s}$ ) and,  $P_g$ , the gas pressure (dry air pressure plus vapor pressure, in  $\text{Pa}$ ). The first term after the equality represents air diffusion, and the second one air convection. Therefore, the dry air transport can be described as a function of the partial gas and vapor pressure driving potentials so that the air balance can be written as:

$$\frac{\partial \rho_a}{\partial P_g} \frac{\partial P_g}{\partial t} + \frac{\partial \rho_a}{\partial P_v} \frac{\partial P_v}{\partial t} + \frac{\partial \rho_a}{\partial T} \frac{\partial T}{\partial t} = \nabla \cdot \left( -\delta_v \nabla P_v + \rho_a \frac{kk_{rg}}{\mu_g} \nabla P_g \right). \quad (9)$$

Due to the presence of low temperature gradients, heat transfer has been attributed to both conductive and convective effects only. The conductive transport is calculated by Fourier's law:

$$\mathbf{q}_{\text{cond}} = -\lambda \nabla T, \quad (10)$$

while the convective transport can be written as:

$$\mathbf{q}_{\text{conv}} = \mathbf{j}_l c_{pl} T + \mathbf{j}_a c_{pa} T + \mathbf{j}_v L + \mathbf{j}_v c_{pv} T, \quad (11)$$

where  $\lambda$  is the thermal conductivity ( $\text{W}/(\text{m.K})$ ),  $c_{pa}$ , the dry-air specific heat at constant pressure ( $\text{J}/(\text{kg.K})$ ),  $c_{pl}$ , the liquid water specific heat ( $\text{J}/(\text{kg.K})$ ),  $c_{pv}$ , the vapor specific heat at constant pressure ( $\text{J}/(\text{kg.K})$ ) and,  $L$ , the vaporization latent heat ( $\text{J/kg}$ ). The terms after the equality represent liquid flow, dry air flow, phase change, and vapor flow, respectively.

The energy balance equation based on the first law of thermodynamics can be written in this case as:

$$c_m \rho_0 \frac{\partial T}{\partial t} = -\nabla \cdot \mathbf{q}, \quad (12)$$

where  $c_m$  is the specific heat of the structure ( $J/(kg.K)$ ) and  $\rho_0$ , the dry-basis material density ( $kg/m^3$ ). In this way, assuming  $0^\circ C$  as the reference temperature, the energy conservation equation can be rewritten in terms of the three driving potentials as:

$$c_m \rho_0 \frac{\partial T}{\partial t} = -\nabla \cdot \left\{ \begin{array}{l} \left[ \lambda - K \frac{\partial P_{suc}}{\partial T} c_{pl} T \right] \nabla T - \left[ K \frac{\partial P_{suc}}{\partial P_v} c_{pl} T + -\delta_v (L + c_{pv} T) \right] \nabla P_v + \\ \left[ \rho_a \frac{k k_{rg}}{\mu_g} c_{pa} T + \rho_v \frac{k k_{rg}}{\mu_g} (L + c_{pv} T) \right] \nabla P_g + K \rho_l c_{pl} T \mathbf{g} \end{array} \right\} \quad (13)$$

### 2.3. Mould Growth Model

The mould growth prediction has been verified through the model presented in [40]. The mould has been measured applying an existing standard index based on the visual appearance of the surface under study. This model was initially presented in [41] considering the analysis in wood material, and improved in a project carried out at VTT (Technical Research Center of Finland) and Tampere University of Technology, where experiments included large sets of steady-state and dynamic laboratory experiments for typical building materials. This mould index assumes the values presented in Table 1.

**Table 1.** Mould growth index ( $M$ ) for experiments and modeling [40]. Reproduced with permission from FEUP, 2011.

$M$	Description of the Growth Rate
0	No growth
1	Small amounts of mould on surface (microscopic), initial stages of local growth
2	Several local mould growth colonies on surface (microscopic)
3	Visual findings of mould on surface, <10% coverage, or, <50% coverage of mould (microscopic)
4	Visual findings of mould on surface, 10–50% coverage, or, >50% coverage of mould (microscopic)
5	Plenty of growth on surface, >50% coverage (visual)
6	Heavy and tight growth, coverage about 100%

The index  $M$  is presumed to increase linearly in time:

$$\frac{dM}{dt} = \frac{1}{7e^{(-0.68\ln T - 13.9\ln\phi + 0.14W - 0.33SQ + 66.02)}} k_1 k_2, \quad (14)$$

where  $k_1$  represents the intensity of growth,  $k_2$  is the moderation of the growth intensity when the mould index ( $M$ ) level approaches the maximum peak value,  $W$  is the timber species (0 = pine and 1 = spruce) and  $SQ$  is the term for surface quality ( $SQ = 0$  for sawn surface,  $SQ = 1$  for kiln dried quality),  $T$  is the temperature ( $^\circ C$ ) and  $\phi$ , the relative humidity (%). The terms  $k_1$  and  $k_2$  are modified for other materials than wood and  $SQ = 0$  is used. The idea is to compare the mould growth of other materials to that of the pine sapwood (original work). More details can be seen in [40,42].

### 3. Simulation Using the Mathematical Model

This section describes the simulation procedures and the data set obtained and adopted for the system identification using LS-SVM technique. The next subsection describes the roof composition and parameters followed by explanations about the data set.

#### 3.1. Simulation Procedures

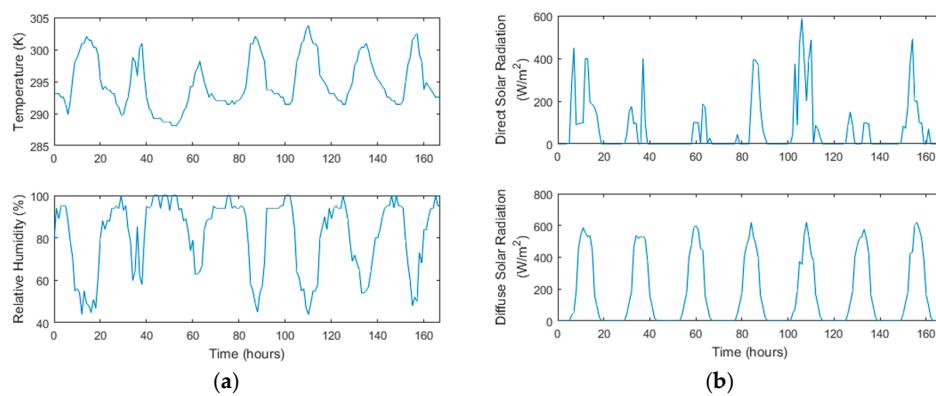
The roof analyzed is composed of concrete tile (1.5 cm) and two rifters (2.5 cm) as shown in Figure 3. The hygrothermal properties have been obtained from [43] for concrete and from IEA Annex 14

Report [44] for timber (rifiers). The internal emissivity was considered equal to 0.9 for both concrete tile and timber roofs, and the solar absorptivity equal to 0.6 has been adopted.

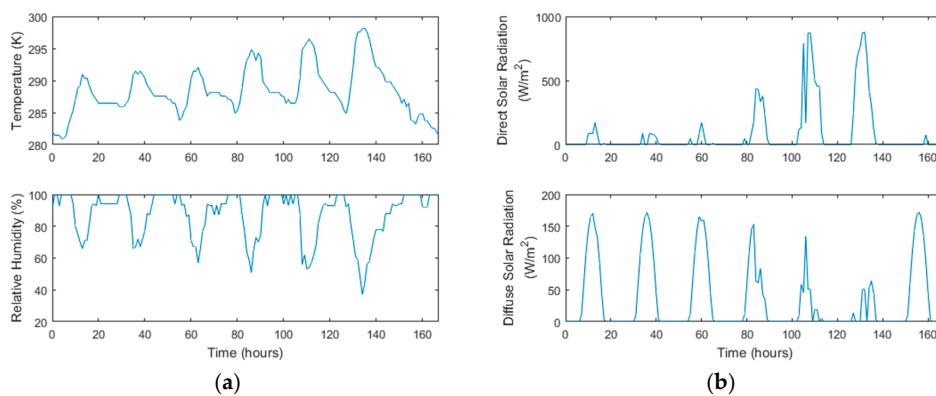
In order to evaluate extremal conditions, where higher incidence of solar radiation is desired, a flat roof at a horizontal angle has been considered during simulations. In terms of moisture transport, this inclination also indicates extremal transport level as the flow is in the opposite direction of the roof normal. Additionally, thermal insulation is not commonly adopted in many regions of Brazil due to elevated costs. The building roof configuration presented in this work is typically adopted residential buildings.

A regular 2-D mesh ( $2.5 \text{ mm}^2$ ) for the discretization using the finite-volume method and a 30 s constant time step have been applied for all simulations. The computational code was implemented in C programming language in order to enable dynamic memory allocation, and the sample time for data generation was set to 6 h. Due to high computational time consumption, the simulation using the hygrothermal model was performed for almost 1 year and 8 months, when the index  $M = 3$  was reached.

A temperature of  $24^\circ\text{C}$  and a relative humidity of 50% (conditioned environment) were considered for indoor conditions. The outdoor climate conditions were represented by the TRY (Test Reference Year) weather data for the city of Curitiba—Brazil, which can be found in [45], and are presented in Figure 3 for the first week of January (summer period), and in Figure 4 for the first week of July (winter period). Constant convective heat transfer coefficients of 3 and  $10 \text{ W}/(\text{m}^2 \cdot \text{K})$  have been used at the internal and external surfaces. The external and internal convective water vapor transfer coefficients are calculated by Lewis's relation for each control volume. The other surfaces were considered adiabatic and impermeable. The sky temperature correlation presented in [46] has been adopted in this work. Gas (moist air) pressure has been considered constant at all surfaces.



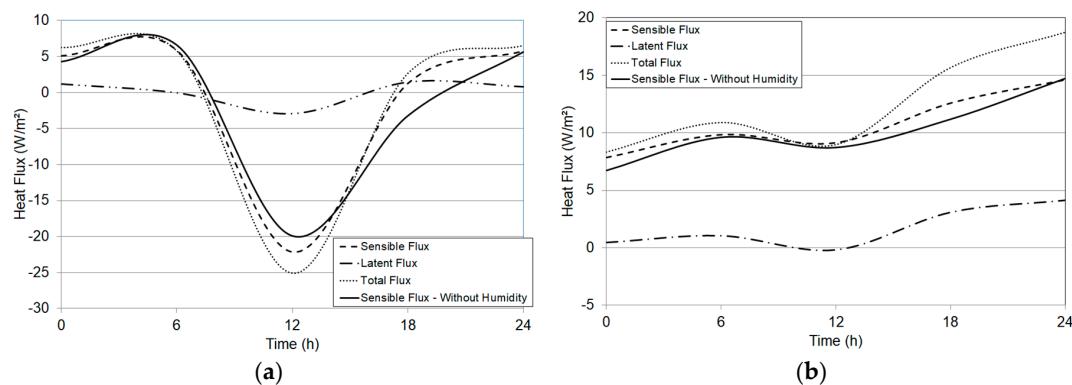
**Figure 3.** Weather data for the first week of January (summer period) for Curitiba, Brazil: (a) temperature and relative humidity; (b) direct and diffuse solar radiation.



**Figure 4.** Weather data for the first week of July (winter period) for Curitiba, Brazil: (a) temperature and relative humidity; (b) direct and diffuse solar radiation.

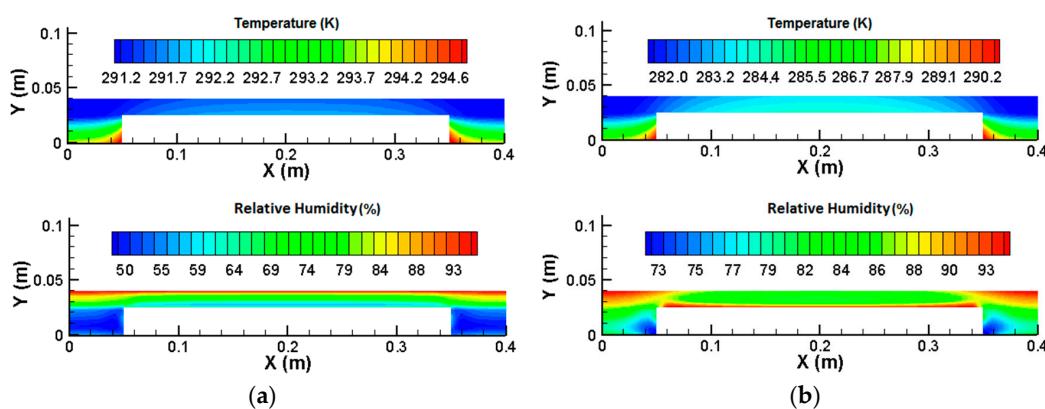
### 3.2. Data Analysis

Heat fluxes at the internal roof surfaces are presented in Figure 5a. The mass transport effect is verified when a roof was simulated without the moisture content and air conservation equations. As observed in Figure 5, the latent effect was small when compared with the sensible heat flux. This fact is attributed to the analyzed period (January). This effect is increased in the winter (Figure 5b), where the sensible heat flux losses dramatically increase its magnitude and the Sun does not dry out the roof as it does in the summer. However, a difference of 20% in the heat flux is reported for the peak values when the mass transport is considered.



**Figure 5.** Heat fluxes at the internal roof surface: (a) on 7 January (summer period); (b) on 7 July (winter period).

Figure 6 presents the temperature and relative humidity profiles within the roof at midnight on 7 January (Figure 6a) and 7 July (Figure 6b). In Figure 6a the low temperature at the external surface increases the relative humidity and, consequently, the mould index in this period, as described by Equation (14). Although the indoor latent heat flux is small when compared to the sensible flux, the mass transport can play an important role in the heat exchanged at the external surface, mainly, in the daytime, when the adsorbed moisture at nighttime is released, decreasing the concrete tile temperature due to the outward evaporation.



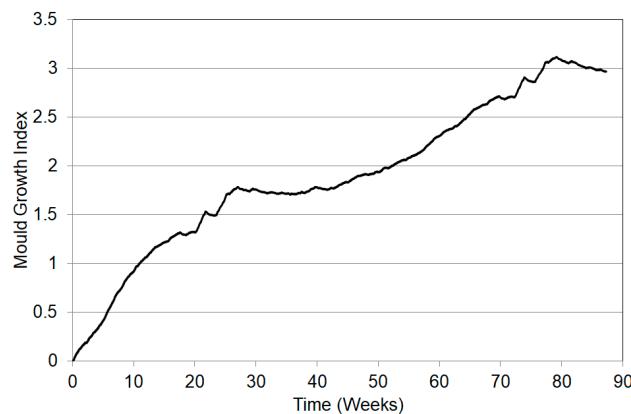
**Figure 6.** Temperature and relative humidity profiles in the roof at midnight: (a) on 7 January; (b) on 7 July.

In order to confirm the possibility of mould growth on concrete tile, the mould index evolution is reported in Figure 7 as the results of the simulation using the numerical model presented in Section 3. Some mould growth can be detected visually after 76 weeks (Curitiba climate). A small decrease in the mould index is observed after the winter period. This work has been limited to 85 weeks due to

the high computational time required and to the fact that according to the mould growth risk model (Section 2.3), mould would visually appear in less than 85 weeks.

Beyond damage and the roof appearance, the mould growth can change the solar absorption of the external roof surface, causing a high building thermal load. The results indicate that a cleaning preventive maintenance is suggested every two years. This problem can be mitigated by using impermeable coating on the tiles.

To conclude this section, it is important to emphasize both the complexity and time consumption of the previously mentioned computational code. It took almost one week to conclude the 1 year and 8 month simulation period.



**Figure 7.** Mould index evolution on the concrete tile.

As it can be seen in Figure 7, it takes almost 78 weeks to reach  $M = 3$  in terms of mould growth, which means, according to Table 1, that there are visual findings of mould at the roof surface, or 50% coverage of mould can be found through microscopic analysis. This graphic also provides information about the mould growth model dynamic, where a time delay of acceptable hygrothermal conditions for growth is considered. Just after this time delay, mould starts growing in order to overcome  $M = 1$  and  $M = 3$  values.

#### 4. Least Squares Support Vector Machines (LS-SVMs)

This section introduces the machine learning technique adopted to identify the hygrothermal dynamic of building roofs. Computational intelligence has been widely applied to identify building thermal dynamics and their subsystems, especially in control applications. The LS-SVM can be classified within a class of models that are used for pattern recognition, those that use a set or subset of training data in the prediction stage based on kernels. These methods perform predictions from combinations of the outputs of functions centered on each of the points available. The functions used for weighting a given set of training data are called kernels.

At first, SVMs were used to train classifiers based on the concept of structural risk minimization [47]. Besides, the SVMs were developed using the method known as statistical learning. Statistical learning theory was developed for solving problems where a small amount of data and little prior knowledge about the system are available, which differs from the traditional methods.

The SVM technique is designed to adjust the vectors defined for supporting a hyperplane, which aims to separate the input data. The SVM estimated the relationship between output  $y_i$  and an input pattern  $x$  by the following equation:

$$y_i = w\varphi(x) + b, \quad (15)$$

where  $b$  is a bias term,  $w$  is a weighting vector and  $\varphi$  is a nonlinear function that maps the input pattern  $x$  into a higher-dimensional feature space. The coefficient vector  $w$  and bias term  $b$  are unknown and can be obtained by solving an optimization problem.

When LS-SVM is applied for system identification tasks, the following optimization problem linked with the minimization of the risk function  $J$  can be defined [29]:

$$\min J = \frac{1}{2} \|W\|^2 + \gamma \frac{1}{2} \sum_{i=1}^n \varepsilon[y_i, f(x_i)], \quad (16)$$

subject to

$$\varepsilon[y_i, f(x_i)] = \begin{cases} 0, & |y_i - f(x_i)| \leq \varepsilon \\ |y_i - f(x_i)| - \varepsilon, & \text{others.} \end{cases} \quad (17)$$

where  $W$  is the vector of weights,  $\varepsilon$  is a given real number and  $\gamma$  is a regularization parameter that provides balance between model complexity and training error.

The first part of the objective function given by Equation (16) is used to regulate the weights and penalize those with higher values. Due to regularization, weights tend to converge to smaller values. This is necessary because heavy loads cause excessive variance in the model dynamic, deteriorating the generalization ability of LS-SVM. The second part of Equation (16) represents the regression error of the training data. The equality constraint imposed by Equation (17) provides the definition of the regression error.

In the case of nonlinearly separable patterns, the model needs to add variables to the problem, by introducing loss variables,  $\zeta_i$  and  $\zeta_i^*$ . In this case, it is possible to transform Equation (16) into a primal objective function given by:

$$\min J = \frac{1}{2} \|W\|^2 + \gamma \frac{1}{2} \sum_{i=1}^n (\zeta_i + \zeta_i^*), \quad (18)$$

subject to

$$\begin{cases} y_i - W\varphi(x_i) - b \leq \varepsilon + \zeta_i \\ W\varphi(x_i) + b - y_i \leq \varepsilon + \zeta_i^* \end{cases}, \quad \text{where } i = 1, \dots, N \text{ and } \zeta_i, \zeta_i^* \geq 0 \quad (19)$$

By introducing the Lagrange multipliers  $\alpha_i$  and  $\alpha_i^*$  (support vectors), the regression function given by Equation (15) can be written as:

$$f(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) G(x_i, x_j) + b \quad (20)$$

where  $G(x_i, x_j)$  is the core function, and vectors  $\alpha_i$  and  $\alpha_i^*$  are obtained solving the linear system of equations, following Karush-Kuhn-Tucker. The vector  $G(x_i, x_j)$  equals the inner product of two vectors  $x_i$  and  $x_j$  in the space of characteristics,  $\varphi(x_i)$  and  $\varphi(x_j)$ ; i.e.,  $G(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$ . The fact of adopting nucleuses to replace the calculation of  $\varphi(x_i)$  and  $\varphi(x_j)$  is complex and can be done in a simpler way by means of an approximate function.

These nucleuses generate a mapping between the input space and a high dimensional space, called the feature space. The SVM hyperplane generated by this space of characteristics, to be mapped back to input space, becomes a non-linear surface. Finally, the separation hyperplane becomes no longer a linear function of the input vectors, but a linear function of the space vector of characteristics [39].

In this work, a Radial Basis Function (RBF) kernel was adopted, which is given by:

$$G(x_i, x_j) = \exp\left(\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (21)$$

where  $\sigma$  is the spread of Gaussian kernel. In this application, to solve the linear programming training problem of LS-SVM, the SIMPLEX method was adopted [48].

A Nonlinear AutoRegressive with eXogenous inputs (NARX) model structure was adopted in this work. Four inputs have been considered in the system identification procedures: external temperature (in K), external relative humidity (in %), direct solar radiation ( $\text{W/m}^2$ ), and diffuse solar

radiation ( $\text{W}/\text{m}^2$ ). Four outputs have been identified in a one-step ahead prediction considering a Multiple-Input, Single-Output (MISO) structure. The identified outputs are: sensible heat flux ( $\text{W}/\text{m}^2$ ), latent heat flux ( $\text{W}/\text{m}^2$ ), vapor flux ( $\text{kg}/\text{m}^2 \cdot \text{s}$ ), and the Mould Growth Risk index.

A NARX model can be defined as a product (Equation (22)) to create a nonlinear form presented in Equation (23).

$$\hat{y}(t) = [a_1, a_2, \dots, a_{n_a}, b_1, b_2, \dots, b_{n_b}] * \\ [y(t-1), y(t-2), \dots, y(t-n_a), u(t-n_d-1), u(t-n_k-2), \dots, u(t-n_d-n_b-1)]^T \quad (22)$$

$$\hat{y}(t) = f \left( \begin{array}{c} y(t-1), y(t-2), \dots, y(t-n_a), u(t-n_d-1), \\ u(t-n_d-2), \dots, u(t-n_d-n_b-1) \end{array} \right) \quad (23)$$

where  $t$  represents the current time, and  $d$  the delayed sample.

The nonlinear function  $f$  can be expressed in terms of the model regressors, and the nonlinear mapping can be performed using nonlinear estimators. In Equations (21) and (22),  $y(t)$  represents the current output of the model,  $y(t-d)$  is a finite number of past outputs,  $u(t-d)$  the inputs,  $e(t)$  is a white-noise error that is introduced in the difference equation, and  $\hat{y}(t)$  is the predicted output of the system. The model structure is entirely defined by three integers, where  $n_a$  represents the number of poles,  $(n_b - 1)$  is the number of zeros, and  $n_d$  is the time delay of the systems.

## 5. Results

This section was divided into three parts. The first describes the simulation parameters adopted for the LS-SVM method on the system identification procedure. The second part of this section shows the analysis performed on the data set in order to define the percentage of data used for both training and test phases. Finally, Section 5.3 presents the prediction results considering the four MISO models proposed in this study.

### 5.1. Simulation Parameters

At the beginning of this analysis, one year and six months of data were collected (2444 samples sets of inputs and outputs) considering the physical domain and the numerical model presented in Section 2.

The training set was divided into 10 subsets following the 1-fold cross validation method [49], in order to train the classifier 10 times, each time leaving out one of the subsets from training, but using the omitted subset to compute the classification errors using the Mean Absolute Error (MAE) presented in Equation (24) as minimization criterion.

$$MAE = \frac{\sum_{k=1}^N |y_i(t) - \hat{y}_i(t)|}{N} \quad (24)$$

In terms of the model structure,  $n_a$ , and  $n_b$  were set equal to 2 according to previous analysis of the number of regressors using the neighborhood component analysis for regression [50]. Additionally, no time delay was considered between the inputs and the outputs ( $n_d = 0$ ).

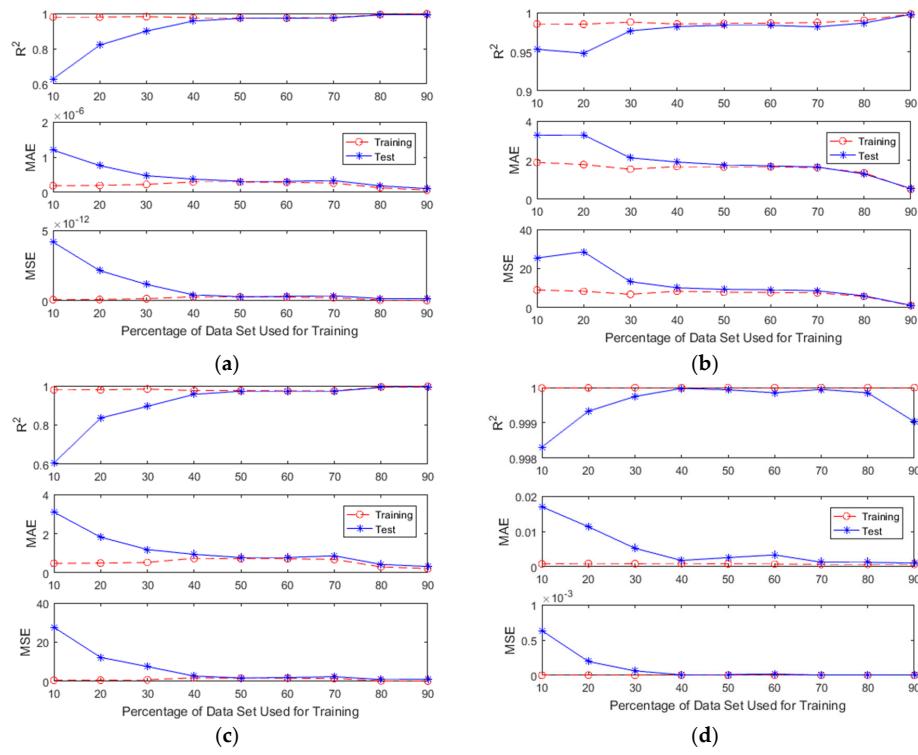
### 5.2. Definition of Training and Test Data Sets

In order to define the percentage of data used for training, an analysis using 10% to 90% of data for training was performed. In this case, besides MAE, the Multiple Correlation Coefficient ( $R^2$ ) and the Mean Square Error (MSE) were also adopted.

$$R^2 = 1 - \frac{\sum_{t=1}^N (y_i(t) - \hat{y}_i(t))^2}{\sum_{k=1}^N (y_i(t) - \bar{y}_i)^2} \quad (25)$$

$$MSE = \frac{\sum_{k=1}^N (y_i(t) - \hat{y}_i(t))^2}{N} \quad (26)$$

Figure 8 presents the training percentage related to  $R^2$ , MAE, and MSE for LS-SVM output prediction. As it can be verified in Figure 8, the test set was also presented in order to evaluate the balance between the quantity of data used for both training and validation procedures.

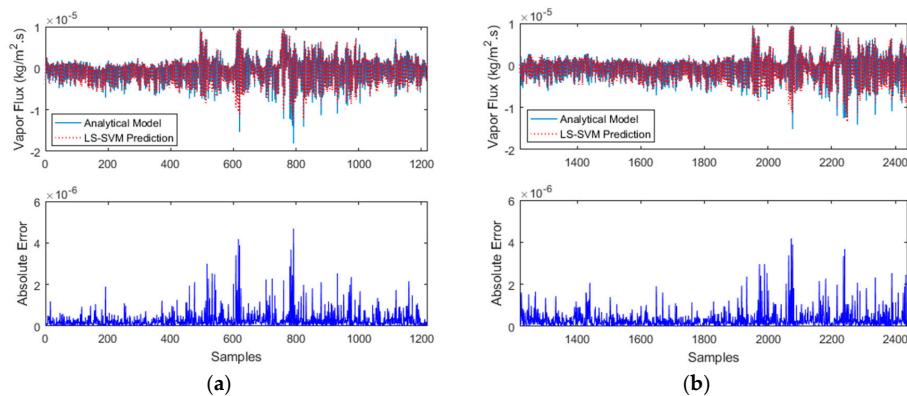


**Figure 8.** Analysis of the data percentage used for training: (a) output 1: vapor flux; (b) output 2: sensible heat flux; (c) output 3: latent heat flux; (d) output 4: mould growth risk.

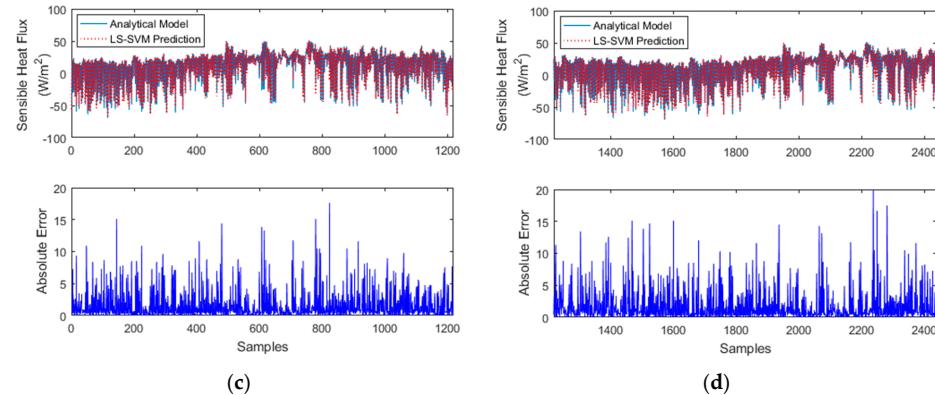
### 5.3. LS-SVM Prediction

According to the results presented in the previous subsection, 50% of data were selected for training, as they present reasonable approximation for all the outputs for both training and test procedures.

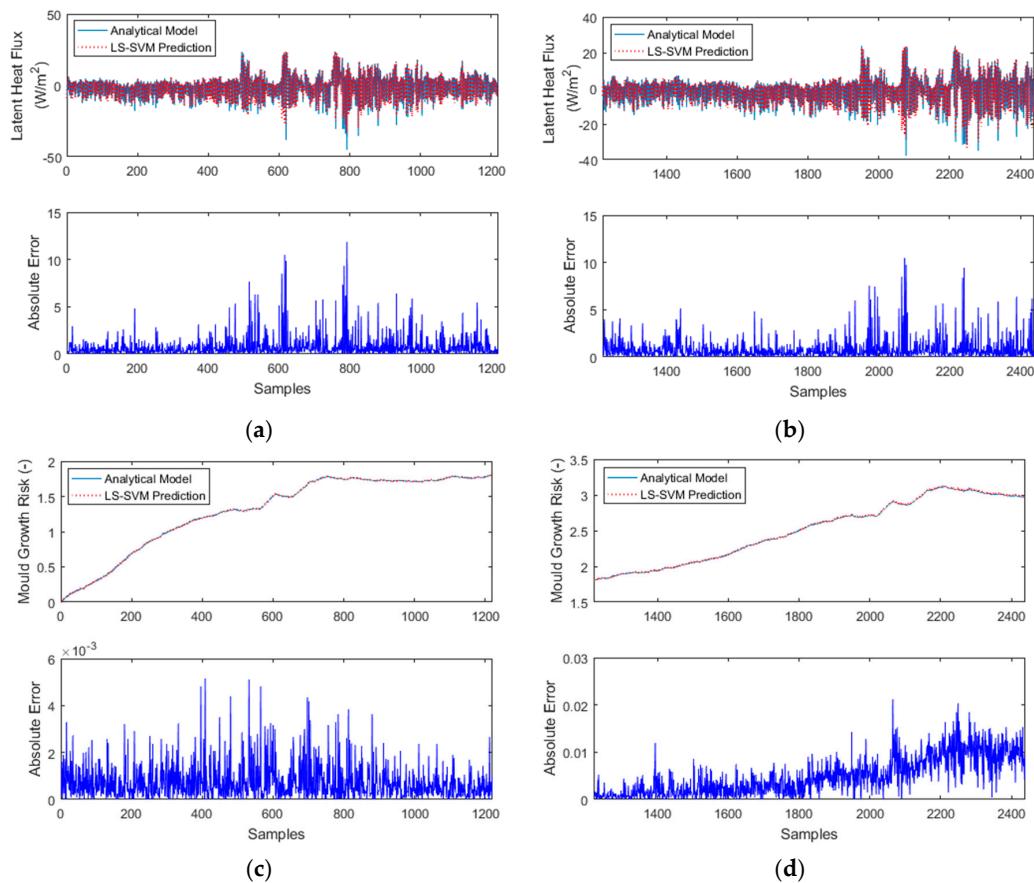
Figures 9 and 10 present the comparison between the LS-SVM model and the numerical method in terms of prediction. Figure 9 presents both training and test phases for outputs 1 (vapor flow) and 2 (sensible heat flow), while Figure 10 shows the results for outputs 3 (latent heat flow) and 4 (mould growth risk). The absolute error is also presented in these figures. Additionally, Table 2 reports the values of the Multiple Correlation Coefficient for both training and test phases.



**Figure 9. Cont.**



**Figure 9.** Comparison between the numerical and the LS-SVM models (a) output 1: vapor flux—training phase; (b) output 1: vapor flux—test phase; (c) output 2: sensible heat flux—training phase; (d) output 2: sensible heat flux—test phase.



**Figure 10.** Comparison between the numerical and the LS-SVM models (a) output 3: latent heat flux—training phase; (b) output 3: latent heat flux—test phase; (c) output 4: mould growth risk—training phase; (d) output 4: mould growth risk—test phase.

As can be observed in Figures 9 and 10, the highest absolute error values can be found in the mould growth index approximation, which can be justified by the different in the dynamic of the mould growth model. As the  $M$  index provides different growth behavior in distinct stages of growth, those defined by distinct equations, all these stages should be used in the LS-SVM training stage. As the LS-SVM training data set presented in this work adopted only the behavior between

the values of  $0 \leq M < 2$  (Figure 9c), the model was still capable of reproducing mould growth with considerable precision.

**Table 2.** Multiple Correlation Coefficient evaluations for both training and test phases.

Output #	Output Name	R <sup>2</sup> Training	R <sup>2</sup> Test
1	Vapor Flux	0.97401	0.97086
2	Sensible Heat Flux	0.98589	0.98376
3	Latent Heat Flux	0.97348	0.97049
4	Mould Growth Risk	1.00000	0.99979

As can be viewed in Table 2, the model presented consistent approximation for all four outputs. Higher values of  $R^2$  for the mould growth risk can be justified by its prior behavior in exceeding the  $M = 2$  limit, as the multiple correlation index is a cumulative measure.

In terms of computational effort, the whole simulation considering both training and test phases did not take more than 30 s, while the traditional method took about 120 h.

## 6. Conclusions and Future Research

This article presented an approach to predict vapor flux, sensible heat flux, latent heat flux and mould growth risk for concrete tiles based on the external weather conditions, considering as inputs for four MISO (multiple-input, single-output) models external temperature, relative humidity and direct and diffuse solar radiation; all inputs were obtained from a Test Reference Year (TRY) weather file.

Roofs are subjected to both thermal and moisture gradients, so that an accurate heat transfer determination requires a simultaneous calculation of both sensible and latent effects. Therefore, a mathematical model considering the combined two-dimensional heat, air and moisture transport through an unsaturated roof was presented, and the effects of moisture adsorption and desorption on the thermal performance of concrete tiles was shown. However, this type of numerical model is complex and both hardware and time-consuming when used for long periods of simulations, which is the case of mould growth evaluation.

Besides its effect on heat transfer, moisture can cause damage to the building structure and can promote both mould and mildew growth. The mould growth on roof surfaces can increase the solar absorptivity, decreasing its hygrothermal performance.

By considering the whole-building simulation, a tendency in the building physics area due to energy policies and the search for thermal comfort in indoor environments, fast and precise techniques, that could be coupled to building simulation software, are proposed and present consistent results, as shown in this work.

For future research, the authors intend to include the changes on roof solar absorptivity in the presented model, so that mould growth and the effects on the hygrothermal performance of the building can be verified more precisely. Moreover, an additional roof painting layer will be considered during simulations in order to provide a consistent efficiency analysis in order to reduce mould growth.

**Acknowledgments:** The authors would like to thank the National Council of Scientific and Technologic Development of Brazil—CNPq (grants number: 404659/2016-0, 303908/2015-7-PQ, 443350/2014-0, 305508/2014-8), and “Fundação Araucária” (grants number: 117/2014 and 17/2015) for the financial support of this work.

**Author Contributions:** Gerson Henrique dos Santos proposed and implemented the model for moisture, air and energy balances presented on Section 2.1, provided data for the system identification procedures and wrote part of the article. Leandro dos Santos Coelho conceived and designed the experiments; Roberto Zanetti Freire performed the experiments; analyzed the data, and wrote the other part of the article.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Sadat, Z.; Tahsildoust, M.; Hafezi, M. Thermal comfort in educational buildings: A review article. *Renew. Sustain. Energy Rev.* **2016**, *59*, 895–906. [[CrossRef](#)]
2. Ciulla, G.; Lo Brano, V.; D’Amico, A. Modelling relationship among energy demand, climate and office building features: A cluster analysis at European level. *Appl. Energy* **2016**, *183*, 1021–1034. [[CrossRef](#)]
3. U.S. Energy Information. A Annual Energy Outlook 2013. *Off. Integr. Int. Energy Anal.* **2013**, *1*, 1–244.
4. Brazilian Energy Research Company (EPE). Brazilian Energy Balance 2016. Available online: [https://ben.epe.gov.br/downloads/Relatorio\\_Final\\_BEN\\_2016.pdf](https://ben.epe.gov.br/downloads/Relatorio_Final_BEN_2016.pdf) (accessed on 4 April 2017).
5. Crawley, D.B.; Lawrie, L.K.; Winkelmann, F.C.; Buhl, W.F.; Huang, Y.J.; Pedersen, C.O.; Strand, R.K.; Liesen, R.J.; Fisher, D.E.; Witte, M.J.; et al. EnergyPlus: Creating a new-generation building energy simulation program. *Energy Build.* **2001**, *33*, 319–331. [[CrossRef](#)]
6. Mendes, N.; Oliveira, R.C.L.F.; Santos, G.D. Domus 2.0: A Whole-Building Hygrothermal Simulation Program. In Proceedings of the Eighth International IBPSA Conference, Eindhoven, The Netherlands, 11–14 August 2003; pp. 1417–1422.
7. Loonen, R.C.G.M.; Favoino, F.; Hensen, J.L.M.; Overend, M.; Attia, S. Building Performance Simulation Tools: Selection Criteria and User Survey. *J. Build. Perform. Simul.* **2010**, *1493*, 1–19. [[CrossRef](#)]
8. Ogando, A.; Cid, N.; Fernández, M. Energy Modelling and Automated Calibrations of Ancient Building Simulations: A Case Study of a School in the Northwest of Spain. *Energies* **2017**, *10*, 807. [[CrossRef](#)]
9. Kim, Y.-R.; Kang, H.J. Development of a Mobile Application for Building Energy Prediction Using Performance Prediction Model. *Energies* **2016**, *9*, 160. [[CrossRef](#)]
10. Min, H.; Zhang, W.; Gu, X.; Robert, C. Coupled heat and moisture transport in damaged concrete under an atmospheric environment. *Constr. Build. Mater.* **2017**, *143*, 607–620. [[CrossRef](#)]
11. Zhang, M.; Qin, M.; Rode, C.; Chen, Z. Moisture buffering phenomenon and its impact on building energy consumption. *Appl. Therm. Eng.* **2017**, *124*, 337–345. [[CrossRef](#)]
12. Dos Santos, G.H.; Mendes, N. Hygrothermal bridge effects on the performance of buildings. *Int. Commun. Heat Mass Transf.* **2014**, *53*, 133–138. [[CrossRef](#)]
13. Mazzeo, D.; Oliveti, G.; Arcuri, N. A Method for Thermal Dimensioning and for Energy Behavior Evaluation of a Building Envelope PCM Layer by Using the Characteristic Days. *Energies* **2017**, *10*, 659. [[CrossRef](#)]
14. Freire, R.Z.; dos Santos Coelho, L.; dos Santos, G.H.; Mariani, V.C. Predicting building’s corners hygrothermal behavior by using a Fuzzy inference system combined with clustering and Kalman filter. *Int. Commun. Heat Mass Transf.* **2016**, *71*, 225–233. [[CrossRef](#)]
15. Moon, J.W.; Kim, K.; Min, H. ANN-Based Prediction and Optimization of Cooling System in Hotel Rooms. *Energies* **2015**, *8*, 10775–10795. [[CrossRef](#)]
16. Moon, J.W.; Jung, S.K.; Lee, Y.O.; Choi, S. Prediction Performance of an Artificial Neural Network Model for the Amount of Cooling Energy Consumption in Hotel Rooms. *Energies* **2015**, *8*, 8226–8243. [[CrossRef](#)]
17. Moon, J.W.; Chung, M.H.; Song, H.; Lee, S. Performance of a Predictive Model for Calculating Ascent Time to a Target Temperature. *Energies* **2016**, *9*, 1090. [[CrossRef](#)]
18. Suehrcke, H.; Peterson, E.L.; Selby, N. Effect of roof solar reflectance on the building heat gain in a hot climate. *Energy Build.* **2008**, *40*, 2224–2235. [[CrossRef](#)]
19. Al-Sanea, S.A. Thermal performance of building roof elements. *Build. Environ.* **2002**, *37*, 665–675. [[CrossRef](#)]
20. Ran, J.; Tang, M.; Jiang, L.; Zheng, X. Effect of building roof insulation measures on indoor cooling and energy saving in rural areas in Chongqing. *Procedia Eng.* **2017**, *180*, 669–675. [[CrossRef](#)]
21. Boixo, S.; Diaz-vicente, M.; Colmenar, A.; Alonso, M. Potential energy savings from cool roofs in Spain and Andalusia. *Energy* **2012**, *38*, 425–438. [[CrossRef](#)]
22. Uemoto, K.L.; Sato, N.M.N.; John, V.M. Estimating thermal performance of cool colored paints. *Energy Build.* **2010**, *42*, 17–22. [[CrossRef](#)]
23. Belusko, M.; Bruno, F.; Saman, W. Investigation of the thermal resistance of timber attic spaces with reflective foil and bulk insulation, heat flow up. *Appl. Energy* **2011**, *88*, 127–137. [[CrossRef](#)]
24. Saber, H.H.; Maref, W.; Sherrer, G.; Swinton, M.C. Numerical modeling and experimental investigations of thermal performance of reflective insulations. *J. Build. Phys.* **2012**. [[CrossRef](#)]

25. D’Orazio, M.; Di Perna, C.; Di Giuseppe, E.; Morodo, M. Thermal performance of an insulated roof with reflective insulation: Field tests under hot climatic conditions. *J. Build. Phys.* **2013**, *36*, 229–246. [CrossRef]
26. Buonomano, A. Code-to-Code Validation and Application of a Performance Analysis. *Energies* **2016**, *9*, 301. [CrossRef]
27. Cheng, H.; Wang, X.; Zhou, M. Optimized Design and Feasibility of a Heating System with Energy Storage by Pebble Bed in a Solar Attic. *Energies* **2017**, *10*, 328. [CrossRef]
28. Vapnik, V.N. Statistical Learning Theory. *Interpreting* **1998**, *2*, 736.
29. Souto, G.; Guilherme, L.; Luvizotto, J.; Cocco, V.; Coelho, S. Least squares support vector machines with tuning based on chaotic differential evolution approach applied to the identification of a thermal process. *Expert Syst. Appl.* **2012**, *39*, 4805–4812. [CrossRef]
30. Friedman, N.; Kumar, P.; Zander, E.; Matthies, H.G. Bayesian calibration of model coefficients for a simulation of flow over porous material involving SVM classification. *PAMM* **2016**, *16*, 669–670. [CrossRef]
31. Dai, C.; Zhang, H.; Arens, E.; Lian, Z. Machine learning approaches to predict thermal demands using skin temperatures: Steady-state conditions. *Build. Environ.* **2017**, *114*, 1–10. [CrossRef]
32. Khan, U.; Schmidt-Thieme, L.; Nanopoulos, A. Collaborative SVM classification in scale-free peer-to-peer networks. *Expert Syst. Appl.* **2017**, *69*, 74–86. [CrossRef]
33. Yildiz, B.; Bilbao, J.I.; Sproul, A.B. A review and analysis of regression and machine learning models on commercial building electricity load forecasting. *Renew. Sustain. Energy Rev.* **2017**, *73*, 1104–1122. [CrossRef]
34. Zhang, F.; Deb, C.; Lee, S.E.; Yang, J.; Shah, K.W. Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique. *Energy Build.* **2016**, *126*, 94–103. [CrossRef]
35. Chen, Y.; Xu, P.; Chu, Y.; Li, W.; Wu, Y.; Ni, L.; Bao, Y.; Wang, K. Short-term electrical load forecasting using the Support Vector Regression (SVR) model to calculate the demand response baseline for office buildings. *Appl. Energy* **2017**, *195*, 659–670. [CrossRef]
36. Suykens, J.A.; Van Gestel, T.; De Brabanter, J. *Least Squares Support Vector Machines*; World Scientific Pub. Co.: Singapore, 2002; ISBN 9812381511.
37. Gestel, T.V.; Suykens, J.A.K.; Baesens, B.; Viaene, S.; Vanthienen, J.; Dedene, G.; Moor, B.D.; Vandewalle, J. Benchmarking Least Squares Support Vector. *Mach. Learn.* **2004**, *54*, 5–32. [CrossRef]
38. Dos Santos, G.H.; Mendes, N. Analysis of Hygrothermal Performance and Mould Growth on Concrete Tiles. In Proceedings of the 23rd ABCM International Congress of Mechanical Engineering (COBEM 2015), Rio de Janeiro, Brazil, 6–11 December 2015; pp. 1–8.
39. Mendes, N.; Philippi, P.C.; Lamberts, R. A new mathematical method to solve highly coupled equations of heat and mass transfer in porous media. *Int. J. Heat Mass Transf.* **2002**, *45*, 509–518. [CrossRef]
40. Viitanen, H.; Ojanen, T.; Peuhkuri, R. Mould growth modelling to evaluate durability of materials. In Proceedings of the 12DBMC—International Conference on Durability of Building Materials and Components, Porto, Portugal, 12–15 April 2011; pp. 1–8.
41. Hukka, A.; Viitanen, H.A. A mathematical model of mould growth on wooden material. *Wood Sci. Technol.* **1999**, *33*, 475–485. [CrossRef]
42. Vereecken, E.; Roels, S. Review of mould prediction models and their influence on mould risk evaluation. *Build. Environ.* **2012**, *51*, 296–310. [CrossRef]
43. Künzel, H.M.; Holm, A.H.; Krus, M. Hygrothermal Properties and Behaviour of Concrete. *WTA Alm.* **2008**, *1*, 161–181.
44. Hens, H. IEA Annex 14: Condensation and Energy. *J. Build. Phys.* **1992**, *15*, 261–273. [CrossRef]
45. Lamberts, R. Wheather Files in Format TRY, SWERA, CSV and BIN. (In Portuguese). Available online: <http://www.labeee.ufsc.br/downloads/arquivos-climaticos/formato-try-swera-csv-bin> (accessed on 28 June 2017).
46. Swinbank, W.C. Long wave radiation from clear skies. *Q. J. R. Meteorol. Soc.* **1963**, *89*, 339–348. [CrossRef]
47. Burges, C.J. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Min. Knowl. Discov.* **1998**, *2*, 121–167. [CrossRef]
48. Sentelle, C.; Georgopoulos, M. A Fast Revised Simplex Method for SVM Training. In Proceedings of the 19th IEEE International Conference on Pattern Recognition (ICPR2008), Tampa, FL, USA, 8–11 December 2008; pp. 3–6.

49. An, S.; Liu, W.; Venkatesh, S. Fast cross-validation algorithms for least squares support vector machine and kernel ridge regression. *Pattern Recognit.* **2007**, *40*, 2154–2162. [[CrossRef](#)]
50. Goldberger, J.; Roweis, S.; Hinton, G.; Salakhutdinov, R. Neighbourhood Components Analysis. *Science* **2000**, *290*, 2323–2326.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).