buildings

MDPI

Article

Artificial Intelligence Approach for Bio-Based Materials' Characterization and Explanation

Ahmed Alami [1], Lala Rajaoarisoa [2], Nicolas Dujardin [3], Ali Benouar [4], Khacem Kaddouri [5], Khedidja Benouis [1] and Mohammed-Hichem Benzaama [6,*]

[1] Laboratory of Process Engineering, Materials and Environment, Faculty of Technology, University of Djillali Liabes, P.O. Box 89, Sidi Bel Abbes 22000, Algeria; alami.ahmed21@gmail.com (A.A.); khedidjabenouis@gmail.com (K.B.)
[2] CERI Systèmes Numériques, IMT Nord Europe, Université de Lille, F-59500 Douai, France; lala.rajaoarisoa@imt-nord-europe.fr
[3] Univ Paris Est Creteil, CERTES, F-94010 Créteil, France; nicolas.dujardin@u-pec.fr
[4] Laboratory of Complex Systems (LCS), Higher School of Electrical and Energy Engineering (ESGEE), Oran 31000, Algeria; benouardz@yahoo.fr
[5] Laboratory of Mechanics Physics of Materials (LMPM Laboratory), University of Djillali Liabes, Sidi Bel Abbes 22000, Algeria; kkhacem@gmail.com
[6] Institut de Recherche, ESTP, 28 Avenue du Président Wilson, F-94230 Cachan, France
* Correspondence: hbenzaama@estp.fr

Abstract: This paper introduces a numerical methodology for classifying and identifying types of bio-based materials through experimental thermal characterization. In contrast to prevailing approaches that primarily focus on thermal conductivity, our characterization methodology encompasses several thermal parameters. In this paper, the physical characteristics of seven types of bio-based concrete were analyzed, focusing on the thermal properties of palm- and esparto-fiber-reinforced concrete. The proposed method uses artificial intelligence techniques, specifically the k-means clustering approach, to segregate data into homogeneous groups with shared thermal characteristics. This enables the elucidation of insights and recommendations regarding the utilization of bio-based insulation in building applications. The results show that the k-means algorithm is able to efficiently classify the reference concrete (RC) with a performance of up to 71%. Additionally, the technique is more accurate when retaining only six centroids, which, among other things, allows all the characteristics associated with each type of concrete to be grouped and identified. Indeed, whether for k clusters k = 7 or k = 5, the technique was not able to predict the typical characteristics of 2% or 3% esparto concrete (EC).

Keywords: experimental thermal characterization; bio-based materials; k-means clustering; artificial neural networks; bio-based insulation

check for updates

Citation: Alami, A.; Rajaoarisoa, L.; Dujardin, N.; Benouar, A.; Kaddouri, K.; Benouis, K.; Benzaama, M.-H. Artificial Intelligence Approach for Bio-Based Materials' Characterization and Explanation. *Buildings* **2024**, *14*, 1602. https://doi.org/10.3390/buildings14061602

Academic Editor: Biao Hu

Received: 30 April 2024
Revised: 24 May 2024
Accepted: 27 May 2024
Published: 1 June 2024

Copyright: © 2024 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 (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The global technological and socio-economic developments have led to excessive energy consumption, resulting in the emission of large amounts of carbon dioxide CO_2/m^2 into the environment. According to current statistics from the International Energy Agency (IEA), the construction sector is a major contributor to energy consumption and carbon emissions. It is responsible for almost 36% of total energy consumption and 37% CO_2/m^2 emissions in the world [1]. The use of renewable energy has been the subject of many research works by Ahmed A. et al. [2,3]. The building and construction industry is one of the main energy-intensive industries, accounting for about 40% of the world's annual energy consumption according to A. Hamidat et al. and A.P. Olukoya et al. [4,5]. Plant fibers in concrete production can lead to a low incarnates the carbon envelope of the building and can even act as a carbon casting. Indeed, about 1 kg of bio-based building materials can store at least 1.44 kg CO_2 at 12% relative humidity, as shown by T. Lecompte et al. [6]. Furthermore, in the building scale, walls built with 0.25 m hemp concrete

blocks can insulate 14–35 kg CO_2/m^2 over a 100-year lifetime, as shown by M. Boutin et al. [7]. Therefore, using plant fibers can be beneficial as an agro-aggregate for bio-sourced concrete. The desire to design modern buildings with low environmental impact ("low carbon") and high energy efficiency has led to an increase in research in the field of concrete construction, reflecting the growing trend towards the use of concrete in passive buildings. In this case, the main objective of this work was to test a composite material, in the form of concrete, stabilized with lime and mixed with date palm and esparto waste aggregates. This was carried out to study the effect of these wastes at different contents (0, 2, 3 and 5% by weight) concerning cement on the thermal properties of the concrete samples obtained. The reported results showed an improvement in thermal insulation performance, from 1.48 W/m K for the test without fiber (the control concrete) to 0.75 W/m K for the test with 5% of palm fiber and 1.21 W/m K for alfa fiber. This improvement is accompanied by a reasonable decrease in mechanical properties while maintaining the minimum performance thresholds required by earthwork construction standards. In recent years, there has been a growing interest in incorporating plant fibers into concrete for construction purposes. Plant fibers, such as hemp, sisal, and bamboo, have been found to improve the mechanical properties and durability of concrete, as well as reduce its environmental impact. These fibers can increase the tensile strength and ductility of concrete, which can help prevent cracking and increase its resistance to impact and fatigue according to M.R. Ahmad et al., and J.P. Holman. et al. [8,9]. Furthermore, incorporating plant fibers can reduce the carbon footprint of concrete production by replacing some of the cement content, which is a significant contributor to greenhouse gas emissions. Additionally, using plant fibers as a reinforcement material can provide a sustainable alternative to traditional steel or synthetic fibers. Overall, the introduction of plant fibers in concrete has the potential to improve the sustainability and resilience of construction materials while also providing economic benefits through reduced maintenance and repair costs according to M. Charai et al., and Y. Shang et al. [10,11]. M. Mehravar et al. have carried out research on saving energy [12]. A recent study showed that the use of straw bales as a building material reduces construction energy consumption and the embodied carbon emissions were reduced by 93.12% and 76%, respectively. These results are generally consistent with the results observed in the majority of research on bio-based concrete buildings, as shown by T.H. Mai et al. [11,13,14]. This paper also provides new data related to the next generation of a comparison between two plant fibers (palm, esparto) in concrete, mainly in relation to thermal properties (conductivity, diffusivity and heat capacity), and appropriate operations for energy efficiency in buildings. Bio-sourced materials have gained increasing attention in recent years as a sustainable alternative to traditional synthetic materials. These materials are produced from renewable resources and can have lower environmental impacts compared to their synthetic counterparts Manal Bouasria et al. [15]. However, the thermal properties of bio-based materials play a crucial role in determining their suitability for various applications such as insulation, packaging, and energy storage, according to Yassine El Mendili et al. [16]. Accurate modeling of these properties is essential for the optimal design of these materials. Traditional modeling techniques for thermal properties are often based on physical and mathematical models that require assumptions and simplifications. These approaches may not be able to capture the complex relationships between input variables and output properties of bio-sourced materials. Moreover, the experimental data needed to validate these models can be scarce and costly to obtain. In recent years, data-driven modeling techniques have emerged as a promising alternative to traditional modeling techniques for thermal properties of materials, as shown by Singh Ramvir et al. [17].

Data-driven modeling techniques, such as machine learning, can analyze large sets of experimental data and develop predictive models that can accurately estimate the thermal properties of bio-sourced materials [18,19]. These models can handle complex relationships between input variables and output properties, as well as incorporate data from various

sources [19]. Several studies have demonstrated the potential of machine learning in predicting thermal properties of materials BOUASRIA, as shown by Manal et al. [20].

Machine learning approaches have shown promise in predicting the thermal properties of materials, including bio-based materials. These models can capture complex relationships between input variables and output properties, and can learn from large dataset of experimental measurements. Artificial neural networks (ANNs) have been widely used for predicting the thermal conductivity of building materials. For example, Chen L. et al. [21] describe a new method for identifying the thermal resistance of exterior walls of buildings using (ANNs) based on numerical experiments. The two authors developed a database of simulated building models with different wall configurations and used it to train the ANNs. The performance of the ANNs was evaluated and compared to that of traditional methods. The results showed that the ANN method was accurate and efficient, and can be used for practical applications. Furthermore, Sharo A. et al. [22]. presented a new (ANN) model to predict the thermal properties of concrete using different neurons and activation functions. The authors trained the ANNs using experimental data and compared the performance of the different models. The results showed that the ANN models with certain neurons and activation functions performed better than the others. The proposed ANN model can be used for accurate prediction of the thermal properties of concrete. OZEL Cengiz et al. [23] proposed an ANN approach for predicting the effective thermal conductivity of moist porous materials. The authors trained the ANNs using experimental data and compared the performance of the ANN models to that of traditional methods. The results showed that the ANN approach was more accurate and efficient than the traditional methods. The proposed ANN model can be used for predicting the thermal conductivity of moist porous materials in various applications. Mohsen Hajihassani et al. [24] compared the performance of two different artificial intelligence techniques, ANFIS (adaptive neuro-fuzzy inference system) and ANN (artificial neural network), for estimating thermal conductivity coefficients of construction materials. The authors conducted experimental tests to measure the thermal conductivity coefficients of various construction materials and used the data to train both ANFIS and ANN models. They compared the performance of the two models in terms of accuracy and computational efficiency. The results showed that both models were effective in predicting thermal conductivity coefficients, but the ANFIS model outperformed the ANN model in terms of accuracy. The findings of this study can be useful in developing more accurate and efficient models for predicting the thermal properties of construction materials. ANNs have also been used for predicting the specific heat capacity of building materials, which is necessary for predicting their mass heat. For example, Grieu S. et al. [25] conducted experimental tests to measure the thermal conductivity, thermal diffusivity and specific heat capacity of limestone. The performance of the model was validated using independent experimental data. The results showed that the model is accurate and can be used for predicting the thermal properties of limestone in various applications in the building industry. ANNs have also been used for predicting the thermal diffusivity of building materials. The authors Wang D et al. [26] proposed new and efficient approaches using artificial intelligence tools (artificial neural networks and neuro-fuzzy systems) and inverse methods to estimate the thermal diffusivity of building materials. They used a non-destructive photothermal method to obtain thermograms and estimate the in situ thermo-physical properties of the materials. Overall, machine learning models have shown great promise in predicting the thermal conductivity, mass heat, and thermal diffusivity of building materials. ANNs, in particular, have been widely used due to their ability to capture complex relationships between input variables and output properties. However, other machine learning models, such as SVR, RF, decision trees, Gaussian processes, and k-nearest neighbor algorithms, have also shown potential in predicting these thermal properties. These models can be useful for optimizing the thermal performance of building materials and designing energy-efficient buildings. In this context, OZEL Cengiz et al. [23] evaluated five empirical models and seven machine learning (ML) algorithms (i.e., Decision Tree (DT), Random Forest (RF), Gradient Boosting Decision Tree

(GBDT), Linear Regression (LR), K -nearest neighbors (KNN), Neural Network (NN), and Gaussian Process (GP)) to estimate soil thermal conductivity using 1602 measured soil thermal conductivity values. The ML algorithms, including Gradient Boosting Decision Tree (GBDT), Neural Network (NN), and Random Forest (RF), outperformed the empirical models, with an average RMSE of 66% and 82% of the empirical model values on validation and test sets, respectively. Soil moisture content and soil bulk density were identified as the most critical factors affecting soil thermal conductivity, with more than 80% of the influence importance value. The RF algorithm was recommended for selecting features due to its more consistent results.

Most of the studies cited are based on artificial neural networks. This study is based on a k-means algorithm data classification method that is not based on an ANN. The k-means algorithm does not use an activation function like ANNs. K-means is an unsupervised machine-learning method used for data classification and clustering. Its operation is based on iteration between updating cluster centers and assigning data points to these clusters without the need for an activation function.

In k-means, cluster centers are initially defined at random, then iteratively updated to minimize the sum of the distances between data points and their nearest cluster center. Data points are then assigned to the cluster whose center is closest based on a predefined distance measure such as Euclidean distance.

Furthermore, most of these articles focus on thermal conductivity and specific heat to the maximum, and they overlook thermal diffusivity. In this paper, we aimed to investigate the thermal properties of palm- and esparto-fiber-reinforced concrete to determine their suitability for building insulation. To achieve this, we set up seven types of bio-based concrete, including a reference concrete, and tested the replacement of a percentage of concrete with esparto or palm fibers. We subjected the samples to thermal stresses and collected data on their thermal conductivity, thermal capacity, and thermal diffusivity, among other parameters. We then applied a clustering approach based on k-means clustering to classify the data into homogeneous groups based on their characteristics. So, the research questions for this paper are as follows: What is the effectiveness of the k-means clustering method in classifying different types of bio-based concrete based on their physical characteristics? What is the recommended choice between using palm fibers or esparto fibers as insulation materials considering the trade-off between their performance?

This paper will be organized as follows. In Section 2, we will introduce the methodology and the conceptual plan to perform the characterization of bio-sourced materials. In Section 3, we will discuss the results and provide explanations of the observed phenomena. Finally, the conclusions drawn from this experiment will be given in Section 4.

2. Experimental and Numerical Methodology

After conducting a review of the recent literature, we devised a research strategy that outlines our procedures. The conceptual study strategy for this research is divided into four sections. Each section comprises various phases, all of which are described in detail in the following sections. The following is an overview of these phases:

- Benchmark design: We developed an experimental test bench that includes multiple measurements such as thermal diffusivity, thermal conductivity, and heat capacity. We test in this step two types of bio-sourced materials: concretes based on esparto and palm tree fibers.
- Data collection: experimental setting and test for data gathering.
- Data analysis and classification: this step consists of applying a clustering approach on data to separate it into homogeneous groups with common characteristics.
- Bio-materials' characterization: during this stage, we compare the observed characteristics of different concrete types with a reference concrete to analyze their physical properties' evolution.

2.1. Benchmark Design

2.1.1. Fibers

The primary plant fibers are represented in Figure 1. The fibers of palm flower, straight leaf fiber (b), and left esparto leaf fiber (a), which were the subject of our study, come from the Algerian steppe region. The plant material must be cleaned and freed from foreign matter, including particles of soil and dust that could affect the test results. For the analysis of esparto and palm, 10 g of finely ground grass, prepared plants with uniform-sized particles was sieved through the n° 24 and n° 27 sieves.



(**a**) (**b**)

Figure 1. Plant fibers esparto and palm.

2.1.2. Fibers

The chosen composition of the prepared concrete is presented in Table 1. Based on water absorption and water content, it is composed of 57% aggregates and 43% sand, with a water–cement ratio of E/C = 0.5. In this formula, a C25/30 class and a dosage of 350/m^3 were used.

Table 1. Ordinary concrete composition.

Actual Final Composition (According to Absorption and Water Content)	**%**	**Origin**	**Laboratory Weight (Kg) (15 × 15 × 15) cm^3**
Sand MIX (S0/3rafi + S0/1mechria)	43.0	SECH	9.64
Sand 0/1	0.0	0	0.00
Gravel 3/8	12.0	SECH	2.69
Gravel 8/15	35.0	SECH	7.86
Gravel 15/25	10.0	SECH	2.24
Cement		ZAHANA	4.20
Real Water		AEP	2.07
Adj 2 ”S,plasticizer” 9WG		Teknachem	0.05

2.2. Production of Test Samples

The samples were meticulously prepared according to the formulation specified in Table 1, where 2%, 3%, and 5% of esparto and palm fiber were added, respectively, compared to the substitution of the total weight of cement. Each addition was precisely measured to ensure accuracy and consistency in the mixture. This formulation variation was chosen to analyze the impact of different fiber concentrations on the properties of the concrete.

The test samples were of a standardized size, measuring (15 × 15 × 15) cm^3, as illustrated in Figure 2. This uniform size ensured consistency in testing conditions and facilitated an accurate comparison of the results.



Figure 2. Samples preparation.

To maintain controlled environmental conditions, the concrete specimens were submerged in water at a constant temperature of 20 °C. This immersion was facilitated using a thermostatic bath, which effectively regulated the temperature throughout the preservation period.

The preservation period in the thermostatic bath was carefully set at intervals of 7, 14, and 28 days. These durations were chosen to observe any changes or developments in the properties of the concrete over time under the influence of the added fibers.

At the conclusion of each preservation period, a compression test was conducted on the samples. This test evaluated the concretes' ability to withstand compressive forces, providing valuable insights into its strength and durability.

The steps to produce the samples are schematically outlined in Figure 3, providing a clear visual representation of the procedure. Additionally, detailed methodologies were provided to guide the exact procedures to be followed during sample preparation and testing. The methods followed are as follows:

- Introduce the equivalent binder (cement + plant fibers) and mix with the aggregate for 6 min;
- Introduce the rest of the mixing water and some of the additives; then, mix for 90 s;
- Introduce the rest of the additive and mix for 2 min;
- Form a cube;
- Weigh the sample after demolding and after storage in a constant temperature bath for a few days;
- Grind the sample with a vertical press.



Figure 3. Samples' preparation.

2.2.1. Thermal Characterization Devices

The device used in our study is the Hot Disk Thermal Constants Analyser (TPS 2500, Hot Disk AB, Gothenburg, Sweden). Measurements were taken at 23 °C (23 ± 1 °C).

The hot-disk method is described in reference [27]. It is a non-destructive transient method that uses a flat nickel probe (e ≈ 10 µm) placed between two layers of Kapton

(25.4 µm) as both a heat source and temperature sensor. In practice, the probe is often sandwiched between two identical samples. An alternative is to place the probe between the sample under study and an insulator with known thermal characteristics, as seen in Figure 4. The increase in the average temperature of the probe (ΔT) is measured through the total electrical resistance of the probe R:

$$R = R_0(1 + a\Delta T(t)) \quad (1)$$

where R_0 is the resistance temperature of the probe at t = 0, and α is the thermal expansion coefficient of nickel.

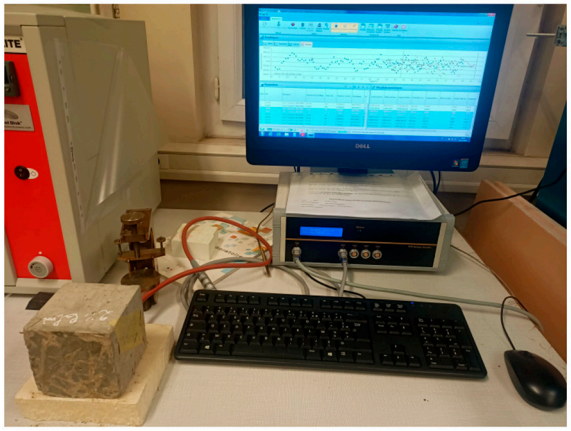


Figure 4. Hot-disk device.

2.2.2. Influence Parameters

During hot-disk experiments, the first step is to determine the best parameters to obtain optimal results on the various tested samples. These key parameters that require prior optimization are the probe size, experiment duration, sample size, and power delivered by the probe. Currently, there are several commercially available probes for performing thermal measurements using a hot-disk. Their radii range from 0.5 mm to 30 mm. The optimal size of the probe should allow for a characteristic time (τ) (see Equation (2)) between 0.33 and 1. The characteristic time is defined as follows:

$$\tau = \sqrt{(t \times \alpha)/r^2} \quad (2)$$

where t is time, (α) is the thermal diffusivity of the sample, and r is the radius of the probe. The heating caused by the probe must penetrate the studied material sufficiently for the measurement to be representative. However, the penetration depth must remain smaller than the smallest dimension of the sample. The manufacturer recommends a thickness at least equal to the diameter of the probe. The penetration depth is expressed as follows:

$$d = \sqrt{4 \times a \times t} \quad (3)$$

The power level of the probe also needs to be determined. Power influences temperature elevation but not penetration depth. It should be a few degrees according to the manufacturer (<5 °C). The thermal conductivity and diffusivity are then estimated from the thermal model integrated into the software. The estimation is performed only on a portion of the data points since the experimental conditions require removing the first points, which correspond to the contact resistance between the probe and the sample, and the last points so that the penetration depth does not exceed the sample thickness. These conditions are described in detail in [28]. The manufacturer recommends that the residues present an average deviation of 10^{-3} K at most. No uncertainty on the estimated parameters is provided by the measuring device. Repeatability studies generally show a

dispersion of thermal conductivity values of about 5%. In this study, the parameters (power, measurement time, temperature elevation, total characteristic time) were chosen to meet these criteria. The probe used was the 5501 (r = 6.4 mm). The values and uncertainties of conductivity and diffusivity were averaged over five experiments and associated standard deviation.

2.3. Data Analysis and Classification

Clustering is one of the most common exploratory data analysis techniques used in machine learning to obtain an insight about the structure of the data. Indeed, it is a particular discipline of machine learning that aims to separate data into homogeneous groups with common characteristics. In other words, it can be defined as the task of identifying subgroups in the data such that data points in the same subgroup (two or multiple clusters Figure 5) are very similar, while data points in different clusters are very different. In this paper, we will cover only the k-means technique, which is considered one of the most used clustering algorithms due to its simplicity.

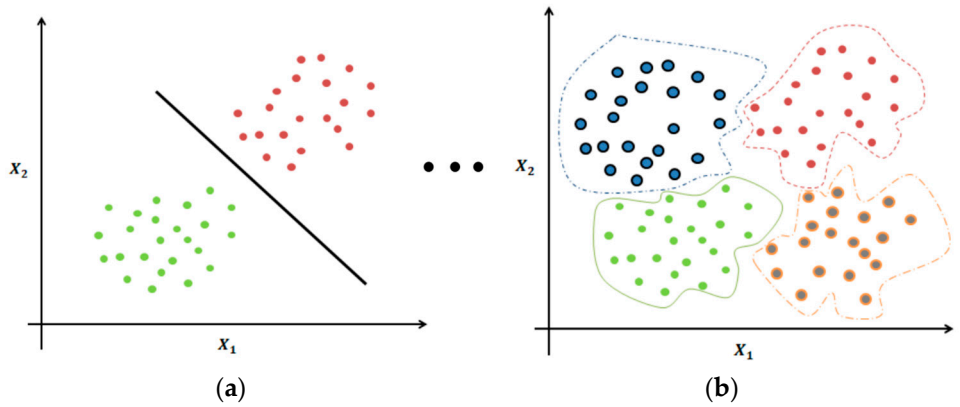


Figure 5. Clustering approach (**a**) with two identified classes; (**b**) with multiple identified classes.

2.3.1. The k-Means Approach

Essentially, the k-means method is a widely used ad well-known unsupervised clustering algorithm that seeks to minimize the average squared distance between points in the same cluster. Let us consider $p = (p_1, \ldots, p_n)$ and $q = (q_1, \ldots, q_n)$; the Euclidean distance, for instance, can be expressed as follows:

$$d(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \ldots + (p_i - q_i)^2 + \ldots + (p_n - q_n)^2} \quad (4)$$

It allows us to evaluate the distance between each point p and the centroids q. For each point, the Euclidean distance between this point and each of the centroids is calculated and then associated with the nearest centroid, i.e., the one with the smallest distance. Thus, the centroid of a finite set of s points $p_1, p_2, \ldots, p_s \in R^d$ can be defined as follows:

$$q = \frac{p_1 + p_2 + \ldots + p_s}{s} \quad (5)$$

So, for the k-means problem, we are given an integer k and a set of n data points $X \subset R^d$. We wish to choose k centers of C to minimize the objective function:

$$J = \sum_{i=1}^{m} \sum_{k=1}^{K} \omega_{ik} \|x^i - c_k\|^2 \quad (6)$$

where K and m specify, respectively, the number of clusters and the step. ω_{ik} = 1 for data point x^i if it belongs to cluster k; otherwise, ω_{ik} = 0. Also, c_k is the centroid of x_i's cluster.

The minimization problem is subdivided into two stages. The first one is the minimization of J by considering ω_{ik} as variable and c_k as fixed. The second one is the minimization of J by considering c_k as variable and ω_{ik} as fixed. Therefore, for each step, the following differentiated terms hold:

$$\frac{\partial J}{\partial \omega_{ik}} = \sum_{i=1}^{m} \sum_{k=1}^{K} \omega_{ik} \left\| x^i - c_k \right\|^2 \tag{7}$$

With

$$\omega_{ik} = \begin{cases} 1 & if \quad k = \text{argmin}_j \|x^i - c_j\|^2 \\ 0 & otherwise \end{cases}$$

In other words, the data point x_i is assigned to the closest cluster judged by its sum of squared distance from the cluster's centroid.

And according to each step, we have:

$$\frac{\partial J}{\partial \omega_{ik}} = 2 \sum_{i=1}^{m} \omega_{ik} \left\| x^i - c_k \right\|^2 = 0 \tag{8}$$

Consequently

$$c_k = \frac{\sum_{i=1}^{m} \omega_{ik} x^i}{\sum_{i=1}^{m} \omega_{ik}} \tag{9}$$

Which translates to recomputing the centroid of each cluster to reflect the new assignments. Finally, to validate the performance of the algorithm, we use a typical metric as *F1score* defined as follows:

$$F1score(i) = 2 * \frac{Precision(i) * Recall(i)}{Precision(i) + Recall(i)} \tag{10}$$

where

$$Precision\ (i) = \frac{Item\ well\ attributed\ to\ the\ class\ i}{Item\ attributed\ to\ the\ class\ i} \tag{11}$$

and

$$Recall\ (i) = \frac{Item\ well\ attributed\ to\ the\ class\ i}{Item\ belonging\ to\ the\ class\ i} \tag{12}$$

For a better understanding, Recall indicates the percentage of actual positives correctly predicted by our model. Specifically, it is calculated as the number of true positives (correctly predicted positives) divided by the total number of actual positives (true positives plus false negatives).

Precision, though similar to recall, focuses on a different aspect. It measures the accuracy of positive predictions. In other words, it is the number of true positives divided by the total number of predicted positives (true positives plus false positives).

2.3.2. The k-Means Algorithm

The k-means clustering [28] is an iterative, data-partitioning algorithm that assigns n observations to exactly one of the k clusters defined by centroids, where k is chosen before the algorithm starts. The algorithm proceeds as follows:

- Step 1: Choose k initial cluster centers (centroid) $C = \{c_1, \ldots, c_k\}$;
- Step 2: For each $i \in \{1, \ldots, k\}$, compute point to cluster centroid distances of all observations to each centroid c_i;
- Step 3: Assign each observation to the cluster with the closest centroid;
- Step 4: Compute the average of the observations in each cluster to obtain k new centroid locations;
- Step 5: Repeat steps 2 through 4 until cluster C assignments do not change, or the maximum number m of iterations is reached.

It is standard practice to choose the initial centers uniformly at random from X. For Step 2, ties may be broken arbitrarily as long as the method is consistent. Steps 2 and 3 are both guaranteed to decrease J; so, the algorithm makes local improvements to an arbitrary clustering until it is no longer possible to do so. To see that Step 3 does decrease J, it is helpful to recall a standard result from linear algebra [29].

2.4. Bio-Materials' Characterization

This step consists of examining the raw data and the clustering results, which will allow us to classify the samples according to their thermal properties. In this way, the data will be analyzed to make better decisions for the design of insulation bio-materials. In addition, it will also be possible to test theories or refute existing models to appreciate the thermal properties of bio-sourced materials well, which play a crucial role in determining their suitability for various applications such as insulation, packaging, and energy storage.

3. Results and Discussion

As mentioned above, in this experiment, we set up seven types of bio-based concrete in order to analyze the physical characteristics of the observed materials. The thermal properties of palm- and esparto-fiber-reinforced concrete are investigated, which include the thermal conductivity, thermal capacity, and thermal diffusivity, as well as the power and temperature used for the characterization. Among the samples, we have a reference concrete (control concrete) and from this reference, we tested the replacement of a percentage of concrete with esparto or palm fibers, ranging from 2% to 5%. The content of these plant fibers in concrete is relative to the weight of cement in order to have less carbon concrete. The samples were then subjected to thermal stresses and different parameters such as conductivity and capacity, and the thermal diffusivity of each type of concrete was collected for further analysis. This provides us with control concrete classes to validate the classification approach based on the data and subsequent analysis. Figure 6 shows the measurements collected during the experiments. For example, on the left side (from top to bottom), we have the electrical power inflicted, the thermal diffusivity and the thermal conductivity for the seven types of concrete. On the right side (top to bottom), we have the heat capacity and the ambient temperature. The thermal conductivity ranges from 0.8 to 1.43 W/m·K, the thermal diffusivity ranges from 0.3 to 1.5 mm^2/s, and the thermal capacity ranges from 1 to 4 $MJ/m^3 \cdot K$.

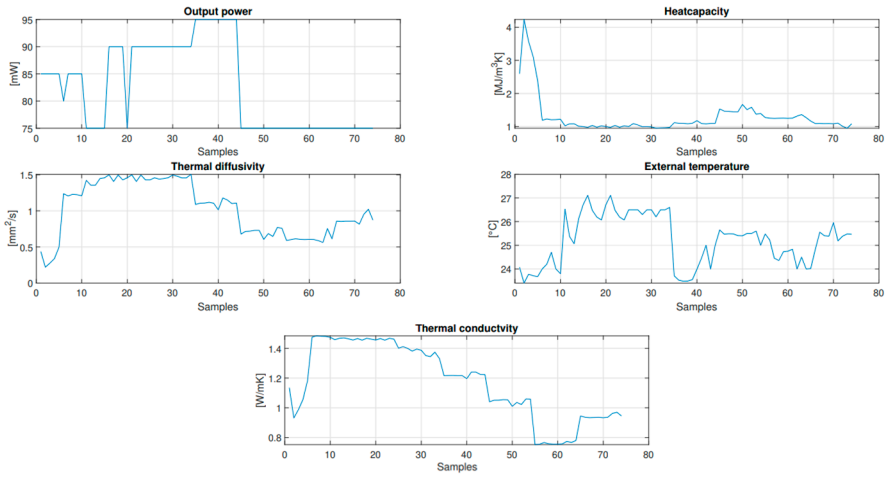


Figure 6. Parameters' behavior; example obtained owing to data collection.

From these measurements, we apply the classification method detailed in the previous section. We tested the power of this technique for the five, six and seven centroids of the concrete classes that will be identified with the k-means tool. So, the last clustering

approach allows us to obtain an insight about the structure of the data and separate the data into homogeneous groups with common characteristics.

Figure 7 shows us the validation of the classification result considering seven, six and five clusters' centroids. The diagonal (in blue) illustrates that the data are well classified with the right type of concrete by considering the characteristics found in the test specimens. For example, it may assign each element of the reference concrete (RC) to the class associated with that type of concrete (last line colored in blue of Figure 7 for k = 7 or 6 or 5). We see in the same figure that with seven centroids, the approach has difficulties in predicting the type of concrete with 2% of the esparto sheets and that there are several measurements that it fails to classify correctly (red and orange cells). That is, for example, the case when we look at the first and last row of the confusion matrix (right side of Figure 7).

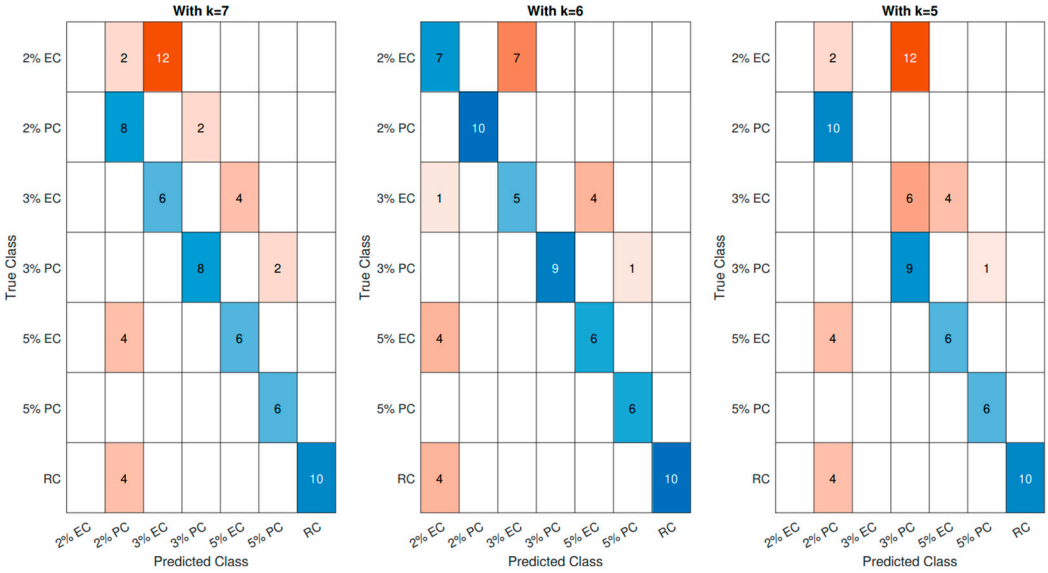


Figure 7. Confusion matrix for different centroid number k = 7, 6 and 5.

To quantify the approach's effectiveness, the overall performance for each centroid is given in Table 2. The results show that the k-means algorithm is able to efficiently classify the reference concrete (RC) with a performance of up to 71%. Also, we see the technique is more accurate by retaining only six centroids which, among other things, allows all the characteristics associated with each type of concrete to be grouped and found. Indeed, whether for k = 7 or k = 5, the technique was not able to predict the typical characteristics of 2% or 3% esparto concrete (EC).

Table 2. Classification efficiency index in percent [%] based on F1score criterion.

Concrete Type	$k = 7$	$k = 6$	$k = 5$
2% esparto concrete (EC)	0	50	0
2% palm concrete (PC)	57.14	71.43	71.43
3% esparto concrete (EC)	42.86	0.3571	0
3% palm concrete (PC)	57.14	64.29	64.29
5% esparto concrete (EC)	42.86	42.86	42.86
5% palm concrete (PC)	42.86	42.86	42.86
Reference concrete (RC)	71.43	71.43	71.43

Based on these performance indicators, we present in Figure 8 the illustration of the classification and characterization of the bio-materials with six groups identified.

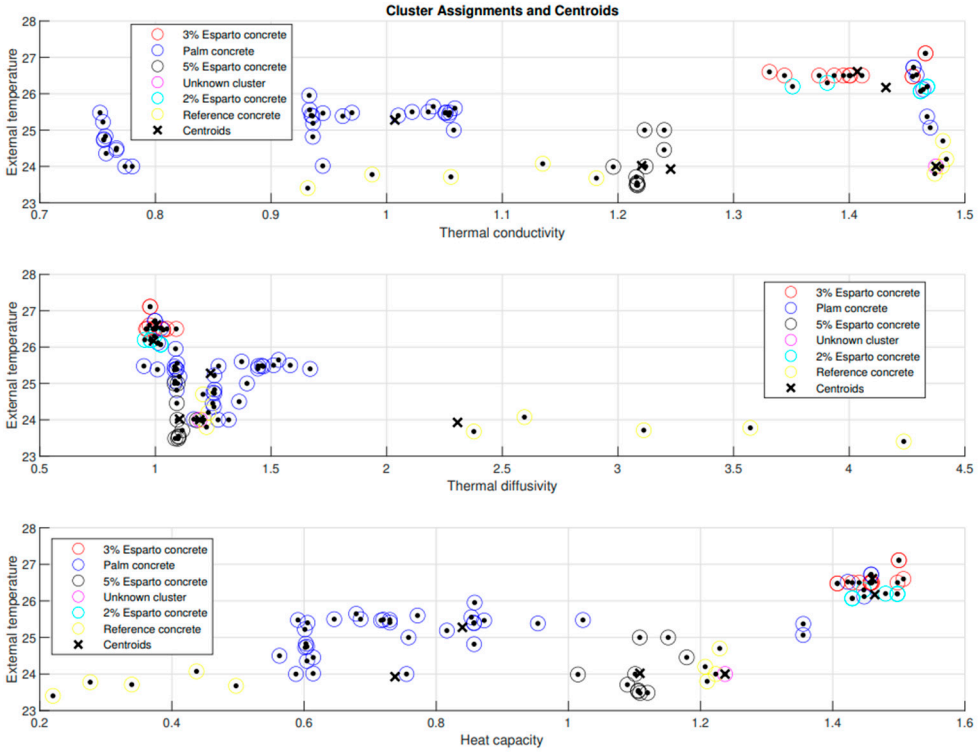


Figure 8. Bio-materials' classification and characterization with six identified clusters.

First, the result shows that varying the content of esparto fibers between 2% and 5% does not make a big difference regarding palm concrete's physical characteristics as all of palm concrete types were regrouped by the *k*-means algorithm into only one cluster (blue circles in the top of Figure 8). Meanwhile, we recommend this type of bio-material for building insulation because the results show that the addition of palm fibers to concrete decreases the thermal conductivity of concrete comparing it with the reference concrete (yellow circles at the top of Figure 8). Indeed, the higher the thermal conductivity, the more heat the material conducts; the lower the thermal conductivity, the more insulating it is. Thus, adding esparto fibers to concrete is not recommended as it significantly increases the thermal conductivity properties of insulation materials (black, red and magenta circles at the top of Figure 8).

Secondly, regarding the thermal diffusivity property (middle of Figure 8), adding bio-materials (without distinction between palm or esparto) reduces the ability of concrete to transfer heat. However, the lower figure reflects the great capacity of bio-materials to accumulate energy in the form of heat, for a given mass, as its temperature increases. This is even more pronounced for esparto fibers when compared to palm fibers. In summary, the choice to use palm and/or esparto in insulation materials would be a trade-off between the performance one wishes to give the material. Indeed, these can improve or deteriorate certain parameters. According to this study, we recommend palm fibers over esparto fibers.

We recognize the difficulty presented by the restricted worldwide accessibility of some categories of bio-based fibers that we investigated in this research. In fact, differences in fiber can greatly affect the classification criteria, which is why we tested two different types of fiber. We showed in this paper that our categorization scheme is workable for these two types of fibers. We shed light on how fiber type affects algorithm parameterization and accuracy, as seen in Table 2. As with any simulation technique, accurate sensitivity analysis and calibration are necessary to properly customize the model for every kind of material.

4. Conclusions

In conclusion, this study underscores the potential of bio-based concrete for building insulation, with a focus on analyzing the thermal properties of palm- and esparto-fiber-reinforced concrete. The findings reveal the following significant insights:

- Thermal conductivity: the addition of palm fibers to concrete exhibits a decrease in thermal conductivity, indicating enhanced insulation capabilities.
- Fiber selection: it is recommended to prioritize the use of palm fibers over esparto fibers for improved material performance, balancing efficiency and sustainability objectives.
- Clustering analysis: the application of k-means clustering proves effective in categorizing concrete types based on their thermal characteristics, though challenges are noted with predicting certain compositions.

Furthermore, the novelty of this research lies in its contribution to both theoretical and practical domains:

- Research novelty: this study sheds light on the thermal behavior of bio-based concrete, offering valuable insights for the development of sustainable building materials.
- Methodological innovation: the utilization of k-means clustering as a classification tool provides a novel approach to categorizing concrete compositions based on thermal properties.

Future research could explore other clustering algorithms or feature engineering techniques to improve model sensitivity for low-fiber concretes.

Author Contributions: Conceptualization, A.A., L.R., N.D. and M.-H.B.; Methodology, A.A., N.D. and M.-H.B.; Software, L.R. and M.-H.B.; Validation, A.A., L.R., N.D. and M.-H.B.; Formal analysis, A.B., K.K. and K.B.; Investigation, A.A., L.R., N.D. and M.-H.B.; Resources, A.A. and N.D.; Data curation, A.A. and N.D.; Writing—original draft, A.A., L.R., N.D., A.B., K.K., K.B. and M.-H.B.; Writing—review & editing, L.R., N.D., A.B., K.K., K.B. and M.-H.B.; Visualization, L.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are contained within the article.

Acknowledgments: This work was supported and benefited from the contribution of the company BTPH-HASNAOUI, in Sidi Bel Abbes, Algeria.

Conflicts of Interest: The authors declare that they have no conflict of interest regarding the content of this article.

References

1. U. N. E. Global Status Report for Buildings and Construction, Available. Global Report. 2021. Available online: https://globalabc.org/roadmaps-buildings-and-construction (accessed on 17 October 2021).
2. Alami, A.; Makhlouf, M.; Lousdad, A.; Khalfi, A.; Benzaama, M.H. Energetic and exergetic analyses of adsorption heat transformer ameliorated by ejector. *J. Braz. Soc. Mech. Sci. Eng.* **2016**, *38*, 2077–2084. [CrossRef]
3. Alami, A.; Boucham, B.; Gouareh, A. Investigation on the Energy Efficiency of a Geo-sol Adsorption Heat Transformer in the Algerian Context. *Int. J. Heat Technol.* **2019**, *37*, 820–830. [CrossRef]
4. Missoum, M.; Hamidat, A.; Loukarfi, L.; Abdeladim, K. Impact of rural housing energy performance improvement on the energy balance in the North-West of Algeria. *Energy Build.* **2014**, *85*, 374–388. [CrossRef]
5. Olukoya Obafemi, A.P.; Kurt, S. Environmental impacts of adobe as a building material: The north Cyprus traditional building case. *Constr. Mater.* **2016**, *4*, 32–41. [CrossRef]
6. Lecompte, T. Biosourced Materials for Construction and Temporary Carbon Storage. Engineering Technique. 2021. Available online: https://www.techniques-ingenieur.fr/base-documentaire/construction-et-travaux-publics-th3/environnement-et-construction-42552210/materiaux-bio-sources-pour-le-batiment-et-stockage-temporaire-de-carbone-c8124/conclusion-c8124niv10006.html/ (accessed on 17 October 2023).
7. Boutin, M.; Flamin, C.; Quinton, S.; Gosse, G. Analyse du cycle de vie: Compounds thermoplastiques chargés fibres de chanvre et Mur en béton de chanvre banché sur os-sature bois. *Rapport d'Étude INRA Lille Réf. MAP* **2005**, *4*, B1.
8. Ahmad, M.R.; Chen, B.; Haque, M.A.; Oderji, S.Y. Multiproperty characterization of cleaner and energyefficient vegetal concrete based on one part geopolymer binder. *J. Clean. Prod.* **2020**, *253*, 119916. [CrossRef]

9. Holman, J.P. Heat Transfer; Tata McGraw-Hill Education. 2008. Available online: https://archive.org/details/JackP.HolmanHeatTransferTenthEdition/page/n5/mode/2up (accessed on 17 October 2023).
10. Charai, M.; Sghiouri, H.; Mezrhab, A.; Karkri, M. Thermal insulation potential of nonindustrial hemp (*Moroccan cannabis* sativa L.) fibers for green plaster-based. building materials. *J. Clean. Prod.* **2021**, *292*, 126064. [CrossRef]
11. Shang, Y.; Tariku, F. Hempcrete building performance in mild and cold climates: Integrated analysis of carbon footprint, energy, and indoor thermal and moisture buffering. *Build. Environ.* **2021**, *206*, 108377. [CrossRef]
12. Mehravar, M.; Veshkini, A.; Veiseh, S.; Fayaz, R. Physical Properties of Straw Bale and its Effect on Building Energy Conservation and Carbon Emissions in Different Climatic Regions of Iran. *Energy Build.* **2021**, *254*, 111559. [CrossRef]
13. Bruijn, P.S.; Donarelli, A.; Balksten, K. Full-scale studies of improving energy performance by renovating historic Swedish timber buildings with hemplime. *Appl. Sci.* **2019**, *9*, 2484. [CrossRef]
14. Mai, T.H.; Wurtz, E.; Collet, F.; Tran, A.D.; Maalouf, C. Transient hygrothermal behaviour of a hemp concrete building envelope. *Energy Build.* **2010**, *42*, 1797–1806. [CrossRef]
15. Bouasria, M.; El Mendili, Y.; Benzaama, M.H. Valorisation of stranded Laminaria digitata seaweed as an insulating earth material. *Constr. Build. Mater.* **2021**, *308*, 125068. [CrossRef]
16. El Mendili, Y.; Bouasria, M.; Benzaama, M.H.; Khadraoui, F.; Guern, M.; Chateigner, D.; Gascoin, S.; Bardeau, J.F. Mud-Based Construction Material: Promising Properties of French Gravel Wash Mud Mixed with Byproducts, Seashells and Fly Ash as a Binder. *Materials* **2021**, *14*, 6216. [CrossRef]
17. Ramvir, S.; Bhoopal, R.S.; Kumar, S. Prediction of effective thermal conductivity of moist porous materials using artificial neural network approach. *Build. Environ.* **2011**, *12*, 2603–2608. [CrossRef]
18. Asgarkhani, N.; Kazemi, F.; Jakubczyk-Gałczyńska, A.; Mohebi, B. Seismic response and performance prediction of steel buckling-restrained braced frames using machine-learning methods. *Eng. Appl. Artif. Intell.* **2024**, *128*, 107388. [CrossRef]
19. Kazemi, F.; Shafighfard, T.; Yoo, D.Y. Data-Driven Modeling of Mechanical Properties of Fiber-Reinforced Concrete: A Critical Review. *Arch. Comput. Methods Eng.* **2024**, *31*, 2049–2078. [CrossRef]
20. Bouasria, M.; Benzaama, M.H.; Pralong, V.; El Mendili, Y. Mechanical and hygrothermal performance of flyash and seashells concrete: In Situ experimental study and smart hygrothermal modeling for normandy climate conditions. *Arch. Civ. Mech. Eng.* **2022**, *22*, 100. [CrossRef]
21. Chen, L.; Zhan, C.; Li, G.; Zhang, A. An artificial neural network identification method for thermal resistance of exterior walls of buildings based on numerical experiments. In *Building Simulation*; Tsinghua University Press: Beijing, China, 2019. [CrossRef]
22. Sharo, A.A.; Rababah, S.R.; Taamneh, M.O.; Aldeeky, H.; Akhrass, H. Mathematical modelling for predicting thermal properties of selected limestone. *Buildings* **2022**, *12*, 2063. [CrossRef]
23. Cengiz, O.; Alper, T. Comparison between ANFIS and ANN for estimation of thermal conductivity coefficients of construction materials. *Sci. Iran.* **2015**, *22*, 2001–2011.
24. Hajihassani, M.; Marto, A.; Khezri, N.; Kalatehjari, R. Indirect measure of thermal conductivity of rocks through adaptive neurofuzzy inference system and multivariate regression analysis. *Measurement* **2015**, *67*, 71–77. [CrossRef]
25. Wang, D.; Horton, R.; Zhao, T.; Liu, S.; Xu, J.; He, H.; Liu, G. Comparative analysis of seven machine learning algorithms and five empirical models to estimate soil thermal conductivity. *Agric. For. Meteorol.* **2022**, *323*, 109080. [CrossRef]
26. Grieu, S.; Faugeroux, O.; Traore, A.; Claudet, B.; Bodnar, J.L. Artificial intelligence tools and inverse methods for estimating the thermal diffusivity of building materials. *Energy Build.* **2011**, *43*, 543–554. [CrossRef]
27. Gustafsson, S.E. Transient plane source techniques for thermal conductivity and thermal diffusivity measurements of solid materials. *Rev. Sci. Instrum.* **1991**, *62*, 797–804. [CrossRef]
28. Stuart, P. Lloyd. Least squares quantization in pcm. *IEEE Trans. Inf. Theory* **1982**, *28*, 129–136.
29. Har-Peled, S.; Sadri, B. How fast is the k-means method? *Algorithmica* **2005**, *41*, 185–202. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.