XLIII CILAMCE
21–25, November, 2022
FOZ DO IGUAÇU – PR – BRAZIL

A methodology to predict the effective thermal conductivity of a granular assembly using deep learning

Osvaldo D. Quintana-Ruiz[1], Eduardo M.B. Campello[1]

[1]*Dept. Structural and Geotechnical Engineering, Polytechnic School, University of São Paulo*
P.O. Box 61548, 05424-970, São Paulo, SP, Brazil
quintana.ruiz@usp.br, campello@usp.br

Abstract. In this work, an Artificial Neural Network (ANN) is employed to predict the effective (i.e., bulk) thermal conductivity of a granular assembly. The ANN is trained with the help of computed thermal conductivities of various different assemblies, obtained through several simulations with our in-house DEM (Discrete Element Method) code. Convection and radiation are not considered as to isolate the conduction problem and allow for a better estimate of the assembly's effective response. The methodology enables the effective thermal conductivity of a granular assembly over a wide range of parameter values, including particles' size and their material's conductivities.

Keywords: machine learning, artificial neural networks, particles, thermal conductivity, granular materials, discrete element method (DEM)

1 Introduction

Granular systems display different characteristics on multiple spatial and temporal scales. While the response of granular materials (from a macroscopic scale) may at times appear to be similar to that of a continuum, when seen from the mesoscale, they are essentially a system of multiple (and interacting) discrete particles. In addition to physical experiments, computational simulation and numerical approaches can be useful to a better understanding of the phenomena involved in those materials. Moreover, regarding to heat transfer in dry particulate media, it can be found in the literature that the conduction contribution is considered the most significant form of heat transfer. It depends on the conductivity of the particles' constituent material, the inter-particle contacts (which provide one source of coupling between the thermal and mechanical fields) and the overall structure of the particle packing. The thermal conductivity of particulate beds is an important property for many industrial handling processes as well as for storage of particulate materials, where parameters such as temperature, particle size, and pressure (to cite just a few) significantly affect the thermal response. The significance of thermal conduction can be present in applications like the usage of geothermal energy, buried earth-structures, radioactive waste disposal [1] or even simply storing particles in a silo after drying [2]. The effective thermal conductivity (k_{eff}) of a granular bed depends on various parameters pertaining to both bulk material and microstructural properties. We can find in the literature several analytical and numerical models for the prediction of k_{eff} of granular assemblies [2] [3] [4] [5] [6] [7], however, they are mostly restricted to certain (simple) types of particle packings, whereas the numerical ones, although more general, are usually relatively computationally expensive and time-consuming. On the other hand, ANN, as a deep learning method, have received a great deal of attention over the last decade and demonstrated excellent performance in material property prediction and material design. Deep learning is a subset of machine learning and refers to any ANN with more than two hidden layers. The aim of the present work is to propose a methodology based on ANNs to predict the effective thermal conductivity of a granular assembly over a wide range of parameter values, in order to reduce the dependency on time-consuming simulations that takes to predict this parameter numerically. The text is organized as follows: in Section 2 we briefly present the reference model that we will use to compare our results; in Section 3 we introduce our ANN model; in Section 4 we present our results along with a brief discussion; and in Section 5 we close the paper with our conclusions.

2 Conduction through a 3D particle assembly

To create the training dataset for the ANN from numerical simulations, we perform seven DEM simulations based on a model problem of a previous works of the authors [8] [9], where a three-dimensional particle assembly of $NP = 3500$ particles are randomly packed under gravity within a cubic box, and subjected to a gradient temperature of 1100 K in its x – direction, as depicted in Figure 1.

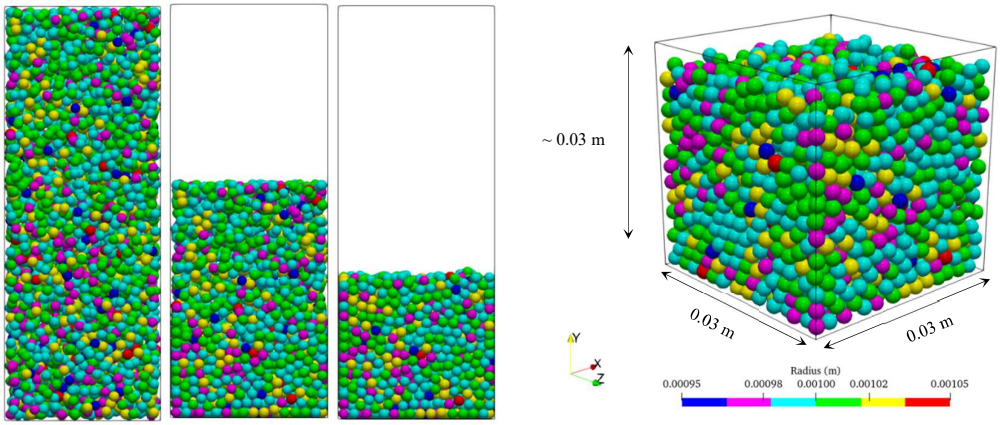

Figure 1. Problem definition. Left: 2D view of the stages during the deposition of the particles. Right: perspective view of the final deposition stage (final configuration)

The box has side dimensions of $L = 0.03\,\text{m}$, whereas the particles have radii following a Gaussian distribution with mean $\overline{r} = 0.001\,\text{m}$ and stv. dev. of 0.00001666 m (the distribution is truncated at three std. dev. from the mean, see Figure 2, left). The thermal gradient (ΔT) is enforced by holding the temperatures of the two opposite walls of the box in the x – direction at 1500 K and 400 K indefinitely, respectively, with the particles' initial temperatures being set at 300 K. Convection and radiation are not considered as to isolate the conduction problem. Other data are summarized in Table 1.

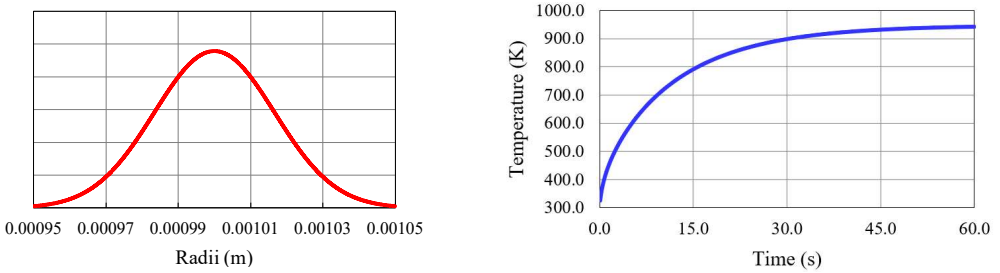

Figure 2. Left: Gaussian distribution of the particles' radii. Right: Average temperature of simulation 7

Table 1. Values used in the simulation

Parameter	Value
Particles' material density	1.000 kg/m^3
Particles' elastic modulus and Poisson's ratio	0.001 GPa and 0.30
Contact and friction damping rate	0.1
Friction and rolling resistance coeffs. between particles	$\mu_s = \mu_d = 1.0$ and $\mu_r = 0.9$
Friction and rolling resistance coeffs. between particles and walls	$\mu_s = \mu_d = 0.65$ and $\mu_r = 0.9$
Simulation time and time-step size[1]	$t_f = 60$ s and $\Delta t = 5.0 \times 10^{-5}$

[1] The criterion adopted is the same of [8] [9], which is based on Hertz' contact theory (see Johnson [11]).

By releasing the system at $t = 0$ and computing the evolution of heat over time, we obtain the effective thermal conductivity in the $x-$direction through

$$k_{\text{eff}} = \frac{\Delta Q}{\nabla T A \Delta t} \quad \text{where} \quad \Delta Q = \sum m_i C_i \Delta\theta_i \quad \text{with} \quad i = 1,..., NP, \tag{1}$$

is the total heat that is provided to the system, in which m_i is the particles' mass, C_i is the particles' specific heat, $\Delta\theta_i$ is the particles' thermal change (w.r.t. the initial time), $\nabla T = \Delta T / L$ is the applied thermal gradient, $A = L \times L$ is the cross-sectional area to the flow and Δt is the time interval considered. Between the seven simulations, we vary two parameters for the training dataset namely, the specific heat and the thermal conductivity (k_i) of the particles. The values of k_{eff} obtained from each simulation through equation (1) are summarized in Table 2. For a typical illustration, we choose simulation seven to show the average temperature of the system's configuration (Figure 2, right) and a sequence of snapshots at selected time instants (Figure 3).

Table 2. Values obtained in our DEM simulations and used to train the ANN

Simulation	C_i (J/kgK)	k_i (W/mK)	k_{eff} (W/mK)
1	100.0	100.0	0.5277
2	110.0	110.0	0.5805
3	120.0	120.0	0.6333
4	130.0	130.0	0.6860
5	140.0	140.0	0.7388
6	150.0	150.0	0.7916
7	160.0	160.0	0.8444

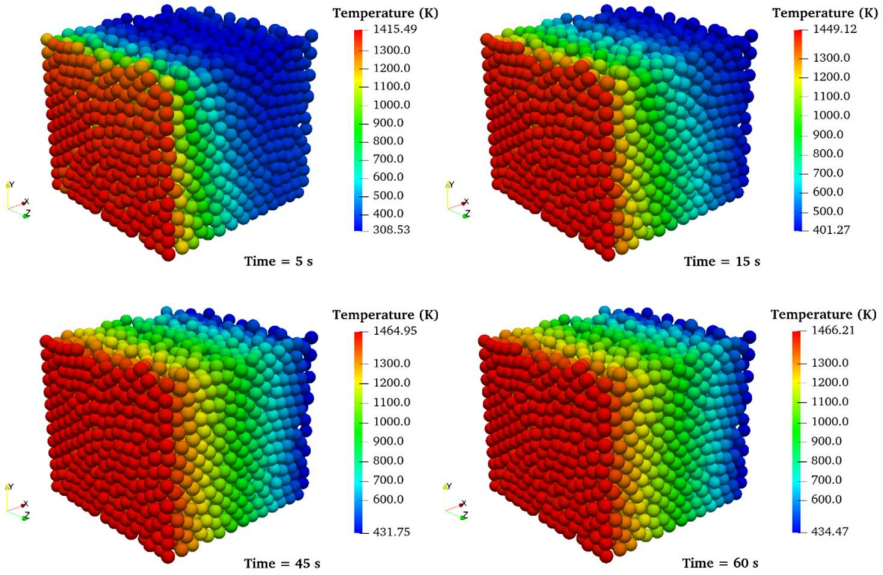

Figure 3. Conduction through a 3D particle assembly. Snapshots from simulation 7. Sequence is from left to right, top to down

3 ANN model

A typical ANN architecture consists of a collection of neurons distributed in several layers, where each layer has a set of neurons connected with the neurons of the previous and next layers through weighted functions. These layers are classified into three main categories: the input layer, which corresponds to the information or experimental observation, the hidden layer and the output layer (which gives the outputs). The information in each level is transformed with the aid of the activation function and further transferred to the next layer sequentially.

Here, the activation function maps the entry numbers into a small range (e.g., 0 to 1) to convert arbitrary real values into ones that can be interpreted as a probability. The neurons in each level are represented as components of vectors where the input layer is propagated through the ANN to the output layer, as depicted in the following expressions

$$\begin{aligned}
\boldsymbol{\psi}_1 &= \varphi_1^*\left(\boldsymbol{\psi}_0\boldsymbol{W}_1 + \boldsymbol{b}_1\right)\\
\boldsymbol{\psi}_2 &= \varphi_2^*\left(\boldsymbol{\psi}_1\boldsymbol{W}_2 + \boldsymbol{b}_2\right)\\
&\vdots\\
\boldsymbol{\psi}_i &= \varphi_i^*\left(\boldsymbol{\psi}_{i-1}\boldsymbol{W}_i + \boldsymbol{b}_i\right)\\
&\vdots\\
\boldsymbol{\psi}_N &= \varphi_N^*\left(\boldsymbol{\psi}_{N-1}\boldsymbol{W}_N + \boldsymbol{b}_N\right), \quad \text{where } i = 1,\ldots,N
\end{aligned} \tag{2}$$

In equation (2), $\boldsymbol{\psi}_i$ represents each layer of the network with $\boldsymbol{\psi}_0$ as the input layer, φ_i^* is the activation function, $\boldsymbol{b}_i$ is the bias, $\boldsymbol{W}_i$ are the weight matrices and N is the total number of layers . The number of rows in $\boldsymbol{W}_i$ must be equal to the number of neurons in $\boldsymbol{\psi}_{i-1}$ and the number of columns equal to the number of neurons in $\boldsymbol{\psi}_i$. Training of the ANN means the evaluation of proper weights and biases to the transformation functions of the neurons, and this is achieved by minimizing the error (E) by the expression (3). Here, E , also known as the cost function (J), is represented by the mean square error (MSE)

$$E = \frac{1}{N}\sum_{j=1}^{NOb}\left(\boldsymbol{y}_j - \hat{\boldsymbol{y}}_j\right)^2 \tag{3}$$

where $\hat{\boldsymbol{y}}_i = \boldsymbol{\psi}_N$, $\boldsymbol{y}_i$ corresponds to a known output value in the training dataset and *NOb* represents the number of observations in the dataset. The main task in the training is to obtain the smallest value of E by updating the weights and biases and this can be performed through the backpropagation (BP) algorithm. Basically, the BP algorithm computes the gradient of the cost function applied to the output of the ANN w.r.t. the parameters in each layer. The design of the ANN architecture usually varies according to the type of problem and performance of the network.

3.1 Creation and training of the ANN

In the present work, we opted for a multilayer perceptron (MLP) neural model implemented in an in-house code written in Fortran with 2, 5 and 1 neurons in each layer where the two neurons in the input layer corresponds to each parameter that we are varying in the analysis (C , k), see Figure 4. To get the weights and biases that take (J) to minimum, we use the gradient descent algorithm where the error is propagated through the network in the backward direction

$$\begin{aligned}
W &= W - \eta\frac{\partial}{\partial W}J(W),\\
b &= b - \eta\frac{\partial}{\partial b}J(b).
\end{aligned} \tag{4}$$

In the equation (4), η corresponds to the learning rate. Since our dataset analysis is small, we decided to use the batch gradient descent where the gradient will be calculated using all the samples of the dataset. The sigmoid (logistic) function was adopted as the activation function for each layer (hidden and output). As we mentioned in the previous section, we train the ANN with the values of k_{eff} obtained through several simulation with our DEM code where the results were presented in Table 2. We remark that we opted for only one hidden layer, taking into account the small size of our problem, but the proposed methodology is generalized for multiple hidden layers.

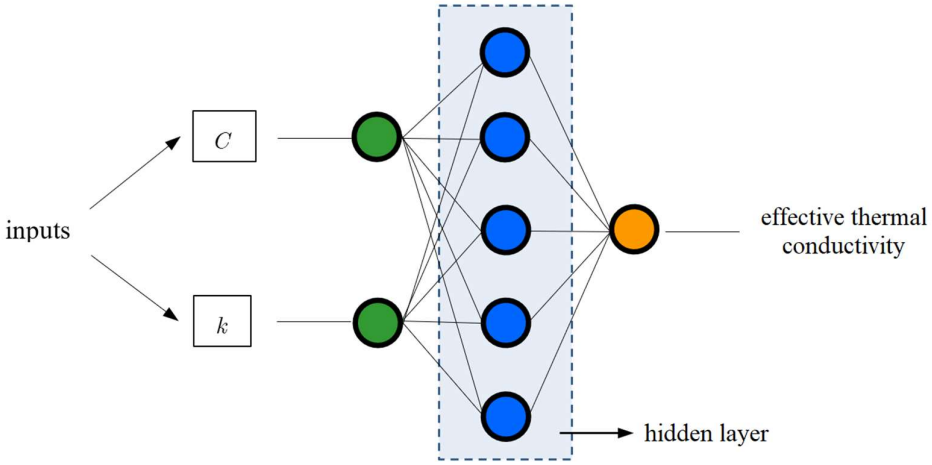

Figure 4. Architecture used to train the ANN

4 Results and discussion

To test our methodology, we run other seven (new) simulations with different values of specific heat and thermal conductivities to obtain new values of k_{eff} and then, we compare the results with those obtained from our trained ANN. The values of k_{eff} obtained from the ANN model and the simulations are summarized in Table 3.

Table 3. Values used to test our methodology for the example with $NP = 3500$

Simulation	C_i (J/kgK)	k_i (W/mK)	$k_{\text{eff,Simulation}}$ (W/mK)	$k_{\text{eff,ANN}}$ (W/mK)
1	105.0	105.0	0.55415	0.55400
2	115.0	115.0	0.60692	0.61500
3	122.0	122.0	0.64387	0.65800
4	137.0	137.0	0.72303	0.74000
5	145.0	145.0	0.76525	0.77600
6	154.0	154.0	0.81275	0.80900
7	165.0	165.0	0.87080	0.84100

Figure 5 shows the regression plot between the ANN model and the numerical simulation presented in Table 3, and Figure 6 plots the cost function w.r.t. the epochs (iterations) where the total epochs in the analysis was 60.000. The R – squared (coefficient of determination) measures how the regression model predicts in comparison with a reference data, accordingly

$$R^2 = 1 - \frac{\sum\left(k_{\text{eff,ANN}} - k_{\text{eff,Simulation}}\right)^2}{\sum\left(k_{\text{eff,Simulation}} - \bar{k}_{\text{eff,Simulation}}\right)^2}. \tag{5}$$

where $\bar{k}_{\text{eff}}$ is the mean of the observed data (results from the simulations). A value of R^2 close to 1 indicates that the prediction is in good agreement with the actual data. As we can see from Figure 5 and Table 3 the results are in good agreement with those from the numerical simulations. We now test again our methodology but now with a different configuration, in this case the number of particles was $NP = 1000$ within a square box of side $L = 0.02$ m, and the particles have radii following a monosized distribution with $r = 0.001$ m. Other parameters are the same of the previous example, see Table 1.

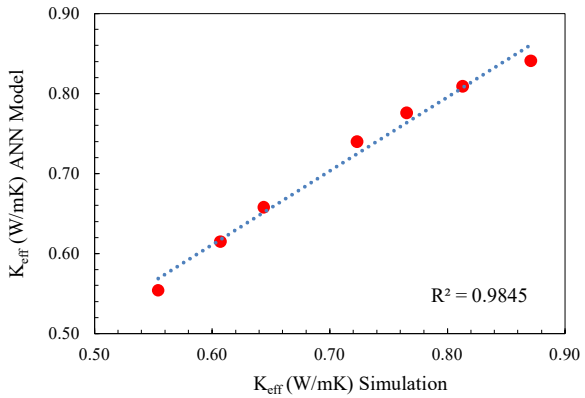

Figure 5. Regression plots for the results from simulations and predictions of ANN

We train the net with different values of specific heat and thermal conductivities (Table 4) and then, we test the model with new values and compare with the trained ANN, see Table 5. Figure 7 shows the regression between the simulated and trained data. With regards to the simulation times, the computation time required for the simulation of the example of $NP = 3500$ is about 3.95 s per each 0.01 s of the problem´s duration (which, for a problem´s duration of $t_f = 60$ s, implies 23700 s) in a standard, single processor laptop computer (at 2.30 GHz) with no parallelization nor usage of the graphics processing unit (GPU), in comparison with 0.828 s that takes to train the ANN with 60.000 epochs. This means a reduction of about 5 orders of magnitude, which we find quite remarkable.

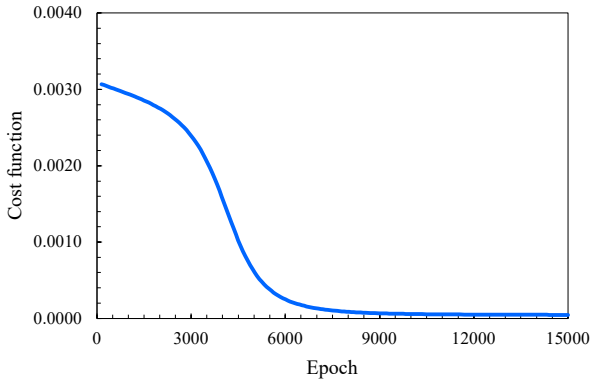

Figure 6. Cost function as a function of epoch for the example of $NP = 3500$

Table 4. Values used to train the ANN for example of $NP = 1000$

Simulation	C_i (J/kgK)	k_i (W/mK)	k^* (W/mK)
1	100.0	100.0	0.24742
2	110.0	110.0	0.27216
3	120.0	120.0	0.29690
4	130.0	130.0	0.32164
5	140.0	140.0	0.34639
6	150.0	150.0	0.37113

Table 5. Values used to test our methodology for example of $NP = 1000$

Simulation	C_i (J/kgK)	k_i (W/mK)	$k_{\text{eff,Simulation}}$ (W/mK)	$k_{\text{eff,ANN}}$ (W/mK)
1	107.0	107.0	0.26474	0.2640
2	114.0	114.0	0.28206	0.2830
3	128.0	128.0	0.31670	0.3200
4	132.0	132.0	0.32659	0.3300
5	146.0	146.0	0.36123	0.3600
6	151.0	151.0	0.37360	0.3690

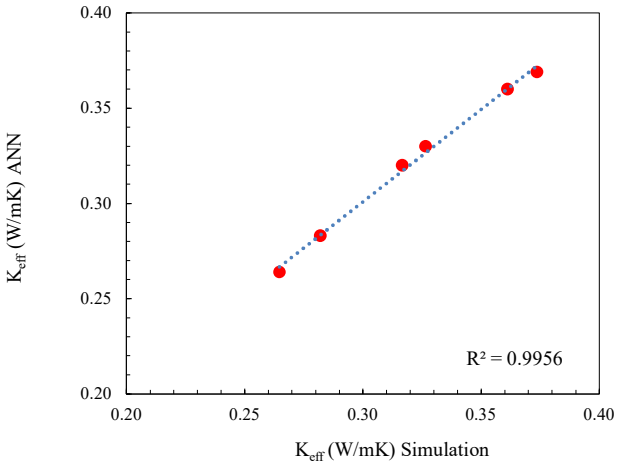

Figure 7. Regression plot for the example of $NP = 1000$

5 Conclusions

In this work, a simple ANN was implemented as a methodology for the prediction of the effective thermal conductivity of a granular assembly. The ANN was trained using the results of a model problem performed in our DEM code following the formulation of [8] [9]. From the obtained results, the methodology proved to work well for the purposes and type of problem we are interested in and proved to substantially reduce the computation time required to numerically estimate the effective conductivity of a granular bed once the net is trained. This truly motivates us to pursue in its advancement and extension. We remark that this work is the first stage of an on-going research and advancements over the present methodology shall appear later in a forthcoming paper. In this regard, the incorporation of more input parameters for training and different architectures of the networks are currently under development by the authors.

Acknowledgements. First author acknowledges the postdoctoral funding from the Government of Paraguay through the National Scholarship Program for Postgraduate Studies Abroad Don Carlos Antonio López (BECAL), under the grant 72/2022. Second author acknowledges support by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), Brazil, under the grant 313046/2021-2.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] A. R. Peeketi, M. Moscardini, A. Vijayan, Y. Gan, M. Kamlah and R. K. Annabattula, "Effective thermal conductivity of a compacted pebble bed in a stagnant gaseous environment: An analytical approach together with DEM," *Fusion Engineering and Design,* no. 130, pp. 80-88, 2018.

[2] T. S. Yun and T. M. Evans, "Three-dimensional random network model for thermal conductivity in particulate materials," *Comput Geotech,* no. 37, pp. 991-998, 2010.

[3] W. Dai, S. Pupeschi, D. Hanaor and Y. Gan, "Influence of gas pressure on the effective thermal conductivity of ceramic breeder pebble beds," *Fusion Eng Des,* no. 118, pp. 45-51, 2017.

[4] M. Panchal, C. Kang, A. Ying and P. Chaudhuri, "Experimental measurement and numerical modeling of the effective thermal conductivity of lithium metatitanate pebble bed," *Fusion Eng Des,* no. 127, pp. 34-39, 2018.

[5] R. Bauer and E. U. Schlunder, "Part I: effective radial thermal conductivity of packings in gas flow. Part II: thermal conductivity of the packing fraction without gas flow," *Int Chem Eng,* no. 18, pp. 181-204, 1978.

[6] G. Weidenfeld, Y. Weiss and H. Kalman, "A theoretical model for effective thermal conductivity (etc) of particulate beds under compression," *Granul Matter,* no. 6(2-3), pp. 121-129, 2004.

[7] T. I. Zohdi, Electromagnetic Properties of Multiphase Dielectrics. A Primer on Modeling, Theory and Computation, vol. Lecture Notes in Applied and Computational Mechanics. Volume 64, Springer, 2012.

[8] O. D. Quintana-Ruiz and E. Campello, "A thermo-mechanical formulation for the modeling of discrete particle systems," in *Ibero-Latin American Congress on Computational Methods in Engineering (XL CILAMCE)*, Natal, Brazil, 2019.

[9] O. D. Quintana-Ruiz and E. Campello, "A coupled thermo-mechanical model for the simulation of discrete particle systems," *Journal of the Brazilian Society of Mechanical Sciences and Engineering,* vol. 42, no. 378, pp. 387(1-21), 2020.

[10] R. Patrick, M. Nicodemi, R. Delannay, P. Ribière and D. Bideau, "Slow relaxation and compaction of granular systems," *Nature Materials,* vol. 4, pp. 121-128, 2005.

[11] K. L. Johnson, Contact Mechanics, Cambridge: Cambridge University Press, 1985.