DEM Calibration Approach: Random Forest

A V Boikov, R V Savelev and V A Payor

Saint-Petersburg mining university, 21 line, Saint-Petersburg, 2, 199106, Russia

e-mail: boikov_av@mail.ru

Abstract. A lot of researchers are developing new DEM parameters calibration approaches based on an experiment plan or the use of learning algorithms. This research is aimed at improving iterative algorithms frequently used for calibration. The big time consumption as a main problem of iterative algorithms is questioned. It is proposed to use Random forest algorithm to determine DEM parameters impact on the measured bulk responses. Measured responses are the parameters obtained by image processing using a technical vision system. As a result of 200 experiments processing, DEM parameters impact values on each bulk response were generated and presented as histograms. Obtained results were interpreted on the basis of the bulk material behavior and its physical properties. There is a discussion on the possibility of developing a universal DEM parameters calibration method based on the iterative algorithm.

1. Introduction

The increase in computing power allowed discrete element method (DEM) to become a widely used tool for bulk materials behavior modeling [1-3]. When user specifies the model parameters in DEM software, the problem of input parameters adequacy (proper calibration) arises. Adequacy means that calibrated model behavior corresponds to actual bulk material responses. The main problem is a calibration of static ($sf_pb$) and dynamic ($df_pb$) friction or restitution ($cor_pb$) coefficients for particle-particle and particle-boundary interaction and also rolling resistance ($rr$) when modeling spheres. Although these parameters can be measured directly using special techniques, such operations are difficult and time-consuming. Moreover, measured values of the parameters do not necessarily correspond to the results in the model, since these parameters are code-dependent [4].

The problem of DEM parameters calibration is severe and many researchers offer solutions for the calibration of bulk material with pre-known properties [5] or conceptual solutions suitable for arbitrary bulk material [6-9]. All calibration methods are usually divided into a bulk calibration approach and a direct measurement approach [10]. However, at the moment there is no universal method that allows one easily and time-effectively to calibrate absolutely any bulk material. As a result, iterative algorithms are traditionally used for bulk materials calibration in a real-case scenario [11, 12]. The iterative algorithms essence is to measure the bulk response (ex. repose angle or expiration time) and then consistently vary DEM parameters until the model bulk responses values become the same as the measured values. As a rule, the sequence of parameters variation is random, and parameters value selection is limited only by setting the variation range. This is the reason why iterative algorithms are very often time-consuming.

Iterative process speed-up requires information about DEM parameters impact on the measured bulk responses. The solution of the problem requires the proper mathematical apparatus (for example, conduct the design of experiment). This research aims to use the Random forest algorithm to evaluate...
the DEM parameters impact on the bulk responses. The original idea is to measure bulk responses using a technical vision system (TVS).

2. Mathematical model

The random forest algorithm is the machine learning methods. It is represented by a multitude of decision trees. Decision tree sensitivity is determined by bias – the average residual value of decision tree, coming out from the ideal value. Each of the decision trees is built independently and the others are built on a randomly formed subset \( m \) from the training set \( M \).

The construction of a single decision tree is carried out according the following algorithm:

1. Splitting training dataset \( (M) \) into subsets \( (m) \) of \( n \) samples.
2. Decision tree construction is performed for a specific subset \( (m) \).
3. Continuous leaf nodes generation for all samples in training subset \( (m) \) without pruning.
4. Each tree in forest gives feature a classification, so called “vote” for some class. The forest chooses the classification that have the most votes over all trees in the ensemble.

Random forest can be used for feature importance measuring in regression analysis. The forest prediction error is calculated by out-of-bag error estimation (OOB). Let \( X_{ni} \) be a bootstrapped dataset for some tree \( bn \). Bootstrapping is a method for training set increasing by random data items duplicating with replacement. In this case some samples may be presented in training set twice and more, while the others be absent at all. OOB, is calculated by:

\[
OOB = \sum_{i=1}^{l} L(y_i, \frac{1}{\sum_{n=1}^{N} x_i \in X_{ni}} \sum_{n=1}^{N} [x_i \notin X_{ni}] b_n(x_i)),
\]

where \( L(y, z) \) is a loss function, \( y_i \) – response on i-th training subset sample. Feature impotence is calculated as the averaged OOB value before and after mixing, normalized to standard deviation. This allows to use Random forest for features (DEM parameters) impact estimation.

3. Preparation and handling the experiment

The DEM parameters impact on bulk responses of a gravel-like bulk material is evaluated using Random forest algorithm. It was decided to simulate spherical particles with 8 mm diameter to simplify the calculations. Hertzian Spring Dashpot + Mindling-Deresiewicz contact model is implemented in Rocky DEM software [15]. The remaining parameters are presented in table 1. The calibrated DEM parameters are \( sf_{pp} \), \( sf_{pb} \), \( df_{pp} \), \( df_{pb} \), \( cor_{pp} \), \( cor_{pb} \). Rolling resistance is also used because the simulated particles are spheres. The experiments were carried out at a specialized rig that was previously used in another research [16, 19].

The bulk material behavior on the rig is captured using technical vision algorithms [17]. The developed technical vision system allows to extract up to 43 different parameters. 8 from 43 most valuable parameters were chosen: the angle of rupture and repose, the expiration time from the funnel and the height and the base length of the formed pile (figure 1). In addition, one of the measured response is the visual image of the “parabola”. Interpretation of this response is performed at a certain moment, when the shape of the bulk material forming the funnel can be described by the parabola equation \( y = ax^2 + bx + c \) (figure 2). The investigated coefficient is the parabola coefficient \( A \).

Data from 200 numerical experiments were used to evaluate the DEM parameters impact on the measured bulk responses. The experiments formation for random forest training was carried out randomly in the DEM parameters value range from 0.1 to 0.9. No data post-processing is required, since all DEM parameters have a real value. The random forest algorithm is provided by scikit-learn library (Python 3.6).
Table 1. Simulation and bulk material parameters

| Parameter                       | Value            |
|---------------------------------|------------------|
| Poisson’s ratio                 | 0.3              |
| Young modulus                   | $10^6$ kPa       |
| Density, kg/m$^3$               | 1300             |
| Shape                           | Spherical        |
| Particle size distribution      | 100% 8 mm diameter |
| Contact model                   | Non-linear       |
| Gravity acceleration            | 9.81 m/s$^2$     |

Figure 1 – Measured bulk responses interpretation using TVS

Figure 2 – Visual image “parabola” interpretation
4. Results and interpretation

Experiment results analysis using Random forest regression algorithm gave DEM parameters impact values on each of the 8 bulk responses. The values are presented as histograms and are grouped as follows: figure 3 shows values on the bulk responses of the top pile (angle of rupture, top base length, top height), figure 4 is the same for the bottom pile (angle of repose, bottom base length, bottom height) and figure 5 shows values for expiration time and parabola A.

DEM parameters impact values were summed separately for the top and the bottom piles. The results of the combined impact are presented in tables 2 and 3.

| Parameter | sf_pp | rr | df_pb | df_pp | cor_pb | cor_pp | sf_pb |
|-----------|-------|----|-------|-------|--------|--------|-------|
| Value     | 0.94  | 0.92| 0.40  | 0.29  | 0.16   | 0.15   | 0.15  |
Comparative analysis yields the expected results. The top pile parameters values are primarily dependent on a static friction for particle-particle interaction. Such a result can be caused by the angle of rupture formation mechanics [18]. The particles mostly do not acquire kinetic energy (a large mass of material remains stationary), and the motion occurs only on the surface of the material. Rolling resistance strong impact has the same explanation. In addition, experiments were held with a spherical particles. The strong impact of RR on the bottom pile parameters confirms the importance of RR calibration when simulating spheres. At the same time, the bottom pile parameters are more dependent on df_pp and df_pb. Before the bottom pile formation, a whole mass of particles flows out of the funnel and then hits the surface of the rig. The collision with the surface can be seen from the greater cor_pp and cor_pb impact on the bottom pile parameters, rather than on the top.

Similarly, DEM parameters impact values on all measured bulk responses were summed (table 4). Table 4 shows in addition the DEM parameters appearance frequency in a three most affect parameters on the each bulk response.

| Parameter | df_pp | df_pb | rr | cor_pp | cor_pb | sf_pb | sf_pp |
|-----------|-------|-------|----|--------|--------|-------|-------|
| Value     | 0.98  | 0.55  | 0.48 | 0.37   | 0.24   | 0.2   | 0.18  |

The values presented in table 4 give an overview of the DEM parameters impact on all measured bulk responses together. The values for individual parameter are confirmed by the particles specificity (spheres were modeled) and by the bulk material behavior at the rig. Most of the simulation time the bulk material was in motion, flowing out one sector into another (a strong impact of dynamic friction). The collision of particles had a significant effect only to the parabola A (a weak impact of coefficient of restitution). Strong sf_pp impact is explained by strong influence on the top pile parameters values, where the angle of repose is formed.

5. Conclusion

The obtained results of the DEM parameters values impact on the measured bulk responses using the Random forest algorithm can be used to design universal iterative algorithm for DEM parameters calibration. The algorithm should start from varying DEM parameters having a greater impact on the entire system (table 4). Then variation should have a sequence that depends both on the DEM parameters impact value and the difference between the actual measured parameter and the result obtained in the model. For the rig and the bulk material used in the research, it will be fair to start the calibration with variation of rolling resistance and dynamic friction, and then continue dynamic friction calibration if the bottom pile bulk responses do not match or calibrate static friction if the top pile parameters do not match. Although the obtained values are valid only for the specific conditions, general regularities are fair for an arbitrary bulk material. The angle of rupture is affected by static friction and the angle of repose or expiration time are affected by dynamic friction. It is worth noting the parabola coefficient A obtained from the visual image of "parabola". This bulk response is affected not only by dynamic friction, but also by cor_pp, which allows adequately calibrating the coefficient of restitution.

Although a lot of researchers dealing with DEM parameters calibration try to use neural networks (NN) and learning algorithms in the universal calibration method development, you should not ignore the iterative algorithms. While the design of a universal method based on NN requires a significant material base accumulation, the universal method based on iterative algorithm requires only standardization in the equipment used (the test rig) and a unified methodology. Such methodology can be bulk responses measurement using a technical vision system. In this case, each bulk response should allow calibrating
a certain DEM parameter, and the set of responses should allow calibrating all of the DEM parameters together.

Acknowledgments
We would like to thank CADFEM CIS and personally Andrey Feoktistov for informational support, ideas and assistance in work on this project.

References
[1] Peters B, Baniasadi M 2017 Coupled CFD-DEM with Heat and Mass transfer to Investigate the Melting of a Granular Packed Bed Chemical Engineering Science 178 136-45
[2] Gospodarikov A P, Vykhodtsev Y N, Zatsepin M A 2017 Mathematical modeling of seismic explosion waves impact on rock mass with a working Journal of Mining Institute 226 1-7
[3] Feoktistov A J, Kamenecki A A, Blekhman L I, Vasilkov V B, Skryabin I N, Ivanov K S 2015 The application of discrete element method to mining and metallurgy process modeling Journal of Mining Institute 192 145-49
[4] González-Montellano C, Fuentes J M, Ayuga-Téllez E, Ayuga F 2012 Determination of the mechanical properties of maize grains and olives required for use in DEM simulations Journal of Food Engineering 111 553-62
[5] Zhao L, Zhao Y, Bao C, Hou Q, Yu A 2017 Optimisation of a circularly vibrating screen based on DEM simulation and Taguchi orthogonal experimental design Powder Technology 310 307-17
[6] Zhou H, Hu Z, Chena J, Xuan L, Xie N 2018 Calibration of DEM models for irregular particles based on experimental design method and bulk experiments Powder Technology 332 210-23
[7] Huy Q D, Aragon A M, Schott D L 2018 A calibration framework for discrete element model parameters using genetic algorithms Advanced Powder Technology 29 1393-403
[8] Cabisco R, Finke J H, Kwade A 2018 Calibration and interpretation of DEM parameters for simulations of cylindrical tablets with multi-sphere approach Powder Technology 327 232-45
[9] Marigo M, Stitt E H 2015 Discrete Element Method (DEM) for Industrial Applications: Comments on Calibration and Validation for the Modelling of Cylindrical Pellets Powder and particle 32 236-52
[10] Coetzee C J 2017 Review: Calibration of the Discrete Element Method Powder Technology 310 104-42
[11] Rackl M, Hanley K J 2017 A methodical calibration procedure for discrete element models Powder Technology 307 73-83
[12] Furuichi M, Nishiura D 2017 Iterative load-balancing method with multigrid level relaxation for particle simulation with short-range interactions Computer Physics Communications 219 135-48
[13] Kursa M B, Rudnicki W R 2011 The All Relevant Feature Selection using Random Forest (University of Warsaw) 1106 5112
[14] Guyon I, Elisseeff A 2003 An Introduction to Variable and Feature Selection The Journal of Machine Learning Research 3 1157-82
[15] Rocky DEM Software official site http://rocky-dem.com
[16] Boikov A V, Savelev R V, Payor V A 2018 DEM Calibration Approach: design of experiment Journal of Physics Conference Series 1015 1-8
[17] Boikov A V, Savelev R V, Payor V A 2018 Technical vision system for analysing the mechanical characteristics of bulk materials Journal of Physics Conference Series 944 1-7
[18] Gospodarikov A P, Vykhodtsev Y N, Zatsepin M A 2017 Mathematical modeling of seismic explosion waves impact on rock mass with a working Journal of Mining Institute 226 1-6
[19] Boikov A V, Savelev R V, Payor V A 2018 DEM Calibration Approach: Implementing Contact Model Journal of Physics Conference Series 1050 1-7