Including stochastics in metamodel-based DEM model calibration

Fransen, Marc Patrick; Langelaar, Matthijs; Schott, Dingena L.

DOI
10.1016/j.powtec.2022.117400

Publication date
2022

Document Version
Final published version

Published in
Powder Technology

Powder Technology 406 (2022) 117400

Contents lists available at ScienceDirect

Powder Technology

journal homepage: www.elsevier.com/locate/powtec

Including stochastics in metamodel-based DEM model calibration

Marc Patrick Fransen [a,*], Matthijs Langelaar [b], Dingena L. Schott [a]

[a] Department of Maritime and Transport Technology, TU Delft, The Netherlands
[b] Department of Precision and Microsystems Engineering, TU Delft

ARTICLE INFO

Article history:
Received 3 February 2022
Received in revised form 31 March 2022
Accepted 12 April 2022
Available online 22 April 2022

Keywords:
Stochastic calibration
Metamodels
Optimization
Random packing
Verification and validation
Repeating experiments

ABSTRACT

In calibration of model parameters for discrete element method (DEM) based models the focus lies on matching the mean key performance indicator (KPI) values from laboratory experiments to those from simulation results. However, due to the stochastic nature of granular processes experimental results can show large variances. To include stochastic behaviour, interpolation-based and regression-based metamodels are trained with stochastic data. These metamodels are used in the standard mean calibration approach and newly introduced mean-variance calibration approach to predict the KPIs mean and variance. In addition, the effect of enriching data on the calibration is investigated up to 50 repetitions of experiments and simulations. Based on a hopper case study, use of regression-based metamodels trained with KPI data repeated at least 20 times is recommended. While differences between mean and mean-variance-based metamodels were minor in the considered case study, regression-based metamodeling clearly showed improved accuracy and stability over interpolation-based metamodels.

1. Introduction

To design reliable bulk handling equipment (BHE) engineers and researchers rely on particle-based models to predict the performance of a new design [16]. The discrete element method (DEM) is used to model granular materials and analyse a wide range of related applications. Accurate predictions can only be made if the input parameter values such as friction coefficients and material properties are chosen adequately. To this end, calibration with experimental findings is typically used. However, granular processes are *stochastic* by nature leading to random results if repeated. This stochastic behaviour is caused by the large variety of particle shapes, sizes, and particle packing compositions in granular systems.

In current calibration approaches the DEM parameters are optimized to ensure that the mean values for the performance parameters are matched to the calibration targets ([6,18,20,24]). However, the mean calibration approach accounts only marginally for the stochastic nature of granular processes originating from distributions of particle shape, size, and microstructural composition. Alternatively, an iterative Bayesian filtering framework in combination with analysis of stress dependency paths can be used to minimize the variance of the solution and obtain more accurate calibration results [3,4]. In this study the initial packing of the bulk material was known through X-ray tomography and exactly represented in the DEM model. In practice, this information is usually available which means that we generally assume random initial packings and repeat experiments to obtain mean and standard deviation values. Experiments for calibration and validation are commonly repeated 3–5 times. This number seems arbitrary because the mean and standard deviations of key performance indicators (KPI) may not have converged to a stable value with this number of repetitions. Additionally, if mean and standard deviations are used in developing BHE designs the common approach is to work with confidence intervals based on standard deviations. For reference, in case of three repetitions the 95% confidence interval with a t-distribution is equal to 1,837 times the standard deviation which approximately induces a 6,5 times wider confidence interval compared to a factor 0,284 when experiments are repeated 50 times. Even though this number of repetitions is not feasible in practice, this study can be used as a reference for a suitable number of repetitions.

Moreover, with increasing irregularity of particle shapes and sizes the variance in experiments increases. Therefore, in experiments with granular materials it is not uncommon to have relatively high standard deviations when material becomes more heterogeneous. For the quartz sand calibration experiments, Derakshani reports standard deviations of 2,64% for the sandglass test with three repetitions [7]. For the gravel calibration experiments with three repetitions, standard deviations between 0.5 and 4% of the mean KPI values for the lifting cylinder, shear box, and drop down test were reported [24]. For wood chips reported standard deviations range from 3;44% to 6;45% for the Angle of Repose

* Corresponding author.
E-mail addresses: m.p.fransen@tudelft.nl (M.P. Fransen), m.langelaar@tudelft.nl (M. Langelaar), d.l.schott@tudelft.nl (D.L. Schott).

https://doi.org/10.1016/j.powtec.2022.117400

and 3;85% to 7;41% for the Angle of Slip [19]. The subsequently tested wood chip feeding system resulted in a standard deviation of 4,5% for eight repetitions. In large scale applications like the grab validation study on spherical iron ore pellets by Lommen resulted in a 2% error for the payload with three repetitions [13]. Another grab validation study for cohesive iron ore reported a standard deviation of the average payload of 5,74% for three repetitions [17]. Combined with the confidence intervals this leads to 95% confidence intervals of up to +/− 10,5% around the mean in the discussed cases. In design of bulk handling equipment, these confidence intervals are generally too high to obtain reliable designs. Therefore, the effect of repetitions on the reliability of experiments with bulk materials is investigated in this study.

Ideally, the DEM model should exhibit the same stochastic behaviour as observed in reality. This requires that the stochastics are included throughout the development of the DEM model. To the author's knowledge, stochastic behaviour of granular processes is rarely included in calibration except for a recent study using random initial packing [10]. Therefore, we propose including the stochastic behaviour in a metamodel which is subsequently used to predict mean and variance in the calibration. To include the stochastics in a metamodel there are two options. Firstly, two individual metamodels are trained for the mean and variance as was demonstrated by [12], here we use a noiseless Gaussian Process Regression (GPR) metamodel which results in a metamodel based on exact interpolation. Secondly, a noise included Gaussian Process Regression metamodel which includes the KPI variance in the noise term [21]. The GPR model has been previously used in a DEM calibration context of bulk material [22]. Furthermore, a mean-variance calibration approach is proposed which includes the variance of the calibration experiments in the objectives of calibration procedure. In mean-variance optimization of DEM parameters, the focus of the optimization algorithm is to find parameter values for which both mean and variance match those of the experimentally obtained targets. In contrast, mean calibration only focusses on finding parameter values for the mean of the targets. In this paper, the noiseless GPR is referred to as the interpolation-based metamodel (IBM) and the noisy GPR is referred to as the Regression-based metamodel (RBM). Both these metamodeling approaches are used and compared for mean and mean-variance calibration of DEM parameters. In addition, we study the influence of the number of repetitions on the obtained calibration results.

We apply the aforementioned approach to a gravel case study where three material parameters are calibrated using a pile forming test, ledge test, and bulk density test. The parameters that are calibrated are the sliding and rolling coefficient, and the particle density. The calibration results are applied to a DEM model of a hopper and the resulting discharge rate is compared to the experimental equivalent. This case study is chosen because of the good measurability of the KPIs and their frequent use for calibration in this field.

The experimental setups in the case study and analysis are described in Section 2. The experiments are followed by the development of the DEM models of the experimental setups in Section 3. Next, Section 4 introduces the metamodeling approaches and the mean and mean-variance calibration, which are demonstrated with the previously described experiments. The results of the calibration and a comparison between the interpolation-based and regression-based metamodel calibration and the proposed mean-variance calibration and mean calibration is given in Section 5. Lastly, this work ends with conclusions and an outline for further work on this topic.

2. Experimental methods and materials

The goal of DEM calibration is to obtain a set of DEM parameters such as friction coefficients and material properties for which the KPI obtained from an experiment is matched to the KPI from a DEM model of the same experiment. In this study, we use the bulk calibration approach (BCA) in which small scale calibration experiments are used to determine the DEM parameters. These calibration experiments provide well-isolated macro properties of the granular material, the KPIs. Ideally, for each DEM parameter an calibration experiment with a specific macro property is desired [1,11,25]. In general, the behaviour of granular material in calibration experiments is assumed representative for the behaviour of the material in large-scale bulk handling applications. In this study three calibration experiments, a pile forming test, ledge test, and bulk density test are used which are discussed in Section 2.1. After finding the DEM parameter set these values are applied to a DEM model of a large-scale hopper model for which the numerical results are compared to the experimental results of the hopper setup described in Section 2.2. Analysis of the experimental results to obtain the calibration (pile forming, ledge, and bulk density test) and validation (hopper) KPIs can be found in Section 2.3. The results for these experiments are presented in Section 2.4.

In this case study gravel is used which is categorised as dry cohesionless bulk material similar to the gravel used in previous studies [22,24]. The particle size distribution (PSD) of this material was found using a Haver and Boecker sieve shaker tester which resulted in a normally distributed PSD with an average radius of 5,02mm and a standard deviation of 1,39mm. The particle density was measured with a submerged mass density test and averages on 2313,4 $\frac{kg}{m^3}$ for ten repetitions.

The experimental setups consist of a combination of stainless steel or acrylic plate material which can have three types of wear. The wall-particle interaction properties have been measured with an inclined surface tester for 25 particles on the wall material and their wear state. These wear states are stainless steel with wear due to sliding, stainless steel and impact wear, and acrylic with sliding wear. The stainless steel with sliding wear is the state of the wall of the bulk density tester. The stainless steel with impact wear is present in the hopper due to the pounding of the particles on the walls during filling. The acrylic with sliding wear is present in the shear box test, pile-forming test, and in the front and back walls of the hopper. The sliding friction properties have been shown in Table 1 and shows the mean and variance values for the experiment. In the DEM models of the experiments the mean value of the friction coefficient is used.

2.1. Laboratory scale experiments

The calibration experiments are a pile test, ledge test, and bulk density test as shown in Fig. 1 (a,b,c). The KPIs are the Angle of Repose (AoR) β, Angle of Movement (AoM) θ, and the bulk density ρ_b obtained from the pile, ledge, and bulk density test respectively. These experiments have been selected because they are representative for different aspects of the final application. The pile-forming test resembles the kinetic behaviour of the material exiting the silo and the formation of the pile beneath the silo. The ledge test is representative for the movement of the material along stagnant zones when the hopper is discharging. The bulk density test is a representative test for the filling of the hopper.

The pile forming test consists of an elevated structure from which gravel is dropped in a container enclosed consisting of acrylic plates, Fig. 1 (a) and Fig. 2 (a). The acrylic plates have a wear profile corresponding Gravel-Acrylic (sliding). The top container is filled up to 110 mm with gravel after which the bottom plate is removed. This starts the flow of material through the orifice of the top container. Consequently, the material fall on the bottom plate of the lower container

Table 1
Wall Gravel interaction properties.

	mean μ_s	*var* μ_s
Gravel-Stainless steel (impact)	0,523	0,05
Gravel-Stainless steel (sliding)	0,456	0,044
Gravel-Acrylic (sliding)	0,446	0,0965

Fig. 1. Laboratory scale tests (a) Pile forming test (b) Ledge test (c) Bulk density test.

and a pile starts forming. After the formation of the pile, a camera is used to photograph the sidewall of the container. This photograph is later used to determine the angle of repose, β.

The ledge test is a container with a retractable sidewall, Fig. 1 (b) and Fig. 2 (b). The container has acrylic walls which have the wear properties of Gravel-Acrylic (sliding). At the edge, the container has a ridge of 2×20 mm covering the entire trailing edge of the container. This ensures that when the sidewall is retracted that there is a stable base of gravel. The container is filled up to 214 mm with gravel after which the sidewall is retracted. After retracting the wall, the material starts to flow which continuous until a stable slope is left in the container. Of this pile a photograph is taken which is analysed to obtain the angle of movement, θ.

The bulk density test involves a cylindrical container with a radius of 82 mm and a height of 237 mm, the bulk density cylinder ISO 17828 (EN 15103), Fig. 1 (c) and Fig. 2 (c). It has a volume of 5 l with a 0,2% error.

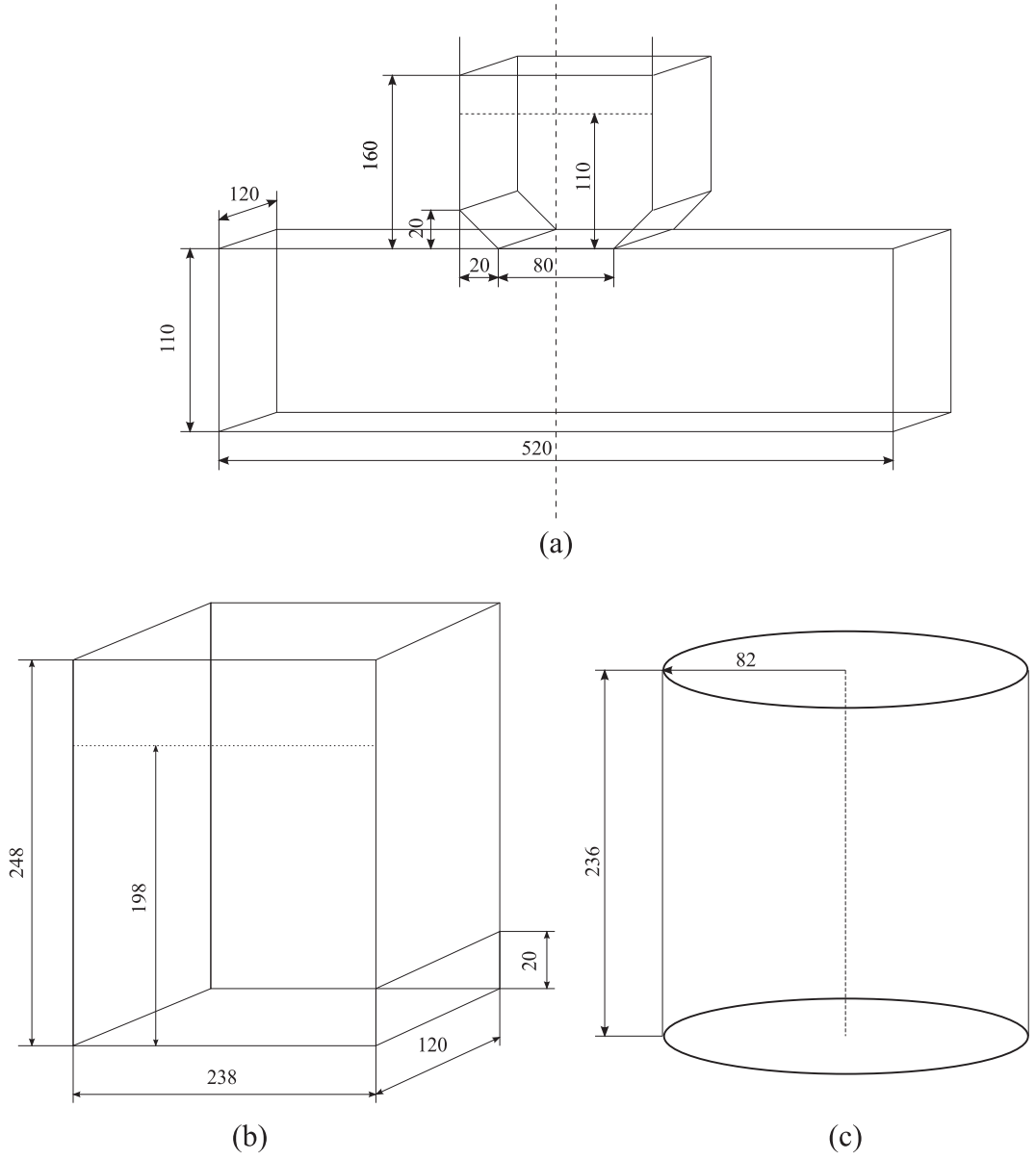

Fig. 2. Dimensions of the experimental setups (a) pile forming test, (b) ledge test, (c) bulk density test, with dimensions in millimetres.

Due to the repetitive use of the container the material has wear properties corresponding to Gravel-Stainless steel (sliding). First the empty cylinder's weight is measured with a scale (Kern EMS 12 K.01). Next, it is filled with gravel until the upper edge of the tester is reached and weighed again to obtain the mass of the bulk.

2.2. Large scale hopper experiment

The semi-two dimensional hopper setup is shown in Fig. 3. The hopper walls are connected to an aluminium frame. The sidewalls of the hopper have worn due to impact which lead to a pitted surface (Gravel-Stainless steel (impact)). These sidewalls are enclosed in the front and back by 5 mm acrylic transparent plates which have been subjected to sliding wear. The angle of the hopper walls is 45° and the size of the discharge opening is 100 mm. The width of the silo is 602 mm and the depth is 50 mm. At the bottom of the hopper a steel bar held by electromagnets closes the orifice of the hopper. To commence discharge of the hopper the steel bar is released by turning off the magnets using a switch. To measure the weight of the bulk material in the hopper the entire structure is positioned on load cells (AEB8D Shear Beam Load cell from AE sensors). The signal of the load cells is sent to a data acquisition box (Texas Instruments) and processed using Labview to a data file. The weight data is measured and stored at a frequency of 50 Hz during the discharge of the hopper. Due to the scale of this setup the number of repetitions is set to five, this number is sufficient because discharge rates are time averaged values which lead to more stable results in the case of a steady-state flow.

2.3. Analysis experimental results

The KPIs of the pile forming and ledge test are determined by analysis of the pictures of the experiments. This analysis follows the same procedure in both experiments, therefore the analysis is only discussed for the ledge test. First, the distortion of the pictures is corrected after which a square grid is layered over the pictures as shown in Fig. 4. This grid has equal spacing in both directions and is equal to two times the average particle radius obtained from the PSD. In the squares along the material edge the particle with the highest location is identified. This is a manual process which comes with errors because of the analyst determining the points, by using squares this error is minimized and the relative distance between points is kept similar. The highest locations are denoted by the blue crosses in Fig. 4 and are used as the input for a linear regression analysis to find the angle of movement. Based on the angles from all experiments the mean and variance of the Angle of Movement are determined.

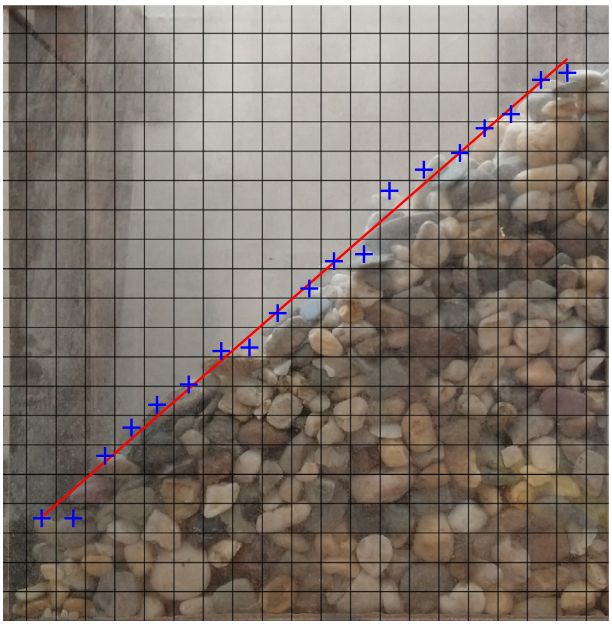

Fig. 4. Image of a ledge test result where the blue crosses are placed on the edge of the pile and a linear regression (red line) is applied to find the angle of movement. This procedure is the same for the pile-forming test to determine the angle of repose.

The bulk density is determined by dividing the bulk mass in the container measured with the scale by the volume of the container.

As we have repeated all the laboratory scale experiments 50 times the distribution of the data can be analysed, especially if the data is normally distributed. Therefore, the Lilliefors test is used to determine if the data from the experiments is normally distributed. These tests turned out to be positive for a 5% significance level which means that the data from the experiments is normally distributed. Based on these observations we can proceed with mean and mean-variance based calibration.

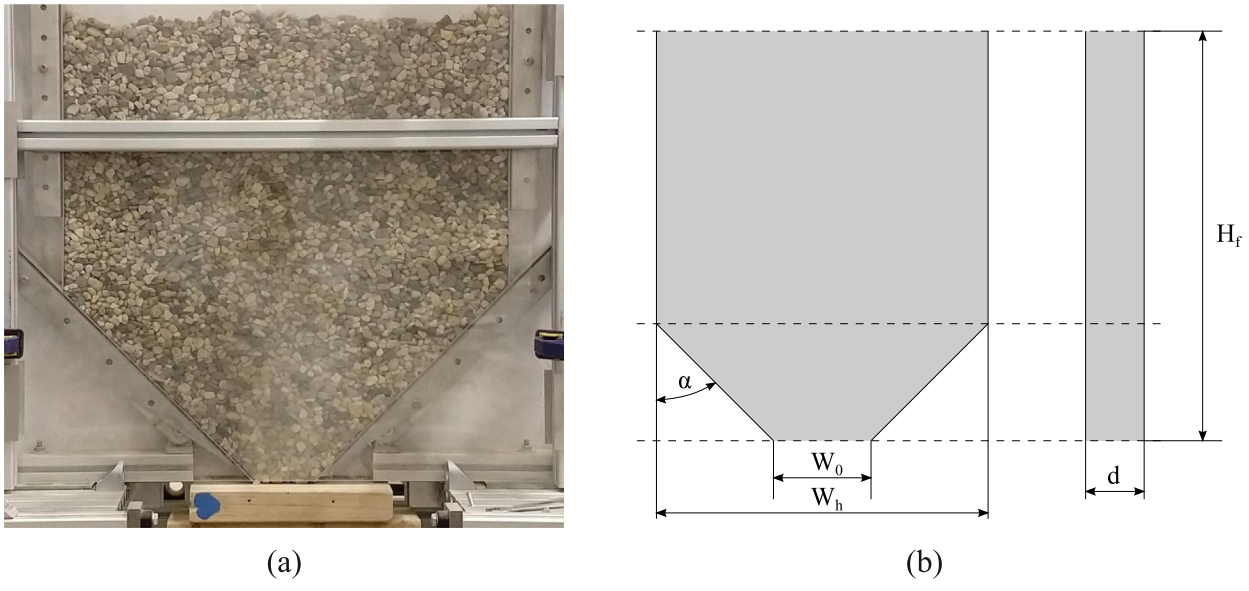

Fig. 3. Experimental setup for the large scale hopper.

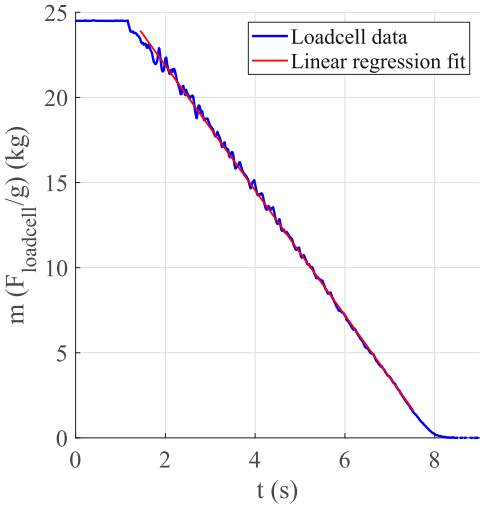

Fig. 5. Linear regression fit to the load cell data, the slope of the fit is the discharge rate of the mass flow in the hopper.

The data obtained from the large-scale hopper experiment is the evolution of the force exerted on the hopper by the material sampled at a frequency of 50 Hz. From this data the discharge rate as a function of time can be determined using linear regression as illustrated in Fig. 5. Assuming that the constant force fit divided by the gravity constant leads to an average discharge rate in kg/s. For each large scale hopper experiment the steady state discharge rate ϕ is calculated. The mean discharge rate and its variance are calculated for the five repetitions of the experiment.

2.4. Experimental results

The experiments for the Angle of Repose, Angle of Movement, and bulk density have been repeated 50 times. On these datasets, a Lilliefors test has been applied to determine if these datasets are likely to be normal with 95% confidence. For all three sets it was found that the data is normally distributed which justifies including the confidence intervals in the analysis of the experimental results.

In Fig. 6 (a,b,c) we can see the way the mean value and 95% confidence interval (CI) of the angle of repose, angle of movement, and the bulk density of the material develops when the number of repetitions increases. In addition, the +/− 1% boundaries are given which indicate the relative size of the error in the mean value. As we can see, the mean stays within the 1% boundaries after the number of repetitions increases over 29 for the angle of repose. The angle of movement reaches a stable bandwidth of +/− 1% around the 41.5 degree angle after nine repetitions where it leaves this bandwidth once at 30 repetitions. For the bulk density we see that the mean is in the 1% bandwidth after three repetitions. However, the mean stabilises when 20 repetitions are reached. The 95% CI of the mean is determined using the t-distribution for a number of repetitions lower than 30 and using the z-distribution for numbers higher or equal to 30 assuming that the central limit theorem holds. Based on the confidence intervals we can say that at 50 repetitions the true mean of the AoR is within 2,2% of the mean with 95% certainty. For the AoM this is equal to within 1% and for the bulk density within 0,2%. In the calibration procedure described in Section 4 we

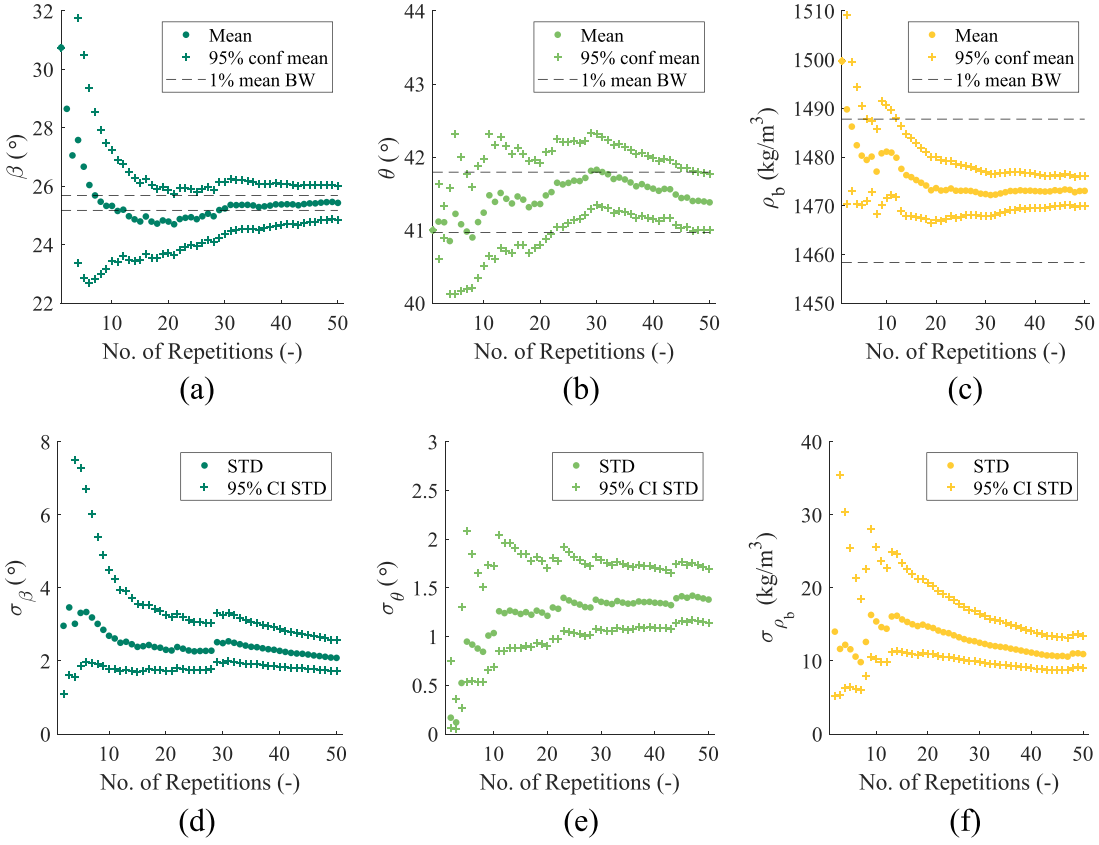

Fig. 6. In figures (a, b, c) the development of the mean value and 95% CI related to the number of repetitions of the calibration experiment is shown for the Angle of Repose, Angle of Movement, and Bulk density, respectively. Additionally a bandwidth of 1% shown around the mean value for 50 repetitions. Figures (d, e, f) show the development of the standard deviation and its 95% CI.

Table 2
Calibration targets for the Angle of Repose, Angle of Movement, and bulk density for 50 repetitions and the validation target for the discharge rate for 5 repetitions.

Calibration Targets	KPI	Number of experiments	Mean ± *CI* 95%	Standard Deviation ± CI 95%
Pile test	β_e (°)	50	25,43 ± 0,58	2,09 ± (−0,24; +0,48)
Ledge test	θ_e (°)	50	41,40 ± 0,36	1,38 ± (−0,36; +0,31)
Bulk density test	$\rho_{b_e}\left(\frac{kg}{m^3}\right)$	50	1473,1 ± 3,0	10,94 ± (−1,88; +2,52)
Validation Target				
Hopper test	$\phi_e\left(\frac{kg}{s}\right)$	5	3,68± 0,029	0,025 ± (−0,011; +0,030)

assume that both the calculated mean and standard deviation are the true values. Therefore, we use these values as the calibration targets. The CIs are only used to indicate the reliability of the experimental results.

In Fig. 6 (d, e, f) the standard deviations and their 95% CI are shown relative to the mean. For the AoR and AoM stabilization of the standard deviation can be observed after 15 repetitions but for the bulk density it keeps declining with the number of repetitions. For the CI of the standard deviation the chi-distribution is used to account for the sample size. Observing the confidence intervals, it is visible that for low numbers of repetitions the confidence intervals for the angle of repose and bulk density are much wider than for a high number of repetitions. These observations indicate the necessity of having sufficient repetitions in calibration experiments for irregular shaped and randomly packed bulk materials.

In this case study the effect of increasing the number of repetitions in a dataset on the calibration and validation result are evaluated. However, for conclusions regarding the most accurate calibration approach the entire dataset of 50 repetitions will be used. In these cases, it is also important to notice that one can consider the standard deviation relative to the mean. These calibration targets are shown in Table 2. It is found the standard deviation is 8,2% for the Angle of Repose, 3,3% for the Angle of Movement, 0, 74% for the bulk density. These percentages show that the variability of each calibration differs significantly. This can be explained by the nature of the calibration experiments. In the bulk density test a container is filled with material which is very confined and restricts high velocity movement of the particles. Hence an experiment that will have less variability. In the ledge test a container is filled as well but when the sidewall is removed the material starts moving. Moreover, the sliding interface in the ledge test causes higher variability of the angle due to the changing orientation of the particles in each repetition of the experiment. In the pile-forming test the material falls from the container on a free surface where the orientation and high velocity of the particles results in a higher variability of the experimental result.

3. DEM models

The DEM models of the hopper and calibration experiments are built based on the schematics shown in Fig. 2. The model assumptions for the granular material and material structure interaction are described in Section 3.1. Next, the procedure for initialisation of the models is described in Section 3.2. Finally, the analysis of the simulation results is discussed in Section 3.3.

3.1. DEM model assumptions and material properties

In general, DEM models of a granular process are simplifications of the actual process. For reproducibility, the procedures followed and assumptions made in creating the model are as follows. In this study we use Mercury DPM, an open source discrete element package [26] to build the three-dimensional DEM models. Although the particle shapes are irregular they are modelled as spheres. As suggested by Wensrich & Katterfeld, the irregular shapes of the particles and its behaviour is assumed to be described by the rolling friction coefficient in the contact model [27]. The size of the spheres follows the experimentally determined PSD but is truncated between −1,5 STD and + 5 STD around the particle size mean. Ideally, the particle size distribution in DEM is identical to the experimentally obtained distribution. However, for reasonable simulation times truncation of the PSD is justified [23].

The contact model used to describe normal, tangential, sliding, and rolling interaction for particle-particle and particle-wall contact is the model developed by Luding [15]. This model is suitable for modelling dry cohesionless granular solids. The contact stiffness k for particle-particle interaction is determined by using the micro-macro relation given in the following equation,

$$k = \frac{KV_p}{C_n r^2} \tag{1}$$

where K is the bulk modulus of the granular material, V_p is the particle volume of a sphere with the average particle radius from the PSD, C_n is the coordination number, and r is the average particle radius [14,15]. The bulk modulus K of the material is set to 70 MPa. The particle contact number, or coordination number, is assumed to be equal to 4 which corresponds to loose material [29]. The damping coefficient of the material is assumed to be equal to 0,3 [9]. The material properties have been tabulated in Table 3 and the contact properties for the particle-particle contact can be found in column two of Table 4. Based on these properties the size of the time-step can be determined. For the time-step size we assume that $\Delta t = \frac{t_c}{10}$ where t_c is the collision time between the particles for the smallest particle size and was used in previous work [8].

For the interaction between particles and three types of walls sliding and rolling friction coefficients are defined. The mean values of the measured sliding friction properties declared in Table 1 are directly included in the DEM model and are assumed uniform over the wall surface. Here it is assumed that the walls have uniform properties and are therefore not spatially dependent. The rolling friction coefficient could not be measured directly. Therefore the relation $\mu_r = 2\mu_s$ is adopted which is common in DEM models for calibration ([5,30]) to ensure sliding is the dominant mode of motion in the simulation. The contact stiffness of the walls is also assumed to be twice the size of the contact stiffness of the particles, $k_w = 2k_p$ ([5]). The properties can be found in the third column of Table 4.

The particle density has been identified experimentally but because of the assumption of spherical particles adjustments to the particle density might be needed. Therefore it is decided to use the particle density

Table 3
Material properties gravel.

Bulk modulus	$K = 70MPa$
Particle density	Calibration parameter
Contact number	$C_n = 4$ [29]
Particle size	5,02 (mm) +/− 1,39 (mm) Sieve test results, in simulation the particle size is limited to −1,5 and + 5 sigma.
Timestep	$\Delta t = \frac{t_c}{10}$

Table 4
Contact Properties.

Stiffness	Particle-particle	Particle-Walls
Contact stiffness k	$k_p = 3{,}67743 \times 10^5\ (\frac{N}{m})$	$k_w = 2k_p$
k_r	$\frac{2}{7}k_p$	$\frac{2}{7}k_w$
k_s	$\frac{2}{7}k_p$	$\frac{2}{7}k_w$
Damping ratio		
γ	0,3 [31]	0,3
γ_s	0,3	0,3
γ_r	0,3	0,3
Friction		
μ_s	Calibration parameter	Experimental sliding friction values (Table 2)
μ_r	Calibration parameter	$\mu_r = 2\mu_s$

as a calibration parameter. In the particle-particle contact, the sliding and rolling friction coefficients are parameters that are subject of the calibration because both these properties cannot be measured directly.

3.2. Initialisation of DEM models and simulations

In Fig. 7 examples of the simulation results for the DEM models of the calibration experiments for μ_s, $\mu_r = 0{,}5$ and $\rho_p = 2750\ kg/m^3$ are shown. Here it can be seen that pile formation occurs in the pile test shown in Fig. 7 (a), a slope in the ledge test result (b), and a filled bulk density container in (c).

To resemble the filling process in the experiments random filling is used. The procedure is illustrated by the filling procedure for the bulk density container. For each DEM model a volume is defined in which the particles will be generated. In the case of the bulk density tester it is the volume of the test apparatus. A circular generation plane is defined at the bottom of the tester and moves up to fill the tester with a predefined volume. The first loop is to generate particles in the circular plane of the bulk density tester. In this loop a random location in the plane is chosen by randomly generation of an angle for the angular location 0 to 360 degrees on the circle and a radius from 0 to the outer radius of the tester. At this location, a particle is generated with a radius drawn from the particle size distribution obtained from the experiments. Before generating it is checked whether the particle is in contact, if so, it is not placed, else, it is placed and the next particle can be placed. With each step in generating a particle, the height of the horizontal generation plane is increased with a step size of $r \times 10^{-4}$ to ensure proper filling of the volume where the average particle radius is r. After adding a particle its volume is substracted from the total volume that needs to be added. When this added volume reaches zero or less the particle generation stops. With this procedure a volume is filled with particles without contact. After initialisation of the particles, the simulation is started and particles settle under the influence of gravity.

The procedure for all simulation setups is similar while the shape of the generation plane is adjusted to the setup. In the DEM models, the mass of material added to the setup is larger than the amount needed to fill the setup. Therefore, when the particles have settled, the deletion boundary that removes excess material above the filling height H_f for each setup is activated. This activation time is different for each model but equilibrium is reached before the deletion boundary is activated (Table 5).

For the hopper simulation material flow starts upon removal of the bottom of the hopper at $t = 1{,}6s$. In the ledge and pile forming test the sidewall and bottom are removed at 1,65 s, respectively. The stopping criterion for all simulations is when ratio between the elastic and kinetic energy of the particles in the system becomes lower than 10^{-6} [8]. The hopper simulation has an additional stopping criterion that ensures that the simulation stops when the mass in the hopper is equal to zero.

The DEM models of the calibration experiments are run on a cluster which uses 2× Xeon E5–2680 v4, 28 core CPUs. The average CPU time for the DEM models of the laboratory scale tests is 5 h for the bulk density test, 7,5 h for the pile forming test, and 12 h for the ledge test. The simulation time is denoted as an average because the packing, value of the friction coefficient, and density influences the behaviour of the model and the simulation time.

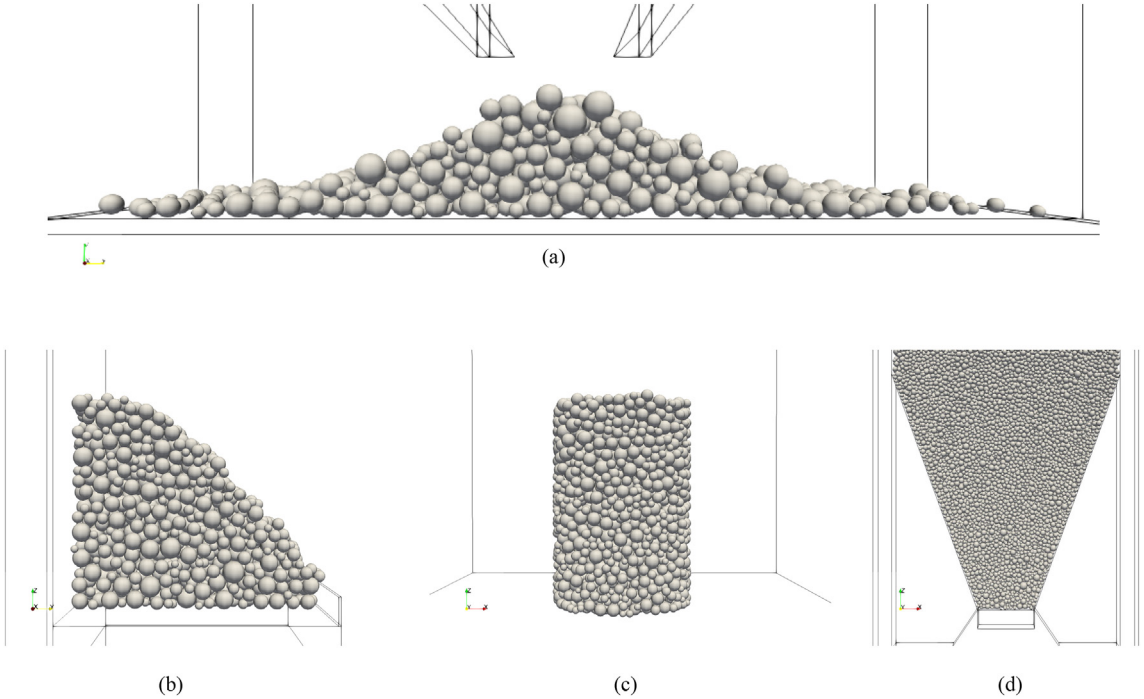

Fig. 7. Simulation results for the pile (a), ledge (b), and bulk density test (c) for sliding and rolling friction coefficient of 0.5 and a particle density of 2750 kg/m3 and a visualisation of a filled DEM model of the hopper (d).

Table 5
Simulation settings.

DEM models	Filling height H_f	Activate deletion boundary
Hopper test	0,66 m	1,5 s
Pile test	0,11 m	1,55 s
Ledge test	0,218 m	1,55 s
Bulk density test	0,236 m	1,5 s

3.3. Analysis simulation results

The DEM simulations provide data files containing the locations and velocities of the particles. This data is used to determine the KPI values for each DEM model. In contrast to the experimental analysis the analysis of simulation results is automated. Therefore this method is not prone to any error from the analyst whereas the experimental analysis relies on picking data points by hand which can induce errors. For each repetition and sample point this value is calculated which in turn is used to determine the mean and variance values. These values are used as training data for the metamodels used in the calibration procedure (Section 4.3).

To obtain the angle of repose from the pile test simulations a grid search is performed to find the locations of the largest combination of particle z-value and its radius in bins. In this search, only the particles located at front are considered analogous to the analysis of the experimental results where only the visible particles are analysed. The pile has two slopes as shown in Fig. 8 (a) for which the approach is to determine the angles at both sides of the pile. The left edge of the search area is d_{s1} which is located 0,1 m left of the center of the setup along the x-axis. The right edge of the search area for the left angle is located at d_{e1} which is located 0,01 m to the left of the setup center on the x-axis. This search area is divided into a number of bins based on the length of the search area divided by the average particle diameter rounded up to an integer. In this case study 9 bins are used. For the right angle the same procedure is followed. After finding the highest particles in each bin, this data is used to determine the angle of repose by linear regression for both the left and right side which are averaged.

After the ledge test simulation finished, the bulk material has formed a stable slope as shown in Fig. 8 (b) for which the angle can be determined. However, the calibration parameter values in the sample do affect the shape of the slope. Increasing the friction values increases the slope of the heap but also forms a horizontal plane at the back of the shear box after which the slope starts. To obtain the accurate angle of movement the particles lying in this plane should not be considered.

To obtain the angle of movement, the first task is to find the particles which have the highest z-values at the front of the simulation. Note that this z-value is the sum of the z-value of the centre of the particle and the radius of the particle. As with the Angle of Repose, we do this by looking for the largest combination of z-value of the particle and its radius in bins. The left edge of the search area is defined by the adopted friction relation, $d_s = \frac{\max(\mu_s, \mu_r)}{0,9} \times 0,06$ and the right side of the bin is located at d_e=0,228 m which is 0,01 m from the leading edge of the shear box. The area between d_s and d_e is divided into a number of bins. The number of bins is determined by dividing the distance between d_s and d_e by the average particle diameter and rounding up. In this way we ensure that the particle in each bin is the highest. After a particle is found in each bin, linear regression is applied to find the angle of movement θ. This method is identical to the experimental analysis.

The result of the bulk density simulation is a volume filled with particles as shown in Fig. 8 c. From the data file the volume of each particle in the container can be calculated, and by multiplying with the particle density and dividing the total mass by the volume of the container the bulk density is obtained.

For the analysis of the hopper results the same procedure is followed as with the experimental hopper results but using the numerical data instead of the experimental (Section 2.4).

4. Metamodel and calibration methods

Our calibration, verification, and validation procedure is shown in Fig. 9. It starts with the calibration experiments shown on the left side of the figure. We use the Pile Test, Ledge Test, and the Bulk density test with the mean and variance of the Angle of Repose, Angle of Movement, and bulk density as output which is discussed in Section 2. For each of those experiments a DEM model has been developed as discussed in Section 3.

The first step in the calibration procedure is to define the bounds and sample the calibration space, described in Section 4.1 and 4.2. The results of the DEM simulations for the sample will be used in the metamodel training step and are subsequently used for DEM parameter calibration. We use two types of metamodels to describe the relation between calibration parameters and KPIs. The first type is an interpolation-based metamodel that describes the mean and variance of the KPIs by two separate metamodels. The second type is a regression-based metamodel, which is a metamodel that incorporates both mean and variance of the KPIs. Details on the models and training procedure are discussed in Section 4.1. For the calibration step, we use

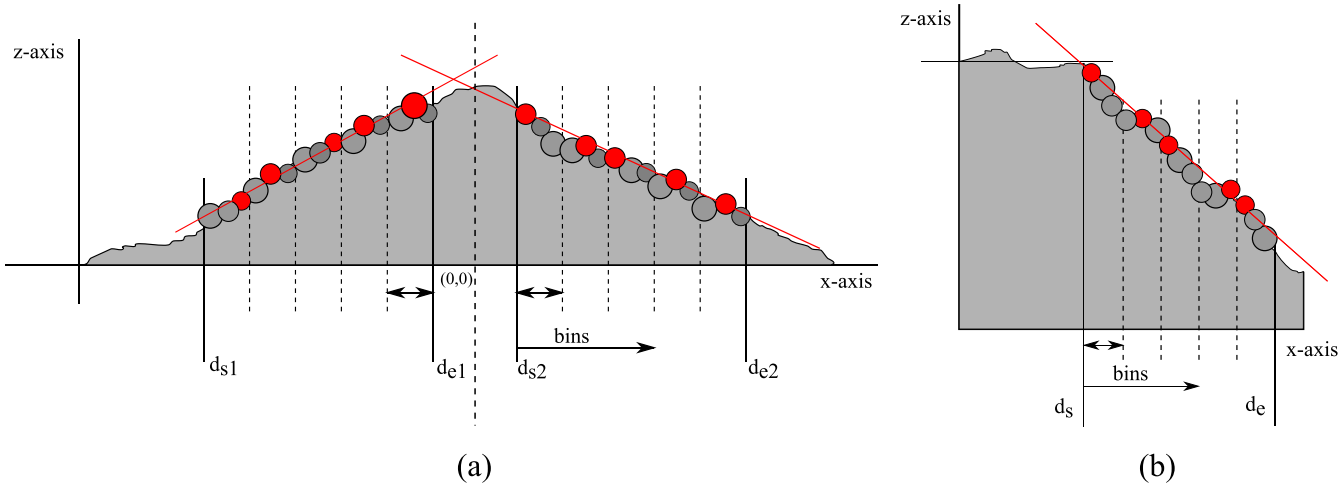

Fig. 8. Visualisation of the analysis of the simulation results where a linear regression is applied to the highest located particles in each bin (a) determination of angle of repose (b) determination of the angle of movement.

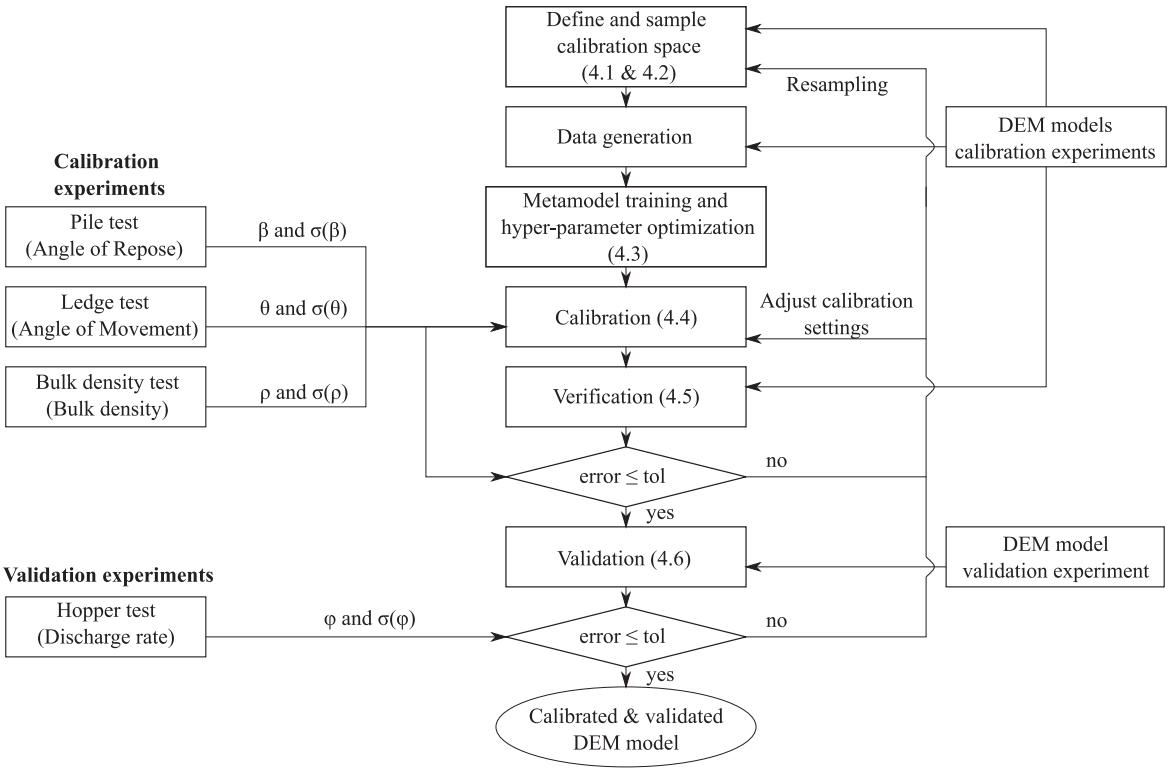

Fig. 9. Overview of steps taken in the calibration approach and the validation.

both the commonly used mean calibration approach and our newly introduced mean-variance calibration approach. The main difference between these methods is that the latter includes the variance of the KPI in the calibration in addition to the mean. In both approaches, using the DEM-based metamodel the DEM parameter values are determined that provide the best agreement with the experimental KPI data. The calibration is further discussed in Section 4.4. After this calibration step, the found optima need to be verified in the verification step as discussed in Section 4.5. Here the found DEM parameter sets are reintroduced to the DEM models of the calibration experiments and the simulation results are compared to the results predicted by the metamodels. If the verification results are not satisfactory, the sample of the calibration space can be enriched by adding samples or the optimization settings in the calibration can be adjusted. In this study resampling based on the calibration and verification results is not considered. After successful verification, the DEM parameter set is used as input to the validation step by executing simulations with the DEM hopper model which is described in Section 4.6. The results of these simulations are compared to the experimental results from the Hopper discharge experiment. If the error between the experiments and simulations is acceptable, a validated DEM model of the equipment has been obtained. If the error is too large revisiting the calibration or sampling step is required.

In this study we show the effect of the mean and mean-variance calibration strategy in combination with the interpolation-based and regression-based metamodels on the evolution of the found DEM parameter optima when the number of repetitions of the calibration sample simulations increases. In addition, the quality of these found optima are evaluated for the calibration and equipment experiments.

4.1. Bounds of the calibration space

Ensuring that there is a solution present in the calibration space that is used is an essential step in the calibration procedure. In this approach we define lower and upper bounds for the calibration parameters which are the sliding friction coefficient, rolling friction coefficient, and particle density. The lower and upper bound of the friction coefficients has been set to 0,1 and 0,9 based on common values in literature involving gravel experiments as well [23,28]. In combination with the measured particle density ρ_{p_e}, the bulk density can be determined by running the bulk density simulation. To determine the bulk density bounds, this is done for two situations, the experimental particle density with the highest friction coefficients and the lowest friction coefficients. Moreover, to ensure stability of the outcome these two situations have been repeated 50 times with random material packings to get an accurate estimate of the average bulk density and its standard deviation. Based on the results from these simulations, the experimental particle density, and the experimental bulk density the lower and upper bound for the particle density can be found by using the following two equations,

$$\overline{\rho_{p_{s\,-\,lb}}} = \frac{\overline{\rho_{b_e}}}{\overline{\rho_{b_s}}\left(\mu_{slb}, \mu_{r_{lb}}, \rho_{p_e}\right)} \overline{\rho_{p_e}} \tag{2}$$

$$\overline{\rho_{p_{s\,-\,ub}}} = \frac{\overline{\rho_{b_e}}}{\overline{\rho_{b_s}}\left(\mu_{sub}, \mu_{r_{ub}}, \rho_{p_e}\right)} \overline{\rho_{p_e}} \tag{3}$$

where $\overline{\rho_{b_e}}$ is the average experimental bulk density, $\overline{\rho_{b_s}}$ is the average simulated bulk density which depends on the friction coefficients and experimental particle density. The average experimental particle density is denoted by the term $\overline{\rho_{p_e}}$. The lower and upper bound are denoted by *lb* and *ub*. With these lower and upper bounds of the particle density determined, they can be used in combination with the sliding and rolling friction coefficient limits to check if the lower and upper limits of the KPIs of the other calibration models, the shear box and pile test, stay in their respective bounds.

After verification of these results the next step is to define the boundary sample. In this case study we have three parameters which results in a three-dimensional calibration space with the following bounds denoted in Table 6.

For these lower and upper bound values we have confirmed that the lower and upper bounds result in values for the angle of repose β, angle of movement θ, and bulk density ρ_b which contain the experimentally found calibration targets. In Section 3 the model assumptions have been declared where it is assumed that the material has a contact number of 4. After the simulations described here the contact number was found to be ≈4,36 with lower limit of 3,6 and an upper limit of 5 depending on high or low friction, respectively. This makes 4 a reasonable assumption for the contact number.

4.2. Sampling the calibration space

After the feasibility check of the calibration space bounds it can be sampled. The sample consists of two parts, a sample covering the edges of the calibration space and a sample covering the internal volume of the calibration space. Sampling of the edges is essential if metamodels are used for the calibration because these have poor extrapolation abilities beyond the space covered by the data points. In this case the boundary sample consists of 27 sample points located at the corners of the calibration space, at the halfway points between the corner points, and the centre points in the plane and cube as shown as purple squares in Fig. 10. To cover the internal space of the calibration space an *internal* sample of 100 points based on Latin Hypercube Sampling (LHS) with the maximin criterion is applied to obtain a sample that covers the calibration space. These sample points are denoted by blue circles in Fig. 10. The total number of sample points is 127 which leads to a sampling density of $\eta = 127^{1/3}$ which is equal to 5,02 samples per unit step in the normalized three dimensional calibration space.

To obtain accurate predictions of the mean and variance of the KPIs the simulations for the sample points are repeated 50 times where in each repetition the packing of the particles is randomly generated. The total number of simulations is equal to 4050 for the boundary sample for the DEM models of the calibration experiments. For the internal sample, the number of simulations is 15000 which totals 19050 simulations training the calibration metamodels. In this study a high number of repetitions is used to find an accurate reference value. Based on the results, recommendations can be given on the amount of repetitions that are needed to reach an accurate calibration result.

Table 6
Bounds of the calibration space.

Calibration parameter	Lower bound	Upper bound
$\mu_s(-)$	0,1	0,9
$\mu_r(-)$	0,1	0,9
$\rho_p\ \left(\frac{kg}{m^3}\right)$	2542,9	2834,7

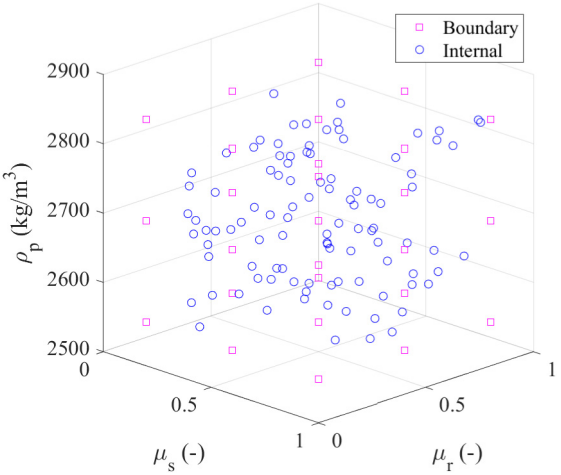

Fig. 10. Three dimensional representation of non-normalized sample.

4.3. Metamodel training

The results from the simulations for the calibration sample are used to train a Gaussian Process Regression metamodel which will be denoted by $f_i(\boldsymbol{x})$ where i is the index of the KPI and $\boldsymbol{x} = [\mu_s\ \mu_r\ \rho_p]$ is a vector containing the calibration parameters as variables for which the GPR metamodel gives a prediction of the KPI. The mean and standard deviation values obtained from the DEM simulation results are normalized such that a normalized mean and coefficient of variation can be used in training the metamodels. As mentioned in the introduction, two types of metamodels are trained.

1. Separate noiseless GPR models based on interpolation for mean and variance of the KPI [12]
2. Single noisy GPR model based on regression including both mean and variance of the KPI [21]

A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution [21]. The Gaussian process *GP* is an approximation of the real process $f_i(\boldsymbol{x})$ that it intends to model. The *GP* consists of a mean function $m(\boldsymbol{x})$ and a covariance function $k(\boldsymbol{x},\boldsymbol{x}')$.

$$f_i(x) \sim GP(m(\boldsymbol{x}), k(\boldsymbol{x},\boldsymbol{x}')) \tag{4}$$

In building a metamodel with Gaussian Processes the correlation function between the training points needs to be chosen which forms the basis of the model. In this case, the correlation function (Eq. (5)) is chosen as a squared exponential Gaussian in $\mathbb{R}^3$ with a shape factor l^2 and an amplitude factor σ_f which in this case are both set equal to 1. To add more flexibility to the correlation function, the contribution of each design parameter is determined based on the Euclidean distance and is weighted with a coefficient a_i for each calibration parameter (Eq. (6)). This makes the shape factor l^2 obsolete and would interfere in the hyper-parameter optimization process. The factor σ_f is set equal to 1 because in the noiseless GPR there is no influence on the solution and in the noise included GPR a value of 1 is suitable because it preserves the ratio between the normalized mean values and coefficient of variation. These coefficients are called hyper-parameters which can be user-defined or optimized. In Gaussian Process Regression it is common to optimize these hyper-parameters by using type II log-likelyhood maximization [21]. In this method the negative marginal log-likely-hood function (Eq. (7)) is minimized to find the hyper-parameter values. Where K is the covariance matrix, σ_n^2 is the coefficient of variation in the training points, and N is the total number of training points. In the noiseless GPR the coefficient of variation in the training points is equal to zero whereas it contains the coefficients of variation in each datapoint in the noisy GPR.

$$\phi(x) = \sigma_f e^{-\frac{1}{2l^2}r^2} \tag{5}$$

$$r = \sqrt{a_1(x_1 - x_{1,c})^2 + a_2(x_2 - x_{2,c})^2 + a_3(x_3 - x_{3,c})^2} \tag{6}$$

$$\log p\,(y|X) = -0.5y^T\left(K + \sigma_n^2 I\right)^{-1} y - \frac{1}{2}\log|K + \sigma_n^2 I| - \frac{N}{2}\log(2\pi) \tag{7}$$

The maximization problem is solved using a constrained optimization solver in $[-1,1]\ \mathbb{R}^3$ using the interior point method on which

theory can be found here [2]. To ensure that the initial guess for this gradient-based solver does not miss the global optimum we have used 100 random initial guesses to rule out this effect. After solving the maximization problem, the found optimum for the hyper-parameters of each KPI is fixed input in training the calibration metamodels used for optimization of DEM-parameters.

The metamodel prediction (Eq. (8)) for the mean value consists of the correlation between the training points and predicted point, $k(x,x^*)$. This is multiplied with the inverse of the summation of the Gram matrix $K(x,x)$ and the variances σ_n^2 and with the reference values y from the DEM simulation results for the training points in the calibration space to obtain a prediction of the mean at x^*. The variance predictor (Eq. (9)) is a correction based on the prior covariance of the predicted points $k(x^*,x^*)$ minus the information the training points give about the function [21].

Predictor mean

$$\overline{f}\,(x^*) = k(x,x^*)^T\left(K(x,x) + \sigma_n^2 I\right)^{-1} y \tag{8}$$

Predictor variance

$$V(f^*) = k(x^*,x^*) - k(x,x^*)^T\left(K + \sigma_n^2 I\right)^{-1} k(x,x^*)^T \tag{9}$$

In this case study we defined three KPIs, the angle of repose β, angle of movement θ, and the bulk density ρ_b for each of which we have developed the interpolation-based and regression-based metamodels. All these calibration metamodels have DEM parameters as their variables, the sliding friction μ_s, rolling friction μ_r, and the particle density ρ_p. For the interpolation-based metamodels this results in six metamodels. A separate mean and variance predictor for each KPI. In the regression-based metamodel the mean and variance predictor are included in a single model which results in three metamodels.

4.3.1. Notation

Interpolation-based metamodels are denoted by IBM, regression-based metamodels by RBM, followed by the abbreviation of the calibration method. M for the mean calibration and MV for the mean-variance calibration.

4.4. Calibration of DEM parameters

In the previous section we have described the two types of calibration metamodels that are developed for the calibration procedure. Both these models can predict mean and variance values for the KPIs of the laboratory scale experiments. Next to these two metamodeling approaches, in the calibration of the design parameters we can choose between using a mean (M) or mean-variance (MV) calibration approach. Here, the mean calibration approach focusses on matching the mean values of the experimental and numerical KPI values. The mean-variance approach matches both mean and variance values of the experimental and numerical KPI values. As discussed in Section 2, the calibration targets are the experimentally obtained mean and variance for which we assume that the found values are the true mean and variance. This means that in the calibration we do not include the confidence intervals around the found mean and variance values. In this section these two approaches and the used solver for the optimization problem are described. The following calibration approaches will be evaluated.

1. Mean calibration with interpolation-based GP model (IBM-M)
2. Mean-Variance calibration with interpolation-based GP model (IBM-MV)
3. Mean calibration with regression-based GP model (RBM-M)
4. Mean-Variance calibration with regression-based GP model (RBM-MV)

The two calibration approaches are applied to each of the two calibration metamodels for an increasing number of repetitions. This means that calibration metamodels are trained with data for all number of repetitions individually to see how the optimal solution changes with increase of the number of repetitions.

To calibrate the DEM-parameters a multi-objective optimization problem is formulated which is solved using constrained optimization with the interior-point-method [2]. In order to find a global optimum, the optimization is run 100 times with different initial guesses in the calibration space to find the optimum DEM-parameter set.

4.4.1. Mean multi-objective optimization problem

For the mean calibration problem the goal is to match the mean KPI values from the experiments to those of the simulations. To calibrate the DEM-parameters the following multi-objective optimization problem is defined.

$$\min \boldsymbol{c}(\boldsymbol{x}) \tag{10}$$

Where $c(\boldsymbol{x})$ is a summation of the objectives for the optimization problem. The goal of calibration is to find DEM parameter values for the sliding and rolling friction coefficient for the interaction between gravel particles and the particle density. For each KPI (Angle of Repose, Angle of Movement, bulk density) an objective is defined according to the following equation,

$$c_i(\boldsymbol{x}) = w_i\left(\frac{(K_{ie} - K_{is}(\boldsymbol{x}))}{K_{ie}}\right)^2 \tag{11}$$

Where K_i represents the KPI, w_i the weight of the objective, and $\boldsymbol{x}$ is a vector containing the DEM parameter values. The subscript e for the KPI denotes that it is the experimental value and the subscript s of the KPI denotes that it is the numerical or simulation value. The term $K_{is}(\boldsymbol{x})$ provides a prediction of the mean and variance of the KPI by the metamodel described in Section 4.3. The objective is normalized by the simulation value of the KPI and squared such that the objective returns a positive value.

For this case study there are three KPIs resulting in three objectives which are summed in $c(\boldsymbol{x})$. Each objective is normalized and is therefore equally important if the weights are equal to one. For the clarity of this study the we have decided to keep the weights equal to 1, $\boldsymbol{w} =$ [1 1 1]. However, in practical applications these weights can be adjusted if the engineer or researcher decides that some KPIs are more important than others.

4.4.2. Mean-variance multi-objective optimization problem

For the mean-variance calibration problem the definition of the optimization problem is the same as in Eq. (10). However, in addition to the mean objectives the relation between the mean and variance is utilized to define six additional objectives. These objectives are the mean plus and minus the standard deviation, which is the square root of the variance, of the KPI. The standard deviation of the KPI is used because it has the same unit as the mean. The upper bound is defined as the error between the sum of the experimental mean K_{ie} and standard deviation value σ_{Kie} and the sum of the numerical prediction for the mean K_{is} and standard deviation σ_{Kis} as described in Eq. (12). The lower bound is defined as the error between the experimental mean minus the standard deviation value and the numerical prediction for the mean minus the standard deviation as described in Eq. (13).

$$c_i(\boldsymbol{x}) = w_i\left(\frac{\left(K_{ie} + \sigma_{K_{ie}}\right) - \left(K_{is}(\boldsymbol{x}) + \sigma_{K_{is}}(\boldsymbol{x})\right)}{K_{ie} + \sigma_{K_{ie}}}\right)^2 \tag{12}$$

$$c_i(\boldsymbol{x}) = w_i\left(\frac{\left(K_{i_e} - \sigma_{K_{i_e}}\right) - \left(K_{i_s}(\boldsymbol{x}) - \sigma_{K_{i_s}}(\boldsymbol{x})\right)}{K_{i_e} - \sigma_{K_{i_e}}}\right)^2 \tag{13}$$

In the mean-variance calibration problem each objective has a weight. Here, the weights indicate the importance of each part of the objective function. For the objective of the mean value the same weights are used as in the mean calibration case. The weights for the mean ± standard deviation objectives is defined by a factor C which scales the influence of these objectives. The value chosen for $C = 0{,}5$ and kept constant in this case study which result in the weight vector $\boldsymbol{w} = [1\ C\ C\ 1\ C\ C\ 1\ C\ C]$. We assume that the lower and upper bound are equally important and that the sum of these should be equally as important as the mean objective of the specific KPI.

4.5. Verification of DEM parameter sets

For each of the four calibration approaches described in the previous section, 50 sets of DEM-parameters for metamodels trained with mean and variance data are obtained. These optimal DEM-parameter sets are applied to the DEM models of the laboratory scale experiments. The results are verified by comparing KPIs from the simulation results to the calibration targets. To reduce the number of simulations, a selection of obtained DEM-parameter sets will be evaluated. This selection is defined by the DEM-parameter sets corresponding to 1, 2, …, 10, 15, 20, 30, 40, 50 repetitions. Simulations for these 15 sets are repeated with the same number of repetitions as the number of repetitions used to train the metamodels. From these simulations the mean and standard deviation values for the KPIs are determined. These values are used to verify the accuracy of the metamodel prediction of the optima of the KPIs mean and mean ± standard deviation. Furthermore, the verification results are evaluated with respect to the calibration targets to get insight on the quality of the calibration results.

4.6. Validation of optimal DEM parameter sets for large scale hopper model

In addition to the verification, the DEM parameter sets need to be applied to the DEM model of the hopper and simulated. The hopper discharge experiments described in Section 2.2 were repeated 5 times so for the validation the DEM parameter sets from the mean and mean-variance calibration were repeated 5 times as well. This procedure as applied to the same solutions as described in the previous section to reduce the number of computations. The results from these simulations are used to determine the error between the experimental results and numerical results by using the following equation.

$$\epsilon_\phi = \frac{\phi_s - \phi_e}{\phi_e} \times 100\% \tag{14}$$

To determine if the interpolation-based or regression-based metamodel with the mean or mean-variance calibration results in the best performing DEM model the results for 50 repetitions are compared in more detail.

5. Results

In this section the results of this study are presented. Section 5.1 starts with an evaluation of the quality of the interpolation-based and regression-based metamodels. Next, the calibration results are evaluated in Section 5.2 followed by the verification of these results in Section 5.3. In this section the quality of metamodel prediction of the KPIs is verified first followed by a comparison between the predicted KPIs and the calibration targets. In Section 5.4 the validation results are presented for the large-scale hopper application.

5.1. Metamodels for calibration

The quality of the metamodel determines the quality of the DEM parameter set found through optimization. To gain insight on the quality of the metamodel the training error of the model can be assessed. Because this error in training does not give insight in the quality of the predictions in between training points, in addition the predictions by the metamodel are compared to a validation set.

5.1.1. Training error

The calibration models have been trained according to the method described in Section 4.3. For these models, the normalized root mean squared error (NRMSE) is calculated such that the accuracy in training of the model can be evaluated and compared for the different KPIs, (Eq. (15)).

$$NRMSE = \frac{\sqrt{\frac{\sum_{i=1}^{N}\left(f_i^*(x) - f_i(x)\right)^2}{N}}}{\bar{f}} \tag{15}$$

In Fig. 11 the NRMSE for the mean, mean ± standard deviation for the interpolation-based and regression-based metamodels are shown. In Fig. 11 (a) to (c) it can be seen that the error of the interpolation-based metamodel is close to machine precision which can be regarded as numerical noise. This can be explained by the fact that the interpolation-based models are exact in the data-points. In Fig. 11 (d) to (f) the RSE for the stochastic models is shown where errors in the mean prediction are small but errors in the mean ± standard deviation are significant. The trend of the RSE of the Angle of Repose and Movement indicates that increasing the number of repetitions leads to a decreasing error up to 20 repetitions after which stabilization occurs. The NRMSE for the bulk density is below 0,0005 which indicates an accurate metamodel. The majority of the improvement of the regression-based metamodel accuracy for the Angle of Repose and Movement occurs up to the first 20 repetitions, which is an indication of the importance of repeating simulations to obtain accurate metamodels. Based on the training error the RBM of the bulk density is best, followed by the RBM of the AoM, and lastly the AoR.

5.1.2. Validation error

The results in Fig. 11 show that the interpolation-based metamodels (IBM) are more accurate in the data-points than the regression-based metamodels (RBM). However, the accuracy of a metamodel is mainly determined by its ability to predict KPI values at locations in between the used training set. To verify this ability a validation strategy such as validation set approach (VSA), Cross-validation, or leave one out cross validation (LOOCV) should be used [8]. Therefore, an additional 32 data point sampling set has been simulated for 50 repetitions such that the 20% validation set approach can be applied. The results in Fig. 12 show the NRMSE error for the interpolation-based metamodel in (a-c) and for the regression-based metamodel in (d-f).

For all KPIs and both the IBM and RBM approach it shows that the addition of repetitions to the dataset leads to a decrease in NRMSE. The difference between the IBM and RBM approach is that the NRMSE of the IBM develops less consistent than the error of the RBM. Considering the mean prediction of the AoR and AoM the NRMSE is 1,5 to 2,5 times lower for the RBM models than for the IBM models. This means that the RBM models are more reliable in predicting the mean. For the mean prediction of the bulk density we see that both the IBM and RBM metamodels lead to similar levels of accuracy. With respect to the bandwidth of the standard deviation around the mean it is observed that for the IBM the error is in the same range as the mean error. This indicates that these models have a small prediction error in the magnitude of the standard deviation but that the error of the mean causes a large shift of the bandwidth. For the RBM of the AoR we see different behaviour. In addition to a shift of the bandwidth due to the mean error it is observed that the mean ± standard deviation has a larger error for the lower limit than for the upper limit. This indicates an additional error introduced by the RBM in the prediction of the variance. This same behaviour is also visible for the AoM but has a smaller magnitude compared to the AoR results. In contrast to the mean ± standard deviation of the bulk

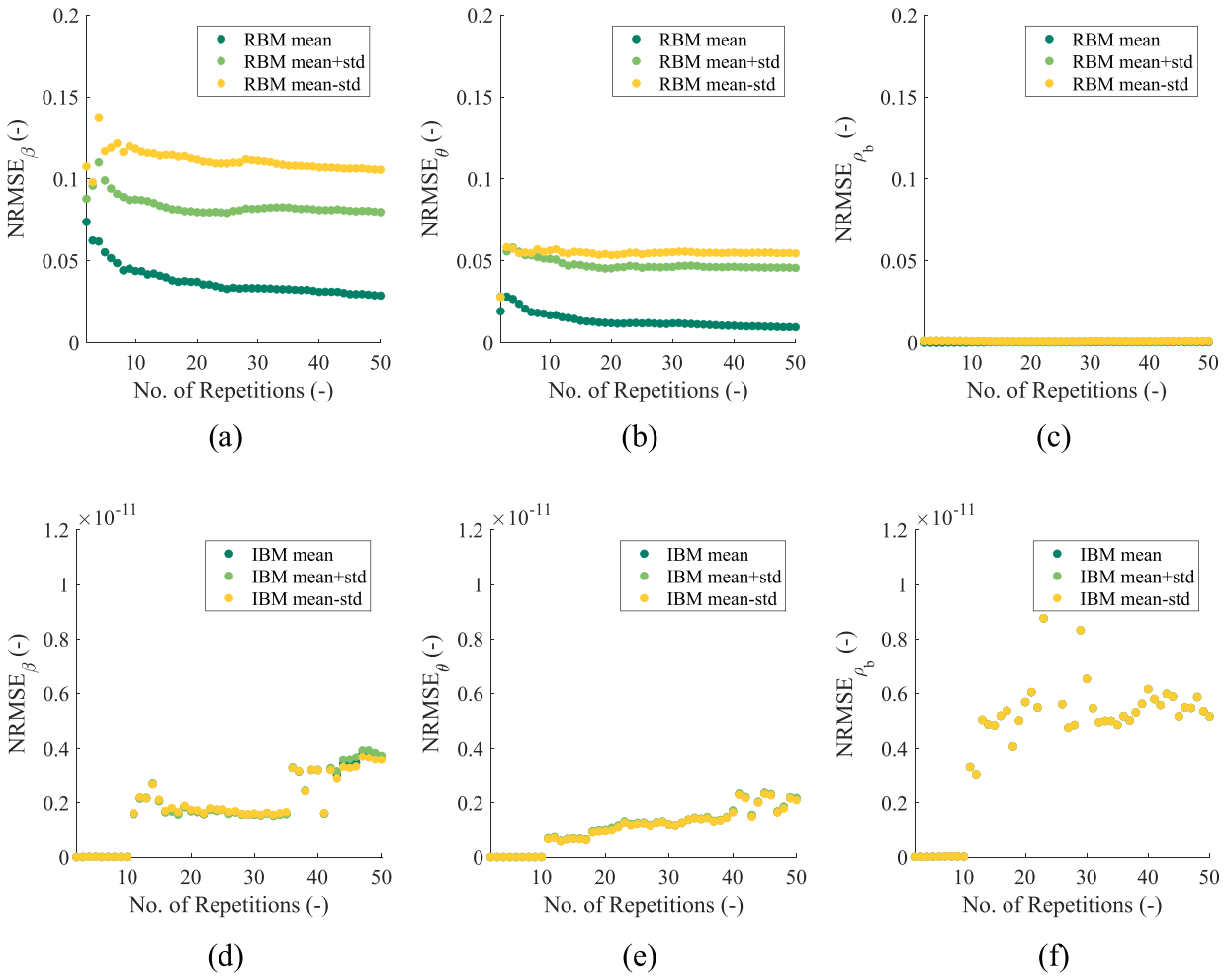

Fig. 11. Normalized Root Mean Squared Error (NRMSE) in the training points for interpolation-based metamodels (a,b,c) and regression-based metamodels (d,e,f) for the Angle of Repose β, Angle of Movement θ, and Bulk Density ρ_b data.

density for the IBM the bandwidth error of the RBM decreases fast in the first 20 repetitions.

Overall, the regression-based metamodel is more reliable in predicting the mean value which results in smaller offsets in the location of the bandwidth of the standard deviation around the mean. Due to the error in mean prediction by both metamodel types the standard deviation bandwidth has an offset. For the IBM the prediction of the standard deviation is accurate in magnitude but has an offset in location due to the error of the mean. The RBM has an inaccurate prediction of the standard deviation magnitude and an induced offset due to the error in the mean prediction leading to an asymmetric bandwidth of the standard deviation.

5.2. Calibration results

The resulting DEM parameter values of the mean (M) and mean-variance (MV) calibration with the interpolation- and regression-based metamodels (IBM & RBM) are shown in Fig. 13. For all three calibration parameters it is observed that the calibration with interpolation-based metamodels results in irregular development of the parameter values μ_s, μ_r, and ρ_p with the increase of the number of repetitions. On the contrary, the regression-based metamodel calibration presents a relatively smooth and steady parameter evolution for an increasing number of repetitions. This indicates that the optimal DEM parameter set is approximately reached at 20 repetitions, which is useful information if this method is applied in engineering practice. Note also that at a more conventional amount of 3–5 repetitions, the parameter values have not stabilized for either method.

In Fig. 14 (a–f) the mean and mean ± standard deviation of the KPI values are presented corresponding to the optimal DEM parameter values in Fig. 13. For all four calibration approaches, Fig. 14 (a) and (b) shows the Angle of Repose as a function of the number of repetitions. Fig. 14 (c) and (d) show the Angle of Movement and (e) and (f) show the bulk density. Figures (a,c,e) show the results obtained with the mean and mean-variance calibration approach with the interpolation-based metamodels, and Figures (b, d, f) show the results obtained with the regression-based metamodels. With respect to the IBM and RBM calibration the clear difference is the consistency of the calibration results obtained with the RBM compared to the IBM calibration results. As was observed with the validation of the metamodel in Section 5.1.2 the IBM models are more sensitive to changes in the dataset than RBM models.

Even though IBM is more sensitive to changes in the dataset, the performance for both mean and mean-variance calibration with the IBM is good. For the mean calibration, the match of the AoR and AoM is good over the entire range and the bulk density has a good match after 36 repetitions. The match for mean ± standard deviation is close for the AoR, twice the target for the AoM and half of the target for the bulk density. This is reasonable because the mean ± standard deviation is not included in mean calibration. For the mean-variance calibration the match of the mean value of the AoR is good. The match to the mean of the AoM is quite good but less than for the mean calibration. The mean of the

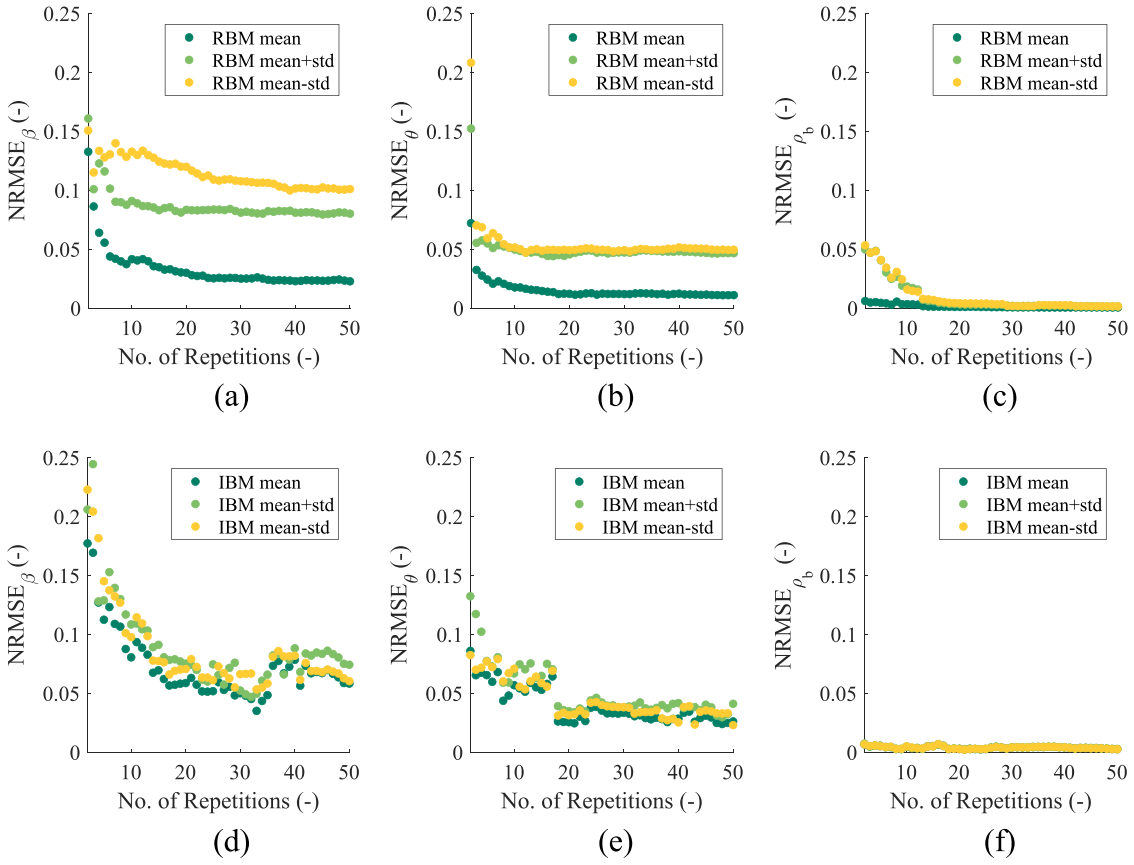

Fig. 12. Normalized Root Mean Squared Error (NRMSE) in the validation points for interpolation-based metamodels (a,b,c) and regression-based metamodels (d,e,f) for the Angle of Repose β, Angle of Movement θ, and Bulk Density ρ_b data.

bulk density fluctuates considerably before converging to a relatively good match. The mean ± standard deviation is included in the mean-variance calibration approach and shows a slightly better match to the targets compared to the mean calibration approach.

The performance of the calibration approaches with the RBM models gives consistent results. However, the error between the calibration results and the target is significant and is around −5% for the AoR and + 5% for the AoM over the entire range of repetitions. For the bulk density a good fit to the calibration target was found after 10 repetitions. With respect to the mean ± standard deviation values the AoR shows that the bandwidth is half of the calibration target. For the AoM the bandwidth is similar to the target in magnitude and half of the target for the bulk density. Comparing the mean and mean-variance calibration approach with the RBM it is observed that there is no significant difference between the calibration results apart from a small shift. This means that in this case there is no added benefit by adding the variance to the calibration.

At 50 repetitions, the dataset for training the metamodels has the highest reliability and therefore we assume that the results of the calibration is the most accurate. In Table 7 the DEM parameter values for

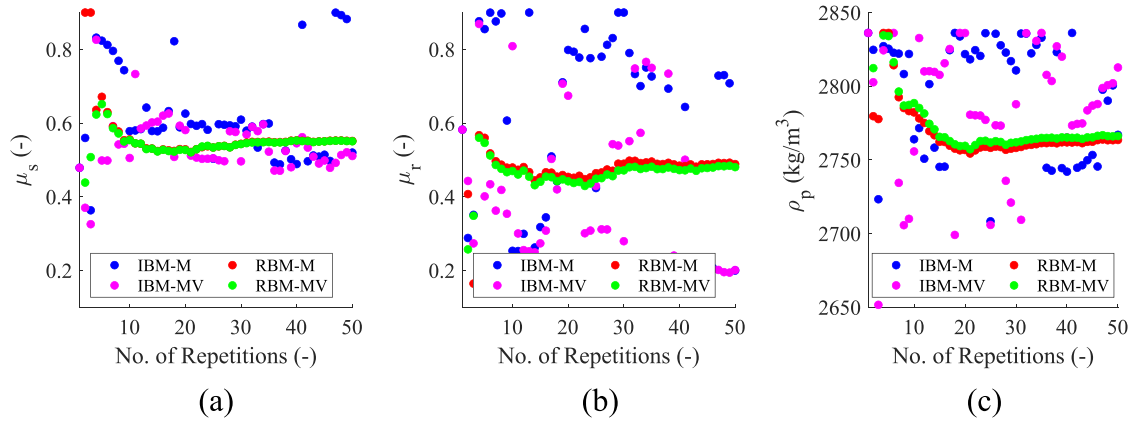

Fig. 13. Development of the DEM parameter values μ_s, μ_r, and ρ_p resulting from mean (M) and mean-variance (MV) calibration with interpolation- and regression-based metamodels (IBM & RBM) with increasing number of repetitions.

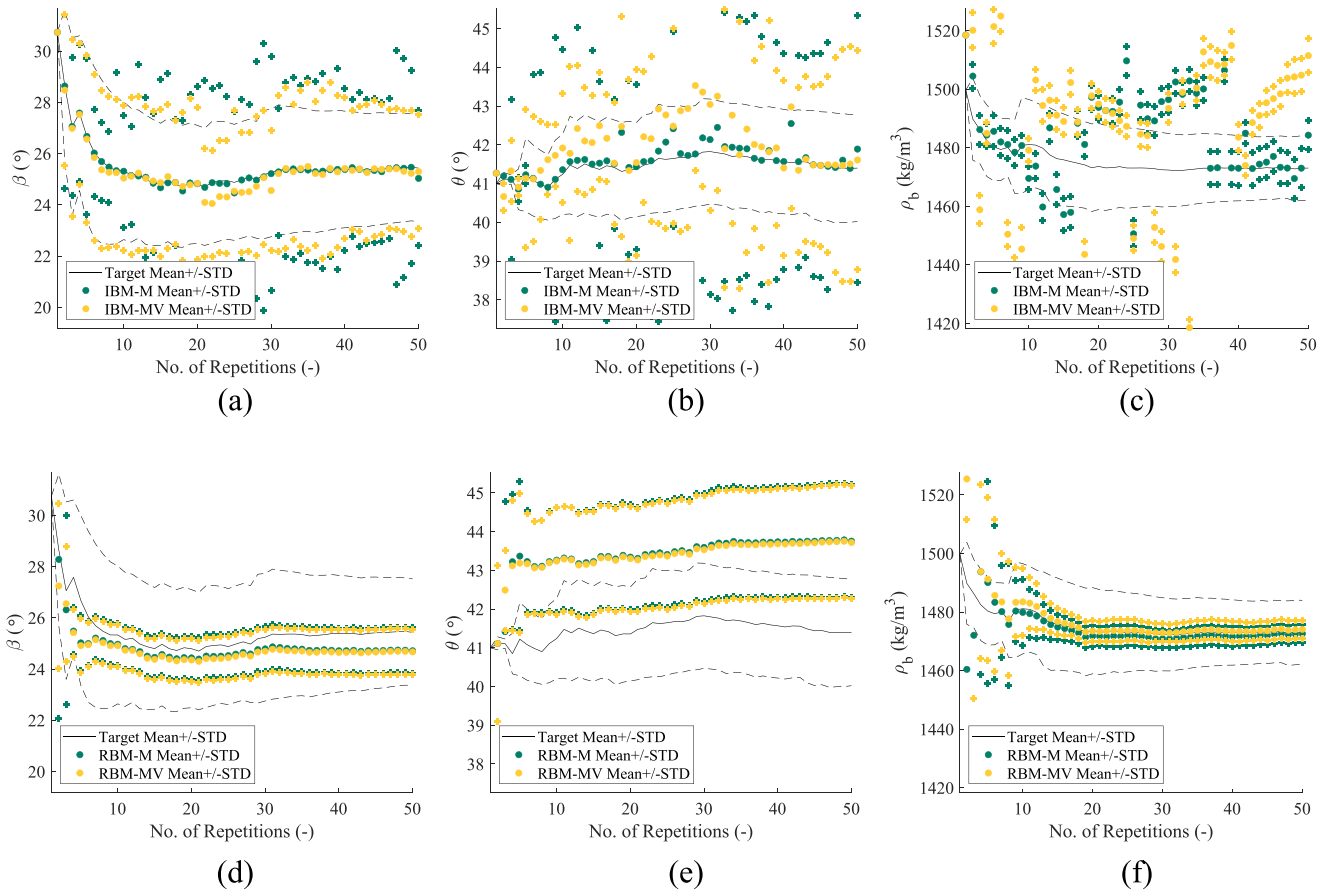

Fig. 14. Mean and mean-variance calibration results for the Angle of Repose β, Angle of Movement θ, Bulk Density ρ_b for the interpolation-based metamodel (a,c,e) and regression-based metamodel (b,d, f) respectively.

calibration with 50 repetitions are shown. For the IBM calibration it can be seen that the DEM parameter values are close. The largest difference can be observed for the particle density. For the RBM calibration we can see that the DEM parameters are more or less similar for the mean and mean-variance calibration. Between the interpolation-based and regression-based metamodel calibration it can be seen that the sliding friction is slightly higher, the rolling friction is almost 2.5 times higher, and the particle density is in the same range.

5.3. Verification of laboratory scale simulations and validation with experimental results

In this section the verification results from the mean (M) and mean-variance (MV) calibration for the interpolation- and regression-based metamodels (IBM & RBM) are presented. In Section 5.3.1 the results from the verification simulations are compared to the KPI values corresponding to the calibration results. This comparison gives an indication of the quality of the metamodel predictions in the calibration procedure. In Section 5.3.2 the results from the verification simulations are compared to the calibration targets which gives insight on the quality of calibration results. For the DEM parameter sets of the cases where [1,..,10, 15, 20, 30, 40, 50] repetitions were used the simulations of the laboratory scale experiments are verified.

Table 7
DEM parameters after calibration for mean and mean-variance calibration with interpolation-based and regression-based metamodels at 50 repetitions of the sample.

	IBM-M	IBM-MV	RBM-M	RBM-MV
μ_s	0,5198	0,5106	0,5513	0,5506
μ_r	0,1996	0,2010	0,4885	0,4804
ρ_p	2767,0	2812,7	2763,5	2766,2

5.3.1. Comparison KPI verification simulations and calibration results

The calibration results are a prediction by the metamodel which can include errors as shown with the training and validation error evaluation in Section 5.1. In Fig. 15 (a–d) the error percentage between the calibration and verification results for the mean of each KPI and mean ± standard deviation of each KPI are shown to indicate the magnitude of this metamodel prediction error. For the IBM-M and IBM-MV results in Fig. 15 (a) and (b) large errors with respect to the calibration results is seen for the Angle of Repose β and the Angle of Movement θ when the number of repetitions is below 10. For the IBM-M calibration results the error in the Angle of Movement stays large until 30 repetitions and reduces for a higher number of repetitions. On the contrary, the Angle of Repose becomes more accurate after 10 repetitions but becomes inaccurate after 30 repetitions. This might be caused by overfitting which is a known risk of interpolation-based metamodels. Fig. 15 (b) shows similar behaviour for the IBM-M and IBM-MV calibration results, however the point of change is located at 20 repetitions instead of 30. Overall, the verification results show that for mean and mean-variance calibration with interpolation-based metamodels large errors can occur even at higher numbers of repetitions. For the error in the mean ± standard deviation bandwidth asymmetry can be observed which is caused by the inaccurate prediction of the mean. The Angle of Repose exhibits an error of −10% for the upper limit and − 20% for the lower limit. The error is within 5% for the Angle of Movement and is below 1% for the Bulk Density.

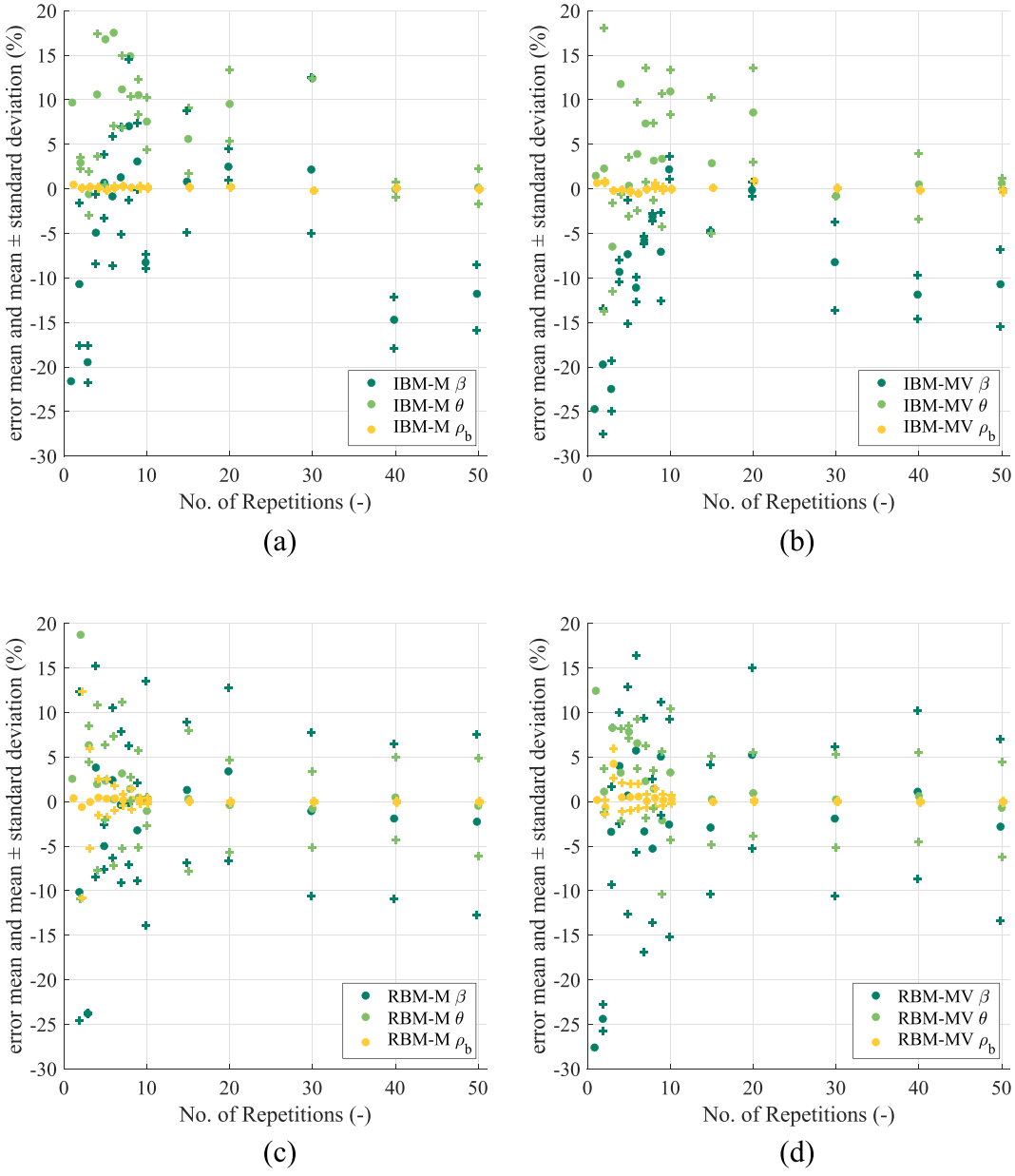

Fig. 15. Percentage error for the mean (·) and mean ± standard deviation (+,Δ) for the Angle of Repose β (dark green), Angle of Movement θ (green), Bulk Density ρ_b (yellow) for (a) interpolation-based metamodel mean calibration (b) interpolation-based metamodel mean-variance calibration (c) regression-based metamodel mean calibration (d) regression-based mean-variance calibration.

In Fig. 15 (c) and (d) the verification error of the regression-based metamodel calibration results are shown. For the mean prediction by both mean and mean-variance calibration results large errors are observed up to ten repetitions but the magnitude is smaller than for the IBM calibration results. The errors for the mean ± standard deviation are significant. After 10 repetitions, the verification error is close to zero for the Angle of Movement and bulk density for both the mean and mean-variance calibration. For the Angle of Repose the error stays within a 5% bandwidth when the number of repetitions is increased. The error in the mean ± standard deviation bandwidth is quite symmetric because of the accurate prediction of the mean value. However, the magnitude shows an error of 15% for the Angle of Repose, 7% for the Angle of movement and below 1% for the Bulk Density. As expected for both IBM and RBM the error in the Bulk density prediction is small because the error in the validation of both metamodels was small.

5.3.2. Comparison KPI verification simulations and calibration targets

The verification simulations show that RBM metamodels give more reliable calibration results. However, the next step is to look at the quality of the calibration results compared to the calibration targets. In Fig. 16 the percentage error between the calibration results and calibration targets are shown for the mean and mean-variance calibration with the IBM and RBM metamodels. While the calibration results for the interpolation-based metamodels showed a good fit to the calibration results in Fig. 14 it can be seen in Fig. 16 (a) and (b) that the error observed with the verification leads to errors with the calibration target. This is especially visible for the Angle of Repose up to 30 repetitions and Angle of Movement after 30 repetitions. For the IBM calibration it is observed that the calibration results are inconsistent with the increase of the number of repetitions. The RBM calibration results in Fig. 16 (c) and (d) show more consistency. Especially after 30 repetitions a

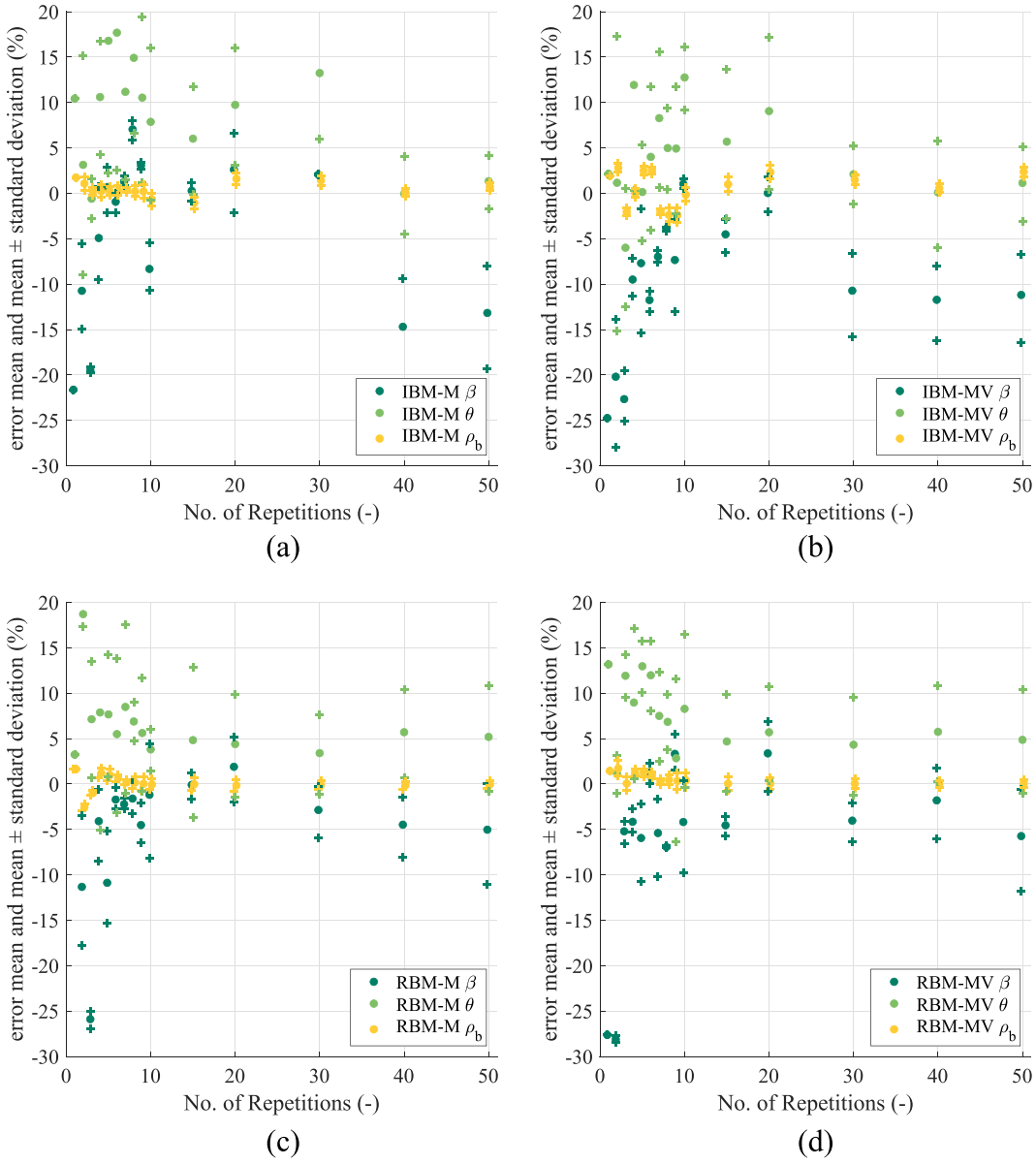

Fig. 16. Error verification results to calibration targets.

balance can be found between the error for the Angle of Repose and Movement where the percentage error stays the same. The stability and consistency in these results make the regression-based metamodel calibration more reliable. There are no significant differences observed between the mean and mean-variance calibration.

To determine which calibration method performs best the results for 50 repetitions are compared to the experimental results for the calibration targets. In Table 8 the percentage error between the verification result and calibration target is shown for 50 repetitions. At 50 repetitions, the percentage error for the calibration with interpolation-based metamodels is between 11 and 14% for the Angle of Repose, 1 and 1,5% for the Angle of Movement, and 0,5 to 2,5% for the Bulk Density. The results for the calibration with regression-based metamodels shows errors between 5 and 6% for the Angle of Repose, 4,5 and 5,5% for the Angle of Movement and 0,1% for the Bulk Density. Here it is clearly visible that even though the errors in matching the calibration targets is quite large the calibration results with regression-based metamodels are more balanced than those of the interpolation-based metamodels. Due to the errors in the mean value of the KPIs the bandwidth of the standard deviation will exhibit an asymmetric error as presented in the bottom section of Table 8. On average, the mean-calibration with the regression-based metamodels presents the best match to the experimental calibration targets and gives the most reliable DEM parameter set for this case.

Table 8
Percentage error between verification results and experimental calibration target values where the underlined results show the most accurate result.

	IBM mean	IBM mean-var	RBM mean	RBM mean-var
Mean KPI % error				
AoR	-13,17	-11,19	-5,02	-5,73
AoM	1,34	1,14	5,19	4,87
BD	0,71	2,37	-0,04	0,06
Mean ± standard deviation KPI % error				
AoR	-8,00 / -19,28	-6,73 / -16,44	0,05 / -11,01	-0,62 / -11,75
AoM	4,21 / -1,73	5,07 / -3,07	10,83 / -0,83	10,36 / -0,99
BD	0,28 / 1,14	1,84 / 2,91	-0,47 / 0,38	-0,39 / 0,52

Based on these results it can be seen that the predictions by regression-based metamodels are more reliable which makes these a better choice for the calibration of the DEM parameters even though in the final results the predictability of the mean ± standard deviation is poor. The discrepancy in the mean ± standard deviation might also be a result of the modelling assumptions for the DEM models of the laboratory experiments themselves, where the stochastic behaviour as seen in experiments is captured to a lesser extent. As shown in the results, there is no significant difference in the outcome of the mean vs. mean-variance calibration using the regression-based metamodels. In addition, the calibration procedure using multiple laboratory setups leads to a multi-objective problem for which it is not certain that there is an optimal solution matching all the targets perfectly. As decided in Section 4.3 the weight of each objective was kept to one for the mean and 0,5 for the variance interval. By changing these values objective weights can be changed, ultimately affecting the outcome of the calibration. An in-depth investigation of this influence is however outside of the scope of this study.

5.4. Validation of hopper DEM model

In this section we present the validation of the DEM model of the hopper described in Section 3 with the experimental results described in Section 2. For this model the discharge rate of the hopper during steady flow is compared for the DEM model of the hopper and the experiments. The calibration of the DEM parameters lead to 4 sets (IBM-M, IBM-MV, RBM-M, RBM-MV) of DEM parameters for each number of repetitions that was used in training the metamodels for the calibration. For the DEM parameter sets of the cases where [1,..,10, 15, 20, 30, 40, 50] repetitions were used the hopper simulations have been carried out five times, the same as the number of experiments. In Fig. 17 the obtained mean and standard deviations for the discharge rate are shown for each of the cases. In these results a large difference between the results obtained with the interpolation-based and regression-based metamodels can be observed. The interpolation-based metamodel calibration results show for both the mean and mean-variance calibration that there is much fluctuation in the results. The mean-calibration mean error follows an irregular path which is close at 30 repetitions but far from the validation target for the other points. In the mean-variance case the mean error tends to decrease up to 15 repetitions after which the error starts increasing continuously. In terms of the bandwidth of the mean ± standard deviation large fluctuations are present for the mean calibration but less for the mean-variance calibration. Indicating that including the variance in the calibration has a certain effect. Compared to the bandwidth of the validation target the bandwidth in the simulation results is larger. For the regression-based metamodel calibration the results are closer to the experimental results for the entire range of repetitions. Both mean and variance calibration cases show a stable offset which starts decreasing after 15 repetitions. For the mean ± standard deviation it is observed that after 10 repetitions the bandwidth stabilises at a bandwidth that is around two times the bandwidth of the validation target. From these results it can be seen that the calibration with the regression-based metamodels leads to more stable output compared to the calibration with the interpolation-based metamodels. Mean-variance calibration shows slightly better performance than the mean calibration but in terms of confidence intervals with five repetitions this observation is not significant.

Table 9
Results for the mean and standard deviation of the discharge rate including the 95% CI of the mean and standard deviation of the design $W_0 = 100$, $\alpha = 45$.

Model	Discharge rate ϕ $\left(\frac{kg}{s}\right)$	Percentage error with respect to the experimental results
	mean (95 % *CI* [*LL*; *UL*]) ± *sigma* (95 % *CI* [*LL*; *UL*])	
Experimental results	3,70 (95 % *CI* [3,67; 3,73]) ± 0,026 (95 % *CI* [0,014; 0,056])	Reference
IBM-M	3,86 (95 % *CI* [3,77; 3,95]) ± 0,081 (95 % *CI* [0,036; 0,260])	*mean error* 4,84 % (*mean* ± *std*) *error* (3,33%; 6,33%)
IBM-MV	3,98 (95 % *CI* [3,92; 4,04]) ± 0,052 (95 % *CI* [0,023; 0,167])	*mean error* 8,15 % (*mean* ± *std*) *error* (7,46%; 8,83%)
RBM-M	3,81 (95 % *CI* [3,76; 3,85]) ± 0,038 (95 % *CI* [0,017; 0,121])	*mean error* 3,34 % (*mean* ± *std*) *error* (3,01%; 3,67%)
RBM-MV	3,76 (95 % *CI* [3,62; 3,80]) ± 0,037 (95 % *CI* [0,016; 0,118])	*mean error* 2,15 % (*mean* ± *std*) *error* (1,84%; 2,46%)

Table 9 shows the mean and the standard deviation of the discharge rate for the calibration results with 50 repetitions. For both these results the 95% confidence intervals are also shown based on a t-distribution where $n = 5$. For these intervals we applied the same method as described in Section 2. In the third column the relative errors to the mean and the mean plus minus the standard deviation are shown.

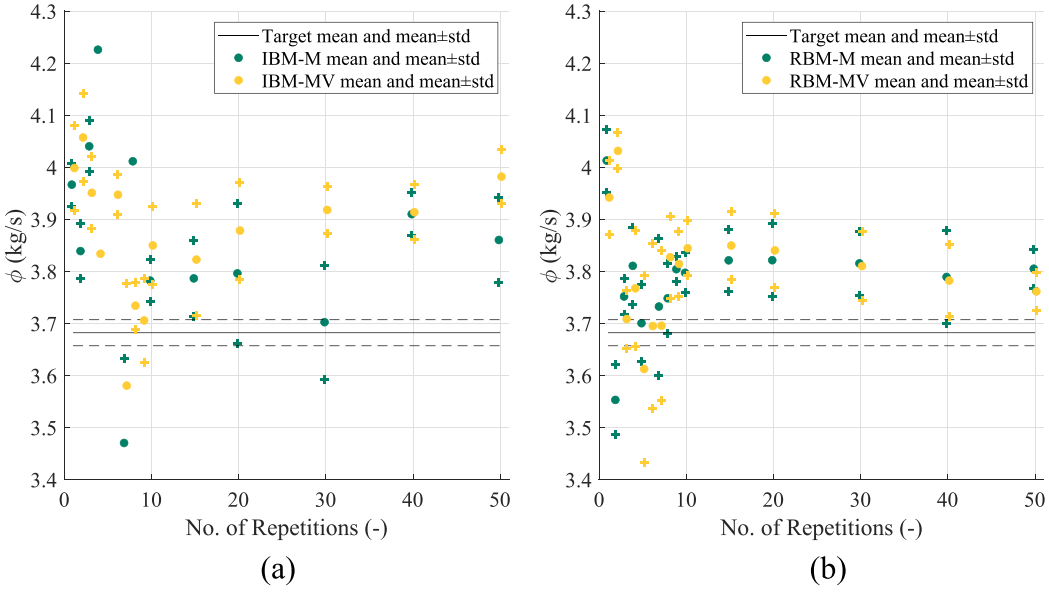

Fig. 17. Results for the validation simulations of the hopper with respect to the number of repetitions.

The results show that for the discharge rate the mean-variance calibration with the stochastic model gives the most accurate prediction of the discharge rate (RBM-MV). This model is followed by the mean calibration with the stochastic model (RBM-M), the mean calibration with the deterministic model (IBM-M), and lastly the mean-variance calibration with the deterministic model (IBM-MV). Differences of up to a factor 6 are found in the mean error percentage, while all methods use the same dataset. This illustrates the importance of using the correct calibration procedure. Overall, it is observed that the variance prediction shows large errors with respect to values found in the experiments. This indicates that the results from the DEM model of the hopper with the optimal DEM parameter sets have a higher variability than those from the experiments. Reasons for this observation might be the simplification of particle shape from irregular to spherical and uniformity assumption of material properties and contact parameters in the DEM model.

6. Conclusions

Experiments with granular materials often exhibit significant variation, which have to the best of the authors' knowledge not been included in calibration approaches. In this paper we investigated the inclusion of stochastics throughout the development of the DEM model, in the calibration, verification with lab experiments, and validation with a hopper application. We used metamodels to predict KPI values and proposed a mean-variance calibration approach that includes the variance of the calibration experiments in the objectives of the calibration procedure. The conclusions of this study are as follows.

- The calibration experiments showed stability for the KPI mean after 5 repetitions for the bulk density, 10 for the Angle of Movement, and 30 for the Angle of Repose. However, in practice 3–5 repetitions is common which means that increasing the number of repetitions should be considered for obtaining reliable calibration targets.
- Stabilization of experimentally obtained variances occurs at a higher number of repetitions than stabilization of the mean, which is an additional motivation to consider the required number of repetitions.
- Based on the verification results in this study, regression-based metamodels prove to be better at predicting mean KPI values than interpolation-based metamodels, both in terms of accuracy and number of required repetitions, and are therefore recommended in metamodel or surrogate model based calibration.
- Both regression-based and interpolation-based metamodels give inaccurate predictions of the variance for the DEM models of both the calibration and the validation experiment. The cause of these inaccuracies might be related to the modelling assumptions in the DEM model or to the used number of repetitions in training the metamodel.
- Accurate prediction of KPI variances by metamodels proves to be harder than predicting KPI means. Therefore, further research is needed on accurately including the variance in DEM model calibration.

Although this study focused on a particular granular material, DEM model and application example, it is expected that the observed trends and the associated conclusions extend to a wide range of use cases, due to the inherent stochastic nature of granular processes. Therefore, it is recommended to conduct further research on including and identification of stochastic behaviour in granular processes and their implications on DEM modelling of particulate systems.

CRediT authorship contribution statement

Marc Patrick Fransen: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Project administration, Visualization. **Matthijs Langelaar:** Supervision, Writing – review & editing. **Dingena L. Schott:** Supervision, Writing – review & editing, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] L. Benvenuti, C. Kloss, S. Pirker, Identification of DEM simulation parameters by artificial neural networks and bulk experiments, Powder Technol. 291 (2016) 456–465, https://doi.org/10.1016/j.powtec.2016.01.003.
[2] R.H. Byrd, J.C. Gilbert, J. Nocedal, A trust region method based on interior point techniques for nonlinear programming, Math. Prog. Series B 89 (1) (2000) 149–185, https://doi.org/10.1007/PL00011391.
[3] H. Cheng, T. Shuku, K. Thoeni, P. Tempone, S. Luding, V. Magnanimo, An iterative Bayesian filtering framework for fast and automated calibration of DEM models, Comput. Methods Appl. Mech. Eng. 350 (2019) 268–294.
[4] H. Cheng, T. Shuku, K. Thoeni, H. Yamamoto, Probabilistic calibration of discrete element simulations using the sequential quasi-Monte Carlo filter, Granul. Matter 20 (1) (2018) 1–19, https://doi.org/10.1007/s10035-017-0781-y.
[5] C. Coetzee, Calibration of the discrete element method: strategies for spherical and non-spherical particles, Powder Technol. 364 (2020) 851–878, https://doi.org/10.1016/j.powtec.2020.01.076.
[6] C.J. Coetzee, Review: calibration of the discrete element method, Powder Technol. 310 (2017) 104–142, https://doi.org/10.1016/j.powtec.2017.01.015.
[7] S.M. Derakhshani, D.L. Schott, G. Lodewijks, Micro – macro properties of quartz sand : experimental investigation and DEM simulation, Powder Technol. 269 (2015) 127–138, https://doi.org/10.1016/j.powtec.2014.08.072.
[8] M.P. Fransen, M. Langelaar, D.L. Schott, Application of DEM-based metamodels in bulk handling equipment design : methodology and DEM case study, Powder Technol. 393 (2021) 205–218, https://doi.org/10.1016/j.powtec.2021.07.048.
[9] C. González-Montellano, Á. Ramírez, E. Gallego, F. Ayuga, Validation and experimental calibration of 3D discrete element models for the simulation of the discharge flow in silos, Chem. Eng. Sci. 66 (21) (2011) 5116–5126, https://doi.org/10.1016/j.ces.2011.07.009.
[10] P. Hartmann, H. Cheng, K. Thoeni, Performance study of iterative Bayesian filtering to develop an efficient calibration framework for DEM, Comput. Geotech. 141 (September 2021) (2022) 104491, https://doi.org/10.1016/j.compgeo.2021.104491.
[11] A. Jensen, K. Fraser, G. Laird, N. Karajan, Z. Han, H. Teng, J. Wang, M.T. Davidson, J.H. Chung, H. Teng, Z. Han, V. Le, M.A. Faraone, J.H. Chung, M.T. Davidson, P.O. Box, M.B. Cil, K.A. Alshibli, A. Jensen, ... J.O. Trierweiler, Improving the precision of discrete element simulations through calibration models, 13 Th International LS-DYNA Users Conference 7 (88) (2014) 405–410.
[12] R. Jin, X. Du, W. Chen, The use of metamodeling techniques for optimization under uncertainty, Struct. Multidiscip. Optim. 25 (2003) 99–116, https://doi.org/10.1007/s00158-002-0277-0.
[13] S.W. Lommen, Virtual Prototyping of Grabs. TU Delft, 2016.
[14] S. Luding, Anisotropy in cohesive, frictional granular media, J. Phys. Condens. Matter 17 (24) (2005) https://doi.org/10.1088/0953-8984/17/24/017.
[15] S. Luding, Cohesive, frictional powders: contact models for tension, Granul. Matter 10 (2008) 235–246, https://doi.org/10.1007/s10035-008-0099-x.
[16] M. Marigo, E.H. Stitt, Discrete element method (DEM) for industrial applications: comments on calibration and validation for the modelling of cylindrical pellets, KONA Powder Part. J. 32 (32) (2015) 236–252, https://doi.org/10.14356/kona.2015016.
[17] M.J. Mohajeri, W. de Kluijver, R.L.J. Helmons, C. van Rhee, D.L. Schott, A validated co-simulation of grab and moist iron ore cargo: replicating the cohesive and stress-history dependent behaviour of bulk solids, Adv. Powder Technol. 32 (4) (2021) 1157–1169, https://doi.org/10.1016/j.apt.2021.02.017.
[18] M.J. Mohajeri, H.Q. Do, D.L. Schott, DEM calibration of cohesive material in the ring shear test by applying a genetic algorithm framework, Adv. Powder Technol. 31 (5) (2020) 1838–1850, https://doi.org/10.1016/j.apt.2020.02.019.
[19] M. Rackl, W.A. Günthner, Biomass and bioenergy experimental investigation on the in fl uence of different grades of wood chips on screw feeding performance, Biomass Bioenergy 88 (2016) 106–115, https://doi.org/10.1016/j.biombioe.2016.03.011.
[20] M. Rackl, K.J. Hanley, A methodical calibration procedure for discrete element models, Powder Technol. 307 (2017) 73–83, https://doi.org/10.1016/j.powtec.2016.11.048.
[21] C.E. Rasmussen, C.K.I. Williams, Gaussian Processes for Machine Learning, Vol. 7, MIT Press, 2006 Issue 5.
[22] C. Richter, T. Rößler, G. Kunze, A. Katterfeld, F. Will, Development of a standard calibration procedure for the DEM parameters of cohesionless bulk materials – part II: efficient optimization-based calibration, Powder Technol. 360 (2020) 967–976, https://doi.org/10.1016/j.powtec.2019.10.052.
[23] T. Roessler, A. Katterfeld, Scaling of the angle of repose test and its influence on the calibration of DEM parameters using upscaled particles, Powder Technol. 330 (2018) 58–66, https://doi.org/10.1016/j.powtec.2018.01.044.
[24] T. Roessler, C. Richter, A. Katterfeld, F. Will, Development of a standard calibration procedure for the DEM parameters of cohesionless bulk materials – part I: solving

the problem of ambiguous parameter combinations, Powder Technol. 343 (2019) 803–812, https://doi.org/10.1016/j.powtec.2018.11.034.
[25] T.A.H. Simons, R. Weiler, S. Strege, S. Bensmann, A ring shear tester as calibration experiment for DEM simulations in agitated mixers - a sensitivity study, Proc. Eng. 102 (2015) 741–748, https://doi.org/10.1016/j.proeng.2015.01.178.
[26] T. Weinhart, L. Orefice, M. Post, M.P. van Schrojenstein Lantman, I.F.C. Denissen, D.R. Tunuguntla, J.M.F. Tsang, H. Cheng, M.Y. Shaheen, H. Shi, P. Rapino, E. Grannonio, N. Losacco, J. Barbosa, L. Jing, J.E. Alvarez Naranjo, S. Roy, W.K. den Otter, A.R. Thornton, Fast, flexible particle simulations — an introduction to MercuryDPM, Comput. Phys. Commun. 249 (2020) 107129, https://doi.org/10.1016/j.cpc.2019.107129.
[27] C.M. Wensrich, A. Katterfeld, Rolling friction as a technique for modelling particle shape in DEM, Powder Technol. 217 (2012) 409–417, https://doi.org/10.1016/j.powtec.2011.10.057.
[28] Y. Yan, R. Helmons, C. Wheeler, D. Schott, Optimization of a convex pattern surface for sliding wear reduction based on a definitive screening design and discrete element method, Powder Technol. 394 (2021) 1094–1110, https://doi.org/10.1016/j.powtec.2021.09.041.
[29] F. Zamponi, Mathematical physics: packings close and loose, Nature 453 (7195) (2008) 606–607, https://doi.org/10.1038/453606a.
[30] Y.C. Zhou, B.H. Xu, A.B. Yu, P. Zulli, Numerical investigation of the angle of repose of monosized spheres, Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics 64 (2) (2001) 8, https://doi.org/10.1103/PhysRevE.64.021301.
[31] González-Montellano, et al., Determination of the mechanical properties of maize grains and olives required for use in DEM simulations, Journal of Food Engineering (2012).