

Data-driven Modelling of Granular Column Collapse

Qinghao Yang¹ and James P. Hambleton, Ph.D., M.ASCE²

¹Department of Civil and Environmental Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60201; E-mail: qinghaoyang2022@u.northwestern.edu

²Department of Civil and Environmental Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60201; E-mail: jphambleton@northwestern.edu

ABSTRACT

Data-driven methods have recently shown their potential in engineering problems. In this paper, a data-driven model is proposed and tested for 2D granular column collapse, a typical problem involving flow-like behavior of granular material. Although many theoretical and numerical models have been developed to investigate granular flows, data-driven methods have yet to be explored. This model can predict granular flows based on the sequential relations over time with data collected from numerical solutions. Two frameworks, a Lagrangian framework and an Eulerian framework, are proposed to give rules for collecting data, and radial basis function networks are used to infer the sequential relations. It is shown that the data-driven model gives reasonable predictions compared to results obtained through direct solution. When the size of data in both frameworks are close, the performance in the Eulerian framework is better, especially for predicting states of columns with different geometric parameters. A sensitivity analysis to explore the influence of time increment and grid spacing is also conducted.

INTRODUCTION

Recently, with the growth of accessible data in various disciplines, data-driven methods have been applied in simulation of different physical and engineering problems. Raissi et al. (2019) used physics-informed neural networks, where neural networks are constrained by physical governing equations, to solve time-dependent partial differential equations and discover partial differential equations from collected data. Kirchdoerfer and Ortiz (2016) proposed a data-driven framework directly using a database composed of stress-strain pairs, avoiding deriving empirical constitutive relations. Hesthaven and Ubbiali (2018) combined proper orthogonal decomposition and neural networks to reduce the computational time in high-fidelity models. Those studies show the potential of data-driven methods, especially when (1) the governing equations are not known or partially known, (2) the equations are difficult to solve numerically, or (3) the computational time is very large.

In terms of geological materials, Wang and Sun (2018) inferred their complex plastic behaviors via deep learning. However, for geological hazards such as landslides and snow avalanches, the prediction of post-failure movement of granular material (e.g., sand, rock, and snow) is as significant as the prediction of the onset of initial failure. Thus, the flow-like behavior of granular material needs to be considered in addition to the solid-like behavior. To investigate granular flow after failure, many experiments have been designed, including runout along an inclined plane, flow in a rotating drum, and granular column collapse (Forterre and Pouliquen,

2008). A variety of numerical methods have also been explored, including discrete element method (Utili et al., 2015), material point method (Fern and Soga, 2016), and smoothed-particle hydrodynamics (Bui et al., 2008).

To explore the possibility of applying data-driven methods to granular flows, this paper proposes a data-driven model that is implemented for the specific problem of granular column collapse. In this problem, restraints around a column of dry granular media are suddenly released, allowing the material to move downwards and spread laterally. The central idea is to collect data from a selected set of numerical solutions and then use this “precomputed data” to forecast the response for alternative configurations. This data-driven model can predict granular flows without knowledge of the underlying physics and without the use of traditional solution strategies.

In this paper, a continuum theory for avalanching of granular material, which has origins in the work of Savage and Hutter (1989), is solved numerically to generate data. In this approach, a rate-independent Coulomb-like friction law and depth-averaged variables (resembling those for shallow water flow) are assumed to derive governing equations from mass and momentum conservation laws. In subsequent research, this theory was extended to the runout of granular material over two-dimensional and curved slopes (Mangeney-Castelnau et al., 2003; Luca et al., 2009; Baker et al., 2016). After analyzing experimental data from granular flows down inclined planes, Pouliquen (1999) proposed an empirical equation describing the rate-dependent friction between granular material and a rigid base, substituting for the previously used rate-independent friction law. This model has also been used to simulate recorded landslides (Kuo et al., 2009; Hungr and McDougall, 2009; Ouyang et al., 2019). For column collapse, Kerswell (2005) derived governing equations and gave numerical solutions for 2D and axisymmetric problems.

With continuing innovations in measurement techniques, especially those in image processing (Lueptow et al., 2000) and remote sensing (Niethammer et al., 2012), a large amount of data that describes the motion of granular flows can be obtained from laboratory experiments and field tests. While the present work focuses purely on data from numerical solutions, the long-term goal of this work is to examine the possibility of integrating other data types, including experimental data.

PROBLEM FORMULATION

The paper focuses on the specific problem of granular column collapse. The boundary conditions are shown in Figure 1. A column is supported by a rigid and smooth wall on one side and released on the other side.

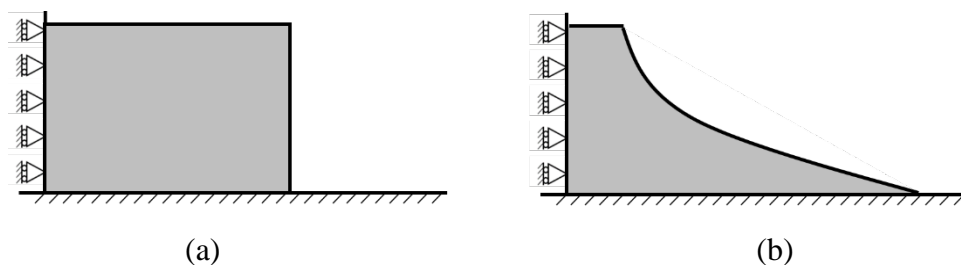


Figure 1: Schematics of walled granular column collapse: (a) initial profile; (b) final profile.

The depth-averaged continuum model for granular column collapse (Kerswell, 2005) is based on the following assumptions:

1. The granular material is cohesionless and obeys the Mohr-Coulomb yield criterion. The internal friction angle of the granular material and the friction angle between the sand and base are both constant.
2. The horizontal component of velocity is much greater than the vertical component, such that geostatic pressure is assumed in the vertical direction.
3. Depth-averaged velocity is used to approximate actual velocity.

For 2D column collapse on a rigid horizontal surface, let $\hat{h}(\hat{x}, \hat{t})$ be the height of the surficial point with horizontal coordinate \hat{x} at time \hat{t} , and let $\hat{u}(\hat{x}, \hat{t})$ be the depth-averaged velocity along the horizontal direction. Under these assumptions, the governing equations can be derived based on mass and momentum conservation laws as (Kerswell, 2005)

$$\hat{u}_{\hat{t}} + \hat{u}\hat{u}_{\hat{x}} + Kg\hat{h}_{\hat{x}} = -g\tan\theta, \quad (1)$$

$$\hat{h}_{\hat{t}} + \hat{u}\hat{h}_{\hat{x}} + \hat{h}\hat{u}_{\hat{x}} = 0, \quad (2)$$

where g is gravitational acceleration and K is an earth pressure coefficient determined by the material's internal friction angle ϕ and the friction angle between the sand and base θ . Let

$$\hat{x} = \frac{\hat{h}^0 K}{\tan\theta} x, \quad \hat{t} = \frac{1}{\tan\theta} \sqrt{\frac{\hat{h}^0 K}{g}} t, \quad \hat{u} = \sqrt{g\hat{h}^0 K} u, \quad \hat{h} = \hat{h}^0 h, \quad (3)$$

where \hat{h}^0 is the initial height of the column. The equations are transformed to dimensionless forms as

$$u_t + uu_x + h_x = -1, \quad (4)$$

$$h_t + uh_x + hu_x = 0. \quad (5)$$

Numerical solutions to Eqs. (4) and (5) can be obtained through a finite difference method (Savage and Hutter, 1989). In the remainder of the paper, results are presented in terms of dimensionless quantities t (time), x (horizontal coordinates), h (height), and u (horizontal velocity).

DATA-DRIVEN METHOD

Overview. The objective of the proposed method is to simulate column collapse directly using a database of “precomputed solutions” obtained from numerical solution of the governing equations. By modeling a number of 2D column collapses with different initial aspect ratios (column height divided by width) and collecting solutions with fixed time increment, the hidden relations between data in adjacent time can be estimated by conventional regression models or machine learning. Consequently, given a new aspect ratio, the granular flow process can be inferred step by step from collected solutions without any physical knowledge. In this paper, two data-driven frameworks are proposed.

Lagrangian framework. In this approach, collected data are described in a Lagrangian reference frame, which means that state variables are associated with material points that move as the material deforms. In contrast, in the Eulerian framework discussed in the next section, state variables are associated with spatial points that are fixed in space. In the initial step of the Lagrangian framework, material points are selected uniformly along the horizontal direction. In subsequent steps, the position (x), height (h), and horizontal velocity (u) of each material point can be collected. Although the positions of these material points are not fixed, it is straightforward to approximate their position in the next step using position and velocity of the current step

$$x_i^{n+1} = x_i^n + u_i^n \Delta t, \quad (6)$$

where the superscript denotes the step, the subscript denotes the label of the material point, and Δt is the time increment between adjacent steps. Therefore, the positions of material points can be tracked in all steps.

If state variables (x , h , u) for all material points in the current step are given, the state variables in the next step must be distinct, but how does one discover the sequential relations? If variables of all material points in the current step are taken as inputs to infer a variable in the next step, the number of input variables is too large to infer the relations with an acceptable number of collected data. This phenomenon is known as the ‘‘curse of dimensionality.’’ Based on the local governing equations, it is reasonable to assume that the state variables of a particular material point in the next step are only affected by the current state of that point and the states of adjacent material points. Therefore, functions for expressing sequential relations can be constructed as

$$h_i^{n+1} = \mathcal{H}^{\mathcal{L}}(x_{i-1}^n, x_i^n, x_{i+1}^n, h_{i-1}^n, h_i^n, h_{i+1}^n, u_{i-1}^n, u_i^n, u_{i+1}^n), \quad (7)$$

$$u_i^{n+1} = \mathcal{U}^{\mathcal{L}}(x_{i-1}^n, x_i^n, x_{i+1}^n, h_{i-1}^n, h_i^n, h_{i+1}^n, u_{i-1}^n, u_i^n, u_{i+1}^n). \quad (8)$$

Given initial conditions, the whole process of column collapse can be modeled sequentially after the functions $\mathcal{H}^{\mathcal{L}}$ and $\mathcal{U}^{\mathcal{L}}$ are inferred from collected data (Figure 2a).

Eulerian framework. In Eqs. (7) and (8), the three inputs ($x_{i-1}^n, x_i^n, x_{i+1}^n$) expressing the locations of material points can be reduced to two inputs ($x_{i-1}^n - x_i^n, x_{i+1}^n - x_i^n$) expressing the relative distance. This reduces the complexity of the sequential relations. Furthermore, if all distances between neighboring points are the same, the sequential relations will not be related to the horizontal coordinates of collected points. Therefore, in the Eulerian framework, the state variables at fixed spatial points with uniform spacing are collected step by step.

For distinction, uppercase letters are used to denote variables in the Eulerian framework. Specifically, X_i denotes the position of the i^{th} spatial point, which will not change over time. H_i^n and U_i^n denote the height and velocity at the i^{th} spatial point and in step n . The sequential relations are

$$h_i^{n+1} = \mathcal{H}^{\mathcal{E}}(H_{i-1}^n, H_i^n, H_{i+1}^n, U_{i-1}^n, U_i^n, U_{i+1}^n), \quad (9)$$

$$u_i^{n+1} = \mathcal{U}^{\mathcal{E}}(H_{i-1}^n, H_i^n, H_{i+1}^n, U_{i-1}^n, U_i^n, U_{i+1}^n), \quad (10)$$

where variables h_i^{n+1} , and u_i^{n+1} are the state variables in step $n + 1$ at the i^{th} material point. To learn the functions $\mathcal{H}^{\mathcal{E}}$ and $\mathcal{U}^{\mathcal{E}}$, we also need data of h_i^{n+1} and u_i^{n+1} . Although h_i^{n+1} and u_i^{n+1} are

not directly collected, they can be calculated from the height and velocity at the spatial points in step $n + 1$ (i.e., H^{n+1} and U^{n+1}). First, position x_i^{n+1} is given by

$$x_i^{n+1} = X_i + U_i^n dt. \quad (11)$$

Then, height h_i^{n+1} and velocity u_i^{n+1} can be mapped by linear interpolation of $H^{n+1} - X$ and $U^{n+1} - X$ at x_i^{n+1} .

After the functions \mathcal{H}^ε and \mathcal{U}^ε are learned from collected data, column collapse can be modeled (Figure 2b). Given the height and velocity at spatial points in the current step n , the height h^{n+1} and velocity u^{n+1} of corresponding material points in the next step $n + 1$ can be obtained from \mathcal{H}^ε and \mathcal{U}^ε . By linear interpolation of $h^{n+1} - x^{n+1}$ and $u^{n+1} - x^{n+1}$ at X , the height and velocity at spatial points in the next step $n + 1$ are calculated. It should be noted that interpolation is required to obtain variables for spatial points from material points, while in the process of inferring the sequential relations, interpolation is required to obtain variables for material points from spatial points.

It should be emphasized that the aforementioned frameworks are only for constructing the sequential relations, such that no limitation is imposed on the underlying technique used to generate data. For example, if data is generated using a numerical method premised on a Lagrangian approach (e.g., large deformation finite element method), solutions can still be used by the data-driven method in the Eulerian framework.

REGRESSION VIA RADIAL BASIS FUNCTIONS

The functions $\mathcal{H}^\mathcal{L}$, $\mathcal{U}^\mathcal{L}$, \mathcal{H}^ε , and \mathcal{U}^ε are approximated with radial basis functions, which can be regarded as single-hidden-layer neural networks with radial basis functions as activation functions (Park and Sandberg, 1991).

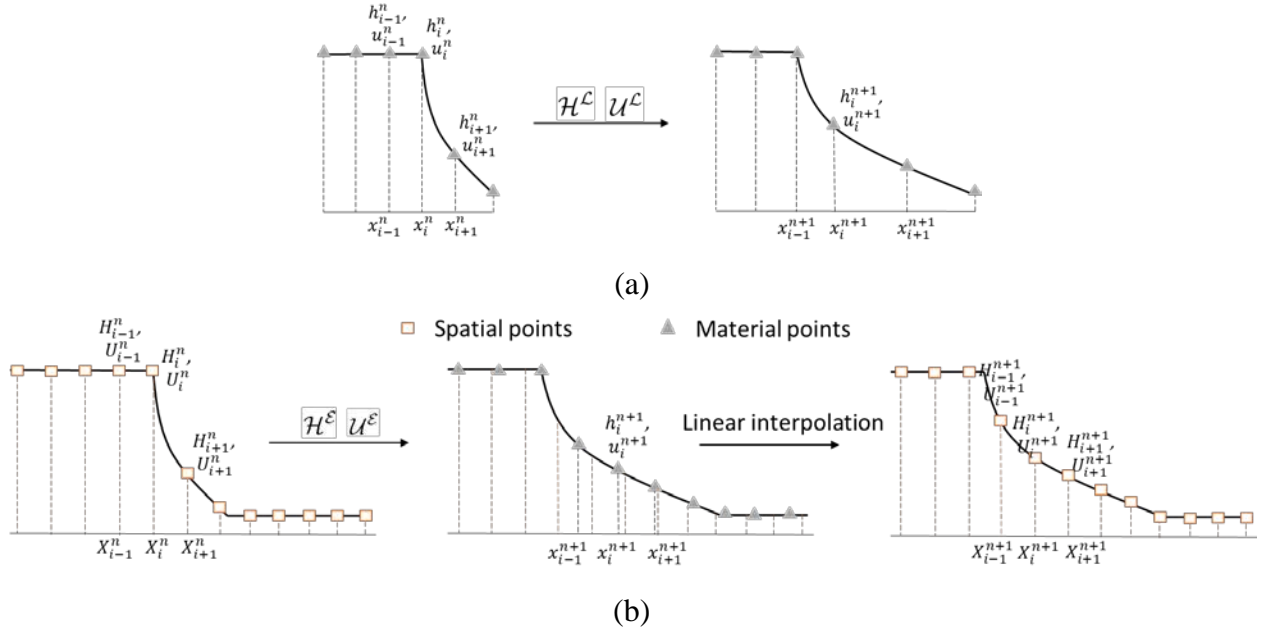


Figure 2: Process of predicting solutions step by step in (a) Lagrangian framework and (b) Eulerian framework.

Taking function $\mathcal{H}^{\mathcal{L}}$ as an example, to simplify the expression, let $h := h_i^{n+1}$, $u := u_i^{n+1}$, and use a vector \mathbf{a} to denote all variables

$$\mathbf{a} := [x_{i-1}^n, x_i^n, x_{i+1}^n, h_{i-1}^n, h_i^n, h_{i+1}^n, u_{i-1}^n, u_i^n, u_{i+1}^n].$$

After constructing a database with K data points, the function can be expressed as a linear combination of K radial basis functions φ :

$$\mathcal{H}(\mathbf{a}) = \sum_k^K w_k \varphi(\|\mathbf{a} - \mathbf{a}_k\|), \quad (12)$$

where the w_k denotes the weight of the radial basis function $\varphi(\|\mathbf{a} - \mathbf{a}_k\|)$. The ordinary radial basis functions includes Gaussian, multiquadric, inverse quadric, etc. Here, multiquadric radial basis functions are used:

$$\varphi(\|\mathbf{a} - \mathbf{a}_k\|) = \sqrt{1 + \epsilon(\|\mathbf{a} - \mathbf{a}_k\|)^2}, \quad (13)$$

where ϵ is a parameter to scale variables. Let

$$\mathbf{A} = \begin{bmatrix} \varphi(\|\mathbf{a}_1 - \mathbf{a}_1\|) & \varphi(\|\mathbf{a}_1 - \mathbf{a}_2\|) & \dots & \varphi(\|\mathbf{a}_1 - \mathbf{a}_M\|) \\ \varphi(\|\mathbf{a}_2 - \mathbf{a}_1\|) & \varphi(\|\mathbf{a}_2 - \mathbf{a}_2\|) & \dots & \varphi(\|\mathbf{a}_2 - \mathbf{a}_K\|) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi(\|\mathbf{a}_K - \mathbf{a}_1\|) & \varphi(\|\mathbf{a}_K - \mathbf{a}_2\|) & \dots & \varphi(\|\mathbf{a}_K - \mathbf{a}_K\|) \end{bmatrix}, \quad (14)$$

$$\mathbf{w} = [w_1, w_2, \dots, w_K]^T, \quad (15)$$

$$\mathbf{h} = [h_1, h_2, \dots, h_K]^T, \quad (16)$$

where $h_k (k = 1, 2, \dots, K)$ denote heights in the database corresponding to the vectors \mathbf{a}_k . Then, the weights w_k can be estimated by minimizing the mean square error

$$MSE = \frac{1}{K} \|\mathbf{h} - \mathbf{A}\mathbf{w}\|^2. \quad (17)$$

Consequently, for a new variable \mathbf{a} , a corresponding h can be calculated by Eq. (12).

VALIDATION OF THE IMPLEMENTATION

To prove that the data-driven method can reproduce numerical (true) solutions successfully,

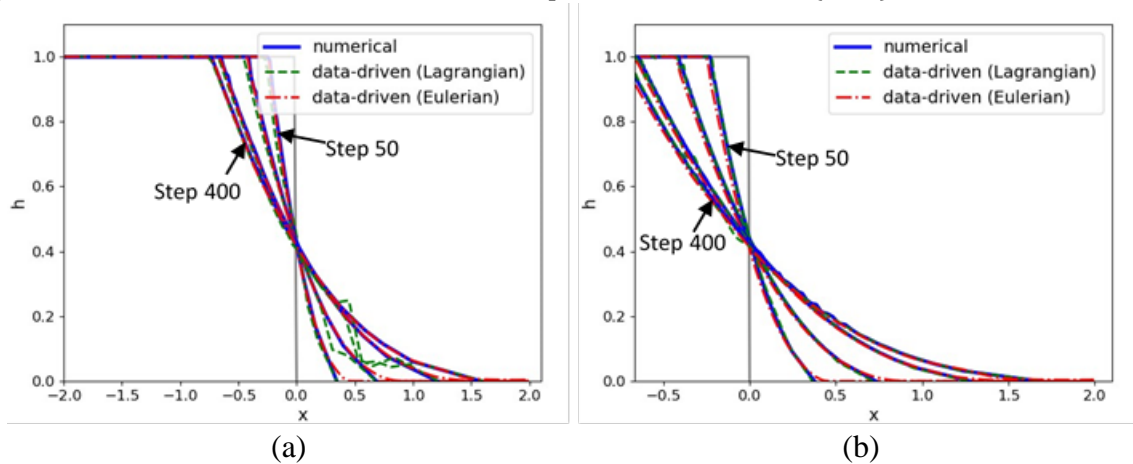


Figure 3: Data-driven results in step 50, 100, 200, and 400 for (a) $A = 0.5$ and (b) $A = 1.5$ using numerical solution obtained for the same configurations ($A = 0.5$ and 1.5).

data are collected from numerical solutions of granular columns with initial aspect ratios $A = 0.5$ and 1.5 . In both the Lagrangian and Eulerian frameworks, the initial spacing (Δx) is 0.01 (the spacing will change with time in the Lagrangian framework), and the incremental time (Δt) of collected data is 0.005 . Consequently, around 2.5×10^5 data are collected, and 1×10^4 data ($1/25$ of the size of data) are selected for inferring the functions \mathcal{H}^L , \mathcal{U}^L , \mathcal{H}^E , and \mathcal{U}^E , i.e., estimating the weights (w_k) for the radial basis functions. After the weights are estimated, the solutions for subsequent steps are inferred sequentially starting from the initial step. Figure 3 compares the data-driven solutions against the numerical (true) solution. Except for the discrepancies in the Lagrangian framework for the column with $A = 0.5$, the curves from data-driven methods are universally close to the numerical (true) solution.

RESULTS

Lagrangian vs Eulerian framework. The next set of data-driven simulations explores whether the method furnishes reasonable predictions when applied to cases with initial geometries that were not included in the collected data. The “training” data are still taken from granular columns with $A = 0.5$ and 1.5 , and the data-driven approach is then applied to granular columns with $A = 0.75, 1.0$, and 1.25 . Figure 4 shows the final profile of column collapse, including the true profile

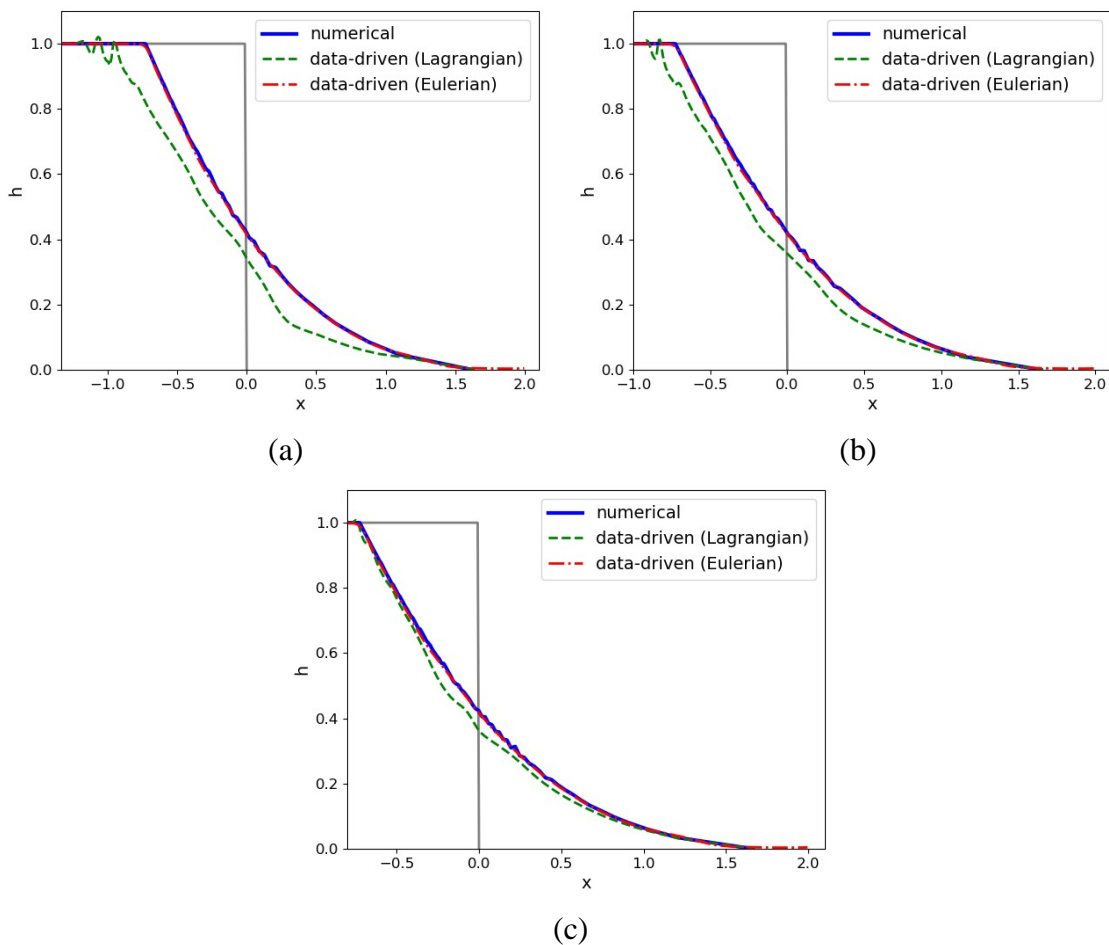


Figure 4: Data-driven results in final profile for (a) $A = 0.75$, (b) $A = 1.0$, and (c) $A = 1.25$ using numerical solution obtained for configurations with $A = 0.5$ and 1.5 .

obtained from direct numerical simulation. As seen, the Eulerian framework gives better predictions compared to the Lagrangian framework. The reason for this might be that the number of input variables (dimensions) in the Eulerian framework is smaller than the Lagrangian framework, so predictions obtained with the Eulerian framework are better when the size of data in these two frameworks are close.

Sensitivity analysis. The goal of this section is to analyze the influence of the spacing (Δx) and the time increment (Δt) chosen when collecting data. The following data-driven simulations are all conducted in the Eulerian framework. The influence of the spacing (Δx) is investigated at first. Taking a column with $A = 2$ as an example, the ratio of time increment to time required for the column to fully collapse (t^{end}) is set at $1/400$. Figure 5 shows the data-driven solutions for different spacings, where $RMSE$ is the root mean square error of predicted heights in the whole process

$$RMSE = \sqrt{\frac{\sum_{n=1}^{N^{points} \times N^{steps}} (h_{data-driven} - h_{numerical})^2}{N^{points} \times N^{steps}}}, \quad (18)$$

where N^{points} and N^{steps} are the number of points and the number of steps, respectively. It is shown that the spacing of collected data has an influence on the accuracy of the data-driven solutions. When $\Delta x/l^0$ is smaller than $1/50$, the predicted error is small ($RMSE/h^0$ is smaller than 0.01).

To obtain the root mean square errors for different time increments (Δt), the spacing (Δx) is set with $\Delta x/l^0 = 1/50$. Figure 6 shows the effect of varying the time increment used for collected data, which also has an influence on the accuracy of data-driven solutions. When $\Delta t/t^{end}$ is smaller than $1/250$, the predicted error is small ($RMSE/h^0$ is smaller than 0.01).

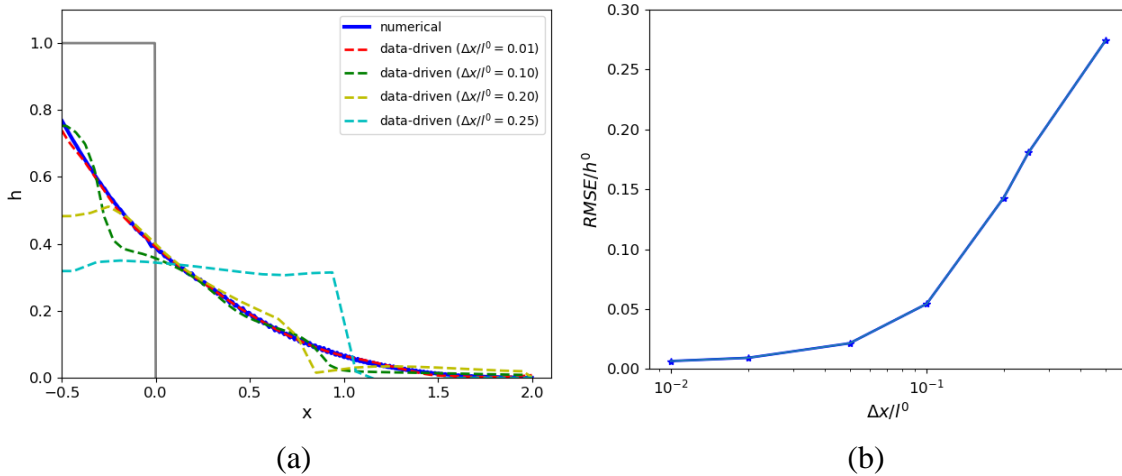


Figure 5: Data-driven solutions of a granular column ($A = 2$) when data are collected under different spacing: (a) final profile of column collapse; (b) root mean square error ($RMSE$) of predicted heights in the whole process (h^0 and l^0 are the initial height and the initial width of the column).

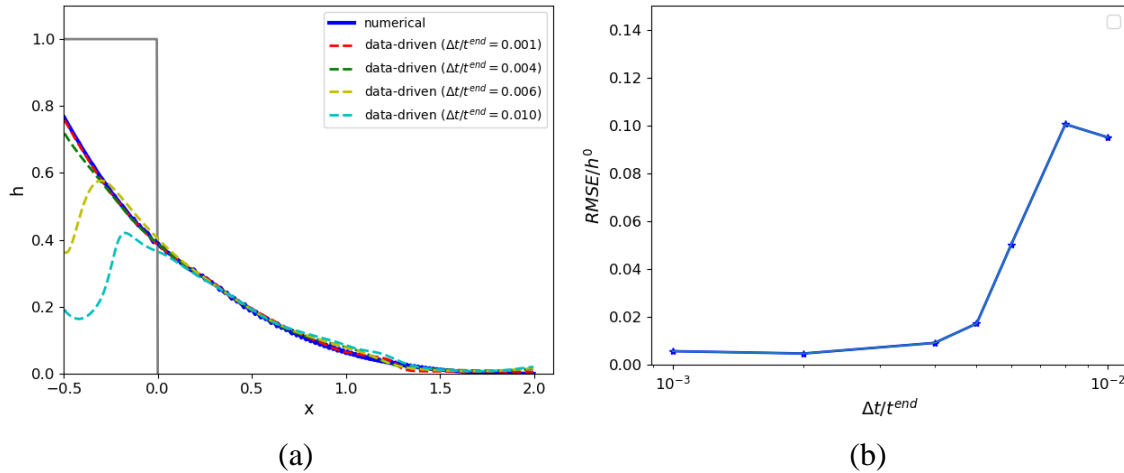


Figure 6: Data-driven solutions of a granular column ($A = 2$) when data are collected under different time increments: (a) final profile of column collapse; (b) root mean square error ($RMSE$) of predicted heights for the whole process.

CONCLUDING REMARKS

In this paper, a data-driven method is designed for predicting granular flow based on two possible frameworks (Lagrangian and Eulerian). The method is applied to predict 2D column collapse using selected data from existing (i.e., precomputed) solutions for particular configurations, and reasonable accuracy is demonstrated, particularly for the Eulerian framework. While these preliminary results are encouraging, problems with different material properties and more complex geometries need to be explored in order to ascertain whether data-driven methods can replace numerical methods for granular flow problems.

This paper focuses on inferring theoretical solutions from existing numerical data, but in fact considerable potential for data-driven methods lies in their ability to give predictions directly from whatever data might be on hand, including theoretical predictions ascertained from numerical simulations or real-world data collected from experiments or field tests. Although all data in this work came from numerical solutions, the proposed approach provides a first step toward generalized data-driven approaches that might incorporate data from other sources, and the sensitivity analysis assessing the influence of grid spacing and time increment for collected data provides guidance on how best to sample such data.

REFERENCES

- J. L. Baker, T. Barker, and J. M. N. T. Gray. A two-dimensional depth-averaged μ -rheology for dense granular avalanches. *Journal of Fluid Mechanics*, 787:367–395, Jan. 2016. Publisher: Cambridge University Press.
- H. H. Bui, R. Fukagawa, K. Sako, and S. Ohno. Lagrangian meshfree particles method (SPH) for large deformation and failure flows of geomaterial using elastic–plastic soil constitutive model. *International Journal for Numerical and Analytical Methods in Geomechanics*, 32 (12):1537–1570, 2008.

- E. J. Fern and K. Soga. The role of constitutive models in MPM simulations of granular column collapses. *Acta Geotechnica*, 11(3):659–678, June 2016.
- Y. Forterre and O. Pouliquen. Flows of Dense Granular Media. *Annual Review of Fluid Mechanics*, 40(1):1–24, 2008.
- J. Hesthaven and S. Ubbiali. Non-intrusive reduced order modeling of nonlinear problems using neural networks. *Journal of Computational Physics*, 363:55–78, June 2018.
- O. Hungr and S. McDougall. Two numerical models for landslide dynamic analysis. *Computers & Geosciences*, 35(5):978–992, May 2009.
- R. R. Kerswell. Dam break with Coulomb friction: A model for granular slumping? *Physics of Fluids*, 17(5):057101, Apr. 2005. Publisher: American Institute of Physics.
- T. Kirchdoerfer and M. Ortiz. Data-driven computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 304:81–101, June 2016.
- C. Y. Kuo, Y. C. Tai, F. Bouchut, A. Mangeney, M. Pelanti, R. F. Chen, and K. J. Chang. Simulation of Tsaoling landslide, Taiwan, based on Saint Venant equations over general topography. *Engineering Geology*, 104(3):181–189, Mar. 2009.
- I. Luca, K. Hutter, Y. C. Tai, and C. Y. Kuo. A hierarchy of avalanche models on arbitrary topography. *Acta Mechanica*, 205(1):121–149, June 2009.
- R. Lueptow, A. Akonur, and T. Shinbrot. Piv for granular flows. *Experiments in Fluids*, 28(2):183–186, 2000.
- A. Mangeney-Castelnau, J.-P. Vilotte, M. O. Bristeau, B. Perthame, F. Bouchut, C. Simeoni, and S. Yerneni. Numerical modeling of avalanches based on Saint Venant equations using a kinetic scheme. *Journal of Geophysical Research: Solid Earth*, 108(B11), 2003.
- U. Niethammer, M. R. James, S. Rothmund, J. Travelletti, and M. Joswig. UAV-based remote sensing of the Super-Sauze landslide: Evaluation and results. *Engineering Geology*, 128:2–11, Mar. 2012.
- C. Ouyang, H. An, S. Zhou, Z. Wang, P. Su, D. Wang, D. Cheng, and J. She. Insights from the failure and dynamic characteristics of two sequential landslides at Baige village along the Jinsha River, China. *Landslides*, 16(7):1397–1414, July 2019.
- J. Park and I. W. Sandberg. Universal approximation using radial-basis-function networks. *Neural computation*, 3(2):246–257, 1991.
- O. Pouliquen. Scaling laws in granular flows down rough inclined planes. *Physics of Fluids*, 11(3):542–548, Feb. 1999. Publisher: American Institute of Physics.
- M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, Feb. 2019.
- S. B. Savage and K. Hutter. The motion of a finite mass of granular material down a rough incline. *Journal of Fluid Mechanics*, 199:177–215, Feb. 1989. Publisher: Cambridge University Press.
- S. Utili, T. Zhao, and G. T. Houlsby. 3D DEM investigation of granular column collapse: Evaluation of debris motion and its destructive power. *Engineering Geology*, 186:3–16, Feb. 2015.
- K. Wang and W. Sun. A multiscale multi-permeability poroplasticity model linked by recursive homogenizations and deep learning. *Computer Methods in Applied Mechanics and Engineering*, 334:337–380, 2018.