Estimate of sediment settling velocities from a theoretically guided data-driven approach

Cao Zhendong[1], Wolfram Phillip[1], Rowland Joel[1], Zhang Yu[2], and Pasqualini Donatella[1]

[1]Los Alamos National Laboratory
[2]Los Alamos National laboratory

November 16, 2022

Abstract

Sediment settling velocities are commonly estimated from process-based or parametric data-driven approaches. The process-based approach has theoretical constraints due to the unclear settling physics; the parametric data-driven approach is limited by its mathematical assumptions. To overcome these limitations, this study compiles an aggregated sediment settling experimental database from literature and develops a non-parametric data-driven model to estimate the non-cohesive sediment settling velocity in water. A cross-comparison against five process-based equations and a parametric data-driven equation demonstrates the higher accuracy and better consistency of the new model in estimating sediment settling velocities under various physical regimes. The data-driven model also shows an easy-implemented self-update capability by assimilating theoretical data generated from the process-based equations. The updated model, leveraging experimental and theoretical data of sediment settling process, further improves the accuracy and reduces the uncertainty in estimating sediment settling velocities. This approach illustrates the value of integrating experimental and theoretical knowledge in estimating the complex process in sediment transport, and provides an alternative framework for future sediment transport exploration.

Estimate of sediment settling velocities from a theoretically guided data-driven approach

Zhendong Cao (caozd999@lanl.gov), Phillip J. Wolfram (pwolfram@lanl.gov),
Joel Rowland(jrowland@lanl.gov), Yu Zhang (yuzhang@lanl.gov) and Donatella Pasqualini (dmp@lanl.gov)

Sediment Settling Velocity Affects

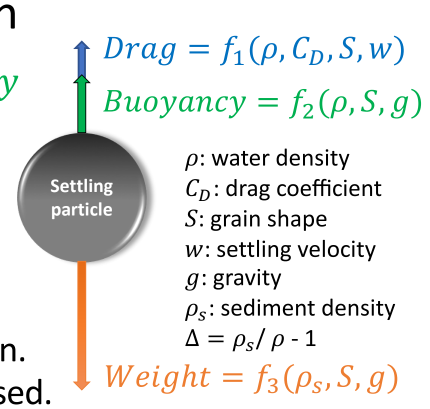
- sediment transport process
- reservoir and harbor designs
- wetland reclamation and restoration
- waterways navigation
- more …

Previous Studies

- Process-based approach

$Weight = Drag + Buoyancy$

$w = f(S, C_D, \rho_s, \rho, g)$



- C_D is a function of particle Reynolds number (Re).
- Laboratory experiments were conducted for $C_D - Re$ relation.
- Many models have been proposed.
- ***Comparisons of different models within integrated data are rare.***
- ***Uncertainty remains in real applications***.

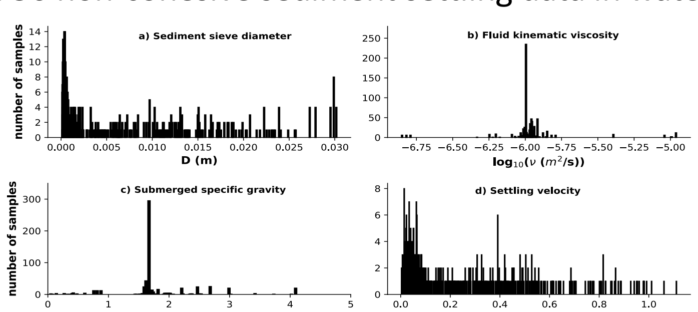
- Parametric data-driven approach

Goldstein and Coco (2014):

- Compiled a multi-source database from literature $w,\ D\ (grain\ size), \nu\ (water\ kinematic\ viscosity)\ and\ \Delta$
- Trained $w = f(D, \nu, \Delta)$ in a genetic programming software
- Picked the best fit from all candidate solutions
- ***Need pre-specify mathematical operators.***

This Study

- Aggregated database
  - 13 distinct experimental data sources
  - 756 non-cohesive sediment settling data in water



- Random forest regression model

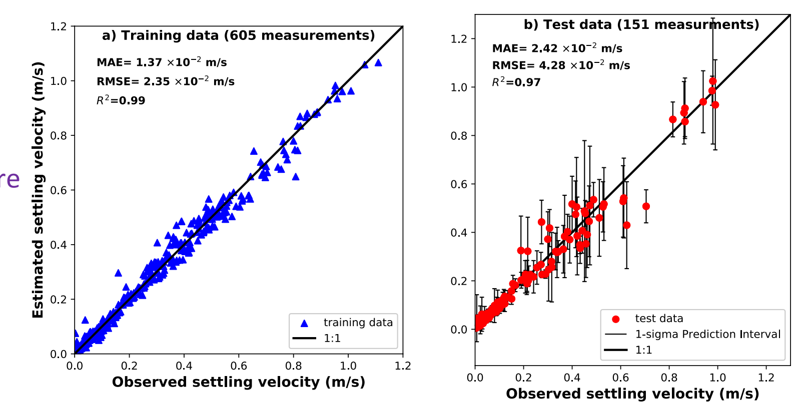
Model Development

- Parameter tuning and model selection
  - test data size, number of trees, maximum depth
  - grid search with 5-fold cross validation

Scenario	Tuning parameters			Training data			Test data		
	$test\_size$	$ntrees$	max_depth	MAE^*	$RMSE^*$	R^2	MAE^*	$RMSE^*$	R^2
1	0.1	100	8	1.57	2.67	0.99	2.54	4.81	0.97
2	**0.2**	**150**	**9**	**1.37**	**2.35**	**0.99**	**2.42**	**4.28**	**0.97**
3	0.3	100	9	1.35	2.30	0.99	2.44	4.38	0.97
4	0.4	350	11	1.17	2.13	0.99	2.60	4.64	0.96
5	0.5	400	13	1.16	2.15	0.99	2.55	4.65	0.96
6	0.6	400	5	2.05	3.31	0.98	3.06	5.55	0.95
7	0.7	150	6	1.62	2.34	0.99	3.30	6.14	0.93
8	0.8	100	5	1.66	2.50	0.99	3.68	6.66	0.92
9	0.9	150	9	1.31	2.29	0.99	3.77	6.78	0.92

* 10^{-2} m/s

- Scenario 2
  - ❖High accuracy
  - ❖No overfitting
  - ❖Simple structure



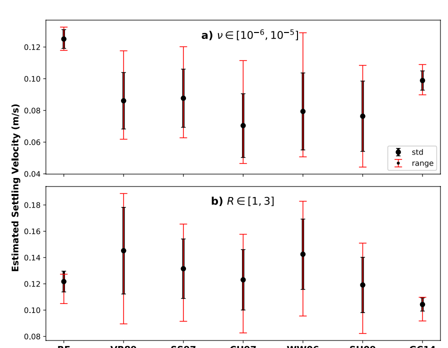
Model Comparisons

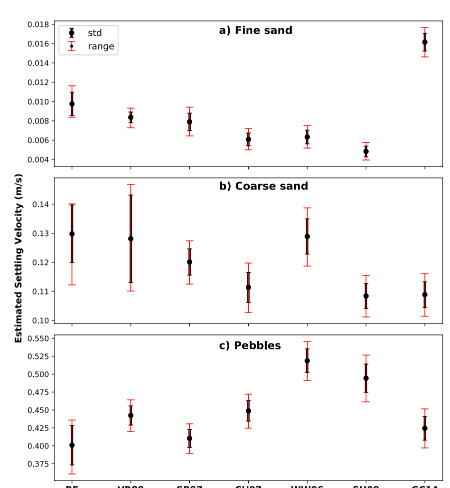
- Accuracy on test data

Model Name	sand			gravel			sand & gravel		
	MAE^*	$RMSE^*$	R^2	MAE^*	$RMSE^*$	R^2	MAE^*	$RMSE^*$	R^2
Van Rijn 1989 (VR89)	1.18	1.62	0.92	7.65	10.60	0.78	3.44	6.39	0.93
Soulsby 1997 (SB97)	1.36	1.93	0.89	8.17	11.40	0.75	3.73	6.90	0.91
Cheng 1997 (CH97)	1.12	1.76	0.91	7.95	10.62	0.78	3.51	6.43	0.92
Wu and Wang 2005 (WW05)	**0.96**	**1.35**	**0.94**	10.34	13.34	0.65	4.23	7.96	0.88
Sadat-Helbar 2009 (SH09)	1.40	2.10	0.86	10.47	13.51	0.64	4.56	8.16	0.88
Goldstein and Coco 2014 (GC14)	1.56	2.23	0.85	**6.67**	**8.60**	**0.86**	**3.34**	**5.39**	**0.95**
This study (RF)	**0.87**	**1.21**	**0.96**	**5.31**	**7.07**	**0.92**	**2.42**	**4.28**	**0.97**

* 10^{-2} m

- Sensitivity tests

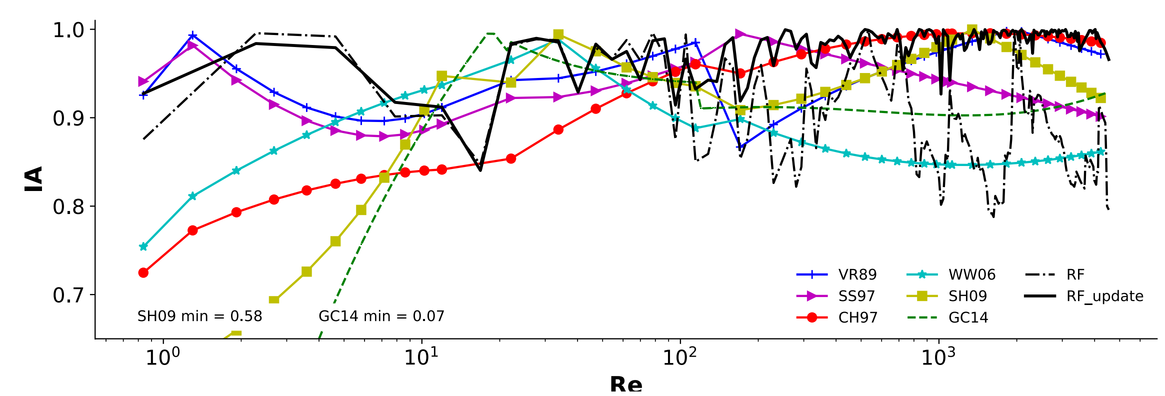




Model Update

- Theoretical data
  - Select 1000 sediment particles with fixed density (2650 kg/m^3) and varying grain sizes evenly distributed between [10^{-4} 10^{-2}] m.
  - Estimate the particle settling velocities in water using the process-based equations (VR89, SB97, CH97, WW05 and SH09).
  - Average the estimate values of each particle as theoretical data.
- Integration of theoretical data
  - Randomly pick 10% of the theoretical data to retrain RF and get RF_update
  - Evaluate model performance on the rest theoretical data via an individual accuracy metric: $IA_{i,j} = 1 - \frac{|w_{new}^i - w_j^i|}{w_{new}^i}$,

    where w_{new}^i is the theoretical value of the ith particle, w_j^i is the estimated value of the ith particle by the jth model.



Conclusions

- Non-parametric data-driven approach reduces the uncertainty in estimating sediment settling velocities under various physical regimes.
- The data-driven model has an easy-implemented update capability by integrating theoretical data.
- Leveraging experimental and theoretical knowledge to enhance estimate performance provides an alternative framework for sediment transport study.

Acknowledgements

Thanks to Drs. Goldstein and Coco for sharing data and for helpful discussions motivating this study. Funding for this study is provided under the Los Alamos National Laboratory Research and Development Directed Research project “Adaption Science for Complex Natural-Engineered Systems” (20180033DR). This publication has been supported by the Los Alamos Laboratory Directed Research and Development project under *LA-UR-19-25788*.