EPSC Abstracts
Vol. 12, EPSC2018-121-2, 2018
European Planetary Science Congress 2018

Application of Machine Learning to Giant Impact Studies

Saverio Cambioni[1]**,** Travis S.J. Gabriel[2], Erik Asphaug[1], Roberto Furfaro[3], Alexandre Emsenhuber[1] and Stephen Schwartz[1]
[1]Lunar and Planetary Laboratory, The University of Arizona, Tucson AZ, USA (cambioni@lpl.arizona.edu)
[2]School of Earth and Space Exploration, Arizona State University, Tempe, AZ, USA,
[3]System and Industrial Engineering, The University of Arizona, Tucson, AZ, USA

Abstract

Giant impacts heavily influenced the final configuration and geochemistry of the terrestrial planets. In this study we use machine learning to explore a rich dataset of giant impact simulations in a supervised fashion. This new methodology produces mappings of giant impact outcomes in an *N*-Dimensional (*N*-D) parameter space, e.g., mass of target, target-projectile mass ratio, impact velocity and impact angle. We discuss the physical insights emerging from this initial analysis and future work.

1. Introduction

Bombardment by large projectiles played a key role in the formation of the inner planets through accumulation of rocky planetesimals. Our group uses Smoothed-Particle Hydrodynamics (SPH) to model giant impacts on planetary bodies such as the Moon, Mercury and Mars (e.g. [3,5,6]). As of today, our dataset is composed of over 1,500 simulations spanning a wide range of parameters (composition, size, mass ratio, impact angle, impact velocity). Each SPH outcome is a complex *N*-D state (consolidated planets, clumps, unconsolidated ejecta, and their characteristics including thermodynamic states, etc.) that requires detailed analysis.

In this regard, state-of-the-art machine learning techniques allow for several advantages: 1) they can streamline the generation of data sets to most efficiently explore regions of interest in a large parameter space; and 2) they can perform accurate mappings of initial conditions and end-states, with associated probabilities, taking into account a high-dimensional parameter space. This is in contrast to human operators that are often limited to a mostly 2-D understanding of the data. Modern machine learning schemes take advantage of this big data problem to spot new and sometimes unexpected correlations.

In this pilot study, we trained, tested and validated an algorithm able to map strictly the pre-impact conditions to the accretion efficiency (defined as the fraction of the projectile mass $M_{p.}$ acquired by the target M_t). The adopted scheme is supervised: the machine learns the correlation between input and output using labelled data: $\{\mathbf{x_i}; y_i\}$, where $\mathbf{x_i}$ is an array of four input parameters (predictors): M_t (mass of the target), $M_{p.}$ (mass of the impactor), θ (impact angle) and the ratio between the impact velocity and the escape velocity; and y_i is the corresponding response. Efficiency, accuracy, and predictive power will be discussed.

2. Classification model

In SPH, the continuous fluid is represented as a Lagrangian set of particles that move with the flow; this allows easy visualization and supports analytical deductions ([4] and references therein). We designed a naive Bayesian classifier [7] based on a visual classification of simulation outcomes. Fig. 1 shows an example of *a posteriori* classification probability for 4 distinct impact outcomes defined and discussed in [1,4,8]: Hit & Run, Graze & Merge, Merge, Disruption. Each of these collisional regimes can be thought of as a 'phase' in which collisional behaviour is similar, that is, a scaling law can be expected to apply. The topology reveals relatively narrow transition regions between the different collisional regimes, in which the collisional outcome associated with a combination of parameters is probabilistic. The classifier can be adopted as a guideline to distinguish "interesting" regions of phase-space to be followed up with further simulations.

3. Surrogate model

While a classifier is able to handle discrete, qualitative responses, a surrogate model is an algorithm able to mimic the SPH physics and to predict continuous (floating point) outputs given the input parameters (predictors). Running the surrogate model drastically reduces the computational time with respect to full

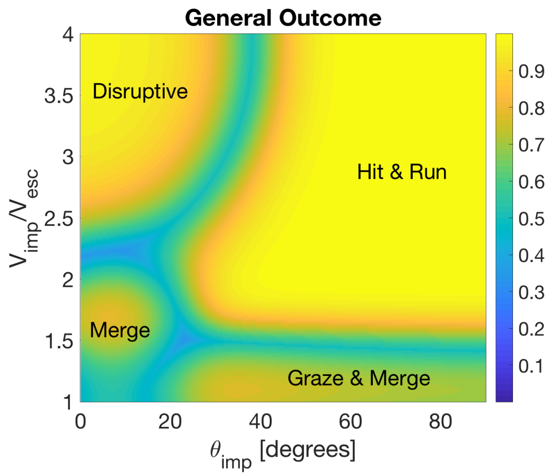

Figure 2. Example of a posteriori probability distribution for the 4 outcome categories ($M_t = 0.1M_{Earth}$ and $\gamma = M_p/M_t = 0.7$).

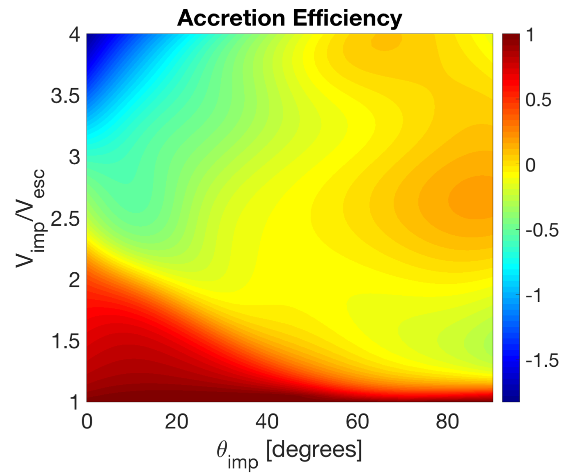

Figure 2. Example of prediction for accretion efficiency using the designed surrogate model ($Mt = 0.1M_{Earth}$ and $\gamma = M_p/M_t = 0.7$).

SPH simulations (from hours to seconds). The surrogate model forecasts accretion efficiency at several times the gravitational timescale, after pressure gradient forces are no longer acting. The accretion efficiency map, Fig. 2, is richer in terms of collisional outcomes with respect to the Bayesian classifier. We note that a significant fraction of the projectile mass M_p can be acquired by the target M_t in 'Hit & Run' collisions too, without a complete merging of the projectile with the target. Conversely, 'Graze & Merge' collisions at low velocity and high impact angle can somewhat be inefficient, implying that a significant fraction of mass ends up forming unbounded remnants.

4. Conclusions and future work

This study shows the potentialities of machine learning in: 1) processing large planetary formation datasets without a heavy involvement/bias from the user; 2) identifying "strange" or interesting regions in the parameter space requiring further study; 3) providing a statistical description to phase boundaries as opposed to the hardline boundaries of traditional scaling laws. One long-term scientific goal is to obtain systematic guidance to solutions of complex problems such as Earth-Moon system formation and Mercury formation [6, 3] that may exist on phase boundaries in outcomes of giant impacts.

In future work we will create a complete surrogate SPH model that can enable additional outputs such as masses, velocities and spin states of the largest remnants. We will also apply this approach to the transition from gravity- to friction-dominated collisions. The surrogate model could be subsequently used as input in *N*-body simulations for post-collisional dynamical studies. Being an invertible function, we can apply the surrogate model to gain a better understanding of the relative likelihoods of any specific giant impact scenario.

References

[1] Asphaug, E., et al.: Hit-and-run planetary collisions, Nature, Vol. 439.7073, pp. 155, 2006.

[2] Asphaug, E.: Impact origin of the Moon? Annual Review of Earth and Planetary Sciences, Vol. 42, pp. 551-578, 2014.

[3] Asphaug, E., and Reufer, A.: Mercury and other iron-rich planetary bodies as relics of inefficient accretion, Nature Geoscience, Vol. 7.8, pp. 564, 2014.

[4] Asphaug, E., et al.: Global scale impacts, Asteroid IV, University of Arizona Press, 2015.

[5] Emsenhuber, A., et al.: SPH calculations of Mars-scale collisions: The role of the equation of state, material rheologies, and numerical effects, Icarus, Vol. 301, pp. 247-257, 2018.

[6] Reufer, Andreas, et al.: A hit-and-run giant impact scenario, Icarus, Vol. 221.1, pp. 296-299, 2012.

[7] Rish, I.: An empirical study of the naive Bayes classifier, IJCAI 2001 workshop, IBM, Vol. 3, n. 22, pp. 41-46, 2001.

[8] Stewart, T., and Leinhardt, Z.: Velocity-dependent catastrophic disruption criteria for planetesimals, The Astrophysical Journal Letters, Vol. 691.2, pp. L133, 2009.