International Society of
the Learning Sciences

A2S: Designing an integrated platform for computational modeling & data analysis for sustained investigations in science classrooms

Adelmo Eloy, Columbia University, adelmo@fablearn.net
Jacob Wolf, Columbia University, jhw2174@tc.columbia.edu
Aditi Wagh, Massachusetts Institute of Technology, awagh@mit.edu
Tamar Fuhrmann Columbia University, tf2464@tc.columbia.edu
Engin Bumbacher, Teacher University Lausanne, engin.bumbacher@hepl.ch
Michelle Wilkerson, Berkeley University, mwilkers@berkeley.edu
Paulo Blikstein, Columbia University, paulob@tc.columbia.edu

Abstract: In this paper we present A2S, an integrated environment that combines building computational models using domain-specific primitives and comparing models with analysis of real-world data for sustained investigations. Targeted towards middle school students and teachers, the design of the environment is guided by three goals: 1) providing accessible entry points for rapid construction of computational models to build scientific explanations; 2) enabling modification of domain-specific primitives to pursue extended investigations through the "unpacking" of blocks; and 3) validating built models through comparison with real-world data from different sources.

Introduction

In this paper, we describe From Access to Sustainability (A2S), an integrated domain-specific computational modeling environment that enables building of agent-based models and comparing them with data of scientific phenomena. In A2S, learners are able to create computational models and make side-by-side comparisons with data of a specific phenomenon. Early versions of A2S have been refined through co-design sessions with eight middle-school science teachers and classroom implementations with 224 students. Below we briefly describe our theoretical framework, an overview of our current design, and a proposed demo session for participants to engage with the environment and provide feedback.

Theoretical framework

Modeling is a key scientific practice that supports learners to represent ideas and explanations for scientific phenomena (NGSS Lead States, 2013). Computational modeling leverages the power and flexibility of computational media to enable students to create, debug, and interact with dynamic models. A2S builds on existing literature in domain-specific modeling that has shown that computational modeling can be made accessible even to novice programmers with domain-specific primitives. By "domain-specific modeling," we mean primitives or building blocks for a model being contextualized to a specific phenomenon or domain and easily understood by learners (Aslan et al., 2020; Kahn, 2007; Martin et al., 2020; Repenning & Ambach, 1997; Wilkerson et al., 2015). While domain-specific environments have significantly reduced barriers for students to engage in computational modeling, they have limited computational expressivity (Reppening & Ambach, 1997). To address this issue, we aim to provide supports for students to "unpack" and modify blocks, so that they can engage in more sophisticated, deeper modeling endeavors.

Finally, we aim to leverage an existing common practice in middle school science classrooms - interacting with experimental data - by juxtaposing data and modeling for students to compare and validate models (Blikstein, 2012; Fuhrmann et al., 2018; Gouvea & Wagh, 2018). Our goal is for this juxtaposition to highlight differences between the model and data ("Bifocal Modeling"). We focus our design on three elements: 1) providing accessible entry points for rapid construction of computational models to build explanations about a real-world phenomenon through domain-specific modeling; 2) enabling modification of domain-specific primitives to pursue extended investigations through "unpacking" of blocks; and 3) validating built models through side-by-side comparison with real-world data from different sources and formats.

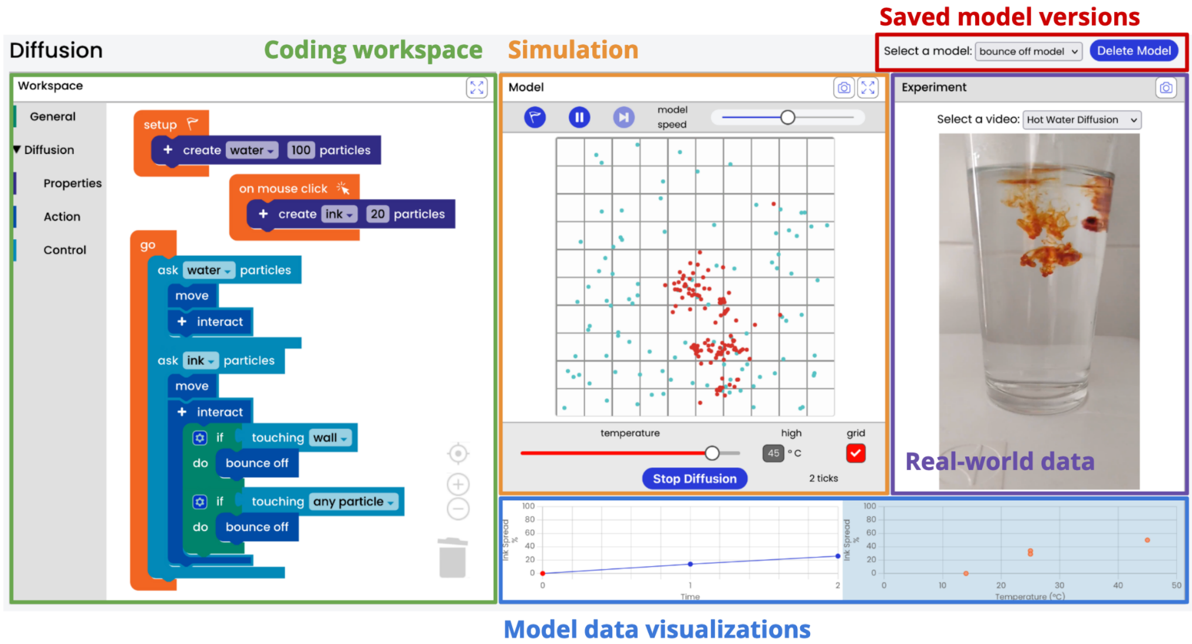
Design

The A2S environment aims to provide a platform for learners to program and validate computational models by comparing them to real-world data. The platform contains five components (Figure 1):

(1) *Code workspace* - This is the area where learners program their models. Learners can code using a custom block-based coding tool built on Google's Blockly library.

(2) *Simulation* - This area allows students to run and manipulate the model they built in the code workspace. Models are generated using a NetLogo engine (Wilensky, 1999).
(3) *Real-world data* - This component contains real-world data in the form of videos, images, tabular data, or data plots from a variety of real-world sources including learners' own experiments.
(4) *Model data visualizations* - This area visualizes select data generated by the model in real time. Data is visualized for each run of the model and also aggregated over multiple model runs.
(5) *Saved model versions* - This feature allows learners to save versions of their models. As learners iteratively develop their models, they can maintain a record of how their ideas changed over time.

Figure 1
A2S modeling environment

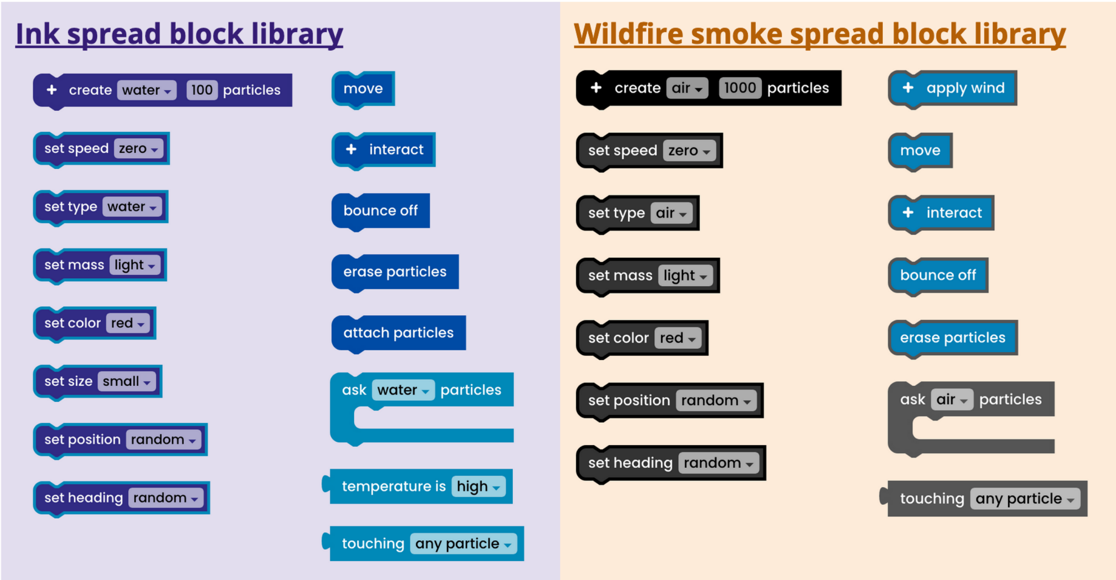


Our goal is to have four curricular units which allow for a sustained investigation into modeling. To date, we have designed two curricular units for middle school science classes. Using a spiral curriculum approach, the two units focus on a common mechanism, diffusion, in two different phenomena. In one unit, learners explore diffusion through the phenomenon of ink spreading through water. In another, they continue exploring diffusion through the phenomenon of wildfire smoke spread and are challenged to consider the effects of forces like wind and gravity/buoyancy on particle diffusion.

Sustained investigations for modeling

Each unit comes with a library of blocks tailored to the specific content explored in that unit. For example, in the wildfire smoke spread unit, an "apply wind" block adds the force of wind to the motion of particles. Given the spiraling curriculum approach we use, we are designing block libraries that maintain a consistent design paradigm across libraries. For example, since the ink spread and the wildfire smoke spread units explore particle motion, *both* block libraries include a "create particles" block which performs a similar function (Figure 2). By developing a consistent paradigm, we intend to make modeling accessible by allowing learners to transfer their understanding of conventions across increasingly complex modeling challenges in the units.

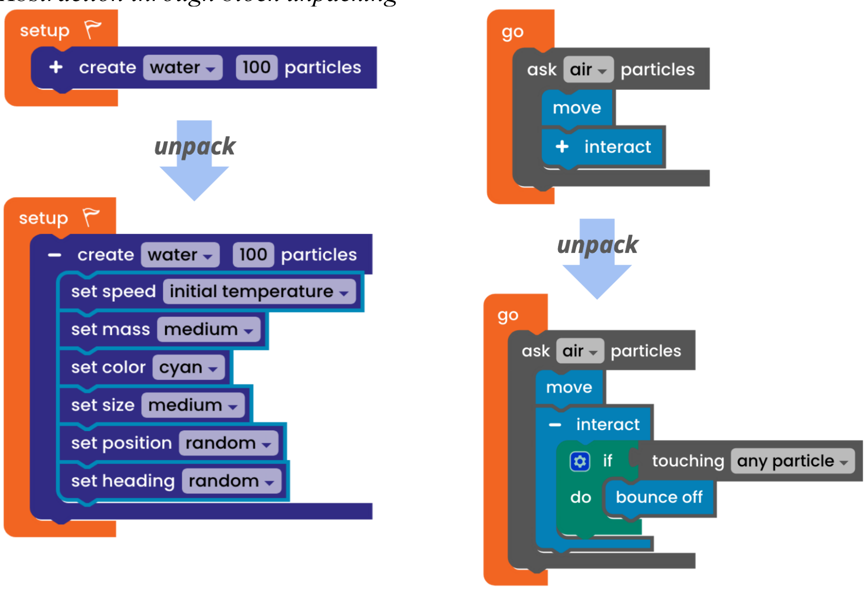
Figure 2
Block libraries from the ink spread and wildfire smoke spread units



"Unpacking" domain-specific blocks

To deepen the levels to which learners can explore a model, we designed a functionality to enable the modification of domain-specific primitives called "*unpacking*." This feature enables learners to "open" an abstracted block to view and modify its code. Unpacking allows students to utilize the high-level functionality of a block and then toggle between layers of abstraction, revealing deeper functionality that is closer to the foundational NetLogo code at each step. For example, learners can unpack "create particles" to see and manipulate the properties which get defined for each particle when it is created. The "interact" block is unpacked to reveal rules for particle collisions (Figure 3). By allowing learners to utilize and manipulate blocks at different levels of abstraction, we seek to draw on the intuitive, natural-language-based semantics of domain-specific primitives without obscuring their lower-level operation and compromising computational expressivity.

Figure 3
Abstraction through block unpacking



Integrating modeling with real-world data analysis

A2S integrates real-world data with the modeling environment side-by-side on the platform, creating an affordance for easily moving between building, running, and validating a model. The platform accommodates a variety of data types including tabular, image, and video data. This data is from multiple sources including experiments conducted by the learner, satellite imagery, online datasets, and physical sensors connected to a learner's computer. While some data will be provided with each unit, learners can also upload their data. Integrating data and modeling in this way can help to highlight the similarities and discrepancies between real-world data and data generated by the model and helps learners develop a better understanding of scientific modeling. Further, by allowing learners to integrate their own data into the process of validation, learners gain agency to explore and resolve the discrepancies which are most pressing to them, supporting them as they make sense of a phenomenon.

Demo session

We propose an interactive demo session where participants explore how A2S can support the understanding of a scientific phenomenon by integrating modeling and data practices. In this session, participants would create, interact, and assess different computational models of a scientific phenomenon and compare models with real-world data. The demo would include three moments: 1) an introduction to the environment; 2) a design challenge, inspired by one of the A2S curricular units; 3) a group reflection on the potential of domain-specific, block-based modeling and the next steps for the design of the A2S environment. The session can be run in an in-person or remote setting. Since A2S is a web-based software, we recommend that participants have an internet-connected computer. Additionally, participants should be able to project or share their screen with an audience.

References

Aslan, U., LaGrassa, N., Horn, M., & Wilensky, U. (2020, June). Phenomenological Programming: a novel approach to designing domain specific programming environments for science learning. In *Proceedings of the Interaction Design and Children Conference* (pp. 299-310).

Blikstein, P. (2012). Bifocal modeling: a study on the learning outcomes of comparing physical and computational models linked in real time. In *Proceedings of the 14th ACM international conference on Multimodal interaction* (pp. 257-264).

Fuhrmann, T., Schneider, B., & Blikstein, P. (2018). Should students design or interact with models? Using the Bifocal Modelling Framework to investigate model construction in high school science. *International Journal of Science Education*, 40(8), 867-893.

Gouvea, J. S. & Wagh, A. (2018). Exploring the Unknown: Supporting Students' Navigation of Scientific Uncertainty With Coupled Methodologies. In Kay, J. and Luckin, R. (Eds.) *Rethinking Learning in the Digital Age: Making the Learning Sciences Count, 13th International Conference of the Learning Sciences (ICLS) 2018*, Volume 1. London, UK: International Society of the Learning Sciences.

Kahn, K. (2007). Building computer models from small pieces. Paper presented at the *Proceedings of the 2007 Summer Computer Simulation Conference*, San Diego, CA.

Martin, K., Bain, C., Swanson, H., Horn, M., & Wilensky, U. (2020). Building Blocks: Kids Designing Scientific, Domain-Specific, Block-Based, Agent-Based Microworlds. In Gresalfi, M. and Horn, I. S. (Eds.), *The Interdisciplinarity of the Learning Sciences, 14th International Conference of the Learning Sciences (ICLS) 2020, Volume 3* (pp. 1641-1644). Nashville, Tennessee: International Society of the Learning Sciences.

NGSS Lead States. (2013). *Next Generation Science Standards: For states, by states*. Washington, DC: The National Academies Press.

Repenning, A., & Ambach, J. (1997). The agentsheets behavior exchange: Supporting social behavior processing. In *CHI'97 Extended Abstracts on Human Factors in Computing System*s (pp. 26-27).

Wilensky, U. (1999). NetLogo. Evanston, IL: Center for Connected Learning and Computer-Based Modeling, Northwestern University. http://ccl.northwestern.edu/netlogo.

Wilkerson, M. H., Wagh, A. & Wilensky, U. (2015). Balancing curricular and pedagogical needs in computational construction kits: Lessons from the DeltaTick project. *Science Education, 99*(3), 465-499.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. #2010413. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

