Scientific Modeling: Using learning analytics to examine student practices and classroom variation

David Quigley
University of Colorado Boulder
Institute for Cognitive Science
Department of Computer
Science
1777 Exposition Drive
Boulder, Colorado
david.quigley@colorado.edu

Jonathan Ostwald
University Corporation for
Atmospheric Research
Digital Learning Sciences
3090 Center Green Drive
Boulder, Colorado
ostwald@ucar.edu

Tamara Sumner
University of Colorado Boulder
Institute for Cognitive Science
1777 Exposition Drive
Boulder, Colorado
tamara.sumner@colorado.edu

ABSTRACT

Modeling has a strong focus in current science learning frameworks as a critical skill for students to learn. However, understanding students' scientific models and their modeling practices at scale is a difficult task that has not been taken up by the research literature. The complex variables involved in classroom learning, such as teacher differences, increase the difficulty of understanding this problem. This work begins with an exploration of the methods used to explore students' scientific modeling in the learning sciences space and the frameworks developed to characterize student modeling practices. Learning analytics can be used to leverage these frameworks of scientific modeling practices to explore questions around students' scientific models and their modeling practices. These analyses are focused around the use of EcoSurvey, a collaborative, digital tool used in high-school biology classrooms to model the local ecosystem. This tool was deployed in ten biology classrooms and used with varying degrees of success. There are significant teacher-level differences found in the activity sequences of students using the EcoSurvey tool. The theoretical metrics around scientific modeling practices and automatically extracted feature sequences were also used in a classification task to automatically determine a particular student's teacher. These results underline the power of learning analytics methods to give insight into how modeling practices are realized in the classroom. This work also informs changes to modeling tools, associated curricula, and supporting professional development around scientific modeling.

CCS Concepts

•Human-centered computing → Collaborative interaction; •Applied computing → Interactive learning environments; Collaborative learning; •Computing

Publication rights licensed to ACM. ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of the United States government. As such, the Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only.

LAK '17, March 13 - 17, 2017, Vancouver, BC, Canada

© 2017 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-4870-6/17/03...\$15.00

DOI: <http://dx.doi.org/10.1145/3027385.3027420>

methodologies → *Classification and regression trees; Scientific visualization;*

Keywords

Scientific Modeling; Collaborative Modeling; Teacher Differences; Classification

1. INTRODUCTION

Scientific models represent ideas, processes, and phenomena by describing important components, their characteristics, and their interactions. Models are constructed across a broad spectrum of scientific disciplines, such as the food web in biology, the water cycle in Earth science, or the structure of the solar system in astronomy. Models are central to the work of scientists for understanding phenomena, and for constructing and communicating theories. Constructing and using models to explain scientific phenomena is also an essential practice in contemporary science classrooms. In *A Framework for K-12 Science Education* [23], developing and using models is one of the eight core practices deemed essential for science learning and instruction. According to the *Framework*, "[s]cientists use models... to represent their current understanding of a system (or parts of a system) under study, to aid in the development of questions and explanations, to generate data that can be used to make predictions, and to communicate ideas to others" [24].

Scientific models can take many forms, such as textual descriptions, visual diagrams, computer simulations, and mathematical equations. For instance, in elementary physical science, Schwarz et al [29] studied the development of students' modeling practices by having students sketch models depicting how light interacts with objects to produce shadows. Bryce et al [2] asked students to construct a clay model of a cell. Even these simple modeling activities push students to represent their current knowledge and to use this knowledge to explain new phenomena. Models are often more complex, involving visual representations or computer simulations. Such models may focus on the complex interactions between components (e.g. predator-prey interactions in a food web) or depict how a substance changes state over time (e.g., how water changes from liquid to gas as it moves through stages in the water cycle).

In this research, we study the development of student modeling practices in secondary biology classrooms. In these classrooms, students used a web-based software tool - Eco-

Survey - to characterize organisms and their interrelationships found in their local urban ecosystem. Students use EcoSurvey to: (1) photograph, map and characterize local species, (2) document how species interact around shared resources such as food, and (3) identify resources and species that are important to the resilience of their environment. EcoSurvey follows in a rich tradition of computer-based modeling tools [32, 14, 5]. These digital modeling tools provide built-in affordances that foreground important scientific modeling practices, and are explicitly designed to scaffold students' modeling activities, through the careful design of the interface and prompts promoting reflection and appropriate action [27, 5]. As such, they support students to develop more complex models that would be difficult to create using traditional tools and these models can be quickly revised thanks to their digital nature.

Digital modeling tools also provide an opportunity for instrumentation to unobtrusively capture usage. Reflecting contemporary software architectures, EcoSurvey is a cloud-based software tool, where all changes and refinements to student models are centrally captured and stored, providing researchers with a fine-grained record of student modeling practices at scale, across potentially thousands of students in a wide range of classroom settings. These rich data offer opportunities for new learning analytic methods to better characterize student scientific modeling practices and to examine classroom level differences. In this paper, we use learning analytics and machine learning techniques to answer the following questions:

- 1) What variation do we see in the models created by students to support explanations of scientific phenomena, in our case, ecosystem biodiversity?
- 2) What variation do we see in student modeling practices across different teachers?
- 3) Can the action sequences used by students during modeling be used to predict each student's teacher?

We analyzed EcoSurvey usage data collected from over 200 secondary students across ten classrooms. We observed large variations in the completeness and complexity of student models, and large variations in their iterative refinement processes. We also observed large differences in student modeling practices across different classrooms and teachers, and we were able to predict a student's teacher based on the observed modeling practices with a high degree of accuracy without significant tuning of the predictive model. These results highlight the value of this approach for extending our understanding of student engagement with an important contemporary science practice, as well as the potential value of analytics for identifying critical differences in classroom implementation. These results shed light on potential improvements in tools and curricula. Before discussing our approach and results further, we first present the education and learning sciences theories underpinning this work and describe our research context and the EcoSurvey tool in more detail.

2. THEORY AND RELATED WORK

A central goal of our approach is to develop theoretically-grounded analytic methods. Education research and the learning sciences offer insights into three areas critical to our approach: the elements of a "good" student model, how to characterize student modeling practices, and variation in classroom implementation across teachers.

2.1 Scientific models that support complete explanations of phenomena

Scientific models are tools for explanation and prediction. A complete scientific explanation should "explain observed relationships between variables and describe the mechanisms that support cause and effect inferences about them" [23]. Thus, to support student explanations, a scientific model of a phenomenon should include important components ("variables"), their interactions ("relationships"), and define the mechanisms involved. When modeling an ecosystem, these correspond to the organisms in the ecosystem (animals, plants, insects, fungi, etc), how these organisms interact with each other and the environment (predator, prey, producer, decomposer, etc), and the involved processes (abiotic, biotic, etc). Professional biologists use this information to measure the biodiversity of an ecosystem in terms of species richness, evenness, and divergence [10, 18, 6].

In this work, we characterize variation in students models by examining the number of organisms present, the variety of types of organisms present, the number of interactions between organisms that students have identified, and the diversity of these interaction types. We also look at how these features are distributed within a model. These measures are used to understand the complexity of a student model. Interestingly, understanding the complexity of an ecosystem has been shown to support students to develop empathy and other affective stances towards nature [13]. Student understanding the flow of matter and energy through ecosystems has also been shown to vary strongly across cultural boundaries [1], providing further motivation for understanding variation in student models and student modeling practices.

2.2 Strong student scientific modeling practices

Constructing scientific models is part of the "inquiry" tradition in science education, where students learn scientific concepts through hands-on "doing" [15]. Understanding what students are doing at a fine-grained level can provide teachers with useful insights into learning processes, as well as provide teachers with feedback as to where and when students need additional assistance. Towards this end, several scholars have developed frameworks characterizing effective student modeling practices [29, 2]. Schwarz et al. [29] identify a series of seven practices: (1) identifying the anchoring phenomena to be modeled, (2) constructing a model, (3) testing the model, (4) evaluating a model, (5) comparing the model against other ideas, (6) revising the model, and (7) using the model to predict or explain phenomena. Bryce et al [2] identify a similar set of practices as important to support student learning during modeling, namely (1) observation (paralleling the anchoring phenomena), (2) model construction, (3) model use, (4) model evaluation, and (5) model revision. Their research suggests that supporting students to engage in these practices can lead to positive learning outcomes [29].

Here, we focus on a subset of these practices - constructing, evaluating, revising, and using models - incorporating them into our analysis framework [3]. We focus on these four practices as they are directly supported through the EcoSurvey interface and can be readily observed and tracked in the usage log. In addition to these four practices, we examine the degree to which students engaged in iterative design of their models. Iteration occurs when students cycle between

the other four modeling practices, where the four practices correspond directly to individual actions in the EcoSurvey interface, such as adding an organism or relationship (construction), editing an organism or relationship (revision), or generating a graph of the entire ecosystem to support explanations (using). Iteration is an important modeling practices that is used to both expand the scope of a model and to improve its accuracy [11, 2]. Learning analytic techniques are used to identify the degree to which students used these practices and to examine variations in student modeling practices. While these usage log analysis methods are an excellent passive way to collect data on student practices [25], it is important to note that these methods do not capture information about how students are reasoning with their models. Exploring student reasoning with models and how they generate explanation using models is beyond the scope of this study, and would require deep exploration of students' cognitive processes using think-alouds, cognitive interviews or other learning and cognitive sciences research methods (e.g. [29]).

2.3 Teacher Differences

Student learning outcomes vary widely across teachers [12, 21]. Students with a top-performing math teacher can be expected to perform .266 standard deviations better on a standardized math test than those with a median teacher [12]. Similarly, McNeill et al [21] evaluated 22 high school ecology classrooms across the US and found that teacher differences accounted for 34.5% of the variance on scores from a multiple choice assessment and 42.5% of the variance on scores from an open ended assessment. Differences in student learning outcomes can be attributed, in part, to differences in their opportunities to learn different topics [20, 22]. For instance, in a classroom setting, the opportunity for iteration can be driven by the structure of the class: students will not expand or refine their model if they are not given the opportunity to do so.

Differences in student learning can also be attributed to differences in the curriculum being utilized, and differences in how teachers implement curriculum in their specific classroom [12]. Large variations in how teachers implement STEM "inquiry-oriented" curriculum have been routinely observed [16, 28], and curriculum integrating modeling is no exception. Windschitl et al [35] conducted a series of studies examining how K-12 teachers integrated student modeling into their classrooms and found significant variance in teacher understanding and adoption. For many teachers, the traditional scientific method notion of generating a hypothesis is deeply ingrained in their views of science practices. Subsequently, these teachers had difficulty adopting a scientific practice that required them to ground ideas and predictions in an initial model. In some cases, they found that teachers simply rejected the model-based inquiry approach, citing that providing students with opportunities to engage in iterative practices took too much classroom time and added unnecessary complexity.

In our analysis, we examine variations in student modeling across classrooms and teachers, analyzing both students' opportunities to learn and variations in the degree to which they engaged in specific modeling practices. For these analyses, we use measures of frequency and variety as features [31]. Frequency characterizes how often students were able to engage in the different modeling practices, whereas vari-

ety captures the breadth of practices that they engaged in. Frequency and variety have been shown to reliably predict the uptake and adoption of new technologies across different groups of users [31, 19]. Here, we use these features to study the different patterns of uptake and adoption of modeling practices across classrooms.

We also explore the ways in which an individual student's modeling processes can be indicative of teacher differences. We use sequence classification techniques [36] to detect recurring patterns, called sequential patterns or action sequence features, in student's modeling practices, as they engage in cycles of creating, evaluating, revising, and using their models. We explore the degree to which automatically extracted and optimized action sequence features are able to correctly predict a specific student's teacher. These pattern mining methods have been used by learning analytics researchers to address questions related to course selection trajectories [4] and group work dynamics [26]. Automatic feature optimization is a common technique used in data mining to identify the features that carry predictive value for classification [9]; the resulting features can reveal insights into processes important for differentiating between categories [7]. In our case, we are using these sequences to detect and understand potential differences in modeling curriculum implementation across teachers.

3. RESEARCH CONTEXT: INQUIRY HUB AND ECOSURVEY

EcoSurvey was developed as part of a larger collaborative design-based research project called the Inquiry Hub, which is focused on supporting teachers in developing student-centered approaches to curriculum and teaching [30]. Inquiry Hub Biology is a digital high school biology curriculum developed in partnership with a large urban school district in the midwestern United States. Within the ecosystems unit of this curriculum, students are asked to choose a tree to plant on their school grounds or other designated site that will improve their local ecosystem's biodiversity and resilience. Classes use EcoSurvey to create a collective model of their local ecosystem. They use these models to provide evidence and construct arguments to support their choice about the type of tree they choose to plant. The recommended type of tree is then planted on the site, in collaboration with the local Parks and Recreation Department, based on the students' arguments and evidence. Thus, the models students create using EcoSurvey support them to construct arguments with real world consequences. To illustrate the use of EcoSurvey within this context, we follow the experience of "Maria", a fictional student in Ms. Smith's 3rd period class.

3.1 Data Collection and Creating the Model

Ms. Smith instructs students to map the ecosystem within a selected site on their school grounds or in the local area, taking pictures and making field notes on the organisms and interactions between organisms that they observe. Maria's group makes observations along the creek that runs next to the school. They find a lady beetle, a honey locust tree, some mushrooms, a gray squirrel, and a few other organisms. Using their smartphones, they take pictures of these organisms and upload them to EcoSurvey, creating a "card" for each organism while out in the field. Each card automat-

ically captures information about the date, time, and location of the observation being recorded. Cards also include a "relations" field to capture interactions between organisms and information about the organism's role in the ecosystem. Students begin entering this information as they observe it in the field, and then continue to augment this information back in the classroom through additional research. In Figure 1, we see Maria's lady beetle card under construction. While in the field, she created the card, uploaded a picture, and added details about interactions they saw. At the same time, her team members are also creating cards for other organisms they are observing.

Figure 1: Edit view for Maria's Lady Beetle card.

3.2 Evaluating the Model

As students create cards, their organisms are added to a shared class "survey". The survey view shows all of the organism cards and their detailed information, ordered by how recently they were edited. Maria can see that her classmates have created many cards, including a Blue Jay card (Figure 2).

Ms. Smith organizes the student groups into pairs and asks each group to review the other's cards for correctness and completeness. Maria's group is paired with Group 2, who completed several cards. Andre, a member of Group 2, asks Maria to first review the red tailed hawk card he created. Maria uses the search feature of the survey view to quickly find the hawk among the cards. She notices that this card is missing many details, including interactions with other organisms.

Figure 2: Main view of Ms. Smith's class survey.

3.3 Revising the Model

Maria recommends that Group 2 do further research into how the hawk contributes to the local ecosystem. She also takes the chance to update her gray squirrel and honey locust cards. She discovered that hawks prey upon squirrels and nest in honey locust trees during her earlier

research. She didn't realize that their school ecosystem included hawks until she reviewed the work of her classmates, as her group did not see a hawk. Once Maria has completed editing her group's cards, she continues her review of Group 2's cards. She uses the group select function to view only the cards created by members of Group 2.

Group 2 notices that two people in Maria's group created duplicate lady beetle cards. Maria decides to add her lady beetle information to the other card, since it is more detailed, and uses the delete function to remove her lady beetle card from the model.

3.4 Iterating the Model

In reviewing Group 2 cards, Maria sees a card for geese, but notices that the group did not add a predatory relationship to grass, even though she observed geese eat the grass on the soccer field. She uses the search functionality and discovers that no one in class created a card to document grass as an observed organism. Maria adds a new card for grass and includes a predatory-prey relationship with geese. By cycling back through earlier modeling practices (creating new cards), Maria is iteratively improving the class model to be more complete and accurate.

3.5 Using the Model

Once the class has created a robust model of their local ecosystem, students use this model to construct arguments for choosing a particular tree to plant. Maria presses the "create relation graph" button, which generates the graph representation of the model and exports it to a digital graphing tool (Figure 3). Maria and her team study the resulting diagram that enables them to visualize the relationships (links) between all the organisms (nodes) they have cataloged. It is clear from looking at her graph that the English Oak trees are an important keystone species in their site, involved in a large number of relationships with a wide variety of organisms. The geospatial locations in the observational data indicate that there are only two English Oak trees located in their site; Maria and her group recommend planting an additional tree of this type.

Figure 3: A section of Maria's final graph.

3.6 Analyzing EcoSurvey Use

Maria's scenario illustrates how EcoSurvey supports students to engage in the practices of creating, evaluating, revising, iterating, and using models. To use a learning analytics approach to study modeling practices, we must map specific actions, or sequences of actions, taken in the EcoSurvey in-

Table 1: EcoSurvey Actions

Modeling Practice	Description	EcoSurvey Actions
Create Model	Create a new entry in the model	New Card
Evaluate Model	Explore the organisms and interactions in the current model	Group Select, Search
Revise Model	Edit or delete organisms and interactions included in the current model	Edit, Delete
Use Model	Export a representation of the model for use (e.g. constructing an argument)	Generate Graph, Download
Iterate	Cycle between creation, revision, and use practices	New Card, Edit, Delete, Generate Graph, Download

terface to specific modeling practices. Table 1 describes the mapping between modeling practices and specific EcoSurvey interface actions that we use in our analyses. As students interact with EcoSurvey, the system captures and logs each of the actions shown in Table 1. Each log entry includes the time, user, survey, and action type.

4. METHODS

Here, we describe data used in our analyses as well as the specific analytic techniques used to answer each of our three research questions. All teachers’ names are pseudonyms.

4.1 Study Data

EcoSurvey usage log data was collected from 262 students, across 10 high school classrooms, during Fall 2015. These 10 classes were taught by three different teachers: Anderson, Baker, and Chavez. Anderson taught two periods of high school biology, which she elected to combine into one group to produce a single ecosystems model. Baker taught three periods, while Chavez taught five. From the sample, we recorded actions for 204 students, while 58 students did not record any activity. All classrooms in this sample followed a 3:1 device deployment where three students used one laptop together; thus it is not surprising that there are students with no recorded activity. A total of 9 models were created, which included 586 organism cards and 545 interactions, generating 3160 action logs.

4.2 Variation in Student Scientific Models

Our first research question examines variation within student models, focusing specifically on the richness of students’ models in terms of the number of organisms and their relationships. We analyze the relative number of organisms and interactions within each class survey. We also look at the balance of interactions per organism by evaluating both the average number of interactions per organism and variance in the distribution of interactions. Examining variance allows us to distinguish different patterns in the assignment of interactions to organisms. Some classes may create models where most organisms have a similar number of interactions, while other classes may create models where only a few organisms have been assigned many interactions.

4.3 Variation in Modeling Practices

Our second research question examines variation in student modeling practices, focusing on action variety, frequency, and iteration. Action variety refers to the range of actions a student performed. For example, some students may have only created and edited cards, while others may have used

the full range of EcoSurvey actions. Frequency refers to the total number of actions completed by an individual student and the number of usage sessions they engaged in. Sessions are defined by a series of actions from a single user without a large break in activity (greater than two hours). Defining a session using a two hour gap allows for any student activity within a long class period to occur within one session; several of our classrooms employ 1.5 hours block periods.

To characterize iteration practices, we look for evidence of design cycles within the log information. Design cycles can be recognized when students engage in multiple sequences of construct-revise-use practices. This focus on a sequence of practices is consistent with Schwarz et al [29], which characterized modeling practices as a series of steps. By extension, a design cycle consists of returning to a previous modeling step after moving on in the sequence (e.g. creating a new card after editing a different card). We counted the number of cycles as a measure of iteration.

Combined, these three metrics - action variety, frequency, and iteration - yield an eight feature vector for each student consisting of total number of EcoSurvey actions, total number of create actions, total number of evaluate actions, total number of revise actions, total number of use actions, total number of EcoSurvey action types taken, number of sessions, and number of iterations. We combined the feature vectors for students with the same teacher, and performed a Kruskal-Wallis H test [17] for each feature to determine differences between teachers. A Kruskal-Wallis H test is a non-parametric adaptation of an ANOVA to compare samples of different sizes, as we have in our groups. We further explored these differences using Tukey’s HSD test [33] to test the significance of pairwise differences between teachers.

4.4 Predictive Value of Modeling Practices

Our third research question examines the degree to which we can use sequences of student modeling actions to predict that student’s teacher. For this prediction task, we use the previously described features of variety, frequency, and iteration as well as automatically extracted sequence patterns. This sequence pattern approach is inspired by the feature-based sequence classification methods summarized by Xing, Pei, and Keogh [36]. In our work, a sequence pattern consists of a series of EcoSurvey actions (e.g. "New Card", "Edit", "Generate Graph") embedded within a student’s complete action log. To extract sequence patterns, we used the Colibri Core [34] software package. This software package, originally designed for natural language processing tasks, treats every action as a token and determines the frequency of consecutive token sequences (n-grams) from

Table 2: Final models for each class.

Survey	# Users	Organisms	Interactions	Interactions Per Organism	Interaction Variance
Anderson 4 & 7	29	155	264	1.7	4.35
Baker 1	28	47	7	0.149	0.297
Baker 2	27	25	5	0.2	0.24
Baker 4	29	19	0	0	0
Chavez 1	27	88	70	0.795	0.663
Chavez 2	29	45	27	0.6	1.31
Chavez 6	30	60	57	0.95	3.78
Chavez 7	31	81	82	1.012	5.72
Chavez 8	32	66	33	0.5	0.826

student usage logs. These token sequences can include wild-card actions (skip-grams). For instance, the software will extract the sequence "New Card", "Edit", "Generate Graph" as either an n-gram or as the skip-gram "New Card", {"*"}, "Generate Graph". This skip-gram will capture similar sequence patterns, where one action occurs between New Card and Generate Graph actions. This yielded 2,893 unique sequence patterns, that occurred at least three times, across all student usage logs. Once we extracted these sequence patterns, we used them as a new series of features to augment each student's existing feature vector. This approach parallels that used by d'Aquin et al [4], where they used sequential pattern mining to study student course enrollment patterns.

To understand which features that characterize a student's modeling actions are most predictive of his or her teacher, we input subsets of each student's feature vector into four Naive Bayes classifiers using Weka [8]. The first classifier used the eight features related to variety, frequency, and iteration of actions. The second classifier used the full set of sequence pattern extracted by Colibri Core for each student. The third classifier implemented a best-first search [9], which automatically reduced the full set of sequence patterns to the eighteen most predictive features. The last classifier combines the eight variety, frequency, and iteration features with the eighteen most predictive sequence patterns. Each test was run using 10-fold cross validation.

5. RESULTS

Results are presented for each of our three research questions.

5.1 (RQ1) What variation do we see in the models created by students?

As shown In Table 2, there are substantial variations in the models created by students in different classrooms. We see that Anderson's students documented many more organisms (155) and interactions (264) than all other classes. Though Anderson had both of her classes work together to create one survey, the total number of students contributing to this model is comparable to the number of students contributing in other classrooms. We also see that students in Baker's three classes each documented significantly fewer organisms and interactions. One class only documented 19 organisms (less than one per student) and did not document any interactions. Chavez's classes exhibit wide variation, particularly in the numbers of interactions documented by each class.

The number of interactions per organism, a broad measure of model complexity, further illustrates apparent classroom differences, with Anderson's class creating more complex models than Baker's and Chavez's classes. To better understand classroom differences, we examine variance in the number of interactions per organism. In Anderson's class, we see a high variance in comparison to the interactions per organism metric, which indicates that there are a small number of organisms with lots of interactions and many organisms with few interactions.

Chavez' P1 and P7 classes provide a particularly interesting case to examine this variation. On reviewing Table 2, we see that the variance in the number of interactions assigned to each organism is significantly lower in P1 than in P7, while the actual number of organisms and interactions are comparable. Further analyses reveal that students in Chavez's P1 did not assign any interactions for 39% of their organisms, while students in P7 did not assign interactions to 74% of their organisms. A similar analysis revealed that 42% of the organisms documented in Anderson's model did not include interactions. In most classes, the majority of organisms have no documented interactions. It appears that students engaged significantly more with describing organisms, and spent far less time consistently documenting interactions.

5.2 (RQ2) What variation do we see in student modeling practices across different teachers?

There are significant differences between the student action sequences of our three teachers on all eight metrics related to variety, frequency, and iteration ($p < .001$). Our Tukey's HSD test for each feature shows that the three groups are each distinct to a significant degree in Create, Revision, and Iteration frequency (Figure 4a, $p < .05$), as well as Overall Actions, Session Count, and Action Variety (Figure 4b, $p < .05$). We also see Anderson's students performed significantly more Evaluate and Use actions than the other two teachers' students (Figure 4a, $p < .05$), though the differences between Baker's and Chavez's students are not significant. Anderson's class also used EcoSurvey twice as much, as measured by session counts. Overall, Anderson's students engaged in more modeling practices than both of the other two groups, and Chavez's students engaged in more modeling practices than Baker's.

There were also differences in the modeling practices that students employed. Students in Baker's classes rarely engaged in three of the five modeling practices we are studying: revisions, iteration, or use. Chavez's class engaged with

(a) The average number of actions by modeling practice type. (b) The average number of actions, types of actions, and action sessions.

Figure 4: Student modeling practices for each teacher's students.

four of the five practices, but appeared to rarely use their models.

5.3 (RQ3) Can the action sequences used by students during modeling be used to predict each student's teacher?

As shown in Table 3, student action sequences can predict their teacher with varying degrees of reliability depending upon the features used. Our baseline assumes that each student is in one of Chavez's classes; almost 52% of the students in this study were in one of his classes. All of the feature sets we studied improved performance over the baseline. Classifying based on all 2,893 sequence patterns improved our classification accuracy by almost 12%, whereas classifying solely based on our variety, frequency, and iteration features improved performance by over 15%. We also trained a model on the best sequence patterns, that is, the 18 most

predictive patterns identified by Weka's Attribute Selection tool [9]; this yielded a nearly 25% improvement in performance. The best performing model was one that combined the most predictive sequence patterns with our variety, frequency, and iteration features. This combination resulted in a 30% improvement over baseline, correctly predicting a student's teacher 80% of the time.

The most useful features for classification accuracy are the 18 "best" sequence patterns (Table 4). A closer examination reveals that these sequence patterns correspond to our five modeling practices in interesting ways. These patterns prioritize model revision, evaluation, and iteration as distinguishing features, which correspond to the differences in classroom modeling practices discussed under research question 2.

To better understand the types of errors that our best performing model makes, we generated a confusion matrix

Table 3: Predictive accuracy of each action sequence feature set.

Feature Set	# Attributes	Naive Bayes Acc
Baseline	0	51.96%
All Sequence Patterns	2,893	63.73%
Variety, Frequency, and Iteration Features	4	67.65%
Best Sequence Patterns	18	75.00%
Combined Features	22	80.39%

Table 4: The most predictive action sequences.

New card, New card, {*} ¹ , New card, {*, {*, New card	New card, {*, Group Select, {*, New card	Group Select
Group Select, {*, Group Select	Group Select, {*, {*, {*, Group Select	Group Select, {*, New card, {*, New card
Group Select, Search	Search, {*, {*, {*, Edit	Edit
Edit, Edit	Edit, {*, Edit	Edit {*} {*} Edit
Edit, Search	Edit, Generate Graph, Download	Edit, Generate Graph, Download, Edit
Generate Graph	Download	Generate Graph, Download

(Table 5). We see that 75% of the errors are due to the misclassification of 30 of Chavez’s students as Baker’s students. One possible reason for this misclassification is that some students in Chavez’s classes performed very few modeling actions overall, similarly to the majority of students in Baker’s classes.

Table 5: Combined features confusion table.

		Classified As		
		Anderson	Baker	Chavez
Correct Class	Anderson	29	0	1
	Baker	1	64	3
	Chavez	5	30	71

6. DISCUSSION

In this study, we demonstrated the utility of learning analytic methods for characterizing variation in students’ scientific models and their modeling practices. We also showed that an individual student’s modelling action sequences can be used to predict his or her teacher. Our results support Windschitl et al’s findings documenting large variations in how teachers implement modeling in their classrooms [35]. While we did not conduct direct classroom observations, our analysis revealed profound, quantifiable differences in the models that students constructed across different classrooms and significant differences in their classroom learning experiences as depicted in the range of modeling practices that they engaged in.

Student models exhibited large variance in the number of organisms and interactions documented. These differences could be due to a variety of factors, such as the time allocated to modeling during class, the degree to which modeling practices were incorporated into instruction, or their teacher’s dispositions and knowledge about scientific modeling. Our results suggest that such teacher level differences do matter. Another source of variation could be differences in ability and knowledge that individual students bring to the modeling task. In our current work, we are revising the Inquiry Hub curriculum to provide better guidance to teachers to integrate modeling into their classroom, and we are providing more opportunities for students to engage in modeling throughout the unit.

Our analysis of student models also revealed a disturbing similarity across all classrooms and teachers: all the models contained significant percentages of organisms that did

¹A {*} refers to a wild card in a skip-gram, which can be completed with any value.

not have a single defined interaction with another organism. Thus, these student models are missing critical elements of a complete and sound ecosystem model. It is unlikely that these models can support students to develop comprehensive explanations and predictions as called out in the *Framework* [23]. There are multiple possible explanations for these behaviors, including weaknesses in the Inquiry Hub curriculum, the associated teacher professional development, or the design of the EcoSurvey tool. As a first step, we have made major changes to the design of EcoSurvey version 2 to make it easier for students to establish relationships from multiple parts of the interface, to visualize established relationships through an integrated graph view, and to see which organisms are not connected to others in the model.

The large variance we observed in student modeling practices provides evidence of significant teacher-level differences. Clearly these teachers are implementing EcoSurvey and the corresponding lessons differently in their classrooms, with wildly varying results. When teachers devoted more time to modeling, as measured by sessions, their students’ engaged in a richer variety of modeling practices. Prior research suggests that there is a linkage between student engagement in modeling practices and future learning outcomes [29, 2]. Thus, it appears that students in several of our participating classrooms lacked critical opportunities to learn [20, 22], that could ultimately impact their academic performance. In future work, we plan to examine the relationships between student engagement in modeling practices and their learning outcomes as measured by end-of-course school district assessments.

Our predictive analysis provided further evidence of significant teacher-level differences. The feature selection algorithm honed in on the presence or absence of three modeling practices - evaluation, revision, and iteration - as the features that best predicted a student’s teacher. This suggests that future professional development and curriculum design should focus on these specific practices, ensuring that all students get an opportunity to participate in these parts of the modeling process. In EcoSurvey version 2, we have expanded features designed to support evaluation, revision, and iteration practices. For instance, we have implemented generating a visual graph of their model directly into the tool, rather than exporting this information into a 3rd party graphic tool. By facilitating students to use (visualize) their models more frequently, we hope that this will prompt them to notice shortcomings and engage in more iterative refinements. The most accurate classifier also benefited from additional features characterizing action variety, frequency (number of actions), and iteration. These features further highlight differences in student engagement, with some stu-

dents missing the opportunity to explore, develop, and use their models over time.

A core aspect of our analytic approach explicitly linked specific user interface actions in the EcoSurvey tool to individual modeling practices identified through prior research: creating, evaluating, revising, using, and iterating [29, 2, 11]. This approach enabled us to work with theoretically and empirically sound features identified through prior classroom research. And, this approach enabled us to interpret the action sequences identified as salient by our algorithms in a theoretically-informed way, enabling us to link our findings back to instructional concerns, such as curriculum design and professional development. This method of linking interface actions to identified modeling practices could support generalizing this analytic approach to other tools that support scientific modeling.

While this study yielded many results that have informed our partnership design work, there are several limitations that are important to note. First, we are working with a limited data set, containing data from only three teachers and 9 models. While we generated interesting insights into differences between these classrooms, it is difficult to generalize our findings to a broader spectrum of classrooms. Second, we cannot attribute our observed variation in models and modeling practices to student-level differences, due to the shared and collaborative nature of the deployment. All our participating classrooms asked students to work in groups and each group shared a single laptop computer; we are actually observing the collaborative modeling practices of small groups rather than individual students.

7. CONCLUSION

We have demonstrated that learning analytics can be used to study student scientific models and student modeling practices at a scale that has previously been impossible. We used quantitative statistical measures to study variation across models and teachers. We also used methods drawn from data mining and machine learning to identify critical differences in student modeling practices and to explore which features of student modeling sequences are useful for classification.

This work opens the door for a wide variety of further research. Future directions could incorporate student demographics and examine potential differences in the uptake of modeling practices across various populations. Future work could also incorporate student assessment data to look at connections between engagement in modeling practices and student learning outcomes. Other work could further explore teacher-level differences, combining classroom observations with learning analytics to better understand the different approaches teachers take during classroom implementation.

The work presented here has already informed the Inquiry Hub partnership's effort. The design-based research team is making evidence-based changes to our curriculum, professional development, and classroom tools based on these results. Other research groups studying student scientific modeling can apply these theories and analytic techniques in their settings to understand variation in models, modeling practices, and classroom implementation.

8. ACKNOWLEDGMENTS

This material is based in part upon work supported by the National Science Foundation under Grant Numbers 1555550 and 1147590. 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. In addition, the authors would like to thank the Gordon and Betty Moore Foundation for their support.

9. REFERENCES

- [1] M. Bang, D. L. Medin, and S. Atran. Cultural mosaics and mental models of nature. *Proceedings of the National Academy of Sciences of the United States of America*, 104(35):13868–13874, 2007.
- [2] C. Bryce, V. B. Baliga, K. de Nesnera, D. Fiack, K. Goetz, L. M. Tarjan, C. Wade, V. Yovovich, S. Baumgart, D. Bard, D. Ash, I. M. Parker, and G. S. Gilbert. Exploring Models in the Biology Classroom. *The American Biology Teacher*, 8(1):35–42, 2016.
- [3] M. Cukurova, K. Avramides, D. Spikol, R. Luckin, and M. Mavrikis. An analysis framework for collaborative problem solving in practice-based learning activities: A mixed-method approach. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, LAK '16, pages 84–88, New York, NY, USA, 2016. ACM.
- [4] M. d'Aquin and N. Jay. Interpreting data mining results with linked data for learning analytics: Motivation, case study and directions. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, LAK '13, pages 155–164, New York, NY, USA, 2013. ACM.
- [5] E. B. Fretz, H.-K. Wu, B. Zhang, E. A. Davis, J. S. Krajcik, and E. Soloway. An investigation of software scaffolds supporting modeling practices. *Research in Science Education*, 32(4):567–589, 2002.
- [6] L. H. Gunderson. Ecological Resilience—In Theory and Application. *Annual Review of Ecology and Systematics*, 31:425–439, 2000.
- [7] I. Guyon and A. Elisseeff. An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research*, 3(3):1157–1182, 2003.
- [8] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The WEKA data mining software. *ACM SIGKDD Explorations*, 11(1):10–18, 2009.
- [9] M. Hall and G. Holmes. Benchmarking attribute selection techniques for data mining. *IEEE Transactions on Knowledge and Data Engineering*, 15(6):1437–1447, 2003.
- [10] C. S. Holling. Resilience and Stability of Ecological Systems. *Annual Review of Ecology and Systematics*, 4:1–23, 1973.
- [11] J. B. Homer. Why We Iterate: Scientific Modeling in Theory and Practice. *System Dynamics Review*, 12(1):1–19, 1996.
- [12] T. J. Kane, D. F. McCaffrey, T. Miller, and D. O. Staiger. Have we identified effective teachers? validating measures of effective teaching using random assignment, 2013.
- [13] S. R. Kellert and E. O. Wilson. *The biophilia hypothesis*. Island Press, 1995.

- [14] D. J. Ketelhut, B. C. Nelson, J. Clarke, and C. Dede. A multi-user virtual environment for building and assessing higher order inquiry skills in science. *British Journal of Educational Technology*, 41(1):56–68, 2010.
- [15] J. L. Kolodner, P. J. Camp, D. Crismond, B. Fasse, J. Gray, J. Holbrook, S. Puntambekar, and M. Ryan. Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting learning by design(tm) into practice. *Journal of the Learning Sciences*, 12(4):495–547, 2003.
- [16] J. Krajcik, P. Blumenfeld, R. Marx, K. Bass, J. Fredricks, and E. Soloway. Inquiry in Project-Based Science Classrooms: Initial Attempts by Middle School Students. *Journal of the Learning Sciences*, 7(3):313–350, 1998.
- [17] W. H. Kruskal and W. A. Wallis. Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, 47(260):583–621, 1952.
- [18] N. W. H. Mason, D. Mouillot, W. G. Lee, J. B. Wilson, and H. Setälä. Functional richness, functional evenness and functional divergence: The primary components of functional diversity. *Oikos*, 111(1):112–118, 2005.
- [19] K. E. Maull, M. G. Saldivar, and T. Sumner. Understanding digital library adoption: a use diffusion approach. *Proceedings of the 11th annual international ACM/IEEE joint conference on digital libraries*, pages 259–268, 2011.
- [20] L. M. McDonnell. Opportunity to learn as a research concept and a policy instrument. *Educational Evaluation and Policy Analysis*, 17(3):305–322, 1995.
- [21] K. L. McNeill, D. S. Pimentel, and E. G. Strauss. The impact of high school science teachers’ beliefs, curricular enactments and experience on student learning during an inquiry-based urban ecology curriculum. *International Journal of Science Education*, 35(15):2608–2644, 2011.
- [22] Y. Mo, K. Singh, and M. Chang. Opportunity to learn and student engagement: A HLM study on eighth grade science achievement. *Educational Research for Policy and Practice*, 12(1):3–19, 2013.
- [23] National Research Council. *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press, 2012.
- [24] National Research Council. *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*, page 57. National Academies Press, 2012.
- [25] R. Pelánek, J. Rihák, and J. Papoušek. Impact of Data Collection on Interpretation and Evaluation of Student Models. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, pages 40–47, 2016.
- [26] D. Perera, J. Kay, I. Koprinska, K. Yacef, and O. R. Zaïane. Clustering and sequential pattern mining of online collaborative learning data. *IEEE Transactions on Knowledge and Data Engineering*, 21(6):759–772, 2009.
- [27] C. Quintana, B. J. Reiser, E. a. Davis, J. Krajcik, E. Fretz, R. G. Duncan, E. Kyza, D. Edelson, and E. Soloway. A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences*, 13(3):337–386, 2004.
- [28] L. Schauble, R. Glaser, R. a. Duschl, S. Schulze, and J. John. Students’ Understanding of the Objectives and Procedures of Experimentation in the Science Classroom. *Journal of the Learning Sciences*, 4(2):131–166, 1995.
- [29] C. V. Schwarz, B. J. Reiser, E. A. Davis, L. Kenyon, A. Achér, D. Fortus, Y. Shwartz, B. Hug, and J. Krajcik. Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, 46(6):632–654, 2009.
- [30] S. Severance, W. R. Penuel, T. Sumner, and H. Leary. Organizing for Teacher Agency in Curricular Co-Design. *Journal of the Learning Sciences*, 25(4):531–564, 2016.
- [31] C.-F. Shih and A. Venkatesh. Beyond adoption: Development and application of a use-diffusion model. *Journal of Marketing*, 68(1):59–72, 2004.
- [32] E. Soloway, A. Z. Pryor, J. S. Krajcik, S. Jackson, S. J. Stratford, M. Wisnudel, and J. T. Klein. Scienceware’s model-it: Technology to support authentic science inquiry. *T.H.E. Journal*, 25(3):54–56, 1997.
- [33] J. W. Tukey. Comparing individual means in the analysis of variance. *Biometrics*, 5(2):99–114, 1949.
- [34] M. van Gompel and A. van den Bosch. Efficient n-gram, Skipgram and Flexgram Modelling with Colibri Core. *Journal of Open Research Software*, 4(1):e30, 2016.
- [35] M. Windschitl, J. Thompson, and M. Braaten. Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education*, 92(5):941–967, 2008.
- [36] Z. Xing, J. Pei, and E. Keogh. A brief survey on sequence classification. *ACM SIGKDD Explorations Newsletter*, 12(1):40, 2010.