

# Project Management, Engagement, and Sustainability Plan (PMESP)

for National Science Foundation Cooperative Agreement EAR-2148762

December 2022

## **1. Introduction**

The Community Surface Dynamics Modeling System (CSDMS) is an NSF-supported facility that serves the US and international Earth-surface geoscience research community. The facility advances research in Earth surface dynamics by developing and sharing computing resources for numerical modeling, engaging and coordinating the community, and providing educational opportunities and resources related to geoscientific computing for the research community as well as broader communities of interest. Primary activities of CSDMS are supported by a Cooperative Agreement from the NSF Geoinformatics program entitled "Engaging a thriving community of practice in Earth-surface dynamics" (EAR-2148762), hereafter referred to as the project. The scientific motivation for the project centers on several broad research themes that have been highlighted in recent NSF Decadal Survey Reports (Box 1). A common challenge across these themes lies in cyberinfrastructure: researchers must be able to work with, adapt, and synthesize computational models and associated digital data across diverse subdisciplines. The project is designed to support research by (1) developing, supporting, and disseminating modern cyberinfrastructure, (2) engaging community in the creation, improvement, use, and sharing of this infrastructure, including models, tools, and related resources, and (3) providing educational opportunities and resources that help advance the community's capacity to take advantage of modern computing technology and best practices. These three areas of effort form the three key pillars of the CSDMS facility: community, computing, and education.

This document describes the plan for managing the project, engaging and building community, and laying the groundwork to sustain impacts beyond the 5-year term of the project. The document is divided into three major sections. The **Management** section describes how the project is governed and coordinated, identifies the roles and responsibilities of project personnel, and details the milestones and

planned schedule for major project activities, together with metrics for progress and impacts. The **Engagement** section describes the scientific communities served by the project, plans for broad engagement with these communities, and the expected science outcomes. The **Sustainability** section describes plans and activities to lay the groundwork for sustaining the project's outcomes beyond the award end date.

Box 1. High-priority science questions related to Earth-surface dynamics.

Questions from Earth in Time report (NRC, 2020) What are the causes and consequences of topographic change? How does the critical zone influence climate? What does Earth's past reveal about the dynamics of the climate system? How is the water cycle changing? How do geological processes influence biodiversity? How can Earth science research reduce the risk and toll of geohazards?

Questions from Sea Change report (NRC, 2015)

How are the coastal and estuarine ocean and their ecosystems influenced by the global hydrologic cycle, land use, and upwelling from the deep ocean?

What are the processes that control the formation and evolution of ocean basins?

How can risk be better characterized and the ability to forecast geohazards like megaearthquakes, tsunamis, undersea landslides, and volcanic eruptions be improved?

## 2. Management

## 2.1 Overview

The management plan builds on several principles: a dynamic, metric-driven approach that adapts to lessons learned; a governance structure that draws together the constituent communities; close ties with allied projects and organizations; leveraging of new developments in cyber science; and designing for sustainability.

## 2.2 Community-Based Governance Structure

CSDMS is founded on the principle of community governance and engagement. Community input guides strategic goals and science initiatives. CSDMS operates with bylaws that are reviewed regularly and revised when needed. Community Members are presently organized into 12 groups: Terrestrial, Coastal, Marine, Education and Knowledge Transfer, Cyberinformatics and Numerics, Hydrology, Carbonates and Biogenics, Chesapeake, Geodynamics, Critical Zone, Human Dimensions, Ecosystem Dynamics. These

groups represent the diverse science communities that compose CSDMS membership. Groups are guided by chairs, who act as liaisons with their domain communities and sponsoring organizations. Groups advise the CSDMS Integration Facility (CIF) on community needs and research directions. Groups may occasionally be created or merged, as trends in community interest and participation evolve.

The **Executive Committee** (EC), comprising the group chairs along with the executive director and the chair of the Steering Committee, represents the community and acts as the primary decision-making body. EC oversees CIF activities, and reviews progress on an on-going basis. The EC also approves reports, plans, co-sponsoring organizations, and other major issues pertaining to the running of CSDMS. During the project, EC will meet semi-annually (normally at the annual meeting in spring, and remotely in autumn), with additional remote meetings as needed.

The Scientific **Steering Committee** (SC) provides high-level guidance, assessing progress, identifying potential new opportunities, and advising on strategy. The SC also approves the bylaws and revisions. The SC will meet at least once annually, with additional meetings as needed. One of the first tasks in project year 1 will be to obtain one or more nominations for a new SC chair, and hold an election.

Growing uptake of CSDMS tools and protocols among US federal agencies motivates an **Interagency Committee** (IC). IC serves as a conduit between CSDMS and government research and operations agencies such as USGS, NOAA, DoE, USACE, EPA, and BOEM. Periodic remote meetings will provide a forum for exchange of information about projects related to cyberinfrastructure for environmental process modeling.

CSDMS operates with a mature set of bylaws available on the CSDMS web portal (<u>https://csdms.colorado.edu/csdms\_wiki/images/Bylaws\_2020-5.pdf</u>). The bylaws are reviewed periodically by the Executive Committee, and proposed revisions are submitted to the Steering Committee for approval and adoption.

#### 2.3 Personnel Roles and Responsibilities

The **Executive Director** oversees the management and operation of CIF and coordinates the project. Gregory Tucker will serve as Executive Director and Lead PI. He will have overall responsibility for project management and will engage in all aspects of CSDMS operations, including governance, CIF management, and interactions with other model development programs. Duties include chairing (Executive) or ex-officiating (Steering, Interagency) committees, and directing CIF (personnel, deliverables, priorities, and activities).

The **Deputy Director** supports the Director in implementing strategic plans and policies, and represents CSDMS at events for which the Director is unavailable. Irina Overeem will serve as Deputy Director. She will also oversee workforce development efforts of the CIF, including the Earth Surface Processes Institute (ESPIn) program. In addition, she will contribute to and direct graduate student efforts for the demonstration projects on 3D hydrodynamics integration.

The **CSDMS Integration Facility (CIF)** team will be responsible for developing cyberinfrastructure, coordinating community-engagement activities, and maintaining and operating CSDMS programs and resources. The **Project Coordinator and Executive Assistant**, Lynn McCready, will be responsible for CIF administrative management, coordination of meetings and workshops, coordinating logistics for ESPIn and Roadshows, general travel coordination, facilitation of governance and Group activities (elections, bylaws revisions, annual progress reviews, etc.), reporting, communications, marketing, and special initiatives. The **Geospatial and Web Services Specialist**, Albert

Kettner, will manage the CSDMS data and model portals and all CSDMS repositories and libraries. He will coordinate integration of geospatial information system capability into the framework, and is the data liaison for CSDMS. He will also have primary responsibility for the CSDMS web portal and coordination of the CSDMS membership base.

The CIF's **Research Software Engineers** (RSEs) conduct cyberinfrastructure development and related educational outreach. The **Chief Research Software Engineer**, Eric Hutton, will oversee and actively support software design, development, implementation, and service of model components and software tools in accordance with the specific needs of the diverse modeling community. He will also provide guidance regarding strategic sustainability of CSDMS software and oversee community high-performance computing (HPCC), cloud computing operations, and user interactions. A **Senior Research Software Engineer**, Mark Piper, will work with Dr. Hutton to design, develop, implement, and maintain CSDMS software products and services. He will also be responsible for the CSDMS Roadshows, developing software-related educational materials, teaching modules in the Summer Institute, and producing materials for webinars. A **Junior Research Software Engineer**, Tian Gan, will assist in implementing, testing, documenting, and maintaining CSDMS software products and services, with a particular focus on Data Components and related technology. She will also contribute research software engineering support to the project on linking 3D hydrodynamics with sediment, carbon, and biota. In addition, Dr. Gan will provide training and contribute to development of educational products.

The project's education and outreach activities will be assisted by an **Education and Knowledge Transfer Specialist**, Matthew Rossi, who will lead the development of online training materials for the educational repository. Dr Rossi will work with the engineering team on interactive tutorials, conduct program outreach to students and members. The ESPIn program will be overseen by Irina Overeem together with collaborating PI Nicole Gasparini (Tulane University), and assisted by RSEs Piper and Gan.

The demonstration project on marine sediment modeling will be led by Julia Moriarty. She will contribute to and direct graduate student efforts for the demonstration projects linking 3D hydrodynamics with sediment, carbon and biota.

The participatory modeling activities will be led by collaborating PI Moira Zellner (Northeastern University), assisted by Hutton, Tucker, **Participatory Complex Systems Research Technician** Dean Massey (also at Northeastern), and an NEU graduate research assistant. A contractor will support ongoing evaluation.

#### 2.4 Partnership with cognate organizations and projects

Partnership with cognate organizations and projects is vital for leveraging resources, for getting the word out about CSDMS products and services, and for connecting community members with resources about which they may be unaware. Partnerships are built into CSDMS' governance structure in part through the mechanism of Group sponsorship. Several groups will continue to be co-sponsored by partner organizations. Here, sponsorship means a commitment by the partner to help identify new group chairs when needed, to encourage their community members to attend CSDMS events, and to maintain an informal line of communication with CIF. As of the project launch in September 2022, the following organizations and projects are affiliated with CSDMS through sponsorship and representation:

• **CUAHSI** sponsors the Hydrology Focus Group. CUAHSI has contributed to past CSDMS annual meetings, and has worked with CIF to identify candidates for chair of that group. In addition, CIF RSE Gan has close ties to the CUAHSI HydroShare development team.

- **Computational Infrastructure in Geodynamics (CIG)**, while not currently an official group sponsor, continues to work with CIF on topics of common interest. CIG and CSDMS have co-convened workshops, and CIG leaders have assisted in identifying candidates for chair of the Geodynamics Focus Group.
- **AIMES** (Analysis, Integration, and Modeling of the Earth System) and **CoMSESnet** (Network for Computational Modeling in Social and Ecological Sciences) serve as the co-sponsors for the Human Dimensions Focus Group, providing a link to communities that focus on modeling of natural-human system interactions.
- CSDMS is a charter member of the **Open Modeling Foundation (OMF)**, an alliance of modeling organizations that coordinates and administers a common, community developed body of standards and best practices among diverse communities of modeling scientists. OMF Executive Director Michael Barton serves on the CSDMS Steering Committee.
- The International Society for Ecological Modeling (ISEM) sponsors the Ecosystem Dynamics Focus Group.
- The International Soil Modeling Consortium (ISMC) sponsors the Critical Zone Focus Group.
- The Chesapeake Community Modeling Program (CCMP) sponsors the Chesapeake Focus Group, and their Steering Committee member Raleigh Hood chairs that group.

In addition to representation on the Executive Committee, the project will foster interaction with relevant organizations, communities, and projects through the mechanism of clinic and keynote presentations at annual meetings. Clinics in particular provide a vehicle for bringing information about new products, data sets, models, and other resources to the attention of CSDMS community members.

## 2.5 Reporting and Administration

As specified in the Programmatic Terms and Conditions, CIF will submit an Annual Report and Work Plan that includes current year deliverables, future year work commitments, and an estimate of the total current year cost and future year budget for each major activity.

## 2.6 Coordination of Project Across Collaborative Organizations

Collaboration among the three participating institutions centers on two activities: the Earth Surface Processes Institute (ESPIn) program, and the participatory modeling activity.

#### 2.6.1 Coordination for ESPIn

Tulane University PI Nicole Gasparini and CU Boulder Co-PI Irina Overeem will co-lead EPSIn. They will coordinate through video-conference meetings as needed throughout each year to plan and execute that year's ESPIn event and to analyze post-event data. During each project year, tasks and milestones in the ESPIn program include:

• Quarter 1 (Sep-Nov): Announce class dates and location. Advertise. Create application with online form. Open application window. Organize applications. Line up science lecture speakers. Make classroom reservation.

- Quarter 2 (Dec-Feb): Close application window. Review and score applications. Make participant selections. Send accept/reject notices. Make participant travel arrangements. Set daily class schedule. Review and test course material.
- Quarter 3 (Mar-May): Send out pre-class instructions and survey. Make JupyterHub logins. Deliver ESPIn. Send out post-class survey. Participant travel reimbursement.
- Quarter 4 (Jun-Aug): Reflect on outcomes. Revise course material based on participant feedback. Interpret survey results.

The ESPIn leads will provide periodic updates to CIF (during weekly team meetings), to the Executive Committee (during the fall and spring meetings), and to the Steering Committee (at its annual meeting).

#### 2.6.2 Coordination for Participatory Modeling

Northeastern University (NEU) PI Moira Zellner will lead the participatory modeling activity, assisted by CIF personnel (primarily Hutton, Tucker, and Rossi). During year 1, the NEU team (Zellner, Massey, and one or more representatives of the Fora.ai developers) and the CIF team will begin with twice-monthly video meetings to choose the first use-case applications, identify (and if need be develop or modify) suitable model(s), design the technical configuration (e.g., communication between the Fora.ai platform and a CSDMS code), and plan the major steps for the first year. Zellner also co-chairs of CSDMS Human Dimensions Focus Research Group, and in that role will participate in the fall and spring meetings of the Executive Committee.

## 2.7 Plans for Scalability

A successful research cyberinfrastructure effort can face scaling challenges as the user community grows. The use of CSDMS' cloud platform for hosting online materials helps mitigate this because the resources can be dynamically adjusted according to need. In-person training (ESPIn) is harder to scale, given staff and budget limitations. However, parallel creation of self-paced online learning materials provides a scalable and accessible way to make part of the curriculum available to an audience of effectively unlimited size. With respect to community all-hands meetings, we hope to achieve a degree of scalability by taking a thoughtful and intentional approach to remote participation options; we expect to draw on lessons from the many similar experiments currently unfolding across the sciences. For shared community software, a key element of scalability lies in capacity building: one outcome we hope to achieve through capacity-building activities is accelerating community-wide contributions to the shared collection of open resources.

## 2.8 Work Plan

#### 2.8.1 Major Project Activities, Timeline, Effort, and Budget

This section outlines plans, timelines, and resources for the major project activities. The three major areas of activity are (1) **community engagement**; (2) **cyberinfrastructure research, development, and maintenance**; and (3) **governance, management, and operations**. The work plan is organized around these three major headings. The major activities under these headings are briefly summarized below, with references to the relevant sections of the proposal. The quarter-by-quarter schedule for the major project

activities is illustrated in Figure 1. The personnel effort dedicated to each of these areas is summarized in Figure 2. The budget allocated to each area is summarized in Figure 3.



Figure 1: Quarter-by-quarter schedule for the major project activities.



Figure 1 (continued)



Figure 1 (continued)



Figure 1 (continued)



Figure 1 (continued)

		Community Engagement	Cyber- infrastructure	Governance, management, operations	Other	TOTAL
	Total person-months by category, year 1	20.905	22.75	15.1	4.61	63.365
	Total person-months by category, year 2	19.95	19.1	12.2	4.61	55.86
Annual Efforts	Total person-months by category, year 3	15.2	9	9.05	4.61	37.86
	Total person-months by category, year 4	16.05	8.9	8.63	4.61	38.19
	Total person-months by category, year 5	16.8	7.95	8.83	4.61	38.19
	Eric Hutton, Chief RSE	3.5	3.5	1	0	8
	Mark Piper, RSE	2.75	2	1.25	0	6
	Lynn McCready, Admin	2.75	0.25	3	0	6
	Irina Overeem, Deputy Director, EKT Lead	0.5	0.2	0.3	0	1
	Greg Tucker, Lead PI, Executive Dir	0.405	0	0.85	0	1.255
	Tian Gan, RSE	1	6	2.5	0	9.5
	Post Doc 2, RSE	2.5	4	3	0	9.5
	Julia Moriarty, Use Case	0	0.4	0.1	0	0.5
Budgeted Y1	Matt Rossi, EKT	3.9	0	0.1	0	4
	Albert Kettner, Web Master	3	2	3	0	8
	INSTAAR IT / accounting tech	0	0	0	4.61	4.61
	Student, Use Case	0.1	4.4	0	0	4.5
	Nicole Gasparini, Tulane PI, EKT	0.5	0	0	0	0.5
	Moira Zellner, NEU PI, Participatory modeling	0	0.5	0	0	0.5
	Research Technician, NEU, Participatory modeling	0	0	0	0	0
	Student, NEU, Participatory modeling	0	6	0	0	6
	Dan Milz (NEU contractor), Participatory modeling	0	0	0	0	0
	Eric Hutton, Chief RSE	3	4	1	0	8
	Mark Piper, RSE	2.5	2.25	1.25	0	6
	Lynn McCready, Admin	3	0.25	2.75	0	6
	Irina Overeem, Deputy Director, EKT Lead	0.6	0.2	0.2	0	1
	Greg Tucker, Lead PI, Executive Dir	0.85	0.1	0.8	0	1.75
	Tian Gan, RSE	1	4	1	0	6
	Post Doc 2, RSE	2.5	1.5	2	0	6
	Julia Moriarty, Use Case	0	0.4	0.1	0	0.5
Budgeted Y2	Matt Rossi, EKT	2.9	0	0.1	0	3
	Albert Kettner, Web Master	3	2	3	0	8
	INSTAAR IT / accounting tech	0	0	0	4.61	4.61
	Student, Use Case	0.1	4.4	0	0	4.5
	Nicole Gasparini, Tulane PI, EKT	0.5	0	0	0	0.5
	Moira Zellner, NEU PI, Participatory modeling	0	0.5	0	0	0.5
	Research Technician, NEU, Participatory modeling	0	0	0	0	0
	Student, NEU, Participatory modeling	0	6	0	0	6
	Dan Milz (NEU contractor), Participatory modeling	0	1.3	0	0	1.3

## Yearly Effort by Task and Person, in Person-Months

*Figure 2: Yearly personnel effort by task (continued on next page).* 

	Eric Hutton, Chief RSE	2.5	4.5	1	0	8
	Mark Piper, RSE	2	1.75	1.25	0	5
	Lynn McCready, Admin	3	0.25	2.75	0	6
	Irina Overeem, Deputy Director, EKT Lead	0.3	0	0.2	0	0.5
	Greg Tucker, Lead PI, Executive Dir	1	0.1	0.65	0	1.75
	Tian Gan, RSE	0	0	0	0	0
	Post Doc 2, RSE	0	0	0	0	0
	Julia Moriarty, Use Case	0	0.4	0.1	0	0.5
Budgeted Y3	Matt Rossi, EKT	2.9	0	0.1	0	3
J	Albert Kettner, Web Master	3	2	3	0	8
	INSTAAR IT / accounting tech	0	0	0	4.61	4.61
	Student, Use Case	0	0	0	0	0
	Nicole Gasparini, Tulane PI, EKT	0.5	0	0	0	0.5
	Moira Zellner, NEU PL Participatory modeling	0	0.5	0	0	0.5
	Research Technician, NEU, Participatory modeling	0	3	0	0	3
	Student, NEU, Participatory modeling	0	6	0	0	6
	Dan Milz (NEU contractor). Participatory modeling	0	1.3	0	0	1.3
	Fric Hutton, Chief BSE	3	4	1	0	8
	Mark Piper BSE	2	1 75	1 25	0	5
	Lynn McCready, Admin	3	0.25	2 75	0	6
	Irina Overeem, Deputy Director, EKT Lead	0.3	0	0.2	0	0.5
	Greg Tucker Lead PL Executive Dir	1 35	0	0.73	0	2.08
	Tian Gan RSE	0	0	0.75	0	0
	Post Doc 2 RSF	0	0	0	0	0
	Julia Moriarty, Use Case	0	0.4	0.1	0	0.5
Budgeted V/	Matt Rossi EKT	2.9	0.4	0.1	0	3
Dudgeted 14	Albert Kettner, Web Master	3	2.5	2.5	0	8
	INSTAR IT / accounting tech	0	0	0	4.61	4.61
	Student Use Case	0	0	0	0	4.01
	Nicole Gasparini, Tulane PL EKT	0.5	0	0	0	0.5
	Moira Zellner, NELL PL, Participaton/ modeling	0.0	0.5	0	0	0.5
	Research Technician NELL Participatory modeling	0	3	0	0	3
	Student NELL Participatory modeling	0	0	0	0	0
	Dan Milz (NEU contractor) Participatory modeling	0	13	0	0	13
	Eric Hutton, Chief RSE	2	1.0	1	0	0
	Mark Piper, PSE	2 75	4	1 25	0	5
	Lypp McCroady, Admin	2.15	0.25	2.75	0	5
	Lina Oversom Deputy Director EKT Load	0.3	0.25	2.75	0	0.5
	Greg Tucker Load PL Executive Dir	1 15	0	0.2	0	2.08
	Tian Can RSE	1.15	0	0.95	0	2.00
	Post Doc 2 PSE	0	0	0	0	0
	Lulia Mariarty, Lian Casa	0.2	0.2	0 1	0	0.5
Rudgeted V5	Matt Possi EKT	2.0	0.2	0.1	0	0.5
Budgeted 15	Albert Ketteer, Web Meeter	2.9	2.5	0.1	0	 
		0	2.5	2.5	1.61	0
	Student Use Case	0	0	0	4.01	4.01
	Nicele Gasparini Tulane PL EKT	0.5	0	0	0	0.5
	Moira Zollaar NELLAL Participaton modeling	0.5	0.5	0	0	0.5
	Passareh Tashnisian NELL Participatory modeling	0	0.5	0	0	0.5
	Student NELL Deticipatory modeling	0	0	0	0	0
	Den Milz (NEU controctor) Derticipatory modeling	0	0	0	0	0
	Dan will (NEU contractor), Participatory modeling	U	U	U	U	U

*Figure 2 (continued): Yearly personnel effort by task.* 

Annual Summary Budget						
ITEM	¥1	Y2	Y3	Y4	Y5	5-yr TOTAL
Community Engagement	\$529,017	\$523,792	\$485,816	\$510,041	\$547,809	\$2,596,475
Cyberinfrastructure	\$469,403	\$460,697	\$345,074	\$285,943	\$209,473	\$1,770,590
Governance, management, and operations	\$304,320	\$274,415	\$233,989	\$231,118	\$243,349	\$1,287,191
Other	\$72,813	\$74,741	\$75,577	\$60,575	\$62,043	\$345,749
TOTAL \$\$\$	\$1,375,553	\$1,333,645	\$1,140,456	\$1,087,677	\$1,062,674	\$6,000,005

Figure 3: Yearly budgeted expenditures by task.

#### 2.8.1.1 Community Engagement

Major community engagement activities and tasks are listed below, with the relevant sections of the proposal given in parentheses:

- 1. **Earth Surface Processes Institute (ESPIn)** (3.1): Plan, organize, and convene an approximately six-day scientific computing workshop for graduate students and early-career scientists. One event per project year.
- 2. **Self-paced learning materials** (3.1): Use ESPIn materials as the basis for an online collection of self-paced learning materials.
- 3. Annual all-hands meetings (3.2): Plan, organize, and convene all-hands meetings each project year.
- 4. **CSDMS Roadshows** (3.3): Provide onsite cyber-skills workshops (1-3 days) at a variety of academic and research institutions around the US, with two workshops per year.
- 5. **Research Software Engineering as a Service (RSEaaS)** (3.4): Provide consulting services by CIF Research Software Engineers. Includes regular videoconference office hours, an online Help Desk, and opportunities for direct project support.
- 6. **Community Forum** (3.5): Set up and operate an online community platform for member-to-member communications and support.
- 7. **Executable Articles** (3.6): Design, trial, and refine a process by which Jupyter Notebooks (and possibly other forms of digital notebook) may be submitted to the annual meetings as peer-reviewed scholarly products.
- 8. **Disseminate Project Resources and Findings:** Broadcast information about products, services, and findings through academic publications, conference presentations, email newsletters, social media posts, and other venues. This category includes personnel travel to conferences and workshops.

#### 2.8.1.2 Cyberinfrastructure Research, Development, and Maintenance

The cyberinfrastructure efforts fall into five broad categories:

- 1. **Next-Generation Code Repositories** (4.1): Improve the CSDMS Model Repository and associated source-code repositories to enhance their adherence to the principles of Findability, Accessibility, Interoperability, and Reproducibility.
- 2. Enhancing the CSDMS Workbench (4.2): Improve the set of cyber-tools and resources that compose the CSDMS Workbench.
- 3. **Participatory Modeling** (4.3): Develop and test a platform for participatory modeling that brings together CSDMS models with the Fora.ai platform.
- 4. **High-Performance Computing** (4.4): Continue to support CSDMS community members in using CU Boulder *Blanca* computing cluster and cloud-based resources, which includes nodes dedicated to CSDMS users. (Blanca support will be limited to its lifespan, and we anticipate a gradual transition to a cloud-based resource.)
- 5. Linking Hydrodynamics with Sediment, Carbon, and Biota (4.5): Conduct a demonstration study that illustrates how to combine 3D hydrodynamic model output with project tools and resources.

#### 2.8.1.3 Governance, Management, and Operations

Activities associated with governance, management, and operations include:

- 1. **Governance** (5.1): Coordinate twice-annual meetings of the Executive Committee, and annual meetings of the Steering Committee. Recruit new Group chairs as needed. Conduct strategic planning.
- 2. **Management** (5.2): Manage personnel and budget, conduct CIF team meetings and other communications, track progress toward project milestones, coordinate among collaborating partners, communicate at least quarterly with NSF representative(s), and prepare project reports.
- 3. **Facility and Operations** (5.2): Maintain and update web portal. Maintain software products and tools, including operating Continuous Integration on source repositories, keeping dependencies up to date, managing user issues and pull requests, responding to community feedback, updating documentation, and managing outreach.
- 4. **Metrics Dashboard** (7.0): Develop and deploy a web-based community metrics dashboard on the CSDMS web portal.

#### 2.8.2 Metrics and Deliverables

The project tracks a set of metrics related to the three major areas of activity. Metrics are listed in Figure 4. Deliverables represent events (such as meetings), products (such as self-paced learning materials), and capabilities (such as the ability to launch an executable demonstration of a model from the Model Repository). The schedule for deliverable events and product planning is indicated on Figure 1.

SO	
R	
Ш	
$\geq$	

(note: * indicates cumulative metric)				YEAR 1	YEAR 2	YEAR 3	YEAR 4	YEAR 5
ACTIVITY / AREA	METRIC	<b>BASIC TARGET</b>	<b>ASPIRATIONAL TARGET</b>					
COMMUNITY ENGAGEMENT								
Membership	number of members	stable @ 2,200-2,300	continued growth	2,200 (2,300)	2,200 (2,400)	2,200 (2,500)	2,200 (2,600)	2,200 (2,700)
ESPIn	enrollment	capacity (25)	capacity (25)	25*	50*	75*	100*	125*
ESPIn	sustained demand	meets capacity	exceeds capacity					
ESPIn	% contact with cohort	10%	25%					
ESPIn	% skills still used	20%	50%					
ESPIn	# products from projects	2	4	2 (4) *	4 (8) *	6 (12) *	8 (16) *	10 (20) *
Roadshows	# site visits	10	10	2 *	4*	6*	8*	10*
Roadshows	# attendees	8 per event	16 per event	16*	32*	48*	64*	80*
Roadshows	# of interviews	1 per event	2 per event	2 (4) *	4 (8) *	6 (12) *	8 (16) *	10 (20) *
Conferences	# convened	1 per year	1 per year	1*	2*	3*	4*	5*
Conferences	# in-person attendees	100 per meeting	140 per meeting	100*	200*	300*	400*	500*
Conferences	# remote attendees	20 per meeting	60 per meeting	20*	40*	e0*	80*	100*
Conferences	clinic enrollment	75 per meeting	105 per meeting	75*	150*	225*	300*	375*
Conferences	# first-time attendees	15 per meeting	30 per meeting	15 (30) *	30 (60) *	45 (90) *	60 (120) *	75 (150) *
Conferences	# posters	40	0.6	40*	80*	120*	160*	200*
Conferences	# software posters	5	0.1	5*	10*	15*	20*	25*
Conferences	meetings held at MSIs	2	0.02	0	1	*	2*	2*
Roadshows	# MSI-affiliated participants	2 per event	4 per event	4*	*8	12*	16*	20*
Conferences	Speaker gender representation	AGU 2020 ratio (1:2 F:M)	parity					
Conferences	Speaker ethnic/racial diversity	similar to AGU 2020	similar to US population					
Conferences	URM scholarships	15	15	15*	30*	45*	60*	75*
ESPIn	Participant gender representation	AGU 2020 ratio	parity					
ESPIn	Participant ethnic/racial diversity	similar to AGU 2020	similar to US population					
ESPIn	participation by MSI faculty/ECRs	З	4	3 (4) *	6 (8) *	9 (12) *	12 (16) *	15 (20) *
Conferences	# MSI-affiliated participants	2	4	2 (4) *	4 (8) *	6 (12) *	8 (16) *	10 (20) *
RSEaaS	# members helped per yr	9	12	6 (12) *	12 (24) *	18 (36) *	24 (48) *	30 (60) *
RSEaaS	products from RSE-assisted projects	-	2	1 (2) *	2 (4) *	3 (6) *	4 (8) *	5 (10) *
Curated notebooks	# of submissions	5	10	5 (10) *	10 (20) *	20 (40) *	30 (60) *	60 (120) *
Curated notebooks	# of citations to notebooks per nb	-	2	n/a	10 (20) *	20 (40) *	30 (60) *	40 (80) *

Figure 4: Summary of project metrics (continued on next page).

CYBERINFRASTRUCTURE								
Code repositories	# launchable models	Q	10	0	1 (1) *	2 (3) *	3 (6) *	5 (10) *
Code repositories	# of user model launches / year	10 per year per model	20 per year per model	0	10 (10) *	30 (40) *	60 (100) *	110 (200) *
Code repositories	# of projects reviewed & enhanced	5	10	1 (2) *	2 (4) *	3 (6) *	4 (8) *	5 (10) *
Workbench	# ABMs available	ъ	9	0	1 (2) *	2 (3) *	3 (4) *	3 (6) *
Workbench	# BMI specs automated	4	ъ	0	*	2 *	3 *	4 (5) *
Participatory modeling	user test sessions	2	ę	1 (1) *	2 (2) *	2 (3) *	2 (3) *	2 (3) *
Participatory modeling	workshops	2	2	0	1 (1) *	2 (2) *	2 (2) *	2 (2) *
Participatory modeling	attendees per workshop	9	12	0	6 (12) *	12 (24) *	12 (24) *	12 (24) *
Participatory modeling	sectors represented	ю	9	0	2 (4) *	3 (6) *	3 (6) *	3 (6) *
HPCC	# of users per year	4	8					
Community usage	community publications	10 per year	20 per year	10 (20) *	20 (40) *	30 (60) *	40 (80) *	50 (100) *
Community usage	community presentations	10 per year	20 per year	10 (20) *	20 (40) *	30 (60) *	40 (80) *	50 (100) *
<b>GOVERNANCE, MANAGEMENT, OF</b>	PERATIONS							
Governance	EC gender representation	AGU 2020 ratio (1:2 F:M)	parity					
Governance	SC gender representation	AGU 2020 ratio (1:2 F:M)	parity					
Governance	EC ethnic/racial diversity	similar to AGU 2020	similar to US population					
Governance	SC ethnic/racial diversity	similar to AGU 2020	similar to US population					
ADDITIONAL SUSTAINABILITY ME	ETRICS							
Community cyber contributions	items posted/answered	20 per year	growing annually	20 (30) *	40 (70) *	60 (120) *	80 (180) *	100 (250) *
Community cyber contributions	unique users posting	15 per year	growing annually	15 (20) *	30 (45) *	45 (75) *	60 (110) *	75 (150) *
Community cyber contributions	user-submitted pull reqs	5 per year	growing annually	5 (6) *	10 (14) *	15 (24) *	20 (36) *	25 (50) *
Community cyber contributions	new codes in model repo	5 per year	10 per year	5 (10) *	10 (20) *	15 (30) *	20 (50) *	25 (50) *
Community EKT contributions	EKT items contributed	4 per year	growing annually	4 (5) *	8 (11) *	12 (18) *	16 (26) *	20 (35) *
Community support requests	# of requests per year	5	10	5 (10) *	10 (20) *	15 (30) *	20 (50) *	25 (50) *
Community orgs represented EC	# of organizations	9	9					
Code transferrability	% code test coverage	80%	95%					
Code transferrability	CI test status	passing	passing					
Code transferrability	% documented modules	80%	95%					

Figure 4: Summary of project metrics (continued from previous page).

## 3. Engagement

Community engagement is a central theme in the project. This section describes the communities served by the project, and the approaches to engaging them.

## 3.1 Engaging the CSDMS research community as a whole

The research community served by the project includes academic and government scientists whose research intersects surface dynamics: transport of sediment and solutes; formation and change of landscapes, seascapes, and deposits; weathering and soils; and similar phenomena. The community research scope encompasses a variety of disciplines and subdisciplines that bear on earth-surface processes, as reflected in the thematic groups discussed below. Based on current membership, and extrapolating to include researchers who might benefit from CSDMS resources but have not signed up for membership, we estimate that the research community served or potentially served by CSDMS numbers a few thousand worldwide. In the context of this project, *engagement* of this community has several facets:

- Providing opportunities to interact with fellow researchers and practitioners across a range of diverse but overlapping fields, and to learn about the latest modeling-related research and technology in these fields.
- Promoting awareness of potentially useful resources and services, including but not limited to those provided by CIF.
- Providing opportunities to learn new technological and methodological skills pertaining to computational modeling and/or earth-surface dynamics.
- Providing expert technical assistance in best-practice scientific software development and related cyberinfrastructure, particularly as pertains to numerical modeling.

Project activities designed to promote these forms of engagement include the annual all-hands meetings, quarterly email newsletters, the Research Software Engineering as a Service (RSEaaS) program, and the online CSDMS Forum. The schedules and milestones associated with these activities are shown in Figure 1. The personnel effort and budget levels devoted to them are shown in Figures 2 and 3, respectively. Metrics are listed in Figure 4.

There are two primary expected outcomes of research community engagement. The first is growing adoption of CSDMS tools and resources for science and engineering applications. One metric indicator of such adoption is scientific products, such as articles, theses, and conference presentations, that make use of CSDMS tools and resources. The second is contributions to the constellation of tools and resources. Community contributions can take many forms, including sharing or updating of codes and associated metadata through the Model Repository; making a model code interoperable by "wrapping" it with the standardized Basic Model Interface (BMI); applying BMI in a project; posting feature requests or bug reports; responding to other users' queries; posting or reviewing Pull Requests (which are online requests to merge a proposed code addition or change into the main version); serving on a product-based steering group such as the BMI Council; developing new components or other functionality in a library such as Landlab; creating and sharing an educational tutorial; and so on. The corresponding metrics for community contribution relate to the volume and growth in the quantity and quality of these contributions over the duration of the award (Figure 4).

#### 3.2 Engaging domain, disciplinary, and cross-cutting groups

One mechanism for community engagement centers around the CSDMS Groups. Currently, CSDMS members are organized into twelve Groups. Three represent major domains of interest: the Terrestrial, Coastal, and Marine Working Groups. Two represent cross-cutting activities: the Cyberinfrastructure and Numerics Working Groups, and the Education and Knowledge Transfer Working Group. In addition to these five Working Groups, there are currently seven Focus Research Groups that represent disciplines, themes, or geographical areas: Hydrology, Geodynamics, Ecosystem Dynamics, Carbonates and Biogenics, Human Dimensions, Chesapeake, and Critical Zone. The Groups are meant to be adaptable, and the themes, numbers, and function of the Groups are expected to evolve as the project progresses.

The CSDMS Groups provide several mechanisms for community engagement. Their chairs are expected to serve as liaisons, bringing relevant issues and developments in their domains to the attention of the CIF and the Executive Committee, representing CSDMS within their own domain communities or organizations. In the case of several Groups, the connection to a particular community organization or project is formalized through the mechanism of sponsorship; for example, CUAHSI sponsors the Hydrology Group, and ISEM sponsors the Ecosystem Dynamics Group. Chairs are encouraged to communicate with their Group members through periodic email newsletters (the CIF manages a mailing list for each group for this purpose), and via formal or informal interaction at annual all-hands conferences.

#### 3.3 Engaging early career scientists

Early-career researchers (ECRs: graduate students, postdocs, and junior faculty and research scientists) are a special focus of the project. Engaging ECRs represents a way to invest in the future, as ECRs bring their knowledge and skills to new research, to future students, and to professional careers. The size of the Earth-surface and Earth-surface-adjacent ECR community is uncertain, but a conservative estimate can be made from demographic data for the American Geophysical Union (AGU). Although the reported demographics do not include career stage, a rough indication can be gleaned from the fact that as of 2020 about 18% of AGU members were under age  $30^1$ . The AGU Earth and Planetary Surface Processes (EPSP) section has (as of 2016) just over 1,000 members who list EPSP as their primary affiliation, and about 7,200 who list it as a secondary affiliation. Given this information, a minimum estimate of the relevant ECR community would be 180, as the expected number of AGU members under 30 who list EPSP as a primary affiliation. This is a minimal estimate because many ECRs are over age 30, and because CSDMS also includes members whose primary interest lies in adjacent domains (such as hydrology or tectonics) as well as people who are not AGU members at all. If one includes the full set of ~8,200 AGU EPSP-affiliated members (both primary and secondary), a rough estimate of the size of the corresponding ECR community would be about 1,500 (18%).

Programs for actively engaging the Earth-surface ECR community include the Earth Surface Processes Institute (ESPIn) (proposal section 3.1) and the Roadshows (section 3.3). Both involve direct, hands-on learning opportunities with CIF staff, in an in-person format (public health conditions permitting). ESPIn is designed to serve approximately 125 ECRs (25 per year) over the duration of the award. The scale of the Roadshows depends on enrollment, which is difficult to predict in advance, but is anticipated to be on the order of 100 ECRs over the project duration. These numbers are consistent with

<sup>&</sup>lt;sup>1</sup> Demographic data from https://www.agu.org/-/media/Files/Join/AGU\_Membership\_Demographics\_2020.pdf

the lower estimate of the Earth-surface ECR community, but fall well below the higher (aspirational) estimate. The online, asynchronous learning materials are designed to increase the reach of these programs by providing an additional mechanism for ECRs to develop scientific computing and numerical modeling skills. In addition, the Student Modeler Award, given at each annual meeting, provides an opportunity for students to gain visibility for their numerical modeling work, and is a mechanism to elevate modeling-related research (including the important aspects of code sharing, quality, and documentation) generally. The ECR community is also served by the RSEaaS program (e.g., online office hours and Help Desk).

An expected outcome from the project's support of the ECR community is a growing group of scientists who have basic training in scientific computing, use recommended practices such as version control and documentation, make use of CSDMS community cyber-resources, and make contributions such as sharing codes, engaging in online dialog, reviewing pull requests, and/or making additions and improvements to existing cyber-tools. Metrics for success include scholarly products (articles, presentations, theses) and contributions to open-source community repositories (Figure 4).

#### 3.4 Engaging students and faculty at minority-serving institutions (MSIs)

One goal of the project is to provide students and faculty at MSIs with opportunities for cyber-skill development and for participation in CSDMS events. The size of this community—consisting of MSI faculty and students who are interested in aspects of Earth-surface dynamics—is difficult to estimate with certainty. MSIs cover a wide range, from R1 schools to community colleges. US Department of Education data from 2022 indicate that there are about 500 4-year MSIs in the United States. Not all of these have programs relevant to Earth and environmental sciences, and of those that do, not all have graduate programs. As a very rough estimate, if we consider that the average 4-year MSI has one junior faculty member, graduate student, or undergraduate interested in learning more about computing and modeling in Earth-surface processes, then the size of the community is several hundred. In fact, gauging the interest level is one of the project objectives. Activities designed to engage this community include MSI scholarships for attending the annual all-hands meeting; the annual meeting itself (which will be held at MSIs in project years 2 and 4); the Roadshow program; and ESPIn. Metrics are based on participation in these activities by MSI-affiliated students and/or faculty (Figure 4).

#### 3.5 Engagement of stakeholders and decision-makers

Another group of interest consists of stakeholders concerned with environmental issues such as urban or post-wildfire flooding. The participatory modeling activity is designed to study how CSDMS models and cyberinfrastructure, in combination with the NEU Fora.ai platform, might be deployed to help inform understanding and decision making among stakeholders. This is an experimental aspect of the project, and is intended to provide insight into what modalities are and are not effective, as well as to establish a foundation for further research. This activity is anticipated to engage one to two dozen stakeholder participants.

#### 3.6 Other engagement beyond the CSDMS membership

One of the objectives of the project is to extend the reach and impact of CSDMS resources and opportunities beyond the current member community. One group of interest consists of potentially interested researchers in the US and internationally who may not know about CSDMS. These are effectively a subset of the broader research community described above. The engagement strategy for this group is based on a combination of standard modes of scientific discourse (including conference presentations and scholarly publications) together with social media posts.

## 4. Sustainability

As the 2020 *Earth in Time* report notes, software is infrastructure, and so too is the social capital that supports it. A realistic sustainability plan begins by acknowledging that no infrastructure can ever be completely maintenance-free. An appropriate goal, therefore, is to minimize maintenance needs while maximizing the likelihood that the resources will continue to be valued and curated. The strategy for accomplishing this—for building a strong foundation to sustain project outcomes—weaves together four threads: capacity building, facility continuity, partnership among allied facilities, and robust, thoughtful software engineering. The time frame considered here for sustainability is 5-10 years beyond the project end date, though the strategy described below should also increase the prospect for sustainability beyond that time frame.

### 4.1 Capacity Building

Community capacity building fosters long-term sustainability by making it as easy and attractive as possible for individuals and teams not just to use the community modeling system, but also to contribute to its ongoing improvement as a natural by-product of their own research. While there will always be a need for coordination and technical expertise, the more deeply involved and invested a community is in looking after its own software, the more cost-effective that coordination becomes. It is hypothesized that the sustainability of digital resources is maximized when a broad user community has (1) a vested interest in the resources, and (2) the know-how to maintain and enhance them. Such a community has the potential to activate a virtuous cycle: the more contributors engage, the more capable and versatile the resources become, which in turn leads to further user contributions and enhancements. Capacity-building plays an essential role in spinning up this virtuous cycle. Cyber-skills training will help equip community members with the baseline level of technical skills needed to contribute. The governance plan fosters breadth by engaging representatives of various communities across the Earth, environmental, and related social sciences. Broad governance applies not just to the overall effort but also to individual tools and resources (for example, the BMI Council is an international user group tasked with overseeing the future evolution of the CSDMS-developed Basic Model Interface). The proposed participatory modeling activities bring further breadth by engaging stakeholders outside the research community.

The metrics associated with the community engagement activities listed above and in Figure 4 also pertain to the sustainability-promoting aspects of capacity building. Of particular interest is evidence of growth in the rate of contributions from community members outside the CIF, where contributions can

include, for example, new codes, enhancements to existing products, educational resources, reviews of pull requests, and commentary on repository forum discussions.

#### 4.2 Facility Continuity

With a 15-year history and broad, committed community support, it is anticipated that the CSDMS organization will endure beyond the lifetime of a single award. One way to promote facility continuity is to nurture and support a committed leadership team within the community, so that there is a deep and broad "bench" of people who are committed to seeking resources to support the facility and its mission. A dynamic and committed leadership team is fostered by having a relatively large and diverse Executive Committee, which should be reflected in the range of career stages, scientific interest areas, and institution types. An additional indication is the use of CSDMS resources by members of the Executive Committee themselves.

Facility continuity is also enhanced when community members who feel invested in the project also feel that they are both benefiting from and contributing to the project's mission. While this is not a simple thing to measure, one indirect indication is the extent to which CIF is asked to partner with community members on projects and grant proposals, which indicates a level of value and trust in the facility (this is covered by the "community support requests" metric in Figure 4).

#### 4.3 Partnership among Allied Organizations

Cyber-oriented organizations and facilities can assist one another in sustaining resources of mutual value. For example, CUAHSI's HydroShare-based JupyterHub now offers their users a version that has CSDMS Workbench software pre-installed; for its part, CSDMS curates CUAHSI-relevant models in the Model Repository. This is an example of mutual leveraging of resources that promotes resilience by striking a balance between independence and interdependence. Achieving this balance is not simple. Modern open-source scientific software involves a complex web of dependencies. The project's strategy for resilient partnership is based on three principles: mutual awareness (cyber-projects and facilities know about one another's work and products), complementary capabilities (resources complement one another), and interoperability (products' standardization, documentation, and transparency lower the bar to transferring stewardship from one entity to another—a vision shared by the Open Modeling Foundation, of which CSDMS is a founding member). In this project, the first two are addressed by the governance structure, and by mutual participation in events (e.g., active members of a sister organization are invited to participate in CSDMS events, while CSDMS team members participate in events hosted by sister organizations). The third is addressed by thoughtful and robust software engineering, as described next.

#### 4.4 Best-practice software engineering

Research software is easier to sustain when it uses a generic, modular, adaptable design; when it takes advantage of standards (like BMI); and when it builds on mature, open-source software tools and languages with large user bases. Overall, the project's focus on building community capacity in parallel with strong cyber connective tissue is designed to place the Earth surface process community in a good position to sustain and advance its modeling infrastructure in the long term. Examples of specific techniques and practices that promote adoption, reuse, and potential transfer include: developing software

tools and educational materials in open repositories; building unit tests into software products; routinely testing the software using continuous-integration technology (in particular, running tests on modified versions before they are merged into the main development branch); providing online documentation for software products; and including substantial inline documentation in a format that can be automatically converted into formatted online documentation. These and other best-practice methods are technical in nature but they can have a substantial impact on software usability, reliability, and longevity. Metrics that reflect these practices include: the percentage of a code base that is automatically tested; the status of these tests on the current release or development versions of code; and the proportion of software products that have accompanying documentation and (ideally) tutorial examples.

### 4.5 Tracking sustainability goals

Progress toward sustainability goals will be monitored in the same way as with other project metrics. Progress will be reported annually to NSF, reviewed at meetings of the Executive and Steering Committees, and reviewed internally by CIF at least quarterly. The Year 3 NSF site review will also present an opportunity to reflect on progress toward sustainability goals, gather feedback from the review panel, and make adjustments accordingly. Year-to-year changes in sustainability-related activities and goals will be reflected in annual project reports. Further insight into the effectiveness of various approaches will be gleaned from breakout discussions at annual all-hands meetings. Responses to findings related to sustainability activities will depend on the nature of those findings. They might include, for example, improvements to documentation, changes in marketing and communications strategy, and/or development of new partnerships.