Physics-Based Virtual Prototyping, p. 4

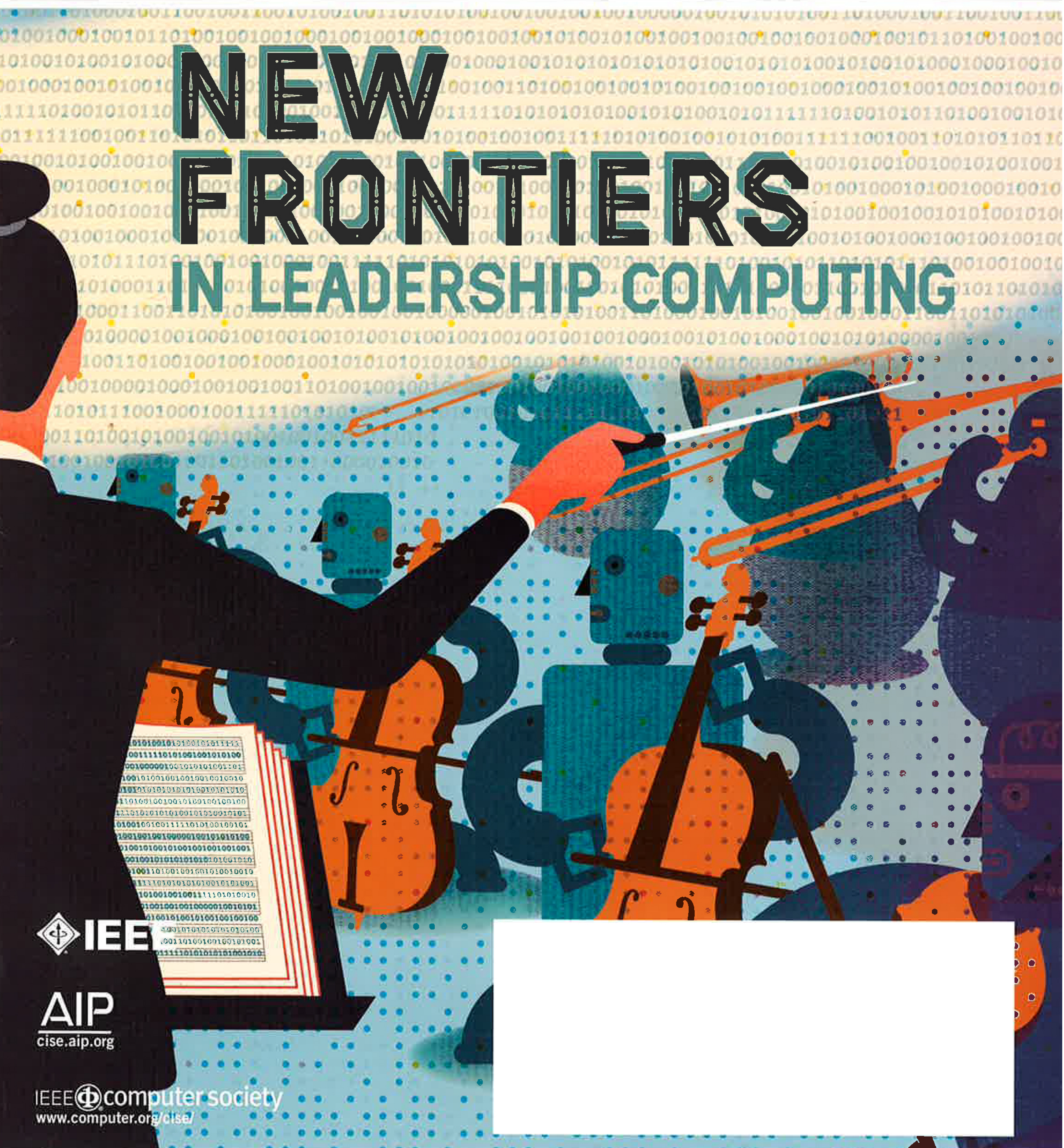
Searching for Moons with Hubble Space Telescope's Data, p. 92

Unintended Payoffs of Scientific Research, p. 104

Computing in SCIENCE & ENGINEERING

Vol. 16, No. 6 | November/December 2014

NEW FRONTIERS IN LEADERSHIP COMPUTING



IEEE

AIP
cise.aip.org

IEEE computer society
www.computer.org/cise/

Training the Next Generation of Computational Geoscientists

Ker Than | School of Earth Sciences, Stanford University

A new graduate program at Stanford University trains students in the use of high-performance computing (HPC) techniques to help analyze, simulate, and predict Earth's myriad systems and interactions.

The Computational Geoscience program, or CompGeo, is a partnership between the School of Earth Sciences (SES) and the Institute for Computational Mathematics and Engineering (ICME). It's aimed at both Earth science students who want to develop expertise in computational research and students interested in computation who want to focus on problems in the geosciences. As Hamdi Tchelepi noted, people are often surprised to learn the role that supercomputing plays in modern Earth sciences, but Earth scientists use more computer resources than almost anybody except the defense industry, and their computing needs can influence the designs of next-generation hardware. Tchelepi is a professor of Energy Resources Engineering and co-director of Stanford's Center for Computational Energy and Environmental Science (CEES), a high-performance facility used by SES faculty.

"Earth science is about understanding the complex and ever-changing dynamics of flowing air, water, oil, gas, CO_2, and heat," Tchelepi said. "That's a lot of physics, requiring extensive computing resources to model."

Tchelepi's own research involves using supercomputers to study interactions between injected CO_2 gas and the complex rock-fluid system in the subsurface. "Carbon sequestration is not a simple reversal of the technology that allows us to extract oil and gas. The physics involved is more complicated, ranging from the micro-scale of sand grains to extremely large geological formations that may extend hundreds of kilometers, and the timescales are on the order of centuries, not decades," Tchelepi said.

For example, modeling how a large plume of CO_2 injected into the ground migrates and settles within the subsurface, and whether it might escape from the injection site to affect the air quality of a faraway city, can require the solving of tens of millions of equations simultaneously.

Scientists today who study the intricate Earth system also have access to an unprecedented wealth of data gathered from different sources, ranging from satellites and aircraft-based sensors to ground-based arrays and an assortment of mobile sensors. "The amount of data that we're collecting in the Earth sciences is exploding," said Biondo Biondi, professor of Geophysics who also co-directs CEES and is the director for CompGeo. Supercomputers are playing an increasingly vital role in weaving all of these data streams together to answer important scientific questions, such as how rising atmospheric CO_2 levels will affect global climate. "Computational mathematics is the backbone of modern geosciences," said ICME director Margot Gerritsen, who is also an associate professor of Energy Resources Engineering.

Through close collaborations with the computer and oil industries, and at national labs that model atmospheric and ocean patterns, Biondi knows from firsthand experience the high demand there is for professionals with the kind of training CompGeo students receive. "Potential employers repeatedly tell me their need to hire people with these very specific sets of skills," Biondi said.

CompGeo was specifically created to address the lack at both the national and global levels of graduate programs that focus on the general interaction of computation and science. This is of particular concern in the field of computational geoscience, where there has traditionally been a lack of scientists able to fully and efficiently use advanced high-performance computational resources.

The CompGeo curriculum draws from a combination of existing courses offered at Stanford through ICME as well several new courses in computational geoscience developed by the School of Earth Sciences. CompGeo students also have access to software and hardware at CEES. Graduates of the master's program are prepared to continue on to a doctoral program or to work as scientific computing professionals in industry and government.

CompGeo student Daniel Blatter envisions computational mathematics as a key driver of geoscience discoveries for many years to come. "Whether you are interested in climate modeling or subsurface imaging, CompGeo is a great foundation for your research," he said.

Specifics

The CompGeo program is offered as a master's in science. This graduate-degree track within the ICME provides students the skills and knowledge required to develop efficient and robust numerical solutions to Earth science problems using HPC. The CompGeo curriculum is based on four fundamental areas.

Modern Programming Methods for Science and Engineering

This first area focuses on object-oriented languages (C++ and Fortran 2003); software development and maintenance tools; and leveraging open source software packages.

Applied Mathematics with an Emphasis on Numerical Methods

The second area focuses on partial differential equation solvers (finite difference, finite element, and finite volume); optimization algorithms (gradient-based and global); adjoint-state methods; and stochastic methods and computation with uncertainties.

Algorithms and Architectures for HPC

A third area focuses on large-scale parallel algorithms and tools (OpenMP, Pthreads, and message passing interface [MPI]); optimization for modern architectures (multicore, GPU, and field-programmable gate array [FPGA]); and distributed computing and large databases.

Computationally Oriented Earth Sciences Courses

The fourth and final area focuses on reservoir simulation; geophysical imaging; tectonophysics and geomechanics; climate, atmosphere, and ocean; and ecology and geobiology.

Research Areas

In addition to the curriculum areas, research can be focused in any of the departments within the School of Earth Sciences: Energy Resources Engineering; Environmental Earth System Sciences; Geological and Environmental Sciences; and Geophysics.

Academics

Now, let's look at the specific academic requirements.

CompGeo Requirements

As previously mentioned, the master's in science CompGeo track is offered through the ICME. For more information about the requirements and admissions process, please see the ICME website (https://icme.stanford.edu).

Students are required to take a total of 45 units of coursework and research credits to earn a degree. The coursework follows the requirements of the traditional ICME master's, degree with additional requirements placed on the general and focused electives. As defined in the general Graduate Student Requirements, students must maintain a grade point average (GPA) of 3.0 or better. All courses must be taken at the 200 level or higher. To continue on to the PhD program in ICME, master's students must maintain a GPA of at least 3.5.

Requirement 1: Foundational Knowledge

Students must demonstrate foundational knowledge in the field by completing the courses in two of the three core parts (see Table 1).

Courses in this area must be taken for letter grades. Deviations from the core curriculum must be justified in writing and approved by the student's ICME adviser and the chair of the ICME curriculum committee. Courses that are waived may not be counted towards the master's degree.

Requirement 2: Breadth Electives in Computational Geoscience

The master's CompGeo track requires 18 units of course work in the geosciences (three units can be applied from a non-computationally focused course). Courses are currently offered, but aren't limited to the following specific areas of the School of Earth Sciences:

- reservoir simulation;
- geophysical imaging;
- tectonophysics and geomechanics;
- climate, atmosphere, and ocean; and
- ecology, and geobiology.

The Earth science courses—offered in the Environmental Earth System Science (EESS), Energy Resources Engineering (ERE), Geological and Environmental Sciences (GES), and Geophysics Departments at Stanford—are selected based on the area of the student's interest and their research/thesis work, along with the advice and consent of the student's adviser. Students are encouraged to choose a range of courses to guarantee breadth of knowledge in Earth Sciences. A maximum of one non-computationally oriented course can be counted towards the master's degree requirements. A list of courses that fulfill these requirements is available at https://pangea.stanford.edu/programs/compgeo/academics/courses.

Requirement 3: Integrative Research in Computational Geoscience

The degree requires nine units of focused research in computational geoscience. Students are required to either

Table 1. Foundational requirements.

Course number	Course title	Units per class
Part 1 (6 units)		
CME 303	Partial Differential Equations of Applied Mathematics	3
CME 306	Numerical Solution of Partial Differential Equations	3
Part 2 (6 units)		
CME 302	Numerical Linear Algebra	3
CME 304	Numerical Optimization	3
Part 3 (6 units)		
CME 305	Discrete Mathematics and Algorithms	3
CME 308	Stochastic Methods in Engineering	3

complete a research project or an internship, as we describe here.

Research Project. Students who plan to apply to the PhD program must take nine units of research. Students will work with the CompGeo program director to find an appropriate advisor and research topic and then enroll in the course EARTHSCI 400: Directed Research (or a similar SES research course). The following could constitute a research project's successful outcome:

- an oral presentation at an international meeting that requires an extended abstract;
- a publication submission in a peer-reviewed journal; or
- a written report.

Internship. As an alternative to the research project, students have the option of an internship, which is recommended for those students interested in a terminal degree. The individual student is responsible for securing and organizing the internship and is required to obtain a faculty advisor and submit a written report on the internship project. Credit for the internship will be obtained through course EARTHSCI 401: Curricular Practical Training (1 unit)—and in this case, only eight units of research are required.

Table 2. Example programs of study for computational geoscience.*

Quarter	Climate modeling	Geophysical modeling and data analysis
Autumn	EARTHSCI 310: Computational Geoscience Seminar CME 302: Numerical Linear Algebra CME 303: Partial Differential Equations of Applied Mathematics CEE 262A: Hydrodynamics	EARTHSCI 310: Computational Geoscience Seminar CME 302: Numerical Linear Algebra CME 303: Partial Differential Equations of Applied Mathematics GEOPHYS 210: Basic Earth Imaging
Winter	CME 212: Advanced Programming for Scientists and Engineers CME 304: Numerical Optimization EESS 245A: Atmosphere, Ocean, and Climate Dynamics: The Atmospheric Circulation	CME 212: Advanced Programming for Scientists and Engineers GEOPHYS 211:Environmental Sounding Image Estimation GEOPHYS 200: Fluids AND Flow in the Earth: Computational Methods CME 500: Computational and Mathematical Engineering Seminar
Spring	CME 213: Introduction to Parallel Computing Using MPI, open MP, and CUDA EESS 246B: Atmosphere, Ocean, and Climate Dynamics: the Ocean Circulation EARTHSCI 400: Directed Research	CME 306: Numerical Solution of Partial Differential Equations GEOPHYS 287: Earthquake Seismology EARTHSCI 400: Directed Research
Summer	EARTHSCI 400: Directed Research	EARTHSCI 400: Directed Research
Autumn	EARTHSCI 310: Computational Geoscience Seminar CEE 263B: Numerical Weather Predictions EESS 211: Fundamentals of Modeling EARTHSCI 400: Directed Research	EARTHSCI 310: Computational Geoscience Seminar GEOPHYS 257: Introduction to Computational Earth Sciences EARTHSCI 400: Directed Research
Winter	CME 305: Discrete Mathematics Algorithms CEE 262B: Transport in Surface Water Flows CME 500: Computational and Mathematical Engineering Seminar	CME 304: Numerical Optimization GEOPHYS 257: Introduction to Computational Earth Sciences GEOPHYS 258: Scientific Data Processing

* The color coding of the courses is denoted as follows: integrative courses, computational and math skills, geosciences, research, and fundamental/elective.

Requirement 4: Programming

This degree also requires four units of programming coursework, to demonstrate programming proficiency. All graduate students in the program are required to complete programming courses at the level of CME 212 or higher.

Requirement 5: Seminar

Additionally, this degree requires three units of ICME graduate seminars or other approved seminars. Additional seminar units may not be counted towards the 45-unit requirement. One of the required seminars for CompGeo must be EARTHSCI 310: Computational Geosciences Seminar (1 unit).

Environmental/Climate/Hydrogeology
- EESS 215: Earth System Dynamics
- EESS 220: Physical Hydrogeology
- EESS 221: Contaminant Hydrogeology
- EESS 246B: Atmosphere, Ocean, and Climate Dynamics
- GES 224: Modeling Transport and Transformations in the Environment
- GES 264: Mathematical Modeling in Biogeochemistry

Geophysical Imaging
- GEOPHYS 204: Spectral Finite Element Method Seismograms
- GEOPHYS 210: Basic Earth Imaging
- GEOPHYS 211: Environmental Soundings Image Estimation
- GEOPHYS 240: Borehole Seismic Modeling and Imaging
- GEOPHYS 280: 3D – Seismic Imaging
- GEOPHYS 287: Earthquake Seismology

General Computational/Mathematical Geoscience
- EESS 214: Introduction to Geostatistics and Modeling of Spatial Uncertainty
- EESS 360: Advanced Statistical Methods
- ENERGY 260: Modeling Uncertainty in the Earth Sciences
- ENERGY 291: Optimization of Energy Systems
- ENERGY 284: Optimization and Inverse Modeling
- GES 240: Geostatistics
- GEOPHYS 258: Scientific Data Processing
- GEOPHYS 281: Geophysical Inverse Problems

Reservoir Simulation/Fluid flow
- GEOPHYS 200: Fluid and Flow in the Earth: Computational Methods
- ENERGY 275: Quantitative Methods in Basin and Petroleum System Modeling
- ENERGY 223: Reservoir Simulation
- ENERGY 224: Advanced Reservoir Simulation
- ENERGY 281: Applied Mathematics in Reservoir Engineering
- ENERGY 290: Numerical Modeling of Fluid Flow in Heterogeneous Porous Media

Subsurface and Reservoir Characterization
- GEOPHYS 202: Reservoir Geomechanics
- GEOPHYS 241: Seismic Reservoir Characterization
- GEOPHYS 260: Rock Physics for Reservoir Characterization

Structural/Tectonophysics/Geomechanics
- GEOPHYS 220: Ice, Water, Fire
- GEOPHYS 251: Structural Geology and Rock Mechanics
- GEOPHYS 288A: Crustal Deformation
- GEOPHYS 290: Tectonophysics

Figure 1. A sample of computational geosciences courses offered in the School of Earth Sciences. (ENERGY = Energy Resources Engineering; EESS = Environmental Earth System Science; GES = Geological and Environmental Sciences; and GEOPHYS = Geophysics.)

Finally, it's important to consider that fundamental courses in mathematics and computing may be needed as prerequisites for other courses in the program (see Table 2 and Figure 1). Check the prerequisites of each required course. Preparatory courses include such subjects as calculus; linear algebra and differential calculus of several variables; integral calculus of several variables; ordinary differential equations (ODEs) with linear algebra; linear algebra and matrix theory; vector calculus for engineers; linear algebra and PDEs for engineers; an introduction to scientific computing; linear algebra with application to engineering computations; PDEs in engineering; computer programming in C++ for Earth scientists and engineers; an introduction to large-scale computing in engineering; numerical linear algebra; programming methodology; programming abstractions; machine learning; introduction to optimization; theory of probability; and data mining and analysis.

For more information, visit http://compgeo.stanford.edu. ■

Ker Than is the associate director of communications for the School of Earth Sciences, Stanford University. His interests span from archaeology to astrophysics. Than has an MS from New York University's Science, Health, and Environmental Reporting Program (SHERP). Contact him at kerthan@stanford.edu.