Newsjournal of the Society for Industrial and Applied Mathematics sinews.siam.org

SIAM NEWS

Volume 53/ Issue 9
November 2020

Special Issue on the Mathematics of Planet Earth

*Read about the application of mathematics and computational science to issues concerning invasive populations, Arctic sea ice, insect flight, and more in this Planet Earth **special issue**!*

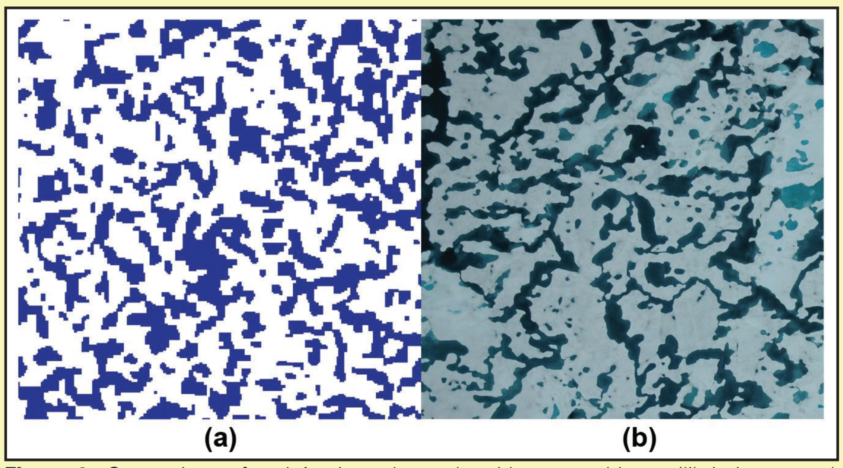


Figure 3. *Comparison of real Arctic melt ponds with metastable equilibria in our melt pond Ising model.* **3a.** *Ising model simulation.* **3b.** *Real melt pond photo. Figure 3a courtesy of Yiping Ma, 3b courtesy of Donald Perovich.*

Vast labyrinthine ponds on the surface of melting Arctic sea ice are key players in the polar climate system and upper ocean ecology. Researchers have adapted the Ising model, which was originally developed to understand magnetic materials, to study the geometry of meltwater's distribution over the sea ice surface. In an article on page 5, Kenneth Golden, Yiping Ma, Courtenay Strong, and Ivan Sudakov explore model predictions.

Controlling Invasive Populations in Rivers

By Yu Jin and Suzanne Lenhart

Flow regimes can change significantly over time and space and strongly impact all levels of river biodiversity, from the individual to the ecosystem. Invasive species in rivers—such as bighead and silver carp, as well as quagga and zebra mussels—continue to cause damage. Management of these species may include targeted adjustment of flow rates in rivers, based on recent research that examines the effects of river morphology and water flow on rivers' ecological statuses. While many previous methodologies rely on habitat suitability models or oversimplification of the hydrodynamics, few studies have focused on the integration of ecological dynamics into water flow assessments.

Earlier work yielded a hybrid modeling approach that directly links river hydrology with stream population models [3]. The hybrid model's hydrodynamic component is based on the water depth in a gradually varying river structure. The model derives the steady advective flow from this structure and relates it to flow features like water discharge, depth, velocity, cross-sectional area, bottom roughness, bottom slope, and gravitational acceleration. This approach facilitates both theoretical understanding and the generation of quantitative predictions, thus providing a way for scientists to analyze the effects of river fluctuations on population processes.

When a population spreads longitudinally in a one-dimensional (1D) river with spatial heterogeneities in habitat and temporal fluctuations in discharge, the resulting hydrodynamic population model is

$$
\begin{gathered}
N_t = -A_t(x,t)\frac{N}{A(x,t)} + \\
\frac{1}{A(x,t)}\Big(D(x,t)A(x,t)N_x\Big)_x - \\
\frac{Q(t)}{A(x,t)}N_x + rN\left(1-\frac{N}{K}\right)
\end{gathered}
$$

$$
\begin{aligned}
&N(0,t)=0 && \text{on } (0,T), x=0,\\
&N_x(L,t)=0 && \text{on } (0,T), x=L,\\
&N(x,0)=N_0(x) && \text{on } (0,L), t=0
\end{aligned}
\qquad (1)
$$

See **Invasive Populations** *on page 4*

Modeling Resource Demands and Constraints for COVID-19 Intervention Strategies

By Erin C.S. Acquesta, Walt Beyeler, Pat Finley, Katherine Klise, Monear Makvandi, and Emma Stanislawski

As the world desperately attempts to control the spread of COVID-19, the need for a model that accounts for realistic trade-offs between time, resources, and corresponding epidemiological implications is apparent. Some early mathematical models of the outbreak compared trade-offs for non-pharmaceutical interventions [3], while others derived the necessary level of test coverage for case-based interventions [4] and demonstrated the value of prioritized testing for close contacts [7].

Isolated analyses provide valuable insights, but real-world intervention strategies are interconnected. Contact tracing is the lynchpin of infection control [6] and forms the basis of prioritized testing. Therefore, quantifying the effectiveness of contact tracing is crucial to understanding the real-life implications of disease control strategies.

Contact Tracing Demands

Contact tracers are skilled, culturally competent interviewers who apply their knowledge of disease and risk factors when notifying people who have come into contact with COVID-19-infected individuals. They also continue to monitor the situation after case investigations [1].

Case investigation consists of four steps:

1. Identify and notify cases
2. Interview cases
3. Locate and notify contacts
4. Monitor contacts.

Most health departments are implementing case investigation, contact identification, and quarantine to disrupt COVID-19 transmission. The timeliness of contact tracing is constrained by the length of the infectious period, the turn-around time for testing and result reporting, and the ability to successfully reach and interview patients and their contacts. The European Centre for Disease Prevention and Control approximates that contact tracers spend one to two hours conducting an interview [2]. Estimates regarding the timelines of other steps are limited to subject matter expert elicitation and can vary based on cases' access to phone service or willingness to participate in interviews.

Bounded Exponential

The fundamental structure of our model follows traditional susceptible-exposed-infected-recovered (SEIR) compartmental modeling [5]. We add an asymptomatic population A, a hospitalized population H, and disease-related deaths D, as well as corresponding quarantine states. We define the states $\{S_i, E_i, A_i, I_i, H, R, D\}_{i=0,1}$ for our compartments, such that $i=0$ and $i=1$ correspond to unquarantined and quarantined respectively. Rather than focus on the dynamics that are associated with the state transition diagram in Figure 1, we introduce a formulation for the real-time demands on contact tracers' time as a function of infection prevalence, while also respecting constraints on resources.

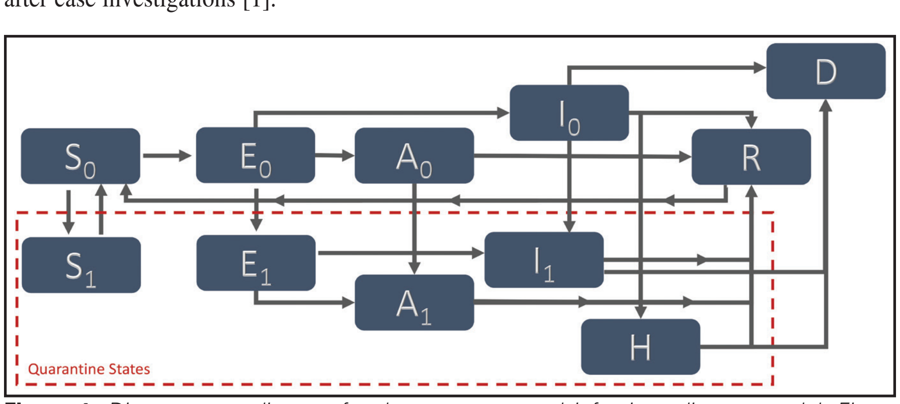


Figure 1. *Disease state diagram for the compartmental infectious disease model. Figure courtesy of the authors.*

When the work that is required to investigate new cases and monitor existing contacts exceeds available resources, a backlog develops. To simulate this backlog, we introduce a new compartment C for tracking the dynamic states of cases:

$$
\frac{dC}{dt} = [flow_{in}] - [flow_{out}].
$$

Flow into the backlog compartment, represented by $[flow_{in}]$, reflects case identification that is associated with the following transitions in the model:

– The rate of random testing: $q_{rA}(t)A_0(t) \rightarrow A_1(t)$ and $q_{rI}(t)I_0(t) \rightarrow I_1(t)$

– Testing triggered by contact tracing: $q_{tA}(t)A_0(t) \rightarrow A_1(t)$, $q_{tI}(t)I_0(t) \rightarrow I_1(t)$, and $q_{tE}(t)E_1(t) \rightarrow \{A_1(t), I_1(t)\}$

– The population that was missed by the non-pharmaceutical interventions that require hospitalization: $\tau_{IH}(t)I_0(t) \rightarrow H(t)$.

Here, $q_{r*}(t)$ defines the time-dependent rate of random testing, $q_{t*}(t)$ signifies the time-dependent rate of testing that is triggered by contact tracing, and τ_{IH} is the inverse of the expected amount of time for which an infected individual is symptomatic before hospitalization. These terms collectively provide the simulated number of newly-identified positive COVID-19 cases. However, we also need the average number of contacts per case. We thus define function $\mathcal{K}(\kappa, T_S, \phi_\kappa)$ that depends on the average number of contacts a day (κ), the average number of days for which an individual is infectious before going into isolation (T_S), and the likelihood that the individual

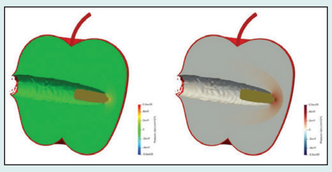
See **COVID-19 Intervention** *on page 3*

Nonprofit Org
U.S. Postage
PAID
Permit No 360
Bellmawr, NJ

SOCIETY for INDUSTRIAL and APPLIED MATHEMATICS
3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 USA

siam news
Volume 53/ Issue 9/ November 2020

6 Scientific Visualization: New Techniques in Production Software
The field of scientific visualization, which is dedicated to data sets with spatial components, has a rich history of producing general-purpose tools. Kenneth Moreland and Hank Childs discuss the innovative work that contributes to scientific visualization software and highlight specific aspects that researchers have translated into usable software for domain scientists.



9 The Institute for Mathematical and Statistical Innovation: An Introduction
The Institute for Mathematical and Statistical Innovation (IMSI) is the newest member of the portfolio of institutes that are funded by the National Science Foundation's Division of Mathematical Sciences. Kevin Corlette, Douglas Simpson, and Panagiotis Souganidis explain IMSI's mission, research objectives, themes, various forms of scientific activity, and emphasis on communication across different research communities.

9 Understanding the New Baseball Statistics
James Case reviews *Smart Baseball* by sportswriter Keith Law, which addresses the quantitative revolution that has swept baseball front offices in recent years. Law explores the inadequacies of traditional baseball statistics, overviews the wide variety of new statistics that are now available due to online databases, and describes how these novel statistics can guide player procurement.

11 SIAM Engages Participants at the 2020 Virtual SciFest
The SIAM Education Committee organized the SIAM booth at the 2020 USA Science and Engineering Festival, which commenced as an all-virtual event this year. Wesley Hamilton and Padmanabhan Seshaiyer recap SIAM's past involvement in the festival and reveal how the online booth engaged visitors and conveyed the importance of mathematics in everyday life.

11 Professional Opportunities

MathJax: The Present and the Future

By Davide P. Cervone and Volker Sorge

At the Joint Mathematics Meetings in January 2010, we announced version 1.0 of the MathJax JavaScript library for accurately typesetting mathematics in web pages. Almost immediately, MathJax became the de facto standard for formatting mathematics on the web. A wide range of online journals, including the American Mathematical Society's (AMS) MathSciNet website and SIAM News Online, employ MathJax. It allows online blogs, wikis, and question-and-answer sites—such as StackExchange and Wikipedia—to include mathematical expressions; provides the mathematics for online homework systems like WeBWorK and learning management systems like Moodle; and is incorporated into e-book readers, screen readers, and other similar products.[1]

In the 10 years since its introduction, MathJax has expanded to include LaTeX, MathML, and AsciiMath input formats,[2] as well as output formats[3] like HTML-with-CSS, Scalable Vector Graphics (SVG), and MathML. The introduction of sophisticated semantic enrichment and speech-generation functionality[4] in June 2016 made MathJax a crucial component in the creation of web pages and e-books that are accessible to blind or low-vision readers who utilize assistive technology, such as screen readers or Braille output devices. In short, MathJax increases the accessibility of online course materials and published research, which is even more valuable in the age of distance learning.

Much has changed in web technology since MathJax's introduction. New web libraries and improvements to the JavaScript language have altered the way in which web-page designers wish to use MathJax, and some of the initial approaches that were built into the library are not compatible with modern workflows. For the last three years, MathJax has been undergoing a complete rewrite with the goal of modernizing its internal infrastructure, increasing its flexibility for use with contemporary web technologies, facilitating its use for preprocessing and server-side support, and expediting the production of typeset mathematics. The release of MathJax version 3.0[5] in August 2019 brought these hopes to fruition.

Version 3 is written in the TypeScript language,[6] a form of JavaScript in which one can add types to variables and functions that a compiler then checks for correctness. This helps identify errors earlier in the process, yields more reliable code that users can easily understand, and makes it easier for others to contribute to MathJax. Version 3 also lets users exploit new features of JavaScript that are part of the latest ES6 standard,[7] while still supporting older browsers that have not implemented these features. For example, we now use ES6's modern class structure[8] and take advantage of asynchronous features like promises.[9] Because these options were not available 10 years ago, v2 implemented its own object system and provided custom signals, queues, and callbacks for asynchronous operation.

Version 3's new internal structure and modern ES6 features—which today's JavaScript interpreters can exploit for runtime optimization—remove performance issues that were inherent in the design of v2, thus improving MathJax's rendering speed. Although it is difficult to make precise comparisons because the two versions operate so differently, tests that render a complete page with several hundred expressions see a reduction of between 60 and 80 percent in rendering time (depending on the browser and operating system).

MathJax v2 used its own loading mechanism for accessing its components, which did not work well with modern JavaScript packaging systems like webpack[10] or Rollup.[11] Version 3 resolves this problem and interoperates more effectively with modern web workflows; for instance, users can make custom single-file builds of MathJax or include it as one component of a larger asset file.

New in version 3 is the ability to run MathJax synchronously; this was not possible in v2, as its operation is inherently asynchronous. In particular, v3 provides functions that can translate an input string (a TeX expression, for example) into an output Document Object Model tree (such as an SVG image).[12] One can apply these functions to individual expressions or entire documents. This is especially important when preparing rendered pages for offline consumption.

MathJax was originally designed for use in a web browser, which left the desire to pre-process mathematics on a server unaddressed. Version 3, which was redesigned to make this possible, can be used within node applications[13] in essentially the same way as in a browser.[14] That is, one can load MathJax components, configure them through the MathJax global variable, and call the same functions for typesetting and conversion as in a browser. This simplifies parallel development for both the browser and the server. Moreover, node applications can access MathJax modules directly (without the packaging needed for MathJax's use in a browser). This provides the most direct access to MathJax's features and the most flexibility in controlling MathJax's actions.

MathJax makes mathematics accessible to readers with disabilities, which distinguishes the library from other math-rendering solutions. While its initial goal was to work with the third-party assistive technology software that was available at the time, all iterations of MathJax since version 2.7 have offered their own accessibility solutions that are meant to operate independently of the user's browser, operating system, and assistive technology solutions (e.g., screen readers). The extension provides support for readers with functional needs, such as blindness, low vision, dyslexia, or dyscalculia.

The accessibility extension uses the speech rule engine library[15] to translate mathematical expressions into speech strings. Its core components comprise the automatic voicing of formulas together with interactive navigation and synchronized highlighting, as well as an abstraction feature that enables visual simplification and summary speech description of formulas.

MathJax v3 greatly extends these capabilities, which make use of its new modular setup to enable a flexible pick-and-mix personalization of accessibility tools. In particular, MathJax now offers a number of different rule sets for voicing mathematics that one can pre-select or switch on the fly. For instance, readers can choose a different rule set while interactively exploring an expression; the expression is thus spoken in a different way, which provides a new view of it. Version 3 also improves low-vision support by enabling magnification, not only for entire formulas but also for sub-expressions. It even provides better accessibility to advanced mathematical material by exploiting information from the original LaTeX code to generate more appropriate speech for different areas of mathematics, as well as subjects like physics, chemistry, and logic.

MathJax was originally meant to be a stop-gap measure, until browsers implemented native math rendering through MathML. Yet after more than 10 years, browser support for MathML is still not universal, and MathJax continues to bring quality math typesetting to all modern browsers. Our work has been supported by grants from the Sloan Foundation and the Simons Foundation, as well as generous contributions from our sponsors, including the AMS, SIAM, Elsevier, the Institute of Electrical and Electronics Engineers, and a variety of professional societies, publishers, and websites.[16] Without their ongoing financial support, MathJax would not be possible. The version 3 rewrite puts MathJax in a strong position to continue making beautiful and accessible mathematics available on the web for years to come.

A similar version of this article also appeared in the November issue of Notices of the American Mathematical Society.

Davide P. Cervone is a professor of mathematics at Union College, where he has taught for nearly 25 years. He has been the lead developer for MathJax since its inception in 2008. Volker Sorge is a professor of document analysis and accessibility in the School of Computer Science at the University of Birmingham. He has served as the Tech Lead for Accessibility of the MathJax Consortium since 2015.

[1] https://docs.mathjax.org/en/v2.7-latest/misc/mathjax-in-use.html
[2] https://docs.mathjax.org/en/latest
[3] https://docs.mathjax.org/en/latest/output/index.html
[4] https://docs.mathjax.org/en/latest/basic/accessibility.html
[5] https://www.mathjax.org/MathJax-v3-0-5-available
[6] https://www.typescriptlang.org
[7] http://ecma-international.org/ecma-262
[8] https://developer.mozilla.org/Web/JavaScript/Reference/Classes
[9] https://developer.mozilla.org/Web/JavaScript/Reference/Global_Objects/Promise
[10] https://webpack.js.org
[11] https://rollupjs.org/guide/en
[12] https://docs.mathjax.org/en/latest/web/typeset.html
[13] https://github.com/mathjax/MathJax-demos-node#mathjax-demos-node
[14] https://github.com/mathjax/MathJax-demos-web#mathjax-demos-web
[15] https://speechruleengine.org
[16] https://www.mathjax.org/#sponsors

siam news

ISSN 1557–9573.

COVID-19 Intervention

Continued from page 1

will recall his/her contacts (ϕ_κ). This yields the following expression:

$$[flow_{in}] := \mathcal{K}(\kappa, T_S, \phi_\kappa)[q_{tE}(t)E_1(t) + (q_{rA}(t) + q_{tA}(t))A_0(t) + (q_{rI}(t) + q_{tI}(t) + \tau_{IH})I_0(t)].$$

The formulation models the rate of increase for the contact tracers' backlog. This leaves us to estimate the rate that is needed to emerge from the backlog, which reflects the competing demands for contact tracers' time.

We begin by emphasizing that the time for contact monitoring (w_m) is independent of the time that is necessary to investigate new cases and their contacts (w_c). Again, we leverage the state variables to derive the contact tracers' total amount of work:

– Total work required to monitor known cases: $w_m(A_1(t) + I_1(t))$

– Total work required to investigate new cases: $w_c C(t)$.

The amount of work that contact tracers accomplish per day will be capped by the available resources and time. Let N_{trace} represent the number of contact tracers who are available to execute the tasks, and q_w be the fraction of a day that consists of the work hours of each tracer.

When enough resources are available to execute the tasks, we expect $work_{applied} \approx work_{demand}$ until we hit our carrying capacity, at which point $work_{applied} \approx q_w N_{trace}$. The bounded exponential function provides a smooth approximation for the relationship between $work_{applied}$ and $work_{demand}$ (see Figure 2). When the demand on work is less than the available resources, the bounded exponential simulates $work_{applied} > work_{demand}$. Though performing more work than demanded depletes the backlog at a higher-than-nominal rate, it does not ultimately overshoot the actual case queue. Recall that the goal is to determine the rate at which contact tracers can investigate new cases, so we further constrain the formulation to only reflect the proportion of work that tracers conduct for new cases:

$$\frac{w_c C(t)}{[w_m(A_1(t) + I_1(t)) + w_c C(t)]} \left[q_w N_{trace}\left(1 - \exp\left(-0.05 \left(w_m\left(A_1(t) + I_1(t)\right) + w_c C(t)\right)\right)\right)\right].$$

To obtain the rate at which tracers remove individuals from the backlog of contacts, we divide by the average workload per contact. Therefore,

$$[flow_{out}] := \frac{C(t)}{[w_m(A_1(t) + I_1(t)) + w_c C(t)]} \left[q_w N_{trace}\left(1 - \exp\left(-0.05 \left(w_m\left(A_1(t) + I_1(t)\right) + w_c C(t)\right)\right)\right)\right].$$

This notation provides a smooth approximation for the rate at which contact tracers can reach and isolate newly-identified COVID-19 cases and quarantine their contacts. It is constrained by the available resources and informed by the current prevalence of infections.

Application to Vaccine Distribution

The bounded exponential formulations also have practical implications in the distribution of vaccines within compartmental model dynamics. The initial vaccine supply will be limited when vaccines first become available, with an increasing rate of availability in the future. By using a monotonically increasing vaccine distribution function $v(t)$, we can formulate the continuously increasing nature of vaccine availability with linear approximations. Since the function is independent of population size, it will be difficult to bound the amount of distributed vaccines when the available vaccines begin to exceed the size of the eligible population.

If we utilize the bounded exponential function, we can approximate the fraction of distributed vaccines, $f_v(t)$, as a function of the current population that is eligible for vaccination. We assume that $S_0(t)$ is the model's only compartment to be vaccinated. If $S_0(t) < v(t)$, we reduce the vaccine distribution to reflect the fraction of the available population with respect to the total amount of available vaccines, $\frac{S_0(t)}{v(t)}$. Doing so results in an expression for the fraction of distributed vaccines:

$$f_v(t) := 1 - \exp\left(-\alpha \frac{S_0(t)}{v(t)}\right).$$

For an appropriate $\alpha > 0$, the product $f_v(t)v(t)$ must have the following properties:

– If the population is greater than the number of available vaccines, we will simulate distribution for all available vaccines.

– If the amount of vaccines exceeds the total population to which they can be distributed, the simulated distribution will only reflect a fraction of the available vaccines.

In closing, we offer a method for incorporating smooth mathematical formulations into the traditional SEIR compartmental modeling structure to simulate constraints on resource allocation. These formulations will provide a framework for the application of methods of dynamically constrained optimization problems in support of optimal resource allocation for a COVID-19 vaccine.

The Adaptive Recovery Modeling Team at Sandia National Laboratories brought together multidisciplinary researchers to provide their unique and experienced perspectives to the challenges of modeling the novel COVID-19 pandemic. In discussions with the New Mexico Department of Health, the practical aspects that public health professionals employ in response to intervention strategy policies have been used to assess baseline model formulations.

This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

Acknowledgments: Research was supported by the DOE Office of Science through the National Virtual Biotechnology Laboratory, a consortium of DOE national laboratories focused on response to COVID-19, with funding provided by the Coronavirus CARES Act. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. User acknowledges that neither the Government nor operating contractors of the above national laboratories makes any warranty, express or implied, of either the accuracy or completeness of this information, or assumes any liability or responsibility for the use of this information.

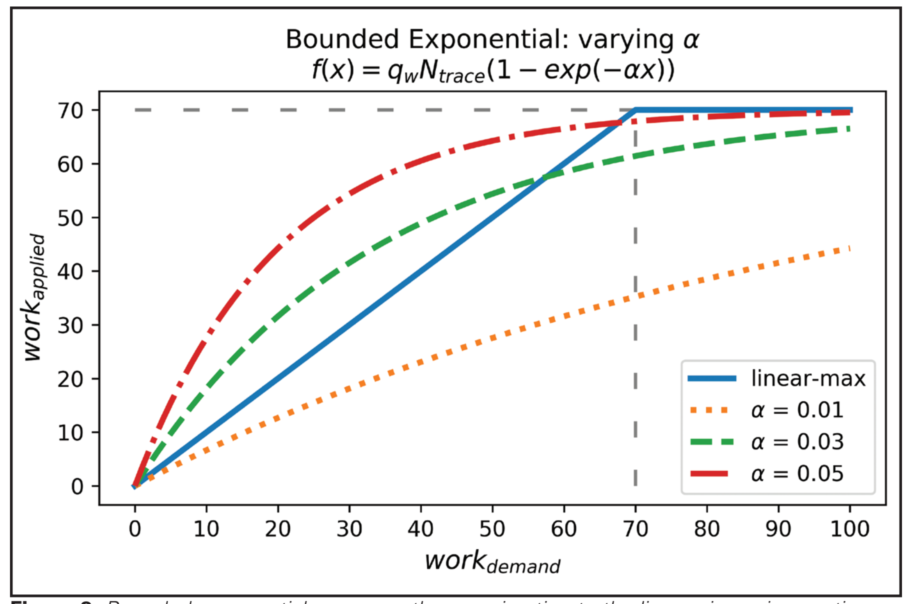


Figure 2. *Bounded exponential as a smooth approximation to the linear piecewise continuous relationship between $work_{demand}$ and $work_{applied}$. Figure courtesy of the authors.*

References

[1] Centers for Disease Control and Prevention. National Center for Immunization and Respiratory Diseases, Division of Viral Diseases. (2020, April 29). *Case Investigation and Contact Tracing: Part of a Multipronged Approach to Fight the COVID-19 Pandemic*. Retrieved from: https://www.cdc.gov/coronavirus/2019-ncov/php/principles-contact-tracing.html.

[2] European Centre for Disease Prevention and Control (2020, March 2). *Resource estimation for contact tracing, quarantine and monitoring activities for COVID-19 cases in the EU/EEA*. Stockholm: ECDC.

[3] Flaxman, S., Mishra, S., Gandy, A., Unwin, H.J.T., Mellan, T.A., Coupland, H., ..., Monod, M. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature, 584*(7820), 257-261.

[4] Gottlieb, S., Rivers, C., McClellan, M.B., Silvis, L., & Watson, C. (2020, March 28). *National coronavirus response: a road map to reopening*. American Enterprise Institute.

[5] Hethcote, H.W. (2000). The mathematics of infectious diseases. *SIAM Rev., 42*(4), 599-653.

[6] Heymann, D.L. (Ed.). (2015). Response to an Outbreak. *Control of Communicable Diseases Manual, 20th Edition* (pp. A8-A13). Washington, D.C.: APHA Press.

[7] Siddarth, D., & Weyl, E.G. (2020). Why We Must Test Millions a Day. *COVID-19 Rapid Response Impact Initiative* (White Paper 6). Cambridge, MA: Harvard University.

Erin C.S. Acquesta of Sandia National Laboratories (SNL) has experience with analyzing the mathematical properties and controllability of dynamical systems that are designed to simulate the spread of infectious diseases. Her broader area of research focuses on the application of methods of formal mathematical analyses to provide evidence of model credibility. Walt Beyeler of SNL has experience developing models of complex systems for the simulation of disease transmission, control, and resource utilization. Application contexts include evaluating strategies for pandemic influenza, detecting and countering epidemics in herd animals in Afghanistan, improving medical laboratory systems to speed the diagnosis of Ebola patients in West Africa, and using medical logistics modeling to anticipate and resolve resource shortfalls that arise from the COVID-19 pandemic. Pat Finley leads biosurveillence and disease modeling efforts at SNL, focusing on rapid-response operational models for developing outbreaks. His current research involves machine learning approaches to predict zoonotic transitions of emergent pathogens in developing countries. Katherine Klise of SNL applies her research experience in infrastructure resilience, sensor placement optimization, and data analytics in the context of water distribution systems, electricity grids, fossil energy, and renewable energy systems to the needs of modeling complex architectures for the simulation of disease transmission. Monear Makvandi of SNL is an infectious disease epidemiologist with experience in emerging/re-emerging infectious disease surveillance and outbreak response. At SNL, she leads the development and implementation of strategic and sustainable biological risk reduction through enhanced laboratory security, improved diagnostic methods, infectious disease control, and health system strengthening. Emma Stanislawski is an epidemiologist with the New Mexico Department of Health, where she specializes in respiratory infectious diseases. She has previously worked as a vaccine-preventable disease epidemiologist and with an immunization advocacy coalition.

Invasive Populations

Continued from page 1

in $\Lambda = (0, L) \times (0, T)$. Here, $N = N(x,t)$ represents population density in the river, N_t and N_x are derivatives of N with respect to t and x respectively, $A(x,t)$ is the cross-sectional area of the river, $D(x,t)$ is the diffusion coefficient that may include biodiffusion and flow-driven diffusion, r is the intrinsic growth rate, K is the carrying capacity, and $Q(t)$ represents the water discharge rate. We assume that there is no population at the upstream boundary $(x = 0)$, as the modeled invasive species is moving from downstream to upstream. However, we do assume a free-flow condition downstream at $x = L$, meaning that an outside source is not increasing the population.

In a 1D river with a constant bottom slope or channels that alternate between riffles (shallow areas) and pools, the water discharge per unit width decreases the upstream spreading speed when it is low but increases this speed when it is high [3]. If the flow fluctuates between high and low, a longer high flow season makes it harder for the invasive population to spread upstream.

When investigating the interplay of spatial heterogeneity and temporal fluctuation, an interesting phenomenon can arise from hybrid hydrodynamic-biological models in 1D and two-dimensional (2D) rivers [3]. We call this an *invasion ratchet*, wherein a species can persist in a favorable patch during adverse times and traverse unfavorable patches in the upstream direction during opportune times. In the long term, this type of phenomenon can ensure a population's upstream spread and consequential persistence in the entire river. Figure 1 depicts an invasion ratchet phenomenon in a 2D (longitudinal-lateral) meandering river. The population spreads upstream when the flow is low and retreats when the flow is high, though it spreads both upstream and downstream in the long run. Scientists have calculated the steady-state water flow with a numerical model called "River2D" [7]; one can then implement the population model—the 2D version of (1)—into River2D's water flow.

An existing conjecture states that "a population can persist at any location in a homogeneous habitat if and only if it can invade upstream" [4]. This conjecture has been verified for some models with temporally-varying flows or spatially-homogeneous habitats — but not necessarily with both [2, 3]. If the time of low discharge in a pool-riffle river with fluctuating flows is not long enough for the population to advance from the riffle to the next upstream pool, the population is then washed back to its foothold in the downstream pool, where it remains until the next low discharge time [3]. The population therefore stalls in the river but cannot spread further upstream, which indicates that the aforementioned conjecture's assumption of homogeneity in space or time is essential [4].

We have now established that the water discharge rate, which is higher during certain seasons and lower during others, affects the population level at various locations. In addition, the water discharge rates and lengths of different seasons also play important roles in the long-term population persistence or invasion in rivers. To investigate the control of water discharge rate as a potential management strategy, one can use optimal control of the water discharge rate $Q(t)$ on model (1) to force the invasive population downstream [5]. The objective functional is to minimize the population's integral (over time and space) with an upstream weight multiplier function to keep the invaders downstream, and with a small cost term to implement control. Thus, the term with the upstream weight dominates the objective. Positive upper and lower bounds on the controls (flow rates) are included in the optimization.

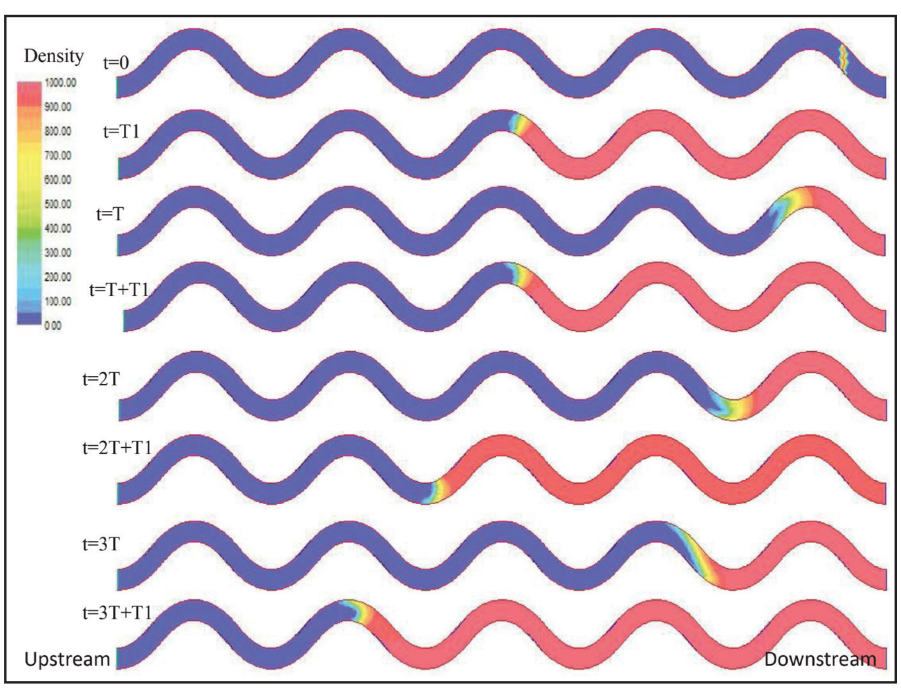


Figure 1. *An invasion ratchet in a meandering river. The colors represent the population density at different locations in the river at different times. The parameters are as follows: flow period is $T = 365$ days, low flow season length is $T_1 = 270$ days, low flow is $1\,\mathrm{m}^3/\mathrm{s}$, high flow season length is $T_2 = 95$ days, and high flow is $20\,\mathrm{m}^3/\mathrm{s}$. Image courtesy of [3].*

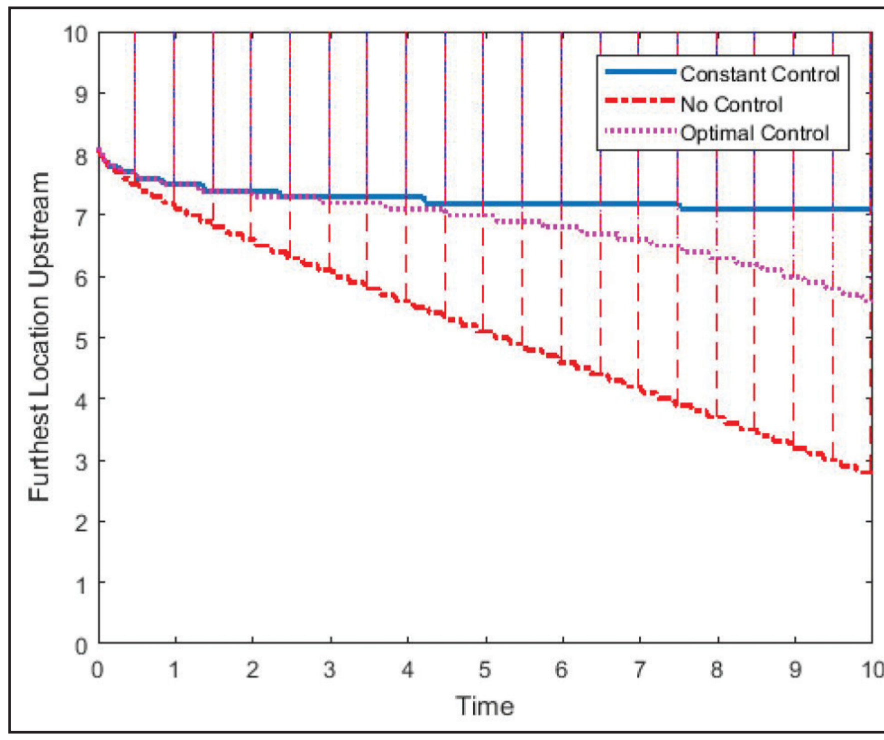


Figure 2. *Location of the population in the river for the no control case (red), the constant control case with Q at its upper bound (blue), and the optimal control case $Q^*(t)$ (magenta) over time. On the y-axis, 0 represents upstream and 10 represents downstream. Image courtesy of [5].*

Figure 2 illustrates the distance of the population's upstream motion with no control, constant control, and optimal control. Using a detection threshold of 0.5 for the population each time, we find the lowest location x (more upstream) with $N(x,t) > 0.5$ and plot that river location. Constant control takes the value of the controls' upper bound. As expected, the population with no control moves further upstream than the population with constant control. Maintaining constant control at the upper bound resulted in the least movement upstream; the population moves further upstream with optimal control than with constant control. Note that optimal control is not at its upper bound for the whole time period due to costs in the objective functional, which balances costs with the desire to keep the population downstream. Our objective functional values for optimal control and constant control cases respectively yield 83 and 74 percent improvements over the value with no control.

A hybrid modeling approach allows us to investigate the impact of river morphology and flow patterns on a population's spatio-temporal dynamics. Optimal control theory can generate strategies for regulating interested parameters to successfully manage an invasive population in a river. Applying the hybrid model and optimal control strategy to a particular invasive species could help scientists design reasonable management plans to control an invasion. For example, we could apply our model and methods to the invasive zebra mussel (*Dreissena polymorpha*), for which researchers have used a homogeneous version of model (1) [6] and a coupled continuous-discrete model [1] to investigate the species' persistence and invasion in a temporally homogeneous habitat. It would be interesting to see how temporal and spatial heterogeneities affect mussel resilience and influence decisions about the invasion's management.

References

[1] Huang, Q., Wang, H., & Lewis, M.A. (2017). A hybrid continuous/discrete-time model for invasion dynamics of zebra mussels in rivers. *SIAM J. Appl. Math.*, *77*, 854-880.

[2] Jin, Y., & Lewis, M.A. (2011). Seasonal influences on population spread and persistence in streams: Critical domain size. *SIAM J. Appl. Math.*, *71*, 1241-1262.

[3] Jin, Y., Hilker, F.M., Steffler, P.M., & Lewis, M.A. (2014). Seasonal invasion dynamics in a spatially heterogeneous river with fluctuating flows. *Bull. Math. Biol.*, *76*, 1522-1565.

[4] Lutscher, F., Nisbet, R.M., & Pachepsky, E. (2010). Population persistence in the face of advection. *Theoret. Ecol.*, *3*, 271-284.

[5] Pettit, R., & Lenhart, S. (2019). Optimal control of a PDE model of an invasive species in a river. *Math.*, *7*(10), 975.

[6] Speirs, D.C., & Gurney, W.S.C. (2001). Population persistence in rivers and estuaries. *Ecol.*, *82*, 1219-1237.

[7] Steffler, P.M., & Blackburn, J. (2002). *River2D, two-dimensional depth averaged model of river hydrodynamics and fish habitat, introduction to depth averaged modeling and user's manual*. Alberta, Canada: University of Alberta.

Yu Jin is an associate professor in the Department of Mathematics at the University of Nebraska-Lincoln. Her research is devoted to mathematical biology, as well as dynamical systems and differential equations. Suzanne Lenhart is a Chancellor's Professor in the Department of Mathematics at the University of Tennessee. Her research involves partial differential equations, ordinary differential equations, optimal control, and biological models.

From Magnets to Melt Ponds

By Kenneth M. Golden, Yiping Ma, Courtenay Strong, and Ivan Sudakov

When the snow on top of Arctic sea ice begins to melt in late spring, small pools of water form on the surface. As the melt season progresses, these simply shaped meter-scale pools grow and coalesce into kilometer-scale labyrinths of cerulean blue with complex, self-similar boundaries. The fractal dimension of these boundaries transitions from one to roughly two as the area increases through a critical regime that is centered around 100 square meters [4]. While the white, snowy surface of the sea ice reflects most of the incident sunlight, the darker melt ponds act like windows and allow significant light to penetrate the ice and seawater underneath. Melt ponds thus help control the amount of solar energy that the ice pack and upper ocean absorb, strongly influencing ice melting rates and the ecology of the polar marine environment. They largely determine sea ice albedo—the ratio of reflected to incident sunlight—which is a key parameter in climate modeling.

When viewed from a helicopter, the beautiful patterns of dark and light on the surface of melting sea ice are reminiscent of structures that applied mathematicians sometimes see when studying phase transitions and coarsening processes in materials science. They also resemble the complex regions of aligned spins, or *magnetic domains*, that are visible in magnetic materials. Figure 1 compares two examples of magnetic domains with similar patterns that are formed by melt ponds on Arctic sea ice. Magnetic energy is lowered when nearby spins align with each other, which produces the domains. At higher temperatures, thermal fluctuations dominate the tendency of the domains' magnetic moments to also align, with no net magnetization M of the material unless one applies an external magnetic field H to induce alignment. However, the tendency for overall alignment takes over at temperatures below the Curie point T_c, and the material remains magnetized even as the applied field H vanishes, where the remaining non-zero magnetization $(M \neq 0)$ is called *spontaneous* or *residual*.

The prototypical model of a magnetic material based on a lattice of interacting binary spins is the Ising model, which was proposed in 1920 by Ernst Ising's Ph.D. advisor Wilhelm Lenz. This model incorporates only the most basic physics of magnetic materials and operates on the principle that natural systems tend toward minimum energy states.

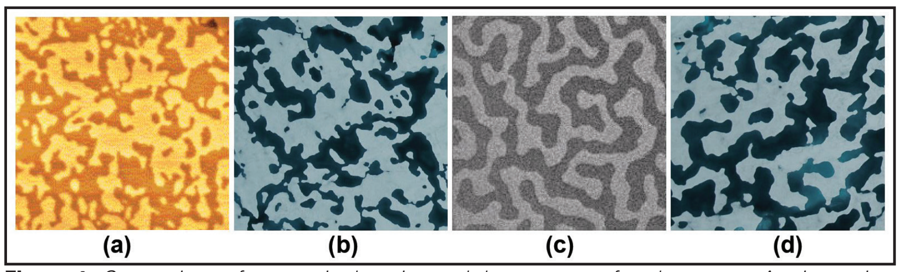


Figure 1. *Comparison of magnetic domains and the patterns of meltwater on Arctic sea ice.* **1a.** *Magnetic domains in cobalt, roughly 20 microns across.* **1b.** *Arctic melt pond, roughly 100 meters across.* **1c.** *Magneto-optic Kerr effect microscope image of maze-like domain structures in thin films of cobalt-iron-boron, roughly 150 microns across.* **1d.** *Similarly-structured melt ponds, roughly 70 meters across. Figure 1a courtesy of [9], 1b and 1d courtesy of Donald Perovich, 1c courtesy of [10].*

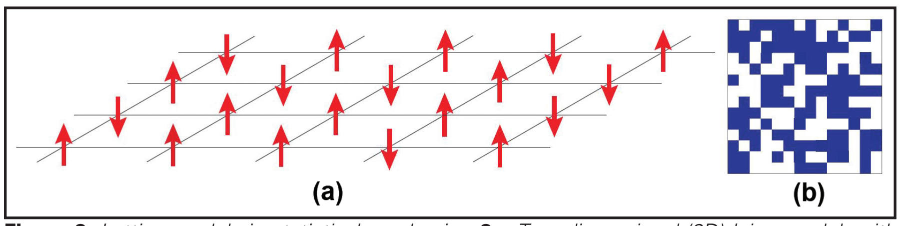


Figure 2. *Lattice models in statistical mechanics.* **2a.** *Two-dimensional (2D) Ising model, with spins either up or down at each lattice site.* **2b.** *Spin configuration. Spin-up sites are blue and spin-down sites are white. Image courtesy of Kenneth Golden.*

Consider a finite box $\Lambda \subset \mathbb{Z}^2$ that contains N sites. At each site, a spin variable s_i can take the values $+1$ or -1 (see Figure 2). To illustrate our melt pond Ising model, we formulate the problem of finding the magnetization $M(T,H)$—or *order parameter*—of an Ising ferromagnet at temperature T in field H. The Hamiltonian $\mathcal{H}$ with ferromagnetic interaction $J \geq 0$ between nearest neighbor pairs is given by

$$\mathcal{H}_\omega = -H\sum_i s_i - J\sum_{<i,j>} s_i s_j$$

for any configuration $\omega \in \Omega = \{-1,1\}^N$ of the spin variables. The canonical partition function Z_N, which yields the system's observables, is given by

$$Z_N(T,H) = \sum_{\omega\in\Omega} \exp(-\beta\mathcal{H}_\omega) = \exp(-\beta N f_N),$$

where $\beta = 1/kT$, k is Boltzmann's constant, $\exp(-\beta\mathcal{H}_\omega)$ is the Gibbs factor, and f_N is the free energy per site: $f_N(T,H) = (-1/\beta N)\log Z_N(T,H)$.

The magnetization $M(T,H) = \lim_{N\to\infty} \frac{1}{N}\sum_i s_i$ is averaged over $\omega \in \Omega$ with Gibbs' weights and expressed in terms of the free energy $f(T,H) = \lim_{N\to\infty} f_N(T,H)$:

$$M(T,H) = -\frac{\partial f}{\partial H}.$$

The model's rich behavior is exemplified in the existence of a critical temperature T_c—the Curie point—where $M(T) = \lim_{H\to 0} M(T,H) > 0$ for $T < T_c$ and $M(T) = 0$ for $T \geq T_c$. Universal power law asymptotics for $M(T) \to 0$ as $T \to T_c^-$ are independent of the lattice type and other local details.

The Metropolis algorithm is a common method for numerically constructing equilibrium states of the Ising ferromagnet. In this approach, a randomly-chosen spin either flips or does not flip based on which action lowers or raises the energy. ΔE represents the change in magnetostatic energy from a potential flip (as measured by $\mathcal{H}_\omega$), and the spin is flipped if $\Delta E \leq 0$. If $\Delta E > 0$, the probability of the spin flipping is given by the Gibbs factor for ΔE. Sweeping through the whole lattice and iterating the process many times attains a local minimum in the system's energy.

We have adapted the classical Ising model to study and explain the observed geometry of melt pond configurations and capture the fundamental physical mechanism of pattern formation in melt ponds on Arctic sea ice [5]. While previous studies have developed important and instructive numerical models of melt pond evolution [2, 3],

See **Melt Ponds** *on page 7*

Scientific Visualization: New Techniques in Production Software

By Kenneth Moreland and Hank Childs

The field of visualization encompasses a wide range of techniques, from infographics to isosurfaces. An important subfield called scientific visualization is specifically dedicated to data sets with spatial components, i.e., (X, Y, Z) locations. This subfield's name is inspired by the fact that the data in question often come from the sciences, like physics simulations or sensor networks.

Scientific visualization has a rich history of producing various general-purpose tools, including ParaView, VisIt, SciRun, Tecplot, FieldView, and EnSight. These tools allow the efforts of relatively few developers to impact numerous stakeholders, with millions of downloads and/or licenses. Furthermore, many other tools—such as MegaMol, for the visualization of molecular dynamics—are dedicated to specific scientific domains or data.

Scientific visualization "tool developers" work closely with a significant research community that regularly generates fundamentally new techniques and improvements for existing methods. Unfortunately, these research works often yield prototypes that domain scientists cannot use. However, when such results are shown to be effective, they are ultimately productized and adopted in scientific visualization tools.

Here we discuss the innovative research that experts have recently integrated into scientific visualization software. Specifically, we aim to highlight some of the top directions from the scientific visualization research community that have been translated into usable software for domain scientists. For the sake of brevity, we assume that readers are already familiar with traditional features, such as contouring, pseudocoloring, glyphing, flow tracing, and clipping.

Topology

Recent work in visualization has focused on the adoption of topological methods for data analysis. These analysis methods can operate on large, mesh-based data structures that are commonly used in science simulations. For example, researchers can utilize topological structures—including Morse-Smale complexes and Reeb graphs—for operations like feature tracking, similarity estimation, and segmentation. These techniques are applicable in numerous fields, such as combustion, materials science, chemistry, and astrophysics [4].

The Topology ToolKit[1] (TTK) has increased the accessibility of topological data analysis. TTK is a library with many topological analysis functions that can integrate into other software. It also provides plugins that assimilate its methods into existing software tools (see Figure 1).

[1] https://topology-tool-kit.github.io

Advanced Flow

Recent advances in parallel hardware and visualization software facilitate the practicality of increasingly more flow visualization techniques. Traditionally, animating particle motion and plotting particle trajectories (i.e., streamlines or pathlines) have been the most common flow visualization procedures. However, recent advances have enabled the calculation of many more particle trajectories that inspire new types of flow visualizations.

Finite-time Lyapunov exponents (FTLE) comprise a noteworthy flow visualization technique that is derived from the tracing of many particle trajectories. This method produces a new scalar field that measures flow separation. To successfully operate, it considers neighborhoods and places multiple particles within a neighborhood. If the particles separate significantly, the scalar field for the neighborhood in question is assigned a high value; if the particles remain close together, the corresponding scalar field is assigned a low value. Figure 2 depicts an example of an FTLE.

SOFTWARE AND PROGRAMMING

Many additional flow techniques use particle trajectories as a foundational step. For instance, stream surfaces operate by seeding particles along a line (or curve) and constructing a surface from the resulting trajectories. Poincaré analysis considers topological structures that form when a particle repeatedly circulates through a volume. Other techniques illuminate underlying Lagrangian coherent structures.

Ray Tracing

Three-dimensional (3D) rendering has remained a staple of scientific visualization ever since graphics hardware became available. However, most rendering in scientific visualization utilizes approximations by independently drawing each small piece. This practice misses "global" effects like shadows, reflections, and diffuse lighting conditions.

Another approach to 3D rendering is ray tracing, which traces the path of light as it bounces off objects (see Figure 3, on page 8). Recent improvements in ray tracing software—as well as increases in computational power—make interactive scientific visualization practical. Intel's OSPRay[2] and NVIDIA's OptiX[3] are software libraries that provide quick and realistic ray-traced rendering. Several visualization tools, including ParaView,[4] VisIt,[5] and VMD[6] are integrating these new rendering capabilities.

[2] https://www.ospray.org

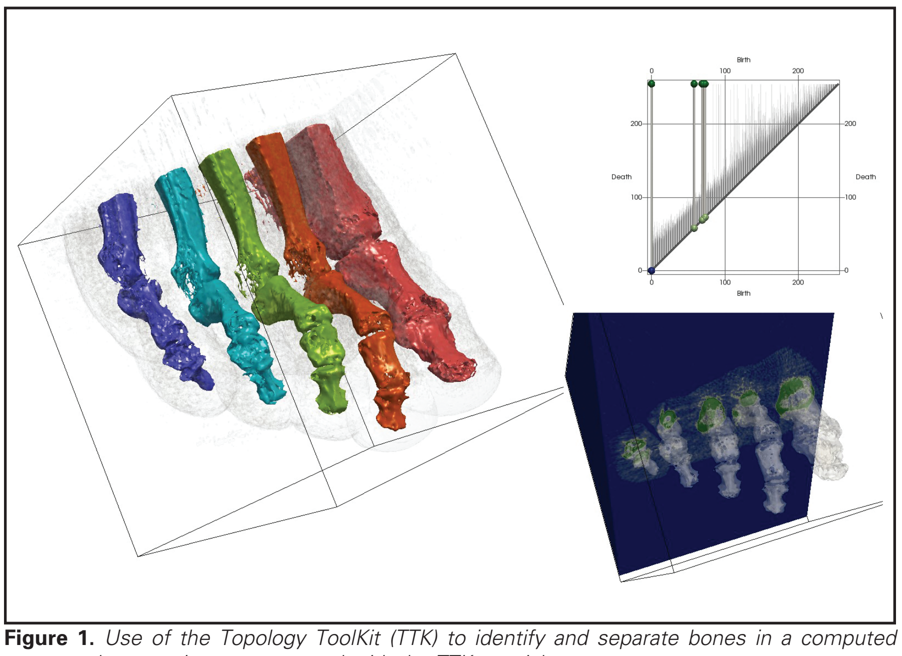


Figure 1. *Use of the Topology ToolKit (TTK) to identify and separate bones in a computed tomography scan. Image generated with the TTK tutorial.*

In Situ Visualization

An increasingly major bottleneck on large, parallel machines is the speed at which data can be written out; the fraction of data that can be written to disk storage is often unacceptably small. To bypass this problem, simulations are turning to *in situ* visualization [1, 2], during which the visualization is run as part of the simulation. In this process, data does not need to be written to disk storage.

Several libraries exist for *in situ* visualization of computational simulations. First, two major post hoc tools can deliver their capabilities in *in situ* form: ParaView provides Catalyst[7] and VisIt offers Libsim.[8] Furthermore, libraries are emerging that are devoted entirely to *in situ*. Ascent[9] is one such library, with foci on flyweight processing—application programming interface (API), memory usage, binary size, execution time—and support for modern supercomputers — central processing units, graphics processing units, and so forth via the VTK-m[10] library. SENSEI[11] is another

See **Scientific Visualization** *on page 8*

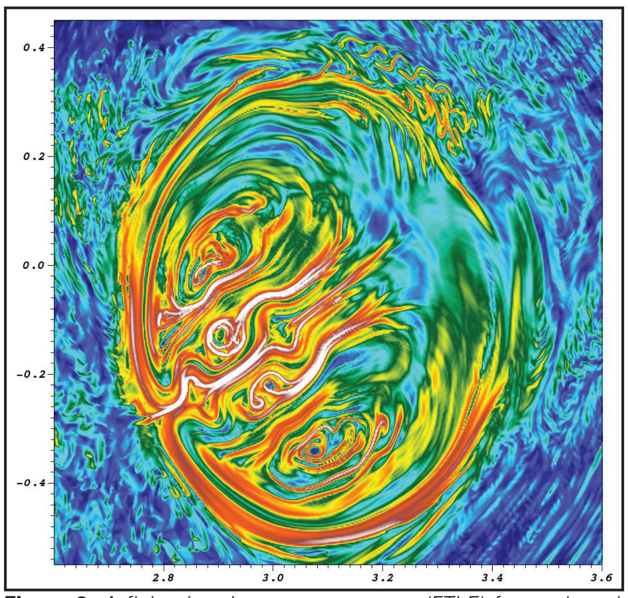


Figure 2. *A finite-time Lyapunov exponent (FTLE) for a tokamak simulation, which determines the separation of flow. Blue and cyan areas are less turbulent, whereas red and white areas are more turbulent. Data courtesy of Linda Sugiyama.*

[3] https://developer.nvidia.com/optix
[4] https://blog.kitware.com/virtual-tour-and-high-quality-visualization-with-paraview-5-6-ospray
[5] https://tacc.github.io/visitOSPRay
[6] https://www.ks.uiuc.edu/Research/vmd/vmd-1.9.3

[7] https://www.paraview.org/in-situ
[8] https://www.visitusers.org/index.php?title=Libsim_Batch
[9] https://ascent.readthedocs.io/en/latest
[10] http://m.vtk.org/index.php/Main_Page
[11] https://sensei-insitu.org

Melt Ponds

Continued from page 5

these models were somewhat detailed and did not focus on the way in which meltwater is distributed over the sea ice surface. Our new model is simplistic and accounts for only the system's most basic physics. In fact, the only measured parameter is the one-meter lattice spacing, which is determined by snow topography data.

The simulated ponds are metastable equilibria of our melt pond Ising model. They have geometrical characteristics that agree very closely with observed scaling of pond sizes [6] and the transition in pond fractal dimension [4]. Researchers have also developed continuum percolation models that reproduce these geometrical features [1, 8].

We aim to use our Ising model to introduce a predictive capability to cryosphere modeling based on ideas of statistical mechanics and energy minimization, utilizing just the essential physics of the system. The model consists of a two-dimensional lattice of N square patches, or pixels, of meltwater $(s_i = +1)$ or ice $(s_i = -1)$, which correspond to the spin-up or spin-down states in the Ising ferromagnet. Configurations $\omega \in \Omega = \{-1, 1\}^N$ of the spin field s_i represent the distribution of meltwater on the sea ice surface. Each patch interacts only with its nearest neighbors and is influenced by a forcing field. However, sea ice surface topography—which can vary from site to site and influence whether a patch comprises water or ice—plays the role of the applied field in our melt pond Ising model. Our model is then actually a *random field* Ising model, and one can write the Hamiltonian as

$$\mathcal{H}_\omega = -\sum_i (H - h_i) s_i - J \sum_{<i,j>} s_i s_j.$$

Here, h_i are the surface heights (taken to be independent Gaussian variables with mean zero) and H is a reference height (taken to be zero in the model's simplest form). The spin field s_i is reorganized to lower the free energy, and the order parameter is the pond area fraction $F = (M+1)/2$, which is directly related to sea ice albedo. We set temperature $T = 0$ and assume for simplicity that environmental noise does not significantly influence melt pond geometry.

Independent flips of a weighted coin determine the system's initial random configuration. A pixel or site has a probability p of its spin being $+1$, or meltwater. The system then updates based on simple rules: pick a random site i and update s_i as follows. If a majority exists among s_i's four nearest neighbors, we assume that heat diffusion drives s_i to agree with this majority. Otherwise we assume water's tendency to fill troughs, as determined by the local value of the random field h_i. This update step, which corresponds to energy minimization via Glauber spin flip dynamics, iterates until s_i becomes steady. The spin-up or meltwater clusters in the final configurations of the spin field s_i exhibit geometric characteristics that agree surprisingly well with observations of Arctic melt ponds (see Figure 3, on page 1). The final configuration is a metastable state — a local minimum of $\mathcal{H}_\omega$. As neighboring sites exchange heat, spins tend to align to minimize energy. In doing so, they coarsen away from the purely random initial state. The emergence of this order from disorder is a central theme in statistical physics and an attractive feature of our approach.

The ability to efficiently generate realistic pond spatial patterns may enable advances in how researchers account for melt ponds and many related physical and biological processes in global climate models (GCMs). Typical GCM grid spacing is tens to hundreds of kilometers, so melt ponds are subgrid-scale and thus too small to resolve on the model grid. Instead, GCMs use *parameterizations* to specify a pond fraction. Specifically, modern parameterizations in GCMs track a thermodynamically-driven meltwater volume and distribute it over the sea ice thickness classes that are present in a grid cell, beginning with the thinnest class since it presumably has the lowest ice height [3]. This yields a pond fraction F and a first-order approximation to sea ice albedo, $\alpha_{sea\ ice} = F\alpha_{water} + (1-F)\alpha_{snow}$, but does not address how the pond's area is organized spatially. Our simple model provides a framework for prescribing a subgrid-scale spatial organization whose realistic fractal dimension or area-perimeter relation could have important influences on pond evolution [7].

At this stage, total agreement between this simple model and the real world is too much to ask. The Ising model is unable to resolve features that are smaller than the lattice constant, and the metastable state also inherits certain unrealistic features from the purely random initial condition. Nonetheless, the model may be able to use more sophisticated rules to reproduce actual melt pond evolution. We anticipate that emerging techniques—like machine learning—will deduce such evolutionary rules from observational data.

References

[1] Bowen, B., Strong, C., & Golden, K.M. (2018). Modeling the fractal geometry of Arctic melt ponds using the level sets of random surfaces. *J. Fract. Geom.*, *5*, 121-142.

[2] Flocco, D., & Feltham, D.L. (2007). A continuum model of melt pond evolution on Arctic sea ice. *J. Geophys. Res.*, *112*, C08016.

[3] Flocco, D., Feltham, D.L., & Turner, A.K. (2010). Incorporation of a physically based melt pond scheme into the sea ice component of a climate model. *J. Geophys. Res.*, *115*(C8), C08012.

[4] Hohenegger, C., Alali, B., Steffen, K.R., Perovich, D.K., & Golden, K.M. (2012). Transition in the fractal geometry of Arctic melt ponds. *The Cryosphere*, *6*(5), 1157-1162.

[5] Ma, Y.P., Sudakov, I., Strong, C., & Golden, K.M. (2019). Ising model for melt ponds on Arctic sea ice. *New J. Phys.*, *21*, 063029.

[6] Perovich, D.K., Tucker, W.B., & Ligett K. (2002). Aerial observations of the evolution of ice surface conditions during summer. *J. Geophys. Res.*, *107*(C10), C000449.

[7] Polashenski, C., Perovich, D., & Courville, Z. (2012). The mechanisms of sea ice melt pond formation and evolution. *J. Geophys. Res.*, *117*(C1), C01001.

[8] Popović, P., Cael, B.B., Silber, M., & Abbot, D.S. (2018). Simple rules govern the patterns of Arctic sea ice melt ponds. *Phys. Rev. Lett.*, *120*(14), 148701.

[9] Portland Waldorf School. (2008). Electromagnetism. Retrieved from http://www.alchemical.org/em/index.html.

[10] Yamanouchi, M., Jander, A., Dhagat, P., Ikeda, S., Matsukura, F., & Ohno, H. (2011). Domain structure in CoFeB thin films with perpendicular magnetic anisotropy. *IEEE Magn. Lett.*, *2*, 3000304.

Kenneth M. Golden is a distinguished professor of mathematics and an adjunct professor of biomedical engineering at the University of Utah. His research is focused on developing mathematics of composite materials and statistical physics to model sea ice structures and processes. Yiping Ma is a Vice Chancellor's Fellow of Mathematics at Northumbria University. His research focuses on nonlinear dynamics and statistical physics with applications to diverse areas, including materials, optics, climate, and fluids. Courtenay Strong is an associate professor of atmospheric sciences at the University of Utah. A substantial component of his research focuses on modeling and analysis of the cryosphere, which includes sea ice and snow. Ivan Sudakov is an assistant professor in the Department of Physics at the University of Dayton and a Kavli Institute for Theoretical Physics Scholar. He specializes in data analysis and mathematical modeling for physical and living systems.

Scientific Visualization

Continued from page 6

example, with an emphasis on tool and method portability (i.e., providing an API that can access other *in situ* libraries) and proximity portability (i.e., able to run on the current resources or in transit resources).

Image Exploration

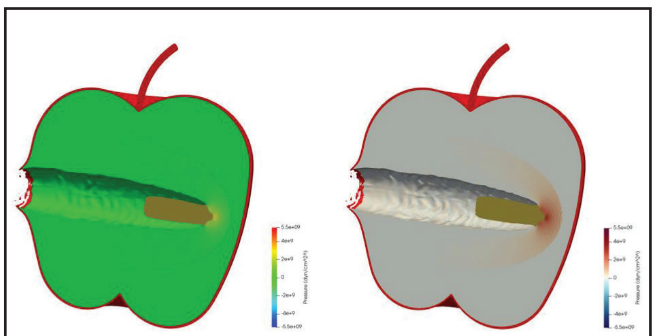
A challenge of *in situ* visualization is that users cannot change the visualization once it is generated. Incorrectly setting parameters can hide important phenomena from the camera, thereby causing these phenomena to be missed. One possible solution involves taking a scattershot approach that performs numerous visualizations of the same data, then organizing these visualization results into a navigable image database.

Cinema[12] is a community project that provides a specification for the organization of an assortment of visualization results. Cinema databases—which comprise image files and text metadata files—are simple to generate, and Ascent, Libsim, and Catalyst directly support their creation. One can then explore these databases via a web browser or programs that run on a desktop.

Color Perception

Most scientific visualization users will recognize the red, yellow, green, and blue colors from the rainbow that are painted on objects to represent data. These classic colors are derived from the natural physical properties of light and can yield some attractive images. Unfortunately, research in human perception suggests that the colors are not ideal for data representation [3]; human vision is complex and does not respond proportionally to changes in light intensity and wavelength. Consequently, use of these rainbow colors can obfuscate visualization data.

Recent work in visualization has built color map functions that are based on models of human perception of color. These new color maps better represent the data they encode (see Figure 4). Perceptual color maps are now easily accessible in many visualization tools.



(a) (b)

Figure 4. *These two images depict identical data wherein only the colors that map numbers vary.* **4a.** *The use of physical rainbow colors works to hide the pressure wave.* **4b.** *The pressure wave is easily visible in the perceptual colors. Data courtesy of Jason Wilke.*

Web Delivery

Most general-purpose scientific visualization programs require the installation of software on a user's computer. But recent years have seen an increased interest in web utilization as a deployment platform for scientific visualization tools. Libraries like VTK.js[13] and ParaViewWeb[14] adapt standard scientific visualization libraries to simplify the building of active web pages, thus enabling the creation of full-featured applications that run in a web browser. Tapestry[15] focuses on nimble delivery, including embedding in general web pages. The user directs new visualizations via web browser interactions; renderings on the cloud are then placed in the browser as if they were generated locally.

In conclusion, scientific visualization researchers have been active in forming newly discovered tools into usable software products. We hope that we have educated readers about recent improvements in scientific visualization that are useable today.

Acknowledgments

We wish to thank Chris Johnson (University of Utah) for his encouragement and support of this paper. We also thank Paul Navratil (Texas Advanced Computing Center) for providing example images. This research was supported by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy's (DOE) Office of Science and National Nuclear Security Administration (NNSA). Sandia National Laboratories is a multimission laboratory managed and operated by the National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the NNSA under contract DE-NA0003525. Any subjective views or opinions that might be expressed in this paper do not necessarily represent the views of the U.S. DOE or the U.S. government.

References

[1] Bennett, J.C., Childs, H., Garth, C., & Hentschel, B. (2019). In situ visualization for computational science (Dagstuhl seminar 18271). *Dagstuhl Reports*, *8*(7), 1-43.

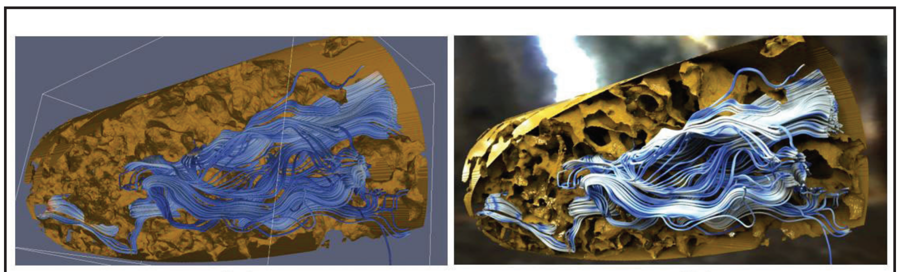
[2] Peterka, T., Bard, D., Bennett, J., Bethel, E.W., Oldfield, R., Pouchard, L., …, Wolf, M. (2019). *ASCR workshop on in situ data management: enabling scientific discovery from diverse data sources*. U.S. Department of Energy, Office of Science.

[3] Rogowitz, B., & Treinish, L. (1998). Data visualization: The end of the rainbow. *IEEE Spect.*, *35*(12), 52-59.

[4] Tierny, J., Favelier, G., Levine, J.A., Gueunet, C., & Michaux, M. (2018). The Topology ToolKit. *IEEE Trans. Visual. Comp. Graph.*, *24*, 832-842.

Kenneth Moreland received his Ph.D. in computer science from the University of New Mexico in 2004. He is currently a principal member of technical staff at Sandia National Laboratories, and his current interests include scientific visualization on large-scale and future-generation computer systems. Hank Childs received his Ph.D. in computer science from the University of California, Davis in 2006. He is currently a professor of computer and information science at the University of Oregon, and his research interests include scientific visualization, high-performance computing, and visualization software.

[12] https://cinemascience.github.io
[13] https://kitware.github.io/vtk-js
[14] https://www.paraview.org/web
[15] https://seelabutk.github.io/tapestry



(a) (b)

Figure 3. *Use of ray trace rendering to improve visualization of a Florida groundwater core sample.* **3a.** *Traditional raster image.* **3b.** *Ray cast image. The porous media is much more discernable in the ray cast image than the traditional raster image. Images courtesy of Paul Navratil and Carson Brownlee. Data courtesy of Michael Sukop, Sadé Garcia, and Kevin Cunningham.*

The Institute for Mathematical and Statistical Innovation: An Introduction

By Kevin Corlette, Douglas Simpson, and Panagiotis Souganidis

The Institute for Mathematical and Statistical Innovation[1] (IMSI) is a new National Science Foundation (NSF)-funded mathematical sciences research institute that is located at the University of Chicago and managed in partnership with Northwestern University, the University of Illinois at Chicago, and the University of Illinois at Urbana-Champaign. It is the newest member of the portfolio of institutes that are funded by NSF's Division of Mathematical Sciences, joining the American Institute of Mathematics, School of Mathematics at the Institute for Advanced Study, Institute for Computational and Experimental Research in Mathematics, Institute for Pure and Applied Mathematics, Mathematical Sciences Research Institute, and Statistical and Applied Mathematical Sciences Institute.

IMSI was founded on the belief that the mathematical sciences are a key enabler of progress in the wider enterprise of research, science, and technology, and the conviction that they must partake in this enterprise in order to thrive. While mathematical tools and insights have always been crucial to research in other disciplines, their usefulness has broadened and intensified in recent years. The National Academies' report on *The Mathematical Sciences in 2025*[2] noted that the mathematical sciences are becoming "an increasingly integral and essential component of a growing array of areas of investigation." It argued that "the mathematical sciences have an exciting opportunity to solidify their role as a linchpin of twenty-first century research and technology" by transforming into "a discipline with a much broader reach and greater potential impact." IMSI aims to be a catalyst for this kind of transformation.

The mathematical sciences' capacity to bring insight to urgent societal challenges is an important dimension of the field's possible impact. Prominent examples include modeling the spread of COVID-19 and its interactions with social distancing policies, the economy, and various forms of infrastructure; modeling climate and the effects of climate change on the conditions that facilitate life on Earth; and investigating the power of ideas and techniques from artificial intelligence and machine learning to transform human society, both for good and ill. The mathematical sciences can make a difference for these types of challenges, many of which demand all hands on deck. Responding to these issues is an essential component of realizing the mathematical sciences' full potential and forms a core feature of IMSI's mission. This active approach is also crucial to the health and diversity of the community; we are less likely to attract new talent to the field if we remain on the sidelines in the face of urgent situations.

This last point speaks to another thread in IMSI's mission: the transformation of the mathematical sciences community. The 2025 report pointed out the benefits that the community could gain if more of its practitioners possessed the skills necessary for interdisciplinary communication and collaboration. Mathematicians with these qualities are critical to the discipline's aforementioned transformation. They are also more likely to model a broad range of career paths in the mathematical sciences, and thus offer further opportunities for a diverse population to envision and define roles for themselves within the profession. IMSI intends to help researchers more fully develop communicative qualities through both research activities and workforce development programs, such as internships for graduate students.

Research at IMSI is organized around a set of scientific themes. These themes will evolve over time, but only on relatively long timescales — on the order of a decade or more. IMSI has identified six initial themes: climate science, data and information, health and medical care, materials science, quantum computing and information, and uncertainty quantification. This organizational scheme corresponds with IMSI's intention that institute activity be driven primarily by applications, rather than specific techniques or areas within the mathematical sciences. We want to facilitate conversations that stem from problems that climate scientists, doctors, economists, and molecular engineers (among others) currently face, and explore ways in which the mathematical sciences might contribute.

We hope that this approach will also spark new conversations across disciplinary boundaries within the field. Organizing our efforts around applications that are relevant to multiple areas of mathematics can bring new perspectives to the connections between these areas. In particular, we expect IMSI to make the boundary between mathematics and statistics more porous by including both mathematicians and statisticians in as many scientific activities as possible.

All of these objectives will require effective communication between different research communities, which connects to another pillar of IMSI's mission. Broadening the usefulness of research requires that we broaden access to the insights it produces. Researchers must be able to describe their work at various levels of accessibility and calibrate their communications to audiences with different technical backgrounds. One of IMSI's ambitions is to develop training programs for institute visitors that prepare them to communicate with diverse audiences, including researchers in other disciplines, funding agencies, policymakers, and the public.

Scientific activity at IMSI will occur in a variety of forms:

– Long programs, typically three months in length, that bring a large interdisciplinary group of researchers together for a period of sustained focus on an area that is ripe for progress

– Workshops, either standalone or attached to a long program, that are typically up to a week in length

– Interdisciplinary research clusters, in which small interdisciplinary teams collaborate on promising projects

See **Mathematical Innovation** *on page 11*

[1] https://www.imsi.institute

[2] https://www.nap.edu/catalog/15269/the-mathematical-sciences-in-2025

Understanding the New Baseball Statistics

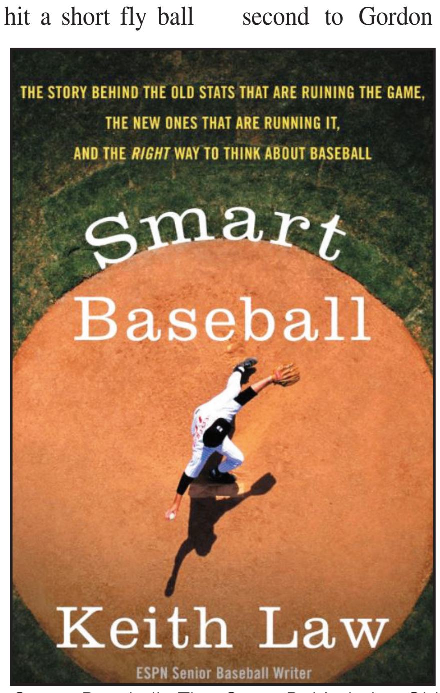
Smart Baseball: The Story Behind the Old Stats That Are Ruining the Game, the New Ones That Are Running It, And the *Right* Way to Think About Baseball. *By Keith Law. William Morrow, New York, NY, April 2017. 304 pages, $27.99.*

BOOK REVIEW
By James Case

Sportswriter Keith Law served as a special assistant to the general manager of the Toronto Blue Jays Baseball Club for several years. His 2017 book, *Smart Baseball*, describes the quantitative revolution that has swept baseball front offices since Michael Lewis published *Moneyball* in 2003. All 30 major league teams now have "Analytics Departments" of technically proficient individuals who attempt to improve player procurement practices. In the baseball world, "analytics" is a catch-all phrase that encompasses the collection and storage of data, as well as the use of that data to generate insight. Before spending five to eight million dollars on a recent high school graduate—or several hundred million on an established player—it pays to invest time, effort, and money assessing that player's potential cash value.

Smart Baseball is divided into three parts. The first section explains how traditional baseball statistics are nearly useless when it comes to evaluating players, while the second describes the bewildering variety of new statistics that have become available in recent years — mainly due to the ever-increasing size and number of online databases. The book's final section explains how one can use the new statistics to guide player procurement.

Law reserves his harshest condemnation for two familiar baseball statistics: runs batted in (RBIs) and pitcher wins (Ws). Teams, not pitchers, get wins, and most of the runs for which batters get credit occur when their teammates are on base ahead of them. It is not that difficult to hit a short fly ball to the outfield with a speedy runner on third base; does the runner deserve less credit than the batter, even though he scored the resulting run? Likewise, a pitcher could pitch badly for five innings—giving up 10 runs while his teammates score 12—before being replaced by a relief pitcher who obliterates the opposition for the next four innings. Why, Law asks, should the inept starting pitcher receive credit for the win instead of the dominating reliever?



Smart Baseball: The Story Behind the Old Stats That Are Ruining the Game, the New Ones That Are Running It, And the Right Way to Think About Baseball. By Keith Law. Courtesy of William Morrow.

Law also considers batting average (AVG)[1] to be overrated and uses Figure 1 to justify his skepticism of this statistic. "GORDON" is Dee Gordon, then of the Miami Marlins, who "led the league in batting" in 2015 with an AVG of .333. "HARPER" is Bryan Harper of the Washington Nationals, whose AVG was "only" .330. While Harper finished a close second to Gordon in AVG, he outperformed him by a substantial margin in every other category of interest, including adjusted batting runs (ABR). So does it really make sense to call Gordon the "batting champion?"

Casual fans are often confused by the so-called "triple slash line" that is frequently appended to a player's name to gauge his prowess as a hitter (i.e., Harper .330/.460/.649). The first number is the usual AVG, which is calculated by the ratio $AVG = \frac{H}{AB}$. Here, H is the number of hits made by the player and AB is the number of his official turns at bat. The middle number is a relatively recent statistic called on-base percentage: $OBP = (H + BB + HBP)/PA$. BB is "bases on balls," HBP is "hit by pitch," and PA is the number of "plate appearances" the player actually made: $PA = AB + BB + HBP + SF$. SF stands for "sacrifice flies," the batter's fly ball outs that allowed a baserunner to score. Finally, the last number is the less familiar slugging average (SLG), which is the ratio TB/AB, where TB is the number of "total bases" generated in the same series of "at-bats:" $TB = 1B + 2 \times 2B + 3 \times 3B + 4 \times HR$. $1B$ is singles, $2B$ is doubles, $3B$ is triples, and HR is home runs.

Statisticians can calculate the three "slash line statistics" for both teams and individual players. Moreover, all three correlate strongly with a team's *runs per game* (R/G), as Law demonstrates in Figure 2 (on page 12). On-base plus slugging (OPS) is as illegitimate as a statistic can be—it is calculated by adding ratios that lack a common denominator: $OPS = OBP + SLG$—but it obviously works. Teams with high OPS scores accumulate a lot of runs.

Later in the book, Law devotes an entire chapter to OBP as the measure of a hitter. His reason, roughly speaking, is that "a walk is as good as a hit." In this new way of thinking, a hitter's first duty is not to make an out. Each team begins every game with a "supply" of 27 outs; when both teams have exhausted their supplies, the game is over.[2] Therefore, a batter who completes a turn at bat without getting out increases the number of runs his team can expect to score in the game. A batter who makes an out diminishes that expectation unless it is a "productive out," such as a sacrifice fly.

In addition to popularizing OBP, baseball analysts have developed a number of "weighted average metrics" to evaluate specific aspects of player performance. One such metric is "batting runs" (BR), which is calculated as follows:

$$BR = .47H + .38D + .55T + .93HR + .33(BB + HBP) - .28\ OUTS.$$

See **Baseball Statistics** *on page 12*

	AVG	OBP	SLG	2B	HR	BB	OUTS	ABR
HARPER	.330	.460	.649	38	42	124	372	79.1
GORDON	.333	.359	.418	24	4	25	447	10.9

Figure 1. *Comparison of Dee Gordon of the Miami Marlins and Bryan Harper of the Washington Nationals in 2015, in terms of batting average (AVG), on-base percentage (OBP), slugging percentage (SLG), doubles (2B), home runs (HR), bases on balls (BB), OUTS, and adjusted batting runs (ABR). Figure adapted from Smart Baseball.*

[1] AVG = number of hits / number of official turns at bat.

[2] Unless of course the score is tied, in which case each team is issued a resupply of three supplementary outs.

Fairness in Voting and Resource Allocation

By Christoph Börgers

The following is a short and timely reflection from the author of Mathematics of Social Choice: Voting, Compensation, and Division, which was published by SIAM in 2010.

American democracy has widely recognized flaws. We vote on a workday, when many people—especially the underprivileged—find it hard to take time off to make it to the polling stations. Lines in front of polling stations are longer in poor neighborhoods than in wealthy ones. Gerrymandering distorts election outcomes. In many elections, voters who support small parties choose one of the two major parties instead, for fear of "throwing their vote away."

Although these are political issues, they do have mathematical aspects. Among them is the question of how to structure a fair election that selects a single winner from a field of candidates. This is the subject of the first half of my book, *Mathematics of Social Choice*. I restrict myself to election methods that are based on voters' *rankings* of candidates. We call a collection of candidate rankings by all voters a *preference schedule*. Denoting the set of all possible preference schedules by $\mathcal{R}$ and the set of all candidates by $\mathcal{C}$, an election method is a map

$$\mathcal{R} \to \mathcal{C}.$$

One can easily think of criteria that this map should satisfy — criteria on which most people would agree. We then ask whether and how a map that satisfies these criteria can be designed.

Nicolas de Condorcet (1743-1794) proposed the most famous such criterion. I use the notation

$$X \succ Y$$

to indicate that more than half the voters rank candidate X over Y. Condorcet's criterion then states that $X \in \mathcal{C}$ should win if

$$\forall Y \in \mathcal{C} \quad X \neq Y \Rightarrow X \succ Y.$$

One calls X the *Condorcet candidate*. A method that satisfies Condorcet's criterion is called *Condorcet-fair*.

FROM THE SIAM BOOKSHELF

The most common election method in the U.S. is the *plurality method*. The winner is the candidate whom the largest number of voters name as their top choice. This is not Condorcet-fair. For example, in the 2000 presidential election in Florida, Al Gore lost to George W. Bush by 537 votes, assuming the votes were counted correctly. However, Green Party candidate Ralph Nader received over 97,000 votes. It seems likely that most Nader voters preferred Gore to Bush, so Gore was probably the Condorcet candidate. Had Gore won Florida, he would have won the presidency.

The economists Eric Maskin and Amartya Sen have advocated for adherence to Condorcet's criterion, which they call "majority rule" [1]. If we agree, we still have to decide what to do when there is no Condorcet candidate. This is not likely, so it is a secondary issue, but it is by no means impossible and therefore must be addressed. There are many possible answers and many fairness criteria to examine, some of which I discuss in my book.

Most voting reform activists in the U.S. favor a different method, called *instant runoff* or *ranked choice voting* (RCV).[1] It is not Condorcet-fair, but it violates Condorcet's criterion less frequently than the plurality method. The strength of RCV is the so-called *later-no-harm property*: voters' second, third, etc. choices cannot affect whether their first choice wins. No Condorcet-fair method has this property.

The second half of *Mathematics of Social Choice* is about fair resource allocation, particularly "cake cutting." The mathematical formalization of this problem is due to Hugo Steinhaus [2], the mentor and collaborator of Stefan Banach. It is based on measure theory, that is, additivity (the language in my book is less formal). This strikes me as a weakness, as the idea that 100 apple pies should please me 100 times as much as one pie seems far-fetched. However, the subject does have real applications to sharing computing resources, assigning seats in sought-after college courses, etc.

I wrote my book for a course at Tufts University that is intended for students who are majoring in non-mathematical subjects; they take the course to satisfy a breadth requirement. Nonetheless, the book contains many theorems and their proofs, along with examples and discussions of the significance of these theorems.

In all my classes, I have always asked myself, "How can I make what I am talking about interesting to my students?" An even better question is, "How can I talk about what my students are interested in?" Many college students care about societal fairness, perhaps now more than ever, which is why fair voting and fair resource allocation are good topics for college mathematics courses, even for mathematics majors.

[1] https://www.fairvote.org

References

[1] Maskin, E., & Sen, A. (2016, May 1). How to let the majority rule. *The New York Times*, p. SR-7.

[2] Steinhaus, H. (1948). The problem of fair division. *Econometrica*, *16*, 101-104.

Christoph Börgers is a professor of mathematics at Tufts University.

CoSIAM 2020 School of Applied and Industrial Mathematics

By Diana Bueno

The Colombia Section of SIAM (CoSIAM)[1] held the inaugural CoSIAM 2020 School of Applied and Industrial Mathematics (EMAI CoSIAM 2020)[2] in late July. This week-long program allowed both national and international institutions to share courses and interdisciplinary experiences related to applied and industrial mathematics. EMAI CoSIAM 2020 served students and professionals who are interested in learning the techniques of applied and industrial mathematics, and provided a platform that highlighted the successful experiences of experts in the math community.

The school featured separate courses for three different topics: machine learning, complex networks, and system dynamics. The courses occurred remotely, in accordance with meeting guidelines during the COVID-19 pandemic. 57 participants—in academic, industrial, and governmental sectors—took part in the event from various regions of Colombia and abroad. During the opening ceremony, organizers explained the CoSIAM ecosystem to clarify the sectors, audiences, and areas of action that the section strives to address in its efforts to develop and promote applied mathematics in Colombia.

The CoSIAM ecosystem uses three axes to frame its actions (see Figure 1). The first axis corresponds to the section's intended audience: undergraduate students, graduate students, and professionals. The second axis covers the areas of action: motivation, training, research, and networking. And the third axis includes the sectors that CoSIAM hopes to impact: academia and the real world. Using this framework, EMAI CoSIAM 2020 academically impacted all audiences in the training area.

Further information about EMAI CoSIAM 2020 and the Colombia Section of SIAM is available online.[3]

Diana Bueno is the vice president of the Colombia Section of SIAM and director of the Applied Mathematics Program at Pontificia Universidad Javeriana Cali in Colombia.

[1] https://www.siam.org/membership/sections/detail/colombia-section-of-siam-cosiam

[2] https://www.cosiam.net/emai-cosiam

[3] https://www.cosiam.net

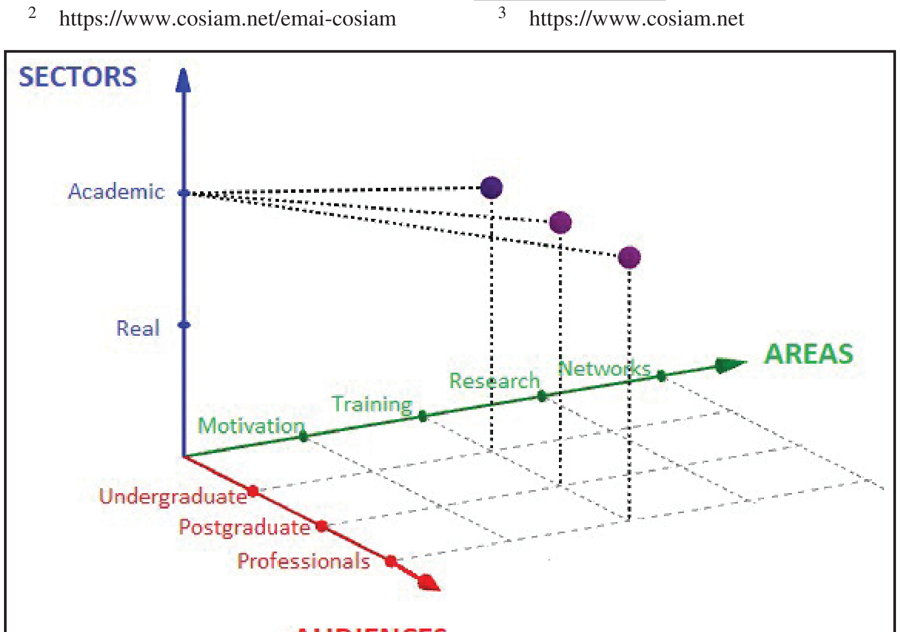


Figure 1. *The CoSIAM 2020 School of Applied and Industrial Mathematics (EMAI CoSIAM 2020) within the ecosystem of the Colombia Section of SIAM. Figure courtesy of Diana Bueno.*

Geometric Sum, Geometrically

MATHEMATICAL CURIOSITIES
By Mark Levi

It is customary to prove the geometric sum formula

$$1+\lambda+\lambda^2+\ldots+\lambda^{n-1}=\frac{1-\lambda^n}{1-\lambda} \qquad (1)$$

by algebra. But a geometric sum deserves a geometric treatment, so here are some geometrical proofs.

Proof by Dilation

Let us subject the segment $[0,1]$ to a linear dilation $x \mapsto \lambda x$ that is repeated n times—as shown in Figure 1—for positive $\lambda<1$. The iterates $\lambda, \ldots, \lambda^n$ break $[0,1]$ into $n+1$ subintervals. And the iterates of the rightmost interval $[\lambda, 1]$ have geometrically decreasing lengths $\lambda^i(1-\lambda)$ with $i=0,\ldots,n-1$, as Figure 1 shows. The combined length of all the intervals, recording from right to left in the figure, is 1:

$$(1-\lambda)+\lambda(1-\lambda)+\ldots+$$
$$\lambda^{n-1}(1-\lambda)+\lambda^n=1,$$

thus resulting in (1).

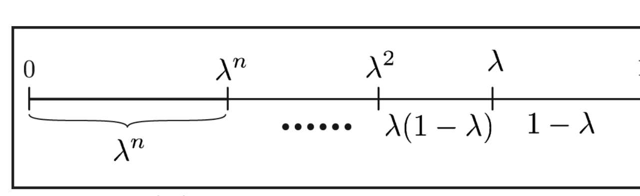


Figure 1. *Proof of (1) by dilation.*

The remaining two proofs are for the infinite sum

$$1+\lambda+\lambda^2+\ldots=\frac{1}{1-\lambda} \qquad (2)$$

and positive $\lambda<1$.

A Pythagorean Proof

The construction of Figure 2 yields a partition of the hypotenuse into an infinite union of segments of lengths 1, $\sin^2\theta, \sin^4\theta, \ldots$ On the other hand, the hypotenuse has length $(1/\cos\theta)/\cos\theta=\cos^{-2}\theta$, so that

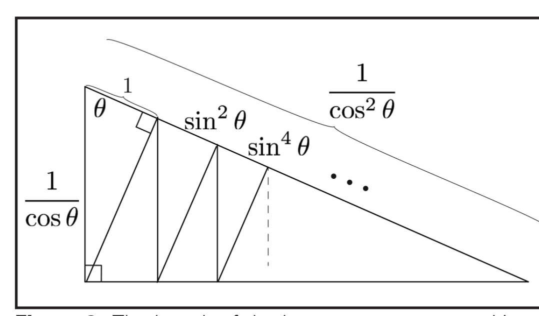


Figure 2. *The length of the hypotenuse computed in two different ways yields (3).*

$$1+\sin^2\theta+\sin^4\theta+\ldots=\frac{1}{\cos^2\theta}. \qquad (3)$$

This implies (2) by choosing θ so that $\lambda=\sin^2\theta$, and by application of the Pythagorean theorem. As a curiosity, reversing the argument—i.e., taking (2) for granted—gives an admittedly strange proof of the Pythagorean theorem.

A Staircase Proof

The two lines $y=1+\lambda x$ and $y=x$, which appear in Figure 1, intersect at height $y=\frac{1}{1-\lambda}$. But this height is also the sum of rises that form a geometric sequence $1, \lambda, \lambda^2, \ldots$, again yielding (2).

The figures in this article were provided by the author.

Mark Levi (levi@math.psu.edu) is a professor of mathematics at the Pennsylvania State University.

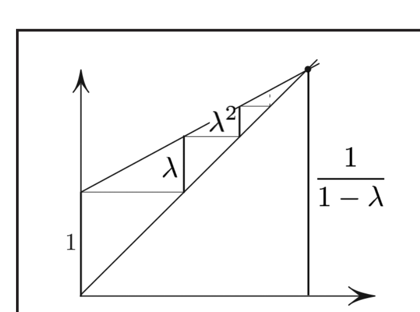


Figure 3. *Proof of (2).*

SIAM Engages Participants at the 2020 Virtual SciFest

By Wesley Hamilton and Padmanabhan Seshaiyer

The USA Science and Engineering Festival[1] (USASEF) is a bi-annual exposition with booths, activities, and invited speakers for K-12 students, parents, teachers, and educators. For the last several years, the event has been held at the Walter E. Washington Convention Center in Washington, D.C., and has attracted over 250,000 visitors during its three-day run. SIAM has been participating in USASEF since it began a decade ago. The 2020 festival was initially scheduled to take place in person from April 24 to 26. But due to COVID-19, the event transitioned to an all-virtual format and commenced online from September 26 through October 31, 2020. Registration was free and STEM enthusiasts of all ages attended. The SIAM Education Committee assembled the SIAM booth. While we could not adapt all of our original plans to the online format, we were still able to construct an engaging experience for attendees.

We designed our virtual booth to educate visitors on the importance of mathematics. The booth incorporated multiple activities, such as a viewing of SIAM's "I use math for…" video;[2] a brief description of the society; and a demonstration and worksheet on internal waves, which was developed by the Joint Applied Mathematics and Marine Sciences Fluids Lab at the University of North Carolina at Chapel Hill (UNC). Padhu Seshaiyer, who has previously served as a "Nifty Fifty" speaker for USASEF, was a STEM Stage speaker this year and discussed the value of learning mathematics in a meaningful way.

During prior festivals, the SIAM booth displayed exhibits of Archimedean solids and boasted activities like origami folding and shape sudoku. SIAM has worked with volunteer student chapters from schools that are both close to Washington, D.C. (such as George Washington University and George Mason University) and further away (such as UNC, North Carolina State University, Temple University, and Virginia Tech). For the would-be spring 2020 festival, we had planned to offer the same activities as before with some engaging new additions, including a Monty Hall problem demonstration with statistics gathered during the event itself, non-transitive dice and unusual probabilities, and tessellations.

The planning committee met roughly once a month beginning in September 2019 and had all of the hands-on activities nearly complete by January 2020. While in theory we could have moved these activities online, we decided as a committee that the experience would not have been the same. Instead of rushing to preserve some of the excitement and engagement in the unprecedented online environment, we chose to save these plans for future in-person festivals. While we were not able to directly see the excitement and curiosity on attendees' faces this year, SIAM's virtual booth no doubt drove visitor interests toward applied mathematics.

Are you near Washington, D.C., and interested in helping to plan and volunteer at future SIAM festival appearances? Email Wesley Hamilton (wham@live.unc.edu) and we'll reach out whenever we begin planning for the next in-person festival!

Acknowledgments: Planning for SIAM's booth at USASEF would not have been possible without the help of several key players. From the SIAM Education Committee, Wesley Hamilton (UNC) led the planning process while Katie Kavanagh (Clarkson University), who is also the SIAM Vice President for Education, marshalled resources and made connections. Richard Moore, SIAM's Director of Programs and Services, helped coordinate SIAM resources and teleconferencing logistics, and marketing representative Kristin O'Neill managed the virtual booth and communicated with USASEF organizers. Finally, a number of volunteers, including Tracey Oellerich (George Mason University), Blain Patterson (Virginia Military Institute), and Grant Innerst (Shippensburg University), helped plan activities that—while not present at our virtual booth—will appear at the next in-person USASEF and other festivals in which SIAM participates.

Wesley Hamilton is a Ph.D. candidate in mathematics and president of the University of North Carolina at Chapel Hill SIAM Student Chapter. He helped to plan and coordinate the 2020 USA Science and Engineering Festival (USASEF). Padmanabhan (Padhu) Seshaiyer is a professor of mathematical sciences and the associate dean for the College of Science at George Mason University, as well as chair of the SIAM Diversity Advisory Committee. He served as a STEM Stage speaker at the 2020 virtual USASEF.

[1] https://usasciencefestival.org
[2] https://go.siam.org/7zByxu

The SIAM booth at the 2020 USA Science and Engineering Festival, which took place virtually from September 26 through October 31, 2020.

Mathematical Innovation

Continued from page 9

– Research collaboration workshops, during which teams of junior and senior researchers work on problems over several months and conclude their collaborations with a workshop.

All of these activities are expected to fall within the scope of IMSI's scientific themes. Standalone workshops, which may explore a wider range of topics, are the sole exception.

IMSI is currently undergoing a ramp-up period, so research activity will primarily take the form of workshops. We expect to host several workshops[3] in 2021 (either virtually or in-person) on the following topics: Mathematical and Computational Materials Science (February 15-19); Confronting Climate Change (March 1-5); The Multifaceted Complexity of Machine Learning (April 12-16); Topological Data Analysis (April 26-30); Verification, Validation, and Uncertainty Quantification Across Disciplines (May 10-14); Decision Making in Health and Medical Care: Modeling and Optimization (May 17-21); Quantum Information for Mathematics, Economics, and Statistics (May 24-28), and Eliciting Structure in Genomics Data: Bridging the Gap Between Theory, Algorithms, Implementations, and Applications (August 30-September 3). A workshop on Dealing with COVID-19 in Theory and Practice took place in late October.[4]

We plan to follow what we expect will be a more typical schedule in 2021-2022, with a long program on Distributed Solutions to Complex Societal Problems in the fall and one on Decision Making and Uncertainty in the spring. Each program[5] will have a linked introductory summer program in 2021.

We are hoping for significant engagement from the mathematics community. If you have ideas, we encourage you to reach out to us to discuss them.

Kevin Corlette is a professor in the Department of Mathematics at the University of Chicago and director of the Institute for Mathematical and Statistical Innovation (IMSI). Douglas Simpson is a professor in the Department of Statistics and the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign. He is the associate director of IMSI. Panagiotis Souganidis is a professor in the Department of Mathematics and a member of the Committee on Computational and Applied Mathematics at the University of Chicago. He is the scientific adviser at IMSI.

[3] https://www.imsi.institute/workshops
[4] https://www.imsi.institute/covid-19
[5] https://www.imsi.institute/programs

Professional Opportunities and Announcements

Send copy for classified advertisements and announcements to ***marketing@siam.org****.*
For rates, deadlines, and ad specifications, visit ***www.siam.org/advertising****.*

*Students (and others) in search of information about careers in the mathematical sciences can click on "Careers" at the SIAM website (****www.siam.org****) or proceed directly to* ***www.siam.org/careers****.*

Dartmouth College

Department of Mathematics

The Department of Mathematics at Dartmouth College welcomes applications for a junior tenure-track opening with initial appointment as early as the 2021-2022 academic year. Exceptional cases can merit appointment at a higher rank. The successful applicant will have a research profile with a concentration in applied or computational mathematics. Current research areas in applied mathematics include complex systems, computational social sciences, network analysis, statistical learning, mathematical biology, stochastic processes, uncertainty quantification, partial differential equations, and signal and image processing.

Applicants should apply online at **www.mathjobs.org**, Position ID: **APAM #16169**. Applications received by **December 15, 2020** will receive first consideration. For more information about this position, please visit our website: **https://www.math.dartmouth.edu/activities/recruiting**.

Dartmouth is highly committed to fostering a diverse and inclusive population of students, faculty, and staff. We are especially interested in applicants who are able to work effectively with students, faculty, and staff from all backgrounds—including but not limited to racial and ethnic minorities, women, individuals who identify with LGBTQ+ communities, individuals with disabilities, individuals from lower-income backgrounds, and/or first-generation college graduates—and who have a demonstrated ability to contribute to Dartmouth's undergraduate diversity initiatives in STEM research, such as the Women in Science Project, E.E. Just STEM Scholars Program, and Academic Summer Undergraduate Research Experience (ASURE). Applicants should state in their cover letter how their teaching, research, service, and/or life experiences prepare them to advance Dartmouth's commitments to diversity, equity, and inclusion.

Dartmouth College

Department of Mathematics

The Department of Mathematics at Dartmouth College is delighted to announce a senior opening in applied mathematics at the rank of professor or associate professor, with initial appointment as early as 2021-2022, as the Jack Byrne Professor or Associate Professor of Applied Mathematics. In exceptional circumstances, we may consider an appointment at the associate professor level. A Ph.D. in mathematics, statistics, or a related field is required. We seek an acknowledged international leader in applied mathematics with an exemplary track record in creating mathematical and statistical methodological advances and their applications. Current applied and computational interests in the department include complex systems, computational social sciences, image and signal processing, mathematical biology, network analysis, statistical learning, stochastic processes, and uncertainty quantification. Our strength in applied mathematics is complemented by strength in several areas of theoretical mathematics.

This position is part of the larger "Byrne Cluster" comprising two positions in the Department of Mathematics and a recent senior hire in decision sciences in Dartmouth's top-ranked Tuck School of Business. The Byrne Cluster represents a new investment in the department's continued efforts to expand its research endeavors and related pedagogy in applied mathematics. We seek a candidate with a demonstrated ability to work across fields and bridge multiple research areas both inside and outside the Department of Mathematics, specifically including the Byrne Cluster member of the Tuck School. The Byrne Cluster comes with programmatic funds to support these interdisciplinary goals. In addition to research qualifications, candidates should have a keen interest and demonstrated excellence in teaching and mentorship of both undergraduates and graduate students.

Applicants should apply online at **www.mathjobs.org**, Position ID: **APAM #16253**. Applications received by **December 15, 2020** will receive first consideration. For more information about this position, please visit our website: **https://www.math.dartmouth.edu/activities/recruiting**.

Dartmouth is highly committed to fostering a diverse and inclusive population of students, faculty, and staff. We are especially interested in applicants who are able to work effectively with students, faculty, and staff from all backgrounds—including but not limited to racial and ethnic minorities, women, individuals who identify with LGBTQ+ communities, individuals with disabilities, individuals from lower-income backgrounds, and/or first-generation college graduates—and who have a demonstrated ability to contribute to Dartmouth's undergraduate diversity initiatives in STEM research, such as the Women in Science Project, E.E. Just STEM Scholars Program, and Academic Summer Undergraduate Research Experience (ASURE). Applicants should state in their cover letter how their teaching, research, service, and/or life experiences prepare them to advance Dartmouth's commitments to diversity, equity, and inclusion.

Institute for Pure and Applied Mathematics

Associate Director

The Institute for Pure and Applied Mathematics (IPAM) at the University of California, Los Angeles (UCLA) is seeking an associate director (AD) for a two-year appointment (renewable for a third year) starting August 1, 2021. The AD will be an active member of the team and is expected to have a research background in mathematics or related fields, with experience in conference organization. The primary responsibility of the AD will be coordinating with the organizing committees to execute IPAM programs. To apply and learn more, go to **https://recruit.apo.ucla.edu/JPF05835**. Applications will receive fullest consideration if received by **February 19, 2021**. UCLA is an equal opportunity/affirmative action employer.

Institute for Pure and Applied Mathematics

Simons Postdoctoral Scholars

The Institute for Pure and Applied Mathematics (IPAM) at the University of California, Los Angeles (UCLA) is seeking to recruit up to three Simons Postdoctoral Scholars (SPD), funded by the Simons Foundation. The appointment will be for one calendar year, beginning August 1, 2021. A Ph.D. in mathematics, statistics, or a related field received in May 2016 or later is required. Women and minorities are especially encouraged to apply. To apply and learn more, go to **https://recruit.apo.ucla.edu/JPF05895**. Applications will receive fullest consideration if received by **January 1, 2021**. UCLA is an equal opportunity/affirmative action employer.

The Aerodynamics and Biology of Insect Flight

By Matthew R. Francis

Many readers have likely heard the common trope that bumblebees should not be able to fly. This is obviously a misleading statement—bumblebees *can* fly, so their ability to do so is self-evident—but it highlights some of the complications that researchers face in their quest to understand insect flight. Due to their small body size and relatively low flight speeds, insects occupy a middle region of aerodynamics that lies between the macroscopic regime of airplanes and the microscopic world of bacteria; in this intermediate realm, the rules that govern each domain do not fully apply.

But insect flight is not simply about aerodynamics. It also pertains to evolution and the biological workings inside insect bodies. "I'm interested in insects as a living system," Jane Wang of Cornell University said. "The balancing act of the insect is very interesting, because it depends on the physics of flight as well as their evolved neurocircuitry. My hunch has been that part of the neural behavior can be understood best if we start from the physical principles of flight."

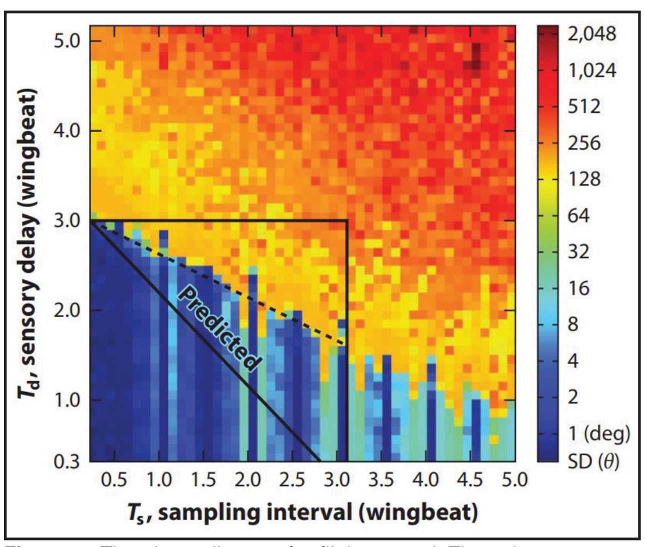


Figure 1. *The phase diagram for flight control. The colors represent the root-mean-square pitch angle of the insect body in degrees. Smaller values (blue) indicate well-controlled flight and higher values indicate insect tumbling. Figure courtesy of [1].*

During her invited presentation at the 2020 SIAM Conference on the Life Sciences,[1] which took place virtually this June, Wang described the wide range of theoretical models and laboratory experiments that are required to ask and answer questions pertaining to insect flight. "It's such a rich problem and contains many elements that we do not normally think about in non-living systems," she said.

Wang and her colleagues examined several species of insect to understand the mechanism that is needed to achieve controlled flight. They measured reaction times, quantified wing and body kinematics, computed flight maneuvers, and even found a potential link between a specific muscle and flight stability in flies.

Insects are the most diverse group of animals on the planet, which complicates matters. Biologists have identified over 900,000 species,[2] and as many as 30 million are still undescribed. From the tiniest fairyflies (wasps that are barely visible without a microscope) to the heaviest Goliath beetles, many insects fly and possess distinct flight mechanics. Bumblebees, dragonflies, and locusts—just to name a few—fly in different ways that reflect their various strategies for survival and feeding. These behavioral aspects accompany more straightforward issues of body size, weight, and shape while tying into the insects' sensory organs and nervous systems.

Flight on the Edge of Instability

However, a few commonalities do exist across species. Despite insects' relatively wide range of sizes, their flight is intrinsically unstable — they must therefore exert effort to maintain their orientation and avoid falling. Wang likens insect flight strategy at the border of stability and instability to our own human bipedal stance. Because we stand upright on two legs without additional balancing organs, we are always very close to falling over (anyone who has ever experienced equilibrium issues knows this well). Even standing in place requires constant tiny adjustments in posture. However, bipedalism means that we can translate our state of perpetual almost-falling into the efficient forward motion known as walking.

Wang postulates that the boundary between stable and unstable flight may provide similar advantages for flying insects [1]. Just as we exploit our unstable human uprightness to walk forward, insects might use flight instabilities to stay aloft and preserve maneuverability.

The starting point when modeling flapping insect wings is the Navier-Stokes equations for incompressible fluids:

$$\frac{\partial \vec{u}}{\partial t} + \vec{u} \cdot \nabla \vec{u} = -\nabla p + \frac{1}{\mathrm{Re}} \nabla \vec{u}$$

$$\nabla \cdot \vec{u} = 0.$$

Here, all physical quantities are expressed as dimensionless quantities and scaled by a characteristic velocity U and length scale L.

This dimensionless form of the Navier-Stokes equation reveals a vital parameter: the Reynolds number Re. $\mathrm{Re} = LU/\nu$, where ν is the kinematic viscosity of the fluid. Re is approximately 150 for a fruit fly and around 3,000 for a dragonfly. Since the transition of smooth fluid flow to unsteady flows occurs roughly at $\mathrm{Re} = 1{,}000$ and the flapping wings move at high angles of attack, many insects literally fly among vortices of turbulent air. Airplanes have Reynolds numbers that are several orders of magnitude higher, while bacteria live in a world of very tiny Re values (because of their small size).

"At this intermediate range of Reynolds numbers, you cannot ignore either the viscous force or the fluid inertia," Wang said. "In the case of flapping flight, the wings reverse periodically. Each time they do this they generate vortices. These vortices interact, so the flow is complex and hard to wrap our heads around."

In fact, Wang believes that it is possible for the vortices to provide some of the extra lift forces that insects need to stay aloft. To test this, she computed the flows and forces by solving the Navier-Stokes equations and studied plates falling through air and water, which have a similar range of Reynolds numbers. This allowed Wang and her students to recreate the types of vortices that are shed by flapping wings and extract theoretical models for the aerodynamic forces exerted on these wings.

Hovering and Maneuvering

The mechanics and aerodynamics comprise only part of the puzzle that Wang wishes to solve. Flight is dynamic; just as humans must actively maintain balance (a mostly reflexive and subconscious process), insects need to constantly react to their surroundings to control their body orientation and keep from falling.

Another important aspect of flight is hovering, which many flying insects can do (even if only for a short time). "This ability to pause has an advantage for maneuvering," Wang said. "Like driving a car, you can change your orientation very effectively if you stop instead of turning in a big loop. In order to hover, insects need neurofeedback algorithms so that they maintain their upright posture."

Wang's control algorithm for her insect model has two timescales: the sensing rate T_s and the reaction delay time T_d. Both govern the way in which the simulated insect adjusts to its body state. In particular, the phase diagram that spans these two timescales (see Figure 1) revealed the presence of a maximum reaction time that enables controlled flight, and demonstrated that effective control occurs when the sensing rate is in half-integer multiples of the rate at which the insects' wings beat.

For fruit flies, the most robust control transpires at the rate of the wing beat frequency (roughly four milliseconds), which is shorter than the flies' visual response time. They thus presumably receive sensory information from balancing organs called halteres, which scientists first identified in the early 18th century as essential for flight in flies. Subsequent research revealed that halteres—which are likely the evolutionary remnants of the second set of wings that many other insect groups possess—act much like gyroscopes and allow insects to detect their bodies' three-dimensional rotational rate. In addition, Wang speculates that a muscle called b1, which fires with every wing beat, is essential for flight stability.

These final points indicate the power of interactions between mathematics and biological studies. Physical analyses inspired the connection to insect neuroanatomy, which in turn informed the experiments' design and helped refine the theoretical models of flight control. Recognizing that insect flight involves a reflexive set of actions and exists on the boundary between stability and instability proved evolutionarily successful. We can observe the flight of the bumblebee, after all; understanding its flight, however, requires the right balance of seemingly disparate fields.

This article is based on Jane Wang's invited talk at the 2020 SIAM Conference on the Life Sciences, which occurred virtually earlier this year. Wang's presentation is available on SIAM's YouTube Channel.[3]

References

[1] Wang, Z.J. (2016). Insect flight: from Newton's laws to neurons. *Ann. Rev. Cond. Matt. Phys., 7*, 281.

Matthew R. Francis is a physicist, science writer, public speaker, educator, and frequent wearer of jaunty hats. His website is BowlerHatScience.org.

[1] https://www.siam.org/conferences/cm/conference/ls20

[2] https://www.si.edu/spotlight/buginfo/bugnos

[3] https://www.youtube.com/watch?v=Q5Izu4SRepA

Baseball Statistics

Continued from page 9

Here, $D = 2B$ and $T = 3B$. This statistic is easy to compute and particularly amenable to historical studies.

In recent years, new statistics have emerged that evaluate pitching performance in addition to hitting. Traditionally, the runs allowed (RA) or earned runs allowed (ERA) per nine-inning game have served as a measure of a pitcher's effectiveness. The latter is obviously superior, as it does not hold the pitcher responsible for runs that are attributable to his teammates' errors in the field. But runs can score even if fielders do not make errors, since speedy and/or agile fielders can produce outs that would have become hits against less mobile defenders. One attempt to correct for this shortcoming in the ERA statistic is known as batting average on balls in play (BABIP):

$$BABIP = (H - HR)/ \\ (AB - K - HR + SF).$$

Home runs (HR) are subtracted from hits (H) in the numerator because they are not balls put in play. For similar reasons, strikeouts (K) and home runs (HR) are subtracted from at-bats (AB) in the denominator, while sacrifice flies (SF) are added because they do not count as official at-bats. Major league teams use metrics like BABIP to separate a pitcher's performance from the effects of defense and luck.

While devising a single metric that evaluates hitting or pitching prowess is undoubtedly challenging, inventing one that measures fielding ability is even more difficult. The traditional revelation that a player compiled a fielding average of 0.973 in 111 "chances," for example, merely discloses that he failed to convert three of those chances into outs. Only those three "errors" count against him; any balls that a more mobile defender would have easily reached do not.

However, these traditional metrics are beginning to change. All 30 big league ballparks are now equipped with Statcast, a system that produces vast quantities of information that Major League Baseball (MLB) Advanced Media converts into a data stream of truly heroic proportions. The result is 1.5 billion rows of data that cover all MLB games beginning with Opening Day 2016. To remain competitive, teams must learn to store, access, and (eventually) analyze this data.

TEAM STATISTIC	CORRELATION TO TEAM *R/G*
Batting average (AVG)	0.749
On-base percentage (OBP)	0.833
Slugging percentage (SLG)	0.903
On-base plus slugging (OPS)	0.936

Figure 2. *The "slash line statistics"—which include batting average (AVG), on-base percentage (OBP), and slugging percentage (SLG)—correlate with a team's runs per game* (R/G). *Figure adapted from Smart Baseball.*

Statcast data include the speed, location, and spin rate of each pitch, as well as the exit velocity, launch angle, and direction of every batted ball. This newfound information allows analysts to deduce the projected direction of a batter's hit with surprising accuracy. Several teams have thus improved their defenses by employing exaggerated shifts against vulnerable opponents.

Perhaps the most portentous aspect of Statcast is the information it provides about movement while the ball is in play. Analysts can now identify a fielder's location when the batter makes contact with the ball and when he intercepts (or fails to intercept) its path. They also know the amount of time that elapses between those two events. At long last, statisticians may be able to reduce a player's defensive prowess to a single number.

Smart Baseball explains far more statistics than is possible to summarize in a single review. Law's tone is at times combative, and the casual fan may find his comparisons of little-known players a trifle tedious, but his explanations are both clear and informative. If you have ever found yourself wondering what modern baseball announcers are talking about on television, then this is the book for you.

James Case writes from Baltimore, Maryland.