Learning from Observations: The Case for a New Generation of Land Surface Models

Bart Nijssen[1], Andrew Bennett[2], and Grey Nearing[3]

[1]University of Washington Seattle Campus
[2]University of Arizona
[3]Google Research

November 24, 2022

Abstract

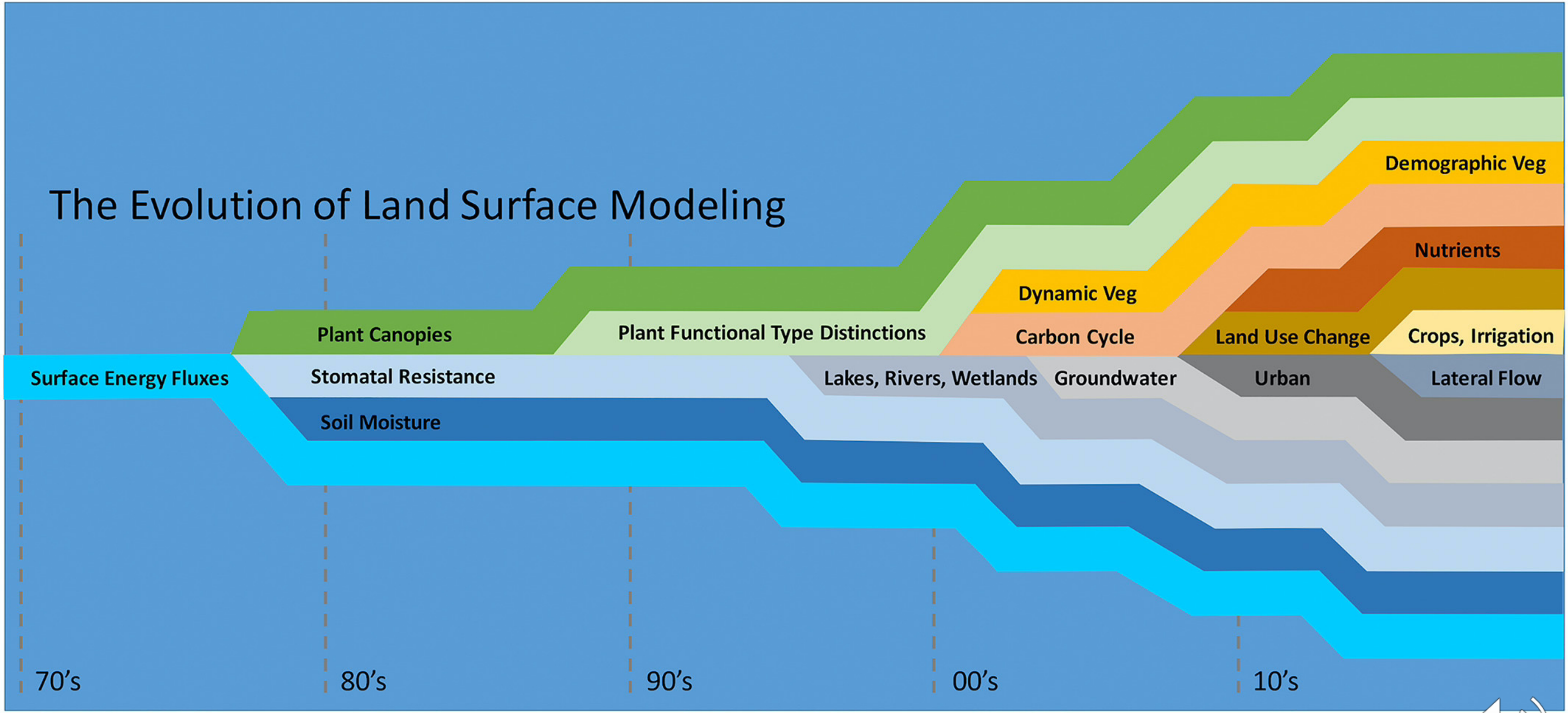
The hydrology community is engaged in an intense debate regarding the merits of machine learning (ML) models over traditional models. These traditional models include both conceptual and process-based hydrological models (PBHMs). Many in the hydrologic community remain skeptical about the use of ML models, because they consider these models "black-box" constructs that do not allow for a direct mapping between model internals and hydrologic states. In addition, they argue that it is unclear how to encode a priori hydrological expertise into ML models. Yet at the same time, ML models now routinely outperform traditional hydrological models for tasks such as streamflow simulation and short-range forecasting. Not only that, they are demonstrably better at generalizing runoff behavior across sites and therefore better at making predictions in ungauged basins, a long-standing problem in hydrology. In recent model experiments, we have shown that ML turbulent heat flux parameterizations embedded in a PBHM outperform the process-based parameterization in that PBHM. In this case, the PBHM enforced energy and mass constraints, while the ML parameterization calculated the heat fluxes. While this approach provides an interesting proof-of-concept and perhaps acts as a bridge between traditional models and ML models, we argue that it is time to take a bigger leap than incrementally improving the existing generation of models. We need to construct a new generation of hydrologic and land surface models (LSMs) that takes advantage of ML technologies in which we directly encode the physical concepts and constraints that we know are important, while being able to flexibly ingest a wide variety of data sources directly. To be employed as LSMs in coupled earth system models, they will need to conserve mass and energy. These new models will take time to develop, but the time to start is now, since the basic building blocks exist and we know how to get started. If nothing else, it will advance the debate and undoubtedly lead to better understanding within the hydrology and land surface communities regarding the merits and demerits of the competing approaches. In this presentation, we will discuss some of these early studies, illustrate how ML models can offer hydrologic insight, and argue the case for the development of ML-based LSMs.

Learning from Observations: The Case for a New Generation of Land Surface Models

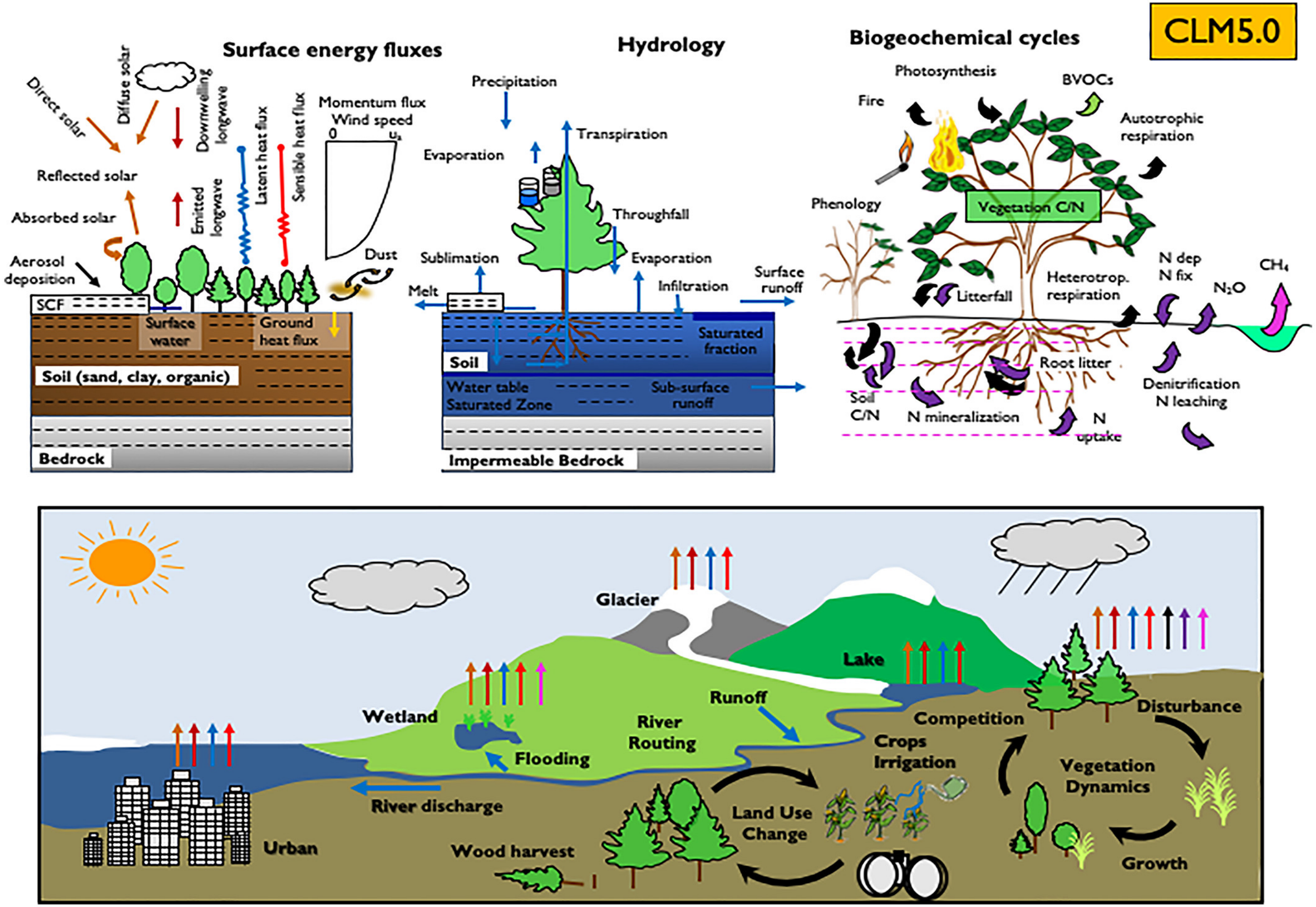
Bart Nijssen, Andrew Bennett, Grey Nearing

AGU Fall Meeting 2021

H33B-04



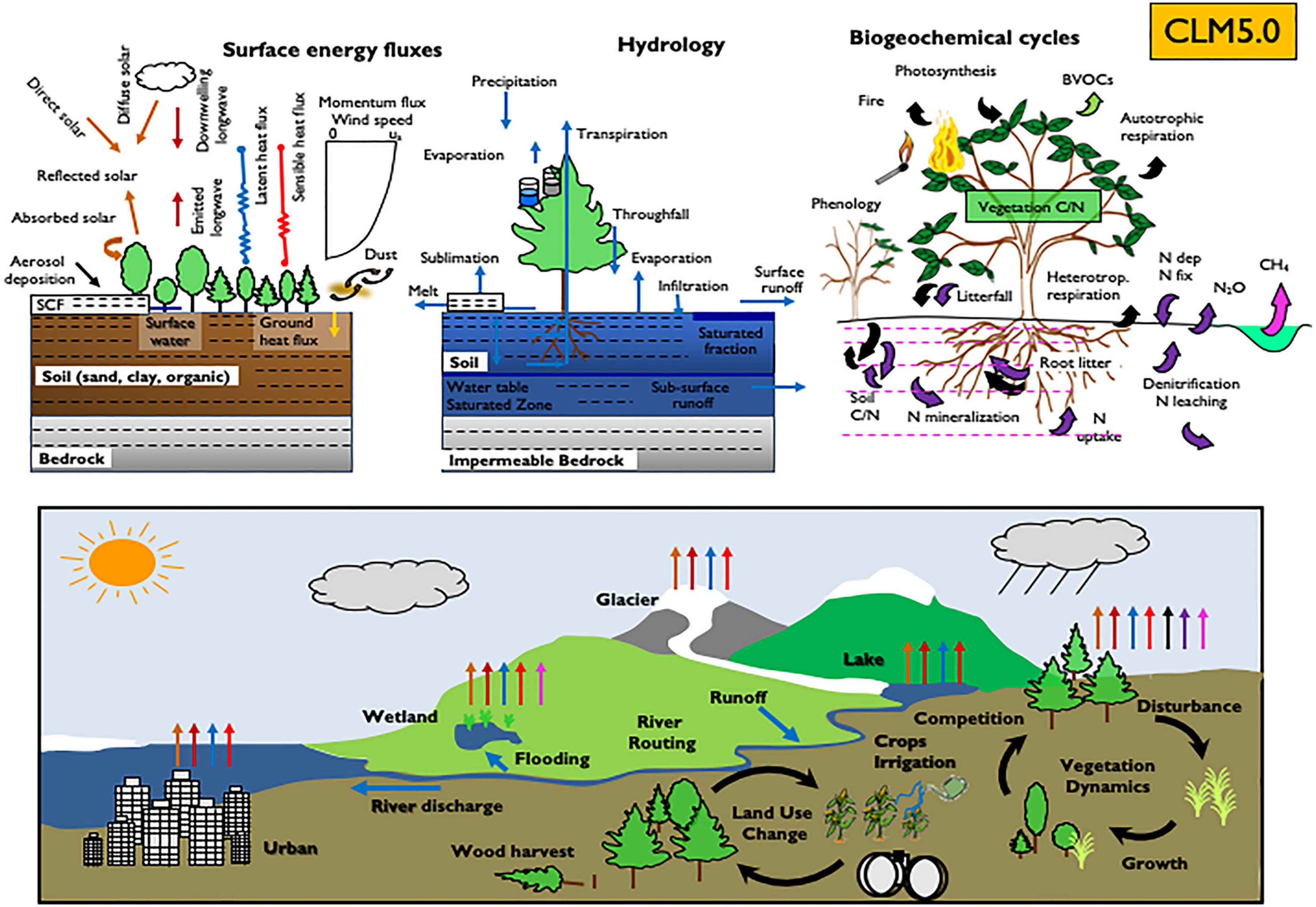
Fisher, R. A., and C. D. Koven (2020), Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems, *Journal of Advances in Modeling Earth Systems*, *12*(4), doi:10.1029/2018MS001453.



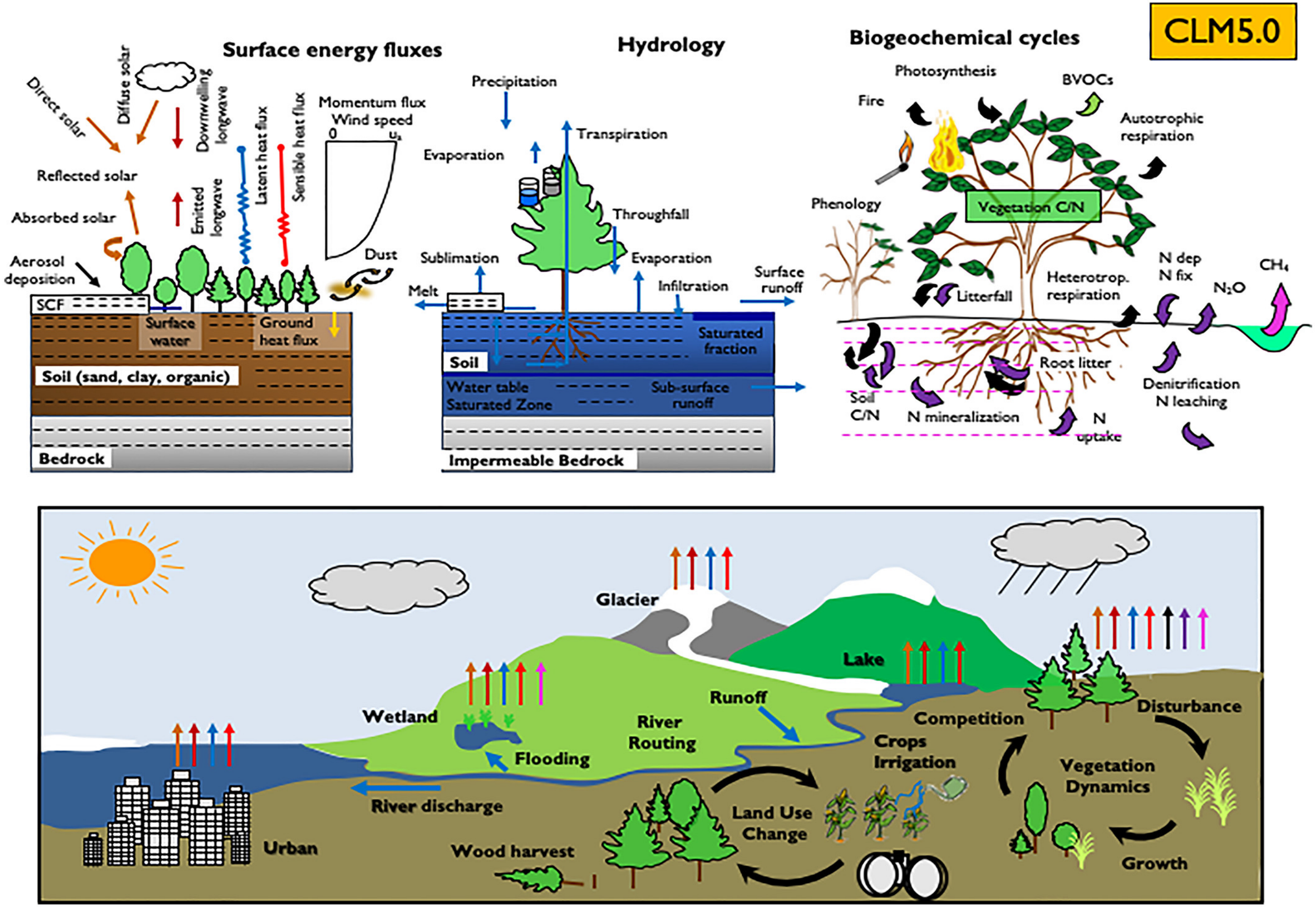
Lawrence, D. M., et al. (2019), The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty, *Journal of Advances in Modeling Earth Systems*, *11*(12), 4245-4287, doi:10.1029/2018MS001583.

big challenges

parameterizations and parameter values



Lawrence, D. M., et al. (2019), The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty, *Journal of Advances in Modeling Earth Systems*, *11*(12), 4245-4287, doi:10.1029/2018MS001583.



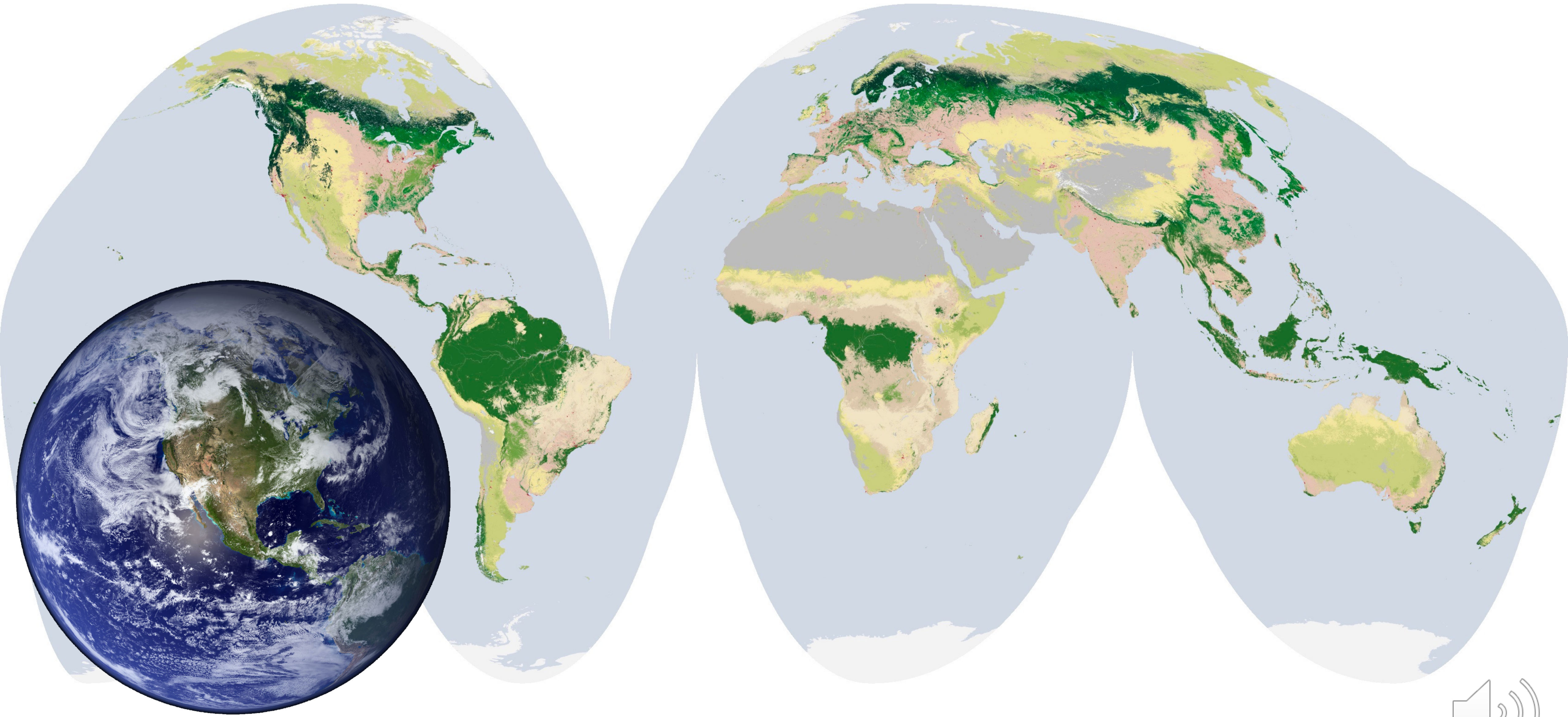
big challenges

parameterizations and parameter values

integrate across the landscape

Lawrence, D. M., et al. (2019), The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty, *Journal of Advances in Modeling Earth Systems*, *11*(12), 4245-4287, doi:10.1029/2018MS001583.

NASA Visible earth: Blue Marble: https://visibleearth.nasa.gov/images/57723/the-blue-marble



NASA Visible earth: Map by Robert Simmon, based on data from the [MODIS Land Cover Group,](#) Boston University.

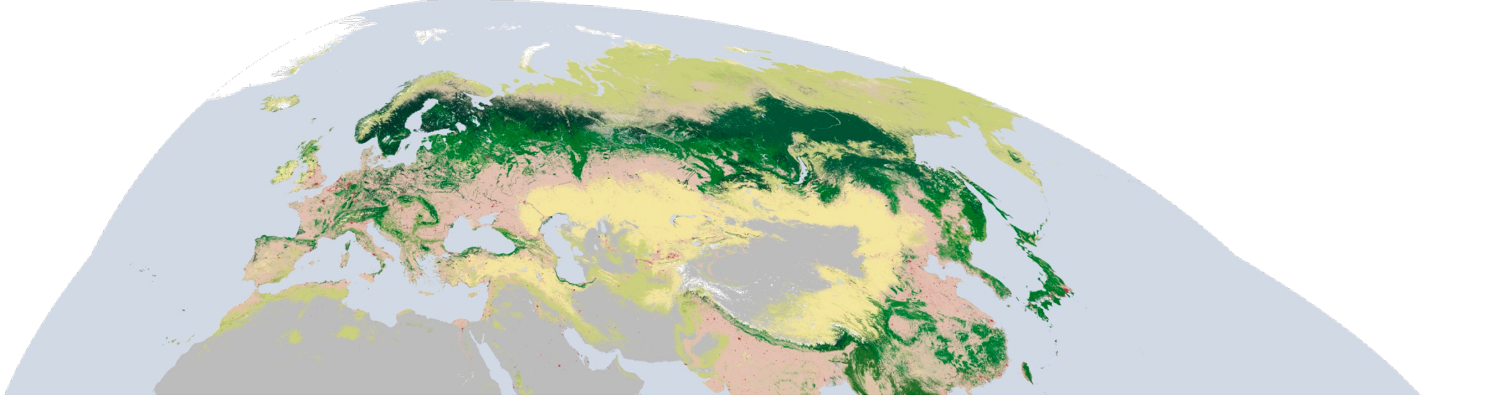


Table 3 Root zone depths (m) and fraction of roots in each zone for IGBP land cover classes.

From: A near-global, high resolution land surface parameter dataset for the variable infiltration capacity model

	Depth(1)	Depth(2)	Depth(3)	Fract(1)	Fract(2)	Fract(3)
Open water	0.1	0.6	0.8	0.44	0.45	0.11
Evergreen needleleaf forest	0.1	0.6	1.1	0.34	0.51	0.14
Evergreen broadleaf forest	0.1	0.6	2.3	0.32	0.44	0.23
Deciduous needleleaf forest	0.1	0.6	1.3	0.34	0.5	0.16
Deciduous broadleaf forest	0.1	0.6	1.3	0.31	0.52	0.17
Mixed forest	0.1	0.6	1.7	0.25	0.52	0.22
Closed shrublands	0.1	0.6	1.8	0.31	0.49	0.21
Open shrublands	0.1	0.6	2.4	0.33	0.43	0.24
Savanna	0.1	0.6	1.7	0.36	0.45	0.19
Woody savanna	0.1	0.6	1	0.37	0.5	0.13
Grasslands	0.1	0.6	0.8	0.44	0.45	0.11
Permanent wetlands	0.1	0.6	0.8	0.44	0.45	0.11
Cropland	0.1	0.6	0.8	0.33	0.55	0.12
Urban	0.1	0.6	0.8	0.44	0.45	0.11
Cropland/natural vegetation mosaic	0.1	0.6	0.8	0.33	0.55	0.12
Permanent snow and ice	0.1	0.6	0.8	0.44	0.45	0.11
Barren	0.1	0.6	3.3	0.22	0.46	0.31

Schaperow, J. R., D. Li, S. A. Margulis, and D. P. Lettenmaier (2021), A near-global, high resolution land surface parameter dataset for the variable infiltration capacity model, *Sci Data*, *8*(1), 216, doi:10.1038/s41597-021-00999-4.

THE WALL STREET JOURNAL.

THE FUTURE OF EVERYTHING | DATA

Climate Change Data Deluge Has Scientists Scrambling for Solutions

As earth-observing satellites, aircraft and ocean buoys churn out ever-rising amounts of information about our planet, data managers turn to cloud computing and artificial intelligence

AUTHOR
ROBERT LEE HOTZ

PUBLISHED
DEC. 5, 2021 11:00 AM ET

https://www.wsj.com/articles/climate-change-data-deluge-has-scientists-scrambling-for-solutions-11638720016

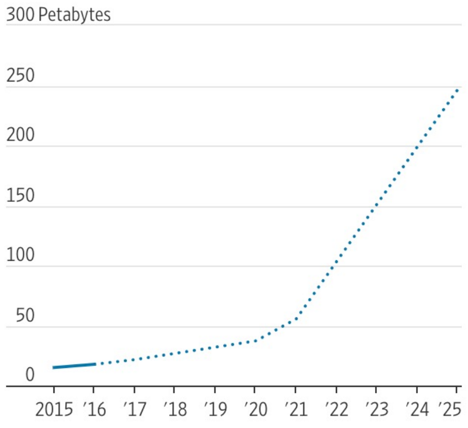
THE WALL STREET JOURNAL.

THE FUTURE OF EVERYTHING | DATA

Climate Change Data Deluge Has Scientists Scrambling for Solutions

Climate Data Explosion

The volume of data in NASA's Earth Observing System Data and Information System is expected to soar in coming years.



Source: NASA

llites, aircraft and ocean buoys churn out ever-rising amounts of information about our urn to cloud computing and artificial intelligence

AUTHOR
ROBERT LEE HOTZ

PUBLISHED
DEC. 5, 2021 11:00 AM ET

Most of our LSMs are data-hungry, yet not well designed to take advantage of the existing and new sources of observations

Most of our LSMs are data-hungry, yet not well designed to take advantage of the existing and new sources of observations

Our process parameterizations don’t integrate well across the land: unclear how to transfer knowledge between locations

the case of hydrology

1986

WATER RESOURCES RESEARCH, VOL. 22, NO. 9, PAGES 46S–58S, AUGUST 1986

Looking for Hydrologic Laws

JAMES C. I. DOOGE

Department of Engineering Hydrology, University College, Galway, Dublin, Ireland

The search for regularities in hydrologic relationships is discussed against the background of the general types of predictive models used in science. The various approaches to the study of water are compared and contrasted. The ideas discussed are illustrated by examples from the development of techniques in flood hydrology and by personal conclusions on the sources for new hypotheses in flood hydrology and the possibility of their verification.

1. INTRODUCTION

1.1. *Relation of Hydrology to Science*

Fifty years ago, there was only a handful of books and a handful of journals available to the hydrologist who wished to establish a sound scientific basis for his practical decisions. Today, there is an embarrassing abundance of books, monographs, journals and symposia proceedings clamouring for his or her attention. Is hydrology now an established science? Is hydrologic practice now firmly based on scientific principles?

This paper deals with the problems raised by the search for regularities and for laws in hydrology. In order to emphasize the challenge implicit in such a search, special attention will be paid to flood hydrology in which the enterprise is particularly difficult. It is proposed to discuss the subject within the context of predictive models and explanatory theories in science generally. Such an approach can be useful not only for the purpose of emphasizing the position of hydrology as one of the earth sciences but also because such an approach could lead to the suggestion of analogies which can be so fruitful in the construction of models and the development of theories [*Polya*, 1954].

Following an introductory section on the nature of scientific method, an outline is given of the contrasting approaches of analytical mechanics and statistical mechanics and the problems involved in dealing with systems of intermediate size. Attention is then turned to the various approaches to the study of water movement and the problem of parametrizing at a macroscale the effect of microscale processes that are not explicitly included in the macroscale model. Finally, the historical development of current methods of flood hydrology is reviewed against the background of the foregoing material. The purpose of the whole exercise is to provide the context for the formulation of a strategy for the development of a body of hydrologic knowledge that is both scientifically respectable and practically useful.

The term model is used to describe a system which is simpler than the prototype system and which can reproduce some but not all of the characteristics thereof. Accordingly, a model is related to those particular aspects of the behavior of the prototype for which understanding or prediction is required. It is important to realize that a model is not a theory and to distinguish beween models that are constructed to provide a prediction of system behavior to some specific accuracy and scientific theories developed to provide insight into the nature of the system operation. Though their function is different, predictive models and explanatory theories can be closely related to one another. An hypothesis or a confirmed theory can be used as the basis for the construction of a predictive model and in turn some predictive models which reproduce a prototype behavior accurately can provide insight for the construction of explanatory theories.

For different scientific workers (or for the same scientific worker at different times) the level of interest in explanatory theories and predictivie models may vary as shown schematically in Figure 1. A research scientist uses observations primarily as a basis of comparison between the predictions based on alternative hypotheses and combines confirmed hypotheses into a theoretical system enabling him or her to understand nature. An engineer uses observations as a check on the predictions he uses in his efforts "to control the materials and forces of nature for the use and benefit of man" (ICE Charter; see *Dennis* [1968]). Hydrology, as one of the earth sciences and as the basis of water resources development, is concerned with both of these functions. Understanding and prediction can aid the control of extreme flood events, but perfect understanding and perfect prediction would be small solace if failure to control resulted in a massive human tragedy.

1.2. *Nature of Scientific Method*

There seems at first sight to be all the difference in the world between the scientific method of the physical scientist and the efforts of the hydrologist to understand and predict extreme flood events or other hydroloical phenomena. On closer examination, however, it becomes clear that while there are very significant differences, there are also similarities and analogies that may be helpful to the hydrologist in his task. It is clearly insufficient to define scientific method as "what scientists do," but it would be equally wrong to think that all scientists under all circumstances act in accordance with what is known as the scientific method.

In his notable work *The Logic of Scientific Discovery*, *Popper* [1959] proposed falsifiability as the criterion of dermarcation of empirical science. He requires of any scientific system that 'it must be possible for an empirical scientific system to be refuted by experience.' Popper goes on to develop such principles and rules as will ensure the testability, i.e., the falsifiability of scientific statements.

Paper number 6W0295.
0043-1397/86/006W-0295$05.00

the case of hydrology

Dooge, J. C. I. (1986), Looking for hydrologic laws, *Water Resour Res*, *22*(9), S46-S58, doi:10.1029/WR022i09Sp0046S.

1986

WATER RESOURCES RESEARCH, VOL. 22, NO. 9, PAGES 46S–58S, AUGUST 1986

Looking for Hydrologic Laws

JAMES C. I. DOOGE

Department of Engineering Hydrology, University College, Galway, Dublin, Ireland

The search for regularities in hydrologic relationships is discussed against the background of the general types of predictive models used in science. The various approaches to the study of water are compared and contrasted. The ideas discussed are illustrated by examples from the development of techniques in flood hydrology and by personal conclusions on the sources for new hypotheses in flood hydrology and the possibility of their verification.

1. INTRODUCTION

1.1. *Relation of Hydrology to Science*

Fifty years ago, there was only a handful of books and a handful of journals available to the hydrologist who wished to establish a sound scientific basis for his practical decisions. Today, there is an embarrassing abundance of books, monographs, journals and symposia proceedings clamouring for his or her attention. Is hydrology now an established science? Is hydrologic practice now firmly based on scientific principles?

This paper deals with the problems raised by the search for regularities and for laws in hydrology. In order to emphasize the challenge implicit in such a search, special attention will be paid to flood hydrology in which the enterprise is particularly difficult. It is proposed to discuss the subject within the context of predictive models and explanatory theories in science generally. Such an approach can be useful not only for the purpose of emphasizing the position of hydrology as one of the earth sciences but also because such an approach could lead to the suggestion of analogies which can be so fruitful in the construction of models and the development of theories [*Polya*, 1954].

Following an introductory section on the nature of scientific method, an outline is given of the contrasting approaches of analytical mechanics and statistical mechanics and the problems involved in dealing with systems of intermediate size. Attention is then turned to the various approaches to the study of water movement and the problem of parametrizing at a macroscale the effect of microscale processes that are not explicitly included in the macroscale model. Finally, the historical development of current methods of flood hydrology is reviewed against the background of the foregoing material. The purpose of the whole exercise is to provide the context for the formulation of a strategy for the development of a body of hydrologic knowledge that is both scientifically respectable and practically useful.

The term model is used to describe a system which is simpler than the prototype system and which can reproduce some but not all of the characteristics thereof. Accordingly, a model is related to those particular aspects of the behavior of the prototype for which understanding or prediction is required. It is important to realize that a model is not a theory and to distinguish beween models that are constructed to provide a prediction of system behavior to some specific accuracy and scientific theories developed to provide insight into the nature of the system operation. Though their function is different, predictive models and explanatory theories can be closely related to one another. An hypothesis or a confirmed theory can be used as the basis for the construction of a predictive model and in turn some predictive models which reproduce a prototype behavior accurately can provide insight for the construction of explanatory theories.

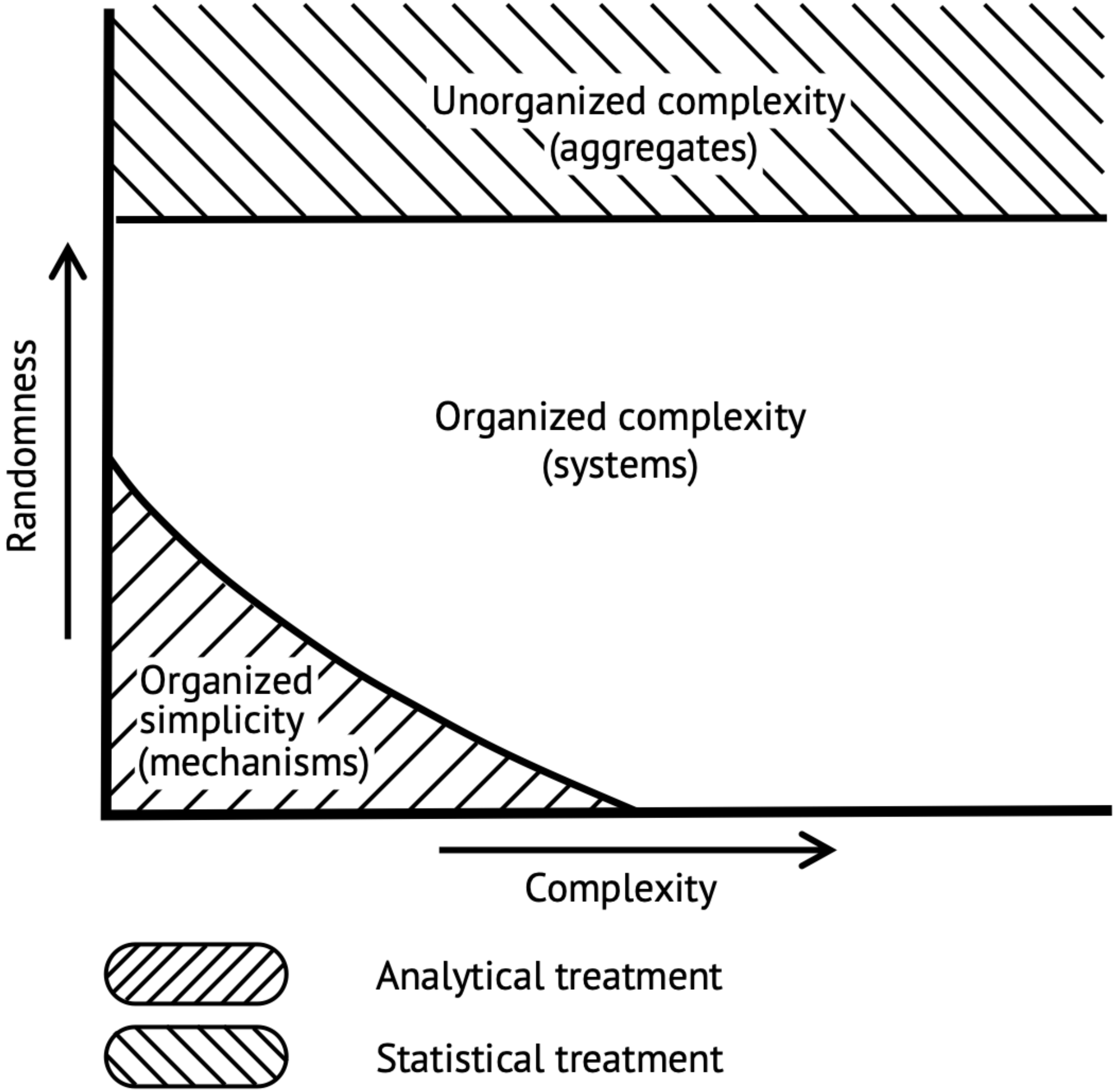
For different scientific workers (or for the same scientific worker at different times) the level of interest in explanatory theories and predictivie models may vary as shown schematically in Figure 1. A research scientist uses observations primarily as a basis of comparison between the predictions based on alternative hypotheses and combines confirmed hypotheses into a theoretical system enabling him or her to understand nature. An engineer uses observations as a check on the predictions he uses in his efforts "to control the materials and forces of nature for the use and benefit of man" (ICE Charter; see *Dennis* [1968]). Hydrology, as one of the earth sciences and as the basis of water resources development, is concerned with both of these functions. Understanding and prediction can aid the control of extreme flood events, but perfect understanding and perfect prediction would be small solace if failure to control resulted in a massive human tragedy.

1.2. *Nature of Scientific Method*

There seems at first sight to be all the difference in the world between the scientific method of the physical scientist and the efforts of the hydrologist to understand and predict extreme flood events or other hydroloical phenomena. On closer examination, however, it becomes clear that while there are very significant differences, there are also similarities and analogies that may be helpful to the hydrologist in his task. It is clearly insufficient to define scientific method as "what scientists do," but it would be equally wrong to think that all scientists under all circumstances act in accordance with what is known as the scientific method.

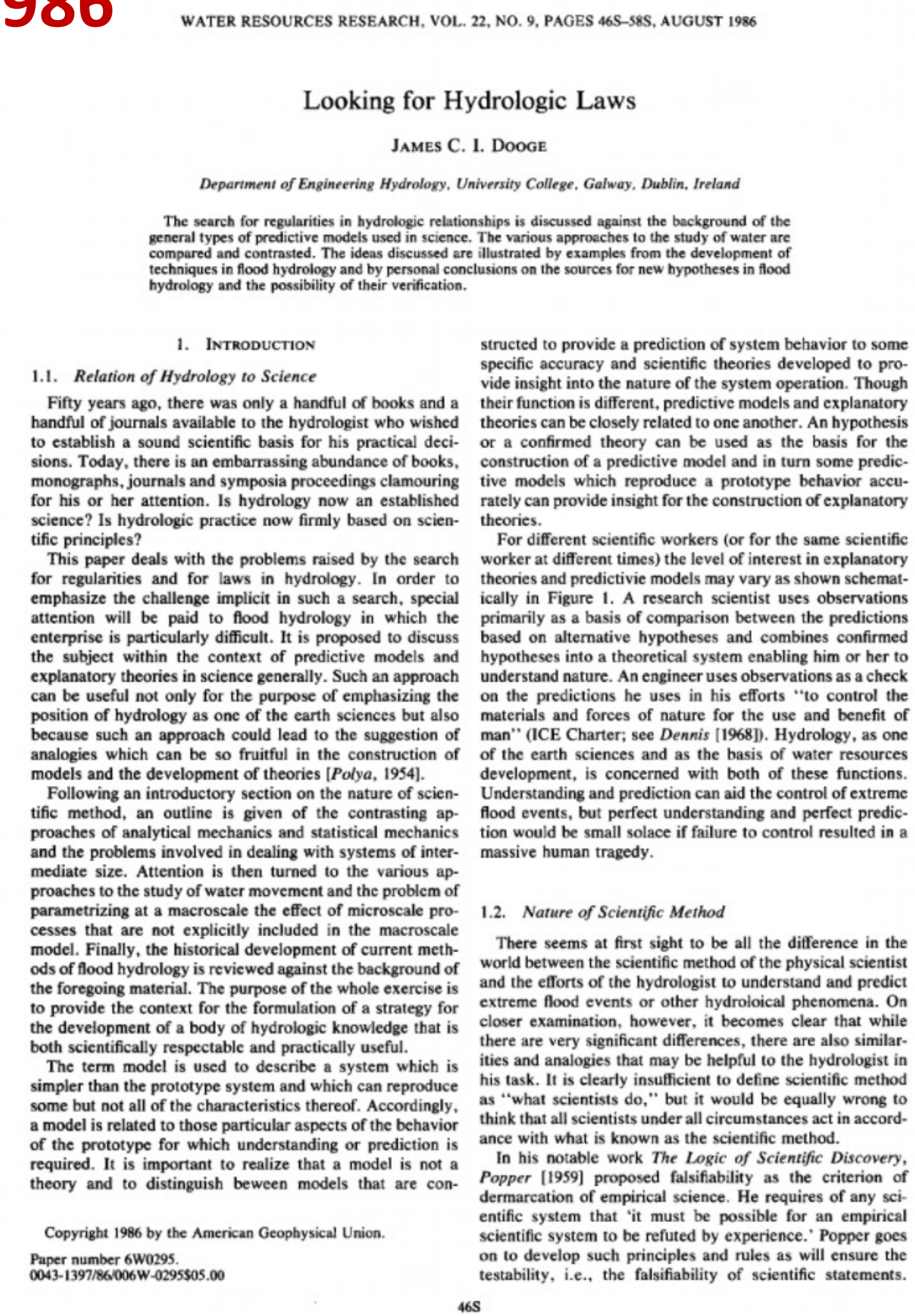
In his notable work *The Logic of Scientific Discovery*, *Popper* [1959] proposed falsifiability as the criterion of dermarcation of empirical science. He requires of any scientific system that 'it must be possible for an empirical scientific system to be refuted by experience.' Popper goes on to develop such principles and rules as will ensure the testability, i.e., the falsifiability of scientific statements.

Paper number 6W0295.
0043-1397/86/006W-0295$05.00



Dooge, J. C. I. (1986), Looking for hydrologic laws, *Water Resour Res*, *22*(9), S46-S58, doi:10.1029/WR022i09Sp0046S.

1986



WATER RESOURCES RESEARCH, VOL. 22, NO. 9, PAGES 46S–58S, AUGUST 1986

Looking for Hydrologic Laws

JAMES C. I. DOOGE

Department of Engineering Hydrology, University College, Galway, Dublin, Ireland

The search for regularities in hydrologic relationships is discussed against the background of the general types of predictive models used in science. The various approaches to the study of water are compared and contrasted. The ideas discussed are illustrated by examples from the development of techniques in flood hydrology and by personal conclusions on the sources for new hypotheses in flood hydrology and the possibility of their verification.

1. INTRODUCTION

1.1. *Relation of Hydrology to Science*

Fifty years ago, there was only a handful of books and a handful of journals available to the hydrologist who wished to establish a sound scientific basis for his practical decisions. Today, there is an embarrassing abundance of books, monographs, journals and symposia proceedings clamouring for his or her attention. Is hydrology now an established science? Is hydrologic practice now firmly based on scientific principles?

This paper deals with the problems raised by the search for regularities and for laws in hydrology. In order to emphasize the challenge implicit in such a search, special attention will be paid to flood hydrology in which the enterprise is particularly difficult. It is proposed to discuss the subject within the context of predictive models and explanatory theories in science generally. Such an approach can be useful not only for the purpose of emphasizing the position of hydrology as one of the earth sciences but also because such an approach could lead to the suggestion of analogies which can be so fruitful in the construction of models and the development of theories [*Polya*, 1954].

Following an introductory section on the nature of scientific method, an outline is given of the contrasting approaches of analytical mechanics and statistical mechanics and the problems involved in dealing with systems of intermediate size. Attention is then turned to the various approaches to the study of water movement and the problem of parametrizing at a macroscale the effect of microscale processes that are not explicitly included in the macroscale model. Finally, the historical development of current methods of flood hydrology is reviewed against the background of the foregoing material. The purpose of the whole exercise is to provide the context for the formulation of a strategy for the development of a body of hydrologic knowledge that is both scientifically respectable and practically useful.

The term model is used to describe a system which is simpler than the prototype system and which can reproduce some but not all of the characteristics thereof. Accordingly, a model is related to those particular aspects of the behavior of the prototype for which understanding or prediction is required. It is important to realize that a model is not a theory and to distinguish beween models that are constructed to provide a prediction of system behavior to some specific accuracy and scientific theories developed to provide insight into the nature of the system operation. Though their function is different, predictive models and explanatory theories can be closely related to one another. An hypothesis or a confirmed theory can be used as the basis for the construction of a predictive model and in turn some predictive models which reproduce a prototype behavior accurately can provide insight for the construction of explanatory theories.

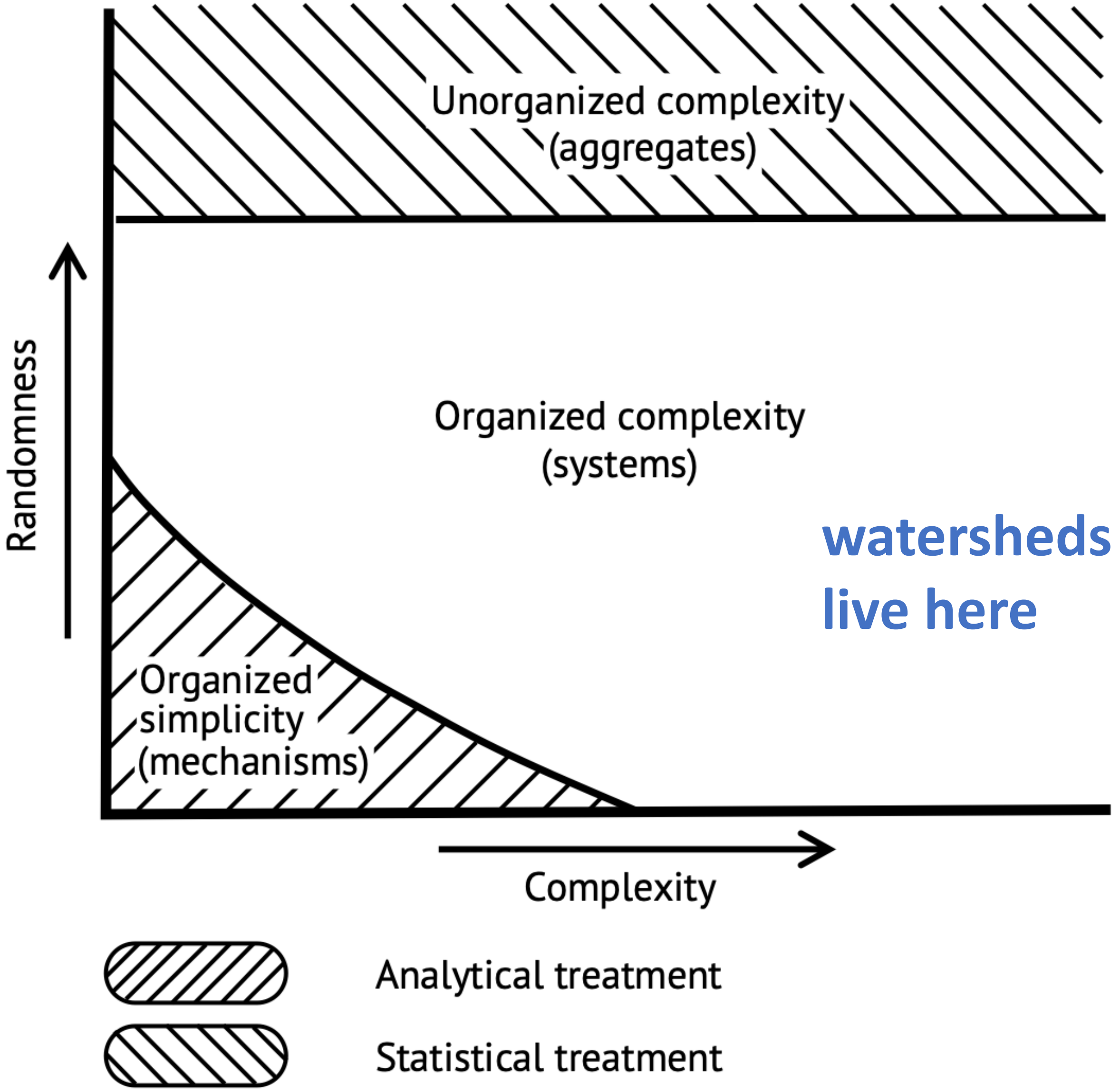
For different scientific workers (or for the same scientific worker at different times) the level of interest in explanatory theories and predictivie models may vary as shown schematically in Figure 1. A research scientist uses observations primarily as a basis of comparison between the predictions based on alternative hypotheses and combines confirmed hypotheses into a theoretical system enabling him or her to understand nature. An engineer uses observations as a check on the predictions he uses in his efforts "to control the materials and forces of nature for the use and benefit of man" (ICE Charter; see *Dennis* [1968]). Hydrology, as one of the earth sciences and as the basis of water resources development, is concerned with both of these functions. Understanding and prediction can aid the control of extreme flood events, but perfect understanding and perfect prediction would be small solace if failure to control resulted in a massive human tragedy.

1.2. *Nature of Scientific Method*

There seems at first sight to be all the difference in the world between the scientific method of the physical scientist and the efforts of the hydrologist to understand and predict extreme flood events or other hydroloical phenomena. On closer examination, however, it becomes clear that while there are very significant differences, there are also similarities and analogies that may be helpful to the hydrologist in his task. It is clearly insufficient to define scientific method as "what scientists do," but it would be equally wrong to think that all scientists under all circumstances act in accordance with what is known as the scientific method.

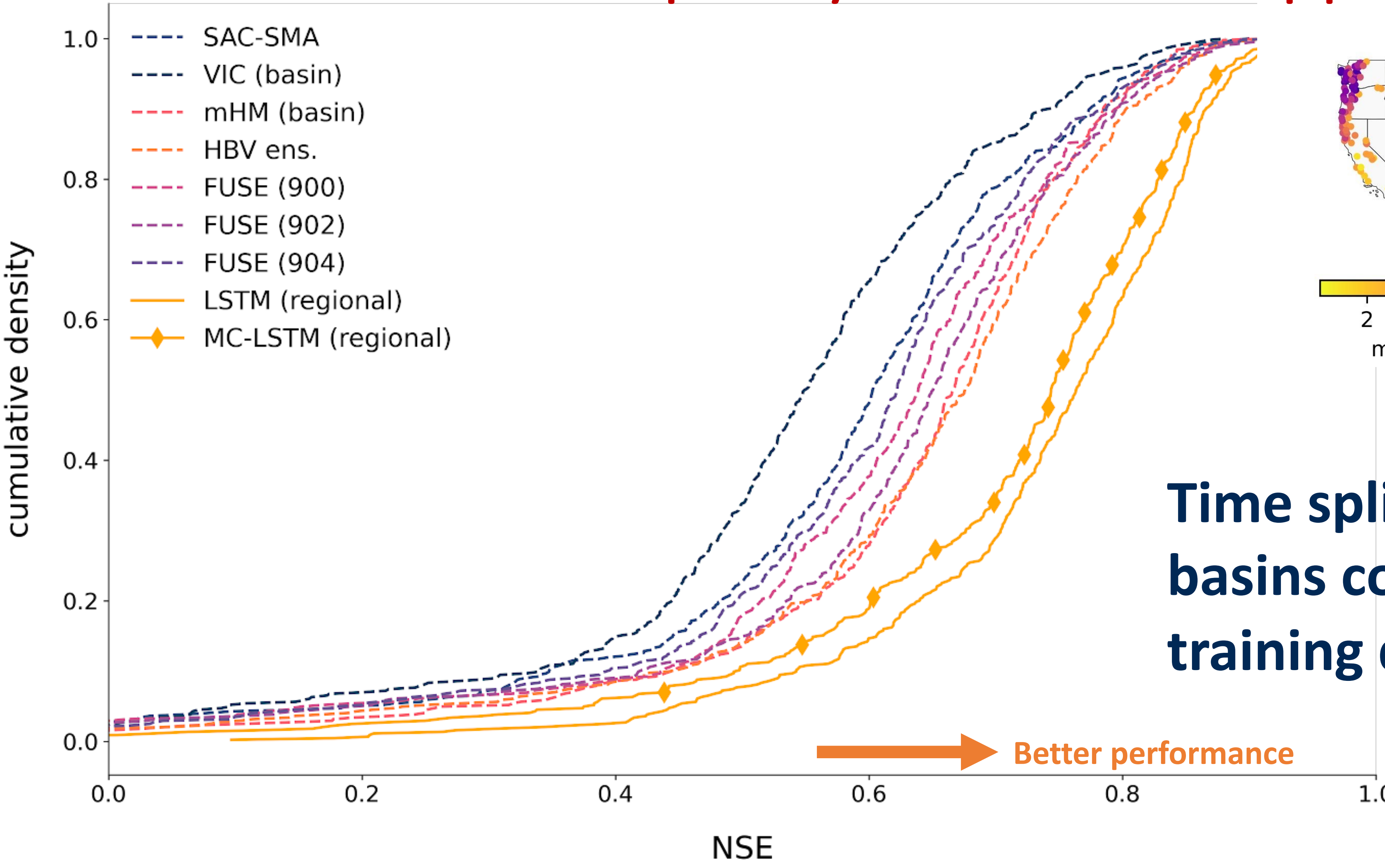
In his notable work *The Logic of Scientific Discovery*, *Popper* [1959] proposed falsifiability as the criterion of dermarcation of empirical science. He requires of any scientific system that 'it must be possible for an empirical scientific system to be refuted by experience.' Popper goes on to develop such principles and rules as will ensure the testability, i.e., the falsifiability of scientific statements.

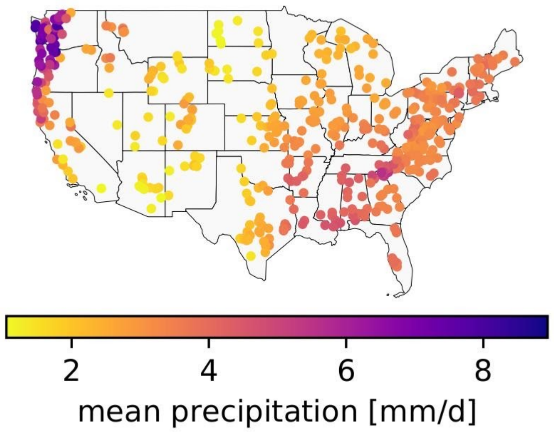
Paper number 6W0295.
0043-1397/86/006W-0295$05.00



Dooge, J. C. I. (1986), Looking for hydrologic laws, *Water Resour Res*, *22*(9), S46-S58, doi:10.1029/WR022i09Sp0046S.

Dataset & model: A purely data-driven approach

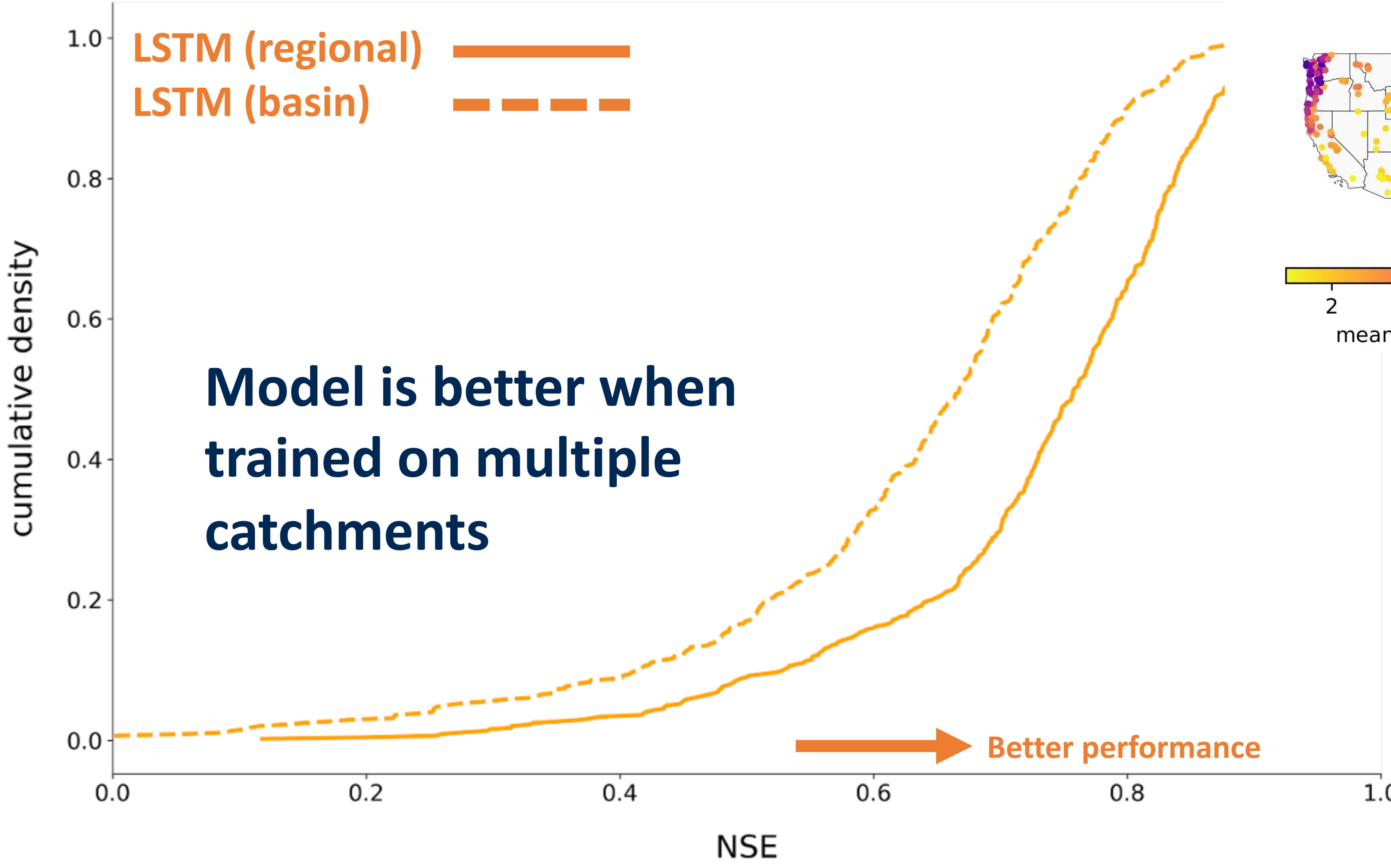


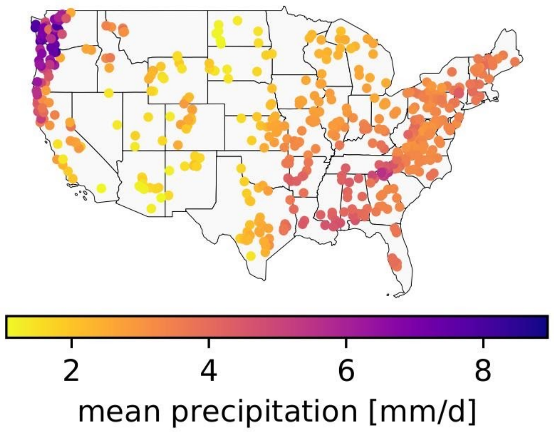


Time split validation: all basins contributed training data.

See for example, Kratzert et al., 2019, doi:[10.5194/hess-2019-368](https://doi.org/10.5194/hess-2019-368)

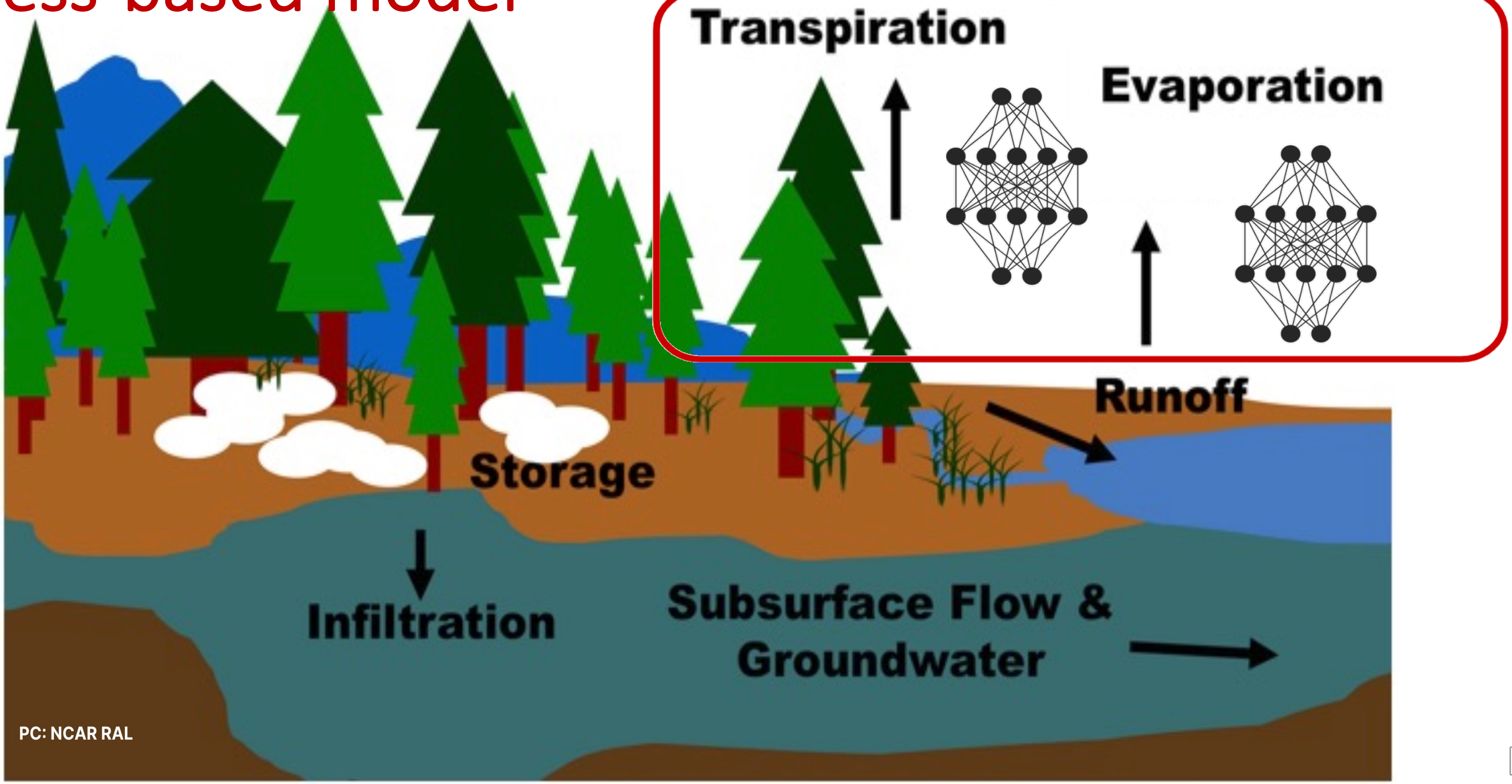
Learning a general model



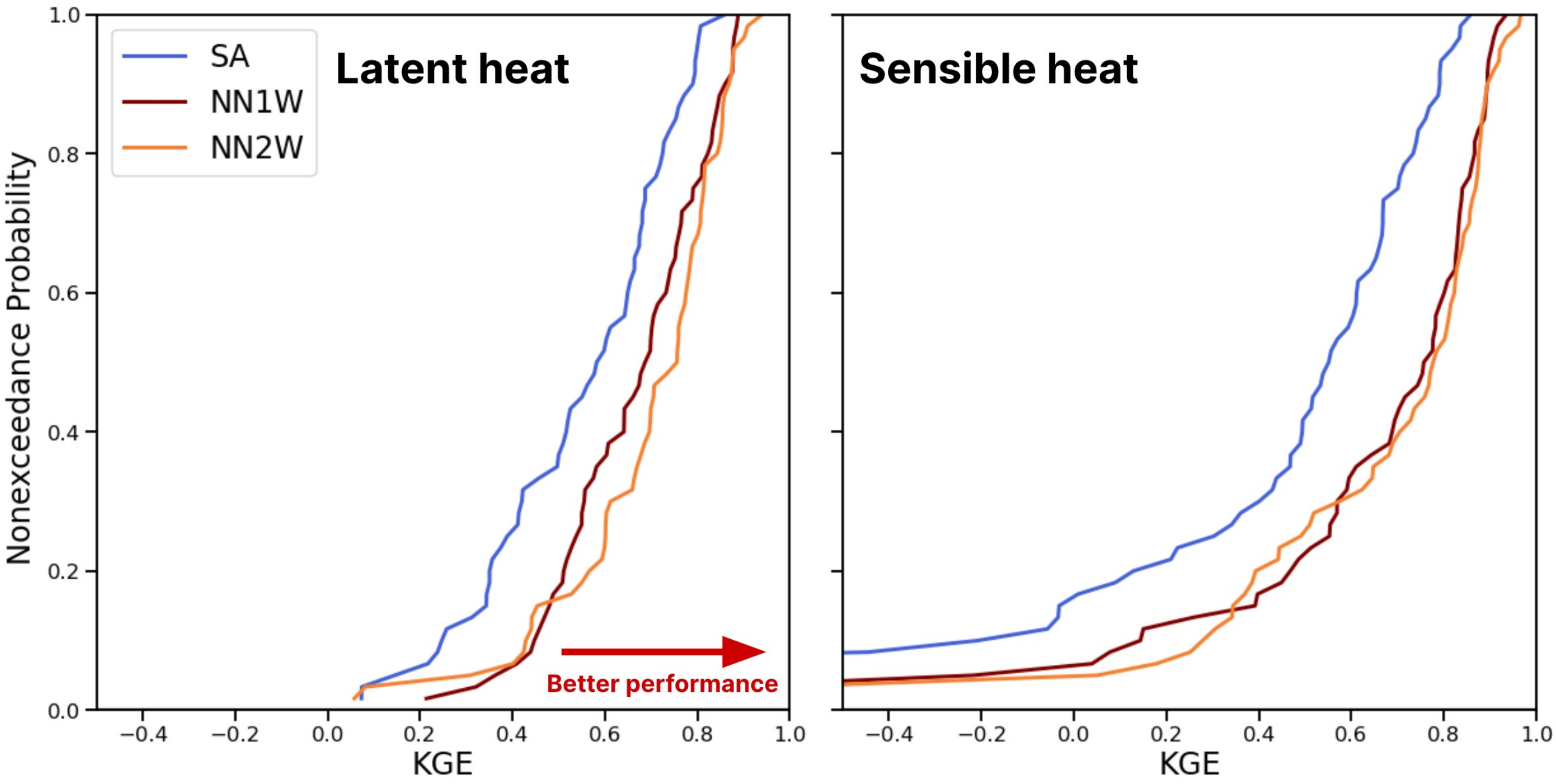


Model is better when trained on multiple catchments

Replacing turbulent heat flux parameterizations in a process-based model

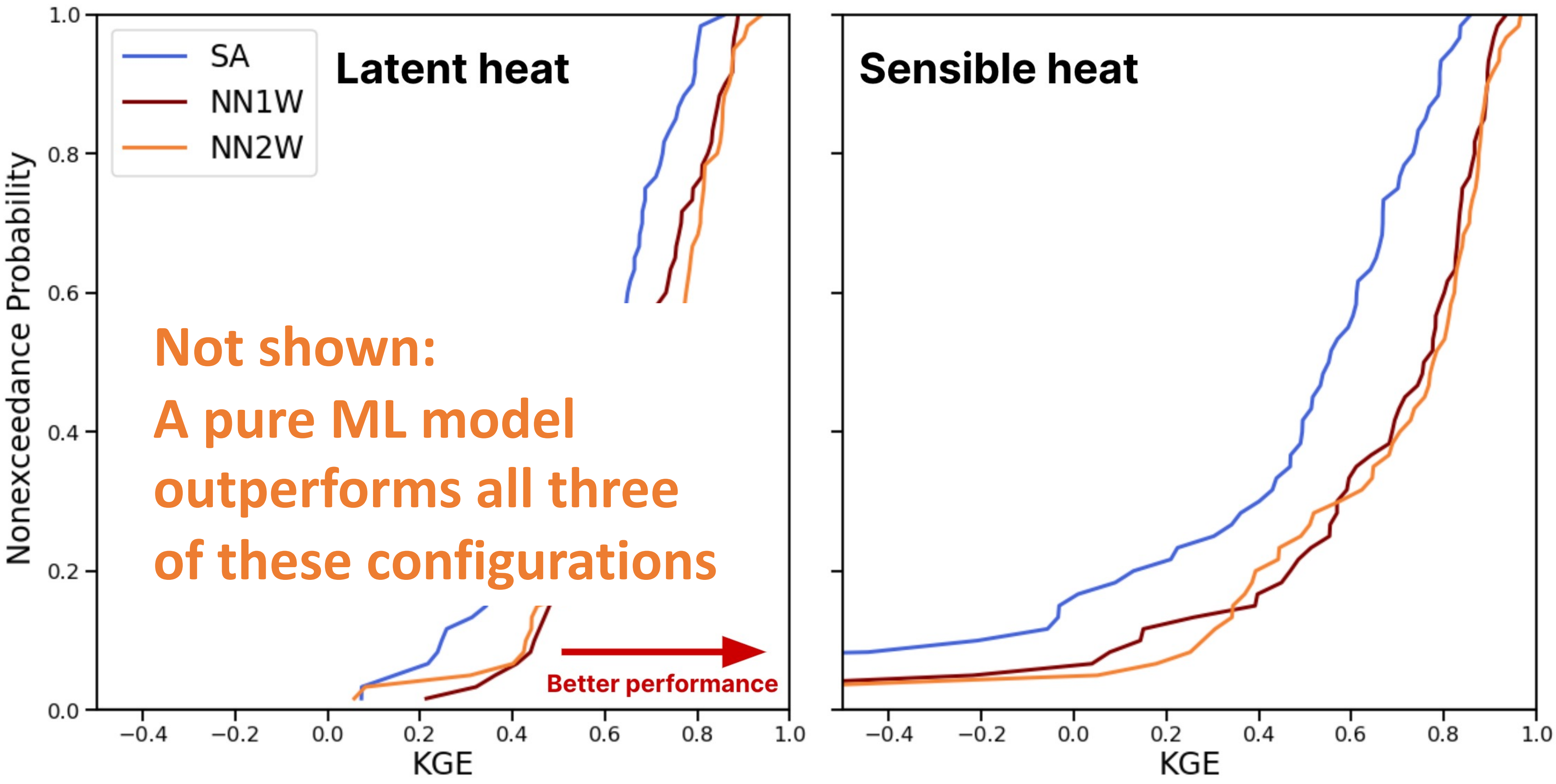


SUMMA & ML turbulent heat parameterization outperforms original SUMMA



Bennett and Nijssen, 2021: doi:10.1029/2020WR029328.

SUMMA & ML turbulent heat parameterization outperforms original SUMMA



Bennett and Nijssen, 2021: doi:10.1029/2020WR029328.

Learning from Observations:

The Case for a New Generation of Land Surface Models

Learning from Observations:

The Case for a New Generation of Land Surface Models

- Take advantage of new observations
- Generalize across locations
- Learn land surface processes at scale

In hydrology: Prediction in ungauged basins

Learning from Observations:

The Case for a New Generation of Land Surface Models

- Take advantage of new observations
- Generalize across locations

  In hydrology: Prediction in ungauged basins
- Learn land surface processes at scale

Make a concerted effort to adopt the use of ML approaches in LSMs

Learning from Observations:

The Case for a New Generation of Land Surface Models

- Take advantage of new observations
- Generalize across locations — In hydrology: Prediction in ungauged basins
- Learn land surface processes at scale

Make a concerted effort to adopt the use of ML approaches in LSMs

FUNDING & RESOURCES

AI4ESP Home White Papers Workshop News Contact Us

AI4ESP Artificial Intelligence for Earth System Predictability

A multi-lab initiative working with the Earth and Environmental Systems Science Division (EESSD) of the Office of Biological and Environmental Research (BER) to develop a new paradigm for Earth system predictability focused on enabling artificial intelligence across fiel[d,] lab, modeling, and analysis activities.

COLUMBIA CLIMATE SCHOOL
Climate, Earth, and Society

State of the Planet

AGRICULTURE CLIMATE EARTH SCIENCES ECOLOGY ENERGY HEALTH SUSTAINABILITY

CLIMATE, GENERAL EARTH INSTITUTE

Columbia to Launch $25 Million AI-Based Climate Modeling Center

BY KIM MARTINEAU | SEPTEMBER 9, 2021 9 Comments

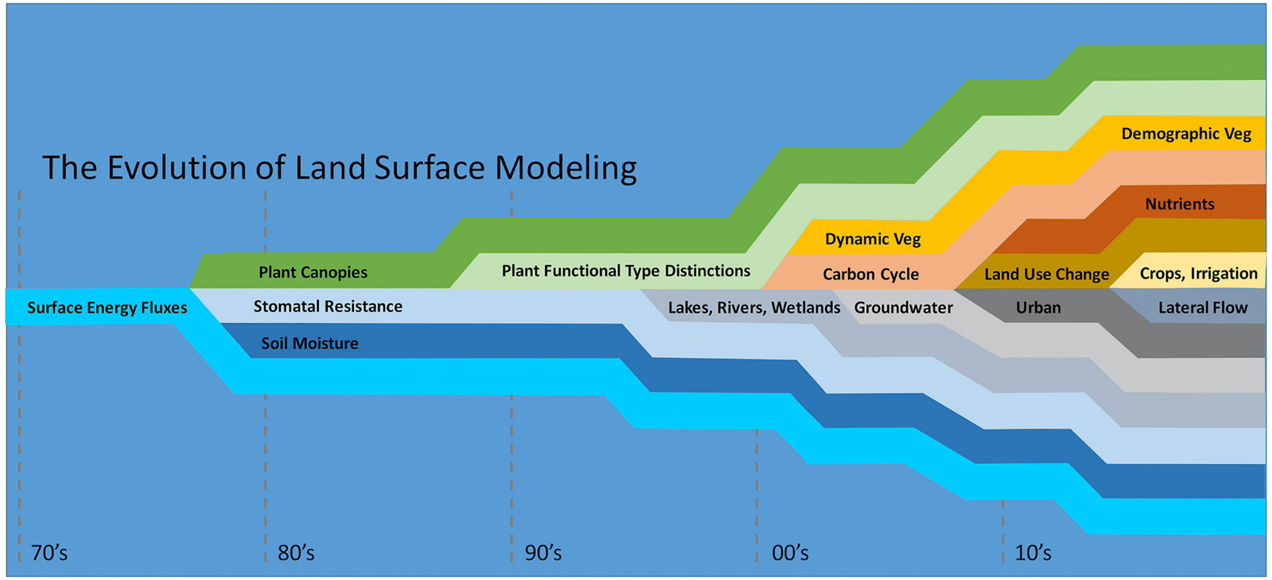
AI2 Allen Institute for AI

Climate Modeling About Papers Team

uses GFDL model

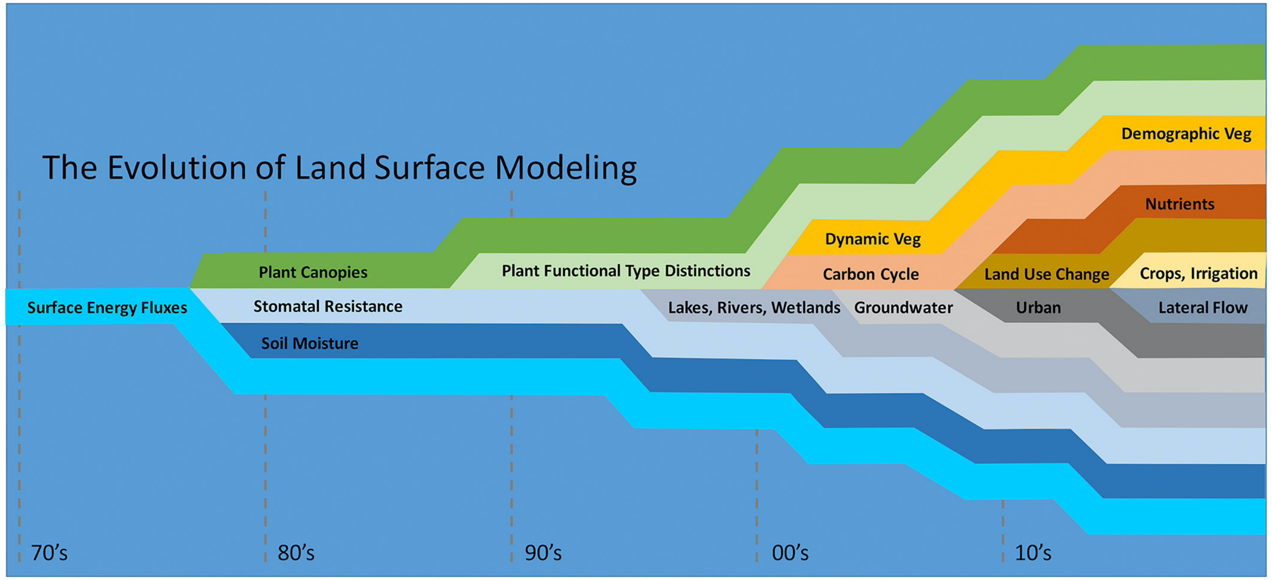
Climate Modeling for the future of the planet

Some of this is already happening



50+ years of development:

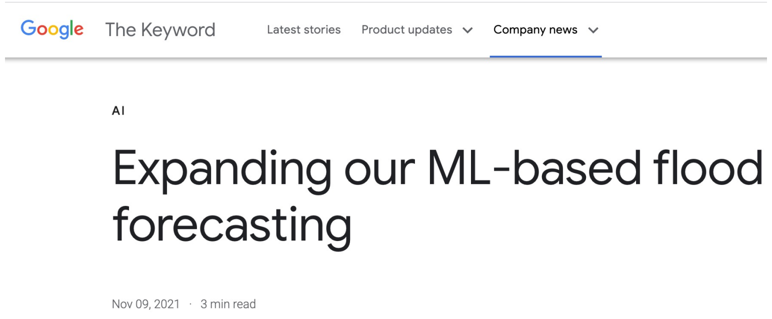
- Understanding
- People
- Models
- Code



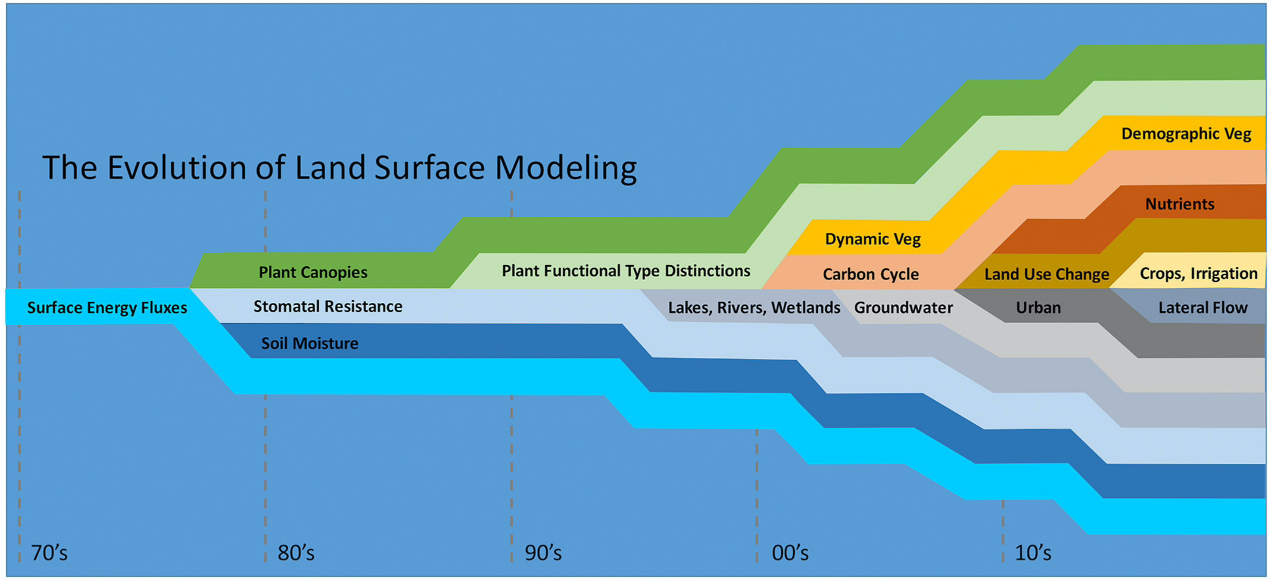
50+ years of development:

- Understanding
- People
- Models
- Code

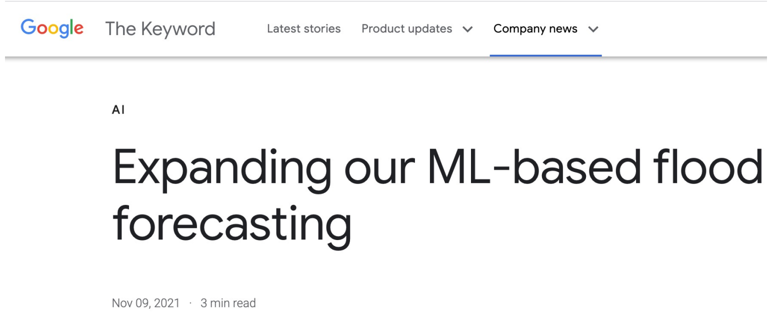
https://www.upstream.tech/hydroforecast



https://blog.google/technology/ai/expanding-our-ml-based-flood-forecasting/



https://www.upstream.tech/hydroforecast

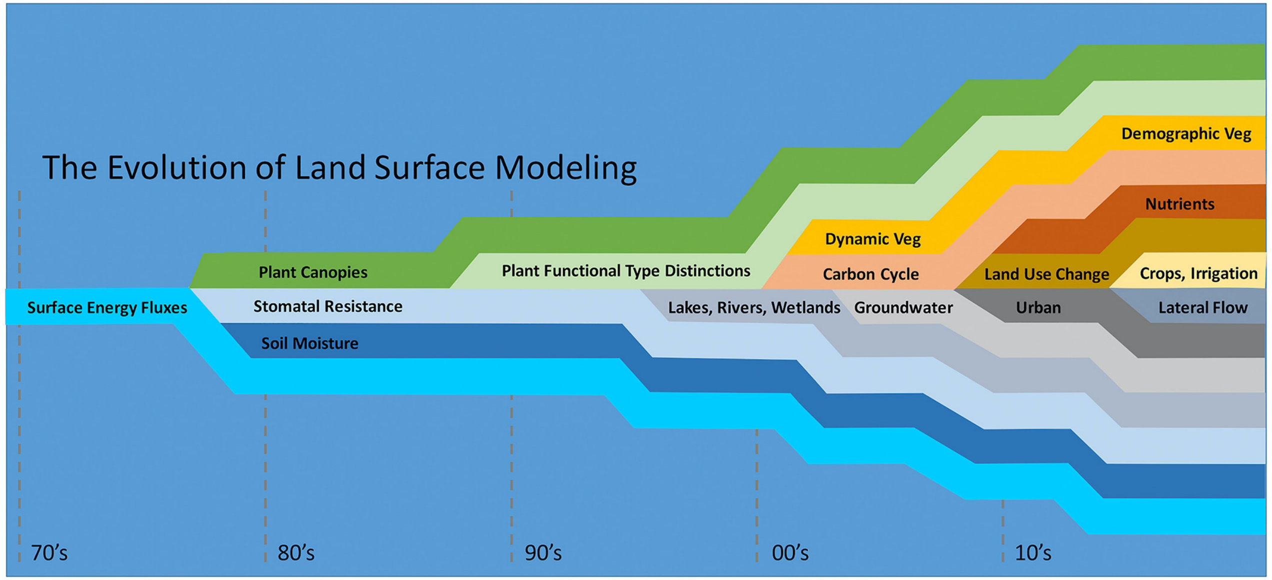


https://blog.google/technology/ai/expanding-our-ml-based-flood-forecasting/

50+ years of development:

- Understanding
- People
- Models
- Code

https://research.ibm.com/blog/icml-predicting-climate-extremes



UPSTREAM TECH

HydroForecast

Water and weather intelligence, at your fingertips

Manage climate and water risk, and see more clearly into the future with AI-powered forecasts for planning and operations.

https://www.upstream.tech/hydroforecast

Google The Keyword Latest stories Product updates Company news

AI

Expanding our ML-based flood forecasting

Nov 09, 2021 · 3 min read

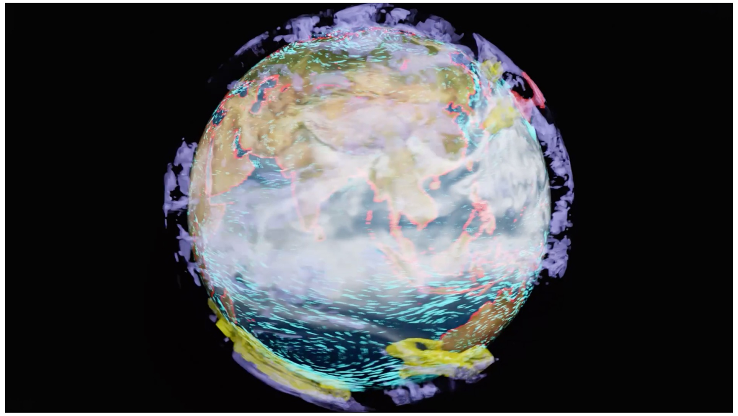
https://blog.google/technology/ai/expanding-our-ml-based-flood-forecasting/

50+ years of development:

Understanding
People
Models
Code

NVIDIA to Build Earth-2 Supercomputer to See Our Future

November 12, 2021 by JENSEN HUANG



https://blogs.nvidia.com/blog/2021/11/12/earth-2-supercomputer/

Research 6 minute read

IBM

How AI can help predict climate extremes

https://research.ibm.com/blog/icml-predicting-climate-extremes

What we need:

Ways to incorporate machine learning to take advantage of new data and existing knowledge

Targeted efforts to explore the best use of new data

Critical evaluation of successes and failures of existing models

Identifying opportunities for each modeling approach

Streamlined model workflows and data sets

Collaborations between earth scientists and data scientists

Training and training resources in data science and machine learning

Learning from Observations:

The Case for a New Generation of Land Surface Models