The challenge of rainfall estimation and prediction across scales

LEARNING FROM PATTERNS

Efi Foufoula-Georgiou
University of California, Irvine

Langbein Lecture
2019 AGU
Walter Langbein – the keen observer

Courtesy of Bill Dietrich
The Concept of Entropy in Landscape Evolution

By LUNA B. LEOPOLD and WALTER B. LANGBEIN

THEORETICAL PAPERS IN THE HYDROLOGIC AND GEOMORPHIC SCIENCES

GEOLOGICAL SURVEY PROFESSIONAL PAPER 500-A

STABLE STATE IN THE STOCHASTIC THEORY OF LONGITUDINAL RIVER PROFILE DEVELOPMENT

A. E. SCHEIDEGGER and W. B. LANGBEIN

River Meanders—Theory of Minimum Variance

By WALTER B. LANGBEIN and LUNA B. LEOPOLD

PHYSIOGRAPHIC AND HYDRAULIC STUDIES OF RIVERS

GEOLOGICAL SURVEY PROFESSIONAL PAPER 422-H

The geometry of a meander is that of a random walk whose most frequent form minimizes the sum of the squares of the changes in direction in each unit length. Changes in direction closely approximately a sine function of channel distance. Depth, velocity, and slope are adjusted so as to decrease the variance of shear and the friction factor in a meander over that in an otherwise comparable straight reach of the same river.
ANNUAL FLOODS AND THE PARTIAL-DURATION FLOOD SERIES

W. B. Langbein

Abstract -- Flood data are ordinarily listed either in annual-flood series or in a partial-duration series. If the expectancy of a flood in the duration series $\epsilon$ is known, then the probability of that flood being an annual flood is shown to be $e^{-\epsilon}$. From this relationship it is possible to transform recurrence intervals in the partial duration series to those in the annual-flood series. It is shown that for equivalent floods, the recurrence intervals in the partial-duration series are smaller than in the annual-flood series, but that the difference becomes inconsequential for floods greater than about five-year recurrence interval.

$$\lim_{n \to \infty} \left[ 1 - \frac{\epsilon}{n} \right]^n = \exp \left[ -\epsilon \right]$$

Transform recurrence intervals in the partial duration series to those in the max annual flood series
A PRIMER ON WATER

1960

Walter Langbein – the science communicator
Walter Langbein – the visionary

USGS Water Program
National network of hydrologic data
Flood Insurance Program
Intern. Hydrologic Decade (1965-1974; UNESCO)
Intern. Association for Hydrologic Sciences (IAHS)
World Meteorological Organization (WMO)
Walter Langbein – the visionary

USGS Water Program
National network of hydrologic data
Flood Insurance Program
Intern. Hydrologic Decade (1965-1974; UNESCO)
Intern. Association for Hydrologic Sciences (IAHS)
World Meteorological Organization (WMO)

…“Science is built up with facts, as a house is with stones. But a collection of facts is no more a science that a heap of stones is a house”
– HENRY POINCARE
Walter B. Langbein (1907–1982)

Walter Langbein was dedicated to science that benefited the public good and was known as a versatile and talented hydrologist. Born in New Jersey in 1907, he obtained his civil engineering degree in 1931 from Cooper Union while attending night classes and working for a construction company. In 1935, he joined the U.S. Geological Survey (USGS) in Albany, but within a year he was transferred to the national headquarters, where he served as a research engineer and senior scientist for the rest of his life.

Langbein's contributions to the field of hydrology are extensive. His 1955 book, Floods, with W. G. Hoyt, was instrumental in the development of the National Flood Insurance Program. He developed methods in flood hydrology and the application of statistical methods to the analysis of hydrologic data. He studied evaporation from water bodies, varying from small stock ponds on the Navajo Reservation to Lake Mead. He studied infiltration in stream channels and its effect on flood wave passage. As early as 1944, Langbein was interested in the use of hydrologic data for the estimation of climate change. With Luna Leopold, he worked to establish a national program in water resources research, which led to the development of the Office of Water Resources Research within USGS. Langbein was instrumental in founding the International Hydrologic Decade (1965–1974), and his participation in the decade focused attention on the determination of the worth of hydrologic data for water resources development. The theory of scientific network design for water data networks evolved from his work.

Walter Langbein was awarded the William Bowie and Robert E. Horton Medals from the American Geophysical Union, the J. C. Stevens Award of the American Society of Civil Engineers, the Distinguished Service Award of the Department of the Interior, and the Warren Prize of the National Academy of Sciences. He and Professor Korzun of the Soviet Union were named corecipients of the International Prize in Hydrology, awarded by the International Association of Hydrologic Sciences.

Langbein once remarked that one's professional career is a race against obsolescence. As noted by others, any hydrologist would claim that Walter B. Langbein clearly won the race.
Precipitation estimation and prediction

Meandering and braided rivers

Tributary and distributary Networks

Human dominated landscapes
Driving scientific questions

1. How do physics organize precipitation systems across spatio-temporal scales?

2. How can this organization be used to improve estimation, modeling and prediction at local to global scales?

3. How can we gain mechanistic process understanding from landscape patterns and form?

4. How do perturbations propagate through a complex ecohydrological system determining its vulnerability to change?
Our data: multi-sensor observations

- NEXRAD
- TRMM/GPM
- LIDAR
- LANDSAT
Our data: Lab experiments

St. Anthony Falls Laboratory, Univ. of Minnesota
Arvind Singh, Vamsi Ganti, Victor Sapozhnikov
Across processes & scales

Plot scale / single river scale
- Meander bends
- Cutoff effects
- Residence/travel time
- Hillslope transport
- Landscape evolution
- Cascade of hydrology to ecology/water quality
- Wetlands for water quality

Watershed scale

Continental scale
- River deltas
  - Topology/Dynamics
  - Process from form
  - Optimality Principle
  - Vulnerability
- Precipitation retrieval
- Trends in extremes
- Large scale dynamics to local precipitation

Global scale
From raw data to quantitative patterns ...

\( \Omega \): Surface described by the regularized LIDAR data through nonlinear filtering.

*Cost function* \( \psi \): cost of traveling on the curve \( C \).

*Geodesic curve* curve with minimal cost, among all possible curved connecting the two point \( a \) and \( b \)

\[
g(a,b) := \arg \min_{C \in \Omega} \int_a^b \psi(s) \, ds
\]

Example of river network extraction on Skunk Creek, South Fork Eel River basin, CA

GeoNet Toolbox

Geomorphologically Inspired image Processing

Paola Passalacqua’s Group

Passalacqua et al. 2010, 2012
Ucayali River

Mining Landsat archives to resolve bend scale river dynamics

RivMAP Toolbox

Jon Schwenk's group

Schwenk et al., 2015, 2016a,b, 2017
The life of a meander bend…

Does the shape of an oxbow lake carry the signature of its forming dynamics?

Does process nonlinearity express itself on the static planform geometry?

How far upstream and downstream of cutoff perturbations propagate?

Schwenk et al., 2015, 2016a,b, 2017
The complexity of river deltas…

-- What physical processes are recorded in delta channel network topology?

-- Can a quantitative framework for delta classification be built based on suitable metrics?

-- Is there an optimality principle behind the self-organization of deltas?

Entropy and optimality in river deltas


Tejedor et al., 2015a,b, 2016, 2017a,b, 2018
The complexity of river deltas...

Coupled processes

Multi-layer Networks

Supra-Adjacency Matrix

\[ \mathcal{A} = \begin{pmatrix} A^c & I \\ I & A^l \end{pmatrix} \]

Supra-Laplacian Matrix

\[ \mathcal{L} = \begin{pmatrix} D_c L_c^c + D_x I & -D_x I \\ -D_x I & D_l L_l^l + D_x I \end{pmatrix} \]

Tejedor et al., 2015a,b, 2016, 2017a,b, 2018
Intensively managed landscapes

Transition from hay and small grains to soybeans changed the eco-hydrology of the system

(Foto: S. Levine, B. Call, P. Belmont)
w/ Belmont, Hansen, Grant, Wilcock, Finlay

Foufoula-Georgiou et al., 2015, WRR

Czuba et al., 2014, 2015, 2017
Today’s focus:

RAINFALL

1. Global estimation from space
2. Seasonal prediction
Walter Langbein – the visionary

“... Precipitation stations are more numerous where people live ... than where precipitation is more variable and therefore most important to record.”
How much of the Earth’s surface is covered by raingages?
Global precipitation

Water and energy cycles
Hydrologic prediction in remote places
Validating climate models
Detecting changes and trends

Credit: NASA
How do we observe precipitation from space?

The GEO-IR constellation
(NOAA-NESDIS, EUMETSAT, JMA)

- 5 IR imagers for a quasi-global coverage

One observation every 15-30 mis
How do we observe precipitation from space?

The LEO-GPM constellation

- 5 conical-scan MW imagers
- 8 cross-track MW sounders
- 1 Dual-frequency Precipitation Radar

One observation every 2-4 hrs

Credit: NASA
Multispectral microwave signature

- 10.6 GHz vertically polarized
- 10.6 GHz horizontally polarized
- 18.7 GHz vertically polarized
- 18.7 GHz horizontally polarized
- 23 GHz vertically polarized
- 37 GHz vertically polarized
- 37 GHz horizontally polarized
- 89 GHz vertically polarized
- 89 GHz horizontally polarized
- 166 GHz vertically polarized
- 166 GHz horizontally polarized
- 183+/3 GHz vertically polarized
- 183+/-7 GHz vertically polarized

- 10.6V
- 10.6H
- 18.7V
- 18.7V
- 23V
- 23V
- 37h
- 89V
- 89H
- 166V
- 166H
- 183+/3
- 183+/-7
Multispectral microwave signature

Brightness Temperature (TB)
Retrieval is an Inverse Problem

Radiometric signature at the top of the atmosphere
Retrieval is an Inverse Problem

Radiometric signature at the top of the atmosphere

atmospheric water content
surface precipitation rate
GPM core satellite
Step 1: Use GPM Satellite to derive set of “Observed” profiles that define an a-priori database of possible rain structures.

Step 2: Compare observed Tb to Database Tb. Select and average matching pairs.

\[ J_i = \exp \left\{ -\frac{1}{2} \left[ \mathbf{t}_b^o - \mathbf{t}_b(R_i) \right]^T (O + S)^{-1} \left[ \mathbf{t}_b^o - \mathbf{t}_b(R_i) \right] \right\} \]
How accurate are these retrievals globally?

https://pmm.nasa.gov/extreme-weather
THE CLASSICAL APPROACH IN COMPARING/VALIDATING RETRIEVALS
THE CLASSICAL APPROACH IN COMPARING/VALIDATING RETRIEVALS
THE CLASSICAL APPROACH IN COMPARING/VALIDATING RETRIEVALS
THE CLASSICAL APPROACH IN COMPARING/VALIDATING RETRIEVALS
THE CLASSICAL APPROACH IN COMPARING/VALIDATING RETRIEVALS
THE CLASSICAL APPROACH IN COMPARING/VALIDATING RETRIEVALS

- point to point or pixel to pixel comparison
How different are these two fields?

MRMS hourly at 1 km shifted by 7 km
How different are these two fields?

$R^2 = 0.54$

$=>$ Quite different at the pixel level!
Effective Resolution (ER)

“The finest scale at which retrievals accurately reproduce the local spatial variability of a reference product”

Global multiscale evaluation of satellite passive microwave retrieval of precipitation during the TRMM and GPM eras: effective resolution and regional diagnostics for future algorithm development

Clement Guilloteau1,*, Efi Foufoula-Georgiou1, and Christian D. Kummerow2
1 Department of Civil and Environmental Engineering, University of California, Irvine
2 Department of Atmospheric Science, Colorado State University, Fort Collins

Guilloteau et al., JHM, 2017; JTech 2018
Variance as a function of the scale

satellite precipitation field

LR → 100Km → 80Km → 40Km → 20Km → 10km → HR
Variance as a function of the scale

satellite precipitation field

LR → HR
Variance as a function of the scale

- **Variance** as a function of the scale
- **var. of sat. precip. field**
- **scale (km)**
- **satellite precipitation field**
- **radar precipitation field**
Variance as a function of the scale

Variance (mm/h)^2

scale (km)

satellite precipitation field
radar precipitation field
Precipitation signal or noise?

Variation of satellite precipitation field

Precipitation signal

Noise

Variation explained: $R^2 \times \text{var(sat.)}$
Precipitation signal or noise?

var. of sat. precip. field

noise

var. explained: $R^2 \times \text{var}(\text{sat.})$

+0.45

precipitation signal

160 80 40 20 10

scale (km)

satellite precipitation field

radar precipitation field

$R^2$
Precipitation signal or noise?

var. of sat. precip. field

noise

var. explained: $R^2 \times \text{var(sat.)}$

+0.45

+0.15

precipitation signal

160 80 40 20 10
scale (km)

satellite precipitation field

radar precipitation field
Precipitation signal or noise?

Variance of satellite precipitation field:

- $R^2 \times \text{var(sat.)}$

Signal:

- +0.45
- +0.30
- +0.20

Noise:

- +0.28
- +0.19

Scale (km):

160  80  40  20  10

Satellite precipitation field

Radar precipitation field
Precipitation signal or noise?

- **var. of sat. precip. field**
  - noise +0.19
  - +0.28
  - +0.30
  - +0.15
  - +0.45

- var. explained: $R^2 \times$ var(sat.)

- noise variance added $>$ signal variance added

- Effective Resolution = 40 km

- These scales are unresolved
• 16,500 GPM orbits: March 2014 to February 2017
• Local values computed from all observations in $3^\circ \times 3^\circ$ boxes.
Effective Resolution of GPROF GMI vs. KuPR

- 16,500 GPM orbits: March 2014 to February 2017
- Local values computed from all observations in 3°×3° boxes.

Guilloteau et al., JHM, 2017
Highly Underdetermined Inverse problem

RETRIEVAL DATABASE

1

nearly identical
spectral
signatures

GMI + DPR

2

channel

TB (K)

140 180 220 260

160 200 240

Channel
Highly Underdetermined Inverse problem

RETRIEVAL DATABASE

1

nearly identical spectral signatures

GMI + DPR

2

orbit #21092
2017-11-14
10:30 UTC

orbit #17177
2017-03-07
17:15 UTC
Highly Underdetermined Inverse problem

RETRIEVAL DATABASE

GMI + DPR

1

nearly identical spectral signatures

76 mm/h

very different surface rain rates

4 mm/h

orbit #21092
2017-11-14 10:30 UTC

orbit #17177
2017-03-07 17:15 UTC
Highly Underdetermined Inverse problem

4,000 neighbors in TB space

Guilloteau et al., 2018

Variance of surface rain $(\text{mm/hr})^2$

Radiometric distance (K)

nugget effect
Highly Underdetermined Inverse problem

1) Increasing the size of the data base will not help

2) Improved inversion algorithms (KNN Bayesian, L1-L2, etc.) limited improvement in retrieval accuracy/extremes

e.g., Ebtehaj et al., 2015, 2016 (L1-L2)
We propose to look beyond the pixel ...

Look at PATTERNS of TB

RETRIEVAL DATABASE

1

76 mm/h

2

4 mm/h
We propose to look beyond the pixel ...

Local depression of the 37 GHz TB = deep convection

RETRIEVAL DATABASE

76 mm/h

4 mm/h
The challenge becomes:

How to extract:

-- the most informative non-local parameters from the TB patterns

-- to increase identifiability and reduce retrieval uncertainty?
Convolution filters to extract spatial information from fields of TB

- Pattern extraction
- Spatial averaging / smoothing
- Spatial differentiation / edge detections / gradients extraction
- Multiscale decompositions (wavelets)
Convolution filters to extract spatial information from fields of TB

“nonlocal” parameter
KNN retrieval from GMI with a 700 000 profile database

**MEAN ABSOLUTE ERROR**

[Graph showing mean absolute error with a trend line indicating decreasing error as k increases, with a label for 'Land' and a note: 13 “pixel” TBs + 2m temp. + surf. type]
KNN retrieval from GMI with a 700 000 profile database

**MEAN ABSOLUTE ERROR**

13 “pixel” TBs + 2m temp. + surf. type

Land

-11%

13 “pixel” TBs + 2m temp. + surf. type + 3 nonlocal param. (at 37 and 89 V GHz)
KNN retrieval from GMI with a 700 000 profile database

**RAINFALL MISS RATE**

- **land**
  - 13 “pixel” TBs + 2m temp. + surf. type
  - -25% miss rate

- **ocean**
  - 13 “pixel” TBs + 2m temp. + surf. type + 3 nonlocal param. (at 37 and 89 V GHz)
  - -20% miss rate
What’s next?

• Is Machine Learning (ML) the solution?
• Eventually maybe, but not without physically-based dimensionality reduction first
• Train Convolutional Neural Networks (CNNs) and by backpropagation methods learn what patterns were retained in the training (attribution methods)
• Could work on specific storm systems, e.g., snowstorms and learn patterns that “detect snow”, etc.
• Error diagnostics for multi-sensor merging (IMERG)
Today’s focus:

RAINFALL

1. Global estimation from space

2. Seasonal prediction
30 years

3 years
<table>
<thead>
<tr>
<th></th>
<th>Minneapolis</th>
<th>Irvine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rain</td>
<td>32 inches</td>
<td>12 inches</td>
</tr>
<tr>
<td>Snow</td>
<td>53 inches</td>
<td>0 inches</td>
</tr>
<tr>
<td>Prec days</td>
<td>112 days</td>
<td>36 days</td>
</tr>
<tr>
<td>Avg T Jan</td>
<td>7 degrees F</td>
<td>46 degrees F</td>
</tr>
</tbody>
</table>

Average Monthly Precipitation in Inches

https://www.usclimatedata.com/
Large interannual variability

Half of annual rain in 5-10 days
Persistent H/L-pressure ridges/troughs over the Gulf of Alaska affect the jet stream diverting it to the N or S relative to its average latitudinal location.

These pressure patterns are typically related to ENSO.

Precipitation in SWUS: It all comes down to Pressure...

Figures are from Lindsey, 2016.
Above normal SSTs in the tropical Pacific increase convergence in the surface which enhances air convection and leads to anomalous divergence in the top of the troposphere.

A quasi-stationary Rossby wave of alternating anticyclonic and cyclonic patterns forms, which is associated with a southward shift of the storm tracks in the subtropical regions.

(Trenberth et al., 1998)
Low predictive ability by ENSO

$\text{Precip (Nov-Mar)}$

$\text{Niño 3.4 (Jul-Oct)}$

$r = 0.35$
Low predictive ability by ENSO

$r = 0.35$

Precip (Nov-Mar)

Niño 3.4 (Jul-Oct)
Low predictive ability by ENSO

$r = 0.35$
Low predictive ability by ENSO

$r = 0.35$

Standardized Anomaly

Year

Precip (Nov-Mar)
Niño 3.4 (Jul-Oct)

Wet miss
Dry miss
False wet alarm
False dry alarm
Low predictive ability by ENSO

- Mega El Niño 2015-16 => dry year
- Strong La Niña 2010-11 => wet year
- ENSO neutral in 1992-93 => one of the wettest years in record

\[ r = 0.35 \]
The increasing importance of Western Pacific

Geophysical Research Letters

RESEARCH LETTER
10.1029/2014GL065978

Probable causes of the abnormal ridge accompanying the 2013–2014 California drought: ENSO precursor and anthropogenic warming footprint

S.-Y. Wang1,2, Lawrence Higgs3, Robert R Gillies1,2, and Jin-Ho Yoon7

1Utah Climate Center, Utah State University, Logan, Utah, USA; 2Department of Physics, Soils and Climate, Utah State University, Logan, Utah, USA; 3Pacific Northwest National Laboratory, Richland, Washington, USA.

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2017JD026575

Remote Linkages to Anomalous Winter Atmospheric Ridging Over the Northeastern Pacific

Daniel L. Swain1,2, Deepak Singh1,2, Daniel E. Horton4, Justin S. Mankin1,3, Tristan C. Ballard1, and Noah S. Diffenbaugh1,3

1Paciﬁc Institute, Oakland, California, USA; 2Department of Earth and Environmental Science, University of California, Los Angeles, California, USA; 3Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA; 4Department of Earth and Planetary Sciences, University of California, Davis, California, USA.

 Causes of Extreme Ridges That Induce California Droughts

HAIYAN TENG AND GRANT BRANSTATOR

National Center for Atmospheric Research, Boulder, Colorado

INTERNATIONAL JOURNAL OF CLIMATOLOGY


Decadal Variations in the Strength of ENSO Teleconnections with Precipitation in the Western United States

GREGORY J. MCCABE** and MICHAEL D. DETTINGER**

**US Geological Survey, Denver Federal Center, MS 412, Denver, CO 80225, USA

DEcadal variations in the strength of ENSO teleconnections with precipitation in the western United States

GREGORY J. McCABE** and MICHAEL D. DETTINGER**

**US Geological Survey, Denver Federal Center, MS 412, Denver, CO 80225, USA

Article

Impacts of Pacific SSTs on Atmospheric Circulations Leading to California Winter Precipitation Variability: A Diagnostic Modeling

Bo Kyeong Myoung1,*, Sang-Wook Yeh2, and Ji-Hee Kim1

1APEC Climate Center, Busan 49098, Korea

2Department of Marine Sciences, Convent of Maryknoll, Busan, Korea

Geophysical Research Letters

RESEARCH LETTER
10.1002/2019GL084021

On the Delayed Coupling Between Ocean and Atmosphere in Recent Weak El Niño Episodes

N. C. Johnson1,2, M. L. Heuer1,2, C.-H. Chang1, and Z.-Z. Hu1

1Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, NJ, USA, 2NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA, 3NOAA/NCEP Climate Prediction Center, College Park, MD, 4Center for Climate/Environment Change Prediction Research, Ewha Womans University, Seoul, South Korea

Drought and the California Delta—A Matter of Extremes

Michael Dettinger1 and Daniel R. Cayan1

*And it never failed that during the dry years the sea level forecast about the wet years and during the wet years, they...
The increasing importance of Western Pacific


“There exists a cross-Pacific pathway of Rossby wave energy, propagating from the western subtropical Pacific toward the Gulf of Alaska…” – Wang et al., GRL, (2014) on the extreme 2013/2014 North American drought

“…there are tropical heating anomalies that do not depend on ENSO that may excite extratropical responses that include extreme west coast ridges.” -- Teng and Branstator, J. Climate, (2017)

A new interhemispheric teleconnection increases predictability of winter precipitation in southwestern US

Antonios Mamalakis, Jin-Yi Yu, James T. Randerson, Amir AghaKouchak, & Efi Foufoula-Georgiou
Above normal precipitation in the southwestern US

Below normal precipitation in the southwestern US

1. Atmospheric bridge
2. Local air-sea couplings (Wang et al. 2000)
3. Deflection of the jet stream (Wang et al. 2011)

May  June  July  August  September  October  November  December  January  February  March  April

Rainy season
Western Pacific pathway hypothesis

Late boreal summer

New Zealand Index

170°E-200°E and 25°S-40°S

Mamalakis et al., 2018, Nat. Communications

Cool NZI: Strengthened southern HC

Warm NZI: Weakened southern HC
Western Pacific pathway

Late boreal summer

Cool NZI: Strengthened southern HC

(Warm-Cool) NZI years

Expect weakened convection in NW Pacific (positive anomalies in zonal mean Omega velocity)

Expect increasing incoming solar radiation in NW Pacific

New Zealand Index

170°E-200°E and 25°S-40°S

Mamalakis et al., 2018, Nat. Communications
Western Pacific pathway

Late boreal summer

New Zealand Index
170°E-200°E and 25°S-40°S

Cool NZI: Strengthened southern HC

Warm NZI: Weakened southern HC

Top of the troposphere

Omega (zonal avg)

Pressure level (mb)

Surface

Weakened convection

Latitudinal shift of ITCZ

Mamalakis et al., 2018, *Nat. Communications*
Western Pacific pathway

Late boreal summer

Increased solar radiation

New Zealand Index
170°E-200°E and 25°S-40°S

Mamalakis et al., 2018, Nat. Communications
Is the WP Pathway “independent” of ENSO?

Late boreal summer

Cascading of NZI SST anomalies in the north Pacific is significant even after accounting for ENSO

$\text{Corr} \left[ \text{NZI(Jul-Sep)}, \text{SST(2, 4 months later)} \mid \text{ENSO(Jul-Sep)} \right]$

2 months later
NZI anomalies cascade to NH

4 months later
NH anomalies sustained

Mamalakis et al., 2018, *Nat. Communications*
Has the WP Pathway amplified?

Late boreal summer

Based on Observations

New Zealand Index
170°E-200°E and 25°S-40°S

-- Internal variability?
-- External forcing?
-- Data quality?

Mamalakis et al., 2018, Nat. Communications
Has the WP Pathway amplified?

Late boreal summer

Based on Models: CESMv1 Large Ensemble

Ensemble mean of correlation $NZI_m$ and $EPI_{m+dm}$

Ensemble st. deviation of correlation $NZI_m$ and $EPI_{m+dm}$

Mamalakis et al., 2018, Nat. Communications

New Zealand Index

170°E-200°E and 25°S-40°S
Has the WP Pathway amplified?

Based on Models: CESMv1 Large Ensemble

Late boreal summer

New Zealand Index
170°E-200°E and 25°S-40°S

 Increases of the Ensemble Mean Correlations

Convergence of ensembles

Has the WP Pathway amplified?

Mamalakis et al., 2018, Nat. Communications
“We are trying to prove ourselves wrong as quickly as possible, because only in that way we can find progress”

Richard P. Feynman
On the Scientific method
Adding Western Pacific SSTs as predictors of Precipitation

**Explained var:** 9%
**Dry success rate:** 28%
**Wet success rate:** 30%

**Explained var:** 22%
**Dry success rate:** 34%
**Wet success rate:** 34%
Is this the best we can do?
Explore the whole Pacific?

Winter precipitation

Weights

Climate predictors (e.g. SSTs, GPHs in Pacific ocean)

\[ y = X\beta + \epsilon \]

Very high dimensional problem

SSTs @ 2x2° x 4 months=> 5612 x4=22,448 predictors
Dimensionality Reduction

$t = 1980$

$t = 2019$
Dimensionality Reduction

Impose constraints that respect the space-time Covariance of SSTs

Stevens et al. 2019
Dimensionality Reduction

Promote similar $\beta$ for highly correlated predictors to enforce sparsity and unravel the most explanatory features w/out specifying them a-priori

Stevens et al. 2019
Data-driven prediction

\[ y = X\beta + \epsilon \]

Climate predictors (e.g. SSTs, GPHs in Pacific ocean)

\[ \hat{\beta} = \arg \min_{\beta} \|y - X\beta\|_2 + \lambda_1 \|\beta\|_1 + \lambda_{TV} \sum_{j,k} |\hat{C}_{j,k}|^{1/2} |\beta_j - \hat{s}_{j,k} \beta_k| \]

- Data fitting
- L1 regularizer (LASSO)
- Graph Total Variation (GTV)

\[ \hat{C} = \text{covariance matrix of } X \]

\[ \hat{s}_{j,k} = \text{sign}(\hat{C}_{j,k}) \]

Stevens et al. 2019
Data-driven prediction

Training period: 1940-1990 (with a non-stationarity filter)
Testing period: 1991-2019

GTV captures almost 40% of the variability in the out-sample period

Stevens et al. 2019
What’s next?

• Is Machine Learning (ML) the solution?
• Eventually maybe, but not without testing the causality of hypothesized mechanisms & predictors
• Perform idealized perturbation experiments designed to understand the process chain of the WP teleconnection (e.g, differentiate between Rossby-wave vs. HC mediated interhemispheric propagation)
• Study CMIP6 outputs (historical and future projections) to understand time-evolving dynamics relevant to prediction, spectral PCA
• Probabilistic prediction for water resources planning
U34B - Data Analytics and Machine Learning Innovation for Climate and Earth Surface Processes

Wednesday, 11 December 2019 - 16:00 - 18:00
Moscone South - 303-304, L3

MARKUS REICHSTEIN
Max Planck Institute

MATTHEW HANCHER
Google Earth Engine

GRÉGOIRE MARIETHOZ
University of Lausanne

Evan B. Goldstein
University of North Carolina

Claire Monteleoni
University of Colorado Boulder

Veronika Eyring
German Aerospace Center DLR

Elizabeth Arnold Barnes
Colorado State University
Measuring the unmeasurable and predicting the unpredictable

Efi Foufoula-Georgiou
University of California, Irvine
Patterns of Life
Patterns of Life
Patterns of Life
Patterns of Life
Efi’s Group -- Positive covariances

✓ Whole > Sum (parts)?

\[ X_1 = \text{contribution of member 1} \]
\[ X_2 = \text{contribution of member 2} \]

\[ X = X_1 + X_2 \]
\[ X = \text{overall contribution} \]

\[ \text{Mean}(X) = \text{Mean}(X_1) + \text{Mean}(X_2); \]

\[ \text{Var}(X) = \text{Var}(X_1) + \text{Var}(X_2) + \text{COV}(X_1, X_2) \]

Whole > sum of its parts \iff \text{COV} (+)
**E.F.I.** BINGO

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;gee&quot;</td>
<td>buys the drinks</td>
<td>not wearing black outfit</td>
<td>makes the room laugh</td>
</tr>
<tr>
<td>😲</td>
<td>🍻</td>
<td>🎈</td>
<td>😂</td>
</tr>
<tr>
<td>&quot;the heck...&quot;</td>
<td>&gt; 4 hour meeting</td>
<td>cooks a meal in &lt;20 mins</td>
<td>sends an emoji</td>
</tr>
<tr>
<td>☹️</td>
<td>😴</td>
<td>🍽️</td>
<td>😁</td>
</tr>
<tr>
<td>tells tryphon to cool it</td>
<td>is bored by your research</td>
<td>high-fives you</td>
<td>&quot;jesus christ&quot;</td>
</tr>
<tr>
<td>😡</td>
<td>😵</td>
<td>😎</td>
<td>😡</td>
</tr>
<tr>
<td>parks illegally/where there is no space</td>
<td>&quot;look&quot;</td>
<td>winks at you</td>
<td>meeting at her home</td>
</tr>
<tr>
<td>🚦</td>
<td>👀</td>
<td>😚</td>
<td>🏡</td>
</tr>
<tr>
<td>you drive her to/from airport</td>
<td>breaks meeting for &quot;my yoga&quot;</td>
<td>your paper is &quot;not there yet&quot; for &gt;6 months</td>
<td>&quot;shit&quot;</td>
</tr>
<tr>
<td>🛫</td>
<td>🤦</td>
<td>😞</td>
<td>🙄</td>
</tr>
</tbody>
</table>

*Efficient Fear Injector

Figure 1. Incoming PhD students must complete a BINGO, defined by marking of five squares in a straight or diagonal line, before a PhD degree may be awarded. Squares with quotation marks indicate precise, standalone phrases that must be directed to you. If you are in doubt, it doesn’t count—you’ll know when you hear it. Pictures of Efi are free squares. The grid was designed to maximize entropy such that each possible bingo has approximately the same probability.
**E.F.I. BINGO**

<table>
<thead>
<tr>
<th>“gee”</th>
<th>buys the drinks</th>
<th>not wearing black outfit</th>
<th>makes the room laugh</th>
</tr>
</thead>
<tbody>
<tr>
<td>🤔 🥈 🍻</td>
<td>🎉</td>
<td>🌈</td>
<td>😃 😃 😃 😃 😃 😃 😃 😃</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>“the heck...”</th>
<th>&gt; 4 hour meeting</th>
<th>cooks a meal in &lt;20 mins</th>
<th>sends an emoji</th>
</tr>
</thead>
<tbody>
<tr>
<td>🙇🏻‍♀️ 🙆♀️</td>
<td>☕️</td>
<td>🥣</td>
<td>😴 😴 😴 😴 😴 😴 😴 😴</td>
</tr>
</tbody>
</table>

**Figure 1.** Incoming PhD students must complete a BINGO, defined by marking of five squares in a straight or diagonal line, before a PhD degree may be awarded. Squares with quotation marks indicate precise, standalone phrases that must be directed to you. If you are in doubt, it doesn’t count—you’ll know when you hear it. Pictures of Efi are free squares. The grid was designed to maximize entropy such that each possible bingo has approximately the same probability.
Thanks to my extended family & sponsors

PhD students
--- Praveen Kumar (1993)
--- Sanja Perica (1995)
--- Alin Carsteanu (1997)
--- Venu Venugopal (1998)
--- Deborah Nykanen (2000)
--- Boyko Dodov (2003)
--- Chandana Gangodagamage (2009)
--- Paola Passalacqua (2009)
--- Arvind Singh (2011)
--- Vamsi Ganti (2012)
--- Ardeshir Mo Ebtehaj (2013)
--- Jon Czuba (2015)
--- Jon Schwenk (2016)
--- Mohammad Danesh (2017)
--- Zeinab Takbiri (2018)
--- Lawrence Vulis
--- Antonios Mamalakis

Post-doctoral fellows
--- Victor Sapozhnikov
--- Daniel Harris
--- Shuxia Zhang
--- Rohan Shreshtha
--- Ian Iorgulescu
--- Kurt Fienberg
--- Bruno Lashermes
--- Stefano Zanardo
--- Mahesh Rathinasamy
--- Zi Wu
--- Yiannis Dialynas
--- Amy Hansen
--- Leichen Guo
--- Anthony Longjas
--- Simon Papalexioou
--- Alex Tejedor
--- Clement Guilloteau
Thanks to all my collaborators
THANK YOU!

“Study hard what interests you the most in the most undisciplined, irreverent and original manner possible”

Richard P. Feynman
2019-20 winter precip. prediction in Irvine?