Causality benchmark: a mechanistic way of diagnosing model fidelity

Qing Zhu

Lawrence Berkeley National Laboratory

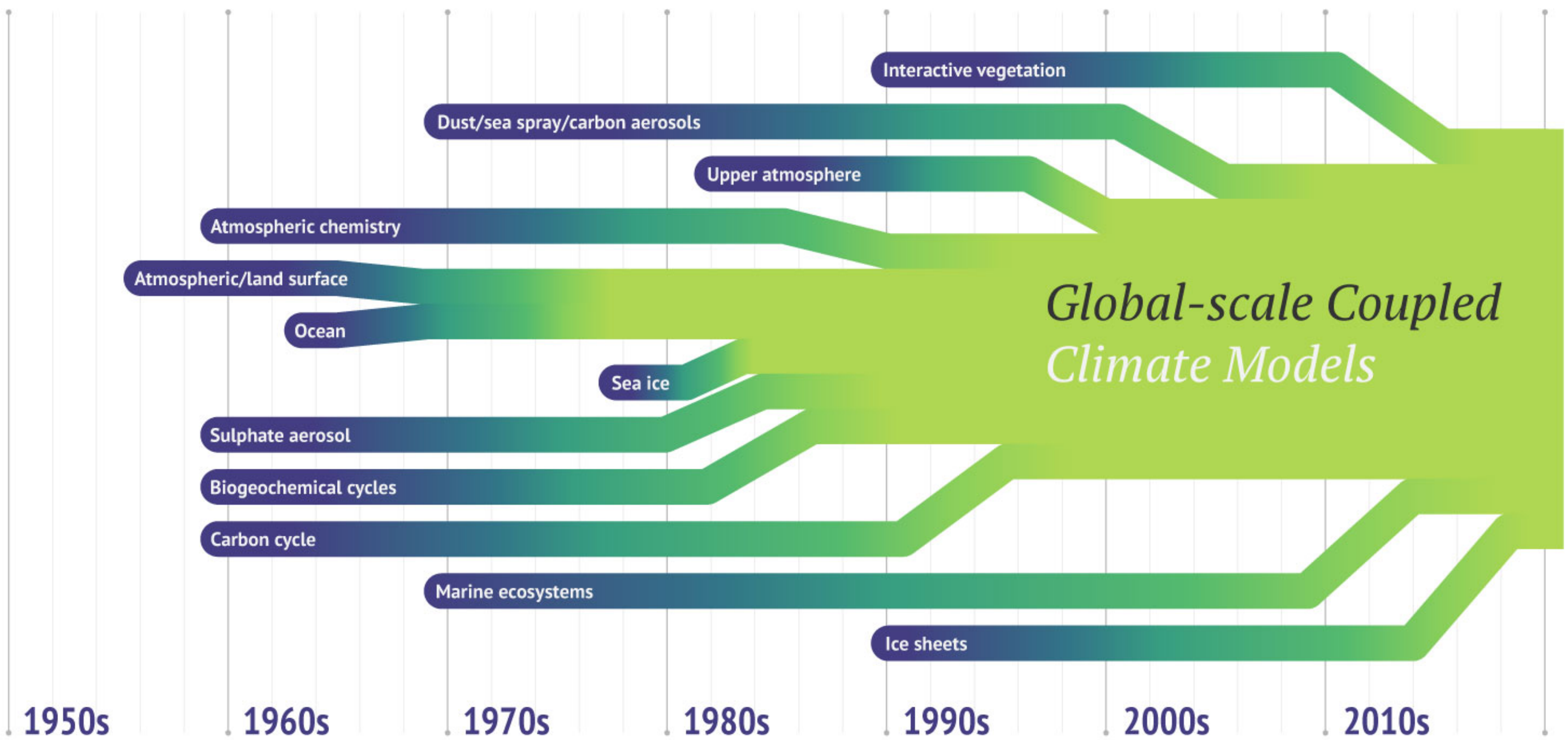
Feb-22-2019

outline

1. Background: benchmarking functional relationship
2. Causality inference: transfer entropy analysis
3. Case study 1: West Sahel precipitation
4. Case study 2: land cover change

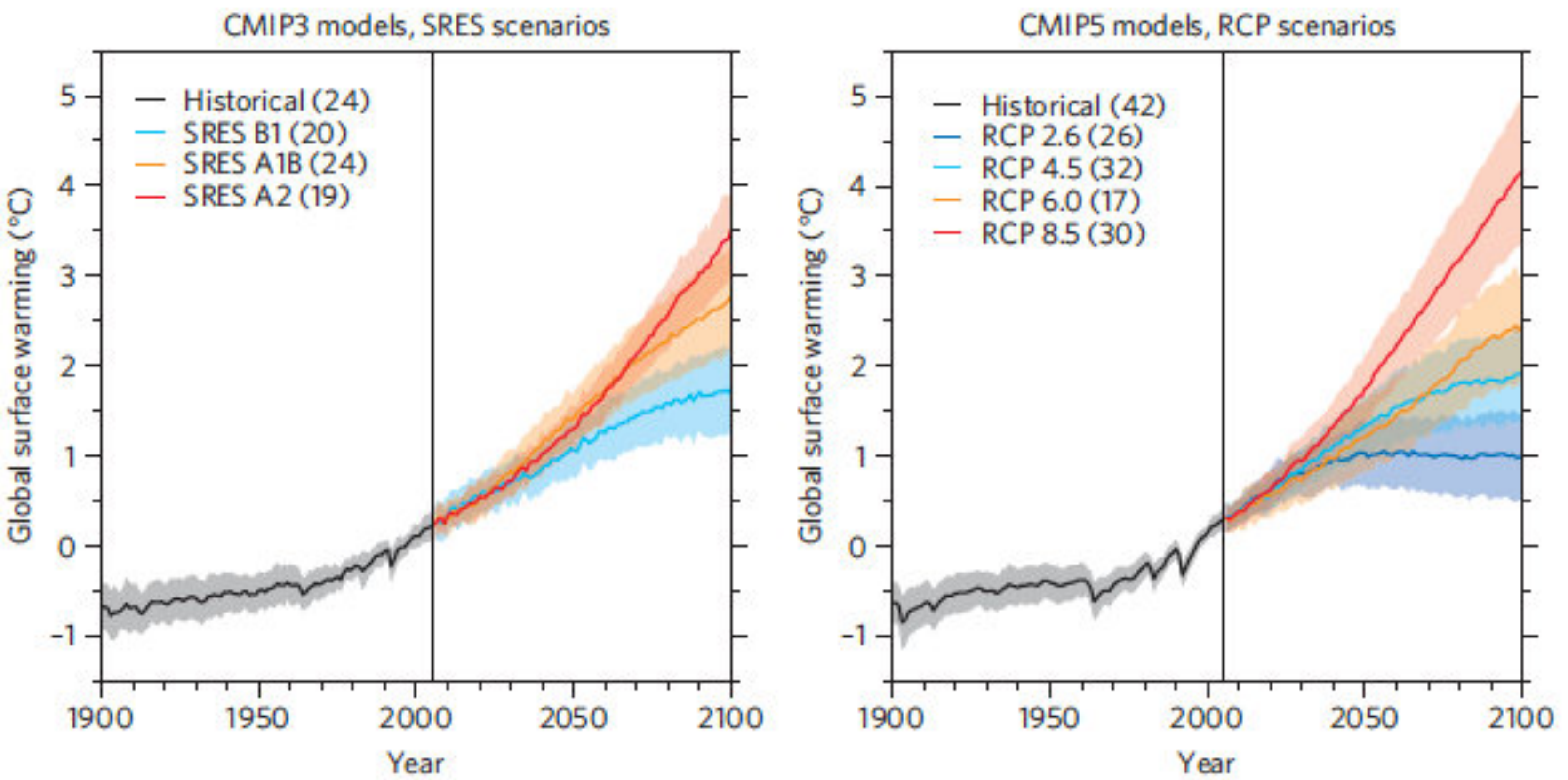
Climate models

For decades scientists have been using mathematical models to help us learn more about the Earth's climate. Known as climate models, they are driven by the fundamental physics of the atmosphere and oceans, and the cycling of chemicals between living things and their environment. Over time they have increased in complexity, as separate components have merged to form coupled systems.



Note: There were some very simplified models before the dates mentioned.

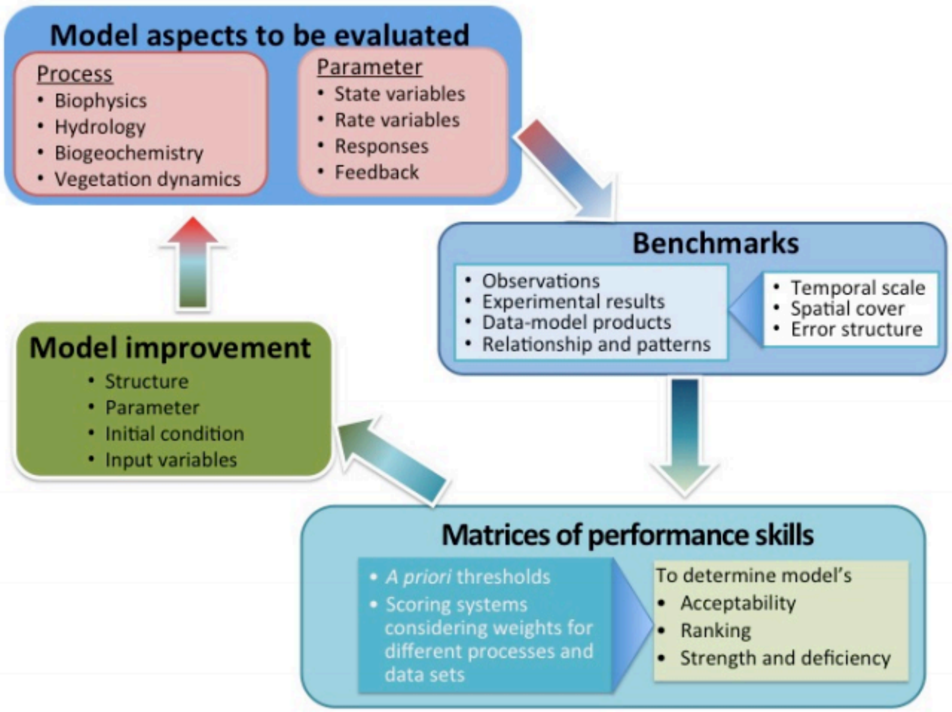
CarbonBrief CLEAR ON CLIMATE



AR4 2007

AR5 2014

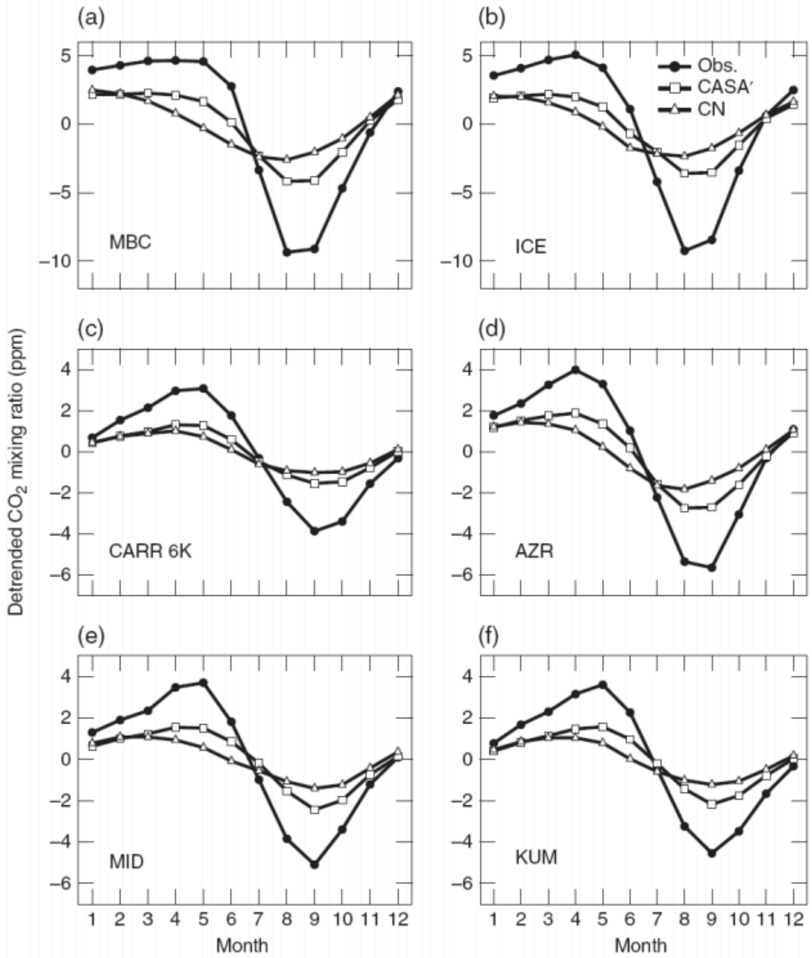
How to use observational benchmark to improve Earth System Model

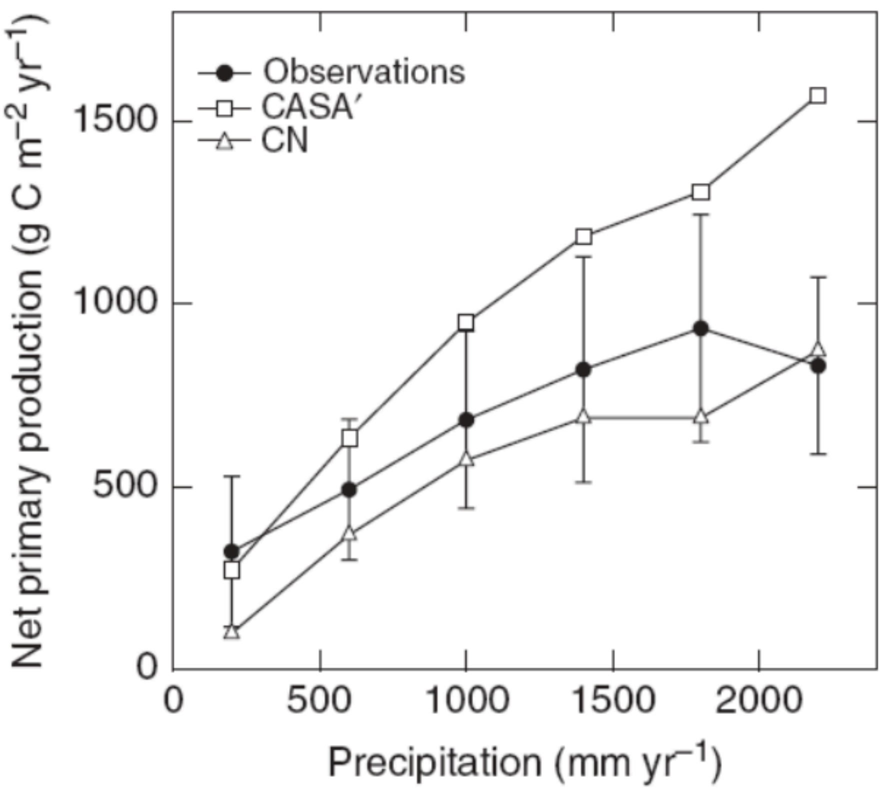


Type	Description	Example
Direct observations	Data from instrument readings with some processing	Atmospheric trace gas mixing ratios, temperature, soil respiration
Experimental results	Data at two or more levels of treatments	Response ratios of biomass and soil moisture
Data–model products	Interpolation and extrapolation of data according to some functions	Global distribution of GPP calculated from satellite and flux data
Functional relationships or patterns	Derived or emerged from data	NPP vs. precipitation, soil respiration vs. temperature

Luo et al., 2012

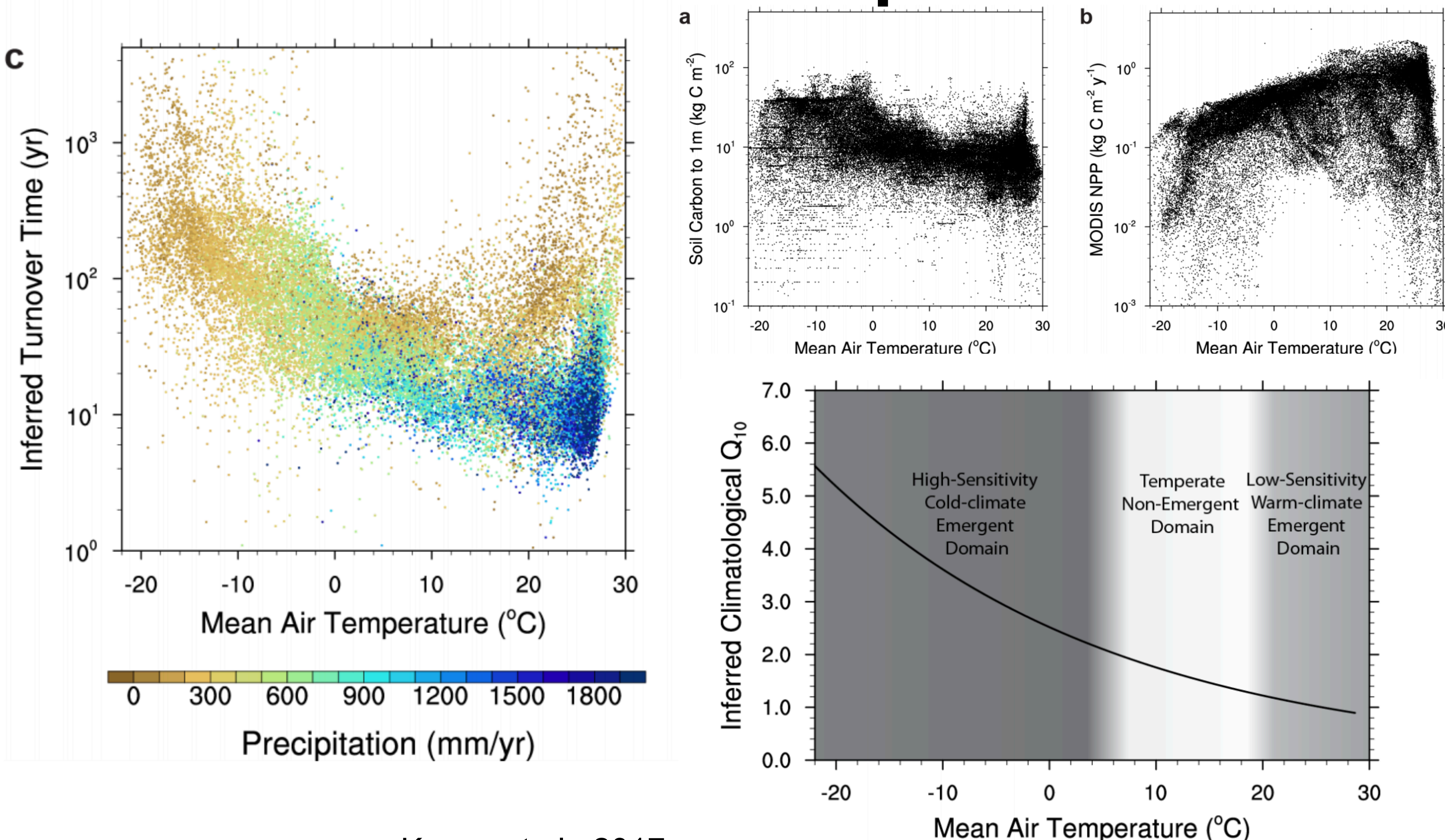
Emerging patterns benchmark





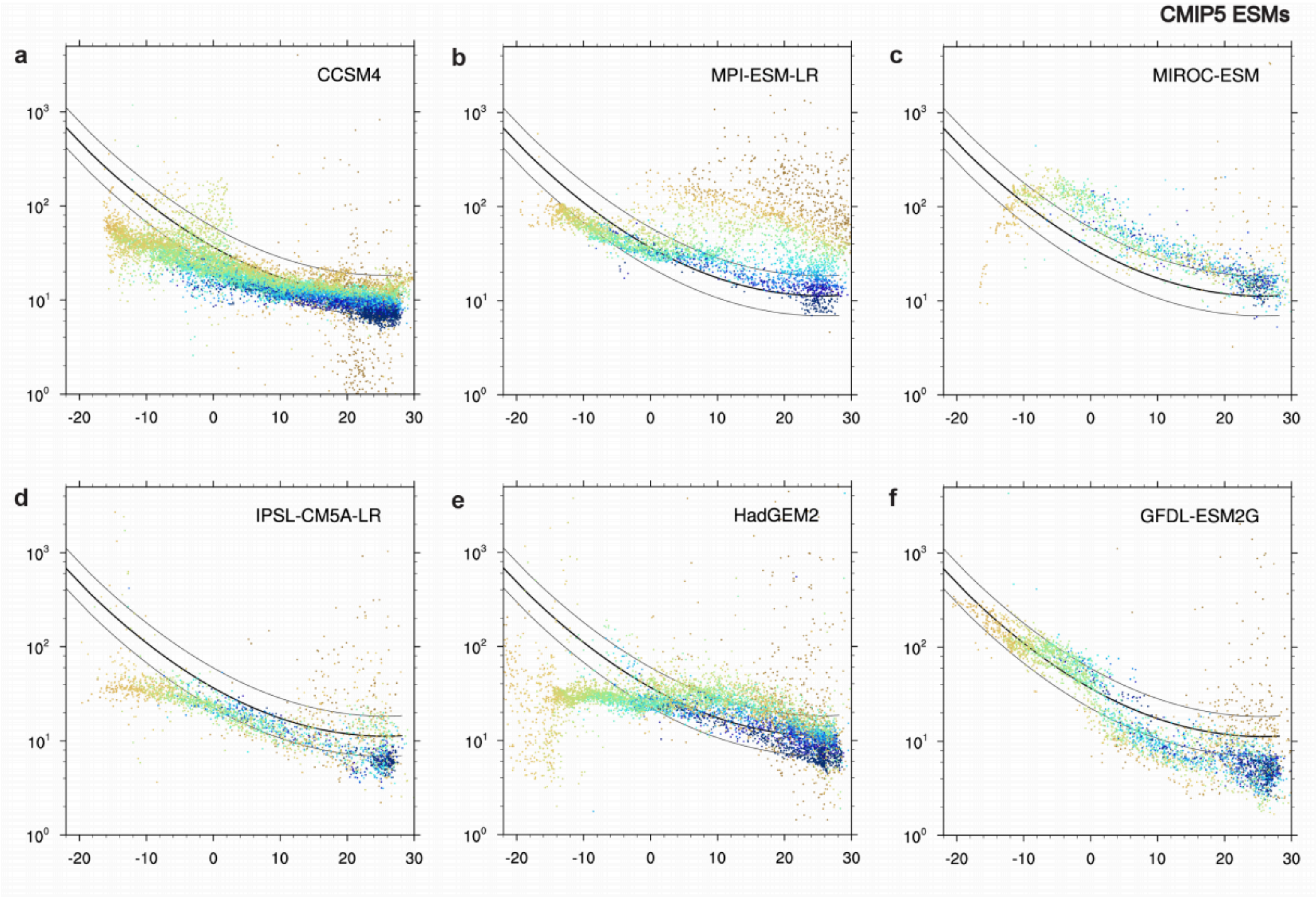
Luo et al., 2012

Functional relationship benchmark



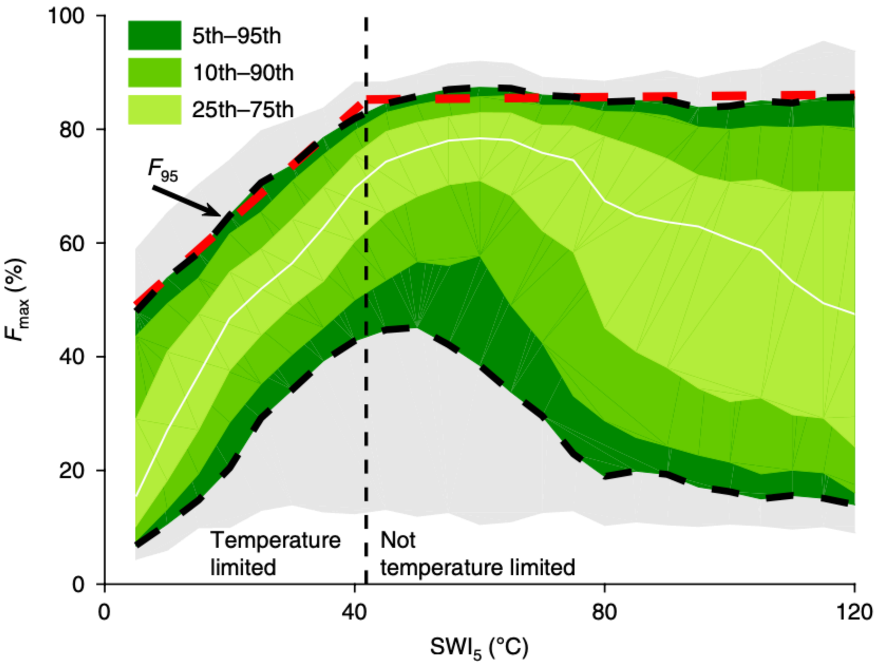
Koven et al., 2017

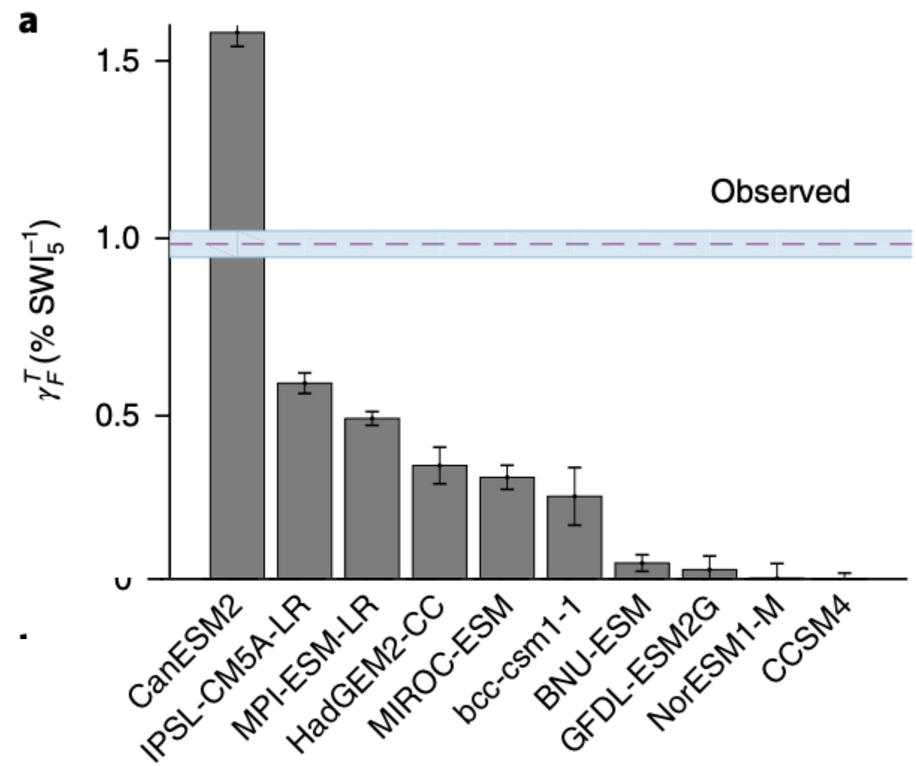
Functional relationship benchmark



Koven et al., 2017

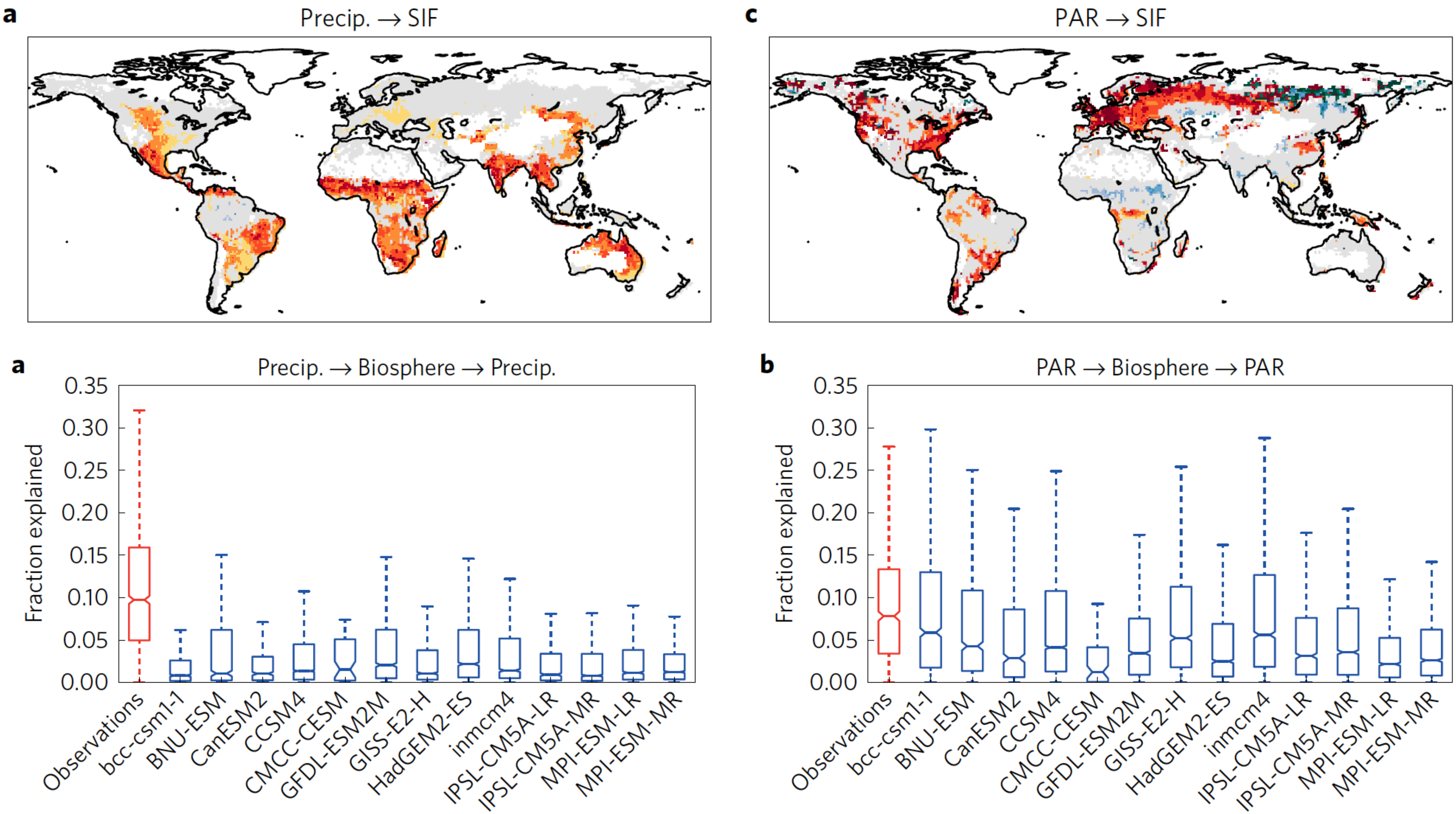
Functional relationship benchmark





Keenan et al., 2017

Causality benchmark



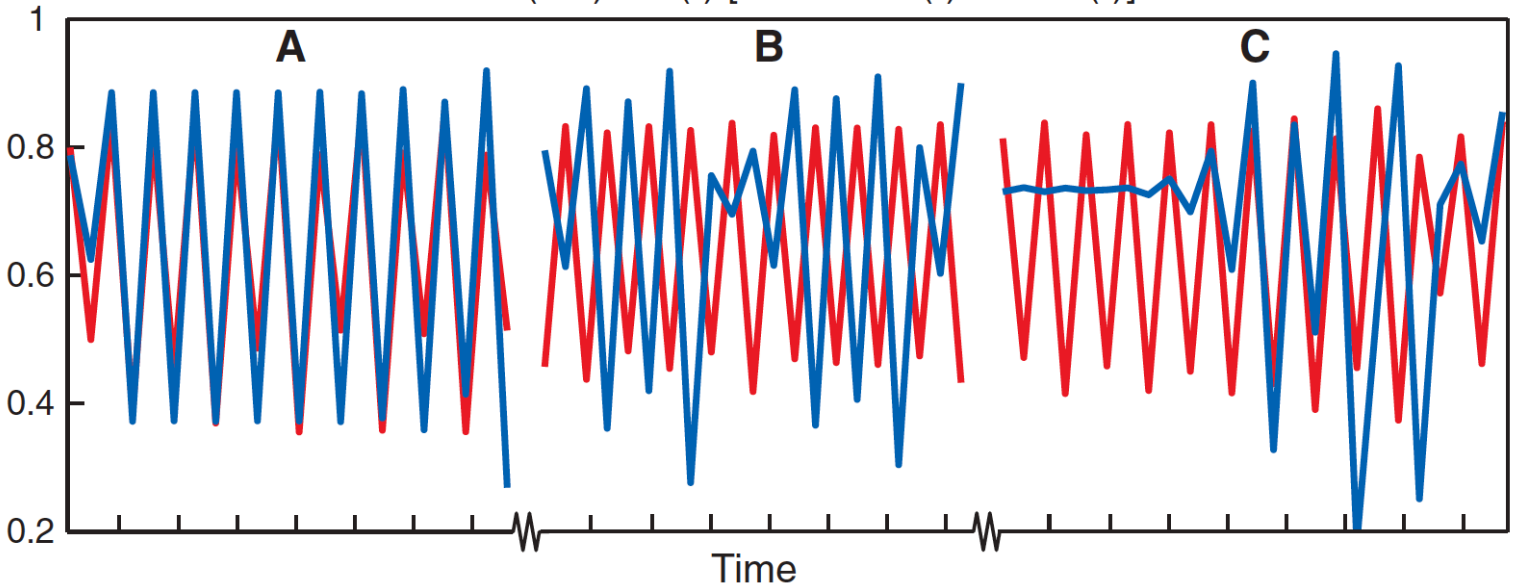
Green et al., 2017

Causality benchmark

- Non-linearity
  - A simplest nonlinear systems (Ephemeral or "mirage" correlations)

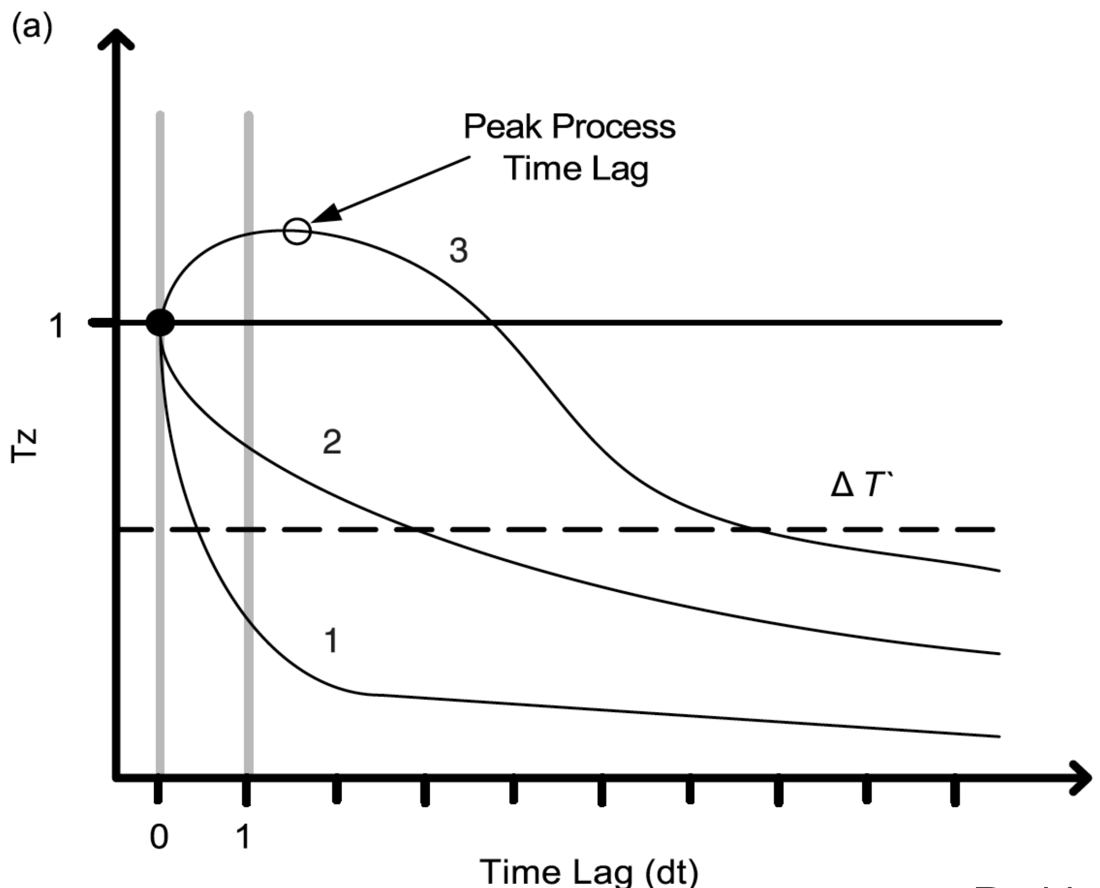
$$X(t+1) = X(t)\,[3.8 - 3.8\,X(t) - 0.02\,Y(t)]$$

$$Y(t+1) = Y(t)\,[3.5 - 3.5\,Y(t) - 0.1\,X(t)]$$



Causality benchmark

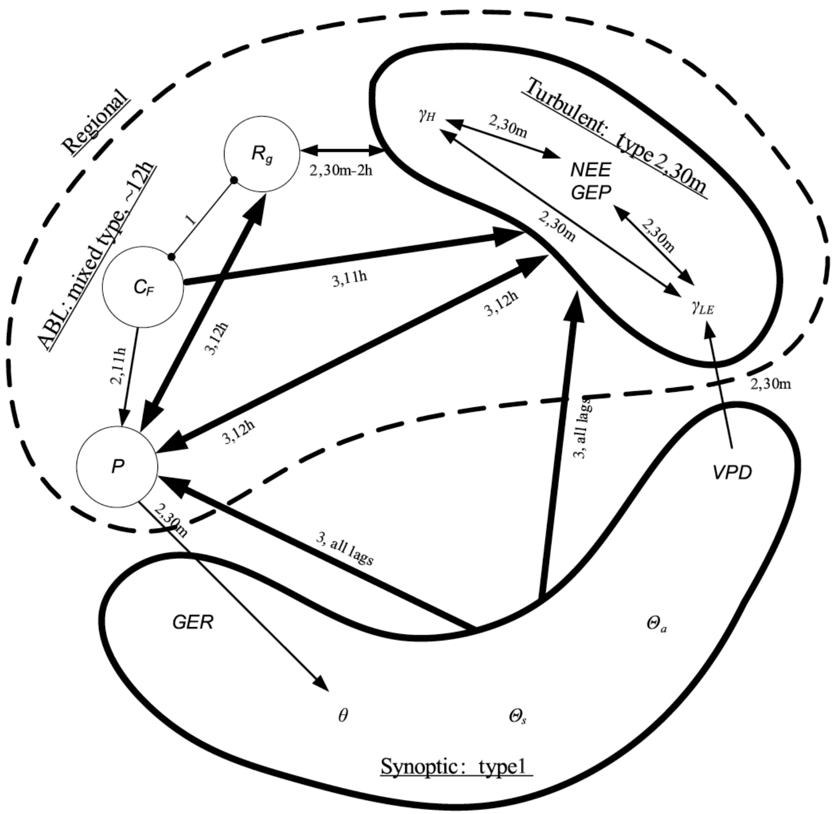
- Nonsynchronous processes



Ruddell et al., 2009

Causality benchmark

- Process network



Ruddell et al., 2009

Contribution to RUBISCO project

- Causation benchmarks
  - Inform underlying mechanisms
  - Better understand model bias
  - Diagnose "wrong reasons -> right answers"

Patterns understanding → Processes understanding

Where？ Why？

Transfer entropy analysis

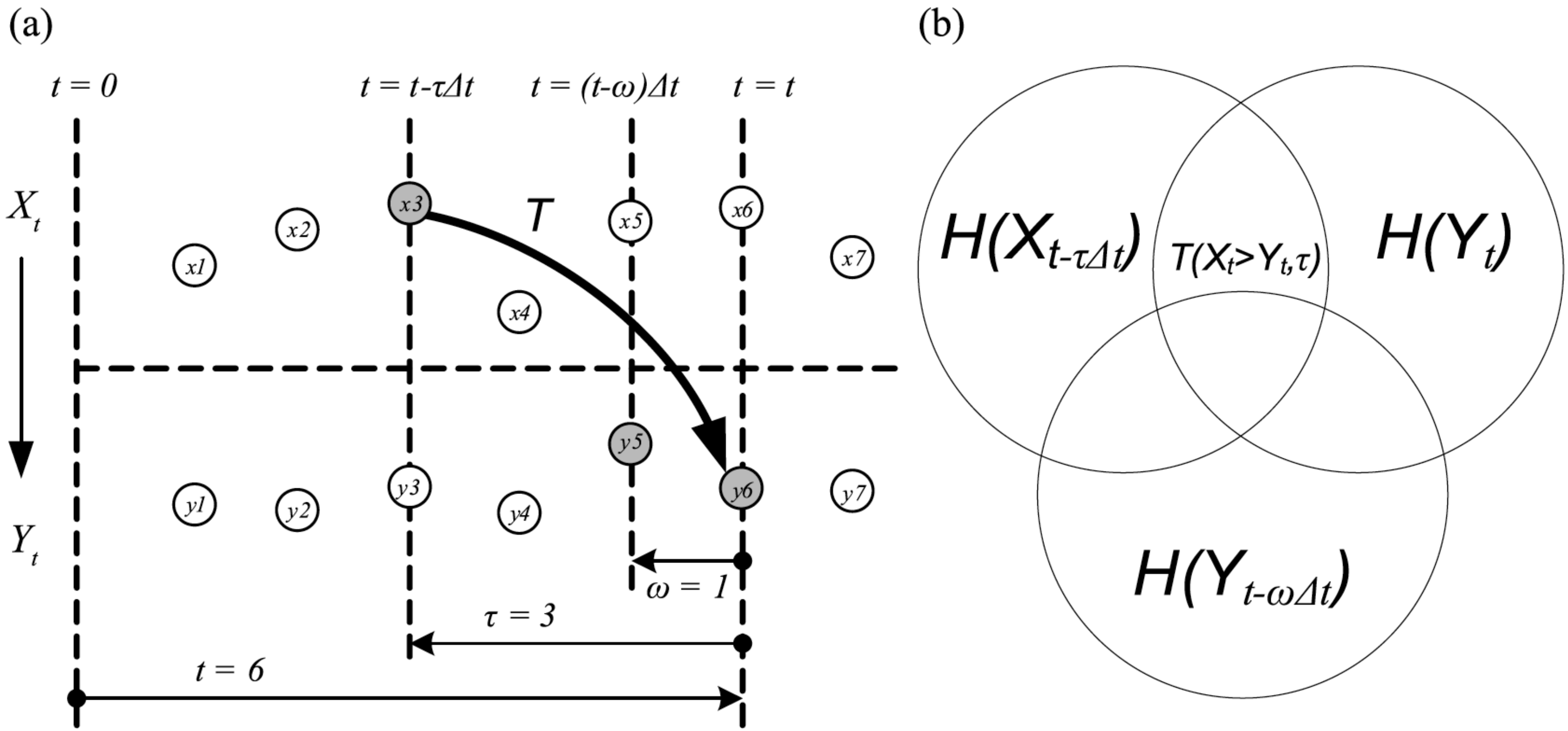
$$
\begin{aligned}
\mathrm{H}(Y|X) &\equiv \sum_{x\in\mathcal{X}} p(x)\,\mathrm{H}(Y|X=x) \\
&= -\sum_{x\in\mathcal{X}} p(x) \sum_{y\in\mathcal{Y}} p(y|x)\,\log\,p(y|x) \\
&= -\sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} p(x,y)\,\log\,p(y|x) \\
&= -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log\,p(y|x) \\
&= -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log\frac{p(x,y)}{p(x)}. \\
&= \sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log\frac{p(x)}{p(x,y)}.
\end{aligned}
$$

$$H_I = -\sum_i p(i)\log_2 p(i)$$

$$I(X_t, Y_t) = \sum_{x_t,y_t} p(x_t,y_t)\log\frac{p(x_t,y_t)}{p(x_t)p(y_t)}$$

$$T^S(X_t > Y_t) = \sum_{y_t,y_t^{[k]},x_t^{[l]}} p\left(y_t, y_t^{[k]}, x_t^{[l]}\right)\log\frac{p\left(y_t \middle| \left(y_t^{[k]}, x_t^{[l]}\right)\right)}{p\left(y_t \middle| y_t^{[k]}\right)}$$

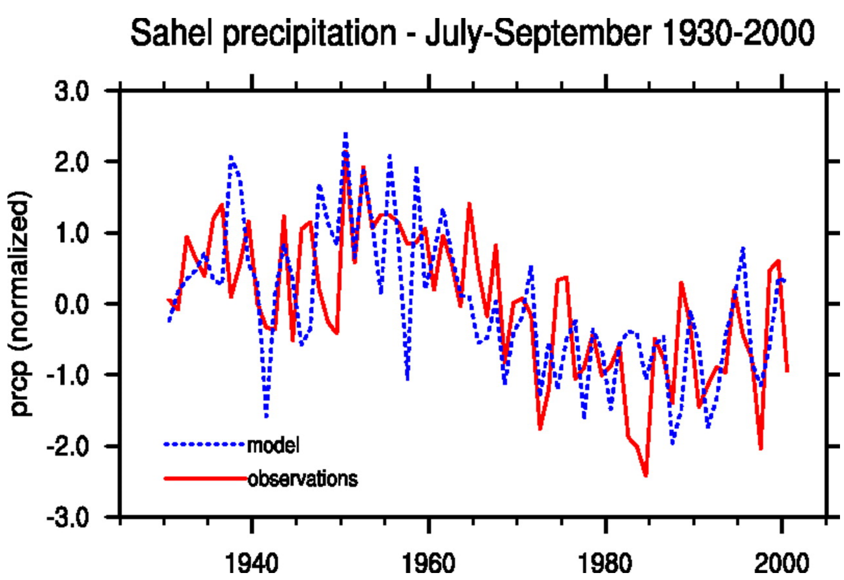
Transfer entropy analysis

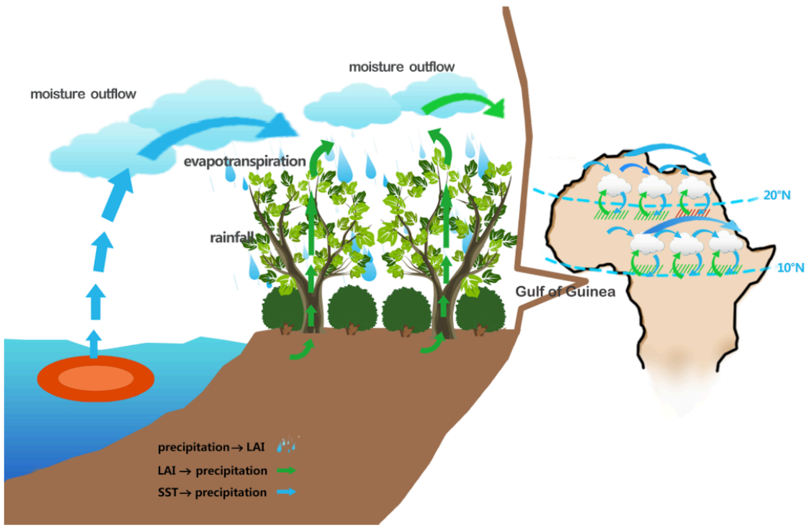


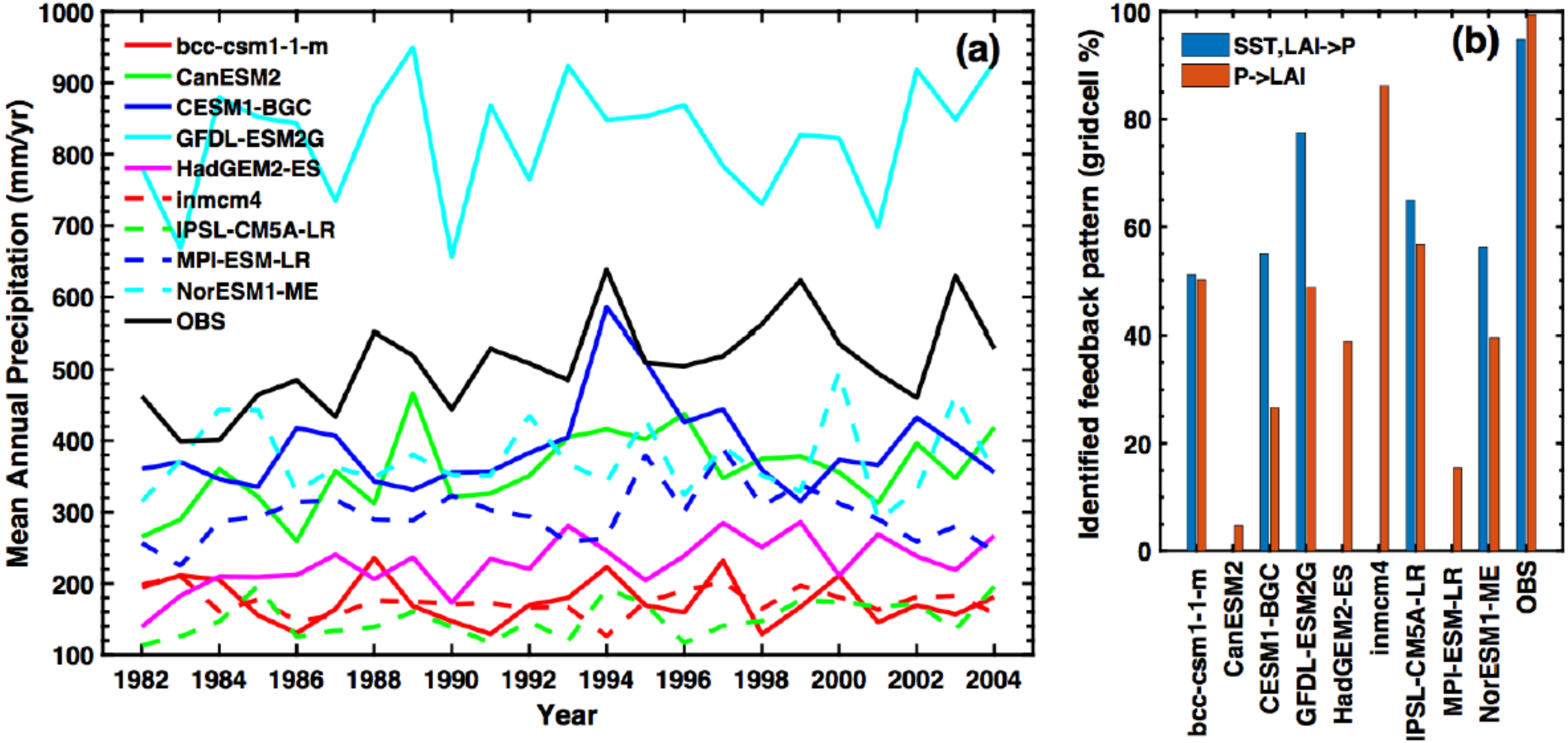
Case 1 West Sahel precipitation

Hypotheses:

- Warm sea surface temperatures weaken land-ocean temperature contrast and migrate deep convection to the ocean, which leads to precipitation decreases over land (Giannini et al., 2003 Science)
- Terrestrial vegetation dynamics control water and energy fluxes into the atmosphere and thus impact local precipitation (Zeng et al., 1999 Science).

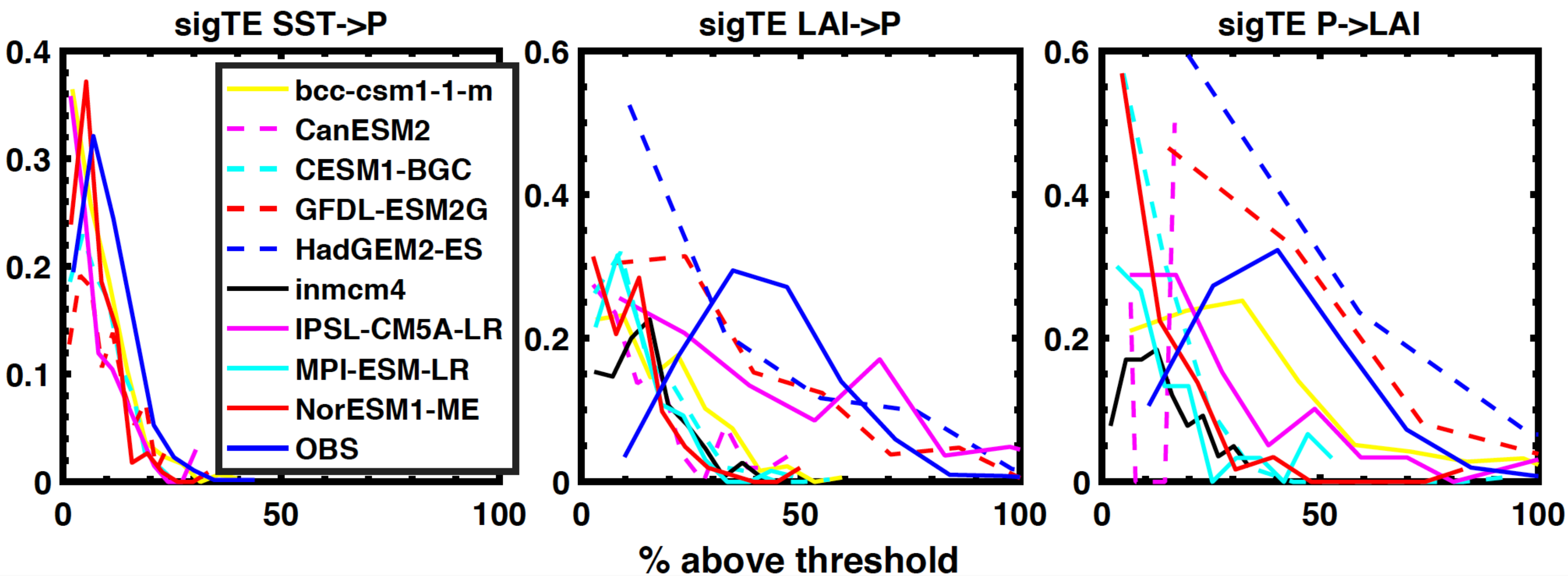


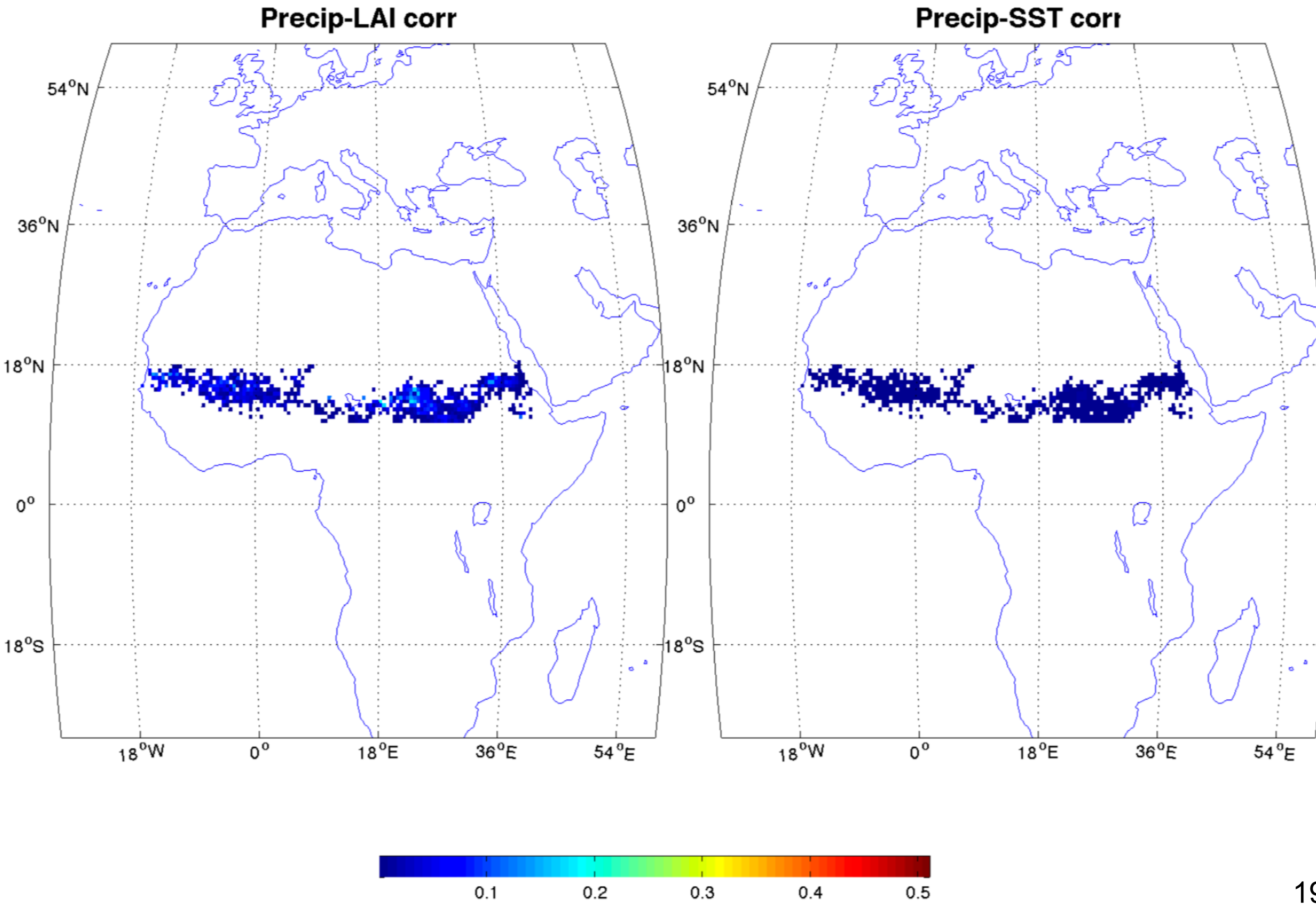




Liu et al., 2019

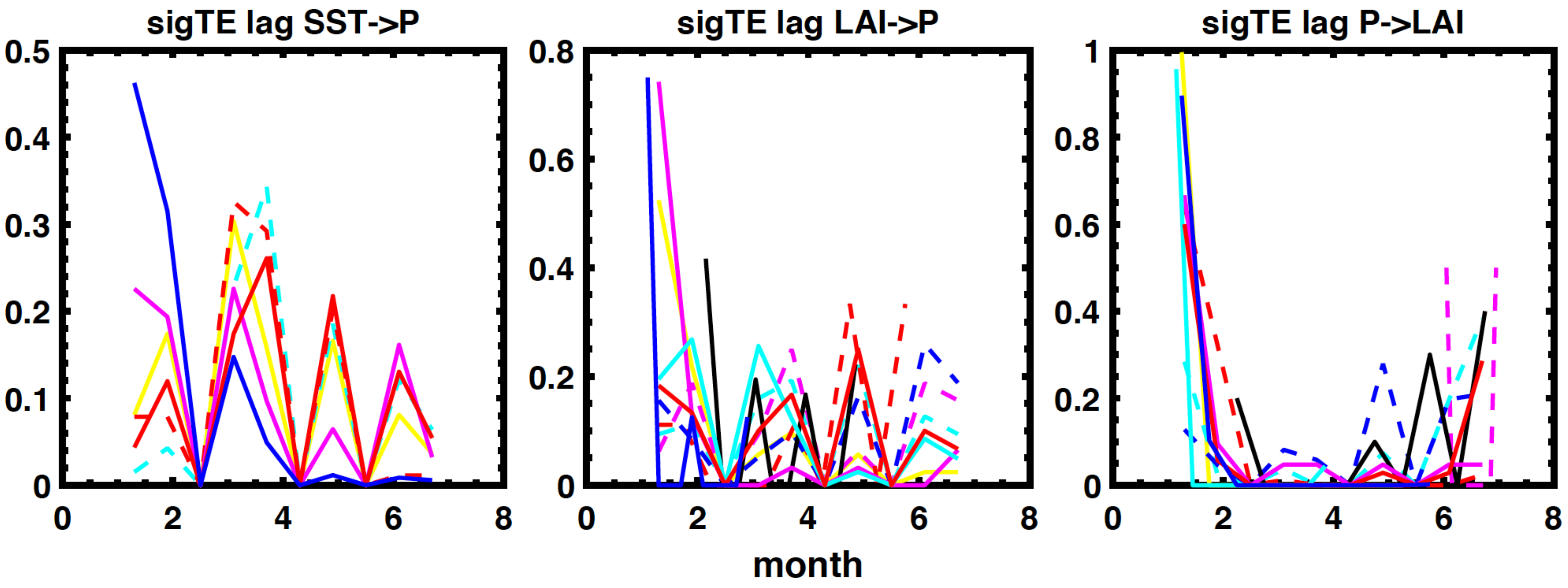
Robustness of identified feedbacks



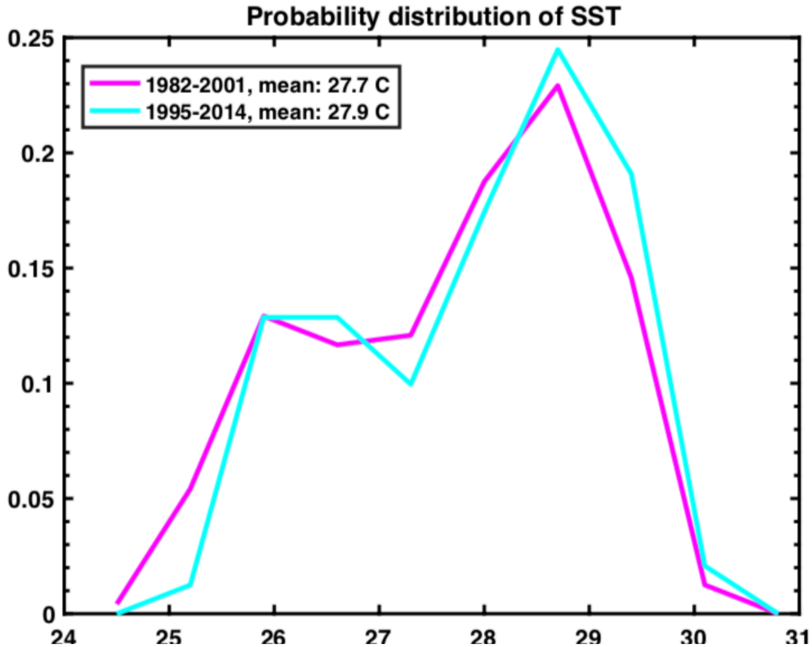


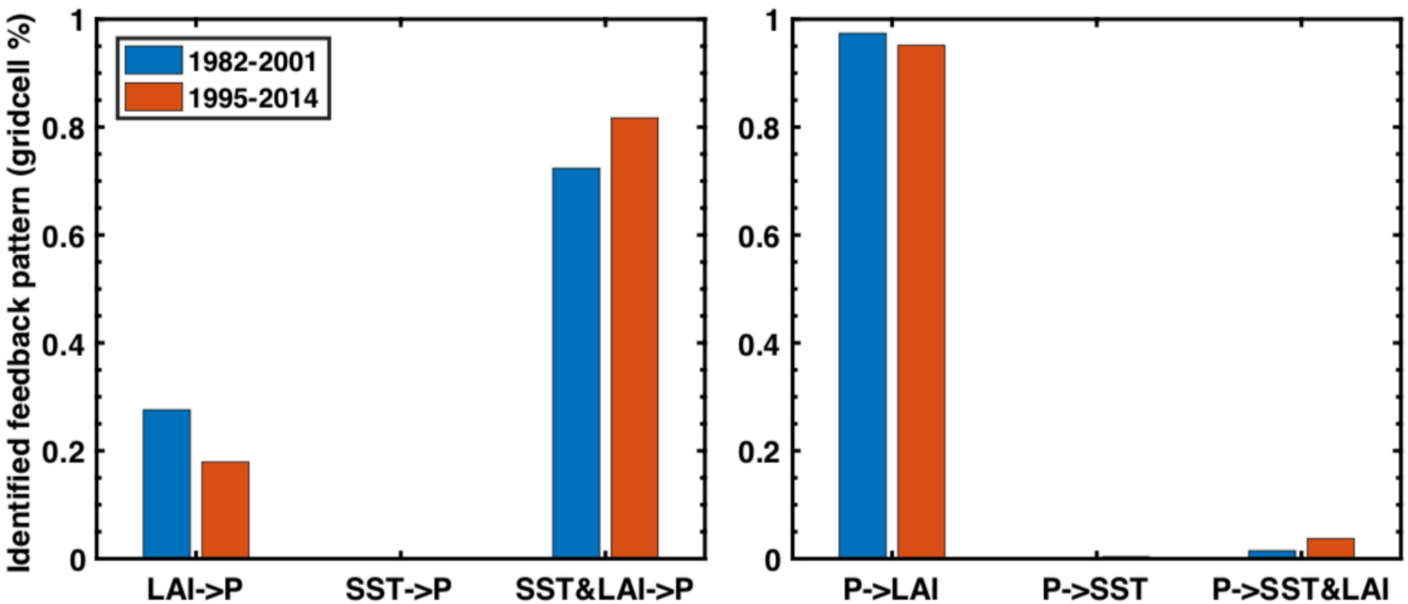
Precip-LAI corr
Precip-SST corr
54°N
36°N
18°N
0°
18°S
18°W
0°
18°E
36°E
54°E
0.1
0.2
0.3
0.4
0.5

Time lag of identified feedbacks

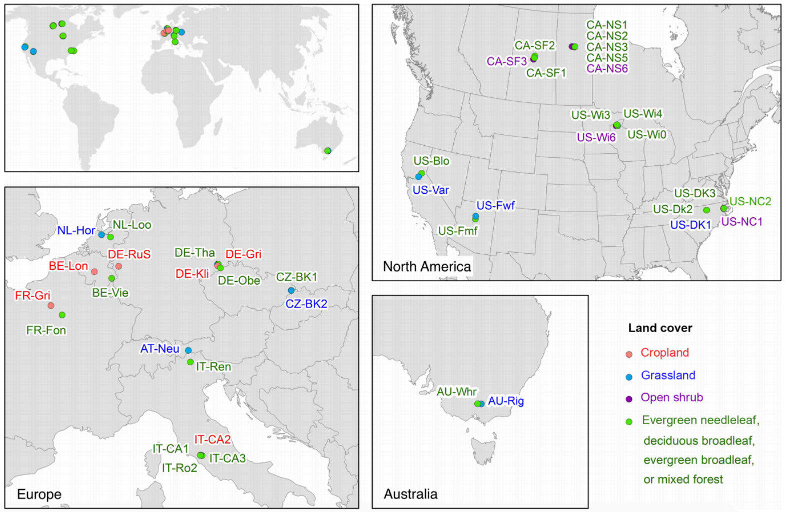


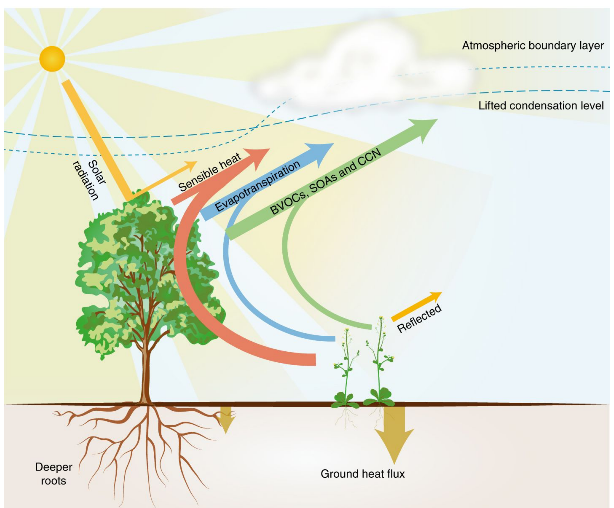
Warming impacts





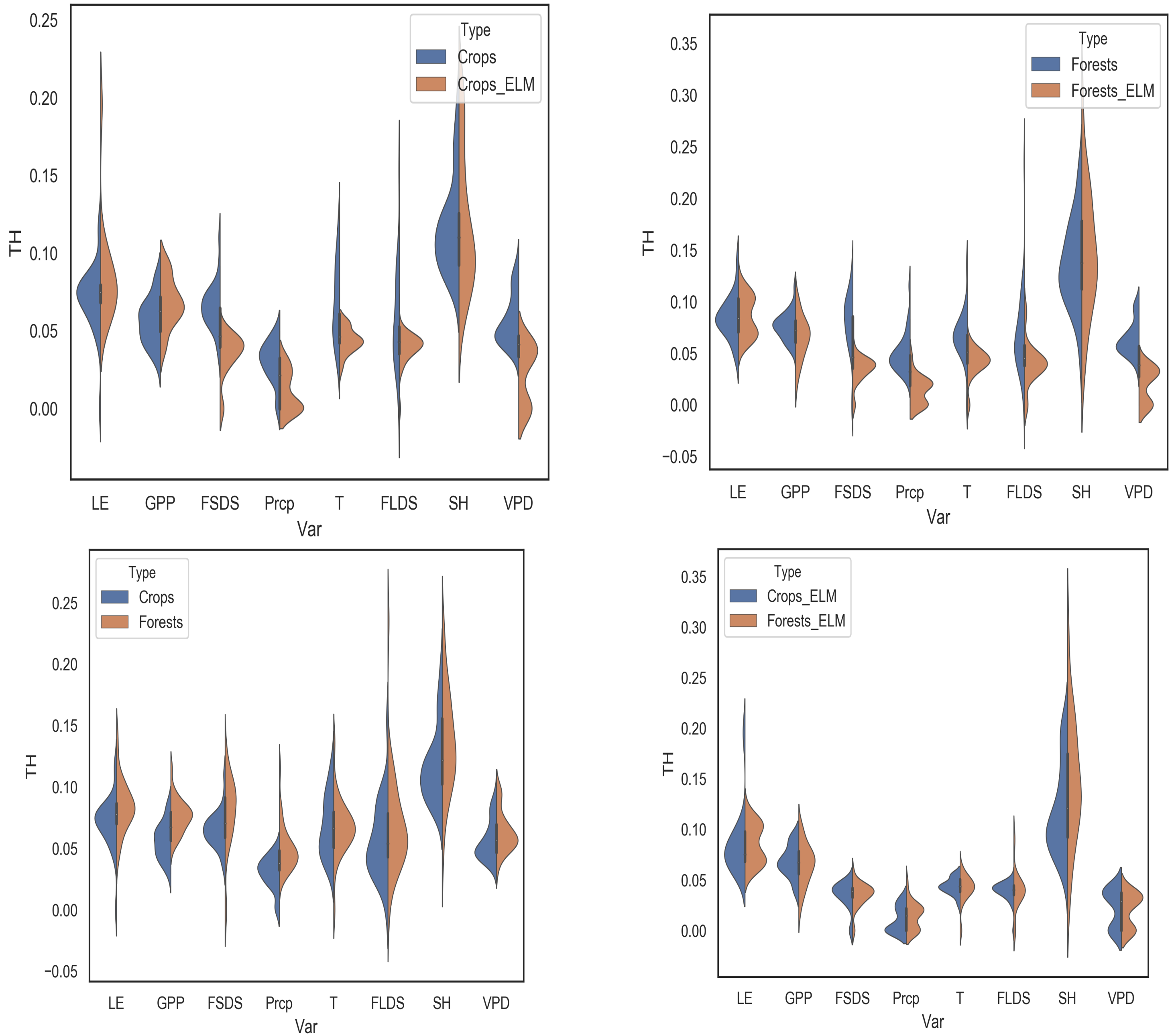
Case 2：deforestation





Land Cover Transition	Pair ID	Site Name	Latitude (degrees N)	Longitude (degrees E)	Elevation (m)	PFT	Distance (km)	Period
Evergreen Needleleaf Boreal Forests → Open Shrublands	1	CA-SF2	54.2539	-105.8775	520	2	19.87	2001-2005
		CA-SF3	54.0916	-106.0053	540	11		
	2	CA-NS2	55.9058	-98.5247	260	2	27.43	2001-2005
		CA-NS6	55.9167	-98.9644	244	11		
	3	CA-NS1	55.8792	-98.4839	260	2	30.25	2001-2005
		CA-NS6	55.9167	-98.9644	244	11		
	4	CA-NS5	55.8631	-98.485	260	2	30.48	2001-2005
		CA-NS6	55.9167	-98.9644	244	11		
	5	CA-NS3	55.9117	-98.3822	260	2	36.29	2001-2005
		CA-NS6	55.9167	-98.9644	244	11		
	6	CA-SF1	54.485	-105.8176	536	2	45.41	2003-2006
		CA-SF3	54.0916	-106.0053	540	11		
Evergreen Needleleaf Boreal Forests → Croplands	7	DE-Tha	50.9636	13.5669	380	2	8.46	2004-2014
		DE-Kli	50.8929	13.5225	480	15		
	8	DE-Obe	50.7836	13.7196	735	2	18.42	2008-2014
		DE-Kli	50.8929	13.5225	480	15		
Evergreen Needleleaf Boreal Forests → Grasslands	9	DE-Tha	50.9636	13.5669	380	2	4.12	2004-2010
		DE-Gri	50.9495	13.5125	385	13		
	10	DE-Obe	50.7836	13.7196	735	2	23.49	2008-2010
		DE-Gri	50.9495	13.5125	385	13		
	11	IT-Ren	46.5869	11.4337	1730	2	59.57	2002-2012
		AT-Neu	47.1167	11.3175	970	13		
Mixed Forests → Croplands*	12	BE-Vie	50.305	5.998	491	2/8	92.82	2004-2014
		BE-Lon	50.5515	4.7461	165	15		
	13	BE-Vie	50.3051	5.9981	493	2/8	69.96	2011-2014
		DE-RuS	50.8659	6.4472	102.76	15		
Evergreen Needleleaf Temperate Forests → Open Shrublands	14	US-Wi4	46.7393	-91.1663	352	1	16.22	2002-2003
		US-Wi6	46.6249	-91.2982	371	10		
	15	US-Wi0	46.6188	-91.0814	349	1	16.57	2002
		US-Wi6	46.6249	-91.2982	371	10		
	16	US-NC2	35.803	-76.6685	5	1	4.03	2005-2009
		US-NC1	35.8118	-76.7119	5	10		
Deciduous Broadleaf Temperate Forests → Open Shrublands	17	US-Wi3	46.6347	-91.0987	411	7	15.27	2002-2003
		US-Wi6	46.6249	-91.2982	371	10		
Evergreen Needleleaf Temperate Forests → Grasslands	18	US-Dk3	35.9782	-79.0942	163	1	0.78	2004-2008
		US-Dk1	35.9712	-79.0934	168	13		
	19	CZ-BK1	49.5021	18.5369	875	1	0.96	2004-2006
		CZ-BK2	49.4944	18.5429	855	13		
	20	US-Fmf	35.1426	-111.7273	546	1	33.91	2006-2010
		US-Fwf	35.4454	-111.7718	2270	13		
	21	NL-Loo	52.1666	5.7436	25	1	46.55	2004-2011
		NL-Hor	52.2404	5.0713	2.2	13		
	22	US-Blo	38.8953	-120.6328	1315	1	60.29	2000-2007
		US-Var	38.4133	-120.9507	129	13		
Deciduous Broadleaf Temperate Forests → Croplands	23	IT-CA1	42.3804	12.0266	200	7	0.36	2011-2013
		IT-CA2	42.3772	12.026	200	15		
	24	IT-CA3	42.38	12.0222	197	7	0.44	2011-2013
		IT-CA2	42.3772	12.026	200	15		
	25	IT-Ro2	42.3903	11.9209	160	7	8.75	2011-2012
		IT-CA2	42.3772	12.026	200	15		
	26	FR-Fon	48.4764	2.7801	90	7	73.3	2005-2013
		FR-Gri	48.8442	1.9519	125	15		
Deciduous Broadleaf Temperate Forests → Grasslands	27	US-Dk2	35.9736	-79.1004	168	7	0.68	2003-2008
		US-Dk1	35.9712	-79.0934	168	13		
Evergreen Broadleaf Temperate Forests → Grasslands	28	AU-Whr	-36.6732	145.0294	165	5	48.82	2011-2013
		AU-Rig	-36.6499	145.5759	152	13		
Deciduous Broadleaf Tropical Forests → C_4 Grasslands	29	PA-SPn	9.3181	-79.6346	78	6	0.5932	2007-2009
		PA-SPs	9.3138	-79.6314	68	14		

Chen et al., 2018


Type
Crops
Crops_ELM
TH
0.25
0.20
0.15
0.10
0.05
0.00
LE
GPP
FSDS
Prcp
T
FLDS
SH
VPD
Var
Type
Forests
Forests_ELM
TH
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−0.05
LE
GPP
FSDS
Prcp
T
FLDS
SH
VPD
Var
Type
Crops
Forests
TH
0.25
0.20
0.15
0.10
0.05
0.00
−0.05
LE
GPP
FSDS
Prcp
T
FLDS
SH
VPD
Var
Type
Crops_ELM
Forests_ELM
TH
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
LE
GPP
FSDS
Prcp
T
FLDS
SH
VPD
Var

Confusion Matrix

	LE->EF	FSDS->EF	T->EF	SH->EF	GPP->EF	Prcp->EF	FLDS->EF	VPD->EF
Accuracy	0.690	0.448	0.414	0.724	0.680	0.310	0.517	0.538

LE->EF
[[5 2]
[7 15]]

FSDS->EF
[[3 3]
[13 10]]

T->EF
[[7 6]
[11 5]]

SH->EF
[[3 2]
[6 18]]

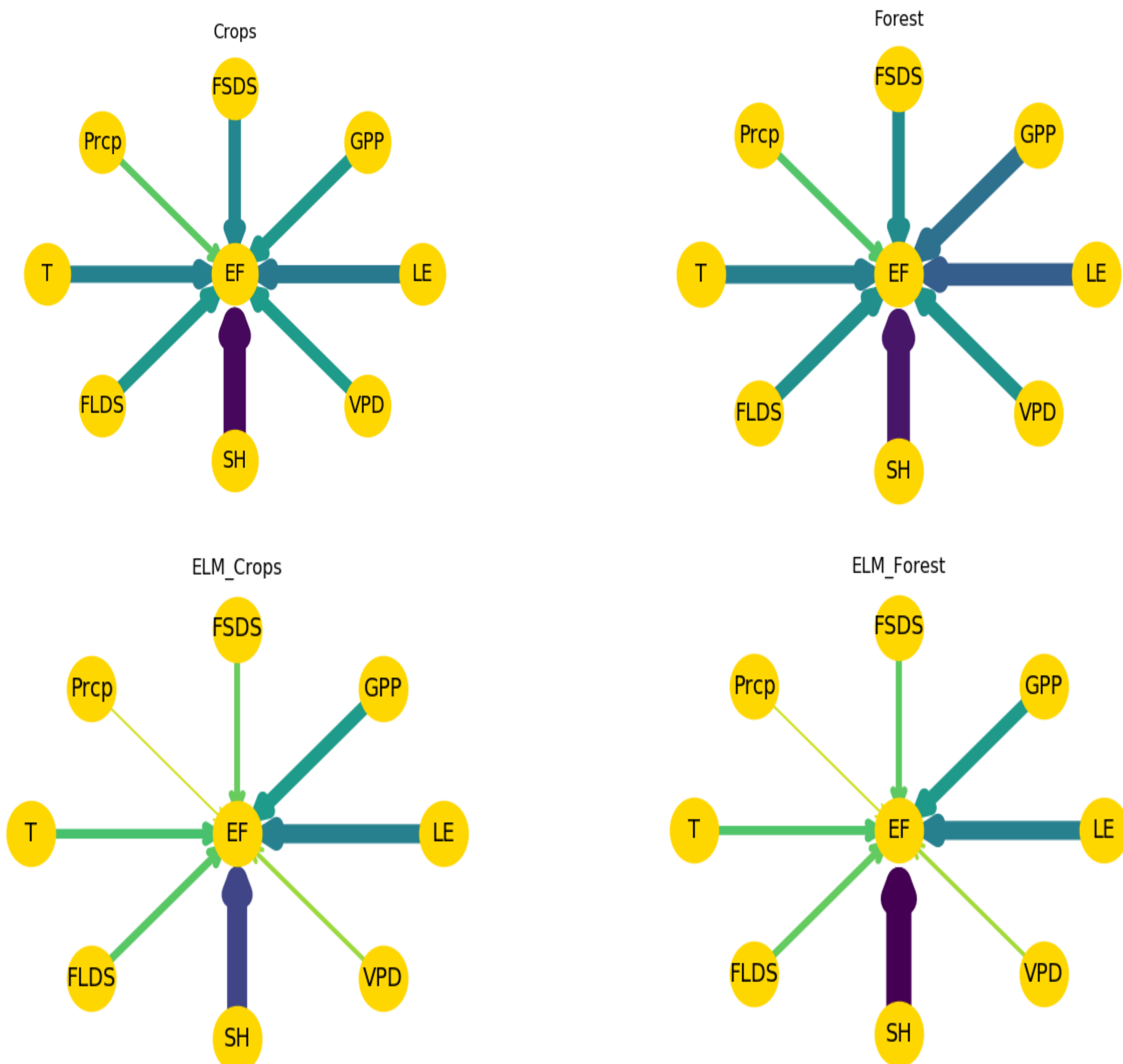
GPP->EF
[[5 0]
[8 12]]

Prcp->EF
[[3 3]
[17 6]]

FLDS->EF
[[6 3]
[11 9]]

VPD->EF
[[10 1]
[11 4]]

Reconstructed Network



Conclusions

1. Earth System Models that capture observed emergent patterns, not necessarily get the right answer with right reasons.
2. Causality benchmark offers a new way to understand model bias
3. Earth system dynamics are often nonlinear, and hard to investigate with linear approach
4. Earth system process to process relationship, often subject to a certain time lag/lead

Ongoing work & next step

- Explore alternative causality metrics
- Apply causality inference to carbon-climate, carbon-concentration feedbacks
- Apply causality inference to FLUXNET2015 dataset, reconstruct whole process network for major biomes
- Integrate causality benchmark metrics in ILAMB framework

Thanks!