Global crop yield prediction using seasonal climate forecast data

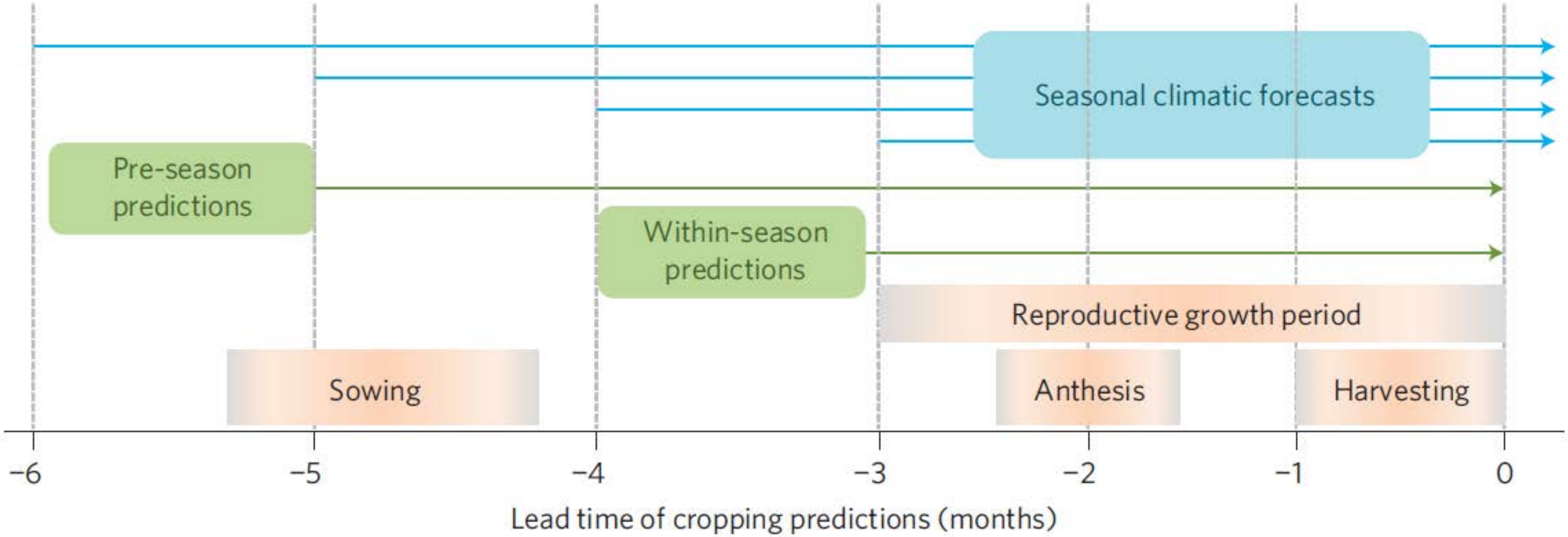
Toshichika Iizumi

National Agriculture and Food Research Organization (NARO)

Japan

Purpose of this talk

- Satellite remote sensing is powerful in monitoring global crop production.
- Seasonal climate forecasting can offer additional information for food agencies to strength their capacity.
- This talk presents recent use of climate forecast data in yield predictions for a large spatial domain (subnational to global).



Climate forecast-based yield prediction

nature climate change

LETTERS

PUBLISHED ONLINE: 21 JULY 2013 | DOI: 10.1038/NCLIMATE1945

Prediction of seasonal climate-induced variations in global food production

Toshichika Iizumi[1]*, Hirofumi Sakuma[2,3], Masayuki Yokozawa[1], Jing-Jia Luo[4], Andrew J. Challinor[5,6], Molly E. Brown[7], Gen Sakurai[1] and Toshio Yamagata[3]

ELSEVIER

Contents lists available at ScienceDirect

Climate Services

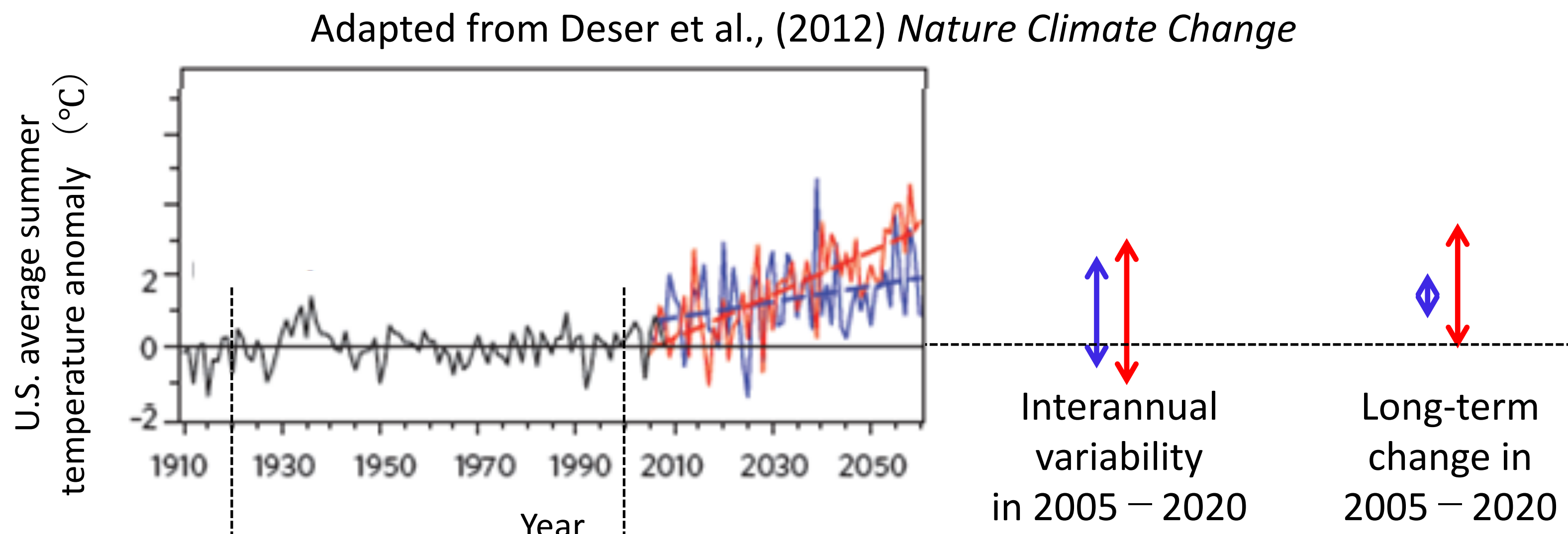
journal homepage: www.elsevier.com/locate/cliser

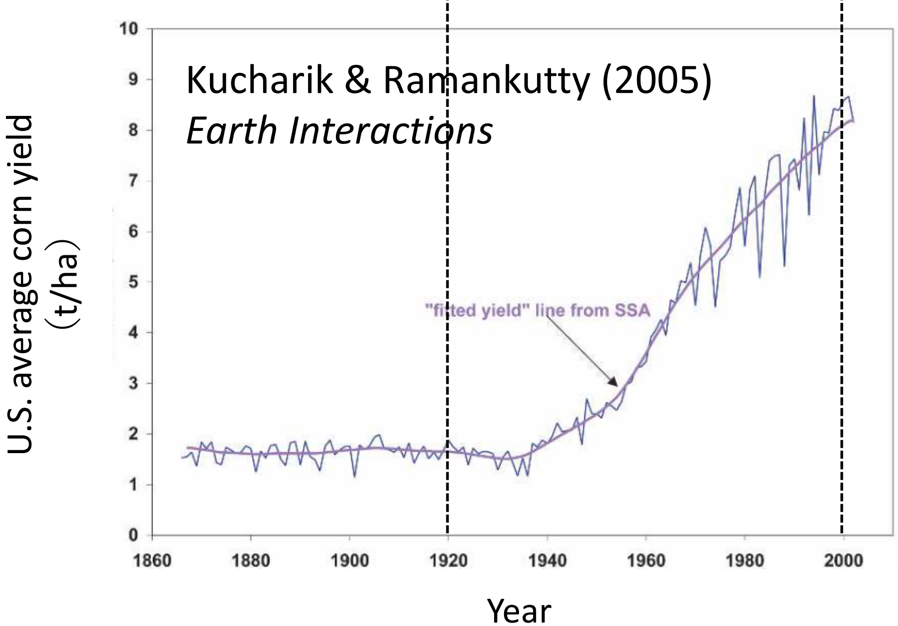
Original research article

Global crop yield forecasting using seasonal climate information from a multi-model ensemble

Toshichika Iizumi[a,*], Yonghee Shin[b], Wonsik Kim[a], Moosup Kim[b], Jaewon Choi[b]

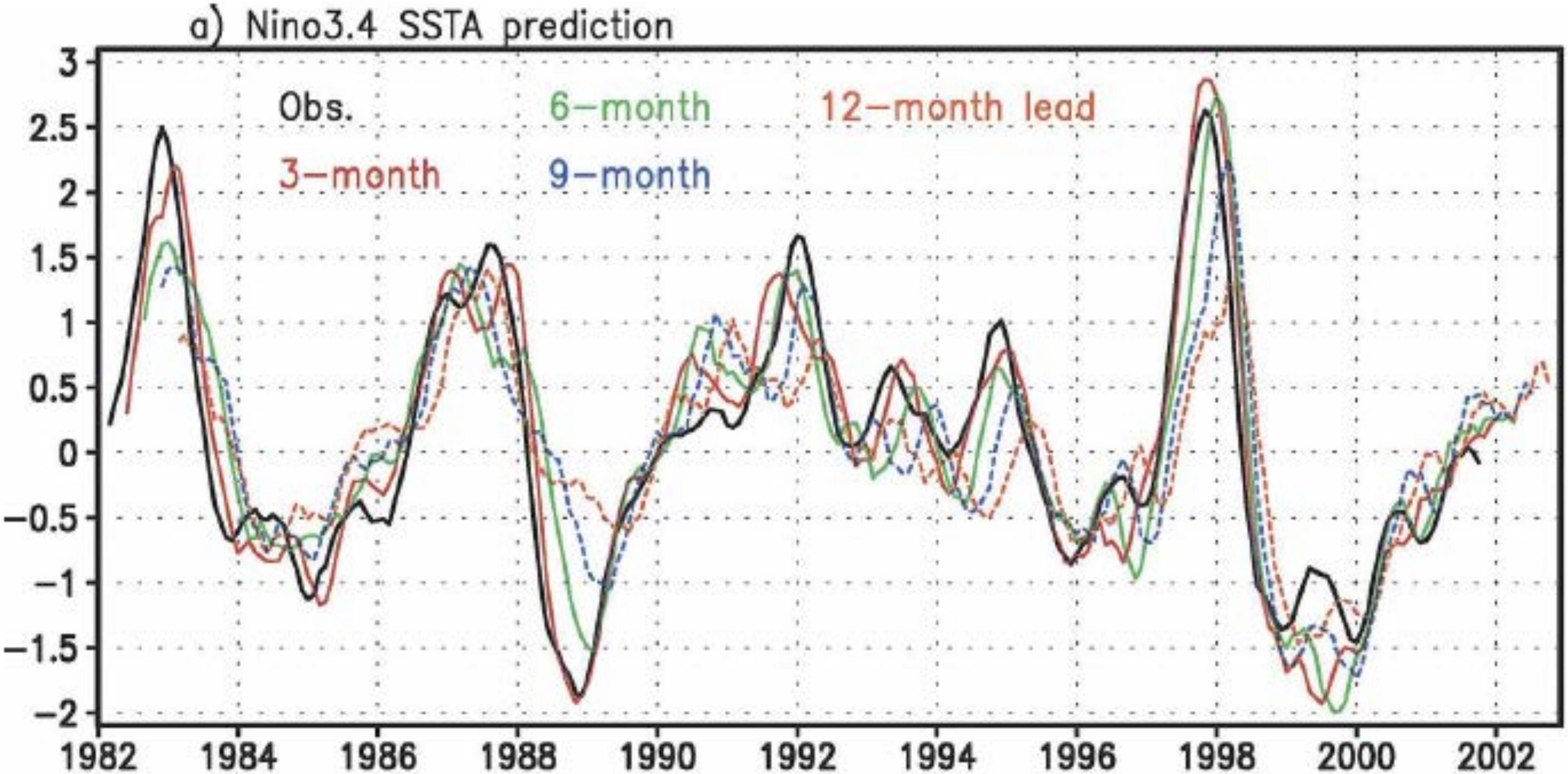
Seasonal forecasting and climate change adaptation





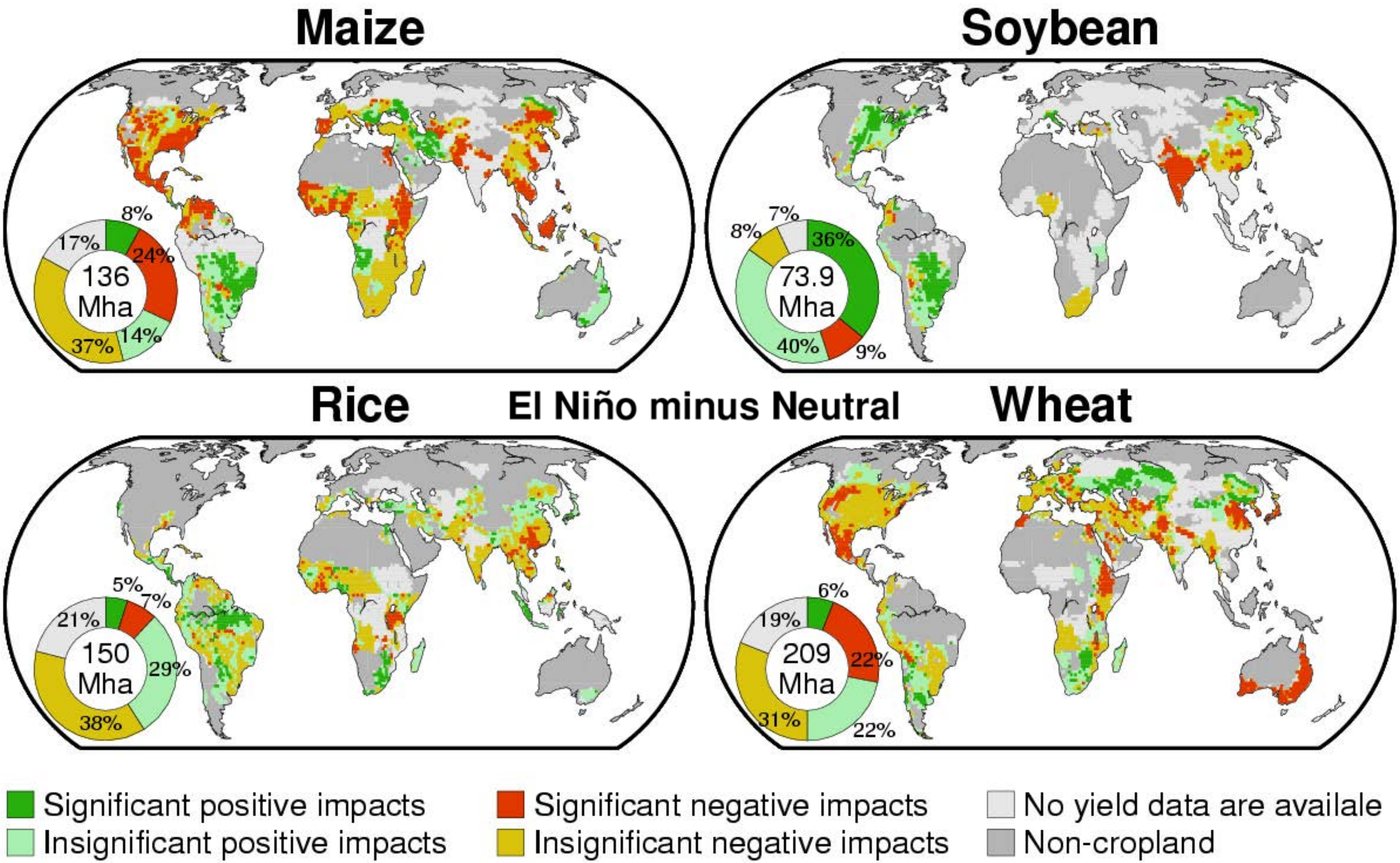
- The amplitude of interannual temperature variability is in general larger than long-term temperature change.
- Responding better to seasonal climate-induced supply shocks will increase society's capability to adapt climate change by up to the middle of this century.

ENSO forecasts are skilful even at 12 months lead time



Luo et al. (2005) *Journal of Climate*, https://doi.org/10.1175/JCLI3526.1

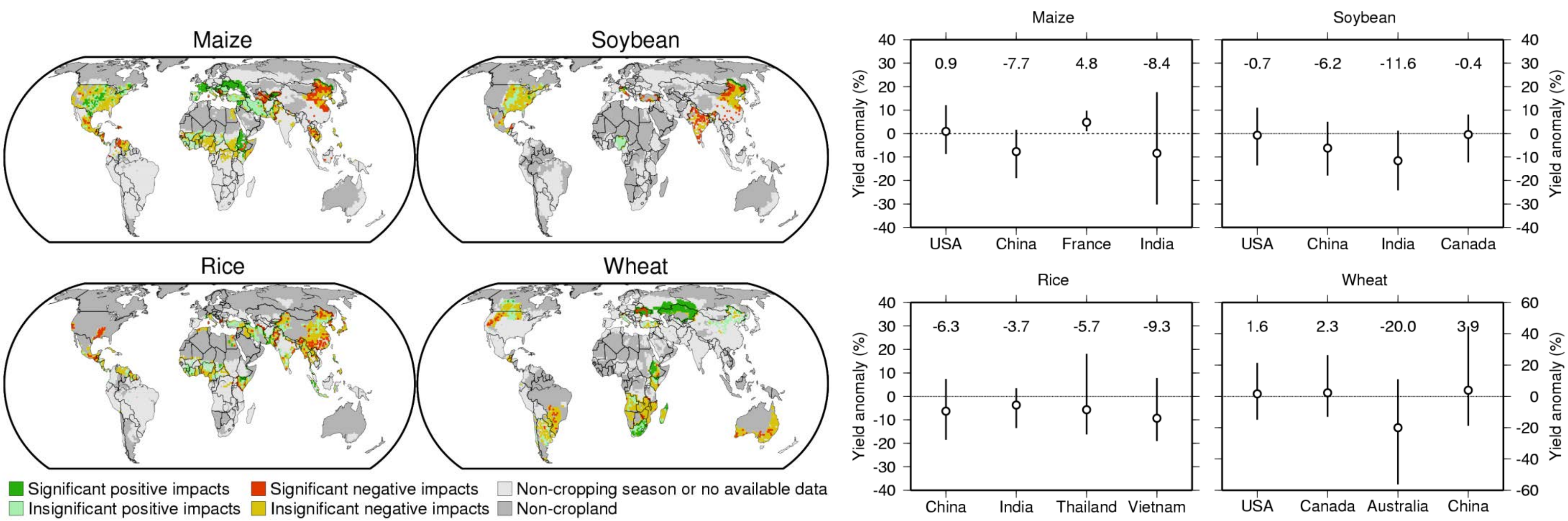
Yield impacts associated with ENSO



Iizumi et al., 2014, Nature Communications, doi:10.1038/ncomms4712

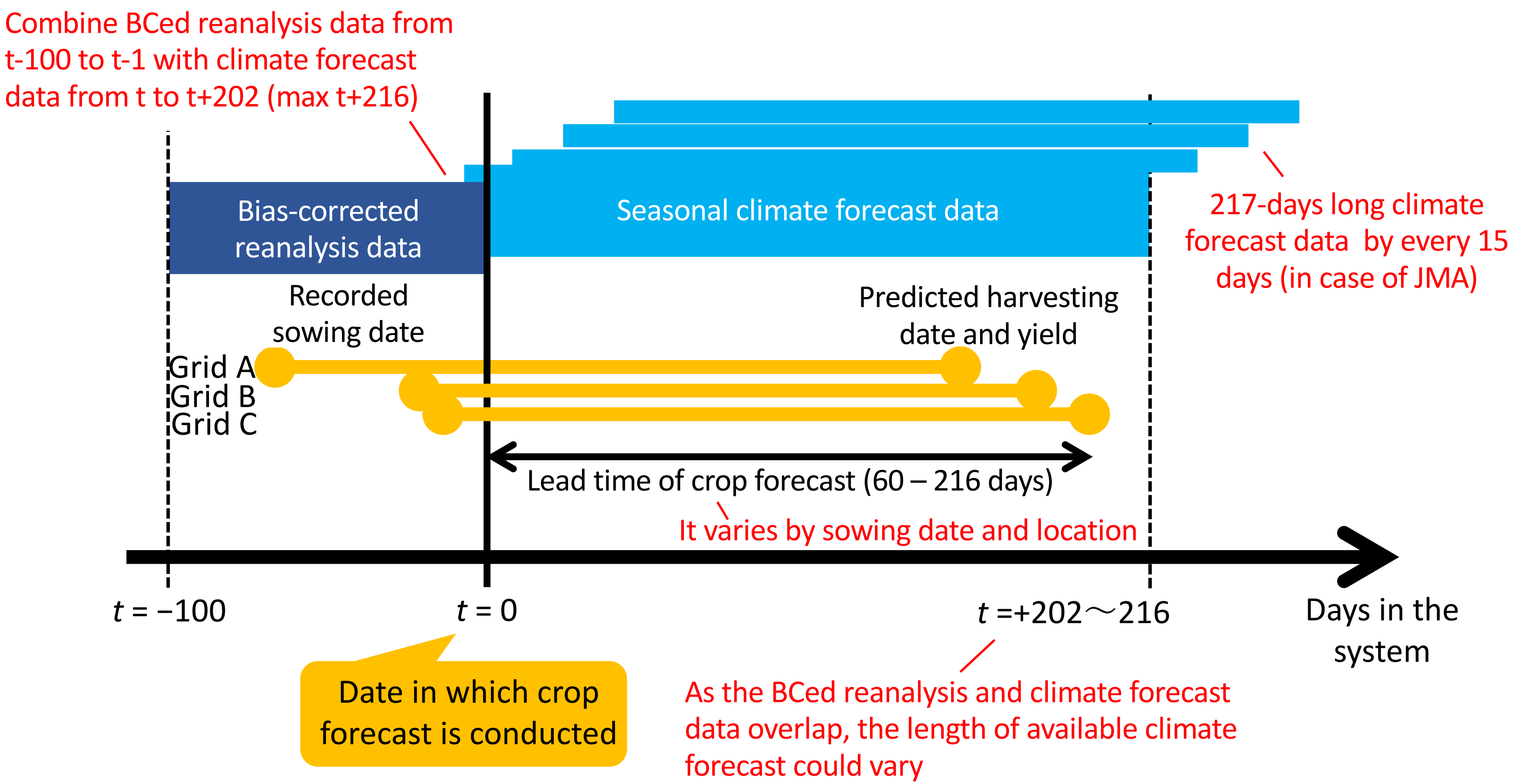
Yield prediction for 2014 El Nino

- The incidence of El Nino was predicted in spring 2014.
- Predictions on possible variations in 2014-fall yields (lead time of +1 to +6 months) were released on July 31, 2014 via the *Monthly Oversea Food Demand & Supply Report* of MAFF.

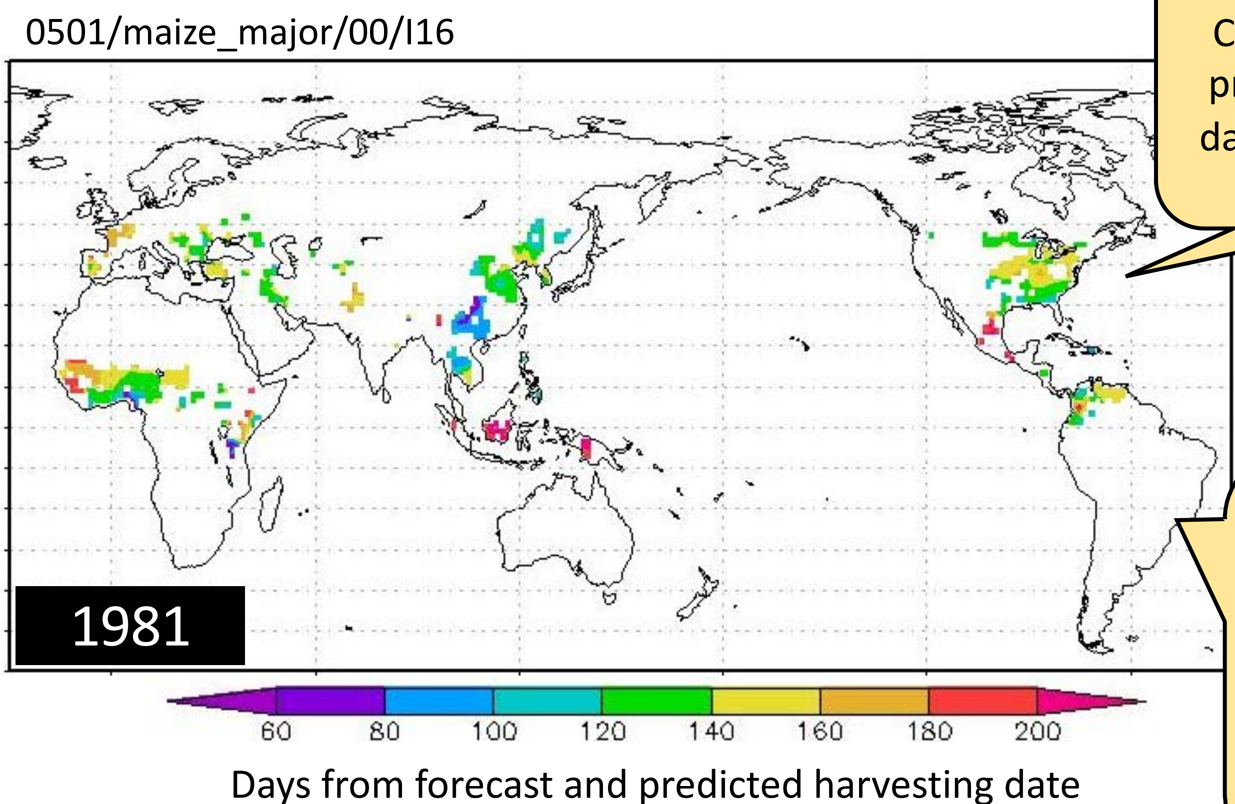


Full text (in Japanese) is available:
http://www.maff.go.jp/j/zyukyu/jki/j_rep/monthly/201407/pdf/21_monthly_topics_1.pdf

Seasonal crop forecasting system (prototype)



Example of maize forecast made on May 1st

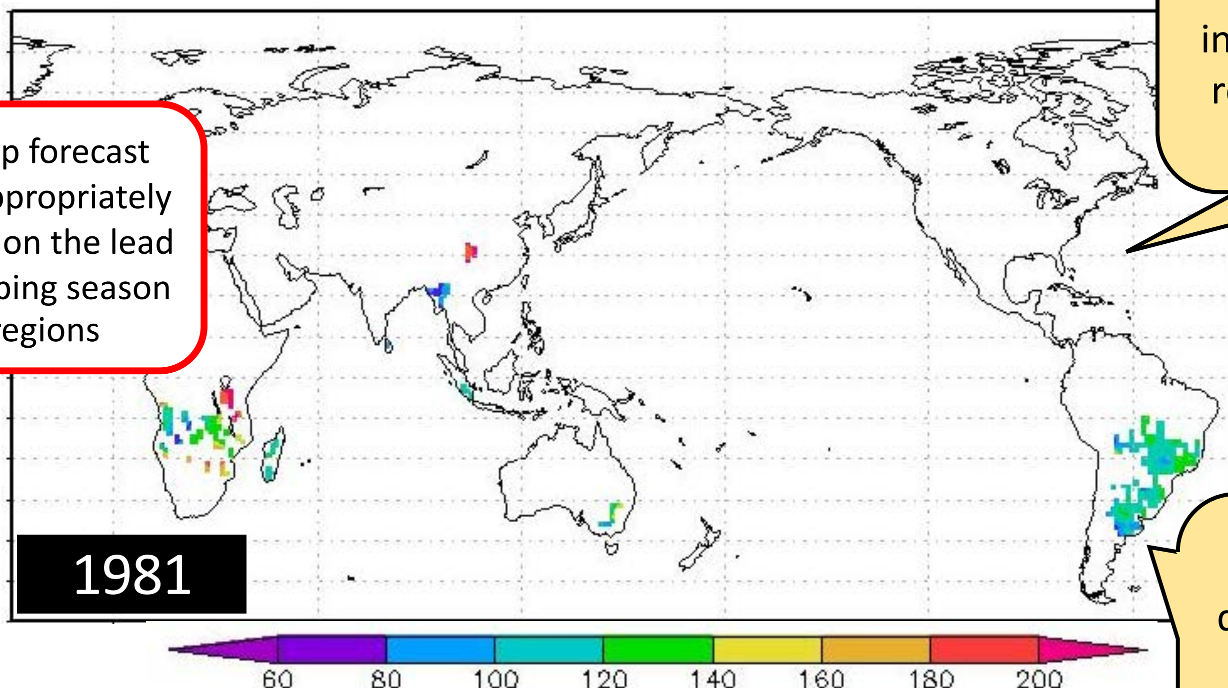


Harvesting date in Cornbelt in the US is predicted 140 to 160 days ahead (Sep 18 to Oct 8)

No crop forecast is made as major maize in South America was just after the harvesting. Secondary maize is separately predicted in this system.

Example of maize forecast made on December 15th

1215/maize_major/00/I16



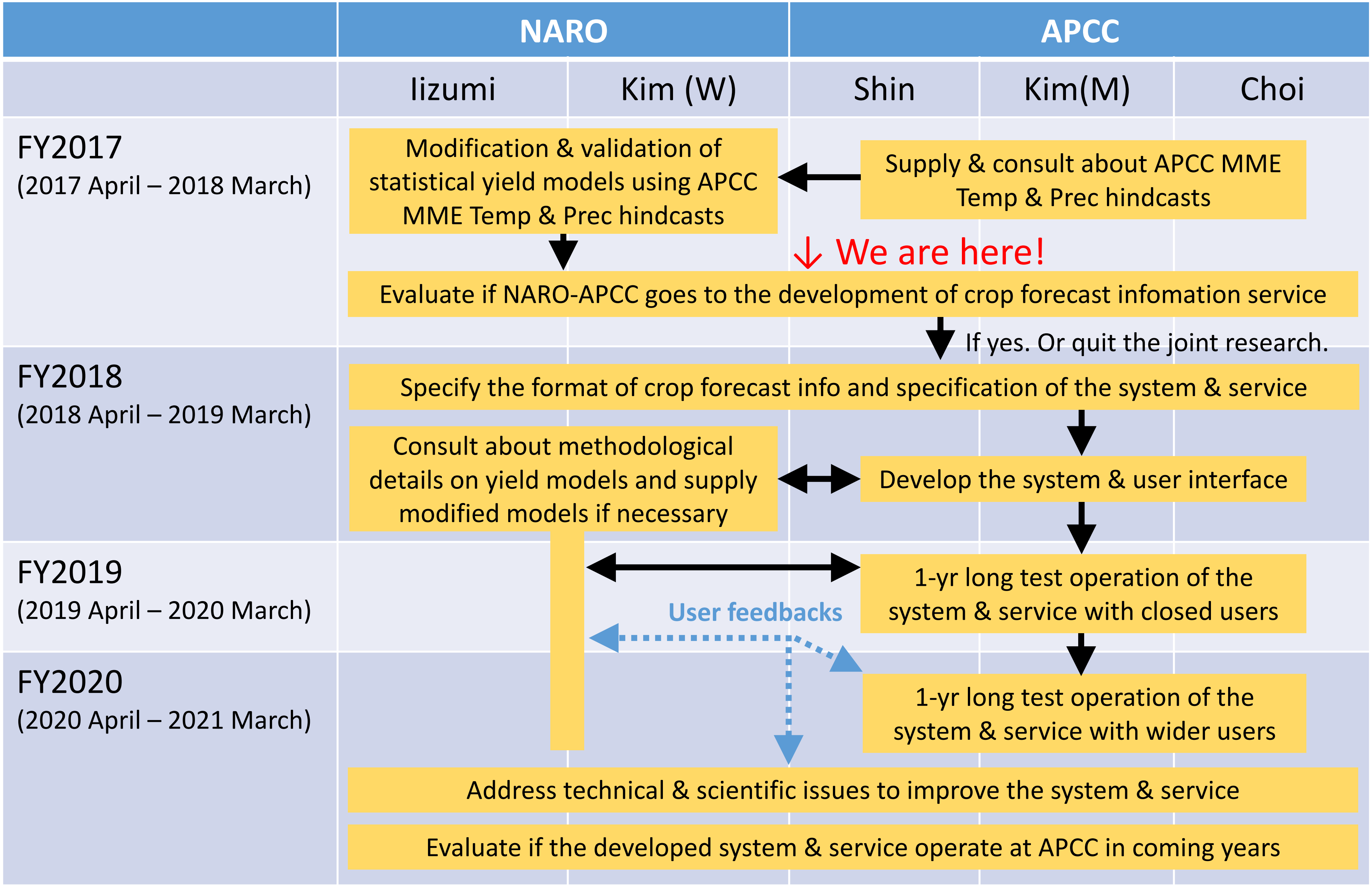
Days from forecasting to predicted harvesting (= lead time)

Timing of crop forecast needs to be appropriately set depending on the lead time and cropping season in target regions

No maize forecast is made for Northern Hemisphere as maize in that region does not reach maturity within 214 days

Predicted harvesting date in South America is 100 to 140 days ahead (March 25th to May 5th, 1982

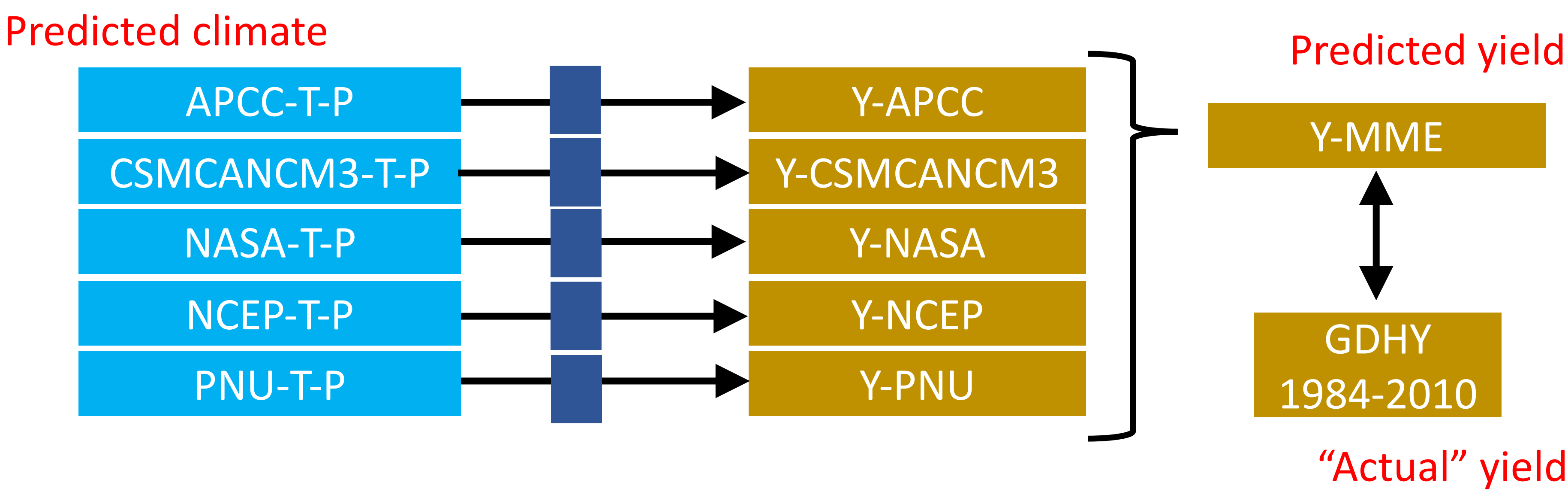
NARO-APCC Joint Research



Hindcast experiments

How good a multi-model ensemble of climate forecasts are to predict yield variability 3- or 6-month before the harvesting?

* Leave-three-out cross-validation technique was used.



* A single best climate model was selected based on independent subset.

Statistical yield model

$$\Delta Y_t = \frac{(Y_t - Y_{t-1})}{\bar{Y}_{t-3:t-1}} \times 100 \quad (1)$$

$$\Delta T_t = T_t - T_{t-1} \quad (2) \qquad \Delta P_t = P_t - P_{t-1} \quad (3)$$

$$\Delta Y_t = a_0 + a_1 \cdot \Delta T_t + a_2 \cdot \Delta P_t + \varepsilon \quad (4)$$

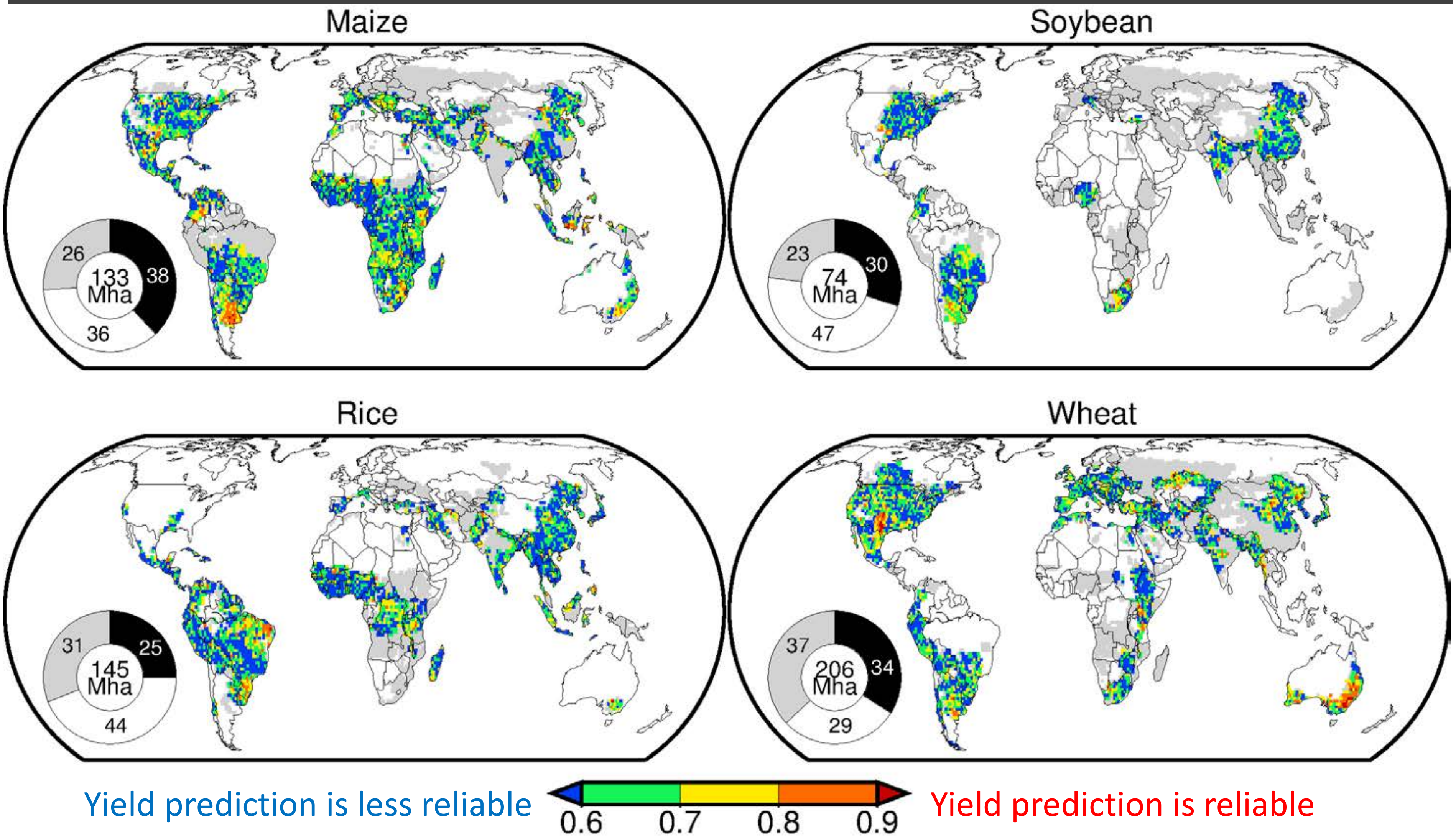
ΔY, yield anomaly (%); Y, yield (t/ha); $\bar{Y}$, average yield (t/ha)

ΔT, RGP mean T anomaly (°C); ΔP, RGP mean P anomaly (mm/d)

a_0, a_1, a_2, regression coefficients; ε, error term

Subscript: t, year; Regression coefficients were estimated by crop, cropping season and grid cell.

Performance of 3-month-lead yield prediction



Good skill (ROC is significant at 10%)

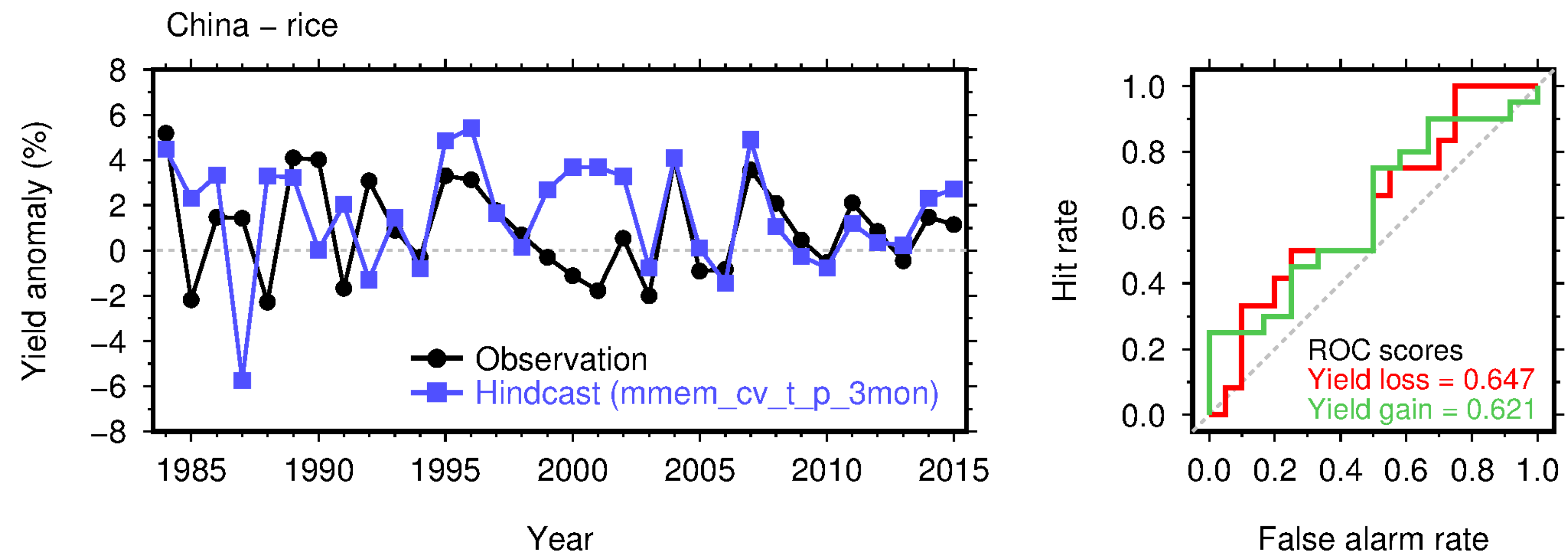
No skill (ROC is not significant)

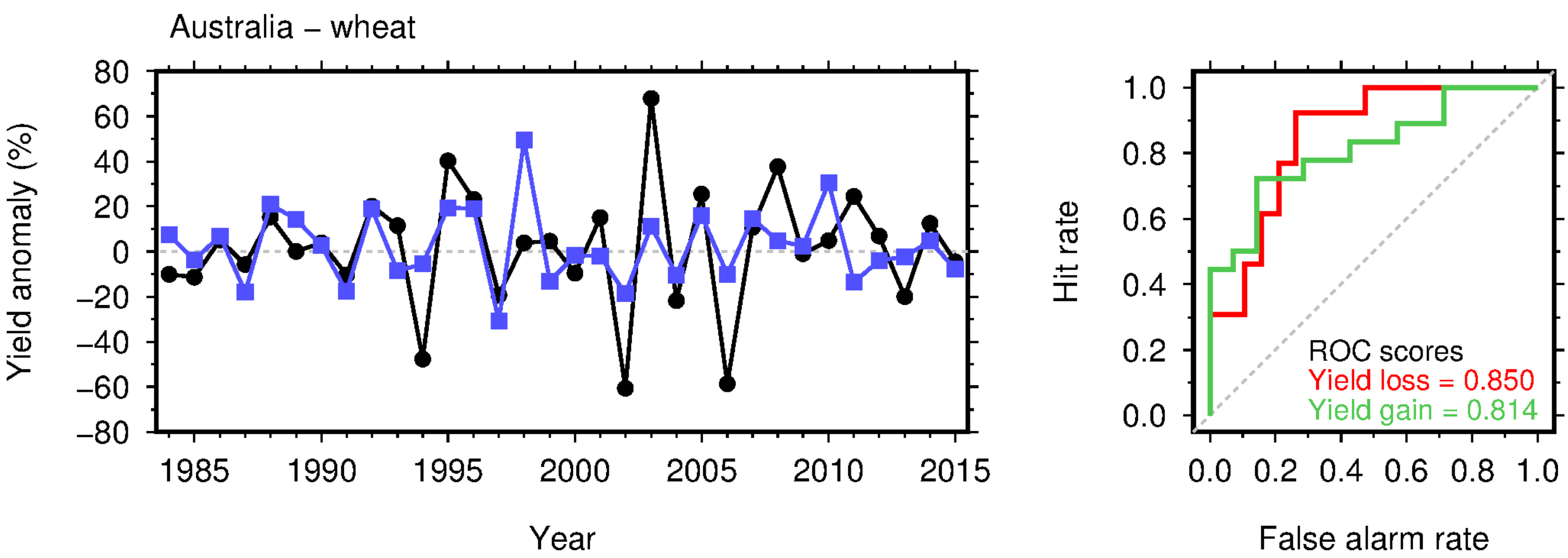
No data available

ROC score

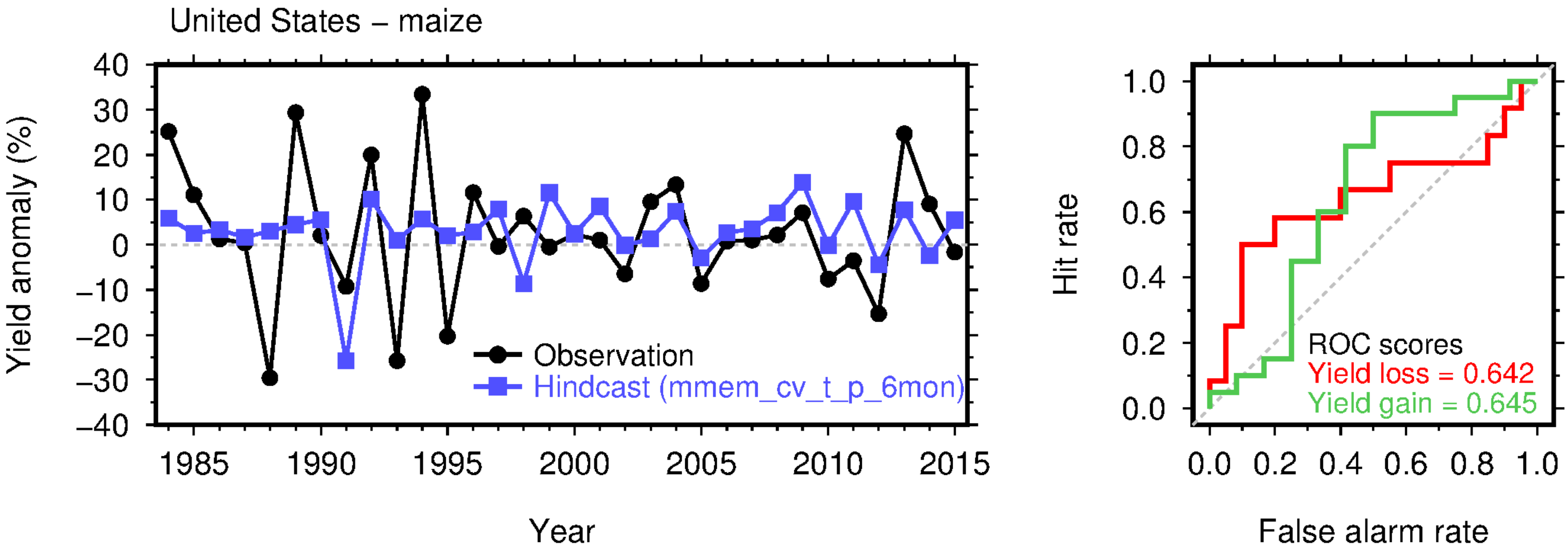
Iizumi et al., 2018, *Climate Services*, https://doi.org/10.1016/j.cliser.2018.06.003

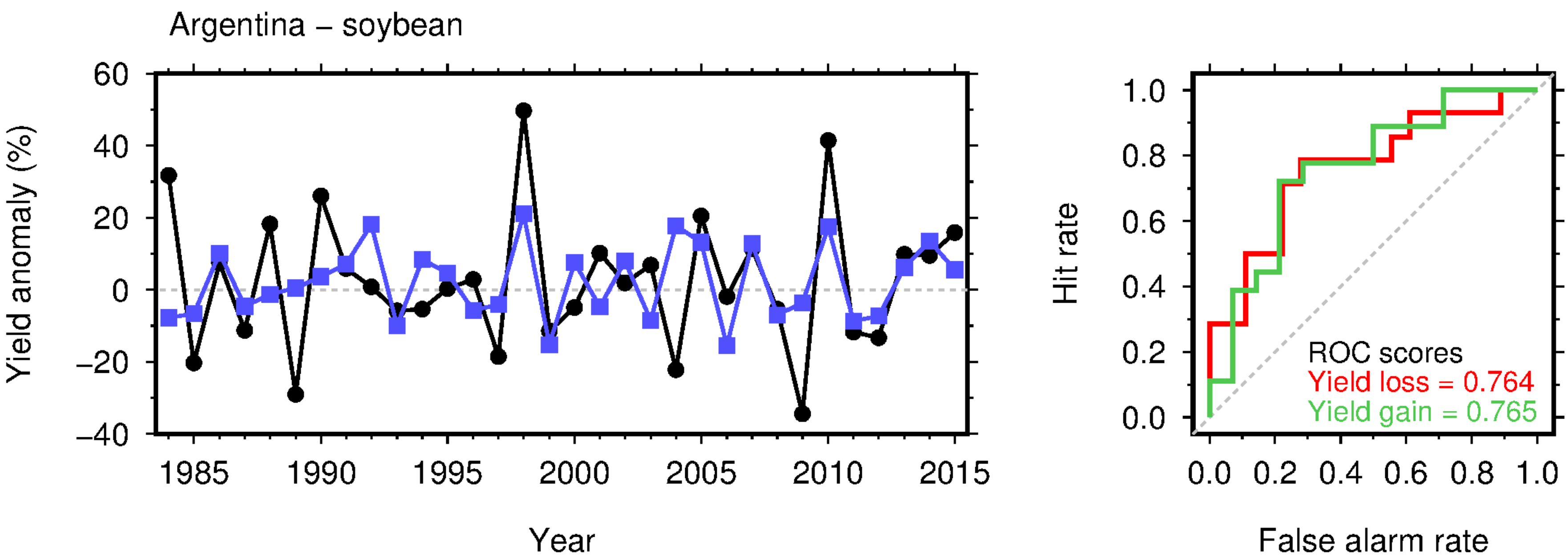
Prediction of country average yield variability (3-mon lead)





Prediction of country average yield variability (6-mon lead)



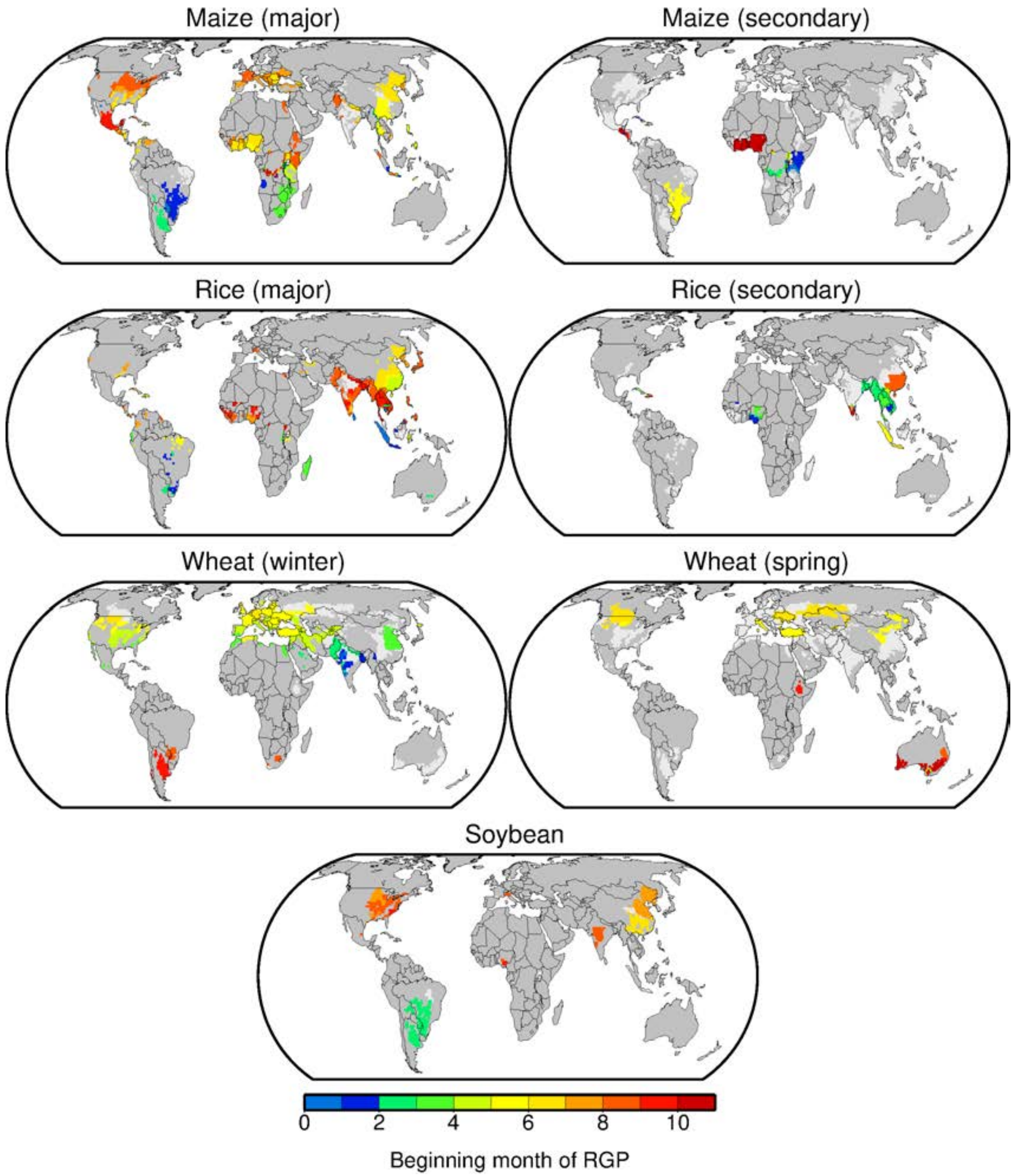


Some thoughts

- Year-to-year variations in yield in many parts of global harvested area can be predicted several months before harvesting.
- There are differences between the two methods:
  - Climate forecasts likely have a longer lead time than RS in expense of spatial resolution (0.5^{o} to 1^{o}).
  - Reliability of climate forecasts varies by season and region (⇔RS has the consistent quality over season).
  - Yield is the only variable for climate forecast, whereas RS can derived multiple variables (area planted).
  - Climate forecasts is useless when non-climatic factors lead to yield loss (e.g., landslides, pests).
- Better understanding pros and cons of these two approaches would benefit food agencies to think about best mix.
- A joint work with JAXA is ongoing to explore better yield prediction methods using satellite and climate data.

Questions?

Reproductive growth period (RGP)



SAGE global crop calendar

(Sacks et al., 2010, *Global Ecology & Biogeography*, doi:10.1111/j.1466-8238.2010.00551.x)

$$m_{\text{end}} = \begin{cases} 12 & d_{\text{h}} \le 15 \text{ and } m_{\text{h}} = 1 \\ m_{\text{h}} - 1 & d_{\text{h}} \le 15 \text{ and } m_{\text{h}} \ge 2 \\ m_{\text{h}} & d_{\text{h}} > 15 \end{cases}$$

$$m_{\text{start}} = \begin{cases} m_{\text{end}} - 2 + 12 & 1 \le m_{\text{end}} \le 2 \\ m_{\text{end}} - 2 & 3 \le m_{\text{end}} \le 12 \end{cases}$$

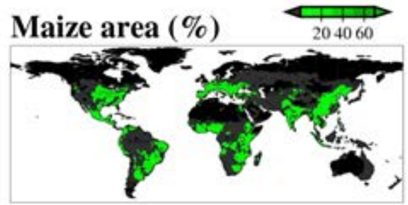
m_{start} & m_{end}, the month in which RGP starts and ends, respectively.

m_{h} & d_{h}, harvesting month and date.

Global dataset of historical yields (GDHY)

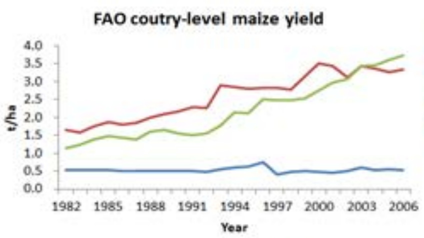
Satellite products (NOAA/AVHRR-NPP)

Harvested area & yield in 2000 (Monfreda et al. 2008)



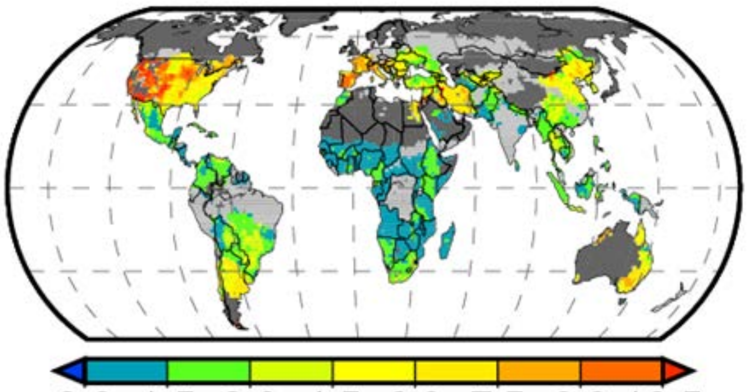
Iizumi et al., 2014,
Global Ecology and Biogeography,
doi:10.1111/geb.12120

FAO country yield statistics

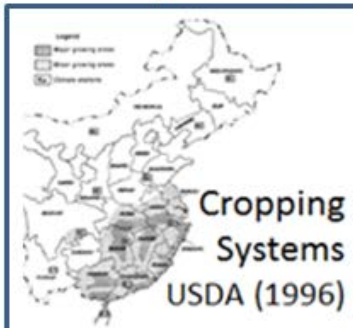


Global Dataset of Historical Yields

Maize

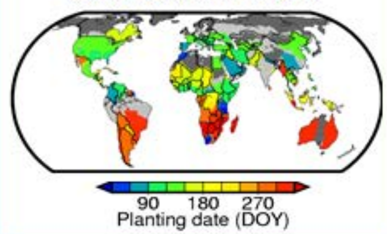


Average yield in 2001-2006 (t/ha)



Cropping Systems USDA (1996)

Global crop calendar (Sacks et al. 2008)



Version 1.2 (0.5°; 1981-2011; doi:10.20783/DIAS.528) is available online
http://search.diasjp.net/en/dataset/GDHY_v1_2

Reliability of grid-cell yield estimates in GDHY dataset

Elliott et al., 2015, *Geoscientific Model Development*, doi:10.5194/gmd-8-261-2015

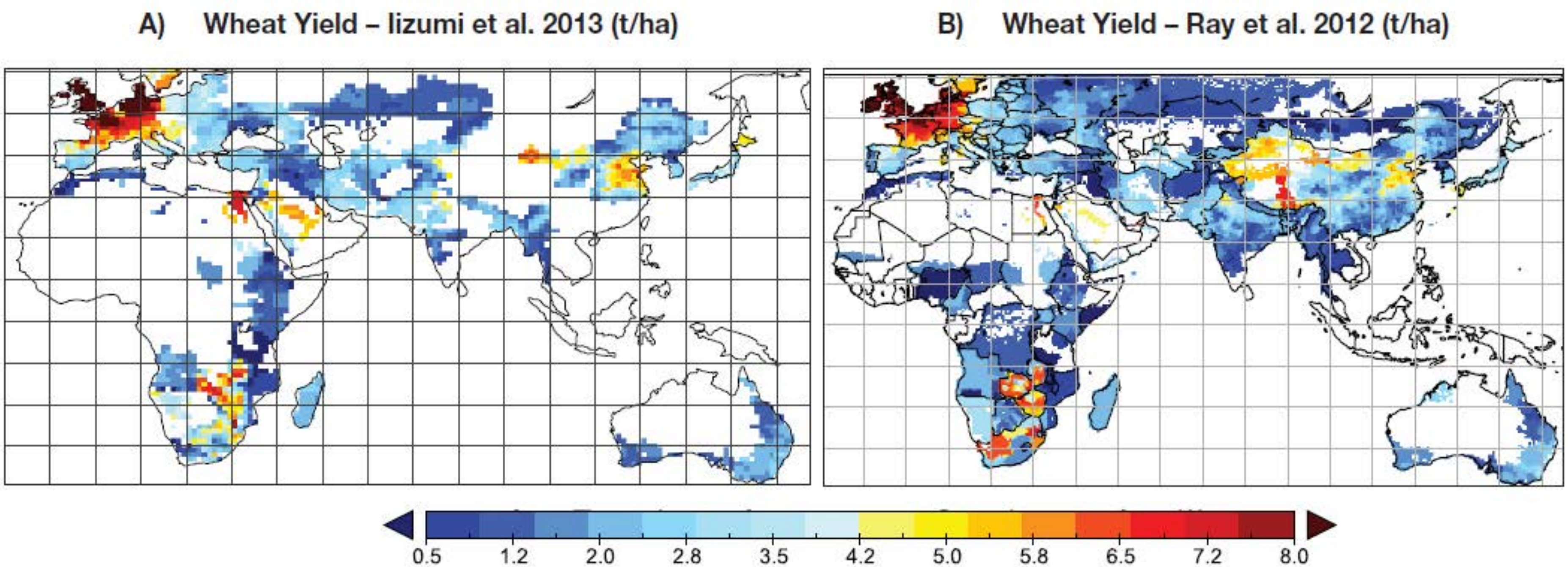


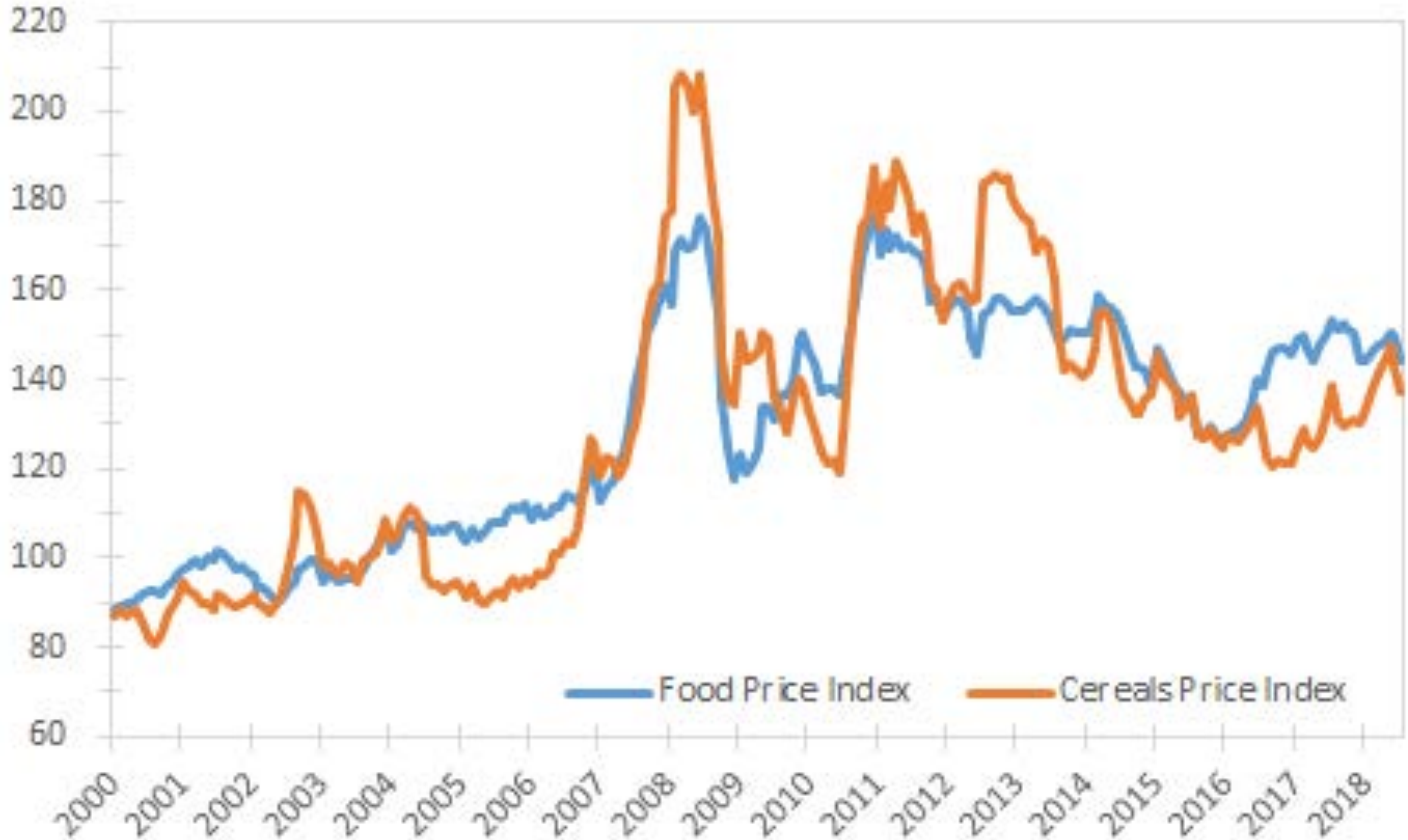
Figure 4. Example of historical evaluation data for year 2000 wheat yields from **(a)** Iizumi et al. (2013) (at 1.125° spatial resolution) and **(b)** Ray et al. (2012) (aggregated from 5 arc minutes to 0.5°).

In-depth validation of GDHY dataset is available soon

Iizumi et al., 2018, *PLOS ONE*, doi:10.1371/journal.pone.0203809 (in press)

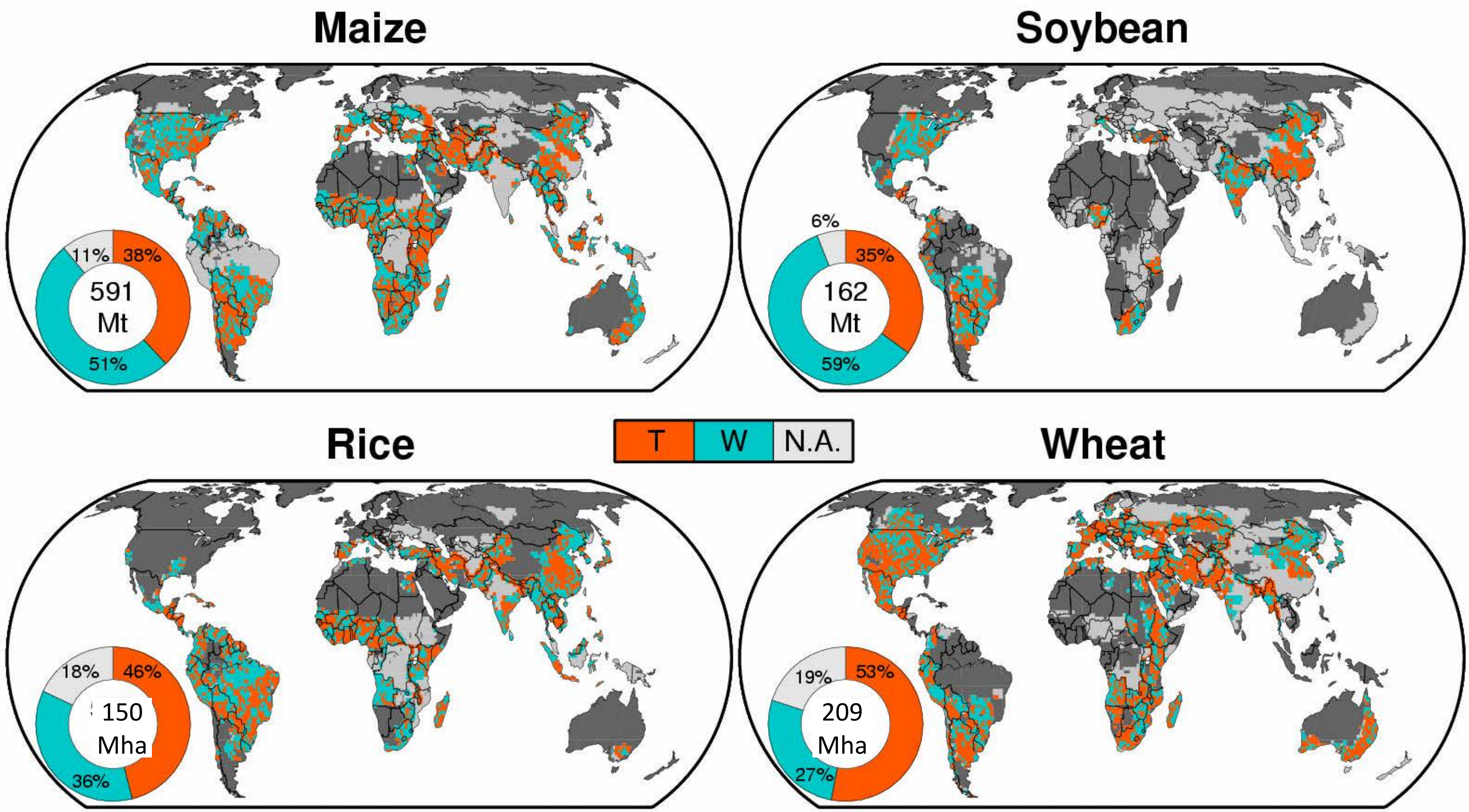
Relatively high food price is persistent

Monthly deflated data; 2002-2004=100



FAO, http://www.fao.org/worldfoodsituation/foodpricesindex/en/

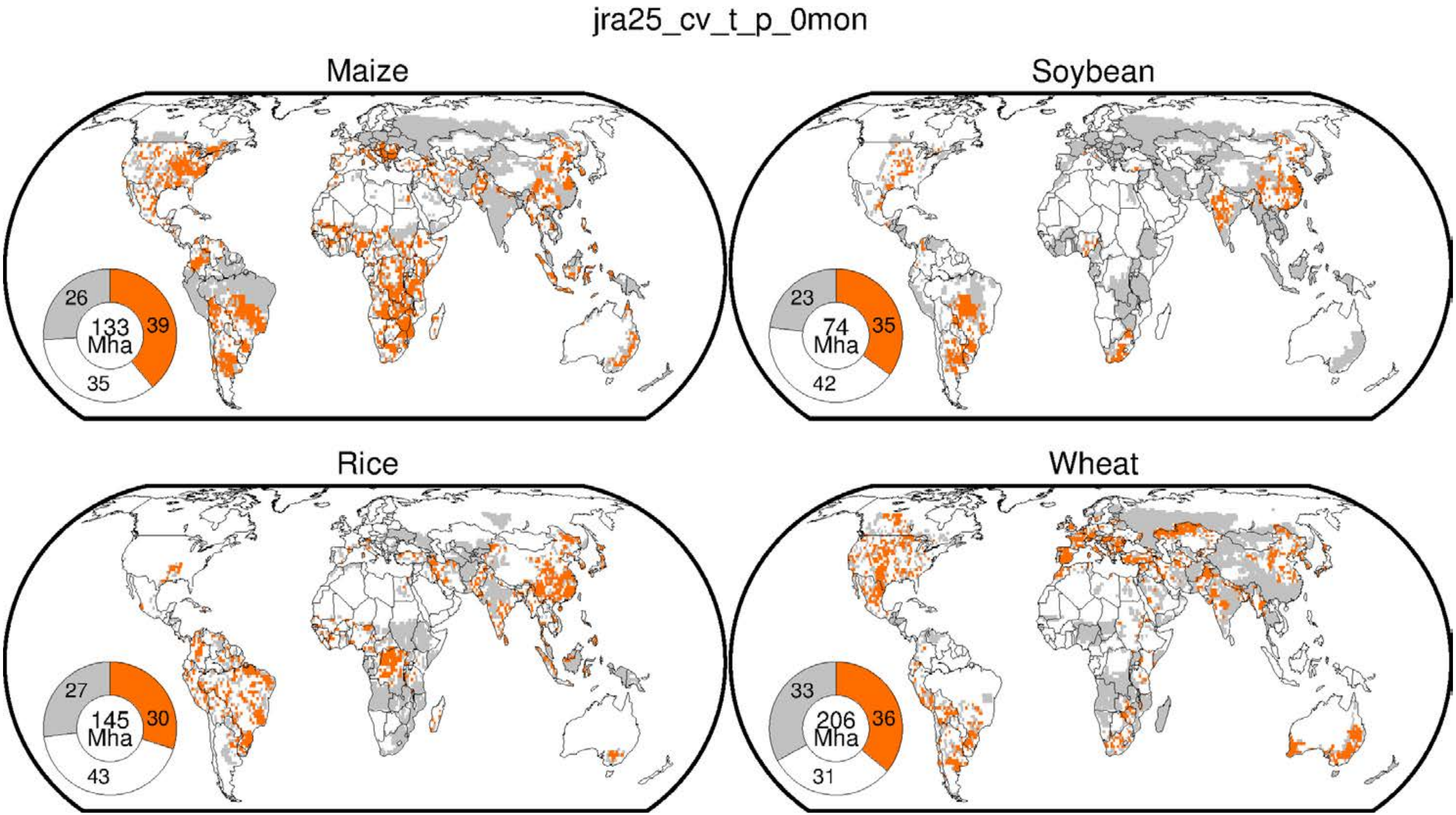
Main climatic driver of yield variability



Temperature-driving / Soil moisture-deriving

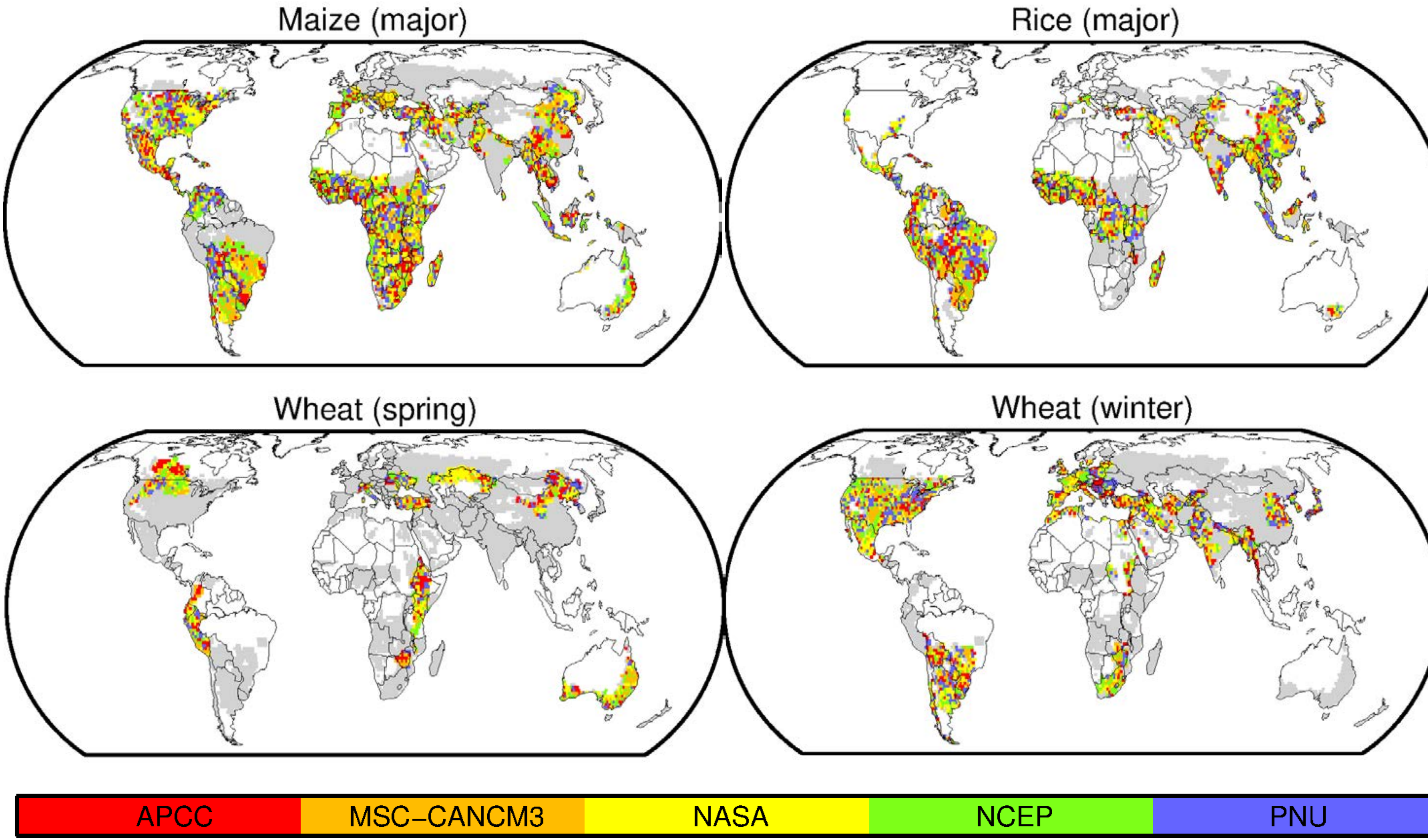
Iizumi et al., 2013, Nature Climate Change, doi:10.1038/NCLIMATE1945

Yield prediction using “actual” climate (0-mon lead)



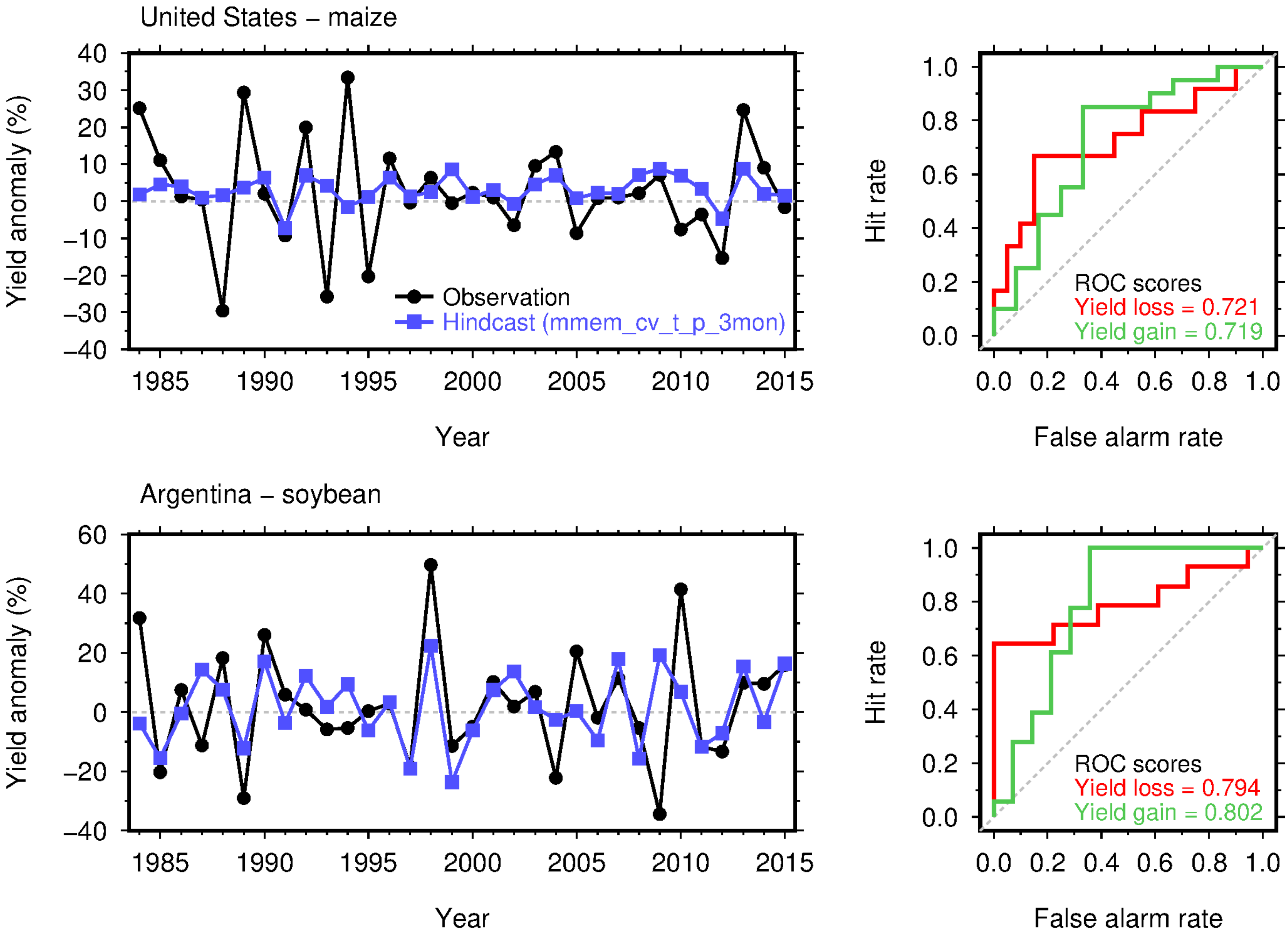
- Good skill (ROC is significant at 10%)
- No skill (ROC is not significant)
- No data available

The mosaic method selects the best-performing GCM by location and cropping season

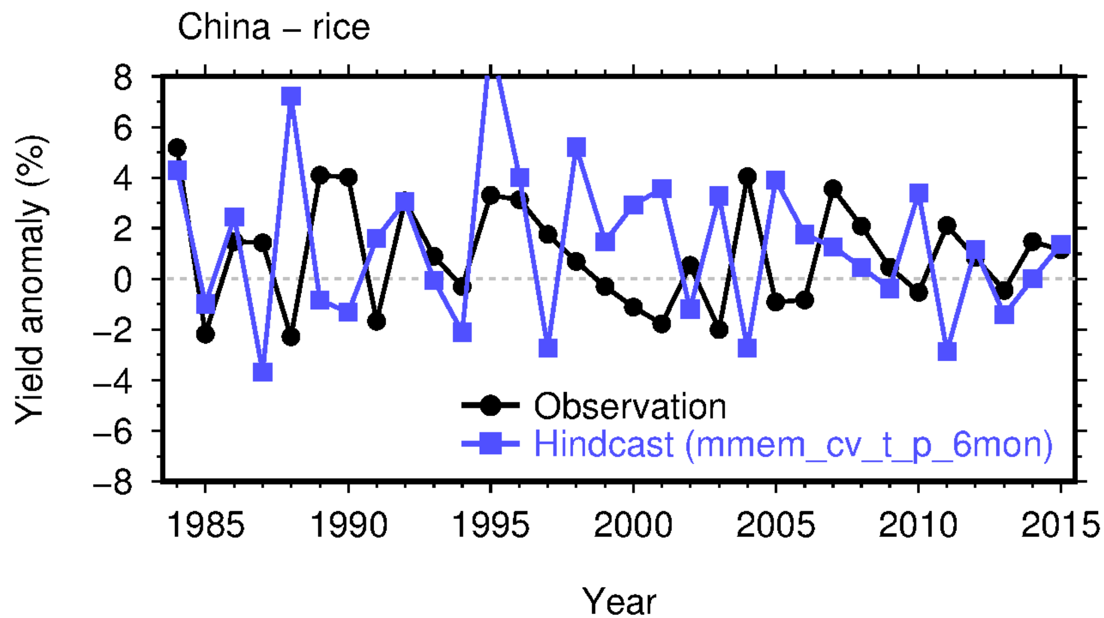


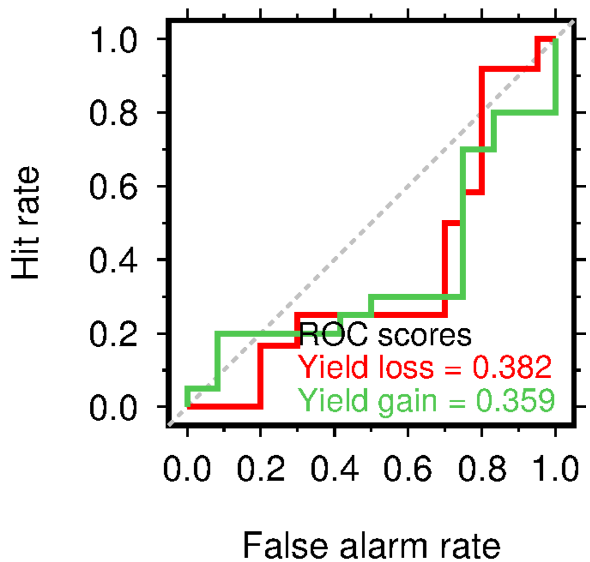
Note that this GCM selection is based on the independent data

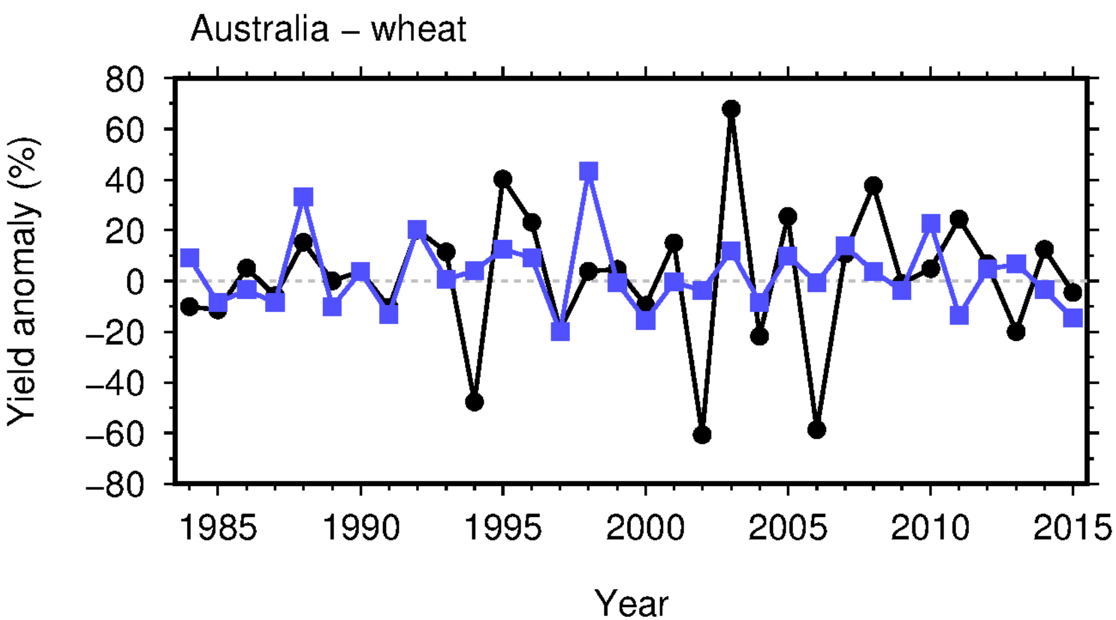
Country-level predictions (3-mon lead), maize & soy

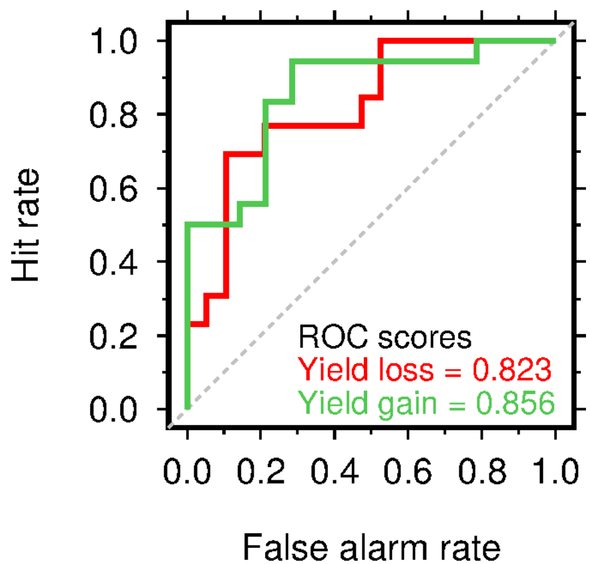


Country-level predictions (6-mon lead), rice & wheat









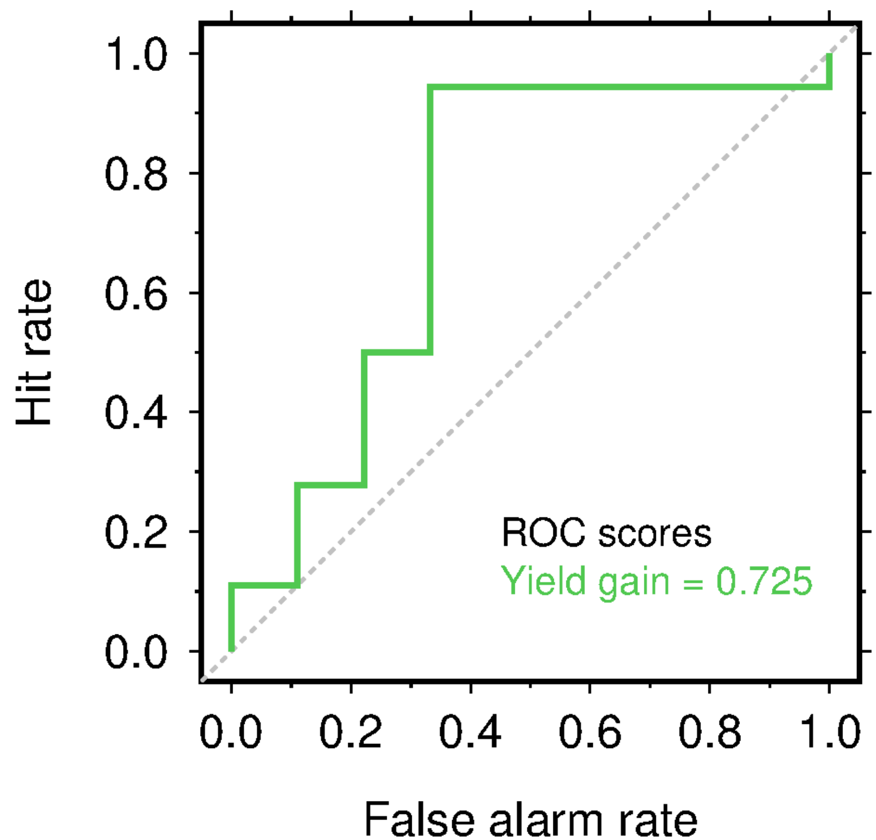
Settings of statistical yield models

	Iizumi et al (2013)	Iizumi et al. (2018)
Period	1983-2006 (24yr)	Grid, 1984-2010 (27yr) Country, 1984-2015 (32yr)
Yield anomaly (normalization)	First difference (average yield t-3:t-1)	Same as I13
Climatic variables	T & S	T & P
Crop calendar	SAGE (Sacks et al., 2010)	Same as I13
Calibration	MCMC	Same as I13
Skill score	R^2	ROC
Yield dataset	Global dataset of historical yields (GDHY) version 1	GDHY version 1.1
Climate model(s)	SINTEX-F1 (average over 9 ensemble members)	5 GCMs & 2 MME methods
Bias correction	Yes (CDFDM)	Same as I13

ROC (Receiver Operatorating Characteristic) score for yield gains

Year	Observed yield anomaly (%)	Hindcasted yield anomaly (%)
1984	25.144	2.52
1985	11.087	4.187
1986	1.338	5.048
1987	0.41	1.201
1988	-29.589	-1.987
1989	29.325	4.822
1990	2.09	5.527
1991	-9.29	0.6
1992	19.979	7.427
1993	-25.745	3.989
1994	33.36	1.1
1995	-20.324	0.607
1996	11.579	6.867
1997	-0.321	0.616
1998	6.325	2.619
1999	-0.497	5.357
2000	2.338	2.386
2001	0.971	3.677
2002	-6.493	-1.845
2003	9.534	5.355
2004	13.288	7.364
2005	-8.615	-0.05
2006	0.768	1.955
2007	1.054	3.816
2008	2.148	6.805
2009	7.143	-2.354
2010	-7.591	7.363

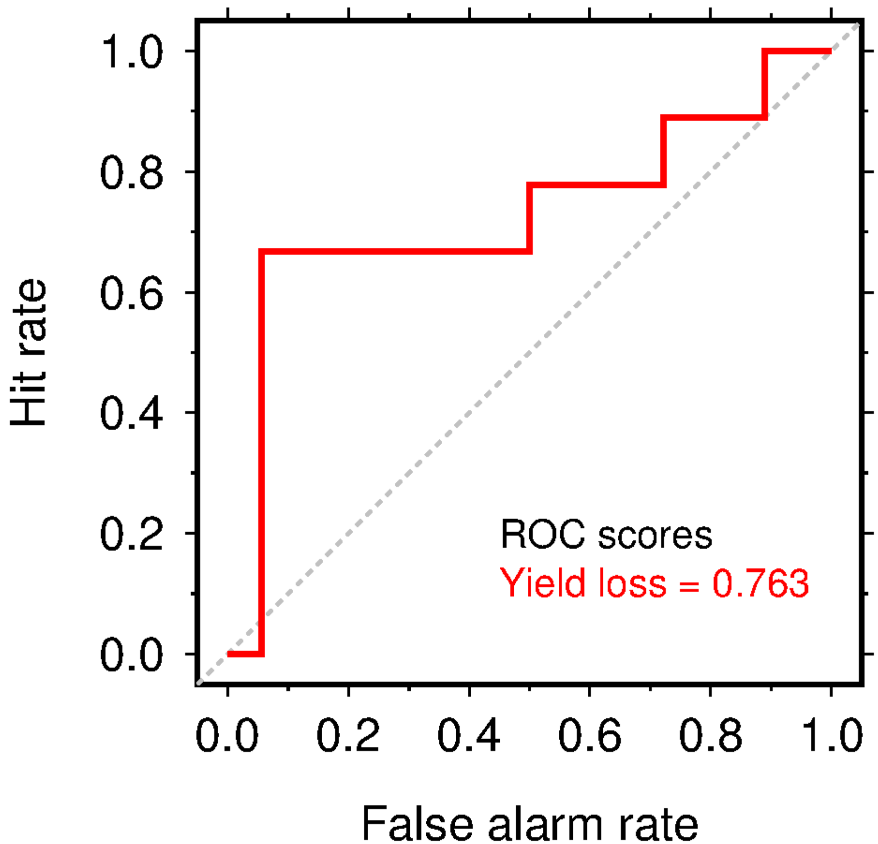
N=27	False N=9	True N=18
Negative (<1.100) N=7 陰性	True negative 0.667 (6/9)	False negative 0.056 (1/18)
Positive (≥1.100) N=20 陽性	False positive 0.333 (3/9)	True positive 0.944 (17/18)



ROC score for yield losses

Year	Observed yield anomaly (%)	Hindcasted yield anomaly (%)
1984	25.144	2.52
1985	11.087	4.187
1986	1.338	5.048
1987	0.41	1.201
1988	-29.589	-1.987
1989	29.325	4.822
1990	2.09	5.527
1991	-9.29	0.6
1992	19.979	7.427
1993	-25.745	3.989
1994	33.36	1.1
1995	-20.324	0.607
1996	11.579	6.867
1997	-0.321	0.616
1998	6.325	2.619
1999	-0.497	5.357
2000	2.338	2.386
2001	0.971	3.677
2002	-6.493	-1.845
2003	9.534	5.355
2004	13.288	7.364
2005	-8.615	-0.05
2006	0.768	1.955
2007	1.054	3.816
2008	2.148	6.805
2009	7.143	-2.354
2010	-7.591	7.363

N=27	False N=18	True N=9
Negative (>0.616) N=20 陰性	True negative 0.944 (17/18)	False negative 0.333 (3/9)
Positive (≤0.616) N=7 陽性	False positive 0.056 (1/18)	True positive 0.667 (6/9)



Improvements to & updates of GDHY datasets

	GDHY1.0	GDHY1.1	GDHY1.2	GDHY1.3
Reference	Iizumi et al. (2014)	Iizumi & Ramankutty (2016)	Iizumi et al. (2018)	In preparation
Period	1982–2006	1981–2011		2000-2016
Resolution	1.125°		0.5° (0.083/1/2)	0.5°
Crops	Maize (major/secondary), soybean, rice (major/secondary), wheat (winter/spring)			
Yield statistics	FAO national yield statistics	Same as the version 1.0, but errors in earlier version were fixed (e.g., Democratic Republic of the Congo)		
Satellite products	2nd generation GIMMS 0.073° bi-monthly NDVI data. The NDVI data were aggregated to 1.125° using harvested area maps and then used to estimate LAI and FPAR at 1.125° resolution. LAI and FPAR were used to derive crop-specific NPP.	3rd generation GIMMS 0.083° bi-monthly LAI and FPAR data. Crop-specific NPP at 0.083° resolution was estimated from LAI and FPAR.		MOD15A2H LAI and FPAR data (1-km 8-day composite data were processed to be 0.083° and daily resolution data)
Radiation	JRA-25 reanalysis			JRA-55 reanalysis
Harv. area	M3-Crops (Monfreda et al., 2008)			
Calendar	SAGE (Sacks et al., 2010)			
Production share by season	USDA (1994)			

Advantages and limitations of GDHY datasets

Advantages

- Yields of a crop for different growing seasons are available.
- Winter and spring wheat are explicitly separated.
- The spatial representativeness of grid-cell yields is more consistent across grid cells locate within an administrative unit.
- Relatively frequently updated (not regularly, but every 2 years)

Limitations

- No separation is available between irrigated and rainfed conditions.
- GDHY datasets offer estimates of grid-cell yield, but not reported (or observed) yields.
- GDHY datasets are largely depend on satellite products, and thus grid-cell yield estimates in minor-cropping areas is less reliable than those in major-growing areas.

Note for users

- A recommended practice is to use subnational (or national) yield statistics in addition to GDHY datasets.
- Analyses for a large spatial domain (continental to global) are suitable for the application of GDHY datasets.
- Keep in mind that stating your conclusions in a qualitative manner rather than in a quantitative manner to be more robust against the uncertainties associated with use of different datasets.

- Good practices are seen:
  - Iizumi & Ramankutty, 2016, Environmental Research Letters, doi:10.1088/1748-9326/11/3/034003
  - Challinor et al., 2016, Nature Climate Change, doi:10.1038/nclimate3061
  - Schauberger et al., 2017, Global Change Biology, https://doi.org/10.1111/gcb.13738

Yield prediction research at NARO

2013

Found the predictability of seasonal climate-induced yield variability for 20% of global harvested area (with JAMSTEC)

2014

Provide global maps of ENSO-yield relationship (with JAMSTEC)

The 1st global crop forecast information

2015

Developed a prototype of crop forecast system at NARO

2016

Detected changes in yield variability associated with climate change

2017

Challenge toward more operational crop forecasting (with APCC)

nature climate change
LETTERS
PUBLISHED ONLINE: 21 JULY 2013 | DOI:10.1038/NCLIMATE1945

Prediction of seasonal climate-induced variations in global food production

Toshichika Iizumi[1]*, Hirofumi Sakuma[2,3], Masayuki Yokozawa[1], Jing-Jia Luo[4], Andrew J. Challinor[5,6], Molly E. Brown[7], Gen Sakurai[1] and Toshio Yamagata[3]

nature COMMUNICATIONS
ARTICLE
Received 10 Oct 2013 | Accepted 24 Mar 2014 | Published 15 May 2014
DOI: 10.1038/ncomms4712

Impacts of El Niño Southern Oscillation on the global yields of major crops

Toshichika Iizumi[1], Jing-Jia Luo[2], Andrew J. Challinor[3,4], Gen Sakurai[1], Masayuki Yokozawa[5], Hirofumi Sakuma[6,7], Molly E. Brown[8] & Toshio Yamagata[7]

Environmental Research Letters
LETTER

Changes in yield variability of major crops for 1981–2010 explained by climate change

Toshichika Iizumi[1,3] and Navin Ramankutty[2]