

# STATISTICAL MODELS FOR QUANTITATIVE SYNTHESIS OF CLIMATE CHANGE IMPACT STUDIES

David Makowski

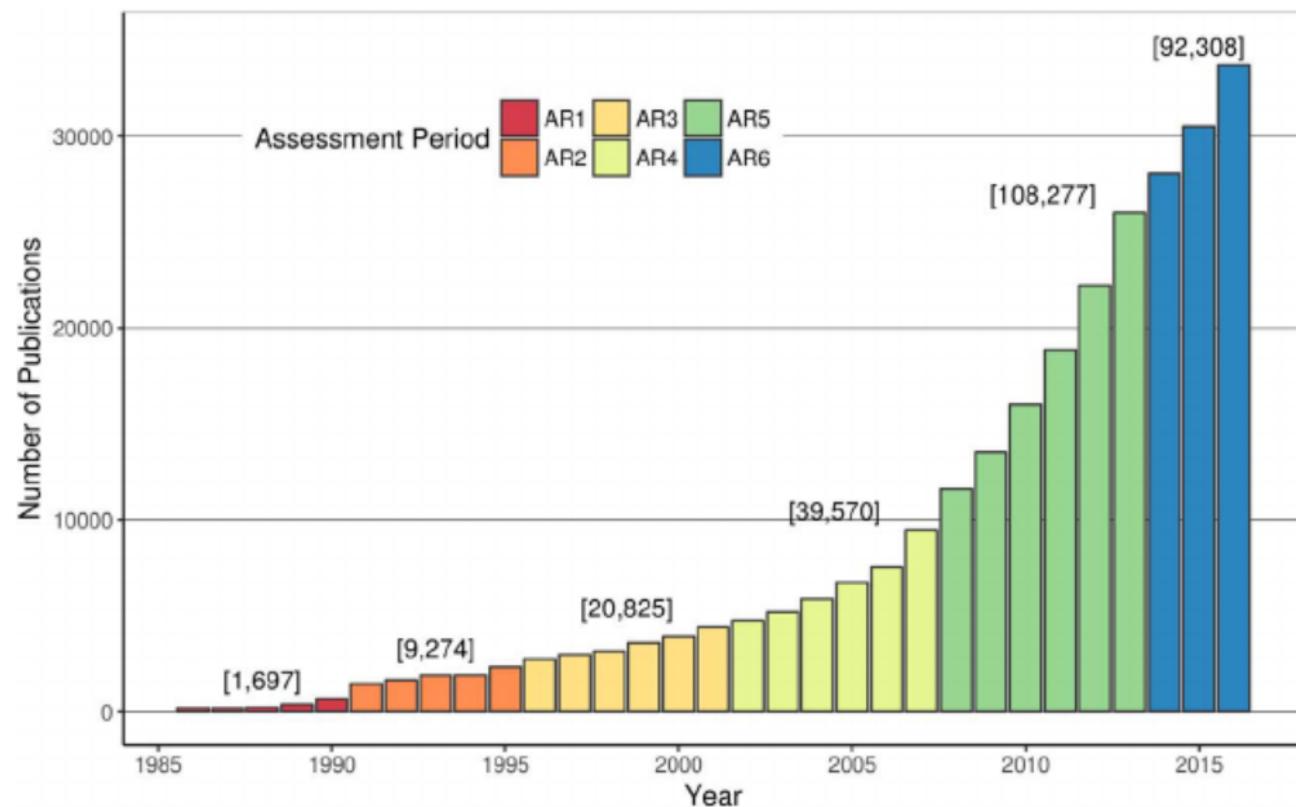
INRA, France

[david.makowski@inra.fr](mailto:david.makowski@inra.fr)

## The ‘data synthesis challenge’

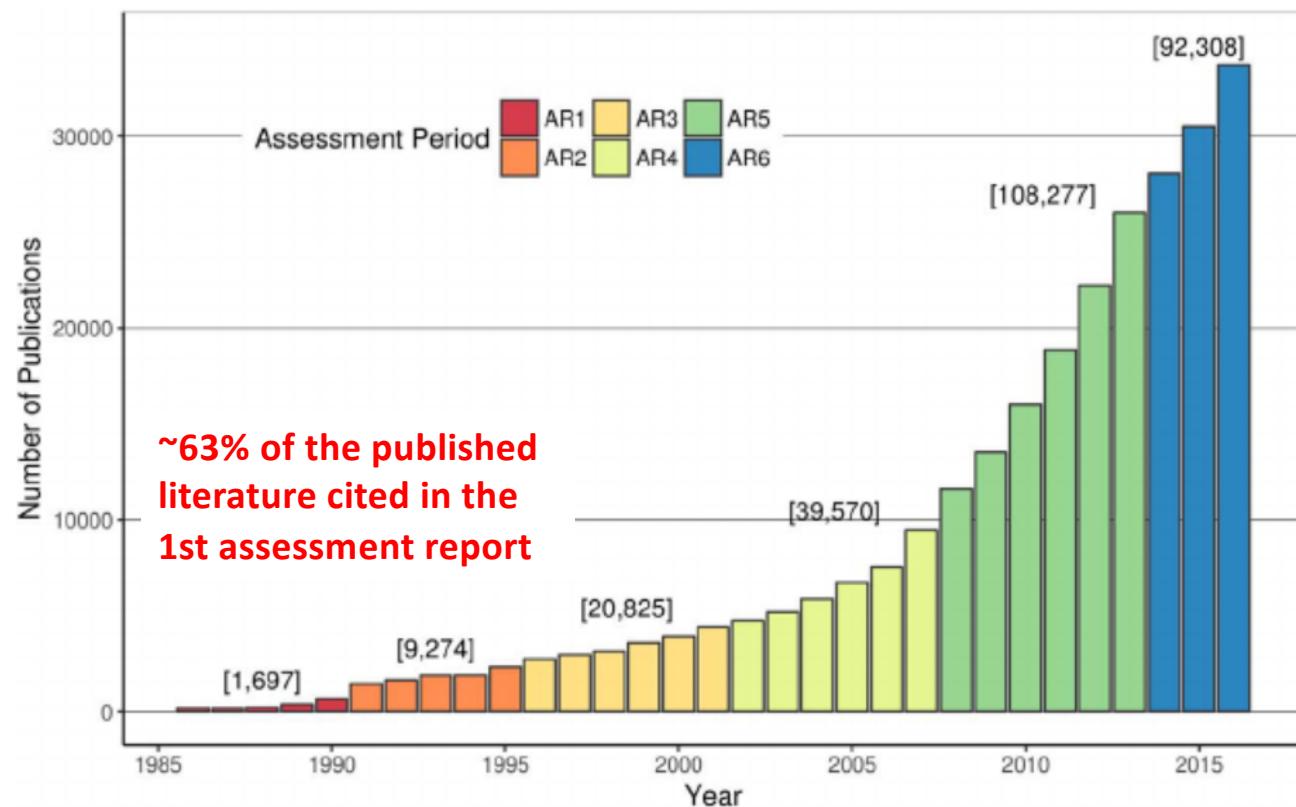
As more and more data become available, how to conduct rigorous and comprehensive assessments on climate change?

The growth of the literature on climate change was much faster than the growth in other areas of research (16% vs. 4%)



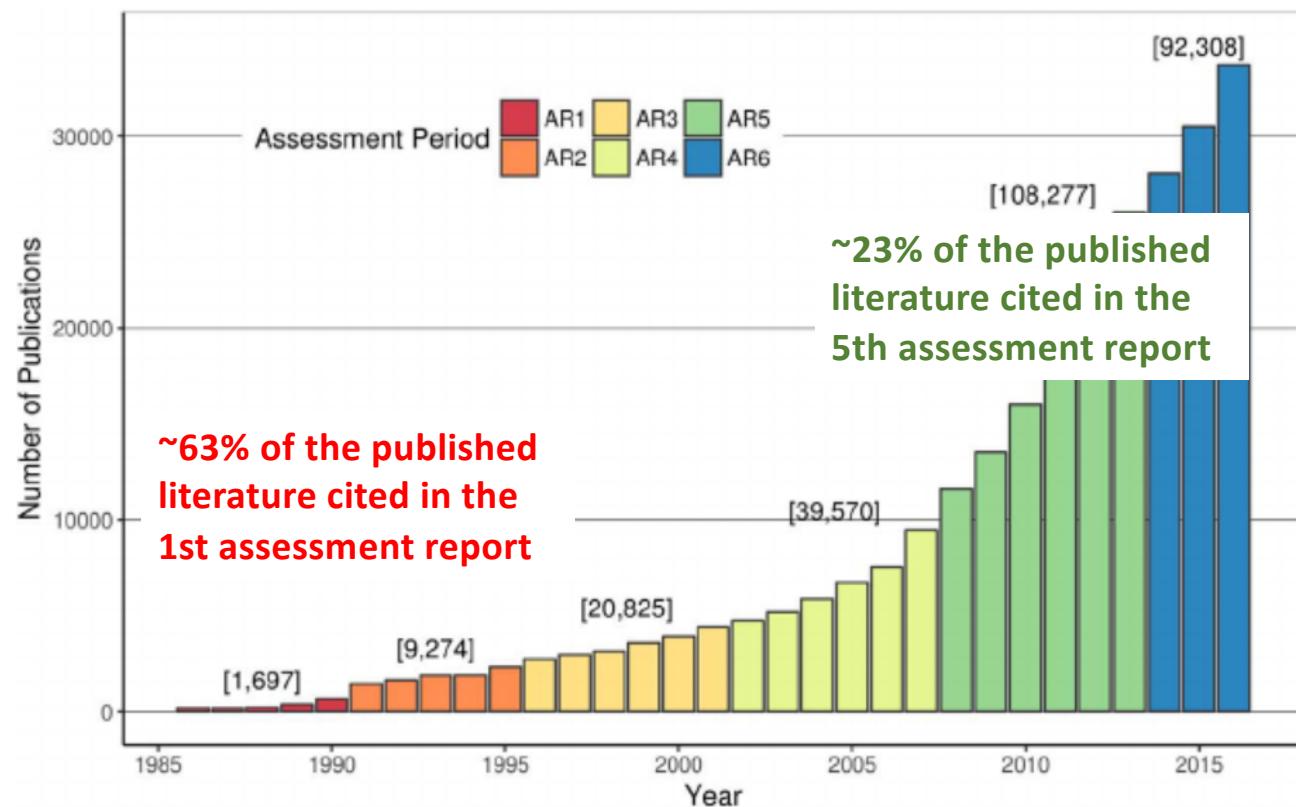
Minx J.C., Callaghan M., Lamb W.F., Garard J., Edenhofer O. 2017. Learning about climate change solutions in the IPCC and beyond. Environmental Science and Policy 77, 252-259

The growth of the literature on climate change was much faster than the growth in other areas of research (16% vs. 4%)



Minx J.C., Callaghan M., Lamb W.F., Garard J., Edenhofer O. 2017. Learning about climate change solutions in the IPCC and beyond. Environmental Science and Policy 77, 252-259

The growth of the literature on climate change was much faster than the growth in other areas of research (16% vs. 4%)



Minx J.C., Callaghan M., Lamb W.F., Garard J., Edenhofer O. 2017. Learning about climate change solutions in the IPCC and beyond. Environmental Science and Policy 77, 252-259

Formal methods are needed to help researchers to conduct rigorous and comprehensive literature synthesis

# Meta-analysis: a statistical approach for quantitative synthesis

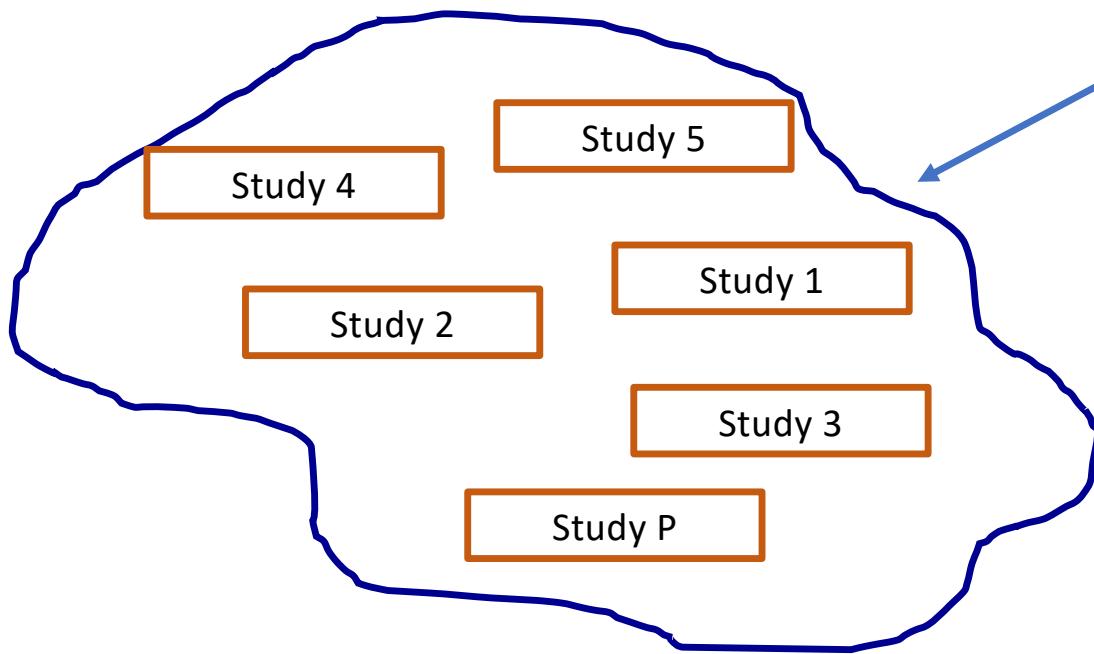
« The analysis of analyses »

« The statistical analysis of a large collection of results from individual studies for the purpose of integrating the findings »

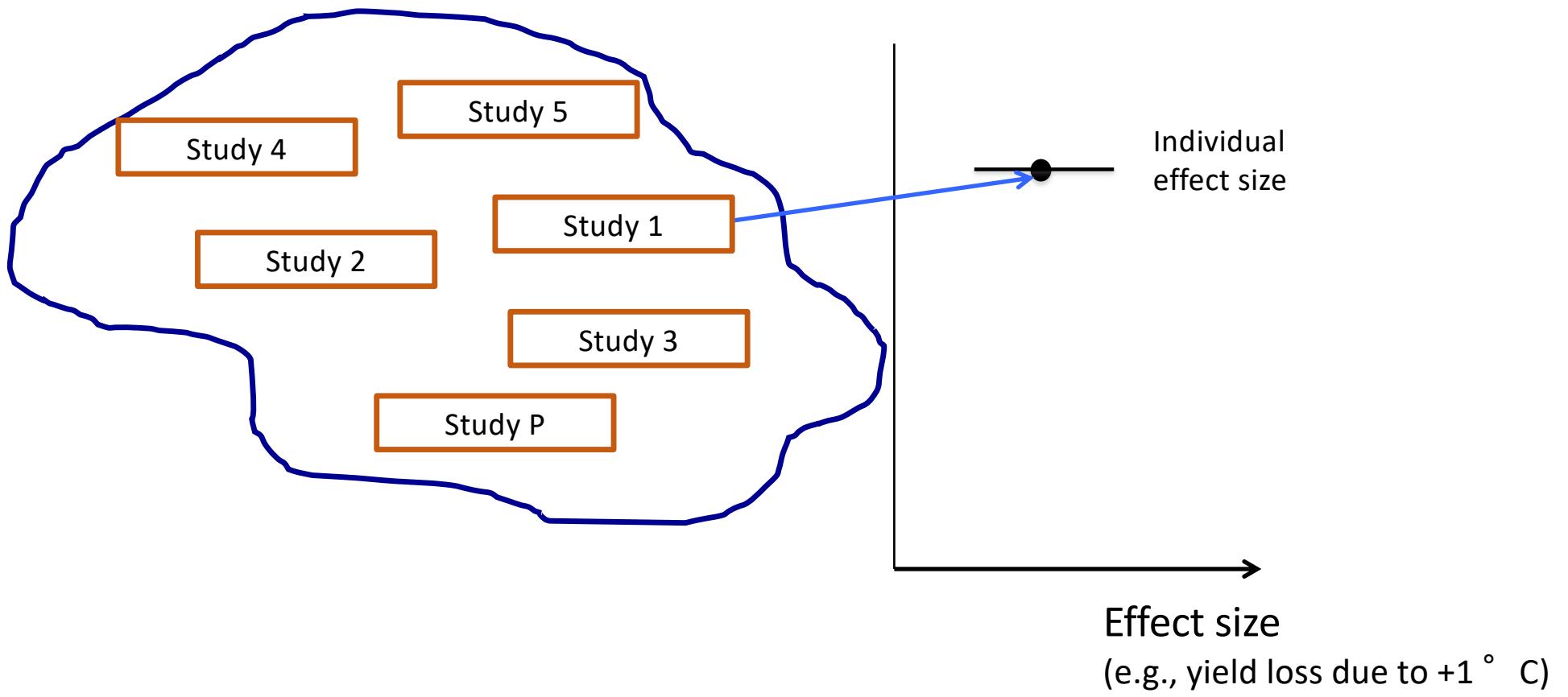
« Systematic review + statistical analysis »

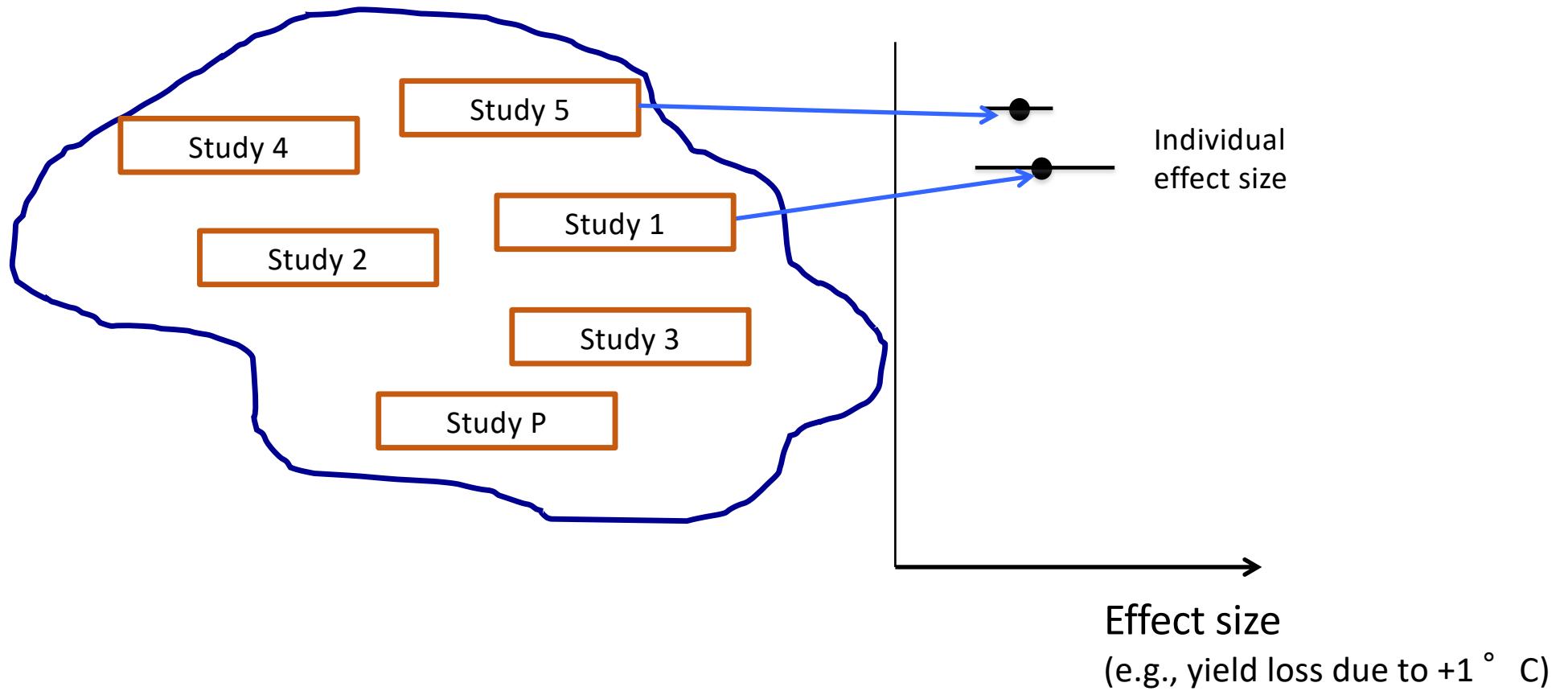
Dictionary of epidemiology, 2001; Chalmers et al., 2002; Glass, 1976; Koricheva et al., 2013

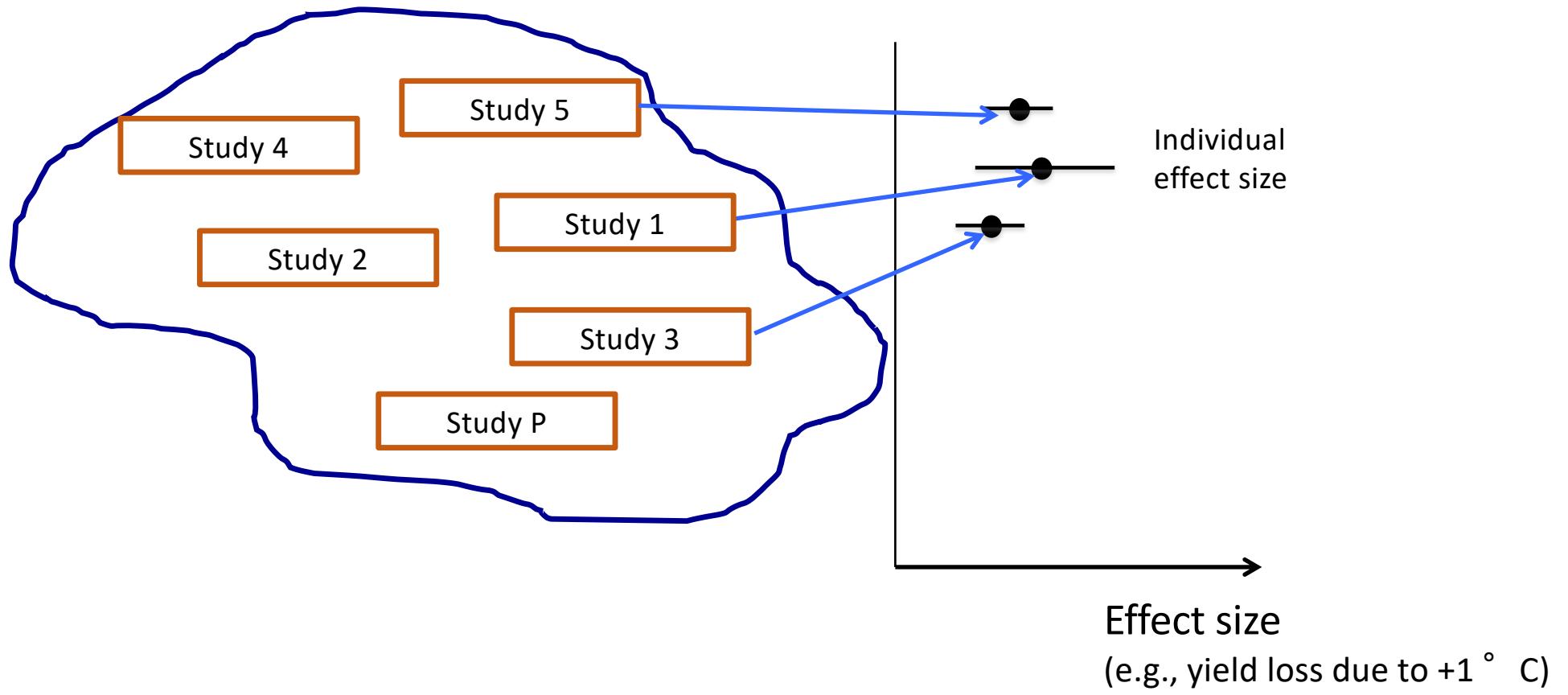
## Systematic review of studies

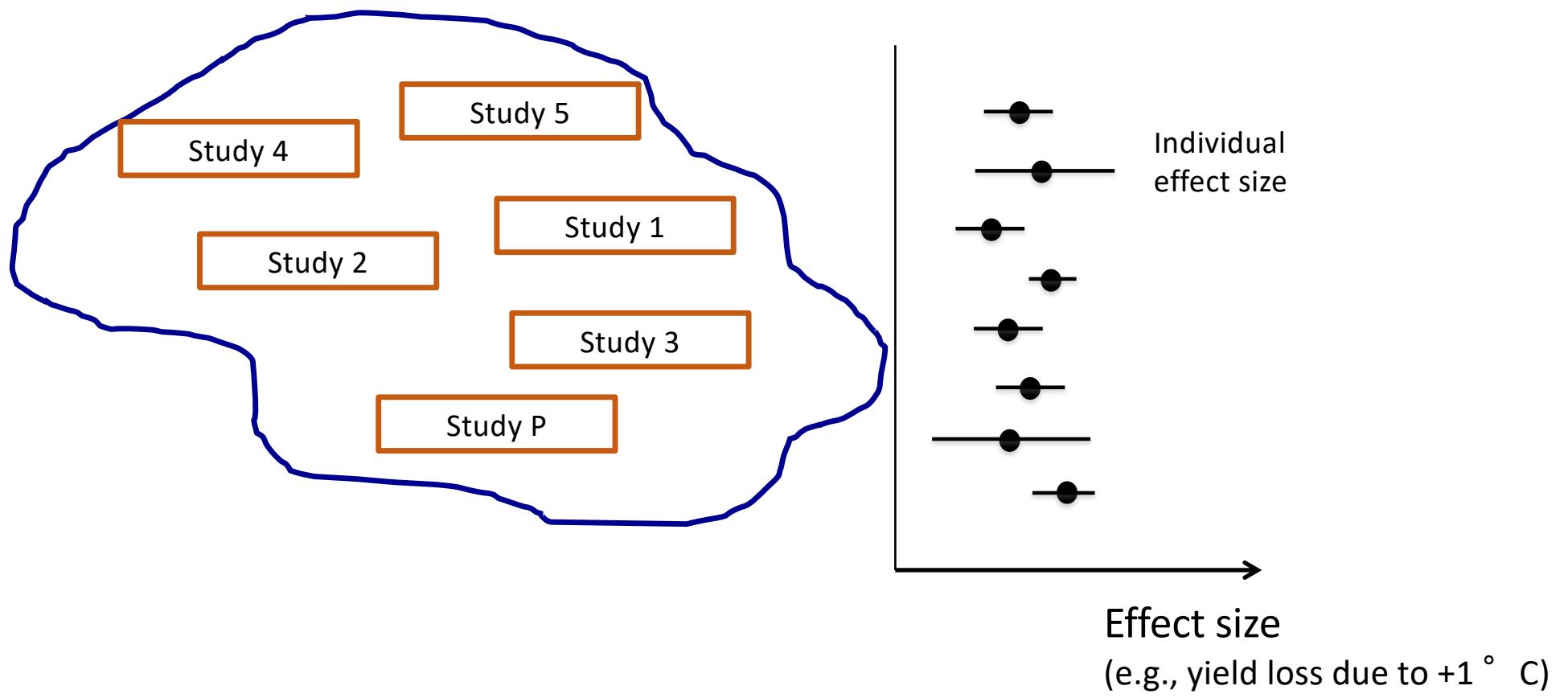


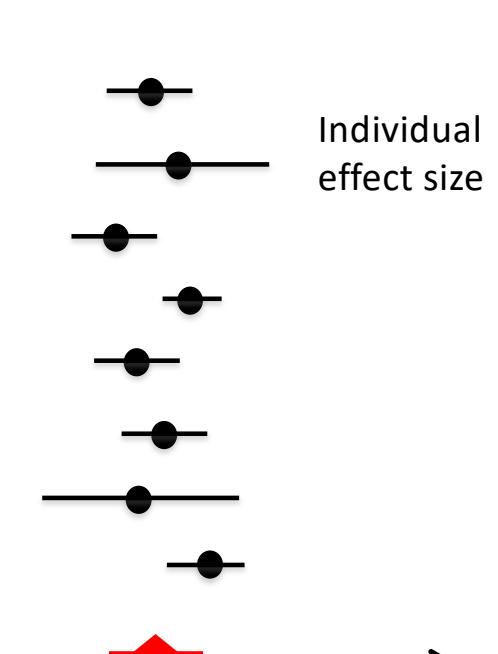
**Set of studies dealing with a specific topic  
(e.g., %yield loss due to +1°C)**



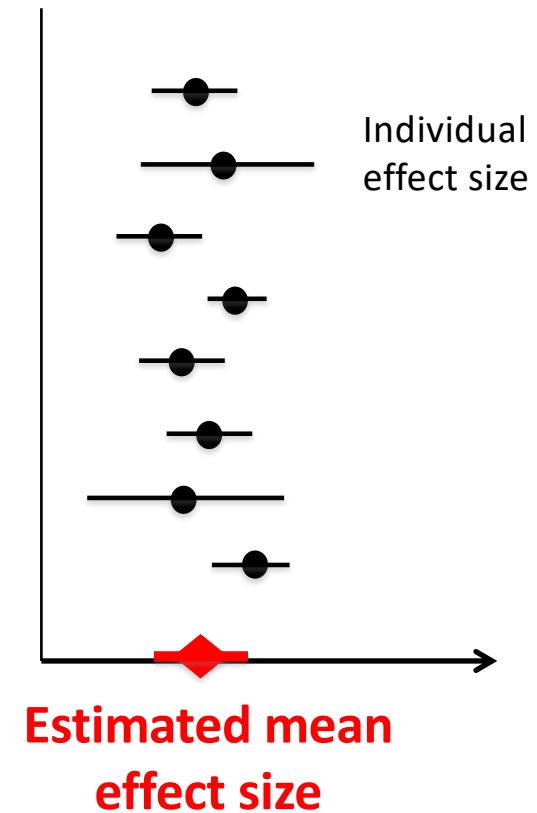
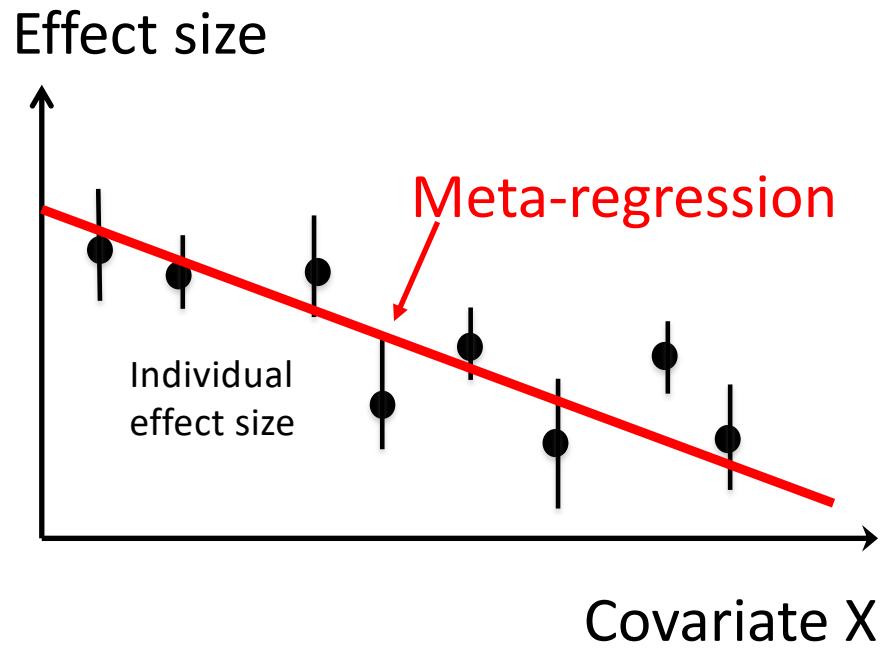








**Estimated mean  
effect size**



Example: Assessment of the impact of temperature increase on crop yield

Two sources of information:

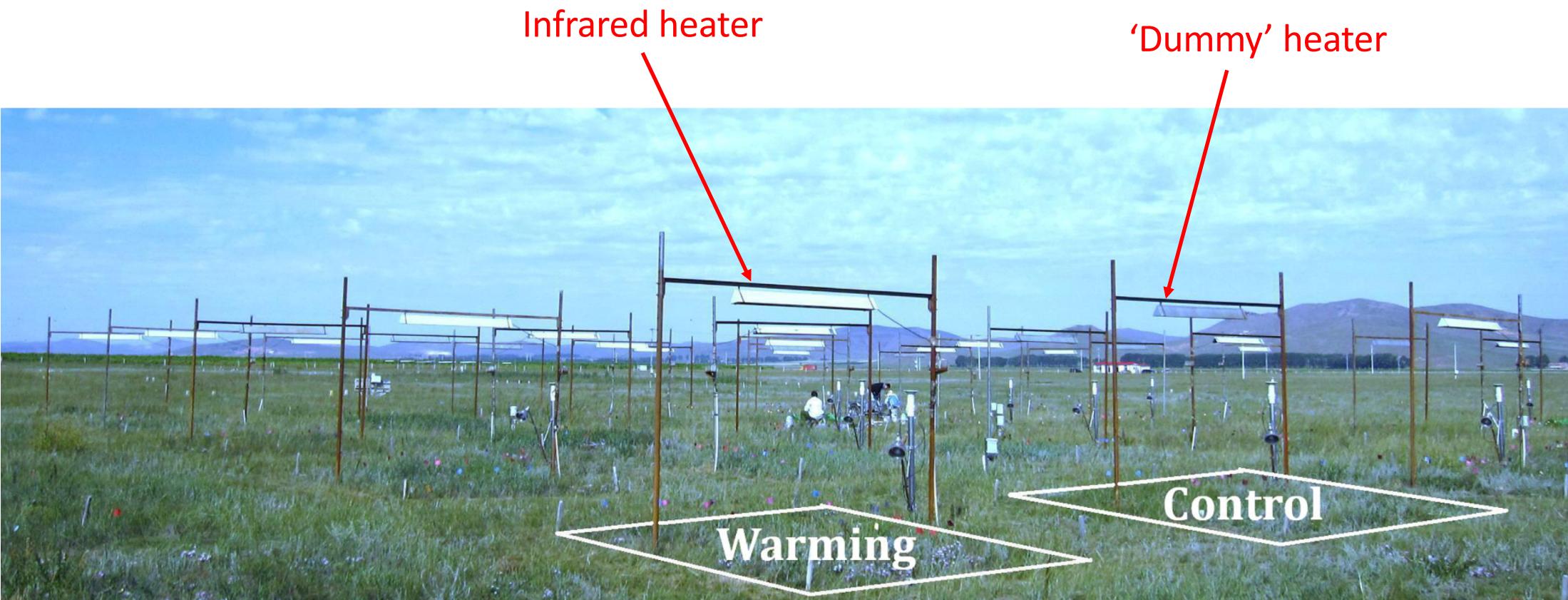
- Experiments
- Crop model simulations

Example: Assessment of the impact of temperature increase on crop yield

Two sources of information:

- **Experiments**
- Crop model simulations

# Field warming experiment



from Chi et al. 2013

[doi.org/10.1371/journal.pone.0056482](https://doi.org/10.1371/journal.pone.0056482)

$$\Delta Y = (Yield_{warm} - Yield_{control})/Yield_{control}$$

$$Sensitivity = Yield \% change per {}^\circ\text{C} = 100 \frac{\Delta Y}{\Delta T}$$

Article | **OPEN** | Published: 17 November 2016

# Field warming experiments shed light on the wheat yield response to temperature in China

Chuang Zhao, Shilong Piao , Yao Huang, Xuhui Wang, Philippe Ciais, Mengtian Huang, Zhenzhong Zeng & Shushi Peng



Compilation of 46 results of field warming experiments located in 11 sites in China

Field experiment  
Ambient CO<sub>2</sub>

# Objectives

- Estimate the mean yield sensitivity to +1°C
- Estimate the effect of the mean local temperature on the sensitivities
- Analyse uncertainties

Site_name	Lat	Long	Design	Temp_increase	Irrigation_mm	Nitrogen_kgha	TGS	PGS	DGS	Sensitivity	Ref
Nanjing	32.03	118.86	Infrared heaters	1.5	0	225	10.3	393	8	11.1	1
Nanjing	32.03	118.86	Infrared heaters	1.5	0	225	10.6	476	7.6	10.5	1
Nanjing	32.03	118.86	Infrared heaters	1.5	0	225	11.6	341	8.3	14.5	1
Nanjing	32.03	118.86	Infrared heaters	1.5	0	225	10.0	387	7.9	6.7	1
Nanjing	32.03	118.86	Infrared heaters	1.5	0	225	10.5	325	8.1	11.8	1
Tongwei County	35.22	105.23	Heating cable	1.4	0	NA	5.7	288	6.6	2.21	2
Tongwei County	35.22	105.23	Heating cable	2.2	0	NA	5.7	288	6.6	1.18	2
Lulu Mountain	35.29	105.26	Heating cable	0.6	0	NA	5.7	288	6.6	0.83	2
Lulu Mountain	35.29	105.26	Heating cable	1.4	0	NA	5.7	288	6.6	2.57	2
Lulu Mountain	35.29	105.26	Heating cable	2.2	0	NA	5.7	288	6.6	2.73	2
Guyuan	36.03	106.46	Greenhouse	0.5	0	NA	12.4	97	11.2	-9.33	3
Guyuan	36.03	106.46	Greenhouse	1.2	0	NA	12.4	97	11.2	-11.03	3
Guyuan	36.03	106.46	Greenhouse	2.0	0	NA	12.4	97	11.2	-7.4	3
Guyuan	36.03	106.46	Greenhouse	0.5	60	NA	12.4	97	11.2	-11.84	3
Guyuan	36.03	106.46	Greenhouse	1.2	60	NA	12.4	97	11.2	-10.65	3
Guyuan	36.03	106.46	Greenhouse	2.0	60	NA	12.4	97	11.2	-6.78	3
Dingxing	39.13	115.66	Infrared heaters	2.9	500	210	8.9	77	9.3	-3.1	4
Dingxing	39.13	115.66	Infrared heaters	2.9	500	210	6.4	113	8.2	8.25	4
Xidatan	38.80	106.30	Infrared heaters	0.5	NA	136	14.3	23	12.2	-1.0	5
Xidatan	38.80	106.30	Infrared heaters	1.0	NA	136	14.3	23	12.2	-0.8	5
Xidatan	38.80	106.30	Infrared heaters	1.5	NA	136	14.3	23	12.2	-3.33	5
Xidatan	38.80	106.30	Infrared heaters	2.0	NA	136	14.3	23	12.2	-8.25	5
Xidatan	38.80	106.30	Infrared heaters	2.5	NA	136	14.3	23	12.2	-7.4	5
Shanghai	31.21	121.13	Infrared heaters	1.5	NA	NA	11.0	433	6.1	3.46	6
Shanghai	31.21	121.13	Infrared heaters	1.5	NA	NA	11.4	343	6.6	3.33	6
Shanghai	31.21	121.13	Infrared heaters	1.5	NA	NA	10.2	476	5.7	3.45	6
Yucheng	36.83	116.57	Infrared heaters	1.3	150	285	8.6	118	10	1.23	7
Yucheng	36.83	116.57	Infrared heaters	1.3	150	285	8.6	118	10	-2.54	7
Yucheng	36.83	116.57	Infrared heaters	1.3	150	285	8.8	160	8.8	-1.15	7
Yucheng	36.83	116.57	Infrared heaters	1.3	150	285	8.8	160	8.8	-4.69	7
Lianyungang	34.55	119.39	Infrared heaters	2.2	0	0	7.9	372	7.3	12.92	8
Lianyungang	34.55	119.39	Infrared heaters	2.2	0	150	7.9	372	7.3	6.31	8
Lianyungang	34.55	119.39	Infrared heaters	2.2	0	225	7.9	372	7.3	2.95	8
Lianyungang	34.55	119.39	Infrared heaters	2.2	0	300	7.9	372	7.3	5.82	8
Luancheng	37.88	114.68	Infrared heaters	2.0	160	240	9.6	131	9.5	-13.0	9
Luancheng	37.88	114.68	Infrared heaters	2.0	160	240	8.9	65	9.8	-13.5	9
Luancheng	37.88	114.68	Infrared heaters	2.0	160	240	8.8	153	8.6	-6.0	9
Luancheng	37.88	114.68	Infrared heaters	2.0	160	0	9.6	131	9.5	-12.5	9
Luancheng	37.88	114.68	Infrared heaters	2.0	160	0	8.9	65	9.8	-3.5	9
Luancheng	37.88	114.68	Infrared heaters	2.0	160	0	8.8	153	8.6	-4.5	9
Luancheng	37.88	114.68	Infrared heaters	2.0	160	240	7.3	167	8.5	15.5	9
Dingxi	35.58	104.62	Infrared heaters	1.0	0	NA	13.6	171	8.6	0.5	9
Dingxi	35.58	104.62	Infrared heaters	2.0	0	NA	13.6	171	8.6	-7.55	10
Dingxi	35.58	104.62	Infrared heaters	3.0	0	NA	13.6	171	8.6	-7.27	10

Site_name	Lat	Long	Design	Temp_increase	Irrigation_mm	Nitrogen_kgha	TGS	PGS	DGS	Sensitivity	Ref
Nanjing	32.03	118.86	Infrared heaters	1.5	0	225	10.3	393	8	11.1	1
Nanjing	32.03	118.86	Infrared heaters	1.5	0	225	10.0	387	7.9	0.1	1
Nanjing	32.03	118.86	Infrared heaters	1.5	0	225	10.5	325	8.1	11.8	1

## Study 1

Site_name	Lat	Long	Design	Temp_increase	Irrigation_mm	Nitrogen_kgha	TGS	PGS	DGS	Sensitivity	Ref
Nanjing	32.03	118.86	Infrared heaters	1.5	0	225	10.0	387	7.9	0.1	1
Nanjing	32.03	118.86	Infrared heaters	1.5	0	225	10.5	325	8.1	11.8	1

## Study 2

Site_name	Lat	Long	Design	Temp_increase	Irrigation_mm	Nitrogen_kgha	TGS	PGS	DGS	Sensitivity	Ref
Lulu Mountain	35.29	105.26	Heating cable	1.4	0	NA	5.7	288	6.6	2.57	2
Lulu Mountain	35.29	105.26	Heating cable	2.2	0	NA	5.7	288	6.6	2.73	2
Guyuan	36.03	106.46	Greenhouse	0.5	0	NA	12.4	97	11.2	-9.33	3
Guyuan	36.03	106.46	Greenhouse	1.2	0	NA	12.4	97	11.2	-11.03	3
Guyuan	36.03	106.46	Greenhouse	2.0	0	NA	12.4	97	11.2	-7.4	3
Guyuan	36.03	106.46	Greenhouse	0.5	60	NA	12.4	97	11.2	-11.84	3
Guyuan	36.03	106.46	Greenhouse	1.2	60	NA	12.4	97	11.2	-10.65	3
...	...	...	...	...	...	...	...	...	...	...	3
Xidatan	38.80	106.30	Infrared heaters	0.5	NA	136	14.3	23	12.2	-1.0	5
Xidatan	38.80	106.30	Infrared heaters	1.0	NA	136	14.3	23	12.2	-0.8	5
Xidatan	38.80	106.30	Infrared heaters	1.5	NA	136	14.3	23	12.2	-3.33	5
Xidatan	38.80	106.30	Infrared heaters	2.0	NA	136	14.3	23	12.2	-8.25	5
Xidatan	38.80	106.30	Infrared heaters	2.5	NA	136	14.3	23	12.2	-7.4	5
Shanghai	31.21	121.13	Infrared heaters	1.5	NA	NA	11.0	433	6.1	3.46	6
Shanghai	31.21	121.13	Infrared heaters	1.5	NA	NA	11.4	343	6.6	3.33	6
Shanghai	31.21	121.13	Infrared heaters	1.5	NA	NA	10.2	476	5.7	3.45	6
Yucheng	36.83	116.57	Infrared heaters	1.3	150	285	8.6	118	10	1.23	7
Yucheng	36.83	116.57	Infrared heaters	1.3	150	285	8.6	118	10	-2.54	7
Yucheng	36.83	116.57	Infrared heaters	1.3	150	285	8.8	160	8.8	-1.15	7
Yucheng	36.83	116.57	Infrared heaters	1.3	150	285	8.8	160	8.8	-4.69	7
Lianyungang	34.55	119.39	Infrared heaters	2.2	0	0	7.9	372	7.3	12.92	8
Lianyungang	34.55	119.39	Infrared heaters	2.2	0	150	7.9	372	7.3	6.31	8
Lianyungang	34.55	119.39	Infrared heaters	2.2	0	225	7.9	372	7.3	2.95	8
Lianyungang	34.55	119.39	Infrared heaters	2.2	0	300	7.9	372	7.3	5.82	8
Luancheng	37.88	114.68	Infrared heaters	2.0	160	240	9.6	131	9.5	-13.0	9
Luancheng	37.88	114.68	Infrared heaters	2.0	160	240	8.9	65	9.8	-13.5	9
Luancheng	37.88	114.68	Infrared heaters	2.0	160	240	8.8	153	8.6	-6.0	9
Luancheng	37.88	114.68	Infrared heaters	2.0	160	0	9.6	131	9.5	-12.5	9
Luancheng	37.88	114.68	Infrared heaters	2.0	160	0	8.9	65	9.8	-3.5	9
Luancheng	37.88	114.68	Infrared heaters	2.0	160	0	8.8	153	8.6	-4.5	9
Luancheng	37.88	114.68	Infrared heaters	2.0	160	240	7.3	167	8.5	15.5	9
Dingxi	35.58	104.62	Infrared heaters	1.0	0	NA	13.6	171	8.6	0.5	9
Dingxi	35.58	104.62	Infrared heaters	2.0	0	NA	13.6	171	8.6	-7.55	10
Dingxi	35.58	104.62	Infrared heaters	3.0	0	NA	13.6	171	8.6	-7.27	10

...

# Steps

- Explore the dataset
- Define and fit Bayesian models
- Conclude

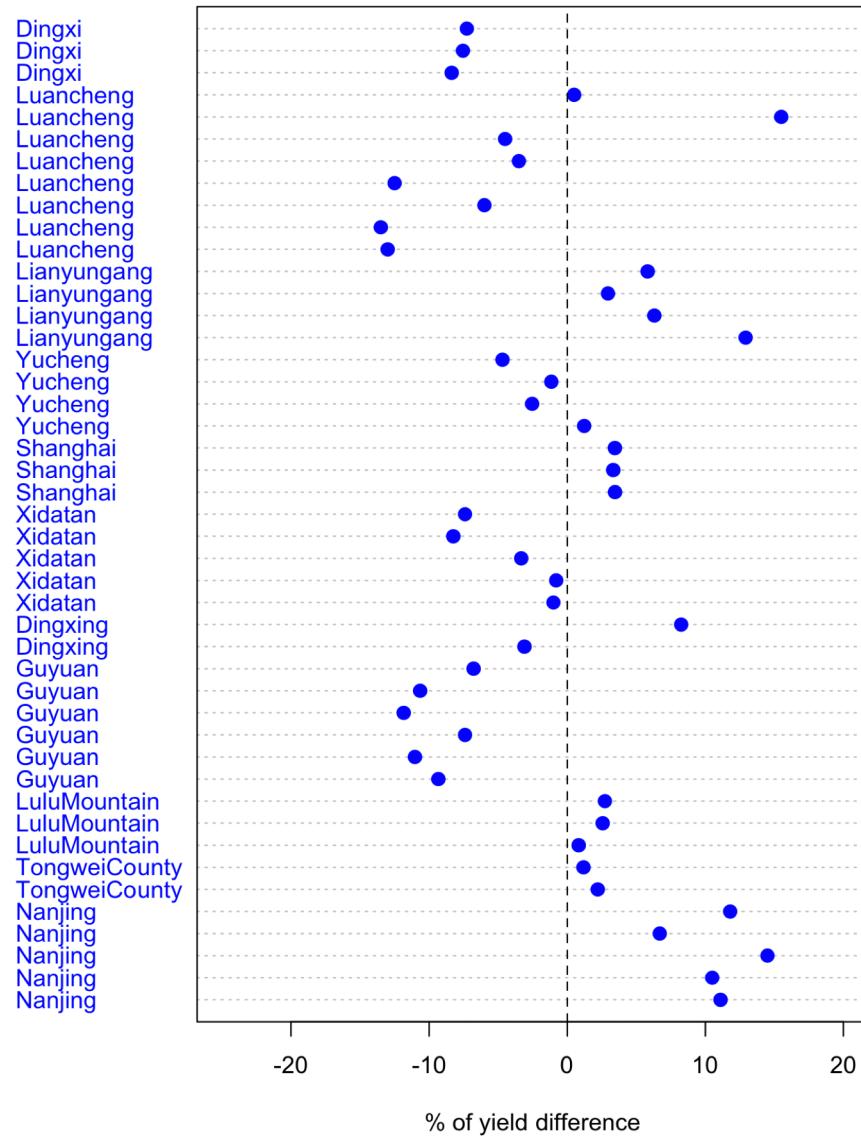
# Steps

- Explore the dataset
- Define and fit Bayesian models
- Conclude

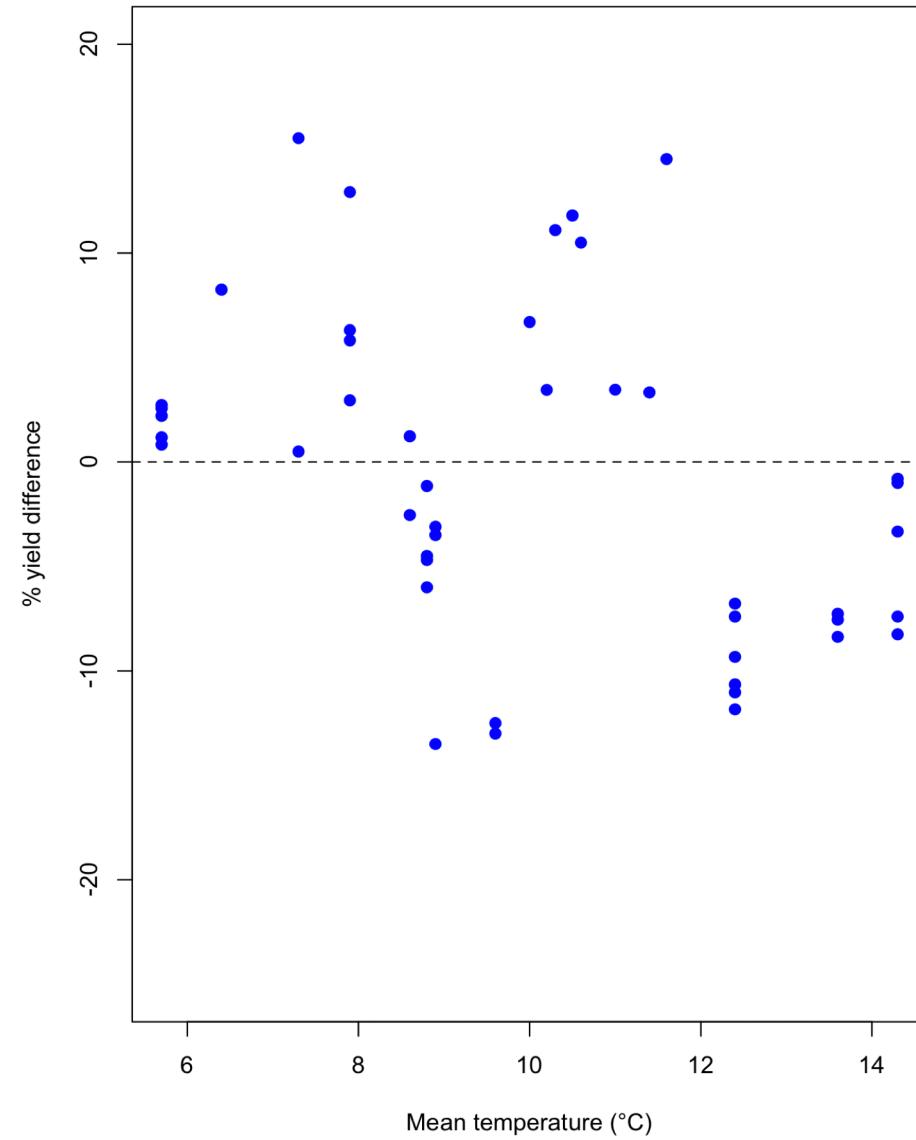
## Open and run Atelier2020\_1.R

Wheat_type	Site_name	Lat	Long	Design	Temp_increase	Irrigation_mm	Nitrogen_kgha	TGS	PGS
Springwheat:14	Luancheng : 8	Min. :31.21	Min. :104.6	Greenhouse : 6	Min. :0.5	Min. : 0.0	Min. : 0.0	Min. : 5.70	Min. : 23
Winterwheat:31	Guyuan : 6	1st Qu.:34.55	1st Qu.:106.3	Heatingcable : 5	1st Qu.:1.3	1st Qu.: 0.0	1st Qu.:136.0	1st Qu.: 7.90	1st Qu.: 97
	Nanjing : 5	Median :36.03	Median :114.7	Infraredheaters:34	Median :1.5	Median : 0.0	Median :225.0	Median : 9.60	Median :160
	Xidatan : 5	Mean :35.86	Mean :112.5		Mean :1.7	Mean : 82.7	Mean :178.6	Mean :10.04	Mean :202
	Lianyungang: 4	3rd Qu.:37.88	3rd Qu.:118.9		3rd Qu.:2.0	3rd Qu.:160.0	3rd Qu.:240.0	3rd Qu.:12.40	3rd Qu.:325
	Yucheng : 4	Max. :39.13	Max. :121.1		Max. :3.0	Max. :500.0	Max. :300.0	Max. :14.30	Max. :476
	(Other) :13					NA's :8	NA's :17		
DGS	Sensitivity	Ref							
Min. : 5.700	Min. :-13.5000	Min. : 1.000							
1st Qu.: 7.300	1st Qu.: -7.4000	1st Qu.: 3.000							
Median : 8.600	Median : -1.0000	Median : 5.000							
Mean : 8.942	Mean : -0.8364	Mean : 5.467							
3rd Qu.:10.000	3rd Qu.: 3.4500	3rd Qu.: 8.000							
Max. :12.200	Max. : 15.5000	Max. :10.000							

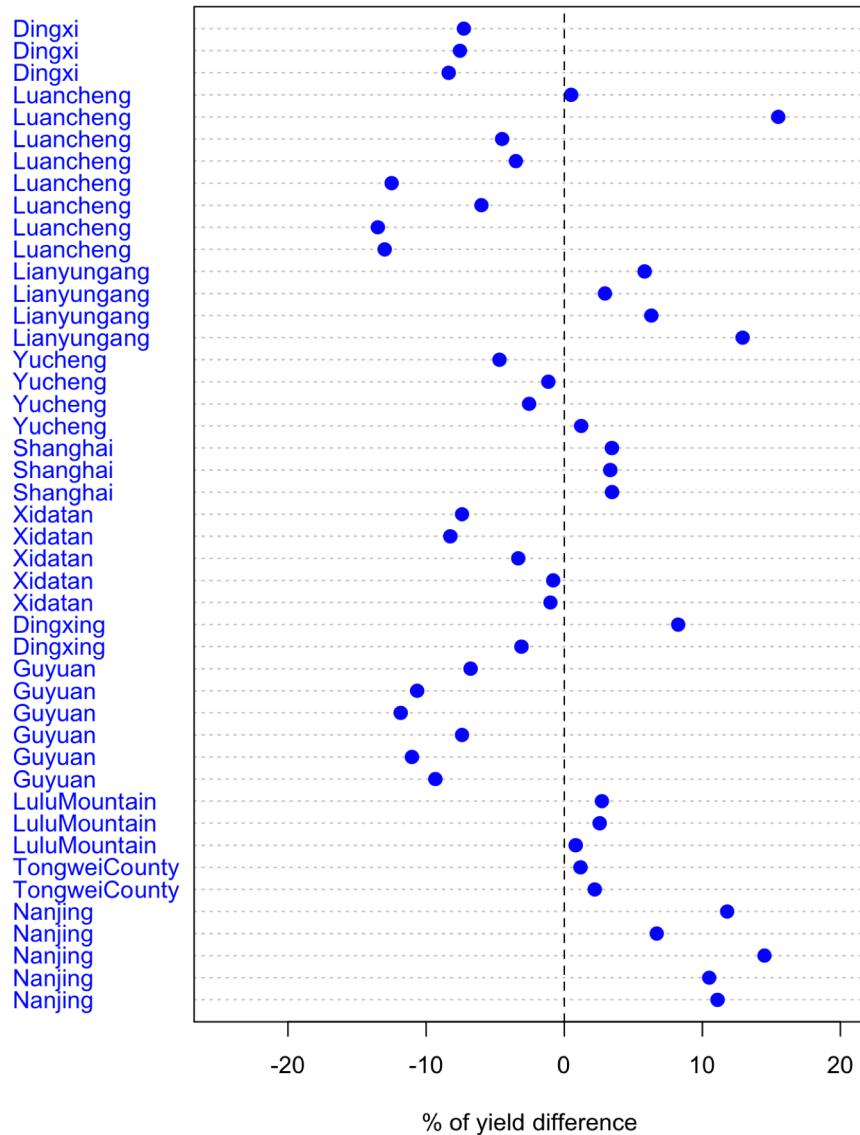
**Individual yield sensitivity values, by site**



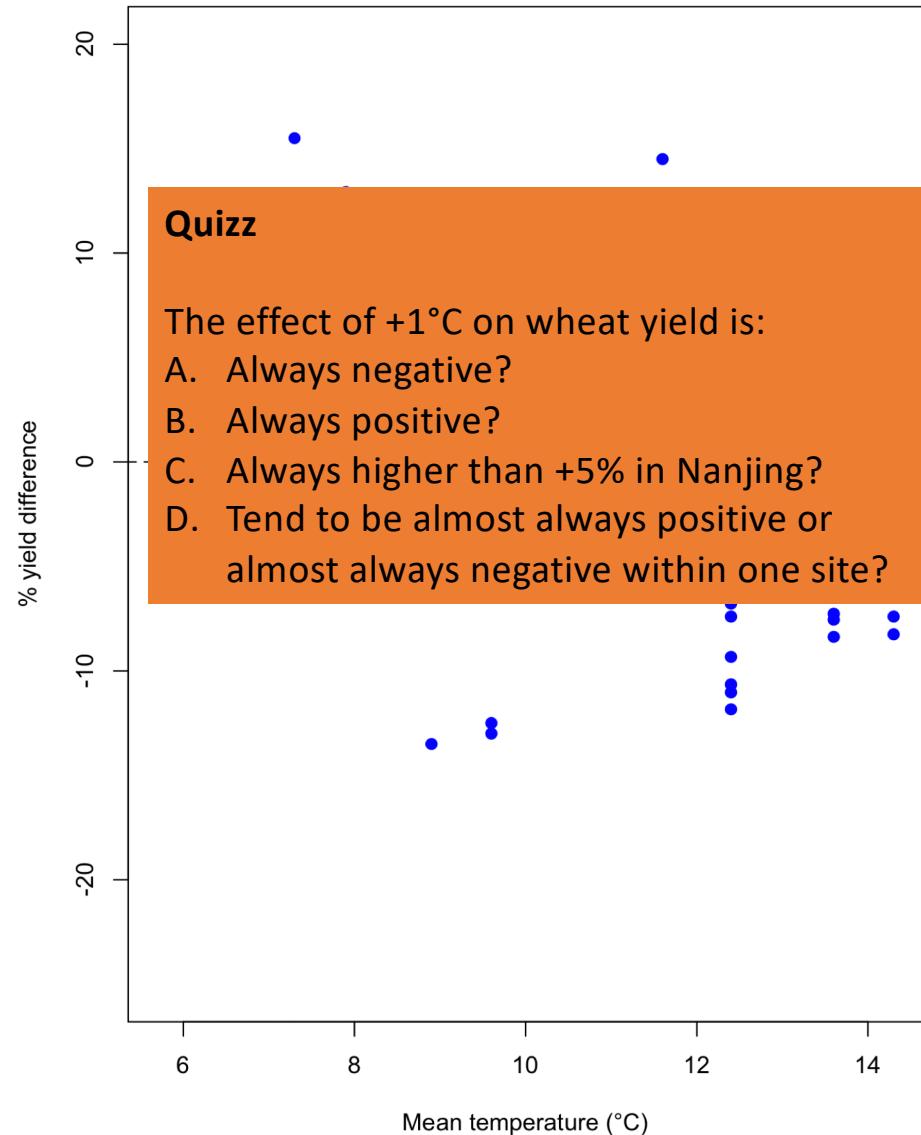
**Individual yield sensitivity values vs. Site average temperature**



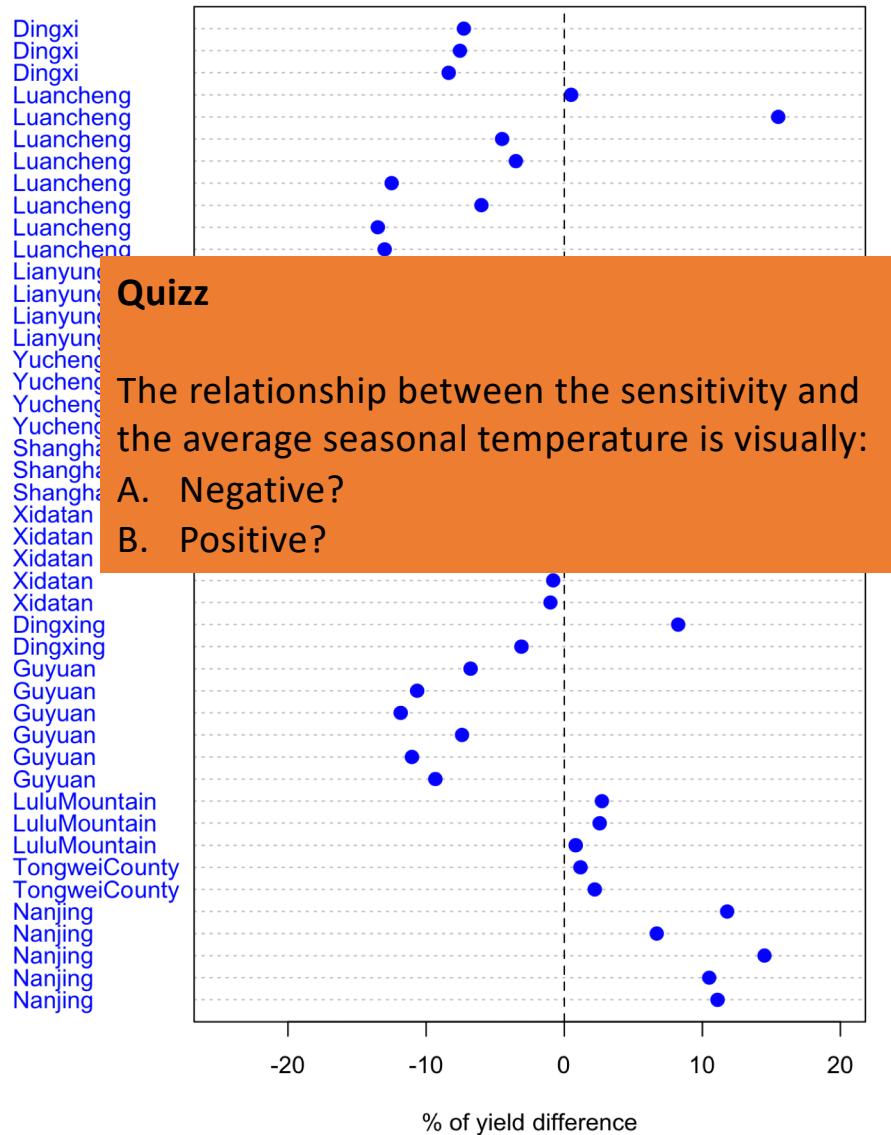
**Individual yield sensitivity values, by site**



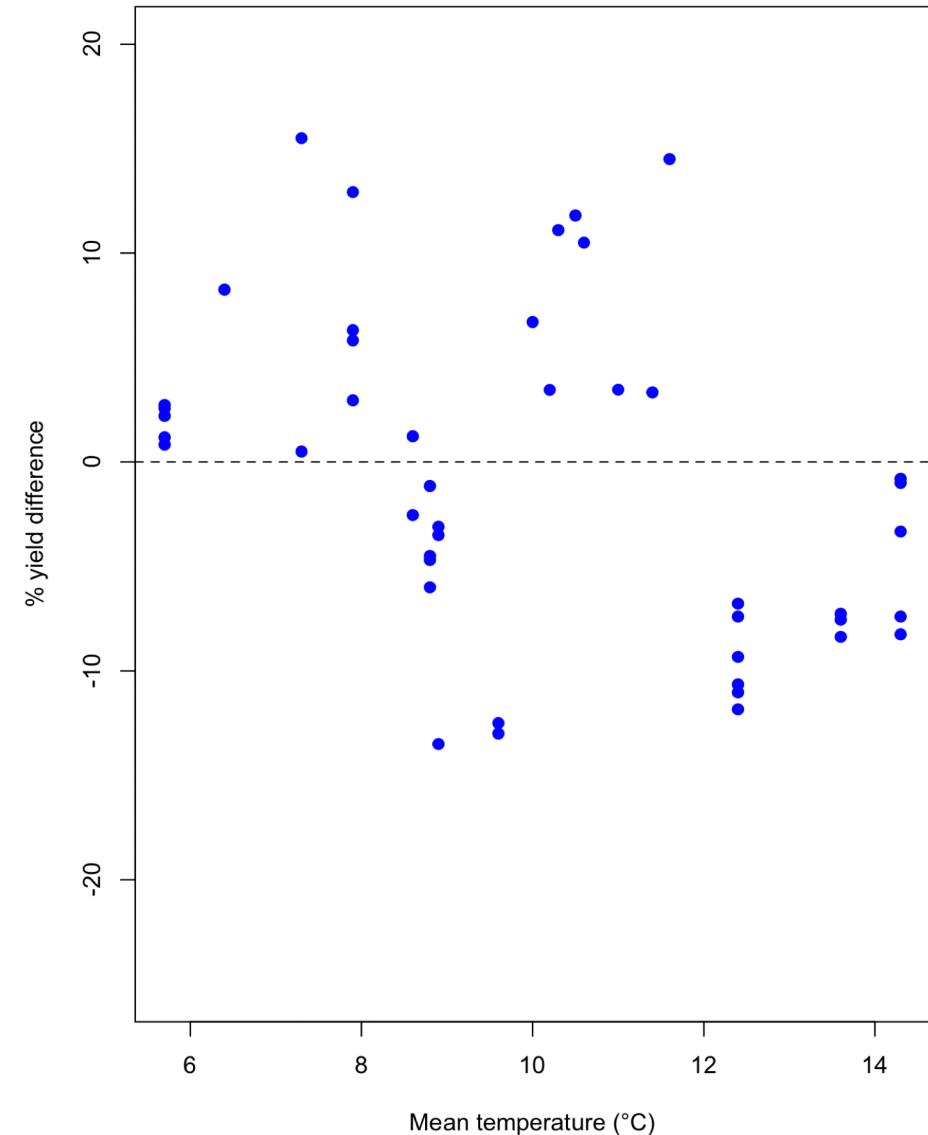
**Individual yield sensitivity values vs. Site average temperature**



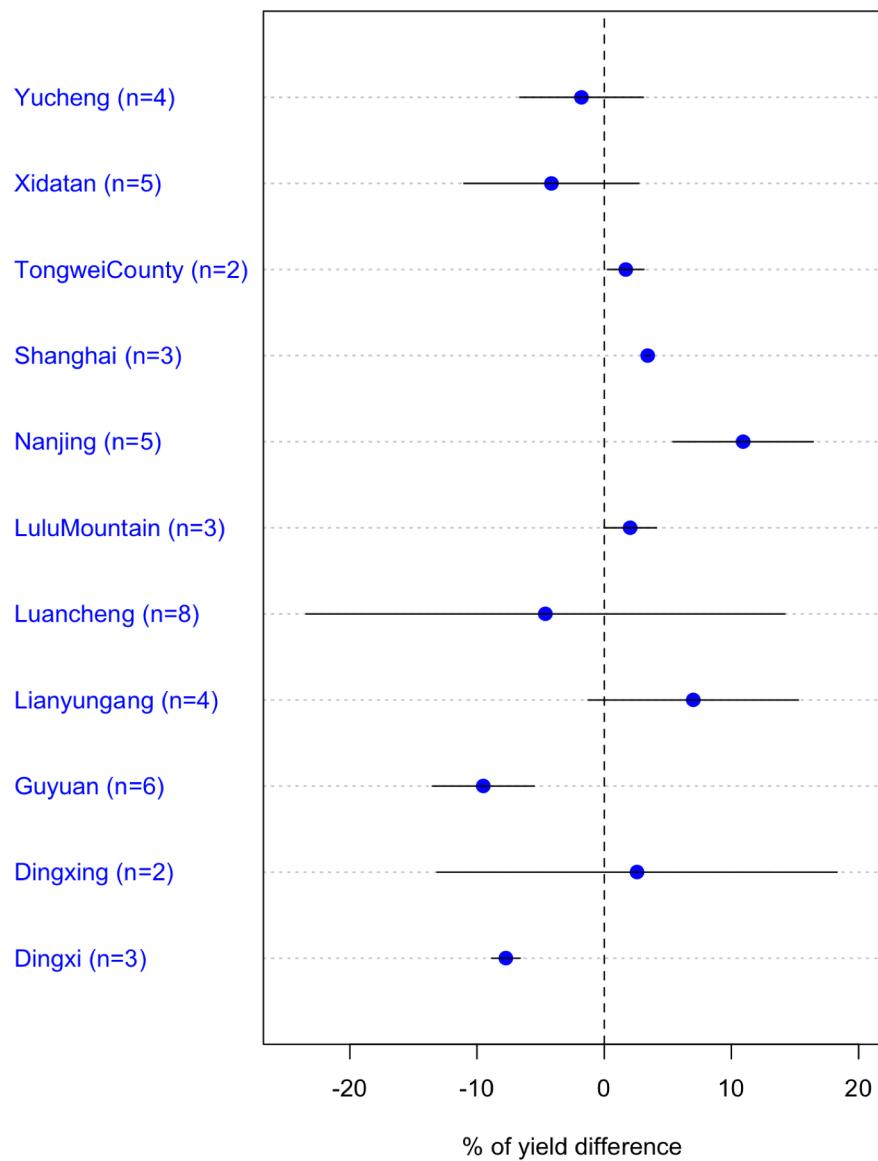
Individual yield sensitivity values, by site



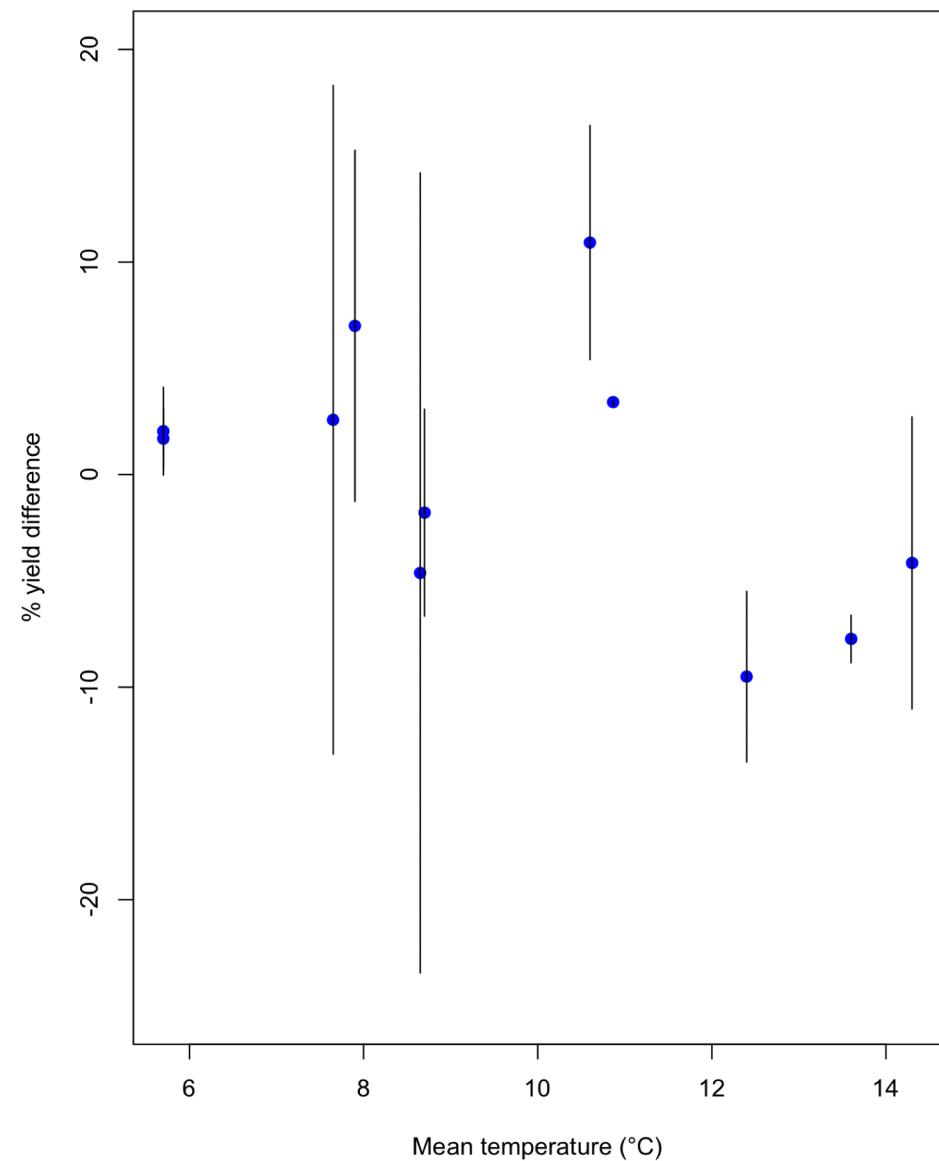
Individual yield sensitivity values vs. Site average temperature



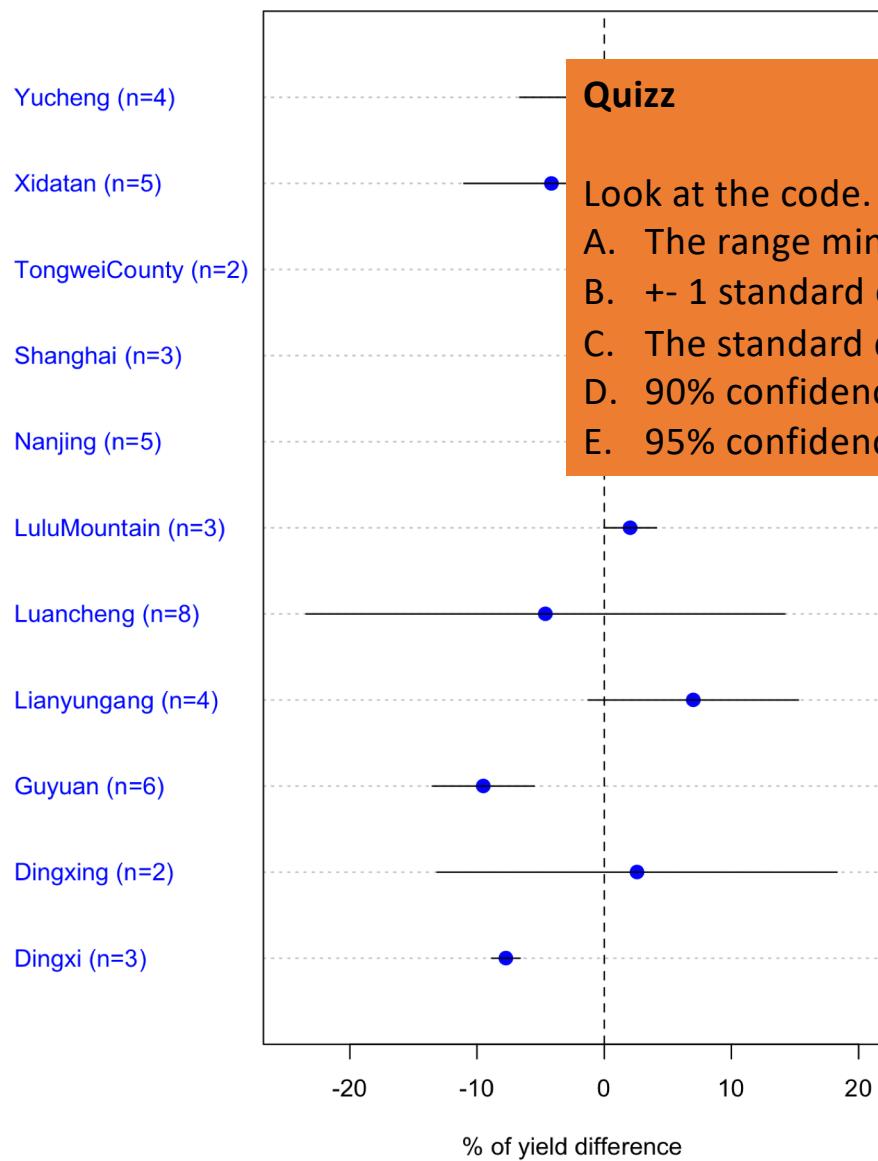
**Yield sensitivities to +1°C by site**



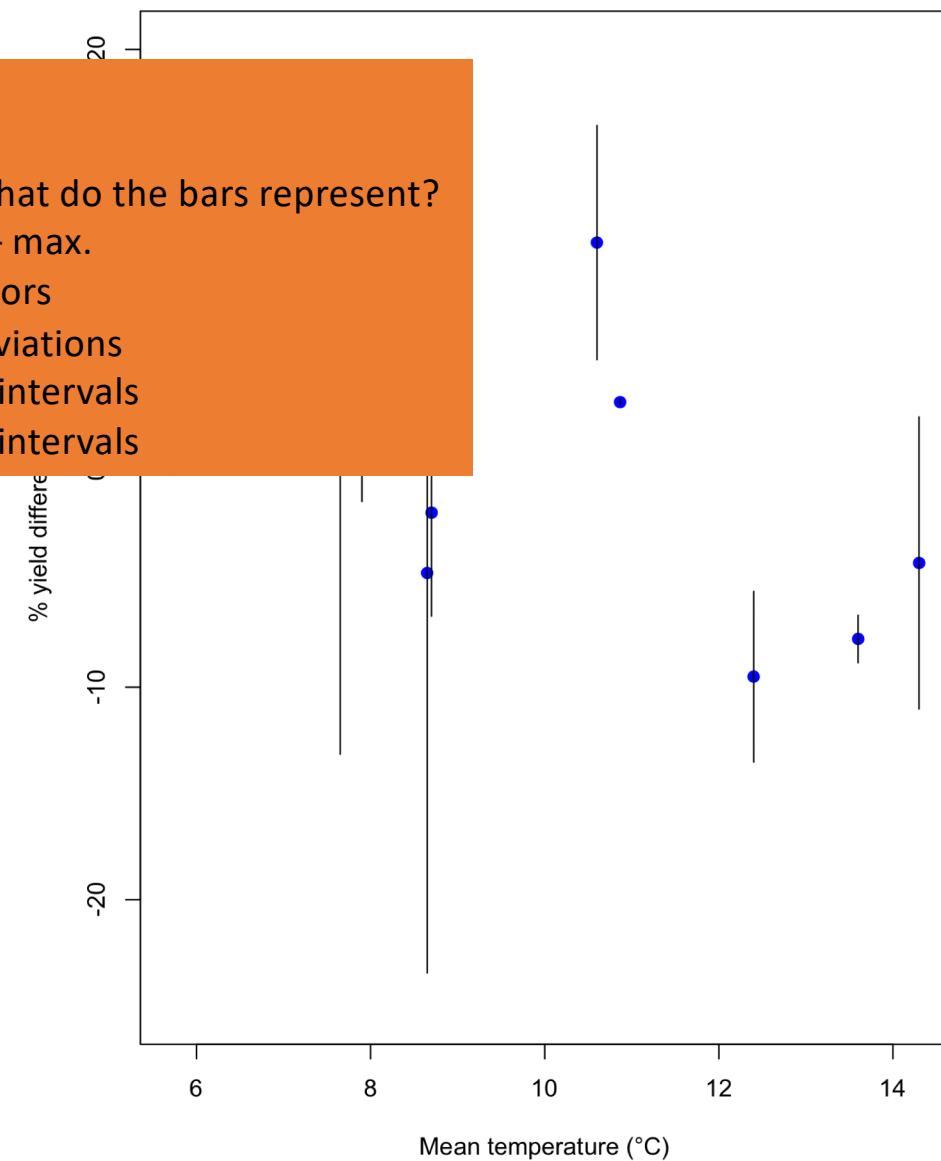
**Yield sensitivities to +1°C vs. Site average temperature**



**Yield sensitivities to +1°C by site**



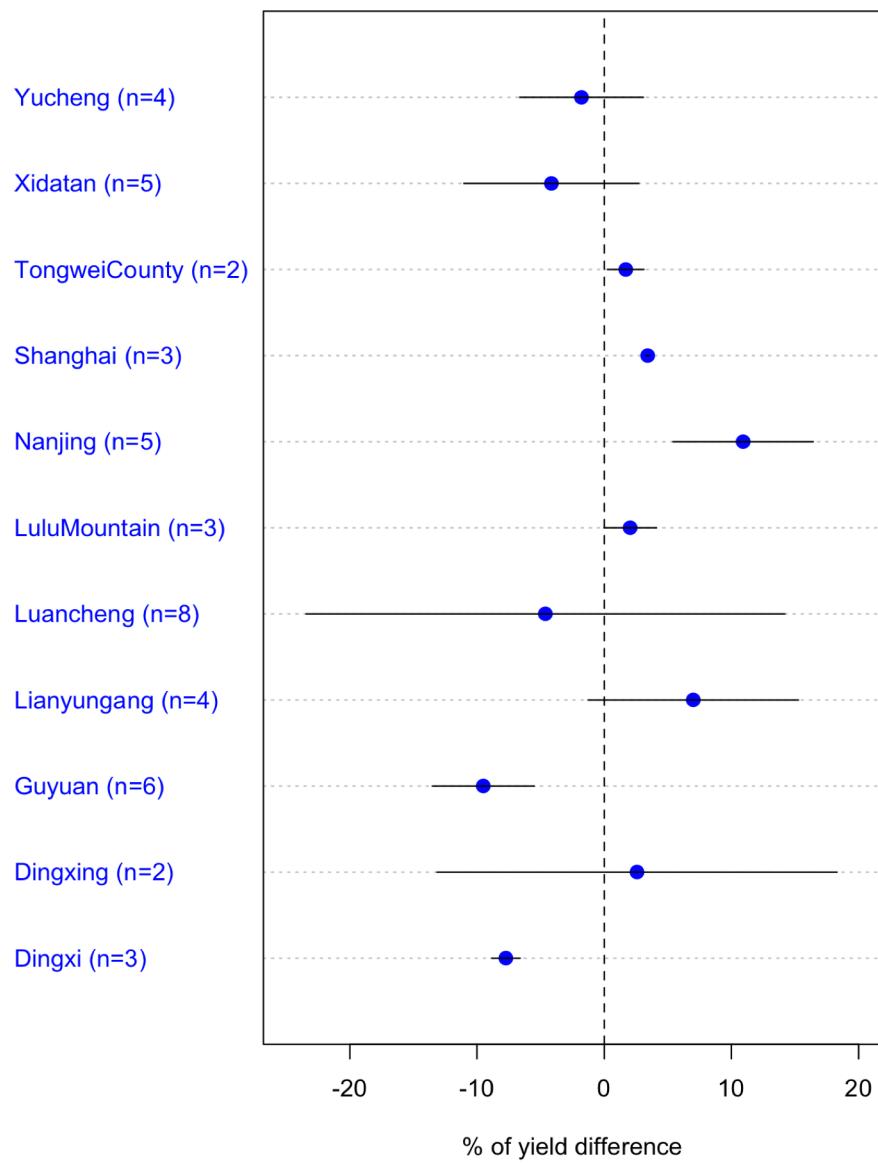
**Yield sensitivities to +1°C vs. Site average temperature**



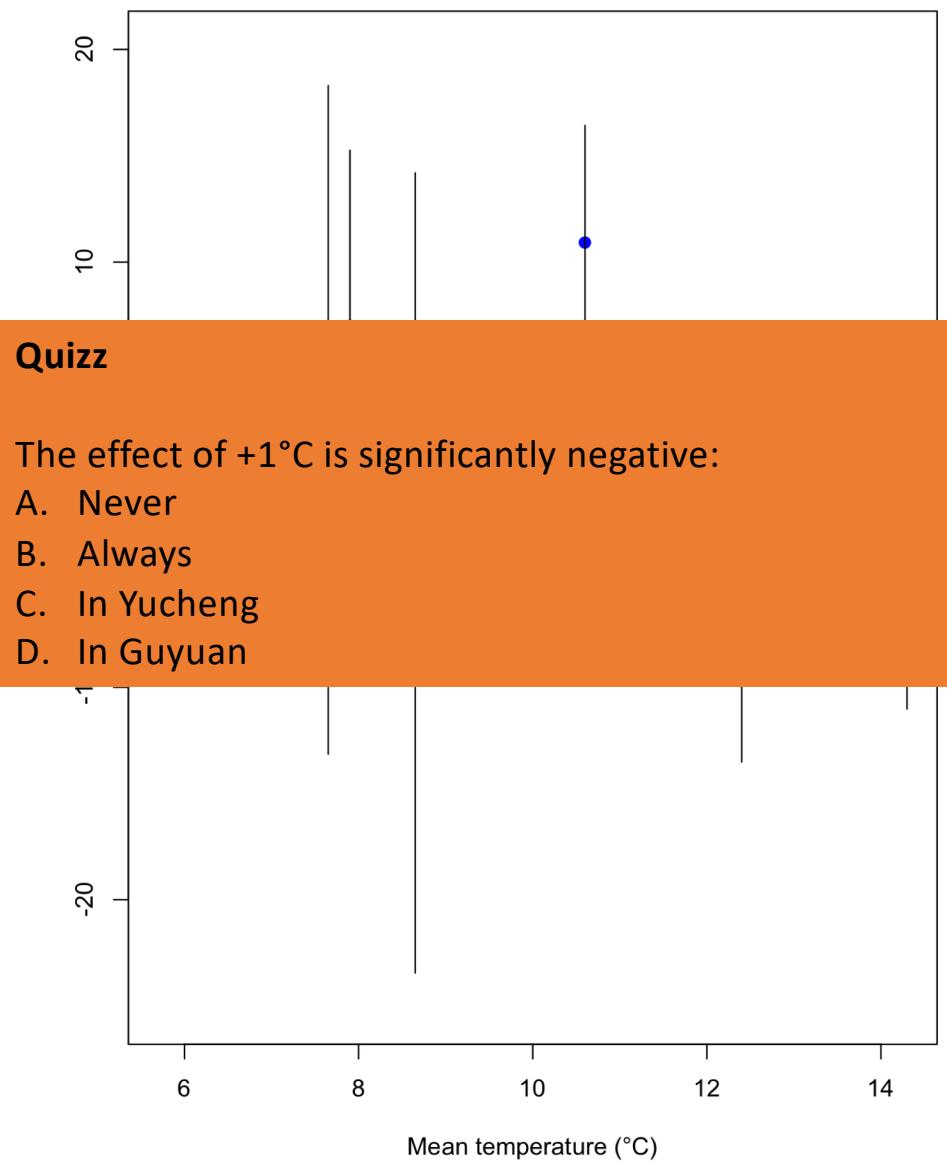
**Quizz**

- Look at the code. What do the bars represent?
- A. The range min. – max.
  - B. +- 1 standard errors
  - C. The standard deviations
  - D. 90% confidence intervals
  - E. 95% confidence intervals

**Yield sensitivities to +1°C by site**



**Yield sensitivities to +1°C vs. Site average temperature**

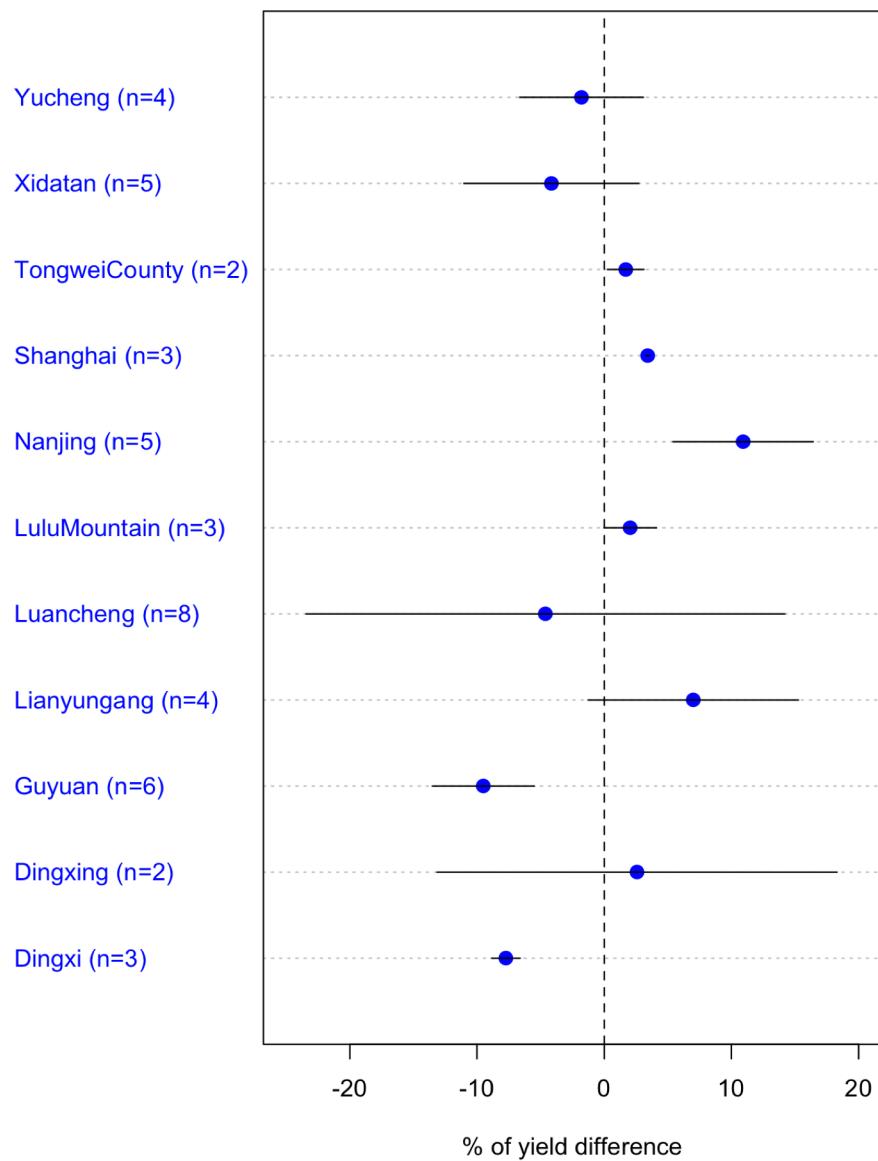


**Quizz**

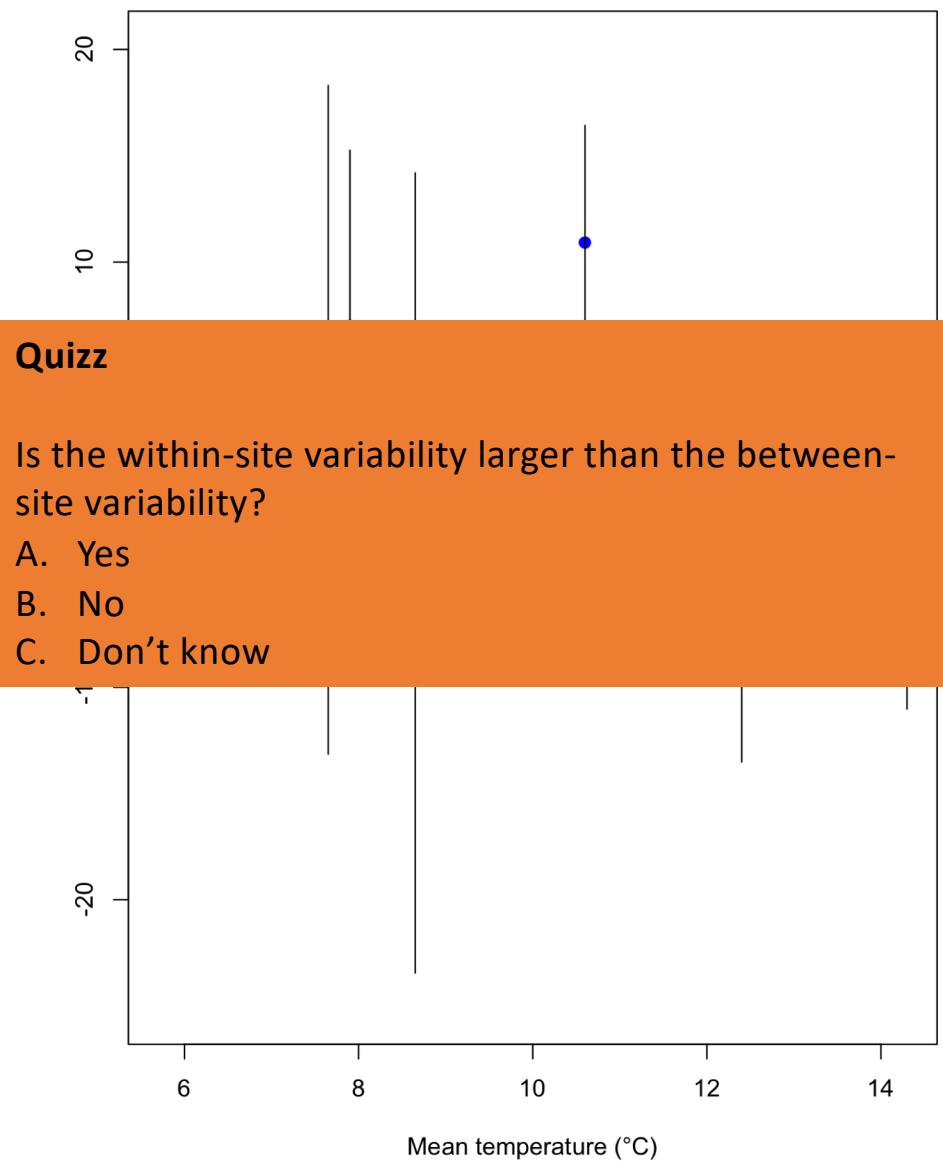
The effect of +1°C is significantly negative:

- A. Never
- B. Always
- C. In Yucheng
- D. In Guyuan

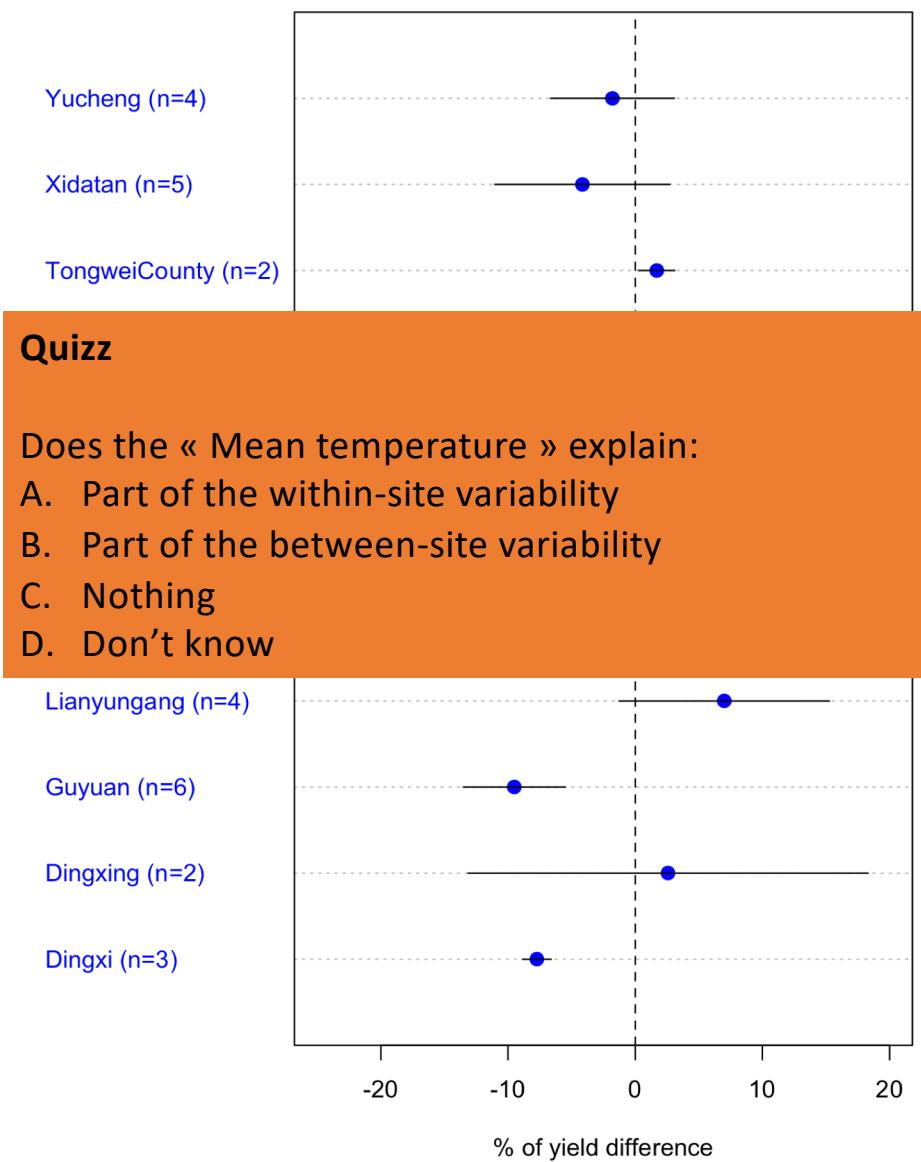
**Yield sensitivities to +1°C by site**



**Yield sensitivities to +1°C vs. Site average temperature**



**Yield sensitivities to +1°C by site**

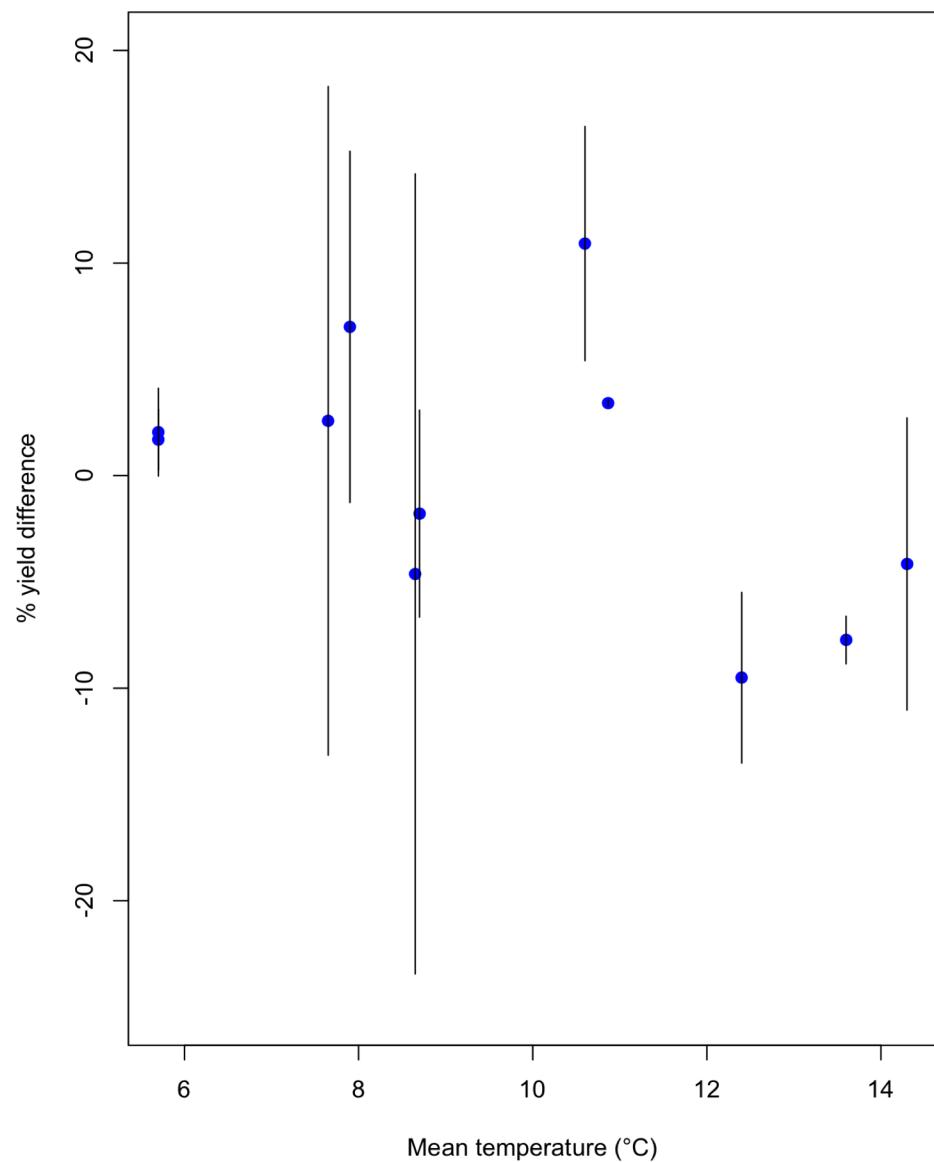


**Quizz**

Does the « Mean temperature » explain:

- A. Part of the within-site variability
- B. Part of the between-site variability
- C. Nothing
- D. Don't know

**Yield sensitivities to +1°C vs. Site average temperature**



# Steps

- Explore the dataset
- Define and fit Bayesian models
- Conclude

# Hierarchical statistical model

## « Random-effect model »

Within-study level:  $S_{ij} = \mu + b_i + \varepsilon_{ij}$        $\varepsilon_{ij} \sim N(0, \sigma_{\varepsilon i}^2)$

Overall mean      Effect of study i  
Yield sensitivity in study i, replicate j      Residual for jth data in study i

Overall mean      Effect of study i  
Yield sensitivity in study i, replicate j      Residual for jth data in study i

# Hierarchical statistical model

## « Random-effect model »

Within-study level:  $S_{ij} = \mu + b_i + \varepsilon_{ij}$        $\varepsilon_{ij} \sim N(0, \sigma_{\varepsilon i}^2)$

Overall mean      Effect of study i

Between-study level:  $b_i \sim N(0, \sigma_b^2)$

Prior: Gaussian and Uniform

# Hierarchical statistical model

## « Random-effect model »

Within-study level:  $S_{ij} = \mu + b_i + \varepsilon_{ij}$        $\varepsilon_{ij} \sim N(0, \sigma_{\varepsilon i}^2)$

Overall mean      Effect of study i

$\downarrow$                      $\downarrow$

Yield sensitivity in study i, replicate j      Residual for jth data in study i

Between-study level:  $b_i \sim N(0, \sigma_b^2)$

Prior: Gaussian and Uniform

Fitting algorithm: MCMC

# Hierarchical statistical model (with covariate)

## « Random-effect model »

Within-study level:  $S_{ij} = \mu + \alpha X_{ij} + b_i + \varepsilon_{ij}$        $\varepsilon_{ij} \sim N(0, \sigma_{\varepsilon i}^2)$

↑  
Yield sensitivity in  
study i, replicate j

↑  
Covariate value for study i replicate j

Between-study level:  $b_i \sim N(0, \sigma_b^2)$

Prior: Gaussian and Uniform

Fitting algorithm: MCMC

## **Open Atelier2020\_2.R**

This file includes four Bayesian models coded with rjags.

- What are the differences between these models?
- Which one is the simplest?
- Which one is the most complex?

# Run model 1

```
Iterations = 100010:2e+05
Thinning interval = 10
Number of chains = 3
Sample size per chain = 10000
```

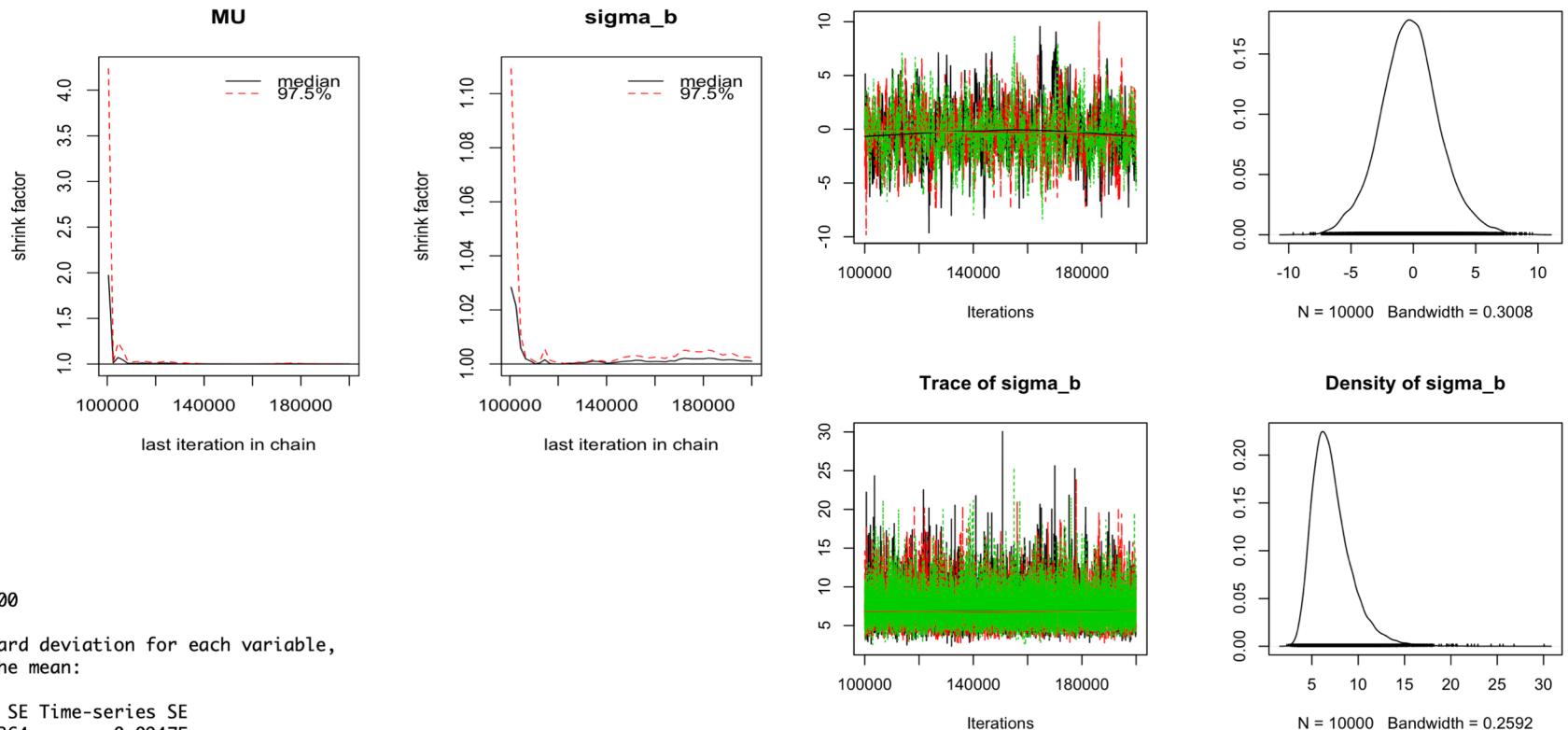
1. Empirical mean and standard deviation for each variable, plus standard error of the mean:

	Mean	SD	Naive SE	Time-series SE
MU	-0.2577	2.362	0.01364	0.09475
sigma_b	7.1907	2.205	0.01273	0.02351

2. Quantiles for each variable:

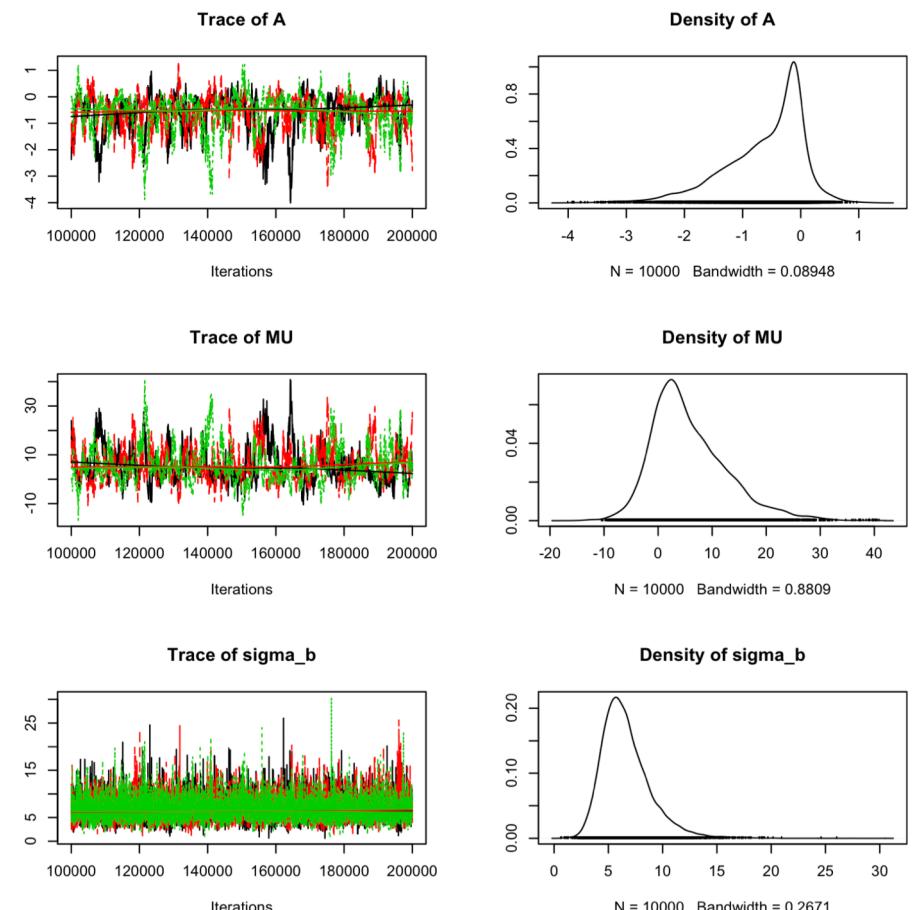
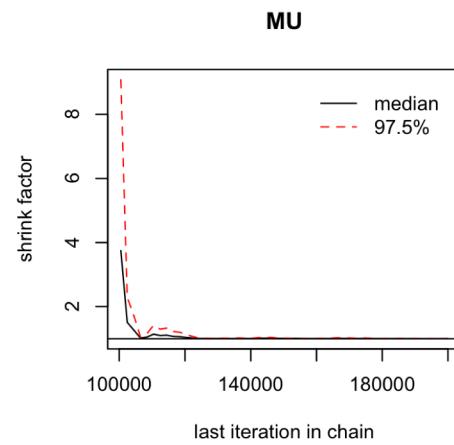
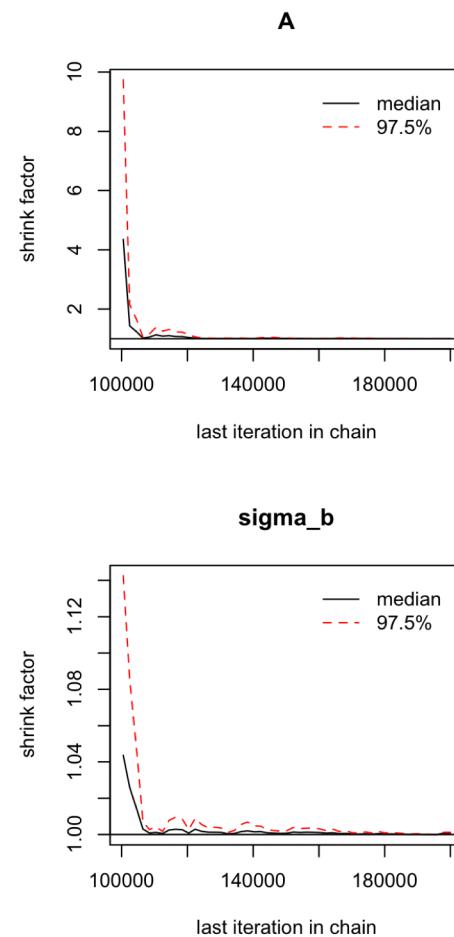
	2.5%	25%	50%	75%	97.5%
MU	-4.985	-1.784	-0.286	1.205	4.668
sigma_b	4.086	5.677	6.798	8.252	12.614

```
> dic.samples(model, 10000)
|*****| 100%
Mean deviance: 236.2
penalty 437.3
Penalized deviance: 673.5
```



- What is the posterior mean of the yield sensitivity?
- What is the 95% credibility interval?
- Practical conclusion?

## Run model 2



- What are the meanings of A, MU, sigma\_b?
- Does the model includes other parameters?

## Run model 2

```
> summary(samples)
```

Iterations = 1000010:2e+05  
Thinning interval = 10  
Number of chains = 3  
Sample size per chain = 10000

1. Empirical mean and standard deviation for each variable, plus standard error of the mean:

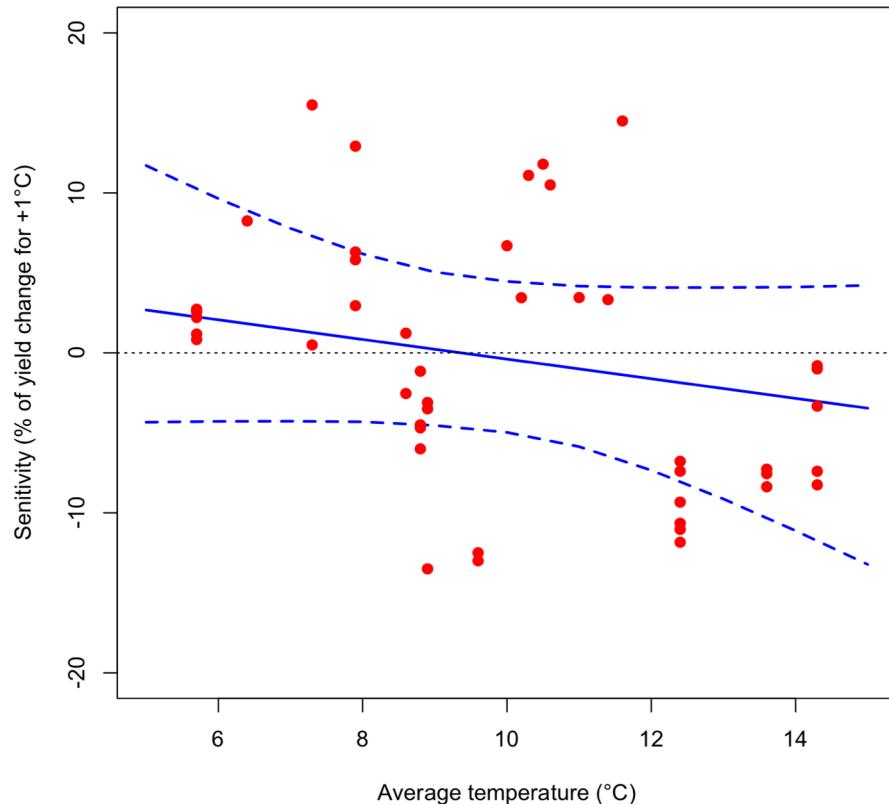
	Mean	SD	Naive SE	Time-series SE
A	-0.6139	0.689	0.003978	0.04979
MU	5.7506	7.015	0.040500	0.49460
sigma_b	6.6276	2.280	0.013165	0.04103

2. Quantiles for each variable:

	2.5%	25%	50%	75%	97.5%
A	-2.293	-1.0024	-0.4249	-0.1132	0.3409
MU	-4.960	0.8774	4.3804	9.6299	22.8198
sigma_b	3.276	5.0932	6.2541	7.7468	12.1035

```
> dic.samples(model, 10000)
|*****| 100%
```

Mean deviance: 243.4  
penalty 1146  
Penalized deviance: 1389



- What is the effect of the average temperature?
- Was this variable able to explain a big part of the between-site variability?

# Run model 1bis

```
> summary(samples)
```

```
Iterations = 100010:2e+05
Thinning interval = 10
Number of chains = 3
Sample size per chain = 10000
```

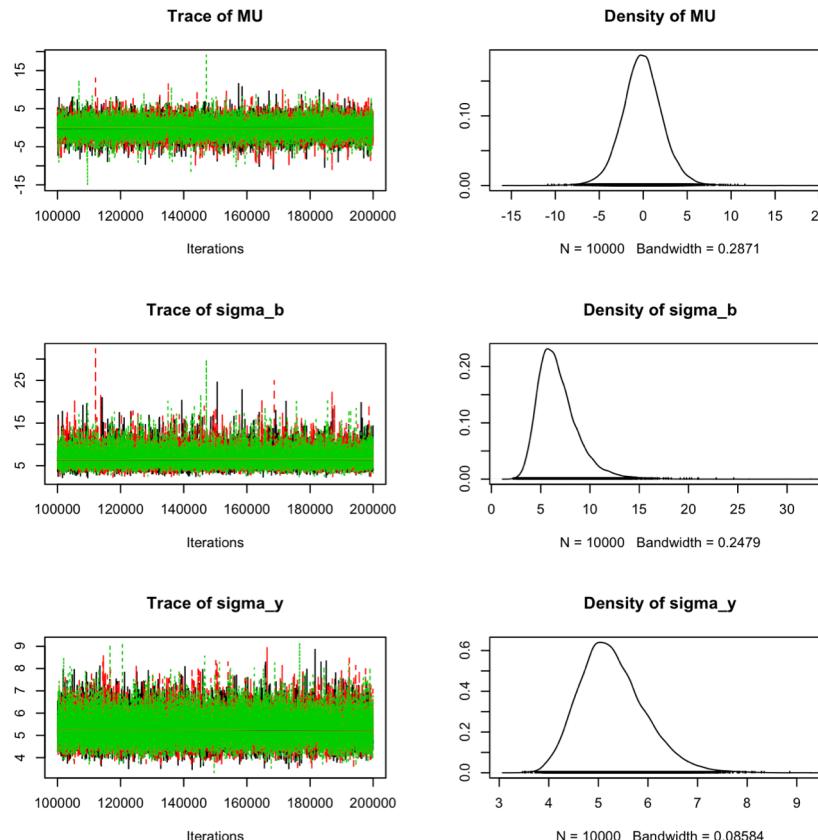
1. Empirical mean and standard deviation for each variable, plus standard error of the mean:

	Mean	SD	Naive SE	Time-series SE	SE
MU	-0.09876	2.2985	0.013270	0.01889	
sigma_b	6.76543	2.1056	0.012157	0.01323	
sigma_y	5.28503	0.6643	0.003836	0.00380	

2. Quantiles for each variable:

	2.5%	25%	50%	75%	97.5%
MU	-4.593	-1.536	-0.1248	1.317	4.528
sigma_b	3.751	5.320	6.4026	7.783	11.902
sigma_y	4.173	4.823	5.2160	5.676	6.772

```
> dic.samples(model, 10000)
|*****| 100%
Mean deviance: 275.9
penalty 11.3
Penalized deviance: 287.2
```



- How does this model compare with model 1?
- Is it relevant to consider different within-site variances?

# Run model 2bis

> summary(samples)

```
Iterations = 100010:2e+05
Thinning interval = 10
Number of chains = 3
Sample size per chain = 10000
```

1. Empirical mean and standard deviation for each variable, plus standard error of the mean:

	Mean	SD	Naive SE	Time-series SE
A	-2.045	0.8503	0.004909	0.027451
MU	19.581	8.3981	0.048487	0.266484
sigma_b	7.273	2.6770	0.015456	0.054430
sigma_y	4.679	0.6271	0.003621	0.006523

2. Quantiles for each variable:

	2.5%	25%	50%	75%	97.5%
A	-3.917	-2.552	-1.974	-1.461	-0.5895
MU	5.001	13.773	18.979	24.668	37.8528
sigma_b	3.719	5.435	6.748	8.501	13.8902
sigma_y	3.632	4.234	4.614	5.055	6.0909

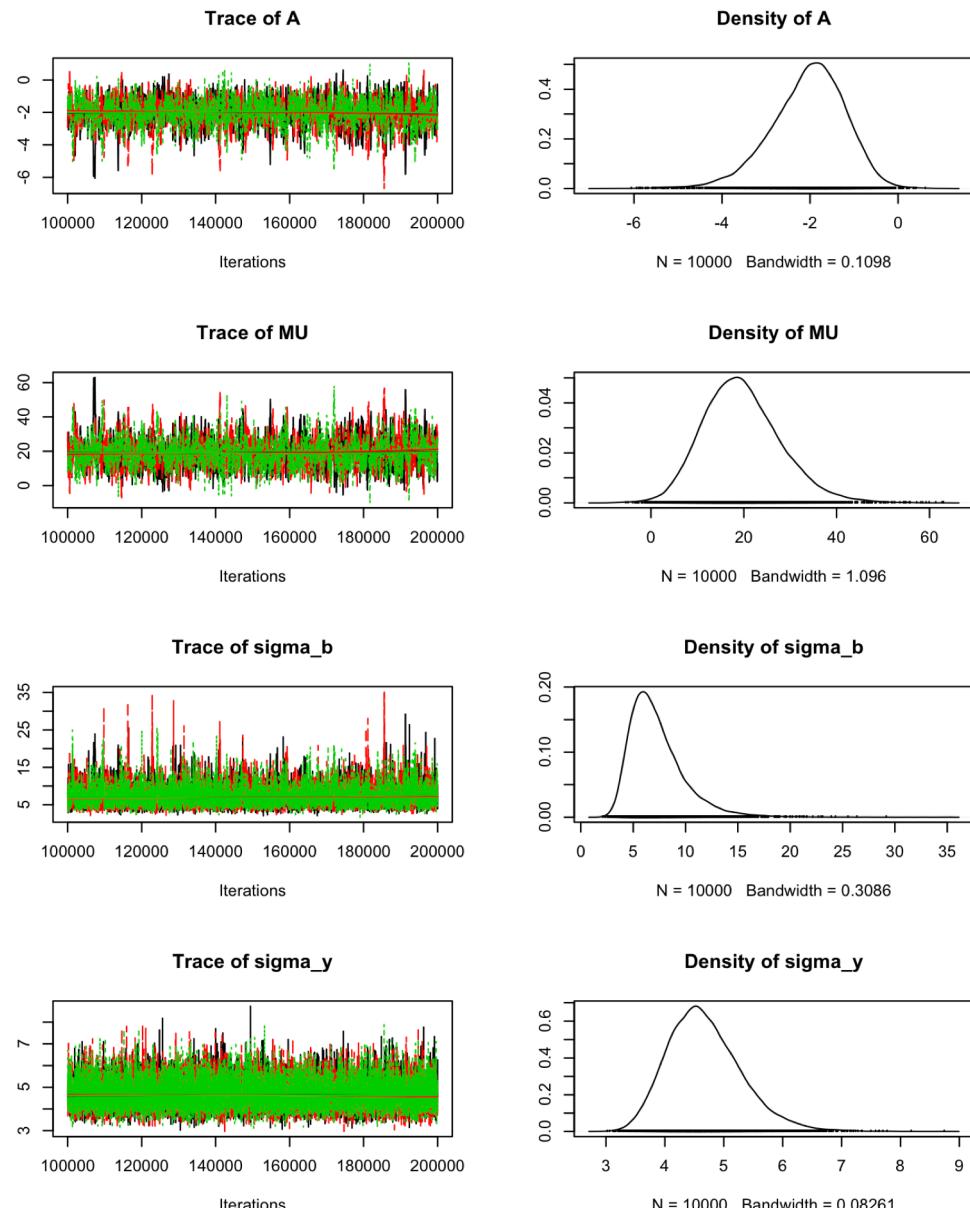
> dic.samples(model, 10000)  
|\*\*\*\*\*| 100%

Mean deviance: 264.7

penalty 12.43

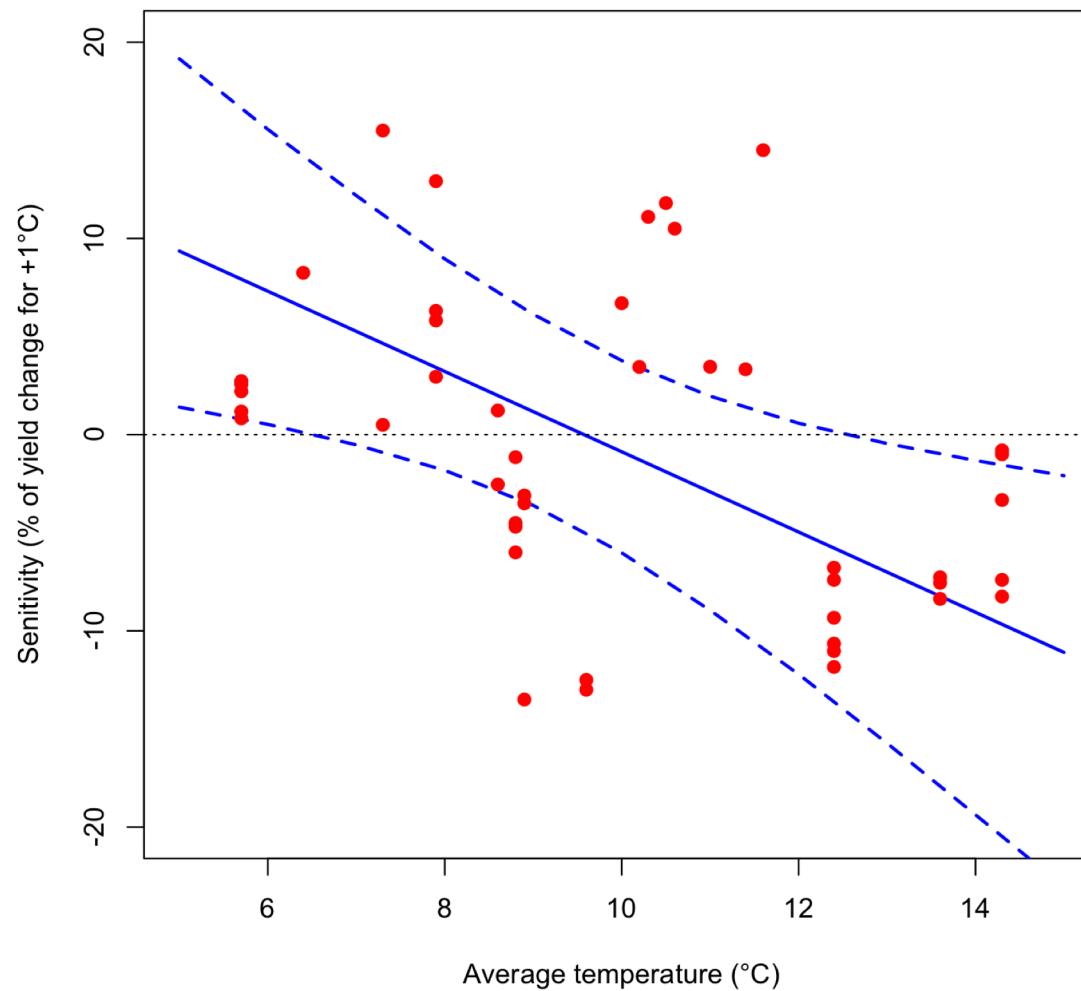
Penalized deviance: 277.2

- How does this model compare with the other models?
- What is the effect of the average temperature?



## Run model 2bis

- What is the yield sensitivity when the average temperature is of 6°C or of 14°C?
- Can you anticipate a shift of the wheat growing area due to climate change based on this model?



If you had 2 days available, how will you improve this model?