2026

Causal inference for assessing the impact of extreme weather conditions on crop yields

David Makowski

MIA-PS
INRAE Université Paris-Saclay

in collaboration with

- Claudia Mendes (INRAE MIA-PS, now PhD student in Montpellier)
- Marc Fermaud (INRAE Bordeaux)
- François Brun (ACTA)
- Marc Raynal (Institut de la Vigne et du Vin)

Outline

- Objective
- Methods
- Application 1: Maize and Sunflower
- Application 2: Vine
- Conclusions

Outline

- Objective
- Methods
- Application 1: Maize and Sunflower
- Application 2: Vine
- Conclusions

Scientists are often asked to answer questions of (strong) societal interest

Aspirin to reduce risk for cardiovascular disease?

LOCAL NEWS ›

Task Force Proposes Adults 60+ Should Not Take Daily Aspirin To Prevent Heart Disease Or Stroke

CBS NEWS BOSTON

OCTOBER 12, 2021 / 12:16 PM / CBS BOSTON

What is the impact of neonicotinoids (pesticide) on honey bees?

The Economist

Menu | Weekly edition | Search

Science & technology | Buzz kill

Neonicotinoids can harm some bees

Two studies show the risks from pesticides

Jul 1st 2017

Share

What is the impact of drought on yield?

Startseite » Erde/Umwelt » Dürre in Deutschland 2018: Gibt es einen neuen Jahrhundertsommer?

News
09.07.2018
Lesedauer ca. 4 Minuten
Drucken
Teilen

DÜRRE IN DEUTSCHLAND 2018

Gibt es einen neuen Jahrhundertsommer?

Die Trockenheit ist nicht so spektakulär wie Sturzfluten - richtet jedoch immense Schäden an. Dabei sind Dürreperioden in Deutschland gar nicht so selten.

von Lars Fischer

© MARCCOPHOTO / GETTY IMAGES / ISTOCK (AUSSCHNITT)

What is the impact of frost on yield?

Bergeracois : « Plus de 50 % des bourgeons touchés » dans les vignes à cause du gel

Lecture 2 min

Accueil • Vin

Sur leurs 50 hectares, les Verdots ont pu en protéger 5 par le système dit de l'aspersion. © Crédit photo : Facebook Vignoble des Verdots

Par Grégoire Morizet - g.morizet@sudouest.fr
Publié le 11/04/2021 à 15h37.
Mis à jour le 11/04/2021 à 19h14.

Écouter Réagir Voir sur la carte Partager

Les estimations des dégâts causés par la vague de froid se poursuivent dans le vignoble. Elles confirment les premières impressions "d'une production qui sera atteinte »

What is the impact of plant diseases?

Filière : Une récolte de betteraves « historiquement basse »

La Confédération générale des planteurs de betteraves (CGB) a dressé un bilan « catastrophique » de la récolte de 2020, estimée à 27 millions de tonnes, largement impactée par la jaunisse de la betterave.

JE M'ABONNE

Publié le 01/12/2020, Mis à jour le 03/07/2025

Partager

Lire plus tard

POLLINIS

FAIRE UN DON

LES POLLINISATEURS | L'EXTINCTION | NOUS CONNAÎTRE | NOS CAMPAGNES | AGIR | PROJETS | ÉTUDES | PUBLICATIONS

PESTICIDES / NÉONICOTINOÏDES

RÉCIT D'UN MENSONGE GOUVERNEMENTAL SUR LES PERTES DE RENDEMENT DE LA FILIÈRE BETTERAVIÈRE

Le gouvernement français a bâti son projet de loi permettant des dérogations pour utiliser les néonicotinoïdes sur des chiffres erronés. La diminution réelle des rendements de betteraves sucrières s'élève à 15 %, très loin des 30 à 50 % relayés en chœur depuis cet été par les betteraviers, les politiques et les médias.

CATÉGORIES : DÉPÊCHES **DATE :** 22 OCTOBRE 2020

What is the impact of pollinators on yields?

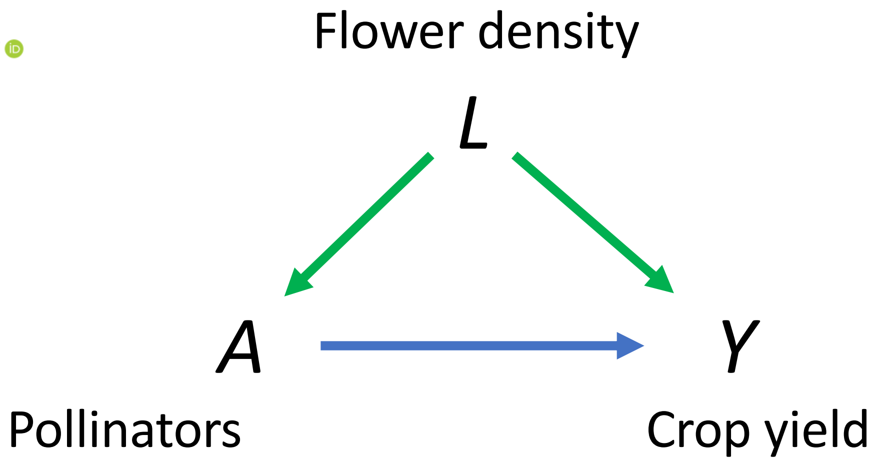
Received: 15 September 2025 | Accepted: 14 January 2026
DOI: 10.1111/1365-2435.70269

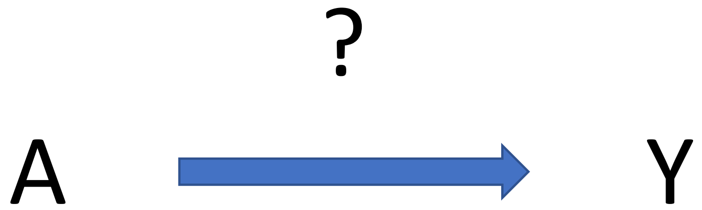
BRITISH ECOLOGICAL SOCIETY Functional Ecology

COMMENTARY

Are crop yields limited by pollinators? Proper assessments using pollinator gradients require measurements of flower density and yield potential

Stan Chabert[1] | James H. Cane[2] | Bernard E. Vaissière[3] | Rachel E. Mallinger[4]




?
A
Y

$$A \xrightarrow{?} Y$$

Three types of causality:

• For one individual	A → Y for an individual
• In average in a population of individuals	A → Y in average
• Conditionnal causality	A → Y under certain conditions

What is an individual causal effect?

A: Treatment variable (either 0 or 1, here)
Y: Outcome for an individual

A has a causal effect on an individual's outcome *Y* if

$$Y^{a=1} \neq Y^{a=0}$$

for the individual

What is an individual causal effect?

A: Exposition to neonicotinoid (0 or 1)
Y: Bee alive, Bee dead (0, 1)

Neonicotinoid has a causal effect on the bee survival if

$$Y^{a=1} \neq Y^{a=0}$$

for the individual bee

What is an individual causal effect?

A: Exposition to neonicotinoid (0 or 1)
Y: Bee alive, Bee dead (0, 1)

Neonicotinoid has a causal effect on the bee survival if

$$Y^{a=1} \neq Y^{a=0}$$

for the individual bee

This is the same bee!

What is an average causal effect?

There is an average causal effect in the population if:

$$\mathrm{E}[Y^{a=1}] \neq \mathrm{E}[Y^{a=0}]$$

Causal effect of drought on crop production

- *A*: Drought occurrence (0 or 1)
- *Y*: Crop yield in a site-year, e.g., wheat field in a given field in 2023

The drought occurrence *A* has a causal effect on an individual's outcome *Y* if

$$Y^{a=1} \neq Y^{a=0}$$

for the site-year considered (same site-year!)

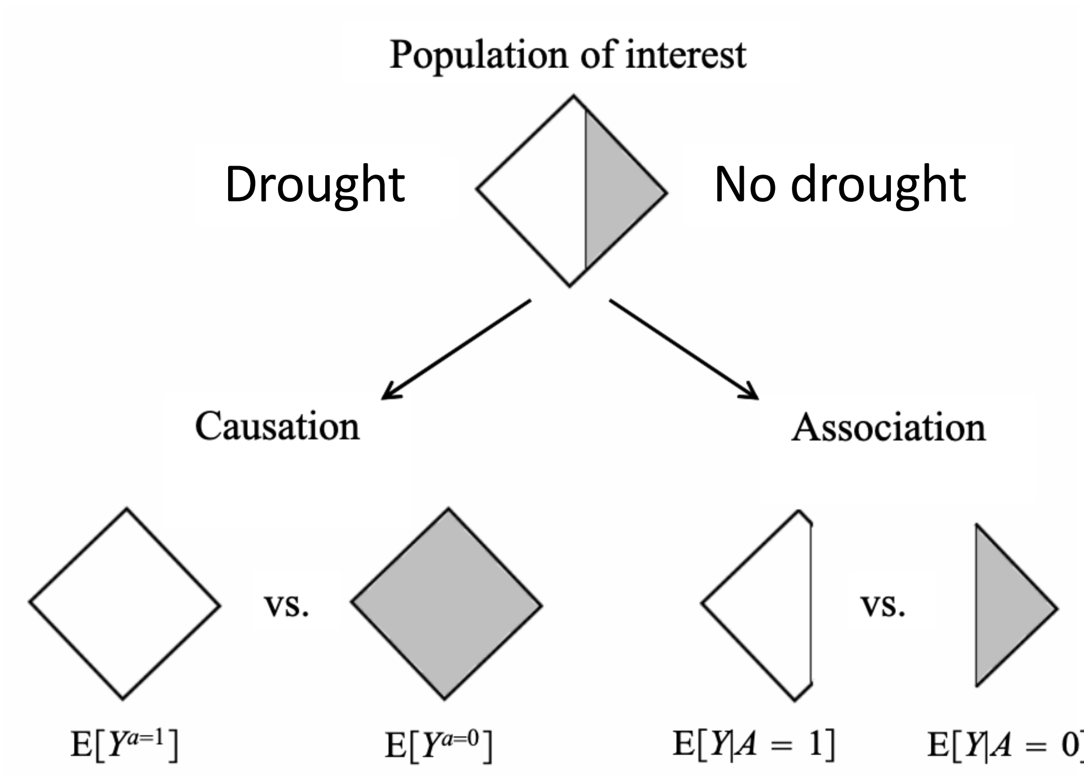
Causal effect of a drought on crop production

- *A*: Drought occurrence (0 or 1)
- *Y*: Crop yield in a site-year
- Population: All wheat site-years in France

There is an average causal effect of the drought on wheat yield in France if:

$$\mathrm{E}[Y^{a=1}] \neq \mathrm{E}[Y^{a=0}]$$

Causation *vs*. Association



$$E[Y^{a=1}] \neq E[Y^{a=0}]$$

$$E[Y|A=1] \neq E[Y|A=0]$$

Causal analysis

Causal discovery

« Learning causal relationships among variables from observational data »

Causal inference

« Estimating the causal effect of a specific variable (treatment) over a certain outcome of interest »

Objective: estimating $E(Y^{a=1}) - E(Y^{a=0})$

https://doi.org/10.1002/widm.1449
https://doi.org/10.1177/25152459241236149

Causal analysis

Causal discovery

« Learning causal relationships among variables from observational data »

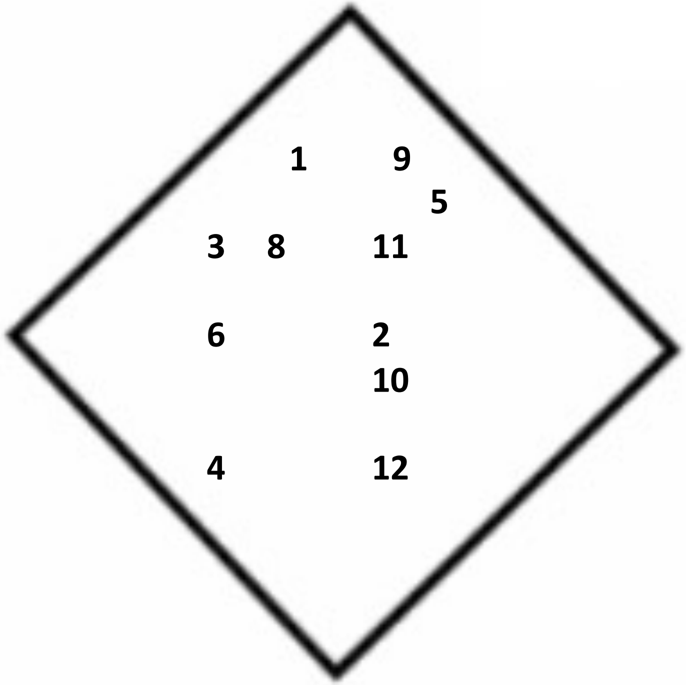
Causal inference

« Estimating the causal effect of a specific variable (treatment) over a certain outcome of interest »

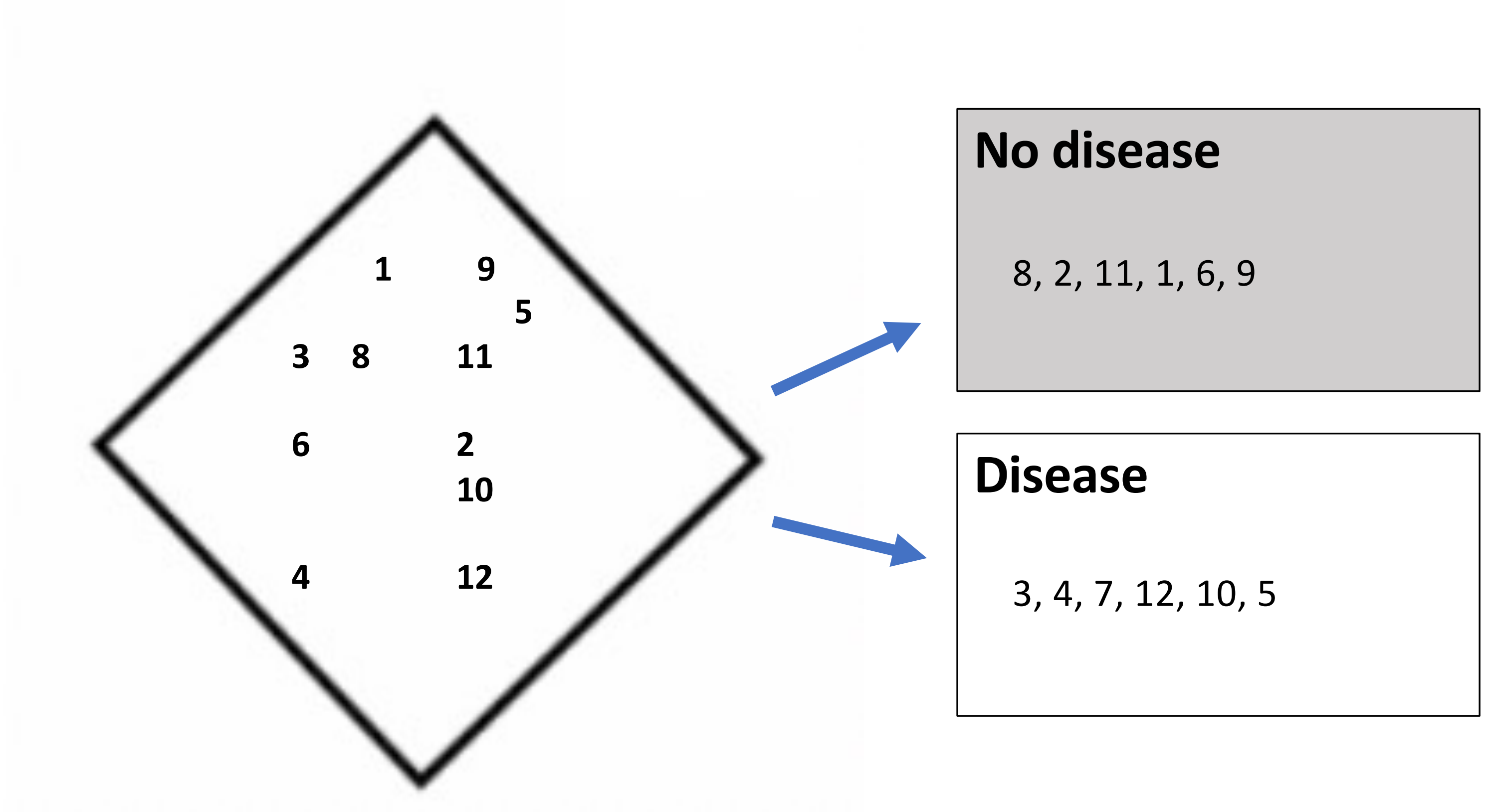
Objective: estimating $E(Y^{a=1}) - E(Y^{a=0})$

https://doi.org/10.1002/widm.1449
https://doi.org/10.1177/25152459241236149

Randomized controlled trial (RCT)

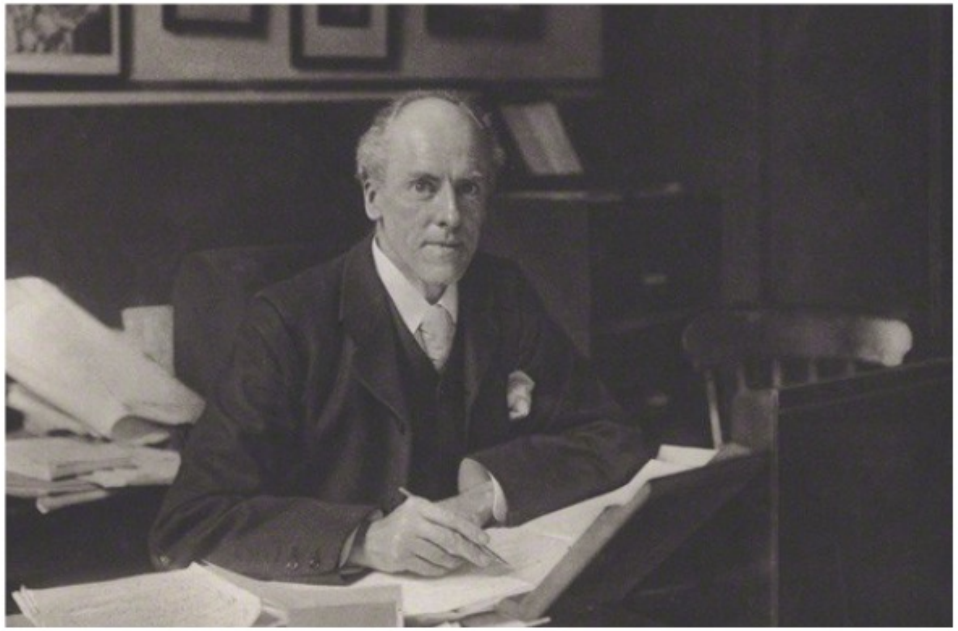


Randomized controlled trial (RCT)



Origins of modern experimental trial

Origins of modern experimental trial



Karl Pearson in 1910
English mathematician

From Gamsjager (2015)

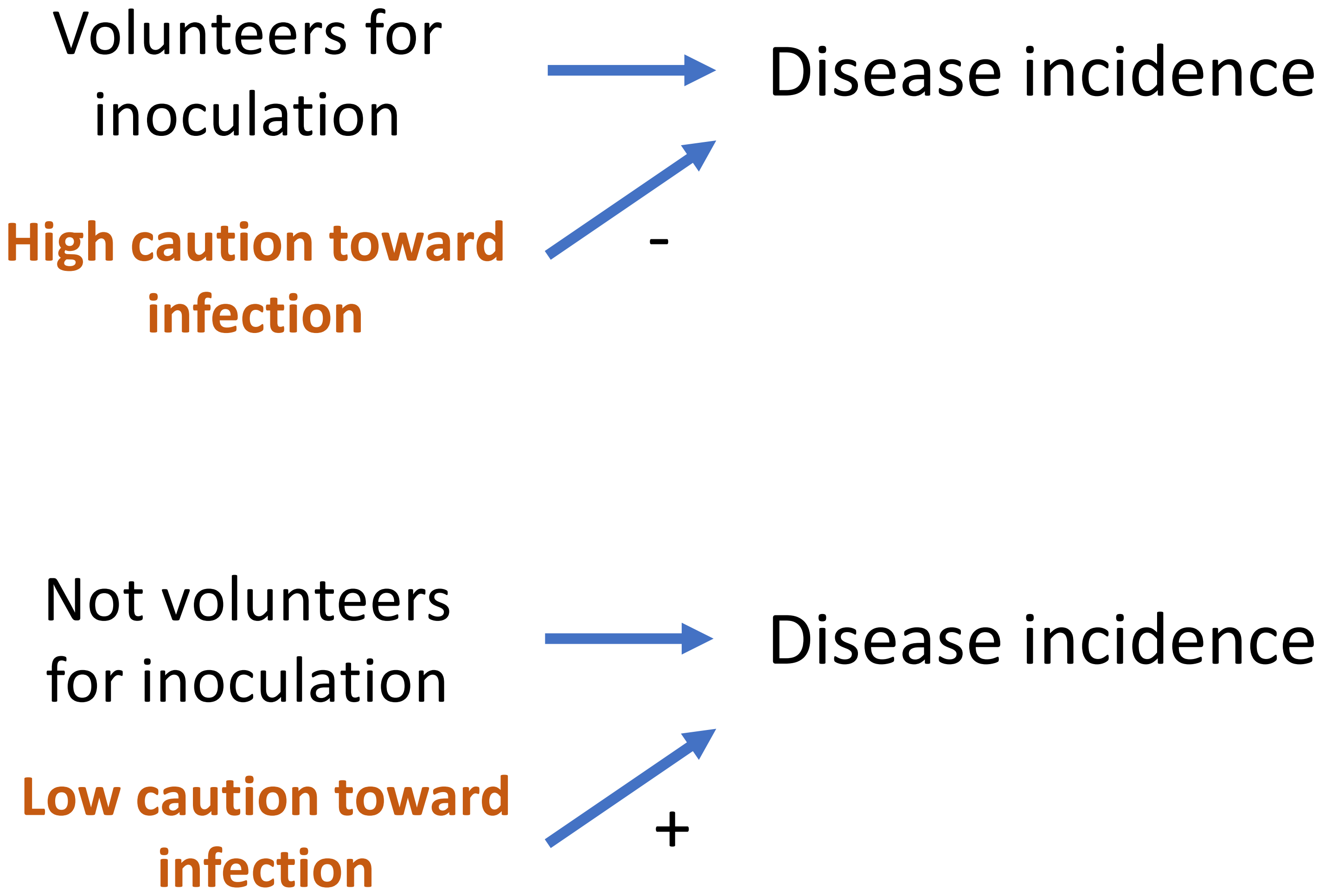
REPORT ON CERTAIN ENTERIC FEVER INOCULATION STATISTICS.

PROVIDED BY LIEUTENANT-COLONEL R. J. S. SIMPSON, C.M.G., R.A.M.C.

By KARL PEARSON, F.R.S.,
Professor of Applied Mathematics, University College, London.

Volunteers for inoculation → Disease incidence

Not volunteers for inoculation → Disease incidence


Volunteers for inoculation
Disease incidence
High caution toward infection
-
Not volunteers for inoculation
Disease incidence
Low caution toward infection
+

Origins of modern experimental trial

1907

THE ACTION OF CAFFEINE ON THE CAPACITY FOR MUSCULAR WORK. By W. H. R. RIVERS and H. N. WEBBER.

(From the Psychological Laboratory, Cambridge.)

https://doi.org/10.1113/jphysiol.1907.sp001215

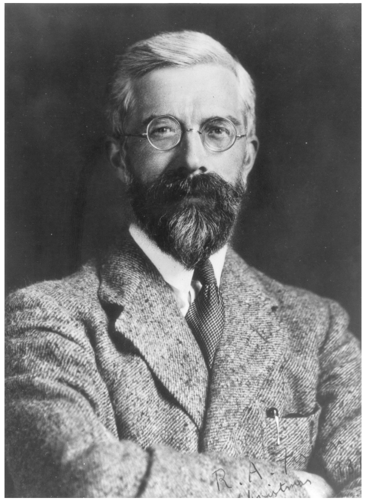
A key tool to avoid confounding:

Randomized controlled trial

Rothamsted (UK), 1914

https://www.adelaide.edu.au/library/special/exhibitions/significant-life-fisher/rothamsted/

Ronald A. Fisher (1890-1962)



Fisher's idea:

Control the identifiable factor(s) in the design and randomise over the others.

Randomized controlled trial is now often viewed as a gold standard in science

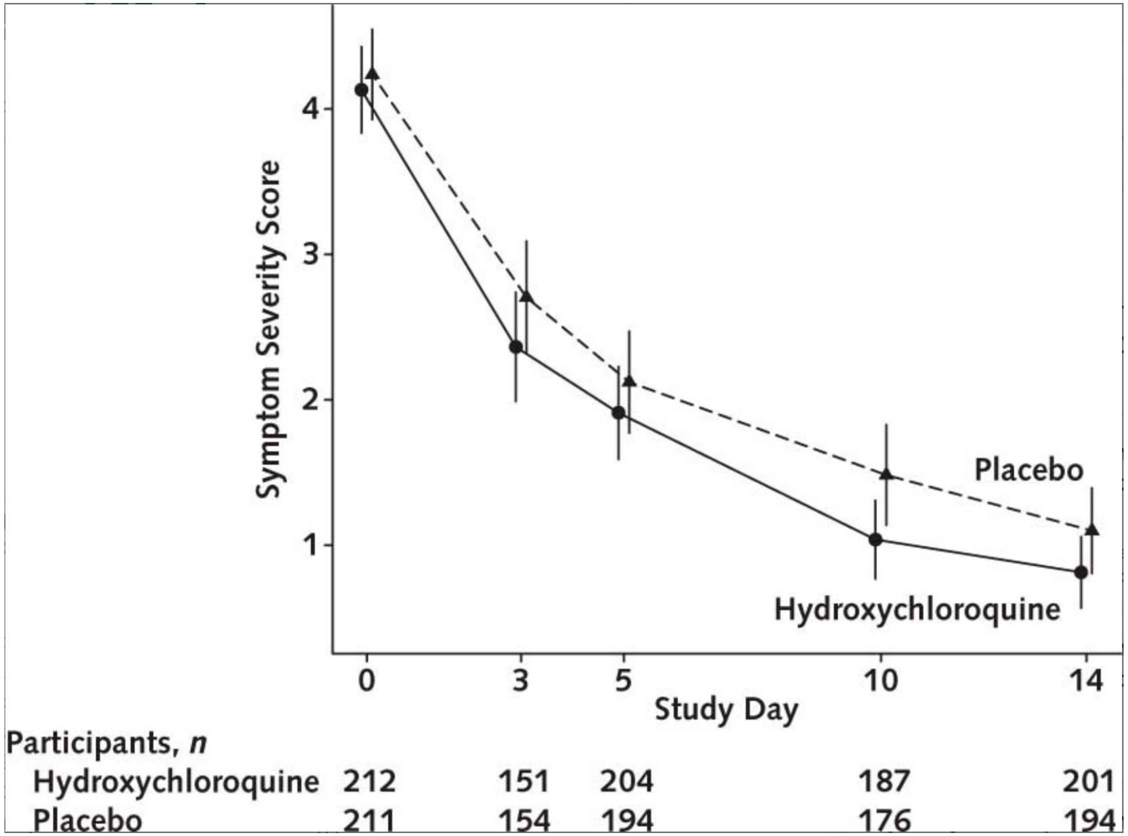
- It is used to control factors not under direct experimental interest.
- It can achieve sufficient control over confounding factors to deliver a meaninful comparison of the treatments studied.

Hydroxychloroquine in nonhospitalized adults with early COVID-19

A randomized trial

Skipper et al. (2020)

https://doi.org/10.7326/M20-4207



Randomized controlled trial is now often viewed as a gold standard in science

Randomized controlled trial is now often viewed as a gold standard in science

Question → Randomized controlled trial → Answer

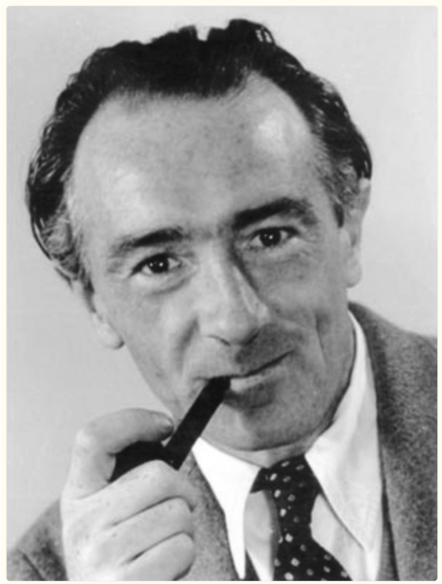
$A \xrightarrow{?} Y$

… Why it does not always work well

… Why it does not always work well

- Lack of power
- Heterogeneity of conclusions
- Restricted conditions

A solution: combining trials by meta-analysis



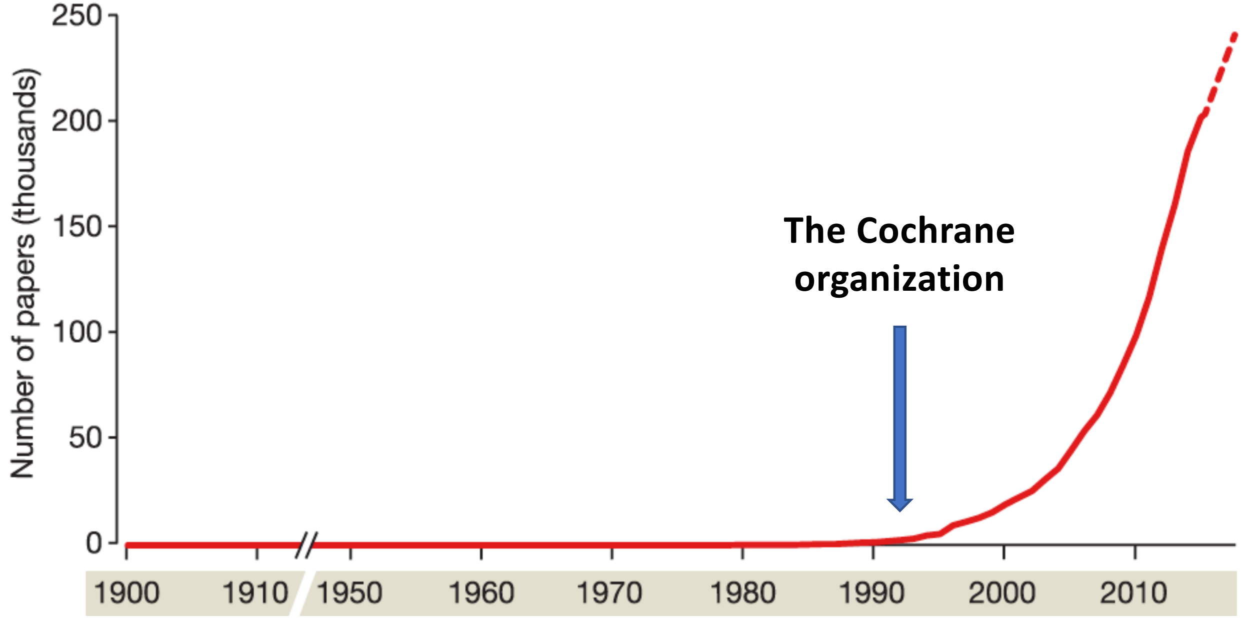
A.L. Cochrane (1909-1988)

Meta-analysis and the science of research synthesis

Jessica Gurevitch[1], Julia Koricheva[2], Shinichi Nakagawa[3,4] & Gavin Stewart[5]

March, 2018

https://www.nature.com/articles/nature25753



Cochrane

About us | Our evidence | Get involved | Learn | Products and services | For authors

Why RCT is not always possible

- Not always possible to apply the treatment A
- Not always easy to randomize
- Costly
- Limited sample sizes
- Unpopular

Non-RCT studies are common

- Non-randomized experiments
- Surveys
- Observational studies
- Expert-based studies

Potential-outcome framework

Table 1. Toy Data Set Illustrating the Potential Outcomes

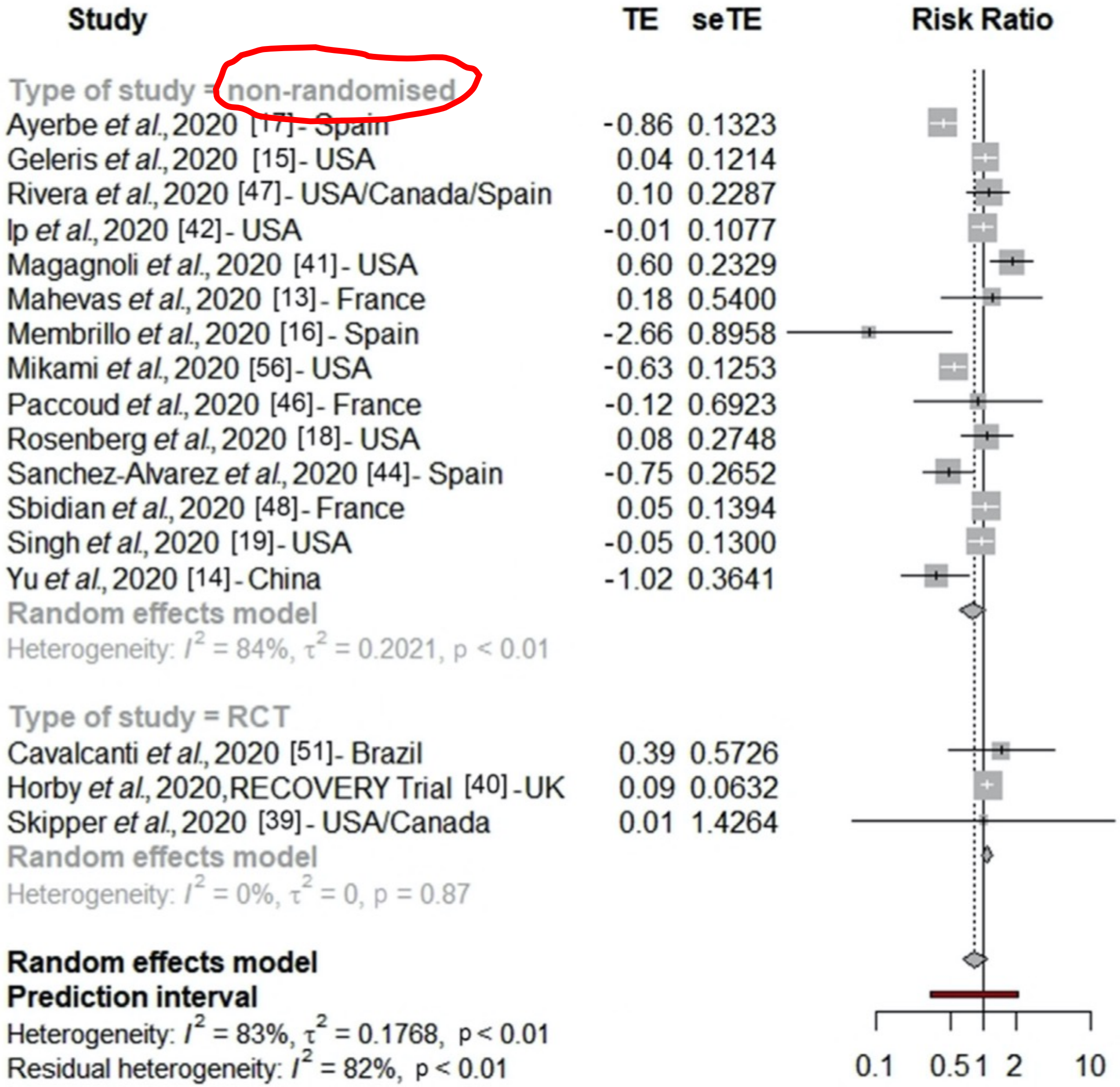
A	Y	Y^1	Y^0
1	1	1	?
1	0	0	?
0	1	?	1
0	0	?	0

https://doi.org/10.1177/25152459241236149

Systematic review

Effect of hydroxychloroquine with or without azithromycin on the mortality of coronavirus disease 2019 (COVID-19) patients: a systematic review and meta-analysis

Thibault Fiolet [1, 2, *], Anthony Guihur [3], Mathieu Edouard Rebeaud [3], Matthieu Mulot [4], Nathan Peiffer-Smadja [5, 6, 7], Yahya Mahamat-Saleh [1, 2]


Study
TE
seTE
Risk Ratio
Type of study = non-randomised
Ayerbe et al., 2020 [17] - Spain -0.86 0.1323
Geleris et al., 2020 [15] - USA 0.04 0.1214
Rivera et al., 2020 [47] - USA/Canada/Spain 0.10 0.2287
Ip et al., 2020 [42] - USA -0.01 0.1077
Magagnoli et al., 2020 [41] - USA 0.60 0.2329
Mahevas et al., 2020 [13] - France 0.18 0.5400
Membrillo et al., 2020 [16] - Spain -2.66 0.8958
Mikami et al., 2020 [56] - USA -0.63 0.1253
Paccoud et al., 2020 [46] - France -0.12 0.6923
Rosenberg et al., 2020 [18] - USA 0.08 0.2748
Sanchez-Alvarez et al., 2020 [44] - Spain -0.75 0.2652
Sbidian et al., 2020 [48] - France 0.05 0.1394
Singh et al., 2020 [19] - USA -0.05 0.1300
Yu et al., 2020 [14] - China -1.02 0.3641
Random effects model
Heterogeneity: I^2 = 84%, τ^2 = 0.2021, p < 0.01
Type of study = RCT
Cavalcanti et al., 2020 [51] - Brazil 0.39 0.5726
Horby et al., 2020, RECOVERY Trial [40] - UK 0.09 0.0632
Skipper et al., 2020 [39] - USA/Canada 0.01 1.4264
Random effects model
Heterogeneity: I^2 = 0%, τ^2 = 0, p = 0.87
Random effects model
Prediction interval
Heterogeneity: I^2 = 83%, τ^2 = 0.1768, p < 0.01
Residual heterogeneity: I^2 = 82%, p < 0.01
0.1
0.5
1
2
10

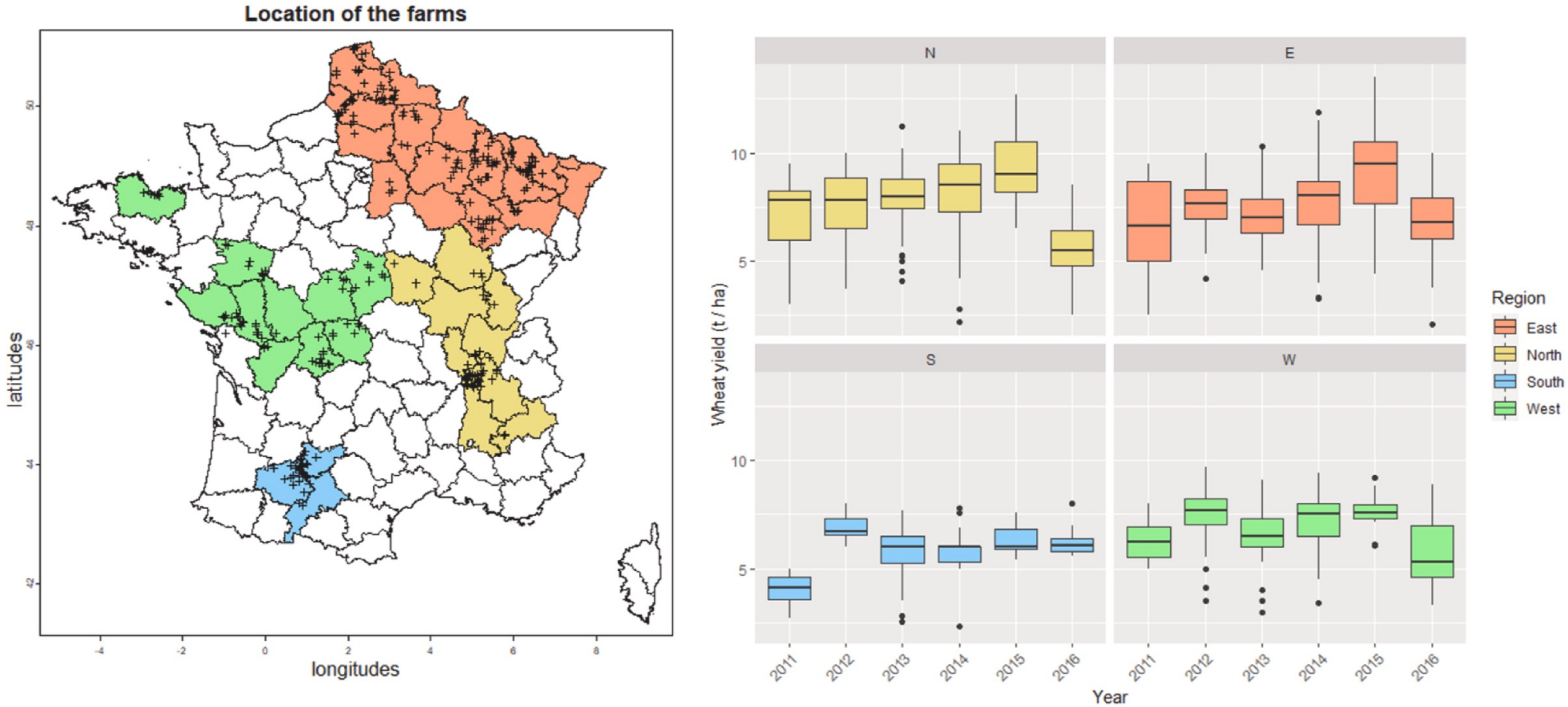


Fig. 1. Left: Location of the farms (crosses) within the four regions defined for this study (N, E, S, W). Right: Boxplots of wheat yields studied in this work, from 2011 to 2016 in the same regions.

https://doi.org/10.1016/j.eja.2024.127254

Effect of drought on yields in France?

A=1 : drought occurrence

A=0 : absence of drought

Y^1 : yield value in a given plot if A=1

Y^0 : yield value in a given plot if A=0

Average drought effect = $E(Y^1) - E(Y^0)$

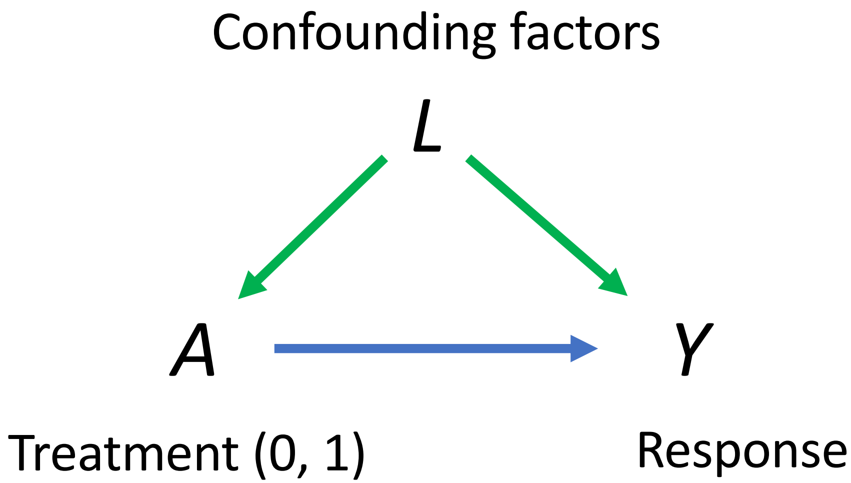
Plot	Drought	Yield after drought Y^1	Yield without drought Y^0	Observed yields
1	No	4	8	8
2	No	5	6	6
3	No	9	10	10
4	**Yes**	6	9	6
5	No	7	7	7
6	No	2	5	5
7	**Yes**	6	8	6
8	No	3	7	7

How can we estimate the average drought effect $E(Y^1) - E(Y^0)$ from observed yield data only, without RCT?

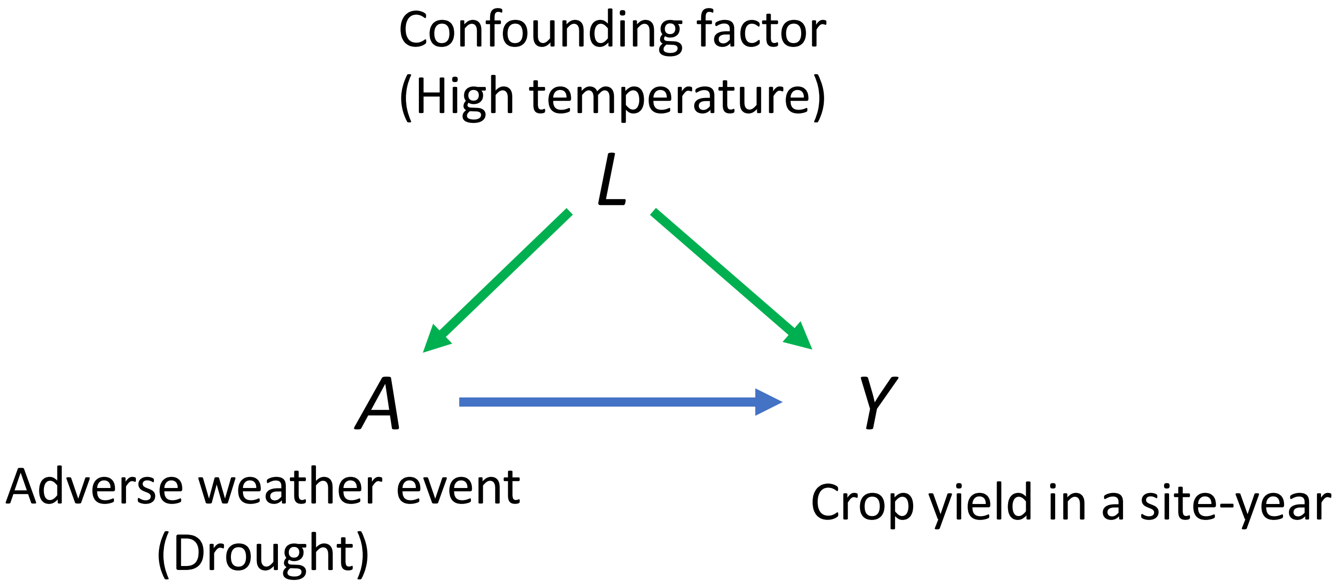
Plot	Drought	Yield after drought Y^1	Yield without drought Y^0	Observed yields
1	No	4	8	8
2	No	5	6	6
3	No	9	10	10
4	**Yes**	6	9	6
5	No	7	7	7
6	No	2	5	5
7	**Yes**	6	8	6
8	No	3	7	7

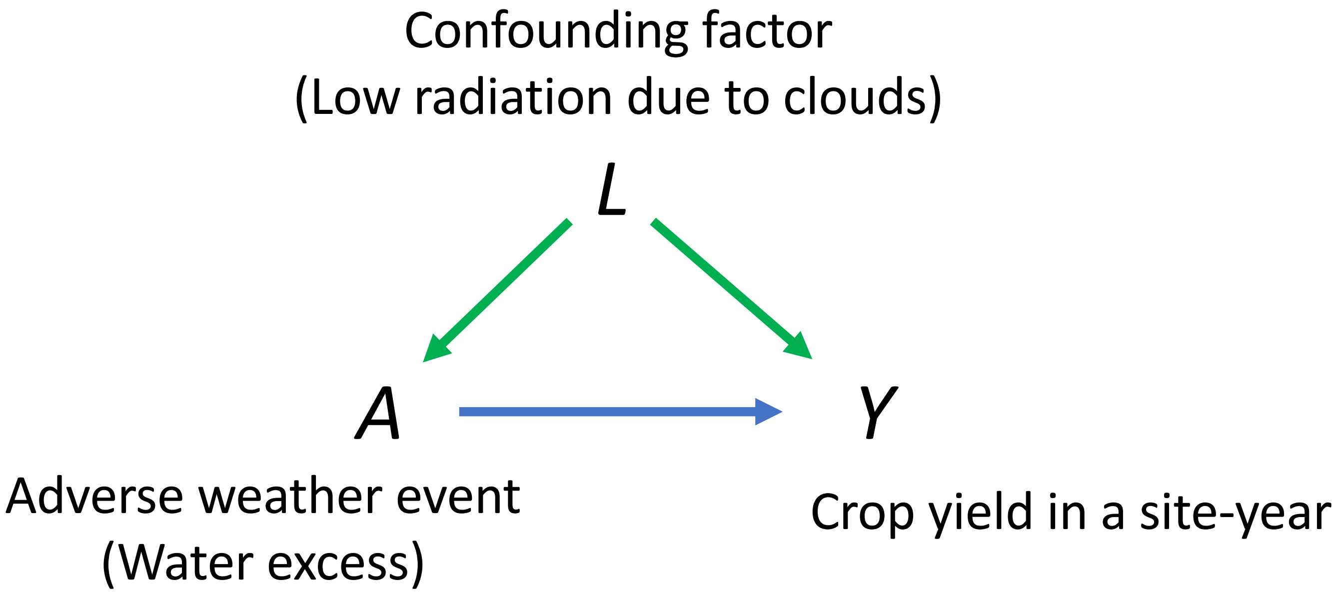
Outline

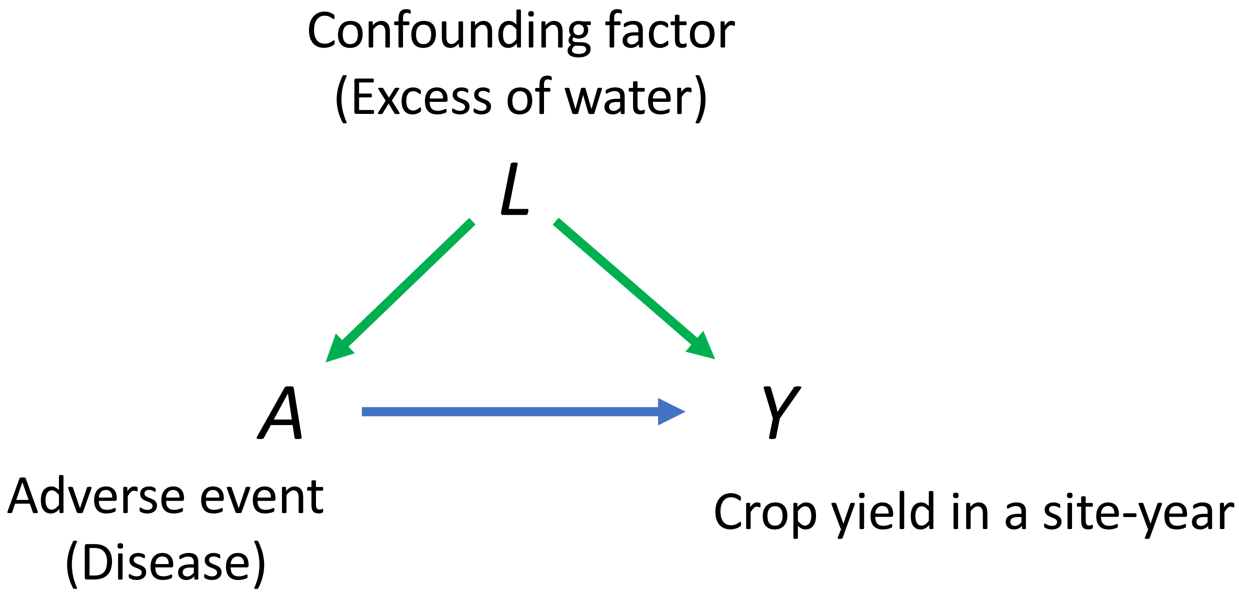
- Objective
- Methods
- Application 1: Maize and Sunflower
- Application 2: Vine
- Conclusions



What is the effect of A on Y?


Confounding factor
(High temperature)
L
A
Y
Adverse weather event
(Drought)
Crop yield in a site-year


Confounding factor
(Low radiation due to clouds)
L
A
Y
Adverse weather event
(Water excess)
Crop yield in a site-year


Confounding factor
(Excess of water)
L
A
Y
Adverse event
(Disease)
Crop yield in a site-year

Causal inference methods to deal with confounding factors to estimate the effect of A on Y

- Inverse probability weighting with propensity scores
- Matching with propensity scores
- Standardization (g-computation)
- Double robust inference

Causal inference methods to deal with confounding factors to estimate the effect of *A* on *Y*

- Inverse probability weighting with propensity scores
- Matching with propensity scores
- Standardization (*g*-computation)
- Double robust inference

Model computing the probability of A (propensity score) as a function of L

Causal inference methods to deal with confounding factors to estimate the effect of *A* on *Y*

- Inverse probability weighting with propensity scores
- Matching with propensity scores
- Standardization (*g*-computation)
- Double robust inference

Model computing Y (outcome) as a function of L and A

Provide unbiased estimation of $E(Y^{a=1}) - E(Y^{a=0})$ if

- Exchangeability
- Positivity
- Consistency
- Noninterference

Inverse probability weighting

$$E[Y^a] = E\left[\frac{I(A=a)Y}{P(A=a|L)}\right]$$

Idea: Compute the mean of *Y* weighted by the inverse probability of *A*=*a*

Implementation

$$\hat{E}\left[\frac{I(Drought)Y}{P(Drought|L)}\right] = \frac{1}{n}\sum_{i=1}^{n}\frac{Y_i\, I(A_i = Drought)}{\hat{P}(A_i = Drought|L_i)}$$

Develop a model $\hat{P}(A_i = Drought|L_i)$: « Propensity score »

- Logistic regression (glm)
- Machine learning for classification (random forest, gradient boosting etc.)

A	L_1	...	L_K	*Y*
0 (no drought)	Irrigated		Temperature=15	9.2
0 (no drought)	Rainfed		Temperature=21	7.2
1 (drought)	Irrigated		Temperature=11	8.5
0 (no drought)	Irrigated		Temperature=24	7.9
1 (drought)	Rainfed		Temperature=14	7.1
…	…	…	…	…
0 (no drought)	Rainfed		Temperature=19	6.8

$$\hat{P}(A = 1|L) = f(L)$$

glm(A~L1+L2+…+LK, family=binomial)
randomForest(A~L1+L2+…+LK)
Xgboost, neuralnet etc.

Implementation

Run the model over all data and compute:

$$\hat{E}\left[\frac{I(Drought)Y}{P(Drought|L)}\right] - \hat{E}\left[\frac{I(No\ drought)Y}{1 - P(Drought|L)}\right]$$

Implementation

Run the model over all data and compute:

$$\hat{E}\left[\frac{I(Drought)Y}{P(Drought|L)}\right] - \hat{E}\left[\frac{I(No\ drought)Y}{1 - P(Drought|L)}\right]$$

- The probabilities of *drought* and *no drought* should be non-zero!
- Numerical issues and instability when these probabilities are too extreme.
- Various tricks +- efficient to deal with this issue.
- Confidence intervals can be computed by bootstrap.

Variant: Matching

- Compute $P(Drought|L)$ for all data
- Create pairs of values of *Y* based on the calculated probabilities
  - Select an observed value Y_d with drought and $P(Drought|L)=P_d$
  - Select an observed value Y_{nd} without drought and $P(Drought|L)=P_{nd}$
  - Match the two values (Y_d, Y_{nd}) if P_d and P_{nd} are « similar »
  - Repeat the procedure for all the observed Y
- Compute the mean difference of *Y* based on the pairs
- Confidence intervals can be computed by bootstrap

Variant: Matching

- Compute $P(Drought|L)$ for all data

- Create pairs of values of Y based on the calculated probabilities
  - Select an observed value Y_d with drought and $P(Drought|L)=P_d$
  - Select an observed value Y_{nd} without drought and $P(Drought|L)=P_{nd}$
  - Match the two values (Y_d, Y_{nd}) if P_d and P_{nd} are « similar »
  - Repeat the procedure for all the observed Y

Many different ways to define « similar » !

- Compute the mean difference of Y based on the pairs

- Confidence intervals can be computed by bootstrap

Standardization

$$\mathrm{E}\left[Y^{a}\right] = \sum_{l} \mathrm{E}\left[Y \mid A = a, L = l\right] \Pr\left[L = l\right]$$

Standardization

$$E[Y^{Drought}] = E[Y|Drought, Irrigated]P(Irrigated) \\ + E[Y|Drought, Rainfed]P(Rainfed)$$

A=Drought
L=Irrigated/Rainfed

Implementation

$$E[Y^{Drought}] = E[Y|Drought, L = Irrigated]\, P(L = Irrigated) \\ + E[Y|Drought, L = Rainfed]\, P(L = Rainfed)$$

Require an « outcome model » $g(Drought, No\ drought, L)$ able to compute $\hat{E}[Y|Drought, L]$

- Linear regression
- GAM
- Machine learning (RF, xgboost, neuralnet etc).

Implementation

$$E[Y^{Drought}] = E[Y|Drought, L = Irrigated]\, P(L = Irrigated) \\ + E[Y|Drought, L = Rainfed]\, P(L = Rainfed)$$

Step 1: Develop a model $g(Drought, No\ drought, L)$ computing $\hat{E}[Y|Drought, L]$

- Linear regression
- GAM
- Machine learning (RF, xgboost, neuralnet etc).

Step 2: Run the model two times over all data, with *Drought* and *No droughts*, successively

Step 3: Compute the average difference

$$\frac{1}{n}\sum_{i=1}^{n} g(Drought, L_i) - \frac{1}{n}\sum_{i=1}^{n} g(No\ drought, L_i)$$

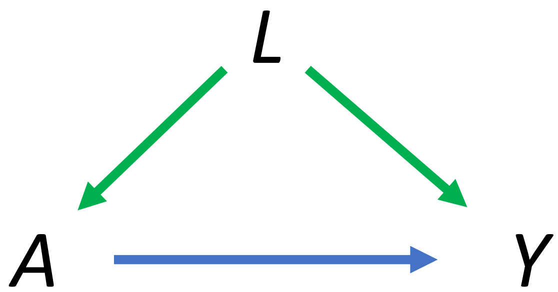
A	L_1	…	L_K	Y
0 (no drought)	Irrigated		Temperature=15	9.2
0 (no drought)	Rainfed		Temperature=21	7.2
1 (drought)	Irrigated		Temperature=11	8.5
0 (no drought)	Irrigated		Temperature=24	7.9
1 (drought)	Rainfed		Temperature=14	7.1
…	…	…	…	…
0 (no drought)	Rainfed		Temperature=19	6.8

$$\hat{E}[Y|A,L] = g(A,L)$$

lm(Y~L1+L2+…+LK)
randomForest(Y~L1+L2+…+LK)
Xgboost, neuralnet etc.

A	L_1	…	L_K	Y	$g(A=0,L)$	$g(A=1,L)$
0 (no drought)	Irrigated		Temperature=15	9.2	9.6	9.1
0 (no drought)	Rainfed		Temperature=21	7.2	8.1	7.9
1 (drought)	Irrigated		Temperature=11	8.5	8.0	7.2
0 (no drought)	Irrigated		Temperature=24	7.9	7.2	6.5
1 (drought)	Rainfed		Temperature=14	7.1	7.9	6.1
…	…	…	…	…	…	…
0 (no drought)	Rainfed		Temperature=19	6.8	5.5	4.1
					↓ Mean yield without drought	↓ Mean yield with drought

Double robust



- Combine Inverse probability weighting and standardization
- Rely on two models

$$\hat{P}(A|L)$$

$$\hat{E}[Y|A, L] = g(A, L)$$

- Unbiased if one of the two models is unbiased

Double robust

$$\hat{E}[Y^{a=1}] = \frac{1}{n}\sum_{i=1}^{n}\left[g(A=1, L_i) + \frac{A_i}{P(A=1|L_i)}\left(Y_i - g(A=1, L_i)\right)\right]$$

Double robust

$$\hat{E}[Y^{a=1}] = \frac{1}{n}\sum_{i=1}^{n}\left[g(A=1,L_i) + \frac{A_i}{P(A=1|L_i)}\left(Y_i - g(A=1,L_i)\right)\right]$$

Predicted effect of *A*=1 on Y

Error of prediction of *Y*

Probability of *A*=1 estimated
as a function of *L*

Double robust

$$\hat{E}[Y^{a=1}] = \frac{1}{n}\sum_{i=1}^{n}\left[g(A=1,L_i) + \frac{A_i}{P(A=1|L_i)}\big(Y_i - g(A=1,L_i)\big)\right]$$

$$\hat{E}[Y^{a=0}] = \frac{1}{n}\sum_{i=1}^{n}\left[g(A=0,L_i) + \frac{1-A_i}{P(A=0|L_i)}\big(Y_i - g(A=0,L_i)\big)\right]$$

A	L_1	…	L_K	Y	$g(A=0,L)$	$g(A=1,L)$
0 (no drought)	Irrigated		Temperature=15	9.2	9.6	9.1
0 (no drought)	Rainfed		Temperature=21	7.2	8.1	7.9
1 (drought)	Irrigated		Temperature=11	8.5	8.0	7.2
0 (no drought)	Irrigated		Temperature=24	7.9	7.2	6.5
1 (drought)	Rainfed		Temperature=14	7.1	7.9	6.1
…	…	…	…	…	…	…
0 (no drought)	Rainfed		Temperature=19	6.8	5.5	4.1

A	***L_1***	***…***	***L_K***	***Y***	$g(A=0,L)$	$g(A=1,L)$	$P(A=1\|L)$
0 (no drought)	Irrigated		Temperature=15	9.2	9.6	9.1	0.4
0 (no drought)	Rainfed		Temperature=21	7.2	8.1	7.9	0.7
1 (drought)	Irrigated		Temperature=11	8.5	8.0	7.2	0.2
0 (no drought)	Irrigated		Temperature=24	7.9	7.2	6.5	0.8
1 (drought)	Rainfed		Temperature=14	7.1	7.9	6.1	0.3
…	…	…	…	…	…	…	…
0 (no drought)	Rainfed		Temperature=19	6.8	5.5	4.1	0.6

A	***L_1***	**…**	***L_K***	***Y***	$g(A=0,L)$	$g(A=1,L)$	$P(A{=}1\|L)$
0 (no drought)	Irrigated		Temperature=15	9.2	9.6	9.1	0.4
0 (no drought)	Rainfed		Temperature=21	7.2	8.1	7.9	0.7
1 (drought)	Irrigated		Temperature=11	8.5	8.0	7.2	0.2
0 (no drought)	Irrigated		Temperature=24	7.9	7.2	6.5	0.8
1 (drought)	Rainfed		Temperature=14	7.1	7.9	6.1	0.3
…	…	…	…	…	…	…	…
0 (no drought)	Rainfed		Temperature=19	6.8	5.5	4.1	0.6

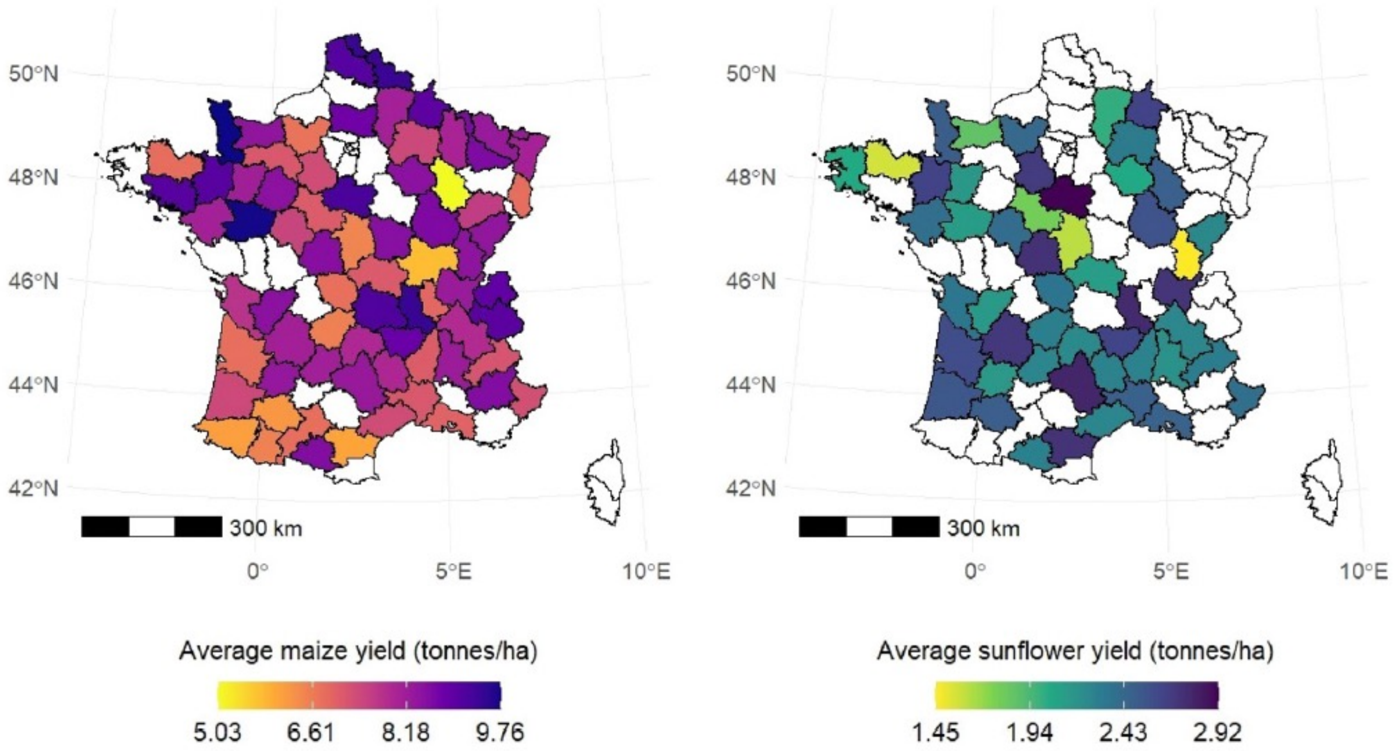
$$\hat{E}[Y^{a=1}] = \frac{1}{n}\sum_{i=1}^{n}\left[g(A=1,L_i) + \frac{A_i}{P(A=1|L_i)}\left(Y_i - g(A=1,L_i)\right)\right]$$

$$\hat{E}[Y^{a=0}] = \frac{1}{n}\sum_{i=1}^{n}\left[g(A=0,L_i) + \frac{1-A_i}{P(A=0|L_i)}\left(Y_i - g(A=0,L_i)\right)\right]$$

Outline

- Objective
- Methods
- Application 1: Maize and Sunflower
- Application 2: Vine
- Conclusions

Causal analysis methods for estimating the impact of drought and cold events on crop yields in France.

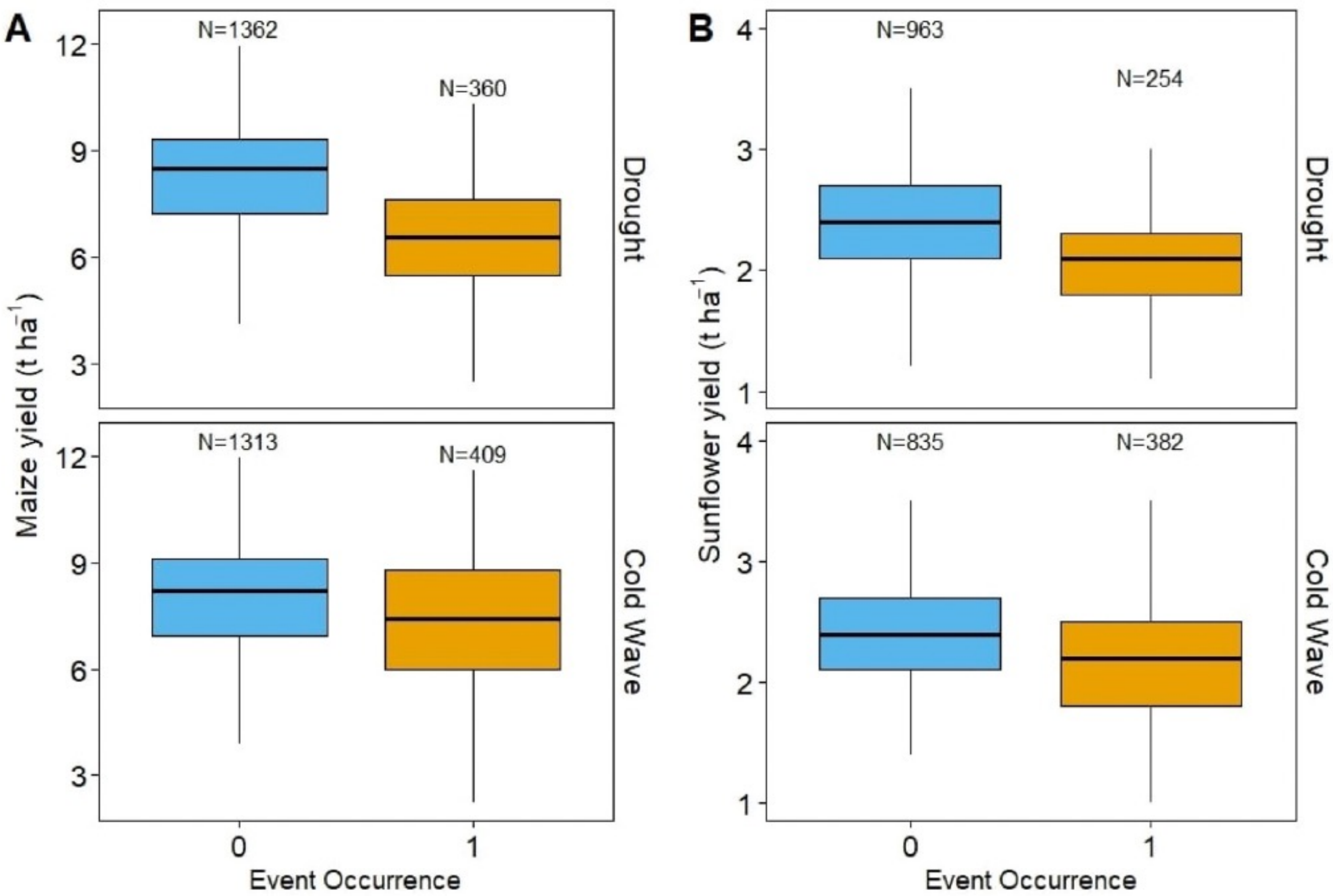


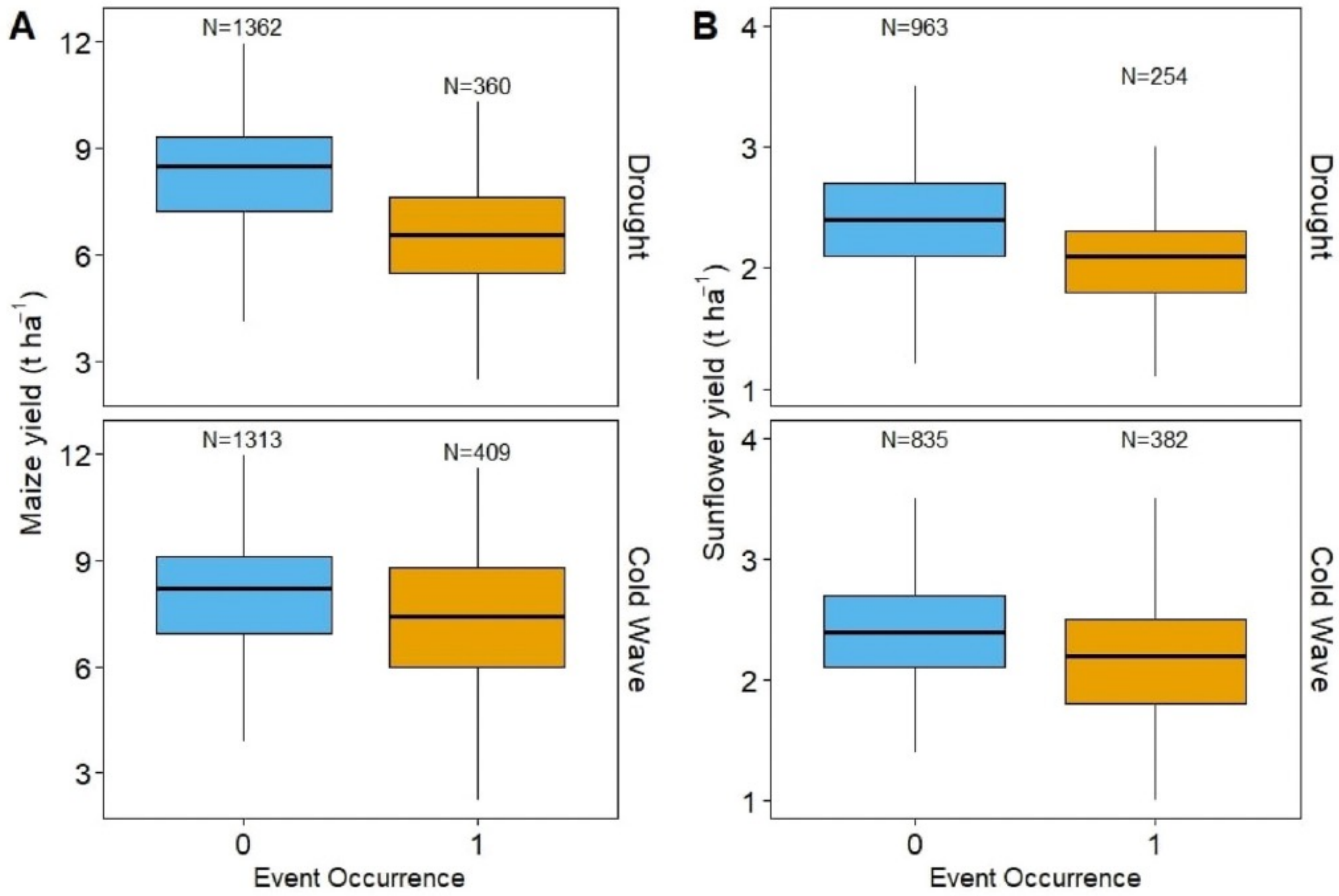
Variable	Description	Unit
NDT0	Monthly number of days with minimum daily temperatures below 0°C	days
NDT30	Monthly number of days with maximum daily temperatures above 30°C	days
P	Monthly average of liquid precipitation per day	mm
RG	Monthly average of visible radiation per day	J/cm^2
ETP	Monthly average potential evapotranspiration (Penman-Monteith) per day	mm
Tmax	Monthly average daily maximum temperature	ºC
Tmin	Monthly average daily minimum temperature	ºC
DE	Drought event in July: 1 if $P_{\text{July}} \leq q_{0.2}$, 0 otherwise	binary
CE	Cold wave event in April: 1 if $NDT0_{\text{April}} \geq q_{0.8}$, 0 otherwise	binary

Confounding variables

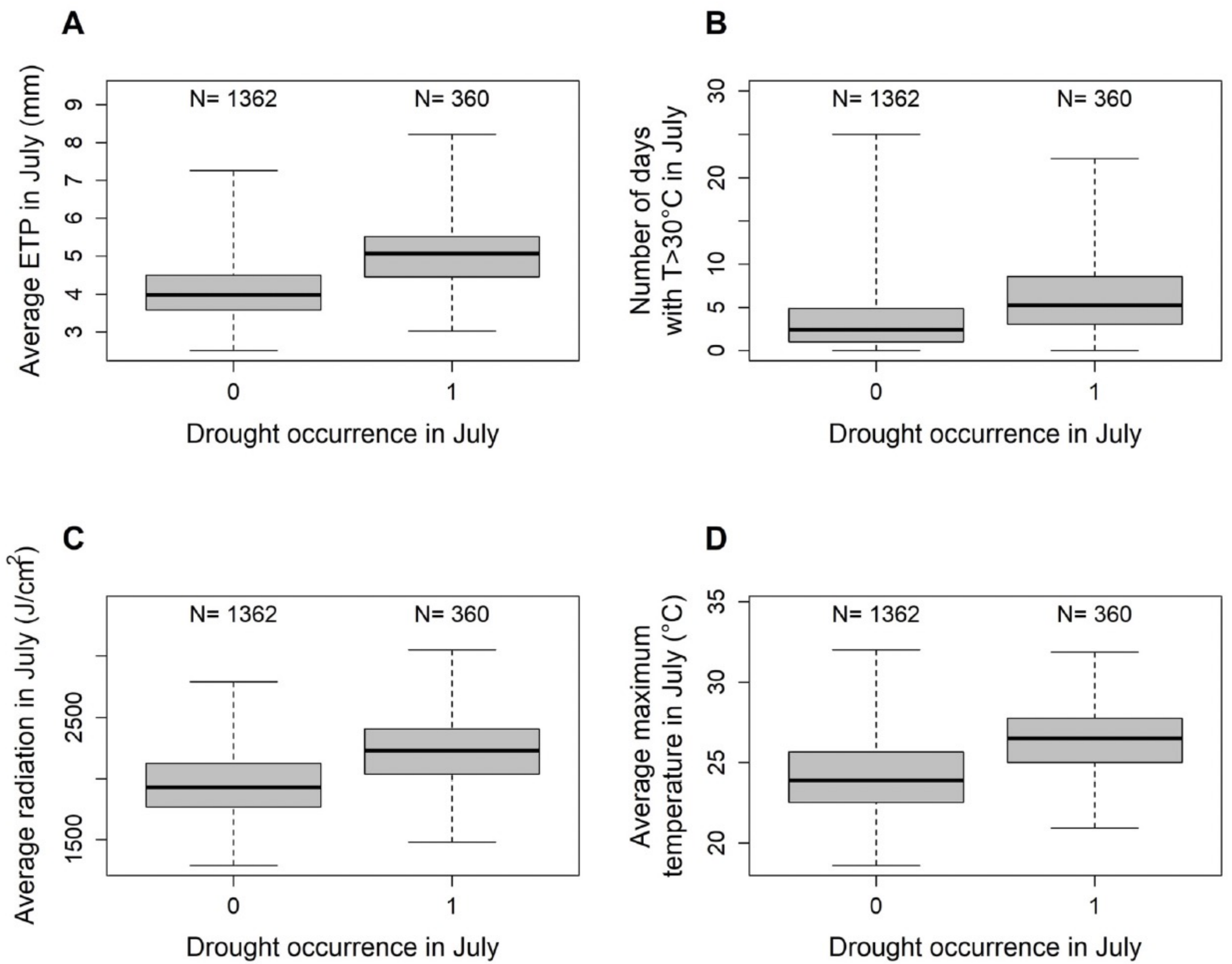
Variable	Description	Unit
NDT0	Monthly number of days with minimum daily temperatures below 0°C	days
NDT30	Monthly number of days with maximum daily temperatures above 30°C	days
P	Monthly average of liquid precipitation per day	mm
RG	Monthly average of visible radiation per day	J/cm^2
ETP	Monthly average potential evapotranspiration (Penman-Monteith) per day	mm
Tmax	Monthly average daily maximum temperature	ºC
Tmin	Monthly average daily minimum temperature	ºC
DE	Drought event in July: 1 if $P_{\text{July}} \leq q_{0.2}$, 0 otherwise	binary
CE	Cold wave event in April: 1 if $NDT0_{\text{April}} \geq q_{0.8}$, 0 otherwise	binary

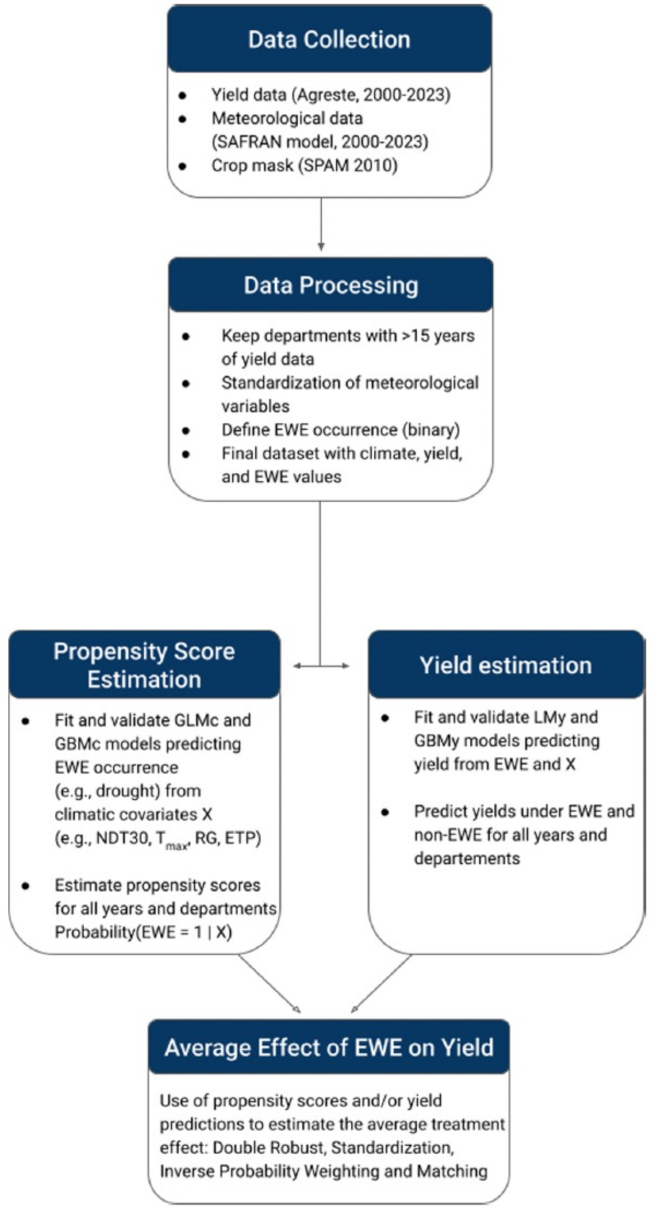
Treatment variables


A
B
N=1362
N=360
N=1313
N=409
N=963
N=254
N=835
N=382
Maize yield (t ha⁻¹)
Sunflower yield (t ha⁻¹)
Drought
Cold Wave
Event Occurrence
12
9
6
3
4
2
1
0

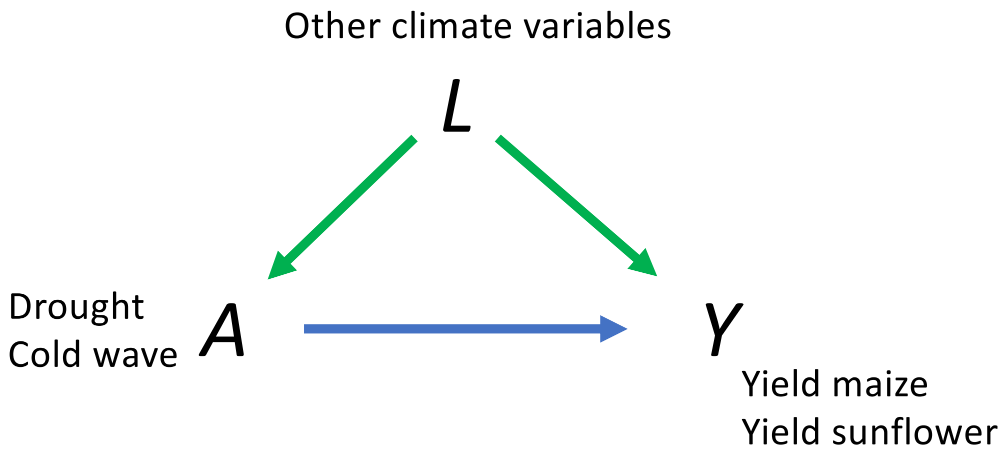


Naive comparison can be wrong because of confounding!


A
B
C
D
N= 1362
N= 360
Average ETP in July (mm)
Number of days with T>30°C in July
Average radiation in July (J/cm2)
Average maximum temperature in July (°C)
Drought occurrence in July
0
1


Data Collection
Yield data (Agreste, 2000-2023)
Meteorological data (SAFRAN model, 2000-2023)
Crop mask (SPAM 2010)
Data Processing
Keep departments with >15 years of yield data
Standardization of meteorological variables
Define EWE occurrence (binary)
Final dataset with climate, yield, and EWE values
Propensity Score Estimation
Fit and validate GLMc and GBMc models predicting EWE occurrence (e.g., drought) from climatic covariates X (e.g., NDT30, T_{max}, RG, ETP)
Estimate propensity scores for all years and departments Probability(EWE = 1 | X)
Yield estimation
Fit and validate LMy and GBMy models predicting yield from EWE and X
Predict yields under EWE and non-EWE for all years and departements
Average Effect of EWE on Yield
Use of propensity scores and/or yield predictions to estimate the average treatment effect: Double Robust, Standardization, Inverse Probability Weighting and Matching

Double robust



- Combine Inverse probability weighting and standardization
- Rely on two models

$$\hat{P}(A|L) = f(L)$$ Model predicting drought and cold wave occurrences

$$\hat{E}[Y|A,L] = g(A,L)$$ Model predicting crop yield

- Unbiased if one of the two models is unbiased

Models predicting crop yields

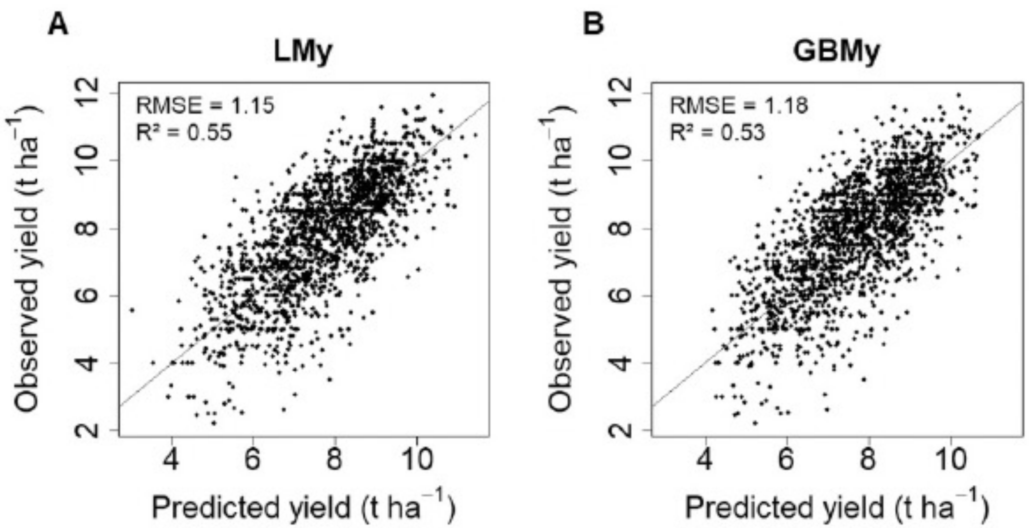


Fig. 6. Performances of models predicting the effect of drought on maize yields. A: Scatter plot showing the relationship between predicted and observed yields for the Linear Mixed-effects Model (LMy). The diagonal blue line represents the 1:1 line. **B:** Scatter plot showing the relationship between predicted and observed yields for the Gradient Boosting Machine Model (GBMy). The diagonal blue line represents the 1:1 line. RMSE and R^2 were obtained using a year-by-year cross-validation.

Model predicting drought occurrences

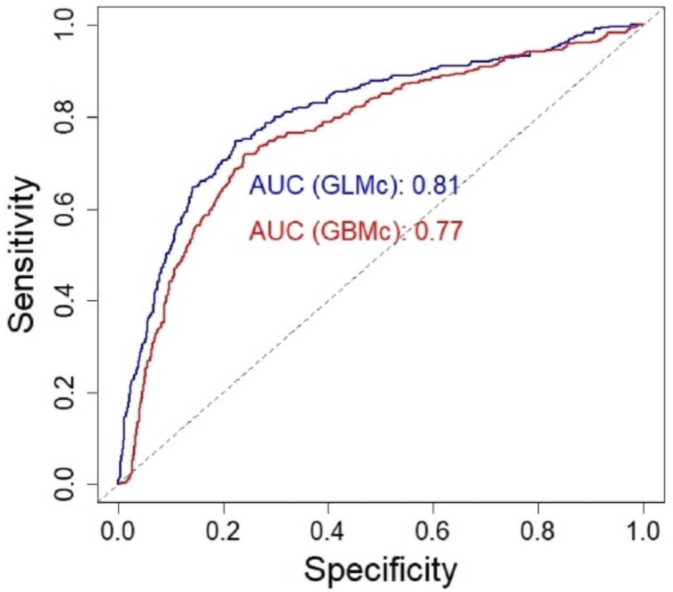


Fig. 7. Receiver Operating Characteristic (ROC) curves comparing the classification performance of the Generalized Linear Mixed-effects Model (GLMc, blue) and the Gradient Boosting Model (GBMc, red) for drought occurrence in maize. The Area Under the Curve (AUC) values indicate the model's ability to distinguish between drought and non-drought conditions (GLMc: 0.81, GBMc: 0.77). The dashed gray diagonal line represents random classification as a baseline. Results were obtained using year-by-year cross-validation.

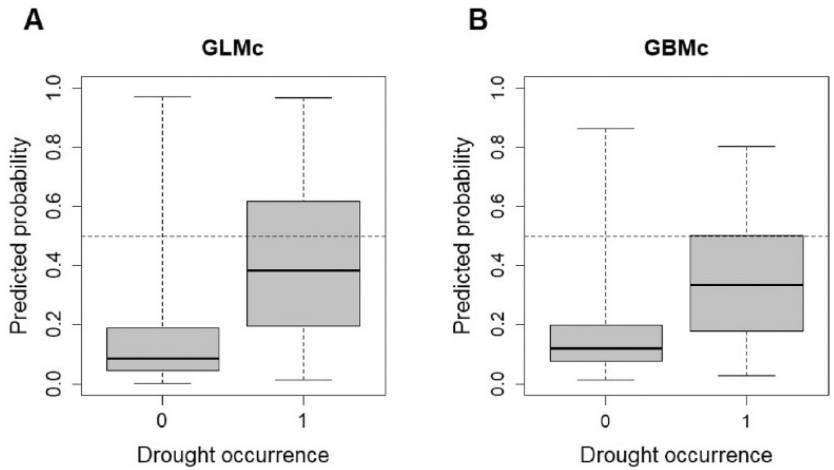
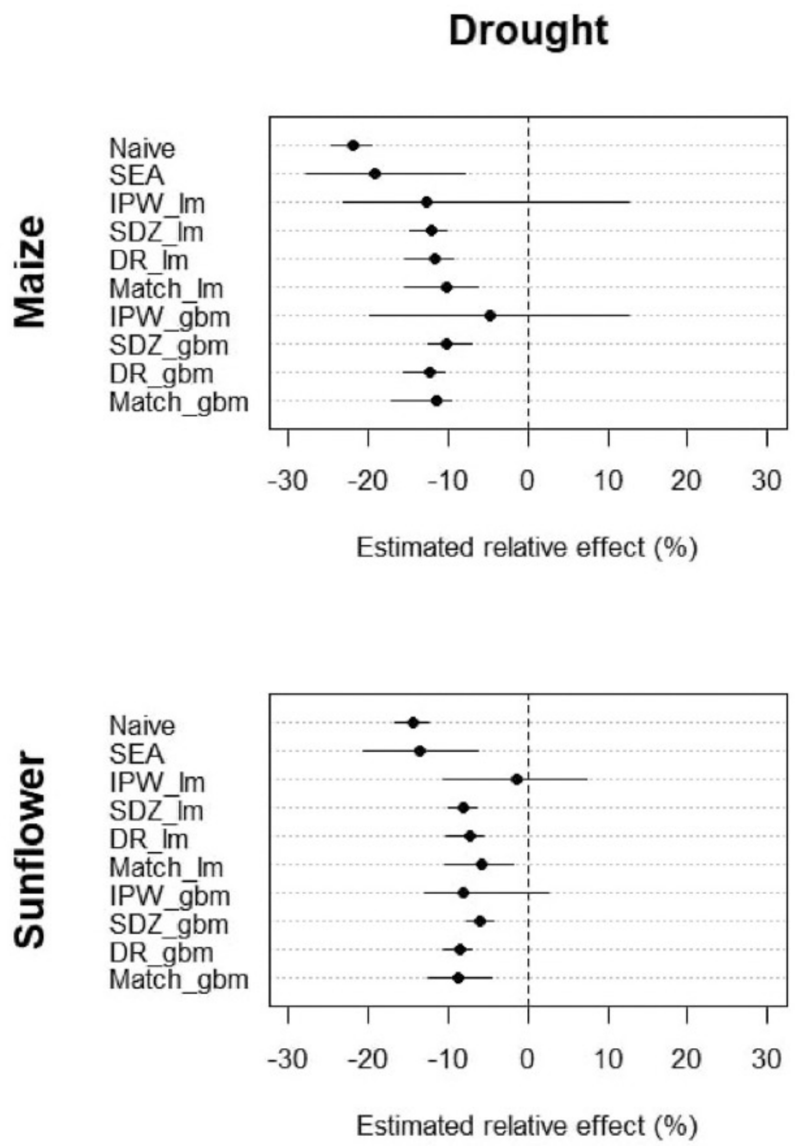
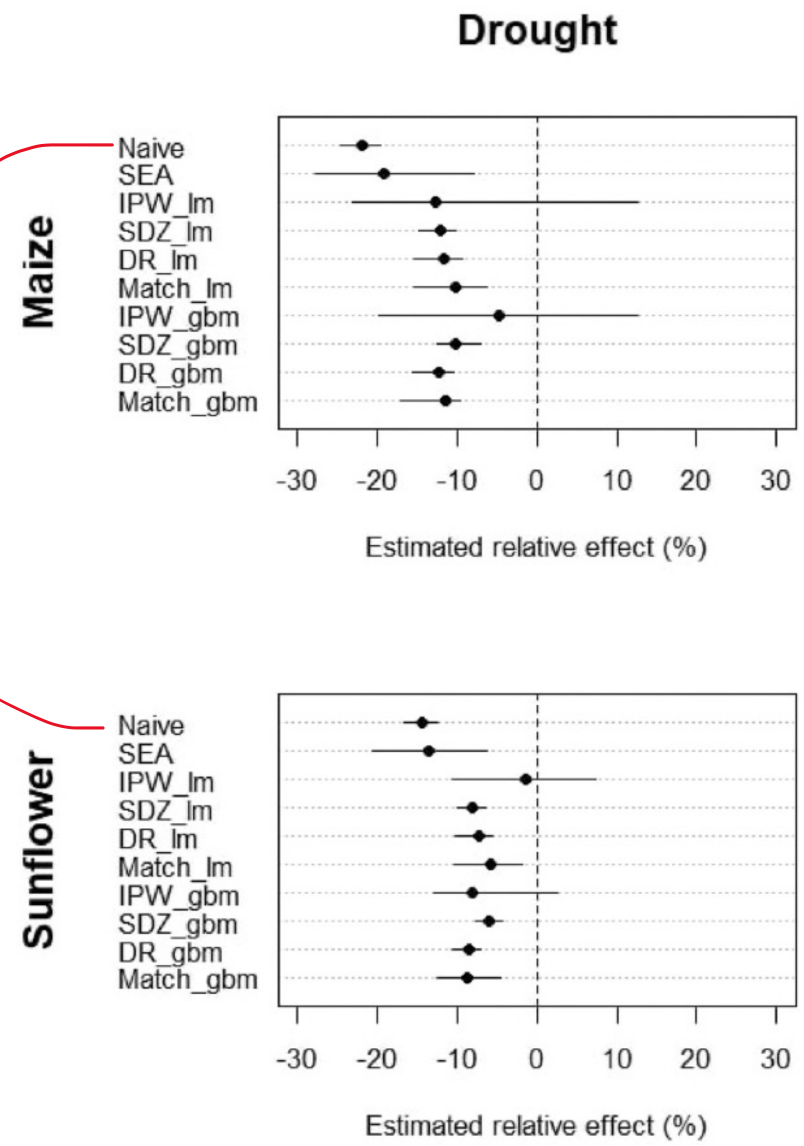


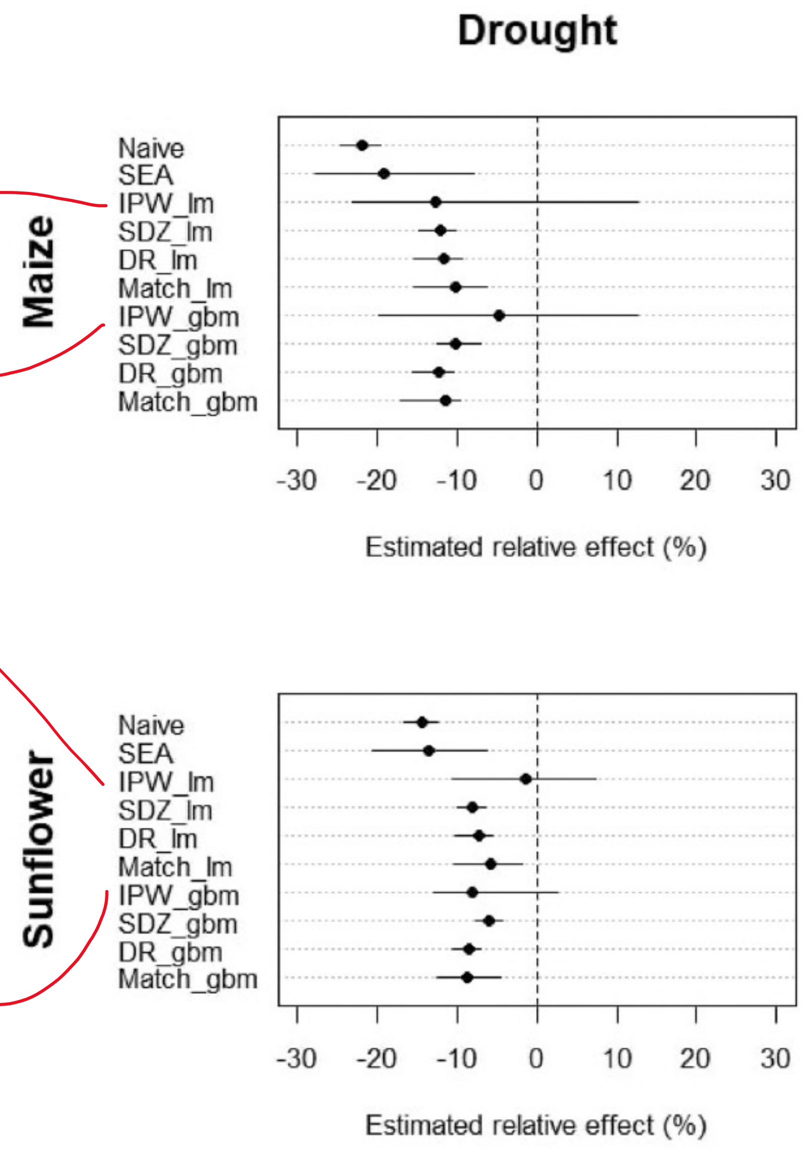
Fig. 8. Probabilities of drought occurrence in maize predicted by the Generalized Linear Mixed-effects Model (GLMc) and Gradient Boosting Model (GBMc). (A:) Boxplots describing the probabilities predicted by GLMc in case of actual drought absence (class 0) and drought occurrence (class 1). **(B:)** Same type of boxplots for GBMc. Horizontal dashed lines at a threshold of 0.5 (50 % chance of drought). The gray boxes represent the interquartile range (IQR), with the thick horizontal line indicating the median predicted probability. Whiskers extend to the minimum and maximum predicted probabilities within 1.5 times the IQR.


Drought
Maize
Naive
SEA
IPW_lm
SDZ_lm
DR_lm
Match_lm
IPW_gbm
SDZ_gbm
DR_gbm
Match_gbm
-30
-20
-10
0
10
20
30
Estimated relative effect (%)
Sunflower
Naive
SEA
IPW_lm
SDZ_lm
DR_lm
Match_lm
IPW_gbm
SDZ_gbm
DR_gbm
Match_gbm
-30
-20
-10
0
10
20
30
Estimated relative effect (%)

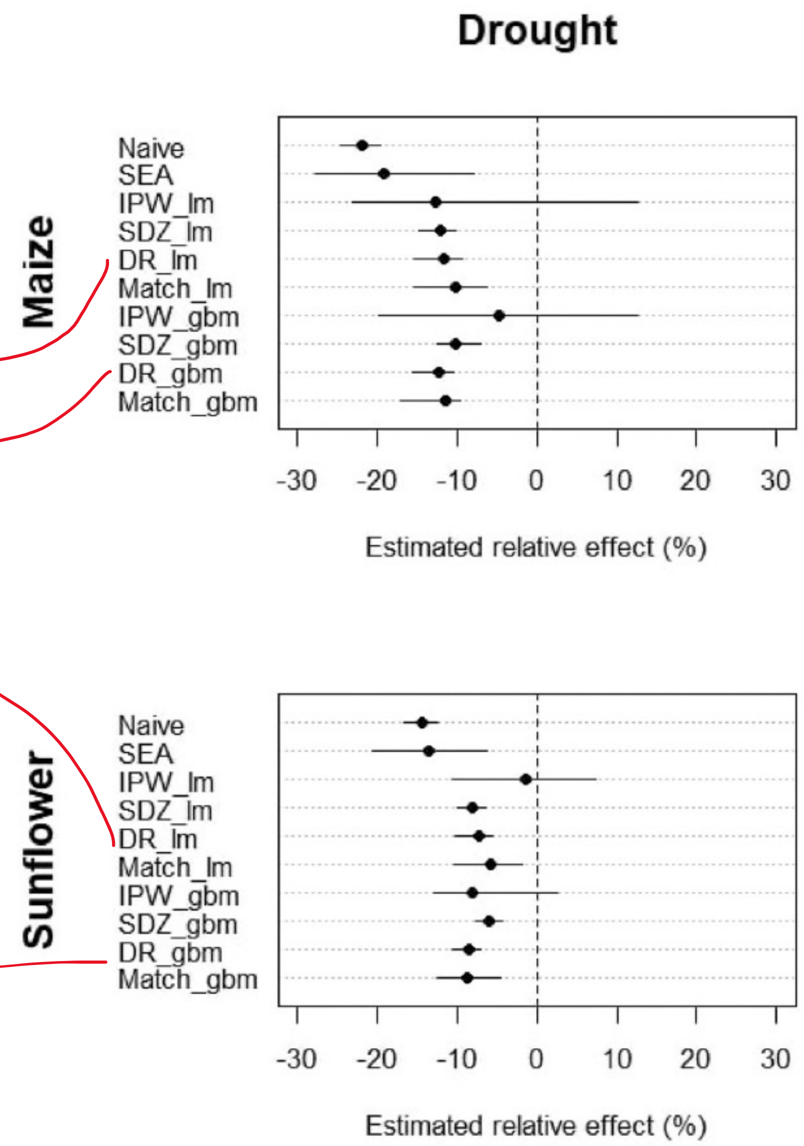
Naive estimates (« direct » comparison of yield mean values) led to stronger yield loss values



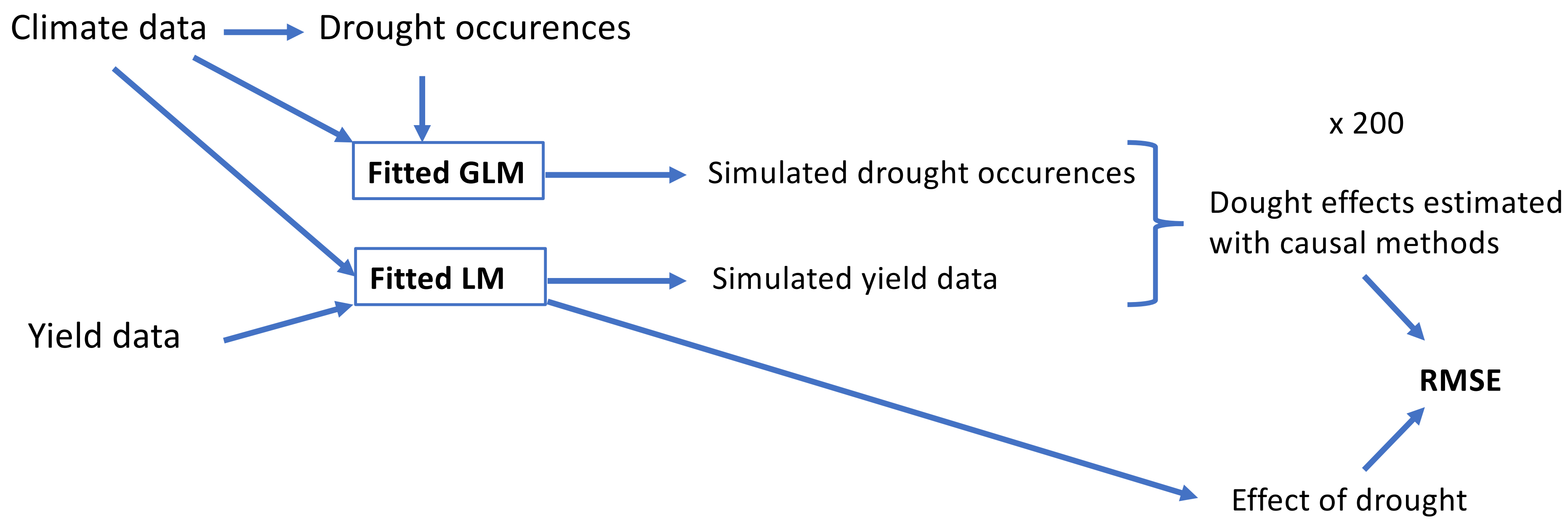
Inverse probablity weighting led to very large confidence intervals (instability)



Double robust provided intermediate estimated and narrow confidence intervals (no instability).



Simulations



Results of simulations for estimating the drought effect on yield

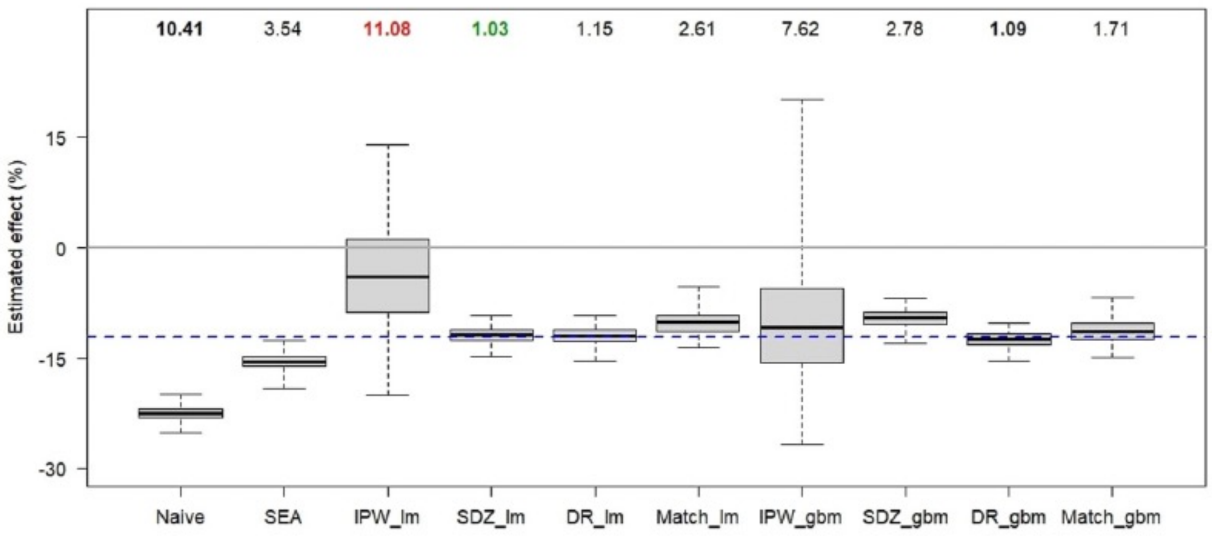


Fig. 10. Results of the data simulation study for estimating yield effects of drought on maize. Boxplots show the distribution of effects from 200 simulated datasets for each method and model type. The blue dashed line represents the true effect (-12.16 %), while the gray horizontal line at 0 differentiates positive and negative estimates. RMSE values at the top quantify method accuracy; the lowest is highlighted in green (SDZ_lm: 1.03), the highest in red (IPW_lm: 11.08), and bold values indicate second-best (DR_gbm: 1.09) and second-worst (naive: 10.41) RMSEs.

Results of simulations for estimating the drought effect on yield

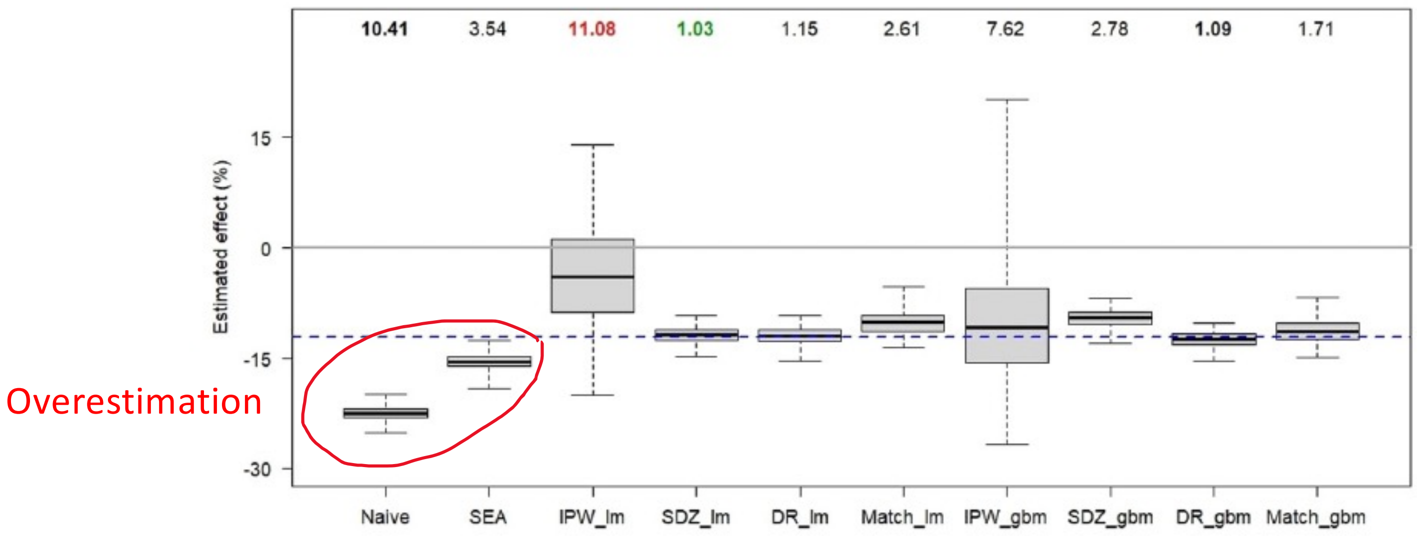


Fig. 10. Results of the data simulation study for estimating yield effects of drought on maize. Boxplots show the distribution of effects from 200 simulated datasets for each method and model type. The blue dashed line represents the true effect (-12.16 %), while the gray horizontal line at 0 differentiates positive and negative estimates. RMSE values at the top quantify method accuracy; the lowest is highlighted in green (SDZ_lm: 1.03), the highest in red (IPW_lm: 11.08), and bold values indicate second-best (DR_gbm: 1.09) and second-worst (naive: 10.41) RMSEs.

Results of simulations for estimating the drought effect on yield

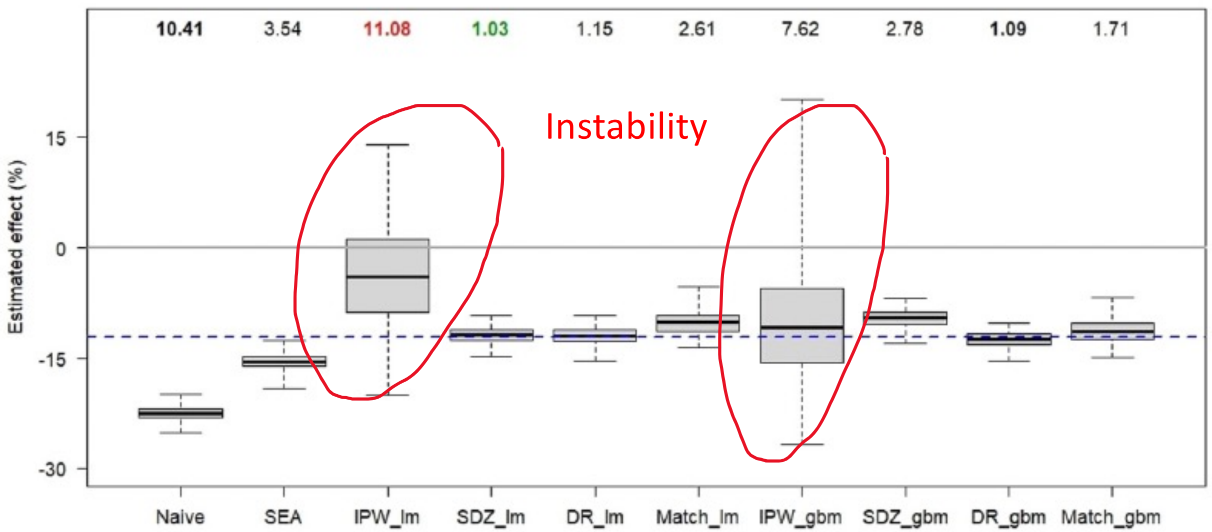


Fig. 10. Results of the data simulation study for estimating yield effects of drought on maize. Boxplots show the distribution of effects from 200 simulated datasets for each method and model type. The blue dashed line represents the true effect (-12.16 %), while the gray horizontal line at 0 differentiates positive and negative estimates. RMSE values at the top quantify method accuracy; the lowest is highlighted in green (SDZ_lm: 1.03), the highest in red (IPW_lm: 11.08), and bold values indicate second-best (DR_gbm: 1.09) and second-worst (naive: 10.41) RMSEs.

Results of simulations for estimating the drought effect on yield

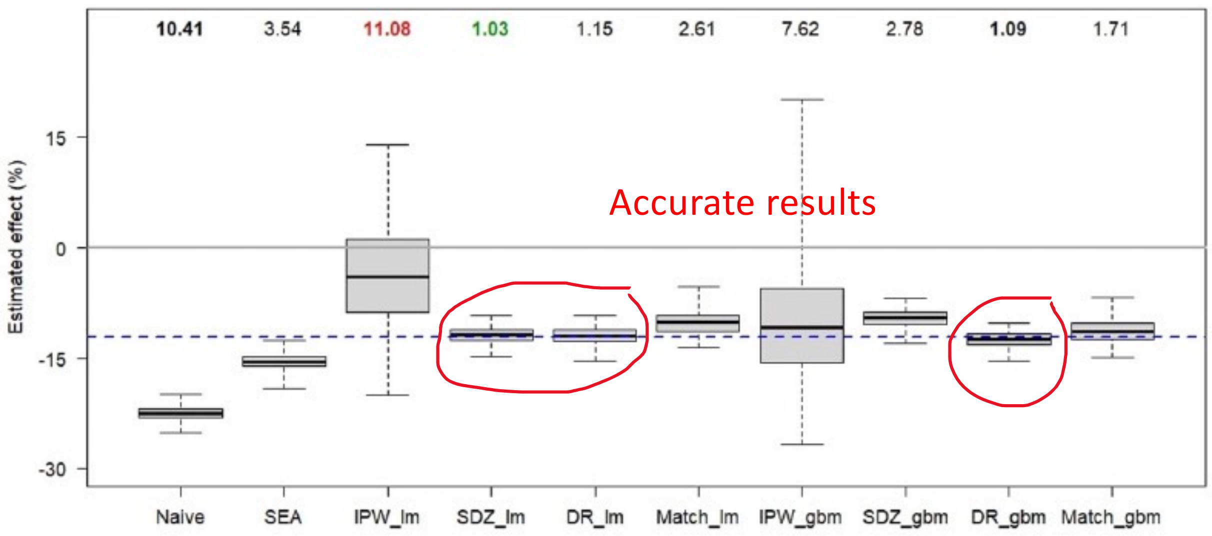
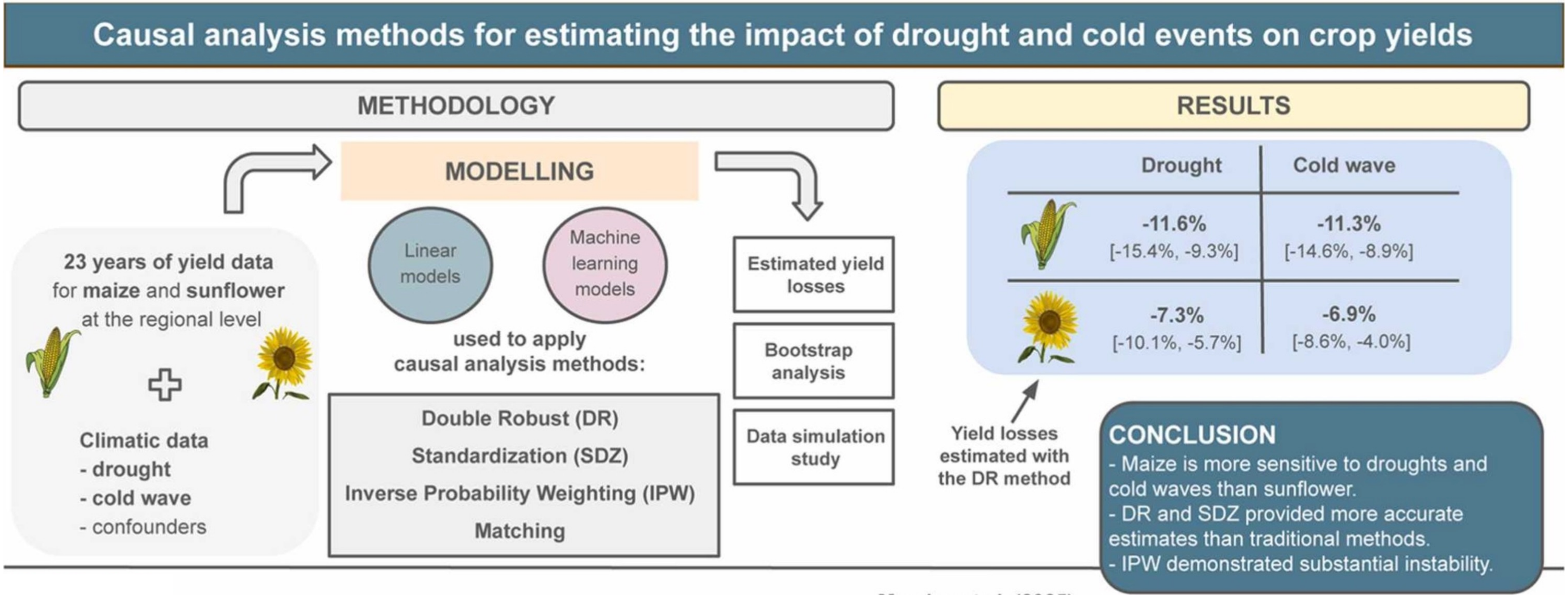


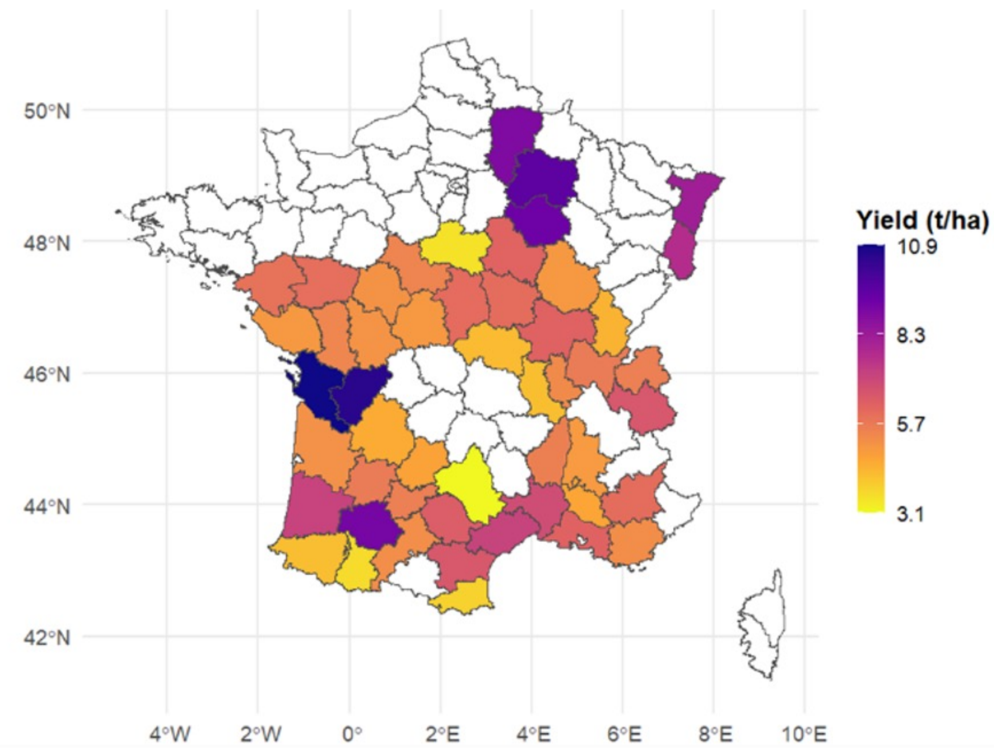
Fig. 10. Results of the data simulation study for estimating yield effects of drought on maize. Boxplots show the distribution of effects from 200 simulated datasets for each method and model type. The blue dashed line represents the true effect (-12.16 %), while the gray horizontal line at 0 differentiates positive and negative estimates. RMSE values at the top quantify method accuracy; the lowest is highlighted in green (SDZ_lm: 1.03), the highest in red (IPW_lm: 11.08), and bold values indicate second-best (DR_gbm: 1.09) and second-worst (naive: 10.41) RMSEs.


Causal analysis methods for estimating the impact of drought and cold events on crop yields
METHODOLOGY
RESULTS
MODELLING
Linear models
Machine learning models
23 years of yield data for maize and sunflower at the regional level
Climatic data
- drought
- cold wave
- confounders
used to apply causal analysis methods:
Double Robust (DR)
Standardization (SDZ)
Inverse Probability Weighting (IPW)
Matching
Estimated yield losses
Bootstrap analysis
Data simulation study
Drought
Cold wave
-11.6% [-15.4%, -9.3%]
-11.3% [-14.6%, -8.9%]
-7.3% [-10.1%, -5.7%]
-6.9% [-8.6%, -4.0%]
Yield losses estimated with the DR method
CONCLUSION
- Maize is more sensitive to droughts and cold waves than sunflower.
- DR and SDZ provided more accurate estimates than traditional methods.
- IPW demonstrated substantial instability.

Outline

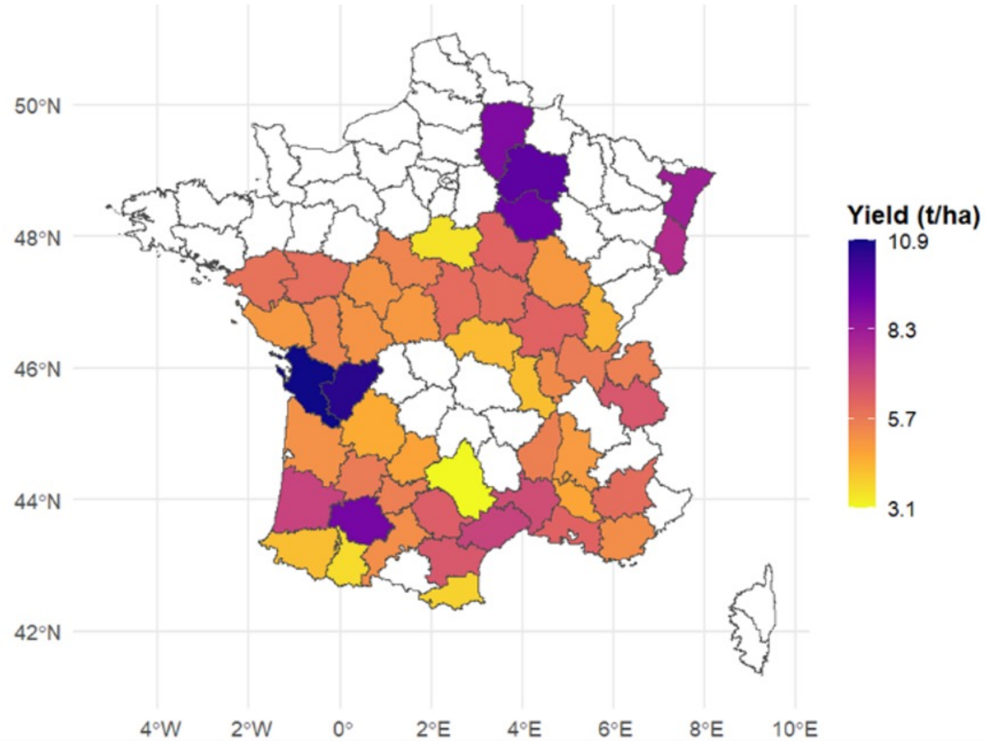
- Objective
- Methods
- Application 1: Maize and Sunflower
- Application 2: Vine
- Conclusions

Effects of high spring rainfall on downy mildew and grapevine yield in France



Effects of high spring rainfall on downy mildew and grapevine yield in France

A Y Y

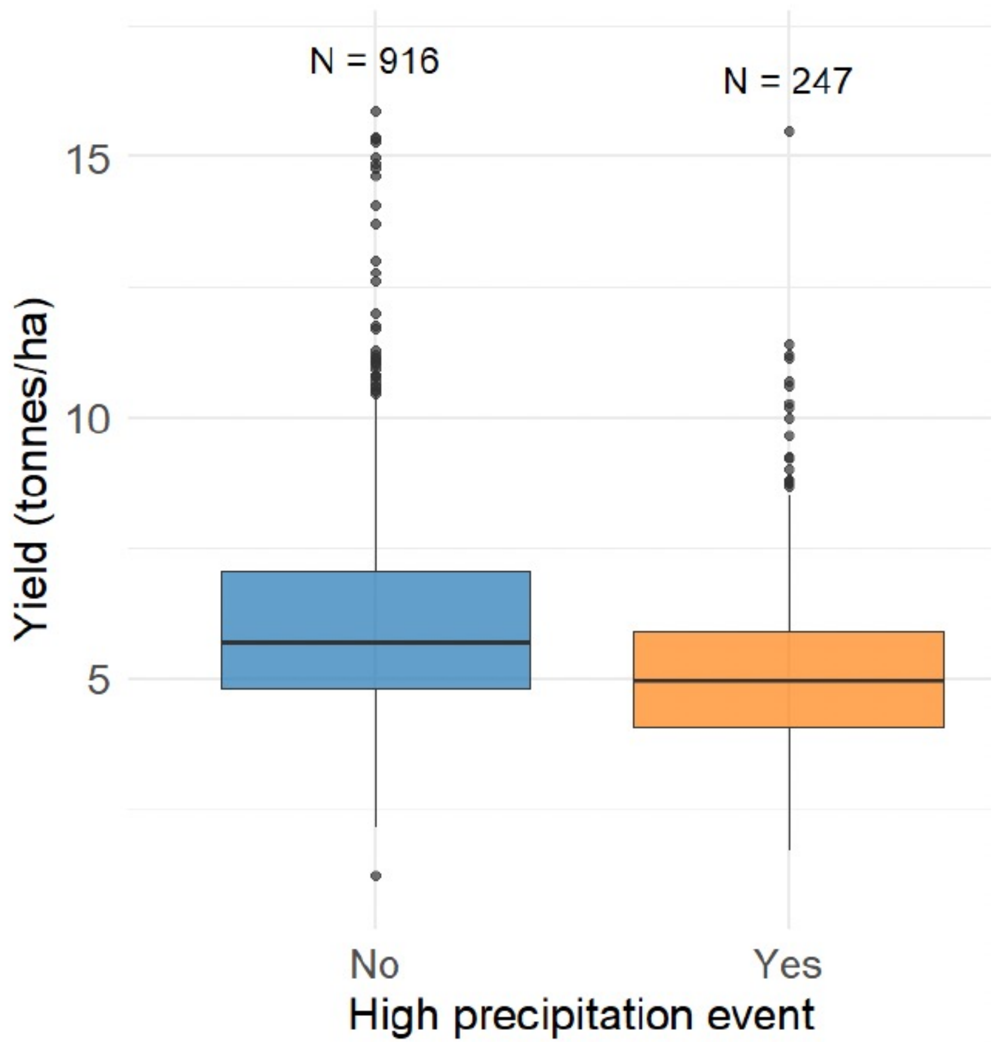


Treatment variable A

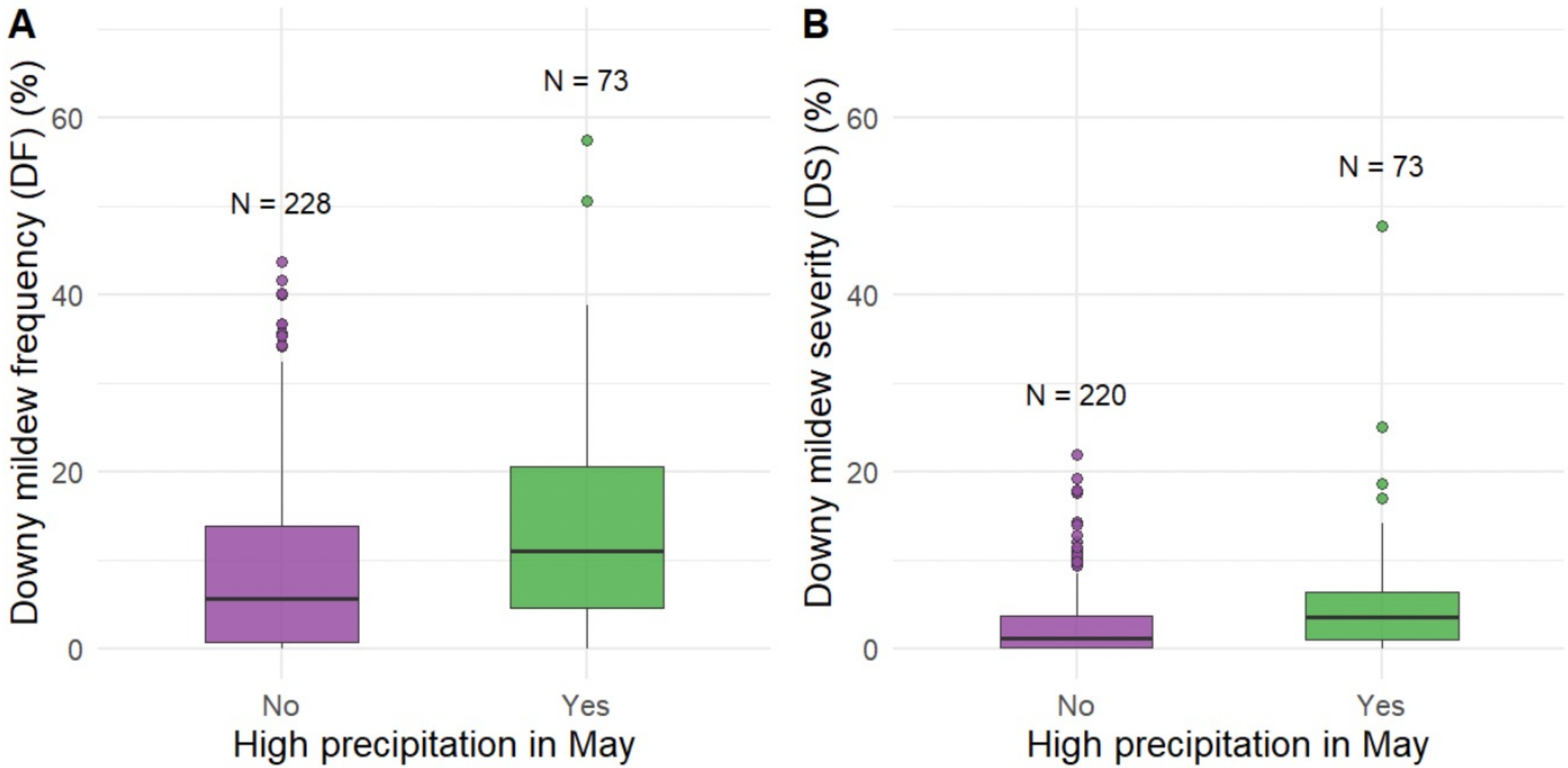
Confounding variables L

Table 1: **List of monthly meteorological variables.** HP corresponds to the high-precipitation event whose impacts on yield are estimated. $q_{0.8}$ corresponds to the 80% percentile of precipitation. The other variables are considered as potential confounding variables.

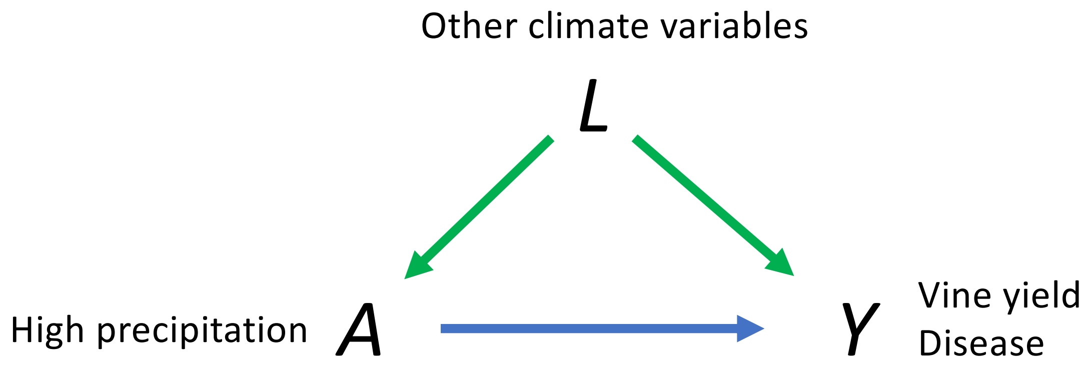
Variable	**Description**	**Unit**
HP	High-precipitation event: 1 if $P \geq q_{0.8}$, 0 otherwise (see section 2.4)	binary
P	Monthly average of liquid precipitation per day	mm/day
T_{max}	Monthly average daily maximum temperature	ºC
NDT30	Monthly number of days with maximum daily temperatures above 30°C	days
ETP	Monthly average potential evapotranspiration (Penman-Monteith) per day	mm/day
RG	Monthly average of visible radiation per day	J/cm^2
SWI	Monthly average soil water index, proxy for soil moisture, per day	%



N = 916
N = 247
15
10
5
Yield (tonnes/ha)
No
Yes
High precipitation event


A
B
N = 228
N = 73
N = 220
N = 73
Downy mildew frequency (DF) (%)
Downy mildew severity (DS) (%)
0
20
40
60
No
Yes
High precipitation in May

Double robust



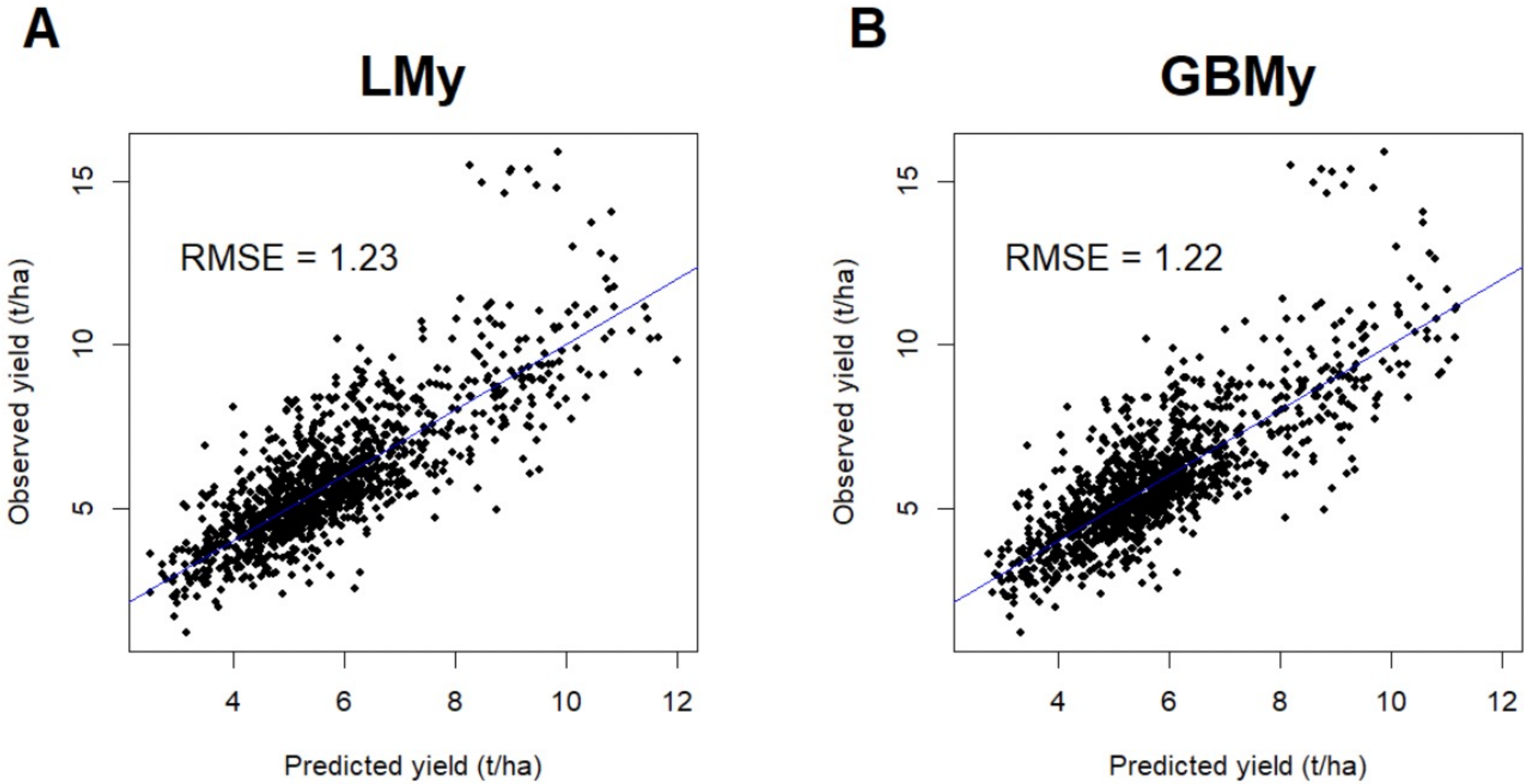
- Combine Inverse probability weighting and standardization
- Rely on two models

$$\hat{P}(A|L) = f(L)$$ Model predicting high precipitation occurrences

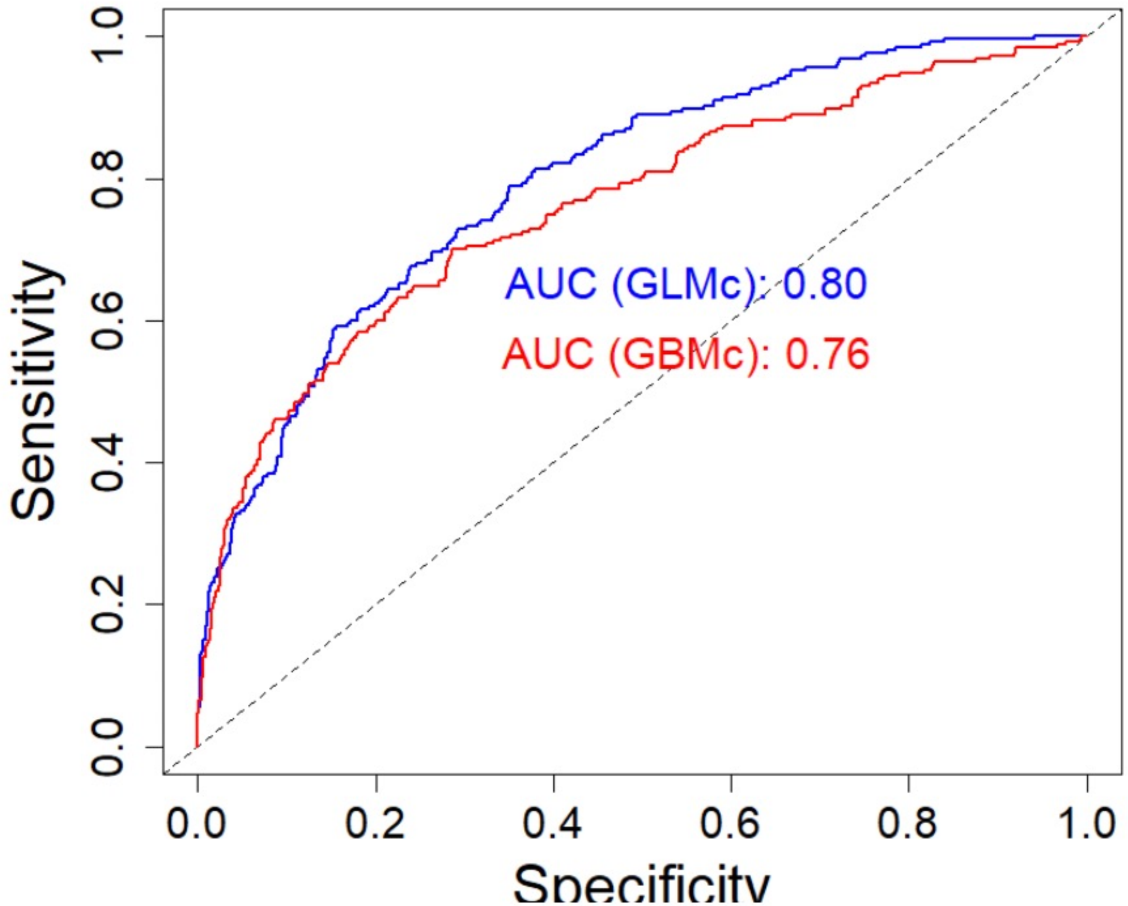
$$\hat{E}[Y|A,L] = g(A,L)$$ Model predicting vine yield

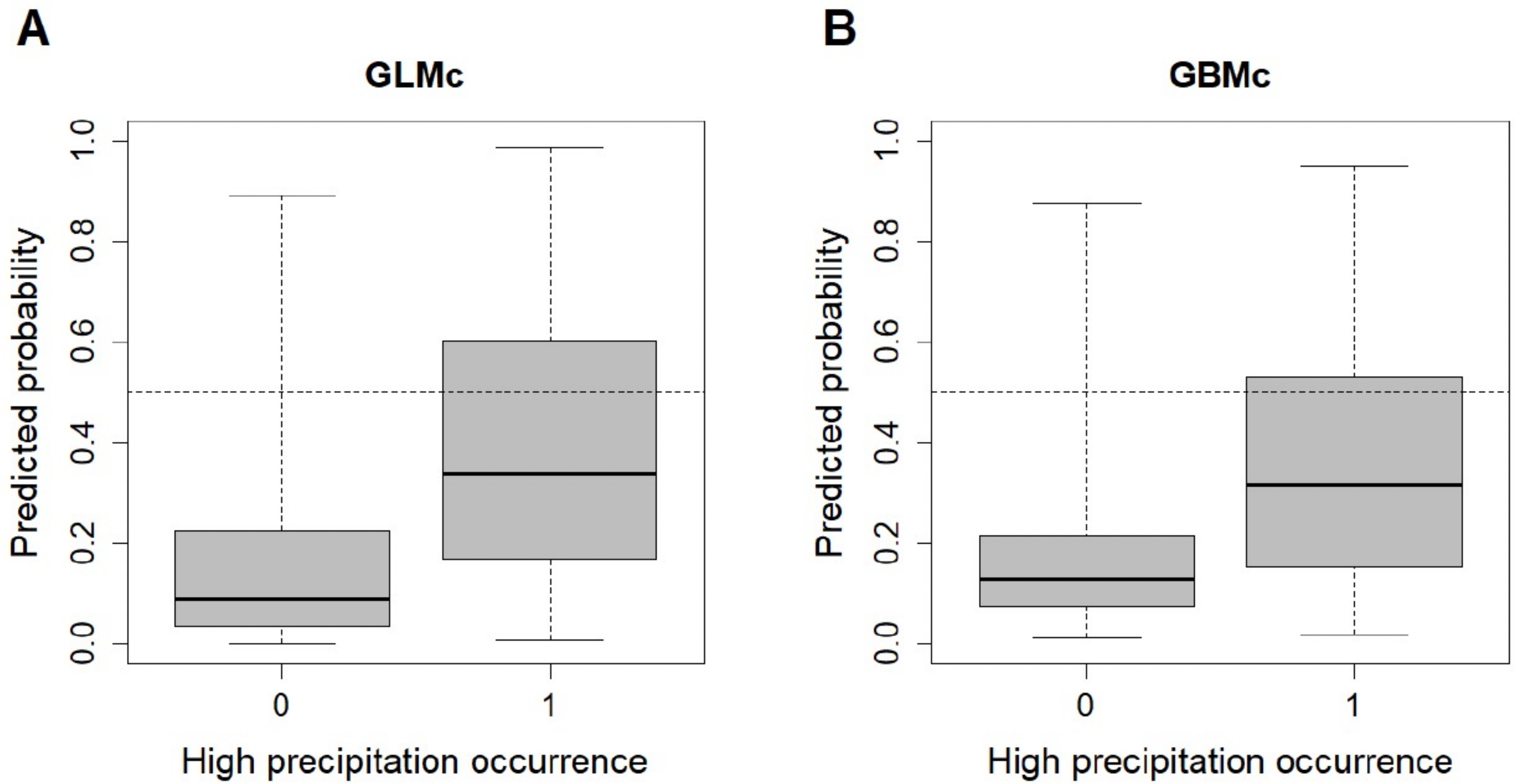
- Unbiased if one of the two models is unbiased

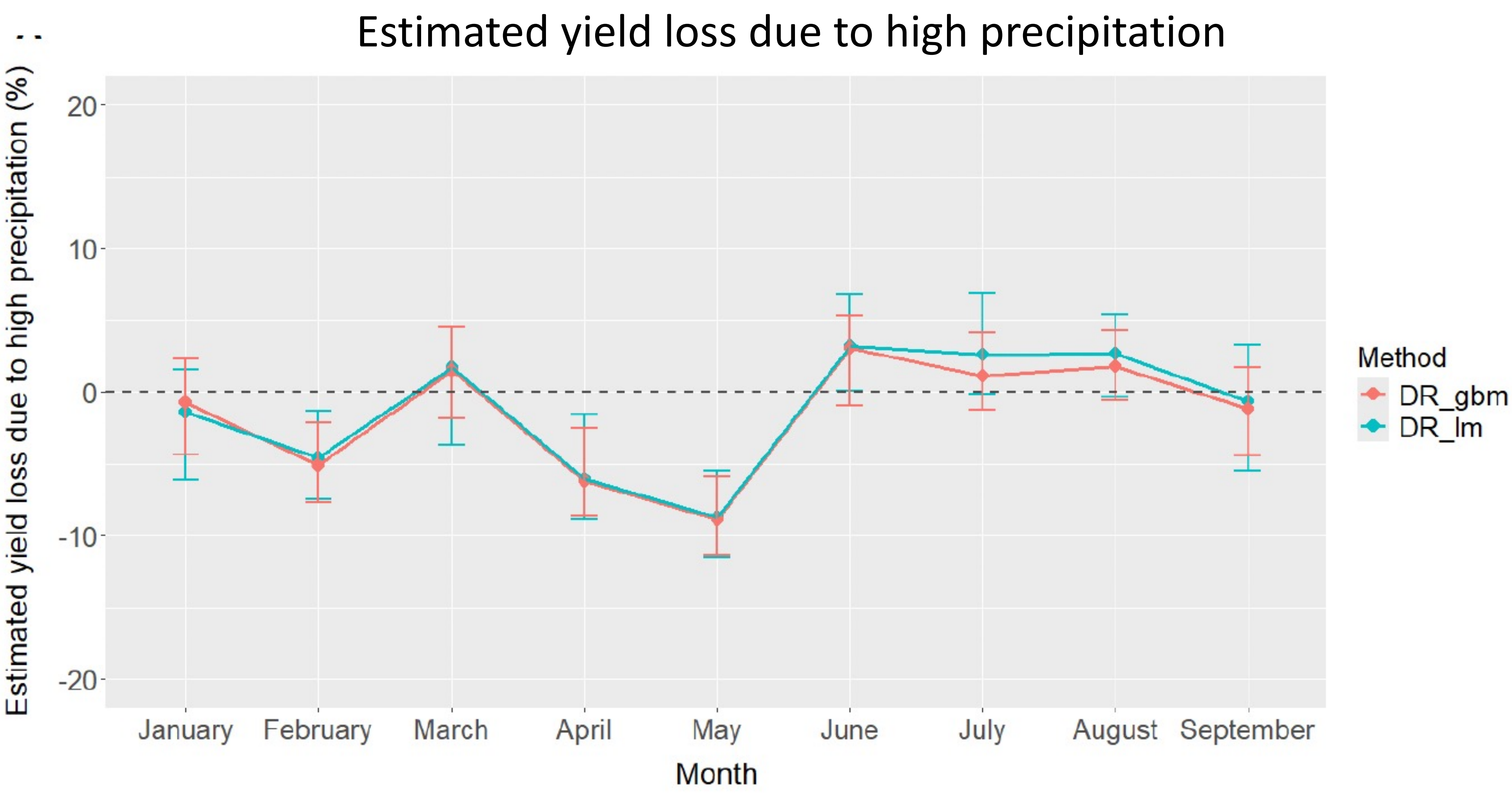
Model predicting yields



Model predicting « high precipitation » occurrence

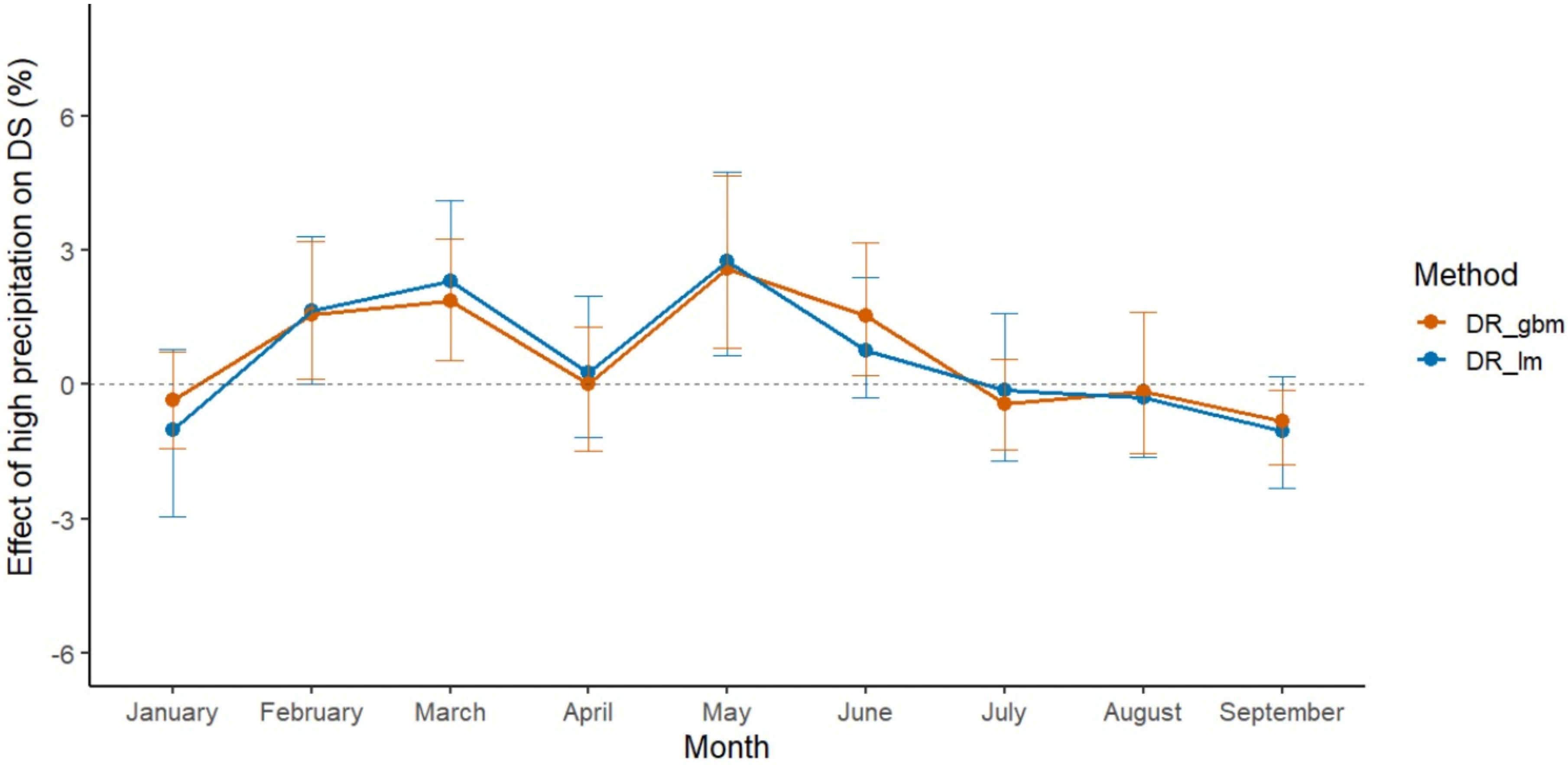


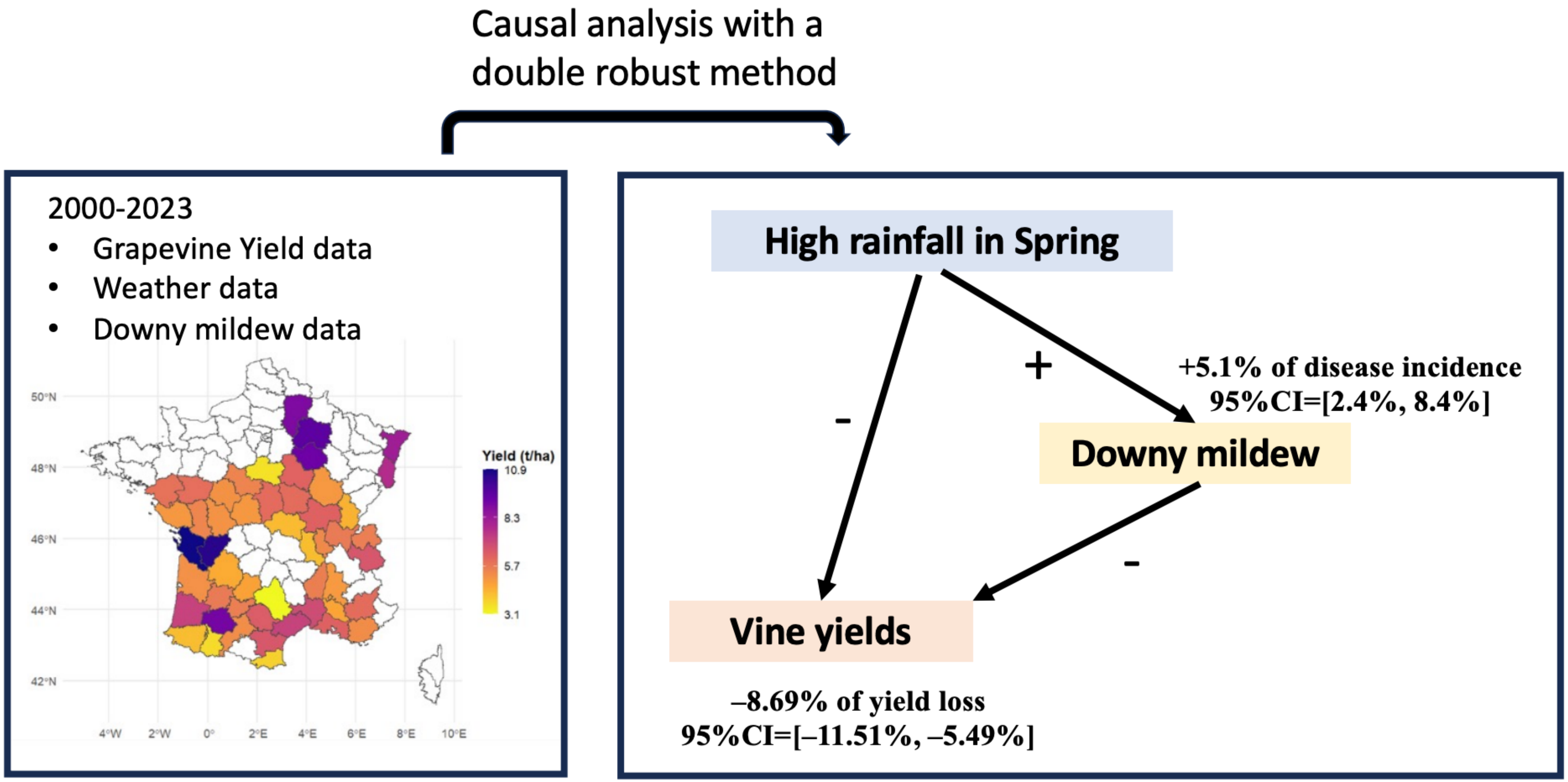

A
GLMc
B
GBMc
Predicted probability
0.0
0.2
0.4
0.6
0.8
1.0
0
1
High precipitation occurrence



Estimated yield loss due to high precipitation
Estimated yield loss due to high precipitation (%)
20
10
0
-10
-20
January
February
March
April
May
June
July
August
September
Month
Method
DR_gbm
DR_lm

Estimated effect of high precipitation on disease severity





Causal analysis with a double robust method
2000-2023
• Grapevine Yield data
• Weather data
• Downy mildew data
Yield (t/ha)
10.9
8.3
5.7
3.1
50°N
48°N
46°N
44°N
42°N
4°W
2°W
0°
2°E
4°E
6°E
8°E
10°E
High rainfall in Spring
+
-
Downy mildew
+5.1% of disease incidence
95%CI=[2.4%, 8.4%]
-
Vine yields
–8.69% of yield loss
95%CI=[–11.51%, –5.49%]

Conclusion

- Randomized control trials (RCTs) are commonly used to identify causal effect
- Meta-analysis is now very popular in agricultural science to synthetize RCTs and increase their statistical power
- However, often, RCTs are not available
- Recently, **causal inference** methods were developed to identify causal effect when **RCTs are not available**
- These methods are promising but not frequently used in agricultural & environmental science
- All these methods rely on one or two **models**, and **model selection** may have strong impacts

Some publications

Chatton A., Rohrer JM. (2024). The causal cookbook: Recipes for propensity scores, g-computation, and doubly robust standardization. *Advances in Methods and Practices in Psychological Science*. Vol. 7, No. 1. https://doi.org/10.1177/25152459241236149

Funk, M. J., Westreich, D., Wiesen, C., Stürmer, T., Brookhart, M. A., & Davidian, M. (2011). Doubly robust estimation of causal effects. *American Journal of Epidemiology*, *173*(7), 761–767.

Hernán, M., & Robins, J. (2020). Causal Inference: What If, Boca R. *FL: Chapman & Hall/CRC*, 257–275.

Mendes C., Aldebert K., Brun F., Makowski D. (2025). Causal analysis methods for estimating the impact of drought and coldevents on crop yields in France. *Eur. J. Agr*. 170, 127743.

Nogueira, A. R., Pugnana, A., Ruggieri, S., Pedreschi, D., & Gama, J. (2022). Methods and tools for causal discovery and causal inference. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,* 12(2), e1449. https://doi.org/10.1002/widm.1449

Sekhon, J. S. (2011). Multivariate and propensity score matching software with automated balance optimization: The matching package for R. Journal of Statistical Software, 42(7), 1–52.

Su Y., Gabrielle B., Makowski D. (2021). The impact of climate change on the productivity of conservation agriculture. *Nature Climate Change*, 11 628–633.

van der Wal, W. M., & Geskus, R. B. (2011). Ipw: An R package for inverse probability weighting. Journal of Statistical Software, 43(13), 1–23.

Yao, L., Chu, Z., Li, S., Li, Y., Gao, J., & Zhang, A. (2021). A survey on causal inference. ACM Transactions on Knowledge Discovery from Data, 15(5), 74:1–74:46.

Zhong Y, Kennedy EH, Bodnar LM, Naimi AI (2021). AIPW:An R Package for Augmented Inverse Probability Weighted Estimation of Average Causal Effects. American Journal of Epidemiology. DOI: 10.1093/aje/kwab207