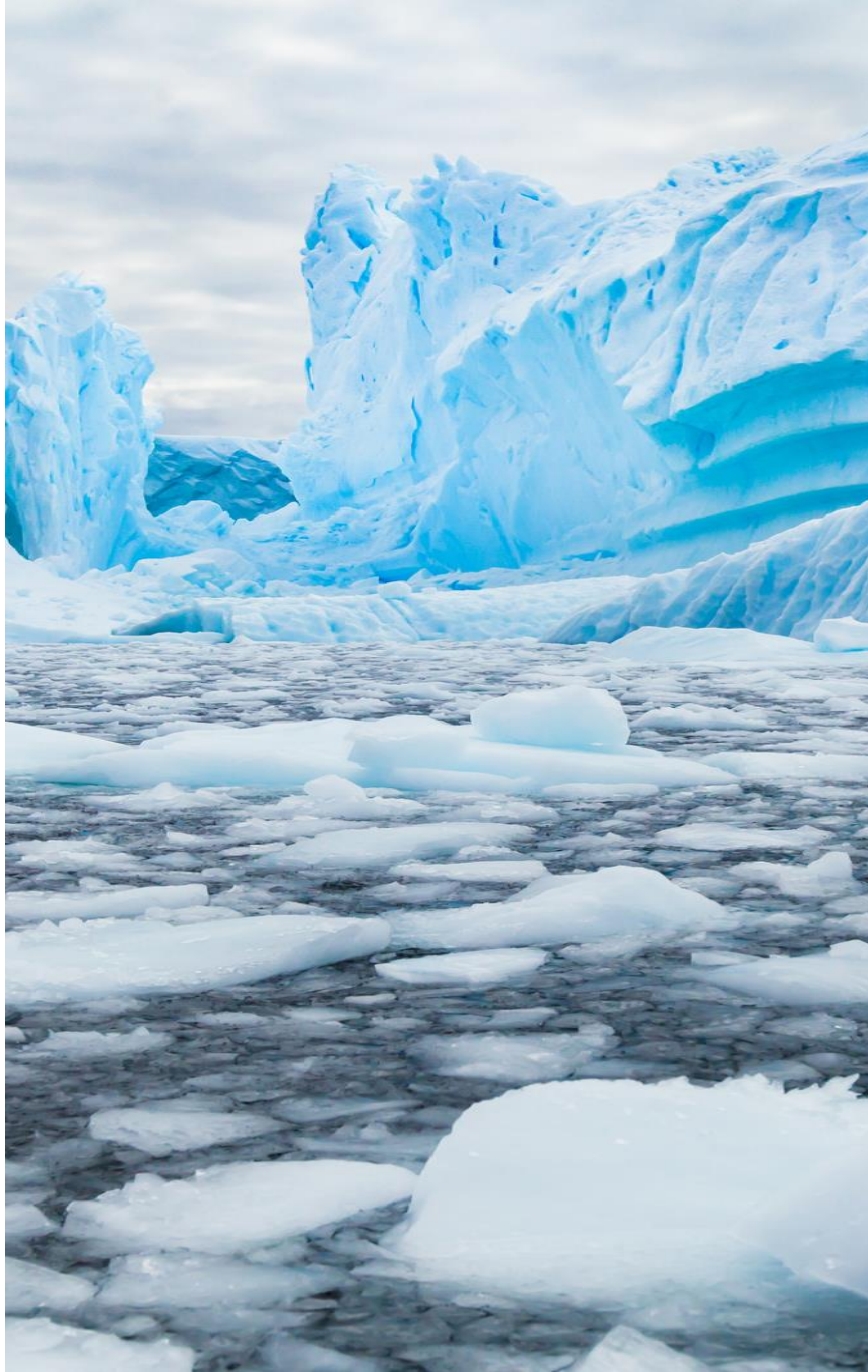
# Improving interpretation of **sea-level projections** through a machine-learning-based **local explanation approach**

Jeremy Rohmer (BRGM), Remi Thieblemont (BRGM), Goneri Le Cozannet (BRGM), Heiko Goelzer (NORSAR), Gael Durand (IGE)

Protect
CRYOSPHERE & SEA LEVEL

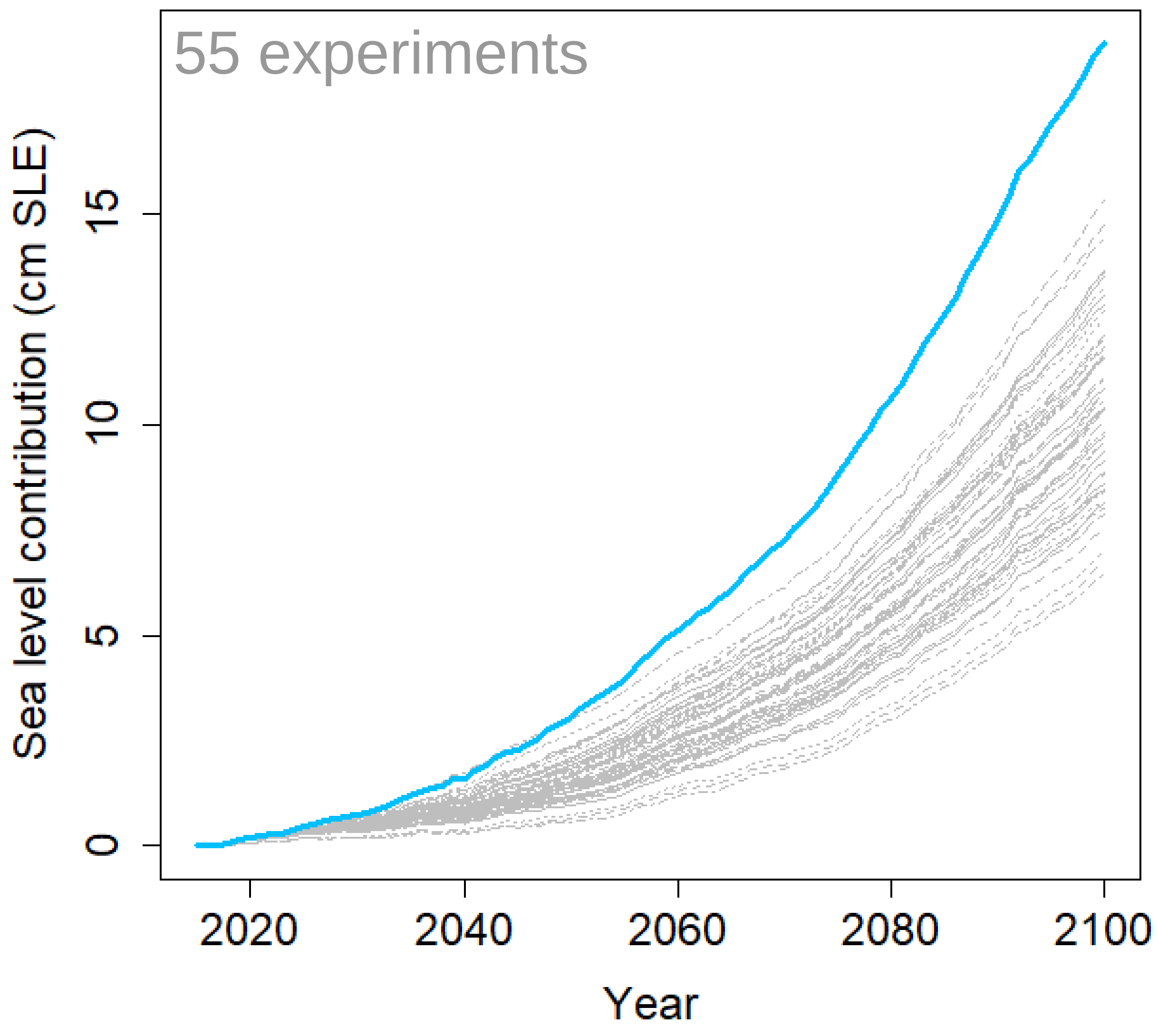


# Motivation

"What is 'easily explained' with words is expected to increase the end-user's level of trust in the model"

# Sea-level projections for the Greenland ice sheet (GrIS) within ISMIP6

**MIROC5,RCP8.5-forced Multi-Model Ensemble MME [1]**




Experiment 1

Modelling choices

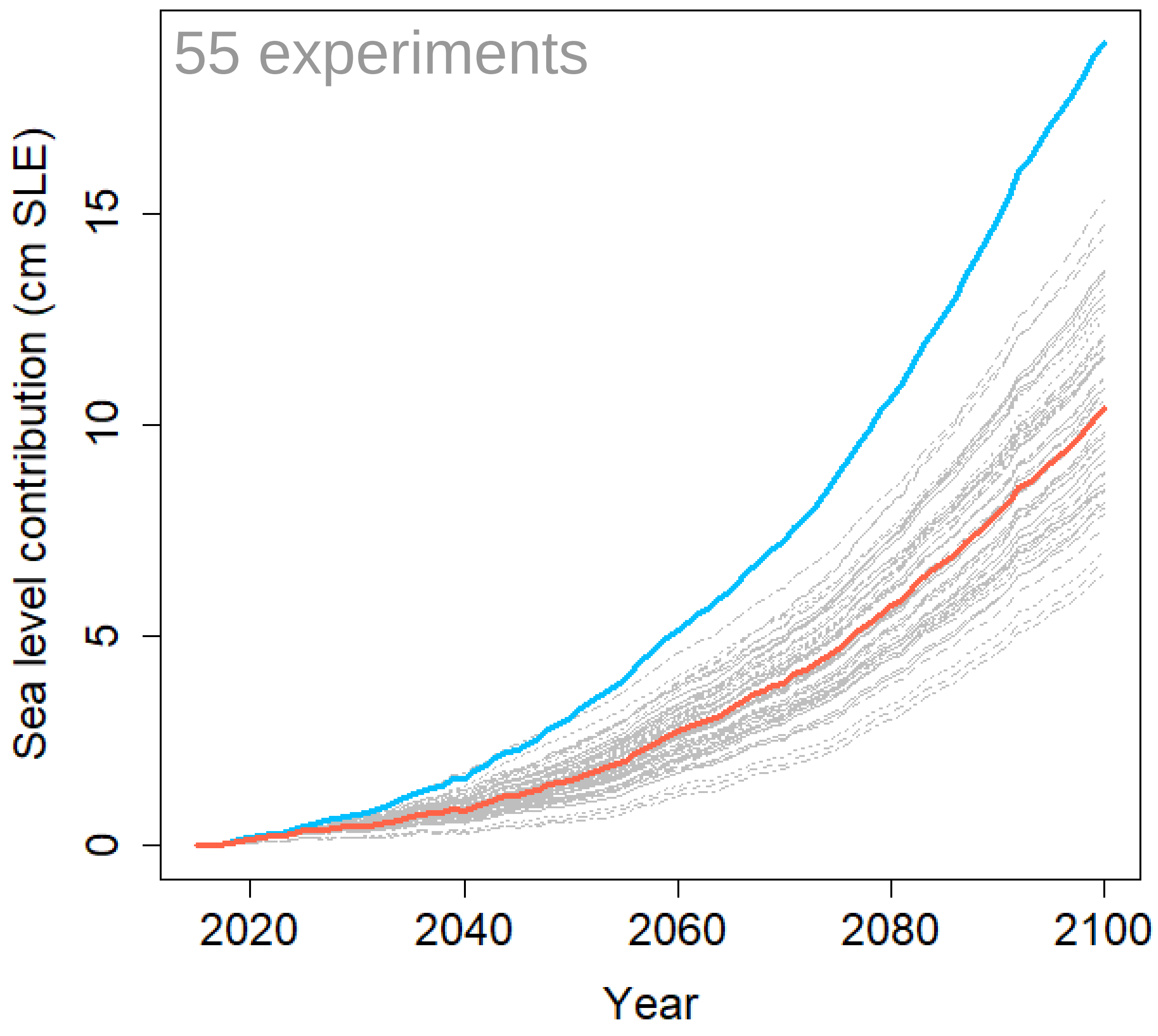
- **Numerical method** = Finite Difference (FD)
- **Ice Flow type** = shallow-ice approximation (SIA)
- **Type of initialization** = nudging to ice mask (NDm)
- **Min. grid size (resolution)** = 16km
- **Initial surface mass balance (SMB)** = RACMO anomalies (RA)
- **Retreat parameter** $\kappa$ = -0.9705 km $(m^3.s^{-1})^{-0.4}$ °C

Protect CRYOSPHERE & SEA LEVEL

[1] modelling assumptions described in Appendix A of Goelzer et al., the Cryosphere, 2020;

# Sea-level projections for the Greenland ice sheet (GrIS) within ISMIP6

**MIROC5,RCP8.5-forced Multi-Model Ensemble MME [1]**




Modelling choices

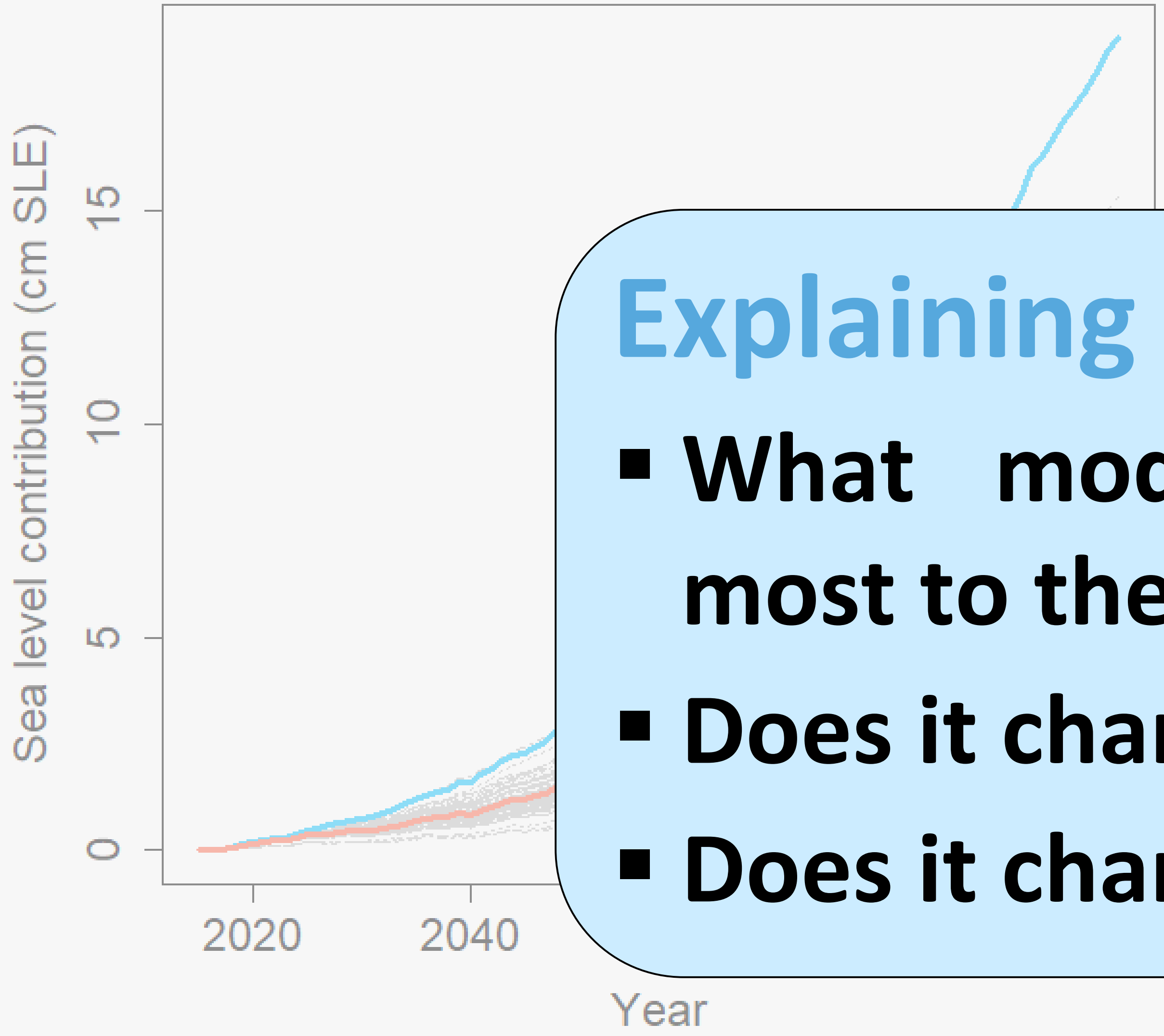
- **Numerical method** = Finite Difference (FD)
- **Ice Flow type** = shallow-ice approximation (SIA)
- **Type of initialization** = nudging to ice mask (NDm)
- **Min. grid size (resolution)** = 16km
- **Initial surface mass balance (SMB)** = RACMO anomalies (RA)
- **Retreat parameter** $\kappa$ = **-0.06 km $(m^3.s^{-1})^{-0.4}$ °C**

**…Etc…**

Protect CRYOSPHERE & SEA LEVEL

[1] modelling assumptions described in Appendix A of Goelzer et al., the Cryosphere, 2020;

# Sea-level projections for the Greenland ice sheet (GrIS) within ISMIP6




## Explaining the projections:

- **What modelling assumptions contribute the most to the sea level value?**
- **Does it change from one experiment to another?**
- **Does it change over time?**

$.s^{-1})^{-0.4}$ °C

…Etc…

[1] Goelzer et al., the Cryosphere, 2020

# Methods

Protect
CRYOSPHERE & SEA LEVEL

# Local explanation approach

For each experiment (i)

Sea level elevation = function f(.) of the modelling choices

$$\text{sea level}^{(i)} = f(\text{Num}^{(i)}, \text{Init}^{(i)}, \text{iceflow}^{(i)}, \text{SMB}^{(i)}, \text{Retreat}^{(i)}, \text{Reso}^{(i)})$$

Modelling choice for the Numerical method

Modelling choice for the initialization

....

Modelling choice for the min. grid size

**Objective:**

$$\text{sea level}^{(i)} = \mu_0 + \mu_{\text{Num}^{(i)}} + \mu_{\text{Init}^{(i)}} + \mu_{\text{iceflow}^{(i)}} + \mu_{\text{SMB}^{(i)}} + \mu_{\text{Retreat}^{(i)}} + \mu_{\text{Reso}^{(i)}}$$

'additive' contribution of each input to the seal level value
= <u>influence of the modelling choice</u>

Protect CRYOSPHERE & SEA LEVEL

## Difficulties

$\boxed{1}$ f(.) is a complex chain of numerical models

$\boxed{2}$ f(.) is long running, and the number of experiments is limited

$\boxed{3}$ Design of experiments is unbalanced + dependence betw. the inputs

## Proposed method

Approximate f(.) by a **Machine Learning ML** [1] model $\tilde{f}(.)$

Use tools of **XAI***:

decompose the ML predictions using SHAP **local attribution approach** [2,3]

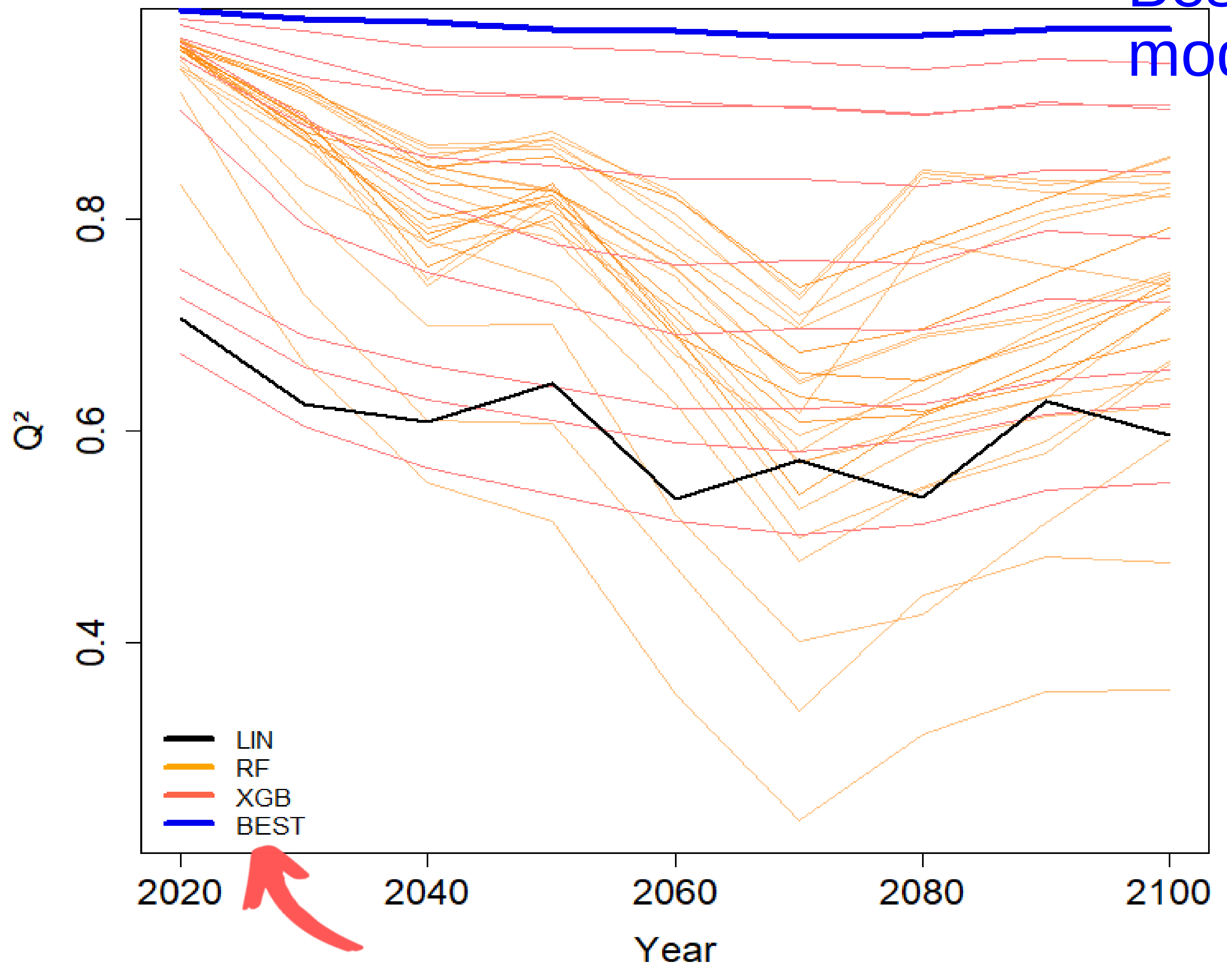
≈ regression coefficients with accounts for **dependence between inputs**



*XAI Explainable Artificial Intelligence;
[1] Random forest, Breiman [2] Lundberg & Lee 2017; [3]: Aas et al. 2019

# Validation of using the machine learning (ML) model

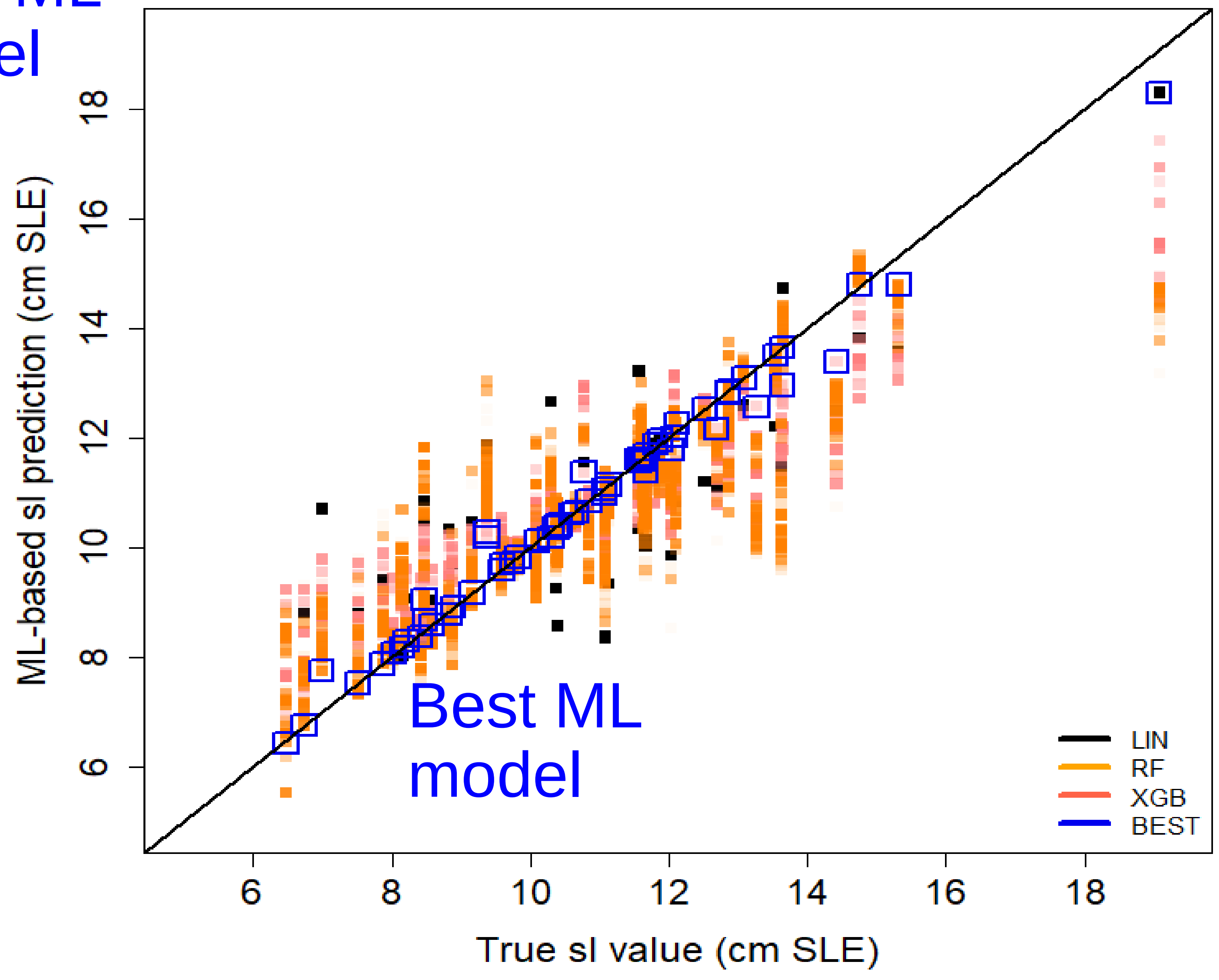
10-fold Cross validation
Performance quality




Best ML model

Series of considered ML models

For 2100

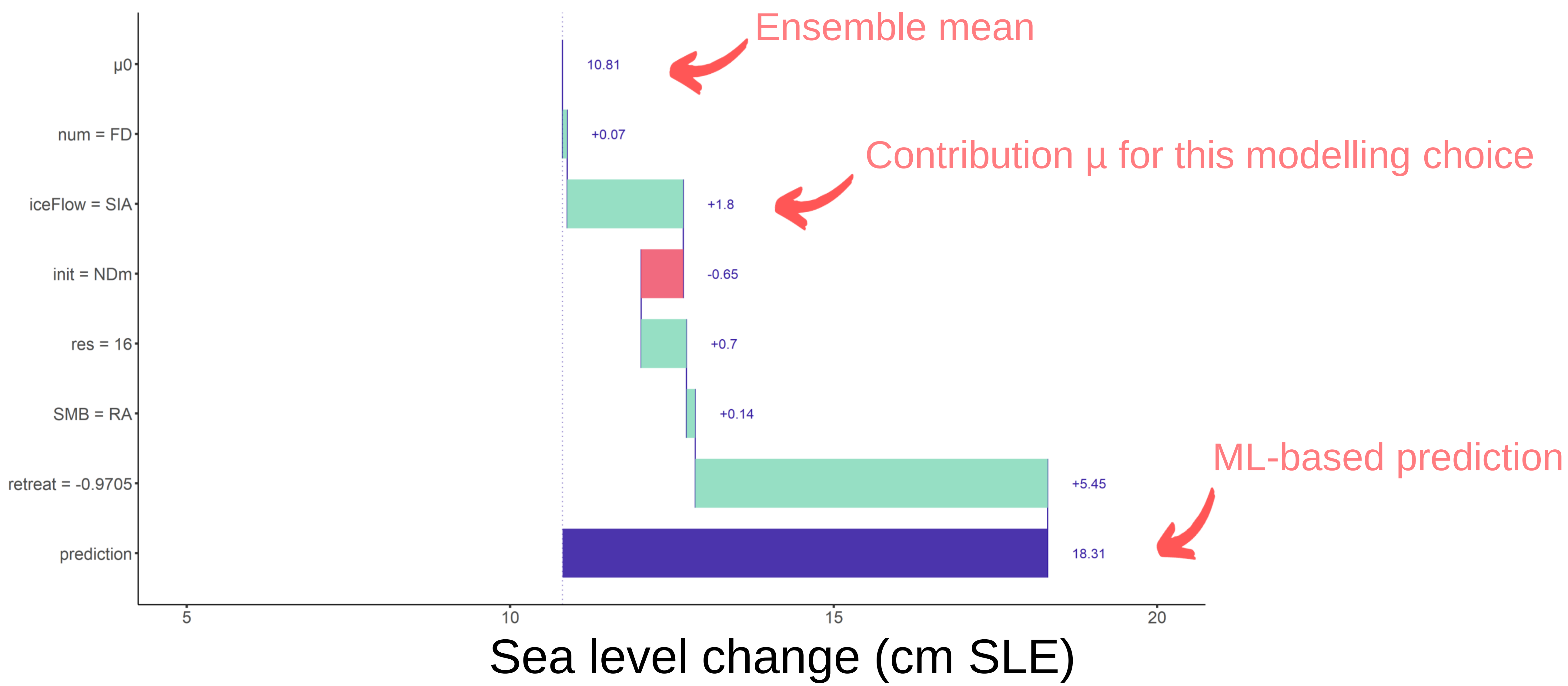



Protect CRYOSPHERE & SEA LEVEL

RF Random forest; XGB eXtreme Gradient; LIN linear regression

# Application

# Experiment (1), prediction time = 2100



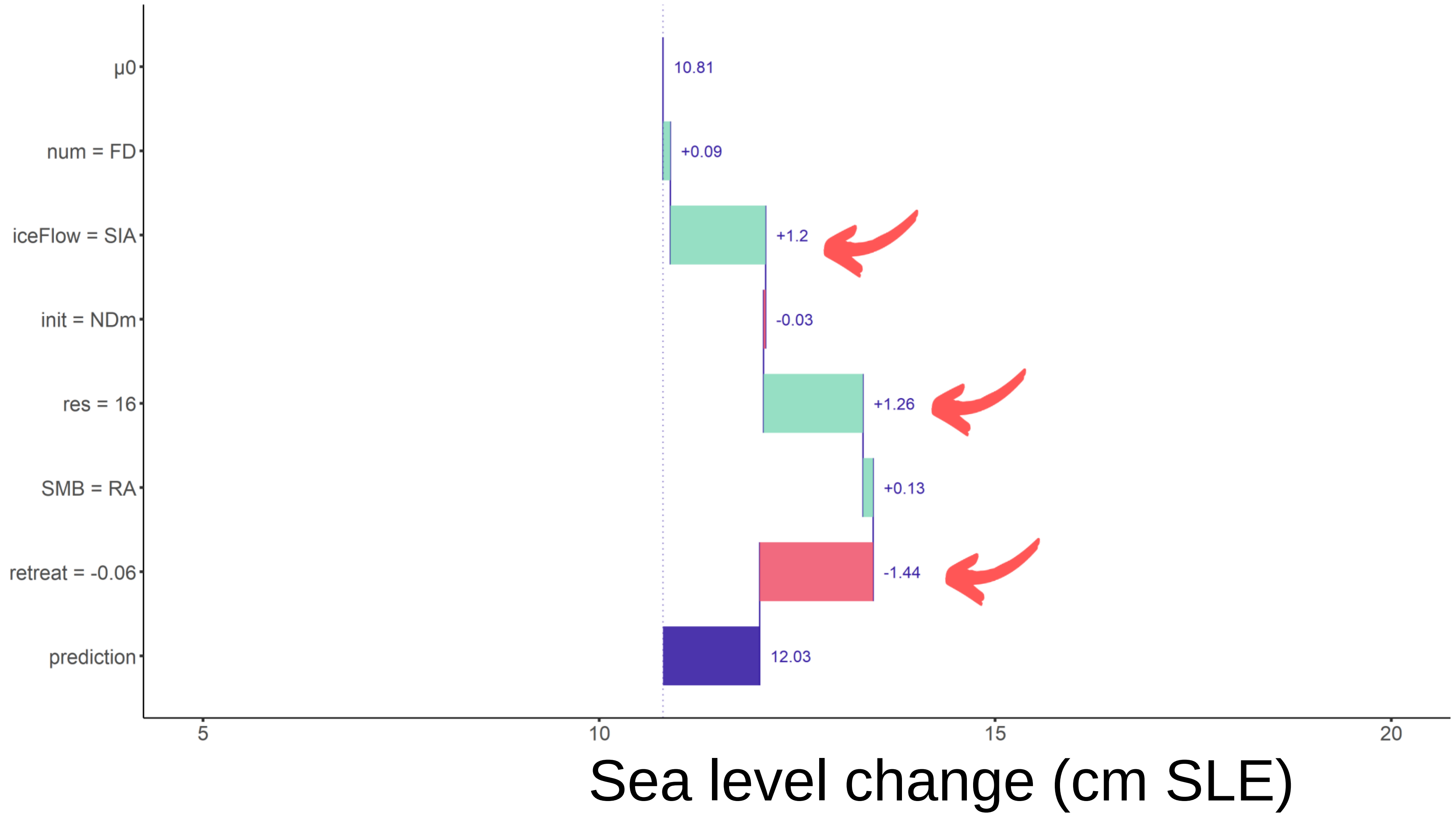

**With words**

"The largest sea level value is mostly influenced by the assumption of a retreat parameter set up at its largest absolute value"

"The contribution of this assumption reaches >5cm"

Protect
CRYOSPHERE & SEA LEVEL

# Experiment (2), prediction time = 2100

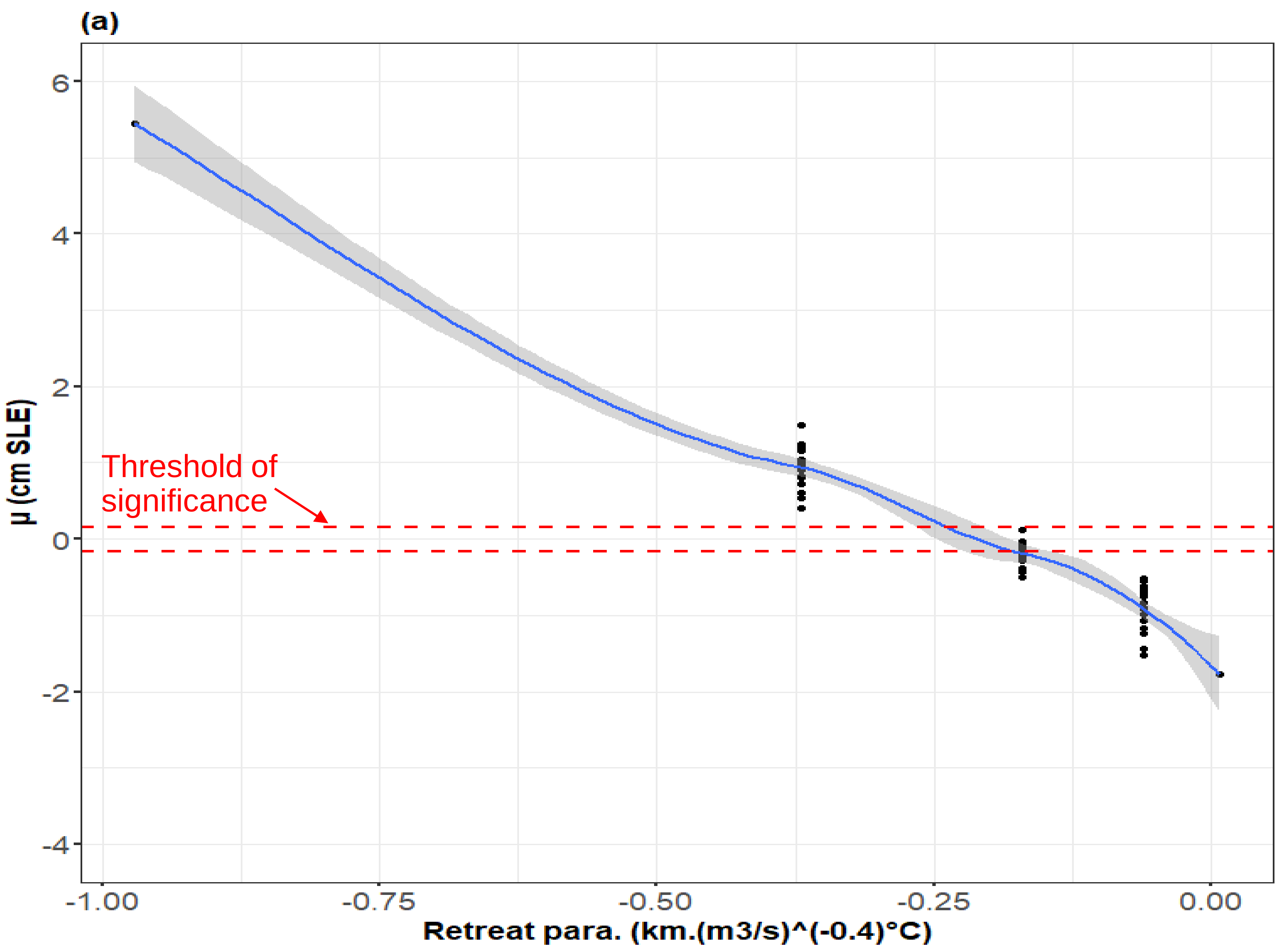



**With words**

"The sea level value of 12cm is mostly controlled by 3 modelling choices with an equal contribution of 1-1.5cm"

Protect
CRYOSPHERE & SEA LEVEL

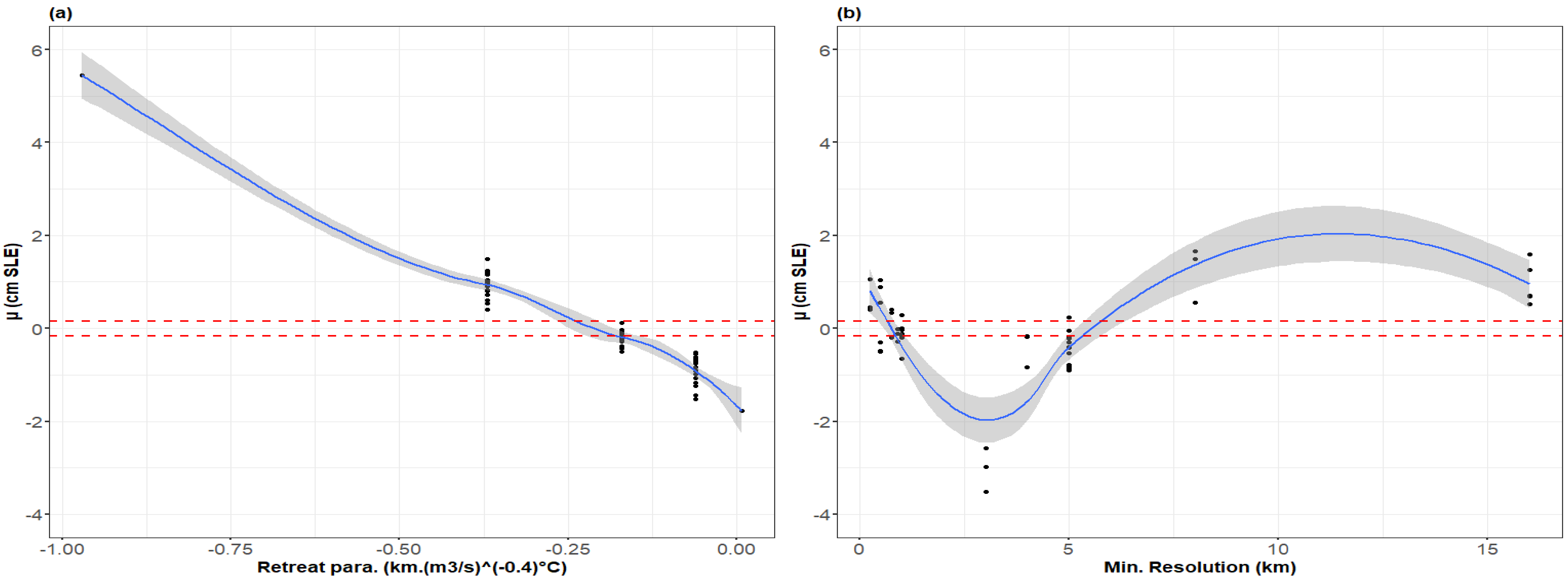
# Aggregation of all 55 experiments, prediction time = 2100




"Increasing the absolute value of the retreat parameter increases its contribution to the sea level value"

Protect
CRYOSPHERE & SEA LEVEL

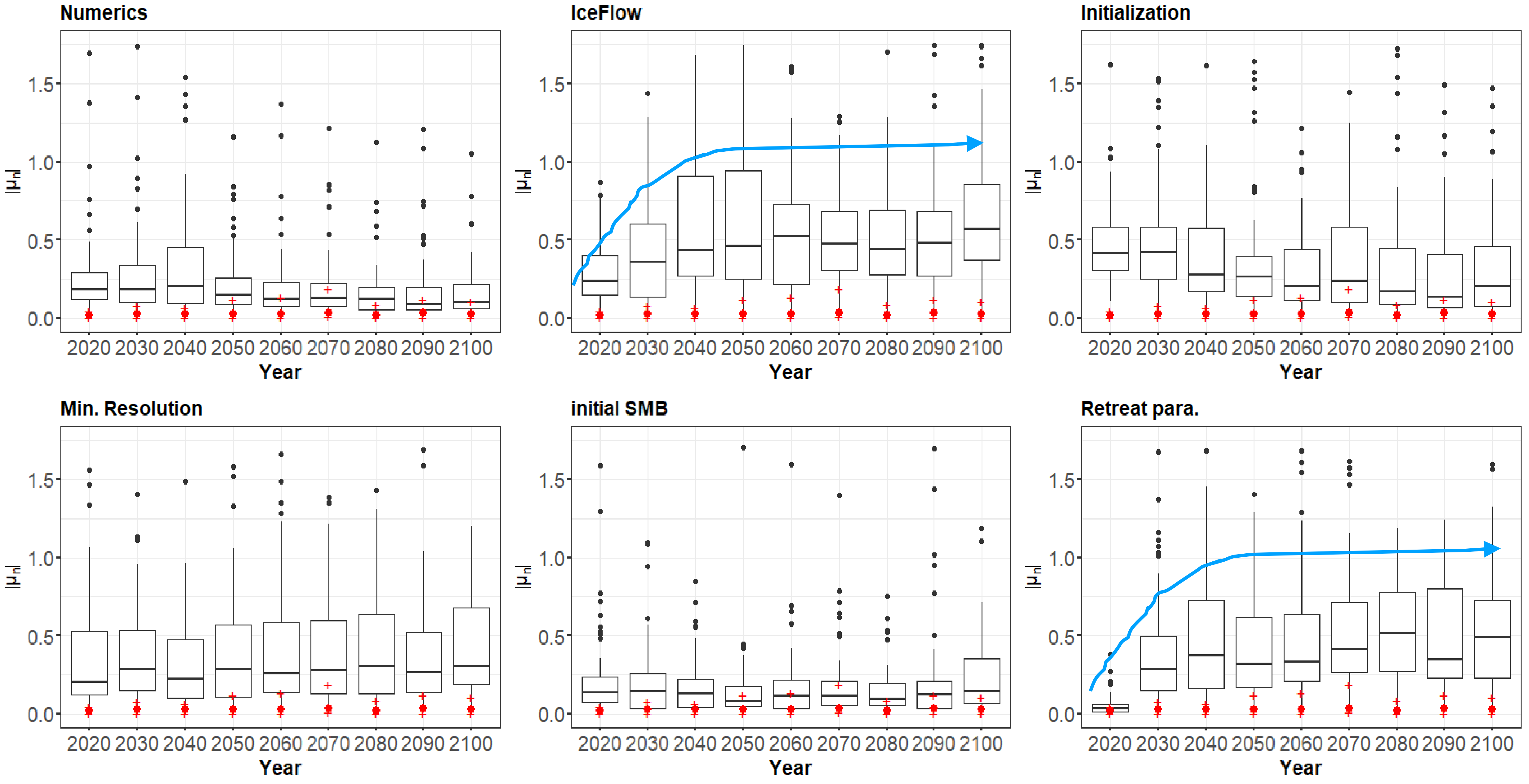
# Aggregation of all 55 experiments, prediction time = 2100




"Too high mesh grid size (>5km) might bias the computed sea level value"

Protect
CRYOSPHERE & SEA LEVEL

# Aggregation of all 55 experiments, over time




"Influence increase of ice flow and retreat parameter over time"

"Influence of the numerical method only in the short term"

"Quasi-constant influence of the min. grid size"

Protect
CRYOSPHERE & SEA LEVEL

# Summary

- **For each experiment (i) of the considered MME** we decompose:

$$\text{sea level}^{(i)} = \mu_0 + \mu_{\text{Num}^{(i)}} + \mu_{\text{Init}^{(i)}} + \mu_{\text{iceflow}^{(i)}} + \mu_{\text{SMB}^{(i)}} + \mu_{\text{Retreat}^{(i)}} + \mu_{\text{Reso}^{(i)}}$$

- Three levels of analysis

1. *μ* (expressed in cm SLE) measures the contribution of the modelling choices in the sea level value = **local importance analysis**
2. Analysing all decompositions at a given prediction time helps understanding how the influence relates to the modelling choices = **model structure**
3. Aggregating all decompositions for any prediction time helps getting insights in the **global influence** of the modelling choices

- Applicable to any MME with
  - a **limited number** of experiments (50-100)
  - **unbalanced** design
  - presence of **dependence** between the inputs

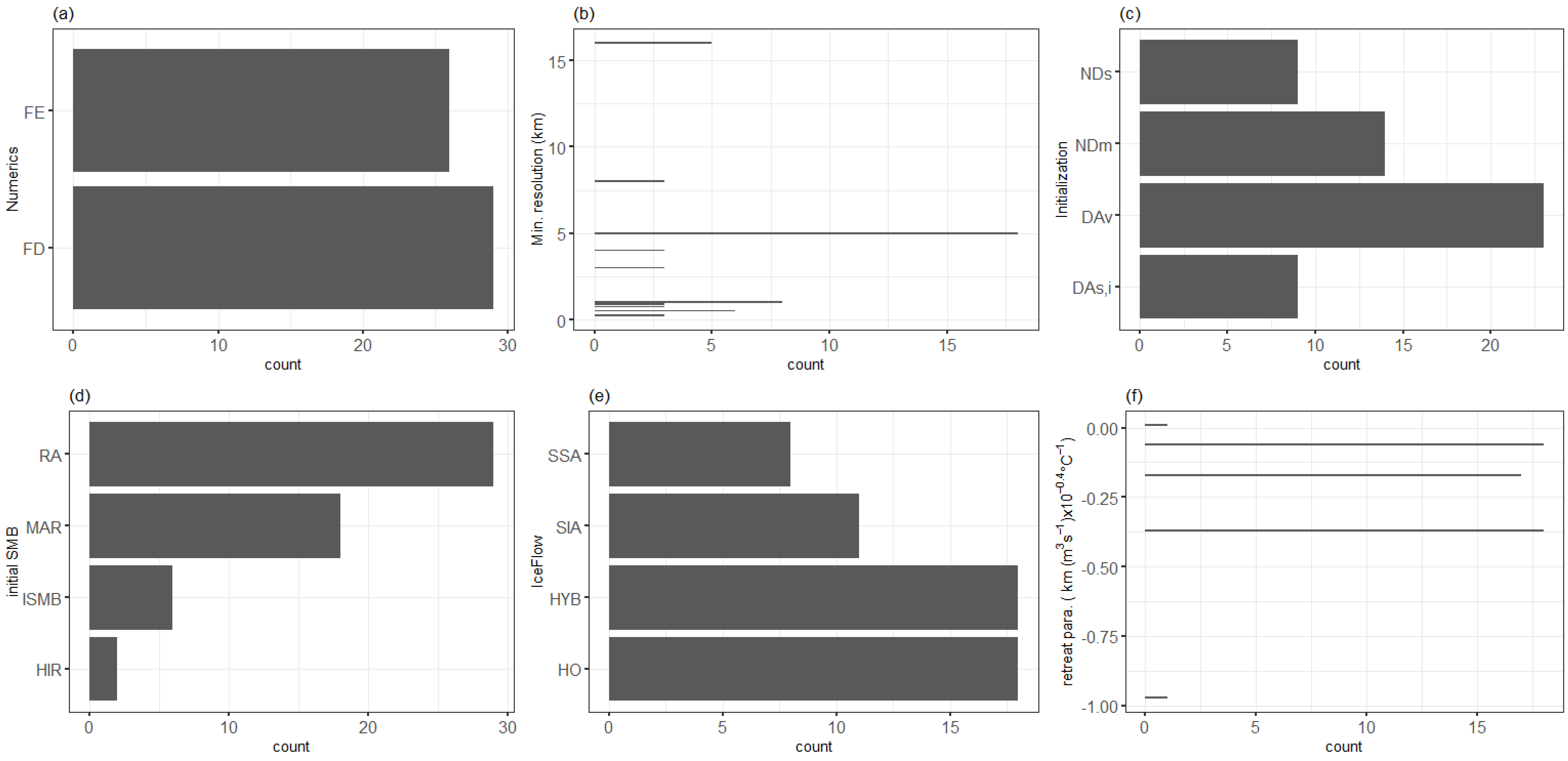
PROTECT CRYOSPHERE & SEA LEVEL

# Acknowledgements

THIS PROJECT HAS RECEIVED FUNDING FROM THE EUROPEAN UNION'S HORIZON 2020 RESEARCH AND INNOVATION PROGRAMME UNDER GRANT AGREEMENT 869304.

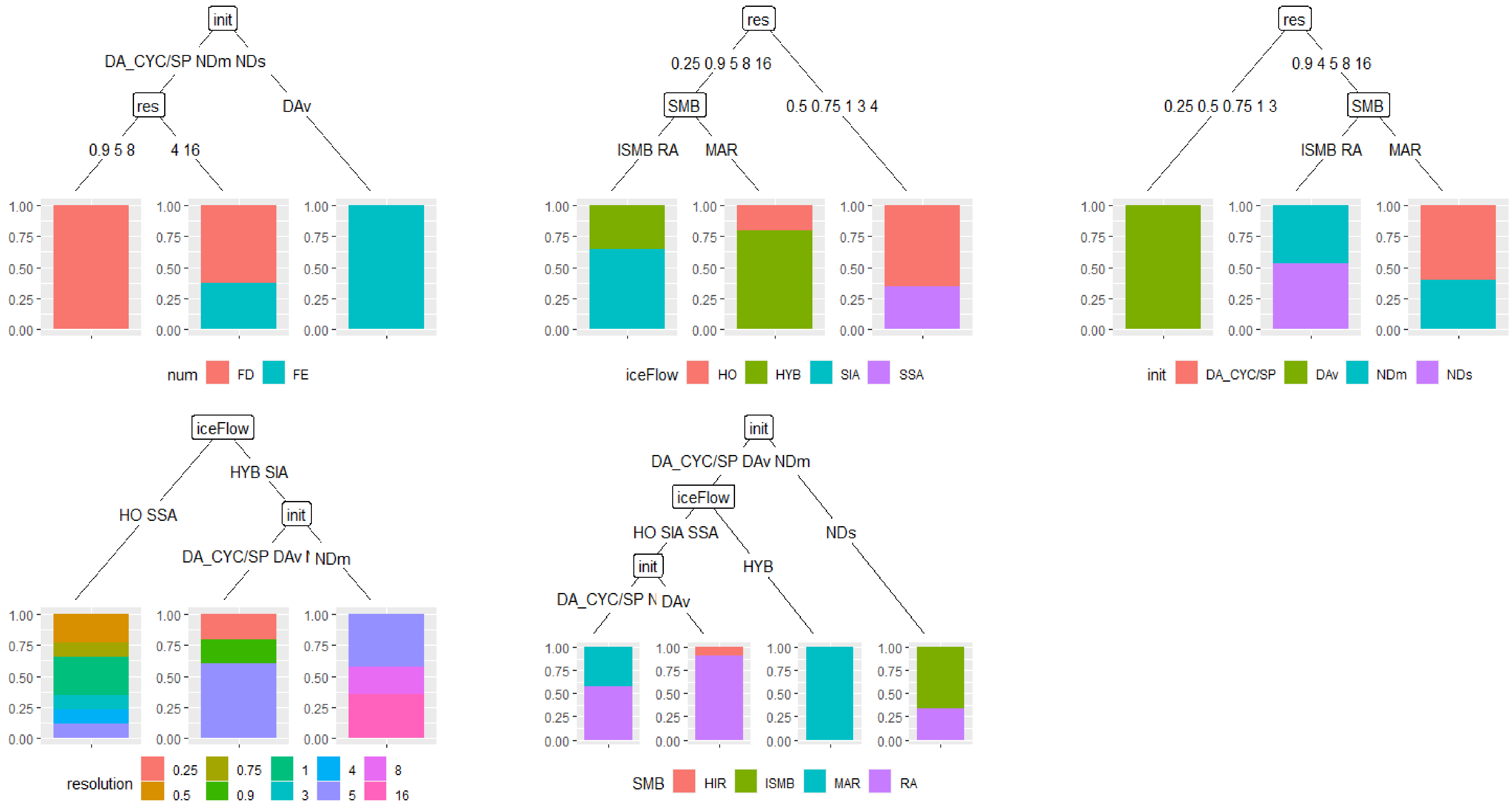
# Appendices

# **Unbalanced** design of experiments in the considered MME GrIS case [1]

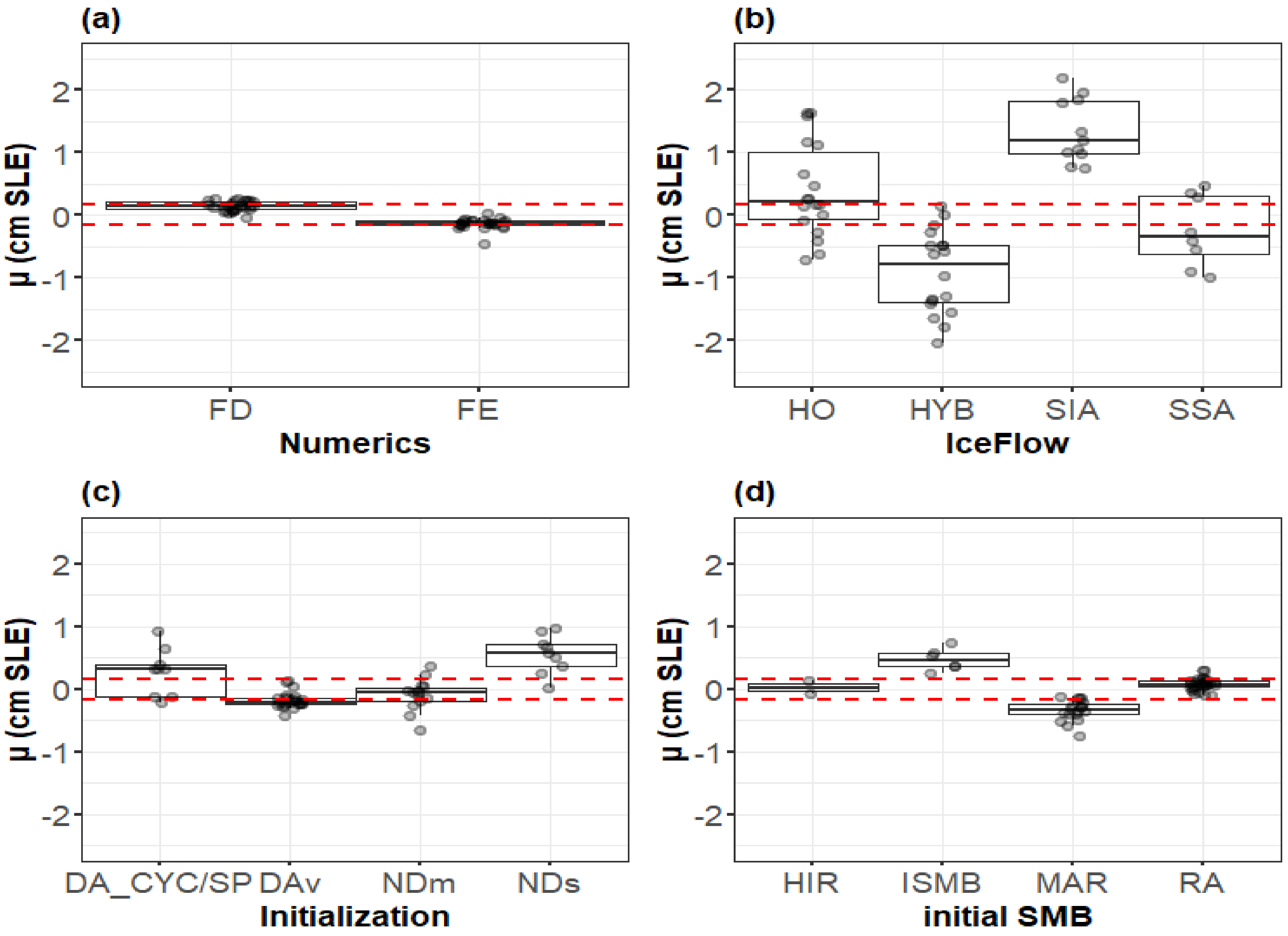



[1] modelling assumptions described in Appendix A of Goelzer et al., the Cryosphere, 2020;

# Evidence of dependence



# Aggregation of all 55 experiments, prediction time = 2100




"The choice in the numerical method does not influence much"

"Setting the ice flow type to HYB or SIA has the largest impact"

Protect
CRYOSPHERE & SEA LEVEL