



Clim4^litis

A PHENOLOGY MODEL INTERCOMPARISON

CHRISTOPH MENZ AND CLIM4VITIS TEAM



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement n° 810176. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Introduction

Motivation

- Model intercomparisons common in other research fields
 - CMIP, ScenarioMIP, CORDEX, AgMIP, ISIMIP, ...

Introduction

Motivation

- Model intercomparisons common in other research fields
 - CMIP, ScenarioMIP, CORDEX, AgMIP, ISIMIP, ...



The good, the Bad and the Ugly, Sergio Leone,
Produzioni Europee Associate/United Artists

Introduction

Motivation

- Model intercomparisons common in other research fields
 - CMIP, ScenarioMIP, CORDEX, AgMIP, ISIMIP, ...
 - Initiating model development
 - Quantify bounds for model uncertainty
 - Guide subsequent impact modelers/practitioners



The good, the Bad and the Ugly, Sergio Leone,
Produzioni Europee Associate/United Artists

Introduction

Motivation

- Model intercomparisons common in other research fields
 - CMIP, ScenarioMIP, CORDEX, AgMIP, ISIMIP, ...
 - Initiating model development
 - Quantify bounds for model uncertainty
 - Guide subsequent impact modelers/practitioners
- Few intercomparison studies available for phenology models for grapevine
 - Models initially developed for single sites/varieties
 - Insufficient observations available
 - No funding



The good, the Bad and the Ugly, Sergio Leone,
Produzioni Europee Associate/United Artists

Introduction

Motivation

- Model intercomparisons common in other research fields
 - CMIP, ScenarioMIP, CORDEX, AgMIP, ISIMIP, ...
 - Initiating model development
 - Quantify bounds for model uncertainty
 - Guide subsequent impact modelers/practitioners
- Few intercomparison studies available for phenology models for grapevine
 - Models initially developed for single sites/varieties
 - Insufficient observations available
 - No funding



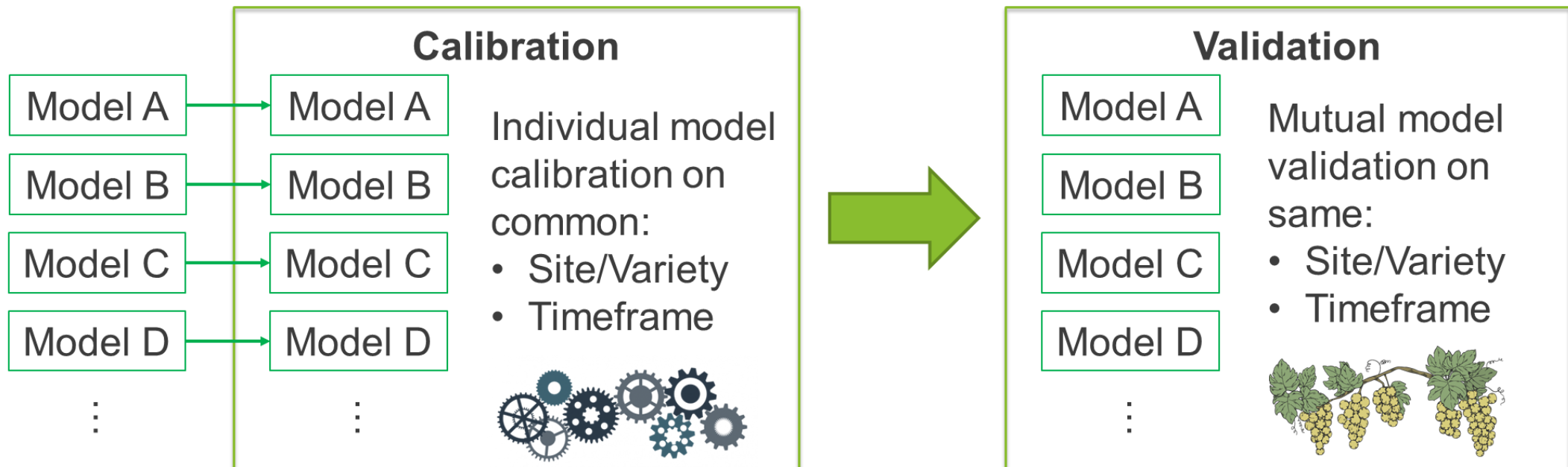
The good, the Bad and the Ugly, Sergio Leone,
Produzioni Europee Associate/United Artists

Introduction

Intercomparison Framework

Create a **Framework** for **unified** model calibration and validation

- **Multiple models** able to simulate the same variety, site and phenological phases
- **Multiple sites/varieties** to calibration and validate the models
- **Long time frames** to conscientiously test model sensitivities (30+ years)



Participating Models

- 5 compound models based on accumulation of GDD

Model A

Model B

Model C

Model D

⋮



STICS	GDD and BRIN	Temperature
vineyard	GDD (add. Thresholds)	Temperature
UniFi-C1	Sigmoid GDD and Gaussian-like GDD	Temperature
UniFi-C2	Sigmoid GDD	Temperature
PhenologyVvMoth	GDD (add. Thresholds), approx. daily cycle	Temperature Precipitation Global Radiation Evapotranspiration

Participating Models

Standard: $GDD = \max(T - T_{lower}, 0)$

- 5 compound models based on accumulation of GDD
- Different complexity of GDD functions

Model A

Model B

Model C

Model D

⋮

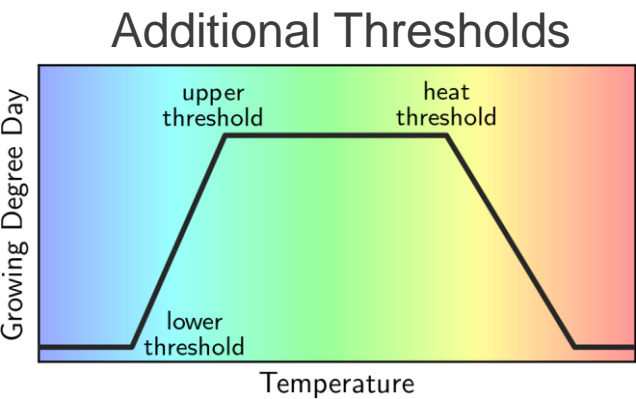


STICS	GDD and BRIN	Temperature
vineyard	GDD (add. Thresholds)	Temperature
UniFi-C1	Sigmoid GDD and Gaussian-like GDD	Temperature
UniFi-C2	Sigmoid GDD	Temperature
PhenologyVvMoth	GDD (add. Thresholds), approx. daily cycle	Temperature Precipitation Global Radiation Evapotranspiration

Participating Models

- 5 compound models based on accumulation of GDD
- Different complexity of GDD functions

Standard: $GDD = \max(T - T_{lower}, 0)$



Model A

Model B

Model C

Model D

⋮

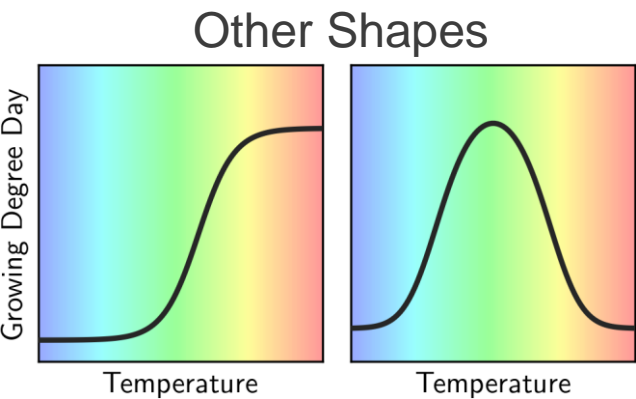


STICS	GDD and BRIN	Temperature
vineyard	GDD (add. Thresholds)	Temperature
UniFi-C1	Sigmoid GDD and Gaussian-like GDD	Temperature
UniFi-C2	Sigmoid GDD	Temperature
PhenologyVvMoth	GDD (add. Thresholds), approx. daily cycle	Temperature Precipitation Global Radiation Evapotranspiration

Participating Models

- 5 compound models based on accumulation of GDD
- Different complexity of GDD functions

Standard: $GDD = \max(T - T_{lower}, 0)$



Model A

Model B

Model C

Model D

⋮

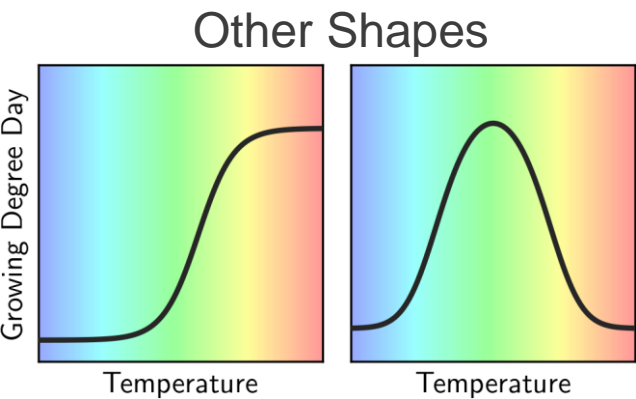


STICS	GDD and BRIN	Temperature
vineyard	GDD (add. Thresholds)	Temperature
UniFi-C1	Sigmoid GDD and Gaussian-like GDD	Temperature
UniFi-C2	Sigmoid GDD	Temperature
PhenologyVvMoth	GDD (add. Thresholds), approx. daily cycle	Temperature Precipitation Global Radiation Evapotranspiration

Participating Models

- 5 compound models based on accumulation of GDD
- Different complexity of GDD functions
- Mostly based on temperature

Standard: $GDD = \max(T - T_{lower}, 0)$



Model A

Model B

Model C

Model D

⋮

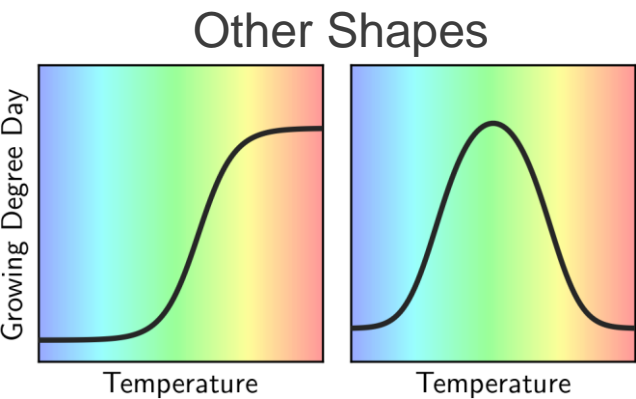


STICS	GDD and BRIN	Temperature
vineyard	GDD (add. Thresholds)	Temperature
UniFi-C1	Sigmoid GDD and Gaussian-like GDD	Temperature
UniFi-C2	Sigmoid GDD	Temperature
PhenologyVvMoth	GDD (add. Thresholds), approx. daily cycle	Temperature Precipitation Global Radiation Evapotranspiration

Participating Models

- 5 compound models based on accumulation of GDD
- Different complexity of GDD functions
- Mostly based on temperature

Standard: $GDD = \max(T - T_{lower}, 0)$



Model A

Model B

Model C

Model D

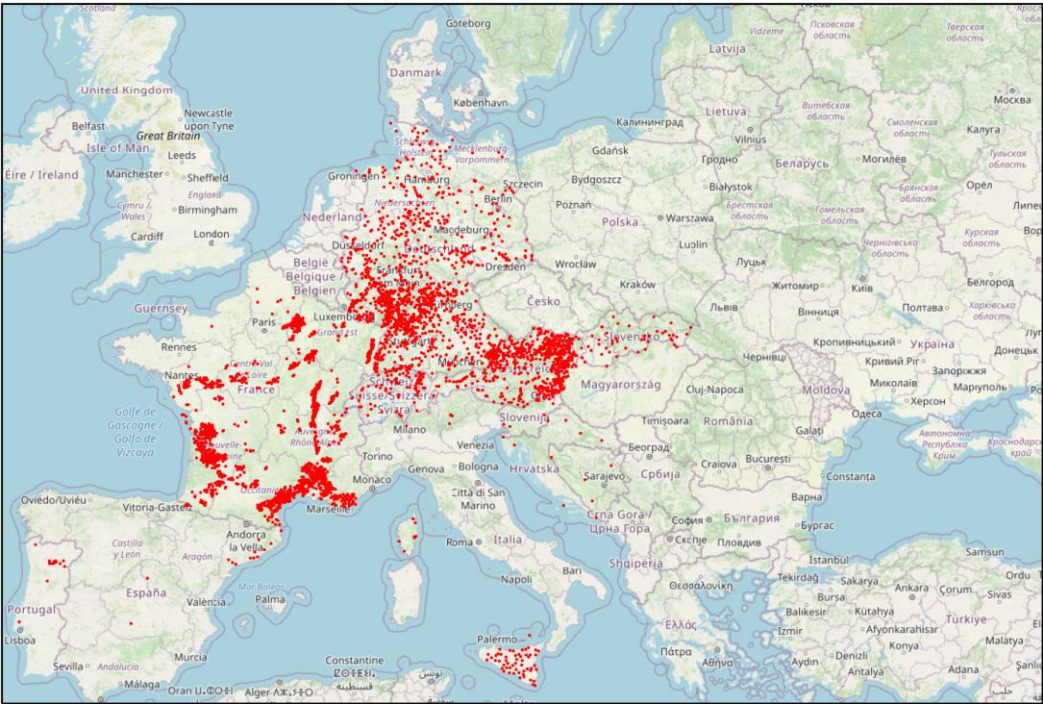
⋮



STICS	GDD and BRIN	Temperature
vineyard	GDD (add. Thresholds)	Temperature
UniFi-C1	Sigmoid GDD and Gaussian-like GDD	Temperature
UniFi-C2	Sigmoid GDD	Temperature
PhenologyVvMoth	GDD (add. Thresholds), approx. daily cycle	Temperature Precipitation Global Radiation Evapotranspiration

Sites and Varieties

- Multiple sites/varieties with long-term observations needed
- Several vineyard observations collected



Calibration

Model A

Model B

Model C

Model D

⋮

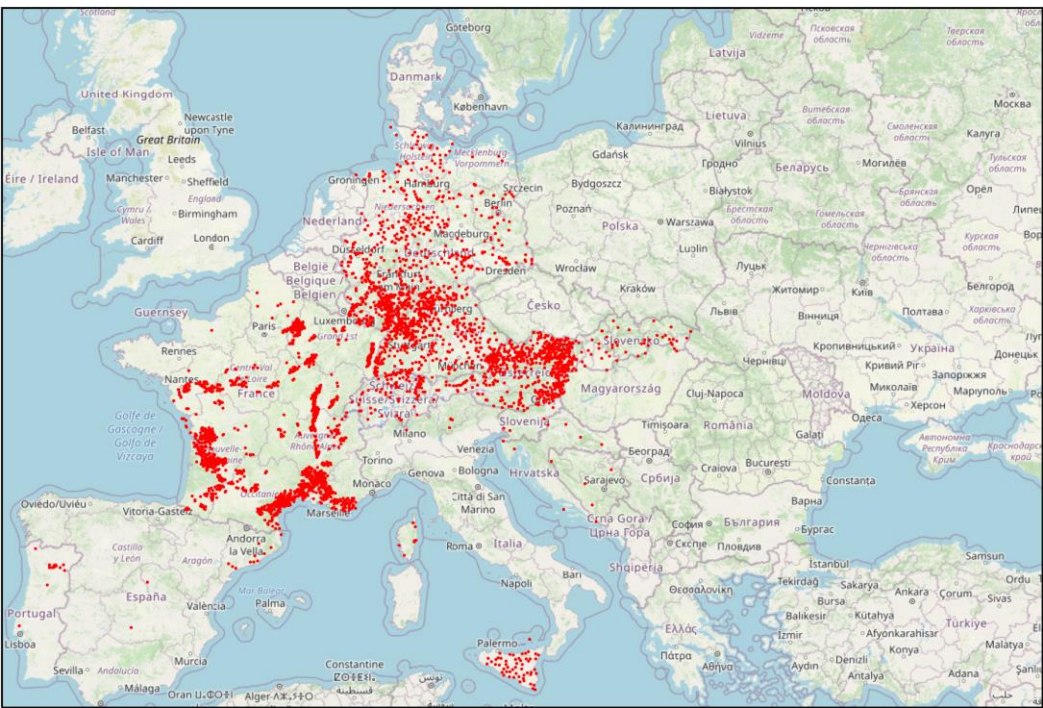
Individual model calibration on common:

- Site/Variety
- Timeframe

Sites	3801
Countries	16
Varieties	123

Sites and Varieties

- Multiple sites/varieties with long-term observations needed
- Several vineyard observations collected



Calibration

Model A

Model B

Model C

Model D

⋮

Individual model calibration on common:

- Site/Variety
- Timeframe

Sites	3801
Countries	16
Varieties	123

- Each station covers 7 BBCH stages on average
- On average 98.9% missing data
- On average less than 7 years

Sites and Varieties

- Multiple sites/varieties with long-term observations needed
- Several vineyard observations collected

Calibration

Model A

Model B


Model C

Model D

⋮

Individual model calibration on common:

- Site/Variety
- Timeframe



- More than 30 years
- Same phenological BBCH stages

Sites	3801
Countries	16
Varieties	123

- Each station covers 7 BBCH stages on average
- On average 98.9% missing data
- On average less than 7 years



Models and Data

Sites and Varieties

- 4 Central European sites representing temperate climate
- Cover almost 50 years of phenological observations for each site
- 3 different white wine varieties

Müller-Thurgau - MT



Rosenzweig (CC BY-SA 3.0)

Riesling - R



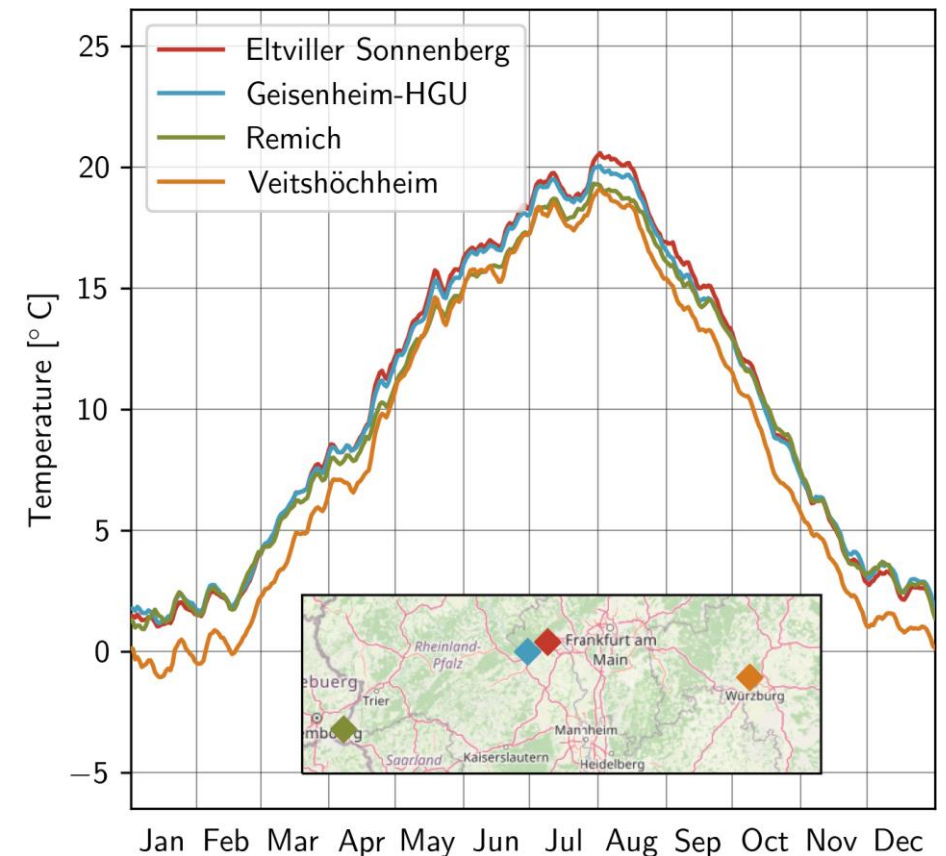
Bauer Karl (CC BY 2.0 AT)

Silvaner - S



Däisd (CC BY-SA 3.0)

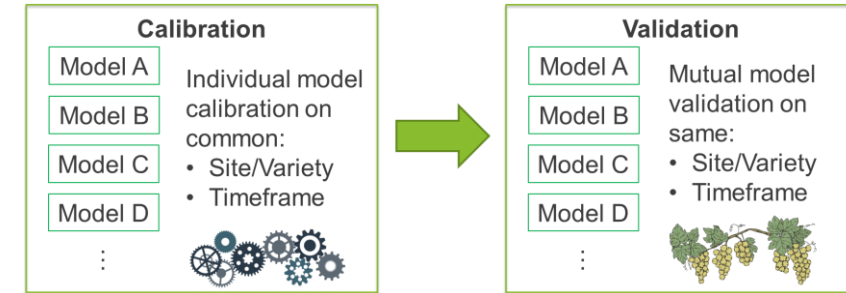
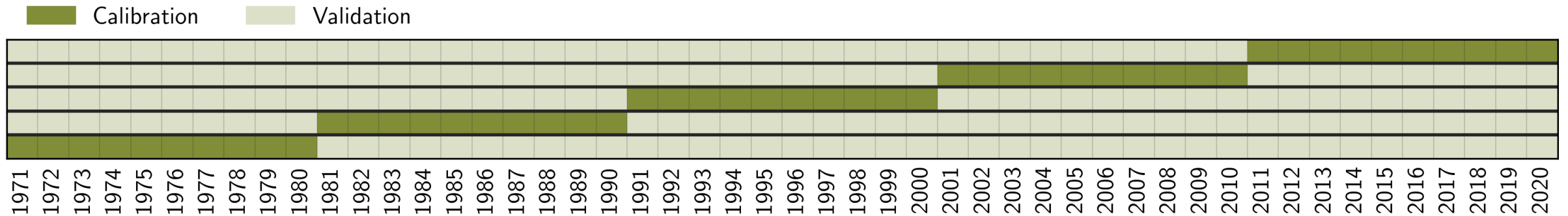
Site	Timeframe	# BBCH	Variety
Eltviller Sonnenberg	1955 – 2020	8	MT, R
Geisenheim-HGU	1951 – 2003	7	R
Remich	1968 – 2020	6	MT
Veitshöchheim	1968 – 2020	21	MT, R, S



Simulations

Cross-Validation

- Multiple separate calibration and validation timeframes



- 5 timeframes for each period length (5-fold cross-validation)
- 10a, 20a and 30a calibration period lengths
- 7 site/variety combinations

= 105 Simulations of each individual models
= 76263 years in total



Results

BBCH

bud break	09
flowering	68
begin ripening	81
end ripening	89



Research Questions

BBCH

bud break	09
flowering	68
begin ripening	81
end ripening	89



Results

Research Questions

Model developer:

- Model performance for different Sites/Varieties
- Model performance for different BBCH phases

BBCH

bud break	09
flowering	68
begin ripening	81
end ripening	89



Research Questions

Model developer:

- Model performance for different Sites/Varieties
- Model performance for different BBCH phases

Practitioners:

- How does the length of calibration period affect model performance?
- How well do the models perform in climate change conditions?

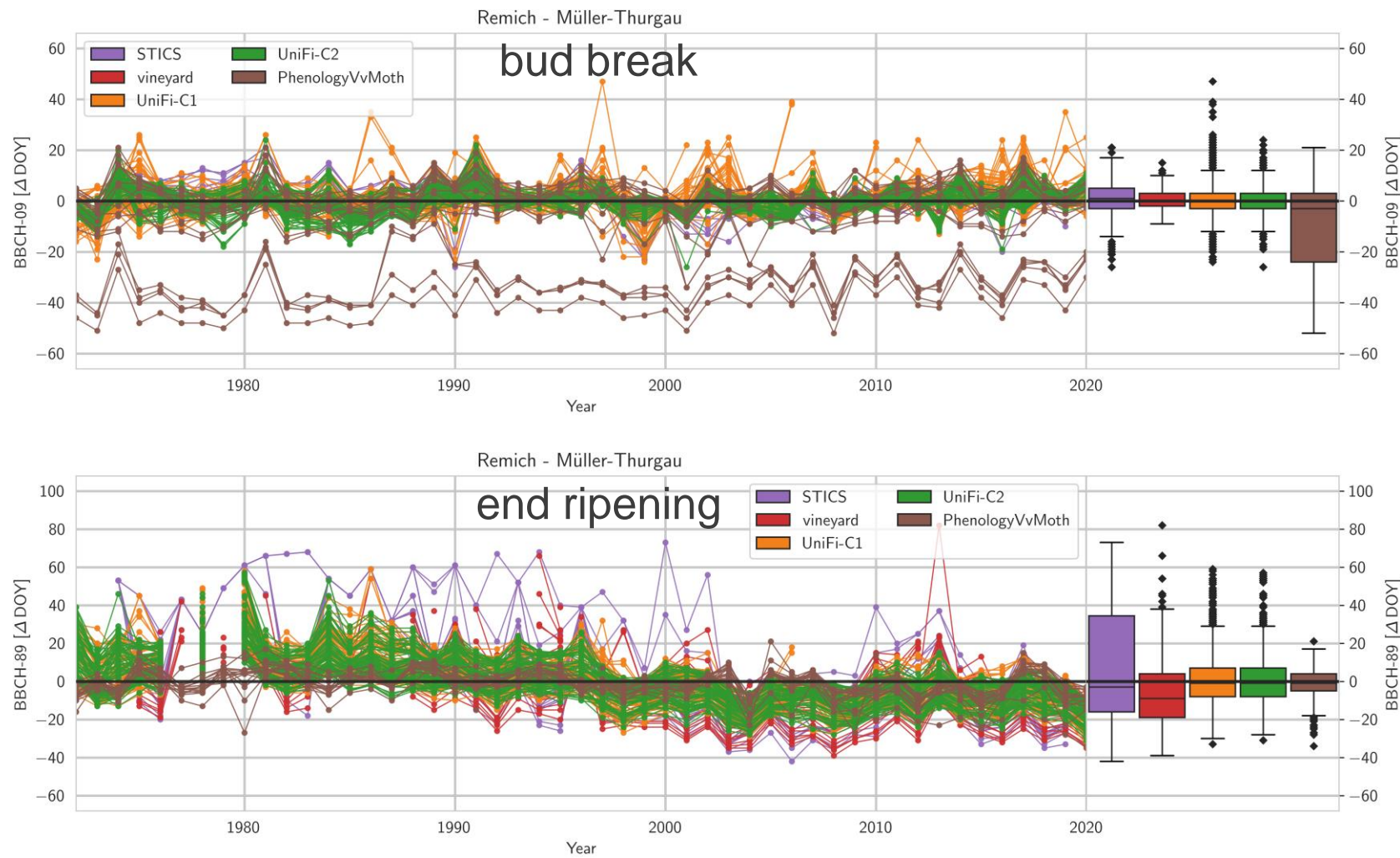
BBCH

bud break	09
flowering	68
begin ripening	81
end ripening	89



Results

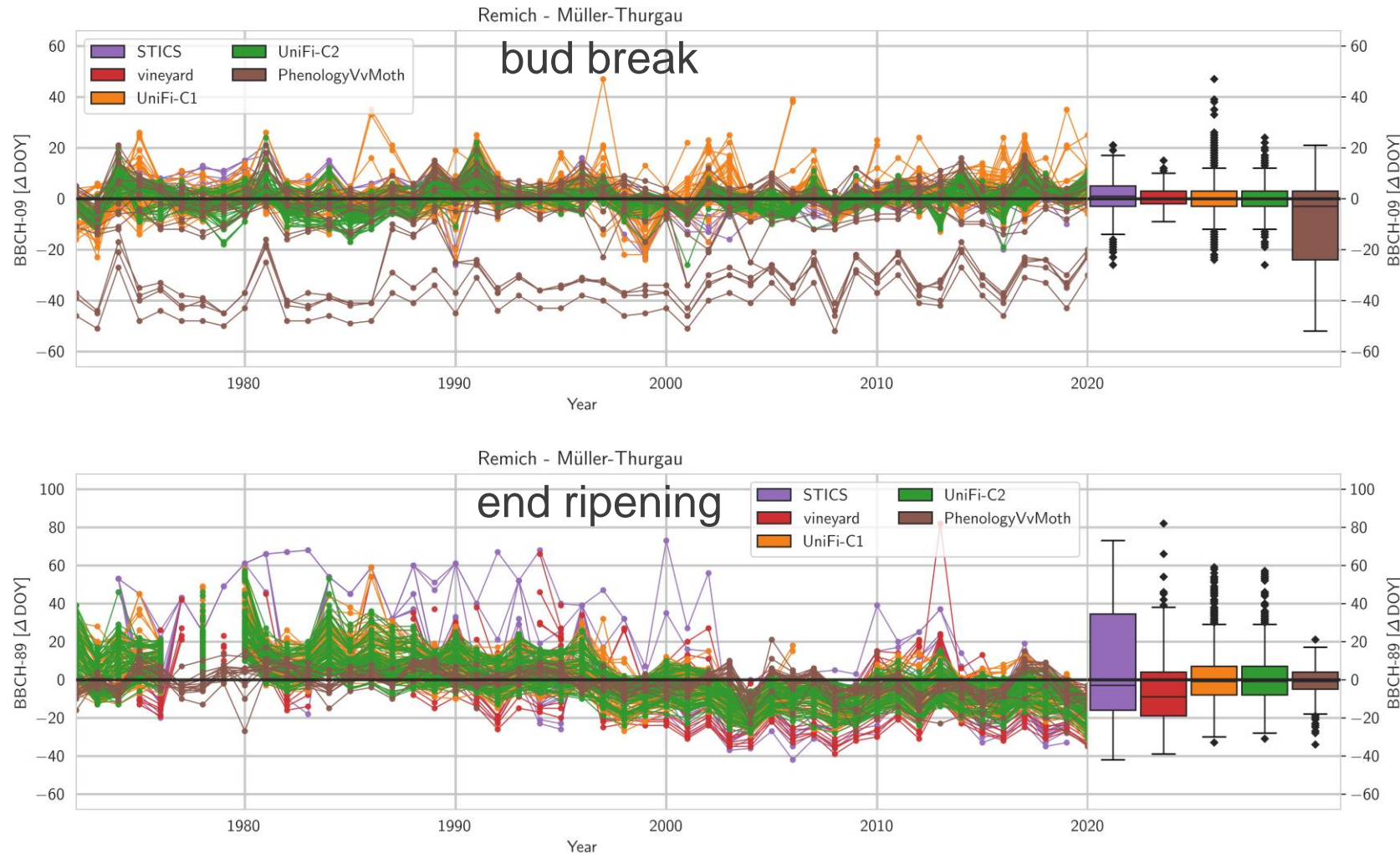
Overview



Results

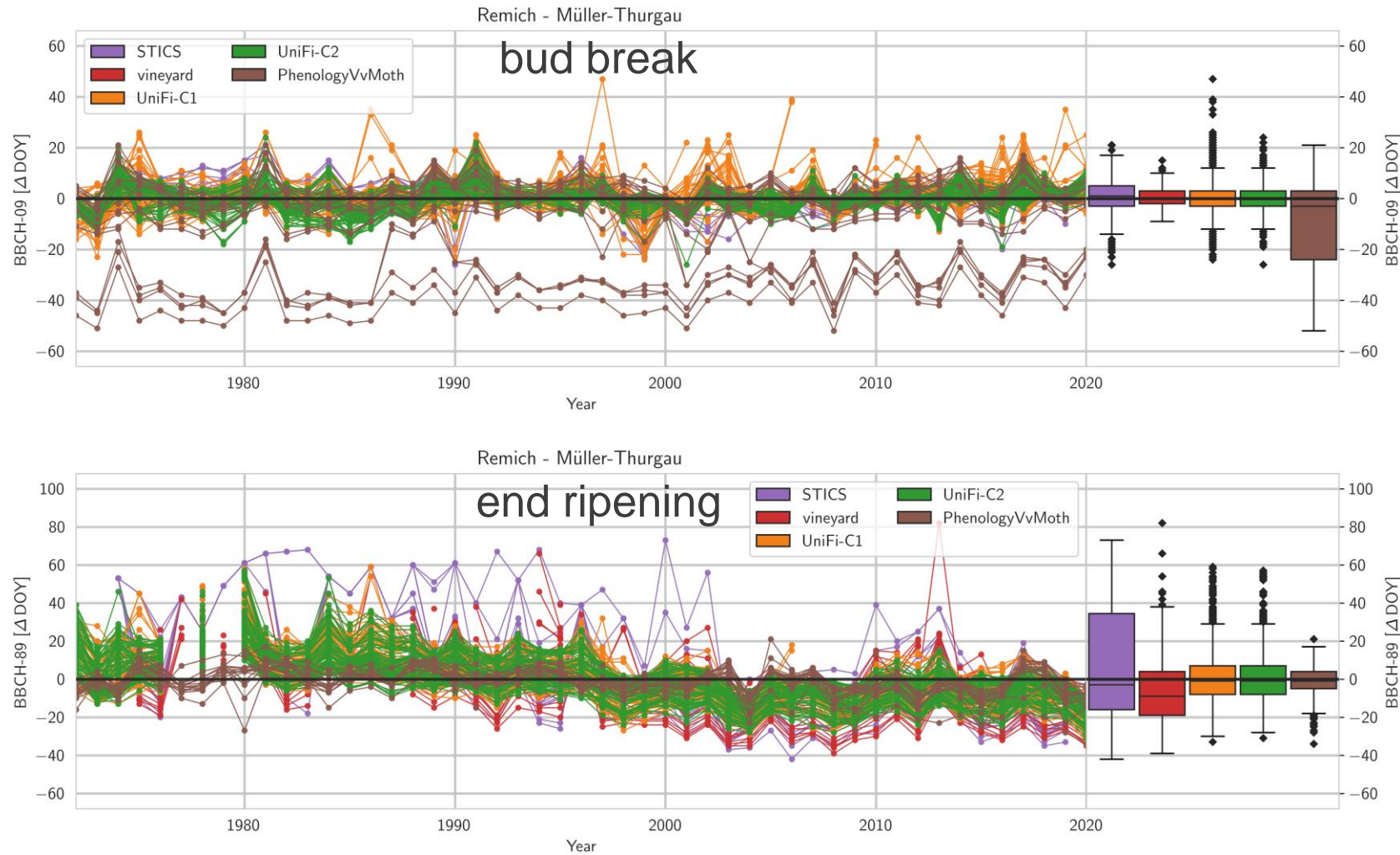
Overview

- Performance differ from model to model



Results

Overview

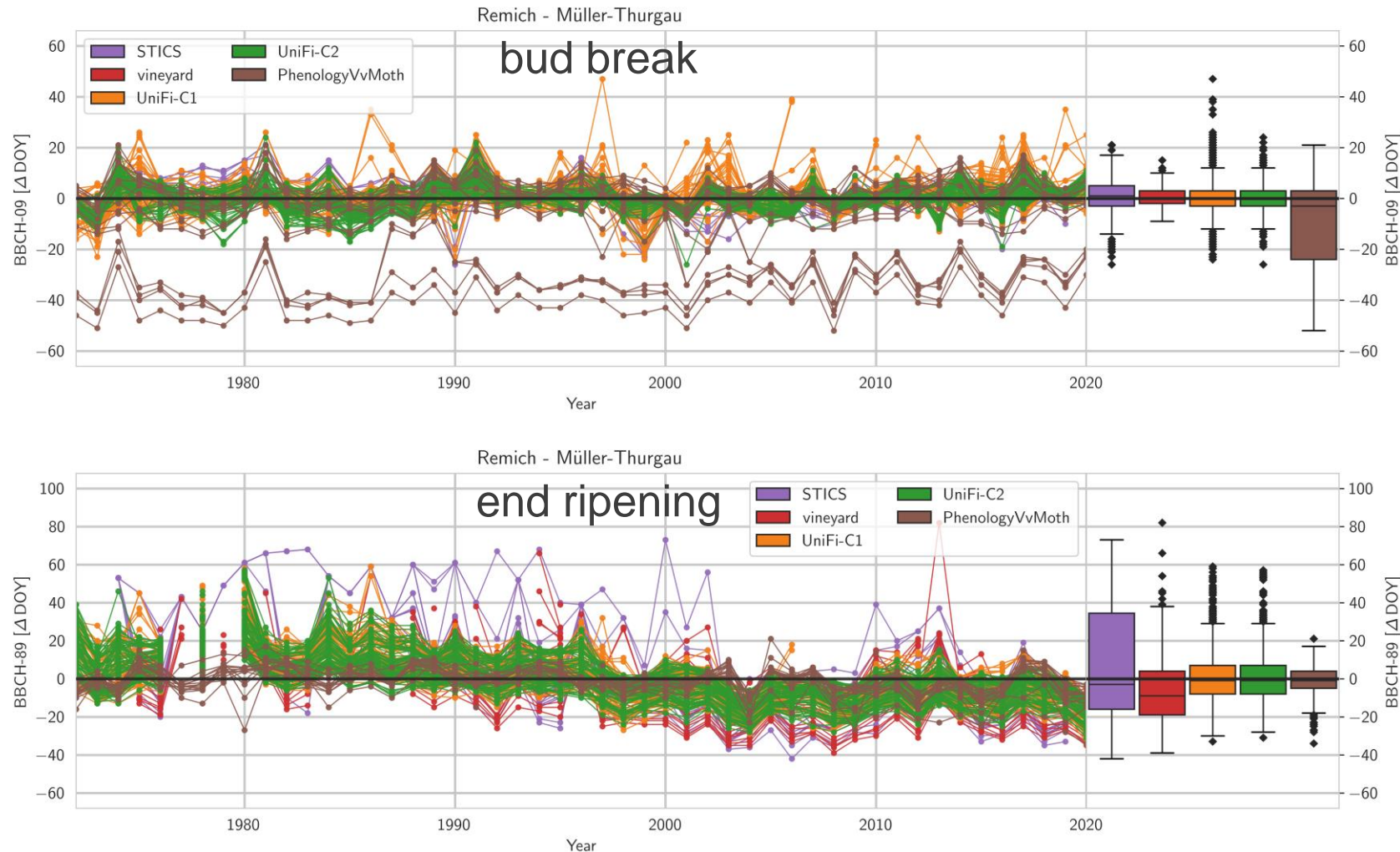


- Performance differ from model to model
- Low bias for most years and models

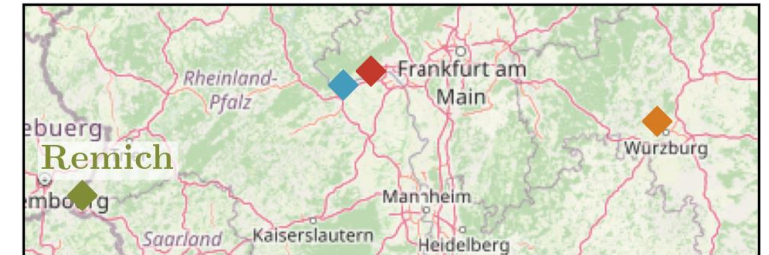


Results

Overview



- Performance differ from model to model
- Low bias for most years and models
- Calibration has a significant impact
- Outliers due to unique years or wrong calibration

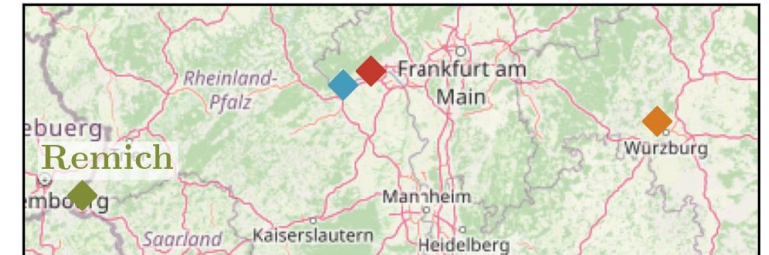


Results

Overview

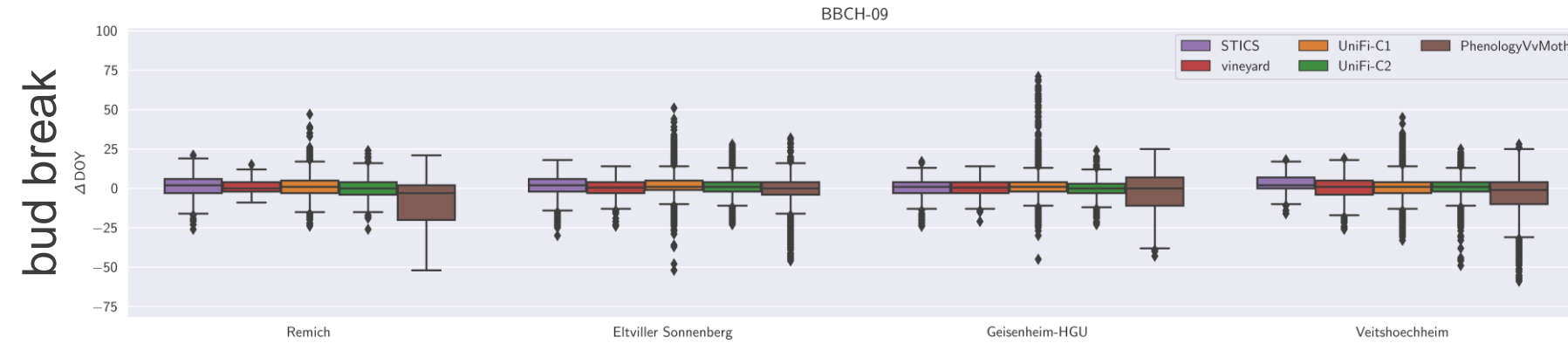


- Performance differ from model to model
- Low bias for most years and models
- Calibration has a significant impact
- Outliers due to unique years or wrong calibration
- Bias trend for BBCH-89



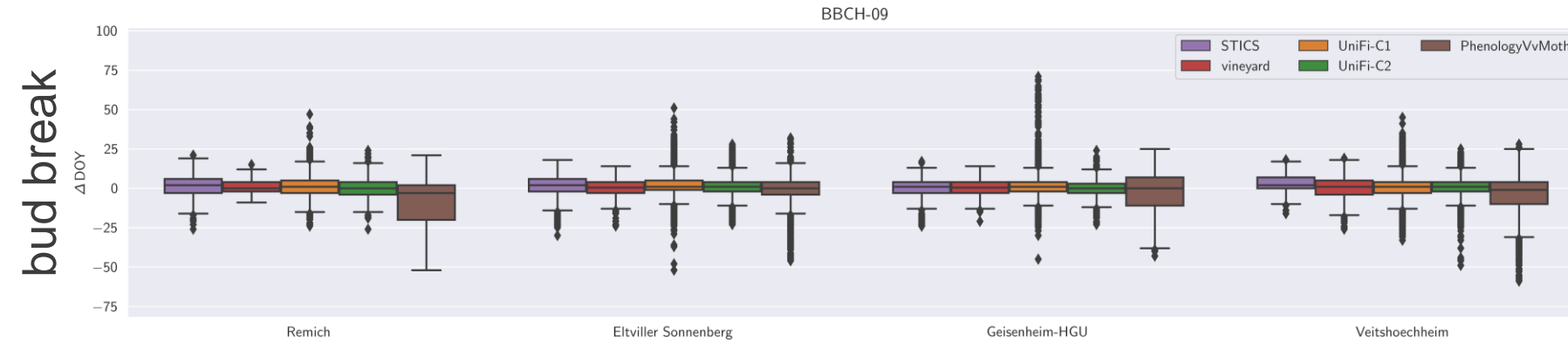
Results

Model Performance for different Sites/Varieties



Results

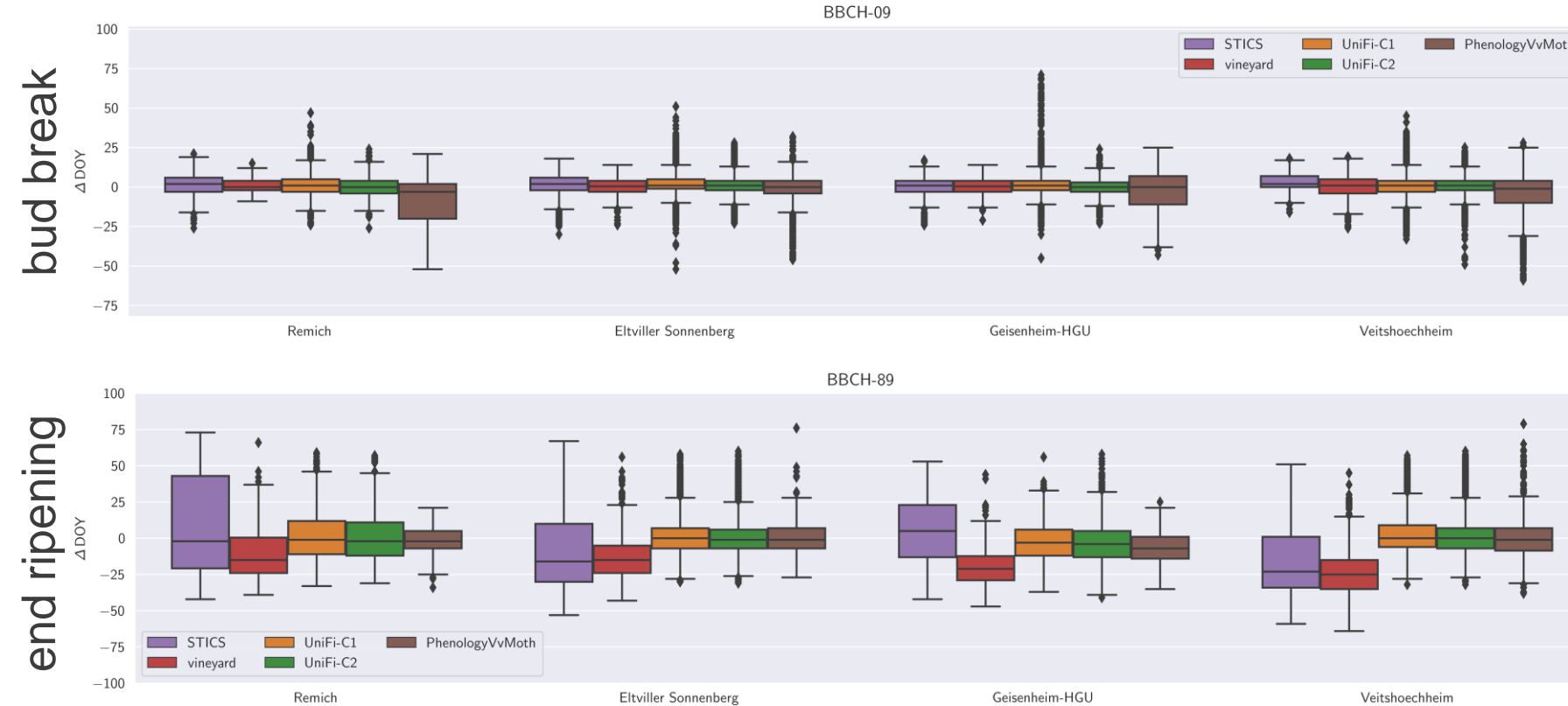
Model Performance for different Sites/Varieties



- Low sensitivity in general
- Mostly seen for outliers

Results

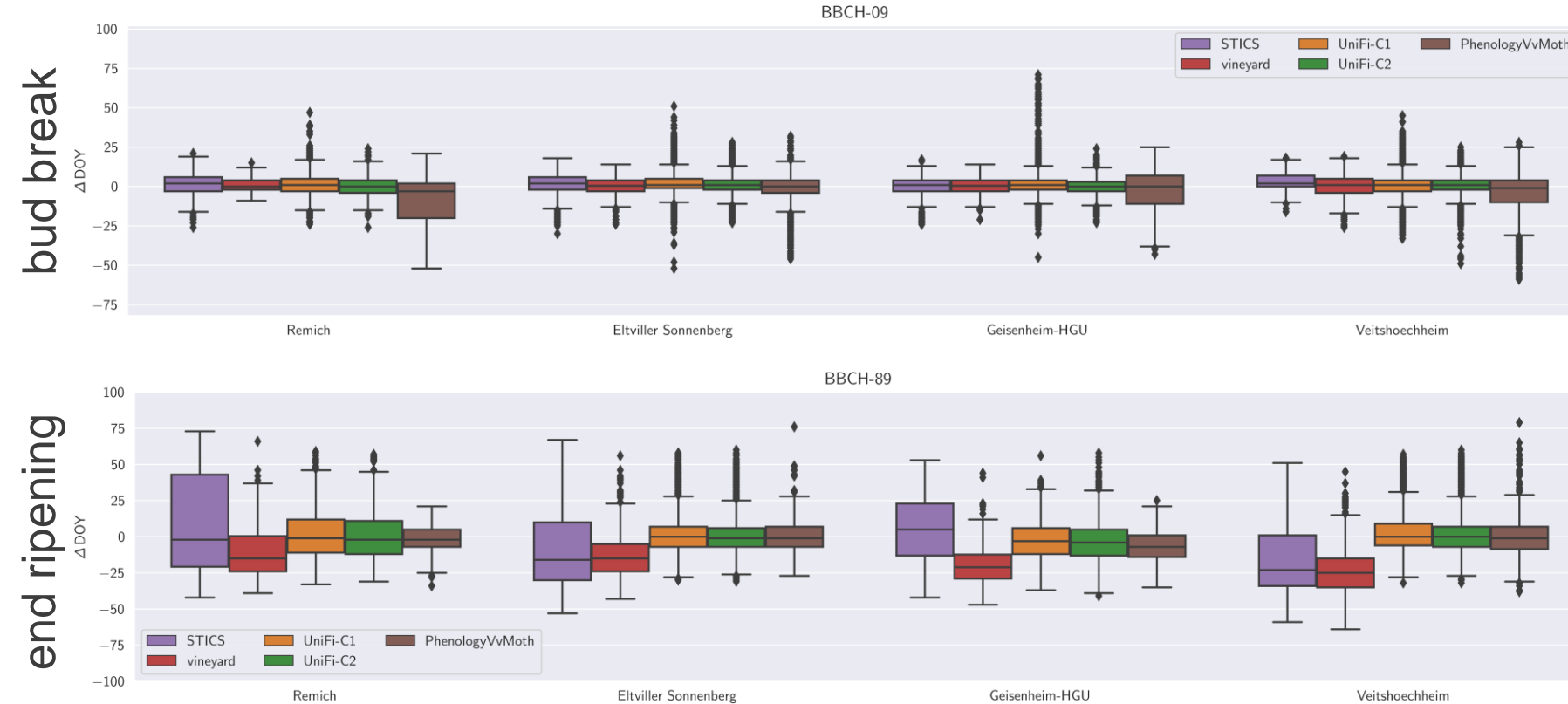
Model Performance for different Sites/Varieties



- Low sensitivity in general
- Mostly seen for outliers
- Higher sensitivity for BBCH-89

Results

Model Performance for different Sites/Varieties



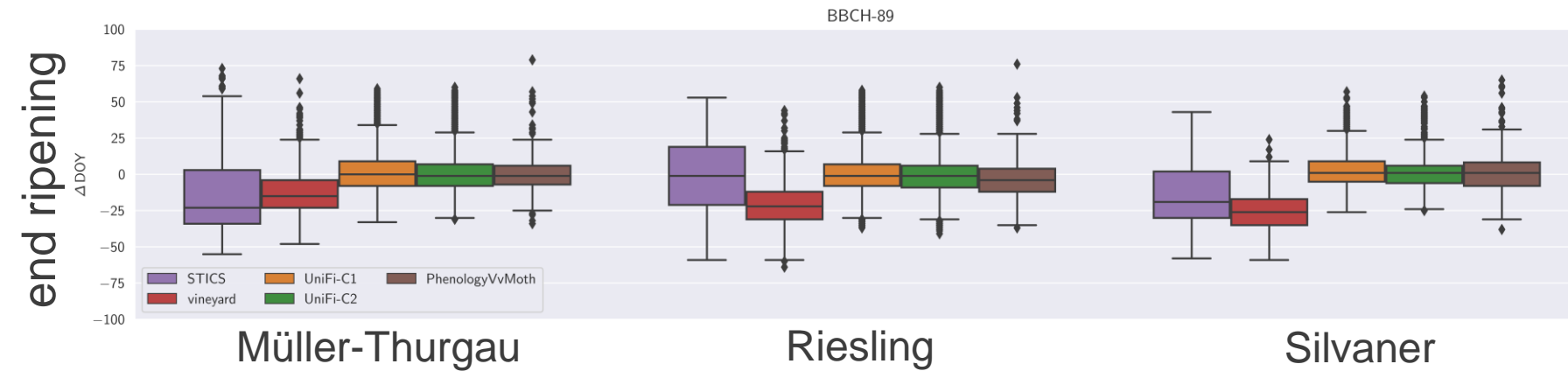
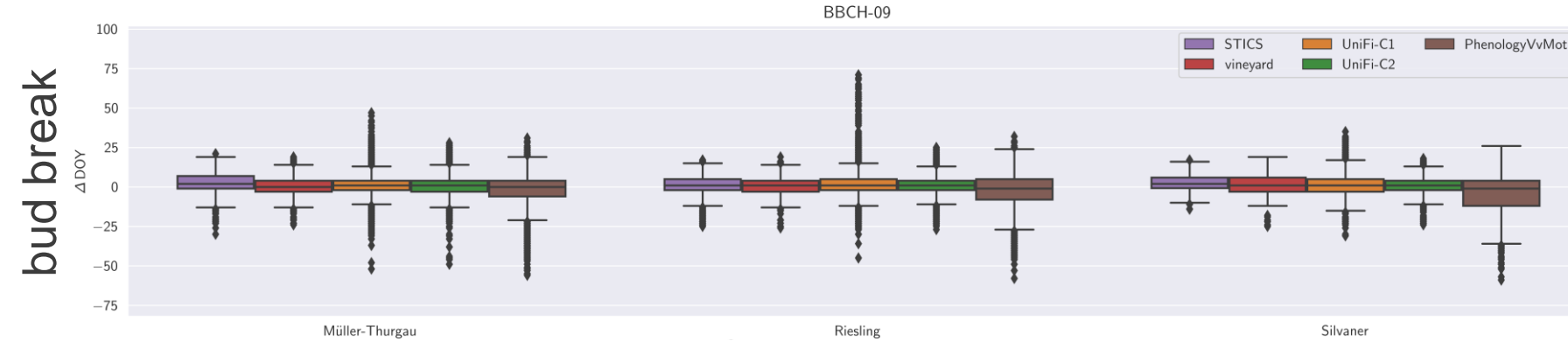
- Low sensitivity in general
- Mostly seen for outliers
- Higher sensitivity for BBCH-89
- STICS, vineyard and PhenologyVvMoth show higher sensitivity

Results

Model Performance for different Sites/Varieties



- Low sensitivity in general
- Mostly seen for outliers
- Higher sensitivity for BBCH-89
- STICS, vineyard and PhenologyVvMoth show higher sensitivity



Rosenzweig (CC BY-SA 3.0)



Bauer Karl (CC BY 2.0 AT)



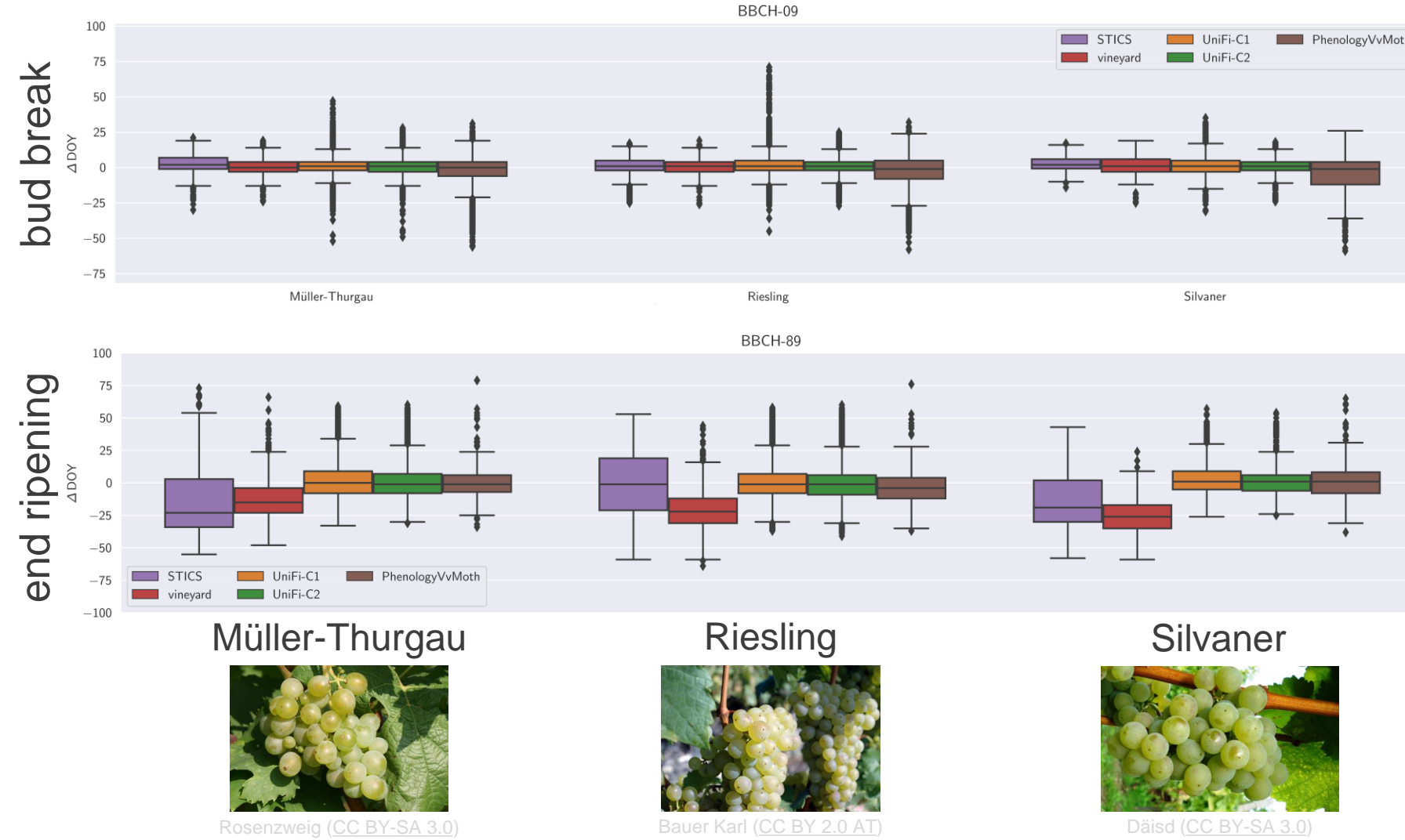
Däisd (CC BY-SA 3.0)

Results

Model Performance for different Sites/Varieties

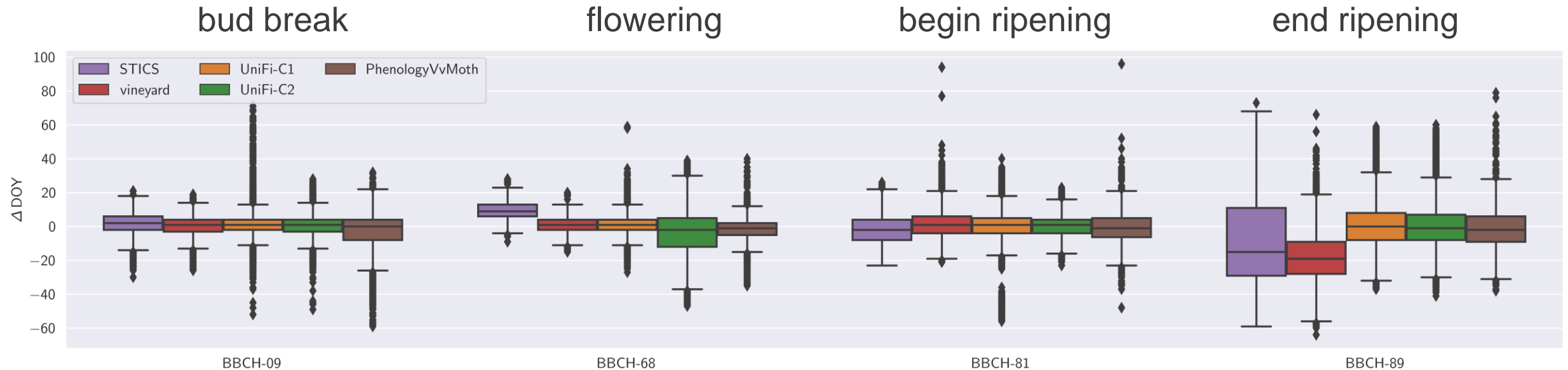


- Low sensitivity in general
- Mostly seen for outliers
- Higher sensitivity for BBCH-89
- STICS, vineyard and PhenologyVvMoth show higher sensitivity
- Models generalize over different sites and varieties



Results

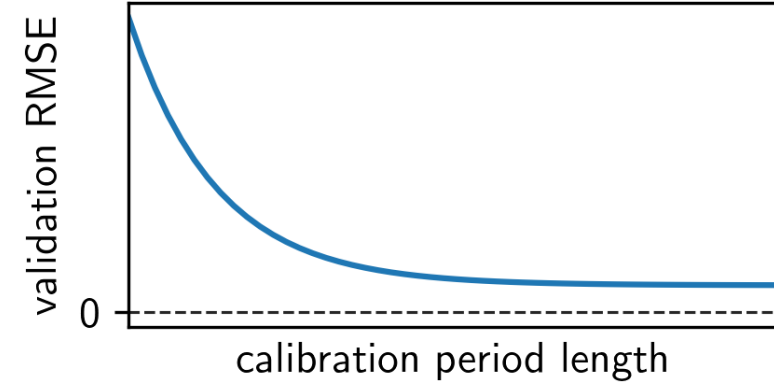
Model performance for different BBCH phases



- Model performance differs from phase to phase
- Bias compensation from one phase to another
- Slight qualitative change from BBCH-09 to BBCH-89

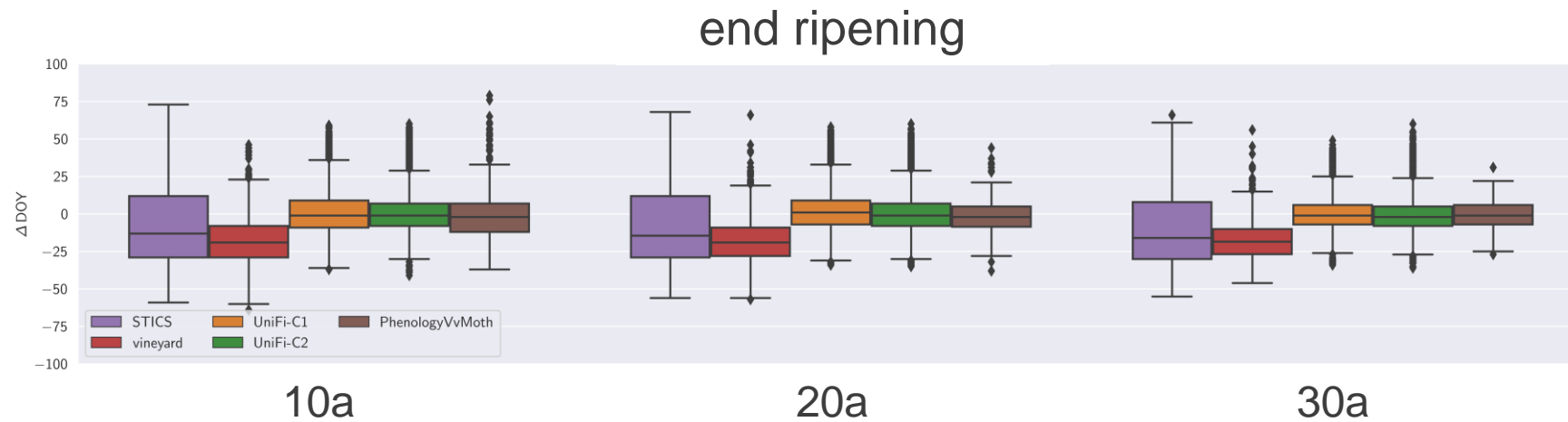
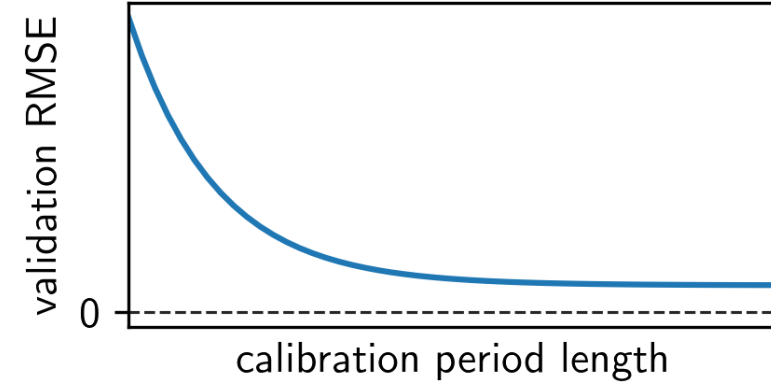
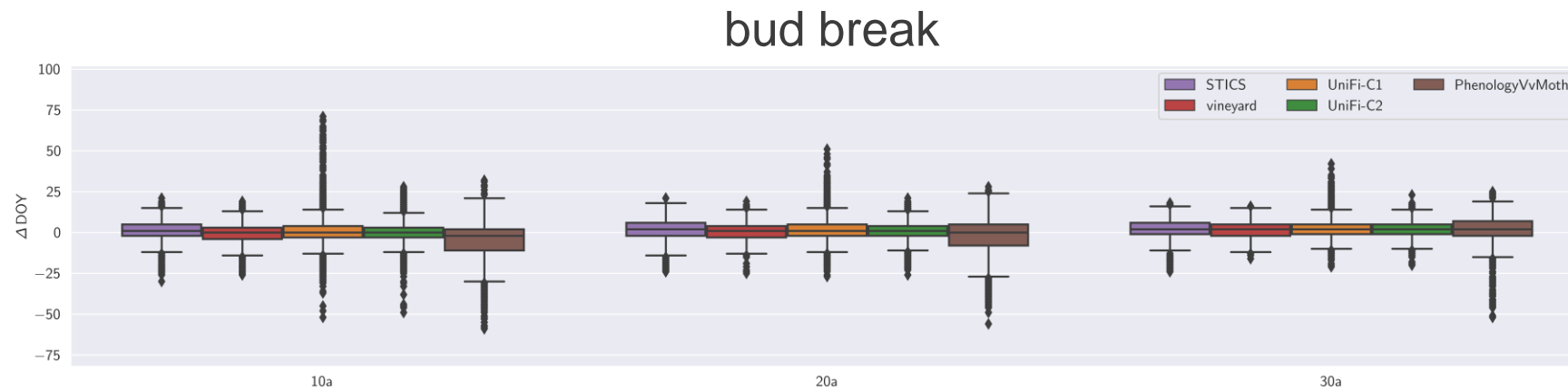
Results

How does the length of calibration period affect model performance?



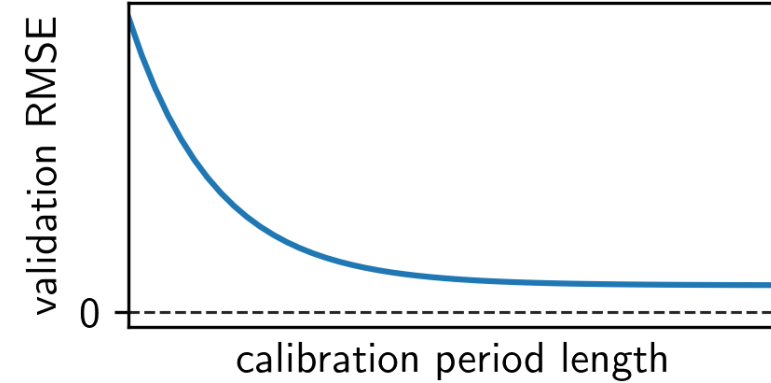
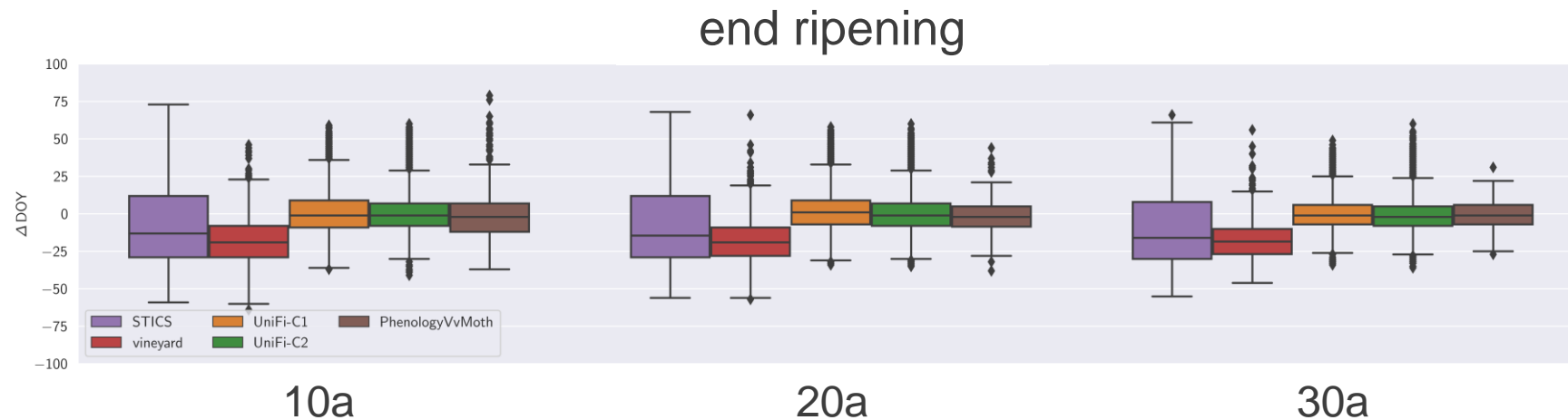
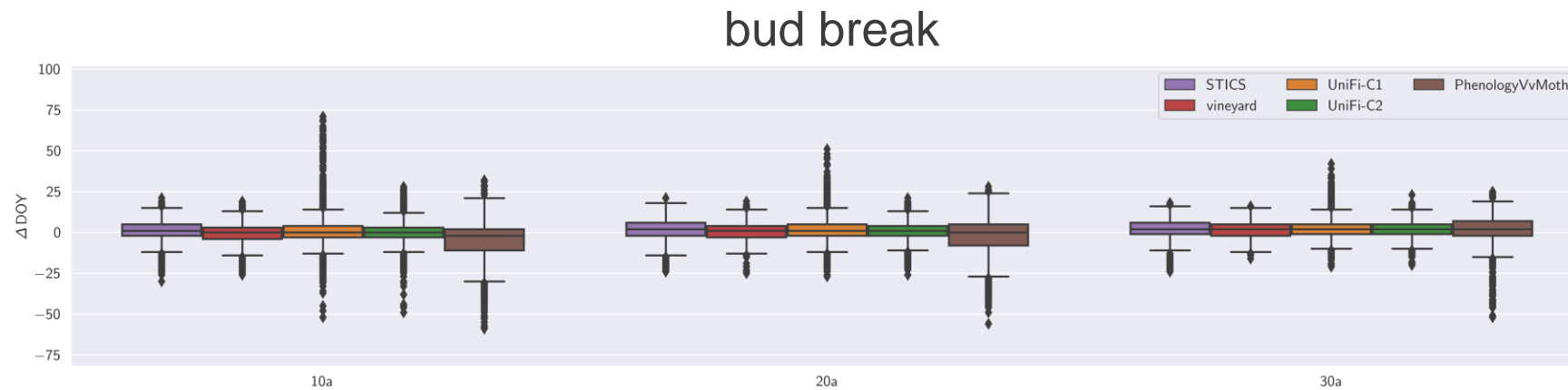
Results

How does the length of calibration period affect model performance?



Results

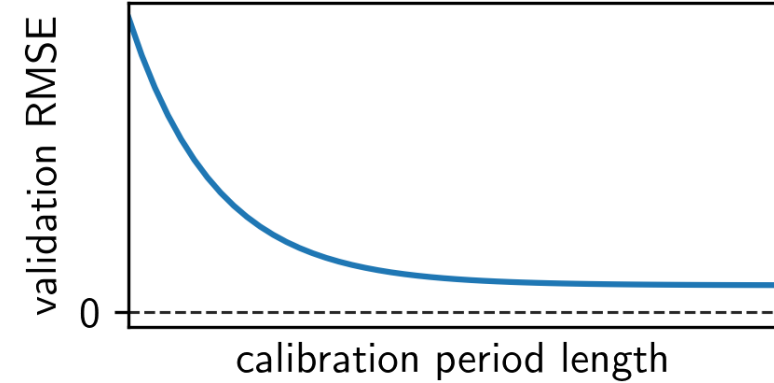
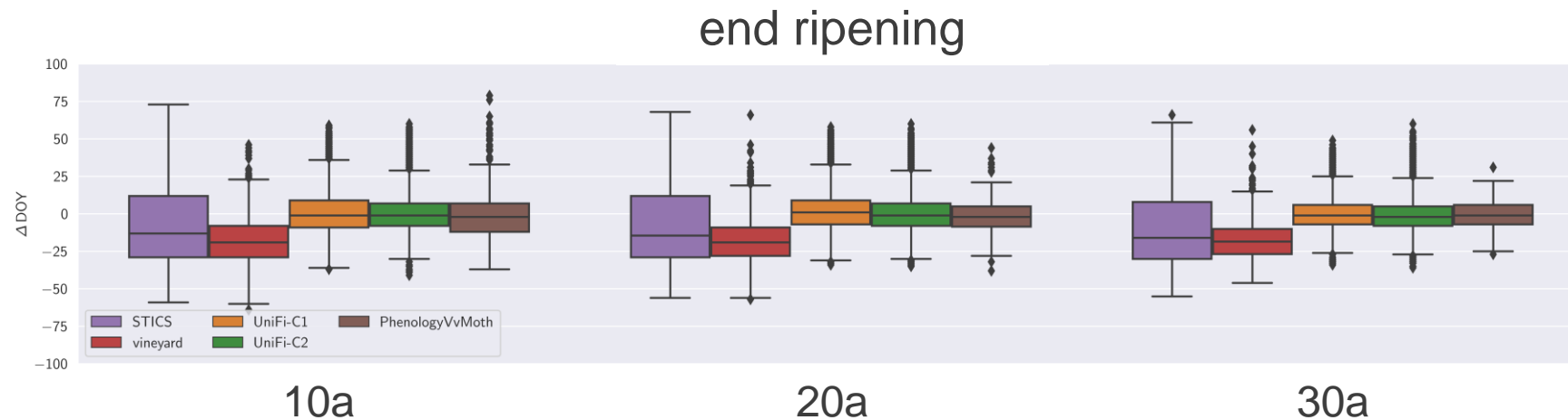
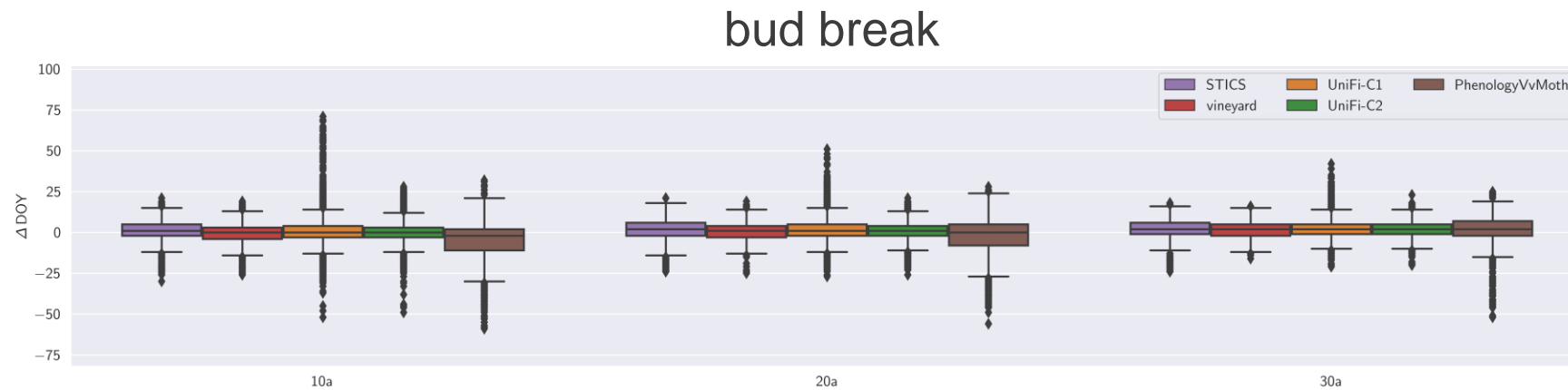
How does the length of calibration period affect model performance?



- Bias does not reduce significantly with increasing calibration period length

Results

How does the length of calibration period affect model performance?

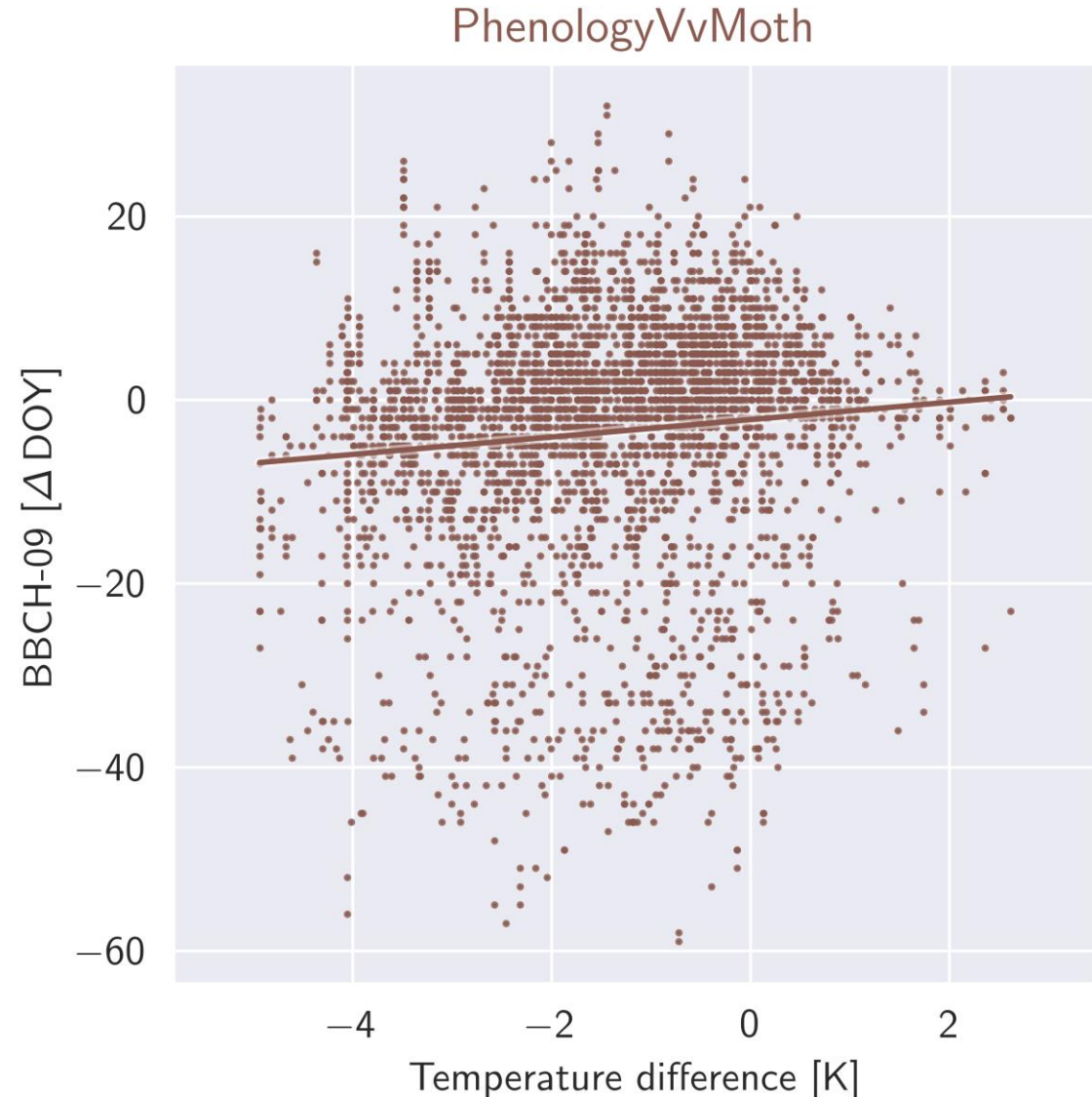


- Bias does not reduce significantly with increasing calibration period length
- Less than 10 years seem to be sufficient

Results

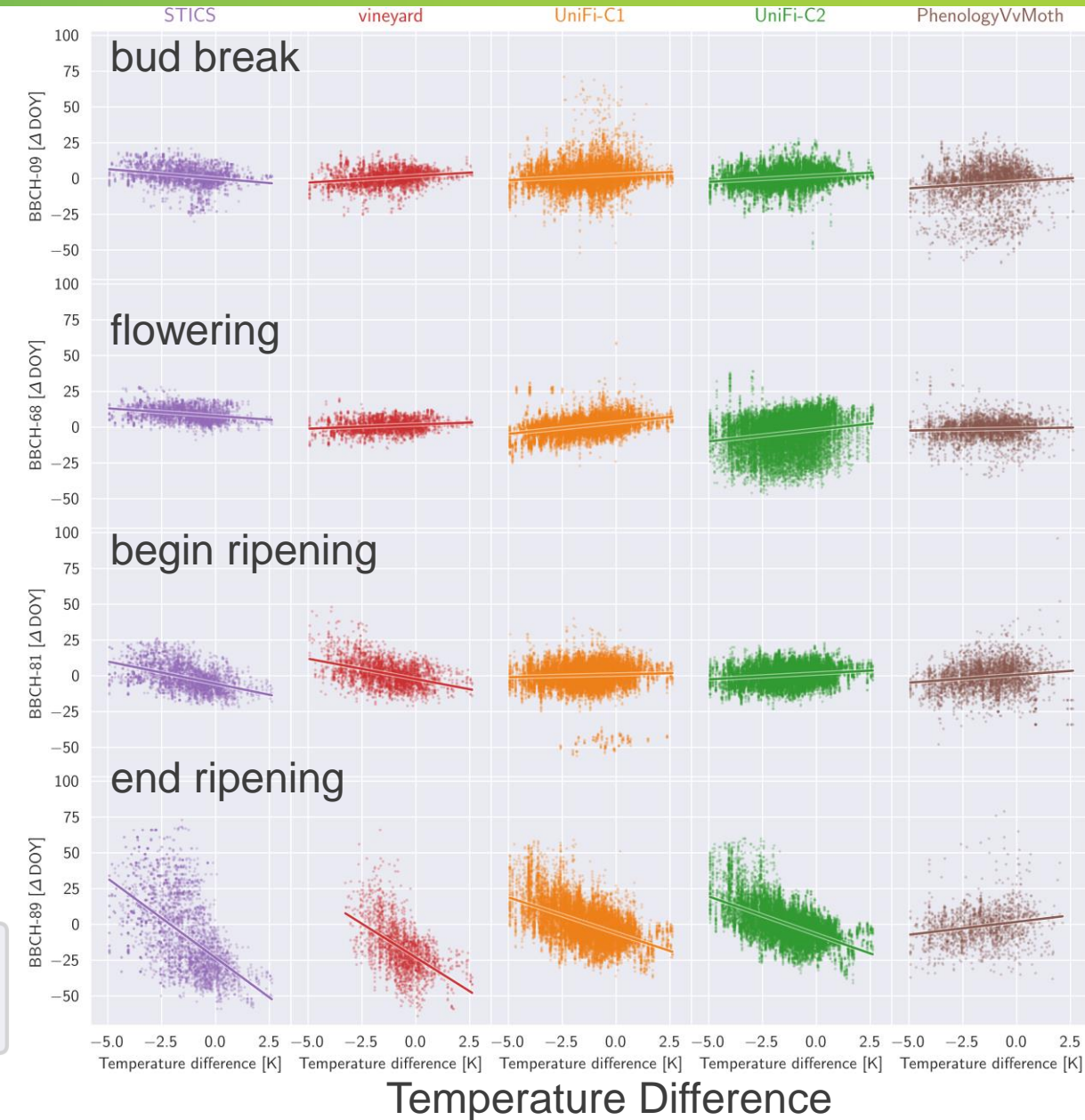
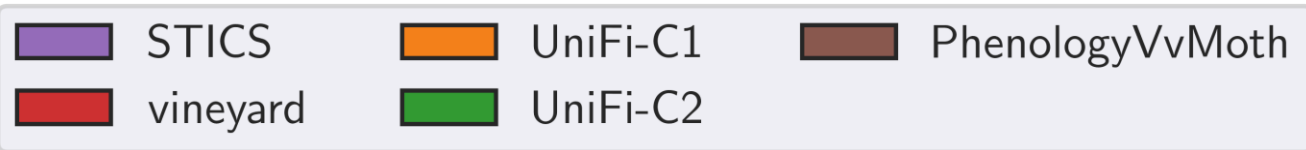
How well do the models perform in climate change conditions?

- Due to the long observation time frame, we can estimate the performance for different climate conditions
- Growing season: AMJJAS
- Temperature difference: individual year vs. hottest calibration year
- Bias should be independent of temperature difference



Results

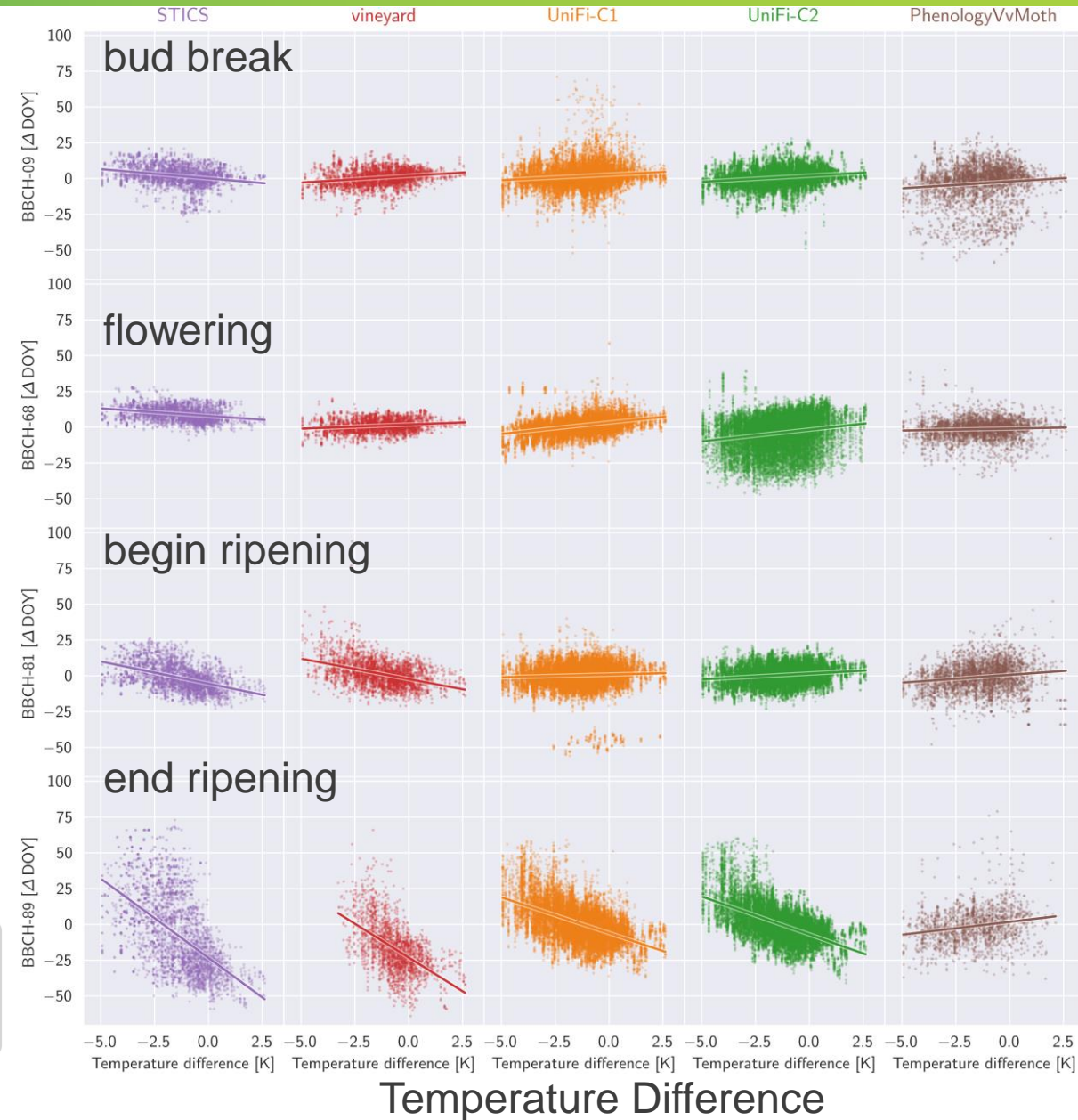
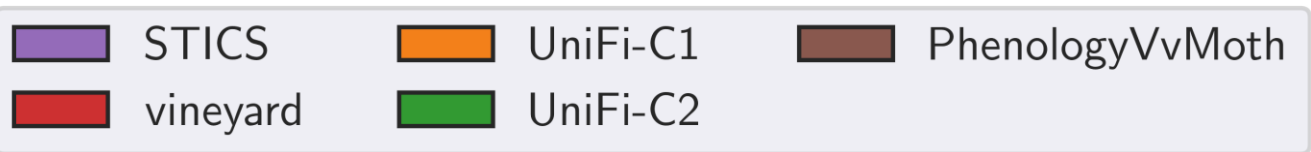
How well do the models perform in climate change conditions?



Results

How well do the models perform in climate change conditions?

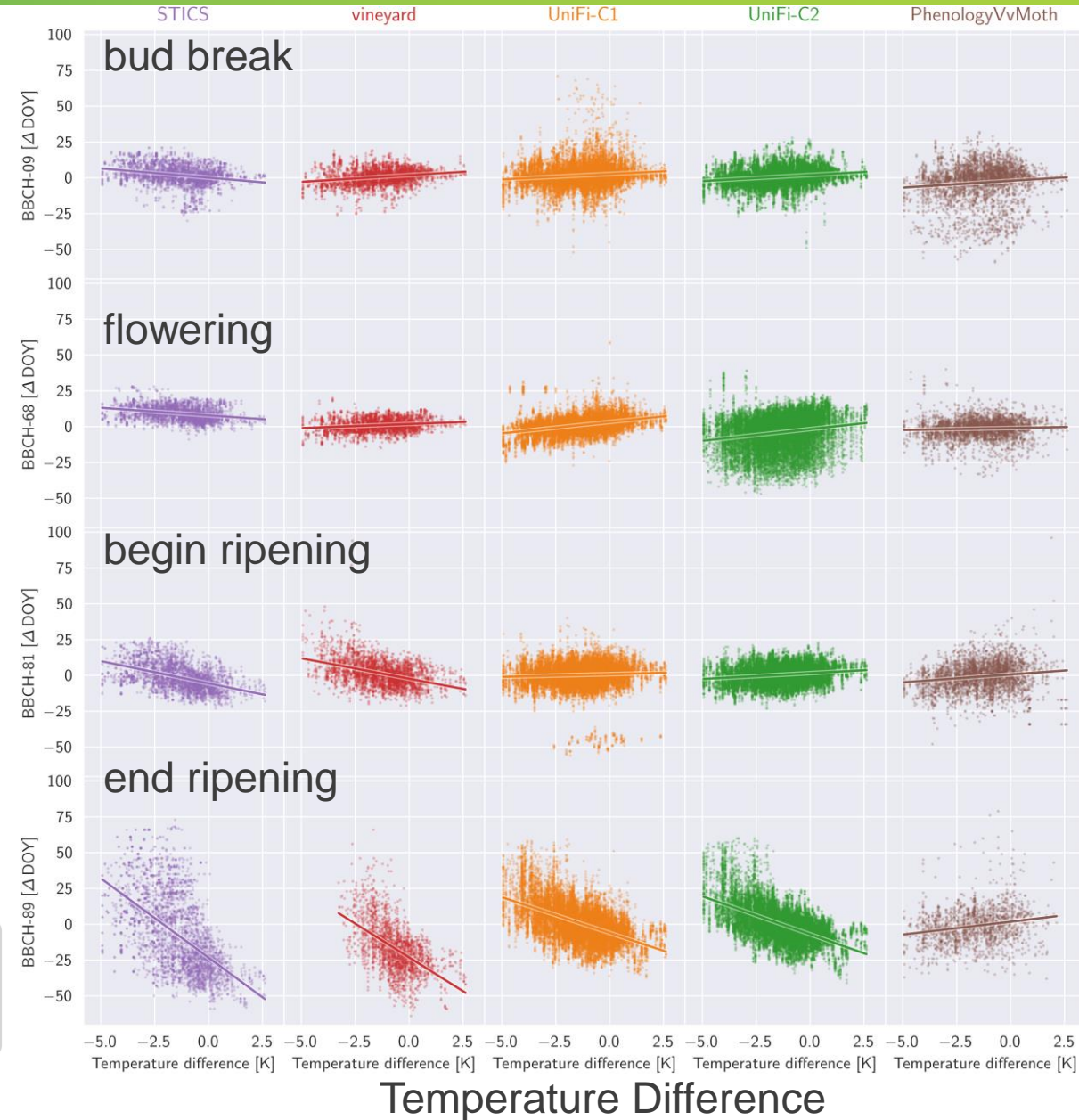
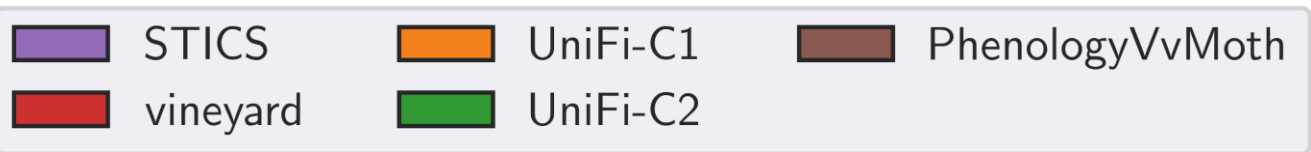
- Models appear to be applicable under climate change conditions for most phases



Results

How well do the models perform in climate change conditions?

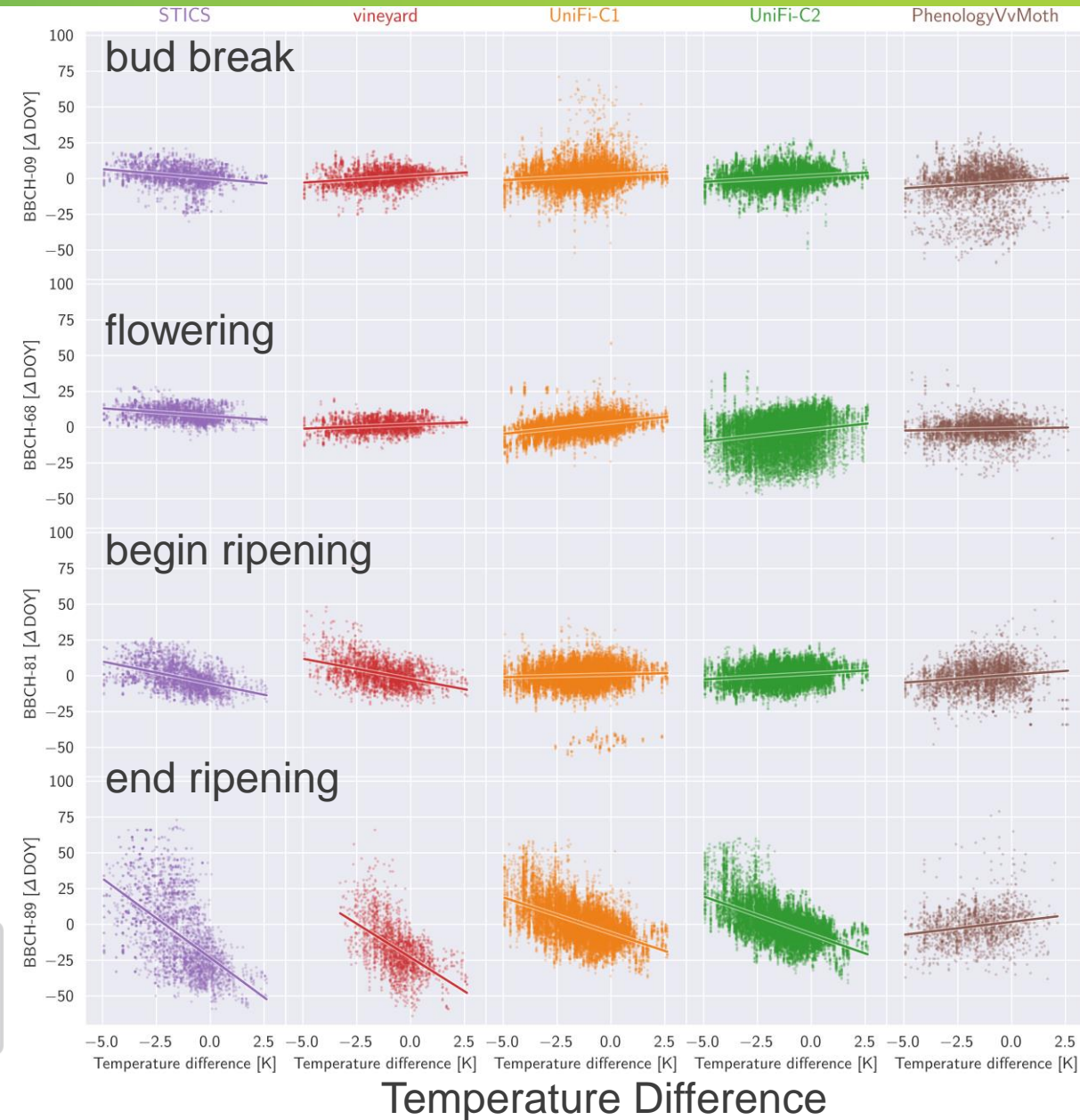
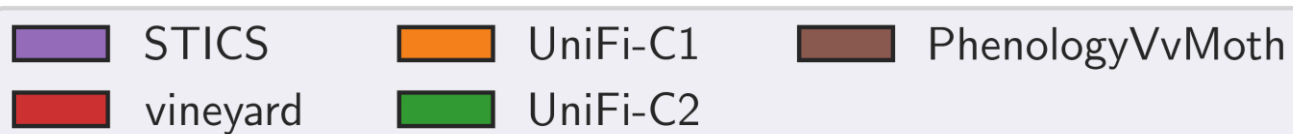
- Models appear to be applicable under climate change conditions for most phases
- Later phases show dependency



Results

How well do the models perform in climate change conditions?

- Models appear to be applicable under climate change conditions for most phases
- Later phases show dependency
- Advance of ripening phase might be overestimated by current models under climate change conditions



Conclusion



- Comprehensive model intercomparison of 5 models for grapevine phenology over long/climatological time frames

Conclusion



- Comprehensive model intercomparison of 5 models for grapevine phenology over long/climatological time frames
- All models perform reasonably well, no model with statistically significant bias in general

Conclusion



- Comprehensive model intercomparison of 5 models for grapevine phenology over long/climatological time frames
- All models perform reasonably well, no model with statistically significant bias in general
- Particular attention should be given to model calibration

Conclusion



- Comprehensive model intercomparison of 5 models for grapevine phenology over long/climatological time frames
- All models perform reasonably well, no model with statistically significant bias in general
- Particular attention should be given to model calibration
- Model development should concentrate on improving representation of phases rather than different sites/varieties

Conclusion



- Comprehensive model intercomparison of 5 models for grapevine phenology over long/climatological time frames
- All models perform reasonably well, no model with statistically significant bias in general
- Particular attention should be given to model calibration
- Model development should concentrate on improving representation of phases rather than different sites/varieties
- 10 years are sufficient for adequate model calibration

Conclusion



- Comprehensive model intercomparison of 5 models for grapevine phenology over long/climatological time frames
- All models perform reasonably well, no model with statistically significant bias in general
- Particular attention should be given to model calibration
- Model development should concentrate on improving representation of phases rather than different sites/varieties
- 10 years are sufficient for adequate model calibration
- Models permit climate change impact analysis for most phases except for ripening phase

Conclusion

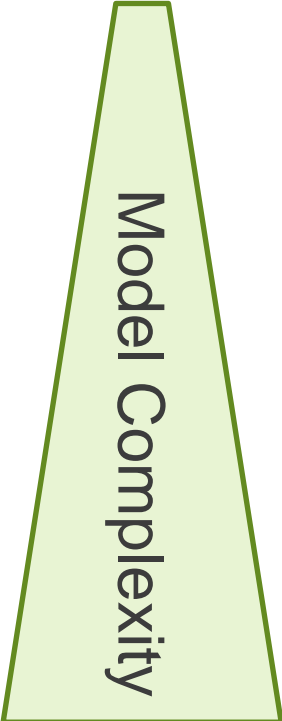
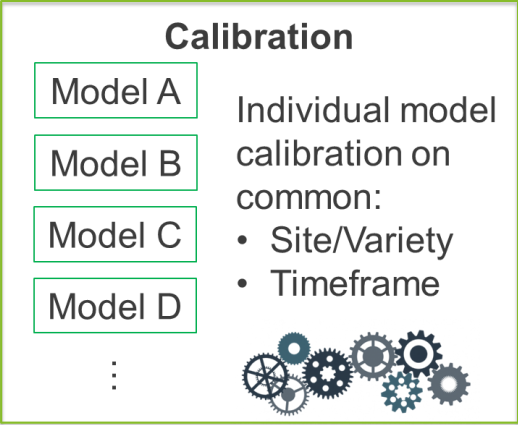


- Comprehensive model intercomparison of 5 models for grapevine phenology over long/climatological time frames
- All models perform reasonably well, no model with statistically significant bias in general
- Particular attention should be given to model calibration
- Model development should concentrate on improving representation of phases rather than different sites/varieties
- 10 years are sufficient for adequate model calibration
- Models permit climate change impact analysis for most phases except for ripening phase
- Current model projection might overestimate the advance of ripening phase

Clim4^{it}itis

Individual Model Calibration

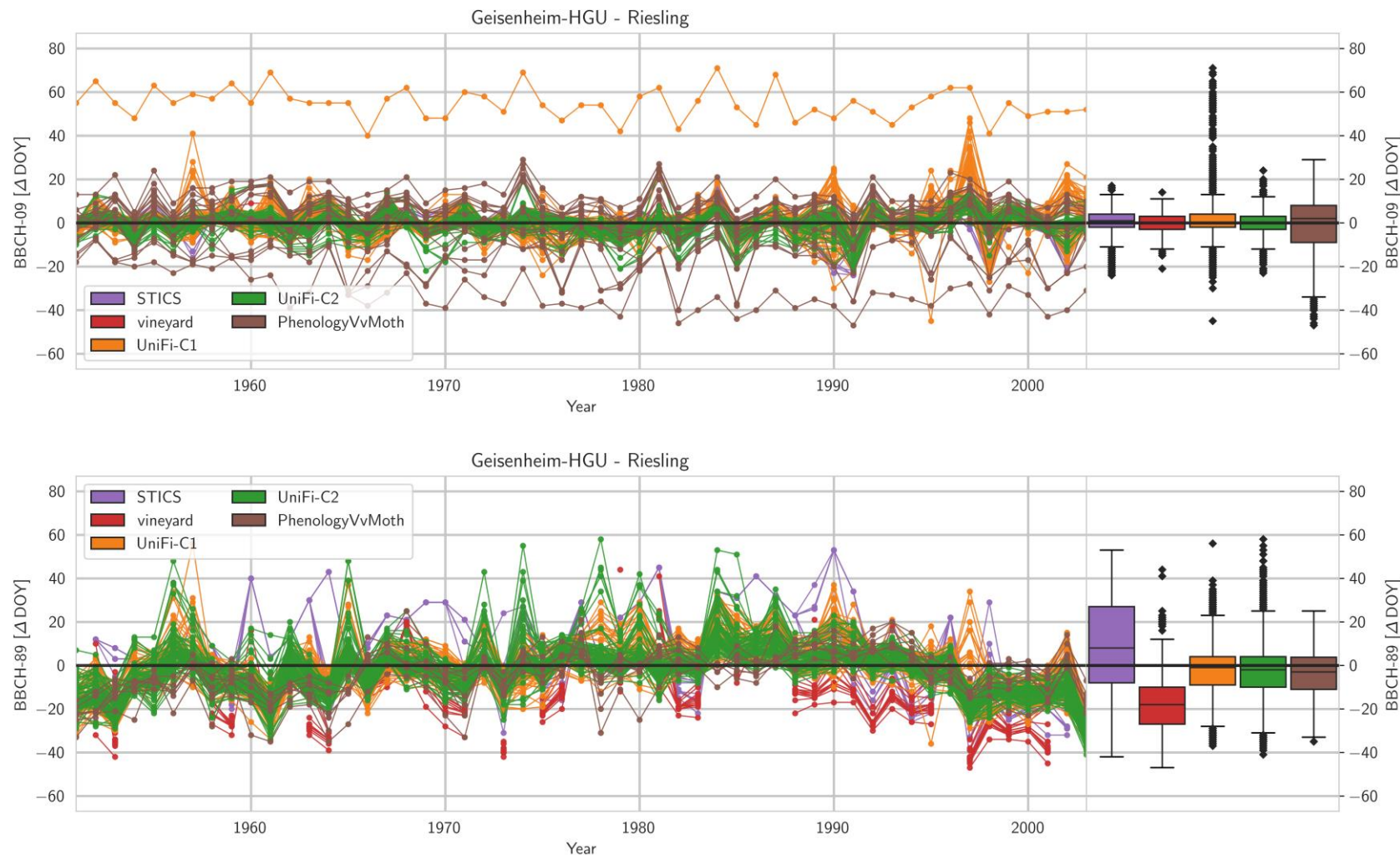
- Framework prescribes sites/variety and timeframes
- No unified calibration scheme prescribed
- Phases, calibration method and objective function chosen by each modeler



Model	# Parameters	Calibration Method	Objective Function
STICS	3	Grid search	MSE
Vineyard	3	SCE-UA	RMSE
UniFi-C2	12	Random search and simulated annealing	Various
UniFi-C1	17	Random search and simulated annealing	Various
PhenologyVvMoth	39	Simulated annealing and Nelder-Mead	MSE

Results

Overview



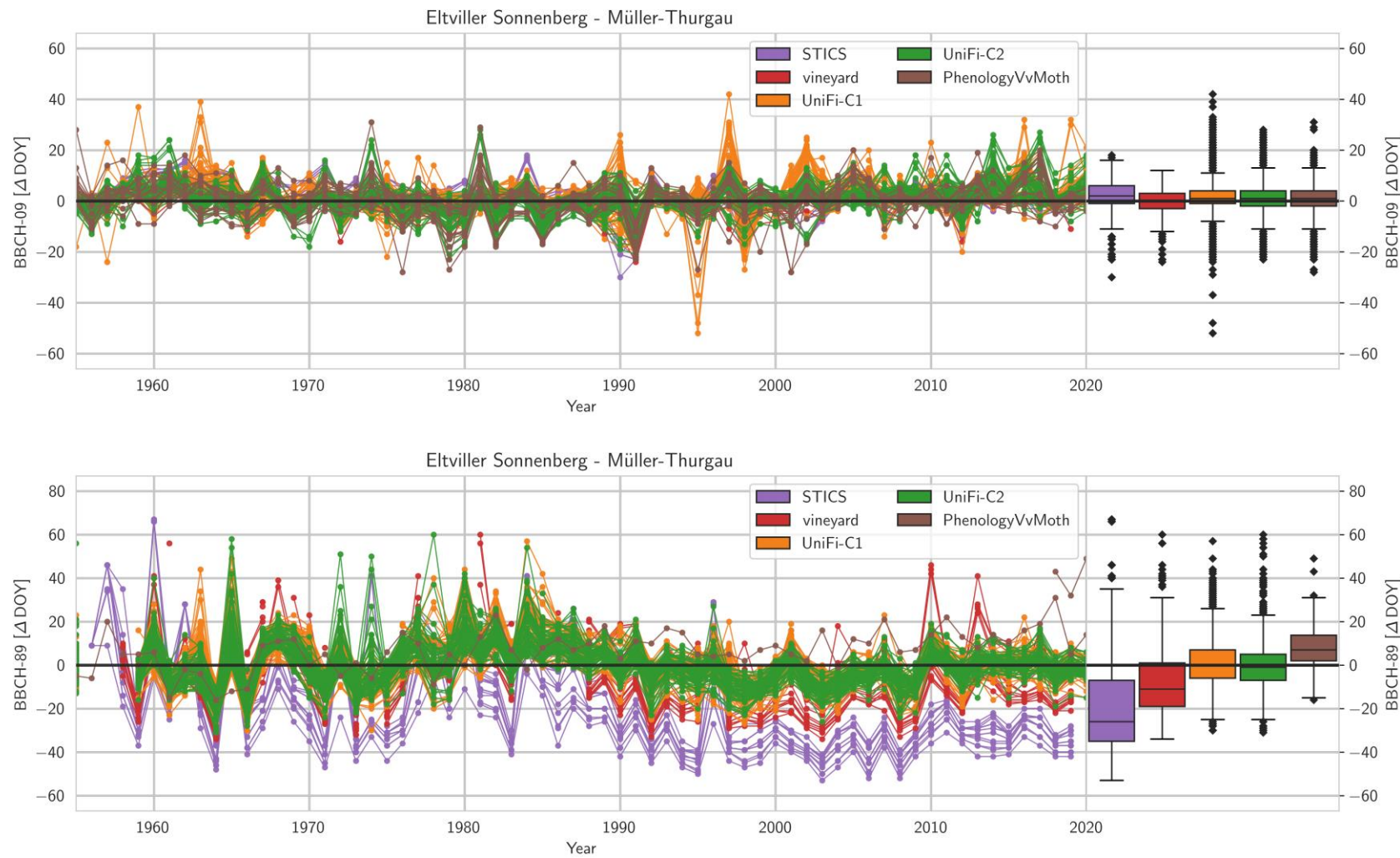
Results

Overview



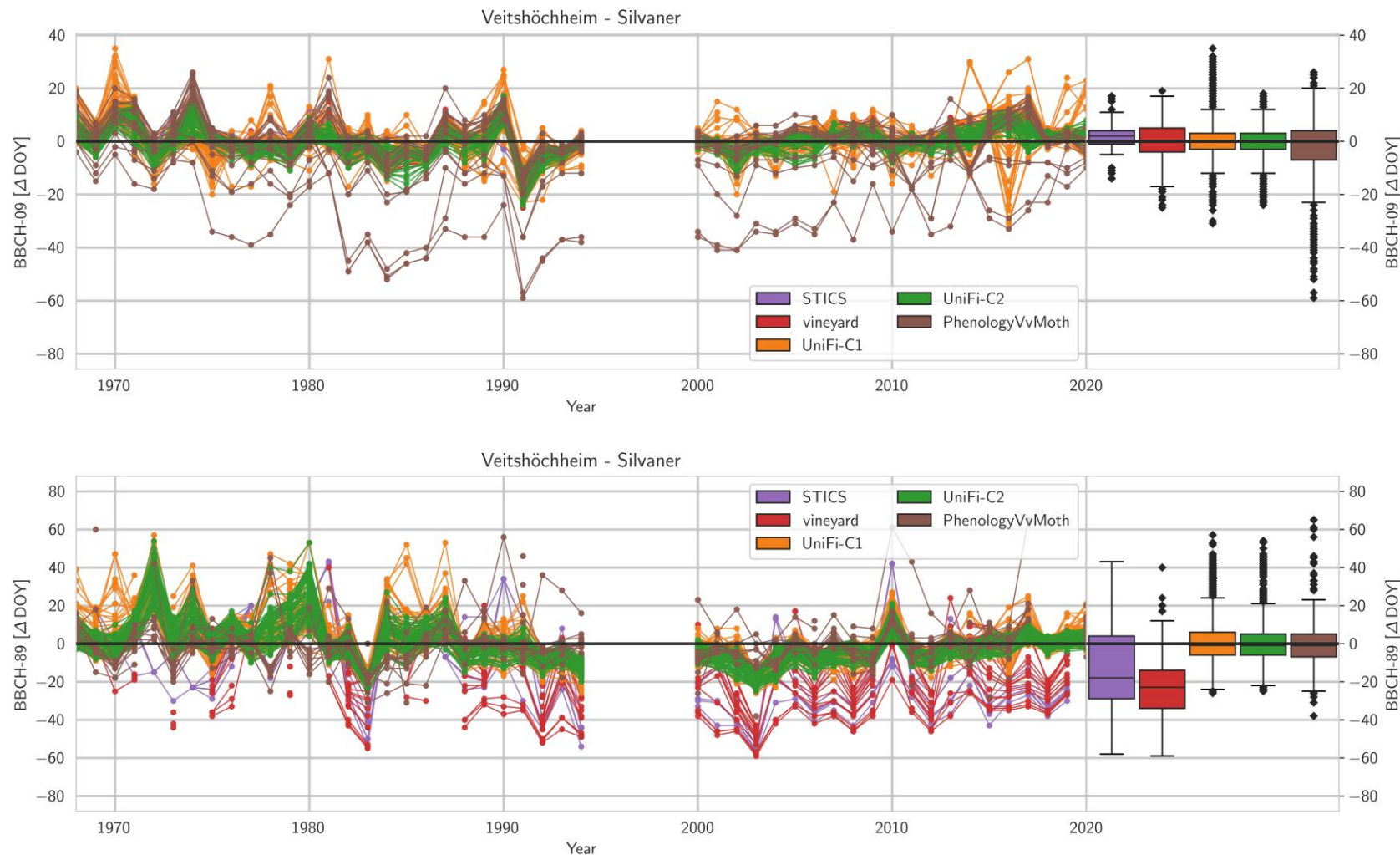
Results

Overview



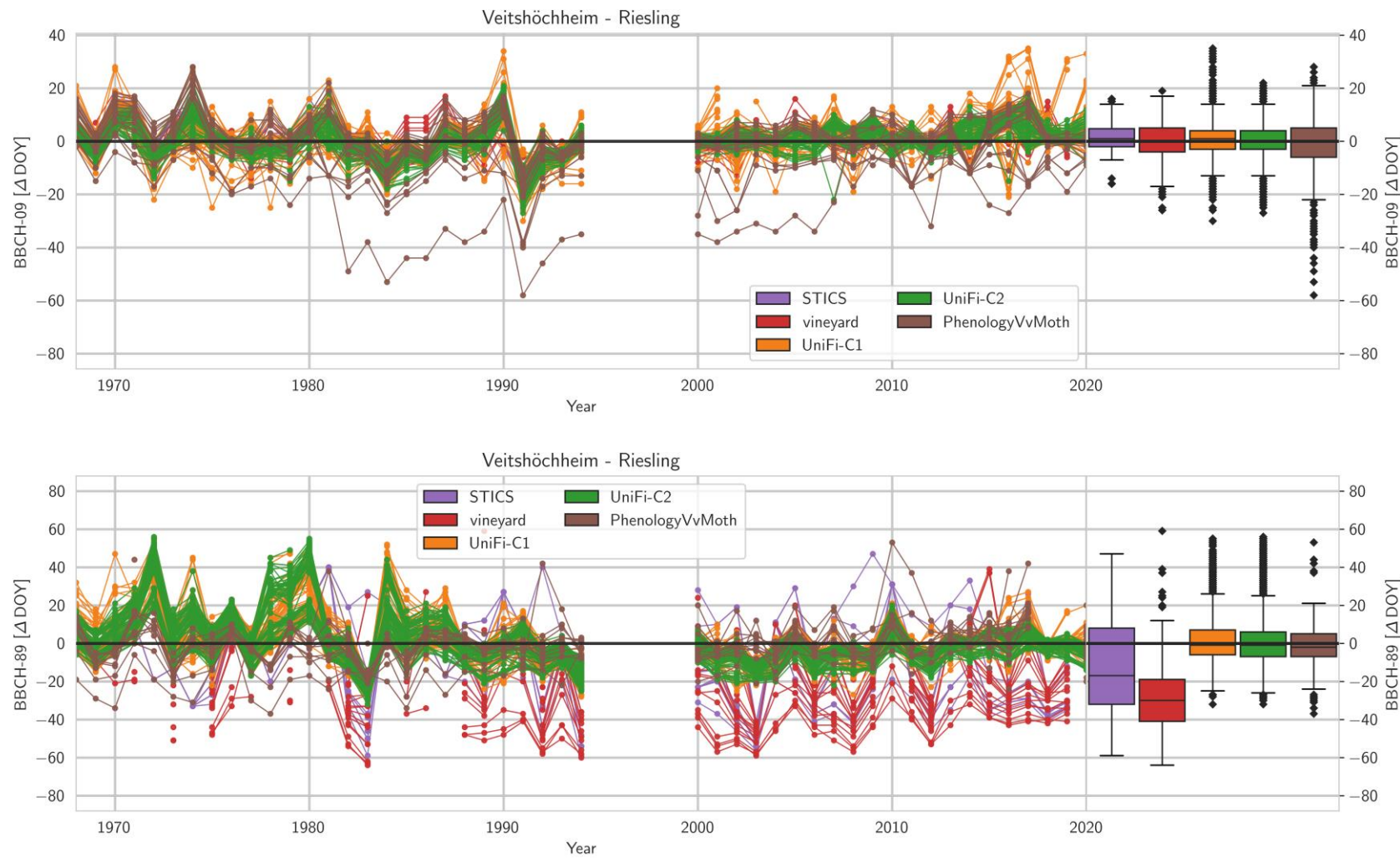
Results

Overview



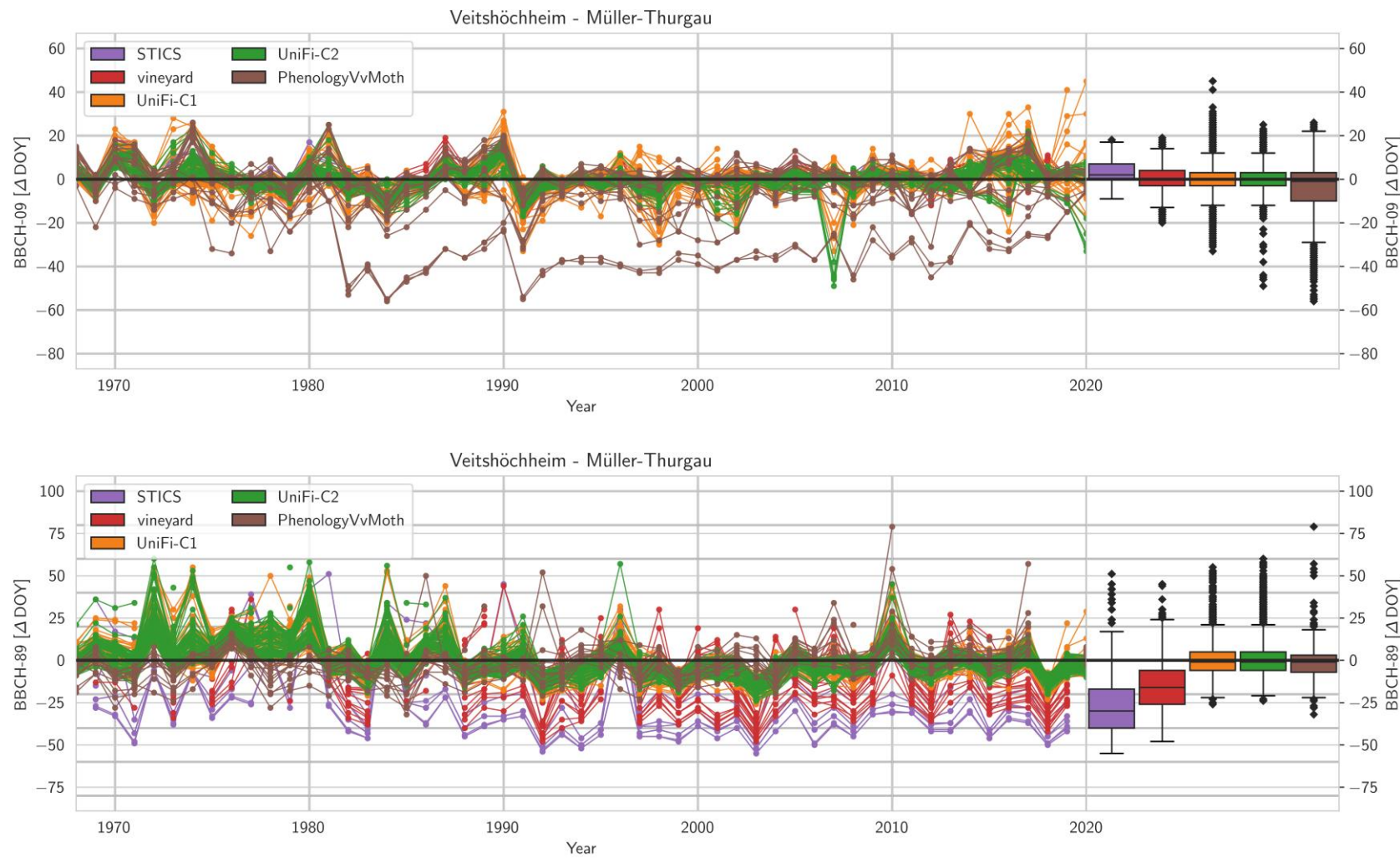
Results

Overview



Results

Overview



Phenology Dataset Overview

Sources	
PEP725	1940 sites, restricted
TEMPO	1640 sites, registered users only
DWD	949 sites, free
UTAD	15 sites and 2 regions
UniFI	8 sites (France) and 54 sites (Sicily)
LIST	2 sites, Luxembourg
PIK	3 sites, restricted
<ul style="list-style-type: none">• ASCII/CSV format	
/p/projects/clim4vitis/data/phenology/observations/preprocessed	



Phenology Dataset Overview

Sources	
PEP725	1940 sites, restricted
TEMPO	1640 sites, registered users only
DWD	949 sites, free
UTAD	15 sites and 2 regions
UniFI	8 sites (France) and 54 sites (Sicily)
LIST	2 sites, Luxembourg
PIK	3 sites, restricted
<ul style="list-style-type: none">• ASCII/CSV format	
/p/projects/clim4vitis/data/phenology/observations/preprocessed	

- Different formats
- Different quality
- Different type



Phenology Dataset Overview

Sources	
PEP725	1940 sites, restricted
TEMPO	1640 sites, registered users only
DWD	949 sites, free
UTAD	15 sites and 2 regions
UniFI	8 sites (France) and 54 sites (Sicily)
LIST	2 sites, Luxembourg
PIK	3 sites, restricted
<ul style="list-style-type: none">• ASCII/CSV format	
/p/projects/clim4vitis/data/phenology/observations/preprocessed	

- Different formats
- Different quality
- Different type
- Separately preprocessing to check/correct and unify each dataset

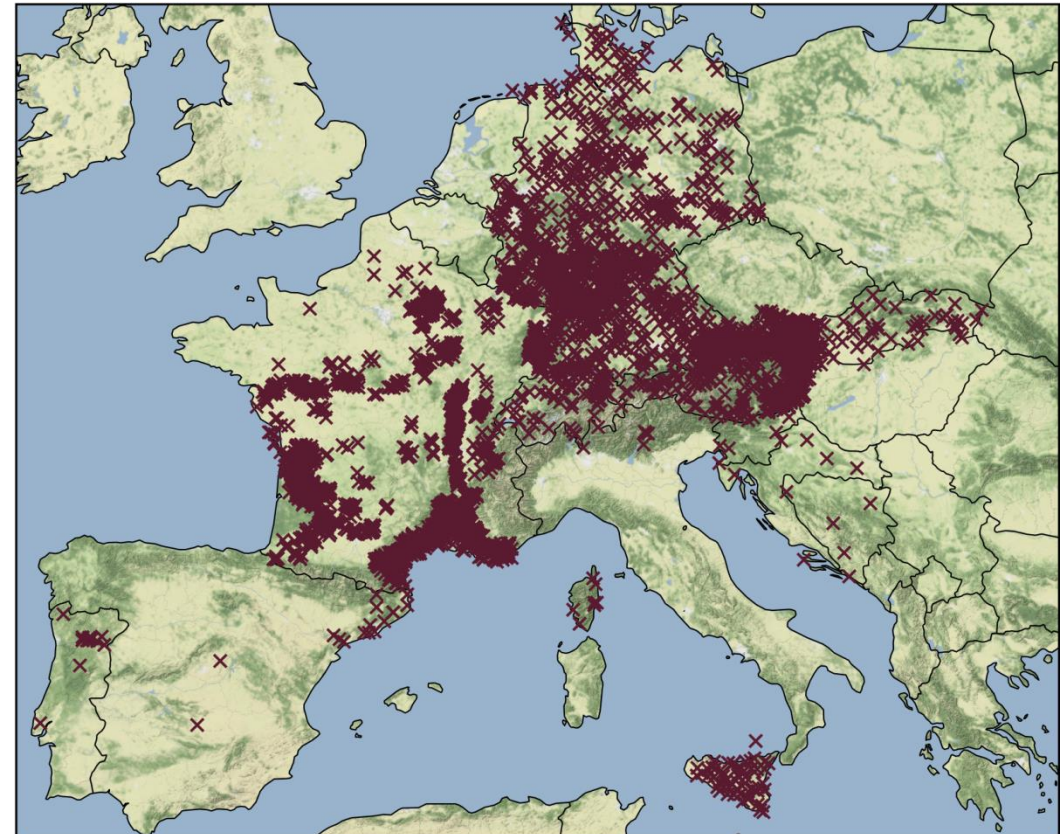


Phenology Dataset Overview

Primary Evaluation Matrix

- Merging of all datasets to get a comprehensive dataset
- Full spatio-temporal coverage and check/complement duplicate sites

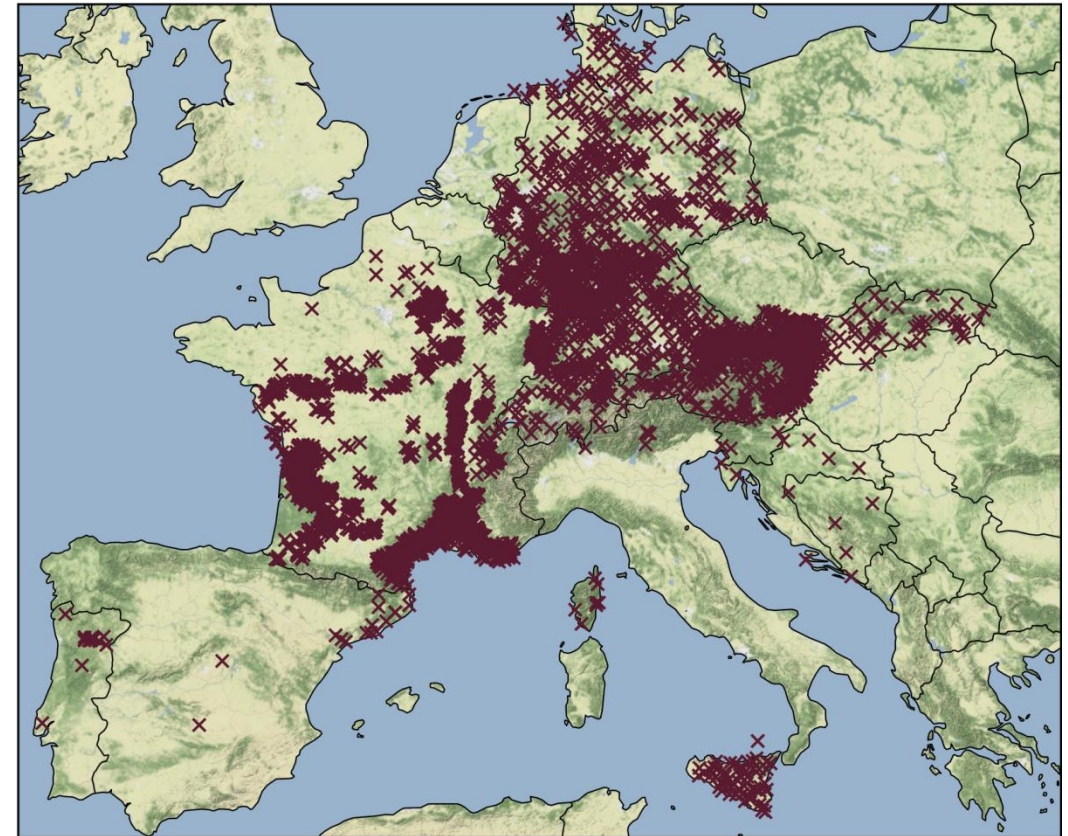
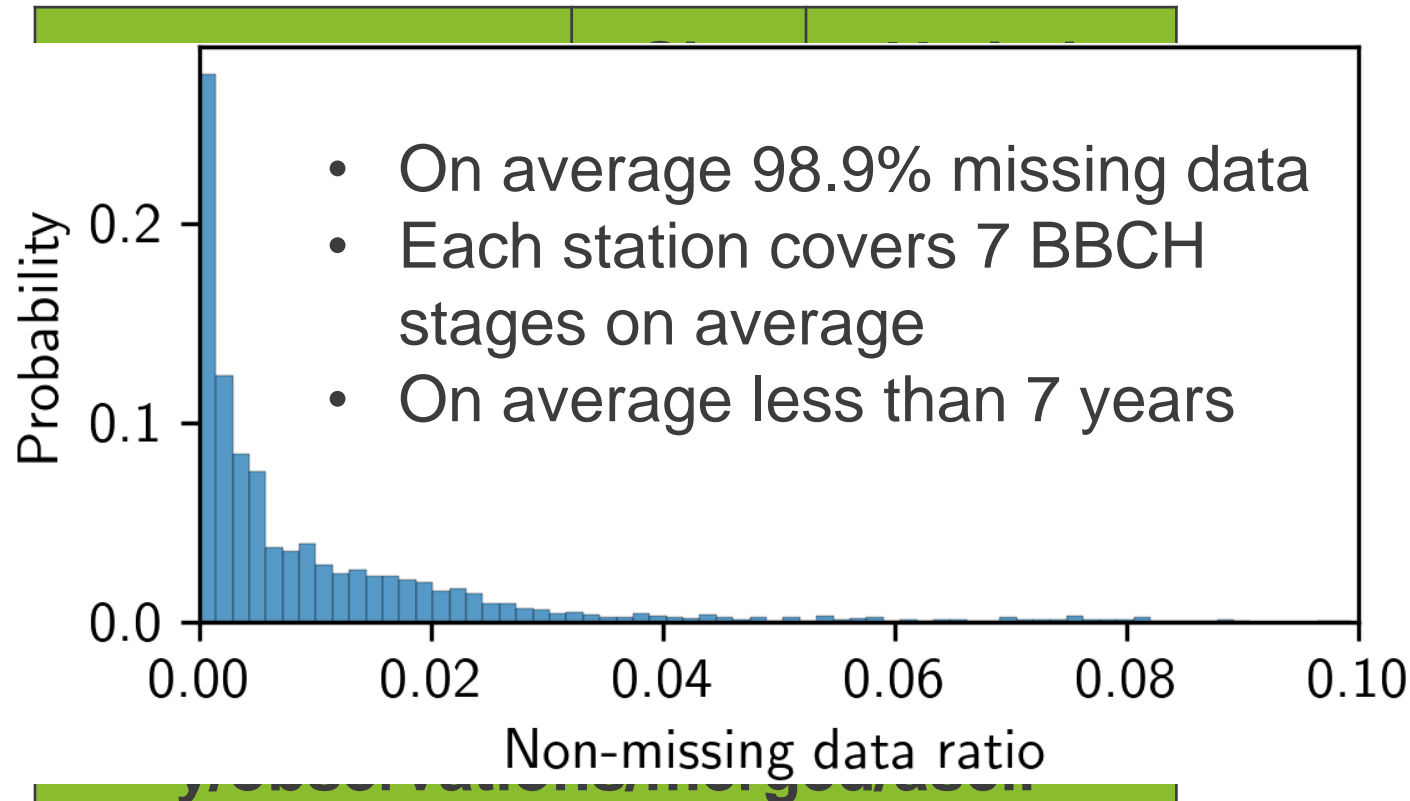
Sources	# Sites	# Varieties
Total	3872	123
Known variety	346	116
Unknown variety	3526	7
• 48 different BBCH stages in total Sparsely covered!		
/p/projects/clim4vitis/data/phenology/observations/merged/ascii		



Phenology Dataset Overview

Primary Evaluation Matrix

- Merging of all datasets to get a comprehensive dataset
- Full spatio-temporal coverage and check/complement duplicate sites



Model Comparison Framework



Model Comparison Framework

Model Comparison Framework

- Calibration and Evaluation of phenology models
- Climate input and phenology data for calibration and validation prepared



Model Comparison Framework

Model Comparison Framework

- Calibration and Evaluation of phenology models
- Climate input and phenology data for calibration and validation prepared
- Needs:
 - At least 30 years of observation
 - Covering budbreak, flowering, end of flowering, ripening and maturity stage
 - Multiple varieties for one site



Model Comparison Framework

Model Comparison Framework

- Calibration and Evaluation of phenology models
- Climate input and phenology data for calibration and validation prepared
- Needs:
 - At least 30 years of observation
 - Covering budbreak, flowering, end of flowering, ripening and maturity stage
 - Multiple varieties for one site



/p/projects/clim4vitis/data/climate/observations/stations/comparison_study
/p/projects/clim4vitis/data/phenology/observations/comparison_study

Model Comparison Framework

Primary Evaluation Matrix

- Primary setup for evaluation
- Robust calibration and statistical significant validation over climatological time scales
- Comparison of phenology model variability

BBCH		7	9	61	63	65	68	69	81	83	85	89
Remich	Rivaner	0	48	0	44	0	44	0	42	0	0	0
Eltviller Sonnenberg	Riesling	0	65	65	0	64	65	0	65	46	45	65
Eltviller Sonnenberg	Müller Thurgau	0	65	65	0	65	65	0	62	47	42	65
Eltviller Sonnenberg	Spätburgunder	0	33	33	0	33	33	0	33	33	32	33
Veitshoechheim	Müller Thurgau	24	52	26	52	52	52	24	52	24	25	52
Veitshoechheim	Riesling	20	47	20	47	47	47	20	47	20	21	47
Veitshoechheim	Silvaner	20	47	20	47	47	47	20	47	20	20	47
Geisenheim-HGU	Riesling	0	53	0	53	53	53	0	53	0	53	53



Model Comparison Framework

Secondary Evaluation Matrix

- At least 20 years
- At least 3 out of 5 man stages covered
- Trade robustness against spatial coverage
- Comparison of regional and variety representativeness

BBCH		7	9	61	63	65	68	69	81	83	85	89
Remich	Gewuerztraminer	0	44	0	44	0	44	0	9	0	0	0
Remich	Riesling	0	44	0	44	0	44	0	9	0	0	0
Remich	Pinot Gris	0	44	0	44	0	44	0	9	0	0	0
Remich	Pinot Blanc	0	44	0	44	0	44	0	9	0	0	0
Remich	Auxerrois	0	44	0	44	0	44	0	9	0	0	0
Remich	Elbling	0	44	0	44	0	44	0	9	0	0	0
Bordeaux Chateau Lafite	Merlot	47	0	4	0	53	0	0	4	0	53	48
Bergheim	Riesling	48	0	0	0	42	0	42	0	0	41	8
Bergheim	Pinot Noir	48	0	0	0	42	0	42	0	0	40	8
Strasburgo INRA Colmar Domaine de Bergheim	Pinot Noir	48	0	0	0	42	0	42	0	0	40	8
44km off Marseillan	Cabernet Sauvignon	30	0	0	0	21	0	0	0	0	19	24
Montpellier Vassal UE	Cabernet Sauvignon	0	39	0	0	26	0	0	0	0	32	0
Montpellier Vassal UE	Syrah	0	41	0	0	26	0	0	0	0	26	0
Lisboa Region	Fernao Pires	0	25	25	0	0	0	0	0	25	0	0
Lisboa Region	Chasselas	0	25	25	0	0	0	0	0	25	0	0
Lisboa Region	Castelao	0	25	25	0	0	0	0	0	25	0	0
Lisboa Region	Aragonez	0	21	21	0	0	0	0	0	21	0	0
Minho Region	Vinhao	0	23	22	0	0	0	0	0	22	0	0
Minho Region	Loureiro	0	23	22	0	0	0	0	0	22	0	0



Introduction

Intercomparison Framework

Create a **Framework** for **unified** model calibration and validation

- Multiple models able to simulate the same variety, site and phenological phases
- Multiple sites/varieties to calibration and validate the models
- Long time frames to conscientiously test model sensitivities

