Stable machine-learning parameterization of subgrid processes for climate modeling at a range of resolutions

Janni Yuval and Paul O'Gorman

Support from Houghton-Lorenz Fellowship

MIT Environmental Solutions Initiative and the NSF

Massachusetts Institute of Technology

Outline

- **Motivation**: Parameterizations of small scale processes are the main cause of uncertainty in climate change **temperature and precipitation** projections
- **Previous studies:** Recent studies on machine-learning parameterization are promising (but still many problems)
- **Results**:
  a. Learning from a high resolution model (SAM)
  b. Example: Sub-grid tendencies are very important
  c. Offline results, conservation of energy, non-negative precipitation
  d. Online results
  e. Is there an optimal length scale for parameterization?

Outline

- **Motivation**: Parameterizations of small scale processes are the main cause of uncertainty in climate change **temperature and precipitation** projections
- **Previous studies:** Recent studies on machine-learning parameterization are promising (but still many problems)
- **Results**:
  a. Learning from a high resolution model (SAM)
  b. Example: Sub-grid tendencies are very important
  c. Offline results, conservation of energy, non-negative precipitation
  d. Online results
  e. Is there an optimal length scale for parameterization?

THE

LONDON, EDINBURGH, AND DUBLIN

PHILOSOPHICAL MAGAZINE

AND

JOURNAL OF SCIENCE.

[FIFTH SERIES.]

APRIL 1896.

XXXI. *On the Influence of Carbonic Acid in the Air upon the Temperature of the Ground.* *By* Prof. SVANTE ARRHENIUS *.

I. *Introduction* : *Observations of* Langley *on Atmospherical Absorption.*

THE

LONDON, EDINBURGH, AND DUBLIN

PHILOSOPHICAL MAGAZINE

AND

JOURNAL OF SCIENCE.

[FIFTH SERIES.]

APRIL 1896.

XXXI. *On the Influence of Carbonic Acid in the Air upon the Temperature of the Ground.* *By* Prof. SVANTE ARRHENIUS *.

1. *Introduction* : *Observations of* Langley *on Atmospherical Absorption.*

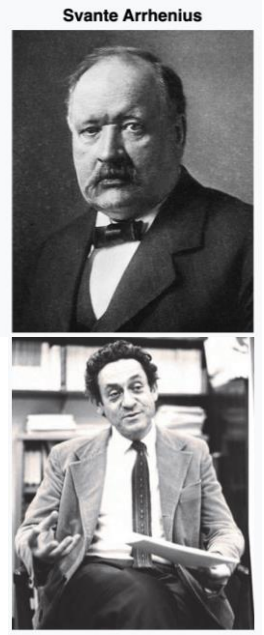
Large uncertainty in the temperature response to increased CO_2 concentrations (slow to non-existent improvement in narrowing the uncertainty)

Arrhenius **(1896)**:
Doubling of CO_2 → global warming of **~5°C**

Large uncertainty in the temperature response to increased CO_2 concentrations (slow to non-existent improvement in narrowing the uncertainty)

Arrhenius **(1896)**:
Doubling of CO_2 → global warming of **~5°C**

Charney Report **(1979)**:
"We estimate the most probable global warming for a doubling of CO_2 to be near **3°C** with a probable error of **± 1.5°C**"

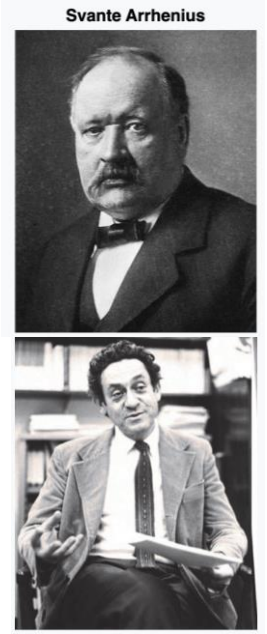


Large uncertainty in the temperature response to increased CO_2 concentrations (slow to non-existent improvement in narrowing the uncertainty)

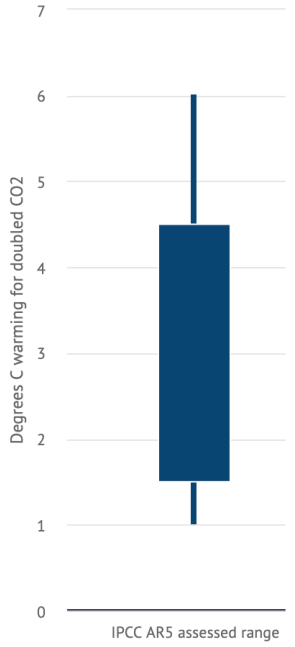
Arrhenius **(1896)**:
Doubling of CO_2 → global warming of **~5°C**

Charney Report **(1979)**:
"We estimate the most probable global warming for a doubling of CO_2 to be near **3°C** with a probable error of **± 1.5°C**"

IPCC report **(2013)**:
"Estimates of the equilibrium climate sensitivity (ECS) based on observed climate change, climate models and feedback analysis, as well as paleoclimate evidence indicate that ECS is likely in **the range 1.5°C to 4.5°C** with high confidence, …. and **very unlikely greater than 6°C (medium confidence)**."

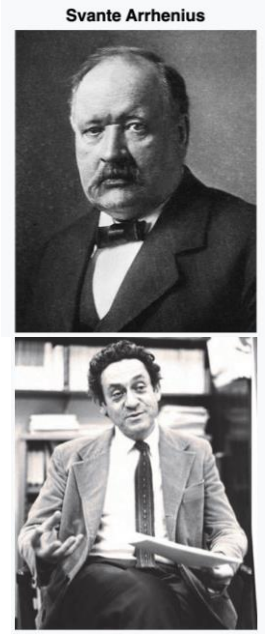


IPCC AR5
1.5-4.5°C



Large uncertainty in the temperature response to increased CO_2 concentrations (slow to non-existent improvement in narrowing the uncertainty)

Arrhenius **(1896)**:
Doubling of CO_2 → global warming of **~5°C**

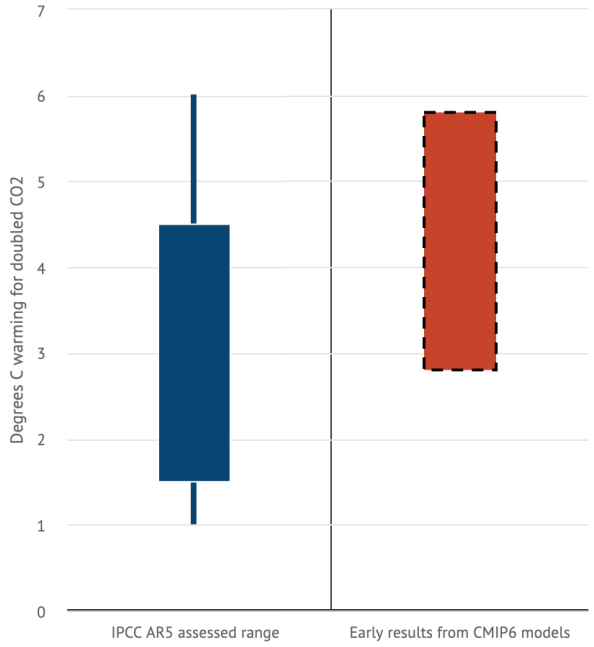


Charney Report **(1979)**:
"We estimate the most probable global warming for a doubling of CO_2 to be near **3°C** with a probable error of **± 1.5°C**"

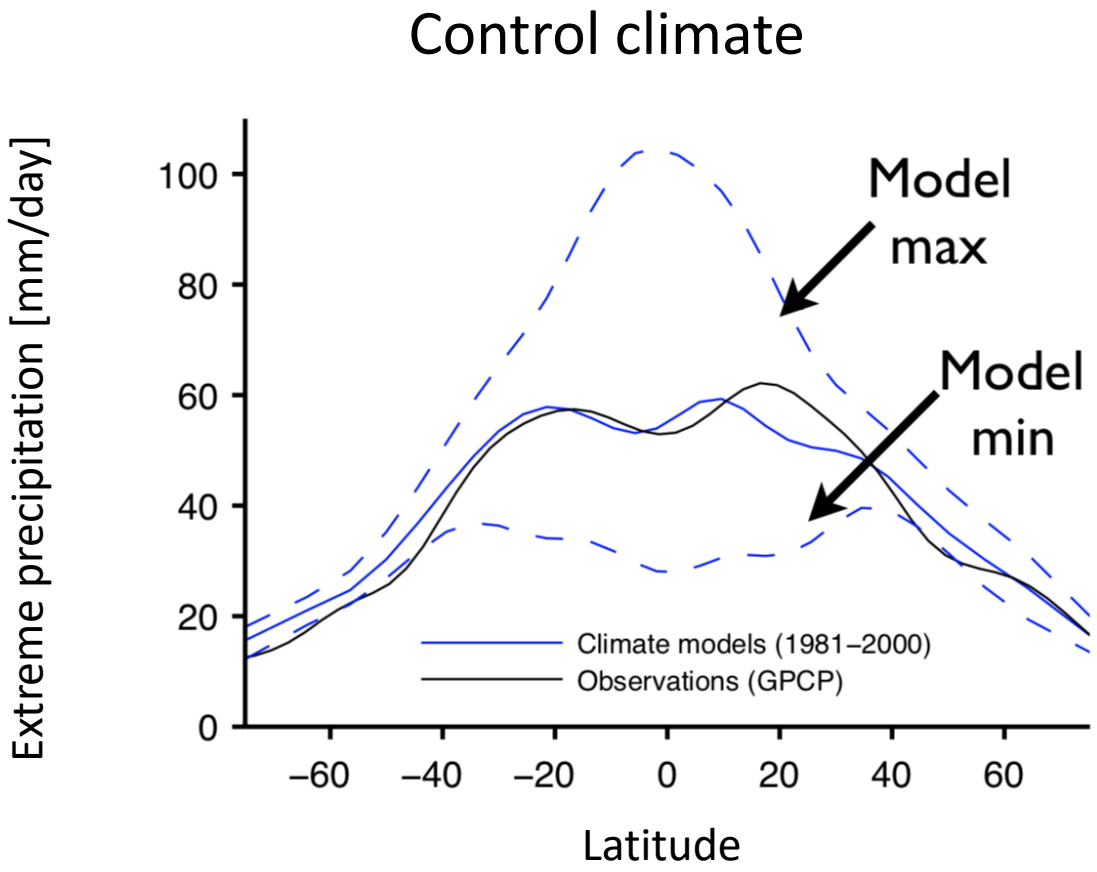
IPCC report **(2013)**:
"Estimates of the equilibrium climate sensitivity (ECS) based on observed climate change, climate models and feedback analysis, as well as paleoclimate evidence indicate that ECS is likely in **the range 1.5°C to 4.5°C** with high confidence, …. and **very unlikely greater than 6°C (medium confidence)**."

IPCC AR5 **1.5-4.5°C**

CMIP6 early results **2.8-5.8°C**



There is a **large spread in the intensity of extreme precipitation events** between climate models

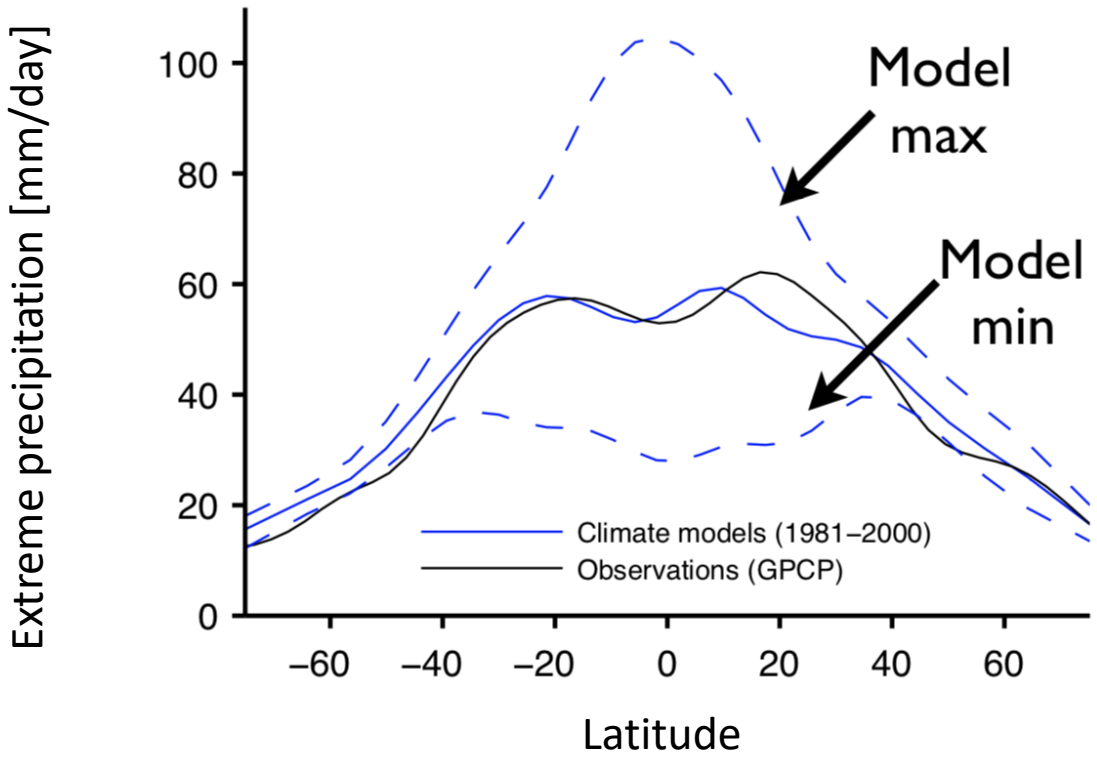


Climate models from CMIP3

Figure adapted from *O'Gorman and Schneider 2009*

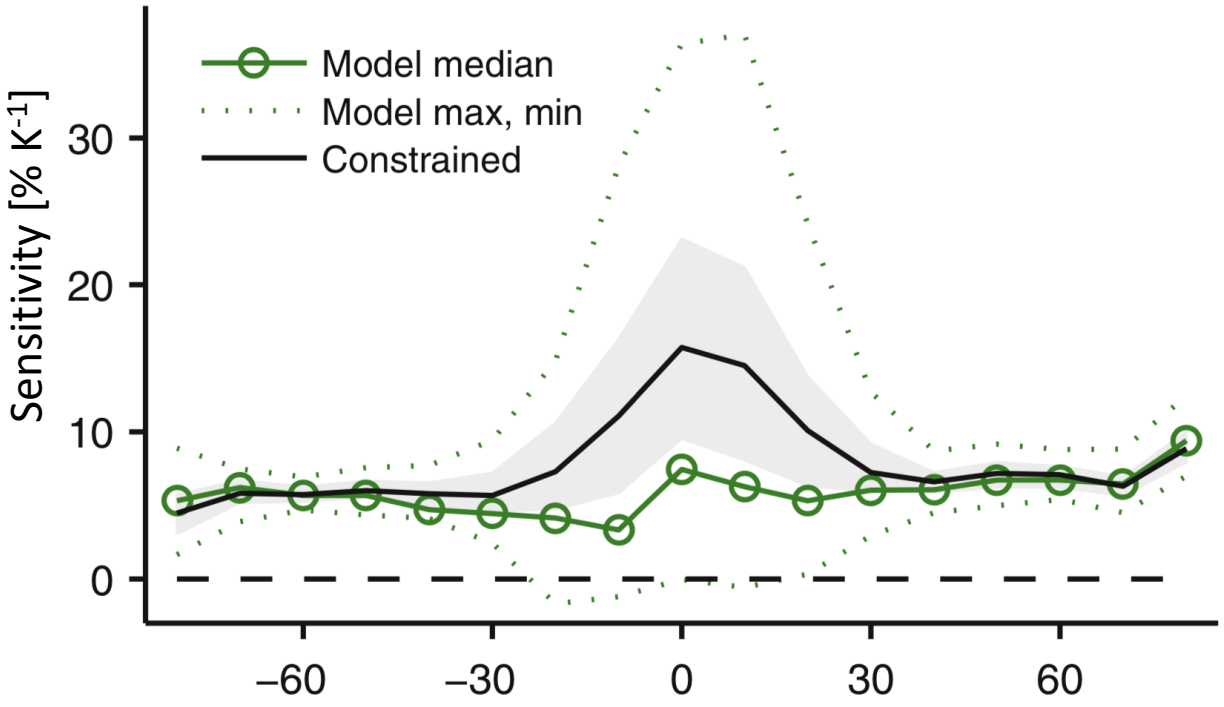
There is a large spread in the intensity of **extreme precipitation response to climate change**

Control climate



Climate models from CMIP3
Figure adapted from *O'Gorman and Schneider 2009*

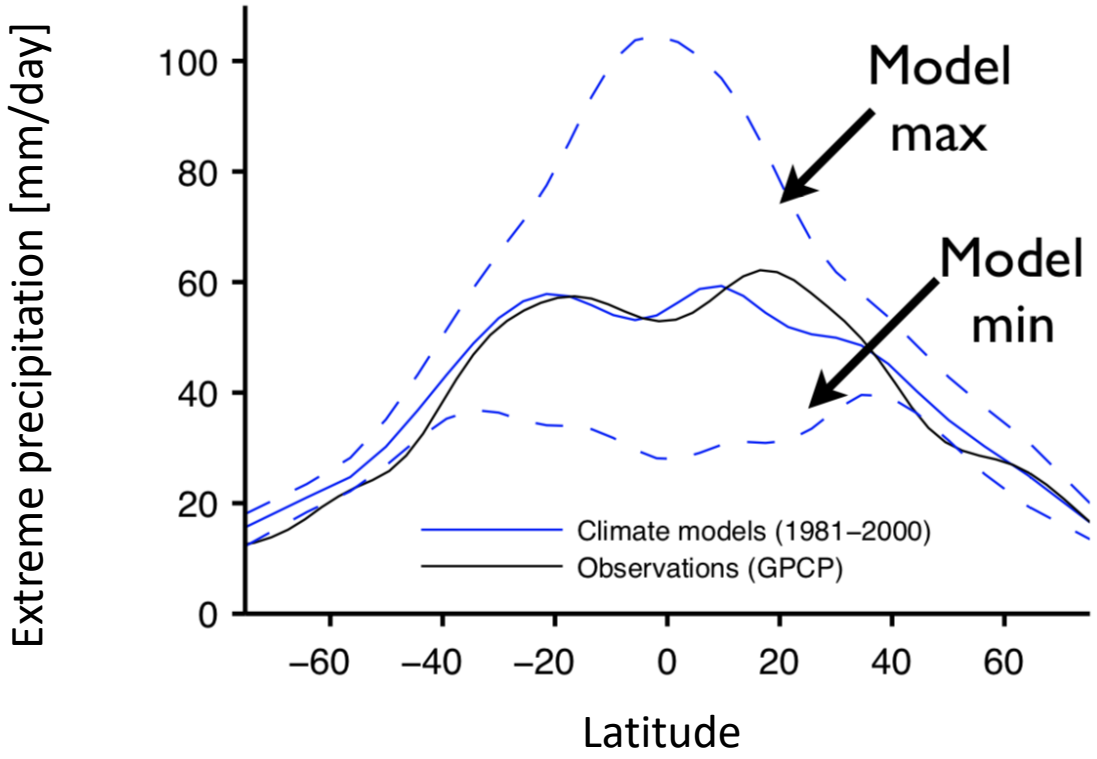
Climate change



Climate models from CMIP5
Figure adapted from: *O'Gorman 2015*

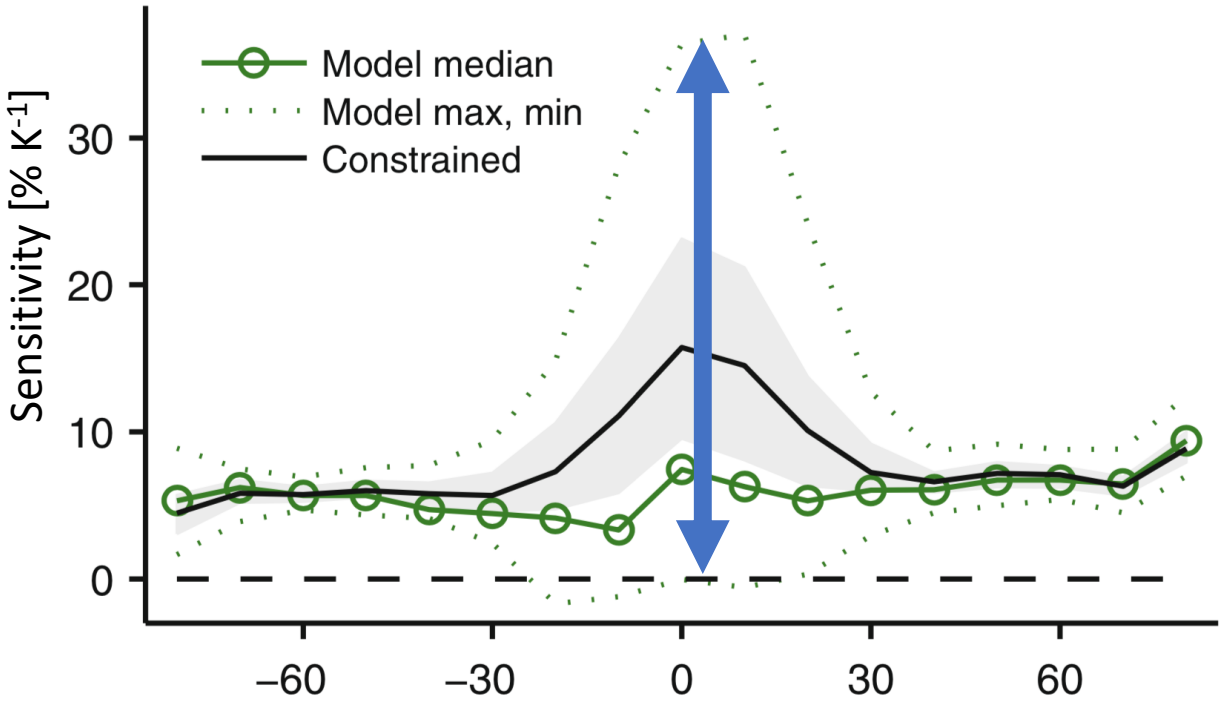
There is a large spread in the intensity of **extreme precipitation response to climate change**

Control climate



Climate models from CMIP3
Figure adapted from *O'Gorman and Schneider 2009*

Climate change



Climate models from CMIP5
Figure adapted from: *O'Gorman 2015*

The large uncertainty in climate projections is mainly caused by (inaccurate) representation of subgrid physics

Law of physics
(e.g., fluid dynamics)

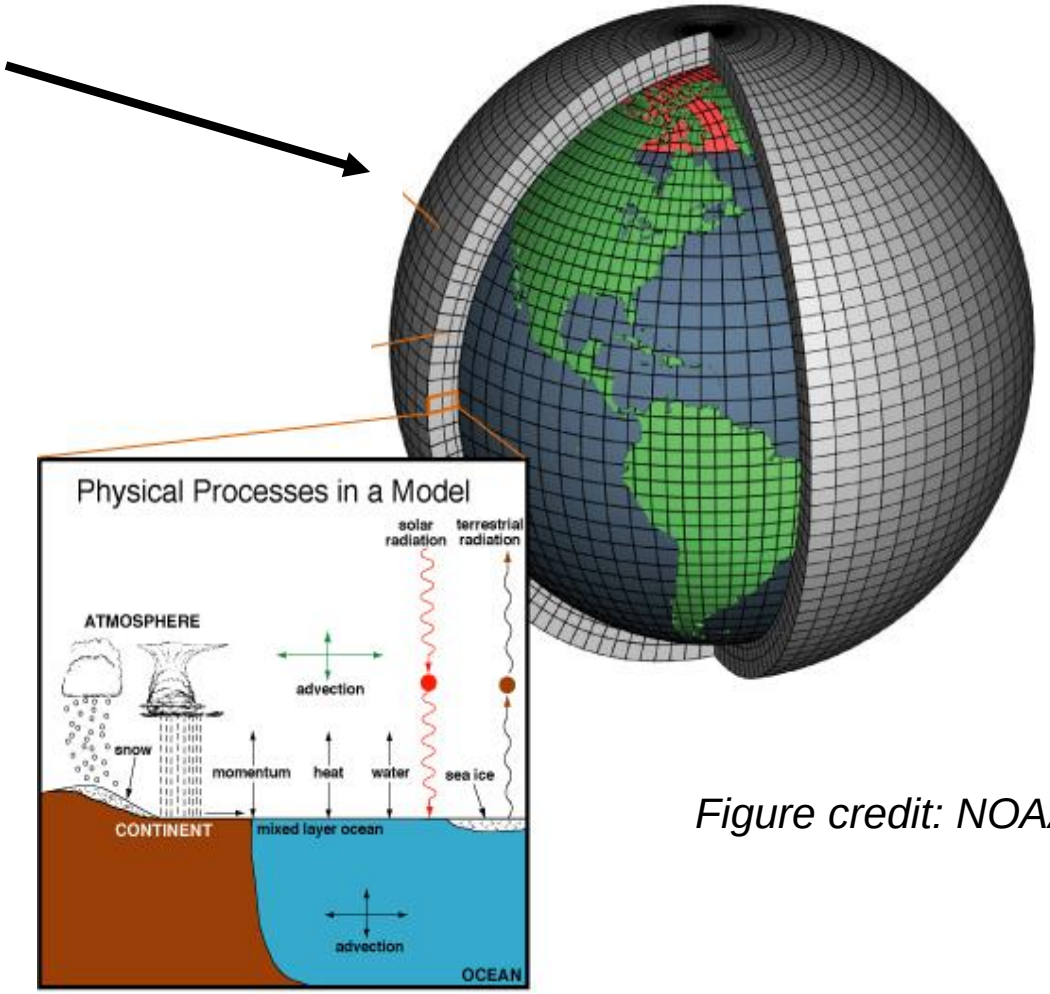
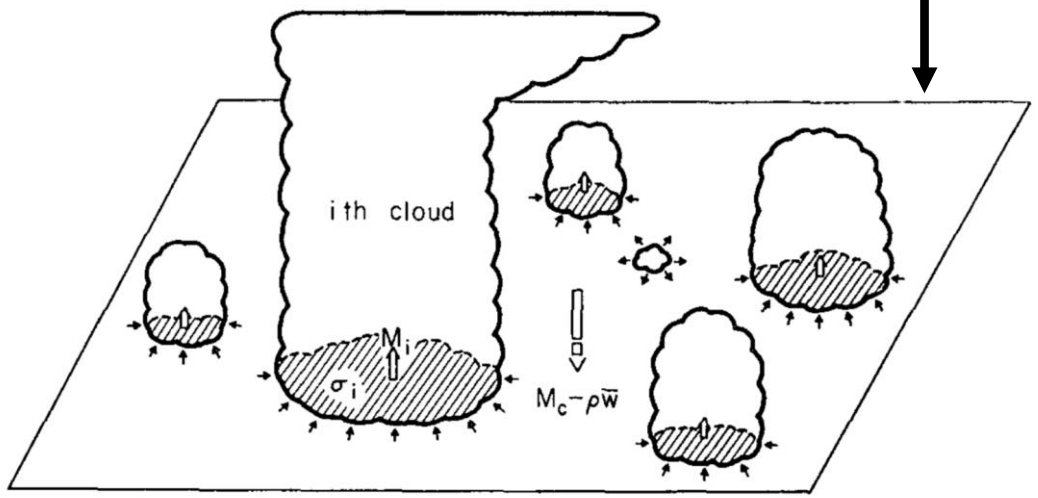


Figure credit: NOAA

The large uncertainty in climate projections is mainly caused by (inaccurate) representation of subgrid physics

Law of physics
(e.g., fluid dynamics)

Parameterizations
represented by simplified
models – (e.g., convection)



Arakawa and Schubert 1974

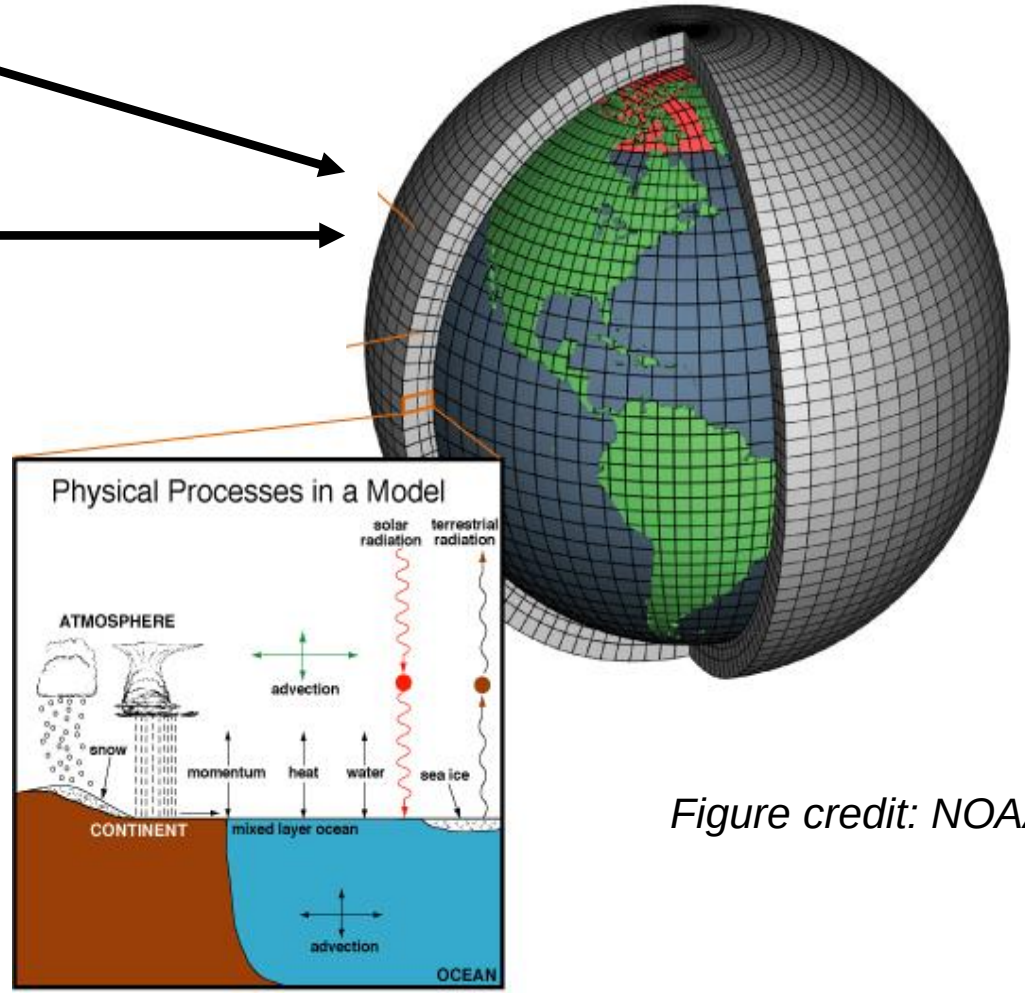
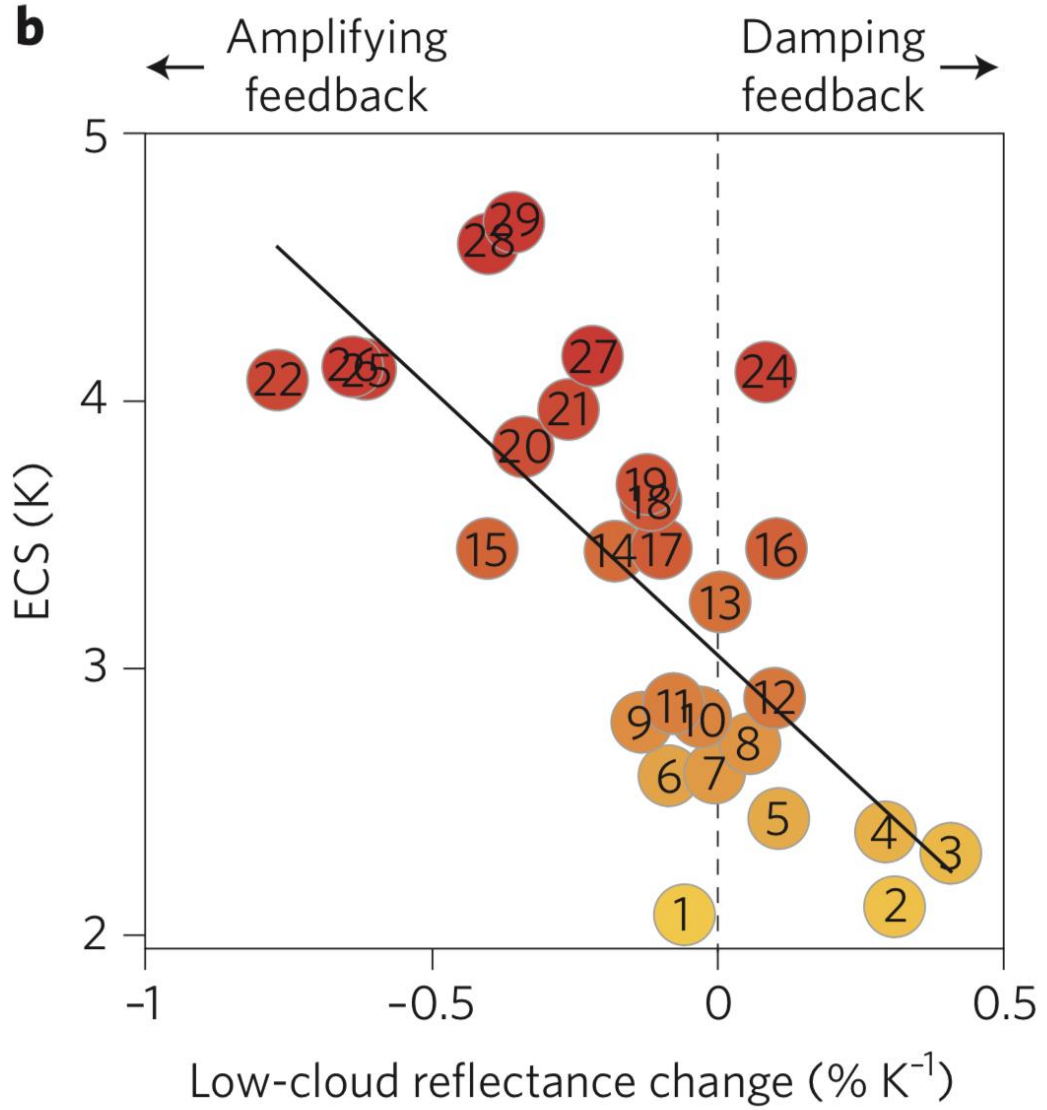


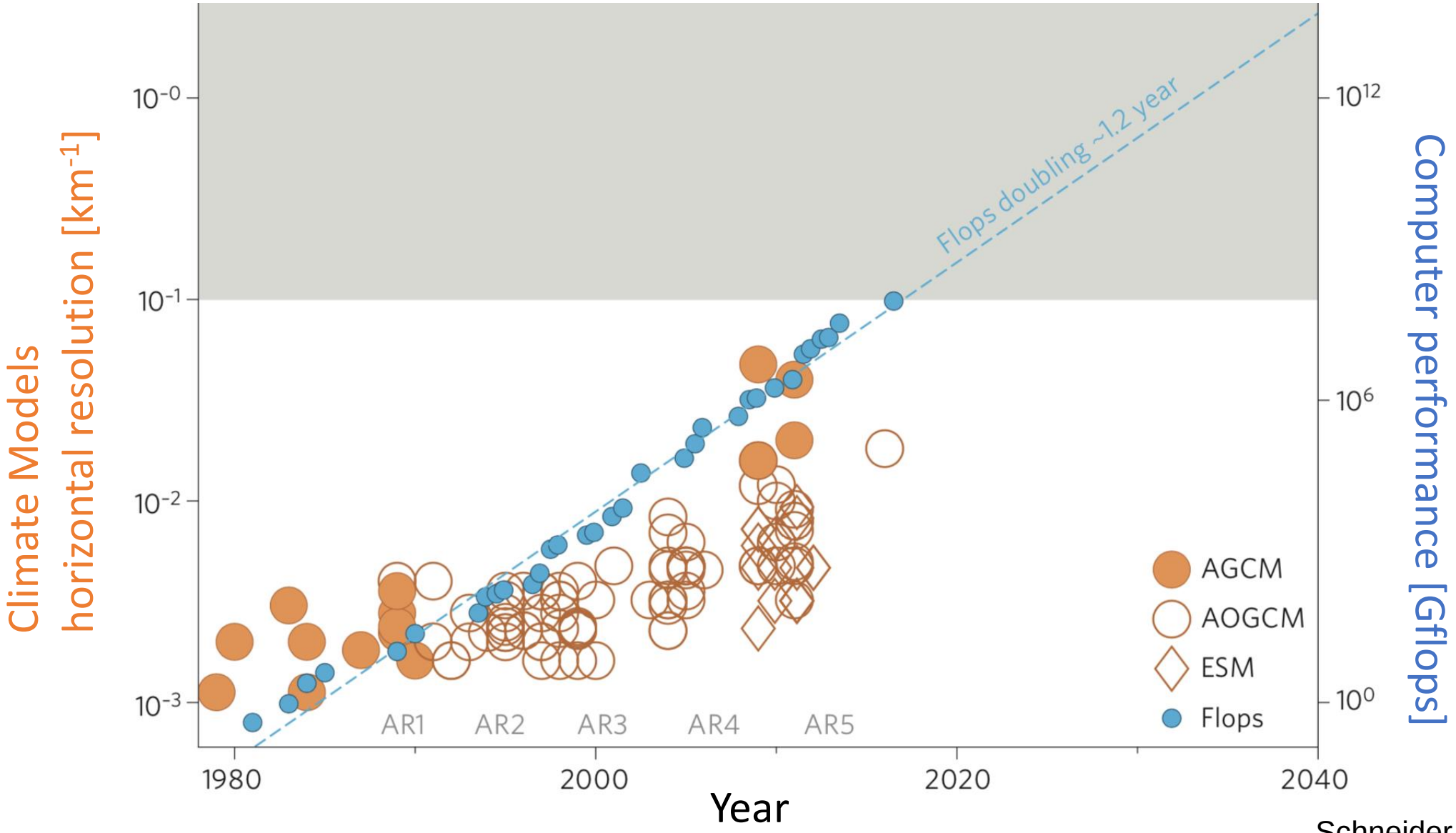
Figure credit: NOAA

Uncertainty in low-cloud feedback leads to large uncertainty in the equilibrium climate sensitivity



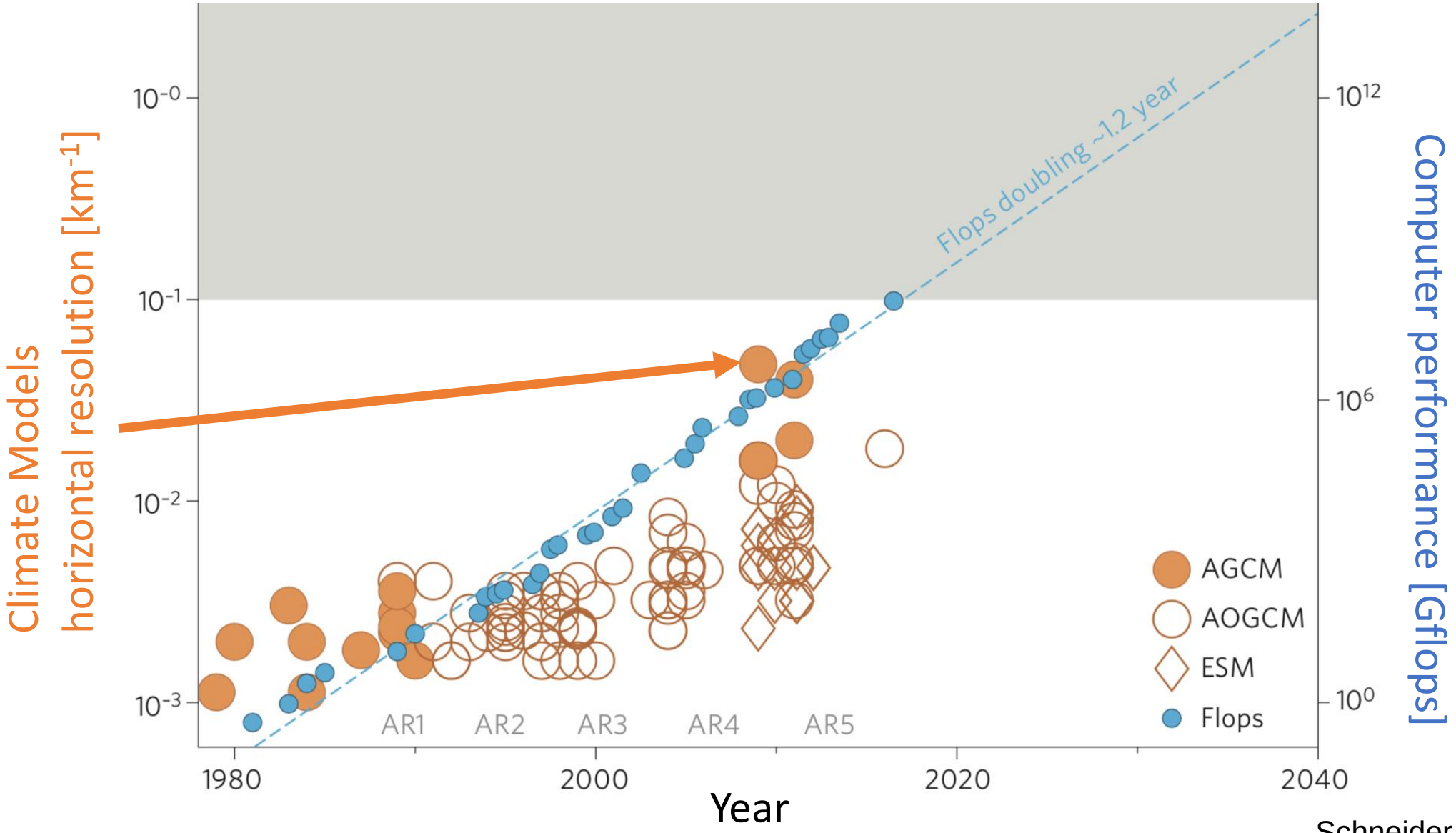
Schneider et al. (2017)

“Important” small scales will not be resolved in the foreseeable future



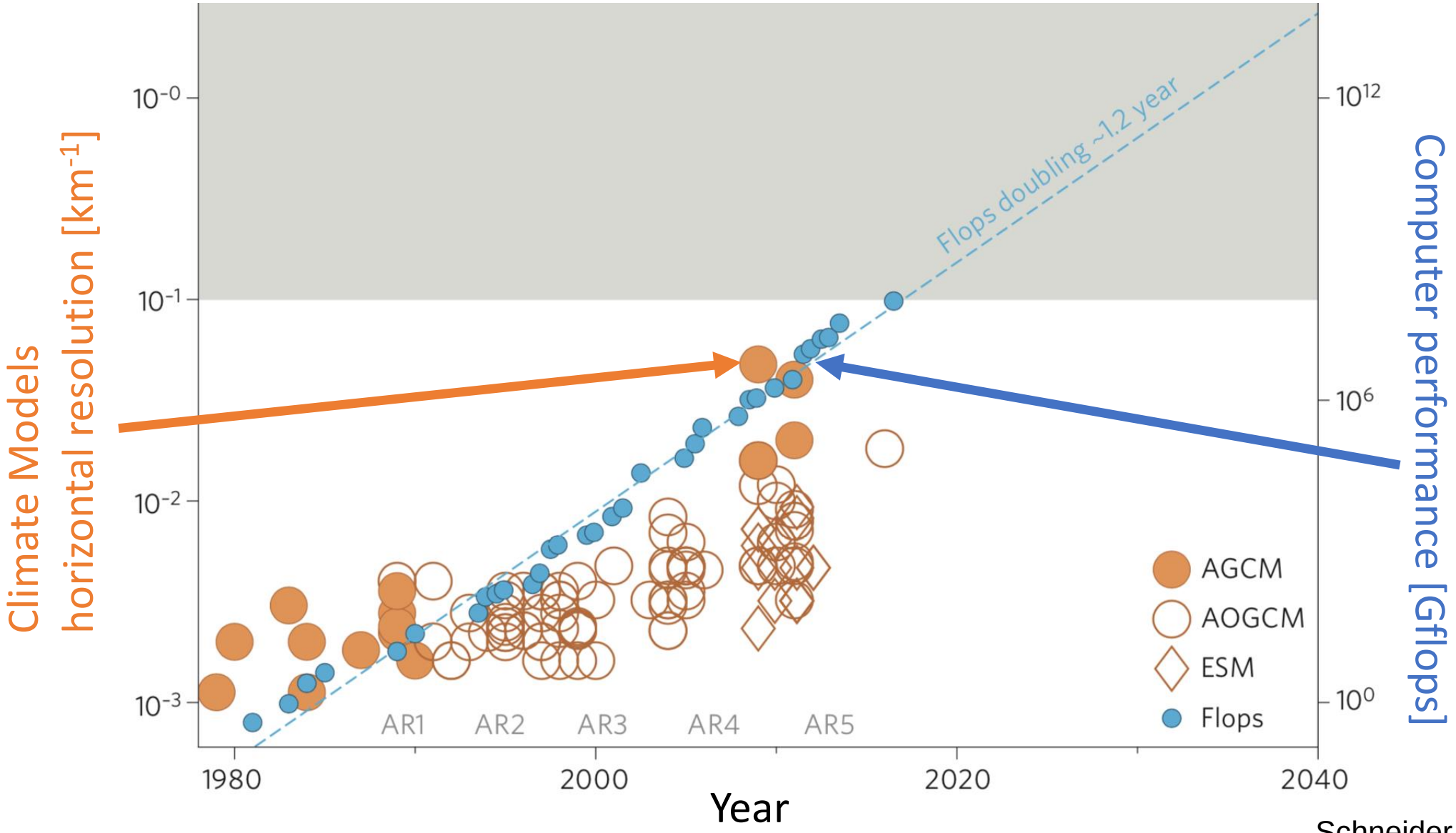
Schneider et al. (2017)

“Important” small scales will not be resolved in the foreseeable future



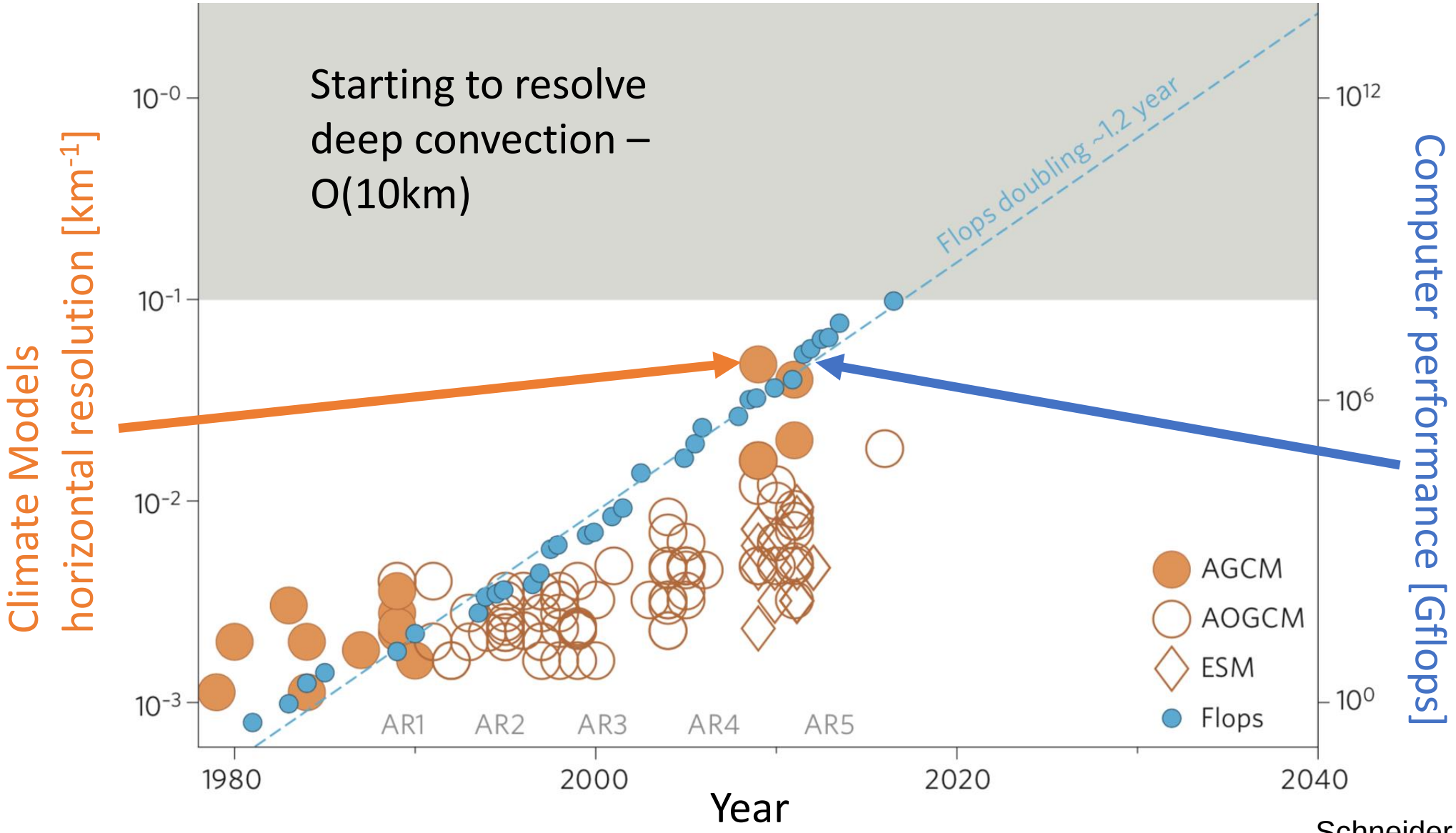
Schneider et al. (2017)

“Important” small scales will not be resolved in the foreseeable future



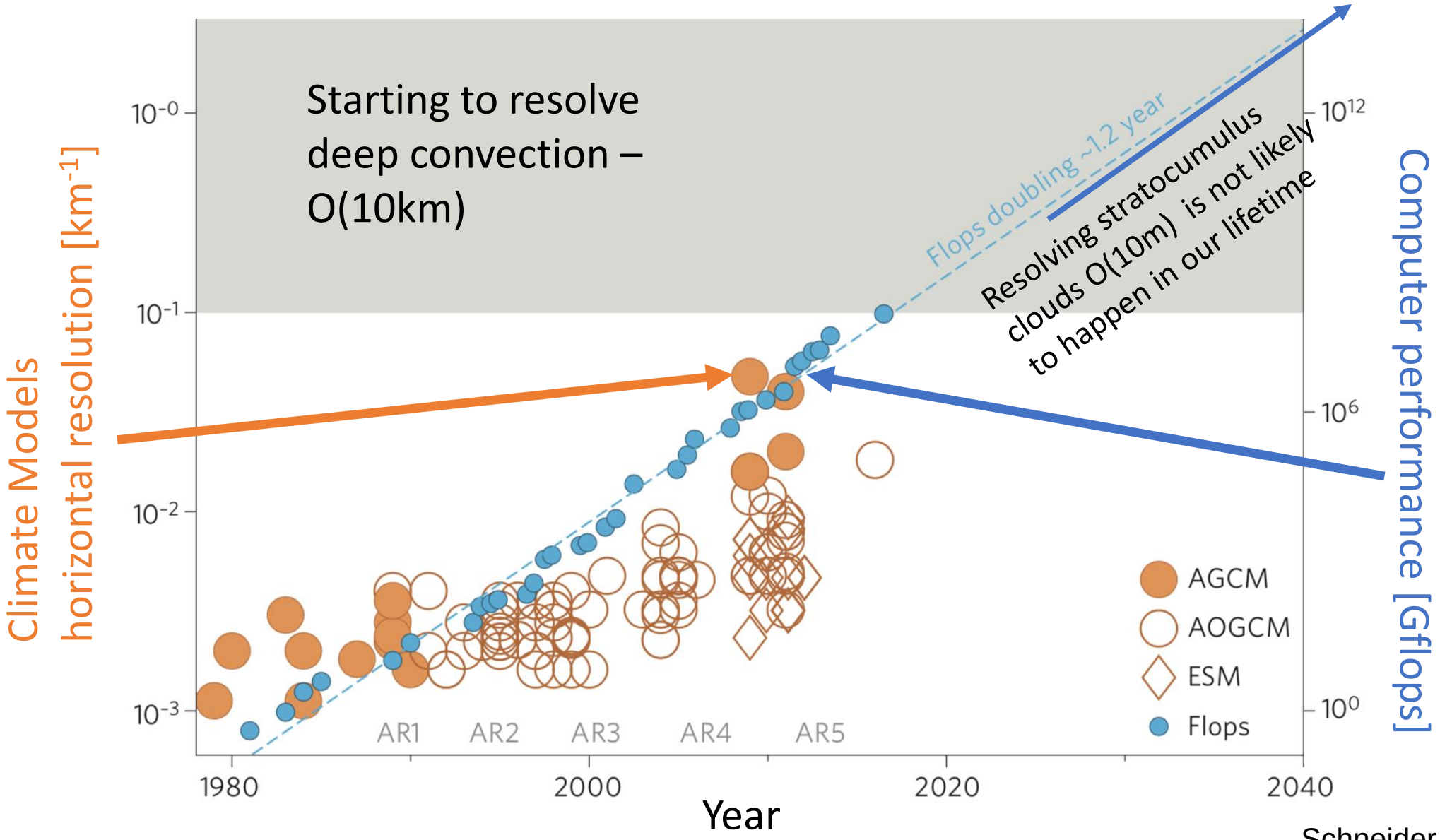
Schneider et al. (2017)

“Important” small scales will not be resolved in the foreseeable future



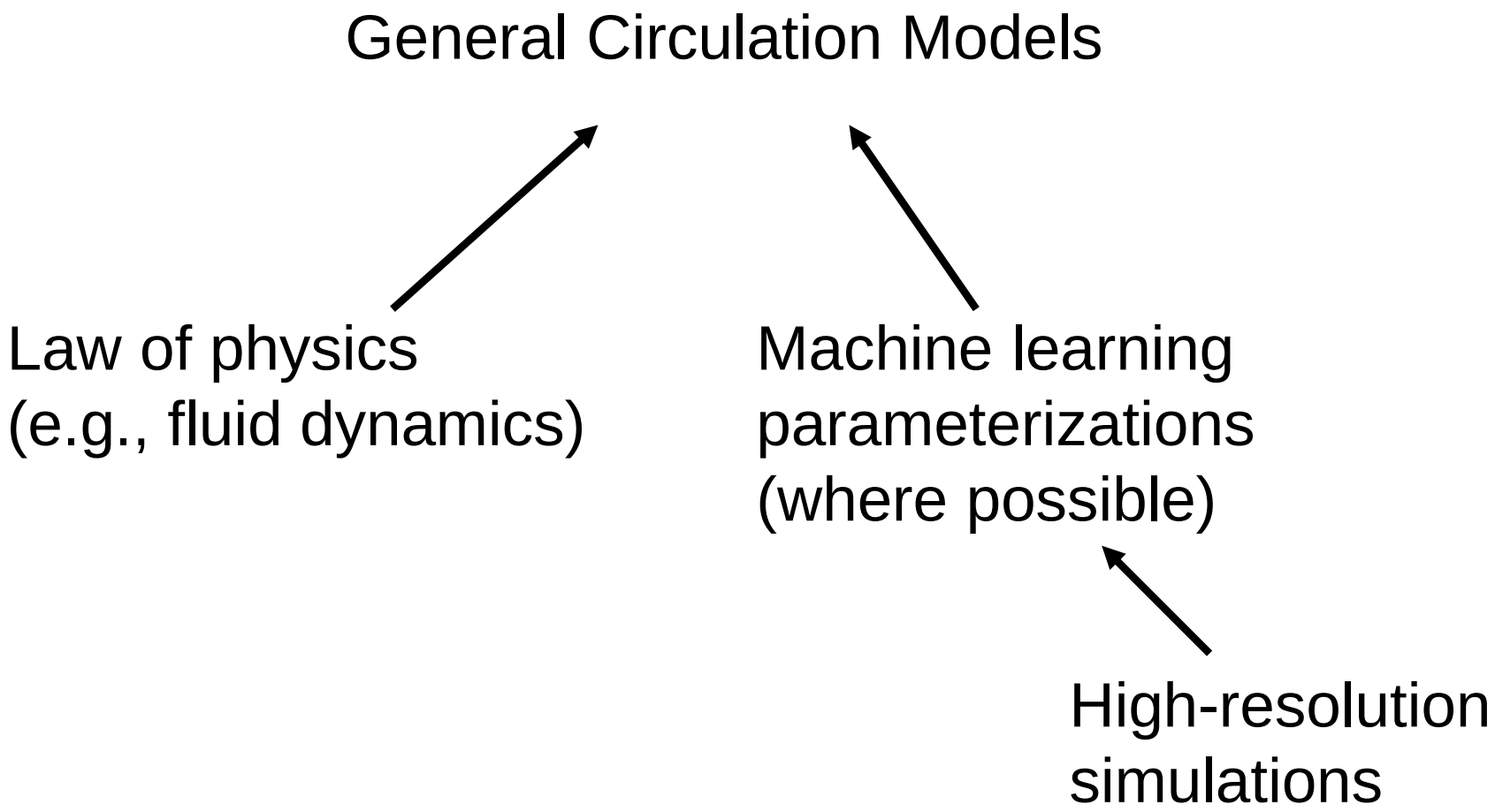
Schneider et al. (2017)

“Important” small scales will not be resolved in the foreseeable future



Schneider et al. (2017)

A different approach to parameterization: Use machine learning to create new parameterizations trained on high-resolution models

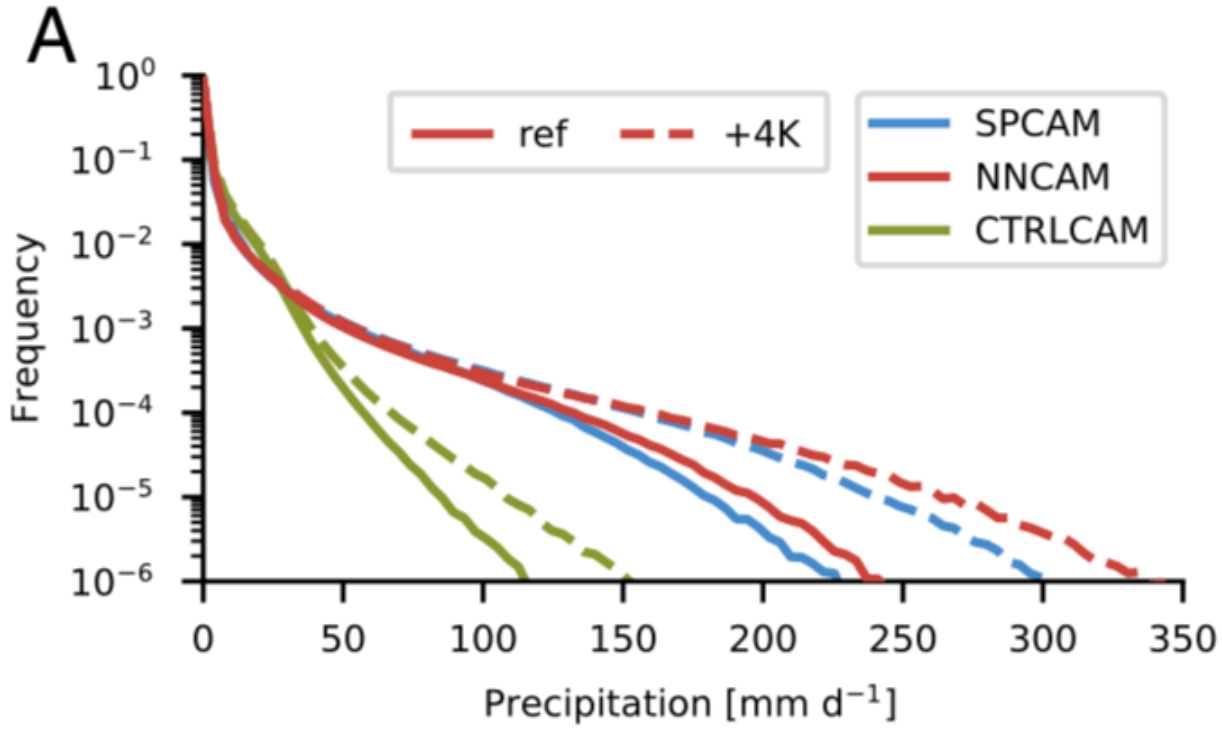


Outline

- **Motivation**: Parameterizations of small scale processes are the main cause of uncertainty in climate change **temperature and precipitation** projections
- **Previous studies:** Recent studies on machine-learning parameterization are promising (but still many problems)
- **Results**:
  a. Learning from a high resolution model (SAM)
  b. Example: Sub-grid tendencies are very important
  c. Offline results, conservation of energy, non-negative precipitation
  d. Online results
  e. Is there an optimal length scale for parameterization?

Recent attempts at machine learning parameterizations had some success but were not always stable/did not obey conservation laws/had climate drift

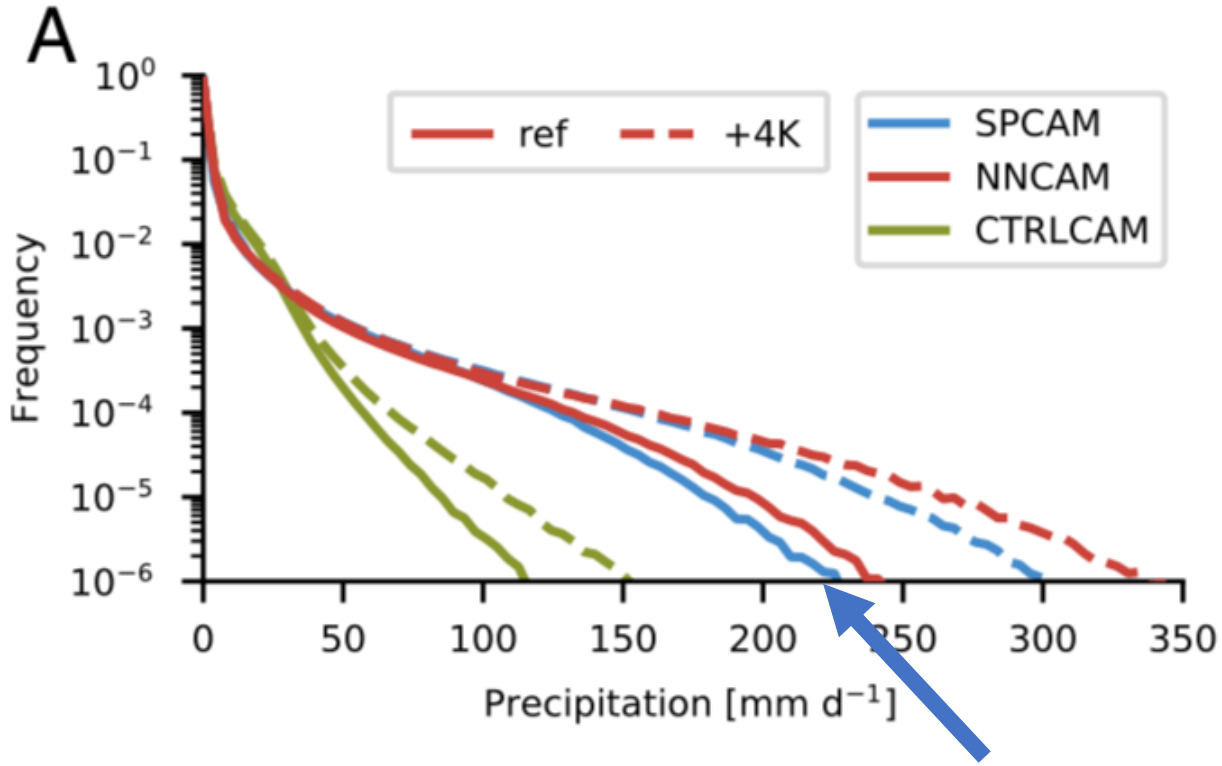
Neural Network to learn SP-CAM (super-parameterized model) sub-grid tendencies.



Rasp et al. (2018)

Recent attempts at machine learning parameterizations had some success but were not always stable/did not obey conservation laws/had climate drift

Neural Network to learn SP-CAM (super-parameterized model) sub-grid tendencies.

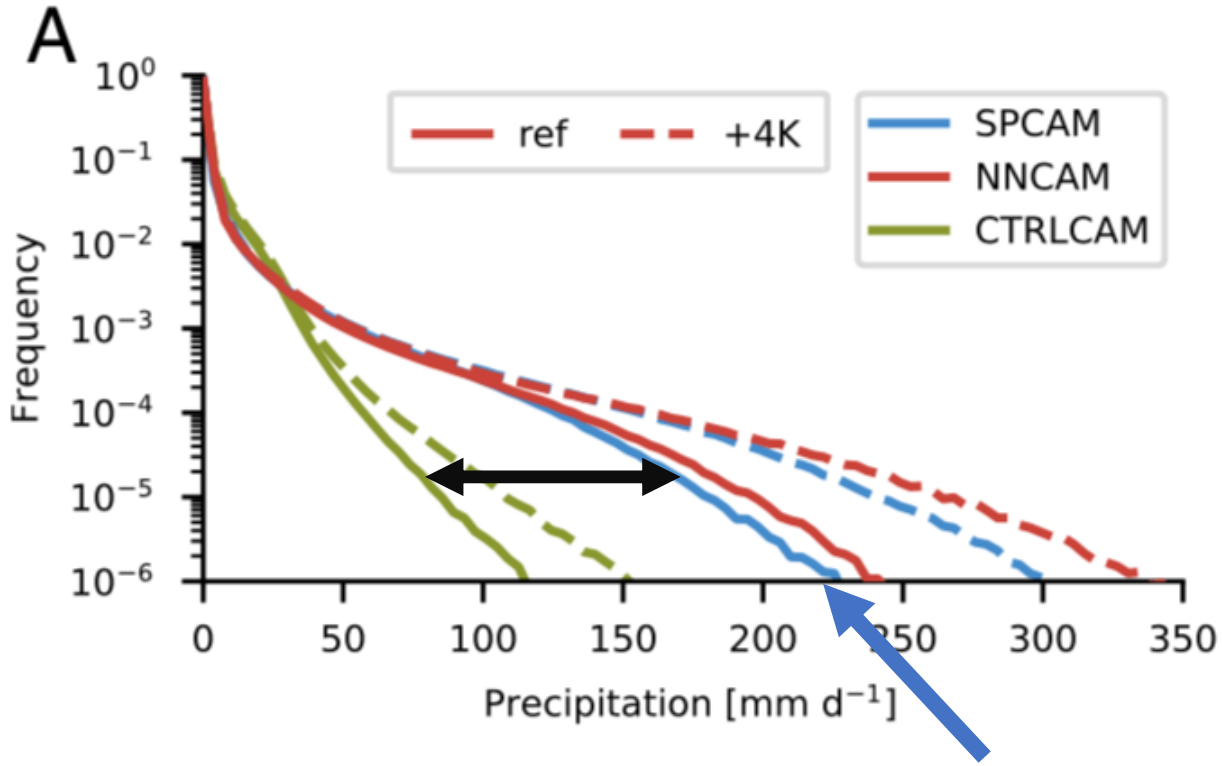


SP-CAM and Neural network-CAM has similar precipitation distribution

Rasp et al. (2018)

Recent attempts at machine learning parameterizations had some success but were not always stable/did not obey conservation laws/had climate drift

Neural Network to learn SP-CAM (super-parameterized model) sub-grid tendencies.

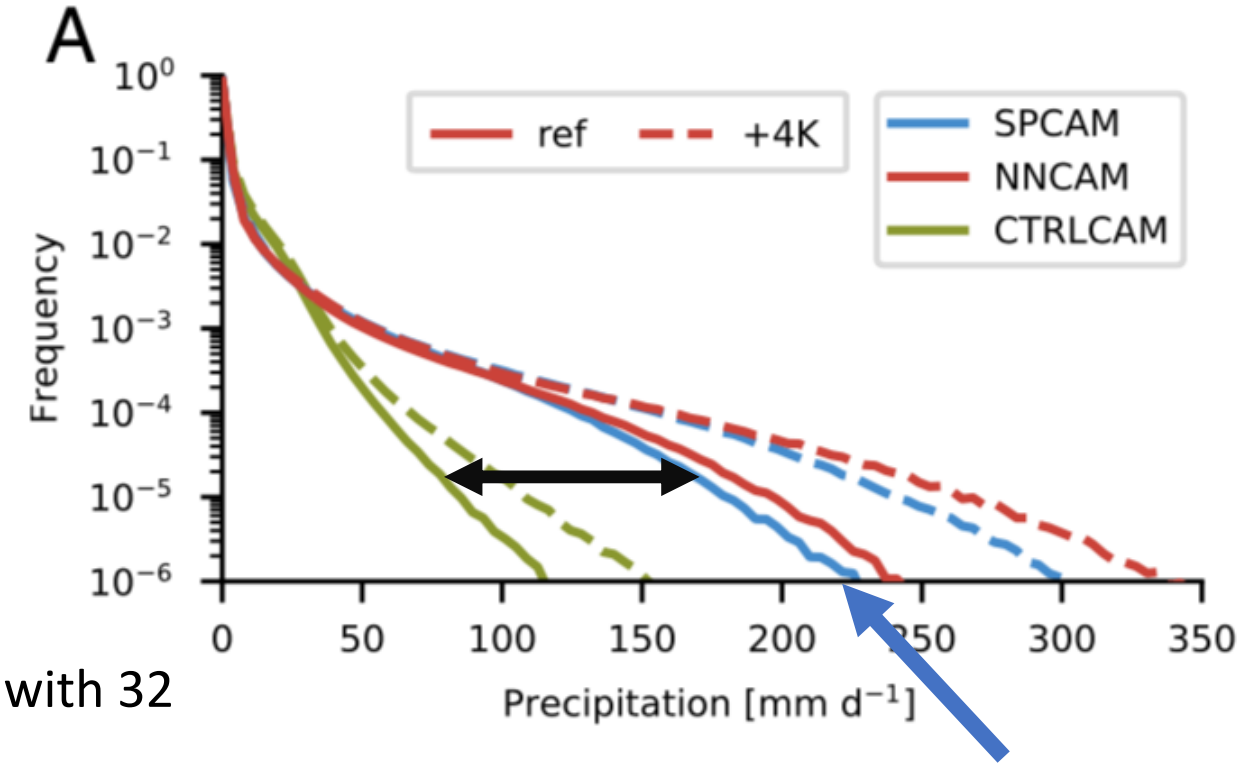


SP-CAM and Neural network-CAM has similar precipitation distribution

Rasp et al. (2018)

Recent attempts at machine learning parameterizations had some success but were not always stable/did not obey conservation laws/had climate drift

Neural Network to learn SP-CAM (super-parameterized model) sub-grid tendencies.



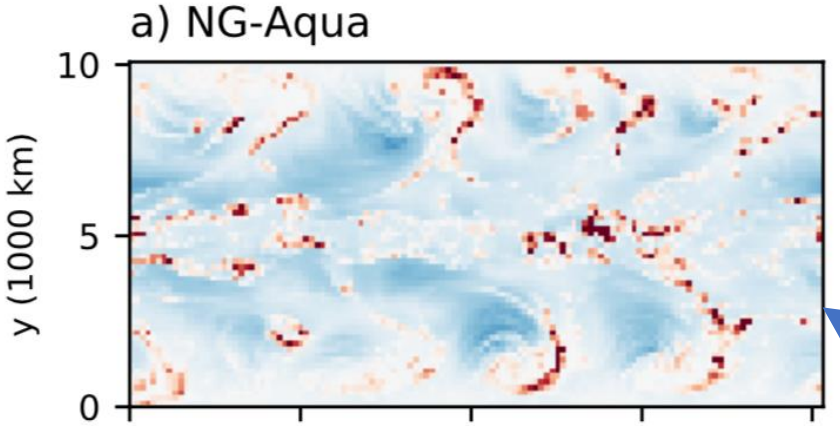
1. 2D CRM with 8 gridboxes (wasn't stable with 32 gridboxes)
2. Changes in the Neural Network architecture can lead to unstable simulations
3. No energy conservation (though see Beucler et.al. 2019)

SP-CAM and Neural network-CAM has similar precipitation distribution

Rasp et al. (2018)

Recent attempts at machine learning parameterizations had some success but were not always stable/did not obey conservation laws/had climate drift

Neural Network to learn a 3D cloud resolving model (SAM, 4km resolution) sub-grid tendencies.



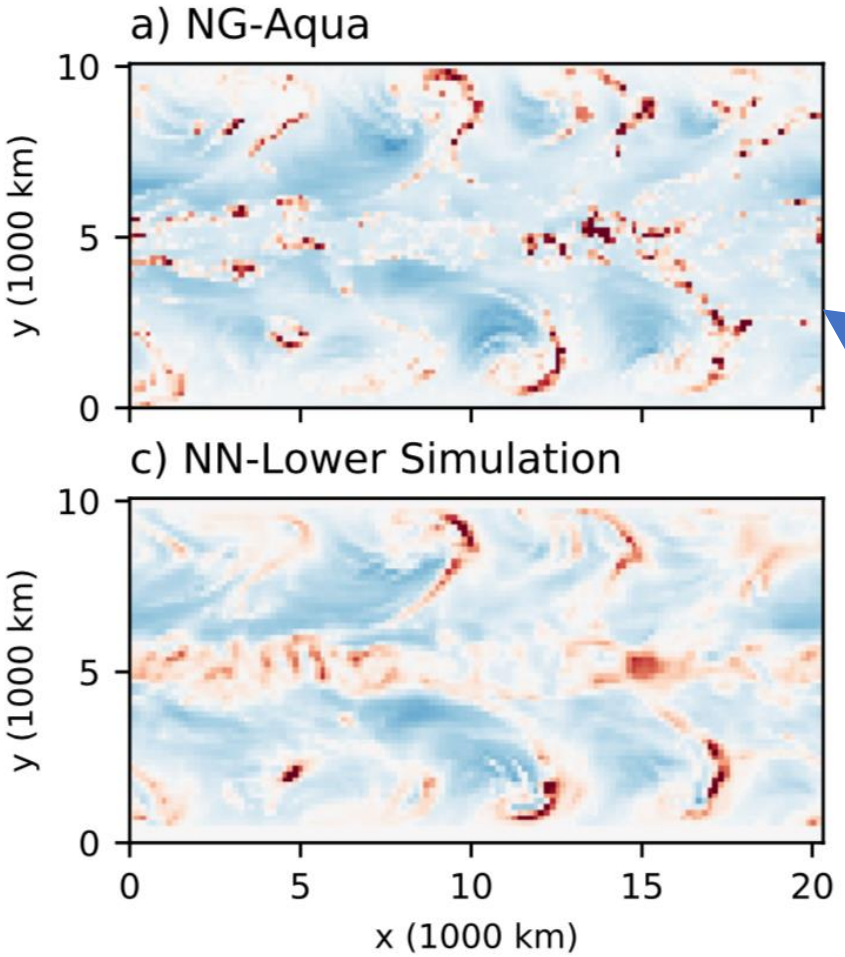
CRM (4 km resolution)

2-day precipitation forecast

Brenowitz and Bretherton (2019)

Recent attempts at machine learning parameterizations had some success but were not always stable/did not obey conservation laws/had climate drift

Neural Network to learn a 3D cloud resolving model (SAM, 4km resolution) sub-grid tendencies.



CRM (4 km resolution)

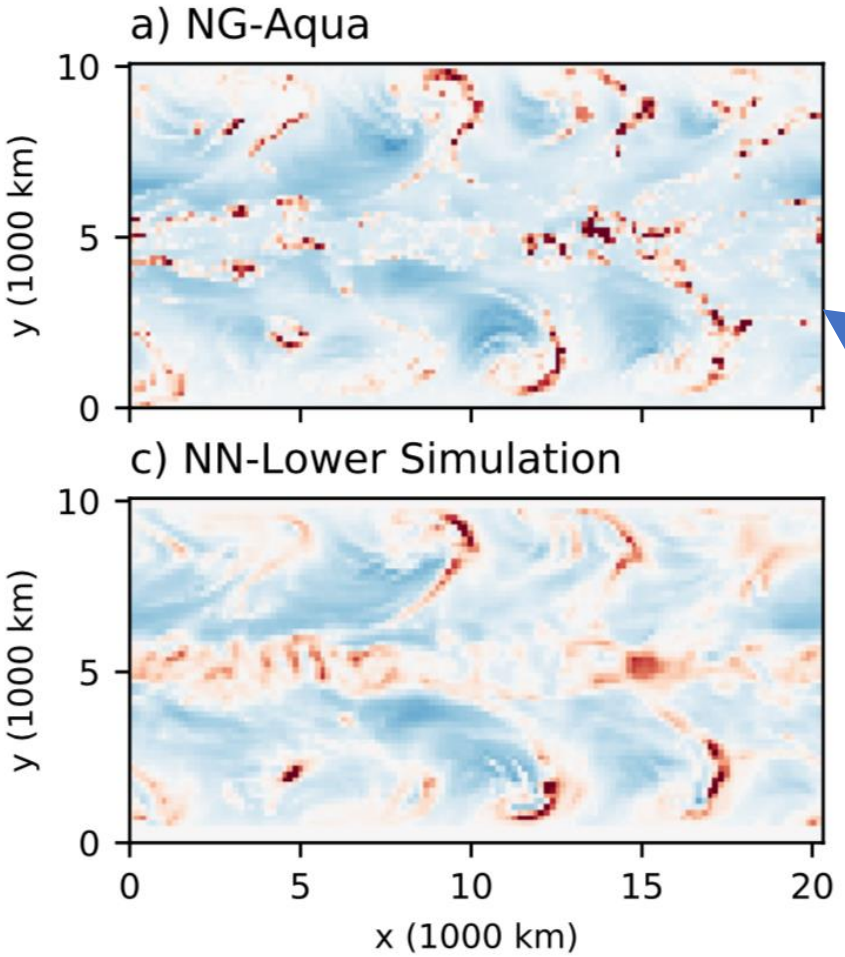
2-day precipitation forecast

160 km resolution + NN parameterization

Brenowitz and Bretherton (2019)

Recent attempts at machine learning parameterizations had some success but were not always stable/did not obey conservation laws/had climate drift

Neural Network to learn a 3D cloud resolving model (SAM, 4km resolution) sub-grid tendencies.



CRM (4 km resolution)

2-day precipitation forecast

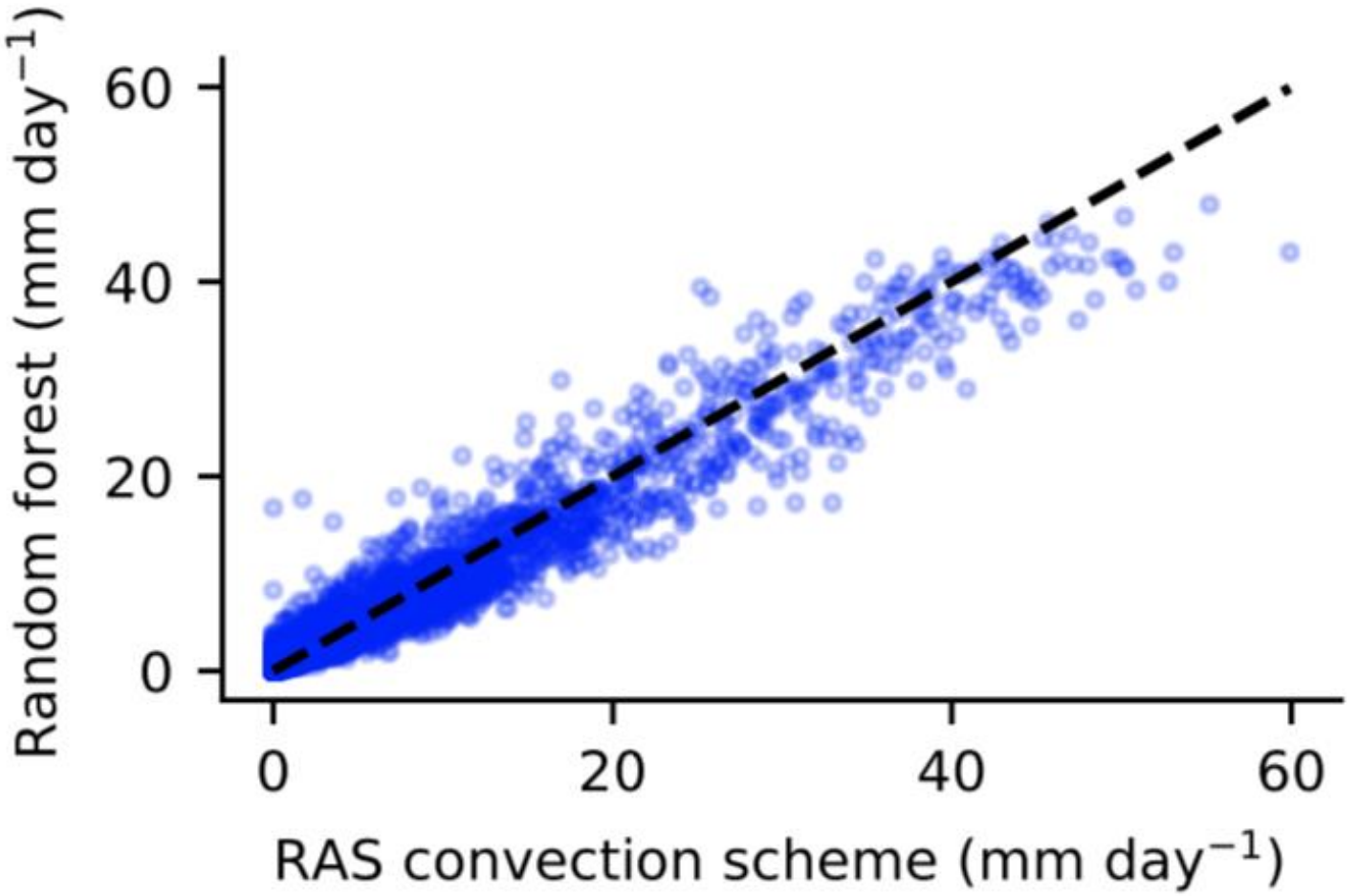
160 km resolution + NN parameterization

1. Drift of the mean climate: only short term prediction
2. Challenging to make it stable
3. No energy conservation

Brenowitz and Bretherton (2019)

Random forests is a machine learning algorithm that can give a parameterization that obey conservation laws

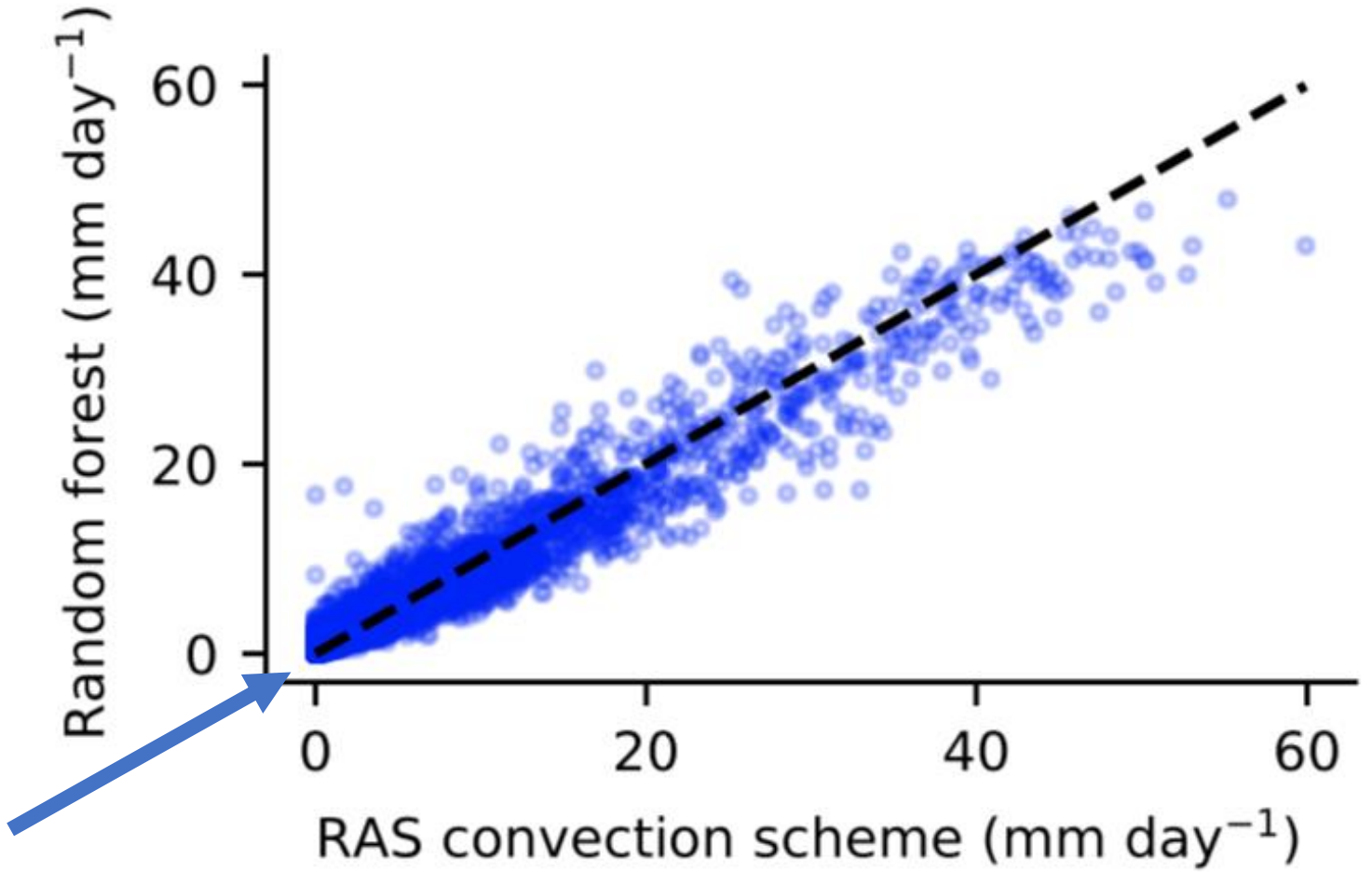
Random Forest could emulate the relaxed Arakawa-Schubert (RAS) convection scheme



O'Gorman and Dwyer (2018)

Random forests is a machine learning algorithm that can give a parameterization that obey conservation laws

Random Forest could emulate the relaxed Arakawa-Schubert (RAS) convection scheme



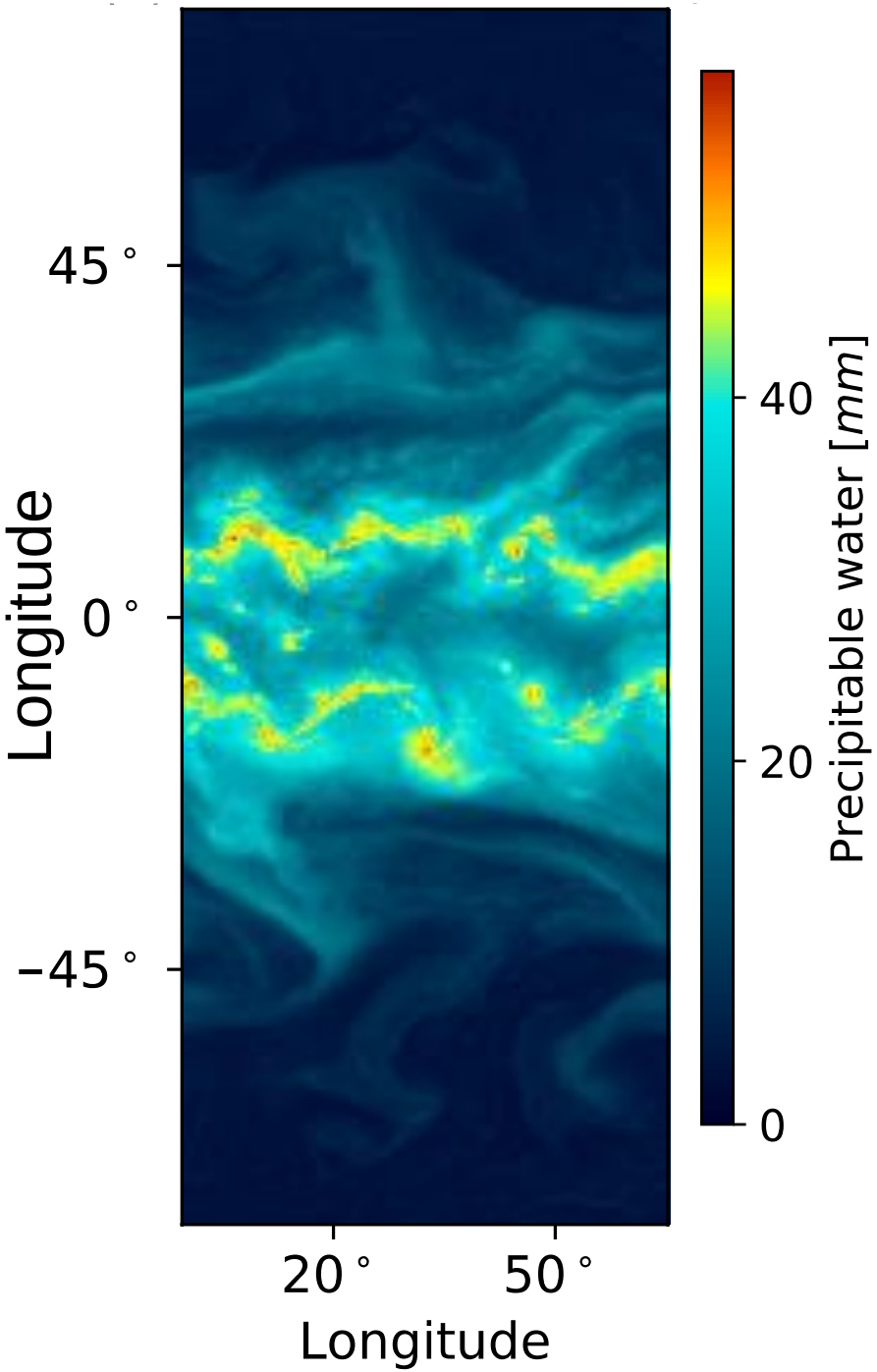
1. Stable when coupled to GCM
2. Conserves energy/non-negative precipitation

O'Gorman and Dwyer (2018)

Goal here: to run stable and accurate climate simulations learning from a fully 3D high-resolution simulation

Outline

- **Motivation**: Parameterizations of small scale processes are the main cause of uncertainty in climate change **temperature and precipitation** projections
- **Previous studies:** Recent studies on machine-learning parameterization are promising (but still many problems)
- **Results**:
  a. Learning from a high resolution model (SAM)
  b. Example: Sub-grid tendencies are very important
  c. Offline results, conservation of energy, non-negative precipitation
  d. Online results
  e. Is there an optimal length scale for parameterization?

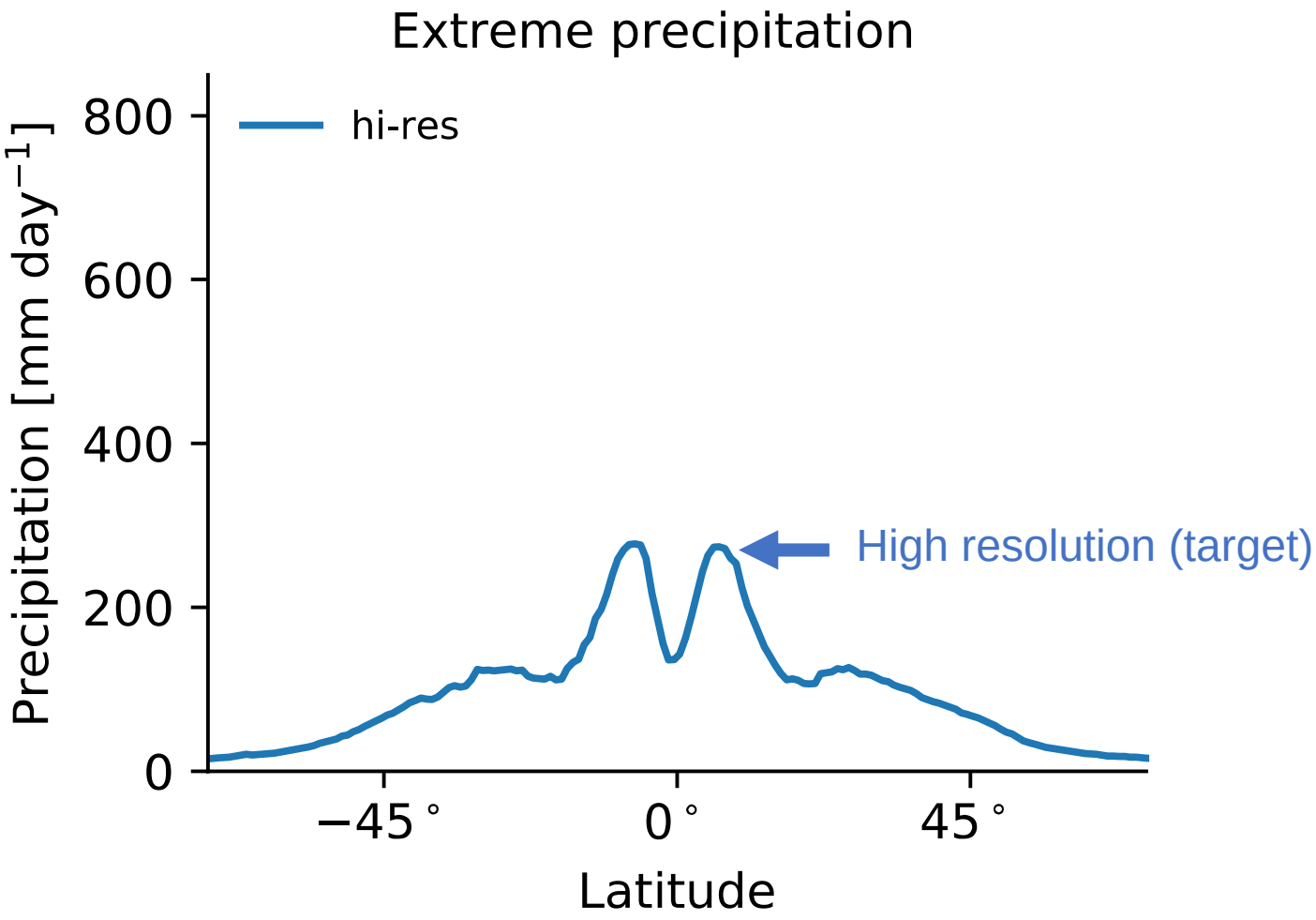


We learn from a quasi-global high-resolution model

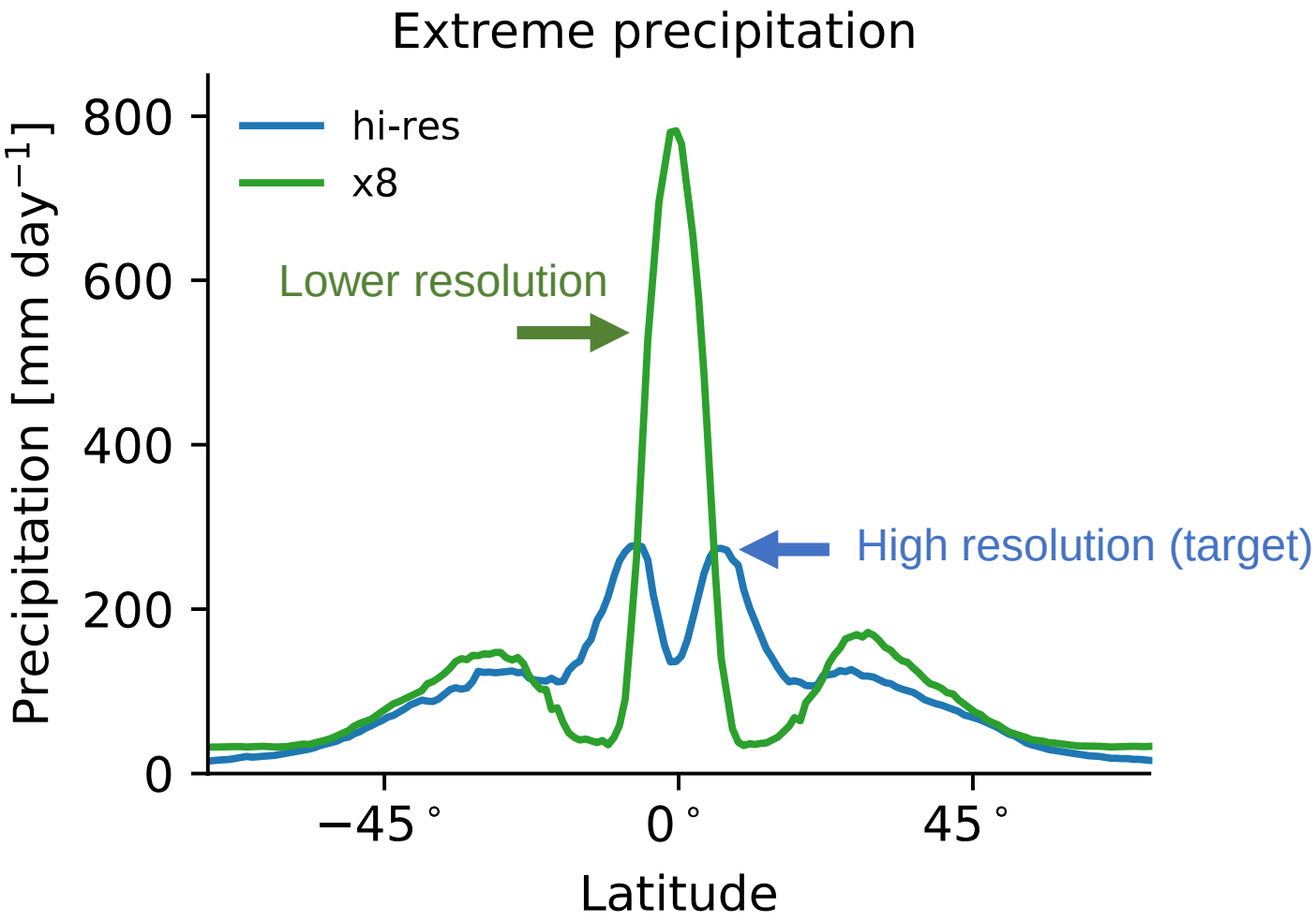
- SAM model with hypohydrostatic rescaling factor of 4 (grid spacing 12km), 48 vertical levels
- Specified SST distribution (qobs)
- Original simulation thanks to Bill Boos

SAM: *Khairoutdinov and Randall, 2003*
Hypohdrostatic/DARE: e.g., *Kuang et al. 2005*

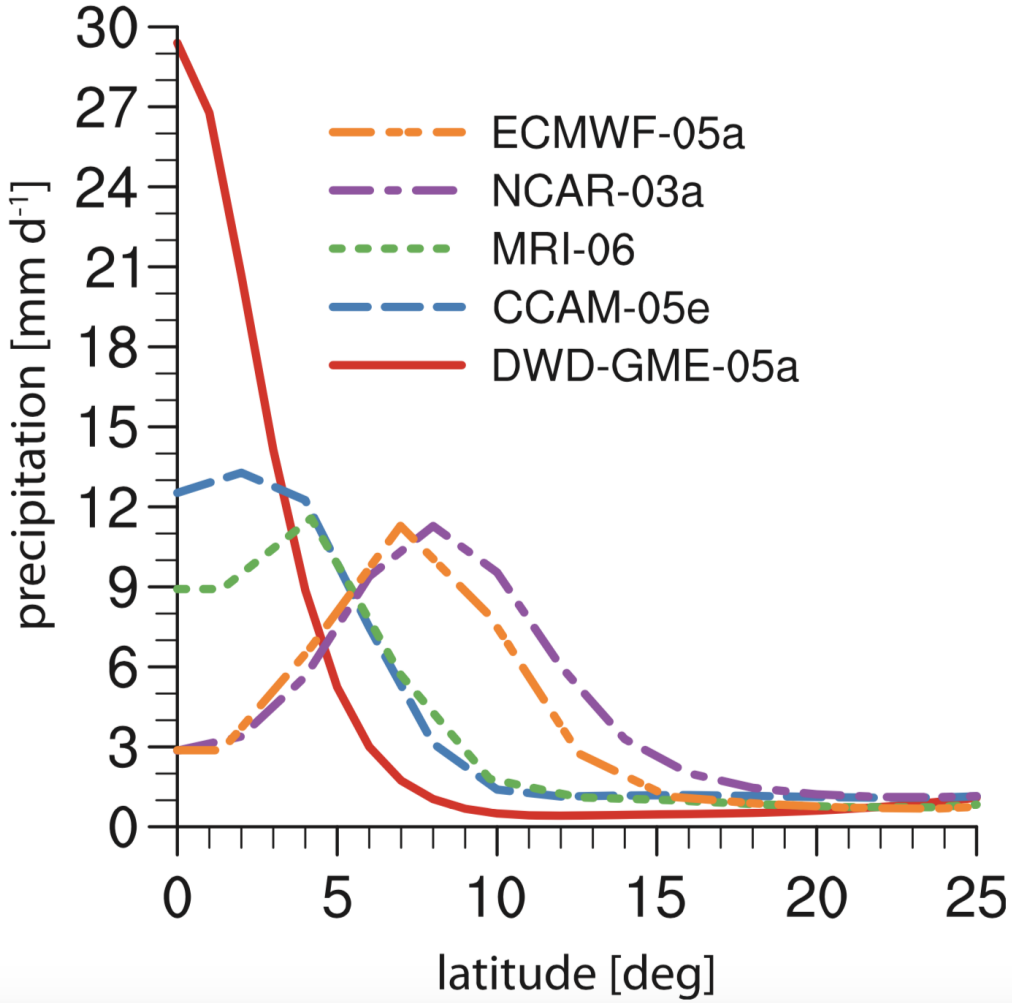
Extreme precipitation is different in high- and low-resolution simulations



Extreme precipitation is different in high- and low-resolution simulations



Privious studies: Aqua-planet simulations with different parameterization schemes lead to very different precipitation patterns



Mobis and Stevens 2012

Goal: to correct the tendencies of the thermodynamic and moisture variables due to sub-grid processes

$$h_L - \text{The liquid/ice water static stability energy}$$

Goal: to correct the tendencies of the thermodynamic and moisture variables due to sub-grid processes

h_L – The liquid/ice water static stability energy

$$h_{\mathrm{L}} = c_p T + gz - L_c(q_c + q_r) - L_s(q_i + q_s + q_g)$$

Liquid

Ice

Goal: to correct the tendencies of the thermodynamic and moisture variables due to sub-grid processes

h_L – The liquid/ice water static stability energy

$$h_{\mathrm{L}} = c_p T + gz - L_c(q_c + q_r) - L_s(q_i + q_s + q_g)$$

Liquid Ice

q_T – total non precipitating water mixing ratio

Goal: to correct the tendencies of the thermodynamic and moisture variables due to sub-grid processes

h_L – The liquid/ice water static stability energy

$$h_{\mathrm{L}} = c_p T + gz - L_c(q_c + q_r) - L_s(q_i + q_s + q_g)$$

Liquid Ice

q_T – total non precipitating water mixing ratio

q_p – total precipitating water mixing ratio

Goal: to correct the tendencies of the thermodynamic and moisture variables due to sub-grid processes

$$\frac{\partial h_L}{\partial t} = \text{advection} + \text{diffusion} + \text{radiation...}$$

$$\frac{\partial q_T}{\partial t} = \text{advection} + \text{diffusion} + \text{microphysics...}$$

$$\frac{\partial q_p}{\partial t} = \text{advection} + \text{precipitation} + \text{microphysics...}$$

Goal: to correct the tendencies of the thermodynamic and moisture variables due to sub-grid processes

$$\frac{\partial h_L}{\partial t} = \text{advection} + \text{diffusion} + \text{radiation...}$$

$$\frac{\partial q_T}{\partial t} = \text{advection} + \text{diffusion} + \text{microphysics...}$$

$$\frac{\partial q_p}{\partial t} = \text{advection} + \text{precipitation} + \text{microphysics...}$$

$$y = \left(\frac{\partial h_L}{\partial t}, \frac{\partial q_T}{\partial t}, \frac{\partial q_p}{\partial t} \right)_{\text{subgrid}}$$

Goal: to use a Random Forest (RF) to predict the effect of subgrid processes on the resolved thermodynamic and moisture variables

$$y = \left(\frac{\partial h_L}{\partial t}, \frac{\partial q_T}{\partial t}, \frac{\partial q_p}{\partial t} \right)_{\text{subgrid}}$$

$$y = \mathrm{RF}(x)$$

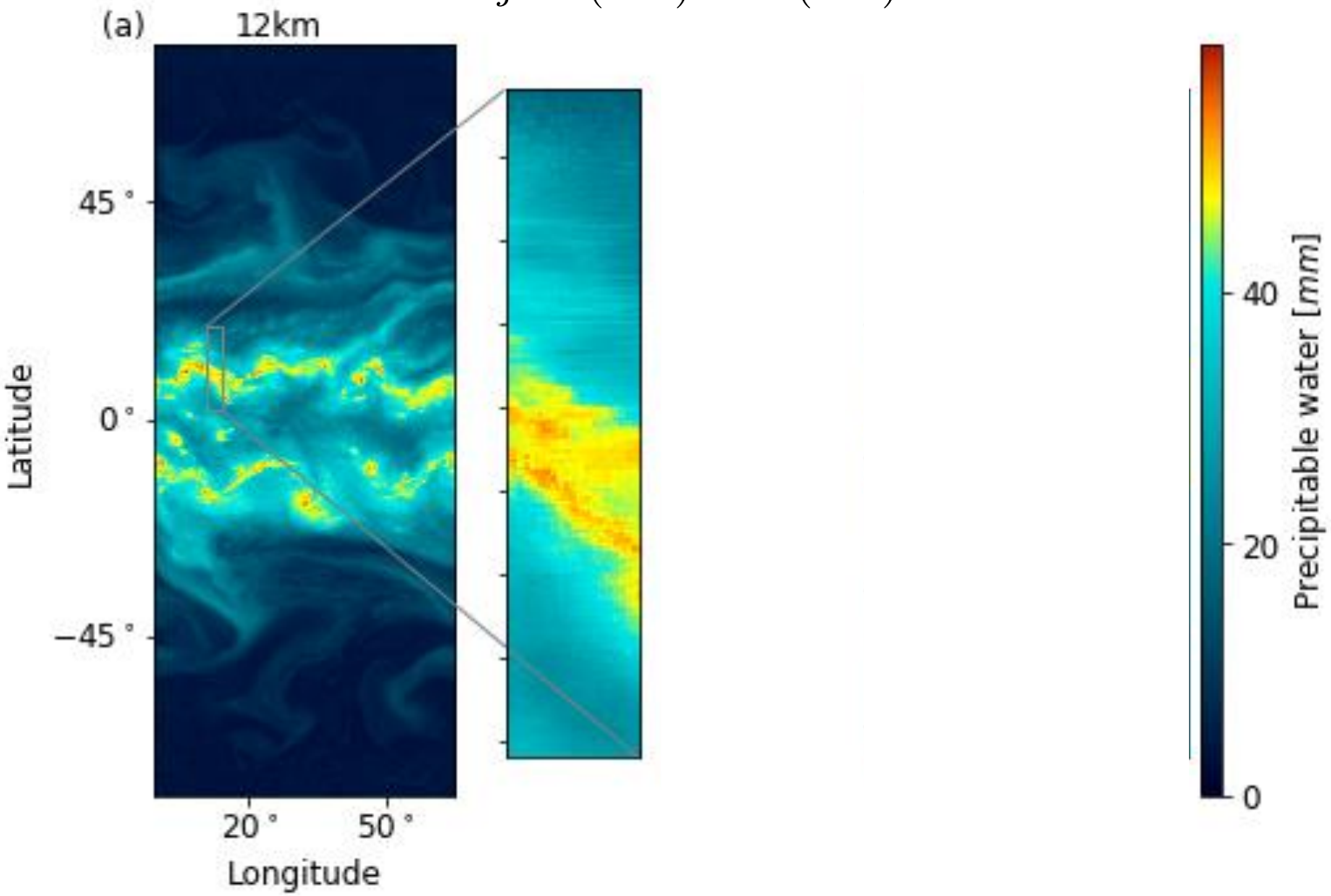
Goal: to use a Random Forest (RF) to predict the effect of subgrid processes on the resolved thermodynamic and moisture variables

$$y = \left(\frac{\partial h_L}{\partial t}, \frac{\partial q_T}{\partial t}, \frac{\partial q_p}{\partial t} \right)_{\text{subgrid}}$$

$$y = \mathrm{RF}(x)$$

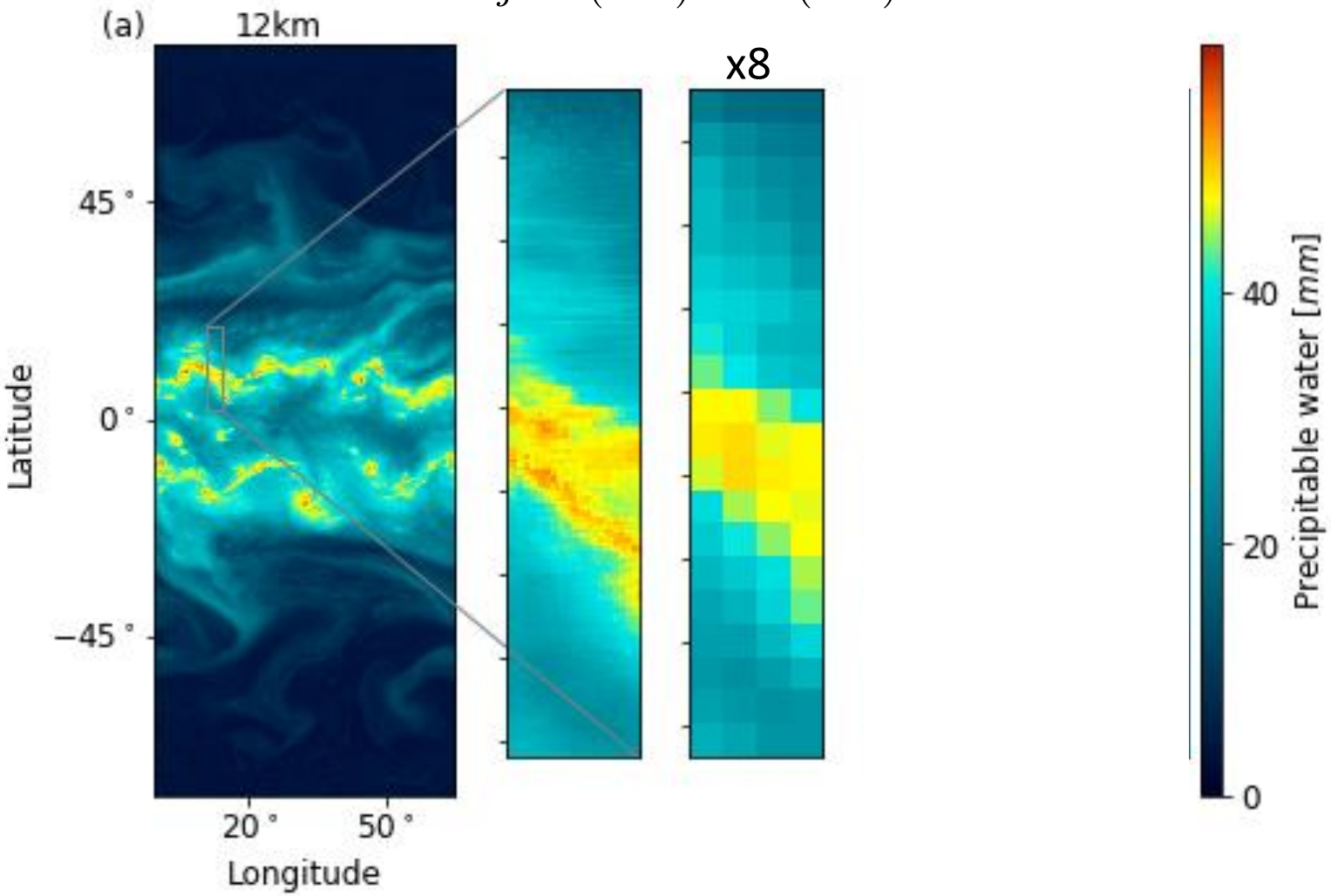
Coarse grained field = average of a field over a number of grid boxes

$$\overline{Q}(x_i, y_l, z_k) = \frac{1}{N^2} \sum_{j=N(i-1)}^{j=Ni} \sum_{h=N(l-1)}^{h=Nl} Q(x_j, y_h, z_k)_{\mathrm{high-res}}$$



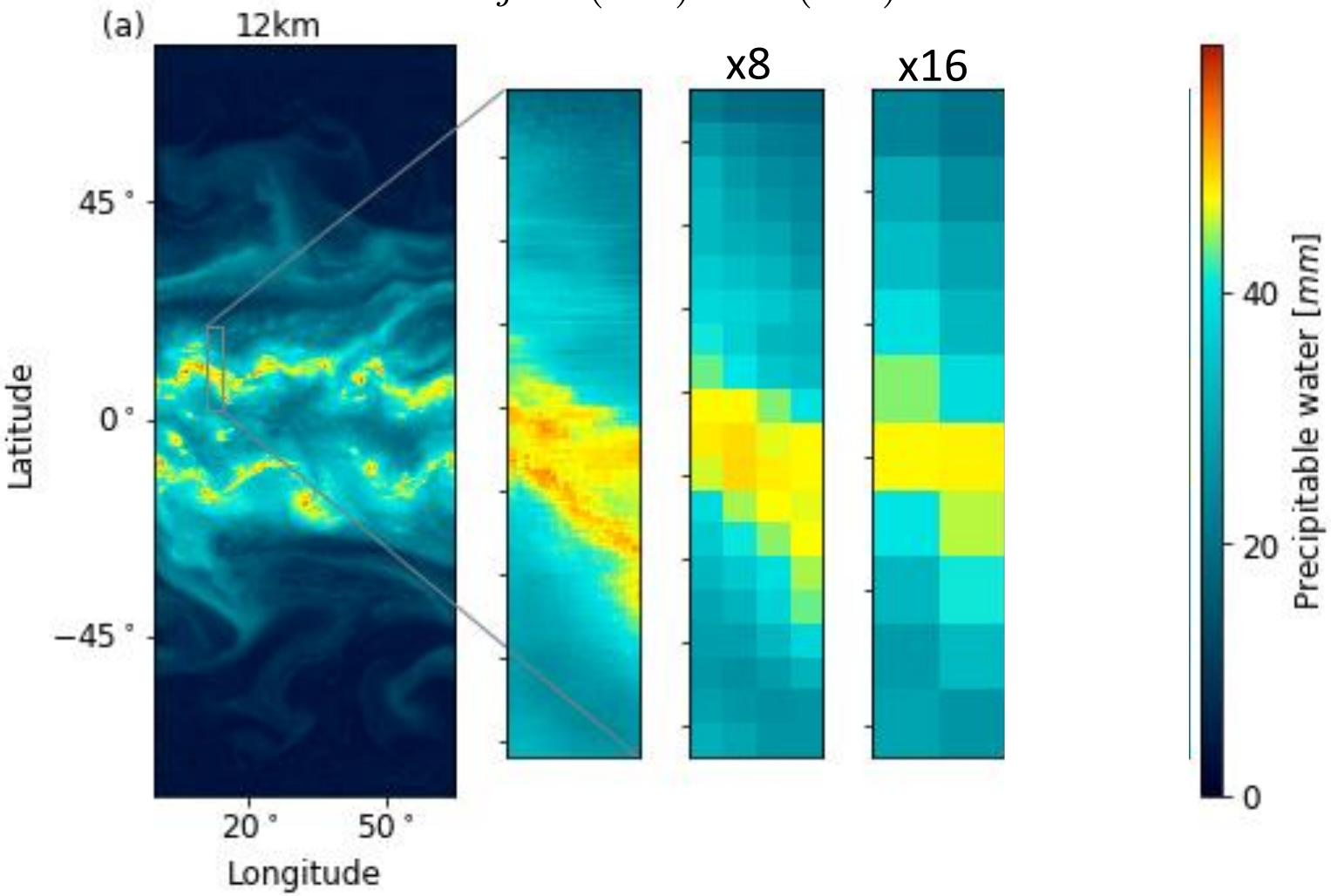
Coarse grained field = average of a field over a number of grid boxes

$$\overline{Q}(x_i, y_l, z_k) = \frac{1}{N^2} \sum_{j=N(i-1)}^{j=Ni} \sum_{h=N(l-1)}^{h=Nl} Q(x_j, y_h, z_k)_{\mathrm{high-res}}$$



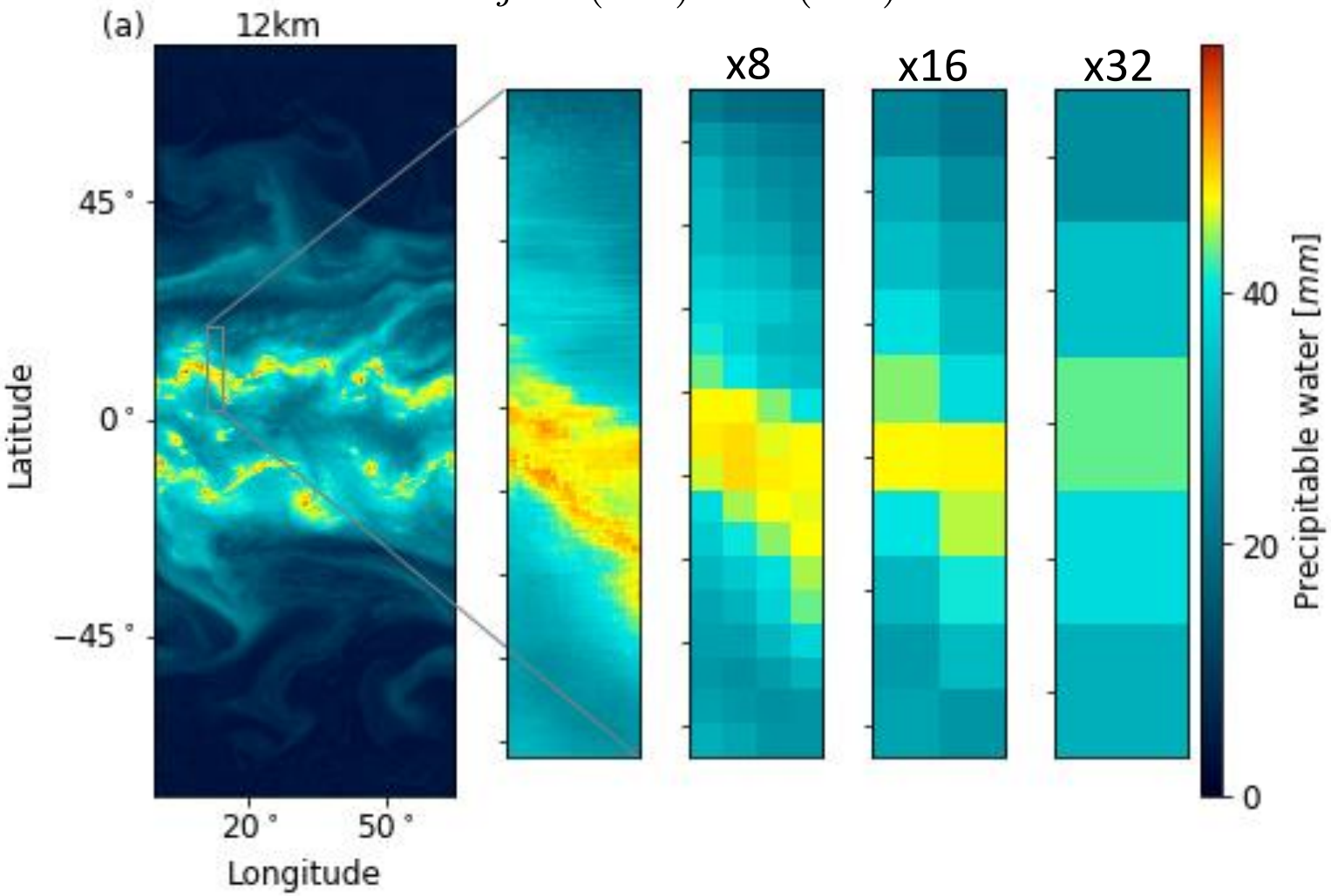
Coarse grained field = average of a field over a number of grid boxes

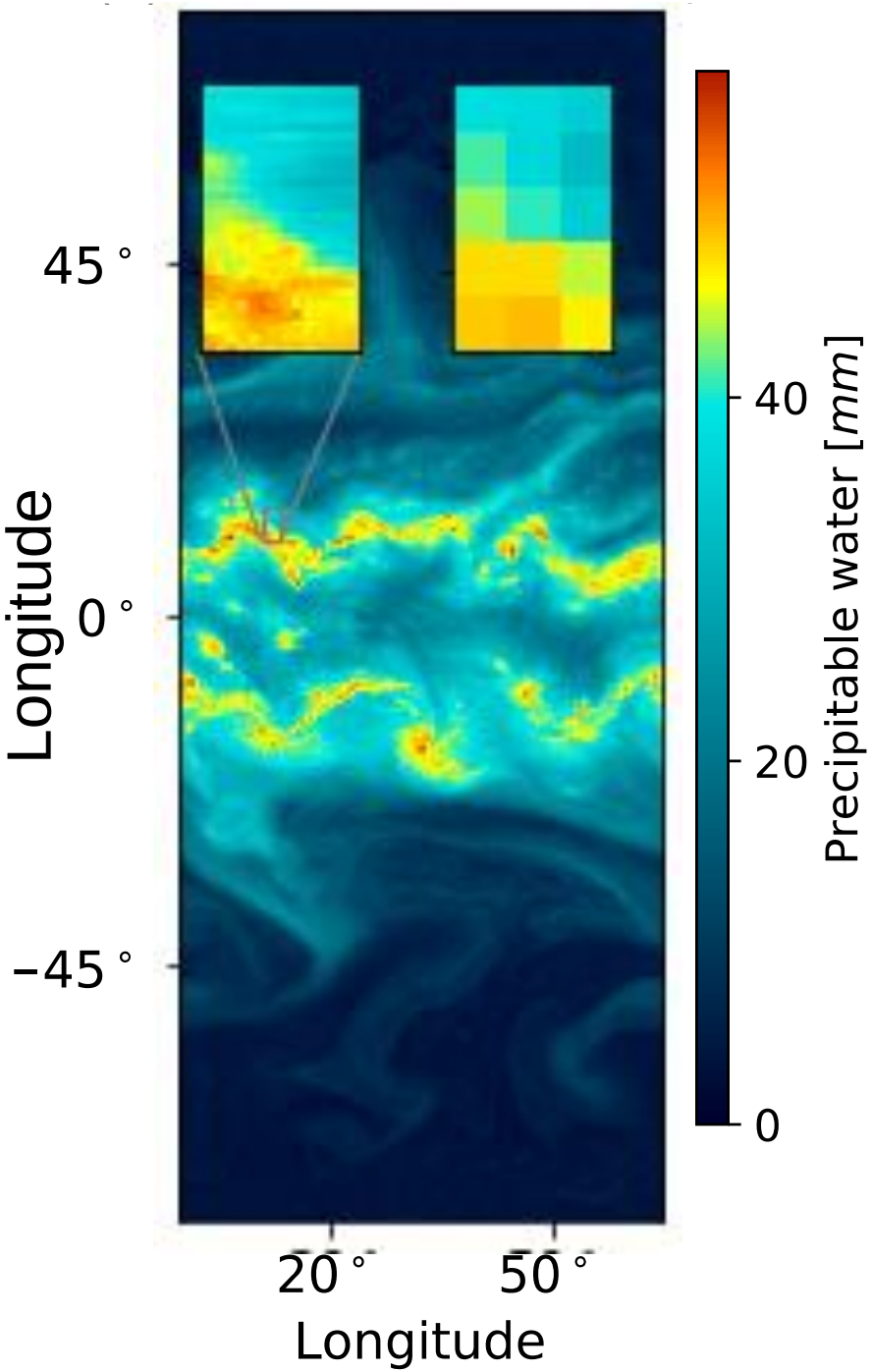
$$\overline{Q}(x_i, y_l, z_k) = \frac{1}{N^2} \sum_{j=N(i-1)}^{j=Ni} \sum_{h=N(l-1)}^{h=Nl} Q(x_j, y_h, z_k)_{\mathrm{high-res}}$$



Coarse grained field = average of a field over a number of grid boxes

$$\overline{Q}(x_i, y_l, z_k) = \frac{1}{N^2} \sum_{j=N(i-1)}^{j=Ni} \sum_{h=N(l-1)}^{h=Nl} Q(x_j, y_h, z_k)_{\mathrm{high-res}}$$



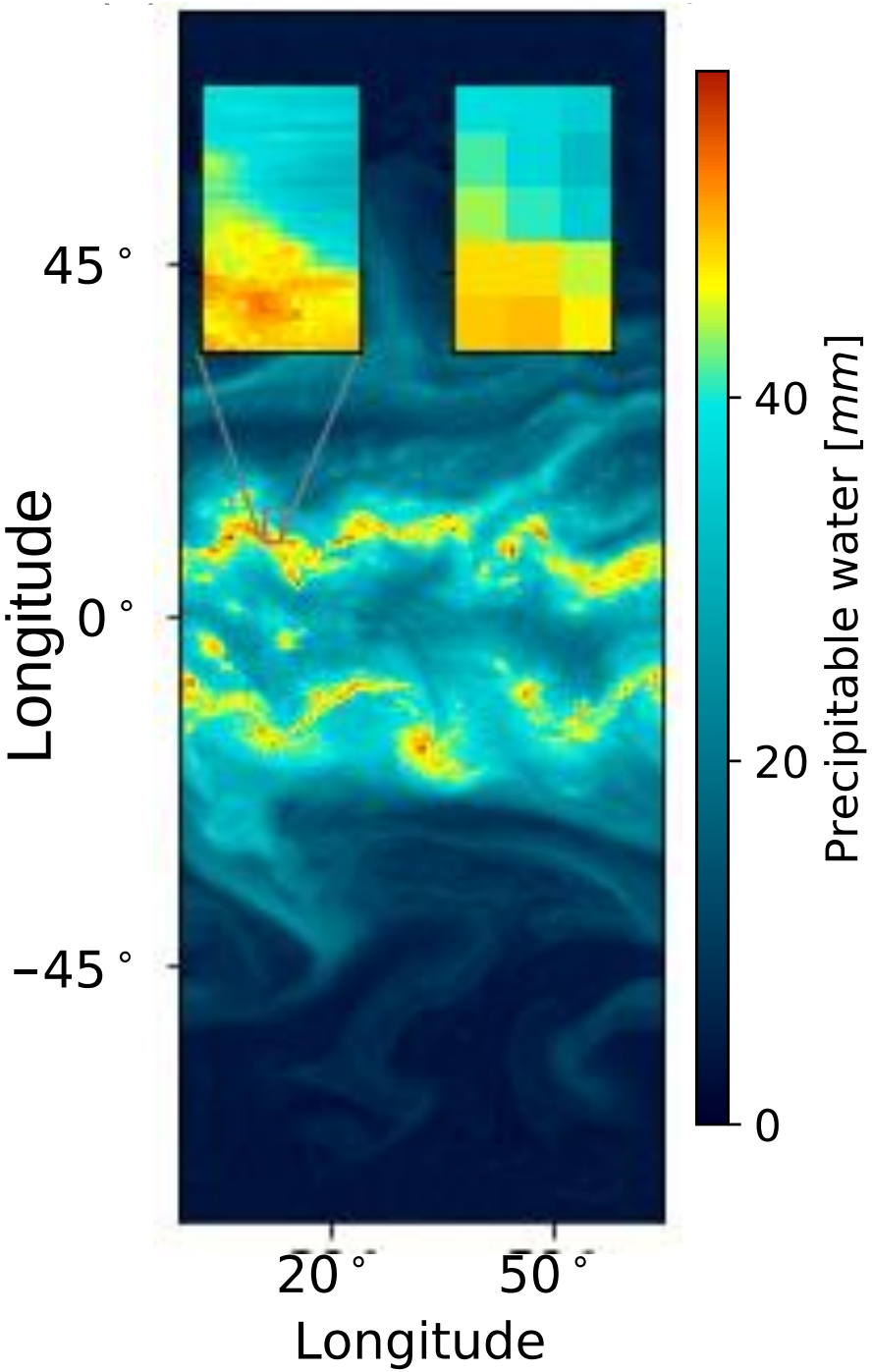


Subgrid tendency is the difference between the coarse-grained (high-res) tendency and the resolved tendency

$$-\overline{\frac{\partial w h_L}{\partial z}}$$

Coarse grained advection from high resolution simulation

Yuval & O'Gorman, Nat. Com. (in press)



Subgrid tendency is the difference between the coarse-grained (high-res) tendency and the resolved tendency

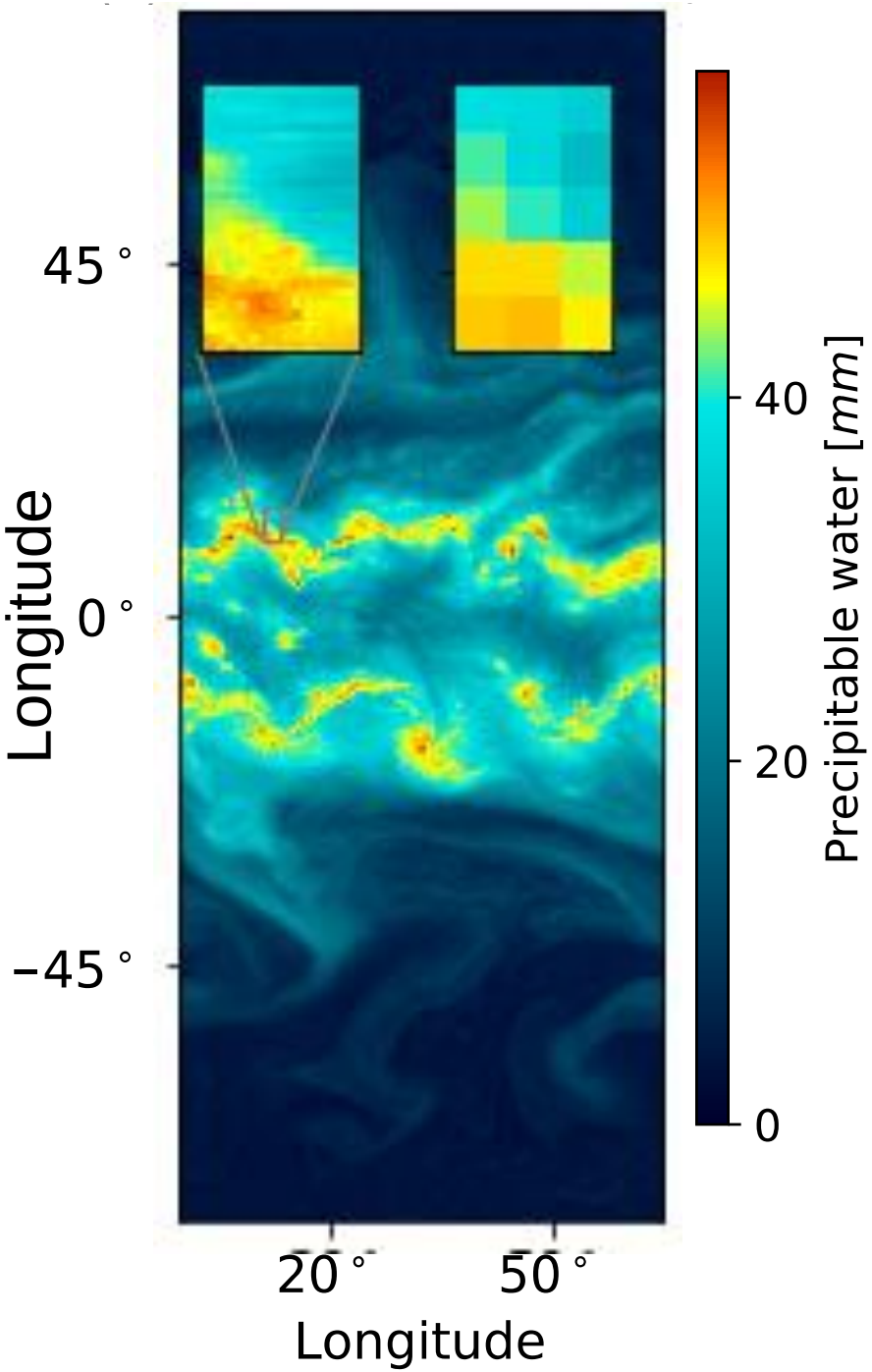
$$-\overline{\frac{\partial w h_L}{\partial z}}$$

Coarse grained advection from high resolution simulation

$$-\frac{\partial \overline{w}\,\overline{h}_L}{\partial z}$$

Resolved advection

Yuval & O'Gorman, Nat. Com. (in press)



Subgrid tendency is the difference between the coarse-grained (high-res) tendency and the resolved tendency

$$-\overline{\frac{\partial w h_L}{\partial z}}$$

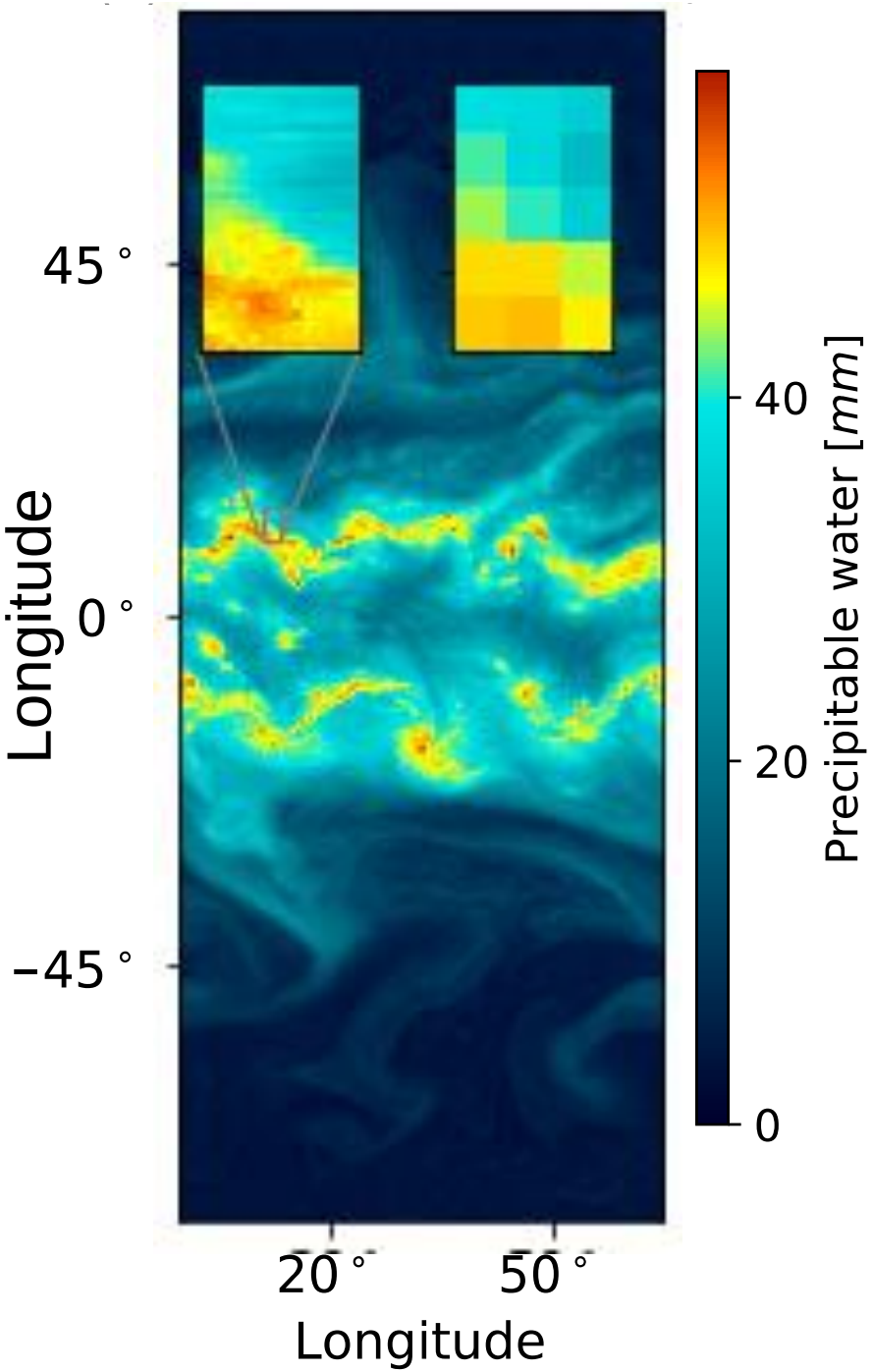
Coarse grained advection from high resolution simulation

$$-\frac{\partial \overline{w}\,\overline{h}_L}{\partial z}$$

Resolved advection

$$\left(\frac{\partial \overline{h}_L}{\partial t}\right)^{\text{advection}}_{\text{subgrid}} = -\left(\overline{\frac{\partial w h_L}{\partial z}} - \frac{\partial \overline{w}\,\overline{h}_L}{\partial z}\right).$$

Yuval & O'Gorman, Nat. Com. (in press)



Subgrid tendency is the difference between the coarse-grained (high-res) tendency and the resolved tendency

$$-\overline{\frac{\partial w h_L}{\partial z}}$$

Coarse grained advection from high resolution simulation

$$-\frac{\partial \overline{w}\,\overline{h}_L}{\partial z}$$

Resolved advection

$$\left(\frac{\partial \overline{h}_L}{\partial t}\right)^{\text{advection}}_{\text{subgrid}} = -\left(\overline{\frac{\partial w h_L}{\partial z}} - \frac{\partial \overline{w}\,\overline{h}_L}{\partial z}\right).$$

- Parameterization includes: microphysics, vertical advection, boundary-layer turbulence and radiation

Yuval & O'Gorman, Nat. Commun. (in press)

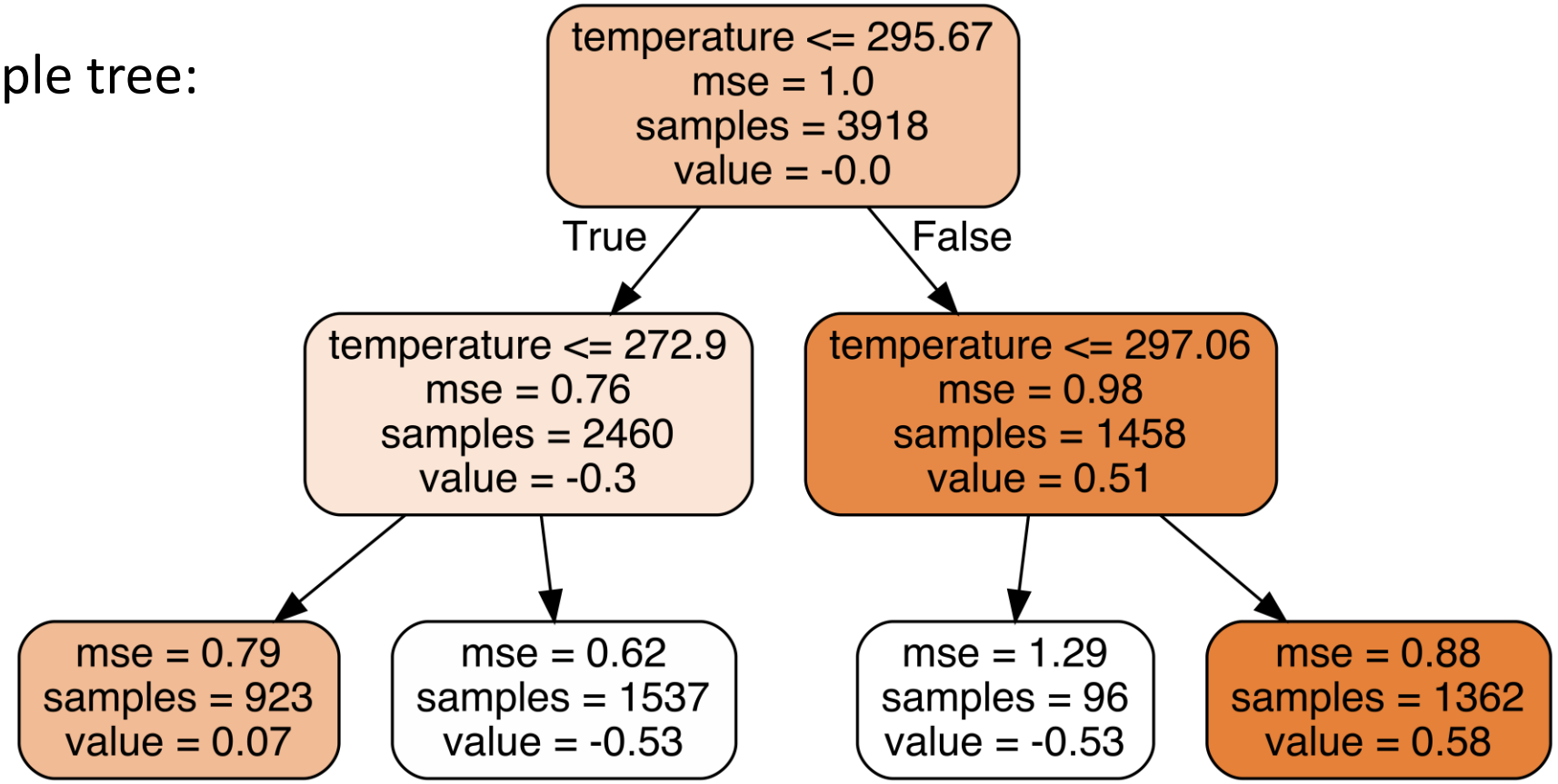
Goal: to use a Random Forest (RF) to predict the effect of subgrid processes on the resolved thermodynamic and moisture variables

$$y = \left(\frac{\partial h_L}{\partial t}, \frac{\partial q_T}{\partial t}, \frac{\partial q_p}{\partial t} \right)_{\text{subgrid}}$$

$$y = \mathrm{RF}(x)$$

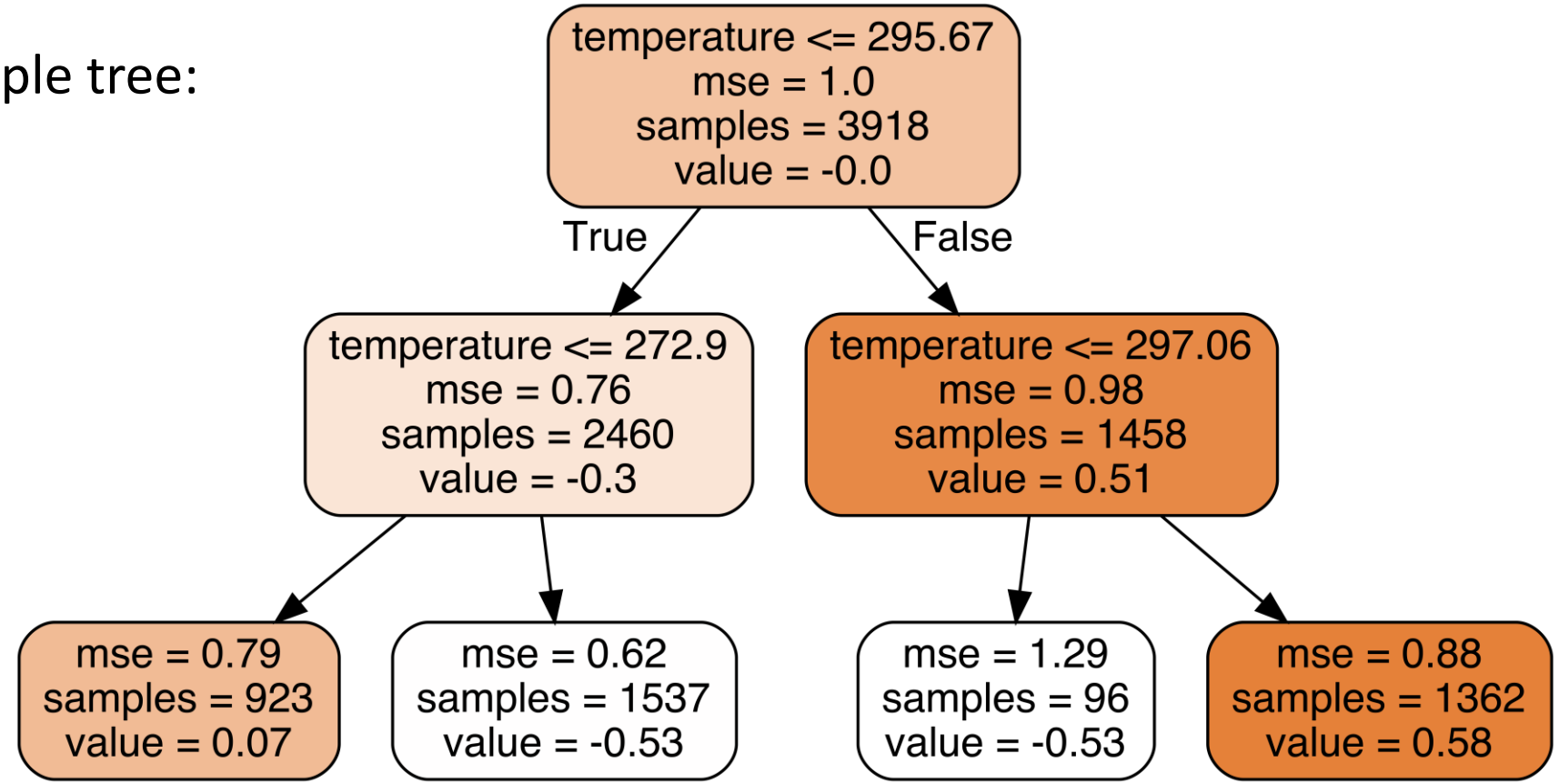
Machine-learning algorithm for parameterization: Random-forest (Breiman 2001)

One simple tree:



Machine-learning algorithm for parameterization: Random-forest (Breiman 2001)

One simple tree:



- Random-forest is stable and robust when implemented in a GCM
- Respects physical constraints (energy conservation, non-negative precipitation)

O'Gorman & Dwyer 2018

RF Inputs: vertical structure of temperature and humidity
RF Outputs: vertical structure of instantaneous tendencies due to subgrid scales

$$RF(X) \rightarrow y$$

$$X = (T(z), q_T(z), q_p(z), u(z), v(z), \text{latitude})$$

RF Inputs: vertical structure of temperature and humidity
RF Outputs: vertical structure of instantaneous tendencies due to subgrid scales

$$RF(X) \rightarrow y$$

$$X = (T(z), q_T(z), q_p(z), u(z), v(z), \text{latitude})$$

Latitude is a proxy for SST, albedo etc.

RF Inputs: vertical structure of temperature and humidity
RF Outputs: vertical structure of instantaneous tendencies due to subgrid scales

$$RF(X) \rightarrow y$$

$$X = (T(z), q_T(z), q_p(z), u(z), v(z), \text{latitude})$$

Latitude is a proxy for SST, albedo etc.

$$y = \left(\frac{\partial h_L(z)}{\partial t}_{\text{sub}}, \frac{\partial q_T(z)}{\partial t}_{\text{sub}}, \frac{\partial q_p(z)}{\partial t}_{\text{sub}}, \text{diffusivity(z)} \right)$$

RF Inputs: vertical structure of temperature and humidity
RF Outputs: vertical structure of instantaneous tendencies due to subgrid scales

$$RF(X) \rightarrow y$$

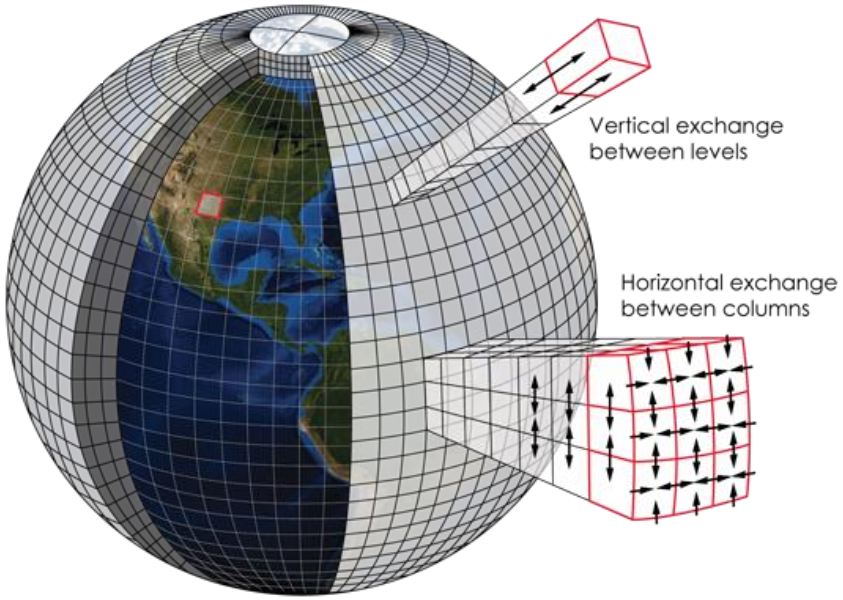
$$X = (T(z), q_T(z), q_p(z), u(z), v(z), \text{latitude})$$

Latitude is a proxy for SST, albedo etc.

$$y = \left(\frac{\partial h_L(z)}{\partial t}_{\text{sub}}, \frac{\partial q_T(z)}{\partial t}_{\text{sub}}, \frac{\partial q_p(z)}{\partial t}_{\text{sub}}, \text{diffusivity(z)} \right)$$

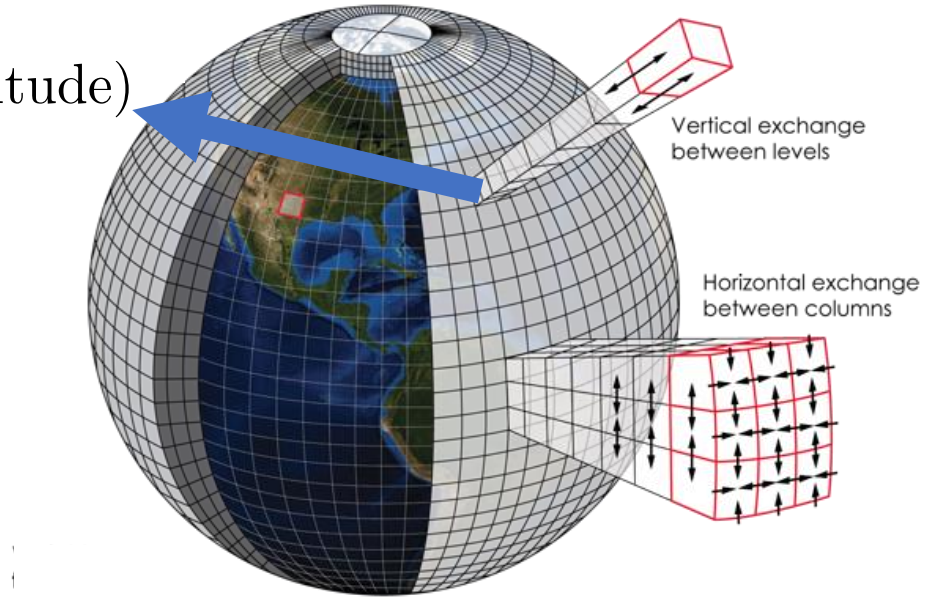
- Single diffusivity for all thermodynamic and moisture variables
- Flux is always down gradient

Coupling the random forest to SAM



Coupling the random forest to SAM

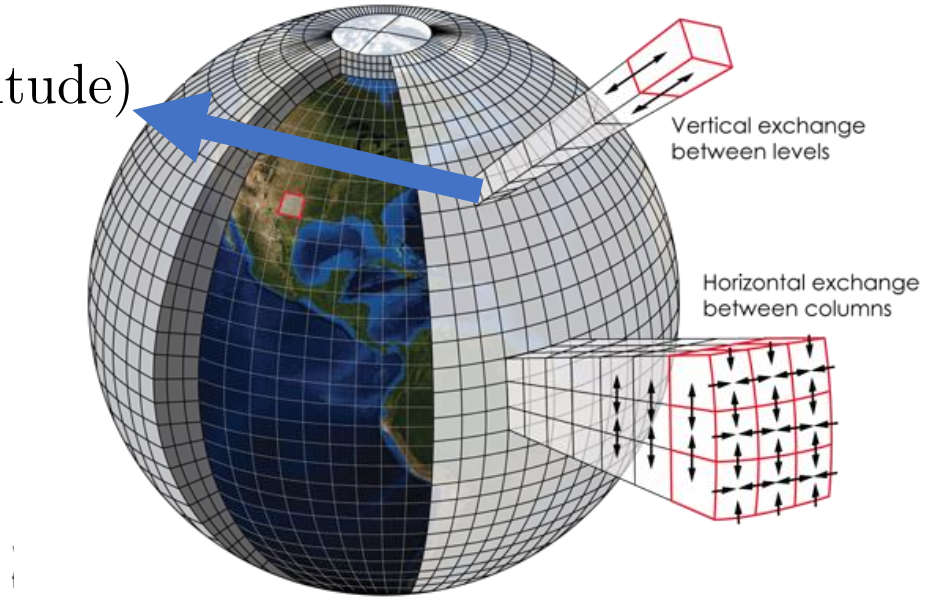
$$X = (T(z), q_T(z), q_p(z), u(z), v(z), \text{latitude})$$



Coupling the random forest to SAM

$$X = (T(z), q_T(z), q_p(z), u(z), v(z), \text{latitude})$$

$$RF(X) \rightarrow y$$

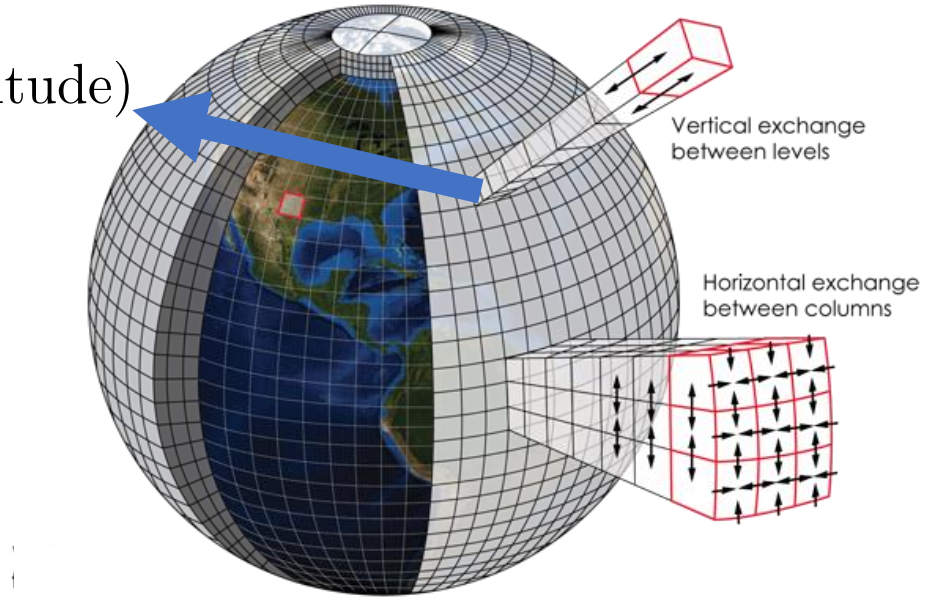


Coupling the random forest to SAM

$$X = (T(z), q_T(z), q_p(z), u(z), v(z), \text{latitude})$$

$$RF(X) \rightarrow y$$

$$y = \left(\frac{\partial h_L}{\partial t}_{\text{sub}}, \frac{\partial q_T}{\partial t}_{\text{sub}}, \frac{\partial q_p}{\partial t}_{\text{sub}}, \text{diffusivity} \right)$$

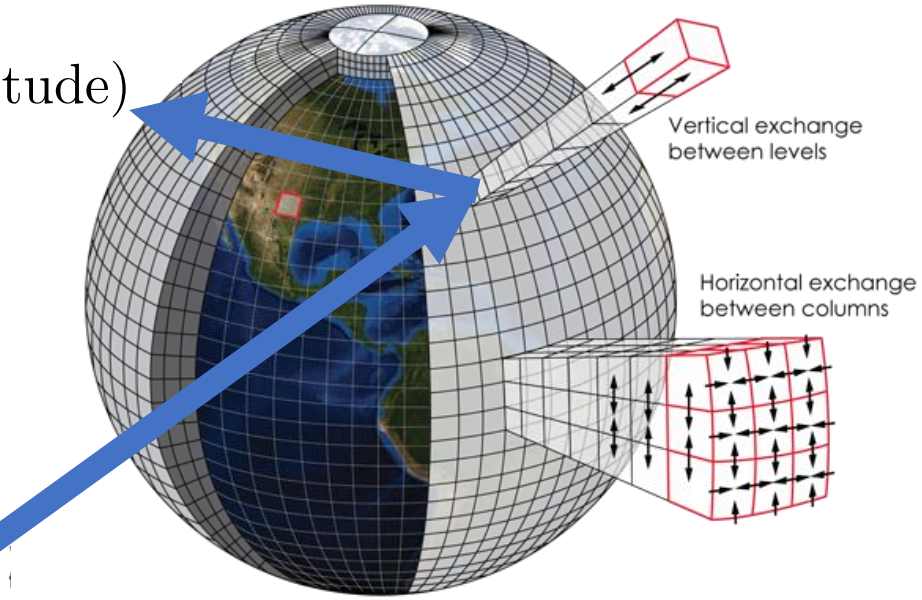


Coupling the random forest to SAM

$$X = (T(z), q_T(z), q_p(z), u(z), v(z), \text{latitude})$$

$$RF(X) \rightarrow y$$

$$y = \left(\frac{\partial h_L}{\partial t}_{\text{sub}}, \frac{\partial q_T}{\partial t}_{\text{sub}}, \frac{\partial q_p}{\partial t}_{\text{sub}}, \text{diffusivity} \right)$$

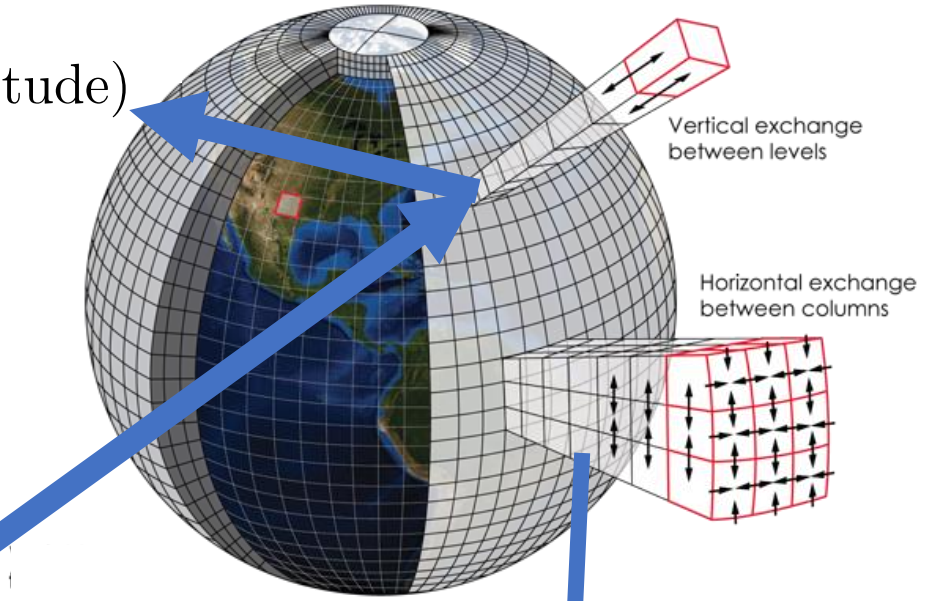


Coupling the random forest to SAM

$$X = (T(z), q_T(z), q_p(z), u(z), v(z), \text{latitude})$$

$$RF(X) \rightarrow y$$

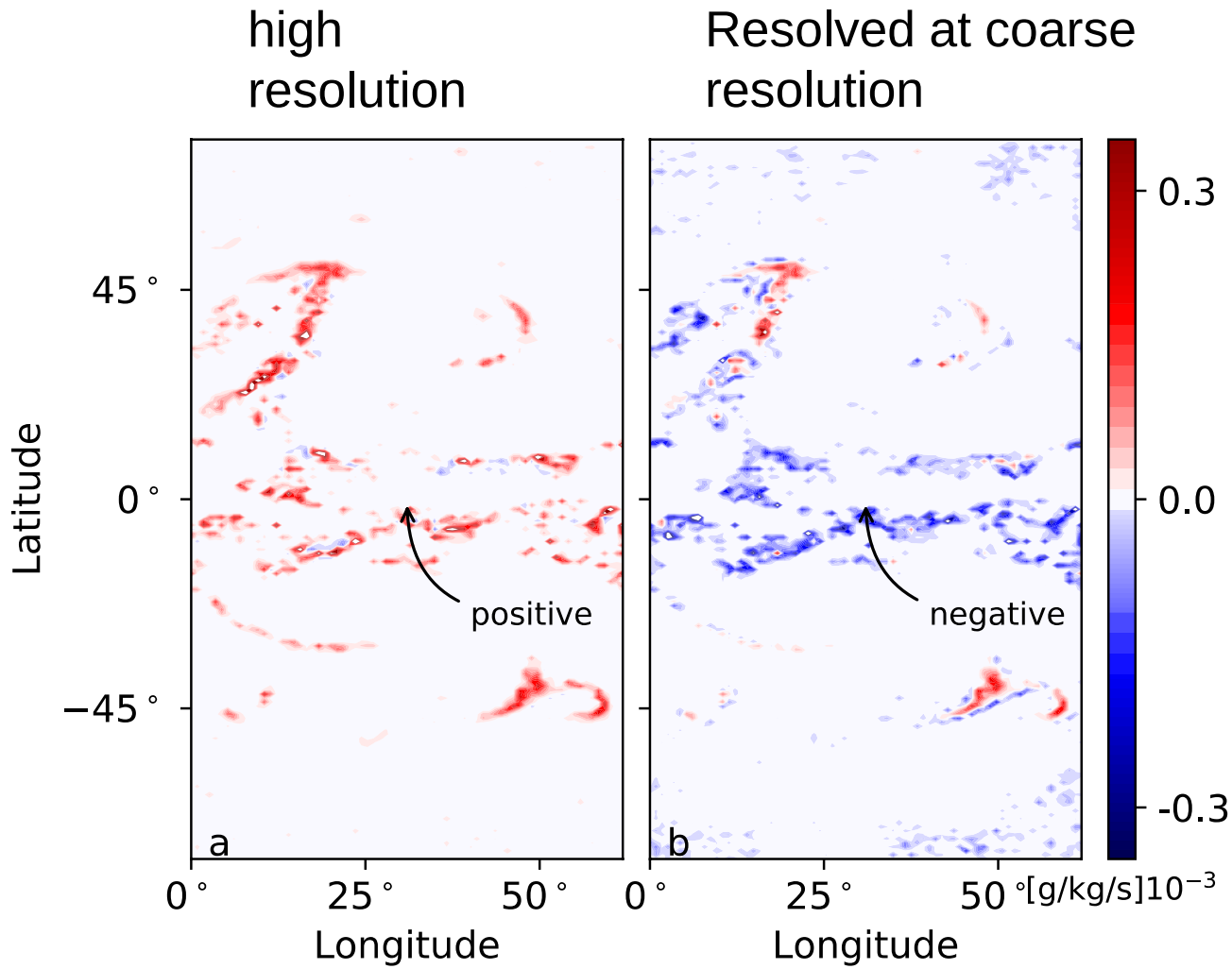
$$y = \left(\frac{\partial h_L}{\partial t}_{\text{sub}}, \frac{\partial q_T}{\partial t}_{\text{sub}}, \frac{\partial q_p}{\partial t}_{\text{sub}}, \text{diffusivity} \right)$$



Outline

- **Motivation**: Parameterizations of small scale processes are the main cause of uncertainty in climate change **temperature and precipitation** projections
- **Previous studies:** Recent studies on machine-learning parameterization are promising (but still many problems)
- **Results**:
  a. Learning from a high resolution model (SAM)
  b. Example: Sub-grid tendencies are very important
  c. Offline results, conservation of energy, non-negative precipitation
  d. Online results
  e. Is there an optimal length scale for parameterization?

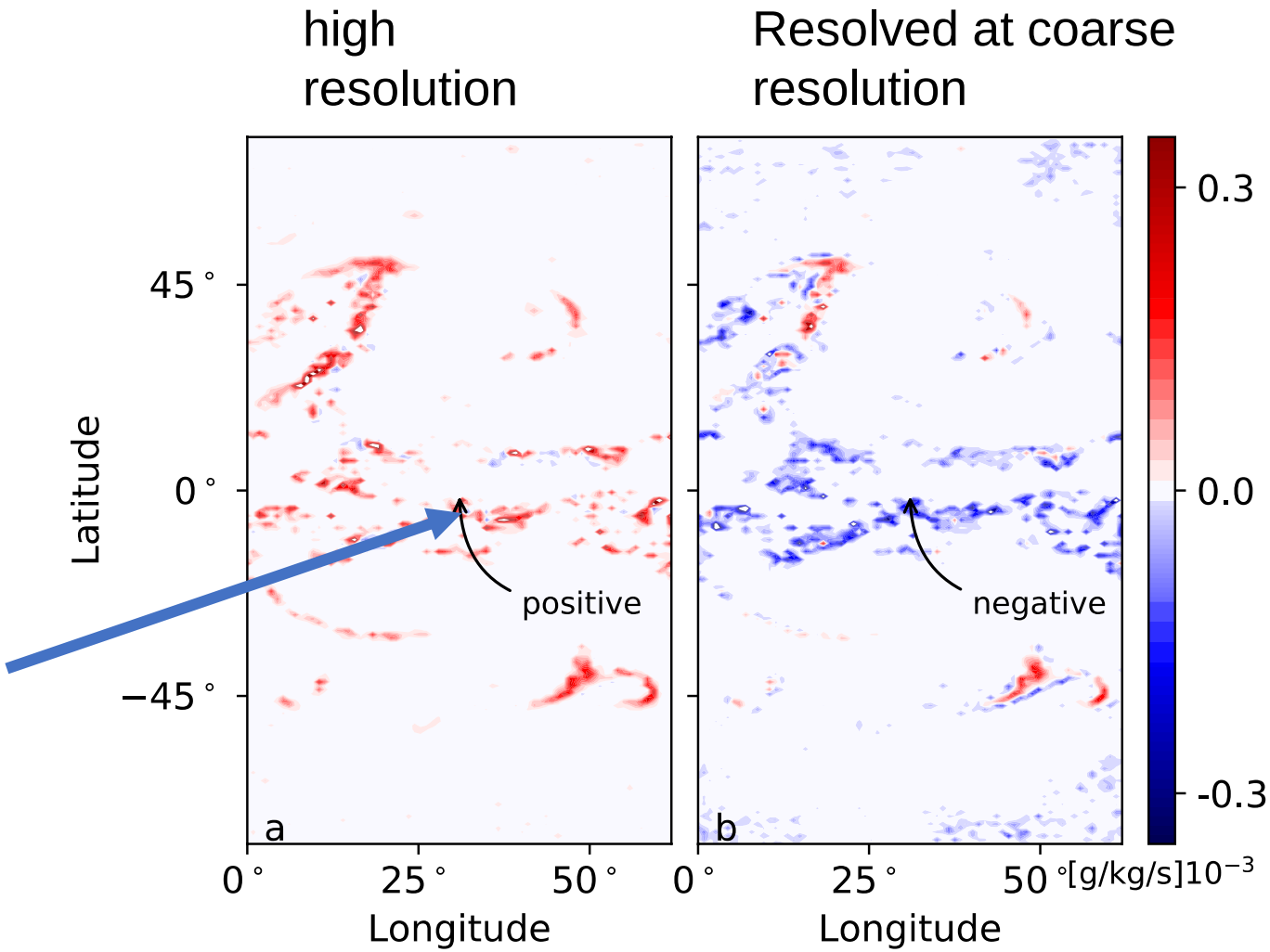
Tendency of precipitating water due to conversion of cloud water/ice and due to evaporation at 3km



Would look the same if subgrid not important

$$\text{subgrid tendency} = \text{coarse-grained} - \text{resolved}$$

Tendency of precipitating water due to conversion of cloud water/ice and due to evaporation at 3km

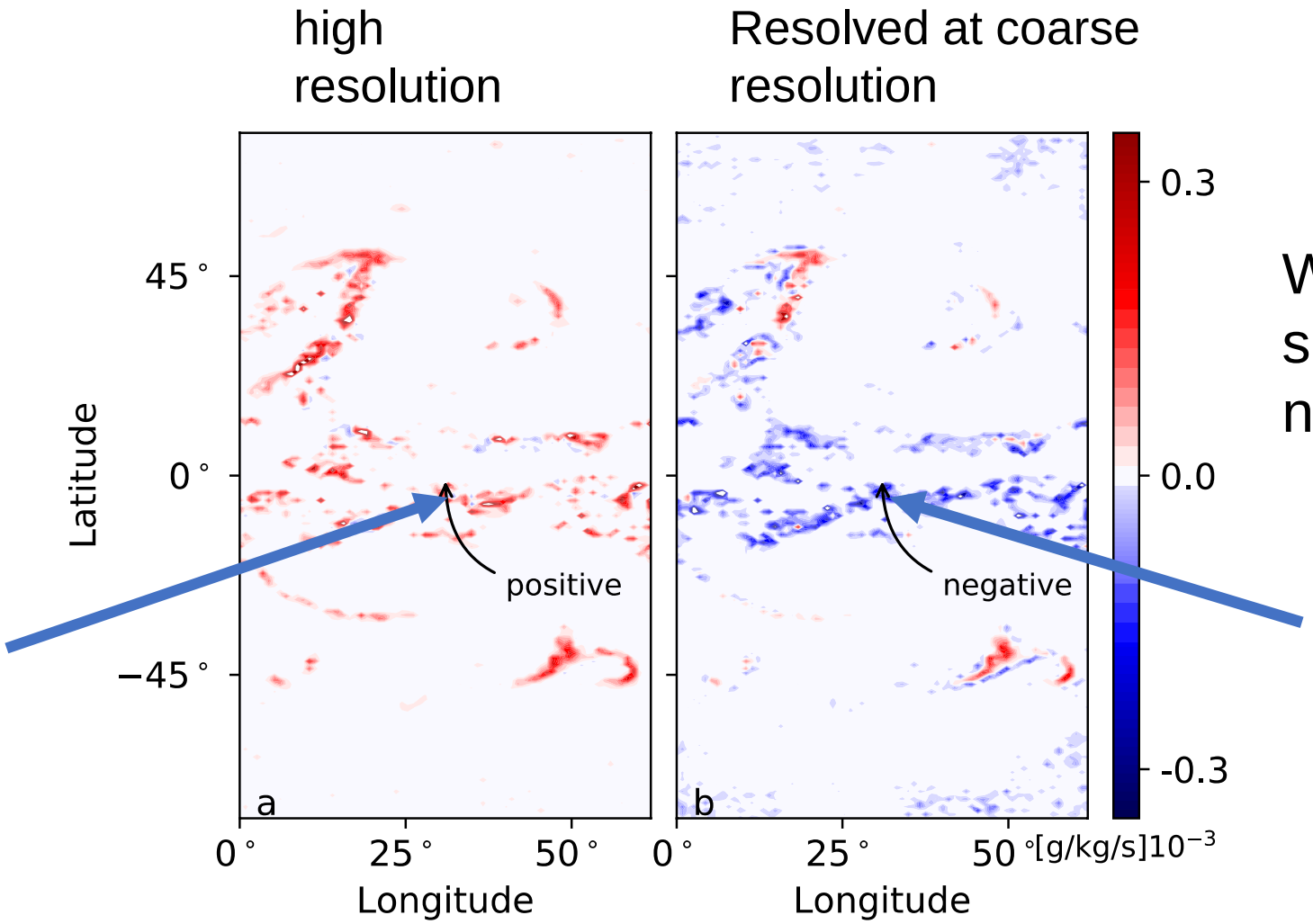


Positive – means autoconversion (droplets grow as they collect water from clouds)

Would look the same if subgrid not important

$$\text{subgrid tendency} = \text{coarse-grained} - \text{resolved}$$

Tendency of precipitating water due to conversion of cloud water/ice and due to evaporation at 3km



Positive – means autoconversion (droplets grow as they collect water from clouds)

Would look the same if subgrid not important

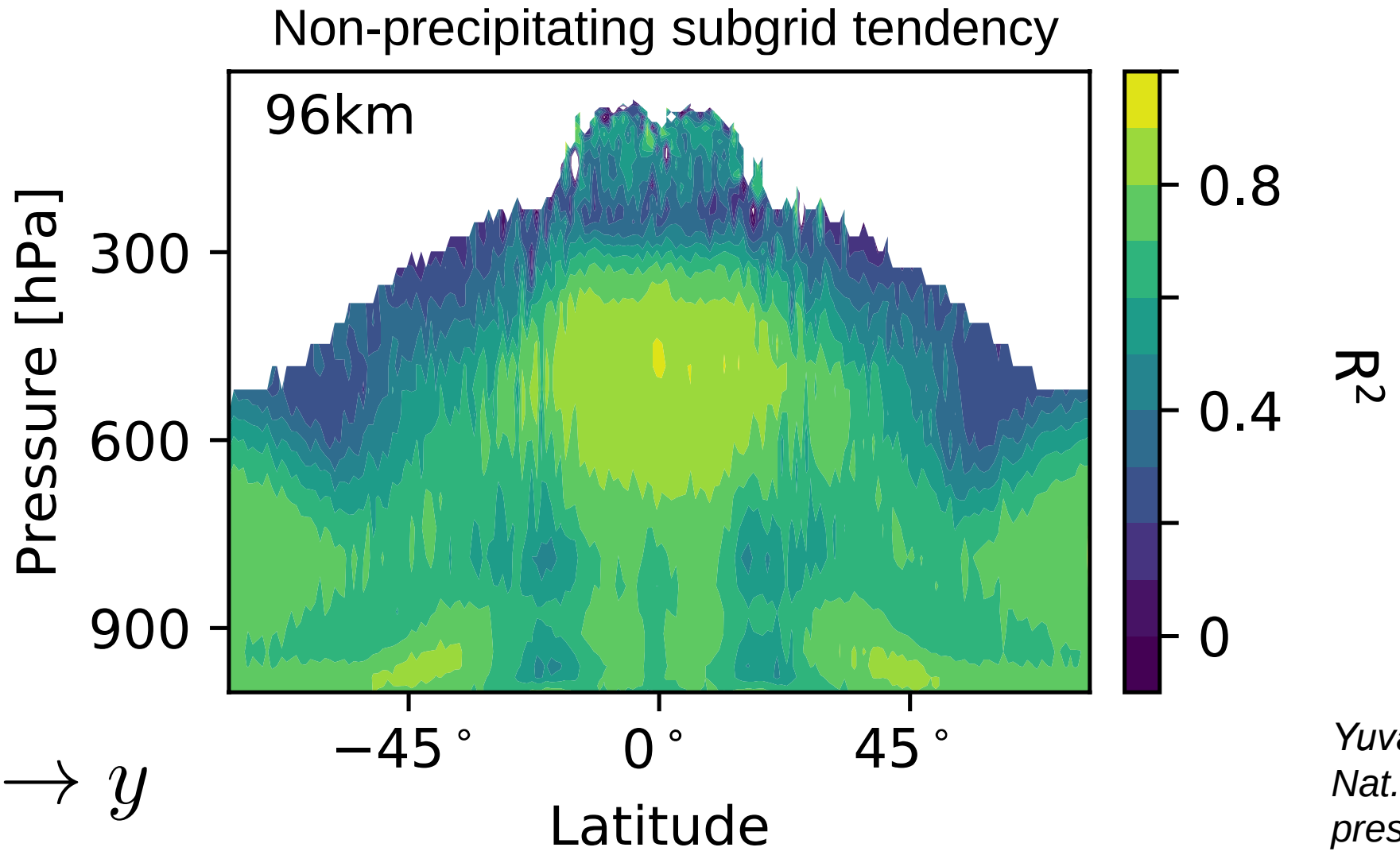
Negative – means Evaporation of precipitating water

$$\text{subgrid tendency} = \text{coarse-grained} - \text{resolved}$$

Outline

- **Motivation**: Parameterizations of small scale processes are the main cause of uncertainty in climate change **temperature and precipitation** projections
- **Previous studies:** Recent studies on machine-learning parameterization are promising (but still many problems)
- **Results**:
  a. Learning from a high resolution model (SAM)
  b. Example: Sub-grid tendencies are very important
  c. Offline results, conservation of energy, non-negative precipitation
  d. Online results
  e. Is there an optimal length scale for parameterization?

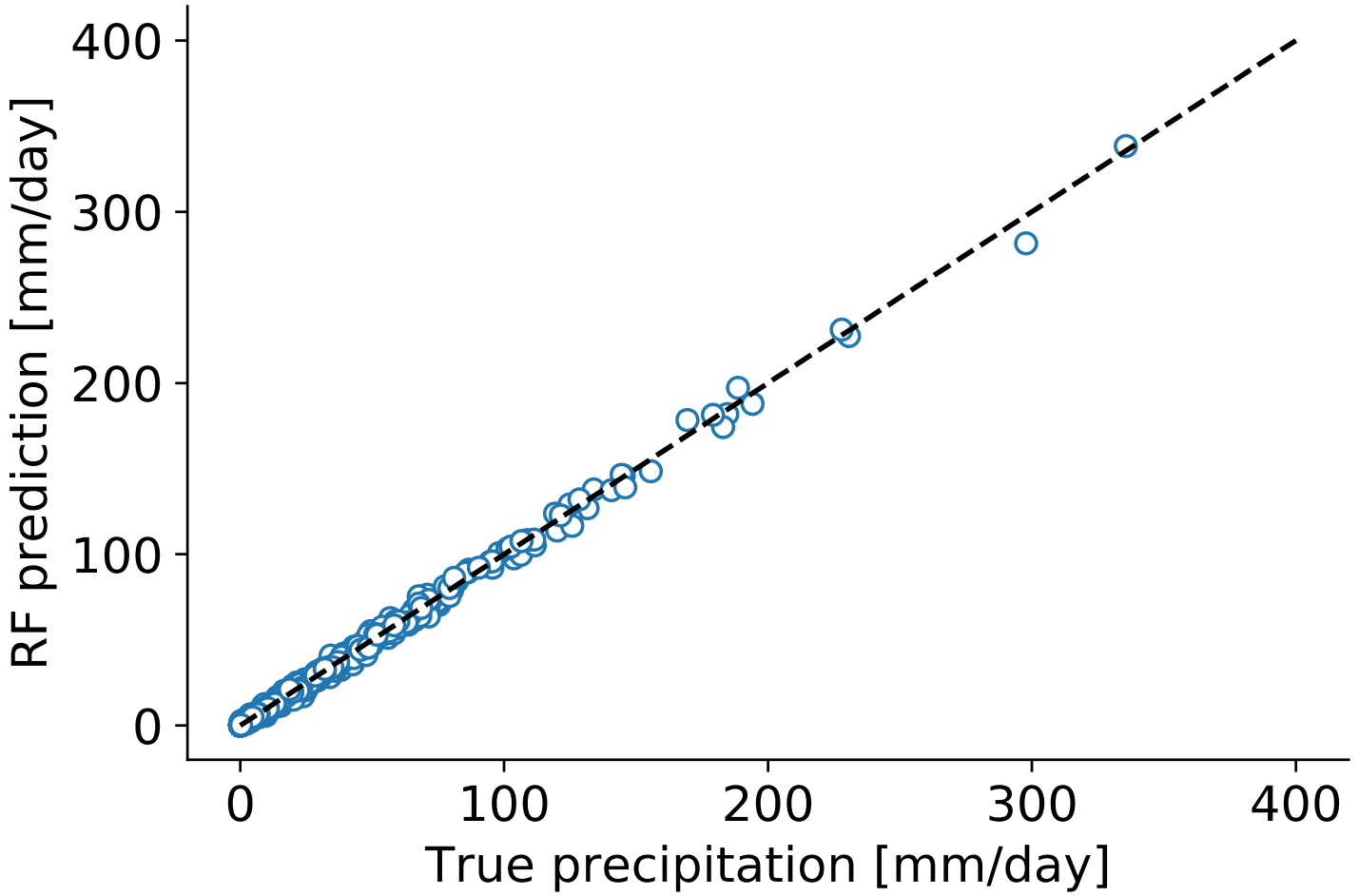
Random forest (96km horizontal grid) predicts reasonably well the subgrid tendencies in a test data set (coefficient of determination $R^2>0.7$)



$$RF(X) \rightarrow y$$

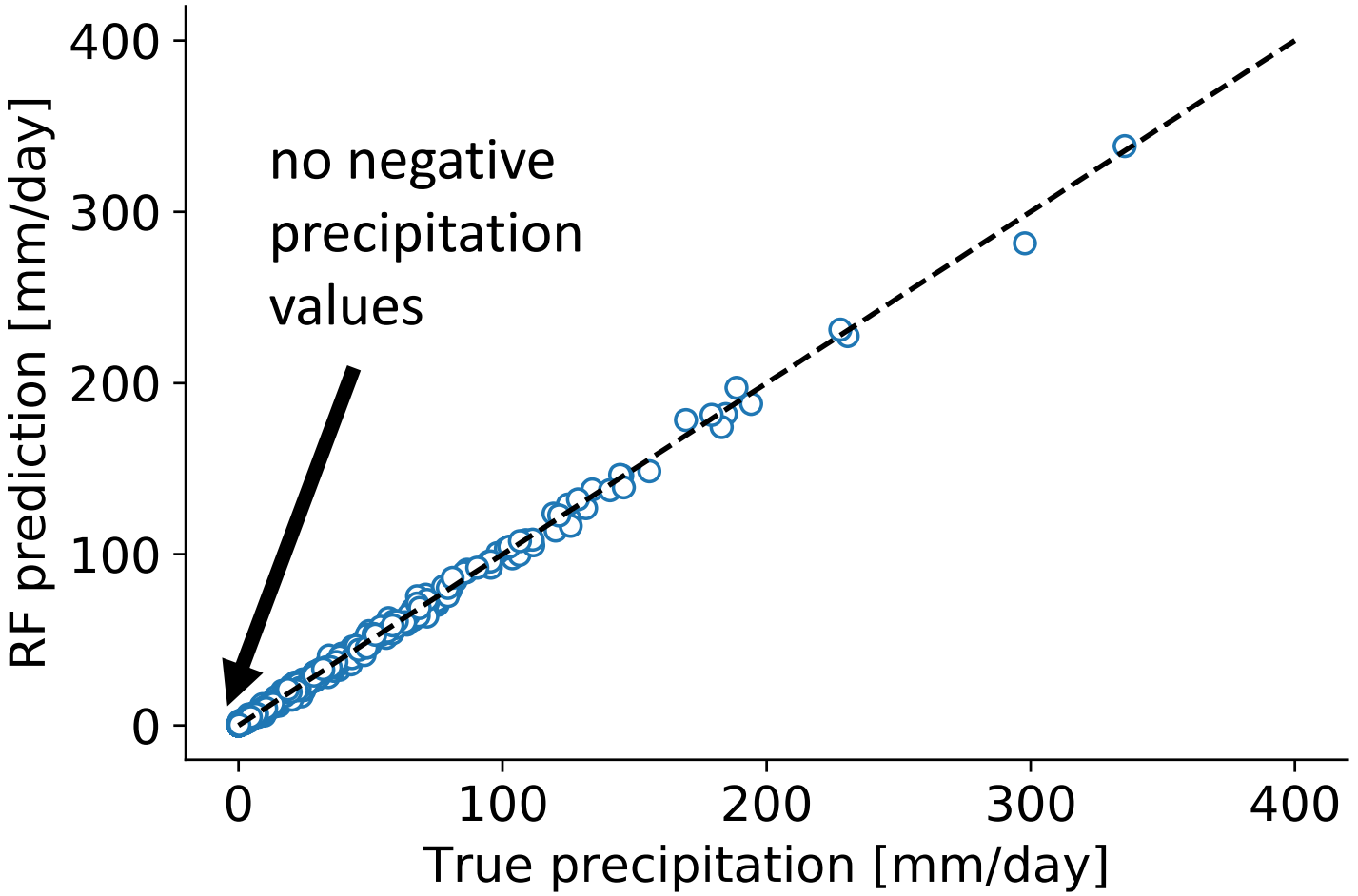
Yuval & O'Gorman, Nat. Commun. (in press)

Precipitation is a diagnostic variable (we do not predict it separately in the RF)



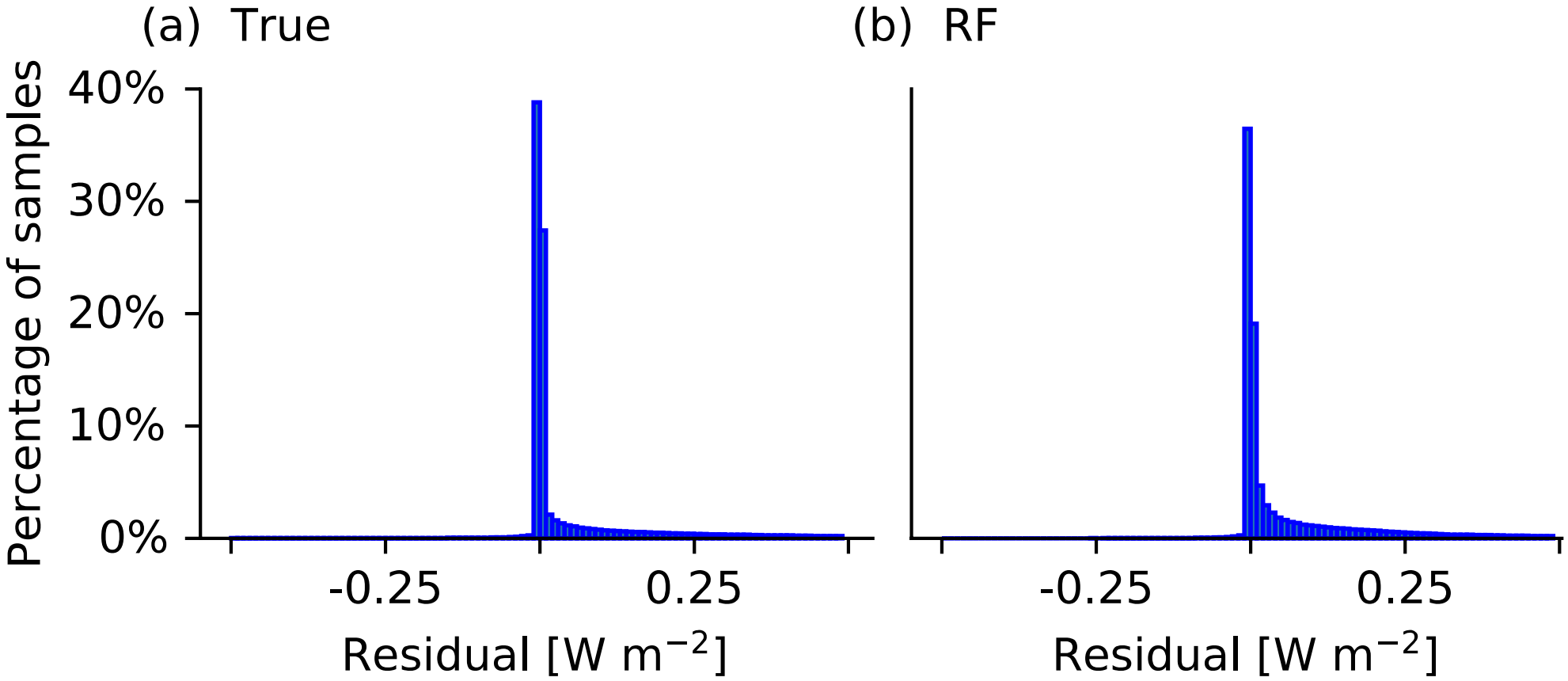
Yuval & O'Gorman, Nat. Commun. (in press)

Precipitation is a diagnostic variable (we do not predict it separately in the RF)



Yuval & O'Gorman, Nat. Commun. (in press)

Random forest (almost) conserves energy in the absence of external forcing



Compared to ~92W/m^2 in Brenowitz and Bretherton (2019) that used neural networks (Though see Buecler et al. 2019)

Yuval & O'Gorman, Nat. Commun. (in press)

Outline

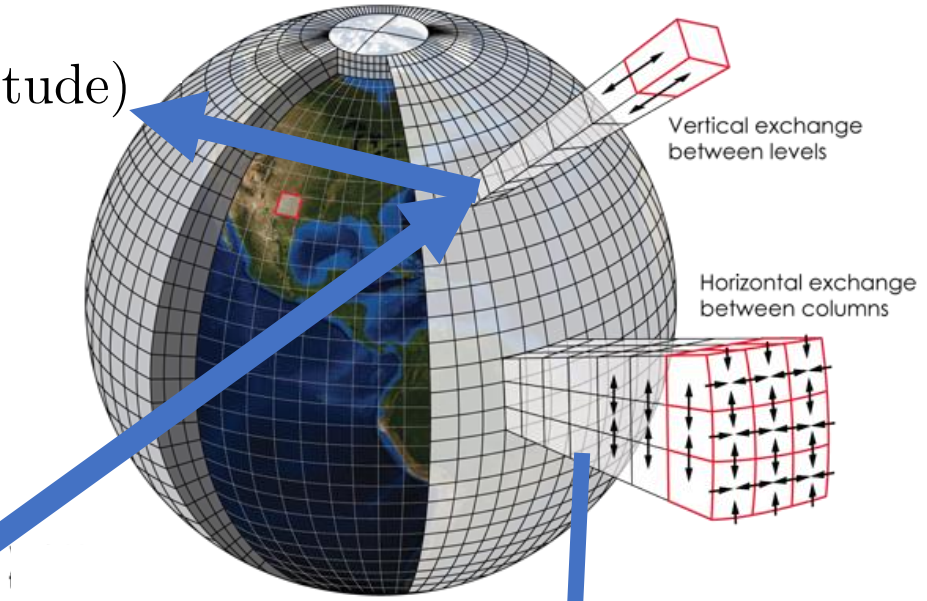
- **Motivation**: Parameterizations of small scale processes are the main cause of uncertainty in climate change **temperature and precipitation** projections
- **Previous studies:** Recent studies on machine-learning parameterization are promising (but still many problems)
- **Results**:
  a. Learning from a high resolution model (SAM)
  b. Example: Sub-grid tendencies are very important
  c. Offline results, conservation of energy, non-negative precipitation
  d. Online results
  e. Is there an optimal length scale for parameterization?

Coupling the random forest to SAM

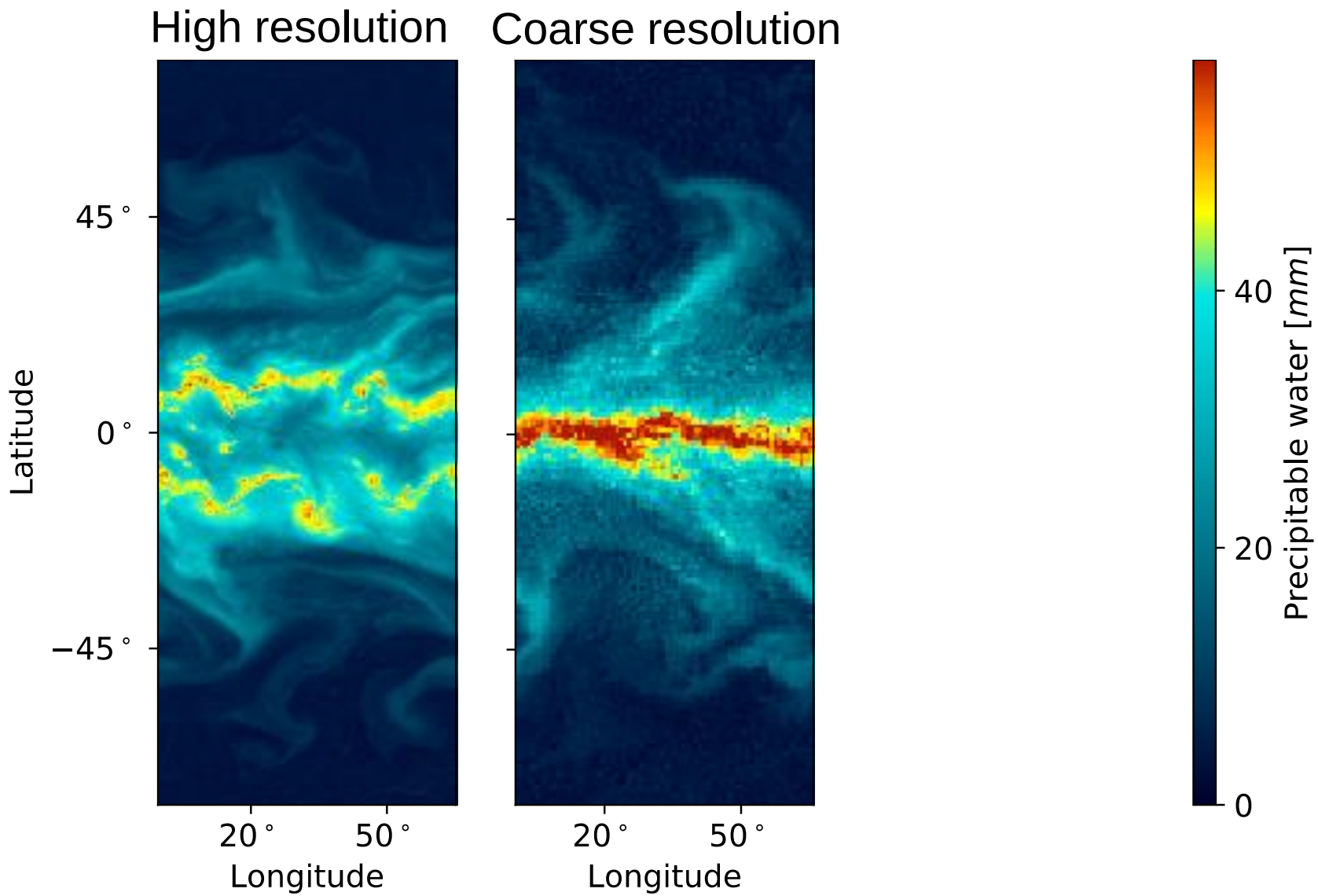
$$X = (T(z), q_T(z), q_p(z), u(z), v(z), \text{latitude})$$

$$RF(X) \rightarrow y$$

$$y = \left(\frac{\partial h_L}{\partial t}_{\text{sub}}, \frac{\partial q_T}{\partial t}_{\text{sub}}, \frac{\partial q_p}{\partial t}_{\text{sub}}, \text{diffusivity} \right)$$

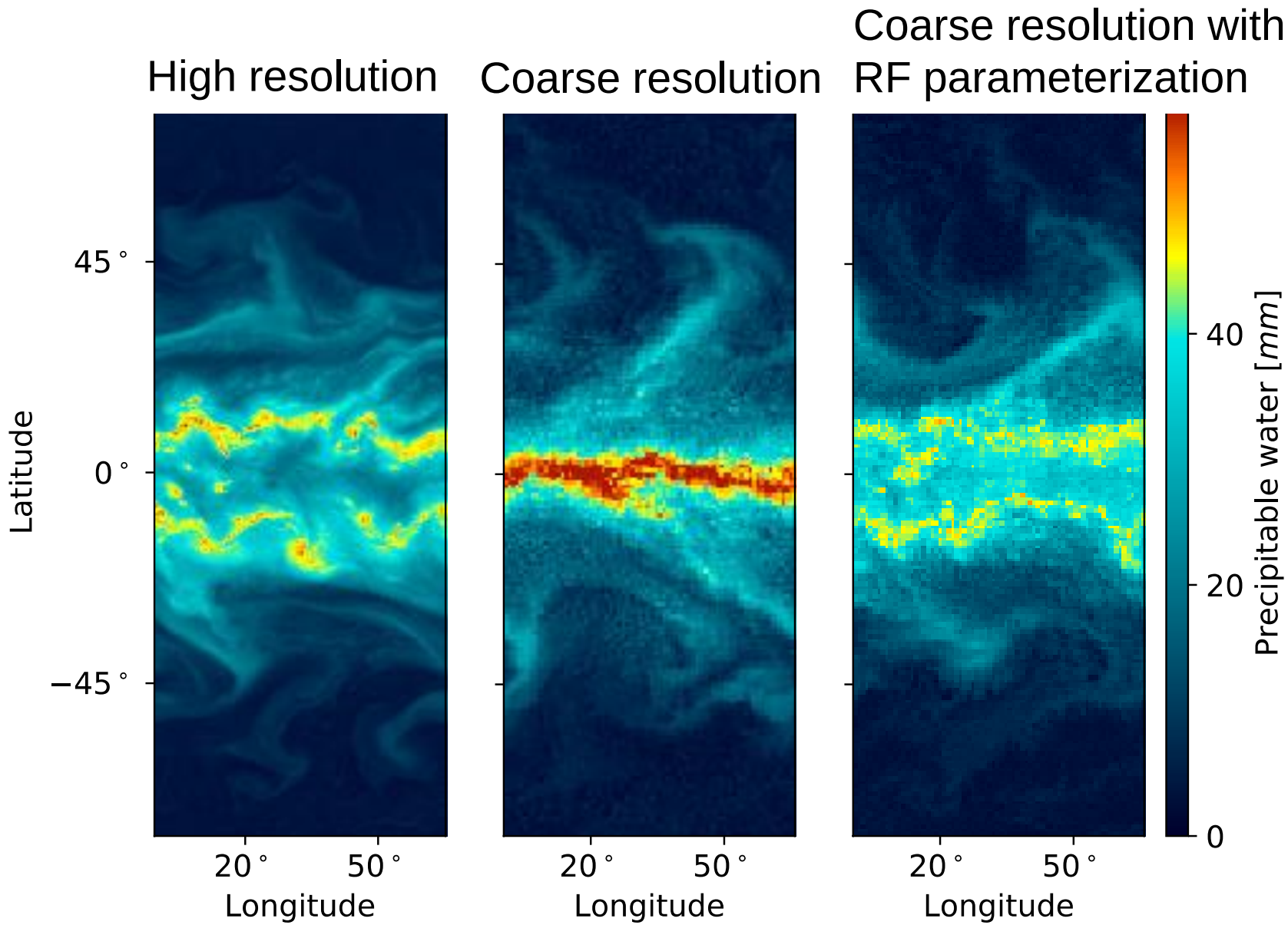


Random-forest parameterization brings the low-resolution model simulation much closer to the high-resolution simulation



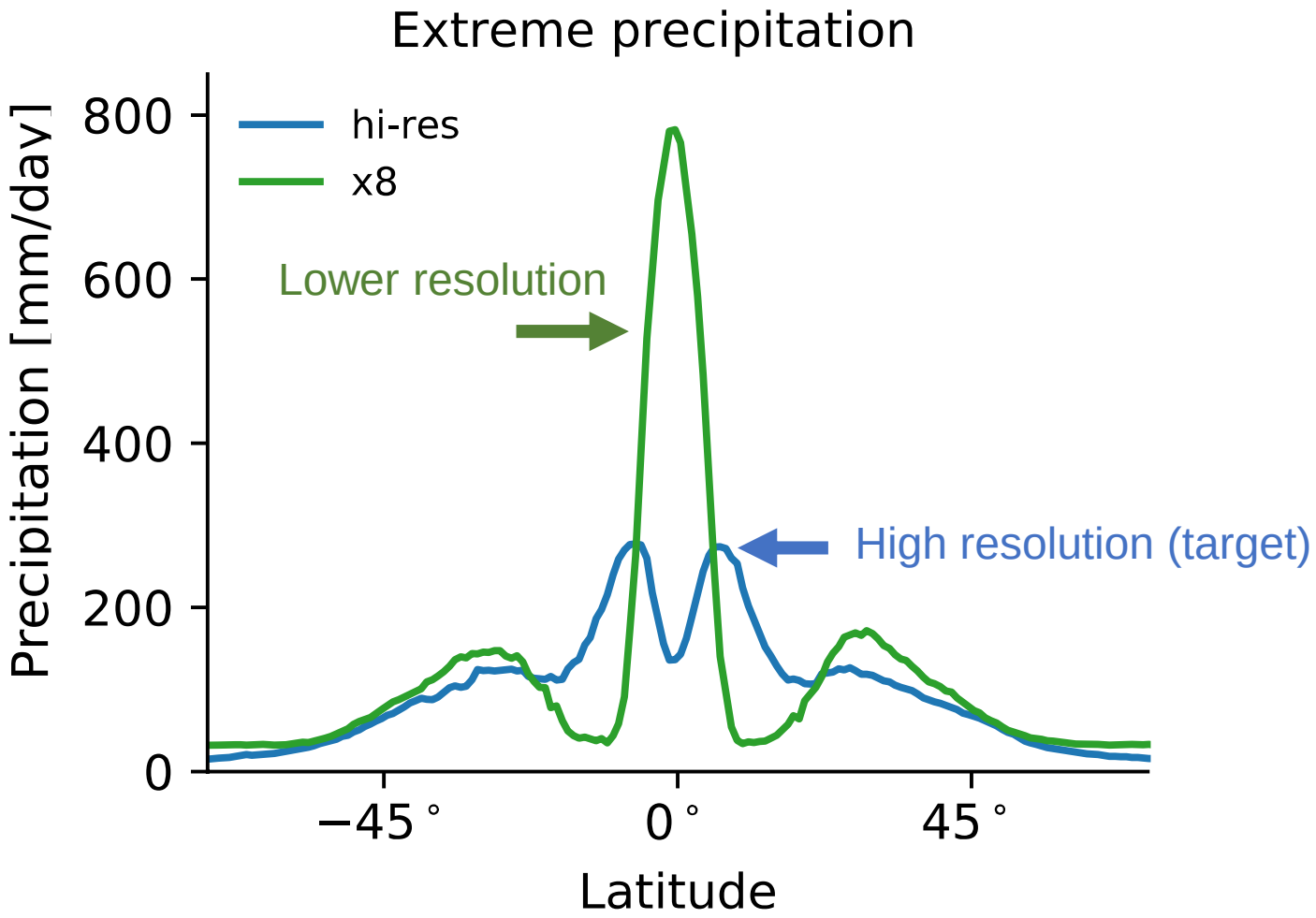
Yuval & O'Gorman, Nat. Commun. (in press)

Random-forest parameterization brings the low-resolution model simulation much closer to the high-resolution simulation



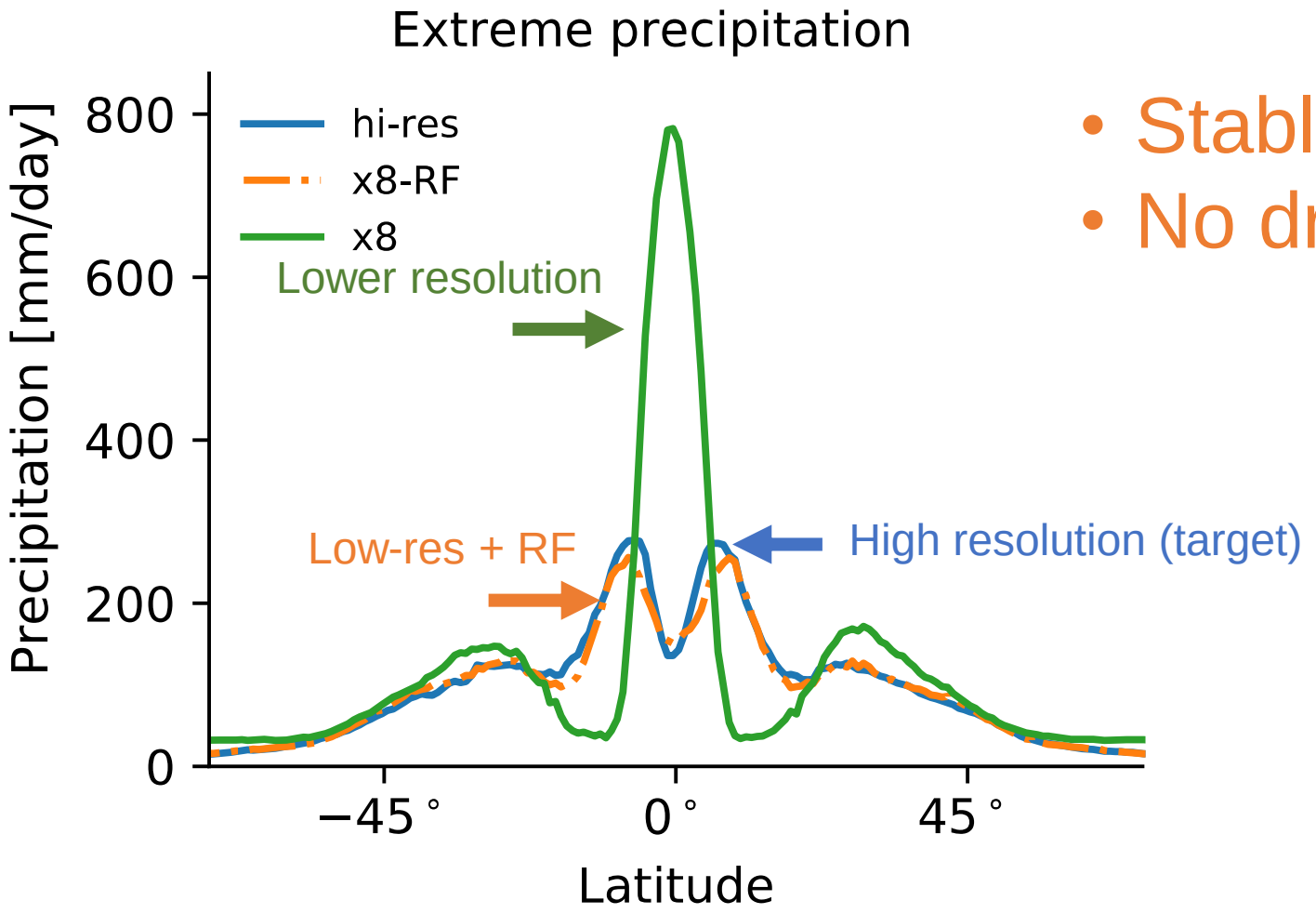
Yuval & O'Gorman, Nat. Commun. (in press)

Random forest parameterization leads to remarkably accurate simulation of **precipitation extremes**



Yuval & O'Gorman, Nat. Commun. (in press)

Random forest parameterization leads to remarkably accurate simulation of **precipitation extremes**



- Stable simulation
- No drift

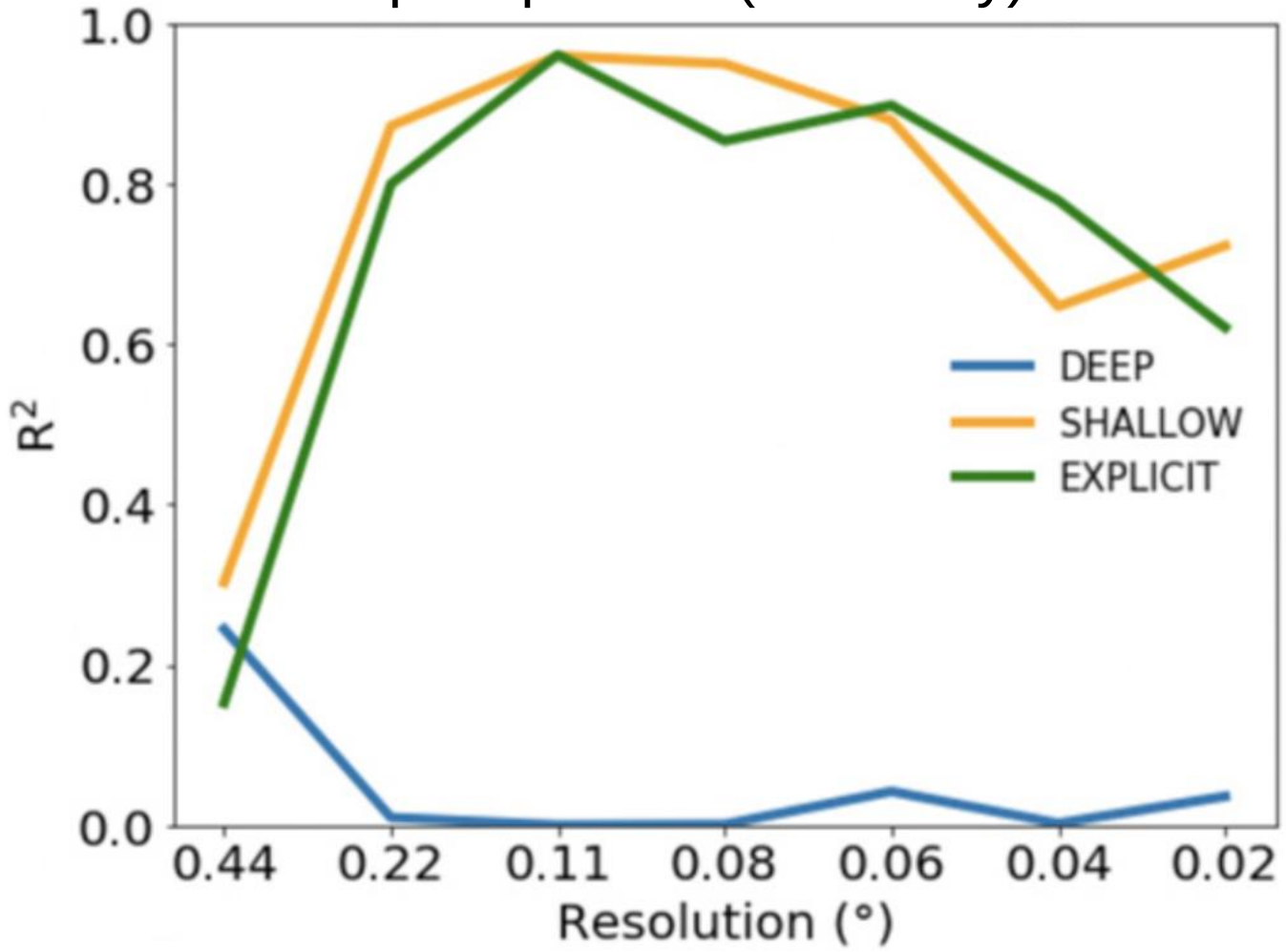
Yuval & O'Gorman, Nat. Commun. (in press)

Outline

- **Motivation**: Parameterizations of small scale processes are the main cause of uncertainty in climate change **temperature and precipitation** projections
- **Previous studies:** Recent studies on machine-learning parameterization are promising (but still many problems)
- **Results**:
  a. Learning from a high resolution model (SAM)
  b. Example: Sub-grid tendencies are very important
  c. Offline results, conservation of energy, non-negative precipitation
  d. Online results
  e. Is there an optimal length scale for parameterization?

Is there an optimal length scale for parameterization?

Test score for diurnal cycle of summer precipitation (Germany)

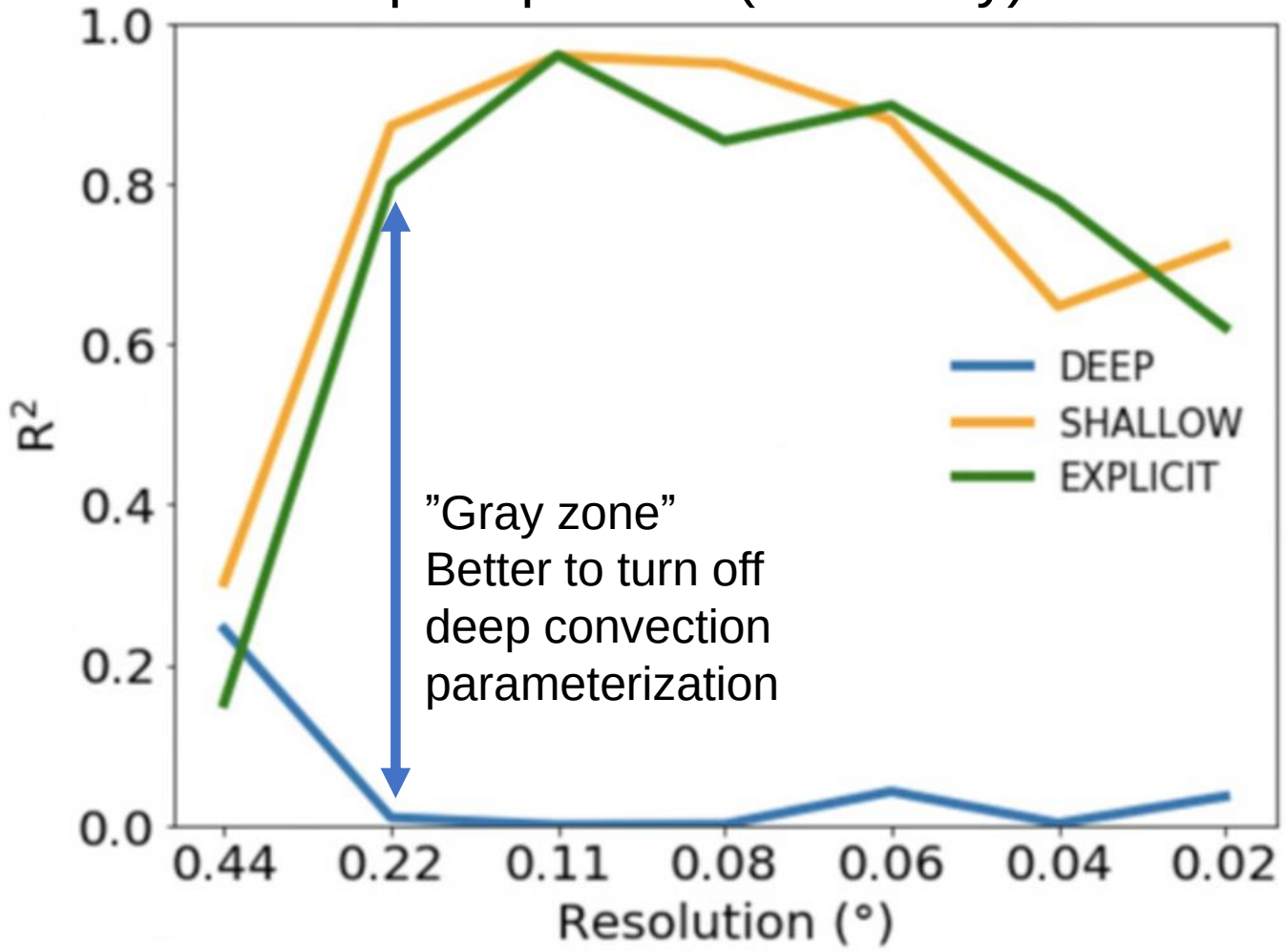


Conventional parameterization (non scale aware)

Vergara-Temprado et al. (2020)

Is there an optimal length scale for parameterization?

Test score for diurnal cycle of summer precipitation (Germany)

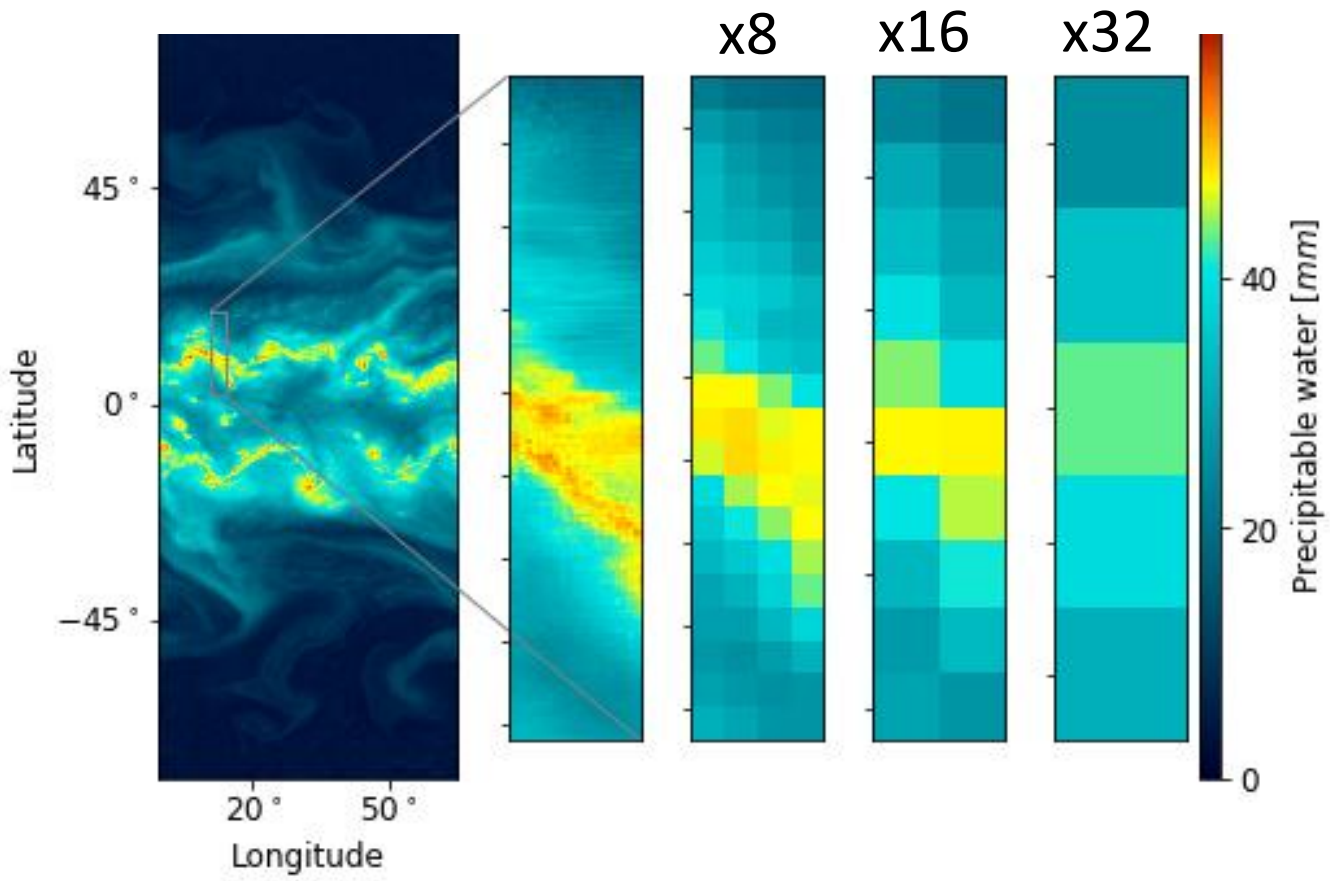


Conventional parameterization (non scale aware)

Vergara-Temprado et al. (2020)

Is there an optimal length scale for parameterization?

When increasing the grid spacing:

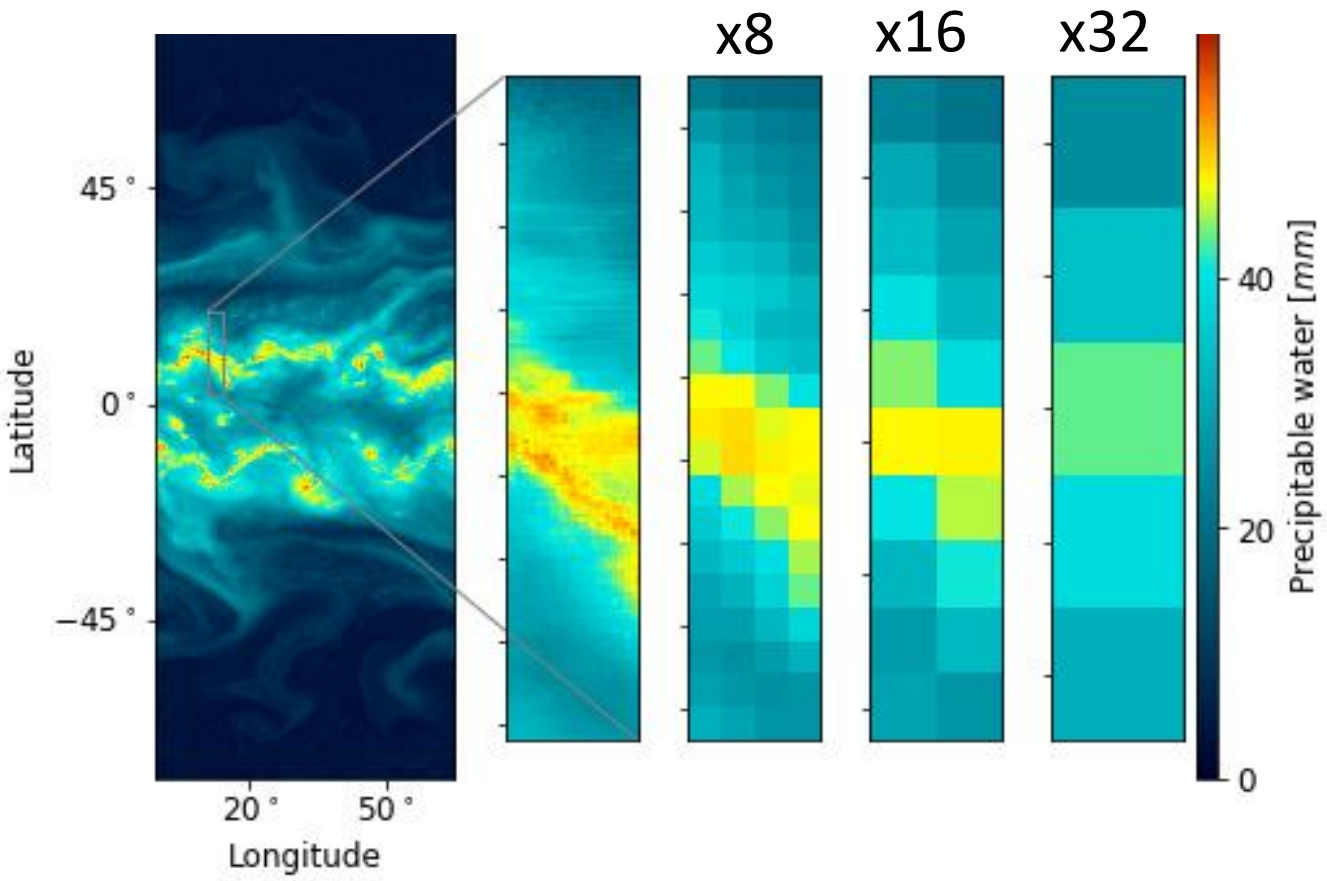


Yuval & O'Gorman, Nat. Commun. (in press)

Is there an optimal length scale for parameterization?

When increasing the grid spacing:

- Averaging over more cloud elements – subgrid tendencies more predictable
- But more of the dynamics becomes subgrid

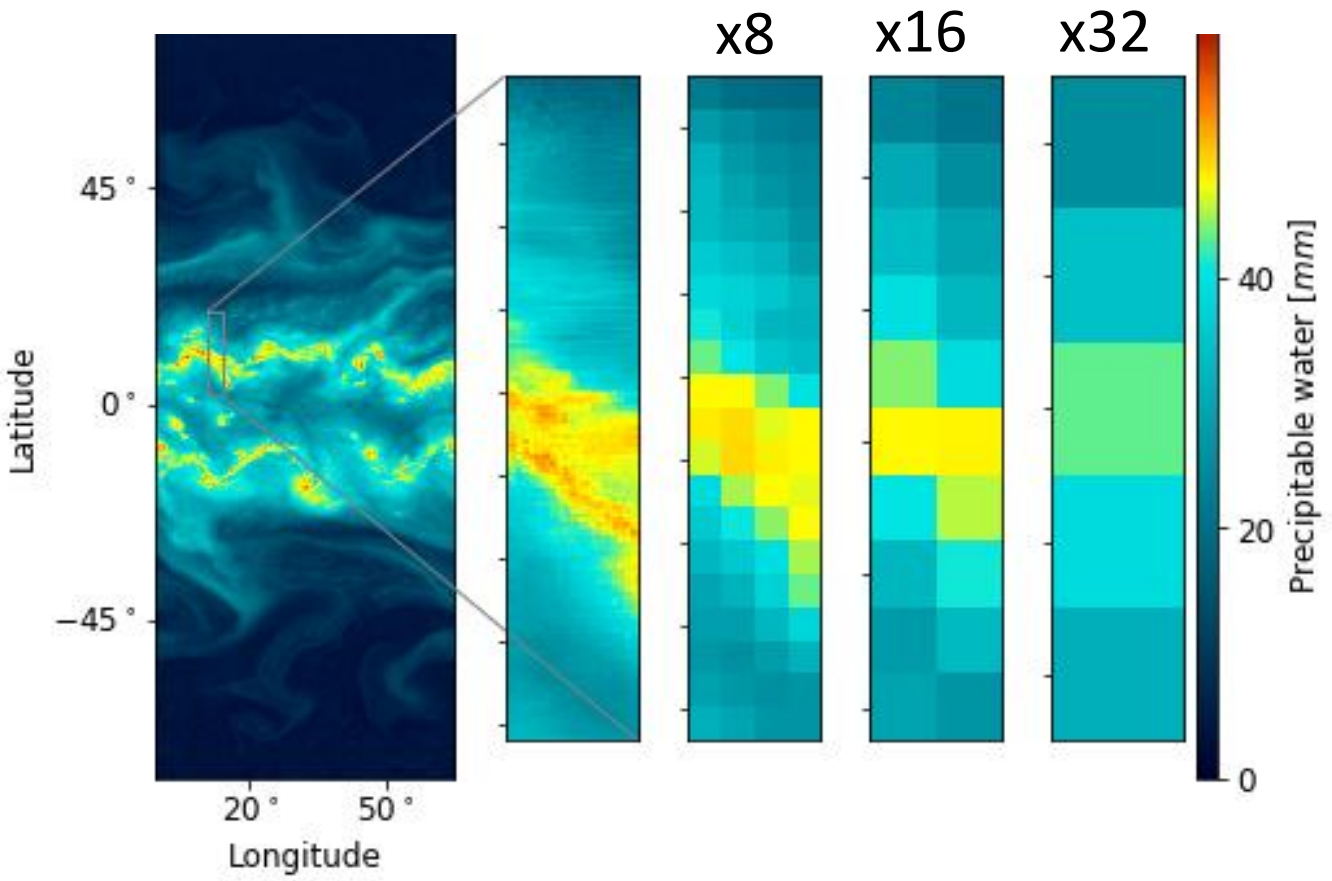


Yuval & O'Gorman, Nat. Commun. (in press)

Is there an optimal length scale for parameterization?

When increasing the grid spacing:

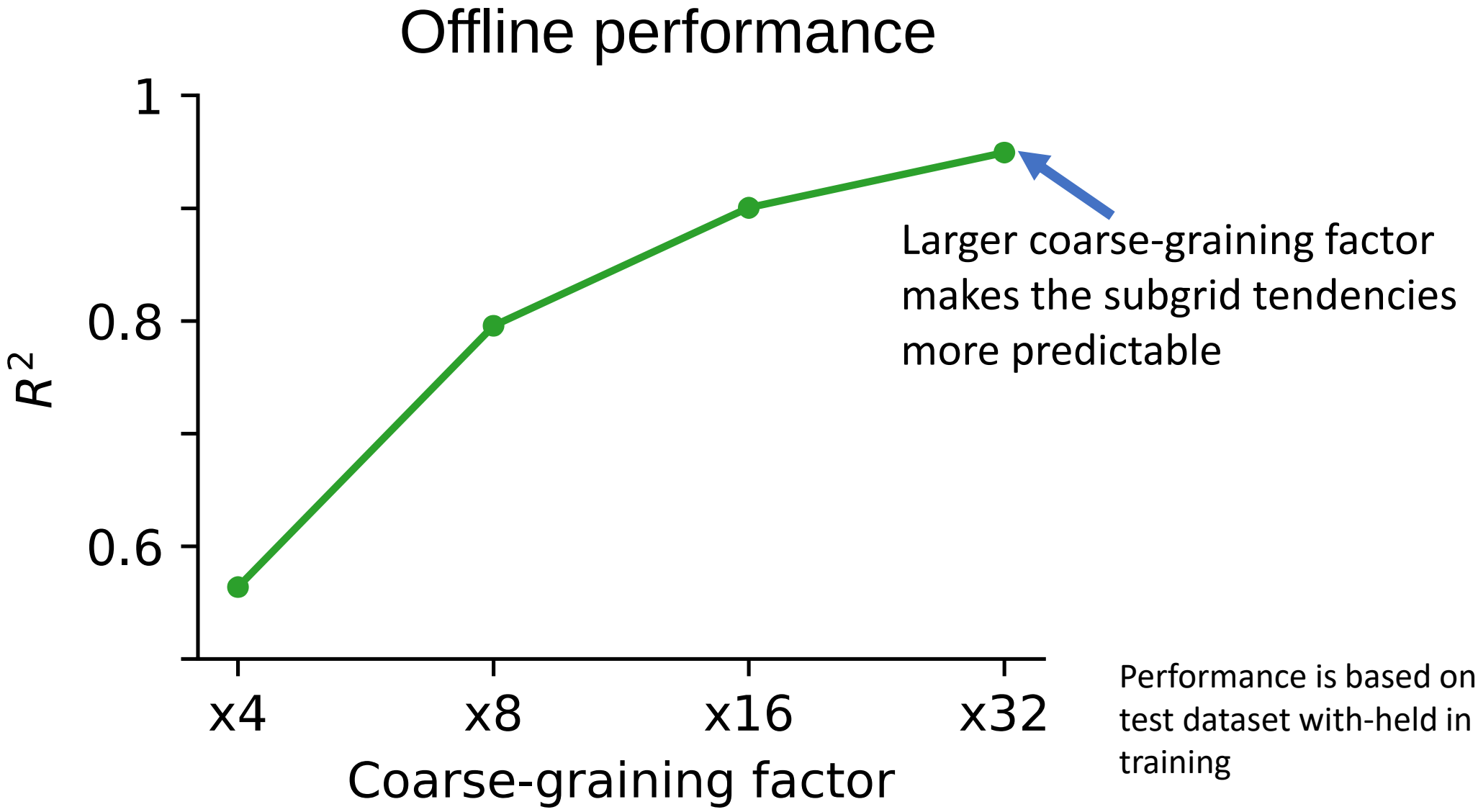
- Averaging over more cloud elements – subgrid tendencies more predictable
- But more of the dynamics becomes subgrid



We train a different random-forest parameterization for different grid-spacing

Yuval & O'Gorman, Nat. Commun. (in press)

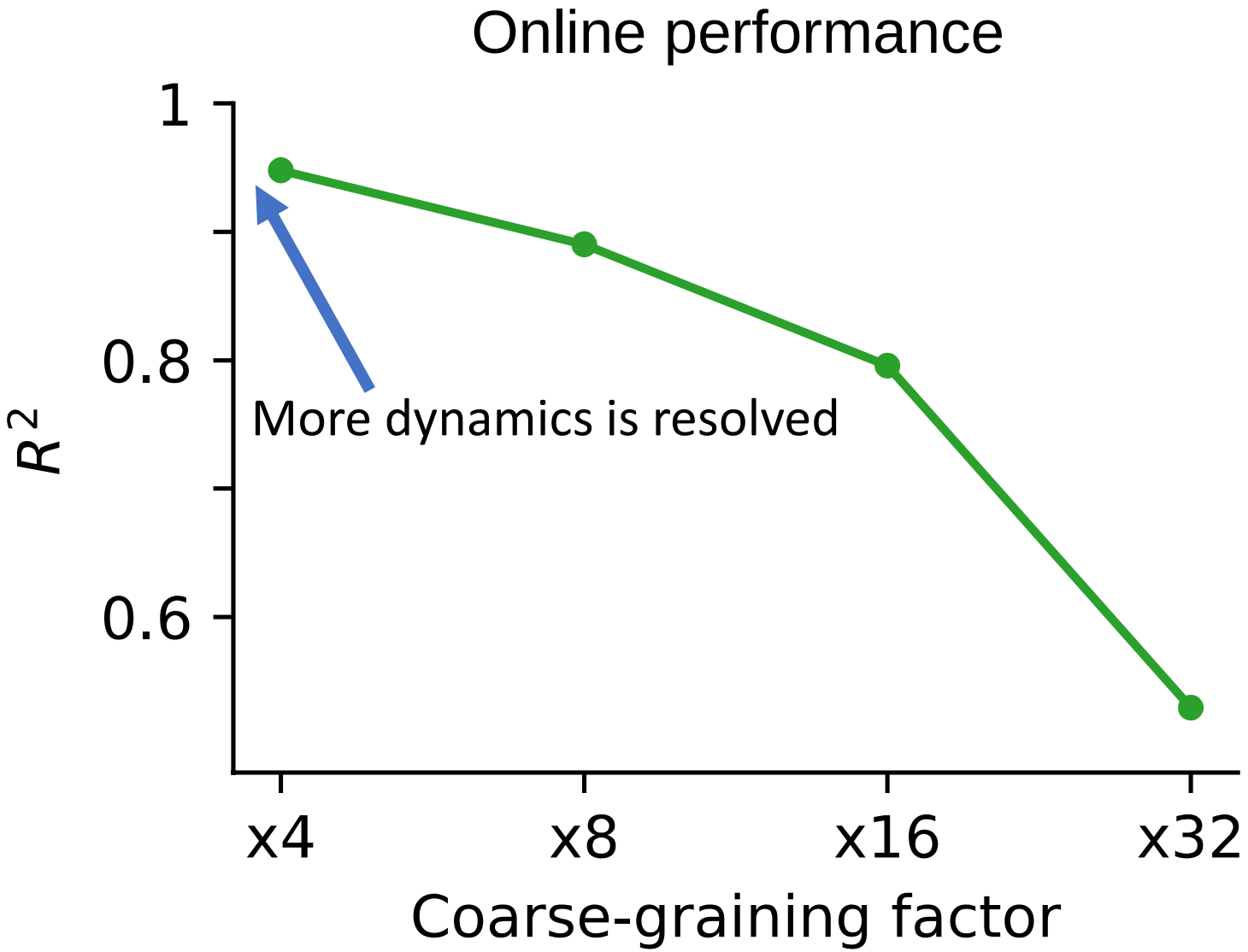
Random-forest parameterization of subgrid tendencies becomes more accurate as grid spacing becomes coarser



Yuval & O'Gorman, Nat. Commun. (in press)

Performance is based on test dataset with-held in training

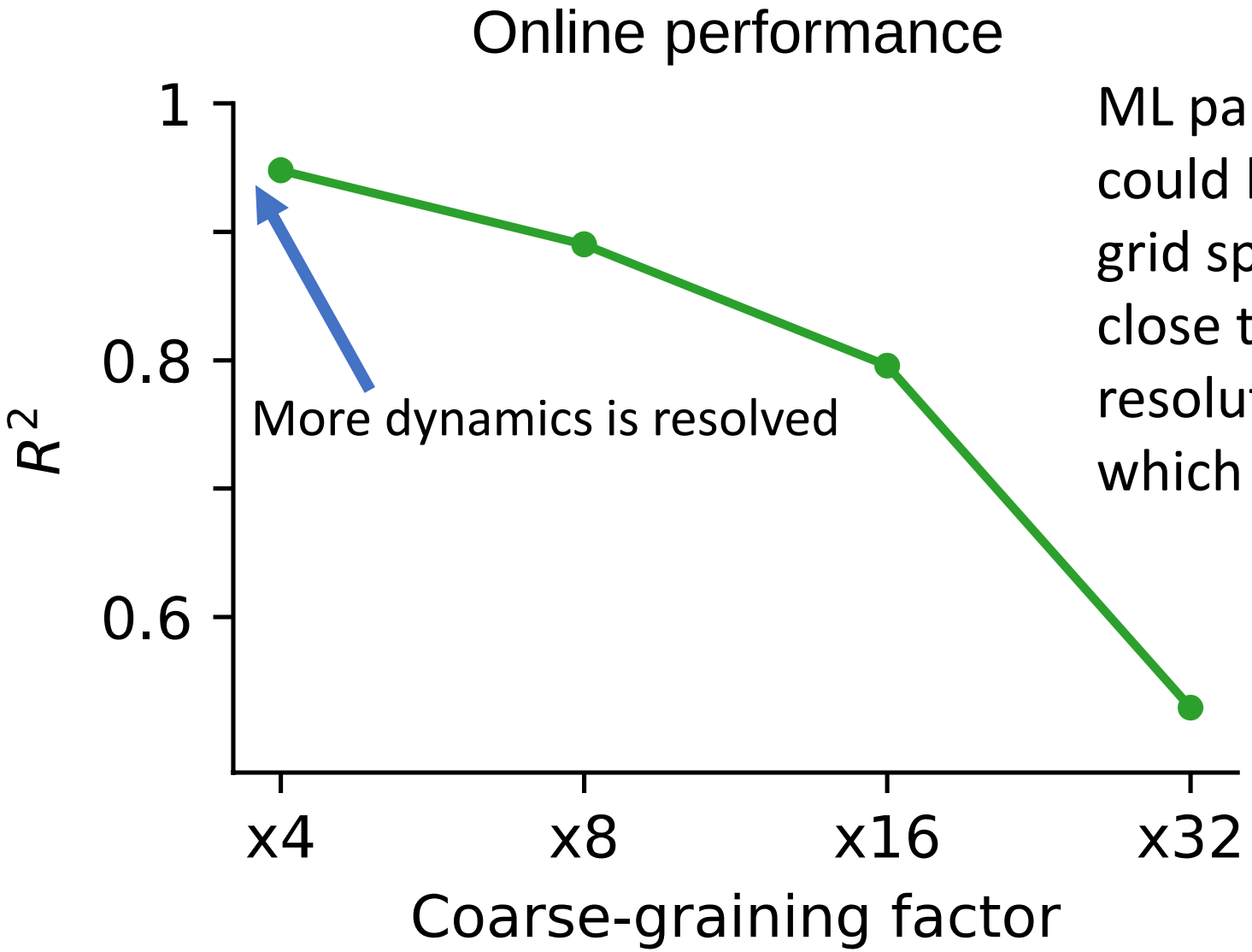
But accuracy of simulations improves with decreasing grid spacing



Yuval & O'Gorman, Nat. Commun. (in press)

Performance is measured by the mean precipitation distribution as compared to high-resolution simulation

But accuracy of simulations improves with decreasing grid spacing

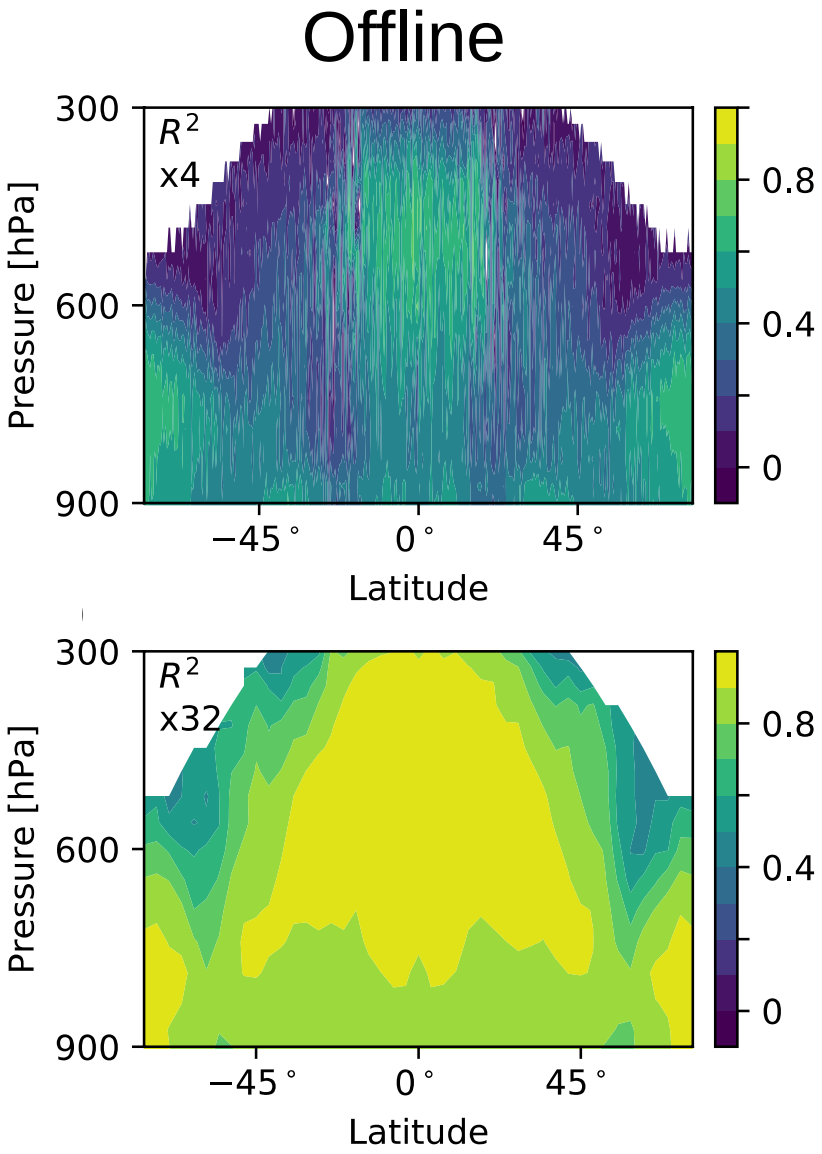


ML parameterizations could be useful for grid spacings that are quite close to that of the high-resolution model from which they are learned

Performance is measured by the mean precipitation distribution as compared to high-resolution simulation

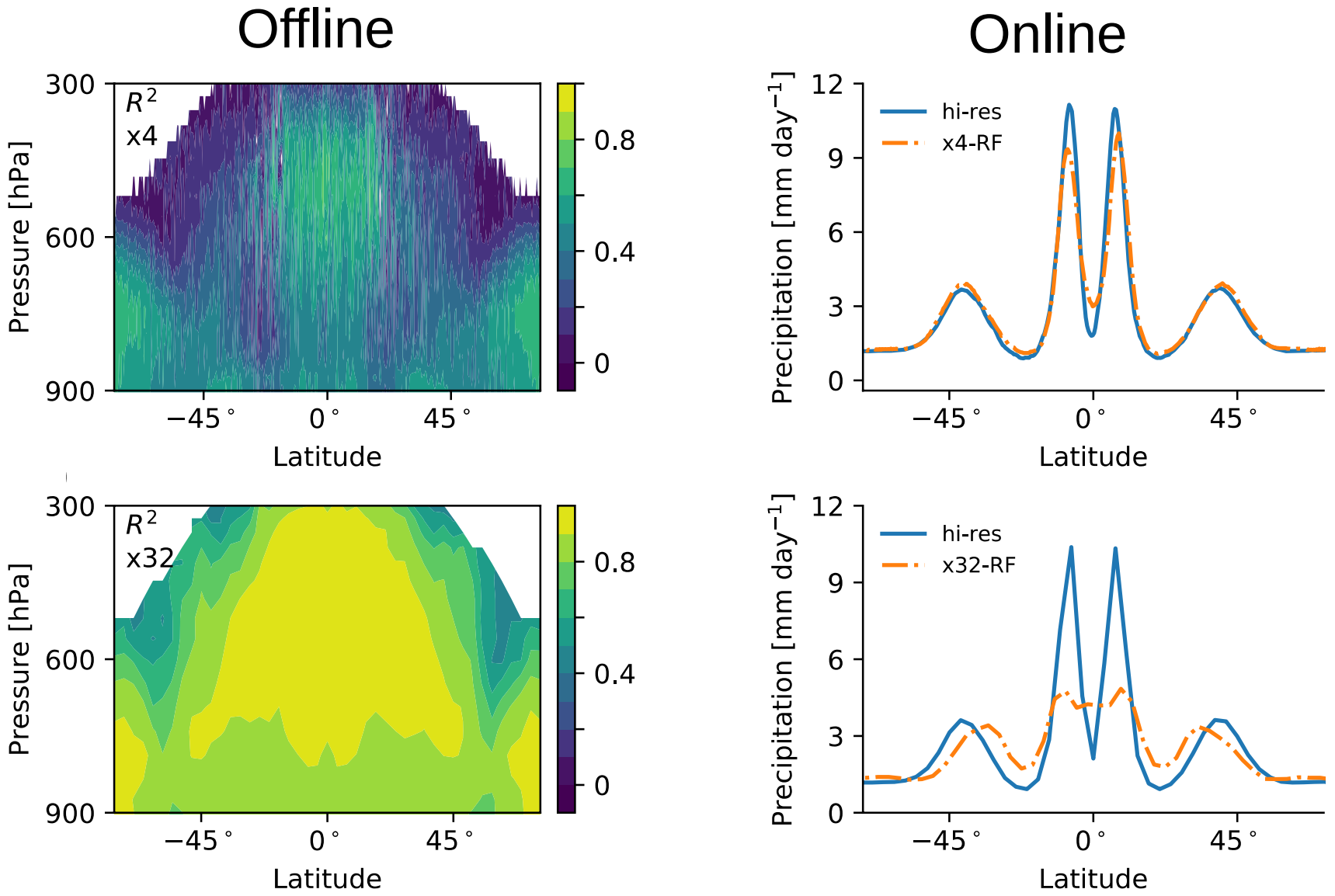
Yuval & O'Gorman, Nat. Commun. (in press)

Why the discrepancy between offline and online performance?



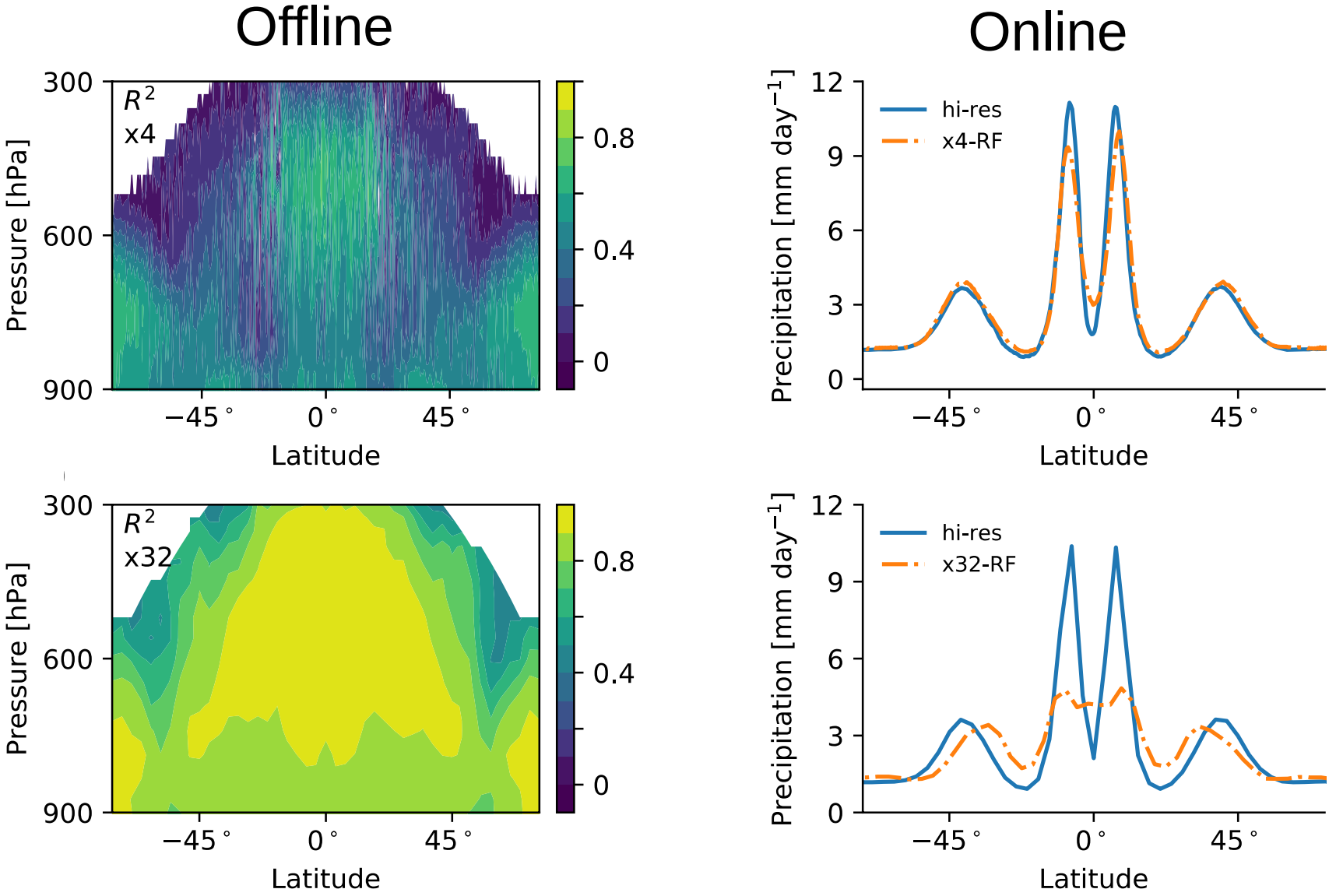
Yuval & O'Gorman, Nat. Commun. (in press)

Why the discrepancy between offline and online performance?



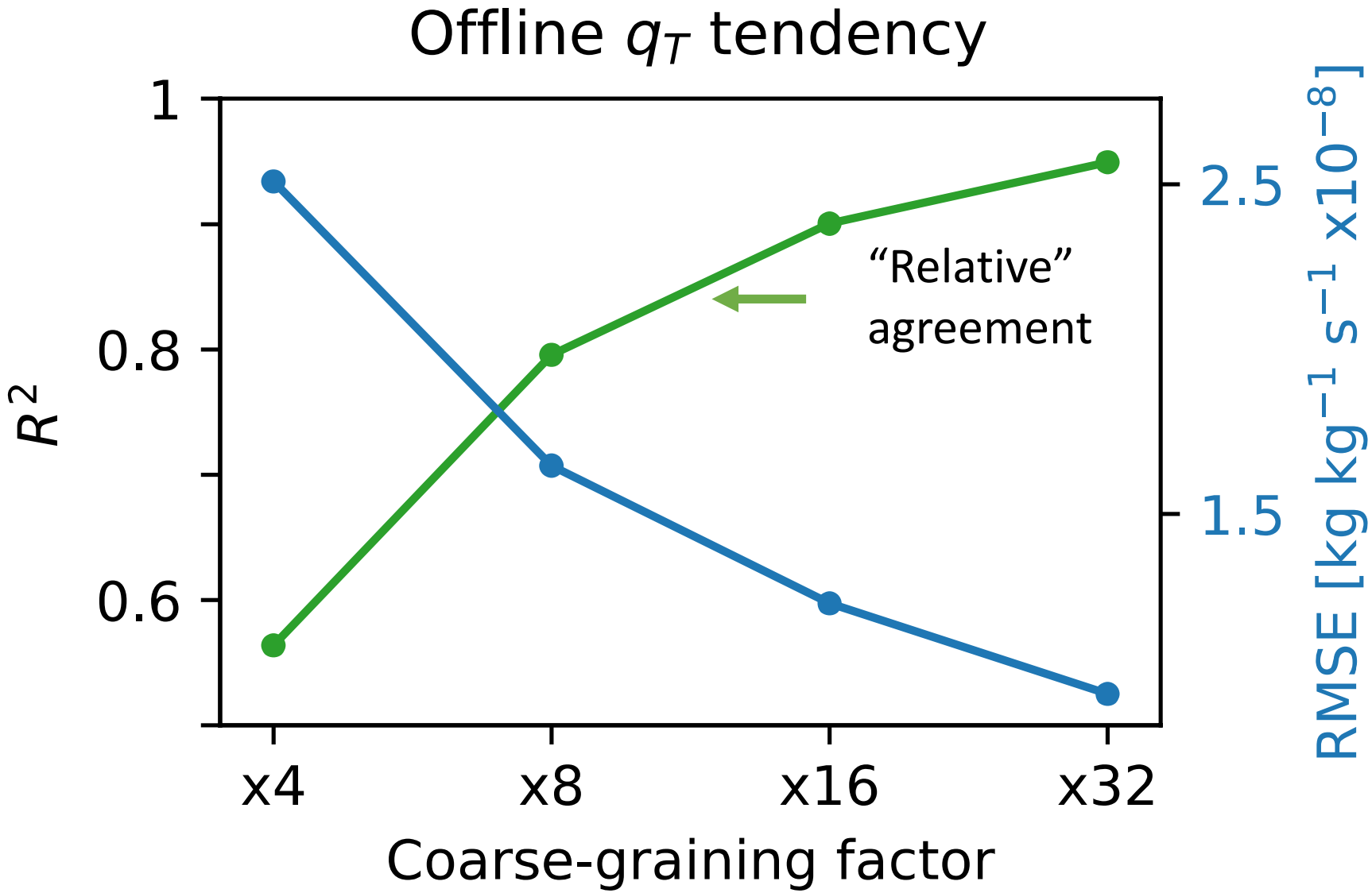
Yuval & O'Gorman, Nat. Commun. (in press)

Why the discrepancy between offline and online performance?

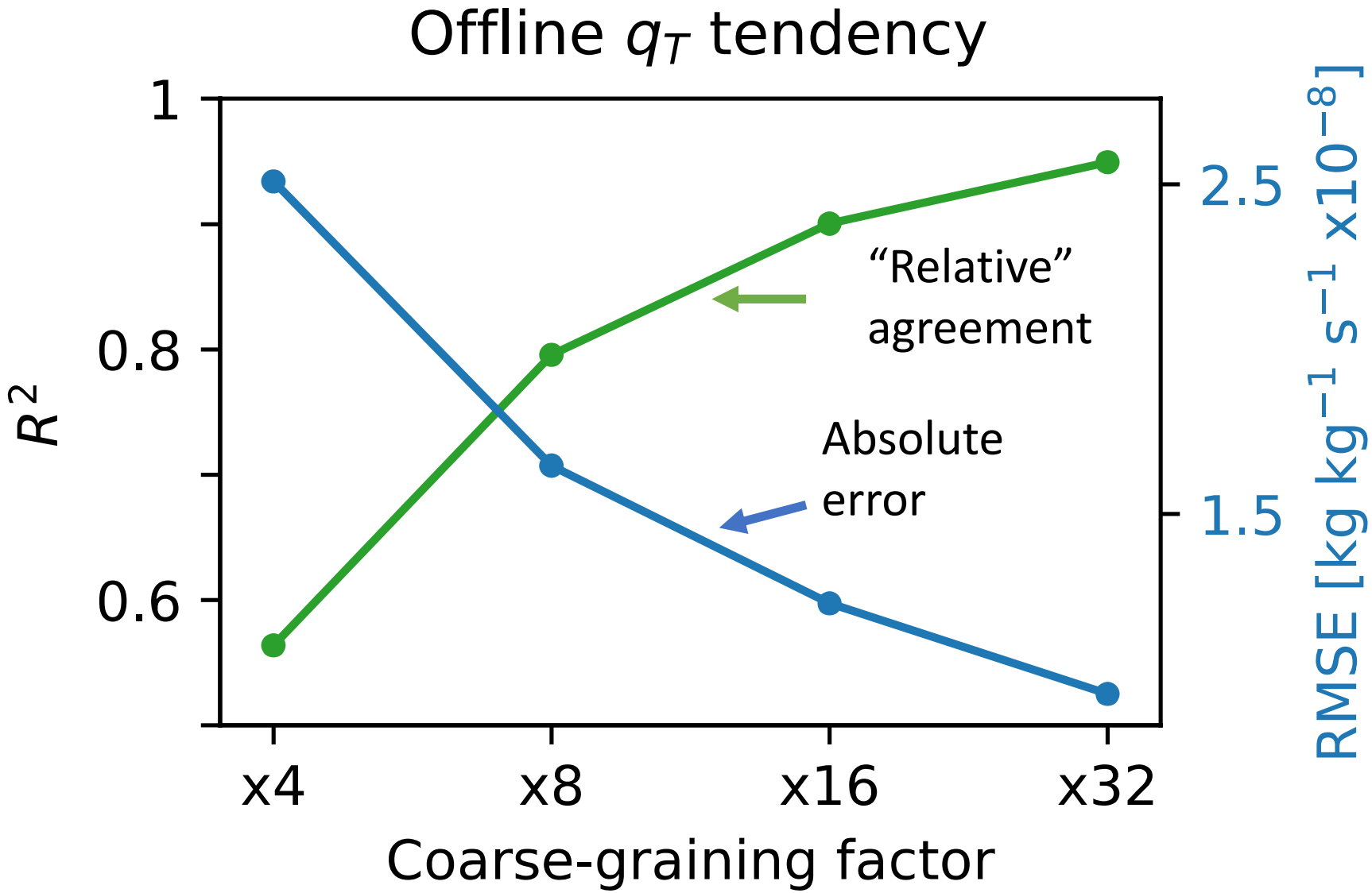


Yuval & O'Gorman, Nat. Commun. (in press)

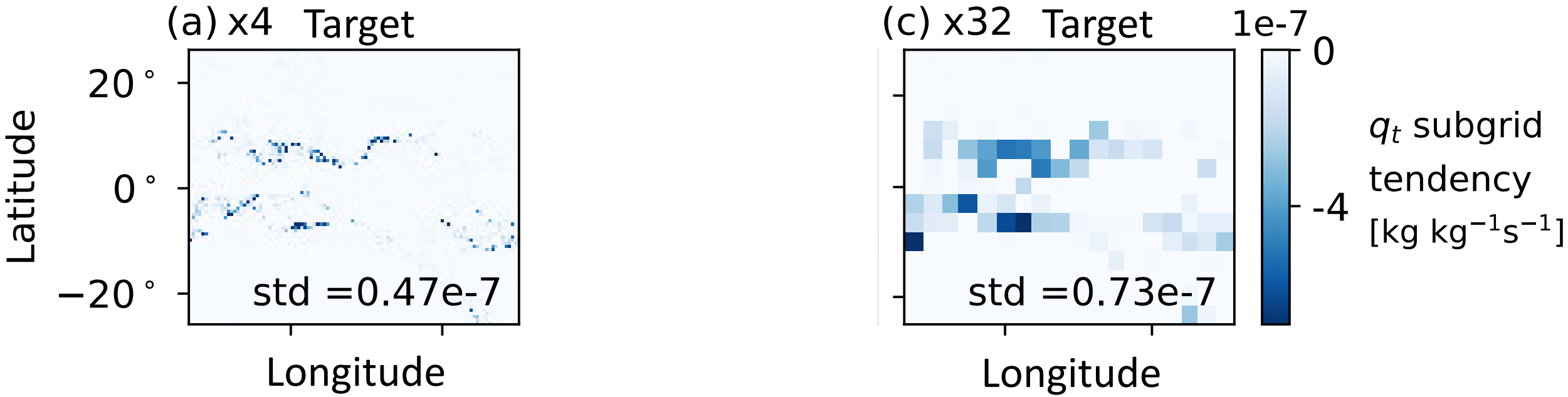
Discrepancy between offline and online performance regardless of whether consider relative or absolute measures of error



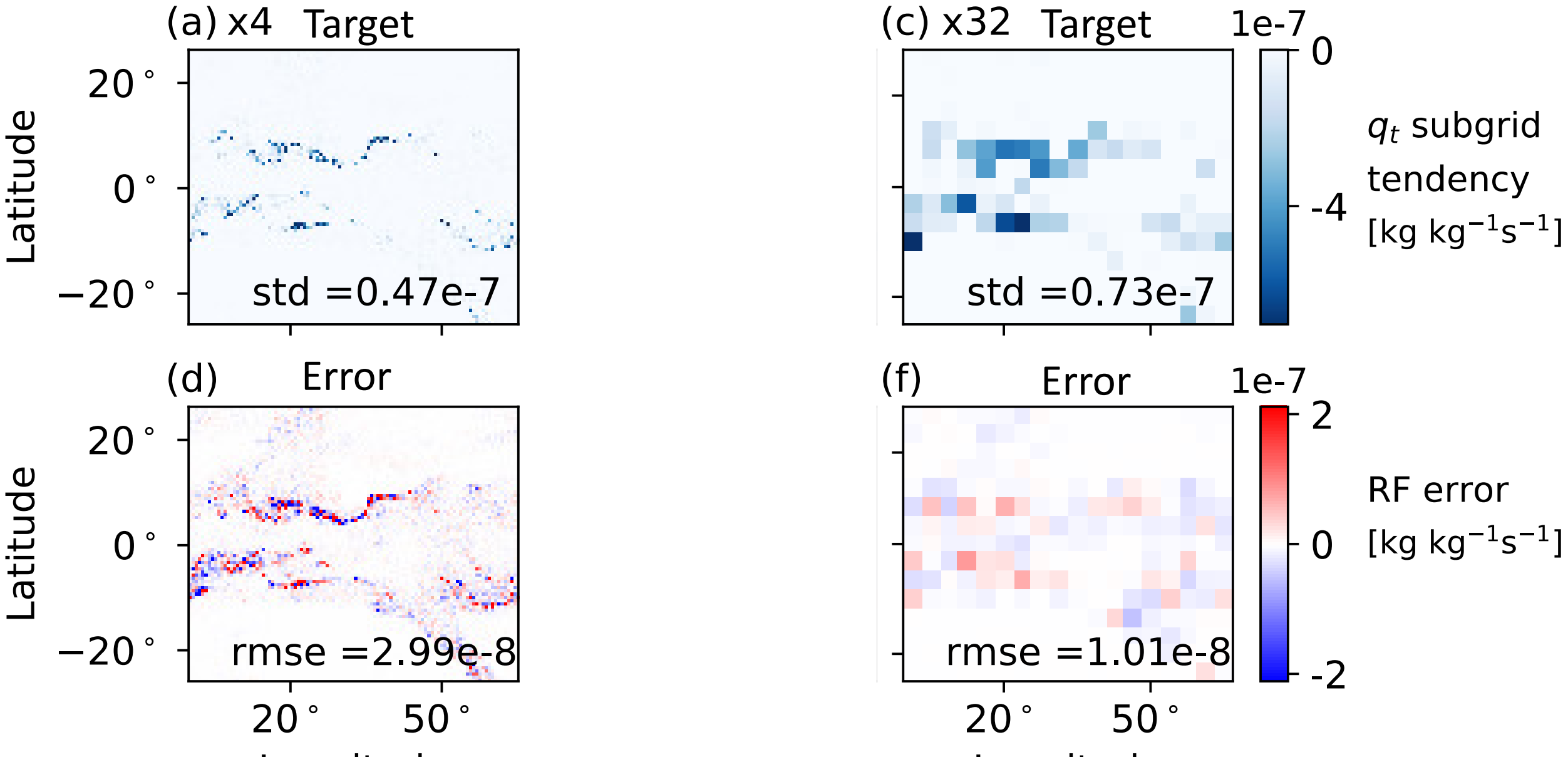
Discrepancy between offline and online performance regardless of whether consider relative or absolute measures of error



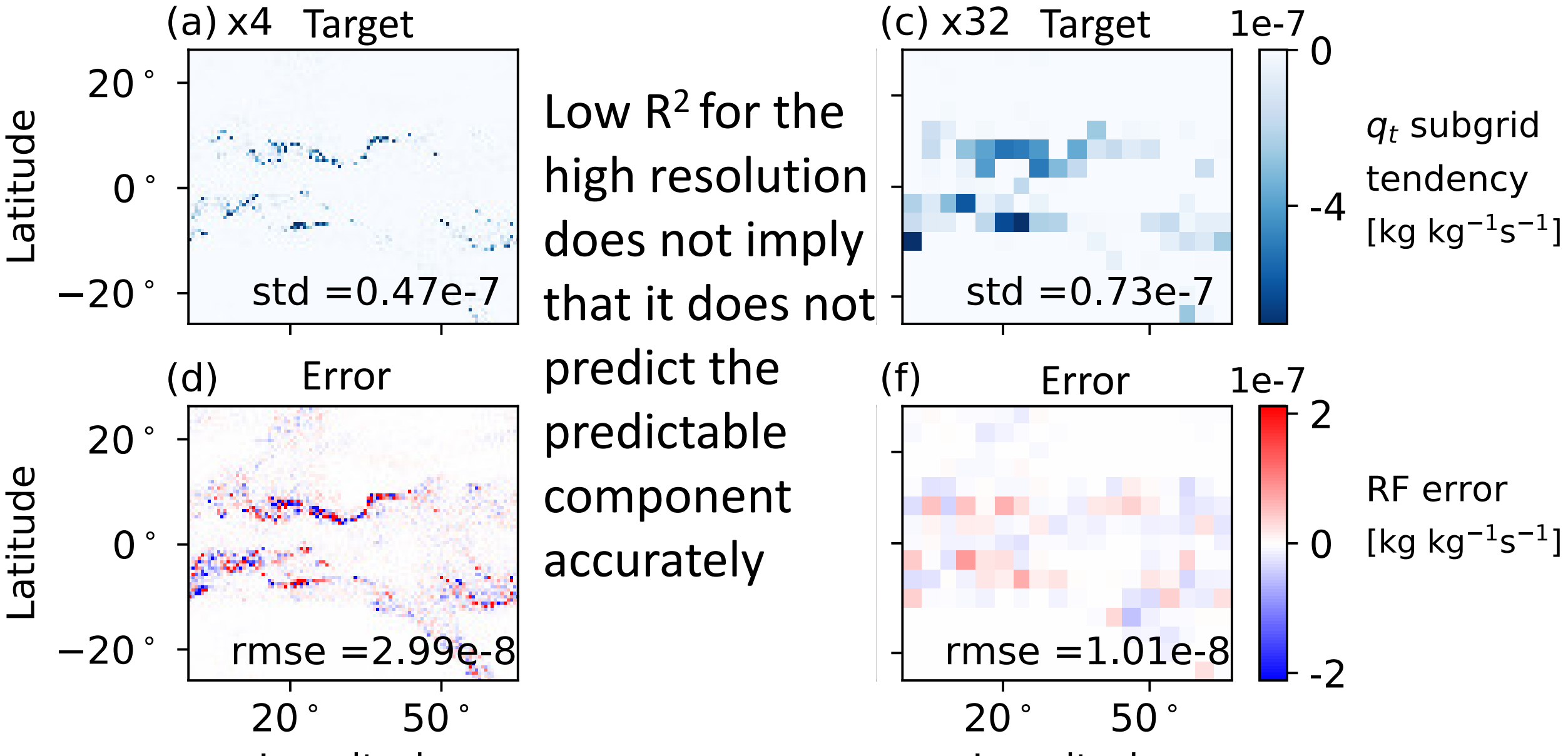
The outputs have (a) a predictable component and (b) a stochastic component



The outputs have (a) a predictable component and (b) a stochastic component

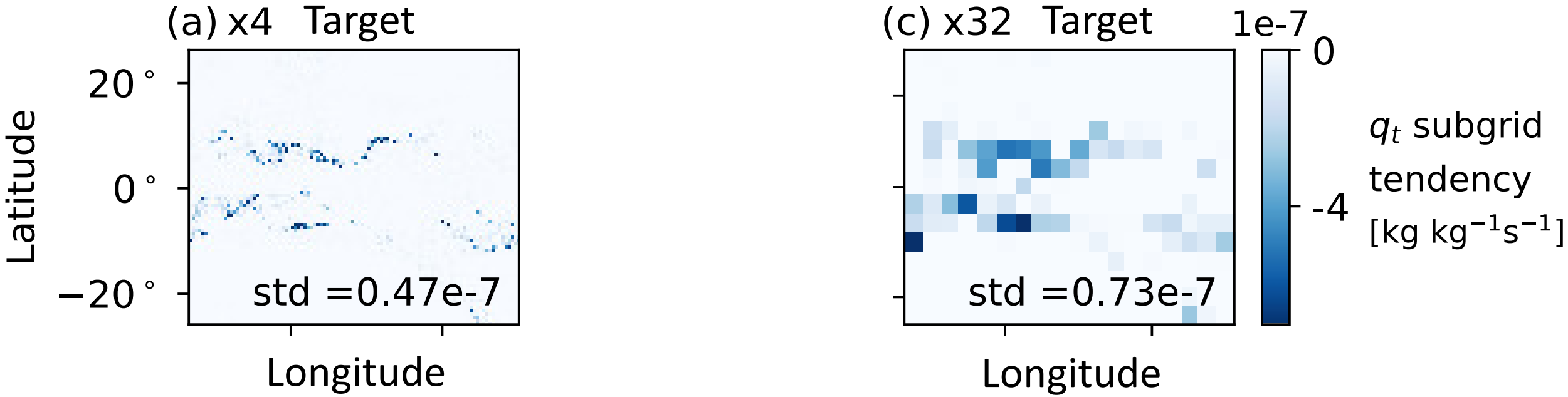


The outputs have (a) a predictable component and (b) a stochastic component

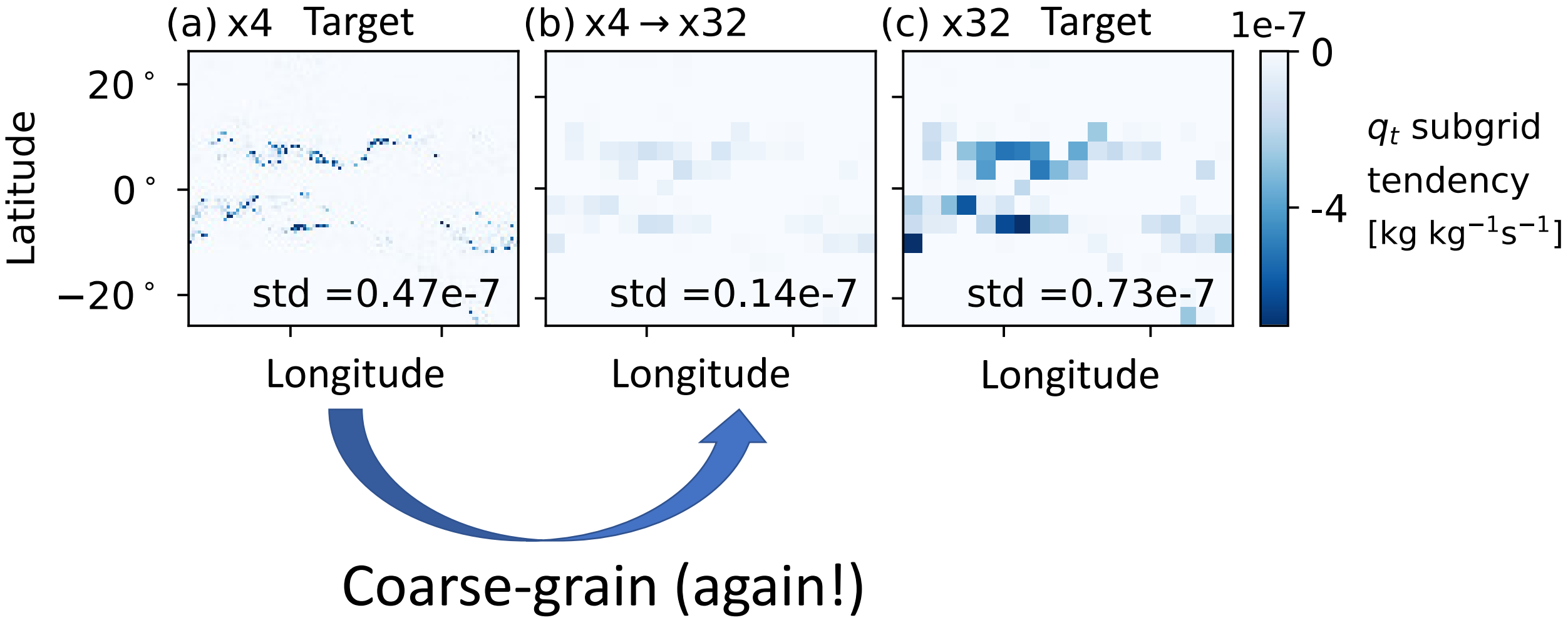


Low R^2 for the high resolution does not imply that it does not predict the predictable component accurately

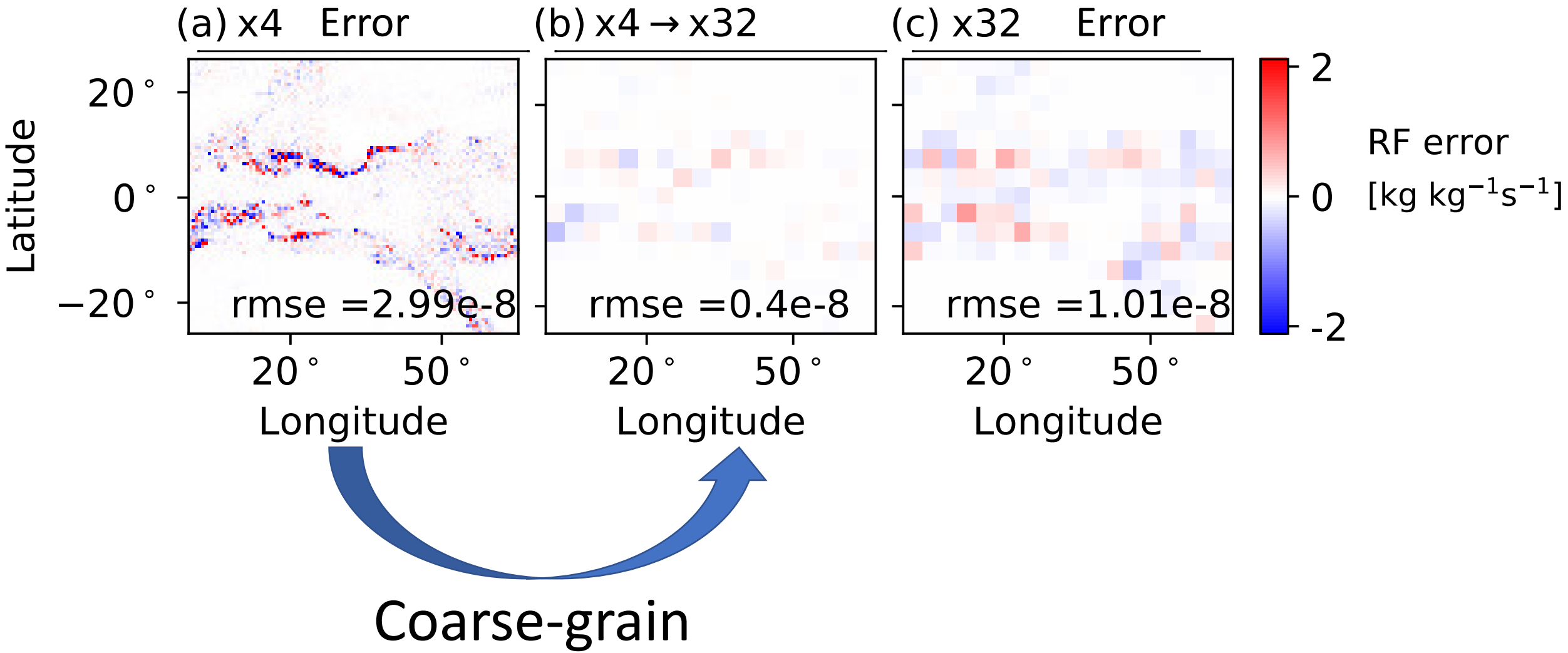
To compare “apples-to-apples” we coarse-grain the higher resolution parameterization



To compare “apples-to-apples” we coarse-grain the higher resolution parameterization

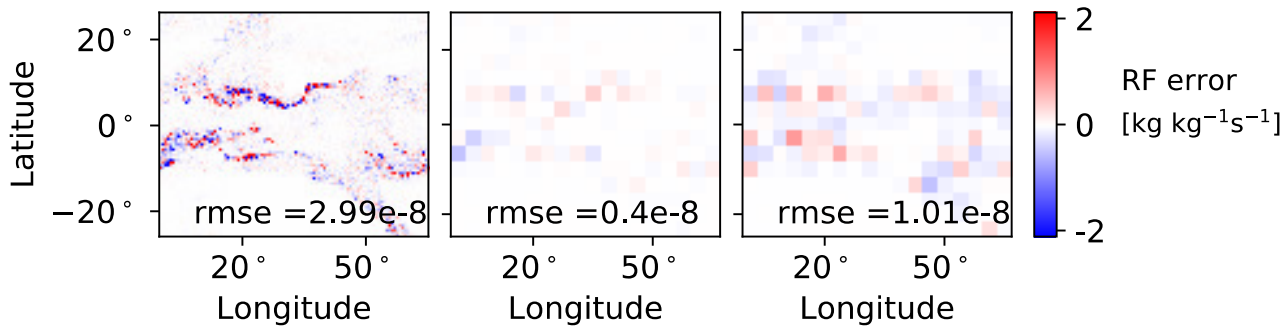


RF with smaller coarse graining performs better in an “apples-to-apples” comparison (at the same length scale)

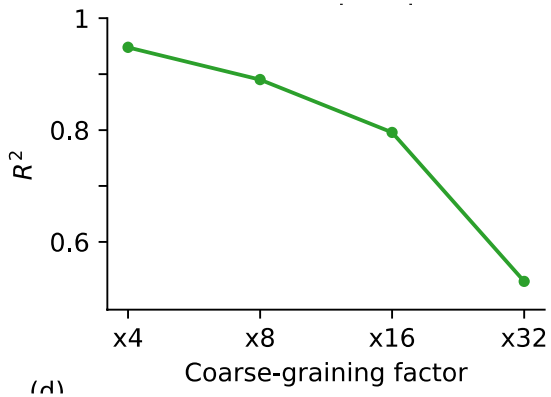


Conclusions

- Offline performance can be misleading in comparisons across length scales unless first coarse-grain to a reference length scale

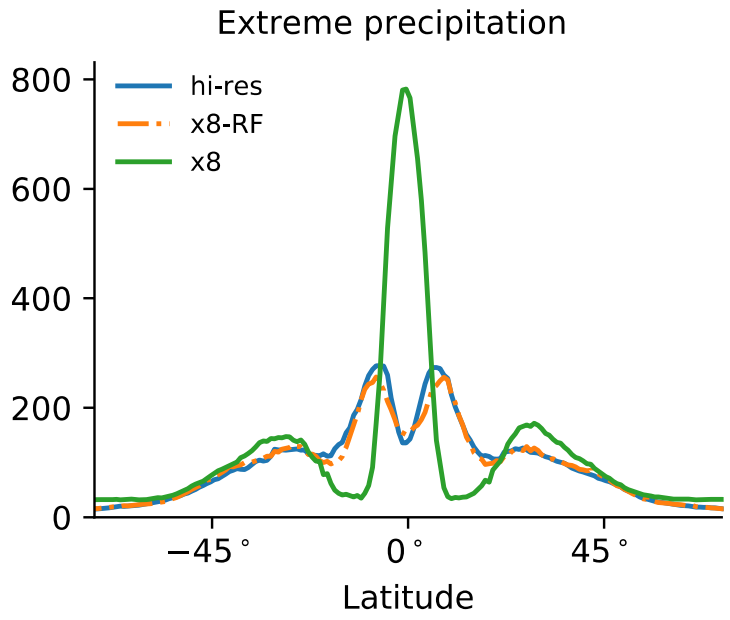


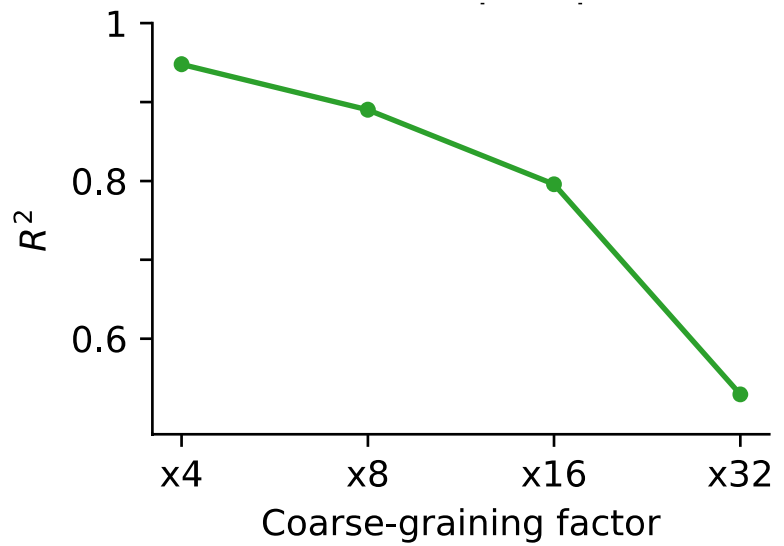
- ML parameterization works better as we decrease grid spacing (no evidence for optimal length scale or gray zone)



Main conclusions

- Random-forest parameterization learned from fully 3-D high-resolution simulation gives stable and accurate simulations at climate-model resolution
- Machine learning can give insights into the parameterization problem (e.g., scale dependence)
- Machine learning parameterization has potential to work well in the "gray zone" of conventional parameterizations





Low resolution simulations with random forest parameterizations are faster than high resolution simulations

- 96km-RF requires **x30** times less CPU time than the high resolution simulation (12km)
- 192km-RF **x120** less resources
- Without changing the time step

A different approach for Random forest parameterization

- Precipitating water (qp) is a variable that changes rapidly (due to precipitation falling with gravity)
- Cannot significantly increase the time step when correcting qp
- Want to avoid using qp as a prognostic variable

Want to avoid using precipitating water as a prognostic variable

h_L – The liquid/ice water static stability energy

$$h_{\mathrm{L}} = c_p T + gz - L_c(q_c + q_r) - L_s(q_i + q_s + q_g)$$

q_T – total precipitating water mixing ratio

~~q_p – total non – precipitating water mixing ratio~~

$$y = \left(\frac{\partial h_L}{\partial t}, \frac{\partial q_T}{\partial t}, \frac{\partial q_p}{\partial t} \right)_{\mathrm{subgrid}}$$

Want to avoid using precipitating water as a prognostic variable

h_L – The liquid/ice water static stability energy

$$h_{\mathrm{L}} = c_p T + gz - L_c(q_c + q_r) - L_s(q_i + q_s + q_g)$$

$$H_L = c_p T + gz - L_c q_c + -L_s q_i$$

q_T – total precipitating water mixing ratio

~~q_p – total non – precipitating water mixing ratio~~

$$\left.\frac{\partial H_L}{\partial t}\right|_{\mathrm{subgrid}}$$

$$y = \left(\frac{\partial h_L}{\partial t}, \frac{\partial q_T}{\partial t}, \frac{\partial q_p}{\partial t}\right)_{\mathrm{subgrid}}$$

+ Change equation of motions in SAM

coarse grained + **simplified** RF parameterization have similar precipitation a to a high resolution simulation