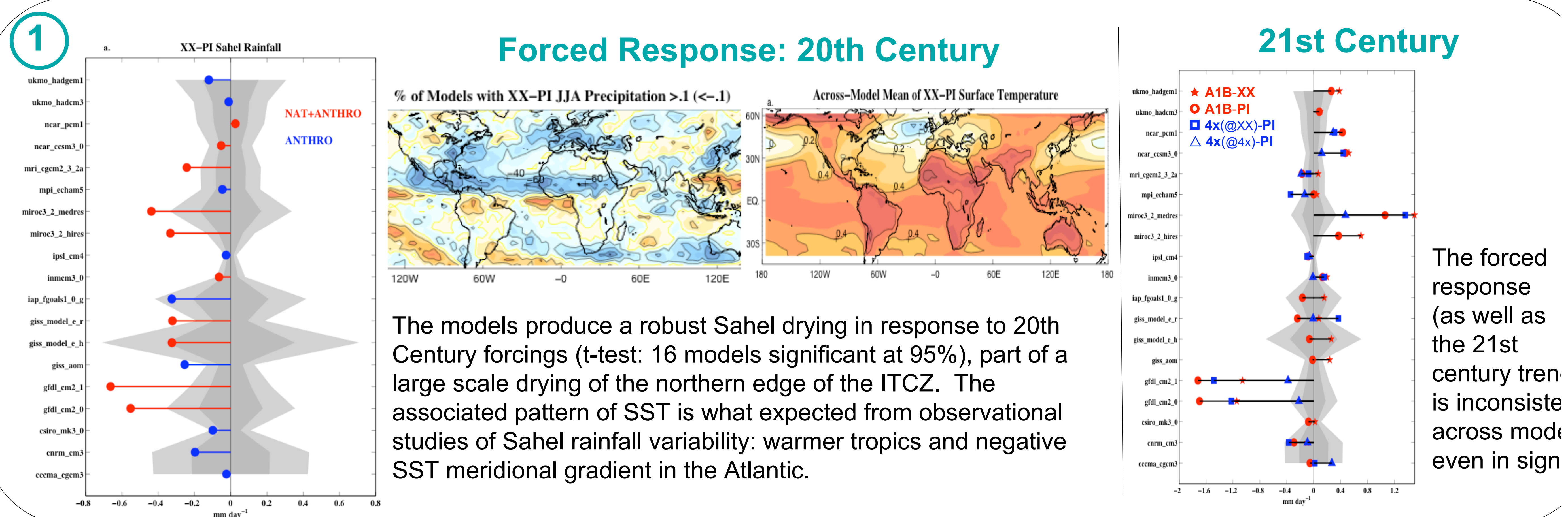


## Abstract:

When should we trust or distrust a model's projection for the future climate? The usual assumption is that a model with lesser biases and a better representation of the past is more trustworthy for the future. The example of Sahel rainfall prediction shows that this is not necessarily the case.

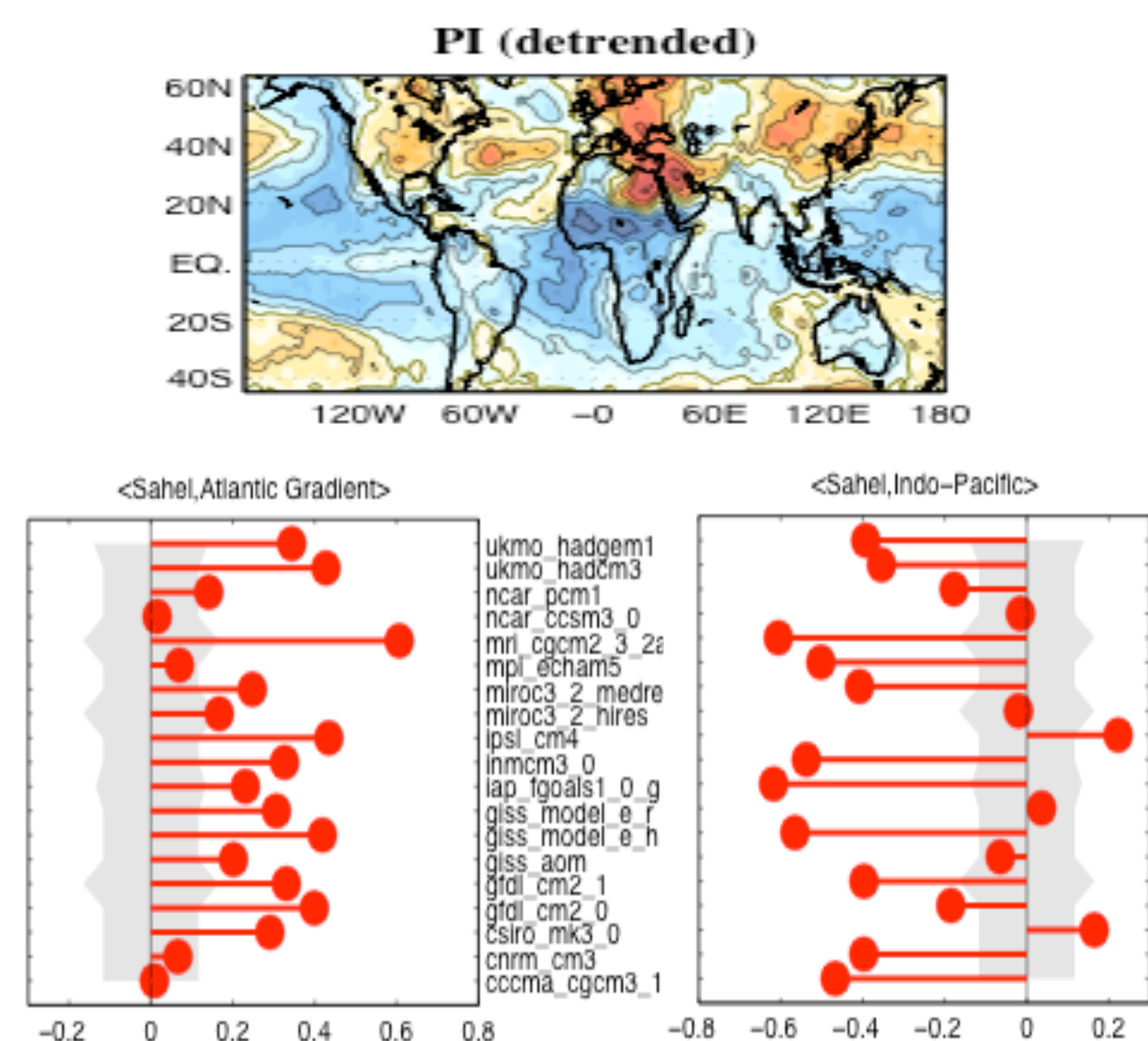
The IPCC models reproduce the late 20th century Sahel drought, but show great disagreement in their outlook for Sahel rainfall in the 21st century (panel 1). Past variations in Sahel rainfall were forced, approximately linearly, by SST variations, but the same rainfall/SST relationship does not explain the projected changes in Sahel rainfall (panel 2).

We use a "perfect model" set-up to show that a metric that privileges small biases and past success is ineffective in selecting trustworthy model predictions (panel 3).

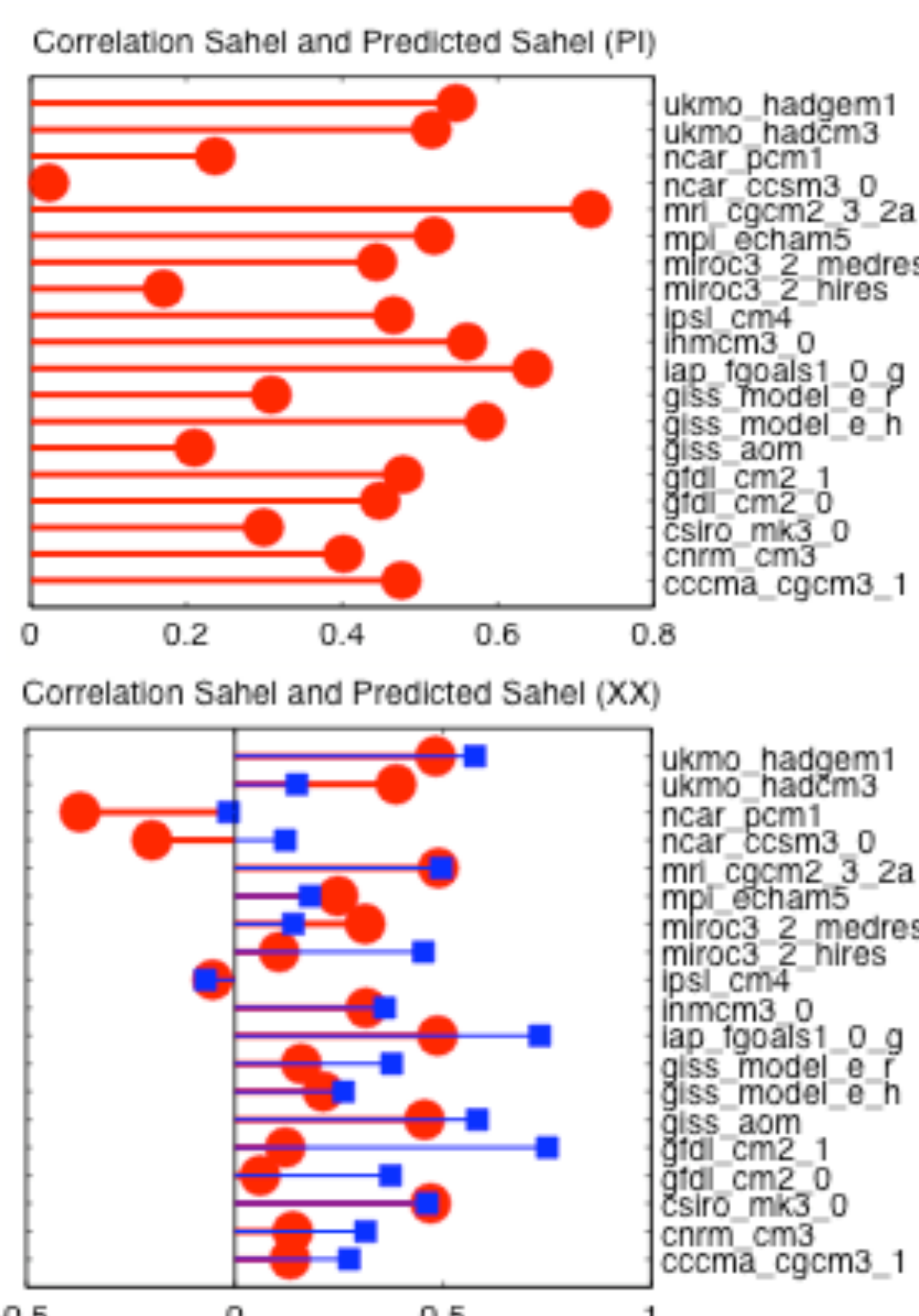


## 2 Sahel rainfall / SST relationship

The relationship between Sahel rainfall and SST in natural variability: positive correlation with the Atlantic meridional gradient and negative correlation with Indo-Pacific SST.



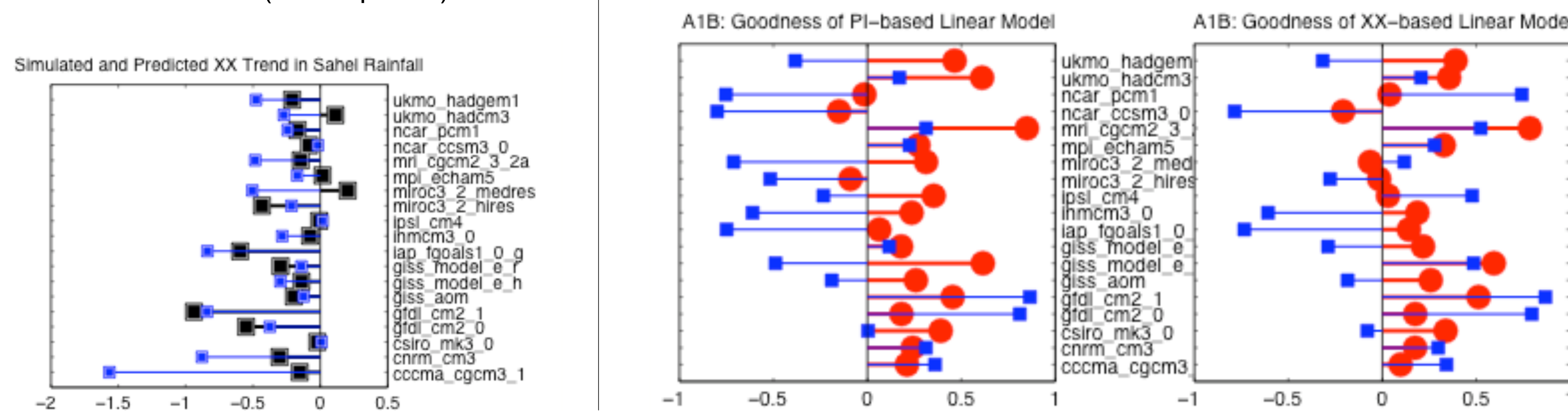
## 20th century SST trends explain the 20th century Sahel rainfall trend via the same relationship seen at interannual timescales in natural variability.



A bi-variate linear regression model trained on the PI runs captures each model's sensitivity to Atlantic Gradient and Indo-Pacific SST. The linear model has predictive skill for both interannual variations (red circles) and the centennial trend (blue squares).

## The interannual SST/Sahel rainfall relationship is maintained during the 21st century, but the trend cannot be explained by SST via the same relationship.

Whether we train our bi-variate linear regression model on the natural variability in the PI runs or on the natural+forced variability in the XX runs, the SST-based linear model has no predictive skill for the centennial trend (blue squares) in a majority of models.

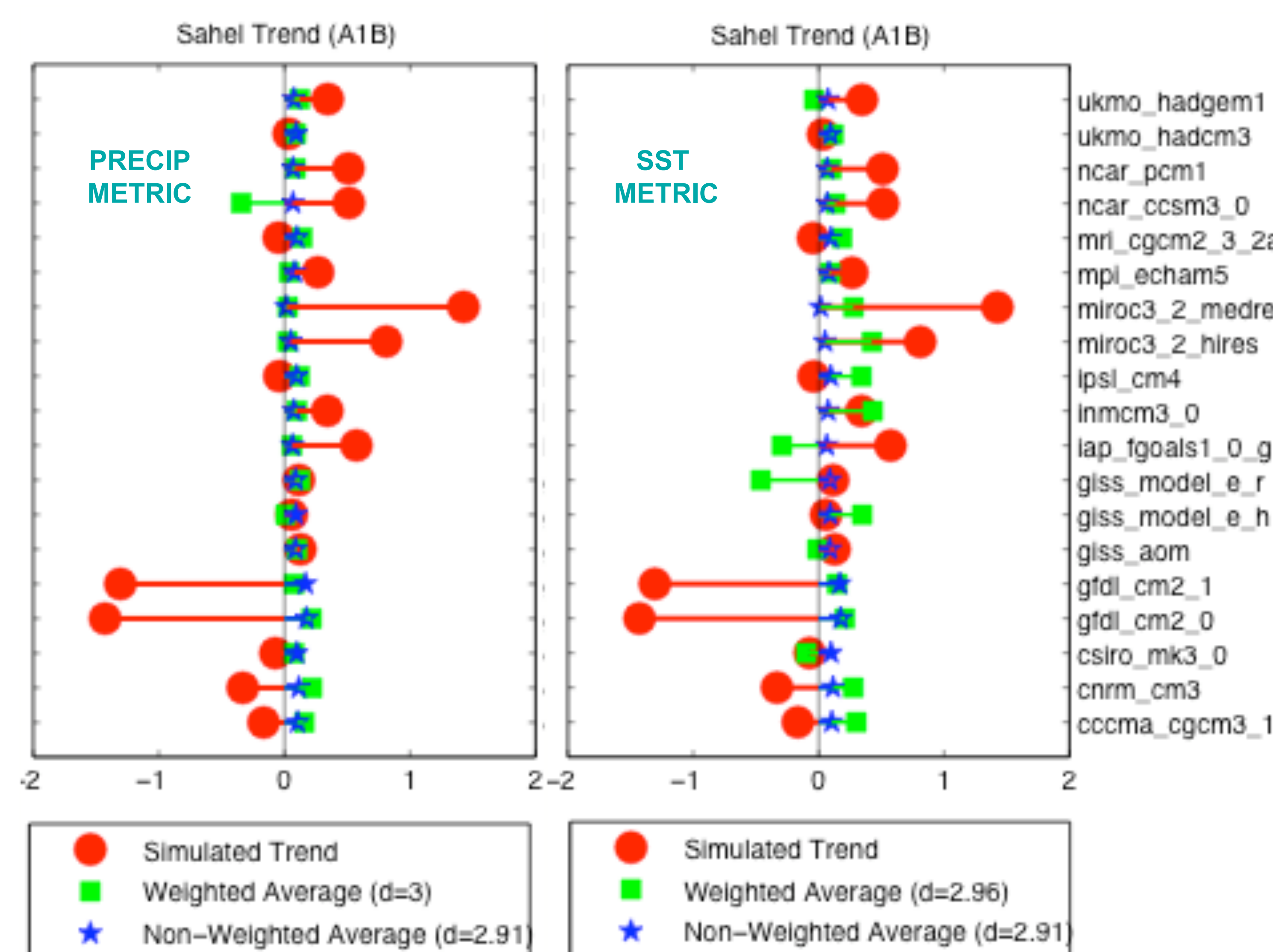


## 3 A problem-specific, objective metric for the "goodness" of models?

We test two common approaches for selecting the most trustworthy projection among many models' disparate ones. One is to compare to observations each model's climatology (for the problem of Sahel rainfall projections see, e.g. Cook and Vizzy, 2005), the second is to compare the natural variability (again Cook and Vizzy, 2005 and Lau et al. 2005). The test is done for a perfect model scenario: we take each 20th century simulation in turn as the truth, weigh all other models according to how well they reproduce such truth in the 20th century, and then compute the weighted mean of the 21st century Sahel trend. If the weighted mean is consistently closer to the 21st century trend in the "true" model than the arithmetic mean, then the metric is effective in isolating more trustworthy models.

Our "precip metric" takes into account the pattern correlation of climatological JAS precipitation, and of its pattern of variability in the African sector and the seasonal evolution of the Sahel rainfall index. Our "SST metric" rewards similar pattern of correlation between Sahel rainfall and SST and similarity in the regression coefficients between Sahel rainfall and Atlantic gradient and Indo-Pacific SST calculated for the bilinear regression model above.

Neither metric does any better than the arithmetic mean in reproducing the "true" trend in the 21st century simulations.



## Conclusions

Variations in SST are the proximate cause for the 20th Century internal variability and the anthropogenically forced changes in Sahel rainfall.

In the 21st century, the historic relation between Sahel rainfall and SST breaks down, and the projections for future Sahel rainfall are very uncertain and model-dependent.

A model's lesser biases in the climatology and success in reproducing the past history of the Sahel are not the right criteria to select a trustworthy future projection.