

Authors' response to interactive comment by Reviewer #1 Juraj Parajka

Black text: Reviewer comment

Blue text: Authors' response

We thank the reviewer for his valuable comments and suggestions that will help us improve our manuscript. Below we reply to each of these and explain how we will incorporate them into the manuscript.

The study evaluates snow and runoff performance of 64 snow routine alternatives based on degree-day approach in large sample of catchments (54) located in Swiss and Czech Republic. The snow routine variants are coupled with HBV conceptual hydrologic model and model simulations are evaluated in terms of observed daily runoff and snow water equivalent observations/estimates. The results indicate that exponential snowmelt function with no refreezing and seasonally variable degree-day factors are the most reliable/robust/accurate variants for snowmelt runoff simulations in selected catchments. Overall, this is an interesting study which is worth to publish. The topic is relevant and within the scope of the journal. The study is clearly written and has a good structure. The analyses and interpretations are based on larger sample of catchments which allows to draw interpretations/conclusions that are relevant for large region of similar physiographic conditions in Central Europe.

Thanks for these kind words.

I have only few comments/notes which can be considered (in my opinion) to add/extend/improve clarity and generality of findings. These include:

1) Perhaps it will be possible to refer here in general to variants of degree-day snow approach, not strictly limits the analysis to HBV variants. The results can be used/implemented in degree-day routines of different hydrological models. In this study, the variants are coupled with HBV concept of rainfall-runoff transformation, but I believe, at least the evaluation of snow efficiency is relevant to general degree-day approach.

We also agree that the evaluation of the snow simulations is relevant to the degree-day approach in general, beyond its use in the HBV model. Indeed, what we present here is a methodology to analyse the impact of using different alternative model structures for a specific purpose (here, snow processes) on the performance of a rainfall-runoff model over a large sample of catchments. We, of course, related the proposed modifications to the HBV model, as this is our tool to conduct the analysis. Nevertheless, the alternative model structures that we explored in this study are not only "HBV variants", but at the same time also variants of the degree-day approach as used in other hydrological models as well. We understand the concern of the reviewer and we will clarify in the manuscript that, while using the HBV model, our study also is an evaluation of the degree-day approach in general.

2) When coupling the 64 snow routine variants with HBV model, there is another interesting question, which can be discussed and this is the robustness/uncertainty of other HBV model

parameters. How consistent/different are the other HBV model parameters for different snow variants? Are, for example, field capacity or nonlinear runoff generation (beta) parameter values similar or compensating some effects of different snow routines?

The reviewer raises an important point here. In models such as the HBV model, model parameters can compensate each other, which makes the interpretation of any modifications rather challenging. While we did not include this in the manuscript, we performed a Monte Carlo sensitivity analysis on the HBV parameters. We provide figures resulting from these analyses for each catchment at the end of this comment (caption only provided for Figure 1). We found that, even if some of the variants (i.e. $T_{p,m}$, ΔP_e , or $C_{0,s}$) produce compensating effects on some parameters (e.g. PCALT, FC, LP or BETA), this effect was only observed for some of the catchments. Overall, parameter values and sensitivity tended to be fairly consistent across all the tested model variants for most of the catchments. We are aware of these potential compensatory effects between model parameters which can mask the real impact of different snow routine variants and, therefore, decided to base the evaluation of the analysis on the ability of the different model variants to reproduce snow water equivalent in addition to stream runoff (which is how hydrological models are traditionally evaluated).

3) It is not clear which part of the snow accumulation/melt phases are described/evaluated by selected snow objective function? For some practical applications, for example, it will be interesting to see the difference in maximum snow water equivalent between the routines, or to what extent the model over or underestimates snow cover duration? To what extent are these aspects covered in current snow efficiency evaluation? Does a good simulation mean well represented maximum SWE or snow cover duration? Perhaps there are some differences in such efficiency between the variants.

Our evaluation of the performance of the snow simulations provides an overall assessment of the snow processes in the selected catchments. As the reviewer correctly states, the evaluation could also be based on more specific aspects such as magnitude or timing of maximum annual snow accumulation. The problem with measures based on specific aspects is that a perfect fit with regard to one single measure does not ensure a good overall performance (similar to individual flow indices or signatures in the case of runoff simulations, see Vis et al., 2015). This implies that a number of measures would be needed to be used together with the challenge to decide on appropriate ways to combine the measures into a single overall performance measure. While we agree that this could allow further assessments and might be valuable for future studies, we are afraid that such an additional analysis would be beyond the scope of this manuscript.

4) In our recent study (Sleziak et al., 2020) we found that there are quite significant differences in snow model performance (by using standard HBV degree-day approach) between lowland and alpine catchments in Austria. (Differences in terms of overestimation of snow cover in alpine and underestimation of snow cover in flatland catchments). Did you observe similar findings here?

Yes, in our study we also observe that model performance is generally lower for lowland catchments than for alpine ones. Our simulations also tended to underestimate snow water equivalent in lowland catchments. Regarding alpine catchments, we observed no clear pattern regarding over- or underestimation. Overall, it was quite frequent for the model to underestimate snow accumulation and delay the timing of the spring snowmelt season. We appreciate the comment of the reviewer and we will expand the related sections in our text and relate our results with the findings from the suggested article (Sleziak et al., 2020).

Specific comments

1) Abstract. Is the last sentence needed?

With this sentence we wanted to highlight some of the limitations of the results obtained in this study. However, as we also discuss these limitations in detail in the manuscript this “general disclaimer” might indeed not be needed. We will remove the last sentence from the abstract.

2) Introduction: It will be interesting to extend somewhat this section by referring to ways how can be/are degree-day routine parameters estimated in hydrological models.

We already do this to some extent in the methods section, where we present the different alternative model structures that we considered in this study, together with references. However, we understand that it would be relevant to refer to these different approaches and implementations in a general way in the introduction as well. We will, therefore, expand the introduction to give an overview of how different hydrological models use the degree-day method (see also response to major comment #1).

3) Data: How close are gridded snow water equivalent data to observations? Is there some bias related to the fact that this dataset is based on some type of degree-day model?

Using a temperature-index (TI) approach in both the runoff model as well as in the snow model providing validation data might indeed lead to some bias. Nevertheless, it has to be taken into account that the snow model makes use of a 3-dimensional sequential data assimilation (DA). The DA itself includes two methods which are both based on spatially correlated error statistics. For snow accumulation, an optimal interpolation approach uses the snow water equivalent station data to correct the simulated snowfall amounts. Concerning snowmelt, an ensemble Kalman filter updates snowmelt rates as well as liquid water content. Finally, the combination of both data assimilation approaches results in corrections of modelled snow water equivalent within all 1 by 1 km grid cells. Magnusson et al. (2014) investigate the performance in predicting snow water equivalent when using this DA approach and compare it to the TI model without DA. Based on 1033 samples from 45 stations, they show that using DA leads to improvements in predicting snow water equivalent.

4) Runoff model efficiency. Why only Nash-Sutcliffe based on logarithmic transformed discharges? It will be interesting to see also the model performance in terms of snowmelt runoff peaks.

When we designed the study, we gave some thought on potential evaluation metrics, among which were the NSE, MARE or snow cover fraction. Nevertheless, since the computational demands for conducting this study were considerable (large array of catchments and model modifications), we decided to just use two objective functions (one for snow processes and another for rainfall-runoff transformation) that would be as relevant as possible. We agree that testing the model performance in terms of snowmelt runoff peaks would be very interesting indeed. Looking at additional performance measures would be a valuable next step, but including this here would make the study overly complex due to the inclusion of too many aspects (see also answer to major comment #3).

5) P.15, l.355: Figure 3 or Figure 4?

This should indeed be Figure 4. We will correct the error.

6) Figure 4. Will it be possible to show such case for a year in the validation period?

The intention behind presenting only the calibration results was to keep a simple story and walk the reader through the results by adding complexity stepwise. Nevertheless, we understand that having some detailed validation results might add valuable information to the reader as well. We will modify the figure to include validation results for the same year.

7) Results: Will it be possible to present runoff and snow model efficiencies for each catchment in the Supplement?

This study includes many different catchments and model variants, in addition to two periods for cross-validation analysis using two different metrics and presenting all these data in a meaningful way is not easy. However, since we only presented absolute results for one catchment in the manuscript, we agree with the reviewer that it would be a good idea to, at least, provide summarised results from all the catchments, periods, and evaluation metrics in an appendix. We will, therefore, include a table showing median efficiency values for each catchment, and objective function, both for model calibration and validation for both periods for the default HBV model. Since we already included relative performance differences for all catchments and metrics in the manuscript (see Figure 5) we argue that this would be sufficient to provide a good overview of the model results.

Additionally, we could provide further figures similar to Figure 4 (including validation results, following the previous comment), for all catchments in an appendix as well. This appendix would, however, be rather extensive and we would appreciate the editor's guidance on whether to include this or not.

References

Sleziak et al. (2020) The effect of the snow weighting on the temporal stability of hydrologic model efficiency and parameters, *Journal of Hydrology*.

Magnusson, J., Gustafsson, D., Hüsler, F., & Jonas, T. (2014). Assimilation of point SWE data into a distributed snow cover model comparing two contrasting methods. *Water resources research*, 50(10), 7816-7835.

Vis, M., Knight, R., Pool, S., Wolfe, W., & Seibert, J. (2015). Model calibration criteria for estimating ecological flow characteristics. *Water*, 7(5), 2358-2381.

Sensitivity analysis

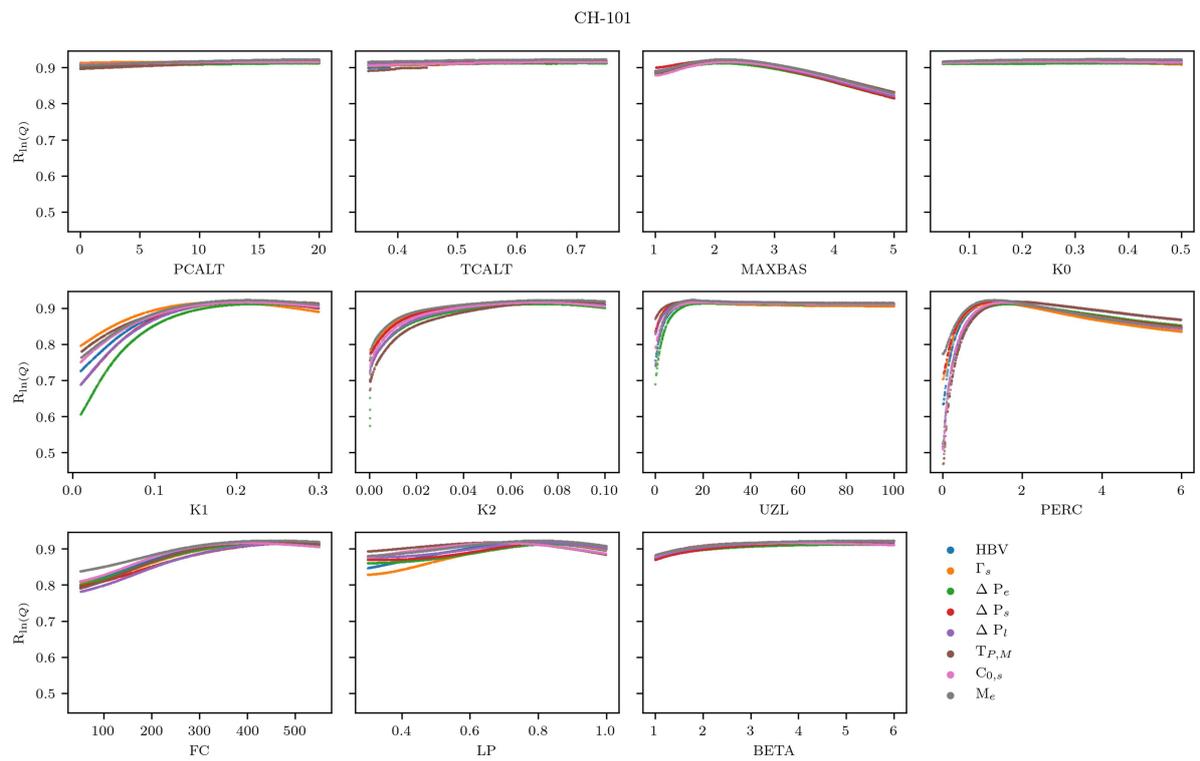
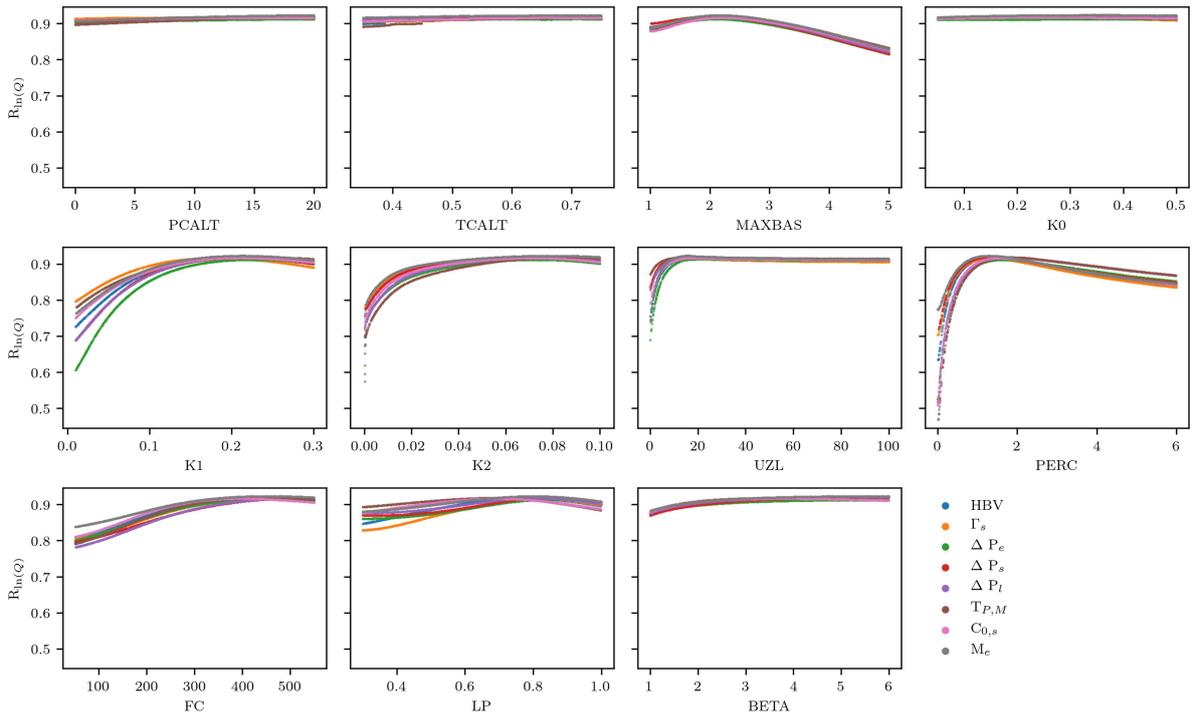
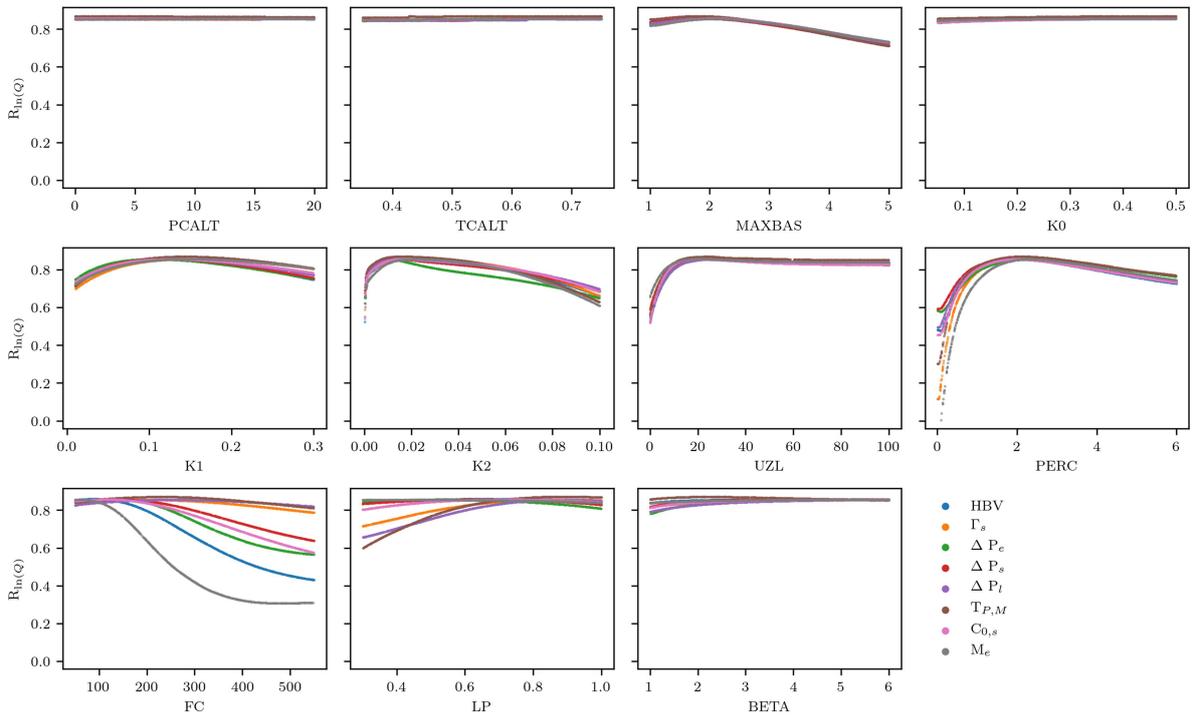


Figure 1. Sensitivity analysis on parameters other than those from the snow routine for a catchment CH-101. Each subplot shows the model performance (y-axis) as a function of the values of a single free parameter (x-axis) with all other parameters set at the median best value of all 10 calibration trials for each of the single modifications to the snow routine of the model (see Table 1 in the manuscript). These results shown here are for model calibration on period 1.

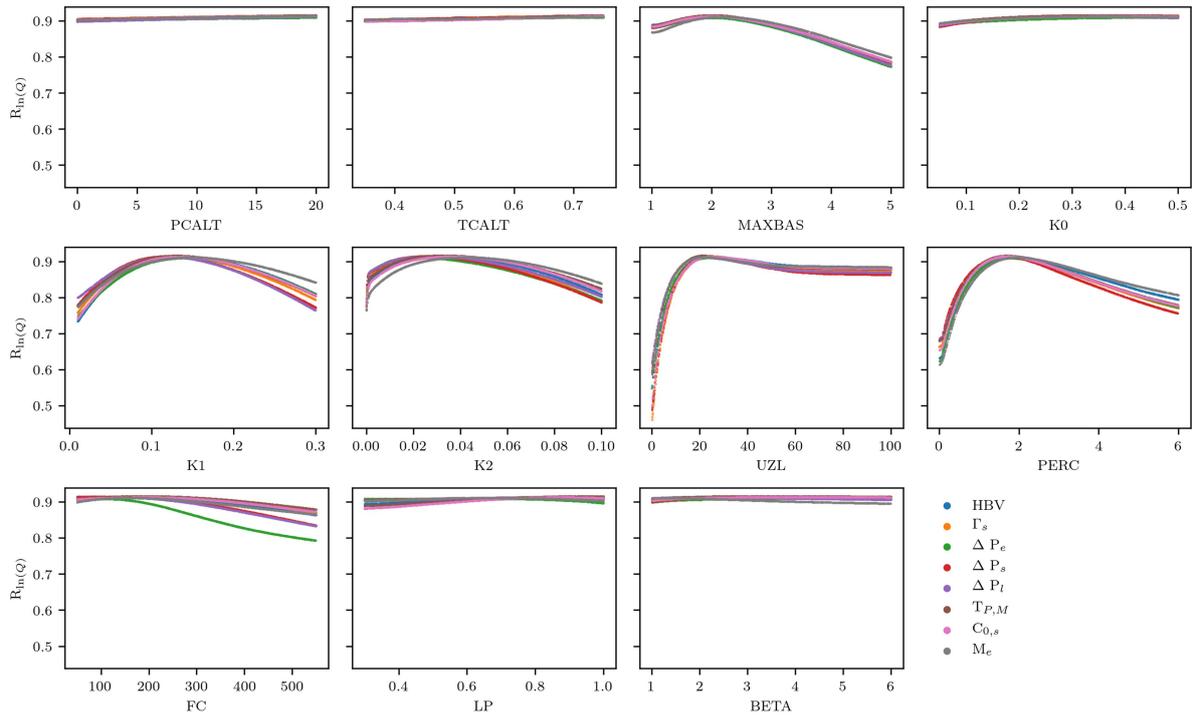
CH-101



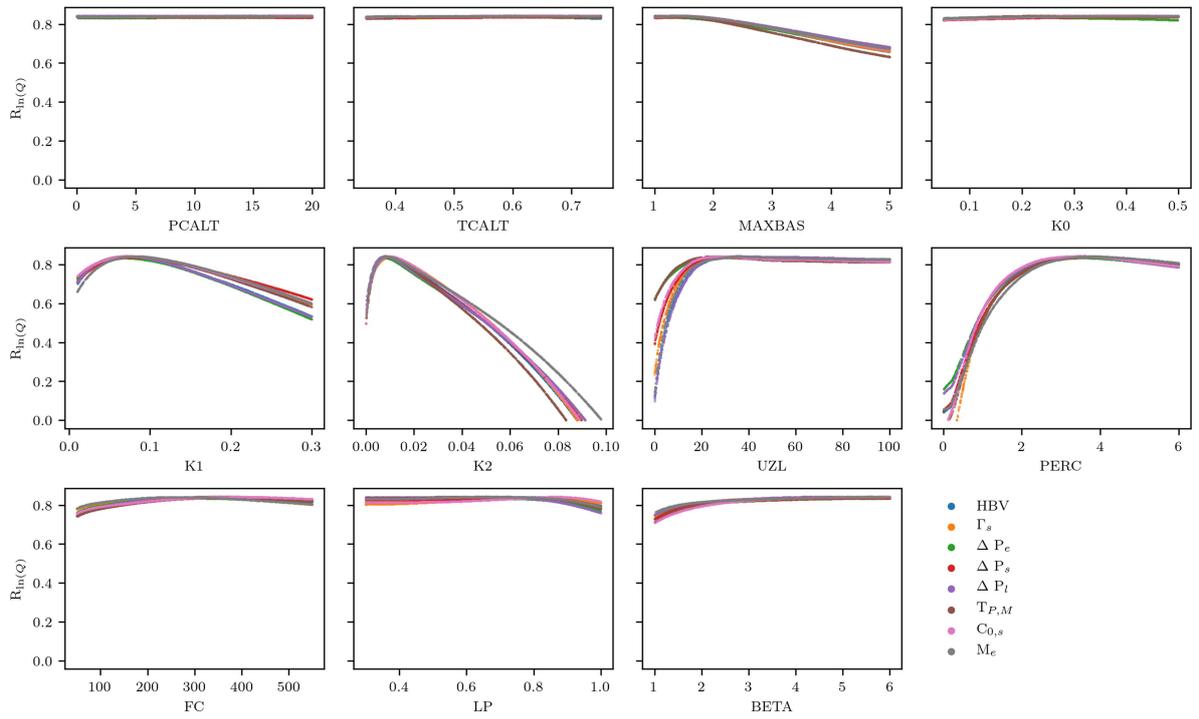
CH-102



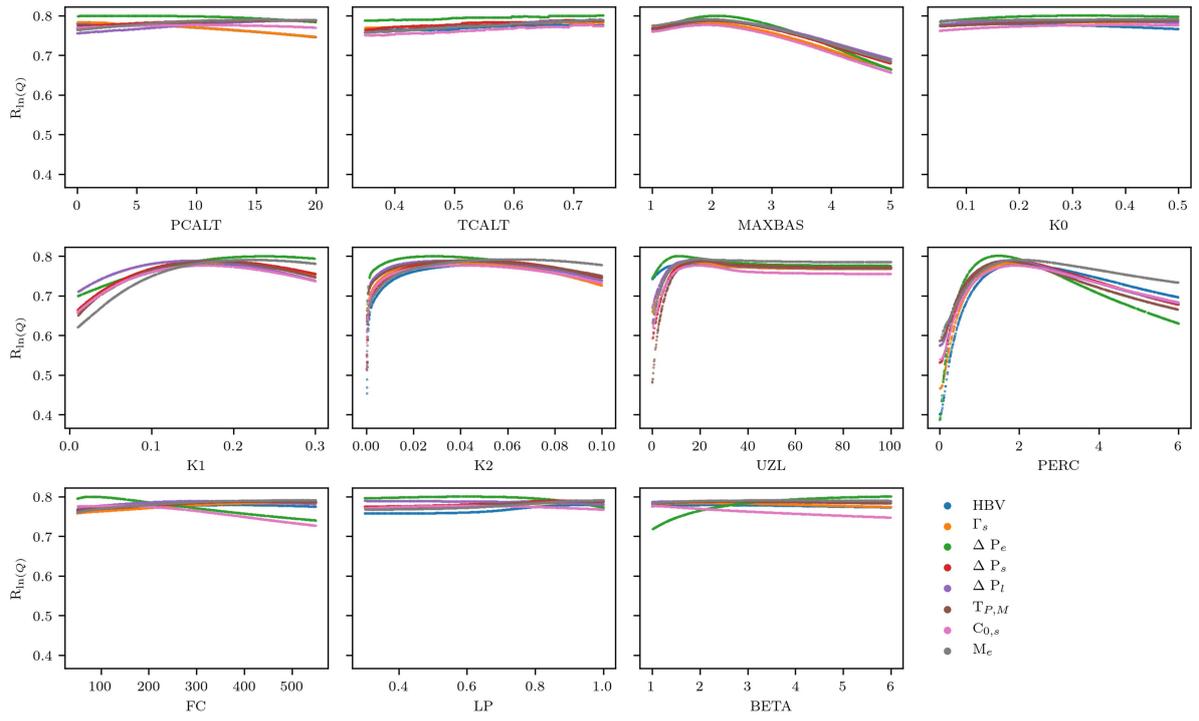
CH-103



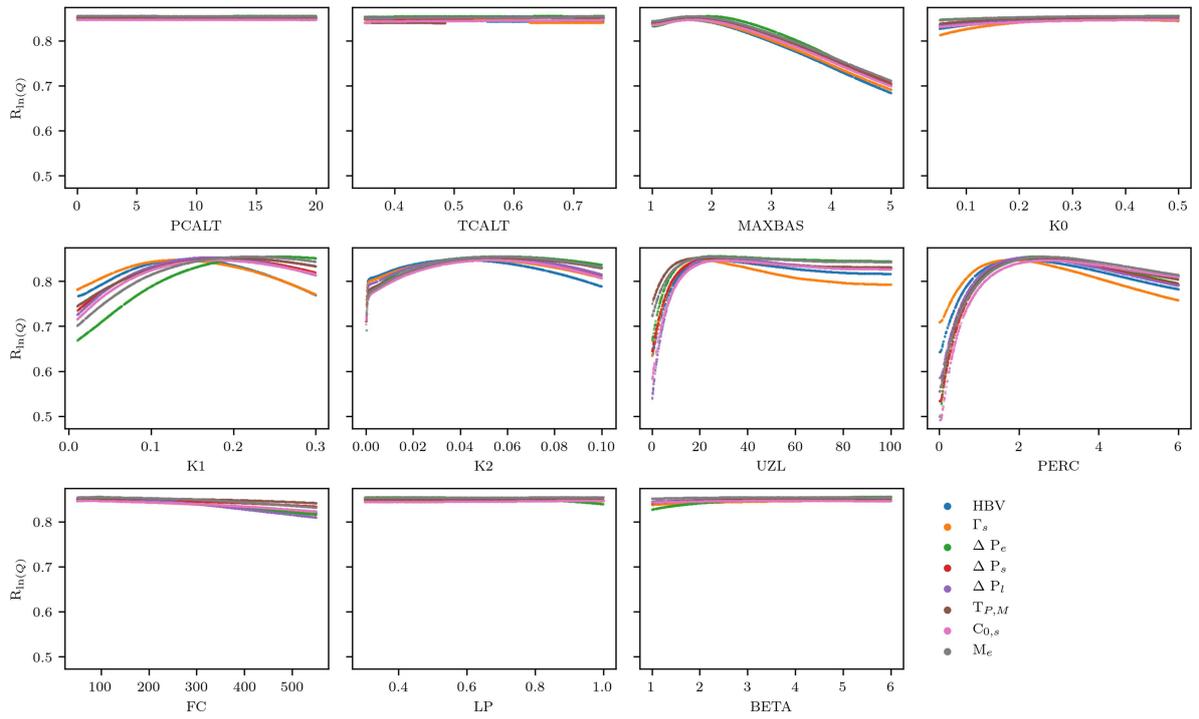
CH-104



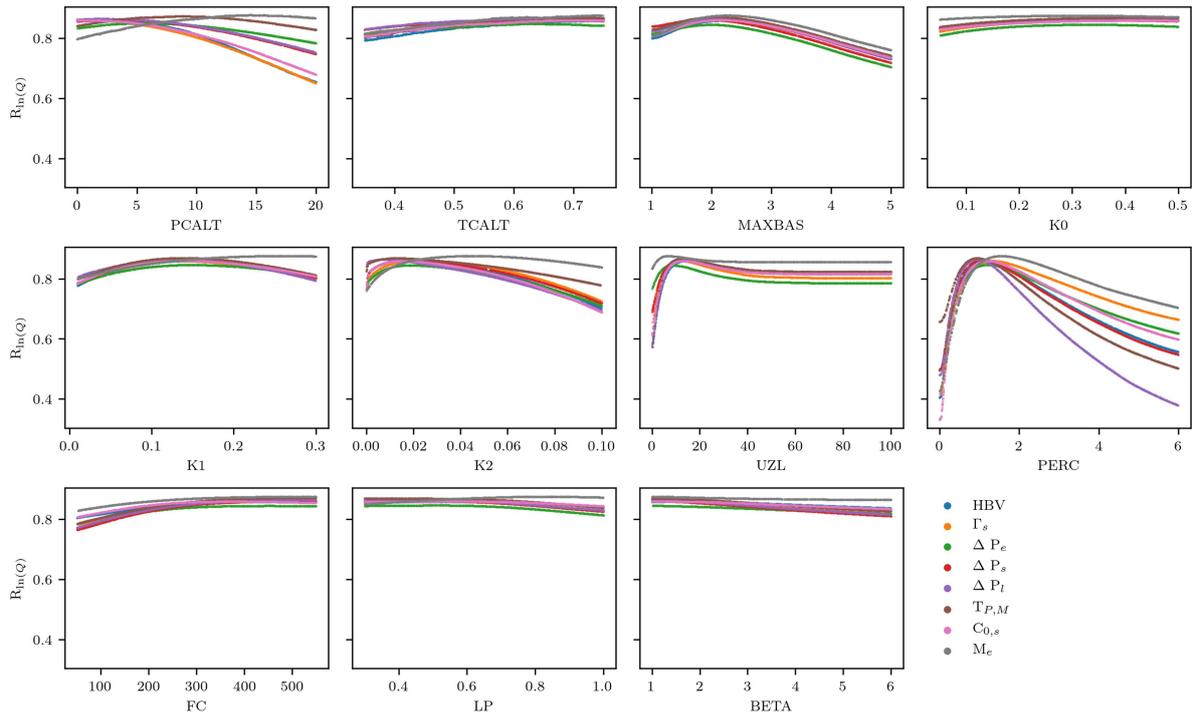
CH-105



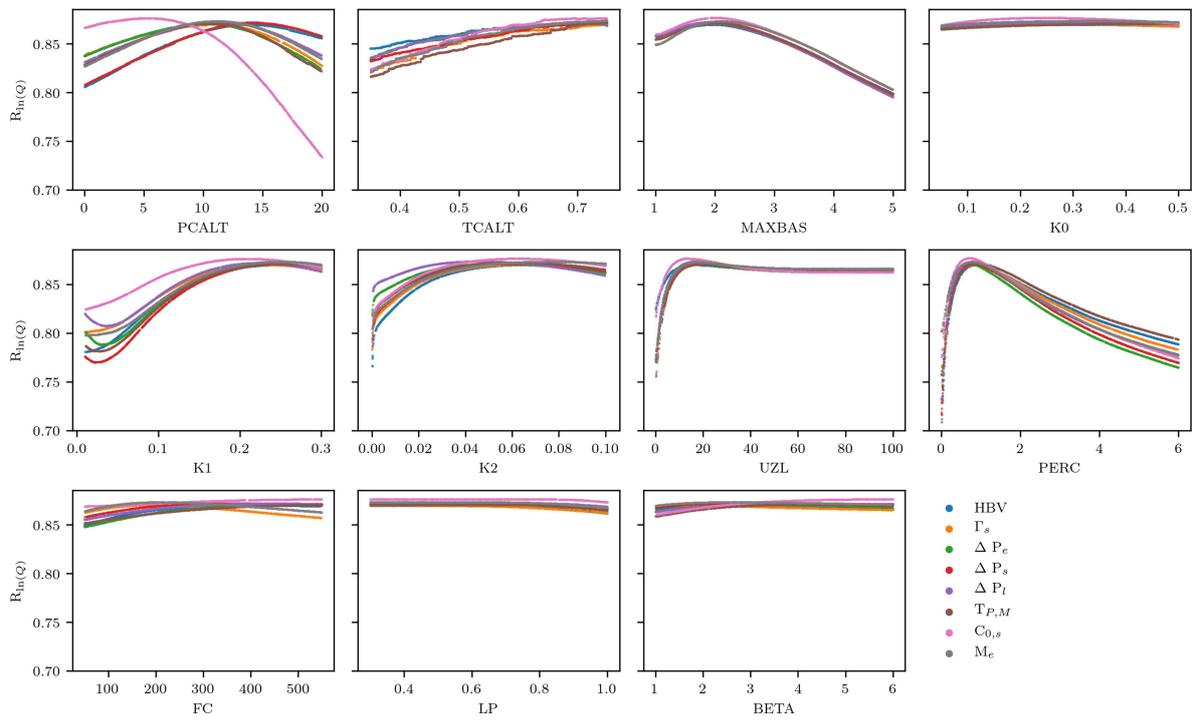
CH-106



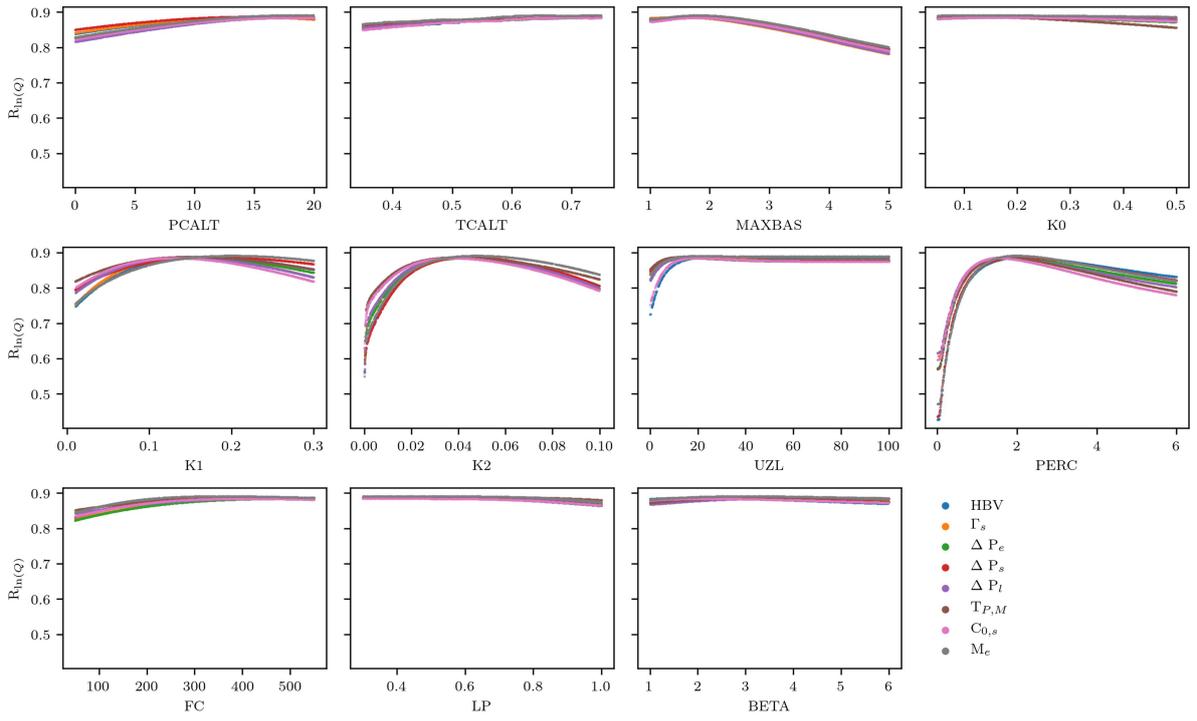
CH-107



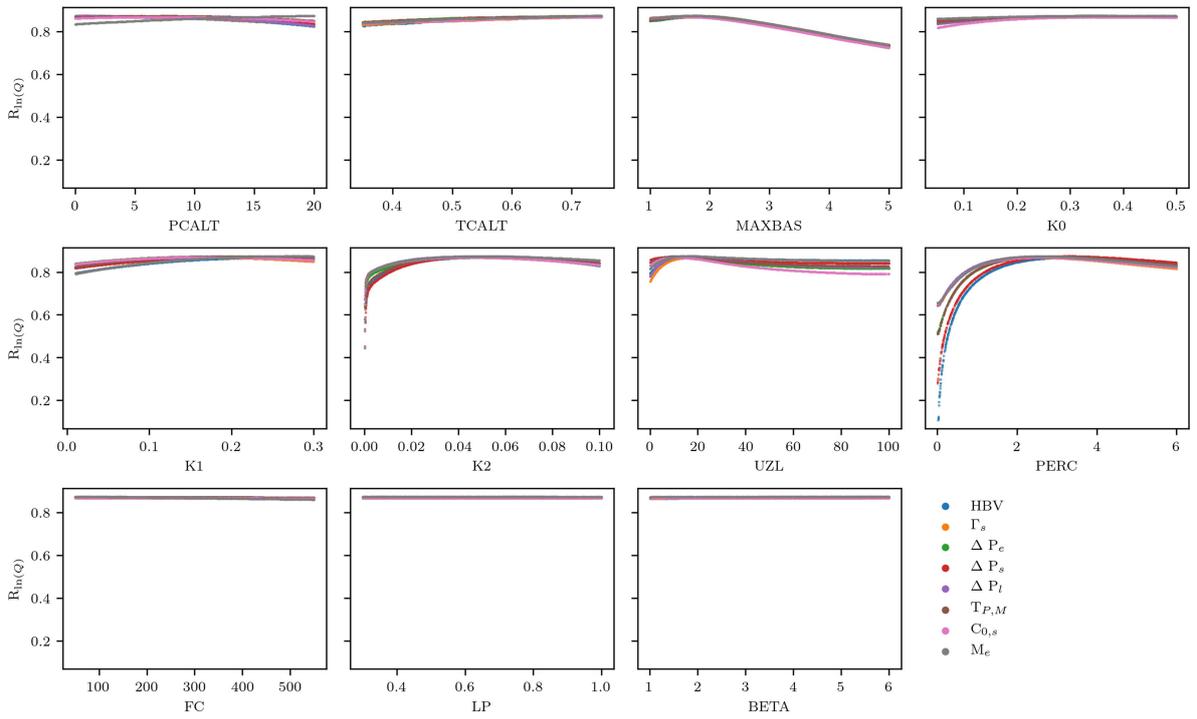
CH-108



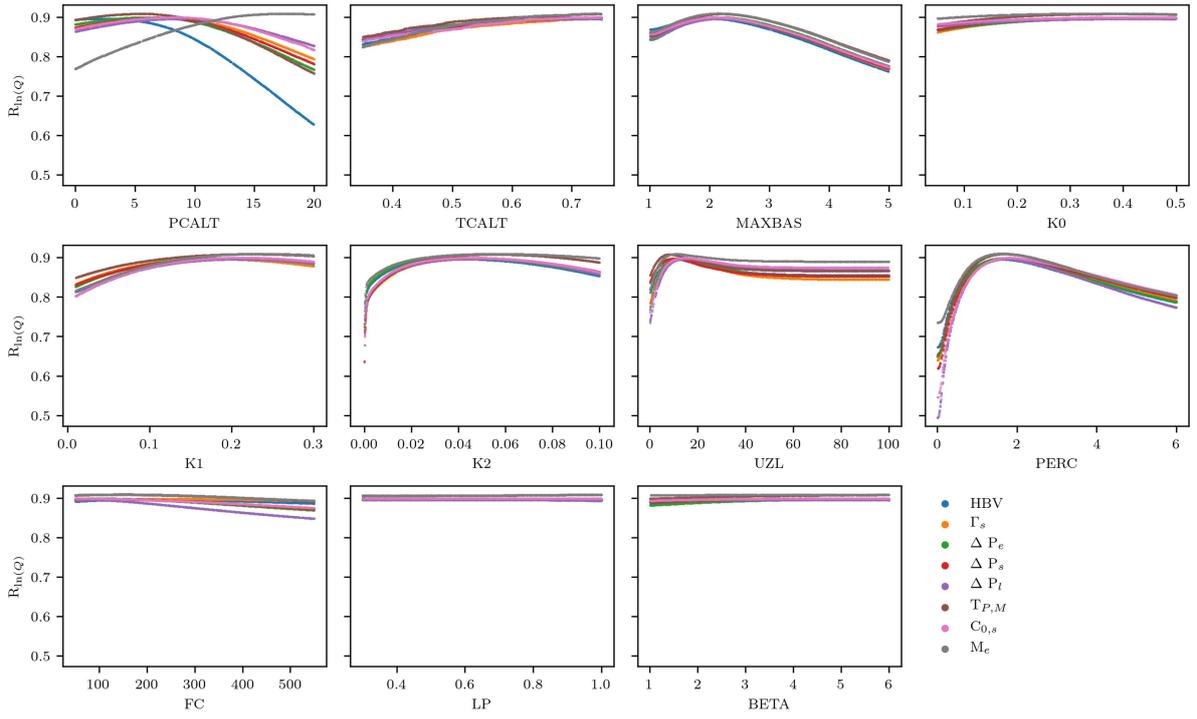
CH-109



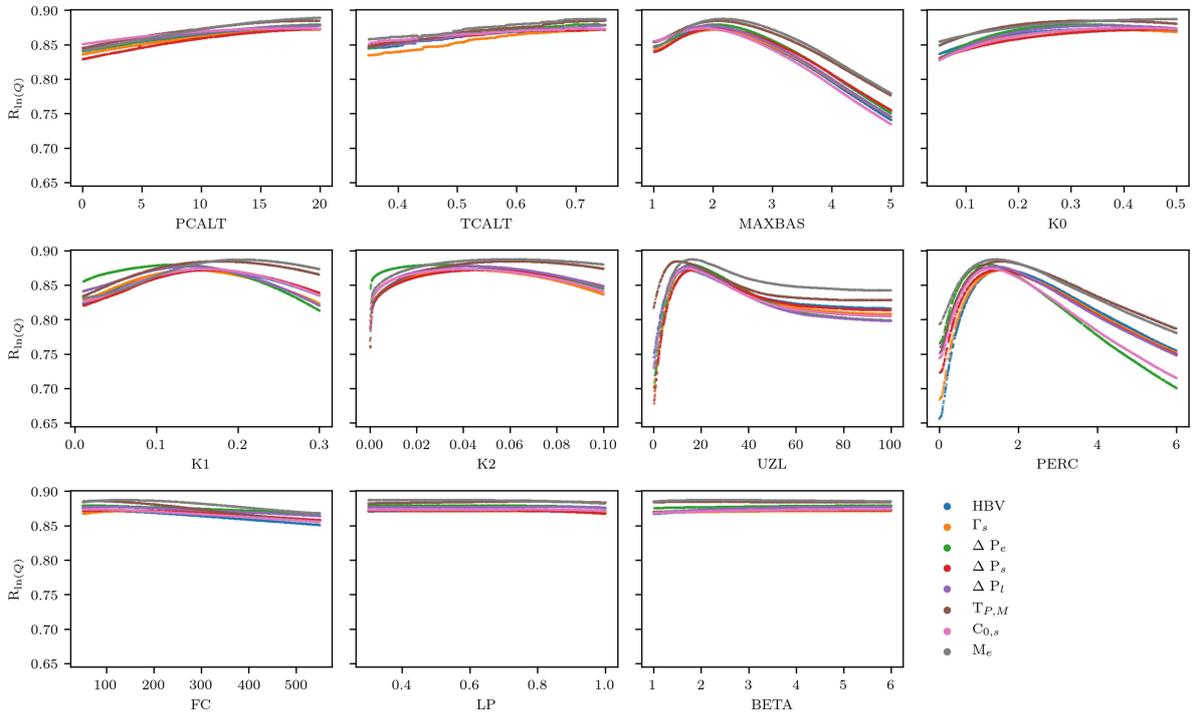
CH-110



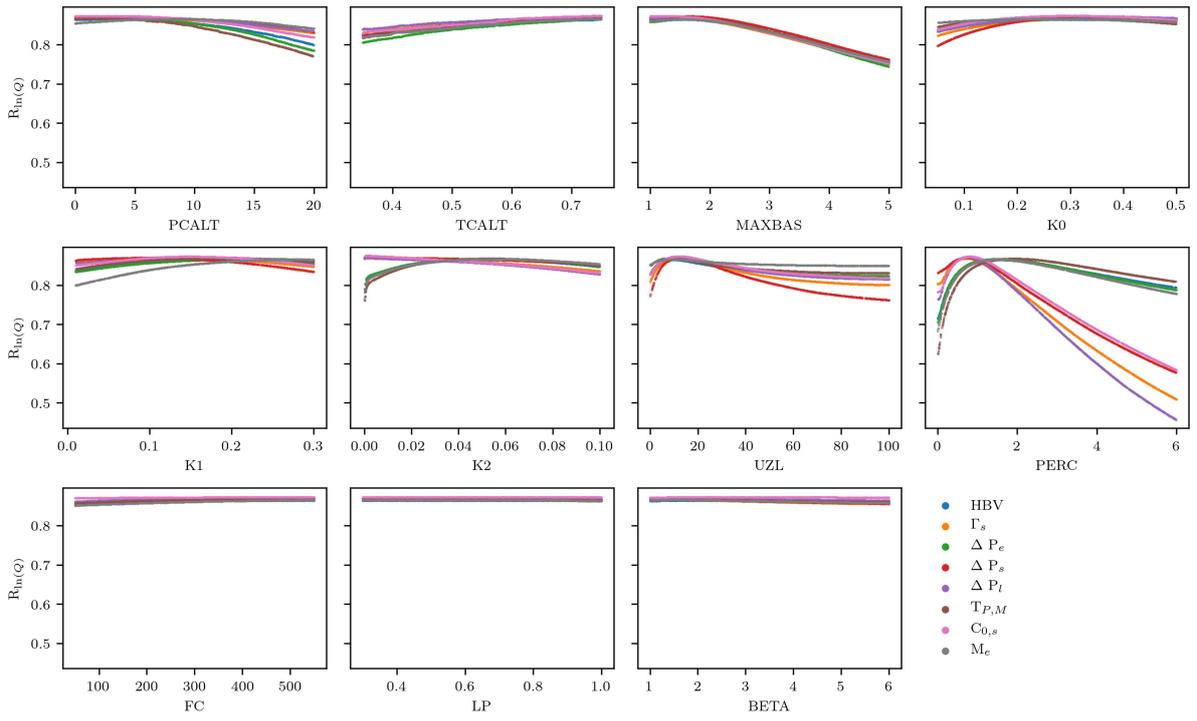
CH-111



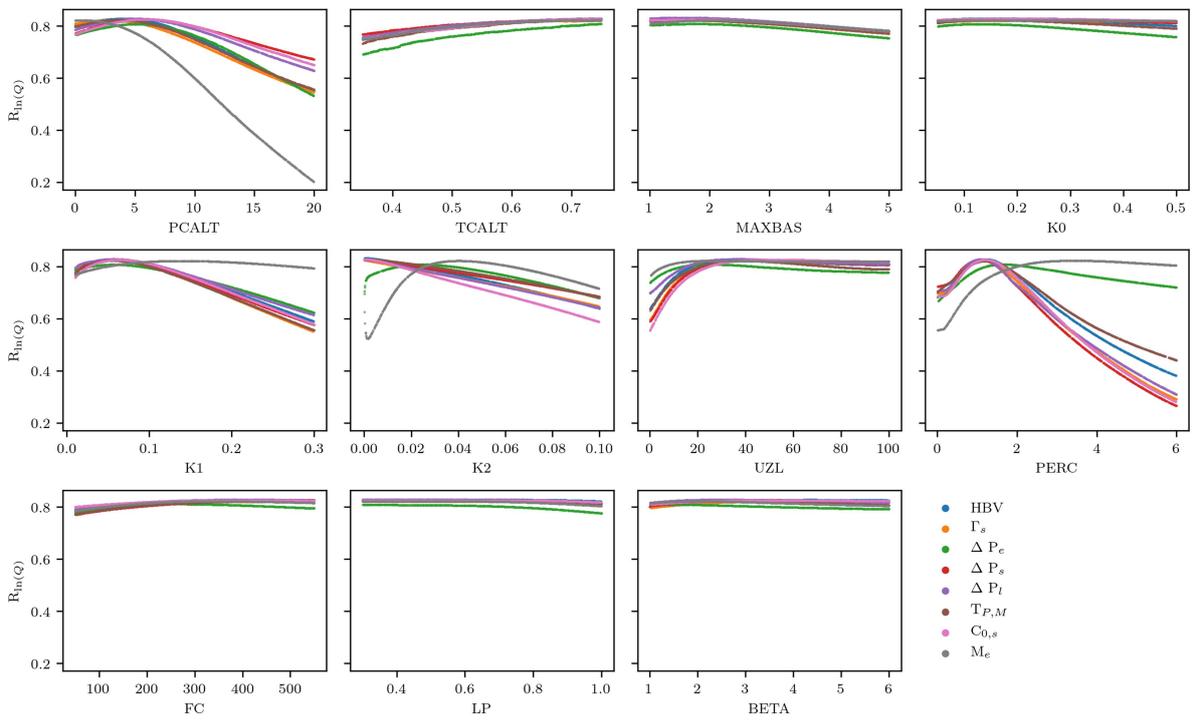
CH-112



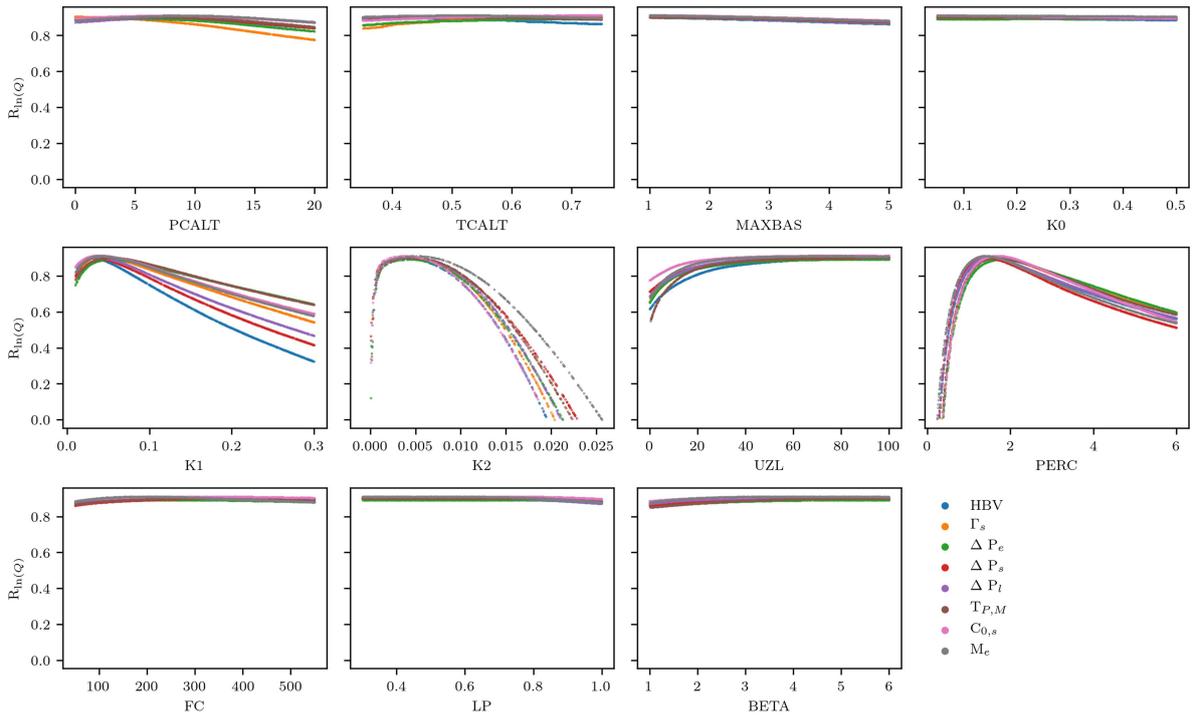
CH-113



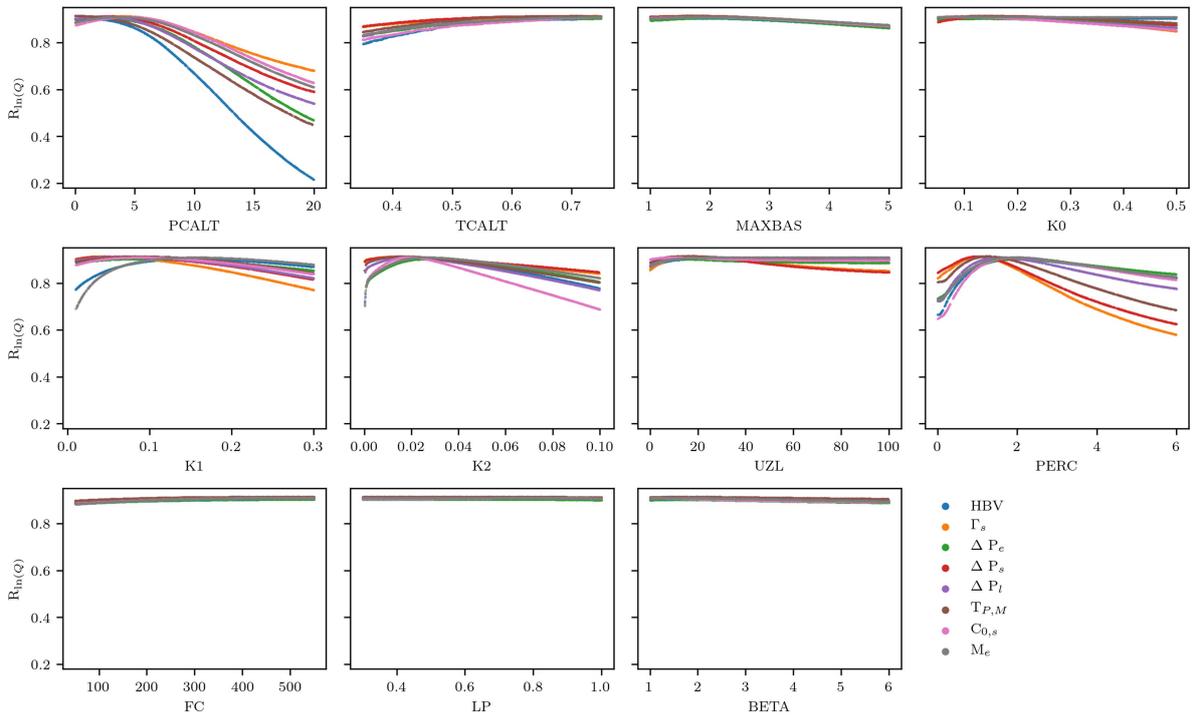
CH-201



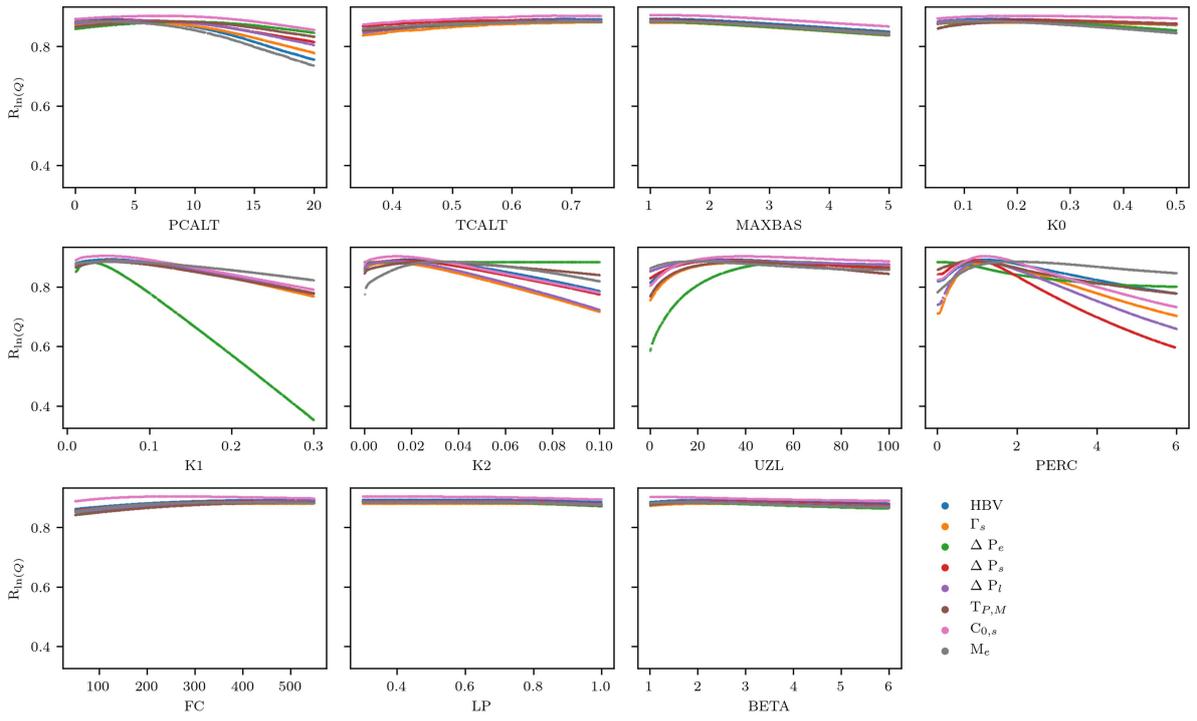
CH-202



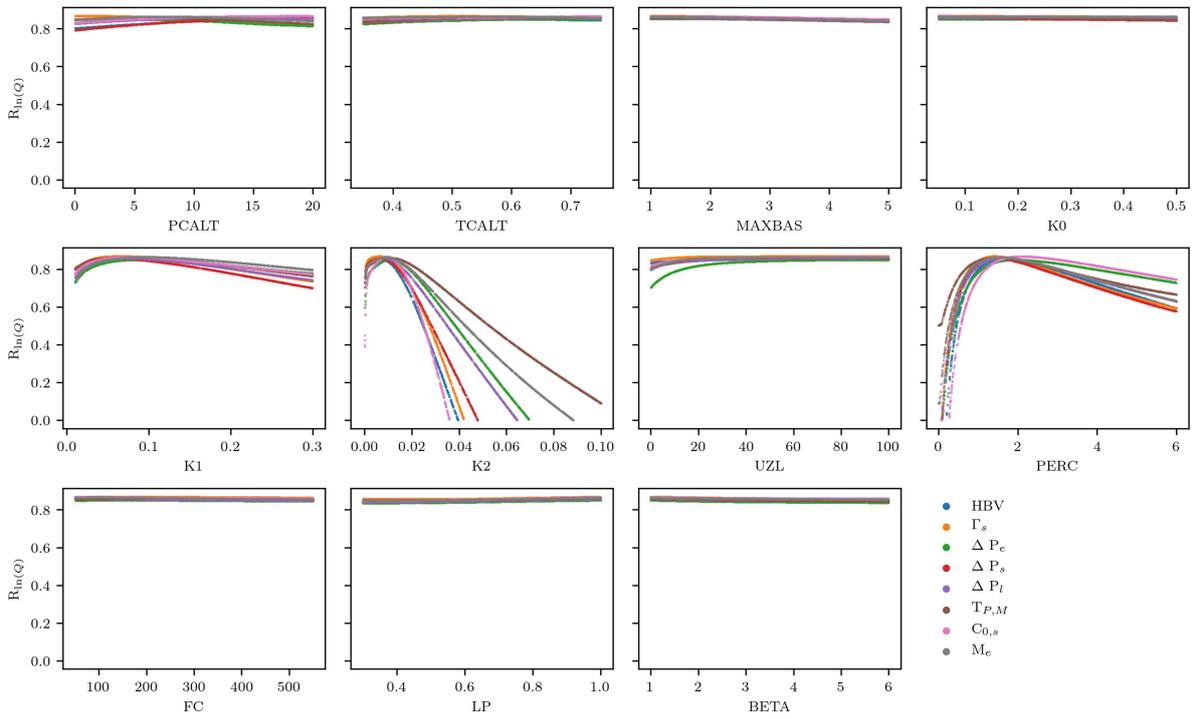
CH-203



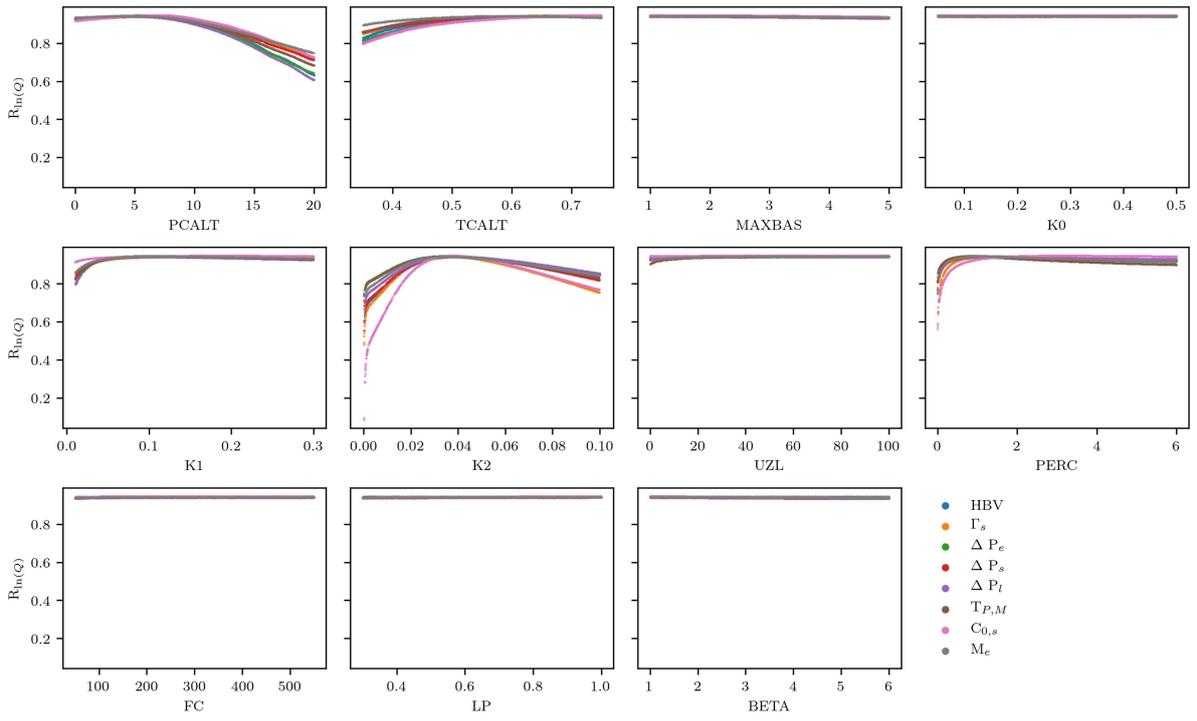
CH-204



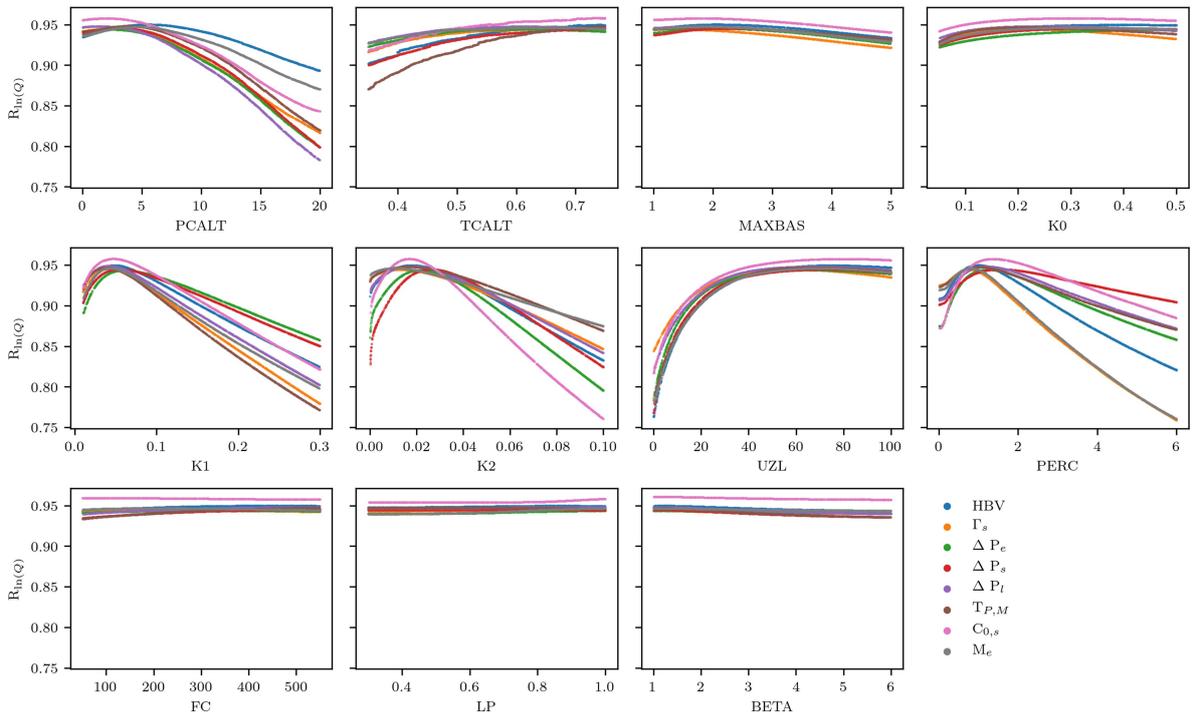
CH-205



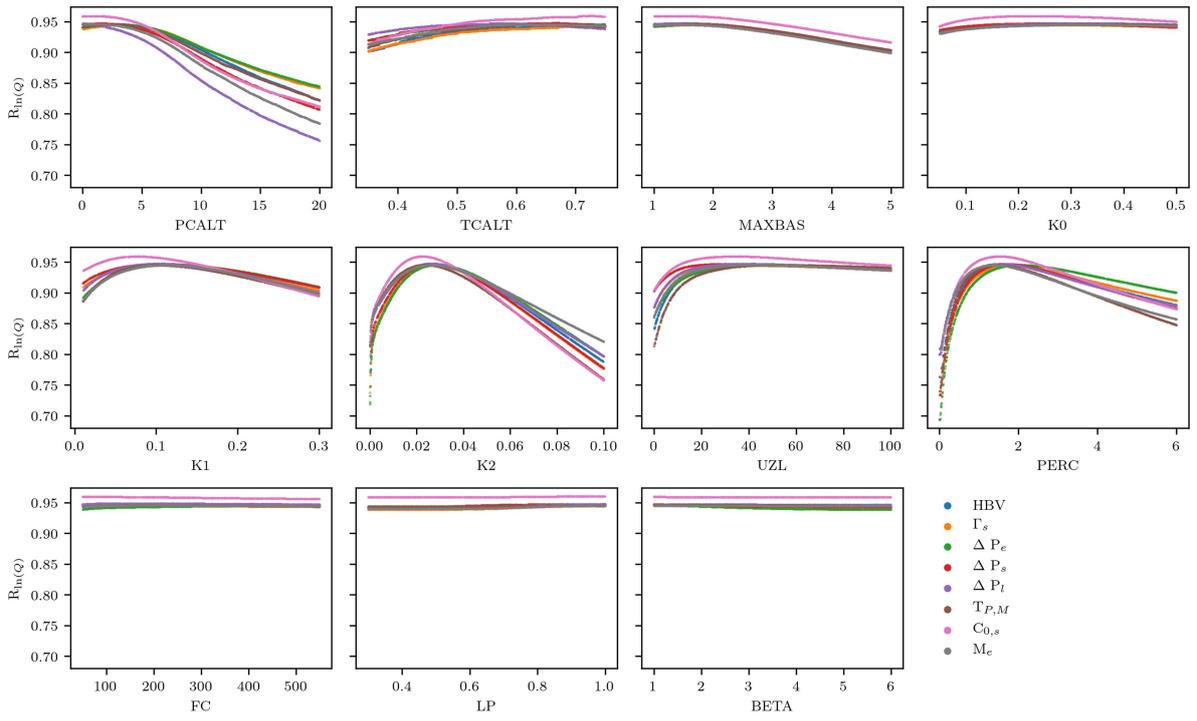
CH-206



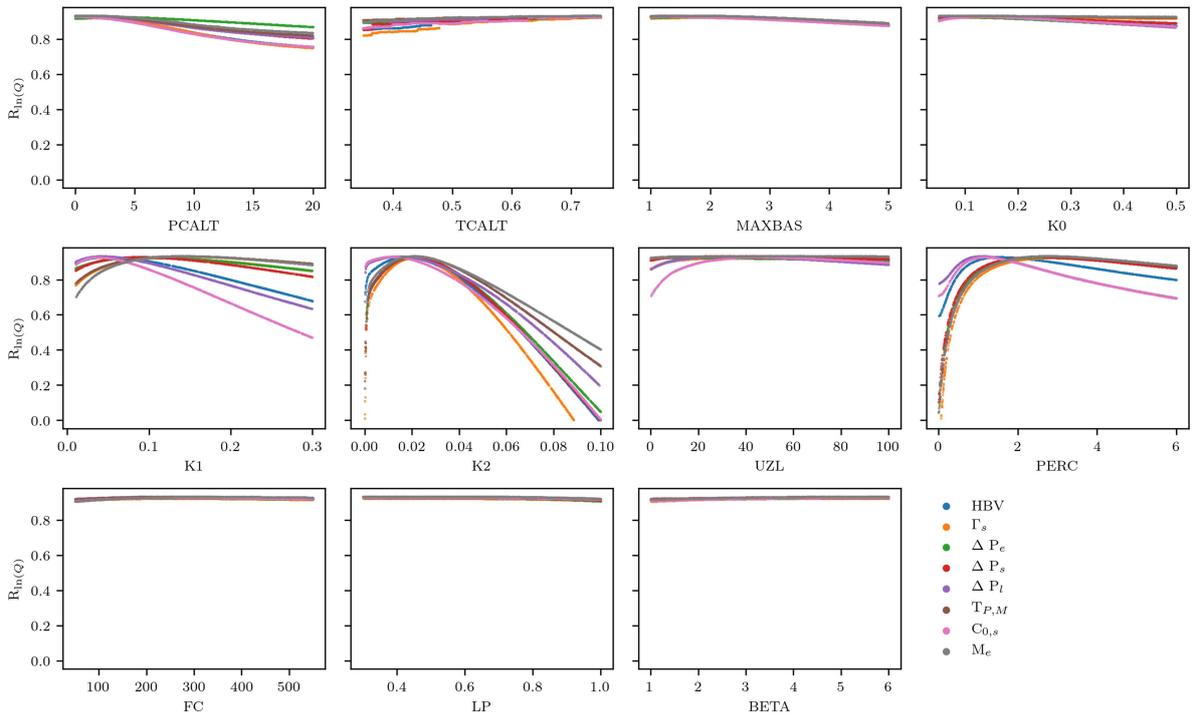
CH-301



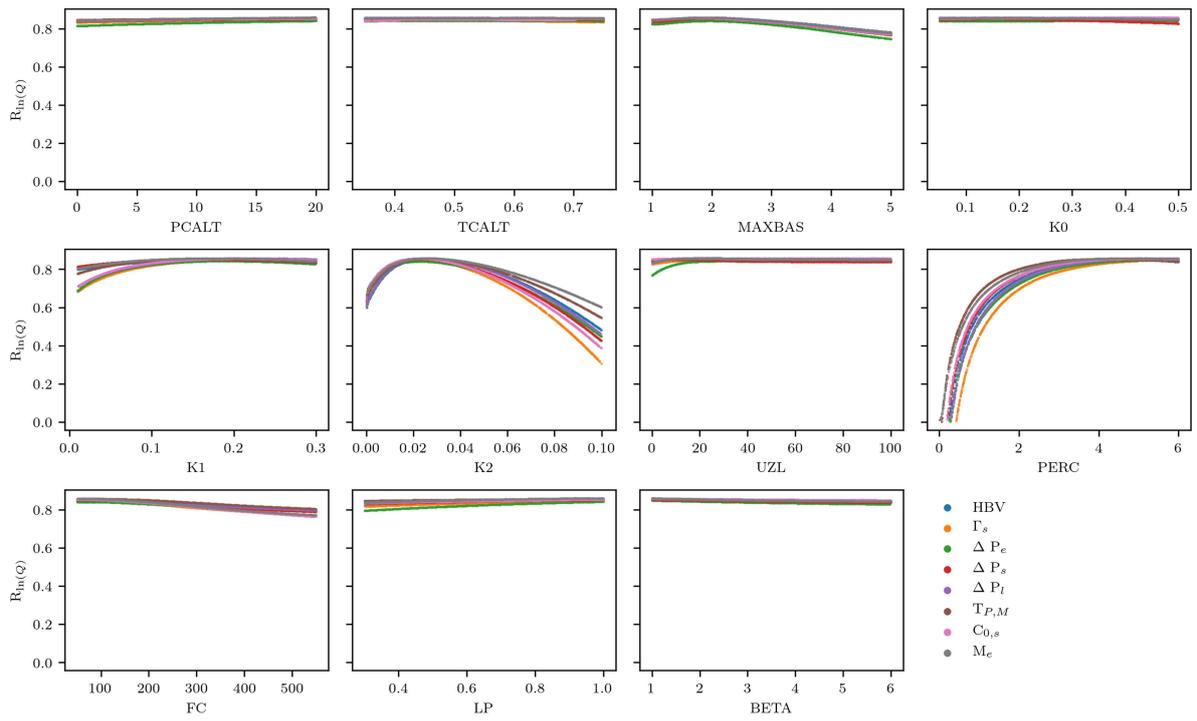
CH-302



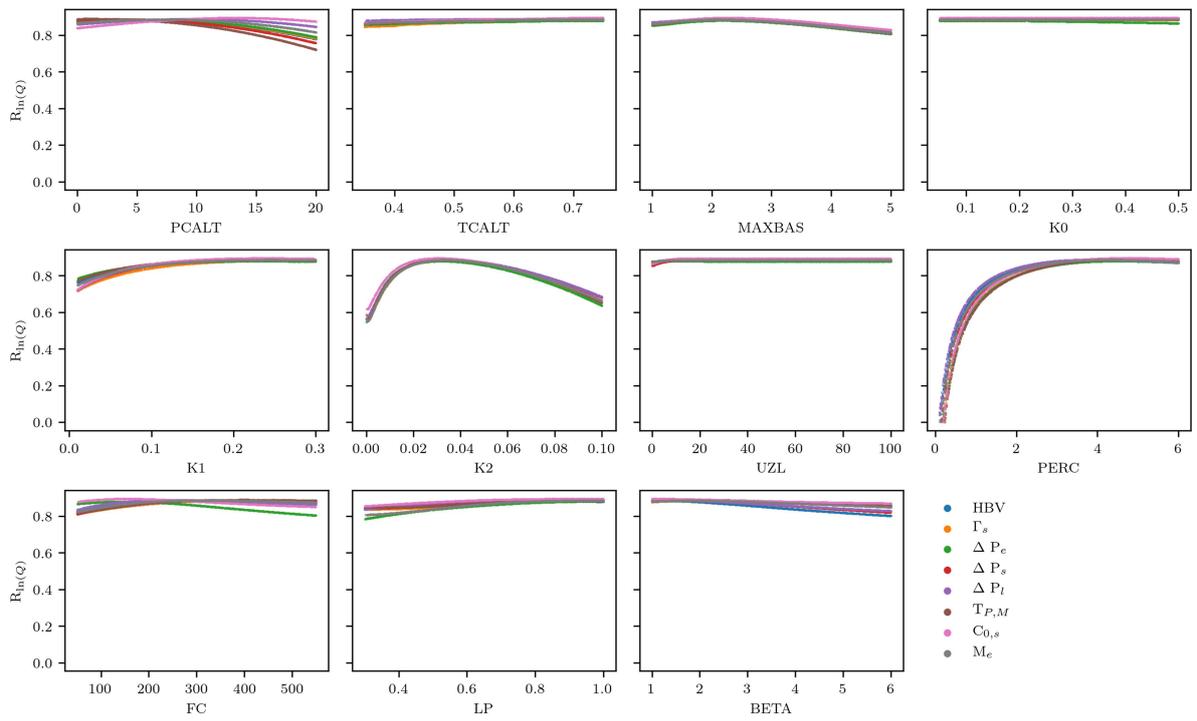
CH-303



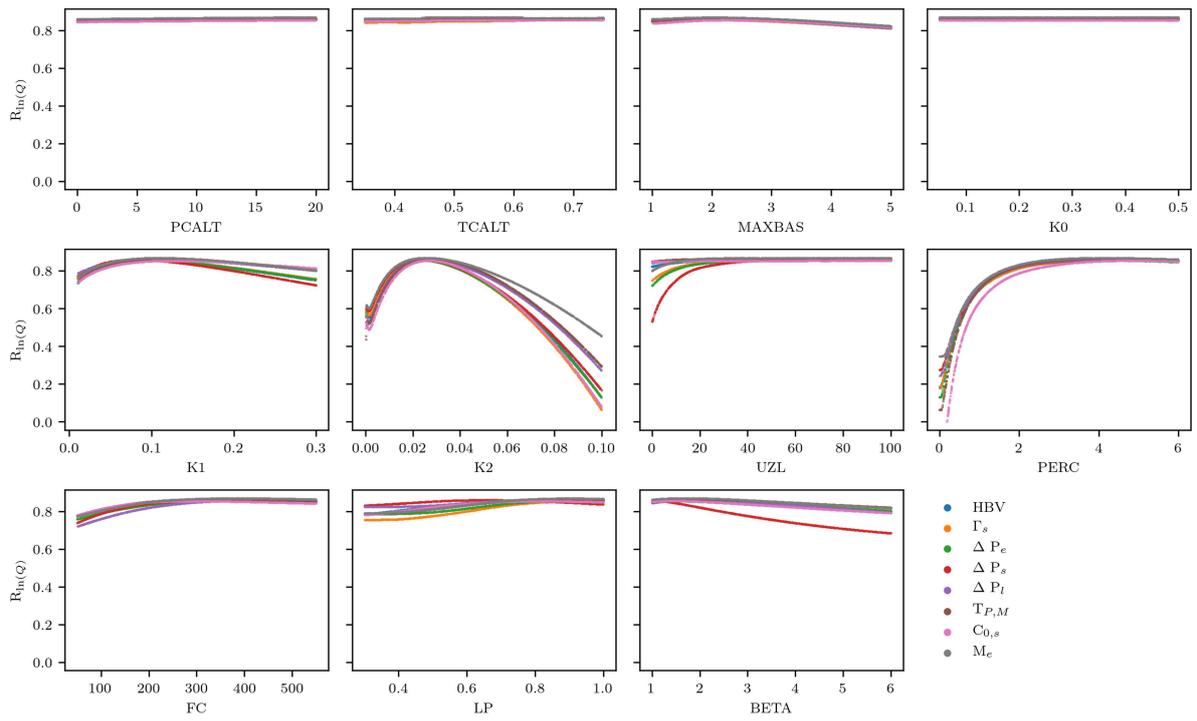
CZ-101



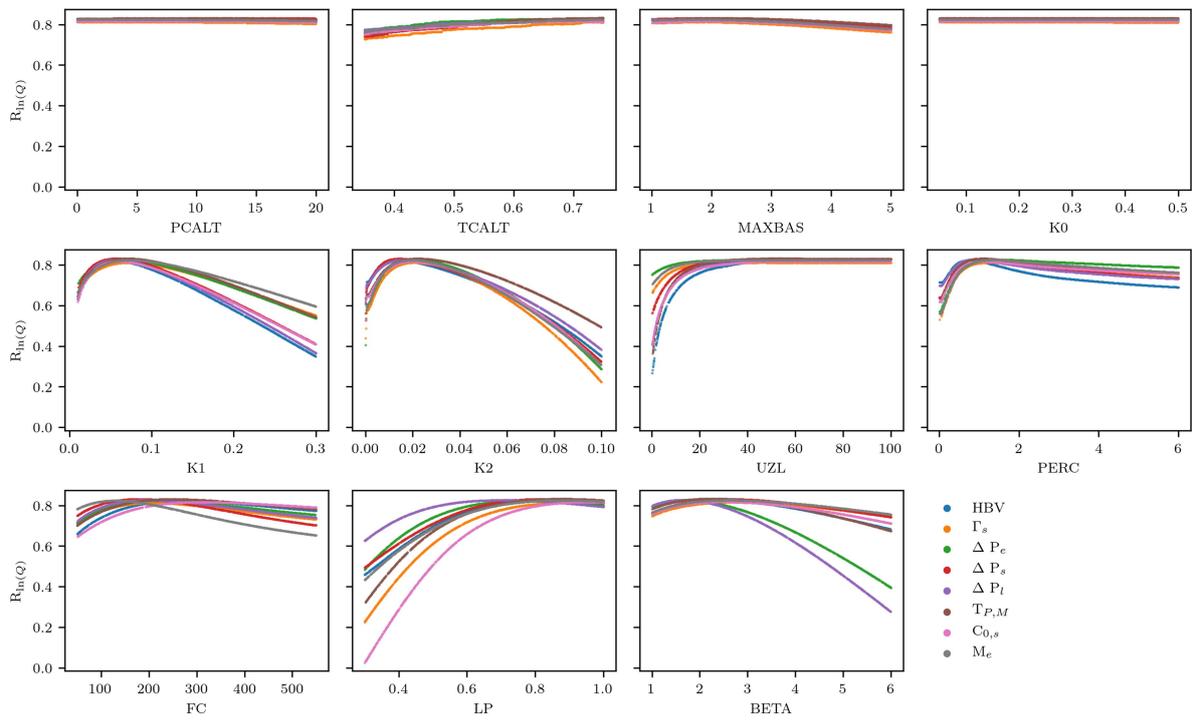
CZ-102



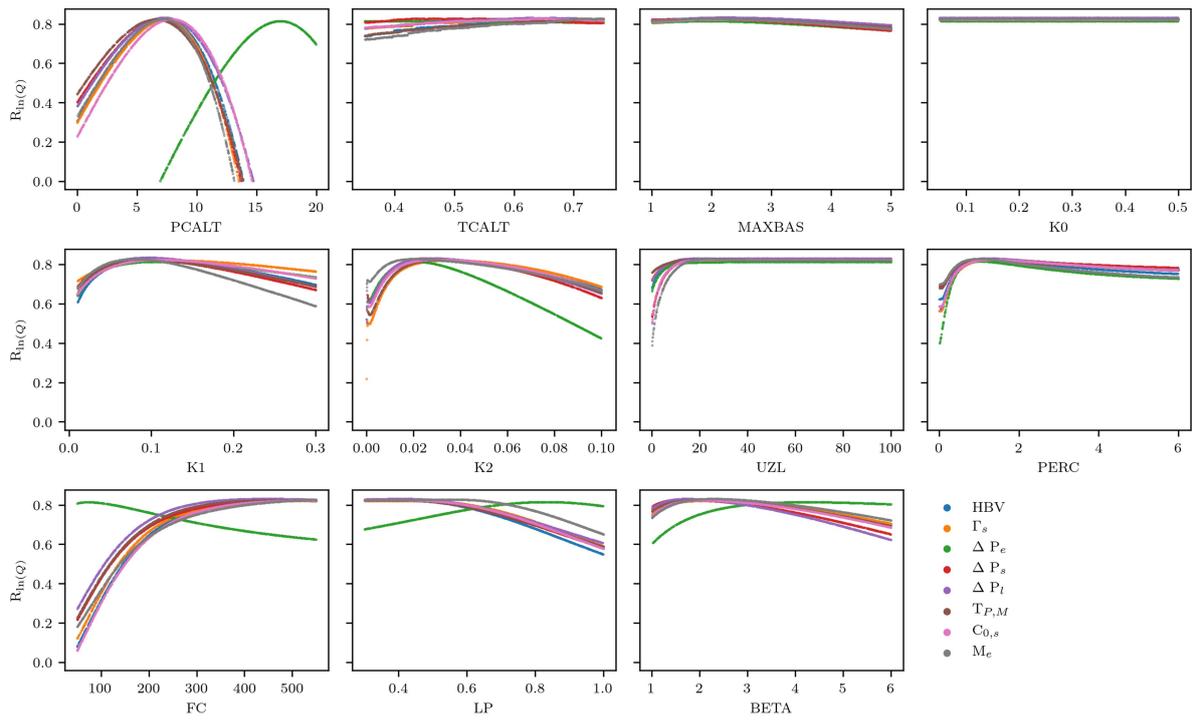
CZ-103



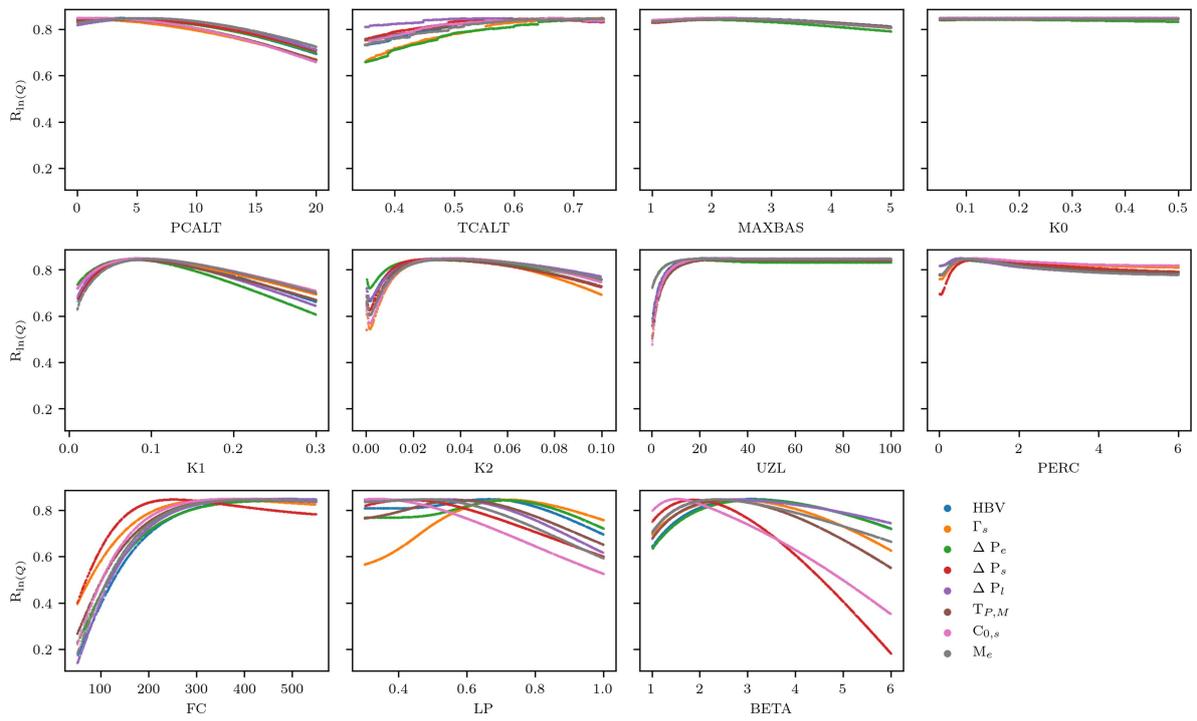
CZ-104



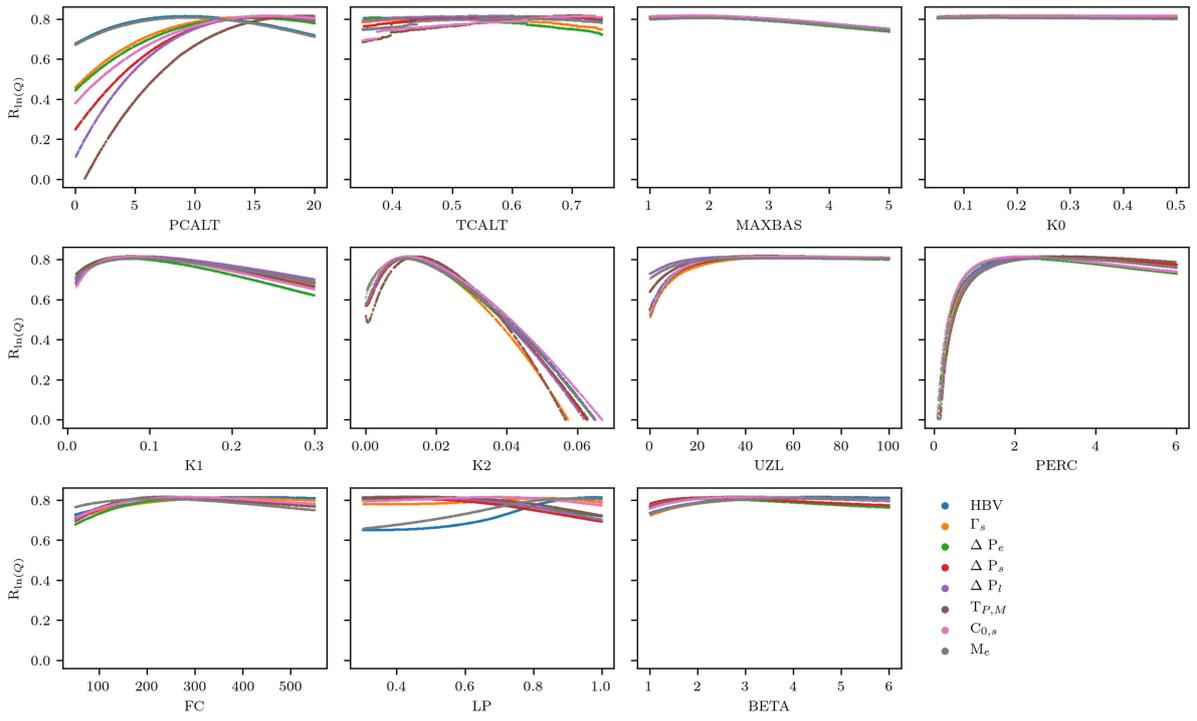
CZ-105



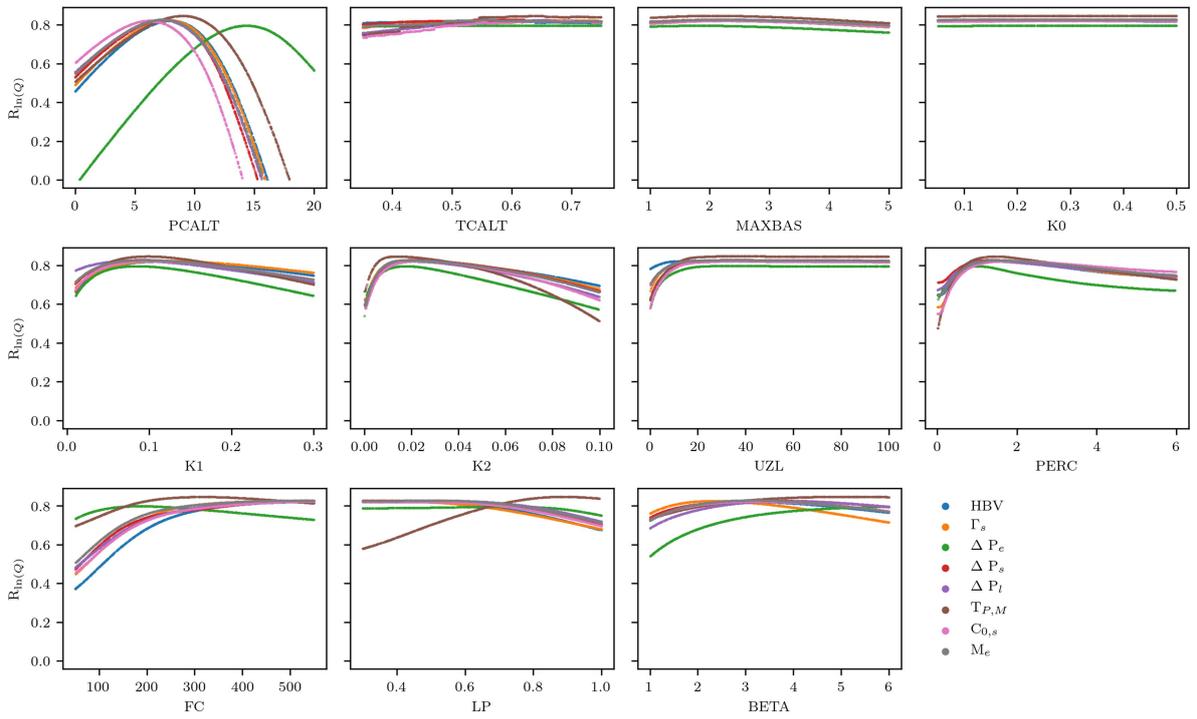
CZ-106



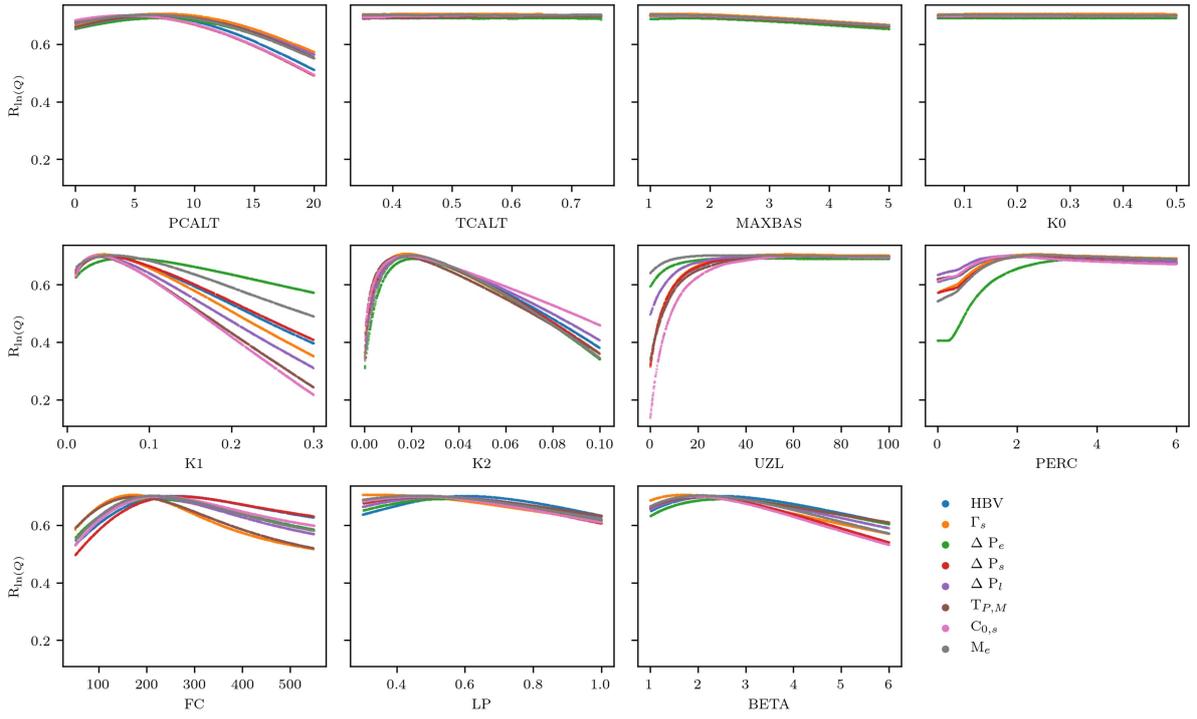
CZ-107



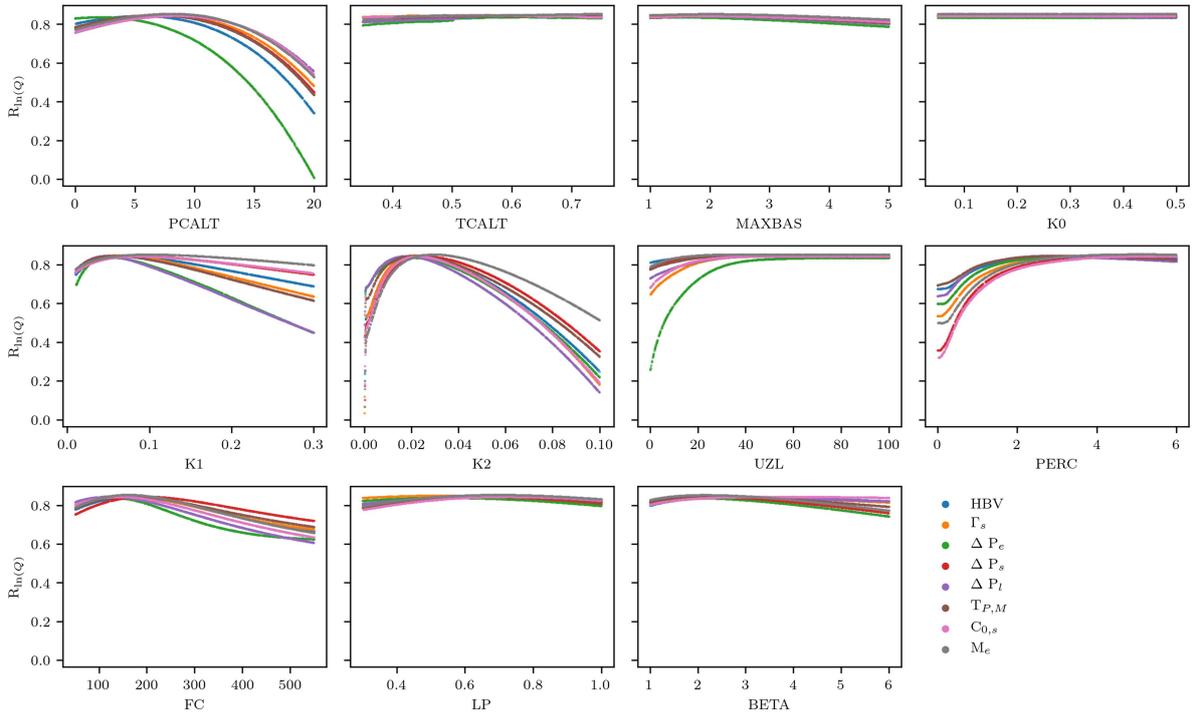
CZ-201



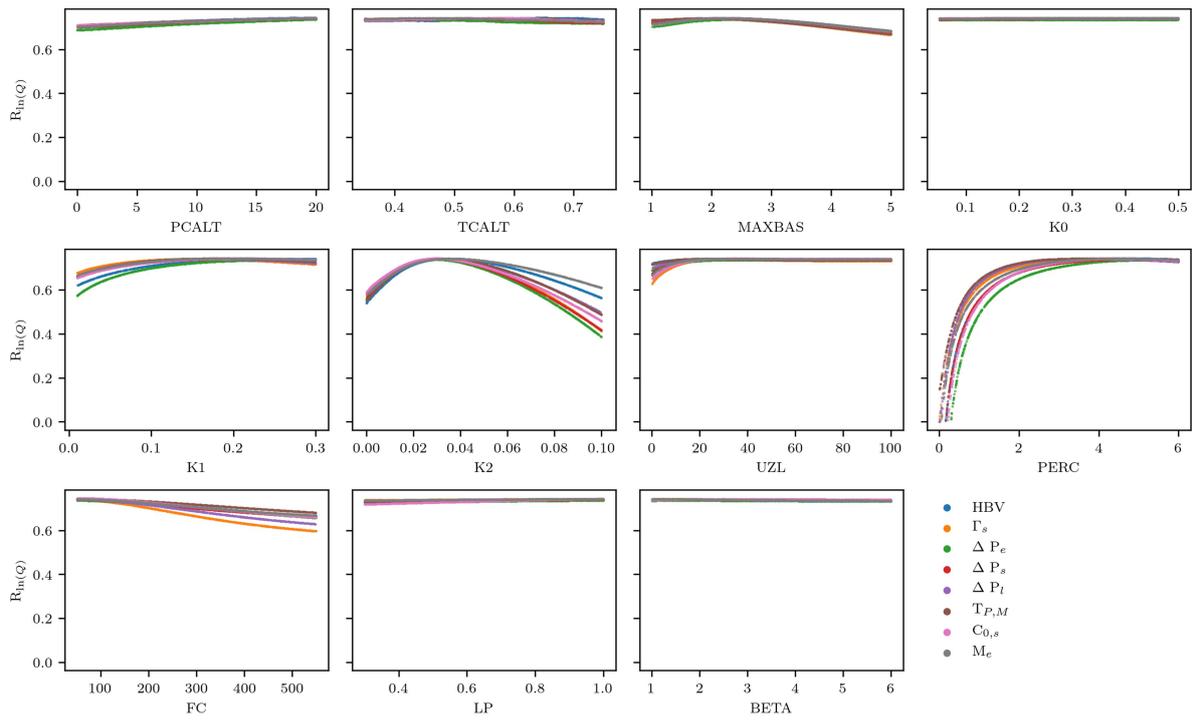
CZ-202



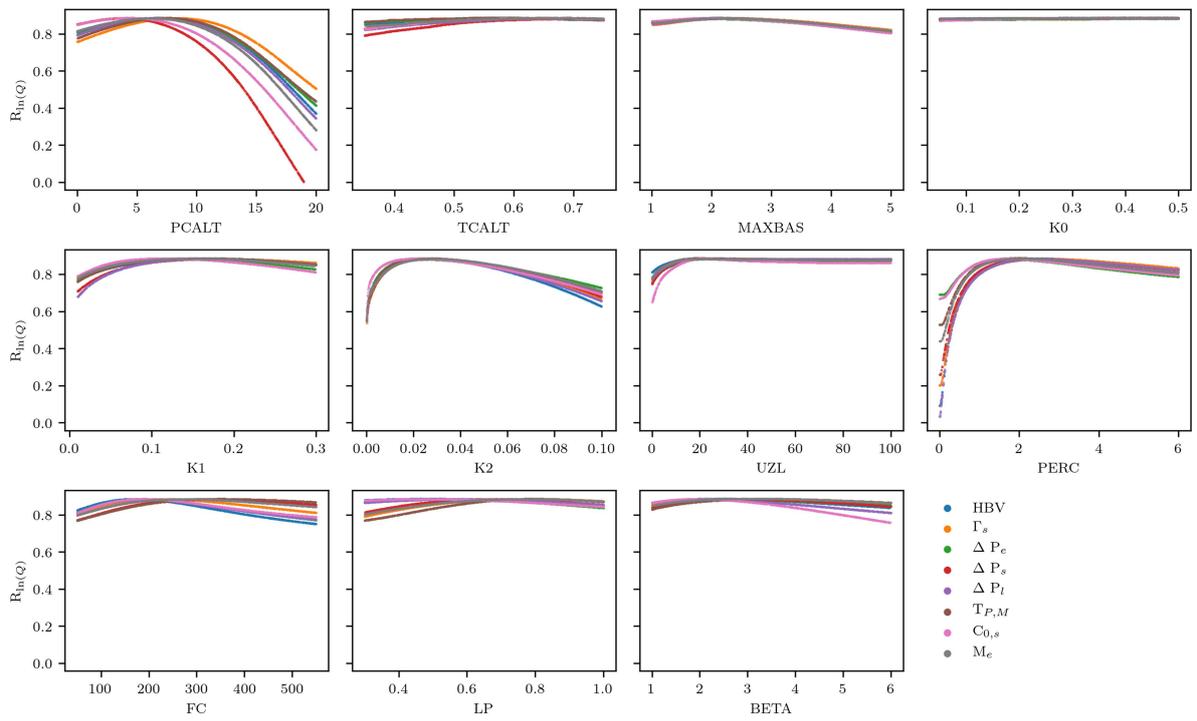
CZ-203



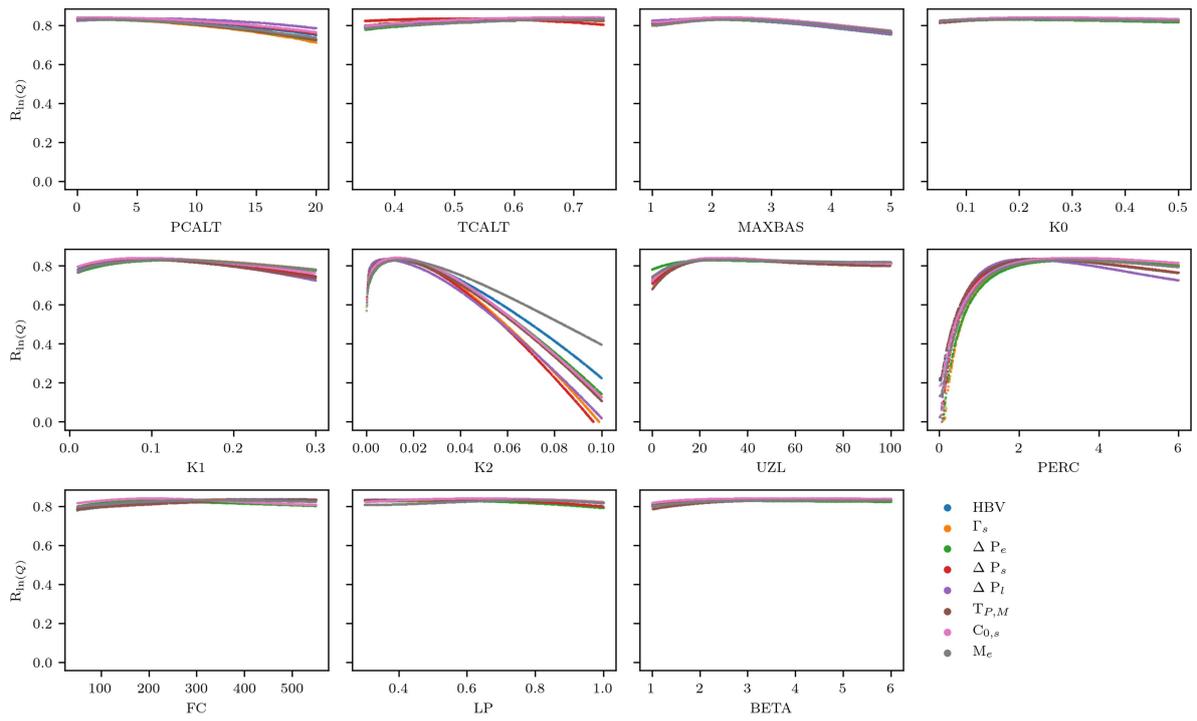
CZ-204



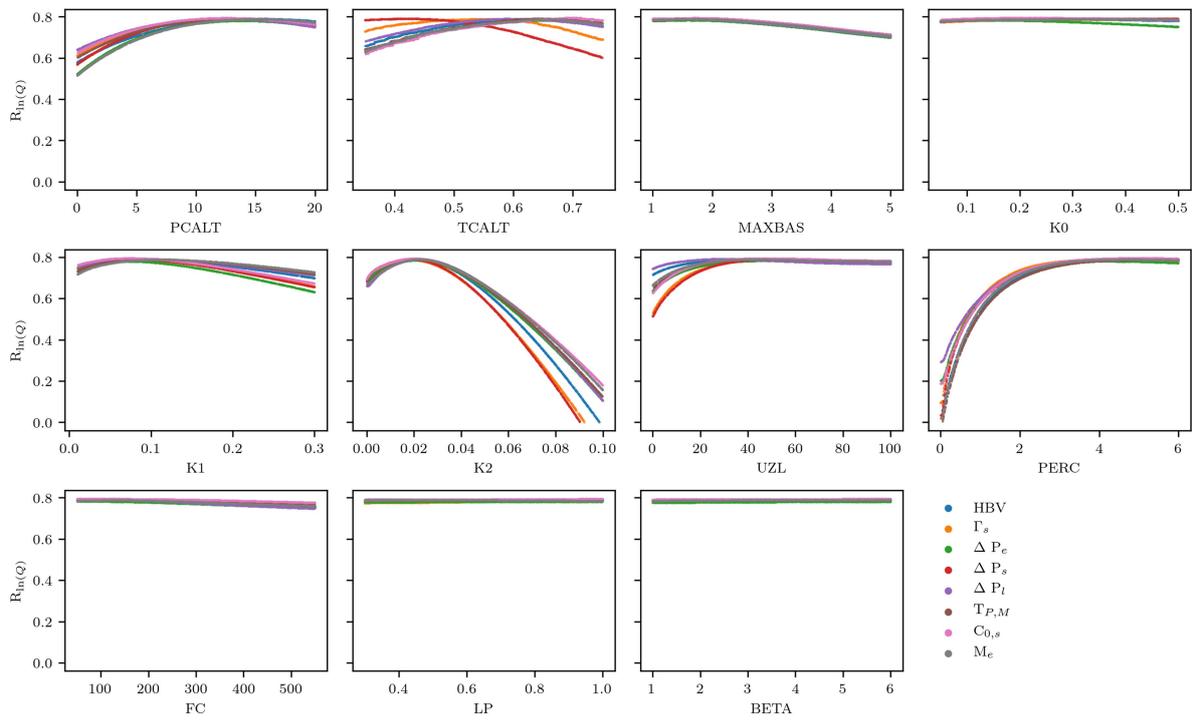
CZ-205



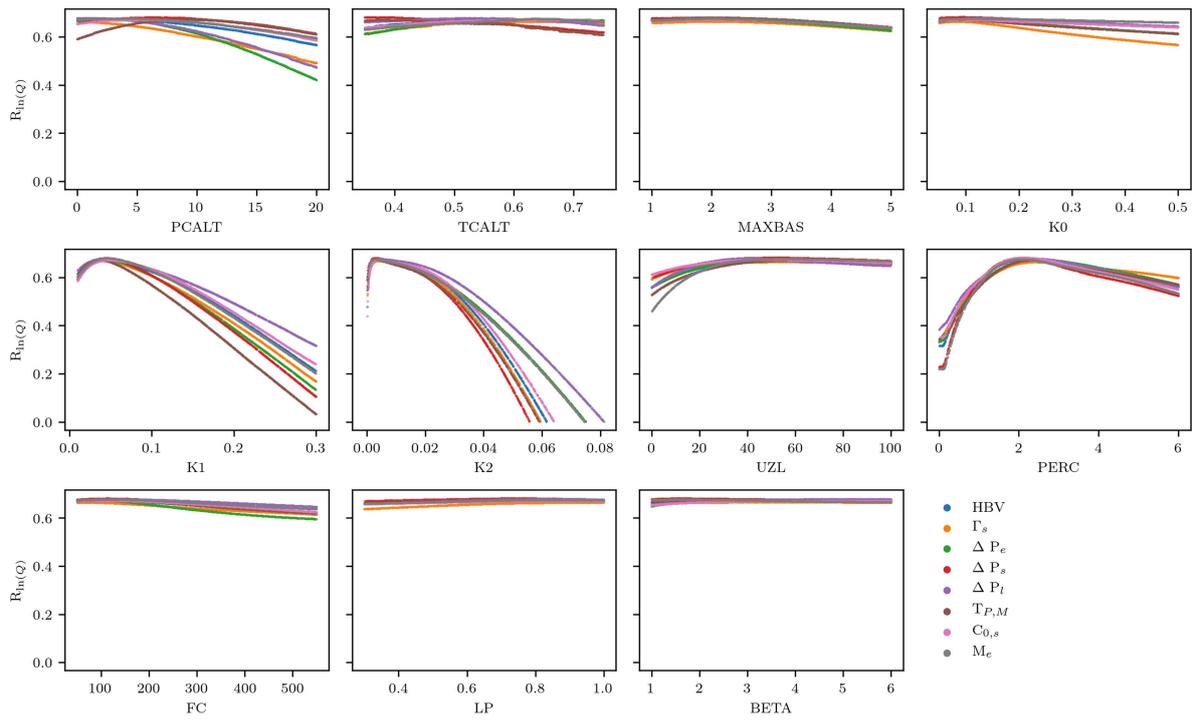
CZ-206



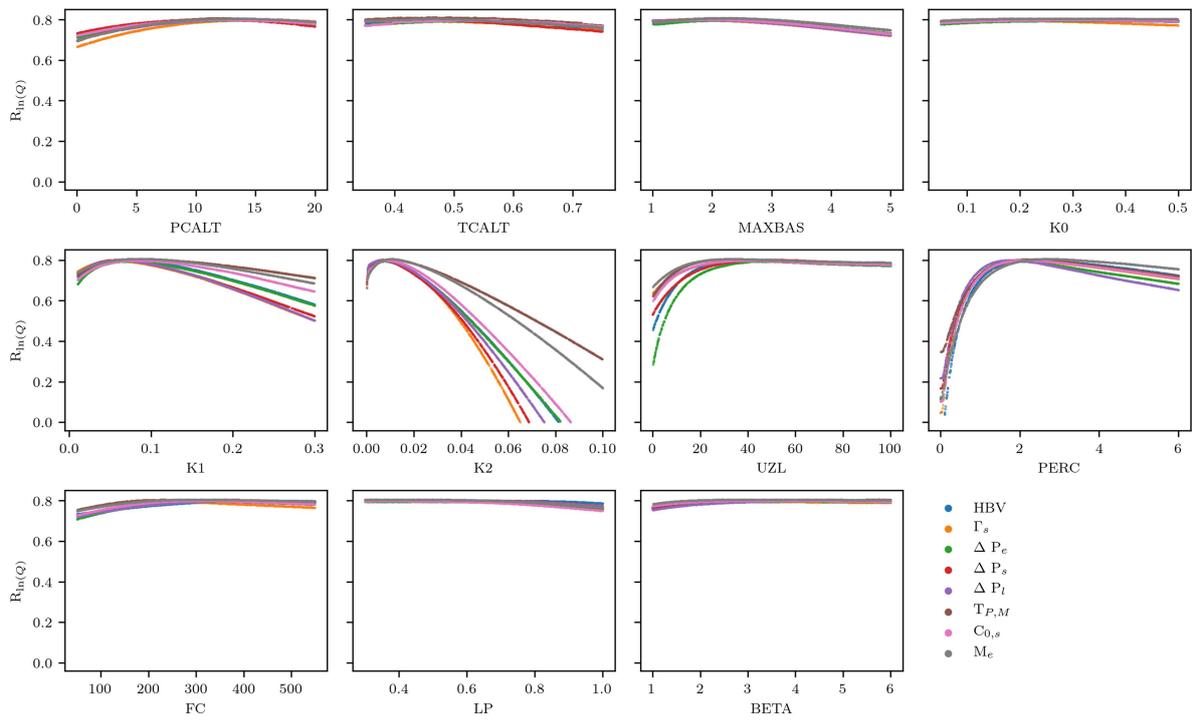
CZ-207



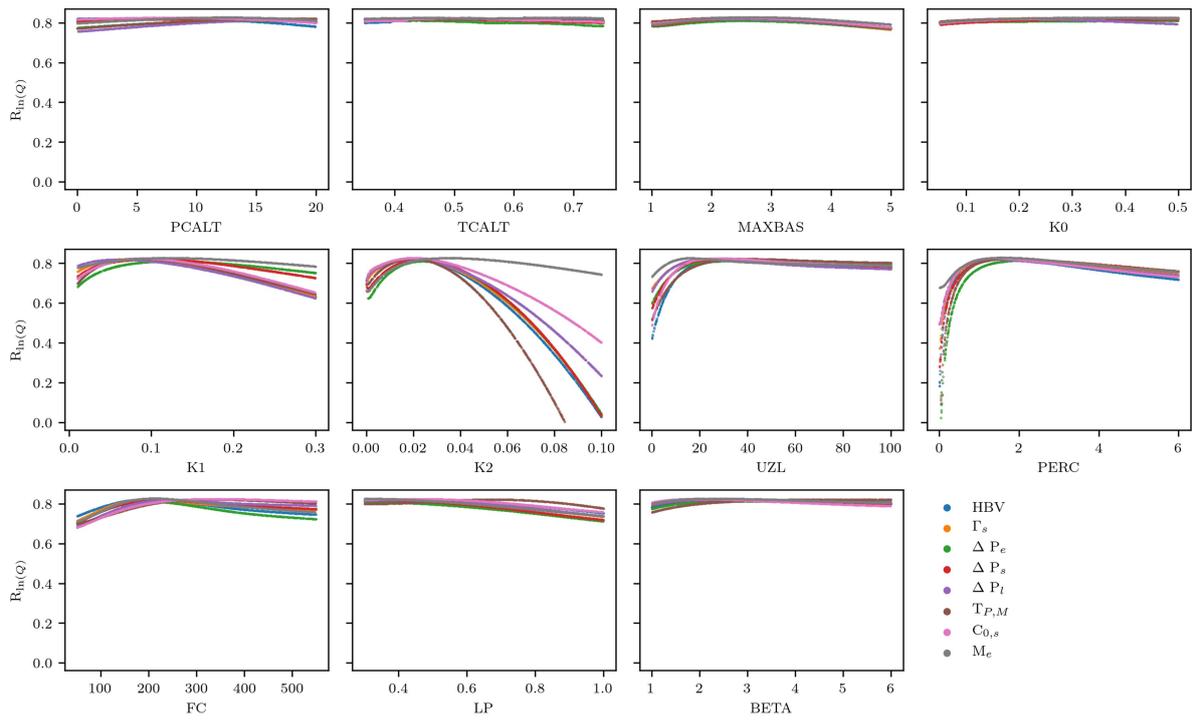
CZ-208



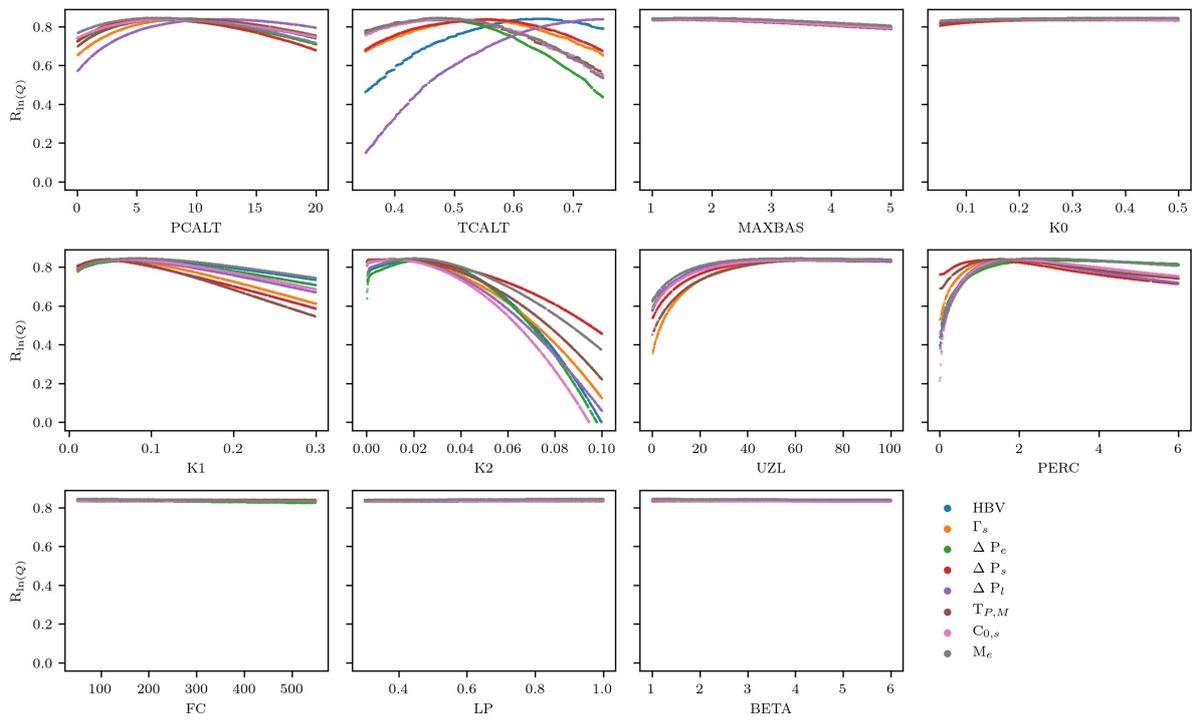
CZ-209



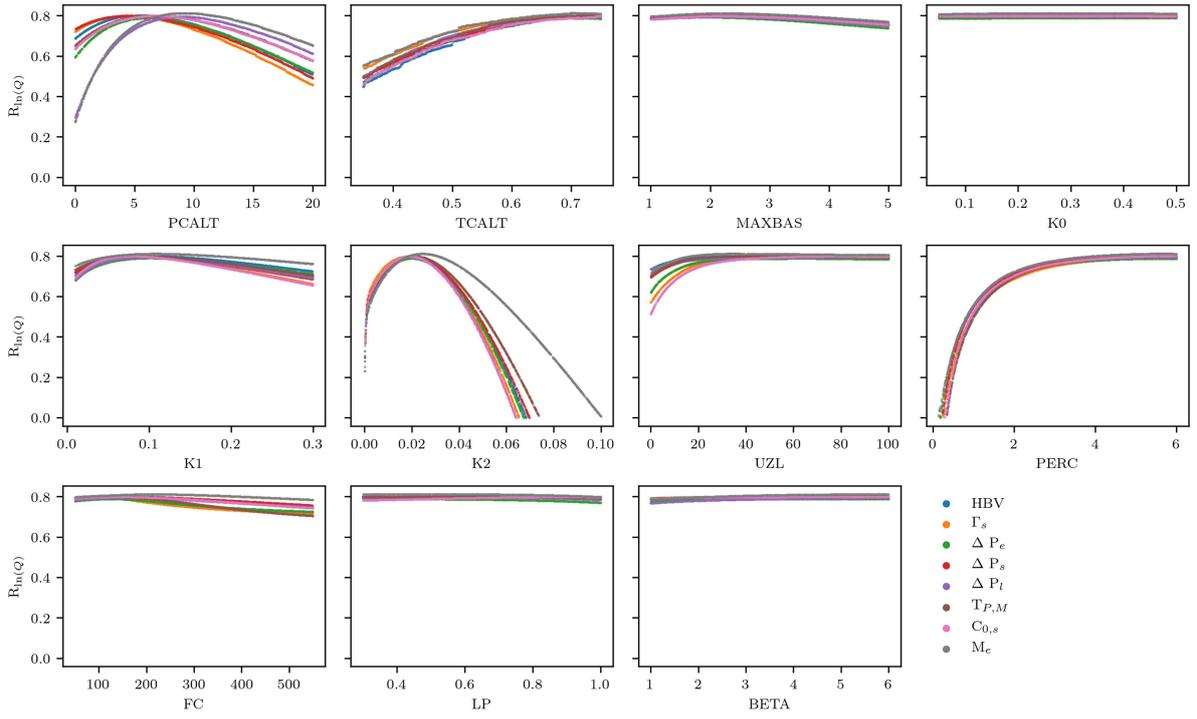
CZ-210



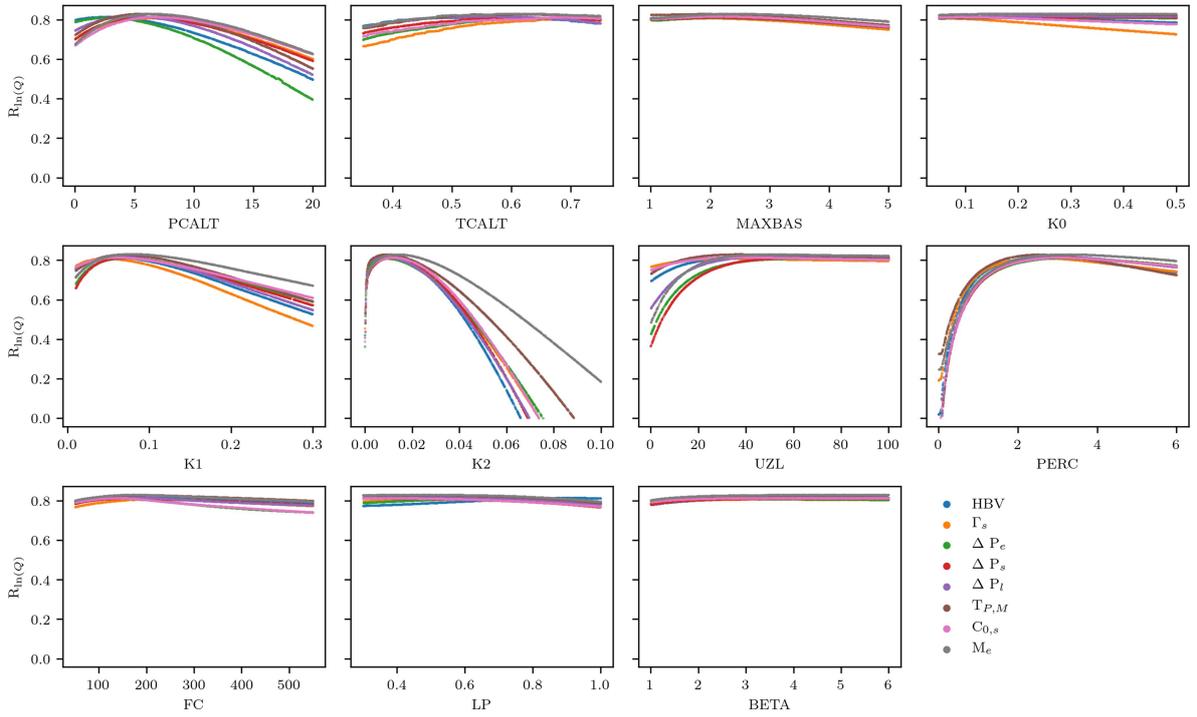
CZ-211



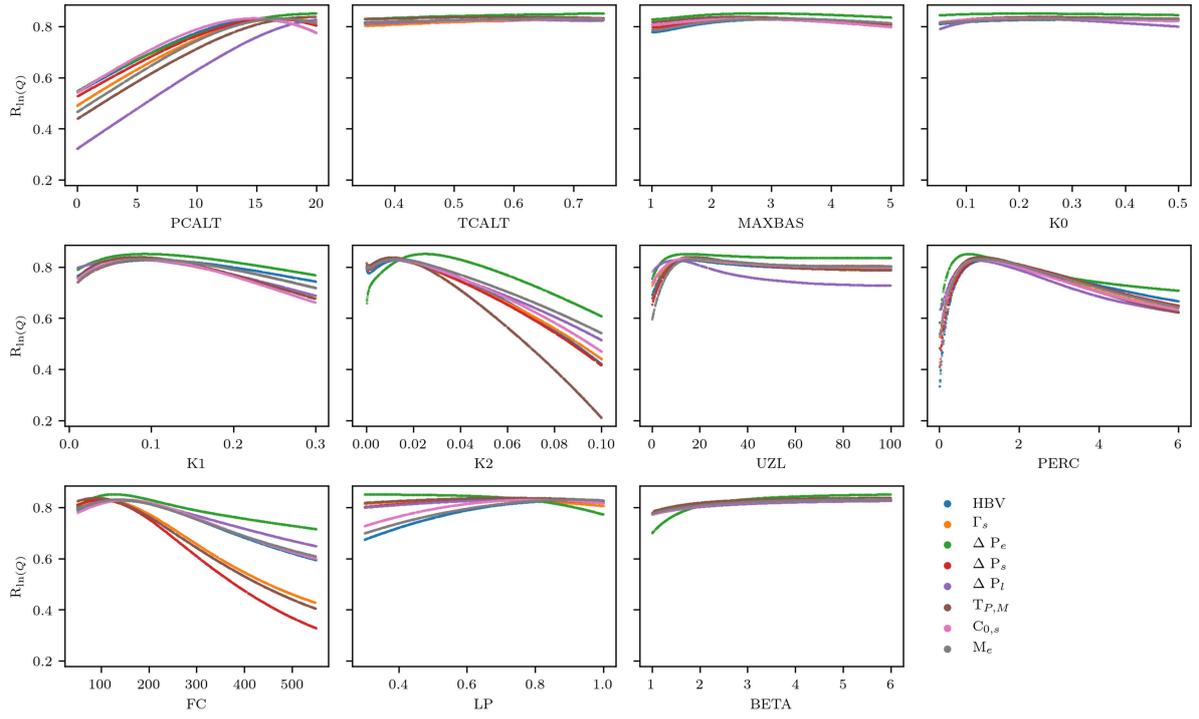
CZ-212



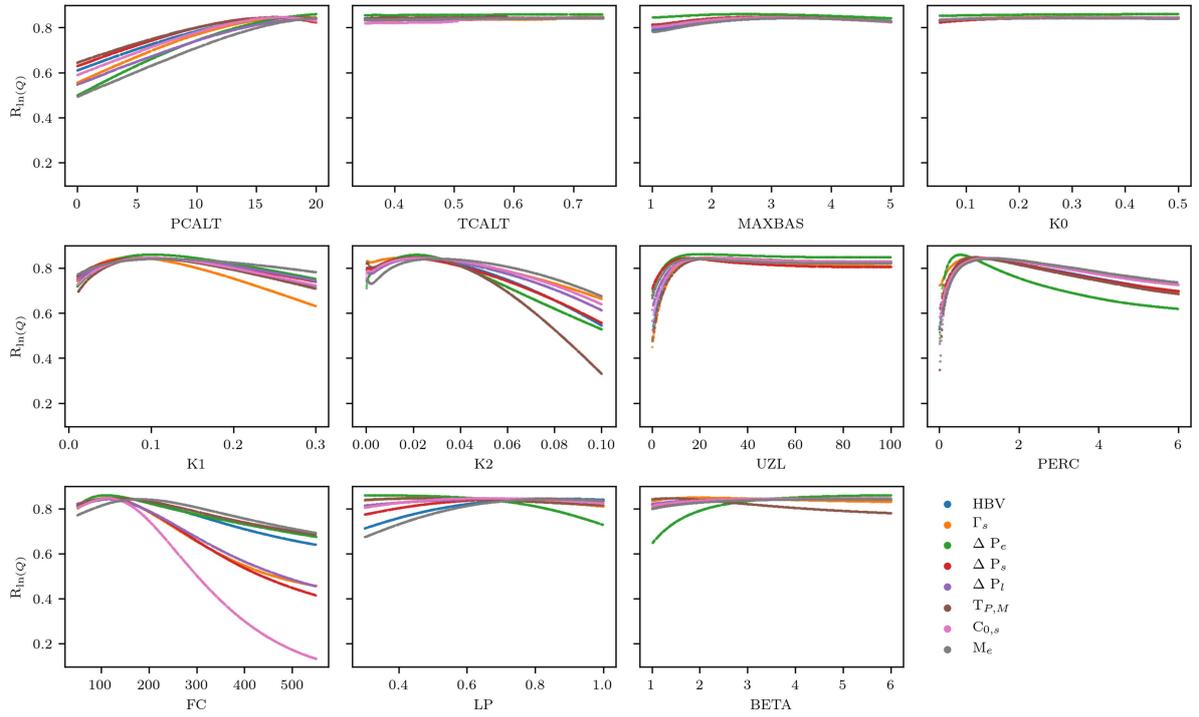
CZ-213



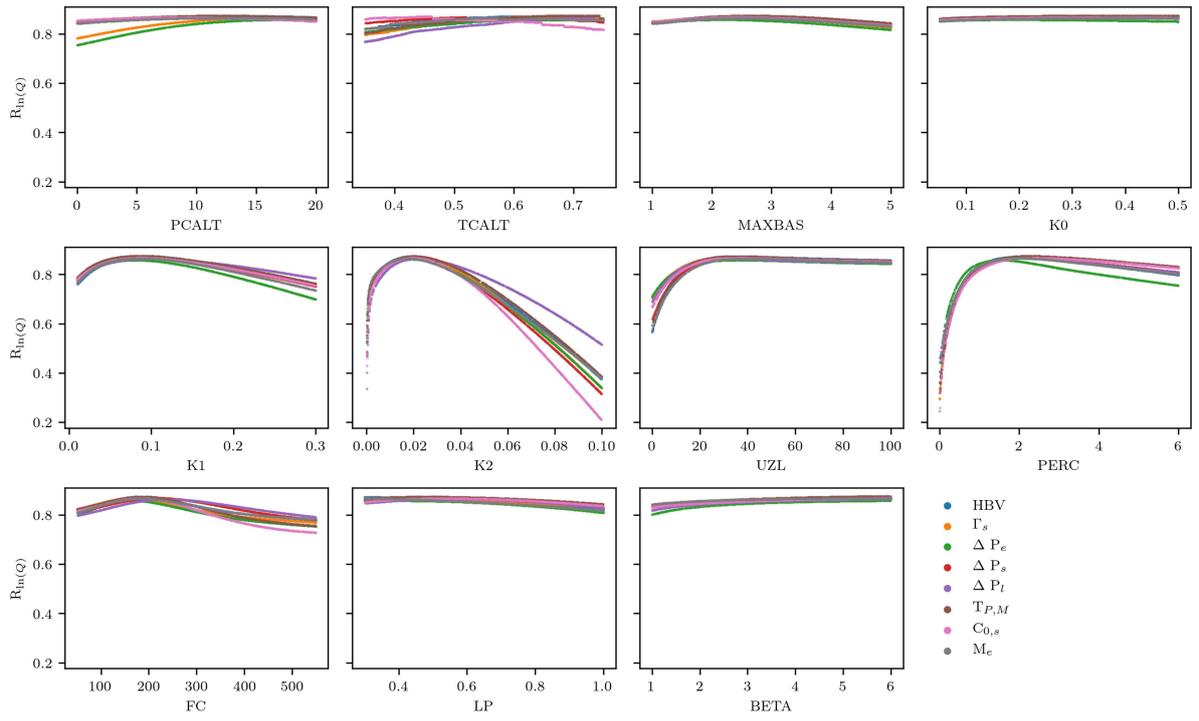
CZ-301



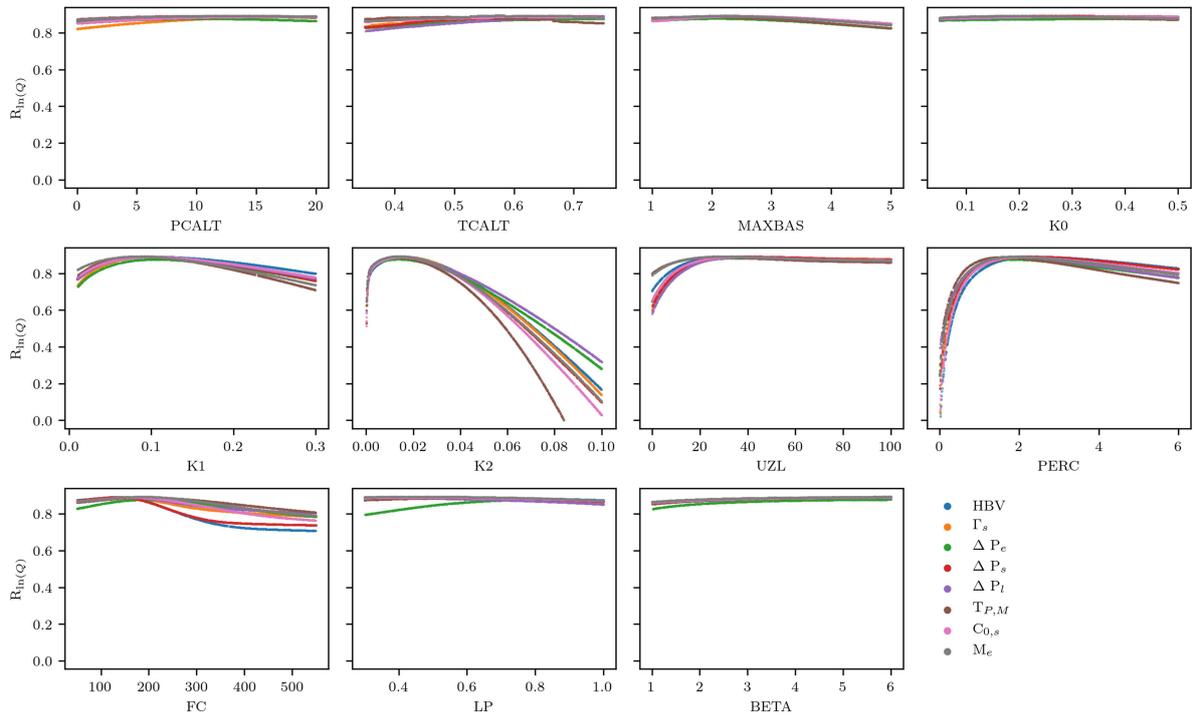
CZ-302



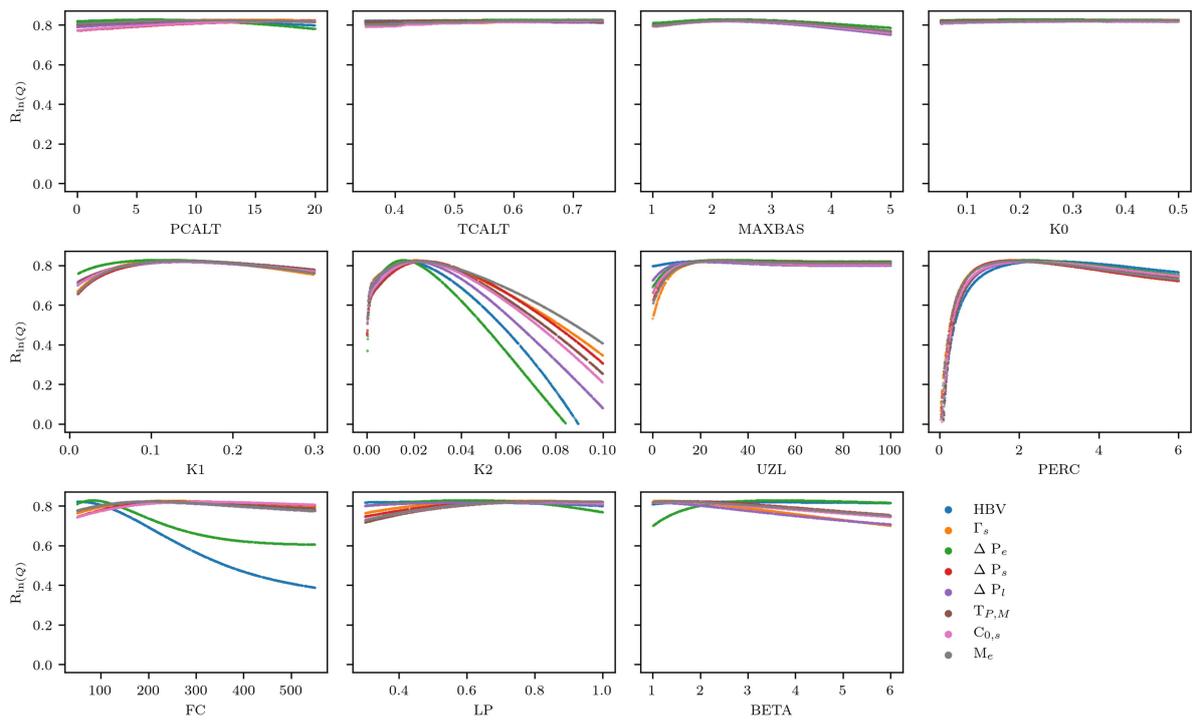
CZ-303



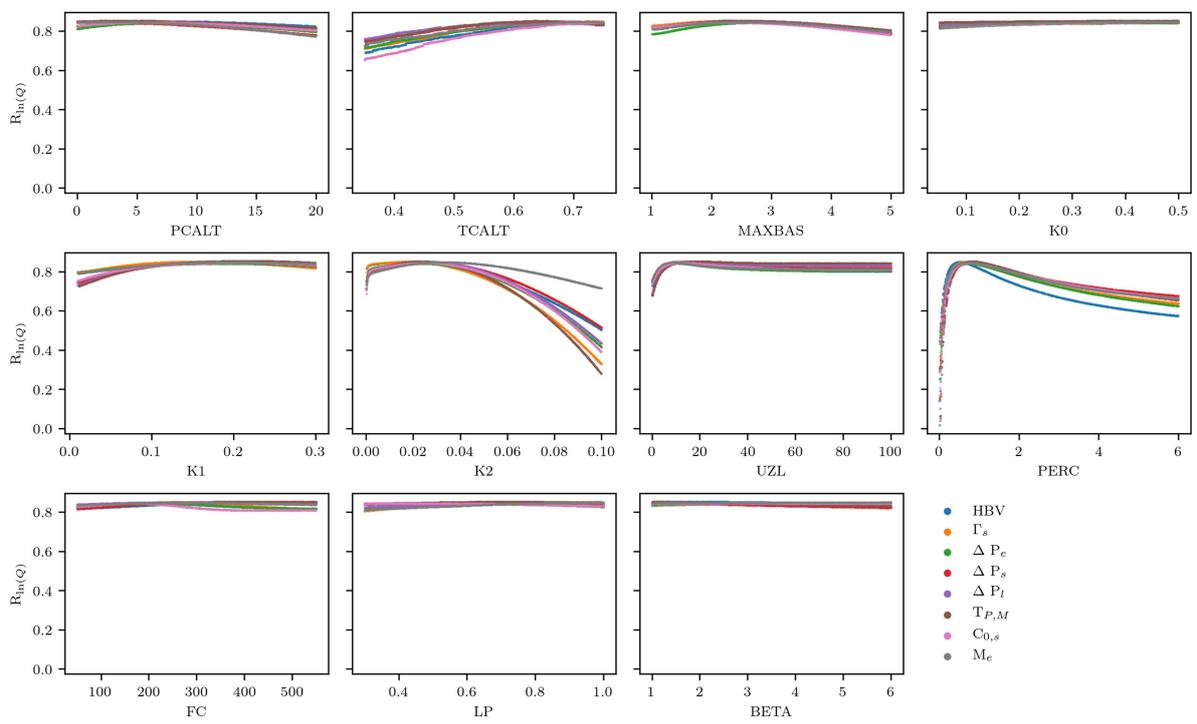
CZ-304



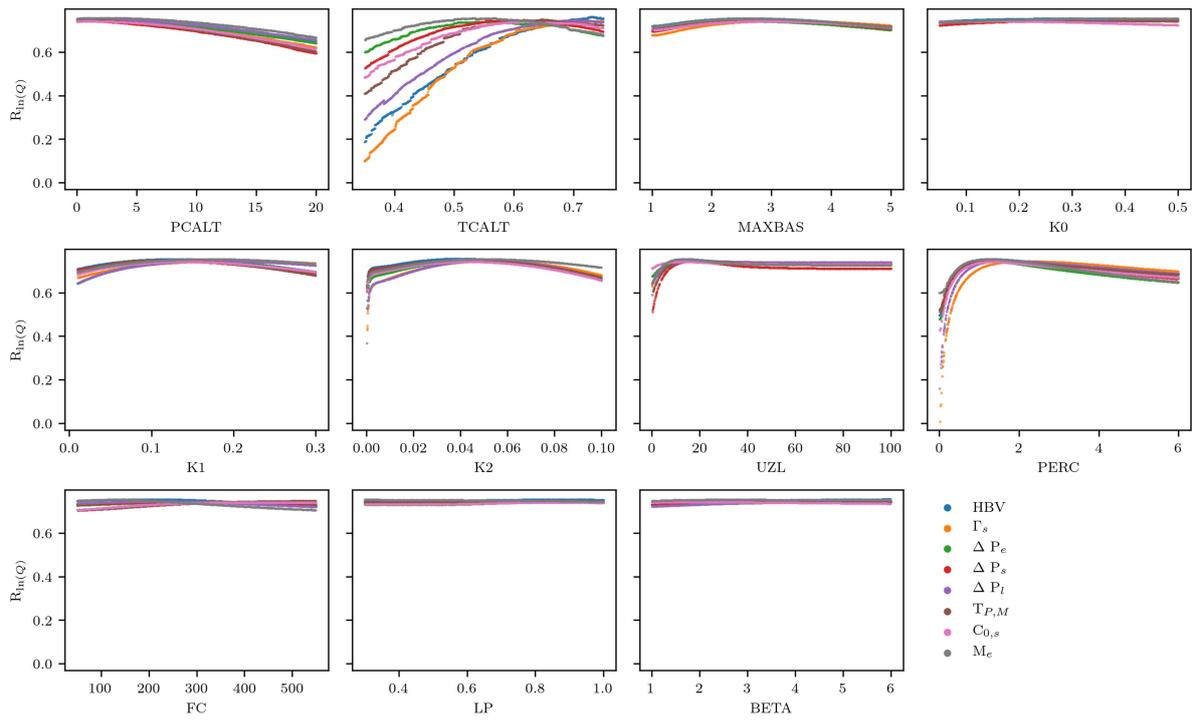
CZ-305



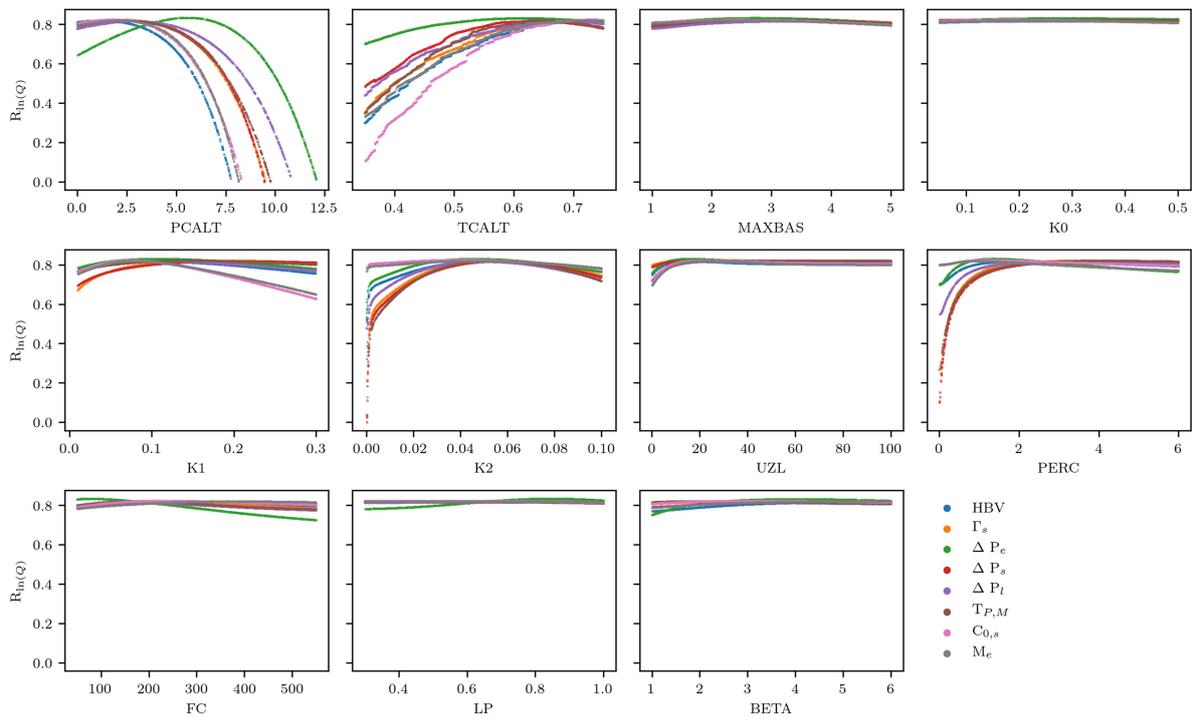
CZ-401



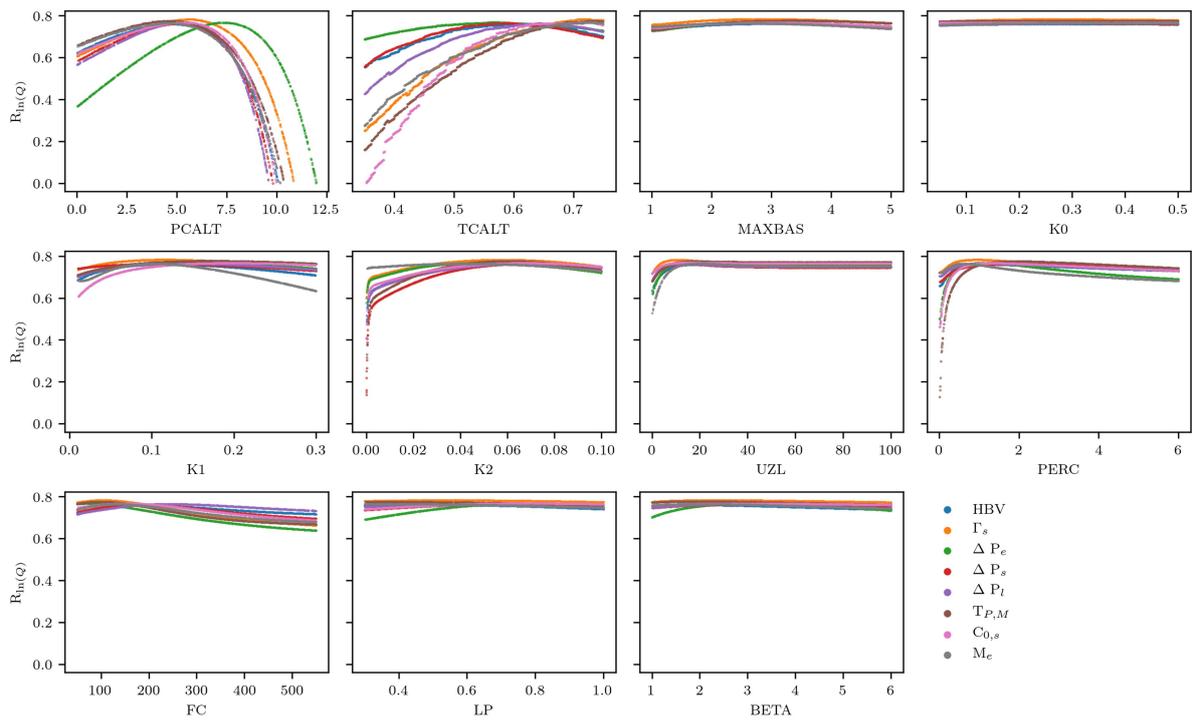
CZ-402



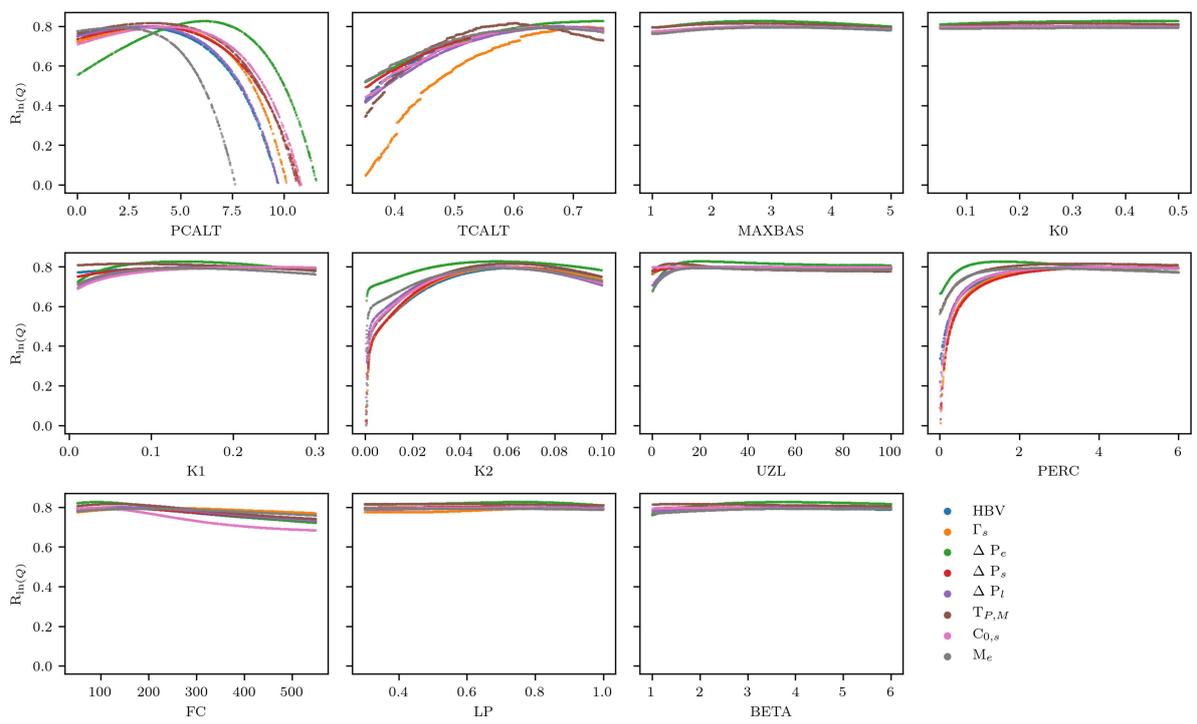
CZ-403



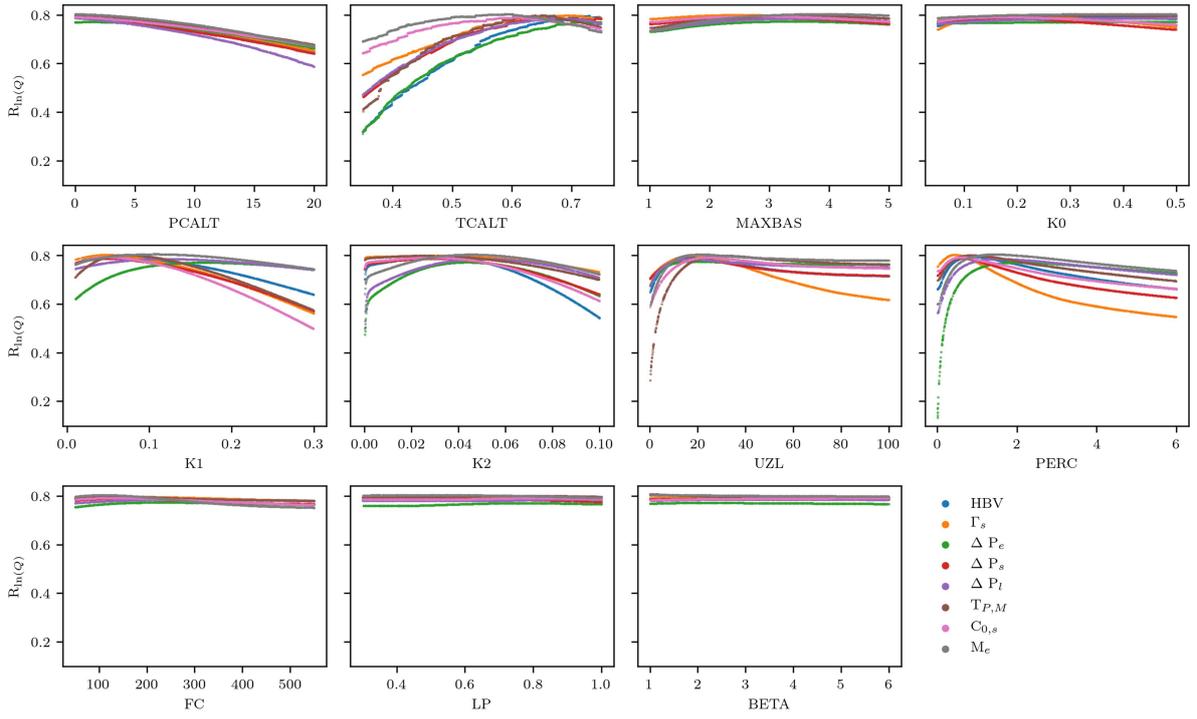
CZ-404



CZ-405



CZ-406



CZ-407

