

To me it seems the main reason you want these slots is just to estimate live performance, and telling us that this need increases with the amount of changes in the tournament.

To estimate live performance we have always had the trouble that we didn't have recent data available, luckily this should soon change with their new planned data release. This should allow us to properly evaluate CORR and FNC(v3) ourselves. Then we still have the problem of not being able to estimate MMC, TC and Corr with meta model ourselves.

Since all data is released anyways, it might make more sense to add this to the validation diagnostics as well. Maybe the computations of TC is a bit more costly to allow unlimited, then either a cap of number of daily evaluations can be added or maybe you can submit some number of predictions each day/hour that then can be processed in batch reducing the computational power required. This seems to me to be a nicer way to evaluate your models then having to wait weeks or months before getting an idea of your new model's TC score.