I hear you and we will revisit related metrics at some point.

But to address your points, TC is in fact not a noisy estimate. TC is calculated by dropping out 1/2 of the stakes 100 times to create 100 portfolios and then calculating the gradient of returns with respect to stakes for all 100 portfolios and then averaging those 100 gradients. Repeating this with a different random seed gives results over 99% correlated.

What is

noisy is stock returns. Their heavy tails effect real life stock returns (and therefore TC) much more than you might guess from the ranked correlation of predictions and stock returns.

Also, it would require pretty substantial stake moves by others to really effect your TC. It's not something I'd worry about or try to attribute on a round to round basis. But if your TC is low for a while you might want to do something more original.

It's basically impossible to evaluate a model based on one round. My advice is to trust your CV and let it ride. I was definitely guilty of lots of switching up models round to round, but that was before there were lots of model slots. So use them!

TC as currently calculated doesn't make sense for only one stake, but we could show your portfolio return for the portfolio created if you were the only one in the metamodel. We experimented with this originally, but iirc it wasn't that helpful for predicting TC as it doesn't capture at all the interaction with everyone else. But I will definitely reevaluate when I revisit TC.