Agree with what you said above. Handling Eras is also something I am still quiet confused about.

I think what most diqualifies the era coloumn as real temporal data is that live data does not reveal its era, so its basically not useable for live prediction and therefore useless. Since validation data contains eras and eras are the following 28 eras after the 120 training eras, a time-series model would be able to yield meaningful predictions. But since this does not apply to live data I don't see, how this holds any value.

I think eras could potentially be used for feature engineering, but I am still figuring out how this might work. This becomes even harder, because feature aggregates across eras are more ore less identical. If you would find a heuristic that allows to infer which era is in live data and what temporal relation it has to train and validation, that would maybe work. But sure no easy thing to do.

IMO the most benefitial ways to use Eras are:

- Cross Validation (e.g. TimeSeriesSplit, GroupedSplit, Purged Time Series, or something similar based on Eras)

- Batch Training (as described above)