Webinar: Weather Data Inputs for Power System Modeling: Mind the Gaps	
Question	**Answer**
How much can regional diversity of intermittent generation alleviate problems of inaccurate forecasting? Different for solar and wind? (edited)	Regional diversity typically reduces the level of aggregate forecast error because errors cancel out between plants. Diversity is also key to reducing overall variability. Unfortunately, this doesn't help as much in the planning arena though since we need to understand the resource at each location in order to correctly diversify in the first place.
Is the WIND Toolkit an example of "good validation" ?	The WIND Toolkit in it's original form was a seven year long dataset produced using a configuration of the WRF NWP model that was specifically designed to produce the best possible representation of wind speeds across the country near hub-height. At the time it was produced it was best in class. It was validated against observation, but users rarely review the validations. The report goes into a lot of detail about this.
What are your thoughts on stochastic optimization to represent and mitigate for weather impacts on power systems?	Representing weather dependency in stochastic optimization processes is a complicated topic and I confess that I don't know enough about the details to be comfortable providing a definitive opinion. However, it is REALLY important to realize that the variables involved are NOT independent, are interrelated in highly non-linear ways, and further the transition from one state to another follows well defined laws that need to be properly represented and this leads me to urge healthy caution.
Are there published studies that compare different sources (HRRR, NSRDB, ERA5, etc.) when used for power system modeling?	Yes. "Critical review of renewable generation datasets and their implications for European power system models" (Kies et al. 2021) mentioned in the presentation is one such review.
How close does the new NCAR/USGS CONUS404 40-year, 4-km resolution reanalysis data set come to filling the gaps you have identified?	It is helpful. However, there are a number of shortfalls. For instance, the model is calibrated to provide results for precipitation processes, and there is no plan to extend it in an ongoing fashion. It has also not been validated with RA in mind, and 4 km is not sufficient in some areas. Lastly, the curation function is missing. The dataset is discussed in more detail in the report.
when you say that the weather data used in power systems are synthesized, do you mean they are reanalysis? (where the synthesis is based on observations) (edited)	Usually, time series data used to estimate wind and solar generation come in power system modeling come are derived from models. They are very rarely based on observations alone (unlike those used to estimate load which often are). Sometimes the models are simple statistical models, but more typically they are physics based, most often numerical weather prediction. Reanalysis data are a subset of NWP based datasets. The report goes into this in detail.

What are your thoughts on using a higher-resolution dataset (e.g., HRRR) to bias-correct a lower-resolution dataset?	There are methods that can use high-resolution output to correct lower resolution output, for instance Generative Adversarial Networks. The different methods works with varied results. My main opinion is that any such adjustments need to be carefully validated by an expert who understands the underlying modeling methods and their limitations. Validation doesn't just mean making sure the MAE is low, it involves making sure distributions are correct and checking that the model doesn't have failure modes that are hidden by averaging of error statistics.
Coming from the power system side - beyond working with cross-disciplinary teams - how can one enhance their skills or gain experience in this hybrid topic?	I think this question is asking how can a power systems person enhance their meteorological skills. Attending webinars like this one is a good start. There are also excellent resources available in places like the UCAR COMET program. The process needs to be two way. PS people need to educate the atmospheric scientists in how data is used too.
CFD (weather forecasting) is extremely resource intensive. How far off are we on achieving <2km<15min res. What can we do to bridge the gap for now?	CFD isn't needed to get down to 2 km. Mesoscale NWP models are able to get to 500 m or so fairly easily with today's compute capabilities. The issue isn't compute resources, it's making sure the data is useful. Fields can look realistic but until they are compared to observations (and not just with bulk metrics) we don't have any confidence in their skill or uncertainty.
For RA modeling, could a dataset calculate probabilities for other wx variables at other locations, given a constraint in one or more variables? ..... Ex: if temp at Portland is 110+ at 5pm, what is wind gen in SPP	Yes. A well configured dataset could dive into this type of analysis. This is the holy grail in my view. If we produce a high-quality well validated re-analysis the world is our oyster.
Thoughts on SARAH?	I'm guessing the question is referring to SARA the ERCOT Seasonal Assessment of Resource Adequacy. This is currently outside of my area of expertize and I don't have time to dig into it at the moment.
You mentioned transmission is key for solving upcoming energy problems. What is your take on decentralization?	It's not clear what is meant by decentralization. If the point is microgrids, then there is an argument that they can increase resiliance, by allowing islanding, but generally for a high share of renewables a strongly interconnected system with multiple failover options (i.e. resilience to a single point of failure) is the way to go.
Considering current limitations in computational capabilities, what can be done to effectively compensate for NWP limitations and inaccuracies?	Compute capability is not the issue at this point. Neither is expense. High resolution reanalysis is easily achievable. The issue is the political will. The limititations and inaccuracies can be reduced by socializing as much observational data as possible and managing the matter of data availability as a requirement of a high RE share system. The costs are trivial compared to the infrastructure costs and inefficiencies. Good data will save billions of dollars. The longer we wait, the more we lock in the inefficiency. It's not that dissimilar to the cost of not building transmission infrastructure, though the investment is much lower.

Given the data assimilation in HRRR would using short-term forecasts from it, say forecast hour +0-1hrs, be useful for resource adequacy questions?	As I mentioned in the presentation, I am opposed to using outputs from high-resolution operational forecast models as proxies for high-resolution reanalysis data to use as weather inputs. I understand the temptation in the absence of a good solution, but operational models are a moving target. Their biases change frequently and they cannot be used for accurate trend analysis.
Are you aware of a study that validates ERA5 vs the public ERCOT data?	Yes. Unfortunately, I was running short of time and didn't dwell on validation of ERA5 using ERCOT data, but there is a slight showing Dev Milsteins work in the deck.
How you thought about using the RTMA/URMA?	RTMA does not provide the fields needed required to estimate wind generation. The wind speed and direction are only given for 10 m AGL, and there is no solar radiation component. If 10 m winds are used to produce wind generation estimates one will overestimate the daytime generation and underestimate night time since surface level winds increase during the daytime due to convective mixing and decrease at night as the surface layer stabilizes. The opposite happens aloft as the momentum balance shifts. The result of this would be invalid resource adequacy studies. The question is a great demonstration of why it is so crucial that data isn't used as a black box and why experts on both sides must communicate. Let's do it right...the cost is trivial in the grand scheme of things and value of the knowledge we'd gain would far exceed the cost, not to mention the value to other sectors.