[Webinar Series]

Deep Learning approaches for weather-dependent businesses

Questions & Answers

	Questions	Answers
1	What NWP models do you use? Are they open-source models?	For solar production forecasting we use European models, such as ECMWF, and American models, such as GFS. We also use regional models, such as HARMONIE-AROME developed by Météo-France. The models mentioned are not open-source, but you can rebuild them from open-source libraries.
2	What's the difference between CNN-LSTM and ConvLSTM?	The CNN-LSTM approach contains two separate architectures used in succession, while in the ConvLSTM network the convolutional operation is applied inside the LSTM cells.
3	Do you have one CNN for each time step or do you apply the same CNN to each?	You have one CNN for each time step.
4	Which data set did you use to train your model? How do I choose the right size of historical data used to train the model? What makes a "good" data set to train the model?	We used NWP data from different models such as GFS, ECMWF or ARPEGE. There is no single answer to choose the right amount of historical data. My advice is to always evaluate the sensitivity of your model performance by varying the size of your training set. For our study, we used 3 years of data for the training validation process. The sensitivity analysis allowed us to show that the first gains in performance compared to the so-called classical approaches appear for a historical size of 6 months. A good dataset should contain variables that are significant in terms of the physics driving your problem. Depending on the data sources, you also need to pay attention to the quality of your inputs, for example data from ground sensors may contain non-physical values that need to be filtered out. Finally, related to your first question, your data set must contain enough data to take advantage of the deep neural network approach. The "enough" depends on the case.

5	How do you handle spatiotemporal correlations in your architecture, when using CNN and LSTM in succession?	We are using a sequence of image derived from numerical weather predictions, e.g. 12 images or 3 hours of data. Those are first passed on to the CNN, which extracts the spatial features for each time step, and subsequently to the LSTM, extracting the temporal information.
6	Did you use the total solar power generation of the Germany to train the model?	Yes, we used the time series (resolution 15 minutes) of the total PV production for Germany.
7	What kind of data do you use from the NWP map?	We use variables such as global horizontal irradiance, temperature, wind speed, snow cover and total column water vapor, for example.
8	What is the benefit of using LSTM for time series prediction compared to MLP for example?	LSTM is able to take into account the temporal correlation that may be present in your time sequence. The MLP just flattens your data and you lose the temporal correlation and the spatial correlation.
9	Did you also try to establish day-ahead forecasts for single solar parks within Germany using this technique? If yes, what were the result?	We have forecasted individual solar park performances with this method, but not in Germany. Our first use case for single solar park forecasting is located in Réunion Island, where solar power is facing very challenging weather conditions due to insularity, and so is production forecasting. Day-ahead forecasts using the CNN-LSTM were not as good as we had hoped, but the conclusion for intraday forecasts (up to several hours ahead) was different: We found the CNN-LSTM to be able to handle both satellite and NWP information correctly and thus to provide a better accuracy than our actual solution.
10	What is the time resolution of the forecasts?	15 minutes for both, input and output.
11	On slide 37 it seems the training with 36 months of data is actually not much better than the so-called Random	It may be a little disappointing to see these results, partially due to the fact that deep neural networks have become known for classification problems where the performance gain is significant, e.g. Imagenet challenge. In the solar production case we are facing a regression problem and none of our results have shown a significant performance gain

	Forest technique. What's the difference?	using deep neural networks in comparison to other statistical approaches. The technology is however advancing and a field of ongoing research.
12	Is the deep-learning approach still good for data with a lower granularity?	There is no reason why this approach shouldn't work for a lower granularity. However, there is probably a limit for very small temporal resolutions if the characteristic time of the problem that you want to solve is too large compared to the temporal resolution of your inputs.
13	Do you have the skill score of your predictions (improvement over persistence model)?	We did not quantify the gain in performance quality against the persistence model, but the skill score is positive in example where we tested it.
14	Can you quantify the improvement you get by taking into account the non-local weather data in addition to the local weather when forecasting for a specific solar park?	Meteorological data around the location of the solar power plant should help for forecasts based on satellite images or NWP forecasts. It can give clues on the evolution of the weather, for example in terms of cloud cover, when using satellite data. For numerical weather prediction, having a large spatial region around the solar park can help the network to identify spatial patterns in terms of irradiation, and link it to the local solar park production pattern. We have performed such qualitative analysis for the country use case and we've seen that the model is able to recognize the spatial distribution of the photovoltaic installed capacity over Germany quite well. However, we have not quantified the improvement we obtained by taking non-local data into account, but it would be clearly interesting to see how the model behaves when using only local weather data or when the analyzed region is extended spatially vs. local data. Your question has a strong link to the concept of the model's explainability mentioned by our co-speakers during the webinar. Our common main idea is that we need to have some tools to understand in a human way what's going on inside our "black-box" model in order to explain why it performs well or not (For an example, check https://xai-aniti.github.io/ethik/).
15	How much does it approximately cost to train such models and how much does the prediction cost? (in terms of electricity consumption) Is it cheaper or more expensive than using NWP?	A smart resource consumption is the key to cost saving. It is not very expensive when you use cloud computing resources, e.g. AWS. Companies should try to use batch resources required to run their jobs, i.e. dynamic compute resources used to balance demand and load, which would be running anyway during times of less demand (and cheaper electricity cost). Atmo rarely does training of networks with typical physical weather models.

16	You mentioned that the flu is influenced by weather. What are effects of weather on Covid-19?	Atmo just published a paper on the effect of weather on Covid-19. You can find the paper here: http://atmo.ai/covid-2020.html
17		
18		