Paper ~~Plan~~ Outline
"Stochastic Wind Power Forecasting"

Waleed Alhaddad* Ahmed Kebaier‡ Raúl Tempone*†

*CEMSE Division, King Abdullah University of Science and Technology (KAUST), Saudi Arabia

†Alexander von Humboldt Professor, RWTH Aachen University, Germany

‡Universit Paris 13, Sorbonne Paris Cit, LAGA, CNRS (UMR 7539) , Villetaneuse , France

Abstract

- *Briefly say the motivation and applications*
- *What is novel here and why its better.*
  - *Fast*
  - *accurate*
  - *computationally efficient*
- *summarize the technique*
  - *Indirect inference*
  - *Moment based approximation of the transition densities*
- *mention the data used.*
  - *Uruguay*
  - *frequency*
  - *Period*
- *mention concluding results.*

1. Introduction

- Importance and interested parties
- Literature review on current forecasting practices
- shortcomings of current deterministic and probabilistic forecasts.
- Mention our ability of simulating production and why it's important.
- Consequence of not taking into account the real world performance of the forecasts. And present proofs where such forecasts fail such as in complex terrains.
- Consequences of not taking into account skewness of the errors this includes asymmetric cost for up and down-regulation of power.
- Emphasize that it works independent of the forecasting technology. Thus we are able to compare different forecasts based on their real world performance.
- Mention past contributions to the topic and show what is novel in this approach.
- Mention the data used in this study.
- Summary of the structure of the paper, describe the flow and what each section contains.

2. Phenomenological Model

- Introduce the intuitive main model in X in a general setting and what class it belongs to.
- motivate such choice by requiring mean reversion and boundedness of the process.
- Introduce derivative tracking and explain why it's needed.
- State the associated Fokker-Plank
- State the iterated expression for all the moments.

2.1. Physical Constrains

- Introduce why we have the physical constrains.
- We want to fulfill the physical constraints with the least number of parameters and assumptions (parsimony).
- Discuss the normalization between $[0, 1]$ and why its needed
- Motivate the choice of diffusion and its behavior at the boundaries.

- Note that this diffusion is state-dependent and why this is a challenge.
- Motivate the time-dependent θ_t
- ~~combine both~~ write the model including the drift and diffusion~~control~~.
- Do a Feller test to show that the process stays almost surely in $[0,1]$
- Show that if we start with zero error in the mean it will stay that way indefinitely. Show this using Ito.
- Motivate the change of variables to avoid numerical differentiation.

2.1.1 State Independent-Models

- ~~say that we have two main approaches based on moment matching~~note that the state-independent diffusion model is favorable to utilize probability theory.
- introduce Lamperti transformation on V, write the new SDE and its moments.
- give explicit expression to compute the moments as a proposition.
- Remark that in the lamperti space:
  - no longer able to obtain the exact moment equations as we had in the original space.
  - inexactness of the moment equations introduces a bias which we didn't have in the original space.
- We can work in either the original space (state dependent diffusion~~and~~) or the Lamperti space (state-independent diffusion).
  - In the original space we use either a Beta distribution as a proxy or a maximum entropy distribution.
  - ~~state-independent diffusion and~~ In the Lamperti space we use the gaussian as a proxy either by linearizing the SDE or its moments equations.
- ~~introduce Lamperti transformation on V, write the new SDE and its moments.~~
- ~~give explicit expression to compute the moments as a proposition.~~
- Say that We may infer in either spaces, the original space or Lamperti space.
- Discuss why we approximate the SDE directly by linearization or why to approximate its moments instead.
- Ability to estimate parameters using the brackets and least squares when high-frequency data is available.

3. Data

- introduce the data sets: observations, forecast, frequency.
- Show histograms exhibiting the skewness.
- Show the correction to the skewness after Lamperti.
- conclude that the diffusion model $\sqrt{x(1-x)}$ is correct as the Lamperti shows.
- Discuss cleaning the data set from instances of curtailing.
- Example plot of forecast and real production from the data set.

4. Parameter Estimation

- Write the likelihood for one path first using markovianity.
- Justify the product of paths by introducing observation error to the recoded production observations. Write that it's independent and small Gaussian noise.
- Mention that the Fokker-Plank written in the model section is computationally formidable for parameter estimation.
- Mention the two ~~approaches: State dependent diffusion and state-independent diffusion SDE.~~ spaces: the original space and Lamperti space.
- ~~discuss that state dependent diffusion SDE approach~~ We assumes a proxy distribution either a Beta or a maximum entropy distribution. That this is the underlaying assumption and reason it based on the skewness and boundedness of the data explained in the above section.
- ~~discuss that~~ In Lamperti space discuss that we either linearize the ~~state-independent approach requires either linearizing the~~ SDE or its moments equations. Justify it by saying that the time intervals are small and such approximations are reasonable.

4.1. Initial Parameter Estimation

- Introduce why this is important to cross check and it may be possible to use it as an estimating procedure for high-frequency data.
- mention that it is important to help initialize the optimization procedure and save on computations.
- estimate the product of the parameters using the brackets.
- estimate the mean reversion parameter using least squares.

Then,

- State the objective function and the optimization technique

5. Results

For each model and technique we will state the following as a results:

- Results from optimization
  - Convergence of the ellipse around the optimal point.
  - confidence intervals on obtained the parameters
  - AIC and BIC information criteria for model comparison.
- Results from MCMC to verify and demonstrate that we have exhaustively searched the space. This is important as we are not able to guarantee the convexity of our objective function.
- Plots of forecast simulations.
- Plots of forecast confidence intervals.

5.1. Comparison Parameter Estimation Technique

State the results above for the following models:

- Moment matching with a Beta proxy.
- Approximate Moment matching with a Gaussian proxy in Lamperti space.
- Moments matching of a Linearized version of the SDE in the Lamperti space.

6. Model Comparison

Choose the best candidate from the different Parameter Estimation Techniques to compare the following models:

- Model 0: This model is the most basic model without derivative tracking.
- Model 1: This model features derivative tracking with a diffusion term that is forecast dependent by including the term $p_t(1-p_t)$.
- Model 2: Model features derivative tracking and excluding the term $p_t(1-p_t)$

7. Forecast Provider Comparison

Choose the best technique form the ”Comparison Parameter Estimation Technique” section. Then choose the best candidate from the ”Model Comparison” section and apply it on forecasts from two different forecasting companies. State results as mentioned before.

- Assigning weights for making forecast ensembles.

8. Model on Disaggregated Data

Choose the best forecasting company, model, technique and apply it on the disaggregated data. State results as before for each wind farm.

- We study it as a network and how it behaves for power transport and load management. This means that we cannot assume that the nodes are independent, but we need to capture the interplay between them with a minimal number of parameters to tune.
- Must incorporate correlation between wind farms in the noise dW_t.

9. Extra: Optimal forecasting update interval

- We have hourly updated forecast from France. Therefore we can choose the best series of predictions (the first point of each update which is at an hourly frequency).
- Allows us to evaluate the uncertainty for the ”next hour” wind power production forecasts.
- redo the same evaluation on series made up of updates every 4 hours and another every 6 hours and so on.
- this allows us to find the optimal update frequency for different operational uses and clients.

- we are able to classify forecasts according to which forecast horizon they are better at. Say forecaster A technology is better at forecasting the next 12 hours and must be recomputed at that frequency. However, forecaster B is better at forecasting the next hour and must be recomputed hourly.

10. Conclusions

- state our achievements
- summarize results briefly.
- recap of the abstract.
- close with a remark for the future

Acknowledgement

References