Main Results
(using new DataSet)

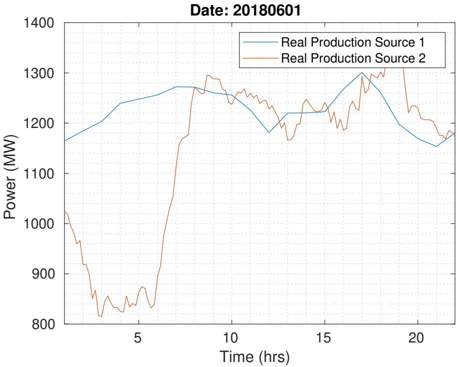
Renzo Miguel Caballero Rosas

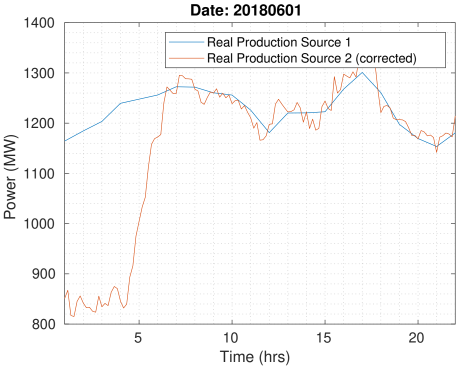
May 21, 2020

About the data:

Curtailing and Delay:

Files: **curt_and_error.eps** and **curt_and_error_corr.eps** (dataConditioner.m, cell (2)).

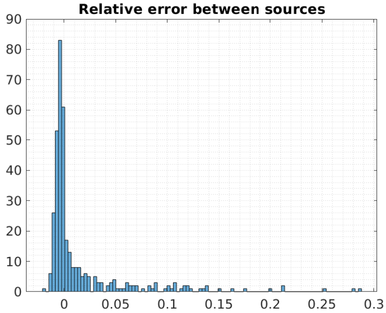


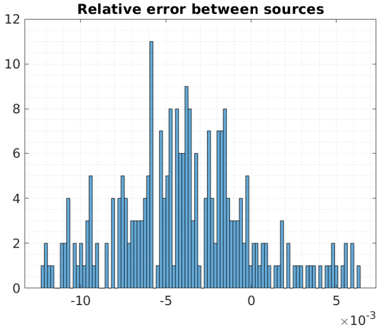


Source 1 has the correct timing; source 2 shows the curtailing and has more frequency. From both sources, we can choose and construct an accurate real production for each day.

Curtailing histograms:

Files: **all2019.eps** and **partially2019.eps** (dataConditioner.m, cell (8)).



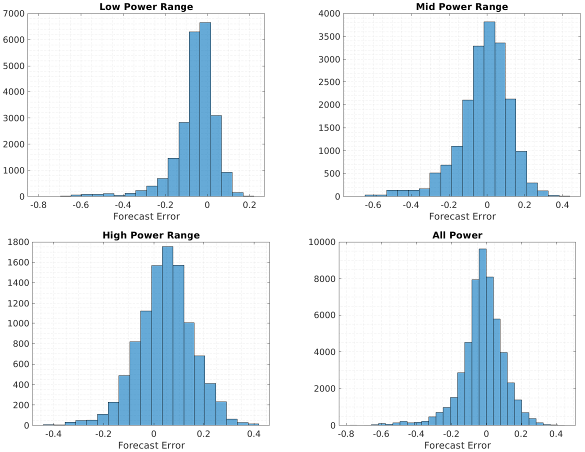


Histograms for the daily mean relative error. On the left, the 365 days. On the right, the corrected data (removed the days with curtailing or other errors).

Error histograms WITH curtailing:

Files: **LP_6.eps**, **MP_6.eps**, **HP_6.eps**, and **AP_6.eps** (dataConditioner.m, cell (12)).

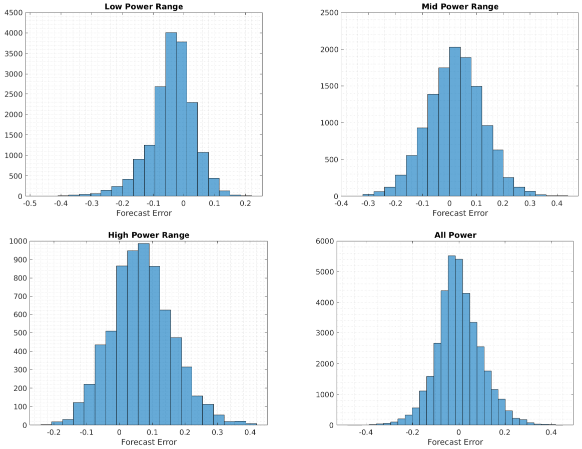
Forecast error for different values of the real production. $LP = [0, 0.3)$, $MP = [0.3, 0.6)$, and $HP = [0.6, 1]$. We have also a histogram for all values of power.



Error histograms WITHOUT curtailing:

Files: **LP.eps**, **MP.eps**, **HP.eps**, and **AP.eps** (dataConditioner.m, cell (11)).

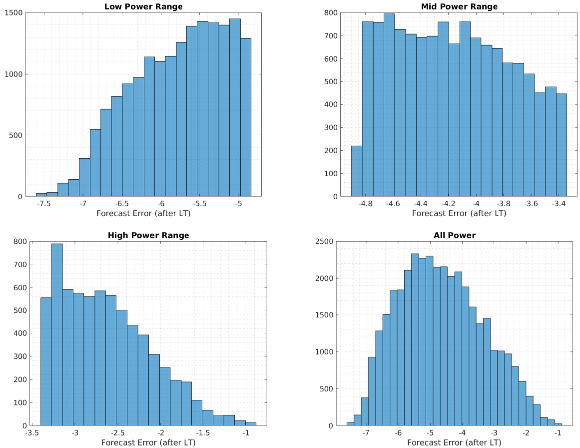
Forecast error after cleaning the data for different values of the real production. $LP = [0, 0.3)$, $MP = [0.3, 0.6)$, and $HP = [0.6, 1]$. We have also a histogram for all values of power.



Lamperti histograms WITHOUT curtailing:

Files: **LP_LP.eps**, **MP_LP.eps**, **HP_LP.eps**, and **AP_LP.eps** (dataConditioner.m, cell (11)).

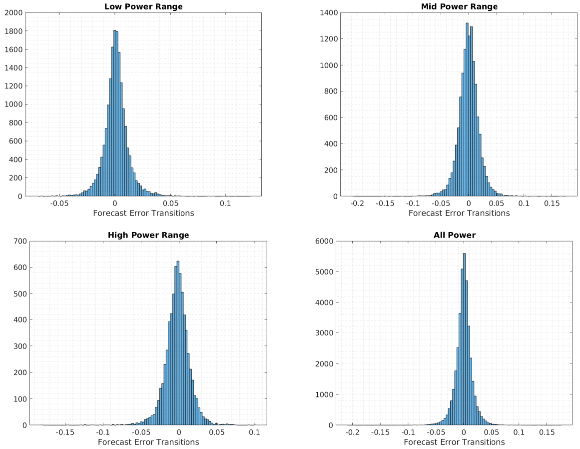
Lamperti after cleaning the data for different values of the real production. $LP = [0, 0.3)$, $MP = [0.3, 0.6)$, and $HP = [0.6, 1]$. We have also a histogram for all values of power.



Error transitions histograms WITHOUT curtailing:

Files: **LP_t.eps**, **MP_t.eps**, **HP_t.eps**, and **AP_t.eps** (dataConditioner.m, cell (11)).

Forecast error transitions after cleaning the data for different values of the real production. $LP = [0, 0.3)$, $MP = [0.3, 0.6)$, and $HP = [0.6, 1]$. We have also a histogram for all values of power.

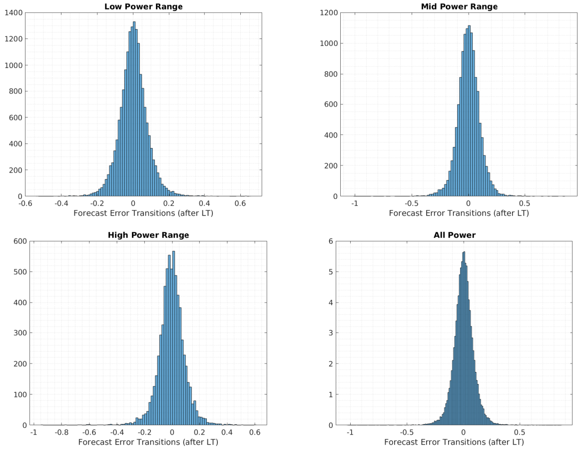


Lamperti error transitions histograms WITHOUT curtailing:

Files: **LP_t_LP.eps**, **MP_t_LP.eps**, **HP_t_LP.eps**, and **AP_t_LP.eps** (dataConditioner.m, cell (11)).

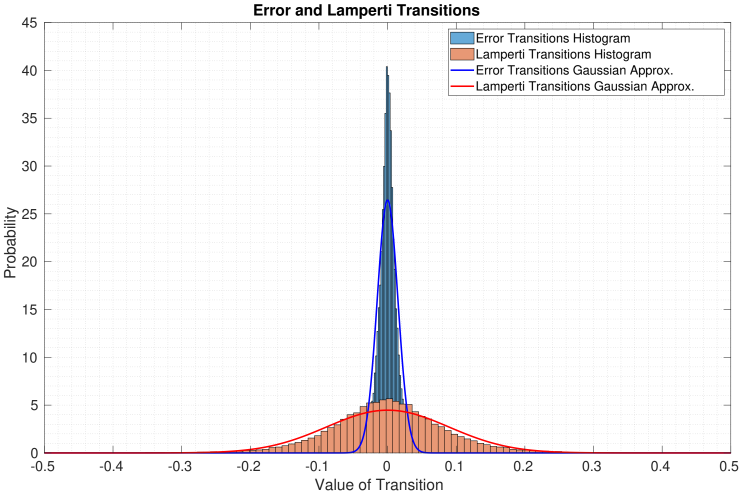
Lamperti error transitions after cleaning the data for different values of the real production. $LP = [0, 0.3)$, $MP = [0.3, 0.6)$, and $HP = [0.6, 1]$. We have also a histogram for all values of power.

We transformed using the initial guesses.



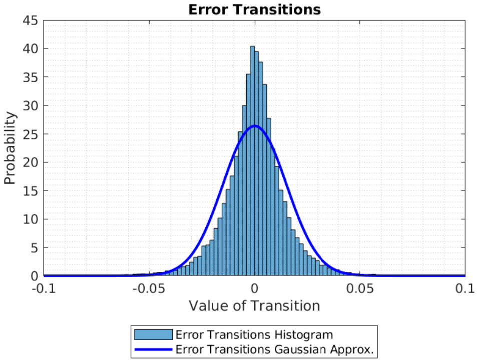
Gaussian approximation for the transitions:

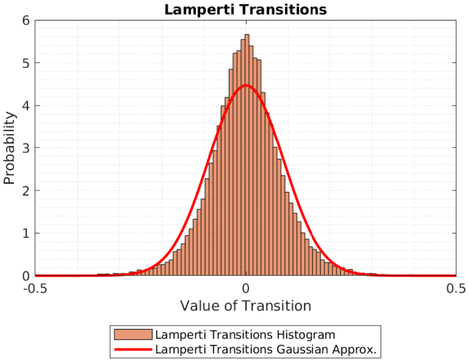
Files: **Gauss_Approx.eps** (dataConditioner.m, cell (11)).



Gaussian approximation for the transitions:

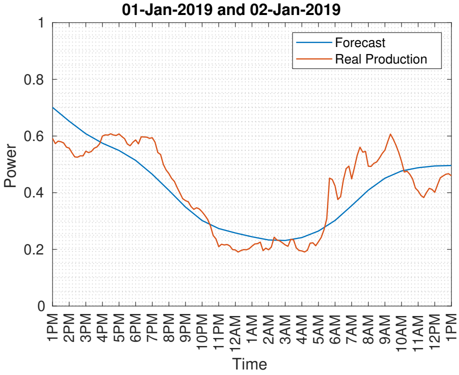
Files: **Gauss_Approx_Err.eps**, and **Gauss_Approx_Lam.eps** (dataConditioner.m, cell (11)).

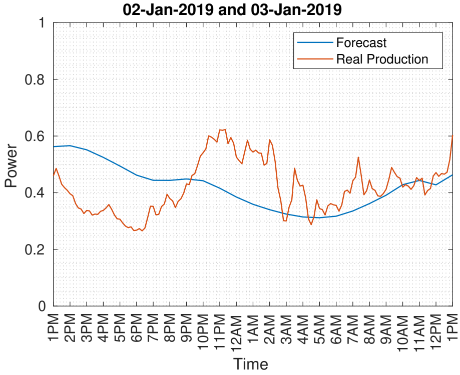




Forecast and production:

Files: **allDaysPlots/1.eps**, and **allDaysPlots/2.eps** (dataConditioner.m, cell (11)).



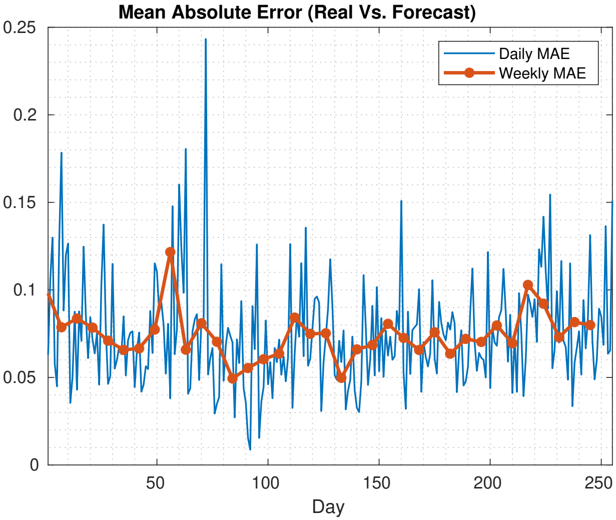


We have this forecast and production plots for the 255 days.

Seasonality effect:

Files: **seasons.eps** (dataConditioner.m, cell (11)).

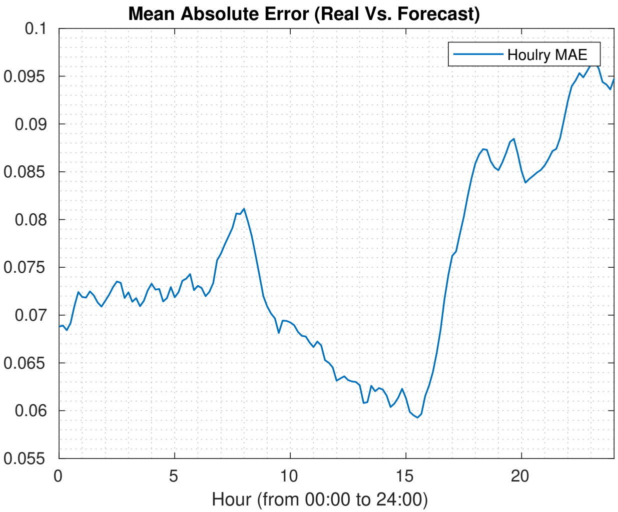
Daily and weakly mean absolute error between the forecast and the real production. We can see no significant seasonality effect.



Hourly effect:

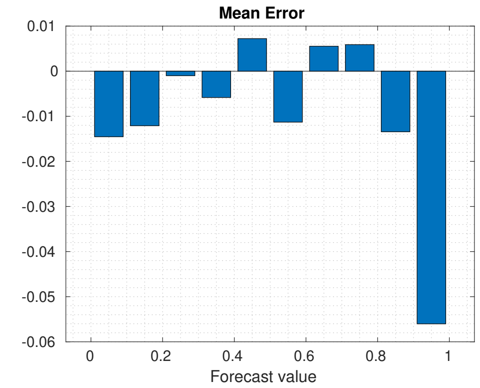
Files: **hourlyEffect.eps** (dataConditioner.m, cell (11)).

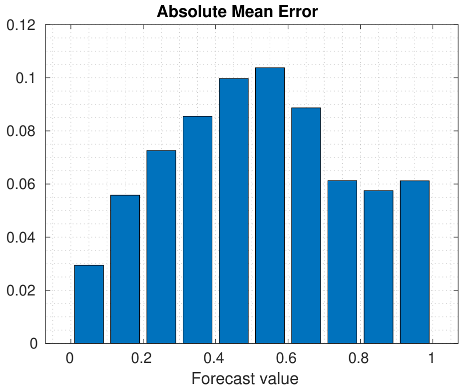
Hourly mean absolute error between the forecast and the real production. We can see a significant reduction in the error during the day.



ME and AME for different intervals of forecast:

Files: **mean_error.eps** and **mean_abs_error.eps** (erroVsForecast.m).

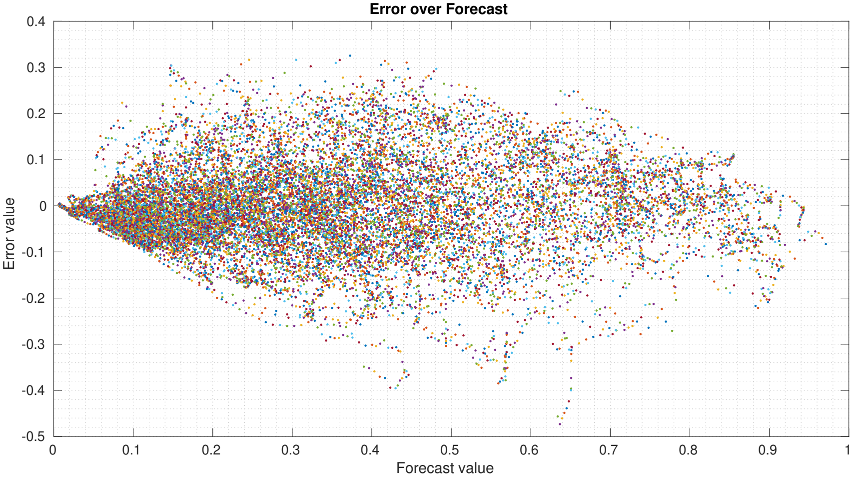




What we are seeing is the **mean error** and **mean absolute error** as a function of the forecast. This is, for each interval with length 0.1 (i.e., [0,0.1), [0.1,0.2), etc.), we average all the errors corresponding to measurement where the forecast was in that intervals, and after we average over the number of elements in each interval.

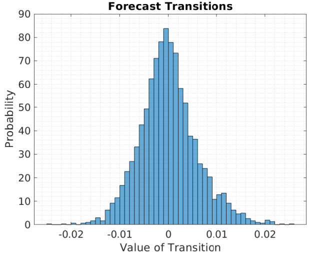
Error Vs. Forecast for all training days (scatter plot):

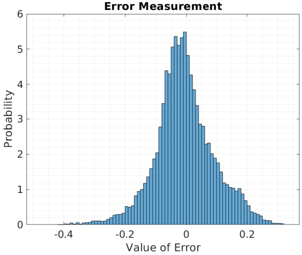
Files: **error_over_forecast.eps** (erroVsForecast.m).

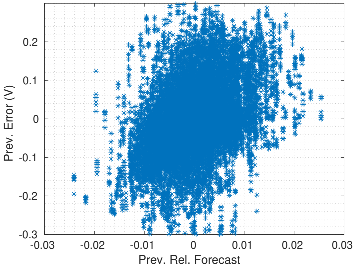


Forecast and error histograms:

Files: **MATLAB_Files/Results/histograms/others** (some_histograms.m).







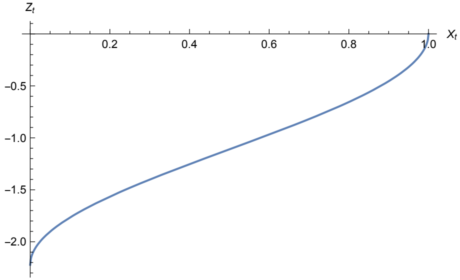
From here we can see that the errors are approximately in the interval $[-0.3, 0.3]$, and the forecast transitions in the interval $[-0.03, 0.03]$.
Then, we want to ensure that the moments are well approximated in the rectangle $[-0.3, 0.3] \times [-0.03, 0.03]$ (for $V \times \Delta p$).

Simulations and Results:

Lamperti transform plot:

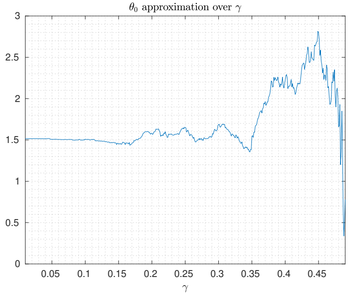
Files: **Mathematica_Files/Range_Z.pdf** (Range_Z.nb).

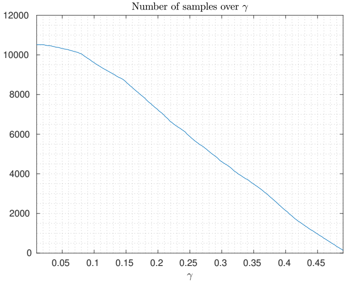
Plot of Z_t as a function of X_t from 0 to 1. We choose $\alpha = 0.06$, and $\theta_t = 1.63$ (initial guess).



Estimation of $(\theta_0, \alpha, \varepsilon)$: θ_0^* using LSM over different sets

Files: **MATLAB_Files/Results/epsilon**: **theta_0.eps** and **num_over_eps_t0.eps** (plot_epsilon.m, third cell).

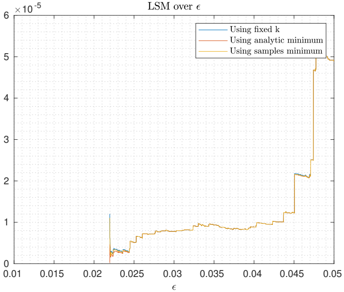


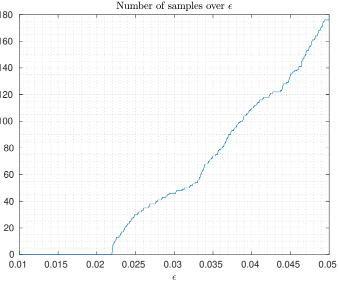


On the right, the number of samples as a function of γ. This samples satisfies that $p_i \in [\gamma, 1-\gamma]$. On the left, the LSM over the samples that satisfies the γ condition.

Estimation of $(\theta_0, \alpha, \varepsilon)$: ε using LSM over different sets

Files: **MATLAB_Files/Results/epsilon**: **LSM.eps** and **num_over_eps.eps** (plot_epsilon.m, second cell).

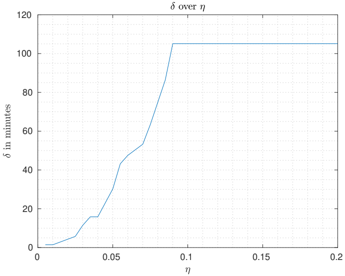


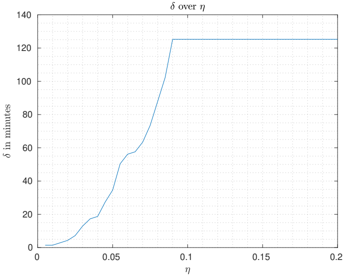


On the right, the number of samples as a function of ε. This samples satisfies that $p_i^{\varepsilon} \in \{\varepsilon, 1-\varepsilon\}$. On the left, the LSM for ε,over the samples that satisfies the ε condition.

δ over η:

Files: **MATLAB_Files/Results/delta**: **eta_min_ini.eps** and **eta_min_opt.eps** (initiasl_time.m, second cell).

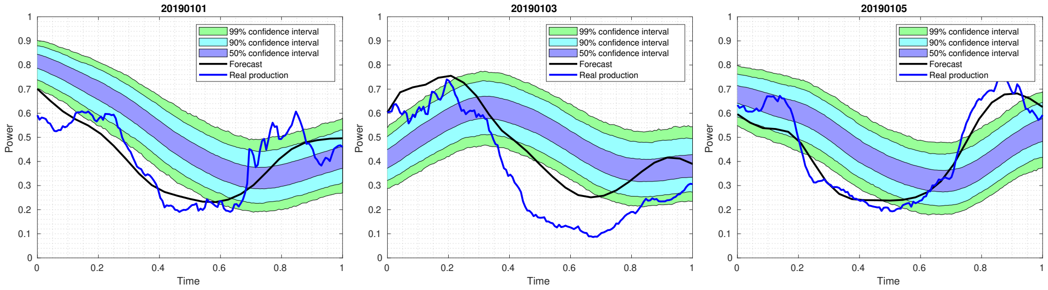




We use the initial transitions that satisfies $|\Delta V| < \eta$. On the left, we use the initial (θ_0, α) $(\alpha\theta_0 = 0.098)$ and we get $\delta = 220\,\text{min}$. On the right, we use the optimal ones. As the optimal product $\alpha\theta_0$ $(\alpha\theta_0 = 0.083)$ is smaller than the initial one, we need either a larger δ to match the variance of the initial error, or to remove some data.

Probability bands for model with delay:

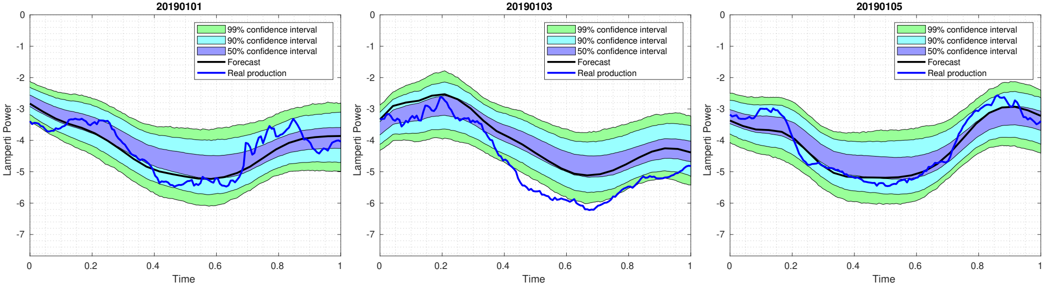
Files: **MATLAB_Files/Results/bands_testing_days/model_1** (path_simulator_OPT_model_1.m).



We used the optimal parameters of the error SDE. We have results for the 128 testing days.

Probability bands for Lamperti:

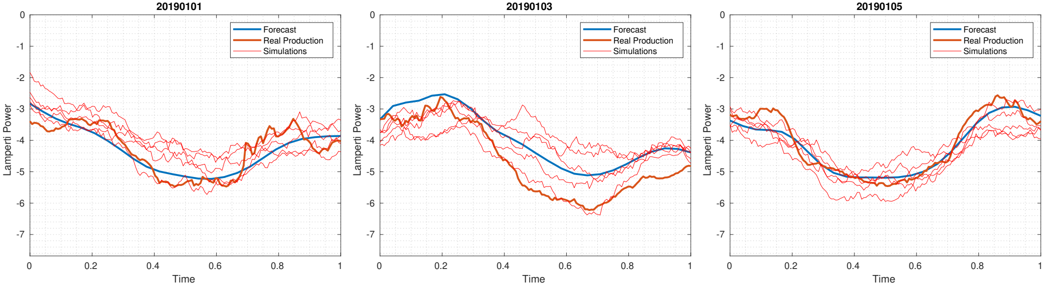
Files: **MATLAB_Files/Results/bands_testing_days/lamperti_optimal** (path_simulator_Lamperti.m).



We used the Lamperti optimal parameters to simulate the Lamperti SDE. We have results for the 128 testing days.

Paths for Lamperti:

Files: **MATLAB_Files/Results/paths_testing_days/lamperti_optimal** (path_simulator_Lamperti.m).

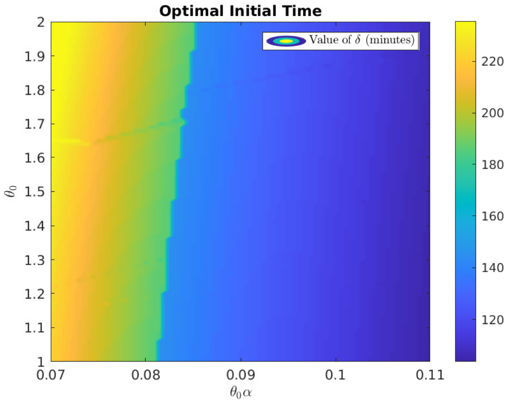


We used the Lamperti optimal parameters to simulate the Lamperti SDE. We have results for the 128 testing days.

Optimal δ:

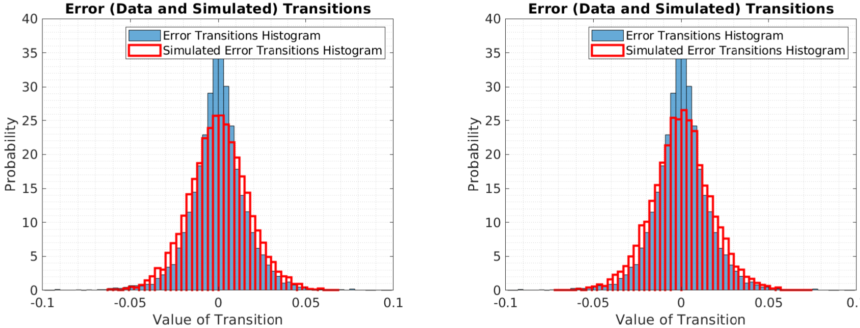
Files: **MATLAB_Files/Results/delta** (initial_time.m).

We can see that the optimal δ is highly dependent on the product $\theta_0\alpha$. This is because, that product represents the diffusion, which is predominant in the estimation of δ.



Error transitions histograms:

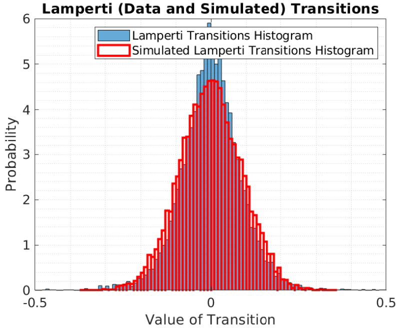
Files: **MATLAB_Files/Results/histograms/classic** (histogram_classic_SDE.m).

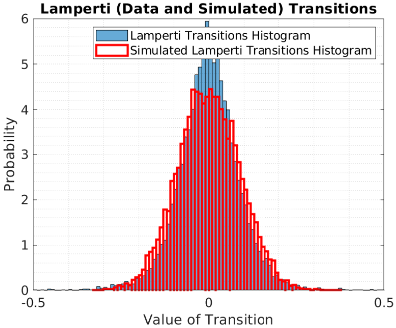


We can see two histograms for the error transitions. On the left, using (θ_0^E, α^E), while on the right, using (θ_0^L, α^L).

Error transitions histograms:

Files: **MATLAB_Files/Results/histograms/lamperti** (histogram_lamperti_SDE.m).





We can see two histograms for the error transitions. On the left, using (θ_0^E, α^E), while on the right, using (θ_0^L, α^L).