Probabilistic Short-Term Low-Voltage Load Forecasting using Bernstein-Polynomial Normalizing Flows

Climate Change AI ICML 2021 Workshop

Marcel Arpogaus[1] Marcus Voss[2] Beate Sick[3] Mark Nigge-Uricher[4] Oliver Dürr[1]

[1]HTWG Konstanz - University of Applied Sciences

[2]Technische Universität Berlin (DAI-Lab)

[3]EBPI, University of Zurich & IDP, Zurich University of Applied Sciences

[4]Bosch.IO GmbH

June 26, 2021

Probabilistic Short-Term Low-Voltage Load Forecasting

- Energy sector is the major contributor to greenhouse gas emissions [6]
- Utilization of renewable energy can be increased by short-term load forecast [3]
- Especially at the low level the distributions are complex and multimodal
- Probabilistic forecasts estimate the distribution to quantify uncertainty for more informed decision-making [4]

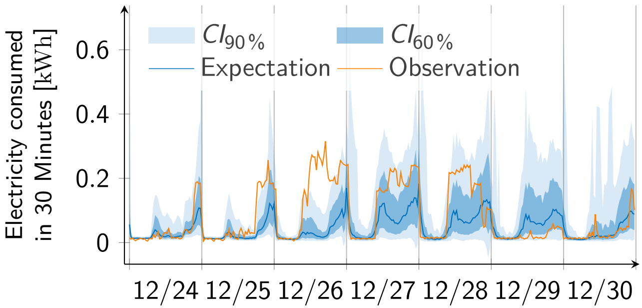
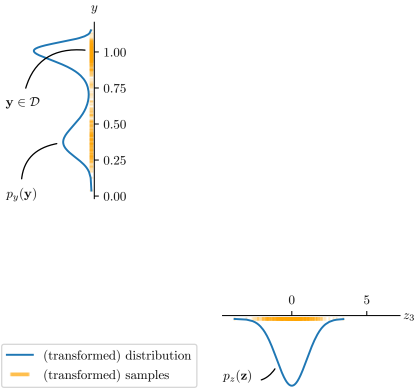


Figure: Exemplary load forecast in Christmas week based on data from [1].

Bernstein-Polynomial Normalizing Flows (BNF)



In a Nutshell

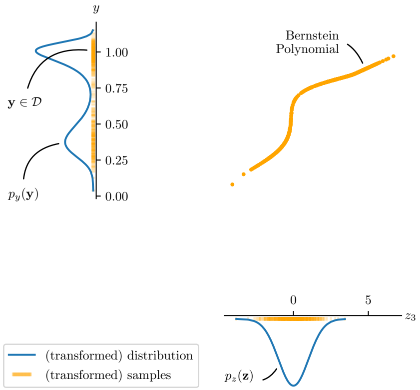
Normalizing Flows fit a parametric bijective function f that transforms between a complex target distribution $p_y(y)$ and a simple distribution $p_z(z)$, often $p_z(z) = N(0,1)$.

Change of Variable Formula

Evaluate the probability $p_y(y)$ from the simple probability $p_z(z)$ as follows:

$$p_y(\mathbf{y}) = p_z(f(\mathbf{y})) \left|\det \nabla f(\mathbf{y})\right|^{-1} \tag{1}$$

Bernstein-Polynomial Normalizing Flows (BNF)

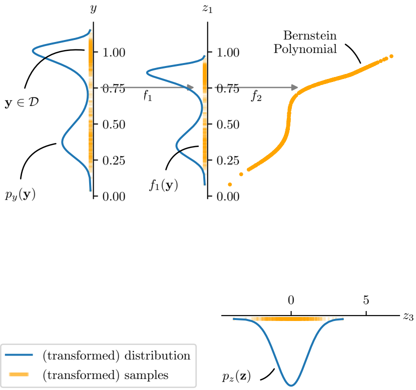


- Bernstein polynomials h^M of order M, approximate any function on $y \in [0, 1]$ [2].
- M controls the flexibility at no cost to the training stability [5].

$$h^M(y) = \frac{1}{M+1} \sum_{i=0}^{M} \mathrm{Be}_i^M(y) \vartheta_i(\mathbf{x})$$

With covariates $\mathbf{x}$ such as lagged power consumption at earlier time steps, holiday indicator and calendar variables.

Bernstein-Polynomial Normalizing Flows (BNF)

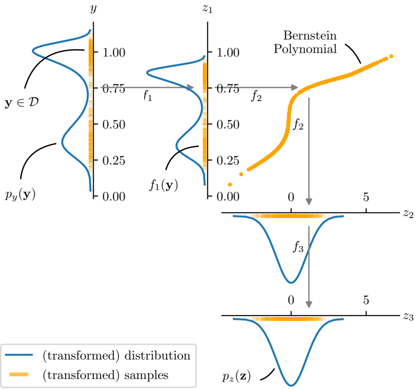


$$f_1 : z_1 = a_1(\mathbf{x}) \cdot y - b_1(\mathbf{x})$$

$$f_2 : z_2 = \frac{1}{M+1} \sum_{i=0}^{M} \mathrm{Be}_i^M(z_1)\vartheta_i(\mathbf{x})$$

With covariates $\mathbf{x}$ such as lagged power consumption at earlier time steps, holiday indicator and calendar variables.

Bernstein-Polynomial Normalizing Flows (BNF)



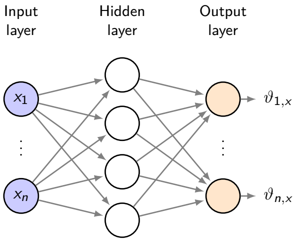
$$f_1 : z_1 = a_1(\mathbf{x}) \cdot y - b_1(\mathbf{x})$$

$$f_2 : z_2 = \frac{1}{M+1} \sum_{i=0}^{M} \mathrm{Be}_i^M(z_1)\vartheta_i(\mathbf{x})$$

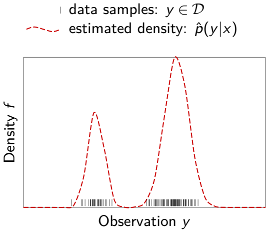
$$f_3 : z_3 = a_2(\mathbf{x}) \cdot y$$

With covariates $\mathbf{x}$ such as lagged power consumption at earlier time steps, holiday indicator and calendar variables.

Maximum Likelihood Estimation



Neural Network:
$n(\mathbf{x}, \omega) = \theta_x$ with $\mathbf{x} = [x_1, \ldots, x_n]^T \in \mathcal{D}$



Conditioned probability density:
$\hat{p}(y, \theta_x)$ with $\theta_x = [\vartheta_{1,x}, \ldots, \vartheta_{n,x}]^T \in \Theta$

Choose the parameters of the model so that the observed data has the *highest likelihood*.

$$\mathcal{L}(\mathcal{D}|\omega) = \prod_{i=0}^{m} \hat{p}(y_i, n(\mathbf{x}_i, \omega))$$

$$\hat{\omega} = \arg\max_{\omega \in \Omega} \{\mathcal{L}(\mathcal{D}|\omega)\}$$

Maximum likelihood estimation

Results of empiric experiments

Architecture	Model	CRPS	NLL	MAE	MSE
FC	BNF (ours)	0.021	-123.157	0.414	9.430
FC	GMM	0.026	-114.892	0.360	0.464
FC	GM	1.314	-75.101	1.027	53.203
FC	QR	0.026	-	0.415	4.351
1DCNN	BNF (ours)	0.017	-132.089	0.342	0.429
1DCNN	GMM	0.019	-125.933	0.384	0.450
1DCNN	GM	0.018	-101.290	0.347	0.366
1DCNN	QR	0.017	-	0.321	0.399

Abbreviations

FC Fully connected neural network
1DCNN dilated 1D-Convocational neural network
BNF Bernstein-Polynomial Normalizing Flow
GM Gaussian model
GMM Gaussian mixture model
QR Quantile regression

CRPS Continuous ranked probability score
NLL Negative logarithmic likelihood
MAS Mean absolute Error
MSE Mean Squared Error

Summary

- BNFs are a powerful and stable method to express complex distributions, with almost no regularization or special tuning.
- This makes them a preferential choice over quantile regression or mixture models for probabilistic load forecasts.
- BNFs allow sampling form the learned distribution to generate synthetic load data for research or scenarios used for grid planning.

Probabilistic Short-Term Low-Voltage Load Forecasting using Bernstein-Polynomial Normalizing Flows

Thanks for your interest, especially also on behalf of my colleagues:

- Marcus Voss
- Beate Sick
- Mark Nigge-Uricher
- Oliver Dürr

Marcel Arpogaus

PhD Student at the University of Applied Sciences Konstanz

eMail `marcel.arpogaus@htwg-konstanz.de`

GitHub `/MArpogaus/stplf-bnf`

arXiv `(not-published-yet)`

Bibliography I

[1] Commission for Energy Regulation (CER).
CER Smart Metering Project - Electricity Customer Behaviour Trial.
http://www.ucd.ie/issda/data/commissionforenergyregulationcer/, 2012.

[2] R. T. Farouki.
The Bernstein polynomial basis: A centennial retrospective.
Computer Aided Geometric Design, 29(6):379–419, Aug. 2012.

[3] S. Haben, S. Arora, G. Giasemidis, M. Voss, and D. V. Greetham.
Review of Low-Voltage Load Forecasting: Methods, Applications, and Recommendations.
arXiv:2106.00006 [stat], May 2021.

[4] T. Hong and S. Fan.
Probabilistic electric load forecasting: A tutorial review.
International Journal of Forecasting, 32(3):914–938, July 2016.

[5] S. Ramasinghe, K. Fernando, S. Khan, and N. Barnes.
Robust normalizing flows using Bernstein-type polynomials.
arXiv:2102.03509 [cs, stat], Feb. 2021.

Bibliography II

[6] World Resources Institute.
World greenhouse gas emissions in 2016, 2020.