# High volatility, high emissions? a hidden-Markov model approach

Pierdomenico Duttilo[1], Marina Bertolini[1], Alfred Kume[2], Stefano Antonio Gattone[3]

[1]*University of Padova, DSS, Via C. Battisti 241, 35121 Padova, Italy*

pierdomenico.duttilo@unipd.it

[2]*University of Kent, SMSAS, Park Wood Rd, CT2 7FS Canterbury, UK*

[3]*University "G.d'Annunzio", DISFIPEQ, Viale Pindaro 42, 65127 Pescara, Italy*

Preprint version

**Abstract**

This study introduces the generalised normal distribution hidden Markov model (GND-HMM) with constrained and unconstrained parameters, where a first-order Markov chain governs the draws of the states from the mixture components. The proposed model is applied to the daily electricity price returns of the electricity market in northern Italy to detect the turmoil periods that occurred during the years 2020-2023. The turmoil periods detected by the GND-HMM model highlight important aggregate events such as the Covid-19 pandemic and the Russia-Ukranian conflict. Furthermore, our study aims to examine the relationship between the identified turmoil periods and the time series of $CO_2$ emissions in the northern Italian electricity market.

**Keywords**: hidden Markov models, electricity prices, $CO_2$ emissions

## 1 Introduction

Due to the rise in geopolitical uncertainties around the world, there have been significant fluctuations in electricity prices throughout Europe [1]. In electricity markets, price volatility, especially when detected in specific market zones, might be heavily influenced by the infrastructures and the energy mix of the area. Therefore, periods of high volatility in the electricity market can result in changes in the generation of electricity at the margin. For example, during periods of high demand or supply disruptions, there may be a greater dependence on certain types of power plants (such as coal or oil plants) that emit higher levels of $CO_2$ compared to cleaner sources such as natural gas or renewables [2]. Analysing turmoil periods in the energy market is crucial to predict its future changes and might be relevant to understand its implications in terms of $CO_2$ emissions.

In this paper, we introduce the generalised normal distribution hidden Markov model (GND-HMM) with constrained and unconstrained parameters, applying it to the Italian electricity market and, in particular, to the day ahead market for the northern part of Italy. We want to start analysing the day ahead market even if it is less exposed to sudden changes in production, but further analysis on balancing markets will be carried out in future research.

After this brief introductory section, we turn to Section 2 where the GND-HMM is introduced. The results are presented in Section 3. Finally, some conclusions are provided in Section 4.

## 2 Data and Methodology

The study is carried out considering the daily electricity prices of the day-ahead market of northern Italy ($MGP_{North}$). Most electricity transactions occur within the MGP, where prices and demand for the following day are established through simultaneous auctions every hour. Data on daily electricity prices are collected from Gestore Mercati Energetici (GME) for the period from January 1, 2020 to January 31, 2024 [3]. The daily returns are computed as follows

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) = \log(P_t) - \log(P_{t-1}), \tag{1}$$

where $r_t$ denotes the log return on electricity price a time $t$, $P_t$ is the electricity price at time $t$, while $P_{t-1}$ is the electricity price at time $t-1$.

A random variable $X_t$ is said to have a $K$-component finite mixture of generalized normal distributions (MGND) if its density is given by

$$f(x_t|\pi_1, ...\pi_K, \theta_1, ..., \theta_K) = \sum_{k=1}^{K} \pi_k f_k(x_t|\theta_k), \tag{2}$$

where $0 < \pi_k < 1$, $k = 1, ..., K$, $\sum_{k=1}^{K} \pi_k = 1$ are the mixture weights, and $f_k(x_t|\theta_k)$, $k = 1, ..., K$, are the component densities taken to be generalized normal distributions (GND),

$$f_k(x_t|\theta_k) = f_k(x_t|\mu_k, \sigma_k, \nu_k) = \frac{\nu_k}{2\sigma_k \Gamma(1/\nu_k)} \exp\left\{-\left|\frac{x_t - \mu_k}{\sigma_k}\right|^{\nu_k}\right\} \tag{3}$$

where $\Gamma(1/\nu_k) = \int_0^\infty t^{1/\nu_k - 1} \exp^{-t} dt$, $\mu_k$ is the $k$-th location parameter ($\mu_k \in$ R), $\sigma_k$ is the $k$-th scale parameter ($\sigma_k > 0$), and $\nu_k$ is the $k$-th shape parameter ($\nu_k > 0$). Figure 1 shows the $k$-th probability density function (Eq. 3) for $\mu_k = 1$, $\sigma_k = 1$. The shape parameter $\nu_k$ controls both the peakedness and the tail weights. If $\nu_k = 1$ the GND reduces to the Laplace distribution and if $\nu_k = 2$ it coincides with the normal distribution. It is observed that $1 < \nu_k < 2$ produces an "intermediate distribution" between the normal distribution and the Laplace distribution. As limit cases, for $\nu \to \infty$ the distribution tends to a uniform distribution, while for $\nu \to 0$ it will be impulsive [4].

Modelling the return density using the MGND model (Eq. 2) is advantageous for two main reasons. First, it is a flexible tool capable of capturing the skewness and excess kurtosis of daily returns [4]. Secondly, from a theoretical perspective, the existence of various mixture components can be attributed to several market states (or regimes). Usually, financial markets are characterised by two market periods: stable and tumultuous. The latter are caused by aggregate events like economic and geopolitical shifts. Since the mixture components could be different with respect to their variances, a daily return belongs to the stability period if it belongs to the mixture component with the lowest variance. Similarly, a daily return belongs to the turmoil period if it belongs to the mixture component with the highest variance [5].

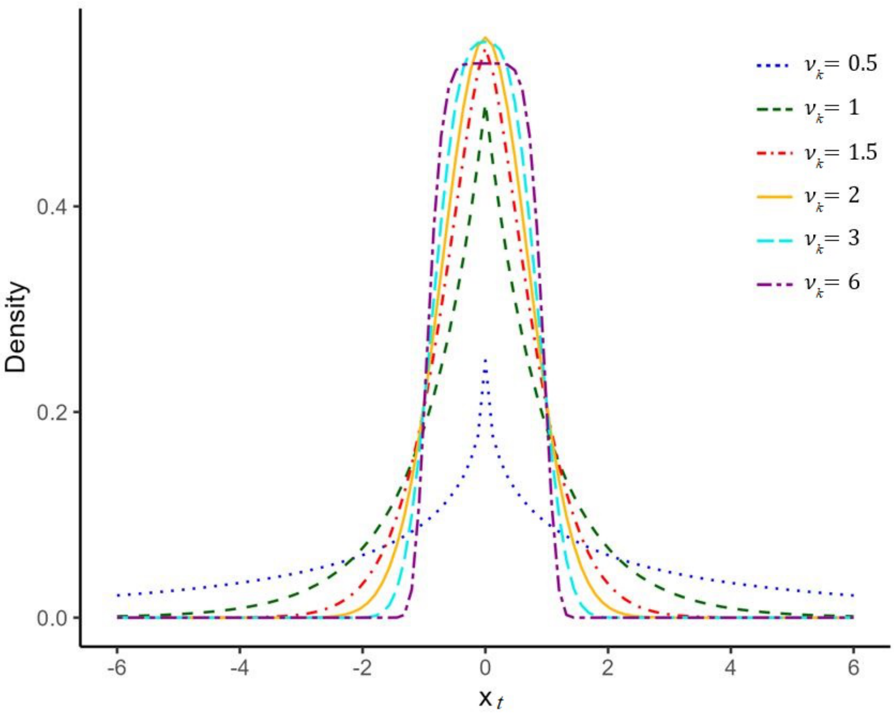


Figure 1: The k-th probability density function for $\mu_k = 0$ and $\sigma_k = 1$.

However, as already identified by Mandelbrot 1963, daily returns exhibit volatility clustering, showing small fluctuations during stable periods and large fluctuations during turmoil periods [6, 7]. In statistical terms, this means heteroskedasticity, i.e., non-constant variance over time. Independent mixture models, such as the MGND model, cannot fully capture heteroskedasticity, since they do not account for temporal dependence [8].

Therefore, it is necessary to define the stochastic process that determines the market states (or regimes), which involves changes in the mixture weights $\pi_k$ for $k = 1, ..., K$. The mixture mechanism should be set in a way that emphasises the probability of choosing the same mixture component at time $t+1$ as the one selected at time $t$. In such situations, Markov chains become essential. A sequence of discrete random variables $\{S_t : t \in \mathbb{N}\}$ is a Markov chain if for all $t \in \mathbb{N}$ the Markov property is satisfied $\Pr(S_{t+1}|S_t, ..., S_1) = \Pr(S_{t+1}|S_t)$. The future state $S_{t+1}$ depends only on its current state $S_t$ and not on the state sequence that preceded it [9]. A Markov chain is represented by important quantities called transition probabilities $\gamma_{ij}(t) = \Pr(S_{c+t} = j|S_c = i)$. An homogeneous ($\gamma_{ij}(t) = \gamma_{ij}$ for all $t$) Markov chain with $K$ states can be defined by a $K \times K$ transition probability matrix (t.p.m.)

$$\mathbf{\Gamma} = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1K} \\ \vdots & \ddots & \vdots \\ \gamma_{K1} & \cdots & \gamma_{KK} \end{pmatrix}, \text{ with } \gamma_{ij} \in [0,1] \text{ for all } ij \text{ and } \sum_{k=1}^{K} \gamma_{ij} = 1. \tag{4}$$

A Markov chain with t.p.m. $\mathbf{\Gamma}$ have stationary distribution $\boldsymbol{\pi} = (\pi_1, ..., \pi_K)'$ if $\boldsymbol{\pi}\mathbf{\Gamma} = \boldsymbol{\pi}$ and $\boldsymbol{\pi}\mathbf{1}' = 1$. The vector $\boldsymbol{\pi}$ is subject to $\sum_{k=1}^{K} \pi_k = 1$ with $\pi_k > 0$. Each element $\pi_k$ represents the probability that the chain is in state $k$ when the chain is in a state described by $\boldsymbol{\pi}$. The MGND model (Eq. 2) together with the Markov chain (Eq. 4) gives the GND-HMM.

In addition, to obtain more parsimonious models, constraints are imposed on the location, scale, and shape parameters: $\mu_k = \mu$, $\sigma_k = \sigma$, $\nu_k = \nu$, for $k = 1, ..., K$. With $K = 2$, taking into account all possible combinations of these constraints would result in a family of eight models. Parameter estimation is performed with the direct optimisation method proposed by Zucchini et al. [9]. The likelihood function of the GND-HMM is given by

$$L_T = \Pr\left(\boldsymbol{X}^{(T)} = \boldsymbol{x}^{(T)}\right) = \boldsymbol{\pi} \boldsymbol{P}(x_1) \Gamma \boldsymbol{P}(x_2) \ldots \Gamma \boldsymbol{P}(x_T) \mathbf{1}', \tag{5}$$

where $\boldsymbol{X}^{(T)} \equiv (X_1, X_2, ..., X_T)$, $\boldsymbol{\pi}$ is the initial distribution of $S_1$ and $\boldsymbol{P}(x)$ is the $K \times K$ diagonal matrix with diagonal element equal to $f_k(x_t | \mu_k, \sigma_k, \nu_k)$, the $k$-th state-dependent distribution.

## 3 Results

Daily returns on $\text{MGP}_{\text{North}}$ are not normally distributed as they present negative skewness (-0.9826) and excess kurtosis (24.1607). Table 1 shows the two-state GND-HMM estimates with and without constraints for the daily returns on $\text{MGP}_{\text{North}}$. The Bayesian information criterion (BIC) is used to select the best model and compare the goodness of fit. According to the BIC the GNDHMM-CCU is the best model. Figure 2 panel a shows the estimated marginal distribution with the two-state GND-HMM-CCU. It can be seen that the stable component (dotted line) describes the central and intermediate values of the data, while the turmoil component (dashed line) describes the more extreme tail behaviours. The stable component is predominant compared to the turmoil component ($\hat{\pi}_1 > \hat{\pi}_2$). The shape parameter of the turmoil component ($\hat{\nu}_2 = 0.8034$) is lower than the shape parameter of the stable component ($\hat{\nu}_1 = 1.5675$). Specifically, the turmoil component has heavier tails than a Laplace distribution. Figure 2 panel b shows the decoded turmoil periods using the Viterbi algorithm [10]. Turmoil periods are highlighted in yellow.

## 4 Final remarks

This study introduces GND-hidden Markov models with unconstrained and constrained parameters, where a first-order Markov chain governs the draws of the states of the mixture components. The proposed model is applied to the daily electricity price returns of the electricity market in northern Italy ($\text{MGP}_{\text{North}}$) to detect recent turmoil periods. The turbulent periods detected in Figure 2 panel b highlight important aggregate events such as the outbreak of the Covid-19 pandemic and the Russia-Ukranian conflict.

Future research will focus on finding out whether the presence of turbolences in the zonal market corresponds to a higher use of pollutant energy sources or vice versa if they represent higher penetration of renewables. In this sense, we want to analyse the trade-off between market stability and the level of emissions. Future extensions will take into account data from all the Italian geographical zones for the MGP market and, later, on the national balancing market.

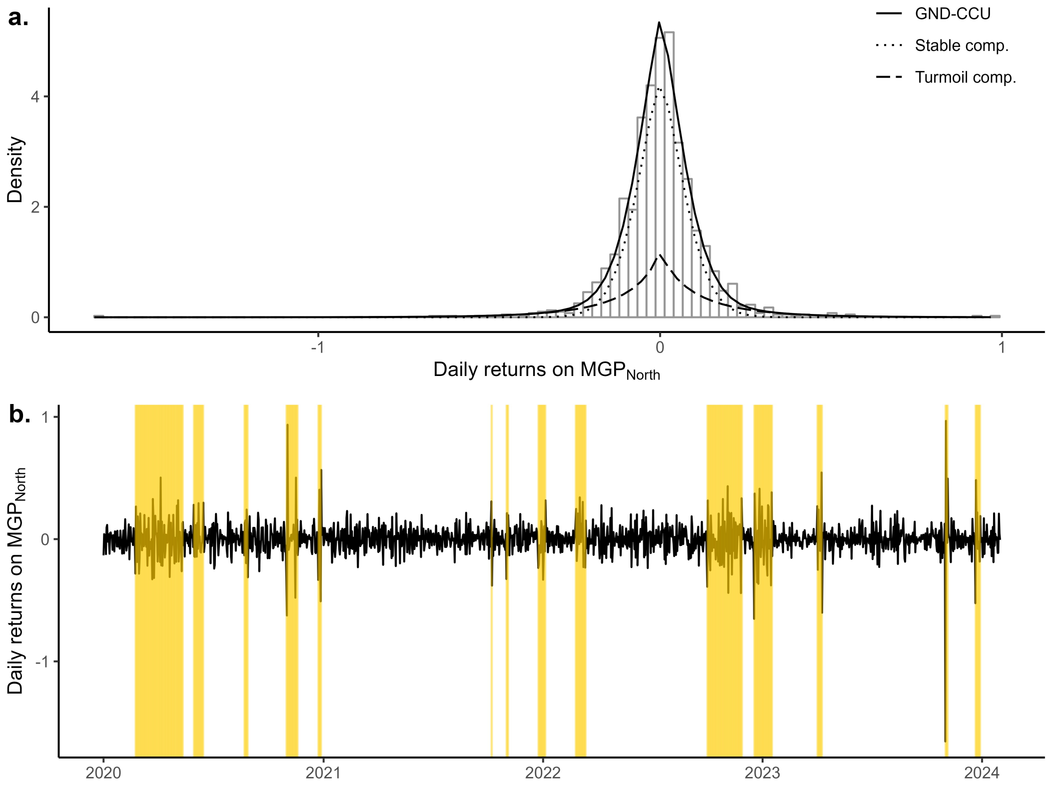


Figure 2: Estimated marginal distribution in panel (a) and turmoil periods (coloured yellow) decoded through the Viterbi algorithm in panel (b).

Table 1: Estimated two-state GND-HMM models for daily returns on $MGP_{North}$.

| | GND-HMM | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $\theta$ | UUU | CUU | UCU | UUC | CCU | CUC | UCC | CCC |
| $\pi_1$ | 0.7360 | 0.7364 | 0.7293 | 0.7729 | 0.7292 | 0.7740 | 0.9987 | 0.5000 |
| $\mu_1$ | 0.0022 | 0.0016 | 0.0019 | 0.0021 | 0.0012 | 0.0017 | 0.0011 | 0.0011 |
| $\sigma_1$ | 0.0882 | 0.0884 | 0.096 | 0.0833 | 0.0962 | 0.0835 | 0.0739 | 0.0707 |
| $\nu_1$ | 1.4942 | 1.4958 | 1.5648 | 1.3517 | 1.5675 | 1.3538 | 0.9327 | 0.9001 |
| $\pi_2$ | 0.2640 | 0.2636 | 0.2707 | 0.2271 | 0.2708 | 0.2260 | 0.0013 | 0.5000 |
| $\mu_2$ | -0.0042 | 0.0016 | 0.0000 | -0.0048 | 0.0012 | 0.0017 | 0.9674 | 0.0011 |
| $\sigma_2$ | 0.1827 | 0.1827 | 0.0960 | 0.2301 | 0.0962 | 0.2310 | 0.0739 | 0.0707 |
| $\nu_2$ | 1.1531 | 1.1522 | 0.8028 | 1.3517 | 0.8034 | 1.3538 | 0.9327 | 0.9001 |
| $\gamma_{11}$ | 0.9591 | 0.9592 | 0.9670 | 0.9627 | 0.9671 | 0.9629 | 0.9987 | 0.8000 |
| $\gamma_{22}$ | 0.9591 | 0.9592 | 0.9670 | 0.9627 | 0.9671 | 0.9629 | 0.9987 | 0.8000 |
| $\log L(\hat{\theta})$ | 1254.39 | 1254.22 | 1247.10 | 1252.03 | 1246.93 | 1251.88 | 1190.42 | 1184.86 |
| BIC | -2443 | -2449.97 | -2435.73 | -2445.60 | **-2452.60** | -2442.70 | -2329.68 | -2325.87 |

*Notes*. In bold the best model according to the BIC.

**Acknowledgments** This study was funded by the European Union - NextGenerationEU, in the framework of the GRINS - Growing Resilient, INclusive and Sustainable project (GRINS PE00000018 - CUP C93C22005270001). The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them.

## References

[1] Liu D, Liu X, Guo K, Ji Q, Chang Y.: Spillover Effects among Electricity Prices, Traditional Energy Prices and Carbon Market under Climate Risk. Int J Environ Res Public Health. **20**, 1–18 (2023).

[2] Sovacool, B. K.: The intermittency of wind, solar, and renewable electricity generators: Technical barrier or rhetorical excuse? Util Policy **17**, 288–296 (2009).

[3] Gestore dei Mercati Energetici (GME). https://www.mercatoelettrico.org/it/

[4] Wen, L., Qiu, Y., Wang, M., Yin, J., Chen, P.: Numerical characteristics and parameter estimation of finite mixed generalized normal distribution. Commun Stat Simul Comput **51**, 3596–3620 (2022).

[5] Duttilo, P., Gattone, S.A., Iannone, B.: Mixtures of generalized normal distributions and EGARCH models to analyse returns and volatility of ESG and traditional investments. AStA Adv Stat Anal (2023).

[6] Mandelbrot, B.: A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices: Comment. Econometrica **41**, 157–59 (1973).

[7] Haas, M. Pigorsch C.: Financial Economics, Fat-Tailed Distributions. In: Encyclopedia of Complexity and Systems Science. Springer New York, 3404–3435 (2009).

[8] Kaehler, J., Marnet, V.: Markov-Switching Models for Exchange-Rate Dynamics and the Pricing of Foreign-Currency Options. In: Econometric Analysis of Financial Markets. Heidelberg: Physica-Verlag HD, 203–230 (1994).

[9] Zucchini, W., MacDonald, I., Langrock. R.: Hidden Markov Models for Time Series: An Introduction Using R, Second Edition. Chapman & Hall. CRC Press (2017).

[10] Viterbi, A.: Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. In: IEEE T Inform Theory **13**, 260–269 (1967).