This article is a preprint. Please cite the published version - DOI: 10.1109/TPWRS.2022.3146314

Information Length Quantification and Forecasting of Power Systems Kinetic Energy

Harold R. Chamorro, *Senior Member, IEEE,* Adrian-Josue Guel-Cortez, *Student Member, IEEE,* Eun-jin Kim, *Senior Member, IEEE,* Francisco Gonzalez-Longatt, *Senior Member, IEEE,* Álvaro Ortega, *Member, IEEE,* and Wilmar Martinez, *Senior Member, IEEE*

***Abstract*—One of the short-coming challenges of power systems operation and planning is the difficulty to quantify the variability of power systems Kinetic Energy (KE) to unveil online additional information for the system operators' decisions support. KE monitoring requires innovative methods to analyse the continuous fluctuations in the KE power's systems. In this paper, we propose the use of information theory, specifically the concept of Information Length (IL), as a way to provide useful insights into the power system KE variability and to demonstrate its utility as a starting point in decision making for power systems management. The proposed IL metric is applied to monthly collected data from the Nordic Power System during three consecutive years in order to investigate the KE evolution. Our results reveal that the proposed method provides an effective description of the seasonal statistical variability enabling the identification of the particular month and day that have the least and the most KE variability. Additionally, by applying a Long Short-Term Memory (LSTM) neural network model to estimate the value of the IL on-line, we also show the possibility of using the metric as data-driven support.**

***Index Terms*—Kinetic Energy Variability, Information Length, Time-series Forecasting, Support Decision Tools, Data Fluctuation Analysis.**

I. Introduction

A. Motivation

Existing power systems governed by conventional energy production sources raise environmental concerns. The development towards a sustainable society involving a high penetration of non-synchronous energy sources such as wind, tidal, and solar power provokes challenges for Transmission System Operators (TSOs) which include the grid's instability and the Kinetic Energy (KE) reduction [1], [2]. In this context, TSOs find themselves in the need to develop real-time inertia analysis tools to help them make safe and cost-effective decisions while promoting regulation rules of the system's strength and KE to accommodate high levels of renewable sources [3], [4].

*The first two authors contributed equally

Harold R. Chamorro is with KU Leuven, Katholieke Universiteit Leuven, Leuven, Belgium. e-mail: hr.chamo@ieee.org.

Adrian J. Guel-Cortez is with Coventry University, Faculty Research Centre in Fluid and Complex Systems, Coventry, UK. (guelcortea@coventry.ac.uk).

Eun-jin Kim, is with Coventry University, Faculty Research Centre in Fluid and Complex Systems, Coventry, UK., e-mail: ejk92122@gmail.com

Francisco Gonzalez-Longatt is with University of South-Eastern Norway, Porsgrunn, Norway. (fglongatt@fglongatt.org).

Álvaro Ortega is with Institute for Research in Technology, ICAI, Comillas Pontifical University. (aortega@comillas.edu).

Wilmar Martinez is with the Department of Electrical Engineering (ESAT) at KU Leuven, Diepenbeek, Belgium, (email: wilmar.martinez@kuleuven.be)

In this vein, research laboratories like the Pacific Northwest National Laboratory (PNNL) have managed to facilitate the KE analysis of, e.g., the Western Electricity Coordination Council (WECC) system by developing KE estimation tools [5]. The California Independent System Operator (CAISO) is also concerned about the operating reserves with different resources [6] by requiring new platforms for monitoring the KE and the status of the network [7]. The European synchronous area and its TSOs have also become aware of the power system's KE reduction in their systems, due to the forthcoming non-synchronous generation integration's impact on operation, and coal power decommissioning [8]. For instance, the Romanian TSO Transelectrica S.A. [9] has observed large imbalances degrading the inertia response after dismantling carbon emitting power plants. Correspondingly, the Spanish TSO, Red Electrica de España (REE) [10], has reported a decrease on the system's KE when trying to develop an improved network plan to contemplate low-inertia future scenarios. Particularly, the Nordic Power System (NPS) is on gradual decommissioning of nuclear and thermal power, which will clearly lead to a substantial reduction of the future physical inertia and its intrinsic response in the system [11].

B. Literature Review

As a way to tackle some of the aforementioned challenges, TSOs are equipped with Supervisory Control And Data Acquisition (SCADA) control centres where the status of all the connected generators is available, and the inertia information can be obtained. By taking advantage of these technologies, the Electric Reliability Council of Texas (ERCOT) in the United States has implemented a procedure for real-time monitoring and evaluation of frequency and inertia events. In their procedure, day-ahead market and the ancillary services are included [12]. An online inertia estimation tool has been implemented by the Nordic TSOs, (e.g. Fingrid, Svenska Kraftnät, Statnett, and Energinet), to keep track of the breaker positions of synchronously connected generators, their apparent power and inertia constant. Thus, the KE in the system can be estimated [13]. Another example is the online tool used for the day-ahead and intra-day KE forecast exhibited by the National Grid of the United Kingdom TSO [14], as a testbed for the enhancement of the system [15]. System's inertia estimations using historical operation data of the Great Britain power system in [16], indicate the peak demand periods to improve the system performance.

So far, we have underlined the importance of monitoring the power system's KE and stressed the existing experiences

employed for such a task. However, it is also important to study other aspects of the system's KE, such as the amount of *fluctuations* or *variations* that KE suffers through time. These studies can provide powerful tools for future decision-making, since a KE variation metric may detect issues like abrupt events in the system [17]. In general, analysing fluctuations in any complex system is a challenging endeavour that remains an open research problem. A possible solution to this problem may consist of the application of concepts from areas such as statistical mechanics, stochastic thermodynamics or information theory [18]–[20]. Up to date, few results have been proposed for the study of fluctuations in the power system's KE which make use of the aforementioned approaches. For instance, authors in [21] propose the use of the Rate of Change of Inertia (RoCoI). This concept is useful when assessing the generation schedule and observing the so-called energy fluctuations in the process. Besides, [22] proposes to use Probability Density Functions (PDFs) to quantify the impact of variability in power systems frequency time series for a yearly period, showing a comprehensive sensitivity analysis and PDFs of the frequency of the Center of Inertia (CoI) measurement under different scenarios.

More recent and promising approaches, also based on the study of the PDFs evolution of the system through time, apply the tools provided by the area of information theory. For instance, a power system risk security assessment based on maximum information entropy is studied in [23]. Similarly, an overall information security assessment is given in [24], by using entropy weight coefficients that improve the assessment factors of the network. Based on the entropy principle [25], the severity of cascading events was determined. A model based on cross-entropy for wind power forecasting is proposed in [26], showed an accurate prediction and information fusion process. Additionally, an optimal generation dispatch strategy that uses the maximum entropy principle for accurate probabilistic distributions of the power flow is presented in [27]. Information transfer is used in [28], for characterising the influence between the various states of a dynamical system, and to identify the generators and states which are responsible for causing instability of the power network related to the participation factors. Authors in [29], propose an approach to reconstruct of the power system security region by using information theory and to evaluate the amount of High Information Content (HIC) in each operating space variable, achieving the effective descriptor variables of the security region separately. In order to describe the overall properties of flow distribution, the notion of flow entropy and its relationship with a blackout in terms of both the dynamic propagation path and the scale of static blackout are investigated in [30]. Based on the entropy theory, the optimal allocation of Flexible AC Transmission Systems (FACTS) is obtained in [31]. Power flow entropy weighted by the regional average load rate is analysed in [32], reflecting the power flow distribution accurately.

When discussing the information theory applications to our field of interest, it is important to highlight that most of the presented information-based results use the so-called Shannon entropy concept. However, Shannon entropy does not provide a full description of the statistical changes in a time series. This has lead to the use of other information theory concepts such as the Fisher information [33]–[35], differential entropy [36], the Kullback-Leibler divergence [37] or the information length (IL) [38], [39]. The latter, being of our main interest in this work, permit us to quantify the total amount of statistical changes in a given range of time. As IL depends on the evolution path between two states (PDFs), it has the advantage over other information metrics (for instance, differential entropy) [40]. Moreover, IL's formulation provides an intriguing link between stochastic processes and information geometry [41]. IL has been successfully applied in different process including quantum, fluid and biological process [38]. In addition, inspired by the IL metric, [42] introduces a new information-geometric measure of causality by calculating the effect of one variable on the information rate of the other variable. Furthermore, [43] elucidates the meaning of IL in light of stochastic thermodynamics, in particular, detailing its connection to entropy production or free energy in the case of the non-autonomous Ornstein–Uhlenbeck process. Finally, [44] and [17] present the analysis of the computation of IL for linear stochastic autonomous processes, facilitating its application to a wide range of engineering scenarios, and proving that IL can be used for abrupt event predictions. The last inspiring the application of IL in the present work.

A secondary result follows if we pursue the forecasting of the IL value, which could prevent the TSOs of future abrupt events in the KE fluctuations. Forecasting is a large area of research, but we can focus on one of the most prominent tools used in power systems time series, the Neural Networks (NNs). As a subset of Machine Learning, NNs applicability has been shown in contributions such as frequency nadir forecasting [45], wind power prediction [46], energy demand forecasting [47], and load forecasting [48]. In addition, probabilistic forecasting has been used in wind and photovoltaic (PV) power forecasting. Equivalently, a wind power probabilistic forecasting optimisation-mixed algorithm with quantile regression is proposed in [49], using data from two wind farms. Two locations were studied, one in Colorado and another one in New Brunswick, showing seasonality variation and high accurate results. A bivariate wind power forecasting algorithm based on a parametric approach is used in [50], with marginal distribution using the data sampled from the Iberian peninsula. In [51], Quantile regression averaging is used with Long-short Term Memory (LSTM) to obtain probabilistic of day-ahead PV forecasting using real measurement data. A probabilistic forecasting of PV generation is proposed in [52], incorporating the variability based on quantile regression and extreme learning machine methods. The impact of such analyses on the operation of power systems is a growing area of research and, as the above contributions show, there is potential of applying and merging both ML and probabilistic techniques for forecasting, such as the proposed IL metric not previously studied.

C. Contributions

The aforementioned works highlight the main issues that TSOs are currently facing, namely the evaluation of the

available KE and its evolution (fluctuations) in the power system as the share of synchronous generation is being gradually reduced. While techniques that provide an accurate estimation of the instantaneous values of such parameters exist and have been successfully implemented, the development of approaches to analyse their evolution is still an intense topic of research. Taking into account this background, an information geometry based approach to analyse the statistical evolution of the KE of a power system has not been presented yet. Such a scheme is the main contribution of this paper, namely, a data-driven approach that applies the promising IL metric over the system KE. The proposed method is tested using power systems' KE time series observables of the Nordic Power System (NPS) obtained from system recording data of three consecutive years, 2018, 2019, and 2020. Then, we discuss how the IL's metric quantifies the variability behaviour by season and month. We also evaluate a scenario during 2015 that includes the total KE and the KE produced by hydropower during March. As a final contribution, we propose a probabilistic forecasting tool that results from the combination of the IL and a LSTM recurrent neural network. Such a result is a proof-of-concept of IL as data-driven support in KE abrupt events prediction.

D. Paper Organization

The remainder of the paper is organised as follows: Section II describes the theoretical framework of the frequency response, and KE definition used. Section III formally discusses the concept of IL, its interpretation and the application to time series variability. Section IV explains the details of the forecasting and IL computation algorithm. Section V shows the time series data study cases performed. The simulation results and discussions are presented in Section VI, specifying the highest and lowest variability months per year. Additionally in this section, a case presenting the KE forecast is given. A brief discussion of the KE forecast is presented in Section VII. Finally, the conclusions and future work are presented in Section in VIII.

II. Kinetic Energy Preliminaries

Synchronous and, to some extent, non-synchronous devices are sources of inertia due to their rotating masses and/or supplementary controllers that allow them to contribute to the power system strength. High values of inertia ensure that regulator controls are initiated before critical values are reached by frequency. Low inertia systems, on the other hand, have a high Rate of Change of Frequency (RoCoF), which is not recommended for stability management. In a low-inertia power system, high RoCoF causes the frequency to easily exceed unsafe values and activate safety before activating the control of the governor [53]. As inertia can vary over time due to the penetration of stochastic non-synchronous renewable energy sources, it is important to track the time-varying KE values in the network. This will help to ensure network planning, operation, stability control and optimisation of system KE values.

Because of load variations, the frequency fluctuates around the nominal value during normal operation. As a result of the mismatch between mechanical and electrical torque, the generator rotor speed changes, releasing the inertial response. System inertia (H_T), which is reliant on the inertia constants (H_i) of all synchronous generators in the system and system rated power (S_T), dictates the system's ability to deal with these frequent shifts in the power balance. This is represented as follows:

$$H_i = \frac{KE_i}{S_i} \tag{1}$$

$$H_T = \frac{\sum_{i=1}^{n} KE_i}{S_T} \tag{2}$$

The rotational inertia (H_i) is a good indicator for the capacity of a synchronous generator dominated power system to cope with active power imbalances. However, the practical use of the rotational inertia is limited due to the need to appropriately relates the individual rotational inertia of each generator to a common base to obtain the total rotational inertia of a system. Instead of using the rotational inertia, we use the KE which is a more convenient dimensional indicator to allow base-free comparisons.

III. Information Length Theory

In this section, we provide a theoretical background of the IL metric. We also present an interpretation of the metric and a practical example where IL is computed from a mathematical model. Besides, to compute IL from a time series, a discrete version of IL is defined. Then, we describe the algorithm to compute IL from a time series. Finally, we describe the major benefits of the metric.

A. The concept of Information length

In mathematical terms, given a time-dependent probability density function $p(\mathbf{x};t)$ of a nth-order stochastic variable $\mathbf{x}$, the Information Length $\mathcal{L}(t)$ of its evolution from an initial time $t_0 = 0$ to a final time $t_f = t$ is defined by

$$\mathcal{L}(t) = \int_0^t \frac{dt_1}{\tau(t_1)} = \int_0^t \sqrt{\mathcal{E}(t_1)}\,\mathrm{d}t_1, \tag{3}$$

$$\mathcal{E}(t) = \int_{\mathbb{R}^n} \left(\frac{1}{p(\mathbf{x};t_1)} \left[\frac{\partial p(\mathbf{x};t_1)}{\partial t_1} \right]^2 \right) \mathrm{d}\mathbf{x}. \tag{4}$$

To understand (3), it is important to note that $\tau(t)$ is a dynamic time unit which gives the correlation time over which $p(\mathbf{x};t)$ changes [40]. $\tau(t)$ also serves as the time unit in the statistical space. In addition, its inverse $\frac{1}{\tau(t)}$ quantifies the (average) rate of change of information in time [43]. Finally, $\mathcal{L}$ is a dimensionless distance that quantifies the total information change in time through the root-mean-squared fluctuating energy rate $\sqrt{\mathcal{E}}$ [38]. In a few words, $\mathcal{L}$ gives the total number of statistically different states the random variable $\mathbf{x}$ passes through in time. Throughout this paper, we will call $\sqrt{\mathcal{E}(t)} = \frac{1}{\tau(t)}$ the information velocity.

B. Information length in a Gaussian process

In engineering, it is common to assume that the studied random variable $\mathbf{x}$ is Gaussian or follows a normal distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$, i.e. $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Hence, it is useful to introduce the following result which allows us to compute $\mathcal{L}$ in any given Gaussian process.

Theorem 1 ([17], [44]): The information length of a nth-order Gaussian random variable $\mathbf{x}$ with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ is given by the following integral

$$\mathcal{L}(t)=\int_0^t \sqrt{\mathcal{E}(t_1)}\,\mathrm{d}t_1, \tag{5}$$

$$\mathcal{E}(t_1)=\dot{\boldsymbol{\mu}}(t_1)^T\boldsymbol{\Sigma}^{-1}(t_1)\dot{\boldsymbol{\mu}}(t_1)+\frac{1}{2}\mathbf{tr}\left((\boldsymbol{\Sigma}^{-1}(t_1)\dot{\boldsymbol{\Sigma}}(t_1))^2\right). \tag{6}$$

From (6), we see that $\mathcal{E}(t) \geq 0$ for all t since $\boldsymbol{\Sigma}$ is symmetric and positive definite and the value of $\mathbf{tr}(\mathbf{A}^2) \geq 0$ for any matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$. Note that in [17], the value of $\mathcal{E}(t)$ has been proved to be an alternative for abrupt events detection in linear Gaussian process.

1) Interpretation of the information length: To have a better understanding of the application of $\mathcal{L}$, we can consider computing its value from the mathematical model of a physical process. To this end, consider a simple linear first-order stochastic process given by the following Langevin equation

$$\frac{dx}{dt} = -\gamma(t)\,(x - f(t)) + \xi, \tag{7}$$

where x is a random variable, F is a deterministic force, ξ is a short correlated random forcing such that $\langle \xi(t)\xi(t_1)\rangle = 2D\delta(t-t_1)$ and $\langle \xi(t)\rangle = 0$. Eq. (7) is commonly used to describe the movement of a particle confined to a harmonic potential $V(x) = \frac{1}{2}\gamma(x - f(t))^2$. According to [54], [55], the value of (6) for (7) is given by

$$\mathcal{E}(t) = \frac{\dot{\beta}^2}{2\beta^2} + 2\beta\dot{y}^2, \tag{8}$$

where $\dot{\beta} = \frac{d\beta}{dt}$ and $\dot{y} = \frac{dy}{dt}$, such that

$$p(x;t)=\sqrt{\frac{\beta}{\pi}}\exp\left[-\beta(x-y)\right]^2,\ y(t)=\langle x\rangle = x(0)e^{-G(t)}+F(t)$$

$$\frac{1}{2\beta(t)}=\langle (x(t)-y(t))^2\rangle = \int_0^t 2De^{-2(G(t)-G(t_1))}\,\mathrm{d}t_1,$$

$$F(t)=\int_0^t e^{-(G(t)-G(t_1))}\gamma(t_1)f(t_1)\,\mathrm{d}t_1, \tag{9}$$

and $G(t)=\int_0^t \gamma(t')\,\mathrm{d}t'$.

From (8), we see that $\mathcal{E}(t)$ depends on changes in the variance and mean defined by the dynamics of (7). If we plot the changes of $p(x;t)$ in the three-dimensional space $(t, x, p(x;t))$, the result would be similar to the schematic shown in Fig. 1 where $\sqrt{\mathcal{E}(t)}$ (information velocity) would fluctuate along the path from the initial PDF $p(x,t_0)$ to the final PDF $p(x,t_F)$ describing the speed limit from the statistical distinguishability of the observables (for further details, see [40]). Thus, the time integral of $\sqrt{\mathcal{E}(t)}$, as mentioned before, gives us the total amount of statistical changes (i.e. changes in the mean y and the variance $\frac{1}{2\beta}$) along the path.

As mentioned in Section I, when using metrics like the differential entropy [36], we may not be able to observe all the statistical changes that occur at any time. This is because of the lack of locality since differential entropy focuses on quantifying the difference between two given PDFs without considering the intermediate states [56]. In other words, it only informs us about the changes that affect the overall system evolution. In contrast, IL $\mathcal{L}(t)$ measures local changes that takes place along the path for the system [38], [39]. Furthermore, IL has been invoked as a new way of mapping out an attractor structure and a useful measure that can link stochastic processes and geometry [44], [56]. For instance, if we define $\lim_{t\to\infty}\mathcal{L}(t)$ as the total IL over the entire evolution of a system, for system (8), $\lim_{t\to\infty}\mathcal{L}(t)$ takes its minimum value at the system's stable equilibrium. Moreover, its value increases linearly with respect to the mean position of an initial PDF $p(\mathbf{x};0)$ from the stable equilibrium point [39], [57]. This linear dependence show that the value of IL preserves the linear geometry of the underlying Gaussian process (8). It is important to note that such a feature is lost when using other metrics [44], [56].

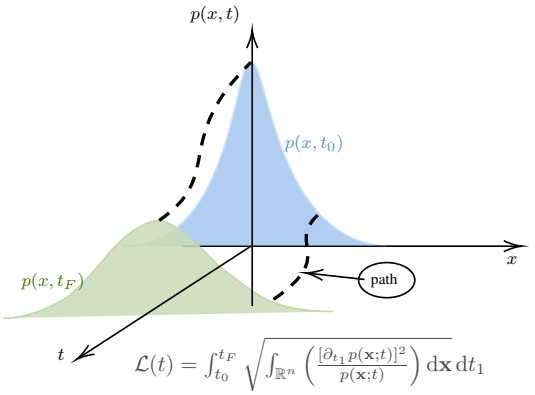


Fig. 1: Schematic of the evolution of $p(x;t)$ over time t. Computing $\mathcal{L}(t)$ gives the total amount of statistical changes on $p(x;t)$ from t_0 to t_F.

Finally, by definition (3), let us emphasise that IL is a model-free and dimensionless metric. Thus, IL can be computed by the estimation of a time-variant PDF of a time series, regardless of the data's nature.

C. The discrete information length

Regarding the nature of our application, it is important to define the methodology that the present work uses to estimate the value of IL for the KE time series in a power system. Specifically, to compute the value of IL for a time series of a random variable x, a discrete version of Equation (3) is applied. Such expression can be defined as follows [56]

$$\mathcal{L}[n]=h\sum_{k=0}^{n}\frac{1}{\mathcal{T}[kT_s]}, \tag{10}$$

$$\frac{1}{\mathcal{T}[kT_s]^2}=\mathcal{E}[kT_s]=\frac{s}{h^2}\sum_j P[j;kT_s]\left(\ln\frac{P[j;(k+1)T_s]}{P[j;kT_s]}\right)^2.$$

where (10) k denotes discrete time with sample period T_s and j denotes partial spatial point. The discrete version of Time-dependent PDF ($p(x;t)$) is denoted by $P[j,kT_s]$; and the time

step and spatial step are denoted as $h = (t_f/n)$ (t_f is the total time) and s respectively. Since, logarithm is used in this calculation (see (10)), when $P[j;kT_s]$ takes the smallest value 0, the $\ln(0)$ gives an undefined value. To overcome this issue the following reasoning is applied:

$$q^2{=}p, \quad \frac{\left[\frac{dp}{dt}\right]^2}{p} = 4\left[\frac{dq}{dt}\right]^2,$$
$$\mathcal{E}[kT_s]{=}\frac{\sum_j 4\left(\sqrt{P[j;(k+1)T_s]}-\sqrt{P[j;kT_s]}\right)^2}{T_s^2},$$
$$\mathcal{L}[n]{=}\sum_{k=0}^{n}\sqrt{\mathcal{E}[kT]}. \tag{11}$$

We use Equation (11) to compute IL of the Kinetic Energy time series. Besides, to estimate each $p[j;kT_s]$, an sliding-window algorithm that moves over the time series is applied. The sliding window can move over the set of real data (see Algorithm 1) or the set of estimated values (see Algorithm 2). The corresponding PDF of the k-th time is found by using the kernel smoothing function method described in [58] (named **ksdensity** in MATLAB®).

1) IL computation from a time series: As mentioned above, implementing an sliding window algorithm is the easiest way to compute IL from a time series. A pseudo-code description for such procedure can be seen in Algorithm 1. From, Algorithm 1, note that at least N measurements are necessary to estimate the first value of $\sqrt{\mathcal{E}}$.

Algorithm 1: Algorithm for the estimation of the Information Length from a time series.

Data: Consider the initial data set: $\mathcal{D}_i := \{D[kT_s]|D[kT_s] \in \mathbb{R} \quad \forall k = i-N, i-N+1, \ldots, i\}$ such that N is the number of samples (window size) in the data set (KE data in our case) sampled with sampling period T_s and $i \in \mathbb{N}$ is the current time. Besides, $n \in \mathbb{N}$ is the final discrete time of the experiment.

Result: The value of $\mathcal{L}[n]$ and $\mathcal{E}[nT_s]$ (see Eq. (11)).

$P_0 =$ **ksdensity**$(\mathcal{D}_0)$ `// Estimate the initial PDF using the function` **`ksdensity`** `from MATLAB® on the initial data` $\mathcal{D}_0$.
$i = 1$
while $i \leq n$ **do**
  $P_i =$ **ksdensity**$(\mathcal{D}_i)$`// Estimate next PDF.`
  `/* Compute IL using Eq. (11). */`
  $[\mathcal{L}[i], \mathcal{E}[iT_s]]$=**InformationLength**$(P_i, P_{i-1})$
  $i = i + 1$
end
return $\mathcal{L}[n], \mathcal{E}[nT_s]$

Finally, the function **InformationLength** in Algorithm 1 corresponds to the programming of the discrete functions in Eq. (11).

IV. IL Forecasting Algorithm

Since IL has been proved to detect abrupt events in the statistical space [17], predicting its future value on real-time can be of great importance in applications like power-systems management. To this end, as a proof of concept, we introduce a forecasting algorithm that implements a basic recursive neural network over a sliding window to estimate the value of the PDF $P[j, kT_s]$ at the discrete time kT_s.

Remark 1: Here, our goal is to provide the basics for applying IL with a forecasting algorithm. We do not intend to provide a full study of the forecasting algorithm methodologies that could be implemented. Instead, we explore a popular method and combine it with the IL metric to analyse the possible implications.

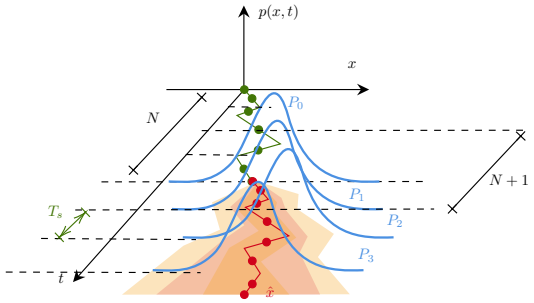


Fig. 2: Schematic describing the data-driven methodology to compute IL through a forecasting algorithm. The method uses the predicted value $\hat{x}$ from the time series x (the present and predicted values presented in green and red colour, respectively) with sample time T_s and includes it in the sliding window of size N to estimate the next value of the PDF (thus, the next value of $\mathcal{E}$).

First, we recall that NNs are a series of architectures and algorithms based on brain behaviour. The goal behind these models is to learn from examples and, in a similar way as human cerebral cells do and to change the interactions between basic units known as neurons [59]. Here, the NN is used as a regressor which incorporates non-linearity and the potential to learn from data. Specifically, the methodology uses a long short-term memory (LSTM) model.

A LSTM is an Recurrent Neural Network (RNN)-based architecture, where the ability to retain part of the information that belongs to the hidden layer can be used for forecasting at particular times [60]. The advantage of LSTM, in relation to common RNN models, is the improvement of the performance over the gradient vanishing problem, which represents a difficulty in the traditional back-propagation algorithm employed for training. However, the comparison of other neural models or the usage of different forecasting methods is out of the scope of this paper.

For our analysis, we have used the implemented LSTM network in MATLAB®, a deep leaning method using 200 hidden units [61]. In Algorithm 2, we provide the pseudocode that computes IL using (11) and the time series forecasted value $\hat{x}$. Figure 2 illustrates the proposed methodology. In brief, as suggested by Fig. 2, in the first prediction Algorithm

2 uses N data (data set $\mathcal{D}_0$) with sample time T_s to train the RNN via the function **TrainingFunc** and estimate the initial PDF P_0 using the function **ksdensity**. Next, it forecasts the next value in the time series $\hat{x}$ using the **PredictandUpdate** function, and adds it to the next sliding window $N+1$ (data set $\mathcal{D}_1$) via the function **UpdateDataSet** to estimate the next PDF P_1. Finally, the value of IL is computed from the initial PDF P_0 and the forecasted P_1. Note that, after the value $\hat{x}$ is predicted by the RNN, we update the network with the real value of the previous prediction using the function **PredictandUpdate** in the next prediction. This process is repeated till we reach the final discrete time n and return the values of $\mathcal{L}[n], \mathcal{E}[nT_s]$.

Algorithm 2: Information length forecasting algorithm.

Data: Consider the normalised initial data set: $\mathcal{D}_i := \{D[kT_s] | D[kT_s] \in \mathbb{R} \quad \forall k = i-N, i-N+1, \ldots, i\}$ such that N is the number of samples (window size) in the data set sampled with sampling period T_s and $i \in \mathbb{N}$ is the current time. Besides, $n \in \mathbb{N}$ is the final discrete time.

Result: The value of $\mathcal{L}[n]$ and $\mathcal{E}[nT_s]$ (see Eq. (11)).

`/* Train the LSTM arquitecture` $\mathcal{N}$. `*/`

$\mathcal{N} =$ **TrainingFunc**$(\mathcal{D}_0)$

$P_0 =$ **ksdensity**$(\mathcal{D}_0)$ `// Estimate the initial PDF using the function` **`ksdensity`** `from MATLAB® on the initial data` $\mathcal{D}_0$.

$i = 1$

while $i \leq n$ **do**

  `/* Insert a new measurement, predict the new value` $\hat{x}$ `and update the network` $\mathcal{N}$. `*/`

  $[\mathcal{N}, \hat{x}] =$**PredictandUpdateNet**$(\mathcal{N}, D[iT_s])$

  $\mathcal{D}_i =$**UpdateDataSet**$(\mathcal{D}_{i-1}, \hat{x})$`// Move the sliding window adding prediction` $\hat{x}$.

  $P_i =$ **ksdensity**$(\mathcal{D}_i)$`// Estimate next PDF.`

  `/* Compute IL using Eq. (11). */`

  $[\mathcal{L}[i], \mathcal{E}[iT_s]] =$**InformationLength**$(P_i, P_{i-1})$

  $i = i + 1$

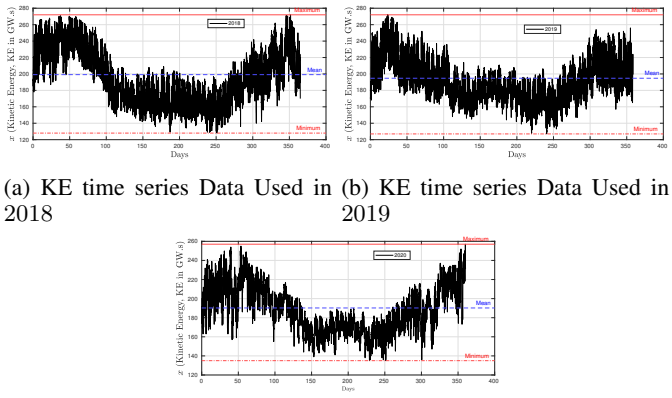
end

return $\mathcal{L}[n], \mathcal{E}[nT_s]$

V. Kinetic Energy Time-series Studied

The NPS is the interconnected and single market area of the Nordic countries that belongs to the region in Northern Europe and the North Atlantic, specifically Sweden, Norway, Finland, and eastern Denmark. For the past ten years, the reduction of rotational inertia has been a concern for the NPS TSOs. One of the short-term measures to ensure the system frequency stability has consisted in installing a measurement and monitoring system to capture the rotational inertia available in the NPS. This monitoring system produces situational awareness alarms to indicate when the levels of the inertia fall below a predefined limit. Using this approach, the TSOs attempt to avoid operational scenarios where the reduced inertia and an $N-1$ contingency criterion can negatively affect the frequency stability.

The NPS used the so-called 'unit commitment method' to calculate the total system rotational inertia, and it is based on adding the rotational inertia of each synchronous machine connected to the system. The TSOs of the NPS have calculated the KE of the NPS in real-time since 2015. This research paper takes advantage of the recorded data of the KE to develop a metric to quantify its variability and unveil hidden information. We utilise the historical data of the KE in the NPS (in GWs)



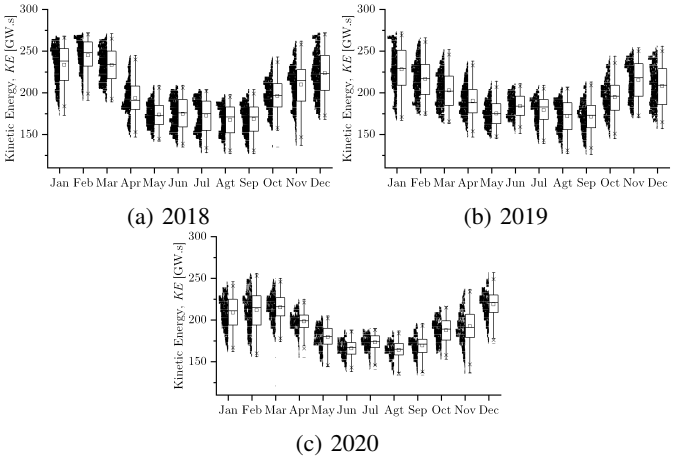
(a) KE time series Data Used in 2018

(b) KE time series Data Used in 2019

(c) KE time series Data Used in 2020

Fig. 3: KE time series of 2018, 2019, and 2020.

recorded during the entire years of 2018, 2019, and 2020. The time series of the KE consists of 44640 samples; it comprises the total data of these years with a resolution of one sample per minute. Figure 3a shows a plot of the KE data where seasonal variation of the KE is evident. For 2018, low values are located during the summer months where the dominant-heating and lighting load is reduced (min: 127 GWs) and as a consequence, the number of generators to cope with the load is minor. As expected, the maximum KE is located during the winter months (max 272 GW.s). Figure 3 reflects the raise of concern about the reduction of KE by comparing the annual averages of KE. Average KE during 2018 had a value of 200 GWs, whereas during 2019, and 2020 is 195 and 190 respectively, representing 5% reduction. A further descriptive using classical statistics of the KE raw data in the form of a monthly box plot is performed in Figure 4 (including distribution of the data as a histogram, the left side of the boxplot). Figure 4 allows identifying the mean and variance per month of the KE during the years studied. From Figure 4a, November 2018 shows the highest variance of the KE with extreme values outside the upper and lower quartiles that almost reach the minimal global inertia reached during summer months. On the other hand, May 2018 exhibits minimum changes in the KE, and it coincides with mild temperature and moderate load in the Nordic countries. For years 2019 and 2020, the histograms show the highest variance during January and November, respectively. In addition, the lowest variance occurs during June, and July respectively. Although these statistical measurements can provide us with information from the KE of the entire month, a day-by-day



(a) 2018 (b) 2019

(c) 2020

Fig. 4: Classic statistical Analysis of the KE Analysed per Month. The charts show a combination of histograms and boxplots per month in the years 2018, 2019 and 2020.

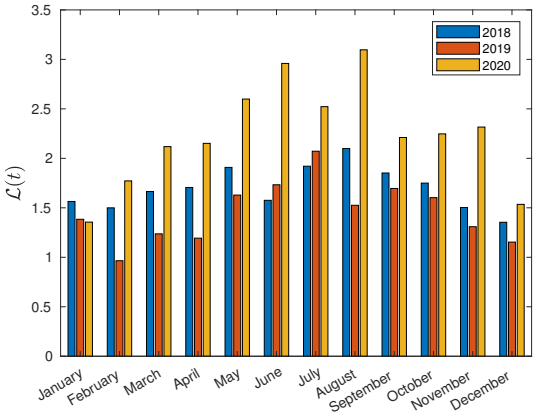


Fig. 5: Monthly IL $\mathcal{L}(t)$ of the KE during 2018, 2019 and 2020.

or hour-per-hour description of the statistical fluctuations is still missing. Here, the IL metric can provide us with such information since, as we have discussed previously, it tracks time series evolution through time-variant PDFs [17], [40].

VI. Information Length Metric Results

In the following, to visualise and analyse the given data, we have assigned the measurements to seasonal groups per year. Typically, in the Nordic countries, spring runs from March/April to May, summer from June to August, fall from September to October/November and winter from November/December to March/February. However, the seasons might have longer winter and summer periods, and the seasons in between, spring and fall, can be shorter. Thus, the demand and power reserves vary accordingly.

From Figure 3, the dramatic effect of seasonality on the KE is perceived, the summer and winter trends are well defined whereas the spring and fall periods can be considered as the decreasing/increasing ramps as the consumption during the months on those seasons decrease/ increase respectively. Additionally, less consumption typically occurs during summer nights. Note that, the load and generation conditions of the KE data are unknown and out of the scope of this work.

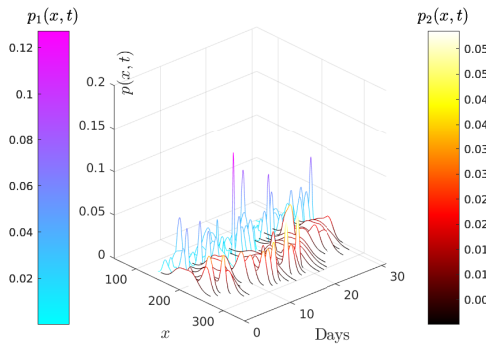
A. Information Length $\mathcal{L}(t)$ per Month during 2018

Figure 5 shows the value of IL $\mathcal{L}(t)$ per month in the years 2018, 2019 and 2020. Here, we start the analysis of the IL metric in the KE from the year 2018. Although, the months with higher load demand (in the Nordic countries are during the winter season due to the lighting and heating households necessity) could be intuitively assigned as the ones with the higher amount of fluctuations. By analysing the value of the IL metric per month during 2018, the highest and the lowest $\mathcal{L}(t)$ are during August, and December, respectively. Thus, indicating that during those months the KE vary the most or remain stiff, respectively. In this regard, from Figure 4a, we can also distinguish an anticorrelation between the variance and IL per month, which persists in the analysis of the two consecutive years. In other words, in comparison to IL, when IL tends to be high the covariance is small. In addition, although in summer the power consumption is reduced (the heating is not needed), the typical load fluctuations during the day show a high $\mathcal{L}(t)$ value. This is because compared to the less variability in winter, where fewer variations in the consumption indicate less variation in the PDF, in summer load fluctuations are more repetitive. This analysis implies that the capacity and reserves need to be adjusted while the day-ahead planning should be carefully optimised. This optimisation process is not analysed in this paper.

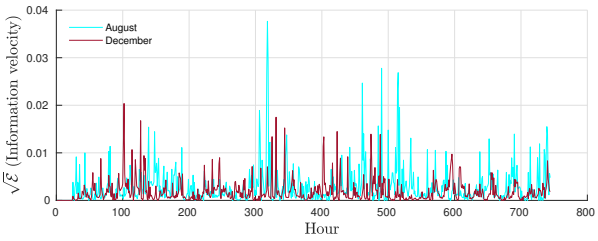
To perform a more detailed analysis of the $\mathcal{L}(t)$ metric, we have selected the months with the highest and lowest IL in the year to create Figure 6, where the evolution over the month of $\mathcal{L}$ and $\sqrt{\mathcal{E}}$ are depicted. When talking about 2018, these are August and December, as we have mentioned before. Figure 6a presents a collection of time-dependent PDFs that describe the KE evolution through the month. Note that, even though all the computations are per hour, the PDFs in Figure 6a are sampled per day to permit a better visualisation of their fluctuations. Besides, Figure 6b shows the value of the information velocity $\sqrt{\mathcal{E}}$ which describes the gradients of the variation of both months PDFs through time. High values and more concurrent peaks during August can be seen, which means that August presents faster and rapid PDF variations. These are depicted by high peaks in $\sqrt{\mathcal{E}}$ on the KE. Lower values of $\sqrt{\mathcal{E}}$ represent slower changes. Let us recall that all quantities are dimensionless.

Furthermore, in Figure 6b, the Information velocity allows to identify the specific days or hours with extreme transitions (abrupt events), as it is seen on the day 11 and, 12 during August, and three subsequent peaks on days 18, 19, and 20. The highest peak in December happens on the third day, and subsequent peaks on the days 12, and 13. Both months tend to have fewer fluctuations at the end of the month.

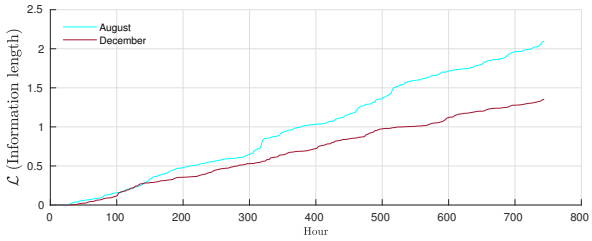
Figure 6c shows the information length $\mathcal{L}(t)$ associated with the results presented in Figure 6b. We see that $\mathcal{L}(t)$ during August increases faster overtime rate than in December,



(a) PDF Evolution per day. August $(p_1(x,t))$ vs December $(p_1(x,t))$ 2018.



(b) Information Velocity per hour. August vs December 2018.



(c) Information Length per hour. August vs December 2018.

Fig. 6: IL Metric Comparison during August and December Months in 2018

specifically in the days 12 and 20 whose rates are considerably ramping up, whereas in December there are fewer fluctuations around the smaller slope. This corroborates how $\mathcal{L}(t)$ can be interpreted as a measure of information changes in PDFs.

B. Information Length $\mathcal{L}(t)$ per Month during 2019 and 2020

To expand the analysis of the IL metric, we explore the KE time series of the next consecutive years 2019 and 2020. Based on Figure 5, for 2019, the months with the highest and lowest variability are July and February, respectively. Besides, for 2020, the months with the highest and lowest variability are August and January, respectively. These months present similar characteristics in comparison to the winter and summer seasons of 2018 mentioned in the previous subsection. Note that February has fewer days than July, for such a reason, we have included a dashed line in Figures 7b and 7c setting the end of February.

For July of 2019, in Figure 7b, the days with the highest variability peaks are the 1st and 14th, which interestingly are at the beginning and middle of the month followed by the increasing consumption. The summer in the Nordic countries is characterised by population movement to summer households which are continuously being modernised, for instance, by including new electricity services. The abrupt and joint activation of these households produce significant changes (strong variations) in power consumption. Two main variability peaks are observed on February 1st and 10th of 2019, while the remainder of the month remains with few strong variations.

Along August of 2020, several KE fluctuations are more visible as seen in Figure 7e. The highest peak is seen on the 11th. However, this month presents a heavily strong variability with high intermittency and irregularity. During January of 2020, several more peaks are seen, especially at the end of the month during the transition to February.

Figures 7c and 7f show a clear difference between the information length $\mathcal{L}(t)$ of the respective months. A month with higher fluctuations will have a higher value at the end of the $\mathcal{L}(t)$ monthly calculation. Thus, the difference between August and January of 2020, since both months are highly fluctuating. The same difference is observed in the final values between July and February of 2019. Similarly, this indicator shows a higher variability for summer and winter of 2020 compared to 2019.

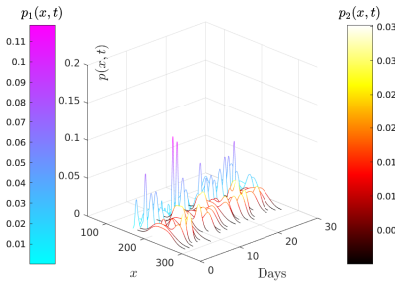
C. Information Length $\mathcal{L}(t)$ from KE Hydropower

As an additional case, we compare the KE variability between the KE provided by a renewable source, the hydropower in this case, versus the total KE during 12 hours of March 27th, 2015. Figure 8a presents both KE time series where we can observe the differences between these KE data. We have also included their estimated PDF evolution in Figure 8b to note the difference on statistical fluctuations. The contribution of hydropower KE to the total KE is approximately 55% during these hours. This is because hydroelectricity is a significant source of energy in the NPS, particularly in Norway, which is almost hydro-dominated. However, this percentage takes into account the other countries in the NPS. Note that wind and solar do not contribute to the KE, and other potential sources are not available. The IL metric applied to these data shows a strong variability change from the 7th to the 9th hour in the KE provided by the hydropower (see Figures 8c and 8d). Since both KE data are correlated, the total KE also presents this variation.

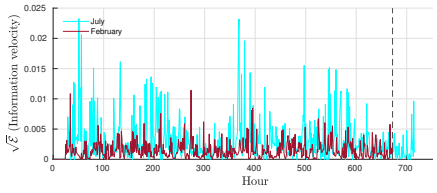
D. Forecasting Results

Now, as a proof of concept, We utilise the probabilistic properties of the KE observables to make predictions in the values of $\mathcal{L}$ and $\sqrt{\mathcal{E}}$. As we have discussed in Section IV and Algorithm 2, the proposed short-term, hour-ahead probabilistic forecast based on LSTM incorporating uses a normalised PDF. Besides, the prediction has an hour-rolling horizon that is being updated with every new estimated value $\hat{x}$ of the KE time series. Here, we test Algorithm 2 using the data of January 2018.

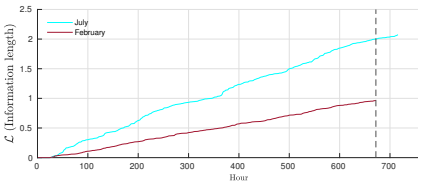
The *a posteriori* multimodal PDFs evolution for the LSTM process are shown in Figure 9a. Note that Figure 9 shows only the second half of the month since the other half of the data have been used for the LSTM training. As a result, we forecast



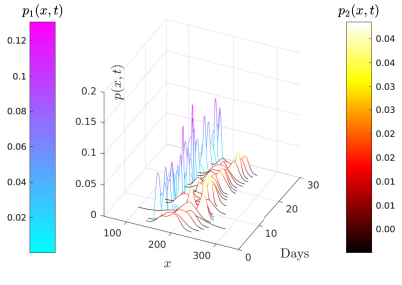
(a) PDF Evolution per day. July $(p_1(x,t))$ vs February $(p_2(x,t))$ 2019.



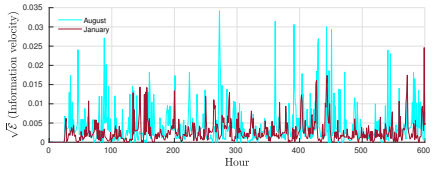
(b) Information Velocity per hour. July vs February 2019.



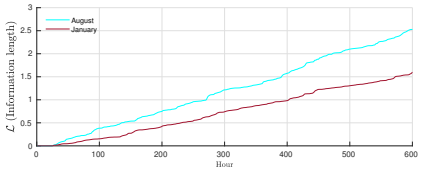
(c) Information Length per hour. July vs February 2019.



(d) PDF Evolution per day. August $(p_1(x,t))$ vs January $(p_2(x,t))$ 2020.

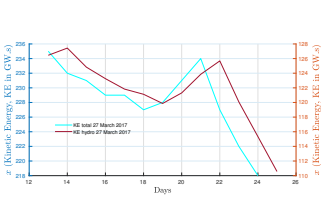


(e) Information Velocity per hour. August vs January 2020.

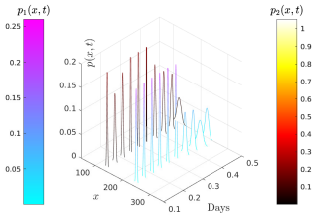


(f) Information Length per hour. August vs January 2020.

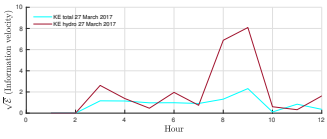
Fig. 7: IL Metric Comparison between the month with the minimum and maximum $\mathcal{L}$ in 2019 and 2020



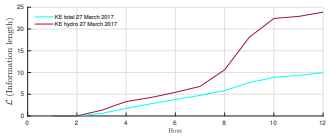
(a) time series per hour of hydro vs total KE (27 March 2015).



(b) PDF Evolution per day of hydro vs total KE (27 March 2015).



(c) Information Velocity per hour of hydro vs total KE (27 March 2015).



(d) Information Length per hour of hydro vs total KE (27 March 2015).

Fig. 8: IL Metric Comparison of total KE vs hydro KE from 12:00-24:00 hrs at 27 March 2015. Here, $N = 2$ and $T_s = 1hr$.

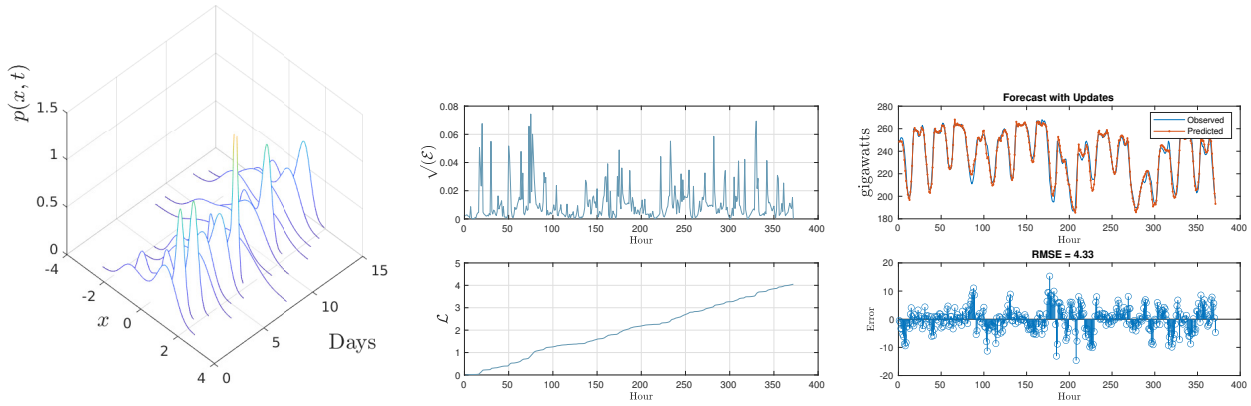
the value of $\sqrt{\mathcal{E}}$ only for the second half of the month. In this regard, Figure 9b shows the forecasted $\mathcal{E}(t)$ and $\mathcal{L}(t)$ metrics. Here, we note that the variability is maintained with various gradients during the month. The highest predicted variability (abrupt event) value is observed during the day 26, however, the general variability values are similar, meaning the same KE trend, a high effort of the system to maintain the heavy consumption since January is in the winter period.

Finally, to quantify the forecasting error, the Root Mean Square Error (RMSE) between the prediction and the observed data is used as a forecasting index. The results of the index are shown in Figure 9c. Although the model gives the highest error of 15%, the forecast data perform significantly well having an RMSE of 4.33%.

VII. Discussion

The potential growth of non-synchronous generation in power systems worldwide is potentially leading to a KE reduction in the system requiring a deep understanding of the trends and fluctuations within months, hours or seasons. The development and application of new metrics can help to design or adjust the generators or controllers with the ability to respond to a peak seasonal demand. To this end, we utilise the IL metric to the behaviour of the KE during the year. Specifically, we measure the time series fluctuations showing the possibility to detect extreme and abrupt events in the system. A clear advantage of the proposed technique is that the availability of specific demand or generation profiles is not required. However, patterns or aggregated annual energy consumption data of the system will potentially help to clarify further detailed aspects when using the proposed metric. This aspect is constrained to availability since such data might require security clearance from TSOs point of view. As the implementation of KE by TSOs is a recently developed monitoring system, the collection of further data will be needed to perform a more exhaustive analysis.

Although, operating at full capacity for long periods of time is unusual for a TSO, anticipatory behaviour and innovative tools that contribute to gain insights on the system are needed to incorporate more flexibility to support grid planning for future irregular or rare events. Moreover, KE analysis, as a



(a) Predicted PDF evolution for the second half of January 2018.

(b) Predicted Information Velocity $\mathcal{E}(t)$ and information Length $\mathcal{L}(t)$ for the second half of January 2018.

(c) Prediction vs observed data error.

Fig. 9: Analysis of the Prediction of Information Length using Algorithm 2 where the first half of January 2018 has been used to train the NN and the second half of the month is predicted.

relatively new power systems topic, requires further understanding to provide operator planning tools that quantify, and extract relevant data.

It is important to underline that traditional statistical analysis should not be understood as erroneous but as complementary to the probabilistic metrics presented in this paper. Both can provide relevant information metrics of the KE periodic variations.

Nonetheless, as we have shown, the IL metric can track the variability through the time series evolution via time-dependent PDFs. This gives the IL metric an advantage over traditional statistical analysis. For the KE annual cases, we consider it more valuable to understand the day-by-day variability since a TSO could use this for its day-ahead operations. Even though we have analysed the highest and lowest variability months of the KE data per year, the proposed metric can be used within other ranges of time.

VIII. Conclusions and Future Work

We have presented the application of the information length metric for the annual Kinetic Energy time series quantification in the NPS during 2018, 2019, and 2020. The proposed metric allows us to identify the variability along the seasons and evaluate the months where the KE fluctuations have abrupt events and the minimum variability. Besides, The metric enables us to detect daily gradient variations that are otherwise difficult to measure for a TSO. Additionally, the proposed forecasting algorithm uses the metric to predict the future KE fluctuations in an hour-ahead horizon, enabling TSOs to adjust the generator's settings accordingly.

Future work will investigate other possible probabilistic and dynamic metrics to measure power system related signals with highly intermittent big data. For instance, we plan to use information length to measure the information flow between the elements in the system by considering its causality properties [42]. We also see that the integration of ML and the IL needs to be explored in more detail. Finally, future work will also focus on studying the practicality of the forecasting algorithm that was presented in this paper, by comparing its performance with other well-assessed forecasting techniques.

Acknowledgment

E.K. acknowledges the Leverhulme Trust Research Fellowship (RF-2018-142-9).

References

[1] T. Ackermann, T. Prevost, V. Vittal, A. J. Roscoe, J. Matevosyan, and N. Miller, "Paving the way: A future without inertia is closer than you think," *IEEE Power and Energy Magazine*, vol. 15, no. 6, pp. 61–69, 2017.

[2] P. Betancourt-Paulino, H. R. Chamorro, M. Soleimani, F. Gonzalez-Longatt, V. K. Sood, and W. Martinez, "On the perspective of grid architecture model with high TSO-DSO interaction," vol. n/a, no. n/a. [Online]. Available: http://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/esi2.12003

[3] H. Gu, R. Yan, and T. Saha, "Review of system strength and inertia requirements for the national electricity market of australia," *CSEE Journal of Power and Energy Systems*, vol. 5, no. 3, pp. 295–305, 2019.

[4] B. Berry, "Inertia Estimation Methodologies vs Measurement Methodology: Impact on System Operations." CIGRE. [Online]. Available: https://e-cigre.org/publication/SYMP_AAL_2019-symposium-aalborg-2019

[5] D. P. Chassin, Z. Huang, M. K. Donnelly, C. Hassler, E. Ramirez, and C. Ray, "Estimation of wecc system inertia using observed frequency transients," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 1190–1192, 2005.

[6] C. Gehbauer, A. Stosic, and D. Black, "Synthetic caiso frequency regulation signal," in *2020 IEEE Power Energy Society General Meeting (PESGM)*, 2020, pp. 1–5.

[7] I. Green, "Caiso experience with impact of high penetration of renewable resources on short-term voltage stability," in *2015 IEEE Power Energy Society General Meeting*, 2015, pp. 1–18.

[8] H. Thiesen, C. Jauch, and A. Gloe, "Design of a System Substituting Today's Inherent Inertia in the European Continental Synchronous Area," vol. 9, no. 8, p. 582. [Online]. Available: https://www.mdpi.com/1996-1073/9/8/582

[9] L. Toma, M. Sanduleac, D. O. Sidea, C. Stanescu, C. Diaconu, M. Albu, and A. M. Dumitrescu, "Frequency dynamics in the romanian power system under large perturbations," in *2020 55th International Universities Power Engineering Conference (UPEC)*, 2020, pp. 1–6.

[10] A. Díaz-García, C. Izquierdo, A. Cordón, F. Rodríguez, and R. Rivas, "Frequency stability model for energy transition studies in spain," in *2020 IEEE/PES Transmission and Distribution Conference and Exposition (T D)*, 2020, pp. 1–6.

[11] M. Persson and P. Chen, "Kinetic energy estimation in the nordic system," in *2018 Power Systems Computation Conference (PSCC)*, 2018, pp. 1–5.

[12] P. Du, N. V. Mago, W. Li, S. Sharma, Q. Hu, and T. Ding, "New ancillary service market for ercot," *IEEE Access*, vol. 8, pp. 178 391–178 401, 2020.

[13] F. Gonzalez-Longatt, M. N. Acosta, H. R. Chamorro, and D. Topic, "Short-term kinetic energy forecast using a structural time series model: Study case of nordic power system," in *2020 International Conference on Smart Systems and Technologies (SST)*, 2020, pp. 173–178.

[14] (2021) Short-term system inertia forecast — nia_ngso0020 — smarter networks. [Online]. Available: https://www.smarternetworks.org/project/nia_ngso0020

[15] P. Imris, M. Bradley, G. Taylor, M. Gordon, and Y. Li, "Enhanced visualisation of fast frequency phenomena as exhibited in the gb transmission system," in *2019 54th International Universities Power Engineering Conference (UPEC)*, 2019, pp. 1–6.

[16] Y. Bian, H. Wyman-Pain, F. Li, R. Bhakar, S. Mishra, and N. P. Padhy, "Demand side contributions for system inertia in the gb power system," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 3521–3530, 2018.

[17] A.-J. Guel-Cortez and E. Kim, "Information geometric theory in the prediction of abrupt changes in system dynamics," *Entropy*, vol. 23, no. 6, p. 694, 2021.

[18] R. E. Spinney and I. J. Ford, "Fluctuation relations: A pedagogical overview," *arXiv preprint arXiv:1201.6381*, 2012.

[19] M. Droz and A. Pękalski, "On the role of fluctuations in the modeling of complex systems," *Frontiers in Physics*, vol. 4, p. 38, 2016.

[20] E. Kim, "Information geometry, fluctuations, non-equilibrium thermodynamics, and geodesics in complex systems," *Entropy*, vol. 23, no. 11, p. 1393, 2021.

[21] Y. Xiao, X. Lin, and Y. Wen, "A framework for assessing the inertia distribution of power systems," in *2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2)*, 2019, pp. 1418–1423.

[22] F. M. Mele, A. Ortega, R. Zárate-Miñano, and F. Milano, "Impact of variability, uncertainty and frequency regulation on power system frequency distribution," in *2016 Power Systems Computation Conference (PSCC)*, 2016, pp. 1–8.

[23] K. Zeng, H. Li, M. Zeng, and P. Liu, "Power system risk security assessment based on maximum information entropy principle," in *2014 International Conference on Power System Technology*, 2014, pp. 421–426.

[24] X. Cheng, M. Li, and H. Zhao, "Fuzzy security assessment of entropy-weight coefficient method applied in electric power information systems," in *2007 International Power Engineering Conference (IPEC 2007)*, 2007, pp. 784–787.

[25] Y. Jia and Z. Xu, "Risk assessment based on information entropy of cascading failure in power systems," in *2012 IEEE Power and Energy Society General Meeting*, 2012, pp. 1–5.

[26] T. Yi, S. Qian, C. Ning, and Z. Lingzhi, "A novel combination model for wind power forecasting with error evaluation parameter based on cross entropy theory," in *2012 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*, 2012, pp. 233–237.

[27] Q. Bian, Y. Qiu, W. Wu, H. Xin, and X. Fu, "Generation dispatch method based on maximum entropy principle for power systems with high penetration of wind power," *Journal of Modern Power Systems and Clean Energy*, vol. 6, no. 6, pp. 1213–1222, 2018.

[28] S. Sinha, P. Sharma, U. Vaidya, and V. Ajjarapu, "On information transfer-based characterization of power system stability," *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 3804–3812, 2019.

[29] D. S. Javan and H. Rajabi Mashhadi, "Wide-area security assessment based on informative variables of power system," vol. 129, p. 106760. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142061520343052

[30] Z. J. Bao, Y. J. Cao, G. Z. Wang, and L. J. Ding, "Analysis of cascading failure in electric grid based on power flow entropy," vol. 373, no. 34, pp. 3032–3040. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0375960109007610

[31] C. Li, L. Xiao, Y. Cao, Q. Zhu, B. Fang, Y. Tan, and L. Zeng, "Optimal allocation of multi-type facts devices in power systems based on power flow entropy," *Journal of Modern Power Systems and Clean Energy*, vol. 2, no. 2, pp. 173–180, 2014.

[32] S. Zhang, Y. Li, T. Jin, H. Yu, G. Yang, and Z. Liu, "Analysis and research on power flow entropy index of power system," in *2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, 2018, pp. 1150–1153.

[33] P. Zegers, "Fisher information properties," *Entropy*, vol. 17, no. 7, pp. 4918–4939, 2015.

[34] A. Ly, M. Marsman, J. Verhagen, R. P. Grasman, and E.-J. Wagenmakers, "A tutorial on fisher information," *Journal of Mathematical Psychology*, vol. 80, pp. 40–55, 2017.

[35] B. R. Frieden, *Science from Fisher information*. Citeseer, 2004, vol. 974.

[36] J. V. Michalowicz, J. M. Nichols, and F. Bucholtz, *Handbook of differential entropy*. Crc Press, 2013.

[37] T. Van Erven and P. Harremos, "Rényi divergence and kullback-leibler divergence," *IEEE Transactions on Information Theory*, vol. 60, no. 7, pp. 3797–3820, 2014.

[38] E. Kim, "Investigating information geometry in classical and quantum systems through information length," *Entropy*, vol. 20, no. 8, 2018.

[39] E. Kim and R. Hollerbach, "Signature of nonlinear damping in geometric structure of a nonequilibrium process," *Physical Review E*, vol. 95, no. 2, p. 022137, 2017.

[40] S. B. Nicholson, L. P. García-Pintos, A. del Campo, and J. R. Green, "Time–information uncertainty relations in thermodynamics," vol. 16, no. 12, pp. 1211–1215. [Online]. Available: https://doi.org/10.1038/s41567-020-0981-y

[41] J. Heseltine and E. Kim, "Comparing information metrics for a coupled ornstein–uhlenbeck process," *Entropy*, vol. 21, no. 8, p. 775, 2019.

[42] E. Kim and A.-J. Guel-Cortez, "Causal information rate," *Entropy*, vol. 23, no. 8, p. 1087, 2021.

[43] E. Kim, "Information geometry and non-equilibrium thermodynamic relations in the over-damped stochastic processes," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2021, no. 9, p. 093406, 2021.

[44] A.-J. Guel-Cortez and E. Kim, "Information Length Analysis of Linear Autonomous Stochastic Processes," vol. 22, no. 11, p. 1265. [Online]. Available: https://www.mdpi.com/1099-4300/22/11/1265

[45] H. R. Chamorro, A. D. Orjuela-Cañón, D. Ganger, M. Persson, F. Gonzalez-Longatt, L. Alvarado-Barrios, V. K. Sood, and W. Martinez, "Data-Driven Trajectory Prediction of Grid Power Frequency Based on Neural Models," vol. 10, no. 2, p. 151. [Online]. Available: https://www.mdpi.com/2079-9292/10/2/151

[46] G. Zhang, B. Xu, H. Liu, J. Hou, and J. Zhang, "Wind power prediction based on variational mode decomposition and feature selection," *Journal of Modern Power Systems and Clean Energy*, pp. 1–10, 2020.

[47] E. Gonzalez-Romera, M. Jaramillo-Moran, and D. Carmona-Fernandez, "Monthly electric energy demand forecasting based on trend extraction," *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp. 1946–1953, 2006.

[48] X. Zheng, X. Ran, and M. Cai, "Short-term load forecasting of power system based on neural network intelligent algorithm," *IEEE Access*, pp. 1–1, 2020.

[49] A. U. Haque, M. H. Nehrir, and P. Mandal, "A hybrid intelligent model for deterministic and quantile regression approach for probabilistic wind power forecasting," *IEEE Transactions on Power Systems*, vol. 29, no. 4, pp. 1663–1672, 2014.

[50] D. Lee, H. Shin, and R. Baldick, "Bivariate probabilistic wind power and real-time price forecasting and their applications to wind power bidding strategy development," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6087–6097, 2018.

[51] F. Mei, J. Gu, J. Lu, J. Lu, J. Zhang, Y. Jiang, T. Shi, and J. Zheng, "Day-ahead nonparametric probabilistic forecasting of photovoltaic power generation based on the lstm-qra ensemble model," *IEEE Access*, vol. 8, pp. 166 138–166 149, 2020.

[52] C. Wan, J. Lin, Y. Song, Z. Xu, and G. Yang, "Probabilistic forecasting of photovoltaic generation: An efficient statistical approach," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2471–2472, 2017.

[53] P. Makolo, R. Zamora, and T. T. Lie, "Heuristic inertia estimation technique for power networks with high penetration of res," in *2020 2nd International Conference on Smart Power Internet Energy Systems (SPIES)*, 2020, pp. 356–361.

[54] J. Heseltine and E. Kim, "Novel mapping in non-equilibrium stochastic processes," *Journal of Physics A: Mathematical and Theoretical*, vol. 49, no. 17, p. 175002, 2016.

[55] E. Kim, U. Lee, J. Heseltine, and R. Hollerbach, "Geometric structure and geodesic in a solvable model of nonequilibrium process," *Physical Review E*, vol. 93, no. 6, p. 062127, 2016.

[56] J. Heseltine and E. Kim, "Comparing Information Metrics for a Coupled Ornstein–Uhlenbeck Process," vol. 21, no. 8, p. 775. [Online]. Available: https://www.mdpi.com/1099-4300/21/8/775

[57] R. Hollerbach, D. Dimanche, and E. Kim, "Information geometry of nonlinear stochastic systems," *Entropy*, vol. 20, no. 8, p. 550, 2018.

[58] A. W. Bowman and A. Azzalini, *Applied smoothing techniques for data analysis: the kernel approach with S-Plus illustrations*. OUP Oxford, 1997, vol. 18.

[59] S. Haykin, *Neural Networks and Learning Machines*, ser. Neural networks and learning machines. Prentice Hall, 2009, no. v. 10. [Online]. Available: https://books.google.com.co/books?id=K7P36lKzI{_}QC

[60] Y. Hua, Z. Zhao, R. Li, X. Chen, Z. Liu, and H. Zhang, "Deep learning with long short-term memory for time series prediction," *IEEE Communications Magazine*, vol. 57, no. 6, pp. 114–119, 2019.

[61] MathWorks. (2021) Time series forecasting using deep learning. [Online]. Available: https://uk.mathworks.com/help/deeplearning/ug/time-series-forecasting-using-deep-learning.html#TimeSeriesForecastingUsingDeepLearningExample-4