Peer-reviewed version available at Energies 2017, 10, 1934; doi:10.3390/en10121934

Article

Principal mismatch patterns across a simplified highly renewable European electricity network

Mads Raunbak [1,‡], Timo Zeyer [2,‡], Kun Zhu [3,‡] and Martin Greiner [3,‡*]

[1] Aarhus University, Department of Physics and Astronomy, Ny Munkegade 120, 8000 Aarhus C, Denmark
[2] DONG Energy A/S, Teknikerbyen 25, 2830 Virum, Denmark
[3] Aarhus University, Department of Engineering, Inge Lehmanns Gade 10, 8000 Aarhus C, Denmark
* Correspondence: greiner@eng.au.dk
‡ These authors contributed equally to this work.

Abstract: Due to its spatio-temporal variability, the mismatch between the weather and demand patterns challenges the design of highly renewable energy systems. A principal component analysis is applied to a simplified networked European electricity system with a high share of wind and solar power generation. It reveals a small number of important mismatch patterns, which explain most of the system's required backup and transmission infrastructure. Whereas the first principal component is already able to reproduce most of the temporal mismatch variability for a solar dominated system, a few more principal components are needed for a wind dominated system. Due to its monopole structure the first principal component causes most of the system's backup infrastructure. The next few principal components have a dipole structure and dominate the transmission infrastructure of the renewable electricity network.

Keywords: renewable energy networks; principal component analysis; large-scale integration of renewables; wind power; solar power; super grid; energy system design

1. Introduction

The weather is the driving force in a highly renewable energy system. The planning for such systems requires an in-depth understanding of the variable mismatch between weather and demand patterns on multiple time and length scales [1,2]. The spectrum of prominent temporal fluctuations of wind and solar power generation ranges from seconds to years [3–6], and impacts the design and operation of power systems [7–13]. The planning of large-scale energy systems is also affected by the spatial variability of the weather. Wind power generation is correlated up to a length scale of about 500km [4,14–16]. Approaches based on spatial correlations, like the optimal portfolio theory [17–19] and the copula method [20], are used for systemic resource assessments of renewables and for analysis of national and continental power grids.

Weather-driven network modelling represents a more direct approach to obtain estimates on the required backup infrastructure of highly renewable large-scale energy systems [21–29]. Weather data covering multiple years are converted into prospective wind and solar power generation with good spatial and temporal resolution [21,22,30–32], and are then used as the driving force in networked electricty system models. This modelling approach has produced estimates on the required amount of conventional backup power plants, transmission lines and storage [22–28,33–38]. Also the optimization of levelized system cost of energy has been addressed and has led to new design concepts like the optimal heterogeneity and the benefit of cooperation [29,39–42].

The weather-driven network modelling takes the high-dimensional weather and demand data "as is", and "computes everything away" by brute force. The resulting infrastructure estimates might not depend on all of the mismatch data. To a large part they might only depend on a few dominant mismatch patterns. This is the central topic of this paper. The central question is: what are the dominant spatio-temporal mismatch patterns between the renewable power generation and the load

which determine most of the required infrastructure of a highly renewable European electricity system? The central method to apply is a principal component analysis (PCA) [43].

Spatio-temporal pattern analyses like the PCA are well known in meteorology. In [44] a PCA has been used to extract the main pattern of variability of wind power generation in Germany, and a wind power forecasting model has been implemented with this approach. In [45] a PCA has been applied to distributed wind power data from Irish wind farms and used as a multivariate dimension reduction scheme to obtain potential wind power production costing simulation efficiency gains, when compared to exhaustive multi-year time series load flow investigations.

This paper has the following structure: Sect. 2 describes a simplified weather-driven European electricity system with a high share of wind and solar power generation, including several of its infrastructure measures. A short description of the PCA is also given. Sect. 3 presents the results on the dominating principal mismatch components and their dependence on the share between wind and solar power generation. Sect. 4 reveals how the principal components contribute to the backup and transmission infrastructure estimates. A conclusion and outlook is given in Sect. 5.

2. Modelling and Methods

2.1. Modelling of a simplified highly renewable European electricity network

We follow the simplified approach introduced by [22,23,26,46] to model a highly renewable European electricity system. The key variable is the mismatch

$$\Delta_n(t) = G_n^R(t) - L_n(t) \tag{1}$$

between the renewable power generation $G_n^R(t)$ and the load $L_n(t)$ at time t for country n in Europe. With a flipped sign the mismatch is often also denoted as the residual load. The renewable power generation

$$G_n^R(t) = G_n^W(t) + G_n^S(t) \tag{2}$$

is composed of wind and solar power generation only; other forms of renewable power generation are discarded. The average wind and solar power generation amount to

$$\langle G_n^W \rangle = \alpha_n \gamma_n \langle L_n \rangle \ , \quad \langle G_n^S \rangle = (1-\alpha_n) \gamma_n \langle L_n \rangle \ , \tag{3}$$

and define two design parameters, the renewable penetration γ_n with respect to the average load $\langle L_n \rangle$ and the renewable mix α_n between the average wind and solar power generation.

Penetration parameters around one describe highly renewable electricity networks. For simplicity, we choose $\gamma_n = 1$ for all countries. With this setting the average renewable power generation $\langle G_n^R \rangle = \langle L_n \rangle$ is equal to the average load. For more simplicity, the renewable mixing parameter $\alpha_n = \alpha$ is also chosen to be the same for all countries. These homogeneous layouts have been first discussed in [22,23,33]. Recently also heterogeneous layouts with differing renewable penetration and mixing parameters between the countries have been investigated [41].

The time series $G_n^W(t)$, $G_n^S(t)$, $L_n(t)$ for the country-aggregated wind power generation, solar power generation and load are taken from Refs. [22,23], where wind velocity and solar irradiation data with good spatio-temporal resolution have been converted into wind and solar power generation; see also Refs. [30–32] for related approaches. The data cover the years 2000 - 2007 with hourly resolution for all European countries. It is assumed that this data is also statistically representative for future years. The average loads $\langle L_n \rangle$ have been fixed to the values of the year 2007 and are assumed to remain constant for future years. They are illustrated in Figure 1. Note that the average loads set the absolute scales for the simplified modelling of a highly renewable European electricity network. If future values should increase or decrease, but such that the relative load strengths between the countries remain unchanged, then the infrastructure measures (to be discussed in the next subsection) will linearly scale

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 September 2017 doi:10.20944/preprints201709.0092.v1

Peer-reviewed version available at Energies 2017, 10, 1934; doi:10.3390/en10121934

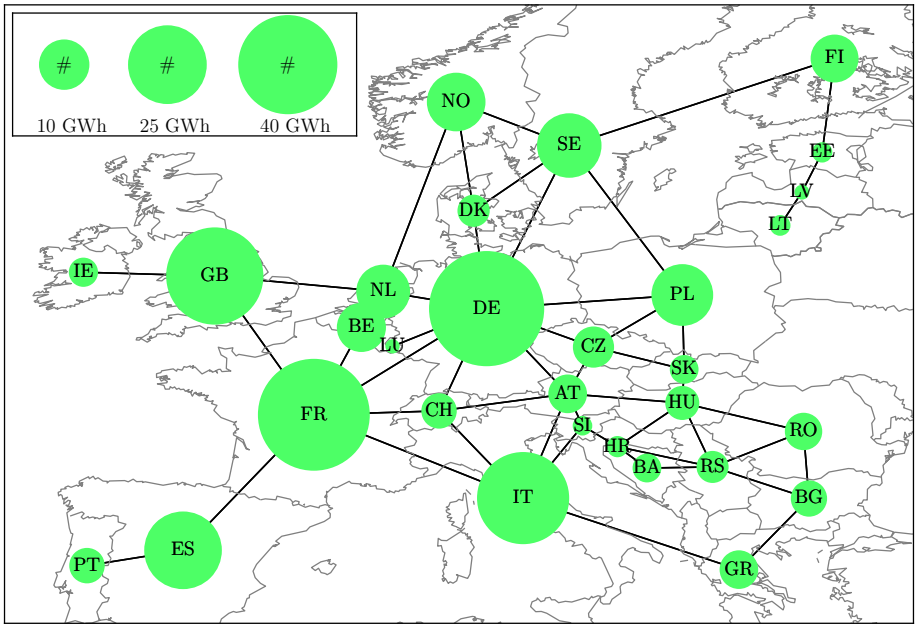

Figure 1. Simplified European electricity network, where countries are represented as nodes and linked by interconnecting transmission lines. The nodal disc areas are proportional to the average loads of the countries.

with changes of the average loads. Figure 1 also shows the interconnecting transmission lines between the countries.

With the chosen setting $\gamma_n = 1$ the average mismatch is $\langle \Delta_n \rangle = 0$. However, most of the times the actual mismatch $\Delta_n(t)$ is non-zero. It is either positive when the wind and the solar radiation are strong, or negative when they are weak. The overall electricity system needs to respond to these mismatch fluctuations. In simplified form, the response can be written as

$$\Delta_n(t) = B_n(t) + P_n(t) \; . \tag{4}$$

The nodal balancing $B_n(t)$ determines the backup power generation and the curtailment,

$$G_n^B(t) = -\min\left(B_n(t), 0\right) \; , \quad C_n(t) = \max\left(B_n(t), 0\right) \; . \tag{5}$$

The nodal injection $P_n(t)$ into the networked system describes the exports ($P_n > 0$) and imports ($P_n < 0$), and has to fulfill $\sum_n P_n(t) = 0$. The nodal injections determine the power flows

$$F_l(t) = \sum_n H_{ln} P_n(t) \tag{6}$$

on the interconnecting transmission lines l between the countries. The linear relationship between the flows and the injections is described by the matrix H_{ln} of the power transfer distribution factors [47], which are constructed from the Moore-Penrose pseudo inverse of the graph Laplacian of the underlying electricity network topology [48]. In principle, more terms can be added to the right hand side of the nodal response equation (4), like for example storage, but for the present purpose we leave this equation as is.

The response (4) allows for various schemes which divide differently between balancing and injection. For sure the simplest scheme is $P_n(t) = 0$ for all nodes and all times, but has the disadvantage that each node has to fully balance its mismatch alone without the benefits of importing and exporting across the network. Market schemes could be used to settle the dispatch between renewable and

backup power generation, but due to the simplicity of the current modelling approach this would be like shooting on pigeons with canons. We adopt the simple scheme of synchronised nodal balancing

$$B_n(t) = \frac{\langle L_n \rangle}{\sum_m \langle L_m \rangle} \sum_k \Delta_k(t) \,, \tag{7}$$

which has been proposed in [46] to minimise the overall backup capacities. It distributes the actual overall mismatch onto all nodes in proportion to their average load.

2.2. Infrastructure measures

If the mismatch (1) was zero for all times and all countries, the infrastructure response (4) would simply be zero, and no backup power plants and no transmission lines would be required. Of course, this is almost surely never the case. This sets the stage for the mismatch variance to become a first infrastructure measure:

$$\begin{aligned} \mathrm{Var}(\vec{\Delta}) &= \sum_{n=1}^{N} \left\langle (\Delta_n - \langle \Delta_n \rangle)^2 \right\rangle & (8) \\ &= \mathrm{Var}(\vec{B}) + \mathrm{Var}(\vec{P}) + 2\,\mathrm{cov}(\vec{B}, \vec{P}) \,. & (9) \end{aligned}$$

The components of the mismatch / balancing / injection vectors are given by the nodal mismatches Δ_n / balancings B_n / injections P_n. The second step, which uses (4), relates the mismatch variance to the variances of the balancings and the injections as well as to their covariance. This emphasises that a small mismatch variance would require a small backup and transmission infrastructure, and a large variance would lead to a bigger infrastructure.

The mismatch variance (8) represents only a rough infrastructure measure. More specific measures are given by the total backup energy, the total backup capacities and the total transmission capacities:

$$\begin{aligned} E^B &= \sum_{n=1}^{N} \left\langle G_n^B \right\rangle \,, & (10) \\ \mathcal{K}^B &= \sum_{n=1}^{N} \max_t{}_q(G_n^B) \,, & (11) \\ \mathcal{K}^T &= \sum_{l=1}^{L} \max_t{}_q(|F_l|) d_l \,. & (12) \end{aligned}$$

The backup capacities and the transmission capacities are defined via the $q = 0.99$ quantiles of the time-sampled distributions $p_n(G_n^B)$ and $p_l(|F_l|)$, respectively. N and L represent the number of nodes and links illustrated in Figure 1. The link length is denoted as d_l and is approximated as the distance between country capitals.

2.3. Principal Component Analysis

We follow the notation of Ref. [43] and sketch the main steps of the Principal Component Analysis (PCA). The mismatch vector is rescaled to

$$\vec{x}(t) = c\,\vec{\Delta}(t) = \sum_{n=1}^{N} c\,\Delta_n(t)\vec{e}_n \,, \tag{13}$$

where the normalisation c is chosen such that $\langle \vec{x} \cdot \vec{x} \rangle = 1$. This turns $c = c(\alpha)$ into a function of the mixing parameter α. Note that the new variable $\vec{x}$ is usually defined with the centred mismatch $\vec{\Delta} - \langle \vec{\Delta} \rangle$, but with our choice $\gamma = 1$ for the renewable penetration parameter the average mismatch

is zero. The normalised mismatch vector (13) is represented in an N-dimensional coordinate system, where the axes with unit vectors $\vec{e}_n$ correspond to the countries.

The PCA changes this coordinate system by rotation. The normalised mismatch vector

$$\vec{x}(t) = \sum_{k=1}^{N} a_k(t)\vec{p}_k \; . \tag{14}$$

is then expressed in terms of the rotated unit vectors $\vec{p}_k$. Those are determined from the diagonalisation

$$\hat{R}\,\vec{p}_k = \lambda_k \vec{p}_k \tag{15}$$

of the covariance matrix

$$R_{nm} = \langle x_n x_m \rangle \; . \tag{16}$$

Since the covariance matrix is real and symmetric, the eigenvectors are orthonormal, i.e. $\vec{p}_k \cdot \vec{p}_l = \delta_{kl}$, and all eigenvalues $\lambda_k \geq 0$ are real and positive. Furthermore, $\sum_{k=1}^{N} \lambda_k = 1$ because of the normalisation introduced in Eq. (13).

The temporal amplitudes $a_k(t) = \vec{x}(t) \cdot \vec{p}_k$ have the properties $\langle a_k \rangle = 0$ and $\langle a_k a_l \rangle = \lambda_k \delta_{kl}$. The total variance (8) of the mismatch can now be written as

$$\mathrm{Var}(\vec{\Delta}) = \frac{1}{c^2} \sum_{k=1}^{N} \langle a_k^2 \rangle = \frac{1}{c^2} \sum_{k=1}^{N} \lambda_k \; . \tag{17}$$

Given the ordering $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$, the last equation reveals that the first eigenvector $\vec{p}_1$ with the largest eigenvalue λ_1 contributes most to the total variance. This implies that the mismatch can be approximated by a truncation to the first K eigenvectors with the largest eigenvalues:

$$\vec{\Delta}(t) \approx \frac{1}{c} \sum_{k=1}^{K} a_k(t)\vec{p}_k \; . \tag{18}$$

This approximation explains the naming of the PCA: the K eigenvectors are called the principal components and carry most of the overall variance.

3. Results: principal mismatch components

The dependence of the mismatch variance (8) on the renewable mix α is shown as the thick black curve in Figure 2. The renewable penetration has been set to $\gamma = 1$ for all European countries. The variance takes a minimum at $\alpha = 0.7$. It is slightly larger in the wind-only limit $\alpha = 1$ and significantly larger in the solar-only limit $\alpha = 0$. This can be seen as a rough indication that the response infrastructure becomes smallest at intermediate renewable mixes, and agrees nicely with the results on the other infrastructure measures (10) - (12) obtained in [23,46].

The eigenvalues resulting from the Principal Component Analysis reveal the dominant contributions explaining most of the mismatch variance. They are shown in Figure 3, again as a function of the renewable mix. In the solar-only limit $\alpha = 0$ the largest eigenvalue is $\lambda_1 = 0.90$. The three largest eigenvalues add up to 0.96, explaining 96% of the mismatch variance. Not much changes for renewable mixes up to about $\alpha \approx 0.3$. After this the largest eigenvalue decreases continuously, until it reaches a value $\lambda_1 = 0.45$ at $\alpha = 0.8$. It roughly remains at this value for even larger renewable mixes. As the first eigenvalue decreases, the next eigenvalues increase. At $\alpha = 0.8$ the first six eigenvalues add up to 0.94, explaining then 94% of the mismatch variance.

A threshold like $\sum_{k=1}^{K} \lambda_k \approx 0.95$ for the first eigenvalues is often used as a criterium to define the number K of principal components. Together they explain 95% of the mismatch variance. For $\alpha = 0$ this number is equal to $K = 3$. The corresponding eigenvectors are shown in Figure 4. The first principal component is a kind of monopole. All countries either have a positive mismatch together or

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 September 2017 doi:10.20944/preprints201709.0092.v1

Peer-reviewed version available at Energies 2017, 10, 1934; doi:10.3390/en10121934

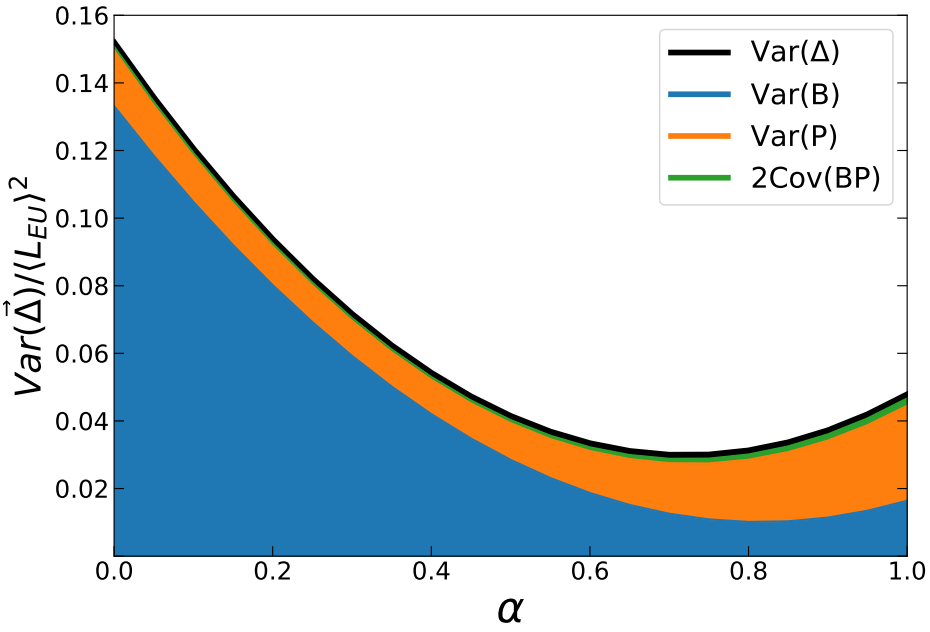

Figure 2. Mismatch variance $\mathrm{Var}(\vec{\Delta})$ (thick black curve) as a function of the renewable mix α for the renewable penetration $\gamma = 1$. The coloured parts illustrate its decomposition (9) into the variances of the balancing (blue) and the injection (orange) as well as their covariance (green).

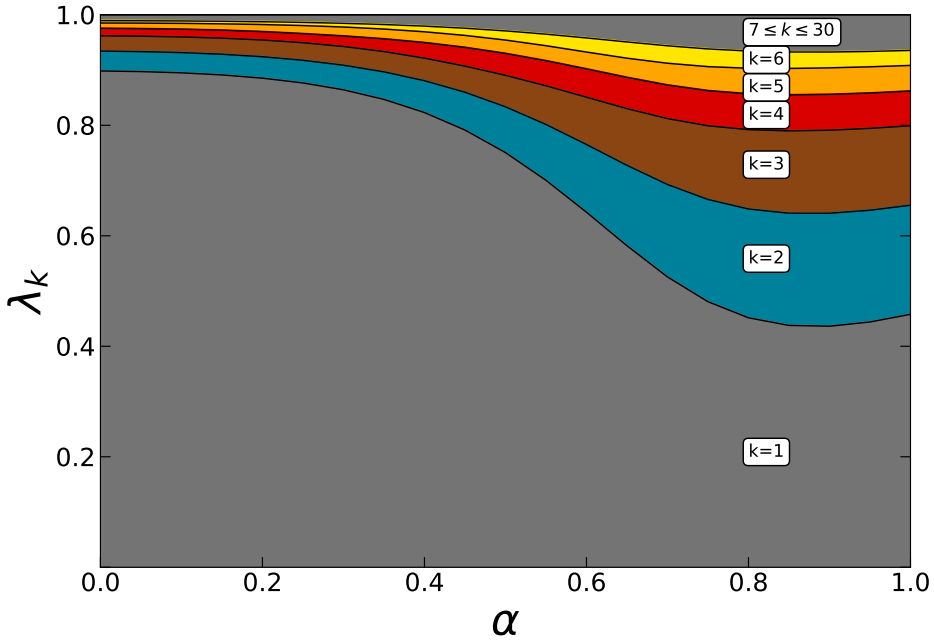

Figure 3. PCA eigenvalues as a function of the renewable mix α.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 September 2017 doi:10.20944/preprints201709.0092.v1

Peer-reviewed version available at Energies 2017, 10, 1934; doi:10.3390/en10121934

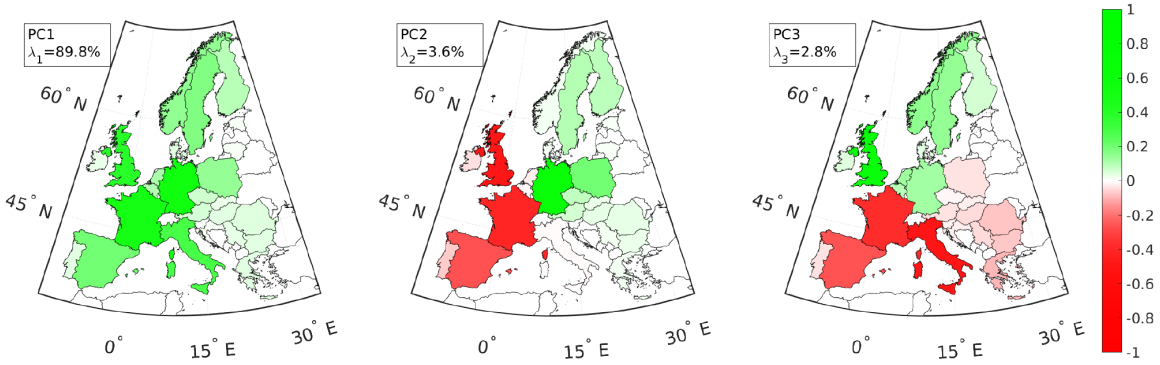

Figure 4. The three principal mismatch components for $\alpha = 0$.

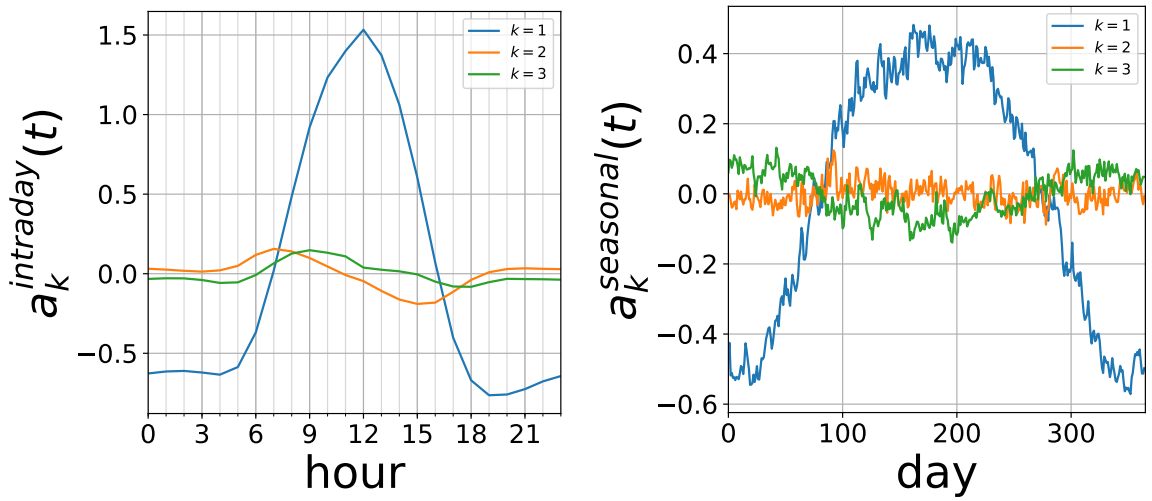

Figure 5. (a, left) Intraday and (b, right) seasonal amplitude profiles of the three principal components for $\alpha = 0$. The intraday profiles $a_k^{\text{intraday}}(t)$ have been obtained by averaging the amplitudes over all days of the underlying eight years while keeping the intraday hour fixed. For the seasonal profiles $a_k^{\text{seasonal}}(t)$ the amplitudes have been first averaged over a day, and then over the eight years while keeping the day of the year fixed.

a negative mismatch. Of course, this is easy to understand. Either the sun is shining for all countries during daytime, or it is dark for all countries during nighttime. The second principal component looks like a dipole in East-West direction. It is a consequence of morning sunrise in the East and of evening sunset in the West. These interpretations are clearly supported by the intraday amplitude profiles $a_k^{\text{intraday}}(t)$, which are shown in Figure 5a.

Also the third principal component has a dipole structure, but with a North-South orientation. This is mostly a consequence of its seasonal amplitude profile $a_3^{\text{seasonal}}(t)$, which is shown in Figure 5b. Figure 5b also reveals that the first principal component has a very pronounced seasonal amplitude profile, whereas no clear seasonal dependence is observed for the amplitude of the second principal component.

Figure 6 depicts the $K = 6$ principal components for the renewable mix $\alpha = 0.8$. The first principal component again looks like a monopole. Roughly it means the wind power generation is above the load for all countries at the same time, or vice versa. Its eigenvalue is smaller than for the $\alpha = 0$ case, and also its structure appears to be slightly different. The other five principal components represent

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 September 2017 doi:10.20944/preprints201709.0092.v1

Peer-reviewed version available at Energies 2017, 10, 1934; doi:10.3390/en10121934

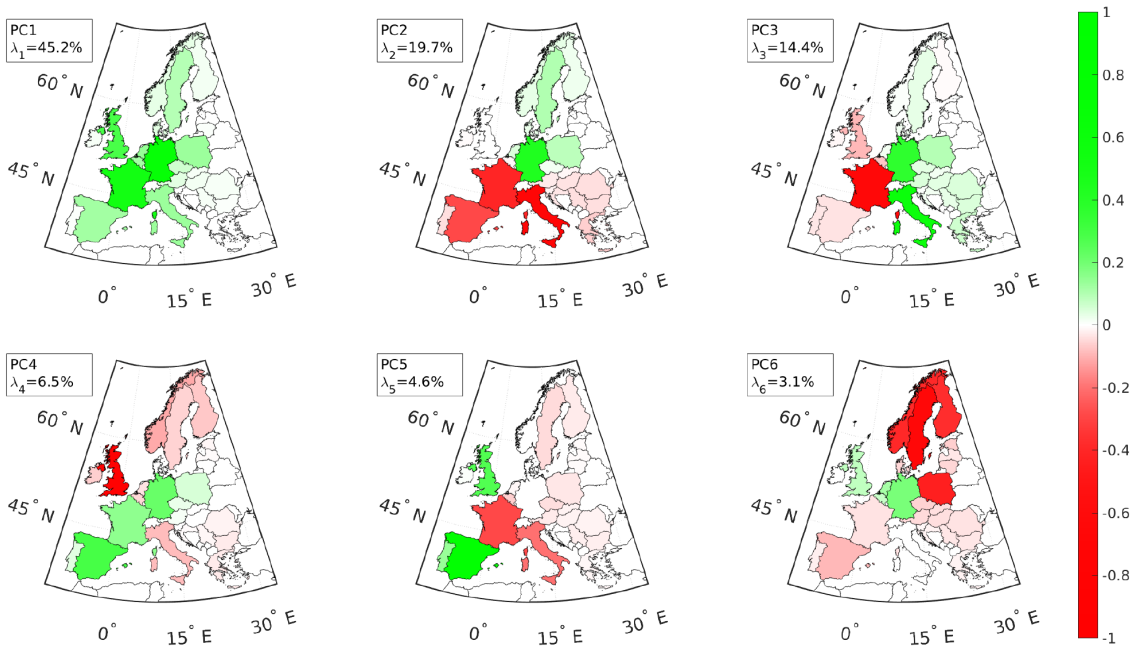

Figure 6. The six principal mismatch components for $\alpha = 0.8$.

different orthogonal spatial mismatch patterns. The second and third one look again like dipoles. The fourth to sixth principle components have a more complicated spatial structure.

Again it is interesting to investigate the time dependence of their amplitudes $a_k(t)$. The intraday and seasonal profiles are shown in Figure 7 for the first three principal components. The intraday amplitude profile of the first principal component has two minima, one during the morning and another one during the evening hours. They are explained by the morning and evening peaks of the load. The maximum at noon traces back to the midday solar power generation, which on average still contributes 20% to the overall renewable power generation due to the setting $\alpha = 0.8$ for the renewable mix. The $k = 2$ amplitude shows a similar intraday profile, although much weaker than for the $k = 1$ amplitude. The seasonal amplitude profiles of the first three principal components all follow the same trend and are correlated to the seasonal wind power generation, which is larger in winter and smaller in summer. The fluctuations on the synoptic time scale of the order of a few days are also clearly visible.

4. Discussion: contribution of principal mismatch patterns to the balancing and transmission infrastructures

According to Eq. (9) the mismatch variance splits into the balancing variance, the injection variance and the covariance between balancing and injection. The three contributions are also shown in Figure 2. For a small renewable mix α the balancing variance dominates, but for large mixes the injection variance becomes bigger. The covariance remains very small for all mixes. It is now interesting to investigate how the different mismatch principal components contribute to the balancing and injection variances.

The PCA eigenvalues can be split into the same three contributions. Making use of the results of Subsection 2.3 and of the response equation (4), the PCA amplitudes can be written as $a_k = c\,\vec{\Delta} \cdot \vec{p}_k =$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 September 2017 doi:10.20944/preprints201709.0092.v1

Peer-reviewed version available at Energies 2017, 10, 1934; doi:10.3390/en10121934

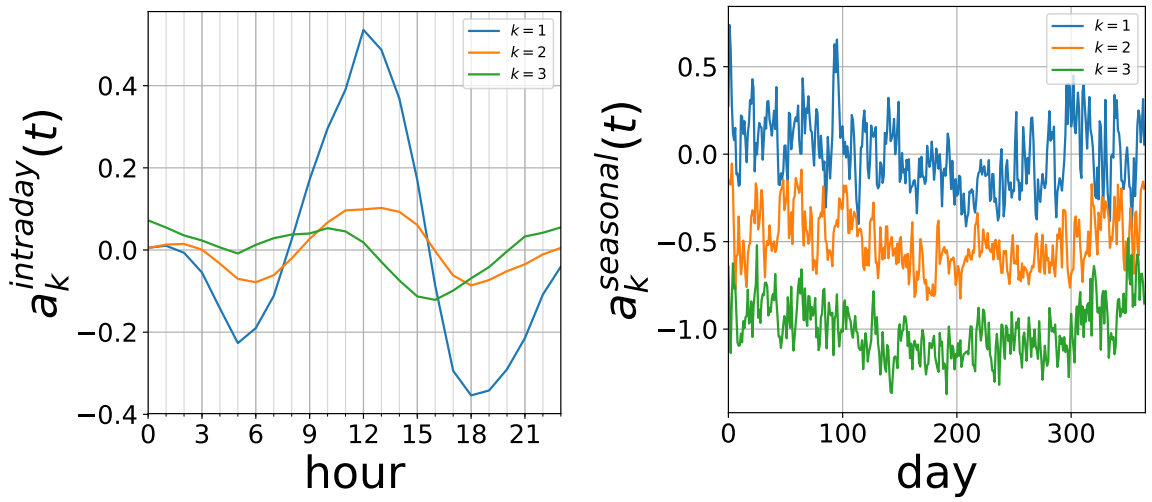

Figure 7. (a, left) Intraday and (b, right) seasonal amplitude profiles of the first three principal components for $\alpha = 0.8$. The seasonal amplitude profiles for $k = 2$ and $k = 3$ have been artificially shifted by -0.5 and -1.0, respectively.

$c\,(\vec{B}+\vec{P})\cdot\vec{p}_k = c\,(B_k + P_k)$. In the last step the abbreviations $B_k = \vec{B}\cdot\vec{p}_k$ and $P_k = \vec{P}\cdot\vec{p}_k$ have been introduced. This allows to write the eigenvalues as

$$\begin{aligned}\lambda_k = \langle a_k^2\rangle &= c^2\left(\langle B_k^2\rangle + \langle P_k^2\rangle + 2\langle B_k P_k\rangle\right)\\ &= \lambda_k^B + \lambda_k^P + \lambda_k^{BP}\,.\end{aligned} \tag{19}$$

This decomposition reveals how much a principal component is contributing to the balancing and injection variances.

Figure 8 illustrates this decomposition. The balancing eigenvalues λ_k^B are plotted from the bottom, and the injection eigenvalues λ_k^P are arranged from the top. The narrow white stripe in the middle represents the sum $\sum_{k=1}^N \lambda_k^{BP}$ and is not specified further into the different k contributions. The first principal component dominates the balancing variance, independent of the renewable mix. This is because of its monopole-like structure. The other components do not contribute much, and only with a small margin for renewable mixes close to one. On the contrary, the first prinicipal component does not contribute much to the injection variance. Here the largest contributions come from the second and third principal component. Again this is intuitive due to their dipole character. For large renewable mixes also the next three principal components contribute to the injection variance to a larger extend.

These results are also documented in Table 1 for the specific choice $\alpha = 0.8$ of the renewable mix. The mismatch has been truncated according to (18). The full ($K = 30$) balancing variance $\mathrm{Var}(\vec{B})$ is fully reproduced with the $K = 6$ principal components. A severe truncation to $K = 1$ is still able to reproduce 88% of the full balancing variance, emphasising again the dominant role of the first principal component for the mismatch-induced balancing. For the injection variance $\mathrm{Var}(\vec{P})$ a different finding is observed. The first principal component alone, reflected by the $K = 1$ truncation, does play only a small role. The second and third principal component contribute the most, but are not able to explain all of the variance. The first $K = 6$ principal components taken together are able to explain 89% of the full ($K = 30$) injection variance. Apparently, higher-order components with $k \geq 7$ still play a small role.

Table 1 lists also the other, more realistic infrastructure measures introduced in Subsection 2.2. The balancing energy E^B and the backup capacity $\mathcal{K}^B$ are related to the balancing variance $\mathrm{Var}(\vec{B})$.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 September 2017 doi:10.20944/preprints201709.0092.v1

Peer-reviewed version available at Energies 2017, 10, 1934; doi:10.3390/en10121934

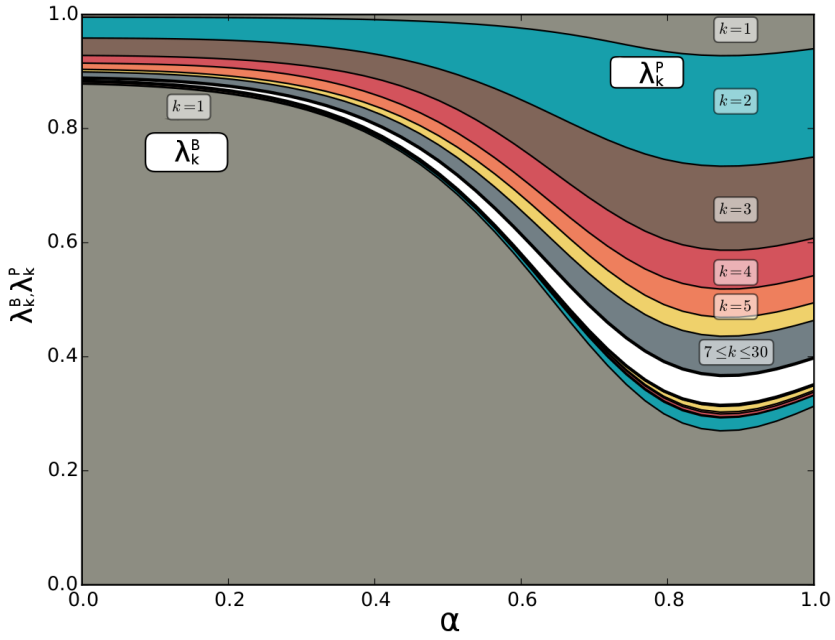

Figure 8. Decomposed eigenvalues λ_k^B (from the bottom) and λ_k^P (from the top) as a function of the renewable mix α. The white band in the middle represents $\sum_{k=1}^{N} \lambda_k^{BP}$.

Table 1. Various infrastructure measures as a function of the truncation parameter K used for the PCA approximation (18).

$\gamma = 1, \alpha = 0.8$	$K = 30$	$K = 6$	$K = 5$	$K = 4$	$K = 3$	$K = 2$	$K = 1$
$\mathrm{Var}(\vec{\Delta})$	0.0301	0.0281	0.0272	0.0258	0.0238	0.0195	0.0136
$\mathrm{Var}(\vec{B})$	0.0101	0.0101	0.0098	0.0098	0.0096	0.0095	0.0089
E^B	0.1402	0.1388	0.1371	0.1368	0.1365	0.1348	0.1318
$\mathcal{K}^B$	0.686	0.682	0.659	0.656	0.650	0.643	0.613
$\mathrm{Var}(\vec{P})$	0.0184	0.0164	0.0154	0.0139	0.0119	0.0077	0.0020
$\mathrm{Var}(\underline{E})$	0.0120	0.0101	0.0091	0.0074	0.0061	0.0040	0.0007
$\mathcal{K}^T$	1.812	1.486	1.372	1.250	1.156	0.904	0.356

Again, the first principal component dominates in both cases, and the $K = 6$ truncation is able to explain 99% of the full $K = 30$ results. The flow variance

$$\mathrm{Var}(\underline{F}) = \sum_{l=1}^{L} \langle F_l^2 \rangle \tag{20}$$

and the transmission capacity $\mathcal{K}^T$ are related to the injection variance $\mathrm{Var}(\vec{P})$. The findings for (20) are very similar to those for the injection variance. The $K = 6$ truncation is able to explain 84% of the full $K = 30$ flow variance. For the transmission capacity (12) the first principal component appears to play a slightly larger role. The $K = 6$ truncation is able to explain 82% of the full $K = 30$ result.

5. Conclusions and Outlook

A Principal Component Analysis has been applied to a simplified highly renewable European electricity system to learn about the most important mismatch patterns between the weather-driven renewable power generation and the load. For a solar dominated system three principal components are enough to explain most of the mismatch variance across the continent. For a wind dominated system six principal components are needed. The spatial structure of the first principal mismatch component resembles that of a monopole, and the respective temporal amplitude dynamics reveals strong fluctuations on the diurnal and seasonal time scales. It determines most of the required backup energy and backup capacity of the overall system. The second and third principal components have a dipole-like spatial structure; their temporal amplitude dynamics is dominated by diurnal, synoptic and seasonal time scales. Together with the fourth to sixth principal components they cause most of the required transmission infrastructure. These results demonstrate that the required backup and transmission infrastructure of a highly renewable European electricity system are to a large degree caused by only a small number of most important mismatch patterns between the weather-driven renewable power generation and the load.

The model used for a highly renewable European electricity system had been quite simple. In particular, its spatial resolution has been scaled to country sizes. As a consequence, the covariance matrix underlying the Principal Component Analysis is dominated by the larger countries like Germany, France, United Kingdom, Spain and Italy. It is interesting to observe that in the most important principal components the smaller countries appear to be grouped into blocks like Scandinavia and East Europe. Nevertheless, the next step to take will be to apply the Principal Component Analysis to larger, fine-grained continental electricity networks, with the additional inclusion of hydro power and storage. In view of the higher dimensionality, the resulting number of principal components is certainly going to increase, but per se it is not clear by how much and how the fine-grained spatial structure of the principal patterns is going to look like. In this respect it will also make sense to investigate related data-reduction approaches like Independent Component Analysis and Dynamic Mode Decomposition [49].

Another future step to take will be the stochastic modelling of the amplitude times series related to the principal components. This can be used for the forecasting of power mismatches, and it will also allow to estimate the additional infrastructure due to the operational forcast uncertainties. An example would be line congestion in transmission systems under uncertainty [50]. Another playground for the application of the Principal Component Analysis could be the impact of climate change on the design of future energy systems, in particular smart energy systems including all cross-sector couplings [51].

Acknowledgments: The authors thank Smail Kozarcanin, Fabian Hofmann, Jonas Hörsch and Tom Brown for helpful discussions. Kun Zhu and Martin Greiner are fully / partially funded by the RE-INVEST project (Renewable Energy Investment Strategies – A two-dimensional interconnectivity approach), which is supported by Innovation Fund Denmark (6154-00022B). The responsibility for the contents lies solely with the authors.

Author Contributions: M.R., T.Z. and K.Z. performed the model simulations and analysis, and produced the figures; M.G. wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

PCA Principal Component Analysis

References

1. Graabak, I..; Korpås, M.. Variability characteristics of European wind and solar power resources – a review. *Energies* **2016**, *9*, 449.
2. Engeland, K..; Borga, M..; Creutin, J.D..; François, B..; Ramos, M.H..; Vidal, J.P.. Space-time variability of climate variables and intermittent renewable electricity production – a review. *Renewable and Sustainable Energy Reviews* **2017**, *79*, 600–17.
3. Holttinen, H.. Wind power variations in the Nordic countries. *Wind Energy* **2005**, *8*, 173–95.
4. Apt, J.. The spectrum of power from wind turbines. *J. Power Sources* **2007**, *169*, 369–74.
5. Milan, P..; Wächter, M..; Peinke, J.. Turbulent character of wind energy. *Physical Review Letters* **2013**, *110*, 138701.
6. Grams, C..; Beerli, R..; Pfenninger, S..; Staffell, I..; Wernli, H.. Balancing Europe's wind-power output through spatial deployment informed by weather regimes. *Nature Climate Change* **2017**, *7*, 557–62.
7. Holttinen, H..; Milligan, M..; Kirby, B..; Acker, T..; Neimane, V..; Molinski, T.. Using standard deviation as a measure of increased operational reserve requirement for wind power. *Wind Energy* **2008**, *32*, 355–78.
8. Holttinen, H..; Meibom, P..; Orths, B..; Lange, M..; O'Malley, M..; Tande, J.O.L..; Estanquerio, A..; Gomez, E..; Söder, L..; Strbac, G..; Smith, J.C..; Van Hulle, F.. Impacts of large amounts of wind power on design and operation of power systems – results of IEA collaboration. *Wind Energy* **2011**, *14*, 179–92.
9. Tarroja, B..; Mueller, F..; Eichman, J.D..; Samuelsen, S.. Metrics for evaluating the impacts of intermittent renewable generation on utility load-balancing. *Energy* **2012**, *42*, 546–62.
10. Lannoye, E..; Flynn, D..; O'Malley, M.. Evaluation of power system flexibility. *IEEE Transactions on Power Systems* **2012**, *27*, 922–31.
11. Huber, M..; Dimkova, D..; Hamacher, T.. Integration of wind and solar power in Europe: assessment of flexibility requirements. *Energy* **2014**, *69*, 236–46.
12. Ulbig, A..; Andersson, G.. Analyzing operational flexibility of electric power systems. *International Journal of Electrical Power & Energy Systems* **2015**, *72*, 155–64.
13. Schlachtberger, D.P..; Becker, S..; Schramm, S..; Greiner, M.. Backup flexibility classes in emerging large-scale renewable electricity systems. *Energy Conversion and Management* **2016**, *125*, 336–46.
14. Kiss, P..; Jánosi, I.M.. Limitations of wind power availability over Europe: a conceptual study. *Nonlinear Processes in Geophysics* **2008**, *15*, 803–13.
15. Baïle, R..; Muzy, J.F.. Spatial intermittency of surface layer wind fluctuations at mesoscale range. *Physical Review Letters* **2010**, *105*, 254501.
16. Widén, J.. Correlations between large-scale solar and wind power in a future scenario for Sweden. *IEEE Transactions on Sustainable Energy* **2011**, *2*, 177–84.
17. Roques, F..; Hiroux, C..; Saguan, M.. Optimal wind power deployment in Europe – a portfolio approach. *Energy Policy* **2010**, *38*, 3245–56.
18. Rombauts, T..; Delarue, E..; D'haeseleer, W.. Optimal-portfolio-theory-based allocation of wind power – taking into account cross-border transmission-capacity constraints. *Renewable Energy* **2011**, *36*, 2374–87.
19. Thomaidas, N.S..; Santos-Alamillos, F.J..; Pozo-Vazquez, D..; Usaola-Garcia, J.. Optimal management of wind and solar energy resources. *Computers & Operations Research* **2016**, *66*, 284–91.
20. Hagspiel, S..; Papamannouil, A..; Schmid, M..; Andersson, G.. Copula-based modeling of stochastic wind power in Europe and implications for the Swiss power grid. *Applied Energy* **2012**, *96*, 33–44.
21. Czisch, G.. Szenarien zur zukünftigen Stromversorgung. PhD thesis, Universität Kassel, 2005.
22. Heide, D..; Von Bremen, L..; Greiner, M..; Hoffmann, C..; Speckmann, M..; Bofinger, S.. Seasonal optimal mix of wind and solar power in a future, highly renewable Europe. *Renewable Energy* **2010**, *35*, 2483–89.
23. Heide, D..; Greiner, M..; Von Bremen, L..; Hoffmann, C.. Reduced storage and balancing needs in a fully renewable European power system with excess wind and solar power generation. *Renewable Energy* **2011**, *36*, 2515–23.

24. Schaber, K..; Steinke, F..; Hamacher, T.. Transmission grid extensions for the integration of variable renewable energies in Europe: Who benefits where? *Energy Policy* **2012**, *43*, 123 – 35.
25. Schaber, K..; Steinke, F..; Mühlich, P..; Hamacher, T.. Parametric study of variable renewable energy integration in Europe: Advantages and costs of transmission grid extensions. *Energy Policy* **2012**, *42*, 498–508.
26. Rodriguez, R.A..; Becker, S..; Andresen, G..; Heide, D..; Greiner, M.. Transmission needs across a fully renewable European power system. *Renewable Energy* **2014**, *63*, 467–76.
27. Becker, S..; Rodríguez, R.A..; Andresen, G.B..; Schramm, S..; Greiner, M.. Transmission grid extensions during the build-up of a fully renewable pan-European electricity supply. *Energy* **2014**, *64*, 404–18.
28. Becker, S..; Frew, B.A..; Andresen, G.B..; Zeyer, T..; Schramm, S..; Greiner, M..; Jacobson, M.Z.. Features of a fully renewable US electricity system: Optimized mixes of wind and solar PV and transmission grid extensions. *Energy* **2014**, *64*, 404–18.
29. Gils, H.C..; Scholz, Y..; Pregger, T..; De Tena, D.L..; Heide, D.. Integrated modelling of variable renewable energy-based power supply in Europe. *Energy* **2017**, *123*, 173–88.
30. Andresen, G.B..; Sœndergaard, A.A..; Greiner, M.. Validation of Danish wind time series from a new global renewable energy atlas for energy system analysis. *Energy* **2015**, *93*, 1074–88.
31. Staffell, I..; Pfenninger, S.. Using bias-corrected reanalysis to simulate current and future wind power output. *Energy* **2016**, *114*, 1224–39.
32. Pfenninger, S..; Staffell, I.. Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. *Energy* **2016**, *114*, 1251–65.
33. Rasmussen, M.G..; Andresen, G.B..; Greiner, M.. Storage and balancing synergies in a fully or highly renewable pan-European power system. *Energy Policy* **2012**, *51*, 642–51.
34. Steinke, F..; Wolfrum, P..; Hoffmann, C.. Grid vs. storage in a 100% renewable Europe. *Renewable Energy* **2013**, *50*, 826–32.
35. Jensen, T.V..; Greiner, M.. Emergence of a phase transition for the required amount of storage in highly renewable electricity systems. *European Physics Journal Special Topics* **2014**, *223*, 2475–81.
36. Bussar, C..; Moos, M..; Alvarez, R..; Wolf, P..; Thien, T..; Chen, H..; Cai, Z..; Leuthold, M..; Sauer, D.U..; Moser, A.. Optimal allocation and capacity of energy storage systems in a future European power system with 100% renewable energy generation. *Energy Procedia* **2014**, *46*, 40–47.
37. Bussar, C..; Stöcker, P..; Cai, Z..; Moraes, L..; Alvarez, R..; Chen, H..; Breuer, C..; Moser, A..; Leuthold, M..; Sauer, D.U.. Large-scale integration of renewable energies and impact on storage demand in a European renewable power system of 2050. *Energy Procedia* **2015**, *73*, 145–53.
38. Weitemeyer, S..; Kleinhans, D..; Vogt, T..; Agert, C.. Integration of renewable energy sources in future power systems: The role of storage. *Renewable Energy* **2015**, *75*, 14–20.
39. Budischak, C..; Sewell, D..; Thomson, H..; Mach, L..; Veron, D.E..; Kempton, W.. Cost-minimized combinations of wind power, solar power and electrochemical storage, powering the grid up to 99.9% of the time. *Journal of Power Sources* **2013**, *225*, 60–74.
40. Rodríguez, R.A..; Becker, S..; Greiner, M.. Cost-optimal design of a simplified, highly renewable pan-European electricity system. *Energy* **2015**, *83*, 658 – 68.
41. Eriksen, E.H..; Schwenk-Nebbe, L.J..; Tranberg, B..; Brown, T..; Greiner, M.. Optimal heterogeneity in a simplified highly renewable European electricity system. *Energy* **2017**, *133*, 913–28.
42. Schlachtberger, D..; Brown, T..; Schramm, S..; Greiner, M.. The benefits of cooperation in a highly renewable European electricity network. *Energy* **2017**, *134*, 469–81.
43. Haken, H.. *Principles of Brain Functioning*; Springer Verlag Berlin Heidelberg, 1996.
44. Von Bremen, L..; Saleck, N..; Heinemann, D.. Enhanced regional forecasting considering single wind farm distribution for upscaling. *Journal of Physics Conference Series* **2007**, *75*, 012040.
45. Burke, D.J..; O'Malley, M.J.. A study of Principal Component Analysis applied to spatially distributed wind power. *IEEE Transactions on Power Systems* **2011**, *26*, 2084–92.
46. Rodriguez, R.A..; Dahl, M..; Becker, S..; Greiner, M.. Localized vs. synchronized exports across a highly renewable pan-European transmission network. *Energy, Sustainability and Society* **2015**, *5*, 21.
47. Wood, A.J..; Wollenberg, B.F..; Sheblé, G.B.. *Power Generation, Operation and Control*; John Wiley & Sons, New York, 2014.

48. Schäfer, M..; Tranberg, B..; Hempel, S..; Schramm, S..; Greiner, M.. Decompositions of injection patterns for nodal flow allocation in renewable electricity networks. *European Physics Journal B* **2017**, *90*, 144.
49. Schmid, P.. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics* **2010**, *656*, 5–28.
50. Deladreue, S..; Brouaye, F..; Bastard, P..; Péligry, L.. Using two multivariate methods for line congestion study in transmission systems under uncertainty. *IEEE Transactions on Power Systems* **2003**, *18*, 353–58.
51. Connolly, D..; Lund, H..; Mathiesen, B.. Smart Energy Europe: the technical and economic impact of one potential 100% renewable energy scenario for the European Union. *Renewable and Sustainable Energy Reviews* **2016**, *60*, 1634–53.