On evaluation of scientific writing

The sensor assesses for each point the extent to which the candidate has achieved these goals:

1. Professional anchoring
2. Theoretical insight
3. Target description
4. Skill Level
5. The work
6. Analysis and discussion
7. Critical reflection
8. Own contribution / goal achievement
9. Structure
10. Language
11. Form

Language

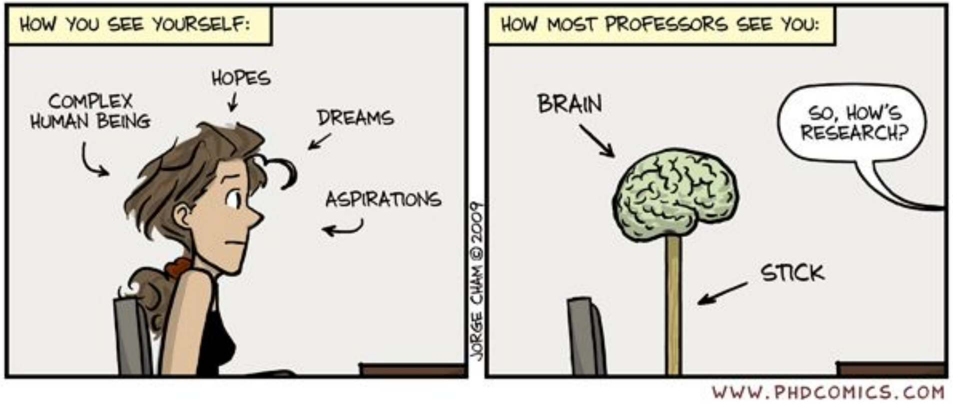
From the guide: “Can the candidate present problem and results with the necessary professional precision? Is it highly readable with high quality in the language used?”

(Usually not a problem mostly because fellow students help with cleaning up the language)

Skills

Do candidates apply relevant methods and use them in their own work in an appropriate and integrated manner?

(Usually not a problem because the supervisor has a pretty good idea of appropriate methods)



Effort

Does the work show creativity and / or contribute to innovation / innovation? Does the work give an impression of being particularly extensive? How is the quality and importance of new knowledge / new results generated in the work assessed?

(usually the effort is substantial – don´t put too much weight on the innovative and creative parts)

Professional anchoring

Is the theoretical and academic foundation well described so that the work is included in the field of international research?

(quite often a problem: too little time spent on reading to get an overview)

Analysis and discussion

Is analysis, interpretation / synthesis and discussion professionally founded and justified and clearly linked to the issue? Is the discussion at a high academic level? Can the candidate apply his / her knowledge and skills in new areas and place the results in a larger context?

(quite often a problem: requires an overview)

Critical reflection

Does the candidate give a reasonable assessment of the importance of the results? Is the candidate critical of different sources of information? Are uncertainties, such as method errors, measurement errors, and others considered and discussed? Are relevant academic, occupational and research ethical issues analyzed?

(quite often a problem: again how can one assess the importance and limitations of the work without an overview of the field)

Preliminary conclusion

A lot of the criteria concern “your overview of the field”:

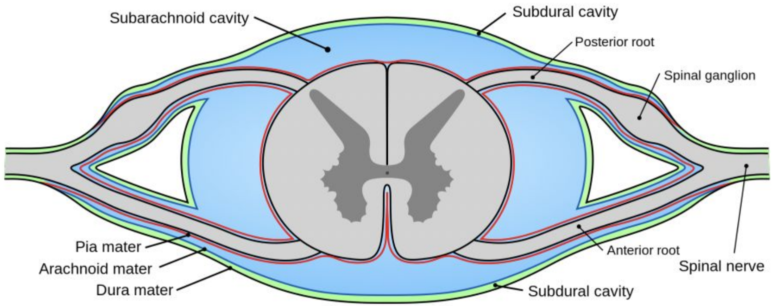
- anchoring
- analysis and discussion
- critical reflection

It is not sufficient to just read the few articles or papers before starting your work

Introduction

The Cerebrospinal Fluid (CSF) is a clear, colorless fluid occupying the space around the brain and spinal cord, namely the Spinal Subarachnoid Space (SSS). The CSF is produced in the cerebral ventricles in the brain, and pulsates through the SSS with nearly zero net velocity. Since the CSF surrounds the Central Nervous System (CNS), drugs are often distributed via the CSF. Due to large patient specific variations of the geometry in the SSS and the complexity of the flow in the CSF, it is a difficult task to predict and control the concentration of the drugs [15]. A poor measurement access in the SSS also restricts our knowledge of the spreading. The uncertainty of the uptake and spreading of the drugs may lead to an increased risk of complications. This is the motivation for developing numerical simulations where the input parameters are the patient-specific variables, and the output is the concentration of the drug. Such simulation could help to optimize the medical procedure and to improve the knowledge of the physiological dynamics.

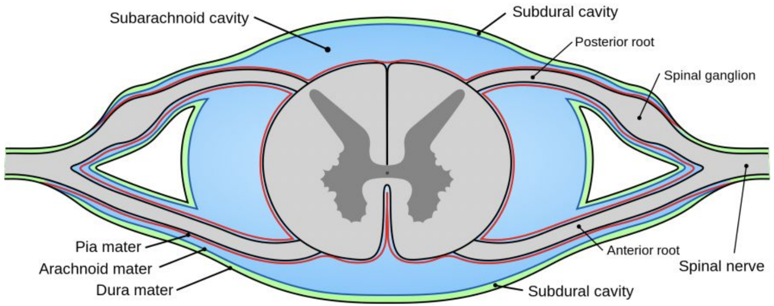
Throughout the years, many studies have been done on the CSF dynamics both *in vivo*, *in vitro* and *in silco*. Studies using different MRI techniques to measure the CSF flow *in vivo* includes [4, 10, 41, 28]. Many of these studies are related to investigating CSF flow in patients with Chiari malformation. Experiments have been perfomed *in vitro* in e.g. [23, 22, 24] to map the CSF flow related to the the syringomyelia disorder. In recent years, researches have begun to apply Computational Fluid Dynamics (CFD) models to improve the knowledge of the CSF flow. Some studies [18, 21, 31] used simplified/idealized models of the SSS to simulate the CSF flow. The first study to include microstructure in the SSS was Stockman [32], where idealized models of NRDL and trebeculae were used with Lattice Boltzman simulations. Heidari Pahlavian et al. [11] and Tangen et al. [37] followed



Introduction

Big picture

The Cerebrospinal Fluid (CSF) is a clear, colorless fluid occupying the space around the brain and spinal cord, namely the Spinal Subarachnoid Space (SSS). The CSF is produced in the cerebral ventricles in the brain, and pulsates through the SSS with nearly zero net velocity. Since the CSF surrounds the Central Nervous System (CNS), drugs are often distributed via the CSF. Due to large patient specific variations of the geometry in the SSS and the complexity of the flow in the CSF, it is a difficult task to predict and control the concentration of the drugs [15]. A poor measurement access in the SSS also restricts our knowledge of the spreading. The uncertainty of the uptake and spreading of the drugs may lead to an increased risk of complications. This is the motivation for developing numerical simulations where the input parameters are the patient-specific variables, and the output is the concentration of the drug. Such simulation could help to optimize the medical procedure and to improve the knowledge of the physiological dynamics.



Throughout the years, many studies have been done on the CSF dynamics both *in vivo*, *in vitro* and *in silco*. Studies using different MRI techniques to measure the CSF flow *in vivo* includes [4, 10, 41, 28]. Many of these studies are related to investigating CSF flow in patients with Chiari malformation. Experiments have been perfomed *in vitro* in e.g. [23, 22, 24] to map the CSF flow related to the the syringomyelia disorder. In recent years, researches have begun to apply Computational Fluid Dynamics (CFD) models to improve the knowledge of the CSF flow. Some studies [18, 21, 31] used simplified/idealized models of the SSS to simulate the CSF flow. The first study to include microstructure in the SSS was Stockman [32], where idealized models of NRDL and trebeculae were used with Lattice Boltzman simulations. Heidari Pahlavian et al. [11] and Tangen et al. [37] followed

What others have done

Information Sciences 590 (2022) 74–89

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

A multivariate dependence analysis for electricity prices, demand and renewable energy sources

Fabrizio Durante [a,*], Angelica Gianfreda [b,c], Francesco Ravazzolo [d,e], Luca Rossini [f]

[a] Dipartimento di Scienze dell'Economia, Università del Salento, Lecce, Italy
[b] Department of Economics "Marco Biagi", University of Modena and Reggio Emilia, Modena, Italy
[c] Energy Markets Group, London Business School, UK
[d] Department of Data Science and Analytics, BI Norwegian Business School, Norway
[e] Faculty of Economics and Management, Free University of Bozen-Bolzano, RCEA, Italy
[f] Department of Economics, Management and Quantitative Methods, University of Milan, Italy

1. Introduction

In recent years, the electricity generation from renewable energy sources (RES) has increased in importance in the economies of all countries, especially in Europe, due to stringent regulations to reduce carbon emissions and to provide incentives for investments in clean technologies. However, the interrelationships between RES and demand, and their combined effect on electricity prices have been under-investigated and there are still few works focusing on this multivariate dependence. These relations are particularly important since RES can reduce the demand for electricity if weather conditions allow. Indeed, it has been largely proved that wind generation reduces the mean (and the skewness) of the distribution of electricity price while increasing the price variability. In contrast, there is no clear understanding of the effect of solar power generation, especially regarding its interactions with demand, and eventually with wind power generation. Therefore, this paper aims at exploring these interdependencies in details.

To this aim, a new database is compiled using hourly electricity prices determined on the day-ahead German market together with predictions for both RES and demand. This allows to consider the dependence between these variables and the effects of their different combinations across all 24 h and across a sample of years, going from 2011 (a year in which

1) General introduction
2) What has been done on the subject?
3) What is not know about the subject?
4) The problems addressed in the current report/paper

The actual introduction in the present paper is quite long

RES were at their early introduction) to 2019. Note that Germany is the largest European electricity market for traded volume and production (see [15]). Moreover, it is a leading country for the total wind power capacity per inhabitant (jointly with Denmark) and solar PV capacity per inhabitant (recently flanked by Italy and Spain).[1] Therefore, studying the German market allows us to understand the dependence structure among prices, demand, and RES, which could provide useful guidance for policymakers. In particular, the uncovered multivariate dependence structure could display important effects due to the increasing RES penetration and could provide support for further investments to reduce carbon emissions.

Here, the dependence among prices, demand, and RES is investigated by using a copula approach. This method is appropriate since it allows for a careful description of the multivariate stochastic behavior and for an accurate analysis of different types of association and tail dependence. This is particularly important since, for example, situations in which high wind generation is coupled with high demand levels, together with high solar production, may represent co-movements in extreme behavior that are not easily detected with other methodologies. In particular, copulas allow to proceed in two steps: first, individual variables are modelled according to their features; and, then, the dependencies between price, demand, wind, and solar generation are described with a greater flexibility.

Several papers have applied copula models for modelling energy markets. [21] adopt copulas to evaluate investment decisions regarding the placement of wind turbines with respect to wind speed in order to reduce output fluctuations and stabilize the supply. [7] use copulas to model and investigate the complementarity between hydro and wind, aiming at reducing the risk of shortages in water inflows. Multivariate copulas are instead considered to inspect the integration of wind energy in the European grid; see [22]. [48] implement a multivariate non-normal copula model for studying the behavior of wind speed, solar radiation, and load profiles of a network.

Moreover, copulas have been used for the relationships between electricity prices observed over different regions, or to depict the relationships between prices and fundamental variables. For example, robust partial correlations are estimated between changes in electricity prices in the connected zones of New York state in [10]. In addition, [24] examine the dependence structure of electricity spot prices across Australian regional markets. Several regime-switching AR–GARCH copulas are proposed in [38] to study the pairwise behavior of electricity prices over interconnected European markets (Germany, France, Netherlands, Belgium, and Western Denmark). In particular, the skewed t distribution is considered because it describes the marginal dynamics better than the normal distribution and can also capture the pair-wise tail dependence.

Regarding the study of the dependence between electricity prices and/or renewable energy sources, the literature has focused largely on bivariate models, mainly by considering prices and wind generation. For instance, the dependence between wind power production and electricity prices is examined in [29,39,43,47]. [14] develop a stochastic simulation model able to capture the full spatial dependence structure of wind power by using copula models incorporating also demand and supply information.

Regarding solar power, [36] show that it decreases price volatility and more recently, [18] show that both wind and solar power reduce mean electricity prices, but increase their volatility. More importantly, they provide new insights regarding the negative effect of wind on the skewness of price distributions, hence suggesting to control for the behavior of the tails.

Therefore, this paper extends the recent literature on the multivariate dependence of electricity prices by providing first new methodological tools for the joint tail behavior and then new empirical results based on a novel dataset.

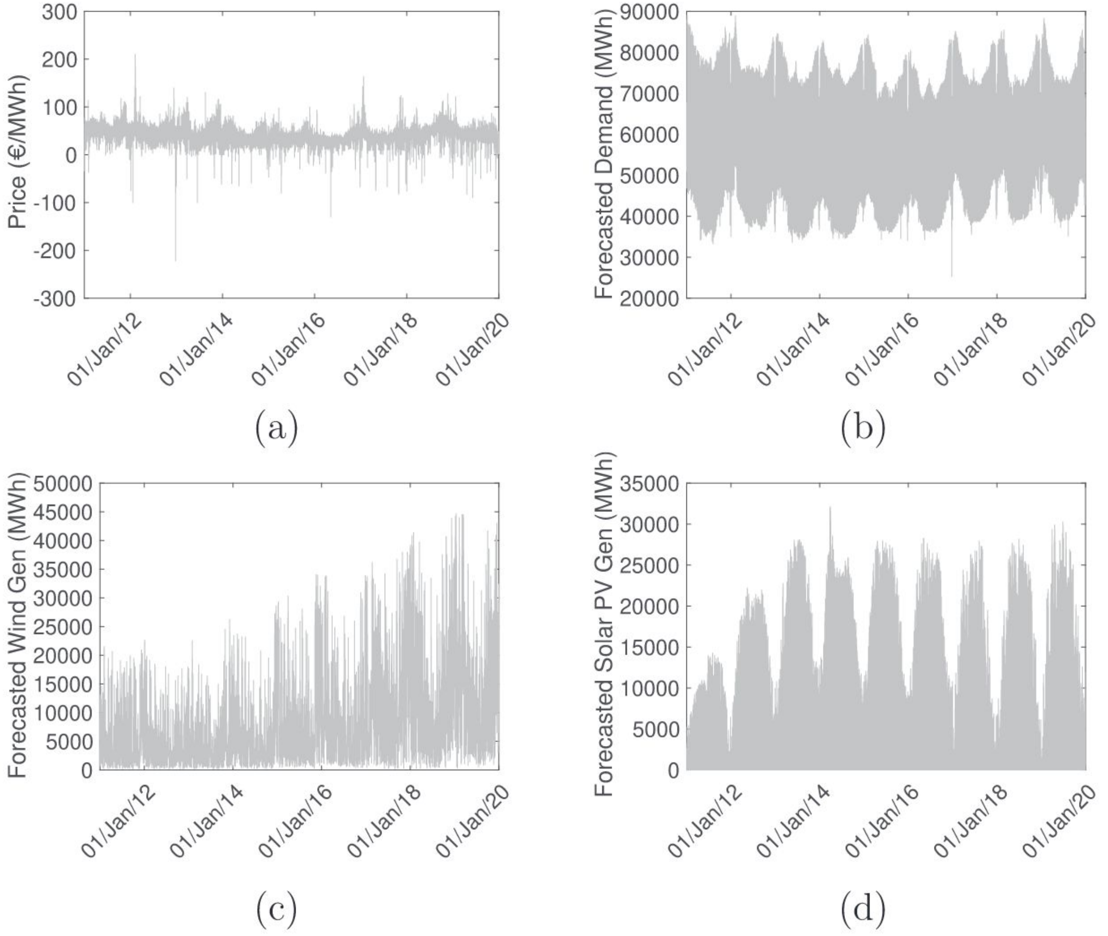
NB: Your introduction has to contain some form of references. You could include academic papers, data science reports, news articles and so on

2. Data description

This empirical study relies on a new hourly dataset consisting of German electricity prices, forecasted demand, forecasted wind, and forecasted solar PV generation from January 1, 2011, to December 31, 2019. Electricity prices are quoted in €/MWh on a daily basis. They have been pre-processed for time-clock changes, that is the 25th hour in October has been excluded, whereas the missing 24th hour in March has been interpolated. Hence, there are no missing observations.

The hourly auction prices in Germany are determined on the day-ahead market before noon, and then, in practice, they are forward prices for delivery during the predetermined hours on the following day. These prices have been collected directly from the German power market, *European Energy Exchange* (EEX). In addition, by considering the day-ahead determination of prices, the forecasted values for demand, wind, and solar PV generation have been used, as provided by Thomson Reuters with an hourly frequency. Specifically, the forecasts used in this analysis are those obtained by the European Centre for Medium-Range Weather Forecast (ECMWF), which result from the running of the operational model at midnight (technically, the model is said to *run at hour 00*). This represents the latest information available to market operators before they submit their bids/offers, because this model updates from 05.40 a.m. to 06.55 a.m.

Short data description. Here, I am missing a clear reference (e.g. website, repository) for the data.



Part of the data description. You can plot raw data here. You may plot the raw data on a map for instance. You should not perform much data post processing before showing these figures

Fig. 1. Hourly Time Series for Electricity Day-ahead Prices (panel a), Forecasted Demand (panel b), Forecasted Wind Generation (panel c) and Forecasted Solar PV Generation (panel d) observed in Germany from 01/01/2011 to 31/12/2019.

3. Methodology

The main purpose of the paper is to develop a joint stochastic model that characterizes the marginal behavior of electricity prices, demand, and renewable energy sources by capturing the related dependence structure. To this end, we exploit the advantages of the copula methodology, which has been used for economic and financial applications in a number of works (e.g., [4,6,31], references therein). Specifically, an n-dimensional copula is a distribution function supported on the unit cube $[0,1]^n$ with a uniform marginal distribution. As well-known, an n-dimensional joint distribution function can be decomposed into its n univariate marginal distributions and an n-dimensional copula, which is unique when the marginal distributions are continuous. For more details, see also [13,34].

Specifically, in view of Sklar's theorem, given an n-dimensional distribution function F with marginals F_j, for $j = 1, \ldots, n$, a copula $C : [0,1]^n \to [0,1]$ exists that satisfies

$$F(\mathbf{y}) = C(F_1(y_1), \ldots, F_n(y_n)) \tag{1}$$

for every $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$. If F is continuous, then the copula is uniquely determined by

$$C(\mathbf{u}) = F\left(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)\right), \quad \mathbf{u} \in [0,1]^n,$$

In the methods section you describe what models you plan to use and why. If appropriate you may describe methodology/models in detail. You can also describe post processing steps (and why you did it) performed on your data material

Results are often quite descriptive (boring)

Then, the induced pairwise (Spearman's) correlation is computed for each hour and presented in Fig. 5. Note that the link between the forecasted solar power generation and the other variables is included only from hour 8 to hour 16. In particular, Fig. 5 shows a positive dependence between the electricity prices and the forecasted demand during the entire 24 h. The correlation falls during the early morning (i.e. from 5 to 6 approximately at 0.2), and in the late evening after 20; hence, confirming the known fact that prices follow the intra-day dynamics of demand, being higher during peak hours and lower in off-peak hours. When the relation between electricity prices and forecasted wind generation is instead considered, a negative correlation is detected, recalling the reverse dynamics of the intra-daily wind profile. Indeed, the negative correlation is larger when wind generation is high (during early or late hours) and it diminishes, keeping its sign, when wind generation decreases (during peak hours, as shown on the left side of Fig. 2). As expected, the correlation between forecasted demand and wind fluctuates around zero and indeed this is not of concern for this analysis since both variables are influenced by weather conditions.

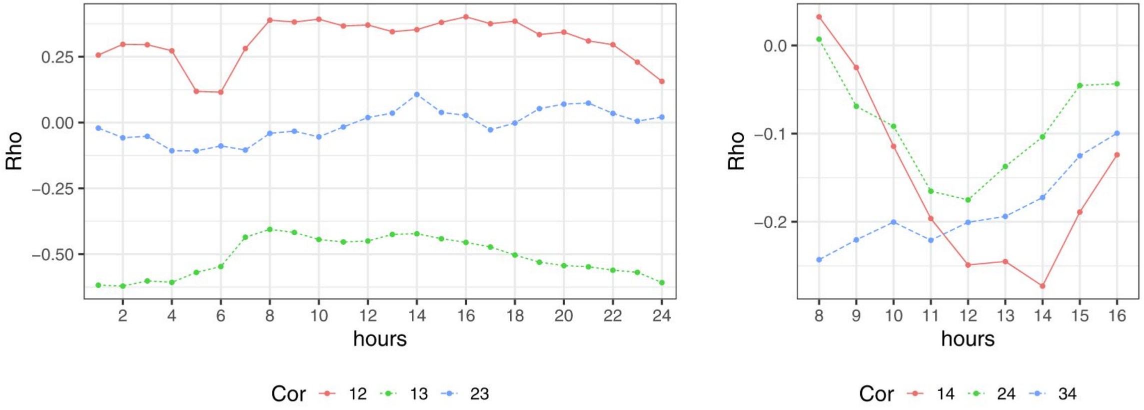


Fig. 5. Pairwise Spearman's correlations induced by the R-Vine Copula model specification over the 24 hours between Electricity Prices (1), Forecasted Demand (2) and Forecasted Wind (3) on the left; and, among Forecasted Solar PV (4) and the other variables on the right.

Results are often quite descriptive (boring)

Remember to reference all figures!

Then, the induced pairwise (Spearman's) correlation is computed for each hour and presented in Fig. 5. Note that the link between the forecasted solar power generation and the other variables is included only from hour 8 to hour 16. In particular, Fig. 5 shows a positive dependence between the electricity prices and the forecasted demand during the entire 24 h. The correlation falls during the early morning (i.e. from 5 to 6 approximately at 0.2), and in the late evening after 20; hence, confirming the known fact that prices follow the intra-day dynamics of demand, being higher during peak hours and lower in off-peak hours. When the relation between electricity prices and forecasted wind generation is instead considered, a negative correlation is detected, recalling the reverse dynamics of the intra-daily wind profile. Indeed, the negative correlation is larger when wind generation is high (during early or late hours) and it diminishes, keeping its sign, when wind generation decreases (during peak hours, as shown on the left side of Fig. 2). As expected, the correlation between forecasted demand and wind fluctuates around zero and indeed this is not of concern for this analysis since both variables are influenced by weather conditions.

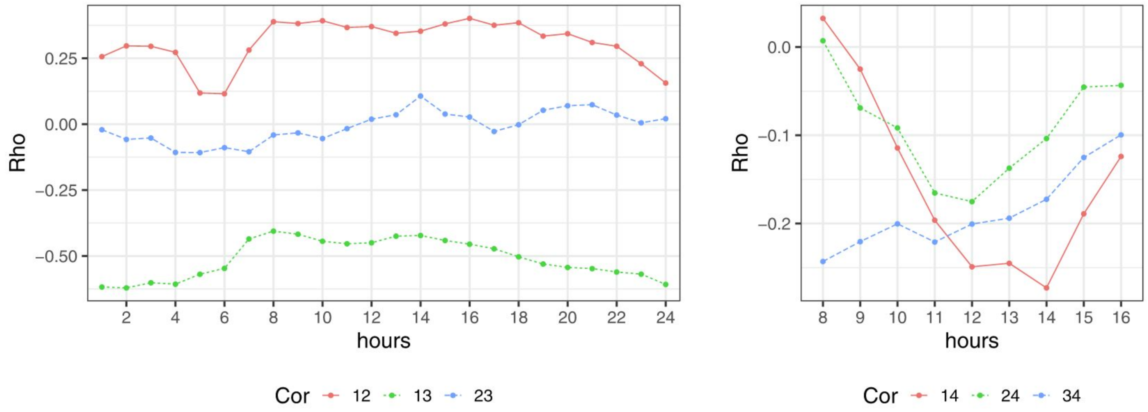


Fig. 5. Pairwise Spearman's correlations induced by the R-Vine Copula model specification over the 24 hours between Electricity Prices (1), Forecasted Demand (2) and Forecasted Wind (3) on the left; and, among Forecasted Solar PV (4) and the other variables on the right.

But can include some discussion/reflection as well…

Overall, these results confirm the well-known *merit order effect*, according to which RES (wind and solar) decrease the electricity prices because they enter the supply curve before the other generation sources and, consequently, they shift the supply curve towards the right, thus decreasing the equilibrium price. However, the results presented in this specific analysis do refer to correlations when considering a multivariate dependence model, that is when all possible interactions between involved variables are considered. More explicitly, correlations are studied when prices interact with individual RES and when RES interact with demand as well. The same results may not occur, for instance, when simple pair-wise correlations or regression models are considered, in which the marginal effects of RES are hypothesized *ceteris paribus*.

Apart from the global correlation, measures of tail dependence have been considered and computed, as described before, between the different variables during the whole 24 h. Fig. 6 shows the model-based pairwise upper and lower tail dependence coefficients (UTDC and LTDC, in short) to capture the extra effect of one variable on the high/low values of the other variable in a pairwise tail dependence, resulting from the multivariate structure.

It can be easily observed that independently from the tails, the coefficient of tail correlation between prices and demand is always positive and varying over the day with dynamics recalling the intra-daily profiles: lower correlations early in the morning and in the evening, higher ones during the middle of the day. This comes with no surprise apart the magnitudes expected to be higher over the right tail when demand pushes power plants under pressures, hence resulting in higher equilibrium prices. However, here the multivariate dependence detects also the interaction between demand and wind, which is indeed higher on the left tail during central hours, and thus resulting in a higher influence on prices. Instead, the most striking result is the asymptotic tail independence between prices and wind on both tails and across all hours, since previous studies have shown how wind instead does influence the left tail of prices at finite, i.e. non-extreme, quantile levels.

5. Conclusions

Using a new compiled dataset, this paper investigates the multivariate dependence between hourly electricity prices, demand, and two different sources of renewable energy (wind and solar PV) in one of the largest producing countries of renewable energy in Europe, i.e., Germany. However, considering multivariate dependence structures is important in all countries for driving policy decisions, since increasing RES generation immediately affects both prices and demand. Therefore, identifying and adopting the appropriate methodology are two important tasks not only for the market studied in this analysis but also for all countries wishing to increase their green generation and reduce carbon emissions.

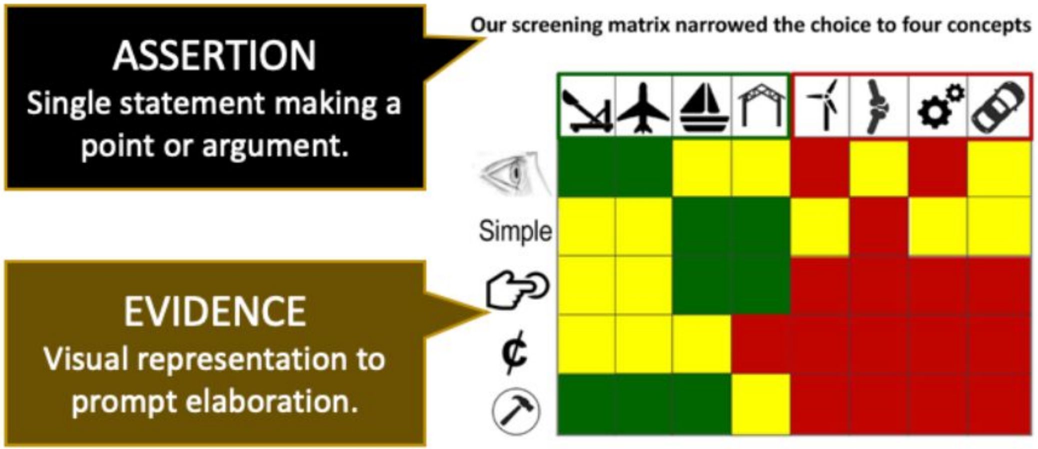
By considering forecasted wind, solar PV generation, demand, and electricity prices, this work studies their joint dependence with a flexible copula approach. Moreover, the introduced multivariate tail dependence coefficients (depending on more than one variable) provide additional insights in the understanding of these relationships in the tail of their joint distribution. Indeed, applying suitable copula-based models for time series, a strong dependence is depicted and mapped between electricity prices, demand and RES during the day with important intra-daily and seasonal patterns.

Apart from the methodological contribution related to the study of tail behavior in a multivariate setting, from an applied point of view, this paper contributes to the literature by filling the gap regarding the interrelationships between RES and demand and their combined effect on the electricity prices, given that there was no clear understanding of the effect of solar, especially its interactions with demand, and, eventually, with wind during central hours; however, here, this issue is addressed, and answers are provided.

I think the conclusion should have a greater focus on the implications and importance of the actual data analytics results than what is given here. Be clear about what you found versus what was known from before and how you link the two to generate more knowledge about the subject

On the use of subheadings

- Subheadings can be very useful in methods, results and discussion.
- Often the assertion-evidence method can be quite useful.



On the use of subheadings

- Subheadings can be very useful in methods, results and discussion.
- Often the assertion-evidence method can be quite useful.

Ex:

Some machine learning algorithms can predict weather with high accuracy. Of the four different machine learning models described in the Methods section, weather conditions were adequately predicted (>95% accuracy) in two cases (Figure 1). Models X and Y revealed high accuracy (96 and 98%) while models Z and W performed poorly on test data (64 and 67% accuracy). All models perform relatively well on training data (93-99% accuracy) suggesting overfitting in models Z and W.

Recap

- Show that you have an overview of the field. This is done in introduction and discussion/conclusion.
- You are allowed to use sub-headings throughout the report. Use it where appropriate to tell the most compelling story.