



HAL
open science

Impact of climate variability on optimal wind-solar energy mixes - The Italian Case

Alexis Tantet, Marc Stéfanon, Philippe Drobinski, Jordi Badosa, Silvia Concettini, Anna Creti, Claudia d'Ambrosio, Dimitri Thomopoulos, Peter Tankov

► **To cite this version:**

Alexis Tantet, Marc Stéfanon, Philippe Drobinski, Jordi Badosa, Silvia Concettini, et al.. Impact of climate variability on optimal wind-solar energy mixes - The Italian Case. 2019. hal-01962044v2

HAL Id: hal-01962044

<https://hal.science/hal-01962044v2>

Preprint submitted on 15 Feb 2019 (v2), last revised 14 Nov 2019 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Impact of climate variability on optimal wind-solar energy mixes — The Italian Case

Tantet A.^{a,*}, Stéfanon M.^a, Drobinski P.^a, Badosa J.^a, Concettini S.^{b,c},
Creti A.^{c,d}, D’Ambrosio C.^e, Thomopoulos D.^e, Tankov P.^f

^a*LMD/IPSL, École polytechnique, Sorbonne Université, ENS, PSL University, CNRS, Palaiseau, France*

^b*IRJI, Université de Tours, Tours, France*

^c*Département d’Économie, École polytechnique, Palaiseau, France*

^d*Université Paris Dauphine, PSL, Leda-CGEMP, Paris, France*

^e*LIX, École polytechnique, CNRS, Palaiseau, France*

^f*CREST, ENSAE, École polytechnique, Palaiseau, France*

Abstract

We develop a renewable energy mix optimization program. The main novelty is to take into account and evaluate the impact of climate variability, from multi-decadal to hourly time scales, on the mix. This impact is resolved by plugging regional climate simulations over the 1989–2012 period into renewable energy production and demand models at regional scale. The optimal mix is then inferred from a mean-variance analysis with as objectives both to maximize the mean of the total renewable penetration and to minimize the variance, or risk, of the latter. We consider two cases: in the first one, the analysis takes cross-region correlations in the production and the demand into account. In the second one, synergies between regions are ignored. As a first case study, we apply the numerical model to Italy, a country with a relatively high share of renewables. We focus on the mix maximizing the ratio of the mean renewable penetration over the risk for the same renewable capacity as installed in 2015. Our main findings are that (i) the optimal technological and spatial capacity distribution — with about two thirds wind and one third solar — differs significantly from the actual mix (with a substantial overinvestment in photovoltaics); and (ii) ignoring the impact of interannual and intraday climate variability on the risk strongly impacts the optimal mix.

*Corresponding author (alexis.tantet@lmd.polytechnique.fr)

This study shows the importance of basing renewable energy development incentives on optimal strategies taking into account both the risk and the impact of climate variability on the latter.

Keywords: Energy, Renewable, Climate, Variability, Mediterranean

Highlights

- Taking into account the risk strongly impacts the optimal Italian renewable mix.
- The risk should account for climate variability from hours to decades.
- The actual Italian mix could yield a higher renewable penetration at a lower risk level after geographical optimization.

1. Introduction

The world net electricity generation is expected to increase by 45% between 2015 and 2040 (IEA, 2017). In view of climate change and energy security concerns, the renewable energies will inevitably play a major role in satisfying this growing demand. Non-hydropower Renewable Energy Sources (RES) are the fastest-growing energy sources for new generation capacity and their share is expected to grow from 7% of total world generation in 2015 to 15% in 2040, with more than half of this growth coming from the wind power (IEA, 2017).

However given the variable nature of the RES production and the need for a constant supply-demand balance, increasing penetration of renewables raises structural, technological and economical issues. On one hand, variability may lead to local power shortages or increased transmission congestion. This must be compensated at all times by an increased flexibility of the conventional generation systems such as coal plants or combined cycle gas turbines (Huber et al., 2014). On the other hand, it brings higher price instability along with a reduction of the wholesale prices. In the long run falling prices may ‘erode’ the returns of both renewable and conventional producers, eventually pushing the latter out of the market while they are essential to smooth out the fluctuations of renewable power output and ensure system stability. Thereby the possibilities for a future large-scale renewable capacity are limited (Hirth, 2013; Spiecker and Weber, 2014).

Technological and spatial diversification are possible strategies to circumvent the problem of intermittency. In Europe, wind and solar-generated electricity roughly have negatively correlated seasonal cycles, solar generation being maximal in summer and wind generation in winter (Heide et al., 2010). Spatial diversification is only applicable at large scale, whenever the RES variability is sufficient (see Widén (2011) for a study focusing on Sweden and Tsuchiya (2012) analyzing Japan). In light of those considerations, questions can be asked about how to implement this double strategy of diversification.

Technological and geographical optimization of renewable energy systems within a multi-objective framework has been discussed by several authors at continent and country scales. Complete electrical systems have been designed to quantify the requirements in installed power, transmission grid and storage capacity for a 100% of renewable energy scenario over Europe. For example, at the European scale, Heide et al. (2011) optimize the wind/solar mix in a fully renewable future European power system to reduce the storage and balancing needs; Rodríguez et al. (2014) do the same for the cross-border transmission capacities in the future; and Becker et al. (2014b) investigate the change in the optimal wind/solar mix in Europe as the transmission grid is enhanced. Becker et al. (2014a) optimize the wind/solar mix in the US to reduce storage needs and Nelson et al. (2012) simulate how a range of generation technologies, storage and transmission may meet the projected energy demand in the US at the least societal cost. Finally, Elliston et al. (2012) analyze how the Australian renewable mix should change in order to reduce the need for backup generation; and Lund and Mathiesen (2009) discuss feasible energy mix scenarios for a fully renewable electricity supply in Denmark.

Other conceptual frameworks with less ambitious energy targets have been explored at continental and regional scale by repowering the current installed renewable energy capacity. Repowering consists in fully decommissioning current renewable energy capacity and in re-allocating this capacity according to specific objectives (Del Río et al., 2011). For example, Beltran (2009) applies the mean-variance optimization techniques to infer the optimal energy mix; Roques et al. (2010) use similar methods to determine optimal wind power deployment among 5 European countries; Thomaidis et al. (2016) and Santos-Alamillos et al. (2017) use mean-variance optimization to assess the optimal wind and solar deployment and repowering actions in Spain. These studies use the Markovitz mean-variance portfolio theory or

analogous methodologies to define the optimal full re-allocation of existing power plants among regions. It relies on a trade-off between maximizing the mean renewable productivity while minimizing the aggregate renewable energy supply risk (i.e. variability). Note that these studies evaluate the risk of a given renewable energy mix using power production data only, whereas this risk clearly depends not only on power production, but also on the power consumption, which is also sensitive to climate. Our methodology presented in this paper allows to take this issue into account.

The majority of assessments of the optimal renewable energy mix are based on the statistical properties of the historical production and demand. Due to the only recent deployment of wind and solar energy systems the length of regional production and demand time series is often limited to a few years. This is not sufficient to properly take into account the effect of interannual climate fluctuations at time scales of the life cycle of a wind or a solar farm (Jourdier, 2015). In particular, it does not allow for the reliable estimation of the mean and the covariance of the production used in mean-variance analyses. Other studies, such as by Bremen (2010), rely on weather data to estimate the renewable production, but on periods still too short to resolve low-frequency climate variability.

In this article, we develop a methodology to take into account climate variability from decadal to hourly time scales in the optimization of wind and solar mixes. We determine the optimal geographical distribution of each renewable energy source using a mean-variance optimization. By designing models of electricity generation and consumption featuring climate variables such as the wind speed, the solar irradiance and the temperature, the former can be estimated from long climate time series. Applying supervised statistical learning methods together with intraday parameterizations from daily data, we are able to obtain robust estimates of the production and the demand for 23 years, at an hourly resolution. The climate time series are obtained from a hindcast of the 1989–2012 period over the Mediterranean region (Fig. 1b) performed in the framework of two international programs — the Hydrological Cycle in the Mediterranean Experiment (HyMEX Drobinski et al., 2014) and the Coordinated Downscaling Experiment for the Mediterranean (Med-CORDEX Ruti et al., 2016). For the first time, our methodology allows mean-variance analyses to take the impact of decadal to intraday climate variability on the renewable production and the demand into account.

As a first case study, we apply the model to Italy, allowing full re-commissioning of the actual renewable power plants. Italy offers an interesting

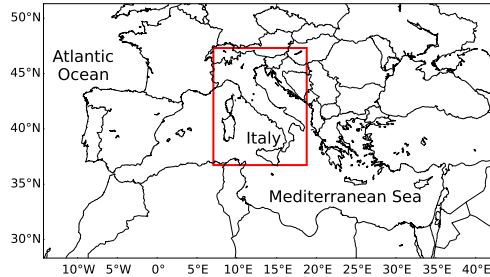
case study of a market with high renewables penetration as it has reached its quota of 17% renewables in final energy consumption in 2014, therefore implementing the 2009 Climate Package six years ahead of the 2020 horizon (GSE, 2015). This is the result of ambitious support policies for RES development that have generated and still subsidize a significant amount of new investments, notably in solar and wind power plants. These policies, however, do not take into account the impact of the variability of the RES production on the electricity system.

The analysis of the Italian renewable mix is based on wind and solar photovoltaics (henceforth PV) production and electricity consumption computed from the aforementioned regional climate simulation. Two cases for the optimization of the renewable energy mix are investigated: in the first case, the Italian electrical network is considered as unique and the overall mix is optimized; in the second one, each interconnected zone minimizes its own risk in priority, ignoring potential benefits from taking covariance and interconnection between zones into account. We show that, in the case of Italy, resolving such a large spectrum climate of time scales is essential to design a renewable energy system that remains optimal over its full life cycle. Considering the strong midlatitude variability of the European climate, we suspect this finding to generalize to other regions of Europe. This article thus gives a first proof of concept that it is possible to design optimal renewable energy mixes taking into account climate variability over a large range of temporal scales. Moreover, by replacing the hindcast simulations with future climate projections it is possible to use our methodology to study the impact of future climate evolution including anthropogenic climate change on the optimal renewable mix, however this is left for further research.

The remainder of the paper is structured as follows. Section 2 details the datasets, the simulation and the production and demand models on which is based the analysis of the optimal mix. The mean-variance optimization problem and its mathematical properties are presented in Section 3. In Section 4 we calculate the optimal geographical distribution of wind and solar generation in Italy and analyze the properties of these mixes in terms of weight given to each technology and in terms of occurrence frequency of shortage and saturation situations. The impact of climate variability on the optimal mixes is discussed in Section 5, together with their comparison with the actual wind-solar mix in Italy. In Section 6 we draw conclusions. The robustness of the numerical results to the climate data, the sampling and the model is tested in Appendix A.



(a) Italian electrical regions.



(b) Domain of the HyMeX/MED-CORDEX simulation covering Europe and the Mediterranean region. The rectangle indicates the domain of investigation of this study.

2. Data

2.1. GME and GSE databases

Time series of the hourly Italian regional electricity demand and of the yearly regional renewable capacity factors are used to design the demand and generation models. These variables are extracted from two publicly available databases provided respectively by the market operator GME¹ and the energy operator GSE². For this reason, we first briefly comment on the structure of the Italian electricity market and next describe the databases we use.

The Italian power market consists of 7 foreign virtual zones, 6 regional sub-markets, or bidding zones, and 5 poles of limited production. The 20 administrative regions composing the Italian territory are aggregated in the 6 bidding zones (Fig. 1a): Northern Italy (NORD), Central-Northern Italy (CNOR), Central-Southern Italy (CSUD), Southern Italy (SUD), Sardinia (SARD) and Sicily (SICI). Each zone has its own generation mix determined by historical and geographic reasons and characterized by a given level of efficiency. For instance, the Northern regions have larger hydroelectric production due to the proximity to the Alps. Inter-zonal transmission capacities are not equally distributed either.

¹ Gestore del Mercato Elettrico: <https://www.gse.it/dati-e-scenari/statistiche>

²Gestore dei Servizi Energetici: <https://www.gse.it/dati-e-scenari/statistiche>

The Italian power exchange, which is managed by the GME³ is composed of a spot market, a forward market and a platform for the physical delivery of contracts concluded on the financial derivatives segment of the Italian Stock Exchange. The spot market is composed of three sub-markets: the day-ahead, the intraday and the ancillary services markets. We focus on the day-ahead submarket. The liquidity of the day-ahead market, calculated as the ratio of volumes traded on the day-ahead market to the total volumes (including bilateral contracts) of the Italian power system, has increased between 2010 to 2015, passing from 62.6% with 198 operators in 2010 to 67.8% with 259 operators in 2015. The peak liquidity has been reached in 2013 with a 71.6% liquidity and 214 operators (GME, 2017).

The GME database encompasses hourly bids and offers in the wholesale electricity market from 2004 to 2017; the offers are identified by supplier's technology. The hourly electricity demand is appraised from this source. GSE annual reports (e.g. GSE, 2016) contain information about the yearly electrical production and the associated installed capacity⁴ detailed by region and sources from the beginning of 2008 to the end of 2016. At the beginning of this period, the installation of renewable energy capacity has shown a very rapid increase.

The regional time-mean capacity factors for PV and wind are calculated from this source. Since 2013, the PV and wind capacity factors are relatively stable. The demand and the capacity factors for PV and wind for the 2013–2017 period are presented in Table 1.

Table 2 summarizes the regional information on electrical installation — in particular the current installed capacity and transmission lines — at the end of 2015. In addition, Figure 2 represents the geographical distribution of the installed RES capacity. The PV (wind) installed capacity is 18.8 GW (8.9 GW). The Northern cross-border region contains the majority of the photovoltaic plants as well as the power lines. The Southern region has most of the Italian wind turbines and has the second transmission capacity, to foreign markets. Note that the electricity demand to Italy from other countries is not taken into account in this study. The share of the renewable energy production in the electricity demand over the six regions in 2015 is found to

³ The GME manage as well the OTC Registration Platform for forward electricity contracts that have been concluded off the bidding system.

⁴In the GSE reports, the capacity for a particular year is the installed capacity at the end of this year.

Region	Electrical demand (GWh/day — %)	Capacity Factor (PV — Wind)
NORD	312.2 — 56.5	12.1 — 20.4
CNOR	50.4 — 9.1	13.3 — 19.2
CSUD	82.0 — 14.9	14.1 — 18.8
SUD	44.2 — 8.0	15.6 — 20.9
SARD	26.7 — 4.8	14.5 — 19.6
SICI	36.8 — 6.7	15.9 — 18.7

Table 1: Regional electrical demand (from GME) and capacity factors for PV (blue) and wind energy (green, both from GSE) averaged over the 2013–2017 period.

be 19.4%.

Region	Transmission lines (MW)	PV installed capacity (MW)	Wind energy installed capacity (MW)
NORD	53400	8241	113
CNOR	4550	2256	133
CSUD	12720	2631	1582
SUD	34100	3600	4351
SARD	595	721	1001
SICI	10100	1302	1753

Table 2: Characteristic of regional electrical infrastructure based on GSE database.

2.2. Regional climate simulations

A third variable employed in our study is the multi-year series of RES production. The deployment of RES capacity being relatively recent (starting around 2008 in Italy), available time series of observed RES production are not sufficiently long to estimate statistics taking into account low-frequency climate variability. To take into account climate variability, RES production is instead computed using regional climate simulations covering the historical 1989 to 2012 period.

We use the version 3.1.1 of the Weather Research and Forecasting Model (WRF). WRF is a limited area model, non-hydrostatic, with terrain following eta-coordinate mesoscale modeling system designed to serve both operational forecasting and atmospheric research needs (Skamarock et al., 2005). The

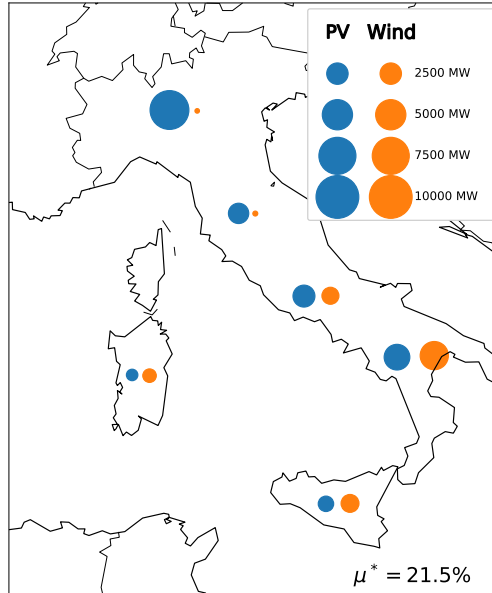


Figure 2: Geographical and technological distribution of the RES capacity installed by the end of 2015 in Italy (GSE database).

WRF simulation has been performed in the framework of HyMeX and MED-CORDEX programs with a 20 km horizontal resolution over the domain shown in Fig. 1b between 1989 and 2012 with initial and boundary conditions provided by the ERA-interim reanalysis and updated every 6 hr (Dee et al., 2011). The WRF simulation has been relaxed towards the ERA-I large scale fields (wind, temperature and humidity) with a nudging time of 6 hr (Salameh et al., 2010; Omrani et al., 2013, 2015). A detailed description of the simulation configuration can be found in e.g. Flaounas et al. (2013).

The simulation has been evaluated against ECA&D gridded precipitation and precipitation at the Mediterranean basin scale (Flaounas et al., 2013), and have been used to study heatwaves (Stéfanon et al., 2014; Chiriaco et al., 2014), heavy precipitation (Lebeau-pin Brossier et al., 2013, 2015; Berthou et al., 2014, 2015, 2016) and offshore wind energy potential assessment (Omrani et al., 2017) in a configuration coupled or not with a regional ocean model for the Mediterranean Sea (Drobinski et al., 2012). The simulation is

available on the HyMex/MED-CORDEX database⁵.

In the following sections, we describe the models used to estimate the wind and solar production and the electricity demand from the daily climate data. A large fraction of the variance of the production and of the demand is contained in shorter periods than a day and significantly impact the results of the mean-variance analysis (see Appendix A). The production and demand models thus all include a parameterization of the intraday variability.

2.2.1. Electricity production model

To compute wind energy production simulated daily-mean horizontal wind-speeds are interpolated at hub height (101 m) using an empirical power-law with exponent 1/7 (Justus and Mikhail, 1976). A transfer function based on the power curve of a particular wind turbine, the relatively representative Siemens SWT-2.3 MW-101m, is applied to the wind speed to compute the electrical production at each climate-data gridpoint (Omrani et al., 2017)⁶ Before applying the transfer function, the wind speed at hub height is multiplied by a factor $(\rho/\rho_0)^{(1/3)}$ accounting for deviations of the daily-mean air density ρ from the standard density ρ_0 for which the power curve has been obtained. The air density ρ is computed from the air temperature, pressure, and specific humidity at the surface from the WRF dataset using the ideal gas law for moist air⁷.

In addition, it is essential in the mean-variance analysis to take intraday fluctuations of the wind production into account. In order to take into account intraday wind fluctuations in the variance of the wind capacity factors, hourly realizations of the wind speed to be fed to the transfer function are obtained by randomly drawing samples from a Rayleigh distribution with mean given by the daily-mean speed at hub height. In other words, the wind magnitude $V(d, h)$ at day d and hour h at some gridpoint is drawn from the Rayleigh distribution

⁵ <ftp://www.medcordex.eu/MED-18/IPSL/ECMWF-ERAINT/evaluation/r1i1p1/IPSL-WRF311/v1/day/>

⁶ Note that, due to the bias correction (see Sect. 2.2.2), only the variability of the wind production may be sensitive to this choice of power curve.

⁷This correction is applied to the wind speed rather than directly to the wind production in order to shift the power curve horizontally rather than scale it vertically and hence preserve the cut-in and cut-out behavior of the turbine.

$$f(x|\sigma(d)) = \frac{x}{\sigma^2(d)} e^{-\frac{x^2}{2\sigma^2(d)}}, \quad x \geq 0 \quad (1)$$

$$\text{with } \sigma(d) = \sqrt{\frac{2}{\pi} \bar{V}(d)}, \quad (2)$$

where the mode $\sigma(d)$ is defined such that the mean of the Rayleigh distribution is equal to the available daily-mean wind-speed $\bar{V}(d)$ at the gridpoint. The effect of this parameterization of intraday wind fluctuations on the wind capacity factor of the north region is shown in Figure 3a and 3b, for a sample week in winter and another in summer 2010, respectively. One can see how the capacity factor increases with the daily-mean wind-speed, as well as the variance of intraday fluctuations due to the parameterization. However, it is clear from the figure that the intraday variability of the capacity factor is underestimated, a discrepancy that should be improved in future versions of the model.

We simulate the PV production for arrays at each gridpoint composed of multi-crystalline silicon solar cells. The crystalline silicon solar cell occupies about 90% of the PV market, among which multi-crystalline solar cells have the highest share at 53% and mono-crystalline solar cells have a 33% share (Hosenuzzaman et al., 2015). Each module has a nominal power of 250 Wm^{-2} for an area of 1.675 m^2 , resulting in a reference efficiency of about 15%⁸. The real efficiency of the cell is, however, dependent on its temperature, which is itself dependent on the air temperature and the wind from the WRF dataset and on the global tilted irradiance (see below). This dependence is modeled using the thermal model described in Duffie and Beckman (2013, Chap. 23)⁹. The efficiency of the overall electrical installation behind the modules is assumed to be of 86%.

Solar radiation from WRF is partitioned into direct, diffuse and reflected components (Duffie and Beckman, 2013, Chap. 2.16) at every gridpoint. This partitioning depends on the clearness index \bar{K}_T and elevation angle of the sun at the gridpoint. The quantity $\bar{K}_T(d)$, for some day d , is defined as the

⁸The nominal power itself is not important here, as only capacity factors are used

⁹ The thermal model is configured for relatively common parameter values for crystalline cells, i.e., for a temperature coefficient of 0.004 K^{-1} , a reference temperature of $25 \text{ }^\circ\text{C}$ and a cell temperature at nominal operating cell temperature of $46 \text{ }^\circ\text{C}$ (Skoplaki and Palyvos, 2009).

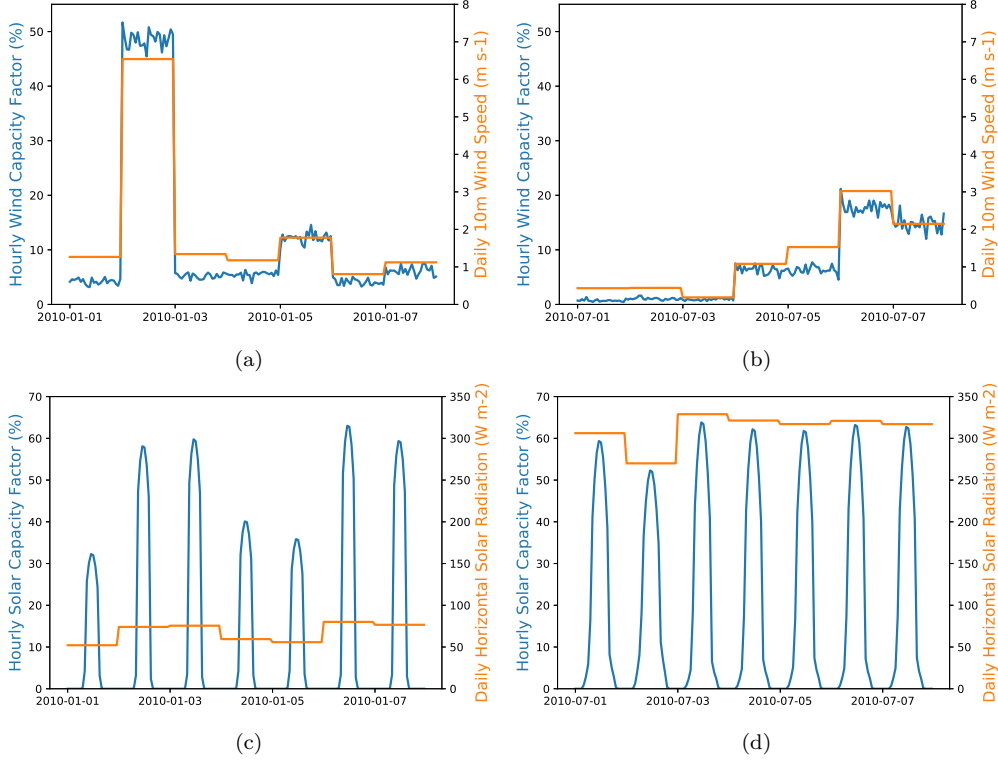


Figure 3: Illustration of the intraday parameterization of the wind (top) and solar (bottom) generation for the north region, the first week of January 2010 (left) and of July 2010 (right). Top panels represents both the computed hourly wind capacity factor (blue) and the daily-mean wind-speed (orange) from which the former is calculated. Bottom panels represents both the computed hourly solar capacity factor (blue) and the daily-mean horizontal surface radiation (orange) from which the former is calculated.

ratio of the horizontal radiation at ground level, $\bar{I}(d)$, to the corresponding radiation available at the top of the atmosphere, i.e. the extraterrestrial radiation $\bar{I}_0(d)$.

In order to take into account the effect of the diurnal cycle on the tilted irradiance, which accounts for most of the variance of the solar production (see Sect. Appendix A), the hourly extraterrestrial solar radiation, $I_0(d, h)$, is computed for every hour h from the calendar information. The hourly horizontal radiation at the surface, $I(d, h)$, for the hour h of the day d is then computed by multiplying the hourly extraterrestrial radiation $I_0(d, h)$ by the clearness index $\bar{K}_T(d)$, assumed constant throughout the day. In other words,

$$I(d, h) = \overline{K}_T(d) I_0(d, h) \quad (3)$$

$$\text{with } \overline{K}_T(d) = \frac{\overline{I}(d)}{\overline{I}_0(d)}. \quad (4)$$

Each array is assumed to be tilted by an angle equal to the latitude of the array and to face due South. To separate the diffuse component from the direct component of the global horizontal irradiance, the model from Reindl et al. (1990a) is used. For solar elevations below 10° and when the sun is behind the array, the direct horizontal irradiance is set to zero. The diffuse component of the tilted irradiance is computed following the model of Reindl et al. (1990b). The reflected component of the tilted irradiance depends on the zenith angle and follows the usual formula given by Duffie and Beckman (2013, Chap. 2.16) with an albedo of 0.2^{10} .

The effect of the diurnal cycle on the solar capacity factor of the north region is shown in Figure 3c and 3d, for a sample week in winter and another in summer 2010, respectively. One can see how the solar production varies with the diurnal cycle and how this cycle is modulated by the clearness of the atmosphere. On the other hand, no variability associated with intraday changes in, e.g., the cloud cover is present since the clearness index remains fixed throughout the day. The variance of the solar capacity factor may thus be underestimated. This discrepancy should, however, remain limited to the extent that intraday variations of the clearness index are averaged out by the regional averages.

2.2.2. Aggregation and bias correction

The regional wind and PV capacity factors are obtained by dividing the computed production by its nominal value and then aggregating at regional level on an hourly basis. In so doing, a strong bias (up to 100%) is found between the yearly-averages of the computed capacity factors and the region's capacity factors computed from the GSE data (Table 1). Since the second moment of the capacity factors roughly scales with their mean, we re-scale the computed capacity factors so that their average over the climate-data period (1989–2012) coincide with the GSE averages over the 2013–2017 period.

¹⁰The global tilted irradiance tends, however, to be dominated by its direct and diffuse components.

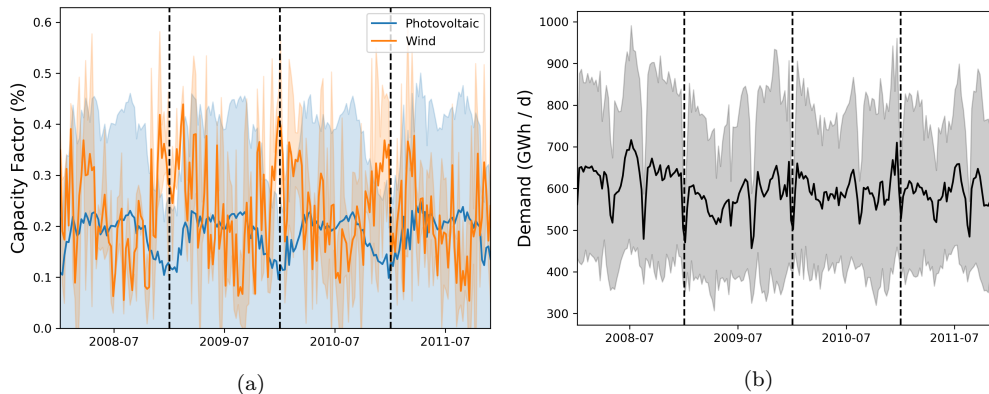


Figure 4: Time evolution of weekly-averaged PV (orange) and wind (blue) capacity factors (a) and demand (b). The shadings represent the standard deviation of the time series.

Figure 4a represents the weekly and regionally-averaged PV (orange) and wind (blue) capacity factors for a few sample years. Seasonal cycles of PV and wind energy production are phase shifted, with wind energy production (PV) peak yield in winter (summer). Wind energy production is also characterized by a stronger sub-seasonal variability when compared to PV at large-scale. Both the mean and the variance of the capacity factors for the wind production tend to be larger than those of the photovoltaic production. Albeit complementary over the year the wind and solar production requires additional energy inputs from other sources to counterbalance some recurrent short term deficit between the demand and the wind and solar production.

Let us insist that the present bias correction only corrects for differences in the first moment of the capacity factors with the observed values. For the analysis of Section 4, higher moments, in particular the variance and covariance, are also important. Although one expects the variance of the capacity factors to scale with their mean, discrepancies may persist. Our computations are tested against observations for bias in the variance in the Appendix A.

Note finally that the conventional production, which includes here thermal as well as hydropower plants, is not explicitly modeled in this study and is left for future work.

2.2.3. Electricity demand model

Since we are primarily interested in the impact of climate variability and change on the demand, the objective of the model is to predict the part of the

daily regional demand depending on climate, in particular on the surface air temperature¹¹, while preserving the statistics associated with other factors.

We follow a statistical learning approach (Hastie et al., 2009) whereby the model is trained against the regionally-averaged temperature data from the WRF model as input and the regionally-averaged demand data from GME as output, from the beginning of 2005 to the end of 2011 (i.e. the intersection of the climate with the demand record).

Let the electricity demand D_{in} for the region i at the time step n and for a particular type of day (see below) be given by

$$D_{in}(T_i) = f_i(T_{in}) + \epsilon_{in}. \quad (5)$$

where f_i , $1 \leq i \leq N$ is some real-valued function of the daily-mean temperature T_{in} in the region i at the time step n (T_{in} is the same for two hours of the same day), and the residual ϵ accounts for other factors impacting the demand, such as changes in the population, the economy, tourism, individuals decisions, etc.

It is known that the demand has a nonlinear dependence on the temperature. Electric heating is switched on only for lower temperatures, while air conditioning is switched on only for higher temperatures. This can be seen in Figure 4b for Italy. The demand has two main peaks per year, one during winter and one during summer, and lows in spring and fall and during holidays. In winter, the consumption peak is due to heating, especially in the northern part of Italy (see below). In summer, the consumption peak is due to tourism and air conditioning (Terna, 2016). Figure 4 shows that, except for the summer period, wind energy production is well correlated with the demand, whereas PV production is negatively correlated with the latter.

Once an individual appliance is switched on, its electricity consumption is to a first approximation linear in the temperature. Assuming, that all consumers behave in the same way and that a consumer switches the heater (air conditioning) for a constant temperature threshold T_H (T_C), we define the functions f_i as a piecewise-linear function of the temperature. In addition, the behaviour of the consumers differs significantly for the week days, Saturdays, and Sundays and holidays (respectively marked *work*, *sat* and *off*,

¹¹ Other variables, such as the specific humidity, the wind, or the irradiance were not found to significantly affect the demand, in this case.

in the following) and the demand is known to strongly depend on the hour of the day (see Appendix A). We thus choose to modulate the daily demand by a composite cycle which only depends on the hour of the day and the day type. The resulting model is given by,

$$f_i(T_{in}) = \begin{cases} f_i^{\text{work}}(T_{in}) & \text{if the day at } n \text{ is a working day} \\ f_i^{\text{sat}}(T_{in}) & \text{if the day at } n \text{ is a Saturday} \\ f_i^{\text{off}}(T_{in}) & \text{if the day at } n \text{ is a holiday,} \end{cases} \quad (6)$$

$$\begin{aligned} \text{with } f_i^{\text{work|sat|off}}(T_{in}) = & a_H^{\text{work|sat|off}} \Theta(T_H - T_{in})(T_H - T_{in}) g_n^{\text{work|sat|off}} \\ & + a_C^{\text{work|sat|off}} \Theta(T_{in} - T_C)(T_{in} - T_C) g_n^{\text{work|sat|off}} \\ & + a_0^{\text{work|sat|off}} g_n^{\text{work|sat|off}} \end{aligned}$$

where Θ is the Heaviside step function¹² and the coefficients $g_n^{\text{work|sat|off}}$ are given by the average — over all days of the same day type and all hours of the same hour of the day — of the observed demand on which the model is trained. The model (6) has a total of 9 parameters $a_H^{\text{work|sat|off}}$, $a_C^{\text{work|sat|off}}$ and $a_0^{\text{work|sat|off}}$ to be adjusted, for each region.

The resulting linear model is fitted assuming that the thresholds T_H and T_C are constant over all regions and all day types. These thresholds constitute two hyper-parameters that we select via a grid-search with a cross-validation (Hastie et al., 2009) over seven blocks of one year.

The linear model is fitted using the Bayesian ridge regression method (MacKay, 1992) both to avoid over-fitting and to take into account the variance arising from factors that are not fully resolved by the deterministic part of the model¹³. A time series of the hourly regional demand over 1989–2012 is predicted from the full length of the temperature record by randomly drawing samples from the posterior distribution of the model at each time step. The way the Bayesian model operates is illustrated in Figure 5 where

¹² The relationship between the demand and the ambient temperature in European countries is smoother than a piecewise-linear function (Bessec and Fouquau, 2008), in part due to the non-homogeneous behaviour of the consumers. Here, however, we prefer to keep the model as simple as possible using the above linear basis.

¹³ The implementation from *scikit-learn* (Buitinck et al., 2013) of the Bayesian ridge regression is used, whereby the residual and the weights are given zero-mean isotropic Gaussian priors. The variances of the latter are given as priors gamma distributions.

the predicted demand for the north region the first week of January 2010 is shown (plain blue line) together with the input daily-mean temperature T_{in} (plain orange line) and the series of composite daily cycles $g_n^{\text{work|sat|off}}$ (dashed blue line).

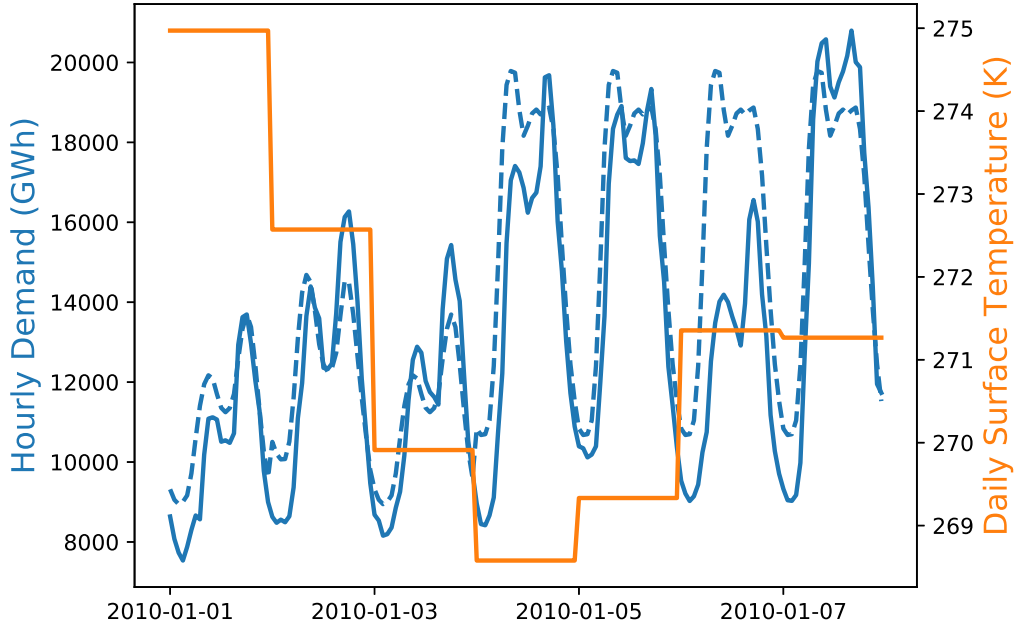


Figure 5: Illustration of the intraday parameterization of the demand for the north region the first week of January 2010. The computed hourly demand (plain blue line) is obtained by modulating the series of composite daily cycles $g_n^{\text{work|sat|off}}$ for each day type (dashed blue line) by the function (6) of the daily-mean temperature (plain orange line) and by adding random perturbations drawn from a normal distribution whose variance depends both on the noise in the demand and the uncertainty in the parameters of the model.

The resulting prediction of the regional demand is represented (only for daily-means) in Figure 6 versus the input temperature. The overall coefficient of determination is 0.73. One can see that the temperature and type of day dependence of the demand is most clear for the economically most dynamic north region. This is also true, yet to a lesser extent, for the central south region. The shaded regions show that the part of the demand that is not explained by the temperature model is compensated by the Bayesian perturbations, although in a random fashion.

To conclude, this method, as apposed to using the 13 years of observed

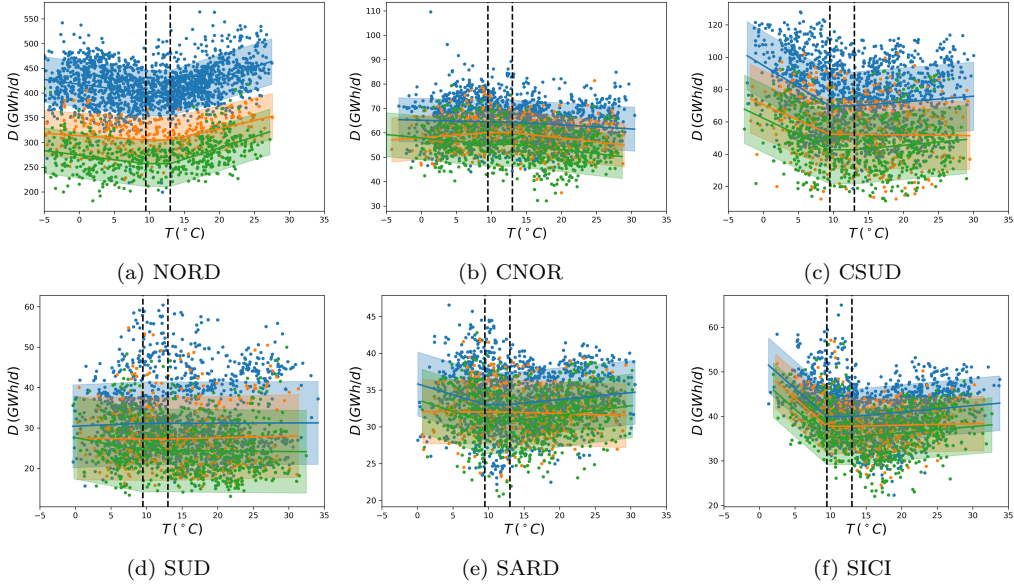


Figure 6: Daily-mean electricity demand for each zone versus the surface temperature. Each point is an observed realization of temperature and demand. The lines represent the functions $f_i^{\text{work|sat|off}}$ of the demand model, while the associated shaded regions represent the variance of the prediction. Blue, orange and green data points and functions correspond to working days, Saturdays, and Sundays and holidays, respectively. The two vertical dashed lines represent the temperature thresholds $T_H = 9.5$ and $T_C = 13.0$.

demand provided by GME, allows to estimate the demand over the longer record of the climate data so as to take into account variations in the demand due to low-frequency temperature variability¹⁴.

3. Mean-variance analysis

Geographic and technological diversification of renewable power plants is based on a mean-variance analysis that is inspired by Markowitz’s modern portfolio theory¹⁵.

¹⁴Note that, assuming that the demand-temperature relation in (6) is valid, low-frequency climate variability may be responsible in changes in the coefficients of the model over periods longer than the observed demand record over which the model is fitted.

¹⁵ See, for example, Mencarelli and D’Ambrosio (2018) for a survey on mathematical programming approaches for the portfolio selection problem.

In our context, the mean-variance analysis refers to the process of finding optimal spatial and technological distributions of renewable energy production achieving a trade-off between the mean penetration rate and some measure of the variance in the renewable energy production. The variance is a proxy for the risk: minimizing the variance corresponds to maximizing the diversification of the renewable configuration, which in turn lowers the variability of renewable energy penetration and improves the flexibility of the system and its resistance to shocks. In particular, a lower variance in the renewable energy mix is less demanding in services from conventional production (for which start up and shutting down services have a cost) or demand management.

Each renewable mix may then be represented in a mean-variance chart. As a bi-objective optimization problem (Miettinen, 1999), mixes are to be optimal in the Pareto sense. A solution is said to be Pareto optimal if there exists no feasible solution with a better or equal value for each of the objective functions (with at least one of these values being strictly better).

Let us consider as an example Figure 7. The points under or to the right of the frontier are by definition suboptimal and will be discarded by a rational investor. The area above or to the left of the frontier cannot be reached.

Two variations of the mean-variance analysis are considered here. In both, the mean RES penetration μ is given by the fraction of the expected total production over the expected total demand¹⁶, i.e.

$$\mu := \frac{\mathbb{E}[\sum_{\mathbf{k}} w_{\mathbf{k}} \eta_{\mathbf{k}}]}{\mathbb{E}[\sum_i D_i]} = \frac{\sum_{\mathbf{k}} w_{\mathbf{k}} \mathbb{E}[\eta_{\mathbf{k}}]}{\mathbb{E}[\sum_i D_i]}, \quad (7)$$

where $\mathbf{k} = (i, j)$ is the multi-index composed of a regional index i in {NORD, CNOR, CSUD, SUD, SARD, SICI} and a technological index j in {PV, wind}, the $w_{\mathbf{k}}$ are the installed capacities for each region and technology, the $\eta_{\mathbf{k}}$ are the corresponding predicted time-dependent capacity factors (Sect. 2.2.1) and the D_i are the predicted regional demands (Sect. 2.2.3). Note that, in the following numerical applications, statistics such as the expectation or the covariance are replaced by sample estimates from the full records (1989–2012).

¹⁶The normalization by the mean total demand in the definition (7) of the mean penetration has no effect on the relative distribution of the capacities in the optimal problem.

In the first strategy, also called *global*, it is assumed that each region produces electricity to satisfy the total demand in priority with no consideration for their local demand or for transmission constraints between regions. In other words, the electricity produced at a given location is immediately available to meet the overall Italian demand. In this case, the risk squared $\sigma_{\text{global}}^2(\mathbf{w})$ is defined as the variance of the sum of the regional RES productions normalized by the total demand, i.e.

$$\sigma_{\text{global}}^2(\mathbf{w}) := \mathbb{V} \left[\frac{\sum_{\mathbf{k}} w_{\mathbf{k}} \eta_{\mathbf{k}}}{\sum_i D_i} \right] = \sum_{\mathbf{k}} \sum_{\mathbf{l}} w_{\mathbf{k}} G_{\mathbf{k},\mathbf{l}}^{\text{global}} w_{\mathbf{l}}, \quad (8)$$

where $G_{\mathbf{k},\mathbf{l}}^{\text{global}} := \text{Cov}[\eta_{\mathbf{k}}/\sum_i D_i, \eta_{\mathbf{l}}/\sum_i D_i]$ is the covariance matrix between the capacity factors normalized by the total demand for each pair of regions and technologies.

In the second strategy, also called *regional*, each region attempts to satisfy its local demand in priority. If the electrical production is larger than the demand, electricity can be exported. The risk squared $\sigma_{\text{regional}}^2(\mathbf{w})$ is thus defined as the sum of the variances of the regional production normalized by the regional demand.

$$\begin{aligned} \sigma_{\text{regional}}^2(\mathbf{w}) &:= \sum_i \mathbb{V} \left[\frac{w_{(i,\text{PV})} \eta_{(i,\text{PV})} + w_{(i,\text{wind})} \eta_{(i,\text{wind})}}{ND_i} \right] \\ &= \sum_{\mathbf{k}} \sum_{\mathbf{l}} w_{\mathbf{k}} G_{\mathbf{k},\mathbf{l}}^{\text{regional}} w_{\mathbf{l}}, \end{aligned} \quad (9)$$

where $G_{\mathbf{k},\mathbf{l}}^{\text{regional}} := \text{Cov}[\eta_{\mathbf{k}}/(ND_i), \eta_{\mathbf{l}}/(ND_j)]$ if $i = j$, 0 otherwise¹⁷.

Note that in the global and the regional definitions of the risk, the RES production is normalized by the total and by the regional demand, respectively. This is particularly important for the regional optimization as normalizing by the local demand favors installing RES capacity in proportion to the latter.

The goal of this study being to assess the optimal recommissioning of the Italian renewable energy mix, we also consider constraining the total installed

¹⁷ The division by the number of regions N in (9) is there for comparison with the global risk (8).

RES capacity to its observed value $w_{\text{total}} = 27.7$ GW in 2015 (Sect. 2.1). The mean-variance analysis, thus, consists in solving, the optimization problem

$\min_{\mathbf{w}} \sigma_{\text{global regional}}^2(\mathbf{w})$	(10a)
$\max_{\mathbf{w}} \sum_{\mathbf{k}} w_{\mathbf{k}} \mathbb{E}[\eta_{\mathbf{k}}]$	(10b)
$\text{subject to } \sum_{\mathbf{k}} w_{\mathbf{k}} = w_{\text{total}}$	(10c)
$w_{\mathbf{k}} \geq 0 \quad \forall \mathbf{k}.$	(10d)

Assuming that the share of the demand that is not satisfied by RES is supplied by the conventional production normalized by the demand, the optimization problem (10) is equivalent to minimizing both the mean and the variance of the conventional production. Taking into account the full power flow of the transmission network together with the conventional production is left for future work.

The classical method used in the following Section 4 to approximate the optimal frontier numerically is explained in Appendix B. In this study, we refer to the optimal frontier as the curve $(\sigma_{\text{global|regional}}(\hat{\mathbf{w}}), \mu_{\text{global|regional}}(\hat{\mathbf{w}}))$, where the $\hat{\mathbf{w}}$ are the optimal solutions, although, strictly speaking, it is the risk squared that is minimized rather than the risk itself. The numerical results of the following section suggest that the so-defined frontier of the bi-objective problem (10) without the total capacity constraint (10c) is a half line with a positive slope that we refer to as the *mean-risk ratio* $\alpha_{\text{global|regional}}$. In other words, the optimal mixes for this problem are such that

$$\mu_{\text{global|regional}}(\hat{\mathbf{w}}) = \alpha_{\text{global|regional}} \sigma_{\text{global|regional}}(\hat{\mathbf{w}}), \quad (11)$$

In the following, we assume that this is indeed the case and the mean-risk ratio $\alpha_{\text{global|regional}}$ is used to diagnose the variants of the optimization problem. The proof of a rigorous mathematical result is left for future work.

4. Optimizing the distribution of wind and solar generation

Global and *regional* strategies are represented in Figure 8 (upper and lower panels respectively). Each point of the frontier represents an optimal combination of the capacities that maximizes the penetration for a given risk,

while satisfying the constraints. The mixes in the region below or to the right of the efficient frontier are suboptimal. The straight black line represents an approximation of the optimal frontier of the same problem, but with the total capacity constraint (10c) removed.

One may first observe that the optimal frontiers of the *regional* strategy (Fig. 8c) are shifted towards lower risk values compared to the frontiers of the *global* strategy (Fig. 8a). This shift is due to the different definitions (8) and (9) of the risk for the two strategies, respectively. It does not provide an objective argument to choose a regional policy over a global one. Instead, a trade-off exists between a strategy in which all regions cooperate for the national welfare and a strategy in which each region attempts to meet its own demand in priority¹⁸.

The mean-risk ratio $\alpha_{\text{global|regional}}$ (the slope of the black curves) of the optimization problems without the total capacity constraint is also given in the sub-captions. It is of 1.69 for the *global* strategy and of 3.71 for the *regional* strategy. Increasing the mean penetration of the optimal mix necessarily comes at the price of an increased risk. For the same level of total mean penetration, adding the constraint on the present total capacity (blue curve) necessarily deteriorates the risk.

The point at which both frontiers join (in black) corresponds to the optimal mix for which the total capacity constraint is inactive. This means that, for this level of mean penetration, the mix minimizing the risk naturally satisfies the total capacity constraint. It is thus the optimal mix satisfying the total capacity constraint that has the maximum mean-risk ratio. If no preference is put on maximizing the mean penetration or minimizing the risk, this optimal mix is attractive. In the following, we refer to this mix as the *maximum mean-risk ratio scenario*.

Yet, one may be interested in allowing for the deterioration of the mean-risk ratio in order to either decrease the risk or increase the total penetration. The blue dots in Fig. 8a and 8c correspond to the optimal mixes minimizing the risk. Their mean-risk ratio is then lowered to a value of 1.61 for the global strategy and of 3.53 for the regional strategy. We refer to this mix as the *minimum risk scenario*. For comparison with the actual mix the blue diamond in Fig. 8a and 8c represents the optimal mix that satisfies the same

¹⁸ Recall from Sect. 3, that in this study, the constraints imposed by the conventional production and transmission network capacities are not considered.

level of risk as the actual mix (gray dot) while maximizing the penetration rate. We refer to this mix as the *high penetration scenario*.

Some important properties of the optimal mixes with the constraint on the total capacity are represented on the right panels of Fig. 8. The PV ratio, i.e. the fraction of photovoltaic capacity in the mix, is plotted in orange. A value of 100% (0%) corresponds to 100% (0%) PV production 0% (100%) wind energy production. The plain and dashed green curves represent the frequency of occurrence of shortage and saturation, respectively. Shortage situations are associated with insufficient energy production. This situation corresponds to large scale blocking atmospheric patterns associated with cold or heat waves and low renewable energy production, especially from wind. These configurations lead to underproduction of electricity from PV and wind farms. Here, it is assumed that conventional generation units are able to meet up to 80% of the maximum demand modeled. Shortage then occurs if the photovoltaic and wind generation is not able to meet the rest of the demand. In this case, electricity needs to be imported from the neighboring countries (which is not explicitly accounted for in our modeling framework). Moreover, shortage situations result in an increase in the electricity market price and add pressure on power networks at continental scale. The second critical situation corresponds to network saturation, when electricity production from wind and PV plants is too large for the network and exceeds the technical limit of renewable energy fraction in the energy mix. In this study, saturation is defined to occur if more than 40% of the demand is met by photovoltaic and wind sources. High probability of occurrence of such situations may jeopardize the funding system for renewable energy infrastructures and may lead to network instability. It can also generate very low or even negative electricity prices (in Italy the prices cannot be negative, the floor is 0 euro), jeopardizing the profitability of the conventional power plants, which are essential for network security. Figures 8d and 8d show the fraction of installed PV capacity and the frequency of occurrence of shortage and saturation situations as a function of the mean penetration for the global and the regional strategy, respectively.

One can first see that the PV ratio is a decreasing function of the mean penetration reached by the mix. This is explained by the fact that the capacity factors from the wind generation are higher than those from the photovoltaic generation (cf. Tab. 1). Moreover, the shortage and the saturation curves (in green) have a distinct global minimum. The convexity of both curves is due to the fact that the probability of occurrence of extremes

increases with the risk and that the latter increases faster for both smaller and larger values of the mean penetration. The vertical lines in Fig. 8b and 8d represent the level of mean penetration for the minimum risk, maximum mean-risk ratio and high penetration scenarios. The minimum risk, the maximum mean-risk ratio and the high penetration scenarios respectively include 45%, 33% and 19% of PV capacity in the mix. The minimum risk scenario is relatively close to the minimum of saturation occurrence, while the maximum mean-risk ratio and the higher penetration scenarios are close to the minimum of shortage occurrence. Thus, favoring the minimization of the risk also allows to avoid saturation situations, while increasing the mean penetration allows to avoid shortage situations. Overall, the regional problem suggests installing more photovoltaic capacity than wind capacity compared to the global problem.

We represent in Figure 9 the resulting spatial distributions corresponding to the three scenarios for both the global and the regional problems. Focusing on the global problem, the minimum risk scenario distributes all the PV capacity in the north region, while achieving higher levels of mean penetration requires to move the PV capacity to the south and to Sicily. For all scenarios, the wind capacity is relatively spread over all regions but the central south one.

The resulting capacity distribution for the regional strategy (right panels of Fig. 9) is dramatically different from that for the global strategy. Indeed, most of the RES capacity is installed in the north region, whatever the scenario, with more wind capacity to increase the mean penetration. Only minimal capacity is installed in the central north, Sardinia and Sicily. This can be understood from both facts that the definition of the risk for the regional problem favors minimizing the risk of each region individually before minimizing the total risk and that the capacity factors in the risk are normalized by the demand. Since most of the demand occurs in the north (cf. Table 1), the variance of the capacity factors in that region is given less weight than the others in the definition of the regional risk. This corresponds to an incentive for the north region to satisfy its demand with its own RES resource first.

The results obtained so far show a strong dependence of the optimal mix on the level of risk that should be achieved. A mix favoring a high level of mean production at the expense of the risk — like in the high penetration scenario — yields a very different PV ratio and geographical distribution of the renewable capacity than in mixes favoring a lower level of risk — like in

the the maximum mean-risk ratio and minimum risk scenarios.

5. Discussion

The methodology developed in this article has specifically been designed (i) to exploit correlations between the RES production and the demand and between regions to optimize the renewable energy mix and (ii) to take into account the impact of the climate variability on this mix. This allows us to discuss in this section the role played by climate variability and the potential for improvement of the actual Italian mix.

5.1. Impact of climate variability on the mix

To assess the impact of interannual climate variability (as found in the CORDEX data) on the mix, we repeat the mean-variance analysis successively using data blocks of one year, from 1989 to 2010. In other words, the mix of the 22 mixes that we obtain is optimized for the climatic conditions of a given year. As a result, the optimal mix for one year may be different from the optimal mix for another year due to low-frequency climate variability.

We focus on the mixes maximizing the mean-risk ratio α_{global} for the global strategy. From the 22 values of the mean-risk ratios associated with each year, we obtain an estimation of their mean value and of an interval containing 95% of their realizations (see Appendix A for more explanations). The mean-risk ratio α_{global} averages to 1.71 and 95% of its distribution belongs to the centered interval [1.57, 1.85]. Thus, even though the average of the yearly mean-risk ratio is close to the one of 1.69 obtained in Section 4 using the full record, interannual climate variability in the CORDEX data is responsible for year-to-year variations of the mean-risk ratio of up to 8.2%.

To go further, we represent in Figure 10 the geographical and technological distribution of the mixes for the year 1989, with a particularly low mean-risk ratio of 1.60, and for the year 1996, with a particularly high mean-risk ratio of 1.85. Due to the steeper optimal frontier for 1996, the mix for the maximum mean-risk ratio scenario achieves both a higher mean total penetration and a lower risk than that for 1989. The distribution of the RES capacity for 1996 also differs from that for 1989 as more wind capacity is installed in the central south and Sardinia regions.

It is thus clear that the low-frequency climate variability has a large impact on the optimal mix and that the latter should be taken into account

in order to assess future optimal RES mixes that are robust to changes in the climate.

5.2. Comparison with the 2015 Italian mix

The 2015 (actual) Italian mix is composed of 68% PV and 32% wind energy capacity (see Table 2). As can be seen from Figure 2, the largest fraction of installed PV capacity is in the North of Italy, whereas most of wind capacity is located in the South. Indeed, Italy has started investing in renewable energy resources since 1991 (with the feed in tariff CIP6), with the objective of developing national energy sources and so decreasing the dependency on imported gas. Historically, the bulk of PV has been developed in Northern Italy, where the entrepreneurial background has favored local business exploiting renewable energy resources subsidies. Wind farms have been installed in Sicily and Sardinia, where regional specific incentives have been set. Southern Italy has invested in renewable energy more recently, after the reinforcement of the main North-South transmission line.

To compare the actual mix with the optimization results, the actual mean penetration and risk are computed for the actual capacity distribution using the same capacity factor and demand data as used to obtain the optimal frontiers of Fig. 8. The corresponding mix is represented by the gray dot in the same figure. This point is not visible in Fig. 8c, for the regional problem, because the value of its risk is several orders of magnitude larger than the x-axis limits. For both the global and the regional problem, the gray point is to the right of the optimal frontiers. The actual mix installed in Italy is thus sub-optimal. For the global problem, this mix reaches a level of mean penetration comparable to that of the minimum risk scenario, but its risk is about 29% larger than that of the latter and its PV ratio about 49% higher. The capacity distribution of the optimal mixes in Figure 9 may also be compared to that of the 2015 Italian mix in Figure 2. The actual mix appears to be closest to the minimum risk and maximum mean-risk ratio scenarios of the global problem (Fig. 9a). However, the actual mix favors photovoltaic over wind capacities, especially in the north, and that more RES capacity is installed in the central south region. This is in strong contrast with all optimal scenarios. For instance, the maximum mean-risk ratio scenario yields about two third (one third) of wind (solar) capacity. However, these differences may partly be attributable to biases in the intraday variance, as the model tends to underestimate intraday variance

of the wind capacity factors and to overestimate the intraday variance of the solar capacity factors (see Table A.3)¹⁹.

6. Conclusion

This work is aimed at developing a proof-of-concept of an integrated modelling framework dedicated to the elaboration of optimal scenarios of renewable energy mix. The proposed framework relies on regional climate simulations. It is shown to be of practical interest for both short and long term renewable energy management, as the model is able to take into account variability in the renewable production and electricity demand from hourly to interannual time scales. The model allows to derive different scenarios consisting in either maximizing the total renewable energy penetration or minimizing the total risk by taking advantage of (anti-)correlations between regions and technologies in an optimal way. Different optimization strategies have been chosen to establish the renewable energy mixes: the first optimization strategy, the global one, maximizes the national Italian welfare by taking full advantage of correlations between region and by ignoring any potential network constraint; the second strategy, the regional one, assumes that each region maximizes its own welfare in priority, disregarding the potential benefit of exploiting cross-regional correlations to improve the national welfare.

The main results have been obtained by relying on the full length of the climate record (23 years), in order to resolve the impact of interannual climate variability on the optimal mixes as well as possible. By computing additional mixes using only one year of climate data for all available years and by assessing the differences between these mixes, we could show that the impact of interannual climate variability on the optimal renewable energy mix should not be neglected.

Comparison with the actual Italian renewable energy mix shows that the actual mix is closer to the global strategy than to the regional strategy as the renewable capacity is relatively evenly spread among regions. However, the scenario maximizing the ratio of the total mean penetration over the risk yields about two thirds (one third) of wind (solar) capacity, in strong contrast with the actual mix containing one third of installed wind capacity. This

¹⁹ Improving the resolution of the intraday variability of the production in the model may require to develop a better parameterization or to rely on additional data. This is left for future work.

result suggests that incentives should be given for the development of wind and solar plants to better approach the optimal distribution. However, the reasons for such differences are difficult to identify as the actual renewable energy capacity deployment did not follow an optimization elaborated at country scale, but relied on regional policies. Conversely, our optimal energy mix scenarios rely on selected optimization strategies and may be prone to biases in the resolution of the intraday variance of the production.

The current framework takes the variance, or risk, of the fraction of the demand covered by the renewable production as a proxy for the flexibility service needed from the hydro and conventional production. Our modeling framework would benefit from the translation of this risk into an economic or a climate cost. For that purpose, the hydro and the conventional production would have to be modeled, taking into account reserve constraints and priority orders between these energy sources. The transmission network would also have to be modeled to take network constraints into account as well as the arbitrage between producing locally and importing/exporting between regions.

The current framework is also adapted to consider optimal strategies in a warming climate using 21st century projections from general circulation models. Finally, the generalization of such integrated modeling tool at Euro-Mediterranean scale is a priority. Our framework opens the way for the study of energy transition scenarios at the European scale based on precise modeling of climate variability and climate change. The interconnected regions will then be replaced with interconnected countries.

7. Acknowledgments

AT is grateful to Mathilde Mougeot and Jérôme Collet for insightful discussions on modeling the electricity demand. This work was conducted in the framework of the TREND-X research program on energy transition at Ecole Polytechnique, which benefited from the support of the Ecole polytechnique fund raising – TREND-X Initiative. This research was also supported by the ANR project FOREWER (ANR-14-414 CE05- 0028) and the HyMeX project (HYdrological cycle in The Mediterranean EXperiment) through the working group Renewable Energy, funded by INSU-MISTRALS program. The TREND-X model is open-source and available on demand. The full documentation of the model is available at <https://www.lmd.polytechnique>.

`fr/~atantet/TRENDX/`. It includes a reference guide and a tutorial, allowing one to easily reproduce the results of this study.

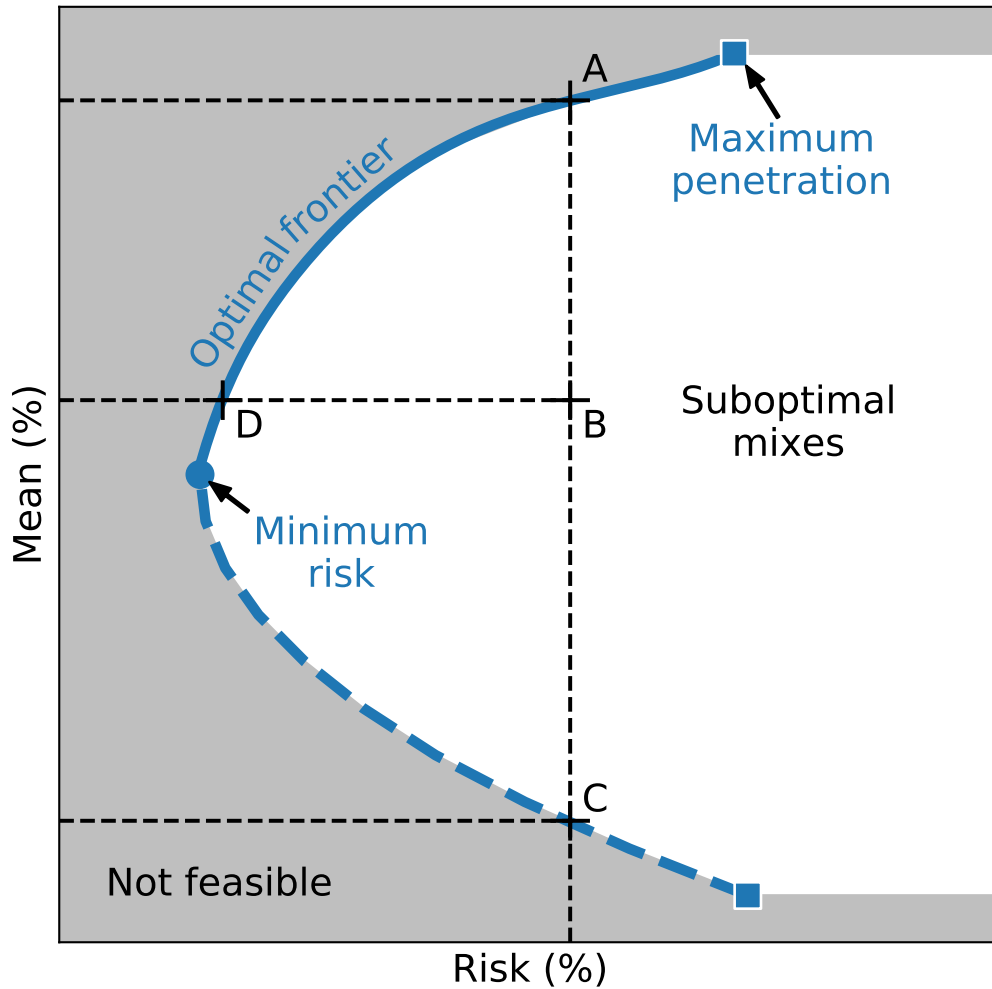


Figure 7: Example of the optimal frontier of a mean-variance bi-objective optimization problem. The optimal frontier is one-dimensional and represented by a plain blue line. Mixes in the white region to the right of the frontier are suboptimal. Points in the gray region to the left of the frontier are not feasible. In this example, the optimal frontier is bounded below by a minimum-risk optimal-mix (blue dot) below which the risk may only increase. The optimal frontier is bounded above by a maximum-penetration optimal-mix above which higher penetration mixes are not feasible due to the constraints of the problem. The point B is an example of suboptimal mix, since a higher mean penetration is achievable for the same risk (point A) and a lower risk is achievable for the same mean penetration (point D). The dashed blue line is obtained by minimizing the risk for a range of target mean penetration values. These solutions are, however, not Pareto optimal. For instance, point C yields the same risk as point A but achieves a lower mean penetration. Thus, A “dominates” C.

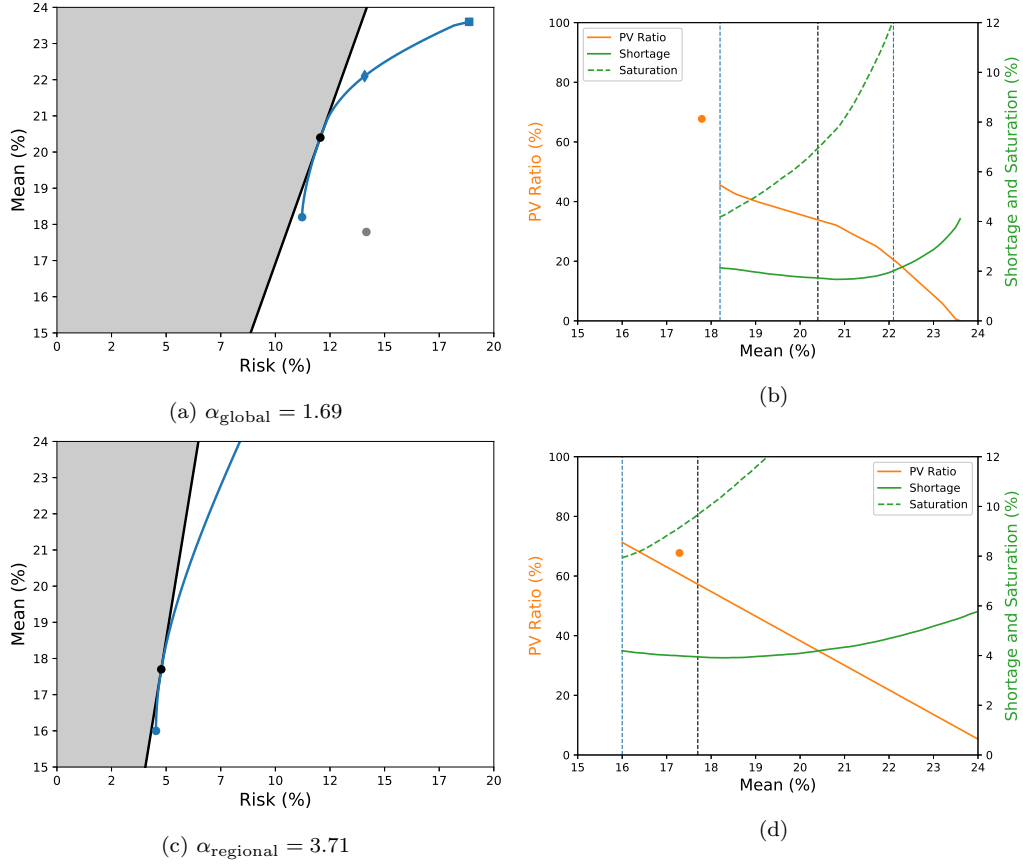


Figure 8: Approximations of the optimal frontiers (left) and of the corresponding electricity mix characteristics (right) for the global (top) and the regional (bottom) optimization problems. The blue curve in the left panels represents the approximation of the optimal frontier of the optimization problem (10). The straight black line is the optimal frontier of the same problem, but with the total capacity constraint (10c) removed. The approximations were obtained using a discretization step of 0.1%. The black dot where the black frontier is tangent to the blue one corresponds to the optimal electricity mix for which the total capacity constraint is inactive (i.e. where adding this constraint has no effect on the results of the optimization problem). The blue dots in panels (a) and (c) correspond to the optimal energy mix for which the risk is minimized while satisfying the total capacity constraint. The gray dot in panel (a) is obtained from the same capacity factor and demand data but applying the actual capacities installed in Italy in 2015 (cf. Tab. 2). The blue diamond corresponds to the optimal mix achieving the same level of risk as the actual mix in gray while maximizing the mean penetration. The blue squares correspond to limits beyond which it is not possible to further decrease or increase the mean total penetration while satisfying the total capacity constraint and the bounds. The values of the mean-risk ratio of the unconstrained optimal frontiers, given by (11), are also reported in the sub-captions. On the right panels are represented the fraction of photovoltaic capacity in the mix (plain orange line), or PV ratio as well as the shortage (plain green line) and saturation (dashed green line) frequencies versus the mean penetration. The blue and black dashed vertical lines mark the mean penetration values corresponding to the blue and black dots and the blue diamond on the left panels. The orange dot represents the PV ratio for the actual capacities installed in Italy in 2015.

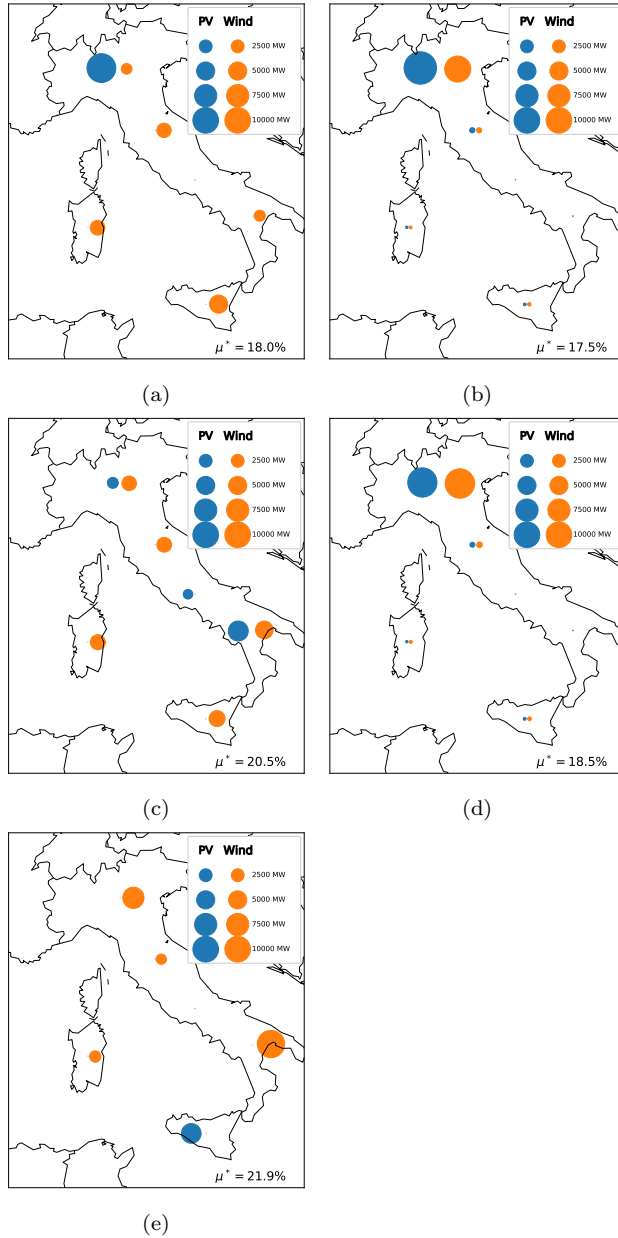


Figure 9: Geographical and technological distribution of the RES capacity for the global (left) and regional (right) optimal mixes respecting the total capacity constraint. The top, middle and bottom panels represent the optimal mixes for the minimum risk, maximum mean-risk ratio and higher penetration scenarios, respectively (blue dot, black dot and blue diamond in Fig. 8a and 8c). Note that there is no plot of the higher penetration scenario for the regional strategy since the regional risk of the actual mix is too high to be achieved by an optimal mix.

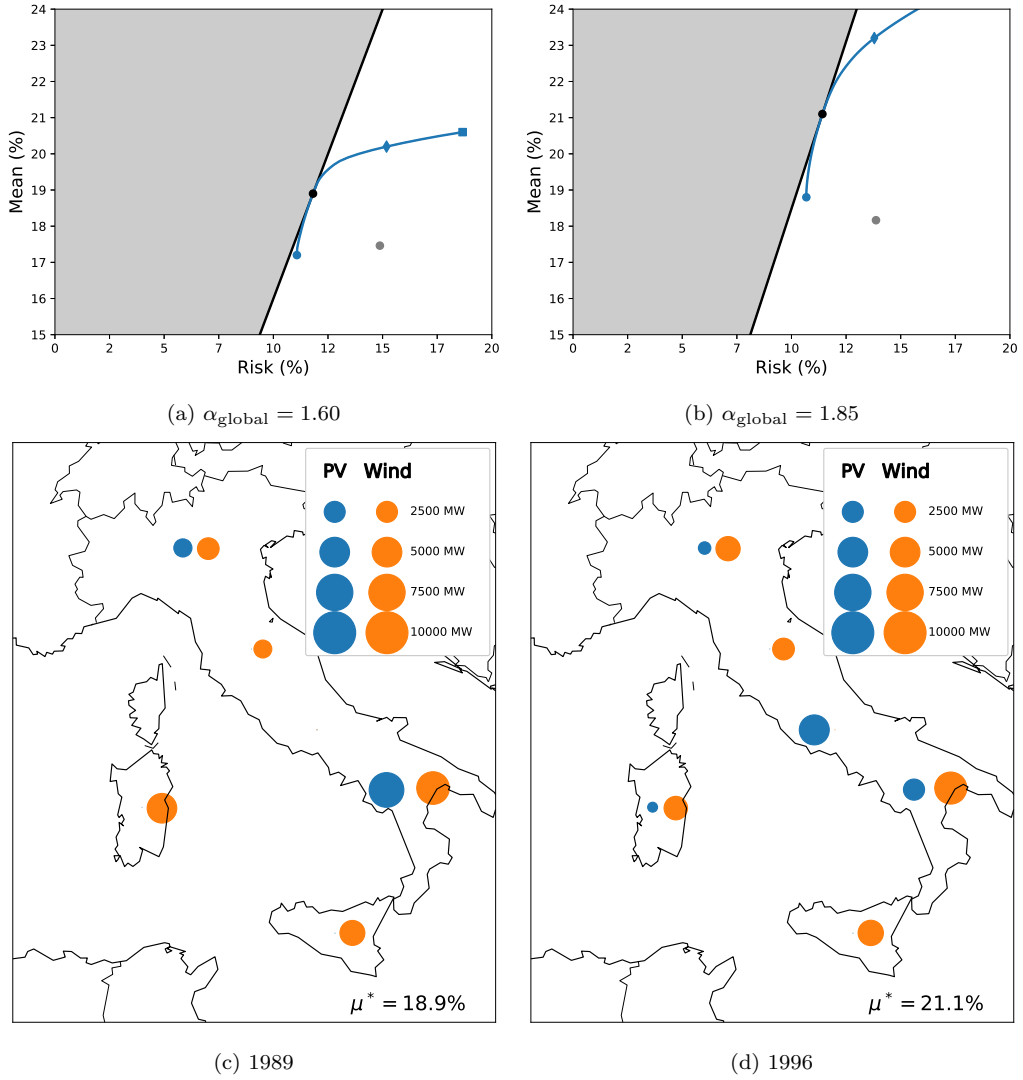


Figure 10: Approximated Pareto frontiers (top) and geographical and technological distribution of the RES capacity (bottom) for the global optimization problem solved for the years 1989 (left) and 1996 (right). The approximations were obtained using a discretization step of 0.1%. The legend is the same as for Fig. 8 and Fig. 9.

Appendix A. Robustness to the climate data and comparison with observations

The robustness of the results presented in Section 4 is tested here. The impact of the choice of the climate data on the optimization problem is first tested together with the importance of using hourly rather than daily data. Next, we check the covariance matrix computed from the modeled production and demand (see Sect. 2) with that obtained from the one year of observed hourly capacity factors available to us. The impact of errors in the computed covariance matrix on the optimization results is also tested.

Recall from Section 2.2.1 that when computing the capacity factors from the climate data, biases in the mean capacity factors are corrected using the observed mean capacity factors from GSE. However, this bias correction does not correct for the variance. Unfortunately, the capacity factors from GSE are only sampled each year (see Sect. 2.1), which is not sufficient to compute the total of the variance up to hourly periods. Instead, crossing the hourly production data per region and technology from the ENTSO-E transparency platform²⁰ with the annual capacity factors from the GSE database allowed us to reconstruct hourly capacity factors for the year 2016. This is done by multiplying the yearly-averaged ENTSO-E production data with the GSE capacity factors of the corresponding years to obtain the corresponding installed capacity for each year and then by dividing the 2016 ENTSO-E hourly production by the computed capacity linearly interpolated between the end of 2015 and the end of 2016²¹.

In addition, we can use climate data from hourly simulations of the MERRA-2 reanalysis (from 1980/01/01 to 2018/07/31, Gelaro et al., 2017) and, applying the same methodology as presented in Section 2, we can test the robustness of the results to changes in the climate data and in the sampling, whether hourly (by parameterizing the intraday variability using the CORDEX data) or daily (by down-sampling the MERRA-2 data)²².

²⁰ <https://transparency.entsoe.eu/>

²¹ Using both datasets is needed because the ENTSO-E production and capacity data from ENTSO-E are aggregates from production units of different sizes.

²² It is difficult to assess which of the CORDEX or the MERRA-2 datasets is most appropriate for the present study. On the one hand, the CORDEX data is from the WRF regional model at a high resolution so that it is able to resolve small-scale processes that the MERRA-2 reanalysis cannot resolve. On the other hand, the MERRA-2 reanalysis is one generation early compared to the ERA-Interim reanalysis used to force the WRF

The Figure A.11 shows the approximated optimal frontiers of the global optimization problem obtained by applying the production models to the CORDEX (left) and the MERRA-2 (right) datasets with an hourly (top) and a daily (bottom) sampling. Let us first note the dramatic difference between the frontiers obtained from hourly and from daily data. Indeed, the mean-risk ratio α_{global} is overestimated when using daily data, which translates into an underestimated risk as the mean penetration is increased. This can be understood from the fact that, while the mean capacity factors remain unchanged, the variance in the modeled daily PV and wind capacity factors explains only 8.3% and 68% of the total variance computed from the hourly MERRA-2 data, respectively.

Overall, the qualitative picture remains unchanged when choosing one dataset or the other, but important quantitative differences exist. First, the mean-risk ratio α_{global} tends to be larger when using the CORDEX rather than the MERRA-2 data, resulting in a larger increase of the risk with the mean penetration. Comparing Fig. A.12 — representing renewable capacity distribution for the maximum mean-risk ration obtained from the MERRA-2 data — with Fig. 9c, one can see that using CORDEX data results in more PV capacity than when using the MERRA-2 data and some regional disparities exist.

To better understand the impact of the sampling on the optimal mixes and to assess the capacity of the production and demand models to reproduce the spectral characteristics of the observations, we compare in Table A.3 the percentage of the variance of the Italian average of the electricity demand (top), solar capacity factor (middle) and wind capacity factor (bottom) explained by periods greater than a year, less or equal than a year and greater than a day and less or equal than a day, from the observations and the two reanalysis²³. We can see that demand model does not resolve the little interannual variability of the observed demand. This is to be expected if this variability is due to socio-economical factors that are only taken into account stochastically by the Bayesian model. Moreover, both the model using the

regional model and may thus resolve some meso-scale processes better than the ERA-Interim reanalysis. In this study, the focus is on the CORDEX data in order to pave the way to using future climate projections from CORDEX in future works.

²³ The variance explained by each frequency band is calculated from the variance of the respectively low-pass, band-pass and high-pass filtered time series using running averages as filters.

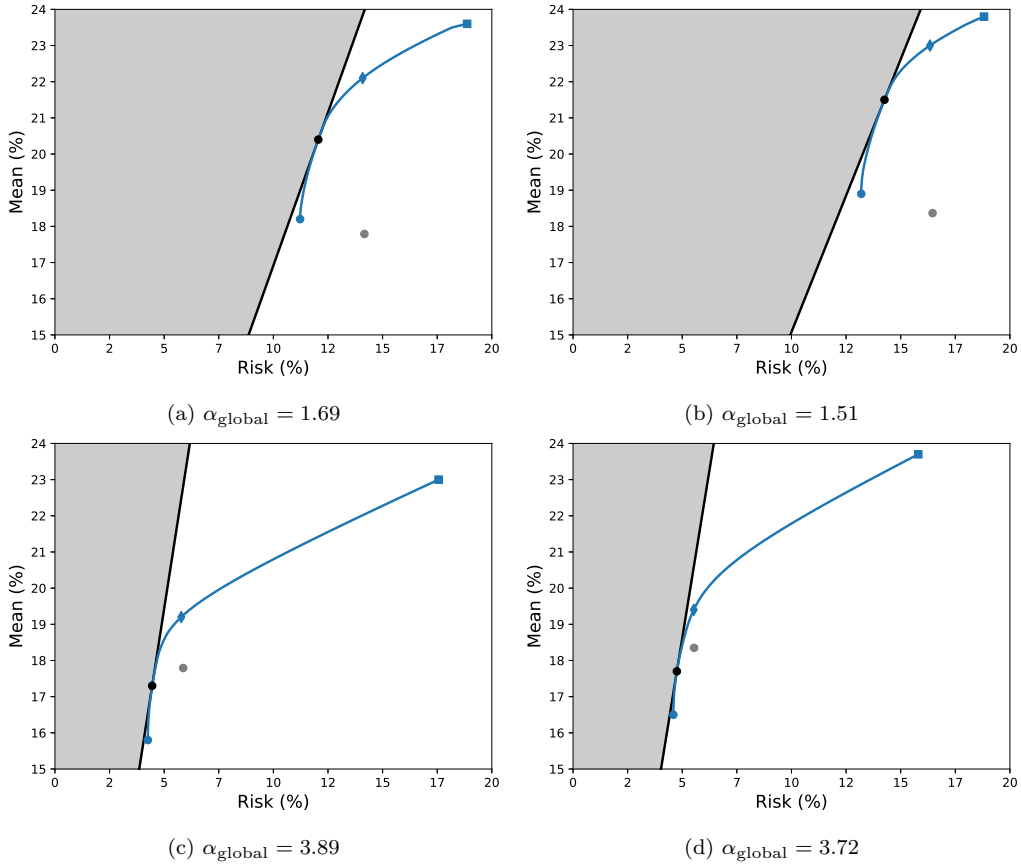


Figure A.11: Mean-risk frontier computed using hourly (top) and daily (bottom) nclimate data from the CORDEX (left) and the MERRA-2 (right) datasets. The approximations were obtained using a discretization step of 0.1%. The legend is the same as for Fig. 8.

CORDEX and the MERRA-2 data underestimate (resp. overestimate) the seasonal (resp. intraday) variability, although this is less true when using the CORDEX data. The intraday variability of the solar capacity factors is overestimated by the model whether the CORDEX or the MERRA-2 is used²⁴.

Finally, the intraday variability of the wind capacity factor is underestimated (resp. overestimated) by the production model using the CORDEX

²⁴ Note that the interannual variability of the capacity factors cannot be evaluated against observations since only one year of observations is available.

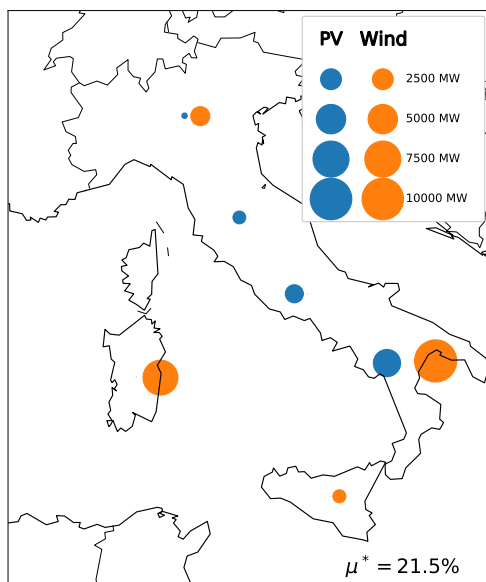


Figure A.12: Comparison of the geographical and technological distribution of the RES capacity for the optimal mixes minimizing the mean-risk ratio for the global strategy obtained from the MERRA-2 dataset with that obtained from the CORDEX dataset in Fig. 9c.

(resp. MERRA-2) data. This suggests that the parameterization of the intraday variability of the wind capacity factors is not entirely satisfactory and should be improved in future works.

Unfortunately, the year 2016 is not part of the climate simulation, so that the variance of the capacity factors computed from the climate data may not be directly compared with the observed capacity factors for that year. Instead, we test if the observed capacity factor variances for the different regions and technologies are significantly different from those computed from the CORDEX and the MERRA-2 data. In other words, we test whether the variability in the climate data is sufficient to explain differences between the variance from the observed and from the computed capacity factors. To do so, the capacity factors computed from the climate data are grouped by year. The standard deviations of the capacity factors for each year, region and technology are then computed. Next, assuming that the realizations of the standard deviations of the capacity factors for each year are independent and identically normally distributed, the 2.5% and 97.5% percentiles of a Student-t distribution are rescaled in order to estimate the

	Interannual	Seasonal	Intraday
<hr/>			
Electricity Demand			
GME	3.6	37.0	59.4
CORDEX	0.0	33.1	66.9
MERRA-2	0.0	26.5	73.4
<hr/>			
Solar Capacity Factor			
ENTSO-E & GSE	No data	8.9	91.1
CORDEX	0.0	2.8	97.2
MERRA-2	0.0	4.0	96.0
<hr/>			
Wind Capacity Factor			
ENTSO-E & GSE	No data	84.6	15.4
CORDEX	0.6	98.9	0.5
MERRA-2	0.5	69.0	30.5

Table A.3: Comparison of the percentage of the variance of the Italian average of the electricity demand (top), solar capacity factor (middle) and wind capacity factor (bottom) explained by periods greater than a year, less or equal than a year and greater than a day and less or equal than a day, from the GME, ENTSO-E & GSE, CORDEX and MERRA-2 datasets.

interval containing 95% of these realizations. Finally, we test whether the standard deviations computed from the observed capacity factors fall within the 95% most likely realizations of the capacity factors computed from the climate data. These results are summarized in Table A.4. One can see that significant differences exist between the observed and the computed values of the standard deviations. In particular, calculations from the CORDEX data tend to overestimate (resp. underestimate) the standard deviations of the solar (resp. wind) capacity factors. Differences exist as well between the results from the CORDEX and the MERRA-2 data²⁵. A possible direction to improve the bias in the variance of the capacity factors is either, whenever possible, to compute the production from hourly data from climate models better resolving the intraday variability, or to use other variables or datasets

²⁵ The quality of this test depends on the sampling and in particular on the length of the record. The CORDEX record being shorter than the MERRA-2 record by 16 years, this test may suffer from insufficiently resolving the low-frequency variability.

	NORD	CNOR	CSUD	SUD	SARD	SICI
Solar Capacity Factor						
ENTSO-E & GSE	17.9	18.6	18.8	20.8	20.2	20.1
CORDEX	17.4 ± 0.7	19.1 ± 0.7	20.4 ± 0.6	22.2 ± 0.7	21.1 ± 0.6	22.9 ± 0.7
MERRA-2	17.8 ± 0.9	19.9 ± 0.9	21.4 ± 0.9	23.0 ± 0.9	21.8 ± 0.7	23.4 ± 0.8
Wind Capacity Factor						
ENTSO-E & GSE	18.8	19.4	19.9	17.9	21.1	17.7
CORDEX	17.7 ± 2.0	16.7 ± 1.5	16.1 ± 1.1	15.4 ± 1.0	15.8 ± 1.0	15.0 ± 1.1
MERRA-2	23.3 ± 3.1	21.9 ± 1.9	21.5 ± 2.3	16.3 ± 1.1	16.4 ± 0.8	17.1 ± 1.1

Table A.4: Comparison of the standard deviations of the computed and observed solar (top) and wind (bottom) hourly capacity factors contained in a year for each region (%). The observed capacity factors are computed from hourly generation data from ENTSO-E and yearly capacity factors from GSE for the year 2016. The sets of computed CORDEX and MERRA-2 values comprise 23 and 39 years, respectively. In each cell, the values and the errors are the mean over all the years and the interval containing 95% of the distribution. The latter are estimated using a Student t distribution (assuming that the yearly standard deviations are independent and identically normally distributed). Computed values for which the observed standard deviation falls outside of these intervals are marked in red.

to design a better parametrization of the intraday variability of the wind speed and the clearness index.

To compare the optimal mixes obtained from direct observations of the capacity factors with those obtained from the climate data the approximated optimal frontiers and the maps of the optimal mix obtained from the observed capacity factors and demand data are represented in Figures A.13 and A.14, respectively. Once again, the qualitative picture is fairly robust, but important quantitative differences exist. Yet, it is not possible to tell from Figures A.13 and A.14, whether these differences are due to errors in the model or due to interannual variability.

Instead, it is possible to directly compare the mean-risk ratio of the optimal mixes obtained from the observed capacity factors for 2016 and those obtained from the climate data, as is done above for the standard deviations of the capacity factors. Solving the optimization problem without the total capacity constraint and using only one year of CORDEX generation data for each year yields a mean-risk ratio α_{global} averaging to 1.71 and with 95% of the distribution estimated to belong to the centered interval [1.57, 1.85]. The value of 1.72 of the mean-risk ratio computed from the observed values of the demand and generation data for 2016 lies within this interval. Thus,

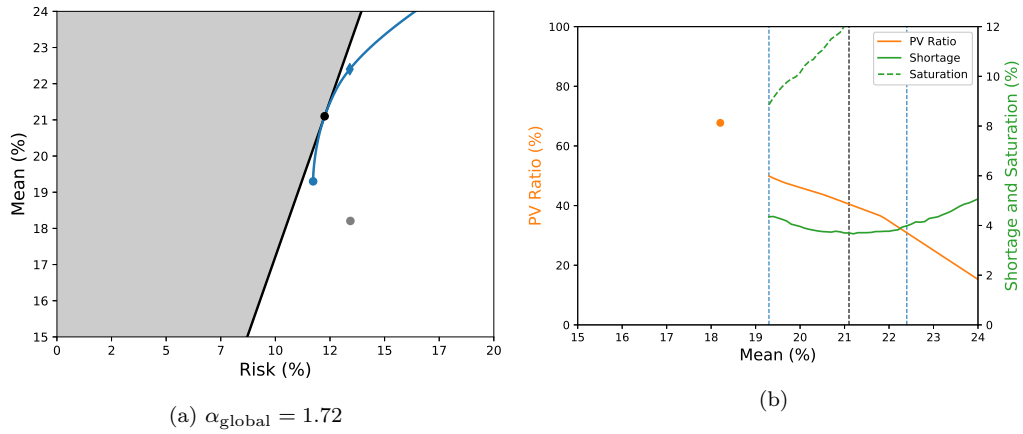


Figure A.13: Approximated optimal frontiers (left) and electricity mix characteristics (right) for the global optimization problem solved from observations, to be compared with the approximated optimal frontiers and mix characteristics obtained from the CORDEX data in Fig. 8.

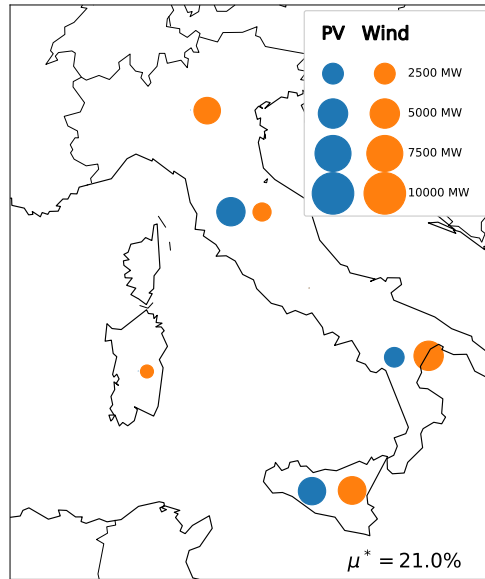


Figure A.14: Geographical and technological distribution of the RES capacity for the maximum ratio scenario of the global optimization problem obtained from observations, to be compared with the RES distribution obtained from the CORDEX data in Fig. 9c.

according to this test, the difference between this observed value and the

CORDEX values are not significant.

Appendix B. Method to find an approximation of the optimal frontier

The results of the following sections are valid for both the *global* and the *regional* strategies. First of all, let us define two single objective sub-problems which represent a restriction of the bi-objective problem we aim at solving. We follow the well-known method called ϵ -constraint, see, for example, Miettinen (1999, Chap. II.3).

The first subproblem $(P)^{\min}$ is defined by

$$\min_{\mathbf{w}} \quad \sigma^2(\mathbf{w}) \quad (\text{B.1})$$

$$\text{subject to} \quad \sum_{\mathbf{k}} w_{\mathbf{k}} = w_{\text{total}} \quad (\text{B.2})$$

$$w_{\mathbf{k}} \geq 0 \quad \forall \mathbf{k} \quad (\text{B.3})$$

$$\sum_{\mathbf{k}} w_{\mathbf{k}} \mathbb{E}[\eta_{\mathbf{k}}] \geq \mu^* \mathbb{E}[\sum_i D_i] \quad (\text{B.4})$$

where μ^* ranges from $l^{of2}/(\sum_i D_i)$ to $u^{of2}/(\sum_i D_i)$ where l^{of2} and u^{of2} are the lower and the upper bound on the value of the second objective function (B.5) (below), respectively.

The second subproblem $(P)^{\max}$ is defined by

$$\max_{\mathbf{w}} \quad \sum_{\mathbf{k}} w_{\mathbf{k}} \mathbb{E}[\eta_{\mathbf{k}}] \quad (\text{B.5})$$

$$\text{subject to} \quad \sum_{\mathbf{k}} w_{\mathbf{k}} = w_{\text{total}} \quad (\text{B.6})$$

$$w_{\mathbf{k}} \geq 0 \quad \forall \mathbf{k} \quad (\text{B.7})$$

$$\sigma^2(\mathbf{w}) \leq (\sigma^*)^2 \quad (\text{B.8})$$

where $(\sigma^*)^2$ is defined in $[l^{of1}, u^{of1}]$ where l^{of1} and u^{of1} are the lower and the upper bound on the value of the first objective function (B.1), respectively.

Appendix B.1. The algorithm

The idea is to find the best value of (B.1) by solving $(P)^{\min}$ for each value of μ^* . As μ^* is continuously defined, it is, of course, impossible to solve it for each possible value of it. Thus, we discretize the possible values of μ^* with a step of 0.1% and find just a subset of the optimal frontier.

Note that solving $(P)^{\min}$ for different values of μ^* is not enough to guarantee that the solutions found are not dominated by any other solution. For

this reason, it is necessary to alternate between solving $(P)^{\min}$ for a given value of μ^* , then solving $(P)^{\max}$ by setting $(\sigma^*)^2$ equal to the objective function value found by solving $(P)^{\min}$, then update μ^* accordingly and solving $(P)^{\max}$ again, and so on, until the values μ^* and $(\sigma^*)^2$ cannot be updated anymore.

If we do not perform this alternating update and solve of the two subproblems it might happen that solving only $(P)^{\min}$ or $(P)^{\max}$ might produce a dominated solution. Take the case of Fig.5(a), for example: for a value of risk equal to 15 there are two points in the light blue curve, i.e., mean equal to 15 and equal to 22. If we solve $(P)^{\min}$ with constraint (B.4) corresponding to mean at least equal to 15, we will find a solution of objective function value 15, but it might correspond either to a mean of value 15 (and in this case the solution would be dominated) or to a mean of value 22.

Appendix B.2. How to find the bound on the RHS of (B.4) and (B.8)

We aim at finding the lower and upper bound of the objective functions (B.1) and (B.5) so as to be able to define the interval over which we can vary the right-hand-side of constraints (B.4) or (B.8).

For the lower bound of (B.1) (resp. the upper bound of (B.5)) it is simple: we drop (B.5) (resp. (B.1)) and solve the corresponding single objective problem, which is a relaxation of original bi-objective problem. To find the upper bound of (B.1) (resp. the lower bound of (B.5)) it is sufficient to we drop (B.5) (resp. (B.1)), invert the direction of (B.1) (resp. (B.5)) and solve the corresponding single objective problem.

References

- Becker, S., Frew, B. A., Andresen, G. B., Zeyer, T., Schramm, S., Greiner, M., Jacobson, M. Z., 2014a. Features of a fully renewable US electricity system: Optimized mixes of wind and solar PV and transmission grid extensions. *Energy* 72, 443–458.
- Becker, S., Rodriguez, R. A., Andresen, G. B., Schramm, S., Greiner, M., 2014b. Transmission grid extensions during the build-up of a fully renewable pan-European electricity supply. *Energy* 64, 404–418.
- Beltran, H., 2009. *Modern Portfolio Theory Applied To Electricity Generation Planning*. PhD Thesis, University of Illinois at Urbana-Champaign.
- Berthou, S., Mailler, S., Drobinski, P., Arsouze, T., Bastin, S., Béranger, K., Flaounas, E., Brossier, C. L., Somot, S., Stéfanon, M., 2016. Influence of submonthly air–sea coupling on heavy precipitation events in the Western Mediterranean basin. *Quarterly Journal of the Royal Meteorological Society* 142 (S1), 453–471.
- Berthou, S., Mailler, S., Drobinski, P., Arsouze, T., Bastin, S., Béranger, K., Lebeaupin-Brossier, C., Dec. 2014. Prior history of Mistral and Tramontane winds modulates heavy precipitation events in southern France. *Tellus A: Dynamic Meteorology and Oceanography* 66 (1), 24064.
- Berthou, S., Mailler, S., Drobinski, P., Arsouze, T., Bastin, S., Béranger, K., Lebeaupin-Brossier, C., 2015. Sensitivity of an intense rain event between atmosphere-only and atmosphere–ocean regional coupled models: 19 September 1996. *Quarterly Journal of the Royal Meteorological Society* 141 (686), 258–271.
- Bessec, M., Fouquau, J., 2008. The non-linear link between electricity consumption and temperature in Europe: A threshold panel approach. *Energy Econ.* 30 (5), 2705–2721.
- Bremen, L. V., 2010. Large-Scale Variability of Weather Dependent Renewable Energy Sources. In: Troccoli, A. (Ed.), *Management of Weather and Climate Risk in the Energy Industry*. NATO Science for Peace and Security Series C: Environmental Security. Springer, Netherlands, pp. 189–206.

- Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Müller, A. C., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., Vanderplas, J., Joly, A., Holt, B., Varoquaux, G., 2013. API design for machine learning software: Experiences from the scikit-learn project. arXiv 1309.0238 [cs.LG], 15.
- Chiriaco, M., Bastin, S., Yiou, P., Haeffelin, M., Dupont, J.-C., Stéfanon, M., 2014. European heatwave in July 2006: Observations and modeling showing how local processes amplify conducive large-scale conditions. *Geophysical Research Letters* 41 (15), 5644–5652.
- Dee, D., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., Van De Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kallberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., De Rosnay, P., Tavolato, C., Thépaut, J. N., Vitart, F., 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* 137 (656), 553–597.
- Del Río, P., Calvo Silvana, A., Iglesias Gómez, G., 2011. Policies and design elements for the repowering of wind farms: A qualitative analysis of different options. *Energy Policy* 39 (4), 1897–1908.
- Drobinski, P., Anav, A., Lebeaupin Brossier, C., Samson, G., Stéfanon, M., Bastin, S., Baklouti, M., Béranger, K., Beuvier, J., Bourdallé-Badie, R., Coquart, L., D’Andrea, F., de Noblet-Ducoudré, N., Diaz, F., Dutay, J.-C., Ethe, C., Foujols, M.-A., Khvorostyanov, D., Madec, G., Mancip, M., Masson, S., Menut, L., Palmieri, J., Polcher, J., Turquety, S., Valcke, S., Viovy, N., Jul. 2012. Model of the Regional Coupled Earth system (MORCE): Application to process and climate studies in vulnerable regions. *Environmental Modelling & Software* 35, 1–18.
- Drobinski, P., Ducrocq, V., Alpert, P., Anagnostou, E., Béranger, K., Borga, M., Braud, I., Chanzy, A., Davolio, S., Delrieu, G., Estournel, C., Boubrahmi, N. F., Font, J., Grubišić, V., Gualdi, S., Homar, V., Ivančan-Picek, B., Kottmeier, C., Kotroni, V., Lagouvardos, K., Lionello, P., Llasat, M. C., Ludwig, W., Lutoff, C., Mariotti, A., Richard, E., Romero,

- R., Rotunno, R., Roussot, O., Ruin, I., Somot, S., Taupier-Letage, I., Tintore, J., Uijlenhoet, R., Wernli, H., Jul. 2014. HyMeX A 10-Year Multidisciplinary Program on the Mediterranean Water Cycle. *Bull. Am. Meteorol. Soc.* 95 (7), 1063–1082.
- Duffie, J., Beckman, W., 2013. *Solar Engineering of Thermal Processes*, 4th Edition. John Wiley & Sons, Hoboken.
- Elliston, B., Diesendorf, M., MacGill, I., Jun. 2012. Simulations of scenarios with 100% renewable electricity in the Australian National Electricity Market. *Energy Policy* 45, 606–613.
- Flaounas, E., Drobinski, P., Vrac, M., Bastin, S., Lebeaupin-Brossier, C., Stéfanon, M., Borga, M., Calvet, J.-C., Jun. 2013. Precipitation and temperature space–time variability and extremes in the Mediterranean region: Evaluation of dynamical and statistical downscaling methods. *Climate Dynamics* 40 (11-12), 2687–2705.
- Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., Zhao, B., May 2017. The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate* 30 (14), 5419–5454.
- GSE, 2015. *Rapporto Statistico 2015: Energia da fonti rinnovabili in Italia*. Tech. rep., GSE.
- GSE, 2016. *Rapporto Statistico 2016: Energia da fonti rinnovabili in Italia*. Tech. rep., GSE.
- Hastie, T., Tibshirani, R., Friedman, J., 2009. *The Elements of Statistical Learning*, 2nd Edition. Springer, New York.
- Heide, D., Greiner, M., von Bremen, L., Hoffmann, C., 2011. Reduced storage and balancing needs in a fully renewable European power system with excess wind and solar power generation. *Renew. Energy* 36 (9), 2515–2523.

- Heide, D., von Bremen, L., Greiner, M., Hoffmann, C., Speckmann, M., Bofinger, S., 2010. Seasonal optimal mix of wind and solar power in a future, highly renewable Europe. *Renew. Energy* 35 (11), 2483–2489.
- Hirth, L., 2013. The market value of variable renewables. The effect of solar wind power variability on their relative price. *Energy Econ.* 38, 218–236.
- Hosenuzzaman, M., Rahim, N. A., Selvaraj, J., Hasanuzzaman, M., Malek, A. B. M. A., Nahar, A., Jan. 2015. Global prospects, progress, policies, and environmental impact of solar photovoltaic power generation. *Renewable and Sustainable Energy Reviews* 41, 284–297.
- Huber, M., Dimkova, D., Hamacher, T., 2014. Integration of wind and solar power in Europe: Assessment of flexibility requirements. *Energy* 69, 236–246.
- IEA, 2017. *World Energy Outlook 2017*. Tech. rep., IEA.
- Jourdier, B., 2015. Wind resource in metropolitan France: Assessment methods, variability and trends. Ph.D. thesis, Ecole Polytechnique, Palaiseau.
- Justus, C. G., Mikhail, A., 1976. Height variation of wind speed and wind distributions statistics. *Geophysical Research Letters* 3 (5), 261–264.
- Lebeaupin Brossier, C. L., Bastin, S., Béranger, K., Drobinski, P., Feb. 2015. Regional mesoscale air–sea coupling impacts and extreme meteorological events role on the Mediterranean Sea water budget. *Climate Dynamics* 44 (3-4), 1029–1051.
- Lebeaupin Brossier, C. L., Drobinski, P., Béranger, K., Bastin, S., Orain, F., 2013. Ocean memory effect on the dynamics of coastal heavy precipitation preceded by a mistral event in the northwestern Mediterranean. *Quarterly Journal of the Royal Meteorological Society* 139 (675), 1583–1597.
- Lund, H., Mathiesen, B. V., 2009. Energy system analysis of 100% renewable energy systems-The case of Denmark in years 2030 and 2050. *Energy* 34 (5), 524–531.
- MacKay, D. J. C., 1992. Bayesian Interpolation. *Neural Computation* 4 (3), 415–447.

- Mencarelli, L., D'Ambrosio, C., 2018. Complex portfolio selection via convex mixed-integer quadratic programming: A survey. *International Transactions in Operational Research*.
- Miettinen, K. M., 1999. *Nonlinear Multiobjective Optimization*. Springer, New York.
- Nelson, J., Johnston, J., Mileva, A., Fripp, M., Hoffman, I., Petros-Good, A., Blanco, C., Kammen, D. M., 2012. High-resolution modeling of the western North American power system demonstrates low-cost and low-carbon futures. *Energy Policy* 43, 436–447.
- Omrani, H., Drobinski, P., Arsouze, T., Bastin, S., Lebeau-pin-Brossier, C., Mailler, S., Feb. 2017. Spatial and temporal variability of wind energy resource and production over the North Western Mediterranean Sea: Sensitivity to air-sea interactions. *Renewable Energy* 101, 680–689.
- Omrani, H., Drobinski, P., Dubos, T., Nov. 2013. Optimal nudging strategies in regional climate modelling: Investigation in a Big-Brother experiment over the European and Mediterranean regions. *Climate Dynamics* 41 (9-10), 2451–2470.
- Omrani, H., Drobinski, P., Dubos, T., Mar. 2015. Using nudging to improve global-regional dynamic consistency in limited-area climate modeling: What should we nudge? *Climate Dynamics* 44 (5-6), 1627–1644.
- Reindl, D. T., Beckman, W. A., Duffie, J. A., Jan. 1990a. Diffuse fraction correlations. *Solar Energy* 45 (1), 1–7.
- Reindl, D. T., Beckman, W. A., Duffie, J. A., Jan. 1990b. Evaluation of hourly tilted surface radiation models. *Solar Energy* 45 (1), 9–17.
- Rodríguez, R. A., Becker, S., Andresen, G. B., Heide, D., Greiner, M., 2014. Transmission needs across a fully renewable European power system. *Renew. Energy* 63, 467–476.
- Roques, F., Hiroux, C., Saguan, M., 2010. Optimal wind power deployment in Europe-A portfolio approach. *Energy Policy* 38 (7), 3245–3256.
- Ruti, P. M., Somot, S., Giorgi, F., Dubois, C., Flaounas, E., Obermann, A., Dell'Aquila, A., Pisacane, G., Harzallah, A., Lombardi, E., Ahrens, B.,

- Akhtar, N., Alias, A., Arsouze, T., Aznar, R., Bastin, S., Bartholy, J., Béranger, K., Beuvier, J., Bouffies-Cloch e, S., Brauch, J., Cabos, W., Calmanti, S., Calvet, J. C., Carillo, A., Conte, D., Coppola, E., Djurdjevic, V., Drobinski, P., Elizalde-Arellano, A., Gaertner, M., Gal an, P., Gallardo, C., Gualdi, S., Goncalves, M., Jorba, O., Jord a, G., L’Heveder, B., Lebeaupin-Brossier, C., Li, L., Liguori, G., Lionello, P., Maci as, D., Nabat, P.,  onol, B., Raikovic, B., Ramage, K., Sevault, F., Sannino, G., Struglia, M. V., Sanna, A., Torma, C., Vervatis, V., 2016. Med-CORDEX initiative for Mediterranean climate studies. *Bull. Am. Meteorol. Soc.* 97 (7), 1187–1208.
- Salameh, T., Drobinski, P., Dubos, T., 2010. The effect of indiscriminate nudging time on large and small scales in regional climate modelling: Application to the Mediterranean basin. *Quarterly Journal of the Royal Meteorological Society* 136 (646), 170–182.
- Santos-Alamillos, F. J., Thomaidis, N. S., Usaola-Garc a, J., Ruiz-Arias, J. A., Pozo-V azquez, D., 2017. Exploring the mean-variance portfolio optimization approach for planning wind repowering actions in Spain. *Renew. Energy* 106, 335–342.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., Powers, J. G., Jun. 2005. A Description of the Advanced Research WRF Version 2. Tech. Rep. NCAR/TN-468+STR, NCAR.
- Skoplaki, E., Palyvos, J. A., May 2009. On the temperature dependence of photovoltaic module electrical performance: A review of efficiency/power correlations. *Solar Energy* 83 (5), 614–624.
- Spiecker, S., Weber, C., 2014. The future of the european electricity system and the impact of fluctuating renewable energy - A scenario analysis. *Energy Policy* 65, 185–197.
- St efanon, M., Drobinski, P., D’Andrea, F., Lebeaupin-Brossier, C., Bastin, S., Mar. 2014. Soil moisture-temperature feedbacks at meso-scale during summer heat waves over Western Europe. *Climate Dynamics* 42 (5-6), 1309–1324.
- Terna, 2016. Sustainability Report 2016. Tech. rep., Terna, Rome.

- Thomaidis, N. S., Santos-Alamillos, F. J., Pozo-Vázquez, D., Usaola-García, J., 2016. Optimal management of wind and solar energy resources. *Comput. Oper. Res.* 66, 284–291.
- Tsuchiya, H., 2012. Electricity supply largely from solar and wind resources in Japan. *Renew. Energy* 48, 318–325.
- Widén, J., 2011. Correlations between large-scale solar and wind power in a future scenario for Sweden. *IEEE Trans. Sustain. Energy* 2 (2), 177–184.