



Estimation and Forecast of PV Generation on a Regional Scale Using Satellite Data and High Resolution WRF Output Combined with Machine Learning Techniques

Alessandro Perotto^a, Enrico Maggioni^a, Francesco Spada^a, Marco Pierro^{b,c}, Cristina Cornaro^{b,d}, David Moser^c, Bruno Fasoli^e, Danilo Pederiva^e, David Tschager^e

^aIdeam Srl, via Frova 34 Cinisello Balsamo, Italy, e-mail: alessandro.perotto, enrico.maggioni, francesco.spada@ideamweb.com

^bDepartment of Enterprise Engineering, University of Rome Tor Vergata, Via del Politecnico 1, 00133 Rome, Italy, e-mail: cornaro@uniroma2.it, marco.pierro@gmail.com

^cEURAC Research, Viale Druso 1, 39100 Bolzano, Italy e-mail: david.moser@eurac.edu

^dCHOSE, University of Rome Tor Vergata, Via del Politecnico 1, 00133 Rome, Italy 7

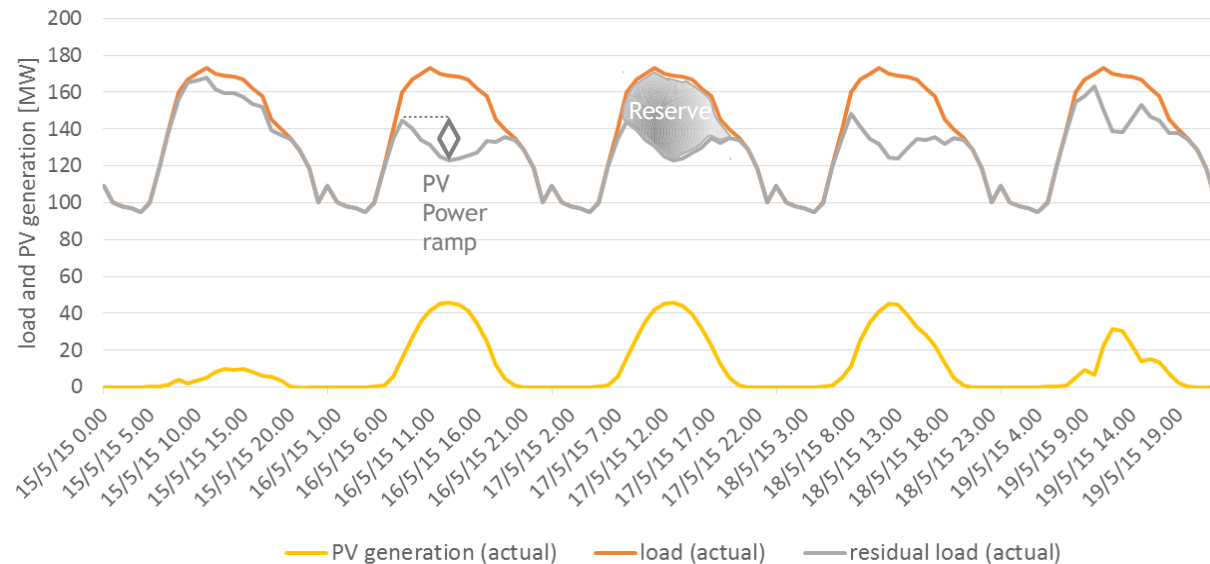
^eAlperia Energy Srl, Via Dodiciville 8, 39100 Bolzano

16th EMS meeting & 11th European Conference on Applied Climatology (ECAC), 12-16 September 2016, Trieste Italy



Introduction and Motivation

Large shares of PV power can compromise the stability of the electrical grid on a regional level with the introduction into the electric load of a stochastic variability dependent on meteorological conditions.

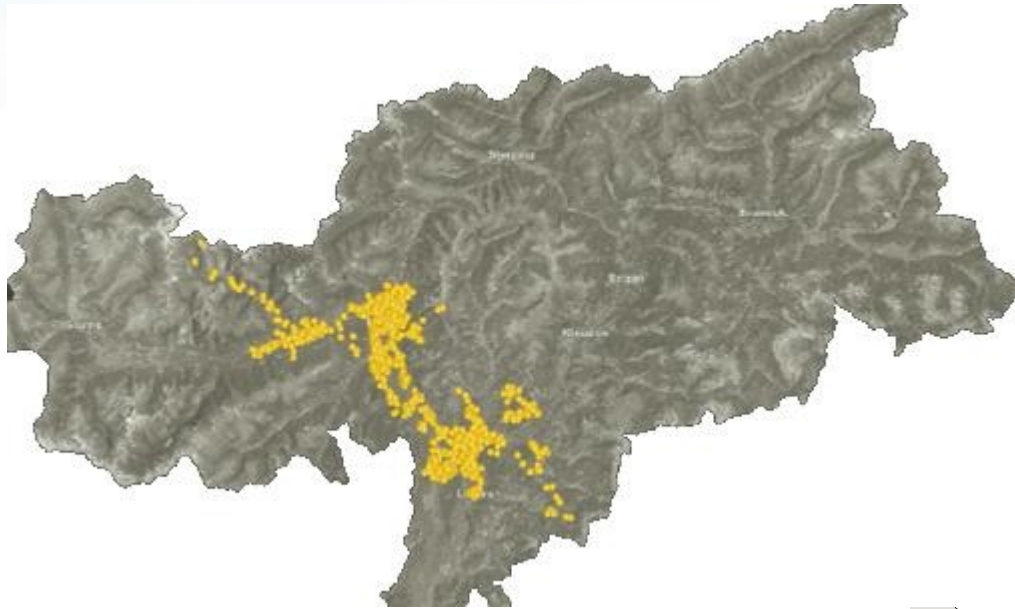


Summer 2015 - Alperia DSO starts collaborating with EURAC to develop a new upscaling method based on data-driven models for regional PV power estimation and mid-term forecast. IDEAM provides NWP model and satellite derived irradiance data.

Estimation allows real time PV distributed power supervision, net load prediction, transmission scheduling and energy trading.

PV Power forecasts improve the capability of residual load tracking and net load supervision. Prediction intervals can be used to reduce uncertainty in the electric demand so that lower energy reserves are needed

Regional upscaling and Spatial Clustering



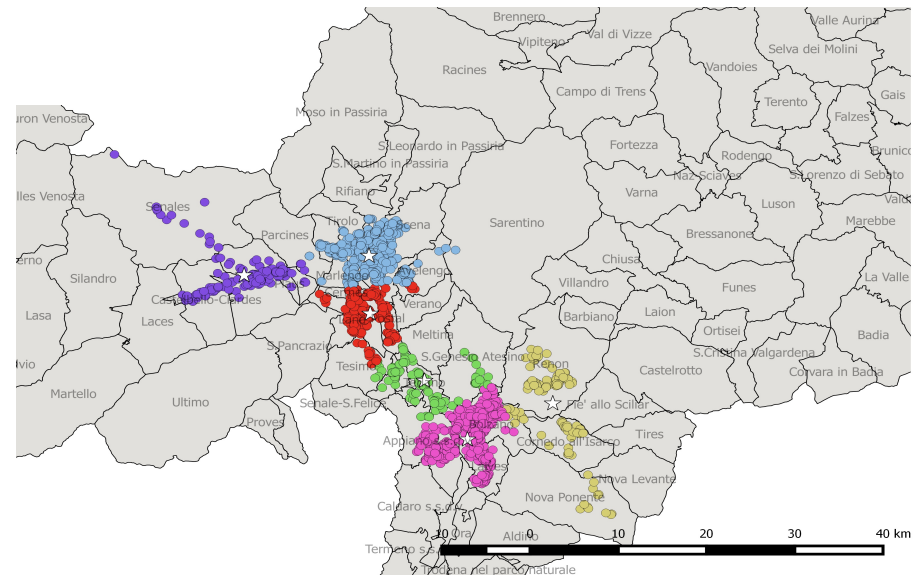
1975 plants distributed over complex terrain.

The K-mean algorithm is used to generate a statistical sample representative of the regional solar irradiance.

6 cluster centroids have been selected to represent the Region.

The advantage is twofold:

- economize computation avoiding single PV plant generation estimation and forecast
- perform spatial smoothing to reduce errors.



Satellite derived irradiance

Source: METEOSAT-9 (MSG3)

OSI SAF SSI algorithm

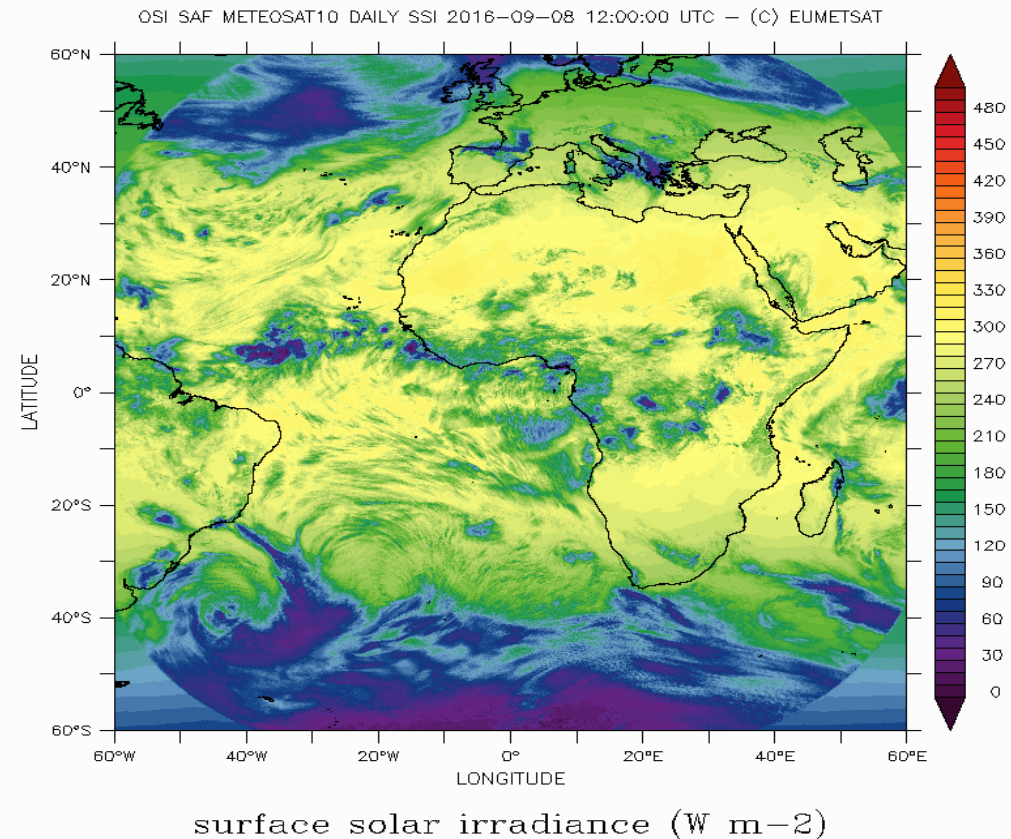
Centered at 0 deg longitude

Window: 60W to 60E, 60N to 60S

Horizontal resolution: 0.05

Temporal resolution= 1 hour

Available near real time (~ 2 hours after measurement)



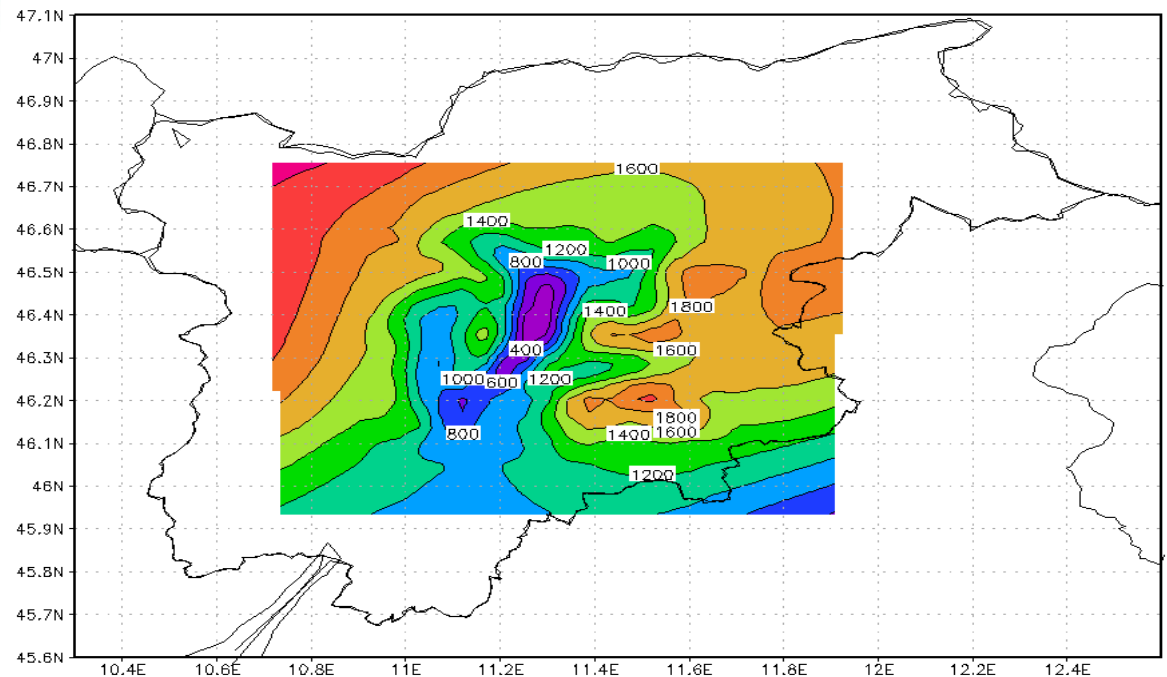
WRF

The model used is WRF-ARW
V3.8

Italian domain with 12km
horizontal resolution

nested domain: 3 km horizontal
resolution

Initial and contour data from
GFS 0.25 deg, 12 UTC
initialization



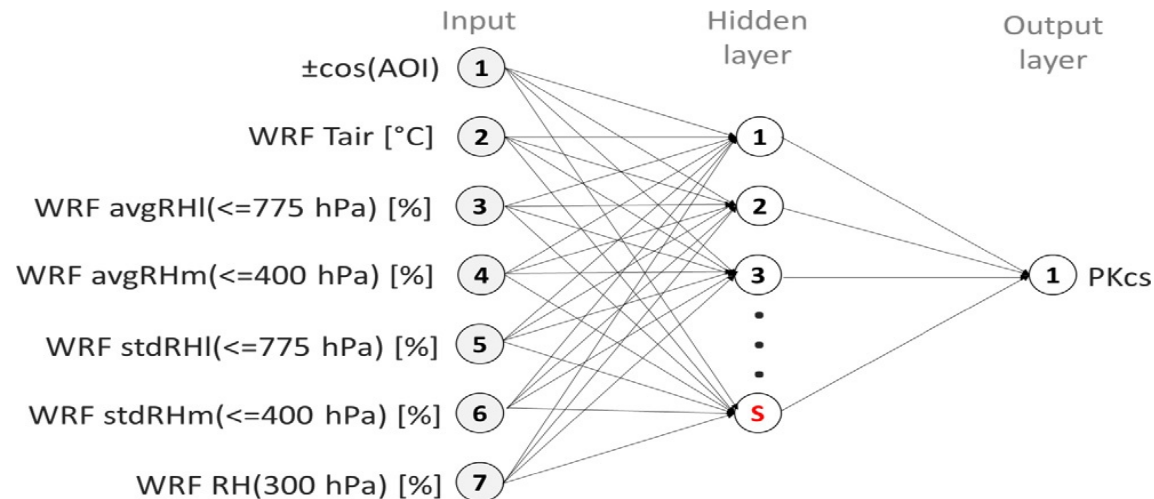
Short wave radiation Scheme: RRTM

- describes sub grid cloud variability through a Monte-Carlo independent column approximation
- distinguishes near-infrared, visible and UV wavelengths.

Concentration and chemical composition derive from parametric values taken from pre-built tables, NO WRF-CHEM is run

Ensemble of ANN: RHNN

RHNN provides the day-ahead forecast of the clear sky performance index (*Pierro et al. 2016. Multi-Model Ensemble for day ahead prediction of photovoltaic power generation, Solar Energy 134 132-146*)



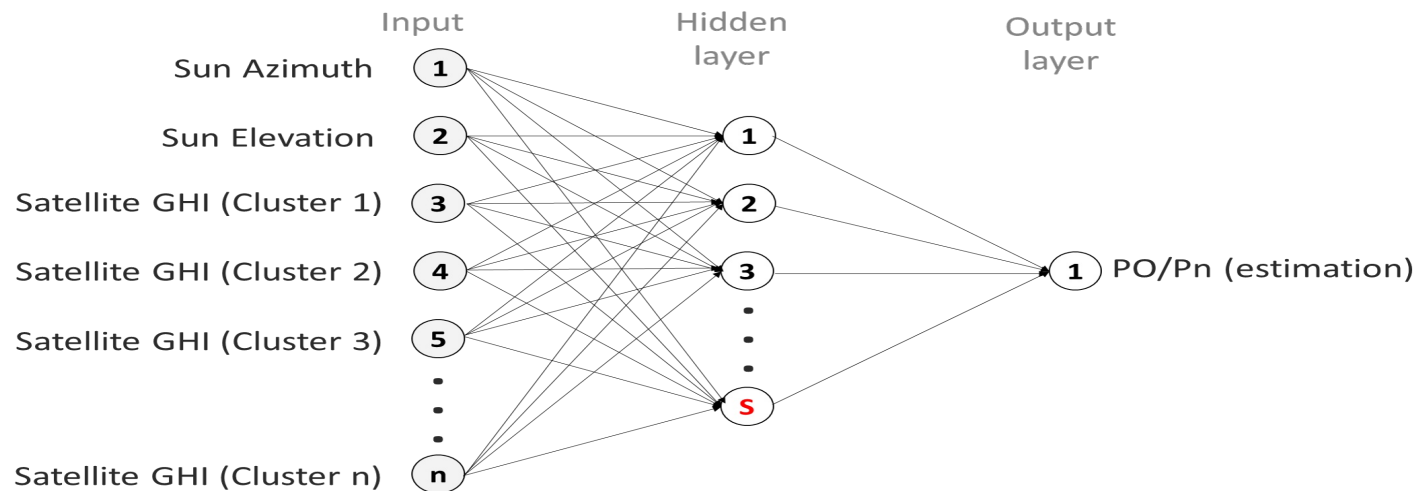
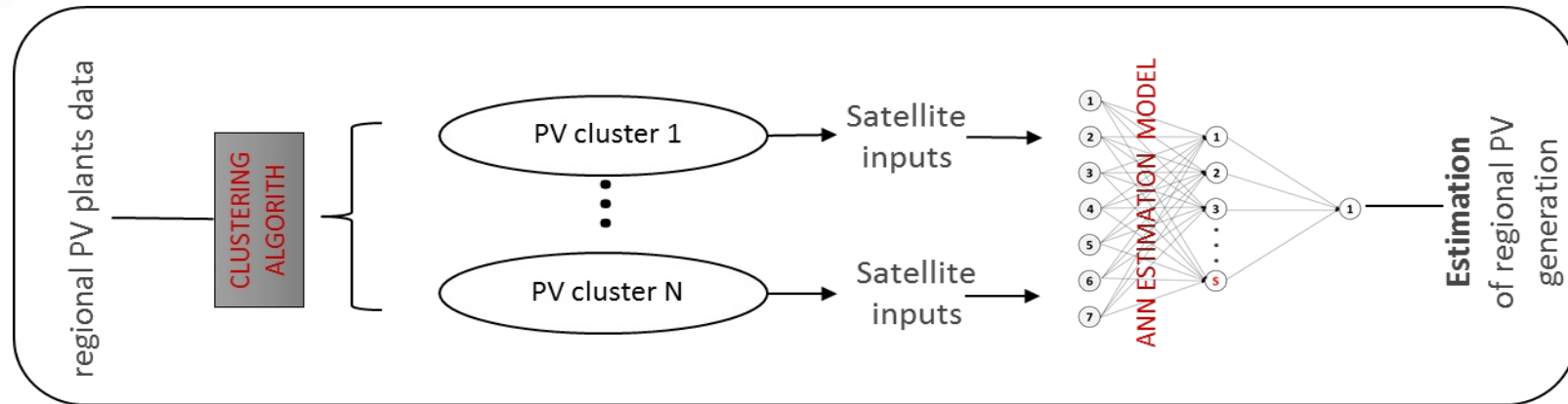
ANN architecture: MLPNN with two layers is adopted. 500 ANNs are generated. Then a qualified ensemble is selected (around 300 ANNs), choosing all the ANN with the MSE lower than the average MSE of the 500 networks. The forecast is obtained by the average on the ensemble outputs

After numerous case studies (*Pierro et al. 2016, Deterministic and stochastic approaches for day-ahead solar power forecasting*) as input for the ANN the WRF 2 m. air temperature, a combination of relative humidity levels, and optionally the **clear sky** model global horizontal irradiance (GHI) are used.

No forecast of GHI reaching the surface is needed.

Estimation of regional PV power

Very few input informations are needed so it can be easily adopted by DSO's.



The power estimation model achieves, using satellite derived irradiance, an RSME of 3% of the installed capacity and a MAE of 2%.

GHI ground measurements lead to an RMSE of 2% and a MAE of 1.5%

RHNN applied to regional day ahead forecast

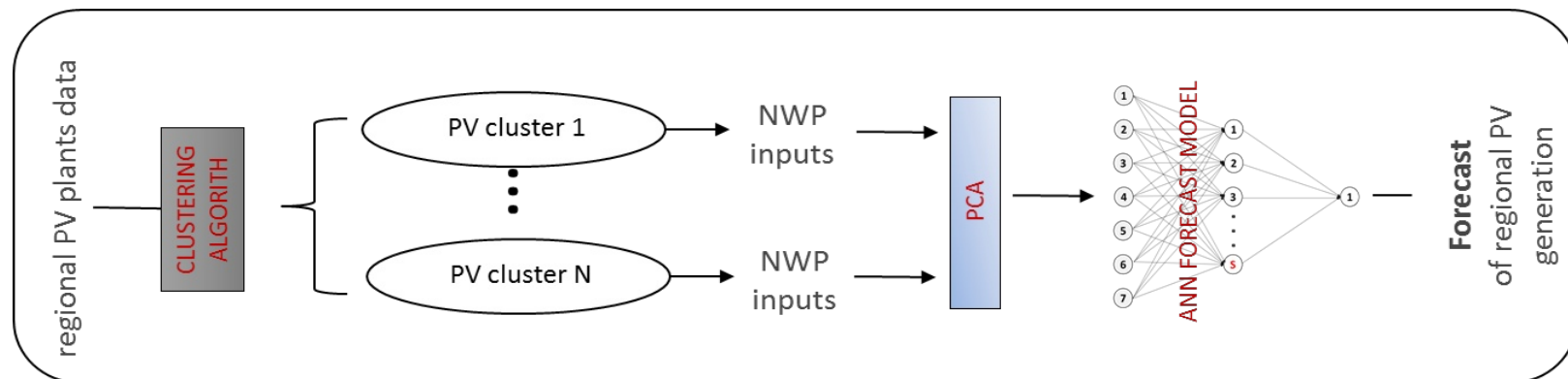
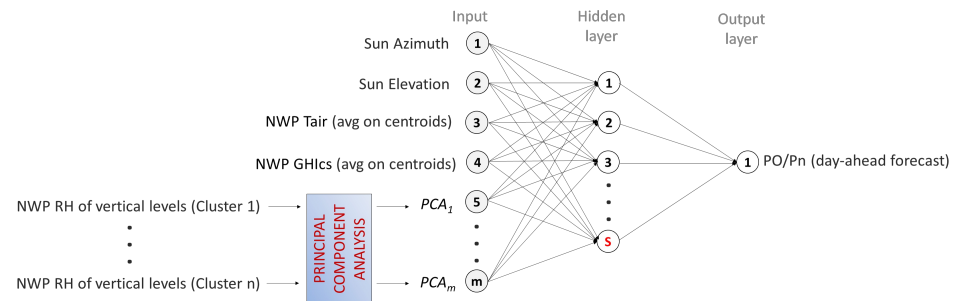
Several approaches were tested.

The most accurate forecasts are obtained by applying a principal component analysis (PCA) to the WRF forecast of relative humidity and subsequently running an ANN, from now on called PCARHNN.

Input for ANN:

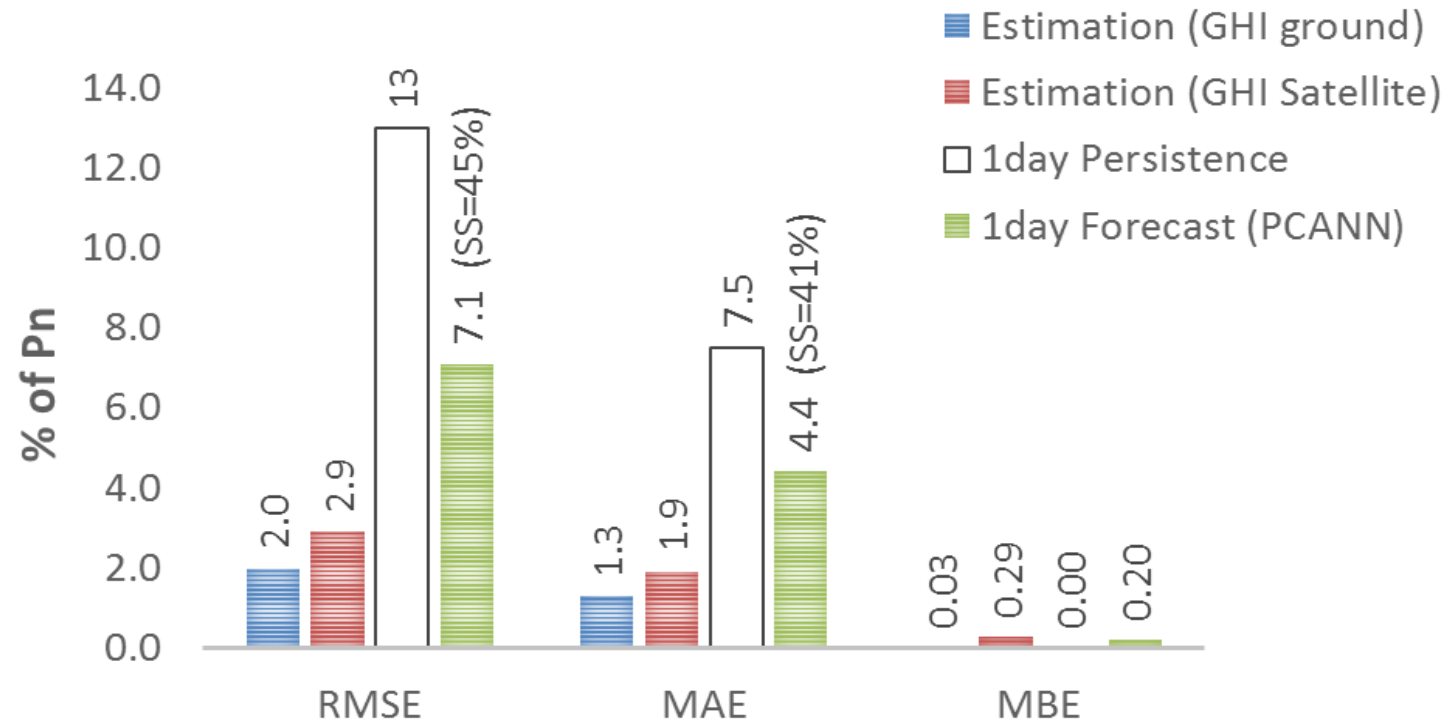
1. SunEL [°](Theoretical Sun Elevation)
2. SunAZ [°](Theoretical Sun Azimuth)
3. NWP Tair [C]
4. NWP GHlcs [W/m2]
- 5-9. $[Z_1, \dots, Z_5] = \text{PCA}(76 \text{ RH level}) \dots$ (19 RH levels for each of the 6 clusters)

PCA reduces input dimensionality of the RH levels from 76 to 5



PCARHNN is also the most straightforward and computationally less expensive of the methodologies tested

Results

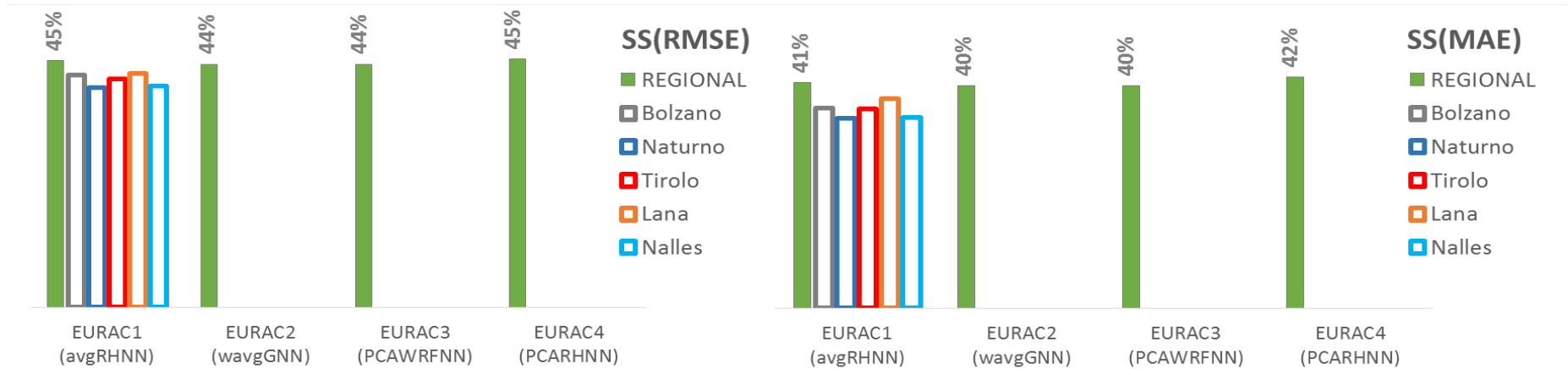


2014 observational data was used for training the NN. 2015 for test and calculation of the accuracy metrics.

Errors are normalized in respect to the installed capacity.

In the figure the results for a simple persistence model are presented.

Results 1 day ahead forecast: smoothing



The skill score is used because it is not site and year dependant. The RMSE Skill Score is defined as follows (MAE SS is similar):

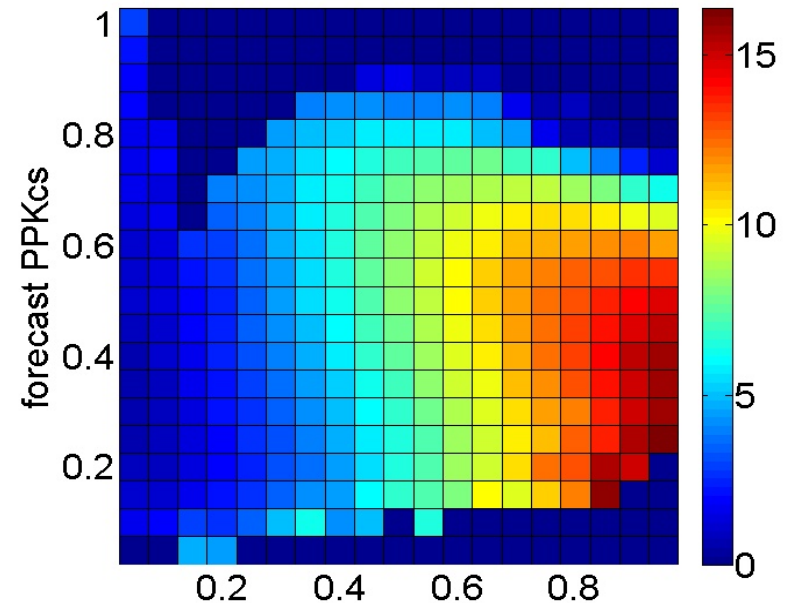
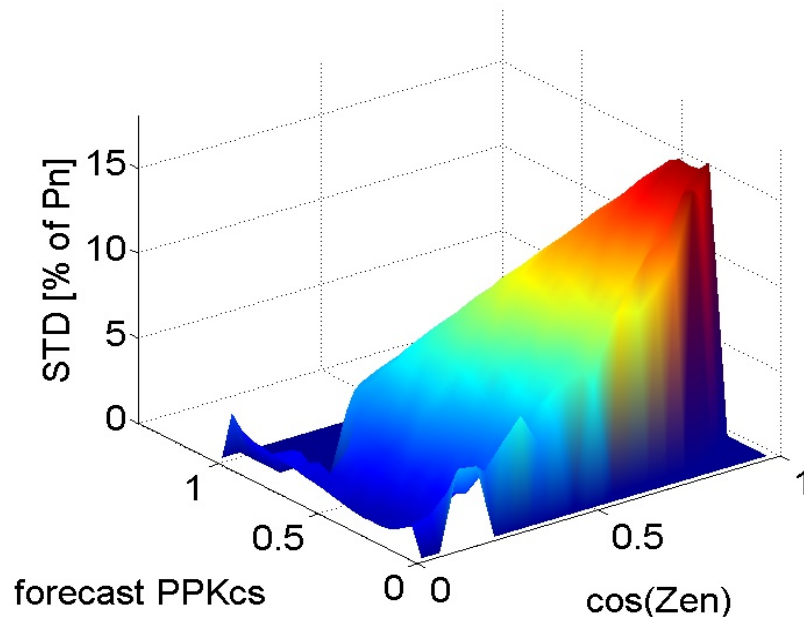
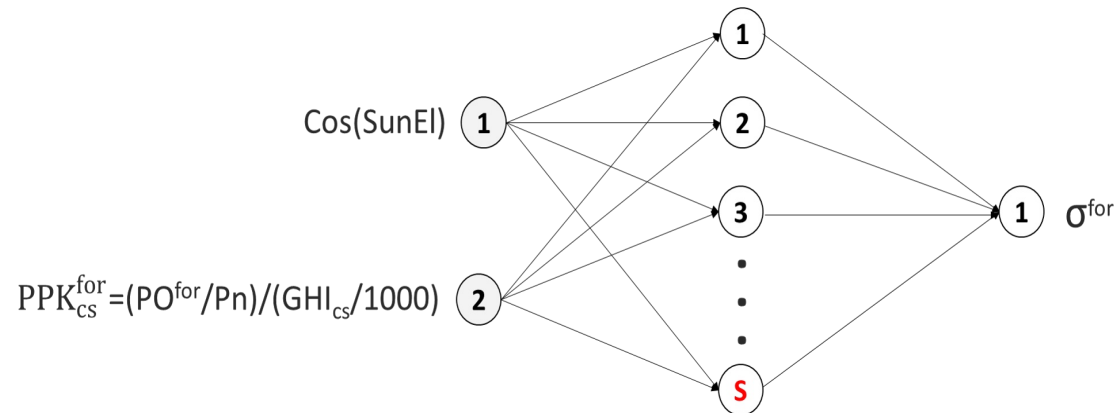
$$SS = 100 \left(\frac{RMSE_{pers} - RMSE_{for}}{RMSE_{pers}} \right)$$

Where $RMSE_{pers}$ is the RMSE for the simple persistence model.

Better results are obtained compared to single site and single cluster centroids forecast

Prediction Intervals

Prediction intervals can be calculated forecasting the standard deviation of the residuals (σ for) under the hypothesis that the residuals are normally distributed with zero expected value



Uncertainty: prediction intervals

Prediction Interval Evaluation:

Pobs= 95% at Confidence Level=95%

Pobs= 81.7% at Confidence Level=75%

Pobs= 59.3% at Confidence Level=50%

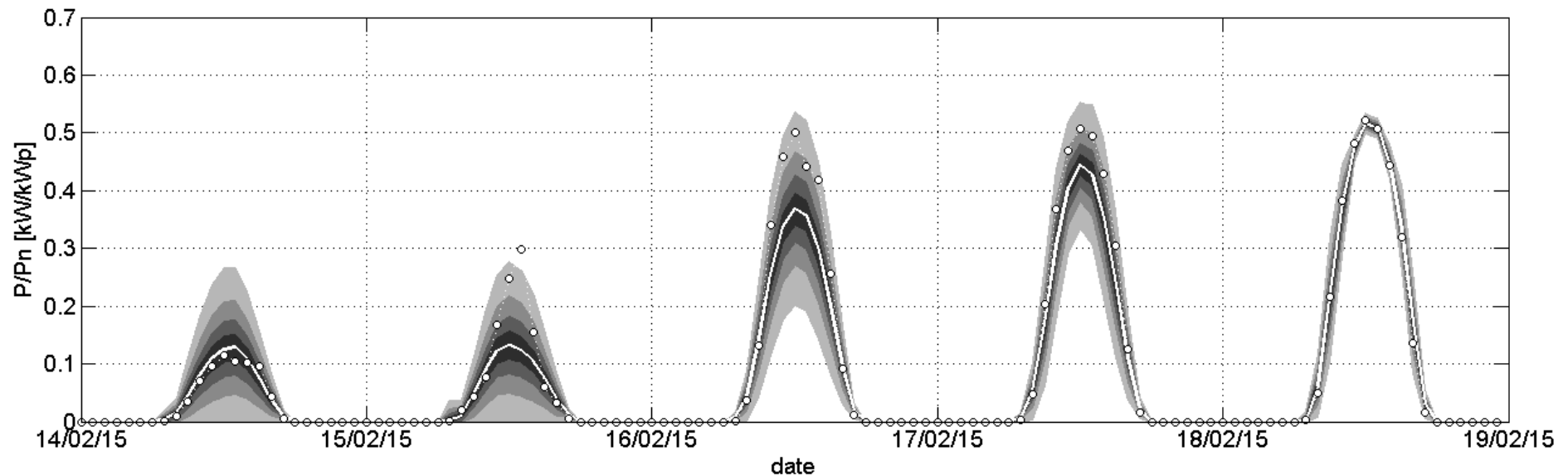
Pobs= 31.6% at Confidence Level=25%

Prediction intervals width evaluation

STDfor/<Xobs>=26.15% with respect to STDobs/<Xobs>= =90.45%

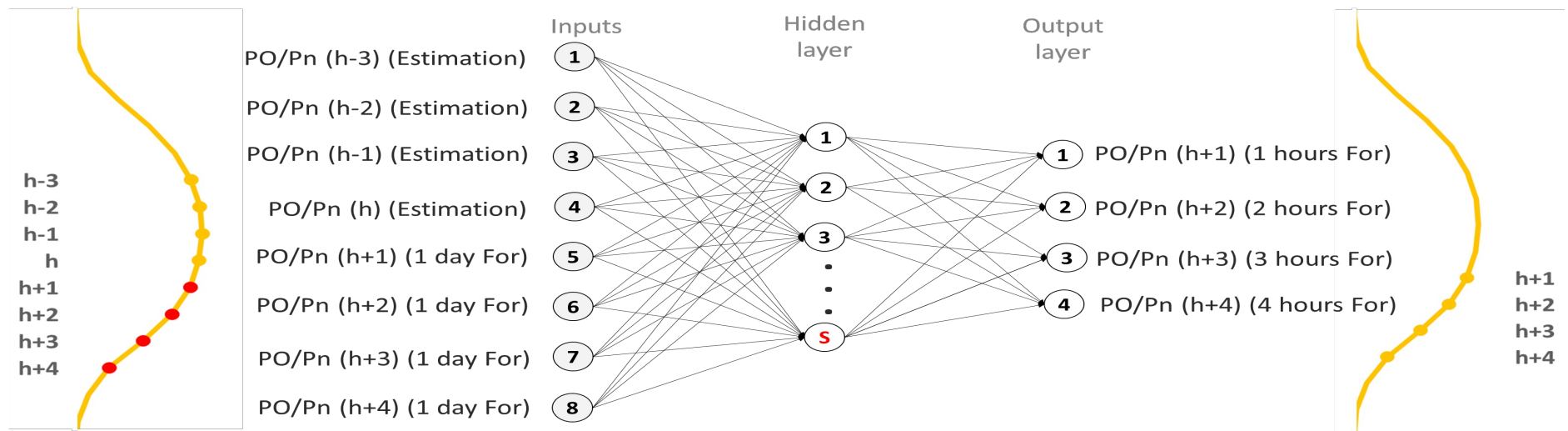
STDfor/STDobs=28.9%

With the transition from overcast to clear sky days the width of prediction intervals decreases



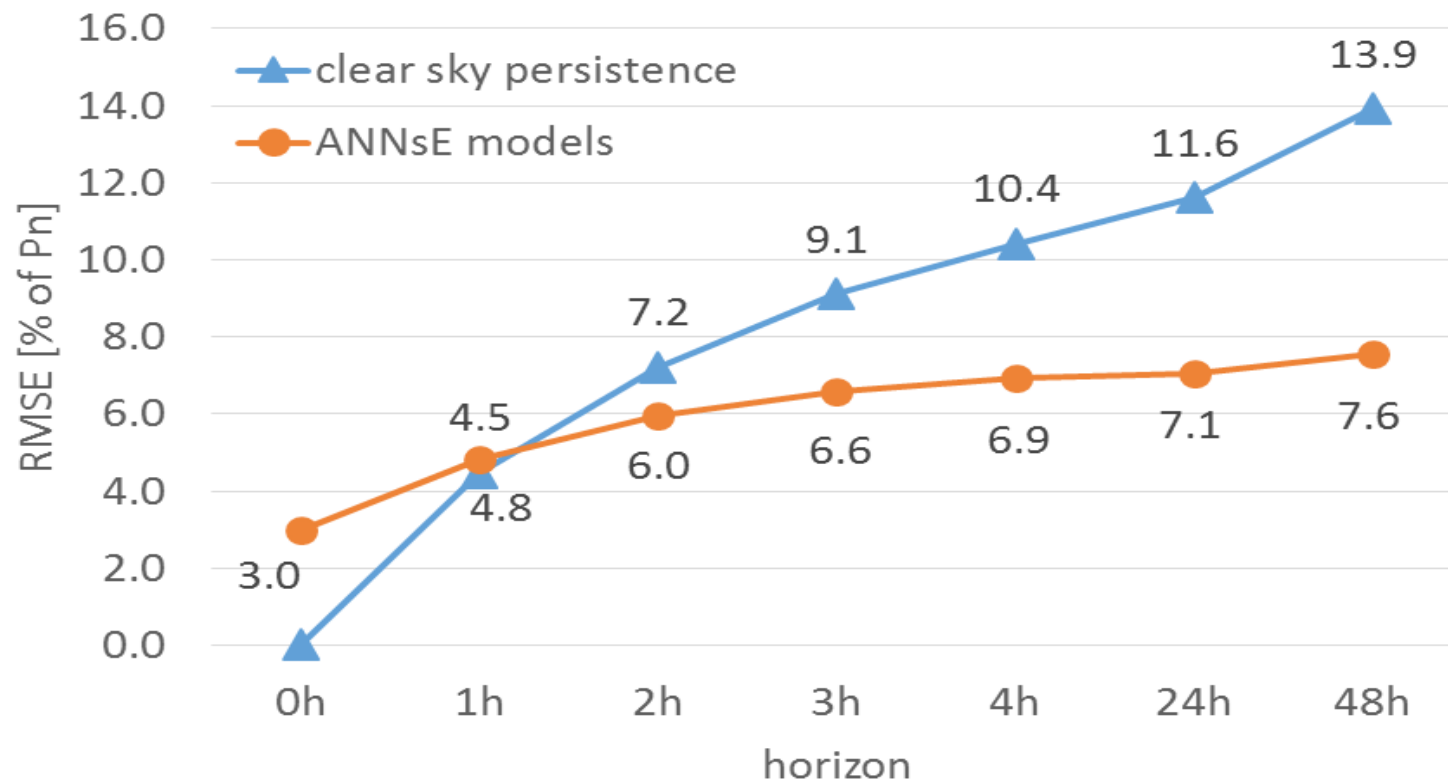
Intraday Forecast

The intra-day forecast model makes use of past power estimation and 1 day forecast obtained by past satellite irradiance and 1 day ahead NWP data.



Estimation of the past 4 hours and the 1 day ahead forecast (PCARHNN) for the following 4 hours are used as input for the specific ANN

Results



The clear sky persistence model performs better than simple persistence, it is derived from the clear sky index.

Satellite data can be used to correct the forecast based on NWP up to a 4 hour horizon.

For longer horizons the intra-day forecast has a lower accuracy compared to the 1 day forecast (PCARHNN).

Conclusions

The power estimation and forecast on a regional scale provide more accuracy than the estimation and forecast of a single site or of each cluster. An ensemble smoothing is obtained by the upscaling performed

Let's review the results:

- the accuracy of the power estimation model achieved using satellite derived irradiance is around 3% of installed capacity thus the upscaling method could be adopted for real time power supervision.
- The intra-day forecast obtains an RMSE between 5% and 7% of installed capacity.
- The day-ahead forecast achieves an RMSE between 7% and 8%.

These values can be considered “state of the art” for a regional site and forecast.

In October 2016 forecast and estimation will become operational

Thank you
for your attention