

# Spatio-temporal models for photovoltaic power short-term forecasting

**X. Ghislain Agoua**, Robin Girard, Georges Kariniotakis  
MINES ParisTech,  
Centre PERSEE, Renewable Energies & Smartgrids Group



# Outline

- Introduction – State of the art
- The proposed models
- Case study
- Evaluation results
- Conclusions & perspectives

# Introduction

## Why forecast PV power production ?

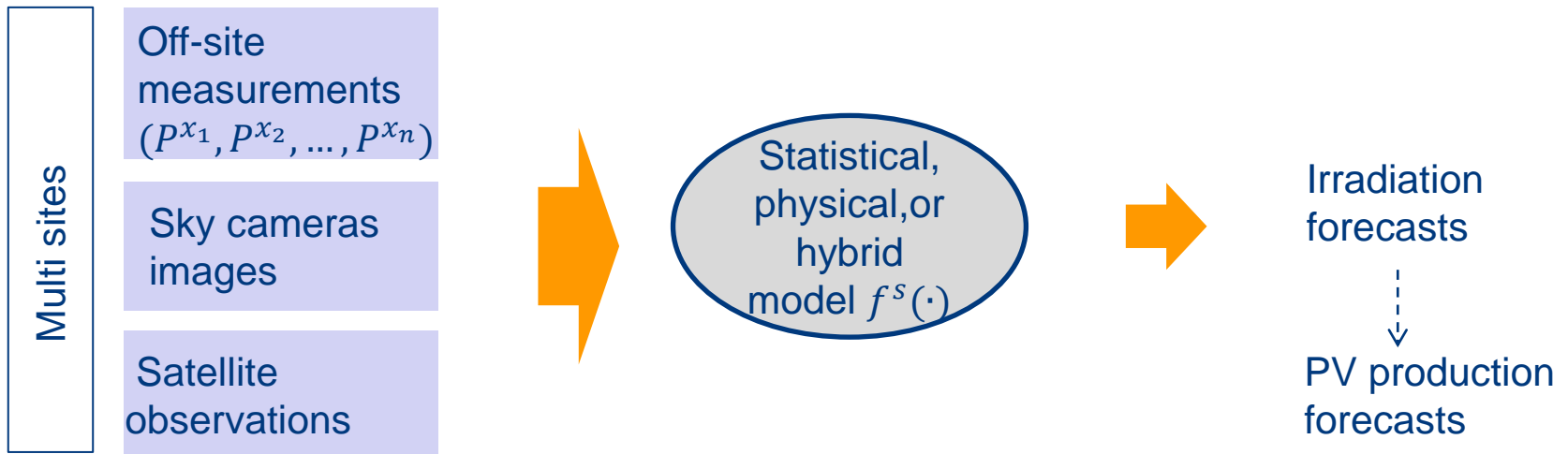
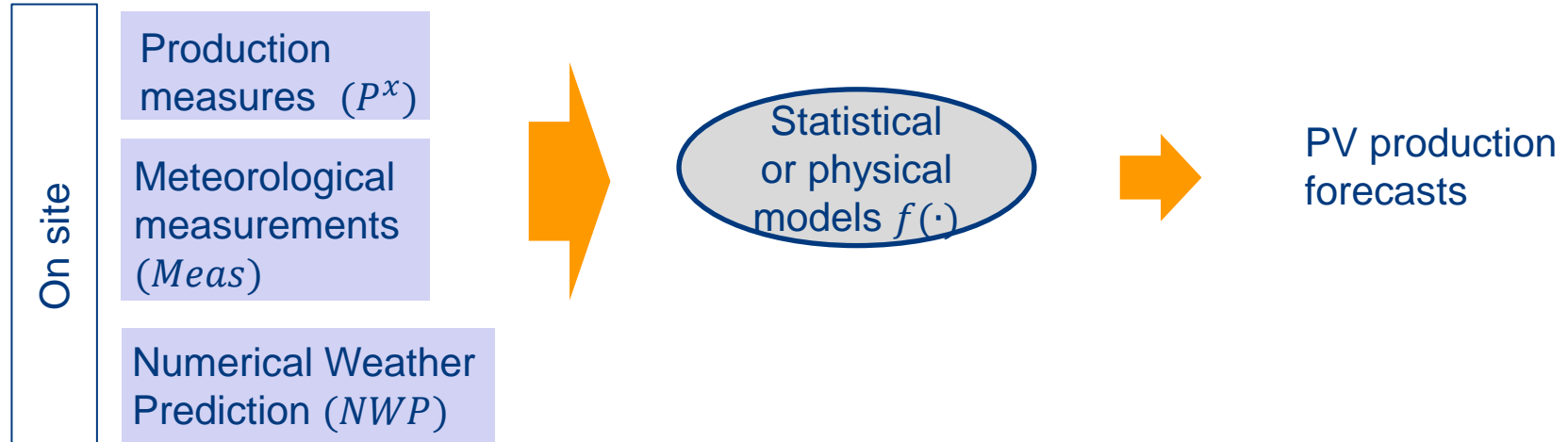
- Management of power system with high PV penetration
  - Network security
  - Balance between production and consumption
  - Reserve management
- Optimise maintenance scheduling of PV plants
- Trading PV generation in the electricity markets
- ...

# Overview of state of the art (1/2)

- “Physical” approach:
  - Based on analytical equations with weather variables and PV system characteristics as inputs
- “Statistical” Approach:
  - Linear regression, regression trees
  - Box and Jenkins methods (ARMA, ARIMA, SARIMA...)
  - Boosting, bagging, random forests
  - SVM, non and semi-parametric models
  - Neural Networks
  - Bayesian inference

# Overview of state of the art (2/2)

$$\hat{P}_{t+h|t}^x = f(P_t^x, \dots, P_{t-l}^x, NWP_{t+h|t}, Meas_t)$$



- Propose a short-term forecasting model that exploits spatio-temporal dependencies between distributed PV plants to provide improved forecasts for a specific PV plant
- Compare the spatio-temporal model with a reference approach that does not use off-site information

# The considered reference models

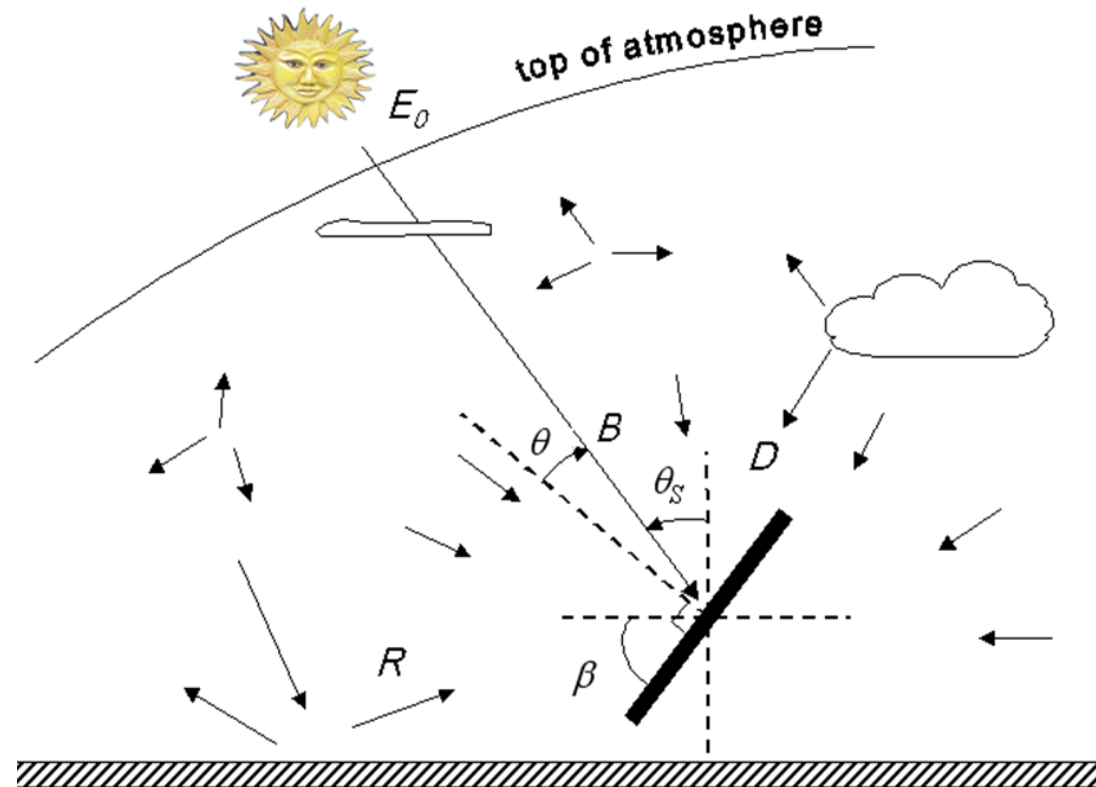
- 2 options considered: Random forests and AR models.
- The AR model was found to outperform and is selected as **reference model** for the power production forecast of each plant  $x$ .

$$\hat{P}_{t+h|t}^x = \hat{a}_h^0 + \sum_{l=0}^L \hat{a}_h^l P_{t-l}^x$$

- Necessity to stationarize the input variables

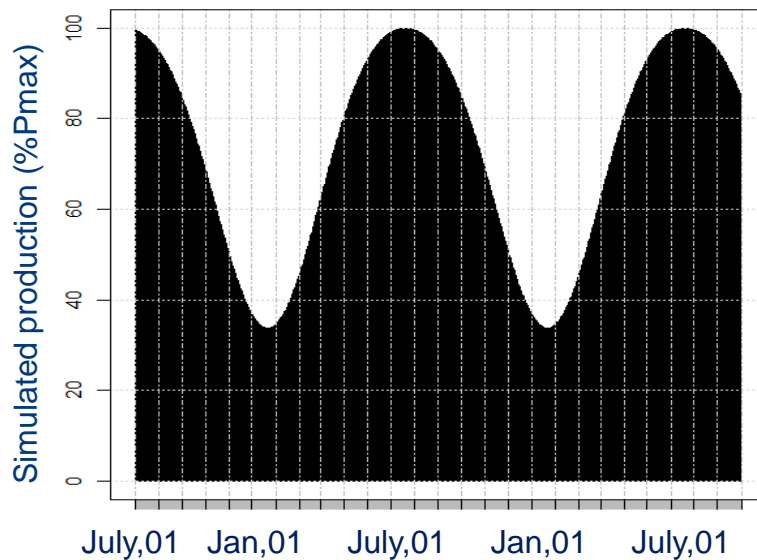
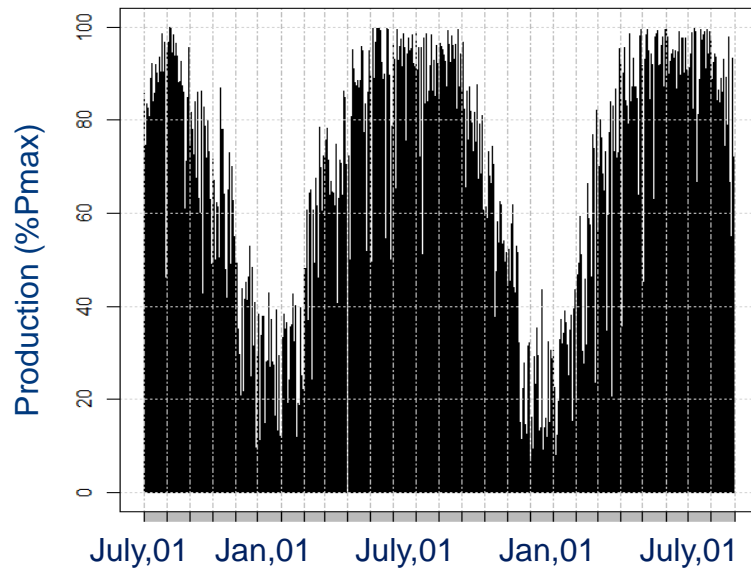
# The proposed stationarization technique

- We simulate the irradiation at the top atmosphere level, with inputs:
  - Longitude and latitude
  - Inclination ( $\beta$ ) and orientation ( $\alpha$ )
  - Solar angle ( $\theta$ )
- Compute the corresponding production ( $TOA_t$ )

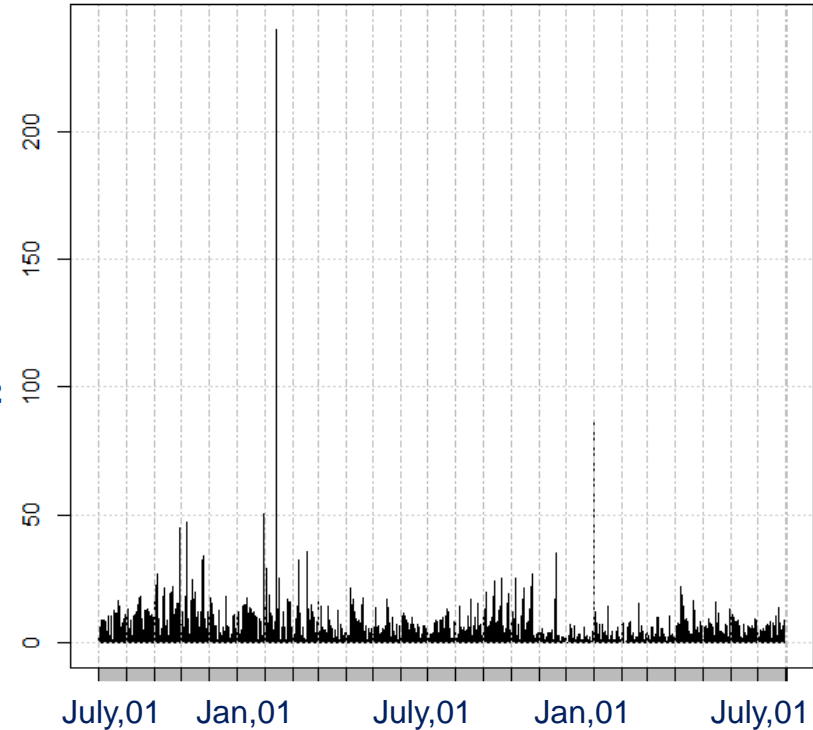




# The proposed stationarization technique



$$\frac{P_t}{TOA_t} = S_t^2$$



Post validation with Dickey-Fuller Test

# The proposed spatio-temporal model

- A linear auto-regressive model chosen

$$P_t^x = a^0 + \sum_{l=0}^{L_s} \sum_{y \in X} a^{l,y} P_{t-l}^y$$

$X$ , the set of all the  $y$  power plants

$L_s$ , the appropriate maximum lag

- Input: power production of all neighbour PV plants as explanatory variables
- The appropriate maximum lag is estimated by AIC criterion
- The model coefficients are estimated by least mean squares regression
- A model is tuned for each prediction horizon

# The evaluation criteria

- Root Mean Square Error  $RMSE_h = \sqrt{\frac{1}{n} \sum_{t=1}^n e_{t+h|t}^2}$

- Mean Absolute Error  $MAE_h = \frac{1}{n} |e_{t+h|t}|$

- Bias  $BIAS_h = \frac{1}{n} \sum_{t=1}^n e_{t+h|t}$

where Error :  $e_{t+h|t} = P^x_{t+h|t} - \hat{P}^x_{t+h|t}$

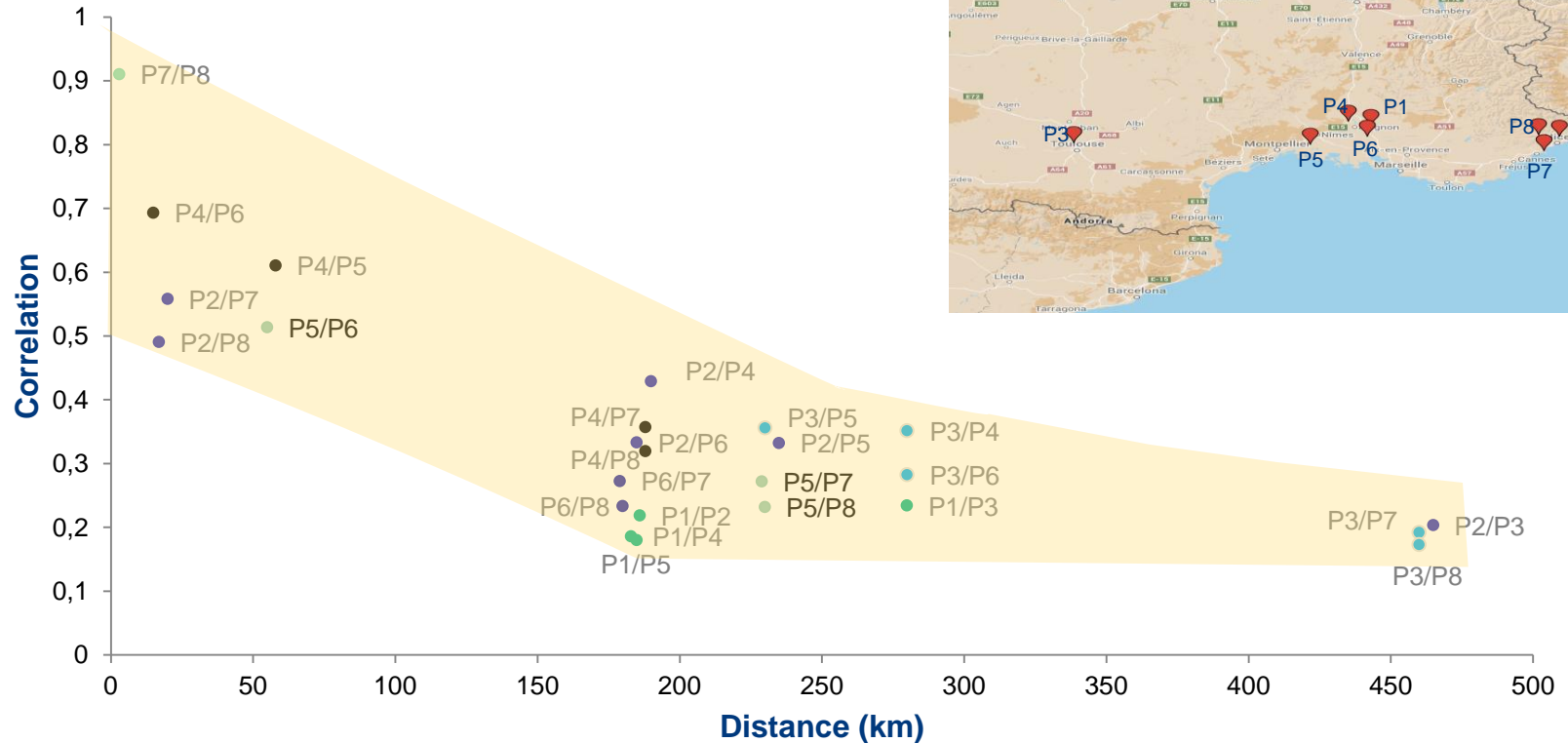
# The case study

- 8 PV power plants in South France
- Peak power ranging from 45 kW to 5 MW
- 20 months measurements starting 7/2013
- 15 min time resolution
- 6 hours horizon with 15 min update resolution



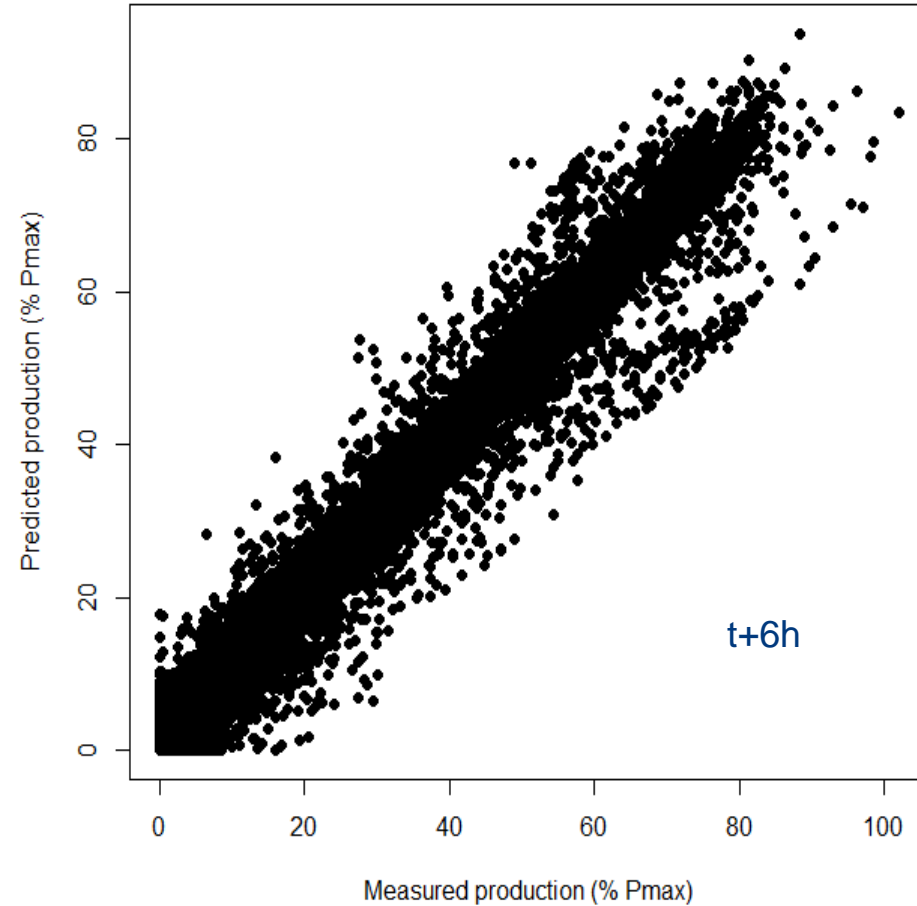
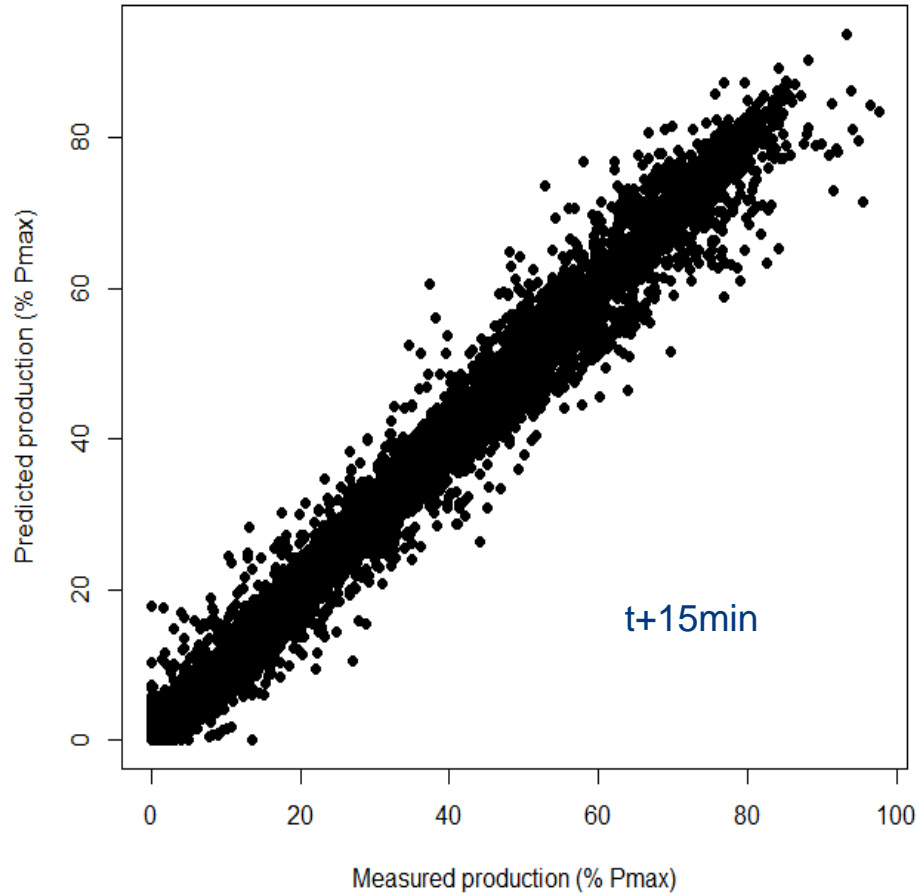
# Correlation of off-site measurements

- The correlation between PV plants with series  $S_t^2$ 
  - The cross-correlation decreases with the distance
  - The cross-correlation values indicate that information from multiple sites can be exploited for prediction



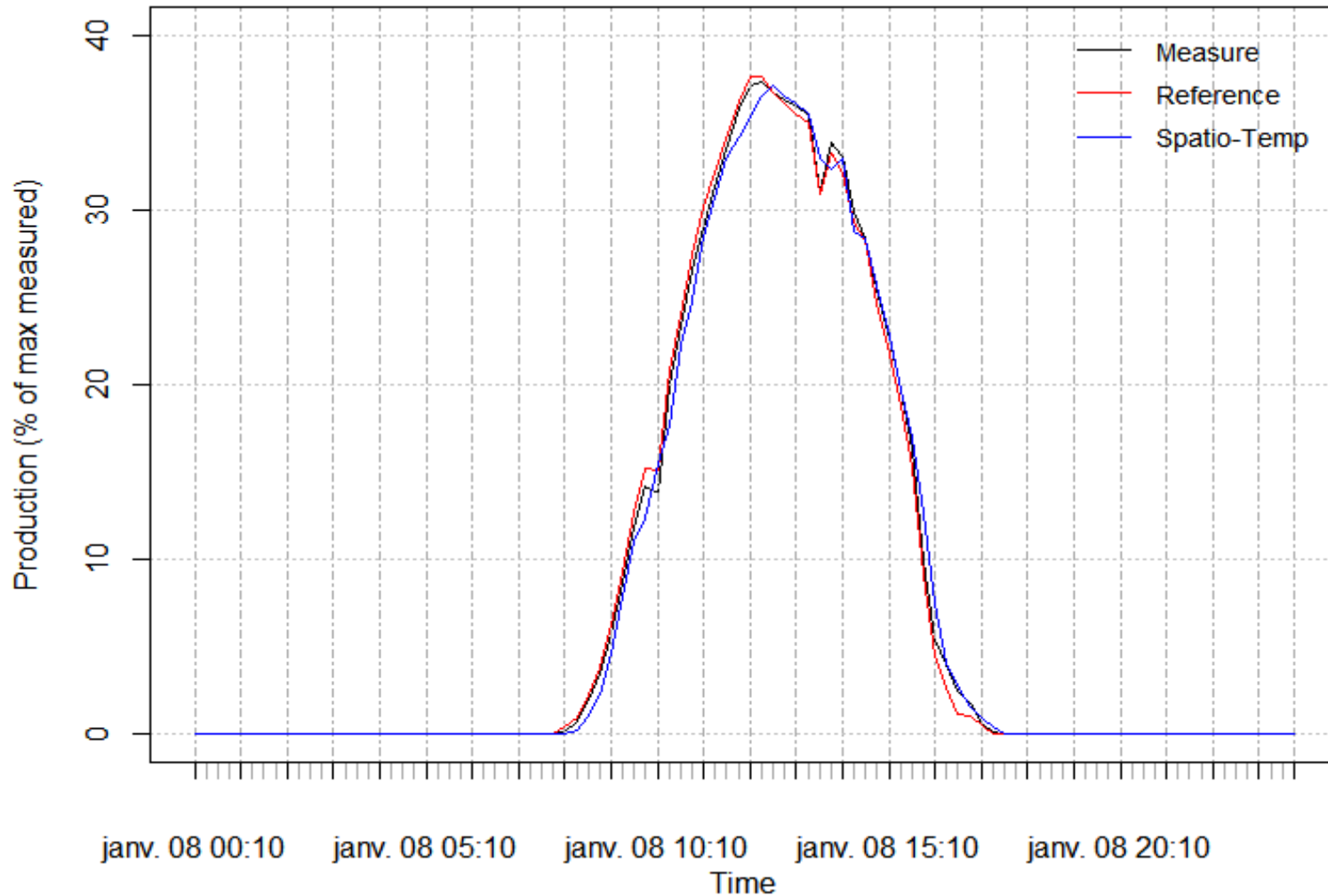
# Evaluation results (1/4)

## Visualisation of forecasts accuracy



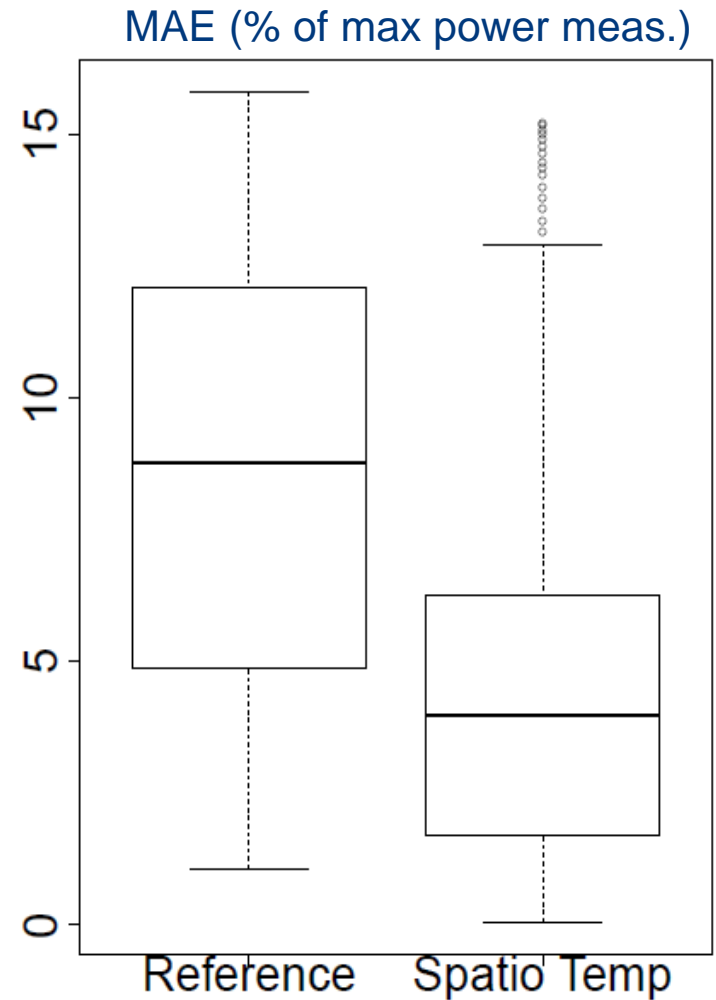
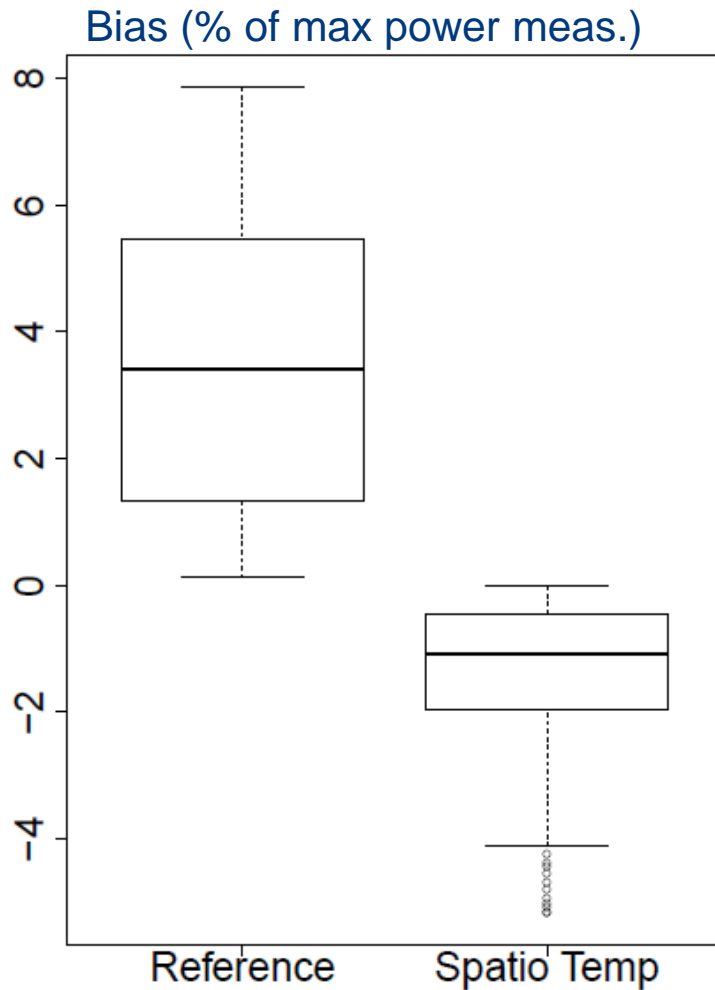
# Evaluation results (2/4)

## Forecasts comparison



# Evaluation results (3/4)

## Bias and MAE performance evaluation (for all time horizons)

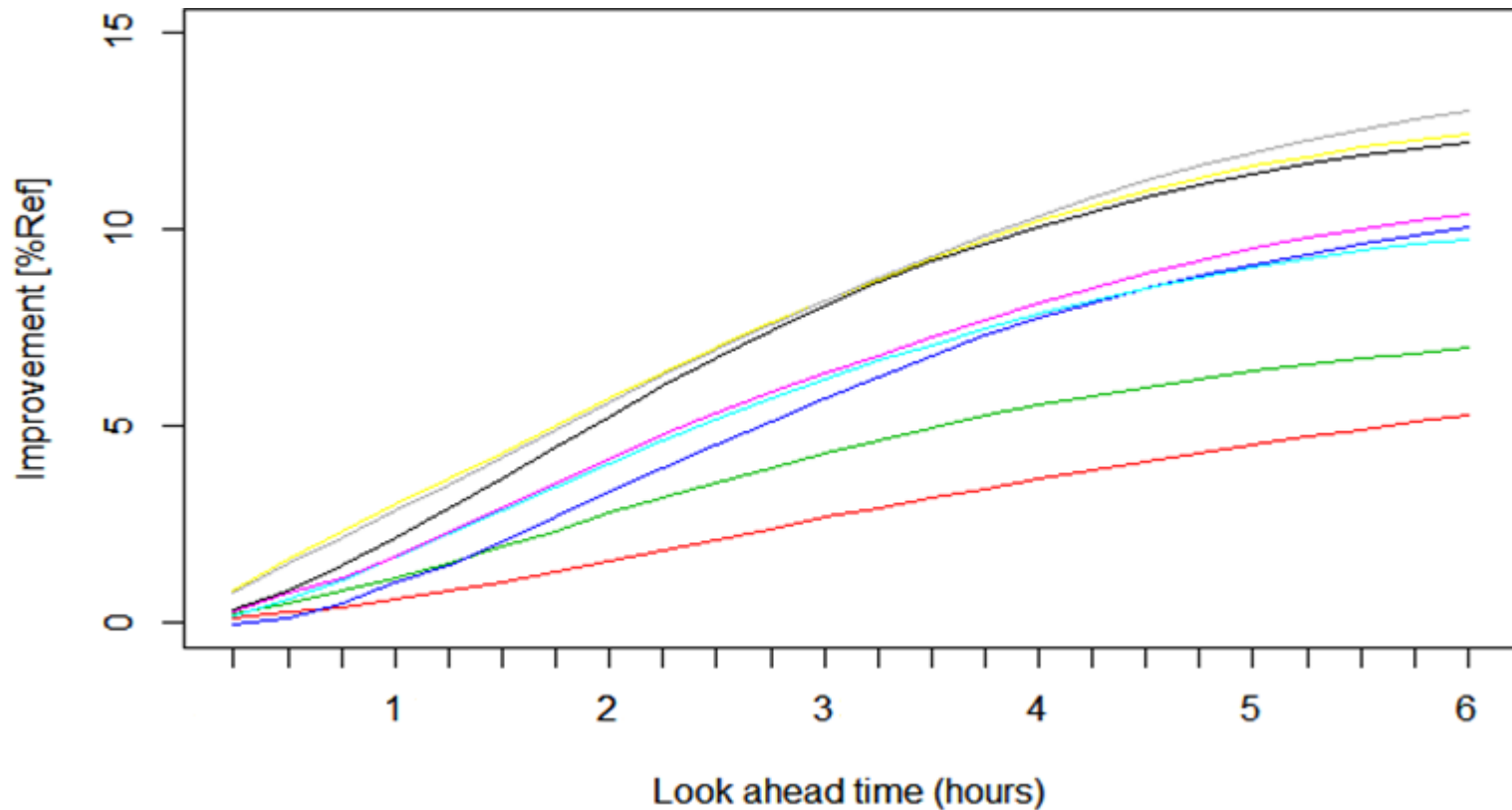


~55%



# Evaluation results (4/4)

- Improvement in terms of RMSE ( $Impr = \frac{RMSE_{ref} - RMSE_{spat}}{RMSE_{ref}} * 100\%$ )
- One line per power plant



# Conclusions and perspectives

- Up to 10% improvement in terms of RMSE
- Significant reduction of Bias and MAE (55%)
- Very short computation time
- Low cost: use of existing data
- No need for weather forecasts (NWP)

## Potential improvements and perspectives:

- Polynomial or splines functions
- Forecasting model parameters as a function of meteorological conditions
- Satellite images or sky cameras
- Probabilistic spatio-temporal forecasting

Thank you for your attention



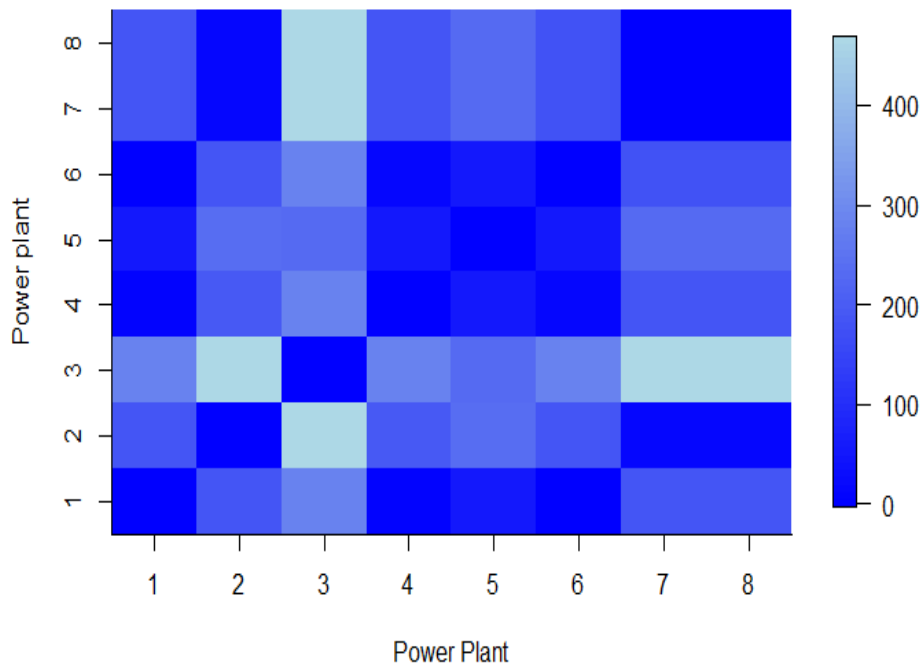
Acknowledgment :



# Correlation of off-site measurements

- The correlation between plants with series  $S_t^2$ 
  - The crossed correlation decreases with the distance
  - The crossed correlation values indicates presence of relations between the multi-sites production

Distance (km)



Cross correlation

