Uncertainty Quantification for long-term Wind Farm Production: A Monte-Carlo Study

Merlin Keller[1], Sanaa Zannane[1]
Nicolas Paul[1]
Cédric Dall'Ozzo[2], Lan Ding[2]

EDF R&D[1] - PRISME, EDF Re[2]

Journes MAS 2022, Rouen

Summary

Section 1

Introduction

Industrial Context

In the context of wind farm development, one of the central quantities of interest is the:

Levelized Cost of Energy

Definition: Average revenue per unit of electricity generated required to recover the costs of building and operating a generating plant during an assumed financial life and duty cycle.

$$\text{LCOE} := \frac{\text{Sum of costs over lifetime}}{\text{Sum of electrical energy produced over lifetime}}$$

Source: `https://en.wikipedia.org/wiki/Levelized_cost_of_electricity`

Here, we focus on quantifying the denominator, or equivalently the Expected Annual Production (EAP), averaged over the lifetime of the windfarm project.

Input data for the study

$$V^{SAT}_{\mathcal{T}} = (v^{SAT}_t)_{t\in\mathcal{T}}$$
$$D^{SAT}_{\mathcal{T}} = (d^{SAT}_t)_{t\in\mathcal{T}}$$

wind-speed & direction satellite reconstructions on long-term period $\mathcal{T}$

$$V^{SIT}_{\mathcal{T}'} = (v^{SIT}_{t,i})_{t\in\mathcal{T}'}$$

wind-speed onsite measures at heights $z_i, i = 0, \ldots, n$ on short-term period $\mathcal{T}'$

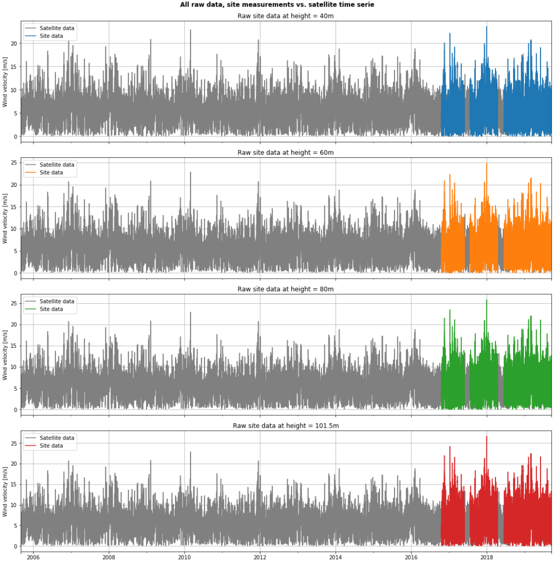


Figure: On-site data for multiple heights, with ERA satellite proxy

Input data, zoom

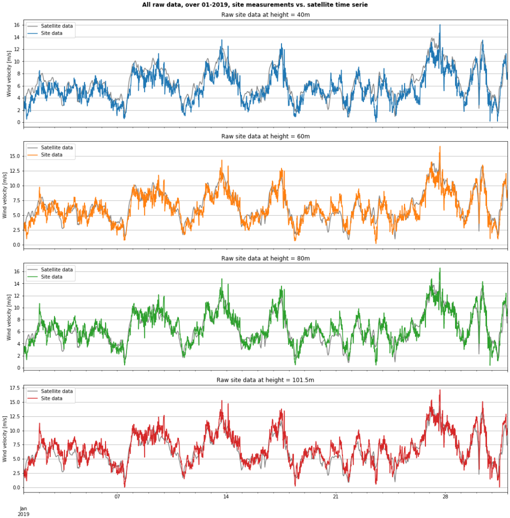


Figure: On-site data for multiple heights, with ERA satellite proxy, zoom on January 2019

Power curve

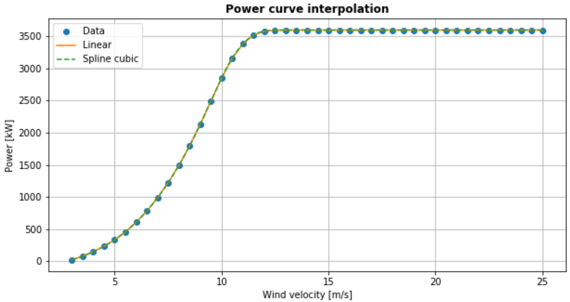


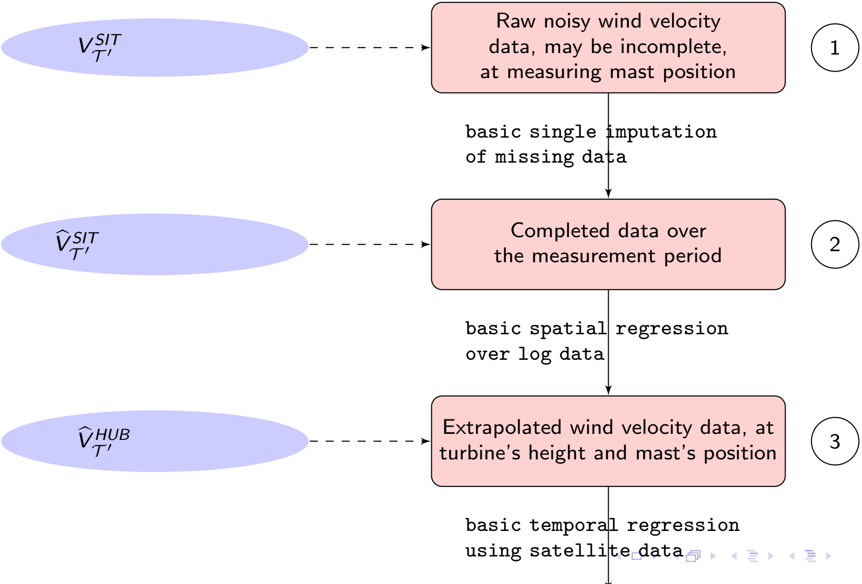
Figure: Reconstructed power function, by means of standard interpolation methods

- Yields instantaneous turbine power, given incoming wind speed:

$$V_{\mathcal{T}}^{TURB} = (v_t^{TURB})_{t\in\mathcal{T}}$$

- Main challenge: extrapolate $V_{\mathcal{T}}^{TURB}$ using available data $V_{\mathcal{T}}^{SAT}, D_{\mathcal{T}}^{SAT}, V_{\mathcal{T}'}^{SIT}$

EDF-R extrapolation procedure (1/2) : statistical modeling



EDF-R extrapolation procedure (2/2) : physical modeling

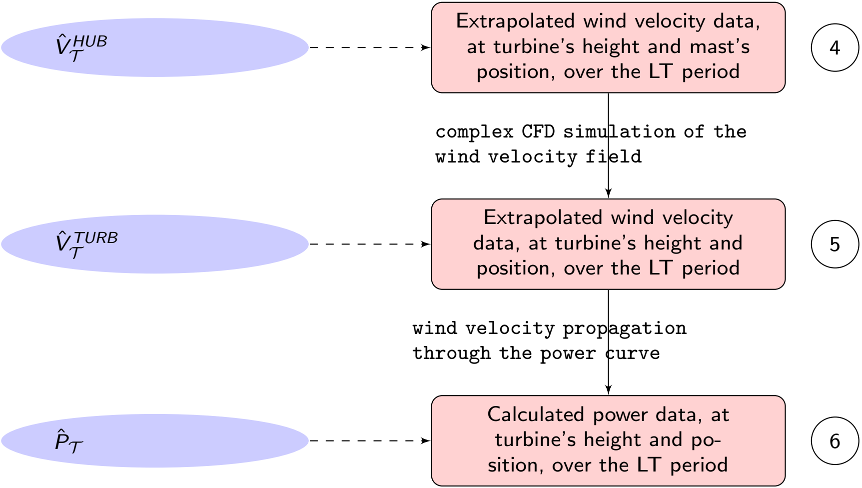


Figure: EDF-R extrapolation chain

Formalization of the industrial problem

We are interested in estimating the EAP (expected annual production), for a given project duration of N_y years :

$$EAP = \frac{1}{N_y} \sum_{t \in \mathcal{T}} \widehat{p}_t \quad (1)$$

Each of the above-described steps adds different sources of uncertainty to the power forecast, in particular:

- Mast measurement errors;
- Vertical extrapolation statistical uncertainty and modeling error;
- Long term-reconstruction extrapolation statistical uncertainty and modeling error;
- Horizontal extrapolation modeling error;
- Power curve modeling error.
- The objective is to estimate the uncertainty surrounding EAP.

Section 2

Proposed approach

Monte-Carlo uncertainty quantification

Current EAP quantification methodology

1. Determine parameter estimates $\widehat{\theta}$ from available data $\mathcal{D} := V_T^{SAT}, V_{T'}^{SIT}$, and deduce long-term hub-height wind-speed prediction $\mathcal{L}(V_T^{HUB}|\widehat{\theta}, \mathcal{D})$
2. Predict EAP from deterministic physical model $\mathcal{F}$ applied to V_T^{HUB} $EAP = \mathcal{F}(V_T^{HUB})$

Proposed parametric bootstrap scheme

Repeat for $n = 1, \ldots, B$:

1. Simulate synthetic dataset $\mathcal{D}_b^* \stackrel{\mathcal{L}}{=} V_T^{SAT}, \widehat{V}_{T'}^{SIT}|\widehat{\theta}$ conditional on the estimated parameters $\widehat{\theta}$
2. Recompute parameter estimate $\widehat{\theta}_b^*$, from synthetic dataset $\mathcal{D}_b^*$
3. Simulate $V_{T,b}^{HUB*}$ conditional on $\widehat{\theta}_b^*$, and deduce EAP: $\widehat{E}AP_b^* = \mathcal{F}(\widehat{V}_{T,b}^{HUB*})$

From bootstrap sample $\left(\widehat{\mathbb{E}}[EAP^*]_b\right)_{b=1,\ldots,B}$, derive bias estimates, confidence intervals, etc.

- In fact, we want to predict the expected EAP: $\mathbb{E}\left[EAP\right] = \mathbb{E}\left[\mathcal{F}(V_T^{HUB})\right]$.
- To keep things tractable, we use first-order Taylor approximation: $\mathbb{E}[EAP^*] \approx \mathcal{F}\left(\mathbb{E}\left[\widehat{V}_T^{HUB*}\right]\right)$

Vertical extrapolation reminder: Shear modeling

Goal: Predict (in practice, simulate) short-term hourly wind speeds at hub height

$$v_{\mathcal{T}'}^{HUB} := v_t^{HUB}{}_{t \in \mathcal{T}'},$$

- Power-law, aka shear, model:

$$v_{t,i}^{SIT} = v_t^{REF} \left(z_i / z_{REF}\right)^{\alpha_{h(t),m(t)}} + \sigma_{h(t),m(t)} \varepsilon_{t,i} \qquad (2)$$

  with:
  - $v_{t,i}^{SIT}$ mast measures time-series at height z_i for $i \in \{1, \ldots, n\}$;
  - $v_t^{REF} := v_{t,0}^{SIT}$ reference time-series (mast measure with height $z^{REF} := z_0$ closest to hub-height z^{HUB})
  - $\varepsilon_{t,i} \overset{iid}{\sim} \mathcal{N}(O, 1)$ the measurement / modeling errors
- "shear" parameter $\alpha_{h,m}$ and variance $\sigma_{h,m}^2$ depend on hour $h \in \{0, \ldots, 23\}$ and month $m \in \{1, \ldots, 12\}$
- OLS estimate $\widehat{\alpha}_{h,m}$ and $\widehat{\sigma}_{h,m}^2$ used in EDF-Re methodology to simulate the short-term hub-height time-series, following:

$$\widehat{v}_{\mathcal{T}'}^{HUB} = \left(v_t^{REF} \left(z^{HUB} / z^{REF} \right)^{\widehat{\alpha}_{h(t),m(t)}} \right)_{t \in \mathcal{T}'} \qquad (3)$$

Long Term reconstruction reminder: MCP regression

Goal: Predict (or in practice, simulate) long-term wind speeds at hub height

$$v_{\mathcal{T}}^{HUB} := (\hat{v}_t^{HUB})_{t \in T},$$

- MCP (matrix correlate predict) linear regression model:

$$\widehat{v}_t^{HUB} = \beta_{s(t),0} + v_t^{SAT} \beta_{s(t),1} + \gamma_{s(t)} \xi_t, \quad (4)$$

  This can be estimated based on the following quantities for time-steps t in the common period $\mathcal{T} \cap \mathcal{T}'$:
  - $(\widehat{v}_t^{HUB}$ simulated short-term wind speeds at hub height (obtained through vertical extrapolation);
  - v_t^{SAT} satellite reconstruction;
  - $\xi_t \overset{iid}{\sim} \mathcal{N}(0,1)$ measurement / modeling error term.
- MCP coefficients β_s and variance term γ_s^2 depend on wind sector $s = 1, \ldots, 12$, computed from direction satellite proxy D^{SAT}
- OLS estimates $(\widehat{\beta}_s, \widehat{\gamma}_s^2)_{s=1,\ldots,12}$ used to predict long-term wind speeds at hub-height:

$$\widehat{v}_{\mathcal{T}}^{HUB} = \left(\widehat{\beta}_{s(t),0} + \widehat{\beta}_{s(t),1} v_t^{SAT} \right)_{t \in \mathcal{T}}, \quad (5)$$

Limits of current modeling and proposed alternative

- The reconstructed hub-height mast measures $\widehat{V}_{T'}^{HUB}$ is modeled according to the MCP model:

$$\widehat{v}_t^{HUB} = \beta_{s(t),0} + \beta_{s(t),1} v_t^{SAT} + \gamma_{s(t)} \xi_t$$

  even though it has been simulated acccording to the vertical extrapolation "shear" model:

$$\widehat{v}_t^{HUB} = v_t^{REF} \left(z_{hub} / z_{ref} \right)^{\widehat{\alpha}_{h(t),m(t)}} + \sigma_{h(t),m(t)} \varepsilon_t,$$

- It is not clear whether MCP and shear models assumptions are compatible, which may result in artifically biased results
- This is why, we propose to apply MCP modeling to (observed) reference time-series $V_{T'}^{REF}$ rather than (simulated) hub-height time-series $\widehat{V}_{T'}^{hub}$, assuming that:

$$v_t^{REF} = \beta_{s(t),0} + \beta_{s(t),1} v_t^{SAT} + \gamma_{s(t)} \xi_t.$$

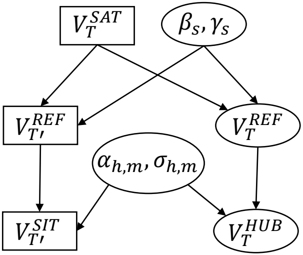
Full uncertainty model

Our final statistical model assumptions reads:

$$
\begin{aligned}
v_t^{REF} &= \beta_{s(t),0} + \beta_{s(t),1} v_t^{SAT} + \gamma_{s(t)} \xi_t \\
v_{t,i}^{SIT} &= v_t^{REF} \left(z_i / z_{REF} \right)^{\alpha_{h(t),m(t)}} + \sigma_{h(t),m(t)} \varepsilon_{t,i} \\
v_t^{HUB} &= v_t^{REF} \left(z_{HUB} / z_{REF} \right)^{\alpha_{h(t),m(t)}} + \sigma_{h(t),m(t)} \epsilon_t .
\end{aligned}
$$

The long-term hub-height expected wind-speed is then easily predicted $\forall t \in \mathcal{T}$ as:

$$
\mathbb{E}\left[v_t^{HUB} \right] = \left(\beta_{s(t),0} + \beta_{s(t),1} v_t^{SAT} \right) \left(z_{HUB} / z_{REF} \right)^{\alpha_{h(t),m(t)}}
$$



Directed acyclic graph (DAG) of the wind speed statistical model

Question: Can we also model and account for $V_{\mathcal{T}}^{SAT}{}'s$ uncertainty?

Signal Processing Intermezzo : the spectral factorization method

Spectral factorization is concerned with the problem of generating a standard Gaussian signal having a given target autocorrelation function $\tau \to \rho^*(\tau)$. It involves designing a suitable linear filter $\mathcal{H}$, used afterwards on a Gaussian white noise sample $(\epsilon_t)_t$. The filter's output should have the same autocorrelation $\tau \to \rho_s(\tau)$ as the target one.

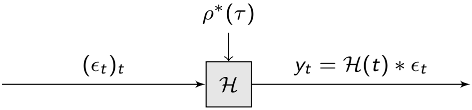


Figure: Block diagram for spectral factorization

Stationarization

The adopted model is mainly derived from this approach. To properly use it, we resorted to breaking down the underlying stochastic process $(V_t^{SAT})_t$ into a stationary process $(X_t^{stat})_t$ and a residual process $(X_t^{res})_t$.

$$V_t^{SAT} = X_t^{res} + X_t^{stat} \qquad \forall t \in \mathcal{T} \qquad (6)$$

We imposed the following further (strong but hopefully reasonable) assumptions :

1. The $(X_t^{res})_t$ process is deterministic,
2. The process $(X_t^{stat})_t$ is a filtered Gaussian White Noise, such as : $X_t^{stat} = \mathcal{H}_t * \epsilon_t$, where $\mathcal{H}_t$ is a linear filter and $\epsilon_t \overset{i.i.d}{\sim} \mathcal{N}(0,1)$, a standard Gaussian white noise.

In practice, we used first Fourier, followed by wavelet, decompositions to identify X_t^{res}. The Fourier decomposition was also used to generate the linear filter $\mathcal{H}$.

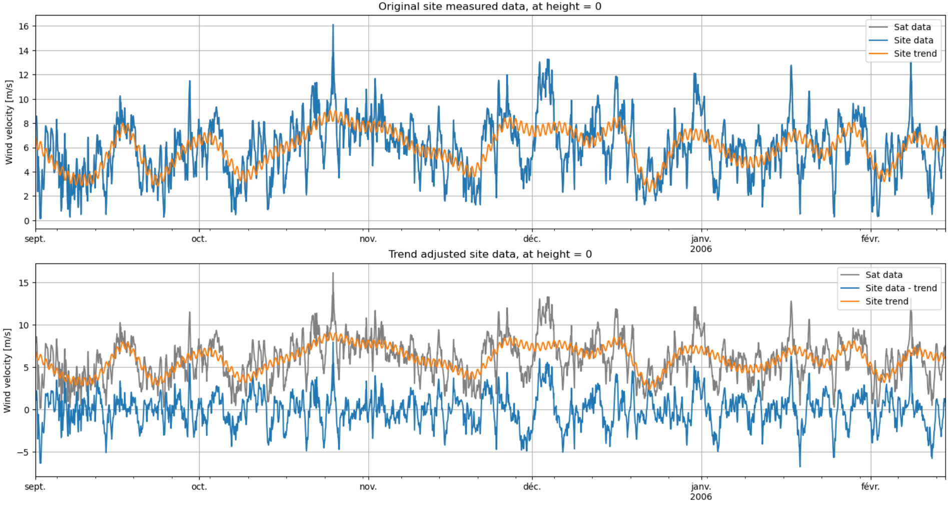
Section 3

Application to case study

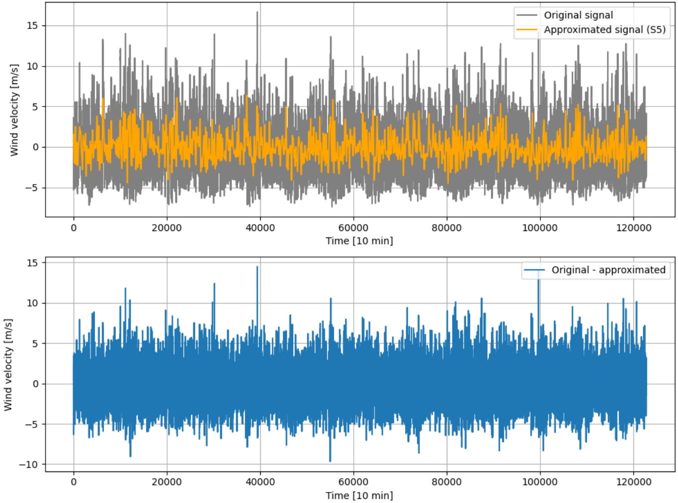
Overview of case study

- We applied our methodology to obtain a bootstrap sample of size 100 from the EAP for a certain windfarm project, both with and without accounting for the uncertainty tainting the long-term satellite data V_T^{SAT}.
- The wind-speed modeling, bootstrap and spectral factorization algorithms were all coded into the experimental Python `winduq` package, which depending on numerous standard packages (`statsmodel`, `scipy`, `pandas`, `openturns`, ...)
- the spatial extrapolation step, enabling to propagate the hub-height, long-term wind speed time-series accross the wind farm, accounting for wake effects, and turbine power curve, was done thanks to the open-source `pywake` Python package.

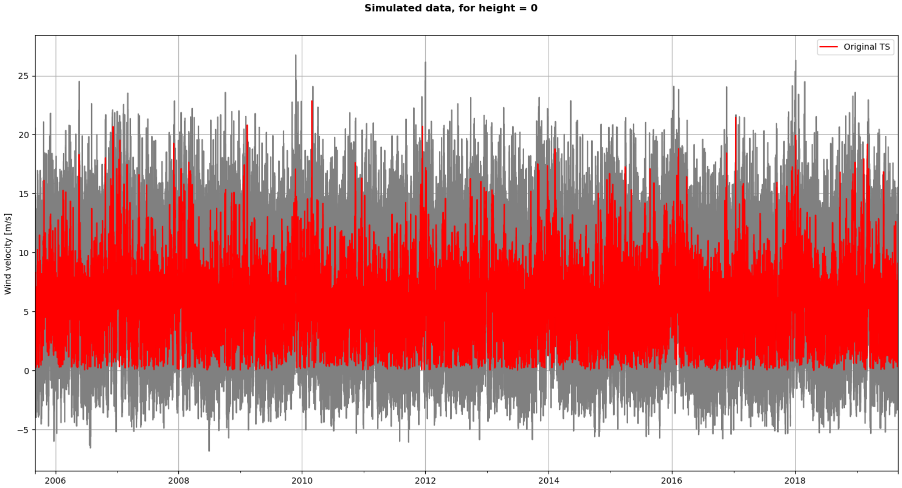
Deseason using 5 Fourier decomposition (5 first modes)



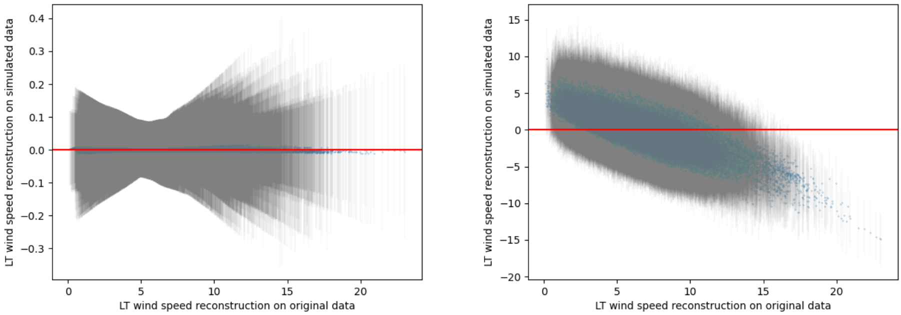
Wavelet approximation at level = 5, at height = 0
(on Fourier filtered data)



Simulated vs original satellite time-series



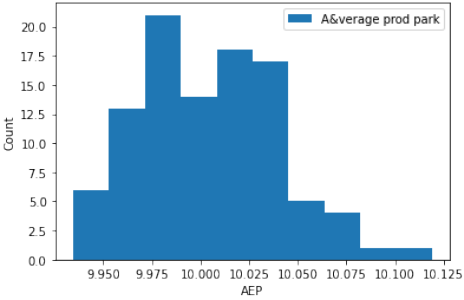
Difference between simulated and original time-series



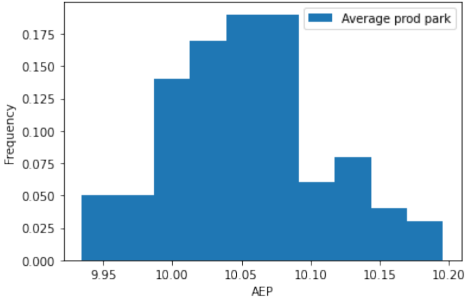
Without UQ on long-term data

With UQ on long-term data

EAP bootstrap sample



Without UQ on long-term data



With UQ on long-term data

Section 4

Discussion

Conclusions and perspectives

What we've done so far

- Propose a fully coherent statistical model for wind-speed data, both onsite measures and long-term proxys;
- Develop a parametric bootstrap approach to quantify uncertainty on the long-term EYA of the windfarm project;
- Illustrate it on a case-study

What's yet to be done

- Resolve negative wind speed simulation issus, due to Gaussian assumption, as well as ensuing bias towards real data
- Quantify uncertainty on long-term wind direction also, not only absolute speed, for instance by considering the 2D speed vector (may solve first point!)
- Elicit priors on the uncertain parameters and calculate Bayesian predictive distribution to solve double Monte-Carlo issue
- Perform sensitivity analysis to identify most influent uncertainty sources;
- More informed long-term weather predictions, accounting e.g. for climate change
- ...

Thank You for your Attention!