KJELLER
VINDTEKNIKK

How to measure representativeness?

Sónia Liléo

25.06.2013, EWEA workshop, Dublin

Why this topic?

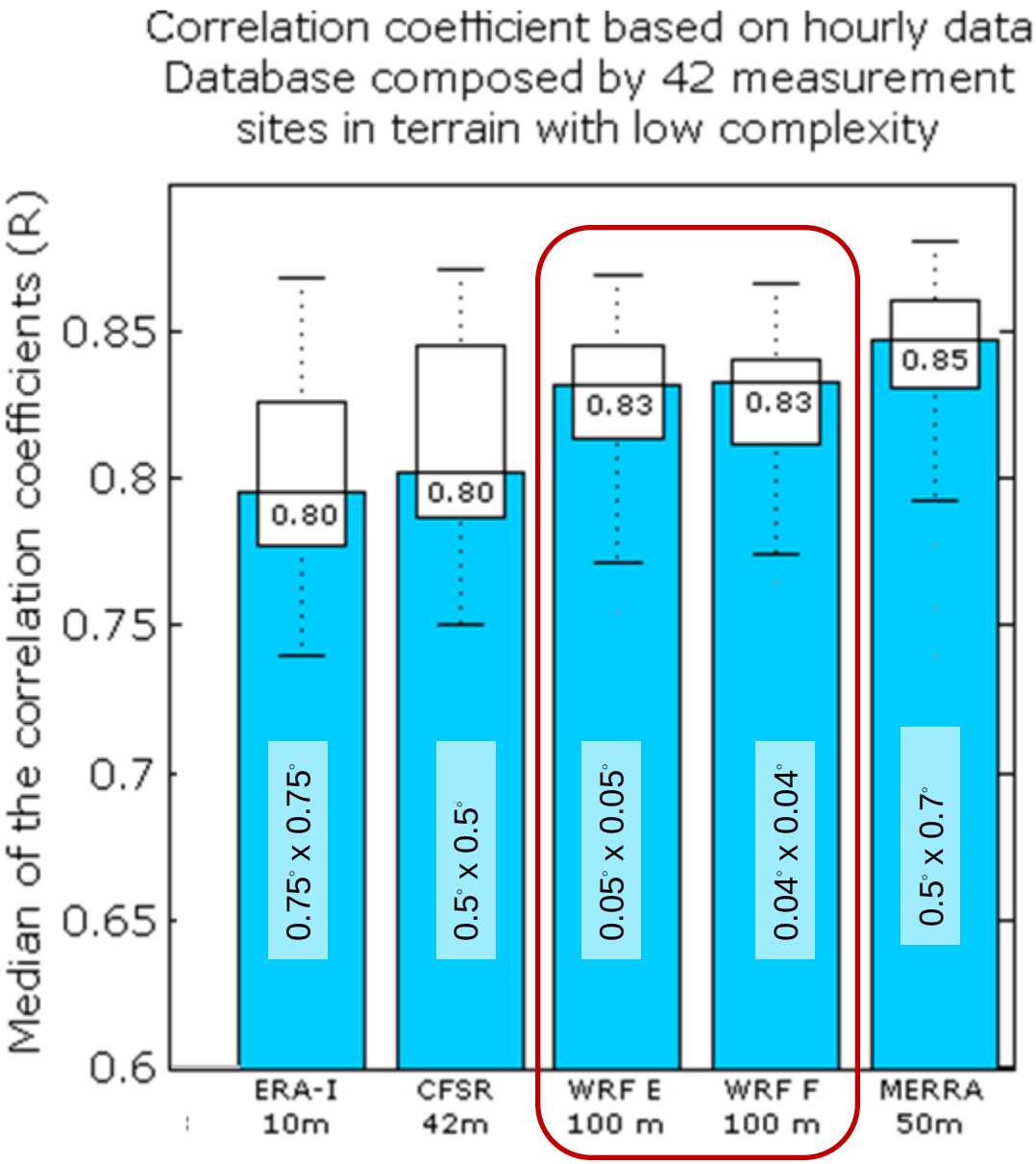
A comprehensive study that describes and compares different reanalysis datasets has been conducted.

”Long-term correction of wind measurements. State-of-the-art, guidelines and future work.”

Liléo et al., EWEA conference proceedings, 2013
Liléo et al., Elforsk report 13:18, 2013 – www.vindteknikk.com

	Reanalysis dataset	Developer	Spatial resolution (deg lat x lon)	Temporal resolution (hours)	Time span
Reanalysis global datasets	ERA-Interim	ECMWF	0.75 x 0.75	6	1979 – on
	MERRA	NASA	0.5 x 0.7	1	1979 – on
	CFSR/CFSv2	NCEP	0.5 x 0.5	6	1979 – on
Reanalysis mesoscale datasets	WRF ERA-Interim	KVT	0.05 x 0.05	1	1992 - on
	WRF FNL	KVT	0.04 x 0.04	1	2000 - on

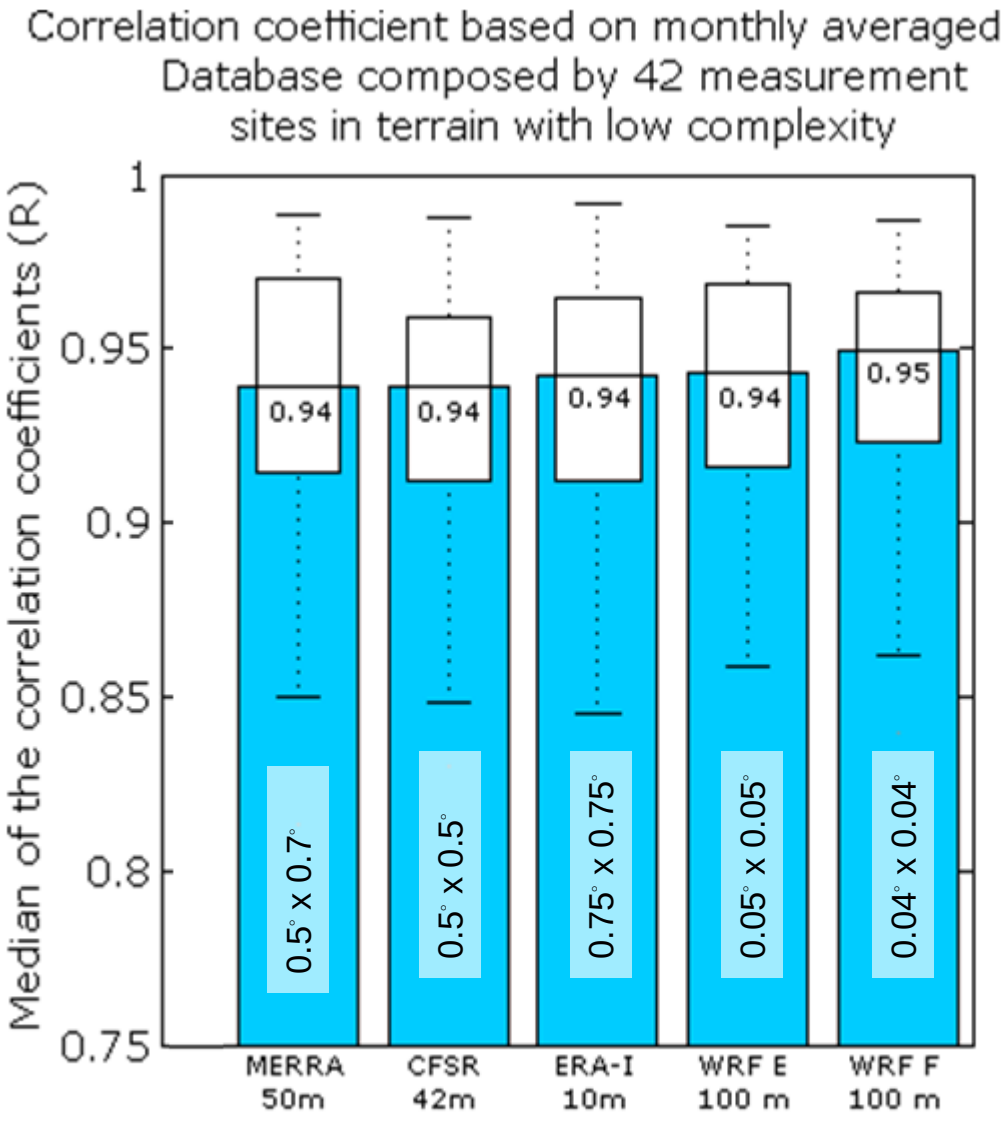
The following results were obtained when using the hourly correlation coefficient:



MERRA data appears to give a more accurate description of the temporal variations of the wind speed on a hourly basis.

KJELLER VINDTEKNIKK

When using the monthly correlation coefficient the obtained results were the following:



Very similar monthly correlation coefficients obtained for the different datasets.

All the datasets appear to describe equally well the monthly variations of the wind speed.

KJELLER VINDTEKNIKK

These results raised the following questions:

Is the correlation coefficient the most adequate parameter to be used when comparing different long-term time series?

A larger correlation coefficient is normally associated with a higher representativeness of the time series.

What is meant by representativeness?

By ”Representativeness” is meant how well a long time series represents the long-term wind conditions at a given site.

How to measure representativeness?

KJELLER
VINDTEKNIKK

What does the correlation coefficient measure?

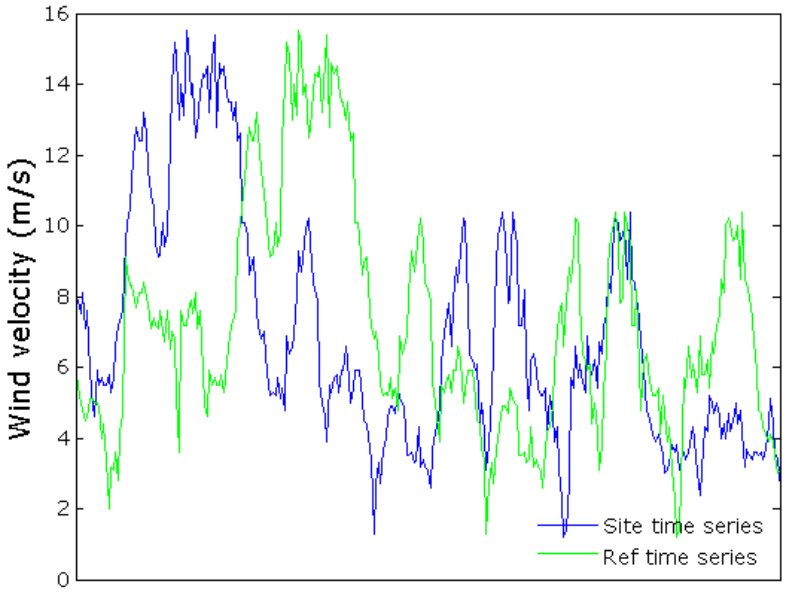
It measures how well the temporal variations of the long-term time series agree in phase with the temporal variations of the short-term time series.

The long-term time series does however not need to agree in phase with the measured time series in order to be representative of the site's long--term wind conditions.

Let's look at a simple example…

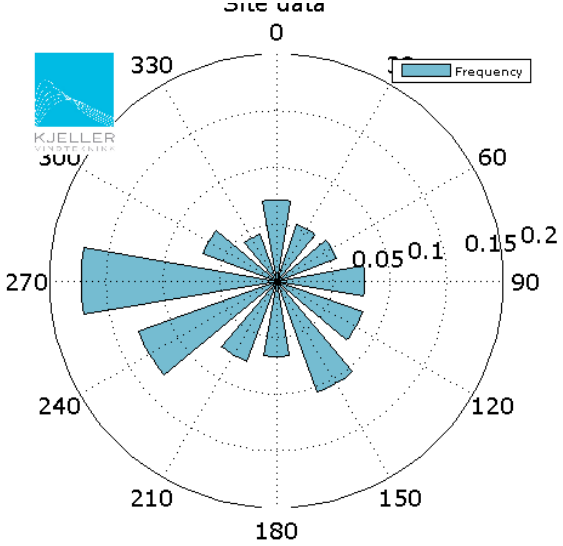
KJELLER
VINDTEKNIKK

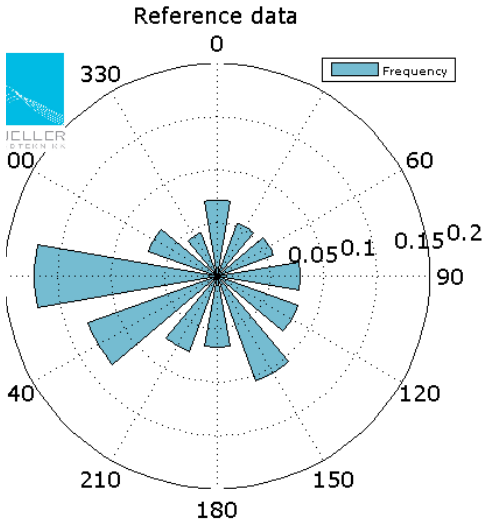
Example: The reference time series is the same as the measured series but just delayed by 2 days.



Mean site wind speed: 6.849 m/s

Mean reference wind speed during the concurrent period: 6.852 m/s

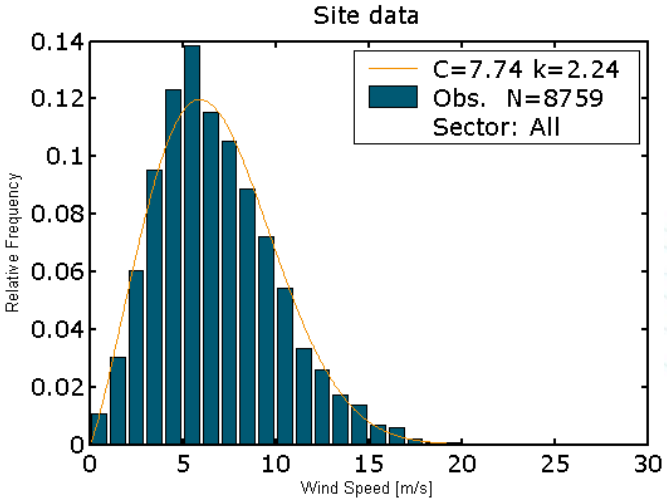


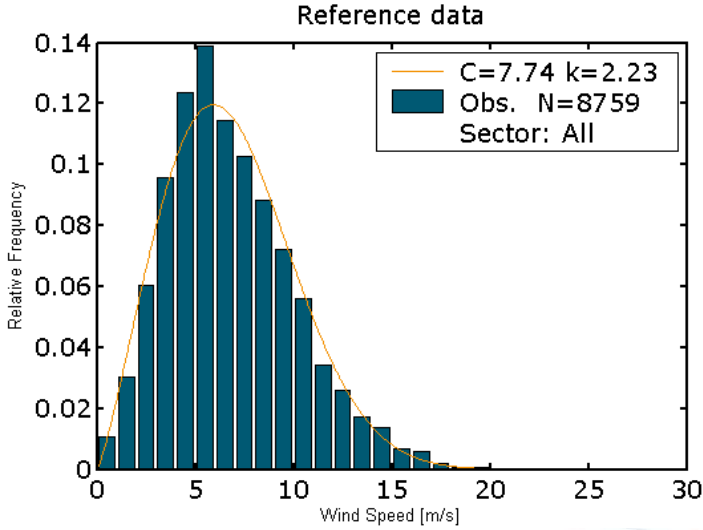


Hourly correlation coefficient: 17 %

Low corr coef

But…Good agreement of the wind rose, wind speed distribution and mean wind speed.





Do different LTC methods consider simultaneity ?

Category		Method's name	Developer	Acts on the following parameters
Regression methods				
	Least squares regression	Regression MCP	EMD, WindPRO	
		Least Squares method	GL GH, WindFarmer	
	Principal component regression	PCA method	GL GH, WindFarmer	
	Quantile regression	U&N method	KVT, internal use	

Do different LTC methods consider simultaneity ?

Category		Method's name	Developer	Acts on the following parameters
Regression methods				
	Least squares regression	Regression MCP	EMD, WindPRO	Wind speed and wind veer. Different moving sectors or non-overlapping sectors. **Considers simultaneity.**
		Least Squares method	GL GH, WindFarmer	Wind speed. Sectorwise or not. **Considers simultaneity.**
	Principal component regression	PCA method	GL GH, WindFarmer	Wind speed. Sectorwise or not. **Considers simultaneity.**
	Quantile regression	U&N method	KVT, internal use	Wind speed and wind direction. Values in ascending order, not temporal. **Does not consider simultaneity.**

Category		Method's name		Acts on the following parameters
Non-regression methods				
	Linear scaling methods	Weibull Scale method	EMD, WindPRO	Weibull scale parameter, Weibull shape parameter, occ.frequency, and mean wind speed. For each of 12 direction bins. **Does not consider simultaneity.**
		Wind Index method	EMD, WindPRO	Power output. Power curve must be chosen. **Does not consider simultaneity.**
		T&N method	KVT, internal use	Wind speed-up. Takes into account the corr coef explicitly and the probability of a given bin at site occurring simultaneously as a given bin at reference. **Considers simultaneity.**
		KH method	KVT, internal use	Wind speed average weighted by the occ freq of each direction bin. **Does not consider simultaneity.**

Category		Method's name		Acts on the following parameters
Non-regression methods				
	Linear scaling methods	Weibull Scale method	EMD, WindPRO	Weibull scale parameter, Weibull shape parameter, occ.frequency, and mean wind speed. For each of 12 direction bins. **Does not consider simultaneity.**
		Wind Index method	EMD, WindPRO	Power output. Power curve must be chosen. **Does not consider simultaneity.**
		T&N method	KVT, internal use	Wind speed-up. Takes into account the corr coef explicitly and the probability of a given bin at site occurring simultaneously as a given bin at reference. **Considers simultaneity.**
		KH method	KVT, internal use	Wind speed average weighted by the occ freq of each direction bin. **Does not consider simultaneity.**
	Probabilistic LTC methods	Matrix Method MCP	EMD, WindPRO	Wind speed-up and wind veer sorted according to the ref wsp and wdir into a matrix. **Considers simultaneity.**

Do different LTC methods consider simultaneity ?

Category		Method's name	Developer	Considers simultaneity ?
Regression methods				
	Least squares regression	Regression MCP	EMD, WindPRO	Yes
		Least Squares method	GL GH, WindFarmer	Yes
	Principal component regression	PCA method	GL GH, WindFarmer	Yes
	Quantile regression	U&N method	KVT, internal use	No
Non-regression methods				
	Linear scaling methods	Weibull Scale method	EMD, WindPRO	No
		Wind Index method	EMD, WindPRO	No
		T&N method	KVT, internal use	Yes
		KH method	KVT, internal use	No
	Probabilistic LTC methods	Matrix Method MCP	EMD, WindPRO	Yes

These methods are not adequate for cases with low corrcoef.

These methods may be applied even if the corrcoef is low.

What should representativeness take into account?

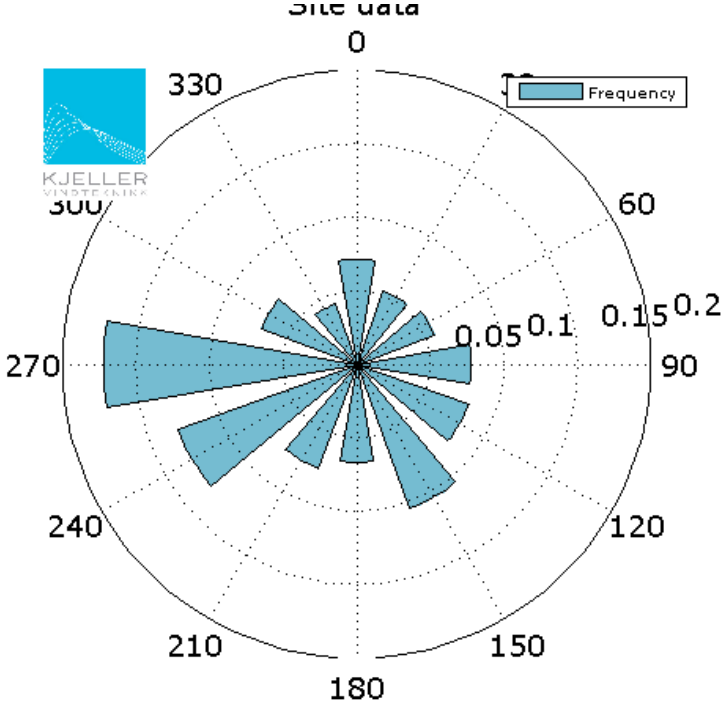
- How well the reference **wind rose** agrees with the site wind rose.

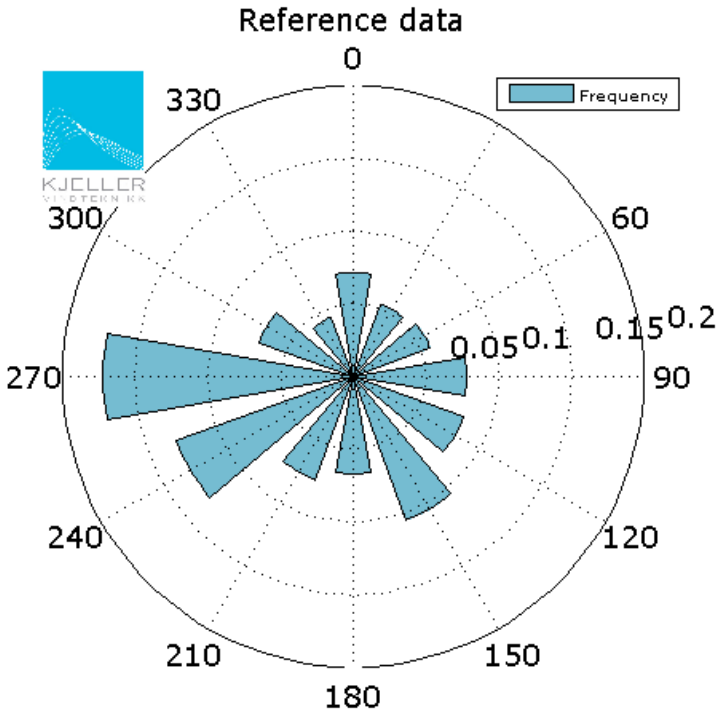
- How well the **frequency distribution** of the reference wind speed agrees with that of the site wind speed.

- How well the **power spectra** of the reference wind speed agrees with that of the site wind speed.

Let’s test it in our example…

KJELLER
VINDTEKNIKK





Good agreement of the site and reference wind roses

How to quantify the agreement?

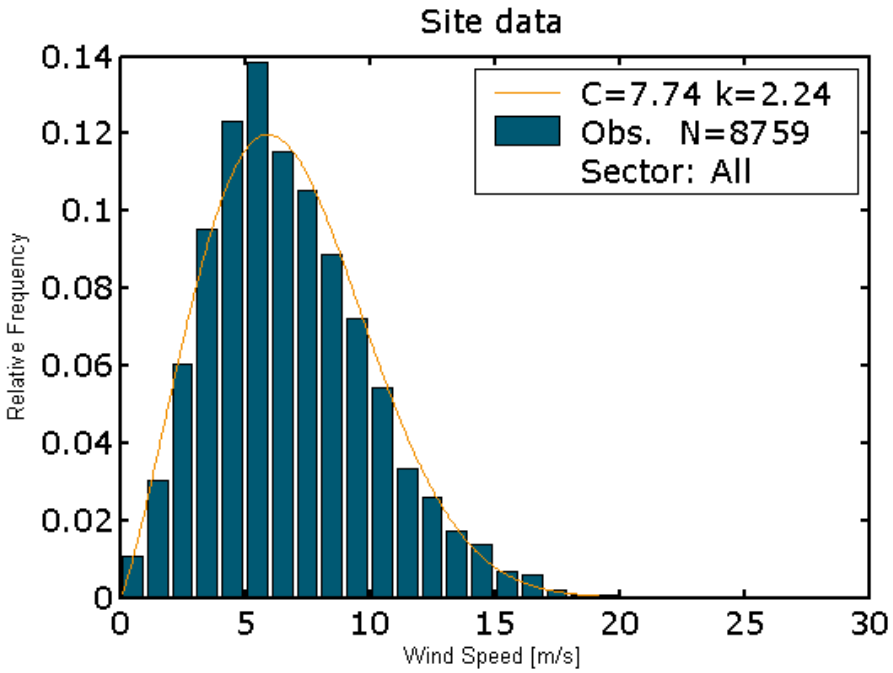
$$Wind\,rose\,difference = \sum_{sector\,i=0}^{all} |occ\,freq\,site_{sec\,i} - occ\,freq\,ref_{sec\,i}|$$

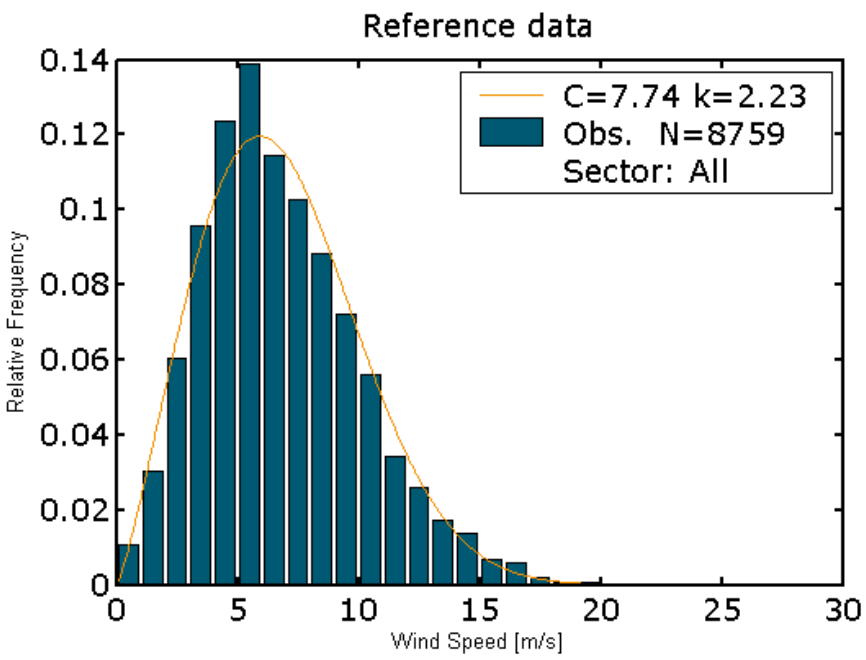
Varies between 0 and 2.

Normalizing

$$Wind\,rose\,difference = \frac{\sum_{sector\,i=0}^{all} |occ\,freq\,site_{sec\,i} - occ\,freq\,ref_{sec\,i}|}{2} * 100$$

Varies between 0 and 100 %.





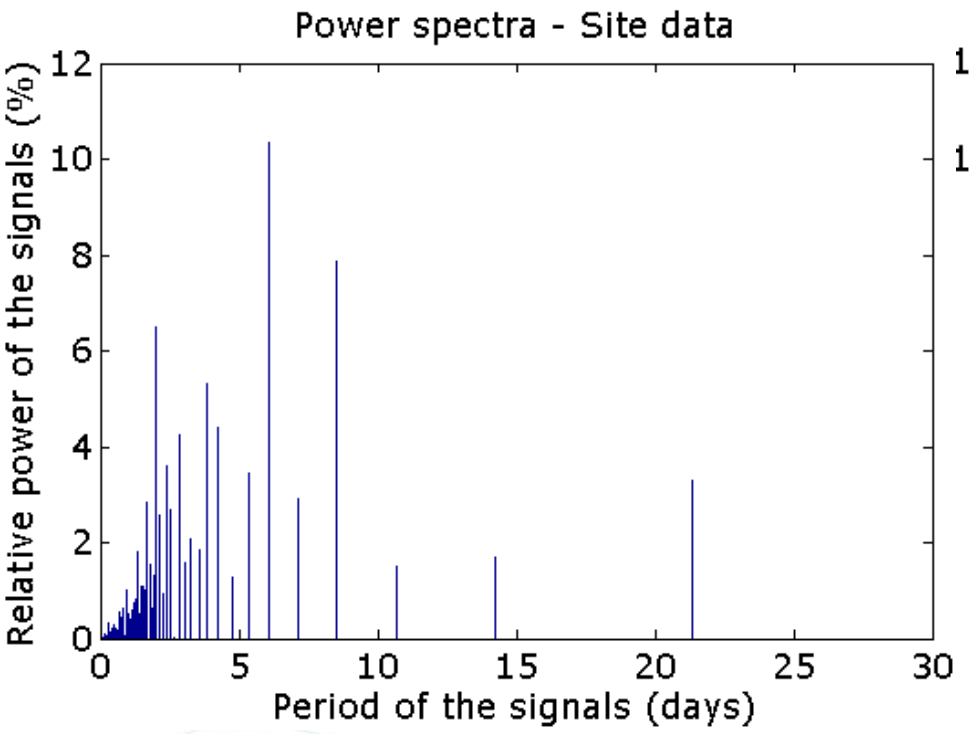
Good agreement of the wind speed frequency distributions

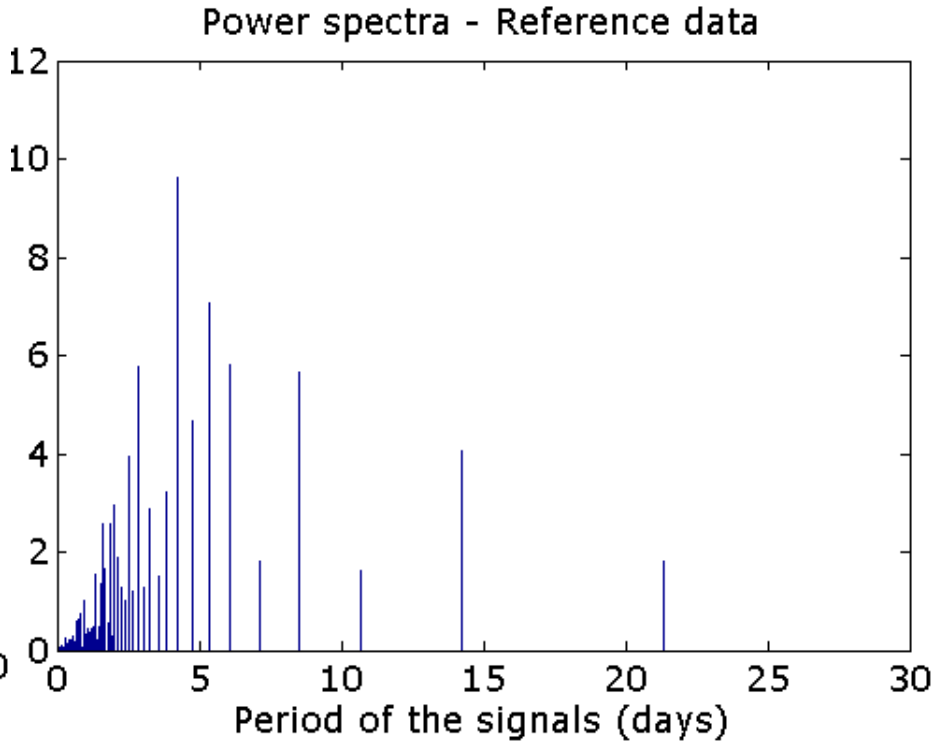
How to quantify the agreement?

$$Weibull\,scale\,difference = \frac{c\,site - c\,ref}{c\,site} * 100$$

$$Weibull\,shape\,difference = \frac{k\,site - k\,ref}{k\,site} * 100$$

KJELLER
VINDTEKNIKK





Good agreement.

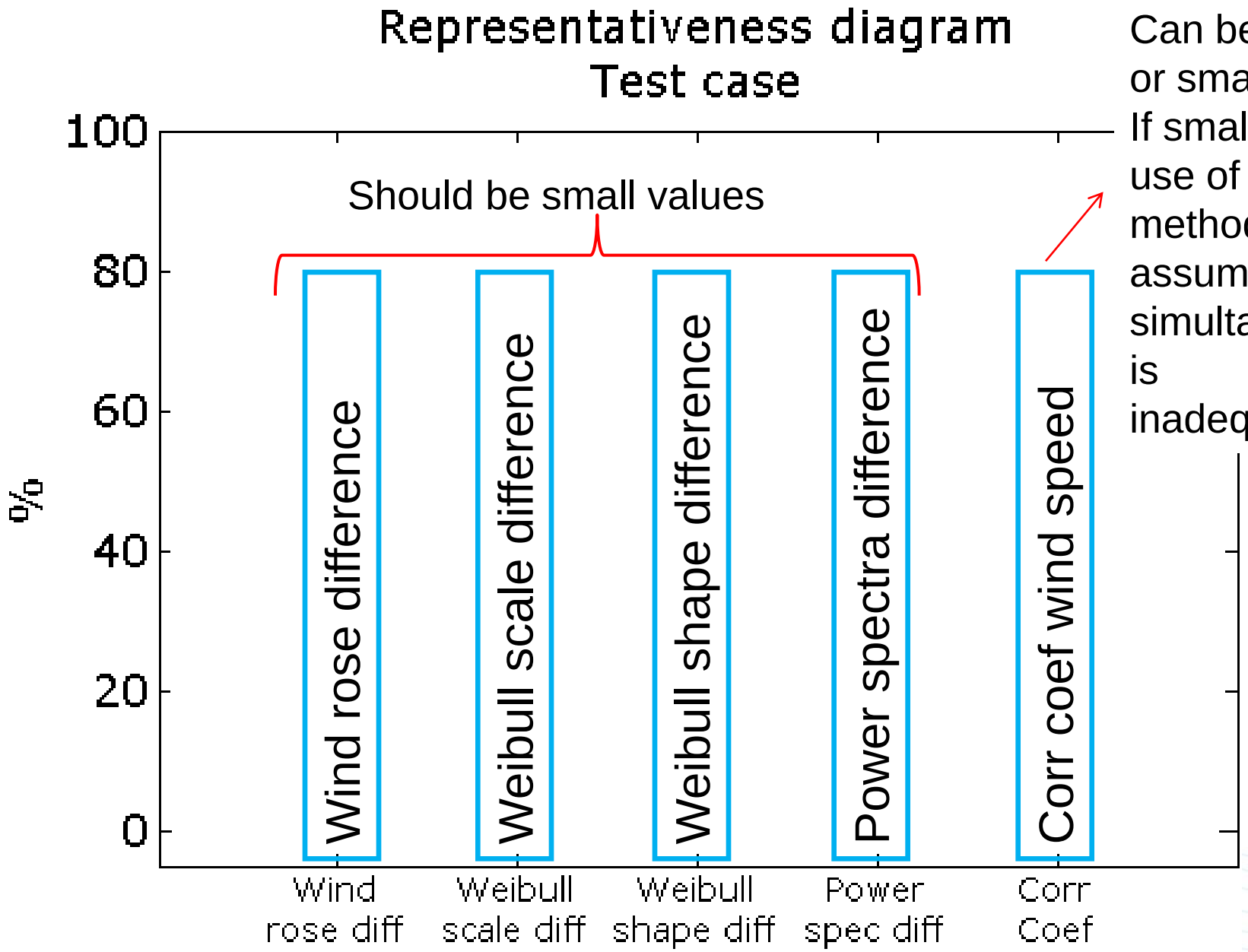
How to quantify the agreement?

$$Power\ spectra\ difference =$$

$$= \frac{\sum_{period\ bin\ i} |relative\ power\ site_i - relative\ power\ ref_i|}{2} * 100$$

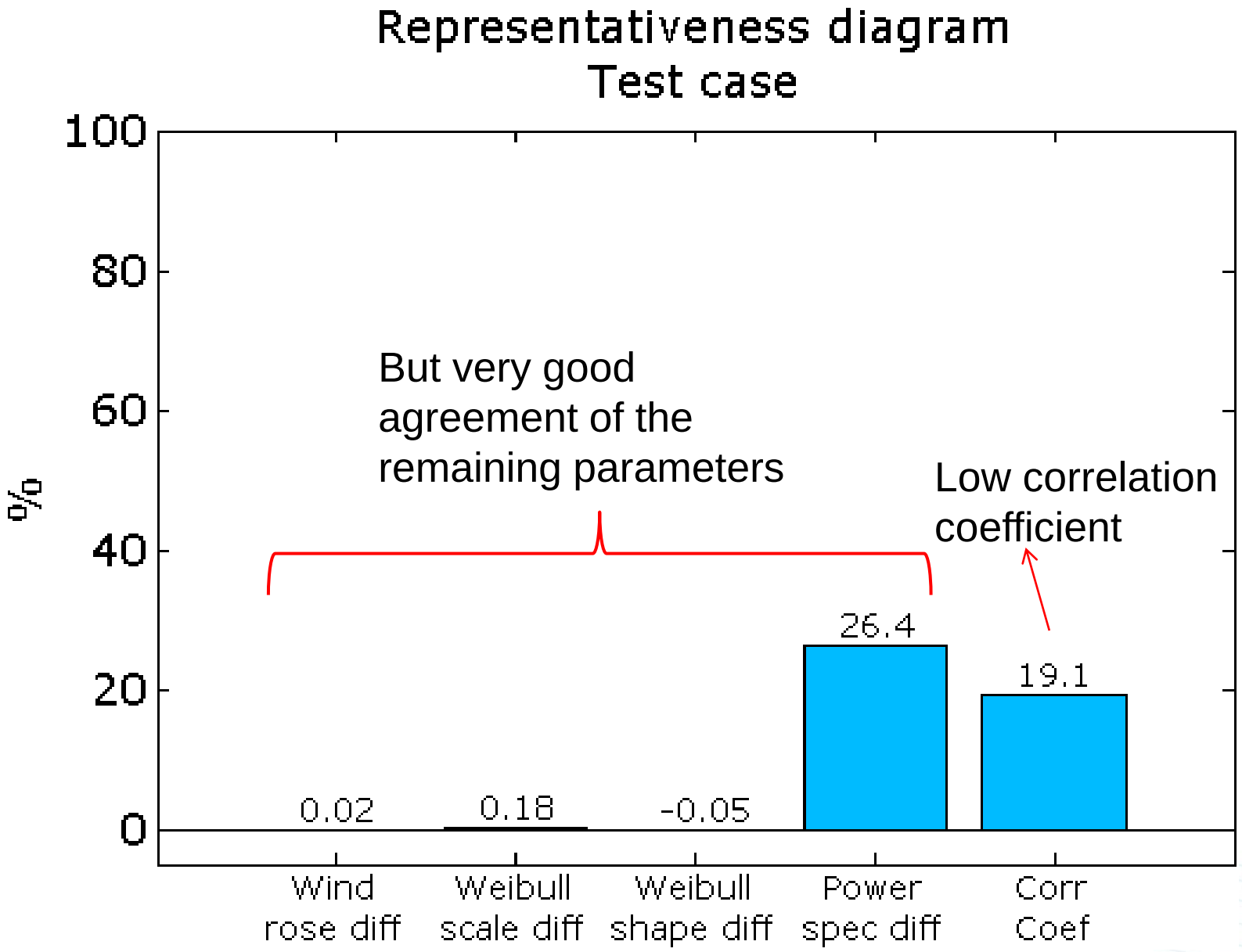
KJELLER VINDTEKNIKK

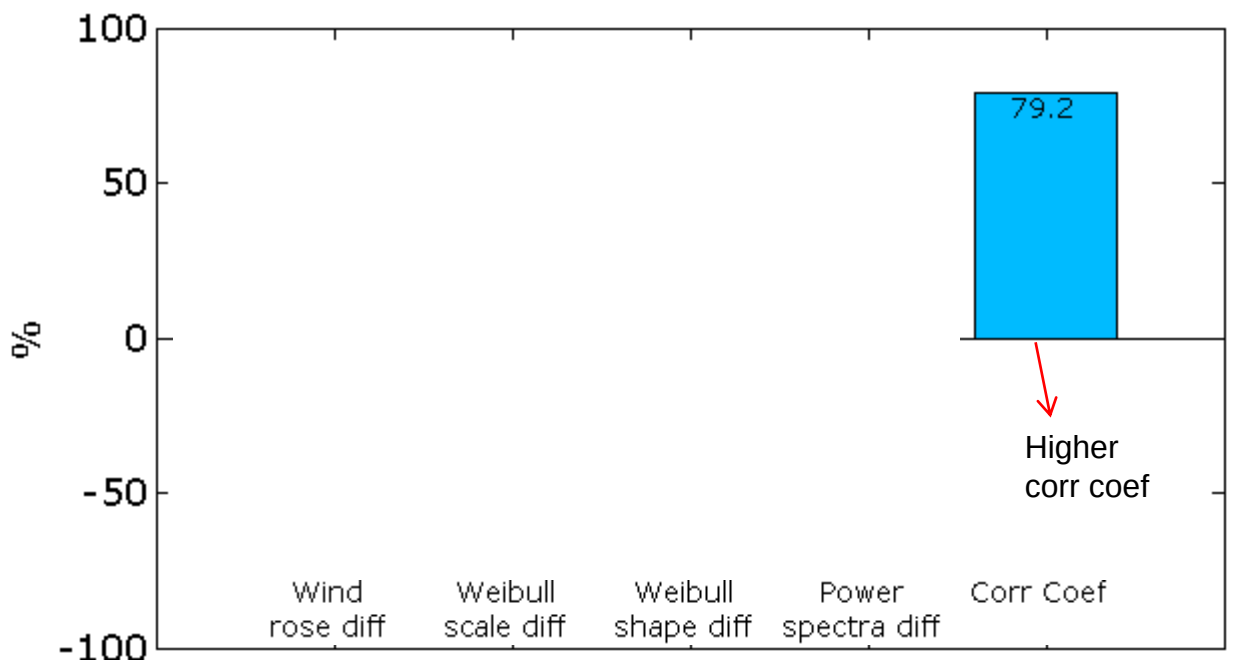
Gathering all the parameters into a diagram



Can be large or small.
If small, the use of LTC methods that assume simultaneity is inadequate.

Results for the test case





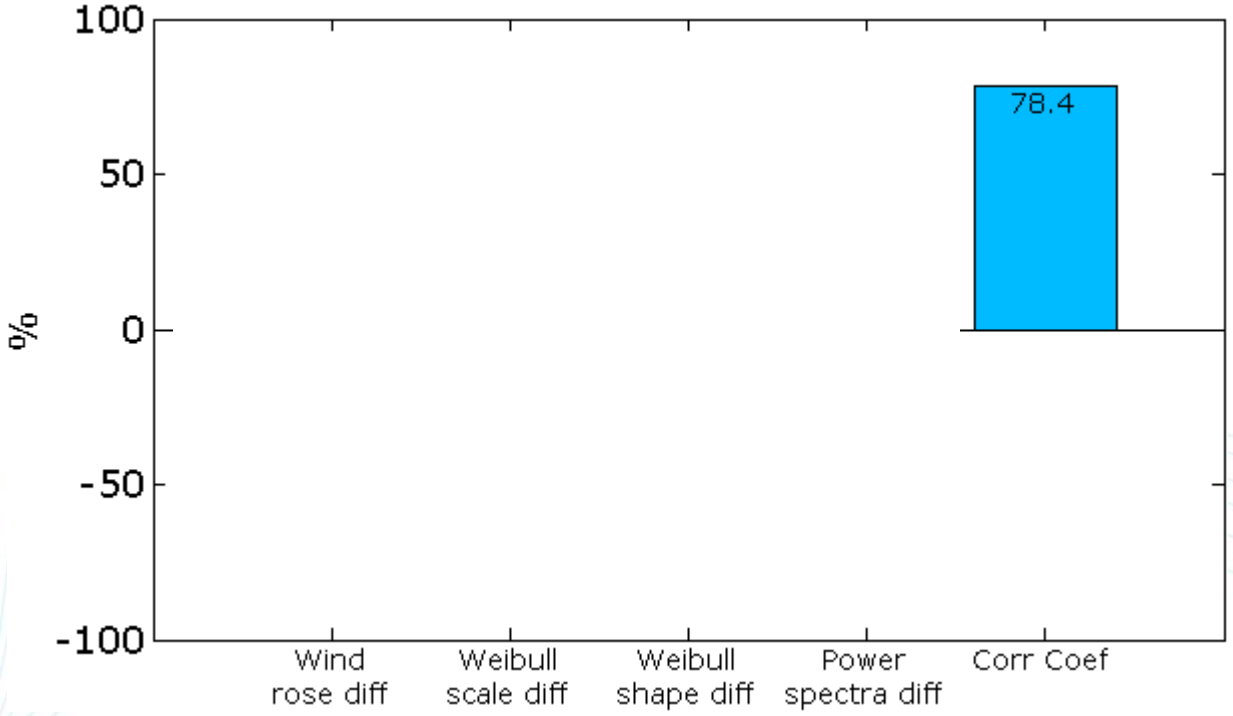
Results for a real case

Met mast series : 2.5 years data

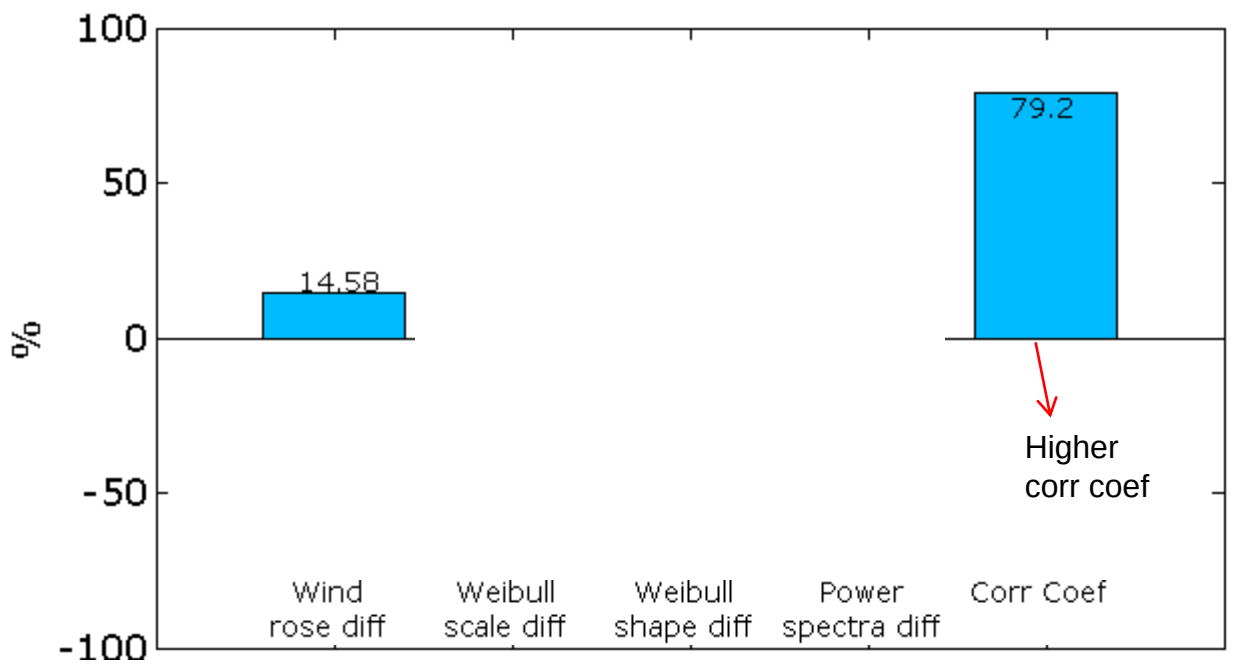
Measurement height: 100 m

Representativeness diagram

Reference data used:
MERRA



Reference data used:
WRF ERA-Interim



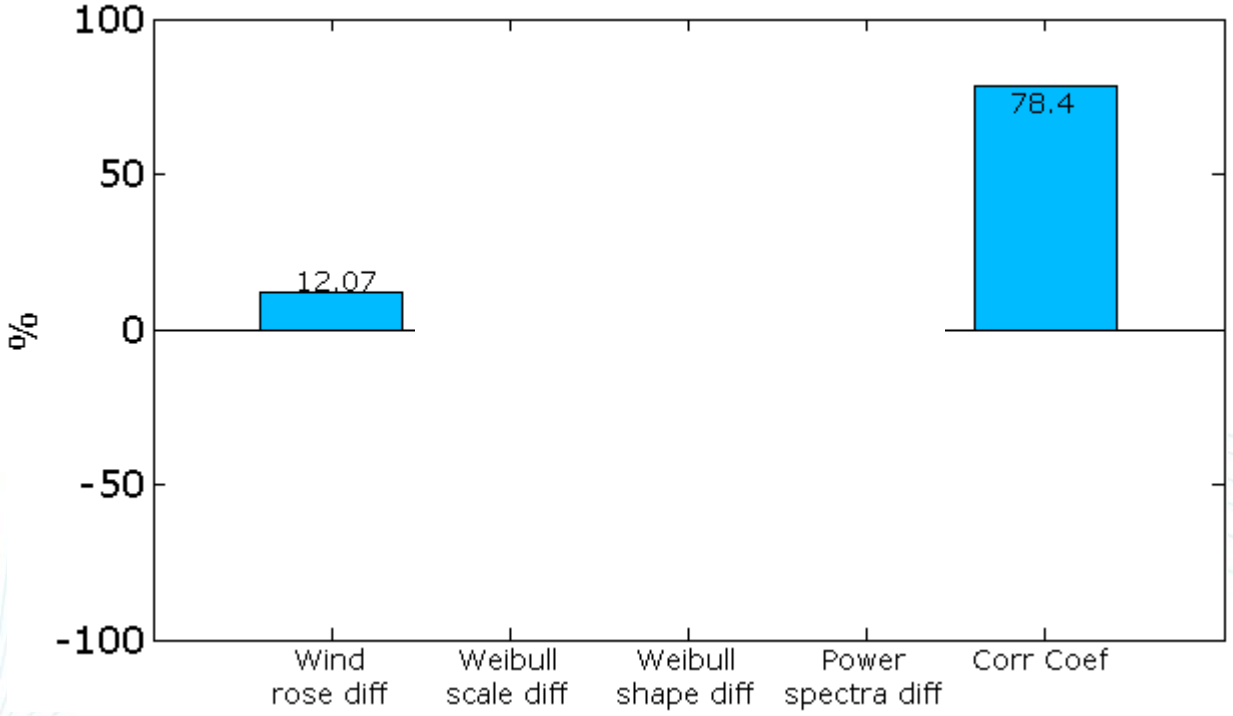
Results for a real case

Met mast series : 2.5 years data

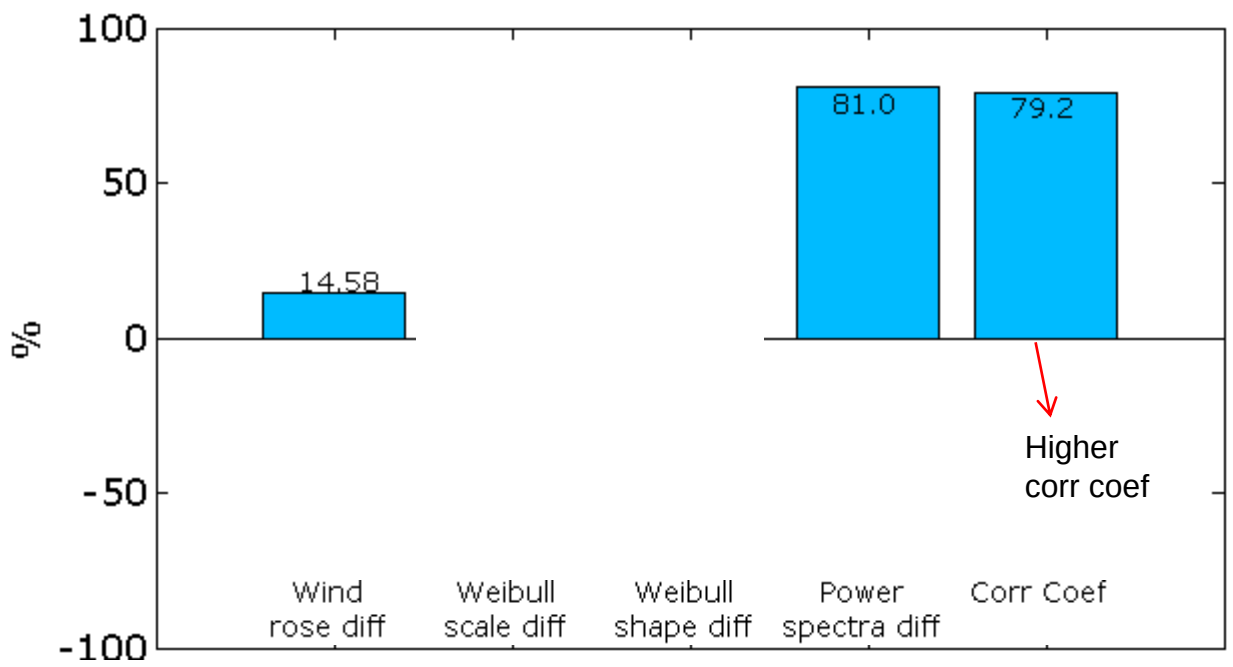
Measurement height: 100 m

Representativeness diagram

Reference data used:
MERRA



Reference data used:
WRF ERA-Interim



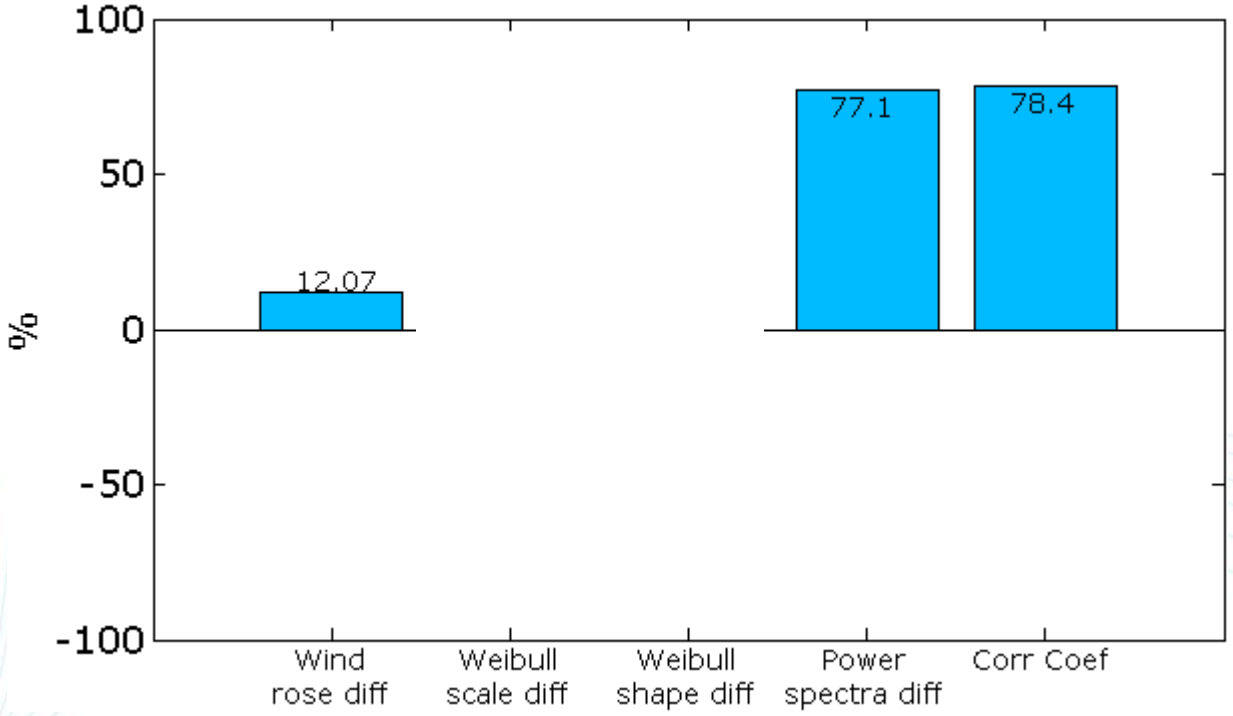
Results for a real case

Met mast series : 2.5 years data

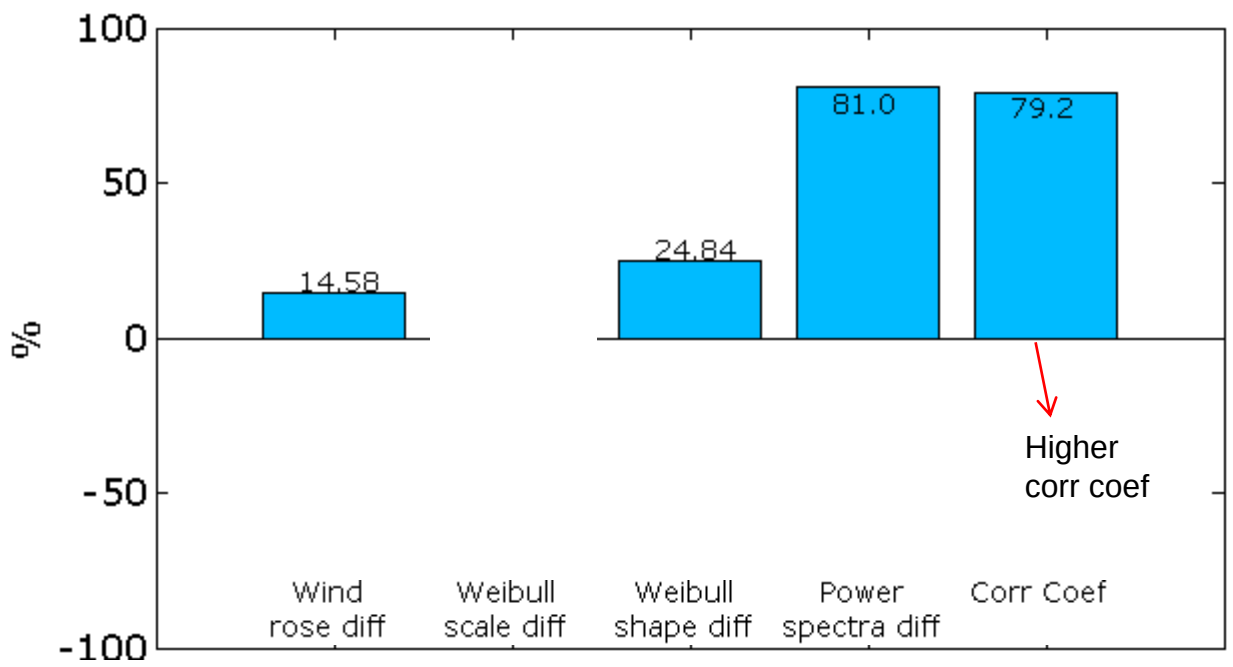
Measurement height: 100 m

Representativeness diagram

Reference data used:
MERRA



Reference data used:
WRF ERA-Interim



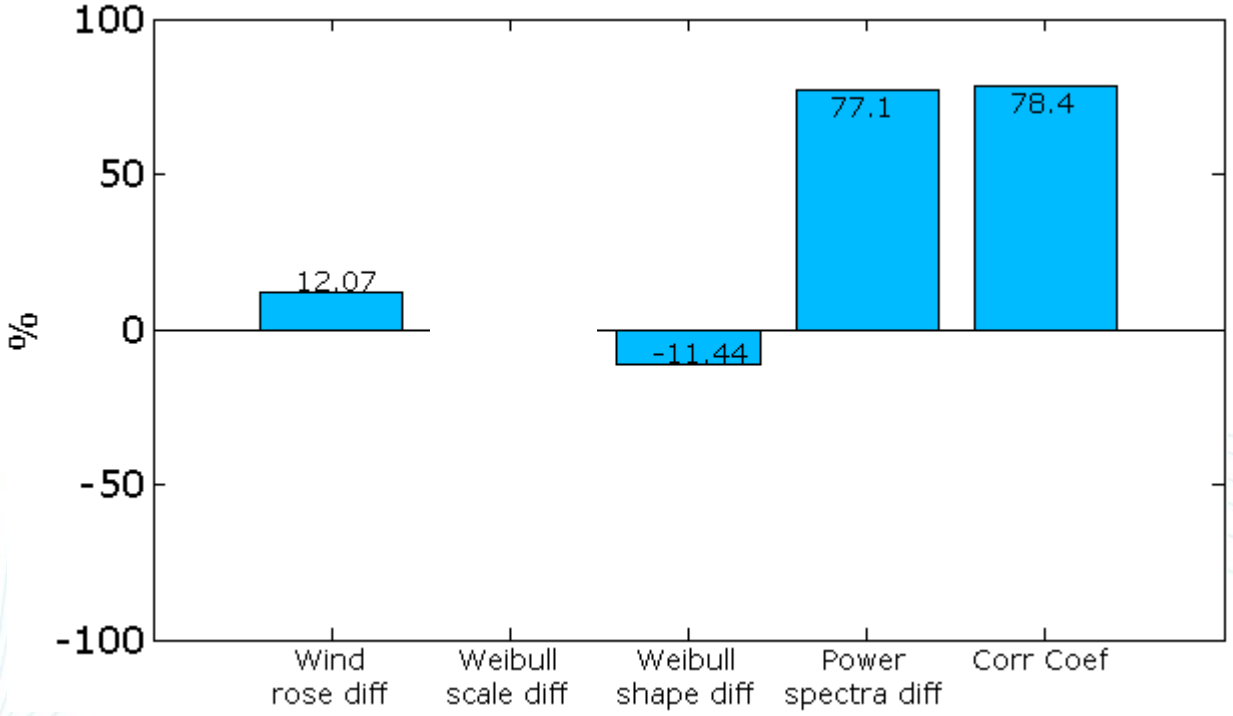
Results for a real case

Met mast series : 2.5 years data

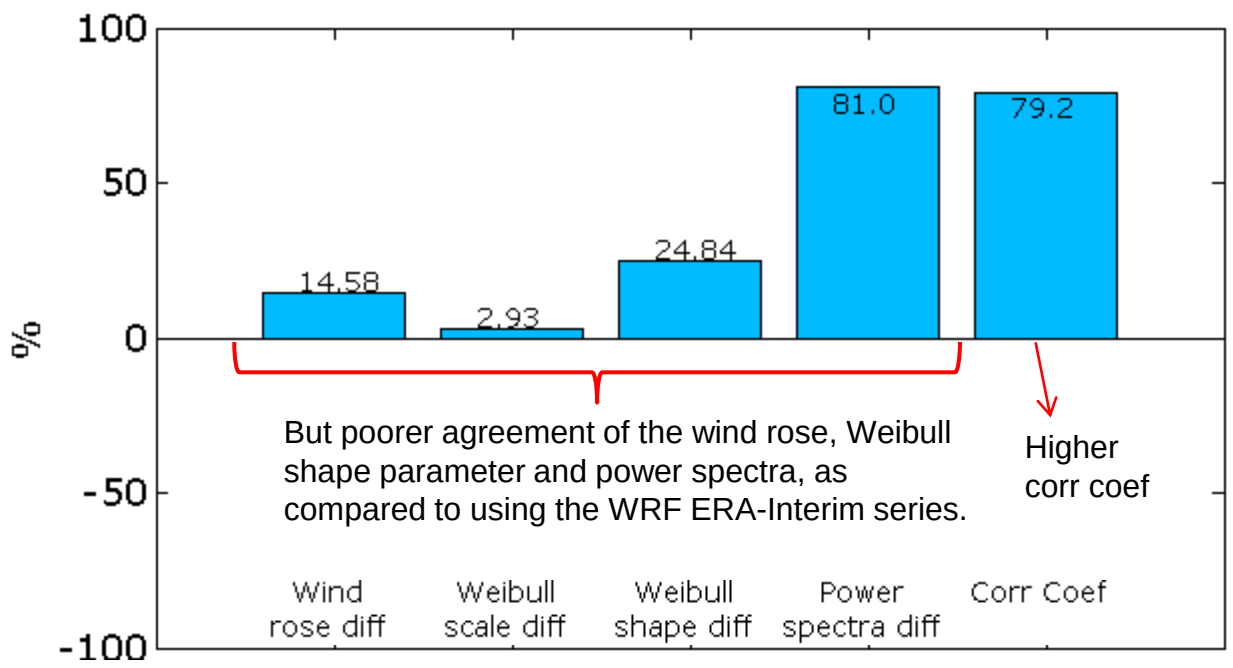
Measurement height: 100 m

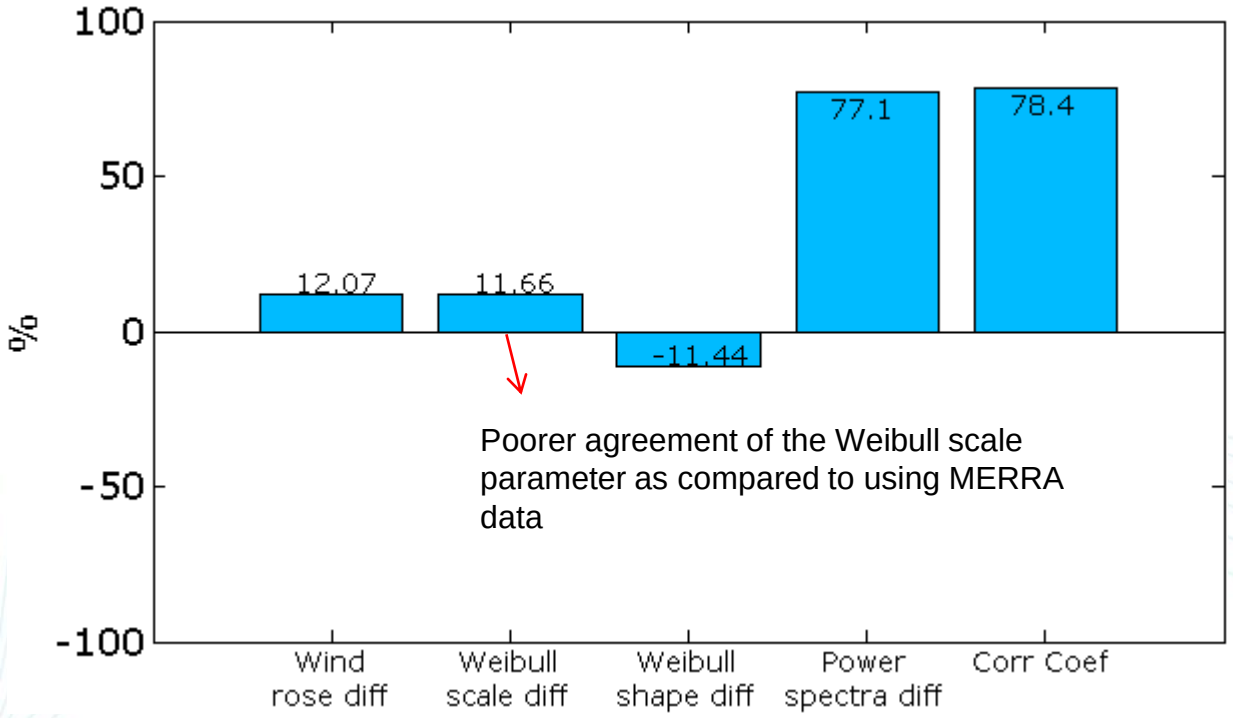
Representativeness diagram

Reference data used:
MERRA



Reference data used:
WRF ERA-Interim





Results for a real case

Met mast series : 2.5 years data

Measurement height: 100 m

Representativeness diagram

Reference data used: MERRA

Reference data used: WRF ERA-Interim

This presentation aimed to discuss the following issues:

Is the correlation coefficient the most adequate parameter to be used when comparing different long-term time series?

Do the LTC methods assume simultaneity of the site and the reference wind variations?

How to measure representativeness?

KJELLER
VINDTEKNIKK

Conclusions

- The representativeness of a long-term time series cannot be described solely by the correlation coefficient.

- The correlation coefficient might be low, although the long-term time series is representative of the site's wind climate.

- Some LTC methods require simultaneity, i.e. high correlation coefficients, some not. The majority of the analyzed methods acts on the wind speed and wind direction, but none focuses on the power spectra of the wind speed.

- The "Representativeness Diagram" compares different parameters that are relevant for the long-term correction procedure. These are the following:

    - Wind direction distribution.
    - Wind speed distribution.
    - Power spectra of the wind speed.
    - Phase of the wind speed variations.

- By comparing the "Representativeness Diagram" of different long-term time series, the most appropriate combination of long-term time series and long-term correction method might be chosen for each specific case.