HOW TO CHOOSE THE BEST SIMULATION FOR A SPECIFIC PURPOSE?

__A. Martilli__, J. L. Santiago (CIEMAT, Spain)
T. G. Reisin, (SOREQ, Israel)
A. Baklanov, R. Nuterman (DMI, Denmark)
J. Bartzis, G. Efthimiou (U. W. Macedonia, Greece)
R. Buccolieri, and S. Di Sabatino (U. Salento, Italy)
A. M. Costa, R. Tavares (U. Aveiro, Portugal)
J. Franke, (U. Siegen, Germany)
A. Hellsten, FMI, Finland)

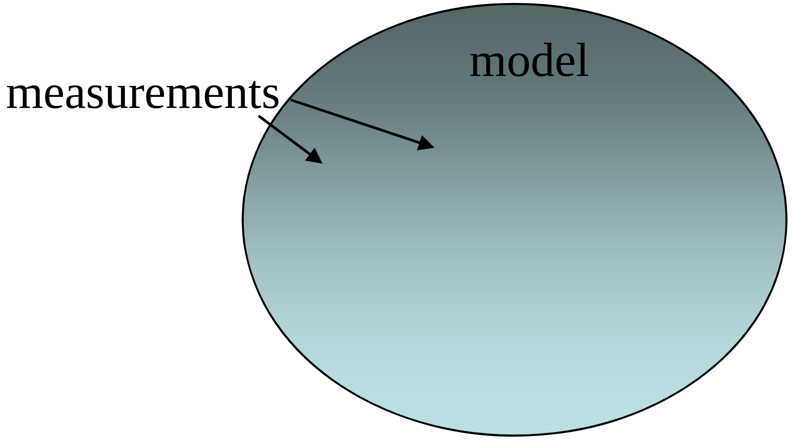
Contact: alberto.martilli@ciemat.es

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA E INNOVACIÓN Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Why do we use models?

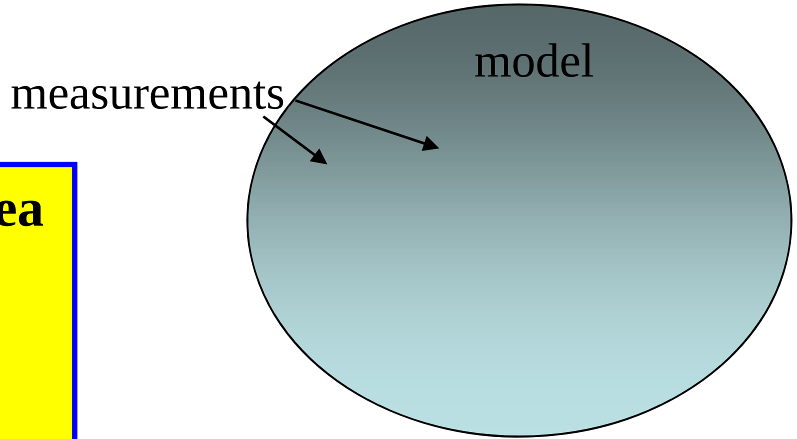
Why do we use models?

Because we want to know the value of quantities of interest (QI) for which there is no experimental information.



Why do we use models?

Because we want to know the value of quantities of interest (QI) for which there is no experimental information.



Examples: 3D distribution of pollutants, area with concentrations above a threshold, maximum concentrations, impact of a reduction strategy, tomorrow’s air quality, etc.

QI depends on the purpose of model use.

In modelling activity there are many uncertainties.

(parameterizations, numerics, initial and boundary conditions, etc.)

In modelling activity there are many uncertainties.
(parameterizations, numerics, initial and boundary conditions, etc.)

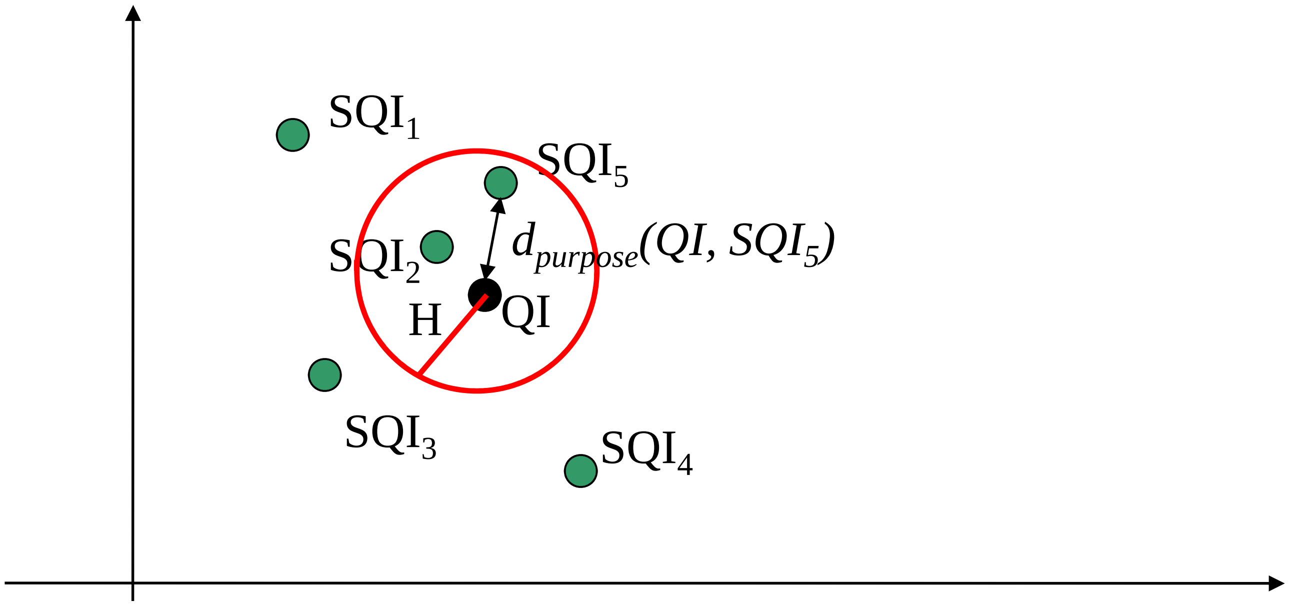
For the same case study an ensemble of simulations can be produced.

In modelling activity there are many uncertainties.
(parameterizations, numerics, initial and boundary conditions, etc.)

For the same case study an ensemble of simulations can be produced.

How to select those simulations that are fit for purpose?

This decision must be based on a measure of the "distance" between the real world value of *QI* and the simulated value *(SQI)* . $d_{purpose}(QI,SQI)$

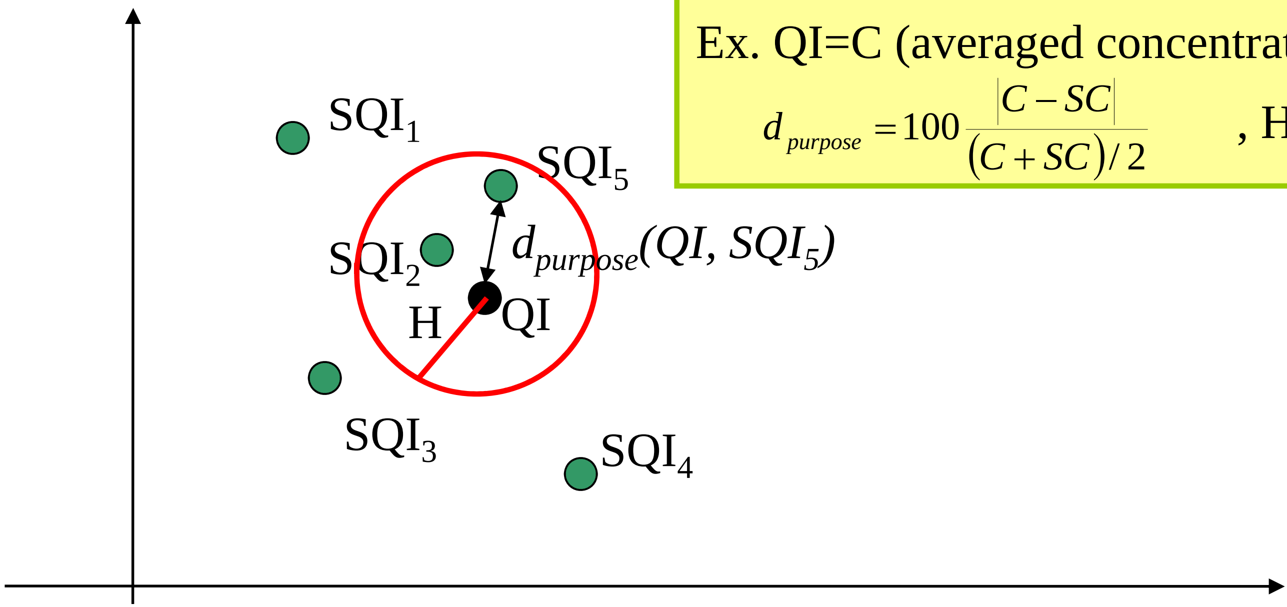


…and a quantitative criterium of acceptance (*H*) based on the purpose.

This decision must be based on a measure of the "distance" between the real world value of *QI* and the simulated value *(SQI)* . $d_{purpose}(QI,SQI)$

Ex. QI=C (averaged concentration ***in a certain area***).

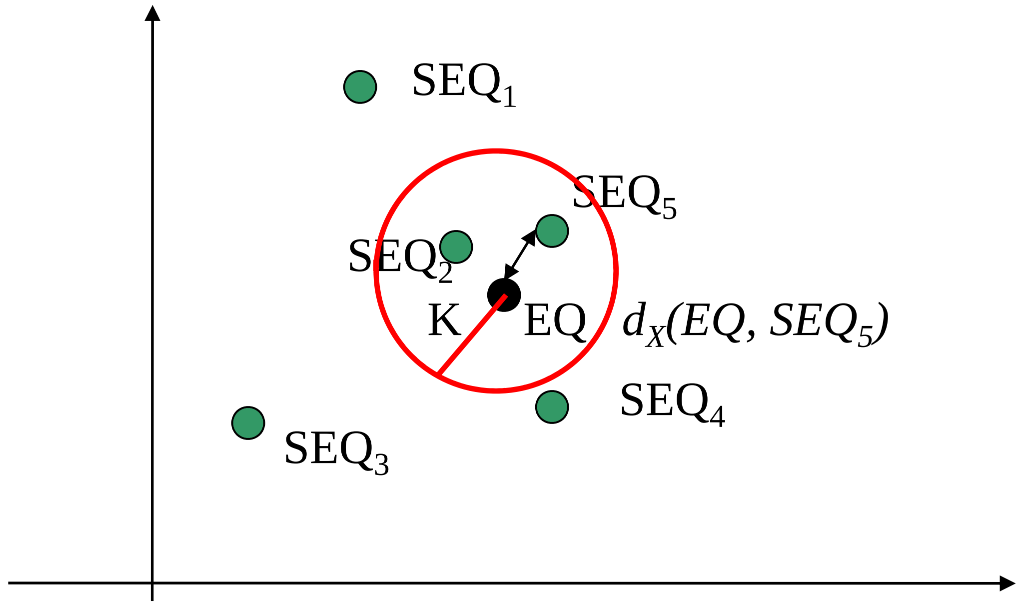
$$d_{purpose} = 100 \frac{|C - SC|}{(C + SC)/2} \quad , \text{H=50\%}$$



…and a quantitative criterium of acceptance (*H*) based on the purpose.

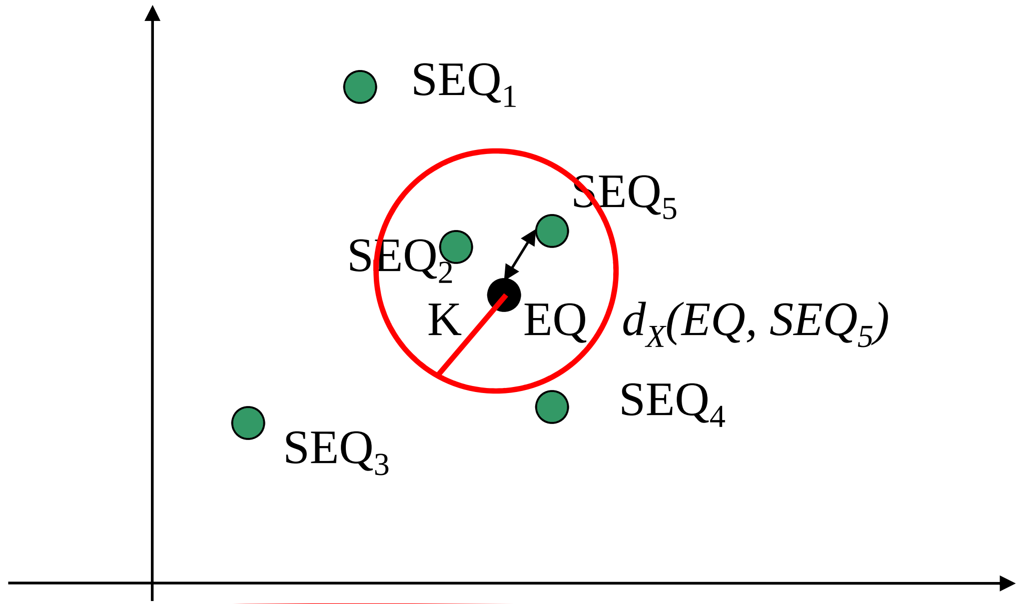
But.. the problem is that $d_{purpose}$ *(QI,SQI)* cannot be computed because there is no experimental information on *QI*.

But.. the problem is that $d_{purpose}$ *(QI,SQI)* cannot be computed because there is no experimental information on *QI*.



What can be computed are distances between experimental quantities and the correspondent simulated values $d_X(EQ,SEQ)$

But.. the problem is that $d_{purpose}$ *(QI,SQI)* cannot be computed because there is no experimental information on *QI*.



What can be computed are distances between experimental quantities and the correspondent simulated values $d_X(EQ,SEQ)$

E. g. standard metrics (RMSE, Hit Rate, Fractional Bias, Factor of 2, etc.) or new ones can be created

Ex. EQ=A (averaged concentration ***at the measurement points***).

$d_X = 100 \frac{|A - SA|}{(A + SA)/2}$ or d_X=RMSE, Fractional Bias, etc.

What's the value of K???

The problem is to select the best metrics d_X that can surrogate $d_{purpose}$,

SEQ_i simulated experimental quantity for simulation ***i***
SQI_i simulated quantity of interest for simulation ***i***

We want a d_X such that:

$$d_X\left(SEQ_i, EQ\right) > d_X\left(SEQ_j, EQ\right) \Leftrightarrow d_{purpose}\left(SQI_i, QI\right) > d_{purpose}\left(SQI_j, QI\right)$$

We want a separation value K such that:

$$d_X\left(SEQ_i, EQ\right) < K \Leftrightarrow d_{purpose}\left(SQI_i, QI\right) < H$$

How to compare metrics?

Use the ensemble of simulations to compare metrics.
For each couple of simulations *(i,j)*, estimate :

$$d_{purpose}\left(SQI_i, SQI_j\right)$$
$$d_X\left(SEQ_i, SEQ_j\right)$$

Advantage: both can be computed

And base the comparison between metrics on the following two techniques:

Models must have passed a scientific evaluation

1)

Kendall's TAU – it measures the similarity between rankings

The highest the value of τ, the most similar are the rankings

$$\tau = \frac{n_t - n_f}{N^4}$$

$$\begin{cases} d_X\left(SEQ_i, SEQ_j\right) > d_X\left(SEQ_m, SEQ_n\right) \\ d_{purpose}\left(SQI_i, SQI_j\right) > d_{purpose}\left(SQI_m, SQI_n\right) \end{cases}$$

or

$$\begin{cases} d_X\left(SEQ_i, SEQ_j\right) < d_X\left(SEQ_m, SEQ_n\right) \\ d_{purpose}\left(SQI_i, SQI_j\right) < d_{purpose}\left(SQI_m, SQI_n\right) \end{cases}$$

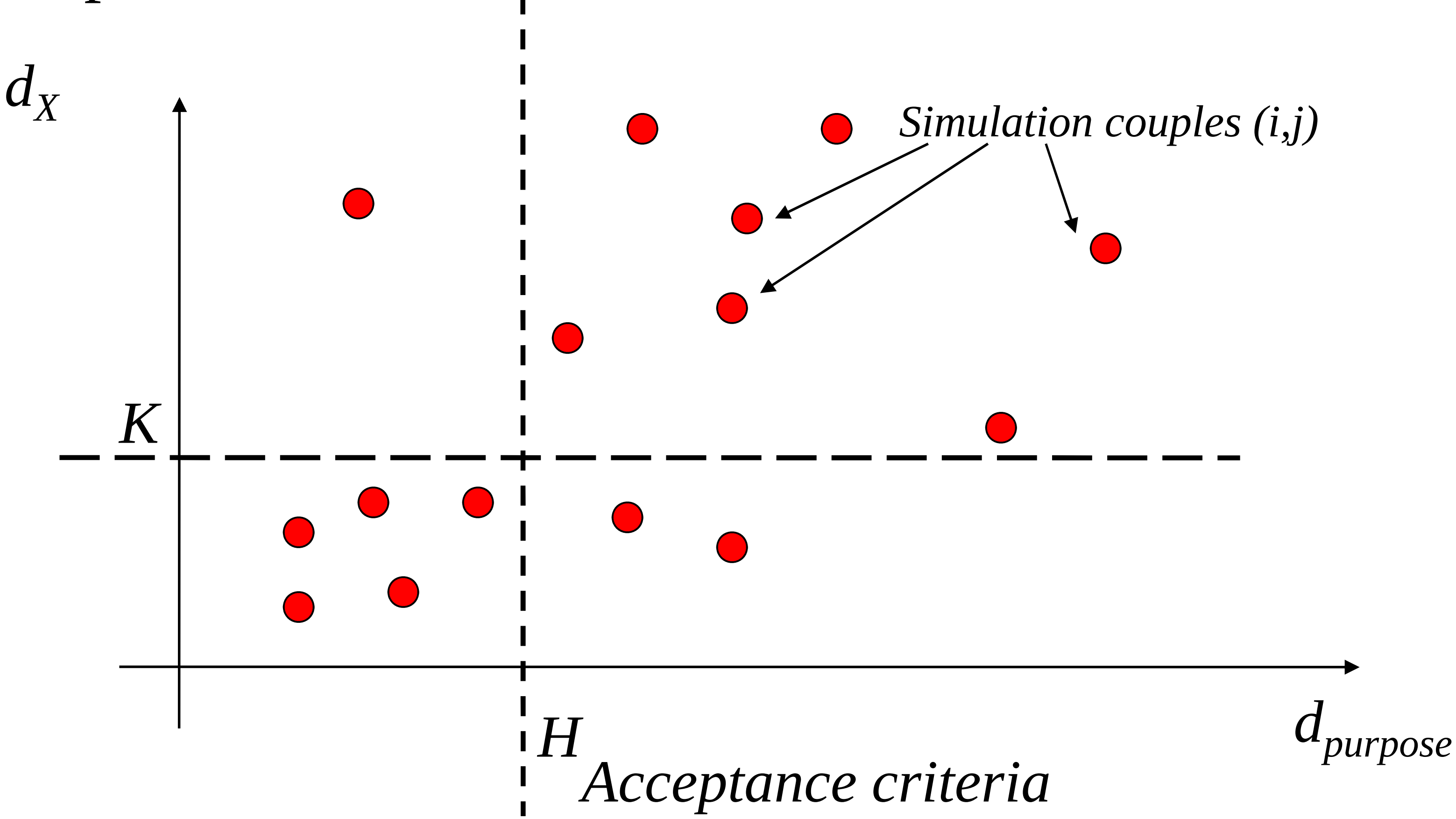
TRUE n_t

FALSE n_f

N = number of simulations

2)

Separation value.



$$m_{ij} = \begin{cases} 1 \Leftrightarrow \left[\left(d_{purpose}\left(SQI_i, SQI_j\right) - H\right)\cdot\left(d_{Xbest}\left(SEQ_i, SEQ_j\right) - K\right)\right] > 0 \\ 0 \quad else \end{cases}$$

$$s(K) = \frac{\sum_{ij} m_{ij}}{N(N-1)}.$$

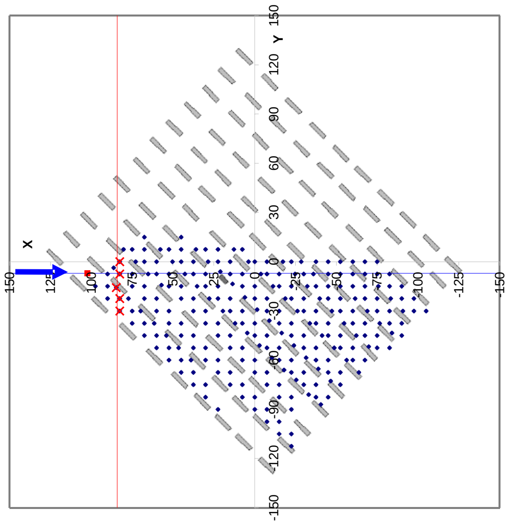
Fraction of points in the upper left or lower right quadrant

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA E INNOVACIÓN

Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Example based on MUST simulations for COST732

Array of obstacles – wind tunnel



- Point release at ground level
- Concentration measurements at H/2 (H=height of the obstacle).
- Flow measurements (velocity components,TKE).

17 simulations (different models, different users, different set-ups)

Model	*Developer*	*Users*
FINFLO	Helsinki University of Technology, Finland	Hellstein (3 sim.)
FLUENT	ANSYS (commercial code)	Franke, Goricsan (2 sim.), Santiago, Buccolieri.
M2UE	Tomsk State University, Russia, and Danish Meteorological Institute	Nuterman, Starchenko and Baklanov
MISKAM	University of Mainz, Germany	Ketzel (2 sim.), Goricsan (3 sim.)
STAR CD	CD-ADAPCO (commercial code)	Brzozwski
VADIS	University of Aveiro, Portugal	Costa and Tavares
ADREA	Environmental Research Laboratory of NCSR “Demokritos”, Greece	Efthimiou and Bartzis

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA E INNOVACIÓN

Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

To test the methdology we need a case where **both** $d_{purpose}$ and d_X can be computed.

So let assume:

$$d_{purpose}\left(SQI_i, SQI_j\right) = 2\frac{\left|\max\left(C_i(x)\right) - \max\left(C_j(x)\right)\right|}{\max\left(C_i(x)\right) + \max\left(C_j(x)\right)}$$

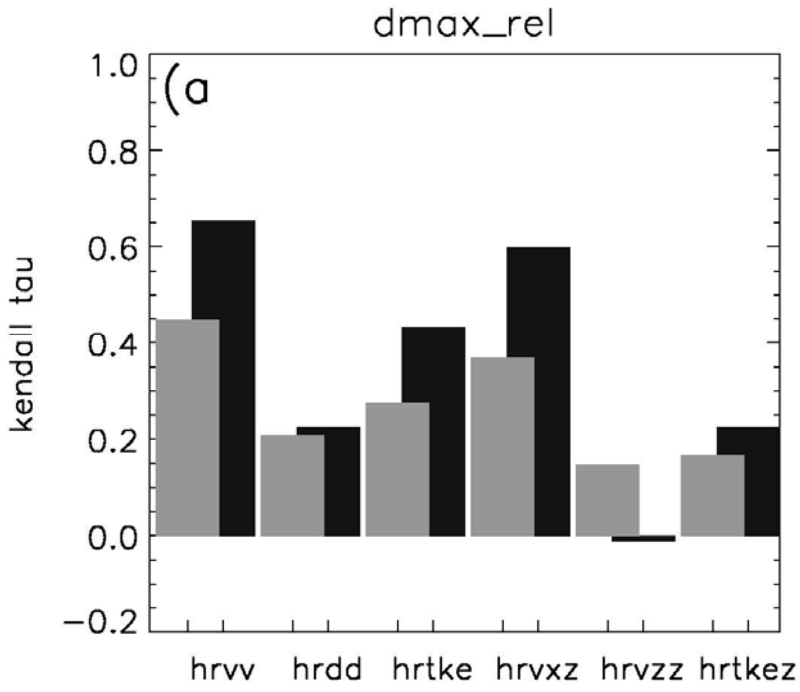
Relative difference of maximum of concentration at the measurement points.

H=0.5

Candidate d_X (not involving concentration measurements)

$d_{hrvv}\left(SEQ_i, SEQ_j\right) = 1. - HitRate\left(vect_i, vect_j\right)$	Horizontal velocity
$d_{hrdd}\left(SEQ_i, SEQ_j\right) = 1. - HitRate\left(dir_i, dir_j\right)$	Wind direction
$d_{hrvxz}\left(SEQ_i, SEQ_j\right) = 1. - HitRate\left(vx_i, vx_j\right)$	X-velocity (from profiles)
$d_{hrvzz}\left(SEQ_i, SEQ_j\right) = 1. - HitRate\left(vz_i, vz_j\right)$	Vertical velocity (from profiles)
$d_{hrtke}\left(SEQ_i, SEQ_j\right) = 1. - HitRate\left(tke_i, tke_j\right)$	TKE
$d_{hrtkez}\left(SEQ_i, SEQ_j\right) = 1. - HitRate\left(tkez_i, tkez_j\right)$	TKE from profiles

τ Kendall



simulation-to-simulation intercomparisons

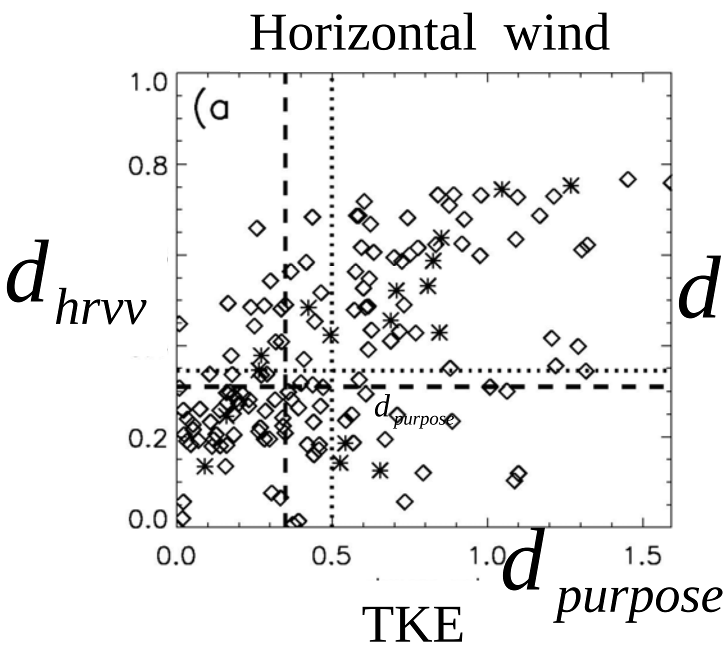
simulations-to-observations

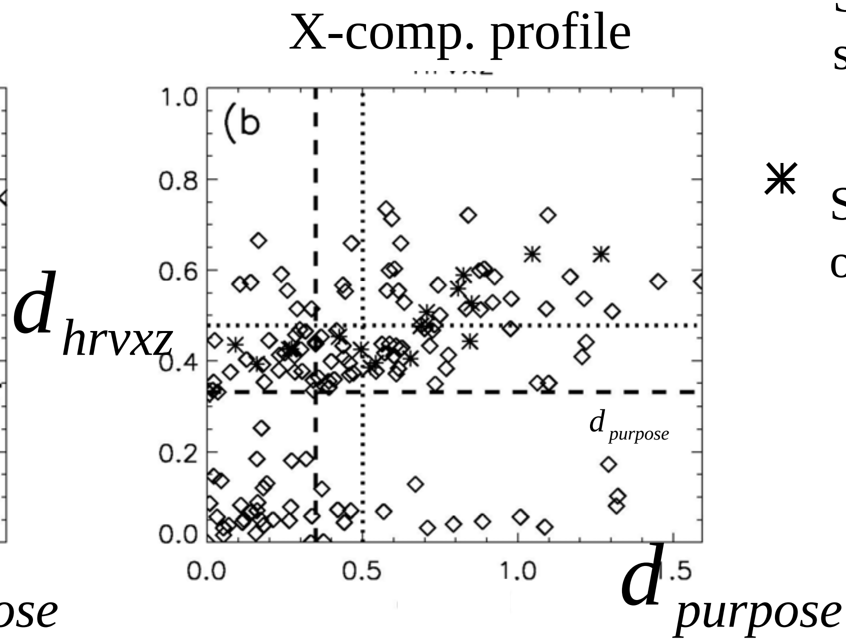
The highest the value of τ, the most similar are the rankings

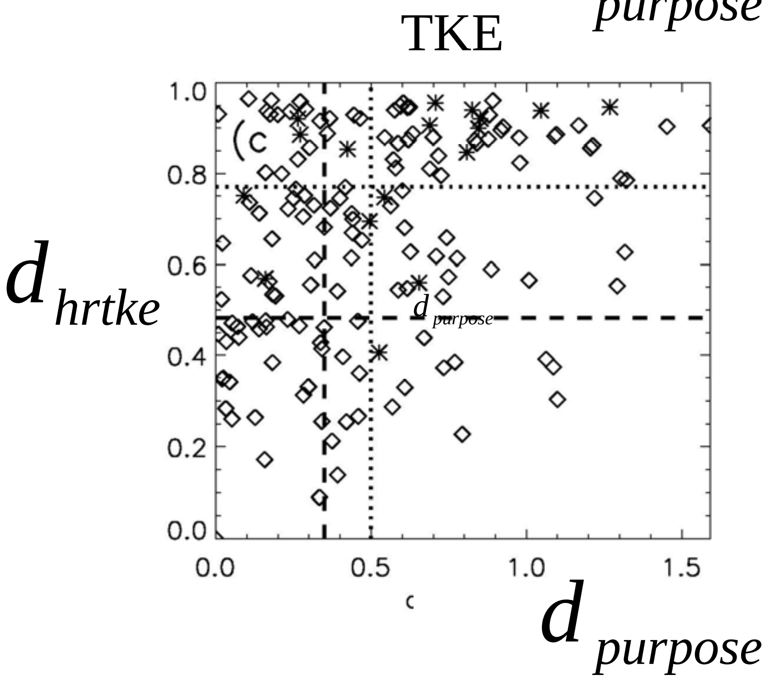
Separation value

◇ Simulation to simulation

✳ Simulation to observation







H=0.5	K_{best}	$s(K_{best})$	$s(K_{best}, Obs)$
d_{hrvv}, d_{hrtke}	0.34	0.77	0.70
d_{hrvxz}	0.77	0.70	0.65

GOBIERNO DE ESPAÑA | MINISTERIO DE CIENCIA E INNOVACIÓN | Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Conclusions

$d_{purpose}$ and the acceptance criteria H depend only on the purpose – cannot be computed directly

Conclusions

$d_{purpose}$ and the acceptance criteria H depend only on the purpose – cannot be computed directly

d_X and K_{best} (those actually used) depend, on the purpose, the specific case, the distribution and type of mesaurements available – can be computed.

Conclusions

$d_{purpose}$ and the acceptance criteria H depend only on the purpose – cannot be computed directly

Methodology

d_X and K_{best} (those actually used) depend, on the purpose, the specific case, the distribution and type of mesaurements available – can be computed.

Conclusions

$d_{purpose}$ and the acceptance criteria H depend only on the purpose – cannot be computed directly.

THANK YOU

Methodology

d_X and K_{best} (those actually used) depend, on the purpose, the specific case, the distribution and type of mesaurements available – can be computed.

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA E INNOVACIÓN

Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas