The European Commission's science and knowledge service

Joint Research Centre

Spatial Representativeness & Harmonization – Do we need a paradigm shift?

Spatial Representativeness of Air Quality Monitoring Stations

Oliver Kracht

FAIRMODE Technical Meeting, 19/21 June 2017, Athens (GR)

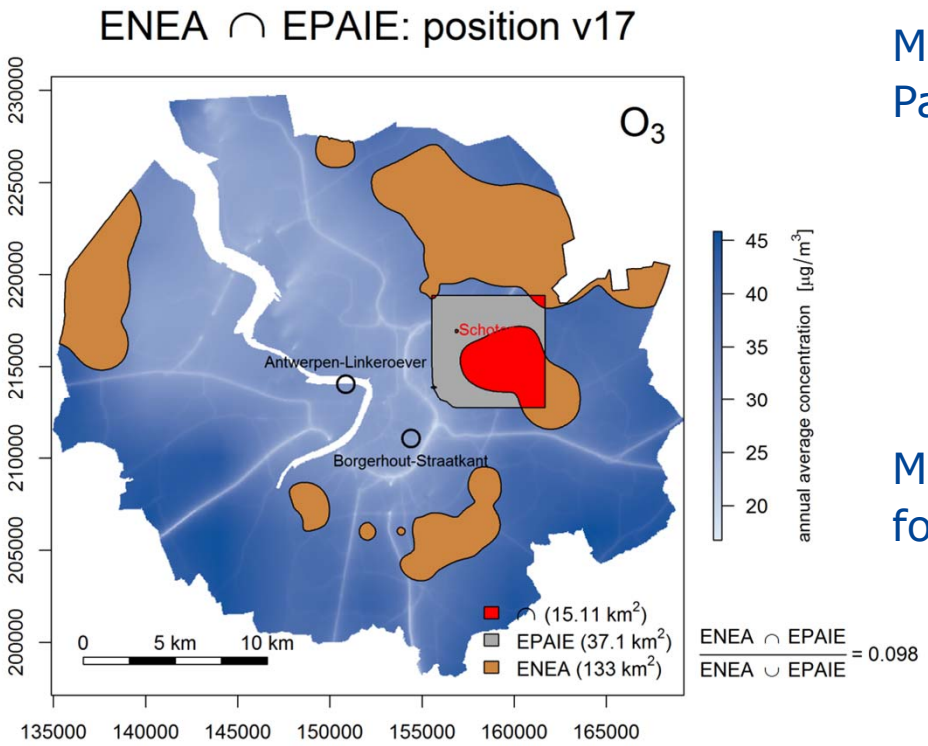
European Commission

Outline

1) Some Summary and Conclusions from the IE

2) Spatial Representativeness & Harmonization – Do we need a paradigm shift?

Mutual Comparisons of the Level of Agreement



Mutual Level of Agreement between Paired Teams

$$MLA = \frac{|SR\ Area\ 1| \cap |SR\ Area\ 1|}{|SR\ Area\ 1| \cup |SR\ Area\ 1|}$$

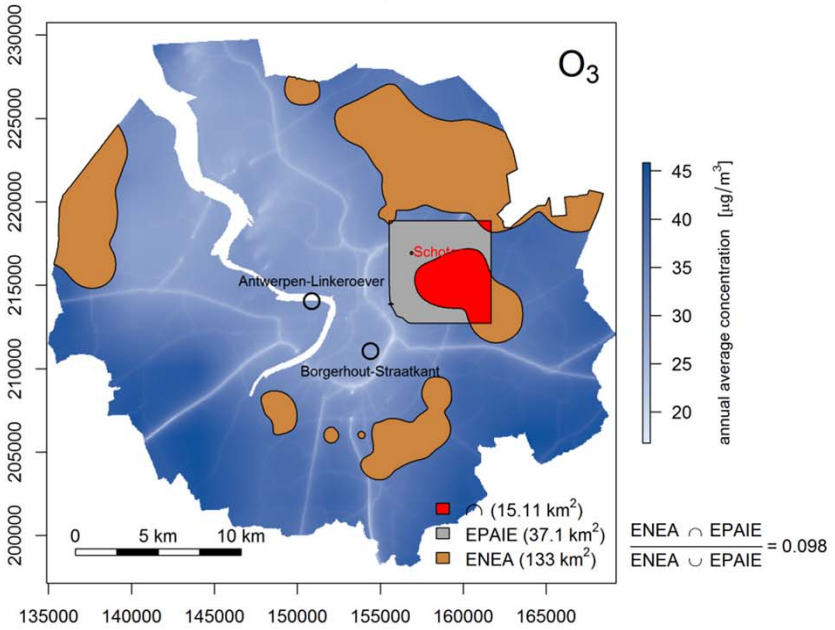
MLA ca 10% between ENEA and EPAIE for the O_3 SR-area at position v17.

Mutual Level of Agreement Indicator (MLA)

- Converges to **1 for full agreement** between Area 1 and Area 2.
- Converges to **0 for no agreement** between Area 1 and Area 2.

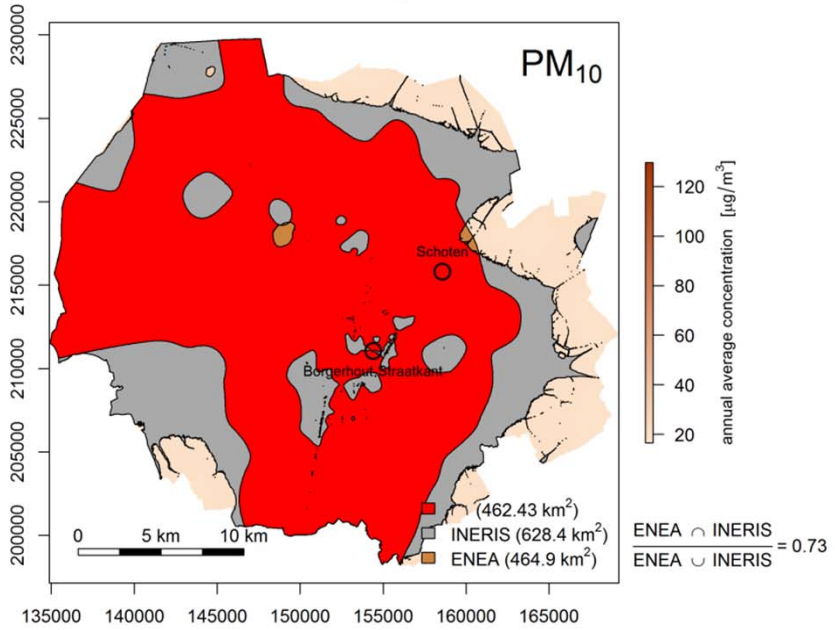
European Commission

ENEA ∩ EPAIE: position v17



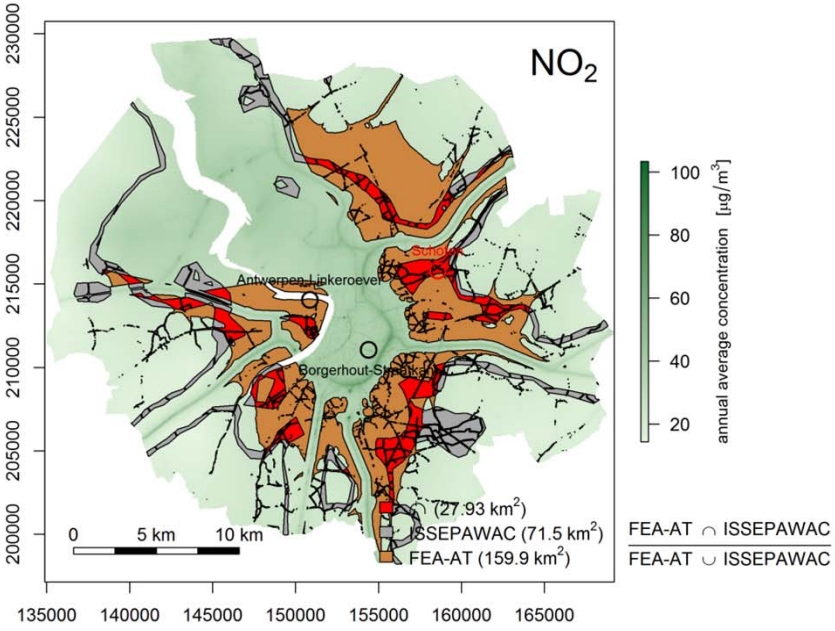
$$\frac{\text{ENEA} \cap \text{EPAIE}}{\text{ENEA} \cup \text{EPAIE}} = 0.098$$

ENEA ∩ INERIS: position v7



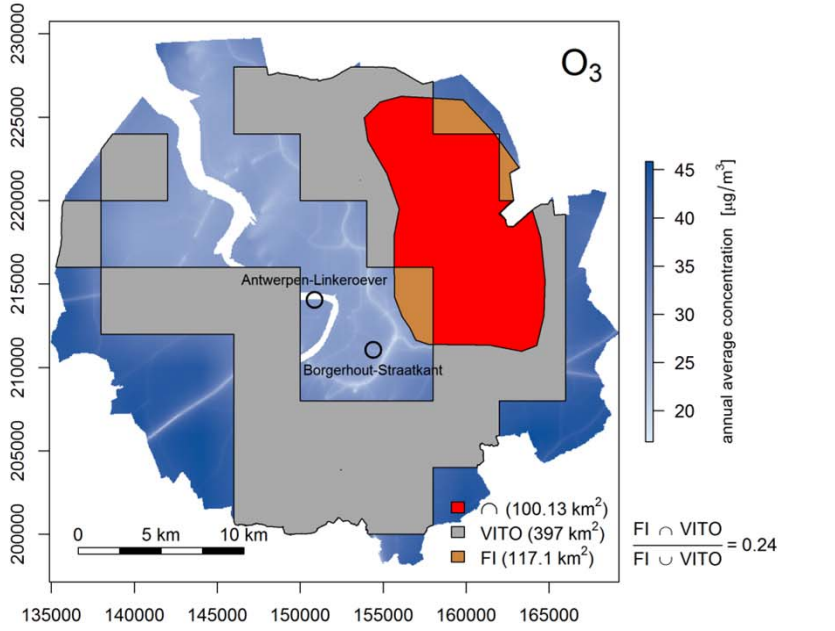
$$\frac{\text{ENEA} \cap \text{INERIS}}{\text{ENEA} \cup \text{INERIS}} = 0.73$$

FEA-AT ∩ ISSEPAWAC: position v17



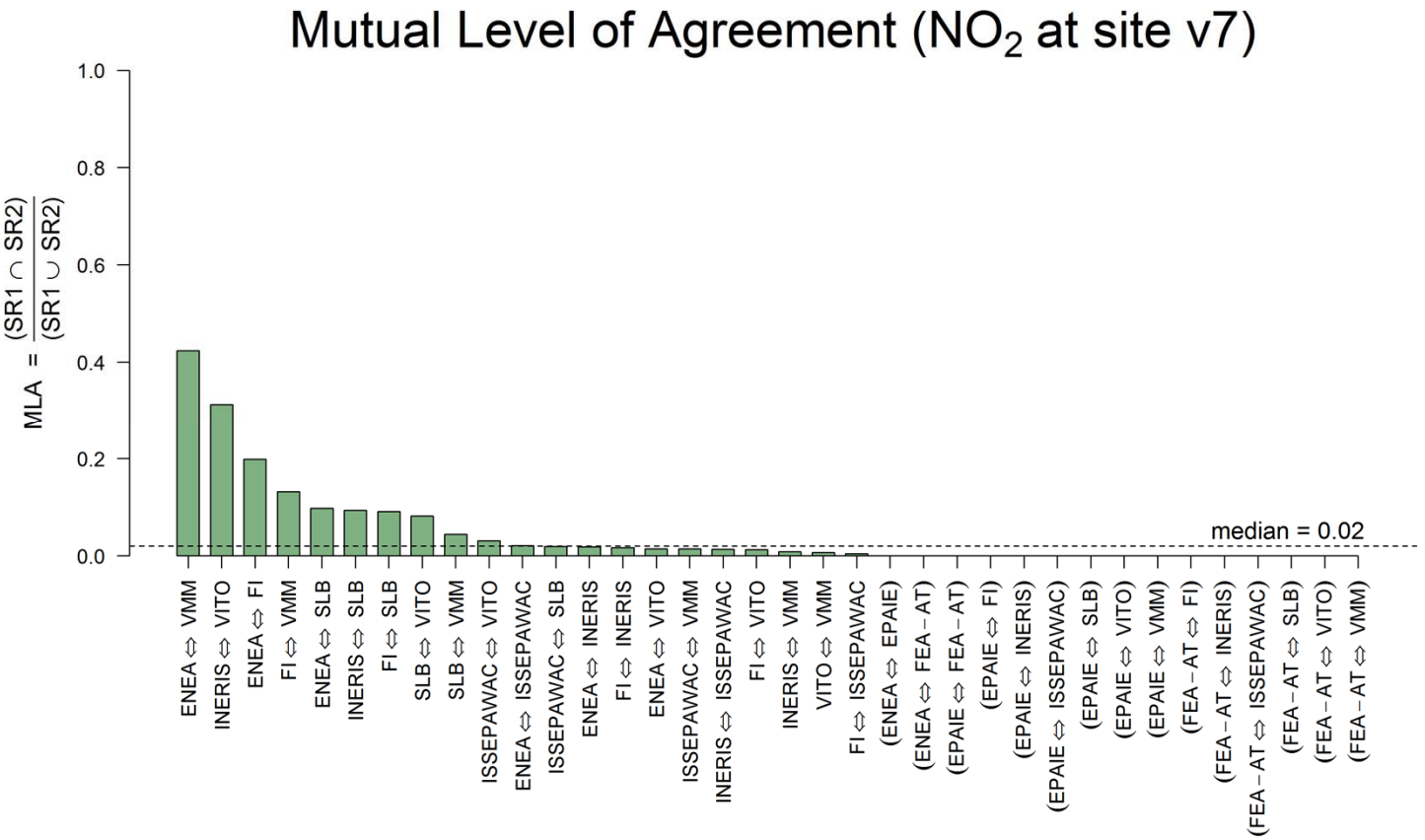
$$\frac{\text{FEA-AT} \cap \text{ISSEPAWAC}}{\text{FEA-AT} \cup \text{ISSEPAWAC}} = 0.$$

FI ∩ VITO: position v17

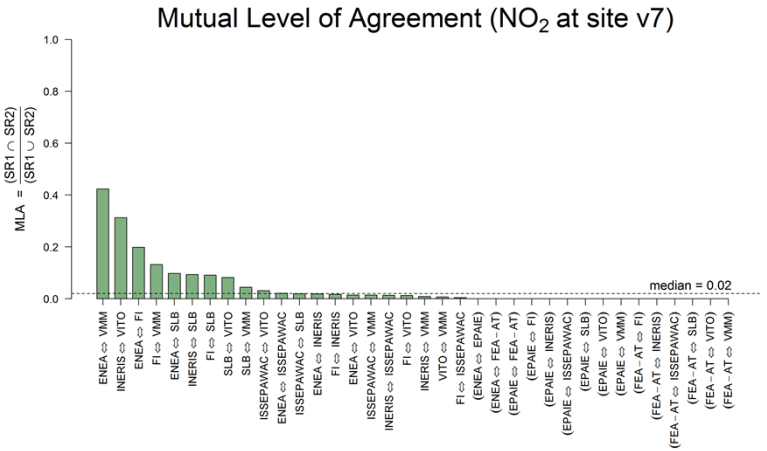


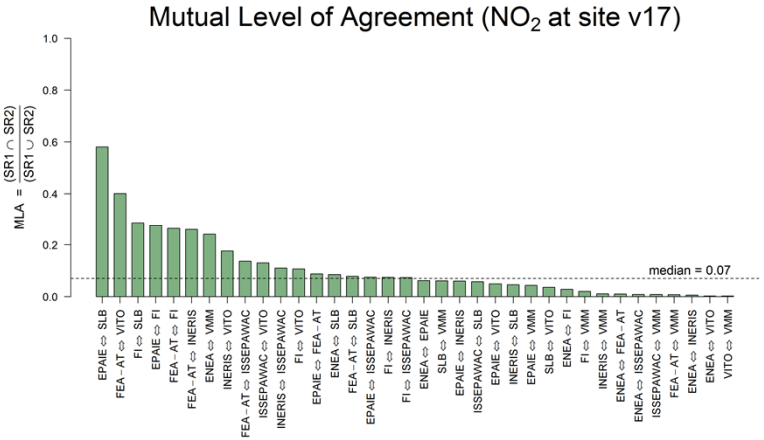
$$\frac{\text{FI} \cap \text{VITO}}{\text{FI} \cup \text{VITO}} = 0.24$$

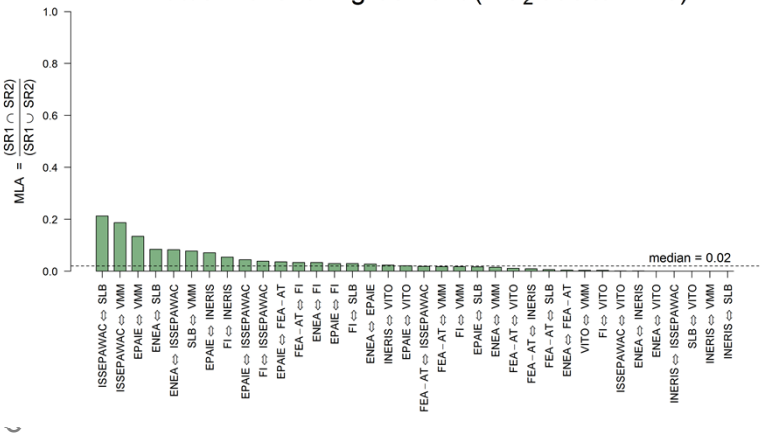
Mutual Comparisons of the Level of Agreement

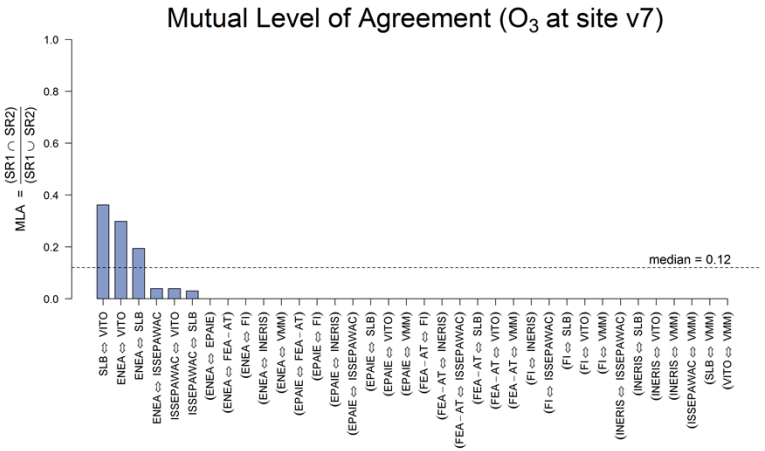


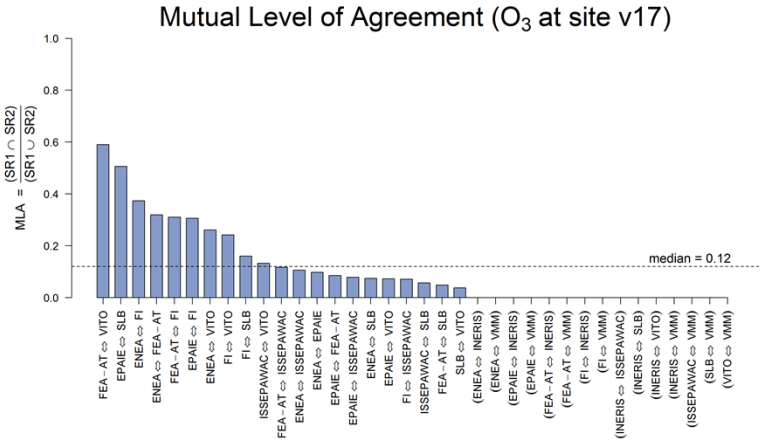
European Commission



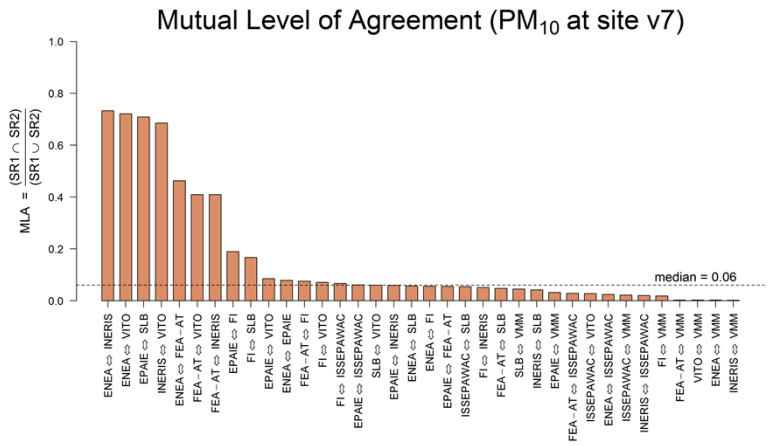


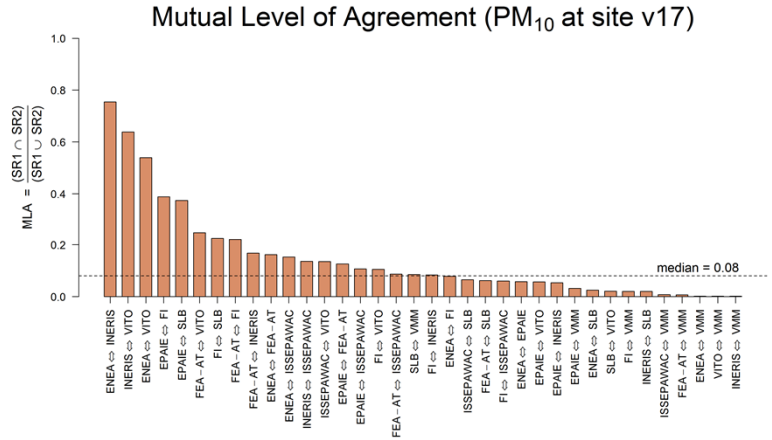


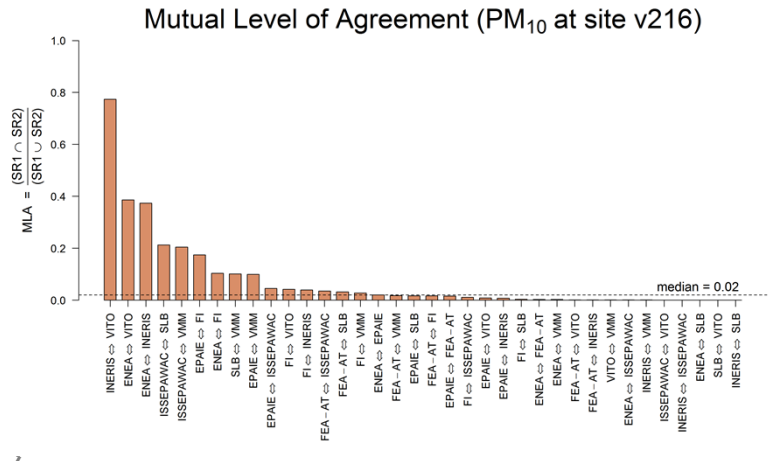




European Commission







- Please refer to the supporting files in order to inspect your particular results and combinations.

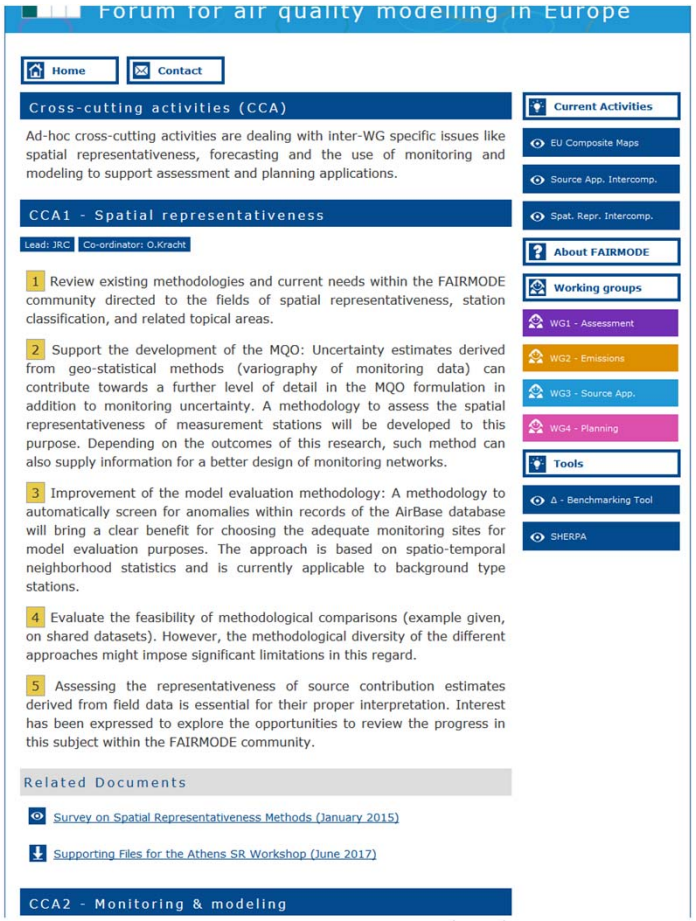
European Commission

Supporting Files

http://fairmode.jrc.ec.europa.eu/

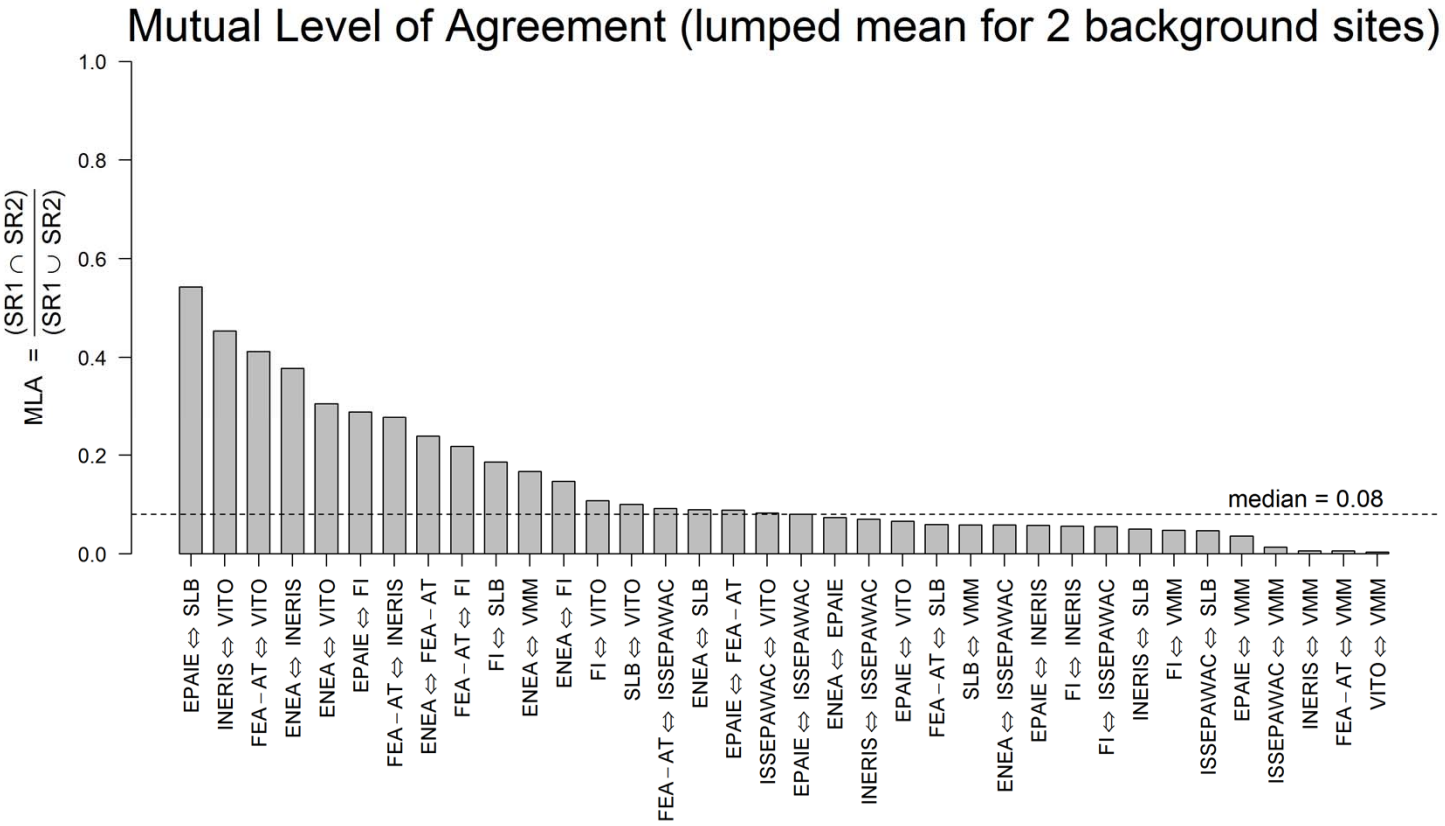
1

2



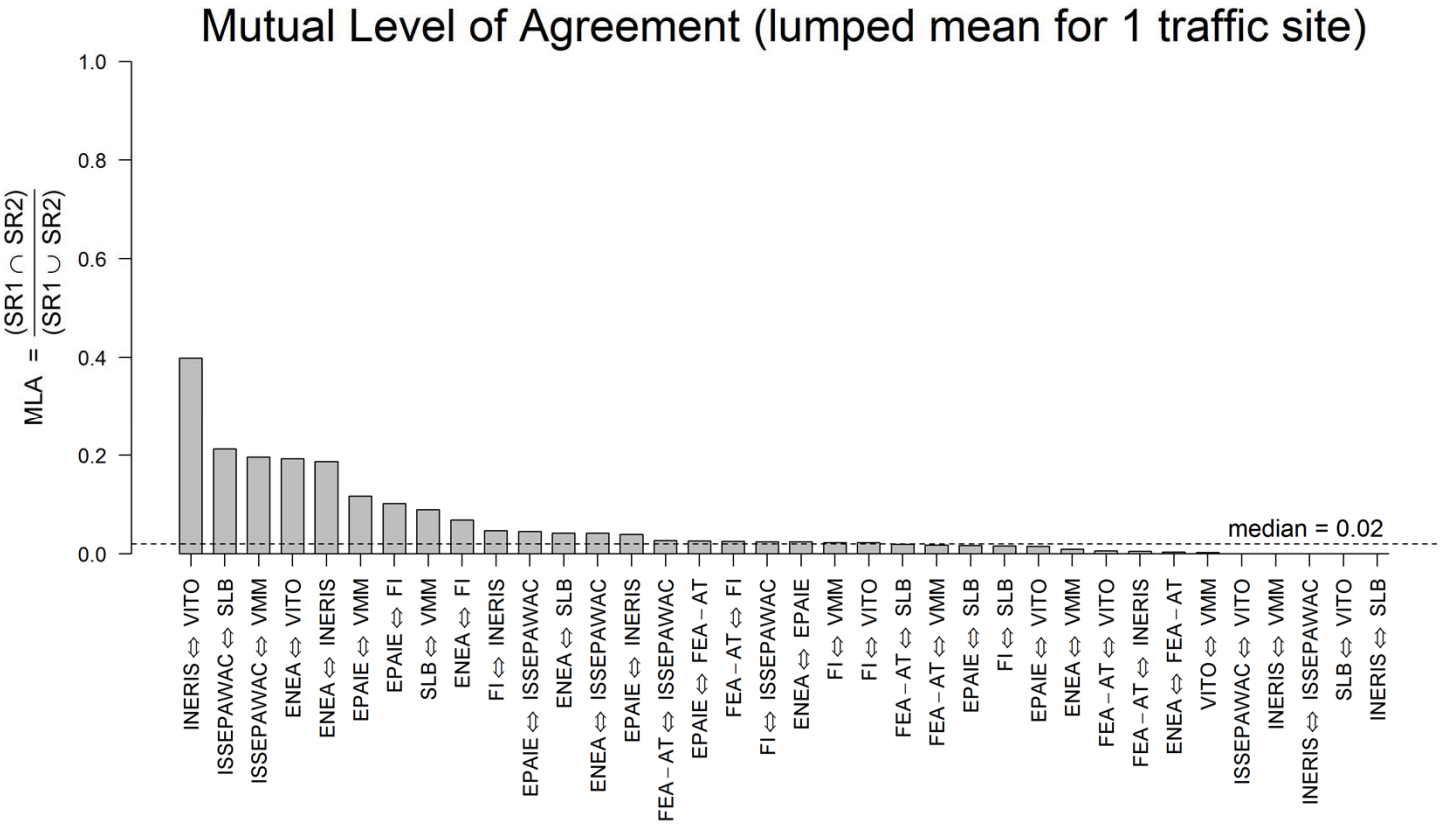
European Commission

Mutual Comparisons of the Level of Agreement



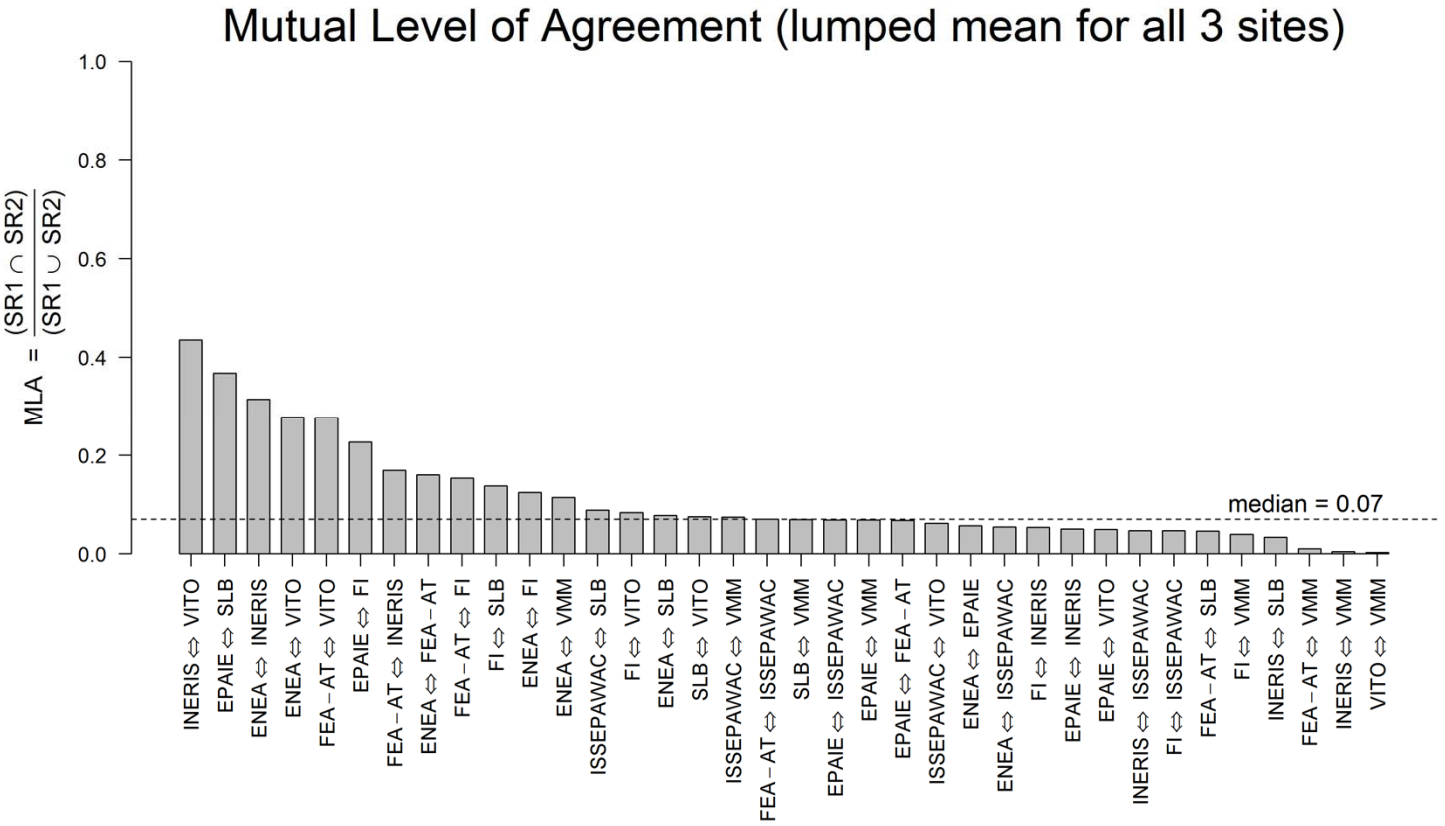
European Commission

Mutual Comparisons of the Level of Agreement

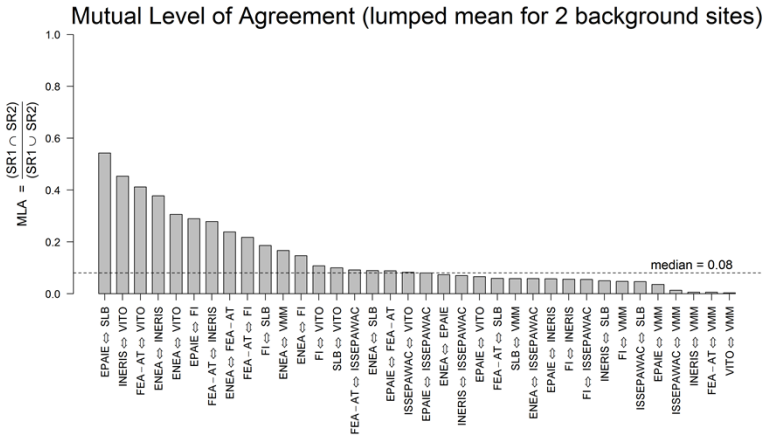


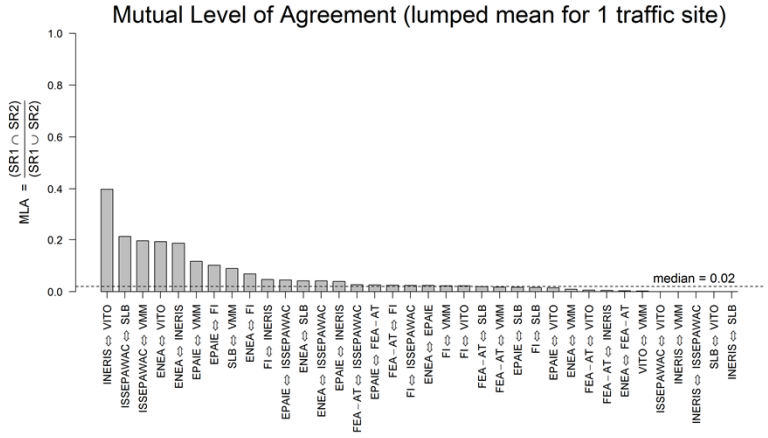
European Commission

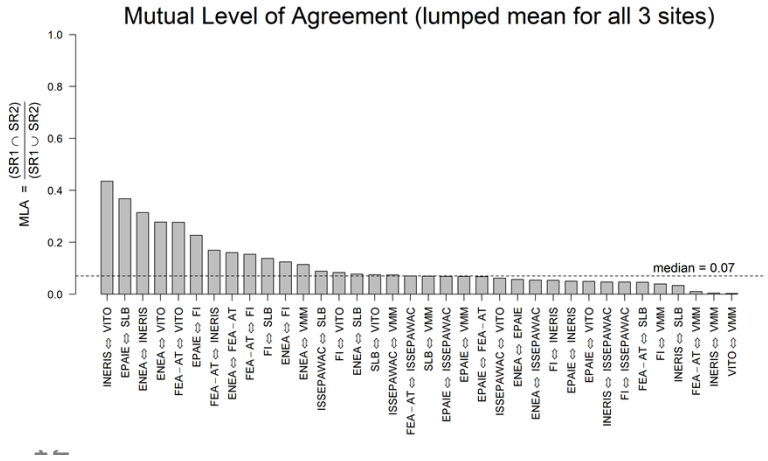
Mutual Comparisons of the Level of Agreement



European Commission







Conclusions from the Mutual Comparisons of the **Level of Agreement**:

- Values for all MLA indicators are quite low.
- Median MLAs always clearly < 10% for the whole set of comparisons.
- A deduced ranking in terms of the order of paired similarities should be taken with care and is probably not reasonably conclusive.

European Commission

Outline

1) Summary and Conclusions from the IE

2) Spatial Representativeness & Harmonization – Do we need a paradigm shift?

Spatial Representativeness & Harmonization
Do we need a paradigm shift?”

- The assessment of the spatial representativeness of air quality monitoring stations is an important subject that has substantial links to several highly topical areas, including risk assessment and population exposure, the design of monitoring networks, model development, model evaluation and data assimilation.

- The concept of spatial representativeness has been discussed intensively within FAIRMODE and AQUILA for many years (>10y). However, no well-established consensus on its definition has been identified so far.

- European directives lack a clear methodology or advise of how to evaluate the spatial representativeness of AQ monitoring stations.

- How can we make progress in the harmonization of assessment procedures of spatial representativeness?

Spatial Representativeness & Harmonization
Do we need a paradigm shift?”

Outline:

- The Objectives and drivers for Spatial Representativeness definitions
- How can we make progress towards a more harmonized quantification of SR ?
- Required elements for a clear definition of a SR characteristic
- Different purposes of estimating SR: A conflict of goals?
- Discussion: Transparency ***/*** Harmonization ***/*** Standardization ?

Spatial Representativeness & Harmonization
Do we need a paradigm shift?”

- **The Objectives and drivers for Spatial Representativeness definitions**
- How can we make progress towards a more harmonized quantification of SR ?
- Required elements for a clear definition of a SR characteristic
- Different purposes of estimating SR: A conflict of goals?
- Discussion: Transparency, Harmonization, Standardization ?

Objectives and Drivers for Spatial Representativeness Definitions

- Spatial Representativeness (SR) is not an intrinsic site feature as such.
- SR can reasonably be defined in a contextual framework only.
- What - in reality - determines the formulation of a SR definition?

Spatial Representativeness: Definitions (Baveno (IT) 2014)

Spatial Representativeness in the Literature

"Representativeness is the extent to which a set of measurements taken in a space-time domain reflects the actual conditions in the same or different spacetime domain taken on a scale appropriate for a specific application."

(Nappo et al. 1982)

[The area of representativeness is …] "… the area in which the concentration does not differ from the concentration measured at the station by more than a specified amount."

(Larssen et al. 1999)

"A monitoring station is representative of a location if the characteristic of the differences between concentrations over a specified time period at the station and at the location is less than a certain threshold value."

(Spangl et al. 2007)

Spatial Representativeness: Definitions (Baveno (IT) 2014)

Spatial Representativeness in the Literature

"A point measurement is representative of the average in a larger area (or volume) if the probability that the squared difference between point and area (volume) measurement is smaller than a certain threshold more than 90% of the time."

(Nappo et al. 1982)

A unified definition?

European Commission

Spatial Representativeness: Definitions (Baveno (IT) 2014)

Possible definitions of Spatial Representativeness

A set of spatial points X is considered the representative area of a monitoring station s_0 located at x_0 if:

$$\left| f\left(x_i\right) - f\left(x_0\right) \right| \leq \delta \quad \forall \quad x_i \in X$$

δ: threshold value (e.g., in µg/m^3)
$f(x_i)$: concentration estimated at x_i
$f(x_0)$: concentration estimated at x_0

Depending on the application, important extensions of such definition can be required to account for:

- The uncertainty of measurement for $f(x_0)$
- The uncertainty for the estimation of $f(x_i)$
- The probability of exceeding the threshold value δ within in a time series
- A combined threshold value absolute / relative

  Example: max of c(± 4 µg/m3 , ± 10%)

Objectives and Drivers for Spatial Representativeness Definitions

What - **in reality** - determines the formulation of a SR definition?

1) The implementation of a **statistical similarity criterion**

2) The **purpose and application** of the user
   - Regulatory purposes and legislation
   - Design of monitoring networks
   - Exposure assessment
   - Statistical evaluations
   - Detection of spatio-temporal outliers
   - Model calibration and model validation
   - …

3) The **methods** and **data available**
   - Might unwontedly lead into a Maslow's hammer trap ("hammer and nail metaphor", "law of the instrument")

Only SR definitions of type (1) can be useful for objective comparisons, whilst types (2) and (3) drivers of course have justifications in their own regards.

In a first step forward towards harmonization one would need to untangle (1), (2) and (3), in this way achieving transparency in reporting SR estimates.

Objectives and Drivers for Spatial Representativeness Definitions

Which are the properties to which the statistical criteria **are usually** applied?

1) The concentration of a pollutant
   - from Measurements
   - from Modeling
   - → both always requiring a certain level of interpolation
2) Emissions
3) Emission Proxies
   - Traffic
   - Population
   - Industry, land-use
   - …
4) Dispersion Conditions
   - Building geometries
   - Geometries of street canyons
   - …
5) Meteorological conditions

… this list is likely not exhaustive

Nevertheless, it is still primarily the **concentration of a pollutant** we should be interested in.

Objectives and Drivers for Spatial Representativeness Definitions

What are the properties to which the statistical criterion **should be** applied?

It is primarily the **concentration of a pollutant** we are interested in.

Nevertheless, the other properties (proxies) can have justifications in their own regards, because of:

- Unavailability of immediate concentration estimates
- Limited confidence in concentration estimates from measurements & modelling
    - In the strict sense, spatially interpolated measurements or modelling results are just as well only proxies of the real concentrations.
- Aiming for a more robust and a more holistic characterization of the SR of a monitoring site

However, **SR estimates** derived by **different principals** and based on different proxy data are **not as comparable** as we might have expected.
(one major outcome of this intercomparison exercise)

There is no straightforward procedure to equally apply the objectives of concentration based similarity criteria to other non-concentration based proxies.

Spatial Representativeness & Harmonization
Do we need a paradigm shift?”

- The Objectives and drivers for Spatial Representativeness definitions
- **How can we make progress towards a more harmonized quantification of SR ?**
- Required elements for a clear definition of a SR characteristic
- Different purposes of estimating SR: A conflict of goals?
- Discussion: Transparency, Harmonization, Standardization ?

It is hardly conceivable to make progress in harmonization without making a clear distinction between 5 different aspects:

1) The co-existence of **different** context related **SR characteristics** (maybe also to be called SR measures, or SR metrics).

   Examples:

   - the SR area for PM_{10} annual averages
   - the SR area for PM_{10} daily averages
   - the SR area for the number of PM_{10} daily averages exceeding 50 $\mu g/m^3$

   Other examples:

   - the SR area for annual averages of NO_2 emissions

   Examples of combinations:

   - the combined SR area estimate for "NO_2 annual averages" and "the number of NO_2 hourly averages exceeding 200 $\mu g/m^3$"

     (basically an intersection of two different SR estimates)

2) The **definition(s) of** these **SR characteristics**, including the specification of a **primary similarity criterion** consisting of:

   - a tolerance criterion for the deviation of concentrations (or counts, …)
   - add-on: a criterion for the level of uncertainty
   - add-on: a maximum permissible frequency of deviations

➢ We need a common agreement and understanding of SR characteristics, nomenclature and taxonomy.

How to make progress towards a More Harmonized Quantification of SR ?

3) The technical **methods for estimating** a particular SR characteristic
   - This might include the specification of one or more **secondary similarity criteria**

Example: A method for the "SR area for NO_2 annual averages " might comprise a secondary similarity criterion related to NO_2 emissions or to traffic conditions.

The primary similarity criterion however needs to remain the tolerance criterion related to NO_2 annual average concentrations.

It needs to be shown that compliance with these secondary criteria can guarantee compliance with the primary criterion.

Otherwise the results should not be named "SR area for NO_2 annual averages" , but a more correct name would then probably be "SR area for NO_2 annual average emissions".

As a side note, this brings up some important issues:

- How should we validate methods which are not immediately based on concentration fields?
- Would it be necessary and reasonable to define an order of preferences?
  - Example: Methods based on concentration fields preferred if such data are available. If concentration fields not available, then use alternative …

How to make progress towards a More Harmonized Quantification of SR ?

(…need clear distinction between 5 different aspects:)

4) The **purpose** of evaluating SR in a specific case of application

Examples:

- Regulatory purposes and legislation
- The local design of monitoring networks
- …

5) The **set of SR characteristics** required for this purpose

- A specific case of application does typically require a set of more than one suitable SR characteristics to be estimated.
- The overall aim of such a set of information might then be called "spatial representativeness characterization".
- The user / local expert / regulator / legislator could specify the set of SR characteristics required in a particular context

➢ A well structured, **transparent** and harmonized **documentation** on the **definitions** of the chosen **SR characteristics** and on the primary and the secondary **similarity criteria** applied should be an attribute feature attached to every SR estimate.

Spatial Representativeness & Harmonization
Do we need a paradigm shift?”

- The Objectives and drivers for Spatial Representativeness definitions
- How can we make progress towards a more harmonized quantification of SR ?
- **Required elements for a clear definition of a SR characteristic**
- Different purposes of estimating SR: A conflict of goals?
- Discussion: Transparency, Harmonization, Standardization ?

Elements for a clear definition of a SR characteristic

A clear **definition of a specific SR characteristic** needs to comprise:

- The **type of pollutant**
  - PM_{10}, NO_2, O_3, …
- The **scope** and the **integration time scale**

  Examples:
  - annual average concentrations
  - number of daily exceedances
  - …
- The **test statistics** applied
  - **tolerance criterion** (minimum requirement)
    - *concentration differences*
    - *differences in counts*
    - *…*
  - criterion for the level of **uncertainty** (add-on)
  - criterion for the maximum permissible **frequency of deviations** (add-on)

Elements for a clear definition of a SR characteristic

A clear **definition of a specific SR characteristic** needs to comprise:

- The **observation time scale**

  Example:

  - ten years time series of annual averages

- If applicable, **further boundary conditions** that might have been employed

  Examples:

  - Shall SR areas be contiguous?
  - Shall SR areas of a set of stations be exclusive?
  - …

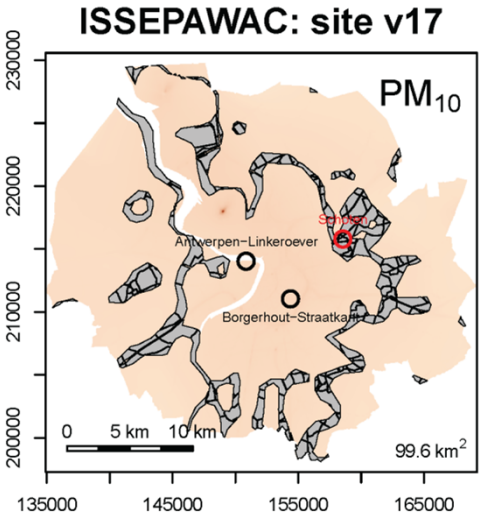
Spatial Representativeness & Harmonization
Do we need a paradigm shift?”

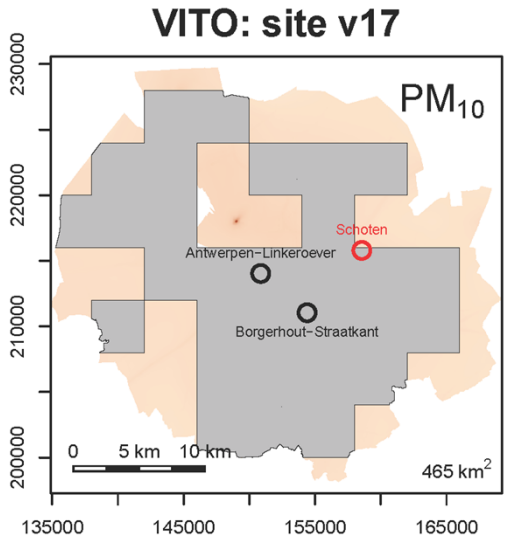
- The Objectives and drivers for Spatial Representativeness definitions
- How can we make progress towards a more harmonized quantification of SR ?
- Required elements for a clear definition of a SR characteristic
- **Different purposes of estimating SR: A conflict of goals?**
- Discussion: Transparency, Harmonization, Standardization ?

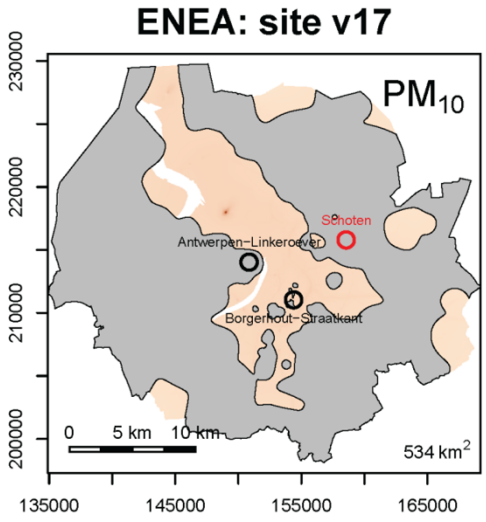
Different purposes of estimating SR: Conflict of goals?

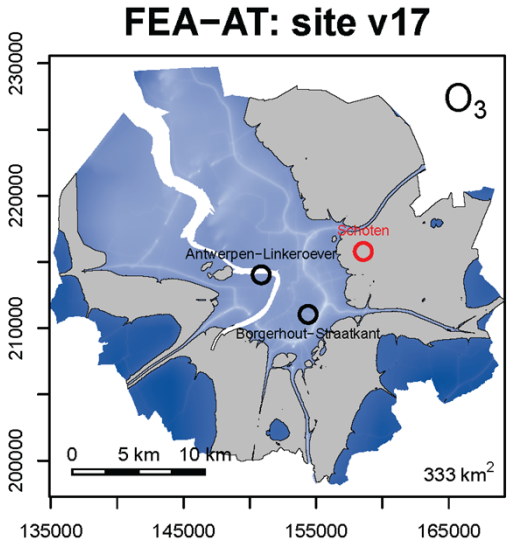
How do SR estimates performed for different purposes relate to each other?

	AQ Monitoring Example	Modelling Example
Evaluation purpose	Coverage within the monitoring network	Association between measured and modelled data
		Needed when Model grid point and monitoring station do not coincide
Similarity criterion	Criterion fit for the regulatory purpose	Similarity criterion likely more strict
		SR concentration tolerance should not be on a larger order of magnitude than the resolution of modelled concentrations







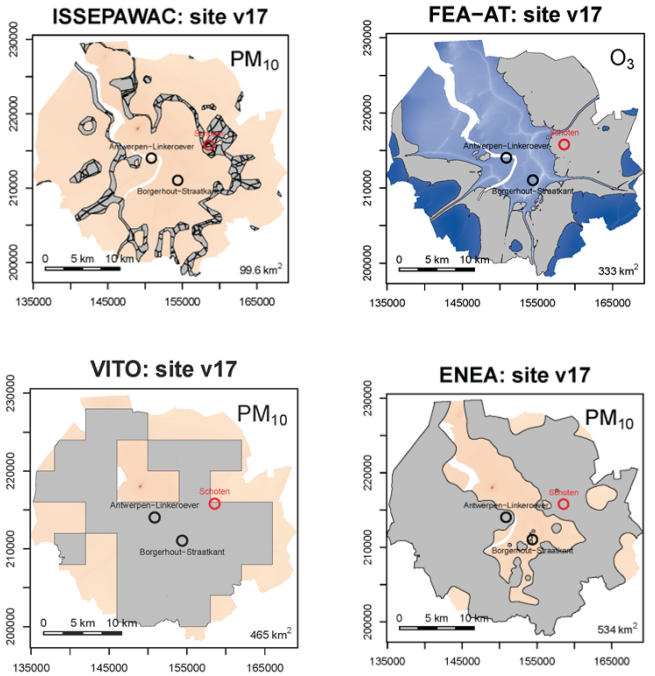


European Commission

Different purposes of estimating SR: Conflict of goals?

How do SR estimates performed for different purposes relate to each other?

	AQ Monitoring Example	Modelling Example
Evaluation purpose	Coverage within the monitoring network	Association between measured and modelled data
		Needed when Model grid point and monitoring station do not coincide
Similarity criterion	Criterion fit for the regulatory purpose	Similarity criterion likely more strict
		SR concentration tolerance should not be on a larger order of magnitude than the resolution of modelled concentrations
Possible conflict of goals, even if the SR estimate is accurate and of good quality.		



It seems rather counterintuitive to compare model data to station data when modelled data-points are located within a rather huge SR area around the corresponding station.

Can this be solved by applying more strict similarity criterion for the second evaluation purpose?

European Commission

In this example, SR is evaluated:

- in a range of the similarity criterion well suited for Purpose 1

  → Fit for purpose for common AQ Monitoring

- in a range of the similarity criterion totally unsuited for Purpose 2

  → Not fit for purpose for Model Calibration or Model Validation

… potentially even more conflicts:

- Adding a buffer could be helpful to obtain a realistic estimate of the population covered by the SR area.
- This would however not at all be helpful if using the SR estimate for model calibration & validation.

Conflicts of goals between these two purposes could exist:

- From a numerical point of view:
  - Different requirements for specifying the similarity criterion
- From a methodological point of view:
  - We need modelled data in order to precisely estimate the SR area.
  - However, we might need a precise SR area estimate in order to calibrate or validate the model.
  - Could this be a vicious cycle?

Spatial Representativeness & Harmonization
Do we need a paradigm shift?”

- The Objectives and drivers for Spatial Representativeness definitions
- How can we make progress towards a more harmonized quantification of SR ?
- Required elements for a clear definition of a SR characteristics
- Different purposes of estimating SR: A conflict of goals?
- **Discussion: Transparency, Harmonization, Standardization ?**

Discussion:
Transparency, Harmonization, Standardization ?

For this afternoon, we would particularly like to trigger the discussion on:

- The future need for **harmonization** and establishing a **common frame of reference** for
  - **SR definitions**
  - **Methods for evaluating SR**

  In this way aiming to **improve** the **transparency** and **consistency** of SR estimations, and to support a better **communication** of results.
- Is there a future need for **standardization**, too?
- Beyond standardization, should the regulators / political bodies make the use of **standards mandatory**?
- Or would it conversely be preferable to have at disposal a **set of guidelines, definitions and standardized tools, but maintaining the freedom** of choosing the most appropriate procedures for the particular case and application?