CENTRE DE RECHERCHE PUBLIC
HENRI TUDOR
www.tudor.lu

Accounting for Geospatial Uncertainties in an Energy - Air Quality Decision Support Tool

Ulrich Leopold, Christian Braun, Laurent Drouet,
Daniel S. Zachary

Resource Centre for Environmental Technologies,
Public Research Centre Henri Tudor,
Esch-sur-Alzette, Luxembourg

FOSS4G, Barcelona – September 8, 2010

Outline

Background

EU directives

for water, air and soil quality suggest today the usage of modelling techniques for sustainable environmental or risk and disaster management.

Modelling tools, such as meta-models or model chains are being developed to

- study physical, economical and social processes,
- provide decision support to stakeholders

in an integrated and optimal way.

Background

EU directives

for water, air and soil quality suggest today the usage of modelling techniques for sustainable environmental or risk and disaster management.

Modelling tools, such as meta-models or model chains are being developed to

- study physical, economical and social processes,
- provide decision support to stakeholders

in an integrated and optimal way.

Background

EU directives

for water, air and soil quality suggest today the usage of modelling techniques for sustainable environmental or risk and disaster management.

Modelling tools, such as meta-models or model chains are being developed to

- study physical, economical and social processes,
- provide decision support to stakeholders

in an integrated and optimal way.

Background

EU directives

for water, air and soil quality suggest today the usage of modelling techniques for sustainable environmental or risk and disaster management.

Modelling tools, such as meta-models or model chains are being developed to

- study physical, economical and social processes,
- provide decision support to stakeholders

in an integrated and optimal way.

Background

EU directives

for water, air and soil quality suggest today the usage of modelling techniques for sustainable environmental or risk and disaster management.

Modelling tools, such as meta-models or model chains are being developed to

- study physical, economical and social processes,
- provide decision support to stakeholders

in an integrated and optimal way.

What about model uncertainties?

Definition

EU directive for air quality allows 50% uncertainty or smaller of model outputs.

Uncertainties associated to model outputs have to be

- assessed,
- accounted for

in the modelling and decision process.

What about model uncertainties?

Definition

EU directive for air quality allows 50% uncertainty or smaller of model outputs.

Uncertainties associated to model outputs have to be

- assessed,
- accounted for

in the modelling and decision process.

What about model uncertainties?

Definition

EU directive for air quality allows 50% uncertainty or smaller of model outputs.

Uncertainties associated to model outputs have to be

- assessed,
- accounted for

in the modelling and decision process.

What about model uncertainties?

Definition

EU directive for air quality allows 50% uncertainty or smaller of model outputs.

Uncertainties associated to model outputs have to be

- assessed,
- accounted for

in the modelling and decision process.

What about model uncertainties?

Definition

EU directive for air quality allows 50% uncertainty or smaller of model outputs.

Uncertainties associated to model outputs have to be

- assessed,
- accounted for

in the modelling and decision process.

Why Spatial Uncertainty?

Definition

Much attention has been given to temporal and parameter uncertainties, **little attention to spatial uncertainties** of geospatial data and model results.

Even **less attention has been paid to accounting for and exchanging uncertainties** in modelling, visualisation and decision making.

Until now **no OGC standard** exists to easily store and exchange uncertainties of geospatial data.

Why Spatial Uncertainty?

Definition

Much attention has been given to temporal and parameter uncertainties, **little attention to spatial uncertainties** of geospatial data and model results.
Even **less attention has been paid to accounting for and exchanging uncertainties** in modelling, visualisation and decision making.
Until now **no OGC standard** exists to easily store and exchange uncertainties of geospatial data.

Why Spatial Uncertainty?

Definition

Much attention has been given to temporal and parameter uncertainties, **little attention to spatial uncertainties** of geospatial data and model results.
Even **less attention has been paid to accounting for and exchanging uncertainties** in modelling, visualisation and decision making.
Until now **no OGC standard** exists to easily store and exchange uncertainties of geospatial data.

Sources of Uncertainties

Table 1: Overview of uncertainties encountered in the modelling process.

Context uncertainty	Boundaries of the system, e.g. environmental, social circumstances.
Input uncertainty	Input data uncertainties, e.g. non-uniform landscape, limitations in land-use identification, meteorological variability
Structural and technical uncertainty	Conceptual errors due to incomplete understanding or simplifications, e.g. approximation in pollutant transport, resolution in space and time.
Parameter uncertainty	Errors related to parameter estimation, e.g. empirical constants.

Setting a specific problem

Definition

Integrated environmental assessment model chains often require spatial disaggregation and/or aggregation of input and output data.

Objective

Model spatial uncertainties associated to emission disaggregation (downscaling) in the Luxembourg Energy Air Quality assessment model to provide error information for decision making.

Approach

Decompose the emission value into its mean and standard deviation and use stochastic simulation to compute a spatially correlated error.

Setting a specific problem

Definition

Integrated environmental assessment model chains often require spatial disaggregation and/or aggregation of input and output data.

Objective

Model spatial uncertainties associated to emission disaggregation (downscaling) in the Luxembourg Energy Air Quality assessment model to provide error information for decision making.

Approach

Decompose the emission value into its mean and standard deviation and use stochastic simulation to compute a spatially correlated error.

Setting a specific problem

Definition

Integrated environmental assessment model chains often require spatial disaggregation and/or aggregation of input and output data.

Objective

Model spatial uncertainties associated to emission disaggregation (downscaling) in the Luxembourg Energy Air Quality assessment model to provide error information for decision making.

Approach

Decompose the emission value into its mean and standard deviation and use stochastic simulation to compute a spatially correlated error.

Luxembourg Energy-Air Quality model

A FOSS approach

The LEAQ model consists of the GEOspatial Emission Calculator (GEOECU) and the AsYmptotic Level Transport and Pollution (AYLTP) model coupled by an optimization routine (OBOE).

- **GEOECU** computes emissions and minimises energy costs for 5 sectors. GLPK
- **AYLTP** computes transport of air pollutants. C++, R spatial
- **OBOE** determines an optimal solution for lowest energy costs with air quality constraints. COIN
- **LEAQ utilities** consist of geospatial tools. R spatial, GRASS
- **Geospatial database** stores inputs, intermediate results and outputs. PostGIS

Luxembourg Energy-Air Quality model

A FOSS approach

The LEAQ model consists of the GEOspatial Emission Calculator (GEOECU) and the AsYmptotic Level Transport and Pollution (AYLTP) model coupled by an optimization routine (OBOE).

- **GEOECU** computes emissions and minimises energy costs for 5 sectors. GLPK
- **AYLTP** computes transport of air pollutants. C++, R spatial
- **OBOE** determines an optimal solution for lowest energy costs with air quality constraints. COIN
- **LEAQ utilities** consist of geospatial tools. R spatial, GRASS
- **Geospatial database** stores inputs, intermediate results and outputs. PostGIS

Luxembourg Energy-Air Quality model

A FOSS approach

The LEAQ model consists of the GEOspatial Emission Calculator (GEOECU) and the AsYmptotic Level Transport and Pollution (AYLTP) model coupled by an optimization routine (OBOE).

- **GEOECU** computes emissions and minimises energy costs for 5 sectors. GLPK
- **AYLTP** computes transport of air pollutants. C++, R spatial
- **OBOE** determines an optimal solution for lowest energy costs with air quality constraints. COIN
- **LEAQ utilities** consist of geospatial tools. R spatial, GRASS
- **Geospatial database** stores inputs, intermediate results and outputs. PostGIS

Luxembourg Energy-Air Quality model

A FOSS approach

The LEAQ model consists of the GEOspatial Emission Calculator (GEOECU) and the AsYmptotic Level Transport and Pollution (AYLTP) model coupled by an optimization routine (OBOE).

- **GEOECU** computes emissions and minimises energy costs for 5 sectors. GLPK
- **AYLTP** computes transport of air pollutants. C++, R spatial
- **OBOE** determines an optimal solution for lowest energy costs with air quality constraints. COIN
- **LEAQ utilities** consist of geospatial tools. R spatial, GRASS
- **Geospatial database** stores inputs, intermediate results and outputs. PostGIS

Luxembourg Energy-Air Quality model

A FOSS approach

The LEAQ model consists of the GEOspatial Emission Calculator (GEOECU) and the AsYmptotic Level Transport and Pollution (AYLTP) model coupled by an optimization routine (OBOE).

- **GEOECU** computes emissions and minimises energy costs for 5 sectors. GLPK
- **AYLTP** computes transport of air pollutants. C++, R spatial
- **OBOE** determines an optimal solution for lowest energy costs with air quality constraints. COIN
- **LEAQ utilities** consist of geospatial tools. R spatial, GRASS
- **Geospatial database** stores inputs, intermediate results and outputs. PostGIS

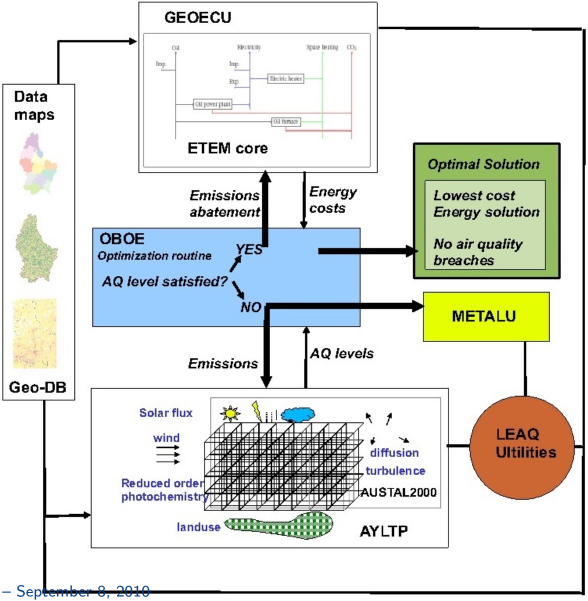
Luxembourg Energy-Air Quality model

A FOSS approach

The LEAQ model consists of the GEOspatial Emission Calculator (GEOECU) and the AsYmptotic Level Transport and Pollution (AYLTP) model coupled by an optimization routine (OBOE).

- **GEOECU** computes emissions and minimises energy costs for 5 sectors. GLPK
- **AYLTP** computes transport of air pollutants. C++, R spatial
- **OBOE** determines an optimal solution for lowest energy costs with air quality constraints. COIN
- **LEAQ utilities** consist of geospatial tools. R spatial, GRASS
- **Geospatial database** stores inputs, intermediate results and outputs. PostGIS

Luxembourg Energy-Air Quality model



The LEAQ model in chain mode

Energy demand/supply (aggregation)

Energy model

Sectoral emissions

Emission allocation

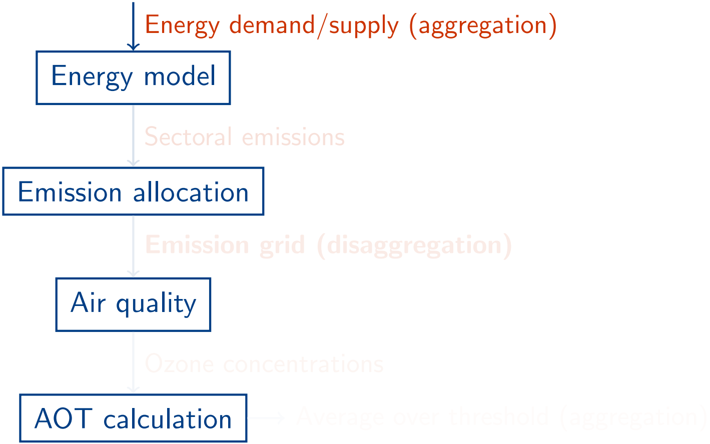
Emission grid (disaggregation)

Air quality

Ozone concentrations

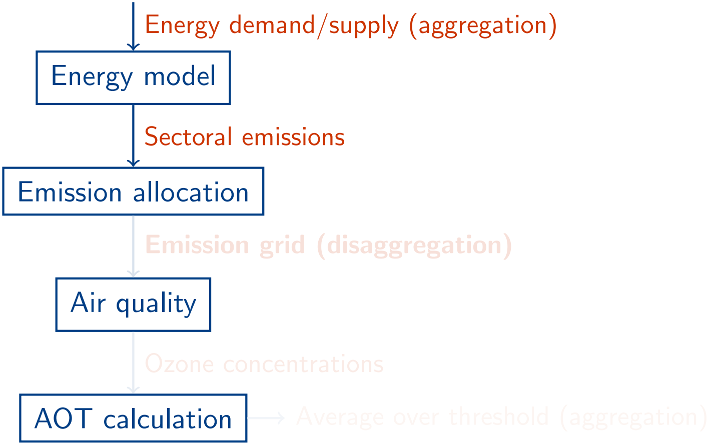
AOT calculation ⟶ Average over threshold (aggregation)

The LEAQ model in chain mode

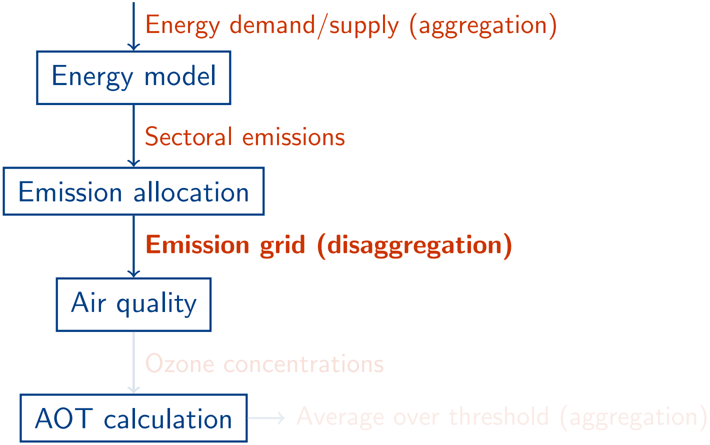


CENTRE DE RECHERCHE PUBLIC HENRI TUDOR www.tudor.lu

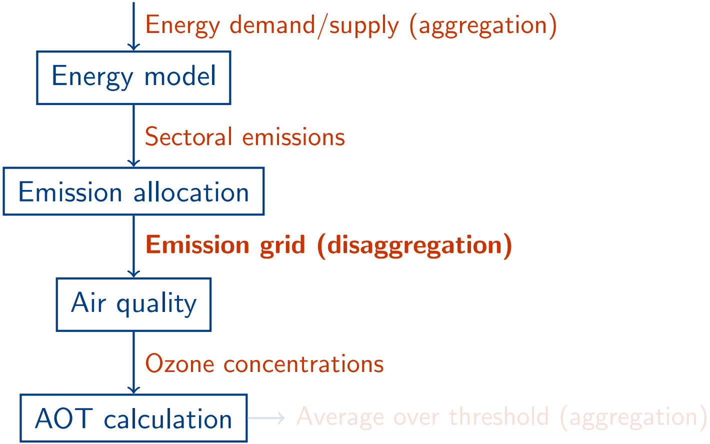
The LEAQ model in chain mode



The LEAQ model in chain mode

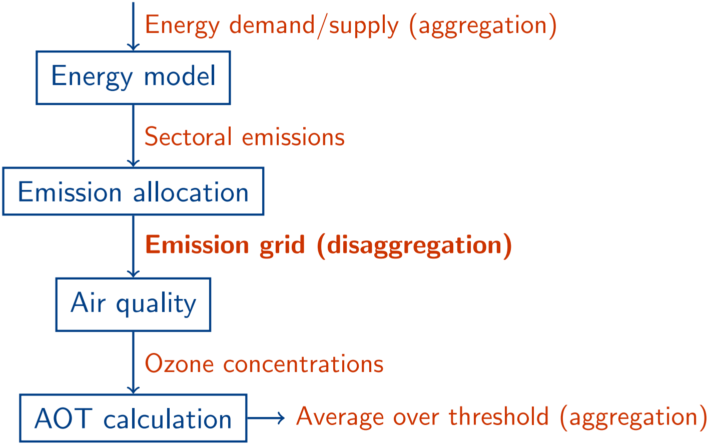


The LEAQ model in chain mode



CENTRE DE RECHERCHE PUBLIC HENRI TUDOR www.tudor.lu

The LEAQ model in chain mode



Steps in emission allocation

1. Compute global emission values per sector.
2. Distribute global mean emission values across sector grid.
3. Compute global standard deviation per sector and allocate to sector grid.
4. Build spatial variogram model based on expert knowledge.
5. Generate 100 realisations for local error using unconditional sequential Gaussian block simulation with Latin Hypercube Sampling based on assumed mean and variogram.
6. Compute local standard deviation from local error and global standard deviation.
7. Compute local emissions from global mean and local standard deviation.
8. Compute statistics, i.e. mean, standard deviation, confidence intervals.

Spatial disaggregation

For each grid cell and sector we decompose the emission value into its mean, standard deviation and spatial error

$$e(x_i) = \mu_{e_s} + \sigma_{e_s} \times \epsilon_e(x_i)$$

$$\sigma_{e_s} = \alpha_s \times \mu_{e_s} \qquad \alpha \in [0,1]$$

$$\epsilon_e(x_i) = \mu_\epsilon + \eta_\epsilon(x_i) \qquad \text{with } \gamma_\eta$$

Estimated global variation coefficient α

The following variation coefficients were chosen for α.

Sector	α
Residential	0.6
Industrial	0.5
Agriculture	0.1
Forest	0.1
Motorways	0.1
National roads	0.3
Municipal roads	0.5

Geostatistical model for spatial error

Definition

For ϵ_e we assume stationarity with known $\mu_\epsilon = 0$ and known variogram

$$\gamma_\eta(h) = E[(Z(x) - Z(y))^2]$$

The variogram model consists of an exponential structure

$$\gamma(h) = (s - n)(1 - \exp(-h/(ra))) + n1_{(0,\infty)}(h)$$

and two spherical structures

$$\gamma(h) = (s - n)\left(\left(\frac{3h}{2r} - \frac{h^3}{2r^3}\right)1_{(0,r)}(h) + 1_{(r,\infty)}(h)\right) + n1_{(0,\infty)}(h)$$

with n being the nugget, s the sill and r the range.

Geostatistical model for spatial error

Definition

For ϵ_e we assume stationarity with known $\mu_\epsilon = 0$ and known variogram

$$\gamma_\eta(h) = E[(Z(x) - Z(y))^2]$$

The variogram model consists of an exponential structure

$$\gamma(h) = (s-n)(1-\exp(-h/(ra))) + n1_{(0,\infty)}(h)$$

and two spherical structures

$$\gamma(h) = (s-n)\left(\left(\frac{3h}{2r} - \frac{h^3}{2r^3}\right)1_{(0,r)}(h) + 1_{(r,\infty)}(h)\right) + n1_{(0,\infty)}(h)$$

with n being the nugget, s the sill and r the range.

Geostatistical model for spatial error

Definition

For ϵ_e we assume stationarity with known $\mu_\epsilon = 0$ and known variogram

$$\gamma_\eta(h) = E[(Z(x) - Z(y))^2]$$

The variogram model consists of an exponential structure

$$\gamma(h) = (s - n)(1 - \exp(-h/(ra))) + n1_{(0,\infty)}(h)$$

and two spherical structures

$$\gamma(h) = (s - n)\left(\left(\frac{3h}{2r} - \frac{h^3}{2r^3}\right)1_{(0,r)}(h) + 1_{(r,\infty)}(h)\right) + n1_{(0,\infty)}(h)$$

with n being the nugget, s the sill and r the range.

Geostatistical model for spatial error

Definition

For ϵ_e we assume stationarity with known $\mu_\epsilon = 0$ and known variogram

$$\gamma_\eta(h) = E[(Z(x) - Z(y))^2]$$

The variogram model consists of an exponential structure

$$\gamma(h) = (s - n)(1 - \exp(-h/(ra))) + n1_{(0,\infty)}(h)$$

and two spherical structures

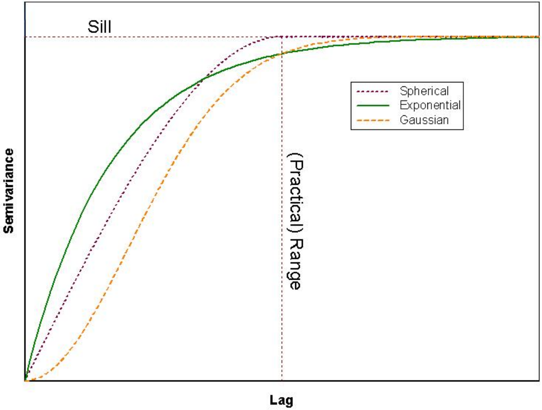
$$\gamma(h) = (s - n)\left(\left(\frac{3h}{2r} - \frac{h^3}{2r^3}\right)1_{(0,r)}(h) + 1_{(r,\infty)}(h)\right) + n1_{(0,\infty)}(h)$$

with n being the nugget, s the sill and r the range.

The Semivariogram

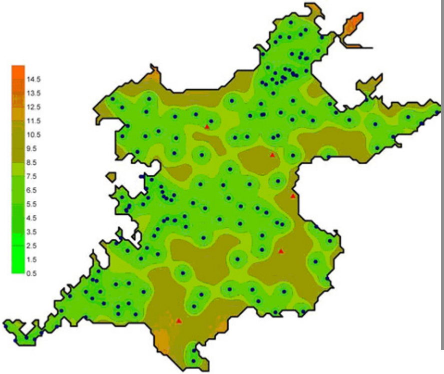
Semivariogram:

- Nugget = intercept
- Sill = total variance
- Range = Lag distance where sill is reached
- Shape = Spherical, Exponential, Gaussian, ...

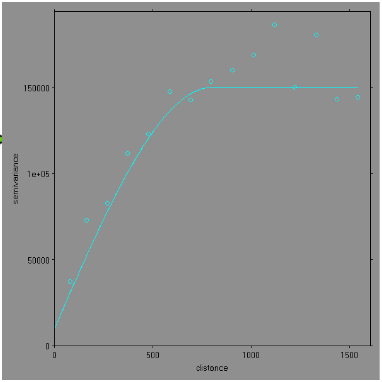


Application in interpolation (Kriging)

Sampling locations and interpolation



Semivariogram and model



Variogram parameters chosen

Definition

The final variogram model has the following components and parameters:

$$\gamma_\epsilon(h) = \gamma_{\epsilon_1} + \gamma_{\epsilon_2} + \gamma_{\epsilon_3}$$

$$\begin{aligned}
\gamma_{\epsilon_1} &:= \{n = 0.1, s_1 = 0.3, r_1 = 100, m = \text{Exp}\} \\
\gamma_{\epsilon_2} &:= \{s_2 = 0.3, r_2 = 1000, m = \text{Sph}\} \\
\gamma_{\epsilon_3} &:= \{s_3 = 0.2, r_3 = 5000, m = \text{Sph}\}
\end{aligned}$$

The above model was used in 100 unconditional sequential Gaussian simulation runs using Latin Hypercube Sampling with $\mu_\epsilon = 0$.

Variogram parameters chosen

Definition

The final variogram model has the following components and parameters:

$$\gamma_\epsilon(h) = \gamma_{\epsilon_1} + \gamma_{\epsilon_2} + \gamma_{\epsilon_3}$$

$$\begin{aligned}
\gamma_{\epsilon_1} &:= \{n = 0.1, s_1 = 0.3, r_1 = 100, m = \text{Exp}\} \\
\gamma_{\epsilon_2} &:= \{s_2 = 0.3, r_2 = 1000, m = \text{Sph}\} \\
\gamma_{\epsilon_3} &:= \{s_3 = 0.2, r_3 = 5000, m = \text{Sph}\}
\end{aligned}$$

The above model was used in 100 unconditional sequential Gaussian simulation runs using Latin Hypercube Sampling with $\mu_\epsilon = 0$.

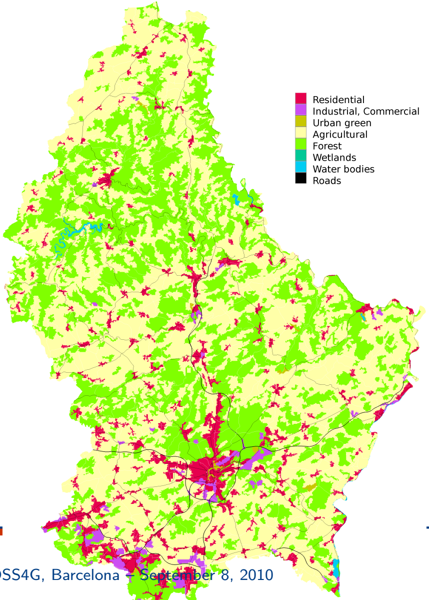
Sequential Gaussian simulation

Definition

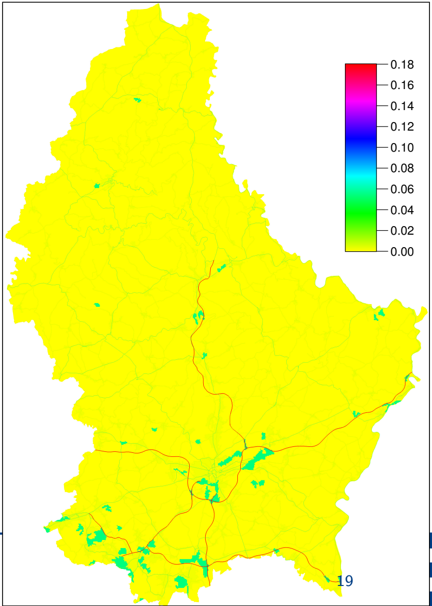
Sequential Gaussian simulation samples randomly a value from a normal Gaussian distribution with a known mean μ and a known semivariogram model γ. Each simulation iteration is called a realisation. After n realisations we can compute summary statistics at each point in space, i.e. mean, standard deviation, median, ... The mean of n realisations equals the kriged mean value.

Allocated initial CO_2 emissions [kt/y]

Sector map

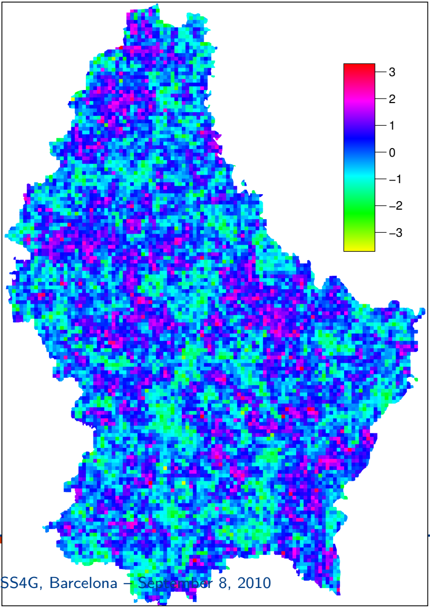


Mean CO_2 emissions

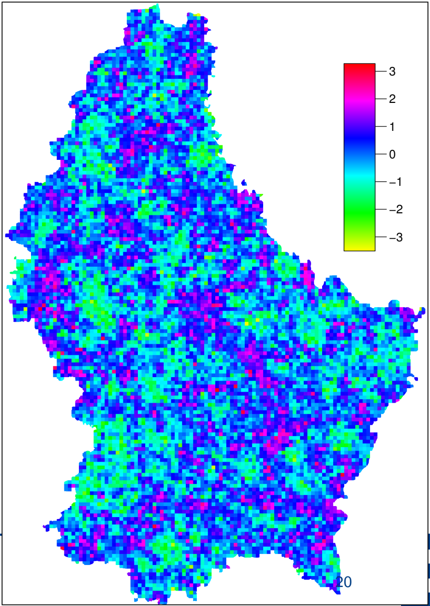


Simulated spatial Gaussian error ϵ

Realisation 1

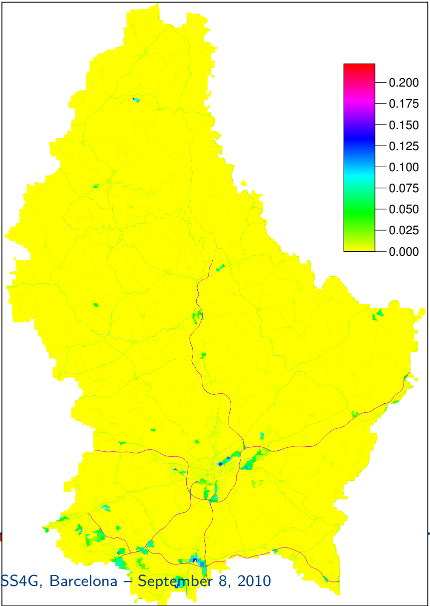


Realisation 5

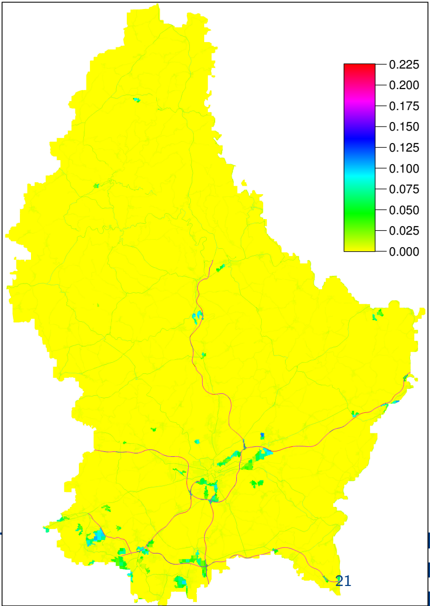


Disaggregated total emissions [kt/y]

Local emissions realisation 1

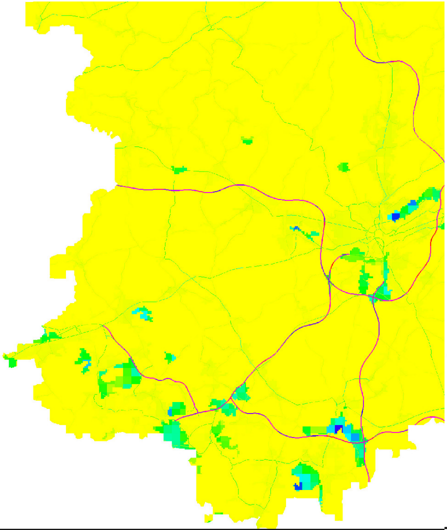


Local emissions realisation 5

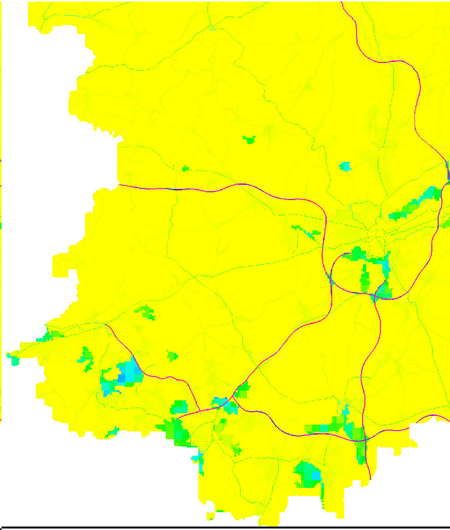


Disaggregated total emissions (zoom)

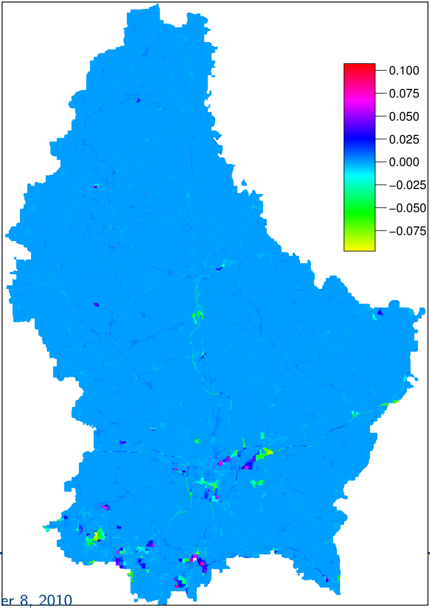
Local emissions realisation 1



Local emissions realisation 5



Difference realisations 1 and 5 [kt/y]



Discussion

- The presented methodology seems to work.
- We can now compute various statistics and use them in the model chain, e.g. for error propagation, robust optimisation or decision making.
- Assumptions made need to be verified, i.e. α, semivariogram model γ, Gaussian distribution of error.
- Validation with emission observations, experts and further analysis of land use/sector patterns would help here.
- Furthermore, we might want to include population as a trend to catch better variation at dense populated areas and industry.

But how to make use of uncertainty information in modelling or decision making?

Discussion

- The presented methodology seems to work.
- We can now compute various statistics and use them in the model chain, e.g. for error propagation, robust optimisation or decision making.
- Assumptions made need to be verified, i.e. α, semivariogram model γ, Gaussian distribution of error.
- Validation with emission observations, experts and further analysis of land use/sector patterns would help here.
- Furthermore, we might want to include population as a trend to catch better variation at dense populated areas and industry.

But how to make use of uncertainty information in modelling or decision making?

Discussion continued ...

- Geospatial models get more complex and implemented in complex spatial software and data infrastructures.
- In the scientific field we can account for uncertainties, but we need to make it more accessible for the non-expert users, i.e. administrations using web processing services.
- UncertML has been proposed as an uncertainty standard (www.uncertml.org).
- What we need now is to implement an uncertainty standard to provide storage, exchange and visualisation of uncertainties via web based services as geospatial information gets more and more complex.
- There is a first serious attempt to make UncertML accessible via web based integration UncertWeb (www.uncertweb.org).

CENTRE DE RECHERCHE PUBLIC HENRI TUDOR www.tudor.lu

Discussion continued ...

- Geospatial models get more complex and implemented in complex spatial software and data infrastructures.
- In the scientific field we can account for uncertainties, but we need to make it more accessible for the non-expert users, i.e. administrations using web processing services.
- UncertML has been proposed as an uncertainty standard (www.uncertml.org).
- What we need now is to implement an uncertainty standard to provide storage, exchange and visualisation of uncertainties via web based services as geospatial information gets more and more complex.
- There is a first serious attempt to make UncertML accessible via web based integration UncertWeb (www.uncertweb.org).

Discussion continued ...

- Geospatial models get more complex and implemented in complex spatial software and data infrastructures.
- In the scientific field we can account for uncertainties, but we need to make it more accessible for the non-expert users, i.e. administrations using web processing services.
- UncertML has been proposed as an uncertainty standard (`www.uncertml.org`).
- What we need now is to implement an uncertainty standard to provide storage, exchange and visualisation of uncertainties via web based services as geospatial information gets more and more complex.
- There is a first serious attempt to make UncertML accessible via web based integration UncertWeb (`www.uncertweb.org`).

Discussion continued ...

- Geospatial models get more complex and implemented in complex spatial software and data infrastructures.
- In the scientific field we can account for uncertainties, but we need to make it more accessible for the non-expert users, i.e. administrations using web processing services.
- UncertML has been proposed as an uncertainty standard (`www.uncertml.org`).
- What we need now is to implement an uncertainty standard to provide storage, exchange and visualisation of uncertainties via web based services as geospatial information gets more and more complex.
- There is a first serious attempt to make UncertML accessible via web based integration UncertWeb (`www.uncertweb.org`).

Discussion continued ...

- Geospatial models get more complex and implemented in complex spatial software and data infrastructures.
- In the scientific field we can account for uncertainties, but we need to make it more accessible for the non-expert users, i.e. administrations using web processing services.
- UncertML has been proposed as an uncertainty standard (`www.uncertml.org`).
- What we need now is to implement an uncertainty standard to provide storage, exchange and visualisation of uncertainties via web based services as geospatial information gets more and more complex.
- There is a first serious attempt to make UncertML accessible via web based integration UncertWeb (`www.uncertweb.org`).

Conclusions

- We have to account for uncertainties
  - to understand results from complex geospatial modelling infrastructures,
  - to support and improve decision making.
- We need to make uncertainties accessible and understandable for the non-expert users, such as policy makers, decision makers.
- Uncertainties need to be integrated in an exchangeable way into web based processing services for exchange in cascading model chains to account for uncertainties.
- First prototypes of uncertainty engines could be developed in R and interfaced via a web processing service.

CENTRE DE RECHERCHE PUBLIC HENRI TUDOR www.tudor.lu

Conclusions

- We have to account for uncertainties
  - to understand results from complex geospatial modelling infrastructures,
  - to support and improve decision making.
- We need to make uncertainties accessible and understandable for the non-expert users, such as policy makers, decision makers.
- Uncertainties need to be integrated in an exchangeable way into web based processing services for exchange in cascading model chains to account for uncertainties.
- First prototypes of uncertainty engines could be developed in R and interfaced via a web processing service.

Conclusions

- We have to account for uncertainties
  - to understand results from complex geospatial modelling infrastructures,
  - to support and improve decision making.
- We need to make uncertainties accessible and understandable for the non-expert users, such as policy makers, decision makers.
- Uncertainties need to be integrated in an exchangeable way into web based processing services for exchange in cascading model chains to account for uncertainties.
- First prototypes of uncertainty engines could be developed in R and interfaced via a web processing service.

Conclusions

- We have to account for uncertainties
  - to understand results from complex geospatial modelling infrastructures,
  - to support and improve decision making.
- We need to make uncertainties accessible and understandable for the non-expert users, such as policy makers, decision makers.
- Uncertainties need to be integrated in an exchangeable way into web based processing services for exchange in cascading model chains to account for uncertainties.
- First prototypes of uncertainty engines could be developed in R and interfaced via a web processing service.

Conclusions

- We have to account for uncertainties
  - to understand results from complex geospatial modelling infrastructures,
  - to support and improve decision making.
- We need to make uncertainties accessible and understandable for the non-expert users, such as policy makers, decision makers.
- Uncertainties need to be integrated in an exchangeable way into web based processing services for exchange in cascading model chains to account for uncertainties.
- First prototypes of uncertainty engines could be developed in R and interfaced via a web processing service.

Conclusions

- We have to account for uncertainties
  - to understand results from complex geospatial modelling infrastructures,
  - to support and improve decision making.
- We need to make uncertainties accessible and understandable for the non-expert users, such as policy makers, decision makers.
- Uncertainties need to be integrated in an exchangeable way into web based processing services for exchange in cascading model chains to account for uncertainties.
- First prototypes of uncertainty engines could be developed in R and interfaced via a web processing service.

Conclusions

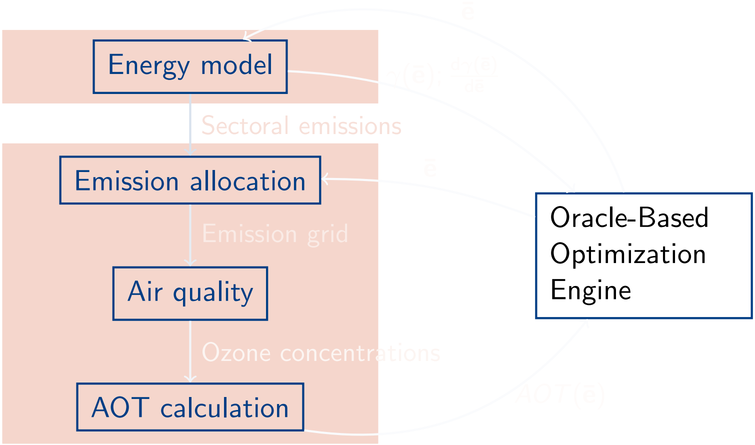
- We have to account for uncertainties
  - to understand results from complex geospatial modelling infrastructures,
  - to support and improve decision making.
- We need to make uncertainties accessible and understandable for the non-expert users, such as policy makers, decision makers.
- Uncertainties need to be integrated in an exchangeable way into web based processing services for exchange in cascading model chains to account for uncertainties.
- First prototypes of uncertainty engines could be developed in R and interfaced via a web processing service.

Thank you for your attention!

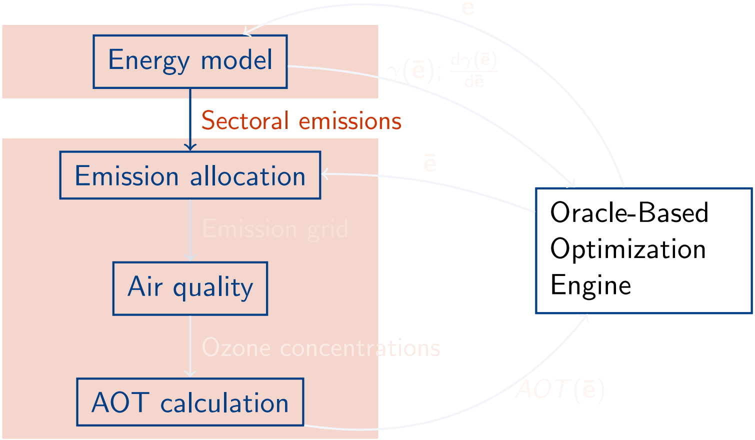
Contact: Ulrich.Leopold@Tudor.lu

Appendix

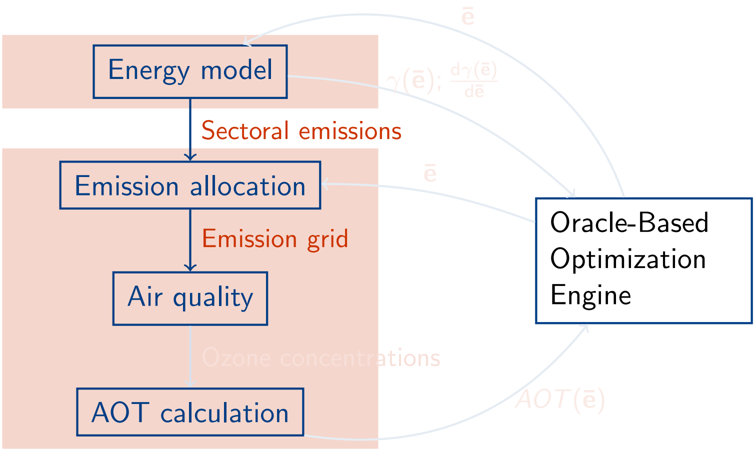
Optimization Framework



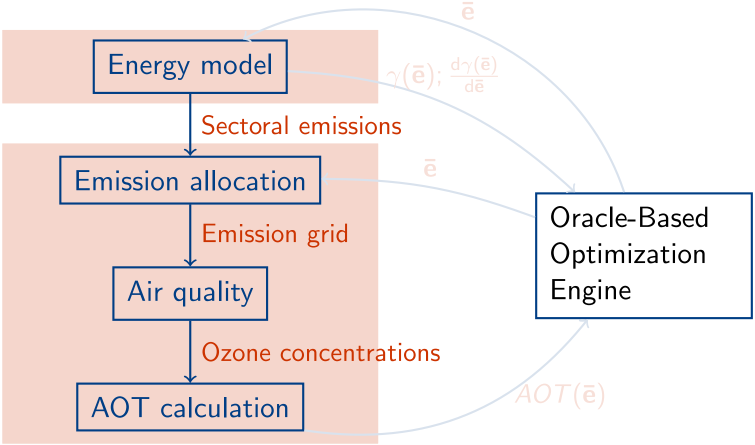
Optimization Framework



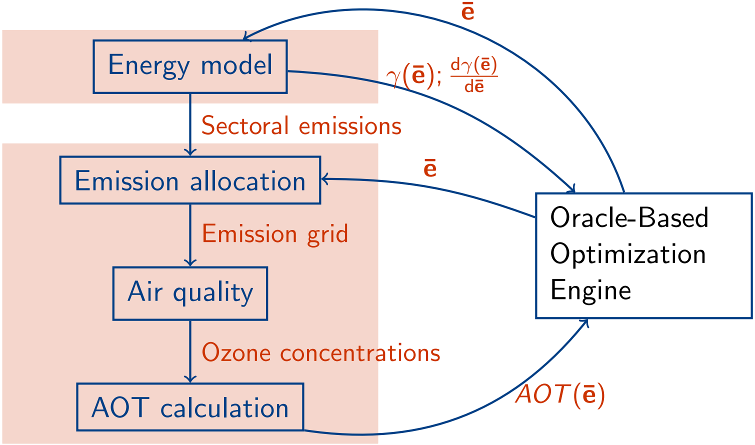
Optimization Framework



Optimization Framework



Optimization Framework



Spatial support problems