GIMA

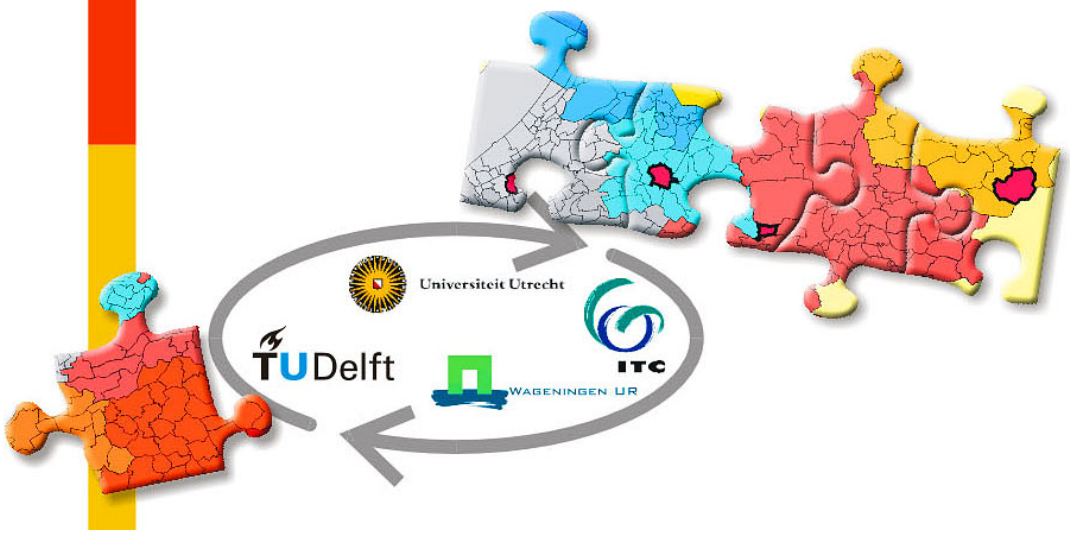
Geographical Information Management and Applications

Uncertainty Propagation Analysis of Nitrogen Fluxes and Methane Emissions from Agricultural Areas in Europe

MSc Thesis Report
July 2011

Vasiliki Ioannidi
vicky.ioannidi@wur.nl

Professor: Prof. Dr. Ir Arnold Bregt (WUR)
Supervisor: Dr. Ir. Gerard Heuvelink (WUR)
Co-supervisors: Dr. Ir. Hans Kros, Prof. Dr. Ir. Wim de Vries (WUR)
Reviewer: Prof. Dr. Ir. Alfred Stein (ITC)



Acknowledgements

Having now completed my MSc thesis on "Uncertainty Propagation Analysis of Nitrogen Fluxes and Methane Emissions from Agricultural Areas in Europe" I would like to express my gratitude to all those who supported me during the implementation of my research. I would like to thank greatly my supervisor Gerard Heuvelink for his excellent guidance and his charisma to provide answers to my questions in such a clear manner. I am also grateful to my co-supervisors Hans Kros and Wim de Vries for their great instruction, for providing the required data, for conveying knowledge about the INTEGRATOR model but most importantly I want to thank them for their kindness and patience. Many thanks go to Gert Jan Reinds as well for his major contribution by adjusting the INTEGRATOR model to the inputs selected for my analysis and thus helping me in this way to carry it out. Finally, I want to thank Jan Cees Voogd, my roommate at Atlas building, for his assistance by running part of my extensive computational work through his PC.

Vicky Ioannidi
Arnhem, July 2011

Summary

The INTEGRATOR model (**I**ntegrated **N**itrogen **T**ool across **E**urope for **GR**een House Gas and **A**mmonia **T**argeted to **O**perational **R**esponses) is a GIS-based, multi-component tool developed to assess the impacts of European-scale changes in land-use, land-management and climate on N fluxes and Green House Gases (GHG). The inputs used in INTEGRATOR are detailed GIS data describing environmental factors related to N and GHG such as land cover, climate, soil type and soil properties and also farming and agricultural applications. These inputs, due to their limited observations, contain always uncertainty. The objective of this research was to analyse how uncertainties in model inputs propagate to NH_3, N_2O, NO_x, CH_4 emissions, and N leaching into surface- and groundwater estimated by INTEGRATOR. The research was limited to agricultural areas in Europe. Uncertainties in categorical inputs, such as land cover and soil type were not included in the analysis. In total 56 Agricultural Parameters (APs) were considered as model inputs concerning:(i) soil properties, (ii) model parameters affecting N inputs to the system, i.e. N fixation, N deposition, N manure input and N fertilizer and (iii) model parameters affecting N and CH_4 fluxes to and from agricultural systems. These APs were divided into four groups according to the four spatial-scale levels of INTEGRATOR in order of increasing size: NitroEurope Computational Units (NCU), EU territorial units (NUTS2/3), EU member states (CNTRY) and EU-25 (EUROPE). The uncertainty in the APs was expressed by defining their probability distribution functions (pdfs) and taking into account their spatial- and cross-correlations. Additionally, three uncertainty scenarios (Optimistic, Reference, and Pessimistic) were incorporated in the research to investigate the robustness of the uncertainty analysis. The outputs were produced at Country and European level for the year 2000, and the propagated uncertainty in them was quantified by applying the Monte Carlo simulation to the model. Results of this research indicate that: (i) when using the Reference scenario, the output uncertainty, expressed as Coefficient of Variation (CV), varies from 10-34% for outputs at European level and from 11-92% for outputs at Country level (ii) the uncertainty increased going from CH_4 and NH_3 emissions to N_2O emissions, to NO_x emissions, to N leaching into surface- and groundwater (iii) the maximum value of CV for the Pessimistic scenario was 55% whereas it was only 12 % for the Optimistic scenario and (iv) the APs at NCU level had the largest contribution to the output uncertainty. In summary the results indicate that it is a quite robust statement to say that the uncertainty is less than 55% for all considered model outputs and is most likely in the range of 12-55%.

Table of Contents

List of Appendices

List of Tables

List of Figures

1. INTRODUCTION

1.1 Context and background

It has been a few decades now since the issue of the greenhouse effect and its impact on climate change gained the attention of the international scientific community. Human-induced burning of fossil fuels and changes in land use patterns, at a broad scale, have resulted in increased annual emission rates of greenhouse gases (GHG) such as carbon dioxide (CO_2), nitrous oxide (N_2O), and methane (CH_4). This has led to an enhanced focus by international organizations with an attempt to achieve coordinated action on the global warming issue. In 1989 the Intergovernmental Panel on Climate Change (IPCC) was set up by the World Meteorological Organization (WMO) and the United Nations Environment Program (UNEP) to provide governments with broad and balanced information on the world's climate. The scientific evidence brought up by the first IPCC Assessment Report of 1990 played a major role in the creation of the United Nations Framework Convention on Climate Change (UNFCCC) fifteen years ago (http://www.ipcc.ch/). The UNFCCC is the key international treaty which many countries have joined in order to face the consequences of climate change (http://www.ipcc.ch/). The Kyoto Protocol, which was adopted in 1997 and entered into force in 2005, is an addition to this treaty and has more powerful measures as it sets legally-binding targets for 37 industrialized countries and the European community. According to Article 7 of the Protocol all signatories are required to submit annual inventories of anthropogenic emissions by sources and removals by sinks of GHG and any steps taken to increase the accuracy of these inventories.

The production of GHG is determined by processes that occur within the biogeochemical cycles of nitrogen (N) and carbon (C), the interactions between these cycles and the degree their equilibrium is perturbed by external factors. The N cycle is crucially important to all living organisms, ecosystem functioning and global change. Additionally it plays a key role in the global C cycle through the effects of primary production and decomposition, influences considerably the terrestrial and aquatic biodiversity, and has major impacts on N_2O and CO_2 fluxes and to a lower extent on CH_4 fluxes from agricultural and natural areas. Besides the GHG fluxes, the increased use of N in European agriculture has raised the NH_3 emissions which affect greatly the plants diversity due to exceedances of critical N loads. Elevated NH_3 emissions lead also to bigger amounts of N leaching to groundwater and runoff to surface water with the related effects on the drinking water quality and the eutrophication of rivers, lakes, wetlands etc. (De Vries et al. 2011a). So it was a matter of major concern that the annual transfer of reactive Nitrogen (Nr which includes all N forms except N_2, i.e. ammonia (NH_3), ammonium (NH_4^+), nitrate (NO_3^-), nitrous oxide (N_2O), nitrite (NO_2), nitrogen oxides (NOx) was estimated to have more than doubled according to Vitousek et al.(1997) and Galloway et al. (2004). The C cycle comparatively was estimated to have been less than 10% disturbed by human activities (IPCC, 2001), but less attention had been paid recently to the quantification of the N cycle than the C cycle. According to Li et al. (2005) climate change approaches that can maximise CO_2 uptake may not necessarily optimize Net Greenhouse gas Exchange (NGE). Consequently, the need for an integrated assessment of NGE, rather than just of CO_2, towards a future strategy development became finally apparent. In 2006 the integrated European Research Project NitroEurope (NEU, http://www.nitroeurope.eu/) was

initiated to deal with the prime issue of European N budgets in relation to C cycling and NGE. The NEU focuses on the integrated analysis in N fluxes and NGE at linked plot-, regional-, and European-scales while at the same time taking into account spatial interactions with other environmental issues such as the interactions with NH_3 emissions and NO_3^- leaching/runoff which are considered indirect sources of N_2O emissions under the IPCC methodology (De Vries et al. 2011c,d).

The main question that NEU is addressing is: *What is the effect of Nr supply on net GHG budgets for Europe?* The objectives of the project are:

- To build vigorous datasets of N fluxes and NGE in relation to C-N cycling of various ecosystems across Europe in order to investigate interactions and assess long-term changes;
- To quantify the effects of past and present global changes in climate, atmospheric composition and land use on C-N cycling and NGE;
- To simulate the observed fluxes of N and NGE, their synergies and their responses to global change decisions;
- To quantify multiple N and C fluxes for contrasting European landscapes, including interactions between farm-scale management, atmospheric and water dispersion;
- To scale up N and NGE fluxes for terrestrial ecosystems to regional and European levels by taking under consideration spatial variability;
- To assess uncertainties in estimates by European models and combine them with independent measurement/inverse modelling approaches for verification of European N_2O and CH_4 inventories and refinement of IPCC approaches.

The INTEGRATOR (**I**ntegrated **N**itrogen **T**ool across **E**urope for **GR**een House Gas and **A**mmonia **T**argeted to **O**perational **R**esponses) model was developed by Wageningen University and Research Centre to tackle the above objectives. INTEGRATOR is a GIS-based, multi-component modelling with an aim to assess:

- Present N (NH_3, NO_x) and GHG (CO_2, N_2O, CH_4) emissions and sinks from terrestrial systems with a special focus on agricultural systems;
- C-N interactions between agricultural and non-agricultural systems;
- Past and future N and GHG emissions and sinks in response to diverse scenarios which reflect past and present land cover and land management decisions and also policies and European Union (EU) directives that affect N emissions in interaction with GHG emissions and climate change (De Vries et al. 2011c).

The inputs used in INTEGRATOR are detailed GIS data describing environmental factors such as climate, land use, soil type and soil properties as well as farming and agricultural management applications. The main concept for the simulation of the processes related to N and GHG emissions is based on:

- The use of relatively transparent calculations from available simple models;
- The inclusion of empirical model approaches with statistical relations between model outputs and environmental variables;
- The focus on the derivation of high resolution spatially explicit input data (De Vries et al. 2011c).

INTEGRATOR assesses N and GHG emissions by sources and removals by sinks, focusing on natural and in particular agricultural systems. However, the uncertainty in estimates of INTEGRATOR inputs such as various agricultural parameters (from now on referred to as APs), determining the emissions from these sources, is rather large

and mainly at local and national scale due to legislation or country-specific management (Oenema et al. 2007). The uncertainty though in model inputs propagates to model outputs and this is why uncertainty assessment is one of the main objectives of the project NEU project.

Since uncertainty propagation analysis is an integral part of NEU modelling, all model developers have agreed to submit uncertainty reports next to the calculated emissions. The most common approach they use in order to analyse how uncertainty in model inputs propagates, through the model, to the outputs is the Monte Carlo simulation. The concept of this approach is to compute repeatedly multiple output realizations by running the model for samples of multiple input realizations which are randomly drawn from the probability distribution of each input (Van Oijen 2006). The multiple output realizations which are produced after the model has been run for all input realizations can be summarized statistically and in this way conclusions can be made about the uncertainty in N and GHG emissions.

1.2 Problem definition

As mentioned above, the inputs used in INTEGRATOR contain sources of uncertainty that propagate to the model outputs. According to Van Oijen 2006 the inputs are defined as all the information needed to run the model which is not incorporated in the model itself. These inputs can be of three types:

1. initial values of state variables at the start of the simulation (e.g. the amount of N already stored in the soil);
2. model parameters (e.g. emission fractions);
3. environmental constants and variables (e.g. soil pH, soil type, precipitation, temperature etc.).

None of these inputs are free or error. For example, emission fractions are mainly obtained by expert judgement or calibration due to lack of hard data. In addition, the model itself also contains errors, because a model is by definition a simplified representation of the reality based on many assumptions. Hence the model output uncertainty is caused by two main sources (Heuvelink et al. 2009):

1. the model input uncertainty;
2. the model structure uncertainty.

Subsequently, the uncertainty assessment of the INTEGRATOR output can be achieved by quantifying the uncertainty in these two sources and analysing its propagation through the model. By this assessment the model can be further improved, the objectives of the NEU can be met and ultimately the GHG inventories according to Kyoto Protocol can be completed with their accuracy quantified.

For this research though, only the model input uncertainty was considered, and from that source only model parameters (APs) that concerned agricultural areas were taken into account. For example, driving forces (e.g. animal numbers), emission fractions (e.g. NH_3 from housing systems), soil properties (e.g C/N ratio) etc. Uncertainties in climatic variables and in categorical data such as land cover and soil type were not included in the analysis.

Within the context of this research, the uncertainty was defined as the lack of knowledge about either the estimated value or the measurement error or the variability

of an AP. Uncertainties about variables are often expressed by characterising their probability distribution functions (pdfs). The most common distribution is the normal or Gaussian distribution which is specified by its two parameters: the mean (μ) and the variance (σ^2). For uncertain variables, however, which are located within a given spatial unit and may be related with each other, their spatial- and/or cross-correlations must be taken into account as well. For example: (i) emission fractions of a plot, e.g. at location x_1, may be related through natural processes with emission fractions at another plot, e.g. at location x_2, within a given neighbourhood and thus covary as a result (i.e. spatial-correlation) and (ii) APs from different N sources, e.g. N excretion rates from cattle and N content in the grass, may covary as well (i.e. cross-correlation). Therefore spatial- and cross-correlations had to be specified within the spatial aggregation levels distinguished in INTEGRATOR which are referring to the spatial extent of the areas under study. These levels are described below:

a. the NCU level (NitroEurope Computational Unit) which is the smallest spatial unit and represents the local areas;
b. the NUTS2/3 level (Nomenclature of Units for Territorial Statistics according to Eurostat) which consists of a number of NCUs and represents the regional areas;
c. the Country level which consists of a number of NUTS regions and represents each EU country;
d. the European level being the largest one.

Truong (2009) used a statistical model to quantify the uncertainty in N_2O emissions from agricultural areas, based on emission fractions which were treated as variable in space but it was assumed their spatial- and cross- correlations were ignored due to lack of data on them. This is not realistic though because if spatial- and cross-correlation are not incorporated in the model, the uncertainty assigned independently to the NCU level will vanish completely when model outputs are up-scaled to the Country or European level (Kros et al. 2010). This will be explained in detail later on in this report.

The problem addressed in this research was the uncertainty quantification of the following model outputs: NH_3, N_2O, NO_x, CH_4 emissions, and N leaching into surface- and groundwater released from agricultural areas. These outputs were estimated by INTEGRATOR:

a. while the uncertainty in the APs was estimated by taking into account their spatial- and cross-correlation within the spatial aggregation levels defined by the model;
b. for each individual EU country and for Europe as a whole.

1.3 Research objective

The general research objective was to analyse the propagation of the uncertainty in the APs to the INTEGRATOR output emissions of NH_3, NO_x, CH_4, N_2O, and N leaching into surface- and groundwater, at Country and European level, for the year 2000.

1.4 Research questions

The research questions that were addressed during this study were the following:

1. How can statistical models (joint pdfs) be built that fully characterize the uncertainty in the APs by taking into account their spatial- and cross-correlations?
2. How can realizations of APs be sampled efficiently from their pdfs by using stochastic simulation techniques?

3. How can the Monte Carlo uncertainty propagation analysis be carried out in batch mode and its results be stored automatically?
4. How can the results of the uncertainty propagation analysis be summarised and visualised and thus efficiently communicated to end-users across Europe?
5. Which APs are the main uncertainty sources contributing to the total output uncertainty?

1.5 Research limitations

This research was based only on the quantification of the uncertainty only in the APs and not in any other variables. For example, uncertainties in climatic variables (e.g. precipitation, temperature etc.), in land cover/land management and soil type data were not considered.

1.6 Structure of the report

In addition to this Introduction Chapter, this report contains another four Chapters:

- Chapter 2 is describing the steps of the Generic Methodology for Spatial Uncertainty Quantification (UQ) / Uncertainty Analysis (UA);
- Chapter 3 is describing how the Generic Methodology for Spatial Uncertainty Quantification (UQ) / Uncertainty Analysis (UA) was applied to the INTEGRATOR model. In particular, explanation is given about the model, about the inputs and the outputs considered for this research, how the uncertainty in the inputs was represented, and how the Monte Carlo simulation was applied to the model;
- Chapter 4 is presenting the Results of this research based on visualizations of calculated statistics on the uncertainty in the inputs and the outputs and Discussion about them is given as well;
- Chapter 5 is providing the General Conclusions of the research, a summary of how the research questions were addressed and also Recommendations about further research steps.

Appendices are also given at the end of the report containing some extra information and the developed R-scripts.

2. GENERIC METHODOLOGY FOR SPATIAL UNCERTAINTY QUANTIFICATION (UQ) / UNCERTAINTY ANALYSIS (UA)

Within NitroEurope, Uncertainty Quantification (UQ) and Uncertainty Analysis (UA) are two important components of the operational methodology for regional scale modelling in order the net N and CHG exchange to be assessed. This methodology is described in Heuvelink et al. (2009). UQ and UA consider the case in which uncertainty in model inputs is analysed as to how it propagates to the output through the model. In particular, the purpose of UQ is to quantify the model output uncertainty, whereas the goal of an UA is to determine how much individual input uncertainties account for the overall output uncertainty. Depending on the available data, several representations of the uncertainty in model inputs can be used such as pdfs, ranked or unranked scenarios, rough stets or fuzzy sets. The pdfs though being rather more comprehensive are the most preferred representations.

2.1 Monte Carlo simulation in spatial modelling

The Monte Carlo (MC) simulation is a rather prevailed approach in spatial modelling for UQ/UA. The consecutive steps taken during the application of this approach are:

1. By using stochastic simulation techniques, a sample drawn from the input's pdf is generated consisting of multiple input simulations (realizations);
2. The MC approach is applied to the model and it is run for each member of the sample;
3. The produced multiple output simulations can be interpreted as a random sample of the output's pdf. By obtaining parameters of this pdf, the level of the output uncertainty can be derived providing the drawn sample of input simulations is big enough to represent the input uncertainty (Van Oijen 2006).

2.2 Selection of model outputs for which the uncertainty is quantified

Most of the time models can have a large number of outputs and each of them may have a different level of uncertainty. Since UQ/UA is a procedure that focuses only on those sources that have the largest contribution to the output uncertainty, the targeted model outputs must be explicitly defined and this in turn will determine which uncertain sources affect these outputs and consequently must be considered in the analysis.

The spatio-temporal support of the model output, which refers to the spatial area and temporal interval over which the variable is aggregated, must also be defined since both the magnitude of the output uncertainty and the degree the uncertain sources contribute to it depend generally on this support. Usually the generated input simulations at which the model calculations are implemented have a much smaller spatio-temporal support than the required output support. However, the simulations of the small support outputs, produced at the end of the model, can be easily aggregated to the larger support by simply computing the average of the small support values contained within the larger support. So the larger the size of the output support, the larger the number of the small supports it can contain and thus the smaller the magnitude of the estimated uncertainty.

2.3 Selection of uncertainty sources included in the analysis

Uncertainty in the model output, as it has been already mentioned, is caused by two uncertainty sources: a. the model input uncertainty and b. the model structure

uncertainty. Initially it is decided which of these sources is included in the UQ/UA and then they are ranked in order of treatment as it would be rather a complex process to address them both at once. So the most important source is considered first and once it is completed the second one is added.

2.4 Selection of model inputs for which the uncertainty is analysed

Models can have hundreds or thousand of inputs which need not be all treated as uncertain. Only the ones that contribute significantly to the model output uncertainty should be taken into account and this can be concluded by checking two factors: a. the sensitivity of the model to changes in inputs and b. the magnitude of uncertainty about the model input.

The above two factors can confirm model experts´ fair ideas concerning the most sensitive inputs of the model and also the difference between the true value of an input and its estimated value available in a database. By computing changes in model outputs caused by fixed changes in inputs (i.e. ± 10%) they can take final decisions about the sensitivity of the inputs. By using historical data on given regions and other additional information they can make assessments about the future values of inputs and their expected uncertainty.

2.5 Uncertainty Quantification in model inputs

Once it is decided which inputs are treated as uncertain, then their uncertainty must be quantified. First a distinction should be made between inputs depending on whether they are measured on a continuous numerical scale or a discrete numerical scale. Next**,** it is checked whether spatial-, temporal- and cross-correlations need to be considered as well, and finally one of the available methods to estimate pdfs´ parameters is decided upon.

The method followed to quantify uncertainty in continuous numerical variables that are spatially- and cross-correlated (this is the case for the APs of this research) is described below:

- Initially, the cumulative distribution function, **F**, of the variables is defined. An uncertain continuous numerical variable X, that varies in space, is characterised by its joint function **F**:

$$F_{X(x_1, s_1 \ldots, x_n, s_n)} = P(X(s_1) \le x_1, \ldots\ldots, x(s_n) \le x_n) \quad (1)$$

which describes the probability that X takes on values less than or equal to numbers x_1, ..x_n at multiple plots $s_1 \ldots s_n$ (n may be any integer value). **F** must be known for each and every combination of the x_i and s_i. The probability density function of X, **P** is the first derivative of **F** and must be non-negative everywhere with a surface area below its curve being equal to one (Figure 1).

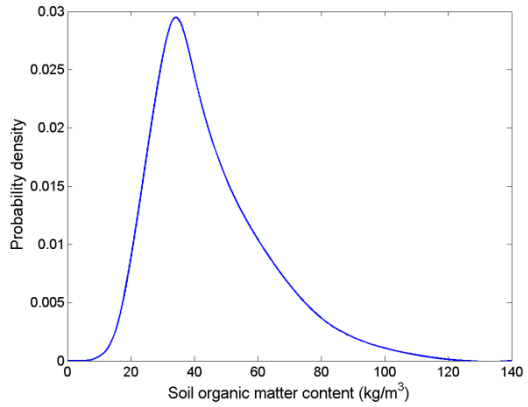


Figure 1. Example of a continuous probability density for soil organic matter content

Equation 1 represents the general case, where the only restriction on **F** and **P** is that are valid functions. Practically these functions can be obtained through different kind of approaches but most of the time several assumptions and simplifications need to be made in order to get reliable estimates of their parameters. One of the assumptions rather imposed, for uncertain input variables that are spatially- and cross-correlated, is that they follow a normal distribution (see Figure 2) otherwise the UQ/UA would be too complex. Another assumption for input variables that are non-normally distributed (i.e. they have a form of skewness) is that some transformation of their values, such as the logarithm or the square root, are normally distributed and the uncertainty in these variables is defined in terms of their transformed values. Subsequently, the parameterization of the pdf of X (either the **F** or **P**) takes place through the normal distribution and X can be denoted as $X \sim N(\mu, \sigma^2)$. The parameters of this pdf (i.e. mean (μ), variance (σ^2), semivariogram and cross-variogram) are derived from literature, observations, or expert knowledge.

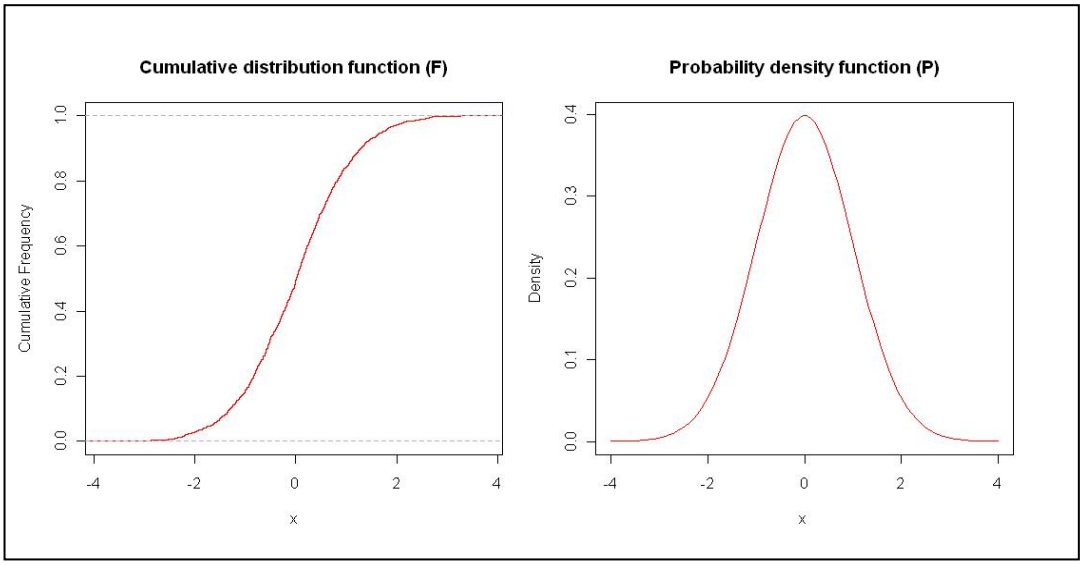


*Figure 2. Graphs of the cumulative distribution function **F** and the probability density function **P**, respectively, of a normal variable X*

- The next step, after the joint pdfs for all inputs have been defined, is to generate samples of multiple input simulations drawn from these pdfs, by using pseudo-random number generators which are incorporated in most statistical software tools and programming languages. When inputs are high-dimensional, the size of their sample must be very large as well in order to be representative across the whole range of allowable values for these inputs. One approach to calculate the required sample size of an input is, to check the cumulative frequency curves produced by the differences between its defined standard deviation (SD) and the SD calculated from a first batch of generated input simulations, for all multiple plots this input is linked to. If these differences are not normally distributed closely around 0 then the size of the sample should be increased, and the new cumulative frequency curves be checked again. This iteration can take place a few times until the above condition is met. Only then can the sample size be considered sufficient to represent the uncertain inputs in the UQ/UA process.

2.6 Uncertainty Quantification in model outputs

After the samples of the multiple input simulations are generated, then the MC method is applied to the model and it is run for each one of these input simulations. At the end of each MC run of the model, each required output simulation is produced and stored at the desired spatial and temporal support. These multiple output simulations form random samples from each output's pdf. By applying statistical sampling theory, estimates of the parameters of these pdfs can be obtained and thus the uncertainty in model outputs be quantified. In Figure 3 it is demonstrated how uncertainty in INTEGRATOR outputs, e.g. $Y_1,\ldots Y_k$, can be quantified through the generation of multiple simulations of the inputs, e.g. AP_1, $AP_2,\ldots$ AP_n.

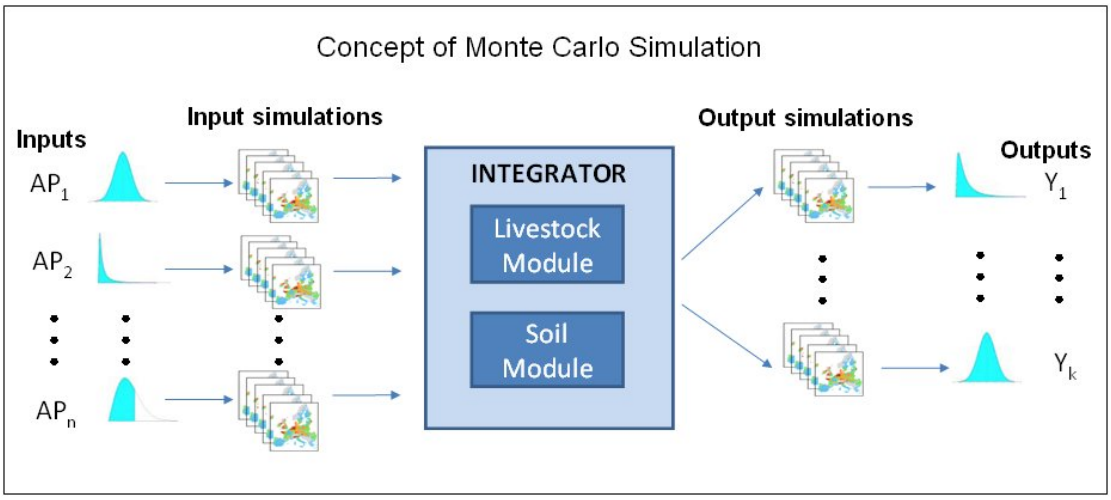


Figure 3. Workflow scheme for estimating uncertainty in INTEGRATOR outputs by applying Monte Carlo Simulation

Before however the model is run, a few provisions should take place to make sure this demanding and time consuming step is properly implemented. First an assessment should be made about the computing time and the storage capacity required for the multiple runs of the model and its outputs respectively. If necessary, the computer recourses may be extended, but in cases this cannot be achieved then some alterations may be considered to reduce complexity. Another important factor to be considered is the streamline of the repeated running of the model which has to be set up in such a way that the amount of work is manageable. Additionally, the model should be run in

batch mode, and all inputs and outputs involved should be stored in an organised and transparent manner. The entire procedure can be checked before the full analysis takes place by running the model with a small subset of input simulations.

2.7 Uncertainty Analysis

The objective of UA is to determine the contribution of the individual sources of uncertainty to the overall output uncertainty.

The simplest method to implement an UA is to apply the MC simulation again, but this time by considering uncertain only those inputs for which the contribution must be estimated. The rest of the inputs are assumed to be certain and fixed on their reference values. By computing the ratio of the variance estimated when only one input is made uncertain to the total variance estimated when all inputs are made uncertain, the relative uncertainty of each input can be derived. This approach though has some disadvantages since uncertainty is addressed based only on the inputs that are being checked while the rest of them having their fixed reference values and not necessarily true values, and also based on correlations between the uncertain and the fixed inputs being ignored. Nevertheless these disadvantages cannot be completely avoided, there are some techniques developed to tackle the problem more thoroughly and they can be found in Jansen et al. (1994), Saltelli et al. (2000), Crosetto and Tarantola (2001), and Jansen (2005).

2.8 Communication of the UQ/UA results

The last step, after parameters of the output pdfs (e.g. mean, SD, CV, interquartile range, percentiles etc.) have been estimated, is the communication of these results to end-users and decision makers in an efficient and comprehensive way, so they can incorporate the information about uncertainty in their tasks (i.e. risk analysis). Several visualizations can be used to this end: box-plots, bar- or column charts, tables, graduated colour maps of uncertainty levels, maps of confidence limits, animations etc. Sufficient explanation about the meaning of these results is also required to help modellers assess the accuracy of the model and draw conclusions on how to enhance it.

3. APPLICATION OF THE GENERIC METHODOLOGY FOR SPATIAL UNCERTAINTY QUANTIFICATION / UNCERTAINTY ANALYSIS TO THE INTEGRATOR MODEL

3.1 INTEGRATOR model

INTEGRATOR is a dynamic model developed to assess the impacts of European-scale changes in land-use, land-management and climate on N fluxes and GHG at a high spatial resolution, in the past and the future, focusing on annual changes in the period 1970-2030. It covers major ecosystems such as grassland, arable land, forest, heathlands and peat lands and it also incorporates interactions between agricultural and natural areas through the emissions of NH_3 and NO_x which account for the N deposition (De Vries et al. 2011c,d).

INTEGRATOR consists of simplified process-oriented and empirical sub-models estimating N and GHG emissions. These sub-models are illustrated in Figure 4 and described below:

a. The adapted MITERRA-Europe model estimating NH_3, NOx, N_2O and CH_4 emissions from housing and manure storage systems and also agricultural soils.
b. Empirical models estimating NO_x, N_2O and CH_4 emissions from natural areas.
c. The YASSO soil model estimating CO_2 emissions from agricultural and natural areas, in combination with EFISCEN for forest, MITERRA-Europe for agriculture and empirical relationships for peat lands.
d. An emission-deposition matrix for NH_3 and NO_x estimating N deposition.

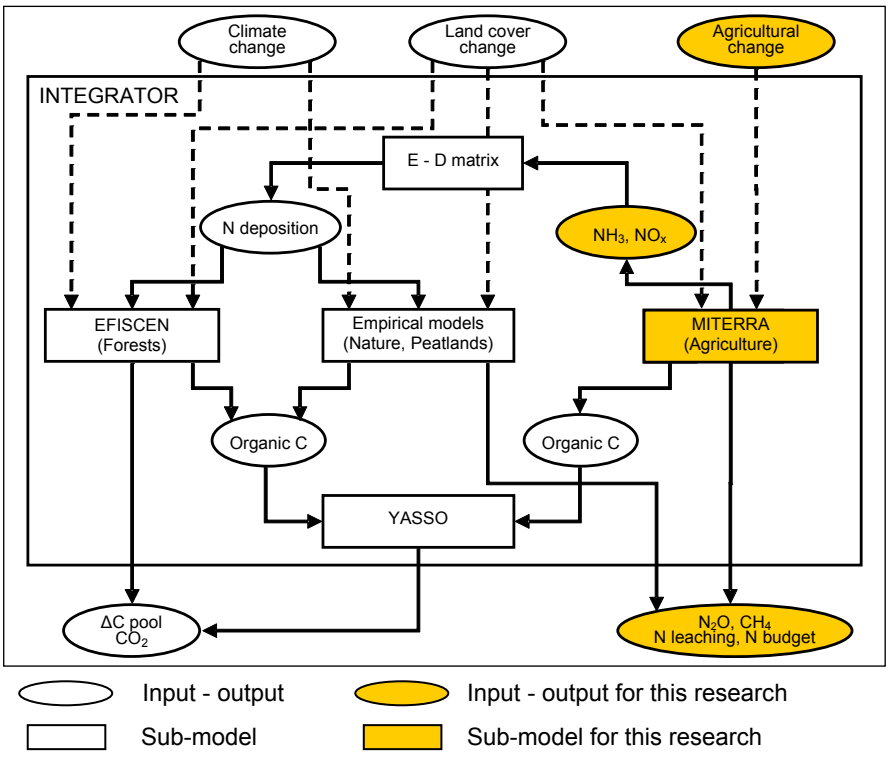


Figure 4. The INTEGRATOR model: its drivers, main modules and main outputs (Source: De Vries et al. 2011c,d)

The above shown sub-models are modular and exchangeable and thus any model components could be simultaneously applied for uncertainty quantification.

Since for this research only the agricultural areas were considered, the only inputs, and sub-models of INTEGRATOR that were incorporated in the analysis were the ones illustrated in gold colour in Figure 4.

The model outputs and their standard units are shown in Table 1.

Table 1. Model outputs

Flux	**Standard unit** [1]
NH_3 emission	Kg NH_3-N ha^{-1} yr^{-1}
N_2O emission	Kg N_2O-N ha^{-1} yr^{-1}
NO_x emission	Kg NO_x-N ha^{-1} yr^{-1}
N leaching into ground water (N_{leGW})	Kg N ha^{-1} yr^{-1}
N leaching into surface water (N_{leSW})	Kg N ha^{-1} yr^{-1}
CH_4 emission	Kg CH_4 ha^{-1} yr^{-1}

1) Kg NH_3-N ha^{-1} yr^{-1} means: the amount of N contained in NH_3 emissions, in Kg, per hectare, per year.

The model inputs, determining the emission and leaching processes of these output fluxes, were the APs described in detail in Section 3.2.

INTEGRATOR was applied for 25 (and not for 27) member countries of EU because for Malta and Cyprus there were no data available to be included in the analysis. The spatial support of the model was the NCU which is a spatially explicit unit consisting of clusters of 1 km^2 grid cells. These clusters resulted from the combined partitioning of a number of input maps, based on the requirement of processes included in INTEGRATOR and their dependence on site and soil characteristics. So the NCUs consist of unique combinations of land use, soil type and agricultural structure (Truong, 2009). For EU-25 there were 35101 NCUs distinguished. The size of the NCUs was highly variable with a mean area of 163km^2 and SD = 557km^2 and this was because in flat areas with uniform soil type the NCUs were larger but in mountainous areas with variable soils they were much smaller. The implicit assumption of using the NCU as the spatial support for the application of UQ/UA was that the spatial variability within an NCU was not (explicitly) taken into account. For each NCU, averages of APs (see Section 3.3 for further details) were used which included spatial variability (Kros et. al. 2010).

The model inputs were prepared as csv (Comma Separated Values) files where all the NCUs within EU-25 were assigned simulated values of APs via their NCU_ids. INTEGRATOR assessed every NCU, and model outputs were produced for every NCU as well. These outputs were finally aggregated (averaged) to: a. the Country level and b. the European level.

The temporal support of the model inputs was referred to the year 2000.

3.2 Defining the uncertain model inputs

The APs that have been considered in the UQ/UA were delivered by the group of the INTEGRATOR developers. Uncertainty was limited to quantitative model inputs: (i)

soil properties, (ii) model parameters affecting N inputs to the system, i.e. N fixation, N deposition, N manure input and N fertilizer and (iii) model parameters affecting N and CH4 fluxes to and from agricultural systems.

Although the spatial support of INTEGRATOR model was the NCU, not all parameters were available for each NCU within EU-25. Some parameters were linked to the NUTS level (e.g. N excretion rates), some to the Country level (e.g. animal numbers), in short CNTRY, and some others were linked to the European level (e.g. N fertilizer distribution). The latter were called GENERIC parameters, in short GEN. There were 56 APs selected and they are provided in Table 2 along with their codes, the spatial levels they were linked to and the output fluxes they affect.

Table 2. Model inputs

Nr	Agricultural Parameters (APs)	Code	Level	Affected fluxes
	Livestock excretion and emission data			
1	Animal numbers, dairy cattle	aninr_ca	CNTRY	NH_3, N_2O, NO_x, NleSW, NleGW, CH_4
2	Animal numbers, other cattle	aninr_oc	CNTRY	NH_3, N_2O, NO_x, NleSW, NleGW, CH_4
3	Animal numbers, pigs and poultry	aninr_pp	CNTRY	NH_3, N_2O, NO_x, NleSW, NleGW, CH_4
4	Animal numbers, other animals	aninr_po	CNTRY	NH_3, N_2O, NO_x, NleSW, NleGW, CH_4
5	N excretion rates, dairy cattle	Nexf_ca	NUTS	NH_3, N_2O, NO_x, NleSW, NleGW
6	N excretion rates, other cattle	Nexf_oc	NUTS	NH_3, N_2O, NO_x, NleSW, NleGW
7	N excretion rates, pigs and poultry	Nexf_pp	NUTS	NH_3, N_2O, NO_x, NleSW, NleGW
8	N excretion rates other animals (horses, sheep and goats and other animals)	Nexf_po	NUTS	NH_3, N_2O, NO_x, NleSW, NleGW
9	C/N ratios manure	Cnam	NUTS	NH_3, N_2O, NO_x, NleSW, NleGW
10	Availability fraction of N deposition compared to N fertilizer (-)	f_Nav_dep	GEN	NH_3, NleSW, NleGW
11	Availability fraction of organic N in animal manure (either applied or excreted by grazing) in crop residues and from soil mineralized N compared to N fertilizer (-) for arable land and grassland	f_Nav_om	GEN	NH_3, NleSW, NleGW
12	Housing fractions, dairy cattle	fhs_ca	NUTS	NH_3
13	Housing fractions, other cattle	fhs_oc	NUTS	NH_3

Nr	Agricultural Parameters (APs)	Code	Level	Affected fluxes
14	Fraction of excreted amount stored as liquid manure in the housing system, cattle	frlam_ca	NUTS	NH_3, N_2O
15	Fraction of excreted amount stored as liquid manure in the housing system, pigs and poultry	frlam_pp	NUTS	NH_3, N_2O
16	NH_3 emission fraction from housing systems	fNemhs_NH3	NUTS	NH_3
17	NH_3 emission fraction from manure storage systems	fNemms_NH3	NUTS	NH_3
18	N_2O emission fraction from housing systems (liquid)	fNemhsl_N2O	NUTS	N_2O
19	N_2O emission fraction from manure storage systems (liquid)	fNemmsl_N2O	NUTS	N_2O
20	NO emission fraction from housing systems (liquid)	fNemhsl_NO	NUTS	NO_x
21	NO emission fraction from manure storage systems (liquid)	fNemmsl_NO	NUTS	NO_x
22	N_2O emission fraction from housing systems (solid)	fNemhss_N2O	NUTS	N_2O
23	N_2O emission fraction from manure storage systems (solid)	fNemmss_N2O	NUTS	N_2O
24	NO emission fraction from housing systems (solid)	fNemhss_NO	NUTS	NO_x
25	NO emission fraction from manure storage systems (solid)	fNemmss_NO	NUTS	NO_x
	Nitrogen input data			
26	Allocation fraction for arable and grassland in the manure and N input assessment procedure	Wamhsara	NUTS	NleSW, NleGW
27	Weighting factor for grassland and fodder in the manure and N input assessment procedure	Wamhsgrass	NUTS	NleSW, NleGW
28	Areas (fractions) of intensively and extensively managed grassland (Area_ext = Area_grass - Area_int)	Area_int	NUTS	NleSW, NleGW
29	National fertilizer N inputs	tNfe	CNTRY	NH_3, N_2O, NleSW, NleGW
30	N deposition data	Ndep	NCU	N_2O, NO_x, NleSW, NleGW
31	N fixation, arable + fodder	Nfix_ar	NCU	N_2O, NO_x, NleSW, NleGW
32	N fixation, grass (int + ext)	Nfix_gr	NCU	N_2O, NO_x, NleSW, NleGW
33	N fixation, legume	Nfix_le	NCU	N_2O, NO_x, NleSW, NleGW
	Nitrogen uptake /immobilization data			
34	Yields, arable	Yieldopt_ar	NCU	NH_3, N_2O, NO_x, NleSW, NleGW

Nr	Agricultural Parameters (APs)	Code	Level	Affected fluxes
35	Yields, fodder	Yieldopt_fo	NCU	NH_3, N_2O, NO_x, NleSW, NleGW
36	Yields, grass intensive	Yieldopt_gi	NCU	NH_3, N_2O, NO_x, NleSW, NleGW
37	Yields, grass extensive	Yieldopt_ge	NCU	NH_3, N_2O, NO_x, NleSW, NleGW
38	Maximum N content in the harvested crops, arable	ctNplmx_ar	NCU	NH_3, N_2O, NO_x, NleSW, NleGW
39	Maximum N content in the harvested crops, fodder	ctNplmx_fo	NCU	NH_3, N_2O, NO_x, NleSW, NleGW
40	Maximum N content in the harvested crops, grass intensive	ctNplmx_gi	NCU	NH_3, N_2O, NO_x, NleSW, NleGW
41	Maximum N content in the harvested crops, grass extensive	ctNplmx_ge	NCU	NH_3, N_2O, NO_x, NleSW, NleGW
42	N index	Nind	NCU	NH_3, N_2O, NO_x, NleSW, NleGW
43	Uptake fraction	Fup	NCU	NH_3, N_2O, NO_x, NleSW, NleGW
44	Ratio between minimum and maximum N uptake	Frmin	NCU	NH_3, N_2O, NO_x, NleSW, NleGW
45	N input at which the yield does not further respond	Ninmx	NCU	NH_3, N_2O, NO_x, NleSW, NleGW
46	Soil C/N ratio	CNso_ms	NCU	NleSW, NleGW
	Soil emission data			
47	NH_3 emission factors from soil systems for all manure types	fNemap_NH3	NCU	NH_3
48	N_2O emission fractions from soil inputs	fNemsi_N2O	NCU	N_2O
49	Ratio between NO_x and N_2O emission fractions	rNON2O	NCU	N_2O, NO_x, NleSW, NleGW
	Leaching, runoff and climatic data			
50	N leaching fractions from the soil	fNle	NCU	NleSW, NleGW
51	N leaching fractions from stored manure	Flems	NUTS	NleSW, NleGW
52	Surface runoff fractions	Fsr	NCU	NleSW, NleGW
53	Sub-surface runoff fractions	Fro	NCU	NleSW, NleGW
	CH4 emissions			
54	a (dairy cattle)	a_CH4_ca	NUTS	CH_4
55	b (dairy cattle)	b_CH4_ca	NUTS	CH_4
56	CH_4 emission other cattle	CH4_oc	NUTS	CH_4

3.3 Quantifying the uncertainty in the APs

All the information regarding the pdfs of the model inputs, with their spatial- and cross-correlations was delivered by the INTEGRATOR group and it was obtained partly from European datasets and partly from expert knowledge.

The approach which was followed for the quantification of the uncertainty in the model inputs was described in Kros et al. (2010). For every AP, it was defined:

1. The distribution type (normal , lognormal)

The normal and the log-normal distribution were chosen to facilitate the UQ/UA procedure. In case of lognormally distributed APs the uncertainty was defined in terms of their $\log_e$ (natural logarithm) transformed values.

2. The standard deviation (SD) or the coefficient of variation (CV = SD/mean)

Following expert judgment, either the SD or the CV was specified, depending on the user's preferences. In some cases it was easier to specify an absolute measure for the error (SD) whereas in other cases it made more sense to specify an error as a proportion of the mean (CV).

3. The Min and Max values

Minimum and maximum values of the APs were set:

- in case of fractions (i.e. N leaching fractions from the soil, fNle) a minimum of 0 and a maximum of 1 were used;
- in other cases minimum and maximum values were not specified but a physical minimum (generally 0, sometimes minus infinity) and a physical maximum (generally infinity) were used instead.

4. The spatial-correlations

The common geostastitical approach to include spatial correlations in the UQ/UA procedure would be the definition of semivariograms and thus for each AP to estimate the sill, the nugget, the range and the model (e.g. spherical, exponential etc.) as shown in Figure 5. The semivariogram is a mathematical function which indicates spatial-correlation between measurements at sample locations. It is commonly represented as a graph that shows the semi-variance in measure with distance between all pairs of sampled locations.

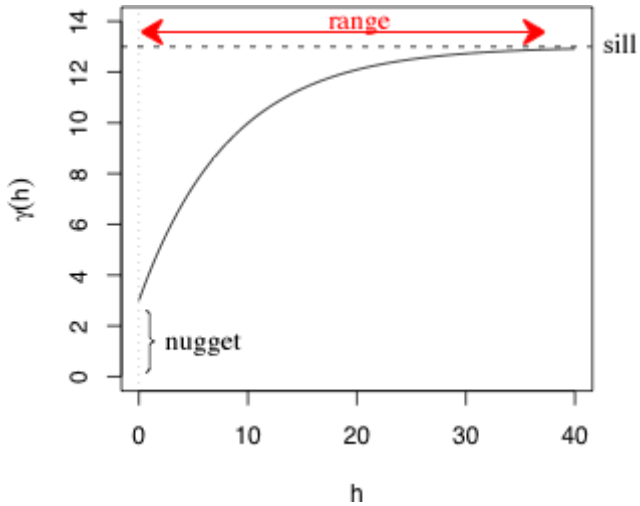


Figure 5. Example of a semivariogram where h=distance and γ(h)=semi-variance (Source:http://www.unc.edu/courses/2007spring/enst/562/001/images/lectures/lecture41/semivariogram.gif)

However, since no data were available to derive these semivariograms, it was decided to include spatial correlations in the UQ/UA in a more pragmatic way as it is recommended by Lesschen et al. (2007). In that paper, it is explained how spatial correlation can be included in an uncertainty analysis of a soil nutrient balance, by assuming that the variance of 1km grid-cell values is constant within a 20km grid-cell.

In accordance with Lesschen et al. (2007) it was assumed that the values of an AP were constant within an NCU, implying that the spatial correlation coefficient ρ_{plot}, between the plots within an NCU, was equal to 1. So, by assuming that the value of an AP at the *i*th plot, Y_i, had a variance (σ^2) that was equal for all plots within an NCU, the aggregated value of the AP for an NCU, $Yagg_{NCU}$, was calculated by formula 2:

$$Y_{agg}{}_{NCU} = \frac{1}{n}\sum_{i=1}^{n} Y_i \qquad (2)$$

where n was the number of aggregated plots within the NCU.

Next, the variance of $Yagg_{NCU}$ was calculated by formula 3, where Lesschen et al. (2007) have incorporated the degree of spatial correlation as an effect on the variance of the aggregated values:

$$\sigma^2_{Y\ agg\ NCU} = \left(\frac{1}{n}\right)^2 \times \left(n\sigma^2 + \left(n^2 - n\right)\rho_{plot}\,\sigma^2\right) \qquad (3)$$

where ρ_{plot} was equal to 1. So formula 3 could be finally written as:

$$\sigma^2_{Y\ agg\ NCU} = \sigma^2 \qquad (4)$$

Since the number of plots within an NCU was extremely large, even for relatively small NCUs, it was assumed that n was infinite. Thus if in formula 3 the ρ_{plot} was 0, then it would be written as:

$$\sigma^2_{Y\ agg\ NCU} = 0 \qquad (5)$$

Formula 5 shows that the variance of the aggregated value for an NCU would have vanished if the spatial correlation between plots had been ignored.

Based on the concept described above, it was decided the spatial correlation to be defined within the spatial levels distinguished in INTEGRATOR: NCU, NUTS2/3, and Country, with Europe as a whole being the model extent (see the spatial levels in an exemplar map of the Netherlands in Figure 6).

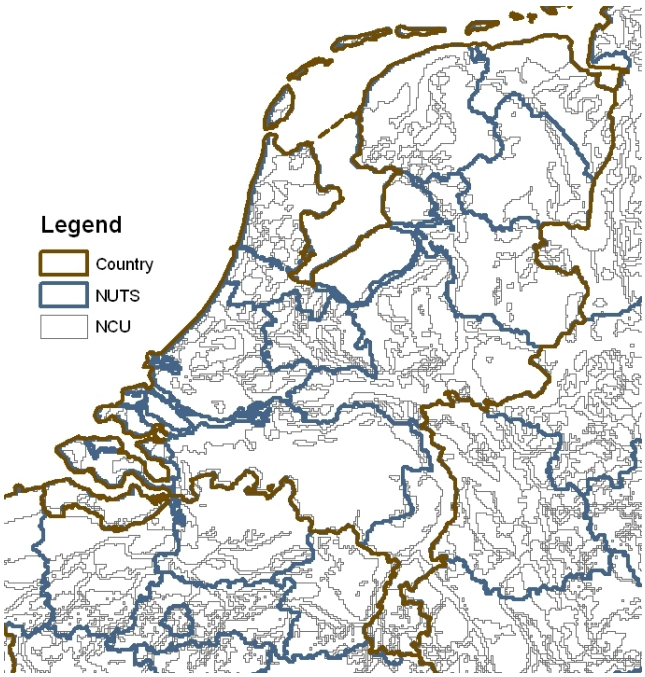


Figure 6. Map of the Netherlands depicting the NCU, NUTS, and Country spatial aggregation levels

For each AP, spatial-correlation coefficients were set between plots in different:

- NCUs within the same NUTS2/3 region, that was ρ_{NCU}
- NUTS2/3 regions within the same country, that was ρ_{NUTS}
- Countries within Europe, that was $\rho_{COUNTRY}$

The closer the plots were located, the stronger the spatial correlation was between them and so the spatial-correlation coefficients satisfied: $\rho_{NCU} \geq \rho_{NUTS} \geq \rho_{COUNTRY}$

In Figure 7 various scatter plots with corresponding spatial-correlation coefficients are illustrated.

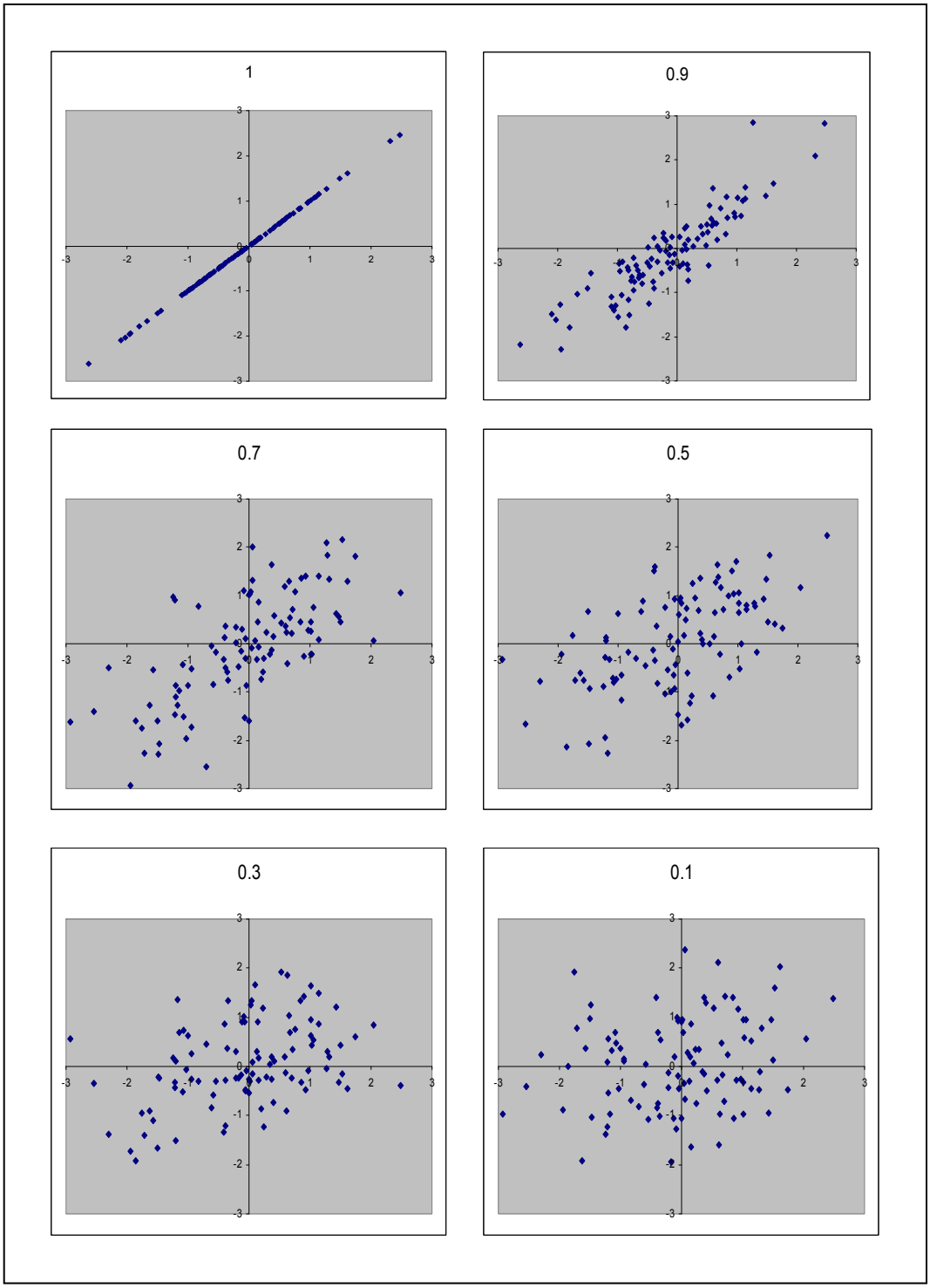


Figure 7. Scatter plots showing the strength of correlation when the correlation coefficient takes on values from 0.1 up to 1 (Source: Kros et. al. 2010)

Depending on the spatial level the APs were linked to, the following approach was followed:

- For NCU parameters, all the ρ_{NCU}, ρ_{NUTS} and $\rho_{COUNTRY}$ were taken into account
- For NUTS parameters, only the ρ_{NUTS} and $\rho_{COUNTRY}$ were taken into account while the ρ_{NCU} was considered to be 1 (i.e. all the NCUs within an NUTS2/3 region were assumed to be perfectly correlated)
- For CNTRY parameters, only the $\rho_{COUNTRY}$ was taken into account while the ρ_{NCU} and the ρ_{NUTS} were both considered to be 1 (i.e. all the NCUs within an NUTS2/3 region, and also all the NUTS2/3 regions within a country were assumed to be perfectly correlated)

The values of ρ_{NCU}, ρ_{NUTS} and $\rho_{COUNTRY}$ are discussed further in Section 3.3.1.

With the above procedure the spatial correlation was preserved during the uncertainty propagation analysis, and the uncertainty assigned to each NCU was finally incorporated in the model outputs which were aggregated to Country and European level.

5. The cross-correlations

Cross-correlations could have been defined by cross-variograms but due to lack of data on them it was decided cross-correlations coefficients, between related APs, to be finally used which were obtained through elicitation of expert knowledge.

An example of an AP correlated with another AP, was the N excretion rate from cattle (fNexf_ca) which was correlated with N content in grass (ctNplmx_gi) because higher N intake through fodder can lead to higher excretions. Another example was the correlation between Yield (YIeldopt_gi) and N content in grass. The product of Yield × N content, representing the N removal from the soil, should not exceed a maximum amount, specified in INTEGRATOR. Therefore a negative correlation was assumed for this pair of parameters.

The selected pairs of APs, that were either positively or negatively correlated, are given in Table 3.

Table 3. APs for which cross-correlations were considered

AP $_i$	**AP $_j$**	$\rho_{cc}(\mathbf{i},\mathbf{j})$
Nexf_ca	ctNplmx_gi	0.5
Yieldopt_gi	ctNplmx_gi	-0.8
fNemhs_NH3	fNemhss_N2O	0.5
fNemhss_N2O	fNemhss_NO	0.8

Since so far, the spatial-correlation coefficients for each AP between different locations: ρ_{NCU}, ρ_{NUTS} and $\rho_{COUNTRY}$ (vectors A in Figure 8) and also the cross-correlation coefficients between different APs at the same location: $\rho_{cc}(\mathbf{i},\mathbf{j})$ (vectors B in Figure 8) had been defined, the next step was to calculate the spatial cross-correlation coefficients between an AP_i at one location with another AP_j at another location (vectors C in Figure 8). In order to limit the number of statistical parameters required for this calculation, it was assumed that the spatial cross-correlation

coefficients: ρ_{NCU}**(i,j)**, ρ_{NUTS}**(i,j)**, and $\rho_{COUNTRY}$**(i,j)** could be calculated by the following formulas:

$$\rho_{NCU}(\mathbf{i},\mathbf{j}) = \rho_{cc}(\mathbf{i},\mathbf{j}) \cdot \sqrt{\rho_{NCU}(\mathbf{i}) \cdot \rho_{NCU}(\mathbf{j})} \quad (6)$$

$$\rho_{NUTS}(\mathbf{i},\mathbf{j}) = \rho_{cc}(\mathbf{i},\mathbf{j}) \cdot \sqrt{\rho_{NUTS}(\mathbf{i}) \cdot \rho_{NUTS}(\mathbf{j})} \quad (7)$$

$$\rho_{COUNTRY}(\mathbf{i},\mathbf{j}) = \rho_{cc}(\mathbf{i},\mathbf{j}) \cdot \sqrt{\rho_{COUNTRY}(\mathbf{i}) \cdot \rho_{COUNTRY}(\mathbf{j})} \quad (8)$$

where **i** and **j** referred to **i** and **j** APs.

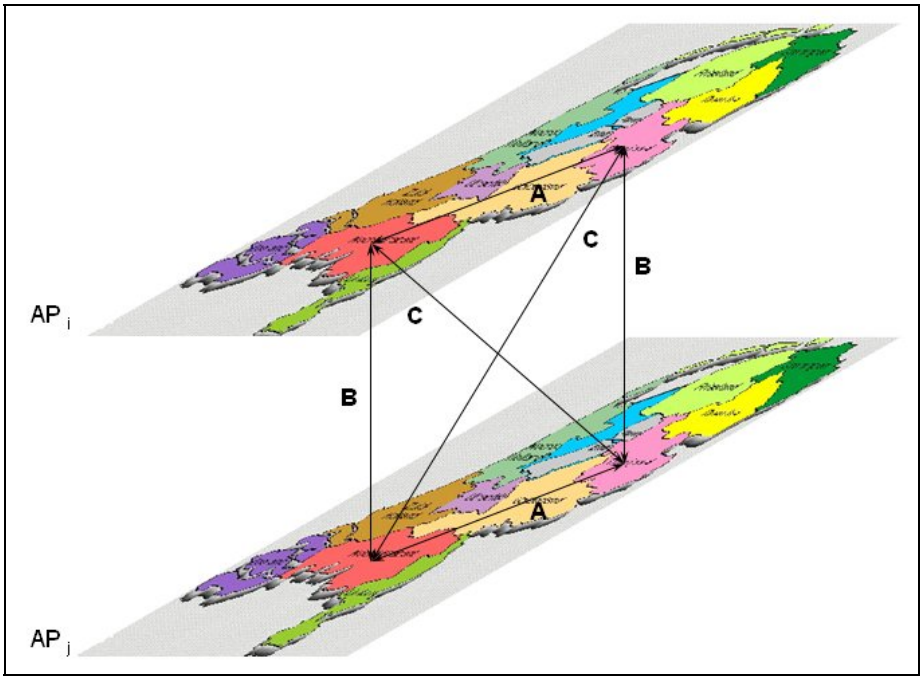


Figure 8. Schematic representation of spatial-correlations (A), cross-correlations (B), and spatial cross-correlations (C) using an exemplar map of the Netherlands where the 12 provincial NUTS regions are depicted in different colours

3.3.1 Building statistical models (joint pdfs) for the uncertainty in the APs

3.3.1.1 Classes of CVs and SDs

As it has been mentioned already, for some APs the CV was defined and for some others the SD. Due to limited quantitative information on the uncertainty in the APs, only three classes of CVs and SDs were used during the definition of their pdfs (for SDs only in case of parameters defined as fractions, in other cases the absolute value of the SD was used):

- Low: 0.10, used for parameters:
  - based on good quality statistics on agronomic data: animal numbers; national N fertilizer inputs
  - N input with no change in yield
  - Uptake fraction
- Moderate: 0.25, used for all parameters, except for these cases mentioned at high and low classes

- High: 0.50, used for guestimated input parameters:
  - N2O emission fractions from solid systems
  - NOx / N2O ratio
  - N fixation for arable and grass

3.3.1.2 Robustness analysis of CVs and SDs

Since the information on the assigned CVs or SDs was rather uncertain, a robustness analysis was carried out by using three uncertainty scenarios with Optimistic (Opt), Reference (Ref) and Pessimistic (Pes) values for the assumed input uncertainties. For the CVs, explicit values were assigned for each scenario (see Table 4). For the SDs, three different factors were used for each scenario respectively (see Table 5). For parameters that were not defined as fractions, the standard values of the SDs were used as Reference values. For parameters that were defined as fractions, the values which were set for the Reference scenario of the CVs were used as Reference values of the SDs as well.

Table 4. Values of CV classes according to three robustness scenarios

Class of CV	Opt	Ref	Pes
Low	0.05	0.10	0.15
Moderate	0.10	0.25	0.30
High	0.40	0.50	0.60

Table 5. Values of SD classes according to three robustness scenarios

Class of SD	Opt	Ref	Pes
Low	0.5×SD	SD	1.5× SD
Moderate	0.5×SD	SD	1.5× SD
High	0.5×SD	SD	1.5× SD

3.3.1.3 Classes of spatial-correlations

Same as with the CVs and SDs, the available information for the spatial correlations was also limited and finally only five classes of spatial-correlation coefficients were decided upon:

- Perfect: 1, was used for ρ:
  - when a perfect correlation could be assumed (i.e. for APs that were not linked to the NCU but to a higher aggregation level, e.g. NUTS or Country level, then the ρ_{NCU} and/or ρ_{NUTS} were assumed to be 1)
- High: 0.85, was used for ρ:
  - when it was obvious that a serious correlation existed
- Moderate: 0.5, used for ρ:
  - when it was obvious that a correlation existed
- Low: 0.2, was used for ρ:
  - when there were slight indications that a correlation existed
- None: 0, was used for ρ:
  - when it was obvious that absolutely no correlation existed

3.3.1.4 Robustness analysis of spatial-correlations

Three robustness scenarios were also applied to the spatial-correlations coefficients in order the robustness of their assigned values to be investigated (see Table 6).

Table 6. Values of spatial-correlations classes according to three robustness scenarios

Class of spatial-correlations	Opt	Ref	Pes
Perfect	1	1	1
High	0.8	0.85	0.9
Moderate	0.3	0.5	0.7
Low	0.1	0.2	0.3
None	0	0	0

The three different tables with the values reflecting the CV, SD and spatial-correlations classes according to three robustness scenarios, for the 56 APs, are given in Appendix 1.

3.3.2 Generating multiple simulations of APs values by repeated sampling from their pdfs

The first step of the MC method is the generation of multiple inputs simulations randomly drawn from the inputs´ pdfs. This could be achieved by using the "**rmultnorm**" function (Multivariate Normal Random Number Generator, see further details below) incorporated in the statistical software R (http://www.r-project.org/) which is an open-source environment for statistical computing and visualization based on S statistical programming language. This generator is based on the multivariate normal distribution which is defined as follows: a multivariate random vector X with k elements, $X = (X_1, X_2,.. X_k)$, that follows a normally distribution is denoted as $X \sim N_k(\mu, R)$ where μ = mean vector and R = variance-covariance matrix (k x k). The R matrix, with all combinations of X_i and X_j pairs, is calculated with formula 9:

$$R[i, j] = \rho[i, j].\sigma_i.\sigma_j \quad (9)$$

where $\rho_{[i, j]}$ is the correlation coefficient matrix for X_i and X_j pairs, and σ_i , σ_j are the SDs of X_i and X_j respectively.

The 56 APs were multivariate random vectors whose values varied over the 35101 NCUs of EU-25. Therefore multiple input simulations could be drawn from the multivariate normal pdf of each AP as long as their mean vectors and their variance-covariance matrices were known.

The APs were either normally or lognormally distributed. So a normally distributed AP could be denoted as x(i), and a lognormally distributed AP could be denoted as $y(i)=e^{x(i)}$ which implied that x(i) was normally distributed. Thus the characterization of all APs (normal or lognormal) could be done through the characterization of the pdf of x(i).

The pdf of the x(i), as explained above, could be defined by the mean $\mu(i)$, the standard deviation $\sigma(i)$, and the spatial-correlation or spatial cross-correlation coefficient matrix.

By applying the common statistical model that is used to convert original observed values of a variable to their standardised values each of the x(i) could be written as:

$$x(i) = \mu(i) + \sigma(i) \cdot \xi(i) \quad (10)$$

where $\xi(i)$ is a standard normal random variable (i.e. normally distributed random variable with $\mu(i) = 0$ and $\sigma(i) = 1$). So a simulation of x(i) could be obtained by multiplying the simulated $\xi(i)$ with the $\sigma(i)$ and then adding the $\mu(i)$ to the product. (This addition could take place in INTEGRATOR and not in R environment because only in INTEGRATOR were the $\mu(i)$ values available).

In the case that CV(i) of x(i) was known instead of the $\sigma(i)$, the above formula could be written as:

$$x(i) = \mu(i) + CV(i) \cdot \mu(i) \cdot \xi(i) = \mu(i) \cdot (1 + CV(i) \cdot \xi(i)) \quad (11)$$

In this case, a simulation of x(i) could be obtained by first simulating the $\xi(i)$, multiplying it with the CV(i), and then adding 1 to it. Next, the $\mu(i)$ (in INTEGRATOR) could be added to the result.

Since the relationship between x(i) and $\xi(i)$ was linear, correlations between the x(i) of different NCUs were equal to those between the $\xi(i)$ of these NCUs. Consequently the correlations as they were specified for the x(i) could be used for the simulations of the $\xi(i)$. These simulations could be generated by using the "**rmultnorm**" function (http://rss.acs.unt.edu/Rdoc/library/MSBVAR/html/rmultnorm.html) which generates multivariate normal random simulations for given mean and variance-covariance matrices. The usage of the function is the following:
`rmultnorm (n, mu, vmat, tol = 1e-10)`, and its arguments are:

- **`n`**: number of simulations to generate
- **`mu:`** vector containing k multivariate means
- **`vmat:`** (k x k) covariance matrix
- **`tol:`** tolerance level used for SVD of the covariance

Since for $\xi(i)$ the $\mu(i) = 0$ and $\sigma(i) = 1$, then for $\xi(i)$ formula 9 could be written as:

$$R[i, j] = \rho[i, j] \quad (12)$$

As a result, the generation of the multiple simulations of $\xi(i)$ could be obtained by using the spatial-correlation or spatial cross-correlation coefficient matrix ρ [i,j] as the **`vmat`** argument, and a vector of k 0s as the **`mu`** argument.

For lognormally distributed parameters the simulated x(i) could be subsequently transformed to simulations of y(i) by calculating $y(i)=e^{x(i)}$. This could be implemented in INTEGRATOR because in order the exponent x(i) to be calculated the $\mu(i)$ is needed.

Also within INTEGRATOR, the simulated values of APs below the Min or above the Max values, as specified for each AP, could be replaced by the Min or the Max value itself. Note that for lognormally distributed values the Min and the Max were specified for the y(i) instead of x(i).

Another important aspect of the MC method is the number of the simulations required to ensure that the broad range of the values for each AP is sufficiently represented in the analysis. The method which was adopted to estimate this number was to check the cumulative frequency curves of the differences between the values of the preset CVs or SDs and the values of the CVs or SDs resulting from a first sample of generated input simulations of the APs from NUTS group. After a check with two samples took place (further information about it is provided in Section 4.1.1) it was decided that a sample of 1000 simulations was sufficient to represent the uncertainty in the APs.

3.3.2.1 Building spatial-correlation coefficient matrices

The values of the 56 APs were linked to each one of the 35101 NCUs so a spatial-correlation coefficient RHO matrix of 35101 rows x 35101 columns had to be built for each AP with the ρ_{NCU}, ρ_{NUTS} and $\rho_{COUNTRY}$ incorporated.

First of all, the values in the RHO matrix were filled in with the $\rho_{COUNTRY}$ which was the spatial-correlation coefficient between plots in different countries within Europe (all NCUs were located in Europe anyway). Then the diagonal of the matrix was replaced with 1s since the correlation of an NCU with itself is always 1. Then it was checked whether individual pairs of NCUs were located within the same country, if they were, then the value in the matrix was replaced with the ρ_{NUTS} coefficient. Finally, it was checked whether the NCU pairs belonged to the same NUTS region, if they were, then the value in the matrix was replaced with the ρ_{NCU} coefficient. This process however could not be finally executed for all NCUs, (a subset of 100 NCUs was initially checked) because it exceeded the computation and memory capacities of the available hardware. Thus it was not feasible to use the "**rmultnorm**" function to produce the RHO matrix for the 35101 NCUs. For this reason it was decided that a different method had to be applied which demanded though high skilled programming that was beyond the scope of this research. The new method was developed by Gerard Heuvelink, Tom Hoogland, and Dennis Walvoort (all three from Alterra, WUR) and the simulated values of APs linked to the NCU level were provided by them.

The above problem however could be limited only to the 23 APs which were linked to NCU level and therefore all the NCUs with their spatial-correlation coefficients (ρ_{NCU}, ρ_{NUTS} and $\rho_{COUNTRY}$) had to be included in the building of the RHO matrices. For the other three groups of the APs, a few assumptions were made which allowed the building of the RHO matrices by using the "**rmultnorm**" function.

For the 26 APs that were linked to NUTS level it was assumed they had constant values between NCUs contained within a NUTS region, this meant ρ_{NCU} =1. So it was sufficient to simulate just one NCU from each NUTS region because all the other NCUs within the same region would have the same value. There were 688 NUTS regions for EU-25. Subsequently the dimension of the RHO matrix was reduced to 688 rows x 688 columns and its generation was feasible by the available computers. Since for these APs the ρ_{NCU} =1, only the ρ_{NUTS} and $\rho_{COUNTRY}$ were incorporated in the RHO matrix. The values in it were filled in with the ρ_{NUTS} or $\rho_{COUNTRY}$ coefficients depending on where individual pairs of NUTS regions were located. Initially all values in the matrix were filled in with the $\rho_{COUNTRY}$ (all NUTS were located in Europe anyway). Then the diagonal of the matrix was replaced with 1s. Then it was checked whether pairs of NUTS regions were located within the same

country, and if they were then the value in the matrix was replaced with the ρ_{NUTS} coefficient.

In the same way, for the five APs that were linked to CNTRY level, it was assumed that they had constant values between the NCUs contained within a NUTS region, (ρ_{NCU} =1), and also that they had constant values between the NUTS regions contained within a country, this meant that ρ_{NUTS} =1. So it was sufficient to simulate only 1 NCU from a NUTS region, and only 1 NUTS region from each country. The number of the countries was 25, so the dimension of the RHO matrix was further reduced to 25 rows x 25 columns and only the $\rho_{COUNTRY}$ was incorporated in the generation of the matrix. The values were initially filled in with the $\rho_{COUNTRY}$ coefficient (all individual CNTRY pairs were located within Europe). Finally the diagonal of the matrix was replaced with 1s.

For the two APs which were linked to European level (GEN), it was assumed that they had constant values between countries, between NUTS regions and between NCUs and that meant all the coefficients: ρ_{NCU}, ρ_{NUTS} and $\rho_{COUNTRY}$ were 1. So the RHO matrix for each AP from GEN group was of 1 row x 1 column dimension and its single value was 1.

Figure 9 provides an overview of the procedure followed for building the RHO matrix of each AP from NUTS group.

Building the RHO matrix of an AP from NUTS group

$$RHO = \begin{array}{c} \\ NUTS_1 \\ NUTS_2 \\ \vdots \\ NUTS_{688} \end{array} \begin{array}{c} \begin{array}{cccc} NUTS_1 & NUTS_2 & \cdots & NUTS_{688} \end{array} \\ \begin{bmatrix} 1 & \rho_{NUTS} & & \rho_{COUNTRY} \\ \rho_{NUTS} & 1 & & \\ & & \ddots & \\ \rho_{COUNTRY} & & & 1 \end{bmatrix} \end{array}$$

Steps followed:

1. Initialize with $\rho_{COUNTRY}$
2. Fill in the diagonal with 1s
3. Fill in with ρ_{NUTS} if pairs of NUTS regions are within the same country (e.g. pair of $NUTS_1$ and $NUTS_2$ regions in the matrix)

Figure 9. RHO matrix for a NUTS parameter and the steps followed for its creation

3.3.2.2 Building spatial cross-correlation coefficient matrices

Due to the fact that among the APs being correlated with others there were two of them (AP_{36}: Yieldopt_gi and AP_{40}: ctNplmx-gi) that were linked to NCU level, it was

decided the cross-correlations to be incorporated in the method particularly developed for the APs at NCU level. Consequently, the creation of the spatial cross-correlation coefficient matrices were not part of this research but spatial cross-correlations were taken into account, during the analysis of the uncertainty propagation, since they were included in the simulated input values for the APs linked to NCU level.

3.3.2.3 Developing R-scripts for the generation of 1000 input simulations for each group of the APs by incorporating the three robustness scenarios

The 1000 input simulations for each AP linked to NUTS, CNTRY, and GEN level, by taking into account the three robustness scenarios, were produced by nine different scripts developed within the R environment: three scripts representing each scenario for each of the three groups of APs. The 1000 input simulations for each of the nine combinations as shown in Table 7 were stored in the so-called MC input files.

Table 7. The nine R-scripts for the simulation of the APs

R_SCRIPTS	Groups of Agricultural Parameters		
	NUTS	CNTRY	GEN
Scenarios	Opt	Opt	Opt
	Ref	Ref	Ref
	Pes	Pes	Pes

The steps of the algorithm for the generation of the MC input files were executed repeatedly for each AP in batch mode. This algorithm (e.g. for the NUTS group, concerning the Reference scenario), as well as the names, and the format of the MC input files required for their linkage with INTEGRATOR are described in Appendix 2.

One of the nine written scripts, i.e. the script for NUTS parameters concerning the Reference scenario, is provided in Appendix 3. The code in the other eight scripts was almost identical with the only difference being the “path” where the data, required for the generation of the MC input files, were stored.

An example with one MC input file for the five APs of the CNTRY group is given in Appendix 4.

3.3.2.4 Running the INTEGRATOR model for the generated input simulations in batch mode and storing its output simulations (UQ)

The 9000 input files, produced by the nine R-scripts mentioned above, along with the 3000 input files for the APs of the NCU group, provided by the Alterra team, were imported to INTEGRATOR. The model was run separately in three rounds, one for each robustness scenario. Within every round, INTEGRATOR was provided with 4000 input files, 1000 from each of the four groups of the APs. At the end of the model, 6000 output files, 1000 for each of the six output fluxes (i.e. NH_3, NO_x, N_{leSW}, N_{leGW}, CH_4 and N_2O) were produced at Country and at European level (see Figure 10).

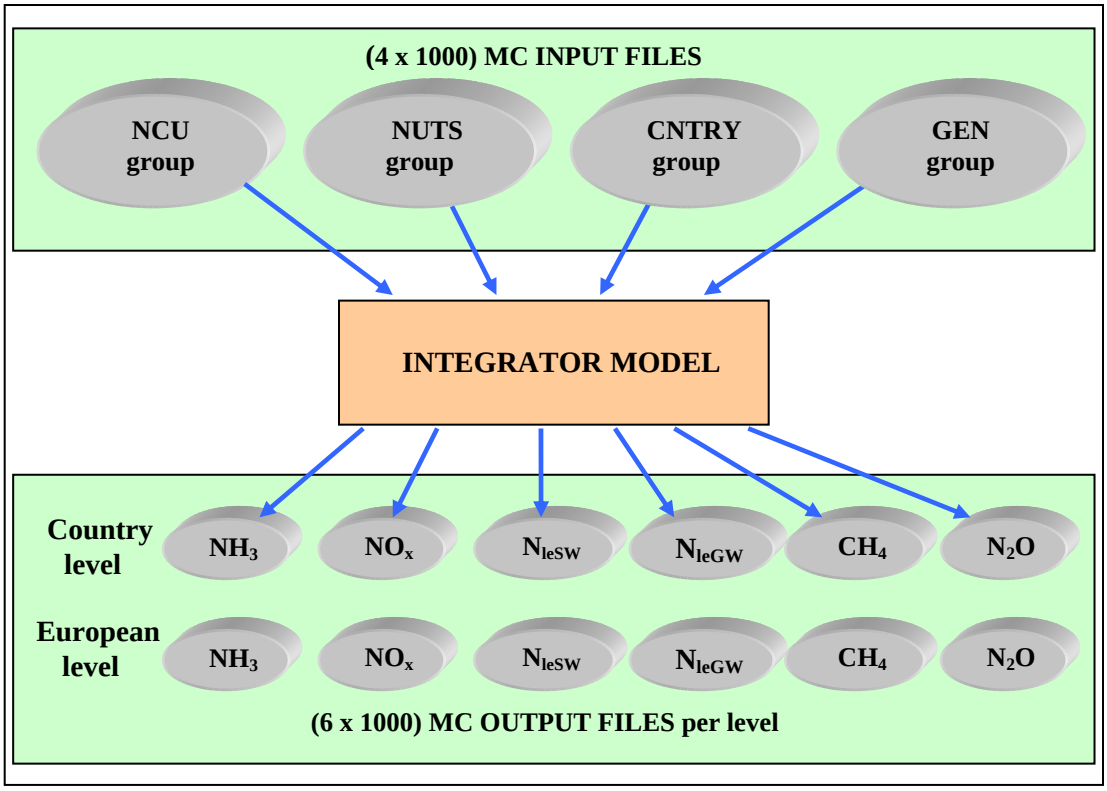


Figure 10. Setting of INTEGRATOR model for each robustness scenario

Three R-scripts, one for each robustness scenario, were written to run INTEGRATOR for the 12000 input files in three rounds. In each of these R-scripts, the INTEGRATOR dll (dynamic link library) was loaded and in this way the model was called within the R-script. This dll was provided by the INTEGRATOR group. The 4000 input files from each individual group were supplied to INTEGRATOR repeatedly within a "for loop". Additionally, the code in the script was built in such a way that the produced output simulations were stored automatically in output files within subdirectories that were structured in an organised manner to ensure a transparent access to them.

The steps of the algorithm followed for running the model in batch mode and storing its MC output files, and also the names and the format of these files are described in Appendix 2.

One of the three written R-scripts, i.e. the script for the Optimistic scenario, is given in Appendix 5.

Two MC output files, e.g. for N_2O emissions, one at County level and one at European level are given in Appendix 6.

3.4 Analysing the contribution of individual uncertainty sources to the overall output uncertainty (UA)

During this step, which concerned the Uncertainty Analysis, the uncertainty contribution of each individual group of the APs to the total output uncertainty was estimated. This was accomplished by running the model again but this time considering uncertain only the group of APs for which the contribution had to be estimated. The other three groups remained "certain" by keeping their default

averaged values which were stored in INTEGRATOR´s database. Therefore the model was run again in four separate rounds, and each time only one of the four groups was made uncertain (see Table 8). The uncertainty contribution of each group was calculated by dividing the variance of the outputs estimated when only this group was made uncertain by the variance of the outputs calculated when all four groups were uncertain during the first time the model was run (i.e. step of UQ). The outcome of each division was the uncertainty contribution of each individual group, respectively, to the overall output uncertainty.

Table 8. Four INTEGRATOR rounds during each of them only one group of APs was made uncertain

INTEGRATOR ROUND 1		INTEGRATOR ROUND 2		INTEGRATOR ROUND 3		INTEGRATOR ROUND 4	
Uncertain	**Certain**	**Uncertain**	**Certain**	**Uncertain**	**Certain**	**Uncertain**	**Certain**
NCU group			NCU group		NCU group		NCU group
	NUTS group	NUTS group			NUTS group		NUTS group
	CNTRY group		CNTRY group	CNTRY group			CNTRY group
	GEN group		GEN group		GEN group	GEN group	

3.5 Summarizing the MC outputs

The final step, after Uncertainty Quantification and Uncertainty Analysis had been completed, was to summarise the simulated values in the MC output files by estimating their statistical measures such as means, CVs, quartiles, interquartile range etc., and then use visualization tools in order to communicate these statistics to end-users. This was achieved by splitting the summation step into the following A and B sub-steps:

A. Check of the effect that:

1. The spatial aggregation to Country, and European level had on the output uncertainty by examining the:
   - Outputs at European level, for the Reference scenario, by means of box-plots and column-line charts;
   - Outputs at Country level, for the Reference scenario, by means of box-plots for five representative EU-25 countries;
   - Outputs at Country level, for the Reference scenario, by means of column-line charts and choropleth maps for all countries of EU-25;
2. The three robustness scenarios had on the output uncertainty by examining the outputs at European level, for each scenario, by means of box-plots and column charts.

B. Estimation of the uncertainty contribution of the four individual groups of APs to the overall output uncertainty for the five chosen EU-25 countries and for the whole Europe as well, for the Reference scenario. The visualizations which were adopted here were stack-columns.

The R programme was mainly used for the calculation and visualization of the statistics of the outputs, except for the choropleth maps that were produced with ArcMap 9.3.1., and the column charts that were produced with Microsoft Excel 2003.

The R-script for the calculation of means, SDs and CVs of the outputs aggregated to Country level, for the Reference scenario, is provided in Appendix 7.

The R-script for the box-plots (and the 5^{th}, 25^{th}, 50^{th}, 75^{th}, and 95^{th} percentiles) of the outputs aggregated to European level, for the three robustness scenarios, is provided in Appendix 8.

The R-script for the stack-columns of the uncertainty contribution, for the five countries, for the Reference scenario, is provided in Appendix 9.

In the following paragraphs a short explanation about the visualizations tools is provided to help readers of this report interpret the statistical information they describe.

One of the tools used for the sub-step "A" was the box-plot (see Figure 13) which is a simple dispersion diagram commonly used in descriptive statistics. It is used to display the entire frequency distribution of a numerical dataset by dividing it into quartiles. The height of the box represents the interquartile range (IQR) which is the difference between the upper quartile (i.e. Q3 = 75th percentile) and the lower quartile (i.e. Q1= 25th percentile). The black thick line within the box depicts the median value (i.e. Q2 = 50th percentile). The two black, thinner horizontal lines below and above the box depict the minimum (Q1-1.5*IQR) and the maximum (Q3+1.5*IQR) values of the distribution respectively. The small circles above the maximum value and below the minimum are considered numerically distant from the rest of the data and they are called outliers. Outliers are often indicative either of a measurement error or a form of skewness (http://en.wikipedia.org/wiki/File:Boxplot_vs_PDF.svg).

All the box-plots are accompanied by tables with values of the 5th, 25th, 50th, 75th, and 95th percentiles which are given in Appendix 11.

Another tool, for sub-step "A", was the column chart depicting the CV level of each output (i.e. proportion of the SD to the mean). The CV is a relative measure of dispersion that allows direct comparison of relative uncertainty in different data sets (i.e. different mean values). In some cases, the CVs of the outputs were combined with their SDs in column-line charts (see Figure 14).

The visualization tool used for sub-step "B" was the "100%" stack-column which depicts in different colours the percentages of different categories that contribute to a total (see Figure 23).

Before the visualization of the uncertainty in the outputs took place, a distinction was made:

- The outputs that contribute to N increase: NH_3, NO_x, N_{leGW}, and N_{leSW} whose magnitude is measured in terms of their amount of N. The unit for N outputs is: N (Kg $ha^{-1}yr^{-1}$).
- The outputs that contribute to CO_2 increase: CH_4 and N_2O whose magnitude is expressed in terms of their Global Warming Potential (GWP). The unit for the GWP outputs is: CO_2 (Kg $ha^{-1}yr^{-1}$). The N_2O produced by INTEGRATOR is expressed as N contained in N_2O and in order to be converted to N_2O itself is multiplied with a factor of 44/28, where 44 is the molar weight of N_2O and 28 the molar weight of 2 × N. Then the result is multiplied with a factor of 298 in order the N_2O to be converted to GWP for a time horizon of 100 years based on 2007 IPCC 4th assessment report (see: http://www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-chapter2.pdf). The CH_4 produced by INTEGRATOR is already expressed as CH_4 (and not as C) and in order to be converted to GWP for a time horizon of 100 years is multiplied with a factor of 25. This means that in a horizon of 100 years, N_2O becomes approximately 19 ((44/28*298)/25) times more effective GHG than CH_4.

4. RESULTS AND DISCUSSION

4.1 Uncertainty in model inputs

In this Section, it is discussed how the test for the required sample size of simulations took place by means of cumulative frequency curves (Sub-section 4.1.1) and also it is checked whether the generated samples of APs were normally distributed and had met their predefined CVs or SDs (Sub-section 4.1.2). The corresponding R-scripts are given in Appendix 10.

4.1.1 Representativeness of the sample size

The group of the 26 NUTS APs, for the Reference scenario, were chosen for the test of the sample size and three of them, the AP_6: Nexf_co, the AP_{13}: fhs_co, and the AP_{28}: Area_int are discussed here. Initially a sample of 100 simulated values for each AP, linked to each one of the 688 NUTS regions, was produced. Then, the cumulative frequency curves of the differences between the values of the preset CVs and the CVs resulting from this sample were generated. These curves are depicted on the left side of Figure 11 along with their 5^{th}, 50^{th}, and 95^{th} percentiles. As it can be seen the sample with the 100 simulations produced cumulative curves where the values of the differences:

- were not yet (completely) normally distributed around 0 because the ranges between the 50^{th} and the 5^{th} percentiles were not yet equal to the ranges between the 95^{th} and the 50^{th} percentiles and the medians (i.e. corresponding values of x when Fn(x) = 0.5) varied between 0.003 - 0.008;
- their 90 % ranges (i.e. the ranges between the 95^{th} and 5^{th} percentiles) varied between 0.055 - 0.066.

These statistics though could be further improved by trying to achieve a much better normal distribution around 0 and also bring the 90% ranges even closer to 0. So a new iteration was done with a generated sample of 1000 simulations. As it can be seen from the curves on the right side of Figure 11, the new sample produced cumulative frequencies where the values of the differences:

- approximated much better a normal distribution around 0 since the ranges between the 50th and the 5th percentiles were almost equal to the ranges between the 95th and the 50^{th} percentiles and the medians varied only between -0.001 - 0.000;
- their 90 % ranges varied between 0.017 - 0.019.

The differences between the values of the preset CVs and the CVs resulting from the sample with the 1000 simulations were considered small enough and thus it was decided that this sample was big enough to describe the uncertainty in the model inputs. In Figure 11, the red vertical dotted lines indicate the shift of the medians much closer to 0 when the sample size was increased from 100 to 1000 simulations.

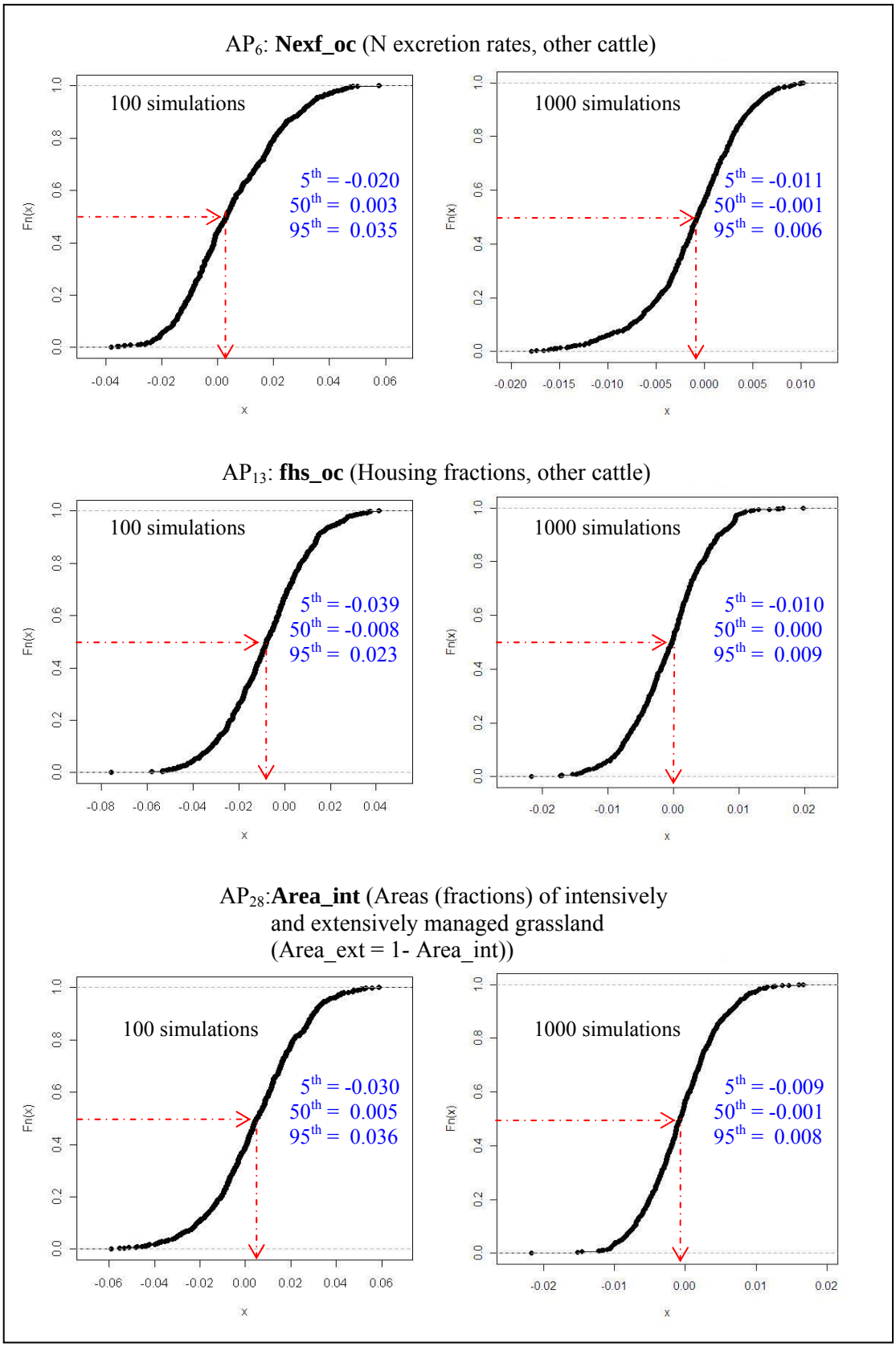


Figure 11. Cumulative frequency curves showing the differences of the preset CVs from the CVs resulting from a sample of 100 and 1000 simulations, respectively, of three APs from the NUTS group, for the 688 NUTS regions

4.1.2 Quantified uncertainty in APs of NUTS, CNTRY and GEN group

Two representative APs from each individual group, for the Reference scenario, were used in this Section, to verify their probability distribution, as well as their CVs or SDs resulting from the generated sample of 1000 simulated values.

In Table 9, the six APs are provided along with their distribution types and the values of their preset CVs or SDs.

Table 9. APs checked for their probability distributions and their CVs or SDs resulting from the sample of 1000 simulations

AP	Group	Distribution	CV	SD
fNemms	NUTS	Normal	0.25	
fNemhsls	NUTS	Lognormal		0.50
aninr_ca	CNTRY	Normal	0.10	
aninr_ca	CNTRY	Normal	0.10	
f_Nav_dep	GEN	Normal	0.25	
f_Nav_om	GEN	Normal	0.25	

The curves illustrated in Figure 12 show that the sample of 1000 simulations, for each of the above APs, was normally distributed, and the resulted CVs or SDs (coded as: "Res.CV or SD") had met their preset CVs or SDs (coded as: "Set CV or SD").

Note that the simulated values for the AP: fNemhsls produced a normal distribution and not a lognormal distribution because they were just products of $\xi(i)$ simulations multiplied with the preset SD (i.e 0.50) and not simulations of $y(i)=e^{x(i)}$ where $x(i) = \mu(i) + \sigma(i) \cdot \xi(i)$. The calculation of y(i) values was implemented in INTEGRATOR (see also Section 3.3.2).

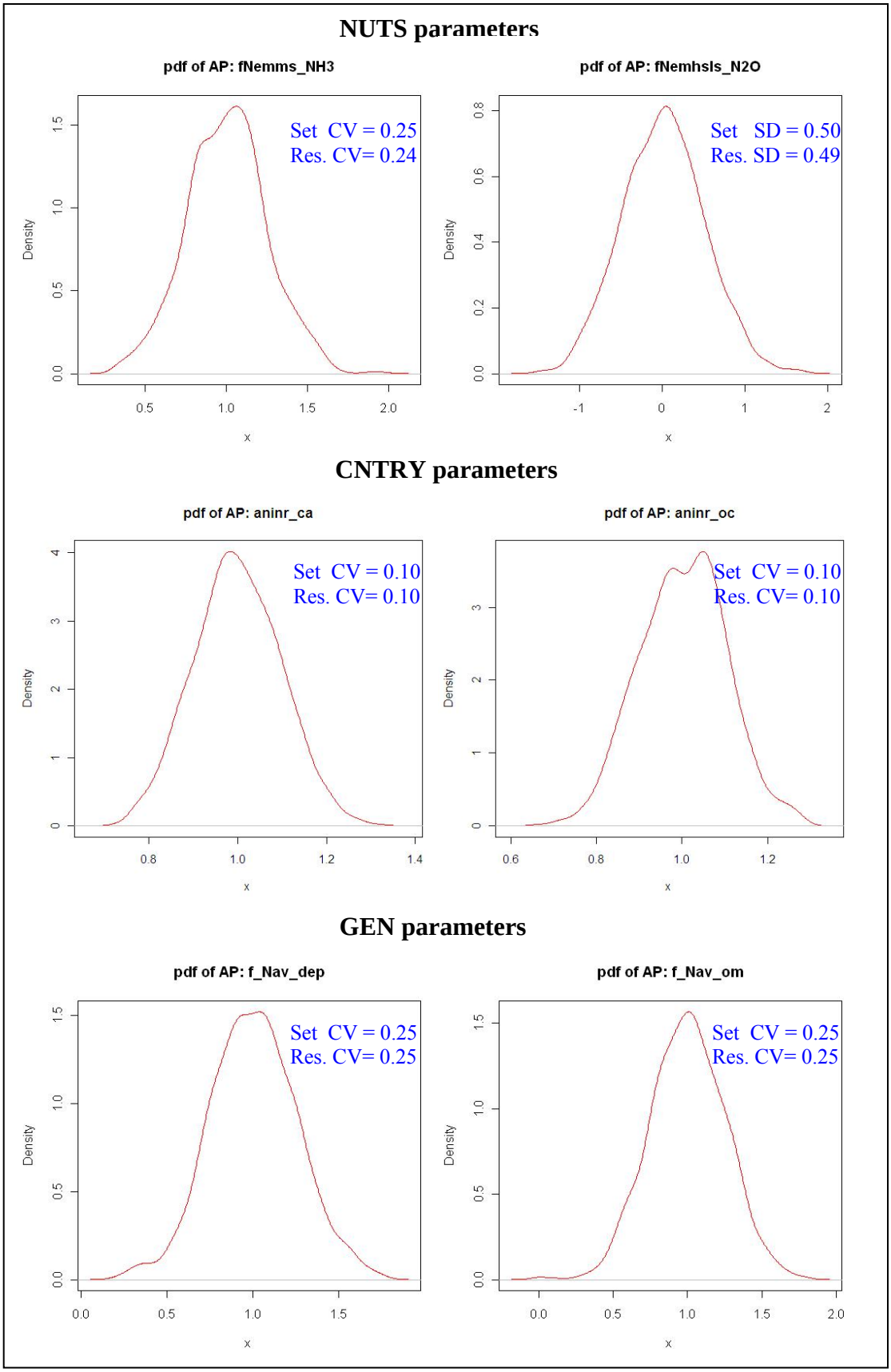


Figure 12. Probability distributions of two APs, from each individual group, produced from a sample of 1000 simulations, depicted along with their preset CVs or SDs and their resulted CVs or SDs respectively

4.2 Uncertainty in model outputs

In this Section, the results from the summation of the model outputs according to sub-step "A" (see also Section 3.5) are presented.

4.2.1 Effect of European level outputs for the Reference scenario

4.2.1.1 Results

The effect of European level on the output uncertainty was visualised by:

- Box-plots provided in Figure 13;
- Column-line charts provided in Figure 14.

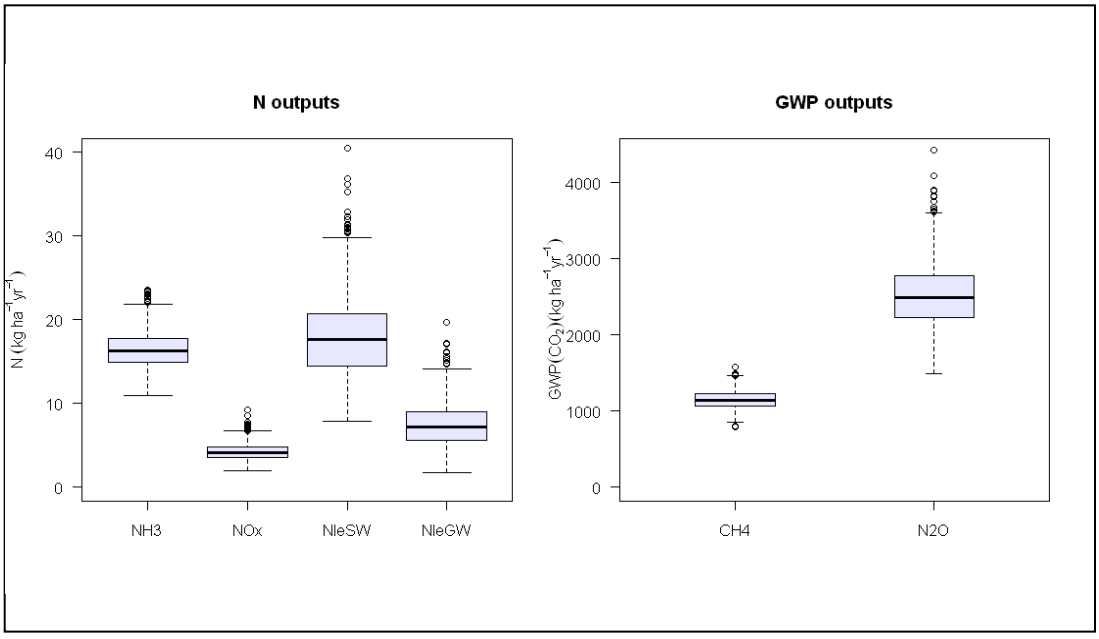


Figure 13. Box-plots displaying uncertainty (expressed as SD) in N and GWP outputs at European level

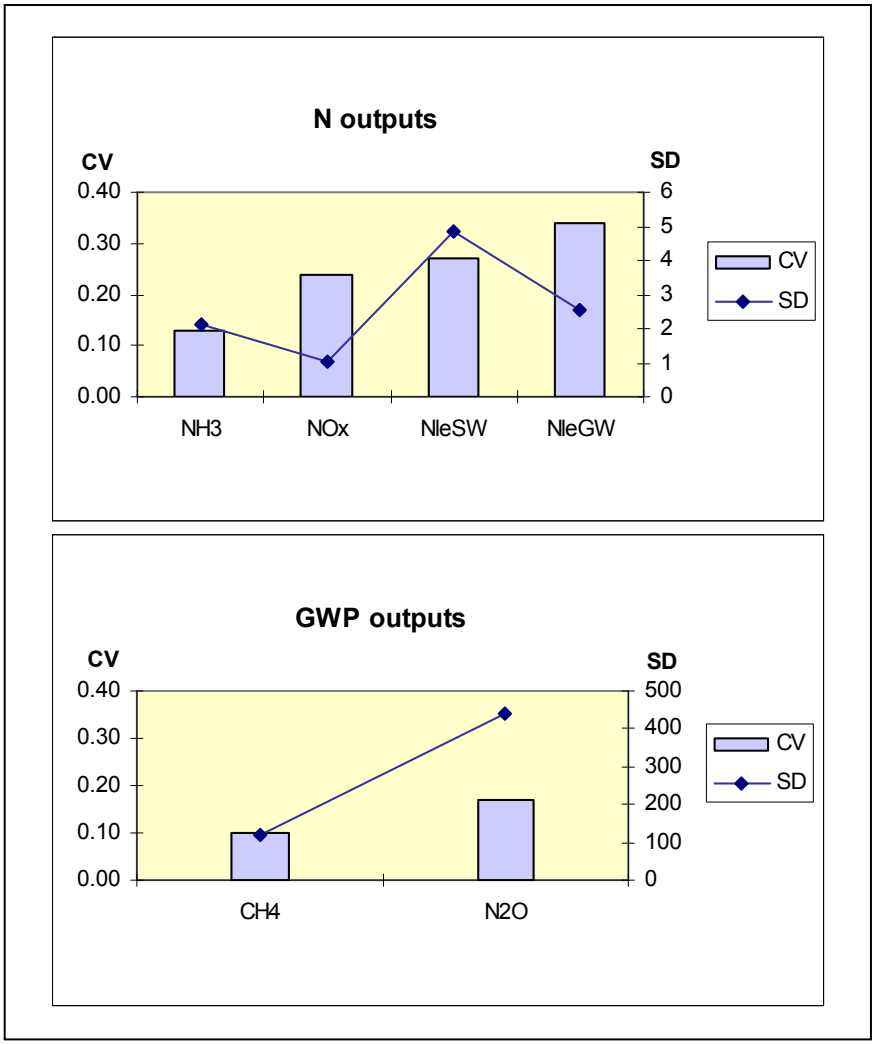


Figure 14. Column-line charts for displaying relative uncertainty (expressed as CV) in N and GWP outputs, in conjunction with their SDs (in kg N ha^{-1} yr^{-1} for N outputs, and kg CO_2 ha^{-1} yr^{-1} for GWP outputs), at European level

The results based on Figure 14 and Table 10 show that the relative uncertainty in the European level outputs varies between 10 – 34 % and increases in the following direction: CH_4, NH_3 < N_2O < NO_x < N_{leSW}, N_{leGW}

4.2.1.2 Discussion

This Section provides a discussion on the uncertainty in the outputs in combination with the quantified uncertainty in the APs which propagated through the model. In addition, a brief explanation is given as to why outputs being affected by the same APs have different uncertainties. Note, however, that a detailed explanation on this kind of differences would require in-depth expertise of INTEGRATOR modellers and/or soil scientists. Additionally, in this Section, estimated values of means and CVs are compared with estimates from other related researches to check the statistical models (pdfs) used to describe the uncertainty in the APs for plausibility.

Results based on box-plots

When examining the N outputs based on the box-plots in Figure 13 , it can be seen that N_{leSW} and N_{leGW} have higher uncertainties than NH_3 and NO_x. Based on Table 2, there are 37 APs affecting N_{leSW} and N_{leGW} whereas there are 31 and 30 APs affecting NH_3 and NO_x respectively. So a simplified explanation for the higher amount of uncertainty, propagated to N_{leSW} and N_{leGW} outputs, could be the larger number of uncertain APs that affected them. Nevertheless this may not be always the case. A better explanation that can be given is that among the 37 APs that affect N_{leSW} and N_{leGW} there are the AP_{31} and AP_{32} which have both high uncertainty (CV=0.50) and the AP_{49} which has also high uncertainty (SD=0.75) whereas the uncertainties in most of the APs that affect NH_3 and NO_x are much lower (CV=0.25).

Another observation that can be made is that N_{leSW} has higher uncertainty than N_{leGW}. The question thus raised here is, how this could be explained since they are both affected by the same APs. According to De Vries et al. (2011c), the leaching from the soil system is distributed between groundwater recharge and subsurface runoff to surface water, and the functions used in INTEGRATOR for the calculation of N_{leSW} and N_{leGW} are the following:

$$N_{leSW} = f_{ro} . N_{le} + N_{sr} + N_{le,ms}$$

$$N_{leGW} = \left(1 - f_{ro} \right) . N_{le}$$

Where:

f_{ro} = The sub-surface runoff fraction (AP_{53})

N_{le} = amount of N for leaching (calculated based on APs related to animals, N input, N uptake, soil emission)

N_{sr} = amount of N for sub-surface runoff (calculated based on AP_{52} and other APs related to N input)

$N_{le,ms}$ = amount of N for leaching from manure storage (calculated based on AP_{51} and other APs related to animal numbers and N excretion: AP_1, ..., AP_8)

Therefore, there are many more uncertain APs involved in the release of the N_{leSW} than the ones involved in the release of the N_{leGW} and this may be the reason why the N_{leSW} has higher uncertainty than N_{leGW}.

Emissions of NH_3 have higher uncertainty than NO_x and as it has been mentioned before there are 31 APs affecting NH_3 and 30 APs affecting NO_x. Therefore almost the same number of APs that affected NH_3 affected NO_x as well. This reveals that by simply counting the number of APs that affect an output, in order to draw conclusions about its propagated uncertainty, is a very simplified manner and may lead to misinterpretations. Since the uncertainties in the APs that affect NH_3 and NO_x have more or less the same level, the explanation for the higher uncertainty in NH_3 could be that NH_3 is more sensitive to variations in the APs that are affecting it. Good knowledge about emissions processes and the related functions incorporated in INTEGRATOR would definitely provide a much better explanation about the higher

level of uncertainty in NH_3 emissions. The acquisition of this knowledge though could not be achieved within the time frame of this research. Therefore the reader of this report should combine the results, discussed in this Chapter, with other specific sources in order to get a sound interpretation.

As far as the GWP outputs are concerned, N_2O has higher uncertainty than CH_4 and this can be explained by the high uncertainties in some of the APs that affect the N_2O: AP_{22} and AP_{23} (SD=0.50), AP_{31}, AP_{32} and AP_{48} (CV=0.50), and AP_{49} (SD=0.75) compared with the lower uncertainties in the AP_1-AP_4 (CV=0.1) and AP_{54}-AP_{56} (CV=0.25) that affect CH_4.

Another observation based on the box-plots is that N_{leSW}, N_{leGW} and N_2O have one-sided outliers. This can be explained by the fact that these outputs have been affected by the high SD (i.e.0.75) of AP_{49} which was lognormally distributed (one-tailed skewed) and as such influenced their frequency distributions.

Results based on column-charts

When examining the outputs based on the column-line charts shown in Figure 14, it can be seen that outputs with high uncertainty (i.e. high SD) and a high mean value too have low relative uncertainty. This explains why the N_{leSW}, which has higher absolute uncertainty (SD) and higher mean value than N_{leGW} in Figure 13, has lower relative uncertainty (CV) than N_{leGW}.

Plausibility of statistical models (pdfs)

The plausibility of the statistical models (pdfs), used to represent the uncertainty in the APs, was checked by comparing the computed mean and CV values against estimates by Schulze et al. (2009) and by Velthof et al. (2009). The first group provides mean and CV values for GWP outputs, referring to the time period between 2000 and 2005, calculated by a synergy of various models (i.e. EFISCEN, ORCHIDEE, LPJ, BIOME-BGC, YASSO and PASIM) For these values it can be assumed that any changes in them, within these five years, are rather small and thus these values can still be compared with results from this research. The second group provides mean values for N outputs, referring to 2000, obtained by MITERRA-EUROPE which is a deterministic and static model (stand alone policy tool). The values from the three different sources are given in Table 10.

Table 10. Comparison between means (in kg N ha^{-1} yr^{-1} for N outputs, and kg CO_2 ha^{-1} yr^{-1} for GWP outputs) and CVs for N and GWP outputs at European level, with corresponding means and CVs obtained from other sources

European level	This research Year 2000		Schulze et al. Year 2000		Velthof et al. Years between 2000 -2005	
N outputs	**mu**	**CV**	**mu**	**CV**	**mu**	**CV**
NH_3	16	0.13			17	
NO_x	4.2	0.24			2.0	
N_{leSW}	18	0.27			16	
N_{leGW}	7.5	0.34				
GWP outputs						
CH_4	1145	0.10	1134	0.51		
N_2O	2524	0.17	1557	0.50		

N.B. Velthof et al. provide one value for N leaching and it refers to the sum of $N_{leSW} + N_{leGW}$

Table 10 shows that some results of this research are similar to those by the other two researches. The higher N_2O emissions from this research may be explained by the use of improved N_2O emission factors (Lesschen et al. 2011). The lower uncertainty in both N_2O and CH_4 may be explained by the fact that during this research only the input uncertainty was taken into account, whereas the model uncertainty was not considered. The difference in the sum $N_{leSW} + N_{leGW}$ can be explained by the fact that Velthof et al. (2009) estimated N leaching by running MITERRA-EUROPE model just once and using average values of APs whereas the values for this research were estimated by averaging over 1000 output simulations of APs with their uncertainties being incorporated in the model. In particular, the skewness of AP_{49} influenced the mean values of N_{leSW} and N_{leGW}. After all, it can be said that the pdfs used for the uncertainty in the APs were rather plausible.

4.2.2 Effect of Country level outputs for the Reference scenario

4.2.2.1 Results

The effect of Country level on the output uncertainty was visualized by:

- Box-plots, provided in Figure 15, describing the output uncertainty for five EU countries that were selected based on different characteristics (i.e. land cover, climate, and pressure of animals):
  - Netherlands: Atlantic climate, combination of grasslands, farmlands and high number of animals;
  - Greece: Mediterranean climate, arable land, low number of animals;
  - Ireland: Atlantic climate, mixed land cover (grazing-dairying-arable), high number of animals;
  - France: Combination of temperate (north) and Mediterranean climate (south), agricultural land, high number of animals;
  - Poland: Continental climate, intensively agricultural land, high number of animals.

- Column-line charts provided in Figure 16 and Figure 17, for all countries of EU-25;
- Choropleth maps, provided in Figure 18, Figure 19 and Figure 20, portraying the variation of the means and the CDs of the outputs over all countries in EU-25, according to five classes.

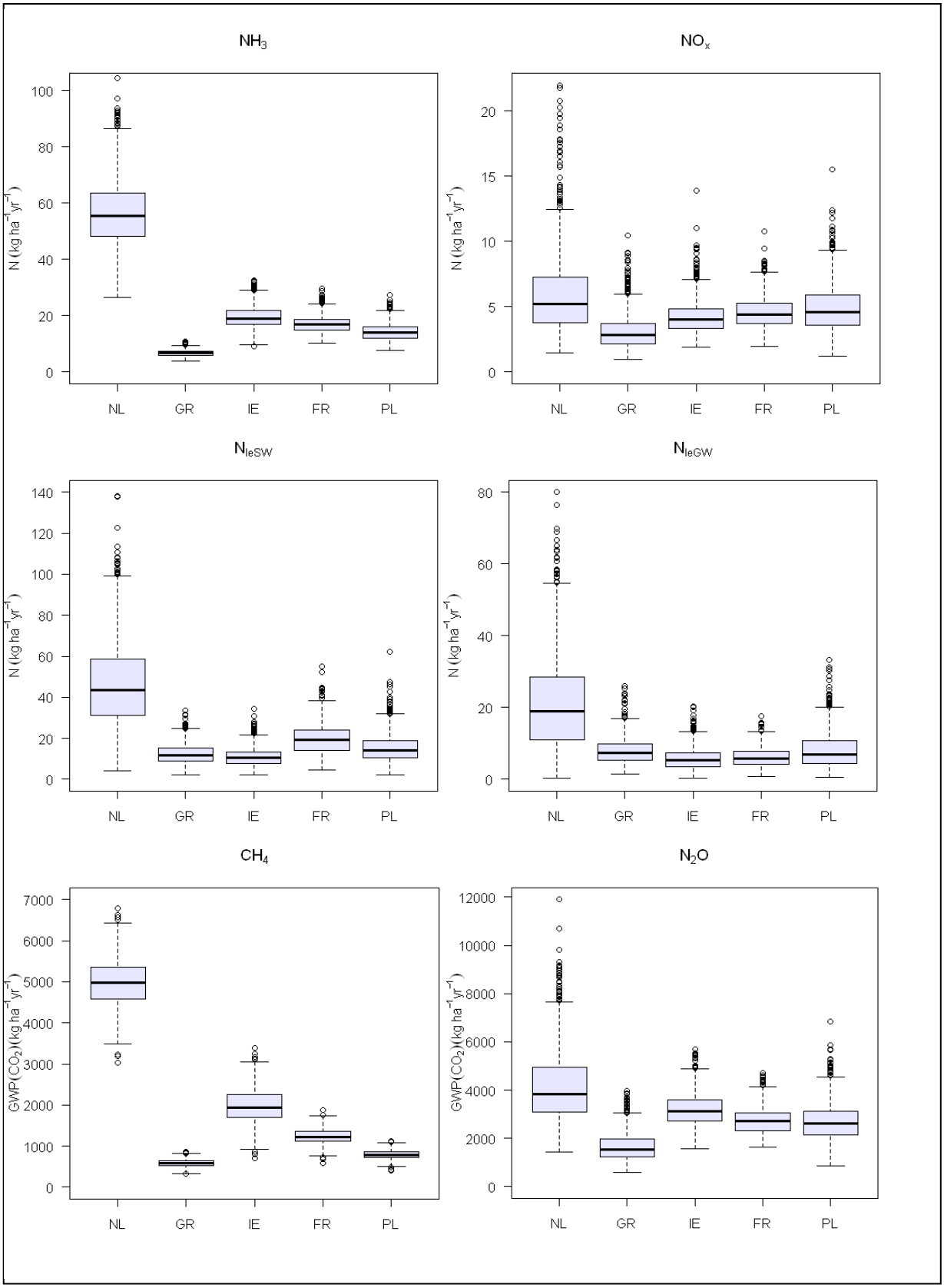


Figure 15. Box-plots for displaying uncertainty in N and GWP outputs for each of the five EU countries

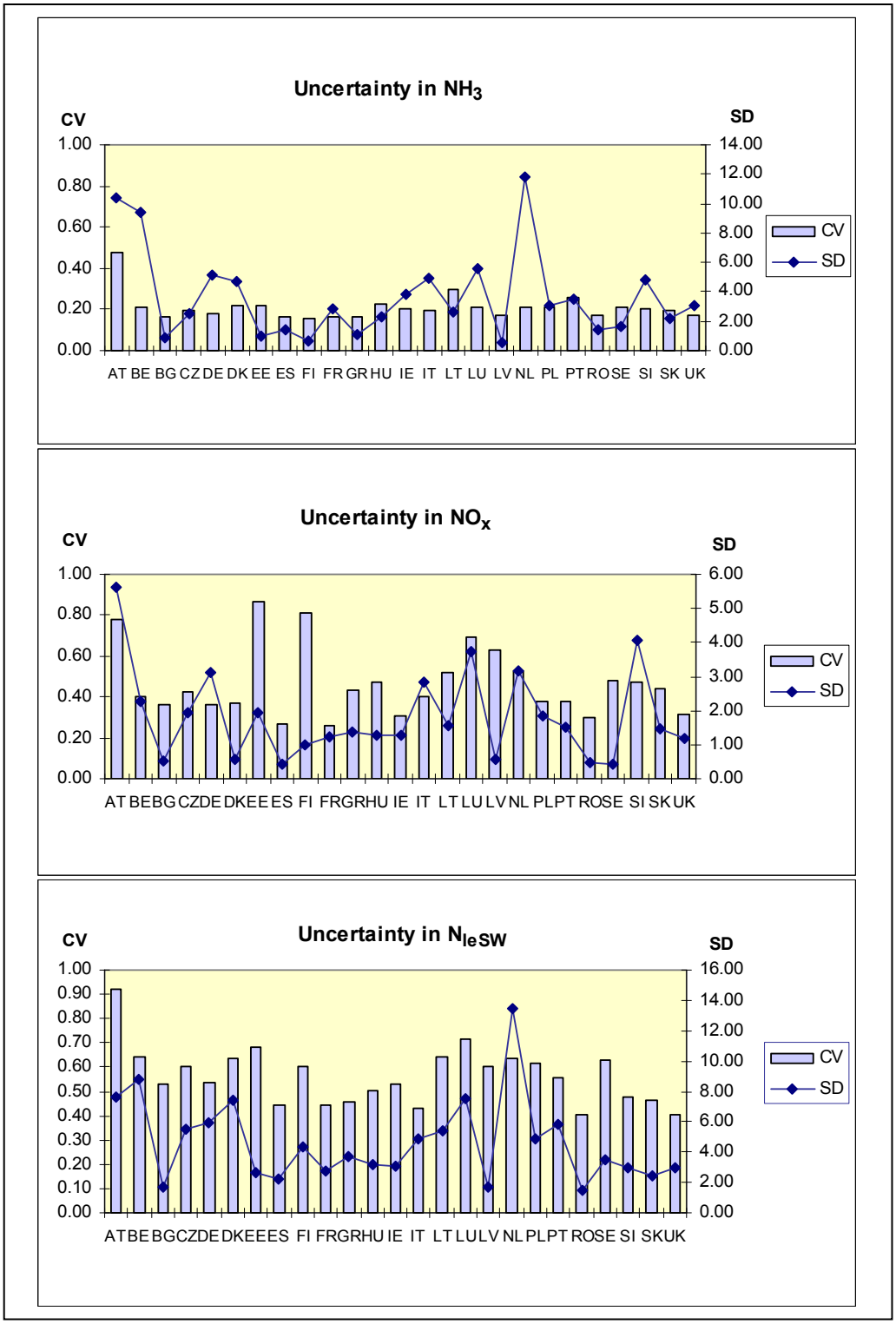


Figure 16. Column-line charts for displaying relative uncertainty in NH_3, NO_x and N_{leSW}, in conjunction with their SDs (in kg N ha^{-1} yr^{-1} for N outputs, and kg CO_2 ha^{-1} yr^{-1} for GWP outputs), for all countries of EU-25

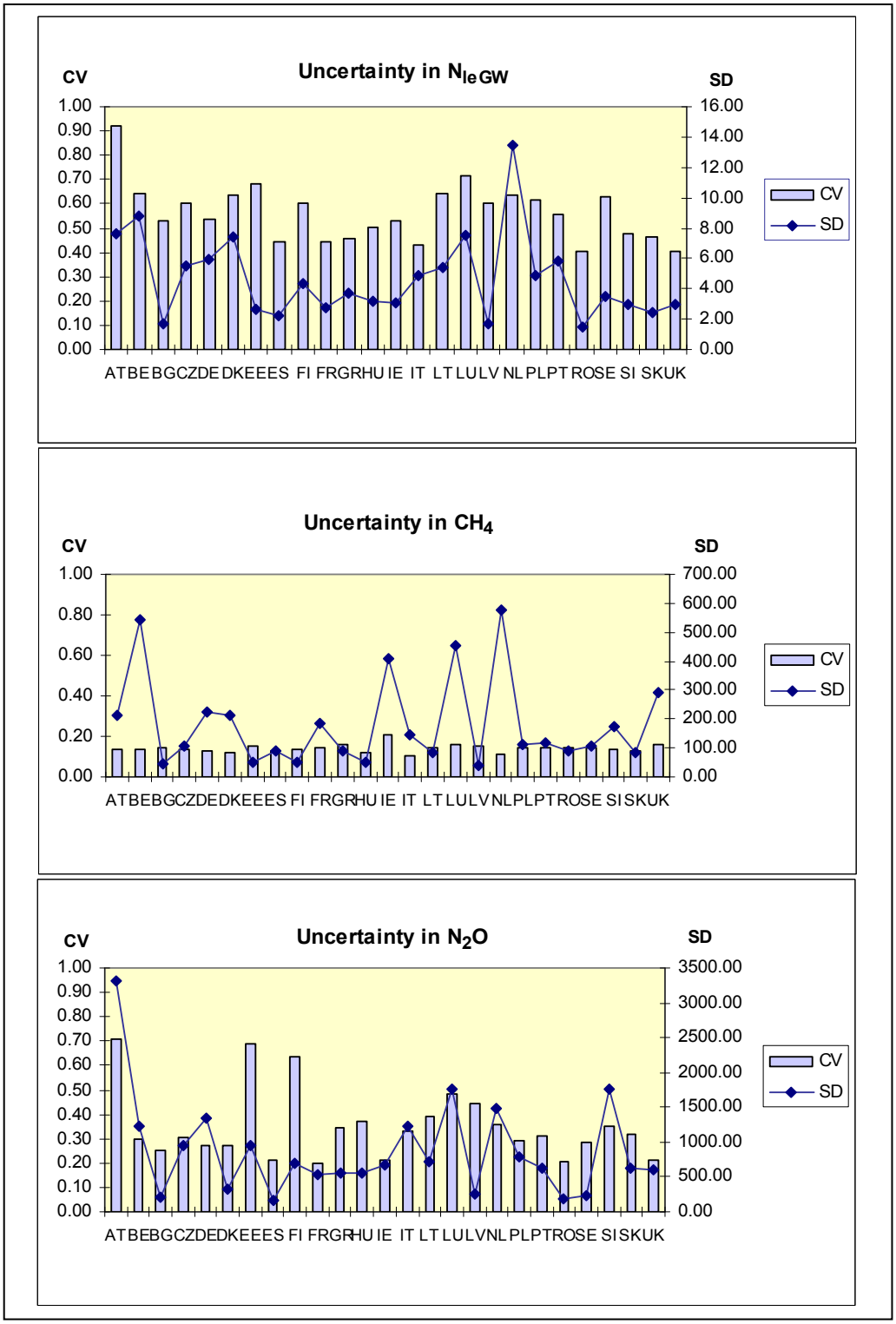


Figure 17. Column-line charts for displaying relative uncertainty in N_{leGW}, CH_4 and N_2O, in conjunction with their SDs (in kg N ha^{-1} yr^{-1} for N outputs, and kg CO_2 ha^{-1} yr^{-1} for GWP outputs), for all countries of EU-25

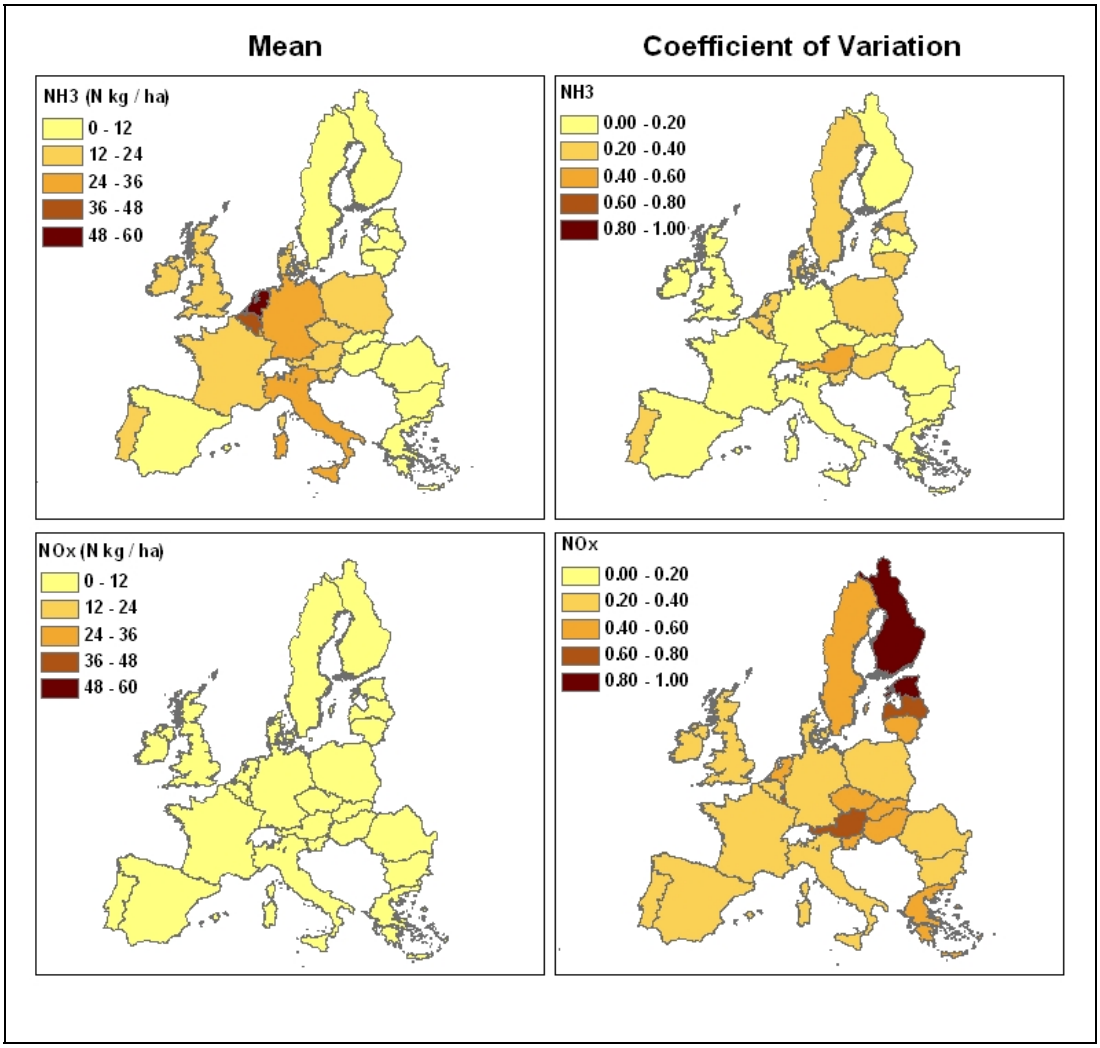


Figure 18. Maps displaying classes of means and relative uncertainty in NH_3 and NO_x over all countries of EU-25

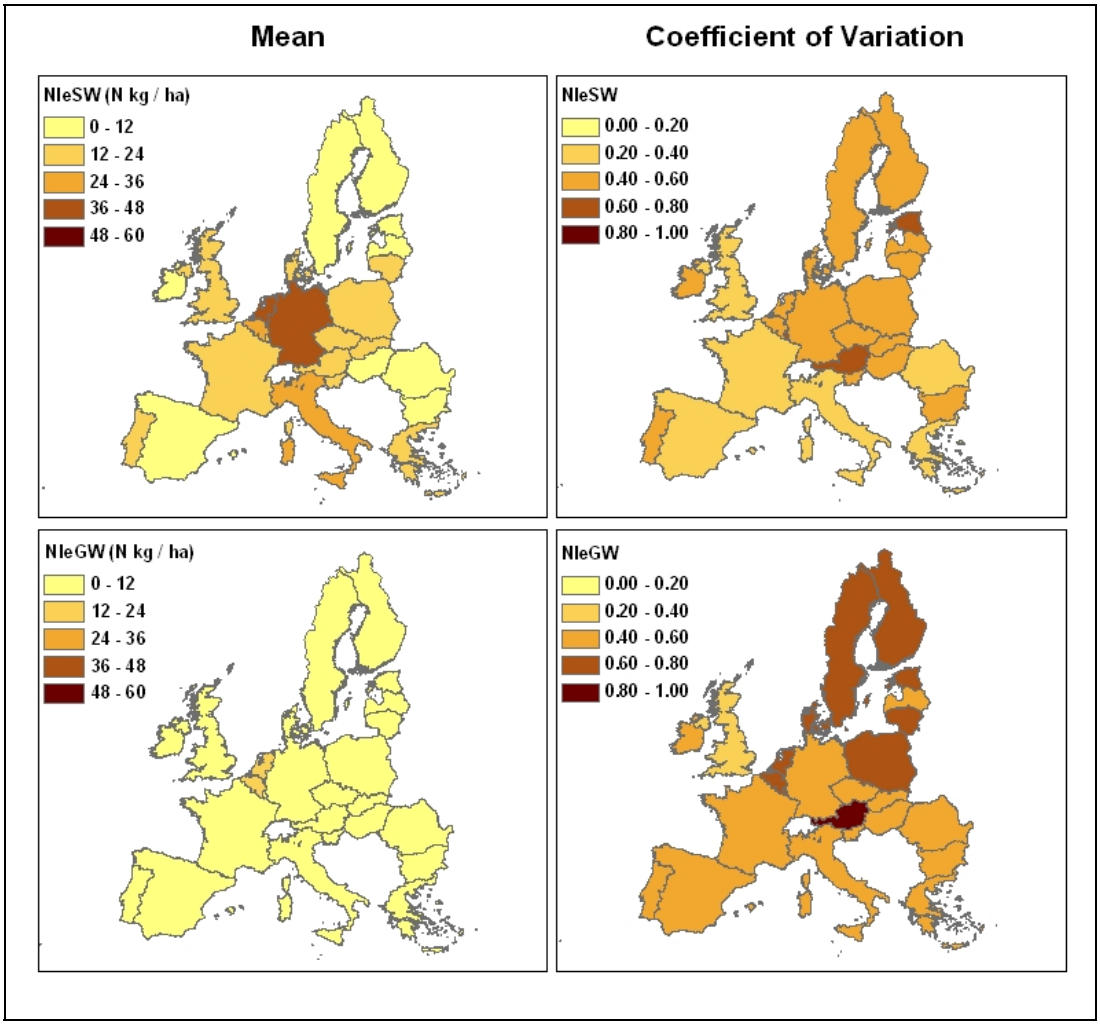


Figure 19. Maps displaying classes of means and relative uncertainty in N_{leSW} and N_{leGW} over all countries of EU-25

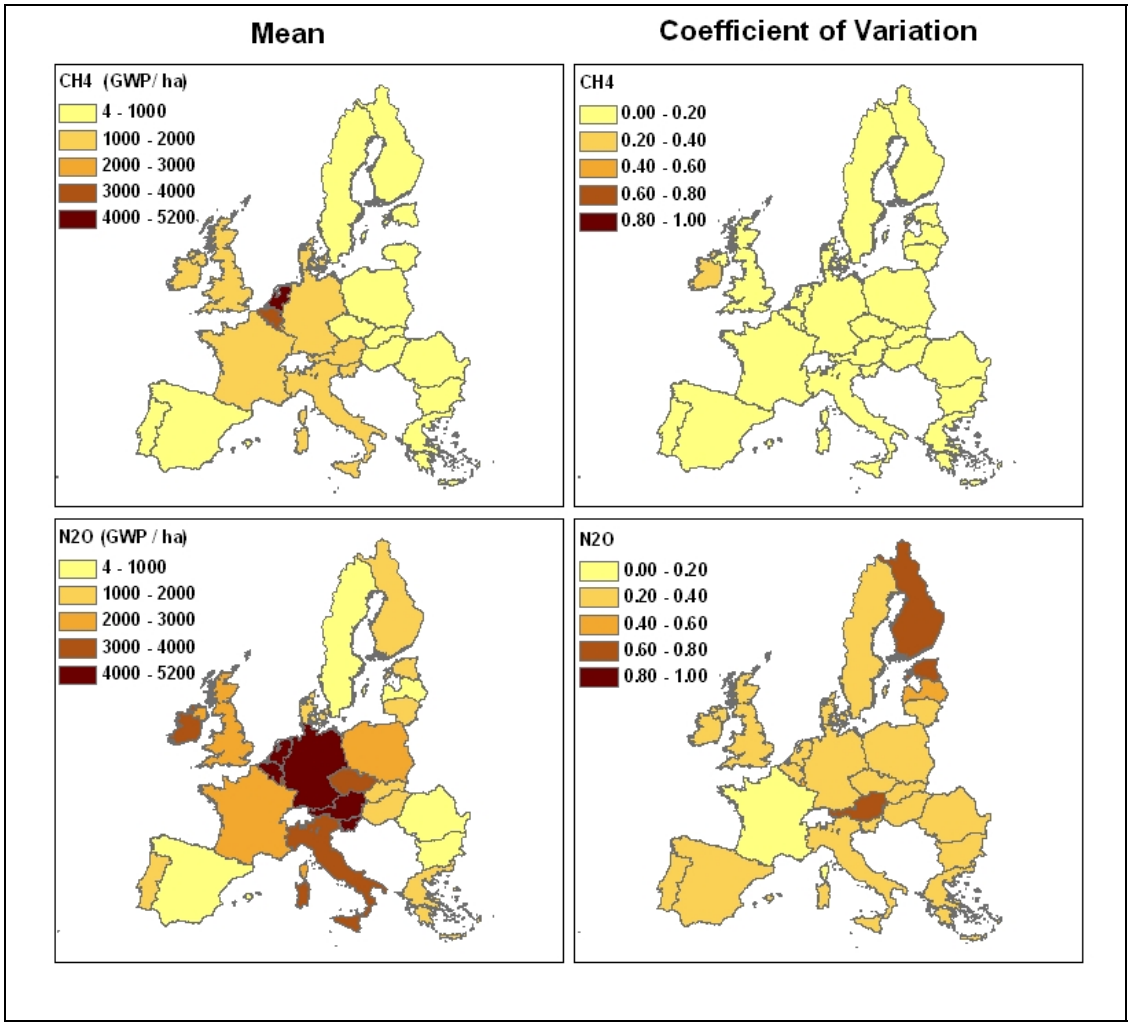


Figure 20. Maps displaying classes of means and relative uncertainty in CH_4 and N_2O over all countries of EU-25

The results based on Figure 16, Figure 17 and Table 12 show that the relative uncertainty at Country level outputs is higher than at European level as it varies between 11 – 92% and it increases in the following direction: CH_4, NH_3 < N_2O, N_{leSW} < NO_x, N_{leGW}.

4.2.2.2 Discussion

Results based on box-plots

When examining the outputs based on box-plots shown in Figure 15, it can be seen that the Netherlands has the highest uncertainty in all outputs. Other countries have more or less comparable levels of uncertainty. According to Oenema et al. (2001), the uncertainty in estimates of N_2O and CH_4 emissions, in agriculture, is rather large and mainly at local and country levels. This is due to the different nature and character of the agricultural sources, among regions and countries, as well as by the limited number and uneven spread of the measurements concerning related emission parameters. The Netherlands is characterised by its extensive areas of agricultural land with crops that demand high use of manure and fertilizers (e.g. potatoes, sugar beets, maize etc.), and also by its high number of cattle. Consequently, all APs related to animals, manure and fertilizer application were involved in the calculation of N and

GWP outputs for the Netherlands and as a result, the amount of uncertainty that propagated through the model has the highest level.

Greece has the lowest uncertainty level in four outputs. This may be explained by the fact that its crops are mainly olives, grapes, fruit, cotton, tomatoes (i.e. little or no use of manure and fertilizers) and that the country has a limited number of animals. Thus, the uncertain APs that affected the N and GWP outputs had much less influence on them.

The higher uncertainty in NH_3, N_2O and CH_4 for Ireland, compared with Greece, France and Poland, may be explained by the large number of animals and also by the crops of barley, sugar beets, wheat and fresh vegetables that are typical of the country. Therefore an increased number of APs, compared with the other three countries may have affected these output emissions.

France and Poland have almost the same level of uncertainty and this may be explained by the similar agricultural characteristics these two countries have.

Results based on column–line charts

When examining the outputs based on column–line charts, shown in Figure 16 and Figure 17, it can be seen that the outputs with the higher relative uncertainty are the NO_x, N_{leSW}, and N_{leGW}. This is due to their low mean values (see Table 12 with mean and CV values for all countries) relatively compared with NH_3. It can also be seen that CH_4 has the lowest relative uncertainty since, as explained previously, its uncertainty has been affected only by seven APs.

Results based on choropleth maps

When examining the outputs based on the choropleth maps shown in Figure 18, Figure 19 and Figure 20 , it can be seen that NO_x, N_{leGW}, and N_2O have high relative uncertainties in Finland, Estonia, Austria and Poland. This is due to the relatively low mean output values for these countries.

Plausibility of statistical models (pdfs)

The plausibility of the pdfs of this research was checked once again by comparing the mean values of the Country level outputs with estimated mean values by Velthof et al. (2009) produced by the MITERRA-EUROPE model. The values from the two sources are provided in Table 11.

Table 11. Comparison between mean values of outputs, for five EU countries, with corresponding mean values obtained from a different source

			Mean values	
		5 EU countries	This research Year 2000	Velthof et al. Year 2000
N outputs (kg N ha^{-1} yr^{-1})	NH_3	NL	57	57
		GR	6.8	8.0
		IE	19	22
		FR	17	19
		PL	14	15
	NO_x	NL	6.0	7.0
		GR	3.1	2.0
		IE	4.2	7.0
		FR	4.6	3.0
		PL	4.9	1.0
	N_{leSW} + N_{leGW}	NL	68	72
		GR	20	7.0
		IE	17	20
		FR	26	17
		PL	23	15
GWP outputs (kg CO_2 ha^{-1} yr^{-1})	CH_4	NL		
		GR		
		IE		
		FR		
		PL		
	N_2O	NL	4157	3746
		GR	1641	937
		IE	3204	1405
		FR	2740	937
		PL	2714	937

N.B. Velthof et.al did not include parameters that affected CH_4 emissions and so there were no values available for them.

The values of NH_3 by Velthof et al. (2009) are similar to the ones obtained in this research. The differences in values of NO_x, N leaching, and N_2O between the two sources stem from the fact that Velthof et al. (2009) (as explained already in the Section about European level outputs) ran MITTERA-EUROPE only once with averages values of APs, and not with their uncertainties incorporated and averaged over 1000 input simulations.

The mean and CV values of N and GWP outputs, for all countries of EU-25 are provided in Table 12.

Table 12. Values of means (in kg N ha^{-1} yr^{-1} for N outputs, and kg CO_2 ha^{-1} yr^{-1} for GWP outputs) and CVs of N and GWP outputs (min CV=0.11, max CV=0.92), for all countries in EU-25

	NH_3		NO_x		N_{leSW}		N_{leGW}		CH_4		N_2O	
EU- 25 COUNTRY	**mu**	**CV**	**mu**	**CV**	**mu**	**CV**	**mu**	**CV**	**mu**	**CV**	**mu**	**CV**
Austria (AT)	21	0.48	7.2	0.78	24	0.71	8.3	0.92	1505	0.14	4690	0.71
Belgium (BE)	44	0.21	5.6	0.40	34	0.50	14	0.64	3996	0.14	4144	0.30
Bulgaria (BG)	5.3	0.17	1.5	0.36	4.2	0.54	3.2	0.53	332	0.14	868	0.25
Czech Republic (CZ)	13	0.20	4.6	0.42	16	0.44	9.1	0.60	805	0.13	3102	0.31
Germany (DE)	28	0.18	8.7	0.36	40	0.41	11	0.53	1730	0.13	4953	0.27
Denmark (DK)	21	0.22	1.5	0.37	20	0.46	12	0.63	1853	0.12	1221	0.27
Estonia (EE)	4.4	0.21	2.2	0.87	4.4	0.63	3.8	0.68	321	0.15	1387	0.69
Spain (SP)	8.9	0.17	1.6	0.27	10	0.36	5.0	0.45	711	0.13	773	0.21
Finland (FI)	4.2	0.16	1.2	0.81	8.9	0.53	7.2	0.61	360	0.13	1100	0.63
France (FR)	17	0.16	4.6	0.26	20	0.38	6.1	0.45	1241	0.15	2740	0.20
Greece (GR)	6.8	0.16	3.1	0.43	12	0.40	8.0	0.46	584	0.16	1641	0.35
Hungary (HU)	10	0.23	2.7	0.47	9.2	0.50	6.3	0.50	436	0.12	1519	0.37
Ireland (IE)	19	0.20	4.2	0.31	11	0.42	5.7	0.53	1976	0.21	3204	0.21
Italy (IT)	25	0.20	7	0.40	27	0.32	11	0.43	1332	0.11	3731	0.33
Lithuania (LT)	8.9	0.29	3	0.52	12	0.48	8.5	0.64	587	0.14	1825	0.39
Luxembourg (LU)	27	0.21	5.4	0.69	22	0.70	11	0.72	2877	0.16	3662	0.48
Latvia (LV)	3.4	0.17	0.89	0.63	3.8	0.47	2.9	0.60	257	0.15	581	0.45
Netherlands (NL)	57	0.21	6	0.53	46	0.46	21	0.64	4987	0.12	4157	0.36
Poland (PL)	14	0.21	4.9	0.38	15	0.45	8.0	0.61	788	0.15	2714	0.29
Portugal (PT)	14	0.26	4.0	0.38	15	0.44	11	0.55	824	0.14	1996	0.31
Romania (RO)	8.5	0.17	1.6	0.30	7.2	0.37	3.7	0.41	603	0.14	956	0.20
Sweden (SE)	7.7	0.21	0.89	0.48	11	0.57	5.5	0.63	750	0.14	790	0.29
Slovenia (SI)	23	0.21	8.6	0.47	13	0.49	6.2	0.48	1265	0.14	4970	0.35
Slovakia (SK)	11	0.19	3.3	0.44	13	0.42	5.4	0.46	660	0.13	1951	0.32
United Kingdom (UK)	18	0.17	3.7	0.32	14	0.33	7.5	0.40	1779	0.16	2778	0.21

4.2.3 Effect of the three robustness scenarios at European level

4.2.3.1 Results

The effect of the three robustness scenarios on the output uncertainty, at European level, was visualized by:

- Box-plots, provided in Figure 21, describing the output uncertainty according to each scenario;
- Column charts, provided in Figure 22, describing the relative output uncertainty according to each scenario.

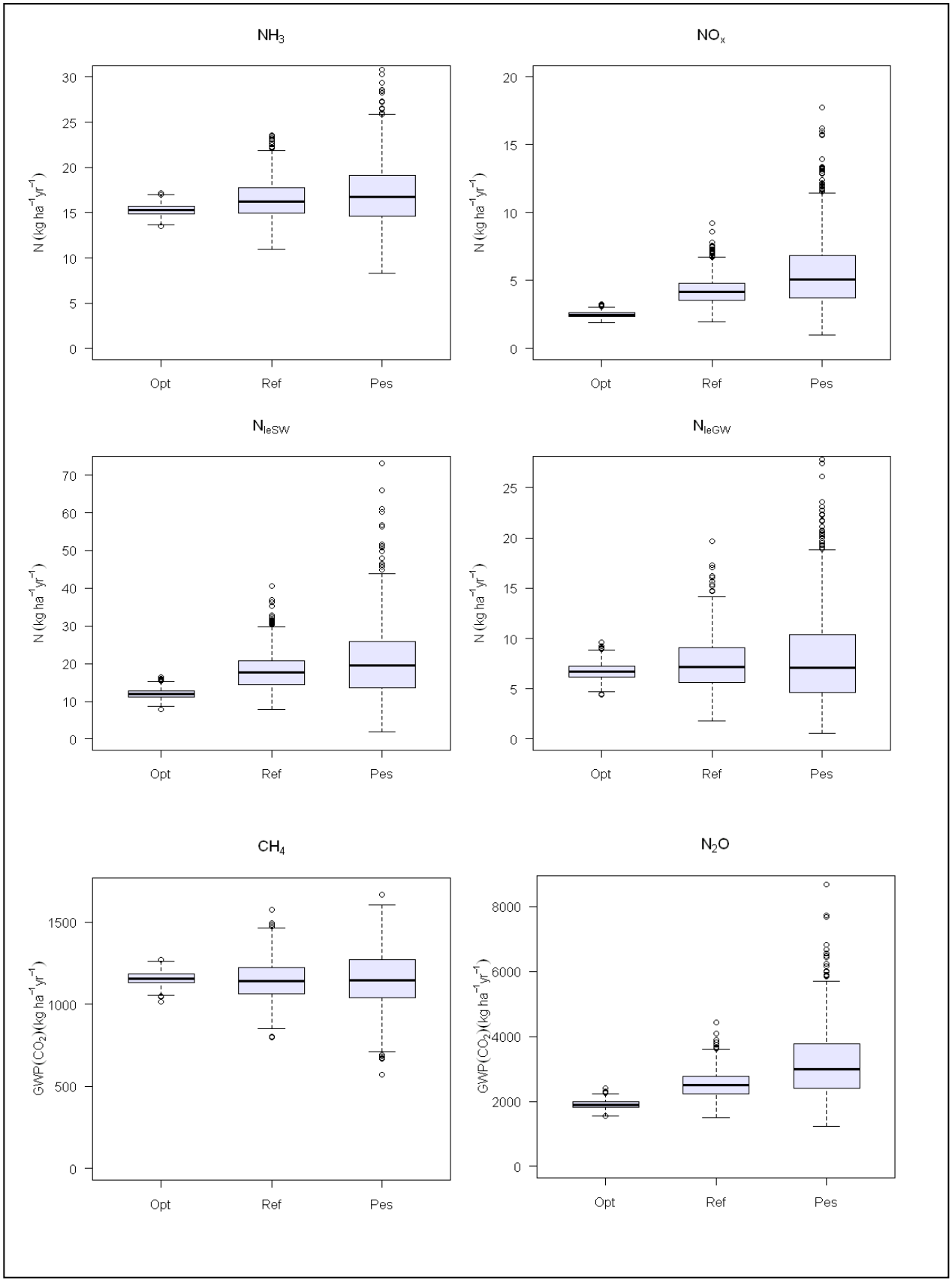


Figure 21. Box-plots displaying uncertainty in N and GWP outputs, for each robustness scenario, at European level

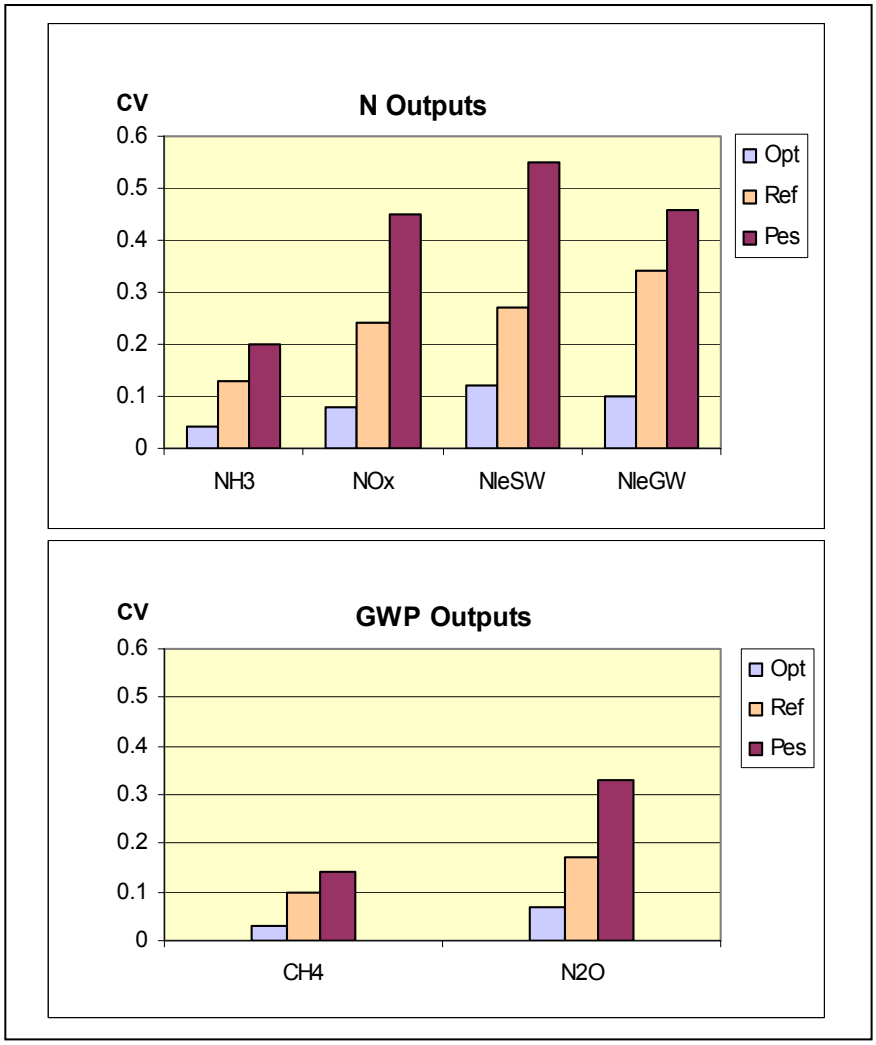


Figure 22. Column charts displaying relative uncertainty in N and GWP outputs, for the three robustness scenarios, at European level

The results based on Figure 22 and Table 13 show that the relative uncertainty in all model outputs, with all three scenarios considered, varies from 12% (max Opt) to 55% (max Pes) and it increases in the same direction as for the Reference scenario.

4.2.3.2 Discussion

Results based on box-plots

When examining the three scenarios based on box-plots shown in Figure 21, it can be seen that the uncertainty in each output increases greatly from the Optimistic to the Pessimistic scenario for all outputs. For NH_3, N_{leGW}, and CH_4 the mean values are kept almost at the same levels whereas for NO_x, N_{leSW}, and N_2O the mean values increase slightly going from Optimistic to Pessimistic scenario. This can be explained by the skewness of the lognormally APs that affected these fluxes and as a result their mean values were influenced as well.

Results based on column charts

When examining the three scenarios based on column charts shown in Figure 22 , the increased relative uncertainty in the model outputs, going from the Optimistic to the Pessimistic scenario, looks even profounder compared with the overview given

by the box-plots. This is because going from the Optimistic to the Pessimistic scenario the relative uncertainty is 4-6 times higher. When going from the Reference scenario, for which the uncertainty analysis was carried out in detail, to the Pessimistic scenario the relative uncertainty is 1.3-2 times higher and inversely when going from the Reference to the Optimistic scenario the relative uncertainty is 2-3.3 times lower. Table 13 is providing an additional overview with the CV values of all outputs for the three scenarios along with their CV_{rob} (robustness) which is the CV calculated from the CV values of Opt, Ref and Pes scenarios for each model output respectively (e.g. 0.53 is the CV of: 0.04, 0.13, 0.20)

Table 13. Values of CVs for N and GWP outputs according to each scenario, and their CV_{rob}, at European level

European level	CV			CV_{rob}
N outputs	**Opt**	**Ref**	**Pes**	**All 3 scenarios**
NH_3	0.04	0.13	0.20	0.53
NO_x	0.08	0.24	0.45	0.59
N_{leSW}	0.12	0.27	0.55	0.57
N_{leGW}	0.10	0.34	0.46	050
GWP outputs				
CH_4	0.03	0.10	0.14	0.51
N_2O	0.07	0.17	0.33	0.56

Based on Table 13, the CV_{rob} (i.e. "the uncertainty in the uncertainty") has a magnitude of approximately 50% and the relative uncertainty in the model outputs, according to the three scenarios, varies:

- between 3 - 12 % for the Optimistic;
- between 10 - 34 % for the Reference;
- between 14 - 55 % for the Pessimistic.

Consequently, the relative uncertainty obtained by this research, which is most likely in the range of 12-55%, can be considered low since for an environmental assessment at EU scale a real maximum of 55% is quite reliable. And this can be explained by the following example: if an output has relative uncertainty 10% and "the uncertainty in the uncertainty" is 50%, then the relative uncertainty for this output still remains in a range of 5-15% only.

4.3 Uncertainty contribution of each group of APs to the total output uncertainty for the Reference scenario

4.3.1 Results

In this Section, the results from the summation of the model outputs according to sub-step "B" (see Section 3.5) are presented.

The uncertainty contribution of NCU, NUTS, CNTRY and GEN group to the overall output uncertainty was visualized by:

- Stack-columns, provided in Figure 23, describing the uncertainty contribution of each individual group to Country level outputs, concerning the five countries of EU-25 that were selected during sub-step "A";

- Stack-columns, provided in Figure 24, describing the uncertainty contribution of each individual group to European level outputs.

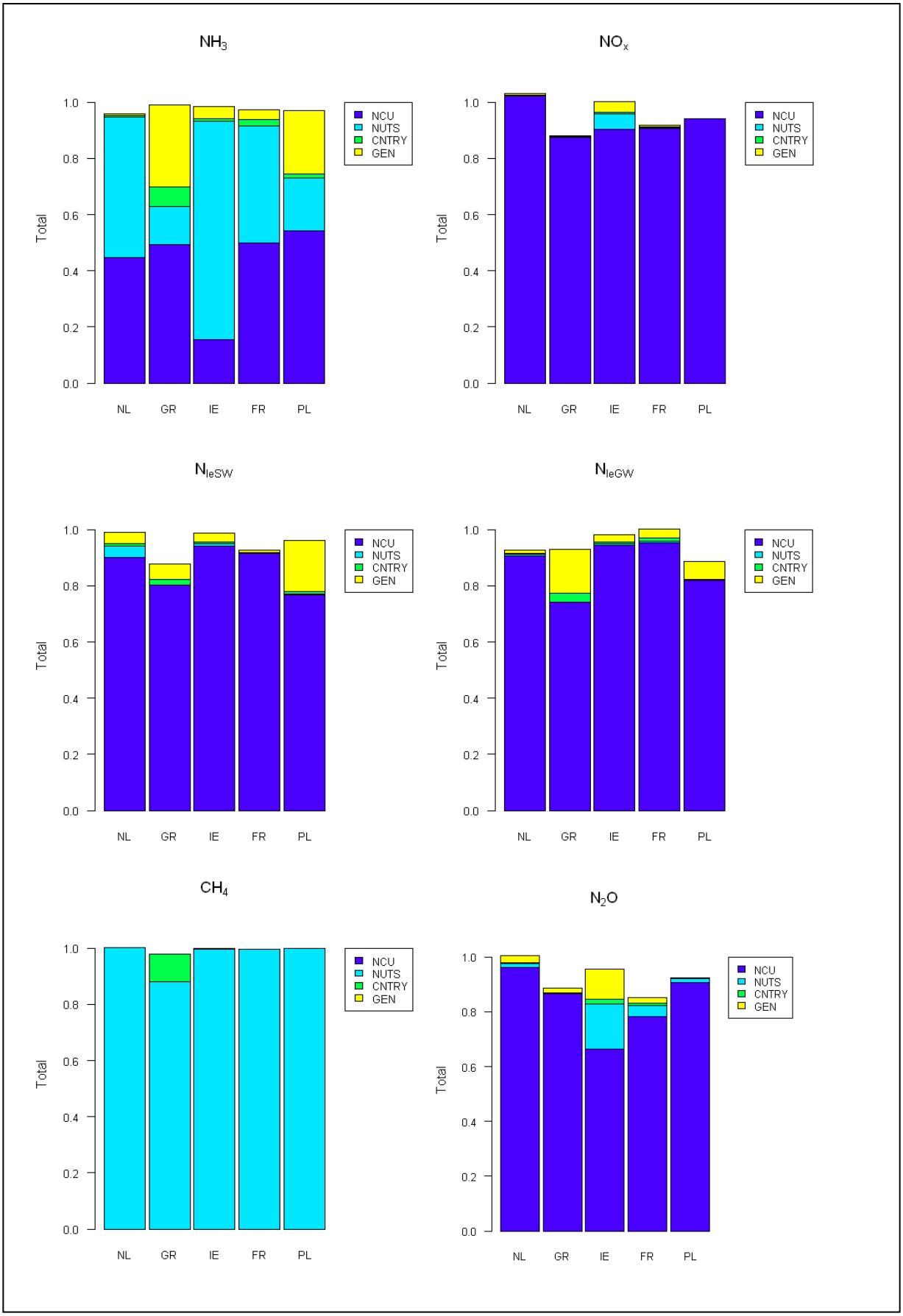


Figure 23. Stack-columns displaying uncertainty contribution of each group of APs to N and GWP outputs, for each of the five EU countries

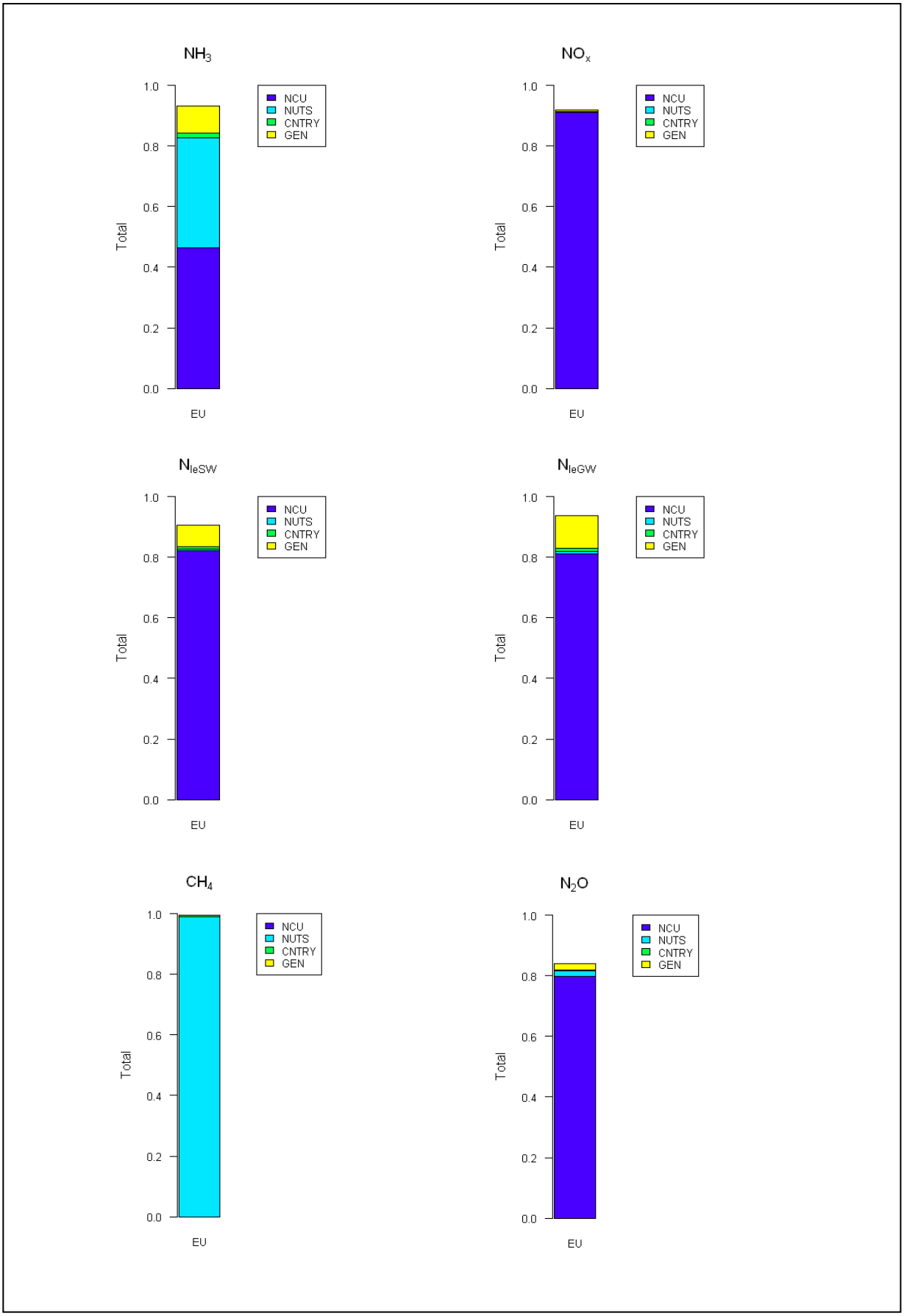


Figure 24. Stack-columns displaying uncertainty contribution of each group of APs to N and GWP outputs, at European level

The results, based on Figure 23 and Figure 24, indicate that the APs of the NCU group have the most significant contribution to the total output uncertainty.

4.3.2 Discussion

When examining the stack-columns for the five countries shown in Figure 23, it can be seen that the group contributing mostly to the uncertainty in NO_x, N_{leSW}, Nl_{eGW} and N_2O fluxes is the NCU group. The GEN group has a small contribution to N_{leSW} and Nl_{eGW} (for Greece and Poland) and this seems rather logical since AP_{10} and AP_{11}, which affect N_{leSW} and Nl_{eGW}, belong to the GEN group. For NH_3 emissions, the NUTS group, and at a lesser extent the GEN group (and slightly the CNTRY group) are contributing as well. The larger contribution of GEN group to NH_3 emissions for Greece and Poland, compared with the other countries, can be explained by the fact that GEN APs affect mainly the amount of NH_3 emissions in countries where chemical fertilizers are used. Mediterranean countries like Greece are using urea as fertilizer which has a relatively high emission factor. This is the reason why in countries where most of NH_3 emissions are released from animal excretion and application of manure, such as the Netherlands, there is hardly any contribution of the GEN group to the uncertainty in the NH_3 emissions.

For CH_4, the NUTS group is the only one that contributes to the output uncertainty (except for the slight percentage of CNTRY group that also contributes to the outputs produced for Greece). The latter could be explained by the fact that AP_{54}, AP_{55}, and AP_{56} belong to the NUTS group and obviously had a heavier influence on CH4 emissions than the CNTRY APs.

When examining the stack-columns for the whole EU-25, shown in Figure 24, it can be seen that the uncertainty percentages from each group have almost the same patterns as the ones discussed above.

By summarizing the results, which showed that the NCU group is the most important source to the total output uncertainty, the conclusion that can be drawn is that by gathering more detailed data on NCU parameters as well as more knowledge about the processes these parameters are associated with the output uncertainty of INTEGRATOR model can be reduced. Furthermore, another conclusion can be also made based on the fact that the verification of the model takes place by checking its estimates against independent measurements and against results from inverse modelling (De Vries et al. 2011b). And this conclusion is that the gathering of field observations that can be used for a calibration may also lead to lower uncertainty ranges in the APs.

5. CONCLUSIONS AND RECOMMENDATIONS

5.1 Main conclusions

After having completed all steps of the uncertainty propagation methodology to the INTEGRATOR model, the general conclusion that can be made is that the objective of this research has been met. The uncertainty in the APs was characterized by defining the parameters of their pdfs and exploring the impact of their uncertainty on N fluxes and CH_4 emissions using Optimistic (Opt), Reference (Ref) and Pessimistic (Pes) scenarios for the assumed input uncertainties (Robustness analysis). The model was run by applying the Monte Carlo simulation. The uncertainty in the model outputs was quantified by computing the statistics of the output´s pdfs. The uncertainty contribution of the APs to overall output uncertainty was also analysed. The following paragraphs provide the main conclusions drawn from the results of this research and their comparison with other independent sources.

By using the Reference values for the assigned uncertainties in the APs, the results at European level showed that the propagated relative uncertainty in model outputs ranges from 10 to 34% and increases in the direction: $CH_{4,em}$, $NH_{3,em}$ < N_2O_{em} < $NO_{x,em}$ < N_{leSW}, N_{leGW}. The results at County level showed that the propagated relative uncertainty in model outputs ranges from 11 to 92% and increases in the direction: $CH_{4,em}$, $NH_{3,em}$ < N_2O_{em}, N_{leSW} < $NO_{x,em}$, N_{leGW}. The results by examining five representatives EU countries showed that differences in uncertainties (expressed as SDs) among countries are caused by APs related to numbers of animals and application of manure and fertilizers. The Netherlands has the highest uncertainty in all outputs.

When comparing the three robustness scenarios, the results showed that the relative uncertainty in model outputs for the Pessimistic scenario is 1.3-2 times higher than for the Reference scenario, it ranges from 12 to 55% and it increases in the same direction as for the Reference scenario. Inversely, the relative uncertainty for the Optimistic scenario is about 2-3.3 times lower than for the Reference scenario and it ranges from 3 to 12% only. In general, the results obtained from the Robustness analysis indicate that a quite robust statement can be made: the relative uncertainty is less than 55% for all considered model outputs and most likely in the range of 12-55%.

The uncertainty contribution of each group of APs to the total output uncertainty showed that the APs from NCU have the highest contribution and therefore additional data on them is required for the reduction of the overall output uncertainty.

The validation of the statistical models (pdfs) used for the uncertainty in the APs showed that statistics calculated during this research were similar to statistics from other related researches and thus the employment of these models can be judged efficient for the objective of this research.

5.2 Addressing the research questions

This Section provides a summary about how each research question (see also Section 1.4) was addressed:

RQ1: How can statistical models (joint pdfs) be built that fully characterize the uncertainty in the APs by taking into account their spatial- and cross-correlations?

The uncertainty in the APs was represented by means of pdfs. The parameters of these pdfs were derived partly from European datasets and partly from expert knowledge by assuming that the values of the APs were normally or lognormally distributed. For each AP the following parameters were defined: the CVs or SDs, the spatial- and cross-correlations coefficients between plots within the four spatial levels of INTEGRATOR (NCU, NUTS, Country and Europe), and the minimum and maximum value. Additionally an analysis on these statistics was performed by three robustness scenarios: Optimistic, Reference and Pessimistic. The APs, depending on the spatial level they were linked to, were divided in four groups: NCU, NUTS, CNTRY, and GEN.

RQ2: How can realizations of APs be sampled efficiently from their pdfs by using stochastic simulation techniques?

After the pdf of each AP was defined, the next step was the generation of a sample of 1000 simulated values for each AP, randomly drawn for their pdfs. This was done by using the "rmultnorm" function (Multivariate Normal Random Number Generator) incorporated in R software. Before this generator was used, the following statistical model for each simulated value x(i) was adopted: $x(i) = \mu(i) + \sigma(i) \cdot \xi(i)$

Where :

$\mu(i)$ was the mean value for each AP

$\sigma(i)$ was the SD for each AP

$\xi(i)$ a standard normal variable

The values of $\sigma(i)$ were estimated with RQ1, and the values of $\mu(i)$ were available in INTEGRATOR´s database, so the simulation of x(i) could be achieved only through the simulation of $\xi(i)$. Since x(i) and $\xi(i)$ had a linear relationship, $\xi(i)$ could be simulated by using the spatial-correlations coefficient matrices as specified for x(i). So an R-script was developed where the spatial-correlations coefficient matrices for each AP were built first and then the simulation of $\xi(i)$ took place by using the "rmultnorm" function. Finally the 1000 simulations, produced for each group of APs, were stored in files with names and format according to INTEGRATOR requirements.

RQ3: How can the Monte Carlo uncertainty propagation analysis be carried out in batch mode and its results be stored automatically?

The Monte Carlo uncertainty propagation analysis took place through the execution of an R-script where the INTEGRATOR dll was loaded. The running of the model for the 1000 input files from each group of APs, and the storing of the output files was implemented within a "for loop". The steps of the loop were the following:

1. Each input simulation from each group was provided;
2. The model was run;
3. One output simulation for NH_3, NO_x, N_{leSW}, N_{leGW}, CH_4, and N_2O respectively was produced at Country and European Level, and it was stored in a corresponding folder.

The above three steps were repeated 1000 times in batch mode.

RQ4: How can the results of the uncertainty propagation analysis be summarised and visualised and thus efficiently communicated to end-users across Europe?

The 1000 output simulations for the six output fluxes, produced at Country and European level, were summarized by checking:

- The effect the spatial aggregation to European level had on the output uncertainty;
- The effect the spatial aggregation to Country level had on the output uncertainty;
- The effect the three robustness scenarios had on the output uncertainty.

The summation of the results took place by visualizing the SD, CV and the mean values with box-plots, column-(line) charts, choropleth maps and tables.

RQ5: Which APs are the main uncertainty sources contributing to the total output uncertainty?

The uncertainty contribution of each group of APs to the total output uncertainty was implemented by running the model again another four times where each time only one of the four groups was made uncertain. The rest of the groups were fixed at their default averaged values. The uncertainty contribution of each group was derived by dividing the variance of the outputs estimated when only this group was made uncertain by the variance of the outputs estimated when all four groups were uncertain during the first time the model was run. An R-script was written for the estimation of these contributions and the results, which were visualised with stack-columns, showed that the NCU group is the main uncertainty source contributing to the total output uncertainty.

5.3 Recommendations for further research

Some steps that can be recommended for further research are the following:

- Include uncertainties in output fluxes released from non-agricultural soils as well;
- Include uncertainty in model inputs such as climatic variables, land cover, soil type, soil pH etc. concerning both agricultural and non-agricultural soils;
- Include model structure uncertainty by quantifying residual errors of modelling approaches used to calculate N emissions (e.g. N_2O from natural areas);
- Assess uncertainty in INTEGRATOR outputs by comparing them with outputs produced from other models with different complexity and data requirements. Such an approach is described in the paper of De Vries et al. (2011a) who compared land N budget (i.e. N inputs – N removals) of agricultural systems in Europe, for the year 2000, that was estimated by four different models: INTEGRATOR, IDEAg, MITERRA-EUROPE, and IMAGE. This approach is an optimal way to get the present insight into the problem of quantifying uncertainties in N output fluxes;
- Investigate how uncertainty in INTEGRATOR outputs can be further reduced by improving the emission and leaching/runoff factors as well as the simple process-based and empirical approaches incorporated in the model to calculate N and GHG emissions.

REFERENCES

Crosetto, M. and Tarantola, S. (2001), Uncertainty and sensitivity analysis: tools for GIS-based model implementation. International Journal of Geographical Information Science 15, pp 415–437.

De Vries, W., A. Leip, G. J. Reinds, J. Kros, J. P. Lesschen and A.F. Bouwman, (2011a). Comparison of land nitrogen budgets for European agriculture by various modelling approaches, Environmental Pollution, DOI 10.1016/j.envpol.2011.03.038

De Vries, W., A. Leip, G. J. Reinds, J. Kros, J. P. Lesschen, A.F. Bouwman, B Grizzetti, F. Bouraoui, K. Butterbach Bahl, P. Bergamaschi and W. Winiwarter, (2011b). Geographic variation in terrestrial nitrogen budgets across Europe. Chapter 15 in: The European Nitrogen Assessment (Eds. Sutton M.A., Howard C., Erisman J.W., Billen G., Bleeker A., Grennfelt P., van Grinsven H. and Grizzetti B.), Cambridge University Press: 317-344.

De Vries, W., J. Kros, G. J. Reinds, H.J.J.Wieggers, G.L. Velthof, D.A. Oudendag, J.P Lesschen, O. Oenema, G. J. Nabuurs, M.J. Schelhaas, M. Perez Soba, W. Rienks, W.P. de Winter, J. van den Akker, A. Leip, M.M. Bakker, P.H Verburg, K. Neumann, J. Liski, B. Eickhout and A.F. Bouwman, (2011c), INTEGRATOR: A modelling tool for European-wide assessments of nitrogen and greenhouse gas fluxes in response to changes in land cover, land management and climate. Wageningen, the Netherlands, Alterra Report (in press).

De Vries, W., J. Kros, G. J. Reinds, H.J.J.Wieggers, G.L. Velthof, D.A. Oudendag, J.P Lesschen, O. Oenema, G. J. Nabuurs, M.J. Schelhaas, M. Perez Soba, W. Rienks, W.P. de Winter, J. van den Akker, A. Leip, M.M. Bakker, P.H Verburg, K. Neumann, J. Liski, B. Eickhout and A.F. Bouwman, (2011d), Assessment of nitrogen and green house fluxes at the European scale in response to land cover, livestock and land management change, Agriculture, Ecosystems and Environment (in prep).

Galloway, J. N. et al. (2004), Nitrogen cycles: past, present and future, Biogeochem, 70 (2), pp153-226.

Heuvelink, G. B. M. et al. (2009), Operational methodology for spatial uncertainty quantification and spatial uncertainty analysis of regional scale NEU models, Alterra, Wageningen UR.

Jansen, M.J.W. (2005), ASSA: Algorithms for Stochastic Sensitivity Analysis. Biometris report, Wageningen UR.

Jansen, M.J.W. et al. (1994), Monte Carlo estimation of uncertainty contributions from several independent multivariate sources: In: Grasman, J. and G. Van Straten (Eds), Predictability and Nonlinear Modelling in Natural Sciences and Economics, Kluwer, Dordrecht: 334-343.

Intergovernmental Panel on Climate Change (IPCC) official Website: http://www.ipcc.ch/

Kros, H. et al. (2010), Uncertainty assessment with INTEGRATOR in NitroEurope, Alterra, Wageningen UR.

Lesschen, J.P. et al. (2007), A spatially explicit methodology to quantify soil nutrient balances and their uncertainties at the national level, Nutrient Cycling in Agroecosystems 78 (2), 111-131

Lesschen, J.P., Velthof, G.L., Kros, J., de Vries, W., (2011), Estimation of N_2O emission factors for soils depending on environmental conditions and crop management, Environmental Pollution, DOI 10.1016/j.envpol.2011.04.001.

Nitro Europe (NEU) official website: http://www.nitroeurope.eu/

Oenema, O. (2001), Technical and policy aspects of strategies to decrease greenhouse gas emissions from agriculture, pp 301-304, Alterra, Wageningen UR

Saltelli, A., K. Chan and E.M. Scott (Eds) (2000), Sensitivity Analysis, Wiley, Chichester.

Schultze et al. (2009), Importance of methane and nitrous oxide for Europe's terrestrial greenhouse-gas balance. 2 (12), 842-850.

Truong, T. N. (2009), Uncertainty Analysis of predicting terrestrial Nitrous-Oxide emissions with INTEGRATOR, Wageningen UR.

United Nations Framework Convention on Climate Change (UNFCCC) official Website: http://unfccc.int/essential_background/items/2877.php

Van Oijen, M. (2006), Protocol – UQ: A protocol for uncertainty quantification and uncertainty analysis of models in NEU

Velthof et al. (2009), Integrated Assessment of Nitrogen Losses from Agriculture in EU-27 using MITERRA-EUROPE. J Environ Qual 38 (2), 402-417.

Vitousek, J. P. W. et al. (1997), Human alteration of the global nitrogen cycle: sources and consequences. Ecol. appl. 7 (3), pp 737-750.

APPENDICES

Appendix 1. Tables with the pdf parameters of each AP according to the three robustness scenarios

Reference scenario										
Code	**Level**	**Distribution**	**CV**	**SD**	**Min**	**Max**	**Unit**	**ρNCU**	**ρNUTS**	**ρCOUNTRY**
aninr_ca	CNTRY	Normal	0.1		0	inf	Head	1	1	0.5
aninr_oc	CNTRY	Normal	0.1		0	inf	Head	1	1	0.5
aninr_pp	CNTRY	Normal	0.1		0	inf	Head	1	1	0.5
aninr_po	CNTRY	Normal	0.1		0	inf	Head	1	1	0.5
Nexf_ca	NUTS	Normal	0.25		0	inf	kg N / head	1	0.85	0.5
Nexf_oc	NUTS	Normal	0.25		0	inf	kg N / head	1	0.85	0.5
Nexf_pp	NUTS	Normal	0.25		0	inf	kg N / head	1	0.85	0.5
Nexf_po	NUTS	Normal	0.25		0	inf	kg N / head	1	0.2	0.2
Cnam	NUTS	Normal	0.25		0	inf	kg C / Kg N	1	0.2	0.2
f_Nav_dep	GEN	Normal	0.25		0	1	-	1	1	1
f_Nav_om	GEN	Normal	0.25		0	1	-	1	1	1
fhs_ca	NUTS	Normal	0.25		0	1	-	1	0.5	0.2
fhs_oc	NUTS	Normal	0.25		0	1	-	1	0.5	0.2
frlam_ca	NUTS	Normal	0.25		0	1	-	1	0.85	0.5
frlam_pp	NUTS	Normal	0.25		0	1	-	1	0.85	0.5
fNemhs_NH3	NUTS	Normal	0.25		0	1	-	1	0.85	0.5
fNemms_NH3	NUTS	Normal	0.25		0	1	-	1	0.85	0.5
fNemhsl_N2O	NUTS	Lognormal		0.25	-inf	0	-	1	0.85	0.5
fNemmsl_N2O	NUTS	Lognormal		0.25	-inf	0	-	1	0.85	0.5
fNemhsl_NO	NUTS	Lognormal		0.25	-inf	0	-	1	0.85	0.5
fNemmsl_NO	NUTS	Lognormal		0.25	-inf	0	-	1	0.85	0.5
fNemhss_N2O	NUTS	Lognormal		0.5	-inf	0	-	1	0.85	0.5
fNemmss_N2O	NUTS	Lognormal		0.5	-inf	0	-	1	0.85	0.5
fNemhss_NO	NUTS	Lognormal		0.5	-inf	0	-	1	0.85	0.5
fNemmss_NO	NUTS	Lognormal		0.5	-inf	0	-	1	0.85	0.5
wamhsara	NUTS	Normal	0.25		0	1	-	1	0.5	0.2
wamhsgrass	NUTS	Normal	0.25		0	inf	-	1	0.5	0.2
Area_int	NUTS	Normal	0.25		0	inf	-	1	0.5	0.2

Reference scenario										
Code	**Level**	**Distribution**	**CV**	**SD**	**Min**	**Max**	**Unit**	**ρNCU**	**ρNUTS**	**ρCOUNTRY**
tNfe	CNTRY	Normal	0.1		0	inf	ton N / country	1	1	0.5
Ndep	NCU	Normal	0.25		0	inf	kg N / ha	0.5	0.2	0.2
Nfix_ar	NCU	Normal	0.5		0	inf	kg N / ha	0.5	0.2	0.2
Nfix_gr	NCU	Normal	0.5		0	inf	kg N / ha	0.5	0.2	0.2
Nfix_le	NCU	Normal	0.25		0	inf	kg N / ha	0.5	0.2	0.2
Yieldopt_ar	NCU	Normal	0.25		0	inf	ton FW / ha	0.85	0.5	0.2
Yieldopt_fo	NCU	Normal	0.25		0	inf	ton FW / ha	0.85	0.5	0.2
Yieldopt_gi	NCU	Normal	0.25		0	inf	ton FW / ha	0.85	0.85	0.5
Yieldopt_ge	NCU	Normal	0.25		0	inf	ton FW / ha	0.85	0.5	0.5
ctNplmx_ar	NCU	Normal	0.25		0	inf	g N / kg FW	0.5	0.2	0.2
ctNplmx_fo	NCU	Normal	0.25		0	inf	g N / kg FW	0.5	0.5	0.2
ctNplmx_gi	NCU	Normal	0.25		0	inf	g N / kg FW	0.5	0.5	0.2
ctNplmx_ge	NCU	Normal	0.25		0	inf	g N / kg FW	0.85	0.85	0.5
Nind	NCU	Normal	0.25		0	inf	-	0.5	0.2	0.2
fup	NCU	Normal	0.1		0	1	-	0.5	0.5	0.2
frmin	NCU	Normal	0.25		0	inf	-	0.5	0.5	0.2
Ninmx	NCU	Normal	0.25		0	inf	Kg N /ha	0.5	0.5	0.2
CNso_ms	NCU	Normal	0.25		0	inf	kg C / kg N	0.5	0.5	0.2
fNemap_NH3	NCU	Normal	0.25		0	1	-	0.5	0.5	0.2
fNemsi_N2O	NCU	Normal	0.5		0	1	-	0.2	0.2	0.2
rNON2O	NCU	Lognormal		0.75	-inf	0	-	0.5	0.2	0.2
fNle	NCU	Normal	0.25		0	1	-	0.5	0.5	0.2
flems	NUTS	Normal	0.5		0	1	-	1	0.85	0.5
fsr	NCU	Normal	0.25		0	1	-	0.5	0.5	0.2
fro	NCU	Normal	0.25		0	1	-	0.5	0.5	0.2
a_CH4_ca	NUTS	Normal	0.25		0	inf	kgCH4/kg milk	1	0.85	0.5
b_CH4_ca	NUTS	Normal	0.25		0	inf	kg CH4 per head	1	0.85	0.5
CH4_oc	NUTS	Normal	0.25		0	inf	kg CH4 per head	1	0.85	0.5

Optimistic scenario										
Code	**Level**	**Distribution**	**CV**	**SD**	**Min**	**Max**	**Unit**	**ρNCU**	**ρNUTS**	**ρCOUNTRY**
aninr_ca	CNTRY	Normal	0.05		0	inf	Head	1	1	0.3
aninr_oc	CNTRY	Normal	0.05		0	inf	Head	1	1	0.3
aninr_pp	CNTRY	Normal	0.05		0	inf	Head	1	1	0.3
aninr_po	CNTRY	Normal	0.05		0	inf	Head	1	1	0.3
Nexf_ca	NUTS	Normal	0.1		0	inf	kg N / head	1	0.8	0.3
Nexf_oc	NUTS	Normal	0.1		0	inf	kg N / head	1	0.8	0.3
Nexf_pp	NUTS	Normal	0.1		0	inf	kg N / head	1	0.8	0.3
Nexf_po	NUTS	Normal	0.1		0	inf	kg N / head	1	0.1	0.1
Cnam	NUTS	Normal	0.1		0	inf	kg C / Kg N	1	0.1	0.1
f_Nav_dep	GEN	Normal	0.1		0	1	-	1	1	1
f_Nav_om	GEN	Normal	0.1		0	1	-	1	1	1
fhs_ca	NUTS	Normal	0.1		0	1	-	1	0.3	0.1
fhs_oc	NUTS	Normal	0.1		0	1	-	1	0.3	0.1
frlam_ca	NUTS	Normal	0.1		0	1	-	1	0.8	0.3
frlam_pp	NUTS	Normal	0.1		0	1	-	1	0.8	0.3
fNemhs_NH3	NUTS	Normal	0.1		0	1	-	1	0.8	0.3
fNemms_NH3	NUTS	Normal	0.1		0	1	-	1	0.8	0.3
fNemhsl_N2O	NUTS	Lognormal		0.125	-inf	0	-	1	0.8	0.3
fNemmsl_N2O	NUTS	Lognormal		0.125	-inf	0	-	1	0.8	0.3
fNemhsl_NO	NUTS	Lognormal		0.125	-inf	0	-	1	0.8	0.3
fNemmsl_NO	NUTS	Lognormal		0.125	-inf	0	-	1	0.8	0.3
fNemhss_N2O	NUTS	Lognormal		0.25	-inf	0	-	1	0.8	0.3
fNemmss_N2O	NUTS	Lognormal		0.25	-inf	0	-	1	0.8	0.3
fNemhss_NO	NUTS	Lognormal		0.25	-inf	0	-	1	0.8	0.3
fNemmss_NO	NUTS	Lognormal		0.25	-inf	0	-	1	0.8	0.3
wamhsara	NUTS	Normal	0.1		0	1	-	1	0.3	0.1
wamhsgrass	NUTS	Normal	0.1		0	inf	-	1	0.3	0.1
Area_int	NUTS	Normal	0.1		0	inf	-	1	0.3	0.1
tNfe	CNTRY	Normal	0.05		0	inf	ton N / country	1	1	0.3
Ndep	NCU	Normal	0.1		0	inf	kg N / ha	0.3	0.1	0.1
Nfix_ar	NCU	Normal	0.4		0	inf	kg N / ha	0.3	0.1	0.1

Optimistic scenario										
Code	**Level**	**Distribution**	**CV**	**SD**	**Min**	**Max**	**Unit**	**ρNCU**	**ρNUTS**	**ρCOUNTRY**
Nfix_gr	NCU	Normal	0.4		0	inf	kg N / ha	0.3	0.1	0.1
Nfix_le	NCU	Normal	0.1		0	inf	kg N / ha	0.3	0.1	0.1
Yieldopt_ar	NCU	Normal	0.1		0	inf	ton FW / ha	0.8	0.3	0.1
Yieldopt_fo	NCU	Normal	0.1		0	inf	ton FW / ha	0.8	0.3	0.1
Yieldopt_gi	NCU	Normal	0.1		0	inf	ton FW / ha	0.8	0.8	0.3
Yieldopt_ge	NCU	Normal	0.1		0	inf	ton FW / ha	0.8	0.3	0.3
ctNplmx_ar	NCU	Normal	0.1		0	inf	g N / kg FW	0.3	0.1	0.1
ctNplmx_fo	NCU	Normal	0.1		0	inf	g N / kg FW	0.3	0.3	0.1
ctNplmx_gi	NCU	Normal	0.1		0	inf	g N / kg FW	0.3	0.3	0.1
ctNplmx_ge	NCU	Normal	0.1		0	inf	g N / kg FW	0.8	0.8	0.3
Nind	NCU	Normal	0.1		0	inf	-	0.3	0.1	0.1
fup	NCU	Normal	0.05		0	1	-	0.3	0.3	0.1
frmin	NCU	Normal	0.1		0	inf	-	0.3	0.3	0.1
Ninmx	NCU	Normal	0.1		0	inf	Kg N /ha	0.3	0.3	0.1
Cnso_ms	NCU	Normal	0.1		0	inf	kg C / kg N	0.3	0.3	0.1
fNemap_NH3	NCU	Normal	0.1		0	1	-	0.3	0.3	0.1
fNemsi_N2O	NCU	Normal	0.4		0	1	-	0.1	0.1	0.1
rNON2O	NCU	Lognormal		0.375	-inf	0	-	0.3	0.1	0.1
fNle	NCU	Normal	0.1		0	1	-	0.3	0.3	0.1
flems	NUTS	Normal	0.4		0	1	-	1	0.8	0.3
fsr	NCU	Normal	0.1		0	1	-	0.3	0.3	0.1
fro	NCU	Normal	0.1		0	1	-	0.3	0.3	0.1
a_CH4_ca	NUTS	Normal	0.1		0	inf	kgCH4/kg milk	1	0.8	0.3
b_CH4_ca	NUTS	Normal	0.1		0	inf	kg CH4 per head	1	0.8	0.3
CH4_oc	NUTS	Normal	0.1		0	inf	kg CH4 per head	1	0.8	0.3

Pessimistic Scenario										
Code	**Level**	**Distribution**	**CV**	**SD**	**Min**	**Max**	**Unit**	**ρNCU**	**ρNUTS**	**ρCOUNTRY**
aninr_ca	CNTRY	Normal	0.15		0	inf	Head	1	1	0.7
aninr_oc	CNTRY	Normal	0.15		0	inf	Head	1	1	0.7
aninr_pp	CNTRY	Normal	0.15		0	inf	Head	1	1	0.7
aninr_po	CNTRY	Normal	0.15		0	inf	Head	1	1	0.7
Nexf_ca	NUTS	Normal	0.3		0	inf	kg N / head	1	0.9	0.7
Nexf_oc	NUTS	Normal	0.3		0	inf	kg N / head	1	0.9	0.7
Nexf_pp	NUTS	Normal	0.3		0	inf	kg N / head	1	0.9	0.7
Nexf_po	NUTS	Normal	0.3		0	inf	kg N / head	1	0.3	0.3
Cnam	NUTS	Normal	0.3		0	inf	kg C / Kg N	1	0.3	0.3
f_Nav_dep	GEN	Normal	0.3		0	1	-	1	1	1
f_Nav_om	GEN	Normal	0.3		0	1	-	1	1	1
fhs_ca	NUTS	Normal	0.3		0	1	-	1	0.7	0.3
fhs_oc	NUTS	Normal	0.3		0	1	-	1	0.7	0.3
frlam_ca	NUTS	Normal	0.3		0	1	-	1	0.9	0.7
frlam_pp	NUTS	Normal	0.3		0	1	-	1	0.9	0.7
fNemhs_NH3	NUTS	Normal	0.3		0	1	-	1	0.9	0.7
fNemms_NH3	NUTS	Normal	0.3		0	1	-	1	0.9	0.7
fNemhsl_N2O	NUTS	Lognormal		0.375	-inf	0	-	1	0.9	0.7
fNemmsl_N2O	NUTS	Lognormal		0.375	-inf	0	-	1	0.9	0.7
fNemhsl_NO	NUTS	Lognormal		0.375	-inf	0	-	1	0.9	0.7
fNemmsl_NO	NUTS	Lognormal		0.375	-inf	0	-	1	0.9	0.7
fNemhss_N2O	NUTS	Lognormal		0.75	-inf	0	-	1	0.9	0.7
fNemmss_N2O	NUTS	Lognormal		0.75	-inf	0	-	1	0.9	0.7
fNemhss_NO	NUTS	Lognormal		0.75	-inf	0	-	1	0.9	0.7
fNemmss_NO	NUTS	Lognormal		0.75	-inf	0	-	1	0.9	0.7
wamhsara	NUTS	Normal	0.3		0	1	-	1	0.7	0.3
wamhsgrass	NUTS	Normal	0.3		0	inf	-	1	0.7	0.3
Area_int	NUTS	Normal	0.3		0	inf	-	1	0.7	0.3
tNfe	CNTRY	Normal	0.15		0	inf	ton N / country	1	1	0.7
Ndep	NCU	Normal	0.3		0	inf	kg N / ha	0.7	0.3	0.3
Nfix_ar	NCU	Normal	0.6		0	inf	kg N / ha	0.7	0.3	0.3

Pessimistic Scenario										
Code	**Level**	**Distribution**	**CV**	**SD**	**Min**	**Max**	**Unit**	**ρNCU**	**ρNUTS**	**ρCOUNTRY**
Nfix_gr	NCU	Normal	0.6		0	inf	kg N / ha	0.7	0.3	0.3
Nfix_le	NCU	Normal	0.3		0	inf	kg N / ha	0.7	0.3	0.3
Yieldopt_ar	NCU	Normal	0.3		0	inf	ton FW / ha	0.9	0.7	0.3
Yieldopt_fo	NCU	Normal	0.3		0	inf	ton FW / ha	0.9	0.7	0.3
Yieldopt_gi	NCU	Normal	0.3		0	inf	ton FW / ha	0.9	0.9	0.7
Yieldopt_ge	NCU	Normal	0.3		0	inf	ton FW / ha	0.9	0.7	0.7
ctNplmx_ar	NCU	Normal	0.3		0	inf	g N / kg FW	0.7	0.3	0.3
ctNplmx_fo	NCU	Normal	0.3		0	inf	g N / kg FW	0.7	0.7	0.3
ctNplmx_gi	NCU	Normal	0.3		0	inf	g N / kg FW	0.7	0.7	0.3
ctNplmx_ge	NCU	Normal	0.3		0	inf	g N / kg FW	0.9	0.9	0.7
Nind	NCU	Normal	0.3		0	inf	-	0.7	0.3	0.3
fup	NCU	Normal	0.15		0	1	-	0.7	0.7	0.3
frmin	NCU	Normal	0.3		0	inf	-	0.7	0.7	0.3
Ninmx	NCU	Normal	0.3		0	inf	Kg N /ha	0.7	0.7	0.3
CNso_ms	NCU	Normal	0.3		0	inf	kg C / kg N	0.7	0.7	0.3
fNemap_NH3	NCU	Normal	0.3		0	1	-	0.7	0.7	0.3
fNemsi_N2O	NCU	Normal	0.6		0	1	-	0.3	0.3	0.3
rNON2O	NCU	Lognormal		1.125	-inf	1	-	0.7	0.3	0.3
fNle	NCU	Normal	0.3		0	1	-	0.7	0.7	0.3
flems	NUTS	Normal	0.6		0	inf	-	1	0.9	0.7
fsr	NCU	Normal	0.3		0	inf	-	0.7	0.7	0.3
fro	NCU	Normal	0.3		0	inf	-	0.7	0.7	0.3
a_CH4_ca	NUTS	Normal	0.3		0	inf	Kg CH4/kg milk	1	0.9	0.7
b_CH4_ca	NUTS	Normal	0.3		0	inf	kg CH4 per head	1	0.9	0.7
CH4_oc	NUTS	Normal	0.3		0	inf	kg CH4 per head	1	0.9	0.7

Appendix 2. Algorithms for: A. Generating the MC input files – B. Running the model and storing the MC output files

A. Generating the MC input files

The steps of the algorithm developed for the generation of the 1000 MC input files for the APs of the NUTS group were the following:

1. The individual NUTS regions with their ids were created
2. The pdf parameters for each AP from NUTS group, for the Reference scenario, were read
3. The spatial-correlation coefficients matrices for each AP from NUTS group were produced within a "for loop"
4. The 1000 simulations of $\xi(i)$ were generated by the "`rmultnorm`" function
5. The simulated value of $\xi(i)$ was post-processed depending on whether the $\sigma(i)$ or the CV(i) value was available
6. The 1000 MC input files for each AP (per NUTS region) were produced within a "for loop" and they were stored in corresponding folders.

With similar algorithms, 1000 Monte Carlo input files were created for each AP from the CNTRY group (per Country), and for each AP from the GEN group (for the whole Europe).

The sequence number of each separate input file was indicated in the file name itself. For instance, the input files for NUTS parameters were:
NUTSPAR0001.CSV
NUTSPAR0002.CSV
…,
NUTSPAR0999.CSV
NUTSPAR1000.CSV

The format of the MC input files, for each of the three groups, was the following:

***NUTS parameters*:** The file had 688 rows with each one of them representing one NUTS region, and 80 columns that are described below:

Column 1: "NUTS_id"
Column 2: "COUNTRY_id"
Column 3: "Dist_id" of AP_1 (Distribution type: **N** for Normal and **L** for Lognormal)
Column 4: "Disp_id" of AP_1 (Dispersion type: **C** for CV and **S** for SD)
Column 5: "ParAP_1" (i) (Simulated value for AP_1 where i the number of the current simulation)
Column 6: "Dist_id" of AP_2
Column 7: "Disp_id" of AP_2
Column 8: "ParAP_2" (i)
Column 9: "Dist_id" of AP_3
Column 10: "Disp_id" of AP_3
Column 11: "ParAP_3" (i)
…
…
Column 2+3n: "ParAP_n" (i) (where n=26)

***CNTRY parameters*:** The file had 25 rows with each one of them representing one country of EU-25, and 17 columns that are described below:

Column 1: "COUNTRY_id"
Column 2: "COUNTRY_nr"
Column 3: "Dist_id" of AP_1
Column 4: "Disp_id" of AP_1
Column 5: "$ParAP_1$" (i)
Column 6: "Dist_id" of AP_2
Column 7: "Disp_id" of AP_2
Column 8: "$ParAP_2$" (i)
Column 9: "Dist_id" of AP_3
Column 10: "Disp_id" of AP_3
Column 11: "$ParAP_3$" (i)
...
...
Column 2+3n: "$ParAP_n$" (i) (where n=5)

***GEN parameters*:** The file had 1 row representing the whole Europe, and 6 columns, bearing the information for the 2 APs of the GEN group, that are described below:

Column 1: "Dist_id" of AP_1
Column 2: "Disp_id" of AP_1
Column 3: "$ParAP_1$" (i)
Column 4: "Dist_id" of AP_2
Column 5: "Disp_id" of AP_2
Column 6: "$ParAP_2$" (i) (where n=2)

B. Running the model and storing the MC output files

The steps of the algorithm developed for running the model in batch mode and storing its output files were the following:

1. The dll was loaded
2. The functions defined in the dll were called
3. The input and output parameters (directories) were set up
4. Each input file from NCU, NUTS, CNTRY and GEN groups was provided
5. The model was run
6. One output file for NH3, NOx, NleSW, NleGW, CH4, and N2O was produced at Country and European level and it was stored in a corresponding folder.

Steps 4, 5 and 6 were being repeated 1000 times, within a "for loop".

The sequence number of each separate output file was indicated in the file name itself. For example:

- for N2O emissions: N2Oemis0001.CSV, N2Oemis0002.CSV,…, N2Oemis0999.CSV, N2Oemis 1000.CSV;
- for CH4 emissions: CH4emis0001.CSV, CH4emis0002.CSV,…, CH4emis0999.CSV, CH4emis 1000.CSV ;
- and so on for the rest of the model outputs.

These output files contained two columns: the “polyID” and the “val”. The “polyID” was linked to each country of EU-25 in the Country level outputs, and to Europe in the European level outputs. The “val” column was bearing the information about the output value for the six fluxes.

Appendix 3. R-script for the generation of 1000 MC input files for the APs of NUTS group for the Reference scenario

```
#SCRIPT FOR THE GENERATION OF 1000 MC INPUT FILES FOR NUTS APs

#clean-up memory:
rm(list = ls())

#read the NCUs with their associated NUTS and COUNTRIES:
NCU = read.csv(file = 'NCU26.csv', header = TRUE, sep = ",",
na.strings = '#N/A')

#group by the NUTS with their COUNTRIES from NCU(indirectly with
#the aggregation function with FUN=min):
NUTS = aggregate(NCU$NCU_NR, list(NUTS = NCU$FSS_NUTS_N, COUNTRY =
NCU$COUNTRY), min)

#read the agricultural parameters with their associated statistical
#info and their correlations:
paramALL = read.csv(file = 'Parameters_Ref.csv', header = TRUE, sep
= ",")

#select only the NUTS parameters:
param = subset(paramALL, paramALL$Level=="NUTS")

#define n as the number of rows of NUTS:
n = dim(NUTS)[1]

#define m as the number of parameters:
m = dim(param)[1]

#load "mtvnorm" package:
library(MSBVAR)

#define the number of MC runs:
k=1000

#create the initial dataframe ds containing the NUTS and the
#COUNTRIES:
ds = data.frame(NUTS$NUTS, NUTS$COUNTRY)

#then assign an index to NUTS parameters in list and make a loop to
#create the RHO matrix for every parameter:
for (parindex in 1:m) {

#initialize correlation matrix with rhoCOUNTRY:
RHO = matrix(data = param$rhoCOUNTRY[parindex], nrow = n, ncol = n)

#replace diagonal with 1s:
for (i in 1:n){
  RHO[i,i] = 1}

#replace correlation of pairs in the same country with rhoNUTS:
for (i in 1:(n-1)) {
 for (j in (i+1):n)  {
   if (NUTS$COUNTRY[i] == NUTS$COUNTRY[j])
     RHO[i,j] = param$rhoNUTS[parindex]
     RHO[j,i] = RHO[i,j] } }
```

```
#simulate the ξ(i)(ksi) from the multivariate normal distribution:
sim = rmultnorm(n = k, rep(0, times=n), vmat = RHO, tol = 1e-10)

#post-process the simulations of ksi depending on parameters
#characterized by CV or SD:
if(!is.na(param$CV[parindex]))
      sim = sim * param$CV[parindex]+ 1
   else
      sim = sim * param$SD[parindex]

#append the transposed matrix "t(sim)" with the simulated values to
#the dataframe ds:
ds = data.frame(ds, t(sim))

# end of parindex loop:
}

# make a loop to generate the MC input files with the required
#format:
for (MC in 1:k)  {

#create a second dataframe ds2 which contains only the 2 first
#columns from the ds:
ds2 = data.frame(ds[,1:2])

#rename the columns of ds2 to the required names:
names(ds2)= c("NUTS_id", "COUNTRY_id")

#make a loop for every parameter to create a dataframe with 3
#intermediate columns containing info for its pdf (N for
#Normal or L for Lognormal), its dispersion(S for SD or C for
#CV)and its simulated values:
for (i in 1:m) {
ds3=
data.frame(substring(param$Distribution[i],1,1),ifelse(!is.na(param$
SD[i]), "S", "C"), ds[,(i-1)*k+2+MC])

#create a short name for the NUTS parameters:
parname = param$Code[i]

#rename the columns of ds3 to the required names:
names(ds3)= c(paste("Dist_id",parname, sep=""), paste("Disp_id",
parname, sep=""), paste("Par", parname, sep=""))

#create a new dataframe ds2 by merging the old ds2 and ds3:
ds2 = data.frame(ds2,ds3)}

#write the final MC input "csv" files:
write.table(ds2, file =
sprintf("E:\\Vicky_Ioannidi\\Reference_Scenario\\INTEGRATOR_ref\\inp
uts\\Nuts_param\\NUTSPAR%04d.csv", MC), row.names = FALSE, col.names
= TRUE, sep = ",")

# end of the MC loop:
}

#END OF SCRIPT
```

Appendix 4. Example of one Monte Carlo input file for the five APs of the CNTRY group

COU	COU	Dist_i	Disp_	Paraninr_ca	Dist_	Disp_	Paraninr_oc	Dist_i	Disp_	Paraninr_pp	Dist_	Disp_	Paraninr_p	Dist_	Disp_	PartNfe
AT	4	N	C	0.93571677	N	C	0.91668107	N	C	1.079533058	N	C	1.0962234	N	C	0.795453
BE	7	N	C	1.08582994	N	C	0.97437262	N	C	1.097869682	N	C	0.9272819	N	C	0.889612
BG	8	N	C	0.98555756	N	C	0.95692836	N	C	1.129105349	N	C	1.0512086	N	C	0.903046
CZ	13	N	C	0.96364461	N	C	1.04755618	N	C	1.145130327	N	C	1.1683666	N	C	0.908529
DE	14	N	C	1.08787191	N	C	1.02266543	N	C	1.094301854	N	C	1.1791157	N	C	0.877948
DK	15	N	C	1.08040041	N	C	0.91037602	N	C	1.020857047	N	C	1.0898923	N	C	0.953339
EE	18	N	C	0.99317769	N	C	0.84702934	N	C	1.071782378	N	C	1.0989936	N	C	0.809936
ES	21	N	C	1.03071022	N	C	0.80414628	N	C	0.91158571	N	C	1.0960063	N	C	0.808397
FI	22	N	C	1.00789564	N	C	0.875619	N	C	0.993971262	N	C	1.1224024	N	C	0.912848
FR	24	N	C	1.01852473	N	C	0.98394369	N	C	1.023291816	N	C	1.1259892	N	C	0.956589
GR	29	N	C	0.93485	N	C	0.83885825	N	C	1.106617577	N	C	1.0159225	N	C	0.783508
HU	31	N	C	1.02416799	N	C	0.88481654	N	C	1.115009573	N	C	1.0262627	N	C	0.908803
IE	32	N	C	1.07378544	N	C	0.84434422	N	C	1.074384027	N	C	1.1636506	N	C	1.023695
IT	38	N	C	1.11739001	N	C	1.06992931	N	C	1.04682987	N	C	1.0897852	N	C	0.837384
LT	44	N	C	1.02533506	N	C	0.89573435	N	C	1.070716078	N	C	0.9746023	N	C	0.773973
LU	45	N	C	1.11256776	N	C	0.88735275	N	C	1.019431453	N	C	1.0410002	N	C	0.728808
LV	46	N	C	1.08990304	N	C	0.99771311	N	C	1.050504825	N	C	1.0968218	N	C	0.880541
NL	54	N	C	0.97461186	N	C	0.7859262	N	C	1.147562434	N	C	1.0626958	N	C	0.896321
PL	56	N	C	1.07875765	N	C	0.93114879	N	C	1.03664067	N	C	1.0514875	N	C	0.723076
PT	57	N	C	1.06101574	N	C	0.95162977	N	C	1.011891745	N	C	1.1508147	N	C	0.903981
RO	58	N	C	0.9819946	N	C	0.98342833	N	C	1.052775182	N	C	1.0976766	N	C	0.86743
SE	61	N	C	1.02198194	N	C	0.91115323	N	C	1.077836002	N	C	1.0553082	N	C	0.978299
SI	62	N	C	0.99931845	N	C	0.86994841	N	C	1.046929392	N	C	1.2526009	N	C	0.912108
SK	64	N	C	0.94642115	N	C	1.01523393	N	C	1.141469067	N	C	1.0602099	N	C	0.833672
UK	70	N	C	1.00035663	N	C	0.9606476	N	C	1.189480293	N	C	1.0239408	N	C	0.861611

Appendix 5. R-script for running the model for the Optimistic scenario and storing its MC output files

```
#SCRIPT FOR RUNNING THE MODEL IN BATCH MODE AND STORING ITS OUTPUTS

# set the working directory
setwd("E:\\Vicky_Ioannidi\\Optimistic_Scenario\\INTEGRATOR_Opt\\newi
nt")

# load the DLL
dyn.load("integratorua.dll")

# get the function addresses defined in the DLL
is.loaded("setmodeldirectory")
is.loaded("initializedll")
is.loaded("provideparamfile")
is.loaded("run")
is.loaded("getvalues")
is.loaded("closeintegrator")
is.loaded("getvalueslu")

# setup input and output parameters
theDir = character(length = 255)
theDir =
"E:\\Vicky_Ioannidi\\Optimistic_Scenario\\INTEGRATOR_Opt\\newint"
theFolder =
"E:\\Vicky_Ioannidi\\Optimistic_Scenario\\INTEGRATOR_Opt\\outputs\\i
npNCU_Nuts_Cntry_Gen\\"
ier = as.integer(0)
res =.Fortran("setmodeldirectory",err=ier,dir=theDir)
str(res)
thePath = character(length = 255)
thePath  =
"E:\\Vicky_Ioannidi\\Optimistic_Scenario\\INTEGRATOR_Opt\\output"
res =.Fortran("initializedll",err=ier, path=thePath)
str(res)

# reserve space for the 4 variables
theNCUFile  = character(length = 255)
theNutsFile  = character(length = 255)
theCntryFile  = character(length = 255)
theGenFile  = character(length = 255)
paramname = character(length=255)

# provide repeatedly the input files from NCU, NUTS, CNTRY and GEN
# groups
for (i in 1:1000){ #start of "for loop"
 theNCUFile = sprintf("E:\\Vicky_Ioannidi\\real_opt2\\real%04d.csv",
i)
 theLevel = as.integer(2) #NCU level
 res
=.Fortran("provideparamfile",err=ier,level=theLevel,file=theNCUFile)
 str(res)
 theNutsFile =
sprintf("E:\\Vicky_Ioannidi\\Optimistic_Scenario\\INTEGRATOR_Opt\\in
puts\\Nuts_param\\nutspar%04d.csv", i)
 theLevel = as.integer(1)#nuts level
 res
=.Fortran("provideparamfile",err=ier,level=theLevel,file=theNutsFile
)
 str(res)
```

```
 theCntryFile =
sprintf("E:\\Vicky_Ioannidi\\Optimistic_Scenario\\INTEGRATOR_Opt\\in
puts\\Cntry_param\\coupar%04d.csv", i)
 theLevel = as.integer(0)#country level
 res
=.Fortran("provideparamfile",err=ier,level=theLevel,file=theCntryFil
e)
 str(res)
 theGenFile =
sprintf("E:\\Vicky_Ioannidi\\Optimistic_Scenario\\INTEGRATOR_Opt\\in
puts\\Gen_param\\genpar%04d.csv", i)
 theLevel = as.integer(3)#generic level
 res
=.Fortran("provideparamfile",err=ier,level=theLevel,file=theGenFile)
 str(res)

 # run the model
 res =.Fortran("run",err=ier)
 str(res)
 # level: 0=EU, 1=country, 2=nuts, 3=ncu
 # Stat: 1 = sum over area (kton), 2=weigthed mean (kg/ha)
 Stat = as.integer(2)


 # WRITE OUTPUT FILES AT COUNTRY LEVEL
 level = as.integer(1)
 # dimension of values MUST be given by number
 valuesC = double(length=38)
 numberC = as.integer(38)

 paramname = ("N2Oemis")
 res
=.Fortran("getvalues",param=paramname,lev=level,stat=Stat,val=values
C,numb=numberC,err=ier)
 str(res)
 res$polyID = seq(1:length(res$val))-1
 write.table(cbind(res$polyID,res$val),
file=sprintf(paste(theFolder,
"\\Cntry_level\\N2Oemis\\N2Oemis%04d.csv", sep=""), i), sep=",",
row.names = FALSE, col.names =c("polyID","val"))

 paramname = ("NH3emis")
 res
=.Fortran("getvalues",param=paramname,lev=level,stat=Stat,val=values
C,numb=numberC,err=ier)
 str(res)
 res$polyID = seq(1:length(res$val))-1
 write.table(cbind(res$polyID,res$val),
file=sprintf(paste(theFolder,
"\\Cntry_level\\NH3emis\\NH3emis%04d.csv", sep=""), i), sep=",",
row.names = FALSE, col.names =c("polyID","val"))

 paramname = ("NOxemis")
 res
=.Fortran("getvalues",param=paramname,lev=level,stat=Stat,val=values
C,numb=numberC,err=ier)
 str(res)
 res$polyID = seq(1:length(res$val))-1
 write.table(cbind(res$polyID,res$val),
file=sprintf(paste(theFolder,
```

```
"\\Cntry_level\\NOxemis\\NOxemis%04d.csv", sep=""), i), sep=",",
row.names = FALSE, col.names =c("polyID","val"))

 paramname = ("CH4emis")
 res
=.Fortran("getvalues",param=paramname,lev=level,stat=Stat,val=values
C,numb=numberC,err=ier)
 str(res)
 res$polyID = seq(1:length(res$val))-1
 write.table(cbind(res$polyID,res$val),
file=sprintf(paste(theFolder,
"\\Cntry_level\\CH4emis\\CH4emis%04d.csv", sep=""), i), sep=",",
row.names = FALSE, col.names =c("polyID","val"))

 paramname = ("NleGW")
 res
=.Fortran("getvalues",param=paramname,lev=level,stat=Stat,val=values
C,numb=numberC,err=ier)
 str(res)
 res$polyID = seq(1:length(res$val))-1
 write.table(cbind(res$polyID,res$val),
file=sprintf(paste(theFolder, "\\Cntry_level\\NleGW\\NleGW%04d.csv",
sep=""), i), sep=",", row.names = FALSE, col.names
=c("polyID","val"))

 paramname = ("NleSW")
 res
=.Fortran("getvalues",param=paramname,lev=level,stat=Stat,val=values
C,numb=numberC,err=ier)
 str(res)
 res$polyID = seq(1:length(res$val))-1
 write.table(cbind(res$polyID,res$val),
file=sprintf(paste(theFolder, "\\Cntry_level\\NleSW\\NleSW%04d.csv",
sep=""), i), sep=",", row.names = FALSE, col.names
=c("polyID","val"))

 # WRITE OUTPUT FILES AT EUROPE LEVEL
 level = as.integer(0)
 # dimension of values MUST be given by number
 valuesE = double(length=1)
 numberE = as.integer(1)

 paramname = ("N2Oemis")
 res
=.Fortran("getvalues",param=paramname,lev=level,stat=Stat,val=values
E,numb=numberE,err=ier)
 str(res)
 res$polyID = seq(1:length(res$val))-1
 write.table(cbind(res$polyID,res$val),
file=sprintf(paste(theFolder,
"\\Europe_level\\N2Oemis\\N2Oemis%04d.csv", sep=""), i), sep=",",
row.names = FALSE, col.names =c("polyID","val"))

 paramname = ("NH3emis")
 res
=.Fortran("getvalues",param=paramname,lev=level,stat=Stat,val=values
E,numb=numberE,err=ier)
 str(res)
 res$polyID = seq(1:length(res$val))-1
 write.table(cbind(res$polyID,res$val),
file=sprintf(paste(theFolder,
```

```
"\\Europe_level\\NH3emis\\NH3emis%04d.csv", sep=""), i), sep=",",
row.names = FALSE, col.names =c("polyID","val"))

 paramname = ("NOxemis")
 res
=.Fortran("getvalues",param=paramname,lev=level,stat=Stat,val=values
E,numb=numberE,err=ier)
 str(res)
 res$polyID = seq(1:length(res$val))-1
 write.table(cbind(res$polyID,res$val),
file=sprintf(paste(theFolder,
"\\Europe_level\\NOxemis\\NOxemis%04d.csv", sep="") ,i), sep=",",
row.names = FALSE, col.names =c("polyID","val"))

 paramname = ("CH4emis")
 res
=.Fortran("getvalues",param=paramname,lev=level,stat=Stat,val=values
E,numb=numberE,err=ier)
 str(res)
 res$polyID = seq(1:length(res$val))-1
 write.table(cbind(res$polyID,res$val),
file=sprintf(paste(theFolder,
"\\Europe_level\\CH4emis\\CH4emis%04d.csv", sep=""), i), sep=",",
row.names = FALSE, col.names =c("polyID","val"))

 paramname = ("NleGW")
 res
=.Fortran("getvalues",param=paramname,lev=level,stat=Stat,val=values
E,numb=numberE,err=ier)
 str(res)
 res$polyID = seq(1:length(res$val))-1
 write.table(cbind(res$polyID,res$val),
file=sprintf(paste(theFolder,
"\\Europe_level\\NleGW\\NleGW%04d.csv", sep=""), i), sep=",",
row.names = FALSE, col.names =c("polyID","val"))

 paramname = ("NleSW")
 res
=.Fortran("getvalues",param=paramname,lev=level,stat=Stat,val=values
E,numb=numberE,err=ier)
 str(res)
 res$polyID = seq(1:length(res$val))-1
 write.table(cbind(res$polyID,res$val),
file=sprintf(paste(theFolder,
"\\Europe_level\\NleSW\\NleSW%04d.csv", sep=""), i), sep=",",
row.names = FALSE, col.names =c("polyID","val"))


#end of "for loop"
}

res2 =.Fortran("closeintegrator")
dyn.unload("integratorua.dll")

#END OF SCRIPT
```

Appendix 6. Two MC output files with values for N_2O emissions at Country and at European level

1. Country level

polyID	val
0	-99999
1	2.491507
2	5.138366
3	0.776849
4	-99999
5	-99999
6	2.410541
7	2.308482
8	1.680739
9	0.998365
10	0.727804
11	1.183139
12	-99999
13	2.536951
14	-99999
15	0.986919
16	1.375434
17	5.807291
18	-99999
19	-99999
20	1.774581
21	-99999
22	-99999
23	1.124262
24	4.551751
25	0.577001
26	-99999
27	-99999
28	4.663023
29	-99999
30	1.107753
31	1.090217
32	0.93709
33	1.033482
34	2.371039
35	1.186418
36	-99999
37	3.968247

Note: The cells with the code "-99999" are referring to countries that were stored in INTEGRATOR´s database but they were not included in EU-25 and thus no output value was produced for them.

2. European level

polyID	val
0	1.93578

Appendix 7. R-script for calculating means, SDs and CVs of the model outputs at Country level for the Reference scenario

```
#SCRIPT TO CALCULATE: means, SDs and CVs for CH4, N2O, NH3, NleGW,
#NleSW, NOx when INTEGRATOR inputs=NCU+NUTS+CNTRY+GEN APs
#and outputs = aggregated to Country Level

#clean up memory
rm(list = ls())

#create the initial dataframes for INTEGRATOR outputs with the
#PolyID and Country_ID
dsCH4 = read.csv(file = 'Countries.csv', header = TRUE, sep = ",",
na.strings = '#N/A')
dsN2O = read.csv(file = 'Countries.csv', header = TRUE, sep = ",",
na.strings = '#N/A')
dsNH3 = read.csv(file = 'Countries.csv', header = TRUE, sep = ",",
na.strings = '#N/A')
dsNleGW = read.csv(file = 'Countries.csv', header = TRUE, sep = ",",
na.strings = '#N/A')
dsNleSW = read.csv(file = 'Countries.csv', header = TRUE, sep = ",",
na.strings = '#N/A')
dsNOx = read.csv(file = 'Countries.csv', header = TRUE, sep = ",",
na.strings = '#N/A')


#make a loop to read the multiple output files for all the N
#emissions and N leaching and create their final dataframes
NMC =1000
for (i in 1:NMC)
{
   CH4emis =
read.csv(file=sprintf("E:\\Vicky_Ioannidi\\Reference_Scenario\\INTEG
RATOR_Ref\\outputs\\inpNCU_Nuts_Cntry_Gen\\Cntry_level\\CH4emis\\CH4
emis%04d.csv", i),
   header = TRUE, sep = ",",na.strings = '-99999')
   dsCH4 = data.frame(dsCH4, CH4emis$val)

   N2Oemis =
read.csv(file=sprintf("E:\\Vicky_Ioannidi\\Reference_Scenario\\INTEG
RATOR_Ref\\outputs\\inpNCU_Nuts_Cntry_Gen\\Cntry_level\\N2Oemis\\N2O
emis%04d.csv", i),
   header = TRUE, sep = ",",na.strings = '-99999')
   dsN2O = data.frame(dsN2O, N2Oemis$val)

   NH3emis =
read.csv(file=sprintf("E:\\Vicky_Ioannidi\\Reference_Scenario\\INTEG
RATOR_Ref\\outputs\\inpNCU_Nuts_Cntry_Gen\\Cntry_level\\NH3emis\\NH3
emis%04d.csv", i),
   header = TRUE, sep = ",",na.strings = '-99999')
   dsNH3 = data.frame(dsNH3, NH3emis$val)

   NleGW =
read.csv(file=sprintf("E:\\Vicky_Ioannidi\\Reference_Scenario\\INTEG
RATOR_Ref\\outputs\\inpNCU_Nuts_Cntry_Gen\\Cntry_level\\NleGW\\NleGW
%04d.csv", i),
   header = TRUE, sep = ",",na.strings = '-99999')
   dsNleGW = data.frame(dsNleGW, NleGW$val)

   NleSW =
read.csv(file=sprintf("E:\\Vicky_Ioannidi\\Reference_Scenario\\INTEG
```

```
RATOR_Ref\\outputs\\inpNCU_Nuts_Cntry_Gen\\Cntry_level\\NleSW\\NleSW
%04d.csv", i),
   header = TRUE, sep = ",",na.strings = '-99999')
   dsNleSW = data.frame(dsNleSW, NleSW$val)

   NOxemis =
read.csv(file=sprintf("E:\\Vicky_Ioannidi\\Reference_Scenario\\INTEG
RATOR_Ref\\outputs\\inpNCU_Nuts_Cntry_Gen\\Cntry_level\\NOxemis\\NOx
emis%04d.csv", i),
   header = TRUE, sep = ",",na.strings = '-99999')
   dsNOx = data.frame(dsNOx, NOxemis$val)
}

#rename the columns of the final dataframes
names(dsCH4)[3:ncol(dsCH4)]=c(paste("CH4emis_MC",(1:NMC), sep=""))
names(dsN2O)[3:ncol(dsN2O)]=c(paste("N2Oemis_MC",(1:NMC), sep=""))
names(dsNH3)[3:ncol(dsN2O)]=c(paste("NH3emis_MC",(1:NMC), sep=""))
names(dsNleGW)[3:ncol(dsNleGW)]=c(paste("NleGW_MC",(1:NMC), sep=""))
names(dsNleSW)[3:ncol(dsNleSW)]=c(paste("NleSW_MC",(1:NMC), sep=""))
names(dsNOx)[3:ncol(dsNOx)]=c(paste("NOxemis_MC",(1:NMC), sep=""))

#STATISTICS for CH4emis

#create a dataframe to store the values for mean, sd and cv
dsCH4_stat = data.frame(dsCH4[,1:2], matrix(data=0, 38, 3))

#rename the 3-5 columnns of dsCH4_stat
names(dsCH4_stat)[3:5]= c("mu", "sd", "cv")

#make a loop to calculate the mean, sd and cv for all countries
for (k in 1:38){
      temp = dsCH4[k,(3:1002)]
      mu=mean(as.numeric(temp))
      sd = sd(as.numeric(temp))
      cv =(sd/mu)
      dsCH4_stat[k,3]=mu
      dsCH4_stat[k,4]=sd
      dsCH4_stat[k,5]=cv
      }

#write the dsCH4_stat dataframe to a csv file
write.csv(dsCH4_stat, file="CH4_inpNCU_Nuts_Cntry_Gen_outCntry.csv")

#STATISTICS for N2Oemis

#create a dataframe to store the values for mean, sd and cv
dsN2O_stat = data.frame(dsN2O[,1:2], matrix(data=0, 38, 3))

#rename the 3-5 columnns of dsN2O_stat
names(dsN2O_stat)[3:5]= c("mu", "sd", "cv")

#make a loop to calculate the mean, sd and cv for all countries
for (k in 1:38){
      temp = dsN2O[k,(3:1002)]
      mu=mean(as.numeric(temp))
      sd = sd(as.numeric(temp))
      cv =(sd/mu)
      dsN2O_stat[k,3]=mu
      dsN2O_stat[k,4]=sd
      dsN2O_stat[k,5]=cv
      }
```

```
#write the dsN2O_stat dataframe to a csv file
write.csv(dsN2O_stat, file="N2O_inpNCU_Nuts_Cntry_Gen_outCntry.csv")

#STATISTICS for NH3emis

#create a dataframe to store the values for mean, sd and cv
dsNH3_stat = data.frame(dsNH3[,1:2], matrix(data=0, 38, 3))

#rename the 3-5 columnns of dsNH3_stat
names(dsNH3_stat)[3:5]= c("mu", "sd", "cv")

#make a loop to calculate the mean, sd and cv for all countries
for (k in 1:38){
      temp = dsNH3[k,(3:1002)]
      mu=mean(as.numeric(temp))
      sd = sd(as.numeric(temp))
      cv =(sd/mu)
      dsNH3_stat[k,3]=mu
      dsNH3_stat[k,4]=sd
      dsNH3_stat[k,5]=cv
      }

#write the dsNH3_stat dataframe to a csv file
write.csv(dsNH3_stat, file="NH3_inpNCU_Nuts_Cntry_Gen_outCntry.csv")

#STATISTICS for NleGW

#create a dataframe to store the values for mean, sd and cv
dsNleGW_stat = data.frame(dsNleGW[,1:2], matrix(data=0, 38, 3))

#rename the 3-5 columnns of dsNleGW_stat
names(dsNleGW_stat)[3:5]= c("mu", "sd", "cv")

#make a loop to calculate the mean, sd and cv for all countries
for (k in 1:38){
      temp = dsNleGW[k,(3:1002)]
      mu=mean(as.numeric(temp))
      sd = sd(as.numeric(temp))
      cv =(sd/mu)
      dsNleGW_stat[k,3]=mu
      dsNleGW_stat[k,4]=sd
      dsNleGW_stat[k,5]=cv
      }

#write the dsNleGW_stat dataframe to a csv file
write.csv(dsNleGW_stat,
file="NleGW_inpNCU_Nuts_Cntry_Gen_outCntry.csv")

#STATISTICS for NleSW

#create a dataframe to store the values for mean, sd and cv
dsNleSW_stat = data.frame(dsNleSW[,1:2], matrix(data=0, 38, 3))

#rename the 3-5 columnns of dsNleSW_stat
names(dsNleSW_stat)[3:5]= c("mu", "sd", "cv")

#make a loop to calculate the mean, sd and cv for all countries
for (k in 1:38){
      temp = dsNleSW[k,(3:1002)]
      mu=mean(as.numeric(temp))
```

```
      sd = sd(as.numeric(temp))
      cv =(sd/mu)
      dsNleSW_stat[k,3]=mu
      dsNleSW_stat[k,4]=sd
      dsNleSW_stat[k,5]=cv
      }

#write the dsNleSW_stat dataframe to a csv file
write.csv(dsNleSW_stat,
file="NleSW_inpNCU_Nuts_Cntry_Gen_outCntry.csv")

#STATISTICS for NOxemis

#create a dataframe to store the values for mean, sd and cv
dsNOx_stat = data.frame(dsNOx[,1:2], matrix(data=0, 38, 3))

#rename the 3-5 columnns of dsNOx_stat
names(dsNOx_stat)[3:5]= c("mu", "sd", "cv")

#rename the 3-5 columnns of dsNOx_stat
for (k in 1:38){
      temp = dsNOx[k,(3:1002)]
      mu=mean(as.numeric(temp))
      sd = sd(as.numeric(temp))
      cv =(sd/mu)
      dsNOx_stat[k,3]=mu
      dsNOx_stat[k,4]=sd
      dsNOx_stat[k,5]=cv
      }

#write the dsNOx_stat dataframe to a csv file
write.csv(dsNOx_stat, file="NOx_inpNCU_Nuts_Cntry_Gen_outCntry.csv")



#END OF SCRIPT
```

Appendix 8. R-script for calculating the box-plots (and the 5th , 25th , 50th, 75th , and 95th percentiles) of the model outputs at European level for the three robustness scenarios

```
#SCRIPT FOR BOXPLOTS AND PERCENTILES: of NH3, NOx, NleGW, NleSW,
N2O, #CH4 on singles windows for the 3 Robustness Scenarios
#when inputs = NCU+NUTS+CNTRY+GEN APs and
#outputs = aggregated to Europe Level

#Clean up memory
rm(list = ls())

#Create 6 dataframes for NH3, NOx, NleGW, NleSW, N2O, CH4 to store
#the 1000 Monte Carlo(MC) outputs
#for Optimistic, Reference, and Pessimistic Scenario

#First create 6 matrices with 1000 rows and 3 columns each
ds_NH3 = matrix(data=0,1000,3)
ds_NOx = matrix(data=0,1000,3)
ds_NleGW = matrix(data=0,1000,3)
ds_NleSW = matrix(data=0,1000,3)
ds_CH4 = matrix(data=0,1000,3)
ds_N2O = matrix(data=0,1000,3)

#Then create 6 dataframes from the previous matrices
ds_NH3 = data.frame(ds_NH3)
ds_NOx = data.frame(ds_NOx)
ds_NleGW = data.frame(ds_NleGW)
ds_NleSW = data.frame(ds_NleSW)
ds_CH4 = data.frame(ds_CH4)
ds_N2O = data.frame(ds_N2O)

#Rename the names of the 6 dataframes to the names of the 3
#Scenarios
names(ds_NH3)= c("Opt", "Ref", "Pes")
names(ds_NOx)= c("Opt", "Ref", "Pes")
names(ds_NleGW)= c("Opt", "Ref", "Pes")
names(ds_NleSW)= c("Opt", "Ref", "Pes")
names(ds_CH4)= c("Opt", "Ref", "Pes")
names(ds_N2O)= c("Opt", "Ref", "Pes")


#Set the paths where the outputs for the 3 scenarios are stored
thePath_Opt =
"E:\\Vicky_Ioannidi\\Optimistic_Scenario\\INTEGRATOR_Opt\\outputs\\i
npNCU_Nuts_Cntry_Gen\\Europe_level\\"
thePath_Ref =
"E:\\Vicky_Ioannidi\\Reference_Scenario\\INTEGRATOR_Ref\\outputs\\in
pNCU_Nuts_Cntry_Gen\\Europe_level\\"
thePath_Pes =
"E:\\Vicky_Ioannidi\\Pessimistic_Scenario\\INTEGRATOR_Pes\\outputs\\
inpNCU_Nuts_Cntry_Gen\\Europe_level\\"


#Make a loop to read the 1000 MC output files for NH3, NOx, NleGW,
#NleSW, N2O, CH4 for Opt, Ref, and Pes scenarios
#and store them in the corresponding dataframes
NMC =1000
for (i in 1:NMC)
{
```

```
   NH3emis_Opt = read.csv(file=sprintf(paste(thePath_Opt,
"NH3emis\\NH3emis%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_NH3[i,1] = NH3emis_Opt$val
   NH3emis_Ref = read.csv(file=sprintf(paste(thePath_Ref,
"NH3emis\\NH3emis%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_NH3[i,2] = NH3emis_Ref$val
   NH3emis_Pes = read.csv(file=sprintf(paste(thePath_Pes,
"NH3emis\\NH3emis%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_NH3[i,3] = NH3emis_Pes$val

   NOxemis_Opt = read.csv(file=sprintf(paste(thePath_Opt,
"NOxemis\\NOxemis%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_NOx[i,1] = NOxemis_Opt$val
   NOxemis_Ref = read.csv(file=sprintf(paste(thePath_Ref,
"NOxemis\\NOxemis%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_NOx[i,2] = NOxemis_Ref$val
   NOxemis_Pes = read.csv(file=sprintf(paste(thePath_Pes,
"NOxemis\\NOxemis%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_NOx[i,3] = NOxemis_Pes$val

   NleGW_Opt = read.csv(file=sprintf(paste(thePath_Opt,
"NleGW\\NleGW%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_NleGW[i,1] = NleGW_Opt$val
   NleGW_Ref = read.csv(file=sprintf(paste(thePath_Ref,
"NleGW\\NleGW%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_NleGW[i,2] = NleGW_Ref$val
   NleGW_Pes = read.csv(file=sprintf(paste(thePath_Pes,
"NleGW\\NleGW%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_NleGW[i,3] = NleGW_Pes$val

   NleSW_Opt = read.csv(file=sprintf(paste(thePath_Opt,
"NleSW\\NleSW%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_NleSW[i,1] = NleSW_Opt$val
   NleSW_Ref = read.csv(file=sprintf(paste(thePath_Ref,
"NleSW\\NleSW%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_NleSW[i,2] = NleSW_Ref$val
   NleSW_Pes = read.csv(file=sprintf(paste(thePath_Pes,
"NleSW\\NleSW%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_NleSW[i,3] = NleSW_Pes$val

   CH4emis_Opt = read.csv(file=sprintf(paste(thePath_Opt,
"CH4emis\\CH4emis%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_CH4[i,1] = CH4emis_Opt$val*25
   CH4emis_Ref = read.csv(file=sprintf(paste(thePath_Ref,
"CH4emis\\CH4emis%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_CH4[i,2] = CH4emis_Ref$val*25
```

```
   CH4emis_Pes = read.csv(file=sprintf(paste(thePath_Pes,
"CH4emis\\CH4emis%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_CH4[i,3] = CH4emis_Pes$val*25

   N2Oemis_Opt = read.csv(file=sprintf(paste(thePath_Opt,
"N2Oemis\\N2Oemis%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_N2O[i,1] = N2Oemis_Opt$val*468.286
   N2Oemis_Ref = read.csv(file=sprintf(paste(thePath_Ref,
"N2Oemis\\N2Oemis%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_N2O[i,2] = N2Oemis_Ref$val*468.286
   N2Oemis_Pes = read.csv(file=sprintf(paste(thePath_Pes,
"N2Oemis\\N2Oemis%04d.csv", sep=""), i),
   header = TRUE, sep = ",",na.strings = '-99999')
   ds_N2O[i,3] = N2Oemis_Pes$val*468.286
}

#Find the max values of the dataframes to define the ylim for the
#boxplots below
max(ds_NH3)
max(ds_NOx)
max(ds_NleGW)
max(ds_NleSW)
max(ds_CH4)
max(ds_N2O)


#BOXPLOTS of the columns of all dataframes on single windows

#Open a new graphical window for the boxplots
X11()

#Set all axis labels horizontal
par(las=1)

#Draw the boxplots
boxplot(ds_NH3, col="lavender", names=names(ds_NH3),
pars=list(boxwex=.6, staplewex=.4), main=expression(NH[3]),
ylim=c(0,30), vertical = TRUE, ylab=expression("N " * group("(", "kg
" * ha^{-1} * yr^{-1}, ")")))
savePlot(filename="BoxplotsNH3_3Scen_Europe", type="png")

boxplot(ds_NOx, col="lavender", names=names(ds_NOx),
pars=list(boxwex=.6, staplewex=.4), main=expression(NO[x]),
ylim=c(0,20), vertical = TRUE, ylab=expression("N " * group("(", "kg
" * ha^{-1} * yr^{-1}, ")")))
savePlot(filename="BoxplotsNOx_3Scen_Europe", type="png")

boxplot(ds_NleGW, col="lavender", names=names(ds_NleGW),
pars=list(boxwex=.6, staplewex=.4), main=expression(N[leGW]),
ylim=c(0,27), vertical = TRUE, ylab=expression("N " * group("(", "kg
" * ha^{-1} * yr^{-1}, ")")))
savePlot(filename="BoxplotsNleGW_3Scen_Europe", type="png")

boxplot(ds_NleSW, col="lavender", names=names(ds_NleSW),
pars=list(boxwex=.6, staplewex=.4), main=expression(N[leSW]),
ylim=c(0,72), vertical = TRUE, ylab=expression("N " * group("(", "kg
" * ha^{-1} * yr^{-1}, ")")))
savePlot(filename="BoxplotsNleSW_3Scen_Europe", type="png")
```

```
boxplot(ds_CH4, col="lavender", names=names(ds_CH4),
pars=list(boxwex=.6, staplewex=.4), main=expression(CH[4]),
ylim=c(0,1700), vertical = TRUE, ylab=expression(GWP (CO[2]) *
group("(", "kg " * ha^{-1} * yr^{-1}, ")")))
savePlot(filename="BoxplotsCH4_3Scen_Europe", type="png")

boxplot(ds_N2O, col="lavender", names=names(ds_N2O),
pars=list(boxwex=.6, staplewex=.4), main=expression(N[2] * "O"), for
Europe, 3 scenarios"), ylim=c(0,8600), vertical = TRUE,
ylab=expression(GWP (CO[2]) * group("(", "kg " * ha^{-1} * yr^{-1},
")")))
savePlot(filename="BoxplotsN2O_3Scen_Europe", type="png")

# 5, 25, 50, 75 and 95 PERCENTILES for NH3, NOx, NleGW, NleSW, N2O,
#CH4 for the 3 Scenarios

percent = read.csv(file = 'Percentiles.csv', header = TRUE, sep =
",", na.strings = '#N/A')
p = c(0.05, 0.25, 0.50, 0.75, 0.95)

#PERCENTILES for NH3

ds_NH3per_3ScenEU = data.frame(percent, matrix(data=0, 5, 3))

names(ds_NH3per_3ScenEU) = c("NH3_3ScenEUpercentiles", "Opt", "Ref",
"Pes")

quant=5
for (i in 1:quant){
            ds_NH3per_3ScenEU[i,2] =
quantile(as.numeric(ds_NH3$Opt),probs=c(p[i]))
            ds_NH3per_3ScenEU[i,3] =
quantile(as.numeric(ds_NH3$Ref),probs=c(p[i]))
            ds_NH3per_3ScenEU[i,4] =
quantile(as.numeric(ds_NH3$Pes),probs=c(p[i]))
}

# write the ds_NH3per_5EU to a csv file
write.csv(ds_NH3per_3ScenEU, file = "NH3_3ScenEUpercentiles.csv")


#PERCENTILES for NOx

ds_NOxper_3ScenEU = data.frame(percent, matrix(data=0, 5, 3))

names(ds_NOxper_3ScenEU) = c("NOx_3ScenEUpercentiles", "Opt", "Ref",
"Pes")

quant=5
for (i in 1:quant){
            ds_NOxper_3ScenEU[i,2] =
quantile(as.numeric(ds_NOx$Opt),probs=c(p[i]))
            ds_NOxper_3ScenEU[i,3] =
quantile(as.numeric(ds_NOx$Ref),probs=c(p[i]))
            ds_NOxper_3ScenEU[i,4] =
quantile(as.numeric(ds_NOx$Pes),probs=c(p[i]))
}

# write the ds_NOxper_5EU to a csv file
write.csv(ds_NOxper_3ScenEU, file = "NOx_3ScenEUpercentiles.csv")
```

```
#PERCENTILES for NleSW

ds_NleSWper_3ScenEU = data.frame(percent, matrix(data=0, 5, 3))

names(ds_NleSWper_3ScenEU) = c("NleSW_3ScenEUpercentiles", "Opt",
"Ref", "Pes")

quant=5
for (i in 1:quant){
            ds_NleSWper_3ScenEU[i,2] =
quantile(as.numeric(ds_NleSW$Opt),probs=c(p[i]))
            ds_NleSWper_3ScenEU[i,3] =
quantile(as.numeric(ds_NleSW$Ref),probs=c(p[i]))
            ds_NleSWper_3ScenEU[i,4] =
quantile(as.numeric(ds_NleSW$Pes),probs=c(p[i]))
}

# write the ds_NleSWper_5EU to a csv file
write.csv(ds_NleSWper_3ScenEU, file ="NleSW_3ScenEUpercentiles.csv")


#PERCENTILES for NleGW

ds_NleGWper_3ScenEU = data.frame(percent, matrix(data=0, 5, 3))

names(ds_NleGWper_3ScenEU) = c("NleGW_3ScenEUpercentiles", "Opt",
"Ref", "Pes")

quant=5
for (i in 1:quant){
            ds_NleGWper_3ScenEU[i,2] =
quantile(as.numeric(ds_NleGW$Opt),probs=c(p[i]))
            ds_NleGWper_3ScenEU[i,3] =
quantile(as.numeric(ds_NleGW$Ref),probs=c(p[i]))
            ds_NleGWper_3ScenEU[i,4] =
quantile(as.numeric(ds_NleGW$Pes),probs=c(p[i]))
}

# write the ds_NleGWper_5EU to a csv file
write.csv(ds_NleGWper_3ScenEU, file ="NleGW_3ScenEUpercentiles.csv")


#PERCENTILES for CH4

ds_CH4per_3ScenEU = data.frame(percent, matrix(data=0, 5, 3))

names(ds_CH4per_3ScenEU) = c("CH4_3ScenEUpercentiles", "Opt", "Ref",
"Pes")

quant=5
for (i in 1:quant){
            ds_CH4per_3ScenEU[i,2] =
quantile(as.numeric(ds_CH4$Opt),probs=c(p[i]))
            ds_CH4per_3ScenEU[i,3] =
quantile(as.numeric(ds_CH4$Ref),probs=c(p[i]))
            ds_CH4per_3ScenEU[i,4] =
quantile(as.numeric(ds_CH4$Pes),probs=c(p[i]))
}

# write the ds_CH4per_5EU to a csv file
write.csv(ds_CH4per_3ScenEU, file = "CH4_3ScenEUpercentiles.csv")
```

```
#PERCENTILES for N2O

ds_N2Oper_3ScenEU = data.frame(percent, matrix(data=0, 5, 3))

names(ds_N2Oper_3ScenEU) = c("N2O_3ScenEUpercentiles", "Opt", "Ref",
"Pes")

quant=5
for (i in 1:quant){
            ds_N2Oper_3ScenEU[i,2] =
quantile(as.numeric(ds_N2O$Opt),probs=c(p[i]))
            ds_N2Oper_3ScenEU[i,3] =
quantile(as.numeric(ds_N2O$Ref),probs=c(p[i]))
            ds_N2Oper_3ScenEU[i,4] =
quantile(as.numeric(ds_N2O$Pes),probs=c(p[i]))
}

# write the ds_N2Oper_5EU to a csv file
write.csv(ds_N2Oper_3ScenEU, file = "N2O_3ScenEUpercentiles.csv")


#END OF SCRIPT
```

Appendix 9. R-script for the stack-columns of the uncertainty contribution of each group of APs to the total output uncertainty, for five EU countries (i.e. NL, GR, IE, FR, PL), for the Reference scenario

```
#SCRIPT FOR THE STACK-COLUMNS: of the Uncertainty Analysis of all
#outputs, for NL, GR, IE, FR, PL, for the Reference Scenario

#Clean up memory
rm(list = ls())

#set the path for COMB variances
thePath  =
"E:\\Vicky_Ioannidi\\Reference_Scenario\\Variances_outCntry\\Statist
ics_Ref\\Stat_inpNCU_NUTS_CNTRY_GEN_Ref\\output_statistics\\inpNCU_N
uts_Cntry_Gen_outCntry\\"

NH3_COMB = read.csv(file =(paste(thePath,
"NH3_inpNCU_Nuts_Cntry_Gen_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

NOx_COMB = read.csv(file =(paste(thePath,
"NOx_inpNCU_Nuts_Cntry_Gen_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

NleGW_COMB = read.csv(file =(paste(thePath,
"NleGW_inpNCU_Nuts_Cntry_Gen_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

NleSW_COMB = read.csv(file =(paste(thePath,
"NleSW_inpNCU_Nuts_Cntry_Gen_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

N2O_COMB = read.csv(file =(paste(thePath,
"N2O_inpNCU_Nuts_Cntry_Gen_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

CH4_COMB = read.csv(file =(paste(thePath,
"CH4_inpNCU_Nuts_Cntry_Gen_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

#set the path for NCU variances
thePath  =
"E:\\Vicky_Ioannidi\\Reference_Scenario\\Variances_outCntry\\Statist
ics_Ref\\Stat_inpNCU_Ref\\output_statistics\\inpNCU_outCntry\\"

NH3_NCU = read.csv(file =(paste(thePath, "NH3_inpNCU_outCntry.csv",
sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

NOx_NCU = read.csv(file =(paste(thePath, "NOx_inpNCU_outCntry.csv",
sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

NleGW_NCU = read.csv(file =(paste(thePath,
"NleGW_inpNCU_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

NleSW_NCU = read.csv(file =(paste(thePath,
"NleSW_inpNCU_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')
```

```
N2O_NCU = read.csv(file =(paste(thePath, "N2O_inpNCU_outCntry.csv",
sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

CH4_NCU = read.csv(file =(paste(thePath, "CH4_inpNCU_outCntry.csv",
sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

#set the path for NUTS variances
thePath  =
"E:\\Vicky_Ioannidi\\Reference_Scenario\\Variances_outCntry\\Statist
ics_Ref\\Stat_inpNUTS_Ref\\output_statistics\\inpNuts_outCntry\\"

NH3_NUTS = read.csv(file =(paste(thePath,
"NH3_inpNuts_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

NOx_NUTS = read.csv(file =(paste(thePath,
"NOx_inpNuts_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

NleGW_NUTS = read.csv(file =(paste(thePath,
"NleGW_inpNuts_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

NleSW_NUTS = read.csv(file =(paste(thePath,
"NleSW_inpNuts_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

N2O_NUTS = read.csv(file =(paste(thePath,
"N2O_inpNuts_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

CH4_NUTS = read.csv(file =(paste(thePath,
"CH4_inpNuts_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

#set the path for CNTRY variances
thePath  =
"E:\\Vicky_Ioannidi\\Reference_Scenario\\Variances_outCntry\\Statist
ics_Ref\\Stat_inpCNTRY_Ref\\output_statistics\\inpCntry_outCntry\\"

NH3_CNTRY = read.csv(file =(paste(thePath,
"NH3_inpCntry_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

NOx_CNTRY = read.csv(file =(paste(thePath,
"NOx_inpCntry_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

NleGW_CNTRY = read.csv(file =(paste(thePath,
"NleGW_inpCntry_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

NleSW_CNTRY = read.csv(file =(paste(thePath,
"NleSW_inpCntry_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

N2O_CNTRY = read.csv(file =(paste(thePath,
"N2O_inpCntry_outCntry.csv", sep="")),
```

```
header = TRUE, sep = ",", na.strings = 'NA')

CH4_CNTRY = read.csv(file =(paste(thePath,
"CH4_inpCntry_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

#set the path for GEN variances
thePath  =
"E:\\Vicky_Ioannidi\\Reference_Scenario\\Variances_outCntry\\Statist
ics_Ref\\Stat_inpGEN_Ref\\output_statistics\\inpGen_outCntry\\"

NH3_GEN = read.csv(file =(paste(thePath, "NH3_inpGen_outCntry.csv",
sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

NOx_GEN = read.csv(file =(paste(thePath, "NOx_inpGen_outCntry.csv",
sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

NleGW_GEN = read.csv(file =(paste(thePath,
"NleGW_inpGen_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

NleSW_GEN = read.csv(file =(paste(thePath,
"NleSW_inpGen_outCntry.csv", sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

N2O_GEN = read.csv(file =(paste(thePath, "N2O_inpGen_outCntry.csv",
sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

CH4_GEN = read.csv(file =(paste(thePath, "CH4_inpGen_outCntry.csv",
sep="")),
header = TRUE, sep = ",", na.strings = 'NA')

#Create the dataframes to store the variances from the 5 input
#groups for all outputs
NH3_Var= data.frame(NH3_COMB$Var, NH3_NCU$Var, NH3_NUTS$Var,
NH3_CNTRY$Var, NH3_GEN$Var)
NOx_Var= data.frame(NOx_COMB$Var, NOx_NCU$Var, NOx_NUTS$Var,
NOx_CNTRY$Var, NOx_GEN$Var)
NleGW_Var= data.frame(NleGW_COMB$Var, NleGW_NCU$Var, NleGW_NUTS$Var,
NleGW_CNTRY$Var, NleGW_GEN$Var)
NleSW_Var= data.frame(NleSW_COMB$Var, NleSW_NCU$Var, NleSW_NUTS$Var,
NleSW_CNTRY$Var, NleSW_GEN$Var)
N2O_Var= data.frame(N2O_COMB$Var, N2O_NCU$Var, N2O_NUTS$Var,
N2O_CNTRY$Var, N2O_GEN$Var)
CH4_Var= data.frame(CH4_COMB$Var, CH4_NCU$Var, CH4_NUTS$Var,
CH4_CNTRY$Var, CH4_GEN$Var)

#Rename the columns of the dataframes of the variances
names(NH3_Var)= c("COMB", "NCU", "NUTS", "CNTRY", "GEN")
names(NOx_Var)= c("COMB", "NCU", "NUTS", "CNTRY", "GEN")
names(NleGW_Var)= c("COMB", "NCU", "NUTS", "CNTRY", "GEN")
names(NleSW_Var)= c("COMB", "NCU", "NUTS", "CNTRY", "GEN")
names(N2O_Var)= c("COMB", "NCU", "NUTS", "CNTRY", "GEN")
names(CH4_Var)= c("COMB", "NCU", "NUTS", "CNTRY", "GEN")


#Create 6 matrices for the relative variances of all outputs with 5
#rows(5 countries) and 5 columns (5 different input groups)
```

```
NH3 = matrix(data=0,5,5)
NOx = matrix(data=0,5,5)
NleGW = matrix(data=0,5,5)
NleSW = matrix(data=0,5,5)
N2O = matrix(data=0,5,5)
CH4 = matrix(data=0,5,5)

#Then create 6 dataframes from the previous matrices
NH3_Rel_Var = data.frame(NH3)
NOx_Rel_Var = data.frame(NOx)
NleGW_Rel_Var = data.frame(NleGW)
NleSW_Rel_Var = data.frame(NleSW)
N2O_Rel_Var = data.frame(N2O)
CH4_Rel_Var = data.frame(CH4)

#Rename the columns of the dataframes of the relative variances
names(NH3_Rel_Var)= c("COMB", "NCU", "NUTS", "CNTRY", "GEN")
names(NOx_Rel_Var)= c("COMB", "NCU", "NUTS", "CNTRY", "GEN")
names(NleGW_Rel_Var)= c("COMB", "NCU", "NUTS", "CNTRY", "GEN")
names(NleSW_Rel_Var)= c("COMB", "NCU", "NUTS", "CNTRY", "GEN")
names(N2O_Rel_Var )= c("COMB", "NCU", "NUTS", "CNTRY", "GEN")
names(CH4_Rel_Var)= c("COMB", "NCU", "NUTS", "CNTRY", "GEN")


#Create a vector with the COUNTRY_NR of the 5 countries
m = c(29, 16, 18, 14, 31)

#Make a "for" loop to calculate the relative variances of all
#outputs for the 5 countries (NL, GR, IE, FR, PL)

for (i in 1:5){
  NH3_Rel_Var[i,1] = NH3_Var$COMB[m[i]]/NH3_Var$COMB[m[i]]
  NH3_Rel_Var[i,2] = NH3_Var$NCU[m[i]]/NH3_Var$COMB[m[i]]
  NH3_Rel_Var[i,3] = NH3_Var$NUTS[m[i]]/NH3_Var$COMB[m[i]]
  NH3_Rel_Var[i,4] = NH3_Var$CNTRY[m[i]]/NH3_Var$COMB[m[i]]
  NH3_Rel_Var[i,5] = NH3_Var$GEN[m[i]]/NH3_Var$COMB[m[i]]

  NOx_Rel_Var[i,1] = NOx_Var$COMB[m[i]]/NOx_Var$COMB[m[i]]
  NOx_Rel_Var[i,2] = NOx_Var$NCU[m[i]]/NOx_Var$COMB[m[i]]
  NOx_Rel_Var[i,3] = NOx_Var$NUTS[m[i]]/NOx_Var$COMB[m[i]]
  NOx_Rel_Var[i,4] = NOx_Var$CNTRY[m[i]]/NOx_Var$COMB[m[i]]
  NOx_Rel_Var[i,5] = NOx_Var$GEN[m[i]]/NOx_Var$COMB[m[i]]

  NleGW_Rel_Var[i,1] = NleGW_Var$COMB[m[i]]/NleGW_Var$COMB[m[i]]
  NleGW_Rel_Var[i,2] = NleGW_Var$NCU[m[i]]/NleGW_Var$COMB[m[i]]
  NleGW_Rel_Var[i,3] = NleGW_Var$NUTS[m[i]]/NleGW_Var$COMB[m[i]]
  NleGW_Rel_Var[i,4] = NleGW_Var$CNTRY[m[i]]/NleGW_Var$COMB[m[i]]
  NleGW_Rel_Var[i,5] = NleGW_Var$GEN[m[i]]/NleGW_Var$COMB[m[i]]

  NleSW_Rel_Var[i,1] = NleSW_Var$COMB[m[i]]/NleSW_Var$COMB[m[i]]
  NleSW_Rel_Var[i,2] = NleSW_Var$NCU[m[i]]/NleSW_Var$COMB[m[i]]
  NleSW_Rel_Var[i,3] = NleSW_Var$NUTS[m[i]]/NleSW_Var$COMB[m[i]]
  NleSW_Rel_Var[i,4] = NleSW_Var$CNTRY[m[i]]/NleSW_Var$COMB[m[i]]
  NleSW_Rel_Var[i,5] = NleSW_Var$GEN[m[i]]/NleSW_Var$COMB[m[i]]

  N2O_Rel_Var[i,1] = N2O_Var$COMB[m[i]]/N2O_Var$COMB[m[i]]
  N2O_Rel_Var[i,2] = N2O_Var$NCU[m[i]]/N2O_Var$COMB[m[i]]
  N2O_Rel_Var[i,3] = N2O_Var$NUTS[m[i]]/N2O_Var$COMB[m[i]]
  N2O_Rel_Var[i,4] = N2O_Var$CNTRY[m[i]]/N2O_Var$COMB[m[i]]
  N2O_Rel_Var[i,5] = N2O_Var$GEN[m[i]]/N2O_Var$COMB[m[i]]
```

```
  CH4_Rel_Var[i,1] = CH4_Var$COMB[m[i]]/CH4_Var$COMB[m[i]]
  CH4_Rel_Var[i,2] = CH4_Var$NCU[m[i]]/CH4_Var$COMB[m[i]]
  CH4_Rel_Var[i,3] = CH4_Var$NUTS[m[i]]/CH4_Var$COMB[m[i]]
  CH4_Rel_Var[i,4] = CH4_Var$CNTRY[m[i]]/CH4_Var$COMB[m[i]]
  CH4_Rel_Var[i,5] = CH4_Var$GEN[m[i]]/CH4_Var$COMB[m[i]]
}

#Write the 6 dataframes to csv files
write.csv(NH3_Rel_Var, file = "NH3_Rel_Var.csv")
write.csv(NOx_Rel_Var, file = "NOx_Rel_Var.csv")
write.csv(NleGW_Rel_Var, file = "NleGW_Rel_Var.csv")
write.csv(NleSW_Rel_Var, file = "NleSW_Rel_Var.csv")
write.csv(N2O_Rel_Var, file = "N2O_Rel_Var.csv")
write.csv(CH4_Rel_Var, file = "CH4_Rel_Var.csv")


#Read the relative variances of all outputs
relat_NH3 =
read.csv("E:/Vicky_Ioannidi/Reference_Scenario/Variances_outCntry/St
ack_Bars/NH3_Rel_Var.csv", header=T, sep=",")
relat_NOx =
read.csv("E:/Vicky_Ioannidi/Reference_Scenario/Variances_outCntry/St
ack_Bars/NOx_Rel_Var.csv", header=T, sep=",")
relat_NleGW =
read.csv("E:/Vicky_Ioannidi/Reference_Scenario/Variances_outCntry/St
ack_Bars/NleGW_Rel_Var.csv", header=T, sep=",")
relat_NleSW =
read.csv("E:/Vicky_Ioannidi/Reference_Scenario/Variances_outCntry/St
ack_Bars/NleSW_Rel_Var.csv", header=T, sep=",")
relat_N2O =
read.csv("E:/Vicky_Ioannidi/Reference_Scenario/Variances_outCntry/St
ack_Bars/N2O_Rel_Var.csv", header=T, sep=",")
relat_CH4 =
read.csv("E:/Vicky_Ioannidi/Reference_Scenario/Variances_outCntry/St
ack_Bars/CH4_Rel_Var.csv", header=T, sep=",")

#Expand right side of clipping rect to make room for the legend
par(xpd=T, mar=par()$mar+c(0,0,0,8))

#Create the stack-columns for all outputs
barplot(t(relat_NH3[,2:5]), main=expression(NH[3]), ylab="Total",
col=topo.colors(4), space=0.1, cex.axis=0.8, las=1,
names.arg=t(relat_NH3[,1]), cex=0.8, ylim = c(0,1.1))
#Place the legend at(6, 1) using topo colors
legend(6, 1, names(relat_NH3[,2:5]), cex=0.8, fill=topo.colors(4))
# Save plot
savePlot(filename="Stackcol_NH3", type="png")

barplot(t(relat_NOx[,2:5]), main=expression(NO[x]), ylab="Total",
col=topo.colors(4), space=0.1, cex.axis=0.8, las=1,
names.arg=t(relat_NOx[,1]), cex=0.8, ylim = c(0,1.1))
#Place the legend at(6, 1) using topo colors
legend(6, 1, names(relat_NOx[,2:5]), cex=0.8, fill=topo.colors(4))
# Save plot
savePlot(filename="Stackcol_NOx", type="png")

barplot(t(relat_NleGW[,2:5]), main=expression(N[leGW]),
ylab="Total", col=topo.colors(4), space=0.1, cex.axis=0.8, las=1,
names.arg=t(relat_NleGW[,1]), cex=0.8, ylim = c(0,1.1))
#Place the legend at (6, 1) using topo colors
legend(6, 1, names(relat_NleGW[,2:5]), cex=0.8, fill=topo.colors(4))
```

```
# Save plot
savePlot(filename="Stackcol_NleGW", type="png")

barplot(t(relat_NleSW[,2:5]), main=expression(N[leSW]),
ylab="Total", col=topo.colors(4), space=0.1, cex.axis=0.8, las=1,
names.arg=t(relat_NleSW[,1]), cex=0.8, ylim = c(0,1.1))
#Place the legend at (6, 1) using topo colors
legend(6, 1, names(relat_NleSW[,2:5]), cex=0.8, fill=topo.colors(4))
# Save plot
savePlot(filename="Stackcol_NleSW", type="png")

barplot(t(relat_N2O[,2:5]), main=expression(N[2]*O), ylab="Total",
col=topo.colors(4), space=0.1, cex.axis=0.8, las=1,
names.arg=t(relat_N2O[,1]), cex=0.8, ylim = c(0,1.1))
#Place the legend at (6, 1) using topo colors
legend(6, 1, names(relat_N2O[,2:5]), cex=0.8, fill=topo.colors(4))
# Save plot
savePlot(filename="Stackcol_N2O", type="png")

barplot(t(relat_CH4[,2:5]), main=expression(CH[4]), ylab="Total",
col=topo.colors(4), space=0.1, cex.axis=0.8, las=1,
names.arg=t(relat_CH4[,1]), cex=0.8, ylim = c(0,1.1))
#Place the legend at (6, 1) using topo colors
legend(6, 1, names(relat_CH4[,2:5]), cex=0.8, fill=topo.colors(4))
# Save plot
savePlot(filename="Stackcol_CH4", type="png")


#END OF SCRIPT
```

Appendix 10. R-script for the generation of: 1. the cdfs of the differences between the preset CVs or SDs and the CVs or SDs resulting from 1000 simulations of each NUTS AP, 2. the pdfs of each NUTS AP, for the Reference scenario

```
#SCRIPT to calculate: mu, sd, and cv resulting from the 1000
#simulations of the 26 NUTS parameters,
#to generate the cumulatives curves of the differences of
#the preset CVs or SDs from the resulted CVs or SDs respectively,
#and also to generate the pdfs of each AP

#clean-up memory
rm(list = ls())

#read the 688 NUTS regions
ds = read.csv(file = '688NUTSids_COUids.csv', header = TRUE, sep =
",", na.strings = '#N/A')

#read 1000 MC input files for the 26 NUTS parameters
for (j in 1:26){
NMC =1000
for (i in 1:NMC)
{
NUTSPAR =
read.csv(file=sprintf("C:\\Robustness_Scenarios_NEW26\\Reference_Sce
nario26\\NUTS_Par_Ref_group26\\1000Simulations_NUTSpar\\nutspar%04d.
csv", i),
header = TRUE, sep = ",", na.strings = '#N/A')
ds = data.frame(ds, NUTSPAR[,j*3+2])
}
}

#create a vector with the names of the parameters
parnames= c("ParNexf_ca", "ParNexf_oc", "ParNexf_pp", "ParNexf_po",
"ParCnam", "Parfhs_ca", "Parfhs_oc", "Parfrlam_ca", "Parfrlam_pp",
"ParfNemhs_NH3", "ParfNemms_NH3", "ParfNemhsl_N2O","ParfNemmsl_N2O",
"ParfNemhsl_NO", "ParfNemmsl_NO","ParfNemhsls_N2O",
"ParfNemmss_N2O","ParfNemhss_NO","ParfNemmss_NO",
"Parwamhsara","Parwamhsgrass", "ParArea_int", "Parflems",
"Para_CH4_ca", "Parb_CH4_ca", "ParCH4_oc")

#rename the columns of the ds according to parnames
for (j in 1:26){
names(ds)[((j-1)*NMC+3):(j*NMC+2)]=
c(paste(parnames[j],"_MC",(1:NMC),sep=""))
}

#build a dataframe to store the above mentioned differences for the
#26 NUTS parameters over the 688 NUTS regions
ds_stat = data.frame(ds[,1:2], matrix(data=0, 688, 26))

#rename the columns of the ds_stat where the differences for each
#parameter will be stored
names(ds_stat)[3:28]= c("ParNexf_ca", "ParNexf_oc", "ParNexf_pp",
"ParNexf_po", "ParCnam", "Parfhs_ca", "Parfhs_oc", "Parfrlam_ca",
"Parfrlam_pp","ParfNemhs_NH3", "ParfNemms_NH3", "ParfNemhsl_N2O",
"ParfNemmsl_N2O", "ParfNemhsl_NO", "ParfNemmsl_NO",
"ParfNemhsls_N2O", "ParfNemmss_N2O","ParfNemhss_NO","ParfNemmss_NO",
"Parwamhsara","Parwamhsgrass", "ParArea_int", "Parflems",
"Para_CH4_ca", "Parb_CH4_ca", "ParCH4_oc")
```

```
#read the info of the NUTS parameters with their set CVs and SDs
param = read.csv(file = 'Parameters_NUTSRef.csv', header = TRUE, sep
= ",")

#make a dataframe to store the mu, sd, cv of the 26 NUTS parameters
#resulted from their samples of 1000 simulations
ds_store = data.frame(param[,1], matrix(data=0, 26, 3))

#rename the names of the ds_store
names(ds_store)= c("Nuts_AP", "mu", "sd", "cv")

#calculate: mu, sd, cv of the 1000 simulations of the 26 NUTS
#parameters and store them in the ds_store,
#and then calculate the differences of the set cv (or sd) from the
#resulted cv (or sd) and store them in the ds_stat
for (k in 1:688){
   for(j in 1:26){
      temp = ds[k,((j-1)*NMC+3):(j*NMC+2)]
      mu=mean(as.numeric(temp))
      sd = sd(as.numeric(temp))
      cv =(sd/mu)
      ds_store[j,2]=mu
      ds_store[j,3]=sd
      ds_store[j,4]=cv
      if(!is.na(param$CV[j]))
         ds_stat[k,j+2]= param$CV[j]- cv
      else
         ds_stat[k,j+2]= param$SD[j]- sd
   }
}

#write the ds_store to a csv file
write.csv(ds_store, file = "MU_SD_CV_NUTS1000sim.csv")


#plot the cdfs of the differences and store them in jpg files
for(j in 1:26){
   plot(ecdf(ds_stat[,j+2]))
   savePlot(filename=sprintf("Cumulative_V2%02d", j), type="jpg")
}

#calculate the 5th, 50th and 95th percentiles of the cdfs for the
#2nd, 7th, and 22nd NUTS parameter
quantile(as.numeric(ds_stat[,2+2]), probs = c(0.05,0.50,0.95), na.rm
= TRUE)
quantile(as.numeric(ds_stat[,7+2]), probs = c(0.05,0.50,0.95), na.rm
= TRUE)
quantile(as.numeric(ds_stat[,22+2]), probs = c(0.05,0.50,0.95),
na.rm = TRUE)

#plot the pdfs of the NUTS APs and store them in jpg files
for (j in 1:26){
plot(density(as.numeric(ds[k,((j-1)*NMC+3):(j*NMC+2)])), type="l",
lty=1, col="red", xlab="x ",
ylab="Density", main=paste("pdf of AP:", parnames[j], sep =" "))
savePlot(filename=paste("pdf_", parnames[j], sep =""), type="jpg")
}

#END OF SCRIPT
```

Appendix 11. Tables with percentiles

Table 14. Percentiles of the outputs, for Europe, for the Reference scenario

EUROPEAN LEVEL OUTPUTS		Percentiles				
		5th	25th	50th	70th	95th
N outputs (kg N ha^{-1} yr^{-1})	NH_3	13	15	16	18	20
	NO_x	2.7	3.5	4.2	4.8	6.0
	N_{leSW}	11	14	18	21	27
	N_{leGW}	4.0	5.6	7.2	9.1	12
GWP outputs (kg CO_2 ha^{-1} yr^{-1})	CH_4	958	1064	1144	1225	1342
	N_2O	1840	2226	2496	2779	3357

Table 15. Percentiles of the outputs, for five EU countries, for the Reference scenario

COUNTRY LEVEL OUTPUTS		5 EU Countries	Percentiles				
			5th	25th	50th	70th	95th
N outputs (kg N ha^{-1} yr^{-1})	NH_3	NL	39	48	56	64	78
		GR	5.2	6.0	6.7	7.5	8.7
		IE	14	17	19	22	26
		FR	13	15	17	19	22
		PL	9.8	12	14	16	20
	NO_x	NL	2.4	3.8	5.2	7.3	12
		GR	1.5	2.2	2.8	3.7	5.6
		IE	2.6	3.3	4.0	4.8	6.6
		FR	2.9	3.7	4.4	5.3	6.7
		PL	2.4	3.6	4.6	5.9	8.6
	N_{leSW}	NL	16	31	44	59	87
		GR	5.7	8.8	12	15	22
		IE	4.7	7.6	10	13	19
		FR	9.3	14	19	24	34
		PL	6.3	10	14	19	28
	N_{leGW}	NL	3.9	11	19	29	47
		GR	3.5	5.3	7.4	10	15
		IE	1.6	3.5	5.3	7.4	11
		FR	2.4	4.1	5.7	7.8	11
		PL	2.2	4.3	6.9	11	17
GWP outputs (kg CO_2 ha^{-1} yr^{-1})	CH_4	NL	4024	4584	4989	5352	5934
		GR	439	520	582	643	741
		IE	1309	1707	1941	2251	2650
		FR	959	1114	1231	1367	1539
		PL	598	716	786	863	976
	N_2O	NL	2317	3107	3850	4932	7036
		GR	901	1231	1539	1962	2729
		IE	2245	2712	3123	3593	4390
		FR	1933	2331	2715	3068	3720
		PL	1630	2148	2628	3137	4172

Table 16. Percentiles of the outputs, for Europe, for the three robustness scenarios

EUROPEAN LEVEL OUTPUTS		Robustness Scenarios	Percentiles				
			5th	25th	50th	70th	95th
N outputs (kg N ha^{-1} yr^{-1})	NH_3	Opt	14	15	15	16	16
		Ref	13	15	16	18	20
		Pes	12	15	17	19	23
	NO_x	Opt	2.2	2.3	2.5	2.6	2.9
		Ref	2.7	3.5	4.2	4.8	6.0
		Pes	2.4	3.7	5.1	6.9	10
	N_{leSW}	Opt	10	11	12	13	14
		Ref	11	14	18	21	27
		Pes	8.4	14	19	26	38
	N_{leGW}	Opt	5.5	6.2	6.7	7.3	8.2
		Ref	4.0	5.6	7.2	9.1	12
		Pes	2.4	4.7	7.1	10	16
GWP outputs (kg CO_2 ha^{-1} yr^{-1})	CH_4	Opt	1093	1130	1155	1184	1217
		Ref	958	1064	1144	1225	1342
		Pes	894	1043	1148	1273	1424
	N_2O	Opt	1698	1814	1904	1989	2117
		Ref	1840	2226	2496	2779	3357
		Pes	1735	2415	2990	3765	5076