Report 550032001/2006

Guidance for uncertainty assessment and communication
Checklist for uncertainty in spatial information and
visualising spatial uncertainty

H. Visser, A.C. Petersen, A.H.W. Beusen, P.S.C. Heuberger,
P.H.M. Janssen

Contact: H. Visser
Information Services and Methodology - IMP
Netherlands Environmental Assessment Agency - MNP
Hans.Visser@mnp.nl

The MNP 'Guidance for uncertainty assessment and communication' series includes:

1. *Mini-checklist & Quickscan questionnaire*, A. C. Petersen, P. H. M. Janssen, J. P. van der Sluijs *et al.*, RIVM/MNP, 2003 (published in Dutch and English)
2. *Quickscan hints & actions List*, P. H. M. Janssen, A. C. Petersen, J. P. van der Sluijs *et al.*, RIVM/MNP, 2003 (published in Dutch and English)
3. *Detailed guidance*, J. P. van der Sluijs, J. S. Risbey *et al.*, Utrecht University, 2003
4. *Tool catalogue for uncertainty assessment*, J. P. van der Sluijs, J. S. Risbey *et al.*, Utrecht University, 2004
5. *Checklist for uncertainty in spatial information & visualising spatial uncertainty*, H. Visser A.C. Petersen, A.H.W. Beusen, P.S.C. Heuberger, P.H.M. Janssen, RIVM/MNP, 2005 (the present document in Dutch)

These publications may be downloaded from *www.mnp.nl/leidraad*

Abstract

Guidance for uncertainty assessment and communication
Checklist for uncertainty in spatial information and visualisation of spatial uncertainty

Developments affecting the environment and nature often include spatial components. In the publications of the Netherlands Environmental Assessment Agency (MNP) this spatial information is presented as maps. To interpret these maps correctly, we need to know how uncertainty might affect the values and patterns presented on the maps. Information on the uncertainty becomes essential if map values approach governmental standards or guidelines. In this report a number of suggestions to researchers have been included about dealing with uncertainty in spatial information. The report supplements the 'Guidance for uncertainty assessment and communication' (http://www.mnp.nl/leidraad).

Suggestions have been condensed into an eleven-point checklist of which the main issues are:
1. options for representing uncertainty in maps
2. static and dynamic presentation methods
3. the role of spatial resolution
4. selection of legend classes
5. the importance of meta data

The checklist is supplemented by examples, ranging from effective to less appropriate methods of presenting uncertainty. The report is not only relevant to researchers at the MNP and other organisations, but because it explains the research methods used by MNP, it will also be useful to decision makers for policy and those preparing policies.

Keywords: maps, meta data, uncertainty, resolution, visualisation

Rapport in het kort

Leidraad voor omgaan met onzekerheden
Checklist onzekerheden in ruimtelijke informatie en visualisaties van ruimtelijke onzekerheid

Ontwikkelingen in milieu en natuur bezitten in veel gevallen een ruimtelijke component. In producten van het Milieu- en Natuurplanbureau (MNP) wordt deze ruimtelijke informatie weergegeven in de vorm van kaarten. Voor een juiste interpretatie van kaarten is het van belang om te weten hoe onzekerheid de weergegeven waarden en patronen kan beïnvloeden. Liggen waarden in een kaart dicht bij wettelijke normen of richtlijnen, dan is onzekerheidsinformatie van essentieel belang. Dit rapport geeft onderzoekers tips hoe om te gaan met onzekerheden in ruimtelijke informatie, en vormt als zodanig een aanvulling op de 'Leidraad voor omgaan met onzekerheden' (http://www.mnp.nl/leidraad).

De tips zijn samengevat in een *checklist,* bestaande uit 11 punten. Aan bod komen o.a.:
1. verschillende manieren om onzekerheden in kaarten te representeren;
2. statische en dynamische presentatievormen;
3. de rol van ruimtelijke resolutie;
4. hoe klassenindelingen in de legenda te kiezen;
5. het belang van metadata.

De *checklist* wordt in het rapport aangevuld met een reeks van praktische voorbeelden van geslaagde (en minder geslaagde) onzekerheidspresentaties. Het rapport is niet alleen van belang voor onderzoekers binnen en buiten het MNP. Voor beleidsmakers en beleidsvoorbereiders geeft het inzicht in de onderzoeksmethoden zoals toegepast bij het MNP.

Trefwoorden: kaarten; metadata; onzekerheid; resolutie; visualisatie

Preface

Preface

Contents

Summary

The 'Guidance for uncertainty assessment and communication' adopted by the Netherlands Environmental Assessment Agency (MNP), helps scientists advising on policies to reflect on uncertainties and how to communicate them (Dutch and English site: http://www.mnp.nl/leidraad). However, this guidance document was seen to require further development with respect to uncertainty in spatial information (i.e. maps), which is now provided by this report. The report is not only of interest to researchers at the MNP and elsewhere, but also helps those developing and deciding on policies to understand the research methods used by the MNP.

How should MNP studies present data at high spatial resolution (grid sizes in the order of tens of metres)? How can uncertainty affect the values shown on a map? And how can uncertainty be represented on a map? What presentation methods are the most appropriate? This report, volume 5 of the 'Guidance for uncertainty assessment' series, aims to answer these and other questions.

The 'Checklist for uncertainty in spatial information' presents a number of suggestions for dealing with uncertainty and maps, and information on how to deal with maps provided to others, and how to present uncertainty in spatial information in MNP publications. The issues discussed include:
- options for representing uncertainty in maps
- static and dynamic presentation methods
- the role of spatial resolution
- selecting legend classes
- the value of meta data and
- communications

The checklist will be discussed in greater detail in the report.

The report also includes a chapter on 'Visualisation of spatial uncertainty' which discusses four methods for presenting uncertainty in spatial information: (i) difference maps, (ii) scenario maps, (iii) ensemble maps and (iv) grid uncertainty maps. The measurement scale (nominal, ordinal ratio or interval scale) is also discussed for each of these methods. A number of visual examples is given. Most of these examples communicate well, others do not.

Unlike the other volumes of the 'Guidance for uncertainty assessment and communication' series, this report should not be considered as a strict protocol. Rather, it aims to provide guidelines to make the reader more aware of the choices to be made.

Once a map is drawn, people tend to accept it as reality

<p align="right">*Bert Friesen*</p>
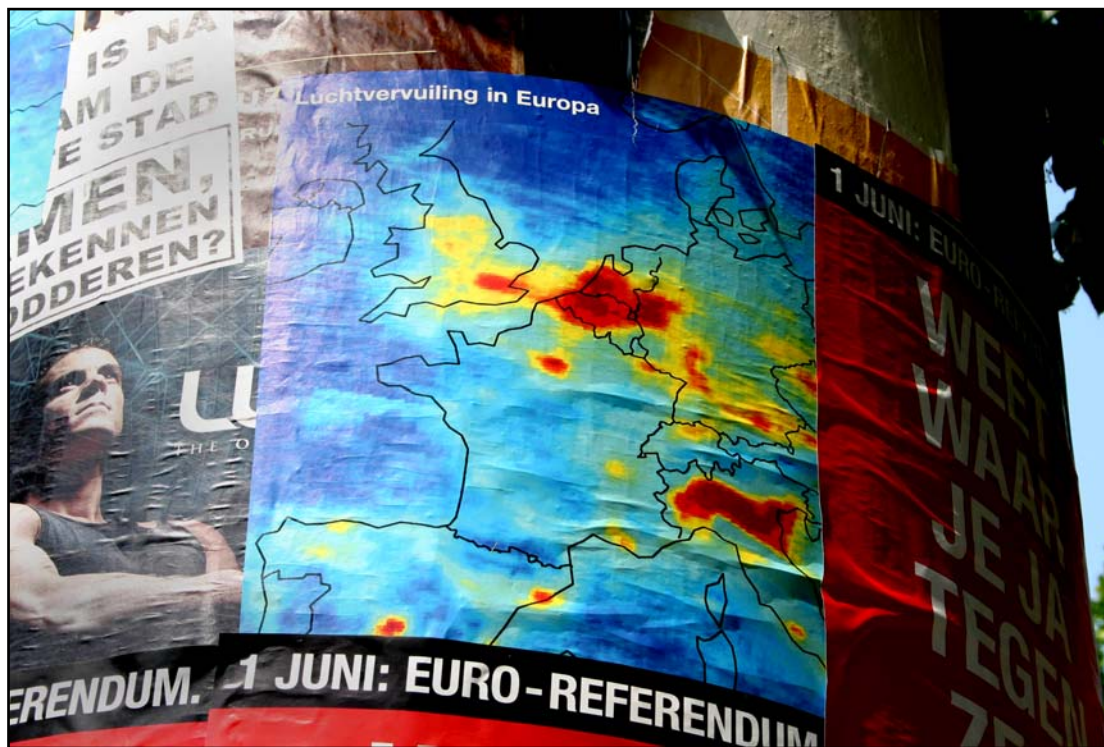
1.   Who, what and why

Background

In 2003, a discussion took place in the context of NO_2 issues about the quality of MNP data at the grid cell level. This concerned the support provided for the implementation of relevant policies. The discussion inspired the MNP to investigate in more general terms how the studies should present data with a high spatial resolution (e.g. grid divisions in the order of tens of metres). This issue is particularly relevant in terms of the presentation of the environmental and health impact. The present report, volume 5 of the 'Guidance for uncertainty assessment and communication' series, is the outcome of a methodological study of dealing with uncertainty in spatial information.

Necessity

The current MNP approach to uncertainty communication is not entirely satisfactory. Numerical data in reports are still not always accompanied by information about the uncertainty. This is an even greater problem for measurement series taken over time. Maps are rarely accompanied by information on uncertainty, as can be seen in examples on pages 108, 115, 147, 149, 151, 166, 179, 181 and 186 of reference [27].

Maps, such as the one in **Figure 1**, look most impressive and on paper appear particularly authoritative, hence the quote at the top of this page. Applied mathematicians would say "Never fall in love with your own data." However beautiful these maps may look, it is still essential to specify the uncertainties!

A psychological explanation for trusting maps "as they are shown" might be that most of us have grown up with beautiful maps in atlases. These maps of regions, provinces, countries, continents and the world as a whole are very accurate. The positional errors are small and normally do not influence the practical usage of the maps. Intuitively we project these early experiences with maps to maps showing concentrations of some pollutant, maps showing effects of climate change or maps showing land-use changes now or in the future (Wim Evers – MNP, personal communication).

Figure 1   *Information in the form of maps is important in policy work, politics and mass*
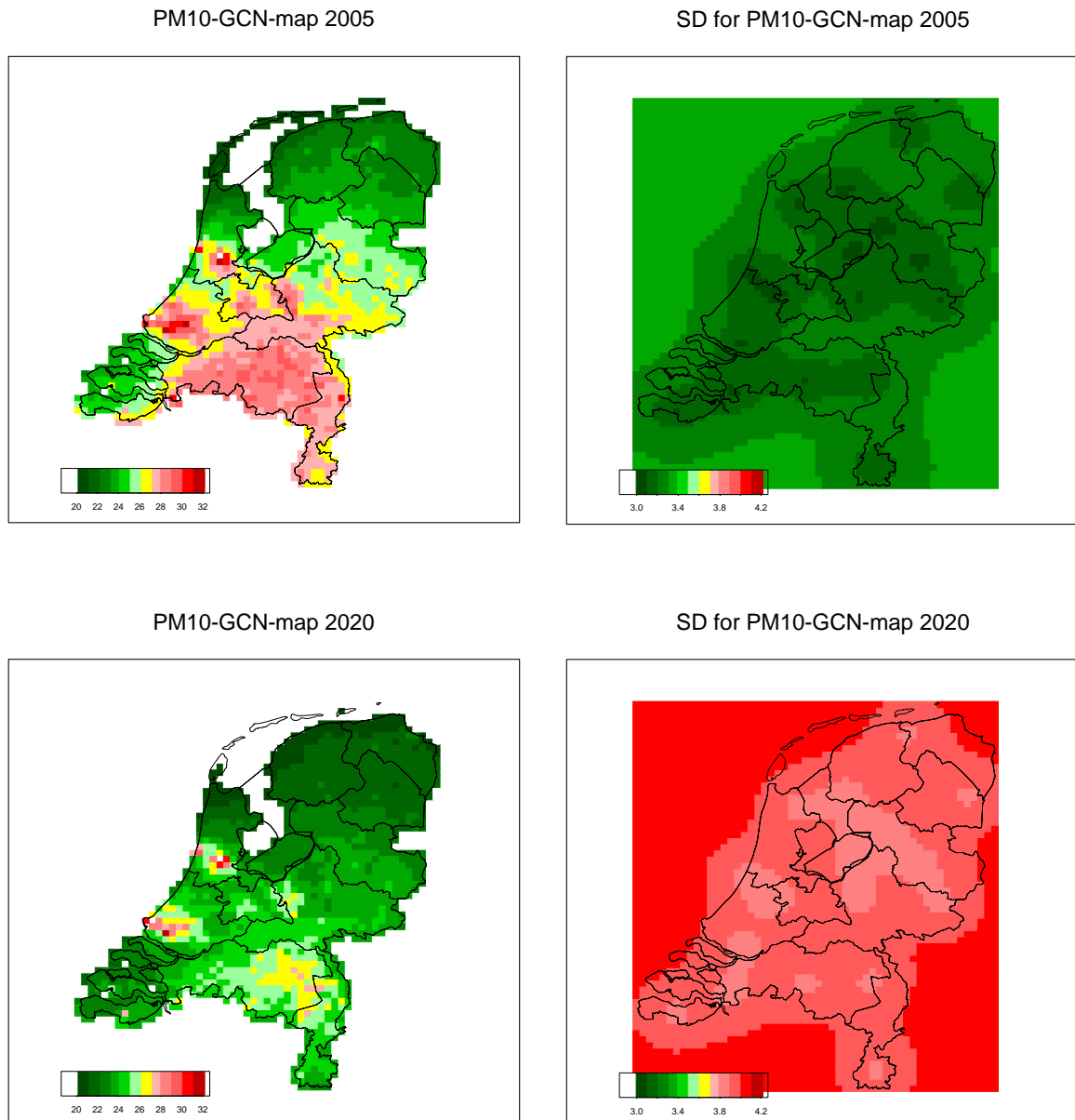   *communications.*
                This map of Europe shows a measure for NO_2 concentrations monitored by the
                SCIAMACHY instrument of the Envisat satellite. The political poster uses this map
                as an argument for voting in favour of the referendum on the European constitution
                (June 1, 2005. The Netherlands voted in majority against the constitution...).
                Figures 9 and 12 provide some interesting comments on this map. Map source:
                UP Heidelberg. Photograph: H. Visser.

Figure 2 presents another example. Shown are particle (PM_{10}) concentrations estimated and
projected respectively for the years 2005 and 2020, and averaged for 5-by-5-km^2 grid boxes.
These maps, denoted by "GCN maps" (Large-scale concentrations maps in the Netherlands),
are generated annually by the MNP for several air-pollution components (NO_2, SO_2, PM_{10}
and O_3) [33].

The PM_{10} maps shown here have strong policy implications because concentrations are just
below the EU standard for daily concentrations. This limit value can be approximated as if it
represented annually averaged PM_{10} concentrations of 30-32 μg/m^3 (seen as red and dark red
in the maps at the right of Figure 2). In the years earlier than 2005 projected concentrations
were found *above or around* the critical range of 30-32 μg/m^3 and a great number of new
building projects were frozen or even cancelled because concentrations were expected to
cross this EU standard. For details please refer to [15, 33].

Figure 2 also shows the uncertainties (right-hand graphs, 1 standard deviation for each 5-by-5-km^2 grid box). Uncertainty bands appear to be very wide due to:

- the uncertainty in PM$_{10}$ measurements (GCN maps for 2005 and 2020)
- the limited number of rural stations (17), leading to interpolation errors (GCN maps for 2005 and 2020)
- the meteorological variability (annual concentrations could be 4 μg/m^3 higher or lower due to particular meteorological conditions) (GCN map for 2020 only)

PM10-GCN-map 2005   SD for PM10-GCN-map 2005

PM10-GCN-map 2020   SD for PM10-GCN-map 2020

Figure 2   *GCN maps for PM$_{10}$ (left) with corresponding uncertainties (right). All concentrations are given in μg/m^3. The colours in the GCN maps have a concentration width of 1.0 μg/m^3, the uncertainty maps of 0.1 μg/m^3. Source: [20].*

It should be noted here that uncertainties are not incorporated into the (juridical) assessment of exceeding the EU PM_{10} standard!

Specifying the uncertainties inherent in maps is not only relevant as the values on the map might well be 25% higher or lower, but also because the *patterns* presented on the map (for example, by isolines) might be quite different in reality. Furthermore, uncertainty also needs to be considered when deciding on the scale and legend.

Difficulties

It is often difficult to determine the uncertainties associated with maps. And it may unexpectedly demand considerable time, for example, when essential information about the uncertainty is not easily available. However, there are several systematic methods that can determine uncertainties.

For instance, if we have measurements from monitoring stations, we can apply geostatistical methods such as Kriging [6, 16, 17] to determine uncertainties and derive ensemble maps. If a map was produced by a model or mathematical operations, we can use Monte Carlo simulation [10, 31] to determine the uncertainties and ensemble maps. However, this requires the uncertainties in the input of the model or processing to be known.
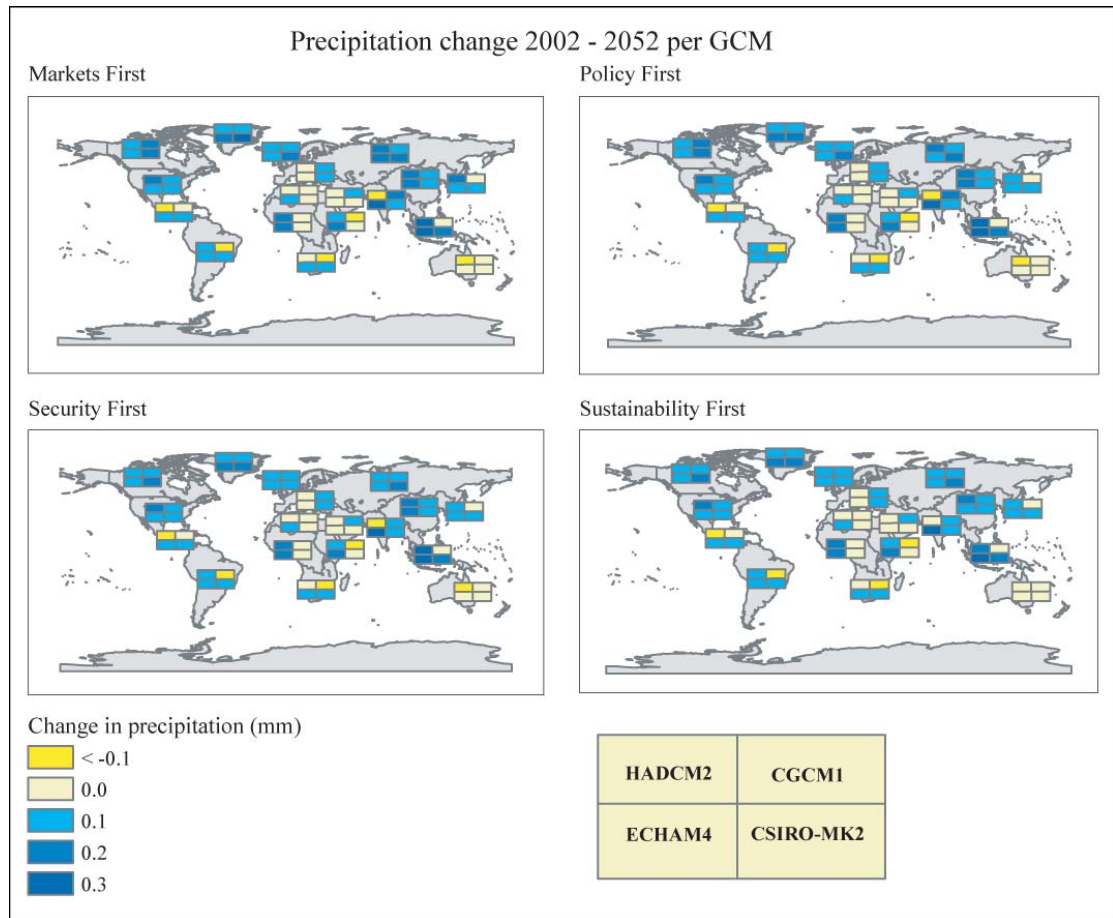
At MNP we often combine maps with different resolutions (grid divisions). This introduces errors which may well be unexpected [2, 11, 12, 13]. Finally, uncertainty may be due to flaws in the model used. To quantify this type of uncertainty we should compare the outcome of this model with the outcome of alternative models. **Figure 3** presents an example of this.

For a recent overview of the many aspects of spatial uncertainty, please refer to [30].

Typology

Each discipline has its own terminology: this applies to GIS as much as it does to statistics and uncertainty analysis. The typology of uncertainty, used at the MNP, is summarised in the RIVM/MNP 'Guidance for uncertainty assessment and communication, Quickscan hints & Actions list' (page 16 ff). The spatial information discipline has also developed its own terminology, such as seen in Burrough's list of various sources of uncertainty [6, Section 9], which is therefore not duplicated here.

We only mention two common terms relating to errors in maps. Frequently, a distinction is made between positional errors (positional uncertainty) and errors in the value of the plotted quantity (attribute uncertainty). Positional uncertainty is less relevant to MNP and RIVM research. One reason for this is that much information about nature and the environment is topographical and therefore includes accurate positions.

Figure 3   *Changes in annual precipitation in the 2002-2050 period.*
          The figure illustrates the scenario uncertainties using four tiles (Markets first, Policy
          first, Security first en Sustainability first), each tile presenting the outcomes of four
          different climate models (HADCM2, CGCM1, ECHAM4 and CSIRO-MK2). In this
          way 16 different changes in precipitation are given for each region. Source: [24,
          Figure 7.12].

A second reason is that there are many techniques for correcting positional errors. Rubber-sheeting methods to correct the locations in satellite observations form a good example. Such a method was used to develop the maps in Figures 1 and 6: the Envisat satellite measured the number of NO_2 particles at an oblique angle, and the software corrected them to ensure that the concentrations shown on the map represent the *vertical* air column.

In so far as positional uncertainty is relevant to the maps discussed here, it is advisable to reduce the impact of this type of uncertainty by analysing the maps using fuzzy methods ([6, Section 11] and [8]).

Checklist

The objective of this checklist is to present a number of leads to assist with map uncertainties, dealing with maps to be provided to third parties and presenting uncertainties in spatial information in MNP publications.

Section 2 includes an 11-point checklist which addresses these issues. The items in the checklist are discussed in more detail in section 3.

Chapter 4 addresses how to represent uncertainty in maps using the following methods:

- difference maps
- scenario maps
- ensemble maps
- grid uncertainty maps

These four basic methods are discussed in section 4.1. Section 4.2 includes examples from work done by the IPCC and MNP, while subsection 4.2.1 covers maps with ratio or interval scales (maps illustrating continuous variables), and subsection 4.2.2 covers nominal scale maps (maps illustrating categories, such as land use). Subsection 4.2.3 briefly discusses hybrid map presentations while subsection 4.2.4 gives examples of maps integrating mean values with information on uncertainty.

Just as other volumes of the 'Guidance for uncertainty assessment and communication', the 'Checklist for uncertainty in spatial information' should not be considered strict protocol. Rather, the checklist aims at making us more *aware* of the choices to be made.

This report is primarily meant for MNP and its partner organisations. Other researchers may want to apply this report to their disciplines or to plan research on communicating uncertainty. Those involved in setting policies can use the report to gain insight into what methods the MNP researchers use.

2.   Checklist for uncertainty in spatial information

1. **Guidance.**   Use the 'Guidance for uncertainty assessment and communication' as a general basis.

2. **Representing uncertainty.**   Of the several methods for representing uncertainty in maps, we will discuss four here. Firstly, uncertainty can be illustrated by a *difference map*. Difference maps are essential to evaluating and presenting the difference between a map generated by a model and measurements (see Figure 5 for an example). Secondly, uncertainties can be illustrated by presenting two or more maps for different *scenarios* (Figure 3). Thirdly, we can use *ensemble maps* to represent uncertainties in spatial information (Figure 6). Finally, maps can be presented with margins for each grid cell. Several formats are available for such *grid uncertainty* maps. For each grid cell we can show the standard deviation or relative error (Figures 2 and 15A), a given percentile value (Figure 15B), a minimum or maximum value, or the probability that a limit is exceeded (Figure 15C). A probability map for each category provides an equivalent presentation for nominal maps (Figure 19).

   Each of the first three methods can show how the *patterns* in the map may change due to uncertainty. The fourth method puts more emphasis on the uncertainty per grid cell and is particularly effective for revealing if the variable under consideration exceed a governmental standard or limit value.

3. **Presentation formats.**   After selecting one of the visualisation methods listed under item 2, we can choose between different *presentation formats*. Some will be more suitable than others, and there is a trend towards integrating map values and their grid uncertainties in a single map (uncertainty given by isolines, 3D maps with the uncertainty plotted on the z-axis, fuzzy category boundaries, etc.). However, these bivariate presentations are not particularly effective in communicating with lay readers and users. This is because an excess of information on a map impedes the communication process. Thus, the map and uncertainty should preferably be presented separately. Another trend is the *dynamic* presentation of the uncertainty on Internet or via a PowerPoint presentation. Animations of ensemble maps are particularly effective. Other methods, such as a higher blink rate for areas with a higher uncertainty, communicate the uncertainty much less effectively.

4. **Calculation and presentation level.**   We should distinguish between the calculation and presentation levels. The uncertainties have to be determined at both levels. We should discuss the uncertainties with the researchers or organisations using our information and together decide about the level of detail in calculations and presentations.

5. **Resolution.** We should be critical and wary of the trend towards using ever higher spatial resolutions in calculations, and avoid using an excessively high spatial resolution in publications.

6. **Classes.** We have to consider the sensitivity of spatial results to subjective choices when deciding on the division into legend classes.

7. **Meta data.** One should never supply data without the associated relevant information (meta data); the meta data should preferably include information about uncertainty. If, when using a fine level of detail, the uncertainties are too high to use the data at that level, this aspect should be discussed with the user and the data should be supplied under strictly agreed terms. If in doubt about this, discuss the issue with the advisors listed under item 10.

8. **Requesting meta data.** We should set clear standards for the quality of the data supplied to us for use in models or maps; this should preferably be accompanied by uncertainty information. Information on the way in which the data was collected or created is also relevant.

9. **Proactively communicating uncertainty.** We should provide relevant supporting information and explain how results were obtained and may be used, even when the user does not request this information. We should ensure that this information is not lost and draw the users' attention to it.

10. **Advisors.** Anyone requiring help on dealing with uncertainties in spatial information should contact (statistical) advisors.

11. **Planning.** Dealing effectively with uncertainty in spatial information takes time and we should allow for that in our schedules.

3.   Explaining the checklist

1.   ***Guidance.***   *Use the 'Guidance for uncertainty assessment and communication' as a general basis.*

The Guidance for uncertainty assessment and communication' [26] aims to create awareness of importance of identifying uncertainties. The checklist also provides suggestions for communicating those uncertainties. However, it does not discuss in detail how researchers should deal with uncertainties in spatial information, which is something this checklist aims to assist with.

This checklist focuses on item 5 of the Guidance (identifying and assessing relevant uncertainties) and on item 6 (reporting uncertainty information) for uncertainties in spatial information. Please note that the main Guidance document should also be used in projects related to spatial information, at both the start and the end of a project.

Figure 4   Expert meeting on uncertainty communication, dealing with the topic of fine particles ($PM_{2.5}$ and PM_{10}) and health. These meetings are organised in the context of the 'Guidance for uncertainty assessment and communication'.
Photograph: H. Visser

*2.   **Representing uncertainty.**   There are several methods for representing uncertainty in maps and we will discuss four such methods here. Firstly, uncertainty can be illustrated by a difference map. Difference maps are essential to evaluating and presenting the difference between a map generated by a model and measurements (see Figure 5 for an example). Secondly, uncertainties can be illustrated by presenting two or more maps for different scenarios (Figure 3). Thirdly, we can use ensemble maps to represent uncertainties in spatial information (Figure 6). Finally, maps with margins for each grid cell can be presented. Several formats are available for such grid uncertainty maps. For each grid cell we can show the standard deviation or relative error (Figures 2 and 15A), a given percentile value (Figure 15B) a minimum or maximum value, or the probability that a limit is exceeded (Figure 15C). A probability map for each category provides an equivalent presentation for nominal maps (Figure 19).*

Each of the first three methods shows how the patterns in the map may change due to uncertainty. The fourth method puts more emphasis on the uncertainty per grid cell and is particularly effective for showing if the variable under consideration exceeds a governmental standard or limit value.

The first visualisation method shows the uncertainty using a *difference map*. Difference maps are essential when assessing model-generated maps and maps of the "real", measured situation. Difference maps are therefore important in calibrating and validating models which produce maps as an output of a mathematical process. Obviously this requires errors in the measured map to be negligible! **Figure 5** gives an example of two nominal maps (top left "measured" and top right "predicted") and a fuzzy difference map (at the bottom). Difference maps can also be informative. For example, when we present two land-use scenarios for 2030, adding a fuzzy difference map can show where major, moderate and minor changes are to be expected in future. Please refer to [8], [34] and [35].

Another method of representing uncertainties in maps is to provide two maps calculated using different *scenarios*. These maps will show the reader how the values in the maps, as well as the patterns, can change under the influence of different assumptions. Figures 3 and 20 provide examples.

A third presentation method uses so-called ensemble maps. These maps are produced with the same model, but using different parameter values. Presenting two or more of these maps clearly shows the reader how the values and patterns may change due to uncertainty. **Figure 6** shows four ensemble maps from a set of one thousand. Another example in the field of soil-terrain models is given in [3].

It can be quite difficult to select the type of maps most appropriate for publication. This aspect is addressed in Section 4.1.
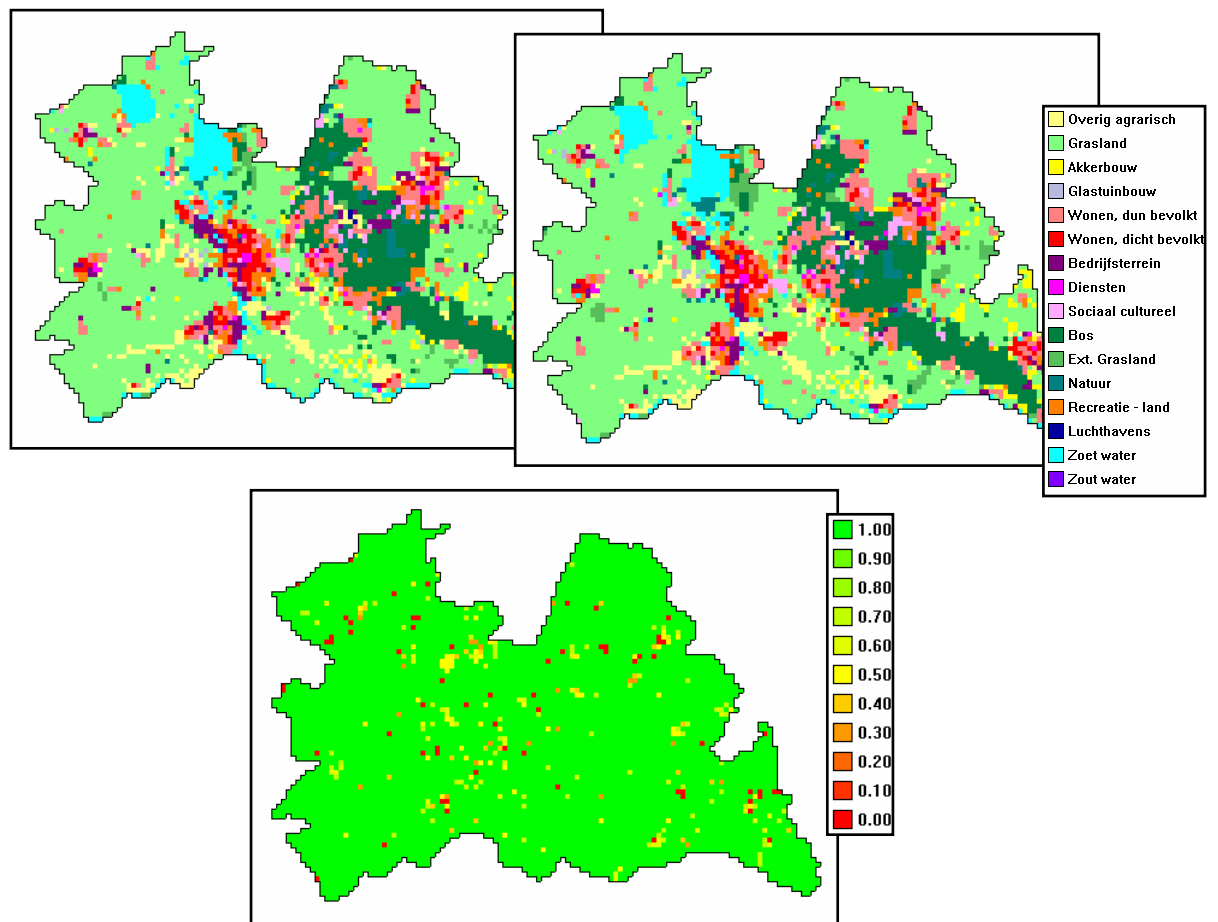
Figure 5   *Land use in the province of Utrecht, 1993, plotted on a 500 x 500 m² grid.*
                 The map at the top left shows the land use based on the Statistics Netherlands map
                 for that year. The map at the top right shows a simulation by the Environment
                 Explorer tool ("LeefOmgevingsVerkenner", LOV). The map at the bottom is a fuzzy
                 difference map produced with the MapComparisonKit-3 software. The difference
                 map shows where there is high (green), moderate (yellow) or poor similarity
                 (orange/red).

Please note that the difference between scenario maps and ensemble maps is that scenario
maps are not based on probability distributions. Instead, they show a range of possible
outcomes based on certain coherent assumptions. By contrast, ensemble maps use
probability distributions to simulate consistent maps, driven by uncertainty in the model
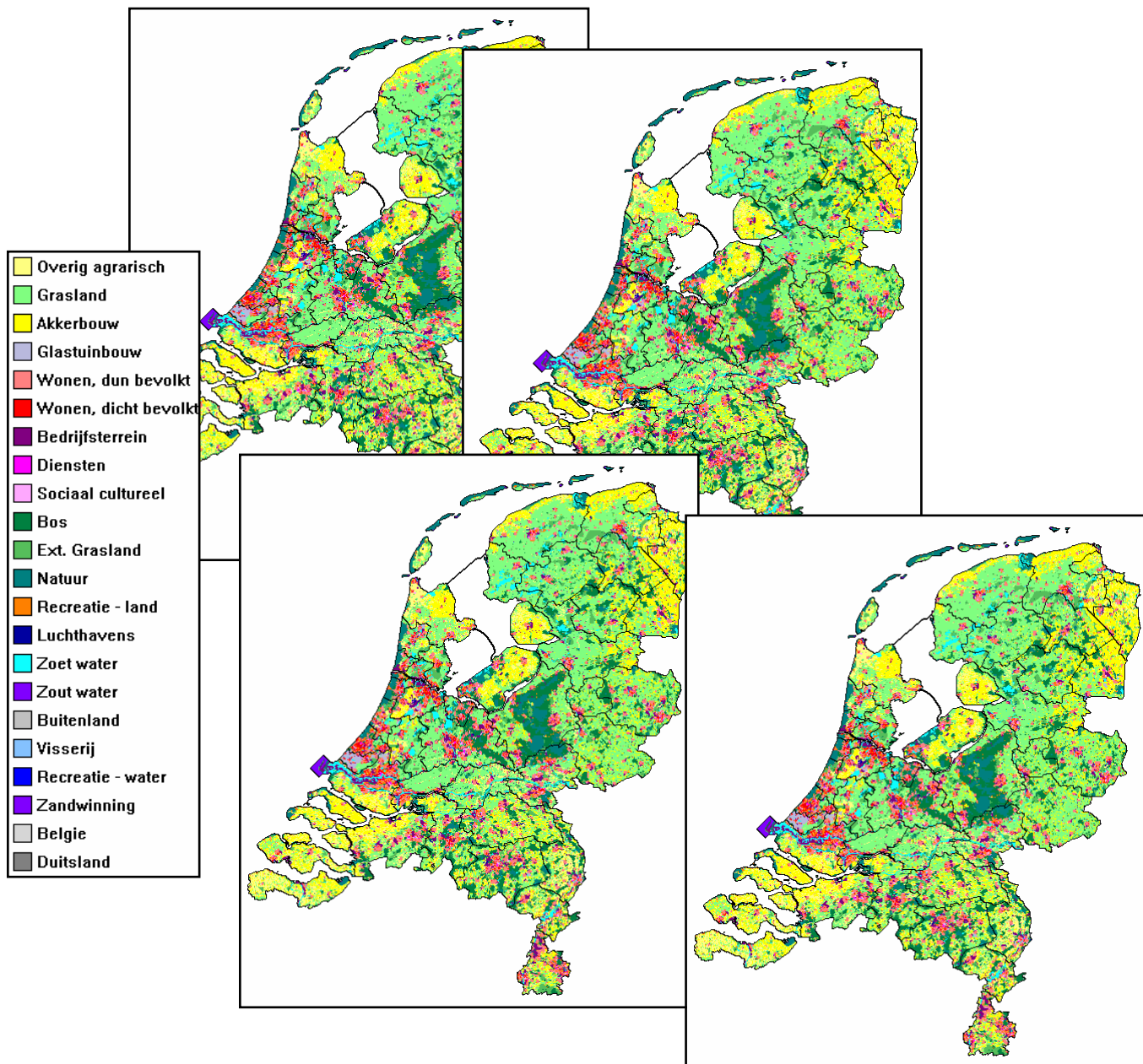parameters and any uncertainty in the inputs of the model.

Legend:
- Overig agrarisch
- Grasland
- Akkerbouw
- Glastuinbouw
- Wonen, dun bevolkt
- Wonen, dicht bevolkt
- Bedrijfsterrein
- Diensten
- Sociaal cultureel
- Bos
- Ext. Grasland
- Natuur
- Recreatie - land
- Luchthavens
- Zoet water
- Zout water
- Buitenland
- Visserij
- Recreatie - water
- Zandwinning
- Belgie
- Duitsland

Figure 6   *Four land-use ensemble maps for the year 2030 based on the Spatial Policy Document.*
Each map is based on a different set of parameter values for the LOV model. There are 20 land-use classes, varying from "other agricultural" (light yellow), to "forest" (dark green) and "recreation" (orange). Source: [21].

Finally, spatial data can be presented with the uncertainty per grid cell. If a probability distribution is known for the grid cells, this may be used to present the standard deviation or relative deviation (100% * standard deviation / mean value). We can also choose an arbitrary percentile, e.g. the 2.5th or 97.5th percentile. An interesting alternative to the standard deviation is the difference in each grid cell for the 95th and 5th percentile or the difference

between the 75th and 25th percentile (also known as the quantile range). See Figures 2, 15 and 16 for examples.

When a probability distribution is known for each grid, it is easy to visualise the probability of a particular threshold being exceeded or not. If there is no probability distribution available for the grid but values for two or more scenarios, it would be possible to determine a minimum value for each grid cell. In this way a "minimum map" and a "maximum map" can be produced. However, the interpretation of the spatial patterns should be considered carefully: the values in the minimum and maximum maps apply to *each grid cell*, hence the maps may not give a coherent presentation of the real situation.

When we are dealing with *nominal* quantities, the equivalent is the probability that a particular category may occur in any given grid cell. These probabilities are easily derived from a set of ensemble maps. We can then present probability maps for relevant categories, see Figure 18 for an example.

If calculation and provision of the maps is intended to serve as *input for another model*, then scenario maps and, especially, ensemble maps are the most appropriate tools to calculate the uncertainty throughout the full chain of models (see also item 7). Grid uncertainty maps are rarely appropriate for this. See also [3], [5, Chapter 10] and [11].

Section 4 includes a further discussion, illustrated with examples.

3.   ***Presentation formats.***   *After selecting one of the visualisation methods listed under item 2, we can choose between different presentation formats. Some will be more suitable than others, and there is a trend towards integrating map values and their grid uncertainties in a single map (uncertainty given by isolines, 3D maps with the uncertainty plotted on the z-axis, fuzzy category boundaries, etc.). However, these bivariate presentations are not particularly effective in communicating with lay readers and users. This is because an excess of information in a map impedes the communication process. Thus, the map and uncertainty should preferably be presented separately. Another trend is the dynamic presentation of the uncertainty on the Internet or using a PowerPoint presentation. Animations of ensemble maps are particularly effective. Other methods, such as a higher blink rate for areas with a higher uncertainty, communicate the uncertainty much less effectively.*

In general, best guess maps and their uncertainty are presented separately in publications. However, some prefer to combine these in one illustration.

Examples include:

- the use of uncertainty isolines in a best guess map (Figure 11)
- 3D maps with the uncertainty per grid point on the z-axis (Figure 22)
- making the category boundaries more or less distinct to indicate the uncertainty of the boundary
- using hatching or colour saturation (texture) methods to make areas with a higher uncertainty less transparent

Further options are discussed in [30, pp. 140-159] and the appropriate literature.

In general, maps will be come more difficult to understand if the information density is too high. **Figure 7** is an example of this. Lay users are unlikely to understand the message communicated by Figure 7b. Hence, it is often better to present the mean values map and the related grid-point uncertainty in different figures.
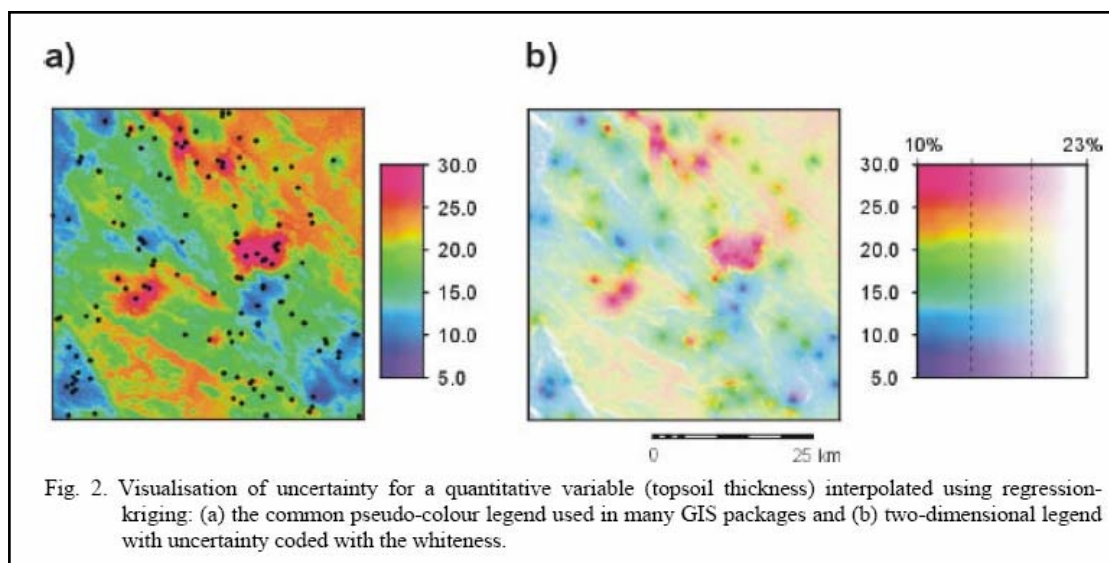

Fig. 2. Visualisation of uncertainty for a quantitative variable (topsoil thickness) interpolated using regression-kriging: (a) the common pseudo-colour legend used in many GIS packages and (b) two-dimensional legend with uncertainty coded with the whiteness.

Figure 7   *Map a) shows interpolated values of a continuous variable (topsoil thickness in cm). Map b) presents the same information, but now combined with the spatial uncertainty.*
The legend is two-dimensional, with the layer thickness on the y-axis, and the relative error on the x-axis. The less saturated (i.e. whiter) the colour, the less certain the interpolation. Source: [9].

There is also a trend towards *dynamic presentation* of uncertainty using the Internet or a PowerPoint presentation. So far, most dynamic presentations use animations. Animated

ensemble maps can clearly show how map patterns may change due to uncertainty. A good example is found in [9].

Other techniques are less effective, some examples include:

- flashing areas, a higher flash rate indicates a higher uncertainty
- flashing areas but with colour changes
- zooming in on areas with a high accuracy is possible, but zooming in on less certain areas is more difficult, corresponding to the increase in uncertainty

See also [30, pp. 140-159] and referenced literature. Again, the information is communicated less effectively than when uncertainty information is presented separately.

More general aspects of map presentation, such as the layout, choice of colour, area symbols and quality of the information are discussed in the Cartography Handbook [1]. These presentation aspects will not be discussed here.

4. ***Calculation and presentation levels.*** *One should distinguish between the calculation and presentation levels and determine the uncertainties at both levels. Discuss the uncertainties with the researchers or organisations using MNP information and decide together about the level of detail in calculations and presentations.*

It is a advisable to distinguish between the calculation and presentation levels. For example the calculations may be based on grid cells of 0.5 x 0.5 degrees, while the results are analysed and presented at a higher level of aggregation, e.g. the catchment area of a river. The choice of the presentation level depends on the highest spatial resolution at which the results can be usefully interpreted and presented. For example, one may identify the input parameters with the lowest resolution. Often, these parameters will determine the adequate resolution of the presentation level. Hence we have to look at the *source* of the input files. At what spatial scale were they calculated? What methods were used for scale conversions?

We consider the LGN land-use map of the Netherlands, with a 25 x 25 m^2 grid as an example. If our model needs an input grid of 250 x 250 m, how should we obtain that from the LGN map? We could use the dominant land use in each grid cell of 250 x 250 m^2. Or we could use a midpoint method to transform the LGN data to a 250 x 250 m^2 grid. If we undertake the calculations with a resolution of 250 m, what would be a good presentation level? At the national level (i.e. for the whole of the Netherlands, the LGN will be represented acceptably if the category frequency distribution is maintained during upscaling (the SPAT-tool software can be used for this purpose). However, this might not be appropriate at the local level.

When using models we have to be aware that if models are calibrated and validated using measurements, these measurements will generally be point measurements. You should determine the level at which the model needs to be calibrated or validated. If you use downscaling or upscaling, you have to determine the sensitivity of the method you use [2].

The problems associated with aggregating the results of calculations to a higher level (i.e. upscaling) are discussed in Heuvelink's article [11].

5.   *Resolution.   One should be critical and wary of the trend towards using ever higher spatial resolutions in calculations, and avoid using an excessively high spatial resolution in publications.*

You should be aware that using a finer scale for the calculations can reduce the reliability at the grid cell level, although there are exceptions to this rule. Normally, aggregation will increase the reliability.
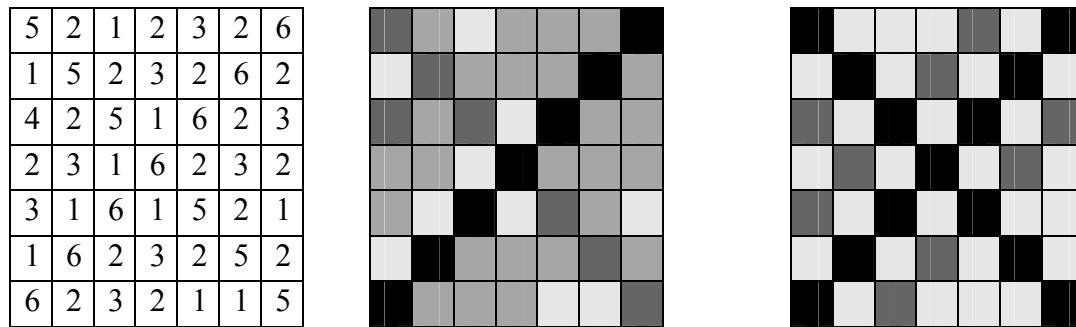
When determining if certain methods or models need to be improved by using a higher resolution we have to balance the required time and resources with the potential improvement in the estimates of the effects. This is because an assumed improvement of the model will not necessarily lead to better overall results. See also [11].

6.   *Classes.   Consider the sensitivity of spatial results to subjective choices when deciding on the division into legend classes.*

A division into classes (intervals) is often used when presenting results as maps. This means that the results to be presented first have to be divided into classes. You should be aware that both the choice of the number of classes and of the class boundaries is always subjective. It is also quite possible that certain patterns will disappear from a map if a different class division is used.

The map on the left in **Figure 8**, with cell values ranging from 1 to 6, provides a good example. It was used to make two maps with different class divisions:

- map in the centre: [1] [2 3] [4 5] [6]
- map on the right: [1 2] [3 4] [5 6]

5	2	1	2	3	2	6
1	5	2	3	2	6	2
4	2	5	1	6	2	3
2	3	1	6	2	3	2
3	1	6	1	5	2	1
1	6	2	3	2	5	2
6	2	3	2	1	1	5

Figure 8   *The map on the left shows an area with values ranging from 1 to 6. The maps in the centre and on the right use different class intervals. The patterns look fairly different, while the underlying data (map on the left) is identical.*

If the results to be presented are uncertain, the choice of classes becomes particularly important. The objective is to limit the loss of information resulting from the choice of class limits. This loss of information is often directly dependent on the uncertainty of the results. Furthermore, the uncertainty is often not the same for all results. Concentration measurements are a good example; test equipment is generally much less reliable at low concentrations than at high concentrations.

Especially for uncertainties with an asymmetrical (e.g. log-normal) distribution, it is inadvisable to choose equidistant classes. It will often be necessary to use classes of increasing size, as discussed in [10]. Another example is found in a symmetrical distribution, where the uncertainty is specified in relative terms, e.g. when all the results have an uncertainty band of ±10%. In this case too, an equidistant classification will result in a greater loss of information for the higher classes.

Figure 9 gives an example of this issue. The map at the top shows the worldwide distribution of NO_2, averaged over cloudless days in 2003 (identical to Figure 4.3.1 in the MNP Duurzaamheidsverkenning [28]). The map is based on measurements by the SCIAMACHY instrument of the Envisat satellite. The map at the bottom gives an enlarged view of Europe. The intervals used for the map at the top increase approximately logarithmically. The areas with low, moderate and high densities are readily apparent.

The map of Europe uses a linear scale. As a result, the areas with moderate concentrations have disappeared, leaving only areas with high densities (red) and safe areas (green, blue). Both the intervals and the use of colour (e.g. green and red) are different.
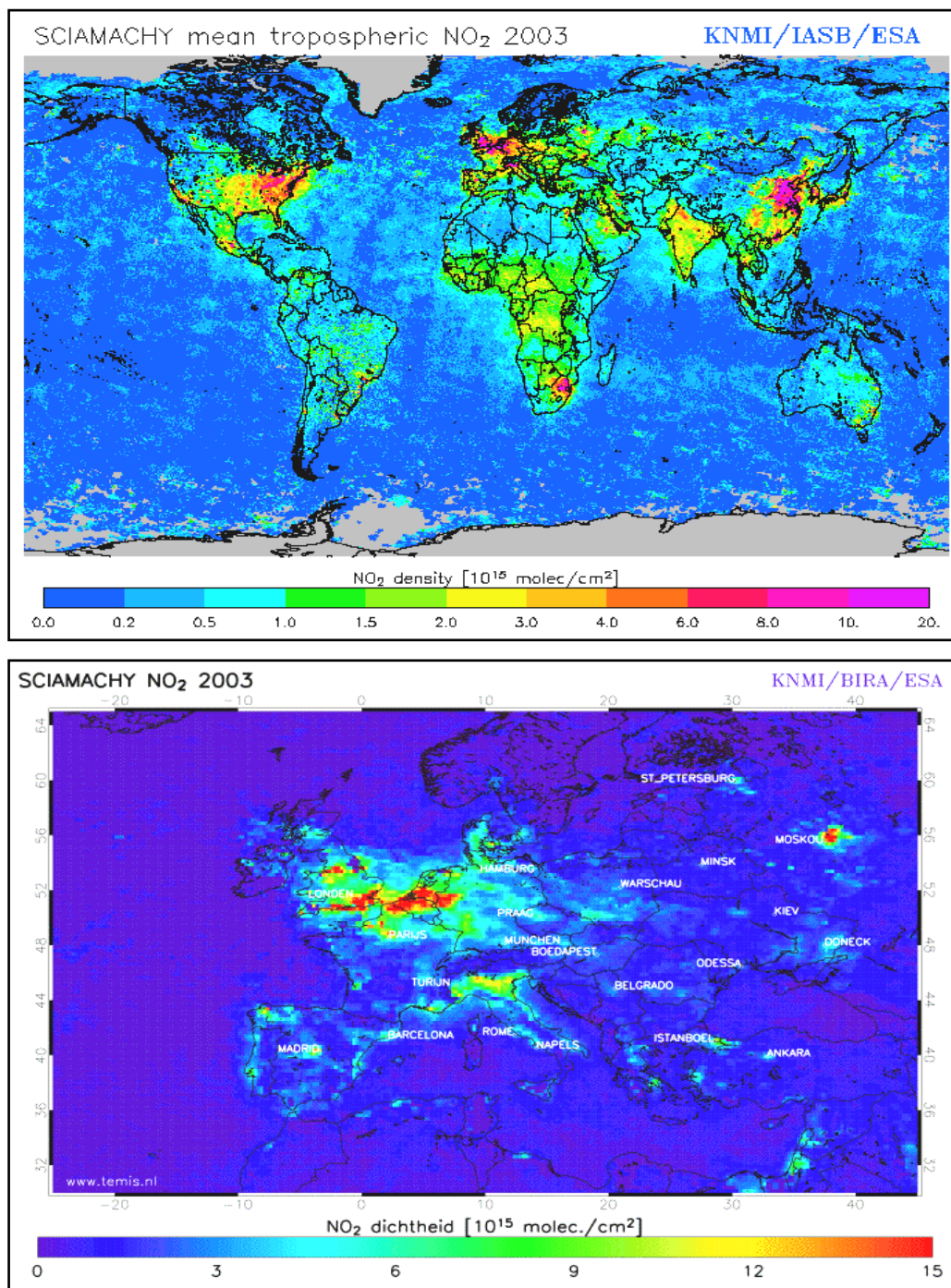
Figure 9   *NO₂ densities (in 10^15 molecule per cm²) measured by the SCIAMACHY instrument of the Envisat satellite.*

It should be noted the data was averaged over cloudless days in 2003. The map at the top shows the densities throughout the world using a **rising interval** scale, while the map at the bottom shows Europe only, using a **linear** scale. It should be noted that densities are for the total column measured, thus for the troposphere plus the stratosphere. Therefore, densities are not synonymous for concentrations at ground level. Source: [18].

7. ***Providing meta data.*** *Never supply data without the associated relevant information (meta data); the meta data should preferably include information about uncertainty. If, when using a fine level of detail, the uncertainties are too high to use the data at that level, then you should discuss this aspect with the user, and the data should be supplied under strictly agreed terms. If you have any doubt about this, discuss the issue with the advisors listed under item 10.*

Normally, when providing data you should not unquestioningly provide data at a particular level if you think that the uncertainty is too high at that level of spatial detail. After discussion with the user the data may be provided on the condition that the uncertainty will be considered when interpreting the results of the calculations using this data. It should be noted that there is an increasing demand for providing highly detailed information to the general public.

In each case you will have to decide, in consultation with the user, at what level we consider the data to be sufficiently accurate to provide it. When providing data to support the implementation of policies (e.g. scenario maps, emission maps and background concentration maps), make sure quantitative information about uncertainties is also included where available.

8. ***Requesting meta data.*** *Set clear standards for the quality of the data supplied for use in models or maps; data should preferably be accompanied by uncertainty information. Information about the way in which the data was collected or created, is also relevant.*

At the MNP many maps are used for chains of models. For example:

$$\textit{Spatial NO/NO}_2 \textit{ emissions} \rightarrow \textit{Dispersal (concentrations)} \rightarrow \textit{Deposition} \rightarrow$$
$$\textit{Soil chemistry} \rightarrow \textit{Plant growth}$$

Try to get an impression of the uncertainties elsewhere in the chain. For example, if the soil model turns out to be the weakest link in the chain, there is no need to use a highly accurate plant growth model. Hence, the various errors should be considered collectively.

If a further analysis is required to identify the uncertainty throughout the chain, then software such as USAtool (this software package has replaced UNCSAM) can be used to determine the progressive effect of errors throughout the chain [10, 26]. However, this will require a significant investment in time and considerable experience.

As many of our environmental models are incorporated into larger chains of models, we have to discuss these issues both with the users of our output data and with the suppliers of our input data. Such discussions to determine what data may or may not be used will sometimes result in actions for improvement which are relevant to the issues under consideration.

We can expect the organisations providing maps to know how to determine the accuracy of their products and provide information about this in writing. Otherwise we will be unable to tell how the accuracy, or rather the lack thereof, will affect our products. One should therefore try to arrange that meta data, and information on uncertainty, is always supplied. Of course, be tactful when broaching this issue. It may well help to discuss a staged plan.

The uncertainty in the data we use as input for our models should be reflected in the choice of spatial detail in the model calculations. A high level of detail is pointless if the input is highly uncertain at that level.

9.   ***Proactively communicating uncertainty.***   *Provide relevant supporting information and explain how the results were obtained and may be used, even when the user does not request this information. You should ensure that this information is not lost and draw the users' attention to it.*

To ensure that our work is transparent, we are responsible for adequate information on the accuracy and uncertainty (or references to this information) remaining closely associated with the data. This information may be included in reports or documents dedicated to the cause and, ideally, all information about the methods used should be available on a website. The backgrounds and limitations of maps should be described clearly (meta data) and safely kept. Aim for the highest degree of transparency to prevent inappropriate use of data.

10.   ***Advisors.***   *Anyone requiring help on dealing with uncertainties in spatial information, should contact (statistical) advisors.*

The contacts at the IMP/MNP are H. Visser, P.H.M. Janssen and P.S.C. Heuberger. The contacts at the RIM/MNP are A.A. Bouwman and A.C.M. de Nijs. They can also help you find reference material. A.C. Petersen of the IMP is the project leader of the 'Guidance for uncertainty assessment and communication' project. For general information on cartographical issues contact C.J. Bartels of the IMP-RPT team.

11. ***Planning.***   *Dealing effectively with uncertainty in spatial information takes time so allow for that in your schedule.*

Project leaders are expected to include resources in their project plannings for considering uncertainty. Using the 'Guidance for uncertainty assessment and communication' and the 'Checklist for uncertainty in spatial information' can help you plan this work. Nevertheless, it will take considerable time and effort!

4.　　Visualising spatial uncertainty

This section provides a more detailed discussion of the formats and methods for presenting uncertainty as discussed in earlier sections. In essence, these formats provide different views of the uncertainties. Section 4.1 uses an imaginary example, while section 4.2 includes a number of examples from actual MNP and RIVM projects.

4.1　　Four methods for representing uncertainty
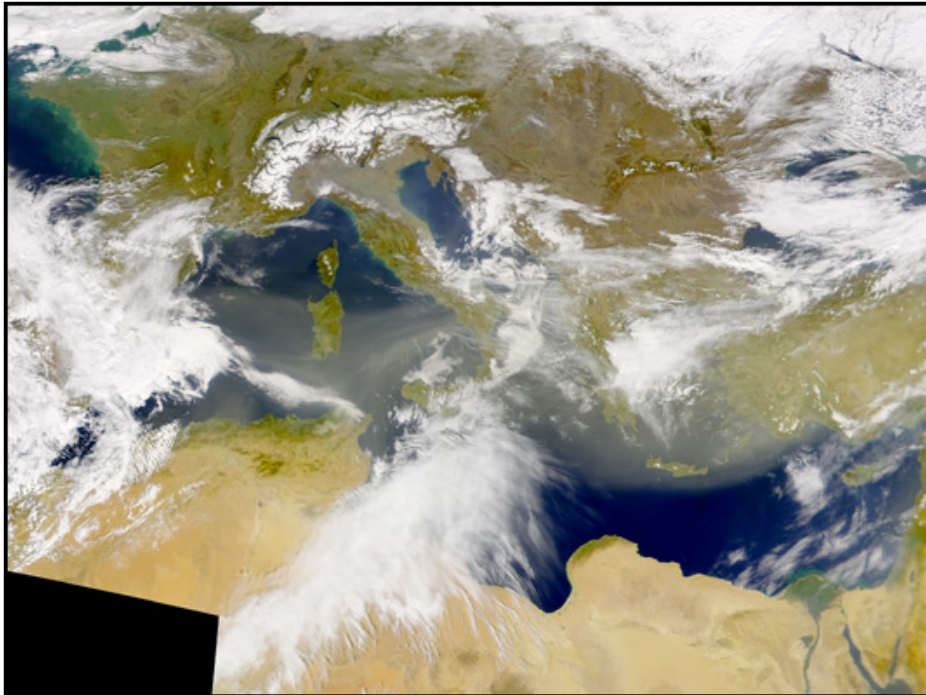
As discussed earlier, there are four methods for presenting uncertainty in maps:
- difference maps
- scenario maps
- ensemble maps
- grid uncertainty maps

A simplified example to explain the differences between these methods follows:

> *Let us assume that we have a model which generates a map showing the concentrations of pollutant y. The concentrations over and around the Mediterranean (**Figure 10**) were calculated using a regular grid. The model has one parameter, the deposition rate D, which is not known with high accuracy, only is that it is found between 1.0 and 3.0. In an article we would normally present a map where y is calculated with the mean value of D, i.e. 2.0. What are the other options, and how can we present the uncertainty?*

Difference maps
If we want to validate the model using measured concentrations of pollutant y then we have to compare the concentrations predicted by the model with measurements of y. There are two ways of doing this. We can use the model to forecast the concentrations at the measurement points and then plot the differences (residues) on a map. Alternatively, we can interpolate the measurements (e.g. using Kriging) to fit the grid used by the model, although this method will result in interpolation errors. The difference map will then show where the model is effective and where its results are poor.

Figure 10   *Photograph of component y over the Mediterranean, 4 March 2002.*
The pollution is brownish and covers a large part of the Mediterranean, the south of
Italy, Albania, Greece and Turkey. Photo: Sea-viewing Wide Field-of-view Sensor
(SeaWiFS) of the American SeaSTAR satellite.

Difference maps can also be used to compare two scenarios, for example, by using one map
based on D = 1.0 and one based on D = 3.0. Presenting only these two maps makes life
difficult for the reader, because it becomes a matter of "spot the differences". This can be
quite a challenge if the maps are very similar. Difference maps make it much easier to spot
the differences. The MapComparisonKit-3 software [8] can be used to generate difference
maps with any type of scale.

Scenario maps
Another option is the scenario method, often used by the MNP and the RIVM. We will
calculate three maps showing the concentrations of y:

- one map based on a high concentration scenario (using the lowest deposition rate:
  D = 1.0)
- one map based on a low concentration scenario (using the highest deposition rate:
  D = 3.0)
- one map based on an average concentration scenario (using the average of the
  deposition rate range: D = 2.0)

It is important to emphasise that we are not assigning probabilities to any of the three scenarios, other than saying that they are all three *equally probable*. It should be noted that there are ongoing discussions in groups such as the IPCC about assigning probabilities to scenarios.

In this publication we can choose to present all three scenario maps, or only the high and low scenario maps to emphasise the extremes. Alternatively, we could only show the middle scenario map and describe the other two maps in the text. For scenario maps it is not relevant if quantity y is a nominal, ordinal, interval or ratio quantity.

Ensemble maps

The third option would be to say that D lies between 1.0 and 3.0. All values between these limits are equally likely, hence we can make a large number of random draws from the interval [1.0, 3.0], e.g. 1000 times (if there is sufficient computer time; otherwise we have to use a lower number). This results in the randomly selected deposition rates D_1, D_2, ... , D_{1000}. Each value of D_k results in a map M_k of quantity y. The 1000 maps M_1, M_2, ..., M_{1000} are known as ensemble maps.

The next problem is which of these 1000 maps we should present in a publication. If the results are presented on the Internet or using a PowerPoint presentation we can use a dynamic presentation, an **animation** of the maps. However, if the results are published in print we will have to limit ourselves to a static presentation by selecting a small number of representative maps, perhaps two or three.

There are several potential selection criteria. If we want to present three maps, we could calculate the average of y for all map grid cells (i,j) in each map M_k. Here, we have 1000 values $y_{average,k}$. We can then present a low percentile map (e.g. a map with $y_{average,k}$ equal to the 5th or 25 percentile), a medium map ($y_{average,k}$ equal to the 50th percentile) and a high percentile map ($y_{average,k}$ equal to the 75th or 95th percentile).

A completely different criterion would be to place all the values in the maps in a vector and then determine the correlation matrix. The correlation coefficient gives an impression of the similarities between the *patterns* in the maps. We can then select one or two maps with the lowest correlation with the other maps (i.e. maps illustrating an infrequent pattern) and a map which is most similar to all the other maps (i.e. the map with the average pattern).

As far as the principle behind ensemble maps is concerned, it is irrelevant if y is a nominal, ordinal, interval or ratio parameter. However, correlation matrixes cannot be calculated for nominal or ordinal maps. The correlation matrix for these maps could be replaced by, for example, a fuzzy-Kappa matrix; see [34] for further details.

Grid uncertainty maps

The last option is to determine the probability distribution of y for each grid cell (i,j). This distribution is easily determined from the 1000 ensemble maps and is determined using the 1000 values $y_{i,j,k}$, where (i,j) is the location of the grid cell on the map and k= 1, ..., 1000. We could test if the values of $y_{i,j,k}$ have a normal or log normal distribution, but this is not actually necessary.

We can now use the probability distributions for each grid cell to present a range of indicators, such as the standard deviation, relative error or an arbitrary percentile value. For example, we could present a 2.5th percentile map and a 97.5th percentile map. Together with a 50th percentile map or a map of the arithmetic means (best guess map) we can now present three maps which give an impression of the uncertainties in the map. Alternatively, we could present two maps, a high percentile map and a low percentile map, which will lend greater emphasis to the extremes.

Grid uncertainty maps have a further use. If it is important to know the probability of the concentration of pollutant y exceeding a limit according to an EU Directive, we can calculate the exceedance probability for each grid cell using the probability distribution. This will produce a map of the probability of exceedance of the limit. Furthermore, we could calculate, for each grid cell, the probability that the limit is not exceeded while the estimated mean indeed exceeds the limit (first order error, α). To accompany this *α map* we can develop a *β map*, showing the probability for each grid cell that the limit will be exceeded while the estimated mean is actually **below the limit**.

These probability maps are useful if y is a continuous quantity (interval and ratio scale maps). However, this method cannot be used if y is a nominal or ordinal quantity. However, there is a similar option, a *probability map* for each category on the map.

Such probability maps are calculated as follows. For each grid point we know whether or not $y_{i,j}$ belongs to a given category c. This means that $y_{i,j}^{k}$ is binary: its value is "1" if $y_{i,j}$ in map k is included in category c and otherwise it is "0". By taking the mean of all 1000 binary values $y_{i,j,k}$ we can estimate the probability that this grid cell is included in category c. Together, the probabilities $p_{i,j}$ form the probability map for the selected category. We can then include maps for relevant categories in the publication.

Note 1: Some models produce probability maps as output without further work being required. In this case there is no need for the simulation. The land-use maps generated by the *Land-use Scanner* form a good example.

Note 2: When presenting a grid uncertainty map, explain to the reader or user that the map is not a result of the model, but rather presents the uncertainty for each grid cell.

4.2   Examples

4.2.1   Maps with interval and ratio scales

Figure 11 gives the first example, which includes two maps of the world, the quantity depicted is the enhanced global warming due to anthropogenic emissions of greenhouse gasses. The maps were produced using a climate model which included a number of assumptions for the emissions through to 2100: the SRES A2 scenario (top) and the SRES B2 scenario (bottom). The legend is divided into classes of integer degrees.

These two *scenario maps* illustrate that depending on the selected emissions, and the underlying assumptions, warming might be very different, both in terms of the absolute values and in terms of the pattern of the warming classes.

Additionally the makers of these two maps increased the information density by adding *grid uncertainty information* to each map. The blue isolines mark areas with larger or smaller ranges per grid cell (i.e. the difference between the minimum and maximum value). The green isolines depict the calculated standard deviations. As mentioned under item 3 of the Checklist, it is advisable to show the uncertainties in separate maps. To readers unfamiliar with the issues, Figure 11 provides an excess of uncertainty information.

Another example is given by **Figure 3** in the introduction. This figure shows the changes in precipitation over the period 2002-2052 [24]. The figure shows the scenario uncertainties in two dimensions:

- the four scenarios *Markets First*, *Policy First*, *Security First* and *Sustainability First*
- the four climate models *HADCM2*, *CGCM1*, *ECHAM4* and *CSIRO-MK2*

The figure itself already has a high information density as the depicted quantity is the difference between two maps (2052 precipitation map minus the 2002 precipitation map) which is presented in 16 versions.

Figure 12 gives another example; as in Figures 1 and 9, the maps show processed satellite data representing NO_2 concentrations. The map at the bottom depicts the grid uncertainty information as the relative error (%) over one day, while the map at the top shows the best-guess values. It is interesting to note that the relative error is high in areas where the concentrations are high (relative error 35 to 60%). See also [4] for further discussion.
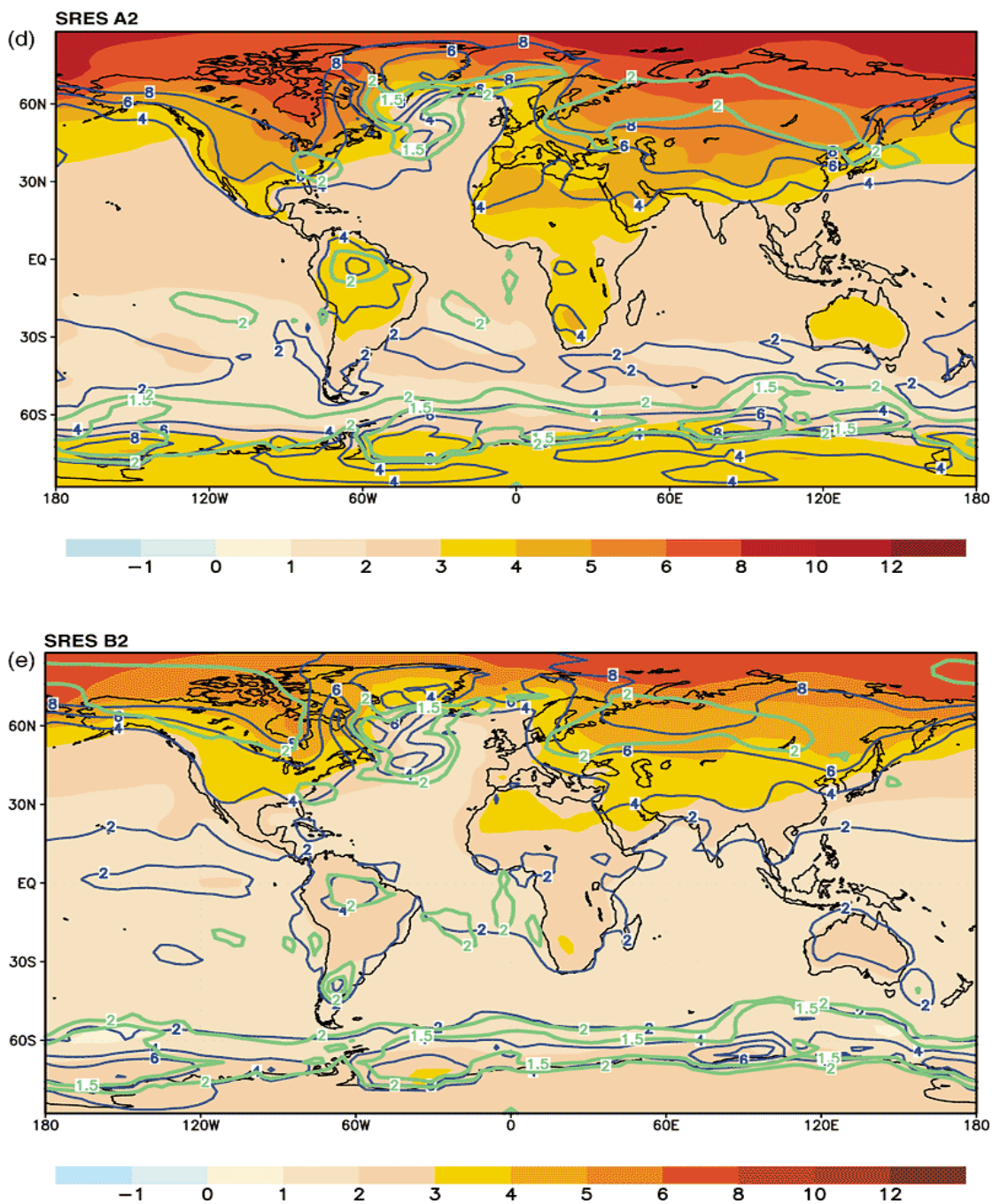
Figure 11   *Annual change in global temperature, in °C.*
The warming is the average over the 2071-2100 period relative to 1961-1990. The temperature classes (ranging from light blue to dark red) are divided by integer degrees. The global warming was calculated using a climate model for the SRES A2 scenario (top), and the SRES B2 scenario (bottom). The maps also contain two sets of isolines. The blue isolines mark areas with larger or smaller ranges (maximum value minus minimum value per grid cell). The green isolines mark areas with higher or lower standard deviations per grid cell. Source: [14, Figures 9.10d/e].
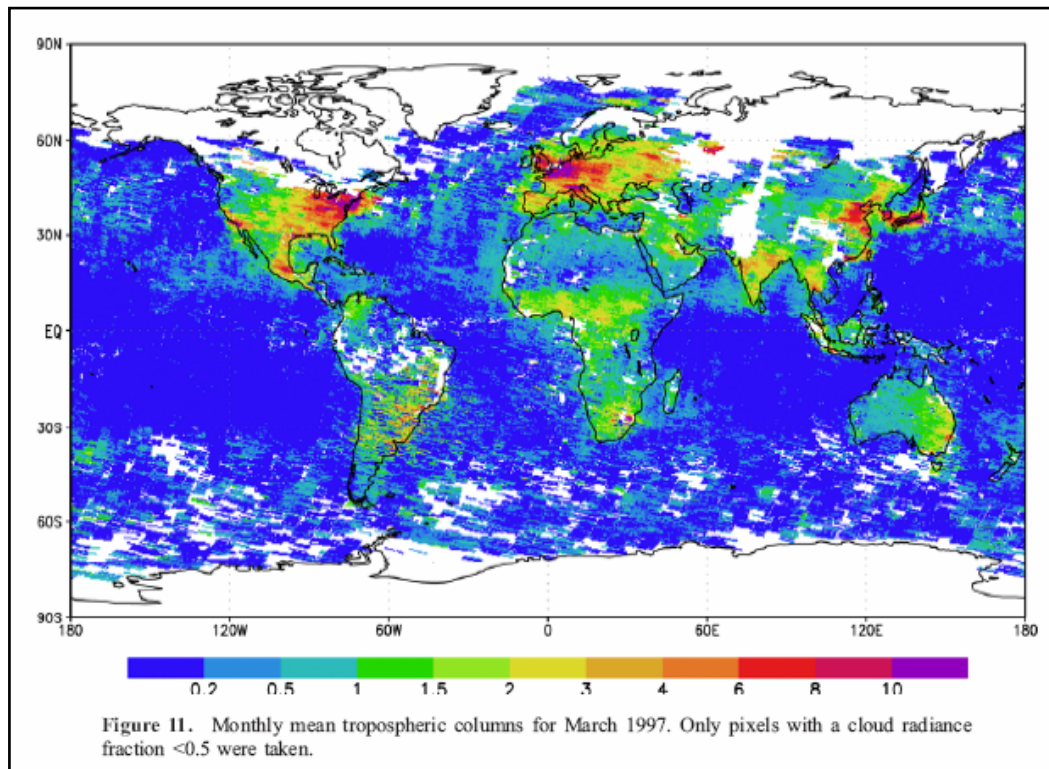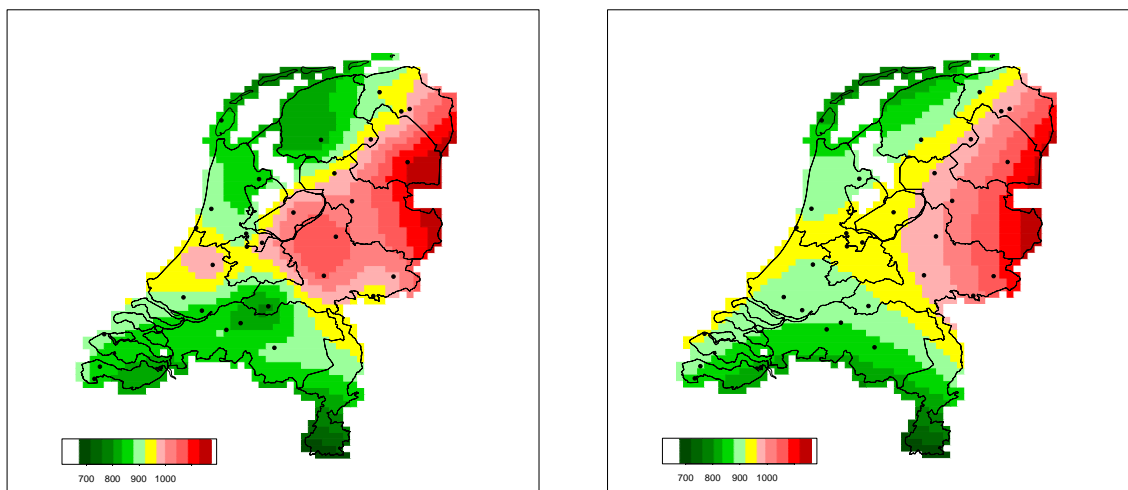
Figure 11. Monthly mean tropospheric columns for March 1997. Only pixels with a cloud radiance fraction <0.5 were taken.

Figure 12. Relative uncertainty in tropospheric columns for March 1997. Only pixels with a cloud radiance fraction <0.5 were taken.

Figure 12   *Error analysis of the determination of NO$_2$ concentrations in the GOME experiment (comparable with the measurements shown in Figures 1 and 9).*
The map at the top depicts the monthly mean concentrations in March 1997. The map at the bottom depicts the grid uncertainty as a relative error (standard deviation/mean * 100%) for one day in that month. (Note: if the mean is calculated over several days, the error will be reduced by $\sqrt{N}$.) Source: [4, used with permission].

Example 4. Let us assume we have 30 measuring locations for the total annual precipitation in the Netherlands. We want to determine the pattern of the annual totals throughout the Netherlands using a trend presented on a map (calculated using a geostatic model (Universal Kriging, see [16])). The Kriging model was used to estimate a quadratic plane where the unexplained element, the noise, was assumed to have a spatial correlation based on a *variogram*. The model estimate (spatially-correlated noise added to the quadratic plane) and the quadratic plane covering the Netherlands are shown in **Figure 13**. The maps were obtained by using the statistical model to make predictions for all values on a regular grid of 5 x 5 km^2 spanning the Netherlands with a total of 1680 grid points.

Figure 13   *A geostatistical model (Universal Kriging) was used to estimate a quadratic plane of the precipitation at 30 stations in the Netherlands, and the uncertainty was modelled with the spatially correlated noise.*
             The map on the left depicts the estimated model, i.e. the spatially correlated noise and the quadratic plane. The map on the right depicts the quadratic plane only. The dots on the maps indicate the locations of measuring stations.

In the estimated Kriging model, Monte Carlo simulations may be used to generate *ensemble maps* (methods discussed in [7]). **Figure 14** includes four ensemble maps, which show that the patterns may vary significantly due to the uncertainty. For example, in the first map the annual total precipitation in Noord-Holland is much lower than in the fourth map. The choices concerning the presentation of ensemble maps were discussed earlier.
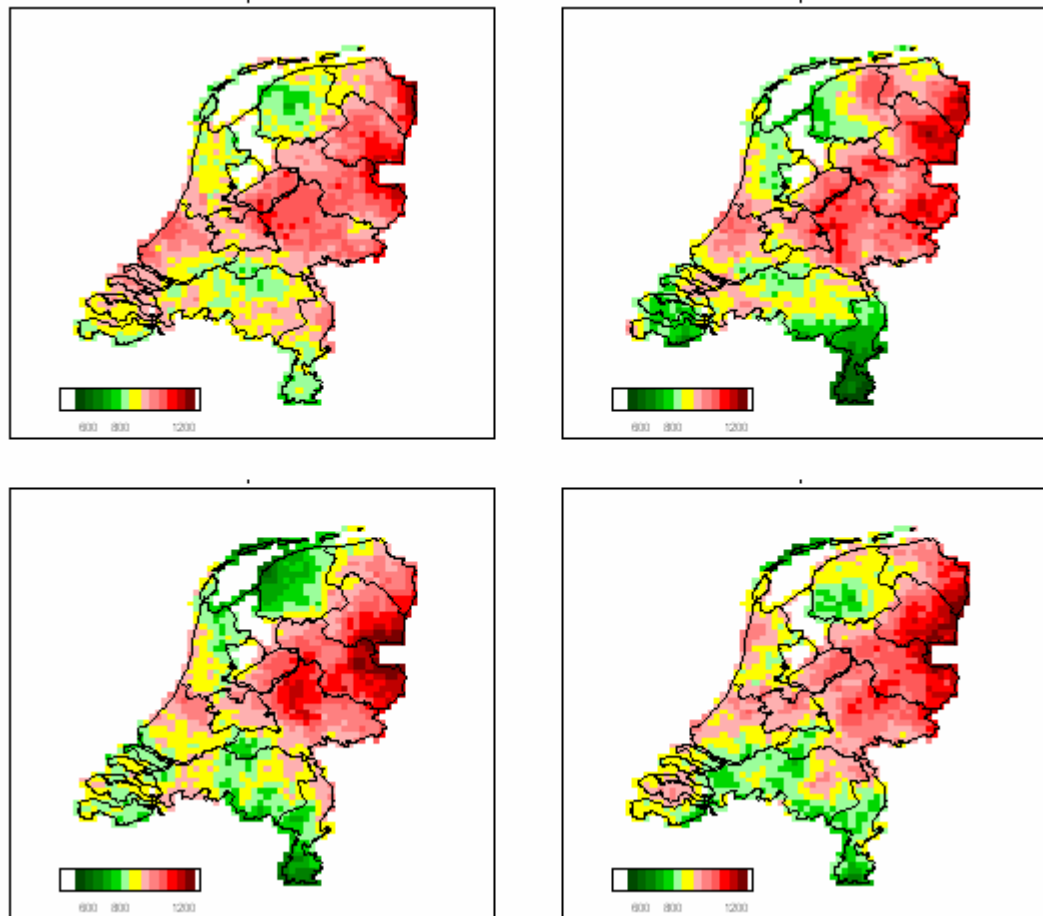
Figure 14   Four ensemble maps, based on the the modeling results shown in Figure 13.

The Kriging model also provides information about the *grid uncertainty*. The model specifies the mean and standard deviation for each grid cell. A statistical test showed the data to have a normal distribution. We can use this information to present the grid uncertainty in several ways.

Firstly, we can show the standard deviation for each grid cell on the map on the left in **Figure 15A**. Secondly, we can show the error relative to the best guess value as a relative error in per cent. This presentation is used for the map on the right. Thirdly, we can calculate any percentile value desired for each grid cell, **Figure 15B** includes the 5th and 95th percentile maps as examples.
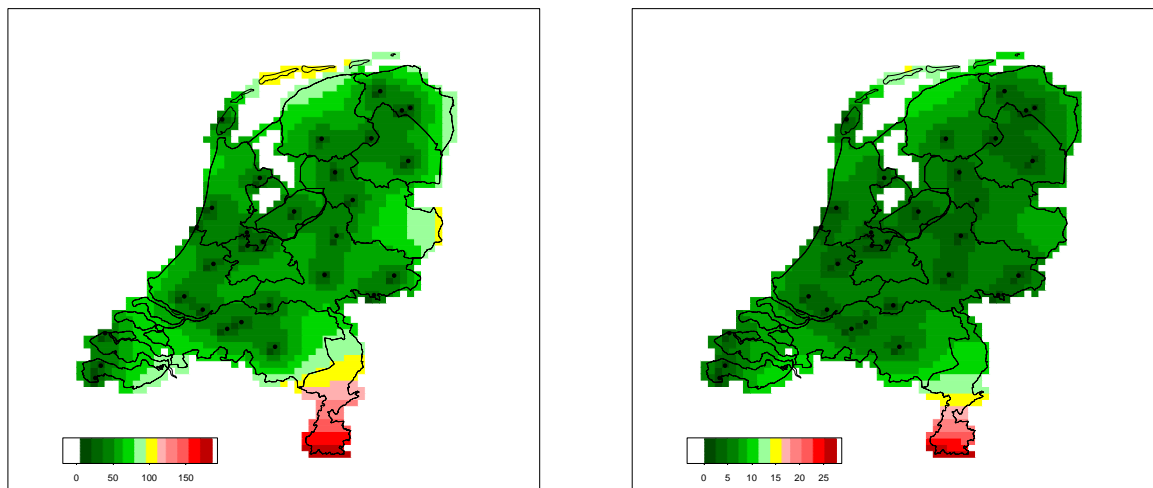
Figure 15A   *Standard deviations (1-sigma expressed in mm, on the left) and relative errors (100%*SD/mean, expressed in %, on the right) per grid cell (i,j) for the example in Figure 13. Both maps present grid uncertainty information.*
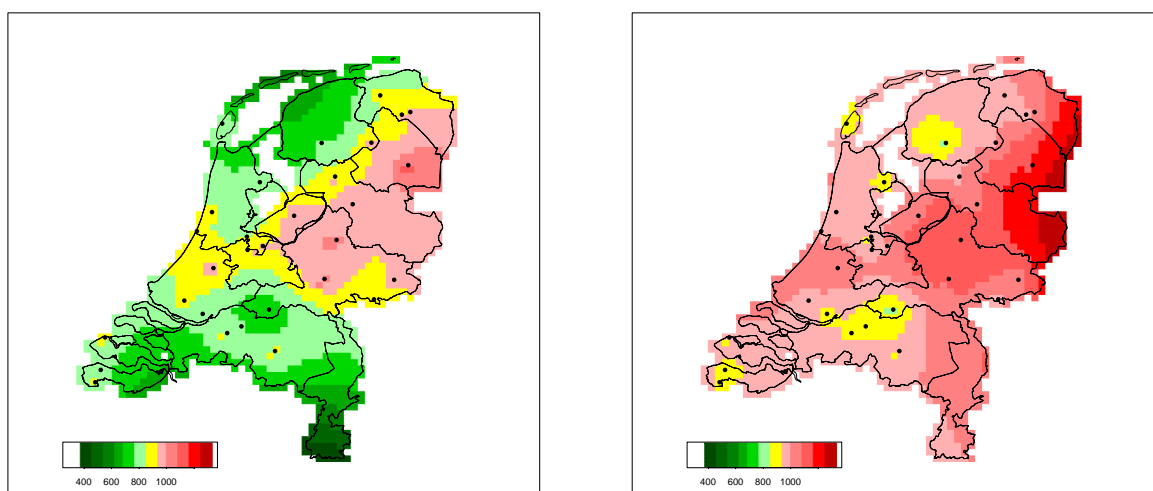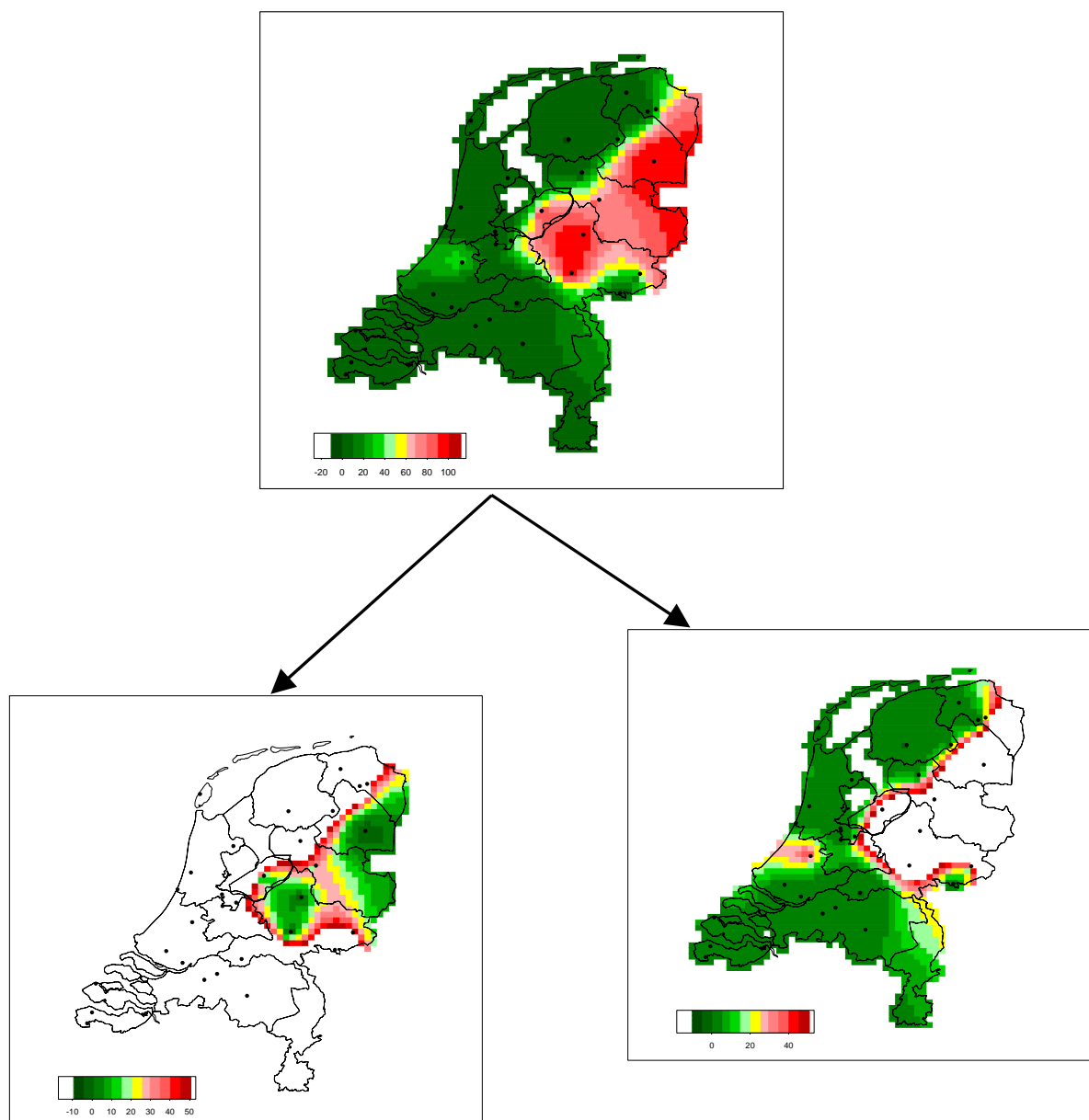

Figure 15B   *Two percentile maps for the example in Figure 13: 5-percentile map (left) and 95-percentile map (right). Both maps represent grid uncertainty information.*

Fourthly, we can show the grid uncertainty for *limit exceedance*. The upper map in **Figure 15C** shows the probability for each grid cell that the threshold of 1000 mm precipitation per year is exceeded. The next step is to mark the areas below the limit (safe) and those above it (unsafe). Of course, this is not really that relevant for a precipitation map. However, when we are dealing with atmospheric pollutants such as PM_{10} this division may

be relevant as both the EU annual limit and EU daily limit are exceeded in some areas in the Netherlands, but not in others.

If we now consider green areas to be safe and red areas unsafe, this division can result in two error types: a grid cell classified as unsafe is **actually safe** (first-order error, α) and, conversely, a grid cell classified as safe is **actually unsafe** (second-order error, β). As such errors can occur in every grid cell, we can make an α and a β map (see also the discussion in Section 4.1), as shown at the bottom of **Figure 15C**.

Figure 15C   *Three grid-uncertainty maps for the exceedance of a limit.*
The map at the top depicts the probability of exceeding the annual precipitation limit of 1000 mm. The 12 probability colour classes have a probability width of 10%. The 50th percentile (in this case equal to the arithmetic mean) is on the edge of the colours light green and yellow. If we interpret green classes as "safe" (e.g. for growing potatoes) and yellow/red areas as "unsafe", this division will result in two error types: a grid cell classified as unsafe is **actually safe** (first-order error, α) and, conversely, a grid cell classified as safe is **actually unsafe** (second-order error, β). As such errors can occur in every grid cell, we can construct an α and β map. The maps at the bottom show the α map (on the left) and the β map (on the right), for example. Note that the 12 probability colour classes in the lower graphs have a probability width of 5%.

Figure 16 shows another presentation form for percentile maps (compare Figure 15B). Upper and lower percentiles are comprised in a squared grid box using triangles. Although the triangle presentation is creative in that it condenses high and low grid percentiles *in one map*, we are not certain how well this presentation form communicates (compared to a presentation of two percentile maps, side by side, as in Figure 15B).
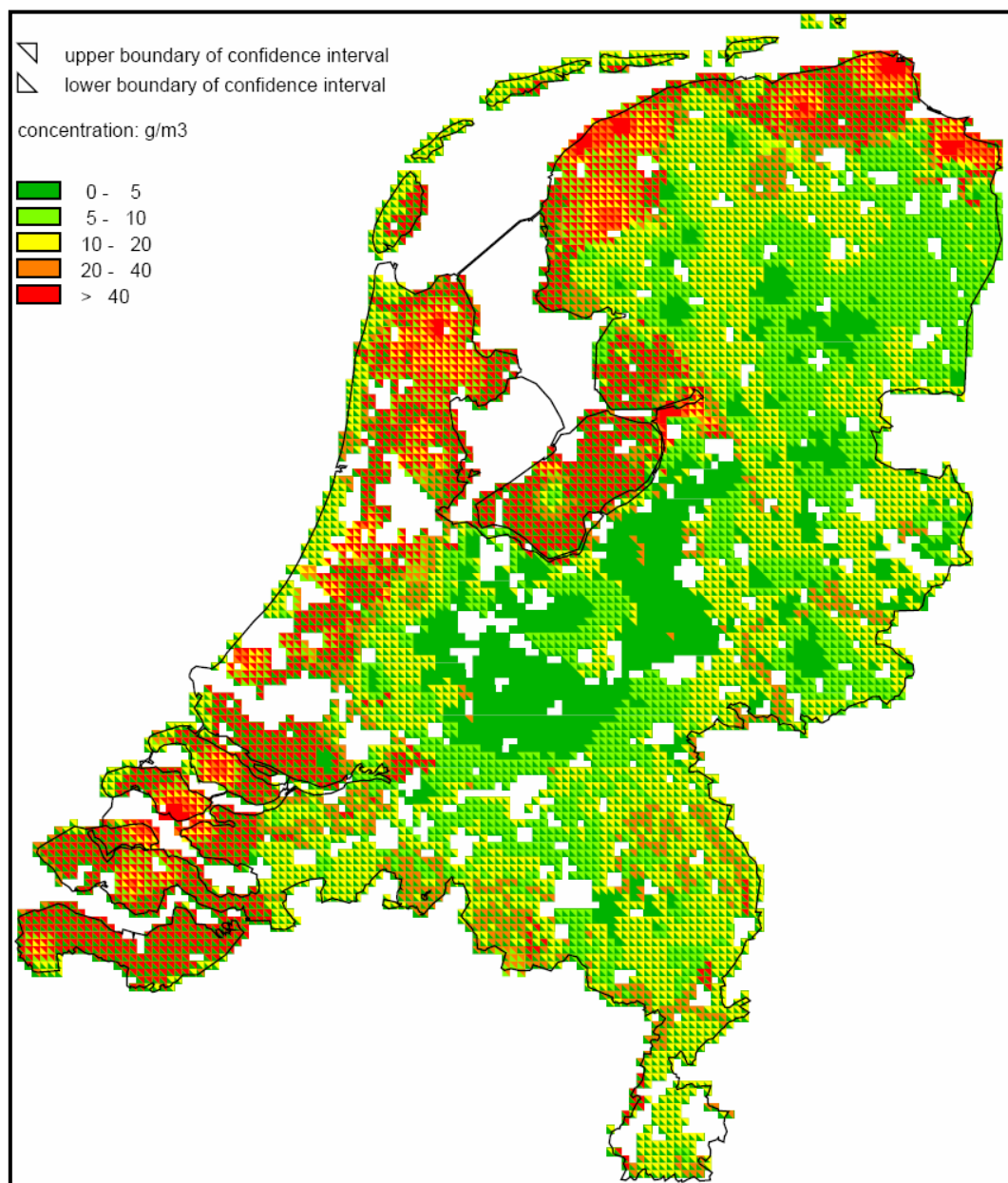

Figure 16   *Presentation of grid uncertainty by using a lower triangle for the 5 percentile and an upper triangle for the 95 percentile. Source: [23].*

A final example. The RIVM ambulance driving-time model uses a 500 x 500 metre grid to generate the time needed to drive from the nearest ambulance station, collect the victim and take the victim to the nearest Accident and Emergency department [34]. **Figure 17** includes two maps for the Netherlands: the one on the left is based on 110 hospitals with an A&E department and the one on the right is based on the assumption that 7 of these hospitals are closed. Hence, these two maps show ambulance driving times for two scenarios: 110 hospitals with A&E departments and 103 of such hospitals.

The map at the bottom is an example of a *difference map* that can be produced in several ways. Two difference maps, depicting the ratio of the driving times per grid cell and the difference in driving time per grid cell are given [34, Figure 24]. We can also present a difference map such as the one at the bottom of Figure 17, called a fuzzy difference map [8, 34]. This difference map emphasises the differences in *patterns* between the two maps. These differences span a continuum from -∞ (completely different maps) to 0.0 (correspondence in patterns that could be expected purely by chance) to 1.0 (identical maps). Here, the difference map clearly presents the differences in patterns between the two driving-time maps.
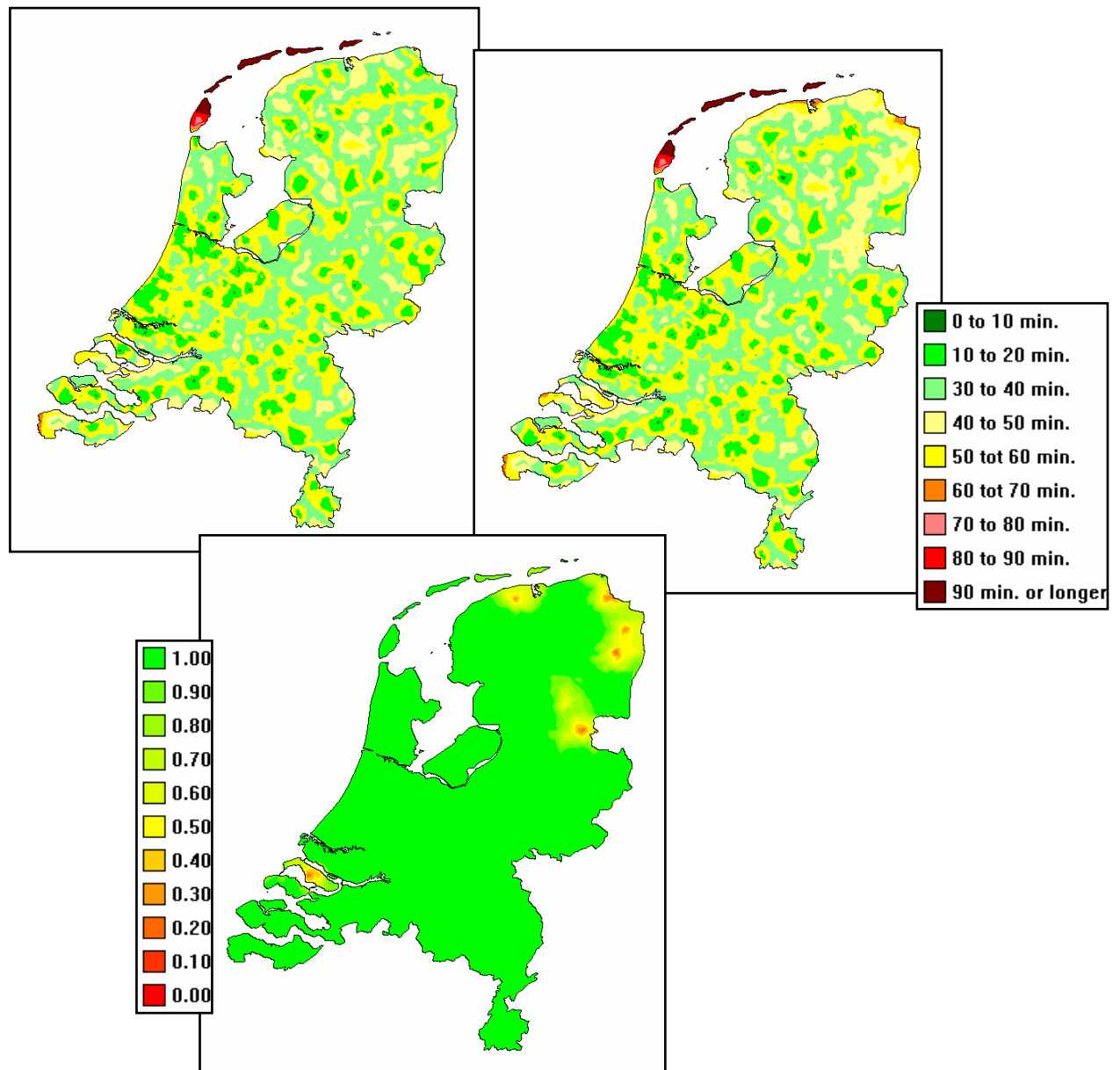
*Figure 17   Ambulance driving times based on 110 hospitals with an A&E department (top left)
               and 103 such hospitals (top right).*
               The fuzzy difference map (bottom) shows the degree of correspondence between the
               two maps (1.0 for maps that are identical and 0.0 to show that the correspondence
               between the maps is equal to that of two randomly generated maps with the same
               category frequencies).

4.2.2   Maps with nominal scales

An example of a *nominal difference map* used for the calibration and validation of the MNP *Environment Explorer* model (LOV) is presented in Figure 5. Extensive discussion and sample maps are included in [21]. The figure includes two land-use maps for the Province of Utrecht as well as a fuzzy difference map. The map on the left is aggregated from the Statistics Netherlands (CBS) map for 1993. The map on the right shows the land use as simulated for that year with the LOV software. A fuzzy difference map was generated to quantify the accuracy of the forecast [8].

Figure 18 shows an example with four *scenario maps* for the year 2030, calculated using the MNP *Land-use Scanner* model. The maps are taken from [5, 28]. The maps show the trend towards urbanisation in four classes: very low, low, average to high and very high. These scenarios, designated A1 (Global Market, map on the top left), A2 (Safe Region, map at the bottom right), B1 (Global Solidarity, map at the top right) and B2 (Caring Region, map at the bottom right) correspond to the assumptions made by the IPCC for the global developments through to 2100. For a discussion of these scenarios, refer to Chapter 2 of [28]. The maps clearly indicate that there is a greater trend towards urbanisation in the A scenarios than in the B scenarios.

An example of *ensemble maps* was given earlier in Figure 6. The LOV model was used to determine land-use maps for the year 2030 [19]. The scenario assumptions included were chosen on the basis of the assumptions of the *Spatial Policy Document*. A total of 1000 maps were simulated. Visually, the differences between the ensemble maps are hardly distinguishable. This means that uncertainties in the LOV will have little impact on the simulated land use in 2030 (please note that this excludes the uncertainties on the assumptions underlying the model!). The differences between the maps could be reveiled with the MapComparisonKit software, using either difference maps or fuzzy-Kappa values.

This set of 1000 maps was then used to determine a set of *probability maps* for different categories. **Figure 19** is a probability map for the aggregated category "urbanisation". Reference [21] also includes probability maps for the categories "housing" (Figure 14 of that document), "industry/commerce" (Figure 15), "offices" (Figure 16), "social and cultural facilities" (Figure 17), "recreation" (Figure 18) and "greenhouses" (Figure 19). The category "urbanisation" combines these six basic categories.
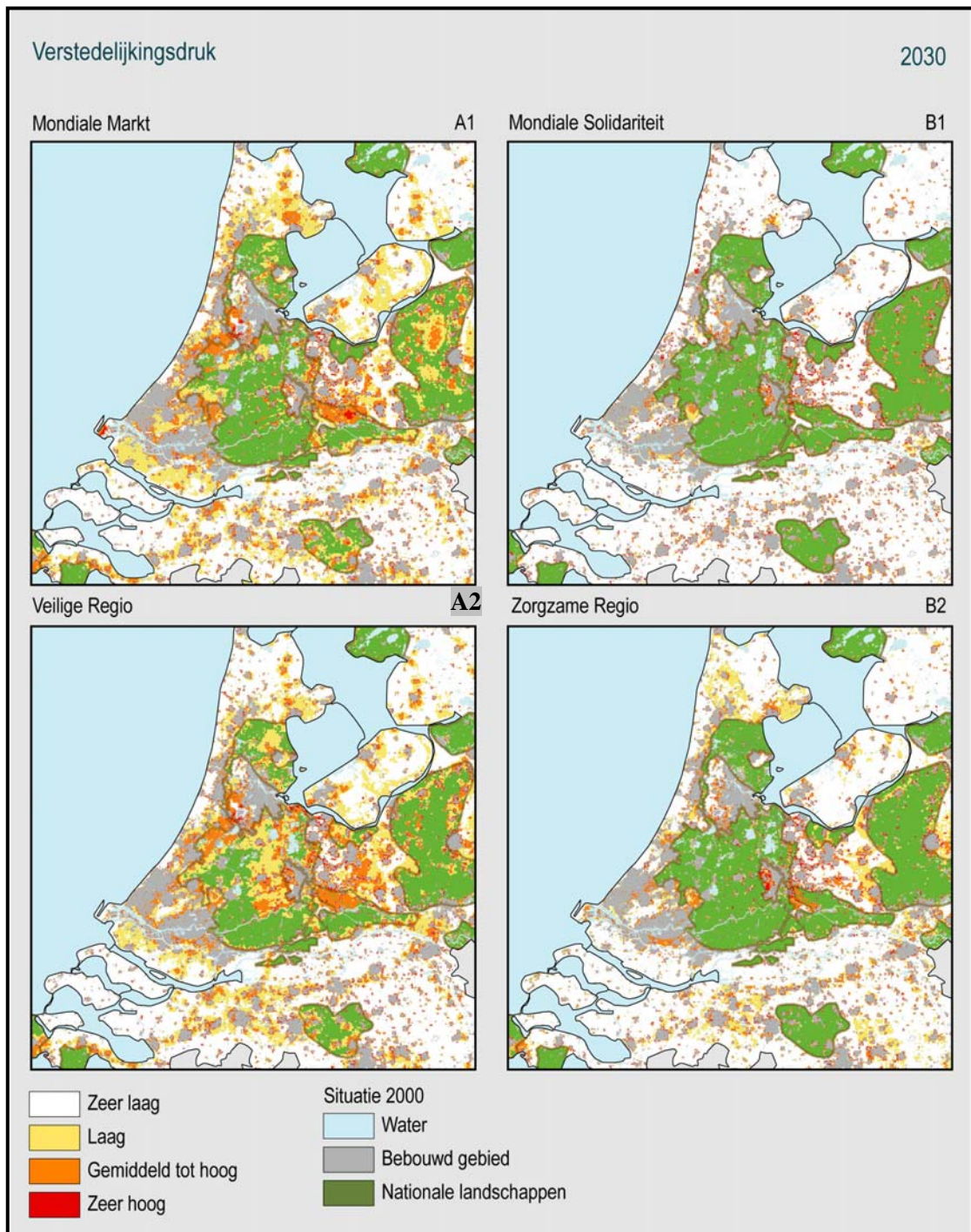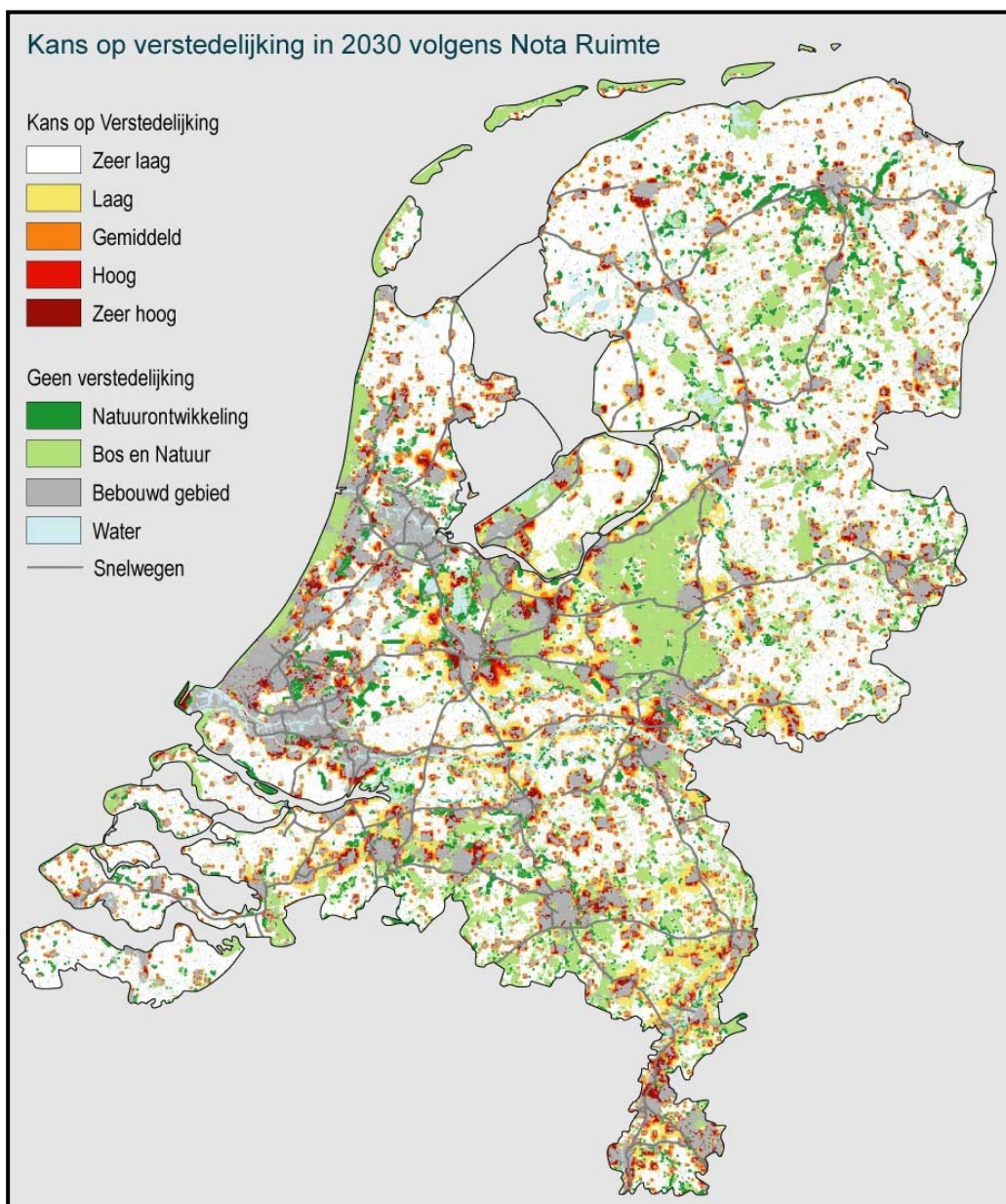
Figure 18   *Trend towards urbanisation in national landscapes calculated for four scenarios: A1 (Global Market, top left), B1 (Global Solidarity, top right), A2 (Safe Region, bottom left) and B2 (Caring Region, bottom right).*

Legend: white= very low urbanisation pressure, yellow= low urbanisation pressure, orange= average to high urbanisation pressure, and red= very high urbanisation pressure. The A1, A2, B1 and B2 scenarios are consistent with assumptions made by the IPCC for global developments through to 2100. Sources: [5, Figure 4.25] and [28, Figure 4.5.2].

Figure 19   *Probability of urbanisation according to the Spatial Policy Document (white= very
low chances, yellow= low chances, orange= average chances, red= high chances,
dark red= very high chances).*
This map is based on 1000 ensemble land-use maps in which the categories housing,
industry/commerce, offices, social and cultural facilities, recreation and greenhouses
were combined to make a new category, "urbanisation". The probability for a grid
cell (i,j) is determined by counting the number of times the cell is included in the
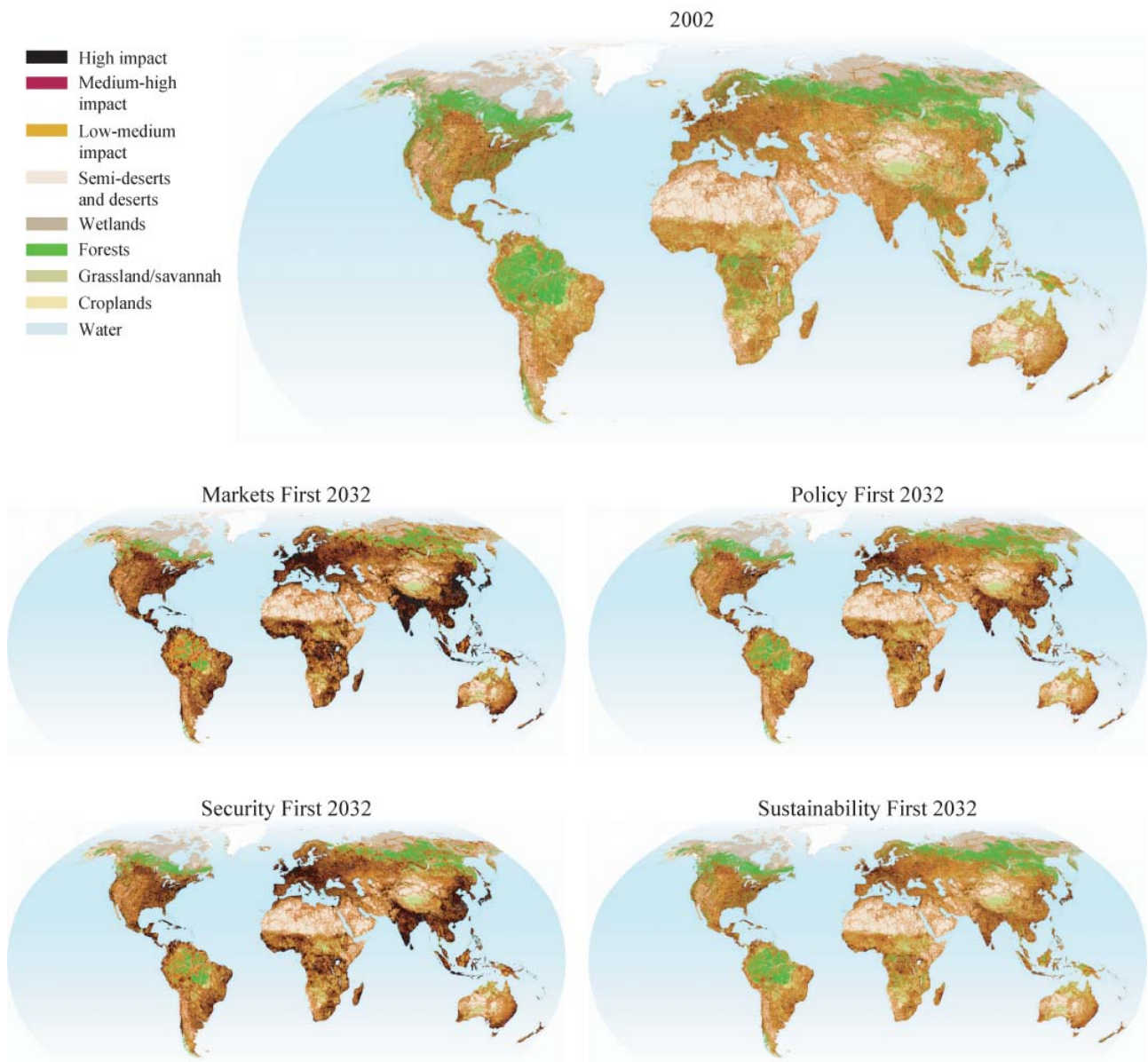category urbanisation, and then dividing the outcome by 1000.
Sources: [21, Figure 10] and [19, Figure 2.3].

4.2.3   Maps with hybrid scales

Occasionally, uncertainty is represented on maps which include data with both interval and ratio scales, and data with a nominal scale. Figure 19 is a good example, as it depicts both probability (ratio scale) and information about different categories of nature (nominal scale), resulting in a hybrid map. The same holds for Figure 18, which depicts urbanisation pressure (ratio   scale)   and   information   land   use   (blue=water,   grey=urbanised   area, dark green=national landscapes).

Figure 20 [24] provides another example of hybrid maps. This set of five maps combines the pressure due to infrastructure developments (ordinal scale) with soil conditions (nominal scale). On the whole, these types of maps do not communicate particularly effectively.

Combining scales makes maps less clear, as the interpretation becomes conceptually more difficult. Hence it is better to avoid composite scales.

Figure 20   *Ecosystems under pressure from infrastructure developments in 2002 (map at the*
                  top) and in 2032 on the basis of four scenarios (Markets First, Policy First, Security
                  First and Sustainability First).
                  The maps include information with ratio (impact), and nominal scales (soil type).
                  Source: [24, Figure 4.1].

4.2.4   Maps with averages and uncertainty combined

In many map-uncertainty presentations average estimates and corresponding uncertainty information are presented in *one map*. However, such presentations have the same danger as the presentation of hybrid maps, as discussed in the preceding section: simply too much information in one map, possibly leading to misunderstanding of "the message". This is particularly the case for dynamic presentations on the Internet, e.g. by hatching or colour saturation methods to make areas with a higher uncertainty less transparent).

Figure 21 shows a simple example where uncertainty is presented by varying gap widths in isolines (average values are presented here in "black"). The example is not very informative. In Figure 10 isolines are applied using numbering next to the contours. Even two types of isolines are used in one graph: isolines for standard deviations (green lines) and isolines for the range (dark blue lines). Two graphs are shown to make the uncertainty information complete: one for the SRES A2 scenario and one for the SRES B2 scenario. We do not think that this presentation will communicate the message due to the overload of information (average values, scenario uncertainty and two forms of grid uncertainty).

Figure 21   Spatial uncertainty is encoded as gaps in contour lines. The more uncertainty, the larger the gaps [22].

Figure 22 shows another example (although one could argue that mean values and standard deviations are presented as *two* separate graphs). Here, standard deviations are presented as 3-D grid uncertainties, and spikes have been coloured according to inter-quartile ranges.

In general we are of the opinion that 3-D presentation of whatever data one has will lead to less informative presentations, despite the fact the 3-D plots look "impressive" (it is very difficult to show "hills" and "dips" without hiding the dynamics lying behind). Thus, we would not advise a presentation similar to the upper plane in the first place. As for the *combination* of the two planes, we find it difficult to project uncertainty information from the upper plane to the right area in the lower plane. In conclusion, the presentation in Figure 22 will not communicate well enough according to us (despite the creative way of presenting this high dimensional data).
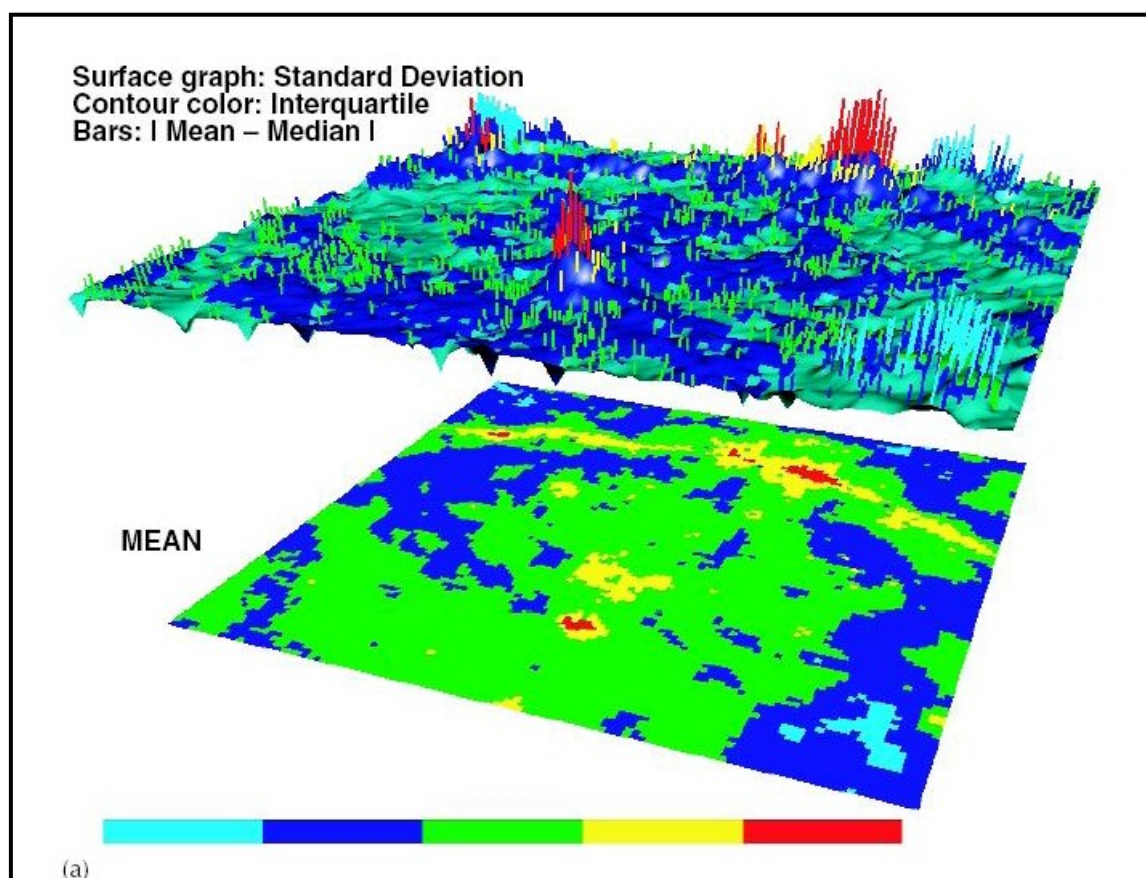

Figure 22   *Presentation of mean values (lower plane) and standard deviations (upper plane).*
                Source: [22].

References

[1]   Bartels, C.J., Veen, A.A. van der, Ormeling, F.J., Koop, R.O., 1994. Kartografisch handboek (Cartographic Handbook) . RIVM report 421504006, Bilthoven (in Dutch).

[2]   Bierkens, M.F.P., Finke, P.A., Willigen, P. de, 2000. Upscaling and downscaling methods for environmental research. Kluwer Academic Publishers, Dordrecht.

[3]   Bishop, T.F.A., Minasny, B., McBratney, 2006. Uncertainty analysis for soil-terrain models. Int. J. of GIS 20(2), 117-134.

[4]   Boersma, K.F., Eskes, H.J., Brinksma, E.J., 2004. Error analysis for troposheric NO_2 retrieval from space. J. Geophys. Res. 109, D04311.

[5]   Borsboom-van Beurden, J.A.M. et al., 2005. Ruimtelijke beelden. Visualisatie van een veranderd Nederland in 2030 (Spatial visions. Visualization of changes in the Netherlands). RIVM report 550016003, Bilthoven (in Dutch).

[6]   Burrough, P.A., McDonnell, R.A., 2000. Principles of geographical information systems. Oxford University Press, Oxford.

[7]   Deutsch, C.V., Journel, A.G., 1998. GSLIB: geostatistical software library and user's guide. Oxford University Press, Oxford.

[8]   Hagen-Zanker, A., 2006. User's Guide Map Comparison Kit 3. RIKS publication, Maastricht. Text and software can be downloaded, free of charge, from http://www.riks.nl/mck/ .

[9]   Hengl, T., 2003. Located at http://spatial-analyst.net/vizualization.php (visited March 2006).

[10]   Heuberger, P.S.C., 2004. USATOOL 1.0, a MATLAB based software package for quantitative uncertainty and sensitivity analysis. IMP/MNP report M005/04, Bilthoven.

[11]   Heuvelink, G.B.M., 1998. Uncertainty analysis in environmental modelling under a change of spatial scale. Nutrient Cycling in Agroecosystems 50, 255-264.

[12]   Heuvelink, G.B.M., 1998. Error propagation in environmental modelling with GIS. Taylor & Francis, London.

[13]   Heuvelink, G.B.M., Pebesma, E.J., 1999. Spatial aggregation and soil process modelling. Geoderma 89, 47-65.

[14]   IPCC, 2001 (eds. J.T. Houghton, Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell, C.A. Johnson). Climate Change 2001, The Scientific Basis. Contribution of Working Group I to the third Assessment Report of the IPCC. Cambridge University Press, Cambridge, UK.

[15]   Koelemeijer, R.B.A., Backes, Ch.W., Blom, W.F., Bouwman, A.A., Hammingh, P., 2005. Consequenties van de EU-luchtkwaliteitsrichlijnen voor ruimtelijke ontwikkelingsplannen in verschillende EU-landen (Consequences of EU air quality directives for spatial development plans in various EU countries). MNP report 500052001, Bilthoven (in Dutch).

[16]   Kaluzny, P.,Vega, S.C., Cardoso, T.P., Shelly, A.A., 1998. S+, Spatial Stats. MathSoft publication, New York.

[17]   Kassteele, J. van de, 2006. Statistical air quality mapping. PhD thesis, Wageningen University, Wageningen.

[18]   KNMI, 2005. See www.temis.nl/airpollution/no2.html (visited March 2006).

[19]   Kuiper, R., De Niet, R. (eds.), 2004. Milieu- en natuureffecten Nota Ruimte (Environmental and nature effects of the Spatial Policy Document). RIVM report 711931009, Bilthoven (in Dutch).

[20]   Matthijsen, J., Visser, H., 2006. Fijn stof in Nederland. Rekenmethodiek, concentraties en onzekerheden (Fine particles in the Netherlands. Models, concentrations and uncertainties). MNP report , Bilthoven (in press, in Dutch).

[21]   Nijs, A.C.M. de, Kuiper, R., Crommentuijn, L.E.M., 2005. Het landgebruik in 2030. Een projectie van de Nota Ruimte (Land use in 2030. A projection for the 'Nota ruimte'). RIVM/MNP report 711931010, Bilthoven (in Dutch).

[22]   Pang, A., 2001. Visualizing uncertainty in geo-spatial data. See http://www7.nationalacademies.org/CSTB/wp_geo_pang.pdf (visited March 2006).

[23]   Pebesma, E.J., DeKwaadsteniet, J.W., 1997. Mapping groundwater quality in the Netherlands. J. of Hydrology 200(1-4), 364-386.

[24]   Potting, J., Bakkes, J., 2004. GEO-3 scenarios 2002-2032. Quantification and analysis of environmental impacts. UNEP/DEWA/RS.03-4 and RIVM report 402001022, Bilthoven.

[25]   RIKS, 2004. Kalibratie van de LeefOmgevingsVerkenner (Calibration of the Environmental Explorer). Final report, July 7, Maastricht (in Dutch).

[26]   RIVM/MNP (2003). RIVM/MNP Leidraad voor omgaan met onzekerheden. National Institute for Public Health and the Environment (RIVM), Bilthoven [Dutch edition of Petersen *et al.* (2003) and Janssen *et al.* (2003); the other two parts (Van der Sluijs *et al.*, 2003, 2004) are only available in English, see www.mnp.nl/leidraad/ and www.nusap.net].

[27]   RIVM/MNP, 2004a. MilieuCompendium (Environmental Compendium). Sdu uitgevers (in Dutch).

[28]   RIVM/MNP, 2004b. Kwaliteit en toekomst. Verkenning van duurzaamheid (Quality and the future. Exploring durability). Sdu uitgevers (RIVM/MNP report 500013009) (in Dutch).

[29]   Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M., 2004. Sensitivity analysis in practice. A guide to assessing scientific models. Wiley & Sons, Ltd, Chichester.

[30]   Shi, W., Fisher, P.F., Goodchild, M.F. (eds.), 2002. Spatial data quality. Taylor & Francis, London.

[31]   Sluijs, J.P. van der, Janssen, P.H.M., Petersen, A.C., Kloprogge, P. Risbey, J.S., Tuinstra, W., Ravetz, J.R., 2004. RIVM/MNP Guidance for uncertainty assessment and communication: Tool catalogue for uncertainty assessment. Utrecht University, Utecht [see www.nusap.net].

[32]   Velders, J.M. *et al*., 2001. Global tropospheric NO_2 column distributions: comparing three-dimensional model calculations with GOME measurements. J. of Geophys. Res. 1006 (D12), 12643-12660.

[33]   Velders, J.M., Aben, J.M.M., Beck, J.P., Blom, W.F., Hoen, A., Jimmink, B.A., Matthijsen, J., Ruiter, J.F. de, Smeets, W.L.M., Velze, K. van, Visser, H., Vries, W.J. de, Wieringa, K., 2006. Grootschalige concentratiekaarten voor luchtverontreiniging in Nederland (Large-scale air quality concentration maps in the Netherlands). MNP report 500093002 (in Dutch), Bilthoven.

[34]   Visser, H. (ed.),   2004. The Map Comparison Kit: methods, software and applications.   RIVM report 550002005, Bilthoven.

[35]   Visser, H. , Nijs, T. de, 2006. The Map Comparison Kit. Environmental Modelling & Software 21, 346-358.

*Figure 23   There could be an explanation for people's tendency to accept maps as "reality", once they are drawn and printed. We all learn geography at school through maps such as those shown for the Netherlands for the years 1950 and 1980. These maps are very accurate and positional errors do not play any role in the practical use. Therefore, a map such as shown here for the year **2010** is easily accepted as "real". Nobody wonders about uncertainties here. And the legend underneath does not contain any information on the land-use model applied and corresponding uncertainties. Photograph: H. Visser.*