DOI: 10.21513/2410-8758-2019-2-103-119 УДК 502/504:327.1+551.583

CHARACTERIZATION OF UNCERTAINTIES IN THE ASSESSMENTS OF THE INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE

S.M. Semenov [1,2)*], *G.E. Insarov* [1,2)], *C.L. Méndez* [3)]

[1)]Yu. A. Izrael Institute of Global Climate and Ecology,
20B, Glebovskaya Str., 107258, Moscow, Russian Federatoin; *SergeySemenov1@yandex.ru*

[2)]Institute of Geography of the Russian Academy of Sciences,
29, Staromonetny Lane, 109017, Moscow, Russian Federation; *insarovg@gmail.com*

[3)]Venezuelan Institute for Scientific Research,
1020-A, Caracas, 21827, Venezuela; *carlos.menvall@gmail.com*

Abstract. Correct use of scientific assessment results in decision-making processes requires information about their uncertainties. The evolution of approaches for assessing the uncertainty of statements (findings, conclusions) of Intergovernmental Panel on Climate Change (IPCC) reports is considered. Recommendations of the IPCC guidance notes of 2000 and 2010 on characterizing the uncertainties are presented. Two approaches for improving uncertainty treatment are proposed. The first pertains to analyzing independence of quantitative data used in the estimation of uncertainty in results of their synthesis. The second pertains to objective assessment of the degree of certainty of qualitative statements.

Keywords. Climate change, consequences, quantitative assessment, qualitative assessment, uncertainties, IPCC.

Introduction

The Intergovernmental Panel on Climate Change (IPCC) was established in 1987-1988 jointly by the World Meteorological Organization (WMO) and the United Nations Environment Program (UNEP). The decision was endorsed later by the UN General Assembly resolution. The task of this group is to periodically assess scientific knowledge about observed and projected changes in the Earth's climate, their consequences for natural and socio-economic systems, options for the mitigation of anthropogenic impacts on the climate, as well as to provide methodologies for inventories of anthropogenic greenhouse gas emissions. The results of IPCC assessments are used in the development of national and international policies to limit anthropogenic impacts on the climate system, as well as at various levels of governance and by businesses and citizens. The group does not conduct research, but rather summarizes findings published in scientific literature.

*Corresponding author

The main goal of IPCC reports is to provide generalized conclusions needed for making decisions that pertain to the climate or relate to adaptations to its changes, and do so in a politically neutral manner, without prescribing specific decisions.

In the formulation of its statements, IPCC experts use scientific publications as a source of data and information. The creation of each generalized statement is a small focused study conducted by IPCC report authors. Since the conclusions are intended for critical decision making, the degree of confidence for the statements should be determined. This determination is intended to provide readers/users with the opportunity to track the data and information on which the conclusion is based, as well as explain how the statement and its degree of confidence was constructed, to ensure the traceability and transparency.

The IPCC reports are a product of the joint work of an international team of experts. At the last stage of the work, the most important statements that are included in the Summary for Policymakers are approved by a consensus of national delegations of IPCC Members. Therefore, it is necessary to have a fairly unified approach to the assessment of knowledge, formulation of key findings, and determination of degree of their confidence. Such approaches are presented in special guidance notes of the IPCC and the methodological publications by its leading scientists. With regard to uncertainty treatment, a detailed history can be found in (Mastrandrea, Mach, 2011).

The objectives of this paper are to briefly present the key points of those approaches and to identify some needs and opportunities for further refinement.

The problem of uncertainty

Estimates of Earth system parameters characterizing its current state, as well as observed and projected changes, always have some degree of uncertainty. Sources of this uncertainty are numerous: variability of measured variables in space and time, sampling errors in field observational data, gaps in observational data series, errors of measuring equipment, a choice of model structure, a choice of model parametrizations, errors of model coefficients, insufficiency of fundamental knowledge about natural processes, and others. These uncertainties are even more pronounced in assessments of climate change impacts on socio-economic systems, options for mitigation of anthropogenic impacts on the climate, and options for adaptation to its changes. The extent of uncertainty range reflects current precision of relevant branches of science.

At the very beginning of the IPCC work, the question of uncertainty of assessments was given considerable attention. Conceptual differences emerged between approaches adopted by Working Group I (Climate Change 1995, 1996a) and Working Group II (Climate Change 1995, 1996b).

Working Group I characterized uncertainties of quantitative estimates using an approach drawn from the physical sciences. An estimate is accompanied with statistical characteristics of its variability, namely, with an actual or possible range of variation, standard deviation, etc. This approach assumes the availability of an ensemble of measurement data or analysis results for the variable (for example,

CO_2 concentration, air temperature, sea level). The parameters of statistical variability are estimated on the basis of such data.

Working Group II often cannot utilize this approach to uncertainty assessment for two reasons. First, sometimes there is no ensemble of measurement or analysis results in the scientific literature, with only indirect historical data on the effects of climate change available. One such example was, until recently, the effect of global warming on the Greenland ice sheet (Alley et al., 2009). Second, information presented in literature may be (1) qualitative or (2) quantitative, but highly heterogeneous and generalizable only in qualitative form due to insufficient elaboration of the model analysis tools. One example is climate change impact assessments on the achievement of sustainable development goals (see Sanchez-Rodriguez, Ürge-Vorsatz, Barau, 2018).

Although in the cycle of the IPCC Second Assessment Report (published in 1996) main conclusions were not yet accompanied by degree of confidence information in the Summary for Policymakers, this information already appeared in the main text and chapter summaries. Moreover, in the preface to the contribution of Working Group II, a special section entitled 'Levels of Confidence' detailed three gradations of degree of confidence for findings: high, medium and low. The following criteria were recommended (Climate Change 1995, 1996b):

• <u>High Confidence</u> — *This category denotes wide agreement, based on multiple findings through multiple lines of investigation. In other words, there was a high degree of consensus among the authors based on the existence of substantial evidence in support of the conclusion.*

• <u>Medium Confidence</u> — *This category indicates that there is a consensus, but not a strong one, in support of the conclusion. This ranking could be applied to a situation in which an hypothesis or conclusion is supported by a fair amount of information, but not a sufficient amount to convince all participating authors, or where other less plausible hypotheses cannot yet be completely ruled out.*

• <u>Low Confidence</u> — *This category is reserved for cases when lead authors were highly uncertain about a particular conclusion. This uncertainty could be a reflection of a lack of consensus or the existence of serious competing hypotheses, each with adherents and evidence to support their positions. Alternatively, this ranking could result from the existence of extremely limited information to support an initial plausible idea or hypothesis.*

This was the first attempt in the IPCC to formalize the determination of a degree of confidence for a statement. Its authors stated that this system and the procedure used to arrive at the certainty determination is just a tool for informing decision-makers about certainty of the report's main conclusions. Importantly, it is an imperfect tool, and the whole procedure is subjective: it strongly reflects expert judgments, and another group of experts may give a different uncertainty estimate based on the same information (Climate Change 1995, 1996b).

During subsequent years, the IPCC repeatedly returned to the uncertainty problem, trying to reduce the degree of subjectivity and the impact of expert judgments in the procedure of estimating uncertainty, to reconcile the approaches employed by Working Groups I and II.

First IPCC guidance document on assessing uncertainty

In the cycle of the Third Assessment Report of the IPCC (TAR, released in 2001), a number of guidance documents were prepared for the report's lead authors including 'Uncertainties in the IPCC TAR: Recommendations to Lead Authors for More Consistent Assessment and Reporting' (Moss, Schneider, 2000). The paper is summarized in this section.

The publication clearly states that the task of authors of the assessment reports is not only to present generalized statements on key topics, but also to accompany them with an assessment of degree of credibility, i.e., of confidence (if possible in quantitative form), based on available research. At the same time, readers, i.e., users of reports, tend to self-assign a degree of confidence to assessment outcomes. Because of this, it is advisable for experts to specify the uncertainty, even qualitatively, if quantitative specification is impossible.

As a theoretical basis for such an approach, R. Moss and S. Schneider (2000) proposed the Bayesian (or subjective) concept of probability. According to this concept, the probability of an event is understood as the degree of confidence (of researchers) in the occurrence of an event[1)]. In this context, when assessing the probability distribution for a system of alternative events, *a priori* probabilities are first assigned, based on the available evidence. The additional evidence obtained in the course of the assessment makes it possible to revise previously assigned probabilities and assign new, improved, *a posteriori* probabilities. As declared in the cited work, science must strive to ensure that theoretical concepts are substantiated, verified by empirical data. However, applied science, 'science for decision-making', often cannot wait for the completion of detailed scientific research, since a decision must be taken within a certain timeframe and under conditions of incomplete information. In this case, degree of uncertainty of scientific conclusions should be available for substantiation of the decision.

Of course, in such a lengthy document as an IPCC assessment report (a typical volume contains 500-1000 pages) it is impossible to accompany each figure and each statement with an estimate of uncertainty. This is practiced in relation to the main statements in the chapter summaries, as well as the Summary for Policymakers and Technical Summary. At the same time, the authors are advised to avoid fuzzy or overly broad statements, for example, "biodiversity may change with warming". This undoubtedly true, but trivial, statement does not bring new information until the level of warming and corresponding level of biodiversity change are indicated.

The guidance document (Moss, Schneider, 2000) recommends sequential steps for assessing level of confidence for any statement. They can be summarized as

[1)] Such an understanding of probability differs from traditional one that is the frequency of a given event in a series of repetitions (playouts) of a single-type experiment. The frequency approach is sometimes not feasible, for example, for unique or very rare events (for example, for the collapse of Meridional Overturn Circulation).

follows: *(1) Identify the main factors and uncertainties that affect the finding. (2) Document the ranges that characterize them, in accordance with the data of scientific publications. (3) Based on the nature of uncertainties and the state of knowledge, determine a type of assessment of a degree of confidence, namely, quantitative or qualitative. (4) Quantitatively or qualitatively characterize the probabilistic distribution of a parameter, variable or result. (5) Prepare a description of obtaining a probability distribution, allowing the reader to track the process*. Moss and Schneider (2000) categorized the main sources of uncertainty using the following typology: problems with data, problems with models, and other sources of uncertainty. The details are given in Table 1.

Table 1. Sources of uncertainty (Moss, Schneider 2000)

Source of uncertainty		
Problems with data	***Problems with models***	***Other sources of uncertainty***
Missing components or errors in the data; *'Noise' in the data associated with biased or incomplete observations;* *Random sampling error and biases (non-representativeness) in a sample.*	*Known processes but unknown functional relationships or errors in the structure of the model;* *Known structure but unknown or erroneous values of some important parameters;* *Known historical data and model structure, but reasons to believe parameters or model structure will change over time;* *Uncertainty regarding the predictability (e.g., chaotic or stochastic behavior) of the system or effect;* *Uncertainties introduced by approximation techniques used to solve a set of equations that characterize the model.*	*Ambiguously defined concepts and terminology;* *Inappropriate spatial/temporal units;* *Inappropriateness of/lack of confidence in underlying assumptions;* *Uncertainty due to projections of human behavior (e.g., future consumption patterns, or technological change), which is distinct from uncertainty due to "natural" sources (e.g., climate sensitivity, chaos).*

Since the IPCC as a whole performs a full cycle of assessments related to anthropogenic impacts on the Earth's climate system, it is important to take into account that, as it moves along the chain from anthropogenic emissions of greenhouse gases to their effects on natural and socio-economic systems, uncertainty will accumulate and grow. R. Moss and S. Schneider (2000) called it the 'cascade of uncertainty' or the 'uncertainty explosion' – see Fig. 1.

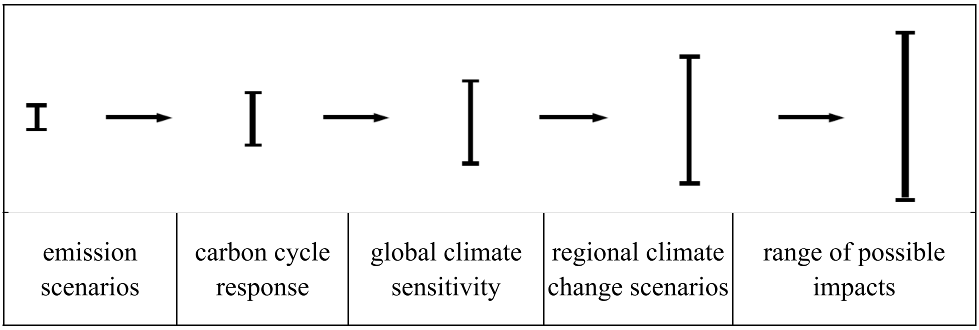

Figure 1. An increase in the range of uncertainties in the chain of estimates from anthropogenic greenhouse gas emissions to possible consequences of physical, economic, social and political nature, including response strategies (Moss, Schneider, 2000)

In addition, terminology can affect perceptions of uncertainty. To characterize a degree of certainty, confidence of a statement, researchers often use a variety of common expressions: 'almost certain', 'probably', 'likely', 'possibly', 'unlikely', 'not possibly', 'doubtfully', etc. Experts imply different meanings in these terms – even for the same assessment tasks. To unify the terminology, R. Moss and S. Schneider (2000) proposed a universal discrete quantitative scale for degree of confidence: from 0.95 to 1.00 = very high confidence; from 0.67 to 0.95 = high confidence; from 0.33 to 0.67 = medium confidence; from 0.05 to 0.33 = low confidence; from 0.00 to 0.05 = very low confidence.

To estimate the amount of knowledge related to a statement, they proposed a system of qualitative terms using two basic characteristics: evidence and agreement (see Tabl. 2). The system aims to describe, using qualitative terms, the sum total of the knowledge on the basis of which a degree of confidence is assigned.

Table 2. Characteristics of the sum of knowledge in qualitative terms (Moss, Schneider, 2000)

Level of agreement/ consensus	*High*	Established but incomplete	Well established
	Low	Speculative	Competing explanations[2)]
		Low	*High*
		Amount of evidence (observations, model output, theory, etc.)	

The guidance note prepared by Moss, Schneider (2000) was employed in the preparation of the Third and Fourth Assessment Reports of the IPCC, issued in 2001 and 2007, respectively. Yet, its use in the contributions of Working Groups I, II and III was uneven and incomplete (Mach et al., 2017). However, even sporadic and subjective use of the approach introduced IPCC experts to the culture of probabilistic assessment of statements (findings, conclusions) reliabilities.

[2)] For example, when describing a phenomenon about half of the evidence are in favor of its anthropogenic origin, and the rest is in favor of natural origin.

Second IPCC guidance document on assessing uncertainties

In the Fifth IPCC Assessment Report cycle, methodological approaches to estimating uncertainties were further developed. On July 6-7, 2010, an IPCC expert meeting in Jasper Ridge, California, USA, resulted in a brief updated version of the guidance document on assessing uncertainties, and a year later, the full journal version was published (Mastrandrea et al., 2010, 2011). This guidance document inherited the philosophy of the previous one (Moss, Schneider, 2000), but was written in simpler scientific language and more focused on the procedure for determining the degree of statement confidence. This section contains some of the main recommendations of this guidance document prepared for the authors of the IPCC Fifth Assessment Report.

Uncertainty (or certainty) is assessed for a statement (finding, conclusion) of the report. The assessment of uncertainty is possible using either qualitative or quantitative terms, depending on the basic information which is generalized. Three metrics are proposed for this purpose: ‘evidence’, ‘agreement’ (to be integrated further into ‘confidence’) and ‘likelihood’ (Mastrandrea et al., 2010; Mastrandrea, Mach, 2011).

If the usual statistical analysis of observational data, experiments, and simulation results or expert judgments is possible, then quantitative probabilistic terms are employed using the likelihood metrics, which are preferable.

Otherwise, the assessment is conducted using qualitative metrics. Namely, confidence (degree of trustworthiness) of a statement (finding, conclusion) is understood as its validity determined by the type, amount, quality and consistency of the evidence, as well as the degree of agreement. ‘Evidence’ is the whole body of facts, data, information or knowledge pertaining to a statement derived from documents (papers, books, etc.) or resulting from expert judgment, commonly referred to as ‘pieces of evidence’. ‘Agreement’ is built on the consistency of the evidence, covering the diversity of competing (or not) explanations or models for phenomenon. The procedure proposed in (Mastrandrea et al., 2011) is displayed in Fig. 2.

The guidance note recommended three gradations for specifying the ‘amount of evidence’: ‘limited’, ‘medium’ and ‘robust’; as well as three gradations for ‘agreement’: ‘low’, ‘medium’, and ‘high’. These specifications of statements are chosen by IPCC authors using their expert knowledge. ‘Confidence’[3] synthesizes ‘evidence’ and ‘agreement’; their relationship is illustrated by Fig. 3.

The guidance note recommends assessing ‘confidence’ using five grades: ‘very low’, ‘low’, ‘medium’, ‘high’ and ‘very high’. If it is not possible to assign a ‘confidence’ grade, uncertainty of the finding can be reported solely with evidence and agreement metrics.

[3] In this context, its meaning differs from the traditional, statistical one. This is a qualitative term corresponding to ‘trustworthiness’.

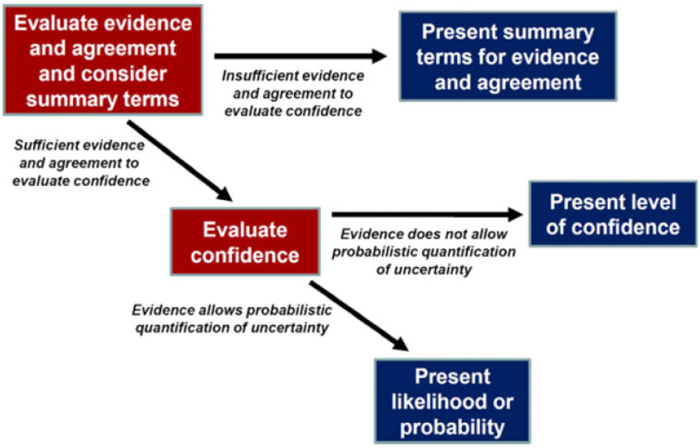

Figure 2. The general procedure for assessing the uncertainty (certainty) of a statement (finding, conclusion) based on a set of available evidence (Mastrandrea et al., 2011)

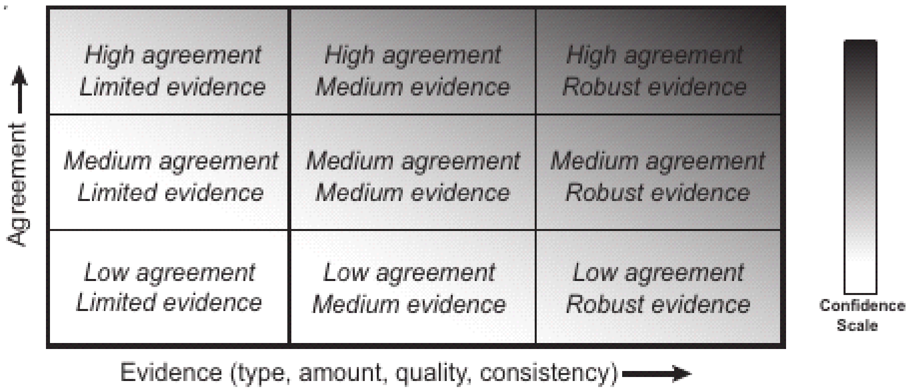

Figure 3. Relationship between evidence and agreement and the resulting confidence scale (Mastrandrea et al., 2010)

The use of quantitative terms in the procedure for assessing uncertainty displayed in Fig. 2, as likelihood or probability, is within the framework of usual statistics and does not require special attention. However, we offer comments on the use of qualitative uncertainty assessment terms.

First, only results of research published in specialized scientific journals should be considered as evidence, in order to ensure scientific quality.

Second, independence of evidence should be explored. For example, consider results of laboratory tests on the effect of temperature on plant growth presented in two publications: one publication only looks at one plant species, and a second publication examines ten. It is not correct to combine these pieces of evidence into a set of 11 without proper analysis. If the laboratory facilities used for the experiments were different, then, in the second case, the results may be interdependent, i.e., may have some common bias. In this case, strictly speaking, there are only two completely independent pieces of evidence, not 11.

Third, the evidence reviewed must use sufficiently consistent methodologies (i.e., field observations, laboratory experiments, model calculations, or expert judgments). Combining diverse pieces of evidence in one set is acceptable with great caution only. The type of evidence should be indicated clearly.

Authors of the IPCC reports are advised to present the material in such a way that a reader can follow the process of evidence assessment and agreement to ensure a traceable account.

If a quantitative characterization of uncertainty (certainty) is possible, the guidance note recommends the use of ordinary statistical characteristics, viz. probability or likelihood (these terms are used almost as synonyms). This applies to those cases where the occurrence of a single event is described by statistical distribution. For example, when it comes to pertaining a climate parameter, an observed trend or an expected change to a certain range of values, quantitative approach is applicable. Table 3 shows the corresponding discrete scale for probability (likelihood).

Table 3. Probabilistic scale for quantifying the certainty of a statement (Mastrandrea et al., 2010)[4)]

Term	Likelihood, %
Virtually certain	99-100
Very likely	90-100
Likely	66-100
About as likely as not	33-66
Unlikely	0-33
Very unlikely	0-10
Exceptionally unlikely	0-1

Assignment of a category listed in Table 3 to a given statement is carried out through usual statistical (frequency) analysis of the ensemble of quantitative data. The sources of such data can be observations, calculations, or expert estimates.

The IPCC 2010 guidance note on the uncertainty treatment (Mastrandrea et al., 2010) and further publications (Mastrandrea, Mach, 2011; Mastrandrea et al., 2011; Mach et al., 2017) clarified the relations and distinctions among uncertainty metrics and facilitated more consistent use of them throughout the IPCC Working Groups than in previous assessment cycles. The guidance note was used extensively by authors of the Fifth Assessment Report of the IPCC and other reports prepared in this cycle, and is still in use in the Six Assessment cycle. It contributes to a clearer characterization of uncertainties of the statements (conclusions and findings) presented for the use in the international climate negotiation process under the

4) When choosing from the upper part of the table, it is recommended to adopt the category with the largest possible lower bound. For example, if the probability (likelihood) of an event is 95%, one should select the category ‘very likely’, not ‘likely’.

auspices of the UN Framework Convention on Climate Change, as well as for the use of international scientific community and relevant international organizations. For example, the approach is adopted by the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) in its own Guide on The Production of Assessments (IPBES, 2018).

However, the necessity for further elaboration of more comparable approaches to uncertainty treatment in different IPCC Working Groups still exists. In particular, this relates to the objectivity of qualitative characterizations of uncertainties and to the independence of assessed data. The latter is especially important for complex systems (Gerlach, Altmann 2019), like the Earth's climate system.

Further development

In 2016, the IPCC Sixth Assessment Report cycle began, and the Report will be released in 2021-2022. Three special reports and one methodological report are to be prepared in this cycle:

− Global Warming of 1.5°C: An IPCC special report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty (released in 2018);

− Climate Change and Land: An IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems (the release is scheduled for 2019);

− Ocean and the Cryosphere in a Changing Climate (the release is scheduled for 2019);

− 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories (adopted in May of 2019).

The guidance note (Mastrandrea et al., 2010) is still in use for the characterization of uncertainties in the preparation of the IPCC Sixth Assessment Report and all reports of its cycle. The IPCC reports scheduled for release in 2019 will use this guidance note in its original form, since the reports obviously cannot await any updated versions. However, before the release of the main, Sixth Assessment Report, there is still enough time to refine the procedures for assessing uncertainties, considering the necessity of improvements mentioned above. Although the current IPCC guidance note (Mastrandrea et al., 2010) emphasizes the necessity to explore the independence of pieces of evidence, there are no methodological recommendations for how to do this. The guidance note also remarks on the importance of type, amount, quality and consistency of evidence. However, no clear indication of how to assess and integrate them is given. In particular, the meaning of 'sufficient amount of evidence' remains vague. This leaves substantial space for subjectivity in the assessment process. Below, we will touch on two problems that need to be resolved: the independence of pieces of evidence and the objectivity of the determination of sufficient amount of evidence.

Independence of pieces of evidence

Initial units of information used in the construction of a summary estimate are not always independent. For example, when estimating the future mean global temperature in the near-surface layer, the results of several models are usually involved (M is their total number). Several (n_m) runs of the m-th model are performed taking into account the random processes embedded in the climate model. Thus, there are $N = (n_1 + n_2 + ... + n_M)$ calculation results. Can they always be considered as independent pieces of evidence?

In general, the answer to this question is negative. Each model has its own parameters (for example, climate sensitivity) whose values affect results of the calculation. One model may somewhat overestimate the results of the calculations (e.g., of a certain variable of the climate system for the end of the 21st century), while the other one may underestimate. In fact, there are M independent pieces of evidence in this example, not N. Correct generalization of the results of such calculations requires preliminary analysis of variance: its partitioning into intra-model and inter-model components (see Fig. 4).

In Fig. 4, Z is the subject of estimation through summarizing the calculation results $\{X_{ij}\}$. The calculation results are of the following origin:

$$X_{ij} = Z + \xi_i + \eta_{ij}, \quad i = 1, 2, ..., M; j = 1,2, ..., n_i .$$

Here ξ_i and η_{ij} are realizations of centered random variables ξ and η having variances D and V, respectively; $Y_i = Z + \xi_i$, $i = 1, 2. ..., M$.

The inter-model variance D and the intra-model variance V can be estimated in different ways. For example, assuming that the distribution of random variables ξ and η is normal, the estimates can be obtained with the maximum likelihood method. This is performed through maximizing the likelihood function quantifying the probability of obtaining $\{X_{ij}\}$ as a set of calculation results.

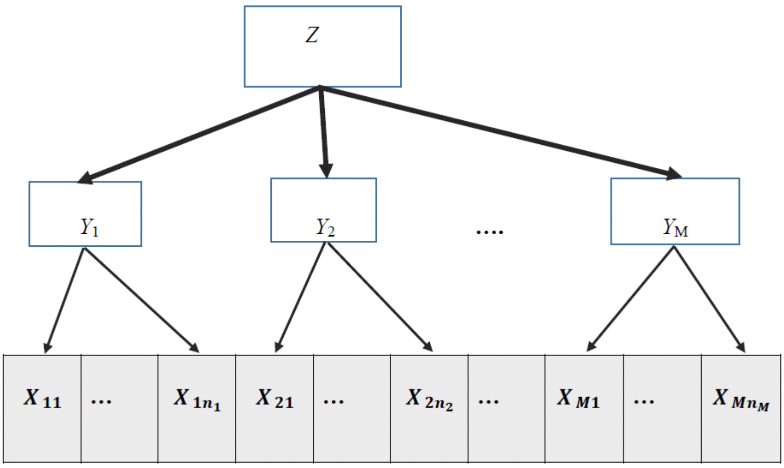

Figure 4. Illustration of the process of generating the results of model calculations to be summarized

Knowing the variances D and V, one can correctly estimate the value of Z and characterize its uncertainty as follows. Estimates of the intra-model mathematical expectations[5)] Y_i, $i = 1, 2, ..., M$, are obtained using simple averaging:

$$\hat{Y}_i = \left(\sum_{j=1}^{n_i} X_{ij} \right) / n_i .$$

These values approximate Z, but, generally speaking, with different accuracy. The variances of their deviations from Z are, respectively, equal to $V_i = D + V / n_i$, $i = 1, 2. ..., M$. Further, an effective estimate of Z can be obtained by summing the partial estimates with the weights:

$$\hat{Z} = \frac{\left(\sum_{i=1}^{M} \hat{Y}_i / V_i \right)}{\left(\sum_{i=1}^{M} 1 / V_i \right)} .$$

Its standard deviation from Z is $\sigma = \left(\sum_{i=1}^{M} 1 / V_i \right)^{-0.5}$.

When presenting the results of assessments, Working Group I sometimes separately presents characteristics of intra-model and inter-model variability of calculation results to offer the reader a clearer understanding of the degree of uncertainty.

This is also important for, but less pronounced in the work of other IPCC teams. For example, Working Group II often assesses change in plant growth with increasing levels of CO_2 in the atmosphere. Even if the data are sufficiently consistent (e.g., all based on laboratory experiments), the results of different studies can be obtained through different techniques and equipment. Therefore they cannot automatically, without special analysis, be considered independent.

Degree of confidence of qualitative statements

In accordance with the current guidelines (Mastrandrea et al., 2010), IPCC authors should include a confidence judgment when summarizing evidence for a statement (conclusion, finding) in qualitative terms. This should be based on the number of pieces of evidence (as well as their quality and consistency) and their agreement (among themselves; see Fig. 3). More pieces of evidence and stronger agreement among them provide greater confidence. However, no guidance is offered for how to perform the assessment and this is left to a decision by the author teams based on their collective expert judgment.

[5)] Intra-model mathematical expectation is conditional, i.e. it is based on the assumption that the parameters of the simulated processes are exactly as in this computational model.

Such an approach often raises questions: How many pieces of evidence need to be considered by an author team?; What minimum percentage should be in favor of the statement in order to make a valid judgment about its confidence?; Do the answers to these questions depend on the 'importance' of the object affected by climate change, or are they universal? Recommendations on these issues could reduce the subjectivity of uncertainty estimates. Below, we offer one possible approach to the issue.

Consider the following (idealized) situation. There are N pieces of evidence on which a statement is based − e.g., publications, for simplicity. The publication authors conducted research with equal degrees of quality and completely independent of each other. The group of experts who assess the body of evidence consider the results of k publications of the whole set of N as supporting the statement, while the others as not.

Random factors affect the assessment process. Even if the statement is solid, a case study based on the current state of knowledge, experimental technique, model means, etc. may not support it. When analyzing a piece of evidence, a group of experts may interpret it in different ways. Therefore, we assume that a result of the assessment of a piece of the evidence is described by binomial probability distribution: the piece of evidence supports the statement with probability p, and it does not support it with probability $(1 - p)$.

If the evidence relevant to the statement consists of numerous (infinitely many, so to speak) pieces, the value of p could be identified precisely. This is the proportion of pieces of the evidence, which the authors group qualified as supportive. At this point, it would be necessary to adopt a threshold value p_0 for accepting the statement: if p exceeds the threshold, then the statement is accepted. For example, $p_0 = 1/2$ ('simple majority') or $p_0 = 2/3$ ('qualified majority').

In fact, the amount of evidence is never very large. A decision must be made on the basis of a pair of numbers (N, k). The procedure for estimating the parameter p of binomial distribution for a given confidence figure ε can be employed for this. This procedure is described in many textbooks and statistical handbooks (for example, in (Müller et al., 1982)). It considers the case of independent playouts of a binary random variable, when TRUE emerges with probability p, while FALSE occurs with probability $(1 - p)$. To estimate the lower confidence limit of parameter p, one can use the probability $S(k, N, p)$ that of the N independent realizations in k or more cases TRUE occurs:

$$S(k, N, p) = \sum_{i=k}^{N} \binom{N}{i} p^i (1-p)^{(N-i)}.$$

Here $\binom{N}{i} = \frac{N!}{i!(N-i)!}$. For given N, k and ε, $0 < k \leq N$, one can solve the equation:

$$S(k, N, p) = \varepsilon.$$

The solution p_{min} is the lower confidence limit for p corresponding to the confidence figure ε: hypothesis $\{p < p_{min}\}$ is rejected with probability $(1 - \varepsilon)$.

Since larger value of k infers the larger value of p_{min}, the inverse task can be also solved: for a given threshold $p_0 < 1$ the smallest k_0 for which $p_{min} \geq p_0$ can be found. For $p_0 = 0.5$ and $\varepsilon = 0.33, 0.10, 0.01$, when the solution exists, the calculation results are shown in Table 4. The gradations for ε correspond to the categories, respectively, 'likely', 'very likely' and 'virtually certain', see Table 3.

Table 4 shows that poor amount of evidence does not allow to adopt the statement for the given confidence figures. One piece of evidence is always insufficient, and for $\varepsilon = 0.10$ and 0.01 at least 4 and 7 pieces are required, respectively. Of course, if the threshold value p_0 is higher, for example, it shifts from 1/2 ('simple majority') to 2/3 ('qualified majority'), then the requirements regarding minimal amount of evidence are strengthened.

Table 4. Minimal amount of pieces of evidence supporting a statement (k_0) of the whole set (N), which infers that $p > 0.5$ for a given confidence figure ε

N	k_0		
	ε = 0.33	**ε = 0.10**	**ε = 0.01**
2	2	-	-
3	3	-	-
4	3	4	-
5	4	5	-
6	5	6	-
7	5	6	7
8	6	7	8
9	6	7	9
10	7	8	10
11	7	9	10
12	8	9	11
13	8	10	12
14	9	10	12
15	9	11	13
16	10	12	14
17	10	12	14
18	11	13	15
19	11	13	15
20	12	14	16
21	12	14	17
22	13	15	17
23	13	16	18
24	14	16	19
25	15	17	19
26	15	17	20
27	16	18	20
28	16	18	21
29	17	19	22
30	17	20	22

The approach presented above allows to conduct more objective assessments of the confidence of statements based on the total amount of pieces of evidence and the amount of supportive ones. The proposed procedure is universal and in no way connected with nature of the objects and impacts, to which the statement relates. However, the independence of pieces of evidence, their consistency and quality are left to the discretion of the expert group.

Conclusions

In the Second Assessment Report cycle, the IPCC had attempted to assess the uncertainty of its main statements (conclusions and findings), which is highly demanded by users of IPCC reports, especially decision-makers. Except for parameter estimates based on quantitative data (where usual statistical methods are applicable), uncertainty is assessed using qualitative terms. The role of expert judgements and the influence of subjective factors on certainty of assessments are significant. In the IPCC guidance notes of 2000 and 2010, and accompanying publications of lead IPCC scientists, the procedure for assessing uncertainty became clearer and better substantiated. It also pointed out the need for analysis of the evidence in regard to amount, independence, quality, consistency and agreement of evidence. However, in the case of Working Groups II and III, even this procedure for assessing the uncertainty of statements remains subjective in many respects, and relies on expert judgments of authors. There are opportunities for further refinement of this procedure to make it more objective and algorithmic. This paper proposes two directions for improvement of uncertainty assessments: an analysis of the independence of quantitative data included in the assessment, and more objective estimation of the confidence of qualitative statements.

Acknowledgements

Our sincere thanks are due to Irina Feygina who has made very helpful comments and graciously edited the final version of the paper.

References

Müller P., Neuman P., Storm R. 1982. Tables on mathematical statistics. − Moscow, Finance and Statistics, 271 p. (in Russian).

Alley R.B., Andrews J.T., Brigham-Grette J., Clarke G.K.C., Cuffey K.M., Fitzpatrick J.J., Funder S., Marshall S.J., Miller G.H., Mitrovica J.X., Muhs D.R., Otto-Bliesner B., Polyak L. 2009. White JWC History of the Greenland Ice Sheet: Paleoclimatic insights. − Quat. Sci. Rev., vol. 29 (15-16), pp. 1728-1756.

Climate Change 1995. 1996a. The Science of Climate Change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change. Editors J.J. Houghton, L.G. Meiro Filho, B.A. Callander, N. Harris, A. Kattenberg and K. Maskell. Published by the Press Syndicate of the University of Cambridge. 572 p. (ISBN 0 521 564336 Hardback; 0 521 564360 Paperback).

Climate Change 1995. 1996b. Impacts, Adaptations and Mitigation of Climate Change: Scientific-Technical Analyzes. Contribution of Working Group II to the Second Assessment Report of the Intergovernmental Panel on Climate Change. Editors: R.T. Watson, M.C. Zinyowera, R.H. Moss. Published by the Press Syndicate of the University of Cambridge, 879 p. (ISBN 0-521-56431-X Hardback. ISBN 0-521-56437-9 Paperback).

Gerlach M., Altmann E.G. 2019. Testing Statistical Laws in Complex Systems. − Physical Review Letters, vol. 122, pp. 168-301.

IPBES, 2018. IPBES Guide on the production of assessments. Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Bonn, Germany. 44 p. Available at https://www.ipbes.net/guide-production-assessments (accessed 20 May 2014).

Mach K. J., Mastrandrea M.D., Freeman P. T., Field C. B.. 2017. Unleashing expert judgment in assessment. − Global Environmental Change, vol. 44, pp. 1-14. http://dx.doi.org/10.1016/j.gloenvcha.2017.02.005.

Mastrandrea M.D., Field C.B., Stocker T.F., Edenhofer O., Ebi K.L., Frame D.J., Held H., Kriegler E., Mach K.J., Matschoss P.R., Plattner G.-K., Yohe G.W., Zwiers F.W. (Core Writing Team). IPCC Fifth Assessment Guidance Note for Lead Authors IPCC Cross-Working Group for Uncertainties. Jasper Ridge, CA, USA 6-7 July 2010. Intergovernmental Panel on Climate Change (IPCC). Available at http://archive.ipcc.ch/pdf/supporting-material/uncertainty-guidance-note.pdf (accessed 20 May 2014).

Mastrandrea M.D., Mach K.J. 2011. Treatment of uncertainties in IPCC Assessment Reports: Past approaches and considerations for the Fifth Assessment Report. − Climatic Change, vol. 108, pp. 659-673.

Mastrandrea M.D., Mach K.J., Plattner G.-K., Edenhofer O., Stocker T.F., Field C.B., Ebi K.L, Matschoss P.R. 2011. The IPCC AR5 guidance note on consistent treatment of uncertainties: a common approach across the Working groups.− Climatic Change, vol. 108, pp. 675-691. DOI 10.1007 / s10584-011-0178-6.

Moss R.H., Schneider S.H. 2000. Uncertainties in the IPCC TAR: Recommendations for Authors For More Consistent Assessment and Reporting..− In: The IPCC. Editors: R. Pachauri, T. Taniguchi and K. Tanaka, Intergovernmental Panel on Climate Change, pp. 33-51. ISBN 4-9980908-0-1.

Sanchez-Rodriguez R., Ürge-Vorsatz D., Barau A.S. 2018. Sustainable development and climate change adaptation in cities. − Nature Climate Change, No. 2. DOI: 10.1038 / s41558-018-0098-9.

Submitted: 03.04.2019 *Revised: 12.06.2019*

Referencing this paper:

Semenov S.M., Insarov G.E., Méndez C.L. 2019. Characterization of uncertainties in assessments of the Intergovernmental Panel on Climate Change. − *Fundamental and Applied Climatology*, vol.2, pp. 103-119. DOI: 10.21513/2410- 8758-2019-2-103-119.

Translation from Russian. The original is published in:
Fundamental and Applied Climatology, 2019, vol. 1, pp. 76-96.