Document ID 1336783	Version 2.0	Status Approved	Reg no	Page 1 (21)
Author Mikael Asperö Lind			Date 2011-11-16	
Reviewed by Maria Lindgren			Reviewed date 2012-04-23	
Approved by Fredrik Vahlund			Approved date 2012-05-22	

SDU-505 - Supplying data for the SR-PSU Data report

Contents

1   Introduction

1.1   Purpose of the instruction

This document is an instruction issued by SKB that should be followed by suppliers and customers of data in the process of developing the SR-PSU Data report.

1.2   Scope of the instruction

This instruction should apply to all suppliers of data to the Data report and to the customer (the SR-PSU team). It should apply to all data of all subject areas of the Data report. A list of all data compiled in the Data report will be given in the Data report. The Data report concerns data that have identified uncertainties which are significant for the SR-PSU safety assessment.

Data that is not covered by the Data report can still be crucial to the safety assessment. There are reasons why data is not included: The data might already be classified as qualified by the SKB quality assurance system. An example of this is the data that originates from the SDM-PSU report. Another reason could be that the data does not have an uncertainty leading to high data variability. For example, the density of water is to be considered as crucial to the safety assessment, but the uncertainty connected with the density data is too insignificant to have any bearing on the final results and therefore will not need the be qualified in the Data Report.

1.3   Background to instruction and need for data qualification

The objective of the Data report is to compile input data, with uncertainty estimates, for the SR-PSU assessment calculations for a wide selection of conditions. Data should be assessed through standardised procedures, adapted to the importance of the data, aiming at identifying the origins of uncertainties and in which the input provided by suppliers is distinguished from judgements made by the assessment team.

All input data used in quantitative aspects of the safety assessment have uncertainties associated with them. The quality of the results of any calculation in the assessment will, among other factors, depend on the quality of the input data and on the rigour with which input data uncertainties have been handled. A methodological approach for the qualification of input data with uncertainties and the subsequent handling of data uncertainty is therefore required.

This instruction has been written to facilitate methodical and traceable data qualification, where comments made by authorities form a basis for the improvements in the data qualification methodology.
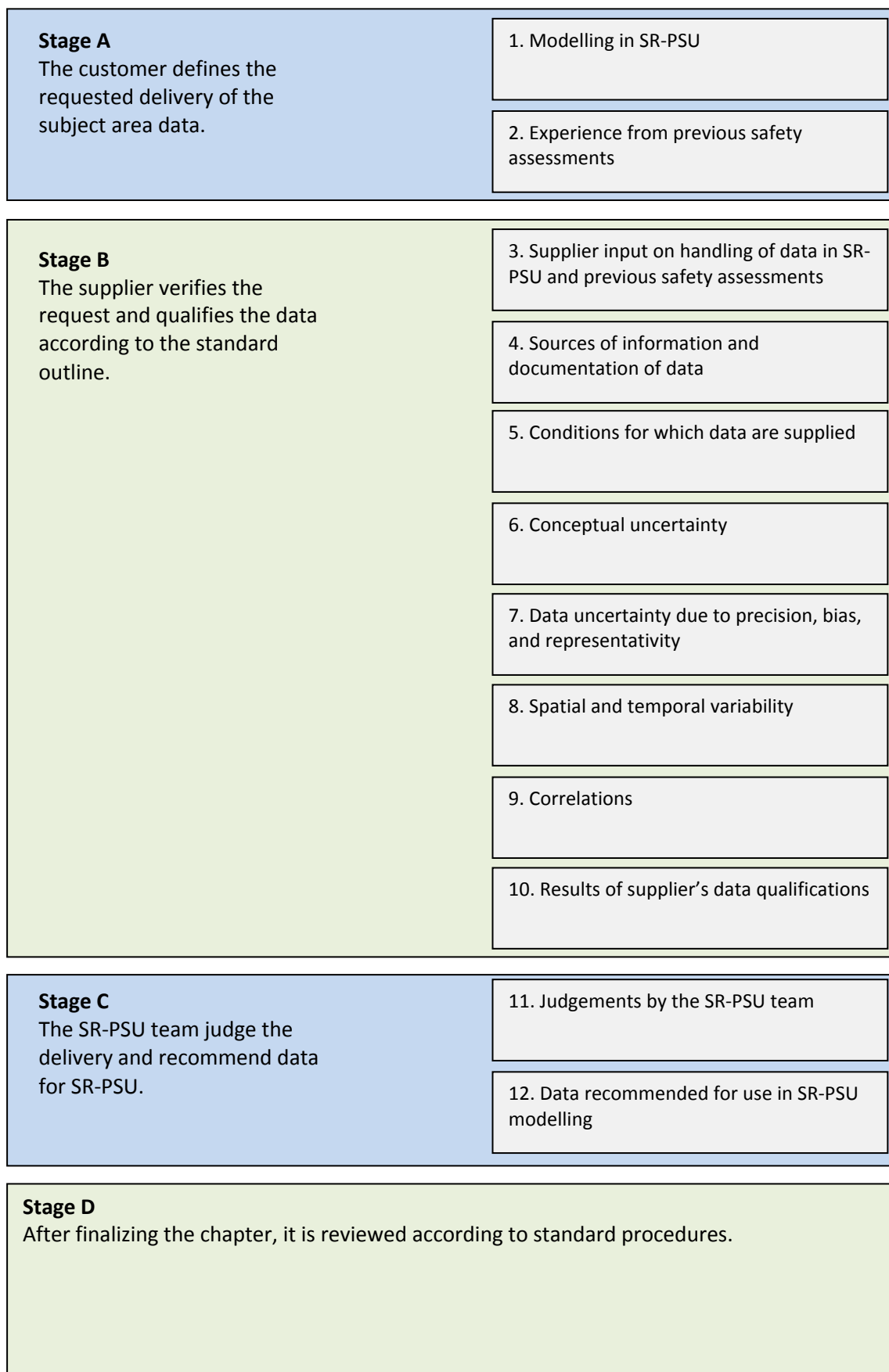
2 Qualification of input data – instruction to supplier and customer

The final objective of the Data report is at performing data qualification including estimates of both conceptual and data uncertainty, as well as of natural variability, for various subject areas. In addition, the traceability of the data is examined. The qualified data are in later stages intended for use as input data in the SR Site safety assessment modelling.

The Data report does not concern all data used in the SR-PSU safety assessment, but those which are identified to be of particular significance for assessing repository safety. Data may concern both measured data from the laboratory and from the field, as well as output from detailed modelling where measured data are interpreted, depending on the subject area. Even though the data may represent both parameters and entities, in this instruction the word data is generally used.

It should be pointed out that in the process of qualifying data, the traceability that is the focus of many quality assurance systems is only one aspect. Perhaps the more important aspect is the scrutinising of the scientific adequacy of the data.

Each data supplied in this report is categorised into one of many different subject areas. For each subject area, the data qualification process comprises a sequence of stages resulting in a text of a standard outline. The sequence of stages and the standard outline are shown in Figure 2-1.

| Stage A
The customer defines the requested delivery of the subject area data. | 1. Modelling in SR-PSU |
| | 2. Experience from previous safety assessments |

Stage B The supplier verifies the request and qualifies the data according to the standard outline.	3. Supplier input on handling of data in SR-PSU and previous safety assessments
	4. Sources of information and documentation of data
	5. Conditions for which data are supplied
	6. Conceptual uncertainty
	7. Data uncertainty due to precision, bias, and representativity
	8. Spatial and temporal variability
	9. Correlations
	10. Results of supplier's data qualifications

| Stage C
The SR-PSU team judge the delivery and recommend data for SR-PSU. | 11. Judgements by the SR-PSU team |
| | 12. Data recommended for use in SR-PSU modelling |

Stage D
After finalizing the chapter, it is reviewed according to standard procedures.

Figure 2-1. Stages of writing and reviewing the Data report. The standard outline of a subject area is shown in the grey boxes.

Svensk Kärnbränslehantering AB
Swedish Nuclear Fuel and Waste Management Co

Below the parties involved in the Data report and the sequence of stages shown in are discussed. The standard outline is described in subsections 2.1.1 to 2.1.12. For each subject area, the Data report team identifies the customer and supplier of data, and assigns a customer representative and a supplier representative that co-author the subject area section[1].

The customer is in broader terms the SR-PSU team that is responsible for performing the SR-PSU safety assessment. However, the entire team is generally not involved in each subject area but it is rather embodied by a group of persons with special knowledge and responsibility. The customer representative should represent the SR-PSU team, and not rely solely upon own opinions.

The suppliers are the teams originating the sources of data, for example the site-descriptive model reports, production line reports, and other supporting documents. The supplier representative should represent the team, and not rely solely upon own opinions.

The intended chronology of the writing of a subject area section is the following.

Stage A: The customer writes the first two subsections defining what data are requested from the supplier, how the data will be used in SR-PSU modelling, and how similar data were used in previous assessment modelling.

Stage B: The supplier writes the following eight subsections that are the core of the data qualification. This is done according to a standard outline where a number of issues such as traceability, data uncertainty, and natural variability should be dealt with. This section should result in sets of qualified data that are the delivery to the customer.

Stage C: The customer, representing the entire SR-PSU team, writes the last two subsections making judgments upon the delivery and recommending data for use in SR-PSU modelling. The text is produced in close cooperation with the supplier and other persons within the SR-PSU team with expert knowledge with the subject area.

The text of each stage should be made available in good time to the person or persons responsible for writing the text of the subsequent stage. Upon the completion of the Data report chapter, it will undergo a review process as part of Stage D.

Stage D: The Data report is reviewed according to standard procedures within the SKB quality framework.

Finally, within the SR-PSU project but outside the scope of the Data report team, a follow-up is made where it is controlled that the correct data are used in SR-PSU modelling. This could be seen as Stage E in the data qualification process, but its falls upon the modellers using the data to carry this stage through. It is therefore not shown in.

In the following subsections, the outline shown in the yellow boxes in Figure 2-1 is described in detail.

2.1.1   Modelling in SR-PSU

In this subsection, the customer should define what data are requested from the supplier, and give a brief explanation of how the data of the subject area are intended to be used in SR-PSU modelling activities.

Defining the data requested from the supplier

Here, the customer should define the data (parameters) that should be part of the supplier's delivery, in a bullet list. If applicable, the parameter symbol and unit should be provided in this list. If the supplier should focus on providing data of certain ranges, or for certain conditions, this should be specified. This text should not only facilitate the task of the supplier, but also assist the reader of the Data report in understanding the scope of the subject area section.

SR-PSU modelling activities in which data will be used

Here the customer representative should give a brief explanation of how the data are intended to be used in different SR-PSU modelling activities. This explanation should cover both how the data are

---

[1] The terms customer and supplier come from standard quality assurance terminology.

used in specific models, and in the SR-PSU model chain. Differences from the use of this type of data in previous safety assessments should be highlighted. The justification for the use of these models in the assessment is provided in other SR-PSU documents, such as the **SR-PSU Main report** and process reports.

As a result of the extensive work that will be conducted up to near completion of the SR-PSU safety assessment, details of the models and the model chain may be modified. As a result, this text may have to be finalised in a late stage of the Data report project. Thus only a preliminary version is provided early on to the supplier representative.

2.1.2   Experience from previous safety assessments

In this subsection the customer should give a brief summary on how the data of the subject area were used in previous safety assessments. The experiences from these assessments should function as one of the bases for defining the input data required in SR-PSU modelling. The summary of how the data were used in previous safety assessments should conform to the following outline:

- Modelling in previous safety assessments

- Conditions for which data were used in previous safety assessments

- Sensitivity to assessment results in previous safety assessments

- Alternative modelling in previous safety assessments

- Correlations used in previous safety assessment modelling

- Identified limitations of the data used in previous safety assessment modelling

More detailed guidance regarding what should be included in the summary in relation to each of these bullets is given below.

Modelling in previous safety assessments

The use of the data in specific previous safety assessment models, as well as in previous safety assessments model chains, should be described. Repetitions from the subsection "Modelling in SR-PSU" should be avoided. If there is no difference between the previous safety assessments and SR-PSU modelling approaches, it is sufficient to state this.

Conditions for which data were used in previous safety assessments

In this subsection, the relevant conditions to which the subject area data were subjected to in previous safety assessment modelling should be outlined. Relevant conditions are only those conditions that significantly influence the data, in the context of demonstrating repository safety. Different subject area data are affected by different conditions. For example, the sorption partition coefficient K_d may be strongly influenced by groundwater salinity. Thus, in characterising the conditions under which K_d values were used, it is likely to be appropriate to give the salinity range during repository evolution, for example as assessed in the previous safety assessments hydrogeochemical modelling. Other types of conditions may include gradients, boundary conditions, initial states, engineering circumstances, etc.

It is sufficient to state the relevant conditions used in previous safety assessment modelling (including those applied in sensitivity analyses, various initial states, different scenarios, and evolution within scenarios) and to refer to the previous safety assessments documents for background information. Justification as to why those conditions were studied is not required. Where appropriate, the relevant conditions should be tabulated. It should be noted that the stated conditions do not restrict qualification of data for use under other conditions, but merely underline the conditions considered appropriate within the modelling context of previous safety assessments.

Sensitivity to assessment results in previous safety assessments

Where appropriate, an account should be given of results from sensitivity analyses performed as part of, or prior to, the previous safety assessment. Such analyses were made in order to prioritise uncertainty assessments for those data and conditions judged to be potentially important for performance, both for overall end-points such as risk and for conditions affecting the state of the system. If such sensitivity analysis was performed, the following issues may be outlined:

- For what ranges of the data was the impact on the previous safety assessment significant and are there ranges where the impact was negligible? If sensitivity analyses show that only part of the range has an impact on repository safety, less effort may be given to quantifying parameter values outside this range.

- Was the impact monotonic, i.e. is there a unidirectional relationship between the data value and performance, is there an "optimal" value, or is the impact dependent in a complicated manner upon the values of other input data?

- What degree of variation in the data is needed to have an impact on safety assessment results (this answer may be different for different data ranges)?

- Were the results applicable to all conditions of interest – or only to some?

In discussing the above, the customer should consider if the cited sensitivity analyses were sufficiently general to provide definitive answers.

Alternative modelling in previous safety assessments

Whenever it applies, the customer representative should summarise alternative modelling approaches studied in previous assessments in which data of this type were used. The following issues should be reflected upon:

- What alternative models exist and what influence did they have on the safety assessment?

- Were conceptual uncertainties, related to the models in which the data were used, identified in previous safety assessments? In that case, what was the impact on assessment results?

Correlations used in previous safety assessment modelling

A correct treatment of probabilistic input data requires that any correlations between those data are identified and quantified. The correlations associated with the subject area data, as accounted for in previous safety assessments, should be briefly described. This includes internal correlations within the subject area and correlations with data of other subject areas. If the same correlations were used as will be used in SR-PSU, it is sufficient to state this.

Identified limitations of the data used in previous safety assessment modelling

If limitations or shortcomings of the data used in the previous safety assessments have been identified, which may significantly have affected the assessment, such should be accounted for. The limitations or shortcomings can be due to, for example, lack of site-specific data or lack of data obtained at conditions representative for the repository. The limitations and shortcomings may have been identified by the authorities, by SKB, or by other parties.

2.1.3   Supplier input on use of data in SR-PSU and previous safety assessments

In this subsection the supplier has the opportunity to comment on the two above subsections. The focus for the supplier should be to help the SR-PSU team in choosing appropriate data and modelling approaches, and avoid repeating errors and propagating misconceptions from previous safety assessments or from earlier safety analyses. Even if a single individual has the roles as both supplier

and customer representative, he or she may still make comment upon the use of data in SR-PSU and previous safety assessments.

2.1.4 Sources of information and documentation of data qualification

This section is devoted to presenting the most important sources of data, as well as categorising different data sets on the basis of their traceability and transparency. Sources of data may include SKB reports, SKB databases, and public domain material. Documents of importance for the data qualification may also consist of SKB internal documents. All underlying documents should be properly sited throughout the Data report.

Sources of information

The supplier is asked to tabulate the most prominent references used as sources of data. In addition, the reference of important documents describing the process of acquiring, interpreting, and refining data may be listed.

If the data qualification process is well documented in supporting documents, it is sufficient to reference these documents and to only briefly summarise the data qualification process. If not, the Data report gives the supplier a chance to appropriately document the data qualification process of the subject area data.

Concerning sources of information, the supplier representative should:

- Fully cite all sources of information throughout the text. It is necessary to keep in mind that the text may have readers with limited in-depth knowledge of the subject. Therefore, what normally would seem as trivial may deserve references for further reading. It is strongly recommended to make an extra effort to refer to the open literature where possible, and not only to SKB documents;

- In case of referring to a document of many pages, for example a site-descriptive model report, give detailed information on the section, figure, table, etc. where the relevant information can be found;

- Properly cite databases, SKB internal documents, etc. even though they may not be available to the general reader. In the case of referring to databases, the precise reference should be given to the individual data set used. For example, it is not sufficient to refer to the SKB database SICADA if not also giving detailed information, such as the activity id. This is to ensure traceability within the SR-PSU project;

- Fully cite advanced modelling tools where the underlying code may have implications for data qualification.

Categorising data sets as qualified or supporting data

The supplier representative should categorise data as either qualified data or supporting data. Qualified data has been produced within, and/or in accordance with, the current framework of data qualification, whereas supporting data has been produced outside, and/or in divergence with, the framework. Data taken from peer-reviewed literature takes a special position in that they may be considered as qualified even though they are produced outside the SKB framework of data qualification. However, such data are not by necessity categorised as qualified, as they may be non-representative or lack in some other aspect.

Data recently produced by SKB, for example in the site investigations, should a priori be considered as qualified. However, before the data are formally categorised as qualified, a number of considerations need to be made as described below. Data produced outside the data qualification framework should a priori be considered as supporting data. This could for example be data produced by SKB prior to the implementation of its quality assurance system, or data produced by other organisations. Before

formally categorise the data as supporting, a number of considerations need to be made as described below.

Data taken from widespread textbooks, engineering handbooks, etc., which are considered to be established facts, need not to be scrutinised. Well-known data that should be excluded from the Data report need not to be categorised as qualified or supporting data, although their exclusion may need to be motivated.

It is outside the scope of the Data report to deal with individual data. Instead the supplier representative should characterise data sets as qualified or supporting. The supplier representative should decide to what extent various data can be included in a single data set for the specific case. The following examples of natural barrier data sets could be used for inspiration:

- Data or part of data, obtained by a specific method at a site, rock volume, borehole, etc.

- Data or part of data, obtained by various methods at certain conditions (e.g. saline water) at a site, rock volume, borehole, etc.

- Data or part of data, taken from an external publication.

Qualified data

The following considerations should be made for data that a priori are identified as qualified, before formally categorising them as qualified. Most of the data that is delivered to the Data report are refinements and interpretations of observed data. Such refinements and interpretations are performed both for engineered and natural barrier data. For example, the multitudes of data acquired within the site investigation are normally refined within the site-descriptive modelling by use of more or less complex models. The supplier should judge whether data acquisition and refinement, and associated documentation, are in accordance with the implemented data qualification framework. The following considerations may form the basis for the judgement.

Considerations concerning data acquisition:

- Is the acquisition of observed data performed in conformance with a widely adopted quality management system (e.g. the ISO 9000 series or equivalent)?

- Is it possible to trace relevant quality assurance documents (for example method descriptions, field notes, etc.) for the measurements? It should be noted that even though the quality assurance documents may not be available for the general reader, they are accessible for the SR-PSU team.

- Is it possible to extract relevant information on the data quality, variability, and representativity from documents reporting the acquisition of data?

- Are concerns associated with the observed data and nonconformities of the measurements transparently described?

- Is the undertaken data acquisition programme sufficient to determine the full range of data uncertainty and natural variability, and do the acquired data appropriately characterise the intended aspect of the system (site, rock domain, waste package, population, etc.)?

Considerations concerning data refinement:

- Are concerns and nonconformities described in the supporting documents propagated to, and addressed in, the data refinement?

- In refining observed data by use of more or less complex modelling, is this done in accordance with documented methods?

- In case of more complex modelling, which may have implication for data qualification, is the details of the modelling described in either in a task description or in the document reporting the modelling results? Furthermore, is the modelling tool developed in accordance with a widespread quality assurance system and/or is its quality tested in other ways?

- Has comparative/alternative modelling been performed to evaluate artefacts induced in the modelling, and to evaluate whether the modelled interpretation of the data is reasonable?

Going through these questions in detail for each data set may be a too extensive task, wherefore the sorting of data to some degree is based on expert judgement. However, in making this judgment, it may be helpful to revisit the above bullet lists.

If appropriate data qualification has been performed and documented in supporting documents, or can be performed and documented as part of the delivery, the data should be formally categorised as qualified data. If the documentation of the data qualification process is inadequate in supporting documents, and appropriate data qualification cannot be performed as part of the delivery, the data must be demoted to the category supporting data.

As mentioned before, data taken from peer-reviewed literature takes a special position in that they may be considered as qualified even though they are produced outside the SKB framework of data qualification. However, before formally categorising them, one needs to judge whether they are representative for the intended KBS-3 repository system and the Forsmark site. A prerequisite for making such a judgement is often that the documents are transparently written. In case the data are non-representative for Swedish conditions, or their degree of representativity is difficult to evaluate, the data may be categorised as supporting data instead of as qualified data.

Supporting data

The following considerations should be made for data that a priori are identified as supporting, before formally categorising them as supporting data. Such data are produced by SKB outside the framework of data qualification, or by other organisations. The supplier representative should consider:

- How well is the method used to acquire the data described? The greater the transparency with which the method is described in the supporting document, the greater the value should be ascribed to the data;

- How well is the method used to interpret and refining the data described? The more transparently the interpretation and refinement is described in the supporting document, the greater the value should be ascribed to the data;

- Is it possible to identify and evaluate the data qualification process used in acquiring and refining the data? If it is shown that a sound data qualification process has been used, the data should be ascribed greater value;

- Judge, based on the above, whether the data can be used as part of the basis for recommending data to SR-PSU safety assessment modelling, as comparative data for other qualified data, or should not be used at all. In some cases the transparency of a document is so poor that crucial information concerning data qualification cannot be extracted. If this renders an assessment of the data's scientific adequacy and their representativity for Swedish conditions impossible, the supplier representative should recommend that the data are dismissed. This can be done even if the numerical values of the data are consistent with other, qualified data.

In case data that a priori are assumed to be supporting are acquired, interpreted, and refined according to a similar data qualification framework as implemented by SKB, and this is transparently described, the supplier representative can promote the data to the category qualified data.

It should be noted that data taken from peer-review literature can be categorised as supporting data. This can be done if, for example, data are only partially representative for the Swedish repository concept and the Forsmark site.

Upon formally categorising the data sets as qualified or supporting, they should be tabulated as exemplified in Table 1. As can be noted, motivations for the sorting are given in the same table for the different items.

Table 1. Qualified and supporting data sets (for parameter X).

Qualified data sets	Supporting data sets
1. /SKB, 20xx/, Section 4.5: All data on parameter X obtained for rock domain RFM029.	5. /Nilsson, 19xx/, Table 1. Data obtained in the pH range 6–9 in sedimentary rock.
2. Data presented in the **Underground construction opening report** in Figure 1.	
3. /Svensson, 20xx/, Table 2: Data between the borehole length 400-452 m in KFM01D, indicating an average value of 2,650 m^3/kg.	
4. All parameter X data presented in SKB Database Y, with the identity number xxx-yyy-zzz.	

1-2, 4: These data have been produced within the site investigation (item 1), within a production report (item 2), or as part of the site-descriptive modelling (item 4). These data are produced within the SKB data qualification framework and are judged as qualified.

3. /Svensson, 20xx/ is a peer-review article and the data are obtained at the Forsmark site and are judged as representative. The data set is judged as qualified.

5. /Svensson, 20xx/ is a peer-review article that is transparent and scientifically sound. However, the data are predominantly representative for sedimentary rock, wherefore they are judge as supporting.

Excluded data

Within the field of nuclear waste management, there are large quantities of data that are of little significance for the SR-PSU safety assessment, as they are less representative for the Forsmark site. than other available data. In general, excluding such data from subsequent use in SR-PSU does not require justification. The exception is if the data constitutes a well-known part of the basis of previous safety assessments (or equivalent tasks), and/or have a significant impact on the perception of the appropriate choice of data value. If it could be seen as a significant inconsistency or omission not to use the data, their exclusion should be explicitly justified. Providing an appropriate justification is particularly important if the excluded data disagree with the presently used data.

2.1.5 Conditions for which data are supplied

The data of the different subject areas are likely affected by different conditions. Conditions refer to initial conditions, boundary conditions, barrier states, and other circumstances, which potentially may affect the data to be estimated. In the process of qualifying data for subsequent use in safety assessment, an important part is to account for the conditions for which data were acquired, and to compare these conditions with those of interest for the safety assessment.

In the subsection "Experience from previous safety assessments" it is stated for what conditions data were used in the previous safety assessments. These conditions should not limit the conditions for which data are examined, but merely point out conditions that are likely to be of importance for a safety assessment. The supplier may have been given instructions from the SR-PSU team, or may have opinions about important conditions, which lead to modifications of the previous safety assessments conditions.

In this subsection, the conditions for which the data have been obtained should be discussed and, as appropriate, justified as relevant to SR-PSU. Such a condition is often a single value (e.g. temperature), a range (e.g. salinity range), or a gradient (e.g. hydraulic gradient). Other factors of

relevance for repository safety may be included as conditions, at the discretion of the supplier. Conditions that are deemed to be of particular importance for repository safety should be highlighted. Other conditions that do not significantly relate to repository safety, but may be of importance for data qualification, are also important to note. Such information is valuable when, for example, crosschecking data sets with those of other studies or evaluations. The supplier representative may list ranges of applied conditions during data acquisition, excluding conditions that are both general (such as the gravitational constant) and self-evident.

In many cases, it is expected that the conditions for which data are supplied will differ from those that apply in the SR-PSU safety assessment. For example, a set of supplied data may not represent the full temperature range required, or may have been obtained at a different pressure than expected in-situ. The differences identified by the supplier representative should be outlined in this subsection. Furthermore, for each deviating condition of importance for the assessment results, the implications should be discussed.

2.1.6   Conceptual uncertainty

This subsection concerns conceptual uncertainty of the subject area data. Two types of conceptual uncertainty should be discussed. The first concerns how well the data, and the models wherein it is used, represent the physical reality, and the second concerns conceptual uncertainties introduced in the acquisition, interpretation, and refinement of the data. Generally data are included of models that represent an idealised reality, which to some degree differs from the physical reality. Therefore, one can expect that a degree of conceptual uncertainty is associated with all data compiled in this Data report.

To the extent possible, the supplier should describe such conceptual uncertainty. This should be done in the context of the models in which the data are used, intended to describe certain postulated processes. Also, it may be appropriate to discuss alternative conceptualisations in which the data may be used in different ways. If comprehensive discussions on the subject have already been documented, such documents may be referred to and a short summary of the conceptual uncertainty will suffice. Aspects of the conceptual uncertainty that are obviously unrelated to repository safety may be disregarded.

Conceptual uncertainty may also be introduced in the acquisition, interpretation, and refinement of the data. For example the data may have been obtained by inverse modelling of experimental results, where conceptual uncertainty is introduced by the model. The data may also have been obtained by using some correlation relationship, where there is conceptual uncertainty in the correlation. Many other sources of conceptual uncertainty are conceivable and may be discussed at the discretion of the supplier. In doing this, the supplier representative should carefully differentiate between uncertainties introduced due to conceptual issues and data uncertainty introduced by measurement errors, etc. Data uncertainty should be discussed in the following subsection.

2.1.7   Data uncertainty due to precision, bias, and representativity

In this subsection data uncertainty should, if possible, be discussed in terms of precision, bias, and representativity, in the context of their application in SR-PSU. Such uncertainty is associated both with the acquisition of data, for example in the site investigations, and subsequent refinement of data, for example in the site-descriptive modelling. Data uncertainty includes neither conceptual uncertainty nor natural variability.

If comprehensive discussions on these matters are documented elsewhere, such documents should be referred to, and a short summary of the discussion will suffice. The supplier should begin with discussing the precision of the supplied data. To the extent possible, data spread due to the precision should be separated from data spread due to natural variability. Precision issues are both associated with the method used in acquiring the raw data and subsequent interpretation of data. Concerning acquiring raw data, limitations in precision are not only associated with the equipment and method used when performing the measurements, but also with the sampling procedure, sample preparation, etc. Precision issues associated with interpretation of the data depend to a large degree on the

procedure used, and should be discussed at the discretion of the supplier. As an example, it may not be straight forward to estimate the precision of data that are a function of other acquired data, with their intrinsic limitations in precision.

Thereafter, the supplier representative should discuss the bias of the supplied data. Similar considerations apply as when discussing precision, both for bias associated with the acquisition of raw data and with their subsequent interpretation. Bias in observed data is often associated with the method used for acquiring data and its calibration, and with effects of sample preparation. Bias is also associated with the sampling procedure, sample size, and differences in conditions for example between those in the laboratory and in-situ. Bias issues associated with data interpretation depend to a large degree on how the interpretation is made, and should be discussed at the discretion of the supplier representative.

Finally the supplier representative should discuss the representativity of the supplied data, both in terms of data acquisition, and data interpretation and refinement. Issues associated with the representativity of acquired data often concern the sampling procedure, the sample size relative to natural variability and correlation length, and differences in conditions between, for example, those in the laboratory and in-situ.

An important issue is whether the data are generic or site and/or technique specific. In the case of access to generic data only, the supplier should discuss whether, and to what degree, the lack of site and/or technique specific data influences the data uncertainty. Representativity issues associated with data interpretation and refinement depend much on the specific interpretation and refinement process, and should be discussed at the discretion of the supplier representative.

As well known, the precision, bias, and degree of representativity often depend on a mixture of the above-suggested sources for data uncertainty, and may not be easily separated. However, the supplier representative is asked to reflect carefully on these issues, as an assessment of data uncertainty is crucial in data qualification. In case data uncertainty cannot be discussed in terms of precision, bias, and representativity, for example as the resolution in data does not allow for such separation, it will suffice to make a general data uncertainty discussion.

Comprehensible illustrations of different data sets are of high value. The objective of the illustrations is not necessarily to provide a detailed basis and description of the numerical values of the individual data. Sometimes the objective may be to give the reader an understanding of how much, and in what ways, the data varies and the data sets differ from each other. An example of presenting different data sets is given in Figure 2-2, where the reader can get an immediate perception about differences between the data sets. Examples of other figures are given in subsection 2.1.10.
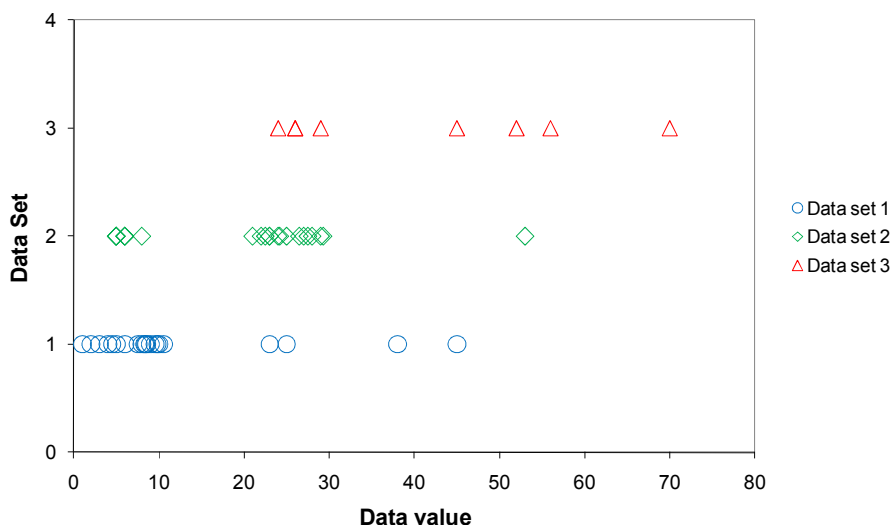

Figure 2-2. Example of presenting differences in data sets.

2.1.8   Spatial and temporal variability

In this subsection the supplier should discuss the spatial and temporal variability of the subject area parameters. The natural variability should as far as possible be separated from data uncertainty, discussed in the above subsection.

The supplier should describe what is known about the spatial variation, sometimes referred to as heterogeneity, of the subject area data. This may result in different data sets for different volumes or elements of the repository system, or for different time periods. If comprehensive discussions on the natural variability are documented elsewhere, such documents should be referred to and a short summary of the natural variability will suffice.

In the process of describing the spatial variability, it may be helpful to reflect on the following line of questions.

- Is there spatial variability of the data, and if so is it of consequence for the safety assessment?

- Is the spatial variability scale dependent? If so, can an appropriate approach of upscaling to safety assessment scale be recommended?

- What is known about correlation lengths from, for example, variograms?

- Can the spatial variability be represented statistically as a mean of data qualification and, if so, how is this done?

- Is there any information about the uncertainty in the spatial variability?

In the process of describing the temporal variability, it may be helpful to reflect on the following line of questions.

- Is there temporal variability of the data, and if so is it of consequence for the safety assessment?

- What processes affect the temporal variability of the data and how is the temporal variability correlated with these processes?

- Does the temporal variability follow any pattern, for example a cyclic pattern?

- Could the temporal variability be represented statistically as a mean of data qualification and if so, how is this done?

- Is there any information about the uncertainty in the temporal variability?

In addition, other relevant issues concerning the natural variability may be addresses at the discretion of the supplier. Comprehensible illustrations of different data sets from different volumes, elements, or time periods are of high value.

2.1.9   Correlations

An appropriate treatment of probabilistic input data requires that any correlations and functional dependencies between those data are identified and quantified. In the extensive work with the FEP database and the Process reports, most correlations and functional dependencies between parameters have been identified. Where appropriate, these correlations and functional dependencies should also be implemented in the safety assessment models. It should be an aim to aid those performing stochastic modelling, by giving well defined and usable information on how to handle correlations between input data.

Correlations and functional dependencies may also have been used when acquiring, interpreting, and refining data. For example, concerning sorption partition coefficients, data have not been acquired for

all relevant radionuclides. For species for which there is a lack in observations, the supplied sorption partition coefficient will have been estimated from data obtained for one of more analogue species. This has implications for how to correlate input data in stochastic safety assessment modelling.

In this subsection the supplier representative is requested to address the following questions:

- For the subject area data, are there correlations or functional dependencies between parameters of the same or of different subject areas? If so, account for these and if possible also for the consequences for the safety assessment.

- If correlations have been used in acquiring, interpreting, and refining data, how is this done? Furthermore, is the outcome based solely upon correlations, or on both measurements and correlations?

- If the data varies in space and time – is anything known about its autocorrelation structure?

- Is there any other reason (apart from already cited correlations and functional dependencies) to suspect correlations between parameters considered as input to SR-PSU modelling?

2.1.10   Results of supplier's data qualification

In this subsection the supplier is requested to present data that are considered to be appropriate as a basis for suggesting input data for use in SR-PSU. Comprehensive information relating to each parameter requested in the bullet list under the heading "Defining the data requested from the supplier" (cf. subsection 2.1.1) should be given. Only one set of data should be delivered for each specified condition, volume, element, time period, etc.

The general process of reducing and interpreting data, valuing different data sets, and finally selecting the recommended data for delivery to the SR-PSU team should be fully accounted for, if not already accounted for in the previous subsections or in supporting documents. In the latter case, it is sufficient to briefly summarise the process of selecting the delivered data.

In case the data presented in supporting documents need reinterpretation and further refinement, in the light of this instruction and/or other information, this should be fully documented. In case the supporting documents give more than one data set for a specified condition, volume, element, time period, etc., further data reduction is required. Such data reduction may include the merging of data sets, and there may be a need to give different weight to different data sets. Much weight should be given to peer-reviewed data judged as representative for the Swedish site and repository system. Generally, more weight should be given to qualified data than to supporting data. The degree to which the data are representative in the context of their application in SR-PSU should also be a factor in the weighting. Exactly how much weight should be given to individual data sets must be decided upon by the supplier. The process of further reinterpretation, refinement, and data reduction should be fully documented. If it increases the readability of the text to also utilise other subsections for such documentation, this is allowed. Also, if this requires much space, some information may be appended.

The data sets that the supplier recommends to the SR-PSU team should be in the form of single point values, probability distributions, mean or median values with standard deviations, percentiles, ranges, or as otherwise appropriate. If the data have significant variability and/or uncertainty, the spread in data could be described as a range. However, the meaning of the range has to be provided, e.g. does it represent all possible values, all "realistically possible" values or just the more likely values? The supplier may provide more than one range, representing different probabilities, as exemplified below:

- The range wherein the likelihood of finding the data is high.

- The range for which the likelihood of finding the data outside this range is very low

All data should be recommended in the context of input data to safety assessment modelling, wherefore the final uncertainty estimate should encompass conceptual uncertainty, data uncertainty,

Svensk Kärnbränslehantering AB
Swedish Nuclear Fuel and Waste Management Co

and natural variability (cf. subsections 2.1.6, 2.1.7, and 2.1.8). If the supplier representative has used some kind of mathematically expression to account for the uncertainty and natural variability, this expression should be provided and justified.

If the data are suggested to be described by a well defined probability distribution, it should be justified on statistical grounds that the data indeed are (sufficiently well) distributed accordingly. The usage of standard deviation is often perceived to imply that the data are normally distributed; even through the definition of standard deviation is unrelated to specific probability distributions. Therefore, when giving the standard deviation, it should be remarked upon whether or not the normal distribution appropriately describes the data. If there are obvious differences between how the data set at hand is actually distributed, and the probability distribution (or range) finally recommended, the reasons for, and implications of, this should be discussed. Outliers should not be dismissed without justification.

It should be noted that in many cases, at some stage probability distributions must be assigned to numerical data being the input to probabilistic safety assessment modelling. If the supplier feels inadequate to deliver a defined distribution, but for example delivers a best estimate, an upper, and a lower limit for data, it may fall on the SR-PSU team to transform such information into probability distributions. This is justified as the SR-PSU team may have a better understanding of how the shape of the assigned distributions (especially in their tails) affects the assessment results. The SR-PSU team may also, in some cases, have a better understanding of the underlying statistics of the suggested distribution.

The above instructions are not applicable to all data, as all data are not necessarily in the form of numerical values. Examples are exit locations for groundwater flowpaths, given as co-ordinates, or information on solubility limiting phases, given as chemical species and reactions.

For a spatially varying function well described by a given stochastic process, e.g. through a variogram or as realised in a Discrete Fracture Network, a potential statement may be that all realisations of this spatially varying function are equally probable.

Finally, it may be impossible to express the uncertainty by other means than a selection of alternative data sets. There are a number of uncertainties that cannot be managed quantitatively in any other rigorous manner, from the point of view of demonstrating compliance, than by pessimistic assumptions. This is allowed, as long as the supplier clearly documents this together with the motivation for adopting this approach.

)
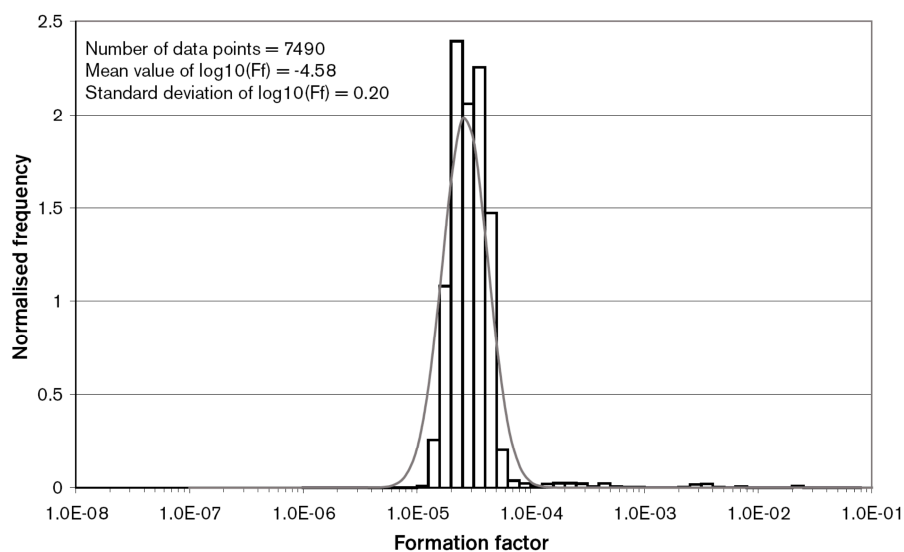
Nuclide/redox state		Non-saline K_d (m^3/kg)
Ni(II)[1]	best estimate	$1.2 \cdot 10^{-1}$
	$K_{d,25\%} - K_{d,75\%}$	$5.5 \cdot 10^{-2} - 3.0 \cdot 10^{-1}$
	$K_{d,low} - K_{d,high}$	$1.8 \cdot 10^{-2} - 5.4 \cdot 10^{-1}$
Sr(II)[1]	best estimate	$1.3 \cdot 10^{-2}$
	$K_{d,25\%} - K_{d,75\%}$	$6.5 \cdot 10^{-3} - 4.1 \cdot 10^{-2}$
	$K_{d,low} - K_{d,high}$	$1.0 \cdot 10^{-3} - 6.1 \cdot 10^{-1}$

b)

	Limiting specie	Reaction
Ag(0)	Ag(s)	$Ag(s) + H^+ + 0.25 O_2 = Ag^+ + 0.5 H_2O$
Ag(I)	AgCl(cr)	$AgCl(cr) = Ag^+ + Cl^-$
Am(III)	Am(OH)$_3$(am)	$Am(OH)_3(am) + 3H^+ = Am^{+3} + 3H_2O$

c)

Figure 2-3. Examples of representations of recommended data, taken from the SR-Can Data report /SKB, 2006b/

For data which are impractical to tabulate in the Data report (for example the co-ordinates of thousands of exit locations for groundwater flowpaths), it is sufficient to precisely refer to a database or equivalent. However, if possible the data should be illustrated in figures or excerpts of tables. Unless published or stored elsewhere, the data should be stored in a database associated with the Data report.

2.1.11 Judgements by the SR-PSU team

In this subsection, the customer representative, on behalves of the SR-PSU team, should document the examination of the delivery provided by the supplier, and make judgment on the data qualification. This text should be produced in close cooperation with persons of the SR-PSU team with special knowledge and responsibility. In case of unresolved issues, the final phrasing should be decided upon by the SR-PSU team. Comments should be made on all the subsections listed below:

- Sources of information and documentation of data qualification

- Conditions for which data are supplied

- Conceptual uncertainty

- Data uncertainty due to precision, bias, and representativity

- Spatial and temporal variability

- Correlations

- Results of supplier's data qualification

Concerning the subsection "Sources of information and documentation of data qualification" the customer should judge if appropriate documents are referenced, and if the categorisation of data sets into qualified or supporting data is adequately performed and justified.

Concerning the subsection "Conditions for which data are supplied" the customer should focus upon whether the conditions given by the supplier are relevant for SR-PSU modelling. If not, it should be accounted for how this is handled in SR-PSU (for example by extrapolating data, using generic data, or assuming conservative values) and what degree of uncertainty such a procedure induces.

Concerning the subsection "Conceptual uncertainty" the customer should judge whether the discussion provided by the supplier is reasonable and sufficiently exhaustive. If the customer sees the need to include additional sources of conceptual uncertainty, such should be described and if possible quantified. Finally, where necessary the impact of the conceptual uncertainty on the assessment should

be discussed, as well as how conceptual uncertainty is handled in SR-PSU modelling (for example by applying conservative corrections factors to the data).

Concerning the subsection "Data uncertainty due to precision, bias, and representativity", the customer should make a judgment on the account provided by the supplier representative. Also, if the customer sees the need to include additional sources of data uncertainty, these should be described and if possible quantified. If necessary the impact of the data uncertainty on the assessment should be discussed, as well as how data uncertainty is handled in SR-PSU modelling (for example by applying data uncertainty distributions or using corrections factors for the data).

Concerning the subsection "Spatial and temporal variability" the customer should focus upon whether the spatial and temporal variability are adequately characterised and whether they are of relevance for SR-PSU modelling. Also, if the customer sees the need to include additional sources of spatial and temporal variability, such should be described and if possible quantified. In necessary, the impact of the spatial and temporal variability on the assessment should be discussed, as well as how this is handled in SR-PSU modelling (for example by applying data distributions or different data for different model times and volumes).

Concerning the subsection "Correlations" the customer should scrutinise the correlations and functional relationships suggested by the supplier. Also, if correlations other than those suggested by the supplier are identified in the SR-PSU programme (for example in Process reports) these should be briefly described where necessary. If appropriate, a summary could be provided concerning which correlations are of actual importance for safety assessment modelling and results.

Concerning the subsection "Result of data qualification" the customer should make judgement on the choice of data by the supplier, based on scientific adequacy, usefulness for the safety assessment, and the data qualification process. Comments should be made on the delivered estimates of data uncertainty and natural variability, as well as on the data reinterpretation/refinement/reduction process. Furthermore, the delivered distributions, data ranges, etc. should be scrutinised from a statistical point of view. It should be judged whether the suggested way of representing data, for example by a log-normal distribution, is adequate for SR-PSU modelling. If the SR-PSU team chooses to promote other data than those suggested by the supplier, the choice should be fully documented.

For all the subsections listed above, supplier statements or supplied data believed to be extra uncertain, dubious, or even erroneous should be highlighted by the customer. These matters should be raised with the supplier and, if possible, resolved and accounted for in this subsection.

2.1.12   Data recommended for use in SR-PSU modelling

The main delivery of the Data report to the SR-PSU modelling is recommendations of data that generally are numerically well defined. Such recommended data should be given in this subsection.

Based on all the available information, but also on the needs from SR-PSU modelling, the customer representative and SR-PSU team should make a final choice of data in form of single point values or well-defined probability distributions, encompassing natural variability, data uncertainty and other uncertainty. These data should be clearly tabulated (or otherwise presented) in this section. Alternatively, precise referencing to tables or equivalent in previous section can be made. For data which are impractical to tabulate in the Data report it is sufficient to precisely refer to a database or equivalent.

Also short guidelines for how to use the data in subsequent modelling should be given, as required. Justifications and guidelines should be kept short so that this subsection mainly contains tabulated data that are easily extractable for SR-PSU safety assessment modelling.

In the process of making the final choice of data, the supplier representative, and potentially also other members of the supplier team, will be consulted one more time in a data qualification meeting. Here the formal decision on the data recommended for use in SR-PSU modelling should be taken, and records of the meeting should be made as part of the SKB quality assurance system. The formal

decision should be acknowledged by those representing the supplier team and those representing the SR-PSU team.

3   Instructions to the customer representative and SR-PSU team

Based on all the available information, but also on the needs from SR-PSU modelling, the customer representative and SR-PSU team should make a final choice of data in form of well-defined probability distributions, including natural variability, data uncertainty and other uncertainty. The choice should be fully documented and the resulting data should be tabulated. Also guidelines for how to use the data in subsequent modelling should be given, as required. Justifications and guidelines should be kept short so that this subsection mainly contains tabulated data that are easily extractable for the SR-PSU safety assessment modelling.

In the process of making the final choice of data, the supplier representative, and potentially also other members of the supplier team, will be consulted one more time in a data qualification meeting. Here the formal decision on the data recommended for use in SR-PSU modelling should be taken, and records of the meeting should be made as part of the SKB quality assurance system. The formal decision should be acknowledged by those representing the supplier team and those representing the SR-PSU team.

The main delivery of the Data report to the SR-PSU modelling is recommendations of data that generally are numerically well defined. Such recommended data should be given in this section. Based on all the available information, but also on the needs from SR-PSU modelling, the customer representative and SR-PSU team should make a final choice of data in form of single point values, ranges, or well-defined probability distributions, encompassing natural variability, data uncertainty, and other uncertainty. These data should be clearly tabulated (or otherwise presented) in this section. Alternatively, precise referencing to tables or equivalent in previous sections can be made. For data which are impractical to tabulate in the Data report it is sufficient to precisely refer to a database or equivalent.

Also short guidelines for how to use the data in subsequent modelling should be given, as required. Justifications and guidelines should be kept short so that this section mainly contains tabulated data that are easily extractable for SR-PSU safety assessment modelling.

4   References

Andersson, J., 1999. SR 97 - Data and data uncertainties. Compilation of data and data uncertainties for radionuclide transport calculations. SKB TR-99-09, Svensk Kärnbränslehantering AB.

Dverstorp, B and Strömberg, B, 2008, SKI:s och SSI:s gemensamma granskning av SKB:s säkerhetsrapport SR-Can. Granskningsrapport, SKI Rapport 2008:19 , SSI Rapport 2008:04

Hedin, A., 2002. Safety Assessment of a Spent Nuclear Fuel Repository: Sensitivity Analyses for Prioritisation of Research. Swedish Nuclear Fuel and Waste Management Co., Sweden, In: Proceedings of the 6th International Conference on Probabilistic Safety Assessment and Management PSAM6. Elsevier Science Ltd. (2002).

Hedin, A., 2003. Probabilistic dose calculations and sensitivity analyses using analytic models. Swedish Nuclear Fuel and Waste Management Co., SKB, Stockholm, Sweden, In: Reliability Engineering & System Safety 79(2003):2 pp. 195-204.

Hora, S., 2002. Expert opinion in SR 97 and the SKI/SSI joint review of SR 97. SSI Rapport 2002 20, Department of Waste Management and Environmental Protection SSI - Statens strålskyddsinstitut (Swedish Radiation Protection Authority) Stockholm Sweden.

Hora, S. and M. Jensen, 2002. Expert judgement elicitation. SSI Rapport 2002 19, Department of Waste Management and Environmental Protection SSI - Statens strålskyddsinstitut (Swedish Radiation Protection Authority) Stockholm Sweden.

SKB, 2004. Interim data report for the safety assessment SR-Can. SKB R-04-34, Svensk Kärnbränslehantering AB.

SKB 2006, Data report for the safety assessment SR-Can. SKB TR-06-25, Svensk Kärnbränslehantering AB.

SDU-507 - Instruction for use of preliminary data used in SR-PSU calculations/modelling

SKI/SSI, 2001. SKI's and SSI's joint review of SKB's safety assessment report, SR 97. Summary. SKI Report 01 3, SSI-report 2001 2, SKI - Statens kärnkraftinspektion (Swedish Nuclear Power Inspectorate); SSI - Statens strålskyddsinstitut (Swedish Radiation Protection Institute).

Wilmot, R. D. and D. A. Galson, 2000. Expert judgement in performance assessment. SKI Report 00 4, SKI Statens kärnkraftinspektion (Swedish Nuclear Power Inspectorate) Stockholm Sweden.

Wilmot, R. D., D. A. Galson and S. C. Hora, 2000. Expert judgements in performance assessments. Report of an SKI/SSI seminar. SKI Report 00 35, SKI - Statens kärnkraftinspektion (Swedish Nuclear Power Inspectorate) Sweden.

Register of revisions

Version	Content of revision	Made by	Reviewed by	Approved by
1.0	New document	Mikael Asperö Lind	Se header	Se header
2.0	Correction in text , from copper canister to waste package in page 9	Mikael Asperö Lind	Se header	Se header