ELSEVIER
Journal of Environmental Radioactivity 42 (1999) 177–190
JOURNAL OF ENVIRONMENTAL RADIOACTIVITY

Effect of user interpretation on uncertainty estimates: examples from the air-to-milk transfer of radiocesium

Gerald Kirchner[a,*], S. Ring Peterson[b], Ulla Bergström[c], Simon Bushell[d], Philip Davis[b], Vitold Filistovic[e], T.G. Hinton[f], Pawel Krajewski[g], Thomas Riesen[h], Paul Uijt de Haag[i]

[a] *University of Bremen, FB 1, Postfach 330440, D-28334 Bremen, Germany*
[b] *AECL Chalk River Laboratories, Chalk River, Ontario K0J 1J0, Canada*
[c] *Studsvik EcoSafe, S-61182 Nyköping, Sweden*
[d] *Quantisci, 45 Station Road, Henley-on-Thames, Oxfordshire RG9 1AT, UK*
[e] *Institute of Physics, Savanoriu av. 231, 2028 Vilnius, Lithuania*
[f] *University of Georgia, Savannah River Ecology Laboratory, P.O. Drawer E, Aiken, SC 29802, USA*
[g] *Central Laboratory for Radiological Protection, St. Konwaliowa 7, PL-03-194 Warsaw, Poland*
[h] *Paul Scherrer Institut, CH-5232 Villingen-PSI, Switzerland*
[i] *RIVM, P.O. Box 1, NL-3720 BA Bilthoven, Netherlands*

Received 24 March 1997; accepted 5 September 1997

Abstract

An important source of uncertainty in predictions of numerical simulation codes of environmental transport processes arises from the assumptions made by the user when interpreting the model and the scenario to be assessed. This type of uncertainty was examined systematically in this study and was compared with uncertainty due to varying parameter values in a code. Three terrestrial food chain codes that are driven by deposition of radionuclides from the atmosphere were used by up to ten participants to predict total deposition of ^{137}Cs and concentrations on pasture and in milk for two release scenarios. Collective uncertainty among the predictions of the ten users for concentrations in milk calculated for one scenario by one code was a factor of 2000, while the largest individual uncertainty was 20 times lower. Choice of parameter values contributed most to user-induced uncertainty, followed by scenario interpretation. Due to the significant disparity in predictions, it is recommended that assessments should not be carried out alone by a single code user. © 1998 Elsevier Science Ltd. All rights reserved.

* Corresponding author.

0265-931X/98/$19.00 © 1998 Elsevier Science Ltd. All rights reserved
PII: S0265-931X(98)00053-8

1. Introduction

Uncertainty in predictions of ecological models arises from many sources (Loehle, 1987). These include the adequacy of the conceptual model, how well the conceptual model has been coded, the details contained in the scenario description, the quality of the input data, the uncertainty in parameter values, and the assumptions made by the user.

The most easily quantified source of uncertainty is that from parameter uncertainty. In recent years, statistical propagation of uncertainties in parameter values has become a commonplace for environmental transfer models (e.g., Bergström and Nordlinder, 1991; Müller and Pröhl, 1993). Uncertainties arising as a result of the design of the conceptual model may also reduce confidence in the results (Ahn and Jin, 1996). In contrast, errors arising from the coding of the conceptual model should be small, since codes can be verified in a number of ways (Sheng and Ören, 1990).

Uncertainties from the user's interpretation of a model, a code, a scenario to be modelled and from the choice of parameter values to be used are rarely taken into account, and usually their potential importance is ignored. However, intercomparison studies of internal dosimetry assessments (Hui et al., 1994) and calibration of a hydrology model (Botterweg, 1995) indicate that model users may contribute significantly to the uncertainty of model predictions. Most likely, uncertainty from user assumptions will increase when modelling occurs under stressful conditions, such as after an accident when calculations must be performed under severe time constraints with limited experience of the model and with incomplete knowledge of the situation/system to be assessed. Generally, there is no mathematical framework to quantify user influence on model predictions.

In a previous BIOMOVS study (Köhler et al., 1991), predictions of various models were compared with observations on the contamination of vegetation and animal produce after the Chernobyl accident in 1986. Two modelling groups independently used the same model and submitted widely differing predictions, indicating that user judgement and interpretation of the given input were important. The present study examined the user's influence on predictions of terrestrial food chain models systematically.

2. Design and implementation of the study

Three terrestrial food chain codes, driven by deposition of radioactivity from the atmosphere, were obtained and run by participants for several scenarios involving ^{131}I and ^{137}Cs transfer from air to pasture and milk. The code manuals and scenario descriptions are documented in a technical report (BIOMOVS II, 1996) that also includes analyses of predictions submitted for both radionuclides. This paper summarizes the results obtained for the ^{137}Cs scenarios. Results and conclusions drawn from the ^{131}I scenarios are almost identical.

In September 1994, one code and two scenario descriptions were sent out to interested users. The third scenario followed six months later. In all, ten participants submitted predictions with the first model. In November 1995, two more codes were distributed, and most participants submitted results from one, while a few submitted results from the other. Three participants made calculations with all three codes. Participants were to use the same assumptions about a given scenario for all three codes, except when necessary to adjust for different parameters used by the individual codes. Participants were asked to explain their assumptions and justify choices of parameter values to assist in understanding of differences between predictions. In March 1996, the working group met for the presentation of predictions and comparison of these predictions with experimental data.

Most participants in this study are familiar with modelling concepts and models of the environmental transport of radioactivity. Discussion amongst participants was forbidden, although seeking expert opinion from non-participants was encouraged. Questions regarding interpretation of codes or scenarios were only answered if this did not have an effect on assumptions. Answers were distributed to all participants.

2.1. Codes

All three codes were adapted for this study so that predicted endpoints corresponded to endpoints of the scenarios: deposition and concentrations in pasture and milk. Manuals directed participants how to run the codes and how to change parameter values, but contained no information on the way the models were formulated or implemented numerically.

CHERPAC (Peterson, 1994) is a stochastic code which predicts the movement of radionuclides through the food chain to dose after an accidental release of radioactivity. Wet and dry depositions, and concentrations on pasture (fresh weight) and in milk, are calculated daily or as monthly averages.

RUINS (Crout, 1991) was developed to investigate soil fixation and plant uptake of ^{137}Cs in upland organic soils of the UK after the Chernobyl accident. It is a deterministic code. Output was modified to include concentrations on pasture (dry weight) and in milk.

CLRP (Krajewski and Pietrzak-Flis, 1995) predicts the movement of radionuclides through the food chain to dose after chronic or accidental releases of radioactivity. Output, with confidence limits, is given for concentrations on pasture (fresh weight) and in milk.

2.2. Scenarios

The starting point for each scenario was time-dependent radionuclide concentrations in air and precipitation. From these, deposition of radionuclides to pasture and transfer to cow's milk occurred. General background data (e.g. yields of pasture, monthly temperatures, generalized diets) were provided when available to assist the user's choices of parameter values. All scenarios included observed data that matched the endpoints asked for in the scenario description. These data were made available only after all calculations had been submitted. Short summaries of the two ^{137}Cs scenarios follow.

The driving data for the *BREMEN* scenario were Chernobyl derived concentrations of ^{137}Cs aerosols in air at Bremen, Germany for 29 April–5 May, 1986. Daily amounts of precipitation were provided for the seven days that the plume was present. Background information was given on the diet and pasture rotation for the one cow from which milk samples were taken. Participants were asked to calculate total deposition, and daily and time-integrated concentrations on pasture (fresh weight) and in milk up to 25 October 1986.

The *FORT COLLINS* scenario utilized monthly (April to August) average air concentrations of ^{137}Cs from nuclear weapons fallout, daily amounts of rainfall, and deposition calculated from the rainfall for 1963 and 1965 in Colorado, USA. Participants were asked to calculate concentrations in dry weight pasture (1963), alfalfa (1965) and milk (1963 and 1965) for the months of May to August. Both monthly average and time-integrated predictions were requested.

3. Results and discussion

3.1. Predictions using the CHERPAC code

3.1.1. BREMEN scenario

Fig. 1 shows the predictions of total ^{137}Cs deposition and time-integrated concentrations on pasture grass and in milk, along with the 95% confidence intervals for measured quantities. Observed deposition was determined from nine samples taken at random from a field during the two months after the Chernobyl accident. Confidence intervals of the observed concentrations on pasture and in milk were obtained from fitting a multi-exponential model to the individual measurements (Kirchner, 1992).

Total deposition tends to be underestimated, while the fraction retained on pasture is overestimated by most predictions. About half the results also overestimate the transfer to milk, and some of the milk concentrations close to the observed values are due to compensation for low prediction of deposition. The spread of predicted total deposition (best estimates) among the ten users is approximately one order of magnitude, but reaches a factor of 40 for concentrations in milk.

About a fourth of calculated 95% confidence intervals do not intersect the range of measurements. Most confidence intervals are smaller than the variability between the best estimate predictions of the participants.

Predictions of daily concentrations on pasture and in milk are shown in Figs. 2 and 3. For most of the predictions on grass the slopes reflect the choice of value for weathering half-life. A deviation from exponentional decline at low concentrations indicates the presence of root uptake, but this is important for participant B only. Milk concentrations should show a steep increase during the first days, since the cow was stabled and received uncontaminated feed until 14 May. Obviously, this was not modelled by all participants. This partly explains some of the high ratios of milk/pasture concentrations apparent in Fig. 1.

The participants' best estimates of parameter values for CHERPAC are shown in Table 1. There is considerable variability among the values, reflecting differing

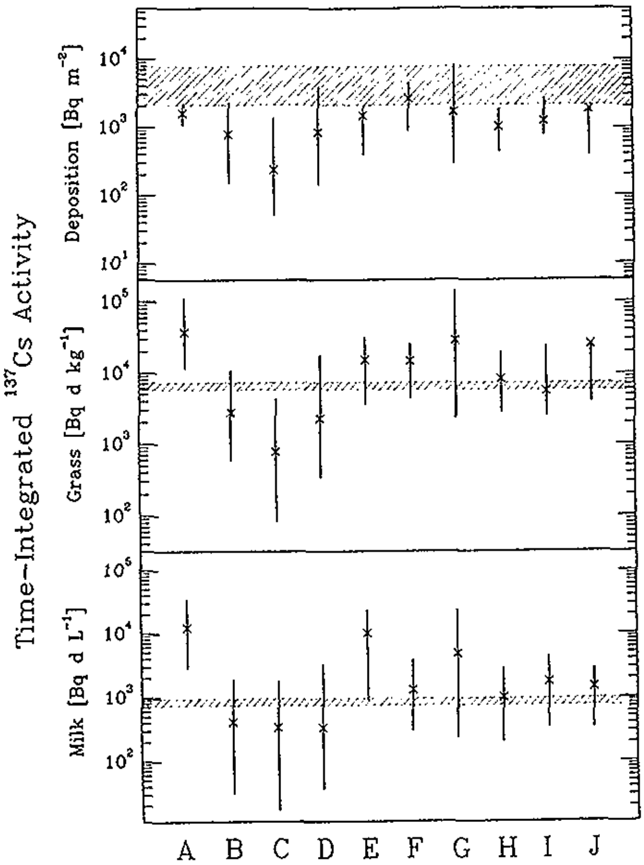

Fig. 1. Total deposition and time-integrated concentrations of ^{137}Cs on pasture grass (fresh weight; 5 May – 27 June) and in milk (14 May – 27 June) for Bremen; capital letters along the abscissa denote the participants; best estimates and 95% confidence intervals on the predictions calculated with CHERPAC are given; hatched areas denote the 95% confidence intervals of the field measurements.

interpretations of modelled processes. For example, values for the yield of pasture vary by more than a factor of 20. After discussing the results, participant A noticed that the value 0.23 kg m^{-2} was in dry weight rather than the required wet weight. Correcting this error improves this participant's predictions. In contrast, participant G chose a low value to account for low grass productivity in early spring based on local climatic conditions in the participant's country, rather than in northern Germany.

Parameter values were based on Chernobyl fallout data, which are directly applicable to the BREMEN scenario. Most of the participants chose data collected in their own countries. Only a few participants explicitly stated that they believed their parameter values to be representative of the Bremen area. This will contribute to some variability of predictions, mainly due to climatic differences. On the other hand,

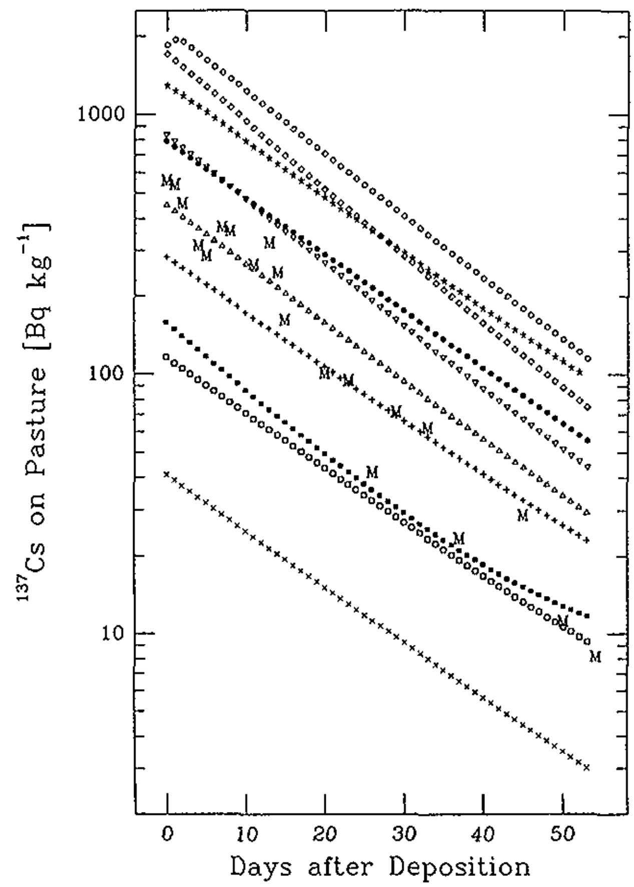

Fig. 2. Daily concentrations (5 May–27 June) of ^{137}Cs on pasture grass (fresh weight) for Bremen; best estimates calculated with CHERPAC are given; "M" denotes measurements; symbols used are: ○ for participant A, ■ for B, × for C, □ for D, • for E, ▽ for F, ◇ for G, △ for H, + for I and * for J.

it indicates that most participants are prepared to use site-specific data for radioecological evaluations in the aftermath of any nuclear event affecting their country.

The influence of parameter values on concentrations of ^{137}Cs in milk, as ranked by the users, is compared in Table 2 with results from a sensitivity analysis performed with CHERPAC using parameter distributions based on the range of best estimate values. This comparison shows that the users as a group had a good a priori understanding of the importance of the parameters in CHERPAC, although some deviations are apparent (e.g., ranking of the particulate dry deposition velocity was high and of the yield of pasture low compared to the results of the sensitivity analysis). A user's choice of parameter values does not always reflect an understanding of the scenario, however. For example, although judging the importance of root uptake as

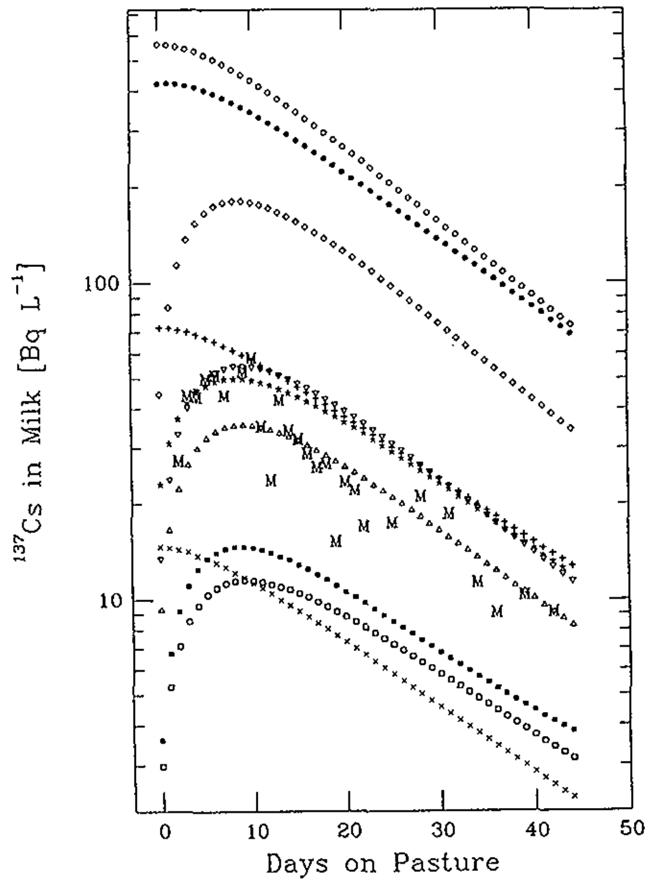

Fig. 3. Daily concentrations (14 May – 27 June) of ^{137}Cs in milk for Bremen; best estimates calculated with CHERPAC are given; "M" denotes measurements; symbols used are: ○ for participant A, ■ for B, × for C, □ for D, ● for E, ▽ for F, ◇ for G, △ for H, + for I and * for J.

low, participant B selected a high value for this parameter. This high concentration ratio was taken from a data base with widely differing soil types, many of which were not relevant for this study.

Distributions of the parameter values had to be specified by each user for uncertainty analysis in CHERPAC. Most participants fit their choice of parameter values to the distributions given in the sample input supplied with the code package rather than change the distribution type. Some participants gave little thought to whether or not the distributions used were reasonable.

After submission of predictions, an uncertainty analysis with parameter distributions based on the range of best estimate values used by the participants (Table 1) was performed with CHERPAC. The 95% confidence intervals corresponded to the range

Table 1
Best estimates of ^{137}Cs specific parameters used with CHERPAC for the BREMEN scenario

Parameter		Participant A	B	C	D	E	F	G	H	I	J
Yield of pasture[a]	($kg\,m^{-2}$)	0.23	1.77	5.0	1.6	1.2	1.5	0.5	1.0	1.204	1.0
Density of 1 cm soil layer	($kg\,m^{-2}$)	15	13.6	14	16	13.5	20	7	20	13.5	15
Hay consumption by cow[a]	($kg\,d^{-1}$)	60	48.5	70	56	112	60	60	70	50	45
Intake of grain in summer	($kg\,d^{-1}$)	1.5	6.2	7	0.5	0	0.5	6	3	0	1
Intake of grain in winter	($kg\,d^{-1}$)	1.5	6.2	7	0.5	0	20	6	2	7	3
Daily soil ingestion	($kg\,d^{-1}$)	0.7	0.002	0.5	0.56	0.8	0.2	1	0.1	0.6	0.1
Breathing rate of cow	($m^3\,d^{-1}$)	150	130	125	170	200	130	100	250	130	200
Particulate dry deposition velocity	($10^{-3}\,m\,s^{-1}$)	2.2	1.0	1.5	1.2	2.0	1.5	2.0	1.0	0.5	7.0
Particulate washout ratio	($10^5\,m^3\,m^{-3}$)	6.2	4.0	0.1	5.8	6.0	3.0	7.0	5.0	5.6	0.5
Concentration ratio for pasture[a]		0.01	0.108	0.01	0.02	0.0018	0.03	0.02	0.05	0.03	0.05
Weathering loss rate	(d^{-1})	0.055	0.0624	0.05	0.04951	0.05	0.057	0.06	0.05	0.0493	0.05
Wet interception factor		0.15	0.25	0.2	0.1	0.6	0.35	0.45	0.3	0.25	0.1
Transfer factor for milk	($10^{-3}\,d\,kg^{-1}$)	7.3	5.4	7.9	4.3	7.9	3.0	5.0	4.0	8.0	2.0

[a] Plant fresh weight.

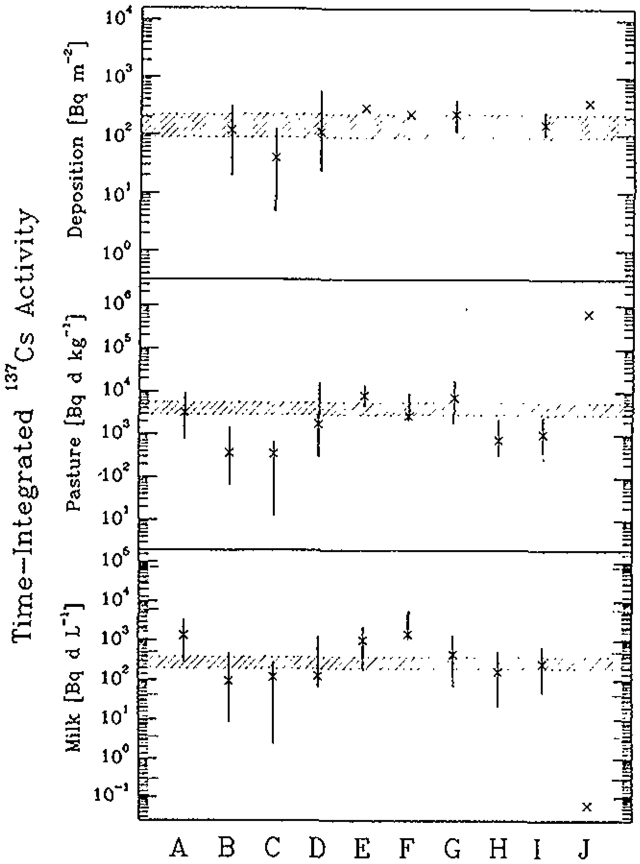

Fig. 4. Total deposition and time-integrated (1 May – 31 August, 1965) concentrations of ^{137}Cs on pasture grass (dry weight) and in milk for Ft. Collins; capital letters along the abscissa denote the participants; best estimates and 95% confidence intervals (if submitted) on the predictions calculated with CHERPAC are given; hatched areas denote the estimated range of measured concentrations.

order of magnitude for time-integrated concentrations, both on alfalfa and in milk. Estimated uncertainties do not always overlap.

Most participants used the same parameter values and distributions for Ft. Collins as they did for Bremen except for yield which most changed. Several participants changed hay consumption, dry deposition velocity and washout ratio. Two participants selected a higher feed-to-milk transfer coefficient based on higher availability of nuclear weapons fallout cesium, and participant F changed all parameter values.

3.2. Predictions using the RUINS and CLRP codes

Due to time constraints, only a limited number of calculations were submitted, so emphasis here will be placed on predictions of time-integrated concentrations for the BREMEN scenario.

Table 2
Comparison of the participants' ranking of the influencial parameters used for the BREMEN scenario to the results of a sensitivity analysis of CHERPAC; parameter distributions for the sensitivity analysis were based on the range of best estimates given in Table 1

Parameter	Influence[a]				Partial rank correlation coefficient
	Negligible	Low	Medium	High	
Yield of pasture	0	1	8	1	−0.80
Density of 1 cm soil layer	4	5	1	0	—[b]
Hay consumption by cow	0	1	4	5	0.59
Intake of grain in summer by cow	5	3	1	1	—
Intake of grain in winter by cow	7	2	0	1	—
Daily soil ingestion	4	4	2	0	—
Breathing rate of cow	8	2	0	0	—
Particulate dry deposition velocity	0	0	5	5	0.36
Particulate washout ratio	0	1	3	6	0.84
Concentration ratio for pasture	2	5	2	1	0.51
Weathering loss rate from pasture	0	1	6	3	−0.41
Wet interception factor for pasture	1	0	4	5	0.52
Transfer factor to milk	0	0	1	9	0.73

[a] Number of votes by the participants.
[b] Correlation coefficient (absolute) < 0.2.

of the best estimate predictions shown in Fig. 1, indicating that variability in predictions can be largely explained by the variation in parameter values.

3.1.2. FORT COLLINS scenario (1965)

Predictions of total deposition and time-integrated concentrations of ^{137}Cs on pasture and in milk from May to August of 1965 are shown in Fig. 4. The confidence interval on observed deposition reflects discrepancies between data from two monitoring systems with different samplers (BIOMOVS II, 1996). Estimated uncertainties on observed time-integrated concentrations on alfalfa and in milk are intended to account for sampling variability, measurement error and missing data. Concentrations on pasture are given in dry weight.

Because of conflicting information on deposition, participants A, B, C and I ignored it and calculated deposition from given air concentrations. The other participants used differing procedures to scale daily deposition measurements to correspond to the monthly data of one of the samplers. These values have no confidence intervals, since they were used as input data to CHERPAC.

Four participants (A, B, F, I) submitted concentrations for wet weight alfalfa. Converting their results to dry weight would improve predictions of participants B and I. The transfer from alfalfa into milk is overestimated slightly by most users. Predictions of participant J reflect difficulties with CHERPAC and this scenario. Ignoring these predictions, the variability between best estimates is still more than one

Predicted to observed (P/O) ratios of total deposition and time-integrated concentrations of ^{137}Cs on pasture and in milk are shown in Table 3 for RUINS and CLRP. P/O ratios are also shown for CHERPAC for comparison.

Deposition and concentrations on pasture generally are underestimated in both RUINS and CLRP. The P/O ratio close to 1 for milk of some RUINS predictions and one CLRP prediction merely result from compensating under- and overpredictions of the individual transfer processes. Variation between mean predictions of deposition for both models was quite small (neglecting participant C's predictions with RUINS) and was about one order of magnitude for time-integrated concentrations on grass and in milk.

Varying amounts of contaminated forage in daily feed cannot be specified in RUINS, unlike CHERPAC and CLRP. Thus, most of the P/O ratios for milk cannot be directly compared among the three models. Participants B and F circumvented this limitation of RUINS by calculating cesium concentrations on pasture and then repeating the calculations assuming now that a single deposition occurred the day before the cow was put to pasture. Deposition was adjusted so that pasture concentrations matched those calculated previously and then used for calculating concentrations in milk.

As requested, participants used single best estimate values for each parameter for all three models when possible. Participant F, whose feed-to-milk transfer coefficient for RUINS was higher than for CHERPAC, was the exception. This explains why predictions of this participant overestimate transfer from pasture to milk using RUINS but underestimate it using CHERPAC.

3.3. General aspects

The three codes used in this study are very different in structure and user-friendliness, and this presented many difficulties to the participants. A comparison of results is not straightforward since assumptions sometimes varied considerably between users (e.g., how diet composition was handled in RUINS). Some users wrestled with a particular code but still had little confidence in their results. To become familiar with a code before starting the studies, three participants performed calculations for a scenario familiar to them. Only one participant used a code not involved in this study to cross-check some of the predictions before submission.

All participants were asked to rank their confidence in the predictions they submitted. Highest confidence was expressed in predictions for the BREMEN scenario using CHERPAC. This is probably due to the amount of information available after the Chernobyl accident. The higher confidence shown in the calculations performed using the CHERPAC code may reflect lack of familiarity with the other two codes because of time constraints. The subjective confidence given to the predictions, however, does not correlate with the results. The Bremen predictions are no closer to observations than the Fort Collins predictions; neither are predictions by CHERPAC more accurate than predictions by RUINS or CLRP.

Table 3
Predicted to observed ratios of total deposition and time-integrated concentrations on pasture and in milk for the BREMEN scenario

Participant	Total deposition			^{137}Cs on pasture[a]			^{137}Cs in milk[b]		
	CHERPAC	RUINS	CLRP	CHERPAC	RUINS	CLRP	CHERPAC	RUINS	CLRP
A	0.32	0.34	0.33	5.69	0.29	0.52	14.56	0.42	0.68
B	0.16	0.44	—	0.42	3.27	—	0.48	0.93	—
C	0.05	0.003	—	0.12	0.18	—	0.40	0.93	—
D	0.17	0.18	0.31	0.34	0.15	0.12	0.39	0.13	0.87
E	0.29	0.35	—	2.30	0.49	—	11.85	2.15	—
F	0.52	0.16	0.43	2.22	0.53	0.21	1.52	1.64	0.12
G	0.33	0.35	—	4.32	2.48	—	5.58	0.90	—
H	0.20	0.23	—	1.27	0.61	—	1.15	0.82	—
I	0.24	—[c]	—	0.84	—	—	2.01	—	—
J	0.36	—	0.33	3.82	—	1.14	1.66	—	0.34

[a] 5 May – 27 June.
[b] 14 May – 27 June.
[c] No prediction submitted.

4. Conclusions and recommendations

1. The global uncertainty (i.e., lowest 2.5% to highest 97.5% confidence limits over all predictions) is greater than the confidence interval of the predictions of any single user.
2. Choice of parameter values contributes most to user-induced uncertainty followed by scenario interpretation. Contribution of code implementation is unknown, but may be low because the majority of the group did not apply the most complex of the three codes.
3. Users either did not try or were not able to get consistent results across codes. This should have been possible given code-dependent assumptions independent from the parameter values held in common among codes.
4. The magnitude of uncertainty due to users' interpretations is enormous, and even for simple scenarios the explanation for differences may not be obvious.
5. Confidence of the user in understanding of a scenario description and/or confidence in working with a code does not necessarily mean that the predictions will be more accurate.
6. Results do not necessarily reflect the difficulties in working with codes or scenarios. "Good" predictions could be the result of chance, and/or of compensation of errors in predictions of individual transport processes.
7. Examples provided in the code manuals may influence code users considerably when preparing their own input files.
8. Mistakes, such as the extremely common one of confusing fresh and dry weights between scenario and model, contribute significantly to variation in predictions and have to be taken into account because since they cannot be separated when results are reviewed.

Conclusions 1–4 above may be restricted to the processes that were addressed in this study, while the others are believed to be general and independent of the type of models and processes studied.

Although the effect of the users' assumptions on predictions was expected to be high, the extreme variability from ten users with one model came as a surprise. It is all the more alarming since the model for all scenarios was very simple with linear transfer between just three compartments, and participants had experience in this sort of modeling. More disparate results could be expected when using a more complex model.

To reduce the effect of user interpretation on model output, assessments should not be carried out in isolation. At the least, calculations should be performed by a team with input from several experts in each modeling area. In some cases it would be preferable for calculations to be carried out by more than one team, perhaps using more than one code, with the final predictions reflecting a consensus reached by understanding and resolving differences in individual results. Each model user should be familiar with the code employed. Units requested in a scenario description and units given in code output should be checked for consistency. Examples provided with a code should be chosen and described very carefully so as not to bias subsequent use. It should be understood that confidence intervals about the predictions calculated from parameter distributions do not include all uncertainties.

Acknowledgements

The study participants are grateful to Neil Crout (Nottingham University, United Kingdom) for making available modified versions of his code and to Marian Scott (University of Glasgow, Scotland) for many valuable discussions. The valuable support given by Frank Brünjes (University of Bremen, Germany) in preparing the figures is gratefully acknowledged. Participants are also grateful for recources made available for meetings and report production through the BIOMOVS II Steering Committee. The general idea of this study was originally suggested by Robert H. Gardner (Oak Ridge National Laboratory, USA).

References

Ahn, K. I., & Jin, Y. H. (1996). A formal approach for quantitative treatment of modeling uncertainties in safety analyses. *Nuclear Technology, 116*, 146–159.

Bergström, U., & Nordlinder, S. (1991). Uncertainties related to dose assessments for high level waste disposal. *Nuclear Safety, 32*, 391–402.

BIOMOVS II (1996). Effect of user interpretation on uncertainty estimates. BIOMOVS II Technical Report No. 7, Swedish Radiation Protection Institute, S 171 16 Stockholm, Sweden.

Botterweg, P. (1995). The users influence on model calibration results; an example of the model SOIL, independently calibrated by two users. *Ecological Modelling, 81*, 71–81.

Crout, N. M. J. (1991). RUINS: A simulation for radiocaesium behaviour in grazing systems – User manual, Report ES0109. Department of Physiology and Environmental Sciences, University of Nottingham.

Hui, T. E., Loesch, R. M., Raddatz, C., Fisher, D. R., & McDonald, J. C. (1994). An internal dosimetry intercomparison study. *Health Physics, 67*, 217–225.

Kirchner, G. (1992). A method for modeling the transfer of radionuclides deposited after the Chernobyl accident via the grass-cow-milk pathway. *Modelling Geo-Biosphere Processes, 1*, 13–22.

Köhler, H., Peterson, S.-R. & Hoffman, F. O. (1991). Scenario A4. Multiple model testing using Chernobyl fallout data of I-131 in forage and milk and Cs-137 in forage, milk, beef and grain. BIOMOVS I Technical Report No. 13, Swedish Radiation Protection Institute, S 171 16 Stockholm, Sweden.

Krajewski, P., & Pietrzak-Flis, Z. (1995) CLRP model descriptions and individual evaluation of model predictions. In *Report of the First Test Exercise of the VAMP Multiple Pathways Assessment Working Group* (pp. 241–252). International Atomic Energy Agency, Vienna, IAEA-TECDOC-795.

Loehle, C. (1987). Errors of construction, evaluation, and inference: a classification of sources of errors in ecological models. *Ecological Modelling, 36*, 297–314.

Müller, H., & Pröhl, G. (1993). ECOSYS-87: A dynamic model for assessing radiological consequences of nuclear accidents. *Health Physics, 64*, 232–252.

Peterson, S.-R. (1994). Model description of CHERPAC (Chalk River Environmental Pathways Analysis Code). Result of testing with post-Chernobyl data from Finland. Report AECL-11089, Atomic Energy of Canada Limited, Chalk River.

Sheng, G., & Ören, T. I. (1990). Computer-aided software understanding systems to enhance confidence of scientific codes. In *On the validity of environmental transfer models* (pp. 275–286) Swedish Radiation Protection Institute, Stockholm.