INTEGRATION OF DATA ASSIMILATION SUBSYSTEM INTO ENVIRONMENTAL MODEL OF HARMFUL SUBSTANCES PROPAGATION

Petr Pecha and Radek Hofman
Institute of Information Theory and Automation, Prague, Czech Academy of Sciences

MOTIVATION

The paper describes present state of implementation of assimilation subsystem within software system HARP designed as decision support tool for fast assessment of radiological consequences of accidental releases of radionuclides into living environment. The product is presented here from viewpoint of its architecture and possibility of its utilization such an educational tool for training purposes in radiation protection field and for decision support staff. A review of assimilation methods implemented in the subsystem is presented here.

ASSIMILATION METHODS

The goal of data assimilation is to provide analysis which relies on so called background field from a model forecast and observations. Other inputs to data assimilation process can be physical constraints on the problem or any additional prior knowledge not included in the model. Merging of these contending sources of information had shown to be very promising in many branches of contemporary Earth sciences. In data assimilation we try to adjust model according to measured values what represents research effort to move from isolated model predictions forward reality. An automatic procedure for bringing observations into the model is called objective analysis. The major progress of objective analysis was achieved in the field of meteorological forecasting techniques that represent efficient tool in struggle with tendency to chaotic destruction of physical knowledge. Advanced assimilation methods are capable to take into account measurement and model errors in form of error covariance matrices.

Palette of assimilation techniques. The assimilation subsystem of HARP system contains implementation of several methods of objective analysis and provides their utilization in radiation protection via user-friendly graphical interface. Palette of assimilation methods is being systematically extended and in current state of art it includes those from pure interpolation methods over empirical method of successive corrections up to statistical method of optimal interpolation (OI) that can handle model and measurement errors. Physical knowledge embodied within the model predictions usually enters into procedure of data assimilation such a vector of background field values $\mathbf{x_b}$ of dimension n (n = number of analysed grid points × number of analysed quantities). Let vector of available measurements $\mathbf{y}$ has dimension p. Analysis in each grid point is represented by vector $\mathbf{x_a}$ of dim n. Common principle of objective analysis can be expressed by relation

$$\mathbf{x_a} = \mathbf{x_b} + \mathbf{Wd} \qquad (1)$$

Equation (1) expresses update step of data assimilation process. It says that we obtain analysis if we take background field vector and add to it product of matrix **W** and vector **d**. **W** is so called weight matrix of dimension $n \times p$ and **d** is vector of innovations given by differences between values of measurement and values of the model. Model values are known in given discrete points and if we want to know value of model in arbitrary point we need to use forward observation operator H, which transforms points from model space into measurement space, so the differences can be then evaluated as:

$$\mathbf{d} = \mathbf{y} - H\mathbf{x} \qquad (2)$$

Modelling in the HARP system is performed on polar mesh comprising of 2800 analysed points and let analyse only one quantity, so dimension of state space is 2800. This number is rather high and may cause some inconveniences in implementation of advanced assimilation methods. Procedure of assimilation is repeated as assimilation cycle in which analysis $\mathbf{x_a}$ in time step t_i serves as new background field for prediction in time step t_{i+1}. The evaluation of analysed quantity in time t_{i+1} according to the quantity value in time t_i is called forecast step.

Classical interpolation. In case of interpolation methods the analysis is constructed only upon measurements of an analysed quantity, so we omit $\mathbf{x_b}$ in eq. (1) which reduces to $\mathbf{x_a} = \mathrm{W}\mathbf{y}.$ Here W represents an interpolation operator. Because of sparse data caused by lack of measurements the analysis problem is often under-determined and good approximation of true state can’t be obtained.

Empirical techniques taking into account model knowledge on a certain level. We obtain much better results when analysis is based on prior information provided by background field. The first method implemented in the HARP code that is capable to consider background field knowledge is Successive Correction Method (SCM). SCM can be considered as old-fashioned but it is still useful because of its low cost computational demands. SCM is empirical assimilation method, so it can’t take into account neither errors of model nor errors of measurements. Only a certain empirical expert knowledge can be introduced through tuning parameter ε which expresses estimated ratio of measurement error dispersion and dispersion of background field error. The first iteration step of analysis $\mathbf{x}^a$ in jth gird point can be obtained as

$$\mathbf{x_j^a} \;=\; \mathbf{x_j^b} + \left(\sum_{\mathbf{k}=1}^{\mathbf{K(j)}} \mathbf{w_{jk}} \cdot (\mathbf{y_k} - \mathbf{x_k^b})\right) \Bigg/ \left(\sum_{\mathbf{k}=1}^{\mathbf{K(j)}} \mathbf{w_{jk}} + \varepsilon^2\right) \qquad (3)$$

Here w_{ik} are empirically determined weights and we sum products of these weights and differences between measurements y_k and model values in locations of measurements over all measurements in so called region of influence around analysed point (with number of relevant observations K(j)). The SCM proved that it can be successfully used for local and preliminary assimilation. Besides inability of SCM to handle model and observation errors, the method has appeared as “blind” with regard to some important local effects of uncertainty propagation (see scenario with local rain region in the poster presentation). Our uncertainty analysis consistently distinguishes between global and local spatial effect of input random parameters what reflects in model error covariance matrix and then the statistical techniques (e.g. OI – see poster) can suppress possible spurious spatial correlations of the results.

The second useful method is based on minimisation procedure of direct search algorithm. Instead of probing classical interpolation techniques (such are polynomial fitting or cubic-spline functions) we have consistently applied a model which naturally respects physical knowledge with direct relation to the Gaussian plume model shape. The model have to be run in probabilistic mode when fluctuations of four selected input parameters influence the shape of the resulted straight-line Gaussian surface over the terrain (e.g. activity deposition of a certain radionuclide) in four permissible manners (Gaussian surface translation, horizontal squeezing, rotation, steepness in advection direction). Minimisation procedure repeatedly calls the model with different sets of the “manipulation” input vector values in order to achieve the best fit (least squares) of modified shape of model surface with observations. Provided that the observations are simulated by the same physical model, the procedure is sometimes called as “twin experiment”. In the poster part of this presentation is shown an extension of minimisation technique to segmented Gaussian plume model (SGPM) reflecting radiological situation just after 3 hours after the release start. In addition, 3 hourly forecasts of wind velocity vectors were modelled as random and then the segmented respond surface shape was function of 9 random input "manipulation" parameters. The tests based on "twin-experiment" scheme have given satisfactory results which predetermine the method at least for preprocessing of initial conditions for better convergence of advanced assimilation methods. It can be interpreted like automatic data tuning device for optimum initial adjustment of model background field with regard to observations. It should increase robustness of management procedures against possible fatal user errors arising from decision under stress during emergency situations).

Constant statistical techniques of objective analysis. Optimal interpolation (OI) method is a representative of frequently used statistical technique working with time-constant model error covariance matrix. Inputs to the procedure of optimal interpolation are vectors of background field and measurements and information on background and measurement errors in form of error covariance matrices. Data update procedure performed by OI can be expressed by equations (1, 2) where form of weight matrix **W** results from least squares analysis:

$$\mathbf{W} = \mathbf{B}\mathbf{H}^{\mathbf{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathbf{T}} + \mathbf{R})^{-1} \qquad (4)$$

Matrices **B** and **R** are covariance matrices of the model error and measurement error and **H** is matrix of linearised observation operator. Optimal interpolation minimizes error of resulting analysis and generates covariance matrix of analysis error $\mathbf{P}_a$:

$$\mathbf{P_a} = (\mathbf{I} - \mathbf{WH})\mathbf{B} \qquad (5)$$

I is identity matrix and upper indices T and –1 mean operators of transposition and inversion.

Advanced statistical assimilation methods. Provided that the time-evolution modelling of forecast error covariance is put into effect, the advanced techniques of sequential assimilation of observations can be done by recurrence over observation times t_i. More advanced methods based on Bayesian approach, variational methods and a certain alternatives of Kalman filter techniques are topics of current investigations supported by the Czech Grant Agency. Specific option of proper methods will emerge from our requirements which consequences of accidental release scenario should be analysed. Let us briefly mention the case when long-term radioactivity deposition of radionuclide Cs137 on the terrain is a matter of interest.

Let initial conditions are related to time t_0 when radioactive plume has just leaved the terrain. The initial conditions are given by complex solution of the previous period of atmospheric dispersion and deposition phase, the result of which are background field $\mathbf{x_b}^{Cs137}(t{=}t_0)$ and background error covariance matrix $B(t{=}t_0)$. Moreover, we can expect measurements incoming just at time t_0 and use objective analysis according to equations (1) and (4). Then, OI analysis of activity deposition is done and its results are represented by "corrected model prediction" $\mathbf{x}_a^{Cs137}(t{=}t_0)$ and its analysis error covariance $P_a(t{=}t_0)$ according to equation (5). Both structures are now the initial conditions for the successive treated long-term scenario of Cs137 activity deposition development with account for incoming measurements (e.g. with yearly period). The new scenario starts to evolve. For step-wise assimilation of incoming observations with model predictions have to be used advanced statistical techniques.

SYSTEM ARCHITECTURE

The aim is to develop modelling, simulation and educational tool with unified user-friendly graphical interface for utilization in radiation protection. The system has 4 principal components. The main component which performs numerical evaluation (numerical module) of all considered radiological quantities utilizes SGPM model of radioactive pollution propagation, which is capable to take into account short-term meteorological forecasts (hourly changes) provided by the Czech meteorological service. Propagation of radionuclides in atmosphere can be simulated up to 100 kilometres from the source of pollution.

At present time deterministic and probabilistic version of HARP system is available. Deterministic version respects variability concept wherever possible. Generated single endpoint values correspond to the best estimation of all input parameters (including effect of variability) of atmospheric dispersion and deposition model (ADM), ingestion model (FCM - Food Chain Model) and dose model (DOS). Successive deterministic consequence assessment insists in testing of single output values against limits given by mandatory regulatory rules.

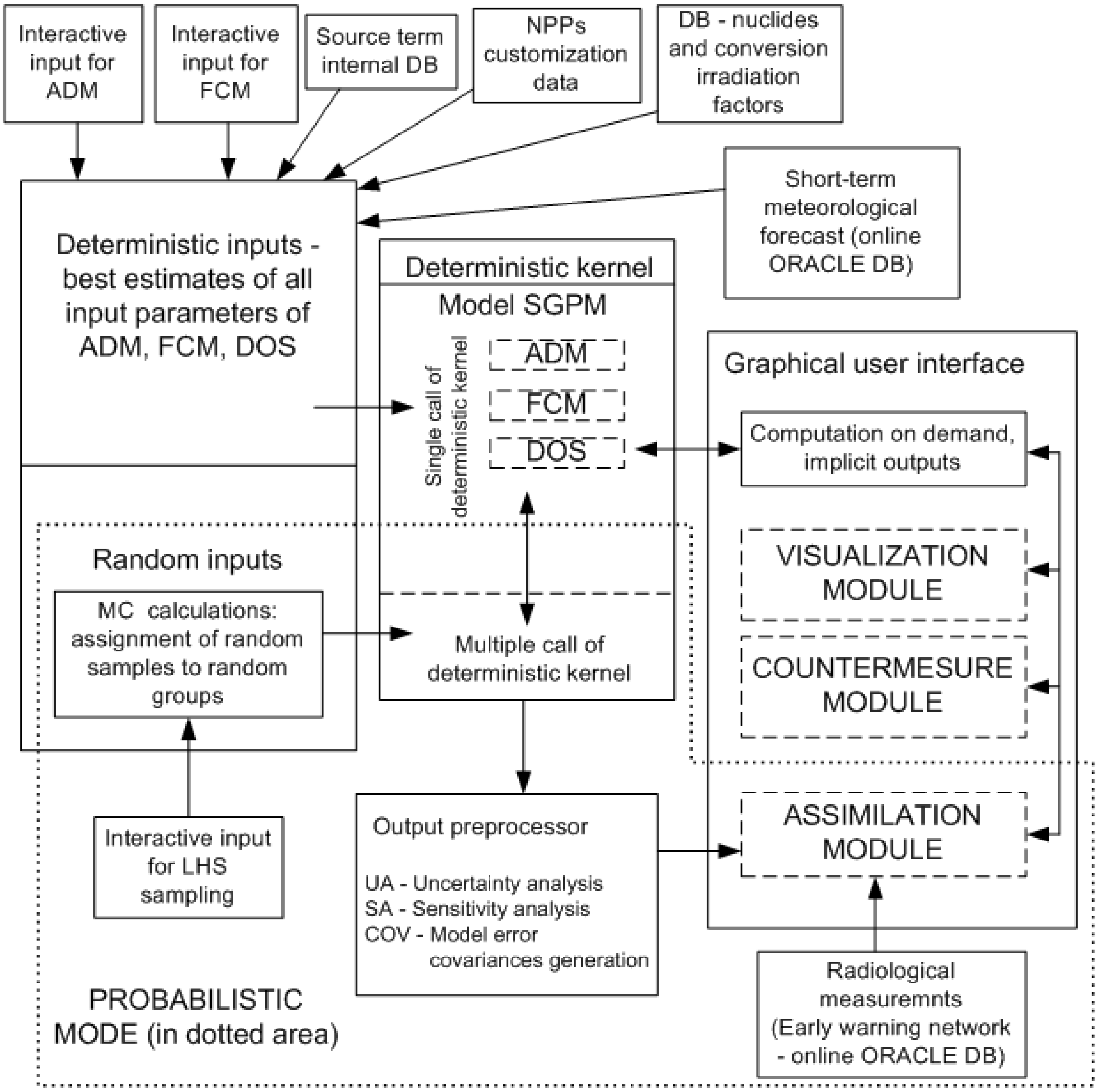

Fig. 1; Integration of assimilation subsystem into the structure of the HARP code.

However the variability concept is accepted, the deterministic risk assessment approach mentioned above did not comply with inherent uncertain character of the problem. Unlike this the probabilistic access takes into account stochastic character of a certain input model parameters, imperfections in description of physical nature of submodels, uncertainties in release scenario formulation, simplifications in computational procedure etc. Probabilistic approach ensures possibility of introduction of modern methods when probabilistic answers can be generated on consequence assessment questions. At the same time, uncertainty treatment represents unavoidable condition for application of advanced assimilation methods.

Numerical module is interconnected to Graphical User Interface (GUI) which binds together remaining three modules: Visualization module, Countermeasure module and Assimilation module. Communication between numerical module and GUI is duplex and besides implicitly evaluated set of output quantities the computations on demand can be also invoked. The user can via graphical form raise a query and transfer it for processing to the main module waiting at background. The system offers various alternative options of input parameters selection for release scenarios in their atmospheric, deposition, ingestion and dose parts. For that reason the software product can serve as a training tool enabling responsible staff to improve their knowledge and perception of the problem details.

Visualization module provides visualization of results on relevant scalable map background. All analysed quantities can be plotted into map as coloured isopleths. Levels of an isopleths can be determined automatically or fully user defined settings can be used. Isopleths are blended to map background and user can select ratio of transparency. Geographical coordinates, distance from NPP and actual level of analysed quantity in certain location are shown in bottom status bar according to current mouse pointer position on the map. For evaluating of quantities besides discrete points of polar mesh is used algorithm of bilinear interpolation. Various objects (measuring stations etc.) can be also drawn to map on user defined coordinates basis and the icons representing these objects can be fully customized (shape, color, hint on mouse move). In case of modelling of hourly changes of a quantity an animation can be projected and for example movement of radioactive plume over terrain can be traced. Map with all visualized quantities can be saved as raster image in JPEG format.

Assimilation module is closely interconnected to visualization part of the system and this binding brings many advantages, which may save a time during emergency situations and make assimilation task clearer. There are also other embedded tools for detail analysis of results. It is possible to adapt the whole system to local conditions of arbitrary NPP, i.e. change map background, coordinates of a nuclear facility, terrain type of surroundings and many other governing local parameters. System is primarily customised for NPPs Temelín and Dukovany.

ACKNOWLEDGEMENTS: This work is part of the grant project GAČR No. 102/07/1596, which is founded by Grant Agency of the Czech Republic. A lot of useful knowledge has been also acquired during RODOS customisation procedure for the Czech territory.

REFERENCES

Drécourt, J-P., 2004: Data assimilation in hydrological modelling. Ph.D. Thesis, DHI Water&Environ. Dpt.

Kalnay E., 2003: Atmospheric modeling, data assimilation and predictability. Cambridge Univ. Press, ISBN 0-521-79179-0.

Hamill, T. M., 2006: Ensemble-based atmospheric data assimilation. In: Predictability of Weather and Climate, T. Plamer and R. Hagedorn edts. Cambridge Univ. Press.

Astrup, P., C. Turcanu, R.O.Puch, C. Rojas Palma and T. Mikkelsen, 2004 : Data assimilation in the early phase: Kalman filtering RIMPUFF. Risø NL, Risø-R-1466(EN).