Hoisting a Red Flag: An Early Warning System for Exceeding Subsidence Limits

Article *in* Mathematical geosciences · February 2010

DOI: 10.1007/s11004-009-9252-2

CITATIONS
8

READS
107

3 authors:

Manuel Nepveu
TNO
55 PUBLICATIONS **298** CITATIONS
SEE PROFILE

I. C. Kroon
TNO
17 PUBLICATIONS **69** CITATIONS
SEE PROFILE

Peter A. Fokker
TNO
102 PUBLICATIONS **644** CITATIONS
SEE PROFILE

Math Geosci (2010) 42: 187–198
DOI 10.1007/s11004-009-9252-2

Hoisting a Red Flag: An Early Warning System for Exceeding Subsidence Limits

M. Nepveu · I.C. Kroon · P.A. Fokker

Received: 29 September 2008 / Accepted: 26 October 2009 / Published online: 9 January 2010

Abstract We present a general framework that enables decision-making when a threshold in a process is about to be exceeded (an event). Measurements are combined with prior information to update the probability of such an event. This prior information is derived from the results of an ensemble of model realisations that span the uncertainty present in the model before any measurements are collected; only probability updates need to be calculated, which makes the procedure very fast once the basic ensemble of realisations has been set up. The procedure is demonstrated with an example where gas field production is restricted to a maximum amount of subsidence. Starting with 100 realisations spanning the prior uncertainty of the process, the measurements collected during monitoring bolster some of the realisations and expose others as irrelevant. In this procedure, more data will mean a sharper determination of the posterior probability. We show the use of two different types of limits, a maximum allowed value of subsidence and a maximum allowed value of subsidence rate for all measurement points at all times. These limits have been applied in real world cases. The framework is general and is able to deal with other types of limits in just the same way. It can also be used to optimise monitoring strategies by assessing the effect of the number, position and timing of the measurement points. Furthermore, in such a synthetic study, the prior realisations do not need to be updated; spanning the range of uncertainty with appropriate prior models is sufficient.

Keywords Monitoring strategy · Decision support · Event probability · Bayes · Ensemble · Subsidence

M. Nepveu · I.C. Kroon · P.A. Fokker (✉)
TNO Built Environment and Geosciences, Princetonlaan 6, P.B. 80015, 3508 TA Utrecht, The Netherlands
e-mail: peter.fokker@tno.nl

1 Introduction

Monitoring is important in situations where certain events must be prevented from happening, or require mitigating measures to be deployed soon or further in the future, or simply need to be reported. Monitoring enables enough information to be obtained so that a warning signal can be given whenever the probability that such an event will occur has reached a predefined critical level. The warning, or red flag, signals the start of an action or a chain of actions. This paper deals with the translation of monitoring data into the probabilities of events occurring at a specified moment in the future, so that a metaphorical red flag can be hoisted when necessary. Our current approach originated from assessing the monitoring strategy for subsidence resulting from gas production in sensitive areas. We will therefore refer to the context of subsidence, although the method proposed is applicable to a wide variety of problems within or outside the field of geosciences. While problems of this nature have been addressed before in different fields of study such as earthquakes, flood forecasting, banking and medicine (for instance, Convertito et al. 2008; Nguyen et al. 2008; Orre et al. 2008; Reggiani and Weerts 2008, Youngblood and Atwood 2005), the methodologies used vary and seem to lack a general basis. This hampers the comparison of results. The method we describe here is conceptually straightforward and general. It entails dealing with the problem in terms of Bayesian probability theory (Jaynes 2003). Another method that might be employed is particle filtering (Arulampalam et al. 2002), but our method seems simpler in its approach.

2 Methodology

For the sake of clarity, let us assume that gas is produced in an area that has particular heritage or environmental value. A licence is issued to the operating company on the condition that the ground should not subside more than a certain specified limit. This can be expressed as

$$U(t, Y) \equiv \text{Limit will be exceeded } Y \text{ years from time } t.$$

In this formulation, Y will typically stand for a period between two extensive measurement campaigns necessary during the production process, or it is a time span enabling mitigation if critical levels are expected. We are interested in the probability, derived from the monitoring data, that the limit will be exceeded, $P(U(t, Y))$. A completely analogous scheme can be set up for all kinds of probabilities; for example, if a limit is not being reached, an aquifer being active, or a fault sealing. To attribute a meaning to this probability, we must ask ourselves in which universe this probability exists. In other words, we must define the space of objects that we have to assess. Before production operations begin, we will make many quantitative models of the subsurface, using a fixed gas production prognosis so as to take account of alternative geological models of the subsurface as well as of a range of geo-mechanical parameters. Key to our method is that our prior assumptions about $P(U(t, Y))$ are updated with measurements—more accurately, they are updated with the probability of each realisation, given the measurements. Suppose our universe, the ensemble of

subsurface realisations, is $\{M_j;\ j = 1, \ldots, N\}$. Then, following standard probability theory, the posterior probability of the event is given as

$$P\big(U(t,Y)|\text{Data}\big) = \sum_i P\big(U(t,Y)|M_i\big)P(M_i|\text{Data}). \tag{1}$$

The fact that we write Data after the vertical bar in this probability signifies that the probability depends on the data hitherto obtained. It is thus a conditional probability.

The relationship between the subsurface models and the subsidence requires a process, or forward, model. We use a reservoir model in conjunction with a published geomechanical method (Fokker and Orlic 2006). By running the forward model on the ensemble of subsurface realisations, we establish an ensemble of subsidence realisations spanning the uncertainty before any measurements are collected. We know for each member whether it exceeds the subsidence limit or not. In other words, we know the definition of $P(U(tY)|M_i)$. Although there will inevitably be computational inaccuracies, we assume them to be such that we can infer from the models whether or not $U(t,Y)$ holds; hence $P(U(t,Y)|M_i) = 1$ or 0. The next question is how to attribute a probability to each realisation in the ensemble. The Bayesian probability of a particular realisation k is

$$P(M_k|\text{Data}) = \frac{P(M_k)\cdot P(\text{Data}|M_k)}{\sum_j P(M_j)\cdot P(\text{Data}|M_j)}. \tag{2}$$

In this formula, the left-hand side is the probability we are looking for (for simplicity, tags for the explicit time dependence have been omitted from the formula). In the right-hand side nominator, we have the prior probability of the realisation, which is our subjective assessment, and the associated likelihood of the data. The denominator is merely a normalising factor, the sum of all the realisation probabilities considered to be unity. Since all models are assumed to have a sound geophysical basis before being admitted to our ensemble, all prior probabilities are assumed equal to $1/N$. As will become apparent from what follows, this is not a major issue.

Next, we discuss how to choose the likelihood $P(\text{Data}|M_k)$? The data obtained at a time $t(z(t))$ must be compared with the result calculated for time t for realisation $M_k(z_k(t))$. We should have a general idea about the accuracy of the data and the M_k combined in a standard deviation. We propose a Gaussian function for this likelihood. The underlying general idea is convincingly explained in Jaynes (2003) [in Chaps. 7 and 11 in particular] and we refer to that publication. We now introduce an objective function $I_k(t)$ at time t, and, hence, state that

$$P(\text{Data}|M_k) = \exp\big[-I_k(t)\big], \tag{3}$$

$$I_k(T) = \int_0^T \frac{(z(t) - z_k(t))^2}{2\sigma_k^2(t)}dt. \tag{4}$$

When the data set is discrete rather than continuous, we must use a summation rather than an integral

$$I_k(T) = \sum_i \frac{(z(t_i) - z_k(t_i))^2}{2\sigma_k^2(t_i)}. \tag{5}$$

Note that the simple summation of the quadratics in the above equations is only acceptable when the cross correlations in the measurements are negligible. In the present study, we assume this to be the case. Using the data acquired, we can now compute the model probability $P(M_k|\text{Data})$ with (2). This probability is time-dependent, as the longer we monitor, the bigger our data set becomes. Finally, we may compute the required updated probability of "the event", using (1).

A number of remarks must be borne in mind. We include in our ensemble only those probabilities that exist in the universe of model realisations that we have adopted. If we are very restrictive in admitting realisations, we may run the risk of insufficiently taking uncertainty into account. We might miss the true realisation by a wide margin. If, on the other hand, we include many models we do not a priori believe in, we run no risk at all. If indeed these models do not stand up against the data, they will be discarded automatically in our procedure. If they are bolstered, we may have to revise our original opinion of them. Problems with many uncertainties may require prohibitively long computation times. However, by contrast with conventional data assimilation, all the model calculations in our red flag method are performed a priori, so when new data come in the models do not have to be updated and no new calculations are necessary. Our proposed red flag method is applicable to a wide variety of problems in geoscience and beyond. Its limitations are mostly in dealing with problems with too many uncertain parameters, or problems for which the model result is very non-linearly dependent on the input. An example of a very non-linear process in geoscience is the prediction of water breakthrough on the basis of pressure measurement. In such a case, a small difference between the pressures predicted by a favoured realisation and the measured pressures may correspond to a large difference in the time of water breakthrough. We feel that the method is most applicable to problems in which the physical relation between the cause and the measured result that triggers the warning is clear; there exists a continuous dependency between cause and effect; measurements can be made which facilitate the updating of the probabilities; and remedial action is possible. Some applications could possibly be found in air pollution aggravated by unfavourable meteorological conditions, such as smog formation (Manders et al. 2009); traffic bottlenecks related to economic activity; flood risks related to weather conditions (Reggiani and Weerts 2008); damage to buildings related to groundwater management, etc. Absolute quality control is possible. For instance, for each piece of data we may expect the relevant term in (5) to be of the order unity. If our estimate of σ_k is sensible—which is likely when we know the measuring procedures, tools and modelling details—that is a result we should expect. This means that the objective function $I_k(T)$ should be of the order N_{data}, if N_{data} is the quantity of data obtained so far. If $I_k(T)$ turns out to be much bigger than the quantity of data included, we know that in an absolute sense realisation k is a bad realisation. Many realisations will indeed be expected to suffer from this inadequacy. It is precisely the point of the method to discover these realisations. If all our realisations suffer from this inadequacy, we know that we have to construct better ones, or enlarge our ensemble. In principle, the set-up of a subsidence criterion is independent from the set-up of the measurements. For example, using measurements taken around the rim of a bowl one could assess the probability of a threshold in the centre of the bowl being exceeded. However, the more closely the measurements and the criterion are connected, the sharper the probability assessment.

Let us recapitulate. We construct model realisations that form our basis of judgment, and we attribute a Bayesian probability to each of them. This probability will generally change as incoming data are incorporated. For each realisation, we will be able to assess whether $U(t, Y)$ is true or not in that particular case. The probability of $U(t, Y)$ can then be computed as the summation over all realisations. When applied to our case, in which subsidence must be monitored, we have the following procedure

1. Define the subsidence criterion.
2. Make various subsurface models, e.g. using a Monte Carlo approach.
3. Calculate the ensemble of subsidence realisations, using forward process models.
4. Define for every time and for each realisation whether or not the criterion is exceeded.
5. Determine the probabilities of the realisations, using the measurements (2).
6. Calculate the probability of exceeding the criterion using (1).

3 Demonstration

We have applied our red flag method to a synthetic case also used in an earlier inversion study (Muntendam-Bos and Fokker 2009). The case is based on an existing gas field in the northern Netherlands. To be able to apply the method, we used a reservoir model and a geomechanical model to generate synthetic measurements. These we used to update the probabilities of the Monte Carlo realisations created.

3.1 Model Set-Up

The predefined, initial model grid represents a tilted reservoir, cut by three nearly vertical faults (Fig. 1). The field has a surface area of $A = 22.2 \times 10^6$ m^2 and an assumed constant thickness of 91 m, distributed over 6 layers. The gas/water interface has a transmissibility of 0.0337; the compaction coefficient is a linear function of the pressure. The basic petrophysical properties of the reservoir are shown in Table 1. The top of the sandstone reservoir is situated at a depth of 2300 m and is capped by

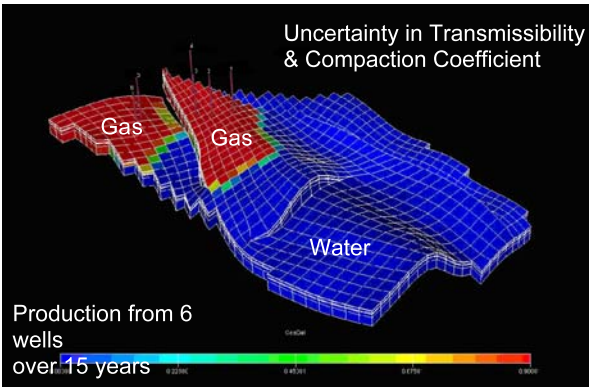

Fig. 1 Synthetic gas field and connected aquifer

Table 1 Characteristics of the synthetic reservoir

Description	Value
Depth	2300 m
Surface area	22.19×10^6 m^2
Thickness	182 m
Number of layers	6
Number of faults	3
Depletion period	15 years
Average total pressure drop	180 bar
Depletion-dependent compaction coefficient	$(0.0022 \times 10^{-5}\ \mathrm{dP} + 0.2 \times 10^{-5})$ bar^{-1}
Transmissibility gas/water interface*	0.0337
Thickness of overlying salt layer	600 m

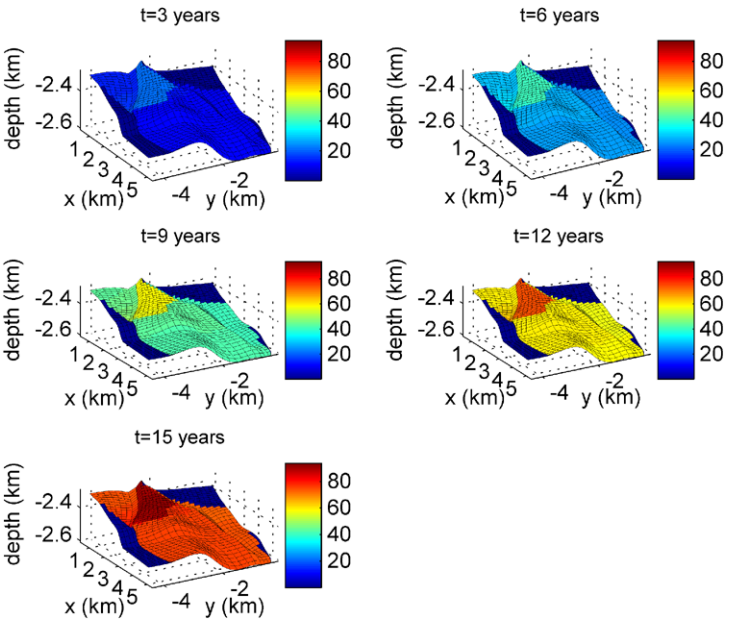

Fig. 2 Calculated pressure depletion of the gas field for the synthetic truth

a 600 m thick layer of rock salt. The reservoir is depleted by six wells over a 15-year period, resulting an average total pressure drop of 90 bar (Fig. 2). The model consists of 1036 grid cells for each layer. To be able to calculate subsidence at the surface, we need to know the total compaction of every layer of each surface grid cell. Therefore, the pressure drop for each surface grid cell at each time step was depth-integrated over the 6 layers. The elastic properties of the subsurface (both overburden and underburden) were not depth-averaged. The salt layer is represented by an intermediate layer that is stronger than the other layers. Synthetic measurements were created in 90 points every 3 years from start of production, using the forward model developed by

Fokker and Orlic (2006). These points mimic consecutive measurement campaigns. At one of the points in the centre of the subsidence bowl, additional measurements were created every year, to mimic the increased monitoring intensity resulting from the installation of a GPS point. All the measurement points had an associated uncertainty σ_k of 5 mm. This figure also includes numerical errors in the dynamic model.

3.2 Limits

We imposed two different types of limits. The first type is associated with a system that requires remediation measures to be taken if the probability that subsidence exceeding a critical value is imminent, meets, or exceeds a pre-defined probability. In that case the limit is a maximum permitted value for subsidence for all measurement points at all times, we chose typical values of 3.3 cm and 5.5 cm. Such a limit is often relevant in areas with artificially maintained water management systems. The second type of limit is associated with the subsidence rate. We applied a maximum average rate of 4 mm/year. This can be seen as a simplification of the limit applied in the Dutch Waddenzee area, where gas is produced from below a protected tidal delta sedimentary system. In order to maintain the morphological system of this protected area, the combined effect of the background subsidence rate, the sea level rise and the induced subsidence rate must be offset by the natural sedimentation rate. Therefore, the induced subsidence rate limit varies with time.

3.3 Prior Uncertainty

In our synthetic case, it was assumed that the set-up of the model, including the depth of the free water level, was known. In the model we used the same uncertainties as in Muntendam-Bos and Fokker (2009): the transmissibility of the gas/water interface in the reservoir, and the compaction coefficient. The prior probability distribution of the transmissibility multiplier resulted in large variation of the pressure in the aquifer—and hence in the compaction too. For the compaction coefficient, both the absolute value at zero pressure and the dependence on the reservoir pressure were assumed to be uncertain. The variation of the value of the compaction coefficient resulted in a general variation of the compaction; its pressure dependence resulted in a variation in the time dependence through the time dependence of the reservoir pressure. The prior uncertainties were mapped using Monte Carlo simulations: 100 realisations were created to cover these. The case used to create the synthetic measurements was not part of this ensemble.

3.4 Forward Calculation

For every simulation, the pressure history and the associated compaction were calculated and stored for the complete time sequence. Then we computed the surface subsidence in response to the reservoir compaction on the same network of 90 observation points every 3 years and on a specific GPS-point every year. These subsidence values were compared with the subsidence criterion to assign $P(U(t, Y)|M_i) = 1$ or 0 for each realisation at each time.

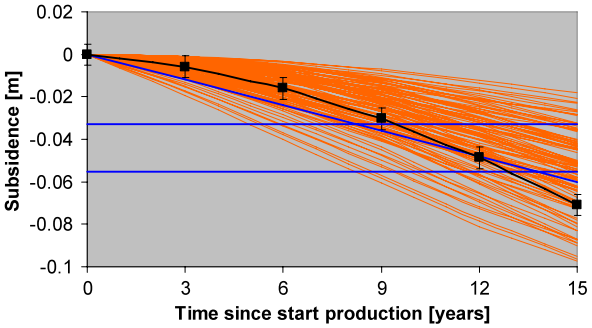

Fig. 3 Ensemble of subsidence realisations in the centre of the subsidence bowl. The *orange lines* represent the Monte Carlo results, the *black symbols/solid line* represents the measurements and their uncertainties obtained with the synthetic truth [as a linear interpolation on the 3-year periods]. The *straight blue lines* indicate the subsidence criteria: the two absolute values as *horizontal lines*, the rate criterion as a *sloping line*

3.5 Results: Updating the Probability of Each Realisation

The ensemble of the 100 Monte Carlo realisations is represented in Fig. 3 for a location near the centre of the subsidence bowl. As an example, the largest absolute-maximum criterion is exceeded for some of the scenarios after 8 years. The maximum-rate criterion is exceeded for some scenarios right from the beginning; other scenarios follow later, due to the compaction coefficient increasing with pressure. At the start of the production, measurements are not yet available and all the realisations have the same probability. After the first measurement campaign, the probabilities of the realisations change. The changing probabilities of realisations with subsequent campaigns result in a higher probability value for realisations showing behaviour close to the truth, while the others decrease in probability. This is indicated in Fig. 4(a). The distinctive power of the measurements increases with time, along with the increase of the absolute value of the subsidence. After 12 years, only a very limited subset of the original suite of scenarios has a probability that is sizable, e.g. larger than 0.01.

The probability can be determined using all the measurements available, not just the ones in one place. We tested the effect of the number of measurements taken. The developments of the probabilities of all the realisations are depicted in Fig. 4. It is clear that with a well-chosen suite of measurements, the quality of the procedure can be much improved. When all the 89 measurements are taken into account, only two of the 100 scenarios retain a reasonable probability; by contrast, in the test with only one measurement in the centre of the subsidence bowl, some 15 scenarios remain probable until the end.

The objective function $I_k(T)$ for the 100 scenarios ends up in the range of 0.67–704 with $N_{\text{data}} = 364$. At least one realisation closely matches the data, and even the worst realisation is still reasonable. This is not surprising, as our case is synthetic and well controlled. If no values close to unity were to occur in real cases, this would signal either that the models had not been adequately chosen (no models predict the

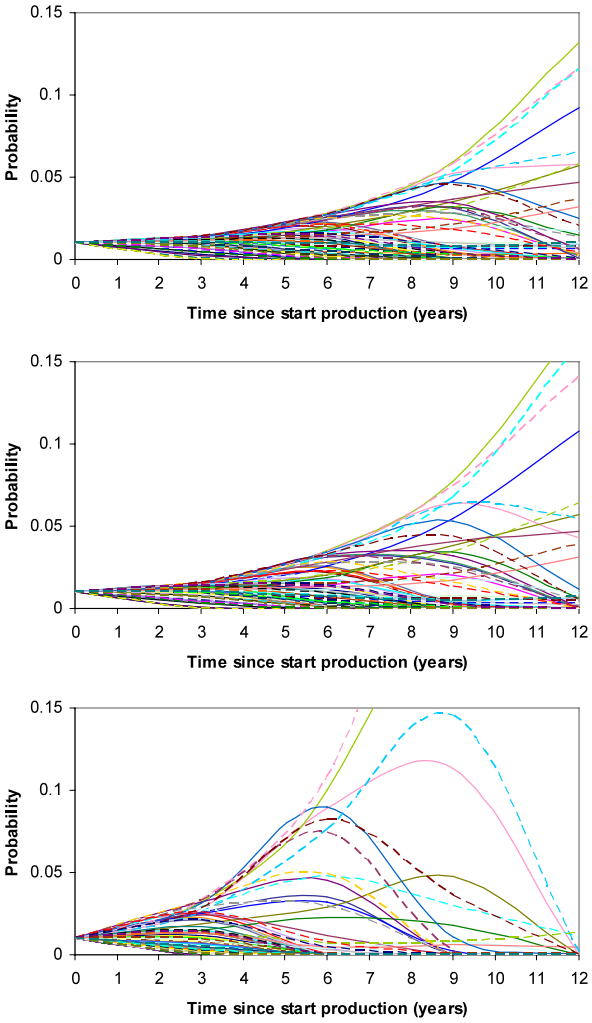

Fig. 4 Competition within the ensemble. With the acquisition of more data with time, the calculated probability of realisations close to the truth increases. When calculating the probabilities, the number of measurement points in each 3-year campaign was: 1 (*top graph*), 7 (*middle graph*) and 89 (*bottom graph*)

measurements closely enough) or that the reliability of the data had been overestimated.

3.6 Results: Updating the Probability of Exceeding the Limit Within the Next 3 Years

The final step in the procedure is the determination of the probability that the subsidence criterion will be exceeded. The results for the probabilities that the criteria will be exceeded in the next measurement campaign are given in Fig. 5. They are compared with the probabilities determined without any measurements, i.e. assigning equal probability to every single realisation during the complete period. This probability (Fig. 5, top) slowly increases with time as more and more realisations exceed the subsidence criterion.

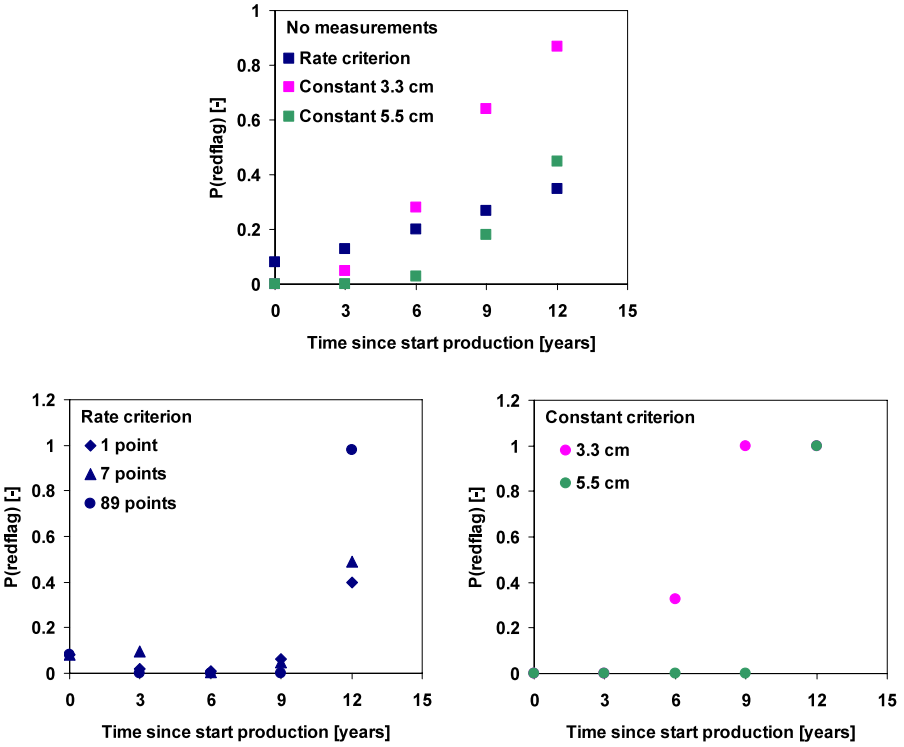

Fig. 5 Probability that a *red flag* should be hoisted, i.e. the probability that in the next campaign the subsidence criterion is exceeded taking account of the number of measurements. *Top*, without any measurements; *bottom left*, rate criterion with results for 1, 7, and 89 points; *bottom right*, absolute-value criteria for 3.3 cm and 5.5 cm, both with 89 points

When there is only one measurement point, there is already a considerable probability that the limit will be exceeded after 6 years (Fig. 5, bottom left), and already even after the first campaign. This is not the case when all 89 measurement points are included, then the determination of the probability is much sharper. The same is true after 9 years. For a single point, the probability of exceeding the limit in year 12 has only risen to about 0.4. With all the 89 points included, the determination of the probabilities is much sharper. The results for the criterion with the absolute value and 89 points included (Fig. 5, bottom right) also show that the procedure results in a realistic estimate of the probability that the criterion will be exceeded.

3.7 Optimising the Monitoring Strategy

The framework that we have developed can be used in a straightforward way to optimise the monitoring strategy. That can be done by assessing the effect of varying the number, position, type and frequency of hypothetical measurements on the probability of exceeding a limit within a certain time window. Note that no actual measurements are needed for such an exercise, which is, in fact, just a sensitivity study.

4 Concluding Remarks

We have presented a general framework for translating monitoring results into the probabilities that an event requiring action will happen, i.e. it is a framework for red-flagging. The probability of such an event is calculated by combining prior information with the likelihood of the measurements for each subsurface model. The prior probability for all realisations under scrutiny is treated equally, as only sensible realisations are allowed. As we wrote earlier, the assessments are heavily influenced by a wealth of measured data. We do not need to be harsher to some models than others at the outset; model realisations of an ill fit are quickly reduced to non-influential status. No updating of the prior models is required. This makes the procedure very fast once the basic ensemble of realisations has been set up. The procedure has been demonstrated with an example where production of a gas field was limited to a maximum amount of resulting subsidence (or rate of subsidence). Starting with 100 realisations spanning the prior uncertainty of the process, the measurements collected during monitoring bolster some of the realisations while refuting others. Inputting more data in this procedure subjects the realisations to a more severe test. Here we must emphasise the importance of understanding the Bayesian interpretation of a probability. Obtaining different probabilities by using more data does not mean that the previous probabilities were wrong. It merely means that we have improved the basis for estimating the probability of exceeding the subsidence limit, $U(t, Y)$. And, indeed, the more measurements we acquire, the more we demand from the underlying model realisations. We note, however, that we consider the measurements to be sufficiently independent to allow the simple definitions in (3) to (5) for the likelihood. These definitions may have to be amended if the data exhibit correlations.

The monitoring strategy in specific cases can be optimised by assessing the effect of the number, frequency and position of the measurement points. This enables us to find out where truly critical measurements must be gathered. Indeed, in such a synthetic study, it is not necessary to update the prior models, because it is assumed that a number of them are realistic. Before beginning the analysis we have to ponder the degree of our uncertainty. If we suffer from tunnel vision, and believe that a universe of just a few models will do, we may get poor answers. A prerequisite for the method described in this paper is the appropriate acknowledgement of what we know and what we do not. This, again, is in perfect agreement with Bayesian thinking.

Acknowledgements The authors wish to thank their colleagues Jaap Breunese and Paul Egberts for first drawing their attention to the issue and Mohamed Louh and Annemarie Muntendam-Bos for carrying out the reservoir simulations and the calculations of the reservoir compaction grids. The authors' editor was Joy Burrough.

References

Arulampalam MS, Maskell S, Gordon N, Clapp T (2002) A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans Signal Process 50(2):174–188

Convertito V, Iervolino I, Zollo A, Manfredi G (2008) Prediction of response spectra via real-time earthquake measurements. Soil Dyn Earthq Eng 28:492–505

Fokker PA, Orlic B (2006) Semi-analytic modelling of subsidence. Math Geol 38:565–589

Jaynes ET (2003) Probability theory. The logic of science. Cambridge University Press, Cambridge
Manders AMM, Schaap M, Hoogerbrugge R (2009) Testing the capability of the chemistry transport model LOTOS-EUROS to forecast PM10 levels in the Netherlands. Atmos Environ 43(26):4050–4059. doi:10.1016/j.atmosenv.2009.05.006
Muntendam-Bos AG, Fokker PA (2009) Unraveling reservoir compaction parameters through the inversion of surface subsidence observations. Comput Geosci 13(1):43–55
Nguyen MN, Shi D, Quek C (2008) A nature inspired Ying-Yang approach for intelligent decision support in bank solvency analysis. Expert Syst Appl 34:2576–2587
Orre R, Lansner A, Bate A, Lindquist M (2008) Bayesian neural networks with confidence estimations applied to data mining. Comput Stat Data Anal 34:473–493
Reggiani P, Weerts AH (2008) A Bayesian approach to decision-making under uncertainty: an application to real-time forecasting in the river Rhine. J Hydrol 356:56–69
Youngblood RW, Atwood CL (2005) Mixture priors for Bayesian performance monitoring I: fixed-constituent model. Reliab Eng Syst Saf 89:151–163