A comparison of models for measurable deterioration:
an application to coatings on steel structures

Robin P. Nicolai[a,*], Rommert Dekker[a] and Jan M. van Noortwijk[b]

ECONOMETRIC INSTITUTE REPORT 2006-28

August 15th, 2006

[a] Econometric Institute and Tinbergen Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands

[b] HKV Consultants, P.O. Box 2120, NL-8203 AC Lelystad, The Netherlands, as well as Delft University of Technology, Faculty of Electrical Engineering, Mathematics and Computer Science, P.O. Box 5031, NL-2600 GA Delft, The Netherlands

Abstract

Steel structures like bridges, tanks and pylons are exposed to outdoor weathering conditions. In order to prevent them from corrosion they are protected by organic coating systems. This paper focuses on modelling the deterioration of the organic coating layer that protects steel structures from corrosion. Only if there is sufficient knowledge of the condition of the coating on these structures, maintenance actions can be done in the most efficient way. Therefore the course of the deterioration of the coating system and its lifetime, which is also of importance for doing maintenance, have to be assessed accurately. In this paper three different stochastic processes, viz. Brownian motion with non-linear drift, the non-stationary gamma process and a two-stage hit-and-grow physical process, are fitted to two real data sets. In this way we are the first who compare the three stochastic processes empirically on criteria such as goodness-of-fit, computational convenience and ease of implementation. The first data set is based on expert judgement; the second consists of inspection results. In the first case the model parameters are obtained by a least squares approach, in the second case by the method of maximum likelihood. A meta-analysis is performed on the two-stage hit-and-grow model by means of fitting Brownian motion and gamma process to the outcomes of this model.

Keywords: Deterioration modelling, Stochastic processes, Brownian motion, Gamma process, Physical process, Maintenance, Estimation, Simulation.
JEL Classification Codes: C13, C44, C51.

---

[*] Corresponding author. Tel.: +31 10 4082524; fax: +31 10 4089162. *E-mail addresses*: `rnicolai@few.eur.nl` (Robin Nicolai), `rdekker@few.eur.nl` (Rommert Dekker) and `j.m.van.noortwijk@hkv.nl` (Jan van Noortwijk).

1 Introduction

In this paper we consider the modelling of the deterioration process of coating systems protecting steel structures, which are exposed to outdoor weathering conditions. The aim of the deterioration modelling is to predict the quality of the coating in time in order to determine when inspections or maintenance should be done and to assess lifecycle costs. In general, deterioration can be modelled in several ways:

- as a black-box statistical time to failure (such as lifetime distribution),
- as a grey-box stress-strength model based on a measurable quantity indicating time-dependent deterioration and failure (such as Brownian motion and gamma process),
- as a white-box model through simulation of the physics of measurable deterioration and failure.

The latter two cases can be split up into cases where the condition of the system is monitored continuously and cases in which (a)periodic inspections have to be performed. In this paper, we consider the latter case, because inspections have to be done to reveal the condition of the coating. Moreover, we consider the case where deterioration can be expressed through a measurable quantity for which failure/intervention limits can be set.

In the literature on deterioration modelling, both the gamma process and Brownian motion have been used to predict the value of a measurable quantity. Dekker et al. (1998) use a Brownian motion process to describe different forms of deterioration of asphalt roads. Van Noortwijk and Klatter (1999) model the scour-hole development in block mats as a gamma process. Van Noortwijk (1996) also uses the (generalised) gamma process to describe sand erosion, crest-level decline and long shore rock transport. Although Dekker et al. (1998) discuss the (sometimes undesirable) property of negative increments of Brownian motion, they do not compare this process with a gamma process. On the other hand, in Van Noortwijk and Klatter (1999) and Van Noortwijk (1996) the possible disadvantages of the gamma process are not mentioned, and this process is not compared with Brownian motion.

Another way of modelling degradation is given in Meeker and Escobar (1998). They consider the degradation of different units, measured at different points in time (resulting in a degradation path for each unit). They propose a regression model, where the regression function depends on (transformed) time. The regression parameters may be random and each unit may have different parameters. The error term is assumed to be normally distributed. However, it should be noted that a regression model cannot capture temporal variability associated with evolution of degradation. As a consequence, the deterioration along a specific sample path is

deterministic in the regression model, whereas it varies probabilistically in the Brownian motion and gamma-process model (Pandey et al., 2006).

There have been written a number of papers on (modelling) degradation of systems exposed to outdoor weathering conditions. For example, Chan and Meeker (2001) relate degradation to environmental factors, such as the weather. These factors are transformed into a degradation rate. A time series modelling approach is proposed to predict daily degradation. Heutink et al. (2004) describe how the maintenance methodology used in the Netherlands is applied to protective paint systems. The lifetime-extending maintenance model, in which deterioration is modelled by a gamma process with expected deterioration non-linear in time, is applied successfully to optimize maintenance of the coating of the Haringvliet storm-surge barrier.

In this paper we do not only model the deterioration of the coating of two specific steel structures by a gamma process and a Brownian motion process; we also simulate a physical process referred to as the two-stage hit-and-grow process. The former two stochastic processes are both in the class of Lévy processes. Restricting this class to processes with continous sample paths leads to the Wiener procsess or Brownian motion. On the other hand, forcing monotonicity leads to jump processes (Van Noortwijk, 1996); the gamma process is a jump process. The simulation of a physical process is likely to give more insight and may perform well since it is more related to the real physical process. Because environmental degradation processes are slow, data is scarce. This paper is an extension of the methodology in an earlier study (Nicolai et al., 2004). In contrast to the earlier study, in this article all processes are fitted to data in two forms: data based on expert judgement and inspection results. By doing so, we compare these processes empirically and determine the advantages and disadvantages for each model. Moreover, we consider the appropriateness of the models as well as the applicability for both types of data, viz. expert judgement and inspection data. The estimation of the parameters of the three processes is discussed in detail. Finally, we do not only show how to model deterioration of coating systems, but we also give detailed results of this modelling. Note that although this paper has a clear focus on the degradation of coating systems, the described methods should be regarded as generally applicable.

The outline of this paper is as follows. In section 2 we clarify the problem. In section 3 the two different data sets are introduced. The first is based on expert judgement; the second consists of inspection results. In section 4 we present three stochastic deterioration models. The topic of section 5 is the parameter estimation for these models. In section 6 the results of the estimation procedures are given. Finally, conclusions are drawn in section 7.

2 Problem description

We study the problem of modelling the deterioration process of the coating on two specific groups of steel objects in the Netherlands. (More information on these objects is given in sections 3.1 and 3.2). In order to get a better understanding of the degradation process, we shall first discuss how steel objects are protected against deterioration in general. The steel base of an object is protected against steel corrosion using a double (duplex) protection coating. The steel is galvanised, i.e. it is covered with a zinc coating that acts as a sacrificial anode. The zinc corrodes first, protecting the steel using the method of cathodic protection[†]. However, in order to prevent a too rapid corrosion of the zinc layer, which would leave the steel unprotected after several years, an organic coating protects the zinc. An intact organic coating shields the zinc from the corrosive environment and prevents it from corroding. Unfortunately, the organic coating is subject to degradation itself. Under the influence of humidity and UV radiation, the protective properties diminish slowly in time, eventually resulting in corrosion of the zinc layer. Too much zinc corrosion is not allowed, because it releases zinc to the environment, causing pollution. Ultimately, the steel corrodes and the structural integrity of the object can be compromised.

Modelling the degradation of coating systems is of importance when taking inspection and maintenance decisions as well as when determining life cycle costs. Not only the amount of deterioration, but also the location of the deterioration (read: corrosion) should be taken into account when planning maintenance. A widespread deterioration is much more difficult to address than a concentrated one. Black- and grey-box processes do not take the location of 'spots' into account, but processes relating more to the physics of the degradation do. For a decision support system for maintenance optimization a visualisation of the degradation process is valuable. On the other hand, black- and grey-box processes may be more convenient to express life time statistics of the coating and to allow all kind of optimizations. In this paper, we therefore also answer the question whether black- and grey-box processes can replace physical processes in a decision support system.

A number of factors have an effect on the degradation process of a coating system. We distinguish between the environment, the type of coating system, the quality of application (of the coating) and the complexity of the object. The environment (covering local weather conditions and the atmospheric composition) and the type of a coating system are judged to have the biggest effect on degradation. According to ISO standard 12944-2 (ISO/EN 12944, 1998) the environments of the two objects we consider are classified as "very high marine",

---

[†] The sacrificial anode is electrically more active than the steel and therefore, the corrosive current will exit from the zinc rather than the steel. The steel is protected while the zinc is sacrificed.

since they are situated nearby sea. This environment is the most severe with respect to the effect on degradation of coating systems. This is the reason why these objects are both protected by thick coating systems. With respect to quality of application, experts rate the steel objects discussed in section 3.1 as 'good' on a scale "bad – mediocre - good - very good" and the steel objects discussed in section 3.2 as 'mediocre / good'. The complexity of the objects is classified as 'complex' on a scale "very simple - simple – complex - very complex". So, the deterioration factors are quite similar for both objects, except for the quality of application.

Visual inspections are used to determine the amount of corrosion deterioration. The amount of deterioration can be expressed as the percentage of corroded surface or, alternatively, as the percentage of surface to be repaired. Corrosion usually starts in small spots which grow in time giving rise to a more-or-less random pattern of spots on a whole surface. Determining the actual percentage of corroded surface of large objects by visual inspection is not easy as one has to do a detailed counting. It is much easier to give a rough estimate of the amount of surface to be repainted. In this paper we consider both deterioration measures and using a simulation model we can link them. There also exist yardsticks to convert one measure to the other one. In both cases the amount of deterioration is expressed on a scale from 0 to 100%. A new coating takes value 0%; an old coating takes value 100%.

3 Data

In sections 3.1 and 3.2 we discuss the two different data sets. In section 3.3 we mention the implications that the different forms of data have on modelling deterioration.

3.1 Expert judgement

We asked experts in the field of painting deterioration to specify the time in years in which they think that the percentage of the coating surface to be repaired (on a specific steel object) exceeds 5%, 15% and 30%. The answers they provided are given in Table 1. Note that they did not specify point estimates but intervals. The experts did not specify the probabilities of exceeding the given deterioration levels in these intervals.

Deterioration level	Interval in years	Numbering
5%	10-11	1
15%	13-14	2
30%	15-17	3

Table 1: Data given by experts.

In order to find a relation between the deterioration level and time we first fitted a number of deterministic curves to the data in Table 1. The middle point of each time interval was chosen as a point estimate for the time that the corresponding deterioration level is exceeded. This way it is possible to apply least squares fitting. We fitted two functions to the data: a power law function of the form $\mu \cdot t^q$ and a logistic function of the form $100(1 - \exp(-\lambda t^\gamma))$. With respect to the choice of the first function we can say that many physical relationships can be modelled by a power function. The latter function is appropriate when the amount of deterioration is bounded or approaches a saturation point, since it can produce an s-shaped curve. Note that the amount of deterioration is bounded by 100% in this case. However, we are mainly interested in deterioration up to 30% (expert judgement), since maintenance should be done when this level is exceeded. Deterioration will only slow down later. So we actually may expect that the power function gives a good approximation of the expected course of deterioration.

It appears that the power function gives a better fit than the logistic function. The errors made by the logistic function are 5 times higher than the errors made by the power function. The parameter estimate of the power q, 4.16, suggests that the deterioration is non-linear in time. Therefore in the next section we assume therefore that the expected value of the stochastic processes follows a non-linear power function.

3.2 Inspection data

The inspection data are taken from Heutink et al. (2004) and concern the coating on the steel gates of the Haringvliet storm-surge barrier in the Netherlands. In 1970, the sea arm Haringvliet was closed off from the sea by the Haringvliet barrier. In the period from 1988 to 1996, the old coatings were replaced by new ones. In 2002, half of these coatings were inspected resulting in different inspection intervals. The available inspection data represent the percentage of the seaside surface of a steel gate that has been corroded due to ageing of the coating. The area to be repaired is about 6.75 times larger than the corroded area. The inspection results are given in Table 2. It is clear that we deal with inspections from different gates, since the inspection value in year 12 is lower than the value observed in year 10.

Time (years)	Corrosion (%)
6	0.27
8	0.41
10	0.84
12	0.75
14	2.10

Table 2: Inspection results for 5 steel gates.

3.3 Comparison of the data

In the previous sections we have seen that the experts provided us time intervals, while the inspections result in point estimates. Both forms of data have their disadvantages. Firstly, in expert judgement effects like anchoring play a role; the estimates are influenced by the initial beliefs of the experts. On the other hand, inspection data may contain errors due to measurement errors and they are costly to obtain. Secondly, estimating the parameters of the stochastic models from these data is done in a different way. For the expert data we will estimate the parameters of the processes by a least squares approach; for the inspection data we apply the maximum likelihood method. In section 5 we return to this subject.

4 Stochastic deterioration models

In this section we introduce three stochastic processes that are suitable for describing the deterioration of organic coatings protecting steel structures. In section 3 we have seen that the expected development of the deterioration process is non-linear in time. Therefore, we first describe deterioration by two grey-box processes with non-linear drift; a Brownian motion (BM) and a gamma process. Subsequently, we present a white-box two-stage physical process.

4.1 Brownian motion with non-linear drift

Let us define $D(v)$ as the stochastic deterioration v time units after the organic coating is renewed, and $D(0) = 0$ with probability 1. Assume that $D(v)$ follows a BM with *linear* drift on the time scale v (see e.g. Karlin and Taylor (1975)). Then we may write

$$D(v) = \mu \cdot v + \sigma \cdot W(v), \; v \geq 0 \tag{1}$$

In equation (1) $W(v)$ is a standard Brownian motion, μ is the drift parameter and σ the volatility parameter of the process. The mean and the variance of $D(v)$ are given by $\mu \cdot v$ and $\sigma^2 \cdot v$, respectively. Note that in the class of Lévy processes, Brownian motion is the only one having continuous sample paths. Brownian motion cannot be monotone, since forcing a Lévy process to be monotone leads to a jump process (Van Noortwijk, 1996).

In many papers on modelling degradation with a Brownian motion a time transformation is proposed, see e.g. Doksum and Hoyland (1992), Whitmore and Schenkelberg (1997), and more recently, Wang and Nair (2005). Doksum and Hoyland (1992) were the first to propose the time transformation in its general form. Real time t is transformed to time $v = \Lambda(t)$, where $\Lambda(t)$ is a

non-decreasing function of time t. In order to model variable levels of stress in an accelerated life time experiment, they apply a time transformation $v = \Lambda(t)$ which is linear with respect to time t. Whitmore and Schenkelberg (1997) consider an exponential time transformation $v = 1 - \exp(-\lambda t^{\gamma})$ and a power time transformation $v = t^q$ for modelling the degradation of self-regulating heating cables. In Doksum and Hoyland (1992) and Whitmore and Schenkelberg (1997) the parameters are estimated by applying the method of maximum likelihood. On the other hand, Wang and Nair (2005) introduce a non-parametric method to estimate (the parameters of) the time transformation function $\Lambda(t)$, without any assumption on its form.

In this paper we apply the power time transformation $v = t^q$. This transformation is suggested by the prevalence of power relationships in physical models generally (see e.g. Çinlar et al., 1977, Sobczyk and Spencer, 1992, Whitmore and Schenkelberg, 1995 and Newby, 1998). We expect that $q > 1$, which means that deterioration is accelerating in time. Applying this transformation to our stochastic process, results in a Brownian motion with non-linear drift on the time scale t (measured in years). A disadvantage of this process is that it has not only positive but also negative increments. This does not seem to be realistic since the deterioration level cannot decrease in practice, unless there are errors within the observations. However, there are no big negative increments as long as the drift is large (i.e. parameters μ and q are large) and the variance is small (i.e. σ is small). Moreover, an occasional improvement is not problematic since maintenance decisions usually depend on observed deterioration, which rarely improves if the time between two inspections is substantial (Dekker et al., 1998). An advantage of this process is that the increments are normally distributed, and this implies that the process itself is normally distributed at every point in time. Therefore, both calculating the distribution of D_v and simulating the process are easy.

Let us first look at some properties of Brownian motion with drift defined by equation (1) and assume that $v = t$. Denote by $T_D(z)$ the first time BM with linear drift exceeds a level $z > 0$, that is $T_D(z) = \inf\{t \geq 0 : D_t \geq z\}$. Now the cumulative distribution function (CDF) $F_{BM}(t, z)$ of this hitting time can be written as (Karlin and Taylor, 1975; Whitmore and Schenkelberg, 1997):

$$F_{BM}(t, z) = \Pr[T_D(z) \leq t] = \Phi\left(\frac{\mu \cdot t - z}{\sigma \sqrt{t}}\right) + \exp\left(\frac{2 \cdot \mu \cdot z}{\sigma^2}\right) \cdot \Phi\left(\frac{-\mu \cdot t - z}{\sigma \sqrt{t}}\right) \tag{2}$$

Here $\Phi(\cdot)$ is the CDF of a standard normally distributed random variable. Note that $T_D(z)$ follows the inverse Gaussian distribution with parameters $z\mu^{-1}$ and $z^2\sigma^{-2}$.

Let us now consider the time transformation in its general form, i.e. $v = \Lambda(t)$. If we denote by $T_\Lambda(z)$ the hitting time of the process $D(\Lambda(t))$, then Doksum and Hoyland (1995) have shown that $T_\Lambda(z)$ and $\Lambda^\leftarrow(T_D(z))$ follow the same distribution. Here, $\Lambda^\leftarrow$ is the inverse function of Λ. Put differently, the CDF of $T_\Lambda(z)$ can be written as $\Pr[T_\Lambda(z) \le t] = \Pr[T_D(z) \le \Lambda(t)]$. So, equation (2) can be used to compute the distribution of $T_D(z)$ and thus also the distribution of $T_\Lambda(z)$ (with $\Lambda(t) = t^q$). Moreover, it can be used to calculate the probability that a certain level is exceeded for the first time in a certain interval in terms of the model parameters. For example, the probability that the deterioration exceeds a value $z > 0$ for the first time between years $t-1$ and t, $p_{t,BM}^z$, is equal to

$$p_{t,BM}^z = F_{BM}(t^q, z) - F_{BM}([t-1]^q, z) \ , \ t \ge 1 \tag{3}$$

If we know this probability for different combinations of v and z, then we can estimate the parameters μ, σ and q.

4.2 Gamma process with non-linear shape function

As far as the authors know, Abdel-Hameed (1975) was the first to propose the gamma process as a proper model for deterioration occurring random in time. For a recent overview of the application of the gamma process in maintenance we refer to Van Noortwijk (2006).

Let us define $X(t)$ as the deterioration level at time t, and $X(0) = 0$ with probability 1, and let the probability density function of $X(t)$ be given by

$$f_{X(t)}(x) = GA(x\,|\,[\mu^2 \cdot t^q]/\sigma^2, \mu/\sigma^2) \tag{4}$$

Here, $GA(x|a,b) = [x^{a-1} b^a \exp(-bx)/\Gamma(a)] \cdot 1_{\{x>0\}}$ is the probability density function of the gamma distribution with parameters a and b. If we assume that $X(t)$ has independent increments, then $X(t)$ is a (non-stationary) gamma process with mean and variance non-linear in time, i.e. $\mu \cdot t^q$ and $\sigma^2 \cdot t^q$, respectively (see e.g. Van Noortwijk and Klatter, 1999).

The process $X(t)$ has the same mean and variance as the above-defined BM $D(t^q)$. However, in contrast to BM, the gamma distributed increments of $X(t)$ are always positive. Actually, the gamma process is a jump process. Jump processes follow from Lévy processes by assuming monotonicity. As a consequence, the CDF of the hitting time can be easily derived from the probability distribution of $X(t)$ itself. We define $T_X(z)$ as the first time the process

$X(t)$ exceeds a level $z > 0$, that is $T_X(z) = \inf\{t \geq 0 : X(t) \geq z\}$. Let us denote the CDF of the first time the gamma process $X(t)$ exceeds a level $z > 0$ by $F_{GA}(v, z)$. This function can be written as follows (see Van Noortwijk and Klatter, 1999)

$$F_{GA}(t, z) = P[X(t) \geq z] = \int_{x=z}^{\infty} p_{X(t)}(x)dx = \frac{\Gamma([\mu^2 t^q]/\sigma^2, \, z\mu/\sigma^2)}{\Gamma([\mu^2 t^q]/\sigma^2)} \tag{5}$$

Here, $\Gamma(a,x)$ is the incomplete gamma function for $x \geq 0$ and $a > 0$, and $\Gamma(a)$ is the gamma function for $a > 0$. Now, we can calculate the probability that a certain level z is exceeded for the first time between years $t-1$ and t ($t \geq 1$) as $p_{t,GA}^z$:

$$p_{t,GA}^z = F_{GA}(t, z) - F_{GA}(t-1, z) \tag{6}$$

Note again that, as for BM, we can use this expression to estimate the model parameters μ, σ and q. Also, equation (5) enables us to compute the probability distribution functions of both $T_X(z)$ and $X(t)$, $t > 0$.

Time transformation of the shape function of the gamma process

We have seen earlier that Brownian motion with non-linear drift function is derived from a standard BM by transforming the time. In a similar way we can show that the non-stationary gamma process can be derived from a stationary gamma process. Below we will first discuss this relation and derive at the same time the relation between the first passage times of these processes. To this end we first define the gamma process $Y_{v,u}(t)$ with shape function $v(t)$ and scale parameter u. Mean and variance of this process are then given by $u^{-1}v(t)$ and $u^{-2}v(t)$. Let us define the first passage time of level z by $T_{v,u}(z)$. Now it is easy to see that the process $Y_{v,1}(t)$ follows the same distribution as $u \cdot Y_{v,u}(t)$. Moreover, we can prove that the first passage time of uz for $Y_{v,1}(t)$, given by $T_{v,1}(uz)$, has the same distribution function as $T_{v,u}(z)$. This means that we may restrict the problem to gamma processes with shape function $v(t)$ and a scale parameter equal to 1.

We will now relate the hitting time $T_{v,u}(z)$ of $Y_{v,u}(t)$ to the hitting time of a gamma process having stationary increments. Let $Y_{id,1}(t)$ be this stationary gamma process with shape function $v(t) = t$ and scale parameter 1 and let $T_{id,1}(z)$ be the first passage time of level z of this process. (Here, the subscript "*id*" indicates that the shape function is equal to the identity

function, which is defined as *id(x)=x*.) Notice that $Y_{id,1}(v(t))$ follows the same distribution as $Y_{v,1}(t)$. So, applying a time transformation $v(t)$ to a stationary gamma process results in the non-stationary gamma process with shape function $v(t)$. Now it is easy to see that $T_{v,u}(z)$ and $v^{\leftarrow}(T_{id,1}(uz))$ follow the same distribution. A direct consequence of this result is that

$\Pr[T_{v,u}(z) \le t] = \Pr[v^{\leftarrow}(T_{id,1}(uz)) \le t] = \Pr[T_{id,1}(uz) \le v(t)]$, which finally relates the hitting times of the stationary and the non-stationary gamma processes.

4.3 Two-stage hit-and-growth process

In this section we present a model that is based on the physics of the deterioration process of coating systems. This model is related to the contact process introduced by Harris (1974). This process models interactions of sites on a lattice, where every site can either be infected or not infected. The sites on the lattice form a so-called $\{0,1\}$-valued random field. In the contact process infected sites infect direct neighbours at a certain rate and they recover at a certain rate. Most of the work on the contact process treats the asymptotic behaviour of the process. We refer the interested reader to chapter 7 of Bremaud (1999) for an introduction to (Markov) random fields. Probably the most well-known model in the context of Markov random fields is the Ising model, introduced by Ising in (1925). This model considers the physics of phase transitions, which occur when a small change in a parameter causes a large-scale, qualitative change in the state of a system (Cipra, 1987). We will model the degradation of a coating layer as a random field, where the coating layer is represented by a grid consisting of squares. The random field is the collection of $\{0,1\}$-valued random variables which present the condition of a square. An infected square is called a spot and such a spot will infect its neighbours after some time. Some people may see similarities between our model and Conway's game of life (Gardner, 1970), but in the game of life the state of the entire system is determined only by its initial state. In our model this is not the case, because spots appear randomly, in time as well as in space, on the grid.

The model presented here is based on the assumption that deteriorated spots arise at random positions on the coating layer. After the appearance of such a spot it expands by 'infecting' adjacent parts of the coating layer. In reality, a spot has approximately the form of a circle and experts stated that its expansion is a quadratic function of the growth of deterioration in one direction. For convenience we model expansion by declaring squares adjacent to spots infected after some (random) time. This results in the desired quadratic expansion. According to the experts in the field of coating deterioration, modelling the expansion in this way is consistent with the way deterioration is measured in practice. An advantage of this hit-and-grow process is

that it is monotonically increasing, i.e. it has positive increments only. Brownian motion does not have this property, but the gamma process has. The main advantages of the model are that

- it gives insight into the physics of the deterioration process, since it is related to the actual physical process,

- the effect of a maintenance action on degradation can be evaluated and quantified.

In this paper we will not incorporate maintenance actions. Below we describe the two stages of the model. In the first stage a spot hits the surface, in the second it starts growing. For this reason we will refer to this model as two-stage hit-and-grow process. For convenience we will use the abbreviation TSHG.

Initiation of spots

We assume that spots hit the surface according to a possibly non-homogeneous Poisson process (NHPP) $N(t)$ with intensity rate function $\eta(t) = \lambda t^p$, where $\lambda > 0$ and $p > -1$. This means that the initiation process is non-stationary (unless $p = 0$). We allow for this since experts stated that the initiation/arrival of spots depends on the permeability of the coating layer, which becomes more permeable as it deteriorates. So, it is likely that the intensity rate function is an increasing function of time ($p > 0$). Note that for this NHPP the expected number of arrivals until time t, i.e. $E[N(t)]$, is equal to $m(t) = \hat{\lambda} t^{p+1}$, with $\hat{\lambda} = \lambda/(p+1)$. Let us introduce a new time scale $y = \varphi(t) = t^{p+1}$. The initiation of spots can be modelled as a HPP with rate $\hat{\lambda}$ on time scale y. This time transformation simplifies the simulation of the NHPP, since the interarrival times of a HPP are exponentially distributed with parameter $\hat{\lambda}$. In this case the arrival times of the NHPP can be retrieved by transforming the time back to time scale $t = \varphi^{\leftarrow}(y) = y^{1/(p+1)}$.

Now imagine the coating layer as a grid consisting of N^2 unit squares, having coordinates (i, j), $i, j = 1, ..., N$. At the arrival time of a spot, it is 'assigned' to a random position on the grid in such a way that each position is equally likely for each spot, that is each position is chosen with probability $1/N^2$. If, at a certain point in time a spot is assigned to a square which was already hit before, then its condition will remain 1, i.e. infected. In such a case we will <u>not</u> select another square.

Note that we can generalize the initiation phase by defining independent (non-homogeneous) Poisson processes $N_{ij}(t)$, which count the number of spots arrived in square (i, j) at time t. (These Poisson processes can be regarded as at most 1 jump processes.) The corresponding

intensity function is given by $\eta_{ij}(t)$. Then, the superposition of the processes $N_{ij}(t)$, defined by $N(t) = \sum_{i,j} N_{ij}(t)$, is again a Poisson process with intensity function equal to $\sum_{i,j} \eta_{ij}(t)$. So, in general the probability that a spot, once it has arrived, is assigned to square (k,l) is given by $\dfrac{\eta_{kl}(t)}{\sum_{i,j} \eta_{ij}(t)}$. Choosing the same intensity function for each square on the grid results in the probability $1/N^2$ mentioned above.

Propagation of spots

The expansion (propagation) of spots is modelled as follows. After a spot has appeared on the grid it infects all adjacent squares after δ time units. Note that a square has a maximum of eight adjacent squares. Infected squares also infect their neighbours δ time units after they were infected themselves. We assume that this interexpansion time (or 'interinfection' time) is constant over time. This seems to be a realistic assumption since the expansion depends on the purity of the surface, which appears to be constant in course of time. We will also experiment with a random interexpansion time $\mathbf{\Delta}$ following an exponential distribution with mean δ. In this case the interexpansion times are realisations of the random variable $\mathbf{\Delta}$. In this way, we also model the variability in the propagation of spots. We suspect that the exponentially distributed interexpansion times result in more interaction between the spots than the model with constant interexpansion time, because due to chance rapid expansions will also occur.

It may have become clear from the preceding that the analysis of the TSHG process can only be done analytically to a limited extent. Asymptotic statements can be made, but we are interested in the time-dependent behaviour. Problems arise when spots start to overlap, as in that case probability statements are difficult to make. Therefore we will rely on a simulation modelling. Every spot corresponds to a square (block) in a grid which is the smallest perceptible amount of corrosion. Notice that the fineness of the grid determines the spot size which is also related to the minimal size that can be detected by inspection. Having many squares results in a more detailed modelling, but is also much more computational demanding. For example, if we consider a sluice door of 100 m^2 and take 10,000 (1,000,000) squares in the simulation model, then each square represents an area of 100 (1) cm^2.

The fact that we consider the initiation of spots as a NHPP makes simulation easy. Simulating the propagation stage is somewhat more difficult since as more spots arise on the grid, the number of potential expansions also increases. This requires a lot of bookkeeping and,

as a consequence, working memory. The simulation was implemented in the statistical software package R (R Development Core Team, 2005).

5   Parameter estimation

In sections 5.1 and 5.2 the parameter estimation of the three processes on the expert data and the inspection data is discussed, respectively. In section 5.3 we introduce a meta-analysis of the TSHG process. With respect to this we discuss the fitting of BM and the gamma process to simulated interval probabilities.

5.1   Expert data

In this paragraph we will give a method for estimating the parameters of the three stochastic processes by expert judgement. In some articles on deterioration modelling expert judgement has been used to estimate (some of) the parameters of a stochastic process. For example, Wang et al. (2002) model the hazard rate of water pumps in the presence of PM as a gamma process. In the absence of PM the mean of the gamma process is taken as the hazard rate of the Weibull distribution. The scale parameter of this distribution is estimated by expert judgement. The method of maximum likelihood is applied to estimate the other parameters, such as the scale parameter of the gamma process itself and the parameters that model the effect of PM on the hazard rate. With the method introduced below all parameters can be estimated simultaneously.

We have mentioned that the answers of the experts in Table 1 include uncertainty. Unfortunately, the experts did not give a good estimate of the probability that deterioration level i is exceeded in interval i, for $i = 1,2,3$. Therefore, we assign this probability to each interval ourselves. Let us define α_i as the probability that deterioration level i is exceeded in the i th interval. Now we can estimate the model parameters by equating α_i (for $i = 1,2,3$) with the probabilities that result from the stochastic processes. For example, estimating the parameters of the gamma process comes down to solving the following system of non-linear equations with respect to the model parameters:

$$
\begin{aligned}
p_{11,GA}^{5} &= F_{GA}(11,5) \ - F_{GA}(10,5) \ = \alpha_1 \\
p_{14,GA}^{15} &= F_{GA}(14,15) - F_{GA}(13,15) = \alpha_2 \\
p_{16,GA}^{30} + p_{17,GA}^{30} &= F_{GA}(17,30) - F_{GA}(15,30) = \alpha_3
\end{aligned}
\tag{7}
$$

For Brownian motion the equations are similar; only the formulas for the hitting time probabilities are different. For given values of the interval probabilities α_i (for $i = 1,2,3$), the

above system can be solved for μ, σ and q. Alternatively, one can also minimise the sum of squared differences (SSD) between the left and right hand sides (LHS, RHS) of the equations with respect to μ, σ and q. We implement the latter approach and apply a standard optimization procedure to find the parameter values that minimize the SSD.

Estimating the parameters of the TSHG process is essentially the same as for the other two processes, but the difference is that this process is simulated. It appears to be impossible to derive characteristics such as the probability distribution of the TSHG process analytically. Estimating the parameters now comes down to finding parameter values for λ, p, and δ such that in $(100 \cdot \alpha_i)$% of the simulation runs deterioration level i is exceeded in the ith interval. To this end we minimise the following stochastic objective function:

$$f(\mathbf{v}) = \sum_{i=1}^{3} (G_i(\mathbf{v}) - \alpha_i)^2 \tag{8}$$

Here, $\mathbf{v}$ is the parameter vector (λ, p, δ) and $G_i(\mathbf{v})$ is the fraction of simulation runs in which the deterioration level corresponding to the ith interval is exceeded. In order to minimise (8) we apply a grid search procedure that evaluates (8) for many combinations of the parameters λ, p and δ. The stochastic objective function is evaluated by simulating the TSHG process 1000 times on a grid of 10,000 squares, for a period of 20 years. All by all, this makes the parameter estimation a time-consuming process.

5.2 Inspection data

In order to fit a Brownian motion with drift or a non-stationary gamma process to inspection data, statistical methods for the parameter estimation are required. For a single component (i.e. steel gate), a typical data set consists of n inspection times t_i, $i = 1, \dots, n$, where $0 = t_1 < t_2 < \dots < t_n$, and corresponding observations of the cumulative amounts of deterioration x_i, $i = 1, \dots, n$, where $0 = x_1 < x_2 < \dots < x_n$.

The parameters μ, σ, and q of the gamma process can be estimated by maximising the likelihood function of the independent increments of deterioration (Çinlar et al., 1977) with respect to μ, σ, and q. The likelihood function of the observed deterioration increments $y_i = x_i - x_{i-1}$, $i = 1, \dots, n$, is a product of independent gamma densities:

$$\prod_{i=1}^{n} f_{X(t_i)-X(t_{i-1})}(y_i) = \prod_{i=1}^{n} GA\left(y_i \left| \frac{\mu^2[t_i^q - t_{i-1}^q]}{\sigma^2}, \frac{\mu}{\sigma^2} \right. \right) \tag{9}$$

This likelihood function can be extended from a single component to multiple components (i.e. steel gates) by considering m independent components, $j = 1,...,m$, for which n_j inspections are performed resulting in n_j independent deterioration increments; that is

$$\prod_{j=1}^{m}\prod_{i=1}^{n_j} f_{X(t_{ij})-X(t_{i-1j})}(y_{ij}) = \prod_{j=1}^{m}\prod_{i=1}^{n_j} GA\left(y_{ij}\left|\frac{\mu^2[t_{ij}^q - t_{i-1j}^q]}{\sigma^2}, \frac{\mu}{\sigma^2}\right.\right)$$ (10)

where x_{ij} is the cumulative amount of deterioration at the ith inspection time t_{ij} for the jth component and $y_{ij} = x_{ij} - x_{i-1j}$ is the ith deterioration increment for the jth component. In a similar manner, the parameters of Brownian motion with non-linear drift can be determined by replacing the likelihood function of independent gamma distributed increments with the likelihood function of independent normally distributed increments. For the Haringvliet steel-gate data, $m = 5$ and $n_j = 1$ for $j = 1,...,m$. Most statistical packages contain optimization procedures for likelihood functions. We have implemented and maximized the likelihood functions for BM and gamma process in EViews 5.0 (Quantitative Micro Software, 2004).

The TSHG process lacks the property of independent increments. This is because the number of spots on the grid determines the growth rate. The more spots on the grid, the higher the probability that another square becomes infected. This means that the likelihood function for this process can not be written as the product of the densities of the increments. We denote the density function of the deterioration $HG(t)$ at time $t \geq 0$ according to the TSHG process by $f_{HG(t)}(x \mid \mathbf{v})$, where $\mathbf{v}$ is again the parameter vector (λ, p, δ). In general, the likelihood function of the data given the model (parameters) can be written as the product of the multivariate density functions (of the components) at the inspection times evaluated at the observed deterioration values:

$$\prod_{j=1}^{m} f_{HG(t_{1j}),...,HG(t_{n_j j})}(x_{1j},...,x_{n_j j} \mid \mathbf{v})$$ (11)

For the Haringvliet steel-gate data we have $m = 5$ and $n_j = 1$, for $j = 1,...,m$. Eq. (11) then changes to

$$\prod_{j=1}^{m=5} f_{HG(t_{1j})}(x_{1j} \mid \mathbf{v})$$ (12)

The problem however is that we do not have explicit expressions for the density functions. We propose the following solution. For a given parameter vector $\mathbf{v}$ we replace the probability density functions in equation (12) by their kernel estimates. (For an introduction to density estimation we refer the interested reader to Scott (1992)). To this end we first simulate k sample paths from the TSHG process and record the realisations of the process at each of the m inspection times $t_{11},...,t_{1m}$. The k realisations for each inspection time are denoted by $z_{j1},...,z_{jk}$ for $j=1,...,m$. Given this sample of realisations the kernel estimate (or the Parzen-Rosenblatt estimate, Parzen (1962)) of $f_{HG(t_{1j})}(x_{1j} \mid \mathbf{v})$ is defined by

$$\hat{f}_{HG(t_{1j})}(x_{1j} \mid \mathbf{v}) = \frac{1}{kh}\sum_{l=1}^{k} K(\frac{x_{1j} - z_{jl}}{h}), \quad j=1,2,...,m \tag{13}$$

Here the function K is a square integrable probability density (e.g. normal) and h is a smoothing parameter. Note that h is usually referred to as the bandwidth parameter. A problem in kernel density estimation is selecting the bandwidth. We follow the literature on the optimal bandwidth problem and apply the procedure proposed by Sheather and Jones (1991).

Maximum likelihood estimation now comes down to maximising the expected value of the *simulated* likelihood function $\hat{L}(\mathbf{v})$ with respect to $\mathbf{v}$. That is, we have the following maximization problem

$$\max_{\mathbf{v}} E[\hat{L}(\mathbf{v})] = \prod_{j=1}^{m} \hat{f}_{HG(t_{1j})}(x_{1j} \mid \mathbf{v}) \tag{14}$$

Considering $\hat{L}(\mathbf{v})$ to be the real likelihood, we can perform a gradient-free optimization procedure to find the parameter estimates numerically. Since the simulation was implemented in the statistical software package R, this package was also used to maximize the likelihood function.

5.3 Meta-analysis of the TSHG process

A meta-analysis is next performed on the results of the TSHG process. This is motivated by the fact that simulating the TSHG process is time-consuming. Therefore it may be better to implement a gamma process or BM in a decision support system for maintenance optimization, instead of or in conjunction with the TSHG process itself. Clearly, the gamma process or the BM should fit well to the outcomes of the TSHG process. Fitting Brownian motion or gamma process to the outcomes of the TSHG can be done in a similar way as described in section 5.1. The most important difference with the estimation approach given there is that we first

determine the interval probabilities by simulation (of the TSHG process). Recall that an interval probability is the probability that the deterioration process exceeds a certain level in a given interval. Another difference is that we can simulate more than 3 interval probabilities such that we have more equations than parameters. Again, the parameters are estimated by minimising the sum of the squared differences between the interval probabilities for the gamma process (or BM) and the estimated probabilities for the TSHG process.

6   Computational results

In this section we discuss the results of fitting the three processes to the expert data and the inspection data. We will also present the parameter estimates of Brownian motion and the gamma process when fitted to the outcomes of the (simulated) two-stage hit-and-grow process.

6.1   Expert data

In order to estimate the parameters of the three stochastic processes to the expert data we used different values for α_1, α_2 and α_3, i.e. the probabilities assigned to each interval.

Process	μ (%/year)	σ (year)	q (-)
BM	1.28E-3	7.40E-3	3.56
GAMMA	1.34E-3	7.37E-3	3.55

Table 3a: Parameter estimates of BM and gamma process for $\alpha_1 = \alpha_2 = 0.75$, $\alpha_3 = 0.8$.

Process	μ (%/year)	σ (year)	q (-)
BM	1.42E-3	5.38E-3	3.52
GAMMA	1.42E-3	5.51E-3	3.52

Table 3b: Parameter estimates for BM and gamma process for $\alpha_1 = \alpha_2 = \alpha_3 = 0.8$.

Process	μ (%/year)	σ (year)	q (-)
BM	9.80E-4	6.93E-3	3.67
GAMMA	1.01E-3	7.10E-3	3.66

Table 3c: Parameter estimates for BM and gamma process for $\alpha_1 = \alpha_2 = 0.8$, $\alpha_3 = 0.95$.

	μ (%/year)	σ (year)	q (-)
BM	1.32E-3	3.89E-3	3.55
GAMMA	1.32E-3	3.87E-3	3.55

Table 3d: Parameter estimates for BM and gamma process for $\alpha_1 = \alpha_2 = 0.9$, $\alpha_3 = 0.95$.

The parameter estimates and the sum of squared differences for both the gamma and the Brownian motion process are given in Table 3. The parameter estimates of these processes do not differ much; the relative difference between the corresponding parameters is highest for the drift parameter μ (2.8% in Table 3c). Note that the values of the power q are smaller than in the deterministic model. The relative difference between the volatility parameters (σ) is at most 2.3% (see Table 3c). In this case the variance of BM (gamma process), given by $\sigma^2 t^q$, grows from 0.22 (0.23) after 10 years and to 2.9 (2.9) in year 20. This indicates that there is little uncertainty in the deterioration process, especially in the first 10 years. The small differences

between the processes are explained as follows. Firstly, the large drift dominates possible negative increments of Brownian motion and secondly, the variance of the processes is very small.
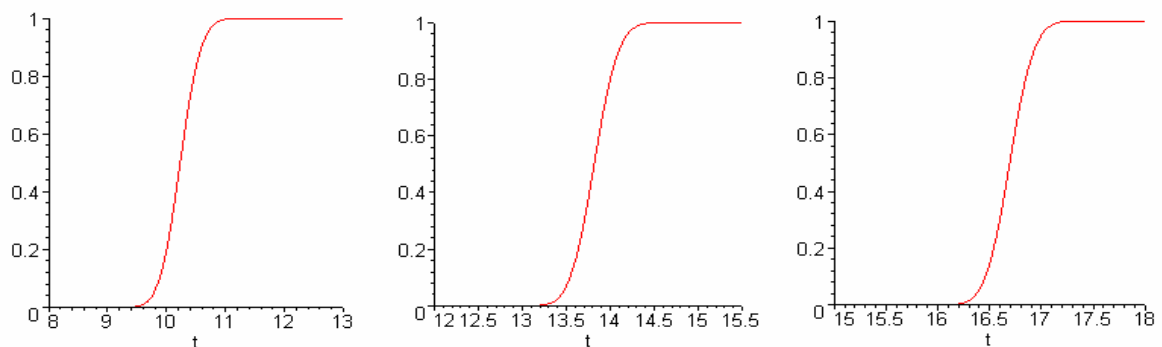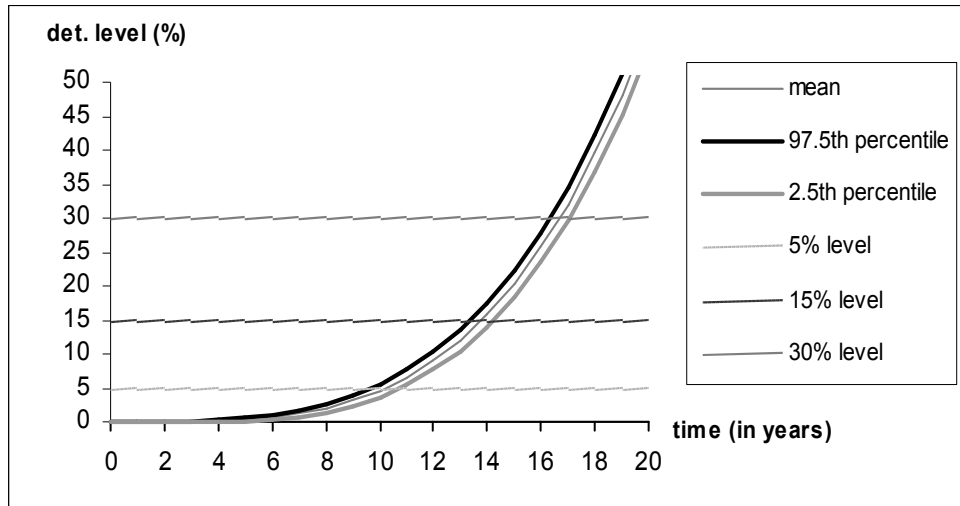

Figure 2: Cumulative distribution functions of the hitting time of Brownian motion for deterioration levels of 5%, 15% and 30%.

Figures 1 and 2 are based on the parameter estimates given in Table 3c. In Figure 1 the expected course of BM and the 2.5% and 97.5% percentiles are plotted against time. The graphs show that the deterioration is between 18% and 22% after 15 years, with probability 95%. Drawing these graphs for the gamma process gives similar results. In Figure 2, cumulative distribution functions of the hitting time of BM are plotted for the levels 5%, 15% and 30%. The graph at the left hand side shows that the probability that the deterioration level exceeds 5% in the first 10.5 years is 83%. Thus, the middle of the interval is not a good point estimate for the time that this level is exceeded. Note that the probability that the deterioration level exceeds 30% before year 16 is 0%. This shows that the experts have mentioned a too wide interval in comparison to the other two intervals; an interval [16,17] would be better, since the probability that the deterioration level exceeds 30% in this interval is 95%.

The TSHG process has been fitted to the expert data for two different sets of interval probabilities: [0.8, 0.8, 0.95] and [0.75, 0.75, 0.8]. The parameter estimates can be found in Tables 4 and 6. The computing time of the estimation procedure is much higher than it is for the gamma and BM. It takes about 10 hours to evaluate the stochastic objective function (based on 10,000 simulation runs) for 10,000 different parameter combinations. The estimates of p in the TSHG processes with non-stationary initiation phase, indicate that the initiation of spots is indeed non-stationary. In Table 5 (Table 7) it is seen that the TSHG process with non-stationary initiation phase and deterministic interexpansion time fits the interval probabilities [0.8, 0.8, 0.95] (0.75, 0.75, 0.8]) best. In this case the value of the objective function for this parameter vector, based on 50,000 simulations, is equal to 1.73E-3 (2.16E-4).

Tables 5 and 7 show that the estimated fractions $G_i(\mathbf{v})$ are not as close to the probabilities α_i as we expected them to be. The differences between $G_i(\mathbf{v})$ and α_i can be explained by the relatively small number of parameter combinations for which the stochastic objective function has been evaluated.

Figures 3, 4 and 5 are based on (simulations of) the TSHG process with the parameter estimates given in the third row of Table 4. It appears from Figure 3 that the 2.5th and 97.5th percentiles of the two-stage hit-and-grow model result in relatively wide 95% probability intervals, as compared to the other two processes. This indicates that the two-stage hit-and-grow model predicts a more volatile deterioration process. Figure 4 shows a realisation of this model. It is seen that the percentage of surface to be repaired, computed as the area of the black spots, is 27.71% after 15 years. Apart from knowing the percentage of surface to be repaired, the distribution of the spots on the surface also provides valuable information. In Figure 5 a histogram of the size of the spots on the surface shown in Figure 4 is given. It follows that there are many spots of size 1, 9 and 25 (measures in unit squares); this is because infected squares infect all their neighbours. Secondly, in this case the total area covered by spots of size less than or equal to (greater than) 25 is 718 (2053) unit squares. So, it is relatively easy to paint about 75% of the deteriorated surface in this case; the other 25% consists of small spots.

Arrival	Expansion	λ (%/year)	p (-)	δ (year)
Stationary	Deterministic	2.14	0	1.89
Non-stationary	Deterministic	0.16	2.11	2.57
Non-stationary	Stochastic	0.028	2.85	3.55

Table 4: Parameter estimates for 3 versions of the simulated TSHG process

for $\alpha_1 = \alpha_2 = 0.8$ and $\alpha_3 = 0.95$.

Arrival	Expansion	$G_1(v)$	$G_2(v)$	$G_3(v)$	$f(v)$
Stationary	Deterministic	0.114	0.312	0.474	0.94
Non-stationary	Deterministic	0.768	0.774	0.953	1.73E-3
Non-stationary	Stochastic	0.628	0.695	0.897	4.35E-2

Table 5: Goodness of fit measures for 3 versions os the simulated TSHG process

for $\alpha_1 = \alpha_2 = 0.8$ and $\alpha_3 = 0.95$.

Arrival	Expansion	λ (%/year)	p (-)	δ (year)
Stationary	Deterministic	2.35	0	2.02
Non-stationary	Deterministic	0.058	2.65	2.55
Non-stationary	Stochastic	0.025	2.90	3.53

Table 6: Parameter estimates for 3 versions of the simulated TSHG process

for $\alpha_1 = \alpha_2 = 0.75$ and $\alpha_3 = 0.8$.

Arrival	Expansion	$G_1(v)$	$G_2(v)$	$G_3(v)$	$f(v)$
Stationary	Deterministic	0.115	0.301	0.443	0.733
Non-stationary	Deterministic	0.751	0.750	0.785	2.16E-4
Non-stationary	Stochastic	0.637	0.681	0.859	2.09E-2

Table 7: Goodness of fit measures for 3 versions of the simulated TSHG process

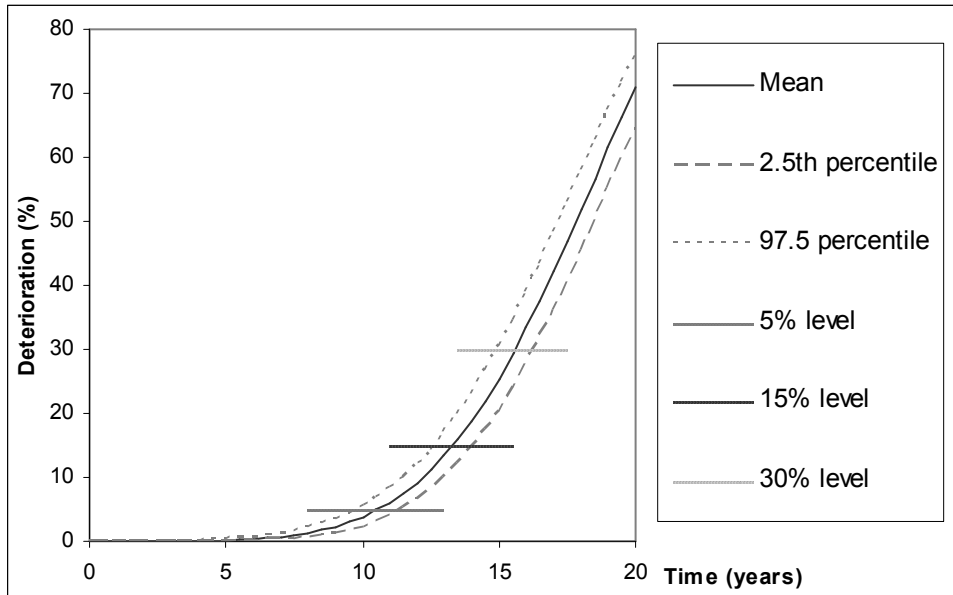for $\alpha_1 = \alpha_2 = 0.75$ and $\alpha_3 = 0.8$.

Figure 3: Expected cumulative deterioration for the two-stage hit-and-grow process with non-stationary arrival process and deterministic inter expansion times, including 2.5[th] and 97.5[th] percentiles.
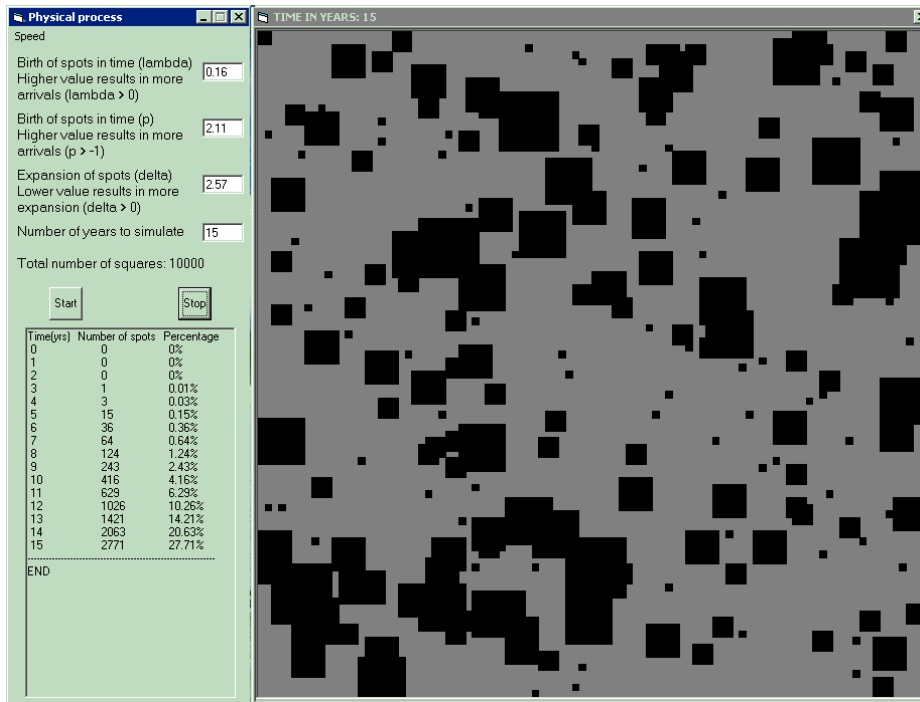
Figure 4: Realisation of the TSHG process with non-stationary arrival process and
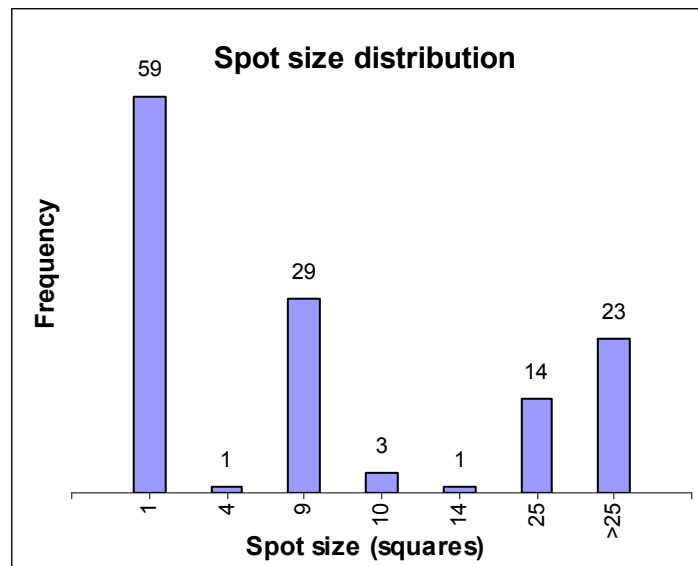deterministic interexpansion times.

Figure 5: Spot size distribution of the realisation of the TSHG process given in Figure 4.

6.2   Inspection data

Estimation results Brownian motion and gamma process

In Table 8 the maximum likelihood estimates of the parameters of Brownian motion and the gamma process are given. The difference between the values of the powers q is striking. To compensate for the high value of q, the parameter values of μ and σ are relatively small for Brownian motion. It appears that for the first 20 years the variance (given by $\sigma^2 \cdot t^q$) of the gamma process is larger than the variance of BM. Comparing the quality of the fit of the two processes is difficult in this case, since we only have 5 data points. The fact that the log likelihood value is higher for BM does not imply that BM fits the data better than the gamma process. After all, the two models are non-nested.

Parameter	Brownian motion	Gamma process
μ	1.76E-3	3.29E-3
σ	9.56E-3	1.42E-2
q	2.63	2.37
Log likelihood	1.31	0.825

Table 8: Parameter estimates for Brownian motion and the gamma process.

Figures 6 and 7 show the expected condition with its 5th and 95th percentiles including the inspection data, the probability density function of the lifetime (time to failure), and the cumulative distribution function of the lifetime according to the gamma process and Brownian motion, respectively. The deterioration failure level is a corroded surface of 3%. The lifetime is the time in years until the failure level is exceeded.

Looking at the expected condition and the 5th and 95th percentiles according to BM and the gamma process, both processes fit the data well. According to the BM the lifetime is between 15.6 and 18.4 years with probability 95%; according to the gamma process it is between 16.0 and 19.5 years with probability 95%. The mean lifetimes are given by 16.9 years for BM and 17.7 years for the gamma process.

As opposed to the gamma process, BM can result in a negative deterioration. This can be seen by comparing the 5th percentile of the deterioration for both processes in Figure 6 (top) and Figure 7 (top). For BM, the 5th percentile of the deterioration can be even negative! The larger the uncertainty in the deterioration process, the higher the probability of negative deterioration for BM, and the less suitable BM is for modelling deterioration.
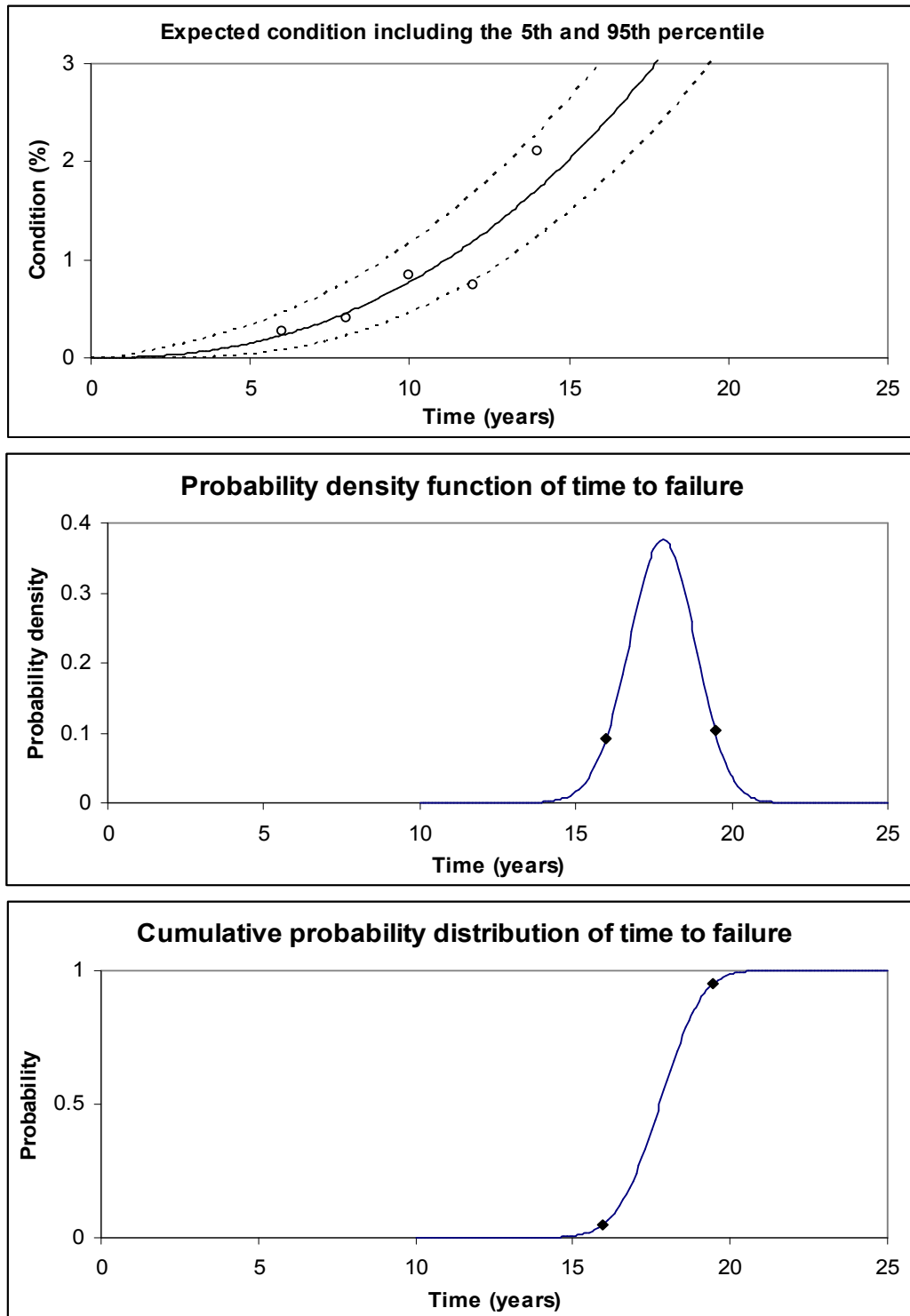
Gamma process

Figure 6: Expected cumulative deterioration, based on the maximum likelihood estimates of the gamma process, including five inspection data (○) and the 5th and 95th percentiles (---); Probability density function of the lifetime and its 5th and 95th percentiles (♦); Cumulative distribution function of the lifetime and its 5th and 95th percentile (♦).
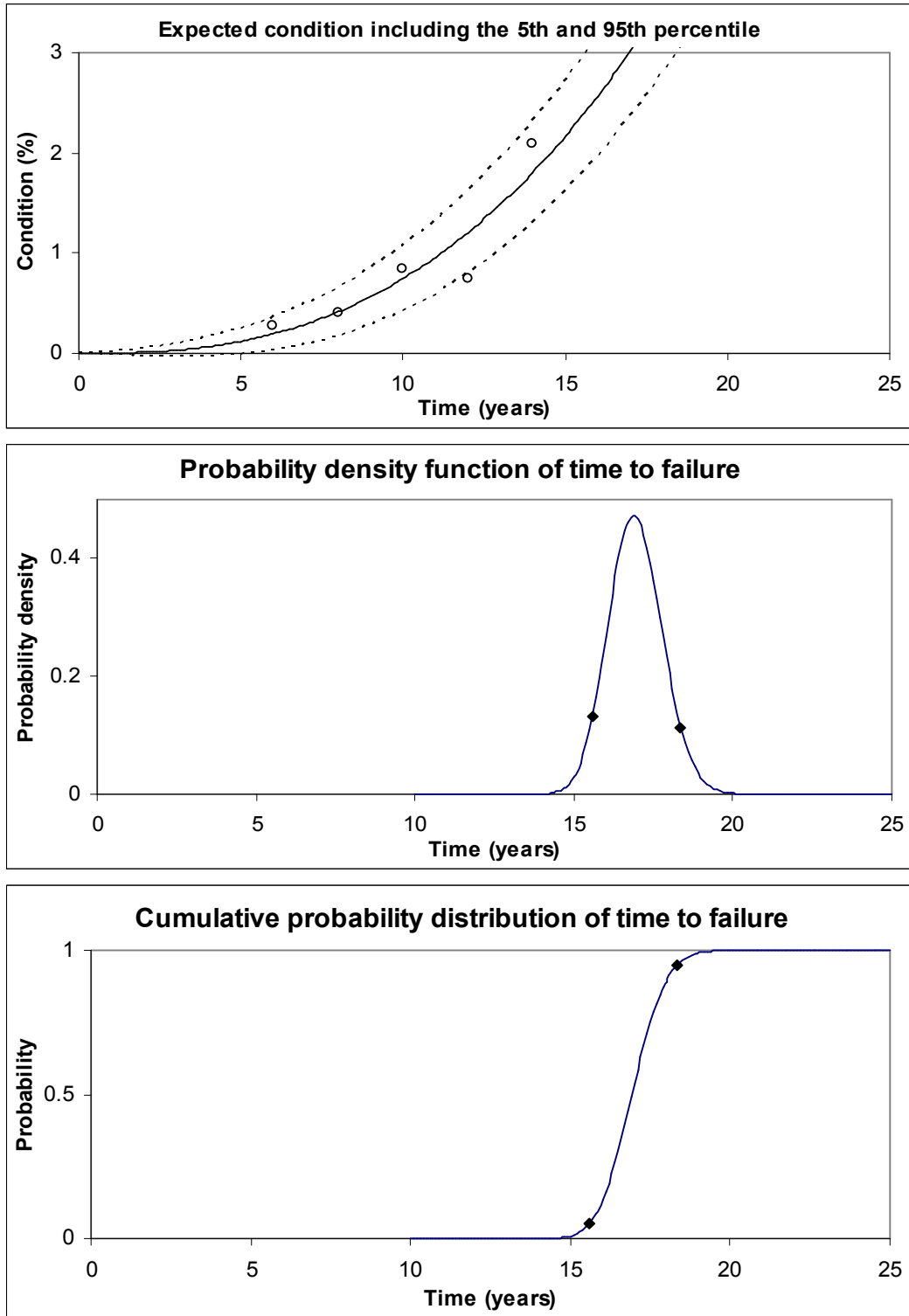
Brownian motion

Expected condition including the 5th and 95th percentile

Probability density function of time to failure

Cumulative probability distribution of time to failure

Figure 7: Expected cumulative deterioration, based on the maximum likelihood estimates of Brownian motion, including five inspection data (○) and the 5th and 95th percentiles (---); Probability density function of the lifetime and its 5th and 95th percentiles (♦); Cumulative distribution function of the lifetime and its 5th and 95th percentile (♦).

Estimation results TSHG process

Because the TSHG process measures the percentage of the area to be repaired, we have multiplied the values in Table 2 by a factor 7 in order to estimate the parameters of the process. The (log) likelihood function corresponding to the TSHG process was approximated by doing 10,000 simulations. This makes the estimation a time-consuming process. The parameter estimates are found in Table 9. The sign of p is conspicious. A negative value of this parameter implies that the intensity rate function decreases with time (the mean value function is increasing, but at a decreasing rate). This contradicts some intuition on this parameter, but it may reflect mediocre/bad application of the paint. The value of parameter δ implies that spots expand every 2.5 years.

Parameter	TSHG process
λ	2.81
δ	2.53
p	-0.42
Log likelihood	-7.96

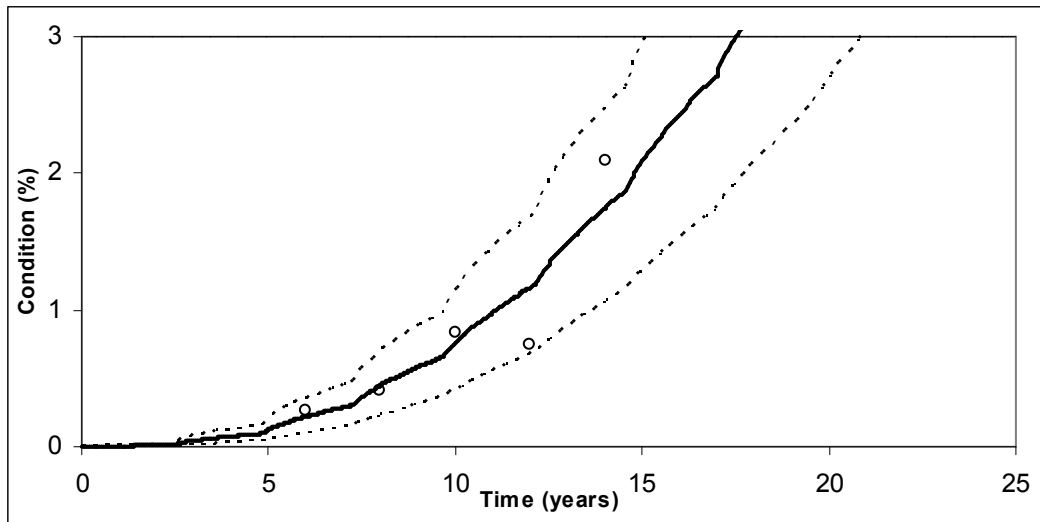Table 9: Parameter estimates for the TSHG process.

Figure 8: Expected cumulative deterioration, based on the maximum likelihood estimates of the TSHG process, including five inspection data (○) and the 5th and 95th percentiles (---).

In Figure 8 the expected cumulative deterioration including the inspection data and 5th and 95th percentiles are shown. The expected cumulative deterioration and the 5th and 95th percentiles are computed by performing 500,000 simulations. Looking at the expected condition and the 5th and

95th percentiles the TSHG process fits the data well. The variance of the TSHG process appears to be larger than the variance of BM and the gamma process.
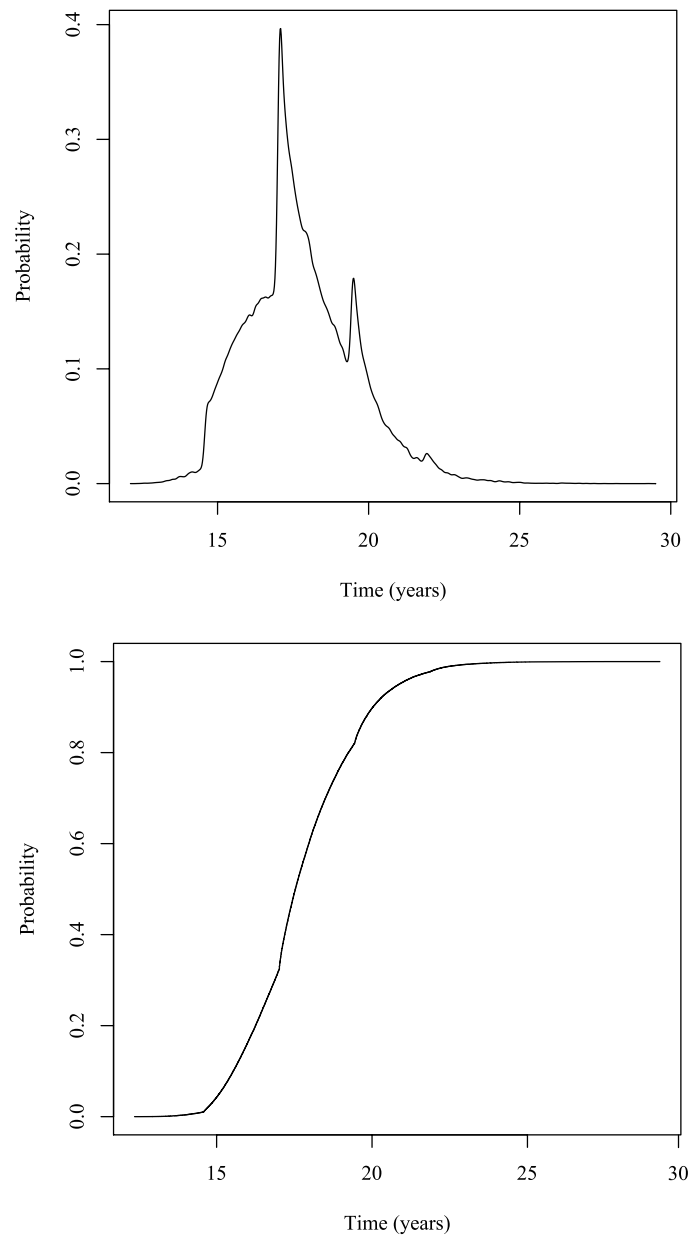

Figure 9: Empirical probability density function and empirical cumulative distribution function of the lifetime, both constructed from 100,000 simulated hitting times. Kernel density estimation is applied to approximate probability density function.

In Figure 9 the empirical probability density function of the lifetime (top) and the empirical cumulative distribution function of the lifetime (bottom) according to the TSHG process are shown. The empirical probability density function of the lifetime is constructed by first simulating 100,000 realisations of the hitting time at 3% and then applying kernel density estimation to the resulting values. The empirical cumulative distribution function of the lifetime

is constructed from the same values. The deterioration failure level is a corroded surface of 3%. The lifetime is the time in years until the failure level is exceeded. According to this model the lifetime is between 15.1 and 20.9 years with probability 95%. This is a much broader interval than we obtained for BM and the gamma process. The mean lifetime is 17.8 years, which corresponds with the result obtained by the gamma process.

6.3 Meta-analysis

In order to determine whether the TSHG process can be replaced by a BM or a gamma process, we have fitted these processes to the interval probabilities resulting from the TSHG process. By simulation we determined the probability that 7 given levels are exceeded in given intervals by the TSHG process with parameter vector $\mathbf{v} = (\lambda, p, \delta) = (0.16, 2.11, 2.57)$ [(0.058, 2.65, 2.55)], and having deterministic interexpansion times. The levels, intervals and the resulting probabilities (and their standard errors) are given in Table 10. The results are obtained by doing 250,000 [1,000,000] simulations, ensuring that the variance of the estimators is at most 10^{-6}.

Level (%)	1	2	5	10	15	20	30
Interval (years)	7-8	8-9	10-11	12-13	13-14	14-15	15-17
$\mathbf{v} = (0.16, 2.11, 2.57)$ **Prob. (st. error)**	0.6817 (9.3E-4)	0.6612 (9.5E-4)	0.7769 (8.3E-4)	0.6617 (9.5E-4)	0.7500 (8.7E-4)	0.6321 (9.6E-4)	0.9349 (4.9E-4)
$\mathbf{v} = (0.058, 2.65, 2.55)$	0.500 (4.4E-4)	0.432 (3.5E-4)	0.752 (4.5E-4)	0.744 (4.6E-4)	0.750 (4.8E-4)	0.481 (3.9E-4)	0.785 (3.1E-5)

Table 10: Estimated interval probabilities and their standard errors.

	$\mathbf{v} = (0.16, 2.11, 2.57)$		$\mathbf{v} = (0.058, 2.65, 2.55)$	
Parameter	**BM**	**Gamma**	**BM**	**Gamma**
μ	7.009E-5	9.780E-5	2.116E-5	3.021E-5
σ	3.216E-3	3.750E-3	2.203E-3	2.388E-3
q	4.750	4.625	5.209	5.080
SSD	1.375E-2	7.891E-3	1.528E-2	6.080E-3

Table 11: Parameter estimates and sum of squared differences (SSD) for BM and the gamma process.

The parameters of Brownian motion and the gamma process are estimated as explained in section 5.3. The estimates and the sum of squared differences are given in Table 11. If we compare these estimates with the parameter estimates of Brownian motion and gamma process fitted directly to the probabilities, we notice that the resulting processes are quite different. For example, the variance term is twice as small, whereas the power has increased by more than 25%.

Level (%)	1	2	5	10	15	20	30
Interval (years)	7-8	8-9	10-11	12-13	13-14	14-15	15-17
BM	0.088	0.028	0.038	0.017	0.052	0.030	0.009
	0.020	0.080	0.052	0.002	0.036	0.040	0.054
Gamma	0.068	0.021	0.032	0.011	0.039	0.017	0.004
	0.020	0.053	0.015	0.002	0.032	0.022	0.034

Table 12: Absolute differences between the interval probabilities of Brownian motion and gamma process, and the probabilities for the TSHG process.

The results in Table 12 show that the absolute differences between the interval probabilities of the hit-and-grow model on the one hand and the probabilities according to BM and the gamma process on the other hand are substantial. Thus, the grey-box processes are no perfect substitutes for the TSHG process.

7   Conclusions

In this paper we have compared three stochastic processes that can be used to model the deterioration of organic coating systems protecting steel structures. We have fitted these processes to expert data (for a specific group of steel objects in the Netherlands) and to inspection data (for the Haringvliet steel-gates). In both cases it is found that the deterioration of organic coatings is non-linear in course of time. It actually follows a power law function with power greater than 1, indicating that the deterioration is accelerating.

Brownian motion with non-linear drift and the gamma process with non-linear shape function describe the deterioration process very well. Both processes arise from the standard form of the stochastic process by applying a time-transformation. It appears that the differences between these processes are small. The reason is that the drift of the processes is high and the variance is relatively small. In spite of this, Brownian motion may have negative increments.

We have also introduced a two-stage hit-and-grow process, which is founded on the physical properties of the deterioration of coatings. The simulation of this process is time-consuming but

necessary, since it is difficult to determine its probabilistic properties analytically. We have fitted different versions of this process to the expert data. It appears that some of these versions fit the data well. For example, choosing the initiation process to be a (non-homogeneous) Poisson process with non-linear expectation gives a good fit. Also, the version with constant (deterministic) interexpansion times performs better than the version with stochastic interexpansion times. All by all, the two-stage hit-and-grow process gives more insight into the physics of the actual deterioration process than grey-box and black-box processes. This should not surprise since the process relates more to the physics of the deterioration of coatings. With respect to the inspection data it is found again that the initiation of spots increases with time, but at a decreasing rate. We have also fitted the TSHG process to the inspection data for the Haringvliet steel gates. To this end, the likelihood of the TSHG process has been estimated by applying kernel density estimation to the simulation output. It appears that the process fits the data well. Surpisingly, the initiation rate of the spots decreases with time.

A disadvantage of the TSHG process is that the probabilistic characteristics have to be determined (approximated) by means of simulation. As a consequence computing the probability distributions of the TSHG process and its hitting time is time-consuming. Therefore it seems better not to use the simulation model in a decision support system for maintenance optimization. A meta-analysis was performed in order to see whether Brownian motion or the gamma process can replace the TSHG process in such a system. We found that fitting Brownian motion and the gamma process to interval probabilities of the TSHG process may result in different results as compared to the direct fitting of these two grey-box processes.

Possible extensions of this work are the introduction of imperfect maintenance actions like painting of the coatings in the deterioration processes. The TSHG process is especially suited for this, since we can define different arrival processes for painted spots and unrepaired spots. The gamma process and BM are also to be modified in order to take into account maintenance actions. A related extension is modelling the maintenance optimization problem for steel structures protected by coating systems. Degradation modelling plays a big role in this problem.

Acknowledgements

References

Abdel-Hameed, M. (1975). A gamma wear process. *IEEE Transactions on Reliability*, 24(2):152-153.

Bremaud, P. (1999). Markov chains: Gibbs Fields, Monte Carlo Simulation, and Queues. Springer-Verlag, New York.

Chan, V., and Meeker, W.Q. (2001). Estimation of Degradation-Based Reliability in Outdoor Environments. Technical report. Iowa State University. Dept. of Statistics.

Çinlar, E., Bažant, Z.P., and Osman, E. (1977). Stochastic process for extrapolating concrete creep. *Journal of the Engineering Mechanics Division* 103(EM6), 1069–1088.

Cypra, B.A. (1987). An introduction to the Ising model. *The American Mathematical Monthly*, 94, 937-954.

Dekker, R., Plasmeijer, R.P., and Swart, J.H. (1998). Evaluation of a new maintenance concept for the preservation of highways. *IMA Journal of Mathematics Applied in Business and Industry*, 9, 109-156.

Doksum, K.A., and Hoyland, A. (1992). Models for variable-stress accelerated life testing experiments based on Wiener processes and the inverse Gaussian distribution. *Technometrics* 34, 74-82.

Gardner, M. (1970). Mathematical games: The fantastic combinations of John Conway's new solitaire game "life". *Scientific American 223*, 120-123.

Harris, T.E. (1974). Contact interaction on a lattice. *Annals of Probability* 2, 969–988.

Heutink, A., van Beek, A., van Noortwijk, J.M., Klatter, H.E. , and Barendregt, A. (2004). Environment-friendly maintenance of protective paint systems at lowest costs. In *XXVII FATIPEC Congress, 19-21 April 2004, Aix-en-Provence, France*, pages 351-364. Paris: AFTPVA.

ISO/EN 12944 (1998). Paints and varnishes – Corrosion protection of steel structures by protective paint systems. International Organization for Standardization (ISO), Geneva, Switzerland.

Ising, E. (1925). Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik*, 31, 253-258.

Karlin, S., and Taylor, H.M. (1975). *A first course in stochastic processes*. Academic Press, San Diego.

Meeker, W.Q., and Escobar L.A. (1998). *Statistical Methods for Reliability Data*. John Wiley & Sons, New York.

Newby, M. (1998). Analysing crack growth. *Journal of Aerospace Engineering (Proceedings of the Institution of Mechanical Engineers part G)*, 212, 157-166.

Nicolai, R.P., Budai, G., Dekker, R., and Vreijling, M. (2004). Modelling the deterioration of the coating on steel structures: a comparison of methods. In *Proceedings IEEE International Conference on Systems, Man and Cybernetics*, pages 4177-4182. Danvers: Institute of Electrical and Electronics Engineers (IEEE).

Pandey, M.D., Yuan, X.-X., and Van Noortwijk, J.M. (2006). The influence of temporal uncertainty of deterioration in life-cycle management of structures. *Structure and Infrastructure Engineering* (in Press).

Parzen, E. (1962). On the estimation of a probability density function and the mode. *Annals of Mathematical Statistics,* 33, 1056-1076.

Quantitative Micro Software (2004). Eviews 5 User Guide. URL http://www.eviews.com/.

R Development Core Team (2005). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-00-3, URL http://www.R-project.org/.

Scott, D.W. (1992). *Multivariate density estimation: theory, practice, and visualization*. John Wiley, New York.

Sheather, S.J., and Jones, M.C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society*, *Series B*, 53, 683-690.

Sobczyk, K., and B.F. Spencer Jr. (1992). Random fatigue: From data to theory. San Diego: Academic Press.

Van Noortwijk, J.M. (1996). Optimal maintenance decisions for hydraulic structures under isotropic deterioration. PhD Thesis, Delft University of Technology.

Van Noortwijk, J.M. (2006). A survey of the application of gamma processes in maintenance (Submitted).

Van Noortwijk, J.M., and Klatter, H.E. (1999). Optimal inspection decisions for the block mats of the Eastern-Scheldt barrier. *Reliability Engineering and System Safety* 65(3), 203–211.

Wang, X., and Nair, V. (2005). A class of degradation model based on nonhomogeneous Gaussian process. Technical report. University of Michigan.

Wang, W., and Scarf, P., and Smith, M. (2000). On the application of a model of condition-based maintenance. *Journal of the Operational Research Society*, 51(11), 1218-1227.

Whitmore, G.A., and Schenkelberg, F. (1997). Modelling accelerated degradation data using Wiener diffusion with a time scale transformation. *Lifetime Data Analysis* 3, 27-45.