# Non-Parametric Infrastructure Deterioration Curves from Differenced Condition Measurements: Method and Examples (Pre-Print)

# Non-Parametric Infrastructure Deterioration Curves from Differenced Condition Measurements: Method and Examples (Pre-Print)

Draft as per 12.04.2021.

Craig Richmond [a]*, Tariq Usman Saeed[b], Bryan T. Adey[c]

[a] Swiss Roads Office FEDRO, Department of Research and Strategy, Bern, Switzerland*;
[b] USDOT Center for Connected and Automated Transportation, Purdue University, West Lafayette, Indiana 47907-2051, USA;
[c] Institute for Infrastructure Management, Swiss Federal Institute of Technology ETH, Zurich, Switzerland

*Corresponding author: Craig Richmond; Swiss Roads Office FEDRO; 3003 Bern; Switzerland; Tel. +41 58 481 29 34; E-mail: crichmond@bluewin.ch

**ABSTRACT**

A new method is presented to develop deterioration curves covering the service life of pavements, but the method also applies of other infrastructure assets. The new method compliments existing methods by permitting the data to express the shape of the deterioration curves without a prior selection of a functional form. The new method is therefore an ideal first step in a traditional estimation process. The new method constructs a curve by linking short linear segments whose slopes are estimated independently. Connecting the short segments end-to-end yields the curves. A number of practical challenges are addressed here to adapt this approach to pavement deterioration data. Solutions are provided and the resulting statistical properties are investigated. An example of the methodology is given using a large data set from Switzerland. The examples show that climate related variable are key to the lengths of service lives in this data set.

**KEYWORDS**

Non-Parametric, Deterioration, Infrastructure, Asphalt, Climate, Filtering

## 1 INTRODUCTION

We present a new method to develop deterioration curves which does not require the a priori specification of a functional form. Thus, the data is provided a greater opportunity to express the shape of the deterioration curve than in previous methods. We call the method the *piecewise linear curves* (PLC) method and it falls into the general category of non-parametric methods. The primary focus of this article is the deterioration of asphalt concrete pavements and our examples use data from a large road network in Switzerland. However, the method itself can be applied to other cases. The method is particularly useful to estimate deterioration behaviour of infrastructure objects because those often have long service lives and therefore generally lack good data series covering the full life cycle. This presentation discusses three ways that the PLC method can be used.

The primary purpose of the PLC method is to retrieve the general shape of the deterioration behaviour from the data before any modelling takes place. The modeler can then, in a second step, select a functional form that is capable of representing the general shape in the first step. From there, the parameter estimation can proceed in the usual way. A second purpose of the PLCs is to explore hypothesis about the data directly. For example, the researcher might hypothesize that rainfall decreases service life to a significant degree. Such hypothesis can be explored directly by comparing how average service life differs when the data set is subdivided by the amount of rainfall. This can be done even if no direct observations of service lives are available. It is even possible in absence of any data on the age of any pavement sections. In addition to the usefulness of such hypothesis testing in its own right, the method can also be used when researchers need to select independent variable for a multivariate regression analysis. The pre-analysis using PLCs can identify those potential explanatory variable which have the greatest impact. Examples of this type are provided in Section 4. A third purpose of the PLC method is to generate a set of dependent variable values to test a model of the length of service life, when such values are otherwise missing. This was the actual purpose that led to the development of the PLC method (C. Richmond, Achilles, & Adey, 2018). Such a need arises when direct observations of the length of service lives are missing, but the model predicts those outcomes. The PLC method can provide a statistically based solution to that problem, if the data set is sufficiently large. An example is provided in Section 4.4. The problem type arises naturally in benchmarking and in life-cycle design problems.

Various fields of study distinguishing between parametric and non-parametric methods. Some well-known examples can be found in statistics (Kvam & Vidakovic, 2007), reliability theory (Kalaiselvan & Bhaskara Rao, 2016), and the economic analysis of production and efficiency (Ray, 2004). The later is similar to the present case in that it also constructs a piecewise linear surface and this provides a useful link to existing literature. In the analysis of production and efficiency, researchers specify a so-called production frontier, which is a hypothesized functional relationship between inputs and outputs of an unknown shape and generally of higher dimension. One solution, referred to as Data Envelopment Analysis (DEA), constructs a piecewise linear function from the data. Both PLC and DEA derive a functional form from a large data set without the use of regression, even though the regression methods are available. Hence these methods should be seen as complements to established parametric means.

There is a very large DEA literature (Fried, Lovell, & Schmidt, 2008), which is familiar to many engineers , so that a brief comparison to the PLC method may be useful to the reader. When confronted with a large set of data, the researcher has many alternative functional forms to choose from. Further, there are many alternative specific parameterisations. The methodology determines how the large set alternatives is reduced to one. In the standard regression case, the reduction occurs by first selecting the regression equation and then selecting a best-fitting parameterization, for instance that one which minimizes the sum of squares. In distinction to this, the DEA reduces the set of alternatives by applying a logical principle to the data, namely, which smallest subset of points can support a connected surface that encloses all of the data points (Charnes, Cooper, & Rhodes, 1978). The resulting *envelope* around the data is a collection of small, connected, piecewise hyper-planes in an n-dimensional space. Each of the individual hyperplane elements has parameters, of course. In that sense, the designation non-parametric is not completely accurate. But the shape of the envelope that the collection of hyperplanes creates cannot be described by some single, parameterizable, functional form. That is the sense of the non-parametric designation.

In an analogous way, the PLC method reduces the set of alternative deterioration curve shapes by applying a logic to the data. It connects end to end a sequence of line segments. What are small hyperplane elements in the DEA n-space, are short line segments in the PLC 2-space. The principle used to reduce the set of possible functions to just one function is this: use the mean slope of deterioration of all the observed road sections within some small range of condition. What results is not a multifaceted, concave, connected envelope, it is a kind of "curve" that looks like model train if it is placed on the floor but left connected and standing on its wheels. In can exhibit many potential curve-like shapes while remaining piecewise linear. The only constrain that is imposed on the curves by the methodology is that it be monotonically increasing. The condition must always get worse as time goes by.

We proceed in what follows in three steps. Section 2 addresses why the new method is needed as an addition to existing methods. Section 3 addresses necessary computations. Section 4 presents examples of results that are not otherwise available. However, before we start the main presentation we wish to point out two topics that will be covered and that are of wider interest to road deterioration researchers than are the specifics of the PLC method.

The first topic is a novel way to treat negative deterioration observations. The observation that a "road section improves all by itself", that is negative deterioration, is a standard problem for empirical deterioration function estimations. Part of these observations derive from measurement errors, but others could be picking up the effects of an unobserved maintenance intervention. In Section 3 we investigate a small set of alternative treatments to handle such data. In doing so, we reject the often-seen practice of excluding negative observations, because it leads to upwardly biased estimates. Our proposed solution avoids this bias, while still eliminating the problem of unobserved maintenance activity. The solution we propose follows the logic of so-called instrumental variables. We provide simulation evidence to support our proposal.

The second topic arises from the demonstrations of the PLC method on actual data. In many cases the data displays a shape that is initially steep, then flatter and then steeper again. We interpret this as the effect of operational maintenance slowing the average speed of deterioration, but this only happens after the condition gets bad enough to require operational maintenance at all. The effect is observable even after excluding the above mentioned issue related to unobserved maintenance interventions. In and of itself, it is not surprising that operational maintenance interventions reduce the speed of deterioration, indeed, that is one of the primary goals of doing the interventions (Gu, Ouyang, & Madanat, 2012; Lee & Madanat, 2015). The surprising observation, at least within the context of established deterioration modelling, is that the magnitude of the effect is strongly interactive with environmental condition. The length in time of the flat section appears to play a key, if not dominating role, in determining the relative length of service life. We observe this only in this data set, which is clearly limited and needs to be confirmed with other data sets. If true, it implies that standard regression models of deterioration should be constructed to allow for this effect. For example, a model could permit three life cycle phases, each with distinct parameters to permit the environmental influences to have a larger impact in the second phase. Details of this are discussed in Section 4.

## 2 WHY A NON-PARAMETRIC METHOD IS NEEDED

### *2.1 Pre-Analysis of the data, as part of the model selection process*

There is extensive literature on the estimation of parametric deterioration curves, all of which must begin with a selection of a functional form. In general, researcher are forced to make some plausible assumption, frequently selecting a functional form that has been used in the literature before. The existing literature offers a number of alternatives which can be initially divided into two main groups. The first group, the empirical models, contains the standard statistical models found not only in the deterioration literature but also in many other literatures across the sciences. These models include linear, log-linear, and the small set of what are referred to as flexible-form models (Barnett, 1983). Examples from the pavement deterioration modelling literature that provide examples of alternative specifications are (Anastasopoulos, Mannering, & Haddock, 2012; Anyala, Odoki, & Baker, 2014; Bardaka, Labi, & Haddock, 2014; Chu & Durango-Cohen, 2008; Gao, Hong, & Ren, 2019; Luo, 2013; S. M. Madanat, Karlaftis, & McCarthy, 1997; Yamany, Saeed, Volovski, & Ahmed, 2020). Markov chain models also belong to this group (Golabi, Kulkarni, & Way, 1982; Kobayashi, Kaito, & Lethanh, 2012; S. Madanat, Mishalani, & Ibrahim, 1995; Tsuda, Kaito, Aoki, & Kobayashi, 2006). The second group is the mechanistic-empirical models. These contains purpose-built mathematical functions based on pavement specific considerations. For example, one might begin with a basic fatigue model and then introduce various additional functions to determine specific parameter values. Examples showing possible specifications within this set are (Archilla, 2006; Hong & Prozzi, 2015; Meegoda & Gao, 2014; Paterson, 1987; Prozzi & Madanat, 2003; Ullidtz, 2005). Some of the major pavement design procedures belong to this group (AASHTO, 1993; Highway Research Board, 1962; Kerali, 2000; Small & Winston, 1988). The PLC method we present can be used as a first step to both sub-sets, either for the model selection or for model construction process. The data is permitted to express the general shape that the selected or constructed functional form should be able to represent.

### *2.2 Construction of observables that can be used to test models*

Infrastructures tend to be large and long-lived. Both attributes make empirical estimation of long-term deterioration behaviour difficult. If experiments are used, these not only takes a long time, but the experiments must also be quite large. An interesting confirmation of this assertion can be seen in the fact that the data from the AASHO experiment, which happened in the early 1960s (Highway Research Board, 1961), is still being used in studies 50 years later (Reger, Christofa, Guler, & Madanat, 2013). That would not be the case if such experiments were not so unique. At the same time, construction methods and standards change. It is indeed unusual to find a time series of internally consistent condition measurements that is longer than the typical service life of the full structure of a road, which may be well over 30 years (C. M. Richmond, Archilla, & Adey, 2020). Typical data sets cover only the recent past, perhaps 10 to 25 years. It is therefore difficult for researchers to find appropriate data sets to estimate or test models that predict the length of in-situ service life. The PLC method was developed to overcome this difficulty (C. Richmond et al., 2018). In that case, the available data covered at most 16 years of history, and generally less. Because the PLC method calculates average deterioration rates for small intervals of condition and then links them end-to-end, the available data was sufficient to construct estimates of average lengths of service life under various conditions, which was sufficient to test the model. We explain the calculation in detail in Section 3. This type of data problem is not unusual for long-lived assets.

### 2.3 *Accommodation of deterioration processes that have multiple phases with distinct behaviour*

Some modelers introduce distinct deterioration behaviours before and after major maintenance intervention, even if the absolute level of condition is the same (Gu et al., 2012; Lee & Madanat, 2015). For example, rates of deterioration of the surface course might be faster if the substructure is older, even if the surface course has been returned to the as-new condition. Further, some maintenance interventions, like crack-sealing or patching, may be primarily directed at reducing the speed of further deterioration. The impact of this is that the resulting rate of change of measured condition may be small or even zero.

Both examples suggest a fundamental question. To what extent is it reasonable to model deterioration without explicitly including maintenance activity as an independent variable? Or even more fundamentally, in what sense is it reasonable to use data from a section of pavement that has been subject to a maintenance intervention – say crack-sealing – together with data from pavement that has not received the same treatment? Note that this question also applies to a single section of pavement over time: after the intervention, the pavement is no longer the same physical object. It divides the observations into pre- and post-maintenance subsets. To model this, one would need to interact a maintenance dummy variable with each of the model's explanatory variable. Consider for a moment what is being predicted by a deterioration equation in a PMS application, if this equation was estimated with in-situ data. Does it describe the expected rate of deterioration with or without maintenance? Does it presuppose a given "standard" level of maintenance? Supposing one decided for the unmaintained version, which is conceptually consistent with questions as to optimal intervention strategies; one needs to know what will happen without maintenance to justify the expenditure on the activity. Where can the required data be found? Presumably, none exists, because no reasonable pavement manager would allow the pavement to deteriorate in that way. We show that maintenance has a large effect on the rate of deterioration even after the jump-improvement is excluded. Hence, deterioration equations that ignore this issue will strongly underestimate the value of maintenance.

If one proceeds from the assumption that the pre-maintenance period is not the same physical process as the post-maintenance period, then models should be constructed to accommodate the difference. In the pre-maintenance period, the behaviour of pavements is natural and undisturbed by interventions. However, once maintenance interventions begin, one is measuring something else. To model this problem parametrically, one must not only make an assumption about the functional form of undisturbed natural deterioration, one must also make an assumptions about how the functional form changes one maintenance begins. Do curves shift additively, become uniformly steeper as by a multiplier, or perhaps decrease exponentially in the intensity of maintenance?

The PLCs offer an interesting tool for exploring the mathematical nature of the interaction between maintenance and the environment to determine the rate of deterioration. One can divide the data by the intensity of the environmental variable of interest: freeze index, traffic levels, mean daily temperature amplitude, or some combination of factors; and then calculate curves for each case. The way the cases differ from one another suggests how the mathematical specification should be formulated. Examples of this technique are provided in Section 4.

In summary, it is far from obvious which functional form can be expected to fit a given in-situ deterioration process. Bottom-up modelling from mechanistic foundations offers a method for modelling the first phase of the life cycle, a time when deterioration is natural and undisturbed by maintenance (Walther & Wistuba, 2012), but it offers an insufficient basis to determine a functional form for the second and third phase of the life cycle. For all these reasons, it is useful to have a method that can allow the data to suggest the general shape that an in-situ, empirical deterioration curve should take.

## 3 CALCULATION OF NON-PARAMETRIC DETERIORATION CURVES

The construction of piecewise linear deterioration curves proceeds in three steps. First, the average amount of time is calculated, that is required for deterioration to cross a small subdivision of the total condition range. This is a value with the units $\Delta t/\Delta C$. Second, that value is inverted to arrive at a value with the units $\Delta C/\Delta t$. This is the slope of the required short line segments. Third, the line segments are linked end-to-end to create the piecewise linear curve across the whole condition range. We demonstrate the method using a condition scale from Switzerland. The method, however, is applicable for other condition scales like the PSI and IRI. The only constraint is that the scales be continuous and monotonic.

### 3.1 *Description of the condition scale used for examples*

The Swiss condition scale (VSS, 2003) resembles the present serviceability index, PSI, in that it is on a bounded interval. However, the direction is reversed. Zero is the "as new" end and five is the worst possible outcome. This scale type is used in five distinct indices, each of which measures a specific condition attribute. In this research, we use only two of the five. The first is called $I^1$. It is a composite index of visible surface defects. The second is the $I^2$ index which expresses the degree of rut damage. The index scales are linked to one another through normative meanings. These are good, fair, adequate, poor, bad for the ranges [0,1], (1, 2], (2, 3], (3, 4], (4, 5] respectively. These normative meanings impose a degree of consistencies across the indices, even if the phenomenon themselves are quite distinct. For example, some form of operational maintenance is unlikely to be needed if any one of the indices reach a values of around 1.5. Major repaving interventions generally occur between levels 2.5 and 3.5.

If the 0 to 5 range is divided into disjoint subdivisions, each having a length of 0.2 condition index units, then 25 distinct subdivisions result. If the data is grouped by the subdivisions, then an average rate of deterioration can be calculated for each. Since the beginning and ending condition levels of the subdivisions are known, combining the slopes and the boundaries results in a small line segment which approximates deterioration behaviour in each subdivision. The construction of the PLCs amounts to linking the segments end-to-end.

### *3.2 Basic calculation*

Let each subdivision of the condition range be indexed by the subscript $i$ and adopt the convention that $i$ is the midpoint. Thus, for the subdivision [0,0.2), $i$ = 0.1. The average rate of deterioration for the subdivision $i$ is

$$d_i = \underset{k \in \{i\}}{Average}\left[\frac{\Delta C_k}{\Delta t_k}\right] \tag{1}$$

where the subscript $k$ indexes road sections, $C$ is the condition level, and $t$ is a time in years. In later analysis, we compare curves for subsets of the total data set, for example, by the number of heavy vehicles per day. The subscript $j$ represents some subset of the data. This leads to a parallel definition

$$d_{i,j} = \underset{k \in \{i,j\}}{Average}\left[\frac{\Delta C_k}{\Delta t_k}\right] \tag{2}$$

Define $s$ as the width of a condition subdivision. The measurement units of $s$ are those of the condition scale. We use $s = 0.2$ throughout. The time required to deteriorate across the subdivision is calculated by inverting $d$.

$$\Delta t_{i,j} = \frac{1}{d_{i,j}} s \tag{3}$$

The time it takes to reach some specific condition level $\chi$ for a subdivision $j$ is

$$t_j^{\chi} = \sum_{i=0.1}^{\chi} \Delta t_{i,j} \tag{4}$$

Figure 1 depicts a non-parametric deterioration curve constructed according to (4) for $\chi \in [0.1, 4.9]$. The depicted damage index, $I^1$, is discussed in more detail later in the text. Our interest here is only to show the general result of the calculation method. Each segment is a short line, but collectively, one can recognize a nonlinear shape. If a parametric regression were to follow, then the functional form would have to be able to represent this shape.

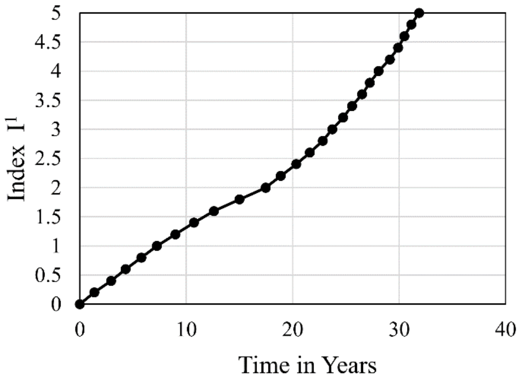


*Figure 1 Example of a non-parametric deterioration curve*

### *3.3 Data procedures*

There is an important caveat in the construction of PLCs due to the fact that $d_{i,j}$ must be inverted and also because it is empirically possible for $d_{i,j}$ to be negative. Negative slopes lead to a loop in deterioration rather than a monotonically increasing curve. Even values that are positive, but close to zero, cause implausibly slow rates of deterioration. Observations of negative rates of deterioration, i.e. that condition gets better over time in the absence of a known maintenance intervention, is common in the pavement deterioration literature (e.g. Chen & Mastin, 2015; Gao et al., 2019; O'Leary & Walsh, 2018; Perera, Byrum, & Kohn, 1998; Yang, Gunaratne, Lu, & Dietrich, 2005). A widespread practice is for researchers to exclude all data observations that are less than zero under the justification that such observations must be wrong: pavements do not get better on their own with age. We disagree with this practice. The justification is fundamentally incomplete. We demonstrate below that this exclusion leads to upwardly biased estimates. The mistake in the reasoning is to confound *the observed data* with *the deterioration* itself. Even if it were true, that the condition of a road could never improve except through interventions – an argument that contradict various kinds of self-healing (Little, Bhasin, & Darabi, 2015) – that still does not rule out the possibility that two sequential *measurements* of condition might show that the condition has improved. It is indeed a simple matter to generate observation of apparently negative deterioration in a simulation, where the true condition always gets worse, but measurement errors are symmetric and sufficiently large. Actual inspection data derives from a complex data generating process that includes many distinct and significant error processes. See (Kobayashi et al., 2012; Lethanh, Richmond, & Adey, 2016) for an extensive list. The problem for the exclude-negative-observations rule is that it treats measurement errors an asymmetric way, excluding only the negative one, but retaining the positive ones. That practice is obviously (!) biased. On the other hand, if there is good cause to believe that a given observation is subject to an asymmetric error, such as an unobserved intervention, and the exclusion process can be carried out in a non-biasing way, then an exclusion of that data is justified. These various possibilities are discussed in detail below.

### *3.3.1 Simulation of the bias caused by excluding all observations of negative deterioration.*

Figure 2 assumes a stochastic data generating process $\Delta C_{k,t} = \Delta C^{*}_{k,t} + u_{k,t}$, where $\Delta C^{*}_{k,t}$ is the true deterioration and $\Delta C_{k,t}$ is what the researcher finds in the data set. The error term $u$ combines three random distributions as might be expected for condition measurements. At location $k$ and time $t$ this is

$$u_{k,t} = \varepsilon_{k,t} - I_{k,t}\,\phi_{k,t} \tag{5}$$

The first distribution, ε, is a symmetrically distributed normal random variate representing various measurement issues. The use of a normal distribution can be justified by the central limit theorem. The second distribution, $I$, is a zero-one variate representing the occurrence or non-occurrence of an unobserved maintenance intervention. The third distribution, $\phi$, is uniformly distributed and always negative. It represents the magnitude of the condition improvement of the unobserved, minor interventions. The general problem faced when estimating deterioration is to calculate an unbiased estimate of $\Delta C^{*}_{k,t}$ given the error process $u$.

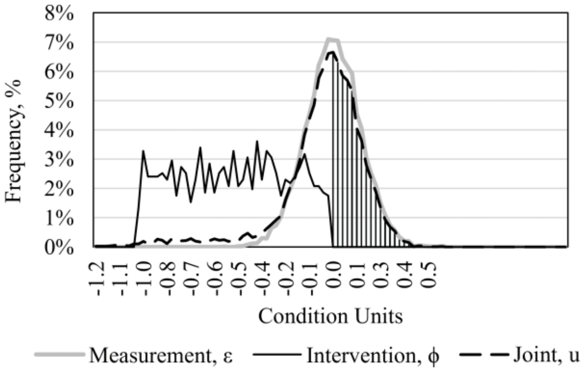


*Figure 2 Illustration of the error types and their combination*

The vertical bars to the right of zero show what is left of the original data after of all the negative values of $u$ have been removed. It is roughly half-normal and clearly biased. Please note that the variate $\phi$ appears large in the figure because it is shown before multiplication with $I$. Once multiplied, the total impact can be seen in the difference between the lines "Measurement" and "Joint".

Two heuristic procedures were developed in an earlier project to resolve this problem (Craig Richmond, Adey, & Achilles, 2019). The first heuristic is to use an alternative mechanism to filter out unobserved interventions. Following the logic of instrumental variables (see G. Judge, W. Griffiths, R.Hill, H. Lütkepohl, & Lee, 1985), one creates a filter based on an observation that is not correlated with the measurement error, but which is correlated with the unobserved interventions. The second heuristic is based on how data is assigned to ranges of $i$. These two procedures are described sequentially below. The effects of the second heuristic are explored in four simulations in Section 3.4.

### *3.3.2 A less biased filter for unobserved interventions*

Frequently, more than one condition index is available for any given section of pavement and this can be used to develop a less biased filter to extract unobserved interventions from the data. The basic idea is that two measurements of different indices will share the impact of maintenance interventions but not share their respective measurement errors. How true this is in practice depends both on the index types and the maintenance intervention type.

Our data set included both a measurement for general surface defects and a measurement for rutting in most cases. The former is the $I^1$ and the latter is the $I^3$. The $I^1$ considers cracking, ravelling, potholes, plastic deformations, and other visible surface defects. These two indexes can largely be assumed to have independent measurement errors although this is not completely true. Rutting is one of the visible defects included in $I^1$. The empirical correlation between the two indexes with respect to observed changes in condition was only 0.19 in our data. This includes any correlation that results from being affected by the same damage factors, for example, by traffic. This means that the measurement errors must have an even lower correlation than 0.19. Therefore, it is reasonable to use the $I^3$ values to filter the $I^1$ values and not to expect a substantial bias in the results. The justification is parallel to the use of instrumental variables in empirical modelling generally.

The filtering process was as follows. If the change in the value of $I^3$ at section $k$ was large and negative, then this was interpreted as evidence that a maintenance intervention had taken place. Therefore, the pair of $I^1$ observations at that location was excluded. Symmetrically, if the $I^1$ data generated a large negative value, then the pair of $I^3$ observations at that location was excluded. Concretely, a change that was more negative than -0.2 was considered large and resulted in the exclusion of the observation of the opposite index.

In addition to this filter, and in contradiction to the previous argument that own-value based filters bias the results, a second filter excluded all very large negative changes of the index itself. To assess the likely biasing impact, we refer to Figure 2 where the standard deviation of the random measurement error is 0.15. The exclusion of negative observations that are, for example, greater than two standard deviations below the mean (-0.3) cannot have a large biasing impact because, at that point, the measurement errors have a minor weight in any weighted average calculation. The same cannot be said of the unobserved interventions, which do not follow a zero mean, normal distribution. This consideration led to a second filter rule: exclude observations where the change was more negative than -0.5.

Table 1 reports summary statistics of our data after the filtering had been applied. One can see that the mean change between two observations is around +0.46 condition units considering both indices. The standard deviation of the joint error around this mean is around 0.57. These values yield potential cut-offs: -0.11 and -0.68; as reference points. These are 1 and 2 standard deviations beneath the mean, respectively. We implemented a standard cut off at -0.5 for both indices. Relative to the values in Table 1, this is 1.67 standard deviations from the mean or the 95$^{th}$ percentile of the distribution.

*Table 1 Descriptive statistics of the filtered data set having both $I^1$ and $I^3$ observations*

| Condition Measure | Damage Type | Observation | Mean Change (Index Units) | Standard Deviation (Index Units) | Average Interval (Years) |
|---|---|---|---|---|---|
| Index $I^1$ | Surface Defects | 216228 | 0.49 | 0.60 | 3.9 |
| Index $I^3$ | Rutting | 216228 | 0.43 | 0.54 | 4.6 |

### 3.3.3 *Heuristic to further exclude unobserved interventions*

The piecewise linear method proceeds by 1) calculating a slope of deterioration between sequential condition observations, 2) assigning slopes to a subdivision of the condition scale, and 3) calculating the average slope within the subdivision. To make the following concrete, imagine two observations represented as the pair $\{0.5,\ 0.9\}$ and a division of the condition scale with endpoints $\{0,\ 0.2,\ 0.4,\ 0.6,\ 0.8,\ 1.0, \ldots\}$. It was natural initially to assign the pairs to subdivisions based on the first value, which is here 0.5. That is, the pair would be assigned to the subdivision [0.4, 0.6). But there is no necessary reason for assigning a pair of values in this way. One could equally well use the ending observations, 0.9, and assign the observation to the subdivision [0.8, 1.0), or use the midpoint, 0.7 and assign to [0.6, 0.8). In other words, there is more than one way to map a sample of observations to a given set of subdivisions. The latter is probably the best choice, all else equal. The other approaches tend to shift the sample means upwards or downwards along the condition scale relative to their subdivision. The direction of the bias depends on the second derivative of the curve.

Under the assumption of linear deterioration, one would expect no impact at all. The average sample slopes should be the same regardless of the location of the sample along the condition scale. However, given either monotonically increasing or decreasing deterioration behaviour, steeper or flatter slopes will be assigned upwards or downwards along the condition scale. One cannot know which case applies until after the shape of the curve has been identified. In the absence of other considerations, an assignment using the midpoint is preferred, because it will be closest to the resulting sample mean.

In testing these options, assigning by the beginning, midpoint, or endpoint, a surprising results became apparent. The assignment rules can be used as a second filter to exclude unobserved maintenance interventions. It works because of two mutually reinforcing features of the data: errors from unobserved interventions are always negative and interventions are concentrated in the mid-condition ranges. A case was shown in Figure 1. The fact that the interventions are concentrated in the middle leads to a shifting of heavily affected mid-range observations to the left using the beginning value to assign pairs and to the right using the ending value. Because the errors from unobserved maintenance are always negative, samples that are more heavily affected will have a lower mean. Since there are two assignment rules, by the beginning value or the ending value, one can create a third rule, which we will call the Max-Value Rule, that selects the larger of the two mean slope estimates for any given subdivision j. The assignment by the beginning value is called the Begin-Point rule. The assignment by the ending value is the End-Point rule.

A simulation framework is adopted to test the impact of assignment rules on the resulting PLCs. Since the underlying deterioration behaviour and the nature of the error terms are known in a simulation, any bias can be calculated or viewed directly in figures.

### *3.4 Simulation study of the heuristics*

The results of four simulations are shown in Figures 3 to 7. The figures proceed from simple to complex. At each step, an additional feature is added. Parameter values were taken from Table 1, index $I^1$. None of the simulations included a filter rule to exclude negative deterioration values. The simulations only test the three assignment rules.

Case 1 combines a deterministic linear deterioration process with a zero mean, normally distributed random measurement errors. Case 2 adds a random variable to the deterioration process, which simulates heterogeneity in the deterioration process. The trend of the deterioration process is still linear, but the rate of deterioration across road sections is stochastic. Case 3 adds a randomly generated unobserved intervention. Finally, Case 4 replaces the linear trend in the deterioration model with a non-linear trend. Case 4 simulates the possibility that some maintenance interventions slow the rate of deterioration in addition to causing a jump shift in condition.

The goal of the simulations is to investigate whether the Max-Value rule gives reasonable estimates of the assumed deterioration process in all the cases. This turns out to be the case.

*Case 1: Linear deterioration, Symmetric measurement error*

We present two figures from this simulation. The first, Figure 3, shows the raw deterioration estimates from the three assignment rules before they are accumulated into a piecewise linear curve. Figure 4 shows the resulting curves.

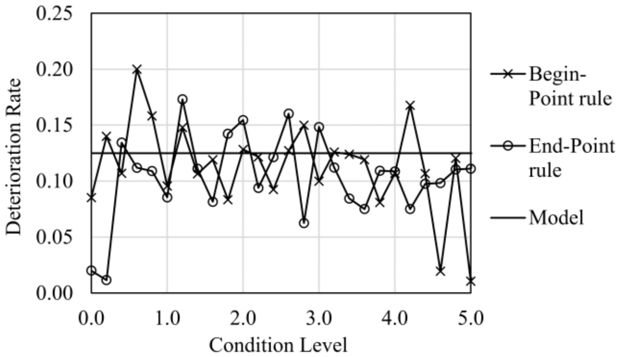


*Figure 3 Estimated deterioration rates per subdivision; Case: Only measurement error*

One can see in Figure 3 that both rules lead to subdivision deterioration estimates that appear to be randomly distributed around the true value, except for behaviour near the bounds of the interval. The boundary effect is not symmetric. At the lower boundary, the End-Point rule underestimates deterioration. At the upper boundary, the Begin-Point rule underestimates the boundary. This is caused by the truncation at the boundary of the assumed normally distributed error term. We did not investigate the added complication of bounded error terms. However, this boundary problem also affects parametric estimation in a parallel fashion if the target condition indices are bounded. We are unaware of a statistical treatment of the issue in the pavement deterioration literature. Although a treatment in a simulation context is straight forward, it would have led to many questions about the specific choice distribution, which go beyond the purpose of this research. We leave the topic to future research. By fortunate coincidence, however, the boundary effect always lowers the average deterioration rate and only affects one estimate at each boundary. The Max-Value rule, therefore, circumvents the problem by selecting the larger of the two. One can see in Figure 3 that one of the two values always lies near the assumed mean.

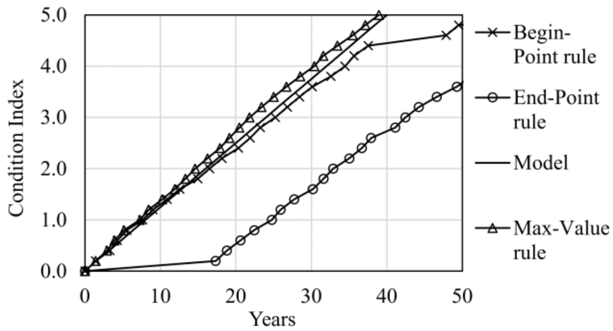


*Figure 4 Non-Parametric deterioration curves using alternative assignment rules*

Figure 4 shows the non-parametric deterioration curves. One can see the effects of the boundedness at either end. The solid line is the true, modelled process. The Max-Value rule has a slight positive bias, but the more important aspect is that it recovers the form of the deterioration process accurately. Here the form is linear. Other than the boundary problems, the other two rules also recover the deterioration model accurately. The bias problem seems reasonably small when compared to the modelled uncertainty, which has a standard deviation of 0.6 units over four years. The maximum observed variation between the Max-Value rule and the true model is 0.27 condition units. The End-Point rule is clearly not viable for the first part of the deterioration curve.

*Case 2: Linear trend, Symmetric measurement error, Stochastic deterioration*

A multiplicatively interactive random variable is introduced relative to the deterioration rate. Nothing else is changed. The random variable follows a lognormal distribution to ensure that the deterioration rate remains positive. The standard deviation is set at 20% and the mean is set to zero. If the standard deviation were reduced to near zero, this model collapses back to Case 1. The results shown in Figure 5 are similar to Case 1.

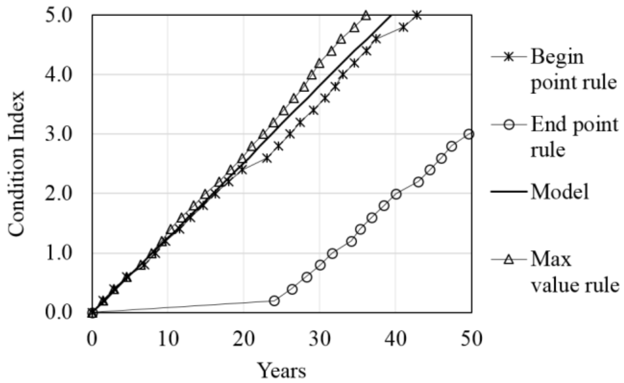


*Figure 5 Alternative assignment rules; adding uncertain rates of deterioration.*

*Case 3: Linear trend, Symmetric measurement error, Stochastic deterioration, Stochastic unobserved intervention*

An unobserved intervention has been added to the previous simulation holding all else constant. The occurrence of the intervention, $I \in \{0,1\}$, and the size of the intervention, $\phi$, have been modelled as stochastic processes as defined in equation (5). The occurrence of the intervention uses a trigger variable, $x$, centred on the condition level of 2 in order to replicate lower rates of interventions before and after the mid-condition range.

$$x_{k,t} = \alpha_1 - Abs\left[2 - C_{k,t}\right] / \alpha_2 \tag{6}$$

The parameters $\alpha_1$ and $\alpha_2$ jointly determine the spread and frequency of the interventions. The determination of the occurrence uses a uniformly distributed variate, $\phi \sim Uniform[0,1]$. The indicator random variable is defined as

$$I_{k,t} = \begin{Bmatrix} 1 \text{ if } x_{k,t} > \phi_{k,t} \\ 0 \quad \text{Otherwise} \end{Bmatrix} \tag{7}$$

This led to an average rate of occurrence over all road sections of 9.6% and an average magnitude of interventions of 0.55 condition units for the values $\alpha_1$ = 0.3 and $\alpha_2$ = 6. The values were selected by trial and error to achieve plausibly realistic values for intervention occurrence and magnitude. A sensitivity analysis that included doubling and halving the parameter levels did not change the qualitative results, which are shown in Figure 6.

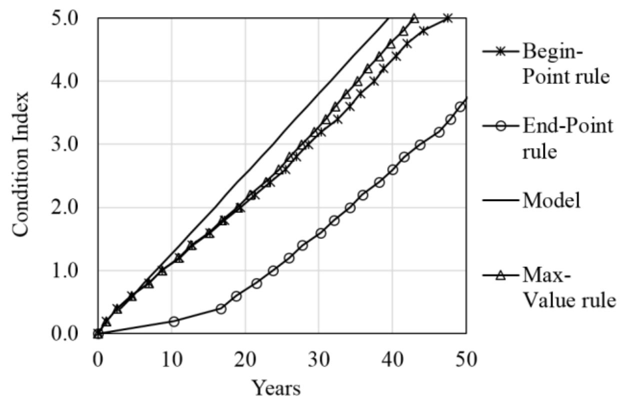


*Figure 6 Alternative assignments; adding unobserved maintenance interventions*

Adding the strictly negative unobserved interventions results in a downward bias for all three rules. Otherwise, the results are similar to the first two cases. Excluding the boundary problem, all rules recover the linear deterioration form. The Max-Value rule has the smallest negative bias.

*Case 4: Non-Linear trend, Symmetric error term, Stochastic deterioration, Stochastic unobserved intervention*

We replace the linear trend with a non-linear, s-shaped curve trend. We choose this particular form because it arises in the empirical cases shown in Section 4. The s-shape has the additional advantage of including a portion of decreasing and a portion of increasing rates of deterioration and thus covers both convex and concave forms. Deterioration itself remains stochastic. The s-shape is implemented by adjusting the mean of the stochastic distribution. As condition approaches the mid-range values, the mean is reduced towards zero. Thereafter, the mean is increased. The simulation is otherwise unchanged from Case 3. Because we are not filtering out unobserved observations, the simulations become susceptible to near-zero or negative deterioration estimations. To avoid this, it was necessary to increase the number of the simulated road sections by a factor of 12 (10000 vs 800). The larger number leads to smoother curves but no change in the qualitative results.

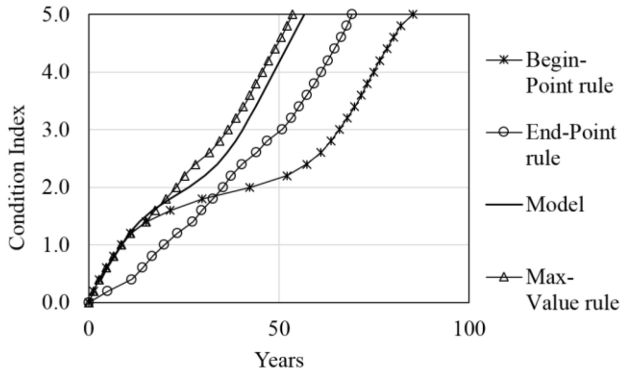


*Figure 7 Alternative assignment rules; adding an S-shaped deterioration curve and a larger sample size*

As in the previous three cases, the Max-Value rule provides a reasonable representation of the modelled deterioration process. It is also closer to the model in absolute terms.

We interpret the simulations first and foremost as a confirmation of the ability of the piece-wise linear curves to recover the form of an underlying deterioration process. We further conclude that the Max-Value rule is to be preferred over the other rules for stochastic deterioration processes like those we have simulated. This rule reduces the unwanted effects of error structures that are likely to exist in infrastructure deterioration data. There is, however, a tendency to produce a small positive bias. We see this as a small defect. The goals of the method are primarily qualitative. We are interested in the shape of the deterioration behaviour as well as relative statement about the impact of influencing factors. We now proceed to show curves constructed from actual in-situ data.

## 4 PIECEWISE LINEAR CURVES TO EXTRACT ADDITIONAL OBSERVATIONS FROM THE DATA

Having demonstrated how the PLC method can recover the shape of deterioration behaviour over the life-cycle of a long lived asset, even if only two inspection observations are available, we now turn to the second and third purposes of the PLC method listed in the introduction. These are the exploration of hypothesis about the data without specifying a functional form and the generation of a set of dependent variable values to test a model of the length of service life, when such values are otherwise missing. After a brief description of the data set used to construct the curves, we provide three examples and discuss the qualities we find novel. The qualities can roughly be summarized as follows: 1) it is possible to recover long term deterioration behaviour even in absence of age data, 2) the impact of increasing exposures to influencing factors can be observed over the life cycle, and 3) the shape of the observable deterioration behaviour suggests a large but differentiated impact of maintenance interventions in interaction with other influencing factors on the length of service life.

The examples share an underlying theme about pavement deterioration. This is, that climate-related factors may have a greater relevance in determining the length of service life than is normally presented in the literature. At least in our data set, the climate factors seem to play an even greater role than traffic. Other data sets also support the importance of climate factors (Anyala et al., 2014; Khattak, Nur, Bhuyan, & Gaspard, 2014; Wu et al., 2019).

### 4.1 *Description of the data set*

The data set used for the non-parametric curves is described in detail in Craig Richmond et al. (2019). We provide a short summary here. Cantons in Switzerland are political sub-divisions that maintain an intercommunal network of roads. The largest canton owns and maintains more than 2000 km of road, the smallest less than 100 km. Switzerland is small relative to most countries, but it has considerable climatic variation. Two mountain ranges divide the country into multiple climatic regions. As with most countries, Switzerland can also be subdivided along the spectrum of population density. There are regions with urban centres and corresponding high traffic loads and regions with dispersed populations and low traffic loads. This two-dimensional differentiation by climate and traffic results in a rich data set that is advantageous for a comparison between the relative effects of axle loading versus climate-related loading. At the same time, all the Swiss cantons share a common set of road design and construction standards. These are conservative standards based on the AASHTO 1993 Guidelines. They result in relatively thick pavements (VSS, 2011). The complete data set had 358,000 records for which four observations were simultaneously available as shown in Table 2. These values are before any filtering was done. Two observations were needed to calculate the difference in the $I^1$ and two were needed to calculate the difference in $I^3$. As previously described, the indices measure general surface defects and rutting respectively. The need for both pairs was for filtering as described in Section 3.3.2. This very large number of observations is particularly useful in the piecewise linear approach.

In a parametric approach, the effect of an influencing factor is captured by the introduction of a variable measuring the exposure to the factor. Each additional variable requires some additional observations to achieve the same certainty of estimation, but one does not normally need to double the data to achieve similar statistical certainty after adding one additional influencing factor. The PLC method is far more data intensive than a multivariate regression. In fact, for each exposure level to an influencing factor, a complete data set is required to construct an independent curve. One advantage of this is that the curves are statistically independent, being based on completely separate data. We return to this aspect in section 4.4. The subscript *j* in equation (4) stands for these divisions of the total data into disjoint sets. Each curve presented below required on the order of 10000 observations. This number would only be 1000 if the road sections were uniformly distributed over the condition range, and if each of the 50 subdivisions required only 20 observations to calculate the mean. In practice, the data is far from uniformly distributed. In order to have 20 observations at the higher condition subdivisions, the lower subdivisions had up to 400 observations. This is an attribute of the empirical road networks and it cannot be changed by the researcher.

*Table 2 Number of observations in the data set*

| Statistic | At least one observation $I^1$ Index | At least one pair $I^1$ Index | One pair of both $I^1$ Index and $I^3$ Index |
|---|---|---|---|
| Count Observations | 647,000 | 574,000 | 358,000 |
| Total Length (km) | 16,494 | 14,809 | 8972 |

### 4.2 *Observing the impact of heavy vehicles on deterioration*

Our first example compares two cases, both of which are subdivided by the degree of exposure to heavy vehicles. These are measured by the average annual daily truck traffic, AADTT. The first case, Figure 8, shows a set of PLCs based on the rutting index, $I^3$. Here the curves create a clear pattern that is in accordance with intuition. The second case, Figure 9, repeats the same exercise except this time using the damage index $I^1$, which considers all visible surface defects – including rutting. In the second case, the pattern does not fit intuition. In fact, no clear pattern arises.

#### 4.2.1 *Heavy vehicles and rutting deterioration*

Figure 8 shows a set of PLCs using the index $I^3$. The data set was first subdivided into five parts by AADTT and then separate curves were constructed according to equation (4) with *j* indexing the degree of exposure to AADTT. The filtering procedure presented in Section 3.3.2 and the Max-Value rule were applied. One can see that it is possible to construct a depiction of deterioration behaviour over the life cycle. These appear regular and non-random. Second, one can observe a reasonable correspondence between the speed of deterioration and the exposure to the influencing factor, heavy vehicles. We use the condition level 3 as the terminal state to calculate the length of service life up to that state. The roads having between 200 and 400 AADTT last 35% longer than roads having between 600 and 800: 37.6 versus 28.0 years respectively. The absolute difference is 11.6 years.

The level of traffic that a road actually experiences is not a random occurrence; it is expected when the road is designed. The design of a road is therefore endogenous. The same is true of the subsequent deterioration behaviour. Thus, there is no *necessary* contradiction in the fact that two extreme loading cases (high and low), ‘0 – 200’ and ‘> 800’ are not also the extreme deterioration cases (fast and slow). This may reflect a design choice. For example, roads with higher traffic loads may also generally have higher speed limits, making the need to prevent rutting of greater concern. The PLCs reflect the net result of all the processes leading to observable, in-situ performance.

Third, patterns can be identified in the shapes of the deterioration curves. All the curves show, to a slight degree, an s-shape. One can also see a separation between the curves roughly from year zero. This may not be surprising, but the separation-from-zero pattern will not be found in all the sets of curves.

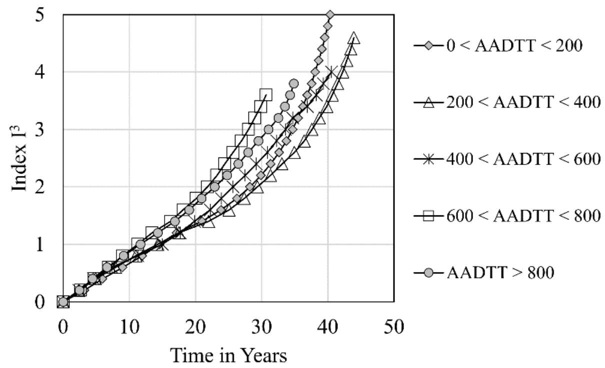


*Figure 8 Non-Parametric deterioration curves of rutting*

### 4.2.2 Heavy vehicles and general surface deterioration

Whereas Figure 8 provides a reasonable match to expectations, Figure 9 is surprising. Figure 9 repeats the same experiment as Figure 8 except that the index $I^1$ replaces $I^3$.

Again, the piecewise linear method can depict the deterioration process over the life cycle and these curves have a consistent, non-random, shape. Second, it is possible to create curves based on the degree of exposure to the influencing factor, heavy vehicles. However, in this case, no consistent ranking results. The length of service up to a condition of 3 is spread over only 3 years from a low value of 21.3 to a high value of 24.3. This is much less than the 11.6-year spread in the rutting example. The mid-ranges of AADTT have the shortest service lives. The second highest exposure and the lowest exposure tie for the longest service life. No clear pattern can be identified.

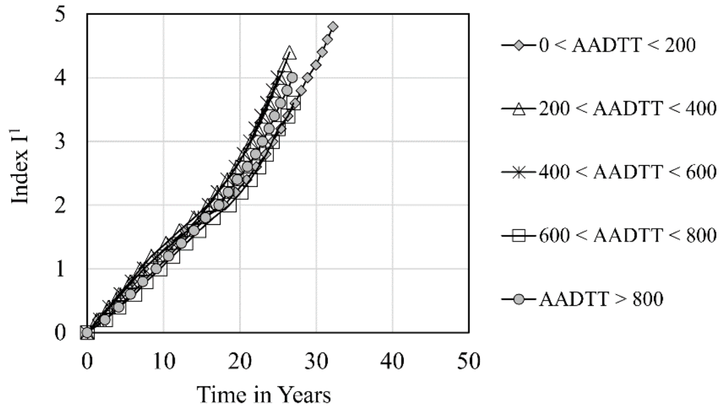


*Figure 9 Non-Parametric deterioration curves of general surface damage*

Third, despite the lack of correspondence to the degree of exposure, the s-shape is clearly visible in all the curves. The s-shape is more clearly defined than in Figure 8. Overall, one is struck by the similarity of the five curves, which is not an obvious result, since these five curves are based on independent data sets. The consistency lends support to the piecewise linear method in general and to the robustness of the finding, that the s-shape reflects a feature of the deterioration process that generates this data set. Surprisingly, in distinction to Figure 8, the curves do not separate from year zero. What little separation does occur, happens later in the life cycle.

We interpret the evidence from Figure 8 and Figure 9 as confirming the ability of the PLC method to recover the shape of the underlying process in a reliable way. The figures also confirm, through their differences and internal consistency, that the method leads to observations that are non-obvious and not otherwise available.

Although intuition suggests that a larger number of heavy vehicles should be strongly correlated with shorter service lives in both cases, this is not what the figures show. Figure 8 demonstrates the ability of the method to identify differences, when they are present. Figure 9 demonstrates that the $I^1$ does not lead to identifiable differences under the same circumstances. For a parametric model of this data set, both figures recommend the use of a mathematical model that is capable of taking an s-shape.

### 4.3 Using natural experiments to explore influencing factors.

One can use the non-parametric method to explore the effect of influencing factors through indirect means. For example, one can use naturally occurring experiments to explore the effects of some influencing factors. We present some examples related to temperature and to temperature change on in-situ deterioration behaviour. One proceeds as above and divides the data along the lines of the natural experiment and then observes the differences in shape and slope of the resulting curves.

Switzerland has many bridges and tunnels as well as many areas where roads are protected from sunlight by either a mountainside or by forest cover. Since traffic is unaffected by these conditions, a natural experiment results where the exposures to climate-related variables differ while traffic is held constant. If the variable factors are important, then the resulting curves will display a clear and consistently differentiated pattern as with the rutting example. If they are not, no clear differentiation will result.

If the tunnel is sufficiently long, then the pavements will not be subject to rainfall, annual temperature variation, or daily temperature amplitude. On bridges, the pavements are less exposed to temperature-induced stresses because the foundation is expanding and contracting along with the pavement. Forest coverage reduces the daily amplitude of temperature by providing shade, but it provides no protection against the seasonal temperature swings or rainfall. None of these cases are the equivalents of laboratory experiments, where only the factor is varied. Their advantage is that they occur naturally and are provided to researchers free of charge. The results are indicative. Negative results, such as the lack of separation in Figure 9, provide stronger evidence. But even the positive cases, for instance those that show a clear and consistent differentiation between curves, give strong evidence that something about the environment matters. But correlation is not causation, further investigations must be made to confirm what is suggested in the PLCs. Examples are given in Figure 10 and Figure 11.

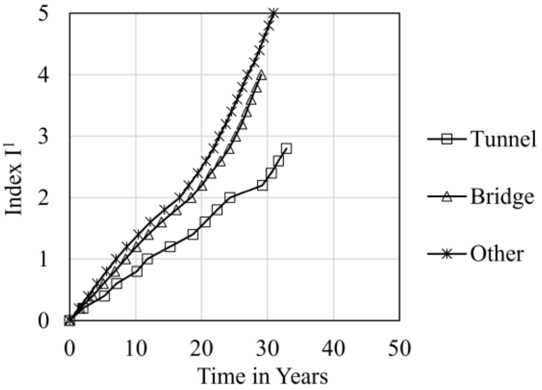


*Figure 10 Deterioration curves for tunnels, bridges and other*

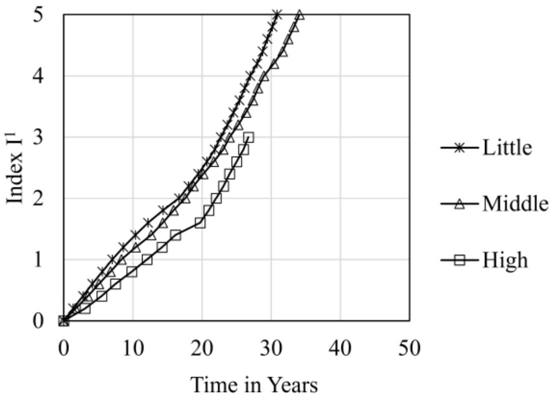


*Figure 11 Deterioration curve for the degree of forest coverage*

Figure 10 provides evidence through the spread of the curves that pavements in tunnels and on bridges deteriorate more slowly than pavements that are neither in tunnels nor on bridges. The pavement in the tunnel group lasts 50% longer than the 'other' group: 33 versus 22 years up to a condition of 2.8. The reason that the curve for the tunnel group ends at the level 2.8 is that there were not enough observations from tunnels with a condition worse than 2.8 to estimate the higher ranges. The pavements in the bridge group last 10% longer. An interesting feature of both the bridge group and the tunnel group is the lower slope from year zero. This is true despite the use of the $I^1$ index to measure damage. In the bridge group and the other group, the s-shapes can be observed.

Figure 11 shows the result of a second natural experiment. It compares average deterioration depending on the degree of forest coverage. The boundaries between the three subdivisions of the data were set at 30% and 70% forest coverage. The spread in Figure 11 is not as wide as in some of the other figures, but the separation between the curves is still distinct. The three groups reach a condition of 3 at 23, 24 and 27 years respectively, a difference of 17% from lowest to highest. The ranking is consistent with the hypothesis that forest coverage reduces daily temperature amplitude and therefore slows deterioration. As with tunnels, the difference in deterioration rates can be seen from the outset. Beginning at a condition level of 1.5, there is a slight s-shape. It is followed by a third phase where the rates of deterioration across the groups appear to be roughly similar. Apparently, whatever advantage the forest cover provides, the effect is lost once a condition of roughly 1.5 has been reached.

### *4.4 Creating data to test models when it otherwise does not exist.*

In Section 2.2 we argued that the PLCs are also sample statistics of the data and these can be used to test predictive models. In this section we present an example. In principle, each node on a curve is a sample statistic. If $\chi = 3$, then $t_j^\chi$ is a data-based observation of the time it takes to reach the condition 3. Like a sample mean, this statistic has stochastic properties. It has a variance; it may be biased. We leave a study of the properties to future research. Our purpose here is only to provide an example of its use.

In a previous project (Craig Richmond et al., 2019), a predictive model for the Swiss $I^1$ index was developed. It predicted the expected time required to reach $\chi = 3$. The modelled value is referred to here as the $I^{1E}$. The value is arrived at by evaluating a closed-form equation. The independent variables of the $I^{1E}$ equation are 1) four climate variables that describe the local stochastic distributions of temperature and temperature change, and 2) the number of heavy vehicles. The modelled values vary continuously over space, which leads to a rich and varied set of predictions in a situation like Switzerland. Further details on the model can be found in English in (C. Richmond et al., 2018).

A strategy for testing the model is to first divide the road condition data by the calculated $I^{1E}$ value as was done in the earlier figures using AADTT. All the locations will be assigned to the same data-subdivision $j$, if they share the same or similar $I^{1E}$ values. That is, the subdivisions reflect the model predictions. If the resulting curves reveal a clear and differentiated ranking, and the rankings correlate well to the mean $I^{1E}$ of each subdivision , then that is evidence in favour of the $I^{1E}$ model. The process is not dissimilar to testing whether higher levels of exposure to heavy vehicles leads to higher rates of rutting damage, by comparing the average rate of rutting between roads with heavy traffic and roads and low traffic. The only difference is that the $I^{1E}$ is a model. If the subdivisions have similar means, then the model is not succeeding in predicting rates of deterioration.

Figure 12 presents the curves for this test. The subdivisions comprise equally wide ranges of $I^{1E}$ values, but the number of observations in each subdivision varies considerably. Group 1 includes all those locations having the shortest predicted service lives and Group 9 the longest. Group 7 is missing because there were insufficient locations with that particular level of $I^{1E}$ values to construct the curve. With the exception of Group 5, the ranking of curves matches the ranking of $I^{1E}$ values almost exactly. For the formal statistical test of the $I^{1E}$ values, the data was divided into groups containing an equal number of road sections, but still ordered by $I^{1E}$. It led to 35 curves and a stronger test. It was possible using this method to demonstrate a highly statistically significant relationship between the $I^{1E}$ values and the empirical speed of deterioration (Craig Richmond et al., 2019).

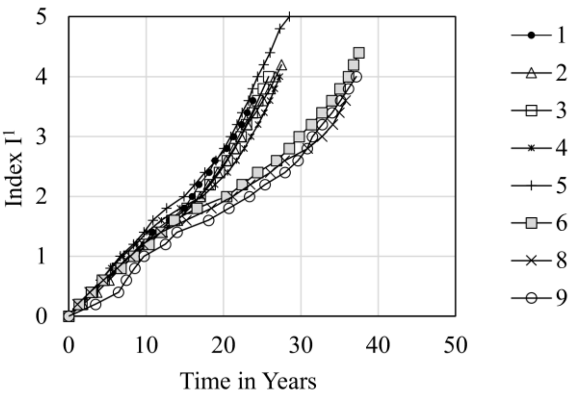


*Figure 12 Non-Parametric deterioration curves using a combined index of climate and axle load as difficulty factors.*

As in the earlier figures, the PLC method permits an analysis of the shape and spread of the curves. The spread between the fastest and slowest at $\chi = 3$ is 11 years. The grouping of the $I^1$ index by AADTT only resulted in a 3-year spread. Apparently, there is a systematic variation in the rates of deterioration in underlying data, but this cannot be found using AADTT alone. By including climate factors as well, a far better identification of geographic locations where road service lives can be expected to be short is possible.

Further, Figure 12 indicates a three-phase, s-shaped, life cycle pattern. In the first phase, between conditions 0 and 1.5, deterioration occurs at similar rates across all groups. Deterioration appears to be roughly linear. In the second phase, between the conditions 1.5 and 2.5, there is a distinct flattening of the deterioration behaviour. In the third phase, after a value near 2.5, the rate of deterioration increases at an increasing rate.

Perhaps the most interesting observation that can be seen in the shapes of the curves is that most of the eventual variation in the length of service life occurs in the second phase. We hypothesize this reflects the differential longevity of smaller maintenance management activities, which are concentrated in the mid-condition ranges. More challenging climates for roads in general, are more challenging *with respect to maintenance activity* in particular. One might call this attribute a difference in *maintainability*. It can be picked up in a parametric model that interacts climate variables with the second service life phase in some fashion. If the three-phase structure is ignored in a parametric model, then estimations will be imprecise and statistically less certain.

We conclude from all three examples, that the PLC method offers researchers a new and interesting tool for exploring data prior to doing a parametric estimation. In particular, it provides a basis for selecting a functional form for a deterioration process, which, in the in-situ case, has many plausible forms. It also provides a means to identify key influencing factors and it can suggest how the interaction should be specified amongst the numerous mathematical options. Finally, the method enables the construction of dependent variables to test models that may otherwise not be available.

## 5 CONCLUSIONS AND IMPLICATIONS

A non-parametric methodology for deriving curves to represent the deterioration behaviour has been presented. It falls into the category of piecewise linear models. The methodology is first applied to a simulated data set. It has been demonstrated that the so-called Max-Value rule leads to reasonable results. The results have been shown to include a small positive bias. The Max-Value rule has an additional advantage as an effective filter on unobserved maintenance interventions, when these are concentrated in the mid years of service life.

The piecewise linear method was then applied to a large data set from Swiss cantonal roads. It led to different shapes and degrees of separation between curves. The curves have an overall consistent, non-random appearance. Some groups of curves separate into distinct and consistently ranked curves that reflect the degree of exposure to the influencing factor. Others do not separate, but still have a consistent overall appearance, suggesting they are random draws from the same basic distribution. Taken together, the method recovers the shapes of the underlying deterioration process and, if an influencing factor has a sufficiently large effect, a ranking according to the exposure to the influencing factor will result.

Factors related to climate were consistently shown to have a significant influence in this data set. Surprisingly, the same could not be shown for heavy vehicles with the exception of rutting. This presumably reflects the endogeneity of road structural design and the fact that traffic loading versus bottom-up fatigue failure has a sufficient engineering solution. It was shown that the method can use natural experiments to investigate the potential importance of influencing factors if a correlation exists to some identifiable feature of the road environment. Tunnels, bridges, and forest cover are examples.

It was demonstrated that the method can generate otherwise missing values of dependent variables to test a model of the length of service life. The specific example generated a set of curves that are widely spread and consistently ranked. The shapes suggest a three-phase model of deterioration behaviour. The greatest differentiation occurs in the second phase when the minor maintenance activity is likely to be most intense. This suggests maintainability as a key variable for future research in deterioration modelling. Whether these characteristics extend to other data sets remains to be tested.

Generally, we conclude that the presented non-parametric method is a useful way to explore the deterioration behaviour within large data sets having a relatively short time coverage and no data on the age of the pavements. An interesting topic for further research would be a direct comparison between non-parametric and parametric calculations if an appropriate data set can be found. It would have to be both very large and contain multiple condition measurements at each location. Further, it would have to contain the pavement's absolute age at each section.

## 6 ACKNOWLEDGMENTS

The authors gratefully acknowledge the comments of a number of anonymous reviewers. These have improved the paper. We also acknowledge the contribution of the previously unpublished figures shown in Section 3. These were generated during a study carried out at the Institute for Infrastructure Management at the ETH in Zürich and financed by the grant VSS 2014 411 provided by the Swiss Federal Roads Office, FEDRO. This research benefits from the contributions of Fabian Achilles who was a co-author on the previous study as well as from comments by its supervising committee. In particular, we wish to thank Dr. Thomas Arn who contributed the observation regarding tunnel pavements.

## 7 REFERENCES

AASHTO. (1993). Guide for design of pavement structures (4 ed.). Washington D.C.: American Association of State Highway and Transportation Officials.

Anastasopoulos, P. C., Mannering, F. L., & Haddock, J. E. (2012). Random parameters seemingly unrelated equations approach to the postrehabilitation performance of pavements. *Journal of Infrastructure Systems*, 18(3), 176-182.

Anyala, M., Odoki, J. B., & Baker, C. J. (2014). Hierarchical asphalt pavement deterioration model for climate impact studies. *International Journal of Pavement Engineering*, 15(3), 251-266. doi:10.1080/10298436.2012.687105.

Archilla, A. R. (2006). Repeated measurement data analysis in pavement deterioration modeling. *Journal of Infrastructure Systems*, 12(3), 163-173. doi:10.1061/(ASCE)1076-0342(2006).

Bardaka, E., Labi, S., & Haddock, J. (2014). Using Enhanced Econometric Techniques to Verify the Service Life of Asset Interventions. Transportation Research Record: *Journal of the Transportation Research Board*, 2431(-1), 16-23. doi:10.3141/2431-03.

Barnett, W. A. (1983). Definitions of 'second order approximation' and of 'flexible functional form'. *Economics Letters*, 12(1), 31-35. doi: 10.1016/0165-1765(83)90108-8.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429-444.

Chen, D., & Mastin, N. (2015). Sigmoidal Models for Predicting Pavement Performance Conditions. *Journal of Performance of Constructed Facilities*. doi:10.1061/(ASCE)CF.1943-5509.0000833.

Chu, C. Y., & Durango-Cohen, P. L. (2008). Incorporating maintenance effectiveness in the estimation of dynamic infrastructure performance models. *Computer-Aided Civil and Infrastructure Engineering*, 23(3), 174-188.

Fried, H. O., Lovell, C. A. K., & Schmidt, S. S. (2008). *The Measurement of Productive Efficiency and Productivity Growth*, Oxford University Press.

Gao, L., Hong, F., & Ren, W. (2019). Impacts of Seasonal and Annual Weather Variations on Network-Level Pavement Performance. *Infrastructures*, 4, 27. doi:10.3390/infrastructures4020027.

Golabi, K., Kulkarni, R. B., & Way, G. B. (1982). A Statewide Pavement Management System. *Interfaces*, 12(6), 5-21. doi:10.1287/inte.12.6.5.

Gu, W., Ouyang, Y., & Madanat, S. (2012). Joint optimization of pavement maintenance and resurfacing planning. *Transportation Research Part B: Methodological*, 46(4), 511-519. doi: 10.1016/j.trb.2011.12.002.

Highway Research Board. (1961). *The AASHO road test: history and description of project*. Washington, D.C..

Highway Research Board. (1962). *The AASHTO Road Test: Pavement Research.* Washington D.C..

Hong, F., & Prozzi, J. A. (2015). Pavement Deterioration Model Incorporating Unobserved Heterogeneity for Optimal Life-Cycle Rehabilitation Policy. *Journal of Infrastructure Systems*, 21(1). doi:10.1061/(ASCE)IS.1943-555X.0000168.

Judge G., Griffiths W., Hill R., Lütkepohl H., & Lee T. (1985). The Theory and Practice of Econometrics (Second ed.), Wiley.

Kalaiselvan, C., & Bhaskara Rao, L. (2016). Comparison of reliability techniques of parametric and non-parametric method. *Engineering Science and Technology, an International Journal*, 19(2), 691-699. doi: 10.1016/j.jestch.2015.11.002.

Kerali, H. (2000). Overview of HDM-4. Vol. 1 of Highway Development and Management. Paper presented at the PIARC 2000, Paris.

Khattak, M. J., Nur, M. A., Bhuyan, M. R. U. K., & Gaspard, K. (2014). International roughness index models for HMA overlay treatment of flexible and composite pavements. *International Journal of Pavement Engineering*, 15(4).

Kobayashi, K., Kaito, K., & Lethanh, N. (2012). A statistical deterioration forecasting method using hidden Markov model for infrastructure management. *Transportation Research Part B: Methodological*, 46(4), 544-561. doi: 10.1016/j.trb.2011.11.008.

Kvam, P. H., & Vidakovic, B. (2007). *Nonparametric Statistics with Applications to Science and Engineering*, Wiley-Interscience.

Lee, J., & Madanat, S. (2015). Jointly optimal policies for pavement maintenance, resurfacing and reconstruction. *EURO Journal on Transportation and Logistics*, 4(1), 75-95. doi:10.1007/s13676-014-0065-1.

Lethanh, N., Richmond, C., & Adey, B. T. (2016). Investigation of the Ability to Estimate Values of Road Section Condition Indicators Based on Their Spatial Correlation. *Journal of Infrastructure Systems*, 22(3). doi:10.1061/(ASCE)IS.1943-555X.0000290.

Little, D. N., Bhasin, A., & Darabi, M. K. (2015). 7 - Damage healing in asphalt pavements: theory, mechanisms, measurement, and modeling. In S.-C. Huang & H. Di Benedetto (Eds.), *Advances in Asphalt Materials* (pp. 205-242). Oxford: Woodhead Publishing.

Luo, Z. (2013). Pavement performance modelling with an auto-regression approach. *International Journal of Pavement Engineering*, 14(1), 85-94.

Madanat, S., Mishalani, R., & Ibrahim, W. H. W. (1995). Estimation of Infrastructure Transition Probabilities from Condition Rating Data. *Journal of Infrastructure Systems*, 1(2), 120-125. doi:10.1061/(ASCE)1076-0342(1995)1:2(120).

Madanat, S. M., Karlaftis, M. G., & McCarthy, P. S. (1997). Probabilistic infrastructure deterioration models with panel data. *Journal of Infrastructure Systems*, 3(1), 4-9.

Meegoda, J. N., & Gao, S. (2014). Roughness progression model for asphalt pavements using long-term pavement performance data. *Journal of Transportation Engineering,* 140(8). doi:10.1061/(ASCE)TE.1943-5436.0000682.

O'Leary, M., & Walsh, J. (2018). Bridge Element Deterioration of Concrete Substructures (WA-RD 893.1). Retrieved from https://www.wsdot.wa.gov/research/reports/fullreports/893-1.pdf on 26.12.2018.

Paterson, W. D. O. (1987). *Road Deterioration and Maintenance Effects: Models for Planning and Management*: Johns Hopkins University Press, Baltimore, Md.

Perera, R. W., Byrum, C., & Kohn, S. D. (1998). *Investigation into the development of pavement roughness*. Federal Highway Administration, Washington D.C.. doi: 10.3141/1655-14.

Prozzi, J. A., & Madanat, S. M. (2003). Incremental nonlinear model for predicting pavement serviceability. *Journal of Transportation Engineering*, 129(6), 635-641.

Ray, S. C. (2004). *Data Envelopment Analysis: Theory and Techniques for Economics and Operations Research*: Cambridge University Press.

Reger, D., Christofa, E., Guler, I., & Madanat, S. (2013). Estimation of Pavement Crack Initiation Models by Combining Experimental and Field Data. *Journal of Infrastructure Systems*, 19(4), 434-441. doi:10.1061/(ASCE)IS.1943-555X.0000148.

Richmond, C., Achilles, F., & Adey, B. T. (2018). Approximating an Analytic Solution for the Optimal Design of Asphalt Pavement: Limited Example. *Journal of Transportation Engineering, Part B: Pavements*, 144(4). doi: 10.1061/JPEODX.0000070.

Richmond, C., Adey, B. T., & Achilles, F. (2019). Ermöglichung eines realistischen Benchmarkings des baulichen Unterhalts der Kantonsstrassen (BORMAC) (in German). Retrieved from www.mobilityplatform.ch on 18.12.2019.

Richmond, C. M., Archilla, A. R., & Adey, B. T. (2020). Are current design service lives for asphalt concrete pavements suboptimal? An analytic argument for longer design service lives. *Sustainable and Resilient Infrastructure*, 1-23. doi:10.1080/23789689.2020.1719730.

Small, K. A., & Winston, C. (1988). Optimal Highway Durability. *The American Economic Review*, 78(3), 560-569. doi:10.2307/1809153.

Tsuda, Y., Kaito, K., Aoki, K., & Kobayashi, K. (2006). ESTIMATING MARKOVIAN TRANSITION PROBABILITIES FOR BRIDGE DETERIORATION FORECASTING. *STRUCTURAL ENGINEERING / EARTHQUAKE ENGINEERING*, 23(2), 241s-256s. doi:10.2208/jsceseee.23.241s.

Ullidtz, P. (2005). Simple Model for Pavement Damage. *Transportation Research Record*, 1905(1), 128-137. doi:10.1177/0361198105190500114.

VSS. (2003) Norm 640925b Zustandserhebung und Indexbewertung (German/French). Zürich: VSS.

VSS. (2011) Norm 640324: Dimensionierung des Strassenaufbaus; Unterbau und Oberbau (German /French). Zürich: VSS.

Walther, A., & Wistuba, M. (2012). Mechanistic Pavement Design Considering Bottom-Up and Top-Down-Cracking. In A. Scarpas, N. Kringos, I. Al-Qadi, & L. A (Eds.), *7th RILEM International Conference on Cracking in Pavements: Mechanisms, Modeling, Testing, Detection and Prevention Case Histories* (pp. 963-973). Dordrecht: Springer Netherlands.

Wu, S., Wen, H., Zhang, W., Shen, S., Mohammad, L. N., Faheem, A., & Muhunthan, B. (2019). Field performance of top-down fatigue cracking for warm mix asphalt pavements. *International Journal of Pavement Engineering*, 20(1), 33-43. doi: 10.1080/10298436.2016.1248204.

Yamany, M. S., Saeed, T. U., Volovski, M., & Ahmed, A. (2020). Characterizing the Performance of Interstate Flexible Pavements Using Artificial Neural Networks and Random Parameters Regression. *Journal of Infrastructure Systems*, 26(2). doi:10.1061/(ASCE)IS.1943-555X.0000542.

Yang, J., Gunaratne, M., Lu, J., & Dietrich, B. (2005). Use of Recurrent Markov Chains for Modeling the Crack Performance of Flexible Pavements. *Journal of Transportation Engineering*, 131(11), 861-872. doi: 10.1061/(ASCE)0733-947X(2005)131:11(861).