Predicting Remaining Useful Life with Similarity-Based Priors

Youri Soons[1], Remco Dijkman[2(✉)], Maurice Jilderda[3],
and Wouter Duivesteijn[2]

[1] Sitech Services B.V., Geleen, The Netherlands
`youri.soons@sitech.nl`
[2] Technische Universiteit Eindhoven, Eindhoven, The Netherlands
`{r.m.dijkman,w.duivesteijn}@tue.nl`
[3] Perfact Group, Munstergeleen, The Netherlands
`mauricejilderda@perfact-group.com`

Abstract. Prognostics is the area of research that is concerned with predicting the remaining useful life of machines and machine parts. The remaining useful life is the time during which a machine or part can be used, before it must be replaced or repaired. To create accurate predictions, predictive techniques must take external data into account on the operating conditions of the part and events that occurred during its lifetime. However, such data is often not available. Similarity-based techniques can help in such cases. They are based on the hypothesis that if a curve developed similarly to other curves up to a point, it will probably continue to do so. This paper presents a novel technique for similarity-based remaining useful life prediction. In particular, it combines Bayesian updating with priors that are based on similarity estimation. The paper shows that this technique outperforms other techniques on long-term predictions by a large margin, although other techniques still perform better on short-term predictions.

Keywords: Remaining useful life · Trajectory based similarity prediction · Bayesian updating · Similarity estimation · Prognostics · Prediction

1 Introduction

Prognostics is the area of research that concerns the prediction of the remaining useful life (RUL) of machines or machine parts. A RUL prediction is a prediction of the time until a machine or machine part must be replaced or repaired. It is important that such predictions are accurate: early predictions lead to unnecessarily frequent maintenance with associated costs, while late predictions increase the risk of a machine break down with associated loss of production time and possibly sales.

Data-driven RUL prediction is based on run to failure data, i.e., observations on what happened to a part or machine in a run from the last maintenance

activity to the next. Figure 1 shows a typical example of run to failure data, in this case data of a filter in a chemical plant. The figure shows condition measurements on the filter over time, in terms of the difference in pressure before and after the filter. It shows that this difference is close to zero for some time. Then, the filter starts to clog up and the pressure builds up, until the filter is replaced and the pressure difference returns to normal. The resulting 'sawtooth' shape is frequently observed in run to failure data.
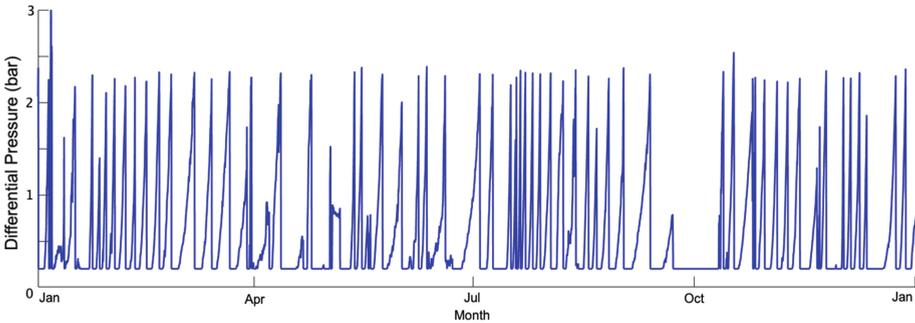
Fig. 1. Example run to failure data.

RUL prediction on run to failure data can be done by fitting a model, such as a regression model or a probability distribution, on the data. Many different techniques exist for those purposes [1]. However, as is evident from Fig. 1, different runs may have very different durations or shapes, and RUL prediction techniques rely on additional data to accurately predict the duration and shape of a particular run. Unfortunately, additional data is often unavailable or hard to relate to the run to failure data [2]. If additional data is unavailable, it is unclear which condition measurements are reliable and of course what their influence is on the RUL. One way to overcome these problems is to use similarity-based techniques, which work based on the hypothesis that, if a curve has developed similarly to some collection of other curves until now, it will likely continue to develop like that, and have a similar remaining useful life.

This paper explores the performance of two similarity-based techniques: trajectory-based similarity prediction, and Bayesian updating. It then adds its own: Bayesian updating with similarity-based priors. The contribution of this paper consists of this technique, described in Sect. 3.4, as well as a detailed evaluation of all three techniques in a case study from practice, described in Sect. 4.

Against this background, the remainder of this paper is structured as follows. Section 2 presents related work on remaining useful life prediction. Section 3 presents similarity-based remaining useful life prediction techniques, including the new technique. Section 4 compares the performance of the various techniques in a case study and Sect. 5 presents the conclusions.

2   Related Work

RUL prediction can be considered a specialized form of survival analysis [10]. Essentially, two types of techniques exist for predicting RUL: model-based and data-driven techniques. Model-based techniques use physical models to accurately represent the wear and tear of a component over time [5]. Data-driven techniques do not presume any knowledge about how a component wears out over time, but merely predicts the RUL based on past observations. Hybrid models, which are a combination of physical and data-driven techniques, also exist [9]. This paper focuses on data-driven models, which are most suited when the physical mechanisms that cause a component to fail are too complex to model cost-effectively, or if they are not sufficiently understood.

A large number of data-driven techniques is available that fall into two classes depending on whether or not a probability distribution of the RUL must be obtained or a point-estimate is sufficient [1]. A probability distribution of the RUL has several benefits [16,17,20]. For example, it facilitates stochastic decision making, where maintenance is done when the probability that a part will fail exceeds a certain threshold, which is in line with the way in which maintenance decisions are made. When it is not necessary to produce a probability density function, several models can be used. The most obvious choices include regression models that use time as the primary independent variable and time-series models. However, regression models require that the behavior of the curve is predictable over time [4,13] and time-series [12] models are only suitable for short-term predictions [3,16] or when the behavior of the curve is predictable over time. Regression models that take other variables into account can also be used [6]. Such models have the benefit that they do not only consider the dependency of the RUL on the time that the part has been in operation, but also on other relevant factors, such as the operational temperature or vibration of the part.

When the RUL depends on other factors beyond time, but data on such factors is not available, one can include them as a black box. While we may not know the values of relevant factors, we can still find historical runs that are similar to the current run. If we assume that the factors that influenced historically similar runs are also similar to the current run, then the future behavior of the current run will also be similar to the behavior of the historically similar runs. This is called Trajectory Based Similarity Prediction (TBSP) [11,18,19]. Bayesian updating techniques use a similar principle [7,8]. Such techniques create a prior probability distribution of the RUL (based on data from historical runs to failure), which updates as more data of the current run is revealed.

3   Prediction Techniques

This section presents similarity-based techniques that can be used for RUL prediction: TBSP and Bayesian updating, which are defined in related work as explained in Sect. 2. Subsequently, Sect. 3.4 presents a novel technique, Bayesian updating with similarity-based prior estimation, which is a combination of TBSP and Bayesian updating.

3.1   Preliminaries

The remaining useful life of a part is defined as follows.

Definition 1 (Remaining Useful Life (RUL)). *Let t be a moment in a run and t_E be the moment in the run at which the part fails. The Remaining Useful Life (RUL) at time t, $r(t)$, is defined as $r(t) = t_E - t$.*

Note that 'failure' can be interpreted broadly. It does not have to be the point at which the part breaks, but can also be the point at which the part reaches a condition in which it is not considered suitable for operation anymore, or a condition in which maintenance is considered necessary. Over time, multiple runs to failure will be observed, such as the runs to failure shown in Fig. 1.

Definition 2 (Run to failure library). *L is the library of past runs to failure. For each $l \in L$, t_E^l is the moment in the run at which the part fails, and $g^l(t)$ is the function that returns the condition of the part at time t of the run.*

The function $g^l(t)$ is created by fitting a curve on the condition measurements of the run. We consider the one-dimensional case here (i.e., the case in which we only measure the condition of the part), but this can easily be extended to a multi-dimensional case (i.e., the case in which we not only measure the condition of the part, but also external factors (i.e., other variables than the condition variable itself), such as the operating temperature or pressure) by considering the observations as vectors over multiple variables. We will also omit the superscript l if there can be no confusion about the run to which we refer.

3.2   Trajectory-Based Similarity Prediction

Figure 2 shows a different (cf. Fig. 1) representation of a run to failure library. It shows all runs in the library, starting from the moment at which the condition variable starts to increase from the base condition. It also shows a 'current' run as a thicker, unfinished curve. The idea of trajectory-based similarity prediction is to find some number k of runs that are most similar to the current run. For each of these k similar runs, we know the time it took until the part failed.
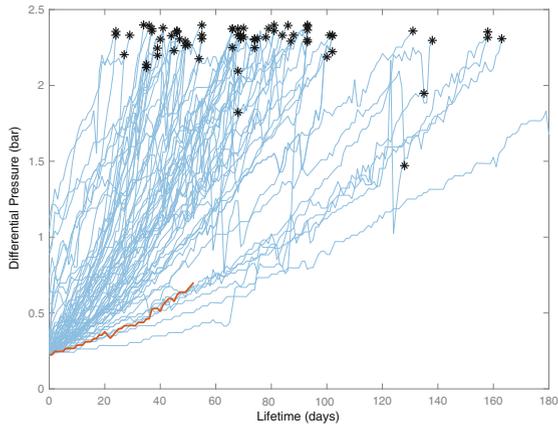

Fig. 2. Example library of runs.

Trajectory-based Similarity Prediction (TBSP) estimates the time until failure as the mean failure time of the similar runs.

Definition 3 (Distance of current run to library run). *At a moment in time t, let I be the number of observations made in the current run, with values $z_1, \ldots, z_I$ observed at times $t_1, \ldots, t_I$, and let $l \in L$ be a library run. We denote by $d^l(t)$ any distance measure contrasting $z_1, \ldots, z_I$ with $g^l(t_1), \ldots, g^l(t_I)$. Let $E^l(t)$ and $M^l(t)$ denote Euclidean and Manhattan distance, respectively.*

Clearly other distance functions can and indeed have been used as well in the context of remaining useful life prediction [21]. An in-depth analysis of the distance function that performs best for TBSP is beyond the scope of this work.

Definition 4 (Fit of current run to library run). *For each library run $l \in L$, let $d^l(t)$ be defined as in Definition 3. The fit of the current run to l is:*

$$S^l(t) = e^{-|d^l(t)|}$$

When, at time t of the current run, the library run l is found that fits the current run best, the remaining useful life of the current run can be predicted as the remaining useful life of that run l: $r(t) = t_E^l - t$. It is also possible to base the prediction of the remaining useful life on the best k runs; sensitivity to k is part of our experiments. If $k > 1$, we can also aggregate RUL predictions by weighted average, where the weights are the goodness of fit of the library runs to the current run.

Definition 5 (Trajectory-based Similarity Prediction). *For each library run $l \in L$, let $S^l(t)$ be the fit of the run to the current run as per Definition 4 and let $r^l(t)$ be the RUL of the run. Let $L' \subseteq L$ be the subset of past runs on which we want to base our RUL prediction. The predicted RUL of the current run, $\hat{r}(t)$, is:*

$$\hat{r}(t) = \frac{\sum_{l \in L'} S^l(t) \cdot r^l(t)}{\sum_{l \in L'} S^l(t)}$$

3.3 Bayesian Updating

A Bayesian updating method has also been proposed to create a probability distribution of the remaining useful life [7,8]. The probability distribution can be updated with each observation of the condition variable that is obtained. The method works by fitting an exponential model to the library runs and subsequently updating that model with observations of the current run.

Intuitively, looking at Fig. 2, Bayesian updating works by fitting a curve to each of the library runs or to a selection of library runs. Based on the resulting collection of curves, a prior probability distribution of the time until the part fails can be created, which represents the 'probable' curve that the current run —or in fact any run—will follow. The prior probability distribution can be updated each time a condition value is observed in the current run. This update leads to a posterior probability distribution that represents the curve that the current run will follow with a higher precision (smaller confidence interval).

Definition 6 (RUL probability density). *For each library run $l \in L$, let $g^l(t)$ be the function that returns the condition of the part at time t of the run. The condition function can be fitted as an exponential model that has the form:*

$$g^l(t) = \phi + \theta e^{\beta t + \epsilon(t) - \frac{1}{2}\sigma^2}$$

Here, ϕ is the intercept, $\epsilon(t)$ is the error term with mean 0 and variance σ^2, and θ and β are random variables.

If we set $\phi = 0$ and take the natural logarithm of both sides, we get:

$$\ln(g^l(t)) = \theta' + \beta t + \epsilon(t)$$

where $\theta' = \ln(\theta) + \frac{1}{2}\sigma^2$. Considering that we have multiple runs $l \in L$, it is possible to fit this equation multiple times to those runs and calculate values for θ', β and σ for each run. With these values, we can compute the prior probability distributions of θ' and β. We assume these distributions are normal distributions with means μ'_0 and μ_1 and variances σ_0^2 and σ_1^2. While the prior distributions are created based on observations from library runs, the distribution can be updated as more observations become available in the current run.

Proposition 1 (RUL probability density updating). *Let $\pi(\theta')$ and $\pi(\beta)$ be the prior distributions of the random variables from Definition 6 with means μ'_0 and μ_1 and variances σ_0^2 and σ_1^2, where $\theta' = \ln(\theta) + \frac{1}{2}\sigma^2$ and σ^2 is the variance of the error term. Furthermore, let there be I observed values, $z_1, \ldots, z_I$, in the current run, made at times $t_1, \ldots, t_I$, and for $i \in I$, let $L_i = \ln(z_i)$ the natural logarithm of each observation. The posterior distribution is a bivariate normal distribution with θ' and β, whose means $\mu_{\theta'}$ and μ_{β}, variances $\sigma_{\theta'}^2$ and σ_{β}^2, and correlation coefficient ρ can be calculated as follows:*

$$\mu_{\theta'} = \frac{\left(\sum\limits_{i \in I} L_i \sigma_0^2 + \mu'_0 \sigma^2\right)\left(\sum\limits_{i \in I} t_i^2 \sigma_1^2 + \sigma^2\right) - \left(\sum\limits_{i \in I} t_i \sigma_0^2\right)\left(\sum\limits_{i \in I} L_i t_i \sigma_1^2 + \mu_1 \sigma^2\right)}{(|I|\sigma_0^2 + \sigma^2)\left(\sum\limits_{i \in I} t_i^2 \sigma_1^2 + \sigma^2\right) - \left(\sum\limits_{i \in I} t_i \sigma_1^2\right)\left(\sum\limits_{i \in I} t_i \sigma_0^2\right)}$$

$$\mu_{\beta} = \frac{(|I|\sigma_0^2 + \sigma^2)\left(\sum\limits_{i \in I} L_i t_i \sigma_1^2 + \mu_1 \sigma^2\right) - \left(\sum\limits_{i \in I} t_i \sigma_1^2\right)\left(\sum\limits_{i \in I} L_i \sigma_0^2 + \mu'_0 \sigma^2\right)}{(|I|\sigma_0^2 + \sigma^2)\left(\sum\limits_{i \in I} t_i^2 \sigma_1^2 + \sigma^2\right) - \left(\sum\limits_{i \in I} t_i \sigma_1^2\right)\left(\sum\limits_{i \in I} t_i \sigma_0^2\right)}$$

$$\sigma_{\theta'}^2 = \sigma^2 \sigma_0^2 \frac{\sum\limits_{i \in I} t_i^2 \sigma_1^2 + \sigma^2}{(|I|\sigma_0^2 + \sigma^2)\left(\sum\limits_{i \in I} t_i^2 \sigma_1^2 + \sigma^2\right) - \left(\sum\limits_{i \in I} t_i\right)^2 \sigma_0^2 \sigma_1^2}$$

$$\sigma_{\beta}^2 = \sigma^2 \sigma_1^2 \frac{|I|\sigma_0^2 + \sigma^2}{(|I|\sigma_0^2 + \sigma^2)\left(\sum\limits_{i \in I} t_i^2 \sigma_1^2 + \sigma^2\right) - \left(\sum\limits_{i \in I} t_i\right)^2 \sigma_0^2 \sigma_1^2}$$

$$\rho = \frac{-\sigma_0 \sigma_1 \sum\limits_{i \in I} t_i}{\sqrt{|I|\sigma_0^2 + \sigma^2}\sqrt{\sigma_1^2 \sum\limits_{i \in I} t_i^2 + \sigma^2}}$$

The proof of this proposition is given in [8]. Consequently, $\ln(g^l(t))$ for the current run to failure l is normally distributed with mean and variance:

$$\mu(t) \cong \mu_{\theta'} + \mu_\beta t - \frac{1}{2}\sigma^2 \qquad \sigma(t) \cong \sigma_{\theta'}^2 + \sigma_\beta^2 t^2 + \sigma^2 + 2\rho t \sigma_{\theta'}\sigma_\beta$$

With this information, the probability that future values of $\ln(g^l(t))$ exceed the maximum acceptable condition at some time t can be computed.

3.4   Bayesian Updating with Similarity-Based Prior Estimation

The RUL probability density function in Definition 6 depends on estimated prior distributions of θ and β. These priors can be set through analyzing previous runs to failure, either based on the complete library of runs, or on a subset of the runs. More precisely, we can determine prior distributions as follows.

Definition 7 (Prior distributions). *For each library run $l \in L$, let $g^l(t)$ be the exponential curve that is fitted to the observations in that run with parameters θ'^l and β^l as in Definition 6. For a subset $M \subseteq L$ of runs, we can determine the mean and standard deviation of θ' and β over all θ'^m and β^m.*

Consequently, our priors depend on the subset $M \subseteq L$ of runs that we use. For example, we can determine our priors based on $M = L$, the complete set of runs. Here, we consider a variant of the Bayesian updating method in which the priors are set based on the runs that are most similar to the current run, using Definition 4 for similarity and thresholds to select the most similar runs. More precisely, we select our priors as follows.

Definition 8 (Similarity-based prior distributions). *Let t be the moment in time at which we determine our prior distributions and k be the number of similar runs on which we base them. Furthermore, let $S^l(t)$ be the similarity of a run l to the observations in the current run until time t as per Definition 4. The set of k most similar runs $M \subseteq L$ at moment t is then defined as the set in which, for all runs $m \in M$, there is no run $l \in L - M$, such that $S^l(t) > S^m(t)$.*

Note that this definition depends on variables t and k, which can therefore be expected to influence the performance of the technique. In our evaluation, we will explore the performance of the technique for different values of t and k.

4   Evaluation

In this section, we put the RUL prediction techniques introduced in Sect. 3 to the test, in a case study with data from practice.

4.1   Case Study

Our data originates from a chemical plant on the Chemelot Industrial Site[1]. The plant we investigate produces a steady flow of various chemical products; whatever the product happens to be, an unwanted byproduct is always generated. Filters have been installed to obtain an untainted final product. These filters have a variable service life, ranging between two and eight days. When the filter performs its function, it withholds residue of the unwanted byproduct. This residue gradually builds up, forming a cake which increases the resistance of the filter. The additional resistance is measured through an increase in differential pressure (δP), as illustrated in Fig. 1. An unclogged filter has a δP of 0.2 bar. When δP reaches a threshold of 2.4 bar, a valve in front of the filter is switched to let the product run through a parallel, clean filter, which returns δP to 0.2 bar and enables engineers to maintain the clogged filter.
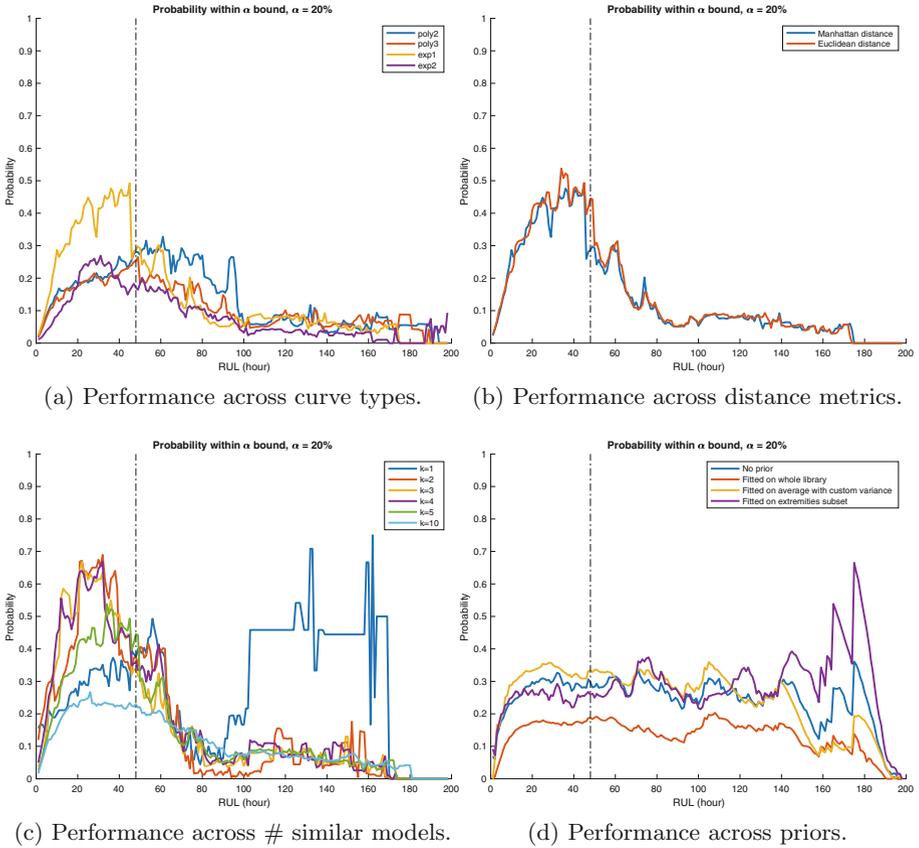
Sensor data, including δP, is stored in a NoSQL database as time series. Preprocessing is needed in several aspects. First, the data has many missing values, which we replace by the last observed value. Second, the sensors generate a data point every second. We established experimentally that resampling the data to the minute barely loses any information from the signal, while still substantially reducing the size of the dataset. Third, to avoid the amplification of clear outliers, they are removed with a Hampel filter [14]. Fourth, we focus on the 'exponential deterioration stage' of the filter's life cycle [5], because—according to the company—the start of that stage is early enough to be able to act on time, and because it provides us with a dataset that is suitable for similarity-based RUL prediction techniques. The start and end of the exponential deterioration stage must be derived from data. We do that by comparing the average pressure over the last hour with its preceding hour. To ensure that every run has only one start per stop, a detected start is ignored if another start was already detected in the same run.

4.2   Results

We quantify our results using an $\alpha-\lambda$ graph. Intuitively, this graph represents the probability that, at a certain moment in the run to failure, the RUL prediction (λ) is within a pre-defined level of precision (α) [15]. We will use a concise representation of the $\alpha-\lambda$ quality: rather than time into the run, we put the RUL on the x-axis, while the y-axis displays the probability. This representation allows us to visually compare different techniques. All analysis is done using 5-fold cross validation. The results presented in the graphs are the averages over the 5 folds.

Figures 3a, b, and c show the performance of the TBSP technique for various parameter settings. Figure 3a compares the performance of TBSP when fitting various types of curves (second ('poly2') and third ('poly3') order polynomials,

---

[1] An anonymized version of the data is made available at: https://surfdrive.surf.nl/files/index.php/s/1dTFFXfZ7woeSUA.

(a) Performance across curve types.

(b) Performance across distance metrics.

(c) Performance across # similar models.
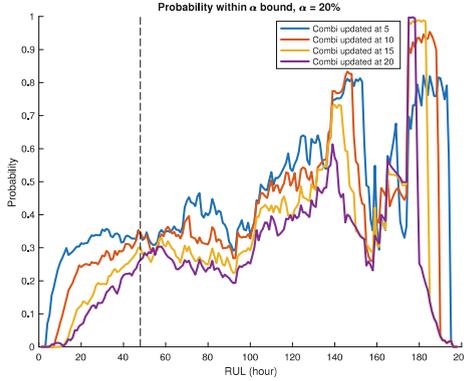
(d) Performance across priors.

Fig. 3. Comparison of hyperparameter settings.

exponential curves ('exp1'), and the sum of two exponential curves ('exp2')),
Fig. 3b compares Manhattan and Euclidean distance, and Fig. 3c shows the sensitivity to the number of similar curves k. The graphs show that TBSP performs
best for an exponential curve in short term (<48 h) predictions, and for $k = 2, 3$,
or 4, while there is little to no performance difference between Manhattan and
Euclidean distance and between $k = 2, 3$, or 4. For those reasons, we parameterize TBSP with exponential curves, using Euclidean distance as a distance
metric, and using 3 similar curves to make the prediction.

Figure 3d shows the performance of the Bayesian updating technique for various prior sets of runs on which the prior is based. We consider four alternatives.
In the first alternative, no prior is defined and the prediction is only computed
based on the current run. In the second alternative, the prior distribution is
based on all runs in the library. In the third alternative, we create a prior distribution by fitting the run with the (closest to) average run to failure time. In
the fourth alternative, we create a prior distribution by fitting the shortest, the

longest, and the average run. The figure shows that for long term predictions, a prior fitted on the 'average', the shortest and longest run performs best, while for short term predictions, a prior fitted on the whole library performs best.

Fig. 4. Performance across moments for setting priors.

Figure 4 shows the performance of Bayesian updating with similarity-based priors for various settings of the moment at which the priors are determined. The best performance is obtained when priors are determined 5 h into the current run to completion; 10, 15, and 20 h were also considered. The number of similar runs on which the priors are based is also a parameter for Bayesian updating with similarity-based priors. The priors are based on the 3 most similar runs. This led to the best results when comparing results for priors based on 1, 2, 3, 4, 5, and 10 similar runs.

Figure 5 shows the results for the various prediction techniques: TBSP, Bayesian Updating, and Bayesian updating with similarity-based priors. The results show a clear distinction in the performance of the different techniques. TBSP performs best for short-term (<48 h before failure) predictions, while Bayesian updating with similarity-based priors performs best in the long term (150–200 h before failure). This is expected, because for long-term prediction, Bayesian updating with similarity-based priors benefits from being based both on similar runs and on general Bayesian behavior, while after some updates the impact of the priors is reduced and the behavior approaches that of normal Bayesian updating. TBSP on the other hand benefits from having a better estimate of the runs to which it is close as time progresses.
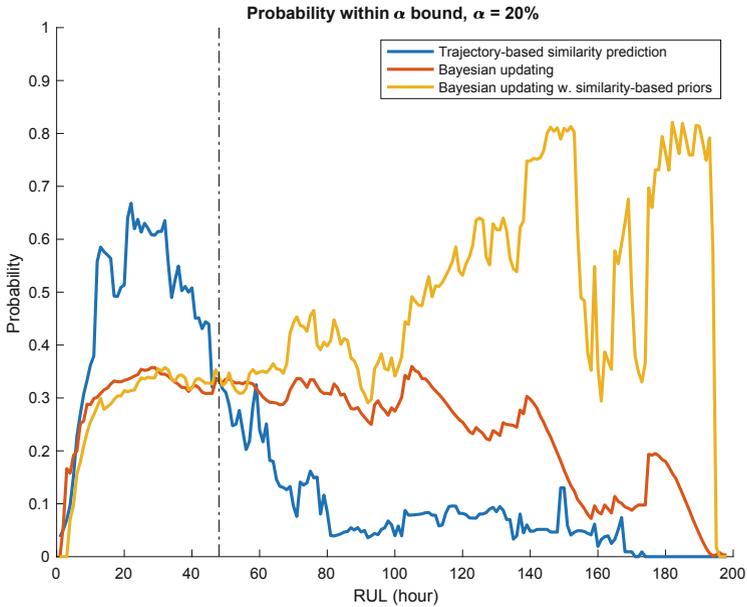
Fig. 5. Overall comparison of techniques.

5 Conclusions

In a case study, we show how techniques from literature can be combined and parameterized to accurately predict the Remaining Useful Life (RUL) of a machine or part. While curves of the degradation of a machine or part over time typically have a similar shape, the challenge is that operational constraints, which may be unknown, influence the exact parameterization of that curve, as evidenced by the real-life runs displayed in Figs. 1 and 2. Therefore, we propose a similarity-based prediction technique: while it makes no sense to compare the current run with all previously observed runs, it is quite likely that there are *some* historical runs that are similar to the current run, because they have similar operational constraints, hence providing us with powerful predictive information.

This paper proposes a new similarity-based prediction technique, in which we obtain a probability distribution of the RUL through Bayesian updating, where the priors of the Bayesian distribution are calculated based on a careful selection of previously seen runs. As evidenced by Fig. 5, our technique outperforms alternative techniques in a case study by a large margin within the long-term region. If we strive to predict the RUL shorter in advance, Fig. 5 clearly indicates that other methods work better.

While we studied the performance of RUL prediction techniques in the context of a particular case study, in many other domains degradation patterns have similar properties. In particular, in many other domains: run to failure data has a 'sawtooth' shape as in Fig. 1, degradation depends on operational conditions

that are unknown (e.g., because they are not measured), and long-term predictions are of interest (e.g., for planning maintenance activities). In such situations our technique can also be expected to work well.

References

1. Aizpurua, J.I., Catterson, V.M.: Towards a methodology for design of prognostic systems. In: Annual Conference of the Prognostics and Health Management Society, pp. 1–13, October (2015)
2. Arif-Uz-Zaman, K., Cholette, M.E., Ma, L., Karim, A.: Extracting failure time data from industrial maintenance records using text mining. Adv. Eng. Inform. **33**, 388–396 (2017)
3. Bleakie, A., Djurdjanovic, D.: Analytical approach to similarity-based prediction of manufacturing system performance. Comput. Ind. **64**(6), 625–633 (2013)
4. Coble, J.B.: Merging data sources to predict remaining useful life-an automated method to identify prognostic parameters. Ph.D. thesis, University of Tennessee (2010)
5. Eker, O.F., Camci, F., Jennions, I.K.: Physics-based prognostic modelling of filter clogging phenomena. Mech. Syst. Sig. Process. **75**, 395–412 (2016)
6. Fink, O., Zio, E., Weidmann, U.: Predicting component reliability and level of degradation with complex-valued neural networks. Reliab. Eng. Syst. Saf. **121**, 198–206 (2014)
7. Gebraeel, N.: Sensory-updated residual life distributions for components with exponential degradation patterns. IEEE Trans. Autom. Sci. Eng. **3**(4), 382–393 (2006)
8. Gebraeel, N., et al.: Residual life distributions from component degradation signals: a Bayesian approach residual-life distributions from component degradation signals: A Bayesian approach. IIE Trans. **37**(6), 543–557 (2005). Research Collection Lee Kong Chian School of Business
9. Goebel, K., Eklund, N.: Prognostic fusion for uncertainty reduction. Soft Comput. (2007)
10. Kleinbaum, D.G., Klein, M.: Survival Analysis, 3rd edn. Springer, New York (2012). https://doi.org/10.1007/978-1-4419-6646-9
11. Lam, J., Sankararaman, S., Stewart, B.: Enhanced trajectory based similarity prediction with uncertainty quantification. In: Proceedings of the PHM, pp. 623–634 (2013)
12. Ling, Y.: Uncertainty quantification in time-dependent reliability analysis. Ph.D. thesis, Vanderbilt University (2013)
13. Liu, J., Djurdjanovic, D., Ni, J., Casoetto, N., Lee, J.: Similarity based method for manufacturing process performance prediction and diagnosis. Comput. Ind. **58**(6), 558–566 (2007)
14. Pearson, R.K., Neuvo, Y., Astola, J., Gabbouj, M.: Generalized hampel filters. EURASIP J. Adv. Sig. Process. **2016**(1), 1–18 (2016). https://doi.org/10.1186/s13634-016-0383-6
15. Saxena, A., Celaya, J., Saha, B., Saha, S., Goebel, K.: Metrics for offline evaluation of prognostic performance. Int. J. Progn. Health Manag. **1**, 1–20 (2010)
16. Si, X.S., Wang, W., Hu, C.H., Chen, M.Y., Zhou, D.H.: A Wiener-process-based degradation model with a recursive filter algorithm for remaining useful life estimation. Mech. Syst. Sig. Process. **35**(1–2), 219–237 (2013)

17. Tobon-Mejia, D.A., Medjaher, K., Zerhouni, N., Tripot, G.: A data-driven failure prognostics method based on mixture of Gaussian hidden Markov models. IEEE Trans. Reliab. **61**(2), 491–503 (2012)
18. Wang, T.: Trajectory similarity based prediction for remaining useful life estimation. Ph.D. thesis, University of Cincinnati (2010)
19. Wang, T., Yu, J., Siegel, D., Lee, J.: A similarity-based prognostics approach for remaining useful life estimation of engineered systems. In: 2008 International Conference on Prognostics and Health Management, pp. 1–6. IEEE, October 2008
20. Yildirim, M., Sun, X.A., Gebraeel, N.Z.: Sensor-driven condition-based generator maintenance scheduling - part I: maintenance problem. IEEE Trans. Power Syst. **31**(6), 4253–4262 (2016)
21. You, M.Y.: A predictive maintenance system for hybrid degradation processes. Int. J. Qual. Reliab. Manag. **34**(7), 1123–1135 (2017)