Warranty claim forecasting based on weighted maximum likelihood estimation

Artur Akbarov, Shaomin Wu*

School of Applied Sciences, Cranfield University, Bedfordshire MK43 0AL, UK

Abstract

Warranty claims reported in recent months might carry more up-to-date information than those reported in earlier months. Using weighted maximum likelihood estimation for estimating model parameters might therefore lead to better performance of warranty forecasting models than maximum likelihood estimation. This paper examines this issue and also presents comparison of the forecasting performance of the parametric models such as Poisson processes and ARIMA models and non-parametric models such as artificial neural networks. It shows that mixed non-homogenous Poisson process models can lead to better forecasting results than other competing methods. The paper also shows that the models built with the weighted maximum likelihood estimation yield smaller error than those based on the maximum likelihood estimation.

Keywords: Warranty forecasting, Poisson processes, weighted maximum likelihood method, overdispersion.

1 Introduction

Warranty claim forecasting is becoming increasingly important for businesses as the financial resources associated with warranty coverage are running into millions of pounds. Warranty has become a marketing tool that is utilised to assure the customer of a superior reliability of the product and the manufacturer's commitment to post sale product service. As a result, many manufacturers offer longer warranty, which leads to larger warranty reserves

*Corresponding author. *E-mail: s.m.wu@kent.ac.uk*

and requires a better resource management through a thorough analysis of warranty data. This study presents some results on improving the accuracy of forecasting warranty claims.

Approaches to warranty data analysis have been addressed by many researchers. For example, Kalbfleisch et al.[1] uses Poisson models to analyse automobile warranty data with reporting delays, Murthy[2] discusses warranty cost analysis based on usage rate. Lawless and Kalbfleisch[3], Kalbfleisch and Lawless[4], Lawless[5], and Suzuki et al.[6] discuss different statistical aspects of warranty data analysis, Fredette and Lawless[7] discusses warranty claims forecasting, Akbarov and Wu[8] uses Poisson models to analyse warranty claim data of electronics products with sales delay. Karim and Suzkui[9] and Wu[10] offers reviews of warranty literature, respectively. Many of these studies use non-homogeneous Poisson processes to model warranty data. Also, see Blischke et al.[11] for more recent discussion of the issues related to the analysis of the warranty data.

Wasserman[12] presents application of dynamic linear models to predicting warranty claims. The author compares the Kalman filter method to the simple linear regression method. Wasserman and Sudjianto[13] presents the results of comparing the forecasting performance of three different modelling strategies. The methods are compared based o nthe analysis of automobile warranty data. The authors consider static predictive models such as ARIMA, Kalman filter and artificial neural networks. This study has shown that the neural networks have resulted in the least forecasting error.

Some studies on the analysis of the warranty data have reported the phenomenon of overdispersion. The overdispersion occurs when the variance of the Poisson random variable is higher than its expectation, where the two should be equal. The phenomenon is thought to occur due to intrinsic discrepancies in the reliability of individual products, heterogeneity of users and operating environments. For more details see Kalbfleisch et al.[1], Kalbfleisch and Lawless[4], Fredette and Lawless[7], and Akbarov and Wu[24].

In this study we consider the application of the following models. Auto-regressive integrated moving average (ARIMA) models as these models are standard models for forecasting time series. Non-homogenous Poisson process (NHPP) models as these models are commonly applied to warranty data. Mixed non-homogenous Poisson process (MNHPP) models as these models are suitable for dealing with overdispersion. Artificial neural networks (ANN) as these models have recently become a popular method for various purposes including the time series forecasting. To our knowledge none of the previous studies on warranty data analysis has looked at comparing Poisson processes against the ARIMA and neural networks models.

Wu and Akbarov[19] have shown that giving higher weights to recent data can lead to better forecasting results and use non-parametric approaches to building forecasting models. In this study we consider a similar approach for estimating the parameters of the Poisson processes. More specifically, we consider the weighted maximum likelihood method. To our knowledge this has not been considered for forecasting warranty claim data so far. We show that weighted maximum likelihood method applied to the mixed non-homogenous Poisson process gives better forecasts than other methods considered here and also better forecasts than when it is fitted using the maximum likelihood method.

2 Parametric models

2.1 Poisson process models

This subsection discusses discrete time Poisson processes parametrised by μ_t. μ_t is the expectation of the increment of the process at time t and depends on the type of the process. As we consider discrete time processes, we define r_t to be the expected number of warranty claims per product unit at time t. r_t can be derived from a continuous function as $r_t = \int_t^{t+\Delta} h(x)dx$, where $h(x)$ can be a continuous function such the hazard rate function. Since we consider monthly data we let $\Delta = 1$. We also denote by N_t the number of products in the market at time t and d_t the number of observed warranty claims in month t.

The *non-homogeneous Poisson process* is one of the most common probability models used to model failure counts of repairable products, see Ascher and Feingold[20]. The NHPP assumes that upon repair the failure rate of a product is restored to the same level as it was just before the failure. Such a repair is referred to as a minimal repair. Many products that are subject to warranty can be thought of as repairable systems. Even in the cases where some part of the system is replaced by a new component, as a whole, the system can often still be viewed as repairable. The NHPP has been applied to model warranty data in studies such as Kalbfleisch et al.[1], Lawless[5], Karim et al.[21], Wang et al.[22], and Majeske[23].

The mean of the NHPP is a deterministic function of time. In our case, the intensity function of the NHPP is given by $\mu_t = N_t r_t$. The increments of the NHPP are independent from each other and distributed according to a Poisson distribution with mean μ_t.

The *mixed non-homogeneous Poisson process* assumes that the intensity function of the process is subject to random changes from its expected value. For mathematical simplicity, such random changes are often modelled using a gamma distribution. Let α be a gamma random variable with $\mathrm{E}(\alpha) = a/b$ and denote $M_t = \sum_{i=1}^{t} \mu_t$. Then the number of events in interval $(0, t]$ is a random variable given by αM_t. The increments of the mixed Poisson processes are not independent, therefore μ_t is given by $\mu_t = (\alpha|D_{t-1})N_t r_t$, where D_{t-1} is the data observed prior to t, for more details, see Fredette and Lawless[7] and Akbarov and Wu[24]. For simplicity, we consider an MNHPP model, where $\mathrm{E}(\alpha) = 1$. The application of MNHPP to automobile warranty data is considered in Fredette and Lawless[7].

2.2 Weighted maximum likelihood estimation

There are several methods that use weighted approach to finding maximum likelihood estimates, namely, local likelihood, relevance weighted likelihood and weighted likelihood. These methods define the log-likelihood function in terms of a weights function $w(z)$ as:

$$\mathrm{lnL}(z; x_1, x_2, ..., x_n) = \sum_{i=1}^{n} w(z)\mathrm{lnL}(x_i, \Theta), \tag{1}$$

where x_i is the i^{th} observation and n is the total number of observations.

The general form of the local likelihood uses a kernel function to concentrate the weights around some value z with bandwidth h, $w(z) = K(\frac{x_i - z}{h})$, for more details refer to Eguchi and Copas[25].

The relevance weighted likelihood function is the weighted function of likelihoods of different data sets assumed to have been generated from distributions similar (at least in some qualitative sense) to the distribution whose parameters are being estimated. $w(z)$, in this case, reflects the degree of similarity, for an example of the use of this approach see Hu and Zidek[26].

The term weighted likelihood function has been used in many different contexts, which are discussed in detail in Wang[17]. Often weighted likelihood methods are used for combining the likelihoods of data from different but somehow similar samples. Asymptotic properties of such estimates can be found in Wang et al.[27].

In this paper, we focus on weighing the likelihoods of the data samples based on their temporal distance from the current point in time. A similar approach has been applied by Wu and Akbarov[19] in the context of warranty claims forecasting using machine learning techniques such as support vector regression and neural networks. They have shown that the weighted approach can yield more accurate forecasts.

In the context of warranty claims forecasting, the necessity for weighing the observed data based on their temporal distances from the current point in time can arise due to the following reasons. It is often happens that product lines undergo some design or other subtle changes to remove problems identified from the failure of earlier production batches. This makes the earlier data less representative of the current failure tendencies. This is one of the most common reasons why more recent data plays a more important role for forecasting future warranty claims. Also, fast paced technological advances can lead to quick obsolescence of products which would have an impact on the propensity of customers to claim warranty due to a more preferable option of purchasing a new, more technologically advanced product.

In this study we consider the weights function, w_t, for discrete time series, given by a normalised cumulative distribution function of the geometric distribution:

$$w_t = \frac{1}{c}(1 - (1-\theta)^t), \tag{2}$$

where the normalising constant $c = \sum_{t=1}^{T}(1-(1-\theta)^t)$ with T being the most recent month, and parameter $p > 0$, which controls the spread of weights. For example $\theta = 1$ leads to equal weights for all t. In general, larger values of θ lead to relatively more equal wights for longer time lags, whereas smaller values of θ lead to more variable weights, see Figure 1. The choice of the above function implies that the weights are in a strictly increasing order from month 1 to month T and $\sum_{t=1}^{T} w_t = 1$. It also insures that there is no over-reliance on the most recent data, where the most recent values have high weights and the rest have very small weights. In this study we choose the parameter θ that maximises the log-likelihood function.

The weighted log-likelihood function of the non-homogenous Poisson process is given by:

$$\text{lnL} = \sum_{t=1}^{T} w_t \Big\{ d_t \text{ln}(N_t r_t) - N_t r_t + \text{ln}(d_t!) \Big\}, \tag{3}$$

where $N_t r_t$ is the expected number of claims in month t and d_t is the number of observed claims in month t.

The weighted log-likelihood function of the mixed non-homogenous Poisson process is given by:

$$\begin{aligned} \text{lnL} = \sum_{t=1}^{T} w_t \Big\{ & \text{ln}(\Gamma(a + A_t + d_t)) + (a + A_t)\text{ln}(b + B_t) + d_t \text{ln}(N_t r_t) \\ & - \text{ln}(\Gamma(a + A)) - (a + A_t + d_t)\text{ln}(b + B_t + N_t r_t) - \text{ln}(d_t!) \Big\}, \end{aligned} \tag{4}$$

where $A_t = \sum_{i=1}^{t-1} d_i$ and $B_t = \sum_{i=1}^{t-1} N_t r_t$, see Akbarov and Wu[24].

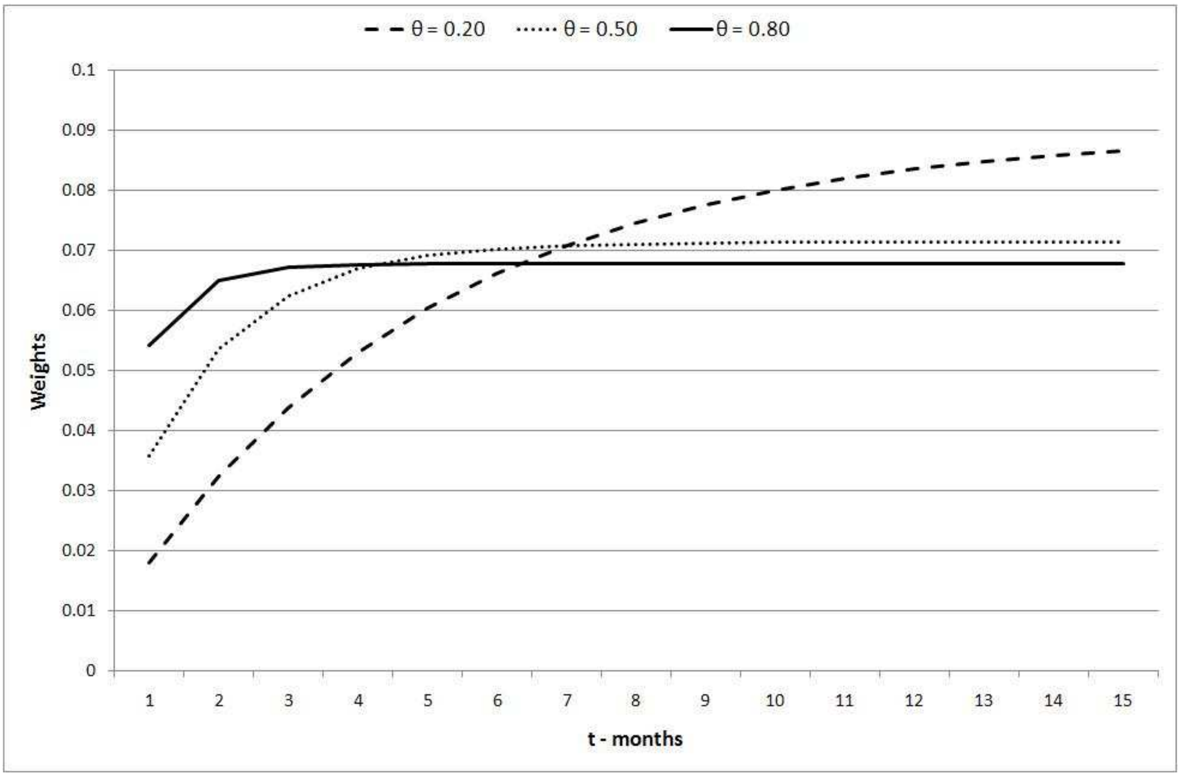

Figure 1: Weights function, w_t, given in continuous form to emphasise the shape of the function.

2.3 ARIMA models

The auto-regressive integrated moving average model (ARIMA) is a widely used time series model (see Montgomery[28] and Chatfield[29], for example). The ARMA (p, q) model for time series X_t with $t = 1, 2, ...$ is given by

$$X_t = \sum_{i=1}^{p} \beta_i X_{t-i} + \sum_{j=0}^{q} \gamma_j \epsilon_{t-j} \quad (5)$$

where the first term is the auto-regressive model (AR) of order p and the second term is the moving average model (MA) of order q and $\gamma_0 = 1$. When the d^{th} difference of a time series follows an ARMA(p, q), the model becomes an ARIMA(p, q, d). The process can be given a mean by adding some constant c to the above equation.

In this paper, the time series under consideration is d_t. The order terms of the models, p, q and d are determined on the basis of the Akaike information criterion (AIC).

3 Neural networks

Neural networks have become a popular method applied to a wide range of problems such as function fitting, classification, regression and time series forecasting, the reader is referred to Bishop[30] for detailed discussion of neural networks. The credibility of the neural networks has been established by the universal approximation property (Hornik et al.[31] and Funuhashi[32]).

Here, we consider a neural network model, or called *multilayer perceptron* (MLP), which is a feedforward artificial neural network model that maps sets of input data onto a set of appropriate output. The time series forecasting problem can be formulated as follows:

$$\hat{x}_{t+i} = f(x_{t-K+i}, x_{t-K+i-1}, ..., x_{t-K+i-p}), \quad (6)$$

where $\hat{x}_{t+i}$ is the forecast for time $t+i$, for $i = 1, 2, ..., K$, and p is the order of auto-regression.

The number, H, of hidden nodes, controls the complexity of the network. A complex model with a large number of nodes can lead to over-fitting, which means the model performs exceptionally well on the training set but has very poor generalisation on the test data.

Here, we consider two forecasting horizons, $K = 3$, and $K = 6$. The available data consists of 24 months of observations for eight products. We use the first 18 months of the data to fit the models and the remaining 3 and 6 months to test the forecasting accuracy of the models. The first 18 months of the data are divided into two parts: the first 15 months are used to train the neural network and the last 3 months, from 16 to 18 are used as the validation set. The model that results in the lowest error on the validation set is then used for forecasting purposes.

The auto-regression order p and the number H of neurons in the hidden layer are chosen in a way that gives the least error on the validation set. H is sought in the range of 2 to 7 neurons, and the range of p is chosen accordingly depending on K. Since training neural networks can results in local optima, the network training is performed 30 times for each combination of H and p.

4 Case study

This section presents the results of data experiments using a data set from electronics industry consisting of eight different products. We use the first 18 months of the data to fit the models and then measure the forecasting accuracy based on forecasts of 3 and 6 months ahead.

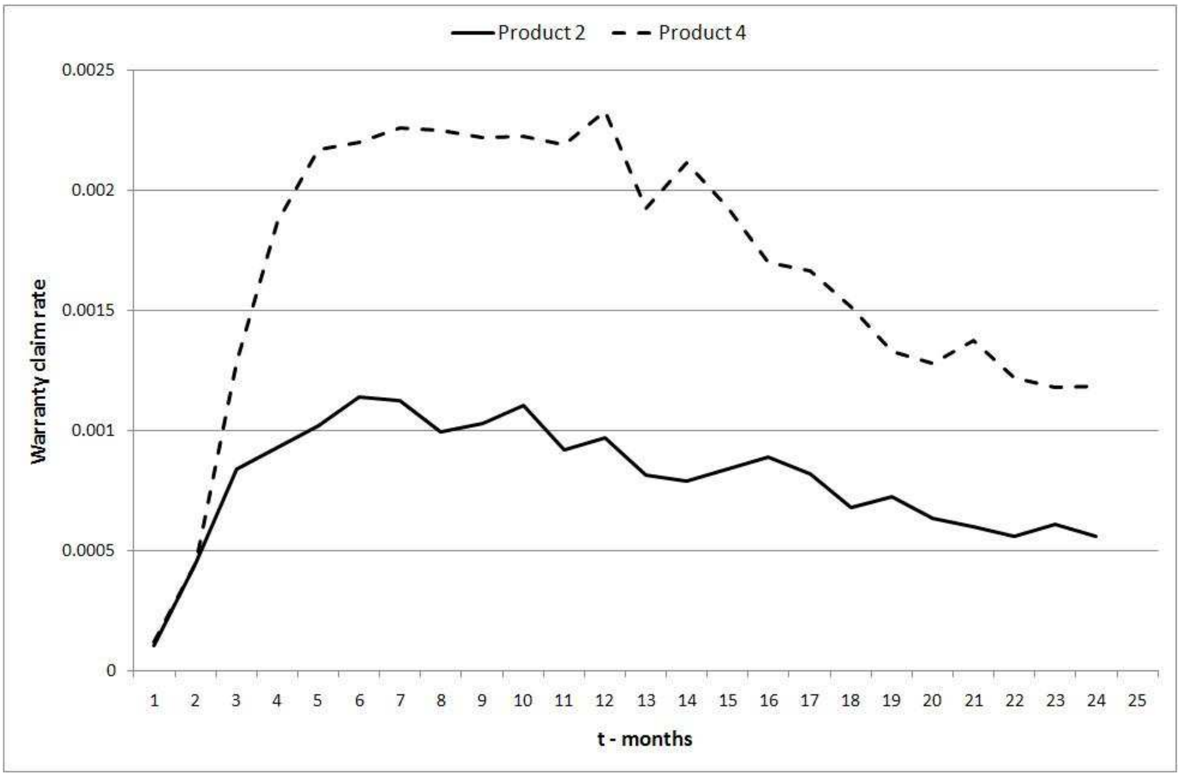

Figure 2: Claim rates for Products 2 and 4.

The data set consists of two main pieces of information, the number of products in the market in month t, N_t, and the number of warranty claims in month t, d_t. The observed claim rate r_t estimated as d_t/N_t for two different products is shown in Figure 2. The figure shows that the claim rates have a single mode, where after initial increase the claim rate start decreasing with time. Other products considered in this study also exhibit a similar behaviour, which can be modelled using the hazard rate function of the inverse-Weibull distribution given by:

$$h(t) = \alpha\beta^{\alpha}t^{-(\alpha+1)}e^{-(\frac{\beta}{t})^{\alpha}}(1 - e^{-(\frac{\beta}{t})^{\alpha}})^{-1}. \quad (7)$$

Thus, the expected number of warranty claims per product unit in month t is given by $r_t = \int_{t-1}^{t} h(x)dx$ for $t = 1, 2, ...$

The forecasting error is measured using the normalised rooted mean squared error (NRMSE) given by:

$$NRMSE = \sqrt{\frac{\sum_{t=T+1}^{T+K}(d_t - \hat{d}_t)^2}{\sum_{t=T+1}^{T+K} d_t^2}}, \quad (8)$$

where $\hat{d}_t$ is the forecasted value of d_t, T is the number of months used for fitting the models, and K is the forecasting horizon.

Table 1 and Table 2 show the results of measuring the forecasting accuracy for all eight products. We can note from these two tables that the MNHPP and the weighted MNHPP have the lowest NRMSE for both forecasting periods. Also, we can note that the weighted MNHPP and the weighted NHPP have better forecasting accuracy than the MNHPP and NHPP respectively.

Table 3 shows the results of the paired two-sample test for means between different models across all eight products for $K = 3$, and Table 4 shows similar results for $K = 6$. The tables should read as comparing the rows to columns, so the positive sign of the t statistic indicates that the mean of the method given in the row is higher than the mean of the method given in the column. The values in the parenthesis represent the associated significance values.

From both tables, 3 and 4, we can note that the difference between weighted MNHPP, NHPP and their counterparts is statistically significant. So, we can conclude that the use of weighted maximum likelihood method can give better forecasting results.

The wMNHPP column in Table 3 shows that the average of the forecasting error is statistically significantly lower than for other methods except for the ARIMA method. Nevertheless, even compared to the ARIMA method the average of the wMNHPP method is 28% lower.

The wMNHPP column in Table 4 shows similar results. In this case, the average of the wMNHPP is statistically significantly lower than for all other methods except for the ANN. Even for the case of ANN, the difference is just outside the 0.05 significance level. However, the wMNHPP average is 51% lower than the average of ANN.

Table 1: Normalised rooted mean squared error for forecasting horizons $K = 3$ and $K = 6$.

Product	**NHPP**		**ANN**		**ARIMA**		**MNHPP**	
	K = 3	**K = 6**	**K = 3**	**K = 6**	**K = 3**	**K = 6**	**K = 3**	**K = 6**
1	0.026	0.065	0.245	0.270	0.226	0.319	0.025	0.065
2	0.293	0.341	0.370	0.417	0.229	0.321	0.134	0.156
3	0.365	0.400	0.241	0.302	0.332	0.490	0.242	0.254
4	0.546	0.609	0.432	0.397	0.089	0.149	0.247	0.260
5	0.186	0.175	0.124	0.130	0.035	0.045	0.067	0.058
6	0.178	0.340	0.230	0.430	0.180	0.290	0.239	0.230
7	0.483	0.552	0.255	0.338	0.210	0.471	0.089	0.105
8	0.254	0.619	0.154	0.442	0.185	0.523	0.136	0.379
Average	**0.291**	**0.388**	**0.256**	**0.341**	**0.186**	**0.326**	**0.147**	**0.188**

Table 2: Normalised rooted mean squared error for forecasting horizons $K = 3$ and $K = 6$ based on weighted maximum likelihood estimation. Prefix "w" stands for weighted.

Product	wNHPP		wMNHPP	
	K = 3	**K = 6**	**K = 3**	**K = 6**
1	0.024	0.059	0.024	0.059
2	0.249	0.293	0.120	0.140
3	0.309	0.338	0.197	0.201
4	0.429	0.481	0.216	0.225
5	0.130	0.119	0.035	0.043
6	0.151	0.278	0.258	0.243
7	0.391	0.450	0.070	0.085
8	0.204	0.547	0.146	0.336
Average	**0.236**	**0.321**	**0.133**	**0.166**

Table 3: Paired two-sample test for means: estimated t-statistic (associated significance level), $K = 3$. Prefix "w" stands for weighted.

	NHPP	ANN	ARIMA	MPP	wNHPP	wMPP
NHPP	-	1.30 (0.12)	1.53 (0.08)	2.77 (0.01)	4.40 (0.00)	2.80 (0.01)
ANN		-	1.15 (0.14)	3.5 (0.00)	- 0.22 (0.41)	3.37 (0.01)
ARIMA			-	0.94 (0.19)	- 0.88 (0.2)	1.29 (0.12)
MPP				-	- 2.14 (0.04)	1.85 (0.05)
wNHPP					-	2.25 (0.03)

Table 4: Paired two-sample test for means: estimated t-statistic (associated significance level), $K = 6$. Prefix "w" stands for weighted.

	NHPP	ANN	ARIMA	MPP	wNHPP	wMPP
NHPP	-	2.78 (0.01)	0.86 (0.21)	3.95 (0.00)	5.22 (0.00)	4.13 (0.00)
ANN		-	-2.08 (0.04)	0.97 (0.18)	-2.07 (0.04)	1.75 (0.06)
ARIMA			-	2.52 (0.02)	0.09 (0.47)	2.88 (0.01)
MPP				-	-3.34 (0.01)	2.96 (0.01)
wNHPP					-	3.61 (0.00)

5 Discussion

The results of this study show that MNHPP models have the best forecasting performance. For short forecasting horizon $K = 3$ the next best method is the ARIMA models. This is to be expected as ARIMA models often perform well for short forecasting horizons. The NHPP model has not given good forecasting results, this can be explained by potential presence of overdispersion in the data, where the NHPP model becomes rather inadequate. The neural networks have not resulted in better forecasts than the ARIMA or MNHPP models. In general, the neural networks are often do not to generalise well on unseen data. Also, neural networks can have many local optima which can make it difficult to achieve consistent results on different runs. Here, we have chosen the neural network model that performs best on the validation set.

The MNHPP models are dependent on the choice of the form of the intensity function. In our examples, we have chosen the inverse-Weibull distribution, because the observed rate of the products can be adequately modelled by this function. This is often useful, as companies produce many similar products and we can choose the intensity function form depending on the observed behaviour of warranty rates for old products. More flexible forms of the intensity function can be contemplated. The reliance on the similarity between the new products and older products can lead to good extrapolations. This is especially true for electronics industry where many products share similar components.

The results of this study emphasise the importance of giving higher weights for more recent data samples when forecasting warranty claims. We have discussed that the reasons for this are often justified by external factors that can have an impact on the warranty claim arrival process.

The field warranty data often exhibits overdispersion, where the variance of the increments of the process is higher than its expectation. From the above results we have seen that models that can deal with overdispersion fit the data better than the non-homogenous Poisson process, where the variance and the expectation of the increments are equal. In practise, the overdispersion can be expected to have a dynamic nature. Mixed non-homogenous Poisson process updates the level of overdispersion as more data becomes available.

6 Conclusions

In this study we have focused on forecasting warranty claims using the following methods, ARIMA, NHPP, MNHPP and neural networks. It is clear from the results that the MNHPP model has the best forecasting accuracy, and that estimating the parameters of the Poisson processes with the weighted maximum likelihood method gives better forecasting results than those with the maximum likelihood method. Based on those warranty claim data we have collected, we can draw the following conclusions from this study:

- Weighted maximum likelihood methods using weights depending on the temporal distance of data samples from the current point in time can yield more accurate forecasts than those obtained by maximum likelihood method for Poisson processes. Although, this can be conditional on the choice of the weights function and its parameters, a suitable choice can have a significant positive impact.

- Mixed non-homogenous Poisson process can often yield better forecasting results than the non-homogenous Poisson process. This is mainly due to the fact that many field warranty data exhibit overdispersion.

We recommend that, when fitting forecasting models to field warranty data, one should consider giving more weights to recent data samples and take into account the phenomenon of overdispersion, which is often present in real life data sets.

Acknowledgement

We are grateful to the anonymous referees for their valuable comments and suggestions that improved the quality of this paper. This research is supported by Engineering and Physical Sciences Research Council (EPSRC) of the United Kingdom (EPSRC Grant reference: EP/G039674/1).

References

1. Kalbfleisch JD, Lawless JF, Robinson JA. Methods for the analysis and prediction of warranty claims. *Technometrics* 1991, **33**: 273–285.

2. Murthy DNP. A usage dependent model for warranty costing. *European Journal of Operational Research* 1992, **57**: 89–99.

3. Lawless JF, Kalbfleisch JD. Some issues in the collection and analysis of field reliability data. In *Survival Analysis: State of the Art*, Klein JP., Goel PK (Eds). Kluwer Academic: Dordrecht, 1992, 141–152.

4. Kalbfleisch JD, Lawless JF. Statistical analysis of warranty claims data. In *Product Warranty Handbook*, Blishke WR, Murthy DNP (Eds). M. Dekker: New York, 1996, 231–259.

5. Lawless JF. Statistical analysis of product warranty data. *International Statistical Review / Revue Internationale de Statistique* 1998, **66**(1): 41–60.

6. Suzuki K, Karim R, Wang L. Statistical analysis of reliability warranty data. In *Handbook of Statistics, vol 20*, Balakrishnan N, Rao CR (eds). Elsvier Science, 2001, 585–609.

7. Fredette M, Lawless JF. Finite-Horizon Prediction of recurrent events, with Application to forecasts of warranty claims. *Technometrics* 2007, **49**(1): 66–80.

8. Akbarov A, Wu S. Warranty claims data analysis considering sales delay. *Quality and Reliability Engineering International* (DOI: 10.1002/qre.1302).

9. Karim MR, Suzkui K. Analysis of warranty claim data: a literature review. *International Journal of Quality and Reliability Management* 2005, **22**(7): 667–686.

10. Wu S. Warranty data analysis: a review. *Quality and Reliability Engineering International* (DOI: 10.1002/qre.1282).

11. Blischke WR, Karim MR, and Murthy DNP. *Warranty data collection and analysis* 2011, Springer.

12. Wasserman G. An application of dynamic linear models for predicting warranty claims. *Computers and Industrial Engineering* 1992, **22**(1): 37-47.

13. Wasserman G, Sudjianto A. A comparison of three strategies for forecasting warranty claims. *IIE Transactions* 1996, **28**:967-977.

14. Lefebvre M. *Applied Stochastic Processes*. Springer: New York, 2007.

15. Seal H. *Survival probabilities: the goal of risk theory*. Wiley: Chichester, 1978.

16. Lu Y, Garrido J. Regime-switching periodic models for claim counts. *North American Actuarial Journal* 2006, **10**(4): 235–248.

17. Wang SX. *Maximum weighted likelihood estimation*. Ph.D. thesis, The University of British Columbia, 2001.

18. Wu S. Warranty claim analysis considering human factors. *Reliability Engineering and System Safety* 2011, **96**(1), 131-138

19. Wu S, Akbarov A. Support vector regression for warranty claim forecasting. *European Journal of Operational Research* 2011, **213**: 196–204.

20. Ascher H, Feingold H. *Repairable systems reliability: Modelling, inference, misconceptions and their causes*. Marcel Dekker: New York, 1984.

21. Karim R, Yamamoto W, Suzuki K. Statistical analysis of marginal count failure data. *Lifetime Data Analysis* 2001, **7**: 173–186.

22. Wang L, Suzuki K, Yamamoto W. Age-based warranty data analysis without date-specific sales information. *Applied Stochastic Models in Business and Industry* 2002, **18**(3): 323–337.

23. Majeske K. A non-homogeneous Poisson process predictive model for automobile warranty claims. *Reliability Engineering and System Safety* 2007, **92**(2): 243–251.

24. Akbarov A, Wu S. Forecasting warranty claims considering dynamic over-dispersion. *International Journal of Production Economics* 2011, (Submitted).

25. Eguchi S, Copas J. A class of local likelihood methods and near-parametric asymptotics. *Journal of Royal Statistical Society: Series B* 1998, **60**: 709–724.

26. Hu F, Zidek JV. The relevance weighted likelihood with applications. In *Empirical Bayes and Likelihood Inference*, Ahmed SE, Reid N (eds). Springer: New York, 1997, 211–235.

27. Wang X, van Eeden C, Zidek JV. Asymptotic properties of maximum weighted likelihood estimators. *Journal of statistical planning and inference* 2004, **119**: 37–54.

28. Montgomery DC, Johnson LA, and Gardiner JS. *Forecasting and Time Series Analysis* 1990, McGraw-Hill: New York.

29. Chatfield C. *The Analysis of Time Series* 1996, Chapman & Hall: London.

30. Bishop CM. *Neural Networks for Pattern Recognition* 1995, Oxford University Press: Oxford.

31. Hornik K, Stinchombe M, White H. Multilayer feedforward networks are universal approximators. *Neural Networks* 1989, **2**:359-366.

32. Funahashi, K. On the approximate realisation of continuous mappings by neural networks. *Neural Networks* 1989, **2**:183-192.