



**HAL**  
open science

## Variance Gamma process for predictive maintenance of mechanical systems

Marwa Belhaj Salem, Mitra Fouladirad, Estelle Deloux

► **To cite this version:**

Marwa Belhaj Salem, Mitra Fouladirad, Estelle Deloux. Variance Gamma process for predictive maintenance of mechanical systems. Congrès Lambda Mu 22 “ Les risques au cœur des transitions ” - 22e Congrès de Maîtrise des Risques et de Sûreté de Fonctionnement, Institut pour la Maîtrise des Risques, Oct 2020, Le Havre (e-congrès), France. hal-03295749v2

**HAL Id: hal-03295749**

**<https://hal.science/hal-03295749v2>**

Submitted on 22 Jul 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Variance Gamma process for predictive maintenance of mechanical systems

## Processus Variance Gamma pour une maintenance prédictive des systèmes mécaniques

Marwa Belhaj Salem  
ICD-M2S

University of Technology of Troyes  
Troyes, France  
marwa.belhaj\_salem@utt.fr

Mitra Fouladirad  
ICD-M2S

University of Technology of Troyes  
Troyes, France  
mitra.fouladirad@utt.fr

Estelle Deloux  
ICD-M2S

University of Technology of Troyes  
Troyes, France  
estelle.deloux@utt.fr

**Abstract**—Remaining Useful Life (RUL) prediction is important for the prognosis and the maintenance of expensive systems. High reliability is an indispensable requirement for advanced systems. Correspondingly, reliability practitioners attempt to investigate the system behaviour in order to mitigate risk as much as possible. The study of the behaviour is insured by collecting the real data of degradation, which is considered as a challenged task. Based on the retrieved data, one can propose a good stochastic process to model the system degradation. In this study, the Variance Gamma process is proposed to model the degradation of the water tank pump. The estimation of the data parameters and their fitting to the model is considered. In addition, the study of the First hitting time (FT) and its distribution is treated. Finally, a good prognostic is conducted based on the FT results.

**Keywords**— Variance Gamma, Stochastic process, Prognostic, Deteriorating system

**Résumé**—La prédiction de la durée de vie résiduelle (RUL) est importante pour le pronostic et l'entretien des systèmes coûteux. Une fiabilité élevée est une condition indispensable pour les systèmes avancés. De même, les praticiens de la fiabilité tentent d'étudier le comportement du système afin d'atténuer le plus possible les risques. L'étude du comportement est assurée par la collecte des données réelles de dégradation, qui est considérée comme une tâche contestée. Sur la base des données récupérées, on peut proposer un bon processus stochastique pour modéliser la dégradation du système. Le but de cette étude est de proposer le processus Variance Gamma comme un modèle de dégradation. La pompe d'un réservoir d'eau est utilisée comme un cas d'application. L'estimation des paramètres de données et leur ajustement au modèle sont considérés. De plus, l'étude du premier temps de passage (FT) et de sa distribution est traitée. Enfin, en se basant sur les résultats du FT un bon pronostic est réalisé.

**Mots clés**— Variance Gamma, processus stochastique, pronostique, système de dégradation

### I. INTRODUCTION

Prediction of Remaining Useful Life (RUL) and maintenance of the system at the appropriate time based on their performance is important in the huge industries and laboratories in the past years [1]. This methodology will assist in predicting the defect in each part of the system based on their degradation by using the historical data retrieved by sensors. This will permit low cost system maintenance by restricting the failure of the whole system. These methodologies are widely used in complex systems such as motors, pumps, generators, etc. [2, 3].

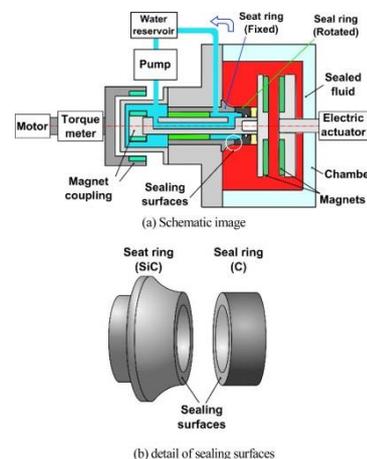


Fig 1. Complex system of centrifugal pump and seal surfaces [4]

This study is mainly centred on the degradation of the industrial centrifugal water pumps (Fig 1). They are widely used in different industrial applications and a small defect in this pump or its part may cause a huge economic loss and catastrophic damages of the entire system. Damage of pump seals is one of the major causes of the defect of the pumps and it is crucial to detect this defect or predict it at the appropriate time. The degradation can be detected by monitoring the level of the leakage in the water tank pump. Due to the development of the industries, more and more complex systems are being

used and it becomes more challenging to collect failure data. Tracking of this data and gathering the degradation data is becoming more expensive due to the systems complexity. As a consequence of this, it is essential to propose a robust mathematical model to model the system and to be able to predict the failure.

Different stochastic processes are being widely used to predict the RUL in the past years. The selection of the stochastic process is mainly identified based on the behaviour of the retrieved data of degradation. Gamma, Wiener and Black & Scholes are the most popular stochastic models used to imitate degradation of systems. Gamma process is a stochastic process, which has an independent and non-negative increment with gamma distribution [5]. Gamma process is convenient to model the monotonic gradual damage, which increases with respect to time such as corrosion, fatigue, crack growth, wear, etc. Additionally, its mathematical calculations are relatively tractable, which is one of the advantages of this process. Even though, the model is restricted to the modelling of only monotonic behaviour, which is one of the major drawbacks as all the systems does not exhibit this behaviour [6, 7]. In order to overcome these limitations, researchers recommended another process called Wiener process. This process is able to describe the dynamic characteristic of a complex system deterioration due to its non-monotonic, infinite divisibility properties and its physical interpretations. It is used extensively in the modelling of degradation in laser generators, bridge beams, light-emitting diodes (LEDs), etc. It has simple computational properties and mathematical calculations, which are great advantages of this model [8, 9]. Nevertheless, as the systems are getting more complex and the industries need better precise models, the requirement of more complex stochastic processes arises.

A non-linear and a non-monotonic process called Black & Scholes (B&S) was considered as deterioration model. The analytical properties of the continuous time Brownian motion and B&S permit the exact computation of related probabilities, which is one of the main advantage of these models. As any other process, B&S has some limitations that restrict them in modelling of different degradation behaviour. Due to this, another stochastic process called Variance Gamma that can better replicate the non-monotonic degradation path is presented and proposed. The first appearance of Variance Gamma process (VG) as an extension of the Brownian motion was in finance and its first presentation as a simplified symmetric model was introduced by Madan and Seneta [12, 13, 14]. The advantage of VG is that it contains two more parameters compared to the Brownian motion, which help in the control of the kurtosis and the skewness. The Variance Gamma process can be obtained by evaluating the time in the Brownian motion by a gamma process. In addition, VG process can be written as a difference of two gamma processes. These two expressions of VG made it a flexible process capable of modelling any kind of non-monotonic degradation phenomenon as far as the increments follow the VG distribution. Since this study assesses the prediction of degradation behaviour of industrial centrifugal water tank pumps, which expresses a non-monotonic behaviour and a heavy tailed with high volatility data. The VG process can be proposed as a model to replicate the behaviour of the industrial centrifugal water pump's degradation and to predict the maintenance of the system.

The aim of this study is to introduce Variance Gamma (VG) process as a stochastic model that can model degradation. The water tank pump was introduced as an application case to study this process in detail. The leakage of water in the water tank is considered here as a reference of degradation. Firstly, the VG model and its properties will be presented. The estimation of the VG parameters will be carried out with the real data retrieved from the sensors. The data will be fitted to VG process and the First Hitting Time Density (FTD) will be identified. Lastly, the prognosis will be performed based on the obtained First Hitting Time (FT) results.

## II. PROCESS AND PROPERTIES

### A. Variance Gamma process

The Variance Gamma process is a Lévy process, which can be written as a time changed Brownian motion (TCBM). It can be presented as an evaluation of the time in a Brownian motion by a stochastic gamma process. Dilip B. Madan and Eugene Seneta [12] initially introduced this process in Finance analysis. It was proposed as an extension of the Brownian motion and its simplified symmetric form were presented [13, 14]. Generally, VG model has two extra parameters, which permit to control the skewness and kurtosis.

### B. Properties

The Variance Gamma process is one of the most flexible stochastic processes. It has two more parameters which help in controlling the skewness and kurtosis which allows for a better fitting the data. VG process can be introduced as an extension of a Brownian motion, it can be obtained by evaluating the wiener process at random times defined by a gamma process. The Variance Gamma process can also be written as a difference of two gamma processes which allow it to model two competitive phenomena. The first presentation of the process as a time changed the Brownian motion start by defining a Brownian motion  $B$  with positive parameters. The drift  $\theta$  and volatility  $\sigma$  and a standard Brownian motion  $W(t)$ . It is defined as an equation (1):

$$B(t; \theta, \sigma) = \theta t + \sigma W(t) \quad (1)$$

The gamma process of independent gamma distributed increments on a time interval  $(t, t+h)$ ,  $\Gamma(t; \mu, \nu)$  with  $\mu$  as the mean rate and  $\nu$  as the variance rate was considered. The VG process is defined as:

$$X(t; \sigma, \nu, \theta) = B(\Gamma(t; \mu, \nu), \theta, \sigma) \quad (2)$$

The other representation of the VG process as a difference of two gamma processes [14] is given by the equation as follows:

$$X^{VG}(t; \sigma, \nu, \theta) = \Gamma(t; \mu_p, \mu_p^2 \nu) - \Gamma(t; \mu_n, \mu_n^2 \nu) \quad (3)$$

$$\text{with } \mu_p = \frac{1}{2} \sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} + \frac{\theta}{2} \text{ and } \mu_n = \frac{1}{2} \sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} - \frac{\theta}{2}$$

As presented, the Variance Gamma process has two presentations which allow it to be more flexible to model different phenomena. In this study, the VG is the mathematical model used to present the degradation of the water tank pump system by fitting to the system degradation data.

### III. DATA

It is always significant to assess a good reliability and to predict failure of systems in order to decrease the risk of failure as much as possible. In general, a system is considered as failed when they are no longer able of performing its function. In this part, the retrieved data are going to be presented as an example of application of the VG process. The historical degradation data (Fig 2) of the water tank pump presented in this study are collected at different periods of the year using different sensors and it describes the level of water leakage with respect to time, i.e. each 4 hours per day. The rate of the leakage is controlled and registered since it will help in defining the distribution of the failure time that helps in defining the prognostic of the systems. The water tank pump is considered as failed when its leakage rate at time  $t$  is denoted by  $X(t)$ , is greater than a specific threshold ( $d$ ), assuming that the threshold  $d$  must be already fixed. As soon as the leakage rate of the pump exceeds a threshold limit, which is set at around 1400 l/h (litre per hour), the pump is stopped and a corrective maintenance action is performed. When such incident happens, the pump will be maintained and its seals must be changed. In order to avoid the failure of the system, reliability engineers and safety practitioners define an alarm threshold in order to propose proactive maintenance decisions. The alarm threshold is set at a level of 1100 l/h. This threshold alerts the operator early enough to a possible degradation of the pump and to carry out some maintenance operations in order to avoid failure. It is recommended in such problems to refer to monitoring, but in this paper the aim is to present the VG process as a degradation model applied on a mechanical system.

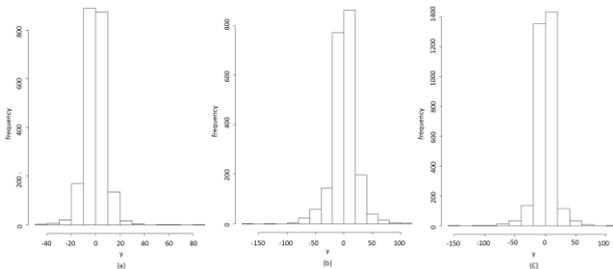


Fig 2. Histogram of data

### IV. RMSE AND R PACKAGES VALIDATION

In this section, the interest is on the calculation of the RMSE in order to compare the two existing R packages used for the estimation. The goal is to select the most efficient package that can estimate the parameters with less estimation error. The estimation of VG parameters is an important step in prognosis, for that it has been recently the subject of several research papers. According to [15], it is impossible to replicate the estimation obtained in [14]. In order to understand this problem, an investigation of the computational problems related to the maximum likelihood estimator is performed. R, MATLAB and a non-standard optimization software such as Ezgrad were used in this investigation. The log-likelihood function is considered as complex because of the presence of many local optima and the presence of the Bessel function of the second kind in its expression. The R Variance Gamma package use the VGfit function which allows the user to employ the Nelder-Mead, the BFGS method (Broyden-Fletcher-Goldfarb-Shanno) or a Newton-type algorithm. While the ghyr R package use the fit.VGuv function which is

based on the Nelder–Mead algorithm. In their paper [16], they developed a new algorithm and the results of the estimation were compared with the results obtained from R Variance Gamma package [17] and the ghyr package [18]. In this paper, the estimation of Variance Gamma parameters was insured using the two R packages: Variance Gamma and ghyr packages. The focus of this part is to study the efficiency of the two already existing R package used to estimate the VG parameters. The RMSE (root mean square error) is calculated and used to evaluate the performance of these two packages. The evaluation is done using a number of 36 real data samples with thousands of observations are used (between 1000 and 4000 observations). The estimation in the two R packages is mainly based on the calculation of the maximum likelihood of the VG process. The maximum likelihood function is introduced as follows:

$$L(\mu, \theta, \sigma, \nu) = \frac{T}{2} \log \frac{2}{\pi} + \sum_{t=1}^T \frac{(X(t)-\mu)\theta}{\sigma^2} - \sum_{t=1}^T \log(\Gamma(a)\theta) + \sum_{t=1}^T \log \left( K_{\nu-0.5} \left( \frac{\sqrt{2\sigma^2+\theta^2}|X(t)-\mu|}{\sigma^2} \right) \right) + \sum_{t=1}^T \left( \mu - \frac{1}{2} \right) \left( \log(|X(t) - \mu| - \frac{1}{2} \log(2\sigma^2 + \theta^2)) \right) \quad (4)$$

The work consists of the estimation of the data parameters using the two R packages results and the comparison between them. The focus was to compare the performance of the two R packages using the RMSE results. It is important to choose the best package which will help in having a good estimation in order to achieve a better prognostic. Among the real data, the choice was on different samples, which describes different behaviour of VG. The job consists on estimating the real parameters of the same data 1000 time and the calculate of the RMSE is obtained using the following equation:

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{n}} \quad (5)$$

TABLE I. RMSE ESTIMATION RESULTS

	VG( $\mu=1, \sigma=2, \theta=2, \nu=2$ )			
	$\mu$	$\sigma$	$\theta$	$\nu$
ghyp	2.2448e-06	0.002060	0.0010353	0.0043101
VGfit	5.3679e-05	0.00369	0.0597441	0.0048243
	VG( $\mu=1, \sigma=1, \theta=2, \nu=2$ )			
	$\mu$	$\sigma$	$\theta$	$\nu$
ghyp	7.4190e-08	0.0006329	0.003265	0.0044128
VGfit	3.3092e-07	0.0126907	0.005723	0.0091818
	VG( $\mu=0, \sigma=2, \theta=3, \nu=1$ )			
	$\mu$	$\sigma$	$\theta$	$\nu$
ghyp	5.6606e-07	0.0017782	0.015155	0.0010810
VGfit	7.2977e-06	0.005696	0.029396	0.004236
	VG( $\mu=0, \sigma=1, \theta=0.5, \nu=0.5$ )			
	$\mu$	$\sigma$	$\theta$	$\nu$
ghyp	0.00034591	0.00053964	0.0011269	0.0009555
VGfit	0.00035332	0.00054489	0.0011668	0.0009612

The results of the RMSE were presented in TABLE I and it is observed that for the different obtained data the ghyr package is more efficient in the estimation of the VG parameters comparing to the Variance Gamma package.

Based on the results presented in TABLE I, the ghy package showed better performance and gave better results than the Variance Gamma package for all the data belonging to different parameters and presenting different system behaviour. The results were better for the Variance Gamma package in the case of a weak VG process where the VG do not show a lot of variation. It can be concluded that, ghy package provides the best estimation comparatively. The estimation of parameters is the first step of the process of determining the best mathematical model. The selection of the process is ensured by applying a goodness of fit tests.

## V. GOODNESS OF FIT AND DATA CALIBRATION

The aim of this section is to choose the best stochastic process between the three processes and this based on the results of the goodness of fit tests. Once the best process is calibrated to the real data, it will be used later to propose a good prognostic of the system. Goodness of fit tests are used to summarize the dissimilarity between observed values and the expected values under the model in question. The observed values are presented by the real retrieved data of the system. The data are presented as vectors ( $V_i$ ) and for each vector its real parameters will be estimated and used to generate new samples (expected) in order to calibrate the good process.

The four goodness of fit tests used in this study are: Kolmogorov-test (KS), Chis-square test (chisq), Anderson-test (AD), Cramer-test (Cramer). The four goodness of fit tests were chosen arbitrary only to confirm the results obtained.

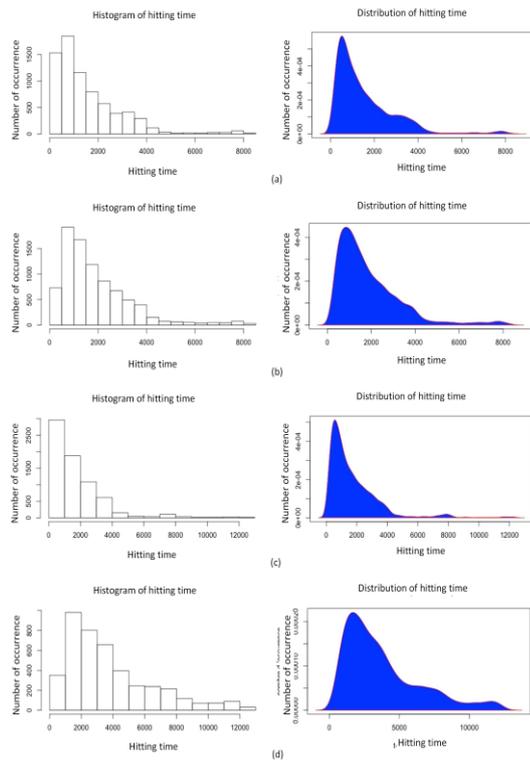


Fig 2. Histogram and distribution of hitting time

These goodness of fit tests are used to calibrate three stochastic processes: Wiener, B&S and VG process. The initial step is to estimate the parameters of the real data of the controlled system. The estimation of the parameters of each ( $V_i$ ) will be done for the three processes: VG, Wiener and B&S. Once the parameters of ( $V_i$ ) are estimated for the Wiener

process, they will be used to generate a new sample of data with the same length of ( $V_i$ ) and then the goodness of fit tests will be applied. The same work will be respectively reproduced for the VG and B&S processes.

The calibration is held with a  $P_{\text{value}} = 0.05$  which means that the error is about 5%. The choice of such  $P_{\text{value}}$  provide us the best reliable results that can be accepted in the industrial world. The results of some chosen data are classified in the TABLE II, which summarize the results of the calibration of the processes.

TABLE II. GOODNESS OF FIT AND CALIBRATION RESULTS

	Process	KS	Chisq	AD	Cramer
V1	VG	0.2439	0.3876	0.3165	0.14236
	Wiener	0.0001854	0.026457	0.001045	0.01002
	B&S	9.158e-10	6.643e-16	0.003193	4.97e-08
V2	VG	0.2472	0.2207	0.1847	0.4246
	Wiener	0.00769	0.001063	0.00907	0.00327
	B&S	0.10382	5.073e-4	0.00513	0.00105
V3	VG	0.2452	0.4309	0.7362	0.11582
	Wiener	8.375e-04	1.282e-10	0.001058	5.51e-04
	B&S	0.081053	1.763e-02	0.00763	0.6354
V4	VG	0.6041	0.3827	0.837	0.2207
	Wiener	4.387e-04	8.361e-08	4.703e-10	0.001063
	B&S	0.001805	0.000951	8.350e-09	0.006301
V5	VG	0.3097	0.5139	0.2106	0.4309
	Wiener	0.00902	0.001753	0.00854	1.76e-08
	B&S	8.045e-09	0.000751	0.005103	0.005395
V6	VG	0.2551	0.7014	0.8104	0.3785
	Wiener	0.04768	0.041981	1.047e-06	0.02451
	B&S	0.01836	4.072e-04	0.03532	0.01849
V7	VG	0.2408	0.063717	0.8105	0.5139
	Wiener	0.046628	0.039015	0.028319	0.02627
	B&S	0.006019	0.0002668	0.000438	0.02451
V8	VG	0.2255	0.2419	0.2412	0.1094
	Wiener	0.033491	0.16179	0.00975	0.007495
	B&S	0.1267	0.030258	0.048609	0.1394
V9	VG	0.2575	0.12123	0.09654	0.2207
	Wiener	3.438e-05	0.010651	0.0085832	2.85e-04
	B&S	0.010341	0.0006654	0.0089969	0.00934

The table resume the results of calibration of the real data ( $V_i$ ) to the three processes. For each ( $V_i$ ), the parameters are estimated from the different processes and then used to generate a new sample. The real data ( $V_i$ ) and the generated data will be tested under the different goodness of fit tests. The result of the calibration is compared to the  $P_{\text{value}} = 0.05$ . For the four tests, the Wiener process cannot model the data of the degradation of the water tank pump. Compared to the  $P_{\text{value}}$ , the results of the calibration of wiener process do not satisfy

the hypothesis of similarity which can lead to conclude that the real data cannot be generated from a Wiener process. The results of the calibration of the B&S for the different ( $V_i$ ) are presented and compared to the  $P_{\text{value}} = 0.05$  affirms that for some ( $V_i$ ), the process is able to model the real data where the results of calibration are greater than the  $P_{\text{value}}$ . But this conclusion is not valid for the all retrieved data. The results of the calibration of the VG process for the data are also presented. The four tests accepted with a  $P_{\text{value}} = 0.05$ , that VG process is the best process to model the water tank pump degradation. For all the data ( $V_i$ ), the results of the calibration obtained using the goodness of fit tests shows that VG process is a good candidate for to model the degradation of the system with similar behaviour. It can be explained because VG process is a stochastic process with four parameters, which gives it more flexibility, but also a good control of the data. Additionally, the representation of VG as a time changed Brownian motion, i.e. the change of the time in the Brownian motion by a gamma process implies more flexibility in fitting the fluctuations and a better reflection of the random speedups and slowdowns in real time. It can be explained because VG process is a stochastic process with four parameters, which gives it more flexibility, but also a good control of the data. Additionally, the representation of VG as a time changed Brownian motion, i.e. the change of the time in the Brownian motion by a gamma process implies more flexibility in fitting the fluctuations and a better reflection of the random speedups and slowdowns in real time.

## VI. ESTIMATION OF FIRST HITTING TIME

According to the presentation of the data and based on the results of the calibration, the threshold of alarm is defined and the aim is to define and estimate the first hitting time (FT). To control the industrial quality, the first hitting time must be defined firstly by defining the quality degradation measures. Accordingly, when the degradation measure reaches a control limit, the first hitting time is established and this information is considered in proposing an efficient maintenance policy. Considering a VG process  $(X_t)_{t \geq 0}$  with initial value  $X_0 = x_0$ , the initial time of maintenance is defined to be the first crossing time of the alarm threshold denoted by ( $d$ ) and it is defined to be the first hitting time which can be written as follows:

$$t_d^* = \inf\{t \geq 0; X_t \geq d\} \quad (5)$$

In this study, the real sample are structured as vectors denoted by ( $V_i$ ) and their real parameters are estimated. Using the estimated parameters for each ( $V_i$ ) a simulation of 10000 samples is carried out and they are structured as vectors denoted by ( $C_i$ ) {i from 1 to 10000} (Fig 4).

The alarm threshold ( $d$ ) is set at a level of 1100 l/h. The density of the FT is defined by defining the density of the time when the samples ( $C_i$ ) crosses for the first time the threshold ( $d$ ). The first time of crossing of ( $C_i$ ) to the alarm threshold is captured and the density is approached and presented. Fig 2 resumes the different densities and histograms for some chosen ( $C_i$ ). The obtained histograms (Fig 2) give a rough sense of the density of the FT. The patterns in the different histograms of the data are right skewed and they have a single-mode. It is always valuable to identify the density of the FT, since it allows to integrate the evolution of the degradation of

the system at the inspection time. Using the four goodness of fit tests, the density of the FT is compared to some empirical distributions known to allow modelling the density of FT. The aim is to fit the distribution of FT to a parametric model such as the exponential distribution, the lognormal distribution, the Weibull and the inverse Gaussian distribution. The exponential and the lognormal distributions are both a right skewed distribution. The Weibull distribution can be right or left skewed or also symmetric. The inverse Gaussian distribution, the Gamma and Beta distribution are also tested. To identify the density of the FT, the goodness of fit tests is used with a  $P_{\text{value}} = 0.5$  and the results are classified in the TABLE III.

TABLE III. FT DISTRIBUTION RESULTS

FT distribution for V1				
	KS	AD	Chisq	Cramer
Exponential	0.04207	0.0158	0.00571	0.02607
Lognormal	0.005867	0.003628	0.000491	0.006827
Weibull	0.033617	0.002151	0.00382	0.00184
Gamma	0.00396	0.00902	0.001753	0.005395
Beta	4.38e-04	0.00047	0.00343	0.01037
Inverse Gaussian	0.00115	0.03835	0.00207	0.000751
FT distribution for V2				
	KS	AD	Chisq	Cramer
Exponential	0.007031	0.004109	0.041981	0.001790
Lognormal	0.01087	0.03668	0.00854	0.000429
Weibull	0.002851	0.001438	0.01095	0.00173
Gamma	0.004762	4.07e-04	0.04259	0.01058
Beta	0.00745	1.05e-07	0.00503	8.32e-05
Inverse Gaussian	0.00694	0.00493	0.02087	0.00047
FT distribution for V3				
	KS	AD	Chisq	Cramer
Exponential	0.1094	0.035009	0.04219	0.00247
Lognormal	0.002601	0.00531	0.000806	0.01553
Weibull	0.00759	0.00458	0.00215	0.000672
Gamma	0.02563	0.00581	2.12e-06	0.001408
Beta	0.01907	2.15e-08	0.001048	0.005024
Inverse Gaussian	0.009262	0.00164	0.071518	2.738e-04
FT distribution for V4				
	KS	AD	Chisq	Cramer
Exponential	0.00507	0.001628	0.00754	0.02593
Lognormal	0.02402	0.01513	0.01307	0.000681
Weibull	0.00287	0.00521	0.00384	0.00156
Gamma	0.00591	0.04529	0.00103	0.0218
Beta	5.01e-05	0.000184	0.00751	1.250e-07
Inverse Gaussian	1.25e-07	0.00162	4.65e-06	0.00372

Using the goodness of fit tests, a comparison between the captured vector of the first hitting time and each parametric model is held. The results of the fitting based on the goodness of fit tests shows that all the parametric models were rejected. For the different ( $V_i$ ), the result of the calibration is compared to the  $P_{value}=0.5$  and as shown in TABLE III all the parametric models were rejected.

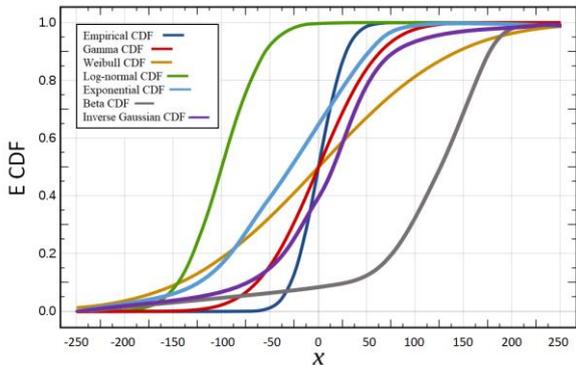


Fig 3. Cumulative distribution function fit of different distributions to FT distribution

The six Cumulative Distributions Function (CDF) were fitted to the retrieved samples, which describes the real first hitting time of the alarm threshold. In this Fig 3, all the proposed distributions were not able to model the FT distribution. The Gamma, Weibull, exponential, lognormal, beta and inverse Gaussian distributions cannot present a distribution that can model the FT distribution of a Variance Gamma process. This implies that the density of the FT of VG process have a specific distribution that cannot be modelled using these candidates. Since there is not a specific parametric model which characterise the FT of the VG process, the simulation method will be used also in proposing prognostic.

## VII. PROGNOSTIC

Based on the results of estimations and fit tests, a good prognostic for the system will be proposed. The prognostic is an important key which provides valuable information for a better maintenance, especially if it is based on predicting the time when the remaining life of a system reaches the predefined maintenance threshold. The degradation of the systems starts before failure occurs by assessing the performance and monitoring the machine degradation will help in avoiding the failure and this by applying earlier corrective actions. Therefore, prognostic focus on monitoring the degradation so that the failure can be predicted. To assure a better prognostic, different type of metrics must be

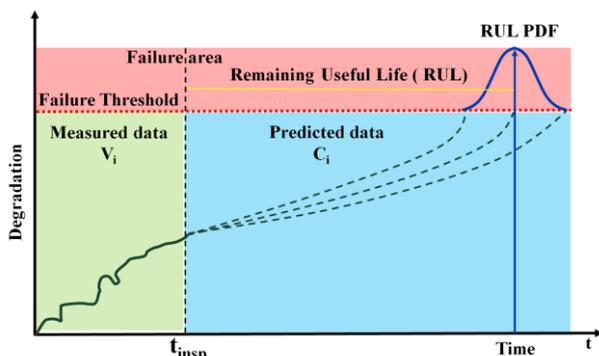


Fig 4. RUL estimation curve

considered such as the RUL (Remaining Useful Life) or the residual time before failure like TTF (Time To Failure). The accuracy in the prediction of the RUL of the defected components is significant to the prognostic of the system and it will assist the operators to replace the components at the appropriate time (Fig 4).

In this study, the calculation of the failure time is performed conditionally in the state of the system at the present time (time of the last inspection  $t_i$ ), which can introduce the RUL as the first exceeding time of indicator  $X(t)$  to the threshold  $d$  follows:

$$t_{t_i}^* := \inf\{h \geq 0; X(t_i + h) \geq d \mid X(t_i) < d\} \quad (6)$$

In this part, for each real data ( $V_i$ ) the parameters are estimated and they are used to generate a 10000 samples representing different degradation paths ( $C_i$ ). An alarm threshold is settled at a level of 1100 l/h and the time of crossing of that level of the degradation paths by the samples is captured. The work focuses on following the evolution of the FT density of some selected data. Some degradation paths are chosen ( $V_i$ ), at  $t=0$  a 10000 samples ( $C_i$ ) are generated and the distribution of the FT is established. To assess the evolution of the degradation at  $t= t_{insp}$ , the same procedure is repeated and the FT is again well established. The objective is to locate the real crossing time of the designed degradation path  $V_i$  (represented as “Time”) compared to the two FT densities obtained at the two designed times. The two densities are obtained using the results of first crossing time of ( $C_i$ ) to the alarm threshold ( $d=1100$  l/h) at the two inspection times  $t=0$  and  $t= t_{insp}$ . The results of the simulation and relative error are presented in the TABLE IV and TABLE V respectively.

TABLE IV. POSITIONING OF THRESHOLD TIME COMPARED TO THE FT

10%	Median	90%	Mean	Time
1072.5	1132.5	1472.5	1267.5	1184
1092.5	1173.5	1494.8	1283.5	1371
1074.5	1279.5	1534.5	1348.5	1126
1685	1769	1881	1750	1793
1745	1829	1927	1857.5	1885
1800	1895	1978	1874	1882
1739	1827	1931	1848.5	1869

TABLE V. RELATIVE ERROR (%) WITH RESPECT TO TIME

10%	Median	90%	Mean
9.42	4.35	24.37	7.05
20.31	14.41	9.03	6.38
4.57	13.63	36.28	19.76
6.02	1.34	4.91	2.4
7.43	2.97	2.23	1.46
4.36	0.69	5.1	0.43
6.96	2.25	3.32	1.1

To evaluate the results of the simulation, the location of the real crossing time of ( $V_i$ ) designed by “Time” is compared to the mean, median, the 10% quantile and the 90% quantile

of the histogram presenting the series of time of exceeding the alarm threshold. The histogram is based on the captured crossing time of the different paths ( $C_i$ ) to the alarm threshold. For that the mean, the median, the 10% quantile and the 90% quantile of the captured vector of time is defined and the position of Time is compared their positions. The results presented in the TABLE IV and TABLE V shows that the real crossing time (Time) is located in the confidence interval of the estimation and is positioned near to the mean and median of the histogram. Based on the results, the real FT is always located in the confidence interval and this can lead us to affirm that VG offers a good prognostic.

### VIII. CONCLUSION

In this study, the VG process is proposed to model the degradation of the centrifugal water tank pump. First, the definition and the properties of the process are presented. After, the parameters of the VG process were estimated using two R packages and a comparison based on the calculation of the RMSE between the two packages is held. Furthermore, an explanation of the water tank pump system degradation is presented and the data retrieved and used in this paper is explained. In order to find the best model for the system, three stochastic processes are evaluated and the calibration is held using goodness of fit tests. These goodness of fit tests was applied and it was concluded that VG process is the best model to fit the data. Moreover, the failure time distribution of the system was also investigated and presented. The prognostic is also suggested and the results demonstrate that VG model provides a good prognostic. It was concluded that VG model is a good stochastic process proposed for modelling the degradation of water tank pump system and proposing a good prognostic. In the further study, the impact of the error of the estimation on the density of the FT and the prognostic of the system will be considered.

### ACKNOWLEDGMENT

The authors would like to acknowledge the valuable financial support of the European Regional Development Fund (FEDER) and the Departmental Council of Aube, France during this research.

### REFERENCES

[1] C. Okoh, R. Roy, J. Mehnen, and L. Redding, "Overview of Remaining Useful Life Prediction Techniques in Through-life Engineering Services," *Procedia CIRP*, vol. 16, pp. 158–163, 2014.

[2] J. Sikorska, M. Hodkiewicz, and L. Ma, "Prognostic modelling options for remaining useful life estimation by industry," *Mechanical Systems and Signal Processing*, vol. 25, no. 5, pp. 1803–1836, 2011.

[3] S. Tang, C. Yu, X. Wang, X. Guo, and X. Si, "Remaining Useful Life Prediction of Lithium-Ion Batteries Based on the Wiener Process with Measurement Error," *Energies*, vol. 7, no. 2, pp. 520–547, 2014.

[4] K. Kanda, H. Sato, T. Miyakoshi, T. Kitano, H. Kanebako, and K. Adachi, "Friction control of mechanical seals in a ventricular assist device," *Biosurface and Biotribology*, vol. 1, no. 2, pp. 135–143, 2015.

[5] M. Mahmoodian and C. Q. Li, "Stochastic failure analysis of defected oil and gas pipelines," *Handbook of Materials Failure Analysis with Case Studies from the Oil and Gas Industry*, pp. 235–255, 2016.

[6] J. V. Noortwijk, "A survey of the application of gamma processes in maintenance," *Reliability Engineering & System Safety*, vol. 94, no. 1, pp. 2–21, 2009.

[7] S. Mercier and I. Castro, "Stochastic comparisons of imperfect maintenance models for a gamma deteriorating system," *European Journal of Operational Research*, vol. 273, no. 1, pp. 237–248, 2019.

[8] Z. Zhang, X. Si, C. Hu, and Y. Lei, "Degradation data analysis and remaining useful life estimation: A review on Wiener-process-based methods," *European Journal of Operational Research*, vol. 271, no. 3, pp. 775–796, 2018.

[9] Y. Zhang, "Wiener and gamma processes overview for degradation modelling and prognostic," *Master's thesis, NTNU*, 2015.

[10] C. Lee, G. Tzeng, and S. Wang, "A new application of fuzzy set theory to the Black–Scholes option pricing model," *Expert Systems with Applications*, vol. 29, no. 2, pp. 330–342, 2005.

[11] T. Nakatsu, "Volatility risk structure for options depending on extrema," *Journal of Computational Finance*, vol. 21, no. 3, pp. 105–122, 2017.

[12] D. Madan and E. Seneta, "The Variance Gamma (V.G.) Model for Share Market Returns", *The Journal of Business*, vol. 63, no. 4, p. 511, 1990.

[13] D. B. Madan and F. Milne, "Option Pricing With V. G. Martingale Components," *Mathematical Finance*, vol. 1, no. 4, pp. 39–55, 1991.

[14] D. B. Madan, P. P. Carr, and E. C. Chang, "The Variance Gamma Process and Option Pricing," *Review of Finance*, vol. 2, no. 1, pp. 79–105, Jan. 1998.

[15] G. P. Cervellera and M. P. Tucci, "A note on the Estimation of a Gamma-Variance Process: Learning from a Failure," *Computational Economics*, vol. 49, no. 3, pp. 363–385, 2016.

[16] M. Bee, M. M. Dickson, and F. Santi, "Likelihood-based risk estimation for variance-gamma models," *Statistical Methods & Applications*, vol. 27, no. 1, pp. 69–89, 2017.

[17] D. Scott, C. Y. Dong, M. D. Scott, and S. RUnit, "Package 'VarianceGamma'", 2018.

[18] W. Breymann, and D. Lüthi. "ghyp: A package on generalized hyperbolic distributions." *Manual for R Package ghyp*