



HAL
open science

Simulation of non-Gaussian stochastic processes with prescribed rainflow cycle count using short-time Fourier transform

Quang Hung Pham, Jérôme Antoni, Antoine Tahan, Martin Gagnon,
Christine Monette

► **To cite this version:**

Quang Hung Pham, Jérôme Antoni, Antoine Tahan, Martin Gagnon, Christine Monette. Simulation of non-Gaussian stochastic processes with prescribed rainflow cycle count using short-time Fourier transform. Probabilistic Engineering Mechanics, 2022, 68, pp.103220. 10.1016/j.probengmech.2022.103220 . hal-03644654

HAL Id: hal-03644654

<https://hal.science/hal-03644654v1>

Submitted on 22 Jul 2024

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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

39 better understand the effect of the excitation on structural fatigue and are for instance used to generate
40 long-term loading (stress, strain, vibration...) needed for fatigue cycle counting. In general,
41 simulations are useful substitutes whenever direct measurements are limited or impossible.

42 The simulation of a NG stochastic process is expected to reproduce some of its critical properties.
43 As far as fatigue analysis is of concern, the engineering practice has focused on the power spectral
44 density (PSD). Spectral-based methods were early introduced to simulate stochastic processes [1-3]
45 with an ascribed PSD, while assuming that the probability density function (PDF) is Gaussian. These
46 popular methods have been used in numerous applications, such as the fatigue analysis of
47 hydroelectric turbines [4] and long-span bridges [5]. A more difficult goal is to jointly prescribe a
48 PSD and a non-Gaussian PDF [6-7]. Although imposing a NG-PDF can be achieved by various
49 transformations, the latter will also modify the PSD in general. A natural approach is therefore to
50 iteratively impose the PSD and the PDF until convergence is reached. Several iterative methods are
51 based on using the translation field proposed by Grigoriu (1995) [9]. A review of such methods is for
52 instance given by Masters et al. (2003), Bocchini et al. (2008), Li et al. (2013) [8-10]. However, in
53 general, reproducing a pair of prescribed PDF and PSD does not guarantee to properly describe every
54 NG stochastic process. This is because the PDF, as it is meant in the aforementioned works, is the
55 marginal distribution over time and not the joint PDF over all-time instants. In other words, processes
56 with different joint PDFs may well have the same marginal PDF. The consequence is that the
57 simulated stochastic processes may not properly account for the actual fatigue damage engendered to
58 a structure. In fatigue analysis, one characteristic that critically matters is the distribution of high-
59 amplitude cycles. Some authors have addressed this problem by extrapolating the rainflow cycles
60 count or by using the extreme value theory [11-13]. The present work follows a different route and
61 introduces a technique for simulating NG processes with an imposed rainflow cycle count (RCC). The
62 RCC – which returns the distribution of load cycles [25] – is indeed one of the most relevant
63 characteristics of a stochastic process in fatigue analysis [4], [26-27].

64 The present work focuses on the simulation of a NG stochastic process that contributes to the
65 fatigue damage of a structure by representing the extreme amplitude fluctuations. The strategy is to
66 simulate a stochastic process from its inverse Short-Time-Fourier-Transform (STFT), where the STFT
67 coefficients are generated to match a given RCC. The method consists in first generating independent
68 and identically random coefficients, then to apply a frequency-dependent nonlinearity, and finally to
69 correlate them by application of a filter. The parameters of the nonlinearity and of the correlation filter
70 are first estimated from a preliminary STFT, as obtained from a partial observation of the stochastic
71 process. The two key elements of the proposed methodology are the use of the STFT, on the one
72 hand, and the use of a specific nonlinear unit, on the other hand.

73 In the literature, the STFT is widely used as an analysis tool in domains such as speech analysis,
74 acoustics, radar/sonar, etc. [14-18]. Its potential for signal synthesis has been early demonstrated, in
75 particular from partial information of the STFT magnitude [15-16] or ridges [17]. However, as far as
76 the authors know, the STFT has never been used to simulate NG stochastic processes. The STFT
77 provides an invertible time-frequency representation of any signal, which by construction is easily
78 amenable to the imposition of a given PSD. At the same time, the STFT makes it possible to introduce
79 a nonlinearity in the time-frequency domain, which may be frequency – dependent as well as dynamic
80 – i.e., involving time derivatives or time lags -- thus resulting in a joint NG PDF.

81 The choice of the nonlinear unit is crucial to match a given RCC or, more generally, a given joint
82 PDF. For instance, static nonlinearity is often used due to their simplicity to match a given marginal
83 PDF. Furthermore, the choice between a pre-nonlinearity model (e.g., Hammerstein model [19]) or a
84 post-nonlinearity model (e.g., Wiener model [20]) also influences the simulation process. The post-
85 nonlinearity model makes it easier to impose a NG-PDF in the simulation process, but it modifies the
86 PSD in the output. On the contrary, the pre-nonlinearity model easily allows the imposition of a
87 prescribed PSD, but it then modifies the PDF. The simulation method proposed in this paper is of the
88 latter type with a dynamic nonlinearity that proceeds from a generalization of the autoregressive
89 conditionally heteroskedastic model (ARCH), coined λ -ARCH. The advantage of introducing an λ -
90 ARCH model in the STFT is its ability to represent a large class of NG stochastic processes, with
91 complex structures in the time-frequency domain. A typical case is a broadband stochastic process
92 with a nonlinearity localized in a frequency band, or one with different nonlinearities in distinct
93 frequency bands.

94 Furthermore, the key parameters of the model used for simulation can be interpolated over
 95 different operating conditions to simulate signals at non-measured conditions.

96 The performance of the proposed method is demonstrated using field test measurements from a
 97 hydroelectric turbine runner. The hydroelectric turbine is a typical example where experimental
 98 measures are difficult to set up to obtain and only possible at a limited number of operating conditions
 99 due to experimental constraints; therefore, unmeasured operating conditions can conveniently be
 100 replaced by simulations to properly characterize the whole operating range.

101 The paper is organized as follows. Section 2 introduces the STFT decomposition of a NG
 102 stochastic process and a method to simulate it based on the principle of λ -ARCH model. Section 3
 103 then briefly explains how to interpolate a simulation model between operating configurations. Finally,
 104 the case study of a hydroelectric turbine runner case study is discussed in section 4.

105 2. Simulation Methodology

106 2.1 Signal Model

107 a) STFT decomposition

108 As mentioned in the introduction, the aim of the present work is to introduce a method to simulate
 109 realizations of a NG stochastic process, as it would be measured in practice, for instance at the *output*
 110 of a system. Therefore, it makes sense to consider discrete time from the onset. Let $s(n) \in \mathbb{R}$ denote a
 111 discrete-time, zero-mean stationary NG stochastic process, sampled with frequency F_s , with
 112 autocorrelation function
 113

$$114 R_s(\tau) = \mathbb{E}\{s(n)s(n - \tau)\} \quad (1)$$

115 (\mathbb{E} is the expected value operator). Its PSD is $S_s(f) = \mathcal{F}\{R_s(\tau)\} = \sum_{\tau \in \mathbb{Z}} R_s(\tau) e^{-2j\pi\tau f/F_s}$. The STFT
 116 of the stochastic process $s(n)$ is defined by the random field

$$117 S_{ik}(R, \Delta) = \sum_{n=iR}^{iR+N_w-1} s(n)w(n - iR)e^{-2j\pi nk\Delta}, \quad S_{ik} \in \mathbb{C}, \quad S_{ik} = S_{i,-k}^* \quad (2)$$

118 or S_{ik} for short, where i, k stand for the time and frequency indices, respectively. In the above
 119 equation, $\varphi_{ik}(n) = w(n - iR)e^{-2j\pi nk\Delta}$ is a time-frequency kernel rooted on a window function
 120 $w(n)$ with length N_w (assumed even), R is the shift between two consecutive kernels, and $\Delta = 1/N_w$
 121 is the frequency resolution. Alternatively, the stochastic process $s(n)$ can be expanded from its STFT
 122 as

$$123 s(n) = \sum_{i \in \mathbb{Z}} \sum_{k=0}^{N_w-1} S_{ik} \varphi_{ik}^*(n) = 2 \sum_{i \in \mathbb{Z}} \sum_{k=1}^{N_w/2} \Re\{S_{ik} \varphi_{ik}^*(n)\}, \quad (3)$$

124 provided that the window satisfies the condition $\sum_{i \in \mathbb{Z}} w(n - iR) = 1$ (assumption H1). The latter
 125 condition is for instance satisfied with a Hann window and $R = N_w/4$, or can be forced in general by
 126 setting $w(n) = v(n)/\sum_{i \in \mathbb{Z}} v(n - iR)$ based on any positive and smooth function $v(n)$, provided that
 127 R is small enough. It is noteworthy that the second equality in Eq. (3) holds for any zero-mean real
 128 stochastic process, with zero Nyquist frequency.

129 The idea is to simulate a stochastic process from its expansion given by Eq. (3), where S_{ik} is directly
 130 sampled in a random field with properties yet to be defined. As explained in the introduction, this is
 131 likely to offer flexibility for generating NG processes. Indeed, under assumption H1, the STFT is a
 132 highly redundant transform – there are about N_w/R as many STFT coefficients as time samples in the
 133 original signal which provides a high number of degrees of freedom for simulating a stochastic
 134 process.

140

141 From now on, the issue then reduces to generating the random field S_{ik} , such as to reproduce a joint
 142 PDF as similar as possible to those of $s(n)$. The approach is to model S_{ik} as a correlated random field
 143 at the output of a nonlinearity.

144

b) *Intrinsic correlation*

145

146 The first constraint when simulating S_{ik} is to match the PSD of $s(n)$. This is equivalent to
 147 imposing a correlation structure to S_{ik} . From Eqs. (1) and (2), the STFT coefficients should verify

148

$$\mathbb{E}\{S_{ik}S_{i-p,k-q}^*\} = \sum_{\tau \in \mathbb{Z}} R_s(\tau) e^{-2j\pi\tau(k-q)\Delta} S_w(\tau, i, p, q) \quad (4)$$

149

150 where

151

$$S_w(\tau, i, p, q) = \sum_{n \in \mathbb{Z}} w(n - iR) w(n - \tau - (i - p)R) e^{-2j\pi nq\Delta}. \quad (5)$$

152

153 Now, applying the change of variable $m = n - iR$, the latter becomes

154

$$S_w(\tau, i, p, q) = e^{-2j\pi iq\Delta R} \sum_{m \in \mathbb{Z}} w(m) w(m - \tau + pR) e^{-2j\pi mq\Delta} = e^{-2j\pi iq\Delta R} C_w(\tau, p, q), \quad (6)$$

155

156 Finally,

157

$$\mathbb{E}\{S_{ik}S_{i-p,k-q}^*\} = e^{-2j\pi iq\Delta R} \sum_{\tau \in \mathbb{Z}} R_s(\tau) e^{-2j\pi\tau(k-q)\Delta} C_w(\tau, p, q), \quad (7)$$

158

159 which shows that the correlation of the STFT coefficients is a periodic function of i with period
 160 $1/(q\Delta R)$, yet independent of k . Reproducing this correlation structure is necessary in simulations for
 161 properly matching the PSD of $s(n)$. When the frequency resolution is set such that $1/\Delta$ is larger than
 162 the effective duration¹ of $R_s(\tau)$ (assumption H2), the following approximation actually holds,

163

$$\mathbb{E}\{S_{ik}S_{i-p,k-q}^*\} \approx \delta_{q0} C_w(\tau, p, 0) \sum_{\tau \in \mathbb{Z}} R_s(\tau) e^{-2j\pi\tau k\Delta} = \delta_{q0} C_w(\tau, p, 0) S_s(f_k), \quad (8)$$

164

165 with δ the Kronecker symbol, which indicates that the correlation between frequency components can
 166 be ignored as compared to correlation between time instants. Under assumption H2, the correlation of
 167 the STFT coefficients is now a function of frequency k only. Eq. (8) draws a direct connection
 168 between the variance of the STFT coefficients and the PSD $S_s(f_k)$ at frequency $f_k = k\Delta F_s$. This
 169 suggests the normalisation

$$S_{ik} = \sqrt{P_s(f_k)} U_{ik} \quad (9)$$

170

171 where $P_s(f_k) = \mathbb{E}\{|S_{ik}|^2\}$, such that the variance of the coefficients U_{ik} is now independent of k .
 172 Based on these results, the correlation structure given by Eq. (8) may be conveniently forced in
 173 simulations by using the autoregressive (AR) model

174

¹ The effective duration is defined as the shortest time d such that $|R_s(\tau)| < \epsilon R_s(0)$ for all $|\tau| > d$, with $0 < \epsilon < 1$ a given small value. This means that magnitude of $R_s(\tau)$ is considered insignificant for time-lags greater than d

$$\begin{cases} S_{ik} = \sqrt{P_s(f_k)} U_{ik} \\ U_{ik} = \sum_{p=1}^P U_{i-p,k} a_p + Z_{ik} \\ \mathbb{E}\{Z_{ik} Z_{i-p,k}^*\} = \delta_{p,0} \delta_{q,0} \sigma_Z^2 \end{cases} \quad (10)$$

175

176

177

178

179

180

with orders P , Z_{ik} an uncorrelated and stationary sequence with variance σ_Z^2 , and $\{a_p; p = 1, \dots, P\}$ a set of AR coefficients. Indeed, substituting the above equation for U_{ik} in the autocorrelation $\mathbb{E}\{U_{ik} U_{i-p,k}^*\}$ verifies Eq. (8) by construction. The free parameters of the above model are $P_s(f_k)$ and a_p . How to properly set their values to match the target PSD will be explained in subsection 2.2.

181

c) Nonlinear Unit

182

183

184

185

186

187

188

189

190

191

Forcing the particular correlation structure (4) by means of the AR model (6) makes it possible to impose a given PSD. The aim is now to impose a given NG PDF by application of a nonlinear unit to the random coefficients Z_{ik} . Introducing a nonlinearity at this stage offers the possibility to generate a large family of joint NG PDFs, since different nonlinear units can be used depending on the frequency index k . The nonlinearity unit used in this paper is inspired by the autoregressive conditional heteroscedasticity (ARCH) model. ARCH is often used to describe leptokurtic stochastic processes (whose PDF has higher tails than the Gaussian) with high fluctuations [21], i.e. with a variance that is conditionally time-varying. In the present context, a generalization is proposed, coined λ -ARCH. It reads

$$Z_{ik} | \{Z_{i-1,k}, \dots, Z_{i-J,k}\} \sim \mathcal{N}(0, \sigma_{ik}^2) \quad (11)$$

192

193

194

195

196

where $Z_{ik} | \{Z_{i-1,k}, \dots, Z_{i-J,k}\}$ stands for the random variable Z_{ik} at time i conditioned on the random variables $\{Z_{i-1,k}, \dots, Z_{i-J,k}\}$ at previous time instants, and σ_{ik}^2 is a time-varying and frequency-dependent variance that follows the AR equation

$$\sigma_{ik}^{2\lambda_k} = \alpha_{0k} + \sum_{j=1}^J \alpha_{jk} |Z_{i-j,k}|^{2\lambda_k}, \quad \lambda_k \geq 0, \quad \alpha_{0k} > 0, \quad \sum_{j=1}^J \alpha_{jk} < 1 \quad (12)$$

197

198

199

200

201

202

The intuition beyond the λ -ARCH model is that there exists a parameter λ and a set of coefficients $\{\alpha_{jk}; j = 0, \dots, J, k = 1, \dots, N_w/2\}$ such that $|Z_{i,k}|^{2\lambda_k} - \alpha_{0k} - \sum_{j=1}^J \alpha_{jk} |Z_{i-j,k}|^{2\lambda_k}$ is a Gaussian stochastic process uncorrelated with respect to time i . It can be shown that the Z_{ik} defined by Eqs. (11) and (12) is a stationary NG process with λ -moment

$$\mathbb{E}\{|Z_{i,k}|^{2\lambda_k}\} = \frac{\alpha_{0k}}{1 - \sum_{j=1}^J \alpha_{jk}}. \quad (13)$$

203

204

205

206

207

208

209

210

211

212

213

As a particular case when $\lambda = 1$, one obtains an ARCH model with variance $\sigma_{Z,k}^2$ returned by Eq. (13). The particular case with $\lambda = 0$ reads $\sigma_{ik}^2 = \exp(\alpha_{0k} + \lambda \sum_{j=1}^J \alpha_{jk} \ln |Z_{i-j,k}|^2)$.

The exact class of NG distribution that can be generated by the proposed model is an interesting but difficult question to answer. It can be easily implied by Eqs. (11-13) that such distributions are characterized by a finite variance and a finite moment of order 2λ . In particular, this implies leptokurtic distributions when $\lambda = 2$. More generally, it should be noted that the large set of free parameters in the λ -ARCH model will provide high flexibility to approximate a diversity of NG joint PDFs. An important particular case is when the ARCH parameters are independent of frequency k , or piece-wise constant with respect to large frequency bands, which results in a substantial reduction of

214 the number of free parameters (assumption H3).

215

216 The simulation of the stochastic process $s(n)$ then proceeds as described in Algorithm 1.

217

Algorithm 1: Simulation of a NG stochastic process from the STFT

Inputs: $P_k, \{a_p; p = 1, \dots, P\}, \{\alpha_j; j = 0, \dots, J\}, \lambda_k$

Loop on k

Initialize i.i.d sequence $\{Z_{J-1,k}, \dots, Z_{0,k}\}$

Loop on i

Evaluate $\sigma_{ik}^{2\lambda_k}$ *from Eq. (12)*

Draw Z_{ik} *in* $\mathcal{N}(0, \sigma_{ik}^2)$

Evaluate U_{ik} *from Eq. (10)*

Set $S_{ik} = \sqrt{P_k} U_{ik}$

Evaluate $s(n)$ *from Eq. (3)*

218

219 2.2 Model identification

220 The signal model introduced in the previous section involves several parameters to identify
221 experimentally, from a given experimental measurement.

222

223 a) Identification of the correlation structure

224 Let the finite length signal $\{x(n); n = 0, \dots, L - 1\}$ denote the partial observation of one
225 realization of the stochastic process $s(n)$. The first step is to compute the STFT coefficients X_{ik} of
226 $x(n)$ from Eq. (1). Note that index i will range from 0 to $I = \lfloor (L - N)/R \rfloor + 1$. Under assumption
227 H2, the variance of the STFT coefficients can then be estimated as

228

$$\hat{P}_s(f_k) = \frac{\sum_{i=0}^{I-1} |X_{ik}|^2}{I}. \quad (14)$$

229

230 Once the spectral signature of the process is captured, the STFT coefficients are equalized:

231

$$V_{ik} = \frac{X_{ik}(R, \Delta)}{\sqrt{\hat{P}_s(f_k)}}. \quad (15)$$

232

233 The coefficients in the AR model of Eq. (10) can now be estimated by least squares. Let $\mathbf{V}_k =$
234 $(V_{P,k} V_{P+1,k} \dots V_{I-1,k})^T$, $\mathbf{a} = (a_1 a_2 \dots a_P)^T$ and \mathbf{B}_k the matrix with elements $B_{ip} = V_{i-1+P-p,k}$.
235 Therefore, the least-square estimator of \mathbf{a} is returned by

236

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a}} \|\mathbf{V}_k - \mathbf{B}_k \mathbf{a}\| = \left(\sum_{k=1}^{N_w/2} \mathbf{B}_k^T \mathbf{B}_k \right)^{-1} \sum_{k=1}^{N_w/2} \mathbf{B}_k^T \mathbf{V}_k. \quad (16)$$

237

238 The next step is to use the estimated autocorrelation coefficients to whiten the random field V_{ik} , such
239 as to produce the uncorrelated sequence

240

$$\hat{Z}_{ik} = V_{ik} - \sum_{p=1}^P \hat{a}_p V_{i-p,k}, \quad (17)$$

241

242 from which the λ -ARCH model can now be identified.

243

244

245 b) *Identification of the nonlinear unit*

246 Let \mathcal{B} denote a set of frequency indices where the nonlinear unit applies uniformly (i.e., where the
247 coefficients of the λ -ARCH model are constant). Then, for a given value of λ , apply the Box-Cox
248 transformation [22]

249

$$E_{ik}^{(\lambda)} = \begin{cases} |\hat{Z}_{ik}|^{2\lambda}, & \text{otherwise} \\ (\ln|\hat{Z}_{ik}|)^2, & \lambda = 0 \end{cases} \quad (18)$$

250

251 for $k \in \mathcal{B}$. From Eq. (12), it can be shown that $E_{ik}^{(\lambda)}$ follows the autoregressive equation

252

$$E_{ik}^{(\lambda)} = \alpha_0 + \sum_{j=1}^J \alpha_j E_{i-j,k}^{(\lambda)} + \eta_{ik}, \quad k \in \mathcal{B} \quad (19)$$

253 where η_{ik} stands for a zero-mean residual error uncorrelated with $E_{ik}^{(\lambda)}$. This offers a direct means to
254 identify the coefficients $\{\alpha_j; j = 0, \dots, J\}$. Taking the average $\langle E_{ik}^{(\lambda)} \rangle = \sum_{i=0}^{I-1} E_{ik}^{(\lambda)} / I$, one has the set of
255 equations

256

$$\begin{cases} \alpha_0 = \langle E_{ik}^{(\lambda)} \rangle \left(1 - \sum_{j=1}^J \alpha_j \right) \\ E_{ik}^{(\lambda)} - \langle E_{ik}^{(\lambda)} \rangle = \sum_{j=1}^J \alpha_j (E_{i-j,k}^{(\lambda)} - \langle E_{ik}^{(\lambda)} \rangle) + \eta_{ik} \end{cases} \quad (20)$$

257

258 The first equation returns an estimate of α_0 given the α_j 's. The second equation can be solved for the
259 α_j 's, for instance by least-square as in the previous paragraph [23]. The complete identification
260 scheme is resumed in Algorithm 2 under assumption H3.

261

Algorithm 2: Identification of model parameters from an observed signal $x(n)$

Inputs: $x(n)$, P , J , λ

(i). Compute the STFT coefficients X_{ik} from Eq. (2)

(ii). Evaluate $\hat{P}_s(f_k)$ from Eq. (14)

(iii). Evaluate V_{ik} from Eq. (15)

(iv). Compute the AR coefficients a_p from Eq. (16)

(v). Evaluate \hat{Z}_{ik} from Eq. (17)

(vi). Get $E_{ik}^{(\lambda)}$ from the Box-Cox transform (18)

(vii). Estimate the AR coefficients α_j from Eq. (20)

Return $\{\hat{a}_p; p = 1, \dots, P\}$ and $\{\hat{\alpha}_j; j = 0, \dots, J\}$

262

263 For the input of *Algorithm 2*, the parameter λ of the Box-Cox transformations can be optimized by
264 maximizing the log-likelihood function, which ensures that Box-Cox transformation distribution is
265 closest to a Gaussian distribution form. Another simpler strategy is to *a priori* assign it a given value,
266 based on the user experience.

267 Other parameters can be modified to optimize the nonlinearity such as $\hat{\alpha}_{0k}$. Therefore, in the next
268 section, an additional step will be proposed to improve the capacity of reproducing a prescribed RCC
269 by optimizing $\hat{\alpha}_{0k}$.

270

271 2.3 Weighted Area Metric for Rainflow Counting

272 So far, the λ -ARCH identified in section 2.2 might not reproduce exactly a prescribed RCC, since
273 the latter does not appear as a target in Algorithm 2. There are several reasons why discrepancies

274 might exist. First, the selected model orders P and J might not coincide with the true (unknown)
 275 values, a situation that will introduce modeling errors. Second, estimation errors will also alter the
 276 model parameters $\{\alpha_p; p = 1, \dots, P\}$ and $\{\alpha_j; j = 0, \dots, J\}$, and will even propagate from the former to
 277 the latter. One solution to mitigate these errors is to use the output of Algorithm 2 as initial values and
 278 then to update these values so as to match a given RCC.

279 It is noteworthy to know that there exist several metrics dedicated to the comparison of PDFs that
 280 can be used in Eq. (16), such as those based on information theoretic criteria. A distance tailored for
 281 fatigue analysis is proposed next.

282 Let θ denote a set of model parameters to be updated. It may include all the free parameters listed
 283 in section 2.1, or a subset of them, such as typically the parameters of the λ -ARCH. Let also R_{cc}^*
 284 denote the target RCC and $R_{cc}(\theta)$ the RCC of simulated signals, which depends on θ . The aim is to
 285 update the values θ as initially estimated by Algorithm 2 by further minimizing some distance
 286 between the target RCC and the simulated one, i.e.

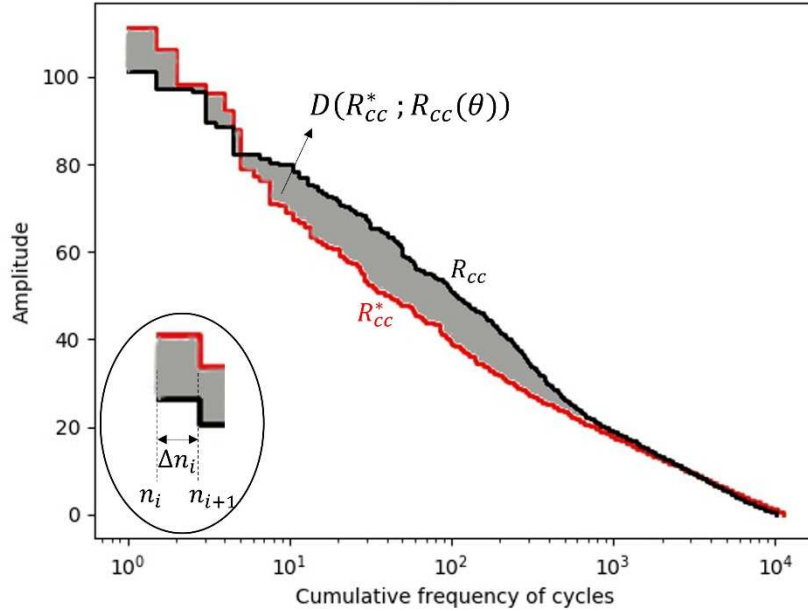
$$287 \quad \hat{\theta} = \arg \min_{\theta} D(R_{cc}^*; R_{cc}(\theta)) \quad (21)$$

288 according to distance D . The Area Metric (AM) is a popularly used to measure the distance between
 289 two cumulative distribution functions [24], The AM between two RCC s is expressed as

$$290 \quad D(R_{cc}^*; R_{cc}(\theta)) = \sum_{i=1}^{i_{max}} \beta_i |R_{cc}^*(\Delta n_i) - R_{cc}(\Delta n_i; \theta)| \Delta n_i \quad (22)$$

291 where Δn_i is the number of cycles corresponding to an amplitude level i and β_i is a user-defined
 292 weight (see *Figure 1*). In order to give more importance to high amplitude cycles, which are more
 293 critical in fatigue, it is proposed to set $\beta_i = \log(1/n_i)$, where $n_i = \sum_{j=1}^i \Delta n_j$ is the cumulative
 294 number of cycles sorted in descending amplitudes.

295



296
 297 *Figure 1. Weighted area metric applied to two Rainflow Counting Counts*

298

299 To reproduce the correct RCC, it is proposed to update the coefficient $\hat{\alpha}_0$ such that

300

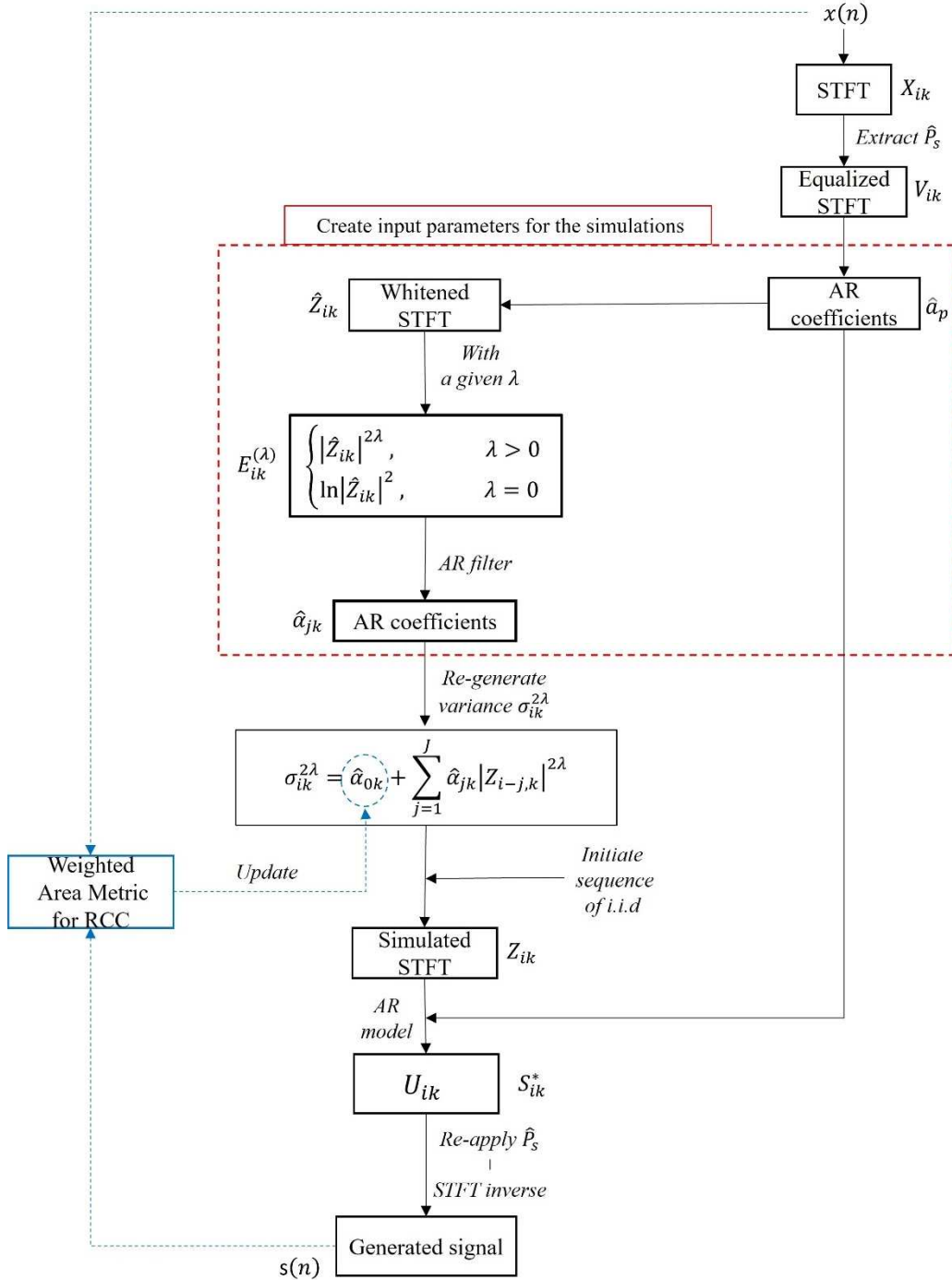
$$301 \quad \sigma_{ik}^{2\lambda} = \hat{\alpha}_0 + \sum_{j=1}^J \hat{\alpha}_j |Z_{i-j,k}|^{2\lambda}, \quad k \in \mathcal{B} \quad (23)$$

302

303 minimizes $D(R_{cc}^*; R_{cc}(\theta))$. Initially, $\hat{\alpha}_0$ is first the value returned from Algorithm 2 when no

303 updating is required, whereas increasing (or decreasing) this $\hat{\alpha}_0$ will increase (or decrease) the effect
 304 of the nonlinear unit. Since only one parameter is involved, the minimization in Eq. (21) can easily be
 305 achieved using grid-search method. As shown in section 4, this simple updating rule will turn out
 306 efficient enough in the case study of this work.

307 Finally, the flowchart of the complete algorithm, including model identification, model updating,
 308 and simulation of the stochastic process is shown in *Figure 2*.
 309



310
 311

Figure 2. Flowchart of the proposed simulation process

312 3. Interpolation Methodology

313 The simulation of a stochastic process is useful when experimental data could not be measured. In
 314 order to deal with this situation, the idea is to interpolate the model parameters (instead of signals)
 315 over different operating conditions where data exist and then to simulate the stochastic process at the

316 operating conditions where data is missing. Although all the model parameters introduced in section
317 2.1 can be jointly interpolated, it is better to keep their number as low as possible in order to guarantee
318 a robust interpolant as well as to reduce the algorithmic complexity. Under the stationary assumption,
319 the autocorrelation coefficients a_p happen to mainly depend on the chosen window parameters.
320 Therefore, they can reasonably be assumed constant over different operating conditions. On the
321 contrary, the interpolation of $P_s(f_k)$ is important if the PSD is likely to vary over different operating
322 conditions. In the case study of section 3, it will similarly be assumed that the coefficients α_j of the λ -
323 ARCH model are reasonably constant, except for the coefficients $\hat{\alpha}_0$, λ introduced in Eq. (23). λ can
324 be fixed using a constant, $\hat{\alpha}_0$ need to be interpolated for the non-linear unit of the model. Since $\hat{\alpha}_0$ is a
325 scalar, it is easily interpolated, for instance with cubic spline interpolation [28-29]. The interpolation
326 of $P_s(f_k)$ is more intricate as it is a function of frequency. For this kind of interpolation, the Spatio-
327 Temporal Kriging (STK) is chosen [30-32]. Let o and f denote the operating condition and frequency
328 variables, respectively. Then, based on known values of the PSD $P_s(o_m, f_l)$ at locations (o_m, f_l) , $m =$
329 $1, \dots, M$, $l = 1, \dots, L$, the STK interpolant reads

$$330 \quad P_s^*(o, f) = \sum_{m=1}^M \sum_{l=1}^L \lambda_{ml}(o, f) P_s(o_m, f_l), \quad (24)$$

331 where superscript * denotes the target value, λ_{ml} is the regression weight, M is the number of known
332 operating conditions and L is the number of studied frequencies at operating condition o_m . With the
333 preservation of the covariance map (or semivariogram) created and modelled from the input set, the
334 weights λ_{ml} in Eq. (24) are estimated by minimizing the error variance $Var(P_s(p_x, o_y) -$
335 $P_s^*(p_x, o_y))$. For more details of STK method, the reader can consult references [30-32]. Finally, a
336 conditional simulation based on kriging parameters is applied to obtain a set of possible interpolated
337 value [33-34].
338
339

340 4. Case Studies

341 Two case studies are used to verify the performance of the proposed method: a numerical
342 experience and an experimental data (hydroelectric domain).
343

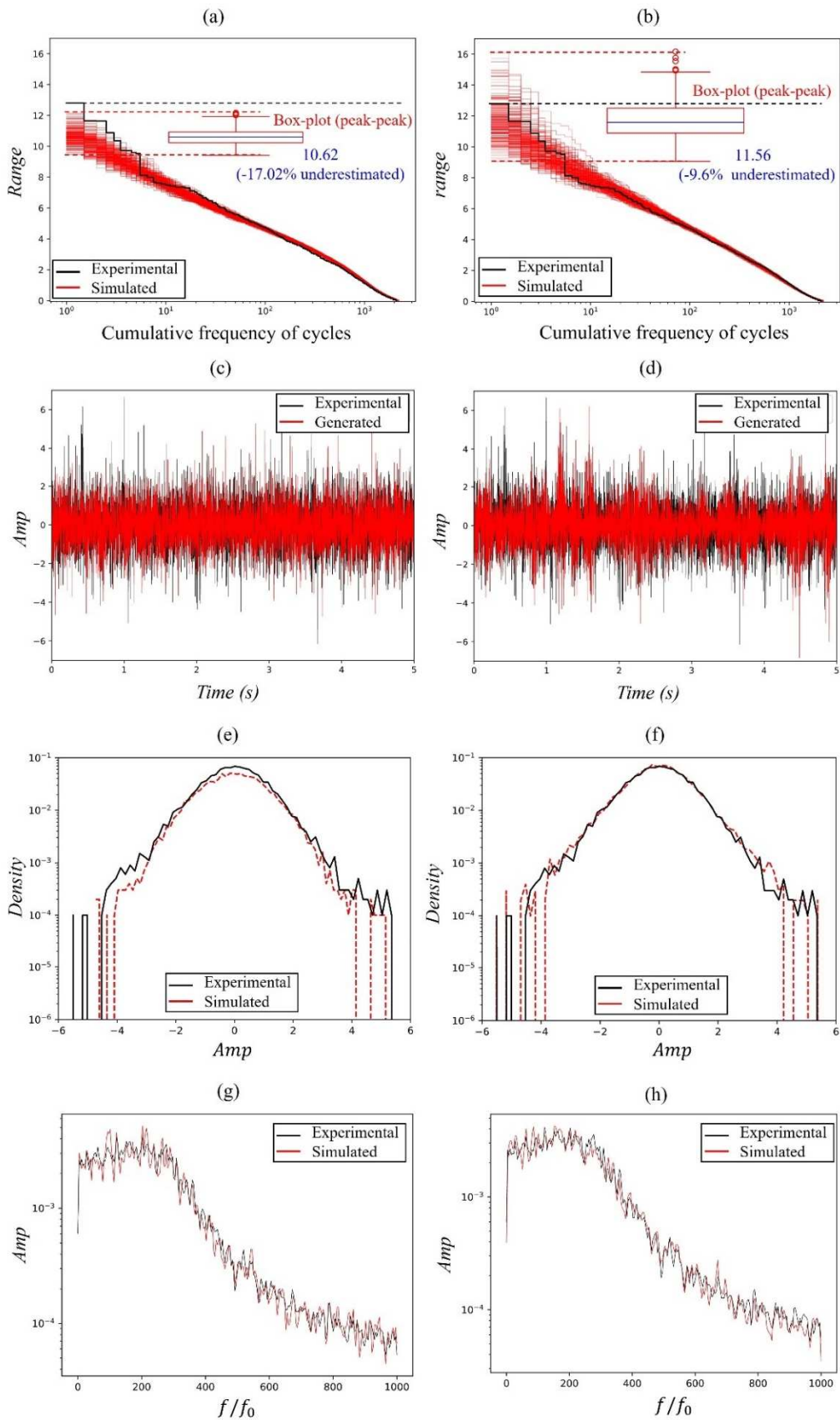
344 4.1 Numerical experience

345 A signal sample is first generated from the t -distribution with zero mean, unitary variance and
346 degree of freedom equal to 4. The t -distribution is used because it provides a leptokurtic behaviour
347 that can highlight the performance of the proposed simulation for the capture of the rainflow count.
348 Then, a numerical filter is applied to the signal in order to assign it a non-constant PSD. The rational
349 transfer function of the filter is:
350

$$350 \quad H(z) = \frac{0.5}{1 - z^{-1} + 0.45z^{-2}} \quad (25)$$

351 The proposed NG simulation is then applied to regenerate realizations of the underlying stochastic
352 process based on the observed signal. *Table 1* displays the parameter set up for the simulation. By
353 default, a Hann window can be safely used as it provides a good compromise between the reduction
354 of frequency leakage and resolution. This comes with 75% overlap ($R = N_w/4$) in order to allow the
355 inversion of the STFT. A critical parameter is the window size N_w , or equivalently the frequency
356 resolution $\Delta = 1/N_w$. The latter should be set fine enough so as to correctly resolve the details of the
357 PSD, which is equivalent to having N_w longer than the effective correlation length of the signal
358 (assumption H2). In practice, some trials and errors might be necessary in setting this value, as is
359 commonplace in experimental spectral analysis. Extensive simulation has shown that an order of 10
360 for the AR model is good enough in most cases, provided that assumption H2 is met. This turned out
361 almost independent of the types of PSD and PDF. As for any parametric model, a difficult setting is

362 the order J in the λ -ARCH model. This value has been set by trials and errors by inspecting the AM
 363 metric, even though more thorough approach may be used [21].
 364



365
 366 *Figure 3 Comparisons of the generation performance of the iterative simulation (left column) and the proposed simulation*

367 (right column) according to the rainflow cycle counts (a, b), one of generated signal (c, d), their PDF (e, f) and their PSD (g,
 368 h)

369

<i>Simulation Parameters</i>	<i>Values</i>
Window function	Hann ($N_w = 2^{10}$)
R	$N_w/4$
Orders of AR model for STFT	$P = 10$
λ	1
J	3
<i>Estimated parameters</i>	<i>Values</i>
$\hat{\alpha}_p$ (real values)	0.0088, -0.0109, 0.0079, 0.0018, 0.0052, 0.0034, 0.0117, 0.0097, 0.0154, -0.0137
$\hat{\alpha}_j$ ($j = 1, \dots, J$)	0.432, 0.226, 0.059
$\hat{\alpha}_0$	0.51

370 *Table 1 Set of parameters for the numerical experience*

371

372

373

374

375

376

377

378

379

380

381

382

383

An iterative simulation scheme with convergence condition², based on the concept of translation field proposed by Grigoriu (1995) [39], that imposes both a prescribed PSD and a marginal PDF, is also applied in this numerical experience to compare with the proposed simulation. *Figure 3* shows comparisons between the two simulations (rainflow count, PDF and PSD of one generated signal). For more detailed comparisons, some distance metrics are calculated in *Table 2* to show the variability (mean \pm standard deviation) of simulation set. In general, one observes that the proposed method performs almost identically to the iterative method in regenerating the PSD and PDF. The advantage of the proposed method is nevertheless obvious when assessing the regeneration of the RCCs, with an AM distance smaller than the iterative simulation. Moreover, as observed in *Figure 3*, the set of RCCs cover the high amplitude (important for fatigue assessment) better than the iterative simulation.

<i>Type of distance (metric)</i>	<i>Iterative simulation</i>	<i>Proposed simulation</i>
Area metric of RCC	365.05 ± 23.73	322.35 ± 25.35
Jensen-Shannon divergence for PDF	0.045 ± 0.004	0.045 ± 0.006
Itakura-Saito distance for PSD	0.028 ± 0.003	0.021 ± 0.003

384 *Table 2 Distance metrics calculated from two simulation sets*

385

386

4.2 Experimental Case Study

387

388

389

390

391

392

In hydroelectric turbines, the water flow entering the runner creates an extreme environment that can make high stress/strain fluctuation [38]. Moreover, with the change of electrical network usage, the turbine must operate more often in operating conditions such as partial load, which increases the high peak in stress/strain signal and the high risk of fatigue on hydraulic runners. Therefore, the hydroelectric turbine case study is a good example to verify the performance of the proposed simulation/interpolation.

393

394

395

396

397

398

399

400

401

The studied data is a set of strain signals obtained from measurements at steady states operating conditions on a Francis turbine runner (a reaction type turbine) in Quebec, Canada [35]. In this case study, to focus on the random part of the signal, hence efforts are directed to simulate the residual part which remains after extracting the synchronous periodic part [4, 36]. Available strain measurements were obtained at different vane opening levels with a relatively constant head. Therefore, the operating conditions in this case study are defined using the vane-opening level (noted as V_o). For example, $20V_o$ corresponds to an operating condition that vanes open about 20% of the maximum opening level (full loading). In the following, the simulation and interpolation based on the proposed methodology will be performed and discussed. Note that some values normalized to respect the

² For each iteration, the convergence point is set when the error (e.g., Itakura-Saito distance) of PSD is smaller than 5% or when a given high number of iterations is reached.

402 confidentiality of such information.

403

404

405

406 *a. Simulation Results and Discussion*

407

408 The first aim is to illustrate and validate the proposed simulation method. Each measurement in
 409 this case study is available in the form of 90-second time series. In particular, the hydroelectric runner
 410 strain measurements by gauges are often available over a limited time window, while some analysis
 411 such as the fatigue assessment, need long-term information. Therefore, to verify the performance of
 412 the proposed simulation, in this case study, only 15 seconds of measurement is used for the training
 413 set to simulate the entire signal of 90 seconds. In this case study, the set \mathcal{B} , where the nonlinearity is
 414 applied uniformly, corresponds to the set of frequencies lower than $100f_0$, with f_0 the frequency
 415 corresponding to the steady synchronous runner rotation. Based on the author's experiences on this
 416 studied signal, this frequency set contains most of the non-Gaussian characteristic of the signal. The
 417 stochastic processes to simulate correspond to the random fluctuations of strain at a steady state
 418 regime in which the runner rotates at a constant speed but with no electrical power output (called
 419 Speed No Load). This operating condition produces complicated random behaviours, which can
 420 significantly contribute to fatigue damage [37,38].

421 Stochastic simulations are run with the parameters shown in *Table 1*. The window type, fraction
 422 of overlap, and order of AR model have all been safety set as in the numerical example of the
 423 previous section. The window length has been set large enough so as return an estimated PSD with
 424 high enough frequency resolution, i.e., such that increasing N_w would not produce any significant
 425 change in the shape of the PSD but only increase its variance. The order of the λ -ARCH and λ have
 426 similarly been set to the smallest values that didn't result in significant changes in the AM metric.
 427 Using these parameters can allow for a minimal calculation cost while still preserving most interest
 428 information.

429

<i>Simulation Parameters</i>	<i>Values</i>
Window function	Hann ($N_w = 2^{12}$)
R	$N_w/4$
Orders of AR model for STFT	$P = 10$
λ	1
J	5
<i>Estimated parameters</i>	<i>Values</i>
$\hat{\alpha}_p$ (real values)	0.0033, 0.008, -0.0014, -0.0093, -0.0007, 0.0024, 0.0003, -0.0030, 0.0031, 0.0019
$\hat{\alpha}_j$ ($j = 1, \dots, J$)	0.189, 0.049, 0.028, 0.039, 0.047
$\hat{\alpha}_0$	4.95

430

430 *Table 3. Set of parameters in the experimental case study*

431

432 *Figure 4* displays one of generated STFTs. Even if the non-linearity is applied uniformly in a wide
 433 set of frequencies, the simulations properly generated the high amplitude zone of STFT. In order to
 434 verify the performance of the proposed methodology to reproduce RCC for fatigue assessment, the
 435 temporal signal generation is compared to two other types of simulation schemes. The first was used
 436 for temporal signal generation of residual strain signal of runners (same signal type with this case
 437 study) by Poirier et al. (2017). They suppose that the residual strain signal, which is remained after
 438 extracting all periodicities, follows a purely Gaussian distribution and then they applied the spectral-
 439 based simulation method (using only a prescribed PSD), as described in [1]. This Gaussian hypothesis
 440 leads to an assumption that certain extreme high peaks are considered as less representative for the
 441 actual deformation of the blade [4]. However, for fatigue analysis, the extreme values cannot be
 442 ignored. In particular, the work of Poirier et al. used a similar type of signal (hydroelectric runner
 443 strain), thus, this comparison can show the improvement in the generation of temporal signal for

444 hydroelectric turbines. The second is the iterative simulation as in the previous section.
 445

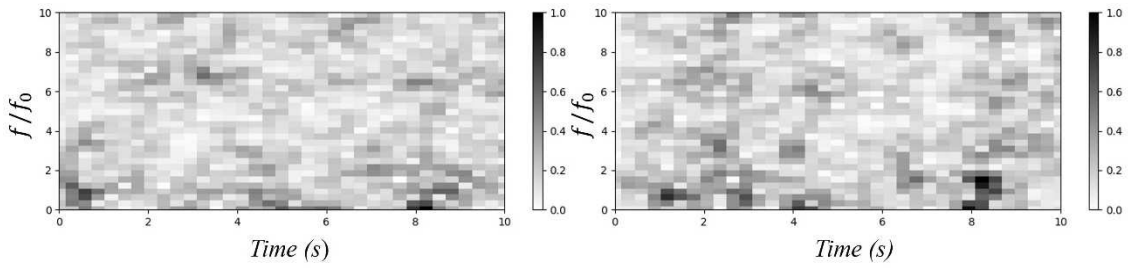


Figure 4. (Left) Experimental STFT and (right) STFT of a simulated signal

446
 447
 448
 449

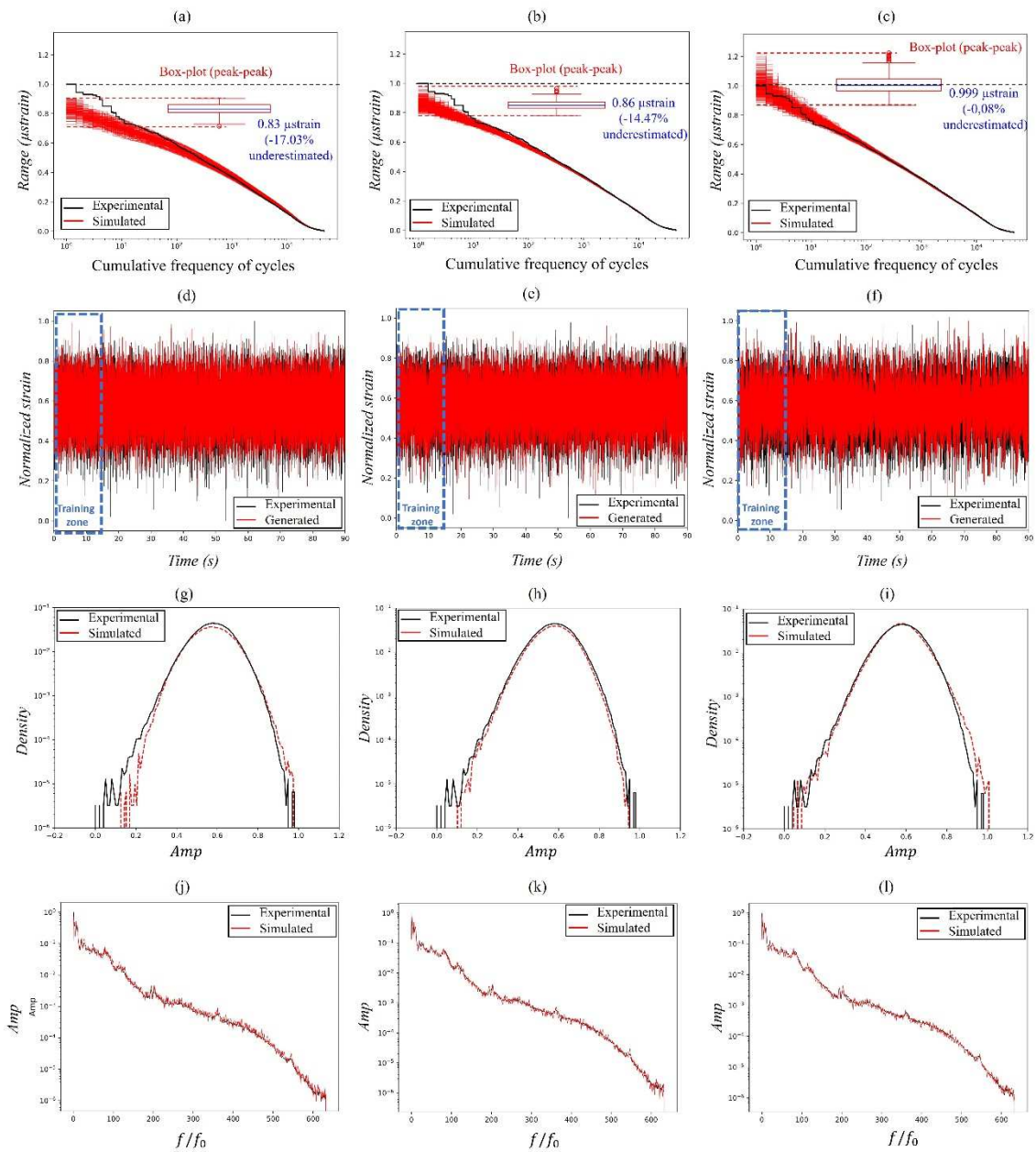


Figure 5. Comparisons of the generation performance of the spectral-based simulation (left column), the iterative simulation (middle column) and the proposed simulation (right column) according to the rainflow cycle counts (a, b, c), one of generated signal (d, e, f), their PDF (g, h, i) and their PSD (j, k, l)

450
 451
 452
 453
 454

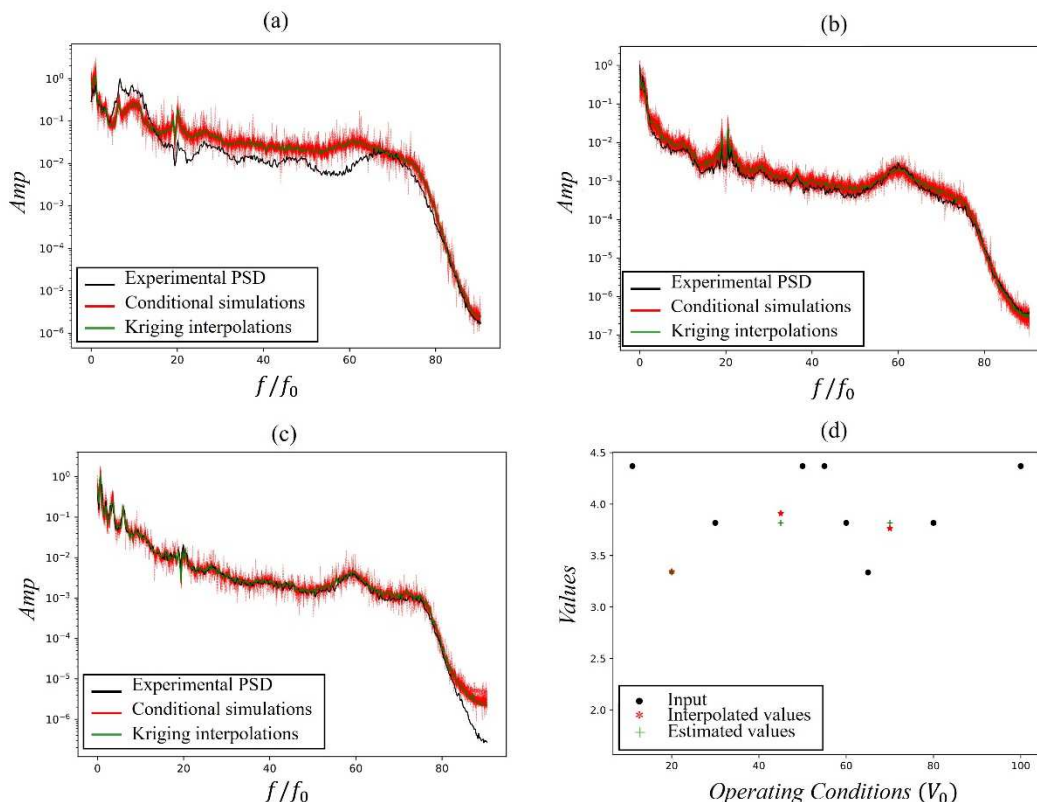
	Area metric of RCC	Jensen-Shannon divergence for PDF	Itakura-Saito distance for PSD
Spectral-based simulation	116.14 ± 15.64	0.040 ± 0.008	$0.017+0.004$
Iterative simulation	90.52 ± 10.89	0.014 ± 0.002	0.024 ± 0.001
Proposed simulation	66.52 ± 9.02	0.034 ± 0.003	0.013 ± 0.001

Table 4. Distance metrics calculated from three simulation sets for the experimental case study

457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474

Figure 5 and Table 4 shows comparisons between three simulations. In this case study, the proposed method has the smallest value of Itakura-Saito distance, which can highlight the advantage of using pre-non-linear model where the PSD is easily imposed. The Jensen-Shannon divergence (JS) of the spectral-based simulation is the highest as the latter is not able to reproduce the NG behaviour of the studied data with presence of extreme values (see distribution tails – Figure 5). The JS variability of the proposed simulation is smaller than the spectral-based simulation and is higher than the iterative simulation. Such value is normal because the PDF is not the target input of the proposed simulation. On the other side, Figure 5 and Table 4 show that, even though the reference methods are based on the reproduction of the target PSD and marginal PDF, none of them cover the high amplitude cycles in the RCC. The proposed method does quite well in this respect with a small error (set mean error < 1% compared to the median of the simulation set, see the boxplot in Figure 5). Moreover, the smallest AM value in the proposed simulation also illustrates the advantage of the optimization using RCC.

b. Interpolation Results and Discussion



475
476
477
478
479
480

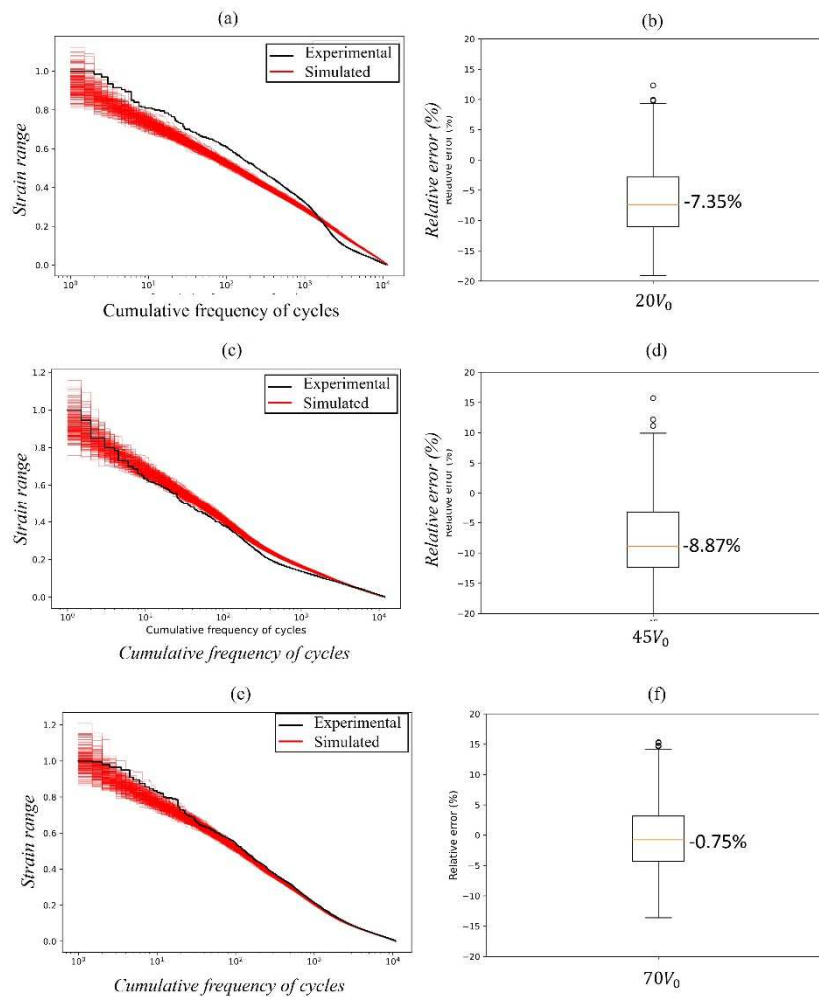
Figure 6. Interpolation sets of PSD at $20V_0$ (a), $45V_0$ (b), $70V_0$ (c), and interpolation of $\hat{\alpha}_0$ (d)

A second objective is to illustrate the capability of the proposed method to simulate strain signals at non-measured operating conditions by interpolating between models. Here, one aims to interpolate the model at $V_0 = 20, 45, 70$ by using the identified parameters at $V_0 = 11, 30, 40, 50, 55, 60, 65, 80$,

481 100. Since there is no information for the fatigue assessment at frequencies higher than $100f_0$, the
 482 signal is downsampled by a factor 7 to reduce the calculation time. The value of λ is fixed as 1 for all
 483 the V_0 .

484 *Figure 6* displays set of interpolations of PSD at three targets operating conditions. The variability
 485 for the lower load condition is higher. The reason is that the $70V_0$ is close to the best efficiency point
 486 of the turbine operation, thus, the fluctuation at this condition is more stable than at $20V_0$ or $45V_0$.

487 Notice that the interpolation quality depends on the number of inputs, while in the case of
 488 experimental measurement, this number is often limited. The quality of the interpolation is assessed
 489 hereafter by means of the RCC s of 200 simulations run with the interpolated model, as shown in
 490 *Figure 7*. The RCCs show errors at low amplitude while the higher amplitude is properly covered by
 491 the simulation set. There are some outliers, which might have been generated because of the error of
 492 interpolation phase. However, the number of the outliers is minor compared to the whole simulation
 493 set.



494 *Figure 7. RCC and relative error for maximum range of simulations set using interpolated parameters at $20V_0$ (a, b), $45V_0$*
 495 *(c, d) and $70V_0$ (e, f)*
 496

497 5. Conclusion

498 This paper has introduced a method for simulating non-Gaussian stochastic processes with a
 499 prescribed RCC, as typically required in reliability and fatigue analysis. The main idea is to generate
 500 the STFT coefficients of the stochastic process after the application of a non-linearity in the form of a
 501 so-called λ -ARCH model. In order to properly reproduce high amplitude cycles, which are important
 502 for fatigue assessment, the simulation parameters are tuned by minimizing the weighted area metric
 503 between the theoretical and the simulated RCCs. Besides, it has been shown that the model used in
 504 simulations can be interpolated between unobserved operating conditions, thus offering the possibility
 505 to predict signals where measurements are missing. The proposed methodology has been applied to

506 analyze a numerical data and an experimental data from hydroelectric turbine runner. These two
507 case studies allow good observations of the proposed simulation performance. In particular.
508 The experimental case study represents a typical type of data measured from a complex structure
509 where the extreme peaks are sensible for the fatigue assessment. This can show the capacity of this
510 proposed method for application in practice. Future work is needed to theoretically delineate the
511 classes of non-Gaussian processes that can be simulated with the proposed method.

512 Obviously, the application of the proposed method can find applications in other domains, since it
513 is flexible enough way to model a large variety of non-Gaussian processes. In particular, the principle
514 used to reproduce a prescribed RCC can be applied to other types of distributions rooted on the joint
515 PDF of a stochastic process.

516

517 Acknowledgments

518 The author would like to thank the Hydro-Québec Research Institute (IREQ), Andritz Hydro
519 Canada Inc, INSA de Lyon (France), École de technologie supérieure de Montréal (Quebec, Canada)
520 and the Mitacs Acceleration program for their support and financial contribution.

521

522 References

523

- 524 [1] Yamazaki, F., & Shinozuka, M. (1988). Digital generation of non-Gaussian stochastic fields. *Journal of Engineering*
525 *Mechanics*, 114(7), 1183-1197.
- 526 [2] Grigoriu, M. (1998). Simulation of stationary non-Gaussian translation processes. *Journal of engineering mechanics*, 124(2), 121-126.
- 527 [3] Shinozuka, M., & Deodatis, G. (1991). Simulation of stochastic processes by spectral representation.
- 528 [4] Poirier, M., Gagnon, M., Tahan, A., Coutu, A., & Chamberland-lauzon, J. (2017). Extrapolation of dynamic load behaviour on
529 hydroelectric turbine blades with cyclostationary modelling. *Mechanical Systems and Signal Processing*, 82, 193-205.
- 530 [5] Cao, Y., Xiang, H., & Zhou, Y. (2000). Simulation of stochastic wind velocity field on long-span bridges. *Journal of Engineering*
531 *Mechanics*, 126(1), 1-6.
- 532 [6] Deodatis, G., & Micaletti, R. C. (2001). Simulation of highly skewed non-Gaussian stochastic processes. *Journal of engineering*
533 *mechanics*, 127(12), 1284-1295.
- 534 [7] Cai, G. Q., & Lin, Y. K. (1996). Generation of non-Gaussian stationary stochastic processes. *Physical Review E*, 54(1), 299.
- 535 [8] Masters, F., & Gurley, K. R. (2003). Non-Gaussian simulation: cumulative distribution function map-based spectral
536 correction. *Journal of engineering mechanics*, 129(12), 1418-1428.
- 537 [9] Bocchini, P., & Deodatis, G. (2008). Critical review and latest developments of a class of simulation algorithms for strongly non-
538 Gaussian random fields. *Probabilistic Engineering Mechanics*, 23(4), 393-407.
- 539 [10] Li, H., & Zhang, D. (2013). Stochastic representation and dimension reduction for non-Gaussian random fields: review and
540 reflection. *Stochastic environmental research and risk assessment*, 27(7), 1621-1635.
- 541 [11] Socie, D. F., & Pompetzki, M. A. (2004). Modeling variability in service loading spectra. *Journal of ASTM International*, 1(2), 1-12.
- 542 [12] Johannesson, P. (2006). Extrapolation of load histories and spectra. *Fatigue & Fracture of Engineering Materials & Structures*, 29(3),
543 209-217.
- 544 [13] Johannesson, P., & Thomas, J. J. (2001). Extrapolation of rainflow matrices. *Extremes*, 4(3), 241-262.
- 545 [14] Schafer, R., & Rabiner, L. (1973). Design and simulation of a speech analysis-synthesis system based on short-time Fourier
546 analysis. *IEEE Transactions on Audio and Electroacoustics*, 21(3), 165-174.
- 547 [15] Nawab, S., Quatieri, T., & Lim, J. (1983). Signal reconstruction from short-time Fourier transform magnitude. *IEEE Transactions on*
548 *Acoustics, Speech, and Signal Processing*, 31(4), 986-998.
- 549 [16] Sturmel, N., & Daudet, L. (2011, September). Signal reconstruction from STFT magnitude: A state of the art. In *International*
550 *conference on digital audio effects (DAFx)* (pp. 375-386).
- 551 [17] Zhu, X., Zhang, Z., Gao, J., & Li, W. (2019). Two robust approaches to multicomponent signal reconstruction from STFT
552 ridges. *Mechanical Systems and Signal Processing*, 115, 720-735.
- 553 [18] Griffin, D., & Lim, J. (1984). Signal estimation from modified short-time Fourier transform. *IEEE Transactions on acoustics, speech,*
554 *and signal processing*, 32(2), 236-243.
- 555 [19] Narendra, K., & Gallman, P. (1966). An iterative method for the identification of nonlinear systems using a Hammerstein model. *IEEE*
556 *Transactions on Automatic control*, 11(3), 546-550.
- 557 [20] Schetzen, M. (1980). The Volterra and Wiener theories of nonlinear systems. Wiley
- 558 [21] Hughes, A. W., King, M. L., & Kwek, K. T. (2004). Selecting the order of an ARCH model. *Economics Letters*, 83(2), 269-275..
- 559 [22] Sakia, R. M. (1992). The Box-Cox transformation technique: a review. *Journal of the Royal Statistical Society: Series D (The*
560 *Statistician)*, 41(2), 169-178.
- 561 [23] Vaidyanathan, Palghat P. "The theory of linear prediction." *Synthesis lectures on signal processing 2.1* (2007): 1-184.
- 562 [24] Liu, Yu, et al. "Toward a better understanding of model validation metrics." *Journal of Mechanical Design* 133.7 (2011).
- 563 [25] Marsh, Gabriel, et al. "Review and application of Rainflow residue processing techniques for accurate fatigue damage
564 estimation." *International Journal of Fatigue* 82 (2016): 757-765.
- 565 [26] Nennemann, B., Morissette, J. F., Chamberland-Lauzon, J., Monette, C., Braun, O., Melot, M., ... & Giroux, A. M. (2014). Challenges
566 in dynamic pressure and stress predictions at no-load operation in hydraulic turbines. In *IOP conference series: earth and*
567 *environmental science* (Vol. 22, No. 3, p. 032055). IOP Publishing.
- 568 [27] Morissette, J. F., Chamberland-Lauzon, J., Nennemann, B., Monette, C., Giroux, A. M., Coutu, A., & Nicolle, J. (2016, November).
569 Stress predictions in a Francis turbine at no-load operating regime. In *IOP Conference Series: Earth and Environmental Science* (Vol.
570 49, No. 7, p. 072016). IOP Publishing.
- 571 [28] Dubrule, O. (1983). Two methods with different objectives: splines and kriging. *Journal of the International Association for*
572 *Mathematical Geology*, 15(2), 245-257.
- 573 [29] Späth, H. (1995). *One dimensional spline interpolation algorithms*. CRC press.

- 574 [30] Montero, J. M., Fernández-Avilés, G., & Mateu, J. (2015). *Spatial and spatio-temporal geostatistical modeling and kriging* (Vol. 998).
575 John Wiley & Sons. (page 266 to 273).
- 576 [31] Pebesma, E., & Heuvelink, G. (2016). Spatio-temporal interpolation using gstat. *RFID Journal*, 8(1), 204-218.
- 577 [32] Sherman, M. (2011). *Spatial statistics and spatio-temporal data: covariance functions and directional properties*. John Wiley & Sons.
- 578 [33] Deutsch, C. V., & Journel, A. G. (1992). Geostatistical software library and user's guide. *New York*, 119, 147.
- 579 [34] Nowak, M., & Verly, G. (2005). The practice of sequential Gaussian simulation. In *Geostatistics Banff 2004* (pp. 387-398). Springer,
580 Dordrecht.
- 581 [35] Breeze, P. (2018). *Hydropower*. Academic Press, Chapter 4.
- 582 [36] Pham, Q. H., Gagnon, M., Antoni, J., Tahan, A. S., & Monette, C. (2019, July). Interpolation of periodic hidden signal measured at
583 steady-operating conditions on hydroelectric turbine runners. In *Surveillance, Vishno and AVE conferences*.
- 584 [37] Liu, X., Luo, Y., & Wang, Z. (2016). A review on fatigue damage mechanism in hydro turbines. *Renewable and Sustainable Energy*
585 *Reviews*, 54, 1-14.
- 586 [38] Seidel, U., Mende, C., Hübner, B., Weber, W., & Otto, A. (2014). Dynamic loads in Francis runners and their impact on fatigue life.
587 In *IOP conference series: earth and environmental science* (Vol. 22, No. 3, p. 032054). IOP Publishing.
- 588 [39] Grigoriu, M. (1995). Applied non-gaussian processes: Examples, theory, simulation, linear random vibration, and MATLAB
589 solutions(Book). *Englewood Cliffs, NJ: Prentice Hall, Inc, 1995*.
- 590 [40]