Title: **A robust-to-parameterization clustering fusion approach for monitoring damage onsets and growth using acoustic emission**

Authors : Emmanuel Ramasso[1]
Vincent Placet[2]
Xavier Gabrion[3]
M. Lamine Boubakar[4]

ABSTRACT

A methodology is presented for acoustic emission (AE) data processing and interpretation, well suited for material characterization of laboratory specimens and in-service structural health monitoring. The methodology relies on the AE streaming in which relevant transients are detected using wavelet-based wave picking, feature vector cleansing with a Mahalanobis-based procedure and a new approach for pattern recognition relying on clustering fusion. The proposed clustering fusion method emphasizes damage kinetics related to the continuous-time definition of AE signals, quantifies uncertainty on clusters (AE sources) evolution, and evaluates the robustness of the results with respect to the change in the parameterization using mutual information. Illustrations concern the early detection and monitoring damage onsets and evolution in a thermoplastic thermostable composite tubular bandage for high-speed rotating engine.

INTRODUCTION

Early detection of damages in structures is of particular importance to ensure the integrity of equipments with the aim to contribute to the reduction of maintenance costs and to the improvement of the availability. Detection of damages can rely on various non-destructive techniques [1, 2] and, in this paper, we especially focus on the acoustic emission (AE) technique which has been shown to be efficient for both material characterization and Structural Health Monitoring (SHM).

The principle of the AE technique is to detect the transient elastic waves originating from the sudden release of energy from localized sources in a material (ASTM E1316). The AE technique is qualified as passive since the sources of AE signals are generally activated by local fractures within a material subjected to loading. In composite materials for instance, fibre-matrix debonding and fibre breakage, delamination and matrix

Authors are with FEMTO-ST institute, 25000 Besançon, France.

[1]E. Ramasso, Department of Automatic Control and Micro-Mechatronic Systems, and Department of Applied Mechanics, Email: emmanuel.ramasso@femto-st.fr.

[2]V. Placet,Department of Applied Mechanics, Email: vincent.placet@univ-fcomte.fr.

[3]X. Gabrion, Department of Applied Mechanics, Email: xavier.gabrion@univ-fcomte.fr.

[4]M.L. Boubakar, Department of Applied Mechanics, Email: lamine.boubakar@univ-fcomte.fr.

cracking are common potential sources of AE signals. Micro-scale damages can thus be detected efficiently which makes this technique suitable for early detection of changes in structural integrity. This technique also allows the detection and the localization of damages in large structures since some components of the stress waves can propagate over long distances [3–5].

The pico/nano-meters displacements induced by the propagation of the stress waves in the material or at its surface can be evaluated by piezo-electric transducers which provides a continuous streaming of transient AE signals. According the sensitivity of the sensors, as well as the operational conditions and the geometry of the structure, the signal-to-noise ratio is generally strongly affected. Data processing methods are thus necessary to extract the most relevant AE signals from the continuous streaming. One difficulty is to cope with the large amount of data, in particular during fatigue tests, due to the high sampling rate required to properly acquire those signals.

SPECIFICITIES OF AE SIGNALS: UNEVENLY-SPACED TEMPORAL DATA

The complexity of the mechanisms involved in the generation and propagation of the AE signals as well as the particular characteristics of those signals require advanced data processing tools, in the ideal case coupled with physics-based approaches to explain wave propagation and fractures [6, 7]. For instance, the stress waves may have components which are strongly impacted by the distance from the sources to the sensors, by the material properties and by the geometry. In [8], we experimentally illustrated the great impact of damage accumulation and change in material properties on the evolution of the signature of AE sources in some CFRP composite materials. This study has emphasized the importance of developing data-driven methods able to cope with the temporal dimension of the AE data generation process. The damage evolution is time-dependent which means that the AE signals are ordered. Therefore the feature extraction or/and the pattern recognition process should take this specificity into account.

A common algorithm for pattern recognition in AE signals used in the literature is called "Kmeans". Proposed some decades ago [9], this method does not take the temporal dimension into account which means that changing the order of the data does not affect the resulting partition. Moreover, this method assumes that the distribution of features is spherical. It is equivalent to considering that clusters are isotropic which means that a change in one dimension (for instance amplitude) implies a similar change in the others (for instance frequency).

The temporal dimension of consecutive AE signals is unfortunately not easily taken into account. The critical point is that the AE signals are unevenly-spaced in time which means that the AE signals are not separated by a constant interval. This specificity prohibits the use of standard algorithms. There are thus two main solutions: 1) use algorithms able to manage continuous-time features, or 2) propose some tricks which allows to take this specificity into account into standard algorithms. The second solution has the advantage to use standard approaches with limited computational resources.

Another critical issue related to the development of dedicated approaches to AE data processing concerns the lack of "ground truths", which are references or baselines to which one can compare the results of algorithms. Turning the problem into novelty detection simplifies in some way the problem by considering "normal" and "faulty" classes

to learn a data-driven approach for SHM [10]. However, in case one is interested in the fine decomposition of the AE streaming into multiple groups (clusters) for a better understanding of the damage onsets and kinetics, this approach is not well suited. Pattern recognition in AE signals is thus generally turned into a clustering or unsupervised classification problem which involves an important challenge in the validation of the partition obtained.

ARCHITECTURE OF THE PROPOSED PATTERN RECOGNITION CHAIN

The methodology follows four main steps and relies on the continuous AE raw data streaming. The first step is called *wave picking* [11–13]. It allows to get two classes of signals: the "relevant" ones, kept for the subsequent steps, and the "noise". For that, a wavelet based denoising has been developed [14]. This approach is of practical interest since it works in real time at high sampling rate on multiple channels. It is important to point out that the denoising step is only used for wave picking which means that the AE signals considered in the subsequent steps are from the raw streaming, without filtering. The choice of the wavelet depends on the streaming content (AE activity and noise).

The second and third steps are *feature extraction* and *data cleansing* respectively. The latter consists in stopping the mechanical sollicitations (if possible, or waiting for pause in loading) and collecting data during this period. Those data are interpreted as background noise used to define an ellipsoid so that feature vectors falling within its boundary are removed. A Mahalanobis distance can be used to define the boundary, as suggested in [10] for SHM applications. This procedure has been applied for composite monitoring during fatigue in [15] where the ellipsoid has been learned on quasi-static tests and adapted with new data during fatigue.

Finally the fourth step is *pattern recognition* which finds out the structure of the data, called partition in the sequel, by exploiting the relationships between features. Pattern recognition may be unsupervised [16], partially supervised [17] or supervised [18] according to the amount of prior available. At this step, it is important to quantify the uncertainty on clusters estimation (at each time) as well as the evaluation of the robustness of the obtained results with respect to changes in the parameterization [19].

THE CLUSTERING FUSION APPROACH: POSITIONING AND PRINCIPLE

Former works on pattern recognition in AE signals made use of *single parameter and single algorithm*. It is frequent to find interpretation of pattern recognition results using a single feature, such as amplitude or frequency. This approach for pattern recognition has found limitations especially for interpretation of AE signals originating from fatigue tests of composite materials which is inherently multivariate.

Then, in the last decades, were developed methodologies based on a *single subset of features and single algorithm*. Questions related to the right number of clusters was tackled using clustering quality indices while feature selection was more recently tackled using "wrapper" [20] and "filter" approaches [15]. The use of clustering quality indices unfortunately makes use of assumptions concerning the distance (generally Euclidean) used in the clustering algorithm and those indices are dependent on the feature space
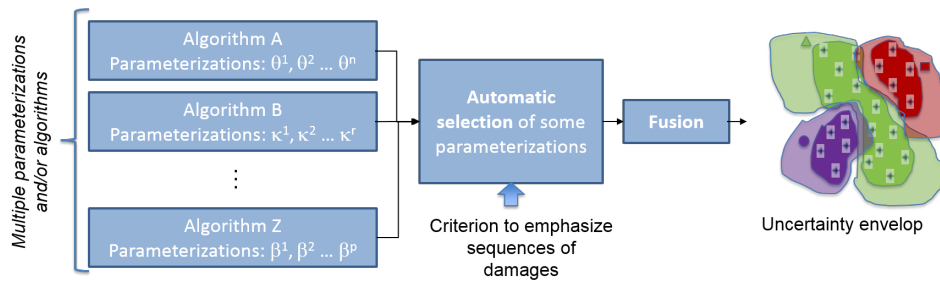
Figure 1. Evaluation of the uncertainty in clustering by information fusion using multiple parameterizations (features, parameters...) and/or multiple algorithms.
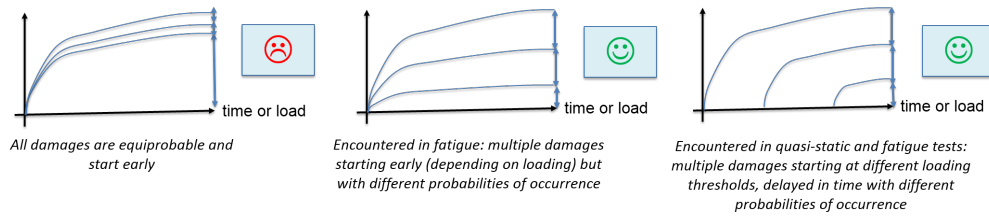

Figure 2. Promote partitions with different probabilities of occurrence of clusters with the aim to emphasize sequences of damages. Each curve represents the cumulated occurrence of a cluster in log-scale (here 3 clusters are assumed for illustration purpose). Arrows on the right-hand side represents the way the probabilities of occurrence are evaluated (by pairs).

(which also makes use of a particular distance).

The clustering fusion approach proposed in [19] relies on *multiple subsets of features and multiple algorithms* (Figure 1). Clustering fusion aims at drawing benefits from multifarious subsets of features and possibly complementary algorithms. The approach is flexible since either multiple algorithms or a single algorithm with different parameterizations can be used. For instance, Kmeans, Hidden Markov Models and Gustafson-Kessel algorithms were considered. Concerning the optimization of the number of clusters, it relies on an information-theoretic criterion (mutual information) that is independent from the features, it only requires various partitions. There is thus no additional distance measures.

The main assumption behind the algorithm of clustering fusion proposed in [19] holds in the probability of occurrence of each cluster which is supposed to be *not* equiprobable (Figure 2). The figure shows how the probabilities are evaluated using pairs of successive clusters sorted by proportions (sorting also allows to perform the fusion of partitions). The assumption is not strong in the sense that the algorithm will give *preference* to subsets of features promoting clusters with different proportions (for instance the figure on the middle or on the right-hand side). However, if all subsets of features agree in providing partitions where the clusters have similar proportions (for instance the left-hand side figure), then the algorithm will not introduce any bias and will provide those partitions as the final result. Besides, if the clusters are assumed to repre-

sent AE sources characteristics, different probabilities of occurrence of clusters means that the activity of each AE source is different. This assumption initially comes from the composite materials considered in our studies where the damage evolution can be decomposed into a sequence involving matrix cracking, delamination, hoop splitting, fibre/matrix debonding and fibres (tows or single) breakage among other damages. The kinetics observed using multi-instrumentation clearly showed a sequence of occurrence of the damages with different proportions. For instance, the matrix was highly emissive generating many AE signals while the hoop splitting initiated and propagated gradually, by saccades, during loading. The methodology has been applied on thermoset, thermoplastics thermostable and various biocomposites where a sequence of damages can be observed experimentally and numerically.

The hypothesis on the probability of occurrence has also an impact on the chronology of the evaluated damages (Figure 2). This is another key point of the algorithm. Indeed, it was experimentally observed that, by compelling the probabilities of occurrence to be different, the damage onsets estimated are generally *delayed* in time compared to standard clustering algorithms (for example on the right-hand side figure) resulting in a clearer chronology. This side effect has been emphasized in simulated data using Markov Chain Monte Carlo sampling in [19]. The temporal dimension of AE data is thus implicitly taken into account.

Finally, the *clusters' parameters can be adapted gradually with new data* in particular with the use of the Gustafson-Kessel clustering. The evolving Gustafson-Kessel method has indeed been developed and exploited in statistical process control [21] and prognostics [22] which makes it a good candidate for SHM [23]. Compared to the Kmeans, for which an evolving variant was proposed by [24] for AE signals, the Gustafson-Kessel method manages arbitrary clusters shape.

A FLAVOUR OF RESULTS ON TUBULAR COMPOSITE MATERIALS

Investigations were performed on ring-shaped specimens produced by cutting 6-layers unidirectional filament wound carbon tubes. A thermoplastic thermostable matrix belonging to the family of polyimide and a high-strength carbon fibre were used [25]. The split-disk test specimens were machined with respect to the ASTM D2290 standard geometry and dimensions, except for the size of the notches. Each specimen has a width of 18 mm and involves two sections of reduced area, which are located 180° apart from each other. The width of the reduced area is 12 mm. The internal diameter is 120 mm. Tensile tests were performed on the split-disk specimen using a home-made test fixture.

AE streamings were recorded during mechanical loading using a four channels data acquisition system form EPA corporation (MISTRAS Group) with miniature piezoelectric sensors (mu-80) with preamplifiers gain equal to 20 dB, a 20-1000 kHz filter and a PCI card with a sampling rate equal to 2 MHz.

Figure 3 pictorially depicts the evolution of the damages deteted by combining 50 partitions resulting from multifarious subsets of features (the number of AE signals collected was close to 350000). The method automatically selected the subsets using an entropy-based criterion computed on the probabilities of occurrence of clusters [19]. It automatically found 6 clusters according to a robustness criterion based on mutual infor-
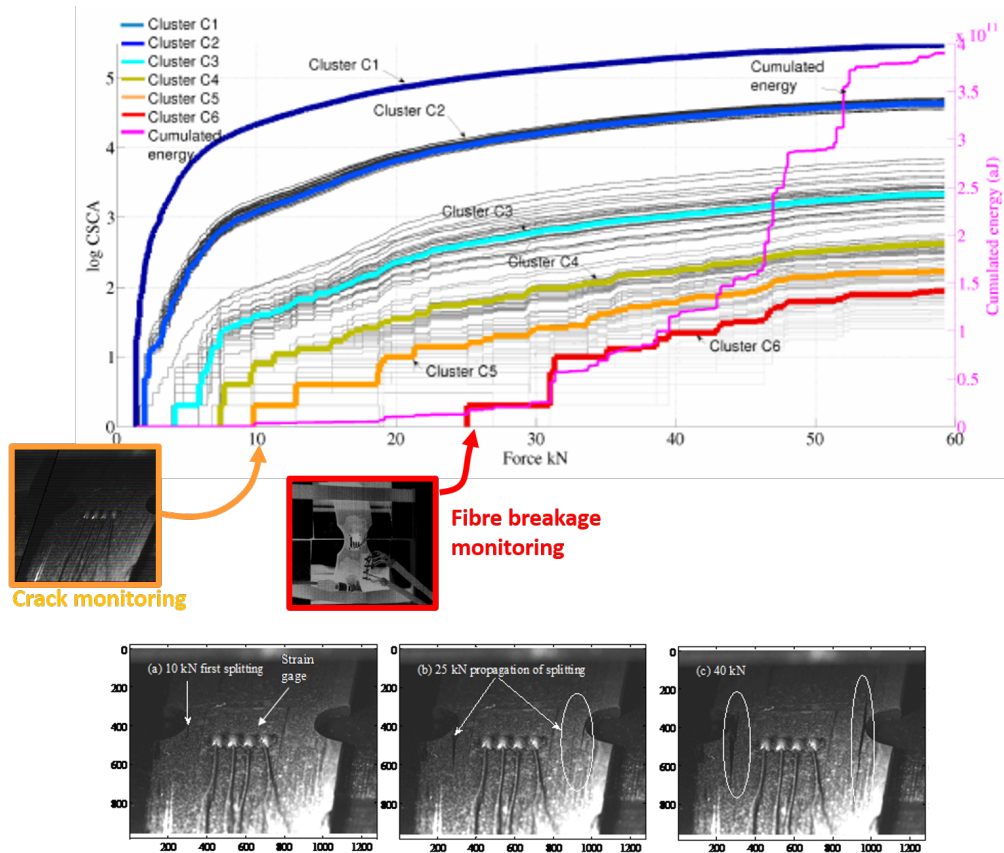
Figure 3. Clustering result using a fusion of 50 parameterizations. Bottom pictures represent CCD and infrared images emphasizing the initiation and propagation of the hoop splitting (CCD) and highly energetic fibre tows breakage (IR images). The curves represent the cumulated occurrence of 6 AE sources determined by the algorithm during loading.

mation computed on around 50 partitions. This criterion has been first proposed in [26] and adapted for AE signals in [19].

Of particular interest, clusters 3, 4 and 5 are highly related to the detection and the propagation of the hoop splitting starting around the notches. Since this damage is quite complex involving a mixing of plies, matrix and fibres failures, it seems relevant to consider that it could be described by multiple clusters. Visually, the evolution of cluster 5, starting around 10kN, was quite well related to the propagation of the splitting observed on the CCD camera (bottom left-hand side picture). In addition, cluster 6, starting around 25kN, fits quite well the evolution of the cumulated energy of AE signals as well as observed events on the infrared camera (bottom right-hand side picture). This means that this cluster is probably related to highly energetic events such as tow breakage.

This figure illustrates the characteristics of the proposed clustering method: different probabilities of occurrence of clusters (the end value of the cumulated damage curves on right-hand side, around 60kN) and onsets of damages are spread on the force axis emphasizing a chronology (cluster 2 is followed by cluster 3, followed by cluster 4 and so on). The chronology also appears in the changes of the slope of each curve. For

instance, the first occurrence of cluster 4 appears when the slope of the curve related cluster 3 changes around 8kN. Moreover, the "steps" observed on a curve represent sudden assignments of multiple AE signals to a given cluster. Those sudden changes emphasize a variation in the kinetics of damages. Those changes could be exploited to early warn against the occurrence of more important failures for prognostics purposes.

CONCLUDING REMARKS

The clustering fusion approach is a new way of computing "natural" clusters in AE signals and focuses on the sequence of damages emphasizing a chronology in place of clusters shape. Any clustering algorithm can be used, however the choice of the Gustafson-Kessel algorithm is justified by the possibility to represent more general clusters shape and clusters parameters can be adapted in real time. The method proposed makes use of multifarious subsets of features in place of selecting a unique one. By optimizing two information-theoretic criteria (entropy and mutual information), it finds out a consensus partition and quantifies its uncertainty and the robustness against variations in the parameterization of the algorithm. Illustrations on thermoplastic thermostable tubular composite characterized by a complex damage evolution show the interest of the approach for early detection of critical damages.

REFERENCES

1. Boller, C., F.-K. Chang, and Y. Fujino, eds. 2009. *Encyclopedia of structural health monitoring*, Wiley.
2. Giurgiutiu, V. 2014. *Structural Health Monitoring with Piezoelectric Wafer Active Sensors*, Academic Press.
3. Baxter, M. G., R. Pullin, K. M. Holford, and S. L. Evans. 2007. "Delta T source location for acoustic emission," *Mechanical systems and signal processing*, 21(3):1512–1520.
4. Eaton, M. J., R. Pullin, and K. M. Holford. 2012. "Acoustic emission source location in composite materials using Delta T Mapping," *Composites Part A: Applied Science and Manufacturing*, 43(6):856–863.
5. Niri, E. D., A. Farhidzadeh, and S. Salamone. 2014. "Nonlinear Kalman Filtering for acoustic emission source localization in anisotropic panels," *Ultrasonics*, 54(2):486–501.
6. Cesnik, C. E. S. and A. Raghavan. 2009. *Fundamentals of Guided Elastic Waves in Solids*, John Wiley & Sons, Ltd, ISBN 9780470061626, doi:10.1002/9780470061626.shm002.
7. Wevers, M. and K. Lambrighs. 2009. *Applications of Acoustic Emission for SHM: A Review*, John Wiley & Sons, Ltd, ISBN 9780470061626, doi:10.1002/9780470061626.shm011.
8. Kharrat, M., V. Placet, E. Ramasso, and M. Boubakar. 2016. "Influence of damage accumulation under fatigue loading on the AE-based health assessment of composite materials: Wave distortion and AE-features evolution as a function of damage level," *Composites Part A: Applied Science and Manufacturing*.

9. MacQueen, J. et al. 1967. "Some methods for classification and analysis of multivariate observations," pp. 281–297.

10. Farrar, C. and K. Worden. 2013. *Structural Health Monitoring: A Machine Learning Perspective*, John Wiley & Sons, Ltd.

11. Kurz, J. H., C. U. Grosse, and H.-W. Reinhardt. 2005. "Strategies for reliable automatic onset time picking of acoustic emissions and of ultrasound signals in concrete," *Ultrasonics*, 43(7):538–546.

12. Grosse, C. U. and M. Ohtsu. 2008. *Acoustic emission testing*, Springer Sc. & Bus. Media.

13. Pomponi, E., A. Vinogradov, and A. Danyuk. 2015. "Wavelet based approach to signal activity detection and phase picking: Application to acoustic emission," *Signal Processing*, 115:110 – 119.

14. Kharrat, M., E. Ramasso, V. Placet, and M. Boubakar. 2016. "A signal processing approach for enhanced Acoustic Emission data analysis in high activity systems: Application to organic matrix composites," *Mechanical Systems and Signal Processing*, 70:1038–1055.

15. Doan, D., E. Ramasso, V. Placet, S. Zhang, L. Boubakar, and N. Zerhouni. 2015. "An unsupervised pattern recognition approach for AE data originating from fatigue tests on polymer–composite materials," *Mechanical Systems and Signal Processing*, 64:465–478.

16. Huguet, S., N. Godin, R. Gaertner, L. Salmon, and D. Villard. 2002. "Use of acoustic emission to identify damage modes in glass fibre reinforced polyester," *Composites Science and Technology*, 62:1433 – 1444.

17. Ramasso, E., V. Placet, R. Gouriveau, L. Boubakar, and N. Zerhouni. 2012. "Health Assessment of Composite Structures in Unconstrained Environments Using Partially Supervised Pattern Recognition Tools." in *Annual Conf. of the Prognostics and Health Management Society*, pp. 1–11.

18. Morizet, N., N. Godin, J. Tang, E. Maillet, M. Fregonese, and B. Normand. 2016. "Classification of acoustic emission signals using wavelets and Random Forests: Application to localized corrosion," *Mechanical Systems and Signal Processing*, 70:1026–1037.

19. Ramasso, E., V. Placet, and M. L. Boubakar. 2015. "Unsupervised consensus clustering of acoustic emission time-series for robust damage sequence estimation in composites," *IEEE Transactions on Instrumentation and Measurement*, 64(12):3297–3307.

20. Sause, M., A. Gribov, A. Unwin, and S. Horn. 2012. "Pattern recognition approach to identify natural clusters of acoustic emission signals," *Pattern Reco. Lett.*, 33:17–23.

21. Angelov, P., D. P. Filev, and N. Kasabov. 2010. *Evolving intelligent systems: methodology and applications*, vol. 12, John Wiley & Sons.

22. Serir, L., E. Ramasso, P. Nectoux, and N. Zerhouni. 2013. "E2GKpro: An evidential evolving multi-modeling approach for system behavior prediction with applications," *Mechanical Systems and Signal Processing*, 37(1):213–228.

23. Placet, V., E. Ramasso, L. Boubakar, and N. Zerhouni. 2013. "Online segmentation of acoustic emission data streams for detection of damages in composites structures in unconstrained environments," in *11th International Conference on Structural Safety & Reliability*, pp. 1–8.

24. Pomponi, E. and A. Vinogradov. 2013. "A real-time approach to acoustic emission clustering," *Mechanical Systems and Signal Processing*, 40(2):791–804.

25. Gabrion, X., V. Placet, F. Trivaudey, and L. Boubakar. 2016. "About the thermomechanical behaviour of a carbon fibre reinforced high-temperature thermoplastic composite," *Composites Part B: Engineering*, 95:386–394.

26. Fred, A. and A. Jain. 2005. "Combining multiple clusterings using evidence accumulation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(6):835–850.

DEStech Publications, Inc.

CONTRIBUTING AUTHOR COPYRIGHT RELEASE FORM

As author of the chapter/contribution titled _**A robust-to-parameterization clustering fusion approach for monitoring damage onsets and growth using acoustic emission**_, to appear in the *Proceedings of Structural Health Monitoring 2017*, I hereby agree to the following:

1.   To grant to DEStech Publications, Inc., 439 North Duke Street, Lancaster, PA, 17602, copyright of the above named chapter/contribution (for U.S. Government employees to the extent transferable), in print, electronic, and online formats. However, the undersigned reserve the following:
a. All proprietary rights other than copyright, such as patent rights.
b. The right to use all or part of this article in future works.

DEStech Publications thereby retains full and exclusive right to publish, market, and sell this material in any and all editions, in the English language or otherwise.

1   I warrant to DEStech Publications, Inc., that I am the (an) author of the above-named chapter/contribution and that I am the (a) copyright holder of the above-named chapter/contribution granted to DEStech Publications, Inc.

2   I warrant that, where necessary and required, I have obtained written permission for the use of any and all copyrighted materials used in the above-named chapter/contribution. I understand that I am responsible for all costs of gaining written permission for use of copyrighted materials.

3   I agree to assume full liability to DEStech Publications, Inc. and its licensee, and to hold DEStech Publications, Inc. harmless for any claim or suit filed against DEStech Publications, Inc. for violation of copyrighted material used in the above-named contribution.

Please sign and date this form and retain a copy for your records. Please include original form with your chapter/paper.
Thank you for your cooperation.

Emmanuel RAMASSO, on behalf of authors

May 12, 2017