Accepted Manuscript

Title: Time since discharge of 9 mm cartridges by headspace analysis, part 2: ageing study and estimation of the time since discharge using multivariate regression

Author: M. Gallidabino F.S. Romolo C. Weyermann

Please cite this article as: M.Gallidabino, F.S.Romolo, C.Weyermann, Time since discharge of 9mm cartridges by headspace analysis, part 2: ageing study and estimation of the time since discharge using multivariate regression, Forensic Science International http://dx.doi.org/10.1016/j.forsciint.2016.12.027

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Time since discharge of 9 mm cartridges by headspace analysis, part 2: ageing study and estimation of the time since discharge using multivariate regression

M. Gallidabino[a,b*] matteo.gallidabino@unil.ch, F. S. Romolo[c], C. Weyermann[a]

[a]Université de Lausanne, École des Sciences Criminelles, Bâtiment Batochime, 1015 Lausanne-Dorigny, Switzerland

[b]King's College London, Analytical and Environmental Science Division, 150 Stamford Street, SE1 9NH London, United Kingdom

[c]Sapienza Università di Roma, Section of Legal Medicine, Department SAIMLAL, Viale Regina Elena 336, 00161 Rome, Italy

[*]Corresponding author: Tel: +41 21 692 46 00; fax: +41 21 692 46 05

Highlights

- Chemometrics were tested on HSSE data for estimating time since discharge.
- Pairwise log-ratio normalisation was found to be a reliable pre-treatment method.
- Random forests and partial least squares were suitable regression approaches.
- Estimation of up to 48 h of ageing was possible with acceptable accuracy.

Abstract

Estimating the time since discharge of spent cartridges can be a valuable tool in the forensic investigation of firearm-related crimes. To reach this aim, it was previously proposed that the decrease of volatile organic compounds released during discharge is monitored over time using non-destructive headspace extraction techniques. While promising results were obtained for large-calibre cartridges (e.g., shotgun shells), handgun calibres yielded unsatisfying results. In addition to the natural complexity of the specimen itself, these can also be attributed to some selective choices in the methods development. Thus, the present series of papers aimed to systematically evaluate the potential of headspace analysis to estimate the time since discharge of cartridges through the use of more comprehensive analytical and interpretative techniques.

Following the comprehensive optimisation and validation of an exhaustive headspace sorptive extraction (HSSE) method in the first part of this work, the present paper addresses the application of chemometric tools in order to systematically evaluate the potential of applying headspace analysis to estimate the time since discharge of 9 mm Geco cartridges. Several multivariate regression and pre-treatment methods were tested and compared to univariate models based on non-linear regression. Random forests (RF) and partial least squares (PLS) proceeded by pairwise log-ratios normalisation (PLR) showed the best results, and allowed to estimate time since discharge up to 48 h of ageing and to differentiate recently fired from older cartridges (e.g., less than 5 h compared to more than 1 – 2 days). The proposed multivariate approaches showed significant improvement compared to univariate models. The effects of storage conditions were also tested and results demonstrated that temperature, humidity and cartridge position should be taken into account when estimating the time since discharge.

3

1. Introduction

In the forensic investigation of crimes involving firearms, the relevance of spent cartridges discovered at a crime scene may be contested during examination. Thus, a suspect may state, for example, that a cartridges was fired long before or after the commission of the alleged offense [1, 2]. In such situations, estimating the time since discharge could be a valuable piece of evidence toward reaching a conclusion [3].

A promising approach towards achieving this is to monitor (over time) the decrease of selected volatile organic substances formed during the discharge, i.e. gunshot residue (GSR) [3-9]. GSR is a chemical trace that is very heterogeneous in composition, consisting mainly of metallic micro-particles, unburnt or partially burnt flakes of smokeless powder, and explosion products [10-13]. The latter are partly composed by volatile organic compounds (such as benzonitrile, naphthalene, and pyrene), which condense on the cartridge walls and disappear over time through evaporation and diffusion phenomena [14]. Solid phase micro-extraction (SPME) coupled to gas chromatography was previously proposed as a sampling method to analyse these residual products, while naphthalene and an unidentified decomposition product of nitrocellulose (referred as "TEA2") were suggested as main target analytes [4-7]. However, unsatisfactory results were obtained in the analysis of small-calibre firearms and cartridges. As reported in the first part of this article, in addition to the natural complexity of the specimens themselves, this problem could also be attributed to three additional factors: the extraction technique, the targeted compounds and the interpretation models. Thus, the present series of paper aimed to further investigate these aspects.

In the previous part, analytical questions were mainly addressed. In this regard, an alternative multi-residue approach based on headspace sorptive extraction (HSSE) was optimised, in order to establish a more comprehensive overview of the ageing processes in spent cartridges in comparison to that allowed by previous single-analyte methods. The specific objective of this second part was, therefore, to investigate efficient solutions to comprehensively interpret the GSR profiles in a dating perspective and evaluate the actual potential of providing helpful information on time since discharge in real cases. In this regard, the implementation of multivariate statistical methods has been explored instead of current one-compound-at-time approaches, in an attempt to implement all sources of information about time since discharge linked to the single compounds into a unique estimation model. Specifically, numerous 9 mm Parabellum cartridges from the brand Geco were analysed at distinct times since discharge and ageing of the single GSR compounds was first studied using heteroscedastic non-linear regression (NLR) analysis. Then, different pre-treatments and multivariate regression methods were evaluated and compared in their ability of estimating time

4

since discharge from analytical data. Regression methods included partial least squares (PLS), multivariate adaptive regression splines (MARS), artificial neural networks (NN), random forests (RF), *k*-nearest neighbors (KNN) and support vector machines (SVM). Concerning pre-treatments, different strategies of normalisation, transformation and scaling methods were applied. These included, for example, total sum normalisation (TSN), probabilistic quotient normalisation (PQN) and pairwise log-ratios normalisation (PLR) as normalisation approaches, as well as logarithm (Log), square root (Sqrt), fourth root (4thr) and inverse (Inv) as transformations. The best multivariate models were selected and compared to compound-specific estimates obtained by NLR. The impact of varying the cartridge storage conditions (i.e. cartridge position, temperature and humidity) was also investigated, and robustness of retained multivariate models to changing in the latter was tested. As it was considered important to evaluate dating possibilities in conditions close to real cases, a casework perspective was generally adopted. Hence, training of models was performed using a limited number of reference cartridges and their validation was carried out on data obtained from different runs.

2. Material and methods

2.1. Chemicals

Five deuterated substances were used as internal standards, i.e. naphthalene-d8 (99%), biphenyl-d10 (98%), acenaphthene-d10 (99%), phenanthrene-d10 (98%) and pyrene-d10 (98%). All deuterated standards were purchased from Cambridge Isotope Laboratories (Tewksbury, USA). Stock solutions were prepared at a concentration of 1 mg mL^{-1} in dichloromethane (Sigma-Aldrich, Buchs, Switzerland). Finally, a mixed solution of 0.6 mg L^{-1} (per analyte) was prepared by successive dilution in methanol (Sigma-Aldrich) to serve as working solution.

2.2. Ammunition and test shooting

All experiments were carried out on Geco 9 mm Parabellum ammunition (RUAG Ammotec, Thun, Switzerland) purchased in 2014. All cartridges belonged to the same production batch. Test shootings were carried out using a SIG Sauer P220 semi-automatic pistol in an indoor shooting range. Cartridges were fired by loading them into the magazine individually. The weapon was never greased during the experimental period.

After discharge, the cartridges were transported to the laboratory and aged in a conditioned oven maintained at pre-set temperature and humidity conditions (Table 1). For ageing in the vertical position, the cartridge were placed in a vial rack inside the oven, while ageing in the horizontal

position was carried out by laying the cartridges on aluminium rails (purchased from a local hardware store). Ageing at 25 °C temperature, 75 % relative humidity (RH) and horizontal position were defined as "standard" conditions.

2.3. Sampling and extraction protocol

Cartridges were sampled, extracted and analysed using the analytical method optimised in the previous part of this work. Different specimens were aged for 1, 3, 5, 7, 24, 31, 48 and 72 h. After ageing, the cartridges were collected and placed into 20 mL HSSE-dedicated crimp glass vials. An extracted internal standard approach was adopted. 5 µL of the internal standard solutions at 0.6 mg L^{-1} were added to the bottom of the vials corresponding to a spiking amount of 3 ng per deuterated molecule. Then, special glass inserts were introduced to the headspace in order to accommodate the stir bars. Stir bars used for extraction were 1 cm in length with a 0.5 mm coating. These were thermo-conditioned before use in a Gerstel tube conditioner (TC) based on the protocol suggested by the manufacturer (30 min at room temperature followed by 90 min at 300 °C and approximately 60 min for cooling down). The vials were finally sealed with 20 mm crimp caps equipped with 3.0 mm PTFE/silicon septa. Cartridges were extracted in a laboratory oven at 70 °C for 24 h. For analysis, vials were re-opened and stir bars were transferred into pre-conditioned desorption tubes. The tubes were capped with special transportation adapters and placed on the GC tray for TD-GC-MS analysis. All materials were supplied by Gerstel (Sursee, Switzerland).

2.4. Thermal desorption – gas chromatography – mass spectrometry (TD-GC-MS) analysis

Stir bars were thermally desorbed in a Gerstel thermal desorption unit (TDU) on-line connected to a Gerstel CIS-4 programmed temperature vaporizing (PTV) injector. These devices were mounted on an Agilent 7890A gas chromatograph coupled to an Agilent 5975C mass selective detector (Agilent Technologies, Basel, Switzerland).

Desorption of stir bars were carried out in splitless mode with a flow of 60 mL min^{-1}. The desorption ramp was programmed as follows: 25°C for 0.5 min, ramped to 300 °C at 720 °C min^{-1} and held at this temperature for 5 min. The transfer line temperature between TDU and CIS-4 was set to 300 °C. Liners for CIS-4 were packed with quartz wool, provided by Gerstel. The cryo-focusing temperature was set to -50 °C. The PTV injection ramp was programmed as follows: -50 °C for 0.1 min, ramped to 300 °C at 720 °C min^{-1}, and held at this temperature for 1.37 min (total injection time: 2 min). Splitless mode was used during injection. The injector was then switch to split mode and a supplementary conditioning time of 7 min at 300 °C was programmed.

GC separation was performed on a HP-5MS column (30 m x 0.25 mm x 0.25 µm) from Agilent. The carrier gas was helium, and column flow was maintained at 1.2 mL min^{-1}. The oven ramp was programmed as follows: 40 °C for 2 min, ramped to 100 °C at 20 °C min^{-1}, ramped to 155 °C at 5 °C min^{-1}, and finally ramped to 250 °C at 12 °C min^{-1} (total chromatographic separation time of 23.9 min). Post-run temperature was set to 320 °C and held for 5 min with a column flow of 3 mL min^{-1}. The transfer line between the column and the MS was set to 280 °C. Ionisation was carried out through electron impact (EI). Selected ion monitoring (SIM) mode was employed for MS analysis. Solvent delay was set to 5.60 min. MS source and quadrupole temperatures were 230 °C and 150 °C, respectively. Detected analytes and respective monitored ions are reported in Table S1 in Electronic Supporting Material (ESM).

2.5. Tested regression approaches

Univariate regression analysis was used to study the ageing behaviours of single extracted analytes as well as to obtain compound-based estimates of the time since discharge. In this regard, inference on time since discharge through the interpretation of ageing profiles of single compounds has been suggested in the literature, but no statistical tool has ever specifically been proposed for (definitive) decision making. Since this, heteroscedastic non-linear regression analysis (NLR) has been adopted in this study as reference univariate approach as it previously proved to be an adequate statistical technique to model ageing profiles of volatile compounds in spent cases [3, 15]. For each analyte, the peak area corresponding to the target ion was extracted. Values were then subject to a two-step normalisation: first, their peak areas were divided by a corresponding deuterated standard (Table S1); second, they were divided by the mean of the values obtained at t = 0 h in order to globally scale down the start of the ageing of all analytes to a nominal value of 1. For fitting, the following equation was adopted:

$$S = \exp(-C \cdot \sqrt{t}) \qquad\qquad (1)$$

where S is the expected peak area mean, t is the time since discharge and C is the characteristic constant of the curve, which is related to the decrease rate. Heteroscedasticity was accounted for using a power-of-the-mean function. Other pre-treatments and fitting procedures were initially tested on data. Particularly, the fitting equation suggested in [15] (which includes two supplementary size constants) was applied on log-transformed data. However, this approach was ruled out because of fitting problems encountered for most of the resampled datasets. Moreover, estimates were missing for values falling beyond the minimal value of the regression curve. Scaling

7

the data using the mean peak area of the respective compounds at t = 0 h (which in turn allowed for eliminating size constants) was found to be a reliable strategy in solving these problems (applicable to real cases).

Concerning multivariate modelling, six regression methods were tested on data, i.e., partial least squares (PLS), multivariate adaptive regression splines (MARS), artificial neural networks (NN), random forests (RF), *k*-nearest neighbors (KNN) and support vector machines (SVM). These were applied after data pre-treatment (see following section) in reverse regression mode [16]. Therefore, the outcome variable was the time since discharge and the predictors were the various volatile organic GSR compounds. Theoretical details of the assessed methods can be found in Table S2 in ESM.

2.6. Pre-treatments before multivariate fitting

The impact of several combinations of pre-treatment methods were tested on the regression performances of multivariate models. These methods included different data normalisation, transformation and scaling approaches. For normalisation, methods for the correction of both, the pure analytical error and the size effect, were tested [17]. The application sequence of the various pre-treatment steps was kept constant in accordance with the following order: pure analytical variation correction, size effect correction, transformation, centring and scaling (Fig. 1).

Dividing the peak areas of the target compounds by those of the corresponding co-extracted deuterated standards was investigated as a method for the correction of pure analytical error. Approaches tested for the correction of size effects included total sum normalisation (TSN), probabilistic quotient normalisation (PQN) and pairwise log-ratios normalisation (PLR). For TSN, the extracted peak areas of each target analyte were divided by their total sum within the same chromatogram, while for PQN, they were divided by a dilution quotient. This quotient was equal to the median value of ratios obtained by dividing the peak areas of single analytes in a selected chromatogram by their respective peak areas in a reference chromatogram acquired at t = 1 h (from the training dataset) [17, 18]. For PLR, all possible pairwise ratios between the peak areas of the different target analytes were calculated and then, their logarithm was determined in order to remove any variance asymmetry caused by inversion of the ratio order [17, 19]. This allowed for the creation of 210 log-ratios per chromatogram which led to excessively long training times for some regression methods, especially RF and NN. Thus, the number of log-ratios was reduced by selecting only those which presented an estimated Spearman correlation coefficient greater than 0.7 with the outcome variable (i.e., the time since discharge) across all the analyses of the training

dataset. This allowed for reducing the number of log-ratios to about 60, depending on the dataset used for estimating correlation coefficients.

Applied transformations included logarithm (Log), square root (Sqrt), fourth root (4thr) and inverse transformation (Inv) [20, 21]. After PLR normalisation, only the logarithm transformation was used, for the reason explained above. Finally, scaling was performed by dividing the peak areas of each target analyte by their respective standard deviation (as estimated by the training dataset). Mean-centring of the variables was always performed before training, independently from scaling.

2.7. Model training and validation

Out of the 29 substances targeted with the developed method, 21 were chosen as predictors for training (Table S1). The three unaltered smokeless powder compounds (i.e., ethylcentralite, dibutyl phthalate and diphenylamine) were discarded because of low repeatability in their detected amounts [14]. Moreover, five explosion products (i.e., o- and m-tolunitrile, 1,3- and 1,4-dicyanobenzene, and isoquinoline) were also discarded due to co-elution issues compromising their detection and relative quantification (see part 1).

For all the different regression models applied in this work (univariate and multivariate), regression parameters were determined based on a training dataset by resampling using bootstrap (100 iterations) [22]. The datasets used for training (Table 1) was dependent on the experiment. Particularly, the two datasets acquired at "standard" conditions (i.e., STD1 and STD2) were used for the initial selection of the best models and the successive study on datable ranges. The remainders were used to test the effect of storage conditions on both, compound ageing and regression performance. All multivariate regression methods used involved at least one tuning parameter (Table S2), which were optimised during the resampling procedure.

Estimation of the predictive performance of fitted models (i.e., validation) was carried out using different approaches depending on the experiment:

- Between-runs validation: in this approach, validation of a model was carried out mainly by predicting analyses from an alternative, but equivalent, dataset. Since only the ageing experiment at "standard" conditions was replicated, this approach was essentially used on STD datasets. Thus, models were trained on one of the two dataset and tested using the other. As STD datasets were composed of analyses acquired several months apart, between-runs validation also simulated the worst case scenario for which reference and questioned cartridges were analysed during different GC-MS runs.

9

- Within-run validation: for comparison purposes and validation of models acquired on non-replicated datasets (i.e., ageing experiments under different conditions), validation was carried out using a mixed leave-one-out cross-validation (LOOCV) and bootstrap approach. Thus, for a given dataset, each observation was recursively left out and a model was trained on the remainders using bootstrap. The excluded observations were then used for validation. This also represented the best case scenario where the reference and questioned cartridges were analysed during the same GC-MS run.

Initial model selection was carried out on the basis of the global estimation performance on the studied ageing interval. As a metric for global estimation performance, root-mean-square error (RMSE) was exploited [22]. An average RMSE was thus calculated as a result of fitting two equivalent models using between-runs validation on STD datasets (one per STD dataset). RMSE is a measure of the mean estimation error on the experimental interval. However, it does not form a specific idea about the actual estimation performance at the different times since discharge. Thus, further evaluation of the local performance was carried out by inspecting the calibration plots of the models (the plots of the estimated versus the actual times since discharge) and their inaccuracy plots (the plots of the inaccuracy versus the actual times since discharge). Particularly, inaccuracy was defined as the average of the absolute errors at a specific time since discharge [23]:

$$I_t = \frac{\sum|\hat{t}-t|}{n_t} \tag{2}$$

where I_t is the inaccuracy at the time t, $\hat{t}$ and t the estimated and actual times since discharge respectively, and n_t the number of observations. Given that inaccuracy was found to follow trends over time for most evaluated models, quadratic linear regression was applied to extrapolate mean values over all the studied ageing interval.

2.8. Software and computing

Chromatograms were processed using the Enhanced Data Analysis software provided by Agilent. Regression methods were fitted using R statistical computing software and available packages (see Table S2 in ESM for details on used packages).

3. Results and discussion

3.1. Ageing of volatile GSR compounds

Several 9 mm Parabellum cartridges from the brand Geco were analysed at various post-discharge intervals using the HSSE-based method developed in part 1, after ageing at different storage conditions. From the analysis of chromatograms obtained at different times, it is already evident that most of the detected compounds decreased noticeably over time independently from the conditions applied (see examples in Figure S1 in Electronic Supporting Material, ESM). Heteroscedastic NLR was used in order to model the ageing of the various target analytes (Fig. 2a-d). It was thus noticed that, while the decrease of each compound was generally rapid in the first hours since discharge, the actual rates seemed to be largely dependent on compound volatility. In this regard, a negative correlation was usually observed between the C coefficients of the inferred ageing curves (which are strictly correlated to the compounds decrease rates) and the compounds boiling points (see also Fig. S2). Consequently, amounts of the most volatile compounds dropped significantly within the first hour and quickly levelled-off, unlike those of the less volatile ones when the levelling-off occurred significantly later. This finding supports the hypothesis that different compounds could actually provide complementary evidence about the time since discharge.

3.2. Selection of best multivariate inferential models

In order to exploit the complementary ageing information provided by the different GSR analytes, multivariate methods were applied. In this regard, different regression and pre-treatment strategies were tested. Particularly, six multivariate regression methods were applied (i.e., PLS, MARS, NN, RF, KNN and SVM) after pre-treatment of data using different combinations of normalisation, transformation and scaling approaches (Fig. 1). This led to the fitting and evaluation of 384 different models (refer to ESM for details).

Screening of the multivariate models was carried out on cartridges aged at "standard" conditions on the basis of their root-mean-square error (RMSEs, i.e., the respective mean errors on the estimated times since discharge over the global experimental interval). Overall, between-runs RMSEs for all models ranged from 8.8 to 27.8 h with a median value of 13.0 h (Fig. S3). The medians for the specific regression methods did not significantly vary from this value, proving a certain coherence between the different regression techniques. Despite that, specific combinations of pre-treatments and regression methods were nonetheless found to perform significantly better than others. Calibration plots for the models with lower RMSEs were inspected. Many of them (especially those

11

based on KNN and NN) seemed to demonstrate moderate symptoms of overfitting, which led to highly variable estimation accuracies over the experimental time range. That is, some data points were perfectly estimated, while others were affected by large estimation errors. In this regard, RF and PLS preceded by pairwise log-ratio normalisation (PLR) proved to be the most consistent (and efficient) combinations, as time-since-discharge estimates were more condensed around their local means in comparison to the other models (Fig. 3a-d). Observed between-runs RMSEs were 10.1 and 10.6 h, respectively. Within-run RMSEs for these models, 9.8 and 12.6 h respectively, were not significantly different from between-runs RMSE values, indeed proving a good robustness with respect to data obtained from same or different runs.

3.3.   Accuracy in time-since-discharge estimation

The RF and PLS models using PLR-normalised data were selected as optimal and exploited for evaluating the actual accuracy in time-since-discharge estimation of the tested Geco 9 mm cartridges as a function of the their age. Table 2 reports inaccuracy (i.e., the average magnitude of absolute error) obtained with these models over the experimental range (see also Fig. 3a-d). Using the RF model, inaccuracy values of 2.5, 6.5, 13.7 and 23.2 h were estimated for 5h-, 24h-, 48h- and 72h-old cartridges, which correspond to percentages of about 50, 27, 28 and 32 % in relation to the actual times since discharge, respectively. The PLS model performed worst at the shortest times since discharge, with an inaccuracy estimated to about 6.5 h across the entire 0 – 24 h range. For the sake of illustration, this corresponded to 126 % of relative error at 5 h since discharge.

The larger average inaccuracy of the PLS model in comparison to the RF one for freshly fired cartridges was surely due to a wider spread of estimation errors obtained on the cartridges actually analysed in the first 0 to 24 h since discharge. This could in turn be attributed to statistical reasons, namely the difficulty of the PLS algorithm in correctly taking into account the considerable non-linear nature of the data in this initial ageing stage [22]. Despite that, a slightly lower average inaccuracy was observed for the PLS model in comparison to the RF one for > 48h-old cartridges. In fact, inaccuracy values of 10.4 h and 18.9 h were estimated for 48h- and 72h-old cartridges, which correspond to 22 and 26 % of the actual times since discharge, respectively (Table 2). In consequence, the two models appeared to be complementary to some extent.

Inaccuracy values for the best models (and particularly RF) were considered acceptable for use in potential dating applications, especially in view of the large random variations which cannot be reduced by internal normalisation (see part 1). Nonetheless, it is noteworthy to mention that the local inaccuracy on the time-since-discharge estimates increased as a function of the actual

cartridge age independently of the model applied. This trend led to systematic negative biases in estimates for cartridges older than 48 h for all the models, which are most probably due to a general level-off in the amount of every volatile GSR compound and the subsequent instalment of a steady state in all the ageing profiles (Fig. 2a-d). Thus, actual maximal post-discharge intervals for the tested Geco 9 mm Parabellum cartridges estimable through multivariate analysis of GSR profiles were determined to be around 48 h, at least through the applied experimental design. Beyond this limit, it is strongly probable that the relative differences in GSR profiles would not be sufficiently significant to allow a reliable time-since-discharge estimation. Improving estimation accuracy for > 24h-old cartridges was attempted by iteratively omitting data from 0h- to 7h-old cartridges during training, in order to evidence any statistical issue related to model estimation. However, no significant improvements were yielded, supporting the hypothesis that the loss in dating accuracy beyond 48 h depended on the lack of significant differences between chemical profiles.

3.4. Differentiation of cartridges at different ageing stages

Even if time-since-discharge estimation attempted on > 48h-old cartridges were found to be potentially affected by significant biases, it is noteworthy that GSR profiles obtained on cartridges aged over this period still presented differentiable distributions of estimates in comparison to those obtained on more recently fired cartridges. For the sake of illustration, Figures 4a-b show the distributions of the time-since-discharge estimates obtained through the RF and PLS models as a function of three actual cartridge ages (i.e., 5 h, 24 h and 72 h). As it can be noted, groups were generally separated. Particularly, 5h-old cartridges could easily be differentiated from 24h- and 72h-old ones, independently from the multivariate model considered. A slight overlap between results obtained on cartridges aged for 24 h and 72 h was observed for the RF model, while they were clearly differentiable for the PLS model. Analysis of raw PLR-normalised data by principal components analysis (PCA), which is an unsupervised modelling method contrary to RF and PLS, gave similar findings (Fig. 5). In fact, good separation between ages was achieved on PC2, after which further separation within age classes was possible by additionally exploiting PC1. Consequently, despite the lost in estimation accuracy as a function of ageing, analyses carried out on > 48h-old cartridges seemed to still convey sufficiently divergent chemical information to be discernible from more recently fired cartridges. In this regard, it is remarkable to note again the complementarity of the two tested chemometric techniques. Indeed, separation between cartridges at different ageing stages was slightly better with PLS than RF.

3.5. Comparison to univariate regression models

For the purpose of comparison, estimation of the time since discharge of the analysed 9 mm cartridges was also attempted using information provided by the single compounds through the use of more traditional univariate regression methods. In this regard, heteroscedastic NLR was again applied and used as reference univariate estimation approach.

Between-runs RMSEs for all univariate NLR models trained on the different GSR analytes ranged from 20.3 to 26.2 h (corresponding to pyrene and indene, respectively), while coefficients of determination (R^2) never surpassed 0.588 (the highest being acenaphthylene; see Table S4 for detailed performance characteristics). These metrics were considerably worse than those obtained from most multivariate models, consequently also denoting worse global performance in estimating times since discharge. Inaccuracy profiles were studied in order to investigate the actual estimation accuracy over time. Inaccuracy values for < 10h-old cartridges attained with the two univariate models considered the best in terms of overall fitting efficiency (i.e., acenaphthylene and fluorene) were in the same order of magnitude of those of the RF model. Despite that, values obtained on cartridges older than 10 h was significantly higher than those of both the PLS and RF models (Table 2 and Fig. 3e-h). Particularly, inaccuracy strongly increased as a function of the actual post-discharge interval after 10 h of ageing and this trend led to systematic negative biases in estimates. Inspection of the results for the other compounds generally gave very similar results. Thus, no analyte was able to provide more accurate time-since-discharge estimates for longer (or different) post-discharge intervals, and maximal intervals which could be estimated through compound-specific regression curves were determined to be around 10 h (vs. 48 h for multivariate approaches). The distribution of time-since-discharge estimates obtained on 5h-, 24h- and 72h-old cartridges additionally showed stronger overlapping (Fig. 4c-d), which significantly decreased the potential to reliably discriminate cartridges that belonged to different ageing groups. Solely exploiting information provided by the single GSR analytes therefore proved to be a less helpful approach for gaining evidence on the time since discharge of spent cases in comparison to multivariate ones. Especially, univariate models seemed to be more sensitive to level-off in the amount of GSR compounds.

3.6. Importance of compounds on time-since-discharge estimation

Simultaneously analysing more compounds at time showed to provide more helpful information than a univariate approach. Nonetheless, some compounds might be more informative than others in characterising ageing of cartridges, and it could be of interest to determine them. In this regard, the

14

RF algorithm involves an internal scoring method which automatically rank predictors as a function of their relative impact on the fitting procedure. Results of this method has thus been used also to assess the importance of each ratio obtained through PLR normalisation for characterizing the ageing of the tested Geco 9 mm cartridges. Of the 210 ratios obtained from all the possible pairwise combinations between retained analytes, a total of 8 ratios generated from different combinations of 8 target compounds were found to be particularly significant in every trained RF model. These were (in order of decreasing importance):

(1) acenaphtylene / phenanthrene,

(2) acenaphtylene / anthracene,

(3) acenaphtylene / acenaphtene,

(4) fluorene / anthracene,

(5) acenaphtylene / 1-naphthalenecarbonitrile,

(6) indole / fluoranthene,

(7) fluorene / phenanthrene,

(8) acenaphtylene / fluoranthene.

On a closer inspection of their ageing profiles, it could be observed that all these ratios decreased over time similarly to the signals of the single analytes (see examples in Fig. 2e-f). Nonetheless, their decrease rates were often slower and yielded more linear ageing curves. This could partially explain the better regression performances, on average, obtained after PLR normalisation in comparison to other pre-treatment approaches.

Concerning the single compounds, acenaphthylene appeared in 5 of 8 most influent ratios. Indole, phenanthrene, anthracene and fluoranthene were each involved in 2 ratios, while acenaphthene, 1-napthalenecarbonitrile and fluorene were each in only 1 ratio. It is interesting that all 8 substances are characterised by boiling points higher than 254 °C (the lowest being indole). Moreover, re-fitting the PLR-RF model using these compounds exclusively, yielded a mean between-run RMSE of 10.3 h, which was not significantly different from that obtained using the full-variables set (RMSE = 10.1 h). Therefore, even if different compounds provided complementary evidence on time since discharge, most of the pertinent information to estimate the time since discharge seemed to be gathered by the least volatile molecules. This may be explained by their slower evaporation and diffusion, allowing for a more comprehensive characterisation of cartridge ageing. Even if strongly exploited in the literature, naphthalene was not highlighted as particularly significant for characterising the ageing of the chosen 9 mm cartridges.

15

3.7. Effect of storage conditions on ageing profiles

In order to test the effect of varying storage conditions on ageing profiles of the single GSR compounds, groups of cartridges were aged at different storage conditions. The latter were obtained by varying three factors: storage temperature, relative humidity and cartridge position. Relative effects on decrease rates were investigated through ANOVA (Table 3).

It was noticed that the temperature had a very significant effect on the decrease of all studied compounds, and also that this was generally positively correlated to the decrease rates. Indeed, at low temperatures, disappearance of compounds was slower and vice-versa (Fig. 6c). Previous works reported similar trends for the decrease of volatile organic compounds found in barrels and other types of cartridges [4, 5]. Though the effects of relative humidity and cartridge position were never previously reported, they were also found to have significant effects on the ageing of most compounds under investigation. Particularly, laying the cartridge horizontally rather than placing it vertically, globally, accelerated the decrease rates (Fig. 6a). This may be due to an increase of the evaporation surface when cartridges were placed horizontally. Such an observation is of particular importance given that cartridges are more commonly encountered in a lying position throughout actual caseworks, and to find them standing is rare. Concerning relative humidity, its effect seemed significant only for the least volatile compounds, but not for the most volatile ones. Furthermore, the correlation between decrease rates and relative humidity in these cases was non-linear, where maximal decrease rates reached approximately 75 % RH for all the influenced compounds (Fig. 6b).

3.8. Robustness of inferential models to storage conditions

Given that ageing profiles of all most GSR compounds were found to be affected by the tested storage conditions, a major concern was to determine if differences in estimation accuracy could equally be encountered between models which were trained using the same regression method but on cartridges aged at distinct conditions. This was tested by refitting the RF model on datasets from cartridges aged at distinct ageing circumstances, and by successively determining estimation accuracy on left-out observations.

No significant shifts in estimation accuracy were observed between all the trained models (Table 4 and Fig. 7). Obtained within-run RMSEs were indeed comparable in all the situations, and ranged from 8.6 to 13.4 h. The inaccuracy profiles were also very similar, proving a good robustness of the selected model at different storage conditions. In this regard, it should be noted that PLR normalisation could have had a significant role in preventing shifts in estimation accuracy, as the decrease rates of most ratios were found to be less affected by influential factors than their

respective composing molecules (Fig. 6d-f). The same tests were carried out on univariate NLR, but highly variable results were observed depending on the storage condition investigated (Fig. 7). For example, using NLR trained on acenaphthylene data, RMSEs ranged from 11.8 h for cartridges stored in vertical position to 41.6 h for those stored at 10 °C.

3.9. Differences between reference and questioned cartridges

Whilst the above results are noteworthy, the main problem in real forensic cases lies in fact that the exact storage conditions encountered by a specific questioned cartridge on a crime scene are rarely known. Thus, differences between the storage conditions applied to reference materials and those actually persisting during the ageing of the questioned cartridges are very likely, making any statistical model trained on reference data not properly pertinent for time-since-discharge estimation in a particular case. The consequences of these situations were investigated by refitting the RF model on datasets from cartridges aged at distinct ageing circumstances, and by successively determining estimation accuracy on datasets from cartridges aged at different conditions.

Suboptimal accuracies were noted in many cases, especially when differences in storage temperature were introduced. Figure 8 and Table 4 show some examples. Thus, when time-since-discharge estimation was attempted on cartridges aged at 10 °C using a model trained with data acquired at 25 °C, outcomes were affected by strong inaccuracies which globally under-estimated the true values. In contrast, predicted times since discharge for cartridges aged at 40 °C (15 °C above the training dataset) were mostly over-estimated (mainly within the first hours of ageing). It was anyway noted that consequences were more limited when differences in cartridge position and relative humidity were introduced.

ANOVA of the RMSEs estimated on all the validation datasets further confirmed these observations. Thus, very limited evidence to supporting the hypothesis that differences in cartridge position between the training and validation datasets are significant on the estimation accuracy of the models were provided (p-value = 0.111). On the contrary, results moderately supported the hypothesis that differences in the relative humidity can have an impact (p-value = 0.021), and strongly the hypothesis that differences in storage temperature have a significant effect (p-value < 0.001).

4. General discussion and conclusion

In this work, the potential of providing helpful evidence about the time since discharge of 9 mm handgun cartridges through the analysis of their headspace has been thoroughly investigated using

highly efficient analytical and statistical techniques. In the first part, a high-capacity extraction method based on HSSE has been optimised to analyse spent cartridges' headspaces in the most exhaustive way. Thus, this second part focused on the application of chemometric analysis to comprehensively interpret GSR profiles in a dating perspective and evaluate the actual potential of providing helpful information on time since discharge in real cases.

Ageing kinetics of 29 GSR compounds in Geco 9 mm Parabellum cartridges was studied. Most of them demonstrated a significant decrease over time, with a generally rapid rate in the first hours since discharge. Despite this, the actual decrease rates were dependent on the compounds respective boiling points, indicating that different compounds could actually provide complementary evidence about the time since discharge. In order to combine all sources of information into a single time-since-discharge estimation model, several multivariate regression techniques were evaluated and combined with different pre-treatment approaches (a total of 384 models were tested). Pairwise log-ratios normalisation (PLR) in combination to RF and PLS regression presented particularly good performance in comparison to other fitted models. Particularly, the accuracy of the obtained outcomes demonstrated potential for estimating the time since discharge in the tested cartridges up to 48 h of ageing or, at least, to differentiate recently fired from older cartridges (e.g., less than 5 h compared to more than 48 h), under known storage conditions.

Whilst these achievements might seem relatively modest at first sight, it should be noted that most firearm-related incidents are discovered and reported in the first hours after commission. Thus, they rather support the hypothesis that useful information on time since discharge might actually be extracted from analysis of the volatile fraction of GSR, as well as the hypothesis that this type of assessment could be helpful in a casework perspective. In fact, a timely analysis providing the aforementioned performance might still help, for example, in identifying those specimens which are most likely to have been fired during the alleged offence (in order to recognise imported evidences and/or focus only on pertinent specimens for further examination), in reconstructing the course of events through situating a shooting in time and/or in discriminating alternative hypotheses brought forth by the prosecution and the defence (such as those reported in [24]).

Comparison of results obtained by multivariate modelling with those obtained by compound-related models fitted with NLR additionally showed that simultaneously interpreting residuals levels of several analytes at the same time allowed providing enhanced selectivity in time-since-discharge estimation in comparison to interpreting single compounds. Hence, approaches combining information from more analytes (such as pairwise normalisation and/or multivariate regression) should be preferred for application in dating. An efficient implementation of the suggested

methodology in casework, however, needs to account for additional questions. Amongst them, we firstly find the effects of factors influencing GSR compounds' ageing kinetics. In this work, cartridge position, temperature and relative humidity have been studied. Thus, while changes in cartridge position and relative humidity were only found to be marginally significant, differences in temperature were prone to introduce more significant errors on the time-since-discharge estimates, proving that estimation accuracy is actually dependent on the quality of the reference material (and, especially, on its degree of consistency in comparison to the questioned cartridges). From a practical point of view, this has two important consequences. Firstly, that a case-by-case approach needs to be adopted for collecting and interpreting data in real casework. Using general reference materials, storage conditions and/or interpretative models does not seem to be a reliable strategy at this stage of research. The same storage conditions as those experienced by the questioned cartridges should be applied to reference material. The type of cartridges used as reference material should also be consistent, as previous research showed a dependence between GSR composition and ammunition [14]. Secondly, that knowledge about the environmental conditions at the crime scene before the questioned cartridges have been sampled should ideally be collected in order to correctly reproduce them in laboratory. On this subject, it is acknowledged that obtaining information on storage conditions experienced by questioned cartridges during their ageing can be a major difficulty in real cases, as the latter are hardly inferable during crime scene investigation. Nonetheless, meteorological data and/or post-crime measurements can be exploited, similarly to what systematically done in forensic entomology [25]. This experimental approach could thus be implemented or adapted to fit with the particular requirements of the forensic investigation of firearm-related crimes. A climatic chamber can then be used to simulate specific storage conditions on reference material.

Another question to be accounted for is the suitability of the analytical protocol. In this regard, it should be noted that a multivariate assessment of the information provided by GSR profiles would basically require a multi-residue analytical method. High-capacity extraction technique as that used in this work are particularly suitable to implement such a method as they allow for high recoveries for a larger range of different compounds, and HSSE was found to be robust and efficient. However, it is acknowledged that it is also a relatively invasive technique, as it does not allow analysing an enclosed specimen without re-opening the collection vial. Hence, HSSE stir-bars should be placed into sampling vials directly on the crime scene in order to obtain reliable results. This might not be feasible in some situations. Adapting the extraction approach for a more effective use in current cartridge-sampling protocols might thus represent an interesting research topic for

future works. In this regard, implementation of HSSE can be improved by using more complex vial-capping solutions. For instance, a "double cap" involving an intermediary space to accommodate HSSE stir-bars can be developed. Otherwise, suitability of SPME for multi-residue extraction can be further investigated and a similar multi-residue method as that suggested in this work developed. Novel and less invasive techniques can also be tested such as cold-fibre SPME (cfSPME), which combines both the advantage of SPME (low invasiveness thanks to a needle-based design) and HSSE (high capacity) [26, 27].

Providing evidence on the time since discharge of handgun cartridges would be a valuable tool for law enforcement authorities. Results reported in this work showed that employment of highly efficient extraction techniques and advanced chemometrical tools for interpreting data allowed obtaining enhanced results in comparison to previously suggested approaches, as well as helpful evidence in simulated conditions for the tested ammunition. Thus, they are promising towards future applications in real casework. Further research is strongly encouraged, and should mainly focus on evaluating the possibility to generalise the findings to other types of cartridges and adapting the solutions developed in this work for current crime-scene operating procedures.

Acknowledgements

References

[1] C. Andersson, J. Andrasko, A novel application of time since the latest discharge of a shotgun in a suspect murder, Journal of forensic sciences, 44 (1999) 211.

[2] R.R. Bridgemon, Cartridge case aging, AFTE journal, 18 (1986) 49.

[3] M. Gallidabino, F.S. Romolo, K. Bylenga, C. Weyermann, Development of a novel headspace sorptive extraction method to study the aging of volatile compounds in spent handgun cartridges, Analytical chemistry, 86 (2014) 4471.

[4] J. Andrasko, T. Norberg, S. Stahling, Time since discharge of shotguns, Journal of Forensic Sciences, 43 (1998) 1005.

[5] J. Andrasko, S. Stahling, Time since discharge of spent cartridges, Journal of forensic sciences, 44 (1999) 487.

[6] J. Andrasko, S. Stahling, Time since discharge of rifles, Journal of forensic sciences, 45 (2000) 1250.

[7] J. Andrasko, S. Stahling, Time since discharge of pistols and revolvers, Journal of Forensic Sciences, 48 (2003) 307.

[8] J.D. Wilson, J.D. Tebow, K.W. Moline, Time since discharge of shotgun shells, Journal of Forensic Sciences, 48 (2003) 1298.

[9] C. Weyermann, V. Belaud, F. Riva, F.S. Romolo, Analysis of organic volatile residues in 9mm spent cartridges, Forensic science international, 186 (2009) 29.

[10] K.H. Chang, P.T. Jayaprakash, C.H. Yew, A.F.L. Abdullah, Gunshot residue analysis and its evidential values: a review, Australian journal of forensic sciences, 45 (2013) 3.

[11] O. Dalby, D. Butler, J.W. Birkett, Analysis of gunshot residue and associated materials - a review, Journal of forensic sciences, 55 (2010) 924.

21

[12] H.-H. Meng, B. Caddy, Gunshot residue analysis - a review, Journal of forensic sciences, 42 (1997) 553.

[13] F.S. Romolo, P. Margot, Identification of gunshot residue: a critical review, Forensic science international, 119 (2001) 195.

[14] M. Gallidabino, F.S. Romolo, C. Weyermann, Characterization of volatile organic gunshot residues in fired handgun cartridges by headspace sorptive extraction, Analytical and Bioanalytical Chemimstry, 407 (2015) 7123.

[15] R.-A. Doong, S.-M. Chang, Y.-C. Sun, Solid-phase microextraction for determining the distribution of sixteen US environmental protection agency polycyclic aromatic hydrocarbons in water samples, Journal of Chromatography A, 879 (2000) 177.

[16] J.N. Miller, J.C. Miller, Statistics and chemometrics for analytical chemistry, 5th ed., Pearson Education Limited, Harlow, UK, 2005.

[17] N. Barco-Bonilla, J.L. Martinez-Vidal, A. Garrido-Frenich, R. Romero-Gonzalez, Comparison of ultrasonic and pressurized liquid extraction for the analysis of polycyclic aromatic compounds in soil samples by gas chromatography coupled to tandem mass spectrometry, Talanta, 78 (2009) 156.

[18] R. Shanahan, Corrosion of new, fired, 0.22-caliber, long-rifle brass cartridge cases buried in soil, Journal of forensic sciences, 22 (1977) 614.

[19] W. Kerkhoff, D. Hazard, M. Lopatka, Atmospheric corrosion of bullets and cartridge cases as an indicator of time since discharge, AFTE Journal, 46 (2014) 51.

[20] K. Andersson, E. Lock, K. Jalava, H. Huizer, S. Jonson, E. Kaa, A. Lopes, A. Poortman-Van der Merr, E. Sippola, L. Dujourdy, J. Dahlen, Development of a harmonised method for the profiling of amphetamine VI: evaluation of methods for the comparison of amphetamine, Forensic Science International, 169 (2007) 86.

[21] S. Lociciro, P. Esseiva, P. Hayoz, L. Dujourdy, F. Besacier, P. Margot, Cocaine profiling for strategic intelligence, a cross-border project between France and Switzerland, part II: validaiton of the statistical methodology for the profiling of cocaine, Forensic Science International, 177 (2008) 199.

[22] M. Kuhn, K. Johnson, Applied predictive modeling, Springer, New York, 2013.

[23] S. Tangmose, P. Thevissen, N. Lynnerup, G. Willems, J. Boldsen, Age estimation in the living: transition analysis on developing third molars, Forensic Science International, 257 (2015) 512.e1.

[24] M. Gallidabino, C. Weyermann, F.S. Romolo, F. Taroni, Estimating the time since discharge of spent cartridges: a logical approach for interpreting the evidence, Science & Justice, 53 (2013) 41.

[25] C. Wyss, D. Cherix, Traité d'entomologie forensique, PPUR, Lausanne, CH, 2006.

[26] A.R. Ghiasvand, S. Hosseinzadeh, J. Pawliszyn, New cold-fiber headspace solid-phase microextraction device for quantitative extraction of polycyclic aromatic hydrocarbons in sediment, Journal of chromatography A, 1124 (2006) 35.

[27] S. Hosseinzadeh, J. Pawliszyn, Cold fiber solid-phase microextraction device based on thermoelectric cooling of metal fiber, Journal of chromatography A, 1216 (2009) 2783.

Figure Captions

Figure 1 – Flowchart of adopted pre-treatment strategy for multivariate regression models. Dashed rectangles indicates steps that have been optimised. "Skipped" means that tests have also been performed by skipping the related step. Centring was always applied.
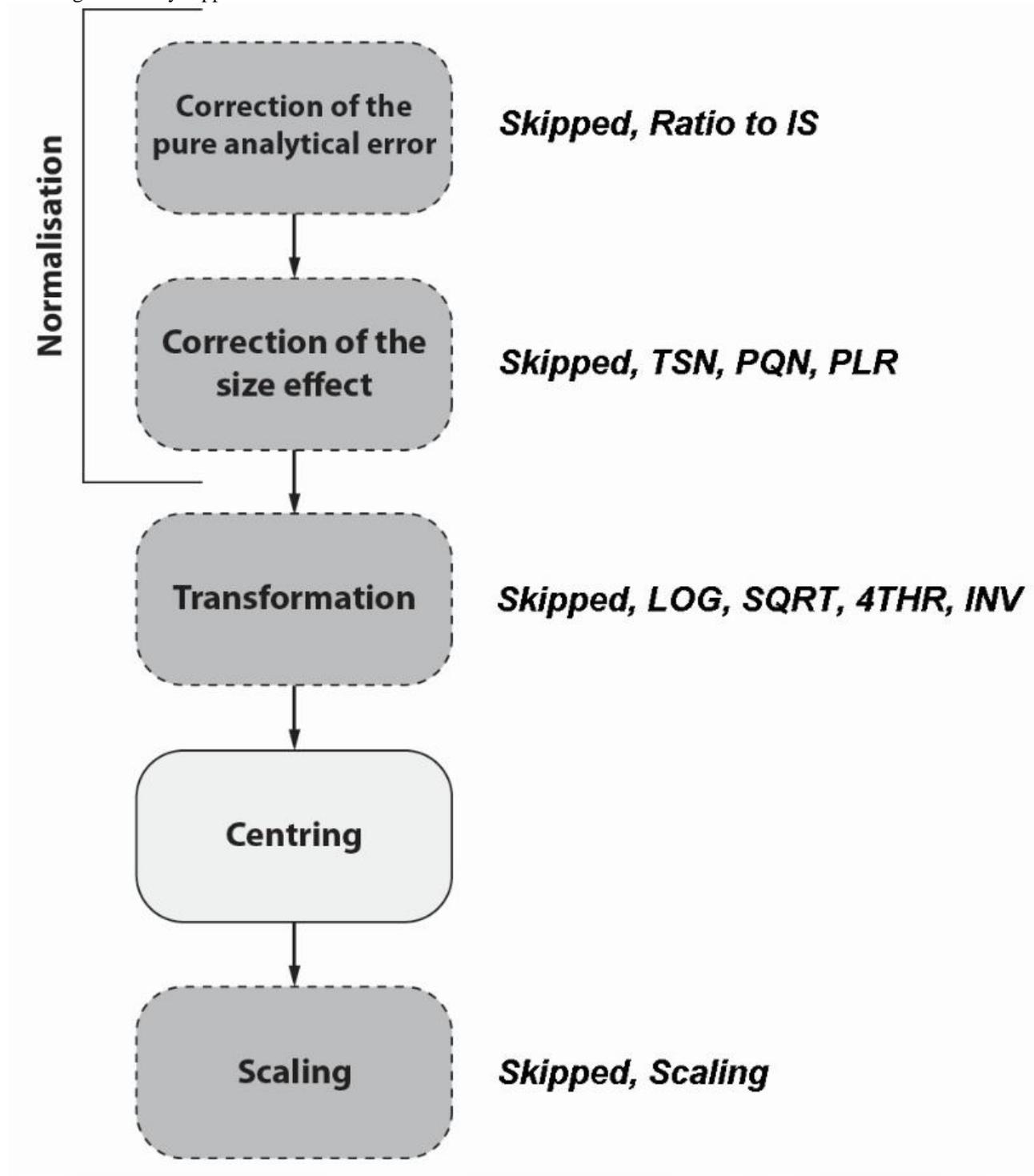
Figure 2 – Examples of ageing curves for 4 selected target compounds and 2 compound-to-compound ratios: (a) naphthalene, (b) biphenyl, (c) phenanthrene, (d) pyrene, (e) acenaphthylene / phenanthrene ratio and (f) acenapthtylene / 1-naphthalenecarbonitrile ratio. "BP" is the boiling point and "C" is the decrease-rate coefficient. For the compound-to-compound ratios, dashed lines indicate ageing profiles for the respective composing molecules.
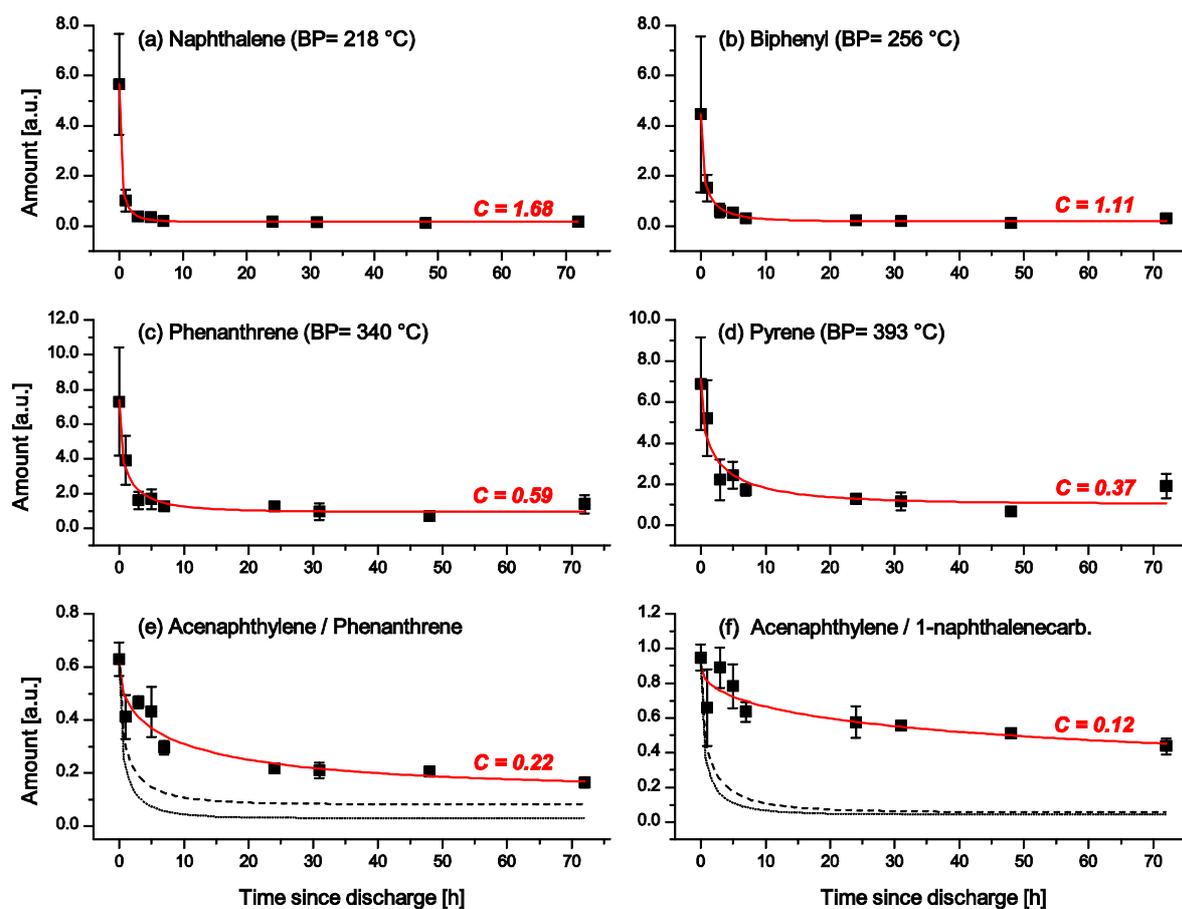
Figure 3 –Calibration and inaccuracy plots for the best multivariate and univariate regression models. For inaccuracy plots, cross-points represent the local means, while overall mean tendencies (approximated by quadratic regression) are represented by solid lines.
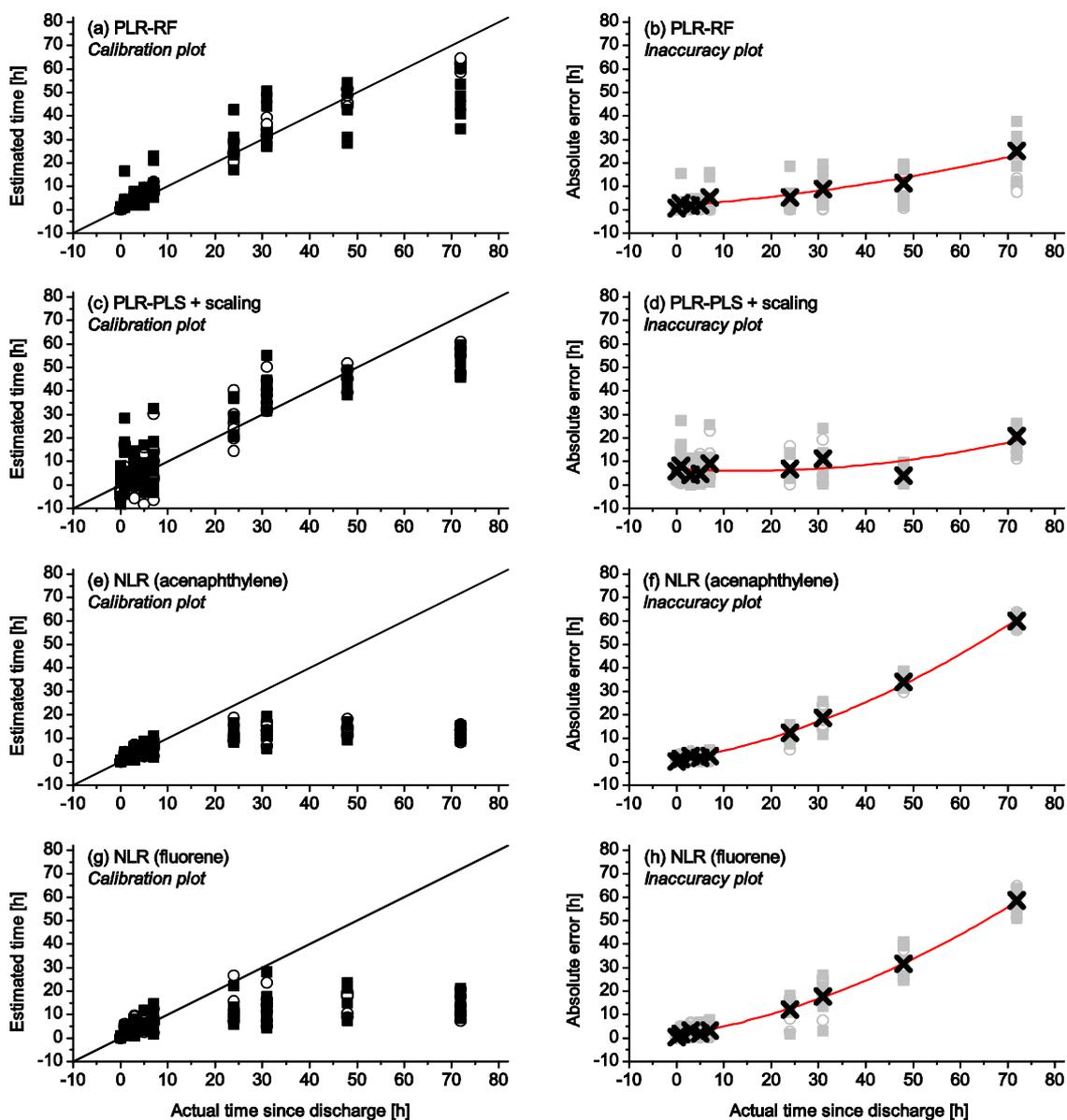
Figure 4 – Notched boxplots (10/25/75/90 percentiles) of the distribution of times since discharge estimated by the best multivariate and univariate regression models as a function of three actual cartridges ages. Estimates were centred and normalized to their standard deviation to improve visualisation.
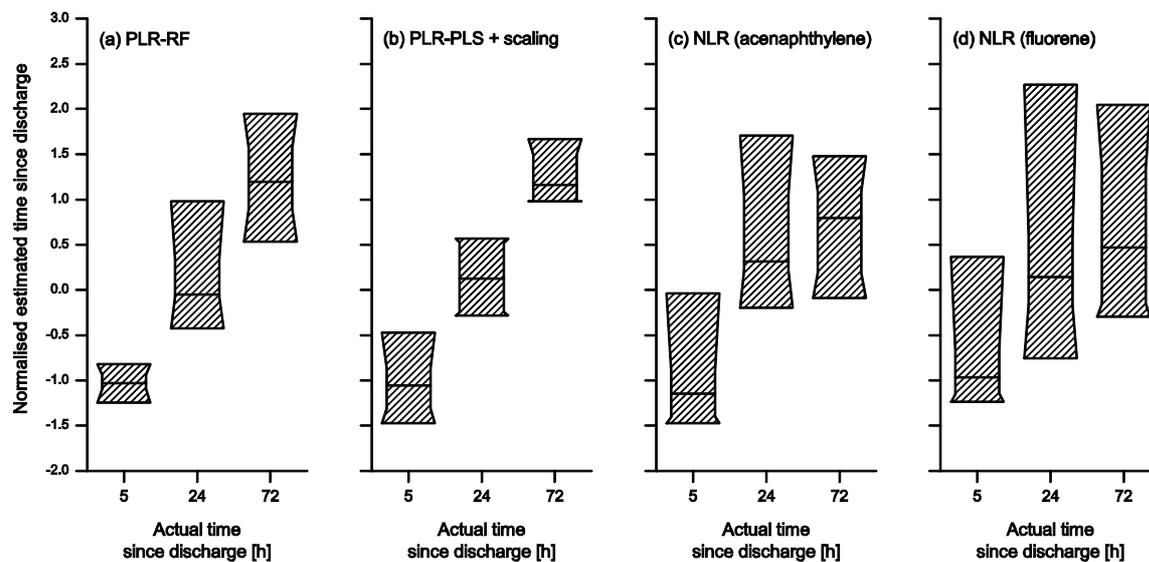
Figure 5 – Principal component analysis (PCA) of the two STD datasets after PLR normalization (all 210 ratios included). Explained variance of each principal component (PC) is reported in brackets.
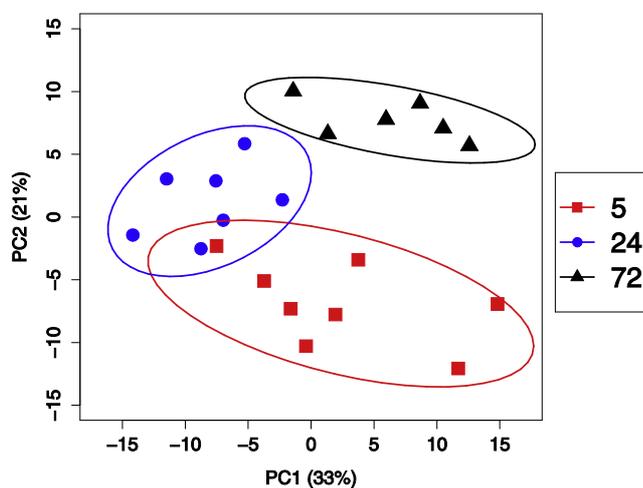
Figure 6 – Evolution of decrease rates (i.e., the *C* coefficients) of three selected molecules and the acenaphthylene / 1-naphthalenecarbonitrile ratio as a function of cartridge position, temperature, and relative humidity. For cartridge position, "1" and "2" indicate horizontal and vertical placement, respectively. For the ratio, the decrease rates of the composing molecules are equally shown for the sake of comparison.

28
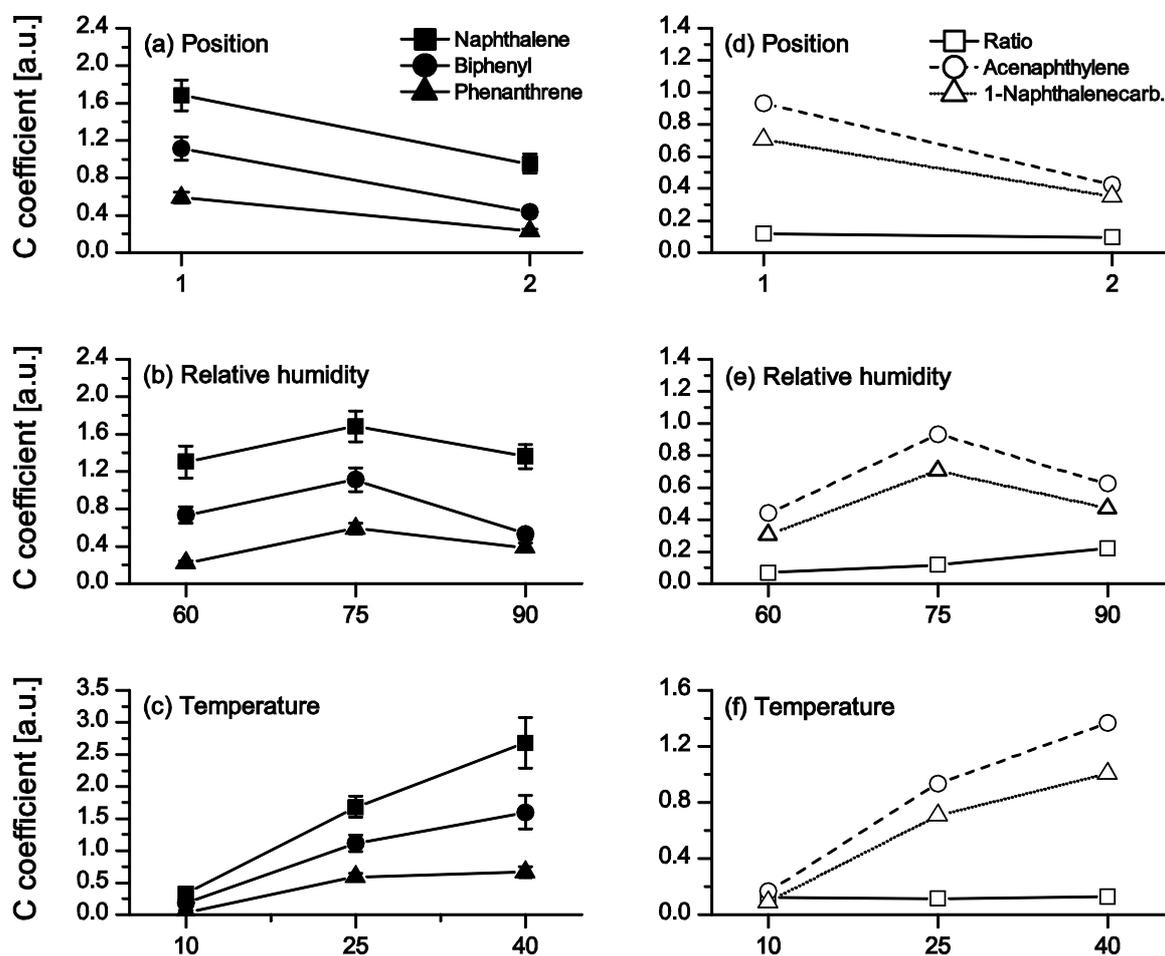
Figure 7 – Calibration and inaccuracy plots obtained with PLR-RF at different storage conditions in the case where the model was trained on reference cartridges aged at the same conditions as the questioned ones (refer to Figure S8 in ESM for complete data). For inaccuracy plots, cross-points represent local means, while overall mean tendencies (approximated by quadratic regression) are represented by solid lines. Dashed lines represent overall mean inaccuracy for univariate NLR models solely trained on the acenaphthylene data.
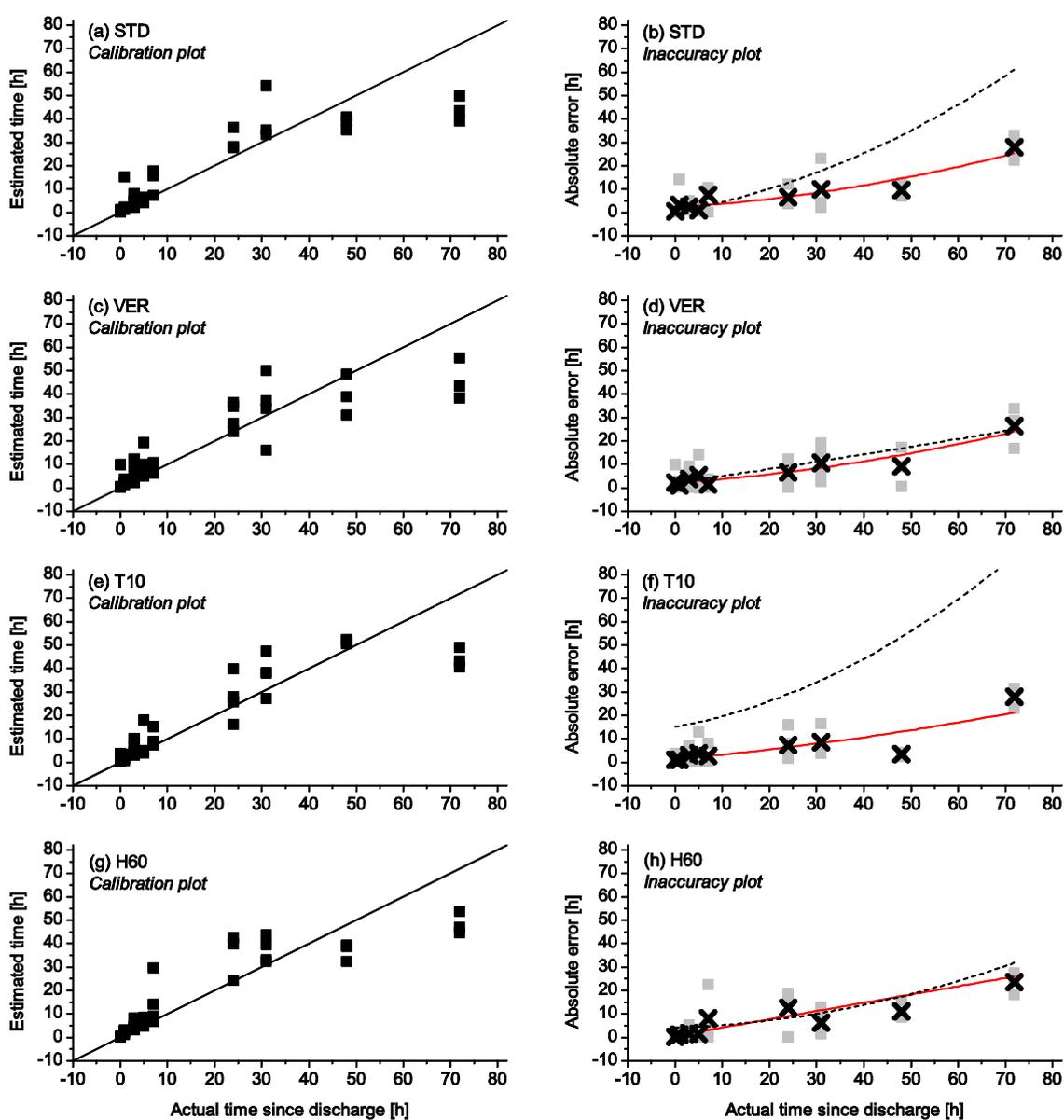
Figure 8 – Calibration and inaccuracy plots obtained by PLR-RF at different storage conditions in the specific case where the model was trained on reference cartridges aged at different conditions as the questioned ones (refer to Figure S9 in ESM for complete data). Particularly, the reference model was trained on cartridges aged at "standard" conditions. For inaccuracy plots, cross-points represent local means, while overall mean tendencies (approximated by quadratic regression) are represented by solid lines.
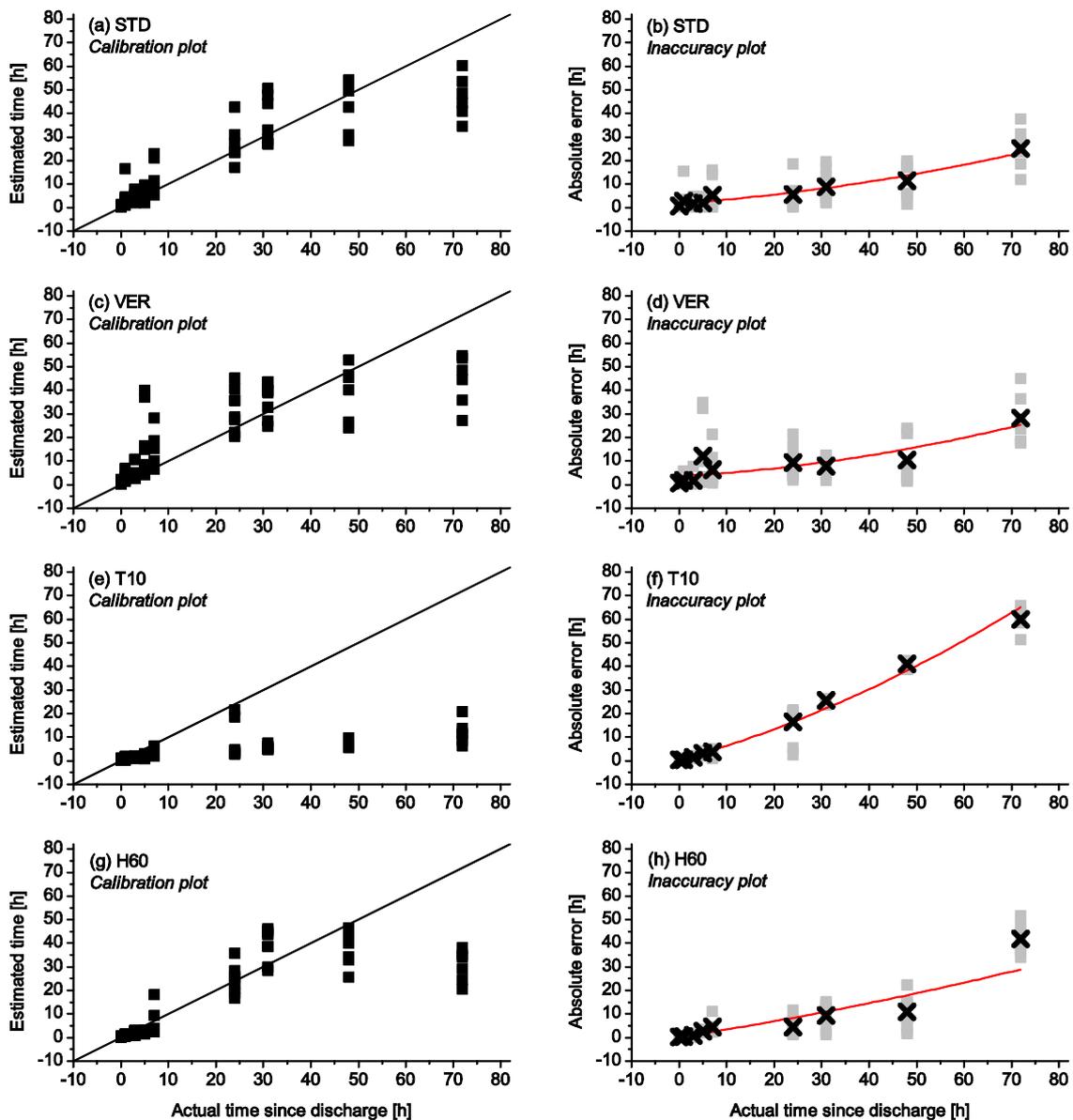
Tables

Table 1 – Ageing conditions for the different experimental series.

Dataset name	Cartridge position	Temperature [°C]	Relative humidity [%]
STD1	Horizontal	25	75
STD2	Horizontal	25	75
VER	Vertical	25	75
T10	Horizontal	10	75
T40	Horizontal	40	75
H60	Horizontal	25	60
H90	Horizontal	25	90

Table 2 – Estimated inaccuracy (I_t) and relative percentage to the actual time since discharge (%) for the best multivariate and univariate models over the experimental range. Inaccuracy values were approximated using a quadratic model.

Actual time [h]	PLR-RF		PLR-PLS + scaling		NLR (acenaphthylene)		NLR (fluorene)	
	I_t [h]	%	I_t [h]	%	I_t [h]	%	I_t [h]	%
1	1.8	183 %	6.7	665 %	0.8	79 %	1.3	132 %
5	2.5	50 %	6.3	126 %	2.4	47 %	2.8	56 %
24	6.5	27 %	6.4	26 %	12.8	53 %	12.7	53 %
48	13.7	28 %	10.4	22 %	32.9	69 %	31.8	66 %
72	23.2	32 %	18.9	26 %	60.8	84 %	58.3	81 %

Table 3 – P-values for ANOVA F-tests concerning the effect of varying storage conditions on compounds decrease rates (i.e., C coefficients). P-values near to 0 support the hypothesis that variations in the considered factor are significant on the decrease rate of the compound, while p-values close to 1 support the hypothesis of no significant effects.

Compound	Cartridge position		Temperature		Relative humidity	
	p-values	α < 0.010	p-value	α < 0.010	p-values	α < 0.010
Benzonitrile	0.001	X	<0.001	X	0.996	
p-Tolunitrile	<0.001	X	<0.001	X	0.480	
Naphthalene	<0.001	X	<0.001	X	0.333	
2-Methylnaphthalene	<0.001	X	<0.001	X	0.013	
1,2-Dicyanobenzene	0.005	X	<0.001	X	<0.001	X
Biphenyl	<0.001	X	<0.001	X	<0.001	X
Acenaphthene	<0.001	X	<0.001	X	<0.001	X
1-Naphthalenecarb.	<0.001	X	<0.001	X	<0.001	X
Phenanthrene	<0.001	X	<0.001	X	<0.001	X
Pyrene	0.002	X	<0.001	X	<0.001	X

Table 4 – Estimated inaccuracy (I_t) and relative percentage to the actual time since discharge (%) obtained with PLR-RF at different storage conditions, in the cases where the model was trained on reference cartridges aged at the same or different conditions as the questioned ones, respectively. In the case of inconsistent conditions, the reference model was trained on cartridges aged at "standard" conditions. Inaccuracy values were approximated using a quadratic model. Abbreviations for ageing conditions can be found in Table 1.

Consistent storage conditions												
Actual time [h]	**STD**		**VER**		**T10**		**T40**		**H60**		**H90**	
	I_t [h]	**%**	**I_t [h]**	**%**	**I_t [h]**	**%**	**I_t [h]**	**%**	**I_t [h]**	**%**	**I_t [h]**	**%**
1	2.3	230 %	2.3	231 %	1.6	155 %	1.5	147 %	1.0	102 %	1.6	155 %
5	2.9	58 %	2.9	58 %	2.3	46 %	2.1	42 %	2.4	49 %	3.0	61 %
24	6.8	28 %	6.7	28 %	6.4	27 %	5.8	24 %	9.1	38 %	10.2	42 %
48	14.5	30 %	14.0	29 %	13.1	27 %	12.2	25 %	17.6	37 %	19.1	40 %
72	25.4	35 %	24.0	33 %	21.3	30 %	20.5	28 %	26.1	36 %	28.1	39 %
Inconsistent storage conditions												
Actual time [h]	**STD**		**VER**		**T10**		**T40**		**H60**		**H90**	
	I_t [h]	**%**	**I_t [h]**	**%**	**I_t [h]**	**%**	**I_t [h]**	**%**	**I_t [h]**	**%**	**I_t [h]**	**%**
1	1.8	183 %	3.6	360 %	0.8	80 %	18.0	1804 %	0.5	52 %	2.5	254 %
5	2.5	50 %	4.1	83 %	3.2	63 %	18.3	366 %	1.8	36 %	3.9	78 %
24	6.5	27 %	7.8	32 %	16.5	69 %	19.7	82 %	8.4	35 %	10.3	43 %
48	13.7	28 %	15.1	31 %	38.2	80 %	21.3	44 %	18.0	37 %	18.5	38 %
72	23.2	32 %	25.4	35 %	65.3	91 %	23.0	32 %	29.0	40 %	26.6	37 %

34