Evidence of seasonal variation of ethyl glucuronide in hair: Modeling a seven-year data series

Evidence of seasonal variation of ethyl glucuronide in hair: modelling a seven-years data-series

Journal:	*Drug Testing and Analysis*
Manuscript ID	DTA-18-0158.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Własiuk, Patryk; Jagiellonian University in Kraków, Department of Analytical Chemistry; Institute of Forensic Research Alladio, Eugenio; Università di Torino , Chemistry; Consorzio Antidoping, Centro Regionale di Tossicologia Salomone, Alberto; Centro Regionale Antidoping, Vincenti, Marco; Università di Torino , Chemistry; Consorzio Antidoping, Centro Regionale di Tossicologia Zadora, Grzegorz; Institute of Forensic Research; The University of Silesia in Katowice, Department of Analytical Chemistry, Institute of Chemistry
Keywords:	Hair EtG, Ethyl glucuronide, Season, Wash-out, Alcohol abuse, Cut-off, Climatic conditions, Modelling
Abstract:	The assessment of chronic excessive alcohol consumption by ethyl glucuronide (EtG) determination in hair is generally based on a cut-off value of 30 pg/mg recognized by regulatory authorities and scientific societies that guide the decision process. The ongoing debate about the risks connected with the straightforward application of this cut-off refers to the factors that may influence the detected EtG concentration. The present contribution to this debate evaluates the seasonal variation of the averaged EtG values along a seven-years period. Over 65,000 data points have been statistically analysed to provide a mathematical model that interprets the data, gives insight into several influencing factors, and forecasts progressive data-points of the time series. This model shows that there is an annual pattern in the data exhibiting lower EtG concentrations during warm seasons and higher values in cold seasons. The estimated EtG cycles are characterized by the seasonal variation of ±2.78 pg/mg above and below the overall mean (with 5.56 pg/mg absolute difference overall). This seasonal factor associated with EtG quantification might result in a potential source of bias, at least in the regional/climatic conditions observed in the samples' collection area. Moreover, the EtG time series reveals that the change in the sample pre-treatment procedure has an effect on the modelled pattern as an abrupt increment (+38%) in the mean value of the EtG concentration. This change corresponds to the time when the former protocol of cutting hair into small segments before their extraction was substituted by their pulverization with a ball mill.

Evidence of seasonal variation of ethyl glucuronide in hair: modelling a seven-years data-series

Patryk Własiuk[a,b], Eugenio Alladio[c,d], Alberto Salomone[c], Marco Vincenti[c,d], Grzegorz Zadora[b,e]

[a] Department of Analytical Chemistry, Faculty of Chemistry, Jagiellonian University in Kraków, ul. Gronostajowa 2, 30-387 Kraków, Poland

[b] Institute of Forensic Research, Westerplatte 9, 31-033 Kraków, Poland

[c] Centro Regionale Antidoping e di Tossicologia "A. Bertinaria", Regione Gonzole 10/1, 10043 Orbassano (TO), Italy

[d] Dipartimento di Chimica, Università degli Studi di Torino, via P. Giuria 7, 10125 Torino, Italy

[e] Department of Analytical Chemistry, Institute of Chemistry, The University of Silesia in Katowice, Szkolna 9, 40-006 Katowice, Poland

Corresponding author:

Prof. Marco Vincenti

Dipartimento di Chimica, Università degli Studi di Torino,

Via Pietro Giuria, 7 – 10125 Torino, Italy

Phone: +390116705264 – Mobile: +393474198878

E-mail: marco.vincenti@unito.it

Patryk Własiuk: patryk.wlasiuk@gmail.com, Eugenio Alladio: ealladio@unito.it, Alberto Salomone: alberto.salomone@antidoping.piemonte.it, Marco Vincenti: marco.vincenti@unito.it, Grzegorz Zadora: gzadora@ies.krakow.pl.

Short title: Seasonal variation of ethyl glucuronide in hair

Keywords: Hair EtG, Ethyl glucuronide, Season, Wash-out, Alcohol abuse

Abstract

The assessment of chronic excessive alcohol consumption by ethyl glucuronide (EtG) determination in hair is generally based on a cut-off value of 30 pg/mg recognized by regulatory authorities and scientific societies that guide the decision process. The ongoing debate about the risks connected with the straightforward application of this cut-off refers to the factors that may influence the detected EtG concentration. The present contribution to this debate evaluates the seasonal variation of the averaged EtG values along a seven-years period.

Over 65,000 data points have been statistically analysed to provide a mathematical model that interprets the data, gives insight into several influencing factors, and forecasts progressive data-points of the time series. This model shows that there is an annual pattern in the data exhibiting lower EtG concentrations during warm seasons and higher values in cold seasons. The estimated EtG cycles are characterized by the seasonal variation of ±2.78 pg/mg above and below the overall mean (with 5.56 pg/mg absolute difference overall). This seasonal factor associated with EtG quantification might result in a potential source of bias, at least in the regional/climatic conditions observed in the samples' collection area.

Moreover, the EtG time series reveals that the change in the sample pre-treatment procedure has an effect on the modelled pattern as an abrupt increment (+38%) in the mean value of the EtG concentration. This change corresponds to the time when the former protocol of cutting hair into small segments before their extraction was substituted by their pulverization with a ball mill.

1. Introduction

The trustworthy assessment of harmful drinking represents a major commitment for forensic and clinical toxicologists and requires both a careful evaluation of suggestive medical evidences and the execution of laboratory analysis aimed at the detection of excessive alcohol consumption biomarkers [1–11]. The latter include indirect biomarkers and direct alcohol metabolites, with a wide range of reliability, depending on the proportion of correct classification of the investigated individuals among harmful drinkers, social drinkers, and teetotallers [3, 12–16].

In recent years, the determination of ethyl glucuronide (EtG) in the keratinous matrices has gained increasing appreciation, since it achieves the highest combination of sensitivity and specificity in the discrimination among alcohol consumers with different drinking habits [2, 16–18]. Thus, the determination of EtG in hair is nowadays widely accepted testing for monitoring chronic excessive alcohol intake and used in different areas of forensic and clinical toxicology, including workplace testing, firearms and driving licence re-granting, post-mortem investigation [19–20]. Since the notion of "chronic excessive alcohol intake" is relative, the EtG analytical results have to be compared with appropriate cut-off values [19] and the accuracy in the quantification of EtG in hair is a fundamental requirement. The current consensus document of the Society of Hair Testing (SoHT) [21] recommends to use a cut-off of 30 pg/mg for excessive drinking and 7 pg/mg for non-contradiction with abstinence. This recommendation has been consistently confirmed in all the revised editions of the consensus document [22].

The appraisal of EtG cut-off values is a subject of an ongoing debate stimulated by the alleged susceptibility of EtG results from several influencing factors. These may condition the final quantitation and affect the comparison with cut-off values. Each of these factors represents a potential source of variability, which relates to either individual, environmental, or methodological causes. In turn, these causes have an impact on the biological, chemical, and physical processes involving the EtG partition in the keratin matrix, and require careful interpretation [23] in the legal, forensic, and medical inquiries. The SoHT agreed to encourage the studies that concern these factors and re-evaluate them in the forthcoming revisions of the SoHT consensus. On the other hand, the very suitability of cut-off values to guide legal decisions has been recently criticized [24], and alternative approaches based on probabilistic evaluative procedures have been suggested to attain a final judgment [17,18,24].

The present study exploits a collection of over 65 thousand hair EtG data records to pursue a twofold purpose. On one hand, analytical evidence is delivered that the cut-off values should

be interconnected with the sample's pre-treatment procedure and the corresponding EtG extraction yield. This evidence confirms the conclusions contained in [25–27]. On the other hand, the climatic effect reported in [28] is expanded and approached by a rigorous time series analysis protocol. These two sources of potential variability are addressed in the present work and approached by meeting the two main goals of time series analysis that include the identification of the nature of the presented phenomenon as well as the task of predicting future values of the time series variable (the EtG concentration).

2. Methods

The hair EtG analytical results used in the present study arose from the samples collected over an eight-year period from subjects who underwent medical examination within either driving re-granting protocols (the large majority of the investigated population) or alcohol abuser's rehabilitation programs. The analyses were commissioned by several medical committees located in Piedmont, north-western Italy. The results constitute part of the daily activity of the forensic toxicology unit at the Regional Anti-Doping and Toxicology Center "A. Bertinaria" in Orbassano, Italy. All analyses were performed by LC-MS/MS using the same instrument since January 2011, described below. Previous results obtained with different instrumentation are not considered in the following statistical analysis.

2.1 Experimental protocol

Sample collection and pre-treatment

All hair samples were cut from the posterior vertex as close as possible to the scalp or the skin surface, using scissors freshly disinfected with glutaraldehyde. The samples were stored at room temperature and analysed within 10 working days. Only the proximal 0 – 3 cm segment was analysed whenever longer head hair samples were collected. Shorter head hair samples and chest hair samples were analysed in their total length. Typically, two-three locks of hair were collected from each subject. A hair aliquot of 40-50 mg was put together from the different hair locks, then weighted and washed twice using methylene chloride and methanol in sequence (1 min. each under shaking). Lastly, the washed hair were crumbled/pulverized.

During the considered activity period (from January 2009 till December 2017), the laboratory adopted two different procedures for the extraction step. In the first period (January 2009 – September 2015), dried hair was cut into small snippets (about 1 mm) with freshly cleaned

scissors. Starting in October 2015, the dried hair was pulverized using a metal beads mill, namely a Precellys 24 Tubes Homogenizer (Bertin Pharma, France) equipped with six 2.8 mm metal beads. In the forthcoming sections, these two hair crumbling procedures are simply denoted as "cutting" and "milling", respectively.

EtG extraction was carried out overnight (≅ 16 h) at room temperature with 720 µL of a 35:1 water–methanol mixture. Then, the samples were sonicated for 90 min., centrifuged at 14000 rpm for 90 s, and a 100 µL aliquot of liquid phase was transferred into a vial for UHPLC-MS/MS analysis.

EtG determination

Analyses were performed using a Shimadzu Nexera UHPLC-system (Shimadzu, Duisburg, Germany) interfaced to a Sciex Triple Quad™ 5500 triple quadrupole mass spectrometer (AB Sciex, Darmstadt, Germany) with an electrospray ionization (ESI) source operating in the negative ion mode. Full details on the instrumental method are reported in previous publications [25, 29].

The method was internally validated and accredited in accordance with ISO/IEC 17025:2005 rules [29]. Further details on the most recent validation procedures for the analytical method are extensively discussed elsewhere [27]. Operational (i.e., fully guaranteed under routine conditions throughout the investigated period) limits of detection (LOD) and quantification (LOQ) were conventionally set at 3 and 10 pg/mg, respectively, even if the actual LOQ was calculated as 1.7 pg/mg and experimentally verified at 2.0 pg/mg [27]. The laboratory performance in hair EtG analysis was constantly monitored through regular participation to inter-laboratory proficiency tests organized by SoHT, the Gesellschaft für Toxikologische und Forensische Chemie (Jena University Hospital, Germany), and the Centre Universitaire Romand de Médecine Légale (University Hospital of Geneva, Switzerland). Quality control samples EGH 2/12-B from ACQ Science GmbH (Rottenburg-Hailfingen, Germany) were periodically analysed.

2.2 Statistical analysis: definition of EtG time series

Over 65 thousand data points that describe the results of EtG quantification in hair were available. These data cover the analyses performed at the Regional Anti-Doping and Toxicology Center over the period of 9 years (since 2009). During this time frame, some individuals were tested several times, but the testing periodicity and total extent of the

observational period were highly variable and depended on the judgements of the Medical Commissions that prescribed the testing. Therefore, all results have been considered as arising from independent samples. The hair EtG results were averaged on a monthly basis, yielding a single data point for each month and a total of 78 data points. In practice, such a sequence of observations defines a time series, in which a pair $(x_i, t_i)_{i\in T}$ is to be recognized as the mean value of EtG concentration x_i, associated with a time-stamp t_i at a i-th discrete time. The available data that span T over Jan. 2009 – June 2017 were taken as the data for the identification step of the phenomenon contained in the time series. However, due to the much smaller number of samples in the first two years in comparison to the subsequent years, the time span was restricted to start in Jan. 2011. A graphical representation of the total number of observations per month and the average sample size associated with the monthly means is reported in Figure 1a. The minimum (2011) and maximum (2017) number of yearly processed samples are equal to 6850 and 15750, respectively, with average sample size of 570 and 1310 samples/month, respectively. Moreover, each result with EtG value exceeding 250 pg/mg was discarded from the dataset since it would otherwise over-affect the corresponding mean and might be thought of as an outlier. These two pre-treatment procedures resulted in the exclusion of merely 0.6% database records. Additionally, it was decided not to discard the samples with a result below the conventional LOQ of 10 pg/mg (but above the actual LOQ of 2 pg/mg). In support to this choice, two time series where created, with and without the data below 10 pg/mg (available in the Supplementary Material). By time-shifting the first dataset with respect to the other many correlation coefficients were calculated. This cross-correlation test of two time series indicates that the correlation is strongest and significant when no shift between datasets is present. Moreover, after running the regression between them and correcting for the serial correlation in the residuals, a positive and significant independent variable (without time shift) was noted. Thus, it was decided to continue the data analysis with the chosen sample set due to the evidence that those two time series are presenting the same pattern. As the time series is a sequence of n $(x_i, t_i)_{i\in T}$ pairs and a month was chosen as a seasonal unit period, the whole data corresponds to 78 periods (n = 78). The data for the remaining six month of 2017 (July-December 2017) were held out to be used in the validation study. The same pre-processing strategy was applied to this test dataset. All statistical analyses were conducted using R software [30].

3. Results and Discussion

The time series described in the preceding section reveals no evident global trend, but the presence of two characteristic features might be noted (Figure 1b). Firstly, the presence of a seasonal pattern is anticipated from a time series profile that exhibits alternating peaks and valleys at about annual intervals. Secondly, a quite abrupt increase in the time series is observed from the end of 2015 onwards, when both peaks and valley reach their maximum point. This is clearly associated with a general change within the sequence of EtG observations that increases the mean values. These two features are inquired in the following sections.

3.1 Breakpoint

An ordered sequence of data points $\{x_1, ...,x_n\}$ is expected to be split into two segments, $\{x_1:x_k\}$ and $\{x_{k+1}:x_n\}$, if there is a change at time t_k. The splitting time-stamp t_k is sought and it is agreed to look for a single change in the mean, suggested by the visual inspection of the time series at hand. The task of estimating the point at which this statistical property changes fits the hypothesis testing framework, highlighted in [31], where the likelihood-ratio approach is used to test the null hypothesis (H_0, hypothesis of stability, which corresponds to no change) versus the alternative hypothesis (H_1, which corresponds to a single change at the time t_k). A parametric model, that assumes the data are independent and normally distributed, employs a test statistic for two segments obtained by a given splitting value of k. Thus, the test statistic requires the calculation of the maximum log-likelihood value under both null (H_0: $\mu_1 = \mu_2 = ... = \mu_n = \mu$, where μ is unknown) and alternative (H_1: $\mu_1 = ... = \mu_k \neq \mu_{k+1} = ... = \mu_n$) hypothesis. Over the sequence of the considered k values ($2 \leq k \leq n-2$) the one yielding the maximum value for this test statistic is taken to represent the estimate of the time point at which the change occurs (the maximum value of the test statistic represents the largest discrepancy between the H_0 and H_1). To find this point an R [30] changepoint package for change-point analysis [32] was used.

The analysis revealed that the change is detected to occur at $\hat{k} = 57$ indicating the October 2015 to be the first month of the "new" period. This result was followed by a bootstrap study on the residuals of time series obtained for 10000 samples each delivering an estimate $\hat{k}^*$. From such a collection of results it was found that the proportion of samples with $|\hat{k}^* - \hat{k}| \leq 1$ is 70% (which is an unconditional bootstrap confidence). This confirmation test ($\hat{k}^* = \hat{k} = 57$) not only agrees with the visual inspection of the times series but it also corresponds well with the knowledge of the procedural change employed in the laboratory in autumn 2015

when the hair sample preparation technique for EtG quantification had been switched from cutting to milling. After the procedure had changed, about 35% rise in the mean value of the times series is observed (the means of all values before and after October 2015 are compared), indicating that on average the detected EtG concentrations were higher after the hair pulverizing strategy was employed with respect to the previous period when the hair specimens were finely fragmented by cutting them with scissors.

3.2 Seasonal pattern

The visual inspection of the time series reported in Figure 1b indicates that the alternating pattern of peaks and valleys is approximately repeated in the same months of the year. The pattern highlights that the EtG profile annually peaks in winter and declines in summer or late-summer. Thus, the mean EtG concentration varies month after month but remains at a fairly constant value in the same month of different years. This shapes the phenomenon of monthly seasonality in the time series, where the seasonal period d is set and fixed to 12 since twelve observations constitute the annually repeated cycle. This repetitive symmetric pattern suggests that the seasonal variation could be described by the sine function that models the seasonality with a smooth description of the variation. Moreover, the final model for the overall time series should be supplemented by an additional term to account for the procedural change detected to occur in October 2015. This calls for the introduction of a dummy variable that codes the 'cutting' and 'milling' period and allows for the intercept of a regression model to be adjusted for this fact. Thus, the proposed final model is expressed in the following form:

$$x_i = \beta_0 + \beta_1 I_i + \sum_{j=1}^{d/2} \left[\beta_{1+j} \sin\left(\frac{2\pi j}{d} i\right) + \beta_{2+j} \cos\left(\frac{2\pi j}{d} i\right) \right] + \epsilon_i$$

where $\beta_0 + \beta_1 I_i$ is the trend expressed as a mean of the time series along with an indicator variable I_i that flags the occurrence of the laboratory procedure change (I_i=0 before the procedure change and I_i=1 after the change from cutting to milling). This component models the intervention effect of known source and known time of emergence. The index i refers to the chosen time scale (ruled by the sequence of integers with $i = 1$ that corresponds to Jan 2011). The sum components, sine and cosine terms, are employed to model the cyclic component of the time series. It was chosen to base the part describing the seasonality pattern solely on the fundamental frequency, $2\pi/12$, since adding its harmonics was associated with

the increase of both Akaike and Bayesian information criteria and did not contribute to the increase of the amount of explained variance. However, after the initial fit, the analysis of residuals revealed that the error part ϵ_i is not a white noise, i.e. $N(0, \sigma^2)$. The Durbin-Watson test as well as the analysis of (partial) autocorrelation plots delivered the evidence that the residuals exhibit an autocorrelation structure. To provide further details about this structure, it can be stated that the errors are positively correlated in the following manner: $\epsilon_i = \phi\epsilon_{i-1} + a_i$, where ϕ is the autocorrelation parameter and a_i corresponds to the white noise. This means that the regression model ought to be fitted keeping such error structure in mind. The final model fitted to the monthly EtG series can be written as:

$$x_i = 10.29 + 3.97I_i + 1.84\sin\left(\frac{2\pi}{12}i\right) + 2.09\cos\left(\frac{2\pi}{12}i\right) + \epsilon_i$$

where the error structure is given by: $\epsilon_i = 0.74\epsilon_{i-1} + a_i$ and a_i is a white noise with the pattern not suggesting that the regression assumptions with respect to the residuals are violated. The autocorrelation component of the model corresponds to some source of inertial effect (for example, an early or late climatic changes in a single year with respect to the average), that makes two consecutive data-points more likely to be placed on the same side of the model line or simply reflects the smoothed variation of the model with respect to the real data in some specific years (for example, 2012 and 2014), or a combination of both effects (Figure 2). However, no additional variable having time-ordered effect on the modelled response was available to resolve this issue in an explicit manner.

Two aspects of the model, presented in the Figure 2a, should be noted. First of all, the effect of the new sample preparation procedure produces a permanent, abrupt, and constant shift in the mean level of the EtG time series. The model indicates that this change describes an increase of the mean EtG level of 38.6% since October 2015. Secondly, the estimate of the amplitude of the wave component describing the seasonal behaviour of EtG concentrations shows that the estimated EtG cycles have peaks and troughs of about 2.78 pg/mg above and below the overall mean, corresponding to an absolute difference of 5.56 pg/mg. Considering solely the part of the model that is responsible for the seasonality component, one might notice that the seasonality given by the wave peaks whenever the time scale corresponds to i = 2, 14, ... and troughs at i = 8, 20, ... Thus, the maximum is associated with February and the minimum with August, respectively. This has an immediate translation to the forecast pattern given in the Figure 2a for the remaining second-half of 2017, where the second month for

which the forecast is delivered, namely August, is in the pit of the whole forecasted sequence. The analysis of the accuracy of this forecast is deferred for the remainder of the next section.

In the average EtG profile reported in Figure 2a, two features show important deviation from the calculated model, namely the long-lasting EtG concentration drop observed in the year 2013 and the delayed winter peaks observed in 2012 and 2014. Although the summer season of the years 2011 and 2013 are recalled as particularly warm, no significant evidence of this anomaly is present in the monthly average temperature of Piedmont (Figure 2b) [33] to explain the EtG data.

3.3 Stability of the model

Once the regression model was reckoned, its stability had to be checked. A twofold approach was implemented for this verification, including the testing proposed in [34] and a forecasting approach aimed to verify both the stability and suitability of the proposed model.

The test presented in [34] is based on the idea that if the model's parameters do not change over time, then building the model on data up to time t ought to be sufficient for forecasting the modelled response at $t+1$. By proceeding with this approach until time T, the standardised one-step ahead recursive residuals may be calculated to yield an empirical process. The analysis of the progressive path for this process will deliver the indication of whether the model suffers from the parameters' inconsistency or not. If the process path crosses the boundaries with probability α, the null hypothesis of "no instability" should be rejected. Here, the initial model was based on the first two years of the EtG time series and then used for forecasting the value of the next response, sequentially from t=25 up to $T-1$, adding one response at a time. Two empirical processes for calculating the cited residuals were considered, according to [34], namely the cumulative sum of standardized residuals (CUSUM test) and the squared standardized residuals (CUSUM of squares test). Figure 3 shows that both patterns fit well the significance tunnels, i.e., the processes capturing the behaviour of recursive residuals are within the boundaries, and a conclusion on the model's stability and its regression parameters might be drawn. It is worth noting (Figure 3, left) that the fit of the model improves further after the 57th observation number. This underbelly point (as inferenced from the CUSUM of squares process) corresponds to the procedural intervention undertaken in Oct 2015 in the laboratory (switching from cutting to milling).

After the successful validation, the model can be proposed as a suitable device for extrapolation. The in-sample prediction delivers root mean square error (RMSE) of 1.61 pg/mg, while the RMSE for the naïve seasonal in-sample prediction (that assumes the forecast is equal to the response value from the adequate month of the previous season) is equal to 4.3 pg/mg. Thus, the delivered evidence consolidates the effort to forecast the EtG values for the remaining six months of 2017 (given by the solid line with triangles in Figure 2a). The data not used for the construction of the model and referring to the monthly EtG means for the second half of 2017 allow the out-of-sample prediction error to be computed. The RMSE for this independent validation set equals 1.39 pg/mg. Moreover, the relative RMSE measure of the forecast accuracy was computed to validate whether there is a room for improvement. This unitless measure is a ratio of the RMSE for the proposed model and the RMSE of the benchmark forecast method. Here, the relative RMSE was computed with respect to two benchmark methods, namely a "random walk" (that assumes a forecast is equal to the last observation) and naïve seasonal (defined as previously). Both values are smaller than one (0.62 for the former and 0.36 for the latter benchmark method) indicating that the developed model works better than the benchmarks methods [35]. Consequently, the proposed model might be termed as a suitable description of the phenomenon contained in the analysed EtG time series. After October 2015, the real data-points apparently show larger fluctuations than the model envisages, but longer data-series are necessary to verify whether also the sinusoidal component of the mathematical model has to be slightly corrected, as a consequence of the introduction of the milling hair treatment.

3.4 Possible explanations of the seasonal pattern

The pictured course of seasonality pattern revealed that the warmer the months the lower the mean EtG concentration, whereas the EtG profile peaks during winter. The plot of the monthly means deduced from the EtG time series along with the data describing the monthly average temperature for the analogous period (2011-2017) clearly advocates this statement (Figure 4). The monthly EtG means were calculated adjusting the 'milling' period to level with the 'cutting' period in accordance with the modelled effect. The monthly temperature means (Figure 2b) were taken from the Agenzia Regionale per la Protezione Ambientale del Piemonte (Turin, Italy) [33], an agency that collects information and data on environmental monitoring in the Piedmont region of Italy, namely the same geographic area from which all hair samples of the present study were collected and submitted to Regional Anti-Doping and

Toxicology Center laboratory for the analysis. Thus, both data series are representative of the same specific region; it is most likely that other geographical areas with different climatic patterns may be arranged with dissimilar trends for mean hair EtG along the monthly sequence. It is worth noting that the EtG minimum is observed in September (August if a seasonality component of the model is considered) while the temperature peak is recorded in July, and, in general, a 0.5-1.5 month shift is observed in the extreme values of EtG with respect to the monthly average temperature reported in Figure 2a: this delay can be easily explained taking into account that the 3-cm hair segment undergoing analysis approximately corresponds to the last 3 months of EtG incorporation into the hair. Interestingly, the fitted wave, stemming from the seasonality component of the developed regression model (Figure 2), seems to be horizontally mirroring the temperature pattern.

Various explanations for the observed seasonality pattern might be postulated, referring to both different consumption habits along the year or a variety of bias-producing factors. For example, it is not unlikely that in the warm periods people tend to refrain from consuming too much alcohol and conversely more abundant intake may be expected during cold season. Under such a hypothesis, the hair EtG concentration eventually reflects the average alcohol consumption of the previous 3-months period, making the seasonal pattern significant for clinical evaluation of alcohol abuse, not for legal purposes. However, it should also be considered that the large majority of the data comprised in the present study refers to subjects tested throughout the driving licence re-granting procedure, who are expected to refrain from excessive alcohol intake anyway, independently from the observational season. Therefore, other factors besides actual exposure to alcohol may be taken into account for the interpretation of the seasonal profile.

Factors possibly associated with a bias include behavioural and environmental causes. As EtG is a hydrophilic compound its elimination/dilution might be associated with the increased perspiration during the warm season. Although EtG is expected to be predominantly transferred to the keratin matrix by incorporation from blood in the upper part of the hair root [15, 36], the relative contribution from sweat in different climatic conditions has not been investigated yet. Another potential source of negative bias that has been consistently claimed in the literature is attributed to the washing-out effect increasingly occurring along the distal portion of the hair shaft [37–39]. Indeed, a quite large portion of the population takes sea-, lake-, and swimming-pool bathing during the summer season and more frequent showering is taken to wash out the sweat and cool the body. Long-lasting immersions in water during

swimming may hypothetically favour the hair keratin swelling and consequent release of part of the incorporated EtG. More specifically designed experiments have to be performed in the future to figure out which one of these possible explanations might chiefly account for the observed seasonal pattern.

4. Conclusions

The present study investigated the factors that may affect the results of hair EtG analysis from the point of view provided by a very large data-set, on the basis of the statistical principle that randomization on large population levels off the individual variability and balances their probability distributions. The first result highlighted by comparing our large data-series is that changing the pre-analytical treatment of hair from "cutting" to "milling" resulted in an increase of the measured EtG mean concentration of about 38%. This confirms the previously reported findings [25,27], where the same average increase was established when the two crumbling techniques were applied to the same hair specimens, suggesting that a more exhaustive extraction of EtG is obtained when the hair matrix is pulverized with a metal ball mill rather than cut into small snippets with scissors. This procedural change within the same analytical framework proved to have a significant impact on the final EtG result. Any comparison of such a result with a specific cut-off value ought to be at least related to the practical details of the analytical protocol. This element of uncertainty highlights the need for the analytical methods to be harmonized across different laboratories whenever a unique cut-off value is considered [22]. More appropriately, interpretation criteria based on probabilistic methods should be adopted [24].

The seasonal effect observed on the time series and based on over 65,000 measurements is intrinsic to the EtG determination, at least in the regional area (Piedmont, Italy) of hair sample collection. The average annual EtG profile peaks at cold months and declines in warm months. The possible reasons might include (1) increased sweating in warm months; (2) increased frequency of showering and sea-bathing during summer; (3) increased intake of ethanol-containing beverages during winter. While the latter alleged factor reflects effective exposure to alcohol, factors (1) and (2) advance the hypothetical occurrence of bias, possibly associated with a washing-out effect. These hypotheses have to be tested further with purposely-designed experiments. In general, the present study demonstrates that behavioural and environmental factors may play a significant role in the outcome of hair EtG results.

The present work adds another point to the debate on the possible sources of variability that should be taken into account when comparison of the value of EtG concentration with a cut-off value is of interest. The very concept of using cut-off values in the forensic context is questionable [24] and the progressive identification of potential influencing factors encourages the adoption of probabilistic approaches to alcohol biomarkers evaluation [17,18] that may use "prior odds" to consider these factors within a Bayesian framework. The clinical setting does not suffer equally from the presence of such sources of bias, since the physicians' judgement is determinative, but even in this case the expression of a correct diagnosis is favoured by a comprehensive evaluation of the factors that may modify the laboratory data. *A fortiori*, the forensic expert ought to work within a proper framework of evidence evaluation, where such sources of variability can be taken into account.

References

[1] Oppolzer, D., Barroso, M., & Gallardo, E. (2016). Bioanalytical procedures and developments in the determination of alcohol biomarkers in biological specimens. Bioanalysis, 8(3), 229–251. https://doi.org/10.4155/bio.15.240

[2] Vincenti, M., Salomone, A., & Pirro, V. (2013). How has screening of harmful drinking changed over the years? Bioanalysis, 5(24), 2981–2983. https://doi.org/10.4155/bio.13.277

[3] Hastedt, M., Büchner, M., Rothe, M., Gapert, R., Herre, S., Krumbiegel, F., Tsokos, M., Kienast, T., Heinz, A., Hartwig, S. (2013). Detecting alcohol abuse: traditional blood alcohol markers compared to ethyl glucuronide (EtG) and fatty acid ethyl esters (FAEEs) measurement in hair. Forensic Science, Medicine, and Pathology, 9(4), 471–477. https://doi.org/10.1007/s12024-013-9416-8

[4] Agius, R., Nadulski, T., Kahl, H.-G., & Dufaux, B. (2012). Ethyl glucuronide in hair – A highly effective test for the monitoring of alcohol consumption. Forensic Science International, 218(1–3), 10–14. https://doi.org/10.1016/j.forsciint.2011.10.007

[5] Pirro, V., Di Corcia, D., Seganti, F., Salomone, A., & Vincenti, M. (2013). Determination of ethyl glucuronide levels in hair for the assessment of alcohol abstinence. Forensic Science International, 232(1–3), 229–236. https://doi.org/10.1016/j.forsciint.2013.07.024

[6] Pragst, F., & Balikova, M. A. (2006). State of the art in hair analysis for detection of drug and alcohol abuse. Clinica Chimica Acta, 370(1–2), 17–49. https://doi.org/10.1016/j.cca.2006.02.019

[7] Niemelä, O. (2016). Biomarker-Based Approaches for Assessing Alcohol Use Disorders. International Journal of Environmental Research and Public Health, 13(2), 166. https://doi.org/10.3390/ijerph13020166

[8] Cappelle, D., Neels, H., De Keukeleire, S., Fransen, E., Dom, G., Vermassen, A., Covaci A., Crunelle C. L., van Nuijs, A. L. N. (2017). Ethyl glucuronide in keratinous matrices as biomarker of alcohol use: A correlation study between hair and nails. Forensic Science International, 279, 187–191. https://doi.org/10.1016/j.forsciint.2017.08.022

[9] Barbaro, M.,& Locatelli, M. (2016). The Markers for Alcohol Abuse: The Good, the Bad and the Ugly. Journal of Alcoholism & Drug Dependence, 4(3). https://doi.org/10.4172/2329-6488.1000242

[10] Oppolzer, D., Barroso, M., Passarinha, L., & Gallardo, E. (2016). Determination of ethyl glucuronide and fatty acid ethyl esters in hair samples. Biomedical Chromatography, 31(4), e3858. https://doi.org/10.1002/bmc.3858

[11] Jastrzębska, I., Zwolak, A., Szczyrek, M., Wawryniuk, A., Skrzydło-Radomańska, B., & Daniluk, J. (2016). Biomarkers of alcohol misuse: recent advances and future prospects. Gastroenterology Review, 2, 78–89. https://doi.org/10.5114/pg.2016.60252

[12] Pirro, V., Oliveri, P., Sciutteri, B., Salvo, R., Salomone, A., Lanteri, S., & Vincenti, M. (2013). Multivariate strategies for screening evaluation of harmful drinking. Bioanalysis, 5(6), 687–699. https://doi.org/10.4155/bio.13.12

[13] Boscolo-Berto, R., Favretto, D., Cecchetto, G., Vincenti, M., Kronstrand, R., Ferrara, S. D., & Viel, G. (2014). Sensitivity and Specificity of EtG in Hair as a Marker of Chronic Excessive Drinking. Therapeutic Drug Monitoring, 36(5), 560–575. https://doi.org/10.1097/ftd.0000000000000063

[14] Boscolo-Berto, R., Viel, G., Montisci, M., Terranova, C., Favretto, D., & Ferrara, S. D. (2012). Ethyl glucuronide concentration in hair for detecting heavy drinking and/or abstinence: a meta-analysis. International Journal of Legal Medicine, 127(3), 611–619. https://doi.org/10.1007/s00414-012-0809-0

[15] Pragst, F. (2015). Alcohol biomarkers in hair, in Kintz, P., Salomone, A., Vincenti, M. (Eds) (2015) Hair Analysis in Clinical and Forensic Toxicology, Elsevier – Academic Press, San Diego, CA, USA

[16] Pirro, V., Valente, V., Oliveri, P., De Bernardis, A., Salomone, A., & Vincenti, M. (2011). Chemometric evaluation of nine alcohol biomarkers in a large population of clinically-classified subjects: pre-eminence of ethyl glucuronide concentration in hair for

confirmatory classification. Analytical and Bioanalytical Chemistry, 401(7), 2153–2164. https://doi.org/10.1007/s00216-011-5314-7

[17] Alladio, E., Martyna, A., Salomone, A., Pirro, V., Vincenti, M., & Zadora, G. (2017). Evaluation of direct and indirect ethanol biomarkers using a likelihood ratio approach to identify chronic alcohol abusers for forensic purposes. Forensic Science International, 271, 13–22. https://doi.org/10.1016/j.forsciint.2016.12.019

[18] Alladio, E., Giacomelli, L., Biosa, G., Corcia, D. D., Gerace, E., Salomone, A., & Vincenti, M. (2018). Development and validation of a Partial Least Squares-Discriminant Analysis (PLS-DA) model based on the determination of ethyl glucuronide (EtG) and fatty acid ethyl esters (FAEEs) in hair for the diagnosis of chronic alcohol abuse. Forensic Science International, 282, 221–230. https://doi.org/10.1016/j.forsciint.2017.11.010

[19] Kintz, P. (2012). Consensus of the Society of Hair Testing on hair testing for chronic excessive alcohol consumption 2011. Forensic Science International, 218(1–3), 2. https://doi.org/10.1016/j.forsciint.2011.10.025, available on-line: http://www.soht.org/images/pdf/2014%20Alcohol%20markers%20revision%2013JUN14%20FINAL.pdf. Last access 26th March 2018.

[20] Salomone, A., Tsanaclis, L., Agius, R., Kintz, P., & Baumgartner, M. R. (2016). European guidelines for workplace drug and alcohol testing in hair. Drug Testing and Analysis, 8(10), 996–1004. https://doi.org/10.1002/dta.1999

[21] Society of Hair Testing: 2016 Consensus for the Use of Alcohol Markers in Hair for Assessment of both Abstinence and Chronic Excessive Alcohol Consumption. http://www.soht.org/images/pdf/Revision%202016_Alcoholmarkers.pdf. Last access 6th January 2018.

[22] Pragst, F., Suesse, S., Salomone, A., Vincenti, M., Cirimele, V., Hazon, J., Tsanaclis, L., Kingston, R., Sporkert, F., & Baumgartner, M. R. (2017). Commentary on current changes of the SoHT 2016 consensus on alcohol markers in hair and further background information. Forensic Science International, 278, 326–333. https://doi.org/10.1016/j.forsciint.2017.07.023

[23] Kintz, P., Salomone, A., Vincenti, M. (Eds.), Hair Analysis in Clinical and Forensic Toxicology, Academic Press (2015).

[24] Biedermann, A., Taroni, F., Bozza, S., Augsburger, M., Aitken, C.G.G., Critical analysis of forensic cut-offs and legal thresholds: a coherent approach to inference and decision. Forensic Science International, in press, https://doi.org/10.1016/j.forsciint.2018.04.030.

[25] Salomone, A., Baumgartner, M. R., Lombardo, T., Alladio, E., Di Corcia, D., & Vincenti, M. (2016). Effects of various sample pretreatment procedures on ethyl glucuronide

quantification in hair samples: Comparison of positivity rates and appraisal of cut-off values. Forensic Science International, 267, 60–65. https://doi.org/10.1016/j.forsciint.2016.08.012

[26] Mueller, A., Jungen, H., Iwersen-Bergmann, S., Raduenz, L., Lezius, S., & Andresen-Streichert, H. (2017). Determination of ethyl glucuronide in human hair samples: A multivariate analysis of the impact of extraction conditions on quantitative results. Forensic Science International, 271, 43–48. https://doi.org/10.1016/j.forsciint.2016.12.011

[27] Alladio, E., Biosa, G., Di Corcia, D., Seganti, F., Salomone, A., Vincenti, M., Baumgartner, M.R. (2018) Systematic optimization of ethyl glucuronide extraction conditions from scalp hair by design of experiments and its potential effect on cut-off values appraisal, Drug Testing Analysis, electronically published, https://doi.org/10.1002/dta.2405

[28] Salomone, A., Pirro, V., Lombardo, T., Di Corcia, D., Pellegrino, S., & Vincenti, M. (2014). Interpretation of group-level factors from a large population dataset in the determination of ethyl glucuronide in hair. Drug Testing and Analysis, 7(5), 407–413. https://doi.org/10.1002/dta.1697

[29] Pirro, V., Di Corcia, D., Seganti, F., Salomone, A., & Vincenti, M. (2013). Determination of ethyl glucuronide levels in hair for the assessment of alcohol abstinence. Forensic Science International, 232(1–3), 229–236. https://doi.org/10.1016/j.forsciint.2013.07.024

[30] R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

[31] Hinkley, D. V. (1970). Inference about the change-point in a sequence of random variables. Biometrika, 57(1), 1–17. https://doi.org/10.1093/biomet/57.1.1

[32] Killick, R., & Eckley, I. A. (2014). changepoint: An R Package for Changepoint Analysis. Journal of Statistical Software, 58(3). https://doi.org/10.18637/jss.v058.i03

[33] http://www.arpa.piemonte.it/rischinaturali/accesso-ai-dati/annali_meteoidrologici/annali-meteo-idro/banca-dati-meteorologica.html. Last access: March 26th, 2018.

[34] Brown, R. L., Durbin, J., & Evans, J. M. (1975). Techniques for Testing the Constancy of Regression Relationships over Time. Journal of the Royal Statistical Society. Series B (Methodological), 37(2), 149–192.

[35] Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting, 22(4), 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001

[36] Schräder, J., Rothe, M., & Pragst, F. (2012). Ethyl glucuronide concentrations in beard hair after a single alcohol dose: evidence for incorporation in hair root. International Journal of Legal Medicine, 126(5), 791–799. https://doi.org/10.1007/s00414-012-0729-z

[37] Luginbühl, M., Nussbaumer, S., & Weinmann, W. (2017). Decrease of ethyl glucuronide concentrations in hair after exposure to chlorinated swimming pool water. Drug Testing and Analysis. https://doi.org/10.1002/dta.2295

[38] Agius, R., Ferreira, L. M., & Yegles, M. (2012). Can ethyl glucuronide in hair be determined only in 3cm hair strands? Forensic Science International, 218(1–3), 3–9. https://doi.org/10.1016/j.forsciint.2011.10.001

[39] Meier, U., Briellmann, T., Scheurer, E., & Dussy, F. (2017). Distribution pattern of ethyl glucuronide and caffeine concentrations over the scalp of a single person in a forensic context. Drug Testing and Analysis, 9(10), 1594–1603. https://doi.org/10.1002/dta.2186

Figure Captions

Figure 1. (a) Number of montly samples analyzed in the laboratory during the period 2009-2017. **(b)** Time series representing the monthly mean EtG concentration values.

Figure 2. (a) Seasonality component of the proposed model (dashed line) employed to describe the data and provide forecasts (six months ahead, solid line with triangles) for the time series representing the monthly mean EtG concentration values (solid line with circles). The October 2015 is the first month when the procedural change in the laboratory protocol has been employed. **(b)** Monthly temperature in Piedmont, averaged for location, day and hour [33].

Figure 3. The stability of regression parameters – cumulative sum of recursive residuals (CUSUM, on the left) as well as cumulative sum of squared residuals (CUSUM of square, on the right) do not wander outside the critical bounds at 5% significance level (indicated by the red dashed line), thus a model might be termed as stable.

Figure 4. The monthly means of EtG across 7 years of analyses performed in the laboratory along with the mean temperature profile in the region of Piedmont. Additionally, the seasonal component of the regression model is superimposed.

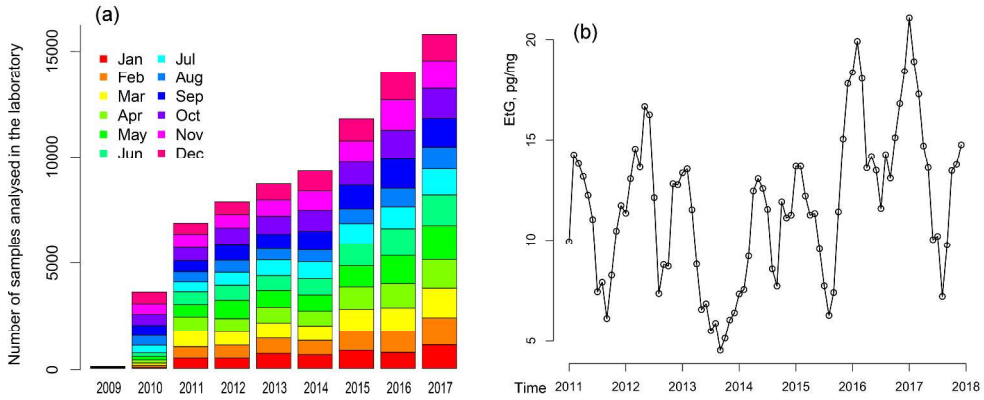

Figure 1. (a) Number of montly samples analyzed in the laboratory during the period 2009-2017. (b) Time series representing the monthly mean EtG concentration values.

393x169mm (300 x 300 DPI)

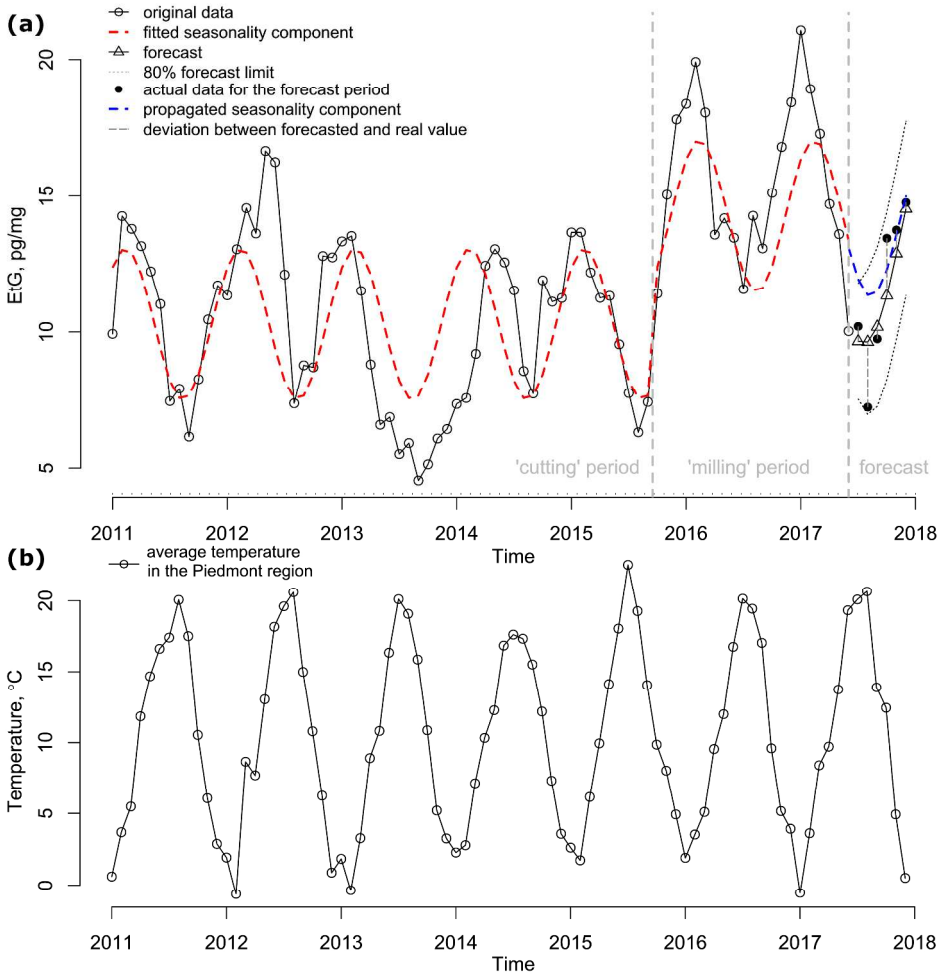

Figure 2. (a) Seasonality component of the proposed model (dashed line) employed to describe the data and provide forecasts (six months ahead, solid line with triangles) for the time series representing the monthly mean EtG concentration values (solid line with circles). The October 2015 is the first month when the procedural change in the laboratory protocol has been employed. (b) Monthly temperature in Piedmont, averaged for location, day and hour [33].

254x254mm (300 x 300 DPI)

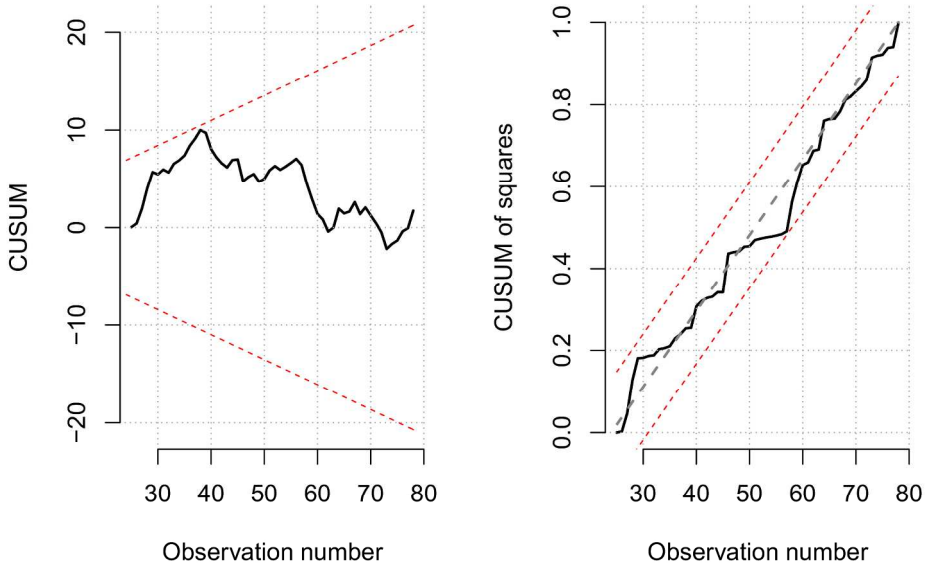

Figure 3. The stability of regression parameters – cumulative sum of recursive residuals (CUSUM, on the left) as well as cumulative sum of squared residuals (CUSUM of square, on the right) do not wander outside the critical bounds at 5% significance level (indicated by the red dashed line), thus a model might be termed as stable.

203x127mm (300 x 300 DPI)

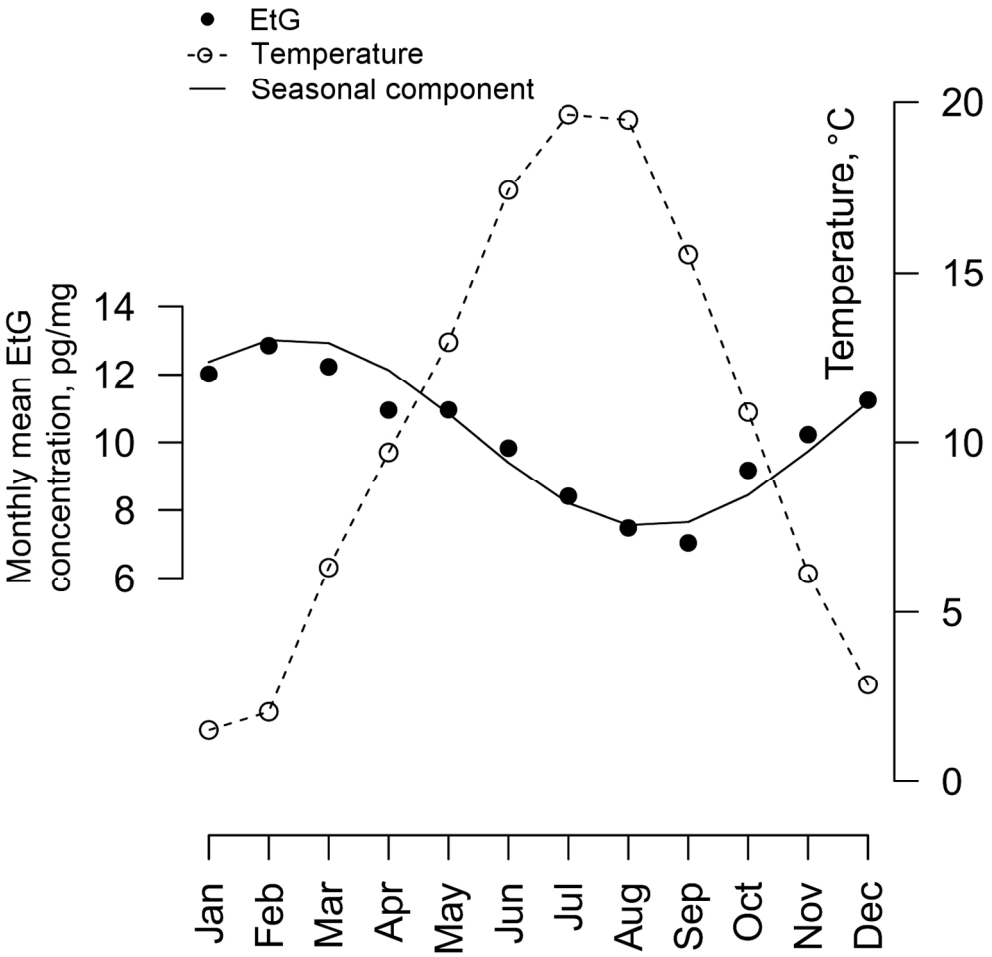

Figure 4. The monthly means of EtG across 7 years of analyses performed in the laboratory along with the mean temperature profile in the region of Piedmont. Additionally, the seasonal component of the regression model is superimposed.

127x127mm (300 x 300 DPI)