**Supplementary material.**

A Hitachi Elite LaChrom chromatograph coupled to a Hitachi L-2490 refractive index detector was used. In line with the objectives of the thesis, an isocratic method was developed with a SHODEX Sugar SH1821 column (ion exchange and size exclusion separation) at 60 ºC, 0.005M H2SO4 mobile phase, 0.5 mL/min flow rate, and RI detector at 40 ºC. It was developed with xylose, glucose, galactose, arabinose, levulinic acid, 5-HMF, furfural, acetic acid and formic acid standards and their respective retention times (Table 1, Figure 1). It was observed that all sugars present had similar retention times and coeluted into a single peak at 16.985 minutes and hence quantification was done by HPLC-IR as "Sugars" and not the individual species of glucose, xylose, fructose, galactose, and arabinose.

**Table 1.** Retention time for standards used.

|  |  |
| --- | --- |
| Platform chemicals | Tiempo de retención (minutos) |
| Glucose | 16,789 |
| Xilose | 17,179 |
| Galactose | 17,255 |
| Arabinose | 17,916 |
| HMF | 35,637 |
| Levulinic acid | 22,127 |
| Formic acid | 20,415 |
| Furfural | 51,349 |



**Figure 1**. Standards used for analytic method development.

For the validation of the chromatographic analytical method, and following the standards established by ISO 17025 in its technical note 17 - "Guidelines for the validation and verification of quantitative and qualitative test methods", the following validation parameters were evaluated: linearity, range, precision, accuracy, limit of detection and quantification, robustness, and sample stability.

**Linearity**: Linearity establishes whether the response of the detector is proportional to the analyte concentration over a given range. To evaluate the linearity of the method, a standard calibration curve for the quantification of sugars, formic acid, levulinic acid, furfural and HMF at 5 concentration levels (0.1, 1, 1, 3, 5, 8, 10 g/L) was performed in triplicate. Linearity was evaluated by means of a linear regression to obtain the equation of the straight line of the data obtained at the different concentrations of the curve. This allows obtaining the correlation coefficient (r), the intercept (b) and the slope of the regression line (m), on the other hand, the sum of residual squares was carried out to determine the homogeneity of the variances.

**Range:** Range evaluation in the validation of chromatographic analytical methods involves determining the concentration range within which the method is linear, accurate and suitable for the intended use. This ensures that the method provides reliable results over a specific range of concentrations. The process involves selecting concentration levels, preparing reference samples or standards, performing experiments, constructing a calibration curve, and evaluating the linearity of the calibration curve using statistical tools. The range is defined as the concentration interval in which the relationship between concentration and detector response is linear and accurate, usually (and in this work) taken as the range in which a coefficient of determination (R2) is greater than 0.99. The upper and lower limits of the range are set according to the results, defining the minimum and maximum concentration for which the method is valid.

For linearity, the correlation coefficients, intercept and slope shown in table 2 and figure 2 were obtained.

**Table 2**. Information of calibration curves for platform chemicals (PC).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **PC** | **Slope (m)** | **Intercept (b)** | **Pearson correlation (R)** | **Coefficient of determination (R2)** |
| **Sugars** | 262400487,8 | 75704583,1 | **0,9978** | **0,9956** |
| **Formic A.** | 41380930,2 | 3784997,3 | **0,9978** | **0,9955** |
| **Levulinic A.** | 150296520,1 | 3912845,5 | **0,9994** | **0,9988** |
| **HMF** | 311724039,2 | 41994984,9 | **0,9968** | **0,9962** |
| **Furfural** | 210362509,3 | 28492710,7 | **0,9981** | **0,9959** |

Table 2 shows the Pearson correlation coefficients and the coefficient of determination (R2), the R2 is above 0.99 in all cases, an aspect that demonstrates linearity in all the calibration curves carried out, and which is a suitable range for the implementation of this curve for quantitative purposes. On the other hand, and applying the data in figure 2, a regression analysis was made applying ANOVA, where the F value and the critical F value were obtained through the sum of squares, the value obtained is less than 0.05 in all cases, which allows us to conclude that the model has statistical significance, as the response coefficient test is less than 5%, we can safely say that it has linearity.



**Figure 2**. Calibration curves done for every PC.

On the other hand, the homogeneity of the variances was evaluated through the residuals of each of the calibration curves (figure 3). This allows us to affirm the homoscedasticity of the method, i.e. for this linear regression model, the estimation errors are constant throughout the different concentration points of the calibration curve, and having a constant variance makes the model more reliable.



**Figure 3.** Residuals of calibration curves.

From the above it can be concluded that the method has linearity (R2 greater than 0.99), homogeneity in its variances, as well as being a model with statistical significance, all this for the working range from 0.1 g/L to 10 g/L.

**Accuracy:**

**Repeatability of the system:** two standards were prepared from a multi-pattern solution at a concentration of 1 g/L and 10 g/L. They were injected into the HPLC-IR apparatus five consecutive times each and the coefficient of variation of the responses obtained in the five repetitions of each standard was calculated.

Tables 3 and 4 present the results for the repeatability of the system, in which a high precision in the retention times of the analytes and an acceptable repeatability within the criteria for the analytical responses of the compounds are observed.

**Table 3.** Repeatability of the system based on the retention time.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Acceptance criteria | Sugars(%) | Formic A. (%) | Levulinic A. (%) | HMF (%) | Furfural (%) |
| Retention time (Minutes) |
| 1 g/L | 10g/L | 1 g/L | 10g/L | 1 g/L | 10g/L | 1 g/L | 10g/L | 1 g/L | 10g/L |
| CV < 2% | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0,12 | 0,07 |

**Table 4.** Repeatability of the system based on the area.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Acceptance criteria | Sugars(%) | Formic A. (%) | Levulinic A. (%) | HMF (%) | Furfural (%) |
| Analytical response (Area) |
| 1 g/L | 10g/L | 1 g/L | 10g/L | 1 g/L | 10g/L | 1 g/L | 10g/L | 1 g/L | 10g/L |
| CV < 2% | 0,44 | 0,53 | 1,79 | 1,25 | 1,78 | 0,57 | 0,98 | 0,97 | 1,69 | 0,07 |

**Repeatability of the method:** It was evaluated by injecting the standard calibration curve and a sample of coffee hydrolysate liquid fraction five times for the quantification of the platform chemicals. Analyte concentrations and standard deviation were determined and compared with the acceptance criteria reported in literature to determine precision.

Table 5 shows the repeatability of the method for each analyte, the results obtained meet the acceptance criteria, implying that the analytical method developed is repeatable.

**Table 5.** Results for the repeatability of the method.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Acceptance criteria | Sugars(%) | Formic A. (%) | Levulinic A. (%) | HMF (%) | Furfural (%) |
| CV < 2% | 1,04 | 1,39 | 1,88 | 0,82 | 1,25 |

**Intermediate precision**: The same procedure for the determination of method precision was replicated, but performed by a different analyst, on a different day than previously performed. Data were taken from the two analysts, the percentage of the platform chemical compounds and the standard deviation of these were determined, and finally compared with the established acceptance criteria.

Table 6 presents the intermediate precision values for the different analytes on 3 different days with two different analysts, all data obtained are below the acceptance criteria and therefore the intermediate precision parameter is accepted.

**Table 6**. Results for intermediate precision for the method developed

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Analyst X | Day | Acceptance criteria | Sugars(%) | Formic A. (%) | Levulinic A. (%) | HMF (%) | Furfural (%) |
| 1 | CV < 2% | 1,04 | 1,39 | 1,88 | 0,82 | 1,25 |
| 2 | 0,81 | 1,52 | 1,64 | 1,24 | 1,56 |
| 3 | 1,20 | 1,22 | 1,52 | 1,15 | 0,74 |
| Analyst Y | Day | Acceptance criteria | Sugars(%) | Formic A. (%) | Levulinic A. (%) | HMF (%) | Furfural (%) |
| 1 | CV < 2% | 1,15 | 1,39 | 1,88 | 0,62 | 1,25 |
| 2 | 0,91 | 1,78 | 1,55 | 1,57 | 1,67 |
| 3 | 0,62 | 1,23 | 0,63 | 0,93 | 0,54 |

**Accuracy**: This was evaluated by enriching a matrix of biomass from a hydrothermal process at low temperature, with a multi-pattern standard, in order to have a known concentration of the analyte that could be quantified, so as to be able to quantify the amount added using the method and confirm that it corresponds to the amount added to the standard of the method. The matrix was enriched with multi-pattern standards of concentrations of 5 and 10 g/L, which were subsequently injected into the equipment together with the calibration curve, in order to quantify the multi-patterns, their recovery percentage and standard deviation of the 3 replicates of each concentration level. The values obtained were compared with acceptance criteria for this criterion reported in literature.

Table 7 shows the recovery percentages versus the concentration added to the matrix for quantification using the method developed. It shows that all the analytes are within the acceptance criteria and therefore the method is accurate.

**Table 7.** Accuracy of the results obtained for the method.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Acceptance criteria (% of recovery) | Sugars | Formic acid | Levulinic acid | HMF | Furfural |
| Concentration added to the matrix |
| 5 g/L | 10g/L | 5 g/L | 10g/L | 5 g/L | 10g/L | 5 g/L | 10g/L | 5 g/L | 10g/L |
| 90 %< x < 110% | 105 | 102 | 106 | 98 | 104 | 101 | 99 | 97 | 107 | 102 |

Limit of detection (LOD) and limit of quantification (LOQ): The calibration equation can be used to estimate the instrumental LOD. Using the estimate of LOD as the blank plus three standard deviations of the blank, the instrument response to a blank is taken as the intercept of the calibration curve (a), and the standard deviation of the instrument response is taken as the standard error of the calibration (sy/x). Therefore, from the calibration equation, **yLOD = a + 3 Sy/x = a + bxLOD** , then xLOD = **xLOD = 3 Sy/x/b.** This equation is widely used in analytical chemistry.

Several conventions have been applied to estimate the LOQ. Depending on the level of certainty required, the most common recommendation is to quote the LOQ as the blank value plus 10 times the repeatability standard deviation, or 3 times the LOD (which gives largely the same figure) or as 50% above the lowest fortification level used to validate the method.

The limit of detection and limit of quantification were calculated and recorded in table 8. The values of both LOD and LOQ are below the values of the measured samples, so it can be said that the detected concentrations were reliably obtained by the proposed analytical method.

**Table 8.** Detection and quantification limit for the method.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| PC | Sugars | Formic acid | Levulinic acid | HMF | Furfural |
| LOD (g/L) | 0,0139 | 0,0129 | 0,0169 | 0,0034 | 0,015 |
| LOQ (g/L) | 0,0436 | 0,0431 | 0,0563 | 0,0114 | 0,051 |

**Robustness:**

To evaluate this parameter, a multi-pattern standard was prepared at a concentration of 1 g/L and injected into the HPLC-IR equipment in duplicate at the chromatographic conditions proposed in the analytical method. After this, slight variations (±2 %) were made to the conditions (flow, temperature and mobile phase) and injected in duplicate to each of the modifications, in order to determine the peak retention times of the standards.

- Composition of the mobile phase: (±1mM)

Condition 1: H2SO4 4mM, Condition 2: H2SO4 6mM, condition 3: H2SO4 5mM

- Chromatographic column temperature (±2 ºC)

Condition 1: 58 ºC, Condition 2: 62 ºC, Condition 3: 60 ºC

- Mobile phase flow rate ( ± 0,1 mL/min)

Condition 1: 0.4mL/min, Condition 2: 0.6mL/min, Condition 3: 0.5mL/min

**Table 9.** Results for robustness of the method

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Condition** | **Sugars** | **Formic A.** | **Levulinic A.** | **HMF** | **Furfural** |
| **Composition MP** | **Retention time (Minutes)** |
| H2SO4 4mM | 17,789 | 21,182 | 23,123 | 35,232 | 53,093 |
| H2SO4 5mM | 16,985 | 20,285 | 22,029 | 34,732 | 51,048 |
| H2SO4 6mM | 15,79 | 18,222 | 19,989 | 33,551 | 48,832 |
| Temp. Column |   |   |   |   |   |
| 58ºC | 16,989 | 20,232 | 21,023 | 34,723 | 50,089 |
| 60ºC | 16,985 | 20,285 | 22,029 | 34,732 | 51,048 |
| 62ºC | 17,001 | 20,320 | 22,103 | 34,791 | 52,099 |
| Flow FM  |   |   |   |   |   |
| 0,4mL/min | 18,934 | 22,238 | 23,183 | 36,239 | 52,8239 |
| 0,5mL/min | 16,985 | 20,285 | 22,029 | 34,732 | 51,048 |
| 0,6mL/min | 13,239 | 16,239 | 17,232 | 30,239 | 45,2392 |

Table 9 shows that the method is robust since the variations in retention times in no case exceed a CV of more than 1%, demonstrating that small variations in the system do not affect the elution of the compounds.

**Sample stability:**

Solution stability was assessed by preparing a sample solution of hydrothermal coffee hydrolysate and multi-pattern solution (10g/L) that was injected into the HPLC-IR equipment every hour for a period of 12 hours, to quantify the amount of platform chemicals at each time, and thus observe the time the solutions can remain prepared without changing their concentration versus the initial time.

All samples taken over the 12 hours maintained their physicochemical characteristics and did not change in the concentration of the analytes or in the retention times of the analytes, so the samples were stable over the injection times.

It can be said that the method developed to quantify PQP by HPLC-IR is linear, repeatable, accurate, precise, robust and stable over time. The LOD and LOQ are also reported to ensure that all measurements obtained are above these values so that they can be reliably reported.