IDENTIFICATION OF CHEMICAL COMPOUNDS RESPONSIBLE FOR THE SPECIFIC FLAVOUR OF PINK TOMATO

Tiago Baltasar Monteiro Rodrigues

Internship report

Master's degree in Agricultural Engineering

Department of Geosciences, Environment and Spatial Planning

2020

Supervisor

Susana Maria Pinto de Carvalho, Assistant Professor, FCUP

Co-supervisors

Anne Marie Schoevaars, Senior Researcher, Enza Zaden

Jos Hageman, Assistant Professor, Wageningen University

Acknowledgments

First of all, thank you to Professor Susana Carvalho for the guidance and backing throughout the years. I would like to extend this to all the teachers, assistants and colleagues of the Master in Agricultural Engineering, for making this possible.

To my family, for the eternal support and for doing their best to provide me with the best working conditions possible when the pandemic turned everyone's world upside down. To my girlfriend, for keeping me honest and hard at work through kindness and compassion, even during the lowest times.

To Marinka van Arendonk, who I never got the pleasure to meet personally but whose work and research on the previous part of this project was crucial and irreplaceable in allowing me to continue carrying the torch.

Finally, I want to thank Enza Zaden B.V. for giving me the opportunity to conduct my Internship and all the Biotechnology colleagues in the person of Doctor Anne Marie Schoevaars, I could not have asked for a better supervisor to lead me in a completely new environment and a different country. For all the dedication, help and encouragement, for always believing in me, thank you. Thank you also to Professor Jos Hageman (Biometris, Department of Plant Sciences, Wageningen University), for his extraordinary practical and problem-solving mindset and for putting up with me when all his teachings became overwhelming

All in all, thank you to everyone that was in any way part of my journey throughout my master's degree and this thesis.

Abstract

Pink tomatoes represent a specialty and novelty within the staple of fresh products that are tomatoes. Their upcoming nature means that there are few bibliographic resources available that explore the biological differences between these and the more common red varieties. As such, development programs aiming to improve consumer traits face a lack of knowledge regarding how their breeding process and the choices of material affect consumer perception. It was with the intent of streamlining and clarifying a custom toolset for the improvement of Pink tomato varieties at Enza Zaden that we developed this project. By collecting data in seven different rounds from non-volatile and volatile variables and coupling it with responses from tasting sessions, we were able to use different data modelling techniques to assess the predictive value of the quantitative traits on consumer traits and recommend an approach to the breeders at Enza Zaden. Due to the volume of variables added by the secondary metabolites we used not only a more common least squares approach, in the form of the well-known Forward Stepwise regression, but also the Least Absolute Shrinkage and Selection Operator or Lasso, representing a more modern penalized regression. The results of both techniques coupled with leave-one-out cross-validation (LOOCV) largely confirm the know dichotomy of the sweet and sour flavour basis of tomato, measured through well-known variables like Degrees Brix or Titratable Acidity and reinforced by the positive effect of Glucose and the negative effect of high Malic acid contents. Some newness comes in the form of the relevance of the positive effect from the content of Citric acid and 2-Isobutylthiazole, with this volatile hinting that there may be more information to be gained by further exploring the secondary metabolism.

Key words

Aroma, chemical compounds, flavour, data modelling, *Solanum lycopersicum*, quantitative traits, volatiles

Resumo

Os tomates cor-de-rosa representam uma tipologia nova dentro do produto fresco elementar que são os tomates. O seu cariz recente implica que existem poucos recursos bibliográficos disponíveis que exploram as diferenças biológicas entre estes e as variedades vermelhas mais comuns. Como tal, os programas de desenvolvimento que visam melhorar as características mais ligadas ao consumo final enfrentam uma falta de conhecimento sobre a forma como o seu processo de reprodução e as escolhas de material afetam a perceção do consumidor. Foi com o intuito de racionalizar e clarificar um instrumento personalizado para a melhoria das variedades de tomate rosa na Enza Zaden que desenvolvemos este projeto. Ao recolher dados em sete rondas diferentes de variáveis não voláteis e voláteis e acoplá-los com respostas de sessões de degustação, pudemos utilizar diferentes técnicas de modelação de dados para avaliar o valor preditivo dos traços quantitativos sobre a experiência do consumidor e recomendar uma abordagem aos *breeders* da Enza Zaden. Devido ao volume de variáveis adicionadas pelos metabolitos secundários, utilizamos não só uma abordagem mais comum do método de mínimos quadrados, sob a forma da conhecida regressão *Forward Stepwise*, mas também o *Least Absolute Shrinkage and Selection Operator* ou Lasso, representando uma regressão penalizada, mais moderna. Os resultados de ambas as técnicas associadas à validação cruzada *Leave-one-out* (LOOCV) confirmam em grande parte a dicotomia conhecida da base de sabor agridoce do tomate, medida através de variáveis bem conhecidas como Graus Brix ou Acidez Titulável e reforçada pelo efeito positivo da glicose e pelo efeito negativo de altos teores de ácido málico. Alguma novidade vem sob a forma da relevância do efeito positivo do teor de ácido cítrico e do metabolito 2-Isobutiliazol, com este indicando que pode haver mais informação a obter explorando o metabolismo secundário.

Palavras-chave

Aroma, compostos químicos, modelação de dados, sabor, *Solanum lycopersicum*, traços quantitativos, voláteis

List of figures

Figure A30 - Results from the Forward Stepwise Regression with LOOCV. On the top response variable Sweetness, on the bottom Sourness. Heated trials data (1 to 4). On the left, the average M.S.E. for each level of model complexity for all n number of models. On the right, the number of times a predictor was selected for the desired level of complexity, decided on basis of the left image, for all n models.

Figure A31 - Results from the Forward Stepwise Regression with LOOCV. On the top response variable Juiciness, on the bottom Firmness. Heated trials data (1 to 4). On the left, the average M.S.E. for each level of model complexity for all n number of models. On the right, the number of times a predictor was selected for the desired level of complexity, decided on basis of the left image, for all n models.

Figure A32 - Results from the Forward Stepwise Regression with LOOCV. On the top response variable Overall Liking, on the bottom Flavour Intensity. Non-heated trials data (5 to 7). On the left, the average M.S.E. for each level of model complexity for all n number of models. On the right, the number of times a predictor was selected for the desired level of complexity, decided on basis of the left image, for all n models.

Figure A33 - Results from the Forward Stepwise Regression with LOOCV. On the top response variable Sweetness, on the bottom Sourness. Non-heated trials data (5 to 7). On the left, the average M.S.E. for each level of model complexity for all n number of models. On the right, the number of times a predictor was selected for the desired level of complexity, decided on basis of the left image, for all n models.

Figure A34 - Results from the Forward Stepwise Regression with LOOCV. On the top response variable Juiciness, on the bottom Firmness. Non-heated trials data (5 to 7). On the left, the average M.S.E. for each level of model complexity for all n number of models. On the right, the number of times a predictor was selected for the desired level of complexity, decided on basis of the left image, for all n models.

Figure A35 - Results from the Forward Stepwise Regression with LOOCV. On the top response variable Overall Liking, on the bottom Flavour Intensity. All trials data. On the left, the average M.S.E. for each level of model complexity for all n number of models. On the right, the number of times a predictor was selected for the desired level of complexity, decided on basis of the left image, for all n models.

Figure A36 - Results from the Forward Stepwise Regression with LOOCV. On the top response variables Sweetness, on the bottom Sourness. All trials data. On the left, the average M.S.E. for each level of model complexity for all n number of models. On the right, the number of times a predictor was selected for the desired level of complexity, decided on basis of the left image, for all n models.

Figure A37 - Results from the Forward Stepwise Regression with LOOCV. On the top response variables Juiciness, on the bottom Firmness. All trials data. On the left, the average

List of tables

Table of Contents

Abbreviations

FI – Flavour Intensity

FIR – Firmness

H – Heated

JU – Juiciness

LASSO – Least Absolute Shrinkage and Selection Operator

LOOCV – Leave-one-out cross validation

MSE – Mean Squared Error

NC - Naringenin chalcone

NH – Non-heated

OL – Overall Liking

PC - Principal Component

PCA - Principal Component Analysis

PCC – Pearson Correlation Coefficient

SO – Sourness

SW – Sweetness

TA – Titratable acidity

Tmm – Tomimaru Muchoo

TSS – Total soluble solids

V - Variety

1. Introduction

1.1. The relevance of flavour in tomato

Sensory proprieties are becoming increasingly important among the quality parameters valued by consumers, especially when it comes to fresh produce. The 'Lexico' dictionary by Oxford University describes flavour as "the distinctive taste of food" [1] and Tomato (*Solanum lycopersicum* L.), as one of the most popular items of that category in supermarkets, has a well-known flavour and consumer experience associated to it. With the focus of the buyers slowly shifting from size to flavour, as the lack of the latter becomes the main source of complaint from customers [2-4], we see a growing market share in the small and medium-size tomatoes with superior taste, cherry tomatoes being the main example of that trend [4].

1.2. Breeding for flavour in tomato

During the long history of domestication and improvement of the crop, its flavour quality has declined, mainly because flavour has not been a high priority of growers and breeders. In fact since growers' income is usually measured by quantity, and since the growers are the breeders' clients, the latter have mainly focused on improving traits like yield, stress resistance and postharvest life, traits that, sometimes, run counter to good flavour, like citrate and malate contents [4], but resulted in cheap, all-year-round produce availability capable of withstanding harsh treatment during harvest, shipping and storage [2]. As an example, historically much effort has been poured into increasing the sugar content of tomatoes, but it has been difficult to break the inverse linkage between size and sugar, as higher sugar content is usually associated with smaller fruit size [2, 5].

This historical trend was very clear until the 1960s when genetic diversity got to its lowest point. This started to change during the 1970s in north-western European countries, following legislation and public pressure to limit the use of plant protection products, leading to the introgression of disease-resistant genes from wild relatives and the breeding of hybrids, with changes being registered across most tomato genes. A second boost in diversity happened in the last decade of the 20th century when German news media labelled Dutch tomatoes as "*Wasserbomben"* (water bombs), to describe the poor, watery-tasting product, impairing the exports of the Netherlands and resulting in an effort from breeders to balance agronomical traits and consumer traits. This caused a diversification of the volatile component's profile in tomatoes noticeable since the 2000s [6].

More recently, organic farmers, foodies and chefs have championed flavour of many fresh produces, tomatoes included [7]. That, associated with the willingness of customers to pay a premium price for full-flavoured tomatoes [3], has resulted in a shift in breeding programs,

product development and reformulation programs, which now put sensory properties amongst the most important parameters [8]. Those properties are, nonetheless, difficult traits to breed for because of their complexity, representing a big challenge for seed companies, who must provide their breeders with new tools, strong enough to resist the environmental impacts [4] and to allow selection for improvement or even maintenance of said traits, as tasting large numbers of fruits in the field is both impractical and, in most cases, impracticable [2].

1.3. Selection tools to breed for flavour

Chemically, flavour is the sum of a large array of texture properties and metabolites, resulting from primary and secondary pathways, that are measured by taste and olfactory receptors [2] and perceived differently by each individual due to differences in the human genome [9, 10]. In the case of tomatoes, sugars and organic acids play a big role in the base notes of this concept and it is believed that an appropriate ratio and concentration is needed as a base for a good tasting variety [3, 7]. Further complexity is provided by other metabolites, especially volatiles, that are mostly perceived by the olfactory system even in the smallest amounts and can have a big impact on the organoleptic characteristics of products, as our receptors have evolved to perceive small concentrations of potentially toxic volatiles [10]. Varieties with enough sugar exist but without improvements in volatile profile they will always be perceived as inferior [2]. Those fragrant compounds are formed throughout ripening as well as during tissue disruption and they originate from many substrates including lipids, carotenoids, amino acids, terpenoids and lignin, who are transformed by enzymes through a variety of pathways, influenced by developmental, physiological and environmental signals until they form the more than 400 know volatiles of tomato. From these, only 15 to 30 are believed to be present in enough amounts to influence flavour, and the ones mentioned vary from author to author [2, 3, 11, 12]. Those mentioned pathways are regulated by two broadly defined classes of genes: those that encode enzymes responsible for synthesis, and those encoding for output regulating factors [3]. Both pathways and genes are still largely unknown and so, while selecting for flavour has clear benefits, for that and all that was mentioned before, it remains a major breeding challenge in tomato [4].

1.4. Pink tomato flavour and the difference between red and pink phenotypes

Along with this innate challenge, comes another one, from the fact that Pink tomato is considered a tasty fresh tomato speciality variety, with a specific desired taste associated with it, especially popular in some Asian countries [5] and in Eastern Europe (Anne Marie Schoevaars, personal communication), adding to the noticeable difference in colour, which is one of the most important consumer traits in tomato [13].

In the past few years, several studies have tried to find the exact origin of the difference between red and pink tomatoes [5, 13, 14]. The phenotype was first described by Lindstrom [15], as a mutation on the Y gene of chromosome 1 responsible for a distinct colourless peel. Since this loci is responsible for the colour of the tomato, and this specific mutation is a recessive one, the phenotype has been referred to as the y phenotype [13]. Further research seems to indicate that the difference in colour is due to a lack of naringenin chalcone (NC), a yellow coloured, ripening-dependant flavonoid, which is the most important flavonoid in red tomatoes, amounting to 1% of the peel's dry weight [14]. The absence of NC means that the peel of the pink tomato is not yellow but almost transparent, resulting in a pinkish phenotype. When looking into the possible genes which encode for the regulatory proteins involved in the pathway of flavonoids in tomatoes, researchers were able to correlate the activity of MYB12 gene with the decrease in the expression of structural flavonoids during ripening [13]. Genetic mapping later located the gene on the chromosome 1 area where the y mutation was shown

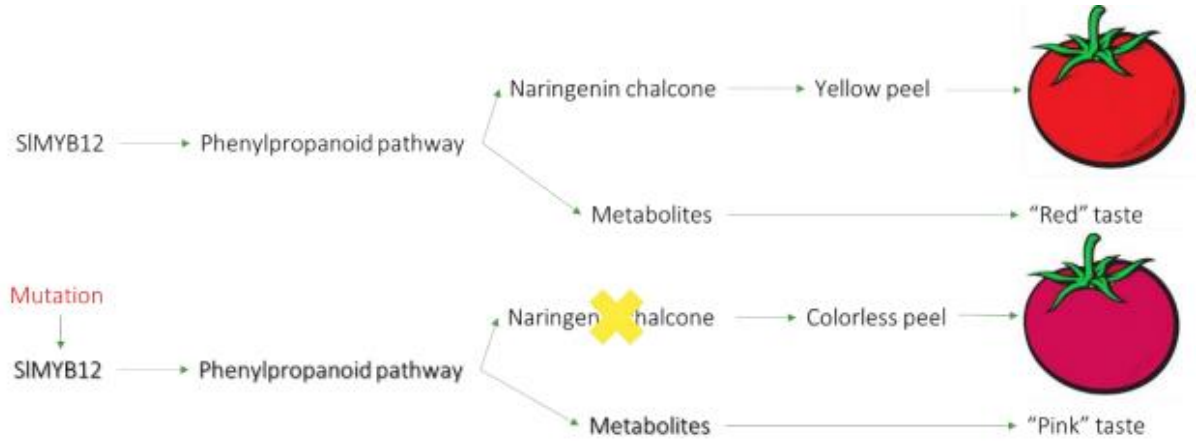


Figure 1 - Simple scheme of the influence of the SlMYB12 mutation. Source: Arendonk [15].

to have occurred and additional research with other colourless-peel mutants confirmed that a lesion in this regulator gene underlies the y phenotype [14].

The differences between both genotypes are not limited to the colour, for example, Adato [14] found that the peel itself is also thinner, less elastic and presents reduced water permeance, compared to red varieties, resulting in diminished shelf life. This acquires increased importance because of the cuticle and flavonoids roles in organ development, pollination and response to stress conditions, like radiation and pathogens [13, 14]. Also, large-scale metabolic and transcript profiling revealed a broad effect on both primary and secondary metabolism, not restricted to the peel and mostly related to the phenylpropanoids' biosynthesis [14].

These differences in metabolism between red and pink tomatoes have proven to be a barrier in the breeding of new pink varieties since the classic selection traits used by Enza Zaden to improve the flavour of tomatoes (consisting mainly of the measurement of brix, sugars, TA and tasting) has shown to be ineffective when breeding for pink varieties and the desired pink-tomato-taste [16].

1.5. Research aim

The upcoming nature of the pink tomato market in Europe and its enormous value worldwide, is an opportunity for Enza Zaden and for the entire breeding industry. Aroma and flavour are important in the acceptance of novelty products by consumers [10] and it is vital to develop this niche market around good tasting pink varieties. For that, a light needs to be shed on the bioactive molecules responsible for the specific desired sensory experience of consuming fresh pink tomato since the identification of a small number of metabolites give a much easier to measure variable than the tasting properties themselves [8]. Using sensory and chemical data we intend to model the complex matrices of metabolomes and identify those biomarkers that significantly track the perception of good 'pink tomato flavour'. The model will hopefully reveal to the breeders a compact toolset that will allow them to efficiently develop new varieties that answer the consumers' quality expectations and understand how their choice of genetic material and the production environment affect these perceptions [8].

2. Methodology

2.1. Trials

The data collection was split into seven trials due to infrastructure and workload limitations, but also to block as much as possible the effect of the growing season and time in sample preparation. The first four trials corresponded to what will be referred as Heated trials since they were grown off-season in a greenhouse with a completely controlled environment and were composed by varieties selected (or in the selection process) to be grown in these same conditions. The other 3 trials, called Non-heated, were also grown in greenhouses but with no temperature control since it was already tomato season. They consisted of material aimed at being produced in less technologically intensive greenhouses, like those found in the majority of Eastern and Southern Europe.

All Heated trials were composed of 21 varieties including 20 pink lines, among which the current pink market leader 'Tomimaru Muchoo', the upcoming variety 'Maluno' and Enza Zaden's new commercialized pink variety 'ENROSA'. The other pink lines were candidate hybrids from Enza Zaden and some of the corresponding parental lines. The trial also included the current red market leader 'Kanavaro' that was used as a benchmark (Table 1). The Non-

heated were composed of 26 varieties, all pink, including the market leader for the Non-heated segment ‘Pink Paradise’, another market reference of the same segment named ‘Thai Pink’ and the Heated market leader ‘Tomimaru Muchoo’. From each variety we collected 15 fruits per trial at the same stage of development, 3 per plant, allowing five biological repetitions consisting of three tomatoes each: one for volatile analysis, one for non-volatile analysis and one for the tasting session.

.

Table 1 - Varietal composition of every trial, R1 to R7. Varieties (V) in bold were selected for the tasting sessions. In pink, the pink market references ('Tomimaru Muchoo' represented by its abbreviation Tmm plus a one or a two to distinguish biological replicas), In red, the red market reference.

Heated				Non-heated		
R1	**R2**	**R3**	**R4**	**R5**	**R6**	**R7**
Tmm	**Tmm1**	**Tmm1**	**Tmm1**	**Pink Paradise**	Pink Paradise	Pink Paradise
	~~**Tmm2**~~	**Tmm2**	**Tmm2**	**Thai Pink**	Thai Pink	Thai Pink
V1	V1	**V1**	V1	**V16**	V16	V16
V2	**V2**	V2	V2	**V17**	V17	V17
V3	V3	V3	**V3**	V18	**V18**	V18
V4	**V4**	V4	V4	V19	**V19**	V19
V5	V5	**V5**	V5	V20	**V20**	V20
V6	**V6**	V6	V6	V21	**V21**	V21
V7	V7	**V7**	V7	V22	**V22**	V22
	V8	V8	V8	**V23**	V23	V23
	V9	V9	V9	**V5**	V5	V5
	V10	V10	**V10**	V24	V24	**V24**
	V11	V11	**V11**	V6	V6	**V6**
	V12	**V12**	V12	**V25**	V25	V25
	V13	V13	**V13**	V26	**V26**	V26
	V14	V14	V14	V27	V27	**V27**
	V15	V15	**V15**	V28	V28	**V28**
	Maluno	**Maluno**	Maluno	**V29**	V29	V29
	ENROSA	**ENROSA**	ENROSA	V30	**V30**	V30
Kanavaro	Kanavaro	Kanavaro	**Kanavaro**	V31	**V31**	V31
				V32	V32	**V32**
				V33	V33	**V33**
				V34	V34	**V34**
				V35	V35	**V35**
				V36	V36	V36
				Tmm	Tmm	Tmm

2.2. Chemical analysis

The chemical analysis and the selection process of the procedures that it incorporates are thoroughly explained by Arendonk [16] on previous research of this project, since it was during

that time that this was made. Nonetheless, below we include a brief description of that methodology, all conform the specific Enza Zaden analysis protocols from the Biochemistry group.

2.2.1. Organic acid analysis

The organic acids (citric, glutamic, and malic) were measured using several kits from Megazyme and the results were expressed in mg of organic acid per mL of blended tomato

2.2.2. Titratable Acidity

We produced a homogenate from which we neutralized the acids with a 0.05 M sodium hydroxide (NaOH) titrant solution, resulting in TA expressed as weight percentage of NaOH.

2.2.3. Total soluble solids

TSS was also measured from tomato homogenate with a refractometer that gave the amount of TSS expressed in Degrees Brix (°Bx).

2.2.4. Sugar analysis

The amounts of glucose, fructose, sucrose and total sugars were measured with High-Performance Liquid Chromatography (HPLC) with Hydrophilic Interaction Chromatography (HILIC) column and Refractive Index Detector (RID). The resulting chromatograms were analysed by Enza Zaden's biochemistry group, who calculated the final sugar concentrations in mg per g of tomato.

2.2.5. Volatile analysis

The volatile analysis was done by Headspace Gas Chromatography-Mass Spectrometry Solid Phase Microextraction (HS-GC-MS-SPME).

Sample preparation was performed according to Tikunov [17] with slight modifications. In short, an ethylenediaminetetraacetic acid (EDTA)-sodium hydroxide (NaOH) solution was prepared by adjusting 100 mM EDTA to a pH of 7.5 with NaOH. The EDTA helps persevering the volatile composition of the sample by preventing oxidation and increases the pH of the solution. Besides that, a 5 M sodium chloride (NaCl) solution and internal standard solution were made. The addition of NaCl serves to stop all enzymatic reactions and to drive the volatiles from the solution into the headspace. Next, a mixture of the EDTA, NaCl and internal standard solution was prepared with the ratio 10 EDTA.: 40 NaCl: 1 internal standard. Subsequently, 1 gram of frozen fruit powder was transferred to a 20 mL headspace vial that was subsequently closed and incubated at 30 °C for 10 minutes. The incubation step is included to thaw the sample and to simulate the chewing process that normally occurs when eating tomatoes and serves to stimulate volatile production. Then, 1 mL of EDTA-NaCl-

Internal standard mix was added. Lastly, the vial was closed and sonicated for five minutes, which improves the reproducibility of the analysis. After the sample processing, the samples were stored at 4 °C until the analysis was performed.

All chromatographs were integrated using a very simple, unfiltered method only to find out, for each trial, which sample had the highest number of peaks. This was done using the Agilent Mass Hunter Qualitative Data Analysis software. With this information, we were able to use the discriminated sample in the construction of a custom method for every trial, albeit keeping the same general properties. The method and this second analysis were made in the Agilent Mass Hunter Quantitative Data Analysis, where some manual integration and removal of compounds was also done to make sure that all compounds were being measured at the same point for every sample and to avoid repeating compounds. After these steps, all areas were divided by the respective area of the internal standard for each sample, to account for any difference that could occur in between readings, resulting in what is known as the response factor. For this, the internal standard needed firstly to be identified in every trial since its acquisition time was found to range between 25 and 27 minutes, and lastly, after the division was made, it could be removed from the data set, as well as any other compounds with Silicate bases that were simply part of the columns preparation.

To create a more robust and compact final data set, and to allow for the comparison and correlation of the data obtained in the different instances, all trials were aligned with one another, using for that effect the retention time of the compounds, the m/z value and identification libraries when the former two were not sufficiently clear. After positively associating the compounds between trials, we removed all those that were not present in more than half of the trials (four out of seven). From this resulted a list of 13 volatiles, including all those that were mentioned in the previous research on this project: Eugenol, 2-Isobutylthiazole, 2,4-Decadienal and 2-Hexenal [16]. 2,4-Decadienal [11, 17-19] and 2-Isobutylthiazole [3, 11, 12, 17, 18], are popular tomato-flavour enhancers, on the other hand, Eugenol [6, 12, 18-20], Hexanal [11, 12, 17, 20] and 2-Hexenal, [11, 17-19], have apparently a negative effect on the consumer perception. The other previously marked as positive volatiles in this final set were: β-Ionone [3, 11, 17, 18], Sulcatone, usually referred as 6-Methyl 5-Heptene 2-one [4, 5, 11, 12, 17, 18, 20, 21] and Geranylacetone [4-6, 11, 12, 17, 21], also a tomato flavour enhancer. Less known volatiles in tomato included Citral [20] and β-Cyclocitral [19, 20]. The remainder volatiles, 2-Octenal [12] and the monoterpenoids Perillene [22] and Carveol [23], although known volatile compounds associated with fruits and vegetables, were not referred as significant flavour compounds in the reviewed tomato literature (Table 2).

Table 2 - Compounds speciffically mentioned as associated with flavour in literature. In the described effects, a (+) or a (-) indicates a positive or negative correlation with the sensory responses from the references that do not further describe the sensory experience.

Compound	Reference	Effect
1-Nitro-2-phenylethane	[12, 20, 21]	floral
1-Nitro-3-methylbutane	[12, 21]	+
1-octen-3-one	[19]	mushroomy
1-penten-3-one	[4, 12, 19, 20]	green, spicy
2,4-Decadienal	[12, 18-20]	enhancer, fatty
2,4-Heptadienal	[18-20]	enhancer, fatty
2,4-hexadienal	[19, 20]	fresh, fruity, green
2,6-nonadienal	[19]	+
2-heptenal	[12, 19, 20]	green
2-Hexenal	[12, 18-20]	green, leafy
2-Isobuthylthiazole	[4, 12, 18-20]	musty, enhancer
2-methyl-1-butanol	[4]	malty
2-methyl-2-butenal	[19]	green, enhancer
2-methylbutanol	[20]	roasted
2-Pentenal	[12, 18, 20]	green
2-Phenylethanol	[6, 12, 20, 21]	floral, fruity
2-phenylnitroethane	[6]	sweet, floral
3-Hexen-1-ol	[12]	+
3-hexenal	[19]	fresh green, sweet
3-hexenol	[19, 20]	green
3-Methylbutanal	[18]	-
3-methylbutanal	[19]	-
3-methylbutanal	[6, 20]	stale, musty
3-methylbutanoic acid	[20]	cheesy
3-methylbutanol	[6, 20]	fermented
3-Pentanone	[12]	+
4-decenal	[12]	+
4-hidroxy-2,5-dimethyl-3(2H)-furanone	[12, 19]	+
6-Methyl-5-hepten-2-ol	[12, 18, 20]	citrus, green
6-Methyl-5-hepten-2-one	[4, 5, 11, 12, 17, 18, 20, 21]	fruity
Benzaldehyde	[12, 18-20]	sweet, almondy, fruity
Benzyl cyanide	[12, 21]	+
β-Ionone	[4, 12, 19, 20]	floral
β-Cyclocitral	[19, 20]	sweet
Butyl acetate	[12]	-
Citral	[20]	citrus
Eugenol	[6, 12, 18-20]	spicy, unpleasant
Geranial	[18, 19]	citrus
Geranylacetone	[4-6, 11, 12, 17, 21]	enhancer, floral

Guaiacol	[4, 6, 18, 20, 21]	smoky, woody unpleasant
Heptaldehyde	[12]	+
Hexanal	[12, 18-20]	green, grassy
Hexenal	[4, 19]	sweet, almondy, apple
Hexyl acetate	[12]	-
Isobutyl acetate	[12]	-
Isovaleric acid	[12]	+
Isovaleronitrile	[12, 21]	+
Linalool	[18-20]	floral
m-Cymene	[20]	chemical
Methional	[6, 19, 20]	potato, enhancer, pungent
Methyl Salicylate	[4, 6, 18-21]	oily, green, minty
Nonanal	[19]	fresh, citrus
Nonyl aldehyde	[12]	+
p-Cymenol	[20]	spicy
Phenylacetaldehyde	[4, 12, 18-21]	green, floral, honey
Phenylalanine	[4]	bitter
Phenyl acetate	[12]	-
Proline	[4]	sweet
Salicylaldehyde	[12]	-
Serine	[4]	sweet

In some of the trials, one or two of the volatiles (Table 2) had not been measured. This has to do with how the method building in the Agilent Mass Hunter Quantitative Data Analysis works since the program only looks at the compounds found in the sample used for the construction of the method. To solve this and not have missing data for an entire compound on an entire trial, we did a new targeted measurement, arranging new methods with the same settings but using samples for the model that we knew had the desired compound. Finally, we identified the missing compounds and labelled the variables accordingly.

2.3. Sensory analysis

The tasting sessions took place at Enza Zaden's tasting facilities and were composed by non-trained panels of tasters. The room was white and well-lit to avoid any visual interference, the panellists were seated separated from each other and behind a door that opened only to receive and return the materials for the taste session, allowing total isolation. Each member of the panel tasted a maximum of 8 selected samples (Table 1) to avoid desensitizing their taste receptors, while intercalating these with water and crackers. Samples were composed of two halves of different tomatoes from the same variety to account for possible differences in texture and ripeness. Because colour is such an important factor in the quality perception, panellists were always required to use glasses with red lenses during the tasting. All sessions

had an approximate number of 40 panellists that were tasting the fruits four days after they had been received, as all the chemical measurements were done prior to the tasting. The formularies filled by the participants asked for six parameters for each sample, Overall Liking on a scale from one to nine, Flavour intensity, Sweetness, Sourness, Juiciness and Firmness all in a penalized scale ranging from one to five where three is the optimal point and higher or lower values mean an excess or lack of the measured response variables, respectively.

Table 3 - List of compound variables with respective retention time for every trial (R1 to R7). Compounds marked as in yellow were individually integrated. Perillene could not be found on the non-heated trials.

	Retention time (minutes)							
Compound	**R1**	**R2**	**R3**	**R4**	**R5**	**R6**	**R7**	**Possible ID**
1	4,613	4,053	4,058	4,043	4,042	4,037	4,038	Hexanal
2	5,702	5,219	2,214	5,214	5,214	5,203	5,209	2-Hexenal
3	9,672	8,901	8,896	8,891	8,896	8,881	8,886	6-Methyl 5-Heptene 2-one
4	11,054	10,242	10,232	10,237	10,232	10,222	10,227	2-Isobutylthiazole
5	11,675	10,869	10,992	10,987	10,987	10,977	10,972	2-Octenal
6	13,113	12,291	12,286	12,288				Perillene
7	16,986	16,318	16,103	16,032	16,038	160,82	16,088	β-Cyclocitral
8	17,176	16,313	16,313	16,303	16,303	16,293	16,293	Carveol
9	17,993	17,135	17,13	17,125	17,12	17,115	17,115	Citral
10	18,625	17,757	17,757	17,746	17,741	17,736	17,731	2,4 Decadienal
11	22,775	21,902	21,907	21,881	21,881	21,871	21,876	Geranylacetone
12	23,648	22,739	22,739	22,729	22,71	22,713	22,719	β-Ionone
13	20,351	19,472	17,473	19,457	19,445	19,447	19,461	Eugenol

2.4. Data analysis

The building of the workflow for the data analysis tried to be as comprehensive as possible, trying to understand and accommodate the specificities of the data set using examples and references related to the specific area of metabolomics [3, 17, 24-28]. Consequently, we decided to divide the process into four different main parts: data pre-treatment, descriptive statistics, model construction and model validation.

2.5. Data pre-treatment

Data pre-treatment methods correct for values that could hinder the biological interpretation of the measured metabolites, by emphasizing and safeguarding the effects of the biological factors and their information, decreasing the influence of noise from non-accounted variables such as errors in data collection [26].

2.5.1. Zero values

This, in a first instance, consisted of the careful analysis of missing values and zero values. It was decided that observations for whom could not be measured any values, would be completely removed. This logic was for example applied to incomplete tasting formularies returned by panellists and to the entire sucrose measurements, as only a very small number of observations presented any reading, and these were always smaller than 0.03 mg/g. This is to be expected, given that we are working with tomatoes, which contain invertase that rapidly converts sucrose in glucose and fructose [29]. This led to the need of updating the formula for the total sugars to simply: *Total sugars* = *Fructose* + *Glucose*. Even though the variable name is technically incorrect we decided to keep it for the reasons mentioned above.

2.5.2. Missing volatile readings

We also agreed that the best approach to deal with eventual readings under the detection limit of any volatile would be to randomly generate values for those observations between zero and the lowest detection value for the entire experiment, as to avoid the negative heavy effects of inserting a hard zero in the models (Jos Hageman, personal communication). The effect of that random input proved to be negligible given that the lower reading found was 50 volts-minute (chromatogram peak area), so the introduced numbers varied between 1 and said value when indeed most of the components reached averages of peak area around the tens or even hundreds of thousands volts-minute.

2.5.3. Data transformation

Some transformations were also made in the data. All metabolites' values were log scaled to approach the data to a normal distribution, reducing the importance of possible outliers and removing heteroscedasticity [26]. In addition to the log transformation, autoscaling was used to centre the data around zero and give all the variables the same range, as to prepare it for Principal Component Analysis.

2.6. Descriptive statistics

The ANOVA (see 2.7) results were used to understand how much of the variation could be explained by the factors Variety and Panellist (Panellist referring to the people who participated in the taste sessions and representing our effort to try and measure the relevance of individual preferences), their effect size, correlation and what was regarded as noise, using the rstatix package [30]. This resulted in pie charts representing the importance of those three simplified dimensions (Figure A8). It also allowed us to identify outlier Panellists and to investigate their specific data to assess if they had understood what was being asked from them during the tasting sessions, or if they even found any differences in flavour.

QQ-plots and scatterplots of residuals against fitted values were used as graphical mean to test the normal distribution of residuals requirement for the aforementioned ANOVAs, as well as the equal variances across the groups (Figure A9 to Figure A15).

To broadly validate the procedures used in data transformation, correlation plots accounting for the p-value were made for the log-transformed data and considering alpha = 0.05. We used the Hmisc and corrplot packages [31, 32]. These were done in every trial and for the entirety of the data set (Figure A16 and A17).

Boxplots with every sensory variable were made for every trial to evaluate the distribution of the sensory data and eventual outliers (Figure A18 to Figure A25). Overall Liking used a 1 to 10 scale while the rest of the response variables varied from one to five, as such we divided all the values from the observations by two.

Lastly, we attempted to represent an eventual trial/batch effect by plotting all trials per variable and later adjusting for differences in average, resulting in a second graph this time adjusted to represent only the differences in variation (Figure A26 to Figure *A28*A28).

2.7. Modelling

After carefully reviewing the state of the art in modelling to solve metabolite related questions, we decided to use multiple methodologies and compare the results. The selected modelling techniques were Principal Component Analysis, Forward Stepwise regression and Lasso regression. These last two coupled with Leave One Out Cross Validation (LOOCV) as an exhaustive way to select the best model complexity at the same time as it allows for model validation.

2.7.1. ANOVA

First, we did some ANOVAs with R Studio on the trials of 2019 to validate a finding and methodology from last year. In three of the four trials, two samples of the pink market reference, ‘Tomimaru Muchoo’ (Tmm), were served to the tasters to measure the reliability and constancy of the tasting panel. We also suspected, from the chemical analysis, that the second sample from the second panel was probably underripe [13]. By doing ANOVAs in which we tried to measure how much difference there was between both samples in a model comprised of every sensory trait, we expected to validate or refute the data from the tasting sessions and the hypothesis of a underripe pink reference.

2.7.2. Principal Component Analysis

PCA is a very powerful multivariate statistical analysis tool, used to model, visualize and measure relations between samples and with and within variables, while retaining as much of

the variation possible through the introduction of new weighted variables named principal components [8, 33]. For our PCAs on the data set, we used the FactomineR and factoextra packages [34, 35], for R Studio. This resulted in biplots (Figure 2 to Figure 7) that use the neatly composed components as an axis to plot the relative positions of the variables and observations that make up the data, summarizing a big part of the variation present in the data set in one single plot, making it easier to check for biological assumptions, find errors and simply represent a big part of the information in as few dimensions as possible [33].

2.7.3. Forward Stepwise regression with LOOCV

Stepwise regression is a statistical method of regression fitting into a final model. This is done by adding or removing the explanatory variables following F-tests with predefined significance levels which assess the amount of added information by a given metabolite to the overall model. The forward variation of the stepwise regression works without allowing the removing of variables, starting with an empty-model and adding the metabolites in decreasing order of significance [28]. These regressions were done using R Studio, for every response variable, with the help of the leaps package [36] and repeated n times (number of observations) to allow for the Leave-one-out Cross-Validation. In LOOCV, every observation is left out of the training set once (n-1) and becomes the validation or testing set. Forward Stepwise regression is applied to the training set, returning a selection of predictors. Using these predictors, the left-out observation will be predicted. When all observations have been left out as a validation set we have n models with an associated prediction error [28]. With this information, we can count the number of times each predictor was used throughout all n models and use the Mean Squared Error (MSE), resulting from the difference between the predicted value and the observed value, to select the desired model complexity (number of predictors composing the model) (Figure A29 to Figure A37). The latter is done to avoid overfitting of the model and, obtain a more accurate model. It is expected that the predictive capacity for training set observations increases with the addition of more variables, its prediction accuracy will usually decrease for validation set observations. To avoid overfitting, we only ran this procedure on all Heated trials, all Non-heated trials or all trials, since this method does not work on data sets with more variables than observations for reasons that we will further elaborate next.

2.7.4. Lasso Regression with LOOCV

The Forward Stepwise Regression uses the ordinary least squares technique to solve the multiple linear regression problem, meaning that it tries to minimize the residual sum of squares. Ordinary least squares estimates are well known and unbiased but are prone to overfitting, especially in cases where the number of variables is somewhat large compared to the number of observations, resulting in poor prediction accuracy. In order to have an

alternative to this, we decided to also use a more recent penalization technique such as the Least Absolute Shrinkage and Selection Operator or LASSO [27], using the glmnet package for R [37]. Penalized or regularized estimates owe their name to the introduction of a penalty parameter that constraints the minimization of the residual sum of squares by fine-tuning the amount of shrinkage. They have risen in popularity because not only do they perform estimations but also, simultaneously, select the variables and the model complexity [27]. As such, the LOOCV was used with this method only to validate the procedure and to count the number of times a predictor was selected in all n models for every response variable (Figure A38 to Figure A43).

3. Results and discussion

3.1. ANOVA

With the ANOVA comparing the effect of Variety and Panellist, we found that, only in the second trial, sample one and two of Tmm were significantly different, for alpha equal to 0.05 (Table 1). This resulted in the removal of the second sample of the second trial, believed to be the underripe one and the validation of last year's methodology for all trials.

The effect size measured with the ANOVA showed that only trials two, five, six and seven had a Variety effect of over 20% (Table 4), meaning that in the other three trials less than one-fifth of the variation could be explained by the biological differences with the Panellist effect being able to predict more accurately the response variables. This could mean that the model is relying more heavily on personal preferences of the panellists or in other variables not accounted for in the ANOVA. This, in turn, could be due to the changing compositions of the panels but also of the material since some trials had more parental lines and first-generation hybrids, inducing flavour variability, compared to some more commercially focused material present in other tasting sessions.

The QQ-plots and scatterplots showed a linear distribution of the residuals, also, the scatterplot shows homogeneous spacing between the lines of fitted values and residuals, both validating the ANOVAs and the conclusions drawn from them (Figure A9 to Figure A15).

Table 4 - Effect Size of Variety and Panellist on each trial based on the ANOVA

	Effect Size		
Trial	**Variety**	**Panellist**	**Residual**
R7	34,6%	49,1%	16,3%
R5	33,8%	31,2%	35,0%
R2	22,7%	31,4%	45,9%
R6	21,3%	29,1%	49,6%
R4	8,7%	48,1%	43,2%
R3	5,7%	38,1%	56,2%
R1	5,2%	43,4%	51,4%

3.2. Correlations

Although the correlation plots are designed to not show statistically insignificant correlations, we see that most of the variables are somehow related to each other. Brix is usually correlated with the sugars, as to be expected, but also with the acid content since it is such an important part of the soluble matters of a tomato. The response variables are correlated with each other but also with sugar content and, more clearly, the citric content. In most of the trials, we also see a grouping of correlation between volatiles. This can be due to a few reasons. Volatiles, as secondary metabolites, are ripeness dependent, meaning that their pathways only start producing noticeable amounts when the fruit is close to its full maturity. So, we usually find high contents of one volatile correlated with high levels of the global volatile content. Secondly, the compounds associated with flavour originate from a very limited source of substrates, mostly terpenoids, lipids and carotenoids [2] so the genes controlling their pathways are somehow related [11]. Lastly, the synthesis of some metabolites is sometimes connected to the same zone of a chromosome activity signal, like it's the case with 6-Methyl 5-Heptene 2-one and Geranyl acetone [5]. We also find that the organic acids and some volatiles present the only clear cluster of negative correlations in the entire plot. This, once again, can be due to the ripeness-dependent nature of the volatiles, meaning that greener, more acidic tomatoes would usually have lower contents of volatiles. Apart from the graphical representation of the correlations, results and further discussion can be found for the correlation between Overall liking and all the other response variables in the summary tables of the models (Table 4 and 5) and modelling sections (2.7.3 and 2.7.4).

3.3. Boxplots

The boxplots for every trial show a constant relative position between response variables, Overall Liking, Sourness and Juiciness being usually higher. Overall Liking, a linear variable, had, as mentioned, its values divided by two due to the scale difference, so the boxes between 2.5 and 3, that we find in most plots, represent a more than average result for at least half of

the data. On the other hand, the rest of the response variables had their perfect value at 3, and more or less than that represents a lack or an excess, respectively, of that trait. We see that most of them, even Sourness, usually are lacking, which could explain why there are no higher values than 7.5 on the Overall Liking. Juiciness is the only attribute with its mean close to the optimal 3. Given the presence of parental lines in the material for the tasting sessions, and that pink tomato is not a common product in the Netherlands, this could explain such results. When it comes to the outliers, they are rare and relatively close to the whiskers, so we decided to keep them since we considered them acceptable given the subjective aspects of a tasting session and the serious implications of removing a data point.

3.4. Trial/Batch effect

While trying to understand the effect of time and batch separation when it comes to measurements and test sessions, it becomes clear that these variables had somewhat of an effect, even correcting for the residuals and aligning the averages, we can see a trial 1 (Figure A26) with increased variability of organic acids and the Non-heated trials with more spread contents of sugars and volatiles (Figure A27 and Figure A28), especially trial five and seven. This explains the behaviour of these trials in the forthcoming principal components analysis results and might also be the reason that there was enough variability to have models with a more relevant predictive value (see 3.7).

3.5. Principal Components Analysis

The Heated trials PCA biplot (Figure 2) represents a total of 48.1% of the variation in the data set composed of trials one to four. Principal component 1 is mainly composed by the volatiles 2,4 Decadienal, 6-Methyl 5-Heptene 2-one, Eugenol, Geranylacetone, Perillene and β-Ionone as well as all the organic acids, all on the positive part of the axis. On the negative side, we find all the sugars. This one component represents 33.1% of the variation while Principal component two takes the remainder 15%. The latter is composed by 2-Isobutylthiazole, β-Cyclocitral, 2-Hexenal and Hexanal on the positive part of the y-axis and by all the response variables negatively, although Sweetness is also close to the positive part of PC1. In this near 45-degree position we also find Brix on the negative part of both PCs and Citral, Carveol and 2-Octenal on the diagonally opposing quadrant of the biplot, influencing both PCs positively.

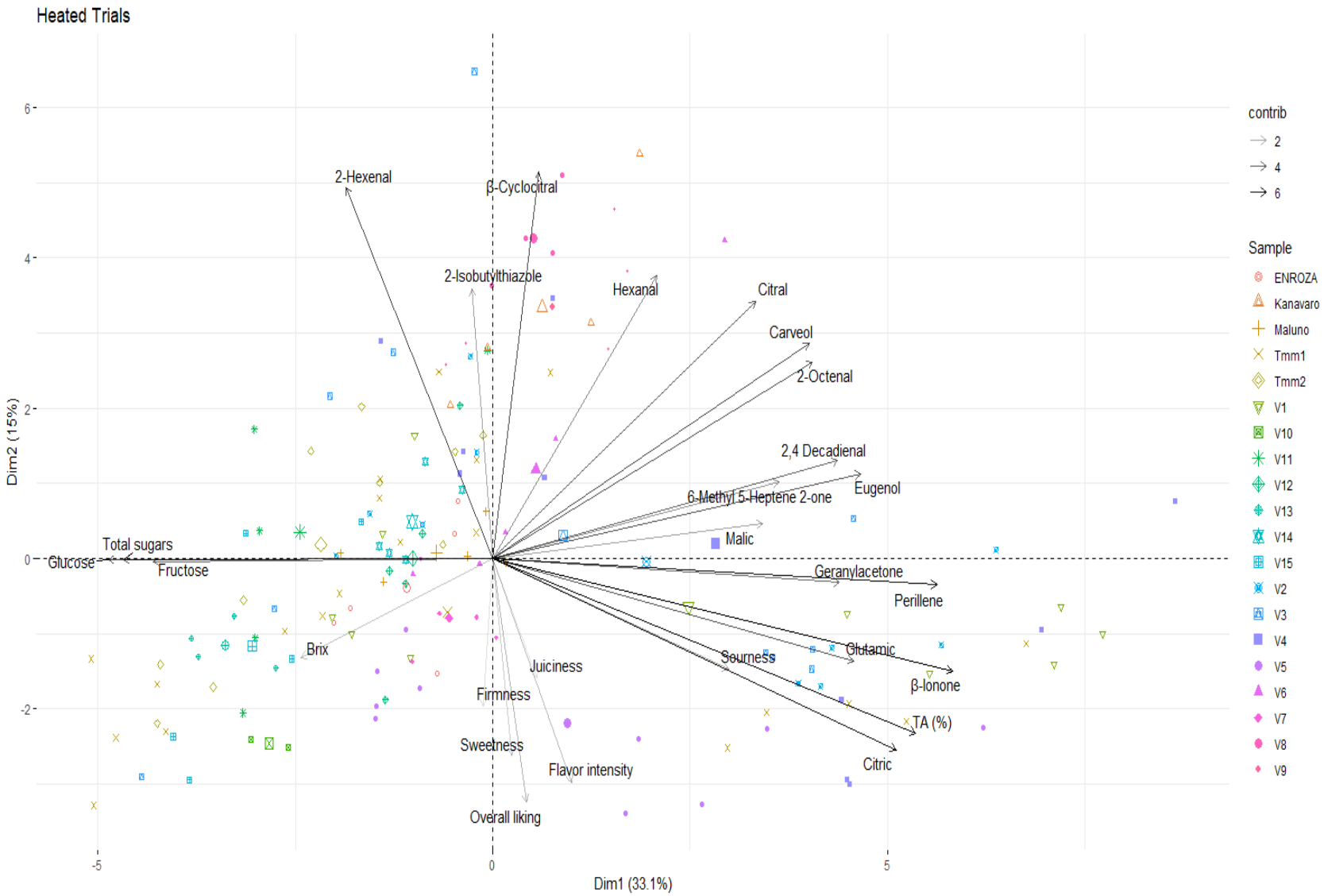


Figure 2 - PCA biplot of the Heated Trials (1 to 4). 20 varieties, including pink reference 'Tomimaru Muchoo' (Tmm). Shade of the arrows indicates the contribution of the variable towards the PCs.

This analysis broadly summarises the variation in the data set through an opposition between sugars and volatiles associated with a higher content of organic acids and by opposing the sensory variables with the remainder volatiles. On the high sugars and positive sensory variables response, we find the pink market references of Tmm (both Tmm1 and Tmm2) and 'ENROSA', closer to the middle. These are mostly spread along with the Brix variable. Both the other pink reference 'Maluno' and the red reference 'Kanavaro' are mostly on the negative side of Overall Liking and have fewer sugars and more organic acids than the first two references. The hybrids, especially the first four of the legend, appear to have the highest concentrations of volatiles and organic acids, consequently, the lower sugars and global appreciation, while the parental lines above Tmm are closer to 'ENROSA' and 'Maluno' on the biplot.

The PCA biplot of the Non-heated trials (Figure 3) summarizes 54.6% of the variation present in trials five to seven. Again, PC1 has a very strong influence of volatiles with Eugenol, 2-Isobutylthiazole and Geranylacetone being closer to the positive part of the axis but less important in the construction of the PC than the cluster composed of the rest of the volatiles, also with a positive effect. This principal component represents 32.7% of the variation and has no negative variables. Principal component two is then composed of all the non-volatile variables, including Glutamic and Malic that are close to a 45-degree angle, showing some more influence in PC1. All these variables are both on the positive part of the PC2 axis and PC1 axis, being in the first quadrant. Some of the variation explained by PC2 (21.9%) comes from the negative effect of the cluster of volatiles mentioned before since these are located in the negative part of the PC2 axis, making a 90-degree angle with most of the positive components of PC2 and ending up in the fourth quadrant.

In this PCA we are not able to find any of the relations previously exposed in the first PCA since most of the variables from the two PCs are at 90-degrees of each other, meaning there is little to no correlation between them. The pink market references are mainly distributed along the Sensory, sugars and organic acids variables, although they also have clear differences in volatile content, as 'Thai Pink' is completely on the left side of the biplot while 'Pink Paradise' is on the right one. The variation in PC1 and the volatiles that compose it seem to come mainly from five to six different hybrids located on the opposing extremities of the axis drawn by these variables.

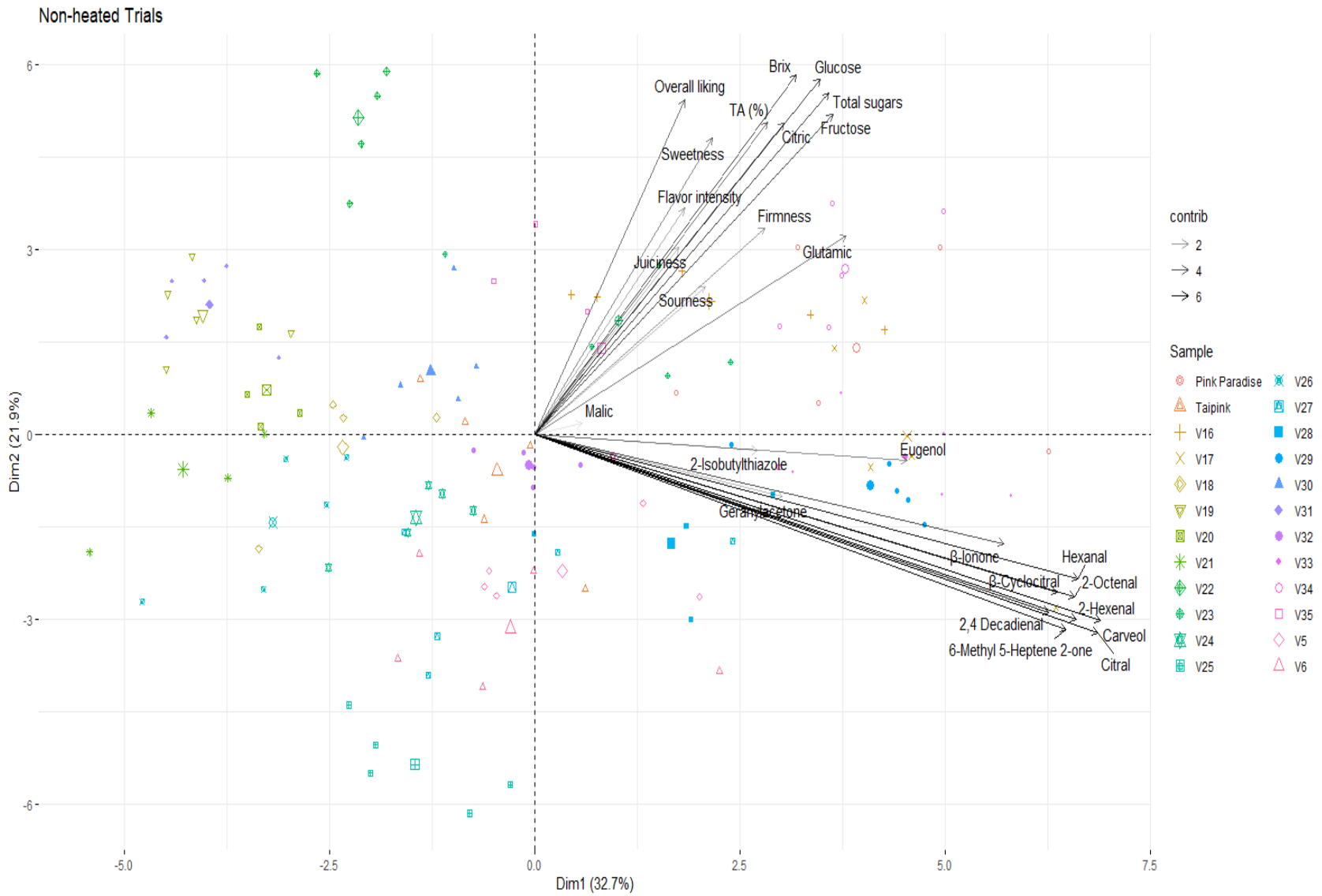


Figure 3 - PCA biplot of the Non-heated trials (5 to 7) with 25 varieties. Shade of the arrows indicates the contribution of the variable towards the PCs

The PCA biplot produced using all the seven trials (Figure 4) summarizes 52.4% of the variation found in the entirety of the data. PC1 is now even more clearly influenced positively by the volatiles since all of them can be found very close to its axis. On the negative part, we have the organic acids and TA, with Juiciness and Flavour intensity drawing close to a 45-degree angle to bring the variation explained by this principal component to a total of 35.6%. The rest of the sensory variables are more related to PC2, joining Brix and the sugars in the positive part of this axis. This PC is only very slightly negatively influenced by the Malic variable, while all response variables are on the negative part of PC1, contrary to the sugars, found symmetrically to these, in the first quadrant.

Now with all the data, we see some of the conclusions of the Heated PCA, like the tendency of the volatile content to somehow influence negatively the response variables. We also find another positive relation between sugar and sensory variables, like we did in the both first PCAs, as well as with these last ones and the acids content, also present in the previous analysis. Once more and unlike the Non-heated PCA, we see an opposition between sugars and organic acids, linking to the sweet and sour dichotomy of tomato flavour [3, 11, 20]. Looking at where the different trials can be found, we see that there is a cluster of the Heated trials (one to four) on the negative part of PC1, meaning higher organic acids content but also higher responses from the panellists and a lower sugar and volatile content. This is clear when compared to trial five and seven, that occupy most of the positive spread along the volatile and sugars axis. Since trial six is still relatively close to the Heated cluster, occupying the middle of the biplot, we decided to do a PCA based on this one plus trial one to four, hoping to keep as many observations as possible while making some less important variations more visible by removing the overwhelming effect of the other two trials.

The resulting analysis (Figure 5) explains 48.5% of the variation in a very balanced split between both principal components. PC1 represents 24.7% of the variation with a positive contribution from the organic acids and the two volatiles that were closer to these variables in the Heated PCA: Geranylacetone and β-Ionone (since Perillene was not found in Trial six). Sourness is also among this group of variables with the remainder of the response variables also influencing positively PC1 but negatively PC2, as such being found in the opposing fourth quadrant. All sugars draw a 45-degree angle with both PCs, influencing both negatively and running counter to the organic acids and associated volatiles. 2-Isobutylthiazole, β-Cyclocitral and 2-Hexenal, like what happened in the Heated biplot, run counter the response variables on the negative part of PC1 and positive of PC2. This is also were we can find Sulcatone, Citral, Carveol and 2-Octenal, although with a more important relation with the 23.8% of variation showcased by PC2. Eugenol and Hexanal have a clear positive effect on PC2 while

Figure 4 - PCA biplot of All Trials (Trial 1 to 7). Shade of the arrows indicates the contribution of the variable towards the PCs

2,4 Decadienal is close but influences more PC1 than the first two variables, completely opposing Brix.

Again, we see a negative effect of most volatiles on the response variables while sugars and acids maintain their apparent dichotomy when it comes to influencing the sensory response. Even if in this PCA, compared to the one with all trials, acids have now a positive and even more pronounced impact on the panellist perception than the sugars' contents. Observing the trials' positions, we see that the observations from Trial 1 are mostly responsible for the positioning of the acids and volatiles cluster. Trial two and three are close together in the middle of the biplot, spreading mainly along the response variables. Trial six still shows some of the symptoms of the Non-heated trials, variating largely along the rest of the volatiles and the sugars. In this last one, it is joined by Trial 4 that seems to constantly have the highest degree brix of all trials.

Backed by the results on the effect size (see 3.1) and the findings of the Lasso Regression (see 3.7), we added two more biplots, one for trial two (Figure 6) and one for trial seven (Figure 7). The PCA on both trials shows the tomato flavour enhancer, 2-Isobutylthiazole, as the only volatile closer to the response variables. Both analyses present the remainder of the volatiles on a close to a 90-degree angle with Overall Liking. Malic is shown to run counter to the sensory variables on both trials, although more clearly in Trial two, while the sugars and the other two other organic acids relate positively to these. Also, on this trial, the pink reference of TMM can be found close to the sensory variables, sugars and 2-Isobutylthiazole, recording the highest value of the volatile.

3.6. Forward Stepwise regression with LOOCV

The results for the individual models can be seen through their graphical representation: in Figures A29 to Figure A37 we see two plots for each response variable, the one on the left shows the evolution of the MSE associated with the predictions of model with different levels of complexity. Ideally, we want the lowest possible error, but we don't want to select too complex models, because their low MSE may be symptomatic of overfitting to this specific data. As such, we see that with Overall liking in Figure A35, there is a visible difference between 1 and two predictors, but a small one between two and three. After three, the error climbs again, peaking at four and slowly coming down until seven, maybe lower than what it was at two, even if only slightly. This behaviour indicates overfitting, so we want to select a model with two predictors since it is the last number with a clear improvement before the rise of the error. We set the complexity to two and the barplot on the right counts the predictors that were selected in all models in the first and second step of the forward Stepwise.

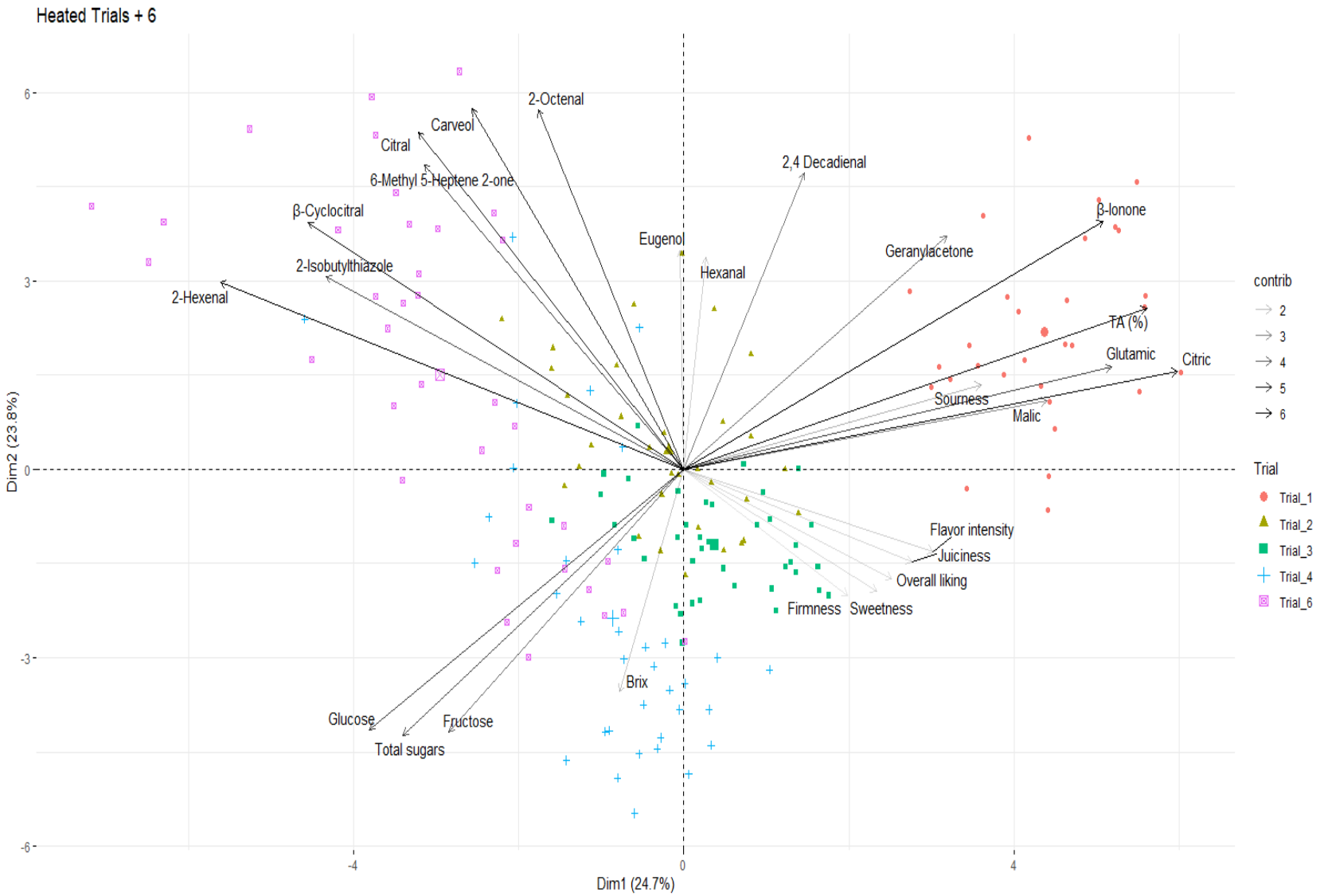


Figure 5 - PCA biplot of the Heated trial (1 to 4) with Trial 6. Shade of the arrows indicates the contribution of the variable towards the PCs

Figure 6 - PCA biplot of Trial 2. Seven varieties, including pink reference 'Tomimaru Muchoo' (Tmm). Shade of the arrows indicates the contribution of the variable towards the PCs

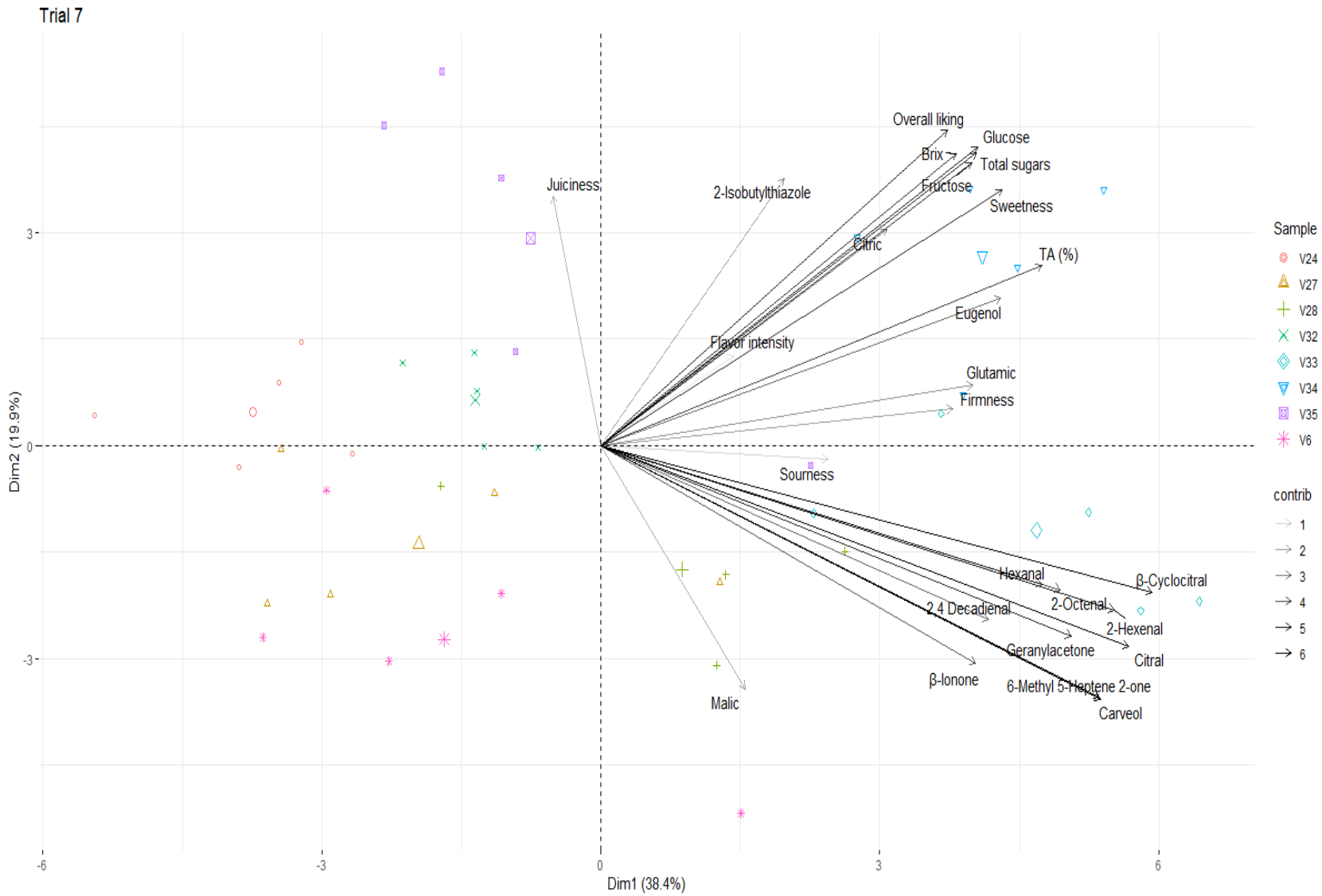


Figure 7 - PCA biplot of Trial 7. Eight varieties, including pink reference 'Tomimaru Muchoo' (Tmm). Shade of the arrows indicates the contribution of the variable towards the PCs

We are left with Glucose and Citric with the same number of selections, meaning that they were in all models always selected as 1 and two. This model has an associated predictive capacity related to the number of selected predictors and each predictor has a coefficient, allowing us to know if the effect they have on the response value is direct (positive) or indirect (negative). In this case, we have a model consisting of two variables with a direct effect on Overall Liking.

Given the amount of information for the combination of every single response variable with the three different data sets (Figure A29 to Figure A37), we opted for a summary table (Table 4). In this table, you may also find the correlation coefficients between the other response variables with Overall liking. This was added to highlight the most important models, those more related to the main response variable. As such, we see in numbers the information that could be found in the correlation of all trials (Figure A16 and Figure A17). The Pearson correlation (r) shows what was also somehow indicated in the different PCA's, this being the high correlation between Overall Liking and Sweetness, followed by a less clear one with Firmness and Flavour intensity. Sourness, although important in explaining the variation in the data, appears, once more, less related to Overall Liking as does Juiciness.

Given that this modelling exercise has a testing set different than the validation set, we decided to use the Q^2 as a way of representing the predictive value of the different models. Q^2 is similar to R^2 but instead of using the residual sum of squares, it expresses the predictive value in terms of the mean square error of prediction. It has a maximum value of one but can assume negative values, meaning that no model could be identified [26].

Unfortunately, low and negative results were all we got from this procedure. The best models only averaging 0.3 when, usually a good model should have values above 0.5 or 0.6, similarly to R^2. These were: predicting Overall Liking with the Heated trials data, predicting Overall Liking with the entirety of the data, or Flavour intensity with only the Non-heated trials. Although no satisfactory numbers of Q^2 were obtained, some attention needs to be drawn to the fact that citric is a part of 12 out of 18 proposed models, including the three mentioned models, leading to believe that this acid might reveal more predictive value in further research. Coupled to Citric we find the predictors Glucose and Geranylacetone. The sugar adding to the sweet and sour dichotomy and this volatile, usually associated with floral/green aroma, being the only slightly relevant compound with a negative effect.

Table 5 - Summary table of the Forward Stepwise Regressions with LOOCV and the Pearson correlation coefficients (r) between the other response variables and the respective OL The predictive value of the models is expressed in Q^2. All the variables included in the models are marked with a (+) or a (-) depending on the values of their coefficients. On the right-most row is a proposed ranking of correlation based on the PCC.

	Brix	TA (%)	Fructose	Glucose	Glutamic	Malic	Citric	Hexanal	Perillene	2-Octenal	Geranylacetone	Eugenol	Q2	Pearson	
OL H				+			+						0,309	1	1
OL NH	+						+						0,237	1	
OL				+			+						0,293	1	
FI H							+						0,165	0,466	4
FI NH							+				-		0,29	0,542	
FI							+				-	+	0,253	0,499	
SW H	+						+						0,106	0,696	2
SW NH	+												0,07	0,752	
SW	+						+						0,198	0,736	
SO H		+											0,17	**0,255**	6
SO NH							+						0,105	0,409	
SO		+											0,141	0,349	
JU H									+	-	+		-0,105	**0,235**	5
JU NH							+						0,136	0,468	
JU		+					+						0,08	0,412	
FIR H								+					0,081	0,495	3
FIR NH							+						0,12	0,592	
FIR							+						0,07	0,559	

3.7. Lasso regression with LOOCV

The Lasso regression reinforces the importance of the Citric and Brix variables and the sweet and sour flavour, although presenting even lower values of Q^2 than the Forward Stepwise Regression. The barplots from Figures A38 to A43 show the number of times a predictor was selected to integrate the model. We then look for the more relevant ones and calculate the predictive value of that model. Here (Figure A42, Overall liking) we see that Brix, Glucose, and Citric are always present in the n models made based on the data from all trials to predict for Overall Liking. There is a very big difference to Carveol on number 4, being selected to less than one-third of the models. From the coefficients, we know how these variables contribute to the model, directly or indirectly and we are also able to calculate the Q^2 for this specific desired model complexity. Again, to summarise this we made a table with the results (Table 6).

To do broader research, less constrained by the presence of the volatiles in most trials, we decided to run the Lasso in all the individual trials trying to predict only Overall Liking. The procedure was able to produce models for three trials, all with relatively higher predictive value

than the multi trials data sets. The results presented in these models must be considered with caution since, contrarily to the bigger sets, these have more variables than observations.

Trial five confirms once more the importance of both sweetness and sourness, with still a very low predictive value. Trial 2 and especially Trial 7 (Figure A44), on the other hand, introduce the known tomato flavour enhancer 2-Isobutylthiazole coupled to, once again, sugars and acids, as was indicated by the PCA on both trials (Figure 6 and Figure 7). The latter trial has a Q^2 value of 0.581 and seems more heavily focused on acids and the mentioned flavour enhancer as more predictive variables than maybe the raw perception of sweetness through sugars.

Table 6 - Summary table of the Lasso Regressions with LOOV. The predictive value of the models is expressed in Q^2. All the variables included in the models are marked with a (+) or (-) depending on the values of their coefficients.

	Brix	TA (%)	Fructose	Glucose	Glutamic	Malic	Citric	Hexanal	2-Hexenal	Sulcatone	Geranylacetone	Q2
OL H	+			+		-	+	-				0,206
OL NH	+						+					0,175
OL	+			+			+					0,238
FIH												-0,021
FI NH							+				-	0,173
FI												-0,01
SW H	+						+				+	0,102
SW NH	+			+								0,232
SW	+						+				+	0,103
SO H		+										0,023
SO NH												0
SO		+										0,011
JU H												0,008
JU NH							+					0,095
JU			+				+		-	-		0,17
FIR H												0,24
FIR NH							+					0,035
FIR		+		+			+				-	0,107

4. Conclusion and recommendations

4.1. Which are the chemical compounds responsible for pink tomato flavour?

As aforementioned, we failed to produce encompassing models with a satisfactory predictive value (Q^2 bigger than 0.5), nonetheless, looking at our results, we think it is clear that the sweet and sour dichotomy of tomato flavour [3, 11, 20] is still very important in pink tomatoes, especially when we look at the prevalence of Citric and Brix in all the models and TA in trial

seven, or how much Sweetness was correlated with Overall Liking and Flavour intensity, in both the PCA (Figure 2 to Figure 7) and the correlation plots (Figure A16 and Figure A17). This gains even more importance since the models were built around response variables with penalized scales (except for Overall Liking), which we know makes it more difficult to model. Also, although trial 7 is a small data set for such a large number of variables, it is clear from the bibliography [3, 11, 17, 18], from the PCAs and also from this specific model, that 2-Isobuthyliazole is a very relevant volatile in the consumer's perception of flavour. So, we believe that further research could find proof to cement these claims and produce a very streamlined toolset for the breeding process of pink tomato, consisting of Citric, either Brix, TA or both, and 2-Isobuthiliazole. This means that the traditional tools used by Enza Zaden for selecting varieties will very likely remain important, especially if we consider that degrees Brix and TA are easier to measure than organic acids or volatiles, but it appears that the extra effort would much likely yield positive results. Measuring organic acids, although it represents the addition of a completely different technique to the process, would seem to be at least the more practical compromise, given how many times and how relevant Citric appeared in our models. On the other hand, for all the current process associated with the GC-MS to be probably considered useful and not only time consuming, definitive evidence of a significative amount of value added by 2-Isobuthyliazole or any combination of volatiles would have to be found.

4.2. Recommendations

Due to the pandemic, we were unable to perform the 2020 trials, so we find ourselves in the peculiar position of recommending what we, in a way, expected to be able to do during this project: collect more data. The question is, how?

Firstly, if we look at the effect size on each trial (Table 4) we can see that all Non-heated trials had higher variety effect/biological effect than most of the Heated trials, excluding trial 2 (Table 1). This may be so for some reasons: the first tasting sessions were conducted on paper, making the sessions longer and increasing the chances of a human error, especially when introducing the data into the data sets. Secondly, some of the Heated tasting sessions repeat varieties leaving others out since they were focused on more commercial material, more available in this segment of production (controlled environment), while the Non-heated presented only once all the varieties for tasting, including more diversity in the sessions. A third point, still related with the volatiles concentrations variability, is that the Non-heated trials, due to the condition that gives them the name, are grown in a less controlled environment, more prone to differences in the environment, leading to different maturity and volatile contents, adding to the already higher diversity, compared to the Heated trials, due to the larger number of varieties (Table 1). Lastly, since the panellists participate on a voluntary

basis, there are different panel compositions during the trials inducing different results by itself, but also may be leading to the increase of variety effect on the latter trials being explained by the effect of a slow increase of tasting experience due to a larger pool of panellists, as it is more likely to find people that participated in more than one session at the end of the experiment.

One of the first solutions would be, as was already done at Enza Zaden, to do all trials directly on a digital medium, having panellists answering, for example, directly on a tablet. Then, it would be important to distribute equally the absolute number of varieties on trial and the relative proportion of parental lines, hybrids and commercial material on all sessions, maybe focusing more heavily on the first two since they should have higher differences in metabolites. In the future, and to obtain more reliable data, we would suggest that a roster of 8 varieties, comprised of genotypes in the most diverse phases of product development possible, should be consistently presented to the same panellists during multiple trials. The selection of varieties could also be done by looking at this project's PCAs and the different profiles of the measured variables and mixing extreme and central varietal clusters. When looking to reduce as much as possible the levels of noise in the data (shown by the importance of residual effects, Table 3) with non-trained panellists, the desired number of tastings per variety should be close to one hundred (Luís Cunha, Faculty of Sciences of Porto University, personal communication). This number can be brought down to half if the members of the tasting panel were to be introduced to different base tastes and would be, in a slight way, trained for the tasting sessions, giving more stable and relevant results that would compensate for less tastings [3]. Alternatively, using an electronic tongue and nose [10] or a highly trained panel [20] is maybe an option, though we believe this is a more suited approach for less market-oriented research.

The already mentioned non-linear scale of most response variables also seems to impede more significant results, constantly having lower predictive value than the linear response of Overall Liking. Testing with all linear scales directly in the formularies would be more desirable.

Our last recommendation would be to randomly attribute the samples to the batches being analysed, especially in the GC-MS. All samples were thoroughly frozen and thawed according to the described methodology, even so, if we look at the batch effect on the volatiles (Figure A27), it is clear to see that trials five to seven have a very pronounced effect. Although we hinted that this variation could be the reason why trial seven produced a relevant model albeit having very few observations, it is also likely that this batch effect present in all the Non-heated trials could have disturbed important patterns in the data, explaining why the models of these

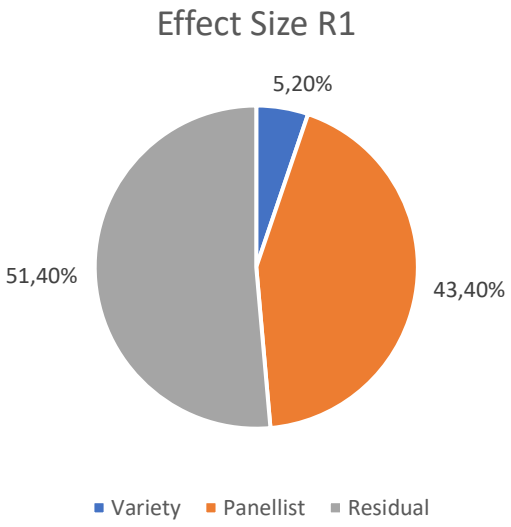
trials always performed worse than the Heated counterparts and why the overall model did not show better predictive ability even though having almost double the observations.

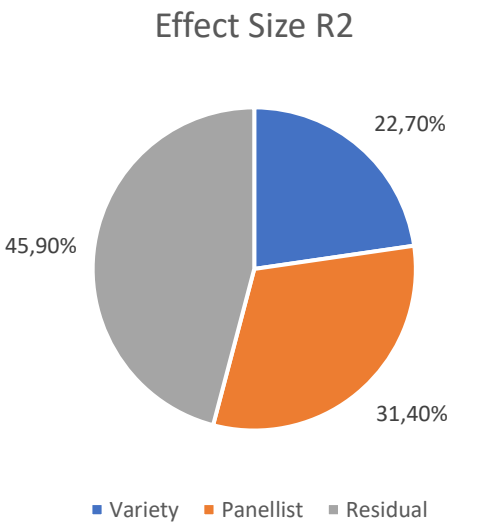
5. References

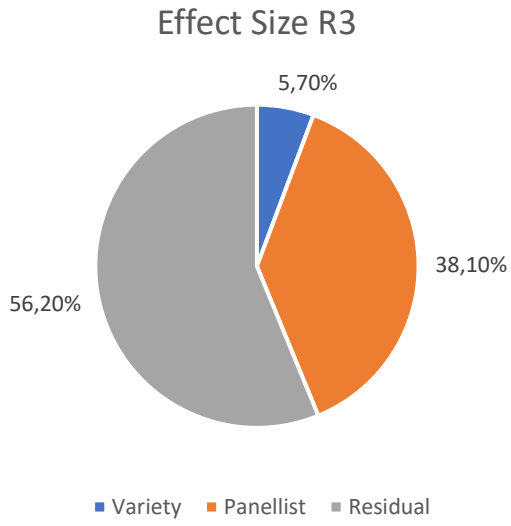
1. *Lexico powered by Oxford*. (accessed on Sepetmber 10, 2020), Oxford Univeristy: https://www.lexico.com/definition/flavour.
2. Klee, H.J., *Improving the flavor of fresh fruits: genomics, biochemistry and biotechnology.* New Phytologist, 2010. **187**: p. 44-56.
3. Tandon, K.S., et al., *Linking Sensory Descriptors to Volatile and Nonvolatile Components of Fresh Tomato Flavor.* Journal of Food Science, 2003. **68**.
4. Zhao, J., et al., *Meta-analysis of genome-wide association studies provides insights into genetic control of tomato flavor.* Nature Communications, 2019.
5. Zhu, G.T., et al., *Rewiring of the Fruit Metabolome in Tomato Breeding.* Cell, 2018. **172**(1-2): p. 249-+.
6. Schouten, H.J., et al., *Breeding Has Increased the Diversity of Cultivated Tomato in The Netherlands.* Frontiers in Plant Science, 2019.
7. Jabr, F., *Why some tomatoes taste better*, in *Sientific American*. 2012.
8. Thissen, U., et al., *A proper metabolomics strategy supports efficient food quality improvement: A case study on tomato sensory properties.* Food Quality and Preference, 2011. **22**: p. 499-506.
9. Shigemura, N., et al., *Genetic and Molecular Bais of Individual Differences in Human Umami Taste Perception.* Plos ONE, 2009. **4**(8).
10. Roelse, M., *Receptomics, design of a microfluidic receptor screening technology*, in *Bioscience*. 2019, Wageningen University: Wageningen.
11. Tieman, D.M., et al., *The Chemical Interactions Underlying Tomato Flavor Preferences.* Current Biology, 2012. **22**(11): p. 1035-1039.
12. Klee, H.J. and D.M. Tieman, *The genetics of fruit flavour preferences.* Nature Reviews Genetics, 2018. **19**(6): p. 347-356.
13. Ballester, A.R., et al., *Biochemical and Molecular Analysis of Pink Tomatoes: Deregulated Expression of the Gene Encoding Transcription Factor S1MYB12 Leads to Pink Tomato Fruit Color.* Plant Physiology, 2010. **152**(1): p. 71-84.
14. Adato, A., et al., *Fruit-Surface Flavonoid Accumulation in Tomato Is Controlled by a SIMYB12-Regulated Transcriptional Network.* Plos Genetics, 2009. **5**(12).
15. Lindstrom, E.W., *Inheritance in tomatoes*, in *Genetics*. 1925, Iowa State College: Ames, Iowa.
16. Arendonk, M.v., *Combination of selection tools is required to define the specific taste of pink tomatoes*. 2019, Master Degree from Wageningen University.
17. Tikunov, Y., et al., *A novel approach for nontargeted data analysis for metabolomics. Large-scale profiling of tomato fruit volatiles.* Plant Physiology, 2005. **139**(3): p. 1125-1137.
18. Kazeniac, S.J. and R.M. Hall, *Flavor Chemistry of Tomato Volatiles.* Journal of Food Science, 1970. **35**.
19. Krumbein, A. and H. Auerswald, *Characterization of aroma volatiles in tomatoes by sensory analyses.* Nahrung, 1998. **42**: p. 395-399.
20. Tikunov, Y., et al., *The genetic and functional analysis of flavor in commercial tomato: the FLORAL4 gene underlies a QTL for floral aroma volatiles in tomato fruit.* The Plant Journal, 2020.
21. Tieman, D., et al., *A chemical genetic roadmap to improved tomato flavor.* Science, 2017. **355**: p. 391-394.
22. *PubChem Database. Perillene, CID=68316*. (accessed on July 21, 2020), National Center for Biotechnology Information: https://pubchem.ncbi.nlm.nih.gov/compound/Perillene.
23. Cui, S.Q., et al., *Development of Fast E-nose System for Early-Stage Diagnosis of Aphid-Stressed Tomato Plants.* Sensors, 2019. **19**(16).
24. Ursem, R., et al., *A correlation network approach to metabolic data analysis for tomato fruits.* Euphytica, 2007. **161**: p. 181-193.

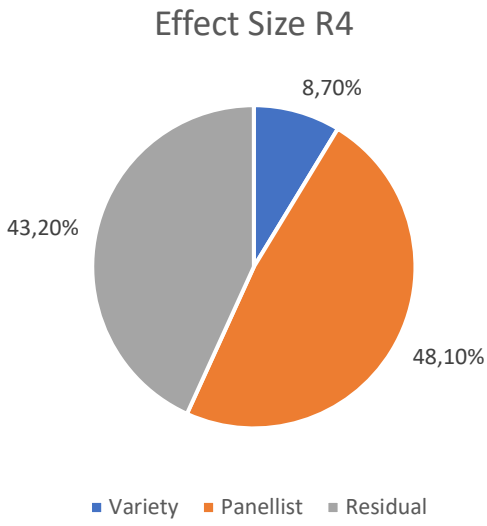
25. Broadhurst, D.I. and D.B. Kell, *Statistical strategies for avoiding false discoveries in metabolomics and related experiments.* Metabolomics, 2006. **2**.
26. van den Berg, R.A., et al., *Centering, scaling and transformations: improving the biological information content of metabolomics data.* BMC Genomics, 2006. **7**.
27. Menendez, P., et al., *Penalized regression techniques for modeling relationships between metabolites and tomato taste attributes.* Euphytica, 2012. **183**(3): p. 379-387.
28. Hageman, J.A., et al., *Robust and Confident Predictor Selection in Metabolomics*, in *Statistical Analysis of Proteomics, Metabolomics and Lipidomics Data Using Mass Spectometry*, S. International, Editor. 2017. p. 239-257.
29. Baldazzi, V., et al., *Inter-species comparative analysis of components of soluble sugar concentration in fleushy fruits.* Frontiers in Plant Science, 2016.
30. Kassambara, A., *rstatix: Pipe-Firendly Framework for Basic Statistical Tests*. 2020.
31. Harrel, F.E., *Hmisc: Harrel Miscellaneous*. 2020.
32. Wei, T. and V. Simko, *R package "corrplot": Visualization of a Correlation Matrix.* 2017.
33. Bro, R. and A.K. Smilde, *Principal Components Analysis.* Analytical Methods, 2014. **6**.
34. Lê, S., J. Josse, and F. Husson, *FactoMineR: An R package for Multivariate Analysis.* Journal of Statistical Software, 2008. **25**(1): p. 1-18.
35. Kassambara, A. and F. Mundt, *facotextra: Extract and Visualize the Results of Multivariate Data Analyses.* 2020.
36. Lumley, T., *leaps: Regression Subset Selection*. 2020.
37. Friedman, J., T. Hastie, and R. Tibshirani, *Regularization Paths for Generalized Linear Models via Coordinate Descent.* Journal of Statistical Software, 2010. **33**(1): p. 1-22.

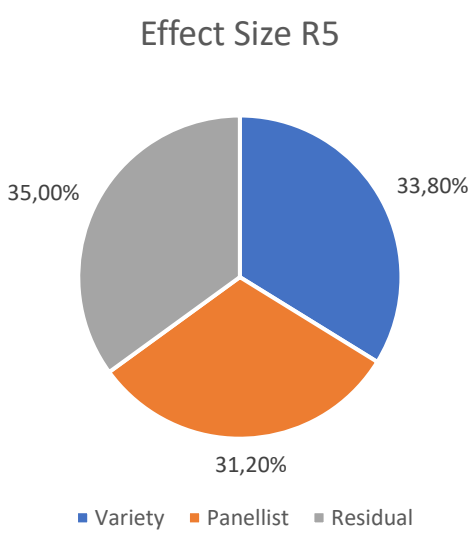
6. Appendices

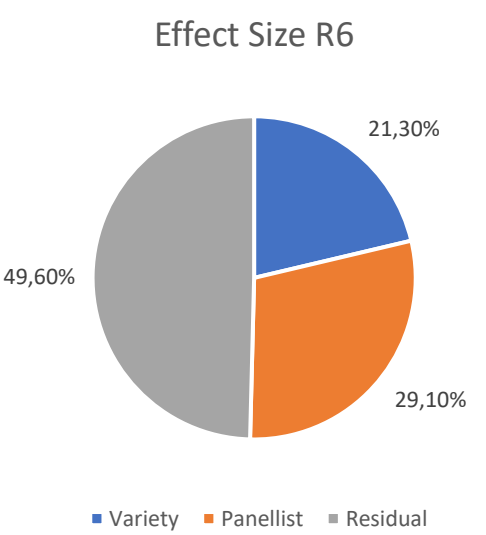












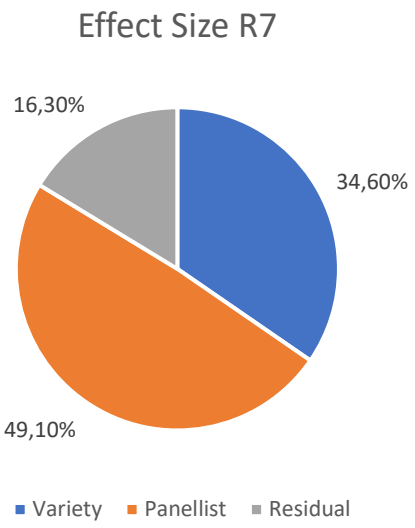


Figure A8 - Effect Size charts for each trial based on the ANOVA results on Overall Liking with the factors Variety and Panellist

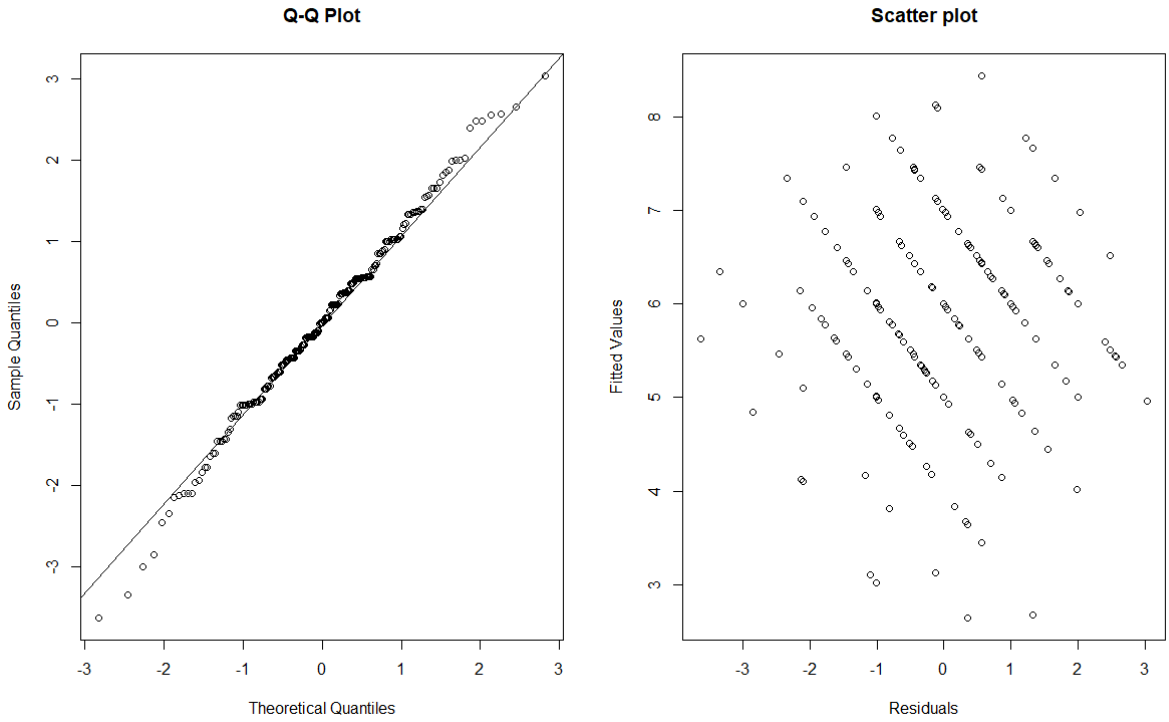


Figure A9 – (Trial R1) Normal Q-Q plot used as direct assessment of the requirement of normal distribution of residuals for validation of the ANOVA. The scatter plot assesses the equal variances across groups, another requirement of the analysis.

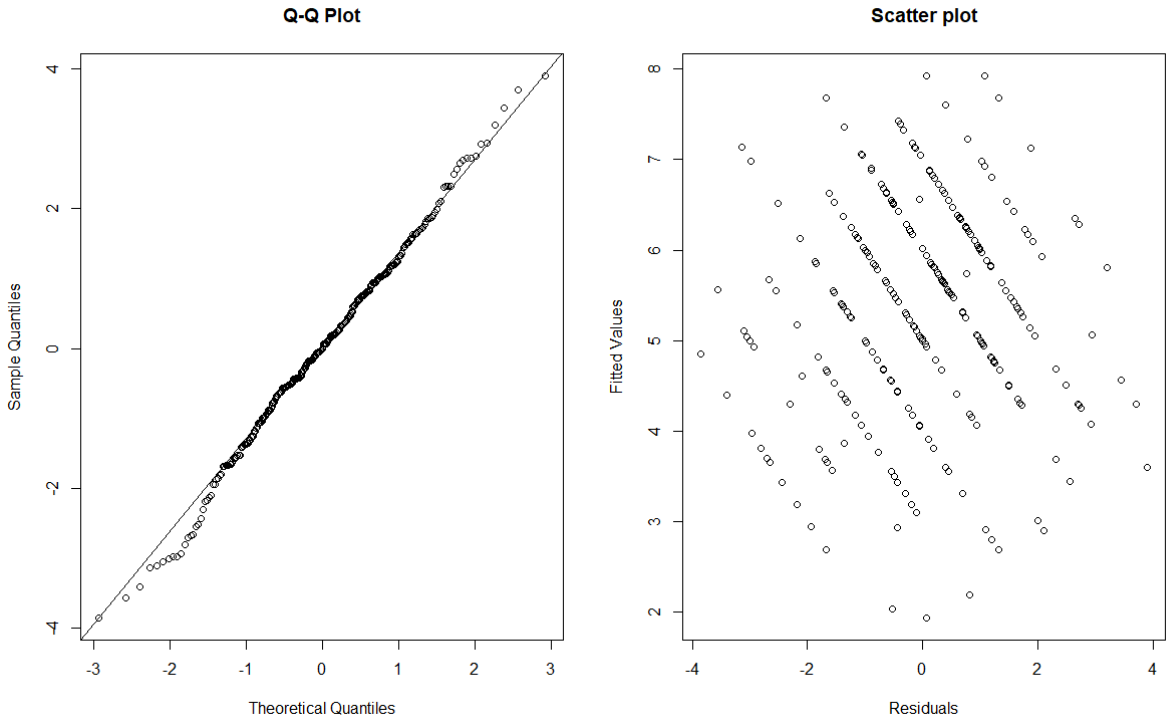


Figure A10 - (Trial R2) Normal Q-Q plot used as direct assessment of the requirement of normal distribution of residuals for validation of the ANOVA. The scatter plot assesses the equal variances across groups, another requirement of the analysis.

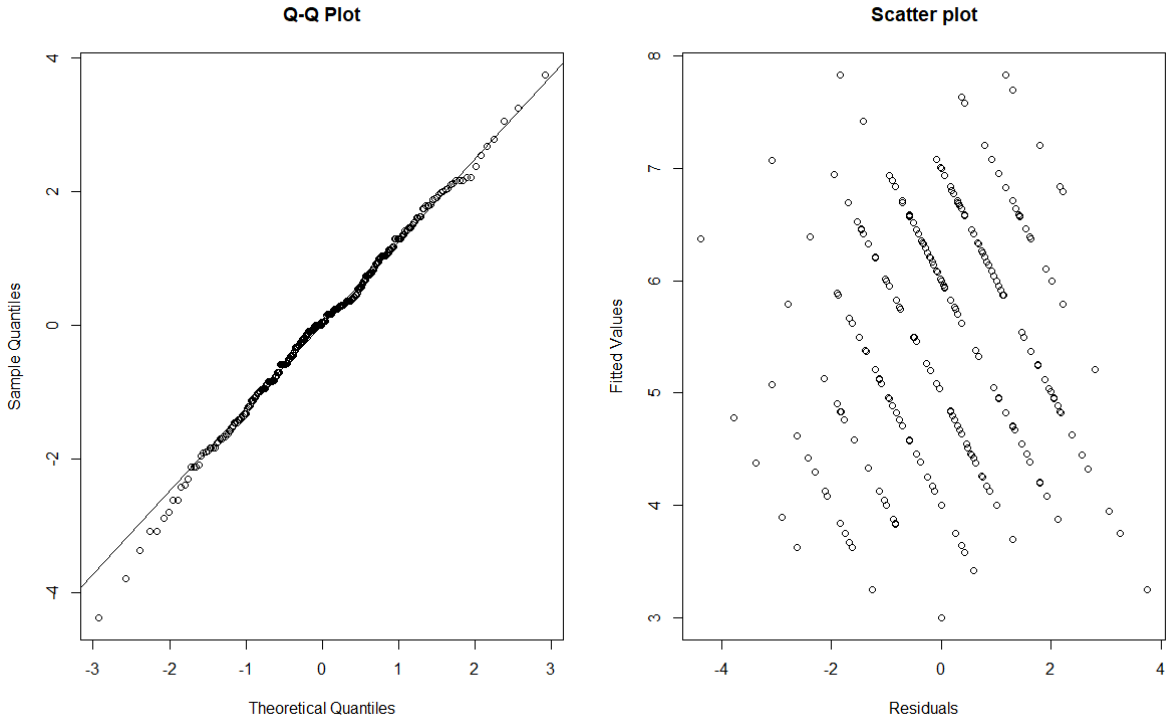


Figure A11 - (Trial R3) Normal Q-Q plot used as direct assessment of the requirement of normal distribution of residuals for validation of the ANOVA. The scatter plot assesses the equal variances across groups, another requirement of the analysis.

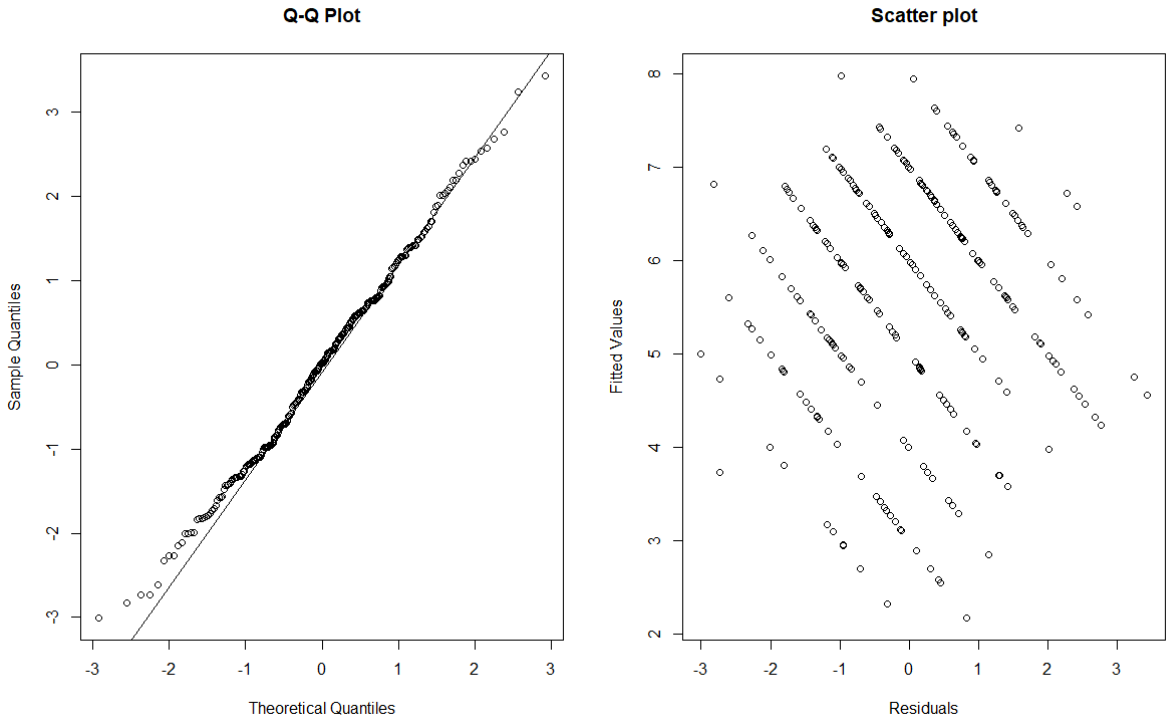


Figure A12 - (Trial R4) Normal Q-Q plot used as direct assessment of the requirement of normal distribution of residuals for validation of the ANOVA. The scatter plot assesses the equal variances across groups, another requirement of the analysis.

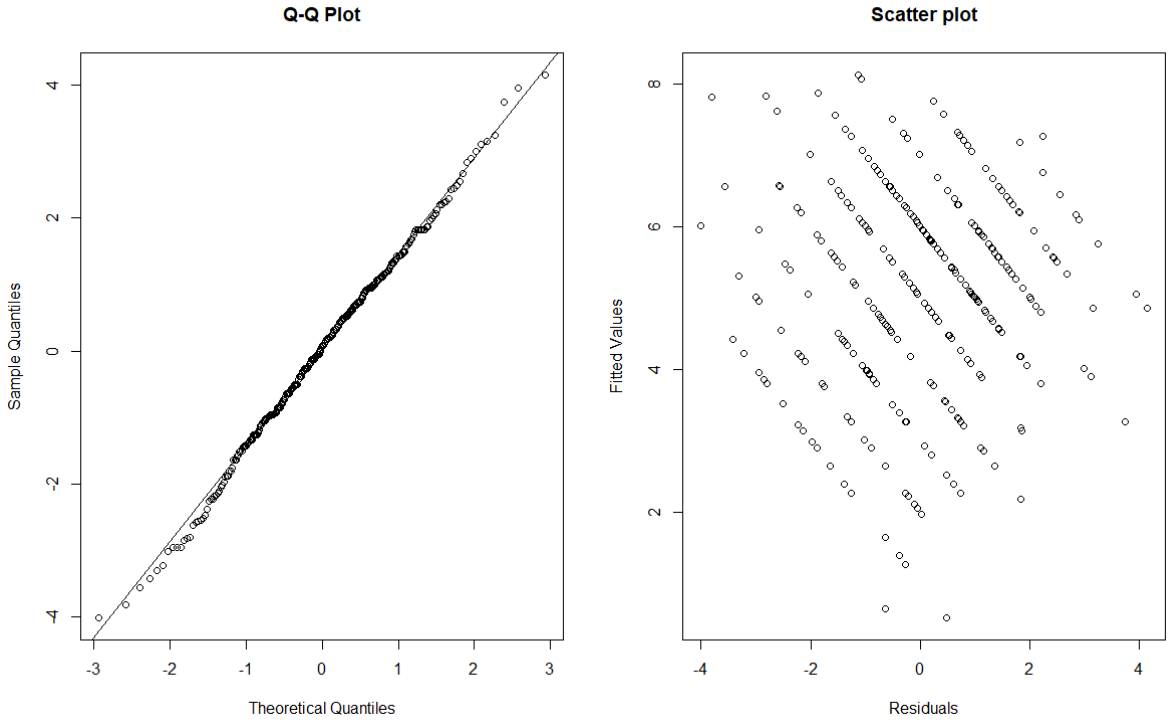


Figure A13 - (Trial R5) Normal Q-Q plot used as direct assessment of the requirement of normal distribution of residuals for validation of the ANOVA. The scatter plot assesses the equal variances across groups, another requirement of the analysis.

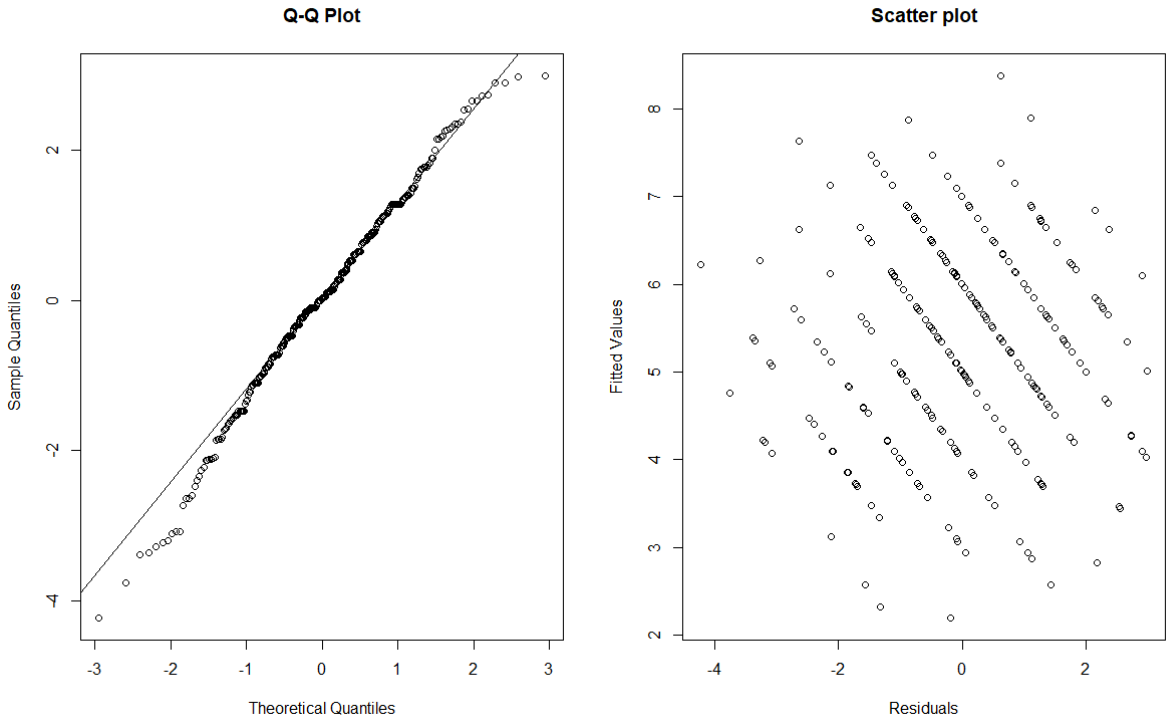


Figure A14 - (Trial R6) Normal Q-Q plot used as direct assessment of the requirement of normal distribution of residuals for validation of the ANOVA. The scatter plot assesses the equal variances across groups, another requirement of the analysis.

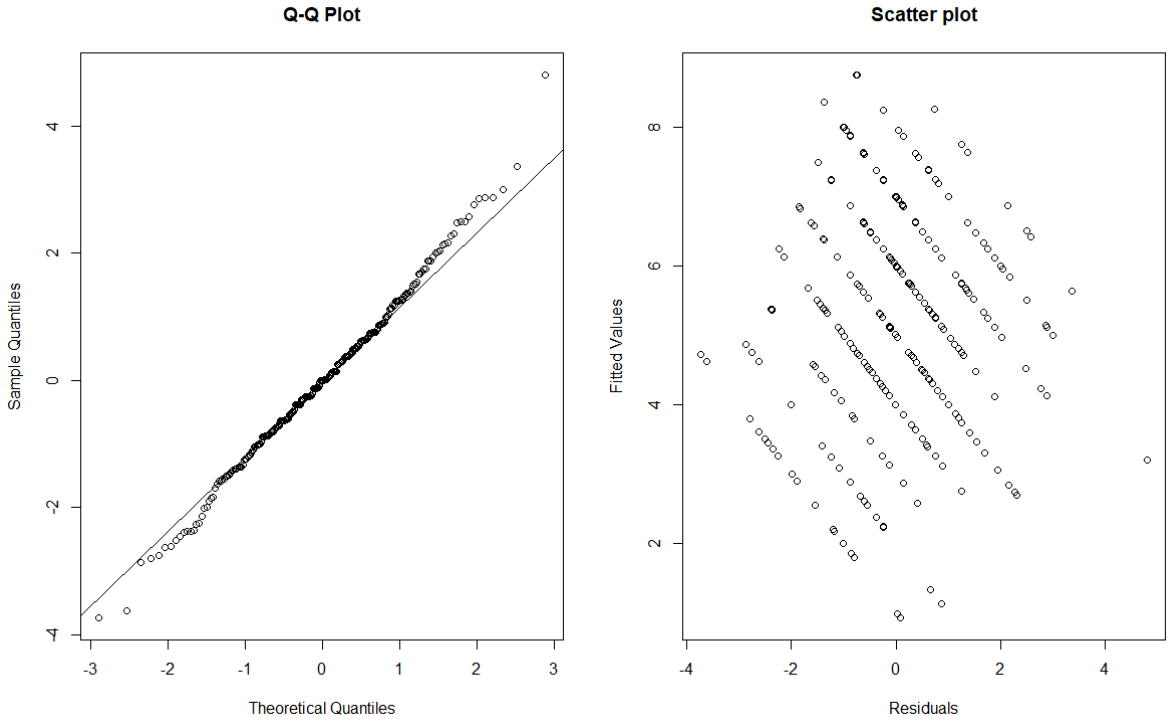


Figure A15 - (Trial R7) Normal Q-Q plot used as direct assessment of the requirement of normal distribution of residuals for validation of the ANOVA. The scatter plot assesses the equal variances across groups, another requirement of the analysis.

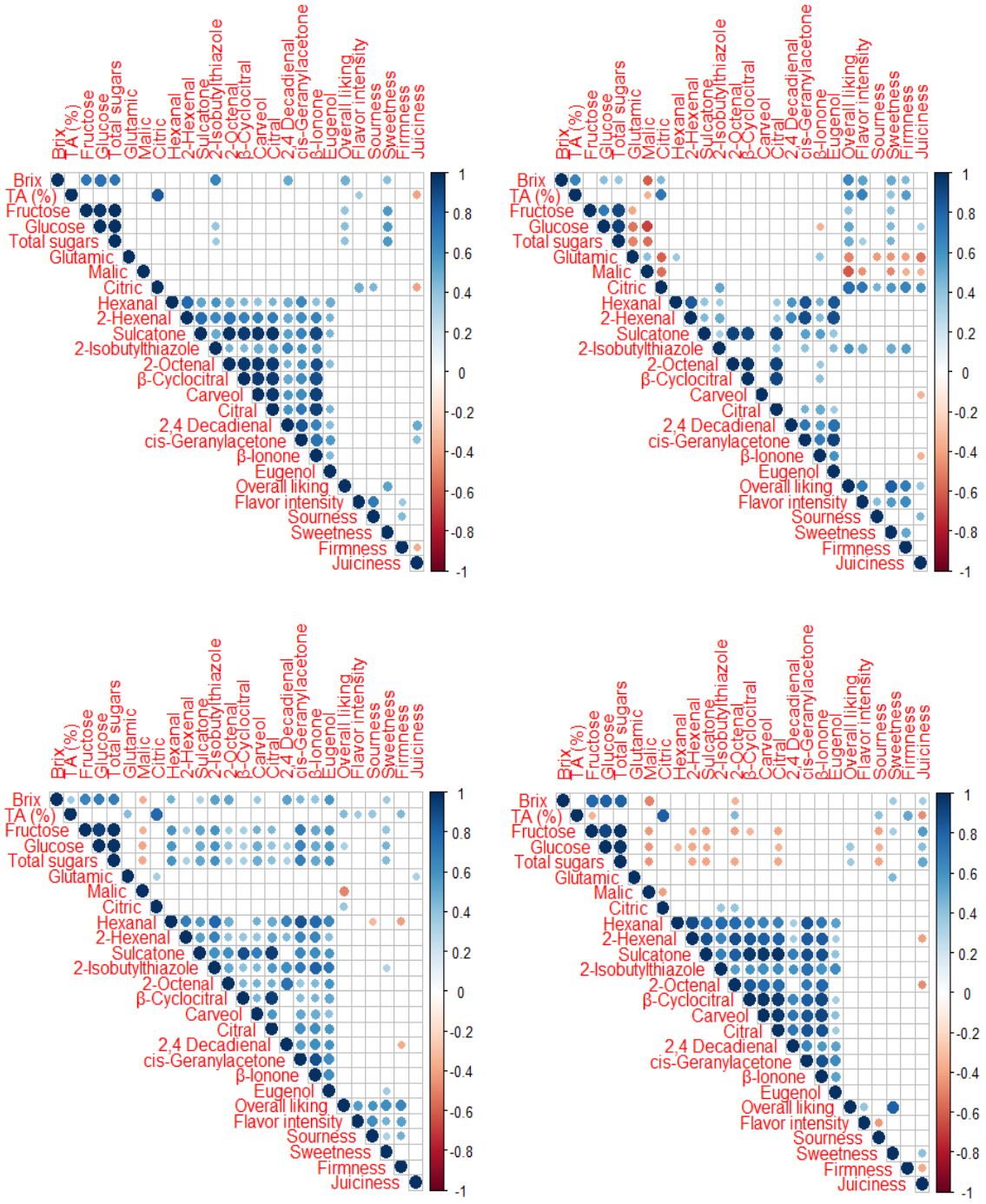


Figure A16 – (From left to right, top to bottom: Trial 1 to 4) Correlation plots with all the variables for alpha=0.05 (statistically insignificant correlation results do not appear)

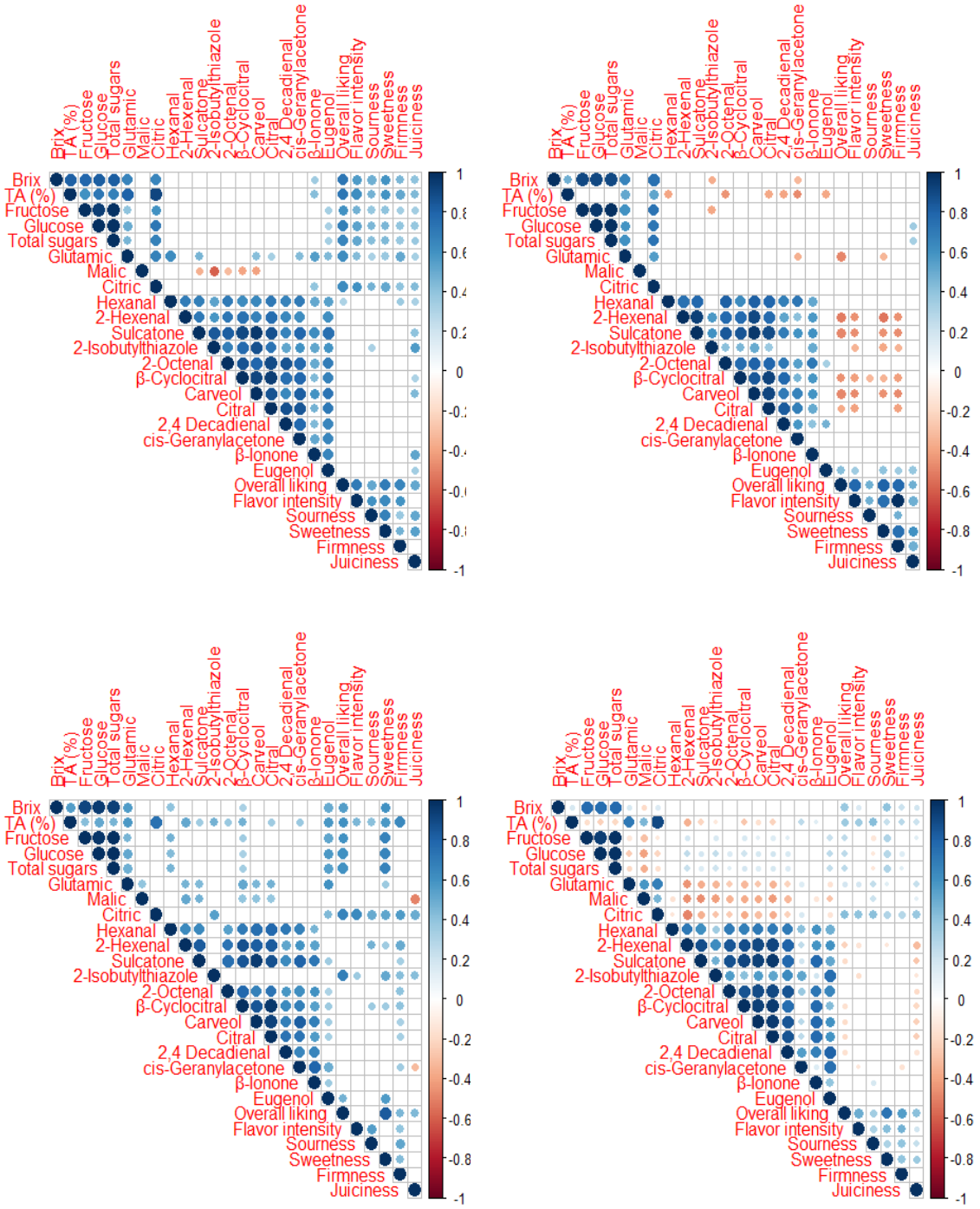


Figure A17 - (From left to right, top to bottom: Trial 5 to 7 and All Trials) Correlation plots with all the variables for alpha=0.05 (statistically insignificant correlation results do not appear)

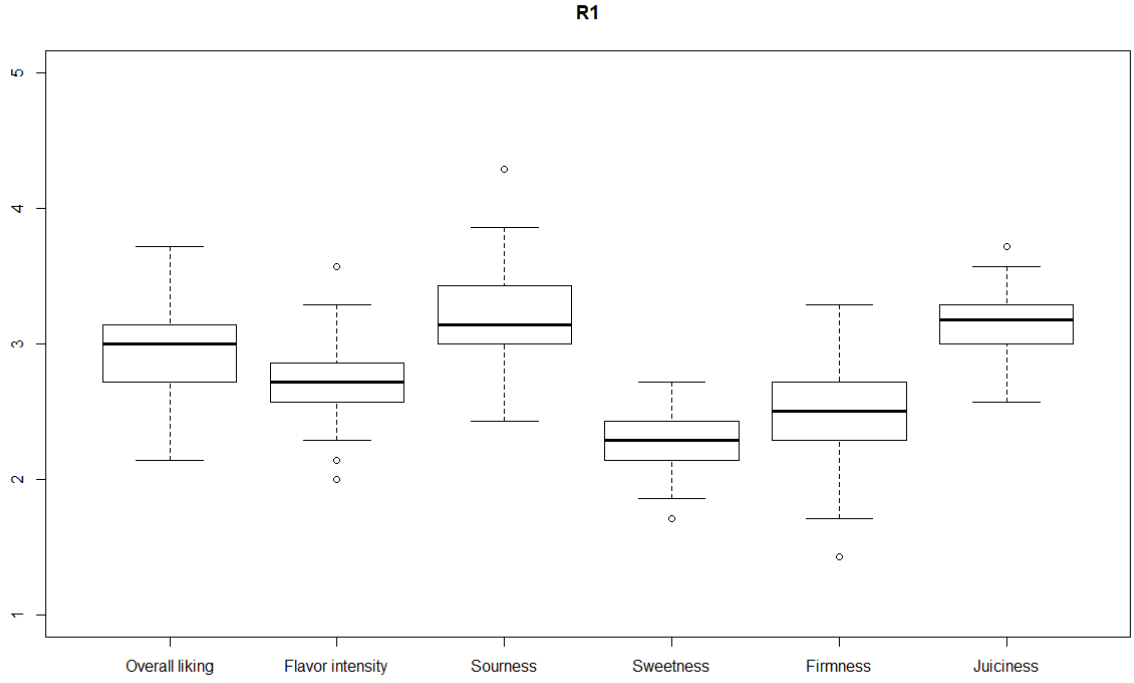


Figure A18 - (Trial R1) Boxplots for all the response variables. Overall liking scale divided by two.

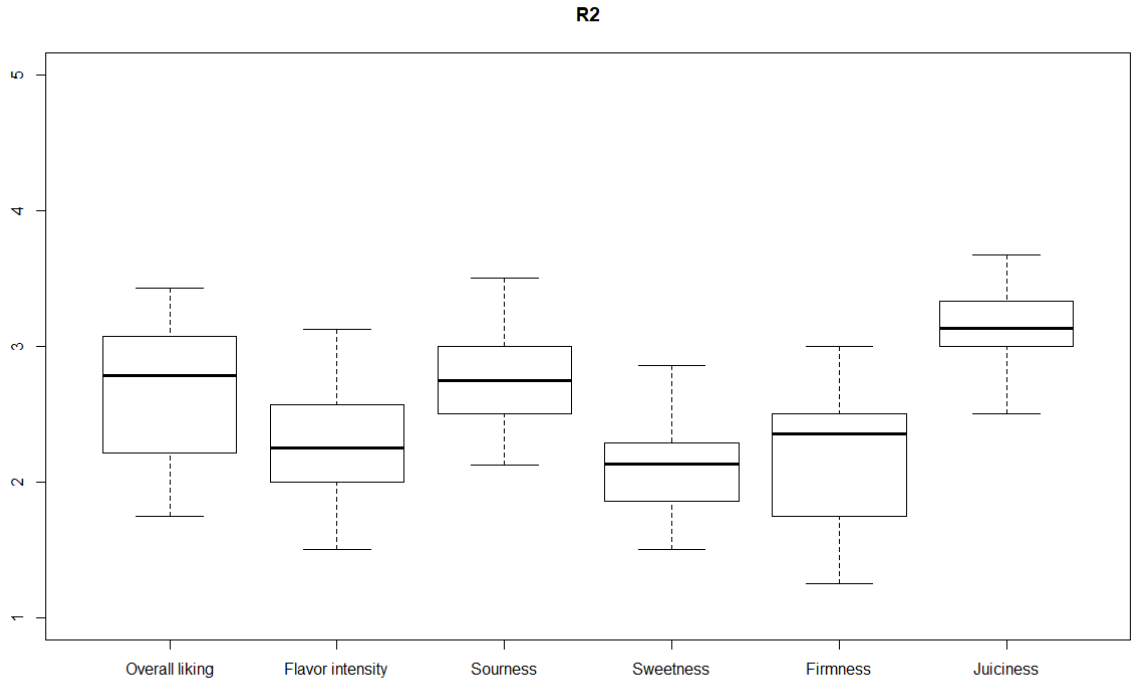


Figure A19 - (Trial R2) Boxplots for all the response variables. Overall liking scale divided by two.

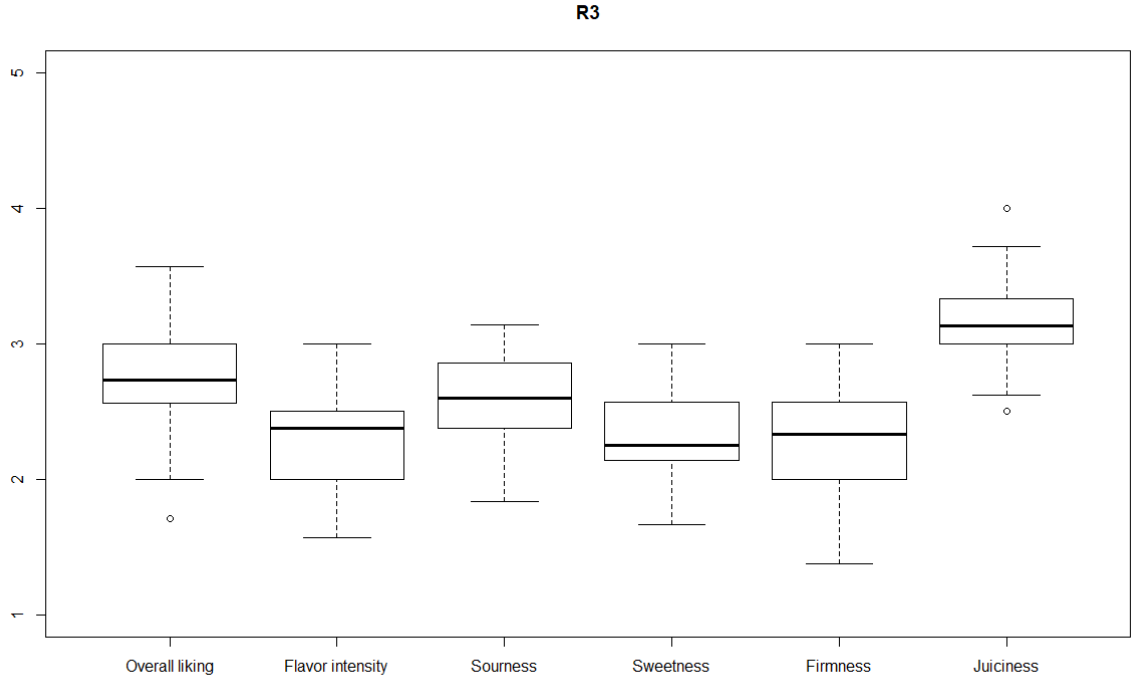


Figure A20 - (Trial R3) Boxplots for all the response variables. Overall liking scale divided by two.

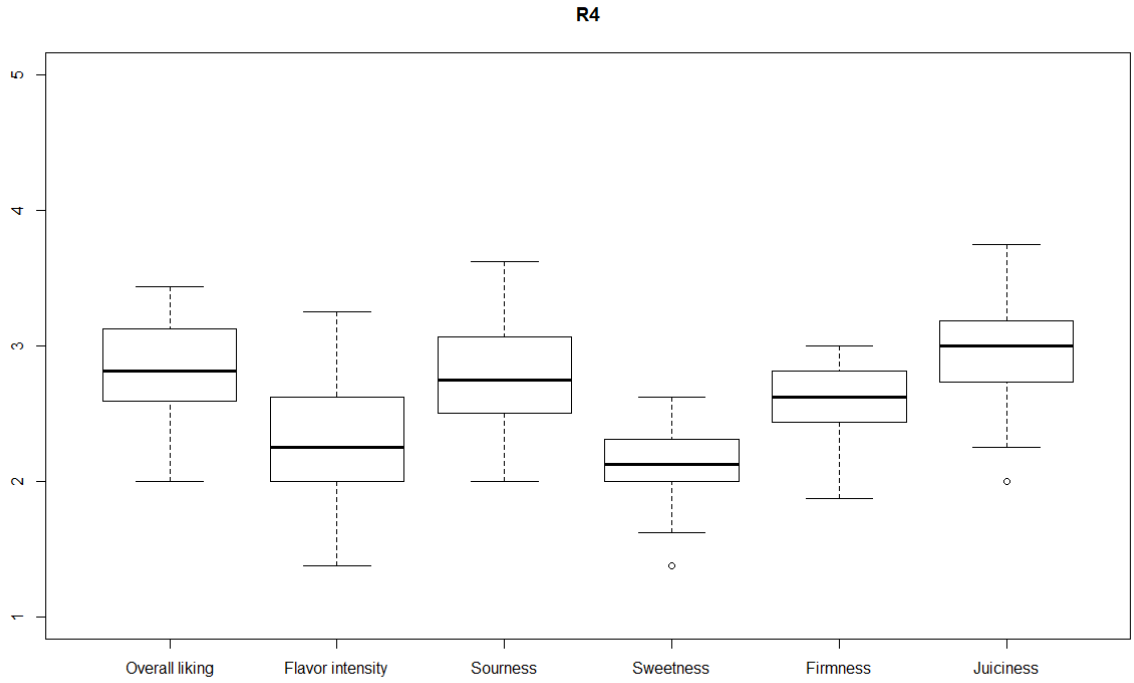


Figure A21 - (Trial R4) Boxplots for all the response variables. Overall liking scale divided by two.

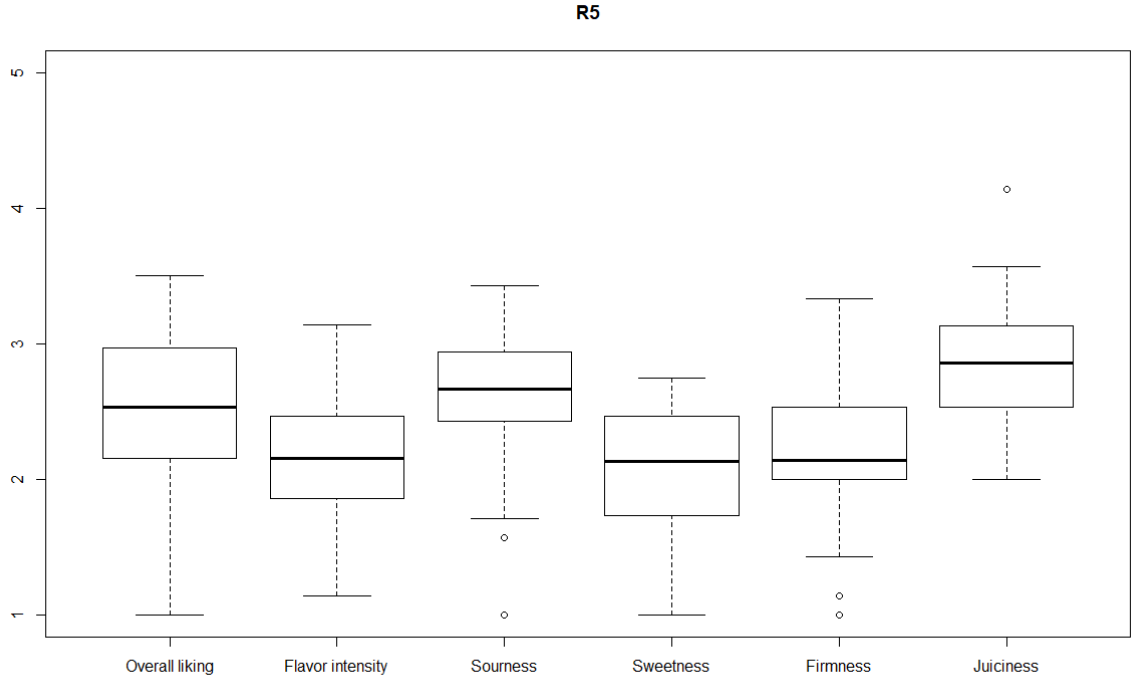


Figure A22 - (Trial R5) Boxplots for all the response variables. Overall liking scale divided by two.

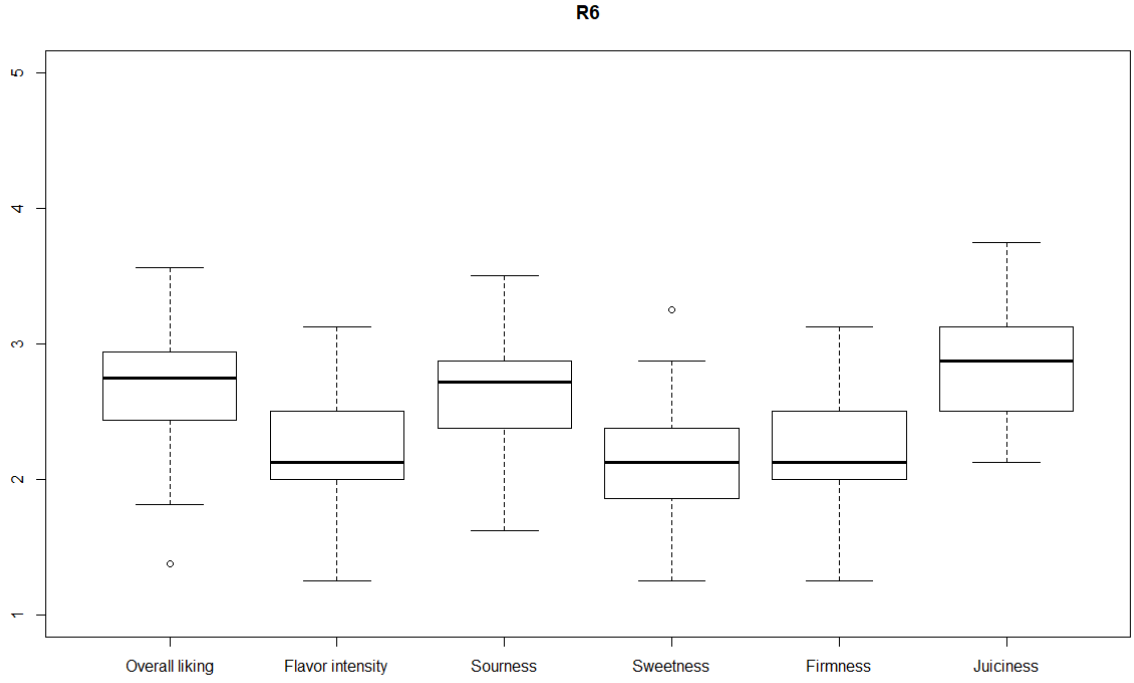


Figure A23 - (Trial R6) Boxplots for all the response variables. Overall liking scale divided by two.

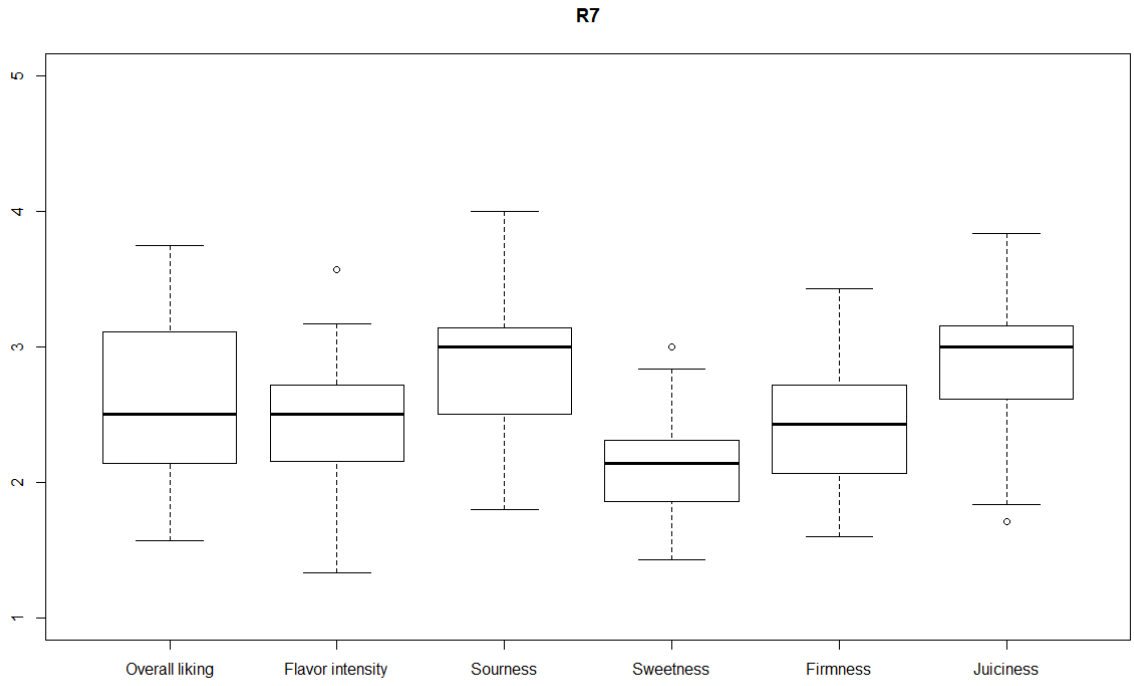


Figure A24 - (Trial R7) Boxplots for all the response variables. Overall liking scale divided by two.

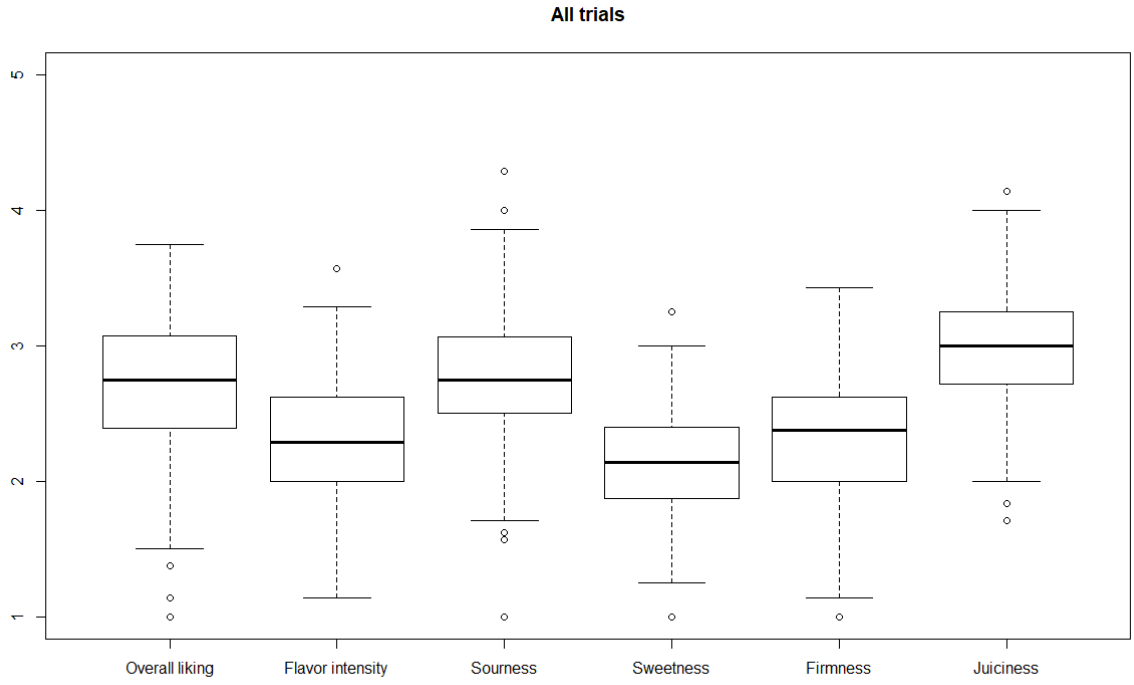
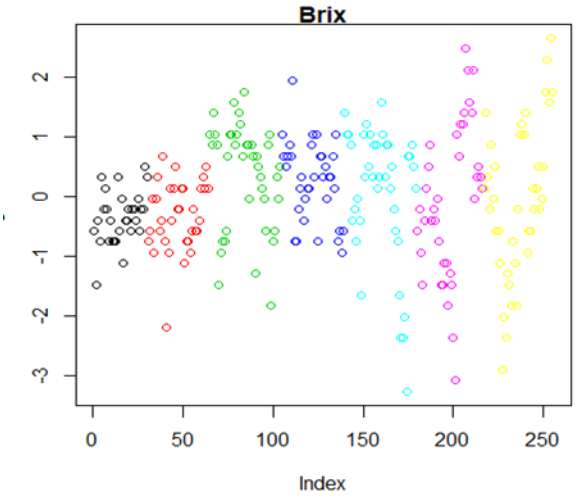
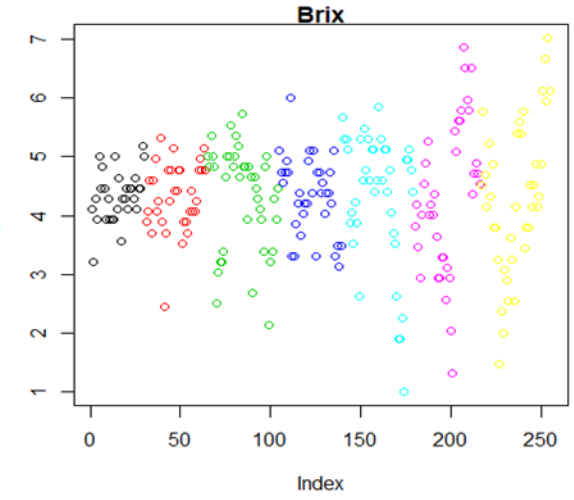
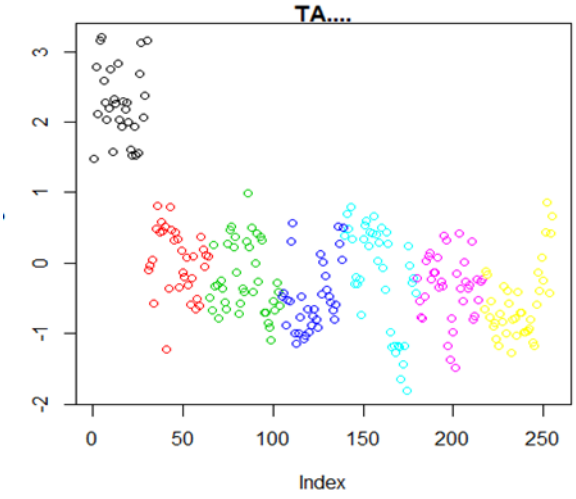
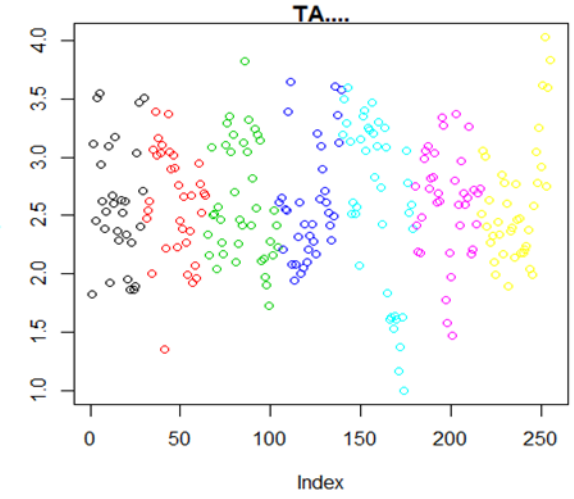


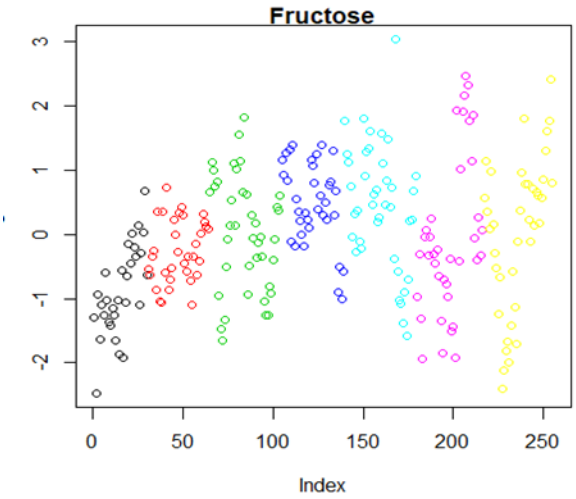
Figure A25 - (All Trials) Boxplots for all the response variables. Overall liking scale divided by two.

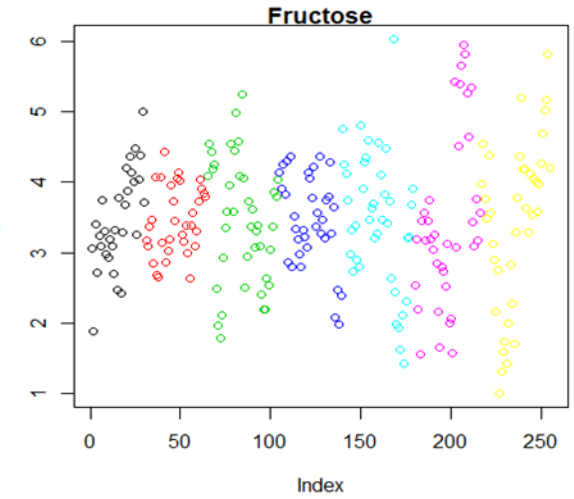

Brix
Index

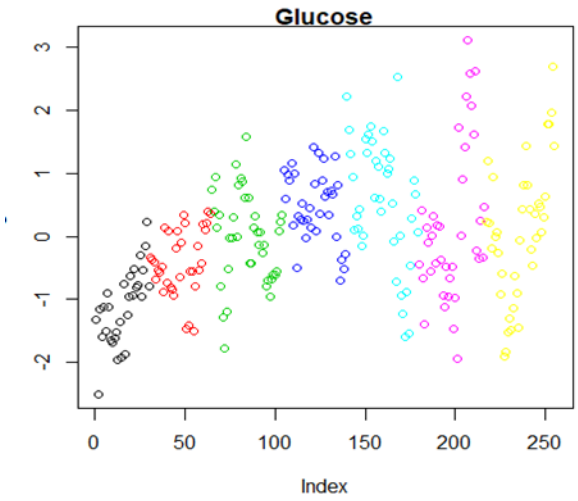

Brix
Index

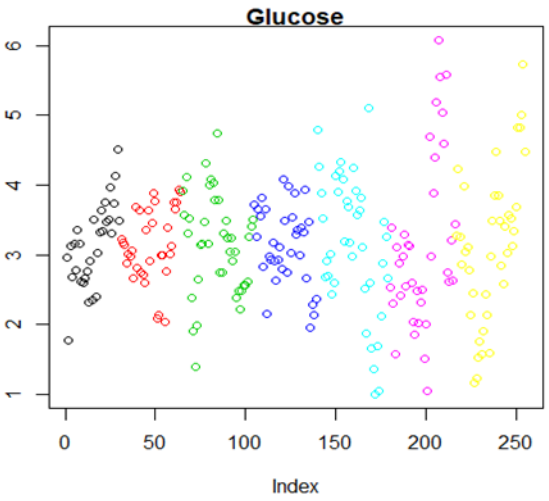

TA....
Index

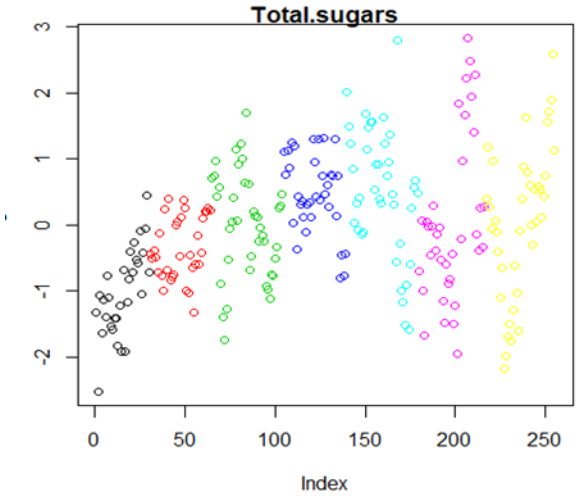

TA....
Index

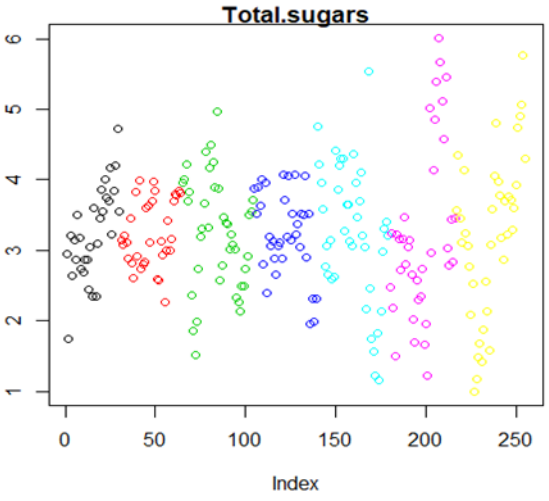

Fructose
Index

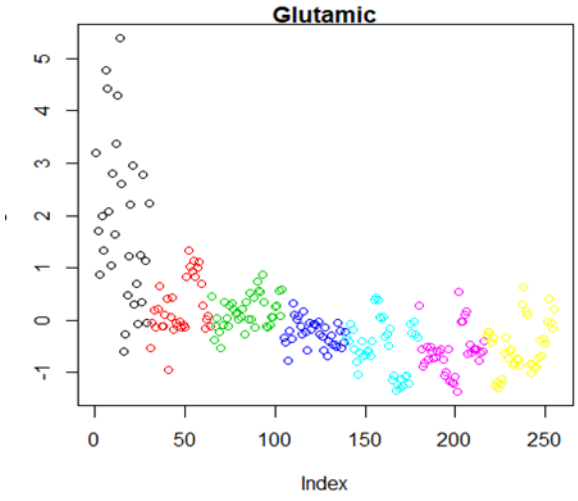

Fructose
Index

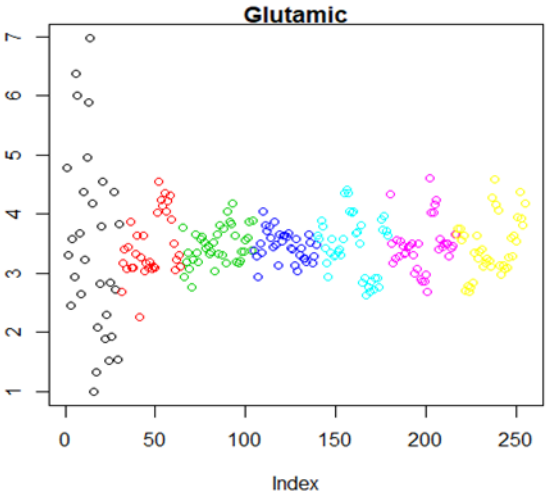

Glucose
Index


Glucose
Index


Total.sugars
Index


Total.sugars
Index


Glutamic
Index


Glutamic
Index

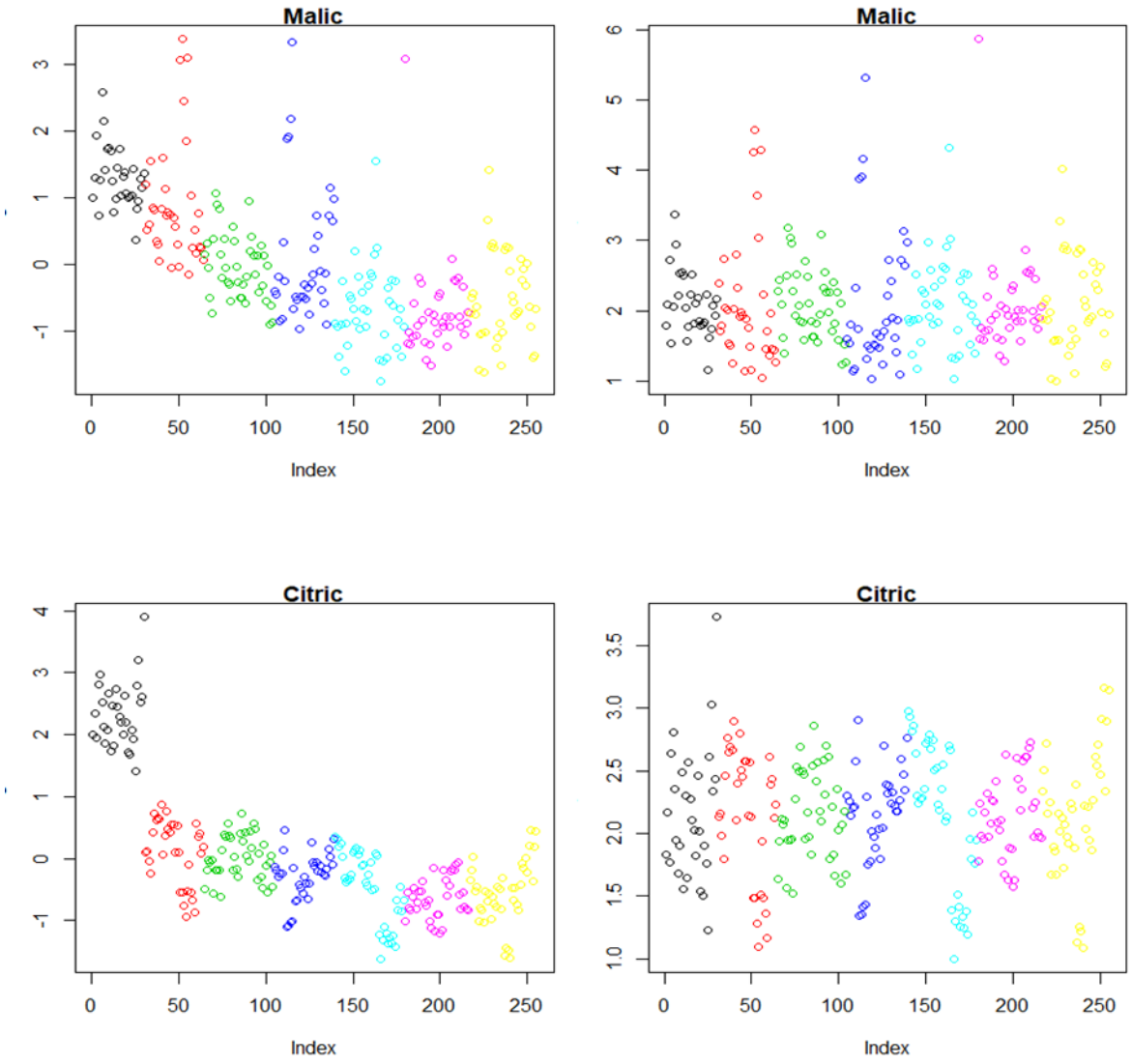
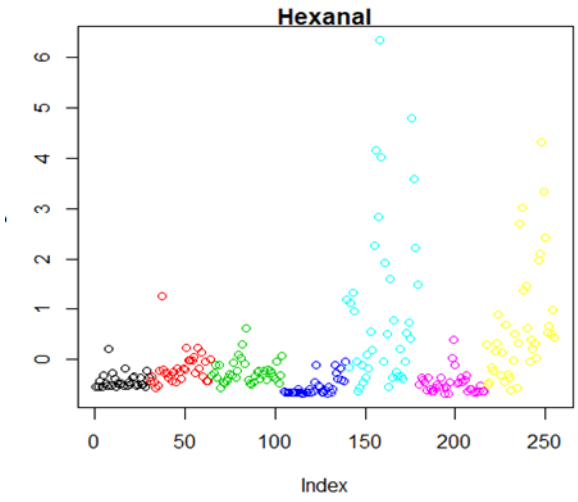
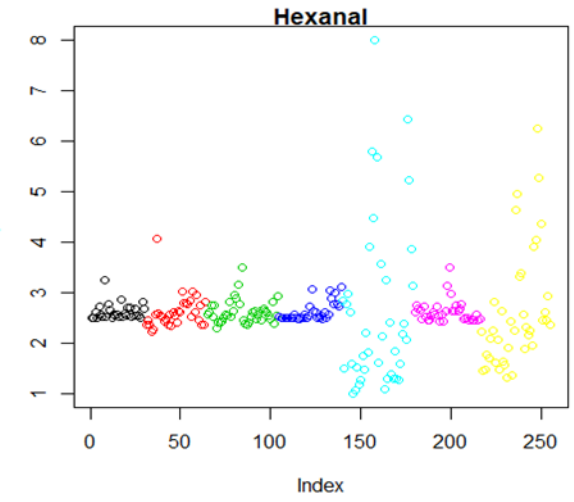
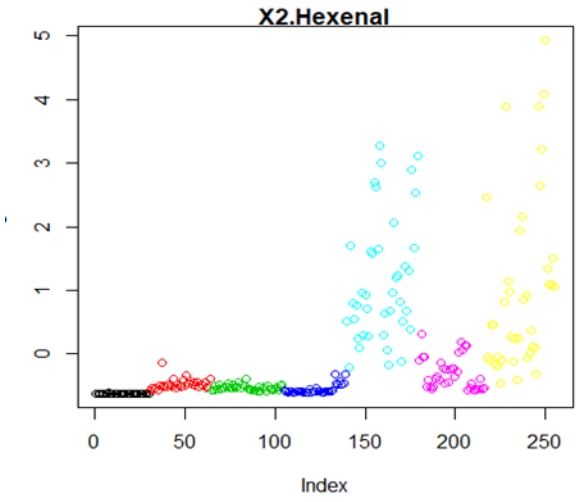
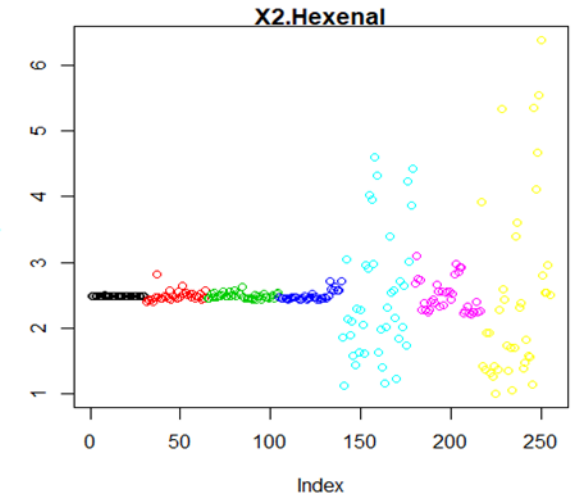


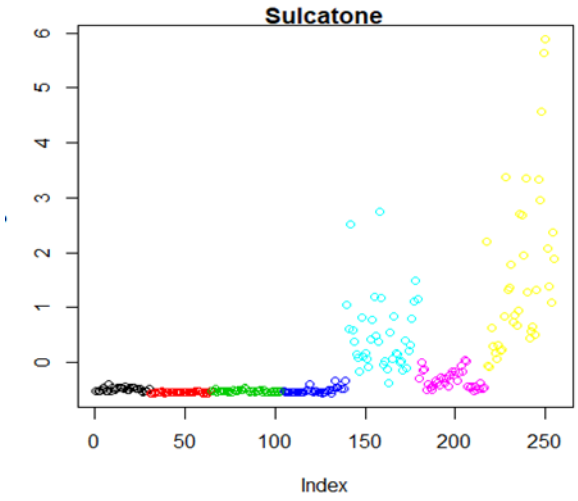
Figure A26 - Batch effect for the Non-volatile metabolites' variables. On the left before the adjustment for the mean, on the right after the adjustment. On each image, from left to right Trial 1 to 7 (represented by different colours)

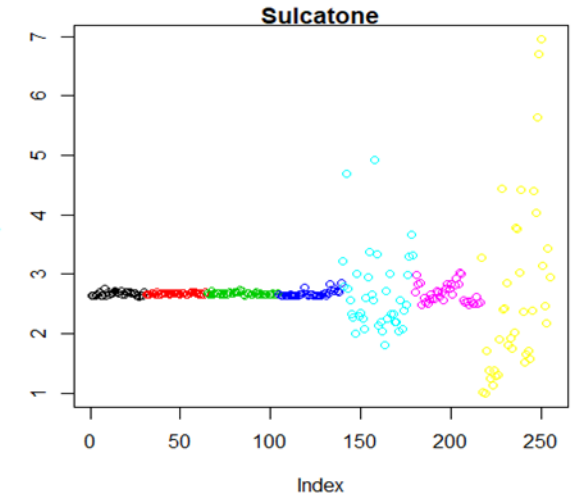

Hexanal
Index

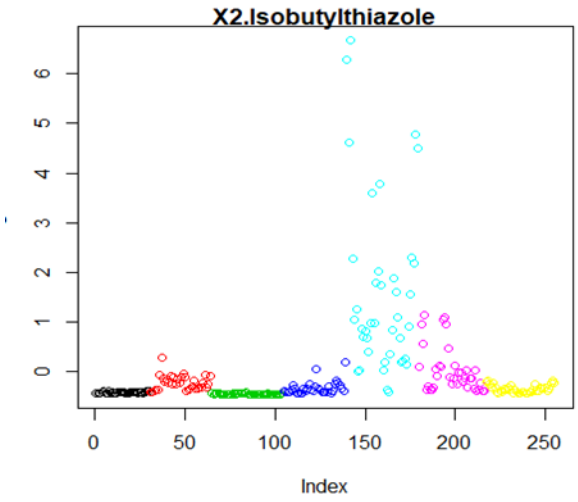

Hexanal
Index

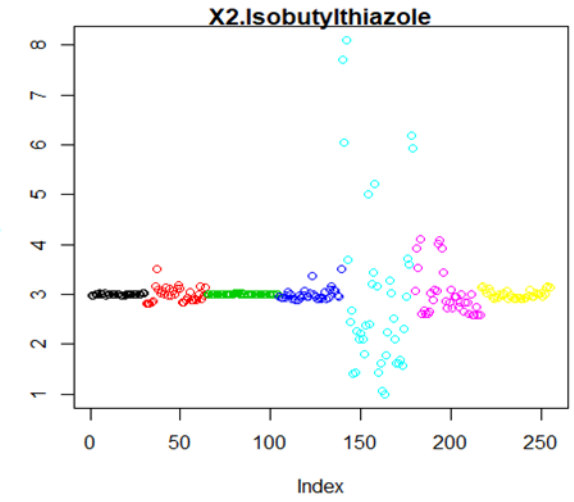

X2.Hexenal
Index

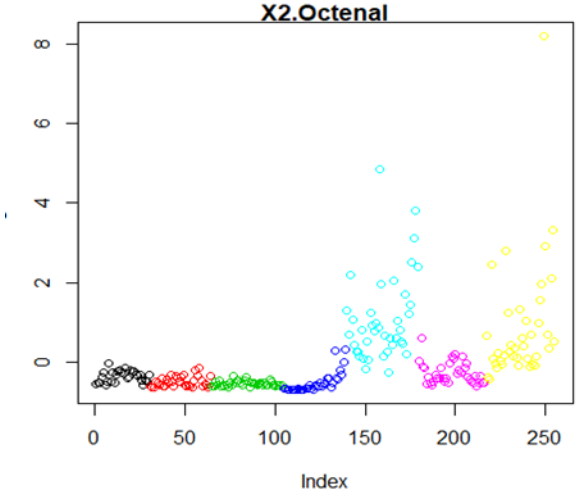

X2.Hexenal
Index

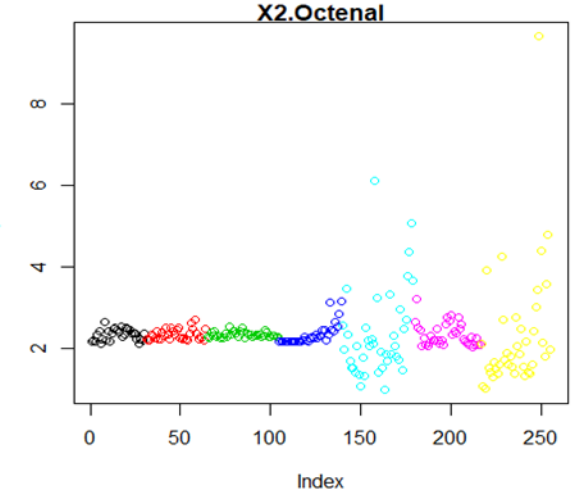

Sulcatone
Index

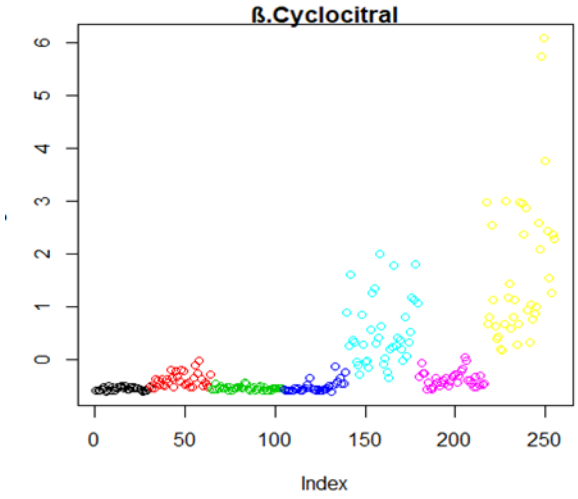

Sulcatone
Index

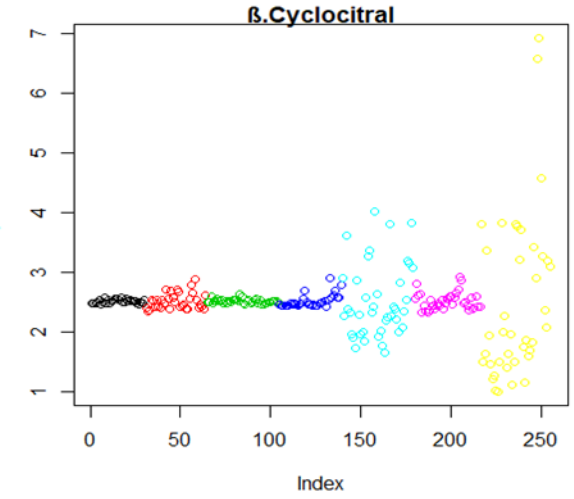

X2.Isobutylthiazole
Index

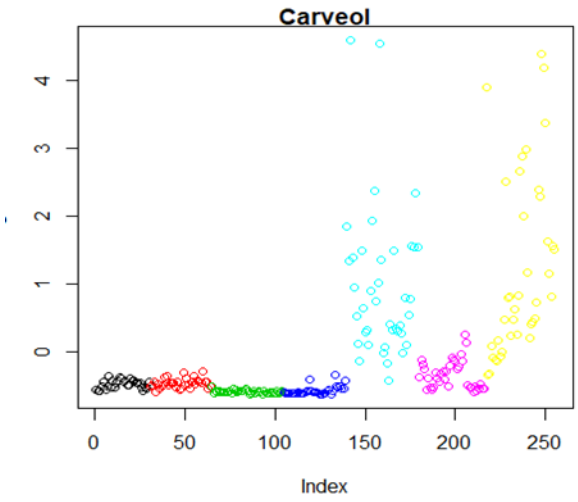

X2.Isobutylthiazole
Index

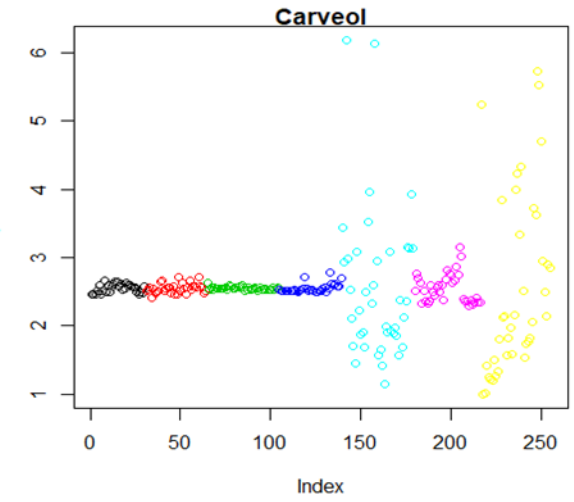

X2.Octenal
Index

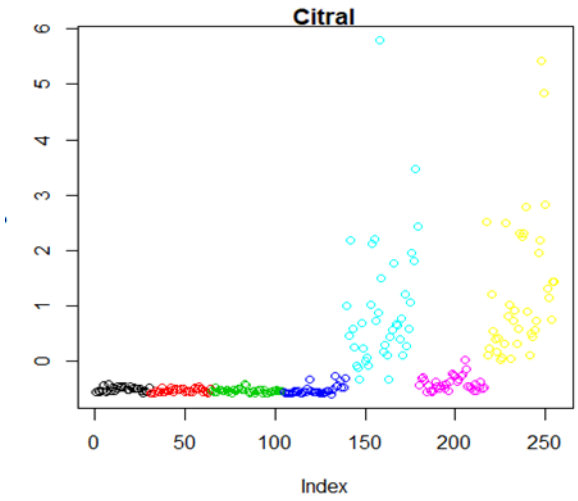

X2.Octenal
Index

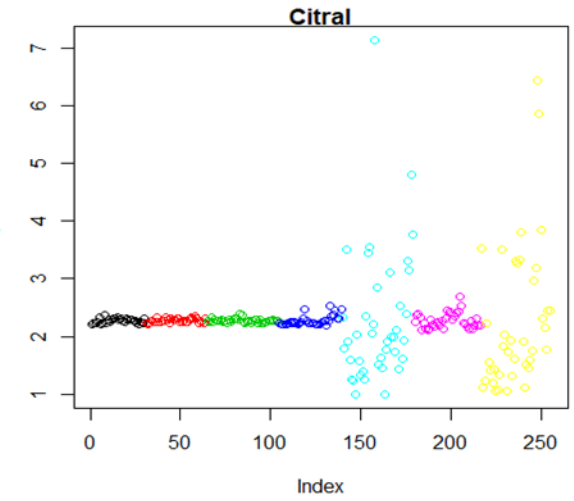

ß.Cyclocitral
Index

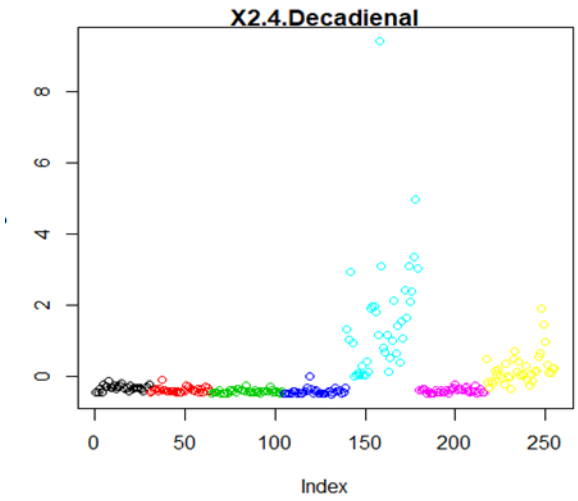

ß.Cyclocitral
Index

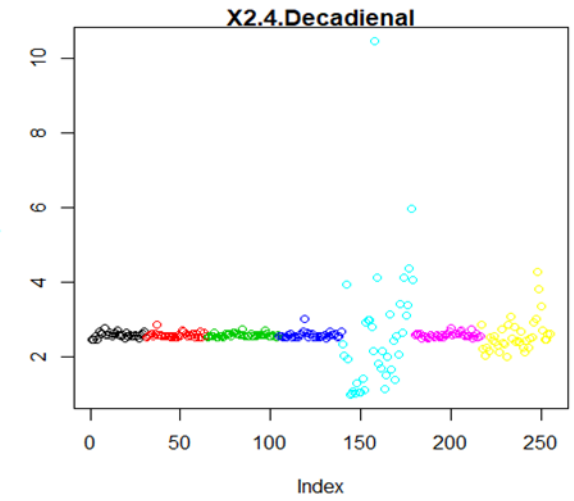

Carveol
Index


Carveol
Index


Citral
Index


Citral
Index


X2.4.Decadienal
Index


X2.4.Decadienal
Index

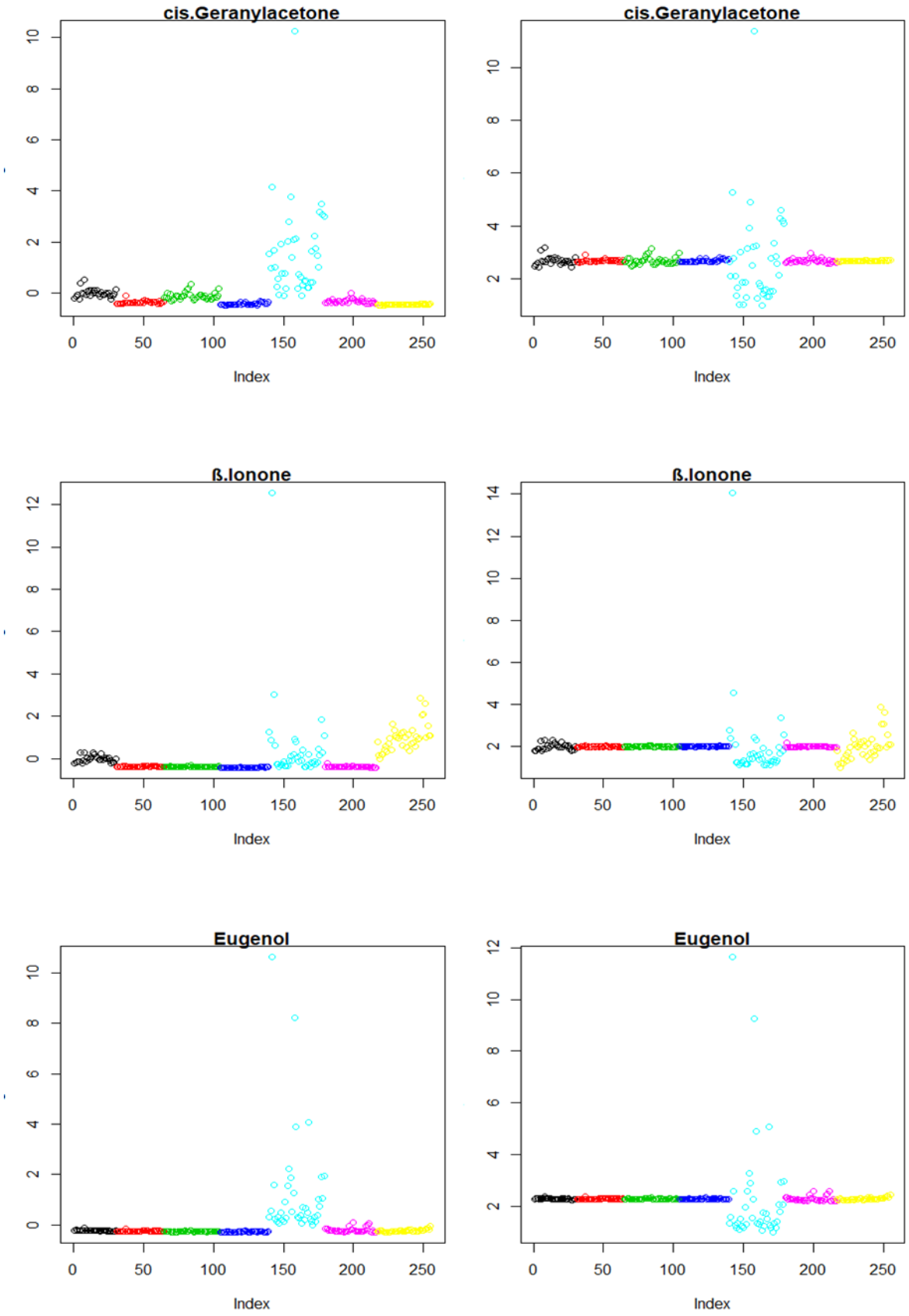
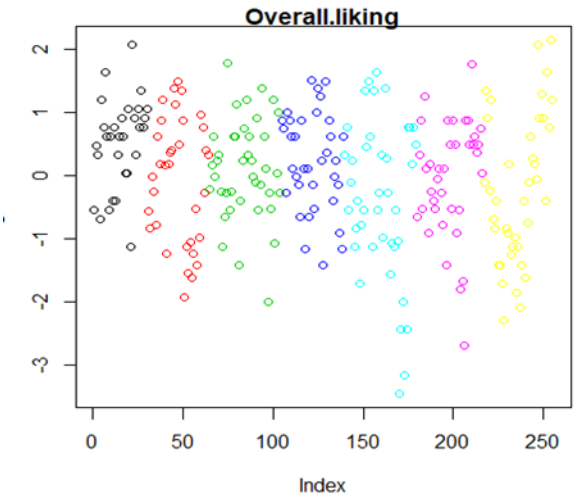
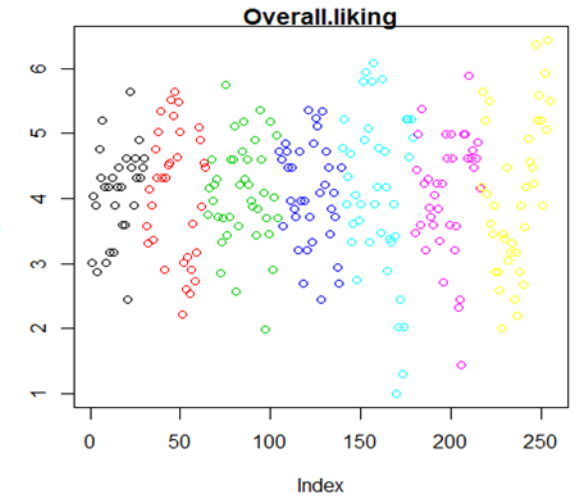
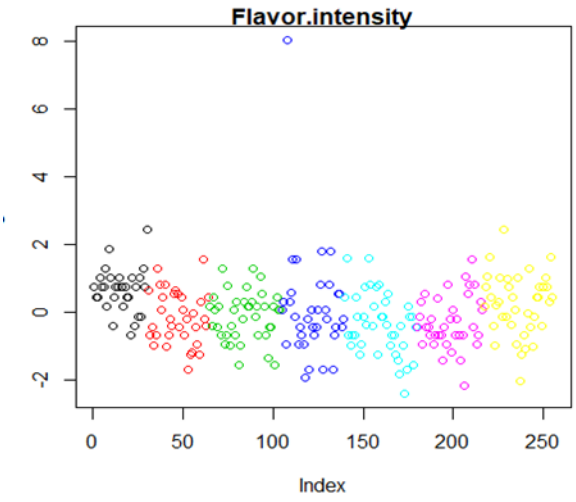
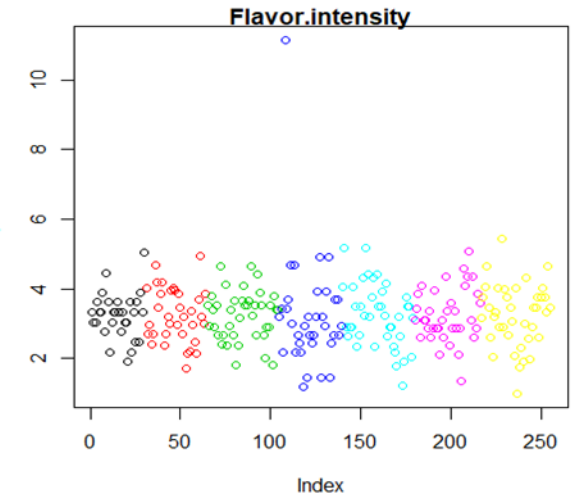


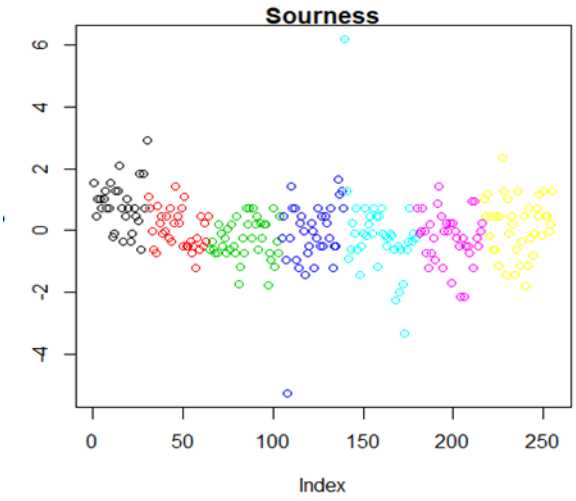
Figure A27 - Batch effect for the volatile variables. On the left before the adjustment for the mean, on the right after the adjustment. On each image, from left to right Trial 1 to 7 (represented by different colours)

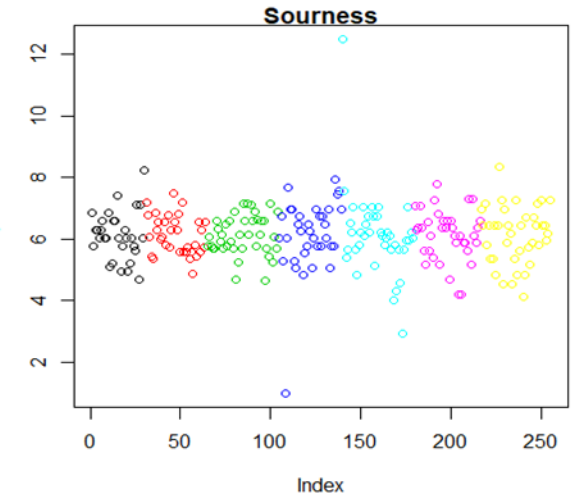

Overall.liking
Index


Overall.liking
Index


Flavor.intensity
Index


Flavor.intensity
Index


Sourness
Index


Sourness
Index

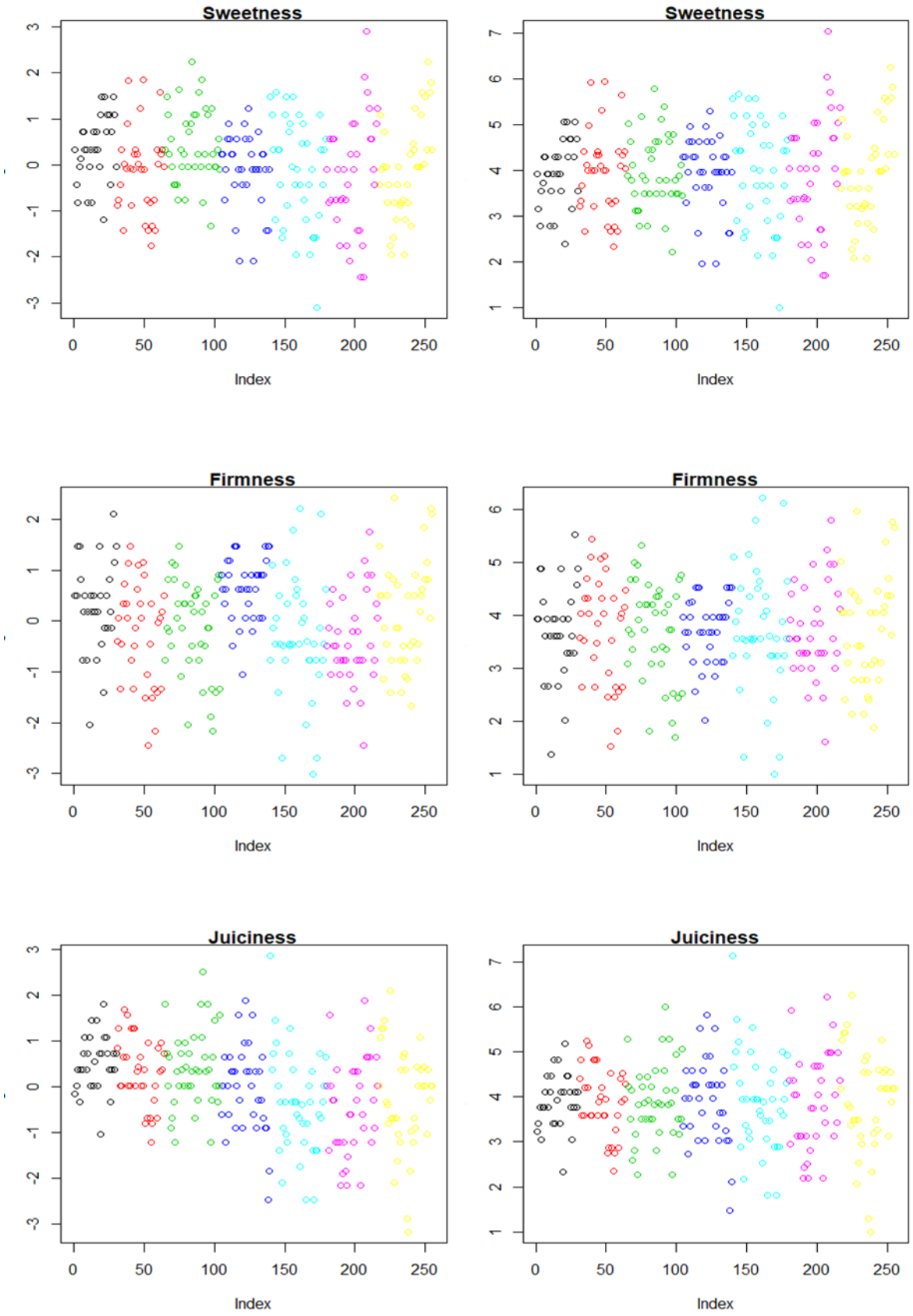


Figure A28 - Batch effect for the response variables. On the left before the adjustment for the mean, on the right after the adjustment. On each image, from left to right Trial 1 to 7 (represented by different colours)

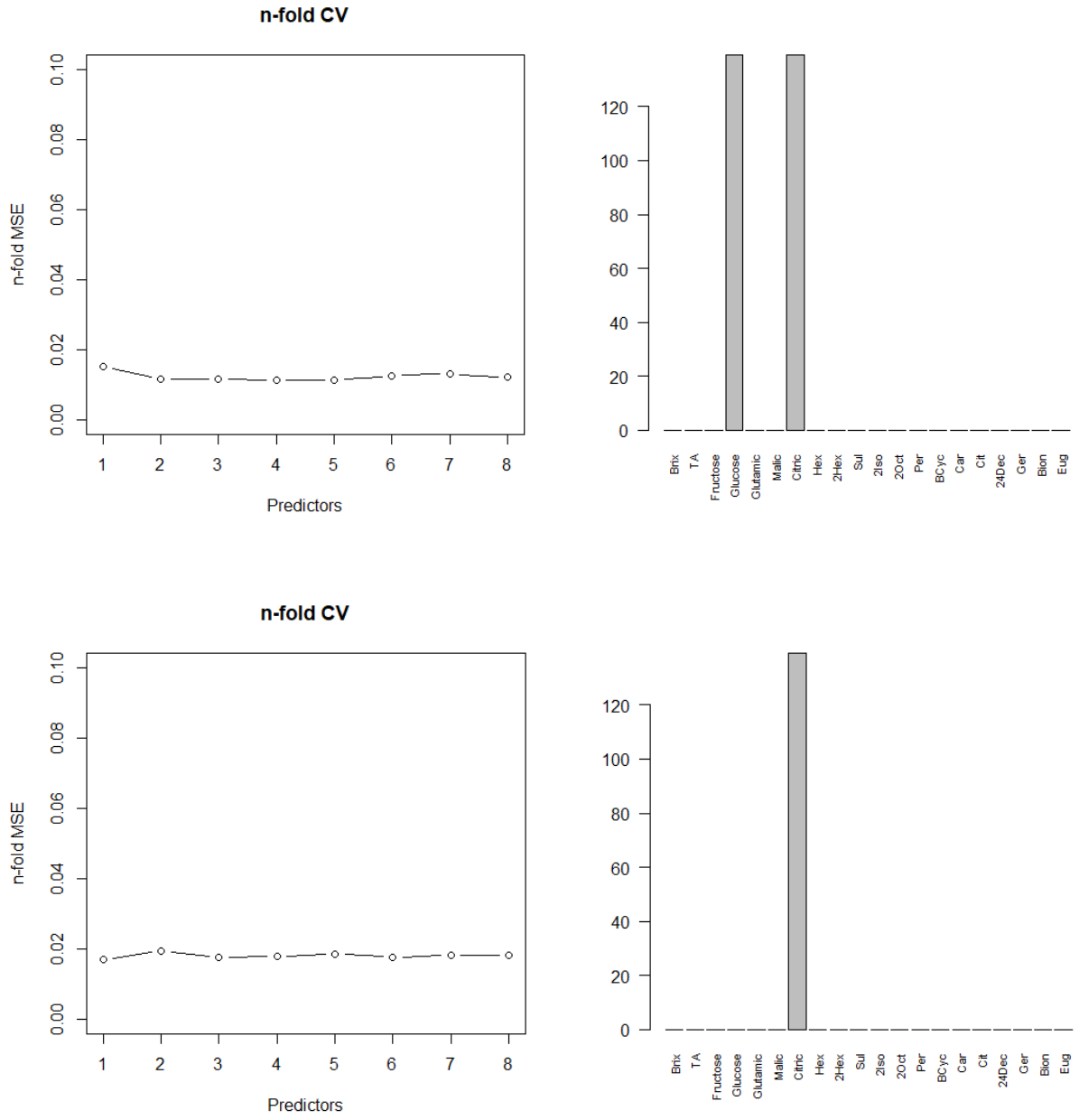


Figure A29 - Results from the Forward Stepwise Regression with LOOCV. On the top response variable Overall Liking, on the bottom Flavour Intensity. Heated trials data (1 to 4). On the left, the average M.S.E. for each level of model complexity for all n number of models. On the right, the number of times a predictor was selected for the desired level of complexity, decided on basis of the left image, for all n models.

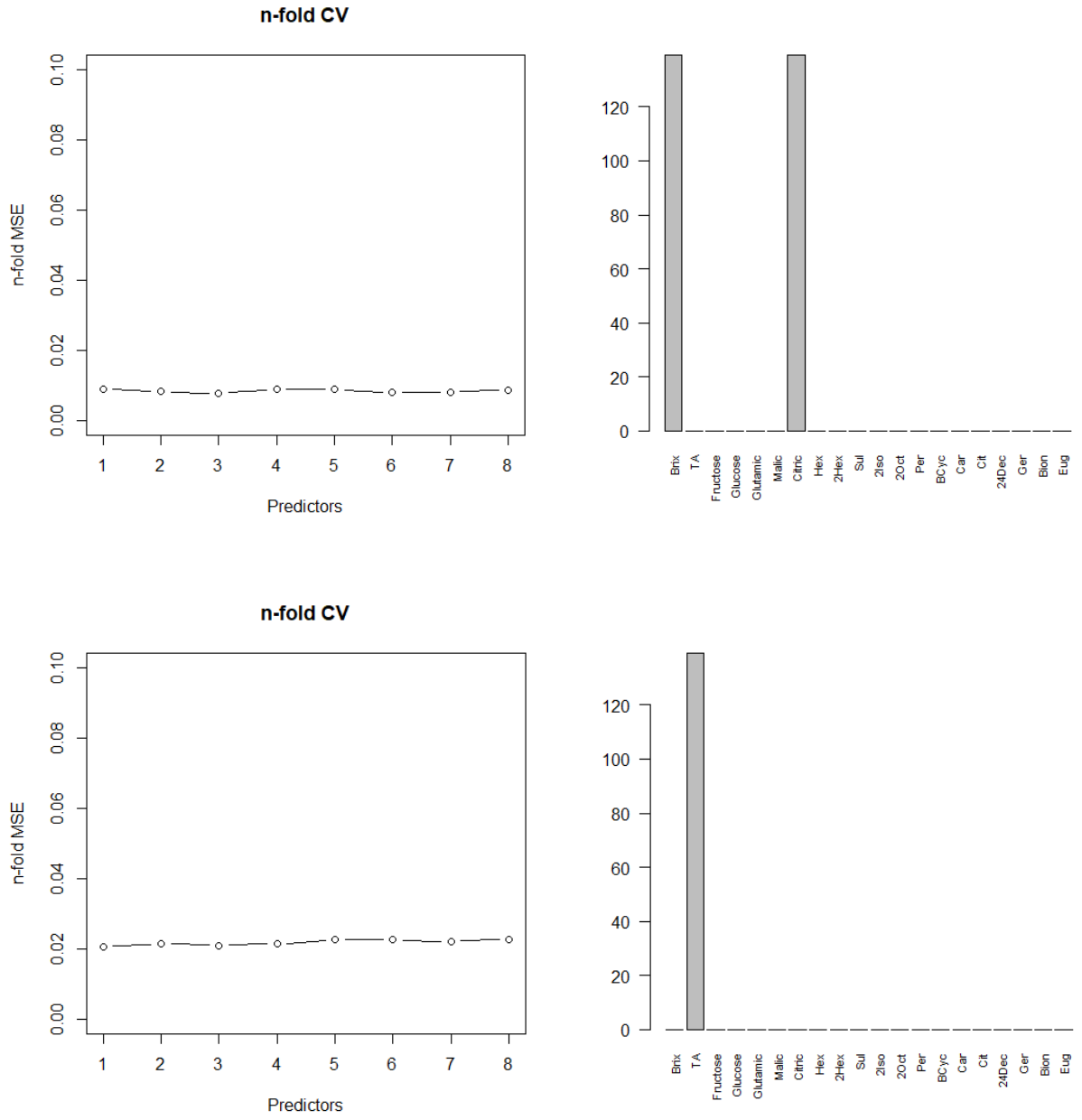


Figure A30 - Results from the Forward Stepwise Regression with LOOCV. On the top response variable Sweetness, on the bottom Sourness. Heated trials data (1 to 4). On the left, the average M.S.E. for each level of model complexity for all n number of models. On the right, the number of times a predictor was selected for the desired level of complexity, decided on basis of the left image, for all n models.

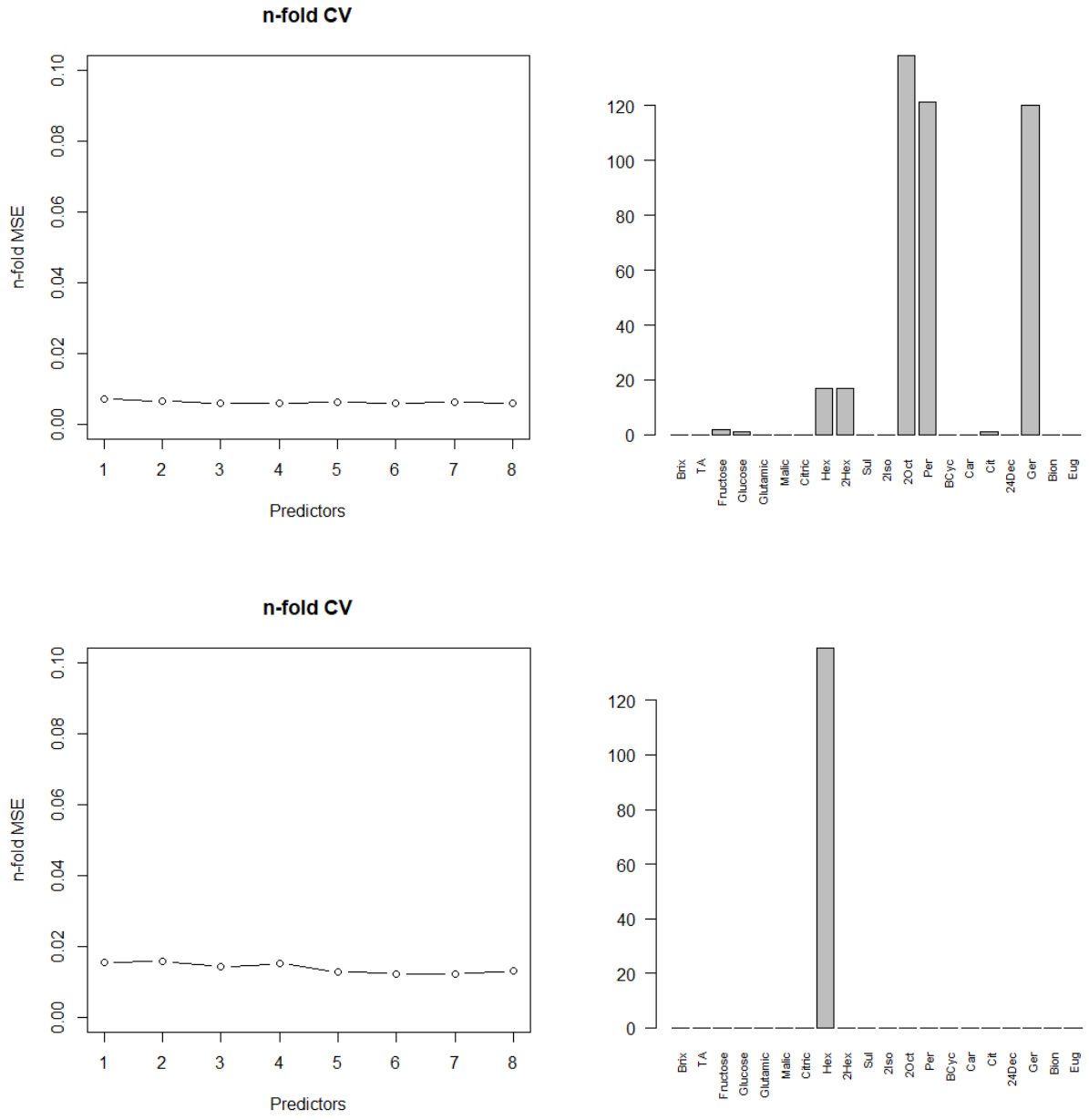


Figure A31 - Results from the Forward Stepwise Regression with LOOCV. On the top response variable Juiciness, on the bottom Firmness. Heated trials data (1 to 4). On the left, the average M.S.E. for each level of model complexity for all n number of models. On the right, the number of times a predictor was selected for the desired level of complexity, decided on basis of the left image, for all n models.

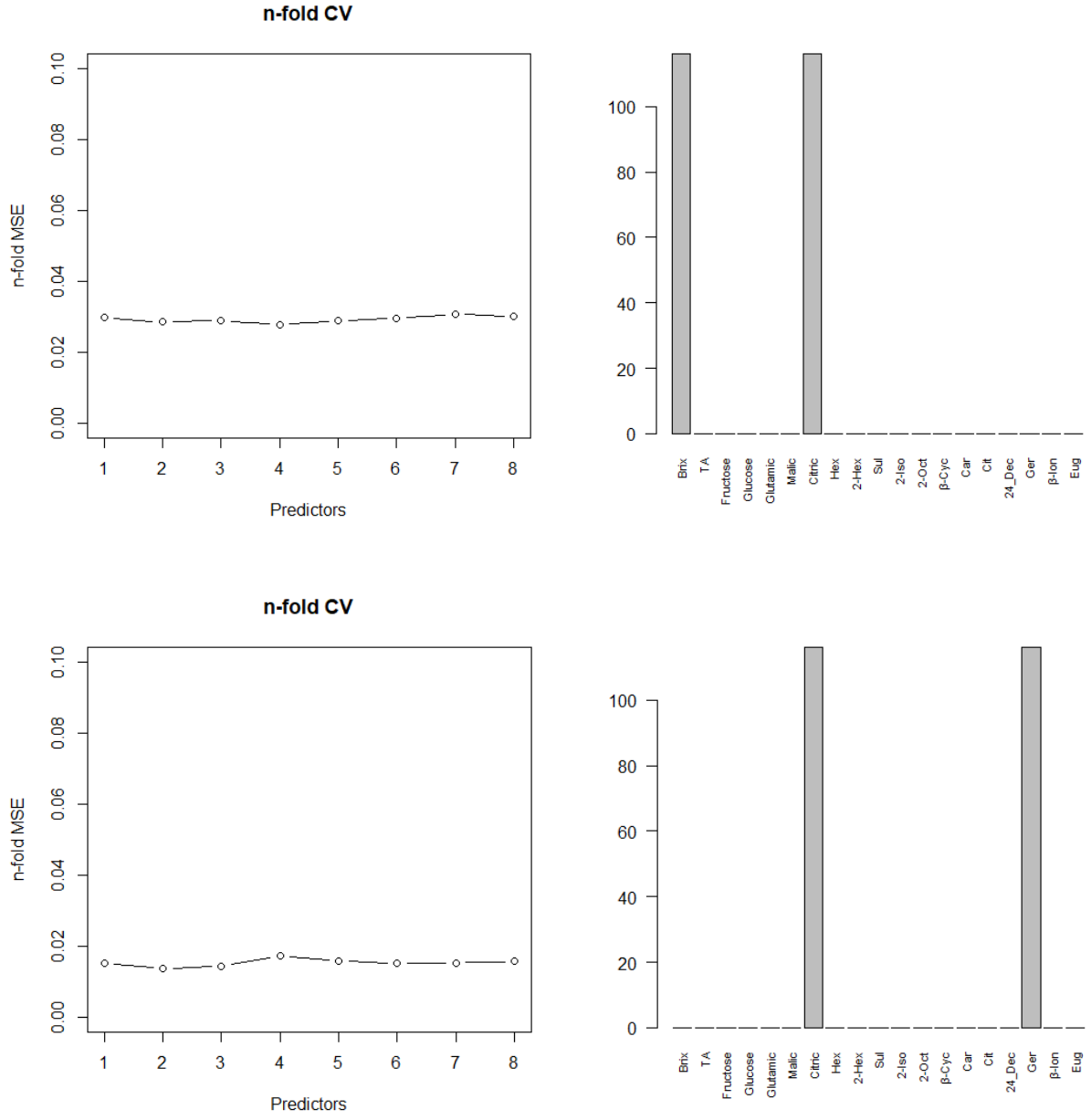


Figure A32 - Results from the Forward Stepwise Regression with LOOCV. On the top response variable Overall Liking, on the bottom Flavour Intensity. Non-heated trials data (5 to 7). On the left, the average M.S.E. for each level of model complexity for all n number of models. On the right, the number of times a predictor was selected for the desired level of complexity, decided on basis of the left image, for all n models.

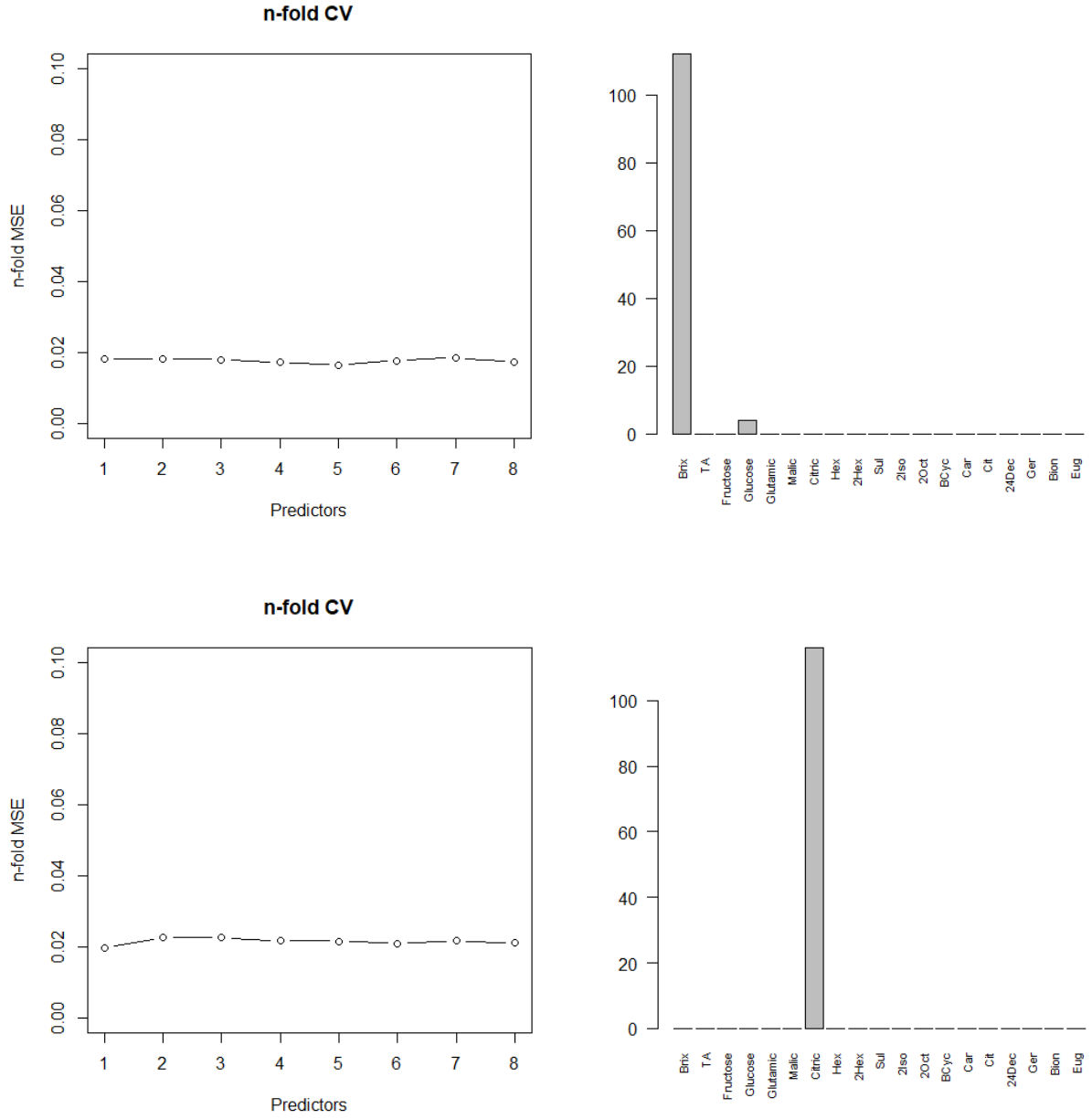


Figure A33 - Results from the Forward Stepwise Regression with LOOCV. On the top response variable Sweetness, on the bottom Sourness. Non-heated trials data (5 to 7). On the left, the average M.S.E. for each level of model complexity for all n number of models. On the right, the number of times a predictor was selected for the desired level of complexity, decided on basis of the left image, for all n models.

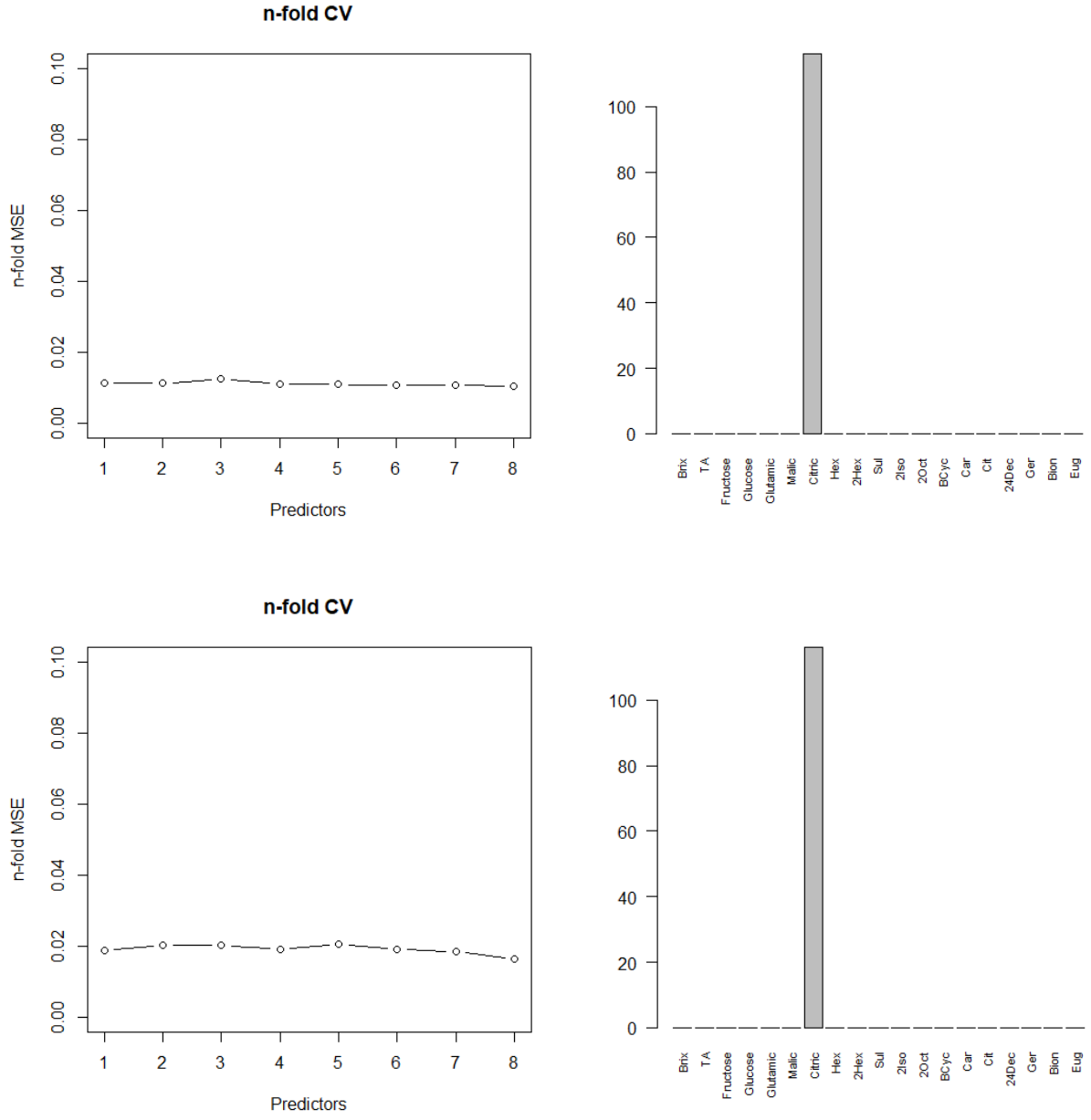


Figure A34 - Results from the Forward Stepwise Regression with LOOCV. On the top response variable Juiciness, on the bottom Firmness. Non-heated trials data (5 to 7). On the left, the average M.S.E. for each level of model complexity for all n number of models. On the right, the number of times a predictor was selected for the desired level of complexity, decided on basis of the left image, for all n models.

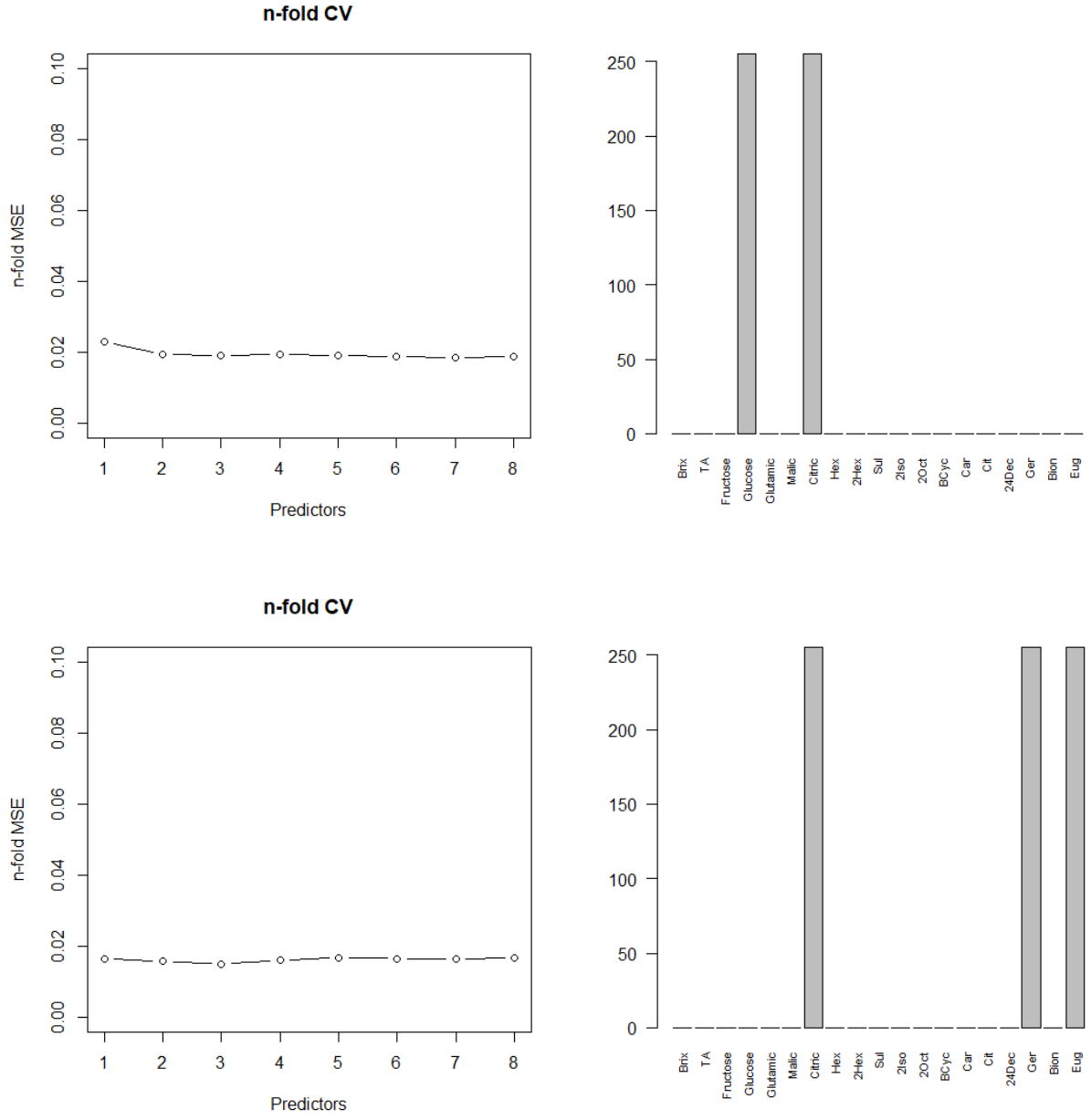


Figure A35 - Results from the Forward Stepwise Regression with LOOCV. On the top response variable Overall Liking, on the bottom Flavour Intensity. All trials data. On the left, the average M.S.E. for each level of model complexity for all n number of models. On the right, the number of times a predictor was selected for the desired level of complexity, decided on basis of the left image, for all n models.

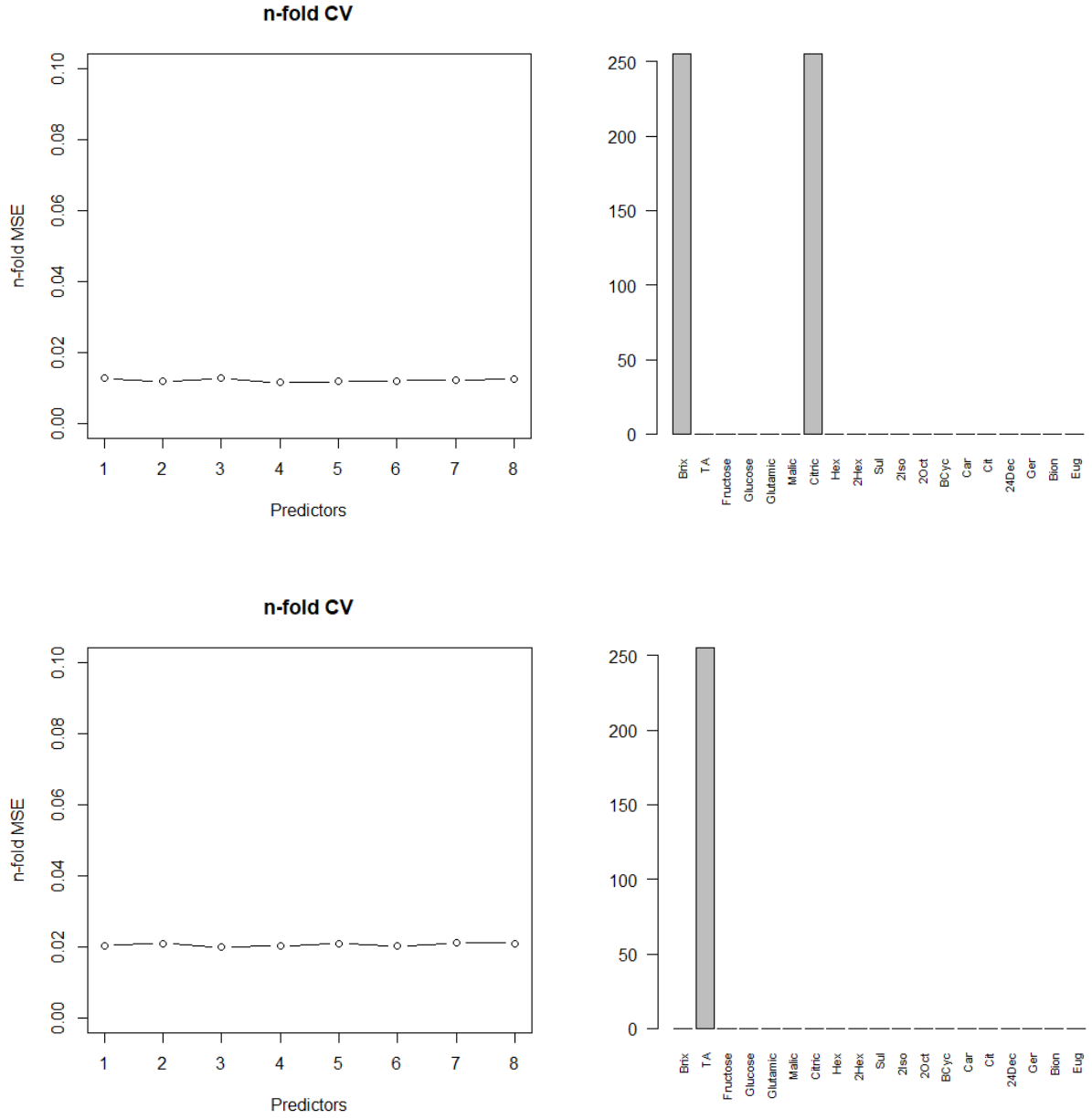


Figure A36 - Results from the Forward Stepwise Regression with LOOCV. On the top response variables Sweetness, on the bottom Sourness. All trials data. On the left, the average M.S.E. for each level of model complexity for all n number of models. On the right, the number of times a predictor was selected for the desired level of complexity, decided on basis of the left image, for all n models.

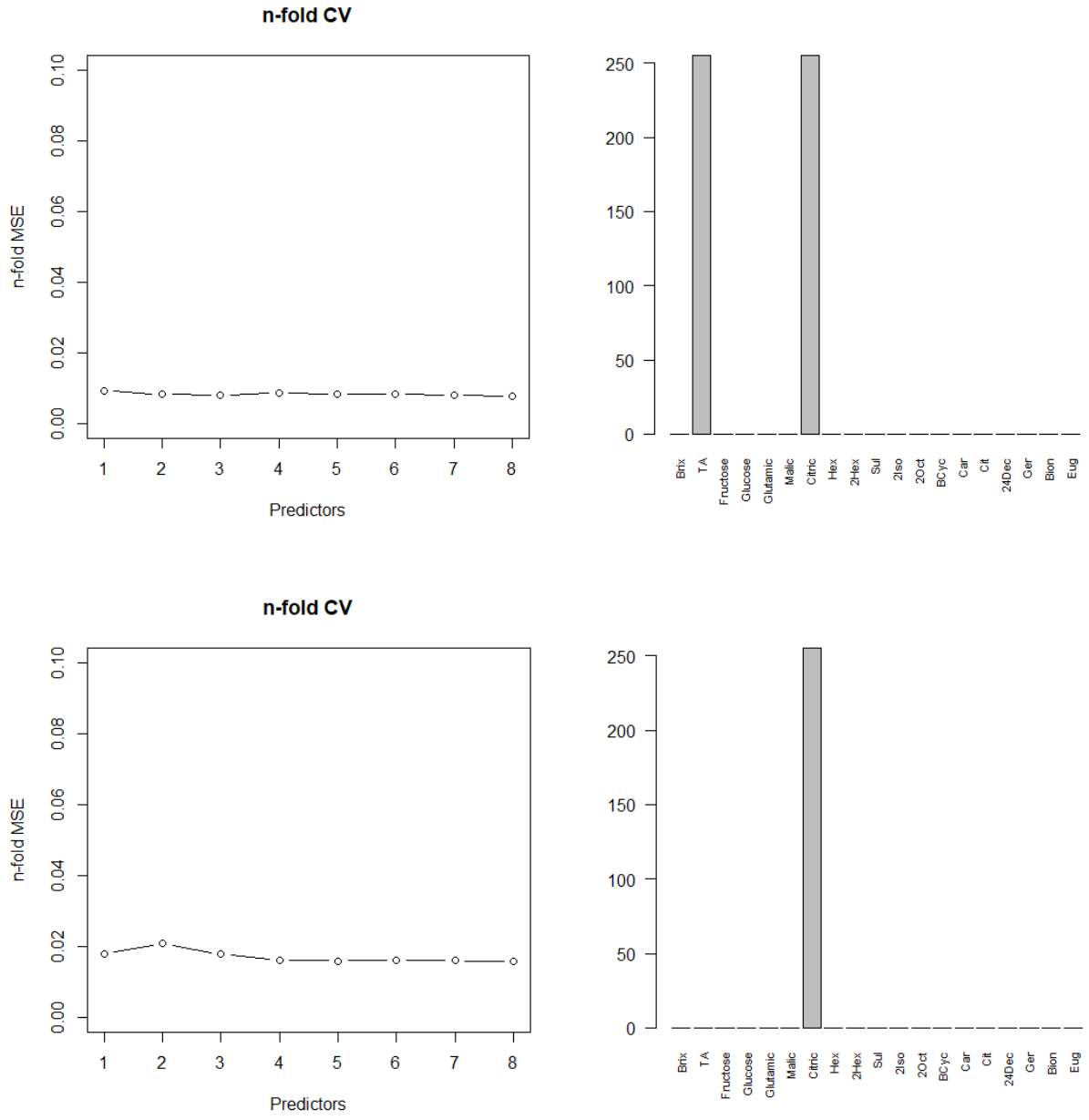


Figure A37 - Results from the Forward Stepwise Regression with LOOCV. On the top response variables Juiciness, on the bottom Firmness. All trials data. On the left, the average M.S.E. for each level of model complexity for all n number of models. On the right, the number of times a predictor was selected for the desired level of complexity, decided on basis of the left image, for all n models.

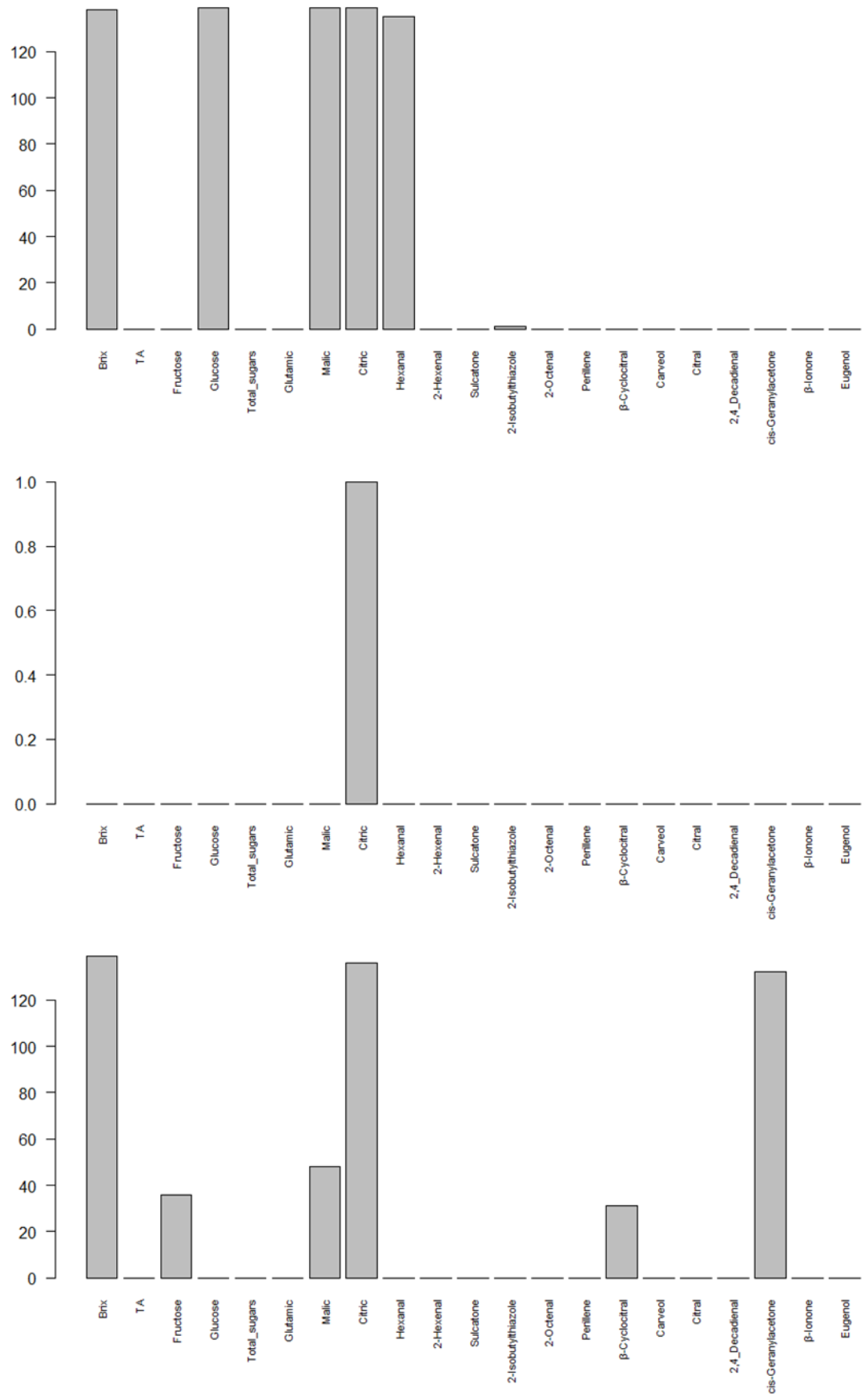


Figure A38 - Results from the Lasso Regression with LOOCV. Heated trials data. The graphs represent the number of times the predictors were selected for the model in all n models. From top to bottom: Overall Liking, Flavour Intensity, Sweetness

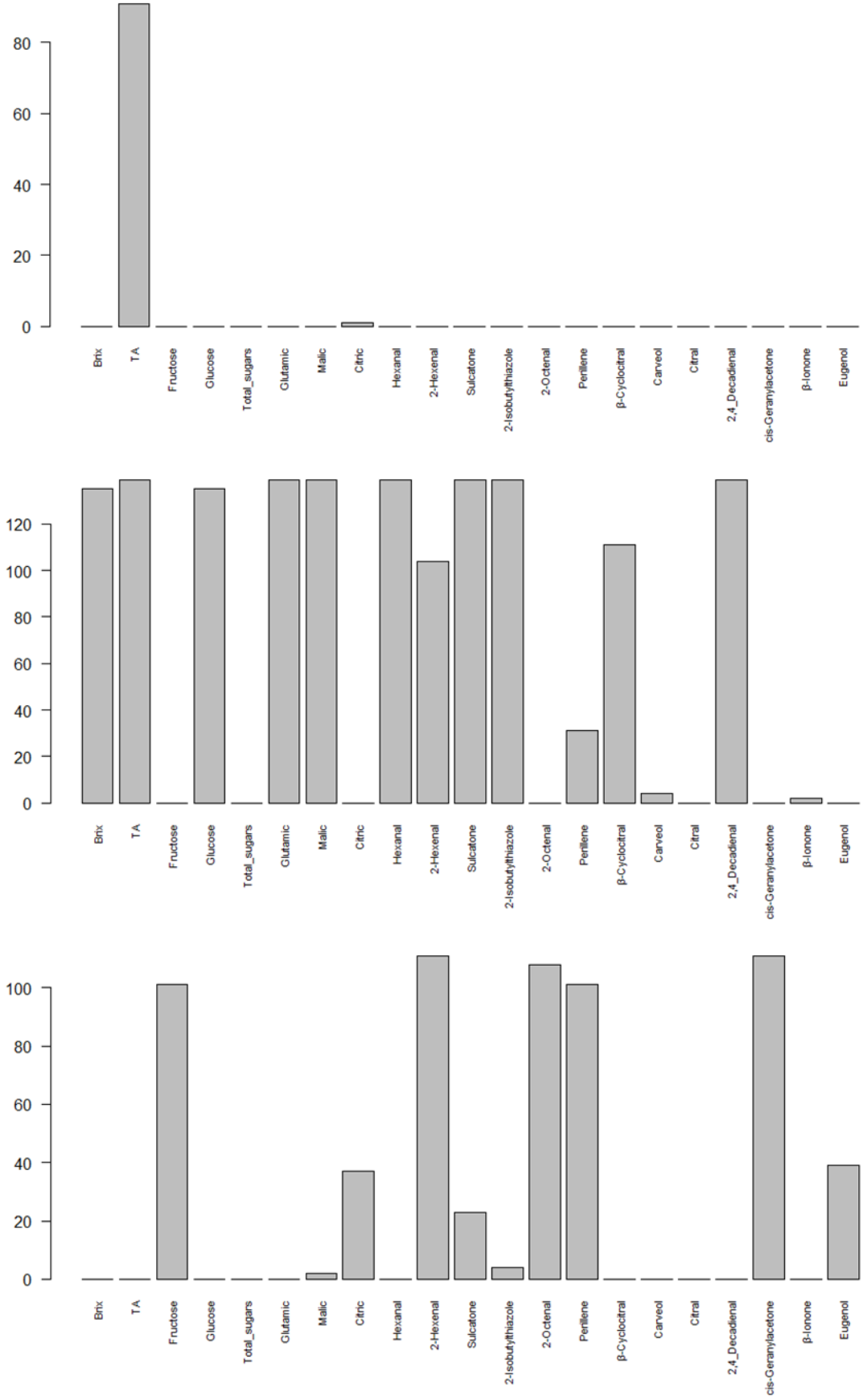


Figure A39 - Results from the Lasso Regression with LOOCV. Heated trials data. The graphs represent the number of times the predictors were selected for the model in all n models. From top to bottom: Sourness, Firmness, Juiciness

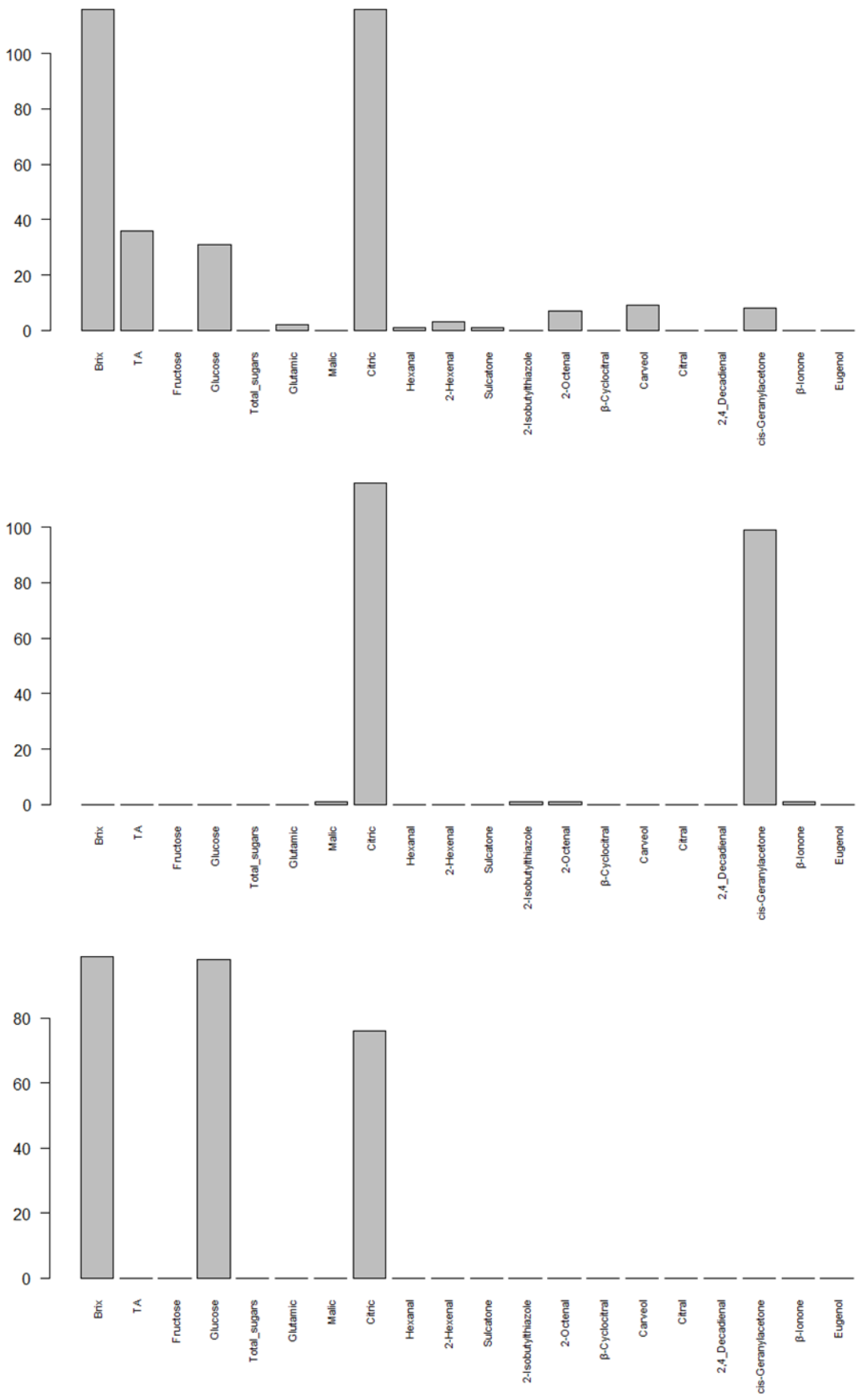


Figure A40 - Results from the Lasso Regression with LOOCV. Non-heated trials data. The graphs represent the number of times the predictors were selected for the model in all n models. From top to bottom: Overall Liking, Flavour Intensity , Sweetness

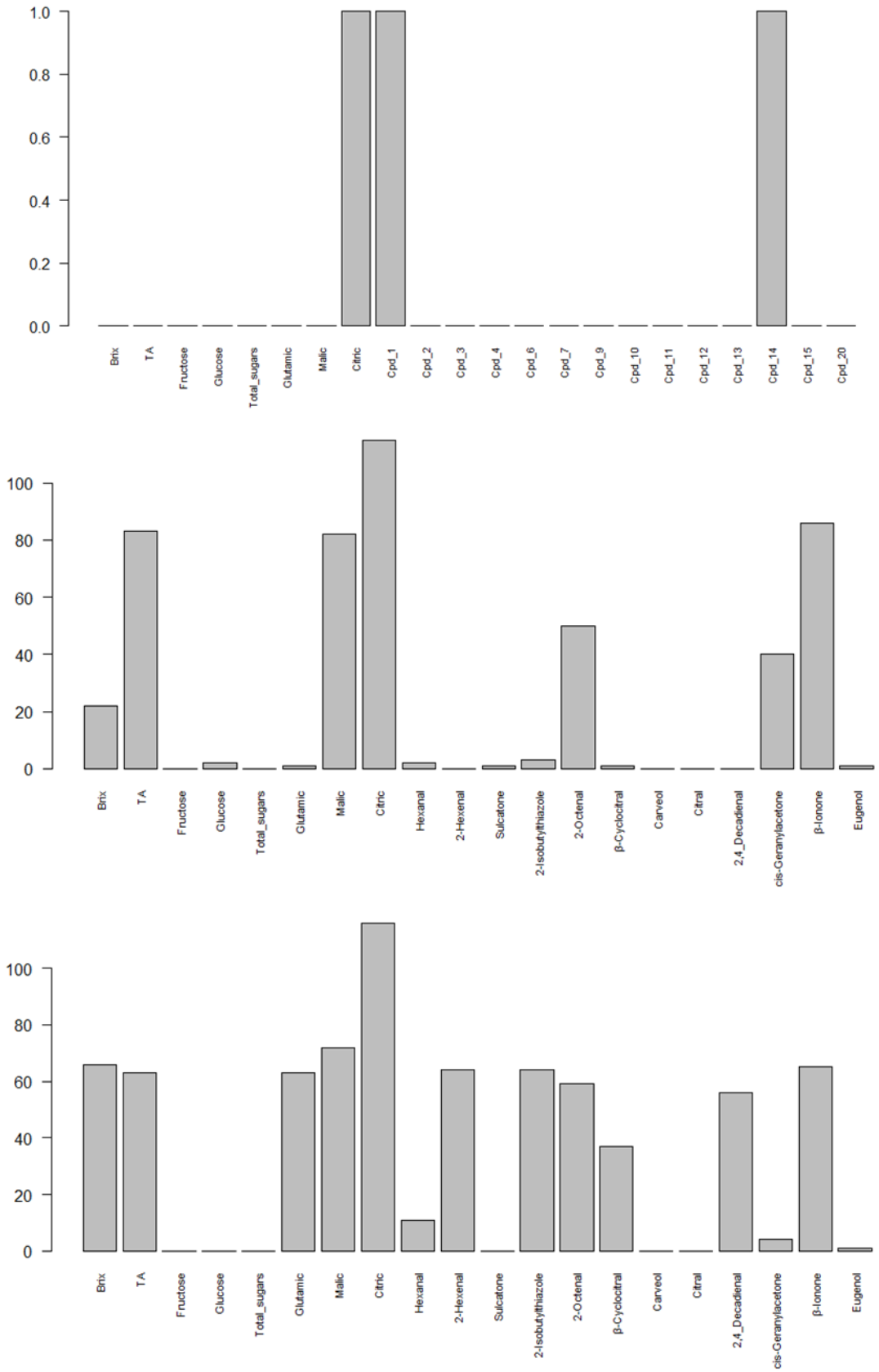


Figure A41 - Results from the Lasso Regression with LOOCV. Non-heated trials data. The graphs represent the number of times the predictors were selected for the model in all n models. From top to bottom: Sourness, Firmness, Juiciness

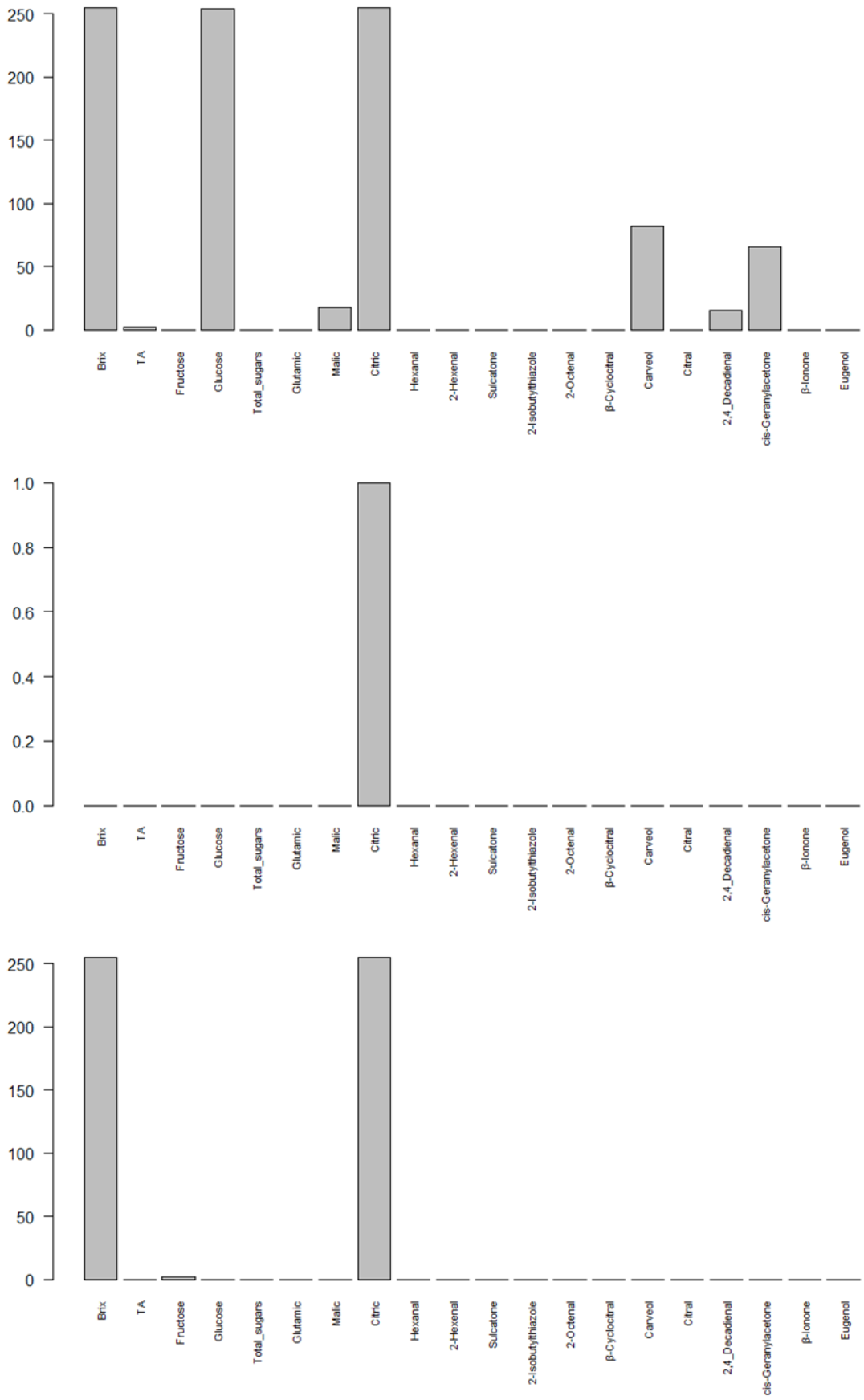


Figure A42 - Results from the Lasso Regression with LOOCV. All trials data. The graphs represent the number of times the predictors were selected for the model in all n models. From top to bottom: Overall Liking, Flavour Intensity , Sweetness

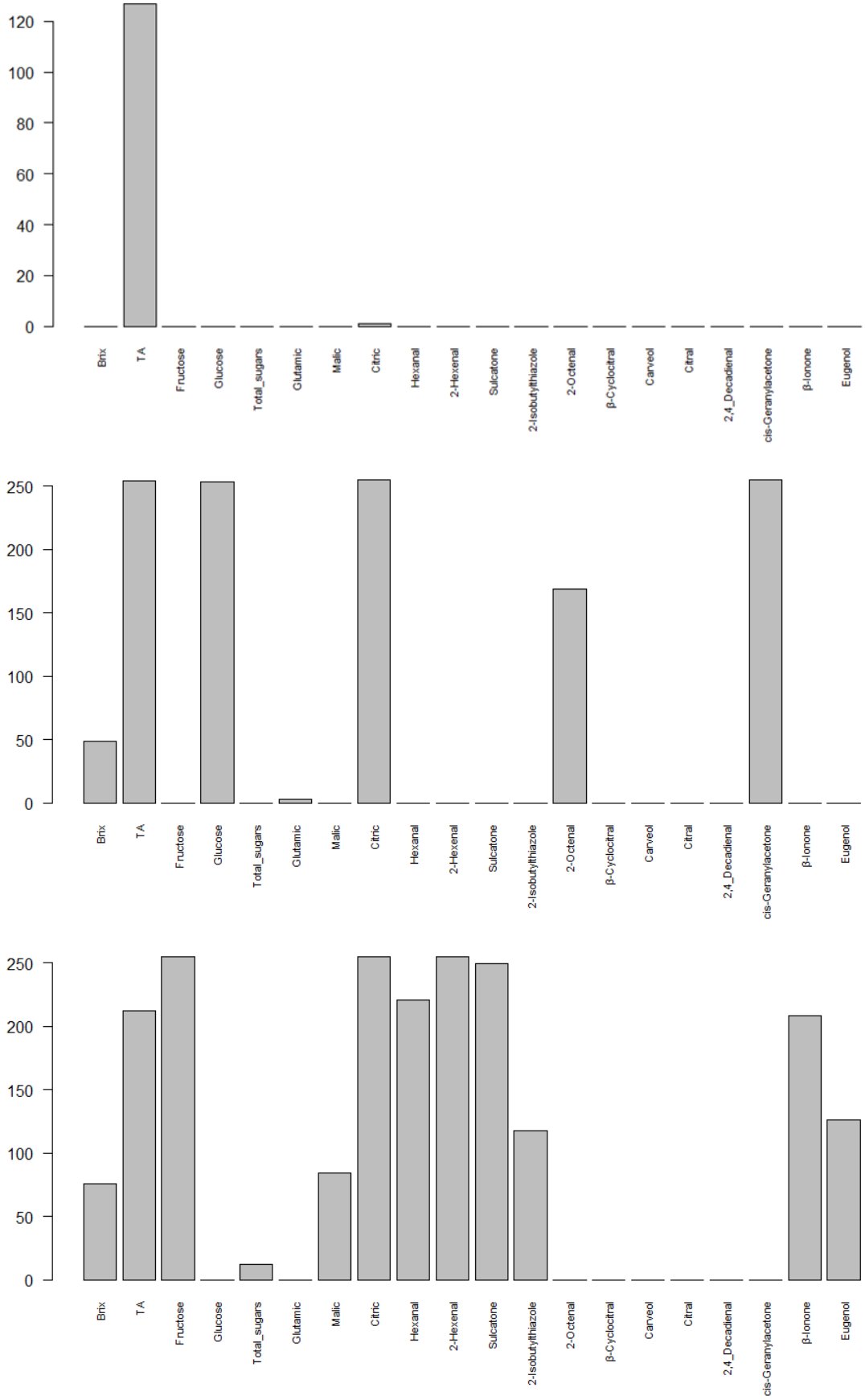


Figure A43 - Results from the Lasso Regression with LOOCV. All trials data. The graphs represent the number of times the predictors were selected for the model in all n models. From top to bottom: Sourness, Firmness, Juiciness

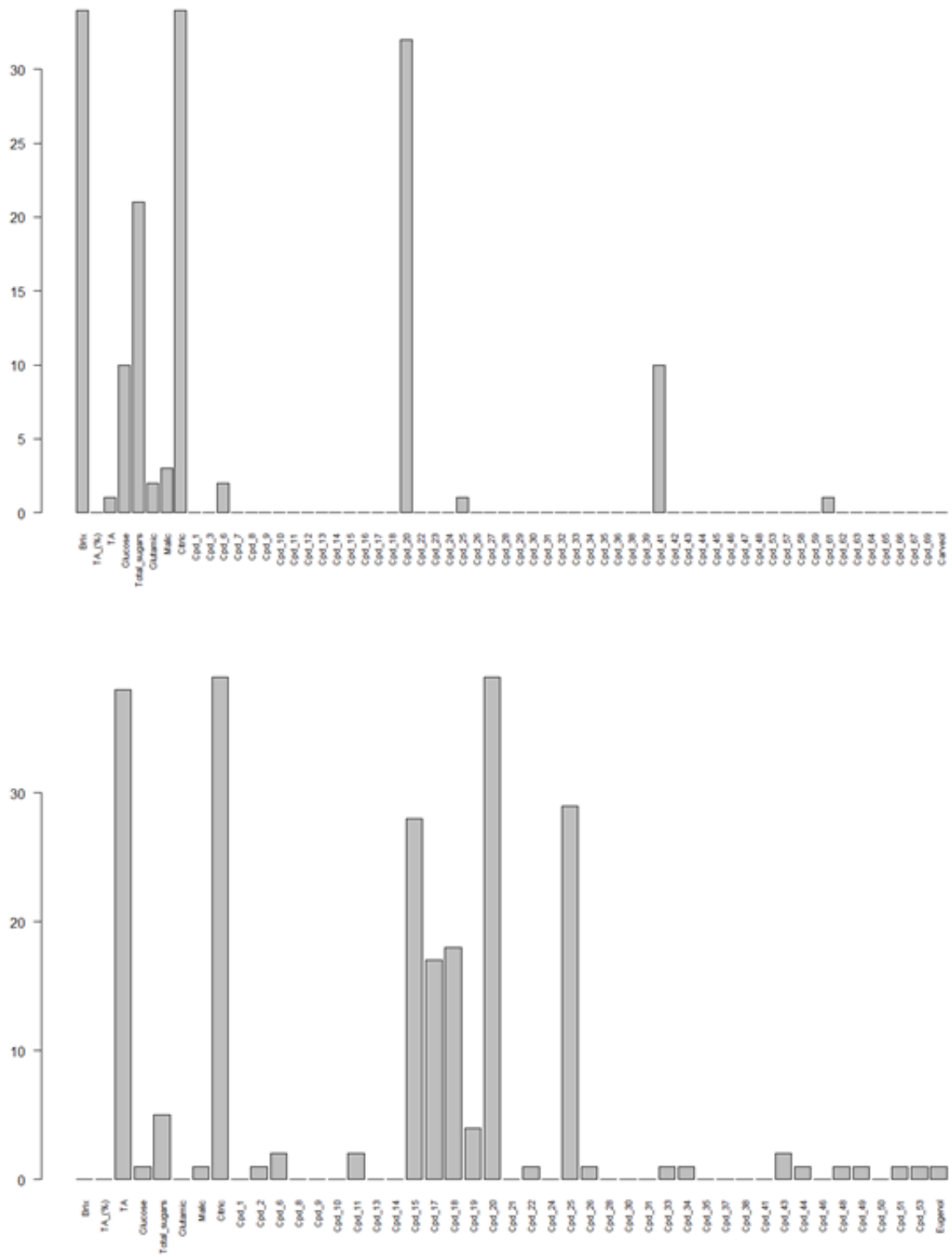


Figure A44 - Results from the Lasso Regression with LOOCV. Trial 2 on top and Trial 7 at the bottom. Both predicting Overall Liking. Cpd 20 corresponds to 2-Isobuthyliazole.