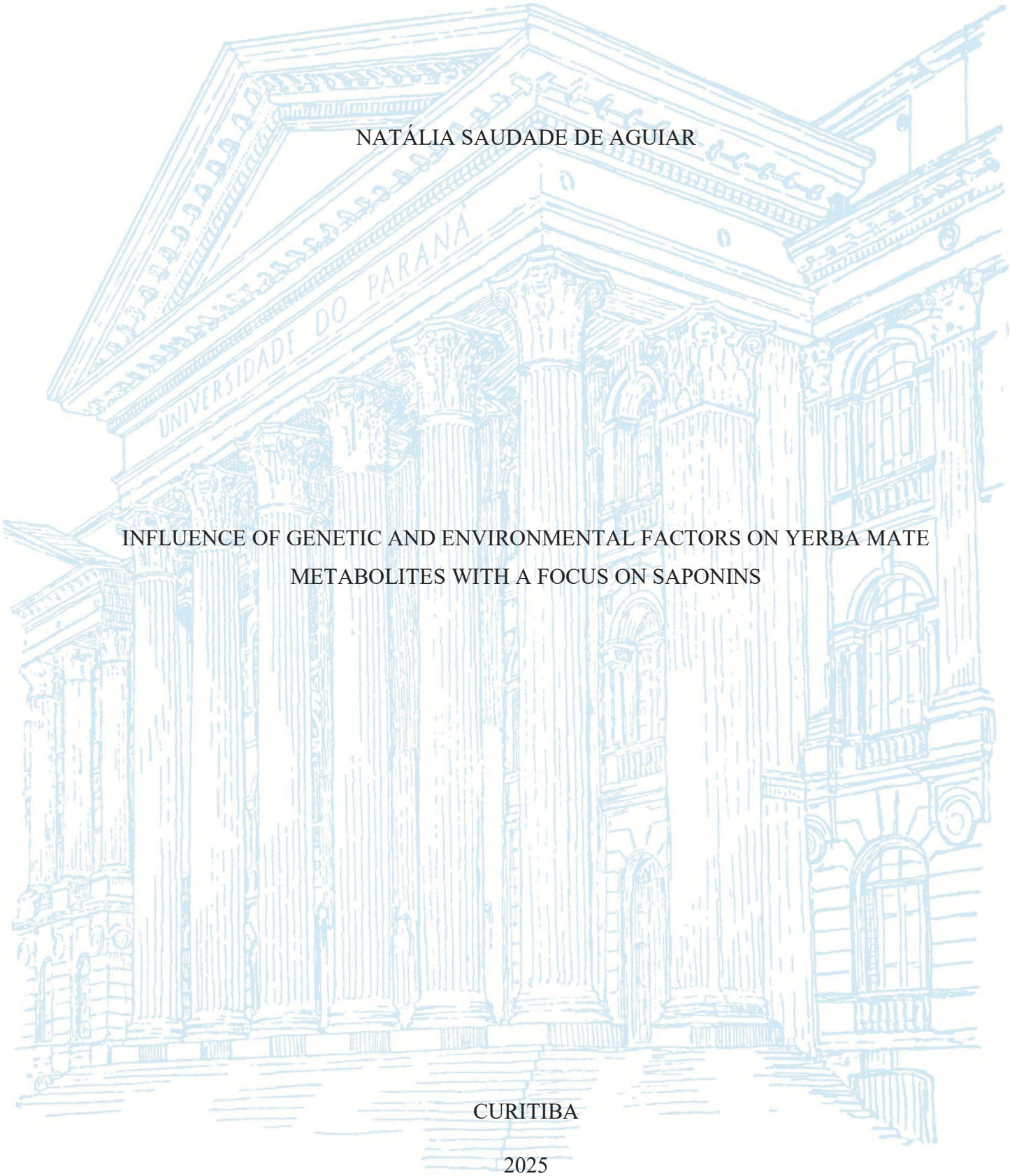
UNIVERSIDADE FEDERAL DO PARANÁ

NATÁLIA SAUDADE DE AGUIAR



INFLUENCE OF GENETIC AND ENVIRONMENTAL FACTORS ON YERBA MATE METABOLITES WITH A FOCUS ON SAPONINS

CURITIBA

2025

NATÁLIA SAUDADE DE AGUIAR

INFLUENCE OF GENETIC AND ENVIRONMENTAL FACTORS ON YERBA MATE METABOLITES WITH A FOCUS ON SAPONINS

Tese apresentada ao Programa de Pós-Graduação em Engenharia Florestal, Departamento de Ciências Florestais, Setor de Ciências Agrárias, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Doutora em Engenharia Florestal.

Orientador: Dr. Ivar Wendling
Coorientadores: Dra. Cristiane Aparecida Fioravante Reis
Dr. Marcelo Lazzarotto

CURITIBA

2025

Ficha catalográfica elaborada pela
Biblioteca de Ciências Florestais e da Madeira - UFPR

Aguiar, Natália Saudade de
Influence of genetic and environmental factors on yerba mate metabolites with a focus on saponins / Natália Saudade de Aguiar. - Curitiba, 2025.
1 recurso on-line : PDF

Orientador: Dr. Ivar Wendling
Coorientadores: Dra. Cristiane Aparecida Fioravante Reis
Dr. Marcelo Lazzarotto
Tese (Doutorado) - Universidade Federal do Paraná, Setor de Ciências Agrárias. Programa de Pós-Graduação em Engenharia Florestal. Defesa: Curitiba, 19/05/2025.

1. Erva-mate. 2. Erva-mate - Composição. 3. Metabólitos. 4. Saponinas. 5. Análise foliar. 6. Espectroscopia de infravermelho. 7. Clones (Plantas). I. Wendling, Ivar. II. Reis, Cristiane Aparecida Fioravante. III. Lazzarotto, Marcelo. IV. Universidade Federal do Paraná, Setor de Ciências Agrárias. V. Título.

CDD - 633.77
CDU - 633.77
634.0.285

Bibliotecária: Marilene do Rocio Veiga - CRB9/424

MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIÊNCIAS AGRÁRIAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO ENGENHARIA FLORESTAL - 40001016015P0

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação ENGENHARIA FLORESTAL da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **NATÁLIA SAUDADE DE AGUIAR**, intitulada: **INFLUENCE OF GENETIC AND ENVIRONMENTAL FACTORS ON YERBA MATE METABOLITES WITH A FOCUS ON SAPONINS**, que após terem inquirido a aluna e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutora está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 19 de Maio de 2025.

Assinatura Eletrônica
28/08/2025 16:26:41.0
IVAR WENDLING
Presidente da Banca Examinadora

Assinatura Eletrônica
28/08/2025 16:32:50.0
DAGMA KRATZ
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica
29/08/2025 13:59:07.0
FABRICIO AUGUSTO HANSEL
Avaliador Externo (EMBRAPA FLORESTAS)

Assinatura Eletrônica
28/08/2025 16:35:47.0
CRISTIANE VIEIRA HELM
Avaliador Externo (EMBRAPA FLORESTAS)

Assinatura Eletrônica
29/08/2025 14:38:34.0
CARLOS ANDRÉ STUEPP
Avaliador Externo (UNIVERSIDADE ESTADUAL DE PONTA GROSSA)

Dedico à todas as mulheres que contribuem para o avanço da ciência.

AGRADECIMENTOS

Primeiramente, quero agradecer ao meu orientador, Dr. Ivar Wendling, por me orientar do estágio da graduação ao doutorado. Realmente foi um prazer ter aprendido tanto com você! Agradeço pelas inúmeras oportunidades oferecidas a mim ao longo desses anos, que me tornaram uma profissional muito melhor. Aos meus coorientadores, Dra. Cristiane Reis e Dr. Marcelo Lazzarotto, que toparam esse desafio e foram essenciais para o desenvolvimento desta tese. Ambos são profissionais dedicadíssimos e eu os admiro muito.

À Dra. Cristiane Helm, por estar sempre à disposição e nunca ter medido esforços para me auxiliar! Obrigada por ceder o laboratório para as análises e pelas inúmeras conversas, oportunidades e trocas de experiências. Ao Dr. Fabricio Hansel que sempre me auxiliou nas análises laboratoriais, estatísticas, e esteve paciente e disposto durante todo esse período. Além de ser uma pessoa divertidíssima e agradável, tens realmente uma mente brilhante! Estendo este agradecimento a todos os profissionais da Embrapa Florestas que me auxiliaram, em especial: Dayanne, Paulino e Jonas. Além disso, o Dr. Gustavo Galo foi indispensável para as análises e discussões referentes ao Capítulo 3; muito obrigada pela ajuda e ensinamentos!

Aos amigos que conheci no Laboratório de Propagação de Espécies Florestais da Embrapa, e que levo pra vida: Manoela, Mônica, Jéssica, Leandro e Renata. Vocês estiveram comigo, longe ou perto, durante toda a pós-graduação e me indicaram o caminho. Somos um time e me orgulho muito de cada um de vocês! Bia e Débora, minhas companheiras de doutorado, agradeço pelo tempo que passamos juntas e pelo apoio nessa caminhada. Desejo todo sucesso do mundo pra vocês!

Além da Embrapa Florestas, sou grata também à UFPR e aos seus servidores, em especial aos professores; são instituições que tem todo o meu respeito e admiração, essenciais para a pesquisa florestal brasileira! Agradeço também à CAPES pela bolsa cedida a mim, que tornou possível minha permanência no doutorado. À Ervais do Futuro, empresa da qual faço parte e que conheci por meio da pesquisa, em especial aos meus sócios Ilo, Raul, Mônica e Sidiano.

Também quero agradecer à minha família, em especial ao meu pai, minha mãe e à minha irmã, por serem minha base. Sou eternamente grata pelo amor, oração e carinho recebidos! Em especial, agradeço ao meu marido Daniel: obrigada pelo carinho, paciência, apoio e incentivo sempre. A vida é muito melhor e mais divertida com você!

A percepção do desconhecido é a mais fascinante das experiências. O homem que não tem os olhos abertos para o mistério passará pela vida sem ver nada.

Albert Einsten

RESUMO

As folhas de erva-mate possuem uma diversidade de compostos, muitos com reconhecidos efeitos bioativos. Entre estes, destacam-se as saponinas, as metilxantinas e os polifenóis. Estudos demonstraram que fatores intrínsecos (genética, idade da folha e sexo da planta, por exemplo) e ambientais (luminosidade, nutrição, entre outros) interferem nos teores de metilxantinas e polifenóis da erva-mate; entretanto, ainda há muitas lacunas de pesquisa referente às saponinas e aos metabólitos primários. Neste estudo, folhas maduras de nove clones de erva-mate provenientes de dois testes clonais (geograficamente distantes e com diferentes sistemas de cultivo – pleno sol e sombreado) foram coletadas em duas estações do ano (inverno e verão). Estas folhas foram secas em micro-ondas e trituradas mecanicamente até atingir granulometria menor que 1 mm, para determinação do teor de saponinas totais e obtenção dos espectros de Infravermelho Médio (MIR) e Próximo (NIR). Para quantificação de saponinas totais nas folhas da espécie, foi necessária a otimização do método espectrofotométrico vanilina-ácido sulfúrico. Dessa forma, o Capítulo 1 desta tese aborda a metodologia de amostragem de folhas, os principais parâmetros para ajustes do método espectrofotométrico e também traz resultados preliminares analisando quatro cultivares clonais de erva-mate, com diferenças significativas entre elas. O Capítulo 2 possui como objetivo analisar os efeitos de clones, épocas de colheita e sítios de cultivo (9 x 2 x 2) sobre o teor de saponinas totais, estimando os parâmetros genéticos e predizendo os valores genotípicos. Os resultados demonstraram forte controle genético sobre este caráter, com elevadas herdabilidade e acurácia de seleção. Os valores genotípicos para os teores de saponinas totais variaram de 28,13 (Yari) a 51,54 (BRS 409) mg g^{-1} de massa seca, e a seleção de clones com baixos ou altos teores de saponinas é indicada dependendo da finalidade industrial da matéria-prima. Este foi o primeiro estudo realizado com avaliação genotípica de compostos bioativos em testes clonais da espécie. E no Capítulo 3, a análise multivariada ANOVA Common Dimensions (AComDim) foi empregada para unir os espectros MIR e NIR em um modelo global e determinar a variação atribuída por cada fator analisado e as interações entre eles. A maior parte da variação foi determinada pela época de colheita e pelo sítio (ambos com 7%), seguida pela interação entre estes fatores (5,7%), e pelo efeito de clones (4%). Por esta técnica de bioespectroscopia foi possível detectar as bandas de absorção mais determinantes para cada fator, sendo estas relacionadas a moléculas do metabolismo primário vegetal. Assim, as técnicas bioespectroscópica, quimiométrica e estatística aplicadas foram efetivas e possuem potencial

de aplicação para padronizar a matéria-prima ainda em campo e no controle de qualidade na indústria.

Palavras-chave: *Ilex paraguariensis*; metabólitos secundários; saponinas triterpênicas; folhas; estação do ano; local de cultivo; teste clonal; bioespectroscopia; NIR; MIR.

ABSTRACT

Yerba mate leaves contain a wide diversity of compounds, many of which exhibit well-documented bioactive effects. Among these, saponins, methylxanthines, and polyphenols stand out. Numerous studies have demonstrated that intrinsic factors (such as genetics, leaf age, and plant sex) and environmental conditions (such as light availability and nutrient supply) influence the levels of methylxanthines and polyphenols in yerba mate. However, there are still significant research gaps concerning saponins and primary metabolites. In this study, mature leaves from nine yerba mate clones, originating from two clonal trials (geographically distant and cultivated under different systems – full sun and shade), were collected during two seasons (winter and summer). The leaves were microwave-dried and mechanically ground into granulometry less than 1 mm to determine total saponin content and acquire Mid-Infrared (MIR) and Near-Infrared (NIR) spectra. Quantification of total saponins in the leaves required optimization of the vanillin-sulfuric acid spectrophotometric method. Accordingly, Chapter 1 of this thesis presents the leaf sampling methodology, key parameters for adjusting the spectrophotometric method, and preliminary results analyzing four yerba mate clonal cultivars, with significant differences between them. Chapter 2 aims to assess the effects of clones, harvest seasons, and cultivation sites (9 x 2 x 2) on total saponin content, as well as to estimate genetic parameters and predict genotypic values for this trait. This represents the first study to analyze genetic parameters of bioactive compounds in clonal trials of the species. The results indicate strong genetic control over this trait, with high heritability and selection accuracy. Genotypic values for total saponin content ranged from 28.13 (Yari) to 51.54 (BRS 409) mg g^{-1} of dry mass, and the selection of clones with either low or high saponin levels is recommended depending on the intended industrial application of the raw material. In Chapter 3, the multivariate analysis technique ANOVA Common Dimensions (AComDim) was applied to combine MIR and NIR spectra into a global model and to determine the variation attributable to each factor and their interactions. The greatest variation was explained by harvest season and site (both 7%), followed by their interaction (5.7%), and the clone effect (4%). This biospectroscopic approach allowed identification of the absorption bands most affected by each factor, which were primarily associated with molecules from the plant primary metabolism. Thus, the applied biospectroscopic, chemometric, and statistical techniques proved to be effective and show potential for standardizing raw material in the field and for quality control in the industry.

Keywords*:* *Ilex paraguariensis*; secondary metabolites; triterpene saponins; mature leaves; seasons; cultivation site; clonal trials; biospectroscopy; NIR; MIR.

LIST OF FIGURES

LIST OF TABLES

SUMMARY

1 GENERAL INTRODUCTION

Traditionally, leaves and fine branches of yerba mate (*Ilex paraguariensis* A. St.-Hil.) are consumed in non-alcoholic beverages, a practice adopted centuries ago by the native peoples of southern South America, including the Guarani, Amerindian, and Caingangue peoples (Bracesco *et al.*, 2011; Croge; Cuquel; Pintro, 2020). From these origins, the consumption of yerba mate has become culturally embedded through infusions such as chimarrão, tereré, and chá mate (mate tea). However, research shows that the potential of yerba mate extends far beyond the traditional uses: species' leaves can be incorporated into beers, soft drinks, baked goods, sweets, cheeses, and many other products (Croge; Cuquel; Pintro, 2020). In addition, there is great potential for pharmaceutical and cosmetic applications, including use as a dietary supplement, antiseptic, capsule, sunscreen, and other products aimed at promoting health benefits (Gerber *et al.*, 2023). This raw material has also been used in the development of nutraceuticals, functional products, antioxidants and food preservatives, natural colorants, among others (Cardozo Junior; Morand, 2016; Gerber *et al.*, 2023).

The wide variety of products derived from yerba mate is primarily due to the presence of secondary metabolites, especially in its leaves. Among the most studied metabolites in the species are the methylxanthines caffeine and theobromine, which are compounds that stimulate the central nervous system. The leaves also contain high concentrations of phenolic compounds, particularly monocaffeoylquinic or chlorogenic acids (5-caffeoylquinic, 3-caffeoylquinic, and 4-caffeoylquinic), as well as dicaffeoylquinic acids (3,4-dicaffeoylquinic, 3,5-dicaffeoylquinic, and 4,5-dicaffeoylquinic) (Meinhart *et al.*, 2017). These compounds are responsible for the species' high antioxidant capacity (Lima *et al.*, 2016; Mateos *et al.*, 2018), with total monocaffeoylquinic acid contents reaching up to 12% of the leaf dry mass (Aguiar *et al.*, 2024a).

Another group of compounds found in yerba mate leaves are terpenes, among which saponins stand out as highly diverse and complex molecules. Saponins, their biosynthetic intermediates, and their derivatives have important applications in the cleaning, cosmetics, food, agronomic, and pharmaceutical industries (Güçlü-Üstündag; Mazza, 2007; Moses; Papadopoulou; Osbourn, 2014). These compounds are widely used as emulsifiers, foaming and detergent agents, feed additives, plant growth regulators, and in soil bioremediation (Güçlü-

Üstündag; Mazza, 2007). They can also be employed in the food industry as natural surfactants and preservatives, contributing to the control of microbial and fungal spoilage (Cheok; Salman; Sulaiman, 2014). Moreover, saponins have been reported to exhibit numerous pharmacological activities, including cytotoxic effects on cancer cells, immune response enhancement, antioxidant, antibiotic, antidiabetic, and cholesterol-lowering properties – effects mainly attributed to their ability to permeabilize and disrupt cell membranes (Augustin *et al.*, 2011; Cheok; Salman; Sulaiman, 2014; Moses; Papadopoulou; Osbourn, 2014).

An example of a plant species that is a source of triterpene saponins is the Chilean tree species *Quillaja saponaria* Molina. Its bark contains a high concentration of saponins, from which important molecules with well-established biological activities are extracted and used across various industries, including agriculture, human and animal nutrition, cosmetics, and pharmaceuticals, especially as adjuvants and immunostimulants (Reichert; Salminen; Weiss, 2019). For this reason, extracts from this species are among the main non-wood forest products of Chile, with a significant increase in export value, particularly in 2020, due to their use in COVID-19 vaccines (Poblete Hernández, 2022). Similar to *Q. saponaria*, yerba mate also has the potential to become a valuable industrial source of saponins, especially for pharmaceutical applications, as suggested by both *in vitro* and *in vivo* studies (Nagatomo *et al.*, 2022; Puangpraphant; Berhow; de Mejia, 2011; Puangpraphant; De Mejia, 2009). It is estimated that over 70% of all drugs approved worldwide between 1981 and 2006 were either derived from or structurally related to natural compounds (Gerber *et al.*, 2023), highlighting the importance of phytochemistry in the discovery of such compounds.

However, in order to meet the demands of these industries, standardization and quality control of the raw material are essential. The levels of secondary metabolites in yerba mate leaves may vary according to factors such as plant and leaf age, harvest season, cultivation environment, and especially genotype (Aguiar *et al.*, 2024a; Benedito *et al.*, 2023; Duarte *et al.*, 2023; Vieira *et al.*, 2021). Genetic parameter analyses have already been conducted with yerba mate for these compounds, and especially for methylxanthines, high heritability is known, indicating strong genetic control (Cardozo Junior *et al.*, 2010; Friedrich *et al.*, 2017; Nakamura *et al.*, 2009). For caffeine, for instance, in the same location, cultivation system, and harvest date, contents can naturally vary from 0.348 to 23.846 mg g^{-1}, and the genetic materials show stable levels over the years (Benedito *et al.*, 2023). As a result, clonal propagation of selected

yerba mate genotypes has been employed to produce standardized raw material with minimally predictable caffeine content even before plantation.

In addition to the strong genetic influence, environmental factors also affect the synthesis and accumulation of secondary metabolites. Regarding the cultivation environment, numerous factors act simultaneously, such as meteorological and edaphic conditions, availability of water, light, and nutrients (Yang *et al.*, 2018). As for the effect of shading, a study on yerba mate clones showed that lower light availability led to higher caffeine accumulation, while theobromine content decreased (Aguiar et al., 2024a). With respect to harvest season, there is a significant influence although not yet fully elucidated on the composition of compounds in the species (Aguiar *et al.*, 2024a; Butiuk *et al.*, 2016; Duarte *et al*., 2023; Ferrera *et al.*, 2016; Rakocevic *et al.*, 2023). Similarly, some compounds are also influenced by the time of day (MELO, 2018). Therefore, the selection of clones combined with appropriate management practices to each industrial purpose is essential for standardizing the raw material.

Many studies have been conducted to understand the effects of intrinsic and environmental factors on the accumulation of polyphenols and methylxanthines in yerba mate leaves; however, there are still knowledge gaps regarding the response of saponins and primary metabolites. Near-Infrared (NIR) and Mid-Infrared (MIR) Spectroscopy are analytical techniques that utilize the infrared region of the electromagnetic spectrum to analyze materials with little or no sample preparation (Berhow *et al.*, 2020). Compared to conventional quantitative analysis methods, these are fast, non-destructive, and environmentally friendly technologies, as they do not require extractions or additional chemical reactions (Hou *et al.*, 2022). When combined with chemometrics and multivariate analyses, these techniques allow us to infer the contribution of each factor and their interactions to the plant metabolome – primary and secondary metabolites. Therefore, these techniques can support the selection of clones based on industrial objectives, the definition of more appropriate management practices, and quality control of the raw material.

To gain a deeper understanding of the impact of genetic and environmental factors on yerba mate, mature leaves were analyzed for total saponin content (through refinement of the spectrophotometric methodology) and MIR and NIR spectral fingerprints. Through univariate and multivariate analyses, it was possible to recommend suitable genetic materials for specific

industrial applications, with a focus on saponins, as well as detect changes in the species' metabolome.

1.1 OBJECTIVES

1.1.1 General Objective

To analyze the effect of clones, cultivation sites, and harvest seasons on the total saponin content in mature yerba mate leaves, genotypically evaluating this trait, and infer the impact of these factors on the plant metabolome through infrared spectral fingerprints.

1.1.2 Specific Objectives

1) Optimization of a rapid, cost-effective, and accurate analytical method for the quantification of total saponin content.
2) To apply the adapted spectrophotometric method to quantify total saponin content in mature leaves collected from different clones, cultivation sites, and harvest seasons.
3) To estimate genetic parameters for saponin content and select high-performing clones for saponin production based on genotypic values.
4) To assess the influence of individual experimental factors on yerba mate leaf powder using MIR and NIR spectral fingerprints, and to identify classes of chemical compounds responsive to these variations.

1.2 THESIS STRUCTURE

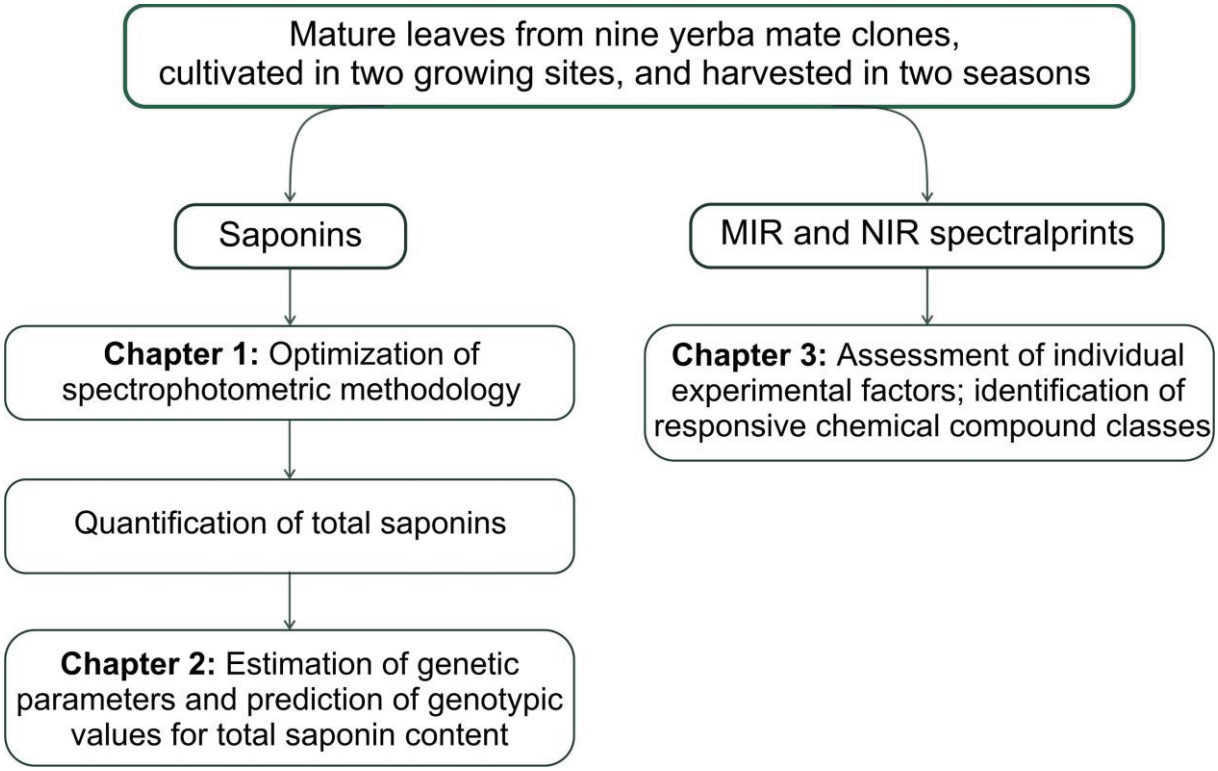
To address the proposed objectives, this thesis includes an initial section with a literature review, followed by three chapters, entitled:

Chapter 1: Optimizing the vanillin-acid sulfuric method to total saponin content in leaves of yerba mate clones (published by Chemistry & Biodiversity in 15 February 2024)

Chapter 2: Total saponins in yerba mate leaves: influence of clones, sites and harvest seasons (published by Crop Breeding and Applied Biotechnology in 31 March 2025)

Chapter 3: Multiblock NIR and MIR spectralprint through AComDim to evaluate the effects of growing site, harvest season, and clone on yerba mate leaves composition (published by Food Chemistry in 20 February 2025)

FIGURE 1 – GRAPHICAL SUMMARY OF THE THESIS



SOURCE: The author (2025).

2 LITERATURE REVIEW

2.1 YERBA MATE

Yerba mate is a species belonging to the Aquifoliaceae family, native to South America, and distributed across parts of Brazil, Argentina, and Paraguay (Oliveira; Rotta, 1983). In Brazil, the species is mainly concentrated in the southern region and displays a

geographical distribution primarily associated with the Atlantic Domain, characterized by a subtropical and humid temperate climate, in areas of higher elevation, with mild summers and rainy winters (Silva; Higuchi; Silva, 2018). It grows spontaneously in the understory of forests with *Araucaria angustifolia* and in subtropical forests of southern Brazil (Oliveira; Rotta, 1983), and is considered a characteristic species of the Araucaria Forest, or Mixed Ombrophilous Forest (Reis *et al.*, 2018). This evergreen tree can reach up to 30 meters in height under natural conditions and is classified as a climax species tolerant to shade (Carvalho, 2003).

The practice of consuming beverages prepared from yerba mate leaves was adopted centuries ago by Indigenous peoples and has become a cultural tradition in Paraguay, Uruguay, northeastern Argentina, and southern Brazil (Bracesco *et al.*, 2011). This forest species drives a significant productive and industrial sector, particularly in the southern states of Brazil (Dallabrida *et al.*, 2016). It is estimated that the yerba mate production chain involves approximately 725 processing companies in Brazil and engages more than 700,000 workers, both directly and indirectly (Ferron, 2016). Unlike other producing countries, where only monoculture plantations of yerba mate exist, 41.7% of Brazil's yerba mate comes from harvesting in native forests (IBGE, 2022a, b). Considering both forms of production – exploration and cultivation – Brazil produced over 1 million tons in 2022, with a production value exceeding R$ 1.49 billion (IBGE, 2022a, b). Currently, yerba mate is the leading non-wood forest product in Brazil in terms of production volume, and the second in production value, surpassed only by açaí. These figures highlight the social and economic importance of yerba mate, especially in southern Brazil.

In addition to Brazil, Argentina and Paraguay are also producers of yerba mate, with a combined production of 1,653,166 tons in 2022, not including Brazil's wild-harvested production (FAOSTAT, 2022). Most of this production is intended for the domestic market or neighboring countries, such as Uruguay; however, the international market has been expanding and increasingly favorable to the commercialization of this product (Landau; Alves Da Silva; Torres, 2020). Beyond South America, yerba mate is also consumed in other countries, including the United States, Germany, Syria, Spain, Italy, Australia, France, Japan, Korea, and Russia (Cardozo Junior; Morand, 2016).

Yerba mate contains a wide diversity of molecules, with more than 200 compounds extracted from its leaves using methanol (Melo *et al.*, 2020). Among its main bioactive compounds are terpenes (such as carotenoids and saponins), polyphenols (including

caffeoylquinic acids and flavonoids), and alkaloids (methylxanthines such as caffeine and theobromine) (Cardozo Junior; Morand, 2016). According to Valduga et al. (2019), few plant species contain such high levels of these secondary metabolites, as observed in yerba mate, which highlights its great potential for exploitation. In addition, its leaves also contain a broad range of other constituents, including minerals, water-soluble vitamins, proteins, and fibers (Cardozo Junior; Morand, 2016; Valduga *et al.*, 2019). Studies have shown that yerba mate possesses a variety of properties, including antimicrobial, antioxidant, anti-inflammatory, antidiabetic, antimutagenic, anti-obesity, diuretic, cholesterol-regulating, central nervous system-stimulating, cardioprotective, analgesic, wound-healing effects, among many others (Cardozo Junior; Morand, 2016; Clemente *et al.*, 2024; Paluch *et al.*, 2021).

For several of these compounds in yerba mate, particularly secondary metabolites, there is a strong genetic influence (Cardozo Junior *et al.*, 2010; Friedrich *et al.*, 2017; Nakamura *et al.*, 2009). Therefore, cloning of selected genotypes is a key tool for the chemical standardization of raw material (Duarte *et al.*, 2023; Vieira *et al.*, 2021). As a forest species cultivated for leaf production, the most commonly used vegetative propagation method is cutting for rescuing mother plants, followed by mini-cuttings for large-scale multiplication (Wendling *et al.*, 2017a). In Brazil, Embrapa (Brazilian Agricultural Research Corporation) has been selecting genetic materials, establishing and evaluating field genetic improvement trials, and registering yerba mate cultivars for commercial release – contributing to increased productivity and raw material standardization. Clones such as BRS 408, BRS 409, Yari, and Aupaba, evaluated in this study, are cultivars already available to rural producers (Wendling *et al.*, 2017a, b). Overall, clonal silviculture of the species is still recent and not yet widely adopted, but it shows strong potential, particularly for industries focused on bioactive compounds.

2.2 SAPONINS: TYPES AND APPLICATIONS

Saponins are secondary metabolites and represent an extremely diverse group of molecules: more than 200 saponins have already been identified in the plant kingdom (Hussain *et al.*, 2019), in addition to those found in marine organisms such as sea cucumbers and starfish (Güçlü-Üstündag; Mazza, 2007). They are considered important components of plant defense

systems, acting against pathogens, insects, and other herbivorous animals. Furthermore, studies indicate that these compounds play significant roles in various stages of plant development, including seed germination, growth and differentiation, fruiting, and nodulation (Moses; Papadopoulou; Osbourn, 2014).

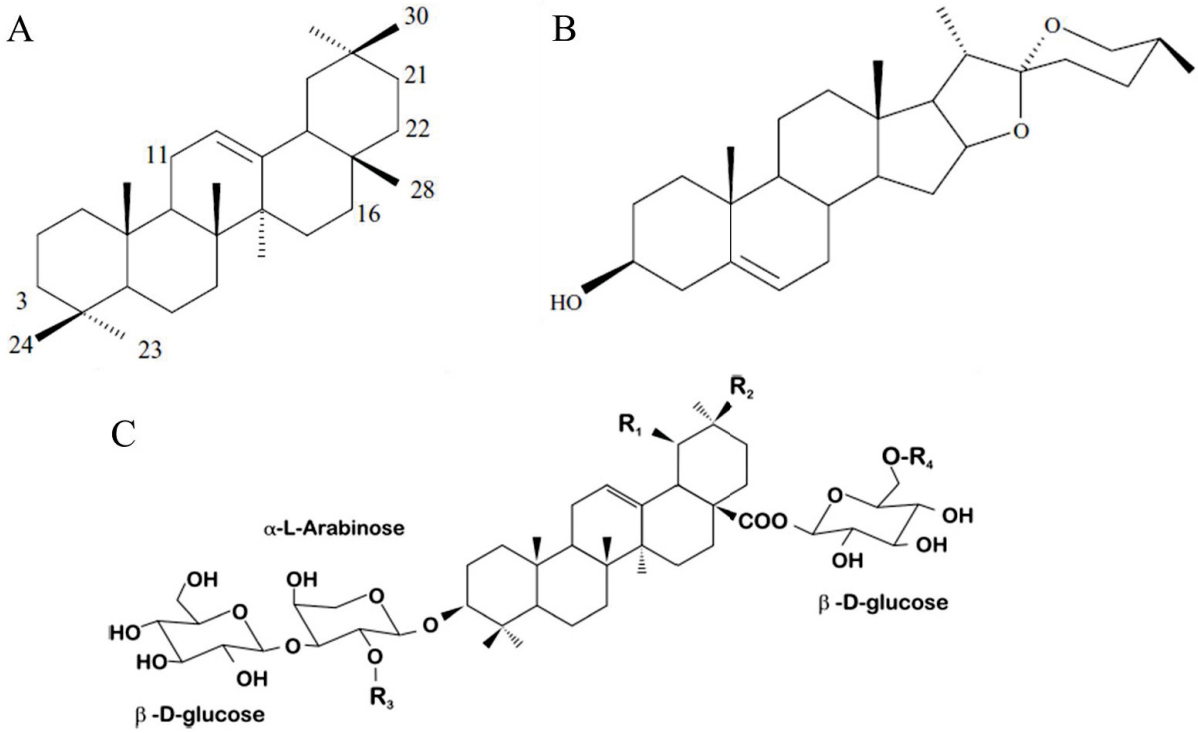
These complex molecules are composed of aglycones (also called sapogenins), which can be triterpenoid or steroidal in nature, linked to sugar chains (Figure 2A, B). Saponins are generally categorized based on the number of sugar chains and are classified as mono-, bi-, or tridesmosidic (with one, two, or three sugar chains, respectively) (Güçlü-Üstündag; Mazza, 2007). The combination of a nonpolar, hydrophobic aglycone and a polar, hydrophilic sugar moiety gives these molecules a highly amphipathic character, as well as emulsifying and foaming properties (Moses; Papadopoulou; Osbourn, 2014). When shaken with water, they produce a soap-like foam hence their name, derived from the Latin word sapo, meaning soap. Due to the great variability in aglycones, chemical linkages, and the number and type of saccharides (Figure 2C), saponins exhibit diverse physical, chemical, and biological properties (Mateos *et al.*, 2017).

Saponins, along with their biosynthetic intermediates and derivatives, have important applications in the cleaning, cosmetics, food, agronomic, and pharmaceutical industries (Güçlü-Üstündag; Mazza, 2007; Moses; Papadopoulou; Osbourn, 2014). These compounds are widely used as emulsifiers, foaming agents, and detergents, as well as in animal feed additives, plant growth regulators, and soil bioremediation (Güçlü-Üstündag; Mazza, 2007). The incorporation of saponins into food products is highly relevant and promising, as they can be used as natural surfactants and preservatives, helping to control microbial and fungal spoilage (Cheok; Salman; Sulaiman, 2014; Jiménez *et al.*, 2021). Artificial surfactants are found in the composition of food products, cleaning agents, cosmetics, textiles, and many other goods. However, these compounds are typically synthesized from petrochemical sources, are non-biodegradable, and may be toxic. Natural surfactants derived from plant saponins represent an ecological and sustainable alternative to their synthetic counterparts (Rai *et al.*, 2021).

Numerous pharmacological activities have also been reported for saponins, including cytotoxic effects on cancer cells, immune response enhancement, antioxidant, antiviral, antibiotic, antidiabetic, and cholesterol-lowering properties. These effects are primarily attributed to their ability to permeabilize and disrupt cell membranes (Augustin *et al.*, 2011; Cheok; Salman; Sulaiman, 2014; Moses; Papadopoulou; Osbourn, 2014). Saponins from

Quillaja saponaria and *Quillaja brasiliensis*, for example, have been used as adjuvants in veterinary vaccines, and the use of saponin fractions from these species in human vaccines is currently under development (Guerra; Sepúlveda, 2020). Another important source of natural saponins is the genus Yucca, with 108 steroidal saponin molecules reported in eight species, particularly *Yucca schidigera*. The high diversity of saponins in these species has led to their broad use in the cosmetic, pharmaceutical, beverage, and animal feed industries, among others, with enormous potential still to be explored (Jiménez *et al.*, 2021).

FIGURE 2 – STRUCTURES OF TRITERPENE (A) AND STEROIDAL (B) AGLYCONES. GENERAL STRUCTURE OF A BIDESMOSIDIC TRITERPENE SAPONIN (C)



SOURCE: Adapted from Güçlü-Üstündag; Mazza (2007) and Coelho et al. (2010).

2.3 SAPONIN ANALYSIS AND NON-DESTRUCTIVE TECHNIQUES FOR LEAF CHARACTERIZATION

The quantification of saponins can be performed using spectrophotometric or chromatographic methods, depending on the objective: spectrophotometric methods provide a

total saponin content, while chromatographic methods quantify specific molecules. The most commonly used spectrophotometric method for determining total triterpene saponins is the reaction of oxidized saponins with vanillin (vanillin-sulfuric acid assay). Although this method is simple, rapid, and low-cost, it requires optimization of parameters such as standards, wavelengths, heating time, and temperature prior to its use (Cheok; Salman; Sulaiman, 2014). However, for the characterization and quantification of individual saponin molecules, more robust methods are necessary, such as High-Performance Liquid Chromatography (HPLC) or Ultra-Performance Liquid Chromatography (UPLC) coupled with mass spectrometry (LC/MS) (Savarino; Demeyer; Colson, 2021).

In complex materials such as plants, spectroscopic techniques especially infrared (IR) spectroscopy has been used for direct analyses, complementing chromatographic methods (Haas; Mizaikoff, 2016). Compared to conventional methods, these are fast, non-destructive, and environmentally friendly technologies, as they do not require chemical reactions (Hou *et al.*, 2022). Techniques such as near-infrared (NIR) and mid-infrared (MIR) spectroscopy can be used to detect differences resulting from a variety of complex effects on plant materials and are collectively referred to as biospectroscopy (Skolik; McAinsh; Martin, 2018). These techniques are capable of analyzing materials with little to no sample preparation, including solid samples, and are sensitive enough to detect variations in both organic and inorganic molecules (Berhow *et al.*, 2020; Haas; Mizaikoff, 2016).

The combination of spectra obtained through NIR and MIR can yield better results than individual models, including in the estimation of saponin content (Liu *et al.*, 2019). In the case of yerba mate, there are records of NIR techniques being used to predict the rooting of minicuttings (Sá, 2018), moisture (Mazur; Oliveira; *et al.*, 2014), and total methylxanthine content (Mazur; Peralta-Zamora; *et al.*, 2014). In the present study, spectra obtained using both spectroscopic techniques were combined into a global model to detect metabolic changes in yerba mate leaves caused by genetic and environmental factors.

2.4 SAPONINS IN YERBA MATE

Saponins are predominantly found in angiosperms, and triterpene saponins occur and accumulate mainly in dicotyledonous plants, including the Aquifoliaceae family, to which

yerba mate belongs (Moses; Papadopoulou; Osbourn, 2014). In yerba mate leaves, triterpene saponins derived from ursolic and oleanolic acids predominate (Mateos *et al.*, 2017; Souza *et al.*, 2011), while in the fruits, pomolic and rotundic acids are most commonly found (Peixoto *et al.*, 2011; Taketa; Breitmaier; Schenkel, 2004). Aglycones of rotundic/asiatic acid (isomers) and pomolic acid have also been detected in the leaves of *I. paraguariensis* (Negrin *et al.*, 2019).

The species exhibits a wide variety of saponins, with 16 distinct molecules reported in a single study (Mateos *et al.*, 2017). However, new molecules continue to be discovered, such as "mateosides 1-3" (Nagatomo *et al.*, 2022). The concentration of saponins in yerba mate leaves ranges from 0.003% to 4.14% (Andrade *et al.*, 2012; Borré *et al.*, 2010; Coelho *et al.*, 2010; Gnoatto; Schenkel; Bassani, 2005; Mateos *et al.*, 2017; Nakamura *et al.*, 2009). This wide variation is attributed to differences in analytical methods, genetic materials, cultivation conditions, and post-harvest processing.

In any case, yerba mate is considered an important dietary source of saponins in South America (Mateos *et al.*, 2017). The species' saponins are regarded as potent agents in terms of anti-inflammatory activity and colon cancer prevention, showing the ability to inhibit cell proliferation *in vitro* (Puangpraphant; Berhow; de Mejia, 2011; Puangpraphant; De Mejia, 2009). Also, an *in vivo* study with rats treated with methanolic yerba mate extract, matesaponins 1 and 2 reduced triglyceride accumulation, suggesting that these molecules have anti-obesity activity (Nagatomo *et al.*, 2022). Another study concluded that different *Ilex* species, including *I. paraguariensis*, exhibit antibacterial effects, which appear to be related to saponins, as well as to a possible synergistic effect between saponins and polyphenols (Paluch *et al.*, 2021).

In addition, oleanolic and ursolic acids the main aglycones found in yerba mate leaves (Mateos *et al.*, 2017; Souza *et al.*, 2011) are isomers with very similar and potent biological activities. These compounds are present in many plant species and, consequently, are part of the human diet. Both are known for their low toxicity and a wide range of biological functions, including antioxidant, antimicrobial, anti-inflammatory, anticancer, antihyperlipidemic, analgesic, hepatoprotective, gastroprotective, anti-ulcer, cardioprotective, immunomodulatory effects, as well as the potential to stimulate muscle growth, improve epidermal barrier function, and induce collagen production, contributing to anti-aging effects (López-Hortas *et al.*, 2018). Although the long-term consumption of yerba mate beverages is considered safe (Andrade *et al.*, 2012; Morais *et al.*, 2009), further *in vitro* and *in vivo* studies on the saponins found in yerba

mate are recommended, particularly with a view to developing alternative forms of consumption, such as capsules.

2.5 GENOTYPE AND ENVIRONMENT INFLUENCE SAPONIN CONTENT

The quantity and composition of saponins can be affected by numerous abiotic factors, such as geographic location, seasonality, light, temperature, water availability, and soil fertility. Biotic factors such as insect and herbivore attacks, plant competition, symbiotic interactions, or pathogens can also lead to variations in saponin synthesis (Moses; Papadopoulou; Osbourn, 2014; Szakiel; Pączkowski; Henry, 2011). In addition, intrinsic factors such as the plant organ or tissue analyzed, as well as the growth and reproductive stages, may also influence saponin levels (Szakiel; Pączkowski; Henry, 2011).

Despite the numerous factors that can influence saponin accumulation, studies have reported strong genetic control over saponin content in certain species. Accessions of *Centella asiatica* (L.) Urb., propagated vegetatively and cultivated under the same environmental conditions, were analyzed for the content of two triterpene saponins: madecassoside and asiaticoside. Significant variation was observed, with madecassoside levels ranging from nearly undetectable to 5.67%; these differences were attributed to genotype, expression levels, or other stages of the biosynthetic pathway of these molecules (Thomas *et al.*, 2010). A study on cassava (*Manihot esculenta* Crantz) also indicated high heritability (0.61) for saponin content (Daemo *et al.*, 2022). In soybean (*Glycine max* (L.) Merrill), high heritability and high genetic gain for saponin content were reported, suggesting low environmental influence and ease of selecting genotypes with desirable saponin levels (Kaur; Gill; Sharma, 2017). In the same species, saponin synthesis was shown to be determined by different allele frequencies, further confirming the relevance of genetic factors in the metabolism of these compounds (Panneerselvam *et al.*, 2013). In yerba mate, high heritability (0.75) at the progeny level was found for saponins derived from ursolic acid (Nakamura *et al.*, 2009); however, the authors recommended further studies involving different yerba mate genotypes to better understand the genetic control of this trait and to enable the selection of genotypes for various industrial purposes.

The recommendation of yerba mate genetic materials based on saponin content depends on the intended industrial use of the raw material. The Brazilian market tends to prefer yerba mate with a mild flavor (Rakocevic *et al.*, 2008), and considering that saponins contribute significantly to the bitterness of the leaves (Pires *et al.*, 1997; Scherer *et al.*, 2007), materials with higher saponin content may not be well accepted in the domestic market. On the other hand, higher concentrations of specific saponins may be desirable for industrial extraction (Cheok; Salman; Sulaiman, 2014), given the great potential for pharmaceutical applications of these compounds. Therefore, depending on the intended product and target market, materials with either low or high saponin levels may be recommended, highlighting the importance of clonal silviculture in yerba mate.

2.6 GENETIC PARAMETER ANALYSIS

To recommend genetic materials, it is necessary to predict the genotypic value of each material, that is, to determine the genotypic mean, free from random environmental effects and from experimental effects. Based on genotypic evaluations, it becomes possible to predict the behavior of materials in future populations, such as in commercial plantations, where environmental conditions will differ from those of the base population, even if located in the same region. The phenotypic mean, which is measured in the field or laboratory, represents a combination of genetic and environmental factors and should not be used as a true value in genetic improvement evaluation (Resende, 2007).

Genotypic evaluations involve the estimation of variance components (genetic parameters such as heritability) and the prediction of genotypic values. The most widely used methodology for obtaining these values is Restricted Maximum Likelihood/Best Linear Unbiased Prediction (REML/BLUP), which relies on linear models and Generalized Linear Models (GLM). BLUP is the optimal selection procedure for determining additive, dominance, and genotypic genetic effects, depending on the type of breeding trial. However, the application of BLUP requires the estimation of variance components, for which REML is an ideal method, as it is well-suited for data from field experiments whether balanced or unbalanced. Moreover, this method allows for the decomposition of phenotypic variation into genetic, environmental, and genotype × environment interaction components. As such, the REML/BLUP procedure has

broad applicability and naturally handles data imbalance, resulting in more accurate estimates and predictions of genetic parameters and values (Resende, 2016).

Due to the complexity of statistical analyses and the variety of models involved, software such as the Brazilian SELEGEN REML/BLUP (Resende, 2016), developed by Embrapa, is highly valuable for the estimation of genetic parameters and the prediction of genotypic values. This open-source software can be applied to analyses of both annual and perennial crops, with different reproductive systems, including clonal propagation, and across a wide range of genetic improvement trials using various experimental designs, whether with balanced or unbalanced data (Resende, 2016).

3 CHAPTER 1: OPTIMIZING THE VANILLIN-ACID SULFURIC METHOD TO TOTAL SAPONIN CONTENT IN LEAVES OF YERBA MATE CLONES

DOI: https://doi.org/10.1002/cbdv.202301883
Natália Saudade de Aguiar, Fabricio Augusto Hansel, Cristiane Aparecida Fioravante Reis, Marcelo Lazzarotto and Ivar Wendling
Chemistry & Biodiversity – JCR (Clarivate) 2023: 0.44

ABSTRACT – Yerba mate (*Ilex paraguariensis*) is a forest species consumed in the form of non-alcoholic beverages in South America, with applications in foods, cosmetics, and pharmaceutical industries. The species leaves are globally recognized for their important bioactive compounds, including, saponins. We adjusted the vanillin-acid sulfuric method for determining spectrophotometrically the total saponin in yerba mate leaves. Seeking to maximize the extraction of saponins from leaves, a Doehlert design combined with Response Surface Methodology (RSM) was used, considering ethanol:water ratios and ultrasound times. In addition, the same methodology was used for the analysis of times and temperatures in the vanillin-sulfuric acid reaction heating. The contents of total saponin in mature leaves were compared in four yerba mate clones. The extraction was maximized using 40% ethanol:60% water and 60 minutes of ultrasound assisted extraction (UAE) without heating. For the reaction conditions, 70ºC for 10 minutes heating is recommended, and UV/Vis reading from 460 to 680 nm. Using the optimized methodology, total saponin contents ranged from 28.43 to 53.09 mg g^{-1} in the four yerba mate clones. The significant difference in saponin contents between clones indicate great genetic diversity and potential for clones' selection and extraction of these compounds from yerba mate leaves.

Keywords: terpenoids; triterpene saponins; *Ilex paraguariensis*; ultrasound assisted extraction (UAE); spectrophotometric method.

3.1 INTRODUCTION

Saponins are amphiphilic molecules composed of a triterpene or steroidal aglycone attached to one or more sugar chains (Moses; Papadopoulou; Osbourn, 2014; Osbourn; Goss; Field, 2011). There are more than 200 saponins molecules reported in the plant kingdom (Hussain *et al.*, 2019); these molecules come from the secondary metabolism, and they are important in plant defense against pathogens and herbivores, and act at various stages of development, including seed germination, growth, and fruiting (Moses; Papadopoulou; Osbourn, 2014). The great variability, with different aglycones, chemical bonds, number and type of monosaccharides, confers distinct physical, chemical, and biological properties to saponins, with applications in the cleaning, cosmetic, food, agronomic, and pharmaceutical industries, for example (Güçlü-Üstündag; Mazza, 2007; Moses; Papadopoulou; Osbourn, 2014; Timilsena; Phosanam; Stockmann, 2023). As biological activities are reported cytotoxic effects in cancer cells, potential to increase immune responses, hemolytic, molluscicidal, antioxidant, anti-inflammatory, antimicrobial, antiviral, antiparasitic, antidiabetic, cholesterol lowering, among others; effects attributed especially to its ability to cross the lipoprotein bilayer of cell membranes (Cheok; Salman; Sulaiman, 2014; Moses; Papadopoulou; Osbourn, 2014). Due to the numerous benefits of saponins and their presence in many foods, including grains, tea and other medicinal species, interest in these molecules has been increasing (Sharma *et al.*, 2023).

Yerba mate is widely consumed in the form of non-alcoholic beverages in countries of South America. Nowadays, in addition to drinks, this species is already part of pharmacological, cosmetics, and food formulations due to its bioactive compounds (Bracesco, 2019; Gerber *et al.*, 2023). In countries that traditionally consume yerba mate infusions, this is considered an important source of saponins in the diet of people (Mateos *et al.*, 2017). There are reports of *in vitro* and *in vivo* biological activities of saponins from yerba mate (Nagatomo *et al.*, 2022; Puangpraphant; Berhow; de Mejia, 2011; Puangpraphant; De Mejia, 2009). In yerba mate leaves, triterpene saponins derived from ursolic and oleanolic acid predominate (Mateos *et al.*, 2017; Souza *et al.*, 2011), in addition to rotundic and asiatic acid (isomers), and pomolic acid aglycones (Negrin *et al.*, 2019). The species leaves present a wide variety of saponin molecules, known as matesaponins, being detected 16 of them (Mateos *et al.*, 2017). However, new molecules continue to be discovered, such as “mateosides 1-3” (Nagatomo *et al.*, 2022).

The concentrations of saponins in yerba mate leaves found in the literature range from 3.65 to 41.4 mg g^{-1} (Andrade *et al.*, 2012; Borré *et al.*, 2010; Coelho *et al.*, 2010; Gnoatto; Schenkel; Bassani, 2005; Mateos *et al.*, 2017; Morais *et al.*, 2009). This wide saponins content range can be attributed to different methods of extraction and analysis, genetic materials, cultivation systems, pedoclimatic conditions, and material processing. Above all, studies indicate a strong genetic influence on the levels of saponins in plant species (Daemo *et al.*, 2022; Kaur; Gill; Sharma, 2017; Panneerselvam *et al.*, 2013; Thomas *et al.*, 2010). There is evidence of large intraspecific genetic diversity for the content of several secondary metabolites in yerba mate, including saponins (Coelho *et al.*, 2010; Nakamura *et al.*, 2009). Therefore, validation of the saponin quantification procedure in yerba mate is necessary to verify the genetic influence and for application in a large number of samples, as breeding programs.

The quantification of saponins can be performed either by chromatographic (HPLC, UPLC, LC-MS, TLC, among others) and spectrophotometric methods (Cheok; Salman; Sulaiman, 2014). Using the spectrophotometric method is possible to quantify the total saponin content (TSC) relative to the triterpene nucleus released after a hydrolysis reaction, with subsequent reaction with vanillin (vanillin-sulfuric acid method) (Hiai; Oura; Nakajima, 1976). This spectrophotometric method is simple, quick, and low-cost when compared to the chromatographic ones; however, is necessary to establish some parameters for the vanillin-sulfuric acid reaction (Cheok; Salman; Sulaiman, 2014), in addition to optimizing the saponins extraction process for yerba mate leaves.

Ultrasound, sonication, or ultrasound-assisted extraction (UAE) is a low-cost, low energy consumption, fast, and environmentally clean method that can increase the efficiency of extracting bioactive compounds; due to cavitation and implosion the cell wall disrupts, the solvent enters, and the intracellular plant material is incorporated in the solvent. Thus, ultrasound alone or combined with other technologies can lead to extremely efficient extraction of bioactive compounds, such as saponins (Ranjha *et al.*, 2021). The solvent also plays a major role in extraction efficiency. Usually, methanol or ethanol is used to extract saponins by different methods, such as maceration, reflux or UAE (Cheok; Salman; Sulaiman, 2014). In this study, we opted for a hydroalcoholic solvent, which has less environmental impact and can be used in human food. The adoption of this method is justified by its reduced toxicity to laboratory personnel. This approach minimizes the exposure of operators to hazardous substances, thereby enhancing laboratory safety. Thus, using ethanol:water and UAE, considered a “green

extraction" (Cheok; Salman; Sulaiman, 2014), a less environmentally damaging, economical, fast and suitable extraction method is proposed for yerba mate leaves.

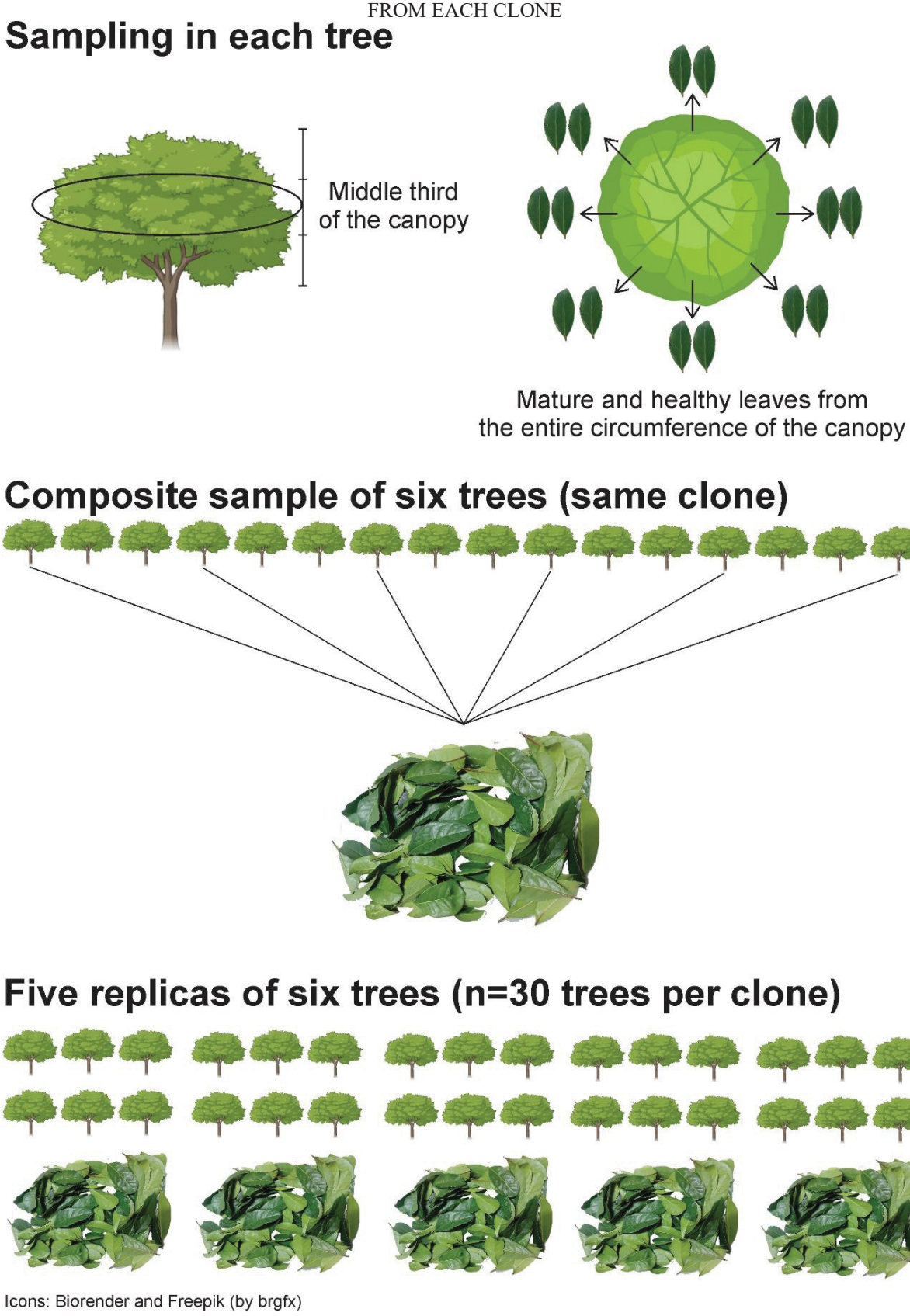
In this study, we addressed important aspects, such as sampling to collect the leaves, the extraction of saponins from them, in addition to the vanillin-sulfuric acid reaction optimization to quantify the total saponin in yerba mate. We believe these protocols can be applied to leaves of other species and adapted to other plants' organs. The optimized spectrophotometric method simple, fast, and low-cost and provided a reliable determination of the total saponin in yerba mate leaves, to be used as a standard method to select genetic materials in breeding programs. The optimized methodology was applied to four yerba mate clones to test its efficiency.

3.2 MATERIAL AND METHODS

3.2.1 Plant Material

The collection of plant material was conducted in a clonal test (monospecific plantation) located in Espumoso, Rio Grande do Sul, Brazil (28º51'59" S, 52º51'39" W). In July 2022, we collected leaves from four yerba mate clones registered by the Brazilian National Register of Cultivars (RNC/MAPA): BRS 408 (34467), BRS 409 (34470), BRS BLD Aupaba (36545), and BRS BLD Yari (36544) (Wendling *et al.*, 2017a, b). Leaves in a mature stage (Aguiar *et al.*, 2022) were collected according to the sampling scheme: six plants were systematically selected to compose each biological repetition, with five repetitions per clone, totaling 30 plants (n = 30) sampled (Figure 3). The leaves were stored in a thermal box and dried in a microwave oven for 5 minutes Tomasi *et al.* (2021), within 24 hours after collection. Subsequently, the samples were ground in a willye type knife mill, with granulometry less than 1 mm (18 mesh sieve) and stored in a -20ºC freezer until the analyses.

FIGURE 3 – SAMPLING SCHEME FOR COLLECTING REPRESENTATIVE YERBA MATE LEAVES FROM EACH CLONE



SOURCE: The author (2025).

3.2.2 Extraction

The extract preparation conditions, ethanol:water proportions and time in ultrasonic bath, were analyzed by Doehlert design (Table 1) and subsequently modeled by response surface (Response Surface Methodology - RSM). The Doehlert matrix was used for two variables (Cerqueira *et al.*, 2021), being determined the range of real values to be analyzed for each variable (times from 25 to 125 minutes and ethanol concentration from 0% to 100%) and the number of central points (3, totaling 9 treatments). The actual values (X) are related to the encoded values (Z) by the following formula:

$$X = Z \times \frac{\Delta X}{\Delta Z} + X_0$$

Where X_0 is the value of the central point (average of the highest and the lowest actual value) and ΔX e ΔZ are the differences between the highest and the lowest values of the actual and coded numbers, respectively (Butiuk *et al.*, 2021).

TABLE 1 – DOEHLERT MATRIX FOR TIME IN ULTRASONIC BATH (1) AND ETHANOL CONCENTRATION (2) IN THE PREPARATION OF PLANT EXTRACT; MATRIX WITH 3 CENTRAL POINTS

	Codified values		Real values	
Treatment	Z1	Z2	X1	X2
1	0.000	0.000	75	50
2	0.000	0.000	75	50
3	0.000	0.000	75	50
4	1.000	0.000	125	50
5	0.500	0.866	100	100
6	-0.500	-0.866	50	0
7	0.500	-0.866	100	0
8	-0.500	0.866	50	100
9	-1.000	0.000	25	50

SOURCE: The author (2025).

To optimize the extraction process, we used samples of the BRS 408 clone. The extracts were prepared using 100 mg of dried yerba mate leaves, adding 10 mL of solvent at room temperature, resulting in a solvent-solute ratio of 100 (mL g^{-1}) (solvent-solute ratio based on saponins extraction studies) (Pham *et al.*, 2018; Sarvin *et al.*, 2018). The extraction was

carried out by homogenizing the sample in a vortex for 10 seconds and in an ultrasonic bath without heating for 25, 50, 75, 100, and 125 minutes, followed by centrifugation for 5 minutes at 4500 rpm. The supernatant was separated and stored in a refrigerator (5ºC) until analyses, within a maximum period of 24 hours.

3.2.3 Reaction

The analysis of total saponin was performed by the vanillin-sulfuric acid method, adapted from Hiai; Oura; Nakajima (1976). It consisted of adding a 100 µL aliquot of extract (or diluted standard) to a test tube, followed by the addition of 250 µL of vanillin solution (8% – 8 g of vanillin in 100 mL of ethanol HPLC grade, prepared on the day of analysis) and 2.5 mL of 60% sulfuric acid. We observed that vortexing before heating results in greater reaction homogeneity.

The intervals for reaction heating were defined: temperatures from 40 to 80ºC and times from 10 to 40 minutes. The design was also performed using Doehlert matrix, factors temperature and time (Table 2), using the BRS 408 sample. We observed that only with ethanol (white, zero point), the reaction already showed changes in absorbance according to heating conditions. Thus, for each reaction time and temperature, a zero point was also analyzed, and its value was subtracted from the sample absorbance, obtaining Δ absorbance.

3.2.4 Reagents, Equipment, and Statistical Analyses

The standards of oleanolic acid (>97%, Sigma Aldrich®, USA) and ursolic acid (>98.5%, Sigma Aldrich®, USA) were diluted in ethanol (1 mg mL^{-1}). We used ethanol HPLC grade (>99.9%) and type 1 ultrapure water. On each day of analyses, a calibration curve was prepared using four aliquots, from 0 to 90 µL of the oleanolic acid standard, completed to 100 µL with ethanol, also totaling 2.85 mL of total reaction volume. The saponin contents found in the yerba mate samples were expressed in mg g^{-1} (dry weight) oleanolic acid equivalent. All readings were performed in quartz cuvettes in UV/Vis spectrophotometer model 1800 (Shimadzu®, Japan) and the "base line" with water. The scan was performed in a wavelength range of 460 to 680 nm. For kinetics, the readings were programmed every 10 minutes, from 0

to 180 minutes. The intraday and interday analyses were performed three times a day for three days (n = 9), using two standard concentrations, 15 and 75 µg of ursolic acid:oleanolic acid (proportion of 1.25:1.00, according to the yerba mate estimate presented by Souza *et al.* (2011). All analyses were performed in triplicate.

Response Surface Methodology (RSM), Kruskal-Wallis non-parametric test, and Wilcoxon test were performed in the R software, as well as the generated graphics.

TABLE 2 – DOEHLERT MATRIX FOR TEMPERATURE (1) AND REACTION HEATING TIME (2); MATRIX WITH 3 CENTER POINTS

	Codified values		Real values	
Treatment	Z1	Z2	X1	X2
1	0.000	0.000	60	25
2	0.000	0.000	60	25
3	0.000	0.000	60	25
4	1.000	0.000	80	25
5	0.500	0.866	70	40
6	-0.500	-0.866	50	10
7	0.500	-0.866	70	10
8	-0.500	0.866	50	40
9	-1.000	0.000	40	25

SOURCE: The author (2025).

3.3 RESULTS AND DISCUSSION

3.3.1 Extraction Conditions

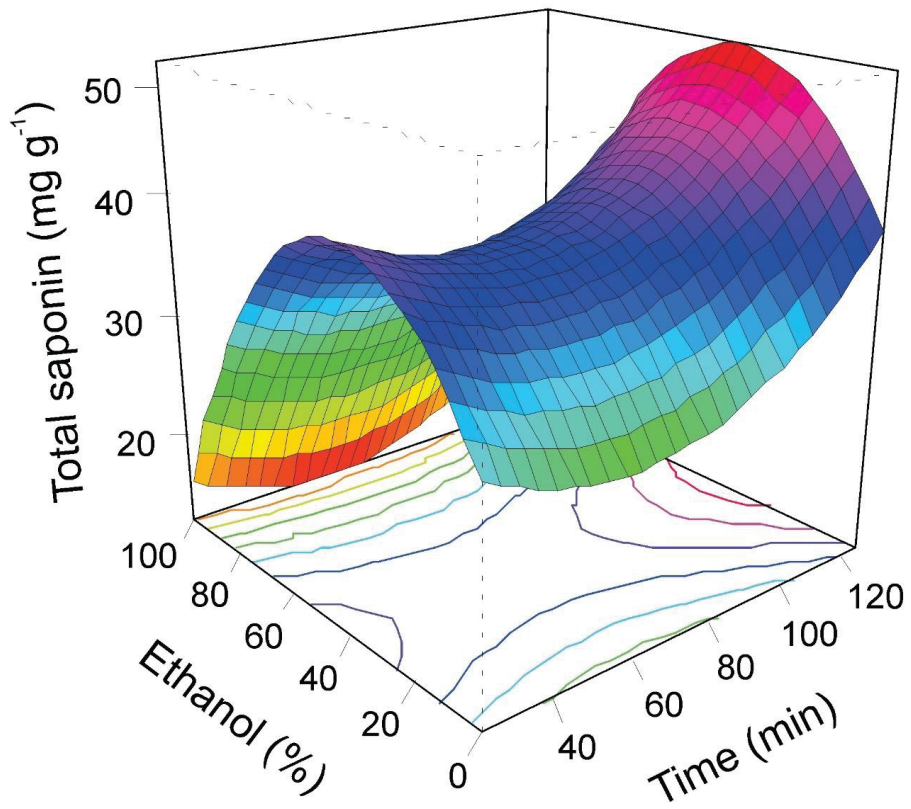
To maximize saponins extraction in yerba mate leaves, different proportions of ethanol:water and ultrasonic bath (without heating) times were tested, using a fixed solvent-solute ratio of 100 and final volume of 10 mL. Both factors were significant but with no interaction between them in the response surface model (Table 3). The stationary points in the original units were 59.55 minutes (minimum point) and 38.3% ethanol (maximum point). To facilitate laboratory analyses, we used 40:60 (EtOH: H_2O) as solvent. Reaction absorbance increased until 120 minutes in an ultrasonic bath (Figure 4), thus the stationary point suggested by the model was used (~60 minutes).

TABLE 3 – RESPONSE SURFACE METHODOLOGY (RSM) FOR TOTAL SAPONINS (MG G^{-1}) ANALYZING TIME AND ETHANOL PROPORTION (ETHANOL:WATER) WITH YERBA MATE SAMPLE FROM CLONE BRS 408. DOEHLER DESIGN WITH 3 CENTRAL POINTS. DATA TRANSFORMED BY SQUARE ROOT

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	4.358	0.039	112.412	$<2.2^{-16}$ *
X1	0.121	0.019	6.222	3.586^{-6} *
X2	-0.546	0.034	-16.256	2.261^{-13} *
X1:X2	0.019	0.034	0.575	0.571
$X1^2$	0.094	0.015	6.131	4.401^{-6} *
$X2^2$	-1.194	0.046	-25.973	$<2.2^{-16}$ *

*Significant at 95% of probability. Adjusted R-squared: 0.977; p-value: $< 2.2^{-16}$; lack of fit (Pr>F): 0.1587
SOURCE: The author (2025).

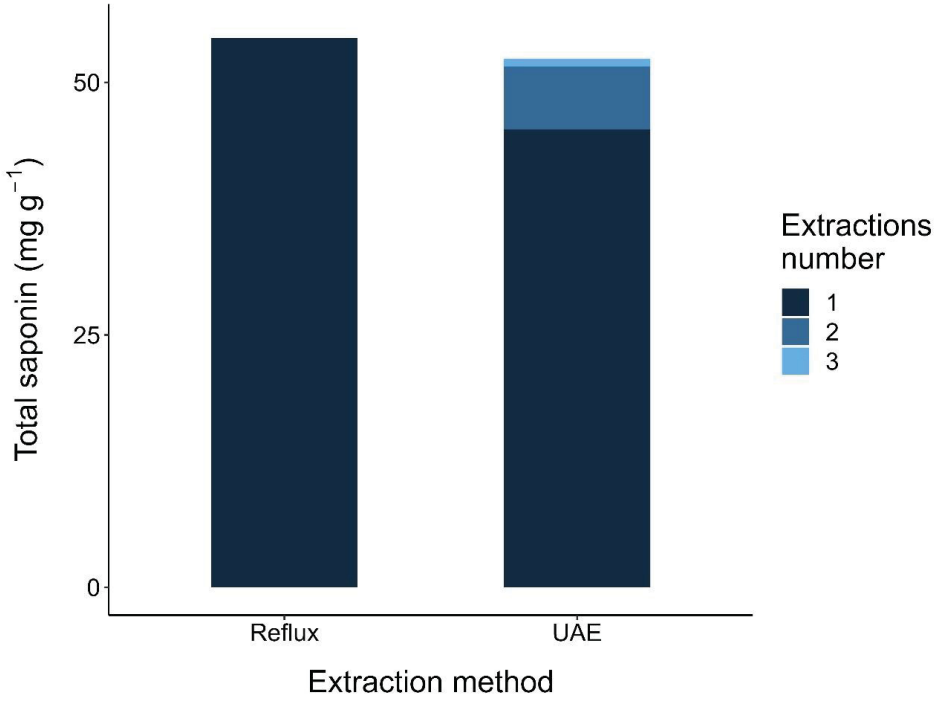
FIGURE 4 – RSM (RESPONSE SURFACE METHODOLOGY) FOR TOTAL SAPONIN CONTENT (MG G^{-1} DRY WEIGHT – OLEANOLIC ACID EQUIVALENT) AS A FUNCTION OF TIME IN ULTRASONIC BATH AND ETHANOL (ETHANOL:WATER) CONCENTRATION FOR THE BRS 408 YERBA MATE CLONE



SOURCE: The author (2025).

Consecutive extractions were made to check the efficiency of saponins extraction. The proportion of saponin yield in the three consecutive extractions was 87:12:1, considering that, the major amount of saponin was extracted in the first step, 60 minutes (Figure 5). In addition, we performed a reflux extraction – exhaustive extraction with heating for 3 hours, solvent-solute ratio of 100, and final volume of 250 mL (40% ethanol:60% water) – and saponin content with reflux was only 16.6% higher than UAE 60 minutes, indicating the efficiency of UAE for saponin extraction. Reflux and consecutive extractions demonstrated that a higher solvent:solute ratio probably would not significantly increase saponin extraction. Therefore, seeking a relative quantification fast and low cost (less solvent and energy), UAE without heating for 60 minutes proved to be an efficient extraction method, making it readily adaptable for industrial-scale applications.

FIGURE 5 – TOTAL SAPONIN CONTENT (MG G^{-1} DRY WEIGHT – OLEANOLIC ACID EQUIVALENT) IN BRS 408 YERBA MATE CLONE AS A FUNCTION OF EXTRACTION METHOD (REFLUX VERSUS UAE, ULTRASONIC ASSISTED EXTRACTION, WITH 3 CONSECUTIVE EXTRACTIONS)



SOURCE: The author (2025).

In the present study, we sought to reduce the environmental impact, by choosing the solvent, lower volume, and clean extraction method. Triterpene acids are insoluble in water and non-polar solvents but are soluble in alcohols (Mateos *et al.*, 2017). Ethanol:cold water is the

most widely used solvent for saponins extraction (Oleszek; Bialy, 2006), not being necessary evaporation of solvents before the vanillin-sulfuric acid reaction, unlike other solvents that interfere with the reaction, such as butanol and methanol (Le *et al.*, 2018). In addition, mixing these solvents can increase extraction efficiency due to the different polarities of saponin molecules (Bitencourt *et al.*, 2014), which was confirmed by our results.

3.3.2 Reaction Conditions

Given the great variation in the spectrophotometric method for total saponin quantification found in the literature, Cheok; Salman; Sulaiman (2014) highlights the importance of defining parameters before the application. According to the same authors, most studies rarely specify the standard used, reaction time, or provide a justification for the chosen wavelength. There are also differences in reagents employed, with variations such as sulfuric or perchloric acid, as well as differences in concentrations and volumes of vanillin and acids. In this study, we used 100 µL of plant extract (or standard), 250 µL of 8% vanillin (w/v in 99.9% ethanol), and 2.5 mL of 60% sulfuric acid.

Regarding times and temperatures of heating in a water bath, both factors were significant in the model, as well as the interaction between them (Table 4). The absorbance delta (subtracting zero point) increased with temperature, not stabilizing until 80ºC (Figure 6). However, at 80ºC for 25 minutes a hypsochromic shift occurred (blue shift, shorter wavelength), distancing from the normally used maximum absorption point (λmax), 544 nm, and a greater data dispersion was noted than 70º for 10 minutes (Figure 7A).

To better evaluate the effect of heating and time factors, we analyzed the BRS 408 sample in a factorial design (2 x 3) with two maximum temperatures (70 and 80ºC) and three reaction times (10, 25 and 40 minutes) (Table 5). With the increase in reaction time, data dispersion also increased: with 80ºC for 40 minutes, for example, the CV was 17.48%, demonstrating the large variation between replicates. The hypsochromic shift was observed at the temperature of 80ºC. Although absorbance was maximized with higher temperatures and heating times, there was a change in its λ_{max} and an increase in data dispersion. This effect is similar to that obtained when the assay was performed with the standard of oleanolic acid (Figure 7B), in which, with 70ºC for 10 minutes (dotted line), the scan became more uniform,

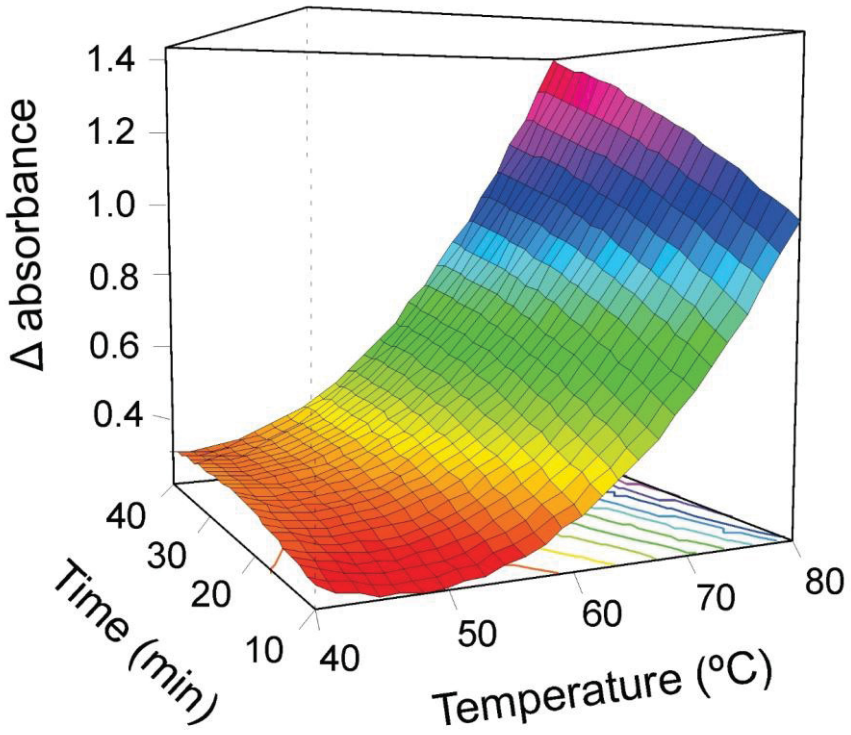
similar to a normal distribution, with λ_{max} 541 nm. Under the same conditions, the standard of ursolic acid presented a λ_{max} of 539 nm; both very close to 538 nm, as described by Hiai *et al.* (1976). Thus, the reaction condition that optimizes the process and is closest to the reference reading value of 544 nm is 70ºC for 10 minutes.

TABLE 4 – RESPONSE SURFACE METHODOLOGY (RSM) FOR DELTA ABSORBANCE (DISCOUNTING ZERO POINT) ANALYZING HEATING TIME AND TEMPERATURE OF THE VANILLIN-SULFURIC ACID REACTION WITH YERBA MATE SAMPLE FROM CLONE BRS 408. DOEHLER DESIGN WITH 3 CENTRAL POINTS

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	0.489	0.024	20.098	3.402^{-15} *
X1	0.231	0.012	18.952	1.097^{-14} *
X2	0.121	0.021	5.765	1.010^{-5} *
X1:X2	0.050	0.021	2.357	0.028 *
$X1^2$	0.074	0.010	7.711	1.477^{-7} *
$X2^2$	-0.045	0.029	-1.564	0.133

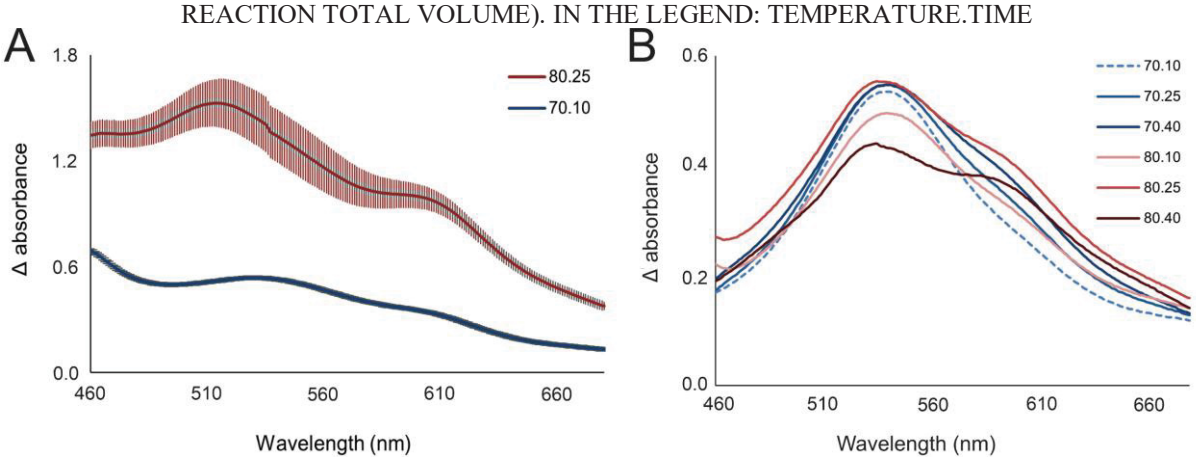
*Significant at 95% of probability. Adjusted R-squared: 0.947; p-value: 1.275^{-13}; lack of fit (Pr>F): 8.539^{-5}
SOURCE: The author (2025).

FIGURE 6 – RSM (RESPONSE SURFACE METHODOLOGY) FOR ABSORBANCE DELTA (SUBTRACTING ZERO POINT), AS A FUNCTION OF TIME AND THE TEMPERATURE OF THE VANILLIN-SULFURIC ACID REACTION HEATING FOR THE BRS 408 YERBA MATE CLONE



SOURCE: The author (2025).

FIGURE 7 – UV/VIS SPECTRUM OF THE VANILLIN-SULFURIC ACID REACTION AT DIFFERENT TEMPERATURES AND HEATING TIMES FOR A) THE BRS 408 YERBA MATE CLONE, MEANS AND STANDARD DEVIATION BARS, AND B) WITH OLEANOLIC ACID ONLY (50 µG IN 2.85 ML – REACTION TOTAL VOLUME). IN THE LEGEND: TEMPERATURE.TIME



SOURCE: The author (2025).

TABLE 5 – FACTORIAL EXPERIMENT OF REACTION HEATING TEMPERATURES AND TIMES, WITH BRS 408 YERBA MATE CLONE SAMPLE

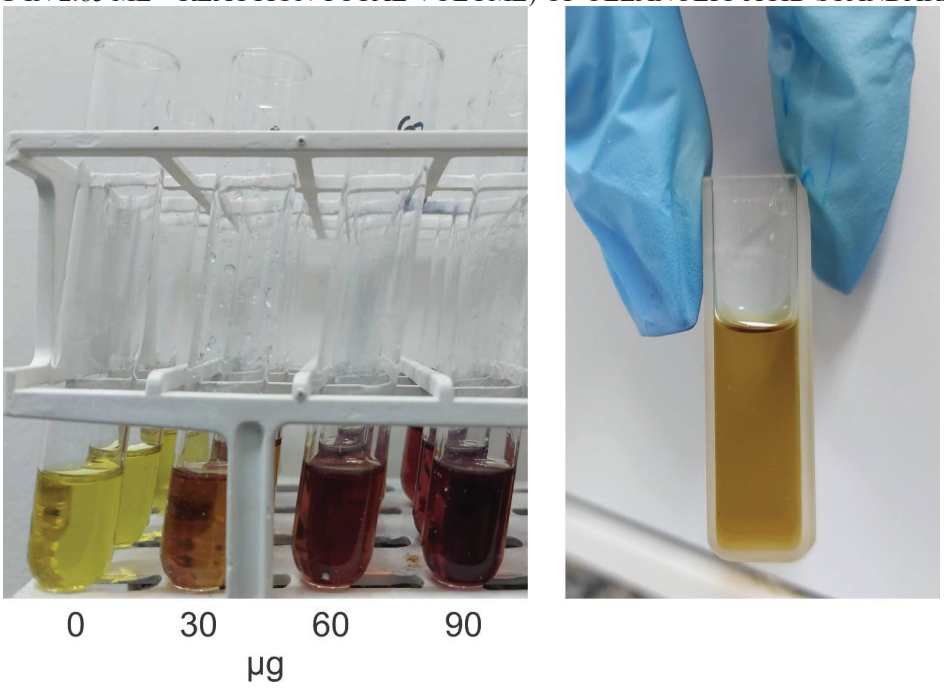
			Δ absorbance (em 544 nm)	
Temperature	Time	λ_{max}	Average	CV (%)
70	10	525	0.241	3.33
70	25	521	0.353	7.07
70	40	520	0.525	13.79
80	10	511	0.296	2.04
80	25	515	0.493	9.32
80	40	515	0.672	17.48

λ_{max}: point of maximum absorption.
SOURCE: The author (2025).

In the original methodology, it is recommended to use 60°C for 10 minutes (Hiai; Oura; Nakajima, 1976), but in literature, there are various adaptations, such as 60°C for 30 minutes (Liu *et al.*, 2012) 70°C for 20 minutes (Li *et al.*, 2010) to ensure full development of color. We observed that at 80°C, absorbance was higher; however, there were changes in the UV/Vis spectrum of the reaction and greater variation between replicates. It is suggested that in this study, the vanillin may have undergone aggregation or self-condensation at high temperature, as indicated in previous research (Broadhurst; Jones, 1978; Hiai; Oura; Nakajima, 1976). In fact, no methodology was found using 80°C, possibly due to factors mentioned earlier. Therefore, it is recommended to use a maximum heating temperature of 70°C. At 70°C for 10 minutes, the UV/Vis spectrum of the reaction became similar to the expected spectrum for

oleanolic acid-vanillin adduct, with more accurate readings. An important aspect of the vanillin-sulfuric acid method is the color of the reaction, which is described as red purple (Hiai; Oura; Nakajima, 1976). While other studies do not mention the visual aspect, our reaction exhibited a color ranging from yellow (zero point - ethanol only) to caramel (Figure 8). The wavelength of 544 nm was adopted for absorbance readings, as this is the most used in the literature (Cheok; Salman; Sulaiman, 2014). However, it is important to read the UV/Vis spectrum from 460 to 680 nm to check for any significant changes in the shape of the reaction.

FIGURE 8 – VANILLIN-SULFURIC ACID REACTION AT INCREASING CONCENTRATIONS (0 TO 90 µG IN 2.85 ML – REACTION TOTAL VOLUME) OF OLEANOLIC ACID STANDARD



SOURCE: The author (2025).

Using these same heating conditions, kinetics was performed, with the reaction considered stable, varying by 7.1% in 180 minutes after the ice bath. We also conducted intraday and interday analyses, using 15 and 75 µg of ursolic acid:oleanolic acid (1.25:1.00). The relative standard deviations (RSDs) for intraday varied from 4.5% to 7.5% at the lower standard concentration and from 5.2% to 8.5% at the higher concentration. Whereas for interday, the RSDs were 15.23% and 12.28% at for 15 and 75 µg, respectively (n = 9). Relative errors (REs) in the estimation of saponin content were 33.3% (15 µg) and 7.6% (75 µg) in interday analyses. Calculated lower limits of detection (LLOD) and quantification (LLOQ)

were 8.01 and 11.94 µg, respectively ($p>0.95$). Determination coefficients (r^2) of the calibration curves were all >0.98.

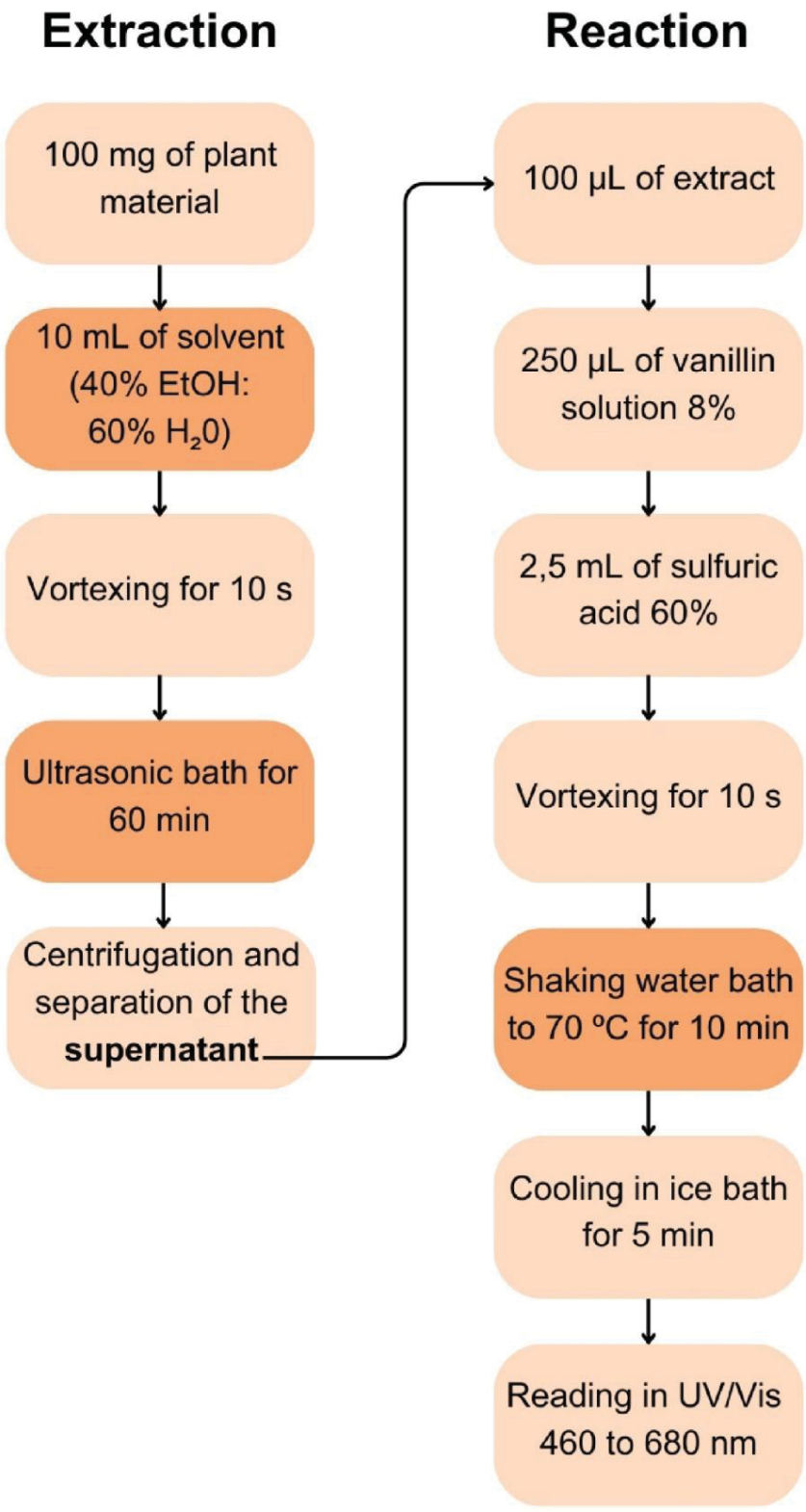
The interday and intraday were within the acceptable limit for precision (± 15%), but this method was not accurate for the lowest concentration, indicating that real samples with a high proportion of ursolic acid may lead to an overestimation of total saponin values. So, the two most abundant aglycones found in yerba mate, oleanolic and ursolic acid have very close molecular masses (Mateos *et al.*, 2017) and react with vanillin at very close wavelengths; however, they have different absorbances in reaction. The present method is relative and the total saponin content is dependent on the used standard, in this case, oleanolic acid equivalent. Although it is not a very accurate method, for genotype selection in a breeding program this method can be very useful, in comparing the total saponin content of different genetic materials (e.g. high x low contents).

After all these optimization and verification steps, a flowchart for optimized spectrophotometric analysis of total saponin in mature yerba mate leaves was developed (Figure 9). The importance of vortex agitation prior to reaction heating is highlighted, as it homogenizes the solution, reducing variations in triplicate.

3.3.3 Total Saponin in Leaves of Yerba Mate Clones

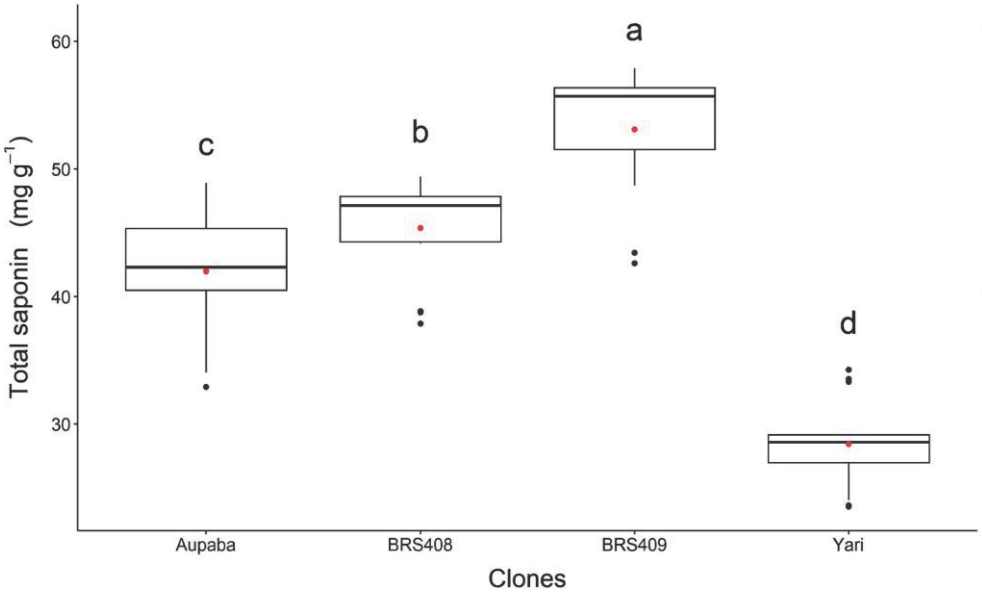
There was low variation between total saponin content within each clone (five replicas and analyses in triplicate), with RSDs ranging from 8.38% to 11.96%. All samples also presented values higher than LLOQ, even considering standard deviation. We found a significant difference among clones ($p = 8.533^{-10}$, Kruskal-Wallis non-parametric test), and all clones differed from each other, with a 95% confidence interval (Wilcoxon test). The BRS 409 clone had the highest saponin content, 53.09 (±9.17) mg g^{-1}, while Yari exhibited only 28.43 (±11.79) mg g^{-1} (Figure 10); on average, BRS 409 showed a total saponin content 86.74% higher than Yari.

FIGURE 9 – FLOWCHART OF THE SPECTROPHOTOMETRIC ANALYSIS FOR DETERMINATION OF TOTAL SAPONIN CONTENT IN YERBA MATE, WITH DARK ORANGE HIGHLIGHTS INDICATING THE PROCESS STEPS THAT WERE OPTIMIZED IN THE STUDY



SOURCE: The author (2025).

FIGURE 10 – BOXPLOT OF TOTAL SAPONIN CONTENT (MG G^{-1} DRY WEIGHT – OLEANOLIC ACID EQUIVALENT) IN MATURE LEAVES OF THE FOUR YERBA MATE CLONES. THE DIFFERENT LETTERS INDICATE STATISTICAL DIFFERENCES BETWEEN CLONES USING THE WILCOXON TEST. RED DOTS INDICATE THE MEAN. RSD: RELATIVE STANDARD DEVIATION



Clones	RSD (%)
Aupaba	11.96
BRS408	8.38
BRS409	9.17
Yari	11.79

SOURCE: The author (2025).

In literature, very variable levels of saponins for yerba mate leaves are found, from 3.65 to 41.4 mg g^{-1} (Andrade *et al.*, 2012; Borré *et al.*, 2010; Coelho *et al.*, 2010; Gnoatto; Schenkel; Bassani, 2005; Mateos *et al.*, 2017; Morais *et al.*, 2009). In our study, the BRS 409 clone reaching 53.09 (±9.17) mg g^{-1} of total saponin. It should be considered that, in most studies, commercial yerba mate was used (composed of leaves and up to 30% of thin branches), and we analyzed only leaves. Borré *et al.* (2010) analyzed leaves and branches of yerba mate separately and found 41.4 (±0.8) mg g^{-1} in leaves, while in branches, the content was 9.4 (±0.8) mg g^{-1}. Therefore, we can assume that the high levels observed in our study were due to the plant part used, only mature leaves; there is still the possibility that young leaves have even higher levels, since these compounds aim, among others, to protect against attack by herbivores. However, studies with other saponin quantification methods must be used to confirm the reported levels. There is great intraspecific chemical diversity in yerba mate for other secondary metabolites, such as caffeoylquinic acids and methylxanthines (Duarte *et al.*, 2023; Vieira *et al.*, 2021). Corroborating with previous studies (Coelho *et al.*, 2010; Nakamura *et al.*, 2009), our results indicate that this diversity is also present in the saponin levels in the species leaves, and this raw material could be promising in terms of saponin yield.

3.4 CONCLUSION

Through the Doehlert design combined with RSM, we optimize ultrasound extraction (ethanol:water ratios and extraction times) and the reaction (temperatures and times of heating) using the vanillin-sulfuric acid method for total saponin of yerba mate leaves. Here we report that high time and/or temperature in vanillin-sulfuric acid reaction can cause larger analytical errors and hypsochromic shift; UV/Vis scanning is important to check for any significant changes in the reaction absorbance. The present spectrophotometric method can separate different yerba mate clones and may be useful for genetic materials selection in species breeding programs. Leaves of yerba mate have the potential for saponins extraction, and the species shows a large saponin variation between genetic materials.

4 CHAPTER 2: TOTAL SAPONINS IN YERBA MATE LEAVES: INFLUENCE OF CLONES, SITES AND HARVEST SEASONS

DOI: http://dx.doi.org/10.1590/1984-70332025v25n2a04

Natália Saudade de Aguiar, Cristiane Aparecida Fioravante Reis, Marcelo Lazzarotto and Ivar Wendling

Crop Breeding and Applied Biotechnology – JCR (Clarivate) 2023: 1.3

ABSTRACT – Yerba mate (*Ilex paraguariensis*) is an important source of bioactive compounds, including triterpene saponins. This study aimed to analyze the effects of genotypes, cultivation sites, and harvest seasons on the total saponin content in yerba mate leaves, as well as estimate the genetic parameters and predict the genotypic values. We harvested mature leaves from nine clones, cultivated in two clonal tests, and harvested in two seasons (winter and summer), and analyzed them using the vanillin-sulfuric acid spectrophotometric method. The total saponin content trait showed high selection accuracy and significant effects of clones, without interactions. High genetic control was observed, with clone mean broad-sense heritability of 0.93. The genotypic value for total saponin content ranged from 28.13 to 51.54 mg g^{-1} on a dry weight basis. Selecting yerba mate clones with low or high leaf saponin levels may be useful for specific industries, such as pharmaceuticals, cosmetics, or beverages.

Keywords: *Ilex paraguariensis*; triterpene saponins; secondary metabolites; heritability; breeding program.

4.1 INTRODUCTION

Saponins are a highly diverse class of secondary metabolites widely found in the Plant Kingdom and serve a general defense function in plants, in addition to possible effects on the regulation of plant growth and development (Costa *et al.*, 2013; Faizal; Geelen, 2013; Hussain *et al.*, 2019). They are high-molecular-weight molecules composed of aglycones (triterpenic or steroidal nucleus) linked to sugar chains, which gives them surfactant amphiphilic properties

(Hussain *et al.*, 2019). Different saponins are present in many plant species, including grains, legumes, and tea (*Camellia sinensis* (L.) Kuntze) (Sharma *et al.*, 2023). Yerba mate (*Ilex paraguariensis* A.St.-Hil.) is a tree species whose leaves are commonly consumed in the form of non-alcoholic beverages and stands out as an important source of triterpene saponins, especially derivatives of ursolic and oleanolic acids (Mateos *et al.*, 2017; Souza *et al.*, 2011). Ursanes, oleananes, and their derivatives have been extensively studied in recent years for their potent anticancer and anti-tumorigenic potential (Biswas; Dwivedi, 2019). The saponins found in yerba mate leaves have *in vitro* preventive effects on inflammation and colon cancer (Puangpraphant; Berhow; de Mejia, 2011), and *in vivo* assays have shown anti-hyperlipidemic and anti-obesity activities. A study with rats also demonstrated that ingested yerba mate aqueous infusions had a protective renal function (Kuropka *et al.*, 2021). Therefore, yerba mate leaves prove to be an important source of saponins for human health and have potential for industrial extraction, reaching a concentration of 53.09 mg g^{-1} or 5.3% (w/w) of total saponins on a dry weight basis (Aguiar *et al.*, 2024b).

The various useful biological properties of triterpenic saponins make them valuable chemicals for industry (Magedans; Phillips; Fett-Neto, 2021). These molecules have different applications, with notable presence in cosmetics, food, agronomy, cleaning, and pharmaceutical industries (Moses; Papadopoulou; Osbourn, 2014; Timilsena; Phosanam; Stockmann, 2023), highlighting the potent antineoplastic activity of some triterpenoid saponins (Biswas; Dwivedi, 2019). However, saponins are recognized to confer a bitter taste to food and beverages. When ingested orally, they appear to be practically non-toxic to humans (Oakenfull, 1981). Still, excessive consumption of these substances should be avoided, as they are considered anti-nutritional factors and can cause indigestion-related disorders (Samtiya; Aluko; Dhewa, 2020). However, few studies have been carried out and it is a very diverse group of molecules, and consequently, difficult to generalize regarding their effects on the human body.

Considering the technical challenges of chemically synthesizing these complex molecules, natural sources of saponins should be sustainably explored to obtain plant material with higher yields of these compounds (Magedans; Phillips; Fett-Neto, 2021). A few studies indicate that genotype highly influences saponin levels in *Manihot esculenta* (Daemo *et al.*, 2022), *Glycine max* (Kaur; Gill; Sharma, 2017; Panneerselvam *et al.*, 2013), *Centella asiatica* (Thomas *et al.*, 2010), and *I. paraguariensis* (Nakamura *et al.*, 2009). However, numerous external (biotic and abiotic) and internal (such as age and phenological phase) factors affect the

synthesis and accumulation of saponins in plants (Magedans; Phillips; Fett-Neto, 2021; Szakiel; Pączkowski; Henry, 2011). Clarifying the impact of environmental factors on saponin levels is crucial, particularly for plants destined for food or pharmaceutical purposes (Szakiel; Pączkowski; Henry, 2011). This endeavor aids in determining optimal growth conditions and ideal harvest time, thereby ensuring the acquisition of high-value plant material.

To explore the potential of yerba mate leaves as a saponin source and enhance industrial yield, we need to comprehend the factors influencing their content. Thus, this study aimed to analyze the influence of genotypes (clones), sites, and harvest seasons on saponin levels in yerba mate leaves. As the main hypothesis, we expect high heritability for the trait saponin content in yerba mate leaves, indicating the predominance of genetic over environmental factors. To validate this hypothesis, we analyzed the total saponin content in mature leaves of nine yerba mate clones cultivated in two clonal tests situated in geographically distant areas and employing different cultivation systems, encompassing both winter and summer harvests.

4.2 MATERIAL AND METHODS

4.2.1 Experimental Design and Saponins Quantification

We analyzed mature leaves of nine yerba mate clones, cultivated in two sites (Espumoso, Rio Grande do Sul State/RS, and São Mateus do Sul, Paraná State/PR, Brazil) and harvested in two seasons (July/2022 – winter, and January/2023 – summer), totaling a 9 x 2 x 2 factorial experiment. For each site and harvest the experimental design was completely randomized with five replications (composed of six plants) of each clone.

These clones were selected previously for interesting traits, such as high productivity and methylxanthine levels. They are from the Southern region of Brazil, and most of them were selected and vegetatively rescued from a provenances and progenies test conducted by Brazilian Agricultural Research Corporation (Embrapa Forestry) in Ivaí, Paraná State (Table 6). Some of them are included in the National Cultivar Registry at the Ministry of Agriculture and Livestock (MAPA – Brazil). The areas were chosen because they are part of or close to traditional yerba

mate production regions, in addition to the large distance (413 linear km) and different cultivation systems. The sites represent the most common yerba mate cultivation systems in each State (Goulart, 2020), with full sunlight predominating in Rio Grande do Sul and shaded cultivation in Paraná. In Espumoso (RS) site, the clonal test was conducted under full sunlight, while in São Mateus do Sul (PR) site, yerba mate was cultivated in the forest understory composed of large native trees, providing an average shading of 70%, considered as an agroforestry system. Shading was determined using a luxmeter, with measurements of light intensity taken in the forest understory and in adjacent open areas. Details of each site and climatic conditions during the harvesting periods are presented in Appendix B.

TABLE 6 – DESCRIPTION OF *ILEX PARAGUARIENSIS* CLONES ANALYZED IN CLONAL TESTS IN ESPUMOSO, RIO GRANDE DO SUL (RS), AND SÃO MATEUS DO SUL, PARANÁ (PR), BRAZIL

Clone code	Wild material origin	Selection site	NCR*
EC22	Ivaí, PR	Ivaí, PR	-
EC25	Quedas do Iguaçu, PR	Ivaí, PR	-
EC37	Ivaí, PR	Ivaí, PR	-
EC40	Barão de Cotegipe, RS	Ivaí, PR	-
EC43	Quedas do Iguaçu, PR	Ivaí, PR	-
BRS 408	Cascavel, PR	Ponta Grossa, PR	34467
BRS 409	Bocaiúva do Sul, PR	Ponta Grossa, PR	34470
BRS BLD Aupaba	São Mateus do Sul, PR	São Mateus do Sul, PR	36545
BRS BLD Yari	Machadinho, RS	São Mateus do Sul, PR	36544

*NCR: National Cultivar Registry (MAPA – Brazil)
SOURCE: The author (2025).

Mature leaves of yerba mate were collected for analyses according to the sampling scheme described by Aguiar *et al.* (2024b), with five biological replicates of each clone, consisting of six plants, totaling n=30. Within 24 hours after harvest, the leaves were dried in a microwave oven for enzymatic inactivation (Tomasi *et al.*, 2021), and ground in a knife mill, with particle size of less than 1 mm. After processing, the samples were kept in a freezer at -20 ºC. Extraction of saponins was carried out using hydroalcoholic solvent in an ultrasound bath. Total saponin content was determined using the vanillin-sulfuric acid spectrophotometric method adapted for yerba mate leaves (Aguiar *et al.*, 2024b), expressed in mg g^{-1} on a dry weight basis – equivalent to oleanolic acid, with triplicate analyses.

4.2.2 Statistical Analyses

Restricted Maximum Likelihood/Best Linear Unbiased Prediction (REML/BLUP) procedure was performed in the Selegen-REML/BLUP software (Resende, 2016) for saponin content trait in all experiments. In individual analyses (for each site and season), Model 20 from Selegen-REML/BLUP was used: $y = Xr + Zg + e$, where y: data vector, r: vector of fixed effects of repetition summed with the overall mean, g: vector of genotypic random effects, and e: vector of errors or residuals (random). Uppercase letters represent the incidence matrices for the respective effects.

In the joint analysis of sites and seasons, Model 155 was used: $y = Xf + Zg + Qgl + Tgm + Wgml + Sp + e$, where: y: data vector, f: vector of fixed effects of the combinations of repetition-site-season summed with the overall mean, g: vector of genotypic random effects, gl: vector of random effects of the genotypes x sites interaction, gm: vector of random effects of the genotypes x seasons interaction, glm: vector of random effects of the genotypes x sites x seasons interaction, p: vector of random permanent effects of plots within sites, and e: vector of errors or residuals (random). Uppercase letters represent the incidence matrices for the respective effects.

From the analyses, deviance values were obtained, and likelihood ratio tests (LRT) were conducted. Estimates of accuracy and genetic parameters of the clones were obtained at the level of sites and seasons, as well as in the overall mean of environments (all sites and seasons). Genotypic values were predicted via BLUP.

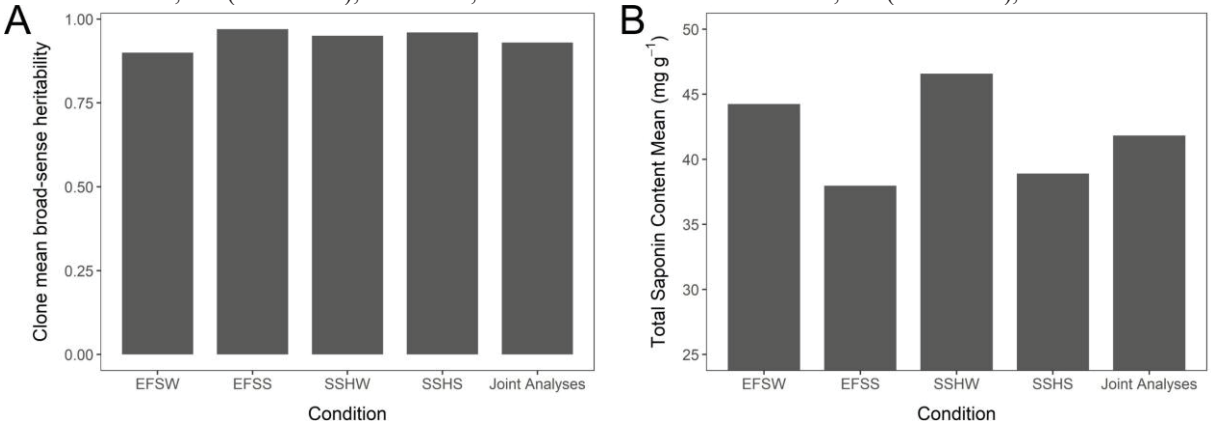
4.3 RESULTS AND DISCUSSION

4.3.1 *Deviance* Analyses

Significant effects of clones were observed in the individual analyses of the experiments conducted in all sites and seasons (Appendix D). The selective accuracies were high, equal to or greater than 0.95 in all analyses. For each experiment, the high proportion of

genotypic variance (V_g) relative to phenotypic variance (V_p) resulted in a clone mean broad-sense heritability greater than 0.90 (Figure 11A). The coefficients of relative variation (CV_{rel}) ranged from 1.39 to 2.68. The mean saponin content was at least 14% higher in winter compared to summer, in both sites, ranging from 37.97 (Espumoso, RS, full sunlight, summer) to 46.58 mg g^{-1} (São Mateus do Sul, PR, shaded, winter) (Figure 11B).

FIGURE 11 – CLONE MEAN BROAD-SENSE HERITABILITY (A) AND TOTAL SAPONIN CONTENT MEAN (B) FOR INDIVIDUAL AND JOINT *DEVIANCE* ANALYSES. EFSW: ESPUMOSO, RS (FULL SUNLIGHT), WINTER; EFSS: ESPUMOSO, RS (FULL SUNLIGHT), SUMMER; SSHW: SÃO MATEUS DO SUL, PR (SHADED), WINTER; SSHS: SÃO MATEUS DO SUL, PR (SHADED), SUMMER



SOURCE: The author (2025).

The higher saponin contents in winter than in summer indicate seasonality in the synthesis of these defense compounds. Seasonality in the production of other secondary metabolites of yerba mate, such as methylxanthines and caffeoylquinic acids, has also been reported (Aguiar *et al.*, 2024a; Duarte *et al.*, 2023). It is known that the synthesis of triterpenic saponins is complex, involving many stages and being highly susceptible to various inducers (Magedans; Phillips; Fett-Neto, 2021). Environmental factors can simultaneously affect these compounds' synthesis (Yang *et al.*, 2018). Possibly, the lower temperatures may have contributed to higher levels of saponins, since these compounds appear to be important in yerba mate cold tolerance (Rakocevic; Janssens; Schere, 2012). Higher levels of saponins were also observed in *Q. brasiliensis* leaves during the winter (Costa *et al.*, 2013). We cannot accurately determine which factors influenced the higher accumulation of saponins in winter, and only studies under controlled environmental conditions can elucidate the interference of these factors in saponin synthesis in yerba mate. In addition to external factors affecting the plant, vegetative rest, flowering, and fruiting can also interfere with the synthesis of these compounds (Szakiel;

Pączkowski; Henry, 2011). The seasons analyzed in this study (summer and winter) are considered periods of species growth pause (Rakocevic; Martim, 2011) and are the usual harvesting times. Furthermore, we only collected mature leaves, minimizing possible interference from their age. The site influenced minimally the saponins content, even with different cultivation systems (full sunlight and shaded), edaphoclimatic differences, and considerable geographical distance (413 linear km).

In the joint analysis of sites and seasons, there was a significant effect of clones (Table 7). The doubles (clones x sites and clones x seasons) and triple (clones x sites x seasons) interactions, as well as the permanent environment effect, were not significant at a 1% error probability. The absence of significant interactions is supported by the high values of genotypic correlations. Additionally, the sum of interaction variances accounted for less than 20% of genotypic variance. As observed in individual deviance analyses, the genotypic variance in the joint analysis accounted for a large part of phenotypic variance, resulting in high clone mean broad-sense heritability, 0.93 (Figure 11A). The accuracy of clone selection was of high magnitude. Repeatability was statistically different from zero according to the confidence interval and was also high.

TABLE 7 – JOINT *DEVIANCE* ANALYSES OF SITES AND SEASONS, FOR TOTAL SAPONIN CONTENT (MG G^{-1} ON DRY BASIS – EQUIVALENT TO OLEANOLIC ACID) OF YERBA MATE CLONES EVALUATED IN WINTER AND SUMMER HARVESTS, IN TWO CLONAL TESTS (ESPUMOSO, RS, AND SÃO MATEUS DO SUL, PR, BRAZIL)

Sources of variation	**Deviance**	**LRT[1]**
Clones^{+}	693.42	10.63*
Clones x sites^{+}	684.89	2.10^{NS}
Clones x seasons^{+}	683.84	1.05^{NS}
Clones x sites x seasons^{+}	687.38	4.59^{NS}
Permanent environment $^{+}$	683.81	1.02^{NS}
Full Model	682.79	-
Individual REML		
Genotypic variance (V_g)	49.72	
Variance of genotypes x seasons interaction	0.66	
Variance of genotypes x sites interaction	3.63	
Variance of genotypes x sites x seasons interaction	5.17	
Variance of permanent plot effects	2.19	
Residual variance	14.62	
Individual phenotypic variance	70.84	
Individual plot broad-sense heritability	0.70 ± 0.18	
Accuracy	0.96	
r^2 of clones x sites interaction effects	0.05	
r^2 of clones x seasons interaction effects	0.09	
r^2 of clones x sites x seasons interaction effects	0.07	
r^2 of permanent plot effects	0.03	
Individual repeatability	0.78 ± 0.19	

Genotypic correlation across sites, valid for any season	0.93
Genotypic correlation across seasons, valid for any site	0.99
Genotypic correlation across sites, in a given season	0.93
Genotypic correlation across seasons, at a given site	0.99
Genotypic correlation across sites, for the mean of all seasons	0.93
Genotypic correlation across seasons, for the mean of all sites	0.99
Genotypic correlation across sites and seasons	0.84

[1] Likelihood ratio test; [+] Deviance of fitted model without mentioned effects; * Significant at 1% error probability, by Chi-square test with 1 degree of freedom; [NS] Not significant at 1% error probability, by Chi-square test with 1 degree of freedom; r^2: coefficient of determination
SOURCE: The author (2025).

4.3.2 Genotypic Values

Since interactions were not significant and genotypic correlations were high, we only present the genotypic values (BLUPs) resulting from joint analysis (Table 8). Clones BRS 409 and BRS 408 stood out with the highest total saponin contents, while Aupaba and Yari showed the lowest contents. To increase saponin levels, the genetic gain would be 16.7% with the selection of the two best clones compared to the mean of the original population composed of nine clones. But if the goal is to obtain lower saponin levels, genetic gain resulting from the selection of two clones would be 22%, that is, a reduction of 22% in saponin content compared to the mean of the original population.

TABLE 8 – GENOTYPIC VALUES (BLUPS) OF CLONES FOR TOTAL SAPONIN CONTENT (MG G^{-1} ON DRY BASIS – EQUIVALENT TO OLEANOLIC ACID) OBTAINED IN THE JOINT ANALYSIS OF SITES AND SEASONS

Order	Genotype	g	u + g	Gain	New mean
1	BRS 409	9.72	51.54	9.72	51.54
2	BRS 408	4.25	46.07	6.99	48.81
3	EC37	3.90	45.72	5.96	47.78
4	EC40	2.69	44.51	5.14	46.96
5	EC43	1.99	43.81	4.51	46.33
6	EC22	-0.10	41.72	3.74	45.56
7	EC25	-4.06	37.76	2.63	44.45
8	Aupaba	-4.70	37.11	1.71	43.53
9	Yari	-13.69	28.13	0.00	41.82

Genotypic effects (g) and predicted values (u + g), free from all interaction with environments.
SOURCE: The author (2025).

The selection accuracies obtained in this study are considered very high, indicating the effectiveness of inferring genotypic values and, consequently, the correct ranking of clones for selection purposes (Resende; Duarte, 2007). Analyses of genetic parameters indicate genetic

variability and high heritability in determining total saponin content in yerba mate leaves, enabling the selection of clones according to industrial objectives. Studies with various annual plant species also indicated strong genetic control over saponin levels (Daemo *et al.*, 2022; Kaur; Gill; Sharma, 2017; Panneerselvam *et al.*, 2013; Thomas *et al.*, 2010). In yerba mate progenies, a high broad-sense heritability of 0.75 for saponins derived from ursolic acid has already been observed using another analysis method (Nakamura *et al.*, 2009). For other secondary metabolites of species, such as methylxanthines (caffeine and theobromine) and phenolic compounds, high heritability values have also been verified in progenies (Cardozo Junior *et al.*, 2010; Friedrich *et al.*, 2017; Nakamura *et al.*, 2009). Therefore, we can infer that genotype strongly influences secondary metabolites in yerba mate leaves. Thus, establishing commercial plantations with selected yerba mate clones is an effective way to obtain standardized raw material with suitable levels of secondary metabolites for industrial purposes.

This is the first study to estimate the genetic parameters of some bioactive compound from yerba mate leaves through clonal tests. The results demonstrate the possibility of increasing or reducing total saponin content through the selection of specific clones. Clones with low saponin content, such as Aupaba and Yari, may be suitable for the market of mild mate beverages. This is because saponins are known to impart a bitter taste to species' leaves (Pires *et al.*, 1997). These yerba mate clones were sensorially analyzed, with Aupaba and Yari being considered mild, while BRS 408 and BRS 409 are moderately mild (Wendling *et al.*, 2017a, b). Thus, there may be a possible negative correlation between saponin levels obtained in our study and the beverages' mildness. However, conducting a sensory panel with all clones is necessary to verify this hypothesis.

The selection of the two clones with the highest saponin content (BRS 408 and BRS 409) would result in a content of 48.81 mg g^{-1}, demonstrating the species' potential as a new, still unexplored source for industrial saponin extraction for pharmaceuticals or cosmetics, for example. All mentioned clones (Aupaba, Yari, BRS 408, and BRS 409) are also highly productive in leaf mass (Wendling *et al.*, 2017a, b). Therefore, whether for low (Aupaba and Yari) or high (BRS 408 and BRS 409) saponin content selection, establishing commercial plantations of these clones would result in large-scale production of raw material with saponin levels suitable for each industrial purpose. Additionally, breeding programs should always work to expand the genetic base of available clones to safeguard against susceptibility to various pests – diseases and insects (Salgotra; Chauhan, 2023).

Comparing the contents of other secondary metabolites is also important for selecting genetic materials that encompass a broader range of bioactive compounds beyond saponins. For most of the clones analyzed in this study, results regarding caffeine, theobromine, and caffeoylquinic acid contents are available (Aguiar *et al.*, 2023, 2024a; Duarte *et al.*, 2020, 2023; Tomasi *et al.*, 2024; Vieira *et al.*, 2021). However, since the leaves were collected from different cultivation systems, seasons, years, and leaf ages, we chose not to present these results in a simplified manner. It is worth noting that only the cultivars BRS 408, BRS 409, Aupaba, and Yari are officially registered and have published data on foliar biomass yield evaluated under field conditions (Wendling *et al.*, 2017a, b). The remaining clones are still under evaluation within the yerba mate breeding program coordinated by Embrapa Florestas.

4.4 CONCLUSIONS

There is a strong genotype effect on saponin content in yerba mate leaves. Clone selection can be for low or high saponin content, with possibilities for different industrial applications. We recommend harvesting yerba mate leaves in winter to maximize saponin yield.

5 CHAPTER 3: MULTIBLOCK NIR AND MIR SPECTRALPRINT THROUGH ACOMDIM TO EVALUATE THE EFFECTS OF GROWING SITE, HARVEST SEASON, AND CLONE ON YERBA MATE LEAVES COMPOSITION

DOI: https://doi.org/10.1016/j.foodchem.2025.143459

Natália Saudade de Aguiar, Gustavo Galo Marcheafave, Elis Daiane Pauli, Manoela Mendes Duarte, Ieda Spacino Scarminio, Roy Edward Bruns, Romà Tauler, Marcelo Lazzarotto and Ivar Wendling

Food Chemistry – JCR (Clarivate) 2023: 8.5

ABSTRACT – The composition of yerba mate implies significant potential in the food, pharmaceutical, and cosmetic industries, which requires standardization of the raw material. This study explores the simultaneous influence of growing sites, harvest seasons, and clones on the spectralprint of leaves through near-infrared (NIR) and mid-infrared (MIR) spectroscopy coupled with ANOVA Common Dimensions (AComDim) multivariate analysis. MIR spectroscopy identifies only the main effects of growing site and harvesting season, and the interaction between these factors. The NIR spectralprint identifies all main effects and interactions. Growing site and harvesting season individually account for approximately 7% of the variance in the chemical composition of yerba mate, with their interaction contributing with 5.7%. Clonal variation significantly affects the spectral profile with approximately 4% variance, which allowed the identification of clones with the highest chemical divergence. The study demonstrates that biospectroscopic and chemometrics can enhance yerba mate quality through clonal selection and optimized agricultural practices.

Keywords: multiblock methods; biospectroscopy; *Ilex paraguariensis*; genetic influence; cultivation systems; environmental factors.

5.1 INTRODUCTION

Yerba mate leaves (*Ilex paraguariensis* A.St.-Hil.), from a tree species of southern South America, are traditionally used in beverages but also have numerous applications in the food, cosmetic and pharmaceutical industries, owing to their bioactive compounds (Alves; Scheer, 2024). Among the more than 200 compounds found in the leaf tissue of this species, secondary metabolites stand out, such as mono- and di-caffeoylquinic acids, caffeine and theobromine, saponins and their aglycones, ursolic and oleanolic acids (Melo, 2018). Due to the high heritability found for its bioactive compounds, such as methylxanthines and saponins (Nakamura *et al.*, 2009), the selection of yerba mate clones has shown to be the most promising alternative to adjust and standardize the compound levels in plant material, depending on the industrial objectives (Aguiar *et al.*, 2024a), thus maximizing yield.

In addition to genetic influence, leaf age, and environmental factors such as seasonality, light, and nutrition, can affect the accumulation of compounds in the species (Aguiar *et al.*, 2024a; Alves; Scheer, 2024; Blum-Silva *et al.*, 2015). Daily fluctuations of some compounds has also already been observed (Melo, 2018). The production of partially stable and standardized yerba mate raw materials is necessary for industrial purposes focusing on specific chemical compounds. Therefore, in this study, we sought to understand the variation in a portion of the metabolome influenced by the clone, cultivation site, and leaf harvest season factors.

To analyze the effect of intrinsic and environmental factors on plants, and enable quality control of this complex raw material, we can use biospectroscopy (Almeida *et al.*, 2022, 2023). Well-established methods such as near-infrared (NIR), and mid-infrared (MIR) offer a broad range of laboratory and industrial applications. They are non-destructive, do not use reagents, are less laborious, and are more environmentally friendly than conventional analytical methods; in addition, they provide fast data acquisition and high spectral resolution (Haas; Mizaikoff, 2016; Skolik; McAinsh; Martin, 2018). Using optical sensors, spectra are generated containing hundreds of variables so unique that they are called spectral biomarkers, making it possible to identify different genetic materials or detect physiological changes through key compounds. Thus, we achieve a greater understanding at a physiological and molecular level of plants' responses to changes in the environment and how these conditions influence valuable

plant substances for human use (Skolik; McAinsh; Martin, 2018). Biospectroscopy systems may become adapted to field applications in various pre- and postharvest scenarios including affecting the species management by determining the ideal harvest time (Beć, Krzysztof B; Grabska; Huck, 2020; Skolik; McAinsh; Martin, 2018). However, difficulty in data analysis due to the large volume of data generated and the complexity of biological systems arises (He *et al.*, 2021).

The simultaneous analysis of various experimental factors and their interactions with different biospectroscopic blocks (e.g., different spectral regions) can be conducted through the application of the ANOVA (analysis of variance) Common Dimensions (AcomDim) method, a multivariate analysis technique introduced by Bouveresse and collaborators in 2011. AComDim successfully replaces separate PCA treatments on different individual instrumental blocks of data, like the ANOVA-PCA method (Harrington *et al.*, 2005), by a single analysis to give an evaluation of significance of the effects (Korifi *et al.*, 2016). This multiblock method has shown its effectiveness in different studies (Korifi *et al.*, 2015, 2016; Rébufa; Dupuy; Bombarda, 2021) and is interesting because it uses all the spectral variance for each level of factors studied.

In this work, AComDim was applied to spectralprints of powdered yerba mate leaves through combined modeling of NIR and MIR spectroscopy to detect significant spectral chemical divergences due to main effects and interactions of cultivation sites (with different cultivation systems), harvest times (winter and summer), and clones.

5.2 MATERIAL AND METHODS

5.2.1 Sample Collection

We analyzed mature leaves (Aguiar *et al.*, 2022) of nine yerba mate clones (EC22, EC25, EC37, EC40, EC43, BRS 408, BRS 409, BRS BLD Aupaba, and BRS BLD Yari), cultivated simultaneously in two geographically distant sites separated by 413 km, with different cultivation systems (Espumoso, Rio Grande do Sul – full sunlight cultivation, and São

Mateus do Sul, Paraná – shaded cultivation in the understory of native forest, both in Brazil), and harvested in two seasons of the year (July/2022 – winter, and January/2023 – summer). More information about the clones, sites, and climatic conditions during the harvest periods can be found in Table 6 and Appendix B.

The leaf collection scheme was represented by Aguiar *et al.* (2024b), with five biological replicates composed of six plants from each clone (n=30). After harvesting, the leaves were stored in a thermal box for transport, and within 24 hours, they were dried in a microwave oven for approximately five minutes (Tomasi *et al.*, 2021). A knife mill was used to grind the leaves to a particle size less than 1 mm. The samples were kept at -20 ºC in a freezer until analysis.

Thus, the multifactorial design incorporated three-fixed factors: factor A has two levels corresponding to the growing sites, factor B has two levels representing the harvest seasons of yerba mate leaves, and factor C is associated with the nine yerba mate clones (Table 9).

TABLE 9 – IDENTIFICATION OF THE COLLECTION OF YERBA MATE TREATMENTS FROM MULTIFACTORIAL EXPERIMENT

Sample identification	**Agronomic experimental factors**		
	Factor A (Growing site)	**Factor B (Harvest seasons)**	**Factor C (Clones)**
1	Espumoso	Winter	EC25
2	Espumoso	Winter	EC37
3	Espumoso	Winter	EC43
4	Espumoso	Winter	EC40
5	Espumoso	Winter	EC22
6	Espumoso	Winter	BRS 408
7	Espumoso	Winter	BRS 409
8	Espumoso	Winter	BRS BLD Aupaba
9	Espumoso	Winter	BRS BLD Yari
10	Espumoso	Summer	EC25
11	Espumoso	Summer	EC37
12	Espumoso	Summer	EC43
13	Espumoso	Summer	EC40
14	Espumoso	Summer	EC22
15	Espumoso	Summer	BRS 408
16	Espumoso	Summer	BRS 409
17	Espumoso	Summer	BRS BLD Aupaba
18	Espumoso	Summer	BRS BLD Yari
19	São Mateus do Sul	Winter	EC25
20	São Mateus do Sul	Winter	EC37
21	São Mateus do Sul	Winter	EC43
22	São Mateus do Sul	Winter	EC40

23	São Mateus do Sul	Winter	EC22
24	São Mateus do Sul	Winter	BRS 408
25	São Mateus do Sul	Winter	BRS 409
26	São Mateus do Sul	Winter	BRS BLD Aupaba
27	São Mateus do Sul	Winter	BRS BLD Yari
28	São Mateus do Sul	Summer	EC25
29	São Mateus do Sul	Summer	EC37
30	São Mateus do Sul	Summer	EC43
31	São Mateus do Sul	Summer	EC40
32	São Mateus do Sul	Summer	EC22
33	São Mateus do Sul	Summer	BRS 408
34	São Mateus do Sul	Summer	BRS 409
35	São Mateus do Sul	Summer	BRS BLD Aupaba
36	São Mateus do Sul	Summer	BRS BLD Yari

SOURCE: The author (2025).

5.2.2 MIR and NIR Spectralprints

For the plant material to reach equilibrium humidity with the medium to carry out the analyses, the powdered yerba mate samples were kept for up to 24 hours at room temperature (± 22 ºC) protected from light. MIR absorbance spectra readings were performed on the Nicolet iS50 FTIR Spectrometer equipment, OMNIC software (Thermo Fisher®), from 4,000 to 10,000 nm with intervals of approximately 3.85 nm. NIR spectralprints were obtained on the NIR-S-G1 portable equipment (InnoSpectra®), with readings from 900 to 1660 nm with intervals of approximately 5 nm. The same glass beaker was used as a sample holder for all MIR and NIR readings; part of the sample was added to the beaker, positioned over the equipment sensor, and covered with aluminum foil to prevent the entry of external light. After reading, the sample was returned to its container and homogenized for a new reading; five consecutive readings were performed for each of the five replicates. Subsequently, we calculated the average of the replicate readings at each wavelength.

5.2.3 Statistical Analysis

The AComDim method was employed in this study to determine whether spectral variations, resulting from changes in factor levels such as growing sites, harvest seasons and clones as well as their interactions, surpass the residual variability. AComDim is an enhanced version of multiblock chemometric analysis that replaces the multiple PCA steps in the

ANOVA-PCA method with Common Components and Specific Weights Analysis (CCSWA), also known as Common Dimensions analysis or ComDim. Hence, the method is termed AComDim (ANOVA Common Dimension) (Bouveresse *et al.*, 2011). Subsequently, a comprehensive multiblock analysis of all matrices is conducted to extract Common Components (CCs), and each block is assigned a specific weight termed salience, linked to each CC (Korifi *et al.*, 2016). In this context, saliences indicate which factor or interaction effect contributes most to each CC (Bouveresse *et al.*, 2011). Since residuals are inherent in all matrices, blocks with substantial contributions to CC1 and CC2 (characterized by high saliences due to two instrumental techniques evaluated in this research) predominantly consist of noise, whereas blocks with lower contributions encompass sources of variability beyond noise (Bouveresse *et al.*, 2011).

In AComDim, the F-test is utilized to assess the significance of main effects and interactions, considering the relationship between saliencies in the blocks. Therefore, we calculate the ratio between the saliency of the residual block in CC1 and the saliency of the specific block in CC1 for each dataset. Subsequently, a statistical comparison is made between these two saliencies using an F-test (Equation 1), with an applied significance level set equal to 0.05. Blocks whose salience reveals a statistically significant difference in relation to the residual salience of the block are identified as associated with significant factor (or interaction) effects. However, sometimes the residuals block is the main contributor to several CCs, not just CC1, reflecting several orthogonal sources of residual variability. In this case, one divides the sum of the residual saliencies on all those CCs by the sum of the saliencies of each block on the same CCs (Figueiredo *et al.*, 2022).

$$F_i = \frac{\lambda_{res}}{\lambda_i} \qquad \text{Eq. (1)}$$

where λ_{res} is the salience of the residual block on CC1, and λ_i is the salience of the i^{th} block on CC1.

In this investigation, the NIR and MIR spectralprint matrices (blocks) of yerba mate powder leaves were analyzed simultaneously, resulting in the creation of a unified global model. The data were organized into two blocks: yerba mate leaves spectra obtained by MIR in the range between 10,000 and 4,000 nm (1557 variables), and NIR in the range between 1,660 and 900 nm (143 variables). MIR and NIR data contained 180 spectra each, 5 for each of the 36 agronomic combinations (Table 9).

Before AComdim, the preprocessing applied to NIR and MIR spectra were assessed individually. For this, we evaluated the normalization (standard normal variate – SNV, multiplicative scatter correction - MSC, and unnormalized data), moving average (Savitzky-Golay filters with widths of 3, 7, and 11), and derivation (1^{st}, 2^{nd}, and without derivation). All spectra were mean-centered. The selection of the preprocessing for AComDim analysis stemmed from an in-depth study of outcomes from ASCA, encompassing 27 combinations. This investigation considered the assessment of primary effects of factors, binary and ternary interactions, and residual variance. Each treatment can be examined in Appendices E and F.

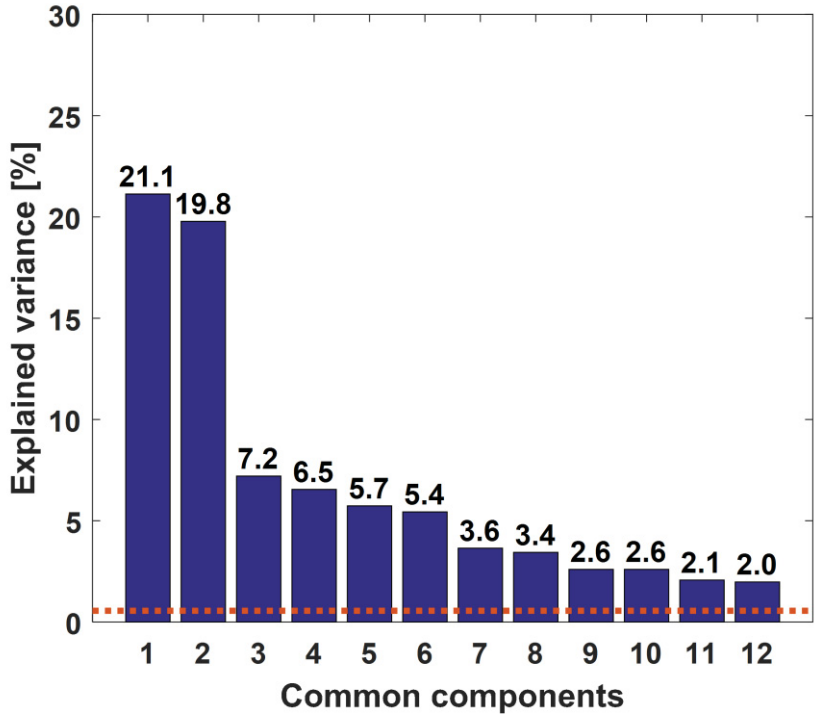
5.3 RESULTS AND DISCUSSION

The impact of pre-processing NIR and MIR spectral data was investigated through a 3^3 design evaluation, guided by the ASCA results that delineated the main effects of factors, as well as binary and ternary interactions, along with the residual variance across the 27 adopted models (Appendices E and F). Prior to engaging in multiblock modeling, the analysis of pre-processing effects was conducted individually. The experimental design, rooted in the primary pre-processing methods utilized in chemometric models, facilitates a systematic evaluation of the individual and concurrent effects of normalization, moving average, and derivation methods applied in this study. Examining the spectralprint in the MIR region (Appendix E), the experiments without derivation (factor 3, level 0) exhibit the lowest residual values.

All experiments at level 0 for factor 3 demonstrate statistical significance for the main effects, encompassing growing sites, harvest seasons, and clones. Using the first and second derivatives accentuates the residual value through noisy broadening of the MIR spectrum. Therefore, experiment 26 (MSC, moving average of 11, and mean center) was selected for MIR preprocessing in the multiblock analysis. Regarding the spectralprint in the NIR region, all models showcase low residual variance. Considering the significance of the main effects, experiment 18 (Second derivative with 2nd-order 11-point Savitzky-Golay and mean center) was chosen for NIR preprocessing in the multiblock analysis (Appendix F). Compared to the MIR spectrum, the NIR spectrum is smoother, which avoids broadening the residual value through derivation.

AComDim was simultaneously applied to the NIR and MIR spectralprint blocks of yerba mate leaves, generating a single global model. The explained variance of each common component, as determined through ANOVA decomposition, reveals that the first component accounts for 21.1% and tends to stabilize, maintaining values lower than 3% from the 9th component onwards, with a variance cutoff of up to 0.56% (indicated by the dotted line) (Figure 12). While all 12 extracted common components demonstrate significant explained variances (all above the cutoff), salience values serve as an appropriate metric for interpreting the sources of variation of different factors within each common component and establishing their relative importance.

FIGURE 12 – PERCENTAGE OF EXPLAINED VARIANCE FOR THE FIRST 12 COMMON COMPONENTS IN A MULTIFACTORIAL DESIGN OF YERBA MATE LEAVES USING MIR AND NIR MULTIBLOCK APPROACH AND CUTOFF AT 0.56% (----)

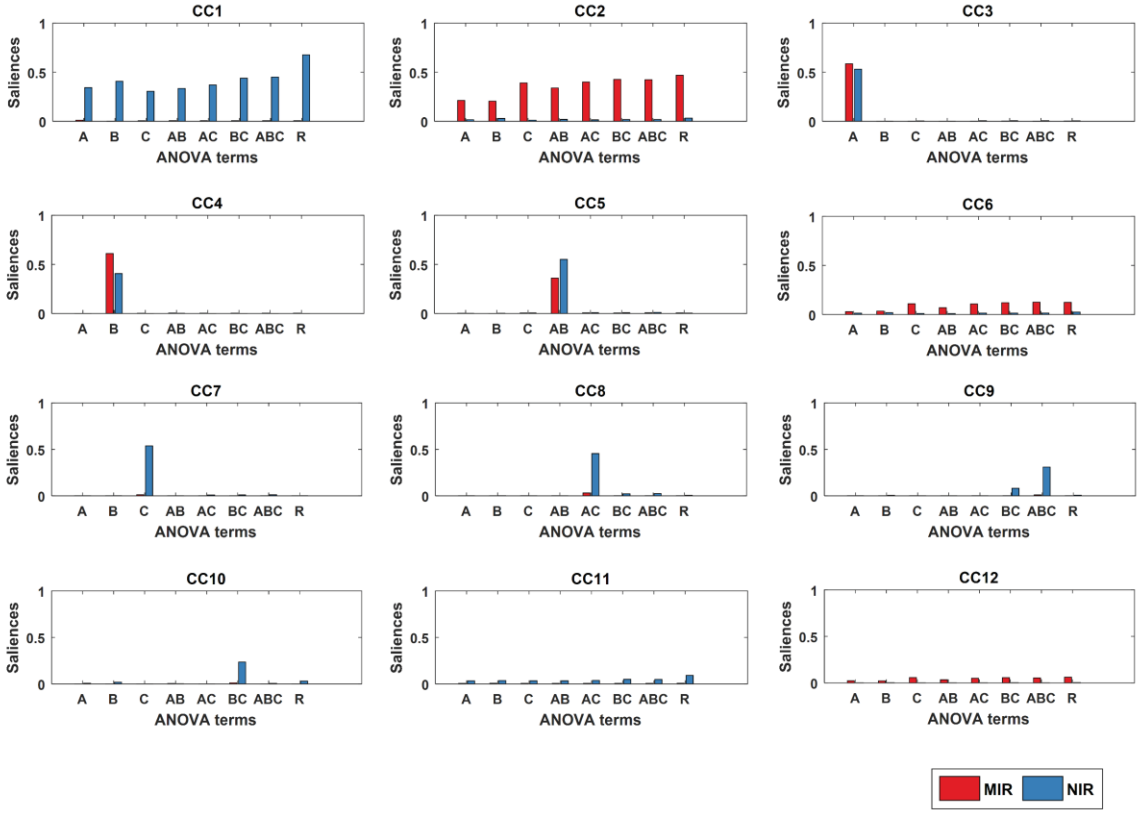


SOURCE: The author (2025).

CC1 and CC2 delineate the common variation attributed to residuals, with CC1 characterizing the residual variation in the NIR block and CC2 expressing the residual variation in the MIR block (Figures 13 and 14). In addition, CC11 represents another orthogonal source of common variation due to residuals in NIR and CC6 and CC12 for MIR spectralprints. Consequently, CCs spanning from 3 to 10, except 6, modulate the contributions of unique

effects within each block. Through the MIR and NIR spectralprint, it becomes possible to identify three main effects: growing sites (effect A), harvest seasons (effect B), and clones (effect C). Furthermore, three binary interactions (AB, AC, and BC) and a ternary interaction (ABC) can be discerned for each spectral block.

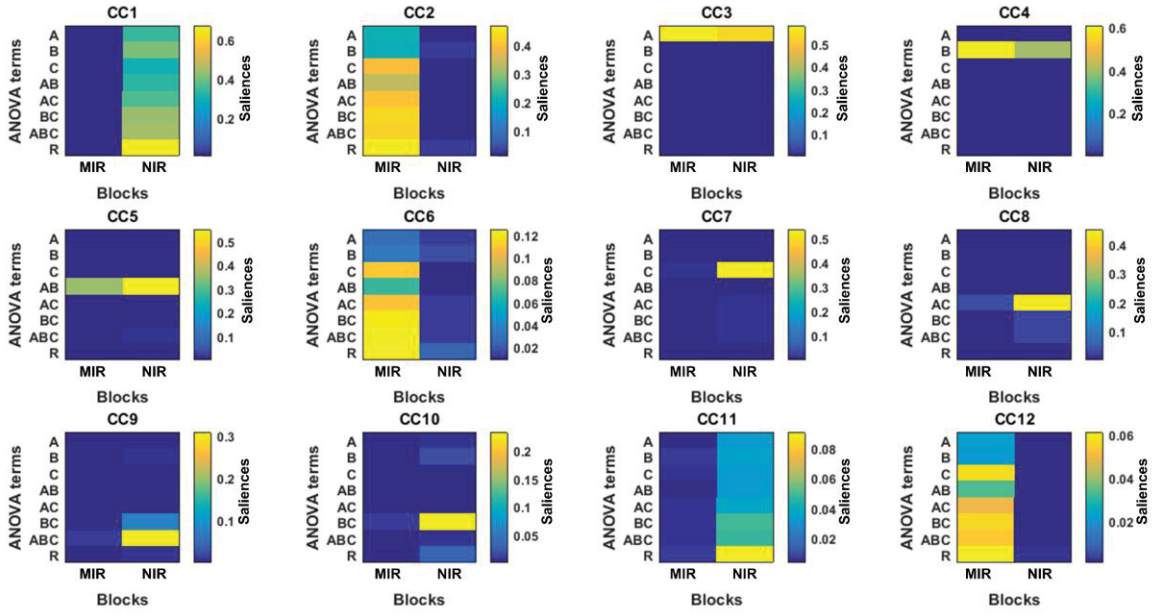
FIGURE 13 – BAR PLOT ILLUSTRATING THE SALIENCES OF EACH ANALYSIS OF VARIANCE (ANOVA) TERM FOR THE FIRST 12 COMMON COMPONENTS (CCS) IN A MULTIFACTORIAL DESIGN OF YERBA MATE LEAVES. FACTORS INCLUDE DIFFERENT GROWING SITES (FACTOR A), HARVEST SEASONS (FACTOR B), AND CLONES (FACTOR C). THE MULTIBLOCK SPECTRALPRINT INCORPORATES BOTH NEAR-INFRARED (NIR) AND MID-INFRARED (MIR) SPECTROSCOPY. ANOVA TERMS ENCOMPASS PURE EFFECTS (A, B, AND C), BINARY EFFECTS (AB, AC, AND BC), TERNARY EFFECTS (ABC), AND RESIDUAL SALIENCE (R)



SOURCE: The author (2025).

The block containing the MIR spectralprint can effectively capture factor A in CC3, factor B in CC4, interaction AB in block 5 (Figures 13 and 14). Nevertheless, the NIR block elucidates factors A, B, AB, C, AC, ABC, and BC in CCs 3, 4, 5, 7, 8, 9, and 10, respectively. Notably, effects A, B, and AB are simultaneously described by both the MIR and NIR blocks in CCs 3, 4, and 5. Through the simultaneous decomposition of the NIR and MIR blocks, AComDim was able to identify all main effects and interactions associated with the cultivation of yerba mate using the spectralprint of leaves.

FIGURE 14 – HEATMAPS DEPICTING THE SALIENCE OF EACH ANALYSIS OF VARIANCE (ANOVA) TERMS FOR THE FIRST 12 COMMON COMPONENTS (CCS) IN A MULTIFACTORIAL DESIGN OF YERBA MATE LEAVES. FACTORS INCLUDE DIFFERENT GROWING SITES (FACTOR A), HARVEST SEASONS (FACTOR B), AND CLONES (FACTOR C). THE MULTIBLOCK SPECTRALPRINT INCORPORATES BOTH NEAR-INFRARED (NIR) AND MID-INFRARED (MIR) SPECTROSCOPY. ANOVA TERMS ENCOMPASS PURE EFFECTS (A, B, AND C), BINARY EFFECTS (AB, AC, AND BC), TERNARY EFFECTS (ABC), AND RESIDUAL SALIENCE (R)



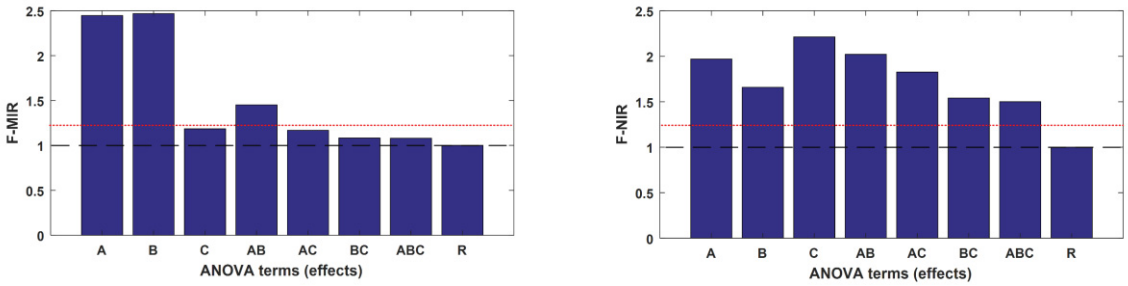
SOURCE: The author (2025).

By considering the salience between the residual blocks (CC1 (NIR); CC2 and CC6 (MIR)), it becomes possible to calculate the significance of the main effects and interactions using the F-test. For the MIR block, the effect A, B and AB demonstrates significance at the 95% confidence level – CC3, CC4, and CC5 contain information specifically related to these effects (Figure 15). Conversely, for the NIR block, all main effects and interactions exhibit significance – CCs 3, 4, 5, 7, 8, 9, and 10 containing significant information about the different effects in the NIR block. Therefore, it becomes feasible to evaluate the dispersion of scores for each significant effect within their respective common components.

Figure 16 projects the seven effects, determined based on their respective common components, as a function of sample indexing. Factor A (growing sites), determined by CC3 with approximately 7% variance, distinctly exhibits score differences between the spectralprints of yerba mate leaves cultivated in Espumoso and São Mateus do Sul (Figure 16). Loading for this component and main effect reveals that wavelengths at 5410 nm, 5122 nm, and 4440 nm contribute significantly to the construction of this CC in the MIR region (Figure 17). The absorption band at 5410 nm corresponds to a characteristic absorption region for carbonyl

groups, commonly found in aldehydes, ketones, acids, esters, and amides (1650-1850 cm^{-1}) (Afonso *et al.*, 2019). Meanwhile, the absorption band at 4440 nm is indicative of C≡N bond (nitriles) (Beć, Krzysztof Bernard; Grabska; Huck, 2022).

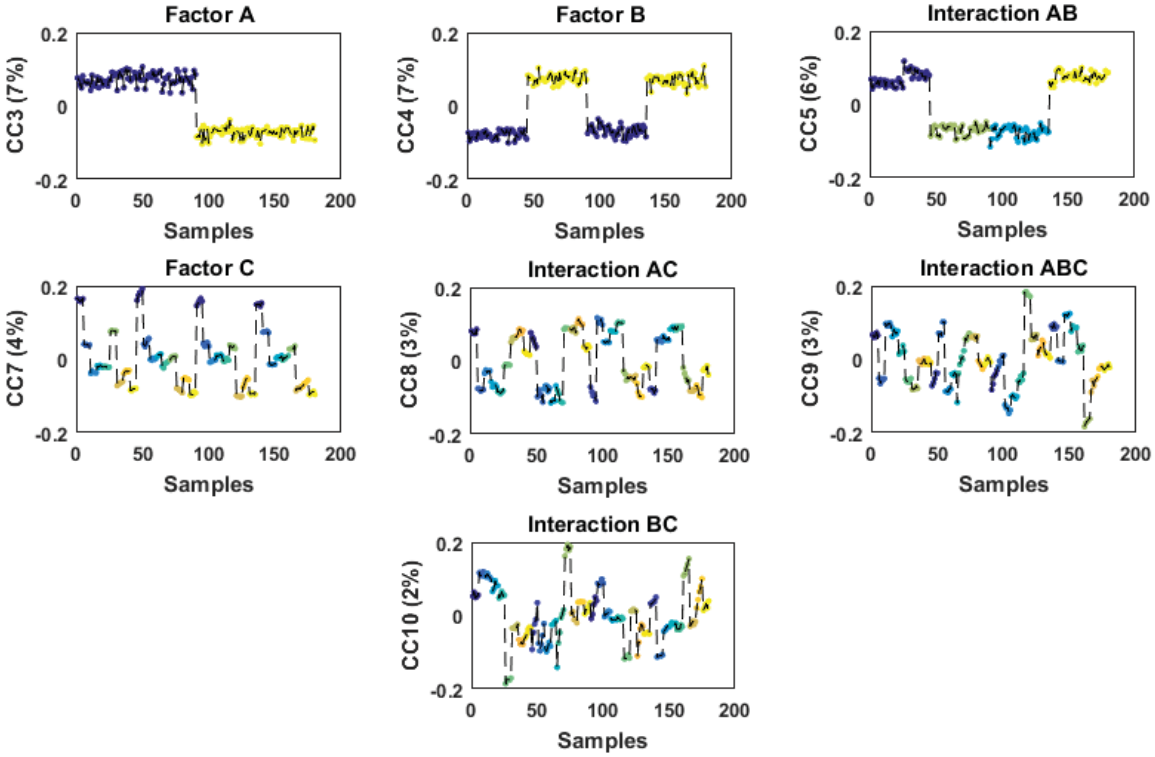
FIGURE 15 – F-VALUES BASED ON COMMON COMPONENTS (CC) 1 TO 12, HIGHLIGHTING HIGH SALIENCIES FOR RESIDUALS CALCULATED FROM ACOMDIM OF MIR (A) AND NIR (B) SPECTRALPRINTS IN YERBA MATE LEAVES. FACTORS INCLUDE DIFFERENT GROWING SITES (FACTOR A), HARVEST SEASONS (FACTOR B), AND CLONES (FACTOR C). ANOVA TERMS ENCOMPASS PURE EFFECTS (A, B, AND C), BINARY EFFECTS (AB, AC, AND BC), TERNARY EFFECTS (ABC), AND THE RESIDUAL MATRIX (R). THE RED LINE (·····) INDICATES THE F-CRITICAL VALUE FOR BOTH THE NIR AND MIR MATRICE



SOURCE: The author (2025).

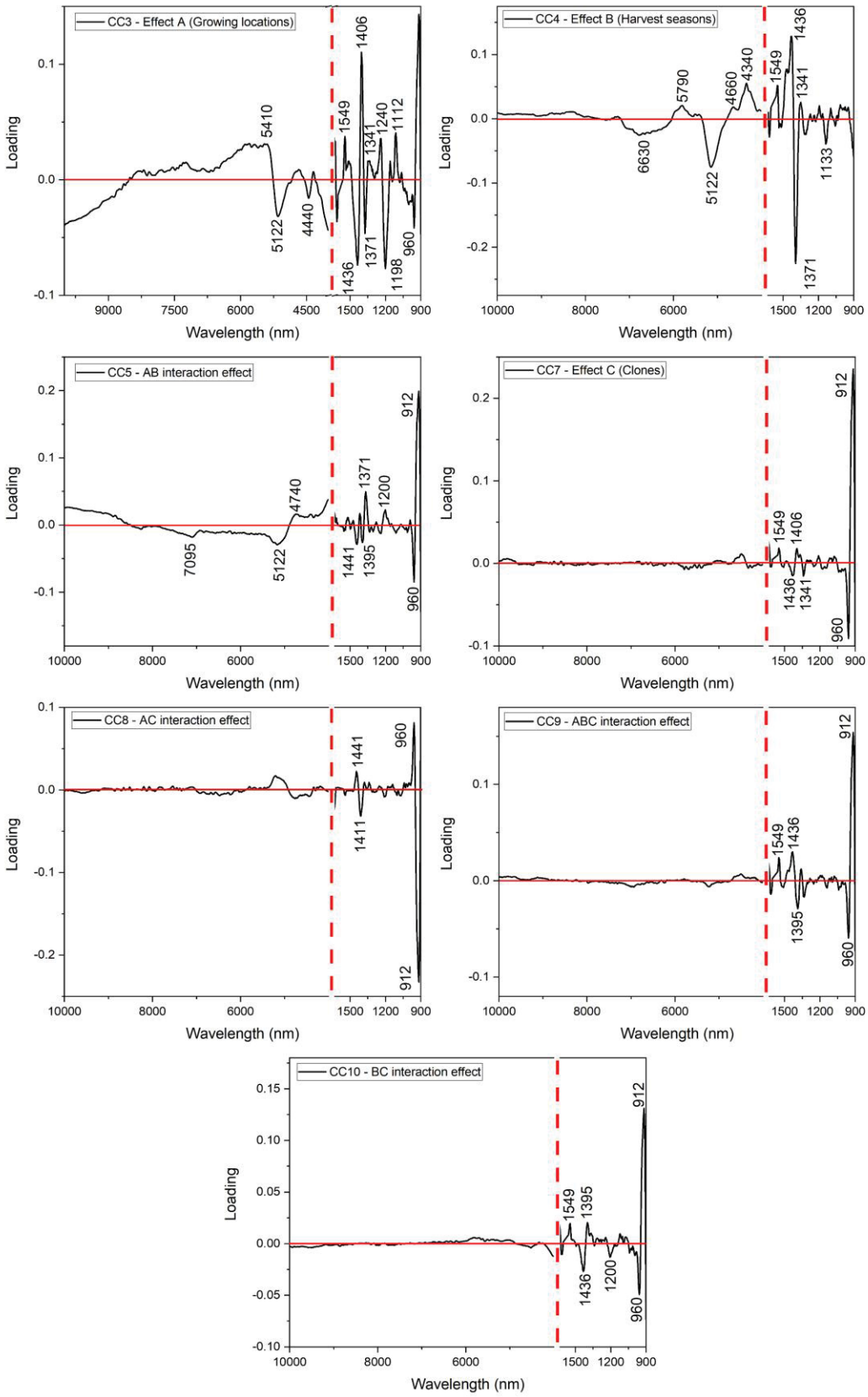
In the NIR region, bands at 1436 nm, 1371 nm, 1198 nm, and 960 nm are associated with yerba mate plants cultivated in São Mateus do Sul (negative region of the scores). The absorption bands at 1436 and 1371 nm are attributed to the first overtone of C-H bond stretching vibrations, which are characteristic absorption bands commonly associated with lipids (fatty acids) and carbohydrates (Li; Zuo; Wang, 2024). Additionally, in this region (7000-8000 cm-1), a broad signal characteristic of the O-H bond present in water can also be observed (Sun, 2009). The band at 1198 nm is characteristic of fatty acid absorptions via C-H bond vibrations (8600-8150 cm-1) (Sun, 2009), but it is also attributed to the C-H stretching of carbohydrates (Türker-Kaya; Huck, 2017). The 960 nm band corresponds to the second overtone of the nonbonded OH stretch (Weyer; Lo, 2006). Conversely, bands at 1549 nm, 1406 nm, 1341 nm, 1240 nm, and 1112 nm contribute to the separation of scores in the positive region, corresponding to Espumoso yerba mate plants (Figures 16 and 17). The bands at 1549, 1406, and 1341 nm are likely characteristic of C-H bonds or O-H stretching vibrations (Li; Zuo; Wang, 2024). The band at 1240 nm may be due to the second overtone of C-H stretching (Peiris *et al.*, 2021).

FIGURE 16 – SCORE PLOT OF THE COMMON COMPONENTS (CCS) ASSOCIATED WITH MAIN EFFECT A (GROWING SITES), MAIN EFFECT B (HARVEST SEASONS), INTERACTION EFFECT AB (GROWING SITES × HARVEST SEASONS), MAIN EFFECT C (CLONES), INTERACTION EFFECT AC (GROWING SITES × CLONES), INTERACTION EFFECT ABC (GROWING SITES × HARVEST SEASONS × CLONES), AND INTERACTION EFFECT BC (HARVEST SEASONS × CLONES). THE MULTIBLOCK SPECTRALPRINT IN YERBA MATE LEAVES INCORPORATES BOTH NEAR-INFRARED (NIR) AND MID-INFRARED (MIR) SPECTROSCOPY. ANOVA TERMS ENCOMPASS PURE EFFECTS (A, B, AND C), BINARY EFFECTS (AB, AC, AND BC), AND TERNARY EFFECTS (ABC). SAMPLE INDEXING FOLLOWS TABLE 9, CONSIDER 5 REPLICATIONS PER SAMPLE



SOURCE: The author (2025).

FIGURE 17 – LOADING PLOTS OF THE COMMON COMPONENTS (CCS) ASSOCIATED WITH MAIN EFFECT A (GROWING SITES), MAIN EFFECT B (HARVEST SEASONS), INTERACTION EFFECT AB (GROWING SITES × HARVEST SEASONS), MAIN EFFECT C (CLONES), INTERACTION EFFECT AC (GROWING SITES × CLONES), INTERACTION EFFECT ABC (GROWING SITES × HARVEST SEASONS × CLONES), AND INTERACTION EFFECT BC (HARVEST SEASONS × CLONES). THE MULTIBLOCK SPECTRALPRINT IN YERBA MATE LEAVES INCORPORATES BOTH MID-INFRARED (MIR) AND NEAR-INFRARED (NIR) SPECTROSCOPY. ANOVA TERMS ENCOMPASS PURE EFFECTS (A, B, AND C), BINARY EFFECTS (AB, AC, AND BC), AND TERNARY EFFECTS (ABC). THE DASHED VERTICAL LINE DIVIDES THE NIR AND MIR LOADING BLOCK



SOURCE: The author (2025).

The differences between the two growing sites determined the chemical differences in the mate leaves through CC3. Cultivation sites have different edaphoclimatic characteristics (Table 6), and may also have numerous other abiotic and biotic variations that were not evaluated. In this case, solar availability can be the main factor responsible for the chemical differences in the metabolism of the plants: in Espumoso, the yerba mate trees are fully exposed to the sun, while in São Mateus, they are grown under the shade of other trees. Light incidence affects the accumulation of secondary metabolities such as caffeine, theobromine, and caffeoylquinic acids in yerba mate leaves, and clones also may have different metabolic responses to light (Aguiar *et al.*, 2024a). Compounds of primary metabolism can also be significantly influenced by light, as is the case with sugars and fatty acids (Melo, 2018).

Factor B (harvest seasons), identified by CC4, exhibits a distinct separation of scores for yerba mate leaves collected in winter (1) and summer (2) with 6.5% of variance explained (Figure 16). Loadings indicate that bands at 5790 nm, 4660 nm, 4340 nm, 6630 nm, and 5122 nm were significant in constructing this CC in the MIR region. The 5790 nm band is characteristic of C=O bonds, typically attributed to proteins or pigments like pheophytin *a* (Nikalje *et al.*, 2019). The 6630 nm band, indicative of proteins, can be attributed to CH_2 and CH_3 groups or to the carbonyl group present in esters, amides, acids, and other components like xanthines and saponins (Marcelo; Pozebon; Ferrão, 2015; Nikalje *et al.*, 2019). Additionally, the C=C aromatic stretch bond characteristic of lignins can be characterized in the region of 1515-1505 cm^{-1} (Türker-Kaya; Huck, 2017). In the NIR region, bands at 1549 nm, 1436 nm, 1341 nm, 1371 nm, and 1133 nm contributed to the modeling of this CC and determination of main effect B (Figure 17).

Growing sites and harvest seasons showed similar variations, with chemical effects of the same magnitude (approximately 7% variance). The yerba mate harvest period has a direct impact on their chemical composition and, in some cases, may be the most significant factor affecting chemical alterations (Tormena *et al.*, 2020), especially when collected in contrasting seasons such as winter and summer. The influence of seasonality on methylxanthines and chlorogenic acids in yerba mate has already been reported (Aguiar *et al.*, 2024a; Duarte *et al.*, 2023). Other compounds, such as boric acid, allantoin, ferulic acid, and ursolic acid were not detected in the leaves of two yerba mate clones in at least one of the four seasons, demonstrating a strong effect of seasonality for these compounds (Melo, 2018).

Primary and secondary metabolites can be affected by abiotic stresses, such as high or low temperatures. During cold seasons, plants synthesize many cryoprotectant metabolites such as soluble sugars, amino acids, polyamines, and sugar alcohol. While high temperatures and solar radiation of summer may increase other compound contents, including proline, alanine, allantoin, myoinositol, rhamnose, and protective compounds like flavonoids, tocopherols, and carotenoids (Parida; Panda; Rangani, 2018). Traditionally, summer and winter are the preferred seasons for harvesting yerba mate, due to the vegetative rest, but it can be extended throughout the year (Aguiar *et al.*, 2022; Aguiar *et al.*, 2024a). Therefore, as confirmed by both MIR and NIR spectral analysis, the chemical composition of yerba mate is influenced by the harvest period, and consequently, raw material destined for industry has variations in its chemical composition depending on the season. The use of portable spectroscopic equipment could, in the future, be very useful to determine the ideal time for yerba mate harvest in the field, by monitoring the content of compounds of interest *in vivo*.

CC5 effectively modeled the AB (growing sites × harvest seasons) interaction, demonstrating a clear separation between scores (Figure 16) with 5.7% of variance explained. In the region of negative scores, yerba mate leaves collected in Espumoso in the summer (3) and São Mateus in the winter (2) were located, with loadings at 7095 nm and 5122 nm in the MIR region and 1441 nm, 1395 nm, and 960 nm in the NIR region (Figure 17). The 7095 nm absorption region is characterized by scissor vibrations of CH_3 and CH_2, as well as the stretching vibration of C-O bonds (esters) (Marcelo; Pozebon; Ferrão, 2015). Additionally, groups such as N=O (nitro compounds), CO-H (aldehydes), and O-H (alcohols) can also be found in this region (Nikalje *et al.*, 2019). It is often associated with the presence of carbohydrates (1500-1200 cm^{-1}). The bands at 1441 and 1395 nm are characteristic of the first overtone of C-H bond stretching vibrations, typically found in fatty acids and carbohydrates (Li; Zuo; Wang, 2024; Türker-Kaya; Huck, 2017). Conversely, in the positive scores, spectralprints of yerba mate leaves collected in Espumoso in winter (1) and São Mateus in summer (4) were observed (Figure 16), with loadings at 4740 nm in the MIR region and 1371 nm, 1200 nm, and 912 nm in the NIR region (Figure 17). The 4740 nm band may be characteristic of terminal alkynes due to the C≡C bond, typically appearing as bands of weak absorption (Beć, Krzysztof Bernard; Grabska; Huck, 2022). The band at 1200 nm may arise from the second overtone of C-H stretching in compounds like fatty acids (Li; Zuo; Wang,

2024; Sun, 2009). The 912 nm region characterizes the third overtone of C-H stretching in alkanes (Weyer; Lo, 2006).

It is important to emphasize the significance of the growing site × harvest seasons interaction effect; substantial evidence suggests that combinations of stresses can pose more potent and realistic threats to plant growth and productivity than individual stresses (Priya *et al.*, 2023). Although the individual effects of growing site and harvest seasons were found to be significant in the CCs 3 and 4, each explaining approximately 7% of the variance, the interaction effect growing site × harvest seasons explains a variance nearly as high as these individual effects and even exceeds the variance explained by the main effect of clones. Several environmental factors simultaneously modulate the synthesis and accumulation of plants' secondary metabolites, for example (Yang *et al.*, 2018). Despite numerous studies on combined stresses, plant molecular responses to such combined stresses remain poorly understood (Priya *et al.*, 2023). As demonstrated in this study, using AComDim, the interaction effect between various environmental stresses can be accurately assessed.

The main effect of clones (C) was determined through CC7 and represents 3.6% of variance explained. Some primary metabolites, such as reducing sugars and fatty acids, are strongly influenced by the environment, while genetic material has little effect on the content of these compounds (Melo, 2018). This might explain the lower variance determined by clones, compared to cultivation sites, harvest seasons, and their interaction. The separation of scores for clones EC25 (1) and EC37 (2) in the positive region (Figures 16 and 17), with loadings at 1549 nm, 1406 nm, and 912 nm contributed significantly to modeling this CC (Figure 17). Clones BRS 409 (7), Aupaba (8), Yari (9) were positioned in the negative region of the scores with loadings at 1436 nm, 1341 nm, and 960 nm. Clones EC43 (3), EC40 (4), EC22 (5), BRS 408 (6) were located at the origin of the graphic (0,0); therefore, the spectralprint of these clones can be considered an average of the positive and negative scores. The clones analyzed come from different municipalities in two states of southern Brazil (Table 6), which may contribute to chemical variation, though genetic diversity is greater within than among yerba mate populations (Friedrich *et al.*, 2017).

Studies with various genetic materials from yerba mate showed great chemical variability and important differences in the contents of proteins, methylxanthines, phenolic compounds, and saponins, for example (Aguiar *et al.*, 2024b; Duarte *et al.*, 2023; Vieira *et al.*, 2021); these results demonstrate the potential of clonal yerba mate cultivations for industrial

purposes with interest in specific compounds. In multiblock spectralprints, the clone EC25 distances itself from the others, in the positive CC7 score plot, indicating a divergent chemical profile; future analyses must be carried out to determine which key compounds distinguish it. We know that clones EC22 and EC40 are very divergent in terms of methylxanthines, with EC40 being classified as having high caffeine content (and low theobromine content) and EC22 as having decaffeinated mature leaves (and a high theobromine content) (Aguiar *et al.*, 2024a). In previous studies with these two clones subjected to five levels of shading, the NIR models could separate clones, and the genetic factor showed a larger effect than shading on spectral variability (Almeida *et al.*, 2022, 2023). We highlight that in previous studies only two clones were analyzed (facilitating their distinction) and that the cultivation system in which the plants were located was protected, with lower environmental variations and a high proportion of young leaves. However, now when we analyzed them in the set of nine clones, EC22 and EC40 were close in MIR and NIR spectralprints of mature leaves, indicating chemical similarity despite different caffeine levels.

Although visualizing the distribution of scores for binary and ternary effects poses challenges, these interactions tell us that each clone responds differently to the environment. The significance of the AC effect can be inferred from the distinctiveness of the score distribution (Figure 16). For instance, considering the same clone and irrespective of the harvest seasons (e.g., clone EC25, Espumoso (1)), its scores consistently fall in the positive region of the graph, while São Mateus do Sul (2) is consistently separated in the negative region, a pattern observed for all different clones with only the positive/negative region of score distribution changing. Another illustrative example involves the scores of the Yari clone in Espumoso (17), which consistently appears in the positive region, whereas in São Mateus do Sul (18) the same genetic material consistently appears in the negative region. The loading plot for the AC binary effect shows that the bands at 1441 and 960 nm (positive values) and 1411 and 912 nm (negative values) contribute to the modeling of this CC (Figure 17). The 1411 nm region is characteristic of C-H bond vibrations, typically found in fatty acids and carbohydrates (Sun, 2009).

CC9 delineated the ABC ternary effect, explaining less than 3% of the variance. Notably, in the negative region, there is a distinctive separation in the spectralprint scores of yerba mate leaves from clone BRS 408 grown in Espumoso and collected in the summer (24), and the winter collections in São Mateus do Sul of clones EC43 (10) and EC40 (14) (Figure 16). Conversely, in the positive region, clones EC43 collected in the summer (12) and BRS 408

in the winter (22) in São Mateus do Sul were grouped. The loadings contributing to the modeling of this CC were 1395 nm and 960 nm, highlighted in the negative region of the plot, and 1549 nm, 1436 nm, and 912 nm in the positive loading plot (Figure 17).

Finally, with less than 3% of the variance explained, the BC effect was determined in CC10. The graphic projection of the scores emphasizes the separation of BRS 408 clone harvested in winter (11; negative region of scores) and summer (12; positive region of scores) (Figure 16). Contributing to the separation of these scores were bands at 1549 nm, 1395 nm, and 912 nm in the positive region of loadings, and 1436 nm, 1200 nm, and 960 nm in the negative region of loadings (Figure 17).

Our study demonstrated that the growing site and harvest season have similar size impacts on yerba mate chemical composition, which in turn has a greater effect than the variability observed among different clones. This chemometric approach enables the identification of factors that cause substantial chemical changes in leaf composition, as well as those that contribute minimal or no variance in plants of commercial interest. This methodology is invaluable for screening and selecting plants, especially when assessing their performance in different environments to ensure the maintenance of desired chemical attributes. Additionally, it allows for the separation and identification of clones with divergent chemical profiles, eliminating variations inherent to the environment. This is achieved through the AComDim method, which separates variances related to main effects and interaction effects. Ultimately, this approach facilitates the identification of plants with unique chemical attributes, that may be of interest to certain industries, such as cosmetics, food, or pharmaceuticals, in addition to determining agronomic strategies to maximize the contents of compounds of interest.

5.4 CONCLUSION

Application of the AComDim multivariate analysis technique, combining MIR and NIR spectroscopy, leads to the conclusion that the growing site, harvesting season, and their interaction, represent the main effects of the variance in yerba mate leaves spectralprints. These effects were effectively captured by both NIR and MIR spectroscopy. Specifically, the MIR spectral region highlighted significant absorption bands linked to carbonyl and nitrile groups, while the NIR region revealed bands associated with lipids, carbohydrates, and water. The

variation due to yerba mate clones was also detected by NIR spectroscopy, accounting for a substantial proportion of the variance in the chemical composition of the leaves, inferring that clones such as EC25 present a divergent spectralprint, when compared with clones BRS 409, Aupaba, and Yari. The application of AComDim allowed for a comprehensive evaluation of the main effects and interactions, revealing the potential for using biospectroscopy as a non-destructive, rapid, and environmentally friendly method for quality control and standardization of the yerba mate for industries.

6 FINAL CONSIDERATIONS

This study is innovative in establishing a rapid and low-cost method for estimating total saponin content in yerba mate; it is also pioneering in analyzing the genetic parameters of bioactive compounds in yerba mate clonal trials, and in applying a robust statistical methodology to combine NIR and MIR spectra into a high-accuracy global model. The influence of genetic factors on saponin accumulation is evident, and specific genetic materials are recommended for different industrial purposes. Cloning is a key tool for standardizing raw material targeting markets interested in saponin content. The results reinforce the great potential of the species as a source of saponins, which may lead to new industrial applications. Moreover, the multiblock analysis of infrared spectral fingerprints revealed significant variations in primary metabolites due to environmental factors, cultivation site, and harvest season. The application of these non-destructive infrared methodologies in field and industrial settings may bring significant advances in the standardization and quality control of yerba mate-derived products.

Further research is recommended using additional analytical techniques, such as liquid chromatography coupled with mass spectrometry (LC/MS), to identify the specific saponins present and quantify each saponin molecule in yerba mate clones. Studies conducted to date have used only one or a few materials for the identification and quantification of saponins. Given the considerable variation in saponin content among genetic materials and the molecular diversity of saponins within the species, it is likely that different saponins predominate in each genetic material representing a broad and promising field for further investigation. Moreover, research on the extraction and application of yerba mate saponins, either individually or in combination, in pharmacological studies is still in a preliminary stage. It is also essential to investigate their interactions with other secondary compounds of the species to fully elucidate their bioactive effects in humans and other animal organisms.

REFERENCES

AFONSO, S.; SILVA, F. B.; MARCHEAFAVE, G. G.; HATUMURA, P. H.; BRUNS, R. E.; SCARMINIO, I. S. Influence of seasonality and sunlight effects on *Rollinia mucosa* leaves fingerprint. **Journal of the Brazilian Chemical Society**, v. 30, n. 5, p. 968–977, 2019. https://doi.org/10.21577/0103-5053.20180242.

AGUIAR, N. S. de; GABIRA, M. M.; DUARTE, M. M.; TOMASI, J. de C.; HANSEL, F. A.; LAVORANTI, O. J.; DESCHAMPS, C.; HELM, C. V.; WENDLING, I. How shading levels affect bioactive compounds in leaves of yerba mate clones. **Biochemical Systematics and Ecology**, v. 113, 2024a. https://doi.org/10.1016/j.bse.2024.104796.

AGUIAR, N. S. de; GABIRA, M. M.; SANTIN, D.; DESCHAMPS, C.; HELM, C. V.; WENDLING, I. Planting seasons and environments in initial field establishment of yerba mate clonal cultivars in Southern Brazil. **Revista Ceres**, v. 70, n. 6, 2023. https://doi.org/10.1590/0034-737X202370060006.

AGUIAR, N. S. de; GABIRA, M. M.; TOMASI, J. D. C.; DUARTE, M. M.; VIEIRA, L. M.; LAVORANTI, O. J.; WENDLING, I. Productivity of clonal *Ilex paraguariensis* genotypes in a semi-hydroponic system is reduced by shading. **Forest Science**, v. 68, n. 4, p. 1–8, 2022. https://doi.org/10.1093/forsci/fxac028.

AGUIAR, N. S. de; HANSEL, F. A.; REIS, C. A. F.; LAZZAROTTO, M.; WENDLING, I. Optimizing the vanillin-acid sulfuric method to total saponin content in leaves of yerba mate clones. **Chemistry and Biodiversity**, v. 21, n. 4, 2024b. https://doi.org/10.1002/cbdv.202301883.

ALMEIDA, A. G. de; TORMENA, C. D.; AGUIAR, N. S. de; WENDLING, I.; RAKOCEVIC, M.; PAULI, E. D.; SCARMINIO, I. S.; BRUNS, R. E.; MARCHEAFAVE, G. G. Direct NIR spectral determination of genetic improvement, light availability, and their interaction effects on chemically selected yerba-mate leaves. **Microchemical Journal**, v. 191, p. 1–7, 2023. https://doi.org/10.1016/j.microc.2023.108828.

ALMEIDA, A. G.; PAULI, E. D.; TORMENA, C. D.; WENDLING, I.; RAKOCEVIC, M.; BRUNS, R. E.; SCARMINIO, I. S.; MARCHEAFAVE, G. G. Portable NIR spectroscopy-chemometric identification of chemically differentiated yerba mate (*Ilex paraguariensis*) clones. **Food Analytical Methods**, 2022. https://doi.org/10.1007/s12161-022-02431-y.

ALVES, F. E. da S. B.; SCHEER, A. de P. Yerba mate (*Ilex paraguariensis*), science, technology and health: A systematic review on research, recent advances and possible paths for future studies. **South African Journal of Botany**, v. 168, p. 573–587, 2024. https://doi.org/10.1016/J.SAJB.2024.04.008.

ANDRADE, F.; ALMEIDA, C.; ALBUQUERQUE, C. De; MARASCHIN, M.; LUIZ, E. Safety assessment of yerba mate (*Ilex paraguariensis*) dried extract: Results of acute and 90 days subchronic toxicity studies in rats and rabbits. **Food and Chemical Toxicology**, v. 50, n. 2, p. 328–334, 2012. https://doi.org/http://dx.doi.org/10.1016/j.fct.2011.08.028.

AUGUSTIN, J. M.; KUZINA, V.; ANDERSEN, S. B.; BAK, S. Molecular activities, biosynthesis and evolution of triterpenoid saponins. **Phytochemistry**, v. 72, n. 6, p. 435–457, 2011. https://doi.org/10.1016/j.phytochem.2011.01.015.

BEĆ, K. B.; GRABSKA, J.; HUCK, C. W. Near-infrared spectroscopy in bio-applications. **Molecules**, v. 25, n. 12, 2020. https://doi.org/10.3390/molecules25122948.

BEĆ, K. B.; GRABSKA, J.; HUCK, C. W. Physical principles of infrared spectroscopy. **Comprehensive Analytical Chemistry**. Amsterdam: Elsevier B.V., 2022. v. 98, p. 1–43. https://doi.org/10.1016/bs.coac.2020.08.001.

BENEDITO, D. C. D.; STUEPP, C. A.; HELM, C. V.; LIZ, M. V. de; MIRANDA, A. C. de; IMOSKI, R.; LAVORANTI, O. J.; WENDLING, I. Bioactive compounds concentrations and stability in leaves of *Ilex paraguariensis* genotypes. **Forests**, v. 14, n. 12, 2023. https://doi.org/10.3390/f14122411.

BERHOW, M. A.; SINGH, M.; BOWMAN, M. J.; PRICE, N. P. J.; VAUGHN, S. F.; LIU, S. X. Quantitative NIR determination of isoflavone and saponin content of ground soybeans. **Food Chemistry**, v. 317, p. 1–9, 2020. https://doi.org/10.1016/j.foodchem.2020.126373.

BISWAS, T.; DWIVEDI, U. N. Plant triterpenoid saponins: biosynthesis, in vitro production, and pharmacological relevance. **Protoplasma**, v. 256, n. 6, p. 1463–1486, 2019. https://doi.org/10.1007/s00709-019-01411-0.

BITENCOURT, R. G.; QUEIROGA, C. L.; MONTANARI, Í.; CABRAL, F. A. Fractionated extraction of saponins from Brazilian ginseng by sequential process using supercritical CO2, ethanol and water. **The Journal of Supercritical Fluids**, v. 92, p. 272–281, 2014. http://dx.doi.org/10.1016/j.supflu.2014.06.009.

BLUM-SILVA, C. H.; CHAVES, V. C.; SCHENKEL, E. P.; COELHO, G. C.; REGINATTO, F. H. The influence of leaf age on methylxanthines, total phenolic content, and free radical scavenging capacity of *Ilex paraguariensis* aqueous extracts. **Brazilian Journal of Pharmacognosy**, v. 25, n. 1, p. 1–6, 2015. https://doi.org/http://dx.doi.org/10.1016/j.bjp.2015.01.002.

BORRÉ, G. L.; KAISER, S.; PAVEI, C.; DA SILVA, F. A.; BASSANI, V. L.; ORTEGA, G. G. Comparison of methylxanthine, phenolics and saponin contents in leaves, branches and unripe fruits from *Ilex paraguariensis* A. St.-Hil (mate). **Journal of Liquid Chromatography and Related Technologies**, v. 33, n. 3, p. 362–374, 2010. https://doi.org/10.1080/10826070903526055.

BOUVERESSE, D. J.-R.; PINTO, R. C.; SCHMIDTKE, L. M.; LOCQUET, N.; RUTLEDGE, D. N. Identification of significant factors by an extension of ANOVA-PCA based on multi-block analysis. **Chemometrics and Intelligent Laboratory Systems**, v. 106, n. 2, p. 173–182, 2011. https://doi.org/10.1016/j.chemolab.2010.05.005.

BRACESCO, N. *Ilex paraguariensis* as a healthy food supplement for the future world. **Biomedical Journal of Scientific & Technical Research**, v. 16, n. 1, p. 15–18, 2019. https://doi.org/10.26717/bjstr.2019.16.002808.

BRACESCO, N.; SANCHEZ, A. G.; CONTRERAS, V.; MENINI, T.; GUGLIUCCI, A. Recent advances on *Ilex paraguariensis* research: Minireview. **Journal of Ethnopharmacology**, v. 136, n. 3, p. 378–384, 2011. https://doi.org/10.1016/j.jep.2010.06.032.

BROADHURST, R. B.; JONES, W. T. Analysis of condensed tannins using acidified vanillin. **Journal of the Science of Food and Agriculture**, v. 29, p. 788–794, 1978. https://doi.org/10.1002/jsfa.2740290908.

BUTIUK, A. P.; MAIDANA, S. A.; ADACHI, O.; AKAKABE, Y.; MARTOS, M. A.; HOURS, R. A. Optimization and modeling of the chlorogenic acid extraction from a residue of yerba mate processing. **Journal of Applied Research on Medicinal and Aromatic Plants**, v. 25, p. 1–7, 2021. https://doi.org/10.1016/j.jarmap.2021.100329.

BUTIUK, A. P.; MARTOS, M. A.; ADACHI, O.; HOURS, R. A. Study of the chlorogenic acid content in yerba mate (*Ilex paraguariensis* St. Hil.): Effect of plant fraction, processing step and harvesting season. **Journal of Applied Research on Medicinal and Aromatic Plants**, v. 3, n. 1, p. 27–33, 2016. https://doi.org/10.1016/j.jarmap.2015.12.003.

CARDOZO JUNIOR, E. L.; DONADUZZI, C. M.; FERRARESE-FILHO, O.; FRIEDRICH, J. C.; GONELA, A.; STURION, J. A. Quantitative genetic analysis of methylxanthines and phenolic compounds in mate progenies. **Pesquisa Agropecuária Brasileira**, v. 45, n. 2, p. 171–177, 2010. https://doi.org/10.1590/s0100-204x2010000200008.

CARDOZO JUNIOR, E. L.; MORAND, C. Interest of mate (*Ilex paraguariensis* A. St.-Hil.) as a new natural functional food to preserve human cardiovascular health - A review. **Journal of Functional Foods**, v. 21, p. 440–454, 2016. https://doi.org/10.1016/j.jff.2015.12.010.

CARVALHO, P. E. R. Erva-Mate - *Ilex paraguariensis*. **Espécies Arbóreas Brasileiras**. Brasília: Embrapa, 2003. v. 1. Available at: https://www.embrapa.br/florestas/publicacoes/especies-arboreas-brasileiras. Accessed on: 4 ago. 2024.

CERQUEIRA, U. M. F. M.; BEZERRA, M. A.; FERREIRA, S. L. C.; DE JESUS ARAÚJO, R.; DA SILVA, B. N.; NOVAES, C. G. Doehlert design in the optimization of procedures aiming food analysis – A review. **Food Chemistry**, v. 364, p. 1–10, 2021. https://doi.org/10.1016/j.foodchem.2021.130429.

CHEOK, C. Y.; SALMAN, H. A. K.; SULAIMAN, R. Extraction and quantification of saponins: A review. **Food Research International**, v. 59, p. 16–40, 2014. https://doi.org/10.1016/j.foodres.2014.01.057.

CLEMENTE, P. A.; MENDONÇA, K. C.; ANDRADE, G. F. de; GONCALVES, L. M.; GODOY, C. M. T.; ROCHA, J. M. P.; SILVA, W. C. F. N. da; REOLON, J. B.; FERREIRA, D. F.; VIEIRA, M. C. U.; SILVA, J. M.; FERREIRA, L. M.; BONINI, J. S. Technological prospection and scientific innovation of *Ilex paraguariensis* Saint-Hilaire in the wound healing process. **Brazilian Archives of Biology and Technology**, v. 67, 2024. https://doi.org/10.1590/1678-4324-2024230086.

COELHO, G. C.; GNOATTO, S. B.; BASSANI, V. L.; SCHENKEL, E. P. Quantification of saponins in extractive solution of mate leaves (*Ilex paraguariensis* A. St. Hil.). **Journal of Medicinal Food**, v. 13, n. 2, p. 439–443, 2010. https://doi.org/10.1089/jmf.2009.0046.

COSTA, F. de; YENDO, A. C. A.; FLECK, J. D.; GOSMANN, G.; FETT-NETO, A. G. Accumulation of a bioactive triterpene saponin fraction of *Quillaja brasiliensis* leaves is associated with abiotic and biotic stresses. **Plant Physiology and Biochemistry**, v. 66, p. 56–62, 2013. https://doi.org/10.1016/j.plaphy.2013.02.003.

CROGE, C. P.; CUQUEL, F. L.; PINTRO, P. T. M. Yerba mate: cultivation systems, processing and chemical composition. A review. **Scientia Agricola**, v. 78, n. 5, p. 1–11, 2020. https://doi.org/10.1590/1678-992x-2019-0259.

DAEMO, B. B.; YOHANNES, D. B.; BEYENE, T. M.; ABTEW, W. G. Biochemical analysis of cassava (*Manihot esculenta* Crantz) accessions in southwest of Ethiopia. **Journal of Food Quality**, p. 1–13, 2022. https://doi.org/10.1155/2022/9904103.

DALLABRIDA, V. R.; DUMKE, C. I.; MOLZ, S.; FURINI, V.; GIACOMELLI, M. B. O. Com erva-mate não se faz só chimarrão! Situação atual e perspectivas de inovação no setor ervateiro do Planalto Norte Catarinense. **DRd - Desenvolvimento Regional em Debate**, v. 6, n. 2, p. 247–272, 2016. https://doi.org/10.1017/CBO9781107415324.004.

DUARTE, M. M.; AGUIAR, N. S. de; GABIRA, M. M.; TOMASI, J. de C.; VIEIRA, L. M.; HELM, C. V.; NOGUEIRA, A. C.; WENDLING, I. Seasonality and genotype influence on *Ilex paraguariensis* cuttings rooting and bioactive compounds. **Plant Genetic Resources: Characterization and Utilization**, v. 21, n. 2, p. 174–181, 2023. https://doi.org/10.1017/S147926212300059X.

DUARTE, M. M.; TOMASI, J. de C.; HELM, C. V.; AMANO, E.; LAZZAROTTO, M.; DE GODOY, R. C. B.; NOGUEIRA, A. C.; WENDLING, I. Caffeinated and decaffeinated mate tea: Effect of toasting on bioactive compounds and consumer acceptance. **Revista Brasileira de Ciências Agrárias**, v. 15, n. 3, p. 1–10, 2020. https://doi.org/10.5039/agraria.v15i3a8513.

FAIZAL, A.; GEELEN, D. Saponins and their role in biological processes in plants. **Phytochemistry Reviews**, v. 12, n. 4, p. 877–893, 2013. https://doi.org/10.1007/s11101-013-9322-4.

FAOSTAT. **Crops.** 2025. Available at: https://www.fao.org/faostat/en/#data/QCL/visualize. Accessed on: 15 jan. 2025.

FERRERA, T. S.; HELDWEIN, A. B.; DOS SANTOS, C. O.; SOMAVILLA, J. C.; SAUTTER, C. K. Substâncias fenólicas, flavonoides e capacidade antioxidante em erveiras sob diferentes coberturas do solo e sombreamentos. **Revista Brasileira de Plantas Medicinais**, v. 18, n. 2, p. 588–596, 2016. https://doi.org/10.1590/1983-084X/15_197.

FERRON, R. M. Situação da erva-mate no Brasil. 2016. **Seminário Erva-Mate XXI** [...]. Curitiba, PR: Embrapa, 2016. p. 25–29. Available at: https://www.researchgate.net/publication/314229862.

FIGUEIREDO, M. de; GIANNOUKOS, S.; WÜTHRICH, C.; ZENOBI, R.; RUTLEDGE, D. N. A tutorial on the analysis of multifactorial designs from one or more data sources using AComDim. **Journal of Chemometrics**, v. 37, n. 7, 2022.

FRIEDRICH, J. C.; GONELA, A.; VIDIGAL, M. C. G.; VIDIGAL FILHO, P. S.; STURION, J. A.; CARDOZO JUNIOR, E. L. Genetic and phytochemical analysis to evaluate the diversity and relationships of mate (*Ilex paraguariensis* A.St.-Hil.) elite genetic resources in a germplasm collection. **Chemistry and Biodiversity**, v. 14, n. 3, 2017. https://doi.org/10.1002/cbdv.201600177.

GERBER, T.; NUNES, A.; MOREIRA, B. R.; MARASCHIN, M. Yerba mate (*Ilex paraguariensis* A. St.-Hil.) for new therapeutic and nutraceutical interventions: A review of patents issued in the last 20 years (2000–2020). **Phytotherapy Research**, v. 37, n. 2, p. 527–548, 2023. https://doi.org/10.1002/ptr.7632.

GNOATTO, S. C. B.; SCHENKEL, E. P.; BASSANI, V. L. HPLC method to assay total saponins in *Ilex paraguariensis* aqueous extract. **Journal of the Brazilian Chemical Society**, v. 16, n. 4, p. 723–726, 2005. https://doi.org/10.1590/s0103-50532005000500007.

GOULART, I. C. G. dos R. **Fatores que afetam a produtividade e a adoção de tecnologias na cultura da erva-mate**. 2020. 107 f. Thesis (PhD in Agronomy) – Universidade Federal do Paraná, Curitiba, 2020.

GÜÇLÜ-ÜSTÜNDAG, Ö.; MAZZA, G. Saponins: Properties, applications and processing. **Critical Reviews in Food Science and Nutrition**, v. 47, n. 3, p. 231–258, 2007. https://doi.org/10.1080/10408390600698197.

GUERRA, F.; SEPÚLVEDA, S. Saponin production from *Quillaja* genus species. An insight into its applications and biology. **Scientia Agricola**, v. 78, n. 5, p. 1–9, 2020. https://doi.org/10.1590/1678-992x-2019-0305.

HAAS, J.; MIZAIKOFF, B. Advances in mid-infrared spectroscopy for chemical analysis. **Annual Review of Analytical Chemistry**, v. 9, p. 45–68, 12 jun. 2016. https://doi.org/10.1146/annurev-anchem-071015-041507.

HARRINGTON, P. D. B.; VIEIRA, N. E.; ESPINOZA, J.; NIEN, J. K.; ROMERO, R.; YERGEY, A. L. Analysis of variance–principal component analysis: A soft tool for proteomic discovery. **Analytica Chimica Acta**, v. 544, n. 1–2, p. 118–127, 2005. https://doi.org/10.1016/J.ACA.2005.02.042. Accessed on: 14 Jul. 2024.

HE, H.; YAN, S.; LYU, D.; XU, M.; YE, R.; ZHENG, P.; LU, X.; WANG, L.; REN, B. Deep learning for biospectroscopy and biospectral imaging: State-of-the-art and perspectives. **Analytical Chemistry**, v. 93, n. 8, p. 3653–3665, 2021. https://doi.org/10.1021/acs.analchem.0c04671.

HIAI, S.; OURA, H.; NAKAJIMA, T. Color reaction of some sapogenins and saponins with vanillin and sulfuric acid. **Planta medica**, v. 29, p. 116–122, 1976.

HOU, Y.; GAO, X.; LI, S.; CAI, X.; LI, P.; LI, W.; LI, Z. Variable selection based on gray wolf optimization algorithm for the prediction of saponin contents in Xuesaitong dropping

pills using NIR spectroscopy. **Journal of Pharmaceutical Innovation**, 2022. https://doi.org/10.1007/s12247-022-09620-6.

HUSSAIN, M.; DEBNATH, B.; QASIM, M.; BAMISILE, B. S.; ISLAM, W.; HAMEED, M. S.; WANG, L.; QIU, D. Role of saponins in plant defense against specialist herbivores. **Molecules**, v. 24, n. 11, p. 1–21, 2019. https://doi.org/10.3390/molecules24112067.

IBGE. **Produção Agrícola Municipal (PAM)**. 2022a. Available at: https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9117-producao-agricola-municipal-culturas-temporarias-e-permanentes.html. Accessed on: 4 ago. 2024.

IBGE. **Produção da Extração Vegetal e da Silvicultura (PEVS)**. 2022b. Available at: https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9105-producao-da-extracao-vegetal-e-da-silvicultura.html?=&t=resultados. Accessed on: 4 ago. 2024.

JIMÉNEZ, G. G.; DURÁN, A. G.; MACÍAS, F. A.; SIMONET, A. M. Structure, bioactivity and analytical methods for the determination of Yucca saponins. **Molecules**, v. 26, 2021. https://doi.org/10.3390/molecules.

KAUR, K.; GILL, B. S.; SHARMA, S. Assessment of genetic variability, heritability and genetic advance in soybean genotypes. **Journal of Crop and Weed**, v. 13, n. 2, p. 84–89, 2017.

KORIFI, R.; AMAT, S.; RÉBUFA, C.; LABED, V.; RUTLEDGE, D. N.; DUPUY, N. AComDim as a multivariate tool to analyse experimental design application to γ-irradiated and leached ion exchange resins. **Chemometrics and Intelligent Laboratory Systems**, v. 141, p. 12–23, 2015. https://doi.org/10.1016/j.chemolab.2014.12.003.

KORIFI, R.; PLARD, J.; LE DRÉAU, Y.; RÉBUFA, C.; RUTLEDGE, D. N.; DUPUY, N. Highlighting metabolic indicators of olive oil during storage by the AComDim method. **Food Chemistry**, v. 203, p. 104–116, 2016. https://doi.org/10.1016/j.foodchem.2016.01.137.

KUROPKA, P.; ZWYRZYKOWSKA-WODZIŃSKA, A.; KUPCZYŃSKI, R.; WŁODARCZYK, M.; SZUMNY, A.; NOWACZYK, R. M. The effect of *Ilex* x *meserveae* S. Y. Hu extract and its fractions on renal morphology in rats fed with normal and high-cholesterol diet. **Foods**, v. 10, n. 4, 2021. https://doi.org/10.3390/foods10040818.

LANDAU, E. C.; ALVES DA SILVA, G.; TORRES, T. Evolução da produção de erva-mate (*Ilex paraguariensis*, Aquifoliaceae). *In*: LANDAU, E. C.; SILVA, G. A. da; MOURA, L.; HIRSCH, A.; GUIMARAES, D. P. (orgs.). **Dinâmica da Produção Agropecuária e da Paisagem Natural no Brasil nas Últimas Décadas**. Embrapa, 2020. v. 4, p. 709–736.

LE, A. V.; PARKS, S. E.; NGUYEN, M. H.; ROACH, P. D. Improving the vanillin-sulphuric acid method for quantifying total saponins. **Technologies**, v. 6, n. 3, p. 1–12, 2018. https://doi.org/10.3390/technologies6030084.

LI, C.; ZUO, Z.; WANG, Y. Optimization of Fourier transform near-infrared spectroscopy model in determining saponin compounds of *Panax notoginseng* roots. **Vibrational Spectroscopy**, v. 130, 2024. https://doi.org/10.1016/j.vibspec.2023.103615.

LI, J.; ZU, Y. G.; FU, Y. J.; YANG, Y. C.; LI, S. M.; LI, Z. N.; WINK, M. Optimization of microwave-assisted extraction of triterpene saponins from defatted residue of yellow horn (*Xanthoceras sorbifolia* Bunge.) kernel and evaluation of its antioxidant activity. **Innovative Food Science and Emerging Technologies**, v. 11, n. 4, p. 637–643, 2010. https://doi.org/10.1016/j.ifset.2010.06.004.

LIMA, J. D. P.; FARAH, A.; KING, B.; PAULIS, T. De; MARTIN, P. R. Distribution of major chlorogenic acids and related compounds in brazilian green and toasted *Ilex paraguariensis* (maté) leaves. **Journal of Agricultural and Food Chemistry**, v. 64, p. 2361–2370, 2016. https://doi.org/10.1021/acs.jafc.6b00276.

LIU, X.; ZHANG, S.; SI, L.; LIN, Z.; WU, C.; LUAN, L.; WU, Y. A combination of near infrared and mid-infrared spectroscopy to improve the determination efficiency of active components in *Radix Astragali*. **Journal of Near Infrared Spectroscopy**, v. 28, n. 1, 2019. https://doi.org/10.1177/0967033519883793.

LIU, Y. W.; ZHU, X.; LU, Q.; WANG, J. Y.; LI, W.; WEI, Y. Q.; YIN, X. X. Total saponins from *Rhizoma anemarrhenae* ameliorate diabetes-associated cognitive decline in rats: Involvement of amyloid-beta decrease in brain. **Journal of Ethnopharmacology**, v. 139, n. 1, p. 194–200, 2012. http://dx.doi.org/10.1016/j.jep.2011.11.004.

LÓPEZ-HORTAS, L.; PÉREZ-LARRÁN, P.; GONZÁLEZ-MUÑOZ, M. J.; FALQUÉ, E.; DOMÍNGUEZ, H. Recent developments on the extraction and application of ursolic acid. A review. **Food Research International**, v. 103, p. 130–149, 2018. https://doi.org/10.1016/j.foodres.2017.10.028.

MAGEDANS, Y. V. da S.; PHILLIPS, M. A.; FETT-NETO, A. G. Production of plant bioactive triterpenoid saponins: from metabolites to genes and back. **Phytochemistry Reviews**, v. 20, p. 461–482, 2021. https://doi.org/10.1007/s11101-020-09722-4.

MARCELO, M. C. A.; POZEBON, D.; FERRÃO, M. F. Authentication of yerba mate according to the country of origin by using Fourier transform infrared (FTIR) associated with chemometrics. **Food Additives and Contaminants - Part A**, v. 32, n. 8, p. 1215–1222, 2015. https://doi.org/10.1080/19440049.2015.1050702.

MATEOS, R.; BAEZA, G.; MARTÍNEZ-LÓPEZ, S.; SARRIÁ, B.; BRAVO, L. LC–MSn characterization of saponins in mate (*Ilex paraguariens*, St. Hil) and their quantification by HPLC-DAD. **Journal of Food Composition and Analysis**, v. 63, p. 164–170, 2017. https://doi.org/10.1016/j.jfca.2017.08.003.

MATEOS, R.; BAEZA, G.; SARRIÁ, B.; BRAVO, L. Improved LC-MSn characterization of hydroxycinnamic acid derivatives and flavonols in different commercial mate (*Ilex paraguariensis*) brands. Quantification of polyphenols, methylxanthines, and antioxidant activity. **Food Chemistry**, v. 241, p. 232–241, 2018. https://doi.org/10.1016/j.foodchem.2017.08.085.

MAZUR, L.; OLIVEIRA, G. A.; BICUDO, M. O. P.; RIBANI, R. H.; NAGATA, N.; PERALTA-ZAMORA, P. Multivariate calibration and moisture control in yerba mate by near infrared spectroscopy. **Acta Scientiarum**, v. 36, n. 2, p. 369–274, 2014. https://doi.org/10.4025/actascitechnol.v36i2.17777.

MAZUR, L.; PERALTA-ZAMORA, P. G.; DEMCZUK JR., B.; RIBANI, R. H. Application of multivariate calibration and NIR spectroscopy for the quantification of methylxanthines in yerba mate (*Ilex paraguariensis*). **Journal of Food Composition and Analysis**, v. 35, p. 55–60, 2014. https://doi.org/10.1016/j.jfca.2014.04.005.

MEINHART, A. D.; DAMIN, F. M.; CALDEIRÃO, L.; DA SILVEIRA, T. F. F.; FILHO, J. T.; GODOY, H. T. Chlorogenic acid isomer contents in 100 plants commercialized in Brazil. **Food Research International**, v. 99, p. 522–530, 2017. http://dx.doi.org/10.1016/j.foodres.2017.06.017.

MELO, T. O. **Metabólitos como ferramenta de seleção de erva-mate (*Ilex paraguariensis* A. St. Hill.) para fins industriais**. 2018. 204 f. Thesis (PhD in Chemistry) – Universidade Federal do Paraná, Curitiba, 2018.

MELO, T. O.; MARQUES, F. A.; WENDLING, I.; KOPKA, J.; ERBAN, A.; HANSEL, F. A. **Compostos presentes em extrato metanólico de tecido foliar de erva-mate, por meio da cromatografia gasosa acoplada à espectrometria de massas – Comunicado Técnico 458**. Colombo, PR: Embrapa, 2020.

MORAIS, E. C. de; STEFANUTO, A.; KLEIN, G. A.; BOAVENTURA, B. C. C.; ANDRADE, F. de; WAZLAWIK, E.; DI PIETRO, P. F.; MARASCHIN, M.; SILVA, E. L. da. Consumption of yerba mate (*Ilex paraguariensis*) improves serum lipid parameters in healthy dyslipidemic subjects and provides an additional LDL-cholesterol reduction in individuals on statin therapy. **Journal of Agricultural and Food Chemistry**, v. 57, p. 8316–8324, 2009. https://doi.org/10.1021/jf901660g.

MOSES, T.; PAPADOPOULOU, K. K.; OSBOURN, A. Metabolic and functional diversity of saponins, biosynthetic intermediates and semi-synthetic derivatives. **Critical Reviews in Biochemistry and Molecular Biology**, v. 49, n. 6, p. 439–462, 2014. https://doi.org/10.3109/10409238.2014.953628.

NAGATOMO, A.; INOUE, N.; KONNO, T.; XU, Y.; SAKAMOTO, C.; SONE, M.; SHIBASAKA, A.; MURAOKA, O.; NINOMIYA, K.; YOSHIKAWA, M.; MANSE, Y.; MORIKAWA, T. Ursane-type triterpene oligoglycosides with anti-hepatosteatosis and anti-hyperlipidemic activity from the leaves of *Ilex paraguariensis* A. St.-Hil. **Journal of Natural Medicines**, v. 76, n. 3, p. 654–669, 2022. https://doi.org/10.1007/s11418-022-01614-5.

NAKAMURA, K. L.; CARDOZO JUNIOR, L.; DONADUZZI, C. M.; SCHUSTER, I. Genetic variation of phytochemical compounds in progenies of *Ilex paraguariensis* St. Hil. **Crop Breeding and Applied Biotechnology**, v. 9, p. 116–123, 2009.

NEGRIN, A.; LONG, C.; MOTLEY, T. J.; KENNELLY, E. J. LC-MS metabolomics and chemotaxonomy of caffeine-containing holly (*Ilex*) species and related taxa in the Aquifoliaceae. **Journal of Agricultural and Food Chemistry**, v. 67, p. 5687–5699, 2019. https://doi.org/10.1021/acs.jafc.8b07168.

NIKALJE, G. C.; KUMAR, J.; NIKAM, T. D.; SUPRASANNA, P. FT-IR profiling reveals differential response of roots and leaves to salt stress in a halophyte *Sesuvium portulacastrum* (L.) L. **Biotechnology Reports**, v. 23, 2019. https://doi.org/10.1016/j.btre.2019.e00352.

OAKENFULL, D. Saponins in food – a review. **Food Chemistry**, v. 6, p. 19–19, 1981.

OLESZEK, W.; BIALY, Z. Chromatographic determination of plant saponins-An update (2002-2005). **Journal of Chromatography A**, v. 1112, n. 1–2, p. 78–91, 2006. https://doi.org/10.1016/j.chroma.2006.01.037.

OLIVEIRA, Y. M. M.; ROTTA, E. Área de distribuição natural de erva-mate (*Ilex paraguariensis* St. Hil.). 1983. **Seminário sobre atualidade e perspectivas florestais** [...]. Curitiba: Embrapa - CNPF, 1983. p. 17–36.

OSBOURN, A.; GOSS, R. J. M.; FIELD, R. A. The saponins-polar isoprenoids with important and diverse biological activities. **Natural Product Reports**, v. 28, n. 7, p. 1261–1268, 2011. https://doi.org/10.1039/c1np00015b.

PALUCH, E.; OKINCZYC, P.; ZWYRZYKOWSKA-WODZINSKA, A.; SZPERLIK, J.; ZAROWSKA, B.; DUDA-MADEJ, A.; BABELEWSKI, P.; WŁODARCZYK, M.; WOJTASIK, W.; KUPCZYNSKI, R.; SZUMNY, A. Composition and antimicrobial activity of *Ilex* leaves water extracts. **Molecules**, v. 26, n. 24, p. 1–35, 2021. https://doi.org/https://doi.org/10.3390/molecules26247442.

PANNEERSELVAM, K.; TSUKAMOTO, C.; HONDA, N.; KIKUCHI, A.; LEE, J. D.; YANG, S. H.; CHUNG, G. Saponin polymorphism in the Korean wild soybean (*Glycine soja* Sieb. and Zucc.). **Plant Breeding**, v. 132, n. 1, p. 121–126, 2013. https://doi.org/10.1111/pbr.12016.

PARIDA, A. K.; PANDA, A.; RANGANI, J. Metabolomics-guided elucidation of abiotic stress tolerance mechanisms in plants. **Plant Metabolites and Regulation Under Environmental Stress**. Amsterdam: Elsevier, 2018. p. 89–131. https://doi.org/10.1016/B978-0-12-812689-9.00005-4.

PEIRIS, K. H. S.; WU, X.; BEAN, S. R.; PEREZ-FAJARDO, M.; HAYES, C.; YERKA, M. K.; JAGADISH, S. V. K.; OSTMEYER, T.; ARAMOUNI, F. M.; TESSO, T.; PERUMAL, R.; ROONEY, W. L.; KENT, M. A.; BEAN, B. Near infrared spectroscopic evaluation of starch properties of diverse sorghum populations. **Processes**, v. 9, 2021. https://doi.org/10.3390/pr9111942.

PEIXOTO, P. M. G.; KAISER, S.; VERZA, S. G.; RESENDE, P. E. De; TRETER, J.; PAVEI, C.; BORRÉ, G. L.; ORTEGA, G. G. LC-UV assay method and UPLC/Q-TOF-MS characterisation of saponins from *Ilex paraguariensis* A. St. Hil. (Mate) unripe fruits. **Phytochemical Analysis**, 2011. https://doi.org/10.1002/pca.1374.

PHAM, H. N. T.; VUONG, Q. Van; BOWYER, M. C.; SCARLETT, C. J. Ultrasound-assisted extraction of *Catharanthus roseus* (L.) G. Don (Patricia White cultivar) stem for maximizing saponin yield and antioxidant capacity. **Journal of Food Processing and Preservation**, v. 42, n. 5, p. 1–12, 2018. https://doi.org/10.1111/jfpp.13597.

PIRES, V. S.; GUILLAUME, D.; GOSMANN, G.; SCHENKEL, E. P. Saponins from *Ilex dumosa*, an erva-maté (*Ilex paraguariensis*) adulterating plant. **Journal of Agricultural and Food Chemistry**, v. 45, p. 1027–1031, 1997.

POBLETE HERNÁNDEZ, P. **Productos Forestales No Madereros - Boletín nº 39**. Santiago: INFOR, 2022.

PRIYA, P.; PATIL, M.; PANDEY, P.; SINGH, A.; BABU, V. S.; SENTHIL-KUMAR, M. Stress combinations and their interactions in plants database: a one-stop resource on combined stress responses in plants. **The Plant Journal**, v. 116, n. 4, p. 1097–1117, 2023. https://doi.org/10.1111/tpj.16497.

PUANGPRAPHANT, S.; BERHOW, M. A.; DE MEJIA, E. G. Mate (*Ilex paraguariensis* St. Hilaire) saponins induce caspase-3-dependent apoptosis in human colon cancer cells in vitro. **Food Chemistry**, v. 125, n. 4, p. 1171–1178, 2011. https://doi.org/10.1016/j.foodchem.2010.10.023.

PUANGPRAPHANT, S.; DE MEJIA, E. G. Saponins in yerba mate tea (*Ilex paraguariensis* A. St.-Hil) and quercetin synergistically inhibit iNOS and COX-2 in lipopolysaccharide-induced macrophages through NFκB pathways. **Journal of Agricultural and Food Chemistry**, v. 57, n. 19, p. 8873–8883, 2009. https://doi.org/10.1021/jf902255h.

RAI, S.; ACHARYA-SIWAKOTI, E.; KAFLE, A.; DEVKOTA, H. P.; BHATTARAI, A. Plant-derived saponins: a review of their surfactant properties and applications. **Sci**, v. 3, n. 4, 1 dez. 2021. https://doi.org/10.3390/sci3040044.

RAKOCEVIC, M.; JANSSENS, M.; SCHERE, R. Light responses and gender issues in the domestication process of yerba-mate, a subtropical evergreen. *In*: BEZERRA, A. D.; FERREIRA, T. S. (Eds.). **Evergreens**. New York: Nova Science Publishers, 2012. v. 1, p. 63–95.

RAKOCEVIC, M.; MAIA, A. de H. N.; DE LIZ, M. V.; IMOSKI, R.; HELM, C. V.; CARDOZO JUNIOR, E. L.; WENDLING, I. Stability of leaf yerba mate (*Ilex paraguariensis*) metabolite concentrations over the time from the prism of secondary sexual dimorphism. **Plants**, v. 12, n. 11, 2023. https://doi.org/10.3390/plants12112199.

RAKOCEVIC, M.; MARTIM, S. F. Time series in analysis of yerba-mate biennial growth modified by environment. **International Journal of Biometeorology**, v. 55, p. 161–171, 2011. https://doi.org/10.1007/s00484-010-0322-4.

RAKOCEVIC, M.; MEDRADO, M. J. S.; LUCAMBIO, F.; VALDUGA, A. T. Intensity of bitterness of processed yerba mate leaves originated in two contrasted light environments. **Brazilian Archives of Biology and Technology**, v. 51, n. 3, p. 569–579, 2008. https://doi.org/10.1590/S1516-89132008000300018.

RANJHA, M. M. A. N.; IRFAN, S.; LORENZO, J. M.; SHAFIQUE, B.; KANWAL, R.; PATEIRO, M.; ARSHAD, R. N.; WANG, L.; NAYIK, G. A.; ROOBAB, U.; AADIL, R. M. Sonication, a potential technique for extraction of phytoconstituents: A systematic review. **Processes**, v. 9, n. 8, p. 1–21, 2021. https://doi.org/10.3390/pr9081406.

RÉBUFA, C.; DUPUY, N.; BOMBARDA, I. AComDim, a multivariate tool to highlighting impact of agroclimatic factors on *Moringa oleifera* Lam. leaf's composition from their FTIR-ATR profiles. **Vibrational Spectroscopy**, v. 116, 2021. https://doi.org/10.1016/j.vibspec.2021.103297.

REICHERT, C. L.; SALMINEN, H.; WEISS, J. *Quillaja* saponin characteristics and functional properties. **Annu. Rev. Food Sci. Technol**, v. 10, p. 43–73, 2019. https://doi.org/10.1146/annurev-food-032818.

REIS, M. S.; MONTAGNA, T.; MATTOS, A. G.; FILIPPON, S.; LADIO, A. H.; DA CUNHA MARQUES, A.; ZECHINI, A. A.; PERONI, N.; MANTOVANI, A. Domesticated landscapes in araucaria forests, southern Brazil: A multispecies local conservation-by-use system. **Frontiers in Ecology and Evolution**, v. 6, p. 1–14, 2018. https://doi.org/10.3389/fevo.2018.00011.

RESENDE, M. D. V. de. **Matemática e estatística na análise de experimentos e no melhoramento genético**. Colombo, PR: Embrapa Florestas, 2007.

RESENDE, M. D. V. de. Software Selegen-REML/BLUP: a useful tool for plant breeding. **Crop Breeding and Applied Biotechnology**, v. 16, p. 330–339, 2016. http://dx.doi.org/10.1590/1984-70332016v16n4a49.

RESENDE, M. D. V. de; DUARTE, J. B. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical**, v. 37, n. 3, p. 182–194, 2007.

SÁ, F. P. de. ***Ilex paraguariensis* A.St.-Hil: Miniestaquia, caracterização anatômica e bioquímica e estimativa do enraizamento por espectroscopia NIR**. 2018. 142 f. Thesis (PhD in Agronomy) – Universidade Federal do Paraná, Curitiba, 2018.

SALGOTRA, R. K.; CHAUHAN, B. S. Genetic diversity, conservation, and utilization of plant genetic resources. **Genes**, v. 14, n. 1, 2023. https://doi.org/10.3390/genes14010174.

SAMTIYA, M.; ALUKO, R. E.; DHEWA, T. Plant food anti-nutritional factors and their reduction strategies: an overview. **Food Production, Processing and Nutrition**, v. 2, n. 1, 2020. https://doi.org/10.1186/s43014-020-0020-5.

SARVIN, B.; STEKOLSHCHIKOVA, E.; RODIN, I.; STAVRIANIDI, A.; SHPIGUN, O. Optimization and comparison of different techniques for complete extraction of saponins from *T. terrestris*. **Journal of Applied Research on Medicinal and Aromatic Plants**, v. 8, p. 75–82, 2018. https://doi.org/10.1016/j.jarmap.2017.12.002.

SAVARINO, P.; DEMEYER, M.; COLSON, E. Mass spectrometry analysis of saponins. **Mass spectrometry reviews**, v. 40, p. 1–30, 2021. https://doi.org/10.1002/mas.21728.

SCHERER, R.; JANSSENS, M. J. J.; MARX, F.; URFER, P.; SCHNEIDER, E. Saponin content and quality-related traits of mass-selected yerba maté (*Ilex paraguariensis* A. St. Hil.) trees. **Journal of Herbs, Spices and Medicinal Plants**, v. 12, n. 1–2, p. 73–85, 2007. https://doi.org/10.1300/J044v12n01_07.

SHARMA, K.; KAUR, R.; KUMAR, S.; SAINI, R. K.; SHARMA, S.; PAWDE, S. V.; KUMAR, V. Saponins: A concise review on food related aspects, applications and health implications. **Food Chemistry Advances**, v. 2, p. 1–9, 2023. https://doi.org/10.1016/j.focha.2023.100191.

SILVA, M. A. F. da; HIGUCHI, P.; SILVA, A. C. da. Impacto de mudanças climáticas sobre a distribuição geográfica potencial de *Ilex paraguariensis*. **Rodriguésia**, v. 69, n. 4, p. 2069–2079, 2018. https://doi.org/10.1590/2175-7860201869437.

SKOLIK, P.; MCAINSH, M. R.; MARTIN, F. L. Biospectroscopy for plant and crop science. *In:* LOPES, J.; SOUSA, C. (Eds.) **Comprehensive Analytical Chemistry**. Amsterdam: Elsevier B.V., 2018. v. 80, p. 15–49. https://doi.org/10.1016/bs.coac.2018.03.001.

SOUZA, L. M. de; DARTORA, N.; SCOPARO, C. T.; CIPRIANI, T. R.; GORIN, P. A. J.; IACOMINI, M.; SASSAKI, G. L. Comprehensive analysis of maté (*Ilex paraguariensis*) compounds: Development of chemical strategies for matesaponin analysis by mass spectrometry. **Journal of Chromatography A**, v. 1218, n. 41, p. 7307–7315, 2011. https://doi.org/10.1016/j.chroma.2011.08.047.

SUN, D.-W. **Infrared Spectroscopy for Food Quality Analysis and Control**. Amsterdam: Elsevier, 2009.

SZAKIEL, A.; PĄCZKOWSKI, C.; HENRY, M. Influence of environmental abiotic factors on the content of saponins in plants. **Phytochemistry Reviews**, v. 10, n. 4, p. 471–491, 2011. https://doi.org/10.1007/s11101-010-9177-x.

TAKETA, A. T. C.; BREITMAIER, E.; SCHENKEL, E. P. Triterpenes and triterpenoidal glycosides from the fruits of *Ilex paraguariensis* (Maté). **Journal of the Brazilian Chemical Society**, v. 15, n. 2, p. 205–211, 2004. https://doi.org/https://dx.doi.org/10.1590/S0103-50532004000200008.

THOMAS, M. T.; KURUP, R.; JOHNSON, A. J.; CHANDRIKA, S. P.; MATHEW, P. J.; DAN, M.; BABY, S. Elite genotypes/chemotypes, with high contents of madecassoside and asiaticoside, from sixty accessions of *Centella asiatica* of south India and the Andaman Islands: For cultivation and utility in cosmetic and herbal drug applications. **Industrial Crops and Products**, v. 32, n. 3, p. 545–550, 2010. https://doi.org/10.1016/j.indcrop.2010.07.003.

TIMILSENA, Y. P.; PHOSANAM, A.; STOCKMANN, R. Perspectives on saponins: food functionality and applications. **International Journal of Molecular Sciences**, v. 24, n. 17, p. 1–22, 2023. https://doi.org/10.3390/ijms241713538.

TOMASI, J. de C.; AGUIAR, N. S.; DUARTE, M. M.; GABIRA, M. M.; VIEIRA, L. M.; PAULETTI, V.; FRANCISCON, L.; HELM, C. V.; DESCHAMPS, C.; WENDLING, I. Bioactive compound production in yerba mate clones with increasing nitrogen in semi-hydroponic system. **Journal of Soil Science and Plant Nutrition**, v. 24, n. 3, p. 5961–5971, 2024. https://doi.org/10.1007/s42729-024-01953-0.

TOMASI, J. de C.; LIMA, G. G. de; WENDLING, I.; HELM, C. V.; HANSEL, F. A.; GODOY, R. C. B. de; GRUNENNVALDT, R. L.; MELO, T. O. de; TOMAZZOLI, M. M.; DESCHAMPS, C. Effects of different drying methods on the chemical, nutritional and colour of yerba mate (*Ilex paraguariensis*) leaves. **International Journal of Food Engineering**, v. 17, n. 7, p. 551–560, 2021. https://doi.org/10.1515/ijfe-2020-0312.

TORMENA, C. D.; PAULI, E. D.; MARCHEAFAVE, G. G.; SCHEEL, G. L.; RAKOCEVIC, M.; BRUNS, R. E.; SCARMINIO, I. S. FT-IR biomarkers of sexual

dimorphism in yerba-mate plants: Seasonal and light accessibility effects. **Microchemical Journal**, v. 158, 2020. https://doi.org/10.1016/j.microc.2020.105329.

TÜRKER-KAYA, S.; HUCK, C. W. A review of mid-infrared and near-infrared imaging: Principles, concepts and applications in plant tissue analysis. **Molecules**, v. 22, n. 1, 2017. https://doi.org/10.3390/molecules22010168.

VALDUGA, A. T.; GONÇALVES, I. L.; MAGRI, E.; DELALIBERA FINZER, J. R. Chemistry, pharmacology and new trends in traditional functional and medicinal beverages. **Food Research International**, v. 120, p. 478–503, 2019. https://doi.org/10.1016/j.foodres.2018.10.091.

VIEIRA, L. M.; MAGGIONI, R. D. A.; TOMASI, J. D. C.; NUNES GOMES, E.; WENDLING, I.; HELM, C. V.; KOEHLER, H. S.; ZUFFELLATO-RIBAS, K. C. Vegetative propagation, chemical composition and antioxidant activity of yerba mate genotypes. **Plant Genetic Resources: Characterisation and Utilisation**, v. 19, p. 112–121, 2021. https://doi.org/10.1017/S1479262121000150.

WENDLING, I.; SANTIN, D.; NAGAOKA, R.; STURION, J. A. **BRS BLD Aupaba e BRS BLD Yari: cultivares clonais de erva-mate para produção de massa foliar de sabor suave - Comunicado Técnico 411**. Colombo, PR: Embrapa Florestas, 2017b.

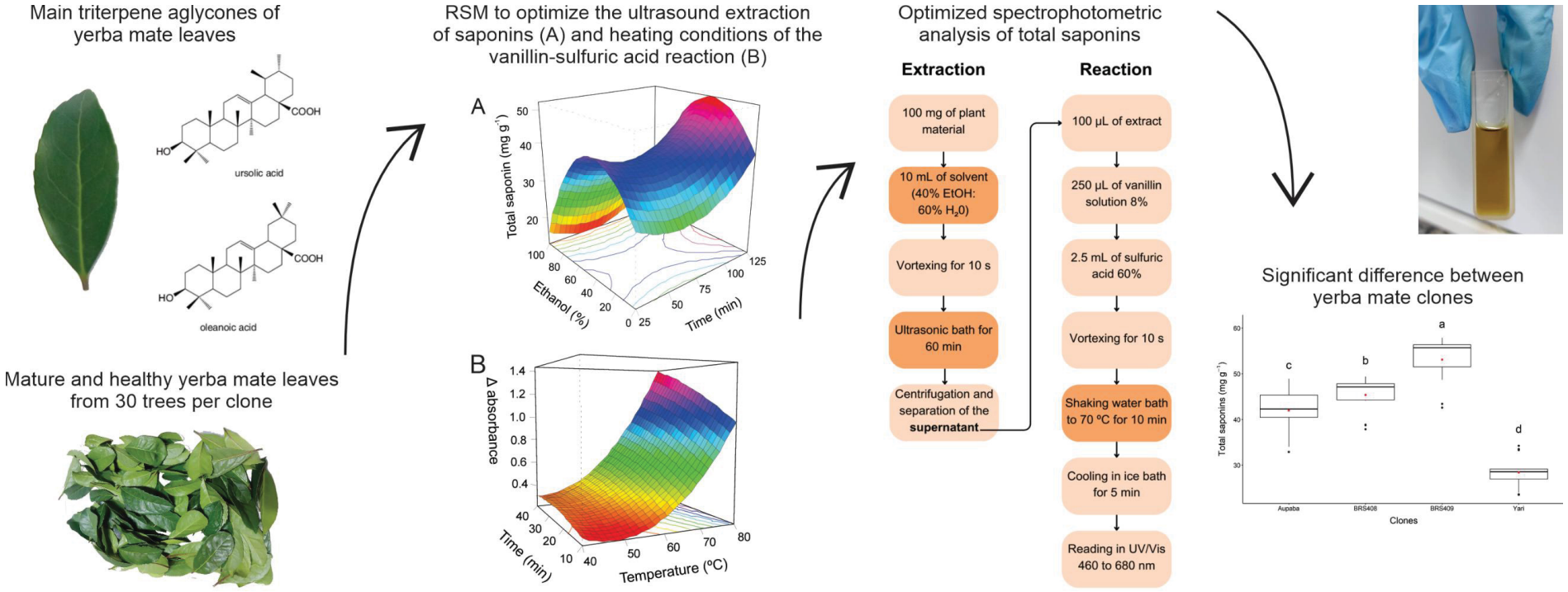
WENDLING, I.; STURION, J. A.; SANTIN, D. **BRS 408 e BRS 409: cultivares clonais de erva-mate para produção de massa foliar - Comunicado Técnico 410**. Colombo, PR: Embrapa Florestas, 2017a.

WEYER, L. G.; LO, S. C. Spectra-Structure Correlations in the Near-infrared. *In*: CHALMERS, J. M.; GRIFFITHS, P. R. (Eds.). **Handbook of Vibrational Spectroscopy**. New York: John Wiley & Sons, 2006. p. 1817–1837. https://doi.org/10.1002/0470027320.

YANG, L.; WEN, K. S.; RUAN, X.; ZHAO, Y. X.; WEI, F.; WANG, Q. Response of plant secondary metabolites to environmental factors. **Molecules**, v. 23, n. 4, 2018. https://doi.org/10.3390/molecules23040762.

APPENDICES

Appendix A – Graphical abstract of Chapter 1



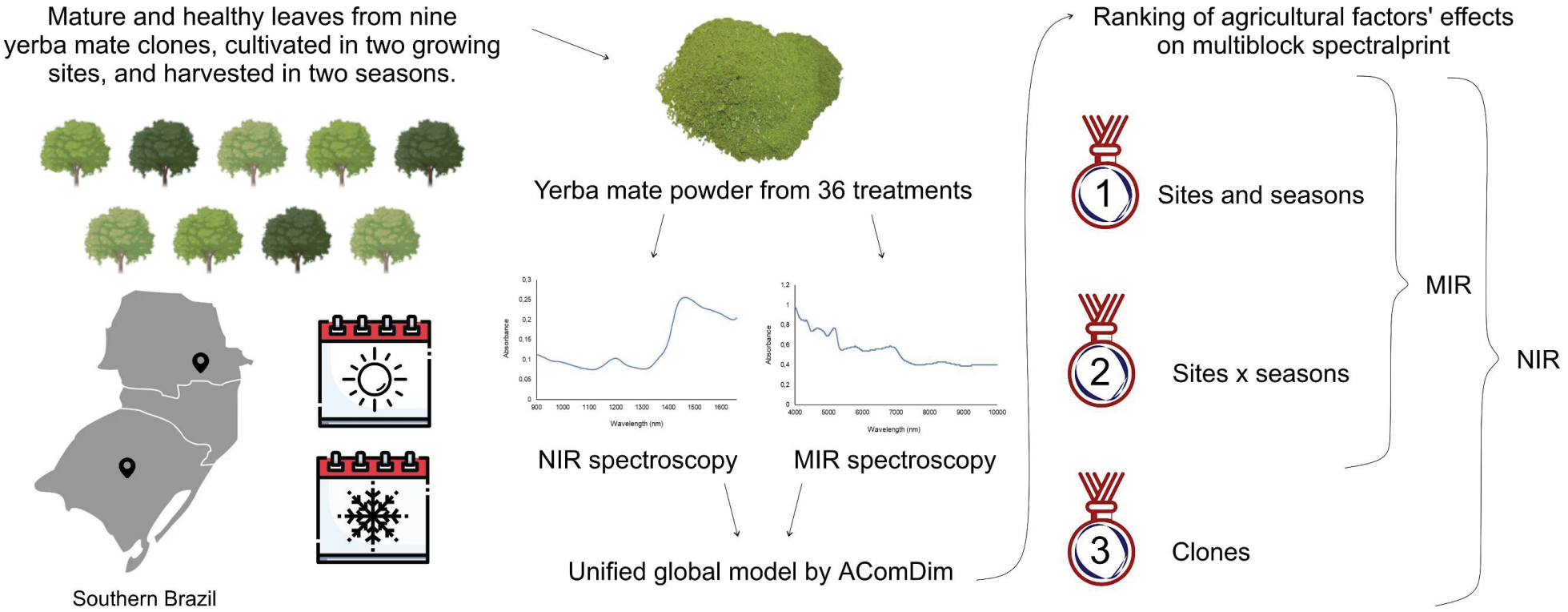
Appendix B – Environmental conditions of the two clonal tests of yerba mate, in Espumoso, Rio Grande do Sul (RS), and São Mateus do Sul, Paraná (PR), Brazil

	Site	
General information	Espumoso, RS	São Mateus do Sul, PR
Geographical coordinates	-28.87 S, -52.86 W	-25.93 S, -50.29 W
Altitude (m)	430	815
Planting date	September/2020	July/2020
Planting spacing (m)	1.5 x 3.0	1.5 x 3.0
Shading (%)	0 (Full sunlight)	70 (Shaded)
Closest meteorological station (INMET)	A837	A874
Geographical coordinates of the meteorological station	-28.86 S, -52.54 W	-25.84 S, -50.37 W
Distance between the cultivation site and the meteorological station (km)	30.4	12.9
Weather		
Average annual temperature (ºC)	17.15	16.70
Annual maximum absolute temperature (ºC)	37.50	34.50
Annual minimum absolute temperature (ºC)	-1.30	-2.30
Annual precipitation (mm)	1,652.60	1,308.50
Average daily solar radiation (kJ/m²)	17,755.77	12,442.53
Average annual relative humidity (%)	75.84	77.45
Weather 30 days before winter harvest (June and July/2022)		
Average temperature (ºC)	15.74	12.90
Absolute maximum temperature (ºC)	27.20	25.70
Absolute minimum temperature (ºC)	2.00	0.40
Precipitation (mm)	48.60	14.20
Average daily solar radiation (kJ/m²)	9,922.05	7,444.85
Average relative humidity (%)	76.83	80.45
Weather 30 days before the summer harvest (December and January/2023)		
Average temperature (ºC)	22.72	20.23

Absolute maximum temperature (ºC)	34.30	31.40
Absolute minimum temperature (ºC)	13.00	8.60
Precipitation (mm)	39.40	233.80
Average daily solar radiation (kJ/m²)	24,492.08	17,642.69
Average relative humidity (%)	65.64	69.98
Soil		
Topography	Gently undulating	Gently undulating
pH in H_2O	5.00	4.68
Clay content (%)	46.00	62.00
Organic matter content (%)	3.00	3.67
Phosphorus (P) content (mg dm^{-3})	2.70	0.66
Potassium (K) content (mg dm^{-3})	82.50	54.74
Exchangeable calcium (Ca) ($cmol_c$ dm^{-3})	3.60	7.18
Exchangeable magnesium (Mg) ($cmol_c$ dm^{-3})	1.50	1.25
Exchangeable aluminum (Al) ($cmol_c$ dm^{-3})	0.80	1.36
CEC* pH 7.0 ($cmol_c$ dm^{-3})	15.00	15.54

*CEC: Cation Exchange Capacity

Appendix C – Graphical abstract of Chapter 3



Appendix D – Individual *deviance* analyses for total saponin content (mg g^{-1} on dry basis – equivalent to oleanolic acid) of yerba mate clones evaluated in winter and summer harvests, in the clonal tests of Espumoso, RS, and São Mateus do Sul, PR (Brazil), corresponding to full sunlight and shaded systems, respectively

	Espumoso, RS (Full Sunlight) Winter		**Espumoso, RS (Full Sunlight) Summer**		**São Mateus do Sul, PR (Shaded) Winter**		**São Mateus do Sul, PR (Shaded) Summer**	
	Deviance	LRT[1]	Deviance	LRT	Deviance	LRT	Deviance	LRT
Clones[+]	217.14	23.28*	196.48	52.61*	217.89	37.61*	217.89	39.10*
Full Model	193.86	-	143.87	-	180.28	-	178.79	-
Individual REML								
V_g	53.92		41.59		58.23		85.18	
V_r	28.03		5.78		15.42		17.53	
V_p	81.95		47.38		73.65		102.71	
h^2_i	0.66 ± 0.35		0.88 ± 0.40		0.79 ± 0.38		0.83 ± 0.40	
Acclon	0.95		0.98		0.97		0.98	
CV_g (%)	16.60		16.98		16.38		23.72	
CV_r (%)	11.96		6.33		8.43		10.76	
CV_{rel}	1.39		2.68		1.94		2.20	

[1]Likelihood Ratio Test; [+] Deviance of fitted model without the mentioned effects; *Significant at 1% level of error probability, by Chi-square test with 1 degree of freedom; V_g: genotypic variance among clones; V_r: residual variance; V_p: individual phenotypic variance; h^2_i: individual plots broad-sense heritability; Acclon: selection accuracy of clones; CV_g%: genotypic coefficient of variation; CV_r%: residual coefficient of variation; $CV_{rel} = CV_g/CV_r$ relative coefficient of variation

Appendix E – Spectral preprocessing selection in mid-infrared (MIR) spectralprint using factorial design and ANOVA – Simultaneous Component Analysis (ASCA) models

Factors	–	0	+
1. Normalization	SNV	Without	MSC
2. Movin average	3	7	11
3. Derivation	1st	Without	2nd

	Factor level			ASCA results							
Factorial experiment	1	2	3	Growing locations (%)	Harvest seasons (%)	Yerba mate clones (%)	F1 × F2	F1 × F3	F2 × F3	F1 × F2 × F3	Residuals (%)
1	–	–	–	0.40	0.65	1.16	0.20	1.04	1.02	1.03	94.50
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	
2	–	–	0	24.31	29.35	3.54	11.73	4.31	2.35	2.85	21.42
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.*0000)	*(1.0000)*	*(0.0001)*	
3	–	–	+	0.12	0.13	1.08	0.11	0.95	1.00	0.93	95.68
				(0.2389)	*(0.0723)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	
4	–	0	–	1.66	3.03	1.69	0.54	1.31	1.10	1.27	89.38
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	
5	–	0	0	24.80	29.92	3.59	11.97	4.37	2.38	2.88	19.95
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
6	–	0	+	0.15	0.18	1.09	0.13	0.98	0.98	0.99	95.51
				(0.0064)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	
7	–	+	–	3.39	6.38	2.33	0.99	1.69	1.29	1.52	82.37
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	
8	–	+	0	25.07	30.21	3.60	12.10	4.40	2.39	2.90	19.17
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
9	–	+	+	0.25	0.35	1.17	0.18	1.00	0.94	1.13	94.97
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	
10	0	–	–	0.40	1.06	1.20	0.21	1.12	1.05	1.14	93.81
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	
11	0	–	0	7.47	42.12	4.81	5.36	4.58	4.30	7.66	23.88
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
12	0	–	+	0.12	0.13	1.09	0.12	0.95	1.01	0.94	95.65
				(0.2163)	*(0.0623)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	
13	0	0	–	1.57	5.16	1.88	0.57	1.66	1.25	1.86	86.00
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	

14	0	0	0	7.48	42.17	4.82	5.36	4.59	4.31	7.67	23.80
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
15	0	0	+	0.17	0.17	1.09	0.14	0.99	0.99	1.01	95.45
				(0.0001)	*(0.0009)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	
16	0	+	–	2.95	10.58	2.67	1.00	2.36	1.58	2.69	76.06
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
17	0	+	0	7.48	42.19	4.82	5.37	4.59	4.31	7.68	23.75
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
18	0	+	+	0.34	0.31	1.18	0.20	1.04	0.95	1.17	94.79
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	
19	+	–	–	0.40	0.65	1.16	0.20	1.04	1.02	1.03	94.50
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	
20	+	–	0	24.31	29.35	3.54	11.74	4.30	2.35	2.85	21.42
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
21	+	–	+	0.12	0.13	1.08	0.11	0.95	1.00	0.93	95.68
				(0.2263)	*(0.0728)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	
22	+	0	–	1.67	3.02	1.69	0.54	1.31	1.10	1.27	89.38
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	
23	+	0	0	24.79	29.92	3.58	11.97	4.37	2.38	2.88	19.95
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
24	+	0	+	0.15	0.18	1.09	0.13	0.98	0.98	0.99	95.51
				(0.0068)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	
25	+	+	–	3.40	6.36	2.34	0.99	1.69	1.29	1.52	82.36
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	
26	+	+	0	25.06	30.21	3.60	12.11	4.40	2.39	2.90	19.17
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
27	+	+	+	0.25	0.35	1.17	0.18	1.00	0.94	1.13	94.97
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	

Appendix F – Spectral preprocessing selection in near-infrared (NIR) spectralprint using factorial design and ANOVA – Simultaneous Component Analysis (ASCA) models

Factors	–	0	+
1. Normalization	SNV	Without	MSC
2. Movin average	3	7	11
3. Derivation	1st	Without	2nd

	Factor level			ASCA results							
Factorial experiment	1	2	3	Growing locations (%)	Harvest seasons (%)	Yerba mate clones (%)	F1 × F2	F1 × F3	F2 × F3	F1 × F2 × F3	Residuals (%)
1	–	–	–	25.83	6.04	16.86	11.91	12.46	8.60	6.73	11.56
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
2	–	–	0	64.56	5.74	3.97	2.12	4.04	2.81	2.55	14.21
				(0.0001)	*(0.0001)*	*(0.0002)*	*(1.0000)*	*(1.*0000)	*(1.0000)*	*(0.6190)*	
3	–	–	+	16.93	6.91	13.80	9.63	14.16	10.98	10.13	17.46
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
4	–	0	–	29.49	5.86	16.25	12.06	10.77	8.44	6.71	10.43
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
5	–	0	0	65.50	5.65	3.68	1.95	3.84	2.67	2.44	14.26
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.6751)*	
6	–	0	+	19.52	2.61	19.67	16.07	14.08	9.85	7.68	10.52
				(0.0001)	*(0.0031)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
7	–	+	–	27.94	6.48	16.55	11.55	11.30	8.53	6.90	10.74
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
8	–	+	0	66.20	5.40	3.54	1.86	3.74	2.60	2.37	14.30
				(0.0001)	*(0.0001)*	*(0.0009)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.7180)*	
9	–	+	+	15.96	3.79	20.92	16.06	14.58	10.14	7.83	10.71
				(0.0001)	*(0.0007)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
10	0	–	–	18.15	14.80	15.46	11.75	11.76	8.58	8.23	11.27
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
11	0	–	0	3.98	5.69	6.49	32.70	11.90	9.85	12.32	17.07
				(0.0035)	*(0.0004)*	*(0.1130)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
12	0	–	+	10.47	10.14	14.23	10.93	14.34	10.93	11.28	17.68
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
13	0	0	–	21.24	15.06	14.73	11.78	10.14	8.47	8.32	10.25
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	

14	0	0	0	3.97	5.68	6.48	32.71	11.91	9.86	12.33	17.08
				(0.0038)	*(0.0006)*	*(0.1108)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
15	0	0	+	16.46	4.02	18.01	17.92	14.34	9.71	9.93	9.62
				(0.0001)	*(0.0002)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
16	0	+	–	19.15	17.14	14.98	10.55	10.55	8.74	8.22	10.68
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
17	0	+	0	3.94	5.65	6.48	32.73	11.92	9.86	12.34	17.08
				(0.0037)	*(0.0007)*	*(0.1172)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
18	0	+	+	14.26	6.58	18.49	16.78	14.31	10.04	9.67	9.87
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
19	+	–	–	26.02	6.09	16.78	11.81	12.42	8.55	6.65	11.68
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
20	+	–	0	64.52	5.75	3.96	2.12	4.05	2.81	2.54	14.26
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.6229)*	
21	+	–	+	17.07	6.91	13.79	9.56	14.15	10.94	10.08	17.50
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
22	+	0	–	29.66	5.91	16.17	11.95	10.74	8.38	6.62	10.56
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
23	+	0	0	65.47	5.66	3.67	1.95	3.85	2.67	2.43	14.30
				(0.0001)	*(0.0001)*	*(0.0002)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.6830)*	
24	+	0	+	19.64	2.63	19.67	15.93	14.09	9.85	7.63	10.55
				(0.0001)	*(0.0026)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
25	+	+	–	28.12	6.54	16.45	11.48	11.25	8.46	6.82	10.88
				(0.0001)	*(0.0001)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	
26	+	+	0	66.17	5.40	3.52	1.86	3.75	2.60	2.37	14.33
				(0.0001)	*(0.0001)*	*(0.0010)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.7165)*	
27	+	+	+	16.07	3.82	20.92	15.94	14.58	10.14	7.79	10.74
				(0.0001)	*(0.0004)*	*(0.0001)*	*(1.0000)*	*(1.0000)*	*(1.0000)*	*(0.0001)*	