UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Geociências
Programa de Pós-Graduação em Análise e Modelagem de Sistemas Ambientais

Pedro Fialho Cordeiro

SPATIALLY EXTENSIVE BIOLOGICAL ASSESSMENTS OF RIVERS IN MINAS GERAIS USING PREDICTIVE MODELING AND RIVER TYPOLOGY APPROACHES

Belo Horizonte
2023

Pedro Fialho Cordeiro

SPATIALLY EXTENSIVE BIOLOGICAL ASSESSMENTS OF RIVERS IN MINAS GERAIS USING PREDICTIVE MODELING AND RIVER TYPOLOGY APPROACHES

Tese apresentada ao Programa de Pós-Graduação em Análise e Modelagem de Sistemas Ambientais da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Doutor em Análise e Modelagem de Sistemas Ambientais.

Orientador: Prof. Dr. Diego Rodrigues Macedo.

Coorientadora: Dra. Maria João Feio.

Belo Horizonte
2023

C794s Cordeiro, Pedro Fialho.
2023 Spatially extensive biological assessments of rivers in Minas Gerais using predictive modeling and river typology approaches [manuscrito] / Pedro Fialho Cordeiro. – 2023.
93 f., enc.: il. (principalmente color.)

Orientador: Diego Rodrigues Macedo.
Coorientadora: Maria João Feio.
Tese (doutorado) – Universidade Federal de Minas Gerais, Instituto de Geociências, 2022.
Área de concentração: Análise, Modelagem e Gestão de Sistemas Ambientais.
Inclui bibliografia.
Inclui anexos e apêndices.

1. Modelagem de dados – Aspectos ambientais – Teses. 2. Recursos hídricos – Desenvolvimento – Minas Gerais – Teses. 3. Ecologia dos rios – Teses. I. Macedo, Diego Rodrigues. II. Feio, Maria João. III. Universidade Federal de Minas Gerais. Instituto de Geociências. IV. Título.

CDU: 911.2:519.6(815.1)

Ficha catalográfica elaborada por Graciane Aparecida de Paula – CRB-6/3404
Biblioteca "Vitória Pedersoli", Instituto de Geociências da UFMG

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ANÁLISE E MODELAGEM DE SISTEMAS AMBIENTAIS

FOLHA DE APROVAÇÃO

SPATIALLY EXTENSIVE BIOLOGICAL ASSESSMENTS OF RIVERS IN MINAS GERAIS USING PREDICTIVE MODELING AND RIVER TYPOLOGY APPROACHES

PEDRO FIALHO CORDEIRO

Tese de doutorado submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ANÁLISE E MODELAGEM DE SISTEMAS AMBIENTAIS, como requisito para obtenção do grau de Doutor em ANÁLISE E MODELAGEM DE SISTEMAS AMBIENTAIS, área de concentração ANÁLISE, MODELAGEM E GESTÃO DE SISTEMAS AMBIENTAIS.

Diego Rodrigues Macedo - UFMG - (Doutor)
Maria Clara Vieira Martins Starling - UFMG - (Doutora)
Raphael Ligeiro Barroso Santos - Universidade Federal do Pará - (Doutor)
Úrsula Ruchkys de Azevedo - UFMG - (Doutora)
Luis Felipe Soares Cherem - Universidade Federal de Minas Gerais - (Doutor)

Belo Horizonte, 30 de junho de 2023.

Documento assinado eletronicamente por **Raphael Ligeiro Barroso Santos**, **Usuário Externo**, em 06/07/2023, às 12:52, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Diego Rodrigues Macedo**, **Professor do Magistério Superior**, em 10/07/2023, às 13:04, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Maria Clara Vieira Martins Starling**, **Professora do Magistério Superior**, em 10/07/2023, às 14:31, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Luis Felipe Soares Cherem**, **Professor do Magistério Superior**, em 12/07/2023, às 11:08, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Ursula Ruchkys de Azevedo**, **Chefe de departamento**, em 03/08/2023, às 15:29, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2446665** e o código CRC **12DD3D23**.

Dedico essa tese à minha esposa Andreza e aos meus pais Cerina e Albérico.

AGRADECIMENTOS

Chegar até aqui é motivo de muito orgulho e gratidão. Para conciliar a rotina de trabalho com o doutorado, tive que abdicar de muitos momentos com pessoas especiais. Por isso, o meu principal agradecimento é pela compreensão de todos devido à minha ausência em diversas ocasiões.

Agradeço à minha esposa Andreza pela sua dedicação e compreensão admiráveis. Embora tivemos de abrir mão de muitos momentos, o seu apoio foi essencial para a conclusão desta pesquisa.

Ao meu orientador, Professor Dr. Diego Macedo, agradeço por todo apoio e incentivo. Você soube me orientar muito bem em todas as etapas, e me incentivou a ampliar o escopo do projeto, incorporando diversos aspectos da biologia, agregando conhecimento à minha formação. Sempre se esforçou em envolver pessoas de credibilidade na pesquisa, e sou muito grato por tudo isso.

À Dra. Maria João, agradeço pela coorientação e pelos feedbacks construtivos, além da paciência. Eu não imaginaria poder contar com o apoio de uma pessoa no outro hemisfério do planeta e referência mundial no tema da pesquisa.

Ao Professor Dr. Marcos Callisto, obrigado pelo apoio e por ter contribuído em todas as etapas do projeto, estando presente durante importantes tomadas de decisão.

Agradeço ao Diego Pujoni pela pronta disponibilidade em me ajudar com o primeiro capítulo da tese, além da estatística e os gráficos maravilhosos feitos no R!

Agradeço também aos meus pais e avós, que apesar da distância continuaram me apoiando durante todo o tempo.

Ao amigo Feliciano, por estar comigo durante toda essa jornada, ajudando-me a resolver questões complexas no SENAI. Sem você seria muito mais difícil conciliar o trabalho com o doutorado.

Ao amigo Samuel Castro que, além de um grande incentivador e conselheiro, durante uma carona para Barbacena me inspirou a dar o primeiro passo desta caminhada.

Ao Bob Hughes, muito obrigado pelas críticas e sugestões, você contribuiu para melhorar esta pesquisa.

Ao Marden e Anderson, e à Luisa e Isadora, agradeço pela ajuda com o banco de dados!

À Mônica Mamão, Marina, Rafael Arcanjo, Carolina, Aylton, Jonathas, Paulo, Mara e Karina, por estarem comigo durante este tempo!

Aos amigos Marcus Manoel, Márcia Melo, Mônica Campos, Sylvia, Marcelo Ávila, Helena e Valéria, vocês possuem grande importância, cada um em uma fase distinta, na minha formação de Pesquisador.

Aos Gerentes e Coordenadores do Centro de Inovação e Tecnologia (CIT) SENAI, pela autorização para a realização do doutorado.

Aos colegas do CIT SENAI e ao Programa de Pós-Graduação em Análise e Modelagem de Sistemas Ambientais, especialmente as secretárias Cora e Isis, e à Profa. Dra. Sônia, que sempre estiveram disponíveis para ajudar.

ABSTRACT

Freshwater ecosystems are among the most threatened by human pressures worldwide. Flow regulation, sedimentation, habitat degradation, invasion of alien species, and poor sewage and waste treatment are the main causes of biodiversity and habitat loss. Given this scenario, ecological status assessment is critical for addressing efficient management practices. In Brazil, assessing the ecological status of aquatic environments is still incipient due to financial and logistical aspects, the lack of trained researchers, and specific legislation. Alternatively, using the experience of the Global-North countries could be extremely useful in large tropical countries such as Brazil to improve freshwater monitoring programs and management. In this context, this Thesis aimed to adapt and validate methodologies for river ecological status assessment in Minas Gerais. Tools such as the European river typology and RIVPACS-type predictive models for ecological status assessment were used, taking advantage of a large biological database gathered over 16 years and several research projects developed along Minas Gerais state. River typology and predictive models based on the assemblage of benthic macroinvertebrates are also helpful for ecological status assessment in Minas Gerais. Besides, the river typology reduces the natural variability and can improve the predictive model performance, reducing the probability of inferring impairment when it does not exist or even not detecting it when it exists. Finally, the tools developed in this Thesis can be used for further development of monitoring programs and management strategies in Brazil, and encourage discussion with the National Water Resources Agency on the importance of biomonitoring development nationally.

Keywords: abiotic typology; benthic macroinvertebrates; water management.

RESUMO

Os ecossistemas de água doce estão entre os mais ameaçados pelas pressões antrópicas em todo o mundo. Alterações no regime de escoamento, sedimentação, espécies invasoras e a ausência de saneamento básico são as principais causas da perda de hábitats e biodiversidade. Dado este cenário, avaliar as condições ecológicas desses ecossistemas é crucial para uma gestão eficiente dos recursos hídricos nacionais. No Brasil, a avaliação do estado ecológico de ambientes aquáticos ainda é incipiente devido a aspectos financeiros e logísticos, à falta de mão de obra capacitada e legislação específica. Dessa forma, modelos de gerenciamento de recursos hídricos de países do Norte-Global, alternativamente, podem ser adaptados para o Brasil e contribuir para o aprimoramento da gestão dos recursos hídricos. Neste contexto, o objetivo desta tese foi adaptar e validar metodologias para a avaliação do estado ecológico de cursos de água de Minas Gerais. Ferramentas como a tipificação fluvial Europeia e os modelos preditivos de qualidade ecológica do tipo RIVPACS foram adaptados, utilizando-se um banco de dados biológico compilado ao longo de 16 anos de projetos desenvolvidos em Minas Gerais. Os resultados demonstram que a tipificação e os modelos preditivos baseados na assembleia de macroinvertebrados bentônicos também são úteis em Minas Gerais para a avaliação do estado ecológico de cursos de água. Foi demonstrado ainda que a tipificação possui potencial para tornar os modelos preditivos menos susceptíveis à variação ambiental natural, o que reduziria a probabilidade de inferir uma perda de qualidade ecológica quando ela não existe ou mesmo não a detectar quando existente. Por fim, as ferramentas desenvolvidas nesta tese podem ser utilizadas para o desenvolvimento futuro de programas de monitoramento e estratégias de gestão no Brasil, bem como encorajar a discussão com a Agência Nacional de Águas – ANA sobre a importância de desenvolver o biomonitoramento nacionalmente.

Palavras-chave: tipologia abiótica; macroinvertebrados bentônicos; gestão de recursos hídricos.

TABLE OF CONTENTS

1. INTRODUCTION

In the 21st century, several countries started a new chapter in environmental quality assessment, incorporating a more comprehensive vision focused on ecological status classification rather than traditional monitoring based on physical and chemical parameters (BUSS et al., 2015; FEIO et al., 2021; PARDO et al., 2012; SANTOS et al., 2021). Recently, an increased focus has been on analyzing the connection between human actions and ecological health by studying how water-related processes occur, such as the interaction between geomorphology, chemistry, and biology aspects. This comprehensive approach aids in comprehending the overall environmental impact of human activity (BROWN; WILLIAMS, 2016; PAULSEN et al., 2008). These assessments include abiotic and biotic attributes that indicate ecological changes caused by human interference compared to reference conditions or least disturbed areas (DAVY-BOWKER et al., 2006; REYNOLDSON et al., 1997).

In Brazil, promoting environmental sustainability requires prioritizing scientific research and carefully crafted public policies that can effectively assess the health of its aquatic ecosystems. However, several obstacles hinder the implementation of these measures, such as financial and logistical constraints, a shortage of skilled workers, and insufficient legislation. Addressing these challenges will ensure Brazil's natural resources' long-term health and stability (JUNQUEIRA; FRIEDRICH; PEREIRA DE ARAUJO, 2010; MACEDO et al., 2016; MELLO et al., 2020). To improve its water resource management, Brazil can consider adopting the models used in the North-Global countries, such as the European Union, United States, Japan, and South Korea (BAPTISTA, 2008; CARDOSO-SILVA; FERREIRA; POMPÊO, 2013; FEIO et al., 2021).

In Europe, The Water Framework Directive (WFD) (nº 2000/60/EC) is a comprehensive strategy adopted by member states, as well as Norway and the UK, to promote the preservation and responsible use of water resources. This framework emphasizes the importance of incorporating ecological principles into policies and regulations at the national level (MELLO et al., 2023). The WFD is a legal framework that promotes research and studies to establish reference conditions and classify aquatic environments. It also encourages the use of predictive models to assess the ecological status of aquatic ecosystems (CARDOSO-SILVA; FERREIRA; POMPÊO, 2013; DAVY-BOWKER et al., 2006; FEIO et al., 2007; SOLHEIM et al., 2019).

Over the past 20 years, the WFD has shifted how water resources are managed by adopting an ecocentric perspective, becoming the primary instrument of the European Union's water policy. In this perspective, instead of humans being the center of focus, water is now viewed as the owner of an ecosystem (CARDOSO-SILVA; FERREIRA; POMPÊO, 2013). Thus, ecological status was established as a new concept and the basis for management decisions related to water quality (SANTOS et al., 2021). The WFD established the following steps for the assessment of ecological status (DQA, 2000): i) characterization of surface waters; ii) establishment of the aquatic environments typology; iii) establishment of monitoring programs; iv) definition of specific reference conditions for each type of water body for biological quality elements; v) classification of all surface water bodies using an Ecological Quality Ratio; and, vi) inter-calibration (IC).

Pardo et al. (2012) suggest that the ecological status assessment should measure how much the biota deviates from minimally disturbed conditions. Following the Annex V of the WFD (AROVIITA et al., 2008; DQA, 2000), the ecological classification system should use an ecological quality ratio (EQR) to compare the conditions observed at a site to the conditions expected at a reference site with minimal impact. The reference conditions mean the absence of disturbance or minimal changes caused by human pressures on the environment and represent an aim for remediation or ecological restoration (STODDARD et al., 2006). Early efforts to establish reference conditions involved identifying specific biological characteristics defining an undisturbed state (BARBOUR et al., 1999; WRIGHT et al., 1984). These studies have widely developed the concept of reference conditions (HAWKINS; OLSON; HILL, 2010; HUGHES; LARSEN; OMERNIK, 1986; PAULSEN et al., 2008; REYNOLDSON et al., 1997; WRIGHT, 1995).

On the other hand, establishing reference conditions is practically impossible since no place on Earth can be considered preserved and has already been exposed to vegetation losses, climate change, deposition of nutrients, and toxic substances (CHESSMAN, 2021).

A critical area of concern is the impact of human activities on large rivers worldwide, for which reference sites are rarely found (GRILL et al., 2019). The unclear understanding of what minimally disturbed conditions represent can lead to the selection of reference sites with some anthropic impact, making comparisons difficult. Stoddard et al. (2006) highlight the multiple meanings that “reference condition” has in a variety of contexts, such as ‘‘minimally disturbed condition'' (MDC); ''historical condition'' (HC); ''least disturbed condition'' (LDC); and ''best attainable condition'' (BAC). The authors also argue that these concepts can be narrowly defined, and each implies specific methods for estimating expectations. Consistency in using terms related to the reference-condition concept is crucial as it significantly affects the outcome of biological assessments. Therefore, to avoid confusing discussions among scientists and managers, the definition of reference conditions should cover, whenever possible, physical, chemical, ecological, and ecotoxicological aspects and specialists' opinions throughout the entire process (PARDO et al., 2012).

To ensure accurate comparisons between biota in different areas, evaluating environments with similar abiotic characteristics is essential because this helps to reduce the natural variability of the biocenosis and provides a baseline for what would be expected in an undisturbed area (LORENZ; FELD; HERING, 2004; SOLHEIM et al., 2019). Based on this idea, the WFD established the minimum requirements, presented in Annex II of the document and structured in two Systems (A and B) to classify water bodies. System A's typology is based on mandatory factors, such as the hydrologic units defined by Abell et al. (2008), the size of the catchments, geology, and altitude. In system B, optional parameters were added to obligatory system A factors (DQA, 2000).

Following systems A or B, several European countries have developed their typologies (BORGWARDT et al., 2019; LORENZ; FELD; HERING, 2004; NOBLE; COWX, 2002) based on characteristics that are not altered by human intervention (e.g., geology, altitude, slope, temperature, and precipitation) (BORGWARDT et al., 2019; LORENZ; FELD; HERING, 2004; SOLHEIM et al., 2019). The typology must reflect the consistency of the biological groups of each community to be effective (FEIO; PINTO, 2009a). According to the WFD, reference conditions are linked to stream typologies, and the population of reference sites should represent the range of conditions expected to occur naturally within the stream type (STODDARD et al., 2006). This way, it is possible to critically analyze the losses of ecological integrity and possible implications for ecosystem functioning (DAVY-BOWKER et al., 2006; FEIO; PINTO, 2009a; NOBLE; COWX, 2002).

In the United States, the policy foundation for USA lotic ecosystem monitoring and assessment is the Federal Water Pollution Control Act (HAWKINS, 2006). The Clean Water Act - CWA led to the establishment of National Programs (e.g., National Rivers And Streams

Assessment), which aimed to use standard protocols to suit biological, physical, and chemical condition indicators. Thus, the raw physical and biological data are converted into metrics and indices (MMI, observed/expected (O/E) models) for reporting at state, ecoregional, and national spatial extents (STODDARD et al., 2006; MELLO et al., 2023). In the European Union (EU), the WFD established ecological assessment programs in its 27 member states. The assessment methods had to go through an Intercalibration Exercise (IC) to guarantee the comparability of classifications among countries. The monitoring networks in the EU are meant to provide complete and organized information on the ecological condition of all waterbodies in every river basin district. These data is crucial for developing River Basin Management Plans (FEIO et al., 2021).

In the Brazilian context, water quality monitoring primarily relies on physical, chemical, and bacteriological parameters, with the option of biomonitoring (BRASIL, 2005), and does not account for natural variability. Nevertheless, these parameters alone might not offer a comprehensive view of the ecological status of aquatic ecosystems (CALLISTO et al., 2019, FEIO et al., 2021).

Recently, in Minas Gerais, Normative Deliberation COPAM/CERH-MG 008/2022 stated that the ecological status must be assessed by biological indicators, using criteria and methodologies recognized by national and international institutions, which is aligned with the recommendations of the European WFD and other Global-North countries directives. However, the tools for its operation still need to be developed for Minas Gerais. In this context, existing approaches such as predictive models can be helpful for the biological assessment based on macroinvertebrate assemblages and may be a useful metric as a constituent of the ecological status of water bodies. The predictive models have been used in water resources management in several national, state, and provincial countries (FEIO et al., 2021).

The first model of this type was developed in England (River Invertebrate Prediction and Classification System - RIVPACS) (WRIGHT, 1995) to classify the ecological status of UK streams based on benthic macroinvertebrate community and environmental variables. Later, this model was adapted by the Australians with the Australian River Assessment Scheme (AUSRIVAS) (SMITH et al., 1999). Other countries also made adaptations, such as Sweden (SWEPAC) (JOHNSON; SANDIN, 2001), the USA (E/O model, VAN SICKLE et al. (2006)), and Portugal (FEIO et al., 2009b). The RIVPACS and AUSRIVAS are the main models/methods used by the United Kingdom and Australian governments to assess the ecological quality of aquatic environments (FEIO; POQUET, 2011; SMITH et al., 1999).

In contrast to the river typology, predictive models are not based on a predefined physical categorization of environments (CLARKE; WRIGHT; FURSE, 2003). The main stages of this model consist of: i) classification of reference sites into groups, exclusively by the criteria of the benthic macroinvertebrate fauna; ii) establishment of equations to relate intervals (ranges) of environmental variables with biological classification, through discriminant analyses; iii) prediction of the fauna that should occur in the absence of environmental stress (expected fauna, E). iv) comparison between the observed fauna (O) and the expected fauna (E), resulting in the observed/expected index (O/E), which is analogous to the Ecological Quality Indices (EQRs) described in the Water Framework Directive (DAVY-BOWKER et al., 2006).

Hawkins (2006) highlights some advantages of using predictive models, such as intuitive outputs, ease of interpretation of biological community results, and their inherent standardization for site conditions compared to other bioassessment tools (e.g., multimetric indices). Furthermore, there is evidence that these models can be predominantly developed using variables not affected by environmental stressors obtained from Geographic Information Systems (HARGETT et al., 2007). Therefore, they are suitable tools for spatially extensive biological assessments (FEIO et al., 2009b; SMITH et al., 1999; SUDARYANTI et al., 2001).

1.1. Research objectives

1.1.1. General objective

To develop and validate tools for river ecological status assessment in Minas Gerais.

1.1.2. Specific objectives

- To develop and validate a river typology for Minas Gerais state based on WFD System B as a point of departure, using abiotic descriptors on a landscape scale and benthic macroinvertebrate assemblages as a biotic indicator.

- To develop and test a multivariate model (MINASPACS) for spatially extensive biological assessments of rivers in Minas Gerais.

1.2. Thesis structure, questions, and hypothesis

The Thesis comprised two chapters, each resulting in original research articles.

1.2.1. Chapter I – Defining river types for establishing spatially extensive biological assessments of Minas Gerais rivers

Chapter I corresponds to the first Thesis's specific objective (topic 1.1.2.). A river typology was developed and validated for Minas Gerais state to achieve the proposed goals based on the Water Framework Directive Typology-B. Two questions were addressed:

- Does family-level identification of benthic organisms suffice for validating and distinguishing different types of Minas Gerais rivers grouped exclusively by abiotic descriptors on a landscape scale?
- What are the most representative benthic macroinvertebrate families of each river type that can be used as sentinel organisms of river degradation?

Assuming that benthic macroinvertebrate assemblage structure responds to landscape characteristics, the following hypothesis was made:

- An agreement between river typology and the structure of aquatic macroinvertebrate assemblages is expected.

1.2.2. Chapter II – A new predictive model (MINASPACS) for spatially extensive biological assessments in Minas Gerais.

Chapter II corresponds to the second specific Thesis's objective (topic 1.1.2.). A RIVPACS-type model was developed, called MINASPACS, for spatially extensive biological assessments of rivers in Minas Gerais. Two questions were addressed:

- Are multivariate predictive models useful for extensive biological assessments of rivers in Minas Gerais?

- What are the main stressors affecting the biological condition of rivers in the Minas Gerais state?

Benthic macroinvertebrates have a wide geographic distribution and high taxa richness with different sensitivity levels. Therefore, we hypothesized that aquatic environments with substantial anthropogenic stressors would simplify macroinvertebrate assemblages and that predictive modeling would represent this impairment.

1.3. Material and methods

This section presents the methods used to develop the two chapters of the Thesis. Both chapters used the same database and reference site selection criteria. The river typology construction is detailed in the methods of Chapter I, and the predictive model (MINASPACS) is explored in Chapter II.

1.3.1. Study sites and environmental characteristics

The study area covers the Pandeiros, Jequitaí, das Velhas, Pará, Araguari, Grande, Paranaíba, and Piracicaba River catchments, which are part of the São Francisco and Paraná basins, covering the main hydrologic units of Minas Gerais (586.528 km²). Single variables characterizing lithological groups, climate aspects, and river basin characteristics were extracted for each of the 381 sites (Figure 1, Table 1 and 2). Those sites were sampled under projects developed by Laboratório de Ecologia de Bentos LEB/UFMG (FEIO et al., 2015; SILVA et al., 2017; MARTINS et al., 2018, 2020; AGRA et al., 2019; CASTRO et al., 2019; GARUANA et al., 2020; LINARES et al., 2021; CALLISTO et al., 2021; MACEDO et al., 2022) e Serviço Nacional de Aprendizagem Industrial SENAI-MG (FERREIRA et al., 2017) between 2003 and 2019. The proportion of land use and cover class were estimated from a Geographic Information System - GIS (WILSON et al., 2007; WALZ; STEIN, 2014). Land use and cover data were obtained from Collection 5 of the MapBiomas online platform (2021), with a spatial resolution of 30 meters (SOUZA et al., 2020). Climatic data regarding temperature and rainfall (50-year climatic reference) were obtained from Worldclim (https://worldclim.org/). The lithological groups were defined from the Geological Map of Minas Gerais (CPRM/COMIG, 2003), scale 1:1,000,000, according to Ferreira et al. (2017). All information was organized in a GIS environment.

Figure 1 - Study area showing the sampling network and the number of samples per site.

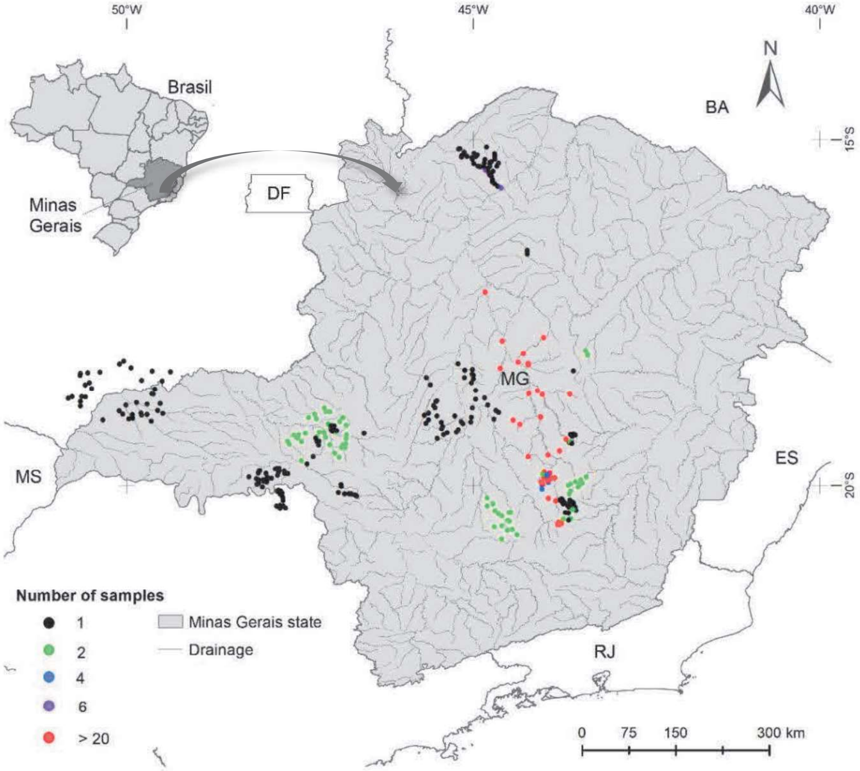


Table 1 - List of variables acquired through field measurements and laboratory analysis with their relative units and sources.

Variable	Source
Eletric conductivity (µS/cm)	Field measurement
Biochemical Oxygen Demand (mg/L)	Laboratory analysis
Phosphate (mg/L)	Laboratory analysis
Total Phosphorus (mg/L)	Laboratory analysis
Total Nitrogen (mg/L)	Laboratory analysis
Nitrate (mg/L)	Laboratory analysis
Nitrite (mg/L)	Laboratory analysis
Total Ammonia Nitrogen (mg/L)	Laboratory analysis
Dissolved oxygen (mg/L)	Laboratory analysis
pH	Field measurement
Total dissolved solids (mg/L)	Laboratory analysis
Total suspended solids (mg/L)	Laboratory analysis
Total solids (mg/L)	Laboratory analysis
Water temperature (ºC)	Field measurement
Air temperature (ºC)	Field measurement
Turbidity (NTU)	Field measurement

Table 2 - List of variables acquired through geospatial tools with their relative units and sources.

Variable	Source
Annual Mean Temperature (ºC)	Worldclim Project (https://worldclim.org/)
Mean Diurnal Range	Worldclim Project (https://worldclim.org/)
Isothermality	Worldclim Project (https://worldclim.org/)
Temperature Seasonality (standard deviation ×100)	Worldclim Project (https://worldclim.org/)
Max. Temperature of Warmest Month	Worldclim Project (https://worldclim.org/)
Min. Temperature of Coldest Month	Worldclim Project (https://worldclim.org/)
Temperature Annual Range (ºC)	Worldclim Project (https://worldclim.org/)
Mean Temperature of Wettest Quarter (ºC)	Worldclim Project (https://worldclim.org/)
Mean Temperature of Driest Quarter (ºC)	Worldclim Project (https://worldclim.org/)
Mean Temperature of Warmest Quarter (ºC)	Worldclim Project (https://worldclim.org/)
Mean Temperature of Coldest Quarter (ºC)	Worldclim Project (https://worldclim.org/)
Annual Precipitation (mm)	Worldclim Project (https://worldclim.org/)
Precipitation of Wettest Month (mm)	Worldclim Project (https://worldclim.org/)
Precipitation of Driest Month (mm)	Worldclim Project (https://worldclim.org/)
Precipitation Seasonality (Coefficient of Variation) (mm)	Worldclim Project (https://worldclim.org/)
Precipitation of Wettest Quarter (mm)	Worldclim Project (https://worldclim.org/)
Precipitation of Driest Quarter (mm)	Worldclim Project (https://worldclim.org/)
Precipitation of Warmest Quarter (mm)	Worldclim Project (https://worldclim.org/)
Precipitation of Coldest Quarter (mm)	Worldclim Project (https://worldclim.org/)
Altitude (m)	Shuttle Radar Topographic Mission – SRTM (USGS, 2005)
Mean river basin altitude (m)	Shuttle Radar Topographic Mission – SRTM (USGS, 2005)
Max. river basin altitude (m)	Shuttle Radar Topographic Mission – SRTM (USGS, 2005)
Min. river basin altitude (m)	Shuttle Radar Topographic Mission – SRTM (USGS, 2005)
River basin altitude range (m)	Shuttle Radar Topographic Mission – SRTM (USGS, 2005)
Mean river basin slope (%)	Shuttle Radar Topographic Mission – SRTM (USGS, 2005)
Max river basin slope (%)	Shuttle Radar Topographic Mission – SRTM (USGS, 2005)
Min river basin slope (%)	Shuttle Radar Topographic Mission – SRTM (USGS, 2005)
River basin slope range (%)	Shuttle Radar Topographic Mission – SRTM (USGS, 2005)
Distance to source (m)	Shuttle Radar Topographic Mission – SRTM (USGS, 2005)
Total area (km^2)	Shuttle Radar Topographic Mission – SRTM (USGS, 2005)
% Forest	MapBiomas (2021)
% Savanna	MapBiomas (2021)
% Reforestation	MapBiomas (2021)
% Grassland	MapBiomas (2021)
% Pasture (%)	MapBiomas (2021)
% Agriculture	MapBiomas (2021)
% Urban infrastructure (%)	MapBiomas (2021)
% Mining	MapBiomas (2021)
% Anthropogenic use (%)	MapBiomas (2021)
Lithological synthesis	FERREIRA et al. (2017)
Hemeroby index	WALZ; STEIN (2014)
Terrain roughness index	WILSON et al. (2007)

1.3.2. Biological and water quality sampling

Sample results comprising 14 projects from the 381 sites for benthic macroinvertebrate assemblages were compiled. Sampling used Surber or kick nets (30 cm aperture, 500 mm mesh). At each site, 3 to 20 samples were collected in the most representative habitats, then aggregated into one composite sample for each site. The samples were fixed in the field with 70% alcohol solution and deposited in the Reference Collection of Benthic Macroinvertebrates at the Institute of Biological Sciences at the Federal University of Minas Gerais and the Center for Innovation and Technology SENAI – CIT. The author collected sampes from one of the 14 compiled projects (Figure 2). For the other 13 projects, secondary data was compiled.

Samples were washed in sieves through 1.00, 0.50, and 0.25 meshes in the laboratory. All individuals were identified mainly at the family level with the aid of taxonomic keys (PÉREZ, 1988; MERRITT; CUMMINS, 1996; WIGGINS, 1996; FERNÁNDEZ; DOMÍNGUEZ, 2001; PÉS et al., 2005; COSTA et al., 2006; DOMINGUEZ et al., 2006; MERRIT et al., 2008; MUGNAI et al., 2009; HAMADA et al., 2014). Only biological data obtained during the dry season (between May and September) were used. The dry season is preferable because it facilitates habitat distinction, and the macroinvertebrate assemblage structure is more stable. Only the data with the highest taxa richness was used for sites sampled multiple times. Water quality data (e.g., total phosphorus - mg/L, total nitrogen - mg/L, and turbidity - NTU) were also compiled for each site. Parameters with more than 20% missing data were excluded from further analyses.

Figure 2 - Benthic macroinvertebrates sampling (a) and identification (b), water sample collection (c), and water quality analysis (d).

1.4. References

ABELL, R. et al. Freshwater Ecoregions of the World: A New Map of Biogeographic Units for Freshwater Biodiversity Conservation. **BioScience**, v. 58, n. 5, p. 403–414, 1 maio 2008.

AGRA, J. U. M. et al. Ecoregions and stream types help us understand ecological variability in Neotropical reference streams. **Marine and Freshwater Research**, v. 70, n. 4, p. 594–602, 2019.

AROVIITA, J. et al. A priori typology-based prediction of benthic macroinvertebrate fauna for ecological classification of rivers. **Environmental Management**, v. 42, n. 5, p. 894–906, 2008.

BAPTISTA, D. F. Uso de macroinvertebrados em procedimentos de biomonitoamento em ecossistemas aquáticos. **Oecologia Brasiliensis**, v. 12, n. 3, p. 425–441, 2008.

BARBOUR, M. T. et al. **Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macriinvertebrates, and Fish, Second Edition**. Washington: [s.n.].

BORGWARDT, F. et al. Ex uno plures – Defining different types of very large rivers in Europe to foster solid aquatic bio-assessment. **Ecological Indicators**, v. 107, p. 105599, 2019.

BRASIL. Conselho Nacional do Meio Ambiente. **Resolução n° 357 de 17 de março de 2005**. Dispõe sobre a classificação dos corpos de água e diretrizes ambientais para o seu enquadramento, bem como estabelece as condições e padrões de lançamento de efluentes, e dá outras providências. Diário Oficial da República Federativa do Brasil. Brasília, DF. 2005.

BROWN, E. D.; WILLIAMS, B. K. Ecological integrity assessment as a metric of biodiversity: are we measuring what we say we are? **Biodiversity and Conservation**, v. 25, n. 6, p. 1011–1035, 27 jun. 2016.

BUSS, D. F. et al. Stream biomonitoring using macroinvertebrates around the globe: a comparison of large-scale programs. **Environmental Monitoring and Assessment**, v. 187, n. 1, p. 4132, 9 jan. 2015.

CALLISTO, M. et al. Beta diversity of aquatic macroinvertebrate assemblages associated with leaf patches in neotropical montane streams. **Ecology and Evolution**, v. 11, n. 6, p. 2551–2560, 7 mar. 2021.

CALLISTO, M.; MACEDO, D. R.; MORENO, P. Biomonitoramento e pressões da urbanização: Uma abordagem integrada entre Ecologia e Geografia na bacia do rio das Velhas. **Revista Espinhaço**, v. 8, n. 1, p. 2–12, 2019.

CARDOSO-SILVA, S.; FERREIRA, T.; POMPÊO, M. L. M. Diretiva Quadro d'Água: uma revisão crítica e a possibilidade de aplicação ao Brasil. **Ambiente & Sociedade**, v. 16, n. 1, p. 39–58, 2013.

CASTRO, D. M. P. et al. Beta diversity of aquatic invertebrates increases along an altitudinal gradient in a Neotropical mountain. **Biotropica**, v. 51, n. 3, p. 399–411, 12 maio 2019.
COSTA, C., IDE, S., SIMONKA, C.E. **Insetos Imaturos - Metamorfose e Identific**ação. Holos, Ribeirão Preto. 2006.

CHESSMAN, B. C. What's wrong with the Australian River Assessment System (AUSRIVAS)? **Marine and Freshwater Research**, v. 72, n. 8, p. 1110, 2021.

CLARKE, R. T.; WRIGHT, J. F.; FURSE, M. T. RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers. **Ecological Modelling**, v. 160, n. 3, p. 219–233, fev. 2003.

CPRM – Serviço Geológico do Brasil. **Mapa Geológico de Minas Gerais (*shapefile*)**. Disponível em <http:..geobank.sa.cprm.gov.br.>. Acesso em: janeiro de 2020.

CPRM/COMIG - Serviço Geológico do Brasil/Companhia Mineradora de Minas Gerais. **Mapa Geológico do Estado de Minas Gerais**, escala 1:1.000.000, 2007.

DAVY-BOWKER, J. et al. A comparison of the European Water Framework Directive physical typology and RIVPACS-type models as alternative methods of establishing reference conditions for benthic macroinvertebrates. **Hydrobiologia**, v. 566, n. 1, p. 91–105, 2006.

DOMINGUEZ, E., MOLINERI, C., PESCADOR, M. L., HUBBARD, M. D., & NIETO, C. Ephemeroptera of South America. In J. Adis, J. R. Arias, G. Rueda-Delgado, & K. M. Wantzen (Eds.), **Aquatic Biodiversity of Latin America** (Vol. 2, 646 pp). Moscow, Russia: Pensoft. 2006.

DQA, 2000. **Directive 2000/60/EC of the European Parliament and of the Council**. [s.l: s.n.]. Disponível em: <https://eur-lex.europa.eu/resource.html?uri=cellar:5c835afb-2ec6-4577-bdf8-756d3d694eeb.0004.02/DOC_1&format=PDF>. Acesso em: 12 de outubro de 2018.

FEIO, M. J. et al. A predictive model for freshwater bioassessment (Mondego River, Portugal). **Hydrobiologia**, v. 589, n. 1, p. 55–68, 2007.

FEIO, M. J. et al. Defining and testing targets for the recovery of tropical streams based on macroinvertebrate communities and abiotic conditions. **River Research and Applications**, v. 31, n. 1, p. 70–84, 2015.

FEIO, M. J. et al. The Biological Assessment and Rehabilitation of the World's Rivers: An Overview. **Water**, v. 13, n. 3, p. 1–46, 2021.

FEIO, M. J.; PINTO, P. Tipologia e cenários biológicos do elemento macroinvertebrados. **Recursos Hídricos**, v. 30, n. 2, p. 19–27, 2009a.

FEIO, M. J. et al. Water quality assessment of Portuguese streams: Regional or national predictive models? **Ecological Indicators**, v. 9, n. 4, p. 791–806, 2009b.

FEIO, M. J.; POQUET, J. M. Predictive Models for Freshwater Biological Assessment: Statistical Approaches, Biological Elements and the Iberian Peninsula Experience: A Review. **International Review of Hydrobiology**, v. 96, n. 4, p. 321–346, 2011.

FERNÁNDEZ, H.R., DOMÍNGUEZ, E. **Guía Para la Determinación de los Artrópodos Bentónicos Sudamericanos**. Universidad Nacional de Tucumán, San Miguel de Tucumán, Tucumán. 2001.

FERREIRA, H. L. M. et al. **Ambientes Aquáticos em Minas Gerais**. 1a edição ed. Belo Horizonte: Prodemge, 2017.

GARUANA, L. et al. Integração de indicadores ecológicos, ambientais e de saúde humana em microbacias urbanas Luziana. **Revista Espinhaço**, v. 9, n. 1, p. 1–16, 2020.

GRILL, G. et al. Mapping the world's free-flowing rivers. **Nature**, v. 569, n. 7755, p. 215–221, 2019.

HAMADA, N., NESSIMIAN, J. L., QUERINO, R. B. **Insetos aquáticos na Amazônia brasileira: Taxonomia, biologia e ecologia**. Manaus, Brazil: Editora INPA. 2014.

HARGETT, E. G. et al. Development of a RIVPACS-type predictive model for bioassessment of wadeable streams in Wyoming. **Ecological Indicators**, v. 7, n. 4, p. 807–826, 2007.

HAWKINS, C. P. Quantifying biological integrity by taxonomic completeness: Its utility in regional and global assessments. **Ecological Applications**, v. 16, n. 4, p. 1277–1294, 2006.

HAWKINS, C. P.; OLSON, J. R.; HILL, R. A. The reference condition: predicting benchmarks for ecological and water-quality assessments. **Journal of the North American Benthological Society**, v. 29, n. 1, p. 312–343, 2010.

HUGHES, R. M.; LARSEN, D. P.; OMERNIK, J. M. Regional reference sites: a method for assessing stream potentials. **Environmental Management**, v. 10, n. 5, p. 629–635, 1986.

JOHNSON, R.; SANDIN, L. **Development of a prediction and classification system for lake (littoral, SWEPAC) and stream (riffle SWEPAC) macroinvertebrate communities**. Department of Environmental Assessment, Uppsala, Sweden: [s.n.]. Disponível em: <http://info1.ma.slu.se/IMA/Publikationer/internserie/2001-23.pdf>.

JUNQUEIRA, M. V.; FRIEDRICH, G.; PEREIRA DE ARAUJO, P. R. A saprobic index for biological assessment of river water quality in Brazil (Minas Gerais and Rio de Janeiro states). **Environmental Monitoring and Assessment**, 2010.

LINARES, M. S. et al. Chronic urbanization decreases macroinvertebrate resilience to natural disturbances in neotropical streams. **Current Research in Environmental Sustainability**, v. 3, p. 100095, 2021.

LORENZ, A.; FELD, C. K.; HERING, D. Typology of streams in Germany based on benthic invertebrates: Ecoregions, zonation, geology and substrate. **Limnologica**, v. 34, n. 4, p. 379–389, 2004.

MACEDO, D. R. et al. Development of a benthic macroinvertebrate multimetric index (MMI) for Neotropical Savanna headwater streams. **Ecological Indicators**, v. 64, p. 132–141, 2016.

MARTINS, I. et al. Regionalisation is key to establishing reference conditions for neotropical savanna streams. **Marine and Freshwater Research**, v. 69, n. 1, p. 82–94, 2018.

MARTINS, I. et al. Are multiple multimetric indices effective for assessing ecological condition in tropical basins? **Ecological Indicators**, v. 110, n. May, p. 105953, 2020.

MACEDO, D. R. et al. Urban stream rehabilitation in a densely populated Brazilian metropolis. **Frontiers in Environmental Science**, v. 10, n. September, p. 921934, set. 2022.

MELLO, K. DE et al. Multiscale land use impacts on water quality: Assessment, planning, and future perspectives in Brazil. **Journal of Environmental Management**, v. 270, n. June, p. 110879, 2020.

MELLO, K. DE et al. Biomonitoring for Watershed Protection from a Multiscale Land-Use Perspective. **diversity**, v. 15, n. 636, p. 1–20, 2023.

MERRITT, R.W., CUMMINS, K.W., BERG, M.B. **An Introduction to the Aquatic Insects of North**. 2008.

MUGNAI, R., NESSIMIAN J. L., BAPTISTA D.F. **Guide for the Identification of Aquatic Macroinvertebrates of Rio de Janeiro State**. Technical Books Editora. Rio de Janeiro, RJ: Brasil. 2009.

MINAS GERAIS. Conselho Estadual de Política Ambiental – COPAM. **Resolução Normativa Conjunta COPAM/CERH-MG nº 08 de 21 de novembro de 2022**. Dispõe sobre a classificação dos corpos de água e diretrizes ambientais para o seu enquadramento, bem como estabelece as condições e padrões de lançamento de efluentes. Diário do Executivo. Minas Gerais. 2022.

NOBLE, R.; COWX, I. **Development of a river-type classification system (D1) Compilation and harmonisation of fish species classification (D2) - Final report**. Europe, n. May 2002, p. 53, 2002.

PARDO, I. et al. The European reference condition concept: A scientific and technical approach to identify minimally-impacted river ecosystems. **Science of the Total Environment**, v. 420, p. 33–42, 2012.

PAULSEN, S. G. et al. Condition of stream ecosystems in the US: An overview of the first national assessment. **Journal of the North American Benthological Society**, v. 27, n. 4, p. 812–821, 2008.

PÉREZ, G. R. Guía para el estudio de los macroinvertebrados acuáticos del Departamento de Antioquia. Universidad de Antioquia Facultad de Ciencias Exactas y Naturales Centro de Investigaciones, CIEN. 1988.

PÉS A. M. O., HAMADA N., NESSIMIAN J. L. Identification keys of larvae for families and genera of Trichoptera (Insecta) of Central Amazonia, Brazil. **Revista Brasileira de Entomologia** 49: 181–204. 2005.

REYNOLDSON, T. B. et al. The reference condition: A comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. **Journal of the North American Benthological Society**, v. 16, n. 4, p. 833–852, 1997.

RUARO, R.; GUBIANI, É. A. A scientometric assessment of 30 years of the Index of Biotic Integrity in aquatic ecosystems: Applications and main flaws. **Ecological Indicators**, v. 29, p. 105–110, 2013.

SANTOS, J. I. et al. Challenges to water quality assessment in Europe – Is there scope for improvement of the current Water Framework Directive bioassessment scheme in rivers? **Ecological Indicators**, v. 121, p. 107030, fev. 2021.

SILVA, D. R. O. et al. An improved macroinvertebrate multimetric index for the assessment of wadeable streams in the neotropical savanna. **Ecological Indicators**, v. 81, n. June, p. 514–525, 2017.

SMITH, M. J. et al. AusRivAS: Using macroinvertebrates to assess ecological condition of rivers in Western Australia. **Freshwater Biology**, v. 41, n. 2, p. 269–282, 1999.

SOLHEIM, A. L. et al. A new broad typology for rivers and lakes in Europe: Development and application for large-scale environmental assessments. **Science of the Total Environment**, v. 697, p. 134043, 2019.

SOUZA, C. M. et al. Reconstructing three decades of land use and land cover changes in Brazilian biomes with Landsat archive and Earth engine. **Remote Sensing**, v. 12, n. 17, p. 2735, ago. 2020.

STODDARD, J. L. et al. Setting expectations for the ecological condition of streams: The concept of reference condition. **Ecological Applications**, v. 16, n. 4, p. 1267–1276, 2006.

SUDARYANTI, S. et al. Assessment of the biological health of the Brantas River, East Java, Indonesia using the Australian River Assessment System (AUSRIVAS) methodology. **Aquatic Ecology**, v. 35, n. 2, p. 135–146, 2001.

USGS (United States Geological Survey). **Shuttle Radar Topography Mission – SRTM**, http://www.srtm.usgs.gov. 2005.

VAN SICKLE, J., HUFF, D.D., HAWKINS, C.P., 2006. Selecting discriminant function models for predicting the expected richness of aquatic macroinvertebrates. **Freshw. Biol**. 51, 359–372. https://doi.org/10.1111/j.1365-2427.2005.01487.x

WALZ, U.; STEIN, C. Indicators of hemeroby for the monitoring of landscapes in Germany. **Journal for Nature Conservation**, v. 22, n. 3, p. 279–289, 2014.

WIGGINS, G. B. **Larvae of the North American Caddisfly Genera (Trichoptera)**, 2nd edn. University of Toronto Press: Toronto, Canada. 1996.

WILSON, M. F. J. et al. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. **Marine Geodesy**. [s.l: s.n.].

WRIGHT, J. F. et al. A preliminary classification of running-water sites in Great Britain based on macro-invertebrate species and the prediction of community type using environmental data. **Freshwater Biology**, v. 14, n. 3, p. 221–256, 1984.

WRIGHT, J. F. Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. **Australian Journal of Ecology**, v. 20, n. 1, p. 181–197, 1995.

2. CHAPTER I - DEFINING RIVER TYPES FOR ESTABLISHING SPATIALLY EXTENSIVE BIOLOGICAL ASSESSMENTS OF MINAS GERAIS RIVERS

2.1. Abstract

Modern spatially extensive programs for the ecological assessment of rivers consider their natural variability as a basis for defining reference values for those assessments. The European Union's Water Framework Directive (WFD) uses a river typology based on common environmental variables that determine different aquatic habitats and biological assemblages. This approach could also be used in the Southern Hemisphere, however, no attempts have been made so far in Minas Gerais in spatially extensive monitoring programs. Thus, we sought to develop and validate a typology for Minas Gerais rivers, using abiotic descriptors on a landscape scale and benthic macroinvertebrate assemblages as biotic indicators. Using a Grouping Analysis tool, the drainage segments of each river type were selected to be as similar as possible, and all other river types were as different as possible. Two markedly different groups of rivers (mountain and lowland) were formed according to the best results of the spatial cluster analysis, which were built with only the continuous variables. Family-level benthic macroinvertebrate data were used to check for statistical differences among the biological assemblages and to validate each river type through multidimensional scaling analyses and ANOSIM tests. The river types were shown to be useful for establishing reference conditions for biological assessment and offer an option for better predicting and managing aquatic ecosystem biodiversity patterns in Minas Gerais rivers and streams.

Keywords: River typology, Water Framework Directive, Benthic macroinvertebrates.

2.2. Introduction

Rivers are dynamic systems influenced by natural and anthropogenic forces that cause constant changes along their longitudinal gradients (CALLISTO et al., 2019b; MACEDO et al., 2016; OMERNIK et al., 2017; VANNOTE et al., 1980). Alterations in their hydromorphology, water quality or quantity can compromise the integrity of these ecosystems, ultimately influencing the ecosystem services provided to humans (MELLO et al., 2020; SANTOS et al., 2021). Distance to the source, drainage area, slope, landform, and type of lithology are also important landscape characteristics that can govern biotic conditions (LORENZ; FELD; HERING, 2004; MOYA et al., 2011). In addition, freshwater biota are affected by several anthropogenic pressures at various spatial and temporal scales, and are often used for assessing river ecological quality (HERLIHY et al., 2020).

The European Water Framework Directive – WFD (Directive 2000/60/EC) established criteria for classifying water bodies, which form the basis of the environmental classification systems of the European Member States (SANTOS et al., 2021). Following the WFD, the European Member States developed river typologies (BORGWARDT et al., 2019; FEIO; PINTO, 2009; LORENZ; FELD; HERING, 2004), seeking to establish groups of rivers with homogeneous natural environmental characteristics (e.g., geology, altitude, slope, temperature, and precipitation) (SOLHEIM et al., 2019). Stream typologies can be developed following either ‘‘top-down’’ or ‘‘bottom-up” approaches. In ‘‘bottom-up’’ approaches, analyses of site-specific biological data are used to group rivers by similarities in their assemblage composition (FERRÉOL et al., 2005; HERLIHY et al., 2020), which requires a considerable data-collection effort a priori. In ‘‘top-down’’ approaches, previous knowledge and human presumptions are used to select candidate parameters. WFD typology-A is based on “obligatory” top-down parameters (hydrologic unit, altitude, catchment area, and geology) (AROVIITA et al., 2008). WFD typology-B includes obligatory parameters (altitude, latitude, longitude, lithology, size) plus 15 optional factors (BORGWARDT et al., 2019; MOOG et al., 2004; PERO et al., 2020; VERDONSCHOT; NIJBOER, 2004)

In a proper classification, the biological variability within the same river type is expected to be lower than that observed in the sum of two or more types (PARDO et al., 2012). Both A and B typologies can be developed using geospatial databases through Geographical Information Systems (GIS). This facilitates the use of both systems in various regions because of the availability of free-use geospatial databases worldwide (MACEDO et al., 2018). River typologies form frameworks for ecological status assessment methods, which can include multimetric indices (HERLIHY et al., 2020; MACEDO et al., 2016, 2014; PONT et al., 2006) and predictive models (DAVY-BOWKER et al., 2006; HARGETT et al., 2007; KAUFMANN et al., 2022). Among the various assemblages used in the ecological assessment of rivers, benthic macroinvertebrates are widely used. They are ubiquitous, have a high taxa richness with different levels of sensitivity to environmental stressors, are relatively sessile, and have relatively long life cycles, which facilitates assessing the effects of changes over space and time (BUSS et al., 2015; CALLISTO et al., 2019b; FEIO et al., 2021, 2022). Although several typology methods have been proposed and tested in temperate ecosystems, few studies have used this approach for rivers and streams in tropical environments except at small spatial extents (e.g., FERREIRA et al., 2017; AGRA et al., 2019; MARTINS et al., 2018). Adapting a European WFD river typology as the basis for spatially extensive ecological monitoring could be extremely useful in large tropical countries such as Brazil, which is composed of different states with different governments, each with various degrees of independence (BUSS et al., 2015).

In Brazil, states and municipalities are allowed to edit laws on the management of the waters under their domain (IGAM, 2019). The law that classifies water quality, CONAMA Resolution Nº. 357/2005 (BRASIL, 2005), requires only the use of physical, chemical, ecotoxicological and bacteriological parameters as water quality indicators, and biological monitoring is optional. It is widely accepted amongst ecologists that assessments based only on such parameters do not provide adequate answers about the ecological quality of aquatic ecosystems (FEIO et al., 2021; KARR, 2006). In Minas Gerais, however, the Normative Deliberation COPAM/CERH-MG number 008/2022 (MINAS GERAIS, 2022) recently established the WFD approach to river typology as a key initial stage of biomonitoring programs.

The aim of this study was to develop and validate a typology for Minas Gerais rivers based on WFD System B as a point of departure, using abiotic descriptors on a landscape scale and benthic macroinvertebrate assemblages as biotic indicators. Assuming that benthic macroinvertebrate assemblage structure responds to landscape characteristics, an agreement between the river typology and the structure of aquatic macroinvertebrate assemblages is expected. Two other two questions were also sought to be answered. i) Does family-level identification of benthic organisms suffice for validating and distinguishing different types of Minas Gerais rivers grouped exclusively by abiotic descriptors on a landscape scale? ii) What are the most representative benthic macroinvertebrate families of each river type that can be used as sentinel organisms of river degradation?

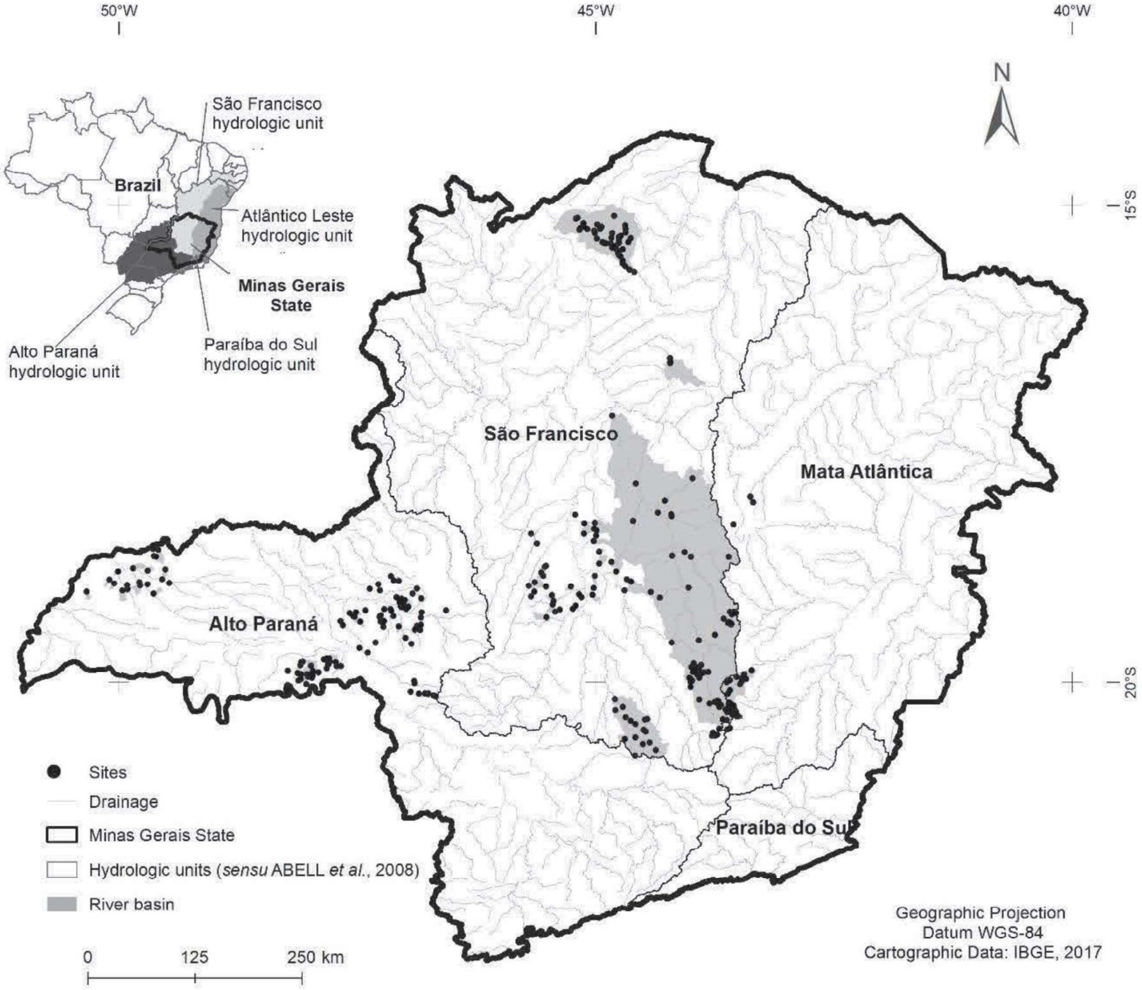
2.3. Materials and methods

2.3.1. Study area

The study area covers the Pandeiros, Jequitaí, das Velhas, Pará, Araguari, Grande, Paranaíba, and Piracicaba River catchments, which are part of the São Francisco and Paraná basins, covering the main hydrologic units of Minas Gerais (586.528 km²) (Figure 1; Supplementary information 1, Table S1). Between 2003 and 2019, benthic macroinvertebrate samples were collected in 348 sites from different research projects of the Laboratory of Ecology of Benthos (AGRA et al., 2019; CALLISTO et al., 2021; CASTRO et al., 2019; FEIO et al., 2015; GARUANA et al., 2020; LINARES et al., 2021; MACEDO et al., 2022; MARTINS et al., 2018, 2020; SILVA et al., 2017) and SENAI-MG (FERREIRA et al., 2017).

Sites are distributed in the São Francisco, Atlantic Forest, and Alto Paraná hydrologic units (sensu ABELL et al., 2008), which cover the Cerrado and Atlantic Forest biomes, with varied landforms and climate (FERREIRA et al., 2017). There is a north-south climate gradient, with sub-hot humid occurrence three dry months in the south of the state, hot semi-humid in the central portion, and five dry months in the extreme north of the area. Industrial and mining activities occur mainly in the southern portion of the São Francisco hydrologic unit, where the “Quadrilátero Ferrífero” region is located (FERREIRA et al., 2017).

Figure 1 - Study area showing the river basins (gray polygons) and sites (black dots).



2.3.2. Study sites and environmental characterization

For each of the 348 sites, single variables were collected characterizing lithological groups, climate, and watershed characteristics (Supplementary Information 1, Table S1). Land use proportions were estimated from a Geographic Information System – GIS. Land use data were obtained from Collection 5 of the MapBiomas online platform (2021), with a spatial resolution of 30 meters (SOUZA et al., 2020). Climatic data regarding temperature and rainfall (50-year climatic reference) were obtained from Worldclim (https://worldclim.org/). The lithological groups (Supplementary Information 1, Table S2) were defined from the Geological Map of Minas Gerais, according to Ferreira et al. (2017). The grouping of cartographic units was based on the similar response of rocks to surface processes such as erosion, weathering, and leaching (FERREIRA et al., 2017). All information was organized in a GIS environment.

2.3.3. Biological samples and water quality

Benthic macroinvertebrate assemblages information were compiled from the 348 sites. Sampling was done using Surber or kick nets (30 cm aperture, 500 mm mesh, and 0.09 m²). At each site, from 3 to 20 sub-samples were collected in the most representative habitats, then aggregated into one composite sample for each site. Samples were fixed in the field with 70% alcohol and deposited in the Reference Collection of Benthic Macroinvertebrates at the Institute of Biological Sciences at the Federal University of Minas Gerais (CALLISTO et al., 2021) and the Center for Innovation and Technology SENAI – CIT. At the laboratory, samples were washed in sieves through 1.00, 0.50, and 0.25 meshes in the laboratory, and all individuals were identified mainly to family with the aid of taxonomic keys (PÉREZ, 1988; MERRITT; CUMMINS, 1996; WIGGINS, 1996; PÉS et al., 2005; DOMINGUEZ et al., 2006; MUGNAI et al., 2009, 2010; HAMADA et al., 2014). Water quality data (total phosphorus - mg/L, total nitrogen - mg/L, and turbidity - NTU) were also compiled for each site. Only biological data obtained during the dry season (between May and September) were used, and in the case of sites sampled multiple times, only the data with the highest taxa richness was used.

2.3.4. River typology

Typology was built based on System B, considering the lithological diversity and climatic complexity in the study area, corresponding to 24 candidate variables. To do so, four steps were followed. i) Hydrologic units identification according to Abell et al. (2008). ii) Multicollinearity reduction of the abiotic variables through Spearman rank correlations (removal of variables with $r > |0.7|$ (DORMANN et al., 2013). iii) Attributes relative to river segments were determined (average length of 4,748. m) with spatial information derived from a surface using GIS. iv) Using the Grouping Analysis tool from ArcGIS 10.4 (ESRI, 2016), a spatial cluster analysis procedure was performed that guarantees that all the drainage sections of each group are as similar as possible, and all the groups themselves are as different as possible. The Grouping Analysis tool uses a K Means algorithm, and grouping effectiveness is measured using the Calinski-Harabasz pseudo F-statistic, which is a ratio reflecting within-group similarity and among-group difference (WARCHALSKA-TROLL; WARCHALSKI, 2022). The Grouping Analysis tool assesses the effectiveness of dividing the features into 2, 3, 4, and up to 15 groups (ESRI, 2016). Therefore, two river typologies for validation were built: i) performing the spatial cluster analysis for all continuous variables (i.e., annual mean temperature and altitude) and ii) joining the spatial clusters with the lithological groups, since lithology is a nominal variable.

2.3.5. Biological validation of the typology

To validate abiotic typologies, reference sites (least disturbed sites) were selected for each river type from available databases (Figure S3). Criteria used for reference site selection were:

1. Land use and occupation in the hydrographic basin: absence of urban infrastructure and percentage of anthropogenic areas < 25%;
2. Exclusion of sites that do not meet the Brazilian legal limits (BRASIL, 2005) for phosphorus, nitrogen, and turbidity, Class II, lotic environments. Class II corresponds mainly to water intended for human consumption after simplified treatment and protection of aquatic communities.

This selection intends to avoid the confounding effect that could occur from alterations of macroinvertebrate assemblages caused by anthropogenic disturbance instead of differences resulting from the different abiotic characteristics, such as geology or climate (STODDARD et al., 2006; WHITTIER et al., 2007). Then, we used the biological data from reference sites to validate the river typology in two stages: (1) considering spatial cluster groups and (2) using the spatial cluster groups joined with river typology lithological groups. In both approaches the following were considered: i) Each site was assigned to a river type. Biological data were used to group sites using non-metric multidimensional scaling (Bray-Curtis coefficient, Past 4.03) based on the fourth root transformation of taxon abundances. ii) Analysis of similarities (ANOSIM) was used to check for statistical differences between the biological assemblages contained in each river type (9,999 permutations, Bray-Curtis coefficient, Past 4.03). iii) Next, SIMPER statistical procedures (similarity/distance percentages, fourth root transformation, Bray-Curtis coefficient, Past 4.03) was used to determine the most representative families (up to 90% of cumulative percentage) for each river type. iv) Six biological and ecological traits (life cycle, food (diet), functional feeding group, mobility, respiration and tolerance) were obtained from literature for the most representative taxa of each river type (ALBA-TERCEDOR, 1996; BIS; USSEGLIO-POLATERA, 2004; JUNQUEIRA et al., 2018; REYNAGA; DOS SANTOS, 2012; TOMANOVA; MOYA; OBERDORFF, 2008). v) Finally, the river types were graphically compared using boxplot and non-metric multidimensional scaling (nMDS) graphics. The Biological Monitoring Working Party (BMWP) index (JUNQUEIRA et al., 2018) was also calculated for each reference site.

2.4. Results

2.4.1. Abiotic typology

Spearman's rank correlation indicated five descriptors for the typology construction: altitude, average annual temperature, annual precipitation, terrain roughness, and lithology (Supplementary information 1, Table S3). Two markedly different groups of rivers (mountain and lowland) were formed according to the best results of the spatial cluster analysis, which were built with only the continuous variables (Supplementary information 1, Table S4). The first group (mountain) refers to rivers predominantly located at higher elevations (average of 858 m), with generally more precipitation (average of 1,459 mm) and lower temperatures (average of 20.1 ºC). Group 2 (lowland) were predominantly located in lower elevations (average of 542 m) with lower rainfall (average of 1,123 mm) and higher temperatures (average of 22.6 ºC) (Figure 2 and 3). Finally, the mountain and lowland river types were combined with the eight lithological classes in Minas Gerais, resulting in 15 river types (Figure 4).

Figure 2 - Box plots showing river-type characteristics: altitude, average annual temperature, annual precipitation, and terrain roughness.

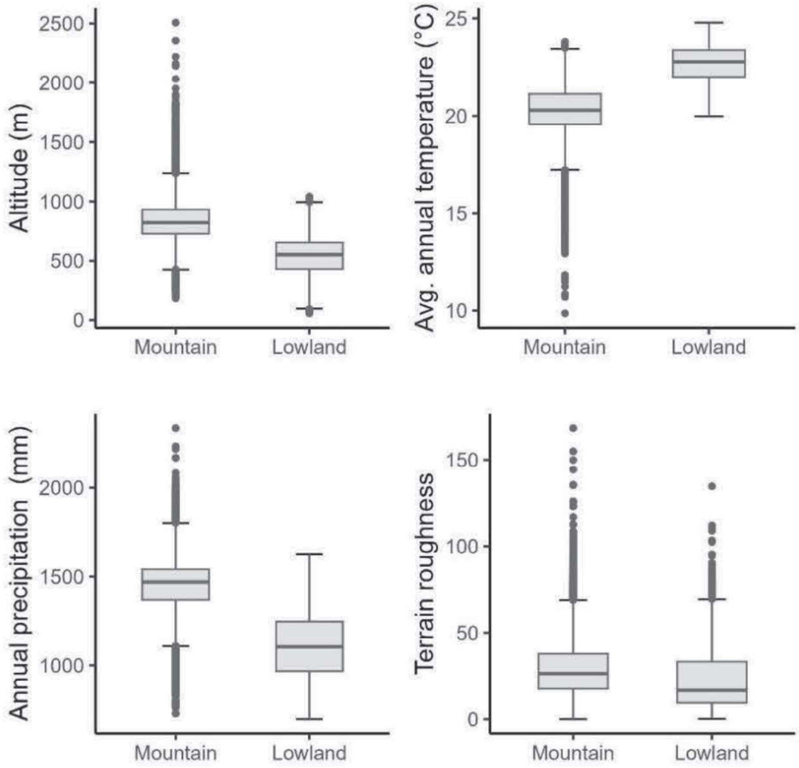


Figure 3 - Spatial distribution of the mountain and lowland river types across Minas Gerais (two river types).

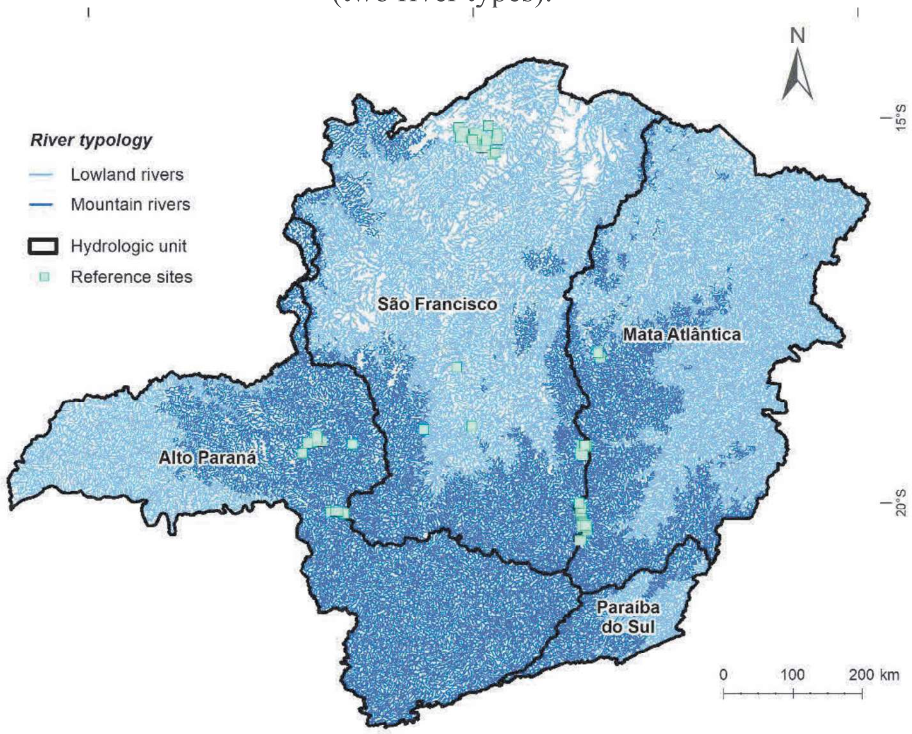
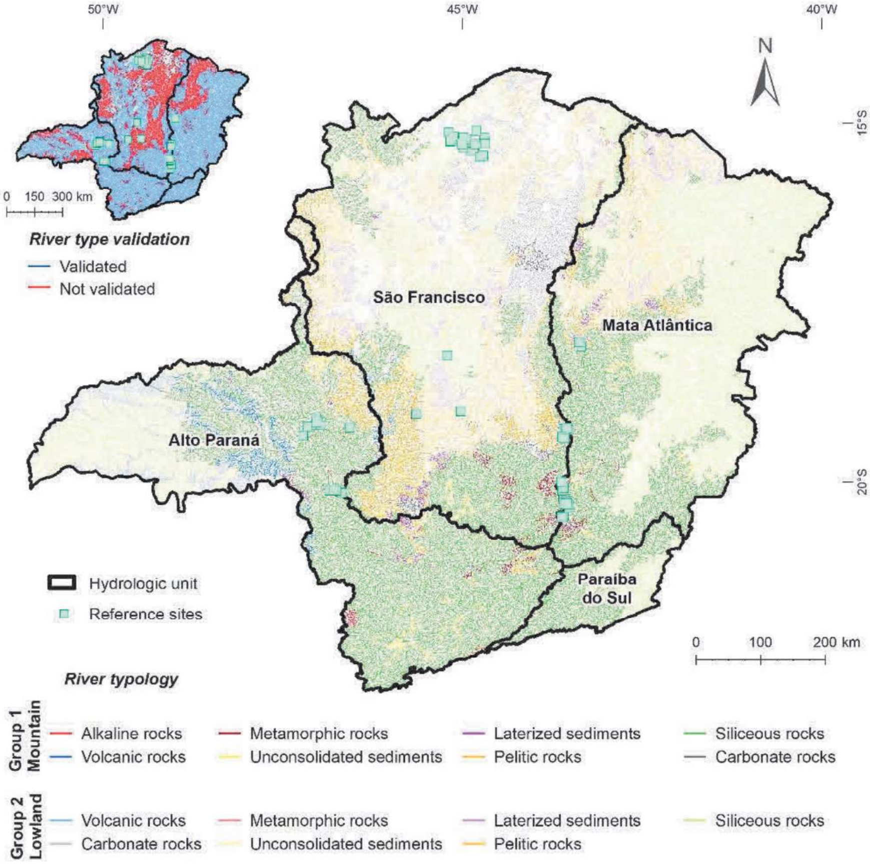


Figure 4 - Spatial distribution of the mountain and lowland river types joined with eight lithological classes (15 river types).



2.4.2. Biological validation of the typology

No site was available in the Paraíba do Sul hydrologic unit, and > 70% of our reference sites are in the São Francisco hydrologic unit. From these, 87 sites met the selection criteria for minimally disturbed sites and were used for the biological validation of the abiotic typology. More than 73% of our reference sites are in small catchments (<100 km²) and only three sites have a drainage area > 1,000 km² (SOLHEIM et al., 2019). A total of 103 taxa were identified, mainly at family level.

The non-metric multidimensional scaling (Bray-Curtis coefficient) for the 87 reference sites significantly differentiated benthic macroinvertebrate assemblages from lowland and mountain rivers (2D stress = 0,256, Figure 5), corroborated by the ANOSIM result (Bray-Curtis coefficient) (R = 0.43, p=0.001). A SIMPER analysis within each type revealed the lowest values for the mountain river type. On the other hand, the average similarity was higher for the lowland river type, despite having fewer samples in this group, indicating a higher consistency in the assemblages found in lowlands (Table 1). These same tests were performed t for the 15 river types built from the integration with lithology (Figure 4), however, the inclusion of lithology actually impeded rather than facilitated the identification of patterns (Fig 5). Thus, subsequent analyses were focused on the lowland and mountain river types. Nonetheless, because lithology is an obligatory variable of the WFD typology-B, results for the lithology integration are presented in Supplementary Material 2 (Table S2, Table S3, Table S4 and Figure S1 and S2).

Figure 5 - Non-metric multidimensional scaling (nMDS) ordination of 87 benthic macroinvertebrate samples classified by river types: two spatial clusters (2D stress = 0,256) and five river types (2D stress = 0,252).

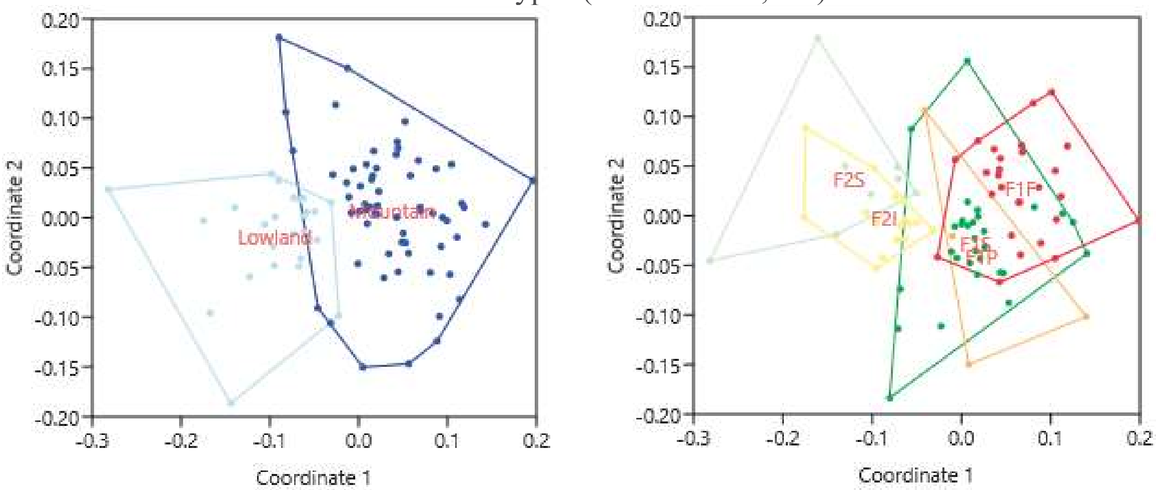


Table 1 - SIMPER analysis based on benthic macroinvertebrate assemblages of river types F1S and F2S (Siliceous rocks); F2I (Unconsolidated sediments); F1P (Pelitic rocks); and F1F (Metamorphic rocks).

Type	Number of sites	Mean similarity
Mountain rivers (F1S + F1P + F1F)	**62**	**56.8%**
Mountain rivers over siliceous rocks (F1S)	30	59.4%
Mountain rivers over pelitic rocks (F1P)	5	55.9%
Mountain rivers over metamorphic rocks (F1F)	27	59.2%
Lowland rivers (F2S + F2I)	**25**	**59.4%**
Lowland rivers over siliceous rocks (F2S)	8	55.4%
Lowland rivers over unconsolidated sediments (F2I)	16	63.7%

A set of common taxa occurred in practically all reference sites and river types, such as Chironomidae, Elmidae, Baetidae and Oligochaeta, families with wide distribution in Minas Gerais, including some of the most tolerant groups to anthropogenic disturbance (JUNQUEIRA et al., 2018). Excluding those common taxa, the most representative families in each river type were: Perlidae, Tipulidae and Hydropsychidae for mountain rivers; Naucoridae, Bivalvia and Hydrobiosidae for lowland rivers (Table 2 and 3). The BMWP index was significantly ($p < 0.01$) higher for the mountain rivers (Supplementary Information 1, Figure S1).

Table 2 - Taxa contributing up to 90% of cumulative percentage to Bray-Curtis similarity for the mountain and lowland river types, in descending order of contribution. Taxa in bold are exclusive between the two river types.

Mountain (Taxa %)	Lowland (Taxa %)
Chironomidae (13.15)	Chironomidae (13.75)
Elmidae (7.89)	Elmidae (7.89)
Leptophlebiidae (7.28)	Ceratopogonidae (6.98)
Baetidae (6.25)	Leptohyphidae (6.13)
Simuliidae (4.68)	Gomphidae (5.02)
Perlidae (4.13)	Leptoceridae (4.87)
Ceratopogonidae (4.07)	**Naucoridae (4.63)**
Leptohyphidae (3.86)	Baetidae (4.6)
Tipulidae (3.7)	**Bivalvia (4.25)**
Hydropsychidae (3.65)	Leptophlebiidae (4.23)
Oligochaeta (3.56)	Libellulidae (4.13)
Coenagrionidae (3.16)	Oligochaeta (4.01)
Libellulidae (2.79)	**Hydrobiosidae (3.98)**
Leptoceridae (2.77)	Coenagrionidae (2.58)
Polycentropodidae (2.45)	Empididae (2.4)
Calamoceratidae (2.01)	Helicopsychidae (2.34)
Gripopterygidae (1.9)	**Pyralidae (2.29)**
Empididae (1.62)	Hydroptilidae (1.62)
Odontoceridae (1.45)	Calopterygidae (1.52)
Gomphidae (1.31)	**Caenidae (1.36)**
Megapodagrionidae (1.27)	**Hydrobiidae (1.27)**
Corydalidae (1.17)	Simuliidae (1.22)
Hydroptilidae (1.12)	-
Helicopsychidae (1.11)	-
Veliidae (1.03)	-
Calopterygidae (0.94)	-
Euthyplociidae (0.84)	-
Dytiscidae (0.77)	-
Philopotamidae (0.68)	-

Table 3 - Description of traits for the most representative families in each river type and references.

River type	Taxon	Life cycle	Food (diet)	Functional Feeding Group	Mobility	Respiration	Tolerance
Mountain	Perlidae	> 1 year[5]	Living macroinvertebrates[5]	Predator[3]	Crawler[3]	Gills[3]	10[1]
	Tipulidae	> 1 year[5]	Detritus (plant)[5]	Shredders[3]	Burrowers[3]	Aerial[3,4]	5[1]
	Hydropsychidae	< 1 year[5]	Detritus < 1mm[5]	Collectors-filters[3]	Crawler[3]	Gills[3]	5[1]
Lowland	Naucoridae	< 1 year[5]	Living macroinvertebrates[5]	Predator[3]	Swimmers[3]	Aerial[3,4]	3[1]
	Bivalvia	> 1 year[5]	Living microphytes[5]	Collectors-filters[3]	Burrowers[3]	Gills[3]	3[2]
	Hydrobiosidae	< 1 year[5]	Living macroinvertebrates[5]	Predator[3]	Crawler[3,4]	Tegumentary[3]	8[1]

1: Junqueira et al. (2018), 2: Alba-Tercedor (1996), 3: Tomanova et al. (2008) , 4: Reynaga; Dos Santos (2012) , 5: Bis; Usseglio-Polatera (2004).

2.5. Discussion

A river typology was developed and validated for Minas Gerais based on WFD Typology-B, which reflected the natural variability of the benthic macroinvertebrate assemblages on a landscape scale, corroborating our hypothesis. The grouping into mountain and lowland rivers (i.e., landform groups) was evident, and similar results were found by Lorenz et al. (2004) in Germany, Pero et al. (2020) in Argentina, Moya et al. (2011) in Bolivia, Fuster et al. (2012) in Chile, and Herlihy et al. (2019; 2020) in the USA. Both river classifications presented in this study should be seen as a first attempt and therefore be used for constructing a quality assessment scheme based on macroinvertebrates, which constitutes an improvement in the present state of the art for Minas Gerais waters.

The reference site selection criteria resulted in 87 minimally disturbed sites, which allowed the validation of the mountain and lowland river types in Minas Gerais, which is unique for Brazil and most South American countries, except Chile (FUSTER et al., 2012). Therefore, it is safe to say that the river typology approach, adapting the European Water Framework Directive (Directive 2000/60/EC), is also useful in Minas Gerais as a basis for developing spatially extensive biological assessments and their classification schemes (BORGWARDT et al., 2019; FEIO; PINTO, 2009; LORENZ; FELD; HERING, 2004; SOLHEIM et al., 2019).

Family-level identification of macroinvertebrates was proven to be efficient for river typology construction also, as previously demonstrated by Martins et al. (2018) in Minas Gerais, Pero et al. (2020) in Argentina, and Gutiérrez et al. (2017) for Colombian Andean rivers. Family identifications were also used worldwide for developing biological quality indices based on macroinvertebrate assemblages (FEIO et al., 2021). The list of the most representative taxa of each river type can help guide future ecological status assessments (FEIO; PINTO, 2009).

The top three uplands indicator taxa are moderately sensitive to anthropogenic disturbances, are predators and omnivores, and occur in stable and depositional substrates. The Tipulidae and Hydropsychidae taxa are generally detritivores, whereas Perlidae feed on living macroinvertebrates. The loss or reduction of riparian vegetation impairs the functional structure of upland rivers (TUPINAMBÁS et al., 2014). Therefore, these taxa should be reduced by anthropogenic disturbances that alter their substrates (such as sedimentation) and reduce their food bases (e.g. insecticides, allochthonus materials). The top three lowlands indicator taxa are moderately tolerant to a wide range of environmental conditions and anthropogenic disturbances (JUNQUEIRA et al., 2018). These taxa consist of predators (Naucoridae and Hydrobiosidae) and collectors-filters (Bivalvia), that can exist in both stable and depositional substrates. Therefore, they should be reduced to a lesser degree than the upland taxa by anthropogenic disturbances that alter their substrates (such as sedimentation) but are still sensitive to disturbances that affect their food sources (such as insecticides and reduced riparian vegetation) or respiratory functions (e.g. inadequate sewage treatment or excess nutrient loadings).

Although the lowland and mountain river types have been validated, increasing the sampling network with high quality sites would allow the definition of more river types or perhaps the use of bottom-up approaches. In Europe, several countries faced problems related to river typology validation because of the lack of reference areas, especially for large rivers with basins >10,000 km^2 but also in smaller rivers and streams in coastal areas, with high population densities and industrialization (BORGWARDT et al., 2019; ELIAS et al., 2016). Other alternatives to establish reference conditions in areas where minimally disturbed sites no longer exist or historical pre-disturbance data are unavailable are the modelling of reference conditions or the adjustment of reference values based on correction factors (e.g, ELIAS et al., 2016). Probability-based, spatially balanced sampling has also proven to be an effective technique for selecting samples that reflect the spatial patterns of study areas (HERLIHY et al., 2020; OLSEN et al., 2008). In Brazil, this is an important approach used only since 2013 (CALLISTO et al., 2019a; FIRMIANO et al., 2017; LIGEIRO et al., 2013; MACEDO et al., 2014; MARTINS et al., 2020; SILVA et al., 2018).

River typologies based on the WFD philosophy generally use environmental descriptors from a landscape scale. Our results showed that altitude is a major driver in defining river types. Similar results were found by Moya et al. (2011) in Bolivia, Fuster et al. (2010) in Chile, Pero et al. (2020) in Argentina, and Lorenz et al. (2004) in German streams, where a clear separation occurred between lowland and upland streams. Lithology, owing to its qualitative nature, does not seem to explain much biological variation (FERRÉOL et al., 2005). It is unclear in the WFD whether the lithology class should be based on the lithology underlying the biological sampling site or the upstream catchment's lithology. Site-scale descriptors, such as substrate composition, current velocity, conductivity, and stream size (AROVIITA et al., 2008; MOYA et al., 2011; BORGWARDT et al., 2019) or landscape predictors such as stream slope, stream volume, distance from the source (PONT et al., 2009, 2006) can be highly correlated with the composition of macroinvertebrate assemblages and can improve the river typology (DAVY-BOWKER et al., 2006; PONT et al., 2009, 2006; FEIO et al., 2007b; MOYA et al., 2011). Some of this information is obtained during the field survey stage, so it is necessary to standardize the sampling protocol and train the researchers to reduce interpersonal variability (HUGHES et al., 2008; JUSIK et al., 2015). The river typology can also be improved using two or more bioindicators because they respond differently to abiotic groupings (FEIO et al., 2007a; HERLIHY et al., 2020). As our data were compiled from different years, sources, research protocols, and independent research teams, criteria for data homogenization were established. Although this type of approach is not ideal because it introduces data variability, it has been successfully used in landscape-scale studies to take full advantage of historical databases (BORGWARDT et al., 2019; FEIO et al., 2022; TAMVAKIS et al., 2014).

River biological assessments are often based on multimetric indices (CALLISTO et al., 2019a; SILVA et al., 2018) and predictive models (MOYA et al., 2011; FEIO; POQUET, 2011; PARDO et al., 2014) that are sensitive to natural variability and anthropogenic pressures (CHEN et al., 2019; FEIO et al., 2021). Using river typology as a preliminary assessment stage can facilitate establishing reference values for any multimetric index, metric or environmental descriptor according to the river type. Approaches that include river typology can provide more accurate answers than methods that do not consider natural environmental heterogeneity

(AGRA et al., 2019). Accurate classifications reduce the probability of inferring impairment when it does not exist or not detecting it when it exists (PERO et al., 2020).

Brazil's National Water Resources Policy defines river basins as the political units for water management. However, smaller units and characteristics, such as river types, should also be considered to monitor water body conservation status at a finer resolution, because abiotic and biotic conditions often vary markedly within river basins (KAUFMANN et al., 2022; OMERNIK et al., 2017). Our results can be used under the new water resources law requirements (MINAS GERAIS, 2022) and encourage discussion with the National Water Resources Agency on the importance of using river typology nationally. In an ideal scenario, this methodology should be expanded to the broadest geographic extent possible (e.g., South America). In this way, it would be possible to compare the ecological status assessments between different regions and countries (BORGWARDT et al., 2019). Therefore, the results can benefit spatially extensive ecological research on the impacts of multiple pressures on rivers by aggregating data comparable across large regions or countries (BORGWARDT et al., 2019; SOLHEIM et al., 2019).

2.6. Conclusions

The results represent a first step for further studies that may use river typology to elucidate aquatic ecosystem biodiversity in Minas Gerais rivers and streams and improve freshwater monitoring programs and management. It is safe to say that the river typology approach, as recommended by the European Water Framework Directive, is also useful in Brazil for improving biological assessment methods. The most representative macroinvertebrate families for the two river types were also listed. By understanding their traits, the processes which most impair each river type can be better understood. Furthermore, the river typology can be effectively used as a tool to improve aquatic ecosystem research and management.

2.7. Acknowledgements

We thank the field crews for their great efforts in generating the large amount of data used in the present study, mainly colleagues from the Laboratório de Ecologia de Bentos

UFMG, Laboratorio de Ecologia de Peixes UFLA e Centro de Inovação e Tecnologia SENAI CIT FIEMG.

2.8. References

ABELL, R. et al. Freshwater Ecoregions of the World: A New Map of Biogeographic Units for Freshwater Biodiversity Conservation. **BioScience**, v. 58, n. 5, p. 403–414, 1 maio 2008.

AGRA, J. U. M. et al. Ecoregions and stream types help us understand ecological variability in Neotropical reference streams. **Marine and Freshwater Research**, v. 70, n. 4, p. 594–602, 2019.

ALBA-TERCEDOR, J. Macroinvertebrados acuáticos y calidad de las aguas de los ríos. **IV Simposio del Agua en Andalucía (SIAGA)**, v. II, n. January 1996, p. 202–213, 1996.

AROVIITA, J. et al. A priori typology-based prediction of benthic macroinvertebrate fauna for ecological classification of rivers. **Environmental Management**, v. 42, n. 5, p. 894–906, 2008.

BIS, B.; USSEGLIO-POLATERA, P. **Standardisation of River Classifications: Species Traits Analysis**. [s.l: s.n.].

BORGWARDT, F. et al. Ex uno plures – Defining different types of very large rivers in Europe to foster solid aquatic bio-assessment. **Ecological Indicators**, v. 107, p. 105599, 2019.

BRASIL. Conselho Nacional do Meio Ambiente. **Resolução n° 357 / 2005**. Dispõe sobre a classificação dos corpos de água e diretrizes ambientais para o seu enquadramento, bem como estabelece as condições e padrões de lançamento de efluentes, e dá outras providências. Diário Oficial da República Federativa do Brasil. Brasília, DF, 2005.

BUSS, D. F. et al. Stream biomonitoring using macroinvertebrates around the globe: a comparison of large-scale programs. **Environmental Monitoring and Assessment**, v. 187, n. 1, p. 4132, 9 jan. 2015.

CALLISTO, M. et al. Multi-status and multi-spatial scale assessment of landscape effects on benthic macroinvertebrates in the Neotropical Savanna. **Advances in Understanding Landscape Influences on Freshwater Habitats and Biological Assemblages**, p. 275–302, 2019a.

CALLISTO, M. et al. **Bases Conceituais para conservação e manejo de bacias hidrográficas**. Belo Horizonte: Cemig - Companhia Energética de Minas Gerais, 2019b.

CALLISTO, M. et al. Beta diversity of aquatic macroinvertebrate assemblages associated with leaf patches in neotropical montane streams. **Ecology and Evolution**, v. 11, n. 6, p. 2551–2560, 7 mar. 2021.

CASTRO, D. M. P. et al. Beta diversity of aquatic invertebrates increases along an altitudinal gradient in a Neotropical mountain. **Biotropica**, v. 51, n. 3, p. 399–411, maio 2019.

CHEN, K. et al. Science of the Total Environment Incorporating functional traits to enhance multimetric index performance and assess land use gradients. **Science of the Total Environment**, v. 691, p. 1005–1015, 2019.

DAVY-BOWKER, J. et al. A comparison of the European Water Framework Directive physical typology and RIVPACS-type models as alternative methods of establishing reference conditions for benthic macroinvertebrates. **Hydrobiologia**, v. 566, n. 1, p. 91–105, 2006.

DORMANN, C. F. et al. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. **Ecography**, v. 36, n. 1, p. 027–046, 2013.

ELIAS, C. L. et al. Predicting reference conditions for river bioassessment by incorporating boosted trees in the environmental filters method. **Ecological Indicators**, v. 69, n. 10, p. 239–251, 2016.

ESRI, 2016. ArcGIS for Desktop version 10.4. Accessed 2 feb 2021

FEIO, M. J. et al. A predictive model for freshwater bioassessment (Mondego River, Portugal). **Hydrobiologia**, v. 589, n. 1, p. 55–68, 2007a.

FEIO, M. J. et al. Diatoms and macroinvertebrates provide consistent and complementary information on environmental quality. **Fundamental and Applied Limnology**, v. 169, n. 3, p. 247–258, 2007b.

FEIO, M. J. et al. Defining and Testing Targets for the Recovery of Tropical Streams Based on Macroinvertebrate Communities and Abiotic Conditions. **River Research and Applications**, v. 31, n. 1, p. 70–84, jan. 2015.

FEIO, M. J. et al. The Biological Assessment and Rehabilitation of the World's Rivers: An Overview. **Water**, v. 13, n. 3, p. 1–46, 31 jan. 2021.

FEIO, M. J. et al. Fish and macroinvertebrate assemblages reveal extensive degradation of the world's rivers. **Global Change Biology**, v. 00, p. 1–20, 17 jan. 2022.

FEIO, M. J.; PINTO, P. Tipologia e cenários biológicos do elemento macroinvertebrados. **Recursos Hídricos**, v. 30, n. 2, p. 19–27, 2009.

FEIO, M. J.; POQUET, J. M. Predictive Models for Freshwater Biological Assessment: Statistical Approaches, Biological Elements and the Iberian Peninsula Experience: A Review. **International Review of Hydrobiology**, v. 96, n. 4, p. 321–346, 2011.

FERREIRA, H. L. M. et al. **Ambientes Aquáticos em Minas Gerais**. 1a edição ed. Belo Horizonte: Prodemge, 2017.

FERRÉOL, M. et al. A top-down approach for the development of a stream typology based on abiotic variables. **Hydrobiologia**, v. 551, n. 1, p. 193–208, 2005.

FIRMIANO, K. R. et al. Mayfly bioindicator thresholds for several anthropogenic disturbances in neotropical savanna streams. **Ecological Indicators**, v. 74, p. 276–284, 1 mar. 2017.

FUSTER, R. et al. Water bodies typology system: a chilean case of scientific stakeholders and policy makers dialogue. **Water resources and wetlands**. Anais...2012

GARUANA, L. et al. Integração de indicadores ecológicos, ambientais e de saúde humana em microbacias urbanas Luziana. **Revista Espinhaço**, v. 9, n. 1, p. 1–16, 2020.

GUTIÉRREZ, J. D. et al. Segment scale typology for colombian andean rivers. **Ecohydrology and Hydrobiology**, v. 17, n. 2, p. 125–133, 2017.

HAMADA, N., NESSIMIAN, J.L.; QUERINO, R.B. **Insetos aquáticos na Amazônia brasileira: Taxonomia, biologia e ecologia**. Editora INPA, Manaus, Brazil, 2014.

HARGETT, E. G. et al. Development of a RIVPACS-type predictive model for bioassessment of wadeable streams in Wyoming. **Ecological Indicators**, v. 7, n. 4, p. 807–826, 2007.

HERLIHY, A. T. et al. The relation of lotic fish and benthic macroinvertebrate condition indices to environmental factors across the conterminous USA. **Ecological Indicators**, v. 112, n. July 2019, p. 105958, 2020.

HUGHES, R. M. et al. Acquiring data for large aquatic resource surveys : the art of compromise among science , logistics , and reality. **Journal of the North American Benthological Society**, v. 27, n. 4, p. 837–859, 2008.

IGAM, Instituto Mineiro de Gestão das Águas. **Gestão e situação das águas de Minas Gerais**. [s.l: s.n.].

JUNQUEIRA, M. V. et al. Índices bióticos para avaliação de qualidade de água de rios tropicais – síntese do conhecimento e estudo de caso: Bacia do alto rio doce. **Revista Brasileira de Ciências Ambientais** (Online), n. 49, p. 15–33, 2018.

JUSIK, S. et al. Development of comprehensive river typology based on macrophytes in the mountain-lowland gradient of different Central European ecoregions. **Hydrobiologia**, v. 745, n. 1, p. 241–262, 18 fev. 2015.

KARR, J. R. Seven Foundations of Biological Monitoring and Assessment. **Biologia Ambientale**, v. 20, n. 2, p. 7–18, 2006.

KAUFMANN, P. R. et al. Physical habitat in conterminous US streams and rivers , Part 1 : Geoclimatic controls and anthropogenic alteration. **Ecological Indicators**, v. 141, n. June, p. 109046, 2022.

LIGEIRO, R. et al. Defining quantitative stream disturbance gradients and the additive role of habitat variation to explain macroinvertebrate taxa richness. **Ecological Indicators**, v. 25, p. 45–57, 2013.

LINARES, M. S. et al. Chronic urbanization decreases macroinvertebrate resilience to natural disturbances in neotropical streams. **Current Research in Environmental Sustainability**, v. 3, p. 100095, 2021.

LORENZ, A.; FELD, C. K.; HERING, D. Typology of streams in Germany based on benthic invertebrates: Ecoregions, zonation, geology and substrate. **Limnologica**, v. 34, n. 4, p. 379–389, 2004.

MACEDO, D. R. et al. The relative influence of catchment and site variables on fish and macroinvertebrate richness in cerrado biome streams. **Landscape Ecology**, v. 29, n. 6, p. 1001–1016, 2014.

MACEDO, D. R. et al. Development of a benthic macroinvertebrate multimetric index (MMI) for Neotropical Savanna headwater streams. **Ecological Indicators**, v. 64, p. 132–141, 2016.

MACEDO, D. R. et al. Development and validation of an environmental fragility index (EFI) for the neotropical savannah biome. **Science of The Total Environment**, v. 635, p. 1267–1279, set. 2018.

MACEDO, D. R. et al. Urban stream rehabilitation in a densely populated Brazilian metropolis. **Frontiers in Environmental Science**, v. 10, n. September, p. 921934, set. 2022.

MARTINS, I. et al. Regionalisation is key to establishing reference conditions for neotropical savanna streams. **Marine and Freshwater Research**, v. 69, n. 1, p. 82–94, 2018.

MARTINS, I. et al. Are multiple multimetric indices effective for assessing ecological condition in tropical basins? **Ecological Indicators**, v. 110, n. May, p. 105953, 2020.

MELLO, K. DE et al. Multiscale land use impacts on water quality: Assessment, planning, and future perspectives in Brazil. **Journal of Environmental Management**, v. 270, n. June, p. 110879, 2020.

MERRITT, R.W., CUMMINS K.W. **An introduction to the aquatic insects of North America**. Kendall/Hunt Publishing Company, Dubuque, Iowa, USA, 1996.

MINAS GERAIS. Conselho Estadual de Política Ambiental – COPAM. **Resolução Normativa Conjunta COPAM/CERH-MG nº 08 de 21 de novembro de 2022**. Dispõe sobre a classificação dos corpos de água e diretrizes ambientais para o seu enquadramento, bem como estabelece as condições e padrões de lançamento de efluentes. Diário do Executivo. Minas Gerais. 2022.

MOOG, O. et al. Does the ecoregion approach support the typological demands of the EU “Water Framework Directive”? **Hydrobiologia**, v. 516, n. 1–3, p. 21–33, 2004.

MOYA, N. et al. Macroinvertebrate-based multimetric predictive models for evaluating the human impact on biotic condition of Bolivian streams. **Ecological indicators**, v. 11, p. 840–847, 2011.

MUGNAI, R., NESSIMIAN, J.L., BAPTISTA, D.F. **Guide for the Identification of Aquatic Macroinvertebrates of Rio de Janeiro State**. Technical Books Editora, Rio de Janeiro, 2009.

MUGNAI, R., NESSIMIAN, J.L., BAPTISTA, D.F. **Manual de Identificação de Macroinvertebrados Aquáticos do Estado do Rio de Janeiro**. Technical Books Editora, Rio de Janeiro, 2010.

OLSEN, A. R. et al. Survey design and extent estimates for the Wadeable Streams Assessment. **Journal of the North American Benthological Society**, v. 27, n. 4, p. 822–836, 2008.

OMERNIK, J. M. et al. How Misapplication of the Hydrologic Unit Framework Diminishes the Meaning of Watersheds. **Environmental Management**, v. 60, p. 1–11, 2017.

PARDO, I. et al. The European reference condition concept: A scientific and technical approach to identify minimally-impacted river ecosystems. **Science of the Total Environment**, v. 420, p. 33–42, 2012.

PARDO, I. et al. An invertebrate predictive model (NORTI) for streams and rivers: Sensitivity of the model in detecting stress gradients. **Ecological Indicators**, v. 45, p. 51–62, 2014.

PÉREZ, G.R. **Guía para el estudio de los macroinvertebrados acuáticos del Departamento de Antioquia**. Universidad de Antioquia Facultad de Ciencias Exactas y Naturales Centro de Investigaciones, CIEN. Medelín, 1988.

PÉS, A.M.O.; HAMADA, N., NESSIMIAN, J.L. Identification keys of larvae for families and genera of Trichoptera (Insecta) of Central Amazonia, Brazil. **Revista Brasileira de Entomologia** 49, 181–204, 2005.

PERO, E. J. I. et al. Ecoregions, climate, topography, physicochemical, or a combination of all: Which criteria are the best to define river types based on abiotic variables and macroinvertebrates in neotropical rivers? **Science of the Total Environment**, v. 738, p. 140303, 2020.

PONT, D. et al. Assessing river biotic condition at a continental scale : a European approach using functional metrics and fish. **Journal of Applied Ecology**, v. 43, p. 70–80, 2006.

REYNAGA, M. C.; DOS SANTOS, D. A. Rasgos biológicos de macroinvertebrados de ríos subtropicales: patrones de variación a lo largo de gradientes ambientales espacio-temporales. **Ecol. Austral**, v. 22, n. August, p. 112–120, 2012.

SANTOS, J. I. et al. Challenges to water quality assessment in Europe – Is there scope for improvement of the current Water Framework Directive bioassessment scheme in rivers? **Ecological Indicators**, v. 121, p. 107030, fev. 2021.

SILVA, D. R. O. et al. An improved macroinvertebrate multimetric index for the assessment of wadeable streams in the neotropical savanna. **Ecological Indicators**, v. 81, n. June, p. 514–525, 2017.

SILVA, D. R. O. et al. Assessing the extent and relative risk of aquatic stressors on stream macroinvertebrate assemblages in the neotropical savanna. **Science of the Total Environment**, v. 633, p. 179–188, 2018.

SOLHEIM, A. L. et al. A new broad typology for rivers and lakes in Europe: Development and application for large-scale environmental assessments. **Science of the Total Environment**, v. 697, p. 134043, 2019.

SOUZA, C. M. et al. Reconstructing three decades of land use and land cover changes in Brazilian biomes with Landsat archive and Earth engine. **Remote Sensing**, v. 12, n. 17, p. 2735, ago. 2020.

STODDARD, J. L. et al. Setting expectations for the ecological condition of streams: The concept of reference condition. **Ecological Applications**, v. 16, n. 4, p. 1267–1276, 2006.

TAMVAKIS, A. et al. Optimizing biodiversity prediction from abiotic parameters. **Environmental Modelling and Software**, v. 53, p. 112–120, 2014.

TOMANOVA, S.; MOYA, N.; OBERDORFF, T. Using macroinvertebrate biological traits for assessing biotic integrity of neotropical streams. **River Research and Applications**, v. 24, n. 9, p. 1230–1239, 2008.

TUPINAMBÁS, T. H. et al. Taxonomy, metrics or traits? Assessing macroinvertebrate community responses to daily flow peaking in a highly regulated Brazilian river system. **Ecohydrology**, v. 7, n. 2, p. 828–842, 2014.

VANNOTE, R. L. et al. The river continuum concept. **Canadian Journal of Fisheries and Aquatic Sciences**, v. 37, n. 1, p. 130–137, jan. 1980.

VERDONSCHOT, P. F. M.; NIJBOER, R. C. Testing the European stream typology of the Water Framework Directive for macroinvertebrates. **Hydrobiologia**, v. 516, n. 1–3, p. 35–54, 2004.

WARCHALSKA-TROLL, A.; WARCHALSKI, T. The selection of areas for case study research in socio-economic geography with the application of k-means clustering. Wiadomości Statystyczne. **The Polish Statistician**, v. 67, n. 2, p. 1–20, 28 fev. 2022.

WHITTIER, T. R. et al. Selecting reference sites for stream biological assessments: best professional judgment or objective criteria. **Journal of the North American Benthological Society**, v. 26, n. 2, p. 349–360, jun. 2007.

WIGGINS, G.B. Larvae of the North American Caddisfly Genera (Trichoptera), 2nd edn. **University of Toronto Press**: Toronto, Canada, 1996.

WILSON, M. F. J. et al. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. **Marine Geodesy**. [s.l: s.n.].

2.9. Appendix A. Supplementary data

2.9.1. Supplementary information 1

Table S1 - Total number of sites in the study area.

Hydrologic unit	River basin	No. of sample sites
São Francisco	Pandeiros	46
	Jequitaí and Pacuí	5
	Três Marias Reservoir	31
	Pará	16
	São Francisco	5
	Paraopeba	10
	das Velhas	101
Mata Atlântica	Jequitinhonha	2
	Araçuiaí	4
	Piracicaba	13
	Santo Antônio	1
	Piranga	4
Alto Paraná	Araguari	64
	Paranaíba	19
	Grande	27

Table S2 - List of variables with their relative units and sources.

Variable	Source
Annual Mean Temperature (°C)	Worldclim Project (https://worldclim.org/)
Mean Diurnal Range ((max. temp – min. temp)) (°C)	Worldclim Project (https://worldclim.org/)
Isothermality	Worldclim Project (https://worldclim.org/)
Temperature Seasonality (standard deviation ×100)	Worldclim Project (https://worldclim.org/)
Max. Temperature of Warmest Month	Worldclim Project (https://worldclim.org/)
Min. Temperature of Coldest Month	Worldclim Project (https://worldclim.org/)
Temperature Annual Range (°C)	Worldclim Project (https://worldclim.org/)
Mean Temperature of Wettest Quarter (°C)	Worldclim Project (https://worldclim.org/)
Mean Temperature of Driest Quarter (°C)	Worldclim Project (https://worldclim.org/)
Mean Temperature of Warmest Quarter (°C)	Worldclim Project (https://worldclim.org/)
Mean Temperature of Coldest Quarter (°C)	Worldclim Project (https://worldclim.org/)
Annual precipitation (mm)	Worldclim Project (https://worldclim.org/)
Precipitation of Wettest Month (mm)	Worldclim Project (https://worldclim.org/)
Precipitation of Driest Month (mm)	Worldclim Project (https://worldclim.org/)
Precipitation Seasonality (Coefficient of Variation) (mm)	Worldclim Project (https://worldclim.org/)
Precipitation of Wettest Quarter (mm)	Worldclim Project (https://worldclim.org/)
Precipitation of Driest Quarter (mm)	Worldclim Project (https://worldclim.org/)
Precipitation of Warmest Quarter (mm)	Worldclim Project (https://worldclim.org/)
Precipitation of Coldest Quarter (mm)	Worldclim Project (https://worldclim.org/)
Altitude (m)	Shuttle Radar Topographic Mission – SRTM
Mean river basin altitude (m)	Shuttle Radar Topographic Mission – SRTM
Total river basin area (km²)	Shuttle Radar Topographic Mission – SRTM
% Land use	MapBiomas (2021) (https://mapbiomas.org/)
% Forest	MapBiomas (2021) (https://mapbiomas.org/)
% Savanna	MapBiomas (2021) (https://mapbiomas.org/)
% Pasture	MapBiomas (2021) (https://mapbiomas.org/)
% Agriculture	MapBiomas (2021) (https://mapbiomas.org/)
% Urban infrastructure	MapBiomas (2021) (https://mapbiomas.org/)
% Mining	MapBiomas (2021) (https://mapbiomas.org/)
% Water bodies	MapBiomas (2021) (https://mapbiomas.org/)
% Anthropic use	MapBiomas (2021) (https://mapbiomas.org/)
Total Phosphorus (mg/L)	Analyzed in the laboratory
Total Nitrogen (mg/L)	Analyzed in the laboratory
Turbidity (NTU)	Field measurement
Lithological synthesis	Ferreira et al. (2017)
Terrain roughness index	Wilson et al. (2007)

Table S3 - Spearman rank correlations for river typology variables.

	Annual Mean Temperature (ºC)*	Mean Diurnal Range (max. temp – min. temp) (ºC)	Isothermality	Temperature Seasonality (standard deviation ×100)	Max. Temperature of Warmest Month	Min. Temperature of Coldest Month	Temperature Annual Range (ºC)	Mean Temperature of Wettest Quarter (ºC)	Mean Temperature of Driest Quarter (ºC)	Mean Temperature of Warmest Quarter (ºC)	Mean Temperature of Coldest Quarter (ºC)	Annual precipitation (mm)*	Precipitation of Wettest Month (mm)	Precipitation of Driest Month (mm)	Precipitation Seasonality (Coefficient of Variation) (mm)	Precipitation of Wettest Quarter (mm)	Precipitation of Driest Quarter (mm)	Precipitation of Warmest Quarter (mm)	Precipitation of Coldest Quarter (mm)	Altitude (m)*	Terrain roughness index*	Land use	Lithological synthesis*
Annual Mean Temperature (ºC)*	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Mean Diurnal Range (max. temp – min. temp) (ºC)	0.28	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Isothermality	0.59	0.28	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Temperature Seasonality (standard deviation ×100)	-0.60	0.02	-0.85	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Max. Temperature of Warmest Month	0.95	0.46	0.45	-0.45	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Min. Temperature of Coldest Month	0.91	-0.10	0.60	-0.74	0.79	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Temperature Annual Range (ºC)	-0.09	0.82	-0.31	0.53	0.18	-0.46	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Mean Temperature of Wettest Quarter (ºC)	0.97	0.29	0.42	-0.41	0.94	0.83	0.03	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Mean Temperature of Driest Quarter (ºC)	0.99	0.22	0.67	-0.72	0.91	0.94	-0.19	0.92	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Mean Temperature of Warmest Quarter (ºC)	0.96	0.26	0.38	-0.39	0.95	0.85	0.02	0.99	0.91	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-
Mean Temperature of Coldest Quarter (ºC)	0.99	0.23	0.67	-0.72	0.91	0.94	-0.18	0.92	1.00	0.92	1.00	-	-	-	-	-	-	-	-	-	-	-	-
Annual precipitation (mm)*	-0.56	-0.10	-0.44	0.44	-0.58	-0.63	0.16	-0.51	-0.55	-0.55	-0.57	1.00	-	-	-	-	-	-	-	-	-	-	-
Precipitation of Wettest Month (mm)	-0.49	0.03	-0.25	0.28	-0.51	-0.58	0.18	-0.49	-0.47	-0.53	-0.49	0.92	1.00	-	-	-	-	-	-	-	-	-	-
Precipitation of Driest Month (mm)	-0.41	-0.39	-0.67	0.61	-0.38	-0.35	0.01	-0.26	-0.45	-0.25	-0.47	0.40	0.07	1.00	-	-	-	-	-	-	-	-	-
Precipitation Seasonality (Coefficient of Variation) (mm)	0.24	0.45	0.57	-0.47	0.25	0.16	0.11	0.12	0.28	0.09	0.30	-0.25	0.12	-0.94	1.00	-	-	-	-	-	-	-	-
Precipitation of Wettest Quarter (mm)	-0.51	0.04	-0.28	0.32	-0.53	-0.61	0.21	-0.50	-0.50	-0.54	-0.51	0.95	0.98	0.12	0.06	1.00	-	-	-	-	-	-	-
Precipitation of Driest Quarter (mm)	-0.41	-0.38	-0.68	0.62	-0.37	-0.35	0.03	-0.26	-0.46	-0.25	-0.47	0.39	0.05	0.99	-0.94	0.11	1.00	-	-	-	-	-	-
Precipitation of Warmest Quarter (mm)	-0.61	-0.11	-0.46	0.58	-0.64	-0.68	0.17	-0.51	-0.63	-0.57	-0.65	0.86	0.77	0.41	-0.24	0.80	0.40	1.00	-	-	-	-	-
Precipitation of Coldest Quarter (mm)	-0.36	-0.36	-0.58	0.54	-0.37	-0.32	-0.01	-0.22	-0.39	-0.24	-0.42	0.50	0.16	0.94	-0.90	0.22	0.95	0.50	1.00	-	-	-	-
Altitude (m)*	-0.84	-0.16	-0.16	0.19	-0.85	-0.74	-0.05	-0.90	-0.77	-0.93	-0.77	0.49	0.53	0.07	0.08	0.53	0.07	0.49	0.08	1.00	-	-	-
Terrain roughness index*	-0.35	-0.36	-0.50	0.34	-0.29	-0.21	-0.08	-0.30	-0.37	-0.24	-0.36	0.03	-0.09	0.40	-0.41	-0.07	0.41	0.05	0.30	0.10	1.00	-	-
Land use	-0.03	-0.05	-0.10	0.11	-0.04	-0.04	0.01	0.01	-0.04	0.00	-0.05	0.11	0.05	0.16	-0.15	0.06	0.15	0.12	0.17	-0.04	0.00	1.00	-
Lithological synthesis*	0.24	0.25	0.37	-0.32	0.21	0.18	0.02	0.16	0.26	0.15	0.27	-0.28	-0.14	-0.46	0.43	-0.16	-0.45	-0.28	-0.43	-0.06	-0.31	-0.08	1.00

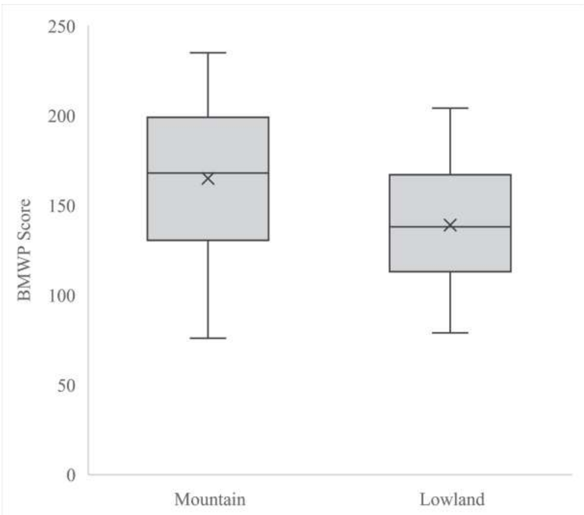
*In green are the variables retained in the database used in river typology construction. In red variables with > |0.7|.

Table S4 - Pseudo F-Statistic Summary.

Number of Groups	Mean	Minimum	Maximum	Median
2	**38,836.90**	**38,836.90**	**38,836.90**	**38,836.90**
3	32,790.35	32,790.35	32,790.36	32,790.35
4	30,320.85	30,320.63	30,320.93	30,320.85
5	28,771.35	28,701.74	28,840.96	28,771.34
6	28,152.37	27,556.94	28,301.31	28,301.31
7	27,371.04	26,906.70	27,680.67	27,680.46
8	26,633.83	24,752.74	26,913.04	26,910.31
9	25,754.86	24,837.60	26,538.79	25,380.53
10	25,193.58	23,942.33	25,722.46	25,419.03
11	24,728.47	24,421.52	24,810.33	24,807.72
12	23,935.06	23,847.21	24,136.37	23,893.13
13	23,437.43	23,232.52	23,519.11	23,481.96
14	22,890.68	22,601.25	23,070.09	22,965.27
15	22,473.83	22,225.44	22,586.82	22,553.47

In bold the selected number of groups used in river typology construction.

Figure S1 - Box plot showing the BMWP score for reference sites in the mountain and lowland river types.



2.9.2. Supplementary information 2

Table S1 - Lithological synthesis of Minas Gerais according to Ferreira et al. (2017).

Group	Description
Siliceous rocks (S)	The siliceous group includes rocks whose chemical composition has silica ($SiO2$) as its main component, such as acidic and intermediate igneous rocks, with more than 52% silica. They include sandy detrital sedimentary rocks, such as quartz-arenites and subarchoses, as well as rich conglomerates and fragments of quartz-arenites and acidic and intermediate igneous rocks. Metamorphic equivalent rocks are also part of it.
Pelitic rocks (P)	Detrital sedimentary rocks formed by fragments in the mud fraction, such as pelites and their metamorphic equivalents.
Metamorphic rocks (F)	Consisting of rocks of igneous and sedimentary origin. Silica content below 52%, which exhibit intermediate to high-grade metamorphism. In this group are Archean and Paleoproterozoic rocks of similar composition.
Carbonate rocks (C)	Sedimentary rocks with a chemical composition rich in calcium, such as limestone and dolomites, belong to the carbonate group.
Volcanic rocks (B)	They consist of basic rocks, mainly extrusive ones, formed by spills. It includes intrusive outcrop rocks of basic composition and their low-grade metamorphic equivalents.
Alkaline rocks (A)	Rocks rich in alkalis, with minerals such as feldspathoids and sodium amphiboles. They include alkaline syenites, phonolites, and dunites. They usually form rocky bodies of small regional expression whose distribution in Minas Gerais territory is restricted to a few occurrences such as Poços de Caldas-MG.
Laterized sediments (L)	Sediments of alluvial, colluvial, and eluvial origin, usually cemented by oxides and hydroxides of iron and aluminum, with occurence in extensive plateaus and some plains.
Unconsolidated sediments (I)	Incohesive sandy and muddy sediments occur along the alluvial plains and terraces.

Table S2 - River typology, hydrologic units, landscape units, lithological synthesis, and area classes.

Hydrologic unit / Landscape unit	River type code	No.Reference sites	Lithological synthesis / Area classes	River basin area			
				0 to 100 km²	100 to 1,000 km²	1,000 to 10,000 km²	> 10,000 km²
São Francisco hydrologic unit (number of drainage sections of each river typology)							
Mountain	F1B	-	Volcanic rocks	68	2	-	-
	F1C	-	Carbonate rocks	296	73	14	-
	F1F	24	Metamorphic rocks	251	59	25	-
	F1I	-	Unconsolidated sediments	101	170	33	2
	F1L	-	Laterized sediments	130	14	-	-
	F1P	5	Pelitic rocks	1917	423	184	-
	F1S	8	Siliceous rocks	3805	642	194	-
Lowland	F2C	-	Carbonate rocks	1466	447	138	-
	F2F	-	Metamorphic rocks	31	5	-	-
	F2I	16	Unconsolidated sediments	786	613	401	427
	F2L	-	Laterized sediments	1014	144	34	50
	F2P	1	Pelitic rocks	3132	954	299	136
	F2S	8	Siliceous rocks	3001	548	169	11
Mata Atlântica hydrologic unit							
Mountain	F1C	-	Carbonate rocks	1	-	-	-
	F1F	3	Metamorphic rocks	129	30	8	-
	F1I	-	Unconsolidated sediments	145	98	64	-
	F1L	-	Laterized sediments	134	11	-	-
	F1P	-	Pelitic rocks	544	104	7	-
	F1S	5	Siliceous rocks	4495	832	217	-
Lowland	F2C	-	Carbonate rocks	3	-	-	-
	F2I	-	Unconsolidated sediments	458	331	213	189
	F2L	-	Laterized sediments	398	35	19	-
	F2P	-	Pelitic rocks	1549	458	241	87
	F2S	-	Siliceous rocks	6469	1315	560	173
Alto Paraná hydrologic unit							
Mountain	F1A	-	Alkaline rocks	73	8	-	-
	F1B	-	Volcanic rocks	625	195	107	41
	F1C	-	Carbonate rocks	18	12	9	2
	F1F	-	Metamorphic rocks	67	27	15	-
	F1I	-	Unconsolidated sediments	76	222	116	1
	F1L	-	Laterized sediments	119	29	-	6
	F1P	-	Pelitic rocks	770	197	106	28
	F1S	17	Siliceous rocks	8021	1704	486	240
Lowland	F2B	-	Volcanic rocks	612	321	248	192
	F2I	-	Unconsolidated sediments	13	5	3	11
	F2L	-	Laterized sediments	1	2	-	-
	F2S	-	Siliceous rocks	1967	266	58	130
Paraíba do Sul hydrologic unit							
Mountain	F1A	-	Alkaline rocks	5	-	-	-
	F1F	-	Metamorphic rocks	4	1	-	-
	F1I	-	Unconsolidated sediments	3	2	-	-
	F1S	-	Siliceous rocks	1197	232	56	-
Lowland	F2S	-	Siliceous rocks	533	152	117	19

Mountain: rivers predominantly located on higher elevations (average of 858 m), with generally more precipitation (average of 1,459 mm) and lower temperatures (average of 20.1 ºC). Lowland: rivers predominantly located in lower areas (average of 542 m) with lower rainfall (average of 1,123 mm) and warmer climates (average of 22.6 ºC).

Table S3 - ANOSIM test for river types indicating significant differences (in red) for benthic macroinvertebrates (global R = 0.4377, p=0.001).

	F1S	F2I	F2S	F1P	F1F
F1S	-	-	-	-	-
F2I	0.0001	-	-	-	-
F2S	0.0001	0.0117	-	-	-
F1P	0.0797	0.0003	0.0023	-	-
F1F	0.0001	0.0001	0.0001	0.0001	-

F1S and F2S (Siliceous rocks); F2I (Unconsolidated sediments); F1P (Pelitic rocks); and F1F (Metamorphic rocks).

Table S4 - Taxa contributing up to 90% of cumulative percentage to Bray-Curtis similarity for the river types, in descending order of contribution. In bold are the taxa exclusive from certain river types. River types: F1S and F2S (Siliceous rocks); F2I (Unconsolidated sediments); F1P (Pelitic rocks); and F1F (Metamorphic rocks).

Lowland		Mountain		
F2S	F2I	F1S	F1P	F1F
Taxa (%)		Taxa (%)		
Chironomidae (14.99)	Chironomidae (12.33)	Chironomidae (12.78)	Chironomidae (15.63)	Chironomidae (11.95)
Elmidae (9.54)	Elmidae (6.83)	Elmidae (8.55)	Leptophlebiidae (7.76)	Elmidae (6.65)
Ceratopogonidae (7.53)	Leptohyphidae (6.54)	Leptophlebiidae (7.18)	Elmidae (7.15)	Leptophlebiidae (6.61)
Bivalvia (5.44)	Ceratopogonidae (6.34)	Baetidae (6.55)	Coenagrionidae (6.17)	Baetidae (6.02)
Gomphidae (5.39)	Baetidae (5.5)	Simuliidae (5.7)	Libellulidae (5.62)	Perlidae (5.37)
Oligochaeta (5.14)	Leptoceridae (5.22)	Leptohyphidae (4.58)	Helicopsychidae (4.24)	Tipulidae (5.19)
Leptoceridae (4.83)	Gomphidae (4.95)	Hydropsychidae (4.09)	Leptoceridae (4.15)	**Gripopterygidae (4.8)**
Leptohyphidae (4.83)	Hydrobiosidae (4.93)	Coenagrionidae (3.38)	Calamoceratidae (3.99)	Ceratopogonidae (4.73)
Coenagrionidae (4.63)	Naucoridae (4.75)	Perlidae (3.3)	Odontoceridae (3.91)	Simuliidae (4.11)
Naucoridae (4.45)	Libellulidae (4.37)	Leptoceridae (3.23)	Polycentropodidae (3.8)	Oligochaeta (3.67)
Libellulidae (3.61)	Leptophlebiidae (4.32)	Ceratopogonidae (3.18)	Ceratopogonidae (3.78)	Hydropsychidae (3.27)
Leptophlebiidae (3.03)	Bivalvia (4.08)	Oligochaeta (3.14)	Hydroptilidae (3.44)	Leptohyphidae (3.11)
Hydrobiosidae (2.9)	Oligochaeta (3.42)	Tipulidae (2.97)	Oligochaeta (3.4)	Polycentropodidae (2.6)
Baetidae (2.66)	Empididae (3.41)	Hydroptilidae (2.4)	Baetidae (2.75)	Libellulidae (2.52)
Caenidae (2.63)	Helicopsychidae (3.2)	Libellulidae (2.39)	**Gerridae (2.34)**	**Megapodagrionidae (2.51)**
Pyralidae (2.35)	Hydroptilidae (2.68)	Polycentropodidae (1.96)	Caenidae (2.02)	Calamoceratidae (2.29)
Hydrobiidae (2.15)	Simuliidae (2.51)	Empididae (1.75)	Gomphidae (1.85)	Coenagrionidae (2.24)
Lutrochidae (1.77)	Pyralidae (2.44)	Dytiscidae (1.56)	Leptohyphidae (1.63)	Leptoceridae (1.84)
Dytiscidae (1.43)	Calopterygidae (1.7)	Pleidae (1.54)	Perlidae (1.59)	Gomphidae (1.77)
Calopterygidae (1.37)	Coenagrionidae (1.56)	Odontoceridae (1.53)	Hydropsychidae (1.37)	**Veliidae (1.57)**
		Calamoceratidae (1.37)	**Notonectidae (1.36)**	Corydalidae (1.56)
		Psephenidae (1.34)	Empididae (1.26)	Empididae (1.36)
		Naucoridae (1.22)	Psephenidae (1.18)	Lutrochidae (1.2)
		Euthyplociidae (1.16)		Bivalvia (1.11)
		Caenidae (1.12)		**Tabanidae (1.01)**
		Pyralidae (1.07)		**Planariidae (0.98)**
		Corydalidae (1.02)		

Figure S1 - Proportion of the river types in Minas Gerais. *River types validated by the benthic macroinvertebrate community.

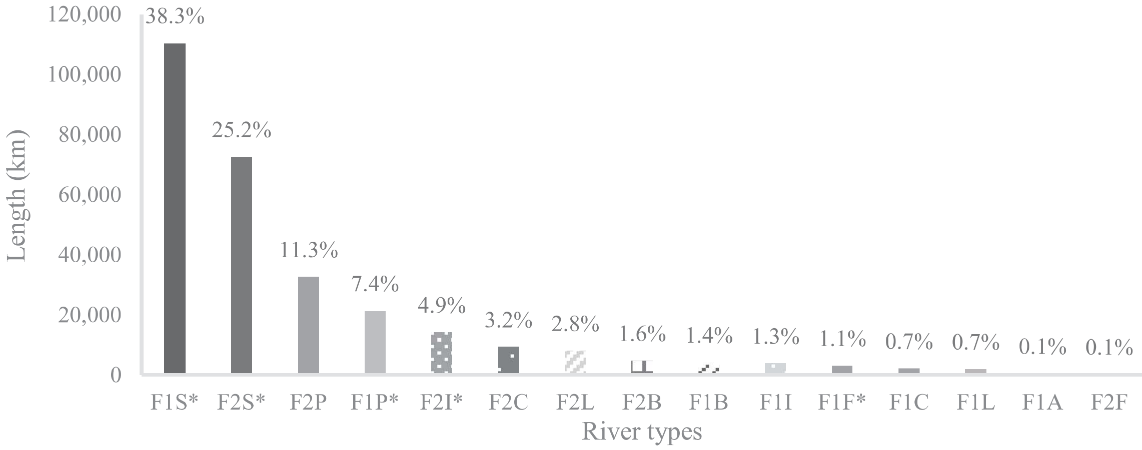


Figure S2 - Box plots showing river-type characteristics: altitude (m), average annual temperature, annual precipitation and terrain roughness.

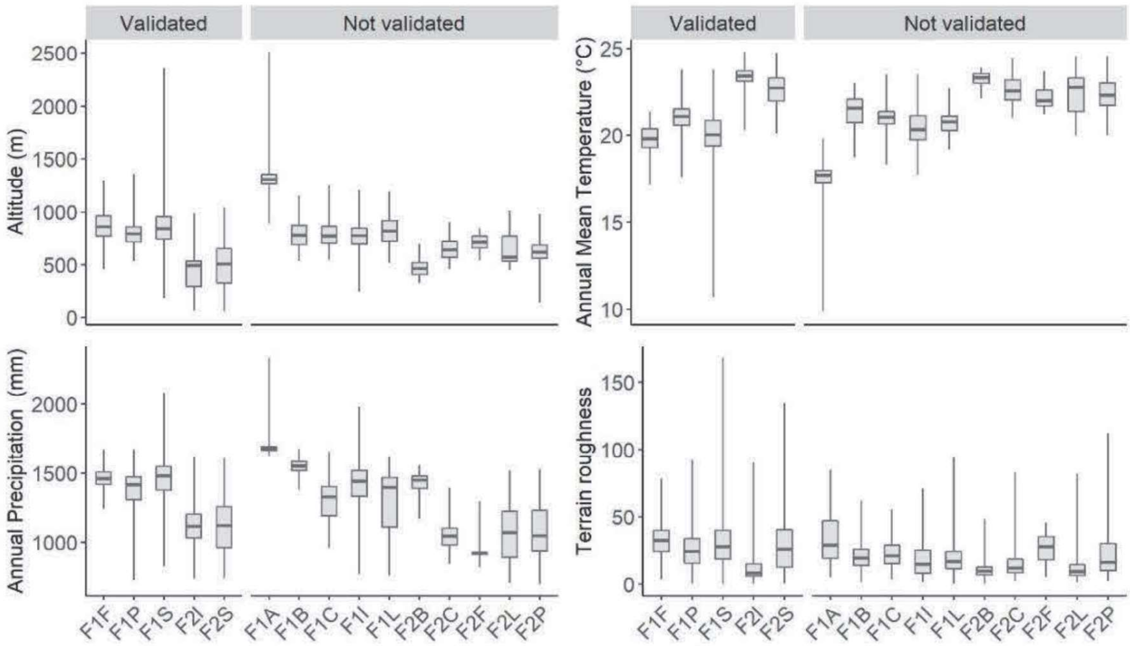
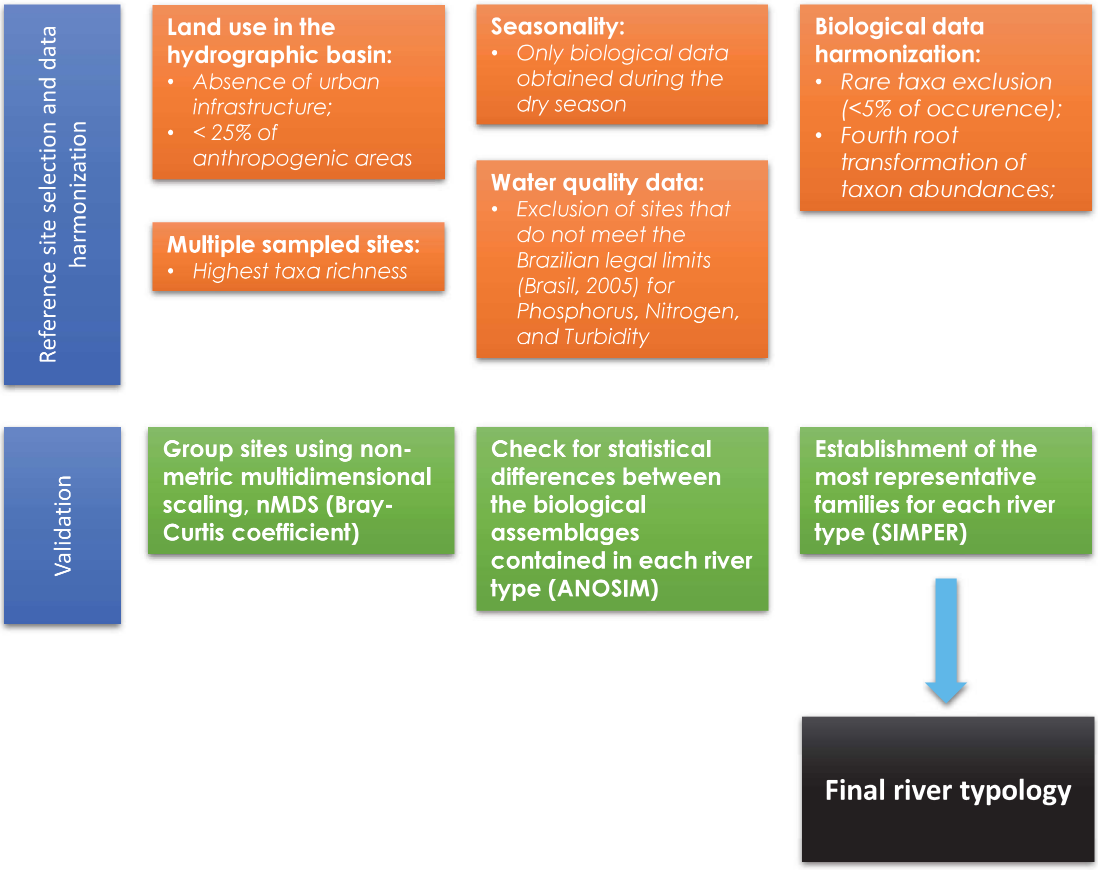


Figure S3 - Reference sites selection and river typology validation.



3. CHAPTER II - A NEW PREDICTIVE MODEL (MINASPACS) FOR SPATIALLY EXTENSIVE BIOLOGICAL ASSESSMENTS IN MINAS GERAIS

3.1. Abstract

Freshwater ecosystems are threatened by flow regulation, sedimentation, habitat degradation, introduction of non-native species, and poor sewage and wastewater treatment. These human pressures have led to a loss of biodiversity and habitats on a global scale. Therefore, it is essential to evaluate the ecological condition of freshwater ecosystems to promote effective management practices. Predictive models based on multivariate analyses are recognized ecological tools that can help monitor and manage freshwater ecosystems worldwide. Meanwhile, only a few studies have used this approach to assess tropical rivers and streams. By adopting existing approaches, such as the RIVPACS predictive model, effective biological assessment models can be develop for large tropical countries such as Brazil, aiming to support recent official recommendations regarding the determination of the ecological condition of water bodies. The primary aim of this study was to develop a RIVPACS-type model based on macroinvertebrate communities, called MINASPACS, for spatially extensive biological assessments of rivers in Minas Gerais, using the river basins of the southeastern Cerrado (neotropical savanna) as a case study. The second objective was to assess the sensitivity of the MINASPACS to the stressors affecting the rivers of Minas Gerais state through the relative risk approach. The MINASPACS model was trained with biological and environmental data from 87 reference sites and showed good accuracy ($R^2 > 0.6$, SDO/E = 0.16). The % urban infrastructure, % anthropogenic use, water turbidity, Total Nitrogen and Total Phosphorus were stressors detected by MINASPACS which represented a risk to the biological condition of Minas Gerais rivers. Because of its accuracy, sensitivity and the ease of usage due to its implementation in Aquaweb platform and use of map-level predictor variables, our model provides a clear, simple and defensible measure of the biological condition of streams in a diverse landscape.

Keywords: Relative Risk Approach, benthic macroinvertebrates, streams, freshwaters, ecological assessment.

3.2. Introduction

Freshwater ecosystems are among the most threatened by human pressures worldwide (REID et al., 2019). Flow regulation and longitudinal barriers (DUDGEON, 2010), sedimentation and habitat degradation (SANO et al., 2019), alien species invasion and poor sewage and wastewater treatment lead to biodiversity and habitat losses (FEIO et al., 2014). Given this scenario, assessing the ecological condition of freshwater ecosystems is critical for addressing efficient management practices (PAULSEN et al., 2016; SILVA et al., 2017). Several methodological approaches based on the use of aquatic organisms as bioindicators have been used in the biological assessment of freshwater ecosystems in North America, Europe and Australia, such as multimetric indices (e.g., HAWKINS et al., 2010; KARR, 1999), relative risk (RR) and relative extent (RE) approaches (e.g., VAN SICKLE; PAULSEN, 2008), and predictive models (e.g., CLARKE et al., 2003; FEIO et al., 2014; REYNOLDSON et al., 1997; WRIGHT, 1995). However, these approaches are not completely explored in other continents, such as Asia (e.g., BLAKELY; HARDING, 2010; CHEN et al., 2019) and South America (e.g., MARTINS et al., 2020; SILVA et al., 2018).

Some of the most recognized ecological tools to monitor and manage freshwater ecosystems are predictive models based on multivariate analyses (FEIO; POQUET, 2011; WRIGHT, 1995), which follow the concept of the Reference Condition Approach (HUGHES et al., 1986; REYNOLDSON et al., 1997; STODDARD et al., 2006). The River Invertebrate Prediction and Classification System (RIVPACS) (WRIGHT, 1995) was the first model of this kind and was developed for the United Kingdom. RIVPACS-type models make site-specific predictions of the benthic macroinvertebrate fauna expected without anthropogenic stressors. Those predictions are based on empirical relationships between individual taxon probabilities of capture and natural environmental features (e.g., latitude, substrate composition, alkalinity, elevation, etc.) derived from data collected from a reference site network (HARGETT et al., 2007). Since its first version, RIVPACS has evolved into a nation-wide bioassessment tool in the UK (WRIGHT, 1995) and was adapted to assess the biological condition of streams in Australia (AUSRIVAS by SMITH et al., 1999), Canada (BEAST, REYNOLDSON et al., 1997), Sweden (SWEPACSRI, JOHNSON; SANDIN, 2001), the USA (O/E, VAN SICKLE et al., 2005), the Czech Republic (PERLA, KOKEŠ et al., 2006), and Portugal (FEIO et al., 2009). In the USA, the RIVPACS-type approach and a probability survey allowed the conclusion that over 44% of the stream length in the conterminous USA have lost >20% of its common macroinvertebrate taxa (USEPA, 2016).

Several aspects of the RIVPACS approach were incorporated into the prescribed methods of the European Water Framework Directive - WFD (Directive 2000/60/EC) (WFD, 2000) for assessing the ecological quality and ecological status of European surface waters (CLARKE et al., 2003). Although several methods have been proposed and tested in temperate regions, few studies have used this approach for rivers and streams in the tropics, except at a single river catchment (e.g., MORENO et al., 2009) or for reservoirs (MOLOZZI et al., 2012). In Brazil, the Minas Gerais state (586,528 km²) recently established the use of quality classes to classify water bodies in terms of ecological condition as one of the stages of biomonitoring programs (Normative Deliberation COPAM/CERH-MG nº 008/2022; MINAS GERAIS, 2022). Using the experience of the European WFD, adopting and adapting existing approaches such as the predictive modeling could be extremely useful in large tropical countries such as Brazil, and contribute to fulfill the recent official requirements (BUSS et al., 2015).

Benthic macroinvertebrates are organisms with a wide geographic distribution and high taxa richness with different sensitivity levels. We hypothesized that aquatic environments with substantial anthropogenic stressors would simplify macroinvertebrate assemblages and that predictive modeling would represent this impairment. Therefore, our primary aim was to develop and test a multivariate model (MINASPACS) for spatially extensive biological assessments of rivers in Minas Gerais. Assessing the sensitivity of the MINASPACS to the major anthropogenic stressors affecting Minas Gerais rivers through the relative risk approach was the second aim (HERLIHY et al., 2020; SILVA et al., 2018).

3.3. Materials and methods

3.3.1. Study area and environmental characterization

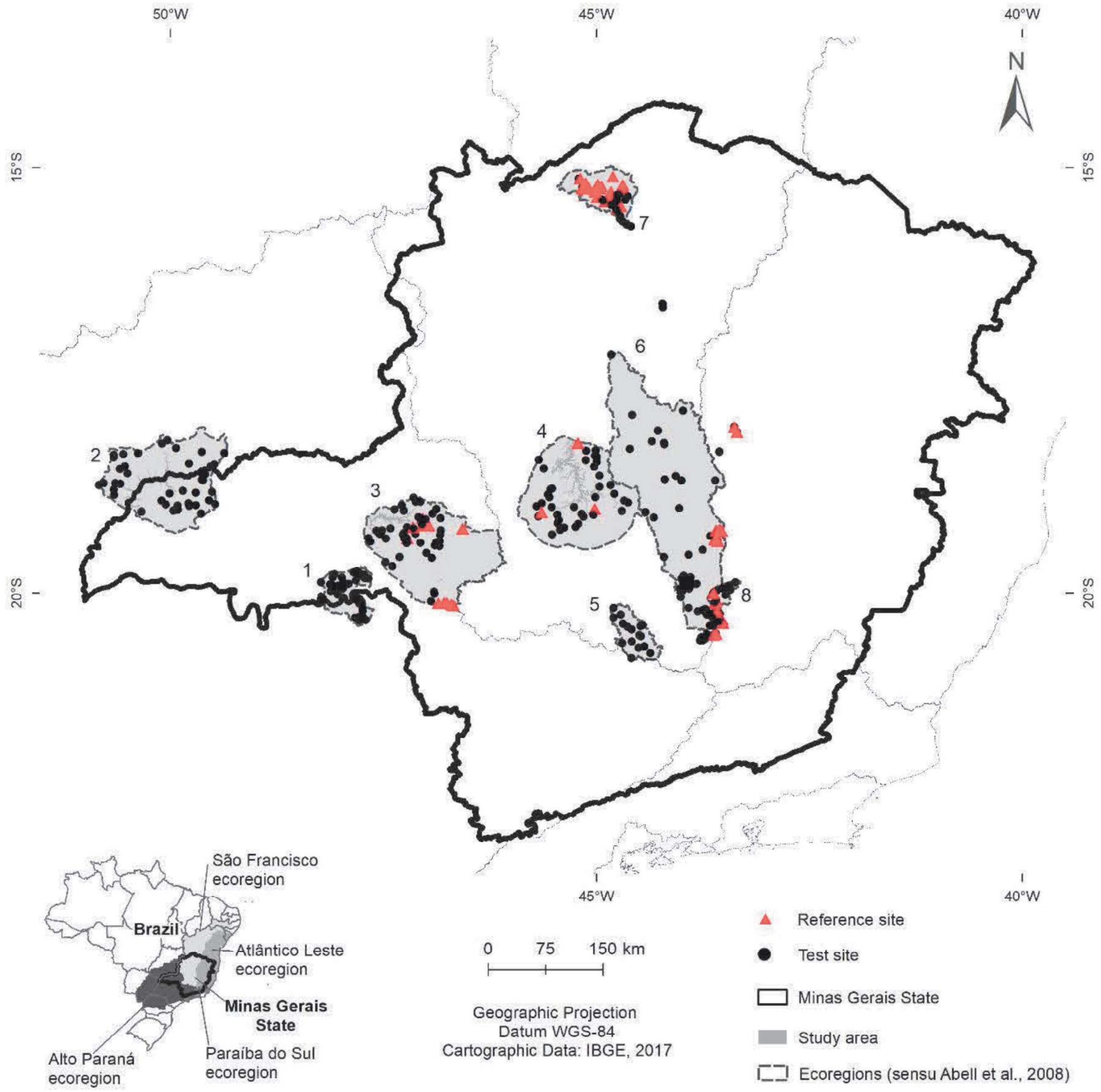
Between 2003 and 2019, benthic macroinvertebrates were collected in 348 stream sites in Minas Gerais from different research projects of the Laboratory of Ecology of Benthos-Universidade Federal de Minas Gerais (AGRA et al., 2019; CALLISTO et al., 2021; CASTRO et al., 2019; FEIO et al., 2015; FERREIRA et al., 2017; GARUANA et al., 2020; LINARES et al., 2021; MACEDO et al., 2022; MARTINS et al., 2020, 2018a; SILVA et al., 2017) and Serviço Nacional de Aprendizagem Industrial SENAI-MG (FERREIRA et al., 2017). Data from 20 stream sites in Goiás state and 13 in São Paulo state were also compiled (CALLISTO et al., 2019), covering a total area of 40,106 km² in 8 hydrological units: 1) Volta Grande Reservoir, 2) São Simão Reservoir, 3) Nova Ponte Reservoir, 4) Três Marias Reservoir, 5) Cajuru Reservoir, 6) das Velhas River, 7) Pandeiros River, and 8) Peti Reservoir (Figure 1). Analyses were conducted as follows: in the Alto Paraná hydrological unit (143 stream sites, grouping the hydrographic basins 1, 2 and 3); and for the São Francisco hydrological unit, 214 stream sites, grouping the hydrographic basins 3, 5, 6 and 7. Additionally, the Alto Paraná and São Francisco sites were grouped with Peti (8) reservoir river basin (13 sites) and sparse sites (3) in the Atlântico Leste hydrological unit, totalizing 381 sites. Each site was characterized according to its lithological group (Supplementary Information 1, Table S1), climate (50-year climatic reference, from Worldclim Project - https://worldclim.org/), and river basin characteristics (Table 1), corresponding to nine candidate variables for predictive model construction. Each river basin's land use proportions (six classes) were estimated from a Geographic Information System (GIS). Land use data were obtained from Collection 5 of the MapBiomas online platform (2021), with a spatial resolution of 30 meters (SOUZA et al., 2020) (Table 1).

Table 1 - Variables with their relative units and sources.

Variable	Source
Latitude (decimal degrees)	Measured on GIS
Longitude (decimal degrees)	Measured on GIS
Annual Mean Temperature (ºC)	Worldclim Project (https://worldclim.org/)
Annual Mean Precipitation (mm)	Worldclim Project (https://worldclim.org/)
Annual Temperature Range (ºC)	Worldclim Project (https://worldclim.org/)
Altitude (m)	Shuttle Radar Topographic Mission – SRTM
Mean catchment slope (%)	Shuttle Radar Topographic Mission – SRTM
Distance to source (m)	Measured on GIS
Lithological synthesis (1-8)*	Ferreira et al. (2017)
Forest (%)	Souza et al. (2020)
Savanna (%)	Souza et al. (2020)
Pasture (%)	Souza et al. (2020)
Agriculture (%)	Souza et al. (2020)
Urban infrastructure (%)	Souza et al. (2020)
Anthropogenic use (%)	Souza et al. (2020)
Catchment area (km²)	Souza et al. (2020)
Total Phosphorus (mg/L)	Analyzed in the laboratory
Total Nitrogen (mg/L)	Analyzed in the laboratory
Turbidity (NTU)	Field measurement

*1) Siliceous rocks; 2) Pelitic rocks; 3) Metamorphic rocks; 4) Carbonate rocks; 5) Volcanic rocks; 6) Alkaline rocks; 7) Laterized sediments; and 8) Unconsolidated sediments.

Figure 1 - Study area showing the river basins (gray polygons) and sampling sites (black dots and red triangles). 1) Volta Grande Reservoir, 2) São Simão Reservoir, 3) Nova Ponte Reservoir, 4) Três Marias Reservoir, 5) Cajuru Reservoir, 6) das Velhas River, 7) Pandeiros River, and 8) Peti Reservoir.



3.3.2. Biological samples and water quality data

Benthic macroinvertebrate assemblages' data collected between 2003 and 2019 in 381 stream sites were compiled. Each sample consisted of a composite sample from 3 to 20 Surber (30 x 30 cm, 500 mm mesh) or D-net samples (30 cm aperture, 500 mm mesh, and 0.09 m²) in the most representative habitats, then aggregated into one composite sample for each site. The samples were fixed in the field with 70% alcohol and deposited in the Reference Collection of Benthic Macroinvertebrates at the Institute of Biological Sciences at the Federal University of

Minas Gerais (CALLISTO et al., 2021) and the Center for Innovation and Technology SENAI – CIT. The samples were washed in sieves in 1.00 and 0.50 meshes in the laboratory. All individuals were identified mainly at the family level with the aid of taxonomic keys (PÉREZ, 1988; MERRITT; CUMMINS, 1996; WIGGINs, 1996; PÉS et al., 2005; MUGNAI et al., 2009, 2010; HAMADA et al., 2014). Only biological data obtained during the dry season (between May and September) were used, and in the case of sites sampled multiple times, only the record with the highest taxa richness was used. Water quality data (Total Phosphorus - mg/L, Total Nitrogen - mg/L, and turbidity - NTU) were also compiled for each site.

3.3.3. Reference sites selection

Screening sites is necessary to avoid the confounding effects of alterations to macroinvertebrate assemblages caused by anthropogenic disturbance instead of differences resulting from the different abiotic characteristics, such as geology or climate (STODDARD et al., 2006; WHITTIER et al., 2007). Therefore, reference sites were selected for predictive model development based on land use and water quality criteria (Table 2):

Table 2 - Criteria for restricting the data to near-natural streams.

Filter criterion	Threshold value	Source
Land use	< 25 % anthropogenic areas / absence of urban infrastructure in the hydrographic basin (LORENZ; FELD; HERING, 2004)	GIS data - MapBiomas (2021) (https://mapbiomas.org/)
Water quality	Exclude sites not meeting the federal limits for Phosphorus, Nitrogen, and Turbidity (SILVA et al., 2017)	CONAMA Resolution nº 357/2005, Class II*, lotic environments.

*Class II water quality corresponds mainly to water intended for human consumption after simplified treatment and protection of aquatic communities.

3.3.4. Predictive model construction (MINASPACS) and validation

The model training was done with biological and environmental data from the reference sites. Sites with fewer than 200 individuals were excluded, and rare taxa with less than 5% occurrence in the stream sites were previously excluded from further analyses. Macroinvertebrate relative abundances were a priori transformed by fourth root. Nine environmental variables (Table 1) were selected as candidate discriminant variables and were previously transformed to ensure normality and homoscedasticity: latitude ($\log_{x+1}$); longitude ($\log_{x+1}$); annual mean temperature (ºC); annual temperature range (ºC); annual mean precipitation (Sqrt) (mm); altitude ($\log_{x+1}$) (m); mean catchment slope ($\log_{x+1}$) (%); distance to source (m); and lithological synthesis. These variables were previously used in similar predictive models based on benthic invertebrate assemblages (e.g., FEIO et al., 2007; HARGETT et al., 2007; PARDO et al., 2014) because they are not easily influenced by anthropogenic activities and are known to reflect the natural distribution of biological assemblages in rivers.

To build the MINASPACS model the AQUAWEB online software was used (http://aquaweb.uc.pt/) (Figure S1). This tool follows the RIVPACS-type approach described in Van Sickle et al. (2006), which was previously used and validated with large datasets of macroinvertebrate assemblages (e.g., AGUIAR et al., 2011; MENDES et al., 2014). Building a RIVPACS-type model contains several steps (FEIO; POQUET, 2011). Briefly, the reference dataset was defined by a priori reference criteria representing the environmental variability present in the study area. Next, the reference sites were classified according to their faunal composition in similar biological groups through a clustering technique (Unweighted Pair Group Method with arithmetic mean, UPGMA) based on the Bray–Curtis similarity and supported by non-metric Multidimensional Scaling (nMDS) (MENDES et al., 2014). Groups had at least five reference sites to generate reliable predictions (WRIGHT, 1995). Later, the biological groups of reference sites and candidate discriminant variables were linked.

A Discriminant Function Analysis (DFA) was used to determine which environmental features best discriminate the biological groups and ranks them using F-tests and Wilks' lambda tests (MENDES et al., 2014). The DFA model produced discriminant functions that maximize the differences among reference biological groups. Then, each taxon occurrence probability at a site was calculated. The frequency of occurrence for each taxon in a reference group was averaged and weighted based on the site's probability of being assigned to that group through discriminant analysis. From this, the number of Observed taxa (O) at a site was divided by the sum of probabilities of occurrence of Expected taxa (E), up to 50% of probability, to obtain O/E50 ratios.

The model performance was assessed from O/E's mean value (MN) and standard deviation (SD) for calibration sites. The MNO/E (mean value of O/E) measures model bias and if its value is equal to one the predictive model is unbiased. The lower the SDO/E, the more precise is the model (MENDES et al., 2014; VAN SICKLE et al., 2005). The model with a high

F-statistic and low Wilks' λ value were targeted for model selection. Furthermore, the selection of the best model were given by its O/E regression ($R^2 \geq 0.5$; $p < 0.05$), intersection close to the origin (a range of −1.5 to 1.5 is acceptable), SDO/E < 0.2, and slope near 1 (acceptable range of 0.85 to 1.15) (LINKE et al., 2005; MENDES et al., 2014; VAN SICKLE et al., 2005).

In MINASPACS, sites were grouped into ecological status classes: high – 1, good – 2, moderate – 3, poor – 4 and bad – 5. The boundary between high and good classes was set at the 25th percentile of the calibration site O/E50 ratios, and the boundaries below were divided into four equal classes (MENDES et al., 2014). Finally, SIMPER statistical procedures (similarity/distance percentages, fourth root transformation, Bray-Curtis coefficient, Primer 6) were used to determine the most representative families (up to 90% of cumulative percentage) of each faunal group created by MINASPACS model.

3.3.5. Sensitivity to stressors - assessing relative risk

Seven stressors were used to evaluate the sensitivity of MINASPACS to the stressors affecting Minas Gerais state. Total Phosphorus (mg/L), Total Nitrogen (mg/L), and turbidity (NTU) results obtained from our database were compiled. Furthermore, % pasture, % agriculture, % urban infrastructure, and % all anthropogenic uses combined for each site's catchment were acquired through geospatial tools. All possible situations of having good or poor macroinvertebrate O/E value given high or low stressor conditions were addressed. Because >50% of our sites were not selected via a probabilistic survey design (STEVENS; OLSEN, 2004), the relative risk approach was used without proportional weighting to estimate stream condition extents (VAN SICKLE; PAULSEN, 2008). For the MINASPACS model, the classes "poor" and "bad" (resulting in "bad") and "good" with "high" (resulting in "high") were joined, and kept the "moderate" class to obtain a 2 × 2 table for the RR calculation. The RR is a conditional probability representing the likelihood that low/bad O/E values are associated with high stressor scores and is calculated as follows (Equation (1)):

$$RR = \frac{\mathrm{Pr}\ (\mathrm{O/Ep}\mid \mathrm{Sh})}{\mathrm{Pr}\ (\mathrm{O/Ep}\mid \mathrm{Sl})} \qquad (1)$$

The numerator is the probability of finding poor biological conditions (O/E value >50% taxa loss) given high stressor scores (Sh), and the denominator is the probability of finding poor biological conditions given low stressor scores (Sl) (SILVA et al., 2018; VAN SICKLE; PAULSEN, 2008). RR scores equal to 1 denote the absence of association between the biological indicator (O/E value) and the stressor (VAN SICKLE; PAULSEN, 2008). For a RR

> 1, we interpret the value as how many times more likely a poor O/E value would occur given high-stressor conditions relative to low-stressor conditions. The 95% confidence intervals for RR estimations using the conditional probability method (ALTMAN, 1991) was calculated, and RR was significant when the lower 95% confidence interval was > 1.

Table 3 - Thresholds of condition classes for human stressor indicators.

Pressure variable	Thresholds		Source
	Good	*Poor*	
% Agriculture	< 60	≥ 60	SILVA et al. (2017)
% Pasture	< 60	≥ 60	SILVA et al. (2017)
% Urban infrastructure	0	> 0	LORENZ et al. (2004)
% Anthropogenic use**	< 25	≥ 25	LORENZ et al. (2004)
Turbidity (NTU)	≤ 100	> 100	CONAMA 357/2005, class II*
Total Nitrogen (mg/L)	≤ 0.2	> 0.2	CONAMA 357/2005, class II*
Total Phosphorus (mg/L)	≤ 0.1	> 0.1	CONAMA 57/2005, class II*

*Class II water quality corresponds mainly to water intended for human consumption after simplified treatment and protection of aquatic communities (BRASIL, 2005).
** All anthropogenic uses combined (pasture, agriculture, monoculture, mining, industrial area, and urban infrastructure).

3.4. Results

Ninety-seven taxa were identified from the dataset in 381 stream sites. The most abundant taxa were Chironomidae (41.99%), Simuliidae (10.36%), Elmidae (7.25%), Baetidae (6.87%), and Oligochaeta (6.47%). Eighty-seven stream sites met the selection criteria for minimally disturbed sites and were used for MINASPACS model construction. Reference and test sites occurred at similar elevations (510.59 – 1,455.69 m and 411.00 – 1,419.56 m a.s.l., respectively), temperature ranges (16.89 – 23.75 ºC and 17.40 – 24.06 ºC), and annual precipitation (1,019.04 – 1,675.43 mm and 972.99 – 1,670.89 mm). However, differences in total area, % land use classes, and water physical and chemical quality were found (Table 4). Overall, test sites had greater catchment areas, Total Phosphorus, Total Nitrogen, turbidity, and % anthropogenic uses than the reference sites (Table 4). Furthermore, more than 70% of our reference sites were in the São Francisco River basin, with small river catchment areas (< 100 km²). Only three sites had a catchment area > 1,000 km².

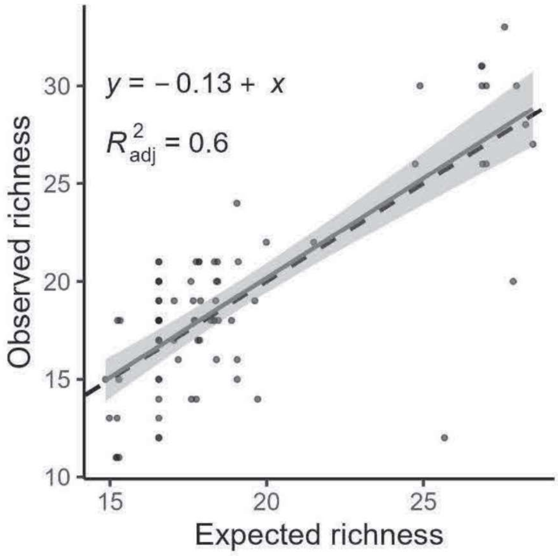
Table 4 - Mean and range of values for selected environmental variables at reference and test sites.

Variable	Reference sites (n = 87)			Test sites (n = 294)		
	Mean	Min	Max	Mean	Min	Max
Annual Mean Temperature (ºC)*	19.96	16.89	23.75	21.30	17.40	24.06
Annual Mean Precipitation (mm)*	1,424.68	1,019.04	1,675.43	1,440.90	972.99	1,670.89
Mean catchment slope (%)	15.5	2.8	48.5	10.2	2.2	46.1
Altitude (m)	943.11	510.59	1,455.69	737.84	411.00	1,419.56
Distance to source (m)	12,429.75	68.71	85,906.92	29,075.95	46.00	706,308.64
Catchment area (km²)	132.28	0.03	1,809.83	606.20	0.00	27,923.96
% Forest	29.01	0.00	100.00	16.09	0.00	75.01
% Savanna	21.71	0.00	83.36	9.50	0.00	79.68
% Pasture	3.97	0.00	21.66	30.03	0.00	86.44
% Agriculture	1.77	0.00	19.35	5.20	0.00	37.45
% Urban infrastructure	0.00	0.00	0.00	6.48	0.00	100.00
% Anthropogenic use	8.32	0.00	24.26	60.63	0.00	100.00
Total Phosphorus (mg/L)	0.01	0.00	0.10	0.28	0.00	11.66
Total Nitrogen (mg/L)	0.08	0.03	0.20	0.25	0.00	16.30
Turbidity (NTU)	6.35	0.10	61.00	15.79	0.30	433.00
Taxa richness	29	16	41	19	0	43
O/E score	1.01	0.47	1.27	0.62	0.00	1.26

The MINASPACS model was built with 87 minimally disturbed stream sites, and 10% of those were used for validation. We defined four reference faunal groups from the cluster analysis of the 78 calibration sites. All reference groups contained at least 14 reference sites. The most representative families in each group were: Psephenidae and Pleidae for group 1, Hydrobiosidae for group 2, Megapodagrionidae and Lutrochidae for group 3. Group 4 had no exclusive representative taxon (Supplementary Information 1, Table S2).

Four variables (mean catchment slope, lithological synthesis, annual mean temperature, and annual mean precipitation) were selected for the final model. Wilks' λ was low (0.148) and F-stat was high (16.561), indicating the model's high discriminatory ability. This is supported by the high accuracy evidenced by the MNO/E = 1.001, SDO/E = 0.16, and the O/E regression was within acceptable values ($R^2 = 0.608$; slope = 1.016; intersection = -0.134). The validation sites had similar O/E values (MNO/E = 1.03; SDO/E = 0.13), which indicated a good evaluation for new sites (Figure 2).

Figure 2 - O/E regression for the MINASPACS model built with biological and environmental data from 87 reference sites.



O/E values ranged from poor/bad conditions (31 % of test sites) to good/high (52 % of test sites), with 17 % in moderate condition (Figure 3). Spearman rank correlations between test sites' O/E values and environmental variables were also examined to assess if a decreasing biological condition was associated with declines in water quality or stressors. Four stressors had significant and negative rank correlations with O/E values: % urban infrastructure (r^2 = -0.43; $p < 0.05$), % anthropogenic use (r^2 = -0.22; $p < 0.05$), Total Nitrogen (r^2 = -0.20; $p < 0.05$) and turbidity (r^2 = -0.35; $p < 0.05$). O/E values were also negatively correlated with Total Phosphorus (r^2 = -0.09; $p < 0.05$) and % pasture (r^2 = -0.04; $p < 0.05$). The correlations between stressors and O/E values were generally weak, except for % urban infrastructure (Supplementary Information 1, Table S3).

The MINASPACS could detect the influence of all seven stressors in the biological condition of rivers based on macroinvertebrate assemblages. RR estimations varied between the Alto Paraná River and the São Francisco River basins (Figure 4). In the São Francisco, only % agriculture was below 1, similar to the regional assessment. Therefore, the other six stressors constitute a risk to biological condition (relative risk > 1). In the Alto Paraná, Total Nitrogen and % of urban infrastructure were the only stressors associated with RR significant for poor/bad O/E values (when the lower 95% confidence interval was > 1). Stressors showing no relative risk resulted from the low association between the stressor levels exceeding the established thresholds and the biological condition.

Figure 3 - O/E classification (5 classes) for all sites.

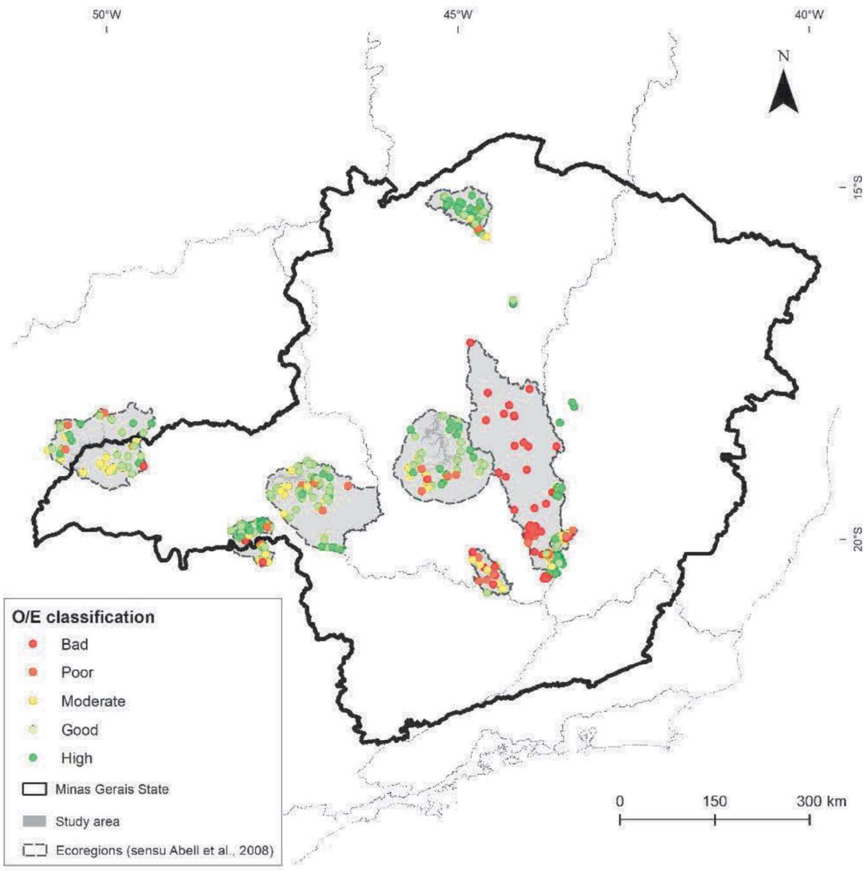
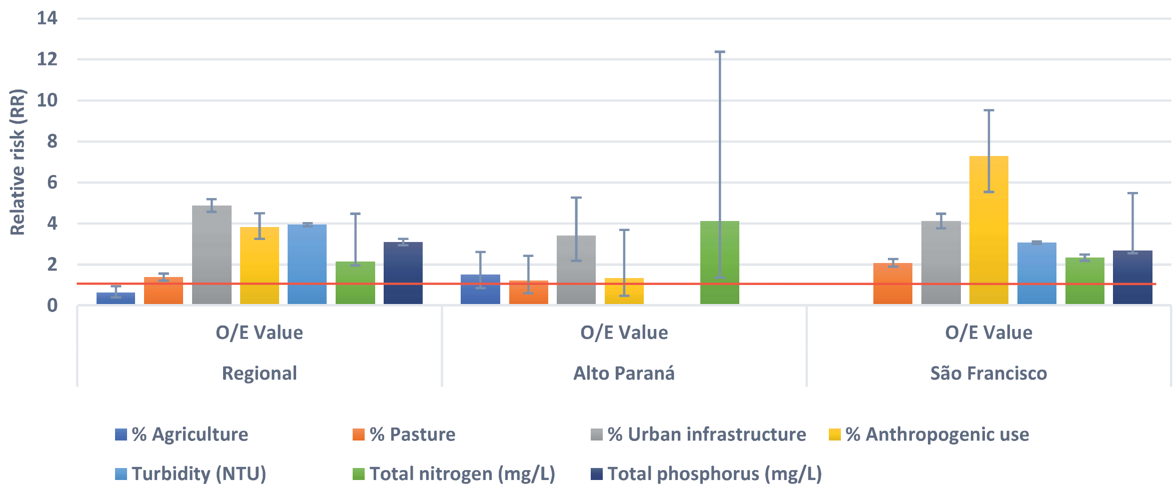


Figure 4 - Estimated relative risk (RR) and 95% confidence intervals for poor/bad O/E values given seven stressors. RR confidence interval below 1 (red line) indicates insignificant association.



3.5. Discussion

Our predictive modeling results corroborate with previous work from Europe (e.g., DAVY-BOWKER et al., 2006; MENDES et al., 2014), North America (e.g., HARGETT et al., 2007; HAWKINS et al., 2000), Asia (CHEN et al., 2019), and South America (e.g., JOVEM-AZEVÊDO et al., 2020; MOYA et al., 2011) which show that predictive models can provide a powerful tool to assess the biological condition of aquatic ecosystems. The relative risk approach confirmed the sensitivity of the MINASPACS to the stressors affecting Minas Gerais streams and rivers.

3.5.1. MINASPACS model construction and validation

Our results present strong evidence that the RIVPACS-type model can be developed based mainly on map-level predictor variables, as noted by Hargett et al. (2007). Moreover, all selected variables are easily obtained through geospatial tools, which supports model development being a useful approach for managers in terms of cost and time for data collection (HARGETT et al., 2007).

The selection of potential predictive variables used in the MINASPACS model took into account not only the statistical measures, but also the experience of model development in other countries such as Great Britain (e.g., WRIGHT, 1995), Australia (SMITH et al., 1999), the USA (HAWKINS et al., 2000), and Portugal (FEIO et al., 2009, 2012, 2007). Variables such as latitude or elevation imply that temperature is a primary factor determining the composition of stream macroinvertebrate fauna. At the same time, alkalinity suggests that either the ionic composition of the water or the geologic origin from which bed materials are derived are also important determinants of biotic structure (HAWKINS et al., 2000). Therefore, the four discriminant variables finally elected for our model (mean catchment slope, lithological synthesis, annual mean temperature, and annual precipitation) are aligned with the results of other studies (e.g., FEIO et al., 2007; HARGETT et al., 2007; MENDES et al., 2014).

MINASPACS was built using four faunal reference groups, which were well discriminated by the environmental variables, covering 40,106 km^2. More groups would result in fewer sites per group, reducing model performance in the test site assessments. For instance, Feio et al. (2007) developed and validated a multivariate model for the Mondego catchment

(6,670 km²) using two reference groups, whereas the RIVPACS study covering the UK (approx. 240,000 km²) used 35 groups. However, the grouping process is one of the most subjective components of the modeling and should be reviewed in the future if new high quality reference sites are added (FEIO et al., 2007). More reference groups could have been defined with more reference sites, although minimally disturbed areas are generally scarce in Minas Gerais, much of Europe (BORGWARDT et al., 2019; LORENZ; FELD; HERING et al., 2004; OLIVEIRA et al., 2016) and much of the USA (HERLIHY et al., 2020; WHITTIER et al., 2007).

Family-level identification of macroinvertebrates was efficient for our model construction, as noted by Sudaryanti et al. (2001) in a previous study. This is an appropriate taxonomic resolution in many tropical regions with high diversity but limited taxonomic knowledge (GODOY et al., 2019). However, for regions with many genera and species per family, important information on species-specific taxon-habitat relationships could easily be lost by adopting family-level taxonomic resolution because of the differing ecological requirements of different species and genera within a family (HAWKINS et al., 2000). In the case of Brazil, predictive models were successfully developed using family-level taxonomic resolution (e.g., JOVEM-AZEVÊDO et al., 2020; MORENO et al., 2009). Nevertheless, Molozzi et al. (2012) highlighted the importance of using genus level Chironomidae (Diptera) in reservoir assessment because different genera have different sensitivities to organic and metal contaminants. Conversely, a higher taxonomic resolution requires taxonomic expertise and is more time-consuming (FEIO et al., 2006; VADAS et al., 2022), which is a critical aspect in Brazil (BUSS et al., 2015).

Regarding the environmental variables selected in the MINASPACS, other environmental factors obtained at the local scale may enhance the accuracy and precision of the model. Relevant variables may include annual runoff, alkalinity, width, depth and flow regime (DAVY-BOWKER et al., 2006). Martins et al. (2018b) showed that taxonomic richness and composition of macroinvertebrate assemblages in Minas Gerais are positively affected by the presence of leaf packs on the streambed. These leaves accumulate on the streambed, forming important habitats for aquatic macroinvertebrates, where they find food and shelter against predators (LIGEIRO et al., 2020). In neotropical ecosystems, Macedo et al. (2014) showed that variables related to stream size (wetted width, bank full width, and wetted area) are positively correlated with macroinvertebrate richness. Castro et al. (2020) demonstrated how biodiversity changes from local to regional spatial extents. Considering regional and local variables could potentially enhance the accuracy and precision of the models.

Minas Gerais is extraordinarily heterogeneous regarding its physical environment and invertebrate biota. Our sampling design is limited and possibly the number of stream sites does not reflect the diversity in the entire area. Thus, further work is needed to identify other local factors that may enhance the accuracy and precision of the MINASPACS model. Quantitative local habitat information obtained during the field survey stage is necessary to standardize the sampling protocol and train the research team to reduce interpersonal variability (HUGHES et al., 2008; JUSIK et al., 2015).

3.5.2. Relative risk assessment

The MINASPACS detected the influence of all seven stressors considered in this study and urban infrastructure posed the most significant risk to biological conditions. Most of our sites are in the das Velhas River basin, which is highly impaired in the Belo Horizonte Metropolitan Region, the capital of Minas Gerais with a population of 2.7 million people. In the past two decades, hydrologic modifications, channelization, sedimentation, nutrient loadings, heavy metals contamination and microplastics are potentially affecting macroinvertebrate assemblages (FEIO et al., 2015). The fact that impairment was prevalent for rivers and streams in lowlands was expected since the population of Minas Gerais mostly occurs in lowland river sections. Areas with high densities of test sites with bad/poor O/E values were confined to the plains and near large cities.

Although the MINASPACS detected the influence of seven stressors in Minas Gerais, other pollutants were not analyzed (e.g., heavy metals) which can limit the presence and development of sensitive organisms (MELLO et al., 2023). Furthermore, no strong correlation between land use and water quality parameters were detected, as Silva et al. (2018) found. In some cases, Total Phosphorus and Total Nitrogen were omitted because they did not exceed thresholds established by the water resources legislation (BRASIL, 2005). Using different O/E classes to represent good and poor biological conditions, as well as different land use classes and nutrient criteria to represent reference conditions would likely affect the biological assessments (FEIO et al., 2014; HERLIHY et al., 2020). Decreased macroinvertebrates richness in small streams could occur with much lower nutrient concentrations than the values set in legislation or regulations (FIRMIANO et al., 2017; HERLIHY et al., 2020).

3.5.3. MINASPACS model as a wide bioassessment tool in Minas Gerais

Grading sites into classes of ‘ecological status’ for surface waters is now a requirement of some Brazilian states (i.e., Minas Gerais, MINAS GERAIS, 2022) and MINASPACS O/E ratio may be a useful metric for assessing the condition of macroinvertebrate fauna as a constituent of the ecological status of rivers. The water resources law (MINAS GERAIS, 2022) also established river typology as one of the stages of biomonitoring in Minas Gerais. Thus, the MINASPACS model can reduce implementation time and cost demands, favoring the environmental efforts in Minas Gerais. To be more useful, the MINASPACS should be robust and sensitive to natural environmental variation. Characterizing different reference conditions for a limited number of river types or within sufficiently homogeneous areas could be a good starting point for a successful type-specific approach. A standard approach aims to facilitate direct comparison of the biological condition of streams and rivers at local, regional, and national scales, thereby yielding improved scientific generalizations, assessments, and regulation (STODDARD et al., 2008). In this way, increasing the number of reference sites, enlarging the sampling area to adjacent basins and states, and multiple-year sampling of a small set of reference sites could improve the model and constitute a powerful tool for a nation-wide bioassessment scheme.

3.6. Conclusion

The MINASPACS predictive model based on macroinvertebrates were developed and validated, which can fulfill all scientific aspects required for classification systems under the present Minas Gerais legislation on water resources. Our model responded to seven human pressures impairing stream and river ecosystems in Southeastern Brazil, providing a clear, simple, and defensible measure of the biological condition of streams in a diverse landscape. The MINASPACS can be effectively used with a relative risk approach to develop biological assessment methods for Brazilian surface waters and, later, to other South American countries.

3.7. Acknowledgements

We thank the field crews for their great efforts in generating the large amount of data used in the present study, mainly colleagues from the Laboratório de Ecologia de Bentos UFMG, Laboratorio de Ecologia de Peixes UFLA e Centro de Inovação e Tecnologia SENAI FIEMG.

3.8. References

AGRA, J. U. M. et al. Ecoregions and stream types help us understand ecological variability in Neotropical reference streams. **Marine and Freshwater Research**, v. 70, n. 4, p. 594–602, 2019.

AGUIAR, F. C.; FEIO, M. J.; FERREIRA, M. T. Choosing the best method for stream bioassessment using macrophyte communities: Indices and predictive models. **Ecological Indicators**, v. 11, n. 2, p. 379–388, 2011.

ALTMAN, D. G. Practical Statistics for Medical Research. **Australian and New Zealand Journal of Surgery**, v. 61, n. 12, p. 963–964, 1991.

BLAKELY, T. J.; HARDING, J. S. The SingScore: a macroinvertebrate biotic index for assessing the health of Sinms and canals. **Nature**, v. 269, n. 5629, p. 55, 2010.

BORGWARDT, F. et al. Ex uno plures – Defining different types of very large rivers in Europe to foster solid aquatic bio-assessment. **Ecological Indicators**, v. 107, p. 105599, 2019.

BUSS, D. F. et al. Stream biomonitoring using macroinvertebrates around the globe: a comparison of large-scale programs. **Environmental Monitoring and Assessment**, v. 187, n. 1, p. 4132, 9 jan. 2015.

CALLISTO, M. et al. Multi-status and multi-spatial scale assessment of landscape effects on benthic macroinvertebrates in the Neotropical Savanna. Advances in Understanding **Landscape Influences on Freshwater Habitats and Biological Assemblages**, p. 275–302, 2019.

CALLISTO, M. et al. Beta diversity of aquatic macroinvertebrate assemblages associated with leaf patches in neotropical montane streams. **Ecology and Evolution**, v. 11, n. 6, p. 2551–2560, 7 mar. 2021.

CASTRO, D. M. P. et al. Beta diversity of aquatic invertebrates increases along an altitudinal gradient in a Neotropical mountain. **Biotropica**, v. 51, n. 3, p. 399–411, maio 2019.

CASTRO, D. M. P. et al. Unveiling patterns of taxonomic and functional diversities of stream insects across four spatial scales in the neotropical savanna. **Ecological Indicators**, v. 118, n. July, p. 106769, 2020.

CHEN, K. et al. Science of the Total Environment Incorporating functional traits to enhance multimetric index performance and assess land use gradients. **Science of the Total Environment**, v. 691, p. 1005–1015, 2019.

CLARKE, R. T.; WRIGHT, J. F.; FURSE, M. T. RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers. **Ecological Modelling**, v. 160, n. 3, p. 219–233, fev. 2003.

DAVY-BOWKER, J. et al. A comparison of the European Water Framework Directive physical typology and RIVPACS-type models as alternative methods of establishing reference conditions for benthic macroinvertebrates. **Hydrobiologia**, v. 566, n. 1, p. 91–105, 2006.

DUDGEON, D. Prospects for sustaining freshwater biodiversity in the 21st century: Linking ecosystem structure and function. **Current Opinion in Environmental Sustainability**, v. 2, n. 5–6, p. 422–430, 2010.

FEIO, M. J. et al. A predictive model for freshwater bioassessment (Mondego River, Portugal). **Hydrobiologia**, v. 589, n. 1, p. 55–68, 2007.

FEIO, M. J. et al. Water quality assessment of Portuguese streams: Regional or national predictive models? **Ecological Indicators**, v. 9, n. 4, p. 791–806, 2009.

FEIO, M. J. et al. AQUAFLORA: A predictive model based on diatoms and macrophytes for streams water quality assessment. **Ecological Indicators**, v. 18, p. 586–598, 2012.

FEIO, M. J. et al. Defining and testing targets for the recovery of tropical streams based on macroinvertebrate communities and abiotic conditions. **River Research and Applications**, v. 30, n. January, p. 132–133, 2014a.

FEIO, M. J. et al. Least Disturbed Condition for European Mediterranean rivers. **Science of the Total Environment**, v. 476–477, p. 745–756, 2014b.

FEIO, M. J. et al. Defining and Testing Targets for the Recovery of Tropical Streams Based on Macroinvertebrate Communities and Abiotic Conditions. **River Research and Applications**, v. 31, n. 1, p. 70–84, jan. 2015.

FEIO, M. J.; POQUET, J. M. Predictive Models for Freshwater Biological Assessment: Statistical Approaches, Biological Elements and the Iberian Peninsula Experience: A Review. **International Review of Hydrobiology**, v. 96, n. 4, p. 321–346, 2011.

FEIO, M. J.; REYNOLDSON, T. B.; GRAÇA, M. A. Effect of seasonal changes on predictive model assessments of streams water quality with macroinvertebrates. **International Review of Hydrobiology**, v. 91, n. 6, p. 509–520, 2006.

FEIO, M. J.; VIANA-FERREIRA, C.; COSTA, C. Testing a multiple machine learning tool (HYDRA) for the bioassessment of fresh waters. **Freshwater Science**, v. 33, n. 4, p. 1286–1296, 2014.

FERREIRA, H. L. M. et al. **Ambientes Aquáticos em Minas Gerais**. 1a edição ed. Belo Horizonte: Prodemge, 2017.

FIRMIANO, K. R. et al. Mayfly bioindicator thresholds for several anthropogenic disturbances in neotropical savanna streams. **Ecological Indicators**, v. 74, p. 276–284, 1 mar. 2017.

GARUANA, L. et al. Integração de indicadores ecológicos, ambientais e de saúde humana em microbacias urbanas Luziana. **Revista Espinhaço**, v. 9, n. 1, p. 1–16, 2020.

GODOY, B. S. et al. Taxonomic sufficiency and effects of environmental and spatial drivers on aquatic insect community. **Ecological Indicators**, v. 107, p. 105624, dez. 2019.

HARGETT, E. G. et al. Development of a RIVPACS-type predictive model for bioassessment of wadeable streams in Wyoming. **Ecological Indicators**, v. 7, n. 4, p. 807–826, 2007.

HAWKINS, C. P. et al. Development and evaluation of predictive models for measuring the biological integrity of streams. **Ecological Applications**, v. 10, n. 5, p. 1456–1477, 2000.

HAWKINS, C. P.; CAO, Y.; ROPER, B. Method of predicting reference condition biota affects the performance and interpretation of ecological indices. **Freshwater Biology**, v. 55, n. 5, p. 1066–1085, 2010.

HERLIHY, A. T. et al. The relation of lotic fish and benthic macroinvertebrate condition indices to environmental factors across the conterminous USA. **Ecological Indicators**, v. 112, n. July 2019, p. 105958, 2020.

HUGHES, R. M. et al. Acquiring data for large aquatic resource surveys : the art of compromise among science , logistics , and reality. **Journal of the North American Benthological Society**, v. 27, n. 4, p. 837–859, 2008.

HUGHES, R. M.; LARSEN, D. P.; OMERNIK, J. M. Regional reference sites: a method for assessing stream potentials. **Environmental Management**, v. 10, n. 5, p. 629–635, 1986.

JOHNSON, R.; SANDIN, L. **Development of a prediction and classification system for lake (littoral, SWEPAC) and stream (riffle SWEPAC) macroinvertebrate communities**. Department of Environmental Assessment, Uppsala, Sweden: [s.n.]. Disponível em: <http://info1.ma.slu.se/IMA/Publikationer/internserie/2001-23.pdf>.

JOVEM-AZEVÊDO, D. et al. Rehabilitation scenarios for reservoirs: Predicting their effect on invertebrate communities through machine learning. **River Research and Applications**, n. April, p. 1–15, 2020.

JUSIK, S. et al. Development of comprehensive river typology based on macrophytes in the mountain-lowland gradient of different Central European ecoregions. **Hydrobiologia**, v. 745, n. 1, p. 241–262, 18 fev. 2015.

KARR, J. R. Defining and measuring river health. **Freshwater Biology**, v. 41, n. 2, p. 221–234, mar. 1999.

KOKEŠ, J. et al. The PERLA system in the Czech Republic: A multivariate approach for assessing the ecological status of running waters. **Hydrobiologia**, v. 566, n. 1, p. 343–354, 2006.

LIGEIRO, R. et al. Choice of field and laboratory methods affects the detection of anthropogenic disturbances using stream macroinvertebrate assemblages. **Ecological Indicators**, v. 115, n. March, p. 106382, 2020.

LINARES, M. S. et al. Chronic urbanization decreases macroinvertebrate resilience to natural disturbances in neotropical streams. **Current Research in Environmental Sustainability**, v. 3, p. 100095, 2021.

LINKE, S. et al. ANNA: A new prediction method for bioassessment programs. **Freshwater Biology**, v. 50, n. 1, p. 147–158, 2005.

LORENZ, A.; FELD, C. K.; HERING, D. Typology of streams in Germany based on benthic invertebrates: Ecoregions, zonation, geology and substrate. **Limnologica**, v. 34, n. 4, p. 379–389, 2004.

MACEDO, D. R. et al. The relative influence of catchment and site variables on fish and macroinvertebrate richness in cerrado biome streams. **Landscape Ecology**, v. 29, n. 6, p. 1001–1016, 2014.

MACEDO, D. R. et al. Urban stream rehabilitation in a densely populated Brazilian metropolis. Frontiers in Environmental Science, v. 10, n. September, p. 921934, set. 2022.

MARTINS, I. et al. Regionalisation is key to establishing reference conditions for neotropical savanna streams. **Marine and Freshwater Research**, v. 69, n. 1, p. 82–94, 2018.

MARTINS, I. et al. Are multiple multimetric indices effective for assessing ecological condition in tropical basins? **Ecological Indicators**, v. 110, n. May, p. 105953, 2020.

MELLO, K. DE et al. Biomonitoring for Watershed Protection from a Multiscale Land-Use Perspective. **diversity**, v. 15, n. 636, p. 1–20, 2023.

MENDES, T. et al. Comparing alternatives for combining invertebrate and diatom assessment in stream quality classification. **Marine and Freshwater Research**, v. 65, n. 7, p. 612–623, 2014.

MERRITT, R.W., CUMMINS K.W. **An introduction to the aquatic insects of North America**. Kendall/Hunt Publishing Company, Dubuque, Iowa, USA, 1996.

MINAS GERAIS. Conselho Estadual de Política Ambiental – COPAM. **Resolução Normativa Conjunta COPAM/CERH-MG nº 08 de 21 de novembro de 2022**. Dispõe sobre a classificação dos corpos de água e diretrizes ambientais para o seu enquadramento, bem como estabelece as condições e padrões de lançamento de efluentes. Diário do Executivo. Minas Gerais. 2022.

MOLOZZI, J. et al. Development and test of a statistical model for the ecological assessment of tropical reservoirs based on benthic macroinvertebrates. **Ecological Indicators**, v. 23, p. 155–165, 2012.

MORENO, P. et al. Use of the BEAST model for biomonitoring water quality in a neotropical basin. **Hydrobiologia**, 2009.

MOYA, N. et al. Macroinvertebrate-based multimetric predictive models for evaluating the human impact on biotic condition of Bolivian streams. **Ecological Indicators**, v. 11, p. 840–847, 2011.

MUGNAI, R., NESSIMIAN, J.L., BAPTISTA, D.F. **Guide for the Identification of Aquatic Macroinvertebrates of Rio de Janeiro State**. Technical Books Editora, Rio de Janeiro, 2009.

MUGNAI, R., NESSIMIAN, J.L., BAPTISTA, D.F. **Manual de Identificação de Macroinvertebrados Aquáticos do Estado do Rio de Janeiro**. Technical Books Editora, Rio de Janeiro, 2010.

OLIVEIRA, U. et al. The strong influence of collection bias on biodiversity knowledge shortfalls of Brazilian terrestrial biodiversity. **Diversity and Distributions**, v. 22, n. 12, p. 1232–1244, 30 dez. 2016.

PARDO, I. et al. An invertebrate predictive model (NORTI) for streams and rivers: Sensitivity of the model in detecting stress gradients. **Ecological Indicators**, v. 45, p. 51–62, 2014.

PAULSEN, S. G. et al. Rivers and Streams: Upgrading Monitoring of the Nation's Freshwater Resources - Meeting the Spirit of the Clean Water Act. **IntechOpen**, n. tourism, p. 13, 2016.

PÉREZ, G.R. **Guía para el estudio de los macroinvertebrados acuáticos del Departamento de Antioquia**. Universidad de Antioquia Facultad de Ciencias Exactas y Naturales Centro de Investigaciones, CIEN. Medelín, 1988.

PÉS, A.M.O.; HAMADA, N., NESSIMIAN, J.L. Identification keys of larvae for families and genera of Trichoptera (Insecta) of Central Amazonia, Brazil. **Revista Brasileira de Entomologia** 49, 181–204, 2005.

REID, A. J. et al. Emerging threats and persistent conservation challenges for freshwater biodiversity. **Biological Reviews**, v. 94, n. 3, p. 849–873, 2019.

REYNOLDSON, T. B. et al. The reference condition: A comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. **Journal of the North American Benthological Society**, v. 16, n. 4, p. 833–852, 1997.

SANO, E. E. et al. Cerrado ecoregions: A spatial framework to assess and prioritize Brazilian savanna environmental diversity for conservation. **Journal of Environmental Management**, v. 232, n. July, p. 818–828, 2019.

SILVA, D. R. O. et al. An improved macroinvertebrate multimetric index for the assessment of wadeable streams in the neotropical savanna. **Ecological Indicators**, v. 81, n. June, p. 514–525, 2017.

SILVA, D. R. O. et al. Assessing the extent and relative risk of aquatic stressors on stream macroinvertebrate assemblages in the neotropical savanna. **Science of the Total Environment**, v. 633, p. 179–188, 2018.

SMITH, M. J. et al. AusRivAS: Using macroinvertebrates to assess ecological condition of rivers in Western Australia. **Freshwater Biology**, v. 41, n. 2, p. 269–282, 1999.

SOUZA, C. M. et al. Reconstructing three decades of land use and land cover changes in Brazilian biomes with Landsat archive and Earth engine. **Remote Sensing**, v. 12, n. 17, p. 2735, ago. 2020.

STEVENS, D. L.; OLSEN, A. R. Spatially balanced sampling of natural resources. **Journal of the American Statistical Association**, v. 99, n. 465, p. 262–278, mar. 2004.

STODDARD, J. L. et al. Setting expectations for the ecological condition of streams: The concept of reference condition. **Ecological Applications**, v. 16, n. 4, p. 1267–1276, 2006.

STODDARD, J. L. et al. A process for creating multimetric indices for large-scale aquatic surveys. **Journal of the North American Benthological Society**, v. 27, n. 4, p. 878–891, 2008.

SUDARYANTI, S. et al. Assessment of the biological health of the Brantas River, East Java, Indonesia using the Australian River Assessment System (AUSRIVAS) methodology. **Aquatic Ecology**, v. 35, n. 2, p. 135–146, 2001.

USEPA. **National Rivers and Streams Asessment, 2008-2009**. Washington, DC: [s.n.].

VADAS, R. L. et al. Assemblage-based biomonitoring of freshwater ecosystem health via multimetric indices: A critical review and suggestions for improving their applicability. Water Biology and Security, v. 1, n. 3, 2022.

VAN SICKLE, J.; HUFF, D. D.; HAWKINS, C. P. Selecting discriminant function models for predicting the expected richness of aquatic macroinvertebrates. **Freshwater Biology**, v. 51, n. 2, p. 359–372, 2006.

VAN SICKLE, J. et al. A null model for the expected macroinvertebrate assemblage in streams. **Journal of the North American Benthological Society**, v. 24, n. 1, p. 178–191, 2005.

VAN SICKLE, J.; PAULSEN, S. G. Assessing the attributable risks, relative risks, and regional extents of aquatic stressors. **Journal of the North American Benthological Society**, v. 27, n. 4, p. 920–931, 2008.

WIGGINS, G.B. Larvae of the North American Caddisfly Genera (Trichoptera), 2nd edn. **University of Toronto Press**: Toronto, Canada, 1996.

WHITTIER, T. R. et al. Selecting reference sites for stream biological assessments: best professional judgment or objective criteria. **Journal of the North American Benthological Society**, v. 26, n. 2, p. 349–360, jun. 2007.

WRIGHT, J. F. Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. **Australian Journal of Ecology**, v. 20, n. 1, p. 181–197, 1995.

3.9. Appendix A. Supplementary data

3.9.1. Supplementary information 1

Table S1 - Lithological synthesis of the Minas Gerais state according to Ferreira et al. (2017).

Group	Description
Siliceous rocks (S)	The siliceous group includes rocks whose chemical composition has silica (SiO2) as its main component, such as acidic and intermediate igneous rocks, with more than 52% silica. They include sandy detrital sedimentary rocks, such as quartz-arenites and subarchoses, as well as rich conglomerates and fragments of quartz-arenites and acidic and intermediate igneous rocks. Metamorphic equivalent rocks are also part of it.
Pelitic rocks (P)	Detrital sedimentary rocks formed by fragments in the mud fraction, such as pelites and their metamorphic equivalents.
Metamorphic rocks (F)	Consisting of rocks of igneous and sedimentary origin. Silica content below 52%, which exhibit intermediate to high-grade metamorphism. In this group are Archean and Paleoproterozoic rocks of similar composition.
Carbonate rocks (C)	Sedimentary rocks with a chemical composition rich in calcium, such as limestone and dolomites, belong to the carbonate group.
Volcanic rocks (B)	They consist of basic rocks, mainly extrusive ones, formed by spills. It includes intrusive outcrop rocks of basic composition and their low-grade metamorphic equivalents.
Alkaline rocks (A)	Rocks rich in alkalis, with minerals such as feldspathoids and sodium amphiboles. They include alkaline syenites, phonolites, and dunites. They usually form rocky bodies of small regional expression whose distribution in Minas Gerais territory is restricted to a few occurrences such as Poços de Caldas-MG.
Laterized sediments (L)	Sediments of alluvial, colluvial, and eluvial origin, usually cemented by oxides and hydroxides of iron and aluminum, with occurence in extensive plateaus and some plains.
Unconsolidated sediments (I)	Incohesive sandy and muddy sediments occur along the alluvial plains and terraces.

Table S2 - Taxa contributing up to 90% of cumulative percentage to Bray-Curtis similarity for the faunal reference groups, in descending order of contribution. Group 1 – Average similarity: 68.13 % (continues).

Taxa	Av. Abund	Av. Sim	Sim/SD	Contrib %	Cum %
Chironomidae	0.74	6.37	5.41	9.36	9.36
Elmidae	0.55	4.74	8.59	6.95	16.31
Leptohyphidae	0.49	4.17	7.27	6.13	22.44
Leptophlebiidae	0.51	3.93	4.23	5.76	28.20
Baetidae	0.45	3.55	4.33	5.21	33.41
Simuliidae	0.47	3.24	2.38	4.75	38.16
Hydropsychidae	0.38	3.20	6.22	4.69	42.85
Perlidae	0.34	2.71	3.85	3.98	46.83
Coenagrionidae	0.29	2.36	5.33	3.47	50.30
Leptoceridae	0.30	2.32	4.34	3.41	53.71
Libellulidae	0.29	2.31	4.35	3.39	57.10
Hydroptilidae	0.30	2.16	2.06	3.16	60.27
Empididae	0.24	1.76	1.78	2.58	62.85
Tipulidae	0.27	1.67	1.53	2.45	65.30
Calamoceratidae	0.26	1.52	1.25	2.23	67.54
Odontoceridae	0.23	1.50	1.72	2.19	69.73
Polycentropodidae	0.22	1.48	1.38	2.18	71.91
Psephenidae	0.23	1.28	1.09	1.87	73.78
Ceratopogonidae	0.23	1.26	0.84	1.85	75.63
Gomphidae	0.21	1.23	1.34	1.80	77.43
Oligochaeta	0.22	1.22	0.97	1.79	79.21
Caenidae	0.21	1.11	0.95	1.63	80.84
Pleidae	0.19	1.10	1.11	1.61	82.45
Naucoridae	0.21	1.04	0.85	1.53	83.98
Dytiscidae	0.18	1.03	0.99	1.51	85.49
Corydalidae	0.16	0.91	0.98	1.34	86.83
Calopterygidae	0.17	0.86	0.85	1.26	88.09
Glossosomatidae	0.16	0.77	0.82	1.14	89.22
Philopotamidae	0.15	0.77	0.84	1.14	90.36

Av. Abund: average abundance, Av. Sim: average similarity, Sim/SD: similarity/standard deviation, Contrib %: percentage contribution, Cum %: cumulative percentage.

Table S2 - Taxa contributing up to 90% of cumulative percentage to Bray-Curtis similarity for the faunal reference groups, in descending order of contribution. Group 2 – Average similarity: 56.83 % (continues).

Taxa	Av. Abund	Av. Sim	Sim/SD	Contrib %	Cum %
Chironomidae	0.78	8.66	4.47	15.24	15.24
Elmidae	0.48	4.83	2.90	8.49	23.73
Ceratopogonidae	0.39	3.88	2.83	6.83	30.56
Leptohyphidae	0.36	3.24	2.01	5.71	36.27
Baetidae	0.32	2.98	2.09	5.24	41.51
Leptoceridae	0.34	2.94	1.60	5.17	46.68
Leptophlebiidae	0.32	2.86	1.74	5.04	51.72
Gomphidae	0.30	2.58	1.67	4.55	56.27
Libellulidae	0.29	2.49	2.02	4.38	60.65
Naucoridae	0.26	2.39	2.13	4.21	64.86
Oligochaeta	0.30	2.35	1.22	4.14	68.99
Bivalvia	0.28	1.95	1.10	3.44	72.43
Hydrobiosidae	0.26	1.77	1.02	3.11	75.54
Coenagrionidae	0.21	1.51	1.02	2.66	78.20
Empididae	0.19	1.39	1.04	2.45	80.66
Helicopsychidae	0.22	1.02	0.62	1.80	82.45
Hydroptilidae	0.19	1.02	0.69	1.79	84.25
Pyralidae	0.16	0.92	0.77	1.62	85.87
Caenidae	0.13	0.70	0.62	1.24	87.10
Calopterygidae	0.13	0.68	0.64	1.19	88.29
Simuliidae	0.19	0.67	0.49	1.19	89.48
Odontoceridae	0.12	0.63	0.52	1.10	90.58

Av. Abund: average abundance, Av. Sim: average similarity, Sim/SD: similarity/standard deviation, Contrib %: percentage contribution, Cum %: cumulative percentage.

Table S2 - Taxa contributing up to 90% of cumulative percentage to Bray-Curtis similarity for the faunal reference groups, in descending order of contribution. Group 3 – Average similarity: 61.71 % (continues).

Taxa	Av. Abund	Av. Sim	Sim/SD	Contrib %	Cum %
Chironomidae	0.71	6.53	5.92	10.59	10.59
Leptophlebiidae	0.54	4.42	2.44	7.16	17.75
Elmidae	0.50	4.31	4.54	6.98	24.73
Perlidae	0.41	3.64	5.68	5.89	30.62
Baetidae	0.38	3.09	2.18	5.00	35.62
Gripopterygidae	0.36	2.90	2.29	4.70	40.32
Tipulidae	0.36	2.87	2.39	4.65	44.97
Ceratopogonidae	0.32	2.81	2.68	4.56	49.53
Oligochaeta	0.32	2.51	2.33	4.07	53.59
Simuliidae	0.33	2.22	1.53	3.60	57.19
Calamoceratidae	0.31	2.17	1.62	3.52	60.70
Coenagrionidae	0.29	2.12	1.67	3.43	64.14
Hydropsychidae	0.30	1.95	1.30	3.16	67.30
Polycentropodidae	0.26	1.85	1.38	2.99	70.29
Megapodagrionidae	0.25	1.72	1.38	2.79	73.08
Libellulidae	0.24	1.44	1.11	2.33	75.41
Leptohyphidae	0.29	1.33	0.72	2.15	77.56
Leptoceridae	0.21	1.08	0.85	1.74	79.31
Lutrochidae	0.19	0.90	0.73	1.45	80.76
Odontoceridae	0.19	0.87	0.72	1.42	82.18
Planariidae	0.17	0.81	0.72	1.31	83.48
Veliidae	0.15	0.80	0.74	1.30	84.78
Euthyplociidae	0.17	0.79	0.62	1.29	86.07
Bivalvia	0.19	0.76	0.62	1.23	87.31
Gomphidae	0.16	0.73	0.63	1.19	88.49
Corydalidae	0.16	0.70	0.62	1.14	89.63
Helicopsychidae	0.17	0.65	0.53	1.05	90.69

Av. Abund: average abundance, Av. Sim: average similarity , Sim/SD: similarity/standard deviation, Contrib %: percentage contribution, Cum %: cumulative percentage.

Table S2 - Taxa contributing up to 90% of cumulative percentage to Bray-Curtis similarity for the faunal reference groups, in descending order of contribution. Group 4 – Average similarity: 53.19 %.

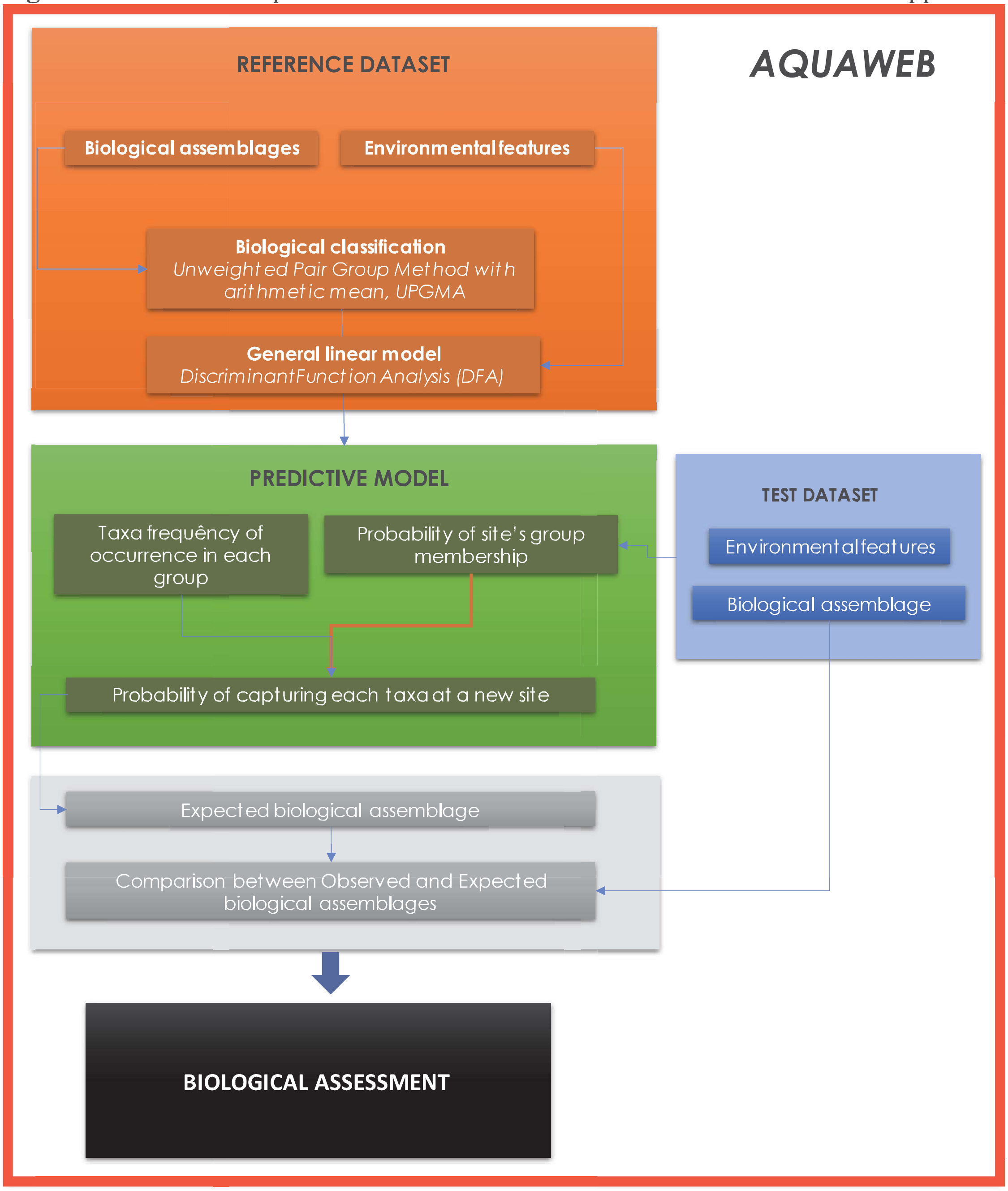
Taxa	Av. Abund	Av. Sim	Sim/SD	Contrib %	Cum %
Chironomidae	0.78	9.67	5.75	18.18	18.18
Simuliidae	0.57	6.23	3.06	11.72	29.90
Baetidae	0.44	4.73	3.12	8.89	38.79
Leptophlebiidae	0.40	4.40	5.22	8.27	47.07
Elmidae	0.42	4.37	1.99	8.22	55.29
Ceratopogonidae	0.34	3.06	1.36	5.75	61.03
Tipulidae	0.30	2.19	1.03	4.12	65.16
Hydropsychidae	0.24	1.90	1.07	3.57	68.73
Perlidae	0.26	1.72	0.87	3.23	71.96
Oligochaeta	0.24	1.71	0.89	3.21	75.17
Leptohyphidae	0.22	1.34	0.68	2.52	77.69
Hydroptilidae	0.19	1.12	0.59	2.10	79.79
Gripopterygidae	0.17	0.98	0.59	1.84	81.63
Veliidae	0.18	0.96	0.59	1.81	83.44
Pyralidae	0.15	0.93	0.59	1.75	85.19
Polycentropodidae	0.17	0.90	0.57	1.68	86.88
Libellulidae	0.15	0.88	0.59	1.65	88.53
Coenagrionidae	0.14	0.79	0.60	1.49	90.02

Av. Abund: average abundance, Av. Sim: average similarity, Sim/SD: similarity/standard deviation, Contrib %: percentage contribution, Cum %: cumulative percentage.

Table S3 - Spearman rank correlations between O/E value, environmental variables, and stressors. Marked correlations are significant at $p < 0.05$. N = 272 (sites with missing data for water quality parameters were excluded).

	O/E Value	Annual Mean Temperature (ºC)	Annual Precipitation (mm)	Annual Temperature Range (ºC)	Altitude (m)	Distance to source (m)	Mean catchment slope (%)	Catchment area (km²)	% Forest	% Savanna	% Pasture	% Agriculture	% Urban infrastructure	% Anthropogenic use	Total Phosphorus (mg/L)	Total Nitrogen (mg/L)	Turbidity (NTU)
O/E Value	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Annual Mean Temperature (ºC)	0.09	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Annual Mean Precipitation (mm)	-0.07	-0.55	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Annual Temperature Range (ºC)	-0.15	0.35	-0.62	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-
Altitude (m)	0.02	-0.94	0.48	-0.41	1.00	-	-	-	-	-	-	-	-	-	-	-	-
Distance to source (m)	-0.17	0.34	-0.42	0.30	-0.32	1.00	-	-	-	-	-	-	-	-	-	-	-
Mean catchment slope (%)	-0.29	-0.64	0.25	-0.16	0.48	-0.11	1.00	-	-	-	-	-	-	-	-	-	-
Catchment area (km²)	-0.13	0.31	-0.35	0.23	-0.27	0.94	-0.11	1.00	-	-	-	-	-	-	-	-	-
% Forest	-0.17	-0.35	0.21	-0.24	0.24	-0.06	0.43	-0.06	1.00	-	-	-	-	-	-	-	-
% Savanna	0.23	0.51	-0.80	0.23	-0.33	0.41	-0.32	0.37	-0.21	1.00	-	-	-	-	-	-	-
% Pasture	-0.04	0.11	0.23	0.12	-0.18	-0.18	-0.13	-0.17	0.09	-0.37	1.00	-	-	-	-	-	-
% Agriculture	0.01	0.02	0.39	-0.32	0.00	-0.16	-0.27	-0.14	-0.02	-0.26	0.00	1.00	-	-	-	-	-
% Urban infrastructure	-0.43	-0.04	0.09	0.06	0.02	-0.03	0.12	-0.04	-0.13	-0.12	-0.15	-0.09	1.00	-	-	-	-
% Anthropogenic use	-0.22	0.11	0.49	-0.08	-0.22	-0.30	-0.23	-0.28	-0.01	-0.62	0.59	0.46	0.24	1.00	-	-	-
Total Phosphorus (mg/L)	-0.09	0.18	-0.20	0.09	-0.13	0.33	-0.11	0.42	-0.08	0.22	-0.15	-0.08	0.09	-0.16	1.00	-	-
Total Nitrogen (mg/L)	-0.20	-0.06	0.03	0.07	0.03	0.00	0.10	-0.03	-0.07	-0.07	-0.09	-0.07	0.49	0.12	0.05	1.00	-
Turbidity (NTU)	-0.35	-0.05	0.03	0.07	0.05	-0.03	0.07	-0.03	-0.04	-0.05	-0.09	-0.07	0.61	0.14	0.11	0.23	1.00

Figure S1 - General steps followed in the MINASPACS for construction and application.



Source: adapted from Feio and Poquet (2011).

4. THESIS CONCLUSIONS

The river typology approach and the predictive modeling based on macroinvertebrate assemblages were helpful for establishing reference conditions for biological assessment and can offer an option for aquatic ecosystem management in Minas Gerais waters. Both tools were developed through a large database gathered over 16 years, using abiotic descriptors on a landscape scale obtained through geospatial tools, which could allow for further development a more inviting endeavor for managers in terms of cost and time.

Our results showed that the macroinvertebrate assemblages responded to different abiotic descriptors, land use conditions, and levels of physical and chemical water quality parameters, corroborating our hypothesis. The reference site selection criteria was proved adequate since environments with more significant anthropogenic alterations caused a simplification of the macroinvertebrate assemblages, and the MINASPACS predictive model represented this impairment.

Regarding the river typology, the conclusions are:

- Lowland and mountain river types have been validated, which reflected the natural variability of the benthic macroinvertebrate assemblages on a landscape scale.
- Family-level identification of macroinvertebrates was efficient for river typology construction.
- The most representative taxa of each river type and their traits were presented. Understanding their traits, the processes most impair each river type can be better understood.

Regarding the MINASPACS predictive model, the conclusions are:

- Family-level identification of macroinvertebrates was efficient for the model construction.
- The MINASPACS can fulfill all scientific aspects required for classification systems under the new Minas Gerais law on water resources (Normative Deliberation COPAM/CERH-MG 008/2022).
- The model detected the influence of all seven stressors considered in this study and urban infrastructure posed the most significant risk to biological conditions.

River typology and predictive models should be seen as complementary tools. The predictive models should be less susceptible to natural environmental variation, which the river typology can reduce. Increasing the number of reference sites and the sampling area for river basins not covered in this study could lead to type-specific ecological assessment. Therefore, for sufficiently homogeneous rivers or streams, only one predictive model, ecological index score range or biological metric can be used in ecological assessments, which would reduce the probability of inferring impairment when it does not exist or even not detect it when it exists.

4.1. Future perspectives

Topics for future research should aim at:

- To harmonize macroinvertebrate sampling methods and habitat protocols on a global scale to ensure comparable results in the future.
- Developing river typology and predictive models with other groups of organisms could allow for a more robust and reliable classification scheme.
- To construct predictive models based on both local scale and map-level environmental factors.
- To test other classifications, such as aquatic ecoregions instead of river typology.
- To create recovery scenarios and analyze the effect of rehabilitation measures on biological quality through predictive models based on machine learning.