

Universidade de São Paulo
Instituto de Astronomia, Geofísica e Ciências Atmosféricas
Departamento de Geofísica

Victor Otávio Salles Moraes Pôrto de Araujo

Otimização do *software* SeisComP para detecção e localização
automática de eventos sísmicos regionais no Brasil

São Paulo
2023

VICTOR OTÁVIO SALLES MORAES PÔRTO DE ARAUJO

Otimização do *software* SeisComP para detecção e localização automática de eventos sísmicos regionais no Brasil

Dissertação apresentada ao Departamento de Geofísica do Instituto de Astronomia, Geofísica e Ciências Atmosféricas da Universidade de São Paulo, como requisito parcial para a obtenção do título de Mestre em Ciências.

Área de concentração: Geofísica.

Orientador: Prof. Dr. Marcelo B. de Bianchi.

Versão corrigida.
O original encontra-se disponível na Unidade.

São Paulo
2023

Autorizo a reprodução e divulgação total ou parcial desse trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

NON DVCOR DVCO

Agradecimentos

Agradeço primeiramente ao meu orientador, Prof. Dr. Marcelo B. de Bianchi, por toda a ajuda, paciência, conhecimento e conselhos fornecidos ao longo de todo o trabalho. Esse projeto não teria sido possível sem a sua atuação nos âmbitos educacional e pessoal, e sou incapaz de exteriorizar a minha satisfação e eterno agradecimento por ter sido seu aluno. Espero, um dia, ter novamente a oportunidade de trabalhar com você.

Agradeço à minha família, pelo apoio e incentivo em todos os momentos e decisões tomadas ao longo da minha vida, bem como não deixar me faltar nada para que eu pudesse seguir minhas escolhas.

Agradeço à minha parceira, melhor amiga e namorada Marina Baseggio, que acompanhou de perto todos os altos e baixos da minha jornada no mestrado. Obrigado pelo apoio incondicional e por entender todas as emoções vividas durante o processo.

Agradeço imensamente aos amigos que, direta ou indiretamente, me ajudaram no desenvolvimento desse trabalho. Especialmente, agradeço ao geofísico mais interdisciplinar que já conheci, Felipe Jesus, por toda ajuda, discussões e ensinamentos. Também incluo nos agradecimentos especiais o meu amigo, veterano, colega de trabalho, gerente e geofísico Daniel Coelho, por todo o apoio, conversas e colaborações ao longo dos últimos anos. Minha sincera admiração por ambos.

Aos amigos e colegas de trabalho da Tetra Tech, em especial ao digníssimo Doutor Leonardo Santana, meu agradecimento pela confiança em mim e no meu potencial para somar à equipe de Geofísica Aplicada.

Ao Prof. Dr. Jesus Berrocal e a todos os membros da Terrafísica, sou eternamente grato pelos anos de aprendizado que se tornaram a base dos meus conhecimentos de sismologia.

Por fim, agradeço a mim mesmo pelas escolhas que me trouxeram até aqui, sejam elas acadêmicas, profissionais ou pessoais. Por não desistir das coisas pelas quais acreditei ser necessário lutar e por desistir do que era necessário desistir. Satisfeito por olhar para trás e ver o caminho percorrido, otimista por olhar para frente e imaginar o que está por vir. Sigo.

Resumo

SALLES, V. **Otimização do *software* SeisComP para detecção e localização automática de eventos sísmicos regionais no Brasil.** 2023. Dissertação de Mestrado – Instituto de Astronomia, Geofísica e Ciências Atmosféricas da Universidade de São Paulo, São Paulo.

A Rede Sismográfica Brasileira (RSBR) está em operação desde 2009, sendo atualmente composta por um conjunto de quase 100 estações sismográficas banda-larga, com capacidade de transmissão de dados em tempo real. O monitoramento automático de ocorrências de eventos sísmicos é realizado através do *software* *SeisComP*, com poucas modificações em seus parâmetros originais de detecção e localização. Mais de 90% dos eventos no catálogo sísmico da RSBR foram detectados manualmente por analistas, em rotinas diárias de análises dos sismogramas. Visando aperfeiçoar a detectabilidade automática de eventos sísmicos no Brasil, foi realizada uma análise da capacidade de detecção automática da RSBR com base em seu catálogo sísmico, seguida de um estudo para otimizar os parâmetros de detecção (filtros de frequência e janelas STA/LTA, limites para detecção e *picker*), utilizando eventos conhecidos. Também buscou-se otimizar os procedimentos de nucleação e associação de origens, inclusive os que estão contidos no código-fonte do *software*, de forma a maximizar a detecção de eventos reais e minimizar os falsos-positivos. Por fim, um novo modelo de velocidades foi proposto, utilizando-se como base o modelo NewBR e rotinas automatizadas de processamento para encontrar o melhor conjunto de parâmetros que minimizam o RMS dos eventos utilizados para a obtenção do modelo.

As análises indicaram, como parâmetros otimizados, um filtro passa-banda com frequências de corte de 4,5 Hz e 10 Hz, em conjunto com um *picker* AIC e janelas de tempo de 0,2 s (STA) e 45 s (LTA), para maximizar as chegadas das ondas P dos eventos regionais. Com relação ao código-fonte do módulo *scautoloc*, foram estabelecidas novas maneiras de se verificar a validade das origens nucleadas, além da remoção de parâmetros e fluxos arbitrários que prejudicam a localização automática de eventos regionais. Também propõe-se a utilização de malhas de pontos cujos números mínimos de leituras e distâncias máximas para nucleação dependem da densidade de estações em um raio de 10 graus a partir de cada nó da grade. Os resultados indicam que as modificações implementadas aprimoram a detecção automática de eventos sísmicos no Brasil, aumentando de 1024 para 5981 as detecções localizáveis entre 2014 e 2021. Dessas detecções, a quantidade de eventos concomitantes ao catálogo sísmico da RSBR passou de 78 para 292, além de possíveis eventos reais que não constam no catálogo oficial. Os parâmetros das origens apresentaram melhorias significativas, tais como a diminuição dos resíduos temporais, a possibilidade de nucleação com contribuição de poucas estações e a detecção de eventos com baixas magnitudes ($\sim M1,5$). Ademais, após as modificações, também destaca-se a possibilidade de detecção de detonações em regiões de minas e pedreiras conhecidas.

Palavras-chave: Sismologia, SeisComP, Detecção Automática, Sismicidade Regional.

Abstract

SALLES, V. **Optimization of the SeisComP software for automatic detection and location of regional seismic events in Brazil.** 2023. Master's thesis – Institute of Astronomy, Geophysics and Atmospheric Sciences, University of São Paulo, São Paulo.

The Brazilian Seismographic Network (RSBR) began operating in 2009 and is currently composed of almost 100 broadband seismographic stations with the capacity for real-time data transmission. The automatic monitoring of seismic events is carried out using the SeisComP software, with few modifications in its original detection and location parameters. More than 90% of the events in the RSBR seismic catalog were detected manually by analysts in daily seismogram analysis routines. Aiming to improve the automatic detectability of seismic events in Brazil, we analyzed the automatic detection capacity of the RSBR based on its seismic catalog. Then, a study was conducted to optimize the detection parameters (frequency filters and STA/LTA windows, limits for detection, and picker) using known events. We also sought to optimize the origins' nucleation and association procedures, including those that are part of the software source code, in order to maximize the detection of real events and minimize false positives. Finally, a new velocity model was proposed, based on the NewBR model and using automated processing routines to find the best set of parameters that minimize the RMS of the events used to obtain the model.

As optimized parameters, the analyzes indicated a band-pass filter with cutoff frequencies of 4.5 Hz and 10 Hz, an AIC picker, and time windows of 0.2s (STA) and 45s (LTA) to maximize P-wave arrivals from regional events. Regarding the source code of the scautoloc module, new ways of verifying the validity of nucleated origins were established. Furthermore, parameters and arbitrary flows that hinder the automatic location of regional events were removed. We also propose using grid points whose minimum number of wave arrivals and maximum distances for nucleation depend on the density of stations within a radius of 10 degrees from each grid node. The results indicate that the implemented modifications improve the automatic detection of seismic events in Brazil: the detections that can be located increased from 1024 to 5981 between 2014 and 2021. Of these detections, the events also present in the RSBR seismic catalog increased from 78 to 292, in addition to real events that are not in the official catalog. The origins' parameters showed significant improvements, such as the decrease in the travel-time residuals of wave arrivals, the possibility of nucleation with the contribution of a few stations, and the detection of events with low magnitudes ($\sim M1.5$). It is also worth mentioning that, after the modifications, it is possible to detect blasts in regions of known mines and quarries.

Keywords: Seismology, SeisComP, Automatic Detection, Regional Seismicity.

Sumário

Lista de Siglas e Termos	ii
Lista de Figuras	iii
Lista de Tabelas	vi
1. Introdução	1
2. Motivação: catálogo sísmico da RSBR	4
3. O SeisComp	9
3.1 Módulos de processamento <i>scautopick</i> e <i>scautoloc</i>	11
3.2 Módulo de configuração <i>scconfig</i>	13
3.3 Parâmetros e fluxos <i>hard-coded</i>	14
3.4 Modelos de velocidades	16
3.4.1 IASP91	16
3.4.2 NewBR	16
3.5 Dados paramétricos	17
4. Métodos e procedimentos	18
4.1 Desenvolvimento de um modelo de velocidades - BRA23	19
4.2 Otimização do módulo <i>scautopick</i>	23
4.2.1 Filtro de frequências	24
4.2.2 Parâmetros de STA/LTA	25
4.2.3 <i>Picker</i>	26
4.3 Otimização do módulo <i>scautoloc</i>	27
4.3.1 Malha de pontos (<i>grid.conf</i>)	28
4.3.2 Configurações de estações (<i>station.conf</i>)	30
4.3.3 Parâmetros <i>hard-coded</i>	31
4.3.3.1 Configurações (<i>config.cpp</i>)	32
4.3.3.2 Associador (<i>associator.cpp</i>)	34
4.3.3.3 Nucleador (<i>nucleator.cpp</i>)	35
5. Resultados	46
5.1 Filtro de frequências e parâmetros STA/LTA	47
5.2 Parâmetros de configuração	49
5.2.1 Módulos <i>scautopick</i> e <i>scautoloc</i>	49
5.2.2 <i>Bindings</i>	51
5.2.3 Parâmetros de configuração <i>hard-coded</i> (<i>config.cpp</i>)	52
5.3 Malhas de pontos propostas	53
5.4 Novo modelo de velocidades - BRA23	55
5.5 Detecções localizáveis (2014 - 2021)	60
5.6 Eventos concomitantes ao catálogo da RSBR	62

6. Discussões	64
6.1 Detecções localizáveis	64
6.2 Eventos concomitantes ao catálogo da RSBR	65
6.3 Estimativas de detectabilidade automática após as modificações	68
6.4 Eventos do catálogo não localizados automaticamente	71
6.5 Possíveis eventos reais não contidos no catálogo	75
7. Conclusões	77
Referências Bibliográficas	80
Apêndices	83
Apêndice 1 - Malhas de pontos (<i>grids</i>) - 2014 a 2021	83
Apêndice 2 - Modelo de velocidades 1D (BRA23)	92

Lista de Siglas e Termos

AIC - *Akaike Information Criterion*. Algoritmo utilizado para refinar as detecções de anomalias de amplitudes realizadas através do algoritmo de razão STA/LTA, agindo como *picker*. Não pode ser utilizado como detector, pois não processa os dados contínuos.

Arrival - Leitura de chegada de onda que foi associada a uma origem de evento sísmico.

Detecção localizável - Origem gerada automaticamente pelo *SeisComp*, mas que não passou por uma revisão manual para validação. Uma detecção localizável indica que a origem em questão passou por todos os testes e critérios do *software*, sendo compreendida pelo mesmo como um evento sísmico real. Entretanto, para se obter a confirmação de que trata-se de um evento sísmico, é necessária a validação de um sismólogo.

Grid - Malha de pontos utilizada para nucleação de eventos. Inicialmente, os hipocentros das origens nucleadas são definidos em um dos pontos discretizados pela malha para, em seguida, serem refinados através das relocalizações.

Hard-coded - Parâmetro ou processo definido no código-fonte do *software*, não sendo acessível para o usuário comum. Para modificar algo *hard-coded*, é necessário editar o código-fonte e recompilar o *software*.

Origem - Conjunto dos parâmetros que determinam a origem de um evento sísmico (coordenadas geográficas, horário de origem, *arrivals* etc.).

Origem preferida - Um evento sísmico pode conter diferentes origens devido a, por exemplo, adição de novos *picks* ou refinamento manual. A origem preferida de um evento é a escolhida para ser a representativa do mesmo, divulgada em boletins e catálogos sísmicos.

Pick - Detecção de anomalia de amplitude (sinal transiente) no sismograma. Um *pick* pode corresponder a uma leitura de chegada de ondas de eventos sísmicos ou a um ruído de alta amplitude que se destaca do *background* local de vibrações.

Picker - Algoritmo utilizado para realizar os *picks* no sismograma.

RMS - *Root mean square*.

SNR - *Signal-to-noise ratio*. Razão sinal-ruído de uma detecção no sismograma (*pick*).

STA/LTA - *Short-term average through long-term average*. Algoritmo utilizado para detectar anomalias de amplitude (sinais transientes) no sismograma. Pode ser utilizado como “detector e *picker*” ou apenas como “detector”. No segundo caso, o algoritmo de STA/LTA fornece uma janela de tempo para que a leitura (*pick*) seja realizada por um *picker* mais robusto (e.g., AIC).

Lista de Figuras

Figura 1.1	Estações sismológicas (triângulos) e instituições (bandeiras) participantes da RSBR.	1
Figura 1.2	Distribuição temporal das magnitudes dos sismos contidos no catálogo da RSBR no período entre 1940 e 2016.	2
Figura 1.3	Média anual da quantidade de eventos sísmicos por magnitude contidos no catálogo da RSBR, nos períodos de 1980 a 2011 e de 2012 a 2016.	3
Figura 2.1	Eventos sísmicos ocorridos entre janeiro de 2014 e dezembro de 2021 na região de interesse.	5
Figura 2.2	Relação entre magnitude e quantidade de <i>arrivals</i> das origens preferidas dos eventos sísmicos ocorridos entre 2014 e 2021 na região de interesse.	6
Figura 2.3	Relações entre magnitudes dos 1717 eventos sísmicos ocorridos entre 2014 e 2021 e distâncias máximas em que os mesmos foram registrados.	7
Figura 2.4	Diagramas de caixas das distâncias entre as estações da RSBR.	8
Figura 3.1	Fluxograma de funcionamento do <i>SeisComP</i> e seus módulos.	9
Figura 3.2	Fluxograma dos principais aspectos dos módulos <i>scautopick</i> e <i>scautoloc</i>	11
Figura 3.3	Exemplo da interface gráfica do módulo de configuração <i>scconfig</i>	13
Figura 3.4	Recortes do código-fonte do módulo <i>scautoloc</i> nos quais há modificações internas dos parâmetros definidos pelo usuário.	15
Figura 3.5	Exemplo da organização hierárquica dos dados paramétricos de um evento sísmico no formato <i>Sc3ML</i>	17
Figura 4.1	Fluxograma indicando os dados utilizados durante cada etapa de otimizações dos parâmetros dos módulos <i>scautopick</i> e <i>scautoloc</i>	18
Figura 4.2	Tempos de percurso para cada um dos 17 eventos utilizados para derivar o modelo de velocidades BRA23, em comparação com as estimativas obtidas através do modelo NewBR.	20
Figura 4.3	Histogramas de resíduos das leituras de chegadas de ondas utilizando-se o modelo NewBR.	21
Figura 4.4	Esquematização do cálculo do tempo de trânsito da frente de onda entre o hipocentro do evento e a estação.	22
Figura 4.5	Análise das combinações de filtros de frequência para um evento sísmico de magnitude M1,5 registrado pela estação NB.NBMA.	24
Figura 4.6	Análise das combinações de janelas de tempo do algoritmo STA/LTA para um evento sísmico de magnitude M1,5 registrado pela estação NB.NBMA.	25
Figura 4.7	Malha de pontos padrão do <i>SeisComP</i> , em comparação com a malha de pontos proposta.	29

Figura 4.8	Fluxograma dos processos contidos no módulo <i>scautoloc</i> que sofreram as principais modificações durante o trabalho.	31
Figura 4.9	Representação gráfica do novo cálculo do parâmetro <i>distScore</i>	39
Figura 4.10	Representação gráfica de um exemplo da obtenção da magnitude MLv preliminar.	40
Figura 4.11	Representação gráfica do novo cálculo do parâmetro <i>amplScore</i>	41
Figura 4.12	Representação gráfica do novo cálculo do parâmetro <i>timeScore</i>	42
Figura 5.1	<i>Corner plot</i> dos principais parâmetros dos 170 <i>arrivals</i> utilizados para se realizar a primeira etapa de otimização dos parâmetros de detecção. .	46
Figura 5.2	Exemplos de filtragens das formas de onda referentes a um evento sísmico de magnitude M2,5 ocorrido em Olho d'Água Grande/AL.	47
Figura 5.3	Exemplos de filtragens e aplicação de STA/LTA nas formas de onda referentes a um evento sísmico de magnitude M2,5 ocorrido em Olho d'Água Grande/AL.	48
Figura 5.4	Disponibilidade de dados em forma de ondas registrados em 2019 pelas estações contidas no inventário do <i>SeisComP</i> do IAG-USP e as localizações das estações com no mínimo 50% de dados disponíveis no período.	53
Figura 5.5	Malha de pontos referente ao ano de 2019.	54
Figura 5.6	Mapa com os 17 eventos e estações utilizadas na construção do modelo BRA23.	55
Figura 5.7	Modelos de velocidade de ondas P derivados da inversão global e local.	57
Figura 5.8	Histogramas do número de testes realizados durante as otimizações globais, referentes a cada uma das variáveis otimizadas.	58
Figura 5.9	Comparação dos resíduos temporais ($t_{model} - t_{obs}$) utilizando-se os modelos BRA23 e NewBR.	59
Figura 5.10	Detecções localizáveis obtidas utilizando-se o <i>SeisComP</i> original.	60
Figura 5.11	Detecções localizáveis obtidas utilizando-se o <i>SeisComP</i> modificado.	61
Figura 5.12	Epicentros das 78 detecções localizáveis concomitantes ao catálogo sísmico da RSBR de 2014 a 2021, utilizando-se o <i>SeisComP</i> original.	62
Figura 5.13	Epicentros das 292 detecções localizáveis concomitantes ao catálogo sísmico da RSBR de 2014 a 2021, utilizando-se o <i>SeisComP</i> modificado.	63
Figura 6.1	Gráficos de setores das porcentagens de detecções localizáveis classificadas, ocorridas no período de 2014 a 2021.	64
Figura 6.2	Número de eventos concomitantes ao catálogo da RSBR após as modificações no <i>SeisComP</i> , considerando diferentes combinações de critérios de concomitância.	65
Figura 6.3	Histogramas de magnitudes dos eventos concomitantes ao catálogo da RSBR.	66

Figura 6.4	Histogramas de RMS das origens dos eventos concomitantes ao catálogo da RSBR.	66
Figura 6.5	Histogramas de resíduos temporais absolutos dos <i>arrivals</i> dos eventos concomitantes ao catálogo da RSBR.	67
Figura 6.6	Histogramas de números de <i>arrivals</i> utilizados (<i>usedPhaseCount</i>) nas localizações dos eventos concomitantes ao catálogo da RSBR.	67
Figura 6.7	Estimativas de detectabilidade automática de eventos sísmicos no Brasil.	68
Figura 6.8	Estimativas de detectabilidade automática de eventos sísmicos no Brasil, com sobreposição dos eventos concomitantes ao catálogo da RSBR obtidos após as modificações no <i>SeisComP</i>	69
Figura 6.9	Exemplo de chegadas de onda emergentes do evento “usp2019jnud” (M2,1) ocorrido em 16/05/2019, as quais são dificilmente diferenciadas do ruído local.	71
Figura 6.10	Exemplo de falta de registros em forma de onda no banco de dados do IAG-USP, acarretando em uma menor quantidade de eventos detectáveis.	72
Figura 6.11	Estimativas de detectabilidade automática de eventos sísmicos no Brasil, com sobreposição dos eventos do catálogo da RSBR que não foram localizados automaticamente após as modificações no <i>SeisComP</i> .	73
Figura 6.12	Histograma de magnitudes dos eventos contidos no catálogo da RSBR e que não foram detectados automaticamente após as modificações no <i>SeisComP</i>	74
Figura 6.13	Exemplo de dois eventos sísmicos localizados automaticamente após as modificações propostas e que não constam no catálogo oficial da RSBR.	75
Figura 6.14	Aglomerados de detecções localizáveis possivelmente associadas a detonações em regiões de minas conhecidas.	76

Lista de Tabelas

Tabela 2.1	Medianas das distâncias entre as estações das sub-redes da RSBR.	8
Tabela 4.1	Parâmetros otimizados durante o procedimento de obtenção do modelo BRA23.	22
Tabela 5.1	Parâmetros de configurações do módulo <i>scautopick</i> acessíveis através da interface gráfica do <i>scconfig</i>	49
Tabela 5.2	Parâmetros de configurações do módulo <i>scautoloc</i> acessíveis através da interface gráfica do <i>scconfig</i>	50
Tabela 5.3	Parâmetros de configurações individuais das estações (“ <i>bindings</i> ”), acessíveis através da interface gráfica do <i>scconfig</i>	51
Tabela 5.4	Parâmetros <i>hard-coded</i> contidos no arquivo <i>config.cpp</i>	52
Tabela 5.5	Solução final dos parâmetros otimizados.	57

1. Introdução

A Rede Sismográfica Brasileira (RSBR) iniciou suas operações após algumas décadas de tentativas de estabelecimento de um estudo contínuo da sismicidade em território brasileiro e sul-americano. Diversas iniciativas se sucederam desde o início dos anos 1900 até que, em 1970, a sismologia no Brasil teve um avanço com a formação de grupos sismológicos em diferentes regiões do país, com a finalidade de se estudar o perigo sísmico relacionado a usinas nucleares e às ocorrências de sismicidades induzidas (Bianchi et al., 2018).

Na atualidade, a RSBR possui quase 100 estações sismográficas banda-larga capazes de transmitir dados em tempo real, alimentando continuamente os computadores de processamento dos quatro principais centros sismológicos do país: a Universidade de São Paulo (USP), a Universidade de Brasília (UnB), a Universidade Federal do Rio Grande do Norte (UFRN) e o Observatório Nacional (ON), como mostra a Figura 1.1.

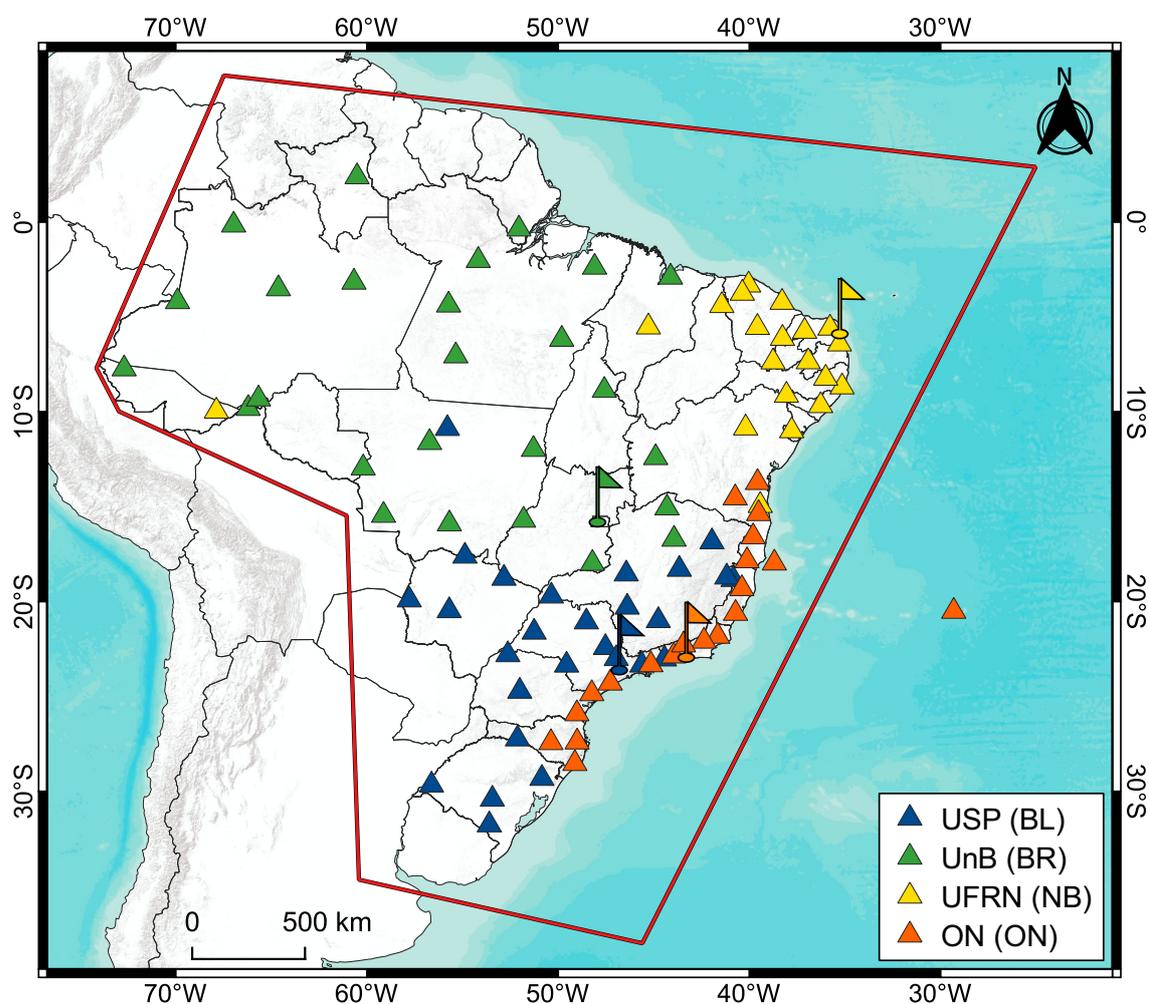


Figura 1.1: Estações sismológicas (triângulos) e instituições (bandeiras) participantes da RSBR. Sub-redes e suas respectivas instituições responsáveis estão indicadas por diferentes cores. Fonte das informações: banco de dados do IAG-USP (outubro/2022).

Cada centro sismológico é responsável pela instalação, operação e manutenção de estações sismográficas em determinadas regiões do território nacional. A rede brasileira é composta por quatro sub-redes (BL, BR, ON e NB), variando conjuntos de instrumentação e tecnologias, mas mantendo um padrão mínimo previamente acordado (Bianchi et al., 2012). É importante ressaltar que, além das estações fixas, existem estações temporárias que são operadas por cada grupo e atuam de forma complementar à RSBR.

Atualmente, o Instituto de Astronomia, Geofísica e Ciências Atmosféricas da Universidade de São Paulo (IAG-USP) realiza a localização de eventos sísmicos para a RSBR, enquanto o Observatório Nacional (ON) realiza a divulgação desses eventos. Todos os eventos registrados pelas estações da RSBR passam por uma revisão manual, realizada pelo IAG, para a definição de sua “origem preferida”. Os dados revisados são enviados para o *International Seismological Centre* (ISC) para compor seu boletim publicado regularmente, juntamente com parâmetros de qualidade das redes.

Com o passar do tempo e a instalação de novas estações sismográficas pertencentes à RSBR, a detectabilidade de eventos sísmicos foi sendo aprimorada. Com isso, Assumpção et al. (2014) inferiu a possibilidade de se detectar, após revisão manual dos dados, sismos em território nacional com magnitudes acima de $M_{3,5}$ a partir do ano de 1980. Por outro lado, Bianchi et al. (2018) aponta que, apesar desse valor ser adequado para a maior parte do Brasil, provavelmente é superestimado para a região da Amazônia, onde há menos estações sismográficas em operação.

A partir da distribuição temporal das magnitudes dos eventos contidos no catálogo da RSBR, Bianchi et al. (2018) estimou a detectabilidade de eventos sísmicos no Brasil entre os anos de 1940 e 2016, como mostra a Figura 1.2. Segundo o autor, a instalação de estações da RSBR na região da Amazônia em 2014 viabilizou a detectabilidade de eventos com magnitudes a partir de $M_{3,5}$ para todo o Brasil.

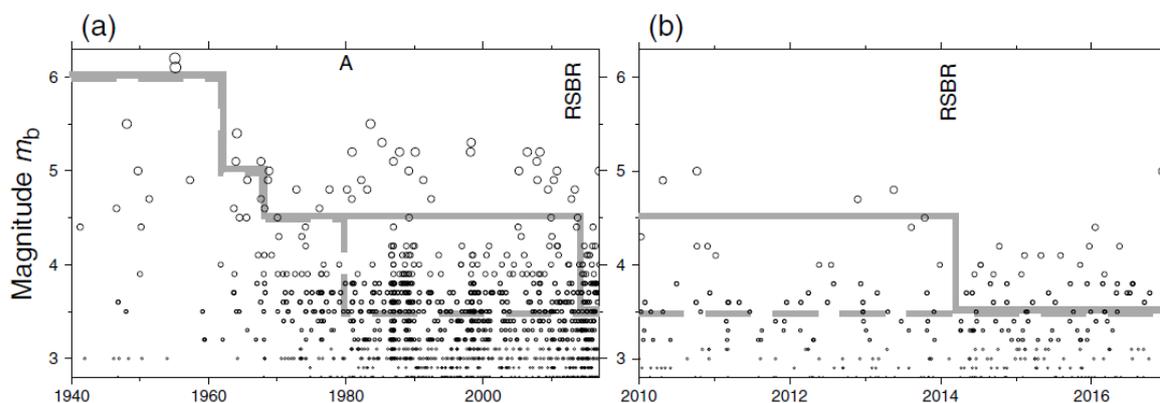


Figura 1.2: Distribuição temporal das magnitudes dos sismos contidos no catálogo da RSBR. (a) Dados entre 1940 e 2016, (b) ampliação entre 2010 e 2016. “A” indica o aumento do número de estações instaladas no Brasil, especialmente no Nordeste, Sudeste e Amazônia. “RSBR” indica a instalação de estações da RSBR na Amazônia em 2014. A linha cinza tracejada indica limites de detecção inferidos por Assumpção et al. (2014), enquanto a linha cinza contínua indica limites de detecção inferidos pelo autor. Fonte: Bianchi et al. (2018).

Bianchi et al. (2018) ressalta que, desde 2012, a RSBR detecta, anualmente, o dobro de eventos que apresentam magnitudes entre M3,5 e M4,5, quando comparado ao que era detectado no período anterior ao estabelecimento da rede sismográfica nacional (Figura 1.3). Para algumas regiões (Sudeste e Nordeste), nas quais a densidade de estações é maior, o limite de detecção é de M3,0 ou até menor.

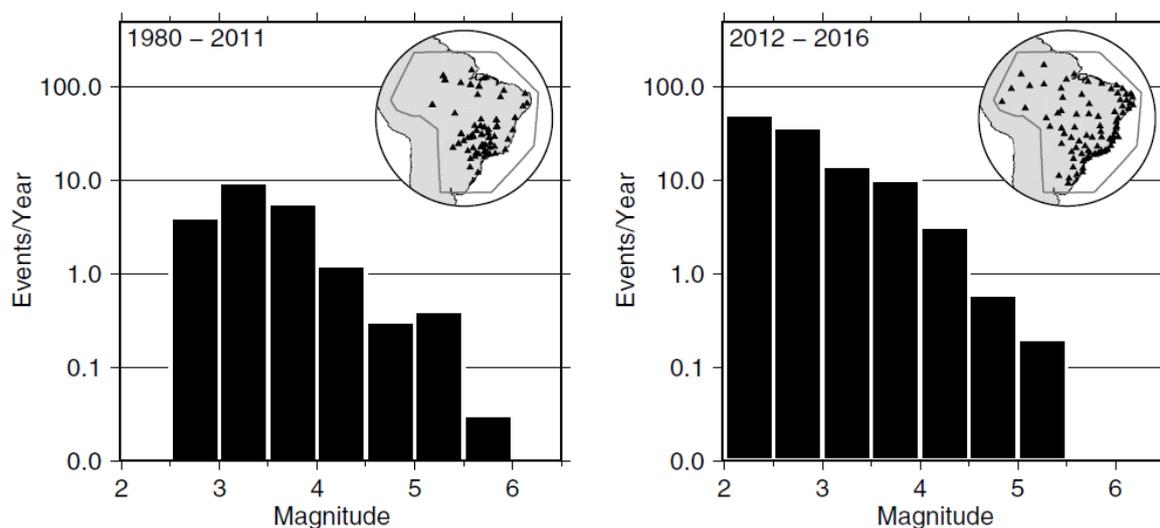


Figura 1.3: Média anual da quantidade de eventos sísmicos por magnitude contidos no catálogo da RSBR, nos períodos de 1980 a 2011 (esquerda) e de 2012 a 2016 (direita), referentes a antes e depois do início da instalação de estações da RSBR, respectivamente. No detalhe, as principais estações operando em cada período. Fonte: Bianchi et al. (2018).

Um dos fatores que podem aumentar a detectabilidade de eventos pela RSBR é a avaliação e correta configuração dos parâmetros e processos realizados pelo sistema de detecção automática utilizado pela rede, de forma a otimizar a detecção e localização de eventos em território brasileiro.

Lopez (2021) realizou uma otimização dos parâmetros de detecção de chegadas de onda nos sismogramas para uma parte das estações da RSBR, mas sem considerar os registros contínuos, ou seja, utilizando apenas recortes de formas de onda que contêm eventos sísmicos conhecidos. Segundo o autor, as otimizações aumentaram a quantidade de *picks* automáticos, contribuindo para o aumento das localizações automáticas dos eventos nas janelas de tempo selecionadas. Entretanto, também destaca-se o aumento da quantidade de *picks* associados a ruídos, os quais prejudicam as localizações em um contexto de processamento em tempo real.

Em testes realizados utilizando os parâmetros propostos por Lopez (2021) em registros contínuos de formas de onda, não foi possível realizar a localização de eventos sísmicos devido ao grande número de “falsos *picks*” gerados. Tal fato indica que, para buscar a otimização da localização de eventos sísmicos em tempo real, os parâmetros devem ser obtidos considerando todos os registros contínuos e não apenas janelas de tempo que contêm eventos conhecidos.

2. Motivação: catálogo sísmico da RSBR

A *International Federation of Digital Seismograph Networks* (FDSN) é uma federação que regula a operação de redes sísmológicas ao redor do mundo, de forma a padronizar o armazenamento e distribuição de dados paramétricos e de formas de ondas (Clark et al., 2014). Ela é responsável por emitir códigos de redes exclusivos e individuais, além de coordenar a cooperação entre grupos de sismologia internacionais.

A RSBR registra os seus códigos de rede na FDSN e adere às suas normas, buscando distribuir seus dados utilizando tais padrões. Os dados em formas de ondas armazenados pela RSBR estão em formato “SEED” (*Standard for Exchange of Earthquake Data*) (IRIS, 2010), enquanto os metadados dos eventos são armazenados no formato *QuakeML* (Schorlemmer et al., 2011), descrito adiante. O acesso aos dados e metadados é realizado de maneira simples e padronizada através do *FDSN Web Services* (FDSNWS), uma interface para a requisição de dados de centros sísmológicos que aderiram aos serviços oferecidos pela FDSN.

Com isso, é possível obter as informações necessárias para avaliar a efetividade do sistema de detecção automática de eventos sísmicos no Brasil. A Figura 2.1 apresenta dois mapas com os eventos ocorridos entre 2014 e 2021 na região de interesse, bem como os histogramas de magnitudes considerando os modos de detecção.

Os eventos sísmicos contidos no catálogo da RSBR apresentam diversas escalas de magnitude (*e.g.*, ML_v , m_R , mb , M_w). Cada escala é adotada de acordo com os parâmetros epicentrais dos eventos e é baseada em diferentes aspectos do sismograma, tais como amplitudes das ondas P, duração do evento, entre outros (Duda and Nuttli, 1974; Kanamori, 1983). No Brasil, os eventos locais possuem apenas a magnitude local (ML_v) computada, enquanto os eventos regionais utilizam a magnitude m_R (Assumpção, 1983) que, no contexto da América do Sul, é equivalente à escala mb telessísmica. Nesse trabalho, para efeito de comparação, as magnitudes de todos os eventos sísmicos são referidas como M , independentemente do tipo de magnitude associada.

A maioria dos sismos confirmados que foram detectados e localizados automaticamente na região de interesse possuem origens associadas a bordas de placas, seja na região do estado do Acre (devido à subducção da Placa de Nazca sob a Placa Sul-Americana) ou na região da dorsal mesoatlântica. Considerando apenas os eventos intraplaca, foram registrados 1717 sismos durante o período, dos quais apenas 106 (6,2%) apresentam ao menos uma origem calculada de forma automática. Os principais aspectos que podem ocasionar a baixa efetividade do sistema de detecção automática são (i) o espaçamento médio entre as estações da RSBR, o qual impede que um evento sísmico de baixa magnitude seja registrado por vários sismógrafos e (ii) os parâmetros e fluxos de processamento do sistema utilizado para detecção automática, os quais devem ser configurados e otimizados considerando-se o contexto da RSBR.

Para a melhoria do primeiro fato destacado, torna-se necessária a adição de novas estações à RSBR. Para isso, é imprescindível investimentos financeiros, a curto e a longo prazo, para a instalação e a manutenção de tais estações, de forma a proporcionar registros de dados com alta qualidade durante longos períodos de tempo. Considerando o segundo aspecto, a detectabilidade automática de eventos sísmicos pode ser aperfeiçoada mesmo mantendo o atual arranjo de estações da rede. Para isso, é necessário avaliar e reestruturar os diversos parâmetros e processos realizados pelo *software* que gerencia e processa os dados registrados pela RSBR, o *SeisComP* (Helmholtz-Centre Potsdam - GFZ German Research Centre for Geosciences and gempa GmbH, 2008).

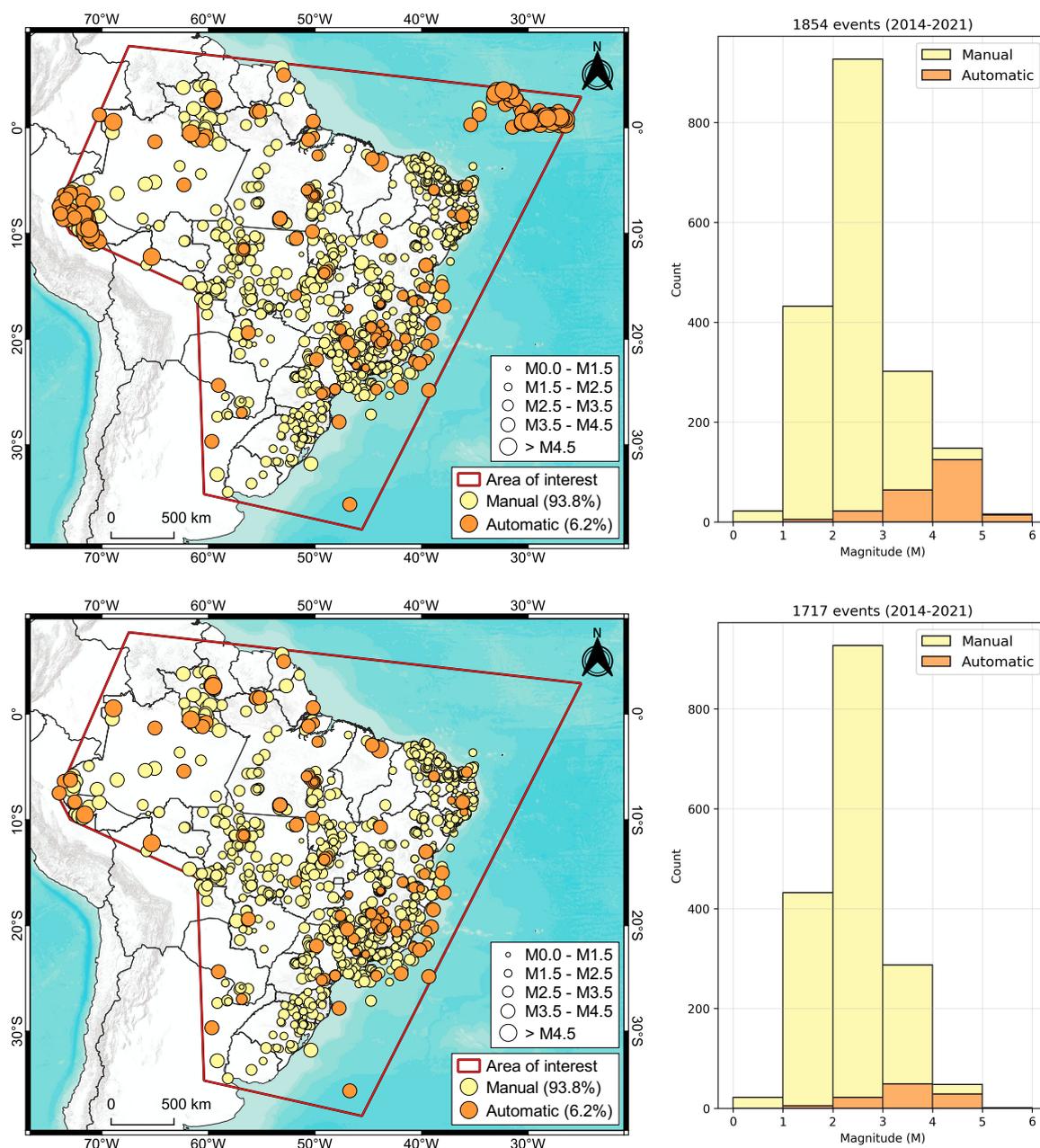


Figura 2.1: Todos os eventos sísmicos ocorridos entre janeiro de 2014 e dezembro de 2021 na região de interesse (acima) e desconsiderando-se os eventos profundos com origens associadas às bordas de placas (abaixo). Fonte das informações: IAG-USP (outubro/2022).

Uma estimativa da detectabilidade máxima que pode ser alcançada com o atual número de estações da RSBR pode ser inferida ao se analisar o catálogo do Centro de Sismologia do IAG-USP. Para isso, o gráfico à esquerda na [Figura 2.2](#) mostra a relação entre magnitude e o número de *arrivals* das origens preferidas de todos os eventos sísmicos registrados entre janeiro de 2014 e dezembro de 2021 na região de interesse mostrada na [Figura 2.1](#). À direita, consta a mesma relação, entretanto, considerando apenas as chegadas de onda (*arrivals*) das últimas origens automáticas dos eventos (quando existentes).

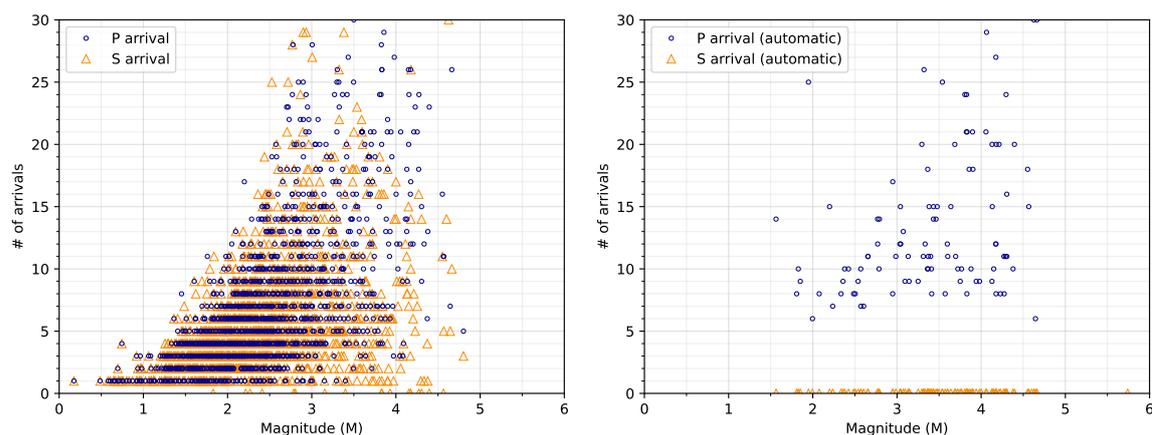


Figura 2.2: Esq.: relação entre magnitude e quantidade de *arrivals* das origens preferidas dos eventos sísmicos ocorridos entre 2014 e 2021 na região de interesse. Dir.: mesma relação, mas utilizando apenas as últimas origens automáticas dos eventos, quando existentes.

Por padrão, o *SeisComP* precisa de no mínimo 6 leituras de chegada de onda P em diferentes estações para poder nuclear um evento automaticamente. Para processamentos manuais, o *software* pode ser capaz de localizar eventos com menos estações, como mostra o gráfico à esquerda da [Figura 2.2](#).

Do gráfico à direita, é possível constatar que os eventos localizados de maneira automática possuem, em sua maioria, magnitudes maiores do que M2,5. Entretanto, comparando com o gráfico à esquerda, é possível verificar que diversos eventos com magnitudes menores do que M2,5 não foram detectados automaticamente pelo *SeisComP*, mesmo possuindo 6 ou mais leituras em suas origens preferidas. Esse fato pode indicar a falha do sistema em identificar automaticamente as chegadas de ondas de eventos sísmicos reais, sendo atualmente necessária uma verificação manual dos dados em forma de ondas para que esses eventos sejam localizados.

Também é importante destacar a grande quantidade de eventos com magnitudes relativamente altas e que possuem poucas leituras, o que possivelmente indica um erro no cálculo de suas magnitudes (*e.g.*, eventos com magnitudes maiores do que M3,0 e com apenas uma leitura de chegada de onda P).

Uma outra maneira de avaliar o limiar de detectabilidade de eventos sísmicos no Brasil é analisando a relação entre magnitude e distância máxima registrada, uma vez que a densidade de estações em território nacional não é uniforme e, por esse motivo, eventos que

não são registrados a grandes distâncias tendem a não serem localizados por não apresentarem chegadas de ondas em várias estações.

Bianchi et al. (2012) estima que um evento sísmico de magnitude M2,5 é registrado a no máximo 150 km de distância, enquanto um evento de magnitude M3,5 é registrado a no máximo 500 km. Finalmente, segundo o autor, um evento de magnitude M4,0 pode ser registrado a uma distância de até 1200 km.

Para verificar as estimativas de Bianchi et al. (2012), a Figura 2.3 mostra a relação entre magnitude e distância máxima registrada, obtida através dos dados de todos os eventos sísmicos contidos no catálogo da RSBR, ocorridos entre 2014 e 2021 na região de interesse.

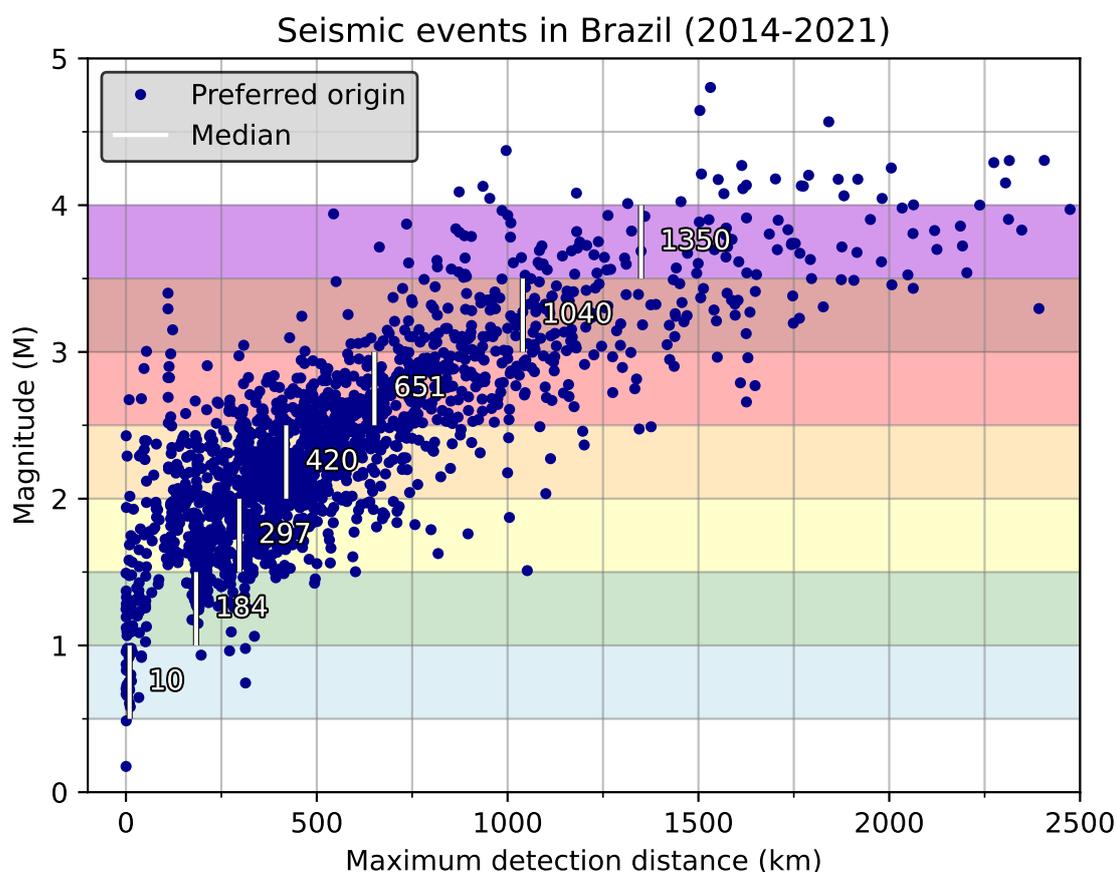


Figura 2.3: Relações entre magnitudes dos eventos sísmicos ocorridos entre 2014 e 2021 e distâncias máximas em que os mesmos foram registrados. As linhas brancas verticais indicam as medianas das distâncias para cada faixa de magnitude entre M0,5 e M4,0.

Nota-se que a mediana das distâncias máximas de registro dos eventos sísmicos com magnitudes entre M1,5 e M2,0 é 300 km, enquanto os eventos sísmicos com magnitudes entre M2,0 e M2,5 apresentam um valor mediano de 420 km. Portanto, o gráfico indica que os eventos sísmicos em território brasileiro são detectados a distâncias maiores do que Bianchi et al. (2012) estimou. Esses dados devem ser considerados ao se definir as distâncias máximas nas quais estações sismológicas podem contribuir com leituras de chegada de onda, visando mitigar a adição de possíveis ruídos em estações muito distantes, nas quais não se esperava registros de um determinado evento.

Adicionalmente, convém mensurar as distâncias entre as estações que compõem as sub-redes da RSBR. Para isso, a [Figura 2.4](#) mostra dois diagramas de caixa: o primeiro refere-se a todas as distâncias entre as estações de cada sub-rede, enquanto o segundo considera apenas as menores distâncias para cada estação.

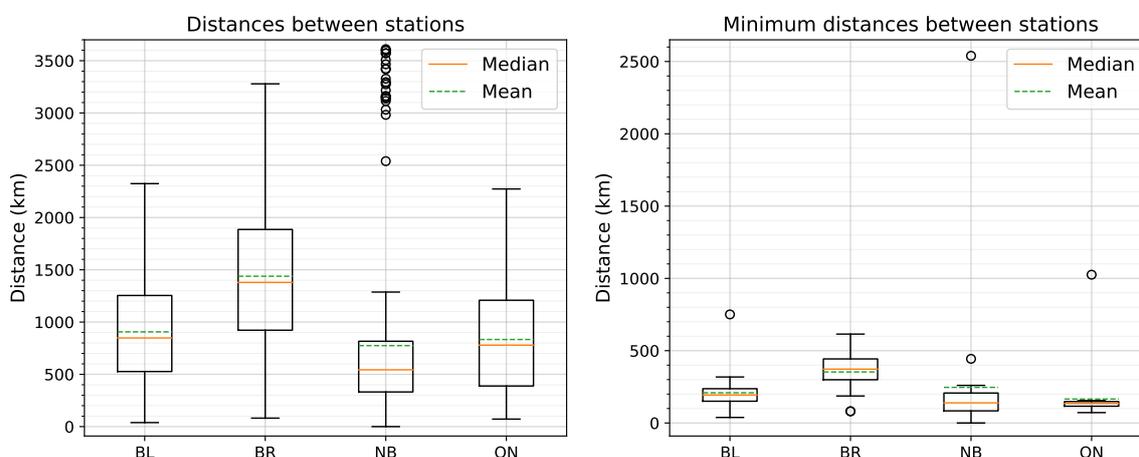


Figura 2.4: Diagramas de caixas das distâncias entre as estações da RSBR. Esq.: todas as distâncias entre as estações de cada sub-rede. Dir.: menores distâncias entre duas estações de cada sub-rede.

A [Tabela 2.1](#) apresenta as medianas das distâncias entre as estações das sub-redes da RSBR, apresentadas graficamente na [Figura 2.4](#).

Tabela 2.1: Medianas das distâncias entre as estações das sub-redes da RSBR. A linha “Total” refere-se às medianas das distâncias entre todos os pares de estações de uma mesma sub-rede, enquanto a linha “Mínimas” indica as medianas das menores distâncias entre duas estações da sub-rede.

	BL	BR	NB	ON
Total (km)	906	1379	543	779
Mínimas (km)	194	373	139	135

Ao se considerar a RSBR como um todo, há uma mediana de aproximadamente 159 km de distância mínima entre as estações. Comparando esse resultado com o gráfico da [Figura 2.3](#), sugere-se que o atual arranjo da RSBR deve ser capaz de detectar automaticamente mais eventos sísmicos com magnitudes menores do que M3,5. Ao analisar as sub-redes de forma individual, infere-se que as sub-redes BL, NB e ON devem possuir capacidades similares de detecção, enquanto a sub-rede BR, responsável por cobrir a Região Norte e parte da Centro-Oeste, apresenta estações mais espaçadas entre si e, conseqüentemente, menor capacidade de detecção de eventos com menores magnitudes.

Dessa forma, o objetivo desse trabalho foi otimizar os parâmetros e procedimentos realizados durante os processos de detecção e localização automática de eventos sísmicos no contexto da RSBR, de forma a se obter uma maior eficiência de detecção automática dos eventos sísmicos regionais ocorridos no Brasil.

3. O SeisComp

Nas primeiras etapas da formação da RSBR, o IAG-USP voltou seus esforços para encontrar soluções de *softwares* que facilitassem a cooperação entre diferentes universidades e institutos na operação da rede sismográfica. O *software* escolhido para a aquisição e processamento em tempo real foi o *SeisComp*, desenvolvido no Centro de Pesquisa Alemão para Geociências (*GeoForschungsZentrum - GFZ*), o qual possui algumas versatilidades como, por exemplo, a possibilidade de aprimoramento de ferramentas e desenvolvimento de novas soluções (Pirchner et al., 2011). A versão do *software* utilizada atualmente pelo Centro de Sismologia da USP e explorada nesse trabalho é a “*SeisComp3 Jakarta 2020.330.06*”, entretanto, as versões subsequentes não sofreram alterações nos módulos de processamento. Com isso, constatou-se que a última versão disponibilizada atualmente (*SeisComp 5.3.0*) também pode ser otimizada utilizando-se os mesmos procedimentos aqui descritos.

O funcionamento do *SeisComp* é baseado em diversos módulos independentes, cada um realizando processamentos específicos que contribuem para a localização de um evento sísmico. O fluxograma de operação desses módulos é mostrado na Figura 3.1.

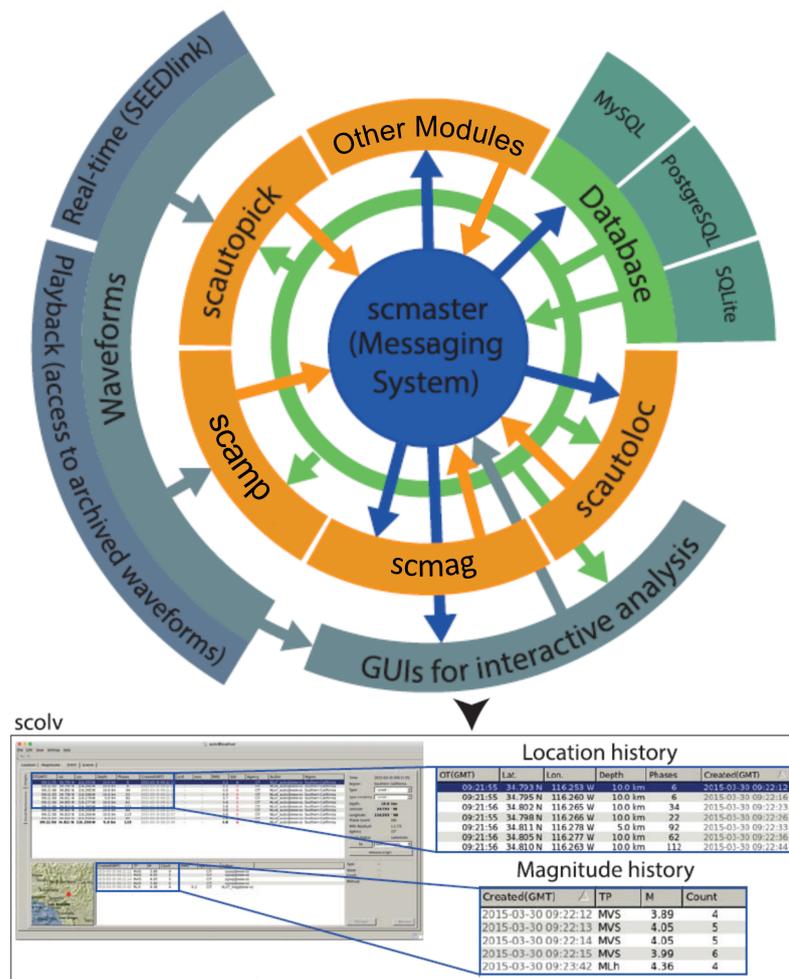


Figura 3.1: Fluxograma de funcionamento do *SeisComp* e seus módulos. As setas indicam a direção do fluxo de informação, coloridas de acordo com o responsável pelo envio. Fonte: adaptado de Behr et al. (2016).

A Figura 3.1 mostra que a detecção, localização e catalogação de eventos sísmicos são realizadas através de uma sucessão de operações realizadas pelos módulos do *software*. Os módulos representados pela cor laranja são de processamento, coordenados pelo módulo de controle (*scmaster*). Os dados são lidos por todos os módulos e escritos, apenas pelo módulo de controle, para diferentes tipos de base de dados suportadas. Análises interativas e revisão dos dados podem ser realizadas através do módulo *scolv*, uma interface gráfica do usuário (GUI) que fornece acesso ao banco de dados e formas de onda, entre outras informações.

Alguns dos principais módulos de processamento e as breves descrições de seus processos executados são:

- ***scautopick***: procura por sinais transientes, em forma de anomalias de amplitude, no dado em forma de onda. Aplica um algoritmo de razão STA/LTA (*short-time-average through long-time-average*) (Allen, 1978) que, quando atinge um determinado valor, cria uma leitura de chegada de onda (*pick*) no instante de tempo em que esse limite é excedido, ou então utiliza algoritmos mais robustos para refinar a leitura inicial;
- ***scautoloc***: responsável por nuclear origens, associar leituras e localizar eventos sísmicos de forma automática e em tempo real. Realiza a análise contínua dos *picks*, tentando nuclear uma nova origem, ou seja, identificar combinações de *picks* que correspondam a um mesmo evento sísmico. Caso seja produzida uma localização consistente com determinados critérios, ela é reportada e passada adiante para outros módulos que recebem origens como dado de entrada;
- ***scamp***: calcula amplitudes das formas de onda utilizando diferentes janelas de tempo, com base nos *picks* previamente obtidos;
- ***scmag***: calcula diferentes tipos de magnitudes a partir das amplitudes calculadas pelo módulo *scamp*. No IAG-USP, esse módulo também está configurado para calcular magnitudes na escala m_R (Assumpção, 1983); e
- ***scevent***: associa uma origem previamente calculada a um evento já existente ou, caso não encontre compatibilidade, cria um novo evento e associa a origem a ele, desde que atingidos os critérios estabelecidos. É importante ressaltar que um único evento sísmico pode conter diversas origens calculadas, uma vez que novas origens podem ser obtidas a partir dos *arrivals* processados após a primeira nucleação.

3.1 Módulos de processamento *scautopick* e *scautoloc*

Os módulos *scautopick* e *scautoloc* são responsáveis por uma série de processos importantes na detecção automática de eventos sísmicos. Neles, diversos parâmetros são definidos para que a automatização seja a mais efetiva possível, evitando falsos-positivos e falsos-negativos. Também são nesses módulos que grande parte da capacidade de detecção de eventos regionais com magnitudes relativamente baixas ($\leq M3,5$) pode ser aprimorada, ajustando parâmetros e métodos de processamento que permitam que o *SeisComP*, inicialmente desenvolvido para detecção de sismos de escalas globais, também localize de forma satisfatória os eventos regionais. A Figura 3.2 apresenta um fluxograma dos principais aspectos de ambos os módulos, bem como a relação entre eles.

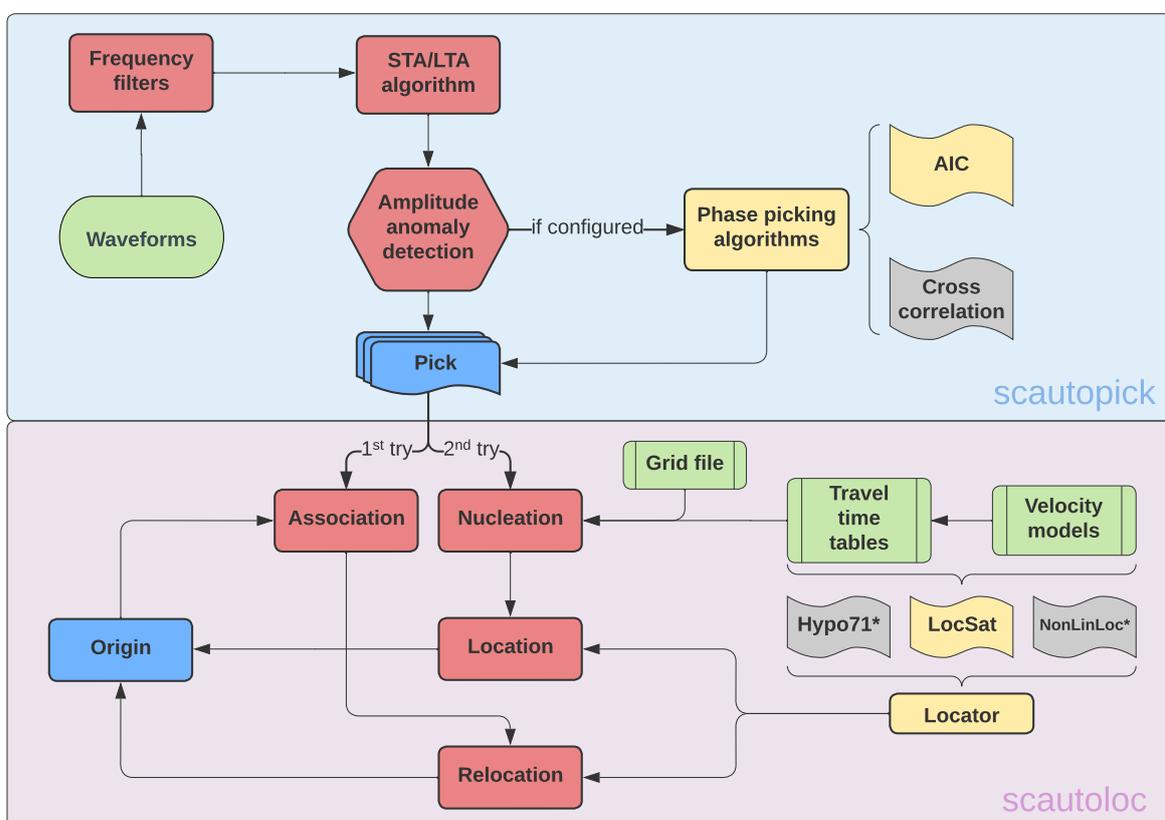


Figura 3.2: Fluxograma dos principais aspectos dos módulos *scautopick* e *scautoloc*. Em verde, são representados os dados de entrada necessários para o processamento. Em vermelho, constam os principais processos realizados por cada módulo. Em amarelo, algoritmos implementados no *software* para contribuir no processo de detecção e localização e, em cinza, algoritmos possíveis de serem utilizados, caso sejam configurados. Em azul, apresentam-se os produtos gerados por cada módulo. (*) Localizadores utilizados apenas durante a relocação de origens existentes.

Os dados em forma de onda são lidos pelo módulo *scautopick*, o qual aplica um filtro de frequências e, em seguida, utiliza um algoritmo de STA/LTA (Allen, 1978) para detectar sinais transientes, em formas de anomalias de amplitude, nos sismogramas. Quando tais anomalias são detectadas, uma marcação no tempo (*pick*) é criada e armazenada para ser utilizada durante a localização dos eventos.

Em seguida, caso esteja configurado, um algoritmo de AIC (*Akaike Information Criterion*) (Akaike, 1971) pode ser utilizado para aperfeiçoar o *pick*, ajustando a marcação, no tempo, para coincidir com o momento de chegada da onda P na estação sismográfica. Algoritmos de AIC são comumente utilizados na sismologia para detecção de chegadas de ondas (Li et al., 2016; Yadav and Mishra, 2014), proporcionando resultados mais acurados do que os obtidos por algoritmos de STA/LTA por detectar alterações de estacionariedade, ou seja, identificando mudanças de comportamento do sinal dentro da janela indicada.

O *pick* é recebido pelo módulo de localização (*scautoloc*), que inicialmente tenta associá-lo a uma origem existente. Caso não seja possível, o módulo utiliza esse e outros *picks* não associados para tentar nuclear uma nova origem, identificando combinações que correspondam a um mesmo evento sísmico. Essa etapa é realizada através de uma busca em grade (*grid search*), na qual todos os *picks* não associados a uma origem são projetados em todos os pontos da malha. Para cada *pick* e cada ponto da grade, projeta-se o horário de origem do evento, utilizando-se a tabela de tempos de percurso fornecida em conjunto com a distância da estação que registrou o *pick* até o ponto da grade em questão. Dessa forma, busca-se encontrar possíveis hipocentros com os menores resíduos de tempo de percurso associados, sendo esse um processo computacionalmente custoso, por realizar milhares de cálculos envolvendo todos os *picks* e pontos do *grid*.

Os principais processos realizados pelo *scautoloc* podem ser resumidos nas seguintes etapas, segundo a documentação oficial do *SeisComP*:

- **Associação:** inicialmente, o *scautoloc* tenta associar um *pick* recebido a uma origem já existente, especialmente com origens que já possuem uma grande quantidade de *arrivals*, de forma a melhorar suas localizações. Caso não seja possível fazer essa associação por incompatibilidade de critérios, é realizado o processo de nucleação;
- **Nucleação:** caso o processo de associação falhe, o *scautoloc* tenta criar uma nova origem a partir dessa e outras leituras anteriormente recebidas e não associadas a nenhuma origem. A nucleação é uma busca em grade, em termos de espaço e tempo, em uma malha de pontos (*grid*) que abrange a área de interesse. Essa malha pode ser modificada de acordo com as necessidades do monitoramento sismológico em questão. Durante o *grid search*, cada ponto da malha é considerado um possível hipocentro, ao mesmo tempo em que verifica-se a compatibilidade das projeções de horários de origem realizadas para cada *pick*, considerando a tabela de tempos de percurso fornecida. Nesse processo, o *software* analisa continuamente, no tempo, a densidade de leituras em cada ponto. Caso o conjunto de leituras projetadas cumpra com todos os requisitos previamente configurados e seja possível obter uma localização, ele é aceito como uma nova origem válida de um evento sísmico;
- **Refinamento de origens:** origens criadas a partir da nucleação ou até mesmo da associação podem estar contaminadas por fases erroneamente interpretadas como chegadas de ondas “P”. No processo de refinamento de origens, o *scautoloc* tenta

melhorar as origens baseando-se, por exemplo, nas razões sinal-ruído das chegadas de onda e suas amplitudes, além de realizar novas tentativas de associar leituras que não foram associadas previamente; e

- **Filtragem de origens:** esse processo resume-se a uma última checagem de consistência e compatibilidade com critérios das origens recém-criadas ou atualizadas. Durante esse processo, as origens não são modificadas, apenas descartadas ou validadas.

Durante os processos realizados pelo *scautoloc*, o *pick* passa por diversas etapas, as quais contém fluxos e parâmetros pré-definidos nos arquivos de configuração do módulo e no código-fonte do *software*. Portanto, para otimizar tais processamentos, torna-se necessário otimizar os parâmetros disponíveis para modificação pelo usuário, bem como realizar ajustes diretamente no código-fonte do *SeisComP* e recompilá-lo. Também é de extrema importância a utilização de uma malha de pontos apropriada para a região de interesse, a qual é utilizada durante o processo de nucleação de novos eventos.

3.2 Módulo de configuração *scconfig*

O *SeisComP* possui um módulo de configuração que conta com uma interface gráfica, o *scconfig*. Através dele, é possível ajustar configurações e parâmetros de todos os outros módulos do *software* de forma fácil e intuitiva.

Os módulos *scautopick* e *scautoloc* possuem seus principais parâmetros acessíveis ao usuário através do *scconfig*, sob o caminho *Modules* ⇒ *Processing* ⇒ *scautopick/scautoloc*, como mostra a [Figura 3.3](#).

The screenshot shows the configuration interface for the *scautoloc* module. The sidebar on the left lists various system components, with *Processing* expanded to show *scautoloc*. The main panel displays the following configuration parameters:

Parameter	Value	Description
global		
scautoloc		
autoloc		
maxRMS [s]	3.5	Max. permissible RMS for a location to be reported.
maxResidual [s]	7.0	Max. individual residual (unweighted) for a pick to be ...
minPhaseCount	6	Minimum number of phases.
maxDepth [km]	1000	Maximum permissible depth for a location to be reported.
maxSGAP [deg]	360	Max. secondary azimuth gap for an origin to be reported ...
maxStationDistance [deg]	180	Stations outside the maximum distance range are ignored.
minStaCountIgnorePKP	30	If the station count for stations at < 105 degrees ...
cleanupInterval [s]	3600	Clean-up interval for removing old/unused objects.

Figura 3.3: Exemplo da interface gráfica do módulo de configuração *scconfig*, com parte das configurações dos parâmetros do módulo *scautoloc*.

As configurações contidas nessas partes do *scconfig* são aplicadas a todas as estações sísmográficas do banco, afetando todos os processamentos realizados por ambos os módulos. Também pode-se optar por modificar esses parâmetros diretamente nos arquivos de configuração de cada módulo (contidos em “/seiscomp/etc/”).

Complementarmente, existem os “bindings” das estações, que contêm informações e parâmetros específicos para cada estação contida no banco, tais como os códigos dos canais utilizados (e.g., HH, BH, EH). Através dos *bindings*, realizam-se ajustes individuais nos parâmetros globais visando, por exemplo, diminuir a quantidade de *picks* gerados por uma estação ruidosa, de forma pontual. Os parâmetros dos *bindings* podem ser modificados pelo *scconfig* ou diretamente pelos arquivos individuais de cada estação, localizados no diretório “seiscomp/etc/key/”.

3.3 Parâmetros e fluxos *hard-coded*

Como mostrado anteriormente, o módulo de configuração *scautoloc* fornece ao usuário a possibilidade de modificar diversos parâmetros que são utilizados ao longo do processo de localização de eventos sísmicos. Entretanto, muitos parâmetros e fluxos de processamento não são acessíveis ao usuário final, sendo modificáveis apenas através do código-fonte do *SeisComP*. Muitos desses parâmetros e fluxos são fundamentais no processo de localização de eventos regionais, sendo que, da forma que estão definidos no código original, impedem que o *software* tenha sucesso em localizar esse tipo de evento sísmico.

Além disso, constatou-se que alguns parâmetros editáveis pelo usuário são sobrescritos ou modificados ao longo do código-fonte, fazendo com que as configurações escolhidas não sejam de fato utilizadas durante os processos. A [Figura 3.4](#) apresenta exemplos de recortes do código-fonte do módulo *scautoloc* nos quais parâmetros definidos pelo usuário são modificados, além de serem realizadas normalizações arbitrárias em informações relacionadas às leituras de chegadas de ondas. Um exemplo de modificação arbitrária no código-fonte de um parâmetro definido pelo usuário é o resíduo temporal mínimo permitido para que uma leitura de chegada de onda seja utilizada, o qual é ajustado pelo usuário através do módulo *scconfig*, mas é multiplicado por números arbitrários em diferentes partes do fluxo de processamento.

Dentre os procedimentos mais importantes para a localização automática de eventos sísmicos, destaca-se a pontuação atribuída a cada origem nucleada. Essa pontuação (*originScore*) visa avaliar se os principais parâmetros da origem (resíduos temporais, amplitudes do sinal e distâncias epicentrais) são coerentes e esperados para um evento sísmico real. Os cálculos das pontuações individuais de cada parâmetro, bem como a pontuação mínima para uma origem ser aceita (*minScore*), são arbitrários e baseados em eventos de escalas globais. Visando uma adequação para a realidade da RSBR, também torna-se necessário realizar ajustes nos cálculos das pontuações, além do valor mínimo para uma origem ser aceita pelo *software*.

```

double minResidual = -minFactor*_config.maxResidualUse;
double maxResidual = maxFactor*_config.maxResidualUse;
(a)

if ( _config.aggressivePKP && is_PKP_arrival(arr) ) {
    minResidual *= 2;
    maxResidual *= 2;
}

if ( is_P_arrival(arr) ) {
    // Autoloc 2 hack for regional phases, to allow use of Pg (sometimes even S)
    // as Pn by increasing maxResidual. Which --in principle-- is bad but may be
    // better than leaving those phases out completely.
    double regionalWeight = 1.+0.7*exp(-arr.distance*arr.distance/50.);
    maxResidual *= regionalWeight;
}

// Weight residuals at regional distances "a bit" lower
// This is quite hackish!
(b)
x = 1 + 0.6*exp(-0.003*delta*delta) + 0.5*exp(-0.03*(15-delta)*(15-delta));

// test minimum number of picks
(c)
if (origin->arrivals.size() < 6) // TODO: make this limit configurable
    continue;

double normalizationAmplitude = 2000.; // arbitrary choice
(d)
const_cast<Pick*>(pick)->normamp = pick->amp/normalizationAmplitude;

double normalizedResidual = arr.residual/_config.maxResidualUse;
// add penalty for positive residuals // TODO: make it configurable?
if (normalizedResidual > 0) normalizedResidual *= 1.5;

```

Figura 3.4: Recortes do código-fonte do módulo *scautoloc* nos quais há modificações internas dos parâmetros definidos pelo usuário. (a) Modificação do máximo resíduo de tempo permitido para uma leitura de chegada de onda; (b) Peso arbitrário para leituras de chegadas de onda; (c) Número mínimo de leituras para definir uma origem sendo um valor fixo; (d) e (e) Normalizações e penalidades arbitrárias de amplitudes e resíduos.

Visando definir um fluxo de processamento e parâmetros que favoreçam a detecção de eventos sísmicos em território brasileiro, todos os processos realizados pelo módulo *scautoloc* foram estudados, permitindo a identificação de pontos de melhorias e até mesmo mudanças significativas. Foram definidas novas metodologias para avaliação da qualidade das origens nucleadas, novos parâmetros que devem ser seguidos durante o processo de nucleação, dentre outras implementações descritas na seção [Métodos e procedimentos](#).

3.4 Modelos de velocidades

Durante o processo de localização de eventos sísmicos, o *SeisComP* utiliza uma tabela de tempos de percurso para realizar uma busca em uma malha de pontos (possíveis hipocentros). Tradicionalmente, o *SeisComP* é fornecido com uma tabela de tempos de percurso obtida através do modelo IASP91 (Kennett and Engdahl, 1991). Nas rotinas de monitoramento da RSBR, um segundo modelo tem sido utilizado para relocalizar eventos regionais, sendo esse o modelo NewBR (Assumpção et al., 2010).

3.4.1 IASP91

O modelo de velocidades IASP91 (Kennett and Engdahl, 1991) foi elaborado pela *International Association of Seismology and Physics of the Earth's Interior* (IASPEI), sendo um modelo 1D que sumariza o comportamento global dos tempos de percurso das ondas P e S para um raio terrestre médio de 6371 km. A partir do modelo de velocidades IASP91, foram propostas tabelas de tempo de percurso que visaram atualizar os dados publicados por Jeffreys and Bullen (1940), amplamente utilizados até então.

No modelo IASP91, a crosta terrestre possui uma espessura de 35 km, com uma descontinuidade em 20 km de profundidade. As velocidades em cada camada são representadas por um modelo paramétrico, sendo que as principais descontinuidades do manto foram definidas em 410 km e 660 km.

O modelo IASP91 é o padrão utilizado pelo *SeisComP* para a localização de sismos. Entretanto, por ser um modelo global, utilizá-lo para realizar localizações de eventos sísmicos regionais pode resultar em resíduos temporais maiores, especialmente quando combinam-se estações em distâncias locais, regionais e telessísmicas. Para evitar a necessidade de se aceitar um maior resíduo temporal por leitura e, conseqüentemente, um maior RMS dos resíduos de tempo do evento, é conveniente utilizar um modelo regional de velocidades para garantir localizações mais precisas e com menores resíduos temporais.

3.4.2 NewBR

O modelo de velocidades NewBR foi proposto por Assumpção et al. (2010) para ser um modelo que leve em consideração a predominância de uma litosfera antiga e cratônica na maior parte do Brasil. Foram utilizados 15 eventos em território brasileiro bem conhecidos, com magnitudes $\geq M4,0$ e registrados em distâncias regionais, para que seus tempos de percurso fossem corrigidos utilizando-se suas respectivas profundidades. O modelo NewBR esférico possui uma espessura crustal de 42 km.

Para a determinação do modelo NewBR, adotou-se o modelo proposto por Herrin (1968) como modelo inicial, pois, segundo Ardito (2009), tal modelo possui tempos de percurso mais próximos aos tempos observados no Brasil. Segundo o autor, as velocidades do modelo NewBR foram ajustadas através de modelagem direta por “tentativa e erro”.

3.5 Dados paramétricos

Durante o desenvolvimento desse trabalho, os dados paramétricos dos eventos sísmicos foram amplamente utilizados para gerar as análises apresentadas. Para esses metadados, o *SeisComp* implementa uma estrutura de representação denominada *Sc3ML*, compatível com o formato *QuakeML* (Schorlemmer et al., 2011), como mostra o exemplo da Figura 3.5.

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <seiscomp xmlns="http://geofon.gfz-potsdam.de/ns/seiscomp3-schema/0.11" version="0.11">
3  <EventParameters>
23 <pick publicID="20190128.135521.56-AIC-NB.NBPB..HHZ">
      [...]
118 <amplitude publicID="20190128.135521.56-AIC-NB.NBPB..HHZ.snr">
      [...]
197 <origin publicID="origin/20220729165733.122099.15657">
198 <time>
199 <value>2019-01-28T14:05:22.432703Z</value>
200 <uncertainty>0.600353539</uncertainty>
201 </time>
202 <latitude>
203 <value>-16.63494873</value>
204 <uncertainty>3.464893818</uncertainty>
205 </latitude>
206 <longitude>
210 <depth>
214 <methodID>LOCSAT</methodID>
215 <earthModelID>newbr3</earthModelID>
216 <quality>
217 <associatedPhaseCount>7</associatedPhaseCount>
218 <usedPhaseCount>6</usedPhaseCount>
219 <associatedStationCount>7</associatedStationCount>
220 <usedStationCount>6</usedStationCount>
221 <standardError>0.2539726416</standardError>
222 <azimuthalGap>183.9771471</azimuthalGap>
223 <maximumDistance>2.752312899</maximumDistance>
224 <minimumDistance>0.08704286814</minimumDistance>
225 <medianDistance>2.379316688</medianDistance>
226 </quality>
227 <evaluationMode>automatic</evaluationMode>
228 <creationInfo>
233 <arrival>
      [...]

```

Figura 3.5: Exemplo da organização hierárquica dos dados paramétricos de um evento sísmico no formato *Sc3ML*. “[...]” indica a existência de conteúdo não mostrado na figura.

A criação do formato *QuakeML* visou padronizar os metadados dos eventos sísmicos e seus compartilhamentos. Existem diversas aplicações que utilizam esse formato, dentre elas, *softwares* de sismologia e serviços de catálogos de grandes centros sismológicos no mundo. Os dados paramétricos nesses formatos são facilmente legíveis tanto por humanos quanto por algoritmos, pois seus elementos são padronizados e bem definidos.

Nos formatos *QuakeML* e *Sc3ML*, o dado paramétrico de um evento sísmico descreve seus elementos de uma maneira hierárquica, por exemplo, associando *picks* às origens que, por sua vez, estão associadas ao evento. Dessa forma, eventos e origens são elementos relacionados entre si através de uma “árvore de dados”. Uma de suas características mais importantes é o armazenamento do histórico completo do evento, desde a nucleação de sua primeira origem. Qualquer parâmetro associado ao evento em algum momento de sua existência é mantido em seu histórico e acessível através de seu arquivo paramétrico.

4. Métodos e procedimentos

Para o desenvolvimento desse trabalho, foram utilizados os dados (registros contínuos em forma de ondas) contidos no banco de dados do Centro de Sismologia do IAG-USP. Apenas os dados registrados por estações contidas no inventário do *SeisComp* do Centro de Sismologia foram considerados, ou seja, nenhuma nova estação foi adicionada.

A Figura 4.1 apresenta um fluxograma indicando os dados utilizados durante cada etapa de otimizações dos parâmetros dos módulos *scautopick* e *scautoloc*. Durante todas as etapas, foram realizados diversas rodadas de processamento buscando verificar a contribuição de cada parâmetro e processo na detecção e localização de sismos regionais brasileiros. Em paralelo aos procedimentos de otimização dos módulos de processamento, também foi desenvolvido um novo modelo de velocidades, como apresentado adiante.

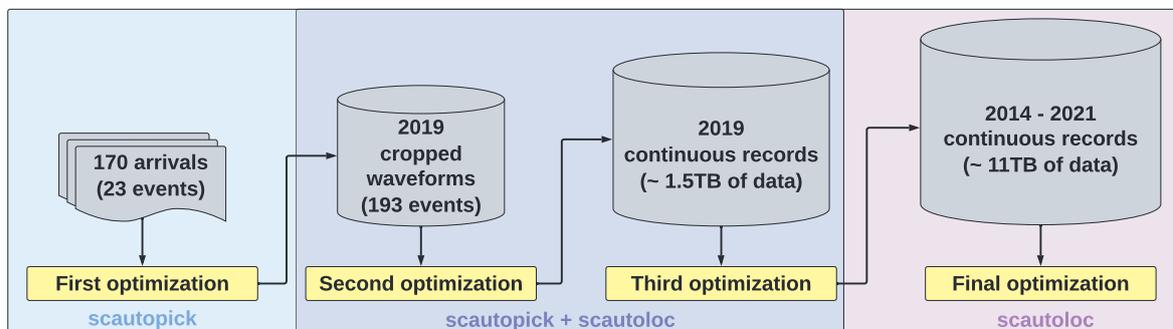


Figura 4.1: Fluxograma indicando os dados utilizados durante cada etapa de otimizações dos parâmetros dos módulos *scautopick* e *scautoloc*.

Inicialmente, foram selecionados 170 *arrivals* de 23 eventos regionais brasileiros contidos no catálogo da RSBR para a etapa inicial de otimização dos dois principais parâmetros do módulo *scautopick*: o filtro de frequências e as janelas de tempo do algoritmo STA/LTA. Buscou-se abranger a maior parte dos *ranges* de magnitudes contidas no catálogo, elegendo os eventos com registros sismográficos de boa qualidade. Em seguida, foram desenvolvidas diversas rotinas em *Python* para realizar as otimizações, favorecendo a detecção das chegadas das ondas P dos eventos sísmicos, como descrito na Subseção 4.2. Outros parâmetros do módulo *scautopick* também passaram pelo etapa inicial de otimização como, por exemplo, o valor da razão STA/LTA necessário para que o *picker* realize uma leitura de chegada de onda.

Posteriormente, os parâmetros do módulo *scautopick* passaram por um segundo processo de otimização, utilizando-se recortes de formas de onda contendo os 193 eventos sísmicos ocorridos no ano de 2019. Nesse segundo momento, também foram iniciados os testes para aprimorar os parâmetros do módulo *scautoloc*, incluindo seus processos *hard-coded*.

Em uma terceira etapa de otimização, os parâmetros de detecção e localização foram refinados através de diversas rodadas de testes, dessa vez utilizando os dados contínuos registrados durante todo o ano de 2019. Dessa forma, buscou-se verificar o comportamento do *software* modificado no contexto de processamento com dados contínuos, e não apenas

recortes de dados. Novamente, foram desenvolvidos algoritmos em *Python* e *Bash* para executar os módulos *scautopick* e *scautoloc* de uma maneira eficiente e sistemática. Os dados paramétricos dos eventos localizados automaticamente, em formato *Sc3ML*, foram constantemente avaliados através de gráficos e análises estatísticas.

Por fim, todo o banco de dados no período de 2014 a 2021 foi processado, de forma a verificar se os parâmetros refinados com os dados do ano de 2019 também geravam os resultados esperados em todo o período, além de realizar ajustes finais no código-fonte do módulo *scautoloc*. Nessa etapa, foram processados mais de 11TB de dados contínuos, registrados entre 2014 e 2021 pelas estações da RSBR e estações temporárias.

Para fins de comparação, o *SeisComP* original (*i.e.*, sem nenhuma modificação em seus parâmetros de detecção e localização) foi utilizado para gerar um catálogo de eventos durante o mesmo período. Ao final, foram obtidos três catálogos de eventos sísmicos para serem comparados entre si: (i) catálogo oficial da RSBR, (ii) catálogo do *SeisComP* modificado e (iii) catálogo do *SeisComP* original. Cabe ressaltar que a versão do *SeisComP* utilizada no IAG-USP para realizar os processamentos para a RSBR é muito próxima da versão original, ou seja, apresenta poucas modificações em seus parâmetros de detecção e localização e nenhuma modificação em seu código-fonte.

Os trabalhos realizados podem ser condensados em três principais tópicos, sendo eles: o desenvolvimento de um novo modelo de velocidades, a otimização do módulo *scautopick* e a otimização do módulo *scautoloc*.

4.1 Desenvolvimento de um modelo de velocidades - BRA23

Para minimizar os resíduos temporais das leituras de chegadas de onda P nas estações, é de extrema importância o emprego de um modelo de velocidades adequado. Para isso, o modelo regional NewBR foi utilizado como base para a geração de um novo modelo 1D de velocidades para o território brasileiro, denominado BRA23. Também foram considerados dois eventos sísmicos mais recentes para a elaboração desse modelo.

Os dados utilizados são as leituras de tempo de percurso para estações regionais, combinadas com os hipocentros dos eventos, ambos determinados de forma acurada. Os 17 eventos regionais utilizados para derivar o modelo BRA23 foram:

- Pacajus/CE (1980);
- Porto dos Gaúchos/MT (1998 e 2005);
- Codajás/AM (1983);
- Nova Ponte/MG (1998);
- João Câmara/RN (2x 1986 e 1989);
- São Caetano/PE (2006);
- Sobral/CE (1988 e 2008);
- Caraíbas/MG (2007);
- Palhano/CE (1988 e 1989);
- Mara Rosa/GO (2010); e
- Taperuaba/CE (1991);
- Montes Claros/MG (2012).

As leituras de chegadas de ondas utilizadas para gerar o modelo BRA23 foram obtidas e revisadas pelo *International Seismological Centre* (ISC), sendo fornecidas através de comunicação pessoal com o Prof. Dr. Marcelo Assumpção.

Os tempos de percurso para cada estação e cada evento são mostrados na Figura 4.2, sobrepostos aos tempos preditos pelo modelo NewBR em cada caso. É possível observar que o modelo NewBR tende a superestimar os tempos de percurso, especialmente para distâncias intermediárias (até aproximadamente 2200 km). Com isso, os resíduos temporais calculados ($t_{model} - t_{obs}$) apresentam uma tendência de serem positivos, como mostrado na Figura 4.3.

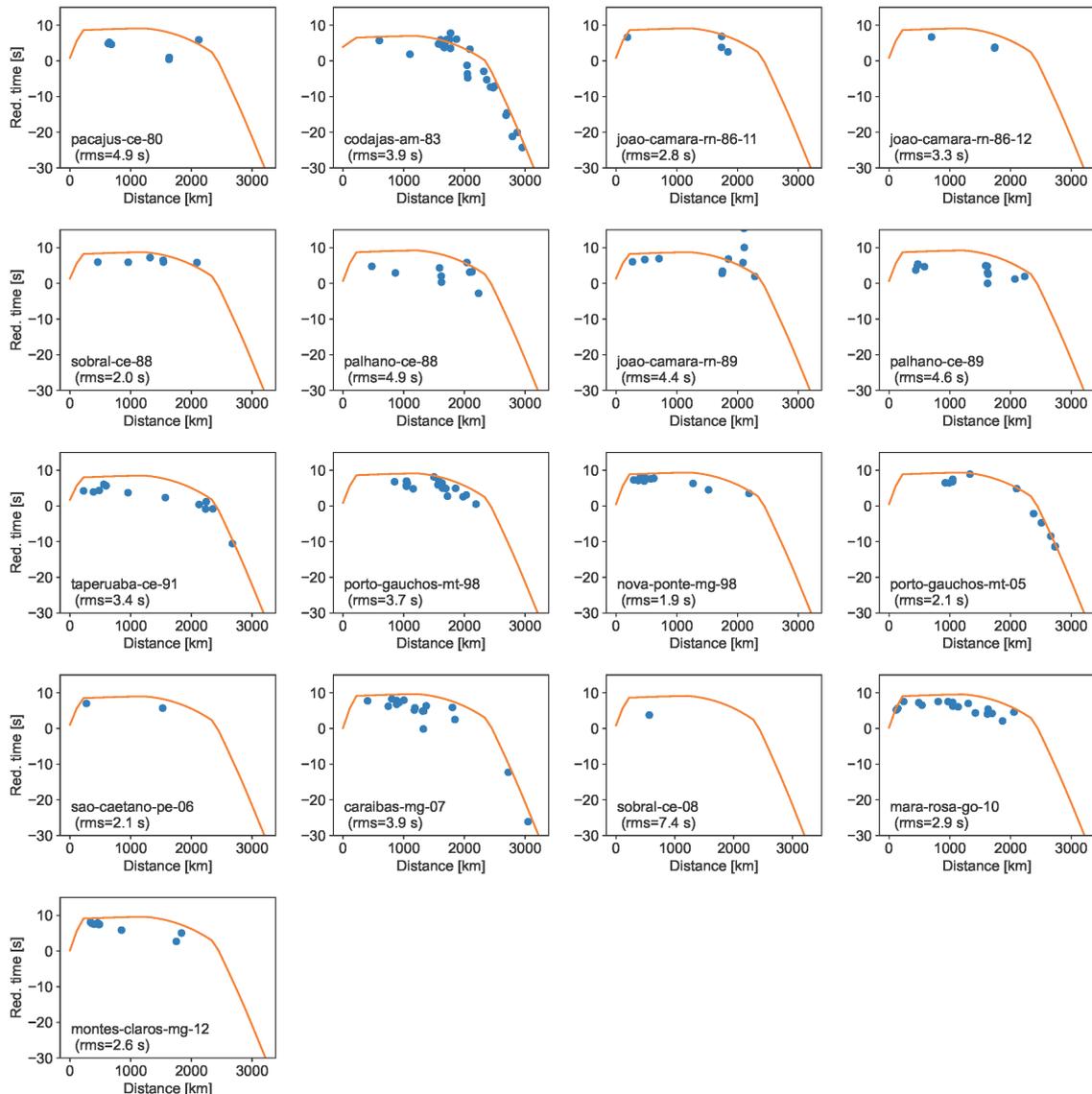


Figura 4.2: Tempos de percurso reduzidos para a velocidade de 8,2 km/s, para cada um dos 17 eventos utilizados para derivar o modelo BRA23. As curvas laranjas sobrepostas aos dados correspondem aos tempos de percurso estimados utilizando-se o modelo NewBR e considerando a profundidade de cada evento. O RMS estimado em cada caso é indicado abaixo do nome/ano do evento.

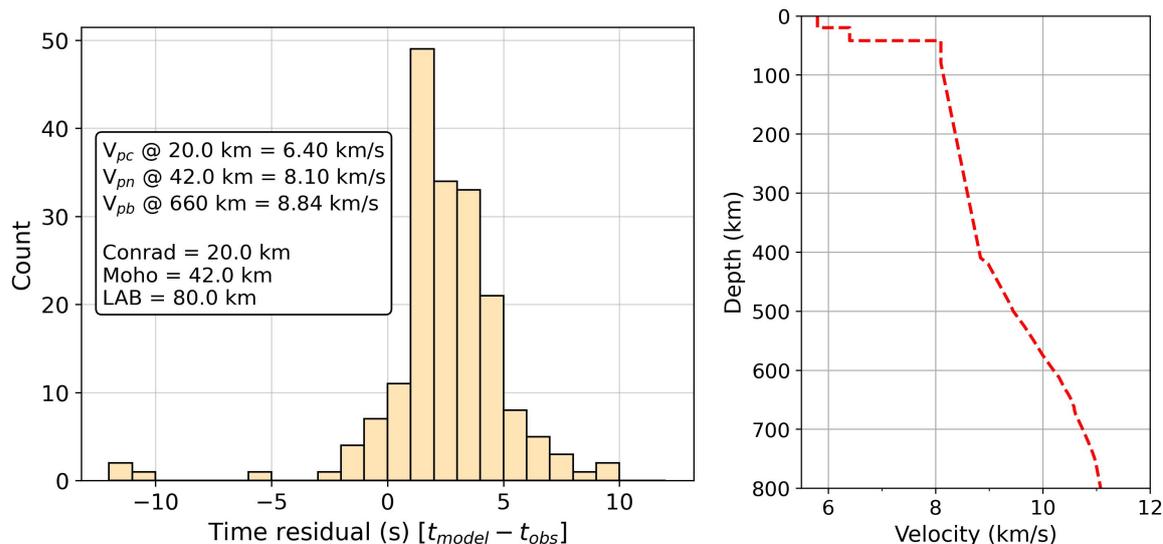


Figura 4.3: Histogramas de resíduos (esquerda) das leituras de chegadas de ondas utilizando-se o modelo NewBR (direita), cujos principais parâmetros são indicados. “ V_{pc} ”: velocidade da onda P na crosta inferior, “ V_{pn} ”: velocidade da onda Pn, “ V_{pb} ”: velocidade da onda P na base do manto superior. Nesse modelo, a descontinuidade de Conrad, a descontinuidade de Moho e a LAB (*lithosphere-asthenosphere boundary*) estão a 20 km, 42 km e 80 km de profundidade, respectivamente.

O novo modelo de velocidades de onda P foi obtido a partir da parametrização de um modelo composto por duas camadas de velocidades uniformes na crosta, um manto parametrizado por uma camada de velocidade uniforme e um gradiente de velocidade que se estende até a profundidade de 660 km. Abaixo de 660 km, adotou-se o modelo AK135-F (Kennett et al., 1995; Montagner and Kennett, 1996).

Os tempos de percurso de cada chegada e de cada evento foram computados de forma individual, utilizando-se o pacote *TauP* (Crotwell et al., 1999) através da interface do pacote *ObsPy* (Beyreuther et al., 2010), de forma a não aplicar nenhuma correção aos dados a serem ajustados. Para cada par evento-estação, foram calculados dois tempos de percurso (t_1 e t_2). No primeiro, considerou-se que o modelo 1D possui a mesma espessura crustal que a da região do hipocentro e, no segundo, a espessura da crosta do modelo foi definida como sendo a mesma da região da estação, como mostra a Figura 4.4. O tempo final para o par evento-estação foi obtido através da média simples dos dois casos. As espessuras crustais consideradas foram propostas por Rivadeneyra et al. (2019) e truncadas em uma casa decimal, de forma a reduzir o número de combinações possíveis (modelos testados) e viabilizar uma inversão.

Para se obter o modelo BRA23, as principais constantes contidas na parametrização proposta pelo modelo NewBR foram otimizadas, visando minimizar a soma de todos os resíduos calculados. As otimizações foram realizadas de forma análoga ao processo realizado na elaboração do modelo NewBR, mas de forma automatizada, através do desenvolvimento de algoritmos de processamento. As constantes otimizadas, bem como suas descrições e intervalos de busca adotados, são apresentadas na Tabela 4.1.

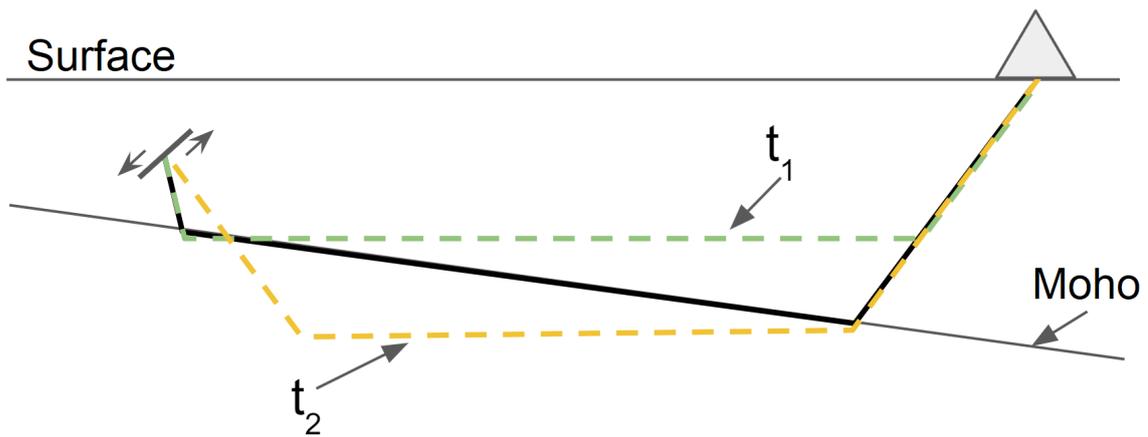


Figura 4.4: Esquemática do cálculo do tempo de trânsito da frente de onda entre o hipocentro do evento e a estação. Como a espessura da crosta não é constante, adotou-se uma média entre os tempos considerando-se as espessuras na região do evento e da estação, simulados por dois modelos 1D distintos.

De forma a obter uma solução estável, a otimização dos parâmetros foi realizada em duas etapas: uma otimização global, utilizando todos os dados e o método de Evolução Diferencial implementado pelo pacote *SciPy* (Virtanen et al., 2020), seguida por uma otimização local pelo método *Nelder-Mead* também implementado pelo pacote *SciPy*. Antes da otimização local, foram removidos dos dados os *outliers* (pontos abaixo e acima de 1,5 a distância interquartil subtraídas e somadas ao primeiro e terceiro quartis). Além disso, a solução da otimização global foi considerada como ponto de partida.

Tabela 4.1: Parâmetros otimizados durante o procedimento de obtenção do modelo BRA23. A otimização dos parâmetros foi realizada considerando os intervalos especificados na coluna “Intervalo de busca”.

Parâmetro	Intervalo de busca	Descrição
V_{pc}	6,4 km/s - 7,5 km/s	Velocidade da onda P na crosta inferior, abaixo da descontinuidade de Conrad.
V_{pn}	7,5 km/s - 8,6 km/s	Velocidade da onda P no manto superior, dentro da LAB.
V_{pb}	8,5 km/s - 10,4 km/s	Velocidade da onda P na base do manto (profundidade de 660 km).
Conrad	10 km - 20 km	Profundidade da descontinuidade de Conrad.
LAB	60 km - 250 km	Profundidade da LAB (ou da camada de velocidade constante no topo do manto superior).

4.2 Otimização do módulo *scautopick*

A primeira etapa de aprimoramento do *SeisComP* envolveu uma avaliação dos parâmetros padrões do módulo *scautopick*. Por ser responsável pela leitura das chegadas de ondas nos sismogramas, o *scautopick* deve ser capaz de identificar anomalias de amplitudes que possam ser oriundas de eventos sísmicos, registrando-as da maneira mais acurada possível, ou seja, na primeira chegada das ondas P.

Inevitavelmente, leituras incorretas associadas a ruídos no sinal são registradas e passadas adiante para os próximos módulos de processamento. A otimização ideal busca aumentar o número de leituras corretas ao mesmo tempo em que diminui as leituras incorretas e, para isso, diversos parâmetros são utilizados a fim de realçar os eventos sísmicos nos sismogramas e atenuar ruídos não relacionados.

Para otimizar as leituras de chegadas de ondas realizadas pelo módulo *scautopick*, os principais parâmetros que foram avaliados e otimizados são descritos a seguir, utilizando-se a sintaxe¹² do *SeisComP*.

- **RMHP(*timespan*)**: média móvel equivalente a um filtro passa-alta, utilizada para remover o *offset* dos dados;
- **ITAPER(*timespan*)**: *taper*, uma janela de suavização unilateral do tipo cosseno, visando diminuir gradualmente as amplitudes nas bordas do dado para evitar descontinuidades nos cálculos;
- **BW(*or, hp, lp*)**: filtro de frequências *butterworth* passa-banda aplicado aos dados contínuos, sendo possível modificar sua ordem (*or*) e seus limites inferior (*hp*) e superior (*lp*);
- **STALTA(*sta, lta*)**: parâmetros para detecção de sinais transientes, compostos pelas janelas de tempo de STA (*short-time window*) e de LTA (*long-time window*);
- **picker**: algoritmo utilizado para realizar as leituras de chegadas de ondas;
- **triggerOn / triggerOff**: limites de ativação e desativação utilizados pelo *picker*;
- **deadTime**: período, em segundos, no qual não haverá novas leituras de chegadas de onda após uma detecção; e
- **amplMaxTimeWindow**: janela de tempo, em segundos, utilizada para calcular a razão sinal-ruído (SNR) de uma leitura.

Esses e outros parâmetros podem ser definidos para todas as estações ou de forma individual para cada estação através do módulo de configurações *scconfig*, como apresentado na [Subseção 3.2](#).

¹<https://www.seiscomp.de/doc/base/filter-grammar.html>

²<https://www.seiscomp.de/doc/base/concepts/configuration.html>

4.2.1 Filtro de frequências

O módulo *scautoloc* utiliza apenas as chegadas das ondas P dos eventos sísmicos para realizar a localização de origens, portanto, deve-se buscar um filtro de frequências que realce essa fase dos sismos regionais brasileiros. Para isso, foi realizada uma análise inicial de filtros passa-banda sobre dados em forma de ondas de 170 *arrivals* de 23 eventos sísmicos, abrangendo a maior parte da variedade de magnitudes dos sismos regionais brasileiros. Essas leituras de chegadas de ondas foram refinadas individualmente, de forma manual, buscando a maior acurácia na marcação. Utilizando-se um algoritmo desenvolvido em *Python*, foram realizadas milhares de combinações de limites de filtros passa-banda, buscando aquelas que fornecessem a maior razão sinal-ruído de cada evento.

A [Figura 4.5](#) apresenta um exemplo de análise de filtros de frequência, realizada sobre os dados registrados pela estação NB.NBMA referentes ao evento “*usp2021load*” do catálogo da RSBR, ocorrido próximo à cidade de Quixeramobim/CE e com magnitude M1,5.

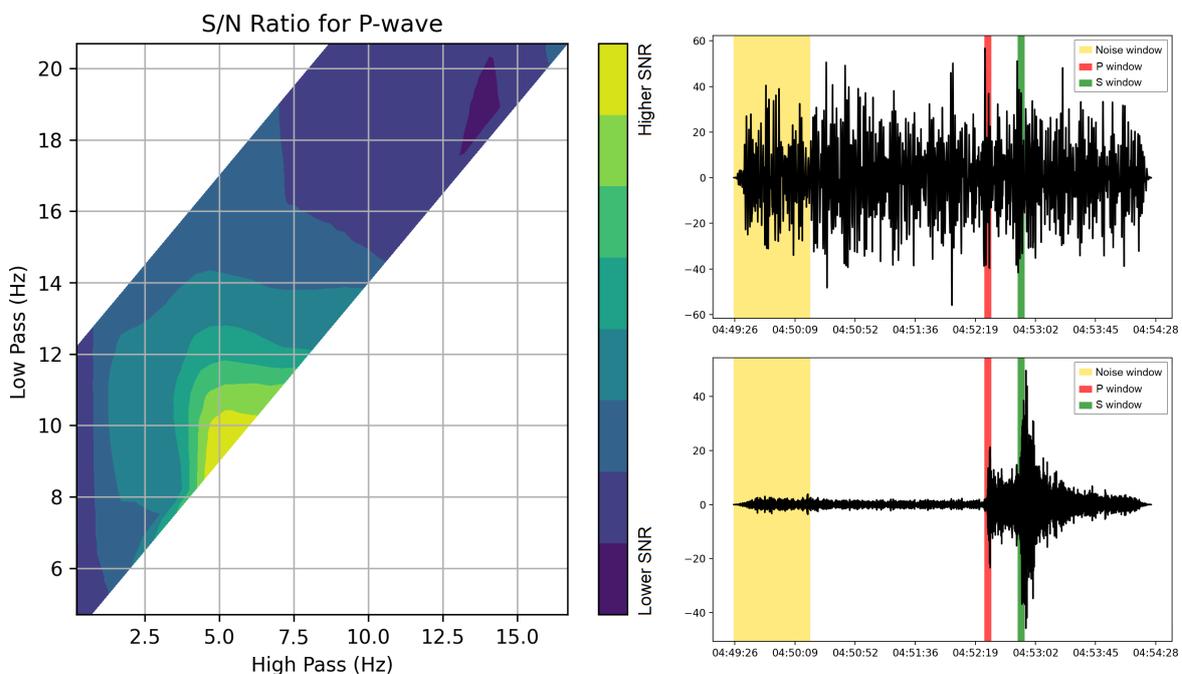


Figura 4.5: Esq.: Análise das combinações de filtros de frequência para um evento sísmico de magnitude M1,5 registrado pela estação NB.NBMA. Dir.: Exemplo de realce das ondas sísmicas provenientes do mesmo evento, através do filtro de frequências adequado. O sismograma superior mostra as formas de onda do evento utilizando-se o filtro padrão do *SeisComP*, enquanto o sismograma inferior apresenta os dados filtrados com os limites sugeridos para esse evento após as análises.

No gráfico à esquerda da [Figura 4.5](#), pode-se notar que o evento em questão apresentou uma maior razão sinal-ruído da chegada das ondas P na estação NB.NBMA quando aplicado o filtro passa-banda com limites de 5,2 Hz e 9,5 Hz, aproximadamente. Os sismogramas à direita mostram as formas de onda desse evento com o filtro padrão do *SeisComP* (acima) e o filtro passa-banda sugerido para esse evento (abaixo), baseado nos resultados na análise.

4.2.2 Parâmetros de STA/LTA

Posteriormente à aplicação do filtro de frequências, o *scautopick* aciona um algoritmo de STA/LTA para realizar a leitura das chegadas de ondas dos eventos sísmicos. Por padrão, as janelas de tempo utilizadas são 2 e 80 segundos para STA e LTA, respectivamente.

Para encontrar os melhores parâmetros de STA e LTA, adotou-se o mesmo procedimento utilizado para definição do filtro de frequências. Dessa forma, foi realizada uma análise de diversas combinações de janelas de tempo, de forma a buscar a qual maximiza a razão STA/LTA da chegada da onda P, após a aplicação do filtro de frequências sugerido. A Figura 4.6 apresenta um exemplo dessa análise, também baseando-se nos dados referentes ao evento “*usp2021load*” e registrados pela estação NB.NBMA.

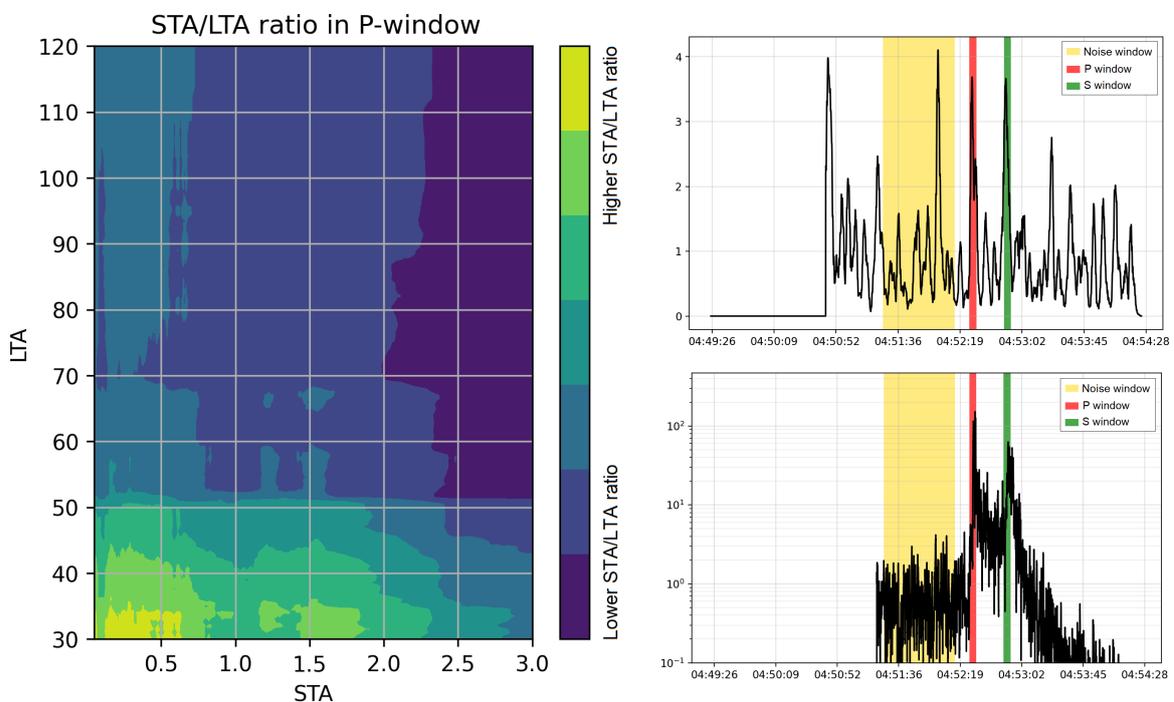


Figura 4.6: Esq.: Análise das combinações de janelas de tempo do algoritmo STA/LTA para um evento sísmico de magnitude M1,5 registrado pela estação NB.NBMA. Dir.: Exemplo de realce das ondas sísmicas provenientes do mesmo evento, através do filtro de frequências e janelas de STA/LTA adequados. O sismograma superior mostra as formas de onda do evento utilizando-se os parâmetros padrões do *SeisComP*, enquanto o sismograma inferior (em escala logarítmica) apresenta os dados com aplicação dos parâmetros sugeridos para esse evento após as análises.

No gráfico à esquerda da Figura 4.6, pode-se notar que o evento em questão apresentou uma maior razão sinal-ruído da chegada das ondas P na estação NB.NBMA ao se utilizar uma janela de STA entre 0,1 e 0,4 segundos e uma janela de LTA entre 30 e 35 segundos. À direita, o sismograma superior mostra as formas de onda com os parâmetros padrões do *SeisComP* (0,7 Hz - 2 Hz; 2 s - 80 s), enquanto o sismograma inferior apresenta os dados utilizando-se os parâmetros sugeridos para esse evento (5,2 Hz - 9,5 Hz; 0,25 s - 33 s).

4.2.3 *Picker*

Na sismologia, o método mais utilizado para leitura de chegadas de ondas P em um sismograma é o que utiliza a razão STA/LTA, como proposto por Allen (1978). Considerando o fato de que essa razão proporciona apenas percepções de variações de amplitudes do sinal, vários estudos já foram realizados visando propor *pickers* mais eficientes, ou seja, que possam avaliar outras informações da forma de onda como, por exemplo, a frequência. Devido ao fato de que a razão STA/LTA relaciona-se apenas com as amplitudes do sinal, não é possível para o algoritmo notar a diferença entre pulsos relacionados a ruído e eventos sísmicos reais, gerando falsos positivos em estações ruidosas.

O *Akaike Information Criterion* (AIC) é um método proposto por Akaike (1971) que, à época, estava estudando a qualidade de ajuste de um modelo estatístico, tentando encontrar a menor ordem que melhor se ajuste aos dados observados. Um dos benefícios colaterais observados é que esse critério é capaz de separar uma série temporal em dois segmentos com propriedades distintas. A implementação desse algoritmo também apresenta robustez nos resultados mesmo com a presença de ruídos de altas amplitudes (St-Onge, 2011).

O módulo *scautopick* conta com um algoritmo de AIC implementado, apesar de não ser utilizado por padrão. Quando ativado, o método STA/LTA continua sendo utilizado como detector, fornecendo uma janela de tempo (a qual contém a detecção de um sinal transiente) para o algoritmo de AIC atuar como *picker*, refinando a leitura inicial. Quando o AIC é desativado, o método STA/LTA é utilizado como detector e *picker* simultaneamente, sendo responsável por gerar os *picks* utilizados para as localizações dos eventos.

Vale salientar que o algoritmo de AIC não pode ser utilizado como detector, pois não é possível aplicá-lo sobre os dados contínuos. Dessa forma, é necessário utilizar o algoritmo STA/LTA para realizar uma detecção inicial para que, em seguida, o AIC atue em uma janela de tempo definida pelo usuário, a qual contém a detecção a ser refinada.

No presente trabalho, o algoritmo AIC foi utilizado como *picker*, o que pode ser configurado através dos *bindings* das estações. Os parâmetros de configuração dos filtros utilizados por esse *picker* seguiram os mesmos obtidos para o STA/LTA. Maiores detalhes sobre todas as modificações propostas e os parâmetros utilizados são apresentados na seção [Resultados](#).

4.3 Otimização do módulo *scautoloc*

A seguir, são detalhadas as melhorias realizadas no módulo *scautoloc*, incluindo as modificações realizadas através do módulo de configurações (*scconfig*) e processos ou parâmetros *hard-coded*, ou seja, contidos no código-fonte do *software*.

Como mostrado na [Subseção 3.3](#), alguns parâmetros editáveis pelo usuário através do módulo *scconfig* são arbitrariamente alterados no código-fonte original do *SeisComP*. Para remover tais arbitrariedades, foram realizadas modificações diretamente no código-fonte e, portanto, os parâmetros acessíveis ao usuário final são respeitados.

Dessa forma, os principais parâmetros de localização disponíveis no módulo *scconfig* e avaliados nesse trabalho foram:

- **defaultDepth**: profundidade padrão dos eventos localizados, utilizada caso o RMS da origem resultante seja menor do que o RMS da origem com a profundidade estimada. Comumente empregada em eventos superficiais sem resolução de profundidade;
- **minimumDepth**: profundidade mínima que o localizador pode associar a uma origem;
- **maxDepth**: profundidade máxima que o localizador pode associar a uma origem;
- **maxRMS**: máximo RMS (*root mean square*) permitido para uma origem ser aceita;
- **maxResidual**: máximo resíduo temporal de um *arrival* permitido para que o mesmo contribua na localização de uma origem;
- **minPhaseCount**: número mínimo de fases contribuindo para a localização (*arrivals*) que uma origem deve ter para ser aceita;
- **maxStationDistance**: distância máxima permitida para que uma estação contribua com *picks* para uma origem;
- **amplTypeAbs**: tipo de amplitude utilizada pelo módulo *scautoloc* (e.g., mb, MLv); e
- **profile**: modelo de velocidades associado à tabela de tempos de percurso utilizada para localizar eventos.

Durante o processo de otimização do módulo de localização *scautoloc*, também foram realizadas otimizações na malha de pontos utilizada para se nuclear novas origens de eventos sísmicos (descritas a seguir), além do próprio desenvolvimento do modelo de velocidades utilizado pelo *software* ([Subseção 4.1](#)).

4.3.1 Malha de pontos (*grid.conf*)

A malha de pontos utilizada pelo módulo *scautoloc* durante o processo de nucleação de novos eventos pode ser modificada pelo usuário através do arquivo *grid.conf* (*gridfile*) contido nos diretórios do *software*. Como descrito na documentação oficial do *SeisComP* (Helmholtz-Centre Potsdam - GFZ German Research Centre for Geosciences and gempa GmbH, 2008), esse arquivo consiste em uma linha para cada ponto da malha, sendo que cada ponto é definido por 6 colunas:

```
-10.00 105.00 20.0 5.0 180.0 8
```

As colunas são, respectivamente, as coordenadas do ponto (latitude, longitude e profundidade), raio, distância máxima e número mínimo de leituras. A linha de exemplo acima cria um ponto nas coordenadas 10°S/105°E, com profundidade de 20 km. Esse ponto da malha é sensível a eventos sísmicos a uma distância de 5° de seu centro, ao mesmo tempo em que estações com distâncias de até 180° podem contribuir com *picks* para a nucleação de eventos. No mínimo 8 *picks* precisam ser utilizados para criar uma origem válida nessa localização.

Todos os parâmetros da malha de pontos são de extrema importância para otimizar a localização automática de eventos sísmicos regionais. Por exemplo, um evento sísmico de baixa magnitude (*e.g.*, < M3,0) ocorrido na Região Norte do Brasil dificilmente será registrado por uma estação sísmica na Região Sul do país. Dessa forma, convém limitar a distância máxima dos pontos da malha, evitando a nucleação errônea de eventuais ruídos.

O adensamento de pontos na região de interesse permite que eventos sísmicos sejam nucleados com mais facilidade, pois, considerando que os pontos da malha são “chutes iniciais” de um hipocentro, quanto mais pontos, maiores as chances de se encontrar um no qual os resíduos de tempo de percurso sejam próximos à realidade. Da mesma maneira, um adensamento exagerado de pontos pode permitir que ruídos sejam nucleados como eventos sísmicos, pois também aumentam-se as chances de encontrar um ponto no qual os resíduos de tempo dos ruídos sejam compatíveis com os resíduos esperados para um evento real.

Por padrão, o *gridfile* do *SeisComP* possui uma malha de pontos que abrange toda a área do globo, com espaçamentos de aproximadamente 5° e com alguns pontos adicionais abrangendo maiores profundidades em locais nos quais hipocentros profundos ocorrem com maior frequência.

Buscando o aperfeiçoamento da detecção automática de eventos sísmicos regionais no Brasil, a malha de pontos foi reformulada de modo a aumentar a quantidade de pontos nessa região e estabelecer parâmetros, para cada ponto, concordantes com a sismicidade brasileira e com a disposição das estações sismográficas em operação.

A malha proposta é composta por 330 pontos equiespaçados de 2° (majoritariamente em território brasileiro) e 12 pontos cobrindo a região andina. Apesar de os eventos andinos não serem o foco do trabalho, é importante que o *software* considere pontos da malha nessa região, evitando que eventos sísmicos originados nos Andes sejam erroneamente nucleados no Brasil.

Para os pontos da malha dispostos na região brasileira e seus arredores, a quantidade de estações em um raio de 10° a partir de cada ponto da malha foi o critério adotado para definir a distância máxima (d_{max}) em que uma estação pode contribuir com *picks* e o número mínimo de *picks* necessários para nuclear um evento no ponto em questão (n_{picks}). Sendo n_{sta} o número de estações em um raio de 10° centrado no ponto, as regras utilizadas para definir os parâmetros foram:

$$d_{max} = \begin{cases} 15, & \text{se } n_{sta} \leq 5 \\ 10, & \text{se } 5 < n_{sta} \leq 15 \\ 8, & \text{se } 15 < n_{sta} \leq 20 \\ 6, & \text{se } 20 < n_{sta} \leq 30 \\ 5, & \text{se } n_{sta} > 30 \end{cases} \quad n_{picks} = \begin{cases} 5, & \text{se } n_{sta} \leq 10 \\ 6, & \text{se } n_{sta} > 10 \end{cases}$$

A Figura 4.7 apresenta dois mapas da região sul-americana, sendo um com a malha de pontos padrão do *SeisComP* e outro com a malha proposta no presente trabalho.

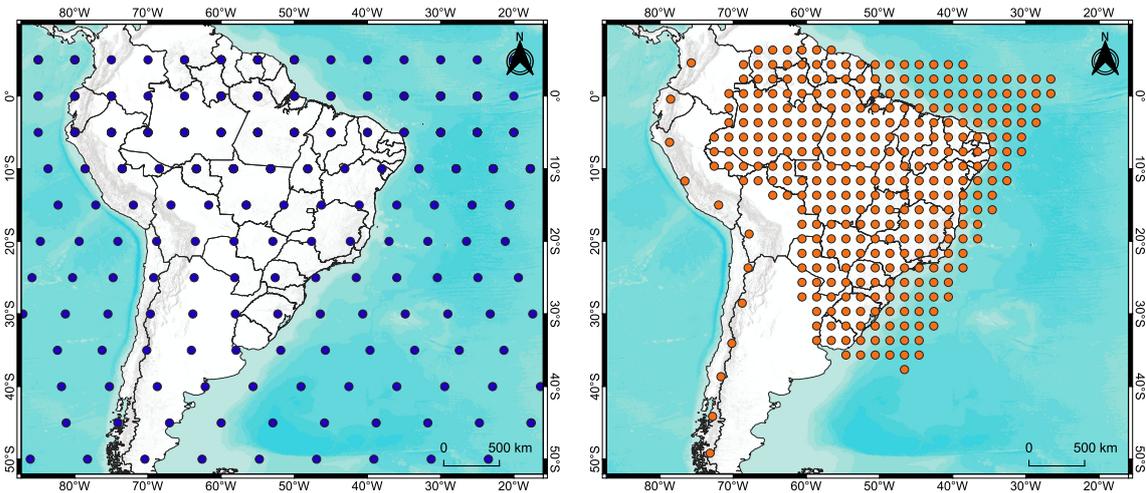


Figura 4.7: Malha de pontos padrão do *SeisComP* (esquerda), em comparação com a malha de pontos proposta (direita).

Os 12 pontos dispostos sobre a região andina não seguem as mesmas regras para a definição de seus parâmetros. Por se tratar de eventos com origens associadas a bordas de placas, com magnitudes majoritariamente maiores do que as apresentadas pelos eventos intraplaca, foi estabelecida a distância máxima (d_{max}) de 35° e no mínimo 9 *picks* (n_{picks}) para que um evento seja nucleado nesses pontos.

Considerando que a quantidade de estações sismográficas no Brasil variou significativamente nos últimos anos, foram elaboradas oito malhas de pontos com diferentes parâmetros, referentes a cada ano entre 2014 e 2021. Para o cálculo da densidade de estações sismográficas em cada ponto, foram consideradas apenas as estações que apresentaram ao menos 50% de dados registrados durante o período de cada ano. Adiante, na seção [Resultados](#) e nos [Apêndices](#), são apresentadas as malhas de pontos utilizadas para cada ano, explicitando as variações anuais da densidade de estações em cada ponto.

4.3.2 Configurações de estações (*station.conf*)

Adicionalmente, mas não menos importante, foi necessário alterar o arquivo *station.conf*, responsável pela configuração de alguns parâmetros das estações contidas no inventário do *SeisComP*. O conteúdo desse arquivo é definido por 4 colunas, como descrito na documentação oficial:

```
* * 1 90
GE * 1 180
GE HLG 1 10
TE RGN 0 10
```

As colunas são, respectivamente, código da rede, código da estação, indicação de uso (0 ou 1), e distância máxima para nucleação. Uma indicação de uso definida pelo número 1 indica que a estação em questão pode ser utilizada durante o processo de localização de eventos. A última coluna indica a máxima distância, em graus, que aquela estação pode contribuir para a geração de novas origens. O exemplo acima indica que todas as estações de todas as redes podem contribuir com leituras de chegadas de onda em eventos a até 90° de distância. As estações da rede GE podem contribuir a qualquer distância, exceto pela estação HLG que está limitada a 10°. A estação RGN da rede TE não será utilizada pelo *scautoloc*.

O arquivo *station.conf* foi modificado de modo a conter uma única linha:

```
* * 1 35
```

Tais parâmetros indicam que todas as estações do inventário devem ser utilizadas e que podem contribuir com leituras de chegadas de ondas de eventos em até 35° de distância. Dessa forma, estabelece-se que as distâncias máximas associadas às origens localizadas pelo *SeisComP* serão definidas apenas pela malha de pontos, configurada através do *gridfile*.

4.3.3 Parâmetros *hard-coded*

Como descrito anteriormente, para aperfeiçoar a detectabilidade de eventos sísmicos regionais em território brasileiro, também é necessário avaliar e reestruturar os parâmetros e fluxos *hard-coded*, os quais não são acessíveis ao usuário comum. Durante o presente trabalho, todo o código-fonte do módulo *scautoloc* foi estudado, de forma a compreender os processos envolvidos na localização de eventos sísmicos e propor mudanças coerentes com o contexto da sismicidade regional do Brasil.

A Figura 4.8 apresenta um fluxograma simplificado da criação de uma origem, envolvendo as funções contidas no código-fonte do módulo *scautoloc* que sofreram as principais modificações ao longo desse trabalho.

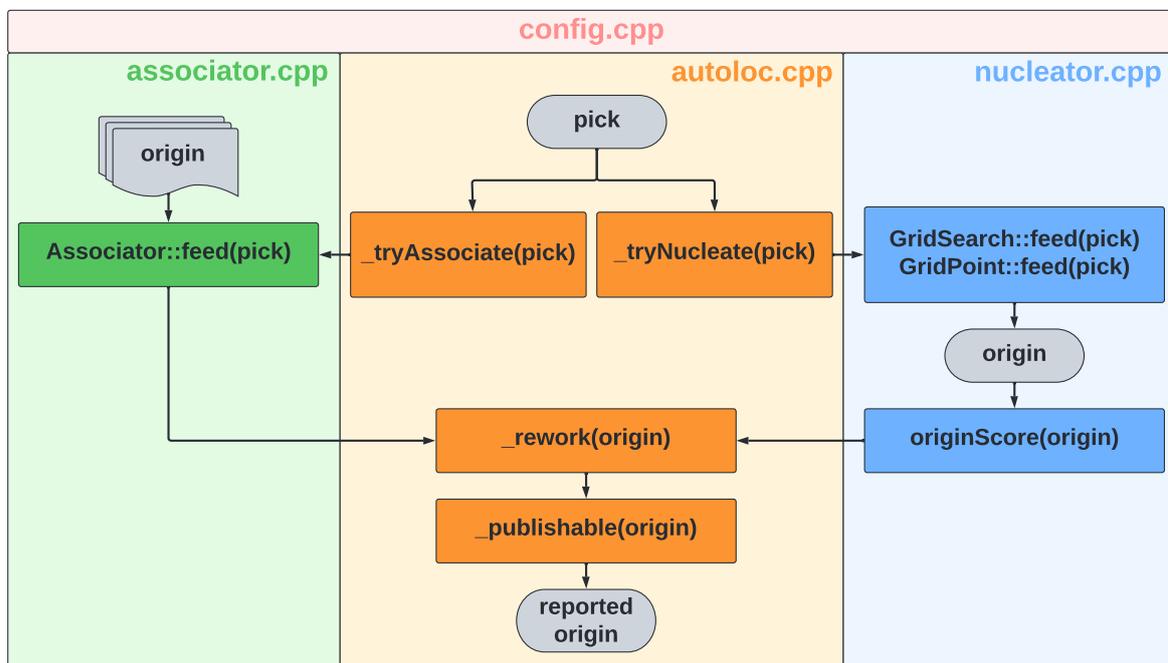


Figura 4.8: Fluxograma dos processos contidos no módulo *scautoloc* que sofreram as principais modificações durante o trabalho. A informação entre parênteses indica o objeto recebido pela função. O arquivo “*config.cpp*” possui parâmetros de configuração que são utilizados em todo o processo.

Ressalta-se que existem diversas funções envolvidas nos processos realizados pelo módulo *scautoloc*, além das apresentadas na Figura 4.8. Dessa forma, a simplificação evidenciada propõe-se a sumarizar as funções que sofreram as principais modificações em parâmetros e/ou fluxos no código-fonte, permitindo que o *SeisComP* detecte mais eventos sísmicos regionais em relação ao *software* original. Tais modificações são apresentadas com maiores detalhes nos tópicos a seguir.

4.3.3.1 Configurações (*config.cpp*)

Dentre os arquivos de códigos do módulo *scautoloc*, há um arquivo de configurações (*config.cpp*) utilizadas durante o processo de localização de eventos sísmicos. Alguns dos parâmetros definidos nesse arquivo não estão disponíveis para serem modificados através do módulo de configuração *scconfig* e, portanto, precisam ser ajustados diretamente no código-fonte do *software* para se obter uma melhor taxa de sucesso nas localizações de eventos sísmicos regionais no Brasil.

A seguir, constam os detalhamentos de cada parâmetro que sofreu alterações diretamente no arquivo *config.cpp*. Os resultados de suas modificações são apresentados na seção [Resultados](#).

- ***goodRMS***: definido no código-fonte como “*typically good RMS in our network*”, esse parâmetro é utilizado em apenas dois momentos durante os processos realizados pelo *scautoloc*, ambos quando já existe ao menos uma origem nucleada. No primeiro momento, esse parâmetro é utilizado para testar se, ao modificar a profundidade da origem para o valor padrão (*defaultDepth*), o novo RMS é melhor do que o RMS com a profundidade calculada. No segundo momento, o parâmetro *goodRMS* é utilizado para verificar a possibilidade de se adicionar novos *picks* a uma origem, realizando a validação através da comparação desse parâmetro (arbitrariamente multiplicado por 2) com os resíduos temporais de cada *pick*. Entretanto, após as modificações propostas, essa verificação foi alterada e passou a utilizar o *maxRMS* para validação;
- ***dynamicPickThresholdInterval***: intervalo de tempo, em segundos, no qual o módulo *scautoloc* começa a descartar *picks* de amplitudes similares de uma certa estação, visando diminuir possíveis utilizações de leituras oriundas de problemas no dado;
- ***maxResidualKeep***: refere-se à tolerância máxima do resíduo temporal de um *pick* para que ele seja associado a uma origem. Diferentemente do *maxResidualUse* (definido no *scconfig*), o *maxResidualKeep* não traduz o maior valor permitido para que um *pick* contribua na localização da origem. Portanto, no caso em que $maxResidualUse < RMS \leq maxResidualKeep$, o *pick* continuará associado à origem em questão, mas não será utilizado no processo de localização. Cabe destacar que, apesar de haver a possibilidade de alterar o valor dessa variável no arquivo *config.cpp*, ela é redefinida para o seu valor original no arquivo *app.cpp*. Dessa forma, é preciso excluir a desnecessária redefinição dessa variável no arquivo *app.cpp* para que a mudança seja realizada de fato;
- ***defaultMaxNucDist***: distância máxima padrão para que uma estação contribua na nucleação de origens. Esse parâmetro é sobrescrito pelo valor definido no arquivo *station.conf*, como apresentado em “[Configurações de estações \(*station.conf*\)](#)”;

- ***minScore***: originalmente, é o valor mínimo aceitável, referente à pontuação arbitrária realizada pelo módulo *scautoloc*, para que uma origem seja aceita. Durante as modificações propostas, esse parâmetro foi renomeado para *minScoreTolerance*, visando uma maior aderência aos novos cálculos da pontuação, sendo descrito em mais detalhes em “Função *_score(origin)*” e “Função *_publishable(origin)*”;
- ***defaultDepth***: valor padrão para ser utilizado como profundidade de uma origem. A profundidade pode ser modificada ao longo do processo, caso os parâmetros de qualidade da origem sejam aperfeiçoados com a mudança;
- ***minimumDepth***: menor profundidade permitida para uma origem. Considerando a sismicidade intra-placa brasileira e a superficialidade dos hipocentros, o valor foi redefinido para 0;
- ***minScoreBypassNucleator***: quando há uma nova leitura de chegada de onda, o módulo *scautoloc* sempre tenta associá-la a uma origem existente antes de adicioná-la no processo de nucleação. Caso o *pick* seja associado a uma origem, o parâmetro *minScoreBypassNucleator* define um valor mínimo de *score* que impede a leitura de seguir para o processo de nucleação, por entender que ela pertence, de fato, à origem que apresentou uma boa pontuação.

4.3.3.2 Associador (*associator.cpp*)

O processo de associação de um novo *pick* a uma origem já existente é iniciado pela função `_tryAssociate(pick)`, contida no arquivo *autoloc.cpp*. Essa função alimenta o associador (*associator.cpp*) com a leitura de chegada de onda, seguindo um fluxo no qual são avaliadas algumas de suas principais informações.

Dentre os processos, destaca-se o cálculo de um parâmetro arbitrário chamado afinidade (*affinity*), o qual determina se o *pick* deverá ser associado à origem em questão. Para que o *pick* seja associado, tal parâmetro deve ser maior ou igual a 0,1, sendo ele definido por:

$$affinity = \cos^2\left(\frac{x \times \pi}{2}\right) \quad (1)$$

sendo x um parâmetro obtido através de uma equação envolvendo o resíduo temporal do *pick* (em segundos) e a distância em relação ao epicentro da origem existente (em graus):

$$x = \frac{residual}{10 \times w} \quad (2)$$

$$w = 1 + 0,6 \times \exp(-0,003 \times dist^2) + 0,5 \times \exp(-0,03 \times (15 - dist)^2)$$

Para reformular o método que define a viabilidade de um *pick* ser associado a uma origem, optou-se por utilizar como referência o resíduo máximo permitido (*maxResidualUse* - configurado previamente pelo usuário) e a distância máxima permitida para a origem (definida pelo *gridpoint* no qual ela foi gerada). Dessa forma, a equação original foi substituída pela seguinte condição:

$$\begin{cases} \text{associar,} & \text{se } |residual| \leq maxResidualUse \text{ e } dist \leq gridMaxDist \\ \text{não associar,} & \text{se } |residual| > maxResidualUse \text{ ou } dist > gridMaxDist \end{cases} \quad (3)$$

Dessa forma, caso o *pick* cumpra com os requisitos de distância e resíduo temporal, ele será associado à origem em questão. Essa foi a principal modificação realizada no processo de associação de *picks*, visando permitir que eventuais novas leituras de chegada de onda sejam associadas às origens de forma menos arbitrária e obedecendo as condições impostas pelo usuário ao invés de cálculos arbitrários.

4.3.3.3 Nucleador (*nucleator.cpp*)

Como descrito anteriormente, caso o processo de associação falhe, o *pick* é introduzido no fluxo do processo de nucleação por meio da função *_tryNucleate(pick)*, a qual aciona o nucleador (*nucleator.cpp*) para realizar tentativas de criação de novas origens.

Funções *GridPoint::feed(pick)* e *GridSearch::feed(pick)*

Originalmente, o nucleador projeta o *pick* em todos os pontos da malha (através da função *GridPoint::feed(pick)*), ao mesmo tempo em que checa se há um número mínimo de *picks* agrupados em cada ponto para dar início às verificações que estabelecem requisitos para que uma nova origem possa ser criada. Esse requisito inicial (quantidade mínima de *picks* agrupados em um determinado ponto) está *hard-coded* e definido com o valor 6, ou seja, são necessárias no mínimo 6 leituras de chegadas de onda para uma origem ser criada. Apesar de haver um parâmetro editável no *sconfig* denominado *minPhaseCount*, o mesmo não é utilizado nessa parte do processo.

Considerando que a RSBR apresenta uma baixa densidade de estações em diversas regiões do território nacional, é prudente modificar o número mínimo necessário para se nuclear novas origens. Entretanto, o efeito adverso é o provável aumento da quantidade de eventos falsos nucleados, ou seja, grupos de *picks* associados a ruídos que foram erroneamente interpretados como chegadas de onda de eventos sísmicos reais. Dessa forma, o número mínimo de *picks* agrupados em um determinado ponto para seguir com o processo de nucleação foi redefinido para 5.

Quando a função *GridPoint::feed(pick)* identifica um grupo de *picks* que atende ao critério inicial supracitado, uma “origem candidata” é criada. Com a introdução de um novo *pick* em todos os pontos da malha, é comum que mais de um ponto atenda ao critério de quantidade mínima, gerando mais do que uma origem candidata. A partir do momento em que uma origem candidata é criada, os *picks* associados à ela são chamados de *arrivals* (chegadas de ondas).

As origens candidatas são retornadas para a função *GridSearch::feed(pick)*, a qual utiliza diversos critérios para definir se, de fato, uma nova origem será reportada. Nessa função, muitos critérios foram modificados a fim de permitir uma maior detectabilidade de eventos sísmicos no Brasil, uma vez que, novamente, vários parâmetros e processos originais são arbitrários e definidos visando a detecção de sismos globais associados a bordas de placas.

Um dos principais motivos para origens candidatas potencialmente promissoras (*i.e.*, as que podem vir a se tornar uma origem de evento sísmico real) serem descartadas em um estágio preliminar é o fato de haver muitas verificações desnecessárias de parâmetros ao longo de todo o processo. Por diversas vezes, o `GridSearch::feed(pick)` verifica se a origem preliminar já atende aos requisitos mínimos de número de leituras para ser publicada, descartando-a em caso negativo. Como há a possibilidade de novas leituras serem associadas à origem após a sua criação, não é ideal descartá-la nesse momento, sendo prudente fazer essa verificação apenas no momento imediatamente anterior à publicação da origem, ou seja, quando ela será de fato aceita como uma origem de evento sísmico.

Um segundo aspecto crucial nessa etapa do processo de localização é o fato de que o nucleador retorna apenas uma origem dentre todas as candidatas, elegendo a melhor com base em alguns critérios, incluindo uma pontuação arbitrária que será explicada adiante. Essa questão está comentada no código-fonte do nucleador, mostrando que é um ponto a ser revisitado pelos desenvolvedores no futuro:

*“Now for all ‘candidate’ origins in tempOrigins try to find the ‘best’ one. This is a bit problematic as we don’t go back and retry using the second-best but give up here. Certainly scope for improvement.”*³

Dessa forma, os processos realizados pelo nucleador foram reformulados, objetivando remover as verificações desnecessárias realizadas em estágios preliminares, bem como permitir que mais de uma origem candidata seja nucleada. Com essas alterações, todas as origens candidatas são passadas adiante e, caso não cumpram com os requisitos definidos pelo usuário, são descartadas em um momento oportuno do processo, quando não há mais possibilidade de melhorias (na função `_publishable(origin)`).

³Comentário extraído do código-fonte do módulo `scautoloc` (nucleator.cpp - linha 653).

Função *_originScore(origin)*

O nucleador possui uma função que atribui uma pontuação arbitrária (*score*) à origem, utilizando-a como parâmetro para manter ou descartar a origem durante o processo. Essa função também é acessada fora do nucleador, de forma a atualizar a pontuação quando for necessário. A pontuação é baseada nas principais informações de cada *arrival* da origem:

- Distância epicentral e profundidade;
- Amplitude e SNR; e
- Resíduo temporal.

Novamente, todos os cálculos realizados para obtenção da pontuação são arbitrários, bem como o limite definido para que a origem seja descartada (*minScore*, por padrão, definido como 8). A pontuação é calculada, originalmente, da seguinte maneira:

$$originScore = \left(\sum_{n=0}^{N_{arr}} arrivalScore[n] \right) \times depthFactor$$

$$arrivalScore = distScore \times amplScore \times timeScore \times phaseScore$$

$$distScore = 1,5 \times \exp\left(\frac{-d \times d}{r^2}\right)$$

$$amplScore = 1,0 + 0,8 \times \left(1,0 + 0,5 \times \log_{10}\left(\frac{amp}{2000}\right)\right) \times \log_{10}(snr) \quad (4)$$

$$timeScore = 0,5 \times \left(\cos\left(\frac{\left(\frac{|residual|}{2 \times maxRMS} - 0,2\right) \times \pi}{1,0 - 0,2}\right) + 1,0 \right)^2$$

$$phaseScore = \begin{cases} 1,0, & \text{se o arrival for P} \\ 0,3, & \text{se o arrival for PKP} \end{cases}$$

$$depthFactor = 1 + 0,0005 \times (200 - depth)$$

Nas equações acima, d é a distância da estação que produziu o *pick* até a origem, r é a distância máxima de nucleação permitida para a estação em questão (definida no arquivo *station.conf*), amp e snr são as amplitudes m_b e SNR calculadas pelo módulo *scautopick*, enquanto $maxRMS$ é o máximo RMS permitido para que uma origem seja reportada.

Todas as constantes, bem como as próprias equações, são arbitrárias e definidas pelos desenvolvedores do *SeisComP*, visando a localização de sismos de escalas globais. Também não há uma normalização do *score* de uma origem com base em sua quantidade de *arrivals*, o que permite que uma origem com muitos *arrivals* “ruins” seja aceita como, por exemplo, em um cenário no qual a somatória dos *scores* baixos de seus *arrivals* atinge a pontuação mínima necessária para a aceitação.

As equações utilizadas para calcular uma pontuação para as origens foram totalmente redefinidas, buscando seguir um padrão que auxiliasse a escolha de um limite aceitável para que a origem seja definida como pertencente a um evento sísmico real. Dessa forma, as seguintes premissas foram adotadas:

- Cada *arrival* de uma origem contribui com um *score* entre 0 e 1;
- As penalizações nos *scores* devem ser, preferencialmente, baseadas nos parâmetros pré-definidos pelo usuário;
- As equações devem ser simplificadas ao máximo; e
- O *score* de cada *arrival* é a média dos *scores* de cada parâmetro, enquanto o *score* da origem é a média dos *scores* de todos os *arrivals*.

As novas equações foram propostas utilizando-se uma base de dados de eventos sísmicos brasileiros, cujas marcações de chegadas de onda foram verificadas e consolidadas previamente. A partir dela, várias rodadas de testes de nucleação de eventos foram realizadas para se encontrar o conjunto de equações que permitisse a maior detectabilidade de eventos reais, ao mesmo tempo em que a ocorrência de falsos-positivos fosse minimizada.

Com isso, todas as partes do cálculo do *score* de uma origem foram reformuladas e são descritas a seguir. O valor de *score* mínimo para que uma origem seja aceita só pode ser alterado pelo usuário, originalmente, através de linha de comando. No código-fonte, esse parâmetro está definido no arquivo *config.cpp* (*minScore*) e possui o valor original de 8.

O novo valor de *score* mínimo é obtido através de um cálculo realizado ao final do processo de criação de uma origem, durante a avaliação dos principais parâmetros da origem a ser publicada, e está detalhado em “Função *_publishable(origin)*”. Também foi adicionada uma nova opção no módulo de configuração *seconfig* para que o usuário possa ter acesso ao *score* mínimo de forma simplificada, através da interface gráfica.

• *Distance Score*

$$distScore = 1,0 - \left(0,25 \times \frac{arrivalDistance}{gridMaxDist} \right) \quad (5)$$

Para o cálculo do *distScore*, optou-se por utilizar uma equação que penalizasse a pontuação conforme a distância da origem à estação que registrou o *arrival* se aproximasse da máxima distância permitida pelo ponto da malha em que a origem foi nucleada. *Arrivals* com distâncias maiores do que a máxima permitida pelo *grid point* não são utilizados no processo de localização, ou seja, o menor *distScore* possível de ser obtido por um *arrival* que contribui com uma origem é 0,75.

A Figura 4.9 apresenta o gráfico referente ao cálculo do parâmetro *distScore*.

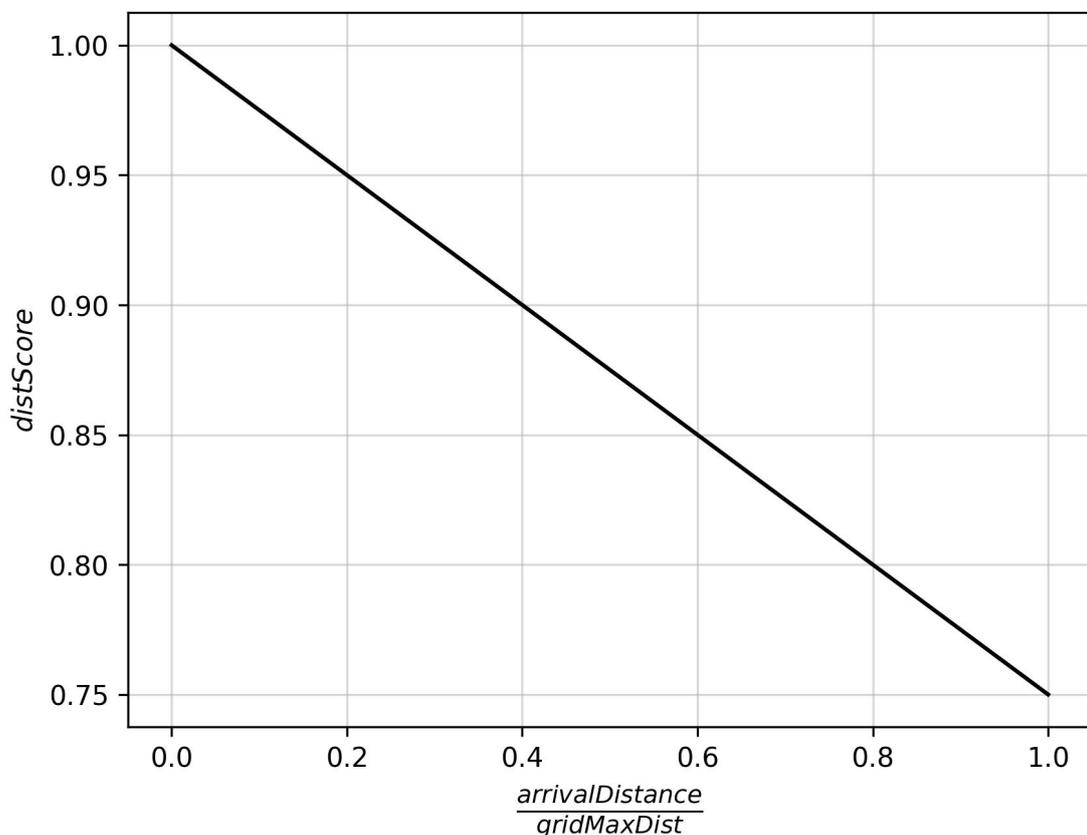


Figura 4.9: Representação gráfica do novo cálculo do parâmetro *distScore*. As maiores pontuações são obtidas com as menores distâncias, enquanto a penalização fica mais severa conforme a distância do *arrival* se aproxima à distância máxima permitida (*gridMaxDist*).

• *Amplitude Score*

O cálculo da pontuação relacionada à amplitude da chegada de onda (*amplScore*) foi severamente reformulado, de forma a introduzir o conceito de magnitude durante o processo. O cálculo da magnitude de um evento sísmico leva em consideração a distância da estação até o epicentro do evento e, portanto, é um parâmetro que auxilia na validação de um *arrival*.

Originalmente, como visto no conjunto de [Equações 4](#), o cálculo do *amplScore* é realizado apenas levando em consideração a maior amplitude do sismograma na chegada da onda P, além de realizar normalizações e operações arbitrárias com esse parâmetro. Nesse trabalho, propõe-se utilizar a amplitude e distância epicentral de cada *arrival* para se calcular uma “magnitude MLv preliminar” do evento.

Para o cálculo da magnitude MLv em cada estação, foi implementada a equação derivada de [Richter \(1935\)](#), em conjunto com a função de calibração proposta por [Alsaker et al. \(1991\)](#), como mostra a [Equação 6](#).

$$MLv = \log_{10}(amp) - \log_{10}(A_0) \quad (6)$$

$$-\log_{10}(A_0) = 0,91 \times \log_{10}(dist) + (0,00087 \times dist) + 1,01$$

A [Figura 4.10](#) apresenta um exemplo gráfico do resultado do cálculo da magnitude MLv utilizando-se dados de 10 estações sismográficas. Cada estação contribui com um valor de magnitude, sendo adotada a mediana dos valores como a magnitude de referência.

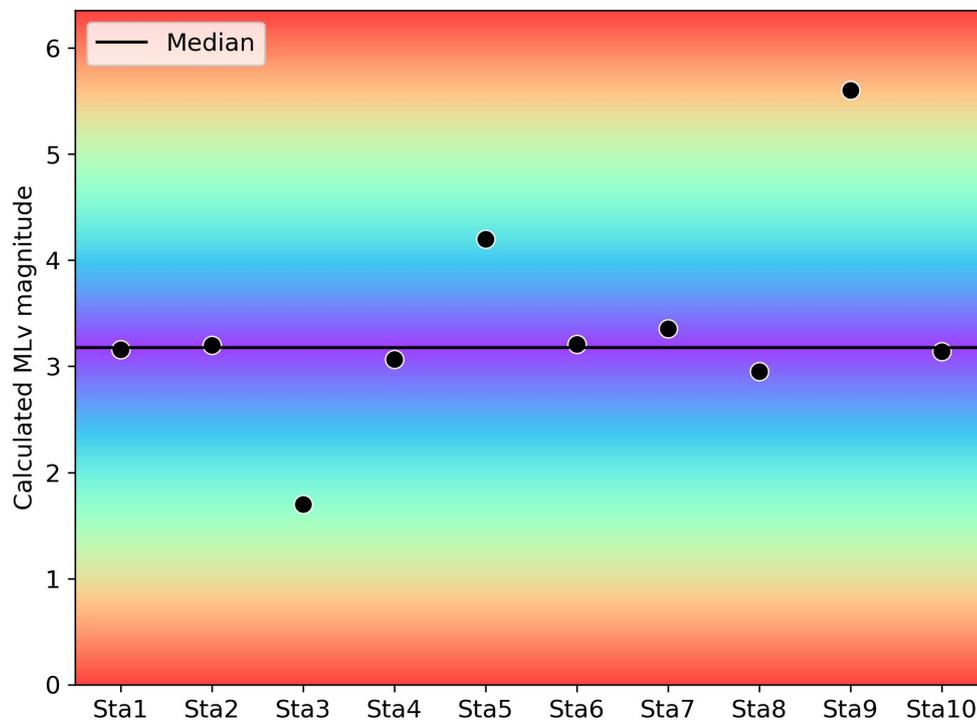


Figura 4.10: Representação gráfica de um exemplo da obtenção da magnitude MLv preliminar (mediana). Quanto mais distante da mediana, maior a penalização no *amplScore* do *arrival* em questão. A escala de cores representa a penalização de forma qualitativa, sendo a cor roxa a representativa das menores penalizações e a cor vermelha das maiores.

Após o cálculo da magnitude MLv em cada estação, o *ampScore* é obtido através da Equação 7, a qual penaliza *arrivals* que apresentam magnitudes distantes da mediana, como ilustra a escala de cores do exemplo da Figura 4.10.

$$\text{ampScore} = \begin{cases} 0, & \text{se } |\text{arrivalMag} - \text{medianMag}| > 4/3 \\ 1,0 - (0,75 \times (|\text{arrivalMag} - \text{medianMag}|)), & \text{caso contrário} \end{cases} \quad (7)$$

Esse procedimento foi adotado utilizando-se da premissa de que estações que tenham contribuído com *picks* associados a um mesmo evento sísmico real tendem a apresentar magnitudes calculadas próximas umas às outras e, portanto, leituras que apresentam magnitudes que diferem significativamente da mediana do grupo possuem maior probabilidade de serem falsas.

A Figura 4.11 apresenta o gráfico referente ao cálculo do parâmetro *ampScore*.

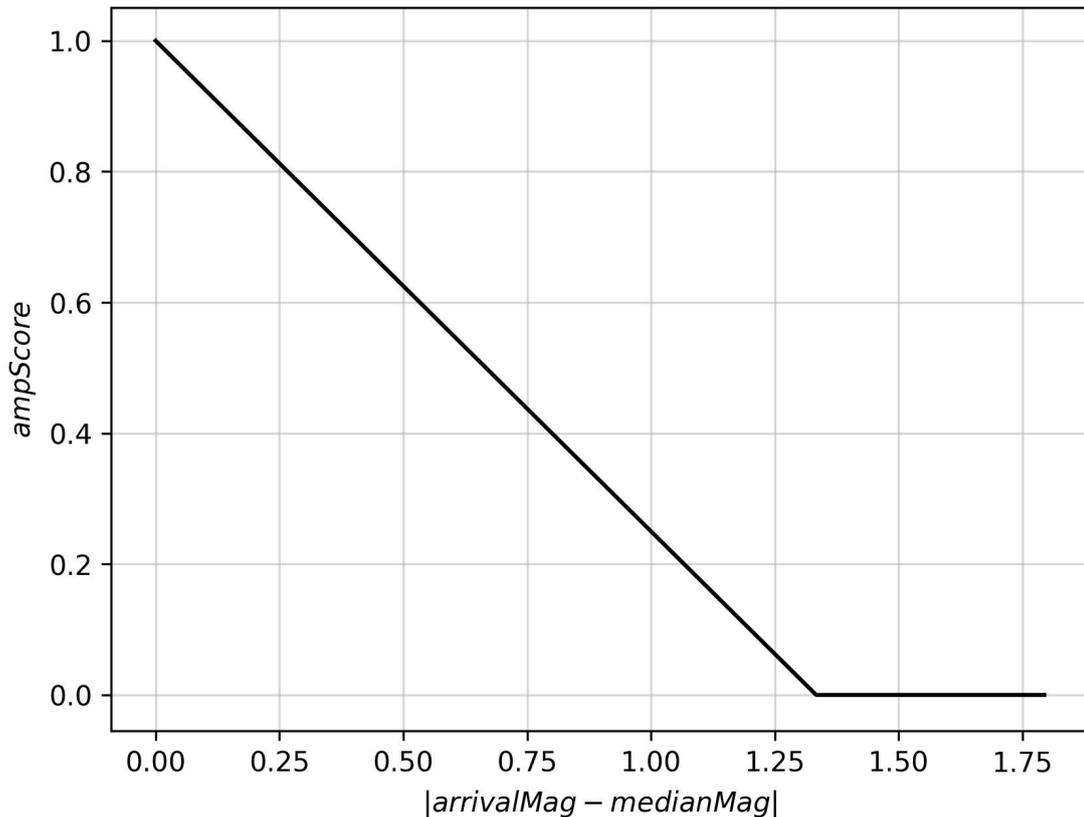


Figura 4.11: Representação gráfica do novo cálculo do parâmetro *ampScore*. As maiores pontuações são obtidas com as menores diferenças entre a magnitude do *arrival* e da mediana do conjunto. A partir de uma determinada diferença entre as magnitudes ($4/3$), a pontuação torna-se zero.

Para que a alteração no cálculo do parâmetro *ampScore* atinja o resultado esperado, é necessário garantir que a amplitude MLv esteja sendo calculada pelo módulo *scautopick* e que o módulo *scautoloc* esteja configurado para utilizá-la como amplitude preferencial, como mostrado na seção Resultados.

• *Time Score*

$$timeScore = 1,0 - \left(0,25 \times \frac{|residual|}{maxRMS} \right) \quad (8)$$

Para o cálculo do *timeScore*, optou-se por utilizar uma equação que penalizasse a pontuação conforme o resíduo do *arrival* se aproximasse do máximo resíduo permitido. *Arrivals* com resíduos temporais maiores do que o máximo permitido não são utilizados no processo de localização, ou seja, o menor *timeScore* possível de ser obtido por um *arrival* que contribui com uma origem é 0,75.

A Figura 4.12 apresenta o gráfico referente ao cálculo do parâmetro *timeScore*.

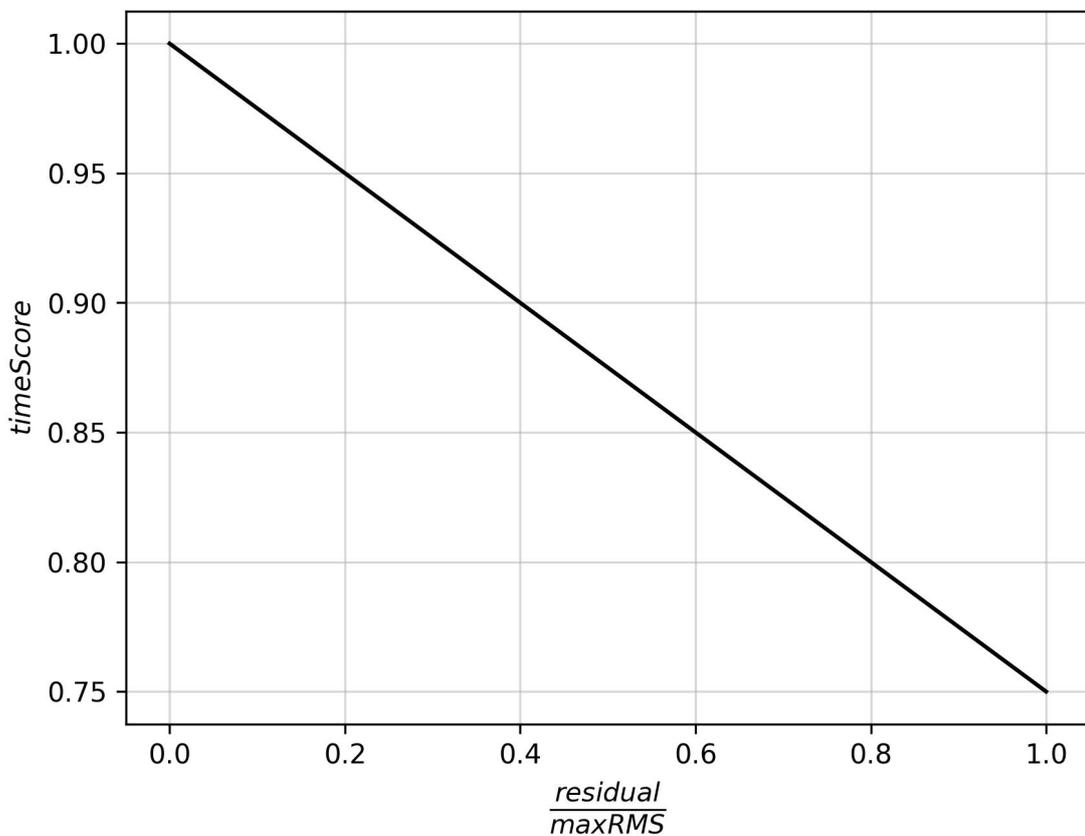


Figura 4.12: Representação gráfica do novo cálculo do parâmetro *timeScore*. As maiores pontuações são obtidas com os menores resíduos de tempo, enquanto a penalização fica mais severa conforme o resíduo do *arrival* se aproxima do máximo resíduo permitido (*maxRMS*).

A utilização do modelo de velocidades BRA23 possui uma relação direta com o parâmetro *timeScore*, uma vez que, ao utilizá-lo para localizar eventos sísmicos no Brasil, observa-se uma tendência de diminuição dos resíduos temporais dos *arrivals*. Com isso, as penalizações propostas para o *timeScore* tendem a não causar grandes impactos na pontuação das origens relacionadas a eventos sísmicos reais.

• **Origin Score**

Ao final das reformulações dos cálculos das pontuações de cada parâmetro, a pontuação final da origem (*originScore*) passou a ser calculada através da seguinte equação, na qual N_{arr} refere-se ao número de *arrivals* que contribuem para a localização da origem.

$$originScore = \frac{\left(\sum_{n=1}^{N_{arr}} arrivalScore[n]\right)}{N_{arr}} \quad (9)$$

$$arrivalScore = \frac{distScore + amplScore + timeScore}{3,0}$$

Adicionalmente, as seguintes penalizações à pontuação final da origem são aplicadas, caso necessário:

$$originScore = originScore \times depthFactor$$

$$depthFactor = \begin{cases} 1,00, & \text{se } depth < 20 \text{ km} \\ 0,95, & \text{se } 20 \text{ km} < depth \leq 30 \text{ km} \\ 0,80, & \text{se } depth > 30 \text{ km} \end{cases} \quad (10)$$

$$originScore = \begin{cases} originScore - 0,10, & \text{se } distancePenalty \text{ for acionado} \\ originScore - 0,05, & \text{se } phaseCountPenalty \text{ for acionado} \end{cases}$$

No que se refere às distâncias dos *arrivals* em relação à origem, foi incluída uma nova penalidade ao *score* (*distancePenalty*). Essa penalidade é aplicada (i) caso a primeira chegada de onda tenha sido registrada em uma estação localizada a uma distância maior do que a metade da máxima distância permitida pelo ponto na malha que a originou (*gridMaxDist*) ou (ii) quando a média das distâncias de todas as chegadas é maior do que $0,65 \times gridMaxDist$.

Baseando-se na premissa de que estações próximas aos eventos são as que os registram em maior quantidade, tal penalidade visa impedir que a nucleação de ruídos em estações distantes gerem uma origem falsa. *Arrivals* com distâncias maiores do que a máxima permitida não são utilizados no processo de localização, ou seja, o menor *distScore* possível de ser obtido por um *arrival* que contribui com uma origem é 0,75.

Com relação ao número de fases associadas à origem (*i.e.*, número de estações que contribuíram para a localização), buscou-se penalizar as origens que apresentam o número mínimo permitido para nucleação (*minPhaseCount*), uma vez que permitir a nucleação com menos estações pode acarretar em um aumento de falso positivos. Essa penalização é opcional, podendo ser desativada pelo usuário através do parâmetro “*phaseCountPenalty*”, implementado ao módulo *scconfig* após as modificações.

Função *_rework(origin)*

Um dos processos mais importantes durante a criação de uma origem é a sua realocização. Inicialmente, a nucleação de uma origem gera um hipocentro em algum dos pontos definidos na malha de pontos, sendo necessário refinar essa localização inicial para que os parâmetros de qualidade sejam otimizados. A remoção de *arrivals* da origem devido a parâmetros de qualidade ruins e a modificação da profundidade do evento são exemplos de momentos nos quais torna-se necessário realizar a realocização de uma origem.

De forma simplificada, a função *_rework(origin)* é responsável por realizar a realocização de uma origem. Assim como em outros locais do código, nessa função também há uma excessiva realização de testes que podem invalidar o prosseguimento de uma origem no processo, tais como a contagem de *arrivals*. Visando o mesmo objetivo de manter tais validações concentradas apenas ao final do processo, a função *_rework(origin)* também foi reestruturada.

No processo de realocização, também excluem-se os *arrivals* com resíduos temporais acima do permitido, além de tentar adicionar possíveis outros *picks* à solução. Leituras realizadas por estações distantes também são descartadas, além de possíveis fases PkP identificadas pelo *software*. Ao final da realocização, os critérios da origem referentes à distância máxima (d_{max}) e número mínimo de *picks* (n_{picks}) são redefinidos com base no ponto do *grid* mais próximo ao novo hipocentro.

Função *_publishable(origin)*

Como mencionado anteriormente, um dos maiores impedimentos de uma origem promissora ser classificada como uma origem válida e conseqüentemente reportada é a quantidade excessiva de testes de qualidade realizados de maneira precoce. Uma origem potencialmente boa pode apresentar, em suas fases iniciais de formação, parâmetros de qualidade abaixo dos configurados pelo usuário para que ela seja aceita.

Por exemplo, devido ao fato de que a nucleação é feita com base em uma malha de pontos discretizada de forma arbitrária, o ponto no qual origem foi inicialmente localizada pode resultar em grandes resíduos temporais por não ser exatamente o epicentro real do evento. Nesse caso, não convém eliminar a origem por causa de seu alto RMS, pois ela deverá apresentar uma melhora nesse parâmetro quando for realocizada para mais próximo de seu epicentro real.

Um segundo exemplo é o caso em que uma origem promissora é inicialmente nucleada com menos estações do que o número definido pelo usuário para que uma origem seja aceita. No código-fonte original, o teste de número de estações é realizado diversas vezes, inclusive imediatamente após a nucleação. Por entender que novas leituras podem ser associadas à origem em diversos momentos ao longo do processo, julgou-se ser imprudente eliminar a origem nos momentos iniciais devido ao número mínimo de *arrivals*. Situações similares ocorrem ao longo de todo o processo, até a efetiva publicação da origem.

Dessa forma, buscou-se remover do código-fonte diversas checagens de validação da origem em momentos considerados desnecessários, concentrando-os preferencialmente no momento em que a origem está prestes a ser publicada, ou seja, aceita como uma origem verdadeira. A função `_publishable(origin)` se tornou a responsável por fazer os últimos testes e conferências de parâmetros, a fim de verificar se a origem é “publicável”. Cabe destacar os seguintes processos, realizados logo após a última realocização efetivada:

- Cálculo das distâncias e azimutes de cada *arrival* da origem, garantindo o armazenamento dos valores corretos;
- Validação do número mínimo de *arrivals*, utilizando como base o valor referente ao ponto da malha mais próximo ao hipocentro do evento;
- Validação do *azimuth gap* da origem;
- Validação do *score* da origem, cujo cálculo está explicado em “`_score(origin)`”. O valor de *score* mínimo para que uma origem seja aceita é explicado a seguir;
- Validação do RMS da origem; e
- Validação da profundidade da origem.

Como mencionado em “`_score(origin)`”, a reformulação da pontuação também incluiu o *score* mínimo necessário para que uma origem seja aceita. Anteriormente, esse valor era fixo e definido no código-fonte (`config.cpp`) ou modificado pelo usuário através da linha de comando (argumento `--min-score`).

Com as modificações propostas, o novo *score* mínimo tornou-se variável e foi atrelado à máxima distância na qual uma estação pode contribuir para a origem (d_{max} , parâmetro definido no arquivo `grid.conf` e explicitado em “Malha de pontos (`grid.conf`)”). Isso se deve ao fato de que origens que permitem a contribuição de estações mais distantes possuem maior probabilidade de nuclear ruídos e, portanto, torna-se necessária uma maior rigidez na avaliação da pontuação.

A variável `minScore` foi renomeada para `minScoreTolerance` e o valor de 0,75 lhe foi atribuído, sendo somado a ela um peso de acordo com a máxima distância permitida, como mostra a Equação 11.

$$\text{minScore} = \text{minScoreTolerance} + \frac{\text{gridMaxDist}}{100} \quad (11)$$

Dessa maneira, uma origem que possui distância máxima de nucleação de 10 graus precisa apresentar um *score* de no mínimo 0,85 para ser aceita. Em contrapartida, origens com chegadas exclusivamente mais próximas (máximo de 5 graus, por exemplo) podem ser aceitas se apresentarem um *score* de valor 0,80.

5. Resultados

Como mencionado na seção “Métodos e procedimentos”, os parâmetros foram ajustados seguindo etapas de processamentos e análises, partindo de um cenário com poucos dados (170 *arrivals* de 23 eventos selecionados) até abranger todos os dados contínuos do período de 2014 a 2021 (\approx 11TB de dados). A Figura 5.1 apresenta um *corner plot* dos resultados da primeira etapa de otimização dos parâmetros de filtros de frequências e STA/LTA, como descrito na Seção 4.

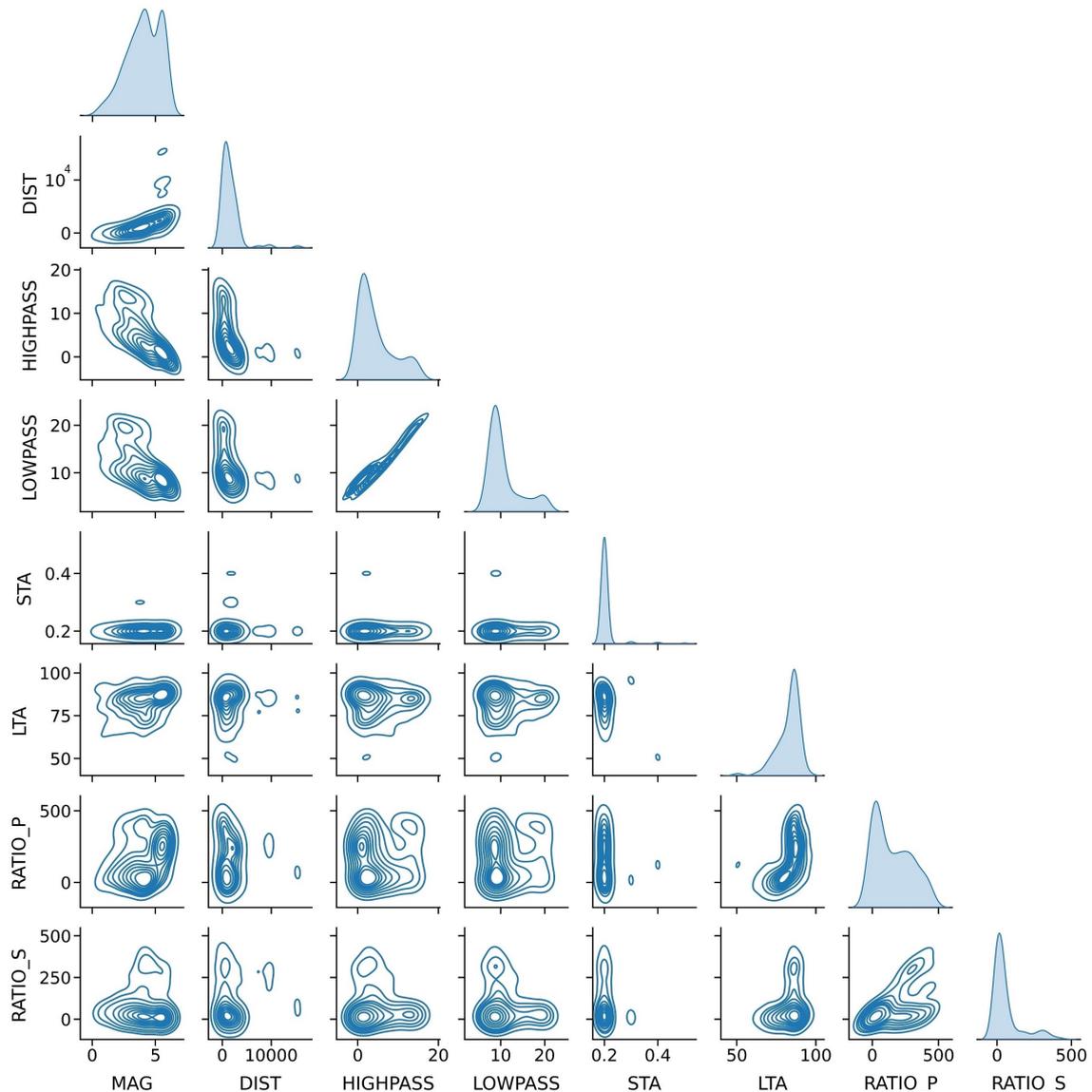


Figura 5.1: *Corner plot* dos principais parâmetros dos 170 *arrivals* utilizados para se realizar a primeira etapa de otimização dos parâmetros de detecção.

Visando a detecção automática das chegadas das ondas P dos eventos sísmicos regionais, buscou-se combinações de filtros passa-banda e janelas de STA e LTA que proporcionassem a maior razão sinal-ruído possível dessas chegadas. Dos gráficos da diagonal da Figura 5.1, foi possível obter os valores iniciais para esses parâmetros, sendo eles: limites do filtro passa-banda de aproximadamente 4 Hz e 9 Hz e janelas de STA e LTA de aproximadamente 0,2

e 80 segundos, respectivamente. Em seguida, diversas etapas de processamento e rodadas de testes foram realizadas buscando refinar esses parâmetros de detecção, em conjunto com outros parâmetros e processos citados a seguir.

Dos gráficos, também destacam-se alguns aspectos dos resultados obtidos inicialmente, como a relação entre a distância epicentral e os limites do filtro de frequências, a qual corrobora com o esperado, uma vez que eventos com distâncias epicentrais menores tendem a apresentar maiores frequências e vice-versa.

5.1 Filtro de frequências e parâmetros STA/LTA

Ao fim de todas as etapas de ajustes de parâmetros, os limites do filtro de frequências ideal foram definidos como 4,5 Hz e 10 Hz, sendo esses os valores que proporcionaram maior realce das ondas P da maior parte dos eventos sísmicos avaliados. A Figura 5.2 mostra a comparação entre as formas de onda filtradas com os parâmetros padrões do *SeisComP* e com os parâmetros propostos.

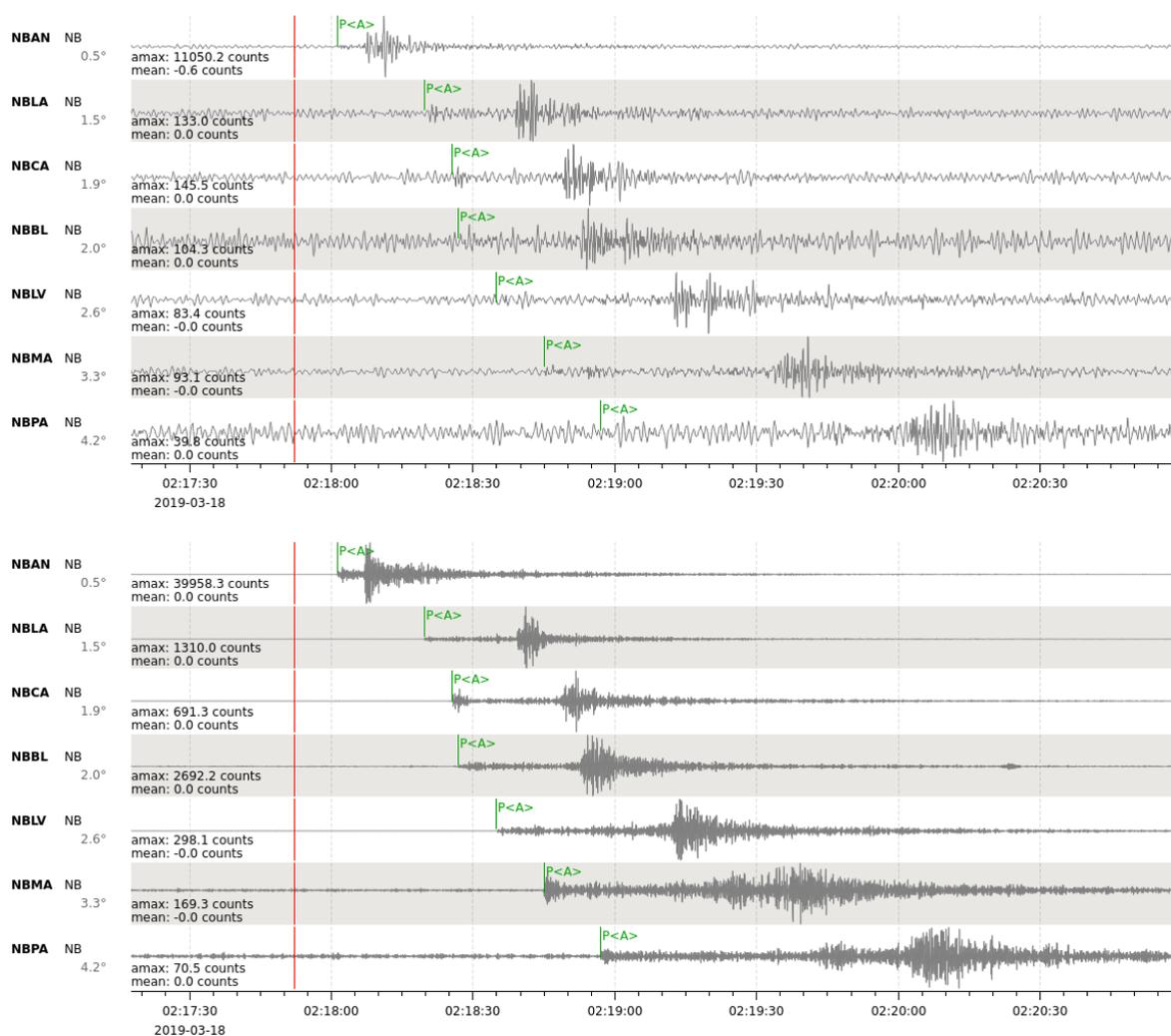


Figura 5.2: Exemplos de filtragens das formas de onda referentes a um evento sísmico de magnitude M2,5 ocorrido em Olho d'Água Grande/AL. Acima, foram utilizados os limites padrões do *SeisComP* (0,7 Hz e 2 Hz). Abaixo, os dados filtrados com os limites propostos nesse trabalho (4,5 Hz e 10 Hz).

Ao fim de todas as etapas de refinamento dos parâmetros, as janelas de tempo para o algoritmo STA/LTA foram definidas como 0,2 e 45 segundos, respectivamente. A Figura 5.3 mostra a diferença entre as formas de onda filtradas com a combinação dos parâmetros padrões do *SeisComP* e com a combinação dos parâmetros propostos.

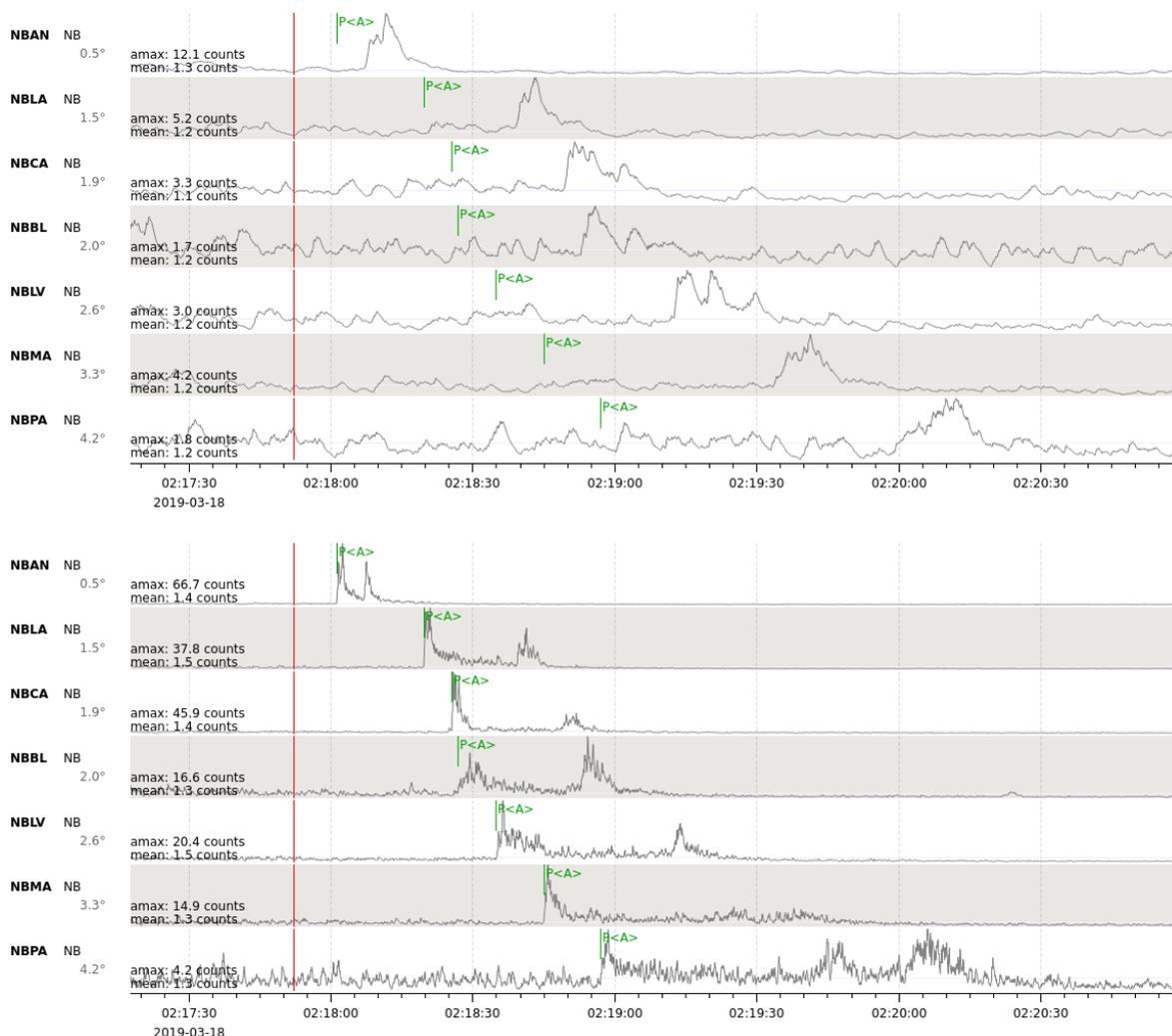


Figura 5.3: Exemplos de filtragens e aplicação de STA/LTA nas formas de onda referentes a um evento sísmico de magnitude $M_{2,5}$ ocorrido em Olho d'Água Grande/AL. Acima, foram utilizados os parâmetros padrões do *SeisComP* (frequências de 0,7 Hz a 2 Hz e STA/LTA de 2 e 80 segundos). Abaixo, os dados processados com os parâmetros propostos nesse trabalho (frequências de 4,5 Hz a 10 Hz e STA/LTA de 0,2 e 45 segundos).

Os dados processados com os parâmetros padrões do *SeisComP* apresentam uma razão STA/LTA deficiente nos momentos das chegadas das ondas P, sendo elas praticamente indetectáveis pelo *software*. Nesse caso, apenas as ondas S ou superficiais poderiam ser detectadas pelo sistema automático de leitura de chegadas de ondas. Em diversos outros casos, nem mesmo essas chegadas são identificáveis pelo sistema.

Utilizando-se os parâmetros propostos, a chegada da onda P e sua razão STA/LTA são consideravelmente realçadas, o que permite sua correta leitura nos sismogramas.

5.2 Parâmetros de configuração

Nessa subsecção, expõem-se as modificações propostas para os diversos parâmetros de configuração, sejam eles acessíveis através do módulo *sconfig* ou apenas através do código-fonte do *software*. As tabelas apresentam os parâmetros modificados utilizando-se a sintaxe do *SeisComP*, a mesma adotada no módulo de configurações e no código-fonte do *software*.

5.2.1 Módulos *scautopick* e *scautoloc*

Tabela 5.1: Parâmetros de configurações do módulo *scautopick* acessíveis através da interface gráfica do *sconfig*.

	Original	Proposto
filter	RMHP(10) ITAPER(30) BW(4,0.7,2) STALTA(2,80)	RMHP(10) BW(4,4.5,10) STALTA(0.2,45)
timeCorrection	-0,8	0
picker		AIC
thresholds.triggerOn	3,0	3,0
thresholds.triggerOff	1,5	0,7
thresholds.maxGapLength	4,5	1
thresholds.amplMaxTimeWindow	10	3
thresholds.deadTime	30	2,5

Da [Tabela 5.1](#), destacam-se as alterações realizadas no filtro de frequências e das janelas de STA/LTA para a detecção de chegada de ondas, bem como a ativação do algoritmo AIC como *picker*. No filtro, optou-se por remover a aplicação do *taper* (*ITAPER(30)*) por ter sido observado que isso mantém o *picker* inativo por um período após uma detecção.

O parâmetro *timeCorrection* foi ajustado para o valor zero, uma vez que o *picker* AIC torna-se o responsável por corrigir a detecção realizada pelo algoritmo de STA/LTA. O parâmetro *amplMaxTimeWindow* refere-se ao tamanho da janela de tempo, em segundos, usada para calcular a razão sinal-ruído do *pick*. Como o dado foi filtrado entre 4,5 Hz e 10 Hz, não há necessidade de uma janela grande para realizar esse cálculo, além do fato de que sua diminuição também contribui para que o *pick* seja reportado mais rapidamente.

Apesar de o valor do parâmetro *triggerOn* se manter o mesmo, o valor de *triggerOff* foi diminuído para 0,7. Em conjunto com o novo *deadTime* de 2,5 segundos, ambos os parâmetros se referem a métricas para que o *picker* não realize diversas leituras seguidas durante um mesmo evento sísmico, desativando-o enquanto a razão STA/LTA permanecer maior do que 0,7 ou durante o intervalo de tempo configurado. Como eventos regionais tendem a ter uma curta duração de tempo, não há a necessidade de manter o *picker* desativado por 30 segundos, como é configurado originalmente.

A Tabela 5.2 apresenta os parâmetros relacionados ao módulo *scautoloc* e que foram modificados através do módulo de configuração *scconfig*.

Tabela 5.2: Parâmetros de configurações do módulo *scautoloc* acessíveis através da interface gráfica do *scconfig*.

	Original	Proposto
locator.defaultDepth	10	0
locator.minimumDepth	5	0
autoloc.maxDepth	1000	50
autoloc.maxRMS	3,5	0,8
autoloc.maxResidual	7,0	1,2
autoloc.minPhaseCount	6	5
autoloc.maxStationDistance	180	35
autoloc.amplTypeAbs	mb	MLv
locator.profile	iasp91	BRA23
autoloc.phaseCountPenalty*		true
autoloc.dropAndeanOrigins*		true
autoloc.minScoreTolerance*		0,75

Considerando que os eventos sísmicos regionais brasileiros ocorrem majoritariamente em profundidades superficiais, buscou-se refinar o intervalo de profundidades permitidas para que o *scautoloc* associe às origens criadas. Adicionalmente, visando diminuir o número de origens falsas, o máximo RMS permitido foi alterado de 3,5 para 0,8, além da diminuição do máximo resíduo temporal permitido para cada *arrival* de 7,0 para 1,2. Assim, origens com grandes resíduos temporais são descartadas por serem, provavelmente, associadas a ruídos.

Para permitir que origens com menos do que 6 *picks* sejam aceitas, o parâmetro *minPhaseCount* foi ajustado para 5, além de ter sido incluído nos processos *hard-coded* de verificação de quantidade mínima de *arrivals*. No *SeisComP* original, mesmo que esse parâmetro seja modificado para um número menor, o valor mínimo de 6 *arrivals* está definido no código-fonte e sobrepõe a modificação do usuário durante o processo.

O parâmetro *amplTypeAbs* define o tipo de amplitude utilizada pelo módulo *scautoloc*. Para que o cálculo de *Amplitude Score* seja realizado corretamente, é necessário utilizar a amplitude MLv. Por fim, o módulo foi configurado para utilizar a tabela de tempos de percurso baseada no modelo de velocidades BRA23 (Subseção 5.4).

Os parâmetros marcados com (*) não estão presentes no *SeisComP* original, sendo propostos pelo presente trabalho. O parâmetro *phaseCountPenalty* habilita a penalização, no *score*, das origens que possuem o número mínimo de chegadas de onda permitido para que a origem seja aceita (*Origin Score*). Por outro lado, o parâmetro *dropAndeanOrigins* faz com que os eventos com origens na região andina (próximos de algum dos 12 pontos mostrados em “Malha de pontos (*grid.conf*)”) não sejam reportados.

5.2.2 Bindings

A Tabela 5.3 apresenta os parâmetros de *bindings* das estações, configurados através do módulo *scconfig*.

Tabela 5.3: Parâmetros de configurações individuais das estações (“*bindings*”), acessíveis através da interface gráfica do *scconfig*.

	Proposto
amplitudes.enableResponses	true
picker.AIC.noiseBegin	-10
picker.AIC.signalBegin	-3
picker.AIC.signalEnd	2
picker.AIC.filter	RMHP(10) BW(4,4.5,10)
picker.AIC.minSNR	3,0
detecFilter*	RMHP(10) BW(4,4.5,10) STALTA(0.2,45)
trigOn*	3,0
trigOff*	0,7
timeCorr*	0

Nos *bindings* de cada estação, buscou-se configurar os parâmetros do *picker* AIC, além de ativar a remoção da resposta de instrumento nos dados, procedimento necessário para que as amplitudes MLv sejam corretamente calculadas.

Os parâmetros marcados com (*) são os mesmos que podem ser definidos para todas as estações através das configurações do módulo *scautopick*, como mostrado anteriormente na Tabela 5.1. Entretanto, pode ser necessário realizar ajustes individuais como, por exemplo, em estações ou redes ruidosas. Durante o desenvolvimento desse trabalho, notou-se a necessidade de aumentar os valores de *trigOn* e *picker.AIC.minSNR* das estações da rede XC, devido ao fato de essas estações produzirem muitos *picks* associados a ruídos. Dessa forma, esses parâmetros foram alterados de 3,0 para 4,5.

5.2.3 Parâmetros de configuração *hard-coded* (*config.cpp*)

A Tabela 5.4 apresenta os parâmetros *hard-coded* do módulo *scautoloc*, modificados diretamente no código-fonte do *software* contido no arquivo *config.cpp*.

Tabela 5.4: Parâmetros *hard-coded* contidos no arquivo *config.cpp*. Os demais parâmetros desse arquivo, não mostrados nessa tabela, são acessíveis ao usuário através do módulo de configuração *scconfig*, não sendo necessário modificá-los no código.

	Original	Proposto
goodRMS	1,5	0,4
dynamicPickThresholdInterval	3600	600
maxResidualKeep	21	3
defaultMaxNucDist	180	35
minScore minScoreTolerance	8	0,75
minScoreBypassNucleator	40	0,97

O parâmetro *dynamicPickThresholdInterval* foi reduzido para 600 segundos, ou seja, o módulo *scautoloc* descartará leituras de uma determinada estação após 10 *picks* de amplitudes similares dentro de 600 segundos. O resíduo máximo permitido para que um *arrival* continue associado a uma origem (*maxResidualKeep*), mesmo sem contribuir para sua localização, foi reduzido de 21 para 3 segundos.

Cabe destacar a renomeação do parâmetro *minScore* para *minScoreTolerance*, uma vez que a pontuação mínima para que uma origem seja aceita passou a ser dinâmica e não mais um valor fixo, como detalhado em “Função *_publishable(origin)*”. O valor mínimo dessa variável dinâmica foi definido como 0,75 considerando que, após as modificações, a pontuação de uma origem varia de 0 a 1.

Originalmente, o parâmetro *minScore* pode ser modificado apenas através das opções de linha de comando (argumento *--min-score*) ou diretamente no arquivo de configuração do módulo *scautoloc*. Entretanto, de forma a facilitar a alteração dessa variável pelo usuário final, as modificações propostas nesse trabalho incluem sua disponibilização através do módulo de configuração *scconfig* (como mostrado na Tabela 5.2). A adição de novos parâmetros ao módulo de configuração *scconfig* pode ser realizada através dos arquivos “*descriptions*” de cada módulo, contidos em “*seiscomp/etc/descriptions*”.

O arquivo *config.cpp* também conta com a definição de outros parâmetros utilizados durante o processo de localização de eventos sísmicos. Entretanto, tais parâmetros (não mostrados na Tabela 5.4) são acessíveis através do módulo de configuração *scconfig* e não precisam ser modificados diretamente no código-fonte, sendo apresentados na Tabela 5.2.

5.3 Malhas de pontos propostas

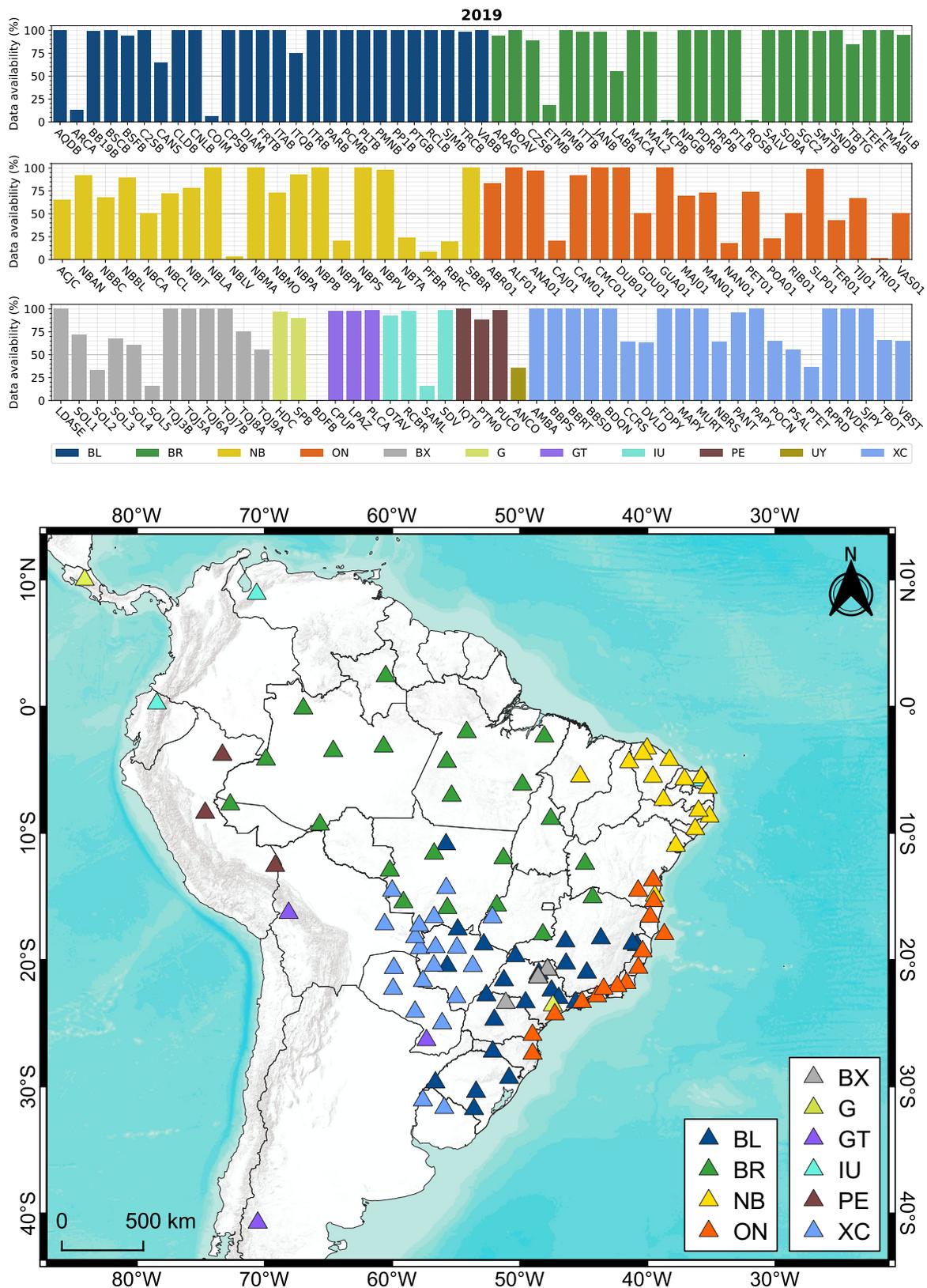


Figura 5.4: Acima: disponibilidade de dados em forma de ondas registrados em 2019 pelas estações contidas no inventário do *SeisComP* do IAG-USP. Abaixo: localizações das estações com no mínimo 50% de dados disponíveis no período.

O processo de criação das malhas de pontos está exemplificado na [Figura 5.4](#): para cada estação, são apresentadas as porcentagens de dados disponíveis no ano de 2019, bem como as localizações geográficas das estações que possuem ao menos 50% de dados registrados no ano, as quais são utilizadas para gerar o *grid* referente a esse período. No mapa da [Figura 5.5](#), apresenta-se a malha de pontos referente ao ano de 2019, com escala de cores baseada na quantidade de estações contidas em circunferências de raio 10° centradas em cada ponto.

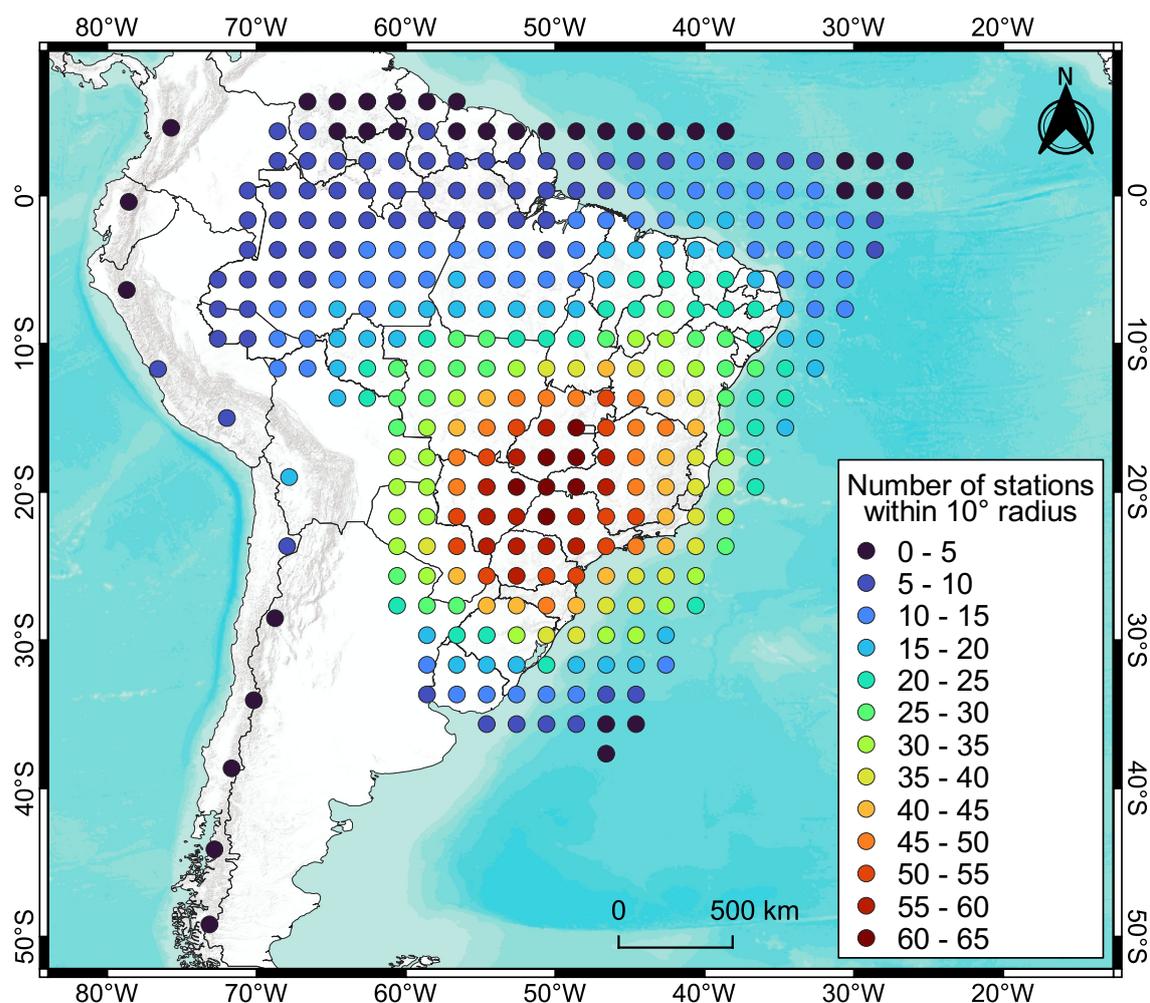


Figura 5.5: Malha de pontos referente ao ano de 2019. A escala de cor refere-se à quantidade de estações em um raio de 10° com ao menos 50% de dados registrados no ano.

Em todos os anos, a Região Sudeste e parte da Centro-Oeste apresentaram as maiores densidades de estações, o que reflete diretamente na quantidade de eventos detectados nessas regiões. A Região Norte do país é a que possui a menor quantidade de estações em operação no período de tempo analisado, seguido pela Região Nordeste. Como detalhado em “[Malha de pontos \(*grid.conf*\)](#)”, a quantidade de estações referente a cada ponto é o fator que determina a distância máxima permitida para uma estação contribuir para uma origem nucleada no ponto e o número mínimo de leituras para que essa origem seja aceita.

As malhas de pontos de cada ano são apresentadas nos [Apêndices](#), bem como a disponibilidade anual de dados de cada estação e suas localizações geográficas.

5.4 Novo modelo de velocidades - BRA23

Durante o processo de nucleação de novas origens, a principal maneira de diminuir falsos eventos é através da utilização de valores consistentes de máximo resíduo temporal permitido para um *arrival* e de máximo resíduo médio da origem, considerando uma sismicidade de escala regional. Os valores de resíduos temporais dependem do modelo de velocidades utilizado durante o processamento, sendo necessária a derivação de um modelo otimizado (BRA23) para reduzir os valores de resíduo e RMS das soluções.

Como apresentado anteriormente, o modelo BRA23 foi obtido através de uma otimização dos mesmos dados utilizados na construção do modelo NewBR, somados a dois eventos regionais mais recentes (Subseção 4.1). No total, foram considerados os dados de 17 eventos, totalizando 183 leituras de ondas P de diferentes regiões do Brasil. A Figura 5.6 apresenta um mapa com a distribuição dos eventos utilizados, bem como as associações com as estações utilizadas para cada evento.

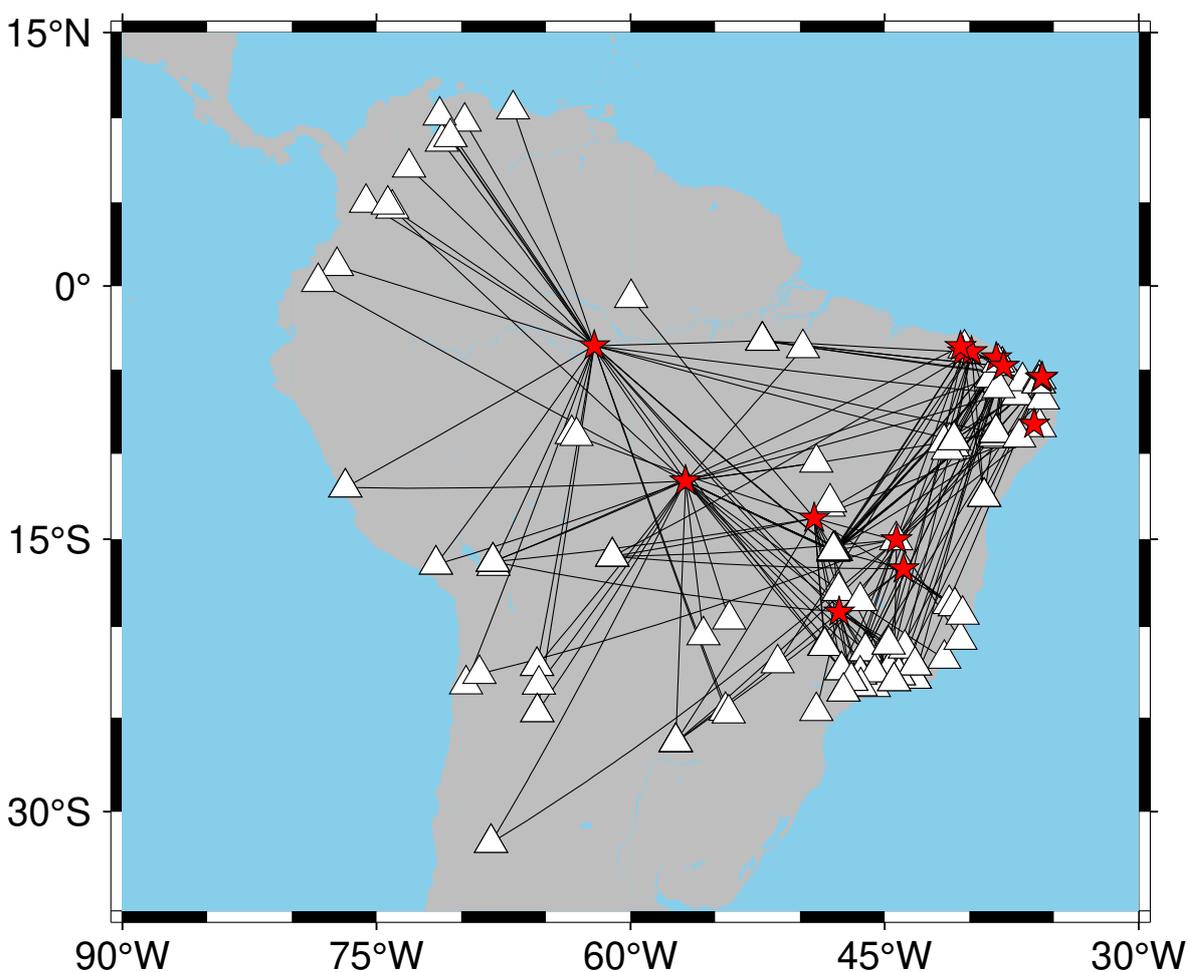


Figura 5.6: Mapa com os 17 eventos regionais (estrelas) e estações (triângulos) utilizadas na construção do modelo BRA23.

A otimização do novo modelo consistiu em minimizar os resíduos dos tempos de percurso para os trajetos mostrados na [Figura 5.6](#). Para cada par evento-estação, a profundidade do evento foi levada em consideração durante o cálculo do tempo de percurso, bem como as espessuras da crosta na região da estação e do evento. As espessuras da crosta utilizadas para corrigir os tempos de percurso foram obtidas de [Rivadeneira et al. \(2019\)](#). Para minimizar o resíduo médio de todos os percursos, os seguintes parâmetros foram otimizados:

- V_{pc} : Velocidade da onda P na crosta inferior;
- V_{pn} : Velocidade da onda P no manto sub-crustal (extrapolada para a LAB);
- V_{pb} : Velocidade da onda P na base do manto;
- **Conrad**: Profundidade da descontinuidade de Conrad, separando a crosta superior (com velocidade fixa de 5,8 km/s) da crosta inferior (com velocidade V_{pc}); e
- **LAB**: Profundidade do limite litosfera-astenosfera que, no modelo BRA23, corresponde à profundidade de influência da velocidade V_{pn} otimizada e na qual a velocidade do manto foi imposta como constante.

Por se tratar de um problema altamente não-linear, e de forma a buscar o mínimo absoluto dos resíduos, optou-se por dividir a otimização em duas etapas. Na primeira etapa, foram realizadas 10 rodadas utilizando um método de otimização global (Evolução Diferencial), sendo esse um método estocástico capaz de buscar, a partir de intervalos de busca ([Tabela 4.1](#)), o mínimo global de uma função. Como resultado desse processo, foram gerados 10 modelos diferentes, apresentados à esquerda da [Figura 5.7](#).

É possível observar que os modelos obtidos são próximos entre si e seguem, de forma aproximada, o modelo NewBR. Os modelos otimizados apresentam uma certa variância tanto nos parâmetros crustais, quanto nos do manto. Em seguida, de forma a refinar a otimização, cada uma das 10 soluções globais foram utilizadas como parâmetros iniciais de uma otimização local, baseada no método de Nelder-Mead. Esse método possui uma convergência para o mínimo local da função objetivo, sendo também sujeito aos limites de busca, e apresenta uma convergência mais rápida do que um algoritmo de busca estocástica para os espaços dos parâmetros desejados.

Diferentemente da otimização global, a otimização local não utilizou todos os dados. Em cada caso de otimização local, foram descartadas as leituras incoerentes com o modelo inicial sendo testado. A quantidade de dados descartados variou entre 6% e 9,2% (11 e 17 leituras do total de 183) nos 10 casos.

Os 10 modelos resultantes da otimização local são apresentados à direita da [Figura 5.7](#). Nota-se que todos os modelos finais obtidos apresentam uma boa convergência para um modelo médio, em especial para os parâmetros V_{pn} , V_{pb} e LAB. Os parâmetros V_{pc} e Conrad apresentam uma maior variabilidade.

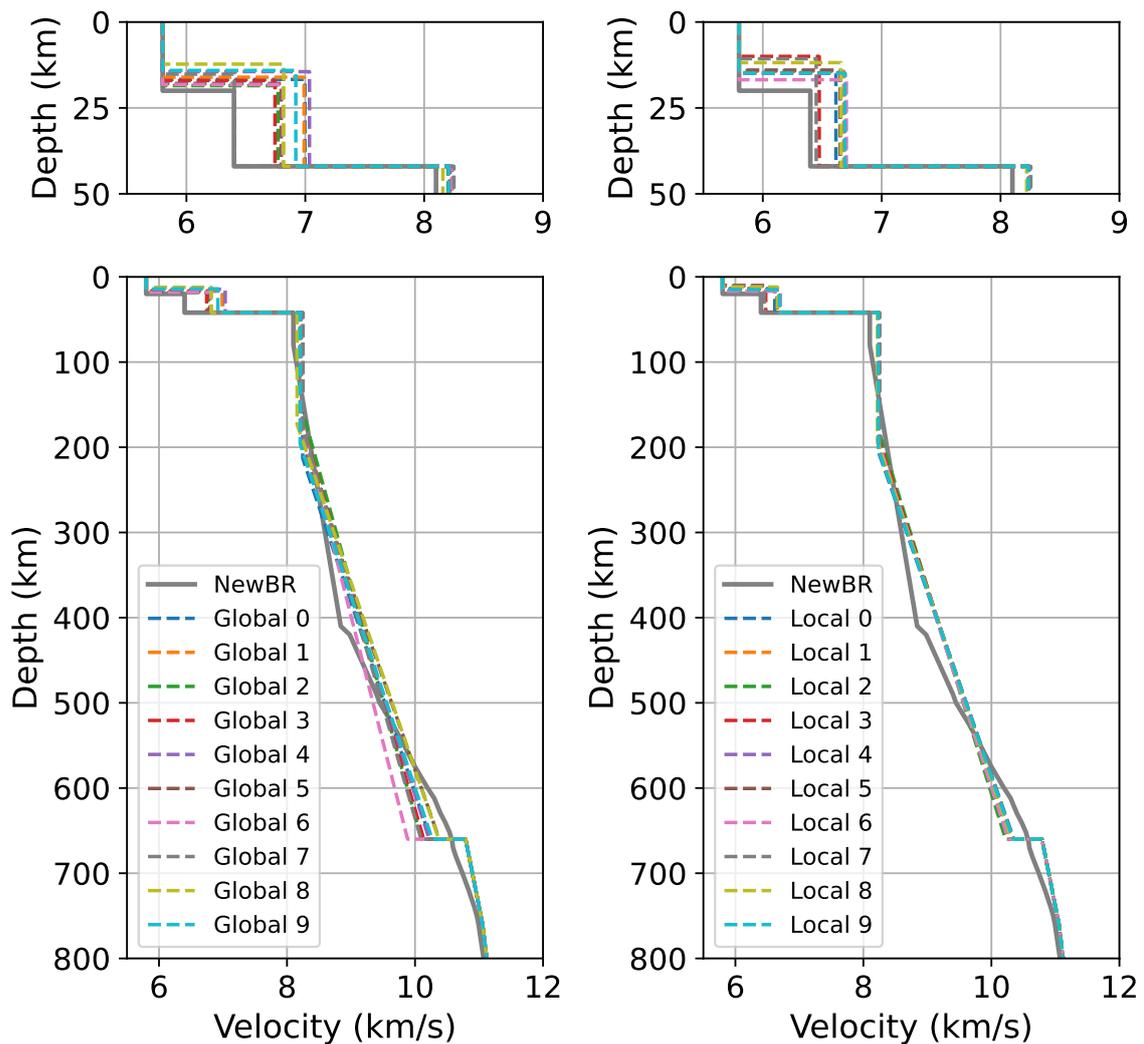


Figura 5.7: Modelos de velocidade de ondas P derivados da inversão global (esquerda) e local (direita). No total, foram realizadas 10 inversões globais e, para cada inversão global, foi derivada uma inversão local. Os painéis superiores apresentam um detalhamento da região crustal. Para comparação, também é apresentado o modelo NewBR original. No geral, os modelos globais mostram uma maior variabilidade no manto quando comparados aos modelos locais, os quais apresentam uma maior convergência dos valores nessa região.

Dada a inviabilidade da realização de testes de mais modelos devido ao tempo total de computação demandado frente ao prazo determinado para a conclusão do presente trabalho, o modelo final adotado foi obtido através da mediana de cada um dos parâmetros estabelecidos pelas 10 otimizações locais. Os valores finais para cada parâmetro, bem como suas incertezas (dadas pelo desvio padrão das soluções), são mostradas na [Tabela 5.5](#).

Tabela 5.5: Solução final dos parâmetros otimizados para o modelo BRA23.

V_{pc}	V_{pn}	V_{pb}	Conrad	LAB
$6,66 \pm 0,09$ km/s	$8,24 \pm 0,01$ km/s	$10,32 \pm 0,04$ km/s	14 ± 2 km	198 ± 8 km

De forma geral, os valores obtidos para todos os parâmetros se mostraram coerentes com o esperado. Um último parâmetro necessário para a construção do modelo é a espessura da crosta, para a qual foi adotado o valor de 40 km, sendo esse o valor médio proposto por Rivadeneyra et al. (2019). Como detalhado anteriormente, durante as otimizações esse valor foi considerado conhecido e variável para cada evento, o que não corresponde à realidade nos casos em que o modelo foi utilizado.

Durante o processo de otimização global, foi possível monitorar os modelos testados e seus respectivos RMS (média quadrática dos resíduos). A Figura 5.8 mostra o histograma, para cada um dos parâmetros, do número de modelos testados, bem como um histograma dos RMS obtidos. O Método de Evolução Diferencial, derivado dos métodos de algoritmo genético (Storn and Price, 1997), tende a concentrar as gerações de modelos a serem testados ao redor das soluções mais atrativas para a função objetiva. Nos gráficos, as barras horizontais vermelhas representam a região correspondente ao valor da mediana mais ou menos um desvio padrão, calculado a partir das 10 soluções das otimizações locais em cada parâmetro. Nota-se que, para todas as variáveis, há uma tendência da solução média ser robusta em relação ao processo de otimização, assim como as próprias incertezas estimadas.

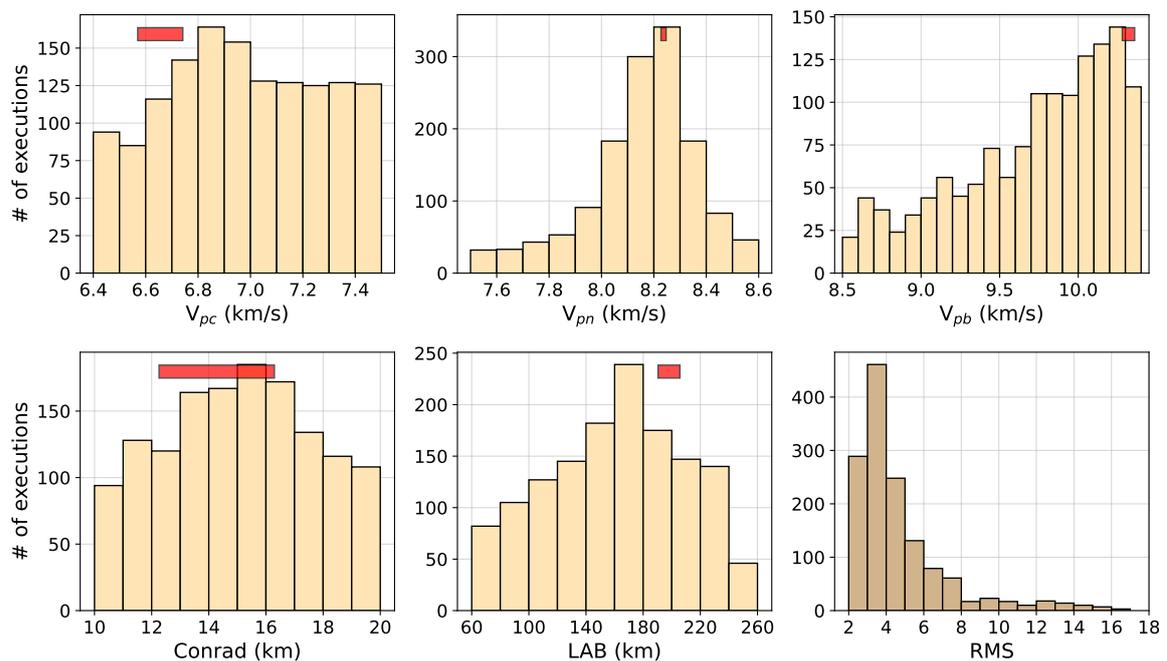


Figura 5.8: Histogramas do número de testes realizados durante as otimizações globais, referentes a cada uma das variáveis otimizadas. Cada barra vermelha indica a região do valor mediano das 10 otimizações locais, somado e subtraído um desvio padrão.

A Figura 5.8 também apresenta o histograma do RMS de todas as soluções obtidas através da otimização global, a qual foi capaz de encontrar soluções com um RMS mínimo de 2,5 segundos. Embora esse valor seja melhor do que o valor do RMS do modelo NewBR (3,39 s), ainda assim foi maior do que a média dos RMS obtidos pelas otimizações locais (1,34 s), os quais variaram entre 1,26 s e 1,48 s.

Para uma validação final do modelo BRA23 obtido, a [Figura 5.9](#) apresenta o gráfico dos resíduos temporais de todas leituras utilizadas para derivar o modelo, comparando-as com os resíduos considerando o modelo NewBR. De forma geral, os resíduos de tempo de percurso para os diferentes eventos apresentam uma média próxima de zero, fato demonstrado no histograma dos resíduos individuais apresentados na mesma figura. O modelo de velocidades BRA23 apresenta um valor de RMS final de 2,462 s, frente ao valor de RMS de 3,390 s do modelo NewBR.

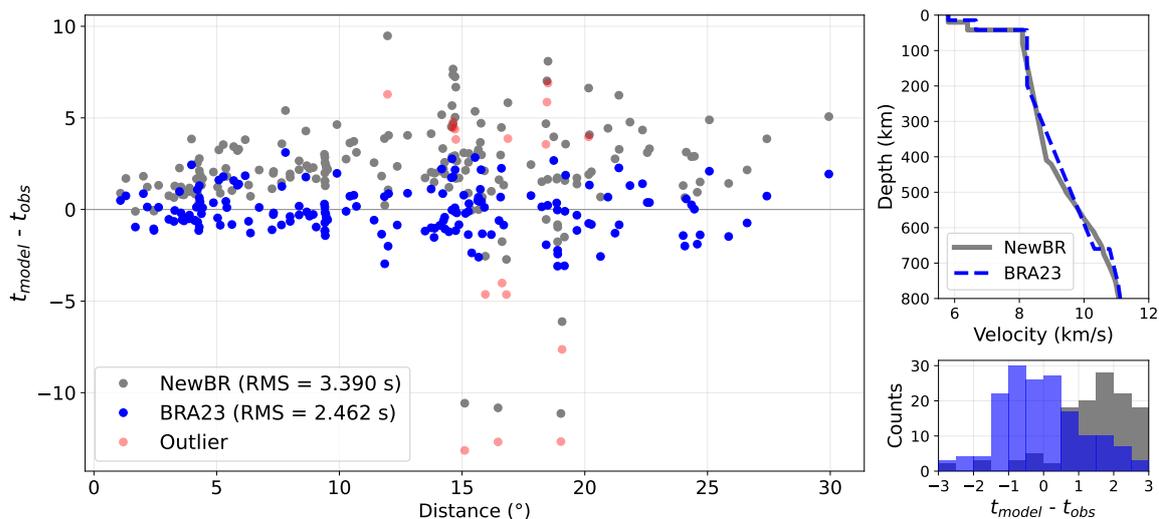


Figura 5.9: Comparação dos resíduos temporais ($t_{model} - t_{obs}$) utilizando-se os modelos BRA23 e NewBR. Nota-se que o novo modelo proposto tende a apresentar uma média próxima de zero, diferentemente do modelo NewBR, o qual apresenta uma tendência de resíduos positivos. Os pontos em vermelho correspondem a *outliers*, não utilizados durante os processos de otimização local.

A dispersão dos pontos é explicada pela grande variabilidade de dados, os quais correspondem a eventos de diferentes regiões do Brasil. Por ser um país continental, o ajuste de todos os dados através de um único modelo 1D resulta em uma maior dispersão, mas viabiliza sua utilização no processo de nucleação de novas origens No *SeisComp*.

Por outro lado, mesmo sendo um modelo baseado apenas em tempos de percurso, obteve-se uma compatibilidade entre a espessura da litosfera do modelo e os resultados de tomografia no Brasil (Ciardelli et al., 2022; Heit et al., 2007). As velocidades da crosta inferior e manto superior, bem como a espessura do limite crosta superior/inferior, mostraram-se compatíveis com os resultados de Mooney et al. (1998), sendo exatamente iguais aos valores propostos pelo autor para os terrenos arqueanos.

O modelo de velocidades 1D BRA23, proposto nesse trabalho, está contido nos [Apêndices](#). A tabela de tempos de percurso utilizada pelo *SeisComp* pode ser obtida através do pacote *TauP* (Crotwell et al., 1999).

5.5 Detecções localizáveis (2014 - 2021)

Uma “detecção localizável” corresponde a um evento automático gerado pelo *software*, podendo esse ser verdadeiro ou falso (como descrito em “Lista de Siglas e Termos”). Trata-se de um conjunto de *picks* que apresentaram parâmetros aceitáveis, os quais passaram por todos os testes e checagens realizadas para que um evento fosse criado.

Cada detecção localizável foi inspecionada visualmente, de forma qualitativa, para definir se a mesma possui bons *picks* (*good picking*), *picks* mal posicionados, mas que ainda assim referem-se a uma chegada de onda de um evento sísmico real (*poor picking*), ou *picks* relacionados a ruídos, sem relação com eventos sísmicos (*noise picking*).

A Figura 5.10 apresenta as detecções localizáveis obtidas através do *SeisComP* original, utilizando-se o banco de dados de 2014 a 2021.

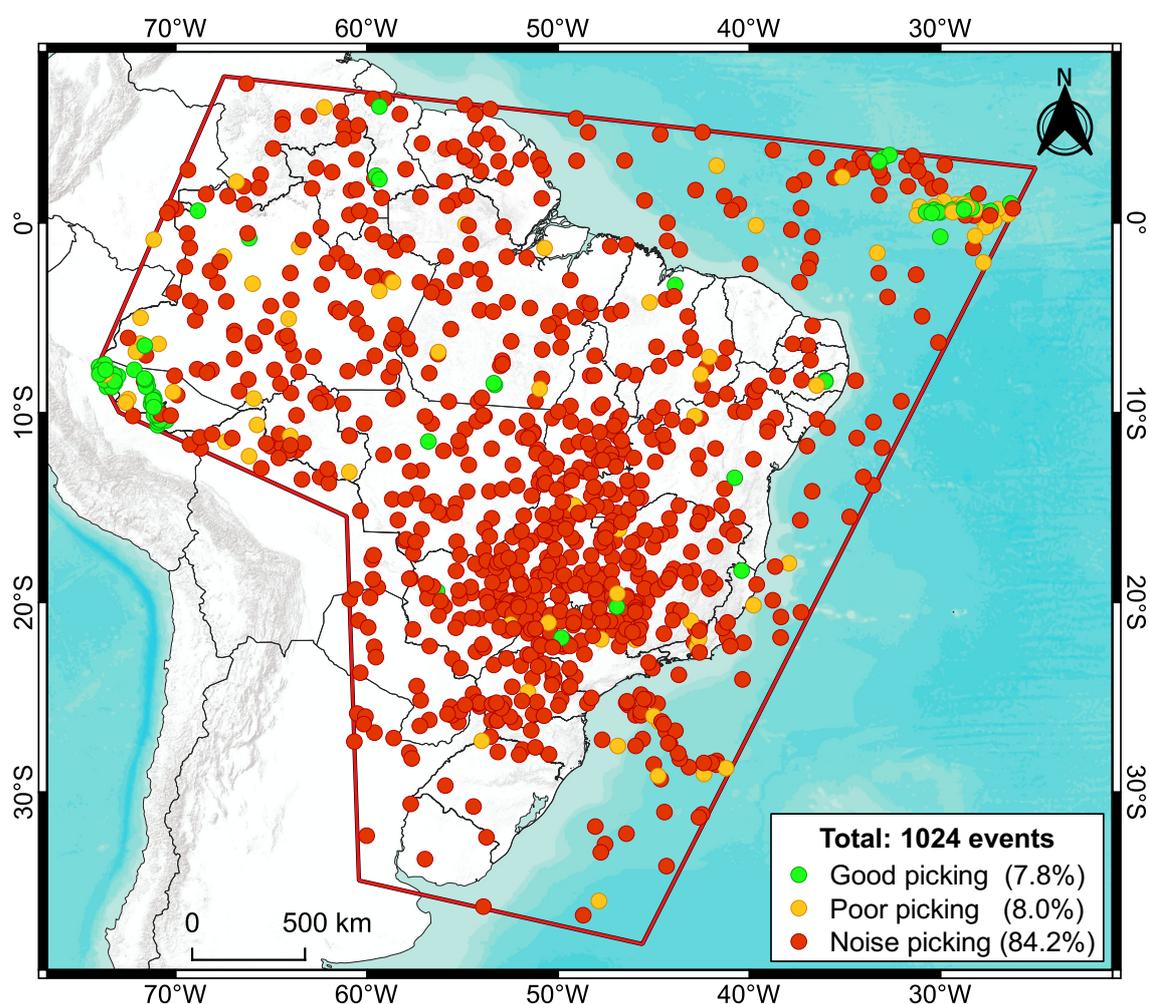


Figura 5.10: Detecções localizáveis obtidas utilizando-se o *SeisComP* original.

Do total de 1024 detecções localizáveis no período, apenas 80 (7,8%) foram avaliadas como eventos sísmicos satisfatoriamente localizados, sendo esses majoritariamente associados a eventos com origens em bordas de placas (sismos profundos no estado do Acre e sismicidade na dorsal mesoatlântica), além dos eventos com maiores magnitudes em

território brasileiro. Também destaca-se o fato de que 862 detecções (84,2%) foram consideradas falsas, associadas à nucleação de ruídos locais.

Para avaliar a performance do *SeisComP* modificado, a Figura 5.11 apresenta as detecções localizáveis obtidas através do *software* com as modificações propostas, utilizando-se o mesmo banco de dados que gerou a figura anterior.

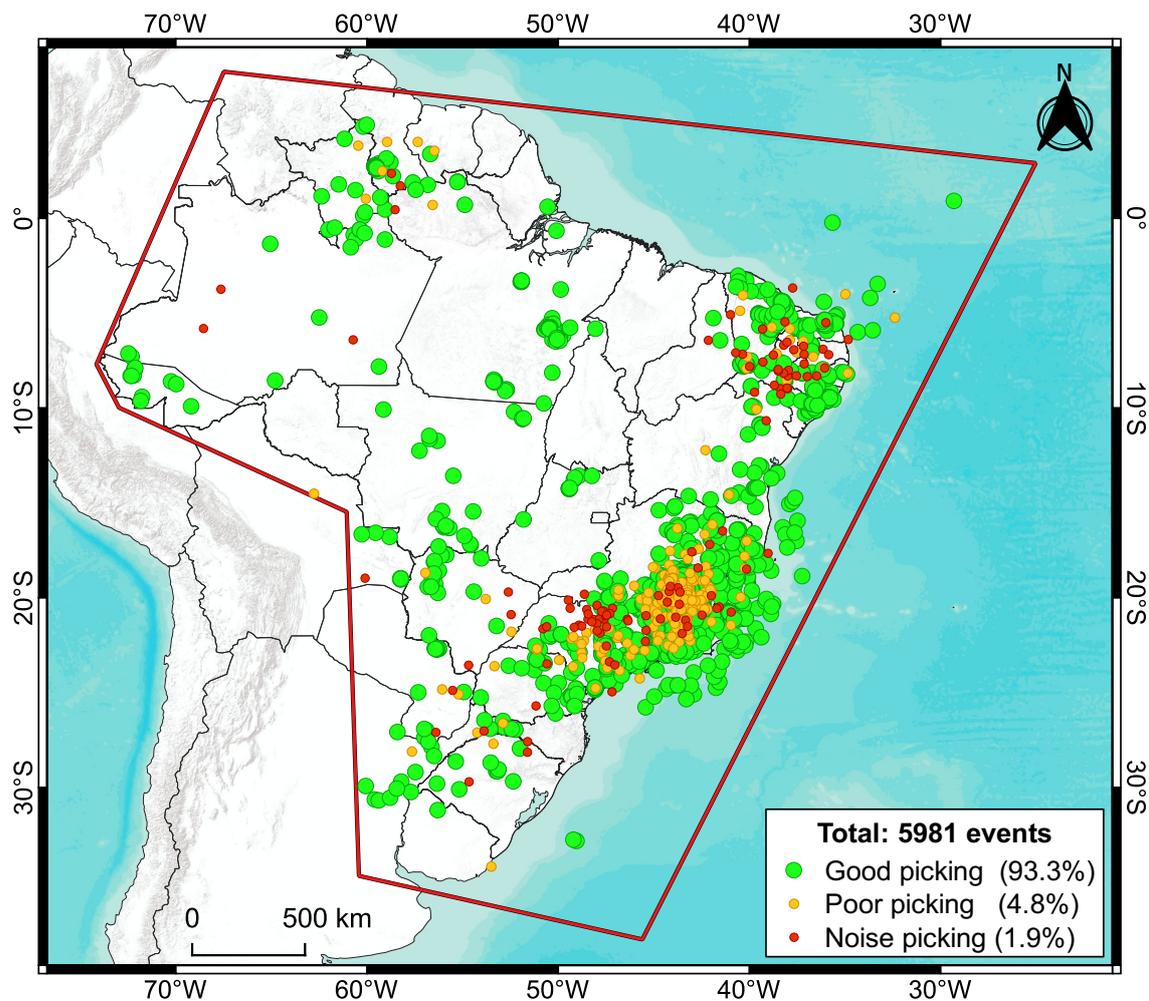


Figura 5.11: Detecções localizáveis obtidas utilizando-se o *SeisComP* modificado.

Do total de 5981 detecções localizáveis no período de 2014 a 2021 utilizando-se o *SeisComP* modificado, 5580 (93,3%) foram consideradas eventos reais, com boas leituras de chegadas de onda e, conseqüentemente, boas localizações epicentrais (“*good picking*”).

Um total de 289 detecções (4,8%) foram categorizadas como “*poor picking*”, *i.e.*, eventos reais, mas com alguns *picks* incorretos que possivelmente afetaram negativamente as localizações. Finalmente, 112 detecções (1,9%) foram associadas a falsos-positivos (“*noise picking*”), ou seja, não estavam relacionadas a eventos sísmicos reais e que contêm *picks* associados a ruídos locais de cada estação.

5.6 Eventos concomitantes ao catálogo da RSBR

Para verificar a concordância das detecções localizáveis obtidas em ambos os cenários (*SeisComP* modificado e original) com os eventos contidos no catálogo sísmico da RSBR, buscou-se encontrar os eventos concomitantes ao catálogo seguindo critérios baseados nos horários de origem e coordenadas geográficas. Foram considerados eventos concomitantes ao catálogo da RSBR aqueles que apresentam até 15 segundos de diferença nos horários de origem e distância máxima de 100 km entre os epicentros.

A Figura 5.12 apresenta o mapa com as detecções localizáveis obtidas através do *SeisComP* original e que são concomitantes a eventos no catálogo sísmico da RSBR.

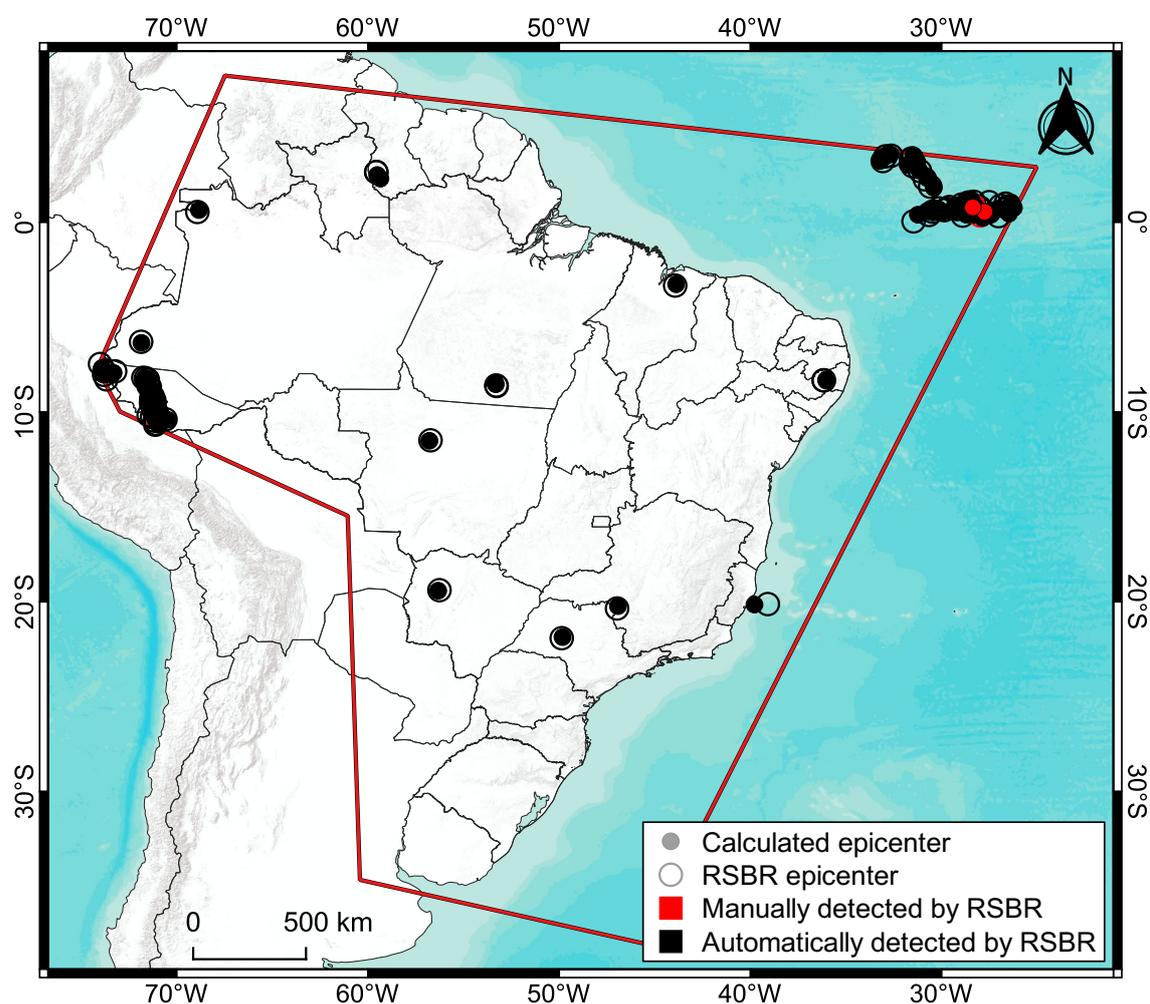


Figura 5.12: Epicentros das 78 detecções localizáveis concomitantes ao catálogo sísmico da RSBR de 2014 a 2021, utilizando-se o *SeisComP* original.

Como esperado, o *SeisComP* original foi capaz de detectar, em sua grande maioria, apenas eventos sísmicos que também foram detectados automaticamente pela RSBR. Ao todo, foram 78 eventos concomitantes ao catálogo da RSBR, sendo que apenas 2 foram originalmente localizados de forma manual. A maior parte desses eventos são de origem tectônica, como os sismos profundos no Acre e os na dorsal mesoatlântica. Alguns eventos de maior magnitude em território brasileiro também foram detectados de forma automática.

Na seção **Motivação**, foi citado que o catálogo da RSBR possui 106 eventos automáticos, número maior do que o mostrado na **Figura 5.12**. Isso se deve ao fato de que, por um determinado período, o *SeisComP* utilizado no Centro de Sismologia do IAG-USP contou com dois *pickers* operando simultaneamente, sendo um deles com parâmetros de filtros alterados pelos mantenedores do sistema. Dessa forma, foi possível detectar alguns eventos a mais do que seria detectado apenas com o *picker* original.

A **Figura 5.13** apresenta o mapa com as detecções localizáveis obtidas através do *SeisComP* modificado e que são concomitantes a eventos no catálogo sísmico da RSBR.

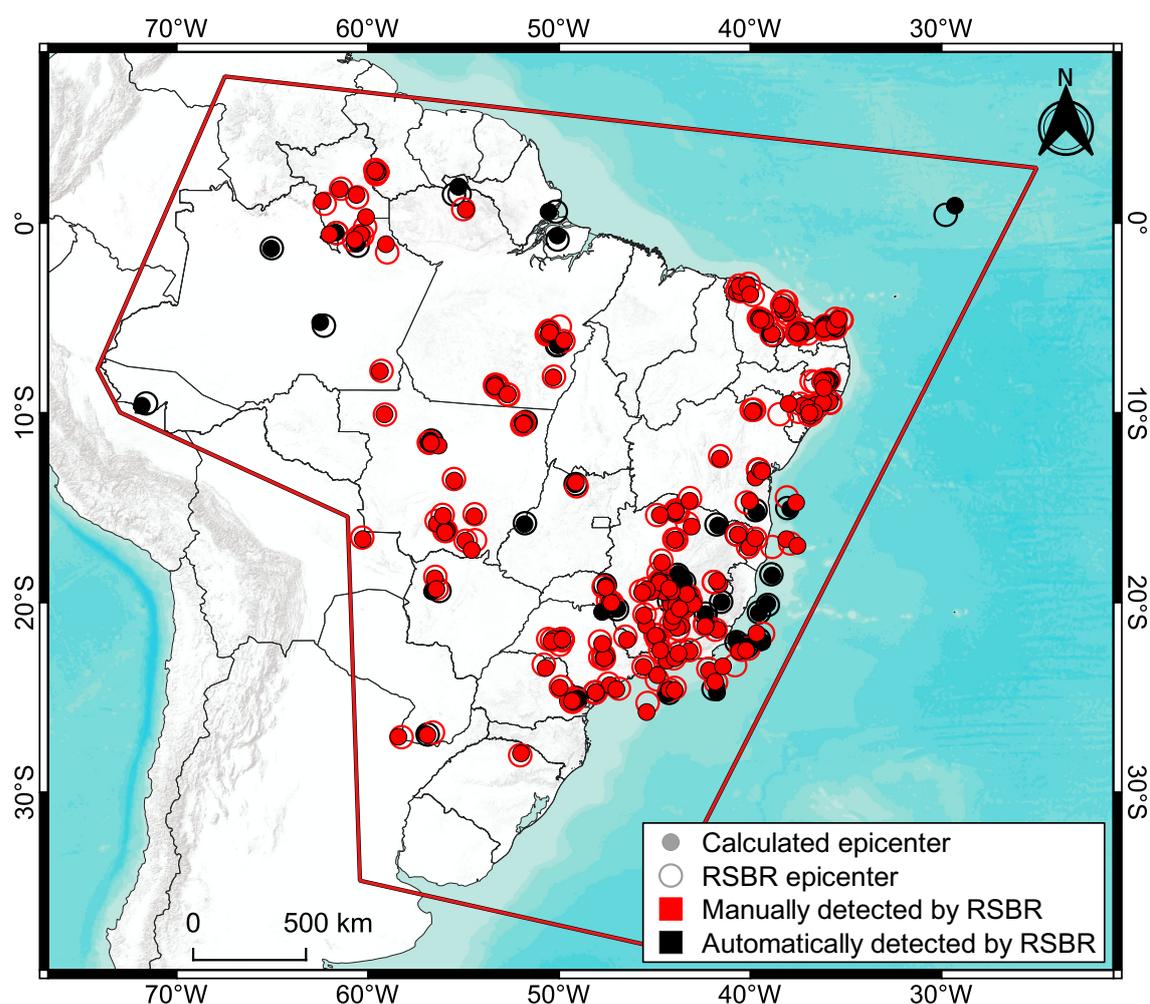


Figura 5.13: Epicentros das 292 detecções localizáveis concomitantes ao catálogo sísmico da RSBR de 2014 a 2021, utilizando-se o *SeisComP* modificado.

No total, somam-se 292 eventos sísmicos concomitantes ao catálogo, sendo que 65 foram detectados automaticamente pela RSBR e 227 foram localizados apenas após análise manual dos dados registrados. Destacam-se os eventos na Região Nordeste que, em sua maioria, foram detectados de forma automática apenas após as modificações propostas. Também há um aumento significativo de detecções automáticas nas Regiões Sudeste e Centro-Oeste, além dos eventos no Norte do país e os recentes eventos sísmicos ocorridos na Guiana.

6. Discussões

6.1 Detecções localizáveis

Com as modificações propostas, foi possível observar um aumento significativo no número de detecções localizáveis durante o período de 2014 a 2021, bem como de eventos concomitantes ao catálogo sísmico da RSBR. A Figura 6.1 apresenta uma comparação das porcentagens de cada classificação qualitativa das detecções localizáveis, considerando o *SeisComP* modificado e o original.

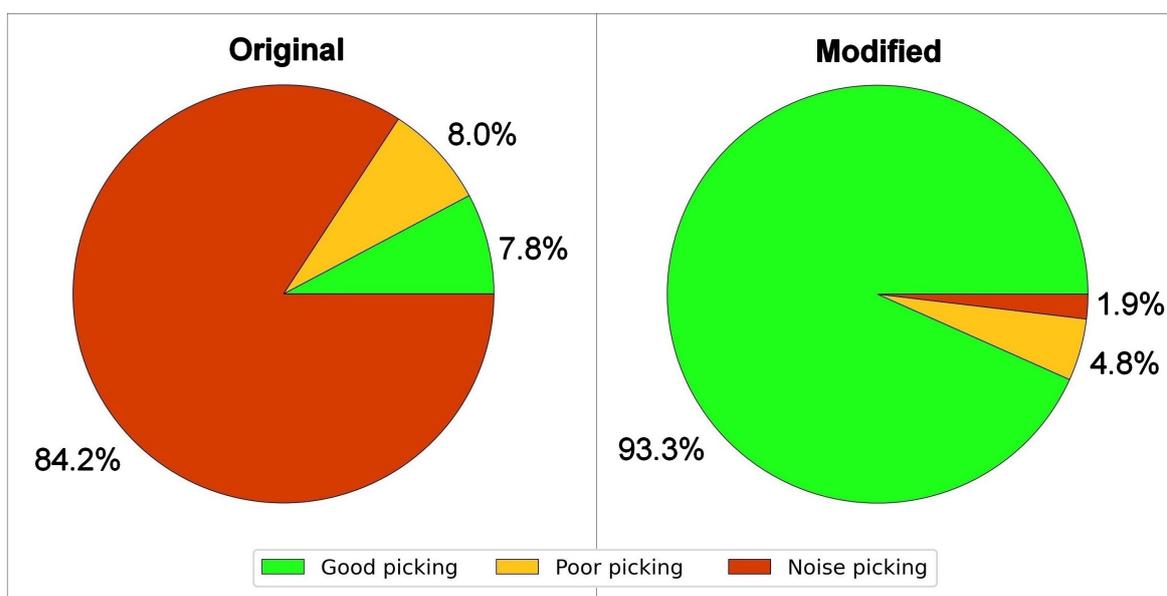


Figura 6.1: Gráficos de setores das porcentagens de detecções localizáveis classificadas, ocorridas no período de 2014 a 2021. À esquerda, as porcentagens referentes às 5981 detecções obtidas através do *SeisComP* modificado. À direita, as porcentagens referentes às 1024 detecções obtidas através do *SeisComP* original.

Além do aumento expressivo da quantidade de detecções localizáveis, destaca-se também o fato de que as modificações propostas aumentaram a quantidade de eventos classificados positivamente. Em números totais, as detecções classificadas como “*good picking*” aumentaram de 80 para 5580, enquanto as detecções classificadas como “*noise picking*” diminuíram de 862 para 112.

Com a diminuição da porcentagem de falsos-positivos, a média anual de eventos errôneos diminuiu de 108 para 14, o que pode facilitar o trabalho dos operadores responsáveis por verificar os eventos localizados de forma automática. Da mesma forma, o aumento de eventos reais localizados de forma satisfatória ou até mesmo os que possuem *picks* incorretos (mas ainda assim associados a eventos reais) proporciona uma melhor efetividade no trabalho de validação manual dos eventos.

6.2 Eventos concomitantes ao catálogo da RSBR

Acerca dos eventos concomitantes ao catálogo da RSBR, as modificações possibilitaram um aumento de 78 para 292 eventos localizados automaticamente e que também estão contidos no catálogo. Como mencionado anteriormente, os critérios para definir a concomitância de uma detecção localizável com um evento contido no catálogo sísmico da RSBR foram suas coordenadas geográficas e seu horário de origem. Definiu-se que seriam considerados eventos concomitantes aqueles que apresentassem uma distância de até 100 km em relação aos epicentros dos eventos do catálogo, bem como uma diferença de até 15 segundos nos horários de origem.

Para verificar se esses critérios são estáveis, ou seja, se não há variação significativa na quantidade de eventos concomitantes ao se variar tais premissas, a Figura 6.2 apresenta um gráfico indicando a quantidade de eventos concomitantes ao catálogo de acordo com a variação de ambos os critérios.

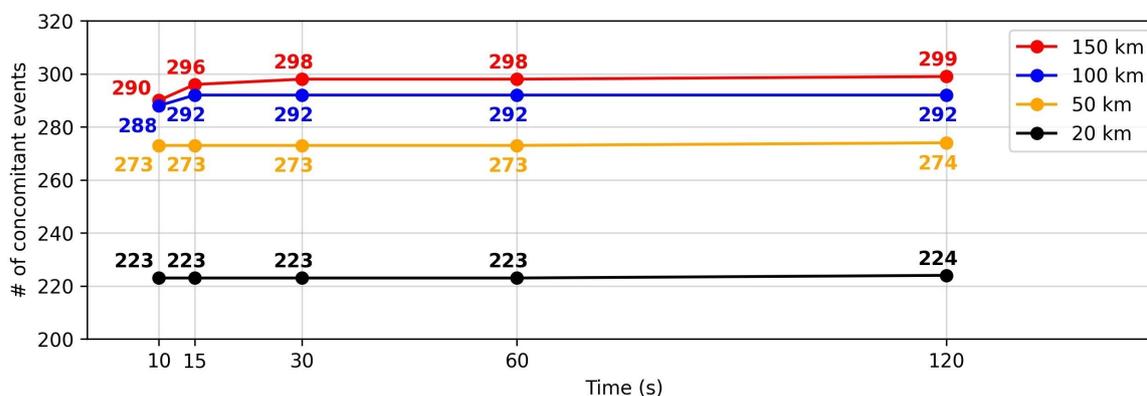


Figura 6.2: Número de eventos concomitantes ao catálogo da RSBR após as modificações no *SeisComP*, considerando diferentes combinações de critérios de concomitância.

Nota-se que a variação da diferença máxima permitida entre os horários de origem não apresenta mudanças significativas na quantidade de eventos concomitantes, enquanto a variação da distância máxima permitida entre os epicentros possui maior impacto na quantidade de eventos. Portanto, torna-se válida a definição da premissa de que uma detecção localizável é concomitante ao catálogo se houver uma distância máxima de 100 km entre seus epicentros e uma diferença de até 15 segundos em seus horários de origem.

A concomitância dos eventos foi considerada utilizando-se apenas os resultados obtidos de forma automática, não havendo quaisquer alterações manuais nas origens após suas localizações. Considerando que *picks* mal posicionados em eventos reais podem influenciar significativamente ambos os parâmetros considerados (coordenadas geográficas e horário de origem), entende-se que, após uma avaliação manual das origens localizadas de forma automática, o número de eventos concomitantes ao catálogo deve aumentar.

Para verificar alguns dos parâmetros dos eventos concomitantes ao catálogo, as figuras a seguir apresentam comparações entre os 292 eventos concomitantes obtidos através do *SeisComP* modificado e os 78 eventos concomitantes localizados com o *SeisComP* original.

A Figura 6.3 mostra que as modificações no *SeisComP* proporcionaram a detecção de eventos sísmicos de magnitudes mínimas de até M1,5, sendo que a maioria dos eventos concomitantes ao catálogo apresentam magnitudes entre M2,0 e M3,5. Sem as modificações, o *software* foi capaz de detectar apenas sismos de magnitudes maiores do que M3,5.

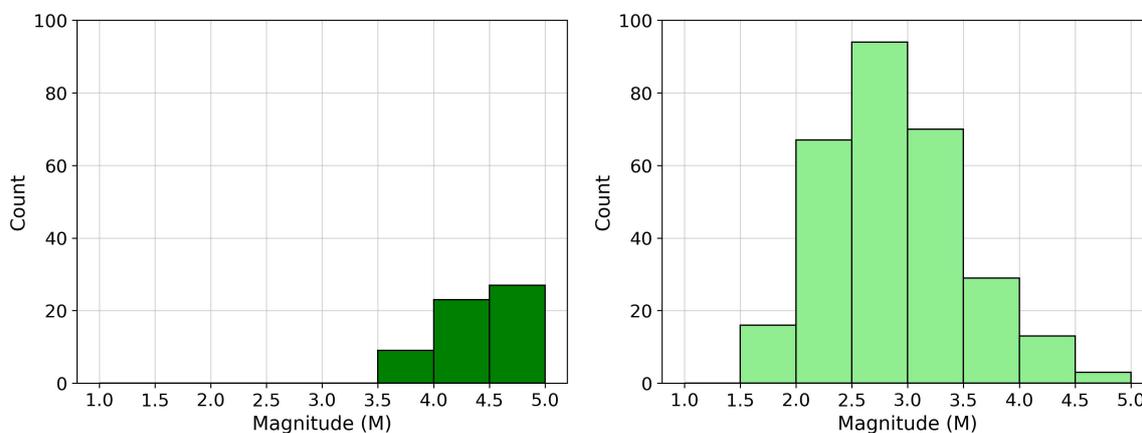


Figura 6.3: Histogramas de magnitudes dos eventos concomitantes ao catálogo da RSBR, obtidos através do *SeisComP* original (esquerda) e modificado (direita).

A Figura 6.4 e a Figura 6.5 apresentam as distribuições dos RMS das origens e dos resíduos temporais de suas chegadas de onda, respectivamente.

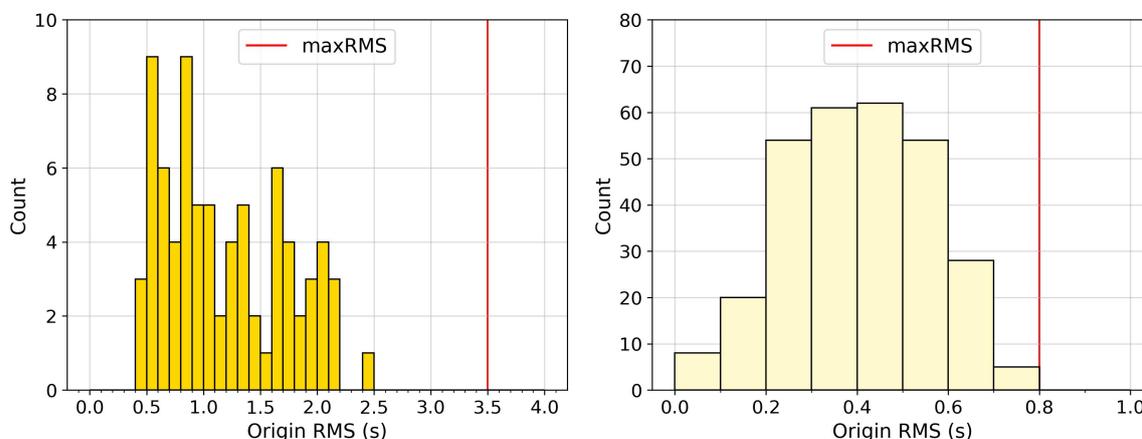


Figura 6.4: Histogramas de RMS das origens dos eventos concomitantes ao catálogo da RSBR, obtidos através do *SeisComP* original (esquerda) e modificado (direita). As barras possuem 0,1 s de largura em ambos os gráficos.

Nota-se, na Figura 6.4, que o *SeisComP* modificado gerou origens concomitantes ao catálogo com RMS entre 0 e 0,8 segundos, enquanto o *SeisComP* original localizou origens concomitantes com RMS entre 0,4 e 2,5 segundos. Com relação aos resíduos temporais dos *arrivals* (Figura 6.5), o mesmo padrão se repete: após as modificações, os *arrivals* apresentaram, majoritariamente, resíduos absolutos abaixo de 1 segundo, enquanto o *SeisComP* sem modificações apresentou diversas chegadas de ondas com resíduos de alguns segundos. A melhoria nos resíduos dos *arrivals* converge com os resultados obtidos durante o desenvolvimento do modelo BRA23, como mostrado anteriormente na Figura 5.9.

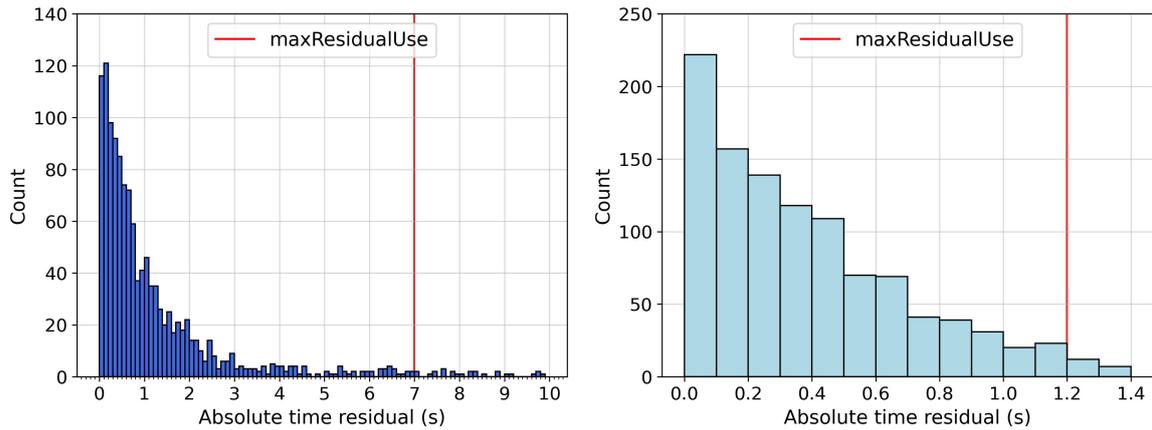


Figura 6.5: Histogramas de resíduos temporais absolutos dos *arrivals* dos eventos concomitantes ao catálogo da RSBR, obtidos através do *SeisComP* original (esquerda) e modificado (direita). As barras possuem 0,1 s de largura em ambos os gráficos.

A alteração do modelo de velocidades empregado para gerar a tabela de tempos de percurso utilizada pelo *SeisComP* e as malhas de pontos propostas foram fatores que contribuíram para a diminuição dos resíduos temporais dos *arrivals* e, conseqüentemente, dos RMS das origens localizadas automaticamente. A maior acurácia nas leituras das chegadas de onda, em consequência das alterações realizadas no módulo *scoutpick*, também tiveram um papel significativo na diminuição desses resíduos.

As linhas verticais vermelhas das Figuras 6.4 e 6.5 correspondem, respectivamente, aos parâmetros *maxRMS* (resíduo máximo para que uma origem seja aceita) e *maxResidualUse* (resíduo máximo de um *arrival* para que possa contribuir na localização de uma origem), os quais foram consideravelmente reduzidos após as modificações.

Por fim, a Figura 6.6 apresenta a distribuição de números de *arrivals* utilizados na localização de cada origem. Com as modificações propostas, foi possível localizar eventos concomitantes ao catálogo com apenas 5 chegadas de ondas, algo que antes não era possível por limitações definidas no código-fonte do *software*.

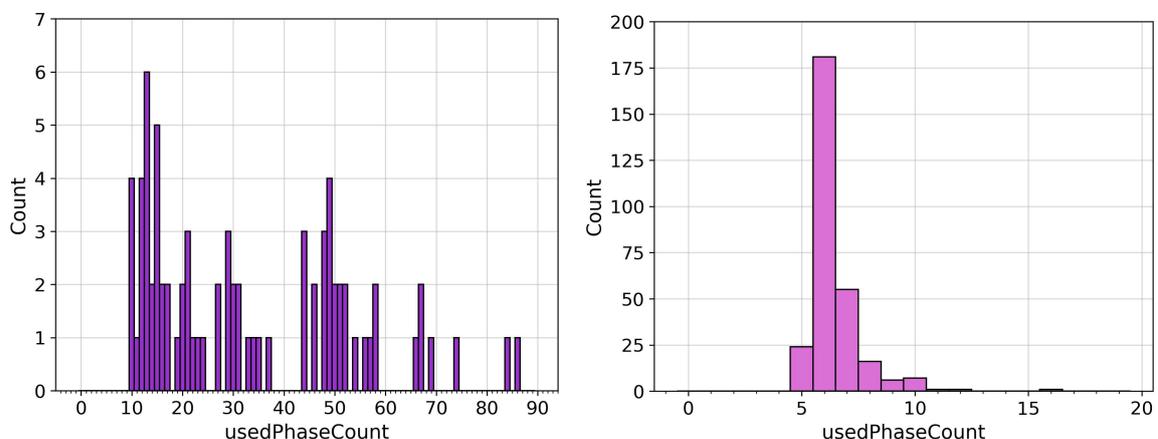


Figura 6.6: Histogramas de números de *arrivals* utilizados (*usedPhaseCount*) nas localizações dos eventos concomitantes ao catálogo da RSBR, obtidos através do *SeisComP* original (esquerda) e modificado (direita).

6.3 Estimativas de detectabilidade automática após as modificações

Combinando as medianas das distâncias máximas em que os eventos sísmicos na área de interesse foram registrados (Figura 2.3) com a densidade de estações sismográficas no Brasil no ano de 2021 (Apêndices), foi possível estimar a atual detectabilidade automática de eventos sísmicos, com base em suas magnitudes, na região de interesse (Figura 6.7).

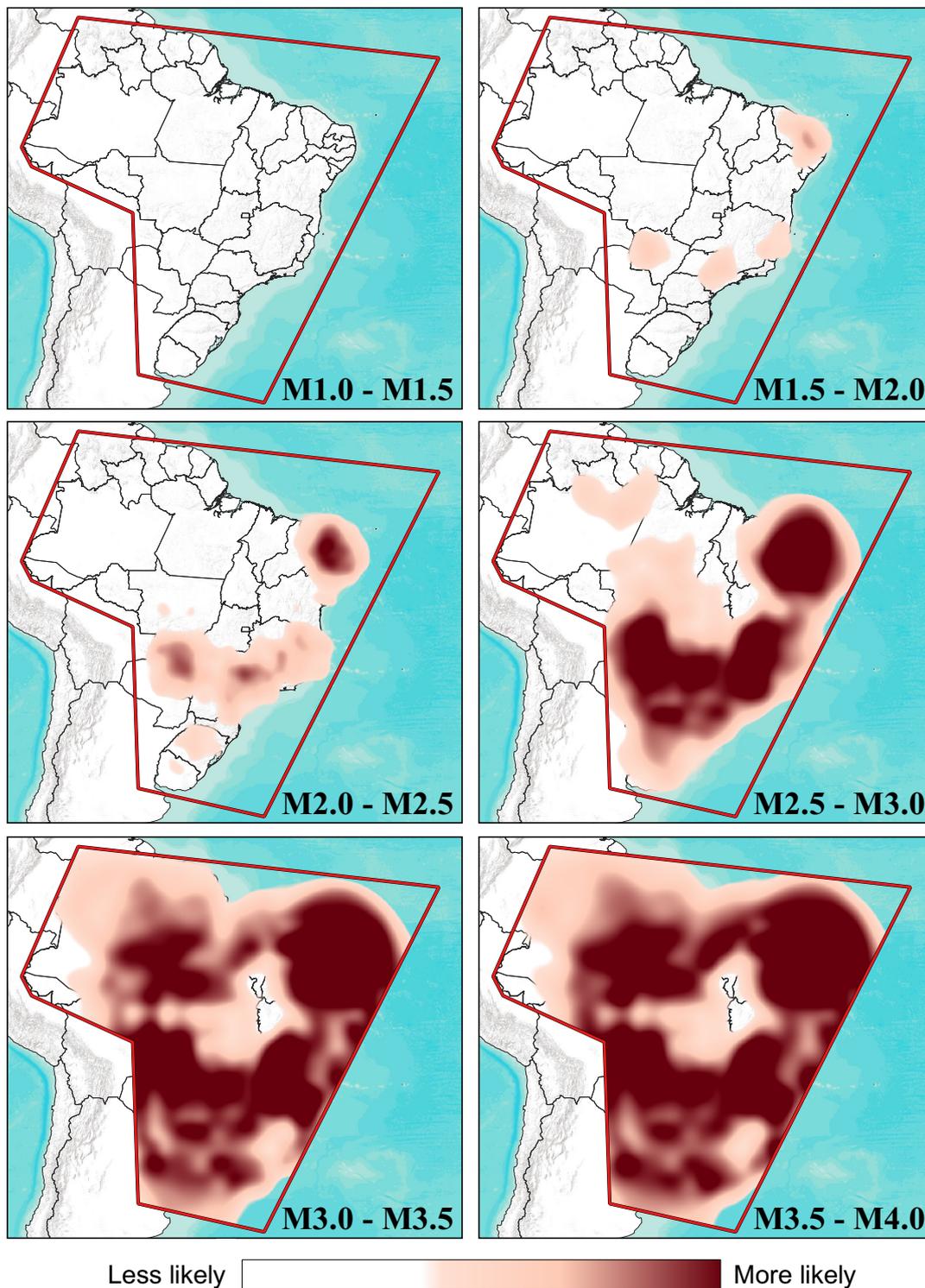


Figura 6.7: Estimativas de detectabilidade automática de eventos sísmicos no Brasil. A escala qualitativa de cores indica a probabilidade de detecção automática.

Para verificar a coerência entre as estimativas de detectabilidade automática e os resultados obtidos, a **Figura 6.8** apresenta os mesmos mapas da figura anterior, mas com a sobreposição dos eventos de 2014 a 2021 concomitantes ao catálogo da RSBR, detectados após as modificações no *SeisComP*. As distribuições dos eventos concomitantes concordam com as regiões em que foram estimadas as maiores probabilidades de detecção.

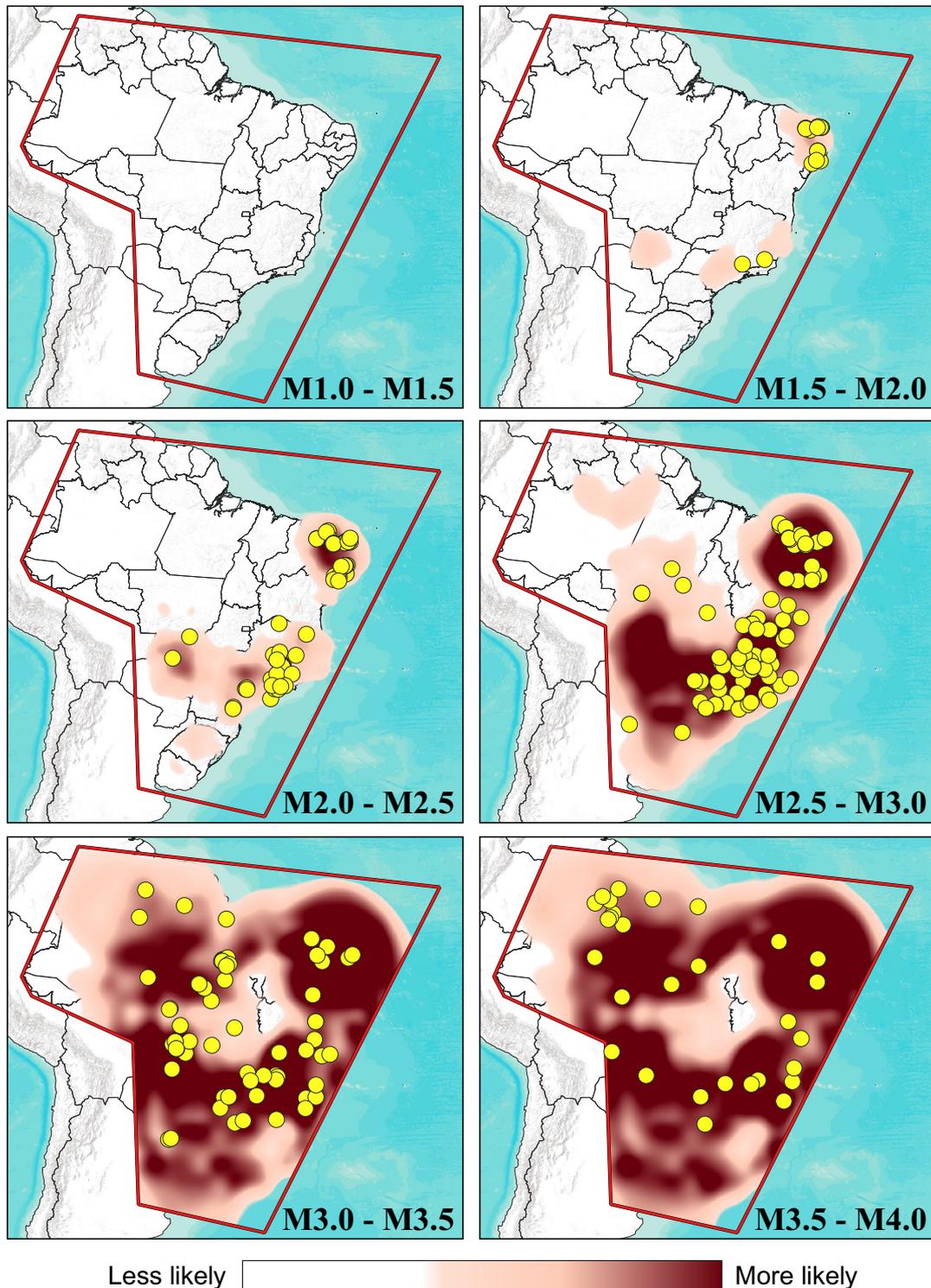


Figura 6.8: Estimativas de detectabilidade automática de eventos sísmicos no Brasil, com sobreposição dos eventos concomitantes ao catálogo da RSBR obtidos após as modificações no *SeisComP*. A escala qualitativa de cores indica a probabilidade de detecção automática.

Os mapas de estimativas de detectabilidade automática foram obtidos através de algumas etapas de processamento. Na primeira, a região de interesse foi discretizada em pontos espaçados de $0,5^\circ$ entre si. Considerou-se que há duas “distâncias limitantes” para que um evento seja detectado por uma estação, sendo elas: (i) a própria distância epicentral e (ii) a máxima distância permitida pela malha de pontos do *SeisComp*, como discutido em “Malha de pontos (*grid.conf*)”. Dessa forma, para cada *range* de magnitude, foram traçadas circunferências centradas nesses pontos, com raios iguais às medianas das máximas distâncias registradas contidas no catálogo da RSBR (Figura 2.3) ou iguais à distância máxima permitida pelo ponto da malha mais próximo (d_{max}), o que for menor (limitante).

Em seguida, calculou-se a quantidade de estações contidas em cada circunferência, associando esse valor ao seu ponto central. Em outras palavras, no caso em que um evento ocorra em um determinado ponto, buscou-se calcular quantas estações seriam capazes de detectá-lo, considerando as limitações de sua distância epicentral e da máxima distância imposta pelo *SeisComp* para esse ponto. Por fim, os mapas foram obtidos através de *krigagens* utilizando-se as quantidades de estações associadas a cada ponto. A escala qualitativa de cores dos mapas não associa nenhuma cor a regiões com contagens menores do que 5 estações distintas (mínimo para que um evento seja detectado automaticamente).

Cabe reforçar que os mapas de estimativas de detectabilidade automática de eventos sísmicos no Brasil foram obtidos utilizando-se o arranjo de estações do ano de 2021. Considerando que a base de dados utilizada nesse trabalho abrange o período de 2014 a 2021, não era esperado que a comparação entre o que foi detectado automaticamente e as estimativas obtidas fossem totalmente aderentes, uma vez que a densidade de estações variou consideravelmente ao longo desse período (vide mapas nos Apêndices).

Por exemplo, o mapa de detectabilidade estimada para o *range* de magnitudes entre 2,5 e 3,0 apresenta uma maior probabilidade de detecção na Região Centro-Oeste, principalmente no estado do Mato Grosso do Sul. Entretanto, após as modificações, não foram registrados eventos sísmicos dessas magnitudes nessa região (Figura 6.8), possivelmente devido ao fato de que a sub-rede temporária XC (responsável por aumentar a densidade de estações nessa região) não esteve em operação durante todo o período de 2014 a 2021. Outro fator que pode ter impactado na eficiência de detecção nessa região foi o aumento do valor de *triggerOn* das estações ruidosas da sub-rede XC, como discutido em “*Bindings*”.

Também deve-se considerar o fato de que os mapas de detectabilidade automática não levam em consideração as taxas sísmicas de cada região do território brasileiro. Desses mapas, apenas inferem-se as regiões nas quais, caso um evento sísmico ocorra, há maiores possibilidades de detecção automática ao se utilizar o *SeisComp* modificado, com base na quantidade de estações contidas em “distâncias limitantes” a partir de pontos discretizados.

Apesar dos fatos mencionados, a coerência entre os mapas de estimativas e os epicentros dos 292 eventos concomitantes ao catálogo (obtidos após as modificações) indica que tais estimativas são compatíveis com a nova realidade de detecção do *SeisComp*.

6.4 Eventos do catálogo não localizados automaticamente

Apesar do aumento significativo das detecções localizáveis após as modificações propostas, não era esperado que todos os eventos sísmicos contidos no catálogo da RSBR fossem localizados automaticamente. O principal motivo é o fato de que muitos eventos possuem chegadas de onda emergentes que dificilmente são diferenciadas do ruído local, mesmo após a aplicação de filtros de frequências, como mostra o exemplo da [Figura 6.9](#). Esses eventos foram localizados e adicionados ao catálogo de forma manual, sendo que, em muitos casos, foram utilizadas menos do que 5 estações para localizá-los.

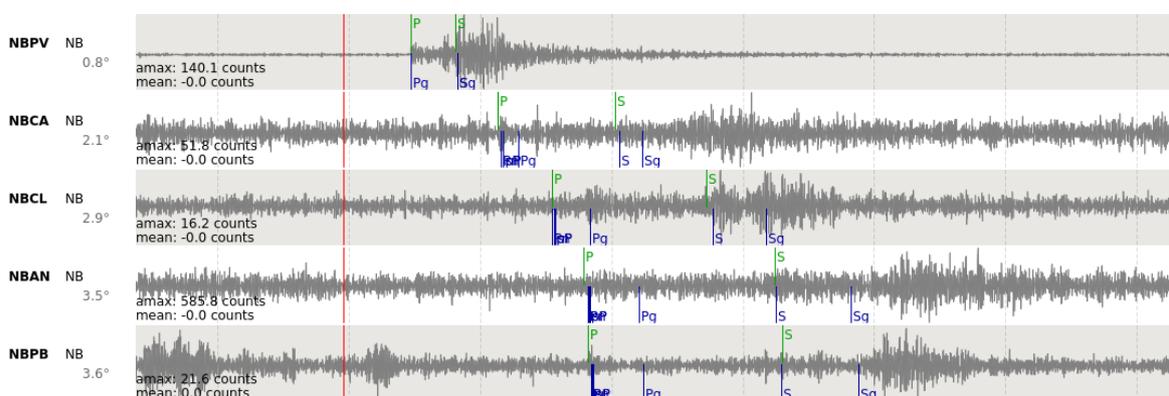


Figura 6.9: Exemplo de chegadas de onda emergentes do evento “usp2019jnud” (M2,1) ocorrido em 16/05/2019, as quais são dificilmente diferenciadas do ruído local. Os *picks* (verdes) foram realizados de forma manual, possivelmente com o auxílio das chegadas teóricas (azuis), baseadas na tabela de tempos de percurso utilizada.

Uma segunda explicação para a falta de expectativa de detecção de todos os eventos sísmicos contidos no catálogo da RSBR é a ausência de parte dos dados em forma de onda no banco do IAG-USP. Durante os trabalhos, foram realizadas sincronizações com os bancos de dados individuais mantidos pelas instituições participantes da RSBR, a fim de se obter a maior quantidade possível de dados durante o período de interesse. Entretanto, constatou-se que diversos eventos sísmicos contidos no catálogo da RSBR ocorreram em momentos em que algumas estações não possuíam dados armazenados no banco utilizado, o que dificulta suas detecções automáticas.

A falta de dados pode estar associada a um período de não funcionamento/transmissão de dados da estação, ou à indisponibilidade desses dados por parte das instituições responsáveis. A [Figura 6.10](#) apresenta um exemplo de falta de dados de 5 estações da rede NB (gerenciada pela UFRN) no momento da ocorrência de um evento sísmico contido no catálogo da RSBR.



Figura 6.10: Exemplo de falta de registros em forma de onda no banco de dados do IAG-USP, acarretando em uma menor quantidade de eventos detectáveis. O evento em questão é identificado pelo código “*usp2019scqm*” e ocorreu em 15/09/2019, com magnitude M1,3, em João Câmara/RN.

Através de uma análise dos registros sísmográficos dos 193 eventos ocorridos no ano de 2019 contidos no catálogo da RSBR, estimou-se que 52 (27%) poderiam ser detectados de forma automática após as modificações propostas. Os eventos presumidamente não possíveis de serem detectados possuem registros em poucas estações ou chegadas de onda P emergentes (com baixa razão sinal-ruído). Após as modificações no *SeisComp*, foram detectados 44 eventos concomitantes ao catálogo no ano de 2019, enquanto o *software* original foi capaz de detectar apenas 7. Cabe reforçar que *picks* automáticos mal posicionados impactam significativamente os critérios de concomitância dos eventos, portanto, uma avaliação manual de todos os eventos registrados automaticamente pode aumentar a quantidade de eventos concomitantes ao catálogo.

Considerando o catálogo do período de 2014 a 2021 (1854 eventos na região de interesse), as modificações propostas resultaram em 16% de concomitância (292 eventos), enquanto o *software* original registrou apenas 4% (78 eventos). Esse resultado indica que as modificações permitiram que a estimativa inicial fosse alcançada de forma satisfatória, reforçando a ideia de que não era esperado que todos os eventos do catálogo pudessem ser detectados e localizados de forma automática.

Os mapas da [Figura 6.11](#) apresentam os epicentros dos eventos de 2014 a 2021 contidos no catálogo da RSBR e que não foram detectados automaticamente pelo *SeisComp* modificado, comparados com os mapas de estimativas de detectabilidade automática apresentados anteriormente.

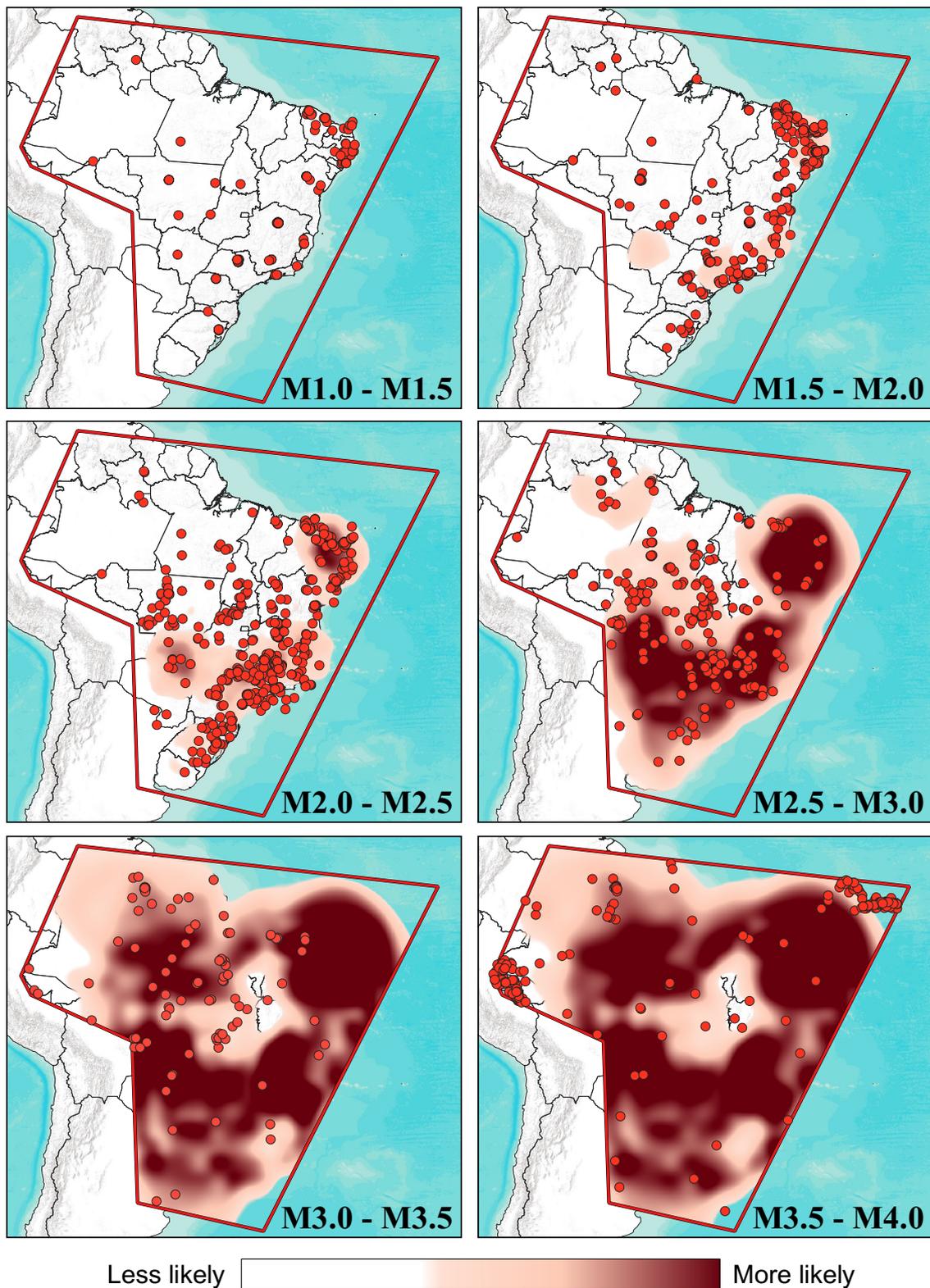


Figura 6.11: Estimativas de detectabilidade automática de eventos sísmicos no Brasil, com sobreposição dos eventos do catálogo da RSBR que não foram localizados automaticamente após as modificações no *SeisComP*. A escala qualitativa de cores indica a probabilidade de detecção automática. Os *ranges* de magnitude indicados em cada mapa são referentes aos eventos contidos no catálogo da RSBR.

Verificou-se que aproximadamente 50% dos eventos não detectados automaticamente estão fora das regiões de maior probabilidade de detecção automática pelo *SeisComP* modificado. Vale frisar, mais uma vez, que os mapas de estimativas de detectabilidade foram calculados com base no arranjo de estações do ano de 2021. Dessa forma, algumas regiões nas quais os mapas indicam maior probabilidade de detecção automática podem não corresponder à realidade dos eventos ocorridos em outros anos.

Uma grande quantidade de eventos na Região Nordeste não foi detectada automaticamente devido à ausência de dados das estações nessa região, como mostrado na [Figura 6.10](#). Os demais eventos, principalmente nas regiões Norte e Centro-Oeste, não foram bem registrados em um número satisfatório de estações.

O histograma da [Figura 6.12](#) apresenta a distribuição das magnitudes dos eventos contidos no catálogo da RSBR e que não foram detectados automaticamente após as modificações no *SeisComP*.

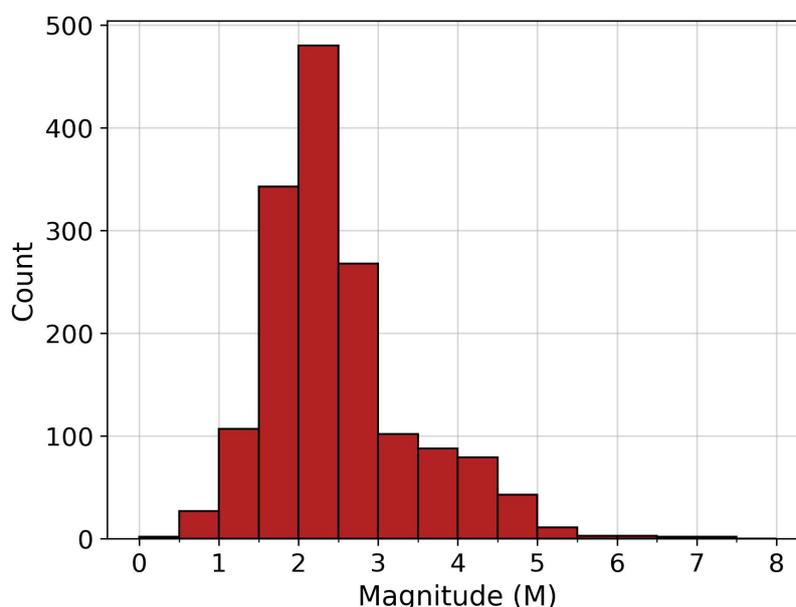


Figura 6.12: Histograma de magnitudes dos eventos contidos no catálogo da RSBR e que não foram detectados automaticamente após as modificações no *SeisComP*.

A maior parte dos eventos não detectados automaticamente apresentam magnitudes entre M1,5 e M3,0, como também pode ser observado na [Figura 6.11](#). Apesar das modificações proporcionarem maior detecção de eventos nesse *range* de magnitudes, os fatos destacados anteriormente impossibilitaram uma maior eficiência do *software*.

As modificações no *SeisComP* foram propostas visando encontrar um equilíbrio satisfatório entre detecções reais e falsos eventos. A quantidade de eventos reais detectados automaticamente pode aumentar se, por exemplo, o usuário final definir um valor de *minScore* menor do que 0,75, como estipulado nesse trabalho e discutido em “[Função *_publishable\(origin\)*](#)”. O efeito colateral é a maior probabilidade de falsas origens serem aceitas, devendo o usuário determinar seu nível de tolerância a falsos positivos.

6.5 Possíveis eventos reais não contidos no catálogo

Por outro lado, foi possível constatar a detecção de possíveis eventos sísmicos que não estão contidos no catálogo da RSBR, indicando que há a possibilidade de se incrementar a quantidade de eventos no Brasil ao utilizar os resultados obtidos através do *SeisComP* modificado, visando diminuir a defasagem do catálogo oficial da RSBR. A Figura 6.13 apresenta dois exemplos de eventos sísmicos detectados após as modificações propostas e que não constam no catálogo da RSBR.

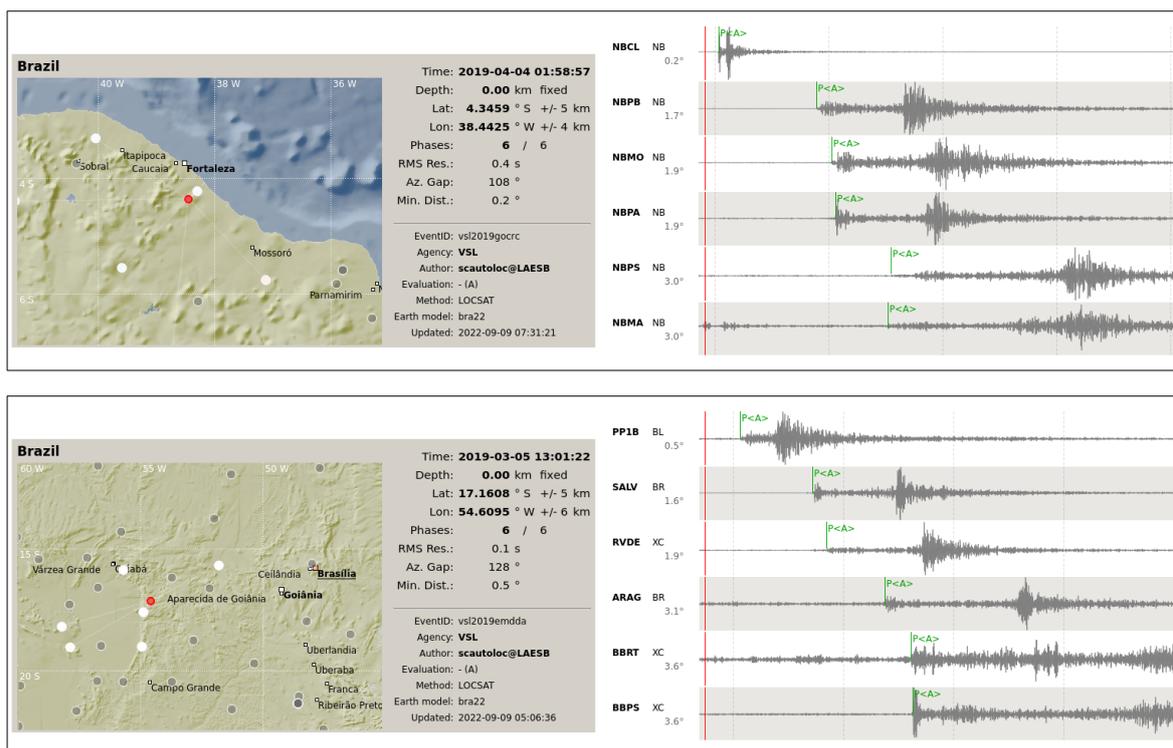


Figura 6.13: Exemplo de dois eventos sísmicos localizados automaticamente após as modificações propostas e que não constam no catálogo oficial da RSBR. Acima, um evento ocorrido em 04/04/2019 próximo à Fortaleza/CE. Abaixo, um evento ocorrido em 05/03/2019 próximo à Aparecida de Goiânia/GO.

Por fim, a grande maioria das 5981 detecções localizáveis obtidas após as modificações no *software* se referem a detonações em minas e pedreiras, as quais não constam no catálogo sísmico da RSBR por não serem eventos sísmicos naturais ou induzidos. Dessas, a maior parte refere-se a detonações próximas à cidade de Belo Horizonte/MG.

A Figura 6.14 apresenta aglomerados de detecções localizáveis consideradas “*good picking*” e provavelmente associadas a detonações, ocorridas próximas às cidades de Parauapebas/PA e Belo Horizonte/MG, em regiões conhecidas por haver detonações de explosivos. No mapa superior, constam 124 detecções localizáveis, enquanto o mapa referente à região próxima a Belo Horizonte/MG apresenta um total de 4051. Juntos, os dois mapas representam cerca de 70% das origens obtidas após as modificações, sendo essa a principal justificativa para a diferença entre a quantidade de eventos contidos no catálogo da RSBR (1854) e a quantidade de origens obtidas (5981).

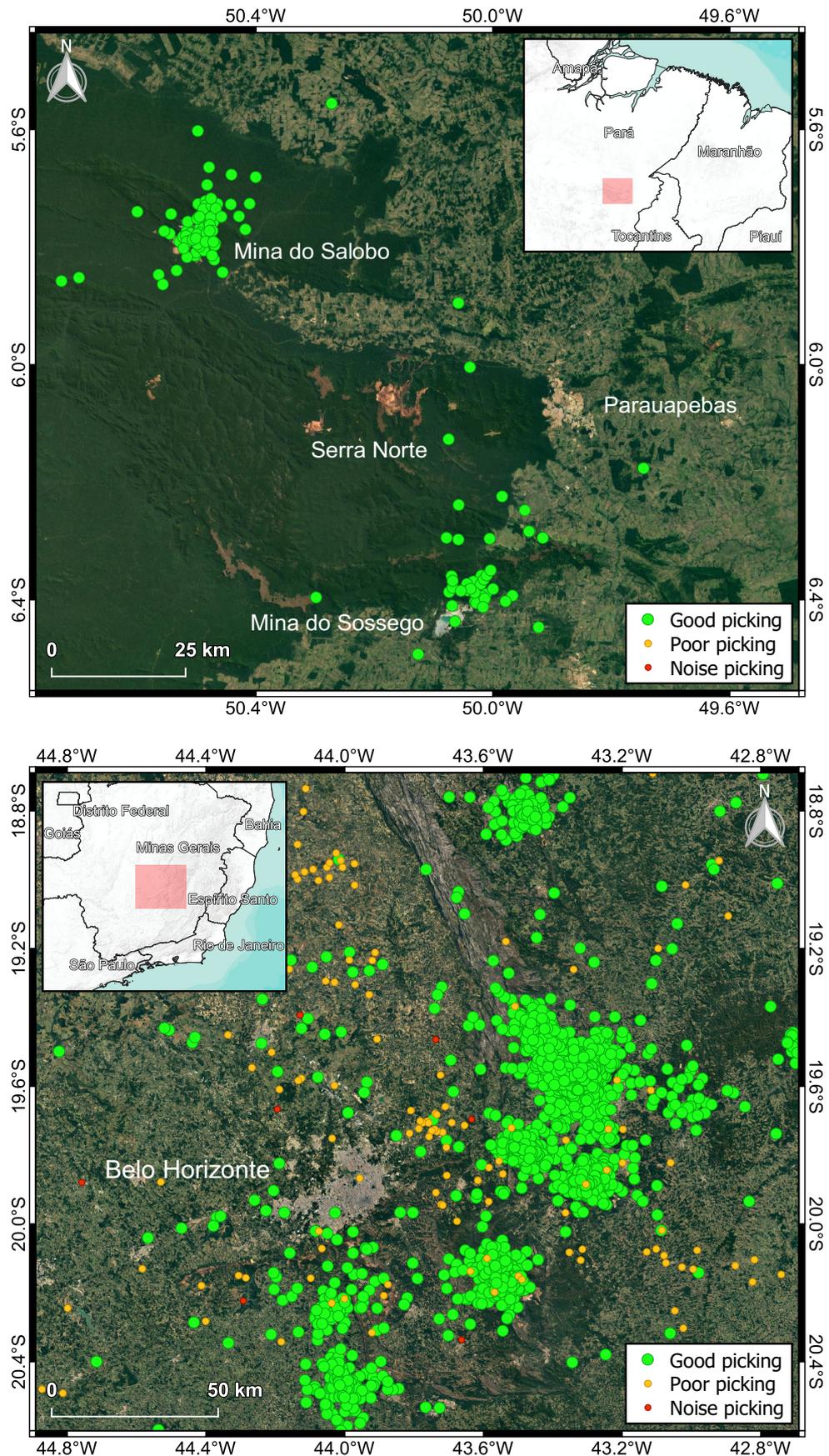


Figura 6.14: Aglomerados de detecções localizáveis possivelmente associadas a detonações em regiões de minas conhecidas. Acima, destacam-se as minas próximas a Parauapebas/PA e, abaixo, aglomerados relacionados às diversas minas próximas a Belo Horizonte/MG.

7. Conclusões

O *SeisComP* é um *software* amplamente utilizado em centros de pesquisas do mundo inteiro, sendo referência na aquisição, processamento e armazenamento de dados sísmológicos. Devido ao fato de que seu desenvolvimento foi realizado visando a detecção e localização automática de eventos sísmicos de escalas globais, torna-se necessário realizar diversos ajustes para que o *software* opere de forma mais adequada em um contexto de sismicidade regional, como é o caso da sismicidade brasileira.

Apesar de diversos parâmetros serem modificáveis através do módulo de configuração do *SeisComP*, alguns são arbitrariamente alterados no código-fonte do *software*, não sendo de fato utilizados como o usuário final os definiu inicialmente. Além disso, no código-fonte também há vários fluxos de processamento que prejudicam a detecção e localização de eventos regionais e de baixas magnitudes, os quais precisam ser avaliados e alterados, recompilando o *software* ao final do processo.

As análises realizadas nesse trabalho indicaram como parâmetros de detecção de chegadas de ondas otimizados, um filtro passa-banda com frequências de corte de 4,5 Hz e 10 Hz, em conjunto com um *picker* de AIC e janelas de tempo de 0,2 s (STA) e 45 s (LTA), visando realçar as ondas P dos eventos. Com relação aos processos *hard-coded* do módulo *scautoloc*, foram estabelecidas novas maneiras de se verificar a validade das origens nucleadas, além da remoção de parâmetros e fluxos arbitrários que prejudicam a localização automática de eventos regionais.

Os cálculos das pontuações que definem se uma origem é aceita (*originScore*) foram reformulados de forma a se basearem em eventos brasileiros para suas definições e remover complexidades arbitrárias das operações realizadas. Dentre as pontuações que definem uma origem, destacam-se as mudanças realizadas no cálculo do parâmetro *amplScore*, relacionado à pontuação da amplitude de uma chegada de onda da origem. As modificações propostas introduzem o conceito de magnitude durante o cálculo dessa pontuação, a fim de levar em consideração o fato de que as amplitudes nos sismogramas decaem com o aumento da distância epicentral. Ainda no contexto do nucleador, também propõe-se a utilização de malhas de pontos cujos números mínimos de leituras e distâncias máximas para nucleação dependem da densidade de estações em um raio de 10 graus a partir de cada nó da grade.

O desenvolvimento de um novo modelo 1D de velocidades, visando a minimização dos resíduos de eventos regionais brasileiros bem conhecidos, foi um fator importante para possibilitar a detecção automática de sismos regionais no Brasil. O novo modelo de velocidades proposto, denominado BRA23, teve como base o modelo NewBR e eventos regionais mais recentes, sendo obtido através de rotinas de processamentos de otimização de parâmetros, implementadas em linguagem de programação *Python*. Os parâmetros otimizados são compatíveis com resultados da literatura e permitiram a redução do RMS dos eventos utilizados, em comparação ao modelo NewBR.

As modificações propostas permitiram o aumento de quase 600% na quantidade total de detecções localizáveis na região de interesse, saltando de 1024 (com o *SeisComP* original) para 5981 durante o período de 2014 a 2021. Através de uma análise qualitativa dessas detecções, observou-se que as localizações satisfatórias aumentaram de 80 (de um total de 1024) para 5580 (de um total de 5981) após as modificações. Da mesma forma, as origens associadas a ruídos diminuíram de 862 (do total de 1024) para 112 (do total de 5981).

Comparando-se as coordenadas geográficas e os horários de origem das detecções localizáveis com os mesmos parâmetros dos eventos sísmicos contidos no catálogo da RSBR, constatou-se a concomitância de 292 eventos após as modificações. Anteriormente, com o *software* original, foram registradas apenas 78 detecções concomitantes ao catálogo. A principal diferença entre ambos os cenários, além do aumento significativo após as modificações, é o fato de que o *SeisComP* modificado possibilitou a detecção de muito mais eventos em território brasileiro, enquanto o *SeisComP* original detectou, em sua maioria, eventos com origens associadas à tectônica de placas (sismicidade profunda no estado do Acre e na dorsal mesoatlântica).

Ainda considerando os eventos concomitantes ao catálogo da RSBR, o *SeisComP* modificado detectou automaticamente, entre 2014 e 2021, eventos com magnitudes entre M1,5 e M5,0. Em comparação, o *SeisComP* original foi capaz de detectar apenas eventos sísmicos com magnitudes entre M3,5 e M5,0 na mesma região. Os resíduos temporais dos *arrivals* e das origens apresentaram reduções significativas após as modificações, além da inclusão da possibilidade de se registrar origens com poucas estações, um cenário comum no contexto da rede sismográfica nacional, a qual não possui uma boa densidade de estações em todo o território brasileiro.

Uma estimativa da detectabilidade automática de eventos com o atual arranjo de estações da RSBR (ano de 2021) foi inferida, considerando o fato de que, com as modificações propostas, tornou-se possível localizar eventos automaticamente com pelo menos 5 estações sismográficas. Foram gerados mapas qualitativos de detectabilidade para diferentes *ranges* de magnitudes, os quais foram posteriormente comparados com os eventos concomitantes ao catálogo da RSBR obtidos após as modificações. A sobreposição desses eventos sobre os mapas de detectabilidade mostrou uma grande concordância com as regiões de maior probabilidade de detecção para cada *range* de magnitude, indicando que as estimativas de detectabilidade automática são coerentes com a nova capacidade de detecção automática do *SeisComP*.

Os eventos que constam no catálogo da RSBR e não foram detectados automaticamente referem-se principalmente a eventos com chegadas de ondas emergentes que não se diferem dos ruídos locais, mesmo após a aplicação de filtros. Também constatou-se a ausência de dados de algumas estações durante a ocorrência de diversos eventos sísmicos contidos no catálogo, outro fator importante para impedir a detecção e localização automática desses eventos. Essa falta de dados pode estar associada a um período de não funcionamento ou

transmissão de dados ou, ainda, à indisponibilidade desses dados por parte das instituições responsáveis. Entretanto, ao se comparar as localizações dos eventos não detectados automaticamente com os mapas de estimativas de detectabilidade automática, nota-se que a maior parte desses eventos está localizada em regiões onde não espera-se que a RSBR seja capaz de realizar detecções automáticas.

Entre 2014 e 2021, o *SeisComP* modificado detectou milhares de detonações de explosivos em pedreiras e minas conhecidas, as quais não constam no catálogo da RSBR por não serem eventos sísmicos naturais ou induzidos. Ao menos 70% das origens obtidas após as modificações são, possivelmente, associadas a detonações de explosivos, sendo essa a principal justificativa para a diferença entre a quantidade de eventos contidos no catálogo da RSBR (1854) e a quantidade de origens obtidas após as modificações (5981). Por outro lado, após as modificações foram identificados possíveis eventos sísmicos naturais detectados automaticamente e que não constam no catálogo da RSBR, indicando que há a possibilidade de incrementar a quantidade de eventos no Brasil ao se utilizar os resultados obtidos nesse trabalho, visando diminuir a defasagem do catálogo oficial da RSBR.

Por fim, os resultados obtidos nesse trabalho indicaram que o *SeisComP* pode ser utilizado de maneira mais eficiente em um contexto de sismicidade regional, com eventos de magnitudes relativamente baixas. Para se obter uma maior efetividade do sistema, é necessário realizar os ajustes propostos com base nos dados da região de interesse, sendo possível reproduzi-los em qualquer região do mundo. No cenário do monitoramento sísmológico brasileiro, a adoção da versão modificada do *SeisComP* poderá facilitar consideravelmente o trabalho dos responsáveis pelas revisões manuais dos dados gerados pelo *software*, além de permitir a detecção de eventos que poderiam ser negligenciados nessas revisões.

A relação de todas as modificações aplicadas diretamente no código-fonte do *SeisComP*, bem como a dissertação (em português e inglês), o modelo de velocidades BRA23 e as malhas de pontos computadas para cada ano estão disponíveis em um repositório público⁴ da plataforma *Zenodo* (Salles, 2023).

⁴<https://zenodo.org/record/7412143>

Referências Bibliográficas

- Akaike, H. (1971). Autoregressive Model Fitting for Control. *Annals of The Institute of Statistical Mathematics - Ann. Inst. Statist. Math*, 23:163, DOI: [10.1007/BF02479221](https://doi.org/10.1007/BF02479221).
- Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces. *Bulletin of the Seismological Society of America*, 68(5):1521, ISSN: 0037-1106, DOI: [10.1785/BSSA0680051521](https://doi.org/10.1785/BSSA0680051521).
- Alsaker, A. et al. (1991). The ML scale in Norway. *Bulletin of the Seismological Society of America*, 81:379, DOI: [10.1785/BSSA0810020379](https://doi.org/10.1785/BSSA0810020379).
- Ardito, J. C. (2009). Determinação de epicentros regionais. *Trabalho de Graduação. Instituto de Astronomia, Geofísica e Ciências Atmosféricas da Universidade de São Paulo (IAG-USP)*.
- Assumpção, M. (1983). A regional magnitude scale for Brazil. *Bulletin of the Seismological Society of America*, 73(1):237, ISSN: 0037-1106.
- Assumpção, M., Ardito, J. C., and Barbosa, J. R. (2010). An improved velocity model for regional epicentre determination in Brazil. *IV Simpósio Brasileiro de Geofísica*, pages 13–16.
- Assumpção, M. et al. (2014). Intraplate seismicity in Brazil. *Intraplate earthquakes*, pages 50–71.
- Behr, Y. et al. (2016). The Virtual Seismologist in SeisComp3: A New Implementation Strategy for Earthquake Early Warning Algorithms. *Seismological Research Letters*, 87:363, DOI: [10.1785/0220150235](https://doi.org/10.1785/0220150235).
- Beyreuther, M. et al. (2010). ObsPy: A Python toolbox for seismology. *Seismological Research Letters*, 81(3):530.
- Bianchi, M. et al. (2012). The Brazilian Seismographic Network: Historical Overview and Current Status. *Summary of the Bulletin of the International Seismological Centre*, 49:70, DOI: [10.5281/zenodo.998851](https://doi.org/10.5281/zenodo.998851).
- Bianchi, M. et al. (2018). The Brazilian Seismographic Network (RSBR): Improving Seismic Monitoring in Brazil. *Seismological Research Letters*, 89(2A):452, DOI: [10.1785/0220170227](https://doi.org/10.1785/0220170227).
- Ciardelli, C. et al. (2022). Adjoint waveform tomography of South America. *Journal of Geophysical Research: Solid Earth*, 127(2):e2021JB022575.
- Clark, A., Evans, P., and Strollo, A. (2014). FDSN recommendations for seismic network DOIs and related FDSN services. *version 1.0. Technical report. International Federation of Digital Seismograph Networks*.

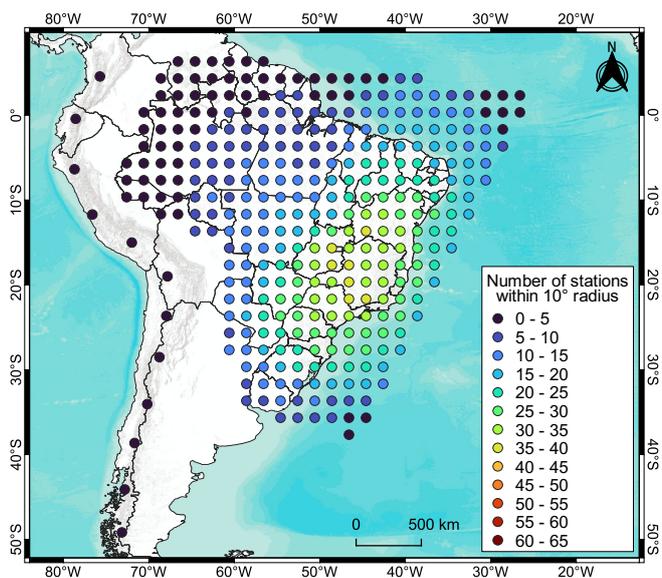
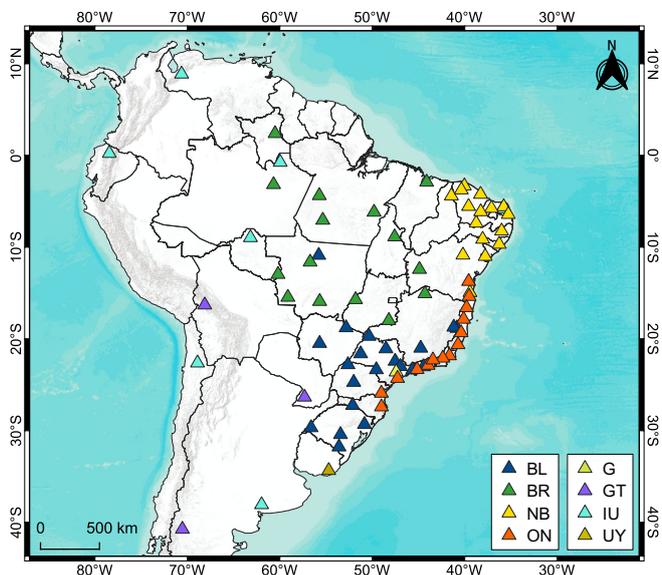
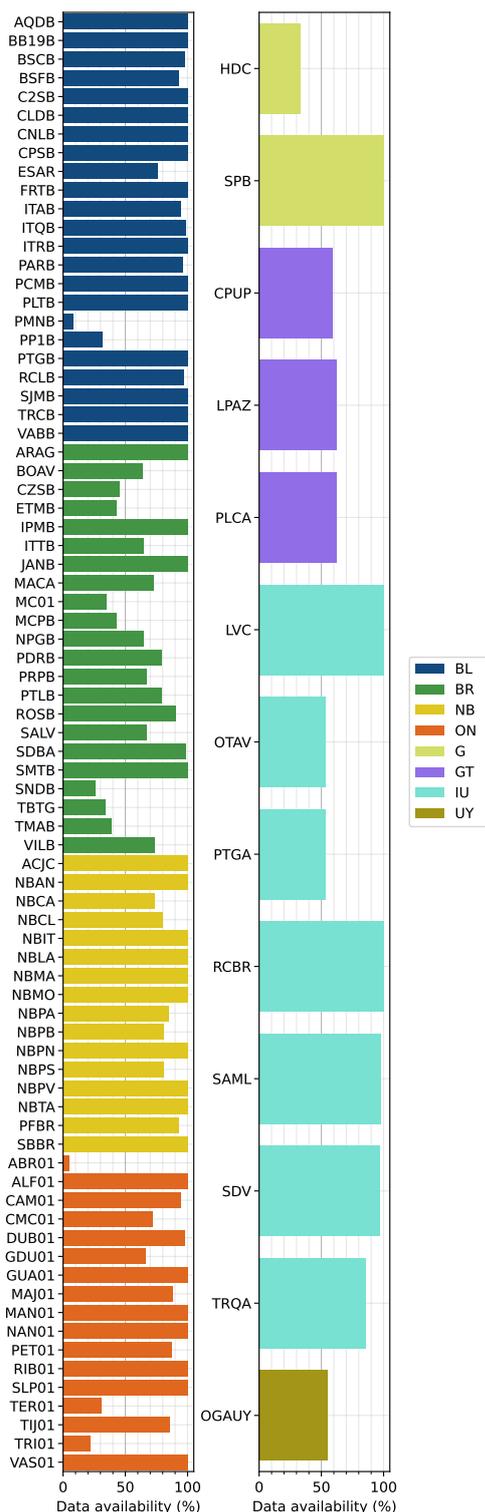
- Crotwell, H. P., Owens, T. J., and Ritsema, J. (1999). The TauP Toolkit: Flexible seismic travel-time and ray-path utilities. *Seismological Research Letters*, 70(2):154, ISSN: 0895-0695, DOI: [10.1785/gssrl.70.2.154](https://doi.org/10.1785/gssrl.70.2.154).
- Duda, S. J. and Nuttli, O. W. (1974). Earthquake magnitude scales. *Geophysical surveys*, 1(4):429.
- Heit, B. et al. (2007). An S receiver function analysis of the lithospheric structure in South America. *Geophysical Research Letters*, 34(14).
- Helmholtz-Centre Potsdam - GFZ German Research Centre for Geosciences and gempa GmbH (2008). The SeisComP seismological software package. GFZ Data Services. DOI: [10.5880/GFZ.2.4.2020.003](https://doi.org/10.5880/GFZ.2.4.2020.003).
- Herrin, E. (1968). Seismological tables for P-phases. *Bulletin of the Seismological Society of America*, 60:461.
- IRIS (2010). SEED Manual - Version 2.4. [FDSN Website](https://www.fdsn.org/). Accessed: March 12, 2021.
- Jeffreys, H. and Bullen, K. E. (1940). Seismological tables.
- Kanamori, H. (1983). Magnitude scale and quantification of earthquakes. *Tectonophysics*, 93(3):185, ISSN: 0040-1951, DOI: [10.1016/0040-1951\(83\)90273-1](https://doi.org/10.1016/0040-1951(83)90273-1). Quantification of Earthquakes.
- Kennett, B. and Engdahl, E. (1991). Traveltimes for global earthquake location and phase identification. *Geophysical Journal International*, 105:429, DOI: [10.1111/j.1365-246X.1991.tb06724.x](https://doi.org/10.1111/j.1365-246X.1991.tb06724.x).
- Kennett, B. L., Engdahl, E., and Buland, R. (1995). Constraints on seismic velocities in the Earth from traveltimes. *Geophysical Journal International*, 122(1):108.
- Li, X. et al. (2016). Identifying P phase arrival of weak events: The Akaike Information Criterion picking application based on the Empirical Mode Decomposition. *Computers and Geosciences*, 100, DOI: [10.1016/j.cageo.2016.12.005](https://doi.org/10.1016/j.cageo.2016.12.005).
- Lopez, C. M. (2021). Parameter optimization of automatic phase detection and picking algorithms - Application in São Paulo University Seismological Center and Colombian National Seismic Network. *Dissertação de Mestrado. Instituto de Astronomia, Geofísica e Ciências Atmosféricas da Universidade de São Paulo (IAG-USP)*.
- Montagner, J.-P. and Kennett, B. (1996). How to reconcile body-wave and normal-mode reference Earth models. *Geophysical Journal International*, 125(1):229.
- Mooney, W. D., Laske, G., and Masters, T. G. (1998). CRUST 5.1: A global crustal model at 5 × 5. *Journal of Geophysical Research: Solid Earth*, 103(B1):727.

- Pirchiner, M. et al. (2011). The BRAZilian Seismographic Integrated Systems (BRASIS): Infrastructure and Data Management. *Annals of Geophysics*, 54(1):17, DOI: 10.4401/ag-4865.
- Richter, C. F. (1935). An instrumental earthquake magnitude scale. *Bulletin of the Seismological Society of America*, 25(1):1, ISSN: 0037-1106, DOI: 10.1785/BSSA0250010001.
- Rivadeneira, C. et al. (2019). An updated crustal thickness map of Central South America based on receiver function measurements in the region of the Chaco, Pantanal, and Paraná Basins, Southwestern Brazil. *Journal of Geophysical Research: Solid Earth*, 124, DOI: 10.1029/2018JB016811.
- Salles, V. (2023). SeisComp software optimization focused on automatic detection and localization of regional seismic events in Brazil. (Complementary Data) to the Master's Thesis, DOI: 10.5281/zenodo.7412143. [Zenodo Public Repository](#).
- Schorlemmer, D. et al. (2011). QuakeML: Status of the XML-based seismological data exchange format. 54:59.
- St-Onge, A. (2011). Akaike Information Criterion applied to detecting first arrival times on microseismic data. *Society of Exploration Geophysicists*, page 1658.
- Storn, R. and Price, K. (1997). Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, 11(4):341.
- Virtanen, P. et al. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261, DOI: 10.1038/s41592-019-0686-2.
- Yadav, A. K. and Mishra, P. K. (2014). Method for detection of first phase in seismogram based on Akaike Information Criterion (AIC) function. *International journal of engineering research and technology*, 3.

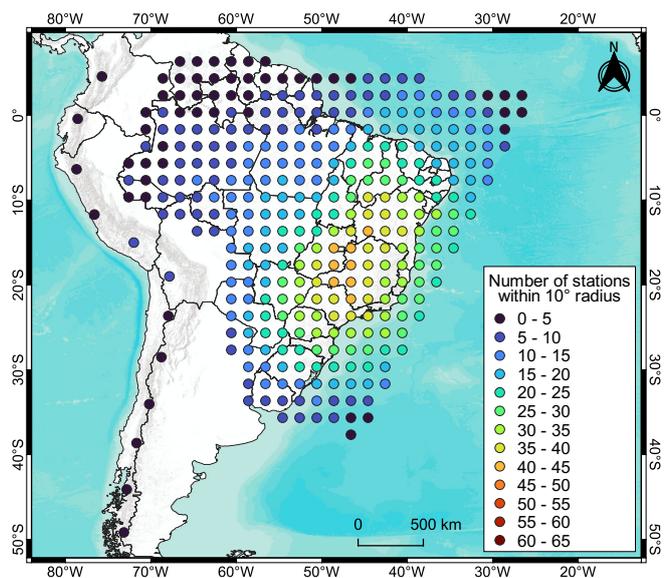
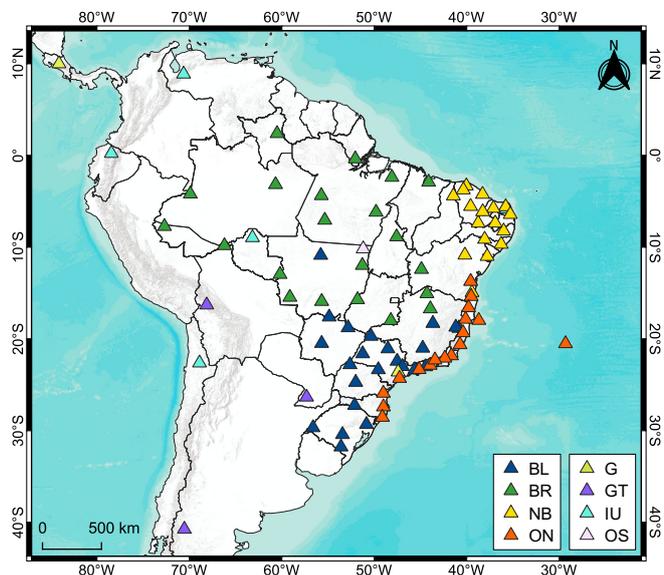
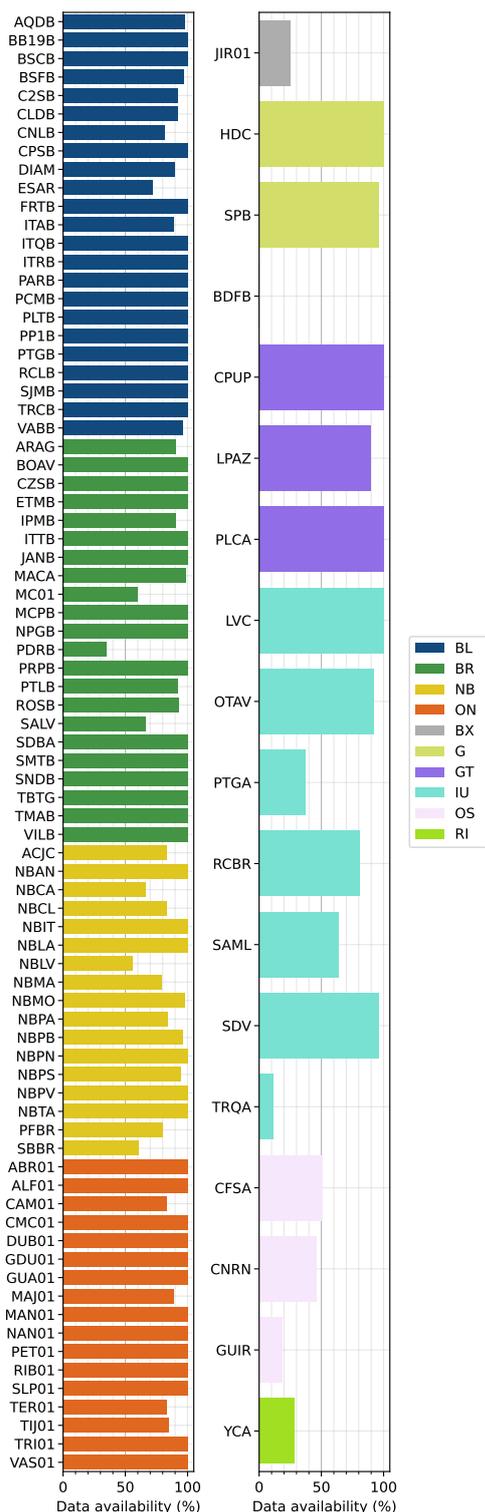
Apêndices

APÊNDICE 1

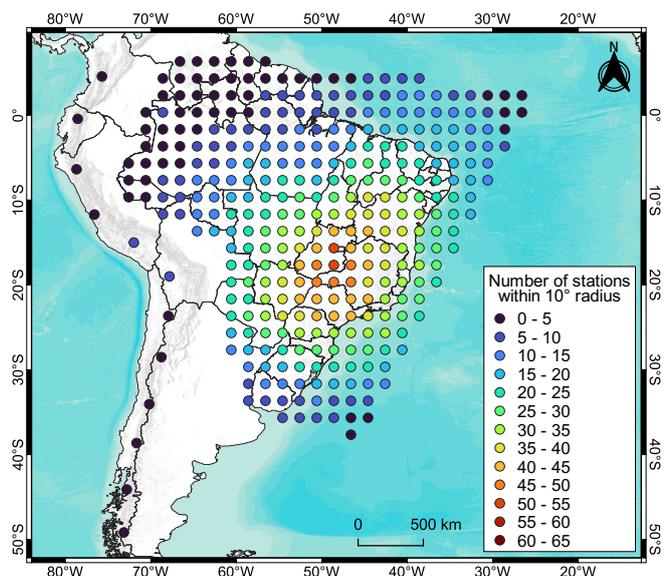
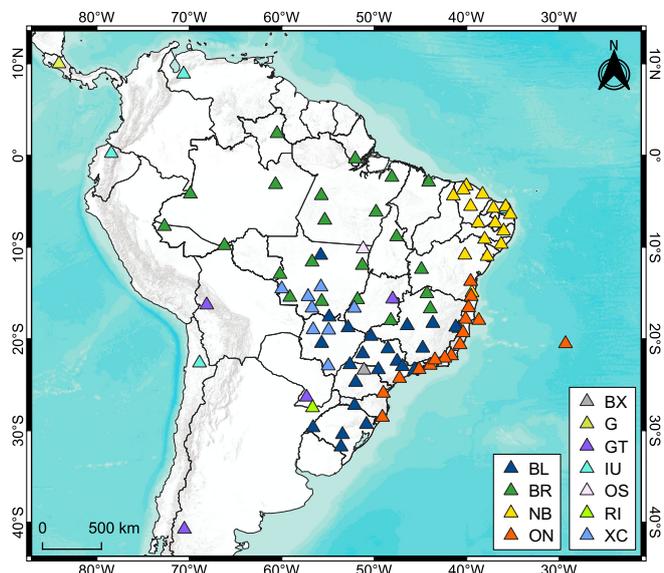
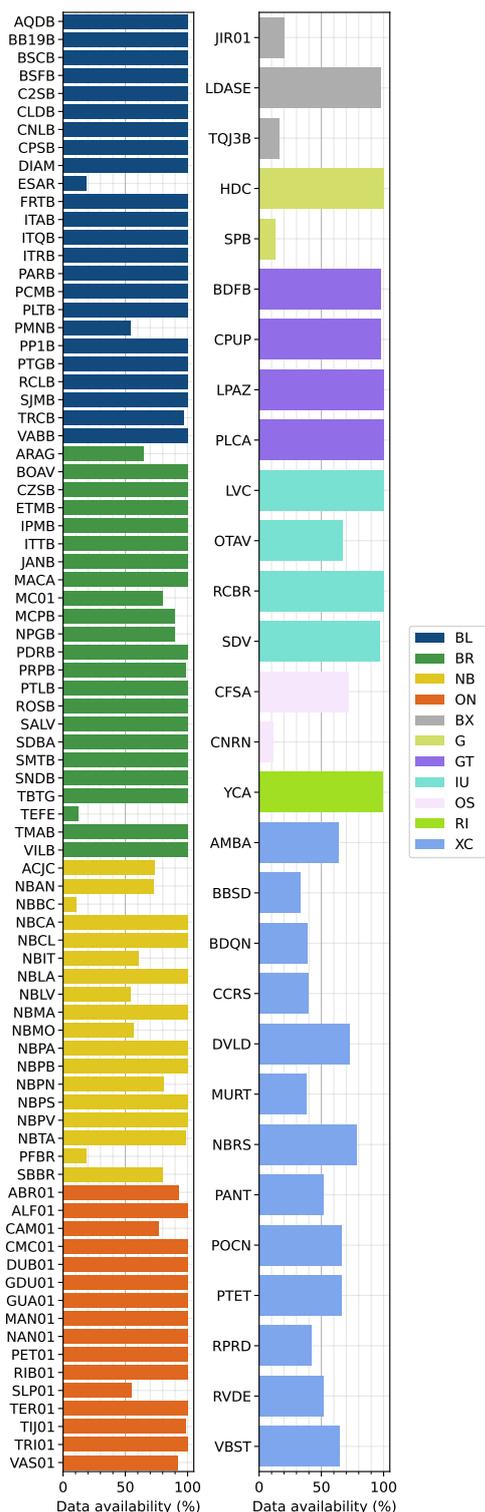
Malhas de pontos (*grids*) - 2014 a 2021



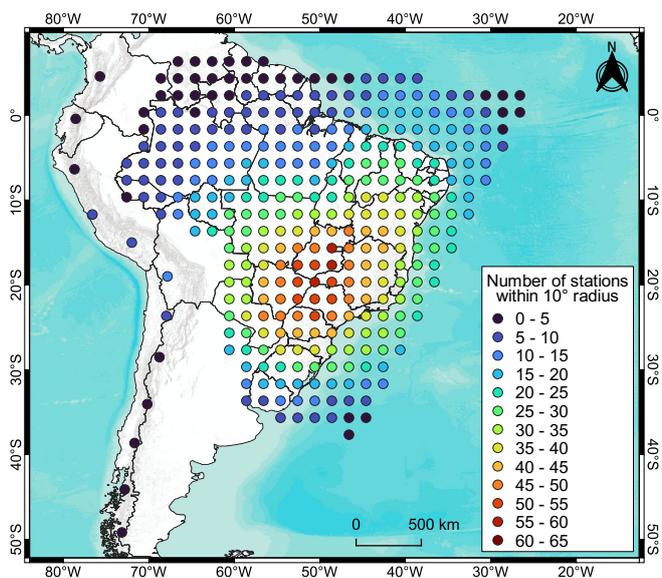
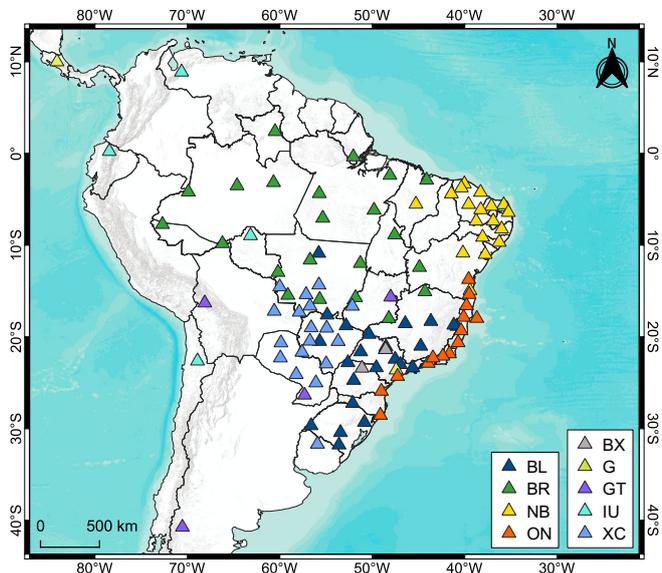
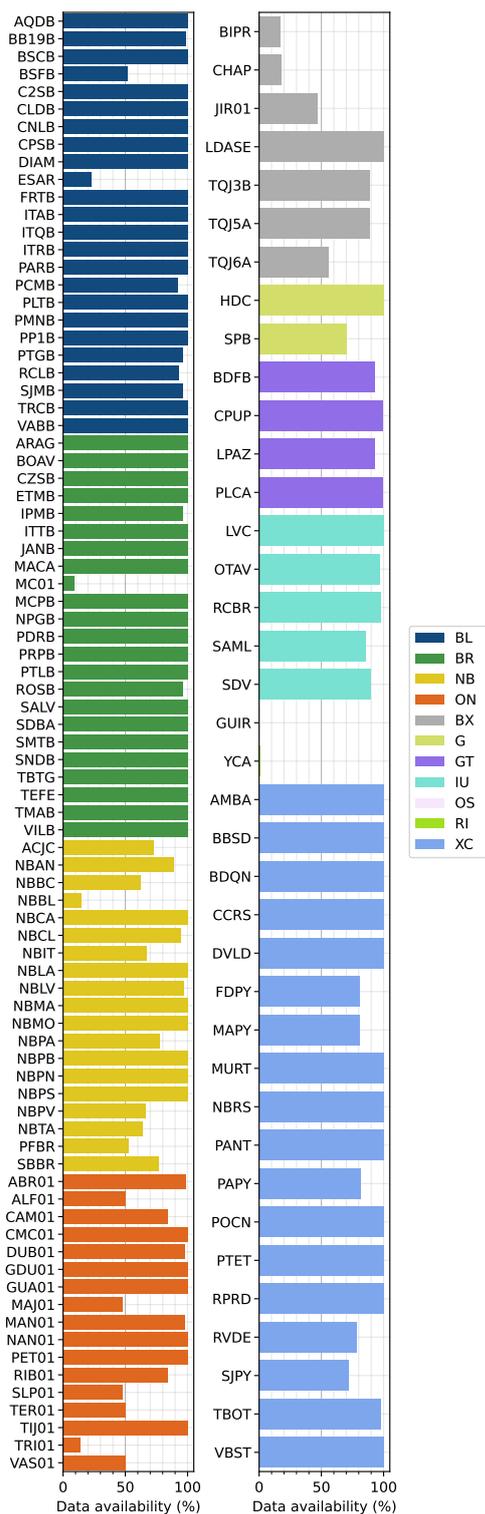
Esq.: Disponibilidade de dados em forma de ondas, registrados em 2014, de cada estação contida no inventário do *SeisComP* do IAG-USP em operação nesse ano. Dir.: Localizações das estações com no mínimo 50% de dados disponíveis no período e a malha de pontos obtida a partir da densidade espacial dessas estações.



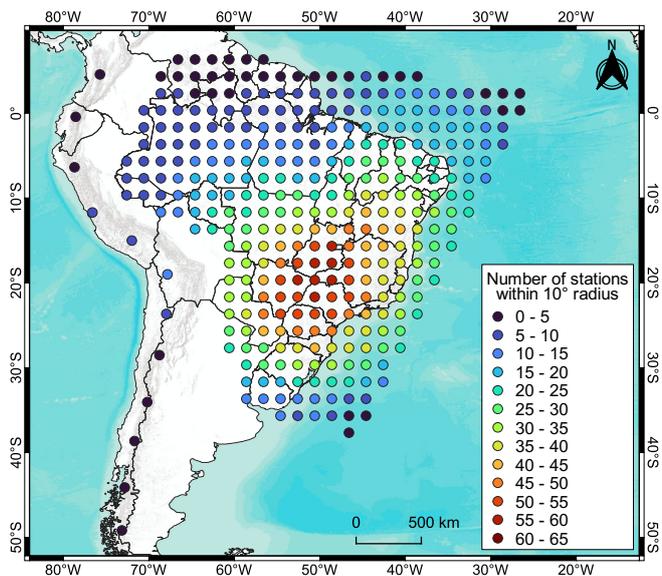
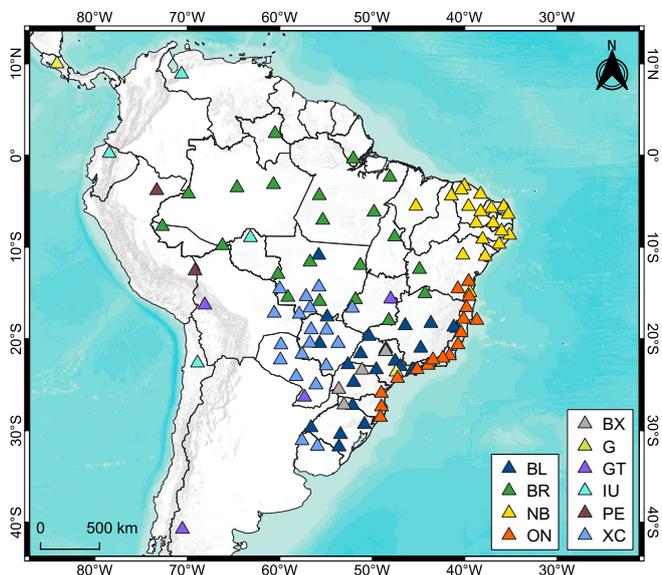
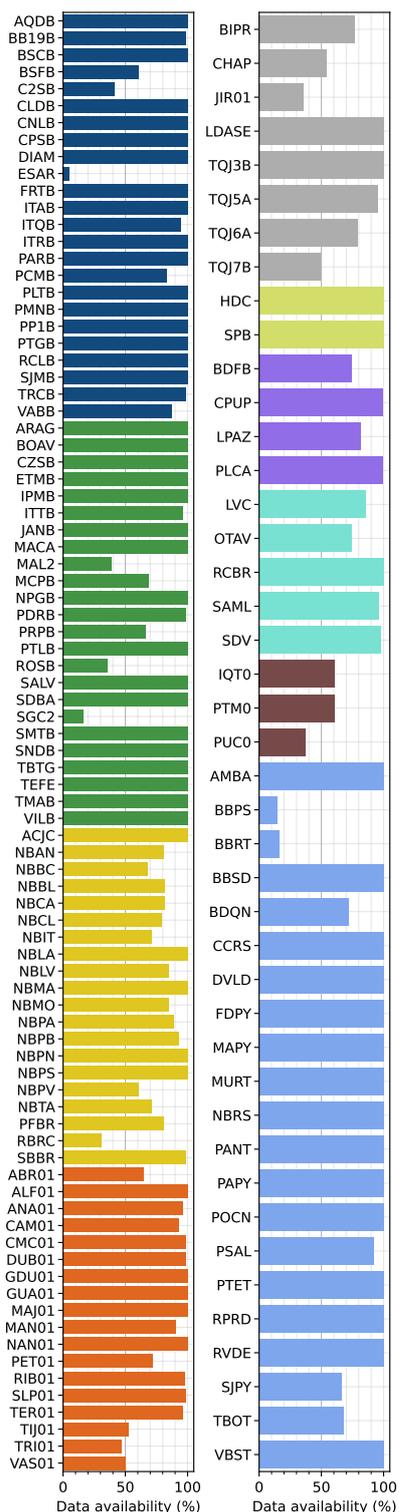
Esq.: Disponibilidade de dados em forma de ondas, registrados em 2015, de cada estação contida no inventário do *SeisComP* do IAG-USP em operação nesse ano. Dir.: Localizações das estações com no mínimo 50% de dados disponíveis no período e a malha de pontos obtida a partir da densidade espacial dessas estações.



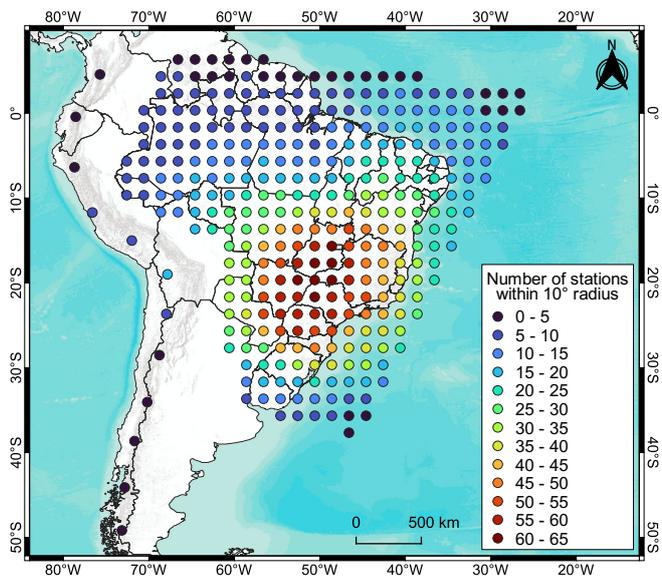
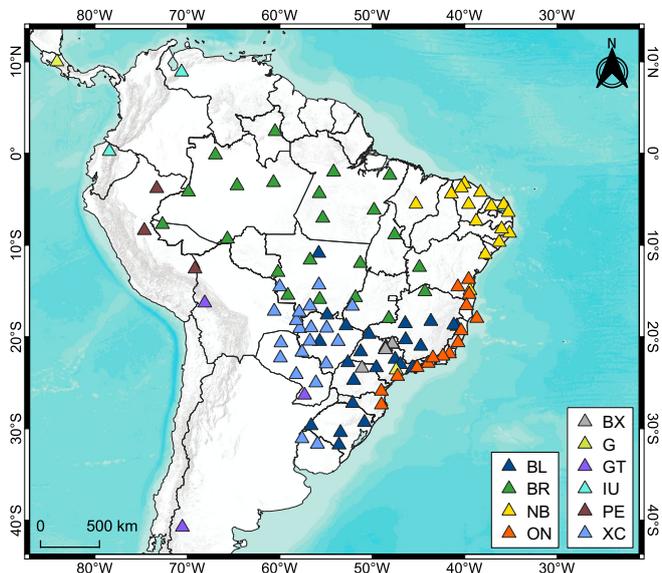
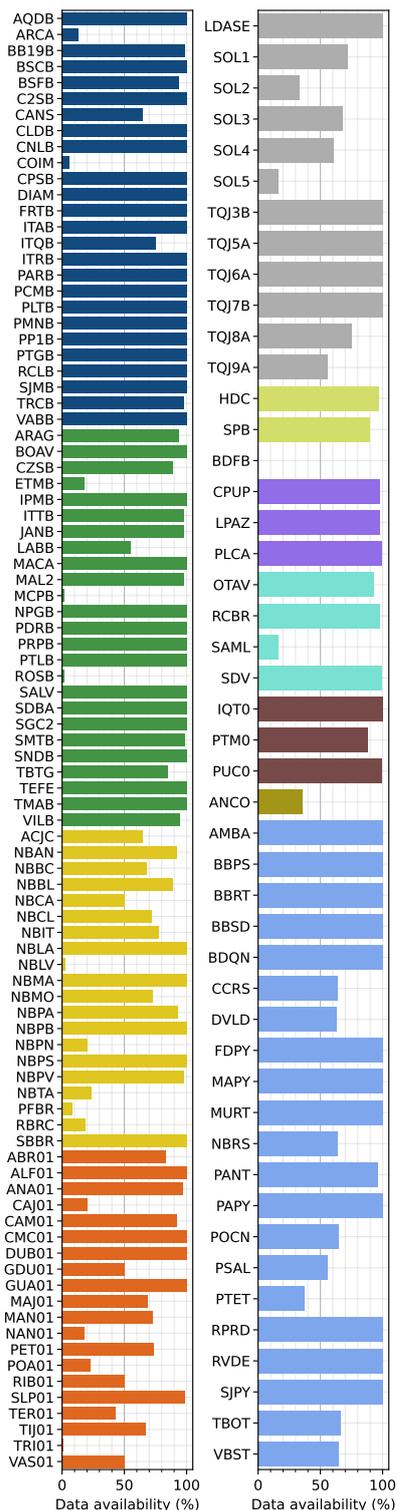
Esq.: Disponibilidade de dados em forma de ondas, registrados em 2016, de cada estação contida no inventário do *SeisComP* do IAG-USP em operação nesse ano. Dir.: Localizações das estações com no mínimo 50% de dados disponíveis no período e a malha de pontos obtida a partir da densidade espacial dessas estações.



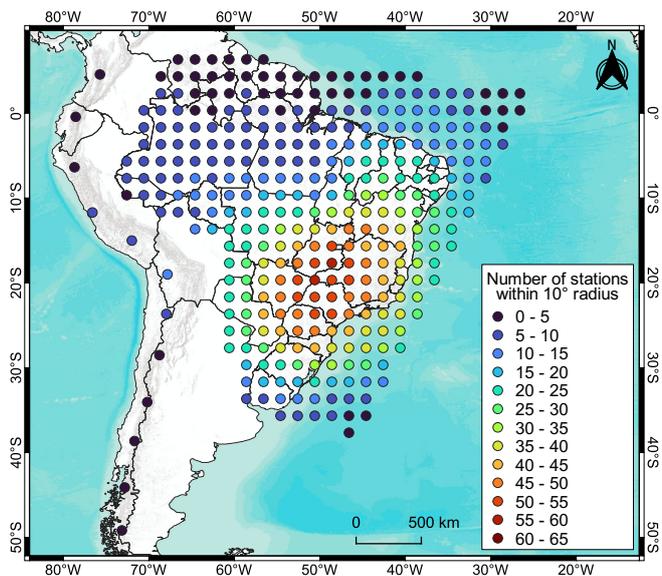
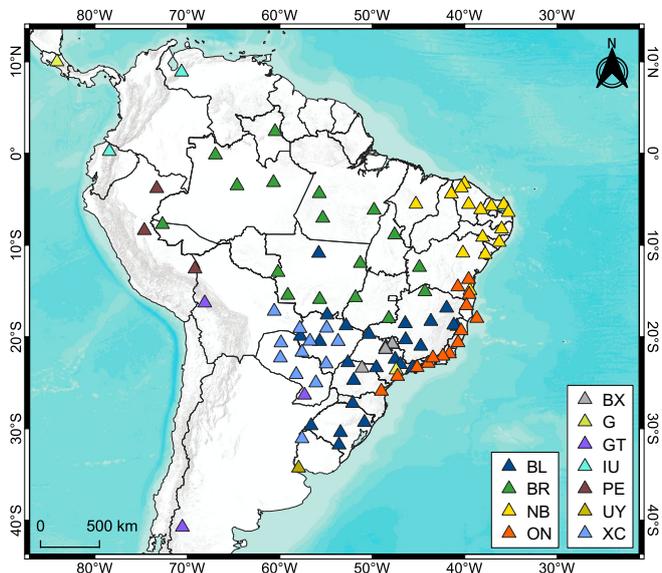
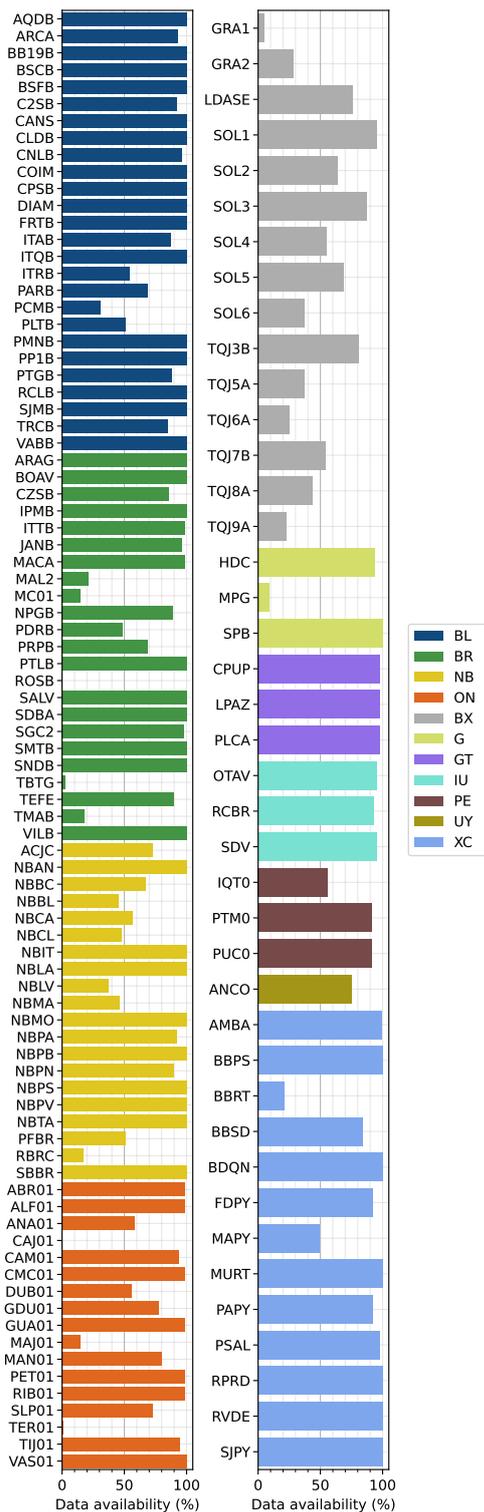
Esq.: Disponibilidade de dados em forma de ondas, registrados em 2017, de cada estação contida no inventário do *SeisComP* do IAG-USP em operação nesse ano. Dir.: Localizações das estações com no mínimo 50% de dados disponíveis no período e a malha de pontos obtida a partir da densidade espacial dessas estações.



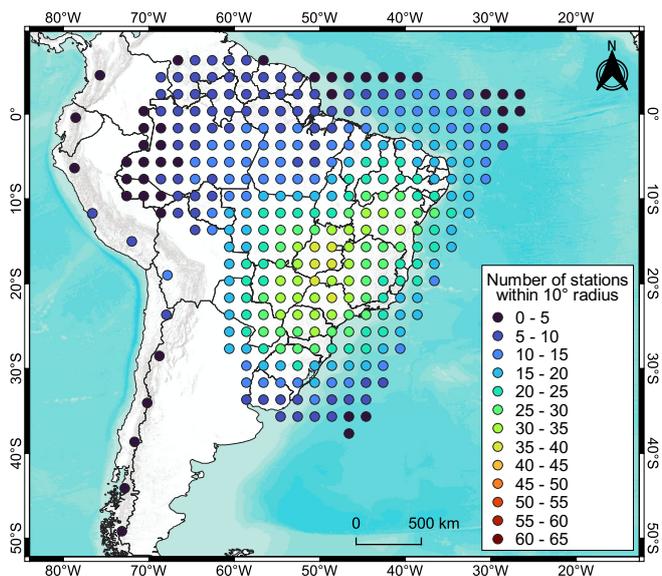
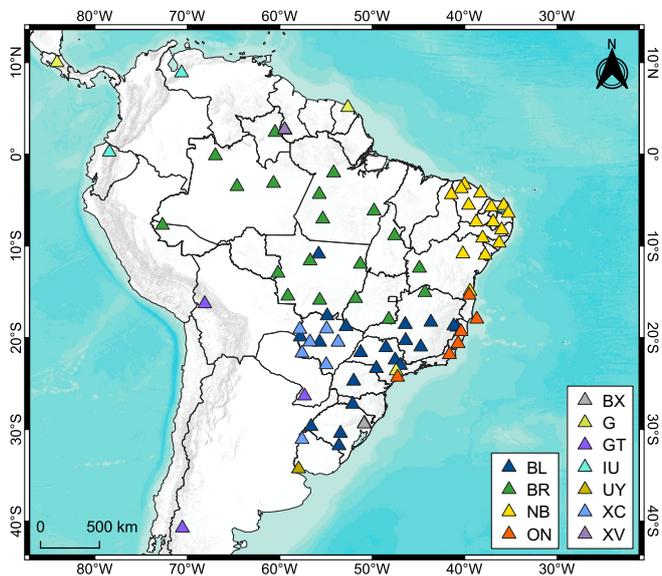
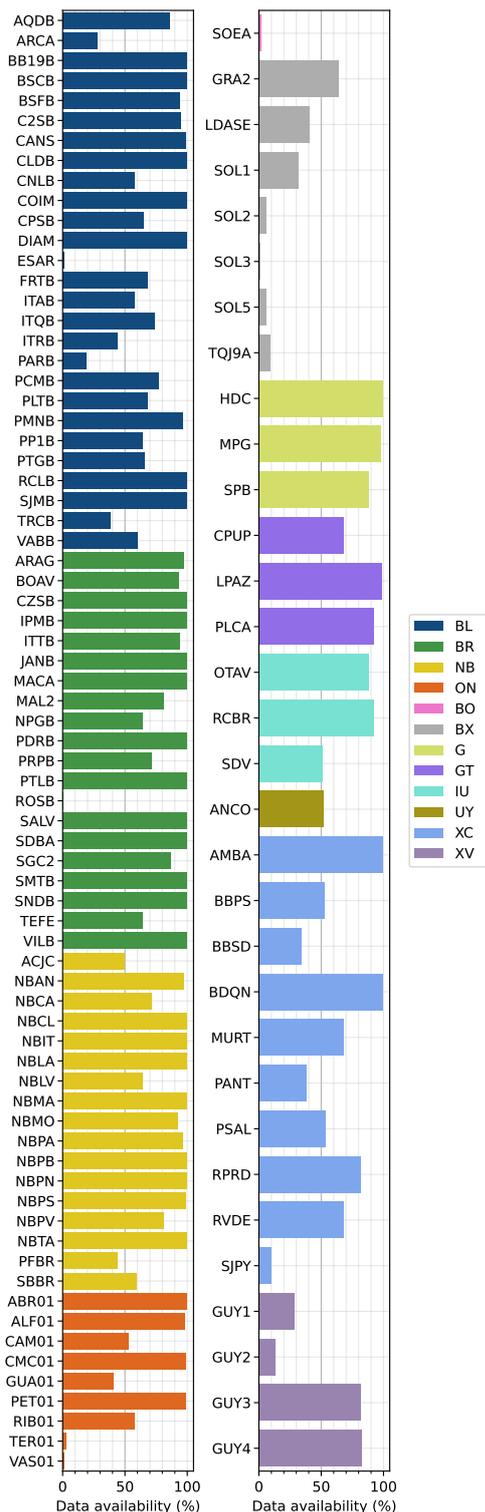
Esq.: Disponibilidade de dados em forma de ondas, registrados em 2018, de cada estação contida no inventário do *SeisComP* do IAG-USP em operação nesse ano. Dir.: Localizações das estações com no mínimo 50% de dados disponíveis no período e a malha de pontos obtida a partir da densidade espacial dessas estações.



Esq.: Disponibilidade de dados em forma de ondas, registrados em 2019, de cada estação contida no inventário do *SeisComP* do IAG-USP em operação nesse ano. Dir.: Localizações das estações com no mínimo 50% de dados disponíveis no período e a malha de pontos obtida a partir da densidade espacial dessas estações.



Esq.: Disponibilidade de dados em forma de ondas, registrados em 2020, de cada estação contida no inventário do *SeisComP* do IAG-USP em operação nesse ano. Dir.: Localizações das estações com no mínimo 50% de dados disponíveis no período e a malha de pontos obtida a partir da densidade espacial dessas estações.



Esq.: Disponibilidade de dados em forma de ondas, registrados em 2021, de cada estação contida no inventário do *SeisComP* do IAG-USP em operação nesse ano. Dir.: Localizações das estações com no mínimo 50% de dados disponíveis no período e a malha de pontos obtida a partir da densidade espacial dessas estações.

APÊNDICE 2

Modelo de velocidades 1D

BRA23

Depth (km)	V_P (km/s)	V_S (km/s)	Density (g/cm³)
0.0000	5.8000	3.4600	2.7200
14.3000	5.8000	3.4600	2.7200
14.3000	6.6556	3.8500	2.9200
40.0000	6.6556	3.8500	2.9200
40.0000	8.2369	4.4800	3.3198
198.0000	8.2369	4.5000	3.3713
660.0000	10.3194	5.6100	4.0646
660.0000	10.7900	5.9600	4.3714
710.0000	10.9229	6.0897	4.4010
760.0000	11.0558	6.2095	4.4305
809.5000	11.1353	6.2426	4.4596
859.0000	11.2221	6.2798	4.4885
908.5000	11.3068	6.3160	4.5173
958.0000	11.3896	6.3512	4.5459
1007.5000	11.4705	6.3854	4.5744
1057.0000	11.5495	6.4187	4.6028
1106.5000	11.6269	6.4510	4.6310
1156.0000	11.7026	6.4828	4.6591
1205.5000	11.7766	6.5138	4.6870
1255.0000	11.8491	6.5439	4.7148
1304.5000	11.9200	6.5727	4.7424
1354.0000	11.9895	6.6008	4.7699
1403.5000	12.0577	6.6285	4.7973
1453.0000	12.1245	6.6555	4.8245
1502.5000	12.1912	6.6815	4.8515
1552.0000	12.2550	6.7073	4.8785
1601.5000	12.3185	6.7326	4.9052
1651.0000	12.3819	6.7573	4.9319
1700.5000	12.4426	6.7815	4.9584
1750.0000	12.5031	6.8052	4.9847
1799.5000	12.5631	6.8286	5.0109
1849.0000	12.6221	6.8515	5.0370
1898.5000	12.6804	6.8742	5.0629
1948.0000	12.7382	6.8972	5.0887
1997.5000	12.7956	6.9194	5.1143
2047.0000	12.8526	6.9418	5.1398
2096.5000	12.9096	6.9627	5.1652

Depth (km)	V_P (km/s)	V_S (km/s)	Density (g/cm³)
2146.0000	12.9668	6.9855	5.1904
2195.5000	13.0222	7.0063	5.2154
2245.0000	13.0783	7.0281	5.2403
2294.5000	13.1336	7.0500	5.2651
2344.0000	13.1894	7.0720	5.2898
2393.5000	13.2465	7.0931	5.3142
2443.0000	13.3018	7.1144	5.3386
2492.5000	13.3585	7.1369	5.3628
2542.0000	13.4156	7.1586	5.3869
2591.5000	13.4741	7.1807	5.4108
2640.0000	13.5312	7.2031	5.4345
2690.0000	13.5900	7.2258	5.4582
2740.0000	13.6494	7.2490	5.4817
2740.0000	13.6494	7.2490	5.4817
2789.6700	13.6530	7.2597	5.5051
2839.3300	13.6566	7.2704	5.5284
2891.5000	13.6602	7.2811	5.5515
2891.5000	8.0000	0.0000	9.9145
2939.3300	8.0382	0.0000	9.9942
2989.6600	8.1283	0.0000	10.0722
3039.9900	8.2213	0.0000	10.1485
3090.3200	8.3122	0.0000	10.2233
3140.6600	8.4001	0.0000	10.2964
3190.9900	8.4861	0.0000	10.3679
3241.3200	8.5692	0.0000	10.4378
3291.6500	8.6496	0.0000	10.5062
3341.9800	8.7283	0.0000	10.5731
3392.3100	8.8036	0.0000	10.6385
3442.6400	8.8761	0.0000	10.7023
3492.9700	8.9461	0.0000	10.7647
3543.3000	9.0138	0.0000	10.8257
3593.6400	9.0792	0.0000	10.8852
3643.9700	9.1426	0.0000	10.9434
3694.3000	9.2042	0.0000	11.0001
3744.6300	9.2634	0.0000	11.0555
3794.9600	9.3205	0.0000	11.1095
3845.2900	9.3760	0.0000	11.1623

Depth (km)	V_P (km/s)	V_S (km/s)	Density (g/cm³)
3895.6200	9.4297	0.0000	11.2137
3945.9500	9.4814	0.0000	11.2639
3996.2800	9.5306	0.0000	11.3127
4046.6200	9.5777	0.0000	11.3604
4096.9500	9.6232	0.0000	11.4069
4147.2800	9.6673	0.0000	11.4521
4197.6100	9.7100	0.0000	11.4962
4247.9400	9.7513	0.0000	11.5391
4298.2700	9.7914	0.0000	11.5809
4348.6000	9.8304	0.0000	11.6216
4398.9300	9.8682	0.0000	11.6612
4449.2600	9.9051	0.0000	11.6998
4499.6000	9.9410	0.0000	11.7373
4549.9300	9.9761	0.0000	11.7737
4600.2600	10.0103	0.0000	11.8092
4650.5900	10.0439	0.0000	11.8437
4700.9200	10.0768	0.0000	11.8772
4801.5800	10.1415	0.0000	11.9414
4851.9100	10.1739	0.0000	11.9722
4902.2400	10.2049	0.0000	12.0001
4952.5800	10.2329	0.0000	12.0311
5002.9100	10.2565	0.0000	12.0593
5053.2400	10.2745	0.0000	12.0867
5103.5700	10.2854	0.0000	12.1133
5153.5000	10.2890	0.0000	12.1391
5153.5000	11.0427	3.5043	12.7037
5204.6100	11.0585	3.5187	12.7289
5255.3200	11.0718	3.5314	12.7530
5306.0400	11.0850	3.5435	12.7760
5356.7500	11.0983	3.5551	12.7980
5407.4600	11.1166	3.5661	12.8188
5458.1700	11.1316	3.5765	12.8387
5508.8900	11.1457	3.5864	12.8574
5559.6000	11.1590	3.5957	12.8751
5610.3100	11.1715	3.6044	12.8917
5661.0200	11.1832	3.6126	12.9072
5711.7400	11.1941	3.6202	12.9217

Depth (km)	V_P (km/s)	V_S (km/s)	Density (g/cm³)
5813.1600	11.2134	3.6337	12.9474
5863.8700	11.2219	3.6396	12.9586
5914.5900	11.2295	3.6450	12.9688
5965.3000	11.2364	3.6498	12.9779
6016.0100	11.2424	3.6540	12.9859
6066.7200	11.2477	3.6577	12.9929
6117.4400	11.2521	3.6608	12.9988
6168.1500	11.2557	3.6633	13.0036
6218.8600	11.2586	3.6653	13.0074
6269.5700	11.2606	3.6667	13.0100
6320.2900	11.2618	3.6675	13.0117
6371.0000	11.2622	3.6678	13.0122