

Otimização do *software SeisComP* para detecção de eventos sísmicos no Brasil

Salles, V., USP; Bianchi, M.; USP

Copyright 2022, SBGf - Sociedade Brasileira de Geofísica.

Este texto foi preparado para a apresentação no IX Simpósio Brasileiro de Geofísica, Curitiba, 4 a 6 de outubro de 2022. Seu conteúdo foi revisado pelo Comitê Técnico do IX SimBGf, mas não necessariamente representa a opinião da SBGf ou de seus associados. É proibida a reprodução total ou parcial deste material para propósitos comerciais sem prévia autorização da SBGf.

Resumo

A Rede Sismográfica Brasileira (RSBR) está em operação desde 2009, sendo 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 seu código-fonte e em seus parâmetros de detecção e localização. Mais de 90% dos eventos no boletim sísmico da RSBR são detectados manualmente por analistas em rotinas diárias. Nesse trabalho, inicialmente realizamos uma análise da capacidade de detecção automática da RSBR com base em seu boletim sísmico, seguido de um estudo para otimizar os parâmetros de detecção (filtros de frequência e comprimento das janelas STA/LTA, bem como os limites de detecção) utilizando eventos conhecidos. Também buscamos otimizar os procedimentos de nucleação, inclusive os que estão contidos no código-fonte do *software*, de forma a maximizar a detecção de eventos reais e minimizando os falsos positivos. A análise indicou como parâmetros otimizados um filtro passa-banda com frequências de corte de 4,5 Hz e 10 Hz e janelas de 0,2s (STA) e 45s (LTA) para maximizar a chegada das ondas P dos eventos. Também propomos a utilização de uma grade de nucleação cujo número mínimo de leituras para nucleação depende do número de estações a uma distância de 10 graus de cada nó da grade. Os resultados preliminares indicam que as modificações implementadas aprimoram a detecção automática de eventos sísmicos no Brasil, passando de 8 para 42 o número de eventos detectados automaticamente no ano de 2019.

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. Diversos projetos sucederam-se desde o início dos anos 1900, até que, em 1970, a sismologia no Brasil teve um avanço significativo com a formação de grupos sismológicos em diferentes regiões do país, com a finalidade de se estudar o perigo sísmico relacionado à usinas nucleares e as ocorrências de sismicidades induzidas por reservatórios de água (Bianchi et al., 2018).

Na atualidade, a RSBR conta com quase 100 estações sismográficas banda-larga com capacidade

de transmissão de dados em tempo real, alimentando continuamente os computadores de processamento nos 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). Cada centro sismológico é responsável pela instalação, operação e manutenção de estações sismográficas em determinadas regiões do território nacional, de modo que a rede é composta por quatro sub-redes (Figura 1), 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 um dos grupos e atuam de forma complementar à RSBR.

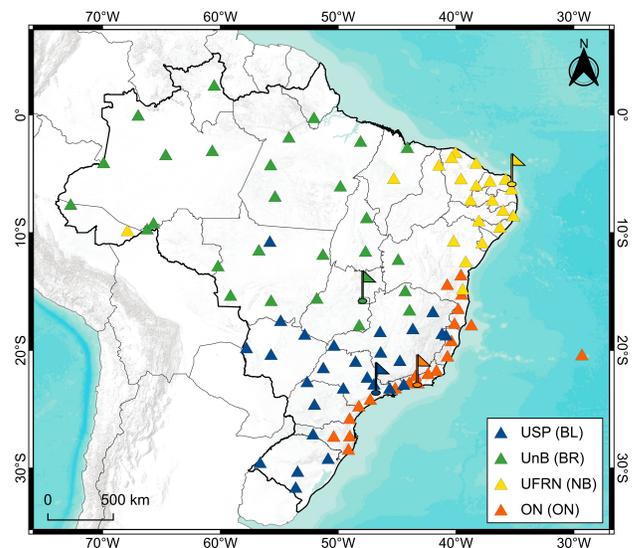


Figura 1 – Estações sismográficas (triângulos) e instituições (bandeiras) participantes da RSBR. Sub-redes e suas respectivas instituições responsáveis são diferenciadas por cores.

A partir da distribuição temporal das magnitudes dos eventos contidos no boletim sísmico da RSBR, Bianchi et al. (2018) estimou a detectabilidade de eventos sísmicos entre 1940 e 2016. Segundo o autor, a instalação da RSBR aperfeiçoou a detectabilidade para eventos de magnitudes a partir de 3,5 m_b para todo o Brasil. Além disso, desde 2012 a RSBR vem detectando o dobro de eventos com magnitudes entre 3,5 m_b e 4,0 m_b por ano, quando comparado ao que era detectado no período pré-RSBR. Para as regiões sudeste e nordeste, nas quais as densidades de estações são maiores, o limite de detecção obtido é de 3,0 m_b ou até menor.

Embora a RSBR utilize o *software* SeisComP desde 2009, apenas uma minoria dos eventos contidos no boletim sísmico brasileiro foram detectados automaticamente, sendo a maioria detectados manualmente através de análises de *dayplots* de estações sismográficas estrategicamente selecionadas.

Um dos fatores que podem aprimorar a detectabilidade de eventos sísmicos é a avaliação e correta configuração dos processos realizados pelo sistema de detecção automática. É fundamental que os principais parâmetros e processos de detecção e localização de eventos sejam estudados a fundo e devidamente aprimorados, de forma a otimizar o funcionamento do SeisComP no contexto da sismicidade brasileira.

Metodologia

Consultamos o boletim sísmico da RSBR, selecionando os eventos sísmicos ocorridos entre os anos de 2014 e 2021 para verificar a atual eficiência de detecção automática do SeisComP. Durante esse período, consta um total de 1550 eventos sísmicos no Brasil (excluindo as atividades sísmicas no Acre e na dorsal meso-oceânica, associadas à tectônica de placas). Algumas variáveis importantes desses eventos foram avaliadas, tais como o número de eventos com ao menos uma origem automática, o número de leituras automáticas de chegadas de ondas e as distâncias máximas em que os eventos foram detectados por uma estação sismográfica. Também foram calculadas as distâncias entre as estações da RSBR, a fim de verificar se o atual conjunto e disposição de estações é o maior fator limitante na detecção automática de sismos. Devido ao alto custo computacional para processar uma média anual de 600GB de dados em forma de ondas, as rodadas de testes sobre todos os dados foram realizadas apenas para o ano de 2019, sendo esse o período recente com maior número de estações em operação na RSBR.

O SeisComP

O funcionamento do SeisComP é baseado em módulos independentes, cada um desempenhando processamentos específicos que contribuem para a detecção e localização de um evento sísmico (Helmholtz-Centre Potsdam - GFZ German Research Centre for Geosciences and gempa GmbH, 2008). Nesse trabalho, focamos na otimização de dois módulos específicos, sendo eles:

- ***scautopick***: procura por anomalias, em forma de variações de amplitude, no dado em forma de onda. Aplica um algoritmo de razão STA/LTA (*short-time-average through long-time-average trigger*) que atua como um detector de anomalias de amplitudes com limite pré-definido. Ao detectar tais anomalias, utiliza um algoritmo AIC (*Akaike Information Criterion*) para criar uma leitura de chegada de onda (*pick*) baseada no instante de tempo obtido pelo STA/LTA; e

- ***scautoloc***: responsável por associar *picks*, nuclear e localizar eventos sísmicos de forma automática e em tempo-real. Realiza a análise contínua dos *picks* para nuclear um evento sísmico, ou seja, identificar combinações de *picks* que correspondam a um mesmo evento, considerando determinados critérios.

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, também buscando evitar ao máximo os falsos positivos. Também são nesses módulos que grande parte da capacidade de detecção de eventos com magnitudes relativamente baixas ($\leq 3,5 m_R$) pode ser aprimorada, ajustando parâmetros e fluxos de processamento que permitam que o SeisComP, inicialmente desenvolvido para sismos de escala global, também processe de forma satisfatória os eventos regionais.

Nesse trabalho, inicialmente otimizamos os parâmetros de leituras de chegadas de ondas, através da criação de um algoritmo que busca encontrar a melhor combinação de filtros passa-banda e valores de STA/LTA que aumente a razão sinal-ruído (SNR) das chegadas das ondas P dos eventos sísmicos. O algoritmo testa diversas variações dos parâmetros (valores de corte, em Hz, para o filtro passa-banda e comprimentos das janelas STA e LTA, em segundos), calculando o SNR e comparando os resultados em um gráfico, permitindo a definição dos valores a serem utilizados no SeisComP.

Em um segundo momento, também buscamos otimizar os fluxos de processamento no código-fonte do módulo *scautoloc*, no qual diversos parâmetros e processos pré-definidos ("*hard-coded*") são executados. Tais modificações foram feitas diretamente no código-fonte do *software*, uma vez que não estão acessíveis através das configurações disponíveis ao usuário final.

Resultados

Dos 1550 eventos analisados, apenas 97 (6,3%) possuem ao menos uma origem automática, o que indica que o SeisComP foi capaz de detectá-los sem interferência humana. A Figura 2 apresenta um mapa com os epicentros de todos os eventos avaliados, bem como o histograma com a distribuição desses eventos em termos de tipo de detecção (manual ou automática) e magnitude.

Os resultados mostram que a faixa de magnitudes que possui o maior número de eventos sísmicos (M1,0 a M3,0) também é o intervalo que possui a menor porcentagem de detecção automática. Dois fatores principais podem limitar a efetividade das detecções automáticas: i) o conjunto atual de estações, o qual pode ser insuficiente para monitorar a sismicidade de um país continental, e ii) as configurações padrão do SeisComP, que favorecem a detecção e localização de eventos telessísmicos.

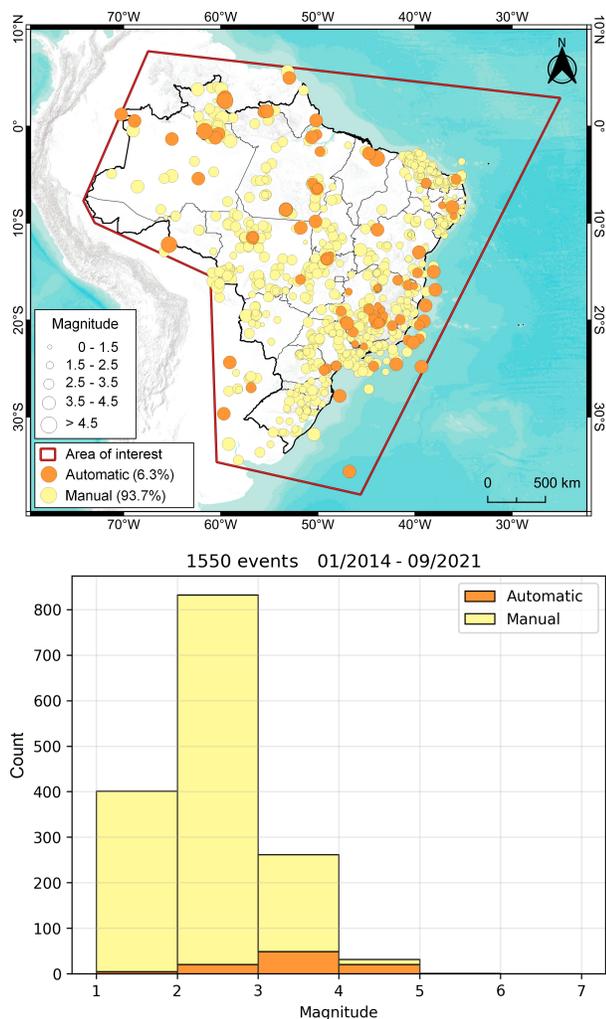


Figura 2 – Distribuição geográfica (acima) e em termos de tipo de detecção e magnitude (abaixo) dos 1550 eventos sísmicos contidos no boletim da RSBR entre 2014 e 2021.

Para verificar o primeiro fator sugerido, avaliamos o boletim sísmico da RSBR em termos das distâncias máximas registradas para cada faixa de magnitude, comparando-as com as distâncias mínimas entre as estações das sub-redes e da RSBR como um todo.

Como mostrado na Figura 3, a mediana das distâncias mínimas entre todas as estações da RSBR é de aproximadamente 159 km. Esse número torna-se mais representativo para as regiões sudeste e nordeste, dadas as maiores densidades de estações nessas regiões. Considerando que a sub-rede BR é praticamente a única cobrindo as regiões centro-oeste e norte, é razoável utilizar a mediana de suas mínimas distâncias (373 km) como referência para essa área.

Analisando as distâncias máximas registradas em relação às magnitudes dos eventos, verificamos que eventos sísmicos com magnitudes entre M1,0 e M1,5 são detectados, em mediana, a 184 km. Para magnitudes entre M2,0 e M2,5, esse valor aumenta para 420 km.

Dessa maneira, indicamos a possibilidade de aumentar a efetividade da detecção automática dos eventos sísmicos com magnitudes entre M2,0 e M2,5, desde que o segundo fator sugerido (otimização do *software* SeisComp) seja considerado.

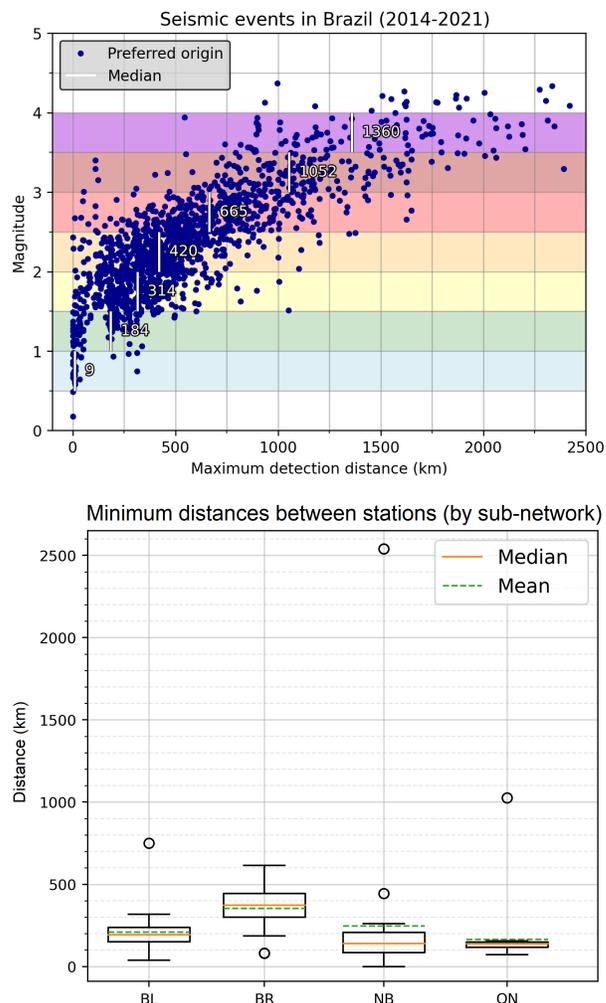


Figura 3 – (Acima) Gráfico de dispersão das distâncias máximas registradas por magnitudes. As linhas verticais brancas indicam os valores medianos de distância para cada faixa de magnitudes. (Abaixo) Diagrama de caixas das distâncias mínimas entre pares de estações das sub-redes da RSBR.

scautopick

Para otimizar a leitura de chegadas de ondas nas estações, utilizamos um algoritmo para buscar a combinação de filtro passa-banda e valores de STA/LTA que aumentem o SNR da chegada da onda P, a única utilizada pelo *scautoloc* para nucleação de eventos. A Figura 4 apresenta um exemplo de resultado de um evento sísmico (1,5 m_R) registrado na estação NBMA a aproximadamente 2,4 graus de distância. Pela análise apresentada, o filtro passa-banda ideal para realçar a onda P desse evento deve estar entre 5 Hz e 10 Hz,

aproximadamente. O comprimento da janela de LTA deve ser, idealmente, entre 30s e 40s, enquanto o comprimento da janela de STA apresentou melhores realces da onda P estando entre 0,12s e 0,4s.

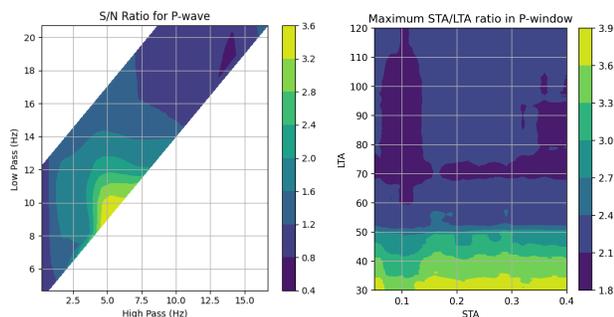


Figura 4 – Análise de combinações de filtros passa-banda (esquerda) e de comprimentos de janela STA e LTA (direita) para um único evento sísmico.

A Figura 5 apresenta dois sismogramas de um mesmo evento sísmico no Brasil, sendo um exemplo de como um filtro de frequências adequado aumenta significativamente a razão sinal/ruído das chegadas de ondas. Valores de SNR maiores aumentam a acurácia e precisão da detecção automática, portanto, o primeiro passo para calibrar o SeisComP para a sismicidade brasileira é propor um conjunto de filtros e parâmetros de STA/LTA que destaquem as anomalias de amplitude referentes a eventos sísmicos.

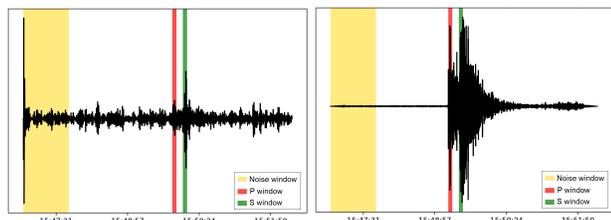


Figura 5 – Registro de um evento sísmico no Brasil, comparando o filtro de frequências padrão do SeisComP (esquerda) com o filtro proposto após a análise explicitada (direita). Nesse caso, os valores de corte para o filtro passa-banda foram 4,5 Hz e 8,5 Hz.

Após a análise das combinações do filtro passa-banda e dos valores de STA e LTA em 54 eventos de magnitudes entre M1,0 e M6,0, calculamos uma média simples dos resultados obtidos e passamos a utilizar os seguintes parâmetros no módulo *scautopick* em nossas rodadas de testes:

- Filtro de frequências passa-banda de quarta ordem, com cortes de 4,5 Hz e 10 Hz;
- Comprimentos de janelas de STA e LTA de 0,2s e 45s, respectivamente;
- Razão STA/LTA para ativar o *trigger*: 3; e
- Razão STA/LTA para desativar o *trigger*: 0.7.

scautoloc

Após a otimização dos parâmetros de leitura de chegadas de ondas, os quais são acessíveis através das configurações disponíveis ao usuário final, iniciamos o estudo do módulo *scautoloc*. Dois dos parâmetros fundamentais para a nucleação de eventos sísmicos são a malha de pontos e o modelo de velocidades que serão utilizados no processo.

A malha de pontos é um conjunto de pontos (a princípio, arbitrários) que cobrem a área de interesse de maneira suficientemente densa. Durante a nucleação, cada ponto é considerado um hipocentro hipotético para todas as leituras de chegadas de ondas. Podemos definir um número mínimo de leituras de chegadas de ondas e uma distância máxima de registro para que um evento seja nucleado em um determinado ponto, sendo que esses parâmetros devem ser ajustados de forma a compensar variações nas densidades de estações em cada região.

A Figura 6 mostra a malha de pontos padrão do SeisComP, comparada com a malha de pontos proposta para o monitoramento da sismicidade no Brasil pelo presente trabalho, com maior densidade de pontos em território nacional. Para a definição dos parâmetros de cada ponto, a malha proposta leva em consideração a densidade de estações sismográficas em uma distância de 10°. Embora não tenhamos interesse nos eventos andinos, é importante que o *software* considere pontos da malha nessa região, evitando que eventos sísmicos originados nos Andes sejam erroneamente nucleados no Brasil.

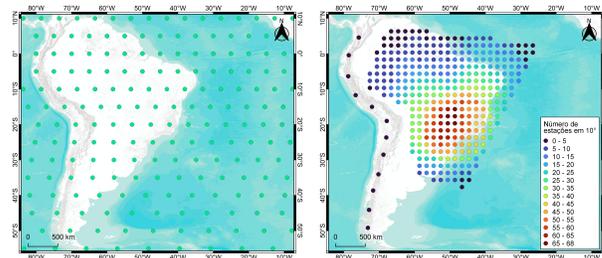


Figura 6 – Malha de pontos padrão do SeisComP (esquerda) e proposta (direita). As cores indicam a densidade de estações sismográficas em um raio de 10° ao redor do ponto.

Inúmeros parâmetros e fluxos de processamento não acessíveis ao usuário final (ou seja, contidos no código-fonte do *software*) possuem escolhas arbitrárias e que podem ser adequadas para eventos telessísmicos, entretanto, acabam prejudicando a localização automática de eventos regionais. Dentre eles, cabe destacar o fato de que o código altera alguns parâmetros inseridos pelo usuário (por exemplo, os limites de resíduos aceitáveis para as chegadas de onda). Também encontramos diversos “pesos” arbitrários para os resíduos e “pontuações” que definem se a origem obtida é satisfatória e deve ser reportada, além de etapas que acabam descartando origens boas desnecessariamente. Essa etapa do trabalho ainda está em andamento e deve ser concluída nos próximos meses.

Discussão

Para verificar se as mudanças propostas até o momento estão surtindo o efeito esperado, realizamos diversas rodadas de testes utilizando os dados em forma de ondas registradas pelas estações sismográficas no ano de 2019. A Figura 7 apresenta a comparação, com o boletim sísmico da RSBR, dos eventos sísmicos detectados automaticamente utilizando o SeisComP padrão e utilizando o SeisComP com as modificações propostas até o momento.

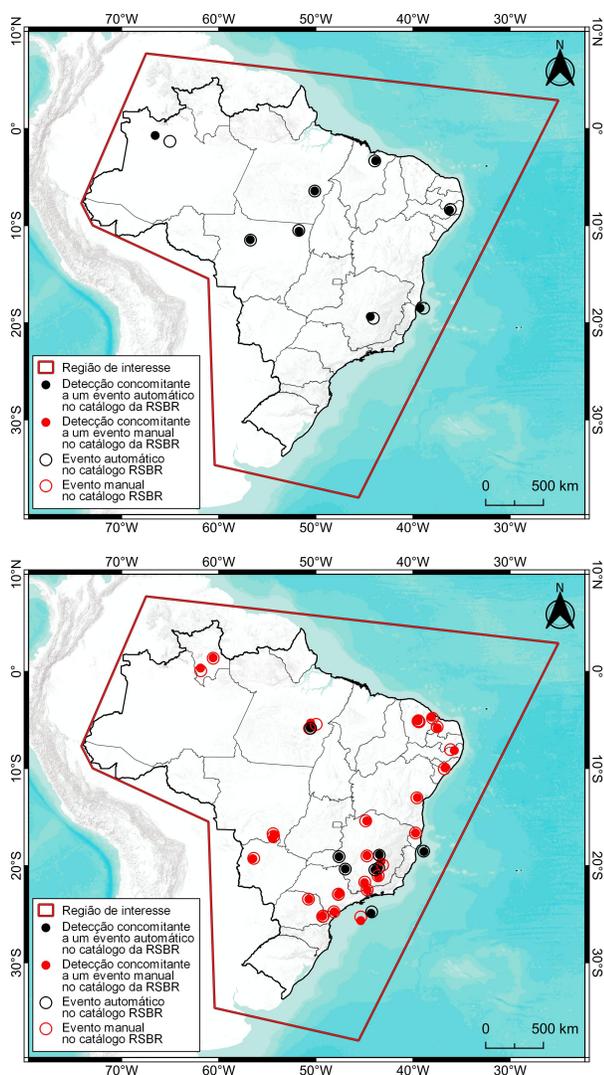


Figura 7 – Mapas de eventos sísmicos contidos no boletim da RSBR e que foram detectados automaticamente em nossos testes, utilizando-se o SeisComP padrão (acima) e o SeisComP modificado (abaixo).

Pode-se notar que alguns eventos detectados pelo SeisComP padrão não foram detectados com o SeisComP modificado. Isso se deve, principalmente, às limitações inicialmente propostas para evitar um grande número de falsos positivos, uma vez que, para aumentar a detectabilidade de eventos de baixas magnitudes,

aumenta-se a tendência de se nuclear eventos falsos. As principais limitações nesse sentido são (i) a distância máxima que um evento nucleado em um determinado ponto pode ser registrado e (ii) o número mínimo de estações em que o evento foi registrado. Esses parâmetros ainda estão em fase de testes, a fim de se aprimorar a razão entre detecções reais e falsos positivos.

Cabe ressaltar que os mapas da Figura 7 apresentam apenas os eventos sísmicos concomitantes ao boletim sísmico oficial da RSBR. Entretanto, verificamos que não constam no boletim da RSBR diversos eventos reais (naturais e detonações) detectados automaticamente pelo SeisComP modificado. Ao fim do projeto, espera-se poder incrementar o boletim de eventos sísmicos no Brasil com os resultados obtidos.

Dos eventos registrados automaticamente pelo SeisComP modificado, destacam-se suas boas qualidades quando seus epicentros são comparados com os do boletim sísmico da RSBR. A mediana das distâncias entre os epicentros calculados automaticamente e os do boletim da RSBR é de 10,5 km, indicando uma boa localização automática dos eventos.

Conclusões

Considerando que o projeto de mestrado que visa a otimização do *software* SeisComP para a detecção de eventos sísmicos no Brasil ainda está em andamento, nós estamos avaliando continuamente os resultados e propondo melhorias aos processos a cada rodada de testes realizada.

Existe uma grande quantidade de parâmetros e processos no SeisComP que devem ser estudados a fundo para garantir que o *software* funcione da melhor maneira possível em um contexto de sismicidade regional intraplaca, como é o caso da sismicidade brasileira.

Apresentamos alguns dos resultados preliminares obtidos até o momento, os quais mostram melhorias promissoras na detecção e localização automática de eventos sísmicos no Brasil. Buscamos diminuir ao máximo o número de falsos positivos, ao mesmo tempo em que aumentamos a acurácia dos eventos detectados.

Verificamos que alguns eventos detectados em nossas rodadas de testes não constam no boletim sísmico da RSBR, portanto, também torna-se um objetivo desse trabalho incrementar o boletim de eventos ocorridos em território brasileiro.

Para os próximos passos, finalizaremos a análise dos processos contidos no código-fonte do módulo *scatoloc* e dos parâmetros da malha de pontos utilizada na nucleação de eventos. Outro ponto importante no qual estamos trabalhando é o desenvolvimento de um modelo de velocidades atualizado para o Brasil, o que permitirá que os eventos detectados automaticamente tenham resíduos de tempo mais compatíveis com os parâmetros que estamos propondo para o fluxo de detecção.

Referências

Bianchi, M. et al., 2012. The Brazilian Seismographic Network: Historical Overview and Current Status, Summary of the Bulletin of the International Seismological Centre, vol. 49: 70., doi:10.5281/zenodo.998851.

Bianchi, M. et al., 2018. The Brazilian Seismographic Network (RSBR): Improving Seismic Monitoring in Brazil, Seismological Research Letters, vol. 89(2A): 452., doi: 10.1785/0220170227.

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, URL <https://www.seiscomp.de>.