

Integration des Seismic Denoiser in den Workflow unter SeisComP

Der *Seismic Denoiser* wird inzwischen beim GD in nahezu Echtzeit eingesetzt um die kontinuierlichen Originaldaten aus verschiedenen seismologischen Netzen zu filtern und in die Auswertung unter *SeisComP* zu integrieren. *SeisComP* ist eine etablierte Software zur Detektion und Auswertung seismischer Ereignisse, die in vielen Einrichtungen verwendet wird. Die routinemäßige Auswertung mit *SeisComP* befindet sich im GD NRW derzeit noch im Aufbau. In diesem Zusammenhang soll auch die routinemäßige Verwendung der mit dem *Seismic Denoiser* gefilterten Daten getestet werden.

Neben der Möglichkeit, den *Seismic Denoisers* zur Unterstützung der Routineauswertung mit dem bisher verwendeten Programm *SeiSan* einzusetzen, wurde ein Workflow entwickelt, der es ermöglicht, den *Seismic Denoiser* kontinuierlich und automatisiert im Ablauf der Seismologischen Auswertung unter *SeisComP* zu integrieren:

- Die an den Erdbebenstationen registrierten Daten werden aus dem seedlink-Buffer (Zwischenspeicher der ankommenden Wellenformdaten mithilfe des sogenannten seedlink-Protokolls) als Input für das *Denoising* gelesen. Hierzu wurde eigens ein Python-Skript entwickelt, das das Einlesen aus der speziellen Datenstruktur des Buffers, die Filterung unter Anwendung des *Seismic Denoisers* und das formatierte Schreiben in einen Zwischenordner organisiert. Mit Hilfe eines Cronjobs wird das Skript und somit das *Denoising* regelmäßig ausgeführt. Da aufgrund von Sicherheitsvorkehrungen der Datentransfer in das Landesnetz nicht per seedlink möglich ist, wurde neben dem Einlesen aus dem seedlink-Buffer auch das Lesen aus einer weiteren Datenstruktur ermöglicht.
- Da eine Mindestlänge des Seismogramms nötig ist und der Buffer nur abschnittsweise gefüllt wird, muss man eine gewisse zeitliche Verzögerung der mittels *Denoiser* gefilterten Daten akzeptieren. Es ist aber mit der Methode möglich, das *Denoising* alle zwei Minuten zu durchlaufen. Somit liegen die gefilterten Daten bereits nach einer geringen Verzögerung von wenigen Minuten vor.
- Gefilterte Daten werden anschließend unter Verwendung des *mseedscan*-Moduls wieder in das Processing mit *SeisComP* integriert. Die Inventory-xml Datei mit den Stations-Metadaten muss dafür zunächst um die Datenkanäle der gefilterten Daten erweitert werden.
- Das *SeisComP* Pipeline Konzept wird verwendet, um die Auswertung der gefilterten Daten in einem gesonderten Ablauf zu testen. Neben der Pipeline für das Picken und Lokalisieren auf Originaldaten, haben wir eine Pipeline erstellt, die sowohl Picks auf Originaldaten als auch Picks auf per *Seismic Denoiser* gefilterten Daten verwendet. Die STA/LTA Filter und andere Einstellungen für das *SeisComP* *scautopick* und *scatoloc* model wurde entsprechend für die gefilterten Daten angepasst. Automatisch lokalisierte Events werden als eigene Origin einem Event zugeordnet. Das heißt, Auswertungen der gefilterten Daten sind separat, aber direkt neben den ggf. vorhandenen Auswertungen auf Basis herkömmlicher Daten verfügbar.

Der Vergleich der gefilterten Daten (*Denoiser*) mit den Originaldaten zeigt eine Vielzahl von Picks, hervorgerufen durch unregelmäßig auftretende Störsignale. Dies hat anfangs zu Fehllokalisierungen geführt. Durch Optimierung der Pickereinstellungen konnte die Anzahl der Fehllokalisierungen jedoch vermindert werden. Eine weitere Optimierung der Einstellungen könnte mit zunehmender Erfahrung in der Routine-Arbeit möglich sein.

Dadurch, dass das *Denoising* im GD mit dem beschriebenen Verfahren kontinuierlich ausgeführt wird, stehen *denoised* Daten auch für die weitere manuelle Bearbeitung unter *SeisComP* Archiv zu Verfügung.

Modelltraining für die jeweiligen Stationen (GD und andere)

Für die Erdbebenstationen des GD NRW werden inzwischen individuell trainierte Modelle verwendet, die anfangs von der RUB berechnet wurden. Für Daten aus anderen Netzen wurde zunächst das von der RUB zur Verfügung gestellte allgemeine Modell verwendet.

Durch Einbinden einer NVIDIA-Umgebung unter der Linux-Maschine und entsprechende Anpassungen des Systems kann das Modelltraining mittlerweile beim GD selbständig und in akzeptabler Zeit durchgeführt werden.

Es ist zu erwarten, dass ein verbessertes Modelltraining ungewünschten Noise besser unterdrücken kann und dass im Fall einer dauerhaften, routinemäßigen Anwendung des *Denoisers* die Modelle regelmäßig nachtrainiert werden müssen, insbesondere an den Stationen, an denen durch bauliche Maßnahmen in der Umgebung neue charakteristische Störsignale hinzukommen. Dies können zum einen zusätzliche Windräder, aber auch andere Baumaßnahmen oder Veränderungen sein. Auch saisonale Einflüsse können über regelmäßiges Retraining berücksichtigt werden. Letztere sind vor allem bei unseren Stationen im Bereich von Talsperren oder Wasserbehältern zu vermuten.

Für die Verwendung des *Seismic Denoisers* im Routinebetrieb sollte das Retraining mit möglichst wenig zusätzlichem Aufwand ablaufen. Hierzu wurde ein automatisiertes Re-Training der Modelle in den Workflow unter *SeisComP* integriert:

- Hierzu werden Daten aus dem *SeisComP* Archiv genutzt, die für einen festgelegten Zeitraum dort vorgehalten werden.
- Das entwickelte Skript extrahiert Noisefenster aus dem *SeisComP* Archiv in einem gewünschten Zeitraum und speichert diese in ein Zwischenverzeichnis. Es wird auf das gleiche csv-File (mit den Kontrollparametern für die jeweiligen Datenkanäle) zurückgegriffen, welches auch vom automatischen Denoising verwendet wird. Auf diese Weise wird der Aufwand für die Organisation des Ablaufs bei sich ändernden oder zusätzlichen Stationen möglichst gering gehalten. Auch hier wird über einen Cronjob die regelmäßige Ausführung des Skripts organisiert.
- Das Retraining wird anschließend für alle verwendeten Modelle ausgeführt.
- Nach Prüfung der Ergebnisse können die alten Modelle überschrieben und von da an für das Denoising verwendet werden.
- Um saisonale Schwankungen am Stationsstandort berücksichtigen zu können, wird das Retraining zunächst einmal im Monat ausgeführt.
- Auch neue Stationen können in den Ablauf eingebunden werden, in dem man das Denoising mit dem allgemeinen Modell startet. Durch regelmäßigen Durchlauf des Retrainings wird das Modell mit der Zeit auf die neue Station angepasst.
- Allerdings fehlt für das Retraining bisher der Nachweis, dass dieses die Rauschunterdrückung effektiv verbessert. Durch Anwendung der Methode über einen längeren Zeitraum wird erwartet, dass der Nutzen des stationspezifische Retrainings überprüft werden kann. Gegebenenfalls sind hierzu weitere Testungen nötig. Hier sind insbesondere solche Stationen von Interesse, deren Stationsqualität sich durch den Zubau von Rauschquellen verschlechtert.

Vorläufige Ergebnisse Denoising

Bisher werden die Ergebnisse des *Denoising* ergänzend beim automatischen Picken und bei der Eventlokalisierung genutzt. Gute Ergebnisse sind beim automatischen Picken auf den denoiseden Spuren erkennbar. Ein beobachteter Vorläufereffekt bei den Phaseneinsätzen, speziell vor dem P-Einsatz, konnten durch eine geänderte Schwellwertsetzung und Zeitkorrektur (scatopick) beim STA/LTA Filter unter Kontrolle gebracht werden. Sprünge in den Zeitreihen, wie anfangs beim *Autodenoiser* der RUB beobachtet, konnten aufgrund einer veränderten Vorgehensweise nicht bzw. nur im geringen Ausmaß in den Daten des GD beobachtet werden. Vor dem Denoisen der Daten werden mehrere Zeitabschnitte (abhängig von der Zeiteinstellung des Cronjob) aus dem Buffer zusammengesetzt und gemeinsam denoised. Anschließend wird nur noch das neu hinzugekommene Zeitsegment per *mseedscan* in *SeisComP* eingelesen. Ein Vorteil dieser Vorgehensweise ist auch, dass die Daten im Buffer früher zur Verfügung stehen als SDS-Archivdaten, für die der Autodenoiser ausgelegt ist. Auf diese Weise ist es möglich, die denoiseden Daten quasi in Echtzeit für die Bearbeitung in *SeisComP* bereit zu stellen.

Automatische Lokalisierungen mithilfe der Pipelines liefern sehr gute Ergebnisse. Manuelle Relokalisierungen führen oft nur zu kleinen Verschiebungen der bestimmten Hypozentren. Auch gibt es einige Events, die nur über die denoised *SeisComP* Pipeline gefunden wurden. Hierzu muss berücksichtigt werden, dass zunächst mit den allgemeinen *SeisComP* scatopick und scatoloc Modulen (open source) gearbeitet wurde. Ein Erwerb des *SeisComP* scanloc Moduls ist geplant. Dieses Modul ist speziell für die Detektion lokaler und somit schwächerer Events entwickelt. Dabei ist auch berücksichtigt, dass kleinere Events auf wenigen Stationen registriert werden. Hier wäre die Frage, ob mit den per *Seismic Denoiser* gefilterten Daten zusätzliche kleinere Events automatisch lokalisiert werden, die über die normale Auswertung mit Originaldaten nicht gefunden werden.

Die Magnitudenbestimmung ist laut Bericht der RUB noch nicht zuverlässig und daher für den Routinebetrieb in der Form nicht nutzbar. Auch die vom GD NRW durchgeführten Auswertungen zeigen eine systematisch zu niedrige Magnitudenbestimmung auf Grundlage der mit dem *Denoiser* gefilterten Daten. Hier müssen weitere Untersuchungen zum Algorithmus unternommen werden. Zurzeit verwenden wir für eine zuverlässige Magnitudenbestimmung ausschließlich Originaldaten.

Um zukünftige Weiterentwicklungen des *Seismic Denoisers* nutzen zu können und um zukünftige Veränderungen im Messnetz zu beachten, müssen im Regelbetrieb Parameterdateien betreut sowie evtl. die programmierten Routinen angepasst werden. Neben der Software muss ebenfalls zusätzliche Hardware (insbesondere CUDA) gewartet/betreut werden. In der manuellen Auswertung sind die zusätzlich Zeitreihen zu beachten.

Fazit

Der *Seismic Denoiser* wurde erfolgreich in die zukünftigen Auswerteroutinen des seismologischen Regelbetriebs im Geologischen Dienst NRW integriert. Hierbei wurde die *Denoiser* Version (Stand Ende 2023) genutzt, die bisher ohne das erweiterte Modelltraining (s.u.) umgesetzt ist. Mit der Integration des *Denoisers* können voraussichtlich auch zukünftige Versionen des Filters auf die seismologischen Daten angewandt werden. Trotz der umfangreichen Automatisierungen werden im Regelbetrieb zusätzliche Arbeitszeit und Kosten für die Betreuung und der zusätzlich anfallenden Datenmenge erforderlich sein.

Anhand einiger Beispielen wurde gezeigt, dass nach einer Anwendung des *Denoisers* vielfach mehr Phaseneinsätze und Wellenformen von Erdbeben zu erkennen sind als ohne Anwendung des *Denoisers*. Jedoch liegt momentan keine Quantifizierung von detektierten Events (einschließlich Fehldetektionen) auf Basis gefilterter Daten im Vergleich zu detektierten Events auf Basis ungefilterter Daten vor. Inwieweit die Einwirkungen von WEA auf die Messungen von Erdbebenmessstationen kompensiert werden können, ist bisher nicht erfasst. Dabei ist auch weiterhin unklar, inwieweit Amplitude und Wellenform nach Anwendung des Filters korrekt sind. Eine Beurteilung hinsichtlich der Nutzung denoisder Daten zur Magnitudenbestimmung steht noch aus und muss nach weiterer Optimierung des *Denoisers* im Rahmen des Projekts erneut untersucht werden.