On-demand custom broadband synthetic seismograms

Journal Article

Author(s):

Krischer, Lion; Hutko, Alexander R.; van Driel, Martin; Stähler, Simon Christian ; Bahavar, Manochehr; Trabant, Chad; Nissen Meyer, Tarje

Publication date:

2017-04-19

Permanent link:

<https://doi.org/10.3929/ethz-b-000130223>

Rights / license:

In Copyright - Non-Commercial Use Permitted

Originally published in:

Seismological Research Letters 88(4), <https://doi.org/10.1785/0220160210>

Funding acknowledgement:

157133 - Seismology on Mars (SNF)

On-demand custom broadband synthetic seismograms

Lion Krischer¹, Alexander R. Hutko², Martin van Driel³, Simon Stähler^{4, 5}, Manochehr
Bahavar², Chad Trabant², and Tarje Nissen-Meyer⁶

¹Department of Earth and Environmental Sciences, Ludwig-Maximilians-University,
Munich, Germany (now: ETH Zurich, Switzerland)

²IRIS - Incorporated Research Institutions for Seismology, Seattle, WA, USA

³Department of Earth Sciences, ETH Zurich, Switzerland

⁴Department of Earth and Environmental Sciences, Ludwig-Maximilians-University,
Munich, Germany

⁵Leibniz-Institute for Baltic Sea Research, Rostock, Germany

⁶Department of Earth Sciences, University of Oxford, Oxford, UK

Corresponding author: Lion Krischer, Institute of Geophysics, ETH Zürich, Sonneggstr. 5, Zürich, 8092,
Switzerland. (lion.krischer@erdw.ethz.ch)

Abstract

We present a new web service, Syngine (<http://ds.iris.edu/ds/products/syngine/>), running at the IRIS DMC, that offers on-demand and custom tailored seismograms over a broad period range from 1 to 100 seconds, served over HTTP. The free service produces full seismic waveforms including effects like attenuation and anisotropy that are calculated in commonly used spherically symmetric Earth models (PREM, ak135-f, iasp91). Users can freely adjust sources and receivers, retrieve seismograms from finite sources, convolve with arbitrary source time functions, and download Green's functions suitable for moment tensor inversions. Syngine extracts and processes seismograms in as fast as fractions of a second making it suitable for applications demanding short iteration times and a large number of waveforms. For the first time, researchers without large computational resources or specialized knowledge can easily access high-quality, custom, broadband seismograms. In this manuscript we present the rationale and basic principles of our method, including its limitations. Additionally we demonstrate the features of Syngine and the included Earth models, showcase several applications, and discuss future possibilities.

1 Introduction

Synthetic seismograms are fundamental for a plethora of tasks in seismology, most notably for comparing to observed seismograms and thus testing hypotheses. A wide array of methods to calculate these synthetics have thus developed over the decades. They range from being very accurate but expensive to calculate to being very fast but using heavy approximations that might not be acceptable for a given purpose. Additionally, many seismological software packages are difficult to use, requiring significant expert knowledge to acquire trustworthy results, especially when significant changes to input parameters are required, such as new Earth models.

In this manuscript, we introduce Syngine (<http://ds.iris.edu/ds/products/syngine/>), a web service to calculate custom-tailored, accurate seismograms sent to users upon request with return times as fast as fractions of a second. Syngine delivers full seismograms for spherically symmetric Earth models with anisotropic and viscoelastic rheologies in a period band between 1 and 100 seconds. The source-receiver geometry can be arbitrarily chosen under the constraint that the receiver is at the surface of the Earth. Further features include seismograms from finite sources, convolutions with arbitrary source time functions, the download of independent components to construct arbitrary source mechanisms with simple calculations, and windowing traces around seismic phases. A major focus of the project is usability, reliability, trustworthiness, and accuracy of the calculated seismograms.

The sole prerequisite to using Syngine is internet access. Thereby, users will not have to perform an expensive numerical simulation or store very large databases. This equips anyone, including researchers and groups that had no previous access to synthetic waveforms, either due to lack of computational facilities or knowledge, with access to broadband, high-accuracy, full synthetic seismic waveforms.

This paper is organized as follows: In section 2 we introduce the mathematical, numerical, and technical

methodology used to generate the synthetic seismograms. Section 3 presents the features of the Syngine service and section 4 discusses the rationale behind the offered Earth models providing guidance for their use. Finally we discuss potential applications in section 5 and the big picture in section 6. The code to (re)create all figures except figures 7 and 9 is part of the electronic supplements in the form of interactive Jupyter notebooks (Pérez and Granger, 2007) and additionally also on <http://seismo-live.org>.

2 Methodology

In this section we present a short and intuitive introduction to the numerical and mathematical methods used to generate and extract the synthetic seismograms. The fundamental methodology has been described at length in the literature and we will refer interested parties to relevant works where appropriate. We aim to provide readers the necessary knowledge to understand the capabilities and also the limitations and trade-offs of our scheme, enabling them to judge its applicability to arbitrary problems. In a nutshell the response of the medium to two forces - a vertical and a horizontal force source at the surface of the Earth - is simulated and recorded in a two dimensional domain. By exploiting the reciprocity of the Green's functions it is possible to swap source and receiver and, assuming a spherically symmetric medium, reconstruct that response between any two points on the planet given that the receiver is positioned at a fixed radius. In most instances this will either be the surface of the Earth or the bottom of the oceans. The method employed by Instaseis (<http://instaseis.net>, van Driel et al. (2015)) assures that the seismic wavefield is a full solution to the elastic wave equation in three dimensions, the only limitation is the required spherical symmetry of both the elastic properties and the domain.

2.1 Generation of the waveform databases

The first step is to compute databases of accurate waveforms using AxiSEM. AxiSEM (<http://axisem.info>, Nissen-Meyer et al. (2007, 2014)) employs a spectral element scheme (e.g. Komatitsch and Tromp, 1999) to propagate global seismic waves in axially symmetric media. This assumption allows the analytic decomposition of the 3-D wavefield into several 2-D wavefields. Thus only a 2-D numerical problem must be solved which is orders of magnitudes cheaper and therefore enables the calculation and storage of global wavefields at high frequencies. The final 3-D wavefield nonetheless contains all effects of a viscoelastic rheology like attenuation (van Driel and Nissen-Meyer, 2014a) and anisotropy (van Driel and Nissen-Meyer, 2014b). An important limitation to keep in mind is that the seismograms only capture the relevant physics up to a period of about 100 seconds. At longer periods, effects such as gravitation and Earth rotation play an increasingly important role (see e.g. Komatitsch and Tromp (2002)) but are, at this point, not taken into account in AxiSEM. The overall effect of the neglected physics can be comparatively small but in some scenarios it does matter and it is thus crucial to be aware of it. The short-period bound is dictated by the numerical mesh and can vary for each database. Strikingly, for a full database encapsulating all possible source-receiver geometries, only two simulations are required due to reciprocity in the wave equation (van Driel et al. (2015)): one for a vertical force source and one for a horizontal force source, both located at the surface of the planet. The displacement response for each is stored at every point of the numerical simulation grid up to a maximum depth. There is no limitation to the recorded depth range but greater ranges result in significantly increased database sizes. As the recorded range determines the possible source depths it is natural to limit it to the maximum depth of naturally occurring seismicity.

2.2 Seismogram extraction

Exploiting the reciprocity now allows one to place a source anywhere in the recorded region where the initial force sources correspond to the different components of a receiver. Recalling that we simulated in a spherically symmetric medium, we can place the reciprocal receiver anywhere on the planet and extract its response to sources throughout the recorded depth range. The stored displacements allow the on-the-fly reconstruction of the full strain tensor and resulting seismograms therefore have all three components of a fully 3-D wavefield. Storing the displacement on the grid of the numerical simulation and subsequent on-the-fly evaluation of the same basis functions as used in AxiSEM means that the spatial interpolation adds no additional error and is as accurate as the spectral-element simulation for arbitrary source-receiver locations independent of the numerical grid. The wavefields are temporally downsampled to four samples per mesh period which captures the spectral content of the wavefield down to around -80dB with respect to velocity. Seismograms are later upsampled again with a tapered sinc function which approximates an optimal reconstruction filter. This extraction is performed by Instaseis, the final seismograms are amended with meta information and serialized to common file formats by ObsPy (<http://obspy.org>, Megies et al. (2011); Krischer et al. (2015)).

This combination of tools allows the extraction of broadband three-dimensional seismograms propagating through spherically symmetric spheres in as little as fractions of a second with negligible spatial and temporal interpolation errors. Downsides of Instaseis as a stand-alone tool are the demanding computational cost for each user to generate the initial AxiSEM databases, and their unwieldy file sizes. Syngine, the topic of this manuscript, tackles both these problems.

3 Features

Syngine is a web service offering convenient and fast access to synthetic seismograms, as described in the previous section, over the HTTP protocol. Waveforms can be downloaded one at a time or in bulk.

3.1 Seismograms

Seismograms constitute the very core of Syngine. Each seismogram represents the response of the medium at one receiver to a particular source. Syngine is used by constructing a web URL that encodes the source, the receiver, and additional optional parameters. Accessing this URL either with a script or a web browser triggers the server-side extraction of the requested seismogram(s) which are then sent to the users. All available parameters are presented in detail in table 1, a few of them are described in more detail in this section. Parameters are roughly grouped into source, receiver, and miscellaneous parameters.

Sources always need a location specified by geographic latitude, longitude, depth values and a mechanism. Mechanisms can be defined by passing the 6 independent moment tensor components M_{rr} , M_{tt} , M_{pp} , M_{rt} , M_{rp} , and M_{tp} or by giving strike, dip, rake, and the scalar seismic moment M_0 . Alternatively the source mechanism can be a vectorial single force which is useful, for example, for noise studies (see e.g. [Basini et al., 2013](#)). For convenience all of these parameters can be replaced by passing an event identifier from an event catalog. Syngine at the time of writing supports event source lookups in the GCMT catalog ([Ekström et al., 2012](#)) and the USGS finite fault model database; other catalogs may be added in the future.

Receivers also require coordinates and must always be located at a fixed radius depending on the database, usually the surface of the Earth. Thus only latitude and longitude coordinates are required for a request. Network and station codes can be passed instead and they will be used to locate coordinates of the corresponding real stations in the IRIS database. Wildcards can be applied to download a large number of synthetics at once.

The following query for example will download seismograms for all stations starting with A from the virtual GSN network: `...&network=_GSN&station=A*&...`

Convenience parameters such as a data scaling factor or a custom file name label are grouped under miscellaneous. The Earth model from which to extract seismograms is also given as a parameter. It specifies the velocity model as well as the frequency content of the database. Additionally any combinations of vertical, north, east, radial, and transverse receiver components can be requested either in displacement, velocity, or acceleration, each in SI units. Last but not least the seismograms can be resampled to any sampling rate larger than the database sampling rate. The resampling employs a very high quality reconstruction algorithm and is careful to avoid edge effects, thus we encourage users to make use of this functionality.

Seismograms are returned either as a ZIP archive of SAC ([Helffrich et al., 2013](#)) files or as a single MiniSEED ([Incorporated Research Institutions for Seismology \(IRIS\), 2012](#)) file. Both have all their important header values filled to aid in data organization and mimicking recorded data.

3.2 Geographical coordinates

Exploiting reciprocity necessitates the use of a spherically symmetric planet unlike the real Earth and Instaseis as well as AxiSEM internally use a spherical geocentric coordinate system. Most seismological applications and data sets on the other hand define geographical coordinates on the WGS84 ellipsoid ([National Imagery and Mapping Agency, 2000](#)) and we follow that convention and adapt it to our internally used spherical symmetry.

All geographical coordinates that users specify are assumed to be given in WGS84 and are, prior to seismogram extraction, converted to spherical, geocentric coordinates. This effectively is a zeroth order ellipticity correction for comparison with real Earth data. Coordinates passed back to the users, for example in the SAC file headers, are also in WGS84. This only affects the latitude which is usually defined as the angle between the surface normal at a point and the surface normal at the equator at the same meridian. This differs for spherical

and elliptical coordinate systems. The difference is usually fairly small and many applications are not affected by it, but it is vital to understand how Syngine handles coordinates.

3.3 Phase relative times

A multitude of applications only require information about certain seismic phases. To this end, the Syngine service supports the specification of start and end times relative to any phase arrival. We use a port of the TauP Toolkit ([Crotwell et al., 1999](#)) in ObsPy ([Beyreuther et al., 2010](#)) which employs the $\tau - p$ method ([Buland and Chapman, 1983](#)) used to calculate theoretical arrival times. It supports arbitrary phase names within its syntactical and semantic limitations and performs the calculations in a spherical coordinate system for better accuracy compared to earlier implementations. Generating phase-windowed seismograms can greatly reduce the amount of data that needs to be transferred. Figure 1 shows a record section with phase relative time settings.

3.4 Source time functions

Syngine relies on synthetics calculated with the spectral element method. Sources for such simulations cannot be true delta functions as this is numerically unstable. A narrow Gaussian smoothly resembling a delta peak is used instead. Considering the band-limited nature of the signal, this is a good approximation. Also, each database contains a time series of the used slip and slip rates to document the embedded source. Calculating seismograms with new slip rate/source time functions (STF) requires the deconvolution of the database's STF from the signal followed by a convolution with the desired STF. Performing this in a single step yields numerical stability under some assumptions regarding the frequency content of both STFs: The new STF must not contain frequencies that the database STF does not have. We disallow STFs that have a finer sampling than the database STF which contains all simulated frequencies and thus almost guarantees this. If the need arises

we'll implement some further stabilizing algorithms.

New seismograms u_j having an STF s_j are calculated from the database seismograms u_i with the database STF s_i as follows:

$$u_j = \mathcal{F}^{-1} \left(\frac{\mathcal{F}(\omega u_i) \mathcal{F}(s_j)}{\mathcal{F}(s_i)} \right) \quad (1)$$

$\mathcal{F}$ and $\mathcal{F}^{-1}$ are the forward and inverse (fast) Fourier transforms, and ω is a tapering function to make sure the initial seismogram ends with zero.

Syngine can perform this process, which we refer to as reconvolution, in a stable manner that also deals with subtle issues such as time shifts. Users can either upload their own source time function (please see table 2 and the Syngine documentation for instructions) or use the `sourcewidth` parameter which reconvolves the seismograms with a Gaussian defined as

$$s = \frac{4}{a\sqrt{\pi}} e^{-\frac{16}{a^2}(t-t_0)^2} \quad (2)$$

with t being the time, t_0 the offset, and a the source width. See figure 2 for plots of this function as well as its effect on seismograms. The peak of the Gaussian will be at the origin time, this results in no apparent time shift but some acausal effects. This for example can be used to approximate the triangular moment rate functions used in the GCMT catalog (Ekström et al., 2012) in which case the `sourcewidth` parameter is twice the half-duration specified in the catalog.

3.5 Finite source seismograms

Although very convenient it is oftentimes insufficient to describe seismic wavefields as originating from single points in space. Real earthquakes, especially larger ones, have rupture surfaces and the point source approx-

imation breaks down. The superposition principle applied to linear elastodynamics allows the calculation of finite-source seismograms for kinematic sources via summation of a number of distributed point sources, each with their independent rise time. Finite source descriptions usually originate from kinematic source inversion, but finite-source models coming from dynamic rupture simulations are emerging as well. The SRCMOD database ([Mai and Thingbaijam, 2014](#)) collects Finite Fault Models (FFM), other groups such as the US Geological Survey (USGS) offer FFM solutions for download. Syngine supports the calculation of seismograms from FFMs. As of the writing of this manuscript it can use models described in the USGS "param" file format; other formats might be supported in the future.

The parameters are largely similar to the standard Syngine functionality for seismograms. Phase-relative times are calculated from the sources' hypo-centers, e.g. the patch with the first onset. A source is completely defined by specifying a USGS param file, no other source parameters are required. For the actual calculation of the finite source seismograms we have made some choices that may not be immediately obvious. Upon providing a FFM as a source, the following sequence of events occurs on the server side:

1. The slip rate of each point source is defined as an asymmetric cosine function with a certain rupture, rise, and fall time. We sample it at 10 Hz for a thousand seconds - this limits the maximum rise and fall times. Rise and fall times smaller than one second will be set to one second to make sure it can be accurately sampled. The cosine function is specific to the USGS param file, other FFM file formats might define a separate STF for each point which would be evaluated here.
2. Each sampled slip rate is zero padded with a number of samples at the beginning and the end (the additional time shift is accounted for later). This is done to avoid running into boundary issues with the following filter.
3. A fourth order Butterworth filter is applied twice (forwards and backwards) resulting in a zero phase

filter. The corner frequency is the dominant frequency of the database. This assures we don't introduce frequencies in the convolution that we cannot propagate in the numerical simulation.

4. The seismograms for all point sources are calculated, time-shifted, convolved, and stacked.

The resulting seismograms are sent to the users in the same fashion as the point source seismograms. This has a profound impact on the seismograms, see Figure 3 for plots of seismograms comparing a point source to a finite source.

3.6 Green's functions

Moment tensor inversions are performed by successively modifying the source mechanism to improve the best fit between data and synthetics until it has converged. This requires the generation of a large number of seismograms from a single origin with varying source mechanisms. A common way of doing this is to reconstruct seismograms as a linear combination of Green's function from fundamental sources. Syngine supports returning Green's functions in the convention introduced by (Minson and Dreger, 2008), which are readily used by SeisComP3 and other software for moment tensor inversion. Seismograms for an arbitrary source mechanism can be calculated by a linear combination of waveforms from these four sources: A vertical strike-slip fault (SS), a vertical dip-slip fault (DS), a dip-slip fault with a dip of 45° (DD), and an explosive source (EP). Formulas for vertical (u_z), radial (u_r), and transverse components (u_t) are then:

$$\begin{aligned}
 u_z = & M_{tt} \left[\frac{u_{Z,SS}}{2} \cos(2az) - \frac{u_{Z,DD}}{6} + \frac{u_{Z,EP}}{3} \right] + M_{pp} \left[-\frac{u_{Z,SS}}{2} \cos(2az) - \frac{u_{Z,DD}}{6} + \frac{u_{Z,EP}}{3} \right] \\
 & + M_{rr} \left[\frac{u_{Z,DD}}{3} + \frac{u_{Z,EP}}{3} \right] + M_{tp} [u_{Z,SS} \sin(2az)] + M_{rt} [u_{Z,DS} \cos(az)] + M_{rp} [u_{Z,DS} \sin(az)]
 \end{aligned} \tag{3}$$

$$\begin{aligned}
u_r = & M_{tt} \left[\frac{u_{R,SS}}{2} \cos(2az) - \frac{u_{R,DD}}{6} + \frac{u_{R,EP}}{3} \right] + M_{pp} \left[-\frac{u_{R,SS}}{2} \cos(2az) - \frac{u_{R,DD}}{6} + \frac{u_{R,EP}}{3} \right] \\
& + M_{rr} \left[\frac{u_{R,DD}}{3} + \frac{u_{R,EP}}{3} \right] + M_{tp} [u_{R,SS} \sin(2az)] + M_{rt} [u_{R,DS} \cos(az)] + M_{rp} [u_{R,DS} \sin(az)]
\end{aligned} \tag{4}$$

$$\begin{aligned}
u_t = & M_{tt} \left[\frac{u_{T,SS}}{2} \sin(2az) \right] - M_{pp} \left[\frac{u_{T,SS}}{2} \sin(2az) \right] \\
& - M_{tp} [u_{T,SS} \cos(2az)] + M_{rt} [u_{T,DS} \sin(az)] - M_{rp} [u_{T,DS} \cos(az)]
\end{aligned} \tag{5}$$

where M_{rr} , M_{tt} , M_{pp} , M_{rt} , M_{rp} , and M_{tp} are the moment tensor components, $u_{X,ZZ}$, $u_{X,DS}$, $u_{X,DD}$, and $u_{X,ES}$ seismograms from one of the elementary sources on the X component, and az is the source-receiver azimuth. See (Minson and Dreger, 2008) for a derivation and further explanation.

Syngine's Green's function service returns these elemental source seismograms for an arbitrary source-receiver geometry. Usage is very similar to requesting normal seismograms except that the source-receiver geometry is specified by epicentral distance and source depth. Table 3 lists all parameters and the supplementary material contains a code example on how to use it. Note that the Green's functions returned by Syngine are band-limited with the delta peak being replaced by a narrow Gaussian moment rate function as discussed in previous sections.

3.7 Meta information and documentation

The service will evolve in the future and it offers an interface to query available models and detailed information for each database. This information includes the frequency range within which each model can accurately deliver seismograms, the embedded source slip rate and other information such as the velocity model and the version of AxisEM used to generate the model. If data is requested as zipped SAC files it will also contain

a log file describing the status for each seismogram requested, including errors for those that could not be generated.

The website (<http://ds.iris.edu/ds/products/syngine/>) contains extensive documentation for the service with detailed information about each of the parameters, example queries, usage guides, tutorials, and an interactive URL Builder to aid in constructing custom queries. Furthermore, all databases are offered for download directly from the IRIS DMC. The heaviest users can therefore run their own Instaseis server to extract a very large number of seismograms without overloading the service and avoiding the latency of sending queries over the internet.

4 Available Earth models

Earth models, no matter if 1D, 2D, or 3D, are constructed upon an inversion of seismic data which requires a number of choices on data, modeling and inversion scheme. Inevitably, different models therefore satisfy different aspects of the inverse problem to a varying degrees. For instance in the spherically symmetric case, the PREM model has been constructed for a reference frequency and includes long period normal modes, whereas the iasp91 model is derived for body waves. Some models may include anisotropic elastic parameters or density, whereas others are isotropic. Even if laterally averaged, some may work better for continental paths, some for oceanic. As such, it is paramount for a generic webservice such as syngine to accommodate the inevitable variety of existent models and therefore waveforms. Syngine can currently calculate seismograms propagating through seven widely used different Earth models. At the time of writing these are `ak135f_1s`, `ak135f_2s`, `iasp91_2s`, `prem_a_2s`, `prem_i_2s`, `ak135f_5s`, `prem_a_5s`, `prem_a_10s`, and `prem_a_20s`. The first part denotes the used Earth model and the final number the smallest period to which synthetics are accurately modelled. `_i_` and `_a_` denote isotropic and anisotropic variants. `ak135-f` is an isotropic variant of the `ak135` velocity model (Kennett et al., 1995) with the attenuation and density model

taken from (Montagner and Kennett, 1996). iasp91 (Kennett and Engdahl, 1991) and PREM (Dziewonski and Anderson, 1981) are other well known Earth models, see figure 4 for comparison of the models and figure 5 for seismograms calculated in all the available models. Currently all models are global and continental - we discuss future plans to include other models in section 7.

5 Applications

Potential applications for the Syngine service are numerous. We hope that providing an extremely low barrier for access to high-quality synthetic seismograms will spark a number of new uses we have not yet considered. In general, Syngine is suitable for problems that require the calculation of a reasonable amount of seismograms. The exact number depends on the required frequency content as well as the available internet connection. Studies that require hundreds of thousands of seismograms, for example fully probabilistic source inversions, are better accommodated by Instaseis directly. This section showcases a few potential uses to demonstrate Syngine’s applicability to real world problems.

5.1 Algorithm test bed

Newly developed algorithms have to be tested and benchmarked against synthetic data with fully known and controlled properties. Most wave solvers have not been designed to be used by non-experts and the oftentimes inadequate usability of scientific software (Brown et al., 2015) brings about substantial efforts in order to actually acquire synthetic seismograms. Furthermore, it is very easy to misconfigure software resulting in data that might initially look fine but contains numerous subtle problems. Syngine and the stack it depends on have been thoroughly vetted and benchmarked against established solutions. Additionally, all input parameters are automatically examined and checked for consistency. Data acquired within the constraints presented in this

paper should be accurate. Algorithms that can be validated with Syngine include array processing, phase picking, event locators, back-projectors, and many others.

5.2 Data quality control

Assessing the quality of recorded data remains a major challenge for network operators and all practicing seismologists handling observed seismograms. Known data problems range from cross-talk amongst recording channels, incorrect sensor orientations to timing errors and erroneous instrument characteristics. Some, like missing or clipped data, are fairly simple to recognize. Others are much harder to detect. Efforts like the IRIS MUSTANG project (<http://service.iris.edu/mustang>) are being pursued with the goal of measuring data characteristics that can be used to quantify data quality. Syngine offers the possibility to augment these quality measurements with comparisons of the observed data to synthetic seismograms where all effects are fully known. Systematic and unexpected deviations from the synthetic seismograms can thus be found and investigated.

An example are small time shifts that are typically introduced by clock errors. The described approach can be applied to various other quality metrics. This is an example demonstrating the feasibility of the approach but it can well be expanded to be performed in a fully-automatic fashion. Figure 6 visually illustrates the chosen strategy.

We compare observed and synthetic P phase seismograms with a cross-correlation technique yielding time shifts with subsample accuracy (Deichmann and Garcia-Fernandez, 1992). The real data's instrument responses are deconvolved and both synthetic and real data are bandpass filtered with an 8th order zero-phase Butterworth filter to a period band between 5.0 to 12.5 seconds. For each event we subtract the mean time shift from the measured time shifts to account for effects like Earth's ellipticity, unknown and unmodelled 3D velocity structure, and faulty origin times and depths. These will affect all seismograms approximately equally

which is a fact exploited by many techniques, e.g. [Waldhauser and Ellsworth \(2000\)](#). We call this the reference data set and the results are average time shifts per station. Differences amongst these are due to topography and local velocity structure.

Measuring the time shifts for other events and comparing these to the reference data should result in very small differences if the station times are consistent. Otherwise there is some mismatch and likely an error with the time information. Figure 6 represent both these cases.

These calculations, including the collection of synthetic seismograms from Syngine, can be performed in a matter of seconds making it feasible to be integrated in a continuously running quality control system. The presented approach is a simplified proof of concept but already works quite well. Including more and especially local events, other phases, and more advanced processing and statistics will likely improve the assessment. Note the similarity between observed and synthetic seismograms including matching amplitudes. The proposed technique could thus be extended to check for correct polarities and changing site-effects.

5.3 Stability testing in source inversion

The modularity of Syngine allows to quickly test seismological codes versus different velocity models. As an example, we show the result of an inversion for moment tensor and source time function based on waveforms of P-waves. The inversion is based on the method described in [Sigloch and Nolet \(2006\)](#); [Stähler et al. \(2012\)](#). It alternates between updating the moment tensor and the source time function by a joint deconvolution of all Green's functions from the measured P-waveforms. The waveforms are used in a 50 second time window around the P-arrival. The code does a grid-search over all plausible depths (1 to 30 km) and choses the one with lowest misfit. The code is available from <https://github.com/seismology/stfinv>. The inclusion of Instaseis allows it to switch between local waveform databases or the ones provided by Syngine. Since the waveform of the P - pP - sP wavetrain, constrains the depth well, but also depends strongly on the

velocity model of the crust, different velocity models will result in different estimates for depth and source time function. Inverting the same earthquake with different velocity models provides a qualitative estimate of the stability of the result, which is shown in figure 7.

5.4 Education

Beyond application as a research resource, Syngine serves as a valuable tool for education. Concepts explained by instructors such as polarity of first arrivals, surface wave amplitudes depending on the hypocentral depth, or phase triplications can be visually and interactively discovered and explored by not only by seismologists and undergraduate students. The only requirement is a working internet connection and the capability to read and plot seismograms. Figure 8 demonstrates two such possibilities. More generally, the intuitive and interactive nature of the tool opens the floor to explore this for school and museum projects.

5.5 Backprojection

For large earthquakes, backprojection rupture imaging (Ishii et al., 2005; Trabant et al., 2012) is a useful complement to finite-fault modeling since the rupture velocity often trades off with length in the later approach while it can be inferred independently using the former. In this example of Syngine's flexibility and ease of use, backprojection results using data are compared with those using equivalent FFM synthetics using the multi-segment USGS model (earthquake.usgs.gov/earthquakes/eventpage/usp000g650#finite-fault). Such a comparison can also provide insight into interpreting backprojection results when they are not clear due to insufficient station geometry or complicated Green's Functions. Downloading the 25 FFM synthetics needed for this analysis can take as little as three minutes with only a single URL request using the appropriate USGS eventid and network name or a users's station list. This type of exercise can be easily repeated to help guide station configurations for future seismic networks or temporary experiments. Similarly,

such an effort can be used to assess resolution issues with perturbations in either the source model or receiver array.

6 Discussion

Services such as Syngine that remotely calculate on-the-fly whatever is needed for a particular application are very likely to become more prevalent in the future. Syngine is, to our knowledge, the first service of its kind that grants access to high-quality and customizable synthetic seismograms for Earth with a simple web interface. ETH Zurich recently launched a service in the same vein to generate synthetic seismograms for various proposed Mars models ([Ceylan et al., 2016](#)). Previously this required high performance computing facilities and considerable technical knowledge, severely limiting its de-facto availability within the research community.

The closest undertaking comparable to Syngine is the ShakeMovie project ([Tromp et al., 2010](#)). ShakeMovie produces more physically accurate synthetic data as it includes three dimensional Earth models, topography of the surface, and the effects of gravity and rotation. The major disadvantage is that ShakeMovie can only offer seismograms from a given list of earthquakes and receivers, and it is limited to comparatively low frequencies. Syngine, on the other hand, grants full flexibility in the source-receiver geometry, accomodates frequencies up to 1 Hz, has numerous different models, but is restricted to spherical symmetry. As the 1D structure explains seismic data, especially body waves, rather well there are a large number of use cases where this limitation is acceptable and the gained flexibility is a worthy trade-off. The combination of capabilities provided by ShakeMovie and Syngine will remain computationally intractable for the foreseeable future.

7 Conclusion and perspective

We present Syngine, a web service to download on-demand, customized, high-frequency, fully three dimensional seismograms calculated through spherically symmetric, laterally averaged, anisotropic, and viscoelastic Earth models. Source and receiver parameters (locations and mechanism) can be freely chosen as long as the receivers are located at the surface of the Earth. Extraction and calculation of seismograms takes less then a second, granting near instant results.

The present state of Syngine is its first realization and it comes with a number of different global Earth models and seismograms in various frequency bands. Future plans include to add more models, some also with water layers, various lithospheric structures, and regional-specific, even higher frequency models/databases. The web interface is sufficiently generic to accommodate synthetics calculated by other means, for example via normal mode summation if gravity effects are of interest. Syngine is currently an IRIS DMC service but we welcome implementations from other institutions, ideally with compatible interfaces. All software components required to a Syngine web service are openly available.

To increase reproducibility, the code to (re)create all figures except figures 7 and 9 is part of the electronic supplements in the form of interactive Jupyter notebooks and additionally also on <http://seismo-live.org>.

Data and Resources

Observed waveform and some used event data can be obtained from the IRIS Data Management Center at <http://www.iris.edu> (last accessed November 2016). This includes waveform data from the Global Seismograph Network (GSN - IRIS/USGS, doi:10.7914/SN/IU), the Southern California Seismic Network (Caltech, doi:10.7914/SN/CI), and the Australian National Seismograph Network. Other earthquake data has been retrieved from the USGS earthquake catalog at <http://earthquake.usgs.gov/fdsnws/event/1/> (last accessed November 2016) and the finite source data for figure 3 is from the USGS event page at <https://earthquake.usgs.gov/earthquakes/eventpage/us20002926#finite-fault> (last accessed October 2016).

Acknowledgements

We thank Editor Zhigang Peng, as well as Mark Panning and an anonymous reviewer for their thoughtful and constructive comments, which helped improve the manuscript. L. Krischer was partially supported by the EU-FP7 VERCE project (number 283543) and also acknowledges support from the EU-funded EPOS project. M. van Driel was supported by grants from the Swiss National Science Foundation (SNF-ANR project 157133 “Seismology on Mars”), S. Stähler by a grant from the Deutsche Forschungsgemeinschaft (project SI1538/2-1 “RHUM-RUM”) and T. Nissen-Meyer by European Union’s Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Grant 641943. Collaborative visits between some of the authors have been generously supported by eCOST (European Cooperation in Science and Technology) Action ES1401-TIDES. Development and implementation of Syngine at the IRIS DMC was supported by U.S.

National Science Foundation awards EAR-1063471 and EAR-1261681.

References

- Basini, P., Nissen-Meyer, T., Boschi, L., Casarotti, E., Verbeke, J., Schenk, O., and Giardini, D. (2013). The influence of nonuniform ambient noise on crustal tomography in europe. *Geochemistry, Geophysics, Geosystems*, 14(5):1471–1492.
- Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., and Wassermann, J. (2010). ObsPy: A Python Toolbox for Seismology. *Seismological Research Letters*, 81(3):530.
- Brown, J., Knepley, M., and Smith, B. (2015). Run-Time Extensibility and Librarization of Simulation Software. *Computing in Science & Engineering*, 17(1):38–45.
- Buland, R. P. and Chapman, C. (1983). The computation of seismic travel times. *Bulletin of the Seismological Society of America*, 73(5):1271–1302.
- Ceylan, S., van Driel, M., Euchner, F., Khan, A., Clinton, J., Krischer, L., Böse, M., and Giardini, D. (2016). From initial models of seismicity, structure and noise to synthetic seismograms for Mars. *Space Sci. Rev.*
- 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):154–160.
- Deichmann, N. and Garcia-Fernandez, M. (1992). Rupture geometry from high-precision relative hypocentre locations of microearthquake clusters. *Geophysical Journal International*, 110(3):501–517.
- Dziewonski, A. M. and Anderson, D. L. (1981). Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors*, 25(4):297–356.
- Ekström, G., Nettles, M., and a.M. Dziewoński (2012). The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, 200-201:1–9.

- Helffrich, G., Wookey, J., and Bastow, I. (2013). *The Seismic Analysis Code: A Primer and User's Guide*. Cambridge University Press, 1 edition.
- Ishii, M., Shearer, P. M., Houston, H., and Vidale, J. E. (2005). Extent, duration and speed of the 2004 Sumatra-Andaman earthquake imaged by the Hi-Net array. *Nature*, 435(7044):933–6.
- Kennett, B., Engdahl, E. R., and Buland, R. P. (1995). Constraints on seismic velocities in the Earth from traveltimes. *Geophysical Journal International*, 122(1):108–124.
- Kennett, B. L. N. and Engdahl, E. R. (1991). Traveltimes for global earthquake location and phase identification. *Geophysical Journal International*, 105(2):429–465.
- Komatitsch, D. and Tromp, J. (1999). Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophysical Journal International*, pages 806–822.
- Komatitsch, D. and Tromp, J. (2002). Spectral-element simulations of global seismic wave propagation - I. Validation. *Geophysical Journal International*, 149(2):390–412.
- Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., and Wassermann, J. (2015). ObsPy: a bridge for seismology into the scientific Python ecosystem. *Computational Science & Discovery*, 8(1):14003–14020.
- Mai, P. M. and Thingbaijam, K. K. S. (2014). SRCMOD: An Online Database of Finite-Fault Rupture Models. *Seismological Research Letters*, 85(6):1348–1357.
- Incorporated Research Institutions for Seismology (IRIS) (2012). *SEED Reference Manual - Standard for the Exchange of Earthquake Data*.
- Megies, T., Beyreuther, M., Barsch, R., Krischer, L., and Wassermann, J. (2011). ObsPy – What can it do for data centers and observatories? *Annals Of Geophysics*, 54(1):47–58.

- Minson, S. E. and Dreger, D. S. (2008). Stable inversions for complete moment tensors. *Geophysical Journal International*, 174(2):585–592.
- Montagner, J.-P. and Kennett, B. L. N. (1996). How to reconcile body-wave and normal-mode reference earth models. *Geophysical Journal International*, 125:229–248.
- National Imagery and Mapping Agency (2000). Department of Defense World Geodetic System 1984: Its Definition and Relationships with Local Geodetic Systems. Technical Report TR8350.2, National Imagery and Mapping Agency, St. Louis, MO, USA.
- Nissen-Meyer, T., Fournier, A., and Dahlen, F. (2007). A two-dimensional spectral-element method for computing spherical-earth seismograms - I. Moment-tensor source. *Geophysical Journal International*, 168(3):1067–1092.
- Nissen-Meyer, T., van Driel, M., Stähler, S. C., Hosseini, K., Hempel, S., Auer, L., Colombi, a., and Fournier, A. (2014). AxiSEM: broadband 3-D seismic wavefields in axisymmetric media. *Solid Earth*, 5(1):425–445.
- Pérez, F. and Granger, B. E. (2007). IPython: a system for interactive scientific computing. *Computing in Science and Engineering*, 9(3):21–29.
- Sigloch, K. and Nolet, G. (2006). Measuring finite-frequency body-wave amplitudes and traveltimes. *Geophys. J. Int.*, 167(1):271–287.
- Stähler, S. C., Sigloch, K., and Nissen-Meyer, T. (2012). Triplicated P-wave measurements for waveform tomography of the mantle transition zone. *Solid Earth*, 3(2):339–354.
- Trabant, C., Hutko, a. R., Bahavar, M., Karstens, R., Ahern, T., and Aster, R. (2012). Data Products at the IRIS DMC: Stepping Stones for Research and Other Applications. *Seismological Research Letters*, 83(5):846–854.

- Tromp, J., Komatisch, D., Hjörleifsdóttir, V., Liu, Q., Zhu, H., Peter, D., Bozdog, E., McRitchie, D., Friberg, P., Trabant, C., and Hutko, A. (2010). Near real- time simulations of global CMT earthquakes. *Geophysical Journal International*, 183(1):381–389.
- van Driel, M., Krischer, L., Stähler, S. C., Hosseini, K., and Nissen-Meyer, T. (2015). Instaseis: instant global seismograms based on a broadband waveform database. *Solid Earth*, 6(2):701–717.
- van Driel, M. and Nissen-Meyer, T. (2014a). Optimized viscoelastic wave propagation for weakly dissipative media. *Geophysical Journal International*, 199(2):1078–1093.
- van Driel, M. and Nissen-Meyer, T. (2014b). Seismic wave propagation in fully anisotropic axisymmetric media. *Geophysical Journal International*, 199(2):880–893.
- Waldhauser, F. and Ellsworth, W. L. (2000). A Double-Difference Earthquake Location Algorithm: Method and Application to the Northern Hayward Fault, California. *Bulletin of the Seismological Society of America*, 90(6):1353–1368.

Tables

Model parameters		
model	prem_ani_5s	Velocity model.
Output parameters		
format	miniseed	Output file format: MiniSEED or a ZIP archive of SAC files.
label	Tohoku	Label to be included in file names.
components	ZRT	Components of the seismograms. Any of Z, N, E, R, and T.
units	velocity	Units: displacement, velocity, or acceleration.
scale	3.3	Amplitude scaling factor.
dt	0.2	Sampling interval in seconds. Only upsampling is allowed.
kernelwidth	8	Width of the sinc kernel used for resampling.
sourcewidth	15	Optionally convolve with a Gaussian STF given the width in seconds.
Time parameters		
origintime	2010-02-27T06:34:14	Source origin time as absolute date and time.
starttime	2010-02-27T06:34:14	Start time of the synthetic traces. Either absolute date and time, a phase-relative offset time, or an offset from the origin time.
endtime	2010-02-27T06:34:14	End time of the synthetic traces. Either absolute date and time, a phase-relative offset time, or an offset from the start time.
Receiver parameters		
network	IU	Network code to lookup station coordinates.
station	ANMO	Station code to lookup station coordinates.
receiverlatitude	47.6	Receiver latitude in degrees on WGS84.
receiverlongitude	-122.3	Receiver longitude in degrees on WGS84.
networkcode	YY	Set network code when directly specifying coordinates.
stationcode	D1Z1	Set station code when directly specifying coordinates.
locationcode	A1	Set location code.
Source parameters		
eventid	GCMT:C201002270634A	Event identifier to lookup earthquakes.
sourcelatitude	-89.99	Source latitude in degrees on WGS84.
sourcelongitude	-179.99	Source longitude in degrees on WGS84.
sourcedepthinmeters	699000	Source depth in meters.
Source mechanism as a moment tensor		
sourcemomenttensor	1e22, -3e22, -1e22, 3e22, -1e22, 9e21	Moment tensor source as M_{rr} , M_{tt} , M_{pp} , M_{rt} , M_{rp} , M_{tp} in Nm .
Source mechanism as a double couple		
sourcedoublecouple	19, 18, 116, 1e19	Double couple source as strike, dip, rake, and the scalar seismic moment in Nm .
Source mechanism as forces		
sourceforce	1e22, 1e22, 1e22	Vectorial force source as F_r , F_t , and F_p in N .

Table 1: Available parameters for the seismograms service of Syngine. Please see the website for a detailed and up-to-date explanation of them. A couple are also elaborated upon in the text of this manuscript and a full example to request vertical component data for the IU.ANMO station and the 2002 Denali earthquake is: <http://service.iris.edu/irisws/syngine/1/query?network=IU&station=ANMO&components=Z&eventid=GCMT:M110302J>

Custom source time function parameters		
cstf-relative-origin-time-in-sec	5.0	Offset in seconds to set the origin time of the seismogram.
cstf-sample-spacing-in-sec	10.0	Sample interval of the STF in seconds.
cstf-data	0.0, 0.5, 0.0	The custom STF data values.
Source parameters		
ffmeventid	USGS:us20002926	Event identifier to lookup finite source model. The name before the colon is the provider of the finite source model, everything thereafter a unique identifier.

Table 2: Additional parameters for the seismogram service which allow the specification of a custom source time function and for finite source calculations. All other, non source related parameters are identical to the seismograms service and are listed in table 1.

Request type parameters		
greensfunction	1	Specify the request type as Green's function.
Model parameters		
model	prem_ani_5s	Velocity model.
Output parameters		
format	miniseed	Output file format: MiniSEED or a ZIP archive of SAC files.
label	Tohoku	Label to be included in file names.
units	velocity	Units: displacement, velocity, or acceleration.
dt	0.2	Sampling interval in seconds. Only upsampling is allowed.
kernelwidth	8	Width of the sinc kernel used for resampling.
Time parameters		
origintime	2010-02-27T06:34:14	Source origin time as absolute date and time.
starttime	2010-02-27T06:34:14	Start time of the synthetic traces. Either absolute date and time, a phase-relative offset time, or an offset from the origin time.
endtime	2010-02-27T06:34:14	End time of the synthetic traces. Either absolute date and time, a phase-relative offset time, or an offset from the start time.
Source/Receiver parameters		
sourcedepthinmeters	15000	The source depth in meters.
sourcedistanceindegrees	72	The epicentral distance between source and receiver in degrees computed on the surface of a sphere.

Table 3: All available parameters for the Green's function service of Syngine. Please see the website for a detailed and up-to-date explanation of them.

Figures

Figure 1: Record section demonstrating a request with phase relative start and end times. This figure shows a plot of data from 8 stations at various epicentral distances. Each vertical component trace contains data from 50 seconds before the theoretical P wave arrival to 100 seconds after the theoretical S wave arrival. The source is a pure M_{rr} source at 600 km depth, the earth model is ak135-f ([Kennett et al., 1995](#); [Montagner and Kennett, 1996](#)), and the database contains periods from 2 to 100 seconds. Each trace is normalized around the first arrival to account for attenuation and geometrical spreading. The phase relative time settings can greatly reduce the amount of data that has to be transferred. Generating this plot, including the database extraction, serialization, sending the data to another continent, and actually plotting it takes only a few seconds.

Figure 2: Illustration of the effect of the `sourcewidth` parameter. The left panel shows seismograms convolved with Gaussian source time functions of different source widths. The seismograms are normalized as the source time function has a profound impact on the amplitude. Note the acausal effects - the event origin time is always defined to be the peak of the Gaussian. The right hand side plots the used source time functions in six shades of grey. The black outline shows the Gaussian with a source width of 25 seconds.

Figure 3: Plots of synthetic seismograms for the 2015-04-25 Mw 7.8 Nepal earthquake recorded at the IU.ANMO station. The seismograms are all on the same scale and show (A) a point source calculation with the database native quasi-Dirac delta STF, (B) a calculation using a finite source model from the USGS summing 400 point sources (see the data and resources section), (C) a point source calculation with a smooth Gaussian STF with a width of approximately 50 seconds, and (D) a point source calculation using an STF derived from the finite source model. The gray seismogram is the observed data and the right panel contains the source time functions used.

Figure 4: Plots of the ak135-f, iasp91, and PREM 1D Earth models for which Syngine can serve seismograms. ak135-f and iasp91 are isotropic models. For PREM, Syngine has an isotropic and an anisotropic variant - only the former is shown here. The left panel shows the full depth range, the right one only the first 50 km as the crust contains most differences. The left curves in each plot are the shear wave velocities, the right ones the compressional wave velocities.

Figure 5: Vertical component seismograms for the same source, mechanism, and receiver combination for all models part of the Syngine service at the time of writing. Seismograms from the first arrivals to the onset of the surface waves are shown in the left panel. The right panel zooms in with the first wave package being a mixture of P, pP, sP, and PcP phases and the second at around 855 seconds the PP phase. Black boxes in the middle denote the model from which each seismogram was extracted. Modelled is a Mw 8.3 earthquake near the coast of Chile from 2015-09-16 with a centroid depth of 17.6 km recorded at the IU.ANMO station in Albuquerque, New Mexico, USA with an epicentral distance of approximately 74° . It is computed with a moment tensor source and no further source time function convolution and the smallest resolvable period is ascending from top to bottom.

Figure 6: Illustration of an application of Syngine to assess the quality of observed data and recording stations. This particular example aims to detect time shifts due to clock errors by cross-correlating observed P phases with synthetic data calculated by Syngine. The left panels show two maps: the top one highlighting the source-receiver geometry, the bottom one zooms in on the receivers, a part of the Caltech Regional Seismic Network (CI). The middle panels show examples of velocity seismograms of observed and synthetic data before and after applying a time-shift equal to the delays measured via cross-correlations. Note that all traces have been processed but no amplitude normalization has been performed - refer to the main text for more details. The rightmost panels are plots of the measured time-shifts. The top one is the reference time shift for each station calculated by using the mean delay from 3 training earthquakes. We subtracted the mean delay of all stations to account for effects such as common unexplained 3D velocity structure, Earth's ellipticity, and incompatible source parameters. The middle one is the difference in delay time for each station to the reference for an event that was not used to calculate that baseline. The measured shifts are very small. The bottom one finally is another event where we applied an artificial delay of ± 1 second to two stations which clearly show up in the final measurements. This is only a proof of concept and requires further studies.

Figure 7: Results of a body-wave source inversion for moment tensor and source time function for the Mw 6.2 earthquake in Italy from 2016-08-24. The inversion is performed using databases with four different velocity models: ak135f_2s, ak135f_5s, prem_a_2s, and iasp91_2s. For each model, the depth with the lowest misfit between measured and simulated waveforms is shown. The resulting moment tensors are relatively similar to each other and to the body-wave solution from the USGS (bottom, [quakeml:us.anss.org/focalmechanism/10006g7d/mwr](https://quakeml.us.anss.org/focalmechanism/10006g7d/mwr)). The best-fitting depth estimate for the body wave solutions varies between 8 and 10 km. The inversion was done using `stfinv` (<https://github.com/seismology/stfinv>) and the Syngine databases. Note that the inversion algorithm does not forbid negative energy in the source time function, if it improves the fit.

Figure 8: Syngine is a very valuable tool to educate future generations of seismologists. Interactively changing source parameters and the geometry of the problem will grant them an enhanced understanding into the behavior of seismograms. This plot demonstrates two examples of this. The left panel shows the effect of nodal planes on seismograms. The source is always the same, except for a differing strike angle. The right panel shows that surface waves are more pronounced for shallower events granting insights into the potential damage of waves originating from different seismic sources.

Backprojection of 2008-05-12 Wenchuan M7.9 using AU network

Figure 9: Backprojection results for the May 12, 2008 Wenchuan earthquake using data from the AU network (left) are compared to results using synthetics using the USGS multi-segment FFM (right). The top panels show the cumulative backprojection energy. The middle panels show the time and locations of local maxima in the backprojection stack amplitude which can be used to infer rupture velocity. The bottom panels show peak backprojection stack amplitude over time. Though local maxima in the FFM backprojection appear to be delayed by about 11 seconds, the results are consistent with the FFM moment rate function which peaks at around 22 sec. Independent of any time offsets, the comparison of the results show agreement in the rupture velocities inferred from the data and present in the FFM; about 2.6 km/s. Small spatial offsets or lack of energy detected late in the rupture when many interfering phases arrive, is not unexpected.