Standardization of seismic tomographic models and earthquake focal mechanisms data sets based on web technologies, visualization with keyhole markup language

Luca Postpischl*, Peter Danecek, Andrea Morelli, Silvia Pondrelli
Istituto Nazionale di Geofisica e Vulcanologia, sezione di Bologna, Via Creti 12, 40128 Bologna, Italy.

A R T I C L E   I N F O

Article history:
Received 8 June 2009
Received in revised form
29 April 2010
Accepted 7 May 2010

Keywords:
Seismology
Geophysics
JSON
Google Earth
KML
Semantic data

A B S T R A C T

We present two projects in seismology that have been ported to web technologies, which provide results in Keyhole Markup Language (KML) visualization layers. These use the Google Earth geobrowser as the flexible platform that can substitute specialized graphical tools to perform qualitative visual data analyses and comparisons. The Network of Research Infrastructures for European Seismology (NERIES) Tomographic Earth Model Repository contains data sets from over 20 models from the literature. A hierarchical structure of folders that represent the sets of depths for each model is implemented in KML, and this immediately results into an intuitive interface for users to navigate freely and to compare tomographic plots. The KML layer for the European-Mediterranean Regional Centroid-Moment Tensor Catalog displays the focal mechanism solutions or moderate-magnitude Earthquakes from 1997 to the present. Our aim in both projects was to also propose standard representations of scientific data sets. Here, the general semantic approach of an XML framework has an important impact that must be further explored, although we find the KML syntax to more emphasis on aspects of detailed visualization. We have thus used, and propose the use of, Javascript Object Notation (JSON), another semantic notation that stems from the web-development community that provides a compact, general-purpose, data-exchange format.

1. Introduction: towards a standard format for Earth model data

Earth models that result from seismic tomographic studies are published by means of their expansion in spherical harmonics functions. Authors typically share the raw coefficients in ascii files, following personal conventions, but having a common formalism would greatly enhance the data dissemination and their usage by other researchers. In practice, any Earth model can be represented by a set of values on a latitude, longitude and depth three-dimensional (3D) grid covering the interior of the Earth. Most models are based on p-wave and s-wave propagation velocities, but in general, many other physical measurables can also be exploited to gain insight into the Earth's interior.

A model typically consists of large quantities of data. Within the Network of Research Infrastructures for European Seismology (NERIES) European Union Research Project ‘JRA1’ activity, which aims at defining a unified reference Earth model for the European region, twenty of the most popular existing global models are being reviewed. The need for an efficient representation of the data sets that can also serve as a common base to visualize and compare models without installing and running Fortran executables provided by original authors have become evident.

We have therefore searched for solutions that are capable of being language independent, easy to parse, and semantic, i.e. solutions that result in self-describing structures that integrate the data and metadata into a single resource. We believe that the Javascript Object Notation (JSON) formalism fits these general requirements well, and we have proposed its adoption for the standardization of Earth models (Postpischl et al., 2008). JSON is a subset of the ECMA-262 specification (Crockford, 2006), and it is based on a very minimal and clean notation. This is currently supported by most of the major programming languages that should already be familiar to most programmers (ActionScript, C, C++, C#, Cold Fusion, D, Delphi, E, Erlang, Haskell, Java, Lisp, LotusScript, Lua, Perl, Objective-C, OCAML, PHP, Python, etc.).
Compared with an equivalent extensible markup language (XML) implementation, a JSON object is significantly more lightweight in file size (Lawrence, 2004). Moreover JSON-formatted structures are fully defined Javascript objects, so their elements are directly parsable by the browser javascript engine without the need for extra middle-ware layers, such as an SOAP, XPath, and SAX that are typically needed for XML processing. JSON is thus very efficient, and it is becoming the preferred data-exchange format for many representation state transfer (REST)ful web services (Richardson and Ruby, 2007). All of the open-source Ajax frameworks developed through the Web Standards Community offer advanced support for JSON, so adopting JSON also means bringing this huge arsenal of software tools into the hands of scientific researchers, potentially transforming a web browser into an advanced data visualization and analysis tool. Moreover the support of JSON by many programming languages guarantees that the conversion of scientific data towards higher level formats, such as the Network Common Data Form (NETCDF, 2007) and Hierarchical Data Format (HDF5, 2010), is easy to implement. This allows advanced plotting with specialized visualization tools, such as the GEON Integrated Data Viewer (IDV, 2007) and the Generic Mapping Tools (GMT, 1988).

KML-generating routines are also easily implemented from the JSON formalism with many programming languages, hence geo-browsers can be directly exploited for visualization of JSON data structures; in our opinion KML, being heavily targeted for presentational tasks, is not a suitable format for general spatial-data storage and transfer.

All revisions of the JSON standardization format proposal for tomographic Earth models are published and discussed openly on the ‘/wiki’ pages of the Istituto Nazionale di Geofisica e Vulcanologia (INGV) Bologna branch website (Eureom, 2008). (Graphic 1).

2. Visualization of tomographic maps with KML

The tomographic models available through the current NERIES JRA1 activity are all expressed in layers of data points that correspond to various depths from ground level down to 600 km below the surface of the Earth. For each depth level, a regular 2D grid with two degrees of resolution of latitude and longitude is defined for the European-Mediterranean region. The visualization of any aspect of the model tightly follows this general organizational scheme of the JSON data sets: in our main KML file for the tomographic models the depth levels are all listed in the ‘places’ sidebar of Google Earth, and by selecting the corresponding check-boxes, the tomographic map for that particular level is displayed, clamped to ground level. We have found this to be the most effective and user-friendly way to actually carry out comparisons across the depth levels, and it is clearly preferable to displaying all of the depths layered according to different altitudes (N.B. Google Earth does not support negative values).

As common in tomographic studies, the data values that are color coded into the KML polygons correspond to the percentage variations from the mean values computed for the depth levels of the single data points. This scalar quantity serves as a common parameter with which to compare different levels, and also different models, even if these are based on physical measurables that are not related.

The most interesting element of the implementation is the use of ‘<NetworkLink’ KML entities inside the main hierarchy of <Folders> (Fig. 1). In this way, all of the KML code generating a particular tomographic map is retrieved asynchronously from the remote server only when the user explicitly requests it, which optimizes bandwidth use. As displayed in the code snippet of Graphic 2, each ‘<NetworkLink’ passes two parameters to a remote server-side script (Graphic 3) that parses the JSON data structure for that model and returns the set of KML ‘<Polygon’ s corresponding to the depth level to the Google Earth client application.

It would be equally easy to use this same technique based on <networkLink> s for importing high-resolution graphical files, batch-produced with specialized scientific routines, as <groundOverlay> s KML entities, an approach which transforms Google Earth into a high-quality, cross-platform software tool for quick, interactive, comparisons of any two models.

As a proof-of-concept of the capabilities of the proposed JSON format, we have here chosen a ‘raw’ implementation from the original data. This has also given us insights concerning the scaling of the performance of the Google Earth platform for the rendering of layers of several thousands of polygons in real time.

As we have not had performance issue limitations in the drawing of the horizontal layers, we believe that our approach can be further exploited for the creation of other kinds of maps that are commonly found in tomographic studies: vertical cross-sections. In this case, the import of batch-produced images would be limiting, because the user would want to create these maps interactively, and along many directions. The maps have to be created in real time once the coordinates of the path start and end points have been specified by the user. To perform such a selection, the standard Google Earth user interface based on the clickable hierarchies of folders in the ‘places’ sidebar would not be sufficient. Yamagishi et al. (2008) used a form on a web page to define the path and the output KML files for cross-sections, although these files have to be imported into Google Earth manually.

The Google Earth plug-in and the corresponding javascript application programming interface (Earth API, 2008) that brings the geo-browser KML support inside the web browser can be exploited to implement vertical cross-sections with a fully interactive user experience. Based on the plug-in, we intend to develop a tool that will let users select a tomographic model from a standard HTML form, and then allow them to define a path by dragging the mouse onto a 3D globe, whereby the mouse-up event of this line-drawing action will trigger the processing of the tomographic data for that particular cross-section.

3. The European-Mediterranean regional centroid-moment tensors

The European-Mediterranean regional centroid-moment tensor (RCMT) catalog collects seismic moment-tensor solutions that have been routinely computed for earthquakes with moderate magnitudes (4.5 < M < 5.5) in the European-Mediterranean regions (Pondrelli et al., 2007). We now have a catalog of centroid-moment tensors that includes more than 900 RCMTs that all together represent the time span from 1997 to the present. This database represents an extension for smaller magnitudes of the Global CMT catalog for the European-Mediterranean area (Ekström et al., 2005).

An RCMT computation is based on the analysis of seismograms that are recorded at regional distances and on the modeling of...
intermediate period surface waves (Arvidsson and Ekström, 1998). Over the past few years, we also inverted simultaneously for body and surface waves, although only for those seismic events with a magnitude greater than 5.0—i.e. when the signal-to-noise ratio at 40–100 s of period is significant enough to contribute to the inversion.

4. The RCMT web-search interface

The RCMT catalog is updated every few months, and reports are published regularly. However, moment-tensor solutions are also being computed on the basis of data that are available in quasi-real time. These preliminary solutions are available within a few hours after the occurrence of an earthquake, and they are published immediately as ‘Quick RCMTs’ in the online version.7

To provide full search capabilities over the moment-tensor solutions, the data set that was previously available as static ASCII files has been imported into a MySQL relational database. This thus provides a PHP web application (Fig. 2) with enhanced user interface controls that allow users to submit queries as a

7 http://www.bo.ingv.it/RCMT/searchRCMT.html.
combination of date, magnitude, depth and geographic coordinate ranges (Pondrelli et al., 2008). These events can be further filtered by two flags: one to distinguish between quick and definitive solutions, and one for the ‘quality’ categorization that we are using for the solutions.

While selecting the latitude and longitude ranges with the slide-bars of the form, the corresponding area of interest is drawn in real time on a zoomable map (Google Maps), which provides precise visual feedback to the user. To guarantee that this can be operated even with older browsers, there are standard input field elements which are synchronized with the slide-bar controls. Once the search parameters have been set and the query is submitted, the map is updated with beach-ball representations of the focal mechanisms, which are positioned at each earthquake epicenter location. A mouse-over event is defined on each beach ball, which generates a dynamic information box that contains the full solution for the event.

Immediately below the map, there is a visual characterization of the resulting data set, given as magnitude, depth and time frequency in-line distribution histograms as in Tufte, 2006. Finally, the full data set is provided and the user can convert it between various formats within the web browser, thus allowing it to be exported directly into other applications. The default output format for the data set through this web service is again JSON, as it intuitively and conveniently integrates the search parameter metadata with the actual records into a single machine and human-readable resource/file.

The other output formats we provide are all generated dynamically on the client-side from the JSON data-object, without further connections to the server. Currently, the form includes the Psmeca and Psvelomeca GMT formats, basic comma-separated values (CSV) for easy integration in spreadsheet applications, and KML for Google Earth. We plan to also include QuakeML, an XML format specifically developed for seismic data exchanges (QuakeML, 2007), and GEON IDV ASCII Point-data format (GIAP, 2007), an ASCII format for which netCDF/HDF5 converters are available.

5. The RCMT KML output

The place-mark icons used in the KML output for the moment tensors are the same 2D bottom-half projections of the beach balls displayed on the in-page Google Map, and they retain the visualization of the further details concerning each earthquake in the dynamic information boxes that are shown on a user click. The beach balls are scaled in proportion to the magnitude of the corresponding earthquake events. (Graphic 4).

This conventional representation that corresponds to the classic one found in the literature poses new problems in the Google Earth 3D-rendering environment: when the user changes the point of view by rotating around the vertical axis or by tilting the view, the orientation of the place-marks remains fixed, and thus no longer correct.

KML would allow the importing of 3D-sphere Collada model files (.dae) for the beach balls, and to further orient them in space into the reference system used in Google Earth, with a simple transformation of the strike, dip and rake angles contained in the data records (i.e. the <heading>, <tilt>, <roll> KML entities) (De Paor, 2008; De Paor and Pinan-Llamas, 2006). (Graphic 5).

Within the 3D-viewing engine, this implementation would clearly display the fault-plane intersection with the ground, a
feature that would be immensely useful in educational contexts. However, after some experimentation and discussion, we finally chose to only provide the classic 2D projection of the beach balls instead, as commonly found for printed maps in the literature (Fig. 3). This is to ease the analysis by seismologists, who will be accustomed to the very counter-intuitive convention for moment-tensor solutions, i.e. when viewed from above; beach balls are displayed as the horizontal projection of their bottom-half, whereas from the same point of view, the 3D spheres would instead be seen as their top half. Providing both representations as two distinct KML folders might be the only solution to this dichotomy, whereby users would be able to choose their preferred representation from the locations sidebar. Also, to give users a better sense of the distinction between these two representations, the 3D spheres can be substituted by other custom, more intuitive, 3D models of the fault planes, as has been used by Labay and Haeussler (2007) and by De Paor and Williams (2006). This will be implemented into further revisions of the RCMT project.

To further simplify and enhance the readability of the map in Google Earth, as for the tomographic model implementation, we

Graphic 2. KML snippet for networkLink entry passing variables to a remote PHP script.
again chose not to position the solutions vertically, and instead we kept them clamped to ground level. In this way, the beach balls are always clearly displayed at their precise epicenter localization points, and they are not affected by perspective parallax effects introduced by the 3D rendering engine on tilted semi-horizontal views. This also by-passes lack of support of Google Earth for negative altitude values, which would otherwise require a specific solution based on extra 3D Collada models, as for that developed by De Paor (2007). Such an implementation is not practicable in our case, since the data sets are formed dynamically at the request of the user, while the .dae Collada files would have to be created explicitly within the Google Sketchup desktop application (Sketchup, 2007), and then referenced in the KML.

The last important feature implemented in these KML code snippets is the inclusion of the `<TimeStamp>` tag, as used by De Paor and Pinan-Llamas (2006). This very simple KML entity auto-triggers the appearance of the time slide-bar in Google Earth, allowing the user to interactively animate and analyze the data set by setting a reference time window and dragging it back and forth within the global 1997-to-present time-frame. This feature is especially important in the full export version of the data set that we

```php
<?php
// get the parameters from the URL (as found in each <NetworkLink>
// of the kml file)
$depth = $_GET["depth"];
$model = $_GET["model"];
// set the hex values for the color-index scale
$color = array
"656D79", "2CF5F9", "3F81AD", "68B5CB", "96CCEE", "CCEEC8", "EEF5DF", "E9F1AD", "E4E37A", "E6CC4C", "E5AB34", "E6768D", "CD7334");
// retrieve the json file for the requested model
$json_file = "http://www.bo.ingv.it";
$json_file = $path.$model.".json.txt";
// parsing of the JSON string into a PHP object
$float = json_decode($string);
// select the index inside the array of depths corresponding to the
// requested depth value
$depths = $float->

dataset]['depth']->


depthIndex = $count;

foreach ($depths as $value) {
    if ($depths[$count] == $depth) {
        $depthIndex = $count;
        $count++;
    }
    $count = 0;
}
// start building the KML output
<kml xmlns="http://www.opengis.net/kml/2.2"><Document><Folder id="Main Folder"><name>Avg. speed $avgSpeed</name>

.setColor="656D79", "2CF5F9", "3F81AD", "68B5CB", "96CCEE", "CCEEC8", "EEF5DF", "E9F1AD", "E4E37A", "E6CC4C", "E5AB34", "E6768D", "CD7334");

$latMin = $obj->

$latMax = $obj->

$lonMin = $obj->

$lonMax = $obj->

$latStep = $obj->

$lonStep = $obj->

$dVal = 0;
```

Graphic 3. PHP snippet receiving parameters from networkLink and processing JSON data into KML color coded polygons.
```bash
for ($i = $lonMin; $i <= $lonMax; $i += $lonStep) {
    for ($j = $latMin; $j <= $latMax; $j += $latStep) {
        $vs = $obj->{'dataset'}->{'measurables'}[0]->{'values'}[$depthIndex][$i][$j];
        echo "$iVal ($j) $vs<br>";
        $percVar = ($vs - $avgSpeed)/$avgSpeed;
        if ($percVar <= -.07) {$colorIndex = $color[0];}
        else if (($percVar > -.07) && ($percVar <= -.055))
            {$colorIndex = $color[1];}
        else if (($percVar > -.055) && ($percVar <= -.045))
            {$colorIndex = $color[2];}
        else if (($percVar > -.045) && ($percVar <= -.035))
            {$colorIndex = $color[3];}
        else if (($percVar > -.035) && ($percVar <= -.025))
            {$colorIndex = $color[4];}
        else if (($percVar > -.025) && ($percVar <= -.015))
            {$colorIndex = $color[5];}
        else if (($percVar > -.015) && ($percVar <= -.005))
            {$colorIndex = $color[6];}
        else if (($percVar > -.005) && ($percVar <= .005))
            {$colorIndex = $color[7];}
        else if (($percVar > .005) && ($percVar <= .015))
            {$colorIndex = $color[8];}
        else if (($percVar > .015) && ($percVar <= .025))
            {$colorIndex = $color[9];}
        else if (($percVar > .025) && ($percVar <= .035))
            {$colorIndex = $color[10];}
        else if (($percVar > .035) && ($percVar <= .045))
            {$colorIndex = $color[11];}
        else if (($percVar > .045) && ($percVar <= .055))
            {$colorIndex = $color[12];}
        $megaVs = $vs*1000;
        $iVal++;
        $boxLonA = $i;
        $boxLatA = $j;
        $boxLonB = $i + 1.99;
        $boxLatB = $j;
        $boxLonC = $i + 1.99;
        $boxLatC = $j + 1.99;
        $boxLonD = $i;
        $boxLatD = $j + 1.99;
        $kml .= '<Placemark><name>Lat: $j Lon: $i</name><visibility>1</visibility><Style><PolyStyle><color>ff$
    $colorIndex</color><outline>0</outline><fill>1</fill></PolyStyle></Style><Polygon><extrude>0</extrude><altitudeMode>clampedMode</altitudeMode><outerBoundaryIs><LinearRing>
        <coordinates>$boxLonA,$boxLatA,$megaVs $boxLonB,$boxLatB,$megaVs $boxLonC,$boxLatC,$megaVs $boxLonD,$boxLatD,$megaVs $boxLonA,$boxLatA,$megaVs
    </coordinates>
</LinearRing></outerBoundaryIs></Polygon>
</Placemark>
    }
}

$kml .= '</Folder></Document>'
print $kml;
```
Indeed, this provides a less cluttered view of the sub-regional clusters of earthquakes.

### Graphic 4.

KML snippet for a basic place-mark for RCMT KML layer.

6. Conclusions

The general approach outlined above for the use of JSON as a semantic data-exchange format for scientific data brings the 'web-as-a-platform' paradigm into the hands of researchers. As well as fitting in with the latest theoretical trends in computer

---

10 http://www.bo.ingv.it/RCMT/fullDataSet.kmz
science, this has enormous practical advantages. Many online tools and advanced javascript frameworks are readily available to build interactive, cross-platform visualizations of data, with no need for users to install or configure anything. Google Earth is just one of these options, and despite its current lack of support for negative values of altitudes, Google Earth still represents a very flexible 3D-rendering environment for the geosciences, as it provides good performance scaling when dealing with large data.

Fig. 3. Focal-mechanism visualization. D beach-ball representation of focal mechanisms is georeferenced in Google Earth in clampedToGround mode. Tilted, semi-horizontal 3D views, such as that shown, would result in a poorly readable visualization of hypocenters if focal mechanisms were plotted at heights proportional to the original inverted depth of the event.

Graphic 5. KML snippet for an enhanced place-mark with importation of 3D-sphere models.
sets. Also of note is the ease with which the geobrowser platform architecture overlays data sets from different sources, allowing users to carry out advanced integrations and comparisons.

Appendix A. Supporting material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejcb.2009.08.001.

References


Earth API, 2008. Google Earth API, Google Incorporated, Mountain View, CA, USA.


