Time Varying Hazards Research Programme Number TVH2/11 Report

M. Gerstenberger

GNS Science Consultancy Report 2009/28
February 2009

BIBLIOGRAPHIC REFERENCE

Gerstenberger, M.; 2008. Time Varying Hazards Research Programme Number TVH2/11 Report, *GNS Science Consultancy Report* 2009/28. 14p.

CONTENTS

FIGURES

EXECUTIVE SUMMARY

Validation of seismic hazard models and forecasts has entered a new era in recent years with a focus on establishing rigorous community developed and accepted testing procedures. These efforts were originally lead by the Regional Earthquake Likelihood Models group of the Southern California Earthquake Centre and are now firmly in place with the work of the Collaboratory for the Study of Earthquake Predicitibilty (CSEP), which includes the New Zealand Earthquake Forecast Testing Centre (NZTC). Testing centres follow a strict protocol

for registering, evaluating and testing of models. The NZTC operates using the officially supplied CSEP computational libraries and is using version 9.1.1 of the library as of February of 2009. The testing is done using a variety of likelihood based procedures which, on a uniform spatial grid of 0.1 degrees, test the numbers of events forecast, the event locations and the relative performance of the models against all other models. In the NZTC we are currently testing models that forecast earthquakes for 3 time scales: 1-day, 3-months and 5-years. The 1-day models are expected to forecast events of magnitude ≥ 4 and the longer term models forecast events of magnitude ≥ 5. All tests are evaluated over a 5 year period, at which point official results will be complete. Currently 11 models including their variations have been submitted to the testing centre. To develop a forecast, all models are dependent on the input of past events from the GeoNet catalogue; currently this catalogue contains missing events from the period April 2007 until July 2008. Until this time gap is completed, official testing cannot begin.

TECHNICAL ABSTRACT

Validation of seismic hazard models and forecasts has entered a new era in recent years with a focus on establishing rigorous community developed and accepted testing procedures. These efforts were originally lead by the Regional Earthquake Likelihood Models group of the Southern California Earthquake Centre and are now firmly in place with the work of the Collaboratory for the Study of Earthquake Predictability (CSEP), which includes the New Zealand Earthquake Forecast Testing Centre (NZTC). The work presented in this report was largely focused on the technical and operational aspects of the NZTC. Important developments have been fully automating all testing procedures, publication of testing-results on a secured web page, http://nz.cseptesting.org/results/, and the installation of forecast models into the NZTC. The NZTC system has been kept in step with CSEP code releases and is currently operating using CSEP v9.1.1. Eleven models, including their variations, have been submitted to the testing centre and seven of these have been installed and are being unofficially tested. The following models have been submitted: NZ national seismic hazard model (Stirling, et. al., 2002), EEPAS (Rhoades & Evison, 2004), PPE (Rhoades & Evison, 2004), SUP (Rhoades & Evison, 2004), STEP (Gerstenberger, et. al., 2005), and Double branching model (Marzocchi and Lombardi, 2008). Of these, the national seismic hazard model, PPE, SUP and STEP are installed and are currently being unofficially tested. The STEP model calculates a new forecast every day which is tested against the daily observations. The other models have provided 1 forecast for a five year period; the tests are updated daily, but the final result awaits five years of data. The five year models are tested against a declustered and a non-declustered catalogue. Official testing is dependent on a complete historical catalogue; currently, the GeoNet catalogue has a completeness gap from April 2007 to July 2008. With current practice, this gap will be filled in about 2 years, at which point official 5-year tests can begin.

1.0 INTRODUCTION

In recent years standards for how testing of earthquake hazard and forecasting models should be conducted have been developed under the auspices of the Collaboratory for the Study of Earthquake Predictibility (CSEP). Based out of California, CSEP is an international group which has developed a detailed procedure and an extensive library of computational codes for testing grid-based earthquake rate forecasts. New Zealand based researchers have played a leading role in the development of CSEP methodologies. In 2006, with funding from the EQC, the beginnings of a New Zealand based earthquake forecast testing centre (NZTC) were established. The overall objective of this study is to firmly establish the NZTC and to facilitate long-term testing of earthquake forecast models. To reach this point, four goals were proposed: development of computational routines to handle automatic processing of earthquake data; installation of a website to display the results of the NZTC; initiation of the first 5 year testing period; and maintenance of the automated system. To ensure credibility of the results of testing forecast models within the NZTC, a methodology has been designed that increases the transparency of testing procedures and allows for reproducibility of both the testing procedures and results (Schorlemmer & Gerstenberger, 2007). Such a system is a departure from traditional earthquake science where individuals use their own testing procedures, or do no testing at all. This has led to a multitude of different testing procedures which are often difficult to understand and can rarely be validated by normal peer review processes. The NZTC includes an extensive computational framework and rigorous standards for the handling of input data (e.g., the earthquake catalogue) and forecast models. For example, all input data must come from an authorised data source and the data must be retrieved in a consistent manner (i.e., automatically, following a predetermined rule set) to ensure reproducibility; for the NZTC all earthquake data is supplied by GeoNet. All models must be installed within the NZTC computational framework prior to the start of official testing. For models that calculate one forecast for the entire 5-year testing period, such as the National Seismic Hazard Model (Stirling et. al., 2002), only the forecast is required to be supplied and stored within the framework. For models that update their forecast throughout the testing period (i.e., forecasts updated every day, or every 3 months), all computer code used to create each forecast must be installed and operational within the computational framework; these models are only provided the official authorised data, no other input data is allowed. All testing procedures used within the NZTC have been developed through a long-term collaborative effort (Schorlemmer, et.al., 2007) which involved researchers from across the earthquake hazard and forecasting community.

2.0 METHODOLOGY

Much of the work in this project was focused on aspects of the computational framework which the NZTC is built upon. The core of the code was provided by CSEP (http://www.cseptesting.org) however, development was necessary so that the code would be functional in the New Zealand environment; this included modifications to the testing procedures as well as the creation of new code. Much of this work was done with the assistance of CSEP programmers.

2.1 Automated Procedures

2.1.1 GeoNet Catalogue Download

The CSEP code is written primarily in the python language with minor components being written in other languages such as Matlab. The first task in getting the code to be fully automatic was to create an interface with the GeoNet catalogue so that the data could be downloaded as needed. Initially no service was in place at GeoNet, however, on request they provided a webservice that allows automated downloads using QuakeML (https://quake.ethz.ch/quakeml/), which is an emerging XML standard for the representation of seismological data. This interface allowed us to create modules within the CSEP code which automatically download the catalogue daily for the collection region (Figure 1), the magnitudes of interest and the time period of interest. For example for tests run for the date of 02/12/08, more than 43,000 earthquakes with magnitudes greater than 3 were downloaded from the catalogue and supplied to the models. More than 7,000 of these were magnitude 4 and greater, and were used to evaluate the performance of the model up to this date.

2.1.2 Declustering

Some long-term models (5-year forecast models) are developed to forecast only mainshocks and do not include aftershocks. To best understand the performance of these models, validation must be done on 2 catalogues: one that includes aftershocks and another where aftershocks have been removed, i.e., the catalogue is declustered. Declustering is not an absolute process and there is no consensus on how a true declustering should be done. The methodology of Reasenberg (1985) is one commonly accepted declustering method and is the method adopted by CSEP. A potential improvement to the Reasenberg methodology is to perform a Monte Carlo search of the declustering parameter space to obtain a probability of each event belonging to a cluster of earthquakes. These probabilities can then be used as uncertainties and allow for many reasonable realisations of the declustered catalogue over which the model can be validated. This procedure is very computationally expensive and, due to the length of the calculations, it was decided that it is impractical to perform the Monte Carlo search for daily catalogue downloads and daily testing. The simulations may, however, be performed for the final evaluation of the long-term models at the end of 5 years. For a more detailed discussion of declustering, see the previous NZTC EQC reports (Rhoades, et. al., 2008).

2.1.3 Archiving

A non-trivial amount of data is produced by the testing centre, where each model can contribute roughly 100 megabytes of data each day. To calculate cumulative tests over a 5 year period, and also to preserve transparency, all data must be archived. All archiving is currently handled using the utilities provided by the core CSEP code and no additional functionality has been added for the NZTC. All forecast data and test results are stored within the computational framework of the NZTC at GNS Science. All results images are duplicated at the CSEP center at SCEC. We are currently purchasing two servers at GNS dedicated to the NZTC, one of which will handle all of the archiving. This data will be backed-up using the standard GNS Science procedures.

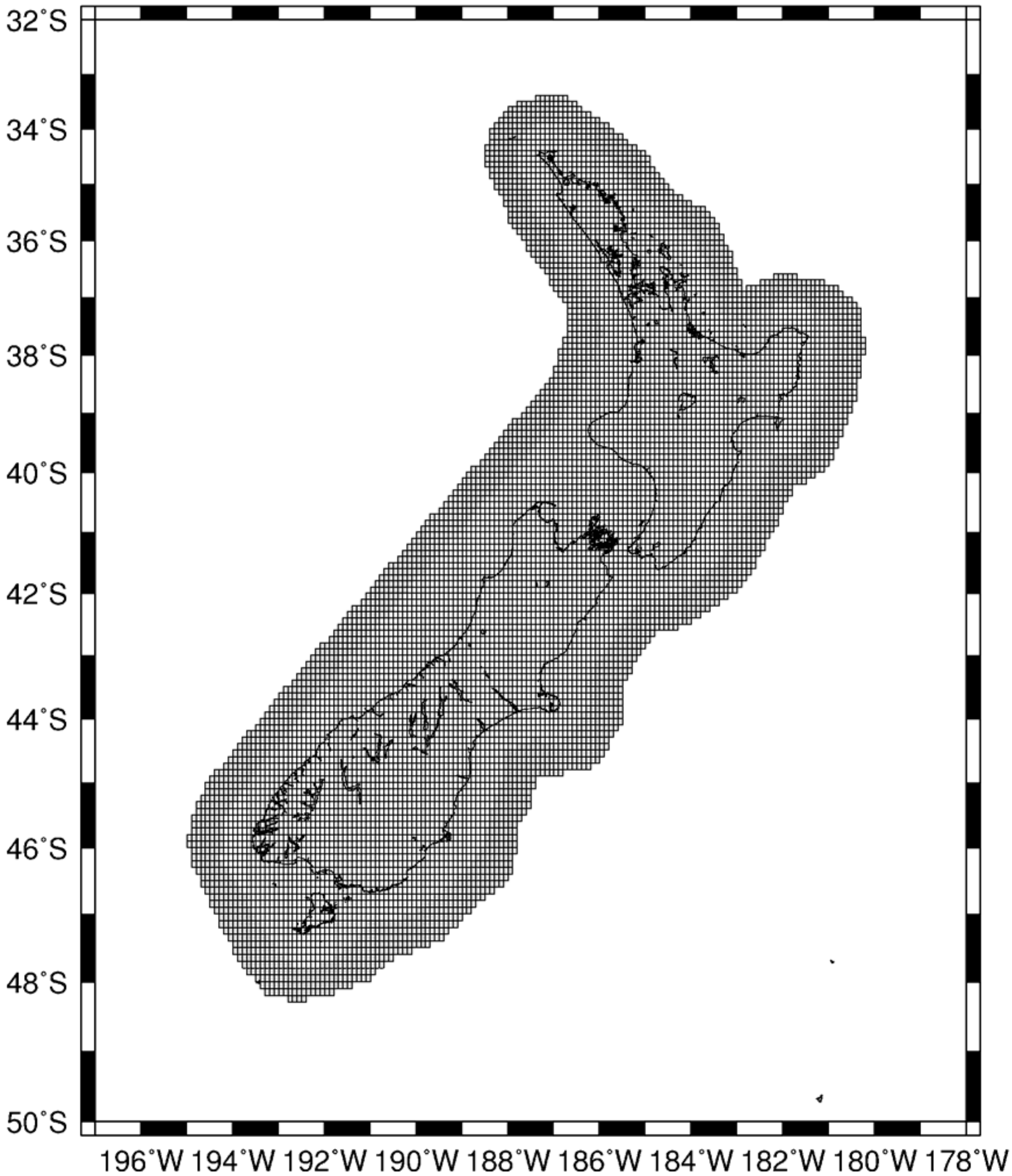

Figure 1 Models are evaluated using events occurring in the testing region which is shown in pink. Models are supplied events from the slightly larger area, the collection region, which includes both the grey and the pink regions.

2.2 Web Platform

The NZTC will ultimately contain forecast models developed by researchers all over the globe. The short-term instances of these forecast models will generate forecasts each day and, although the final official results will require 5 years of data, all forecast models will be evaluated daily. An easy and efficient way to disseminate this information to interested researchers is via an NZTC webpage. The NZTC webpage is now up and running at http://nz.cseptesting.org. While all NZTC computations are performed on dedicated NZTC servers based at GNS Science, the web page hosting is handled by CSEP and the Southern California Earthquake Center (SCEC). The decision to host the pages at SCEC was made due to the fact that the infrastructure to do so was already in place and it reduced the workload on the PI. However, if in the future it is decided that these pages need to be hosted in New Zealand, transferring the web data is not a difficult task. The webpages contain basic information about the NZTC, CSEP and includes automated updates from the international CSEP working groups.

The model forecasts and test results are displayed at http://nz.cseptesting.org/results/. This is a password protected page to avoid misinterpretation of these research forecasts by the public and others. The login name and password will be available to all involved in the project and to others upon application to the NZTC. The page which gives a brief description of the testing methods is shown in Figure 2. The forecast of a particular model and its evaluation is shown in Figure 3a and 3b. Additional detail on how to interpret the plots is displayed when placing the mouse over individual test results.

2.3 System Maintenance

The CSEP operational environment is relatively immature and is rapidly developing. At present, the code developers release a new official version of the code every 3 months, with minor changes occasionally occurring at shorter intervals. Each time a new release is delivered it is required to update the NZTC to the current release as the releases will contain bug fixes and necessary enhancements. Initially the upgrading of the system could be a relatively daunting task and could take some number of days; however, as the code has matured, and as the NZTC operators have become more fluent with the code, the updates have become less difficult and have required less time from the NZTC operators. As of 02/09 the NZTC is using CSEP v9.1.1. The detail of changes to the code and the bug fixes, as well as the release schedule are available on both the NZTC website and the main CSEP website.

In the early stages of real-time operation, system maintenance was primarily dedicated to keeping the system updated and ensuring that all features necessary to the NZTC were in place.

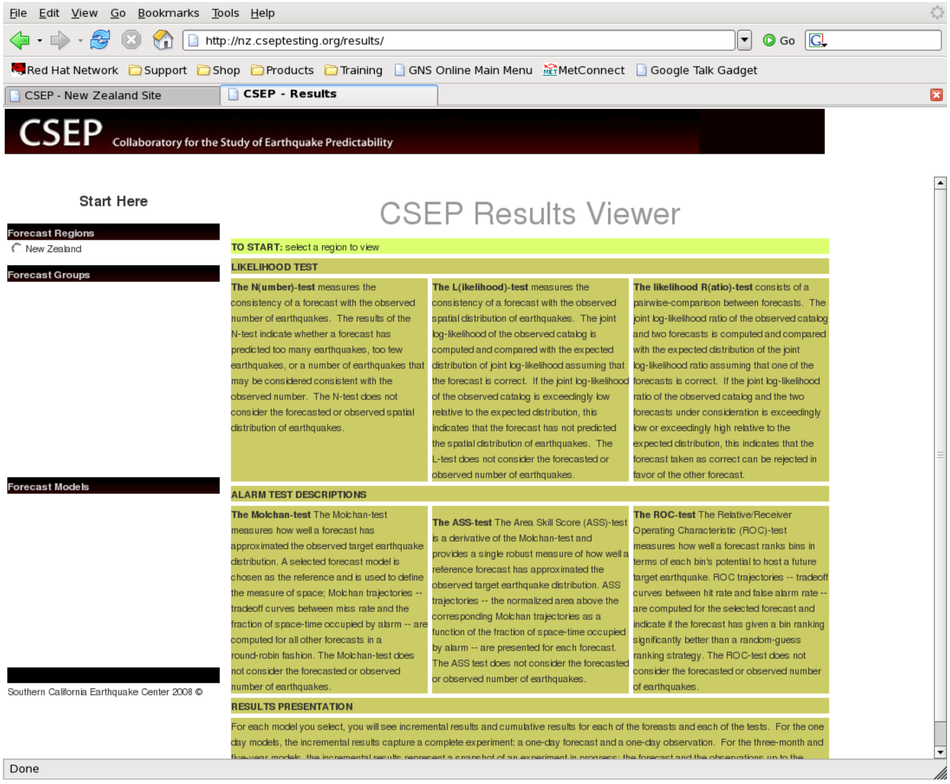

Figure 2 The online test results viewer: explanations of the various tests performed by the NZTC

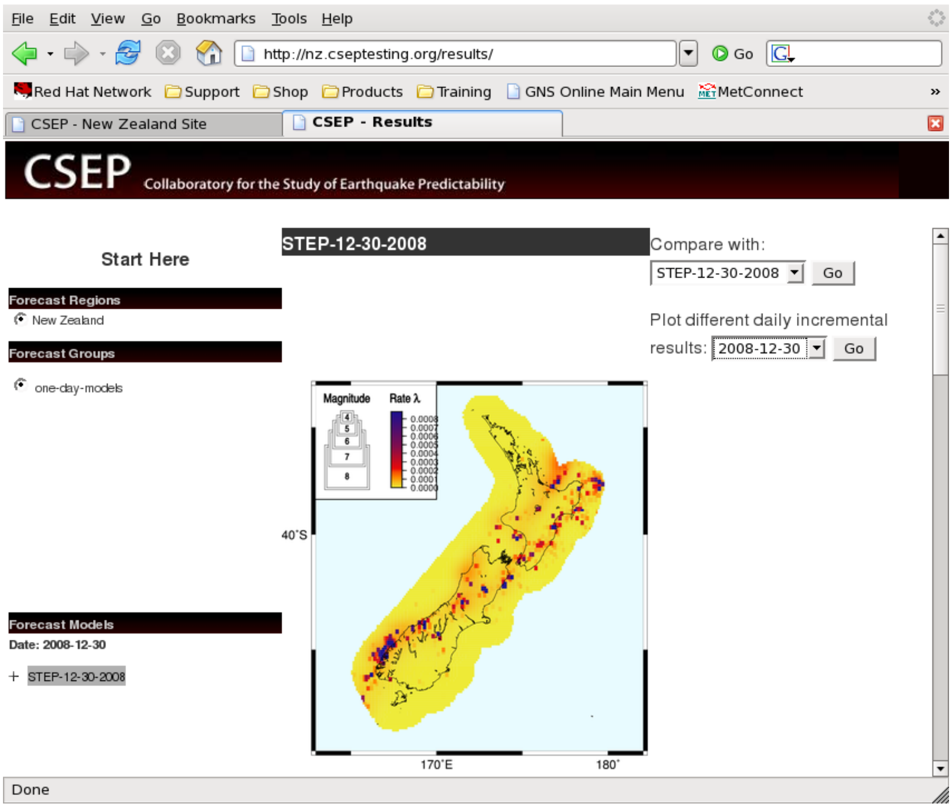

Figure 3a Forecast produced by a 1-day model on 30/12/08. The map shows the number of events ≥ magnitude 4 for the 24-hour testing period.

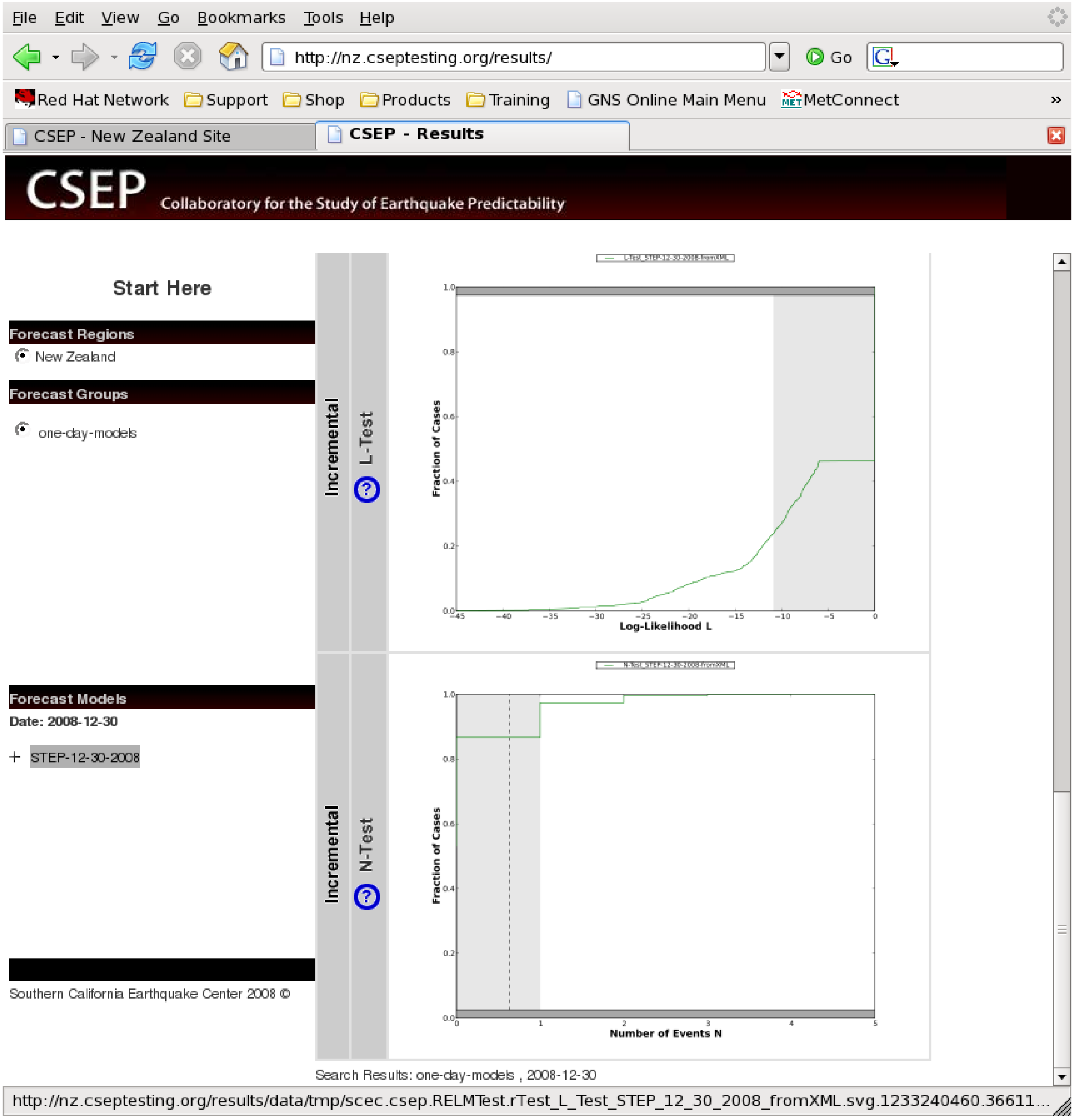

Figure 3b Evaluations of the 1-day forecast, shown in Figure 3a, using two different testing methods: the L-Test, which primarily tests the location of forecast events and the N-Test which tests the total number of events forecast. In both cases the models is consistent with the data.

2.4 Automated Testing: Prospective Testing Phase 1.

The ultimate goal of the NZTC is to test a suite of models, which target 3 different forecast periods (1-day, 3-months, and 5-years) in prospective tests over a 5-year period. To accomplish this aim, each model must be installed within the testing centre as described above. To be installed within the testing centre models must be adapted to the requirements of the testing centre. They must forecast the expected number of earthquakes in the appropriate magnitude range for the prescribed grid. They must also be able to read and write data in the specified formats. The most difficult challenge is that they must operate in a fully automated mode without any input from the researcher. In no cases are these trivial tasks. The final stage is to interact with the NZTC operators to ensure that the model is installed correctly and is behaving as expected. The following models have been submitted to the testing centre: NZ national seismic hazard model (Stirling, et. al., 2002), EEPAS (Rhoades & Evison, 2004), PPE (Rhoades & Evison, 2004), SUP (Rhoades & Evison, 2004), STEP (Gerstenberger, et. al., 2005), and Double branching model (Marzocchi and Lombardi, 2008). Of these, the national seismic hazard model, PPE, SUP and STEP are installed and are currently being tested. The EEPAS model is in the process of being installed and will be undergoing testing by the end of February, 2009. The Marzocchi and Lombardi model will be installed after the EEPAS installation is complete.

For testing we have three main forecast groups: 1-day, 3-month and 5-year. The 5-year group is split into two groups for testing: models tested against a declustered catalogue and models tested against a non-declustered catalogue. Currently, the 1-day category contains only the STEP model. The 3 month group contains EEPAS, PPE and SUP. Both of the 5-year declustered groups contain the national seismic hazard model, PPE and SUP. The Marzocchi and Lombardi model will be tested in the five-year, non-declustered group.

A 5-year forecast map as generated from the national seismic hazard model is shown in Figure 4. The version of the model installed within the NZTC is an interpretation of the original model. To be tested in the NZTC, a model is required to produce a gridded forecast of rates of events. The seismic hazard model consists of two rate components: rates of earthquakes on faults and rates of gridded background earthquakes. A strict interpretation of the fault contribution to the rates does not allow for gridded rates because most faults will pass through multiple grid nodes, but will have only a single expected rate for the entire fault length; however the fault contribution has been evenly divided up onto all grid cells through which the fault passes. This may cause a small penalty on the model during testing because the tested rates for a single grid cell are not consistent with the rate for the entire fault; however, the total rates for the tested model are equivalent to the total rates for the original hazard model. This is a consequence of grid and epicentre based testing and is a problem to be solved in future research.

2.4.1 Data quality issues

All Models tested require some sort of learning period prior to being tested. For example, if the STEP model is required to provide a forecast for tomorrow, the model will use all earthquake-catalogue data from approximately the last 5 years to develop this forecast; for this period the catalogue should not be missing events larger than about magnitude 3. Other models have variable requirements of longer and shorter time periods. Unfortunately, the catalogue maintained by GeoNet currently contains a gap of poor completeness from approximately April 2007 through July 2008. This means any model that includes this time period in its learning may be severely penalised in the testing. The only models potentially excepted from this are the 5-year models which may be less sensitive to the seismicity from a short time period such as that covered by the gap. GeoNet expects potentially two years or more before this gap is filled. Official testing within the testing centre cannot begin until this gap is filled in order to ensure fair and accurate testing; this means, as the GeoNet protocol currently stands, that official testing will not begin for a minimum of two years. Doing otherwise will produce spurious results and damage the credibility of the testing centre.

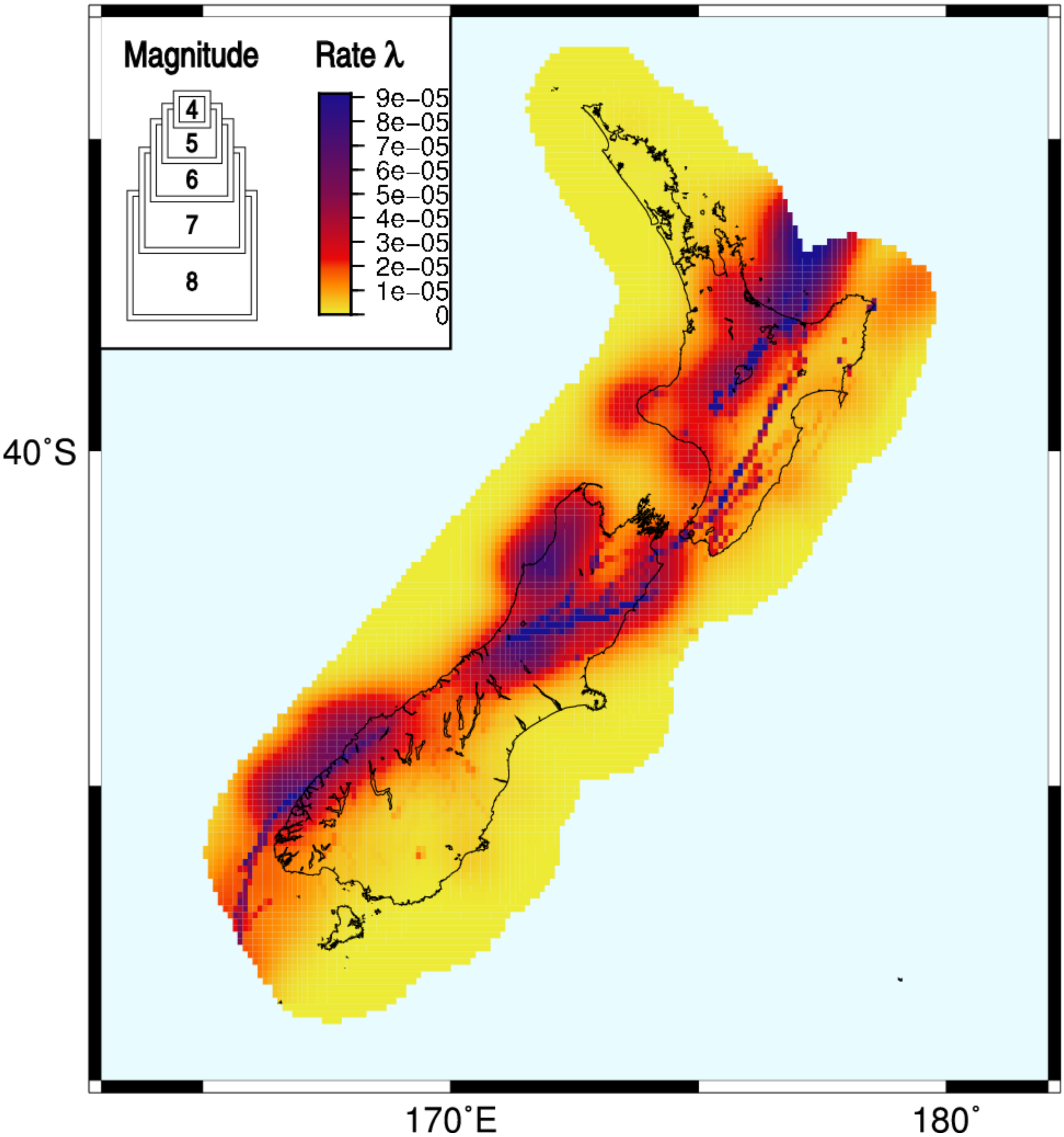

Figure 4 Forecas number of events of magnitude ≥ 5 in a five year period by the New Zealand national seismic hazard model.

3.0 FUTURE WORK

The operations of the testing centre will be an ongoing procedure for some years to come. However, with the completion of the work described in this report the practicalities of setting up a fully automated testing centre are now complete and future work will involve the addition of models to the testing regime and improvements to the testing centre. Improvements will include operational duties such as code upgrades and bug fixes, but will also include new testing methodologies and new model specification procedures being introduced. Several overseas researchers are ready or have expressed interest in submitting models to the NZTC. These models will be submitted and installed in the testing centre as they become ready. Additionally, alternate versions of some models already submitted to the testing centre are in preparation and will be tested. Expected models are: a reworking of the basic STEP model element to better account for generic aftershock triggering; another STEP version that aims to better handle low productivity aftershock sequences by using the Abundance model; a combined STEP and EEPAS model that shows an improvement over either individual model in preliminary investigations; an ETAS model; the 2008 version of the New Zealand national seismic hazard model; and, two alternate version of EEPAS, one that is earthquake rate dependent and another that includes aftershocks.

The first 5-year testing run will begin once a complete catalogue can be obtained from GeoNet where the existing completeness gap is covered. Once this initial testing run has finished, all results will be published in an international journal. While any test is not complete until this 5-year period has finished and will not be published, the interim test results may allow modellers to learn important information about their models as the testing progresses. By analysing the interim test results, modellers may opt to make changes to their model and submit the new model, as an additional model, to the testing centre. In this case, the initial model will be retained in the testing centre in order to maximise the learning potential from it. It is envisioned that operations of the NZTC will continue into the indefinite future.

4.0 ACKNOWLEDGEMENTS

This work was funded by the EQC Research Foundation under project no. TVH 2/11. Software developed by the California CSEP Earthquake Forecast Testing Center has been extensively used in setting up the New Zealand testing centre. David Rhoades and Mark Stirling have provided helpful reviews of the report.

5.0 REFERENCES

Gerstenberger, M., S. Wiemer, L.M. Jones, and P.A. Reasenberg (2005). Real-time forecasts of tomorrow's earthquakes in California, *Nature* **435,** 328-331.

Marzocchi, W., and A.M. Lombardi (2008), A double branching model for earthquake occurrence, *Journal of Geophysical Research*, 113, B08317, doi:10.1029/2007JB005472.

Reasenberg, P.A. (1985). Second-order moment of central California seismicity, 1969-1982. *Journal of Geophysical Research* **90**, 5479–5496.

Rhoades, D. A., and F. F. Evison (2004). Long-range earthquake forecasting with every earthquake a precursor according to scale. *Pure and Applied Geophysics* **161**, 47–71.

Rhoades, D., Gerstenberger, M., Christophersen, A., Savage, M., and Zhuang, J., Testing and Development of Earthquake Forecasting Models, *GNS Science Consultancy Report*, 2008/70.

Schorlemmer, D., and Gerstenberger, M.C., RELM Testing Center, *Seismological Research Letters* **78** (1), 30-36.

Schorlemmer, D.,Gerstenberger, M.C, Wiemer, S., Jackson, D.D. and Rhoades, D.A. (2007). Earthquake likelihood model testing, *Seismological Research Letters* **78** (1), 17–29.

Stirling, M.W., McVerry, G.H., and Berryman, K.R. (2002). A new seismic hazard model for New Zealand. *Bull. Seismol. Soc. Amer*. **92**, 1878–1903.2.0 …

www.gns.cri.nz

Principal Location

1 Fairway Drive
Avalon
PO Box 30368
Lower Hutt
New Zealand
T +64-4-570 1444
F +64-4-570 4600

Other Locations

Dunedin Research Centre
764 Cumberland Street
Private Bag 1930
Dunedin
New Zealand
T +64-3-477 4050
F +64-3-477 5232

Wairakei Research Centre
114 Karetoto Road
Wairakei
Private Bag 2000, Taupo
New Zealand
T +64-7-374 8211
F +64-7-374 8199

National Isotope Centre
30 Gracefield Road
PO Box 31312
Lower Hutt
New Zealand
T +64-4-570 1444
F +64-4-570 4657