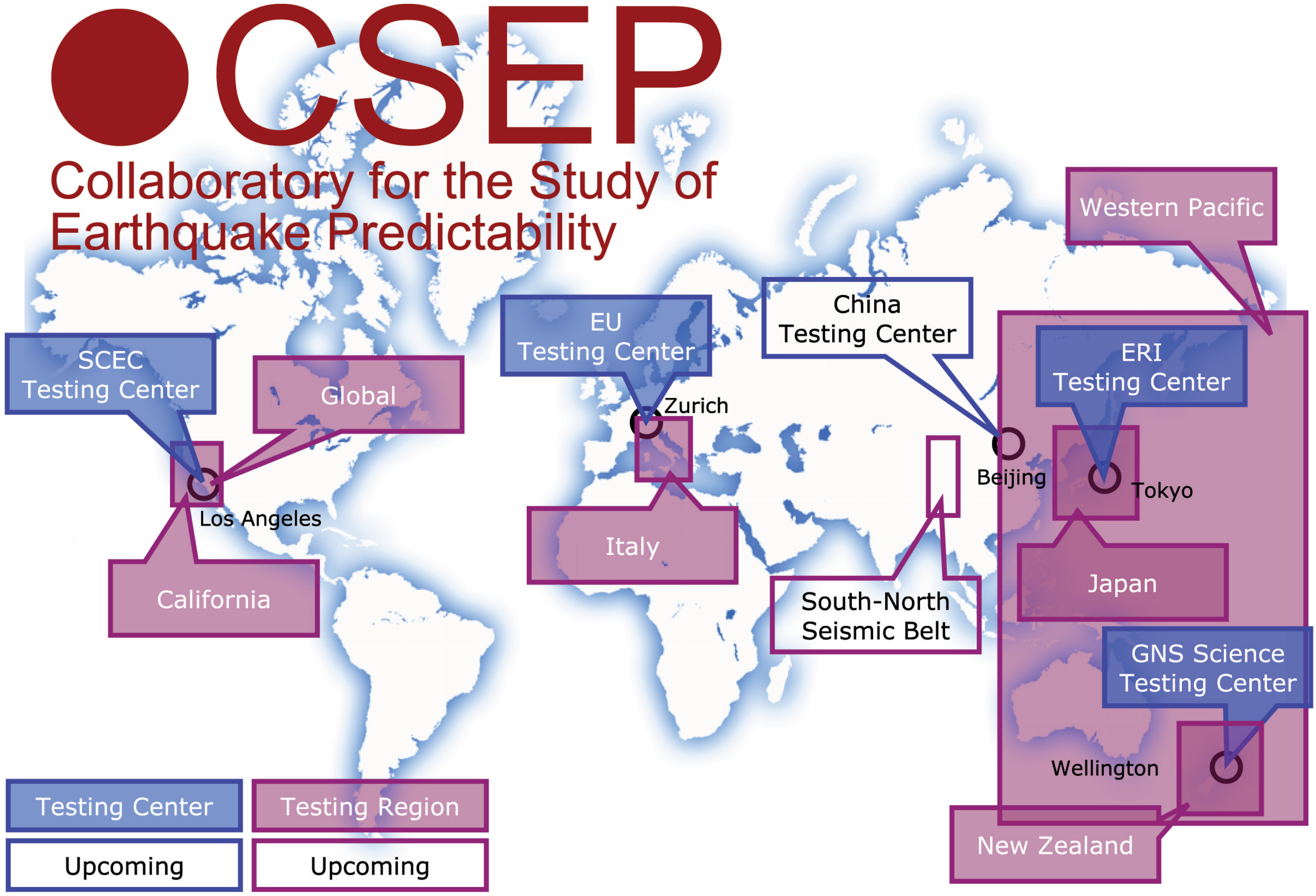

Figure 1 (top). Overview of the international CSEP testing centers and the testing regions.

Preface

Warner Marzocchi[1], Danijel Schorlemmer[2], Stefan Wiemer[3]

[1] *Istituto Nazionale di Geofisica e Vulcanologia, sezione di Roma, Italy*

[2] *Southern California Earthquake Center, University of Southern California, Los Angeles, USA*

[3] *ETH Zurich, Swiss Seismological Service, Zurich, Switzerland*

«A hypothesis that cannot in principle be put to the test of evidence may be interesting, but it is not scientifically useful.» [AAAS 1989]

This statement was reported by the American Association for the Advancement of Science —the largest general scientific society in the World—and it describes well the basic rationale that stands at the root of the Collaboratory for the Study of Earthquake Predictability (CSEP; www.cseptesting.org), a project initiated by the Southern California Earthquake Center. The study of earthquake predictability and forecasting-related research has in past decades been impeded by lack of an adequate experimental infrastructure. There was no possibility of conducting scientific prediction experiments under rigorous, predefined, and controlled conditions [Jordan 2006]. To remedy this deficiency, over the past 5 years, CSEP has been running a large earthquake forecast model testing program. This is structured into four testing centers [Schorlemmer and Gerstenberger 2007], in the USA, Japan, New Zealand, and Europe, which serve a variety of testing regions around the World (Figure 1). CSEP promotes rigorous scientific research on earthquake predictability through: i) an open and international collaboratory infrastructure; ii) rigorous and prospective testing of earthquake forecast/prediction models and hypotheses; iii) a global program covering a variety of tectonic environments.

The practical importance of these kinds of studies has been highlighted by the recent L'Aquila earthquake. On April 6, 2009, a M_w 6.3 earthquake occurred close to the city of L'Aquila, central Italy [Chiarabba et al. 2009]. The earthquake created substantial damage to the city of L'Aquila and its surroundings, causing more than 300 fatalities. The mainshock occurred after a few months of regionally increased seismicity, as a seismic swarm that was characterized by the largest event of local magnitude 4.2 only one week prior to the mainshock. This earthquake and the losses it caused has had a major impact on Italian seismology and politics. The death toll and building damage were high, given that the event was of moderate magnitude and that it took place in an area identified previously as one of high seismic hazard, according to national long-term hazard maps [MPS Working Group 2004]. Beyond that, there has also been a series of reproaches that the increased seismic activity, as well as claims of a radon precursor [Kerr 2009], had been ignored by the authorities, as a warning sign for an imminent disaster. While an international expert commission concluded that at present short-term earthquake prediction lacks scientific credibility, be it based on radon or other diagnostic precursors [Jordan et al. 2009], and while a study of the swarm that preceded the mainshock concluded that evacuations as a mitigation action were not warranted [van Stiphout et al. 2010], the intense controversy following the L'Aquila disaster once again highlighted the lack of a general consensus on earthquake predictability, and ultimately also the lack of progress made towards operational earthquake forecasting and its application in real cases.

In 2006, the European Commission funded the Integrated Infrastructure Initiative project NERIES (Network of Research Infrastructures for European Seismology; www.neries-eu.org). One of the targets of NERIES was the development of an integrated approach to seismic hazard assessment and earthquake forecasting that can span a multitude of spatial and temporal scales. A key requirement for achieving this goal is a community-accepted testing and model-evaluation framework for Europe that supports scientific experiments related to earthquake predictability in a controlled environment. In October 2008, the NERIES community agreed to create an earthquake forecast testing center for Europe, as the European node of the international CSEP efforts. The CSEP EU Testing Center, located at ETH Zurich, in Switzerland (cseptesting.org/centers/eth), will support experiments for a number of testing regions in Europe. The bylaws of the testing center, which are available on the CSEP EU webpage, regulate the operation and governance of the testing center.

In this special volume of the Annals of Geophysics, we report on the first CSEP testing region created within Europe: Italy. The high quality of seismic monitoring run by the Istituto Nazionale di Geofisica e Vulcanologia (INGV) has allowed the definition of a testing area comprising the whole Italian territory [Schorlemmer et al., this issue] and has made available an authorized and calibrated seismic catalog for the testing phase. In general, a CSEP experiment consists of comparing forecasts/predictions with the future target earthquakes in a truly prospective test (zero degrees of freedom). The experiment requires: i) definition of the testing area, characterized by high-quality seismic recordings; ii) exact description of the forecast; iii) exact definition of the input data (authorized and calibrated) ; and iv) definition of the measures of success. In this way, the output of each experiment is reproducible, transparent, and obtained in a controlled environment.

For the Italian testing region, fully prospective testing of models commenced on August 1, 2009. Testing is supported in three distinct testing classes:

1. For 5-year and 10-year forecasts (18 models submitted). These models define a forecast rate for each space/magnitude bin (starting from $M = 4.95$) for the period August 1, 2009 to August 1, 2014. The forecast rates for each bin had to be received by the testing center before August 1, 2009.

2. For 3-month forecasts (3 models submitted). These models define a forecast seismicity

rate for each space/magnitude bin (starting from $M = 3.95$) for consecutive 3-month periods. The 3-month models must be implemented at the CSEP EU Testing Center as software code that can independently and automatically compute seismicity forecast rates, based on one or more predefined authoritative streams of input data.

3. For 1-day forecasts (5 models submitted). These models define a forecast rate for each space/magnitude bin (starting from $M = 3.95$) for consecutive 1-day periods starting at midnight UTC. 1-day models must be implemented at the CSEP EU Testing Center as software codes that can independently and automatically compute seismicity forecast rates, based on one or more predefined authoritative streams of input data.

Table 1 summarizes the models submitted, and additional 3-month and 1-day forecast models are expected to be added with time. Most models are described in detail in this special issue, along with two articles that describe the set-up of the experiment [Schorlemmer et al., this issue] and the retrospective testing of the submitted 5-year and 10-year models [Werner et al., this issue].

Model name	**Paper in this volume** (* not included)	**Testing class**
TripleS-CSI	Zechar and Jordan	5/10 years
TripleS-CPTI	Zechar and Jordan	5/10 years
TripleS-Hybrid	Zechar and Jordan	5/10 years
TripleS	Zechar and Jordan	3 months
RI	Nanjo	5/10 years
RI_S	Nanjo	3 months
RI_L	Nanjo	3 months
DBM	Lombardi and Marzocchi (p. 31)	5/10 years
HZA_TI	Chan et al.	5/10 years
HZA_TD	Chan et al.	5/10 years
HiResSmoSeis-m1	Werner et al. (p. 107)	5/10 years
HiResSmoSeis-m2	Werner et al. (p. 107)	5/10 years
LTST	Falcone et al.	5/10 years
MPS04	MPS Working Group 2004 *	5/10 years
LASSCI	Pace et al.	5/10 years
HAZGRIDX	Akinci	5/10 years
PHM_Grid	Faenza and Marzocchi	5/10 years
PHM_Zone	Faenza and Marzocchi	5/10 years
ALM	Gulia et al.	5/10 years
HALM	Gulia et al.	5/10 years
ALM.IT	Gulia et al.	5/10 years
ETAS_LM	Lombardi and Marzocchi (p. 155)	1 day
ERS	Falcone et al.	1 day
ETES	Falcone et al.	1 day
STEP_NG	Woessner et al.	1 day
STEP_LG	Woessner et al.	1 day

Table 1. The models submitted for the CSEP experiment in 2009.

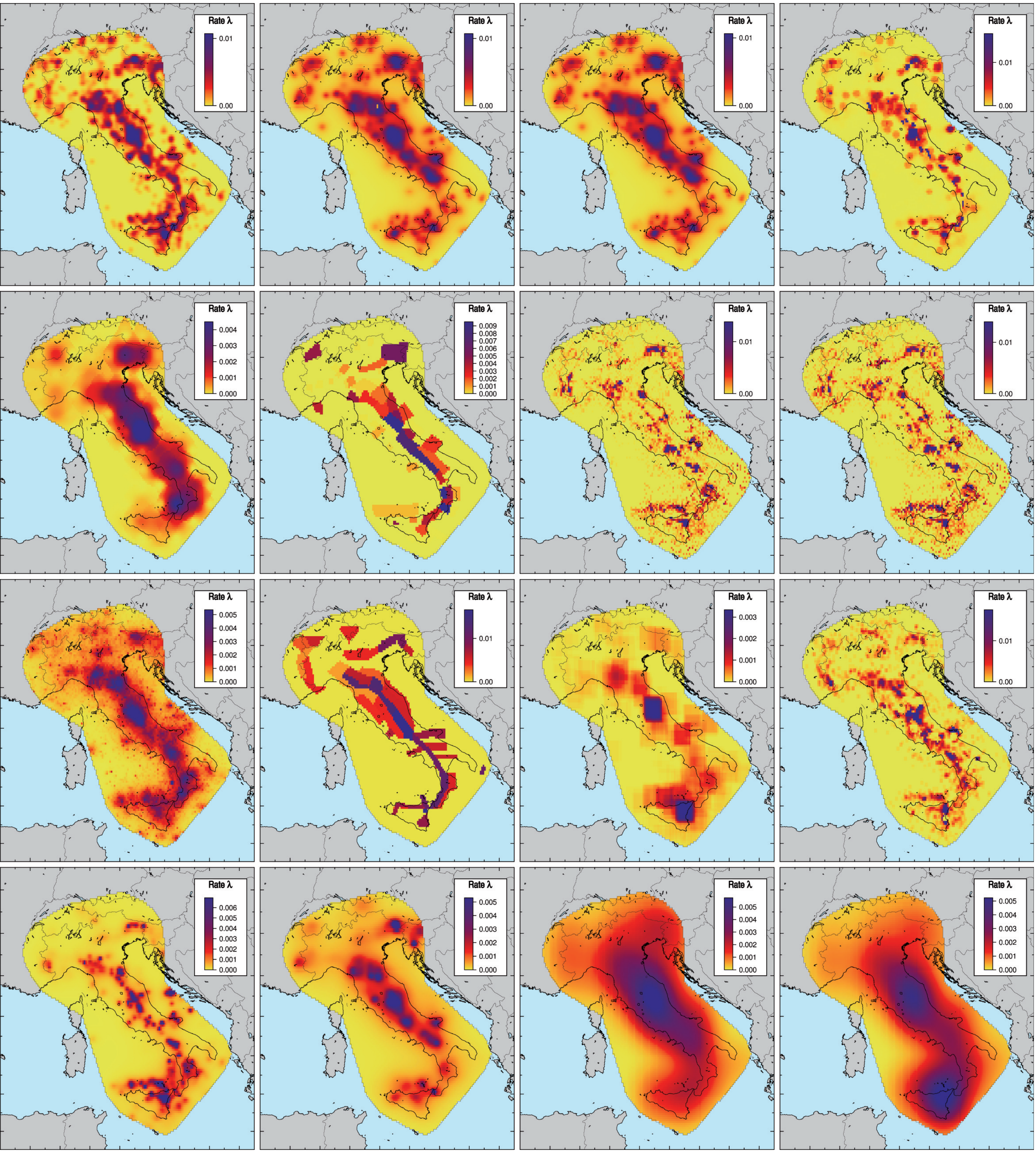

Figure 2. 16 of 18 the forecast models submitted for the Italian testing region. Color coded are the rates of forecast events for the next 5 years (note: scales are not the same). Prospective testing of these models commenced on August 1, 2009, and will last for 5 years. From upper left, moving right and down: HAZGRIDX, HZA_TD, HZA_TI, LTST, PHM_Grid, PHM_Zone, ALM, HALM, DBM, MPS04, RI, ALM.IT, HiResSmoSeis-m1, HiResSmoSeis-m2, TripleS-CPTI, TripleS-Hybrid.

The testing of these models against the observed seismicity will continue for at least 5 years (i.e. to August 1, 2014), as stipulated in a memorandum of understanding between INGV and ETH Zurich, which is also available on the CSEP EU webpage. CSEP has developed a range of tests that can be applied to measure the absolute and relative performances of these models against the data [Schorlemmer et al. 2007, Zechar et al. 2010]. Additional tests can be applied as the need arises, which is one of the advantages of a fully transparent and reproducible testing set-up. Note that the testing of the models is primarily a scientific experiment, and the results of the testing are intended to be analyzed by the scientific community with some care. They cannot serve as unfiltered input into decision-making on operational forecasting, nor can they be used readily by the public and media to educate themselves on the state of the art in earthquake forecasting. Therefore, access to the real-time testing results on the CSEP webpage is password protected and limited to the modelers and CSEP members. The CSEP EU Testing Center will report annually on the testing progress, and a final report will summarize the results of the experiment.

What can we expect from this experiment in the Italian testing region? Over the next 5 years, we expect to observe a sufficient number of target earthquakes to perform a detailed statistical analysis of the models. Figure 2 shows a composite of 16 of the 18 long-term models that were submitted. The most striking feature of the comparisons is the variability between the forecasts, which indicates that the ideas and models of how seismicity over the next 5 years will develop indeed vary greatly. Over these next 5 years, we expect to observe about 9 magnitude $M \geq 4.95$ earthquakes. This is a marginal number of events for a detailed statistical analysis; however, as detailed in Werner et al. [this issue], the first conclusions about how the models perform can be drawn up. The greatest strength of CSEP, however, is that the same models can be implemented in a number of testing regions, thus allowing trading space for time. Therefore, we anticipate that in a few years time, the CSEP process will allow scientist to discriminate between the more successful and less successful models, and ultimately to draw conclusions about the physical processes that govern earthquake occurrence.

References

American Association for the Advancement of Science (1989). Science for All Americans: A Project 2061 Report on Literacy Goals in Science, Mathematics, and Technology, Washington, D.C., 217 pp.

Chiarabba et al. (2009). The 2009 L'Aquila (central Italy) M_w 6.3 earthquake: Main shock and aftershocks, Geophys. Res. Lett., 36, L18308; doi: 10.1029/2009GL039627.

Jordan, T.H. (2006). Earthquake predictability, brick by brick, Seismol. Res. Lett., 77, 3-6.

Jordan, T.H., Y.-T. Chen, P. Gasparini, R. Madariaga, I. Main, W. Marzocchi, G. Papadopoulos, G. Sobolev, K. Yamaoka and J. Zschau (2009). Operational Earthquake Forecasting: State of Knowledge and Guidelines for Implementation, Findings and Recommendations of the International Commission on Earthquake Forecasting for Civil Protection, released by the Dipartimento della Protezione Civile, Rome, Italy, October 2, 2009.

Kerr, R.A. (2009). After the Quake, in Search of the Science—or Even a Good Prediction, Science, 324 (5925), p. 322; doi: 10.1126/science.324.5925.322

MPS Working Group (2004). Redazione della mappa di pericolosità sismica prevista dall'Ordinanza PCM 3274 del 20 Marzo 2003, Rapporto conclusivo per il Dipartimento della Protezione Civile, INGV, Milano-Roma, April 2004 (MPS04), 65 pp. + 5 appendices; http://zonesismiche.mi.ingv.it.

Schorlemmer, D. and M.C. Gerstenberger (2007). RELM testing center, Seism. Res. Lett., 78, 30-36.

Schorlemmer, D., M.C. Gerstenberger, S. Wiemer, D.D. Jackson and D.A. Rhoades (2007). Earthquake likelihood model testing, Seism. Res. Lett., 78, 17-29.

Schorlemmer, D., A. Christophersen, A. Rovida, F. Mele, M. Stucchi and W. Marzocchi (2010). Setting up an earthquake forecast experiment in Italy, Annals of Geophysics, 53, 3 (present issue).

Van Stiphout, T., S. Wiemer, W. Marzocchi (2010). When are mitigation actions warranted:

the case of the 2009 L'Aquila earthquake, Geophys. Res. Lett., 37, L06306; doi: 10.1029/2009GL042352.
Werner, M.J., J.D. Zechar, W. Marzocchi, S. Wiemer and the CSEP-Italy Working Group (2010). Retrospective evaluation of the five-year and ten-year CSEP-Italy earthquake forecasts, Annals of Geophysics, 53, 3 (present issue).
Zechar, J.D., M.C. Gerstenberger and D.A. Rhoades (2010). Likelihood-based tests for evaluating space-rate-magnitude earthquake forecasts, Bull. Seismol. Soc. Am., 100, 1184-1195.

Annals of Geophysics, 53, 3, 2010; doi: 10.4401/ag-4851