Mevr. Dr. Annemarie G. Muntendam-Bos
Senior Specialistisch Inspecteur
Staatstoezicht op de Mijnen

Prof. Dr. Stefan Wiemer
Chair of Seismology
Director, Swiss Seismological Service, ETH Zürich
Switzerland

Prof. Dr. Iunio Iervolino
Professor of Earthquake Engineering and Structural Dynamics
Università degli Studi di Napoli Federico II
Italy

Zurich and Naples, October 27, 2020

Subject: Support request of SodM for advice on model version recommendations

Dear Dr. Annemarie Muntendam-Bos, dear Annemarie

In your e-mail dated October 12, 2020, you requested from us advise to SodM on the Groningen public Seismic Hazard and Risk Assessment (SHRA) 2021 implementation. We are giving this advice as members of the KEM panel and in recognition that a broader KEM sub-panel for the Groningen SHRA is on ongoing process, one part of transitioning of the responsibility for the Groningen SHRA from NAM to TNO.

As an opening remark we like to note that, given the extensive material to review, given the very short amount of time available and considering also our own time constraints, we are unable to comment on the technical issues in greater depths. A number of our comments are thus related to procedural considerations and the overall approach to be followed. We hope that, in the future the sub-panel we will be able to follow the technical discussions more closely while they develop, and with a more defined schedule. This would allow us to be aware of and participate closely in the important technical discussions.

Overall, we note that this is a very special and somewhat delicate time for the Groningen SHRA assessment efforts. On the one hand, there are a number of just emerging scientific and technical advances to all components of the Groningen model train model, based on new data, new understanding and new assessments, but also correcting issues discovered in existing model versions. In that sense, is also rewarding to see that first results emerging from the KEM program are starting to be used. The SHRA model, however, is also growing ever more complex, documentation, peer review and model validation is in some components lagging behind. In addition, there are open scientific question raised with respect to various model components (such as tapers on earthquake size distributions) and on the computational implementation. Finally, and most importantly, this is a period of an ongoing transitioning in ownership, in review and in governance of the SHRA model, and a transitioning toward a fully public and open-access SHRA model. TNO is still in the process of gaining experience, capacity building, implementing quality and review procedures etc. In this stage of the model transition, it seems overall advisable to limit changes in the model components to the required minimum. The current Groningen SHRA is already a well-developed and highly sophisticated model

and modifications to this model should be done with great care, with strict quality control and also considering the impact of these changes on the ultimate model output (in terms of risk). We suggest to not change too many – or even any – model components in this transition time. We rather see model enhancements at a slower pace, but done properly and with sufficient reflection, discussion and testing, in order to spot problems and avoid making mistakes.

It is also to note that, given the quite unique case that Groningen represents in terms of seismic risk and the limited amount of data available from the field; several models are totally, or partially, based on assumptions and choices that reflect expert judgment and opinions. Whilst this is legitimate, if not necessary, it is difficult, if not impossible, to settle the debate on these choices uniquely; the goal should be, therefore, on the one hand to red flag modelling instances that are potentially not sound, do not reflect the state-of-the-art, or carry inconsistencies between all parts of the risk assessment. On the other hand, appropriate uncertainty quantification most be ensured so that alternative views can be integrated. This process also needs to be carefully structured within the new SHRA model governance.

Considering these aforementioned boundary conditions our overall recommendation with respect to the public Groningen SHRA 2021 is to apply only changes that are clearly required and ready to be implemented and otherwise postpone the implementation of new and update model components to the next model releases. This would allow for a thorough peer review, validation and sensitivity analysis of these components and their risk implications and allow TNO to assume ownership without rush and with the due confidence. If now too many changes are implemented at once, and then reversed in the next year, there is a chance that the results of the SHRA vary substantially with time. We feel that the open issues related to the NAM model discussed below are important and need to be resolved with care, we need to achieve a broad community consensus on such issues which does take time.

You asked in your mail a number of specific question that we address below.

1) Seismological Source Model (SSM)

- *What is your scientific opinion on discarding the v6 b-value model of a single b-value with exponential taper based on the analysis of TNO (advice of May 2020), provided the rebuttal of NAM and the reviews of the assurance panel members?*

- *The calibration of the SSM is currently not providing consistent parameter results. TNO was explicitely asked to address the differences and impact of the differences in their report. However, this has not been done. In addition, no description of the calibration procedure of TNO has yet been reported. Despite, TNO recommends using their own calibration. Do you think this is wise without being able to understand their approach first and without having identified where the differences originate from? Or would it better to this year still use the NAM updated calibration for consistency and first have a proper report on the TNO approach and identify the source of the differences? For your reference, neither calibration procedures have been part of the Tessella code assurance.*

Our response

The discussion on tapering and also on parameter calibration are important and potentially can have significant impact on the results. They are, however, also part of a broader discussion on alternative seismogenic source models, on uncertainty quantification and ensemble models and on formalized steps for model validation.

Going through the TNO report, NAM's rebuttal and the assurance panel comments, it appears to us that first of all this is an important and highly technical debate not yet fully resolved. We also perceive that the shortcomings relate in parts to the presentation of the TNO exercises, if not necessarily in the analyses itself. (Currently, the TNO experiments are not sufficiently described in terms of statistical tools and assumptions made. For example, terms such as 'bias' are not used as actually intended in statistics; this apparently cre-

ates some confusion.) Nevertheless, the criticism raised by TNO, and additional suggestions and issues mentioned by the assurance panel members, suggest that a wider debate in the community is needed and that changes to the b-value estimation model may be appropriate. We are not in a position to provide a final answer to issue of the exponential taper. We also lack input on the risk sensitivity of these choices. About this latter issue, we recall that the relevance of the modelling options should always be assessed in terms of their impact on the risk assessment, this being the primary output of the HRA. We have not seen this relevance assessment in the provided material.

We also note that the criticisms raised to the SSM extend beyond the tapering question and relate to alternative source representations and extended uncertainty quantification. There are ongoing research questions in KEM and in the community that may provide alternative source representations. In addition, the question of M_{max} and b-value computation (e.g., its spatial and temporal evolution) overall was supposed to be revisited in a workshop in the fall of 2020, now delayed. Therefore, reviewing the SSM model should be a TNO priority in 2021 and may then flow into the SHRA 2022. We thus embrace and broaden the TNO statement "*to review the Mmax-distribution (values and weights) in the logic tree through an expert elicitation process. This would also be an excellent opportunity to discuss the necessity and desirability of a supplementary (stress-dependent) taper branch point in the logic tree.*» In our assessment, this important workshop should be including alternative source representations and formal model validation and include an evaluation on the effect on risk assessment.

→ We recognize that TNO has raised relevant concerns on the use of exponential tapers and has provided also a relevant and potentially superior calibration procedure. However, in our assessment these criticisms and improvements need to be substantiated, peer reviewed and evaluate in their HRA impact before being used in the next HRA. For the HRA 2021, we therefore recommend using the same approach and the same weights used in the HRA 2020 (SSM v6).

2) Ground Motion Model (GMM)

- *Considering that v7 resolves a number of known problems of v6 (data issues and inconsistency in the damping model) as well as includes the effect of additional increased site response at Wierden, would you consider it wise to discard the update for HRA 2021?*
- *The identified discrepancy between the period-to-period variability needs to be resolved. However, from the report of TNO we derive the current implementation of NAM is conservative and leads to a higher risk than the proper implementation of Bommer et al (2019). Do you agree that the current implementation has therefore so far been conservative and a reparation of the issue would resolve in lower risk estimates?*

Our response

The GMM remains of critical importance in setting the hazard level and will typically have a large impact on risk. The GMM efforts in the Groningen HRA have in our assessment always been exemplary in their depth and thoroughness, and clearly represent the state of the art. Improvements, however, are always possible and also needed, given issues discovered in past data, new data acquired and progress in analysis techniques. It is clear that GMM v7 does offer a number of improvements that in general warrant the transition from v6 to v7. Again, we are here not in position to thoroughly review GMM v7 in depth. However, in our assessment there are three relevant issues related to the implementation of GMM v7 for the SHRA 2021:

1) Documentation and peer review are so far too limited right now. While preliminary reports and power points have been provided, the actual model implementation is not yet final, the model documentation not completed, the sensitivity analysis incomplete and the community review of the model incomplete. This will take time, and due to the ongoing global pandemic possible even more

time than anticipated. The process of completing and fully documenting the model should not be rushed or mistakes are more likely to happen.

2) TNO raises valid concerns on the model, which need to be resolved. Even if the model modifications are conservative in the sense that that tend to increase hazard and risk, we advise again implementing a known problematic model component, because it will undermine the acceptance of the model and may lead to swings in the results that are difficult to justify.

3) TNO also points out in their report that they may not have sufficient time to actually implement and test GMM v7 in their HRA engine, to a level that they are satisfied with the quality control. This is a considerable constrain that needs to be considered.

→ All three considerations suggest that while GMM v7 will certainly be an improvement, it is not yet ready for the SHRA 2021 but should be implemented in SHRA2022. We recommend using GMM v6 for SHRA 2021.

In addition, we like to point out that two statements in the TNO report need elaboration or clarification:

> *"For the vulnerability classes with the highest risk, a reduction up to 40% is observed in the center region. An important cause for the reduction is that the intensity measure used in the V6 and V7 fragility models is defined as the average of the spectral accelerations at 10 spectral periods. This averaging reduces the variability in the intensity measure relative to the variability of the contributing periods. However, this reduction effect is stronger if the variation of the ground motions for the individual periods are less correlated. The literal implementation of Bommer et al. (2019) therefore leads to lower variability in the intensity measure, which ultimately leads to a lower risk."*

The structural risk computed via integration of hazard and fragility, is an application of the total probability theorem. Therefore, in principle, it is insensitive of the chosen intensity measure used as an interfacing variable between the hazard and fragility. In light of this consideration, the statement about the variability of the average acceleration with respect to the variability of the contributing spectral accelerations (not periods) is not clear. We need this to be clarified before possibly giving an advice.

> *"However, the lack of correlation structure in the site response effectively compromises the generic correlation structure (according to Baker and Jayaram (2008)) imposed at the reference level."*

There are three forms of correlation in the ground motion intensity measures in the Groningen region: (i) correlation of spectral accelerations' GMPE residuals at different periods at the same site; (ii) spatial correlation of residuals of spectral accelerations for the same period at different sites; (iii) spatial correlation of residuals of spectral accelerations at different periods and different sites (iii degenerates in ii for the same period). Because the correlation refers to residuals of GMPEs that systematically include terms for site response, and Baker and Jayaram (2008) deal with (i), while this comment seems to deal with (ii) and/or (iii), it needs to be better clarified why this statement leads to questioning GMM V7. We need this to be clarified before possibly giving an advice.

3) Damage Model (DM)

- *Please provide us with your expert opinion on the adaptations (and substantiation of the adaptations) proposed by TNO to the FCM of NAM?*

Our response

TNO suggest using DM v7, but suggest a number of changes. While some of these may be reasonable, they seem in part ad-hoc, purely justified and lacking peer review and sensitivity review. There may in our assessment be not be major obstacle in installing DM v7 for the public Groningen SHRA 2021, but in order to

do so, a significant additional effort on the side of TNO to review, discuss, document and justify the choices is needed.

→ We recommend for TNO to finalize the DM v7 implementation at TNO but implement strict quality and peer review assurances, including a full sensitivity analysis. As a backup, DM v6 should be ready to be used for the SHRA 2021.

Specifically, we comment:

Ad 1) We may agree, but it is recommended to obtain a rebuttal by NAM and the DM assurance panel on the issue before making a final judgement.

Ad 2) *"In the development of the fragility functions NAM discovered that fragility curves resulting directly from MDOF models are in general more fragile than those of the SDOF models by 15% for masonry buildings. Therefore a 15% shift to lower values of the intensity measure is applied by NAM."*

How is a 15% fragility increase is measured (what 15% increase means) and why it is accommodated by a 15% left-lateral rigid translation of the fragility curves? In fact, this is an awkward and not recommended procedure. Modifications in fragilities should only come from modifications in the models (or specimens) the fragility represent a surrogate model of. Finally, these evaluations and modifications should include an analysis on the impact on risk. More in general, the various manual manipulations of fragilities and the logic tree weights on fragility seem somewhat arbitrary within the NAM v7 model in the first place and should in our assessment not necessarily be followed by TNO.

Ad 3) Same as comment to Ad 1 above.

Ad 4) Evaluation needs supporting documentation.

Ad 5) Note that PAGER mainly refers to the existing European building stock, which is built, in most of cases, not for seismic loads until very recently (or with obsolete seismic design procedures in the most fortunate cases). Therefore, this concern seems less relevant than what TNO thinks. A recommendation to NAM, if this is not what has been actually done, could be to make sure that the comparison with PAGER is done for the situations (i.e., structural taxonomy) as similar as possible to those of Groningen.

Moreover, the following statement is somewhat questionable in our assessment:

> *"Here, it must be noted that the consequence model describes probability of loss of life conditional on a certain collapse state. It is irrelevant if a certain collapse state is observed in a model at lower earthquake intensities than in a test. What is important for the consequence model is the debris coverage conditional on a collapse state."*

This, in fact, depends on the way collapse is defined and measured (by the way, *failure* or *damage state* would be more appropriate because *collapse* is ambiguous). Shake table tests are not carried out until debris are observed; i.e., not until actual collapse is observed. They are carried out until a certain response (i.e., some damage state) of the structure is observed, usually very far from collapse, for laboratory safety issues. Therefore it "*It is irrelevant if a certain collapse state is observed in a model at lower earthquake intensities than in a test*" could be questionable if collapse is the used, in turn, to compute fatality probability.

With respect to the Exposure Data Base (EDB): We agree with the TNO proposal to use the NAM Exposure Database v7 with upgrades carried out in 2020. Updating this new model component seems the least problematic and ready to go.

We hope that these comments are useful for your difficult ask of recommending a workflow for the next SHRA for the Groningen area. Please feel free to contact is in case that there are remaining questions or need for clarifications.

Sincerely yours,