Philosophical reflexions following the Lyttleton 2011 New Zealand Earthquake: Ten years after

P. Quintana Gallo
Pontifical Catholic University of Valparaíso, Valparaíso, Chile.

D. M. Carradine
BRANZ Limited, Judgeford, New Zealand.

ABSTRACT

At 12:51 p. m. on 22 February 2011, the M_w 6.3 Lyttelton earthquake struck the Canterbury region of New Zealand. The seismic event was reported as the result of the rupture of a previously unknown geological fault, 5.9 km deep, with an epicentre 6 km SSE of the Central Business District (CBD) of Christchurch. The Lyttelton earthquake followed the M_w 7.1 Darfield event, which occurred 4 September 2010, but did not cause extreme damage, particularly in engineered buildings. This fact satisfied the expectations of the engineering community, creating an apparent feeling of 'complacency'. The consequences of the Lyttelton earthquake, however, were much more severe in terms of destruction compared to Darfield. Amongst these, the intense damage (or collapse) suffered by engineered reinforced concrete (RC) buildings designed with modern codes was particularly worrisome. Several of these aspects deserve attention from a philosophical perspective. These are related to the inevitable fallibility of code provisions, which are continuously modified based on new experiences uncovering their downsides; and the belief that different seismic demands cannot occur at a given place in the future, based on prior evidence. It is shown why the current use of the theory of probabilities within a performance-based design framework might have conceptual pitfalls, and how ethics constitute an unavoidable part of the technical decisions and investigations. After briefly reporting the characteristics of the earthquake and part of the damage caused to one selected building, this paper discusses such philosophical aspects, based on the philosophy of science literature.

1 INTRODUCTION

This paper builds, in part, on personal experience, as the authors were in the city centre of Christchurch, in the Holiday Inn building on High Street, in front of Grand Chancellor (Figure 1), as many of those who were part of the New Zealand engineering community at the time. It presents philosophical questions that emerged after the February 11th 2011 Lyttleton Earthquake, and some of the answers (and further questions) provided in the literature by philosophers of science and of ethics.

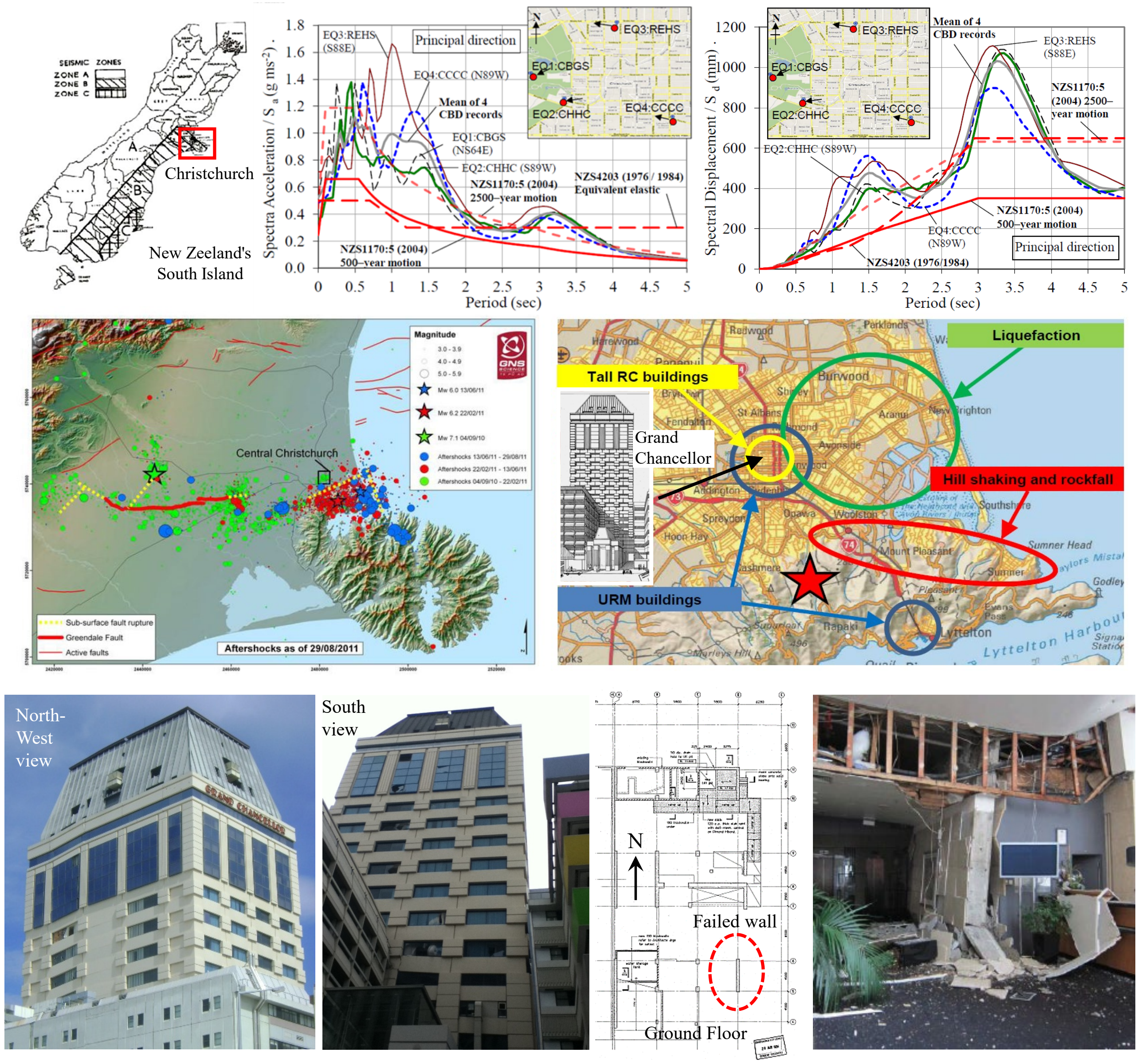

Figure 1: Overview of the Canterbury earthquakes' characteristics and the Christchurch scenario (adapted from Pampanin et al. 2011 and Quintana Gallo 2014); bottom-right picture courtesy of Nigel Priestley.

The Lyttleton 2011 earthquake was triggered by a previously unknown fault, with the only antecedent being the Darfield 2010 earthquake, occurred just six month ago. Nobody really expected this to occur, it was not really possible to anticipate such an event, given the background information on other earthquakes affecting the Canterbury region previously. The largest threat was the Alpine Fault, crossing both main islands of the country from north to south. This fault is approximately 100 km away from Christchurch, which means that the far-field event characteristics of the motion were expected. Large velocity pulses resulting from near-field shallow earthquakes, in turn, were not expected in the hazard scenario of the time. Proof of this is the near-fault factor N(T,D), prescribed by the New Zealand Standard NZS1170.5 (Standards New Zealand 2006), which is equal to 1.0 for Christchurch, and to 2.0 for Wellington, for example. The motion of the ground, as perceived by the first author, was radically different to that of a far-field earthquake, such as most of those that occurred in Chile (e.g. Valparaíso 1985 and Maule 2010). The vertical motion was perceived to be as strong as the lateral motion, and were overlapped, reflecting that principal and secondary seismic waves arrived almost at the same time, regardless of their different velocities. Moreover, for such a small fault-to-site distance, the phenomenon becomes largely complex.

Referring to engineered buildings, the level of destruction, i.e. the consequences of the earthquake, was large. We believe that, prior to the earthquake, there was a perception of safety in light of the rather low intensity of damaged produced in these structure by the Darfield earthquake. In addition, even though the PGA has shown not to correlate solely with damage, as the maximum PGA registered during Darfield at Christchurch stations was equal or smaller than that prescribed by the code requirements (Z = 0.22g). Nevertheless, the linear-elastic displacement spectra were shown to be larger than the design spectra for some natural periods of vibration, even for the most conservative case defined for “2% probability of exceedance in 50 years” (Figure 1).

This contribution covers a series of philosophical matters which are thought to be relevant to seismic engineering, and may help addressing some of the issues above, requiring understanding of the logic involved in the generation of scientific knowledge, the meaning of ‘validation’ in empirical science, the interpretation of probability; and the role of values (ethical, social, political) in technical knowledge.

2 POPPER’S CRITICAL RATIONALISM AND THE PROBLEM OF INDUCTION

In his *Logic of Scientific Discovery*, Karl Popper (1934) proposed a novel approach to the problem of induction (firstly outlined by David Hume) and of the status of scientific theories. Popper was inspired by the work of Albert Einstein in his General Relativity theory, which amongst other things, posed risky and counter-intuitive predictions, such as the curvature of the trajectory of light in space close to massive bodies (Popper 1967). Popper was concerned about the ‘scientific status’ of other theories, such as Marx’s theory of history, Freud’s psychoanalysis, and Adler’s individual psychology, and how they appeared to be different to Einstein’s relativity in the sense that they appeared to admit no counterexample able to refute them: somehow any new event was a confirmation of the theory (Popper 1967). In 1919, sir Edington tested the risky prediction related to the curvature of light implied by Einstein’s theory, providing a strong scientific status to this new and revolutionary theory of gravity, and refuting Newton’s ‘law’ of gravitation. In 1919-1920, Popper proposed a new vision of the nature of scientific knowledge, and what grants the status of ‘scientific’ to a scientific theory and makes them different from ‘pseudo-science’. In Conjectures and Refutations (Popper 1967, pp.47-48), Popper summarized this vision in the following points:

1. It is easy to obtain confirmations, or verifications, for nearly every theory-if we look for them;
2. Confirmations should count only if they are the result of risky predictions;
3. Every ‘good’ scientific theory is a prohibition: it forbids certain things to happen;
4. A theory which is not refutable by any conceivable event is not scientific (infallibility is a vice);
5. Every genuine test of a theory is an attempt to falsify it, or to refute it (testability is falsifiability);
6. Confirming evidence should not count except when it is the result of a genuine test of the theory;
7. It is possible to introduce ad-hoc auxiliary hypotheses to cope with refutations of a theory, at the expense of lowering its scientific status;

3 ON THE VALIDATION OF HYPOTHESES IN ENGINEERING

As in other areas of science and technology, it is easy to find the word ‘validation’ amongst the research publications on the field of seismic engineering. We suggest that the usage of this term is not totally correct, from a methodological or epistemological point of view. For, according to Popper’s falsifiability criterion, a theory is scientific if, in principle, it can be refuted by some empirical finding. Hence, it is not possible to establish the infallibility of any scientific theory or model, neither of a method or procedure. A mathematical model, either deterministic or probabilistic, in general, and a model or framework applied to seismic engineering in particular, cannot be conclusively validated by means of empirical data, and therefore, is always open to criticism, provisory and corrigible. This view might seem contradictory or at odds at first, as we are prepared for producing error-proof methods for solving practical tasks. However, these methods should be perceived as *the best method proposed to date, according to the existing empirical evidence, but refutable, in principle, at some stage in the future by further evidence*. And even if refuted, as Newton's theories of gravity and kinematics, they can still be useful under certain narrower conditions, but at the expense of lowering its scientific status (Point 7 in the list above).

It is also crucial to understand the difference between deductive and inference arguments in terms of logic. In logic, an argument consists of one statement which is the *conclusion* and one or more statements of supporting evidence, named *premises* (Salmon 1963). An inductive argument expands the content of premises by sacrificing necessity, whereas the deductive argument achieves necessity by sacrificing any expansion of content. In an inference (inductive) argument, if all the premises are true, the conclusion is not necessarily true, and have more informative content than the premises. As mathematics deal with deductive arguments, empirical science deals with inference arguments. If one ensures that the conclusions obtained with an inference argument are 'validated', in the sense of deductive correctness (Salmon 1963), one is committing a fallacy known as the fallacy of affirming the consequent or *principio participii*.

The impossibility of indicative probability is proofed as follows (Popper & Miller 1983). Given a hypothesis h, and a possible evidence e, deducible from h in the light of the background knowledge b, it can be shown that the logical statement $h = (h \leftarrow e)(h \vee e)$ holds true (h equal h if e and h or e). With this consideration, the probability of h given e, P(h, e), can be expressed as in Equation (1). As e is given true, $(h \vee e)$ is true, which leads to Equation (2), showing that $P(h \leftarrow e, e) = P(h, e)$.

$$P(h,e) = P\big((h \leftarrow e)(h \vee e), e\big) \tag{1}$$

$$P(h \leftarrow e, e) = P\big((h \leftarrow e)e, e\big) = P(he, e) = P(h, e) \tag{2}$$

Assume that e supports h, such that $P(h, e) > P(h)$. Question: Does e provide any support for the factor $(h \leftarrow e)$, which in the presence of e is alone needed to obtain h? Answer: no, e countersupports $(h \leftarrow e)$ unless one of both, $P(h, e) = 1$ or $P(e) = 1$, which are irrelevant cases. This means that if $P(h, e) \neq 1 \neq P(e)$, *e countersupports* $(h \leftarrow e)$, such that $P(h \leftarrow e, e) < P(h \leftarrow e)$. In terms of *Theorem 2* in Popper and Miller (1983): if $P(h, e) \neq 1 \neq P(e)$, $Exc(h, e)$, the '*degree of counter-support*' of h given e, is such that:

$$Exc(h,e) = P(h \leftarrow e) - P(h \leftarrow e, e) = P(\neg h, e)P(\neg e) > 0 \tag{3}$$

As written by Popper and Miller (1983): "*the factor that contains all of h that does not follow deductively from e is strongly counter-supported by e. It is counter-supported the more the greater the content of e, which may be measured by ct*(e, b) = 1 – P(e, b)… *A result totally devastating for to the inductive interpretation of the calculus of probability.*"

On the grounds outlined above, we argue that it is not correct to claim that a given procedure or method for determining the seismic demand over a structure located at a given place, such as a design response spectrum, is 'validated' (meaning infallibility). In fact, it is also true that any methodology, theory or code prescription, *needs* to be fallible in the sense of being refutable, in principle, to be scientific. Hence, from the fact that a design spectra is found to be conservative compared to those constructed with the ground motions recorded during prior earthquakes, it does not follow that it will be conservative for those concerning future events. Furthermore, to be scientifically meaningful, such design spectra *must* be refuted at some stage. In the case of the New Zealand design spectrum, even for the minimum 2% probability of exceedance in 50 years (see Figure 1), it was found that it was non-conservative compared to the records gathered in 2011, for a large range of periods (further found to also be the case following the Kaikoura 2016 earthquake). Hence, its security, partially confirmed during the 2010 Darfield earthquake, was falsified, and led to a modification to its amplitude by means of an increase of the hazard factor Z from 0.22 to 0.30 after the Canterbury Earthquake Royal Commission (CERC 2011).

Now, what if a methodology does not forbid any possible outcome of a given experiment or natural phenomenon (e.g. throw of a dice or the consequences of an earthquake), but provides different probabilities to all of the possible outcomes? Priestley et al. (2007) seem to have encountered this problem when commenting on probabilistic seismic hazard analysis. On page 55 of Priestley et al. (2006) it reads: "*Consider a rather artificial case where seismic hazard is dominated by a fault which fractures every 1000 years, on average inducing PGAs at the site being considered of 0.8g. The remainder of seismic risk is attributed to background seismicity capable of inducing PGAs of only 0.2g. A probabilistic analysis might then determine the 475 year risk as being characterized by a PGA of 0.4g, since the probability of occurrence of the design earthquake occurring within the 475 year period is approximately 50%. However, the earthquake either occurs during that period, or it doesn't. If the structure is designed for 0.4g and the earthquake occurs, the structure fails, since the intensity is much greater than the design strength. If the earthquake doesn't occur, the structure is over-designed*".

Of course one may further argue that it is also true that for any usable life of a structure, and any probability of exceedance of the nominal collapse (failure) index (PGA in this case), a particular structure in a specific location, either fails or does not fail during that period. This issue is known in philosophy of probabilities as the "single-value problem", addressed in the next section.

4 PROBABILISTIC APPROACH TO PBEE AND THE SINGLE-VALUE PROBLEM

The widely known PEER Performance-Based Earthquake Engineering (PBEE) framework was introduced during the late 90's and early 2000's (e.g. Cornell 1996, Hamburger 1996, Deierlein 2004, Krawinkler et al. 2005, Kunnath 2007). It makes use of conditional probabilities for constructing an equation able to provide the mean annual frequency (MAF) of exceedance (λ) of a decision variable (DV) for a given location (L) and building design (D), given a damage intensity measure (IM), given an engineering demand parameter (EDP), given an intensity measure (IM) of the seismic demand for L and D with MAF of exceedance λ(IM) with a given probability distribution function. This expression, derived from the total probability law (Equation (4)), can be written as in Equation (5) (Kunnath 2007).

$$\lambda(x) = \int_0^\infty G(x|y)\, g(y) dy \tag{4}$$

$$\lambda(DV) = \int_0^\infty \int_0^\infty \int_0^\infty G(DV|DM)\; g(DM|EDP)\; g(EDP|IM)\; g(IM)\; dIM\; dEDP\; dDM \tag{5}$$

In Equation (4): $\lambda(x) = \mathrm{P}[X \geq x]$ = MAF of exceedance of x, $G(x|y) = \mathrm{P}[X \geq x|Y = y]$ = conditional probability of exceedance (CPE) of x given y; $g(y)$ = exceedance probability density function (EPDF) of y; X and Y are random variables. In Equation (2): $G(DV|DM)$ = CPE of DV given DM; $g(DM|EDP)$ = EPDF of DM given EDP; $g(EDP|IM)$ = CPE of EDP given IM; $g(IM)$= EPDF of IM. As it holds that $g(x|y) = dG(x|y)/dx$, the terms $g(x|y)dx$ in Equation (5) can be replaced by $dG(x|y)$. Similarly, as $g(x)dx = d\lambda(x)$, the term $g(IM)dIM$ can be replaced by $d\lambda$(IM). Using these equivalencies, the more compact form of Equation (5), presented in Equation (6), is obtained:

$$\lambda(DV) = \int_0^\infty \int_0^\infty \int_0^\infty G(DV|DM)\; dG(DM|EDP)\; dG(EDP|IM)\; d\lambda(IM) \tag{6}$$

In the particular case where only EDP is evaluated given IM, $\lambda(EDP)$ is calculated with Equation (7):

$$\lambda(EDP) = \int_0^{\infty} G(EDP|IM)\, d\lambda(IM) \tag{7}$$

The MAF of exceedance $\lambda(EDP)$ is used for computing the 'probability of collapse' (e.g. Cornell 1996, Krawinkler et al. 2005, Kunnath 2007), $P_C = P[EDP > EDP_C] = \lambda(EDP_C)$, with EDP_C the value of EDP defining the nominal collapse of a structure. Hence, a given structure (single-case) is said to 'to have a mean annual probability of collapsing equal to P_C', which is equivalent to saying that '*a particular structure (built in L with design D) collapses with a mean annual probability equal to* P_C'.

Can a frequency probability measure be assigned to a sample comprising of a single element (single-case problem), such as the building hosting this conference? The answer is *depends on the philosophical interpretation of probability* (see Hajek 2019). From the many existing interpretations (e.g. classical, logical, and subjective probability, and frequency, propensity and best-system interpretations), to address the single-case problem, we select: (1) von Mises frequentist interpretation (used in the reviewed probabilistic framework), and (2) Popper's propensity interpretation.

4.1 Frequentist interpretation of probabilities and the single-case problem

Single-case probability assignments are well known to be incompatible with a frequentist understanding of probabilities, which include concepts such as mean probability of exceedance, long-run probability (limit probability), etc., as in the probabilistic formulation described above. Miller (1991) in his article "Single-Case Probabilities", provides a sharp review of the problem, remarking that one of the founders of the frequentist school, von Mises (1951), wrote: "*probabilities belong not to individual events, but kinds of events (attributes) in an unending repetition of an experiment or natural recurring phenomenon*." In the frequency theory, the probability of an attribute is defined only relative to a collective, and can be different for different collectives. It makes no sense, according to von Mises (1951), to talk of the probability of a single event. In von Mises (1951) a collective is understood as: "*a potentially infinite sequence in which these limits exist* [the axiom of convergence] *and remain unaltered in subsequences selected by any disinterestedly specified gambling system* [the axiom of randomness]". For example, the attribute 'death before 50' is empirically found to have different long-run frequencies in the class of men aged 49 and in the larger class of men and women aged 49; if these classes are authentic collectives, the probability of death before age 50 varies with the collective to which it is referred. It makes no sense to talk of the probability of a single event, as the death before his 50th birthday of a particular man aged 49 (Miller 1991).

Miller (1991) claims that the main problem of the frequentist interpretation is that because its axioms (convergence and randomness) concern only what occurs in the limit (after an infinite repetition of realizations and therefore outcomes of an experiment), which cannot be reached, statements of probability are "drained from empirical content": a given actual finite sequence of outcomes can reflect an apparent regularity which is not attained in the long run, and vice-versa.

4.2 Popper's Propensity Interpretation of probabilities

Popper (1934) proposed an alternative philosophical understanding of probability, named the "*propensity interpretation*" (PI). In PI, the probability of an outcome is not a measure of any frequency. In turn, it is a measure of the inclination of the current state of affairs to realize that outcome (Miller 1991). In PI, the probability space is the set of all future, completely specified, possibilities that might result from the perpetually changing present (Miller 1991). It is only if the propensity of an outcome remains constant in globally different (though locally similar) situations that we can expect to measure its 'true' value empirically. Therefore, either an ensemble or a repeatable experiment is needed.

In PI, the probability of an outcome is no longer conditional or relative: it is undefined in the absence of a reference frame and can take different values at a time and also assume different values at different times or conditions. It is only *conditional in the sense that it is a function of the conditions*. A conditional probability, $p(a, b) = p(ab)/p(b)$, is then what the propensity of a would be were b to be obtained. True conditional probabilities are relative to the present state of the affairs (absolute probability) as well as to yet unrealized conditions. Within PI, singular probability statements are "statements of the propensity of the balanced state of affairs to eventuate in one outcome rather than another (Miller 1991). A singular statement of probability makes implicit reference to the "objective statistical state of affairs" (and hence to a collective) (Miller 1991). Strictly, every propensity must be referred to the complete situation of the universe at the time: they depend on the situation *today*. Only in this way the problem of the single-case can be addressed.

As an example, the probability of a building surviving collapse one more year from today varies day by day, and is affected by its use, by the progress of geology, and ultimately by inescapable aging of the materials. The propensity *today* of the building hosting *this* specific conference to survive another year without collapsing, given that there will be an $M_w = 6.0$ earthquake in one month time should be different from the propensity of surviving, given that there won't be such event. The term $p(b, a)$ is "the propensity for today's world to develop in one month time into a world in which there is a M_w 6.0 earthquake given that it (today's world) will by the end of the year have developed into a world in which the building is still standing, i.e. did not collapse. In turn, $p(a,b)$ is what the propensity for the earthquake to occur in a month time would be today if something were to happen today that guaranteed the survival of the building for another year. Furthermore, the propensities would be different if the focal distance of the earthquake was 100km or 5km.

5 THE ROLE OF VALUES IN SCIENCE AND ENGINEERING

Given the prior discussions, further issues arise: how to define what is 'strong enough' evidence for accepting a theory, model, prototype, etc., and how to define what is a 'small enough probability' of a given parameter to be exceeded. It is widely accepted in philosophy of science that the answers to these questions require consideration of non-epistemic (social, ethical, political) values. In his seminal work, Rudner (1953) argued that: "*…the scientist* [*qua-scientist*, i.e. while acting as such] *does make value judgments. For, since no scientific hypothesis is ever completely verified, in accepting a hypothesis the scientist* [engineer] *must make the decision that the evidence is sufficiently strong or that the probability is sufficiently high to warrant the acceptance of the hypothesis. Obviously our decision regarding the evidence and respecting how strong is "strong enough"* [or what a 'sufficiently high' probability is], *is going to be a function of the importance, in the typically ethical sense, of making a mistake in accepting or rejecting the hypothesis.*"

Later on, Hempel (1965) claimed that values are required in the definition of the criteria for accepting or rejecting a hypothesis, i.e. when determining whether a hypothesis is presumably true or presumably false. He introduced the concept of 'inductive risk': the risk of error in accepting or rejecting a given scientific hypothesis (to accept a hypothesis as presumably true and it is false, or to reject the hypothesis as presumably false and it is true). To properly construct rules of acceptance, non-epistemic values able to balance the consequences of committing inductive errors are required. This is of particular relevance if the acceptance of a hypothesis may lead to a given course of action with non-epistemic effects (Douglas 2010).

The answer to the issue on account relies on ethical-consequentialist judgements, rather than on epistemic (technical) consideration. In the facts, however, the definition of acceptable limits such as an acceptable probability of exceedance, may simply rely on tradition.

6 CLOSURE

This contribution covered a series of philosophical matters which are relevant to seismic engineering. These can be summarized as: (1) the logic of scientific discovery consists of the postulation of conjectures and the attempts for their refutation (trial and error) (Popper); (2) the status of scientific can only be attributed to :

refutability is needed; (3) validation belongs to deductive arguments only (mathematics): it is not correct to claim that a model, procedure, prototype, etc., is validated in a conclusive sense, as empirical science deals with inference (inductive) arguments which cannot be proofed by any amount of evidence; (4) when a probability measure is assigned to a specific unique entity (e.g. to *a specific building*), the single-vale problem is encountered: this makes no sense within a frequentist perspective (von Mises), such as that reviewed in the paper; (5) if a probability measure is assigned to a specific entity (building), it should be understood as a propensity measure (Popper's propensity interpretation): the probability of the current state of affairs to result into a world where the outcome under investigation develops; and (6) value judgements are involved in technical decision-making: ethical considerations in the sense of the consequences of being wrong provide a reference for determining what strong-enough evidence and what a sufficiently-small probability of exceedance are.

7 REFERENCES

Canterbury Earthquakes Royal Commission, CERC (2012). Final Report – Volume 1: Summary, Christchurch, New Zealand.

Cornell, C (1996). Calculating building seismic performance reliability: a basis for multi-level design norms. *In: Proc. 11th WCEE*, Acapulco, Mexico.

Deierlein G (2004). Overview of a comprehensive framework for Earthquake performance assessment. In: *Performance-Based Seismic Design Concepts and Implementation*, PEER Report 2004/05, Berkeley, CA.

Douglas, H (2010) Inductive Risk and Values in Science. *Philosophy of Science*, 67:559-579.

Hajek A (2019). *Interpretations of Probability*. The Stanford Encyclopedia of Philosophy (Fall 2019 Edition), Stanford University, Stanford, CA.

Hamburger RO (1996). Implementing performance-based seismic design into engineering practice. *In: Proc. 11th WCEE*, Acapulco, Mexico.

Hempel CG (1965). Science and Human Values. In: Aspects of scientific explanation and other essays in the -philosophy of science. New York: The Free Press, 81-96.

Krawinkler et al. (2005) Van Nuys Hotel Building Testbed Report: Exercising Seismic Performance Assessment. PEER Report 2005/11, Berkeley, CA.

Kunnath SK (2007). Application of the PEER PBEE Methodology to the I-880 Viaduct. PEER Report 2006/10, Berkeley, CA.

Miller D (1991). Single-case probabilities. *Foundations of Physics*, 21:1501–1516.

Pampanin S, Kam WY, Akguzel U, Tasligedik AS, Quintana Gallo P (2012). Seismic Performance of Reinforced Concrete Buildings in the Christchurch CBD in 22 February 2011 Earthquake. Part I: Overview. Technical Report University of Canterbury, Christchurch, New Zealand.

Popper K (1934). *Logik der Forschung*, Springer, Berlin, Germany (*original version in German*).

Popper K (1963). *Conjectures and Refutations*, Routledge, New York, NY.

Popper K & Miller D (1983). A proof of the impossibility of inductive probability. *Nature*, 302(21):687-688.

Priestley MJN, Calvi GM, Kowalsky MJ (2007). *Displacement-Based Seismic Design of Structures*. IUSS Press, Pavia, Italy.

Salmon WC (1967). The Foundations of Scientific Inference, University of Pittsburgh Press, PA.

von Mises R (1957). *Probability, Statistics and Truth*. George Allen & Unwin, London, UK.