Journal of Computational Finance **23**(2), 1–27
DOI: 10.21314/JCF.2019.371

Research Paper

A new approach to the quantification of model risk for practitioners

Zuzana Krajčovičová,[1] Pedro Pablo Pérez-Velasco[2] and Carlos Vázquez[1]

[1]Department of Mathematics, University of A Coruña, Campus de Elviña, 15071 A Coruña, Spain; emails: z.krajcovicova@udc.es, carlosv@udc.es

[2]Model Risk Division, Banco Santander, Avenida de Cantabria s/n, 28660-Boadilla del Monte, Spain; email: pedpperez@gruposantander.com

(Received May 29, 2018; revised July 26, 2018; accepted August 30, 2018)

ABSTRACT

Global regulation obliges financial institutions to manage model risk with the same severity as any other risk. Its quantification is therefore essential to meet these requirements and to ensure an institution's basic internal operations are able to run smoothly. In this paper, we address the quantification of model risk by calculating the norm of an appropriate function defined on a Riemannian manifold endowed with a Fisher–Rao metric. Our aim is twofold: to introduce a mathematical framework that is sufficiently general and sound to cover the main areas of model risk, and to illustrate how a practitioner can identify the relevant abstract concepts and put them to work.

Keywords: model risk; uncertainty; Fisher–Rao information metric; model manifold; exponential map.

Corresponding author: C. Vázquez

Print ISSN 1460-1559 | Online ISSN 1755-2850

1 INTRODUCTION

Models are simplified mappings of reality that serve a specific purpose: applying mathematical, financial and economic theories to available data. They deliberately focus on specific aspects of reality and downgrade or ignore the rest. Understanding the capabilities and limitations of underlying assumptions is key when dealing with a model and its outputs. According to the US Federal Reserve (2011), model risk is defined as

> the potential for adverse consequences from decisions based on incorrect or misused model outputs and reports. Model risk can lead to financial loss, poor business and strategic decision making, or damage to a bank's reputation.

Thus, the US Federal Reserve (2011) identifies two main reasons for model risk: inappropriate usage and fundamental errors. Further, it states that model risk should be managed and addressed with the same severity as any other type of risk, and that banks should identify the sources of model risk and assess their magnitude. The US Federal Reserve also emphasizes that expert modeling, robust model validation and a properly justified approach are necessary elements in model risk moderation, although they are not sufficient in and of themselves and should not be used as an excuse to avoid improving models.

Despite the rise in awareness of model risk and an increased understanding of its significant impact, there are no globally recognised industry or market standards regarding its exact definition, management and quantification, even though, as mentioned above, proper model risk management is required by regulators.

Within the finance literature, some authors have defined model risk as

- uncertainty about risk factor distribution (Gibson 2000),
- misspecification of an underlying model (Cont 2006),
- deviation of a model from a "true" dynamic process (Branger and Schlag 2004),
- discrepancies relative to a benchmark model (Hull and Suo 2002), and
- inaccuracy in risk forecasting due to estimation error and using an incorrect model (Boucher *et al* 2014).

Model risk has previously been classified in all asset classes; see Morini (2011) for interest rate products and credit products, Cont (2006) for option-pricing models, Christodoulakis and Satchell (2008) for portfolio applications, Saltelli *et al* (2013) for asset-backed securities and Boucher *et al* (2014) for model risk's relation to measuring market risk.

Quantification, an essential part of model risk management, is required for the consistent management of model weaknesses and limitations as well as their effective communication to decision makers and users. It is also necessary for assessing model risk in the context of an organization's overall position. The quantification of model risk should consider the uncertainty stemming from mathematical technique selection (eg, by focusing on fitting a normal distribution and, hence, leaving aside other distribution families), the calibration methodology (eg, different optimization algorithms may derive different parameter values) and limitations of the data sample (eg, a sparse or incomplete database).

Model risk quantification poses many challenges due to the high diversity of models, the wide range of techniques and the different implementations of models, among others. Some model outputs drive decisions, while other model outputs provide one source of management information. Some outputs are further used as inputs in other models. In addition, model outputs may be completely overridden by expert judgment. Finally, in order to quantify model risk, you need another model, which will also be prone to model risk.

It is common practice to consider the most relevant areas of analysis for the quantification of model risk to be data and calibration, model foundations, model performance, IT infrastructure, model use, controls and governance, and model sensitivity. The model may be fundamentally defective because of errors in its theoretical foundation or conceptual design that have emerged from incorrect assumptions, model misspecification or the omission of variables. Data quality issues, inadequate sample sizes and outdated data contribute to model performance issues such as instability, inaccuracy or bias in model forecasts. Model risk also arises from inadequate controls over model use. Flawed test procedures or failure to perform consistent and comprehensive user-acceptance tests, to name just two errors, can lead to material model risk.

The focus of this paper is on developing a new approach for quantifying model risk within the framework of differential geometry (see Murray and Rice 1993) and information theory (see Amari *et al* 1987). In this work, we introduce a measure of model risk on a statistical manifold, where models are represented by a probability distribution function. The differences between models are determined by the geodesic distance under the Fisher–Rao metric. This metric allows us to use the intrinsic structure of the manifold of densities and to respect the geometry of the space we are working on, ie, it accounts for the nonlinearities of the underlying space.

The rest of this paper is structured as follows. In Section 2, we summarize the central concepts from Riemannian geometry and introduce the terminology used throughout the paper. Our modeling process and a general description of our proposed method for the quantification of model risk are presented in Section 3, which is followed by a detailed discussion of the main quantification steps. Sections 4 and 5

describe the construction of a neighborhood containing material variations of the model and provide a definition of the weight function, respectively. The model risk measure is then defined and explained in Section 6. Section 7 is dedicated to an empirical example of the model risk calculation of a credit risk model. Section 8 provides some final conclusions and directions for future work, and, finally, the online appendix contains the construction of the weight function and the proofs of our main results.

2 BACKGROUND ON RIEMANNIAN GEOMETRY

In this section, we introduce some essential concepts from differential geometry and information theory, and fix the notation and terminology used throughout this paper. For more details, we refer the reader to, for example, Amari *et al* (1987) or Murray and Rice (1993).

Let $\mathcal{M}$ be a statistical manifold consisting of the probability density functions $p(x|\theta)$ of a random variable $x \in \mathcal{X}$ with respect to a measure μ on $\mathcal{X}$, such that every distribution is uniquely parameterized by an n-dimensional vector $\theta = (\theta^i) = (\theta^1, \ldots, \theta^n)$.[1] Specifically, let

$$\mathcal{M} = \{p(x|\theta) \mid \theta \in \Theta \subset \mathbb{R}^n\},$$

with a one-to-one mapping between θ and $p(x|\theta)$. In addition, under the assumptions that the parameterization of $\mathcal{M}$ is differentiable and C^∞ is a diffeomorphism, the parameterization θ forms a coordinate system of $\mathcal{M}$ (Amari *et al* 1987). The local coordinate system $\theta = (\theta^1, \ldots, \theta^n)$ then induces a basis

$$\frac{\partial}{\partial\theta} = (\partial_1, \ldots, \partial_n)$$

of the tangent spaces (∂_i is shorthand for $\partial/\partial\theta^i$).

The structure of $\mathcal{M}$ is specified by a Riemannian metric, $g = (g_{ij})$, which is defined by a local product on tangent vectors at each point $p \in \mathcal{M}$, denoted by $g\colon T_p\mathcal{M} \times T_p\mathcal{M} \to \mathbb{R}$; that is, symmetric, bilinear, positive definite and C^∞ differentiable in p. By the bilinearity of the inner product of g, for any two tangent vectors $u, v \in T_p\mathcal{M}$,

$$g(u, v) = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i u_j g(\partial_i, \partial_j),$$

[1] We describe only the case for a continuum on the set $\mathcal{X}$; however, if $\mathcal{X}$ were discrete, the given framework would still apply if we switched $\int(\cdot)$ with $\sum(\cdot)$.

where $\{\partial_i\}_{i=1}^n$ form the basis elements of $T_p\mathcal{M}$. For a statistical manifold, the Fisher–Rao information metric is given by

$$I_{ij}(p) = g_{ij}(p) = \mathbb{E}\left[\frac{\partial \log(p)}{\partial\theta^i}\frac{\partial \log(p)}{\partial\theta^j}\right] = \int p\frac{\partial \log(p)}{\partial\theta^i}\frac{\partial \log(p)}{\partial\theta^j}\,\mathrm{d}\theta, \quad (2.1)$$

where $I(p) = I(p(x|\theta))$ can be considered as a Riemannian metric. Note that $\det(I(p(x|\theta)))$ represents the amount of information a sample point conveys with respect to the problem of estimating the parameter $\boldsymbol{\theta}$, and so $I(\cdot)$ can be used to determine the dissimilarities between distributions. This measures the ability of the random variable x to discriminate the values of the parameter θ' from θ for θ' close to θ.

EXAMPLE 2.1 (Statistical manifold of normal distributions) The normal distribution with mean μ and variance σ^2 is given by

$$\mathcal{N}(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}, \quad x \in \mathbb{R},\ \mu \in \mathbb{R},\ \sigma > 0.$$

Set $\Theta = \{(\theta^1, \theta^2) \in \mathbb{R}^2 \mid \theta^2 > 0\}$, $\mathcal{X} = \mathbb{R}$ and $p(x|\theta) = \mathcal{N}(x \mid \sqrt{2}\theta^1, (\theta^2)^2)$. Then, the statistical manifold with respect to $\{p(\cdot|\theta) \mid \theta \in \Theta\}$ has a Riemannian metric $g = 2(\theta^2)^{-2}\sum \mathrm{d}\theta^i\,\mathrm{d}\theta^i$ of constant curvature $-1/2$.

The Riemannian metric encodes how to measure distances, angles, areas and curvature on the manifold by specifying the scalar product between tangent vectors at a particular point. If we consider a curve $\gamma\colon [a,b] \subset \mathbb{R} \to \mathcal{M}$ on the manifold, its length $\ell(\gamma)$ can be defined as

$$\ell(\gamma) = \int_a^b \left\|\frac{\mathrm{d}\gamma}{\mathrm{d}t}\right\| \mathrm{d}t = \int_a^b \sqrt{g_{ij}\dot\gamma^i\dot\gamma^j}\,\mathrm{d}t = \sqrt{\langle\dot\gamma,\dot\gamma\rangle},$$

where $\dot\gamma^i$ is the derivative of $\dot\gamma^i = \partial\gamma/\partial\theta^i$. Then, the distance between two points $p, q \in \mathcal{M}$ is defined by the infimum of the length of all smooth curves between these two points. This is given by

$$d(p,q) = \int_{\gamma\in\Gamma} \ell(\gamma),$$

where Γ is a set of all of the smooth curves between these two points. The local length-minimizing smooth curve $\gamma(t)\colon [0,1] \to \mathcal{M}$ is called geodesic and is characterized by the fact that it is autoparallel, eg, the field of tangent vectors $\dot\gamma(t)$ stays parallel along γ (velocity is constant along the geodesic $\nabla_{\dot\gamma}\dot\gamma = 0$ on γ). In local coordinates, a curve is geodesic if and only if it is the solution of the system of n second-order Euler–Lagrange equations:

$$\frac{\mathrm{d}^2\theta^k}{\mathrm{d}t^2} + \sum_{i,j=1}^n \Gamma_{ij}^k\frac{\partial\theta^i}{\mathrm{d}t}\frac{\partial\theta^j}{\mathrm{d}t} = 0$$

for all $k = 1, \ldots, n$, where Γ_{ij}^k are Christoffel symbols of the second kind.

From the results of existence and uniqueness of solutions of differential equations, for each $p \in \mathcal{M}$ and tangent vector $v \in T_p\mathcal{M}$ there exists an open interval I_v with $0 \in I_v$ and a unique geodesic $\gamma_v \colon I_v \to \mathcal{M}$ such that $\gamma_v(0) = p$ and $\dot{\gamma}_v(0) = v$. Therefore, the exponential mapping $\exp_p \colon T_p\mathcal{M} \to \mathcal{M}$ is defined by $\exp_p(v) = \gamma_v(1) = \gamma_{v_1}(\|v\|)$, with $v_1 = v/\|v\|$. For each $p \in \mathcal{M}$, there exists a neighborhood $\mathcal{V}$ of the origin in $T_p\mathcal{M}$ such that $\exp_p$ is a diffeomorphism from $\mathcal{V}$ onto a neighborhood V of p. The neighborhood $\mathcal{V}$ is star-shaped, ie, for any point belonging to $\mathcal{V}$, the line joining the point to the origin is contained in $\mathcal{V}$. The image of a star-shaped neighborhood of the origin under the exponential map is a neighborhood of p on the manifold (also called a normal neighborhood).

A notion of connection ∇ defines a map between any neighboring tangent spaces. The canonical affine connection on a Riemannian manifold is the Levi–Civita connection, and this is defined directly from the covariant derivative, ie, it is an orthogonal projection of the usual derivative of the vector fields onto tangent space. It parallel transports a tangent vector along a curve while preserving its inner product (it is compatible with the metric, ie, the covariant derivative of the metric is zero). The Levi–Civita coefficients are defined in each local chart by the Christoffel symbols of the second kind Γ_{ij}^k, given by

$$\nabla_{ij}^k = \Gamma_{ij}^k = \tfrac{1}{2} g^{kl} \left(\frac{\partial g_{jl}}{\partial \theta^i} + \frac{\partial g_{il}}{\partial \theta^j} - \frac{\partial g_{ij}}{\partial \theta^l} \right)$$

for all $i, j, k, l = 1, \ldots, n$; the Einstein summation convention is used, and g^{kl} is the metric inverse. When the connection coefficients of ∇ with respect to a coordinate system of $\mathcal{M}$ are all identically 0, then ∇ is said to be flat, or, alternatively, $\mathcal{M}$ is flat with respect to ∇.

3 MODELING PROCESS STEPS AND QUANTIFICATION OF MODEL RISK

There are different types and aspects of model risk that tend to easily overlap, co-occur or co-vary. In this context, we propose four rough model creation steps: (1) data, (2) calibration, (3) model selection and testing, and (4) implementation and usage. These steps may occur in an iterative fashion, although they result in a general linear flow that ends with institutional use (implementation and maintenance) to direct decision making (often encoded into an IT system). Limitations in any of these areas can impair the reliability of the model results.

Data refers to the definition of the reason for modeling, the specification of the modeling scope, human and financial resources, and the specification of data and

other prior knowledge as well as its interpretation and preparation. Data may be obtained from both internal and external sources, and it is further prepared by cleaning and reshaping. Model risk may arise from data deficiencies in terms of both quality and availability, including errors in data definition, insufficient sample sizes, inaccurate proxies, sensitivity to expert judgment or misinterpretation.

Calibration includes selecting the types of variables and the nature of their treatment, tuning free parameters and the links between system components and processes. Estimation uncertainty may occur due to simplifications, approximations, flawed assumptions, inappropriate calibration, errors in statistical estimation or market benchmarks, computational or algorithmic limitations, or the use of unobservable parameters.

Model selection and testing involves choosing the estimation performance criteria and techniques, and identifying the model structure and parameters, which is generally an iterative process with the underlying aim of balancing sensitivity to system variables with complexity of representation. Further, it is related to conditional verification, which includes checking sensitivity to changes in the data and to possible deviations from the initial assumptions. In this step, model risk stems from, for example, inadequate and incorrect modeling assumptions, outdated models due to parameter decalibration, model instability or model misspecification.

Implementation and usage refers to deploying the model into production, which is followed by regular maintenance and monitoring. Sources of model risk in this step include using the model for unintended purposes, a lack of recalibration, IT failures, a lack of communication between modelers and users, and a lack of understanding with regard to model limitations.

The quantification of model risk, from a best-practice perspective, should be quick and reliable, without the need to refit or rebuild models or reference particular structures and methodologies, and with prioritizing analysis (getting immediate assurance on shifts that are immaterial).

In this paper, we aim to show that differential geometry and information theory can offer the base for such an approach. For this purpose, in our setting, a model is represented by a particular probability distribution, $p\colon \mathcal{X} \to \mathbb{R}_+$, that belongs to a manifold of probability measures $\mathcal{M}$, the so-called statistical manifold, available for modeling. The manifold $\mathcal{M}$ can be further equipped with the information-theoretic geometric structure that allows us to quantify variations and dissimilarities between probability distribution functions (models), among other things.

The set of probability measures may be further parameterized in a canonical way by a parameter space Θ, $\mathcal{M} = \{p(x;\theta) \mid \theta \in \Theta\}$. This set forms a smooth Riemannian manifold $\mathcal{M}$. Every distribution is a point in this space, and the collection

of points created by varying the parameters of the model, $p \in \mathcal{M}$, gives rise to a hypersurface (a parametric family of distributions) in which similar distributions are mapped to nearby points. The natural Riemannian metric is shown to be the Fisher–Rao metric (see Rao 1945), which is a unique intrinsic metric on the statistical manifold. It is the only metric that is invariant under reparameterization (see, for example, Amari *et al* 1987).

Let us consider a given model p_0, which can be uniquely parameterized using the vector $\theta_0 = (\theta_0^1, \ldots, \theta_0^n)$ over the sample space $\mathcal{X}$, and which can be described by the probability distribution $p_0 = p(x; \theta_0)$. This probability distribution belongs to a set (family) of distributions $\mathcal{M} = \{p(x; \theta) \mid \theta \in \Theta \subset \mathbb{R}^n\}$ that forms a model manifold. We assume that for each $x \in \mathcal{X}$ the function $\theta \mapsto p(x; \theta)$ is C^∞. Thus, $\mathcal{M}$ forms a differentiable manifold, and we can identify models in the family with points on this manifold. Thus, choosing a particular model is equivalent to fixing a parameter vector $\theta \in \Theta$.

We define the model risk for a given model p_0 on the scale of an open neighborhood around p_0, which contains alternative models that are not too far, in a sense that is quantified by the relevance of missing properties and limitations of the model. The model risk is then measured with respect to all of the models inside this neighborhood as the norm of an appropriate function of the output differences over a weighted Riemannian manifold endowed with the Fisher–Rao metric and the Levi–Civita connection.[2] This analysis consists of five steps:

(1) embedding the model manifold into one that considers missing properties in the given model p_0;[3]

(2) choosing a proper neighborhood around the given model;

(3) choosing an appropriate weight function that assigns relative relevance to the different models inside the neighborhood;

(4) calculating the measure of model risk with respect to all models inside the neighborhood through the corresponding norm; and

(5) interpreting the measure with respect to the specific use of the model risk quantification.

---

[2] The Levi–Civita connection transports tangent vectors defined at one point to another and is compatible with the geometry induced by the Riemannian metric (Amari *et al* 1987). In addition, for this choice of connection, the shortest paths are geodesics.

[3] Or properties not appropriately modeled, for which there is no consensus, or which cannot be adequately calibrated, etc.

Each step addresses and aligns different limitations of the model and the uncertainty in various areas related to the model.[4] In the following sections, we further develop these steps and describe the intuition behind them.

4 NEIGHBORHOOD AROUND THE MODEL

Recall that the given model p_0 belongs to a n-dimensional manifold $\mathcal{M}$, where each dimension represents different aspects of information inherited by p_0. In order to incorporate any missing properties, uncertainty surrounding the data and calibration, additional information about the limitations of the model or wrong underlying assumptions, we may need to add new dimensions to $\mathcal{M}$, and thus to consider a higher-dimensional space into which $\mathcal{M}$ can be embedded.[5]

We define a neighborhood around p_0 with the help of the tangent space $T_{p_0}\mathcal{M}$.[6] Note that $T_{p_0}\mathcal{M}$ is a vector space that describes a first-order approximation, infinitesimal displacements or deformations on the manifold at point p_0. From a practical point of view, not all perturbations are relevant; thus, taking into account the materiality with respect to the intended purpose of the model, its usage, business and market, we consider only a small subset of the tangent bundle.

Let $\mathcal{U}$ be the open set around p_0 of a normal neighborhood V such that

$$\mathcal{U} = \{tv \in V \subset T_{p_0}\mathcal{M} \mid 0 < t \leqslant \alpha(v),\ v \in \mathcal{S}(p_0, 1) \text{ and normal coordinates are defined}\},$$

where $\mathcal{S}(p_0, 1) = \{v \in T_{p_0}\mathcal{M},\ \|v\| = 1\}$ is the unit sphere on $T_{p_0}\mathcal{M}$.

The set $\mathcal{U}$ includes the directions of all relevant perturbations of the model p_0 up to a certain level $\alpha(v)$ in direction v. The level $\alpha(v)$ depends on the tangent vectors, since the degree of our uncertainty on p_0 might not be constant across the canonical parameter space; for instance, we could assume more uncertainty in the tails of the distribution p_0 than in its body, or higher sensitivity to one of the parameters than to the others. The level $\alpha(v)$ can be interpreted as a means of controlling the uncertainty surrounding our choice of model p_0 as well as the data and calibration, and it is appropriately chosen based on the usage of the model.

---

[4] For example, data, calibration, model selection, model performance, model sensitivity and scenario analysis, and – most importantly – model usage.

[5] Consider, for example, a case where the underlying model space represents the family of normal distributions, ie, a two-dimensional manifold $\mathcal{M}_0$; we may want to consider the family of skew normal distributions, ie, a three-dimensional manifold $\mathcal{M}$ for which $\mathcal{M}_0 \subset \mathcal{M}$, in order to also examine skewness.

[6] Note that throughout the paper we do not refer to the neighborhood as a strictly topological neighborhood.

Since $\mathcal{U}$ is a subset of the normal neighborhood around p_0, the exponential map is well defined and we can construct a corresponding set of models close to p_0:

$$U = \exp_{p_0}(\mathcal{U}) = \{p \in \mathcal{M} \mid d(p_0, p) \leqslant \alpha(v)\}. \tag{4.1}$$

From now on, we shall require the boundary $\partial U = \{\exp_{p_0}(\alpha(v)v) \mid v \in \mathcal{S}(p_0, 1)\}$ to be continuous and piecewise regular. Moreover, U shall be a compact star-shaped set with respect to p_0, which is defined as follows.

DEFINITION 4.1 A compact subset U of a Riemannian manifold $\mathcal{M}$ is called star-shaped with respect to $p_0 \in U$ if, for all $p \in U$, $p \neq p_0$, there exists a geodesic segment γ, with $\gamma(0) = p_0$ connecting p_0 and p such that $\gamma(t) \in U$ for all $t \in [0, \alpha(v)]$, where $\alpha(v) > 0$ with $v \in \mathcal{U}$.

One advantage of the exponential map in this setting is that we can avoid the calibration of different alternative models inside U. For each unit vector $v \in \mathcal{U}$, there exists a unique geodesic connecting points on the boundary of U with point p_0. This geodesic is given by $\gamma(t) = \exp_{p_0}(tv)$ for $t \in [0, \alpha(v)]$.

5 WEIGHT FUNCTION DEFINITION

Variations of the chosen model are not equally material and they might all take place with different probabilities. By introducing a nonlinear weight function (kernel), K, over the set U, we can easily assign relative relevance to each alternative model and determine the credibility of the underlying assumptions that would make alternative models partially or relatively preferable to the nominal one, p_0. The particular choice of kernel structure depends on various factors, such as model usage, distance from p_0 and sensitivity to different changes.

In what follows, we define a general weight function K and show that under certain conditions it is well defined and unique. In general, we consider K to be a non-negative and continuous function that depends on the local geometry of $\mathcal{M}$ by incorporating a Riemannian volume associated with the Fisher–Rao information metric, which is given by $\mathrm{d}v(p) = \sqrt{\det(I(\theta))}\,\mathrm{d}\theta$. The volume measure is the unique Borel measure on $\mathcal{M}$ (Federer 2014). With respect to a coordinate system, the information volume of p represents the amount of information a single model possesses with respect to its parameters. For example, a small $\mathrm{d}v(p)$ means the model contains much uncertainty and requires many observations to learn.

As the underlying factors that influence the perturbations of the given model happen with some likelihood, we treat all of the models inside $\mathcal{M}$ as random objects.[7]

[7] For example, the uncertainty surrounding data, calibration or model selection.

As a consequence, we require K to be a probability density with respect to the Riemannian volume, ie, $\int_{\mathcal{M}} K \, \mathrm{d}v(p) = 1$. In addition, we state that the right model does not exist and that the choice of p_0 is to some extent a subjective preference.

DEFINITION 5.1 An admissible weight function K defined on $\mathcal{M}$ satisfies the following properties:

$(K1')$ K is continuous on $\mathcal{M}$;

$(K2')$ $K \geqslant 0$ for all $p \in \mathcal{M}$;

$(K3)$ $\int_{\mathcal{M}} K \, \mathrm{d}v(p) = 1$.

Recall that to compute the n-dimensional volume of the objects in $\mathcal{M}$, one considers a metric tensor on the tangent space $T_p\mathcal{M}$. In particular, the Fisher–Rao information metric I on $\mathcal{M}$ maps each $p \in \mathcal{M}$ to a volume $\mathrm{d}v(p)$, which is symmetric, and an n-form that further defines an n-dimensional volume measure on any measurable subset $U \subset \mathcal{M}$ by $\mathrm{Vol}(U) = \int_U \mathrm{d}v(p)$. A smooth probability density K over $\mathcal{M}$ with respect to the Riemannian measure induces a new absolutely continuous probability measure ζ with respect to Vol:

$$\zeta(U) = \int_U \mathrm{d}\zeta = \int_U K \, \mathrm{d}v(p) \tag{5.1}$$

for all measurable $U \subset \mathcal{M}$ and $\zeta(\mathcal{M}) = 1$. The pair $(\mathcal{M}, \zeta)$ is then called a weighted manifold, or a Riemannian metric–measure space, and it is proved to be a nontrivial generalization of Riemannian manifolds (Morgan 2005).

The weight function K of Definition 5.1 represents a general characterization of a probability density over the Riemannian manifold $\mathcal{M}$. In order to tune K for a proper analysis of model risk, we need to impose additional properties that are connected with the specific uncertainties surrounding the given model.

From a practitioner's point of view, models that do not belong to the chosen neighborhood U are not relevant from the perspective of model risk, and so they do not add any uncertainty. Therefore, we assume the weight function K to be nonnegative only over the neighborhood U, and zero elsewhere. Moreover, the translation of changes in various underlying assumptions, data or calibration into changes in output and further usage of the model are going to vary with respect to the direction of the change. Hence, we require K to be continuous along the geodesic curves γ uniquely determined by $v \in \mathcal{S}(p_0, 1) \subset T_{p_0}\mathcal{M}$ starting at p_0 and ending at the points on ∂U. These additional properties motivate the following modifications of $(K1')$ and $(K2')$ in Definition 5.1:

$(K1)$ K is continuous along all geodesics γ starting at p_0 for all unit vectors on $\mathcal{S}(p_0, 1)$ and ending at the points on ∂U;

($K2$) $K > 0$ for all $p \in U \backslash \{\partial U\}$, $K \geqslant 0$ for all $p \in \partial U$, and $K = 0$ for all $p \in \mathcal{M} \backslash \{U\}$.

A weight function satisfying properties ($K1$)–($K3$) takes into consideration and is adjusted according to the different directions of the changes, ie, it prescribes different sensitivities to different underlying factors.

The construction of a weight function on a given Riemannian manifold may become technically difficult since it requires precise knowledge of the intrinsic geometry and structure of the manifold. In order to overcome this difficulty and determine a weight function K that satisfies all of the required properties, we introduce a continuous mapping from a manifold endowed with a Euclidean geometry to the model manifold endowed with a Riemannian geometry that preserves the local properties. In summary, we construct three mappings: the exponential map $\exp_{p_0}$, the polar transform P and a further coordinate transform Λ_ρ. Euclidean geometry is well understood and intuitive, and thus the construction of a function on this space is relatively easy and intuitive. The steps of the construction and the associated proofs are relegated to the online appendix, as we believe this will improve the readability and flow of the paper.

6 MEASURE OF MODEL RISK

In this section, we introduce a definition of the quantification of model risk, relate it to the concepts we have already introduced and study some realistic applications.

Recall that we have so far focused on a weighted Riemannian manifold $(\mathcal{M}, I, \zeta)$ with I the Fisher–Rao metric and ζ as in (5.1). The model was assumed to be identified with a distribution $p \in \mathcal{M}$. More likely, a practitioner would define the model as some mapping $f : \mathcal{M} \to \mathbb{R}$ with $p \mapsto f(p)$, ie, the model outputs some quantity.[8]

We introduce the normed space $(\mathcal{F}, \|\cdot\|)$ such that $f \in \mathcal{F}$. Though not strictly necessary at this stage, we shall assume completeness so that $(\mathcal{F}, \|\cdot\|)$ is a Banach space.

DEFINITION 6.1 Let $(\mathcal{F}, \|\cdot\|)$ be a Banach space of measurable functions with respect to ζ. The model risk Z of $f \in \mathcal{F}$ and p_0 is given by

$$Z(f, p_0) = \|f - f(p_0)\|. \tag{6.1}$$

[8] This is not always the case, but we can proceed along these lines depending on the usage given to the quantification itself. For example, an inter(extra)polation methodology on a volatility surface is a model whose output is another volatility surface, not a number. If we want to quantify the model risk of that particular approach for Bermudan derivatives, we might consider its impact on their pricing.

Note that the measure represents the standard distance. All outcomes are constrained by the assumptions used in the model itself, so the model risk is related to changes in the output while relaxing and changing them.

The quantification of model risk can be thought of as a model with a purpose, such as the allocation of provisions, capital calculation or the comparison of modeling approaches. The possibilities are endless, so we might have started with some $T\colon \mathcal{F} \to \mathcal{F}$ and set $Z(f, p_0) = \|T \circ f\|$;[9] however, we think (6.1) is general enough for our present purposes.

In what follows, we consider four examples of Definition 6.1. Their suitability very much depends on, among other factors, the purpose of the quantification, as we shall see later.

(1) $Z^1(f, p_0)$ for $f \in L^1(\mathcal{M})$ represents the total relative change in the outputs across all relevant models:

$$Z^1(f, p_0) = \|f - f(p_0)\|_1 = \int_{\mathcal{M}} |f - f(p_0)| \, \mathrm{d}\zeta.$$

(2) $Z^2(f, p_0)$ for $f \in L^2(\mathcal{M})$ puts more importance on big changes in the outputs (big gets bigger and small gets smaller). It would allow us to keep some consistency across calibration processes such as the maximum likelihood or least squares algorithms:

$$Z^2(f, p_0) = \|f - f(p_0)\|_2 = \left(\int_{\mathcal{M}} (f - f(p_0))^2 \, \mathrm{d}\zeta \right)^{1/2}.$$

(3) $Z^\infty(f, p_0)$ for $f \in L^\infty(\mathcal{M})$ finds the relative worst-case error with respect to p_0:

$$Z^\infty(f, p_0) = \|f - f(p_0)\|_\infty = \operatorname{ess\,sup}_{\mathcal{M}} |f - f(p_0)|.$$

Further, it can point to the sources of the largest deviances: using $\exp_{p_0}^{-1}$, we can detect the corresponding direction and size of the change in the underlying assumptions.

[9] For example, another possibility is to use

$$\left\| \frac{f}{f(p_0)} \right\| \quad \text{or} \quad \left\| \frac{f - f(p_0)}{f(p_0)} \right\|.$$

These functional forms would allow us to obtain a dimensionless number, which might be a desirable property.

(4) $Z^{s,p}(f, p_0)$ for $f \in W^{s,p}(\mathcal{M})$ is a Sobolev norm that can be of interest in cases where not only f but also its rate of change is relevant:[10]

$$Z^{s,p}(f, p_0) = \|f - f(p_0)\|_{s,p} = \left(\sum_{|k| \leqslant s} \int_{\mathcal{M}} |\partial^k (f - f(p_0))|^p \, \mathrm{d}\zeta \right)^{1/p}.$$

A sound methodology for model risk quantification should at least consider the data used for building the model, the model's foundation, the IT infrastructure, overall performance, model sensitivity, scenario analysis and, most importantly, usage. Within our framework, we address and measure the uncertainty associated with the aforementioned areas and the information contained in the models. Choosing the embedding and proper neighborhood of the given model takes into account the knowledge and uncertainty of the underlying assumptions as well as the data and the model's foundation. The weight function that assigns relative relevance to the different models inside the neighborhood considers the model sensitivity, scenario analysis, the importance of the outcomes with connection to decision making, the business and the intended purpose; it also addresses the uncertainty surrounding the model's foundation. Besides, every particular choice of the norm provides different information about the model. Finally, and most importantly, the model risk measure considers the usage of the model represented by the mapping f.[11]

7 APPLICATION TO CAPITAL CALCULATION

In this section, we contextualize the proposed framework by applying it to a credit risk model used by a commercial bank. More specifically, we employ the proposed methodology for the quantification of model risk to a probability of default (PD) model of high default portfolios used for capital calculation.[12] We analyze different scenarios and risk parameter assumptions in order to assess how these scenarios affect the economic capital based on methodologies commonly applied by internal ratings-based (IRB) institutions.

In general, the purpose of a credit risk model is to estimate the PD of future credit losses on a bank portfolio. For a given time horizon, the model generates a distribution – a probability density function – of future losses that can be used to calculate the losses associated with any given percentile of the distribution. In practice, banks concentrate on two such loss components: expected loss (EL) and unexpected loss (UL). EL is the mean of the loss distribution and represents the amount a bank expects

[10] An example is a derivatives model used for not only pricing but also hedging.

[11] Or, equivalently, by any possible transformation $T : \mathcal{F} \to \mathcal{F}$.

[12] Note that the purpose of this analysis is only to illustrate the proposed framework, so it does not cover all potential risks inherent in PD calculation.

to lose on average on its credit portfolio over a given time horizon. In contrast, UL refers to the portfolio risk that is computed as the losses associated with some high percentile of the loss distribution (eg, the 99.99th percentile); it thus covers all but the most extreme events. For more details about credit risk, see, for example, Bluhm *et al* (2016).

Capital models are usually based on three risk parameters: PD, exposure at default (EAD) and loss given default (LGD). Under the Basel II IRB framework, the PD per rating grade is the average percentage of obligors that will default over a one-year period. The EAD gives an estimate of the amount that would be outstanding should the borrower default, and the LGD represents the proportion of the exposure that will not be recovered after default. These parameters are aggregated from obligor level (risk bucket) to portfolio level, with the correlation set by the regulator or the financial entity.

Let N be the number of borrowers in a given loan portfolio. Assuming a uniform value of LGD, the aggregated expected loss amount, L, can be calculated as the sum of individual L in the portfolio, ie,

$$L = \sum_{i=1}^{N} \mathrm{EAD}_i \mathrm{LGD}_i \mathrm{PD}_i. \tag{7.1}$$

Based on the loss distribution, the capital requirement for a bank under the IRB approach at confidence level α can simply be calculated as the difference between EL and the percentile for the level being considered:

$$C_\alpha(L) = q_L(\alpha) - \mathrm{EL}, \tag{7.2}$$

where $q_L(\alpha)$ is the α-quantile of L defined by $\mathbb{P}(L \leqslant q_L(\alpha)) = \alpha$. In practice, the portfolio is categorized into homogeneous risk buckets, $j = 1, \ldots, M$, and the capital calculation is done on the risk-bucket level. The same default probabilities PD_j and LGD_j are then assigned to all borrowers in each of these buckets.

When modeling credit risk losses, several sources of model risk may arise. Examples are a scarcity of default events, a lack of data driving calibration and backtesting, correlations between failures, wrong-way exposure (a growing utilization of credit lines in case there is an increase in PD), or independence of PD and LGD (in periods of increasing default rate, PD and LGD will not increase).[13]

[13] Empirical evidence has led to developments in credit risk models that take into account the dependence between PD and LGD. It was shown that LGD exhibits not only volatility but also a positive correlation with default rates or PD (Han 2017). One of the reasons for this relation is the business cycle. However, despite strong evidence of PD–LGD dependence, some commercial credit risk packages still rely on constant or independent assumptions regarding LGD.

For the purpose of illustration, we will only focus on the quantification of model risk arising from PD estimation at the capital level, so we assume both LGD and EAD to be known and independent of PD.

7.1 PD model (p_0)

The PD is the likelihood that an obligor will default within one year given all of the currently available information. We consider a particular internal model for the long-run PD estimation that serves our regulatory purposes. The PD model is built on internal behavioral data and bureau information. Each customer account is scored and the portfolio is categorized, with respect to the scoring, into M homogeneous risk buckets. The number of defaults in each bucket is assumed to follow a binomial random variable, in which the defaults are independent across customers and over time, and occur with common probability. The underlying point-in-time (PIT) PD model is calibrated to the latest observed PD, ie, the observed default frequency (ODF), that is, just the approximation of the maximum likelihood estimator of the parameter of the binomial distribution. PIT PD is then adjusted by the central tendency (CT) to generate the long-run PD based on a combination of historical misalignment of the underlying model and expert judgment. The through-the-cycle (TTC) PD, ie, the pooled PD for each risk bucket, is given by the formula below:

$$\mathrm{PD}_i = \mathrm{ODF}_i \frac{\mathrm{CT}}{\mathrm{ODF}}, \tag{7.3}$$

where

- ODF_i is the observed default frequency obtained for each risk bucket i over the most recent quarter of the calibration sample;
- CT refers to the central tendency used for the TTC adjustment, which is based on external macroeconomic data series in order to extend the internal ODF series;[14] and
- ODF is the average observed default frequency over the given segment portfolio.

Thus, the long-run ODFs are calculated for each risk bucket and are adjusted to the average PD observed for each portfolio over a complete economic cycle. The PD therefore gives the likelihood that obligors with a particular rating grade at the start of a given time period will default within this period. The distribution of defaults under our simplifying assumptions would equal the loss distribution of a portfolio

[14] In case there is not enough data history in order to cover the whole economic cycle.

TABLE 1 Normalized PD and frequency of accounts in the portfolio across risk buckets.

	Risk buckets									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
PD	0.276	0.198	0.181	0.090	0.085	0.065	0.037	0.034	0.023	0.012
N (frequency)	0.007	0.012	0.021	0.055	0.066	0.088	0.176	0.280	0.241	0.056

for which all the borrowers had an EAD of 1, an LGD of 0.45 and a maturity of one year.

The empirical analysis is based on a hypothetical portfolio that was elaborated in order to represent the behavior of one particular segment consisting of $N = 9860$ clients. All customers are categorized by risk-scoring into buckets that are heterogeneous at a defined confidence level and represent the grade of credit quality. We assume a portfolio consisting of ten buckets ($M = 10$) with a monotonous PD related to the model scores. Higher scores imply lower ODFs.

All customers within the same risk bucket are assigned the same pooled PD, which can be thought of as the average of individual PDs. This means that the pooled PD assigned to a risk bucket is a measure of the average value of the PDs of customers in that bucket.

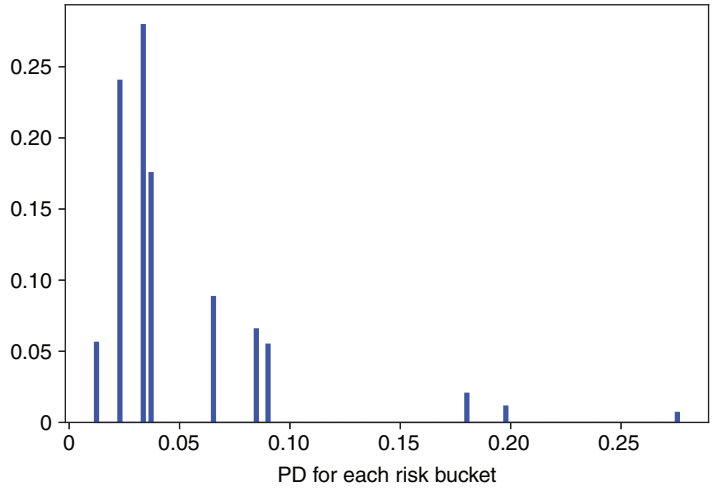
The normalized PDs vary between 0.01143 and 0.27638 and decrease from lower to higher risk ratings, although this decrease is moderate (Table 1). The distribution is highly skewed to the right (as expected), with mean and variance equal to 0.4177 and 0.08768, respectively. The frequency of accounts across buckets ranges from 0.70% to 28.89% (Table 1). The low PD buckets account for the majority of the portfolio, where approximately 75% of the total accounts have PDs lower than 0.04.

There are many factors that influence the direction and extent of the relations between the developments in defaults and the risk factors influencing solvency. These are the macroindicators, market factors and idiosyncratic factors. For example, foreign clients tend to be more risky than domestic ones; the risk of default is expected to grow with increasing interest rates, unemployment rates or percentage deviation between the exchange rate level at which the individual loan was granted and the actual exchange rates.

7.2 Quantification of model risk

We can better understand the impact of the uncertainty surrounding PD on the capital calculation by obtaining a mathematical representation of the underlying statistical distribution over the portfolio (see Figure 1).

FIGURE 1 Distribution of the PD within the portfolio.



A change in the model assumptions or segmentation would represent a shift in the parameters defining the statistical distribution.[15] Based on the goodness-of-fit, we propose predicting the best future values of our given portfolio by using a two-parameter inverse Gaussian distribution (ING), given by

$$p(x,\lambda,\mu) = \left(\frac{\lambda}{2\pi x^3}\right)^{1/2} \exp\left(-\frac{\lambda}{2\mu^2 x}(x-\mu)^2\right)$$

with mean μ and variance $\sigma^2 = \mu^3/\lambda$.[16] This means that we consider our model space to be a manifold of the ING distribution family, $\mathcal{M}$, parameterized by (λ, μ).

The initial model, $p_0 \in \mathcal{M}$, was arrived at via maximum likelihood estimation (MLE), with parameters $\lambda_0 = 0.5418$ and $\mu_0 = 4.8075$, ie, $p_0 = \text{ING}(\lambda_0, \mu_0)$. With this representation, we take into consideration how specific risk buckets perform within the default portfolio or, alternatively, the weight of default of each grade within the portfolio.

The PD model allows us to estimate the PD for a particular segment across different risk buckets within the limits of the defined parameters. The model represents

[15] Note that we can work equally well with empirical distributions (histograms) and use the Fisher–Rao distance between them (Navab *et al* 2010). However, to properly illustrate the aforementioned framework, we approximate the portfolio by a parametric probability distribution.

[16] The inverse Gaussian density function represents a wide class of distributions, ranging from highly skewed distributions to symmetrical ones as λ/μ varies from 0 to ∞. For more details on the inverse Gaussian distributions, see, for example, Tweedie (1957).

a probability distribution of several correlated and mutually interacting events. Note that the parameters of the distribution are nonlinear functions of the individual PD and N for each risk bucket, ie, $\mu(\mathrm{PD}, N)$ and $\lambda(\mathrm{PD}, N)$. For the quantification of model risk, it is important to understand both how the model integrates the information provided by the sample and how sensitive it is to changes in the sample characteristics (see Figure 2).

7.3 Identification of potential sources of model risk

To properly assess model risk, we first need to identify and describe the potential sources of model risk, covering model assumptions, weaknesses and arbitrariness in the development process. For a given PD model, the main risk may arise from the segmentation and risk-scoring applied, bias toward historical experience, unexpected moves in the exchange rate, relevant changes in macrovariables, missing values, incorrect structure and methodology, inappropriate factors used for the TTC adjustment, or the subjective selection of some parameter values.

The impact of these risk sources should be expressed quantitatively and then combined based on whether the separate model risks are independent, or whether some amplify, counteract or absorb each other. In what follows, and for the sake of illustration, we concentrate on the risk inherent in the applied segmentation arising from the PD estimation. The given model p_0 is only one possible description of the data, incorporating expert knowledge and assumptions directed by expected use as well as implementation constraints. Any uncertainty and errors in the estimation of PD influence the required capital.

7.4 Choosing a suitable Riemannian metric

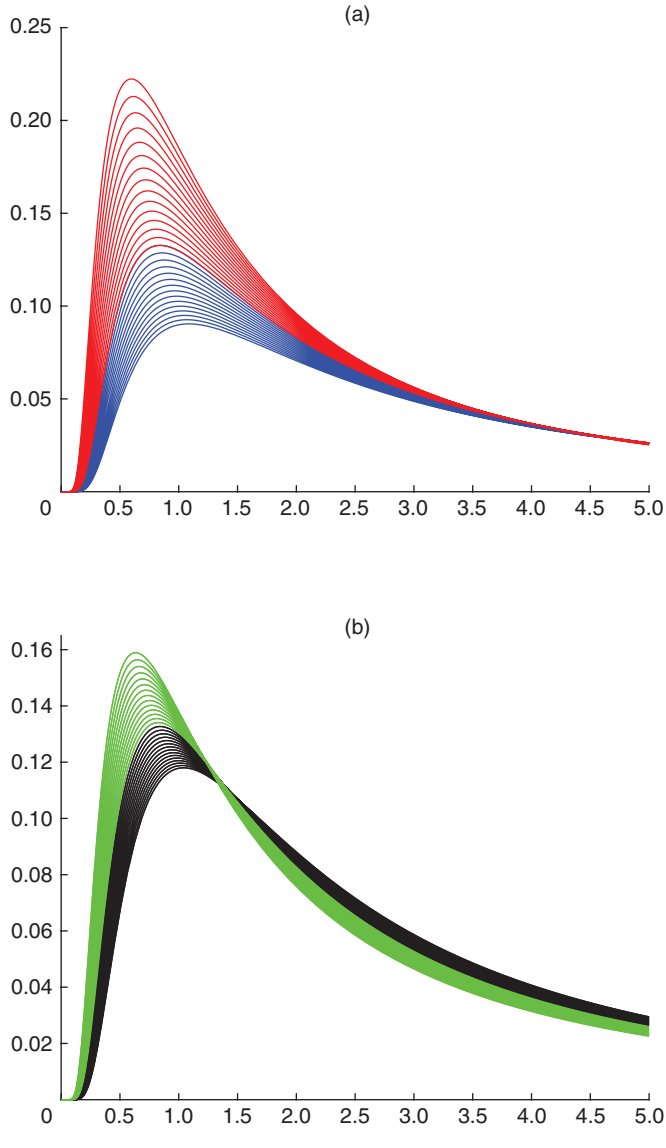
We develop our example on the assumption that the model space should resemble the metric properties of a Riemannian manifold of negative curvature equipped with the Fisher–Rao information metric. Since there are many possible choices of metric on a given differentiable manifold, it is important to consider the additional motivating properties of the Fisher–Rao information metric.

(1) The Fisher–Rao information metric is based on the notion of maximum entropy, which might be geometrically interpreted by the possible trajectory in a statistical manifold that describes its evolution. We can examine the degree of uncertainty by measuring the evolution of entropy in the model space.

(2) The Fisher–Rao information may be thought of as the amount of information a sample supplies with respect to the problem of estimating the parameters.[17]

[17] Since the MLE is asymptotically unbiased, the inverse Fisher information represents the asymptotic fluctuations of MLE around the true value.

FIGURE 2 Probability density functions of inverse Gaussian distributions.



(a) Densities for different μ with $\lambda = 0.5418$. (b) Densities for different λ for $\mu = 4.8075$. The densities are unimodal with a mode between 0 and μ. As μ/λ increases, the distribution becomes more skewed to the right and the mode decreases relative to the mean.

(3) The Fisher–Rao information metric determines a constant negative Gaussian curvature for the ING distribution (see Chen (2014) for details). The metric is given by

$$g(\lambda, \mu) = \begin{bmatrix} \dfrac{1}{2\lambda^2} & 0 \\ 0 & \dfrac{\lambda}{\mu^3} \end{bmatrix}.$$

Negative curvature guarantees that MLE estimates are well defined and unique (Said *et al* 2017).

In this metric, the model space has nonzero curvature and reflects model specification sensitivity accurately. For further discussion on the selection of a proper Riemannian metric, we refer the reader to Krajčovičová and Pérez-Velasco (2018).

7.5 Neighborhood selection

Following Section 4, defining a proper neighborhood around the given model is a trade-off between plausibility and severity in order to ensure that no harmful but plausible risks are missed and irrelevant risk factors are not included. As has already been mentioned, we focus on the sensitivity of the chosen segmentation and the inherent uncertainty. Namely, we analyze the model risk arising from uncertainty about the default counts. We ask: how different could the model be if the default counts were undercounted or overcounted under the assumption of a fixed number of risk buckets?

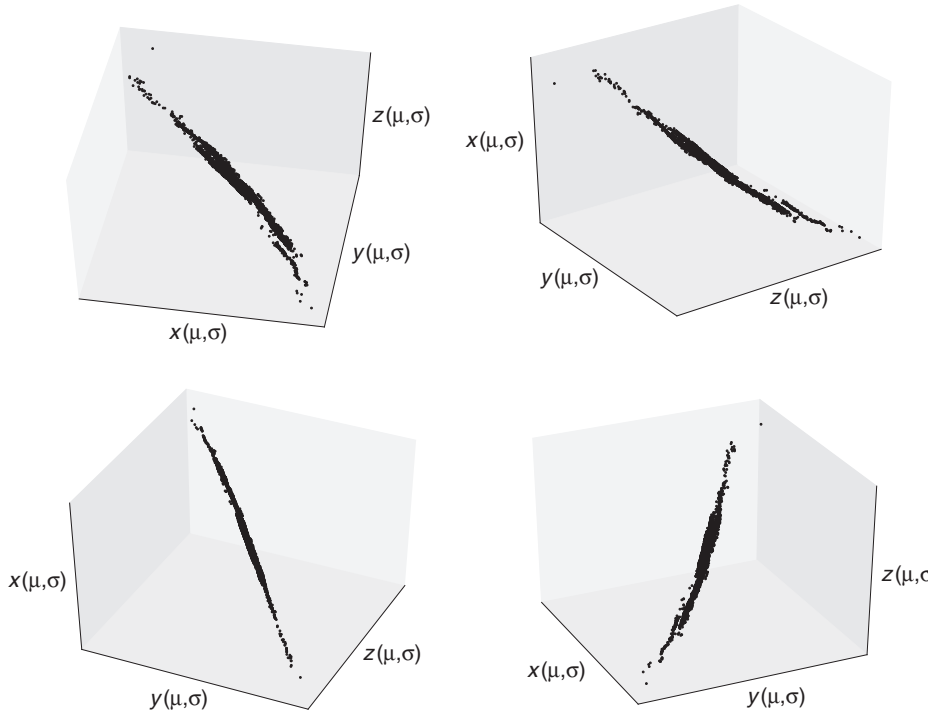
The default rates are calculated at the risk-bucket level, so the misestimation in each bucket is going to contribute to the overall uncertainty. Therefore, we consider all possible deviations within $[-0.3, 0.3]$ standard deviation changes in each of the ODF_i, $i = 1, \ldots, 10$.[18] We take into account all possible shifts in each bucket separately as well as all of the relevant combinations. The limit on the deviation was set to preserve the heterogeneity among risk buckets in terms of PD and to restrict the variability of nearby risk buckets, which could be influenced by the presence of outliers, noise or variation in the density of the points on the manifold.[19]

For the given manifold $\mathcal{M}$ and point $p_0 \in \mathcal{M}$, in order to build the neighborhood U we start by fitting probability distributions to all of the combinations of the maximal ± 0.3 standard deviation changes in the corresponding ODFs (ie, 2^{10} distributions), first to 1 000 000 and then to 10 000 000 random combinations within the interval $(-0.3, 0.3)$. The scale parameter varies in the interval $[\lambda_1, \lambda_2] =$

[18] Recall that ODFs refer to maximum likelihood estimators of the binomial distributions for each risk bucket.

[19] The given portfolio (sample) is going to constrain the specificity and sensitivity possible.

FIGURE 3 Three-dimensional, multidimensional scaling embedding of U.



[0.517300, 0.565403], and the mean is in the interval $[\mu_1, \mu_2] = [4.58638, 4.95968]$. These fitted distributions form our chosen neighborhood.[20] For illustration, Figure 3 represents the three-dimensional embedding of U, the set of alternative models around p_0 determined by the estimated ING distributions.

Within this chosen neighborhood, we can analyze and quantify the possible impact of the individual shifts and all of their possible combinations on the model output that might be capable of causing material model movements. Note the conservatism in the chosen neighborhood.

REMARK 7.1 Different adjusted data may result in the same distributions. Considering only movements in the distribution guarantees that the quantification of

[20] The key idea is to leave the data to determine the neighborhood instead of imposing one. With an increased number of sampled distributions, we can see convergence in the weight function (see Section 7.6) as well as in the overall calculated model risk (see Section 7.7).

model risk will not depend on double-count risk factors that are highly correlated with factors already included in the analysis.

7.6 Choosing an appropriate weight function

In what follows, we show how to estimate the weight function entirely from the data, without imposing any additional constraints or particular structures. Roughly speaking, once we assume that the alternative models are random objects valued on the Riemannian manifold, the weight function estimation consists of having the alternative models within U "contribute" to the estimate at a given point according to their distances from p_0.

More precisely, let $(\mathcal{M}, g)$ be a Riemannian manifold, and let us consider $p_1, \ldots, p_n$ independent and identically distributed random points on $\mathcal{M}$ with density function $K(p)$. The estimate of $K(p)$ is then a positive function of the geodesic distance in U, which is then normalized by the volume density of (U, g) to account for curvature. These estimators are an average of the weights depending on the distance between p_i and p_0. Formally, using alternative models $p_1, \ldots, p_n$, we propose defining the weight function by a map $K_n \colon p \in U \to K_n(p) \in \mathbb{R}$, given by

$$K_n(p) = \frac{1}{\eta_{p_0}(p)} w_i \quad \text{for all } i = 1, \ldots, n, \tag{7.4}$$

where $\eta_{p_0}(p)$ denotes the volume density function and w_i are assigned weights such that $\sum_{i=1}^{n} w_i = 1$.

The weights are determined as follows. First, we calculate the geodesic distance of all of the estimated distributions inside U from p_0. Next, we determine the number of levels m with respect to the maximal distance from p_0, ie,

$$L(\lambda_i) = \{p \in U \mid \lambda_{i-1} < d(p_0, p) \leqslant \lambda_i\} \tag{7.5}$$

for a sequence $\lambda_1, \ldots, \lambda_m$ with $\lambda_m = \max d(p_0, p)$ and $\lambda_i - \lambda_{i-1} = \lambda_m / m$. We set $m = 5000$ in the case of 1 001 024 distributions inside U and $m = 50\,000$ for 10 001 024 alternative models inside U. Next, we examine the number of distributions within all of these m level sets and, based on the concentration within these levels, we calculate normalized weights, ie, w_i. The normalized average frequencies of alternative models within U for 1 001 024 and 10 001 024 with respect to the distance from p_0 are illustrated in Figures 4 and 5, respectively. The resulting normalized weights are then multiplied by the associated values of the volume density at $p \in U$.

REMARK 7.2 The initial model and data single out a particular choice of weight function through the concentration and variations of the fitted probability distributions on the manifold.

FIGURE 4 The weight function given by the average concentration of the alternative models with respect to their geodesic distance from p_0 for 1 000 000 + 1024 alternative models with 5000 level sets.

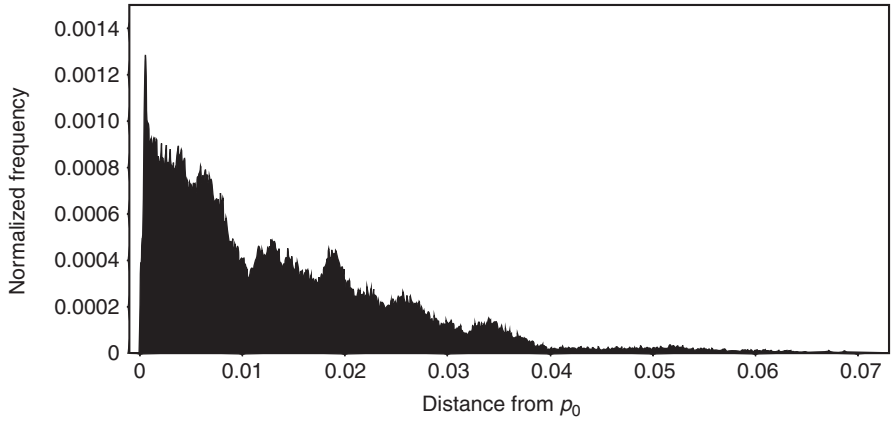
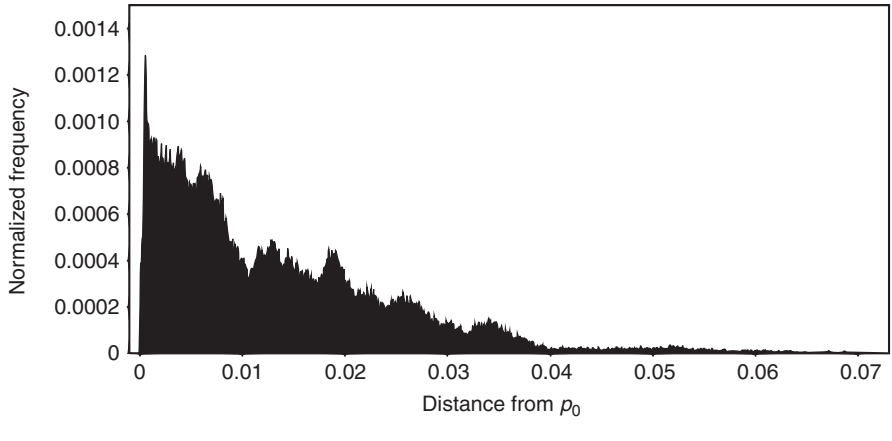


FIGURE 5 The weight function given by the average concentration of the alternative models with respect to their geodesic distance from p_0 for 10 001 024 alternative models with 50 000 level sets.



7.7 Measure of model risk

The last step in our proposed framework is the choice of a proper norm and the evaluation of the identified model risk described in Section 6. We suggest using the $L^2(\mathcal{M})$

norm that in our particular example guarantees consistency with the maximum likelihood used for estimation and amplifies big changes in the outputs, ie, capital level. With this choice, we obtain:[21]

$$Z(C(p), p_0) = \|C(p) - C(p_0)\|_2 = 2.35 \times 10^{-4}.$$

This means the model risk from segmentation represents the 0.0235% variation of capital inside U.[22] The model risk is obtained in terms of capital, thus allowing for easy interpretation.

8 CONCLUSIONS AND FURTHER RESEARCH

In this paper, we introduce a general framework for the quantification of model risk using differential geometry and information theory. We also provide a sound mathematical definition of model risk using weighted Riemannian manifolds, which should be applicable to most modeling techniques using statistics as a starting point.

Our proposed mathematical definition is to some extent comprehensive in two complementary ways. First, it is capable of coping with relevant aspects of model risk management, such as model usage, performance, mathematical foundations, model calibration or data. Second, it has the potential to assess many of the mathematical approaches currently used in financial institutions: credit risk, market risk, derivatives pricing and hedging, operational risk or XVA (valuation adjustments).

It is worth noting that the approaches in the literature, to the very best of our knowledge, are specific in these two ways: they consider very particular mathematical techniques and are usually very focused on selected aspects of model risk management.

There are many directions for further research, all of which we find to be of both theoretical and practical interest. We shall finish by naming just a few of them.

Banach spaces are very well known and have been thoroughly studied in the realms of, for example, functional analysis. Weighted Riemannian manifolds are nontrivial extensions of Riemannian manifolds, one of the building blocks of differential geometry. The study of Banach spaces over weighted Riemannian manifolds will broaden our understanding of the properties of these spaces as well as their application to the quantification of model risk.

Our framework can include data uncertainties by studying perturbations and metrics defined using the sample, which are then transmitted to the weighted Riemannian manifold through the calibration process.

[21] Model risk calculated with 1 001 024 distributions with a weight function based on 5000 level sets is equal to $2.33499704056 \times 10^{-4}$, while model risk based on 10 001 024 distributions inside U with 50 000 distance levels is equal to $2.35152164064 \times 10^{-4}$.

[22] Note that this refers only to one source of model risk inherent in the model for capital calculation.

The general methodology can be tailored and made more efficient for specific risks and methodologies. For example, one may interpret the local volatility model for derivatives pricing as an implicit definition of a certain family of distributions, extending the Black–Scholes stochastic differential equation (which would be a means of defining the lognormal family). For an example on financial derivatives, we refer the reader to Krajčovičová and Pérez-Velasco (2018).

Related to the previous paragraph, and despite the fact that there is already literature on the topic, the calculation of the Fisher–Rao metric itself deserves further numerical research so that more efficient algorithms can be derived.

DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to José Carlos Colas Fuentes, Olivia Peraita Ezcurra and Inmaculada González Peréz from the credit risk models validation department at Banco Santander for providing insight and feedback that greatly enhanced the examples in Section 7.

This work has been partially funded by EU H2020-MSCA-ITN-2014 (WAKEUP-CALL Grant Agreement 643045), Spanish MINECO (Grant MTM2016-76497-R) and Xunta de Galicia (Grant CN2014/044).

REFERENCES

Amari, S. I., Barndorff-Nielsen, O. E., Kass, R. E., Lauritzen, S. L., and Rao, C. R. (1987). Differential geometry in statistical inference. In *Lecture Notes – Monograph Series*, Volume 10. Institute of Mathematical Statistics (https://doi.org/10.1214/lnms/1215467059).

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics* **12**, 171–178.

Bluhm, C., Overbeck, L., and Wagner, C. (2016). *Introduction to Credit Risk Modeling*. Chapman & Hall/CRC Press.

Boucher, C. M., Daníelsson, J., Kouontchou, P. S., and Maillet, B. B. (2014). Risk models-at-risk. *Journal of Banking & Finance* **44**(1), 72–92 (https://doi.org/10.1016/j.jbankfin.2014.03.019).

Branger, N., and Schlag, C. (2004). Model risk: a conceptual framework for risk measurement and hedging. EFMA 2004 Basel Meetings Paper, European Financial Management Association (https://doi.org/10.2139/ssrn.493482).

Chavel, I. (2006). *Riemannian Geometry: A Modern Introduction*. Cambridge Studies in Advances Mathematics, Volume 98. Cambridge University Press (https://doi.org/10.1017/CBO9780511616822).

Chen, W. W. S. (2014). Finding Gaussian curvature of lifespan distribution. *Applied Mathematics* **5**, 3392–3400 (https://doi.org/10.4236/am.2014.521316).

Christodoulakis, G., and Satchell, S. (2008). The validity of credit risk model validation methods. In *The Analytics of Risk Model Validation*, Christodoulakis, G., and Satchell, S. (eds), pp. 27–44. Academic Press (https://doi.org/10.1016/B978-075068158-2.50006-8).

Cont, R. (2006). Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical Finance* **16**(3), 519–547 (https://doi.org/10.1111/j.1467-9965.2006.00281.x).

Federer, H. (2014). *Geometric Measure Theory*. Springer.

Gibson, R. (2000). *Model Risk: Concepts, Calibration and Pricing*. Risk Books, London.

Han, C. (2017). Modelling severity risk under PD–LGD correlation. *European Journal of Finance* **23**(15), 1572–1588 (https://doi.org/10.1080/1351847X.2016.1212385).

Hull, J., and Suo, W. (2002). A methodology for assessing model risk and its application to the implied volatility function model. *Journal of Financial and Quantitative Analysis* **37**, 297–318 (https://doi.org/10.2307/3595007).

Joshi, S., Srivastava, A., and Jermyn, I. H. (2007). Riemannian analysis of probability density functions with applications in vision. In *2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR2017) Proceedings*, pp. 1664–1671. IEEE, Piscataway, NJ.

Krajčovičová, Z., and Pérez-Velasco, P. P. (2018). Model risk and differential geometry applied to sensitivity analysis. Working Paper, Social Science Research Network.

Lang, S. (2012). *Fundamentals of Differential Geometry*. Graduate Texts in Mathematics, Volume 191. Springer.

Morgan, F. (2005). Manifolds with density. *Notices of the American Mathematical Society* **52**, 853–858.

Morini, M. (2011). *Understanding and Managing Model Risk: A Practical Guide for Quants, Traders and Validators*. Wiley (https://doi.org/10.1002/9781118467312).

Murray, M. K., and Rice, J. W. (1993). *Differential Geometry and Statistics*. Monographs on Statistics & Applied Probability, Volume 48. CRC Press (https://doi.org/10.1007/978-1-4899-3306-5).

Navab, N., Pluim, J. P. W., and Viergever, M. A. (2010). *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2010)*. Springer.

Rao, C. R. (1945). Information and accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society* **37**(3), 81–91.

Said, S., Bombrun, L., Berthoumieu, Y., and Manton, J. H. (2017). Riemannian Gaussian distributions on the space of symmetric positive definite matrices. *IEEE Transactions on Information Theory* **63**(4), 2153–2170 (https://doi.org/10.1109/TIT.2017.2653803).

Saltelli, A., Tarantola, S., Campolongo, F., and Ratto, M. (2013). *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. Wiley.

Tweedie, M. C. K. (1957). Statistical properties of inverse Gaussian distributions I. *Annals of Mathematical Statistics* **28**, 362–377 (https://doi.org/10.1214/aoms/1177706964).

US Federal Reserve (2011). Supervisory guidance on model risk management. Supervisory Letter SR 11-7, Board of Governors of the Federal Reserve System, Office of the Controller of the Currency.

Willmore, T. (1993). *Riemannian Geometry*. Oxford University Press.