A Sound Modelling and Backtesting Framework for Forecasting Initial Margin Requirements *†

Fabrizio Anfuso[‡],
Daniel Aziz[§] Paul Giltinan[¶]
Klearchos Loukopoulos[‖]

January 11, 2017

Abstract

The introduction by regulators of mandatory margining for bilateral OTCs is going to have a major impact on the derivatives market, particularly in light of the additional funding costs and liquidity requirements that large financial institutions will face.

The authors propose in the following a simple and consistent framework, equally applicable to non-cleared and cleared portfolios, to develop and backtest forecasting models for initial margin.

Key Words: Initial Margin, BCBS-IOSCO, SIMM, CCP, OTC, Clearing, Counterparty Credit Risk, MVA, XVA, Liquidity, Funding costs

1 Introduction

Since the publication of the new BCBS-IOSCO guidance on mandatory margining for non-cleared OTCs [1], there has been a growing interest in the industry for the development of Dynamic Initial Margin (DIM) models, see e.g. [2], [3], where by DIM model we are referring to any model that can be used to forecast future portfolio Initial Margin Requirements (IMR).

*The views expressed are those of the authors only, no other representation should be attributed

†The authors gratefully acknowledge: (i) Quaternion Risk Management Ltd. for providing us the QRE software package, used for generating some of the results shown in the present paper. (ii) Andrew Hudson for the careful review of the manuscript and the highly valuable feedback.

‡Credit Suisse Securities Limited; dr.fabrizio.anfuso@gmail.com.

§Credit Suisse Securities Limited; dnaziz@gmail.com.

¶Quaternion Risk Management Ltd; paulgiltinan@outlook.com.

‖Credit Suisse Securities Limited; kloukopoulos@googlemail.com.

The business case is at least two-fold: (i) The BCBS-IOSCO IMR (B-IMR) are supposed to cover for potential future exposure at a high level of confidence (99%) and will substantially affect funding costs, XVA and capital. (ii) The B-IMR set a clear incentive towards clearing; extensive margining, in the form of Variation and Initial Margins (VM, IM), is the main element of the CCP risk management model as well.

Therefore, both for bilateral and cleared derivatives, current and future IMR affects significantly the profitability and the risk profile of a given trade.

In the present work, we consider B-IMR as a case study and we show how to include a suitably parsimonious DIM model in the exposure calculation. We propose an end-to-end framework defining also a methodology to backtest the model performance.

The paper is organised as follows: in Sec. 2, the DIM model for forecasting of future IMR is presented; in Sec. 3, methodologies for two distinct levels of backtesting analysis are discussed; finally, in Sec. 4, we draw conclusions.

2 How to construct a DIM model?

A DIM model can be used for various purposes. In the computation of Counterparty Credit Risk (CCR) capital exposure or CVA, the DIM model should forecast on a path by path basis the amounts of posted and received IMs at any revaluation point. For this specific application, the key ability of the model is to associate realistic IMR to a simulated market scenario based on a mapping that makes use of a set of characteristics of the path.

The DIM model is *a priori* agnostic to the underlying Risk Factors Evolution (RFE) models used for the generation of the exposure paths (as we will see, dependencies may arise if e.g. the DIM is computed based on the same paths generated for the exposure).

The story becomes different if the goal is to forecast the IMR Distribution (IMRD) at future horizons, either in the real world P or in the market implied Q measures. In this context, the key feature of the model is to associate the right probability weight to a given IMR scenario; hence, the forecasted IMRD becomes also a measure of the accuracy of the RFE models (that ultimately determine the likelihood of the different market scenarios). The distinction between the two cases will become clearer in Sec. 3, where we discuss how to assess model performance.

In the remainder of the paper, we consider BCBS-IOSCO IM as a case study. For B-IMR, the current industry proposal is the ISDA Standard Initial Margin Model (SIMM) [4], a static aggregation methodology to compute IMR based on first order delta-vega trade sensitivities. The exact replication of SIMM in a capital exposure / XVA Monte Carlo (MC) framework requires in-simulation portfolio sensitivities to a large set of underlying risk

Electronic copy available at: https://ssrn.com/abstract=2716279

factors, which is very challenging for most production implementations[1].

Since the exposure simulation provides portfolio MtM values on the default (time t) and closeout (time $t + \text{MPOR}$, being MPOR the Margin Period Of Risk) grids, Andersen, Pykhtin and Sokol [3] have proposed to use this information to infer path-wise the size of of any percentile of the local $\Delta MtM(t, t+\text{MPOR}, \text{path}_i)$ distribution[2] based on a regression which uses the simulated portfolio $MtM(t)$ as the independent variable. This methodology can be further empowered by adding more descriptive variables to the regression, e.g. the values at the default time t of selected risk factors of the portfolio.

For our DIM model the following features are desirable:

(f1) The DIM model should consume the same paths as generated for the exposure simulation, so as to minimise the computation burden.

(f2) The output of the DIM model should reconcile with the known B-IMR value for $t_k = 0$, i.e. $\text{IM}(\text{path}_i, 0) = \text{IMR}_{\text{SIMM}}(0)\ \forall i$.

Before proceeding, we note some of the key aspects of the BCBS-IOSCO margining guidelines and consequently of the ISDA SIMM model [4]:

(a1) The MPOR for the IM calculation of a daily margined counterparty is equal to $10d$. This may differ from the capital exposure calculation where e.g. $\text{MPOR} = 20d$ if the number of trades in the portfolio exceed 5000 (many bilateral portfolios in scope for BCBS-IOSCO phase 1 fall in this category. The MPOR can even be longer in the case of margin disputes between the counterparties over the last two quarters).

(a2) The B-IMR in [1] prescribe to calculate IM segregating trades from different asset classes. This feature is coherently reflected in the SIMM model design.

(a3) The SIMM methodology consumes trade sensitivities as only input and has a static calibration not sensitive to market volatility (at least until the next model recalibration takes place, i.e. no less than yearly).

These features, together with the requirements (f1) and (f2) stated previously are addressed by our model proposal, as we will see.

For the IM calculation, the starting point is similar to [3], i.e. (i) to use a regression methodology based on the paths $MtM(t)$ to compute the

[1] Two common proposals for enabling the computation of in-simulation sensitivities are (i) Adjoint Algorithmic Differentiation (AAD) and (ii) adoption of GPU technology. Both of these approaches have their own drawbacks however, as discussed in more detail in [5].

[2] The $\Delta MtM(t, t+\text{MPOR}) = MtM(t+\text{MPOR}) - MtM(t)$ distribution is constructed assuming that no cash flows take place in between the default and closeout times. For a critical review of this assumption, see [6].

moments of the local $\Delta MtM(t, t+\text{MPOR}, \text{path}_i)$ distribution and (ii) to assume that the $\Delta MtM(t, t+\text{MPOR}, \text{path}_i)$ is a given probability distribution that can be fully characterised by its first two moments drift and volatility. Additionally, since the drift is generally immaterial over the MPOR horizon, we do not compute it and set it to 0.

There are multiple regression schemes that can be used to determine the local volatility $\sigma(i,t)$. For the present analysis, we follow the standard American MC literature [7] and use a Least Squares Method (LSM) approach with a polynomial basis:

$$\sigma^2(i,t) = \left\langle (\Delta MtM(i,t))^2 | MtM(i,t) \right\rangle = \sum_0^n a_{\sigma,k} MtM(i,t)^k \quad (1)$$

$$\text{IM}^{\text{U}}_{R/P}(i,t) = \Phi^{-1}(0.99/0.01, \mu = 0, \sigma = \sigma(i,t)), \quad (2)$$

where R/P indicates received and posted respectively and in our implementation n in Eq. 1 is set equal to 2; i.e. a polynomial regression of order 2 is used. We observe that LSM performs well compared to more sophisticated kernel methods (such as e.g. Nadaraya-Watson used in [3]) with the advantage of being parameter free and cheaper from a computational stand point.

The unnormalised posted and received $\text{IM}^{\text{U}}_{R/P}(i,t)$ are calculated analytically in Eqns. 1 and 2 by applying the inverse of the Cumulative Distribution Function $\Phi^{-1}(x,\mu,\sigma)$ to the appropriate quantiles, being $\Phi((x,\mu,\sigma)$ the probability distribution that models the local $\Delta MtM(t, t+\text{MPOR}, \text{path}_i)$. The precise choice of Φ does not play a crucial role since the difference in quantiles among different distributional assumptions can be compensated in calibration by the scaling factors applied (see $\alpha_{R/P}(t)$ functions in Eq. 4). For simplicity, we assume Φ to be Normal in the below.

As a next step we should account for the $t = 0$ reconciliation and for the mismatch between the SIMM and exposure model calibrations (see respectively items (f2), (a1) and (a3) above). These points can be tackled by scaling $\text{IM}^{\text{U}}_{R/P}(i,t)$ with suitable normalisation functions $\alpha_{R/P}(t)$:

$$\text{IM}_{R/P}(i,t) = \alpha_{R/P}(t) \times \text{IM}^{\text{U}}_{R/P}(i,t) \quad (3)$$

$$\alpha_{R/P}(t) = (1 - h_{R/P}(t))\sqrt{\frac{10d}{\text{MPOR}}} \times \left(\alpha^{\infty}_{R/P} + (\alpha^0_{R/P} - \alpha^{\infty}_{R/P})e^{-\beta_{R/P}(t)t}\right) \quad (4)$$

$$\alpha^0_{R/P} = \sqrt{\frac{\text{MPOR}}{10d}} \cdot \frac{\text{IMR}^{\text{SIMM}}_{R/P}(t=0)}{q(0.99/0.01, \Delta MtM(0, \text{MPOR}))}. \quad (5)$$

Where:

(i) In Eq. 4, $\beta_{R/P}(t) > 0$ and $h_{R/P}(t) < 1$ with $h_{R/P}(t = 0) = 0$ are four functions to be calibrated (two for received and two for posted IMs).

As it will become clearer in Sec. 3, the model calibration generally differs for received and posted DIM models.

(ii) In Eqns. 4 and 5, MPOR indicates the MPOR relevant for Basel III exposure. The ratio among MPOR and $10d$ accounts for item (a1) and is taken in square root because the underlying RFE models are typically Brownian, at least at short horizons.

(iii) In Eq. 5, $\mathrm{IMR}_{R/P}^{\mathrm{SIMM}}(t=0)$ are the $\mathrm{IM}_{R/P}$ computed at $t=0$ using SIMM, $\Delta MtM(0,\mathrm{MPOR})$ is the distribution of MtM variations over the first MPOR and $q(x,y)$ is a function giving the quantile x for the distribution y.

The normalisation functions $\alpha_{R/P}(t)$ are defined at $t=0$ such as to reconcile the $\mathrm{IM}_{R/P}(i,t)$ with the starting SIMM IMR. Instead, the functional form of $\alpha_{R/P}(t)$ at $t>0$ is dictated by what is shown in the upper left panel of Fig. 1: accurate RFE models, both in the P and in the Q measures, have either a volatility term-structure or an underlying stochastic volatility process that accounts for the mean-reverting behaviour generally observed from extreme low or high volatility to normal market conditions. Since the SIMM calibration is static (see item (a3) above), the $t=0$ reconciliation factor is inversely proportional to the current market volatility and not necessarily adequate for the long-term mean level. Hence, $\alpha_{R/P}(t)$ interpolate between the t=0 scaling driven by $\alpha_{R/P}^{0}$ and the long term scaling driven by $\alpha_{R/P}^{\infty}$, where the functions $\beta_{R/P}(t)$ are the mean-reversion speeds. The value of $\alpha_{R/P}^{\infty}$ can be inferred from a historical analysis on a group of representative portfolios or *ad-hoc* calibrated e.g. by computing a different $\Delta MtM(0,\mathrm{MPOR})$ distribution in Eq. 5 using the long end of the RF implied volatility curves and solving the equivalent scaling equation for $\alpha_{R/P}^{\infty}$.

As we will see, the interpretation of $h_{R/P}(t)$ can vary depending upon the intended application of the model: (i) for capital and risk models, $h_{R/P}(t)$ are two haircut functions that can be used to reduce the number of backtesting exceptions (see Sec. 3) and ensure that the DIM model is conservatively calibrated. (ii) For XVA pricing, $h_{R/P}(t)$ can be fine-tuned (together with $\beta_{R/P}(t)$) so as to maximise the accuracy of the forecast based on historical performance.

Notice regarding item (a2) above that the $\mathrm{IM}_{R/P}^{x}(i,t)$ can be computed on a stand-alone basis for every asset class x defined by SIMM (IR/FX, Equity, Qualified and Not Qualified Credit, Commodity) without any additional exposure runs. The total $\mathrm{IM}_{R/P}(i,t)$ is then given by the sum of the $\mathrm{IM}_{R/P}^{x}(i,t)$ values.

A comparison between the forecasts of the DIM model defined throughout Eqns. 1-5 and the historical IMR realisations computed with the SIMM methodology is shown in the upper right panel of Fig. 1, where alternative scaling approaches are also considered. This comparison is performed at different forecasting horizons using 7 years of historical data, monthly sampling and averaging among a wide representative selection of single trade portfolios for the posted and received IM cases. For a given portfolio/horizon, the chosen error metric is given by $\langle |F_{R/P}(t_k+h) - G_{R/P}(t_k+h)| \rangle_{t_k} / \langle G_{R/P}(t_k+h) \rangle_{t_k}$, where $\langle ... \rangle_{t_k}$ indicates an average across historical sampling dates (for $F_{R/P}$ and $G_{R/P}$, see definitions in Sec. 3). The tested universe is made of 102 single-trade portfolios and the products considered, always at the money and of different maturities, include: cross-currency swaps, IR swaps, FX options and FX forwards (approximately 75% of the population is made of $\Delta = 1$ trades).

As evident from the plot, the proposed term structure of $\alpha_{R/P}(t)$ improves the accuracy of the forecasts by a significant amount. The calibration used for this analysis is provided in the caption of Fig. 1. In Sec. 3, we will discuss further which range of values the haircut functions $h_{R/P}(t)$ are expected to take for a conservative calibration of DIM to be used for regulatory capital exposure.

Finally, as an outlook, we show in the lower left panel of Fig. 1 the error metrics for the case of CCP IMR as well, where the DIM forecasts are now compared to PAIRS (LCH) and HVaR (CME) historical realisations[3]. The forecasting capability of the model is separately tested for PAIRS and HVaR IMR and for 22 single-trade portfolios (IRS trades of different maturities and currencies). The error at any given horizon is obtained by averaging among the 22×2 cases.

Without any further fine-tuning the calibration, the time dependent scaling $\alpha_{R/P}(t)$ drives a major improvement in the accuracy of the forecasts w.r.t. the alternative approaches.

3 How to backtest a DIM model?

So far, we have discussed a DIM model for B-IMR without being too specific on how to assess model performance for different applications such as CVA and MVA pricing, Liquidity Coverage Ratio (LCR) / Net Stable Funding Ratio (NSFR) monitoring [8] and capital exposure. As mentioned above, depending on which application one considers, it may or may not be important to have an accurate assessment of the distribution of the simulated IM requirements values (the IMRD). While a perfect model would serve well all purposes, it could be the case that the performance of a realistic implementation may differ significantly across applications.

The aim of this section is to introduce two distinct levels of backtesting that can measure the DIM model performance in the two topical cases of:

[3] The realisations are based on prototype replications of the market risk components of the CCP IM methodologies.

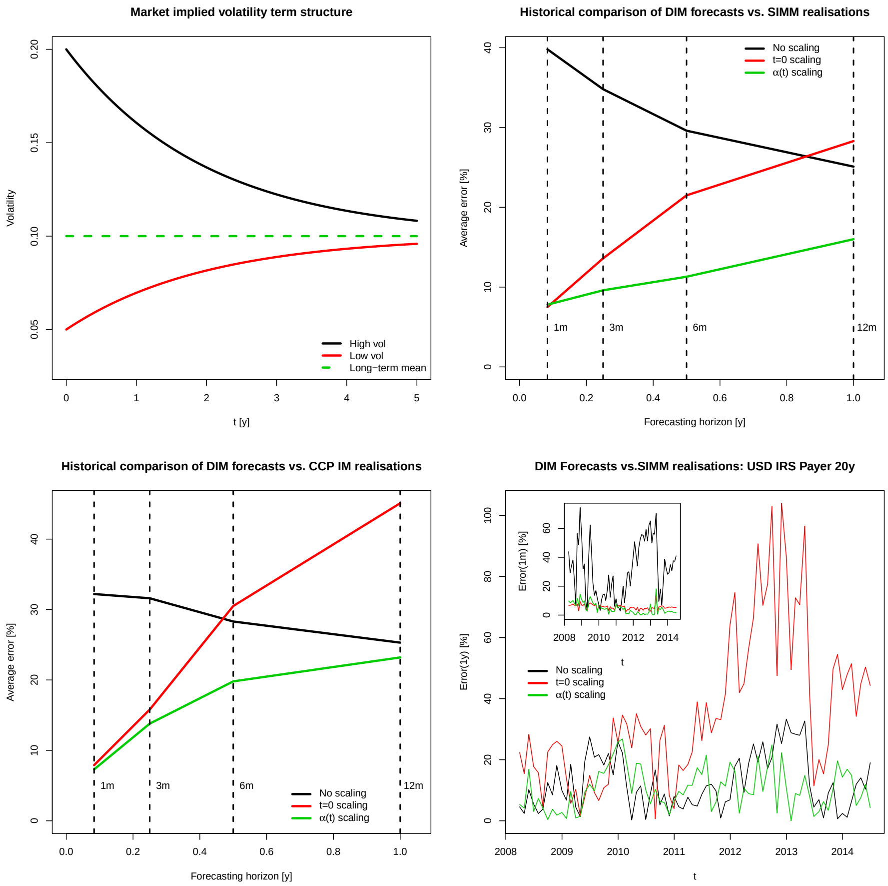

Figure 1: **Upper left panel:** Sketch of the term structure of the RF volatility in the case of low (red line) and high (black line) volatility markets. The dashed green line indicates the long term asymptotic behaviour. **Upper right panel:** The accuracy of the DIM forecasts is measured vs. historical SIMM realisations for three choices of scaling (no scaling: $\alpha(t) = 1$; $t = 0$ scaling: $\alpha(t) = \alpha^0_{R/P}$, as for Eq. 4 with $\beta_{R/P}(t) = 0$ and $h_{R/P}(t) = 0$; $\alpha(t)$ scaling: as for Eq. 4 with $\beta_{R/P}(t) = 1$, $\alpha^\infty_{R/P} = 1$ and $h_{R/P}(t) = 0$). **Lower left panel:** The accuracy of the DIM forecasts is measured for CCP IMR with an equivalent error metric. **Lower right panel:** The realised error of the SIMM DIM for received IM is shown vs. time for a 20y USD IRS payer in the cases of 1m (inbox) and 1y (main plot) forecasting horizons.

(i) DIM applications that do not directly depend on the IMRD (such as capital exposure and CVA) and (ii) DIM applications that do directly depend on the IMRD (such as MVA calculation and LCR / NSFR monitoring). The two corresponding methodologies are presented in sections 3.1 and 3.2

respectively with a focus on P measure applications.

3.1 Backtesting the DIM mapping functions (for capital exposure and CVA)

In a MC simulation framework, the exposure is computed determining the future MtM values of a given portfolio on a large number of forward looking RFs scenarios. To ensure that a DIM model is sound, one should verify that the IM forecasts associated to future simulated scenarios are adequate for a sensible variety of forecasting horizons, initial and terminal market conditions. We should therefore introduce a suitable historical backtesting framework so as to assess statistically the performance of the model by comparing the DIM forecasts with the realised exact IMR (e.g. in the case of B-IMR calculated according to the SIMM methodology) for a representative sample of historical dates / market conditions and portfolios.

Let us define generic IMR for a portfolio p as:

$$\text{IMR} = g_{R/P}(t = t_\alpha, \Pi = \Pi(p(t_\alpha)), \vec{M}_g = \vec{M}_g(t_\alpha)). \tag{6}$$

Where in Eq. 6:

(i) The functions g_R and g_P represents the exact algorithm used to compute the IMR for received and posted IMs respectively (e.g. SIMM for B-IMR or, in the case of CCPs, IM methodologies such as SPAN, PAIRS or HVaR).

(ii) $t = t_\alpha$ is the time at which the IMR for the portfolio p are determined.

(iii) $\Pi(p(t_\alpha))$ is the trade population of the portfolio p at time t_α.

(iv) $\vec{M}_g(t_\alpha)$ is a generic state variable characterising all the $T \leq t_\alpha$ market information required for the computation of the IMR.

Similarly, we define the DIM forecast for the future IMR of a portfolio p as:

$$\text{DIM} = f_{R/P}(t_0 = t_k, t = t_k + h, \vec{r}, \Pi = \Pi(p(t_k)), \vec{M}_{\text{DIM}} = \vec{M}_{\text{DIM}}(t_k)). \tag{7}$$

Where in Eq. 7:

(i) The functions f_R and f_P represents the DIM forecast for received and posted IMs respectively.

(ii) $t_0 = t_k$ is the time at which the DIM forecast is computed.

(iii) $t = t_k + h$ is the time for which the IMR are forecasted (over a forecasting horizon $h = t - t_0$).

(iv) $\vec{r}$ (the "predictor") is a set of market variables whose forecasted values on a given scenario are consumed by the DIM model as input to infer the IMR. The exact choice of $\vec{r}$ is depending on the DIM model. For the one considered in Sec. 2, $\vec{r}$ is simply given by the simulated MtM of the portfolio.

(v) $\vec{M}_{\mathrm{DIM}}(t_k)$ is a generic state variable characterising all the $T \leq t_k$ market information required for the computation of the DIM forecast.

(vi) $\Pi()$ is defined as for Eq. 6.

Despite being computed with the use of stochastic RFE models, f_R and f_P are not probability distributions since they do not carry any information regarding the probability weight of a given received / posted IM value. $f_{R/P}$ are instead mapping functions between the set $\vec{r}$ chosen as a predictor and the forecasted value for the IM. Since the functions $f_{R/P}$ do not have a probabilistic interpretation, the Probability Integral Transformation (PIT) framework [9] is not suitable for backtesting (while it should be used in the case of IMRD forecasting, see Sec. 3.2).

Instead, in terms of $g_{R/P}$ and $f_{R/P}$, one can define exception counting tests. The underlying assumption is that the DIM model is calibrated at a given confidence level (CL); therefore it can be tested as a VaR(CL) model. This comes naturally in the context of real world P applications such as capital exposure or liquidity monitoring where a notion of model conservatism (and hence of exception) is applicable, since the DIM model will be conservative whenever it understates (overstates) received (posted) IM.

For a portfolio p, a single forecasting day t_k and forecasting horizon h, one can proceed as follows:

1. The *forecast functions* $f_{R/P}$ are computed at time t_k as $f_{R/P}(t_0 = t_k, t = t_k + h, \vec{r}, \Pi = \Pi(p(t_k)), \vec{M}_{\mathrm{DIM}} = \vec{M}_{\mathrm{DIM}}(t_k))$. Notice that $f_{R/P}$ depends explicitly only on the predictor $\vec{r}$ ($\vec{r} = MtM$ for the model considered in Sec. 2).

2. The realised value of the predictor $\vec{r} = \vec{R}$ is determined. For the model considered in Sec. 2, $\vec{R}$ is given by the portfolio value $p(t_k + h)$, where the trade population $\Pi(p(t_k + h)$ at $t_k + h$ differs from t_k only because of portfolio aging. Aside for aging, No other portfolio adjustments are made.

3. The *forecasted values* for the received and posted IMs are computed as $F_{R/P}(t_k + h) = f_{R/P}(t_0 = t_k, t = t_k + h, \vec{r} = \vec{R}, \Pi = \Pi(p(t_k)), \vec{M}_{\mathrm{DIM}} = \vec{M}_{\mathrm{DIM}}(t_k))$.

4. The *realised values* for the received and posted IMs are computed as $G_{R/P}(t_k + h) = g_{R/P}(t_0 = t_k + h, \Pi = \Pi(p(t_k + h)), \vec{M} = \vec{M}_g(t_k + h))$.

5. The forecasted and realised values are compared. The received and the posted DIM models are considered independently and a backtesting exception occurs whenever F_R (F_P) is larger (smaller) than G_R (G_P). As discussed above, this definition of exception follows from the applicability of a notion of model conservatism.

Applying the 1-5 program to multiple sampling points t_k, one can detect backtesting exceptions for the considered history. The key step is 3, where the dimensionality of the forecast is reduced (from a function to a value) making use of the realised value of the predictor and hence allowing for a comparison with the realised IMR.

The determination of the test p-value requires the additional knowledge of the Test Value Statistics (TVS) that can be derived numerically also if the forecasting horizons are overlapping (see [10]). In the latter situation, it can happen that a single change from a volatility regime to another may trigger multiple correlated exceptions and hence the TVS should adjust the backtesting assessment for the presence of false positives.

The single trade portfolios of Fig. 1 have been backtested with the above described methodology using SIMM DIM models with the three choices of scaling discussed in the same Figure. The results shown in Tab. 1 confirm the greater accuracy of the term structure scaling $\alpha_{R/P}(t)$. In fact, for the same level of haircut function $h_{R/P}(t > 0) = \pm 0.25$ (positive/negative for Received/Posted), a much lower number of exceptions is detected. We observe in this regard that for realistic diversified portfolios and for a calibration target of $CL = 95\%$, the functions $h_{R/P}(t)$ take values typically in the range (10%-40%)[4].

Notice as well that the goal of the BCBS-IOSCO regulations is to ensure that netting sets are largely overcollateralised (as a consequence of (i) the high confidence level at which the IM is computed and (ii) the separate requirements for daily VM and IM). Hence, the exposure generating scenarios are tail events and the impact on capital exposure of a conservative haircut applied to the received IM is rather limited in absolute terms. See in this regard the right panel of Fig. 2, where the Expected Exposure (EE) at a given horizon t is shown as a function of $h_R(t)$ (haircut to be applied to the received IM collateral) for different distributional assumptions on $\Delta MtM(t, t + \mathrm{MPOR})$.

In the right panel of Fig. 2, $h_R(t) = 0$ and $h_R(t) = 1$ indicate full IM collateral benefit or no benefit respectively and the unscaled IM is taken as the 99th percentile of the corresponding distribution. For different choices of the ΔMtM distribution the exposure reduction is practically unaffected up to haircuts of order $\approx 50\%$.

[4] This range of values for $h_{R/P}(t)$ has been calibrated using $\beta_{R/P}(t) = 1$ and $\alpha^{\infty}_{R/P} = 1$. Both assumptions are broadly consistent with the historical data.

SIMM backtesting: R/P DIM VaR(95%) exceptions / tests					
Scaling type	**1m**	**3m**	**6m**	**12m**	**Total**
$\alpha(t)$ scaling	0/200	2/184	14/152	35/120	51/656
$t = 0$ scaling	0/200	4/184	32/152	42/120	78/656
No scaling	106/200	97/184	79/152	47/120	329/656

Table 1: Historical backtesting for the SIMM DIM model defined throughout Eqns. 1-5, where $h_{R/P}(t > 0) = \pm 0.25$. VaR(95%) exception counting tests are performed for the same single trade portfolios of Fig. 1 and according to the backtesting methodology described in Sec. 3.1. The results are shown in the format (x/y), where y is the number of portfolios tested at a given horizon at x is the observed number of backtesting failures for a p-value acceptance threshold of 5% (received and posted IM cases are considered in aggregate). The backtesting analysis is performed using 7y of historical data and monthly sampling frequency.

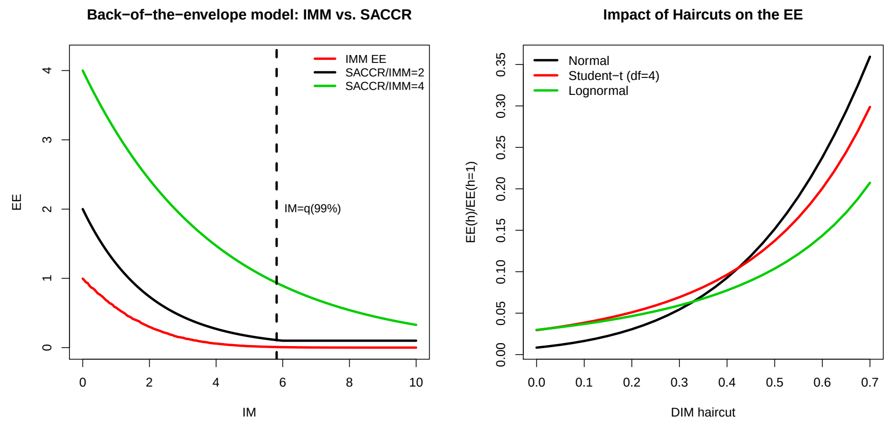

Figure 2: **Left panel:** The EE(t) of a daily VM margined counterparty is shown vs. the received IM for a simplified IMM model (continuous red line) and SA-CCR (continuous black and green lines). The IMM EE(t) is calculated assuming that $\Delta MtM(t, t + \text{MPOR})$ is Normal. Since the SA-CCR EE(t) depends on the size of the trades add-ons, the SA-CCR exposure is computed for the stylised cases of SA-CCR add-ons being two (black line) or four (green line) times larger than the correspondent IMM level. **Right panel:** EE(t) vs. $h_R(t)$ for different distributional assumptions on $\Delta MtM(t, t + \text{MPOR})$: Normal (black line), Student-t (red line) and Lognormal (green line).

3.2 Backtesting the IMRD (for MVA and LCR / NSFR)

The same MC framework can be used in combination with a DIM model to forecast the IMRD at any future horizon (here we implicitly refer to models where the DIM is not always constant across scenarios). The possible business applications of the IMRD are multiple. To mention two that equally apply to the cases of B-IMR and CCP IMR: (i) Future IM funding costs in the Q measure, i.e. MVA. (ii) Future IM funding costs in the P measure, e.g. in relation to LCR or NSFR regulatory requirements [8].

Our focus is on forecasts in the P measure (tackling the case of the Q measure may require a suitable generalisation of [11]). The main difference with the backtesting approach discussed in Sec. 3.1 is that now the model forecasts to be tested are the numerical distributions of simulated IMR values. These can be obtained for a given time horizon by associating to every simulated scenario the correspondent IMR forecast computed according to the given DIM model. Using the notation introduced in Sec. 3.1, the numerical representations of the received / posted IMRD Cumulative Density Functions (CDFs) of a portfolio p for a given forecasting day t_k and horizon h are given by:

$$
\begin{aligned}
\mathrm{CDF}_{R/P}(x, t_k, h) &= \#\left\{ v \in \mathbf{V} \middle| v \leq x \right\} / N_{\mathbf{V}} \qquad (8)\\
\mathbf{V} &= \Big\{ f_{R/C}(t_0 = t_k, t = t_k + h, \vec{r}_\omega, \Pi = \Pi(p(t_k)), \\
&\quad \vec{M}_{\mathrm{DIM}} = \vec{M}_{\mathrm{DIM}}(t_k)), \forall \vec{r}_\omega \in \Omega \Big\}, \qquad (9)
\end{aligned}
$$

where: (i) In Eq. 8, $N_{\mathbf{V}}$ is the total number of scenarios. (ii) In Eq. 9, $f_{R/P}$ are the functions computed using the DIM model, $\vec{r}_\omega$ are the scenarios for the predictor (the portfolio MtM values in the case discussed in Sec. 2) and Ω is the ensemble of the $\vec{r}_\omega$ spanned by the MC simulation.

The IMRD in this form is directly suited for historical backtesting using the PIT framework. Refering to the formalism described in [10], one can derive the PIT time series $\tau_{R/P}$ of a portfolio p for a given forecasting horizon h and backtesting history $\mathcal{H}_{BT}$ as follows:

$$
\tau_{R/P} = \Big\{ \mathrm{CDF}(g_{R/P}(t_k + h, \Pi(p(t_k + h)), \vec{M}_g(t_k + h)), t_k, h), \forall t_k \in \mathcal{H}_{BT} \Big\}. \qquad (10)
$$

In Eq. 10, $g_{R/P}$ is the exact IMR algorithm for the IMR methodology we intend to forecast (defined as for Eq. 6) and the t_k are the sampling points in $\mathcal{H}_{BT}$. Every element in the PIT time series $\tau_{R/P}$ corresponds to the probability of the realised IMR at time $t_k + h$ according to the DIM forecast built at t_k.

As discussed extensively in [10], one can backtest the $\tau_{R/P}$ using uniformity tests. In particular, in analogy to what is shown in [10] for portfolio backtesting in the context of capital exposure models, one can also use test metrics that do not penalise conservative modelling (i.e. models overstating / understating posted / received IM). In all cases, the appropriate TVS can be derived accordingly using numerical MC simulations.

In this set-up, the performance of a DIM model is not tested in isolation. The backtesting results will be mostly affected by:

1. The choice of $\vec{r}$. As discussed in 3.1, $\vec{r}$ is the predictor used to associate an IMR to a given scenario / valuation time point. If $\vec{r}$ is a poor indicator for the IMR, the DIM forecast will be consequently poor.

2. The mapping $\vec{r} \to$ IMR. Similar to the previous item, if the mapping model is not accurate the IMR associated to a given scenario will be inaccurate. For example, the model defined throughout Eqns. 1-5 includes scaling functions to calibrate the calculated DIM to the observed t=0 IMR. The performance of the model is therefore dependent upon the robustness of this calibration at future points in time.

3. The RFE models used for $\vec{r}$. These models ultimately determine the probability of a given IMR scenario. It may so happen that the mapping functions $f_{R/C}$ are accurate or even exact but the probabilities of the underlying scenarios for $\vec{r}$ are misstated, hence causing backtesting failures.

Notice that items 1 and 2 above are relevant also for the backtesting methodology discussed in Sec. 3.1. Item 3 instead is peculiar of this backtesting variance since it concerns the probability weights of the IMRD.

4 Conclusions

In the previous sections, we have presented a complete framework to develop and backtest DIM models. Our focus has been on B-IMR / SIMM and we have shown how to obtain the forward looking IMs from the simulated exposure paths using simple aggregation methods.

The proposed DIM model is suitable both for XVA pricing and capital exposure calculation, where the haircut functions $h_{R/P}(t)$ in Eq. 4 can be used either to improve the accuracy (pricing) or to ensure the conservatism of the forecast (capital).

If a financial institution were to compute CCR exposure using the Internal Model Methods (IMM), the usage of a DIM model could reduce CCR capital significantly, even after the application of a conservative haircut. This should be compared with the regulatory alternative SA-CCR where the benefit from overcollateralisation is largely curbed (see left panel of Fig. 2 and [12]).

As part of the proposed framework, we have introduced a backtesting methodology able to measure model performance for different applications of DIM. The DIM model presented in Sec. 2 as well as the backtesting methodology are agnostic to the underlying IMR algorithm and can be directly applied in other contexts such as CCP IM methodologies, as shown in the lower left panel of Fig. 1.

References

[1] Basel Committee on Banking Supervision, "Margin requirements for non-centrally cleared derivatives," March 2015.

[2] A. D. Green and C. Kenyon, "MVA by replication and regression," *Risk*, May 2014.

[3] L. Andersen, M. Pykhtin, and A. Sokol, "Modeling Credit Exposure for Margined Counterparties," Risk Minds, 2014.

[4] ISDA, "ISDA SIMM Methodology," *draft document*, April 2016.

[5] N. Sherif, "AAD vs GPUs: banks turn to maths trick as chips lose appeal," *Risk*, January 2015.

[6] L. Andersen, M. Pykhtin, and A. Sokol, "Rethinking Margin Period of Risk," *SSRN pre-print*, 2016.

[7] F. Longstaff and E. Schwartz, "Valuing american options by simulation: A simple least- squares approach," *The Review of Financial Studies*, vol. 14, pp. 113–147, 2001.

[8] Basel Committee on Banking Supervision, "Basel III: The Liquidity Coverage Ratio and liquidity risk monitoring tools," January 2013.

[9] F. X. Diebold, T. A. Gunther, and A. S. Tay, "Evaluating density forecasts with applications to financial risk management," *International Economic Review*, vol. 39, no. 4, pp. 863–883, 1998.

[10] F. Anfuso, D. Karyampas, and A. Nawroth, "Credit exposure models backtesting for basel III," *Risk*, September 2014.

[11] L. Jackson, "Hedge backtesting for model validation," *Risk*, September 2013.

[12] F. Anfuso and D. Karyampas, "Capital flaws," *Risk*, July 2015.