

# RISK MANAGEMENT MAGAZINE

**Vol. 17, Issue 2**  
**May – August 2022**

## EXCERPT

<https://www.aifirm.it/rivista/progetto-editoriale/>



**Current and prospective estimate of  
counterparty risk through dynamic neural  
networks**

**Alessio Agnese, Pier Giuseppe Giribone, Francesca Querci**

# Current and prospective estimate of counterparty risk through dynamic neural networks

Alessio Agnese (Duferco Energia), Pier Giuseppe Giribone (Banca Carige - Gruppo BPER, University of Genoa), Francesca Querci (University of Genoa)

Article submitted to double-blind peer review, received on 27<sup>th</sup> April 2022 and accepted on 12<sup>th</sup> July 2022

## Abstract

The estimate of the probability of default plays a central role for any financial entity that wants to have an overview of the risks of insolvency it may incur by having economic relations with counterparties. This study aims to analyze the calculation of such measure in the context of counterparty risk from a current and prospective standpoint, by using dynamic neural networks. The forecasting aspect in the calculation of such risk measure is becoming more and more important over time as current regulation is increasingly based on a "Through the Cycle" and not a "Point in Time" assessment, consequently giving fundamental importance to such estimate. To this end, three different models aimed at calculating the Probability of Default have been investigated: the CDS method, the Z-Spread method, and the KMV method (Kealhofer, Merton and Vasicek). First, the different techniques have been applied to one of the main suppliers of gas and energy in Italy as a reference company. Then, they have been applied to calculate the same risk measure on the 50 companies included in one of the most important European indices, the Euro Stoxx 50.

## Key Words:

Default Probability, Counterparty Risk, Credit Default Swap, Corporate Bond, KMV model, Nonlinear Auto-Regressive (NAR) network, forecasting

## 1) Dynamic neural networks for forecasting a time series

Artificial Neural Networks (ANNs) are a field of Machine Learning and they represent the cornerstone of Deep Learning algorithms. They owe their name and structure to the human brain, as they emulate the way in which biological neurons send signals to each other by acquiring knowledge of the external environment (Arbib, 2003). There are different types of neural networks, and they are classified according to the different purposes for which they are used. The perceptron is the oldest and simplest form of artificial neural net created by Frank Rosenblatt in 1958. It has an input layer, a single hidden layer and an output layer (Rosenblatt, 1958). Feedforward neural networks, also called multi-layer perceptrons (MLPs), are formed by an input layer, two or more hidden layers and an output layer (Haykin, 1994). Convolutional neural networks (CNNs) are similar to feed-forward networks, but their architecture is much more sophisticated (Cun et al., 1990). They are typically used for image recognition, pattern recognition and computer vision (Krizhevsky et al., 2012). This article focuses on recurring neural networks, or RNNs, which are characterized by feedback loops used for modeling time series to make predictions about future results (Kolen and Kremer, 2001). They are widely used in finance, for example for making predictions (Decherchi and Giribone, 2020). In order to make time series predictions using dynamic neural networks, it is necessary to use a sequence of values as input, in the first place, and, subsequently, to set the corresponding network in such a way so that it can use its previous values to best interpret the non-linear relations present in it (Tsay, 2010). We focus on recurring networks of the Non-linear Auto-Regressive network (NAR) type, which exclusively use the endogenous variable to perform this task (Beale et al., 2014). When a neural network is designed with these purposes, the hypothesis behind the reasoning is that the value in  $t$  of the time series can be a function of its past values. Furthermore, the prediction can be carried out at different depth levels: for example, if the goal is to find the value of a share on the following day, a step ahead prediction will be implemented. Likewise, if an analyst is interested in the trend that a certain variable may have over time, he will perform a multiple-step ahead prediction (Giribone et al., 2018). The quantity and quality of available data play a crucial role in the training of the network. Indeed, the available dataset needs to be as deep and complete as possible in order to obtain a statistically reliable forecast (Géron, 2019). The autoregression of the network can be mathematically represented by the following relationship:  $y(t) = f(y(t-1), y(t-2), \dots, y(t-n))$  where  $y(t)$  is the time series to be modeled and  $y(t-1)$ ,  $y(t-2)$ ,  $y(t-n)$  are the past values of the time series itself up to  $t-n$  time lags (Agosto and Giribone, 2019). Similarly to the static version, NARs must also be trained using a gradient method. Through this procedure, the statistical model calibrates its own parameters in order to best interpret the input data. In the design of the architecture, particular attention was paid to the potential problem of over-fitting by dividing the sample of available data into a training-validation-test set (Giribone, 2020). In addition, different network configurations have been examined by putting as a parameter the number of hidden layers, the number of neurons and the lags with which the forecasting problem is to be econometrically faced. The selection measures to reach the best predictive architectures were designed according to the absence of self-correlation of the error, of the  $R^2$  and of the RMSE (Root Mean Squared Error) calculated on the out-of-sample test (Bonini et al., 2019). From an architectural point of view, ANNs are composed of layers of nodes that contain an input layer, one or more hidden layers and an output layer. In a dense artificial neural network, the neurons which belong to the next layer are fully connected with the neurons which are in the current layer. Every connection is characterized by a weight associated to the arch which links the two neurons, and every neuron is itself characterized by an activation threshold and a further parameter called bias. If the output of any single node of the architecture is above a threshold value, this node is activated, sending signals to the next layer of the network. Conversely, if the output is below the threshold value, no information is passed to the next layer (Rojas, 1996). Neural networks rely on training data for their training phase and for improving their accuracy over time. Once optimized, these learning algorithms are very powerful tools and allow to classify and organize data at a very high speed (Fonseca and Lopez, 2017). To understand the working principles of a neural network at an elementary level, each single node can be considered as a linear regression model composed of many inputs, many weights, a threshold and an output (Freeman and Scapura, 1992). Once an input level has been determined, the weights are assigned through a numerical optimization routine: the greater the weight associated to these hyperparameters, the greater the importance of the signal processed by the network. Subsequently, the inputs are multiplied by their respective weights and then added together. Once the process is finished, this first output obtained is passed through an activation function which determines a further output; if the output exceeds a certain significance threshold of the transfer function, then the node is activated, passing the data to the next level of the

network: substantially, the output of one node becomes the input of the next node. This process of passing data from one level to another defines this type of neural network as a feedforward network (Caligaris *et al.*, 2015). During the training of the model, one of the needs is to evaluate its accuracy, for example through the Mean Squared Error (Chollet, 2018). The goal is therefore to reach a convergence point or a local minimum, through a progressive adjustment of the weights included in the algorithm (Principe, 2000). Neural networks are for the most part feedforward, that is, the signal flows in only one direction from the input to the output. However, the model adopts the back-propagation technique during the training, that is, information provided by the minimization of the cost function moves in the opposite direction: from output to input. The backpropagation allows to calculate and attribute the error associated to each neuron, allowing to appropriately adjust and calibrate the parameters (Rumelhart *et al.*, 1986). It is beyond the scope of this article to provide further technical details on the operating principles of a NAR. Interested readers can refer to the works mentioned above.

## 2) Current and prospective estimation of the Probability of Default using CDS premiums

Estimating the Probability of Default is of fundamental importance for financial institutions as it provides a summary information on the creditworthiness associated with the counterparty. The objective of this section is to analytically explain how to estimate the probability of default through CDS premiums and to provide two operational examples of calculation: the first is implemented by estimating the default probability at the present valuation time, so-called time  $t_0$ , and the second is implemented by calculating the same measure, but in a prospective way, that is after a forecasting of the CDS premiums made through dynamic neural networks. Given that the CDS is an insurance that hedges the holder from the loss in case of default of the issuer, it can directly measure the counterparty risk. As a result, this methodology is often preferred by the market players in comparison to others. In other words, if the counterparty in question has listed CDSs, it is preferable to use them for calculating the probability of default as such derivatives are expressly aimed at hedging credit risk (Bottasso *et al.*, 2019) through the following formula (Hull, 2015):

$$PD(T) = 1 - \exp(-\bar{\lambda}(T)T) \quad (1)$$

Where:

$PD(T)$  is the probability of default before  $T$ ;

$\bar{\lambda}(T)$  is the hazard rate

$T$  is the time to maturity expressed in years;

The above stated formula can be derived through the theory of probability by proceeding with the following steps. The hazard rate at time  $t$  is defined in such a way that  $\lambda(t)\Delta t$  is the probability of default between time  $t$  and time  $t + \Delta t$  conditional on the fact that there be no default before. In this sense,  $SV(t)$  is the cumulative probability that the company survives at time  $t$ .

Therefore, the conditional default probability between time  $t$  and time  $t + \Delta t$  is:

$$PD(t, t + \Delta t) = \frac{SV(t) - SV(t + \Delta t)}{SV(t)} \quad (2)$$

Since expression (2) is equal to  $\lambda(t)\Delta t$ , then we have:

$$\frac{SV(t + \Delta t) - SV(t)}{SV(t)} = -\lambda(t)\Delta t \quad (3)$$

$$\frac{SV(t + \Delta t) - SV(t)}{\Delta t} = -\lambda(t)SV(t) \quad (4)$$

By setting  $\Delta t \rightarrow 0$  and assuming  $SV(t)$  differentiable (i.e.,  $SV \in C^1$ )

$$\frac{dSV(t)}{dt} = -\lambda(t)SV(t) \quad (5)$$

And by solving the Ordinary Differential Equation for  $SV(t)$ , the following general solution is obtained:

$$SV(t) = \exp\left[-\int_0^t \lambda(\tau)d\tau\right] \quad (6)$$

Defining  $PD(t)$  as the issuer's probability of default at time  $t$  such that  $PD(t) = 1 - SV(t)$  we therefore have:

$$PD(t) = 1 - \exp\left[-\int_0^t \lambda(\tau)d\tau\right] = 1 - \exp[-\bar{\lambda}(t)t] \quad (7)$$

Where  $\bar{\lambda}(t)$  is the average hazard rate or, equivalently, the default intensity, between time 0 and time  $t$ .

In the case of the CDS market, the premium can be seen as a direct compensation received by the insurer for the potential event of default by the issuer until maturity  $t = T$ . This means that the average loss rate between time 0 and time  $T$  can be annually approximated by the CDS premium  $s(T)$ , typically expressed in basis points.

Assuming now that the average hazard rate during such period is  $\bar{\lambda}(T)$  and taking into consideration that the Recovery Rate ( $RR$ ) in a standard Credit Default Swap is 40% (Giribone *et al.*, 2014), the average loss rate can be expressed by the quantity:  $\bar{\lambda}(T)(1 - RR)$ . This means that the following relationship is reasonably valid:

$$\bar{\lambda}(T)(1 - RR) = s(T) \rightarrow \bar{\lambda}(T) = \frac{s(T)}{(1 - RR)} \quad (8)$$

## 2.1) Pre-processing data for the default probability estimation at the current time

Before calculating the issuer's probability of default, a preliminary analysis of the dataset was necessary in order to verify, and potentially correct, the anomalies present in it. As a first step, after retrieving the CDS premiums at different maturities from Bloomberg®, Matlab was used to check the data and their consistency. The first check consists in analyzing the homogeneity of the dataset by verifying that there are no missing data, or that the CDS premiums are characterized by a daily frequency in the trading days considered in the analysis. In the case of anomalies of this kind, the missing values were linearly interpolated. After having interpolated any missing data, the various time series were grouped monthly to have a single monthly value for each different CDS maturity. To do this, the arithmetic mean of the different CDS Premiums was calculated for each month in order to obtain a dataset that is more aligned with the prospective time span of interest. The last check that was implemented, before the actual calculation of the probability of default, was the one relating to the presence of any potential outlier in the dataset. To search for such potentially odd values, the procedure is as follows: the historical series of the CDS Premiums were grouped two by two in ascending order (i.e., first grouping: 6M and 1Y, second grouping: 1Y and 2Y, etc.), then the difference between the different historical series was calculated and, on such difference, the mean and the standard deviation were calculated. The final step consists in constructing the interval within which the values are expected to fall. If the differences exceed the confidence interval, the premiums of the corresponding CDS are interpolated. The procedure can be exemplified with the following logic of preparation of the time series for CDS with tenors of six months and one year:

$$CDS6M = [CDS6M_1, CDS6M_2, \dots, CDS6M_i, \dots, CDS6M_n]$$

$$CDS1Y = [CDS1Y_1, CDS1Y_2, \dots, CDS1Y_i, \dots, CDS1Y_n]$$

With  $i = 1, \dots, n$  where  $n$  is the length of the time series. This procedure has been repeated for each CDS tenor (6M-1Y-2Y-3Y-4Y-5Y-7Y-10Y).

$$D_{1Y-6M} = [D_1 = CDS1Y_1 - CDS6M_1, \dots, D_i = CDS1Y_i - CDS6M_i, \dots, D_n = CDS1Y_n - CDS6M_n]$$

$$D_{2Y-1Y} = [D_2 = CDS2Y_1 - CDS1Y_1, \dots, D_i = CDS2Y_i - CDS1Y_i, \dots, D_n = CDS2Y_n - CDS1Y_n]$$

This procedure has been repeated for each couple of CDS tenors.

$$\mu = \frac{1}{n} \sum_{i=1}^n D_i \quad (9)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (D_i - \mu)^2}{n}} \quad (10)$$

$$\inf = \mu - 3\sigma$$

$$\sup = \mu + 3\sigma$$

$$\text{if : } \begin{cases} \inf \leq D_i \leq \sup \rightarrow \text{market CDS premium} \\ D_i < \inf \text{ or } D_i > \sup \rightarrow \text{interpolated CDS premium} \end{cases}$$

Where:

$D$  is the vector containing the differences of the values contained in the CDS vectors;

$\mu$  is the mean of the differences of the values vector;

$\sigma$  is the standard deviation of the vector;

$\inf$  is the lower extreme of the interval;

$\sup$  is the upper extreme of the interval;

As a result, the linear interpolation between two contiguous values ( $i-1$  e  $i+1$ ) of the time series has been carried out only when the CDS premiums provided by the market are considered outliers, i.e. they exceed the threshold of their mean  $\pm 3$  standard deviations. Once the control phase on the dataset was completed, the default probability was estimated for the various historical series of the CDS according to formula (1).

## 2.2) Estimation of the forecasted default probability

For the calculation of the prospective default probability, it was first necessary to forecast the CDS Premiums for each selected maturity. For forecasting purposes, NAR artificial neural networks were used. For each CDS maturity, 27,750 networks (for a total of 222,000) were tested in order to select the 8 best networks (one for each CDS reference tenor) for forecasting the data (Table 1). The selection of the most performing networks has been conducted only among the models characterized by the absence of auto correlation error (Econometric Test) and a good  $R^2 = 0.95$  on validation set (Quality Accuracy). Among these architectures the best networks are chosen in accordance with the minimum out-of-sample gap. Table 1 also shows the In-Sample error and the average between the In-sample and Out-of-Sample error. This last quantity has been taken into account in the choice of the best network when the machine learning model with the lowest out-of-sample errors produced unfeasible forecasts (for example prospective negative premiums).

Non Linear Auto Regressive Best Networks					Econometric Test (Boolean)		Quality Accuracy
					TRUE		≥0.95
CDS	Number of Delays	Neurons (First Layer)	Neurons (Second Layer)	Neurons (Third Layer)	In Sample RMSE	Out Sample RMSE	MEAN IN SAMPLE AND OUT SAMPLE RMSE
6M	23	6	5	5	3.618442834	3.71288727	3.66566506
1Y	23	8	10	7	2.042700524	2.76245372	2.59477949
2Y	23	5	9	10	5.448432725	2.17786316	3.81314794
3Y	25	7	10	3	7.080359036	3.85242165	5.46639035
4Y	24	10	10	3	7.665926988	3.82910442	5.74751570
5Y	25	5	4	2	11.01484139	11.19604652	11.10544396
7Y	23	7	10	6	7.294506369	6.95375541	7.12413089
10Y	24	3	8	10	7.411680334	6.04787764	6.72977898

Table 1 Architecture of the Best networks: CDS method

After having trained and selected the various best networks, we proceeded with the forecasting of the various CDS Premiums and with the calculation of the default probability.

The figures  $\tau = \{0.5, 1, 2, 3, 4, 5, 7, 10\}$  have been estimated, as the year fractions  $t$  vary and for each tenor provided by the market, the corresponding listed spread  $s(t, t + \tau)$  and the corresponding  $PD(t, t + \tau)$ .

In particular:

$t = 0$  corresponds to the first sampling date, which is 18<sup>th</sup> June 2008.

$t = 11.84$  corresponds to the date on which the network training set ends, that is 29<sup>th</sup> February 2020. The training set is characterized by the black line in Figure 1 both for the spreads,  $s(t, t + \tau)$  and for the derived probability,  $PD(t, t + \tau)$ .

$t = 12.85$  corresponds to the date on which the network test set ends, i.e., 28<sup>th</sup> February 2021. The test set is characterized by the red line in Figure 1 both for the spreads,  $s(t, t + \tau)$  and for the derived probability,  $PD(t, t + \tau)$ .

For  $t > 12.85$  the forecasting is implemented. Both the predicted values of  $s(t, t + \tau)$  and the value of  $PD(t, t + \tau)$ ; they are characterized by the blue line in Figure 1.

By way of example, the graph shows the premiums of the various CDS with tenor equal to 7 years and the corresponding default probability estimated using the CDS method.

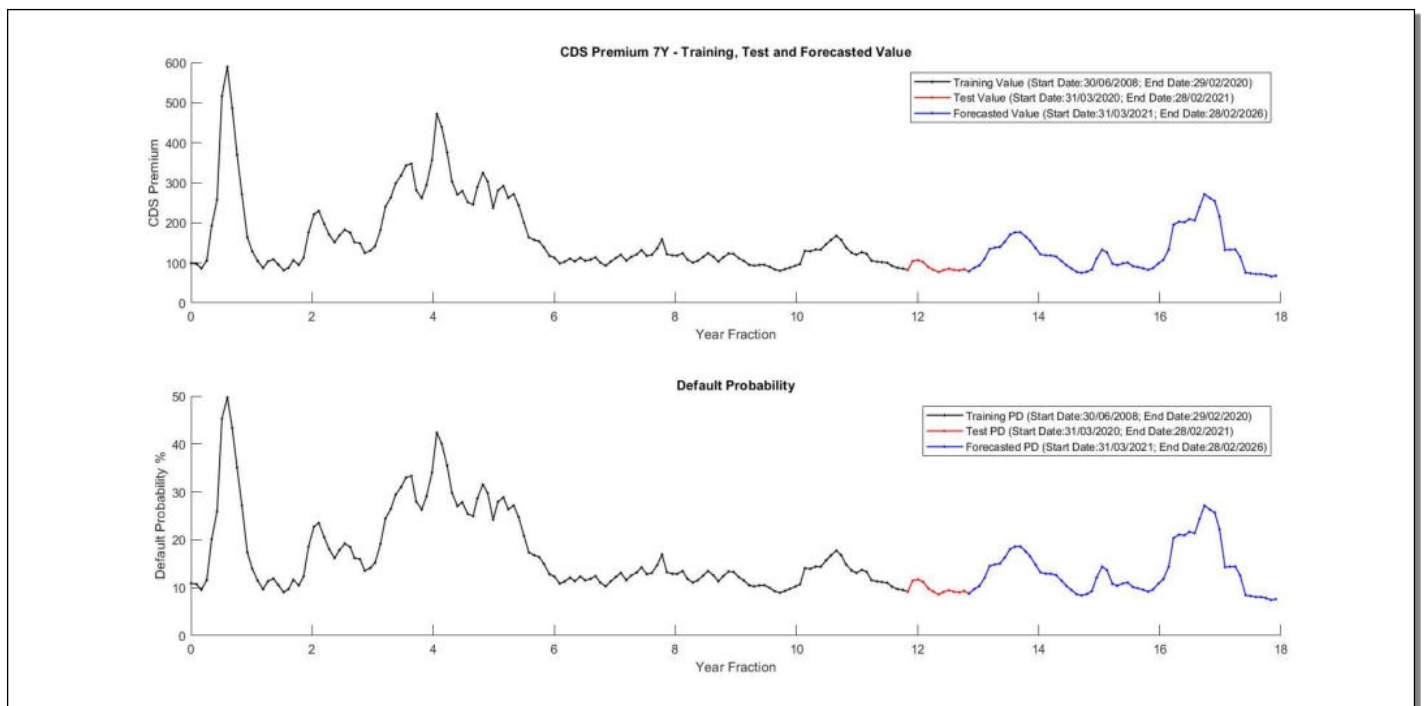
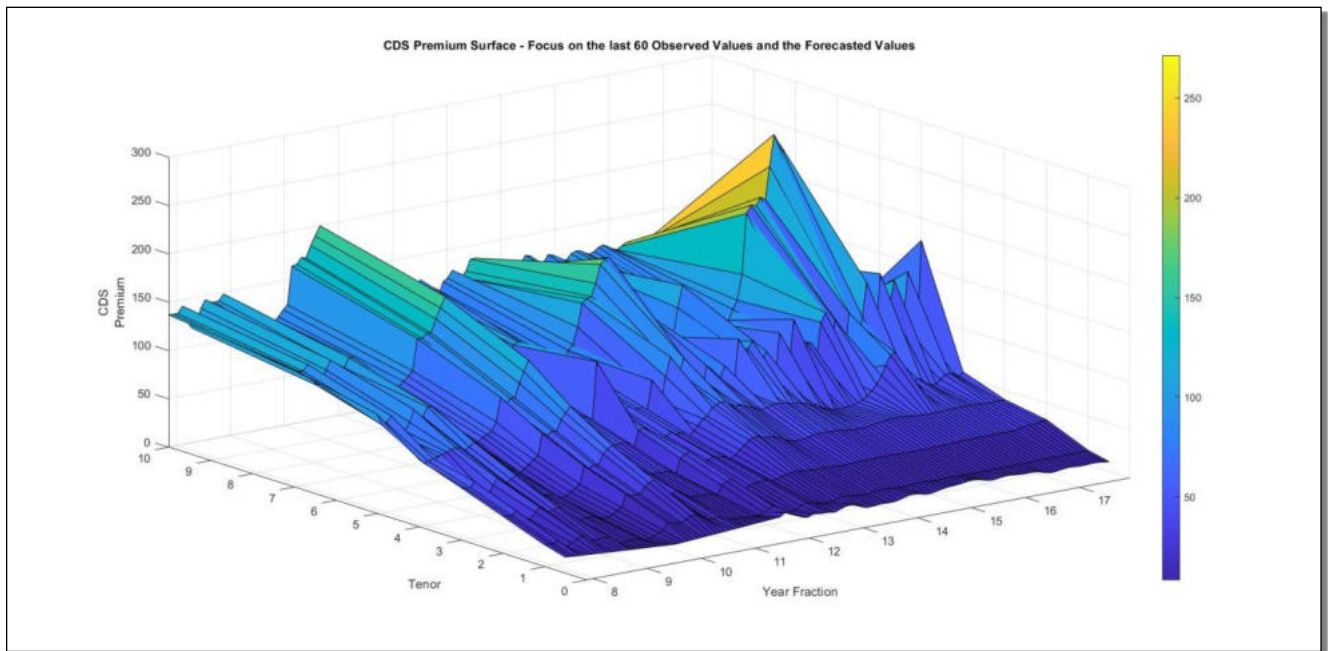


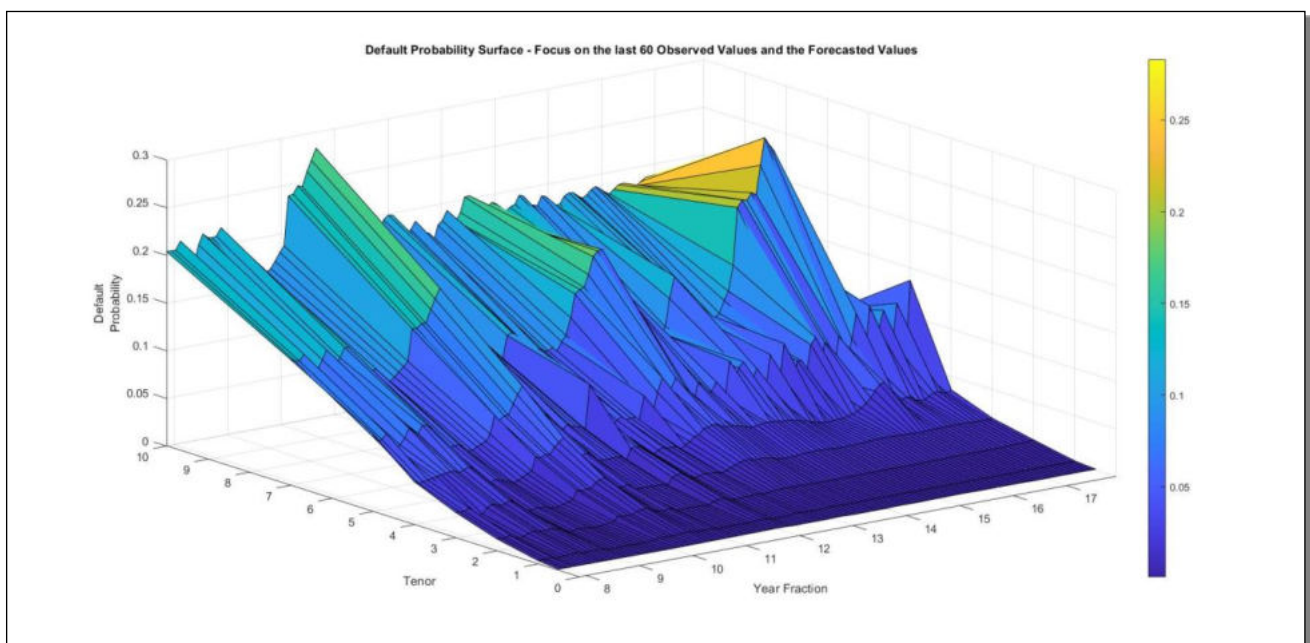
Figure 1 CDS premium (tenor: 7 years) and Default Probability



The further graphs below represent the surfaces of the CDS Premiums and of the Default Probability: the fractions of year are represented on the x-axis, the different tenors are represented on the y-axis and the different premiums are on the z-axis (Figure 2) and the corresponding implied probabilities (Figure 3).



*Figure 2 Observed and Forecasted CDS premium surface*



*Figure 3 Default Probability surface obtained with the CDS method*

### 3) Current and prospective estimation of the Probability of Default using listed bonds

The second approach for calculating the Default Probability consists in estimating the Z-Spread and using such figure as a proxy for a synthetic CDS premium. The Z-Spread indicates the excess return that a bond must give with respect to the risk-free rate and, usually, this yield differential is the compensation for a potential default of the bond issuer (Hull, 2015).

A clearer explanation of this risk transfer can be provided by considering a portfolio composed as follows:

A corporate bond with a 5-year maturity and a yield of 5%;

A long position on a 5-year Credit Default Swap, which has a cost of 250 basis points per year.

This portfolio is approximately equivalent to a long position in a risk-free instrument offering a rate of return of 2.5% per annum. The effect that the CDS has on the composition of the above-mentioned portfolio is to "transform" a risky instrument, such as a corporate bond, into a risk-free instrument. This is easily understood from the fact that, if the issuer of the bond does not incur a default, the return for the holder of the above-mentioned portfolio is equal to 2.5%, or 5% profit deriving from the corporate bond minus 2.5% which is the Credit Default Swap premium.

On the other hand, if the issuer were to go bankrupt, the investor would have a profit of 2.5% until the event of default and the full notional capital would be repaid thanks to the CDS. Furthermore, this capital can be reinvested at the risk-free rate for the time between the default and the maturity of the security.

In theoretical terms, the spread of a T-year CDS ( $s$ ), should be equal to the difference, in terms of yield, between a corporate bond with a maturity of T-years and a risk-free security with the same maturity. In mathematical terms, this can be expressed as follows:

$$s_T = y_T - r_T \quad (11)$$

Where:

$s$  is the excess spread;

$y$  is the yield of the corporate bond;

$r$  is the return on the risk-free security.

$T$  is the maturity.

If the above stated would not happen, arbitrage opportunities would arise.

### 3.1) Procedural example

In order to illustrate the calculation of the Z-Spread, a corporate bond issued on 18<sup>th</sup> May 2020 at a price equal to 99.308 was taken into consideration. The bond expires 6 years from the date of issue (18<sup>th</sup> May 2026) and pays an annual coupon of 1.25%. Based on the Moody's rating scale, this bond is rated Baa, which means that it is considered a security with a moderate credit risk. On the day our analysis was carried out, i.e., 26<sup>th</sup> February 2021, the bond's market price was 105.595. The goal is to price the bond with the characteristics listed above. The feature to be evaluated is the spread composed of the part related to liquidity risk plus the credit risk, however the bonds suitable for this purpose should be selected among the most liquid securities, which means that the liquidity risk can be considered negligible and, therefore, the spread is almost entirely associated with credit risk. In order to understand the impact of credit risk, a theoretical valuation model was constructed and, therefore, the bond was first priced at the risk-free rate. The zero rates were derived from the term structure of the 6-month Euribor which, typically, in the context of pricing a fixed-income instrument, is used as the best proxy for the risk-free rate. The formulation of cash flows is shown in Table 2

Payment Date	Year Fraction	Cash Flows	Zero Rate	Adjusted Zero Rate	Adjusted Discounted Factor	Net Present Value
18/05/2021	0.222	1.25	-0.516%	-0.516%	1.0011487	1.2514359
18/05/2022	1.222	1.25	-0.460%	-0.460%	1.0056497	1.2570621
18/05/2023	2.222	1.25	-0.454%	-0.454%	1.0101677	1.2627096
18/05/2024	3.222	1.25	-0.407%	-0.407%	1.0132148	1.2665185
18/05/2025	4.222	1.25	-0.348%	-0.348%	1.0148243	1.2685304
18/05/2026	5.222	101.25	-0.283%	-0.283%	1.0148966	102.7582784
Z-Spread	0	Dirty Price	109.0645349			
Value Date	26/02/2021	Accrued	0.9726027			
Issue Date	18/05/2020	Clean Price	108.0919322			

Table 2 Risk-free pricing of the corporate bond examined as of 26<sup>th</sup> February 2021

Where:

the Payment Dates are the dates on which the coupon of the bond is paid and, at maturity, the repayment of the notional;

the Year Fractions (YF) are the fractions of the year corresponding to the Payment Dates;

the Cash Flows (CF) are the cash flows generated by the bond, i.e., the coupons until 18<sup>th</sup> May 2025 and the Face Amount plus the coupon on 18<sup>th</sup> May 2026;

regarding the values of the Zero Rate (ZR) they have been interpolated following the corresponding year fractions of the Euribor term structure;

the Adjusted Zero Rate (AZR) is the risk-free rate adjusted according to the Z-Spread (Z);

the Adjusted Discounted Factor (ADF) is the adjusted discount factor which, in this case, is equal to the discount factor as there is no additional spread;

the Net Present Values (NPV) correspond to the present value of the cash flows;

the Dirty Price (DP) is the sum of the NPVs including the accrued interest (Accrued Interest);

the Clean Price (CP) is the sum of the NPVs, net of accrual;

the Valuation Date corresponds to the day on which the pricing of the corporate bond was made;

the Issue Date corresponds to the day on which the corporate bond was issued.

With the bond valued in these terms, i.e. at the risk-free rate (which is why the Z-Spread is equal to 0), a value of 108.092 is obtained.

Since the considered corporate bond is a liquid bond, the market provides a contribution which, in this case, is equal to 105.595, lower than the result of the theoretical model built above.

To take creditworthiness into account, the bond should be priced considering a spread. In order to identify the exact value of the Z-Spread that returns the market value of the corporate bond, the problem should be solved by means of a Goal Seeking method programmed in Matlab (Byrd *et al.*, 1994).

Payment Date	Year Fraction	Cash Flows	Zero Rate	Adjusted Zero Rate	Adjusted Discounted Factor	Net Present Value
18/05/2021	0.222	1.25	-0.516%	-0.057%	1.0001276	1.2501595
18/05/2022	1.222	1.25	-0.460%	-0.001%	1.0000180	1.2500225
18/05/2023	2.222	1.25	-0.454%	0.004%	0.9999053	1.2498817
18/05/2024	3.222	1.25	-0.407%	0.052%	0.9983277	1.2479096
18/05/2025	4.222	1.25	-0.348%	0.111%	0.9953429	1.2441786
18/05/2026	5.222	101.25	-0.283%	0.176%	0.9908704	100.3256296
Z-Spread	45.852	Dirty Price	106.5677815			
Value Date	26/02/2021	Accrued	0.9726027			
Issue Date	18/05/2020	Clean Price	105.5951788			

Table 3 Risk-Adjusted pricing of the examined corporate bond as of 26<sup>th</sup> February 2021

In order to estimate a synthetic term structure of CDS we need to consider several liquid bonds of the same issuer and to repeat the calculation for all days in which the market provides a price. Starting from the summary of this term structure of the CDS premiums, the Probabilities of Default are obtained with the technique described in the previous section. The 16 bonds selected for the considered issuer are shown in Table 4.

ID	Coupon	Maturity	Payment Frequency	Day Basis
A	2.625	22/11/2021	Annual	ACT/ACT
B	0.750	17/05/2022	Annual	ACT/ACT
C	3.250	10/07/2023	Annual	ACT/ACT
D	1.750	18/01/2024	Annual	ACT/ACT
E	0.625	19/09/2024	Annual	ACT/ACT
F	1.000	14/03/2025	Annual	ACT/ACT
G	3.750	12/09/2025	Annual	ACT/ACT
H	1.500	02/02/2026	Annual	ACT/ACT
I	1.250	18/05/2026	Annual	ACT/ACT
J	1.500	17/01/2027	Annual	ACT/ACT
K	1.625	17/05/2028	Annual	ACT/ACT
L	1.125	19/09/2028	Annual	ACT/ACT
M	3.625	29/01/2029	Annual	ACT/ACT
N	0.625	23/01/2030	Annual	ACT/ACT
O	2.000	18/05/2031	Annual	ACT/ACT
P	1.000	11/10/2034	Annual	ACT/ACT

Table 4 Corporate bonds used for the analysis

In this example as well, the probability of default was calculated both at time zero and prospectively. This required the use of NAR neural networks, to implement the various forecasting of the values (de Simon-Martin *et al.*, 2020).



Therefore, 27,750 different network architectures were tested for each single bond, for a total of 440,000, in order to select the 16 best networks used for forecasting the different Z-Spreads, whose architecture is shown in Table 5.

The criteria used for the selection of the most performing networks remain the same that we have previously described for Table 1.

Non Linear Auto Regressive Best Networks					Econometric Test (Boolean)		Quality Accuracy
					TRUE		≥0.95
Bond	Number of Delays	Neurons (First Layer)	Neurons (Second Layer)	Neurons (Third Layer)	In Sample RMSE	Out Sample RMSE	MEAN IN SAMPLE AND OUT SAMPLE RMSE
A	17	9	9	5	1.525955908	0.65624776	1.09110183
B	17	8	7	1	1.387777305	0.63832195	1.01304963
C	11	1	4	2	1.665444521	0.88516771	1.27530612
D	16	8	8	4	1.720082629	0.59825059	1.15916661
E	18	9	7	5	1.587440954	0.63539117	1.11141606
F	18	9	5	6	1.68540112	0.63272568	1.15906340
G	16	8	7	8	1.756757209	0.71813277	1.24244499
H	12	8	5	8	2.071197384	0.59747441	1.33433590
I	19	10	10	8	0.76682641	0.56315685	0.66499163
J	12	6	10	6	1.029262298	0.87700847	0.95313538
K	3	8	1	1	1.359849175	0.93094883	1.14540900
L	19	10	3	9	1.722871502	0.72863780	1.22575465
M	9	8	9	8	2.243363579	0.78521178	1.51428768
N	17	10	8	10	1.426791275	0.61658974	1.02169051
O	20	9	6	9	0.754451999	0.47453179	0.61449190
P	18	8	2	2	1.099202997	0.58227227	0.84073763

Table 5 Architecture of the Best networks: Z-spread method

By way of example, the graph of bond O is shown in Figure 4, in which, in the upper part, the different time series used and the respective forecasting of the synthetic CDS premiums are shown and, in the lower part, a detail of the forecasting and the last 44 observed values.

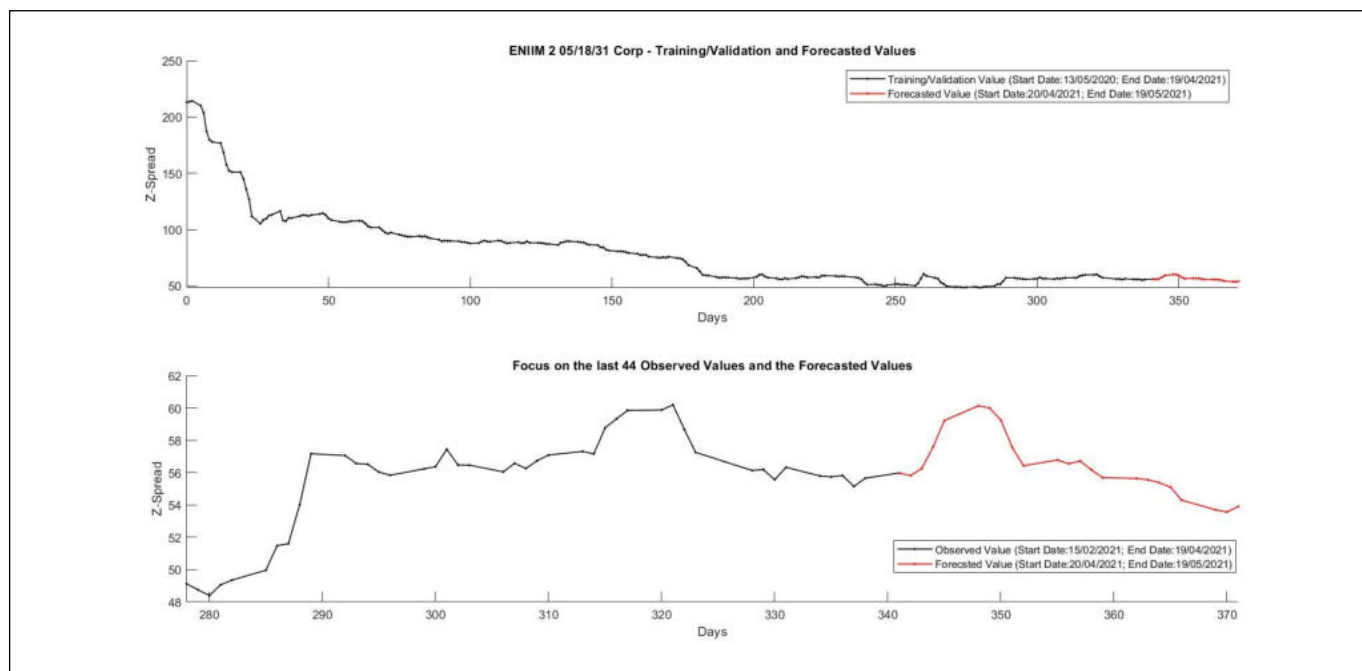


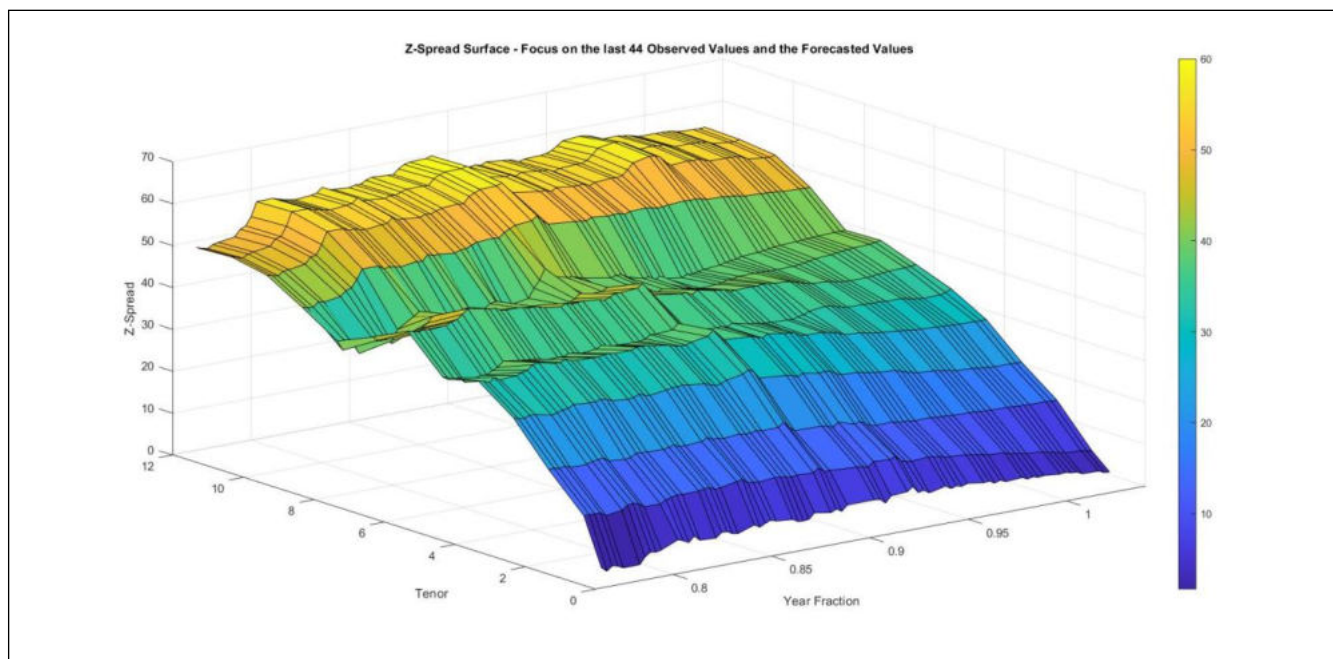
Figure 4 Training / Validation and Forecasting for Bond with ID: O

Once all the necessary forecasting for the prospective estimate of the default probability had been implemented, a further step was required which is not present in the operational example of the Credit Default Swaps. In fact, each single bond is characterized by its own time to maturity (as shown in Table 4).

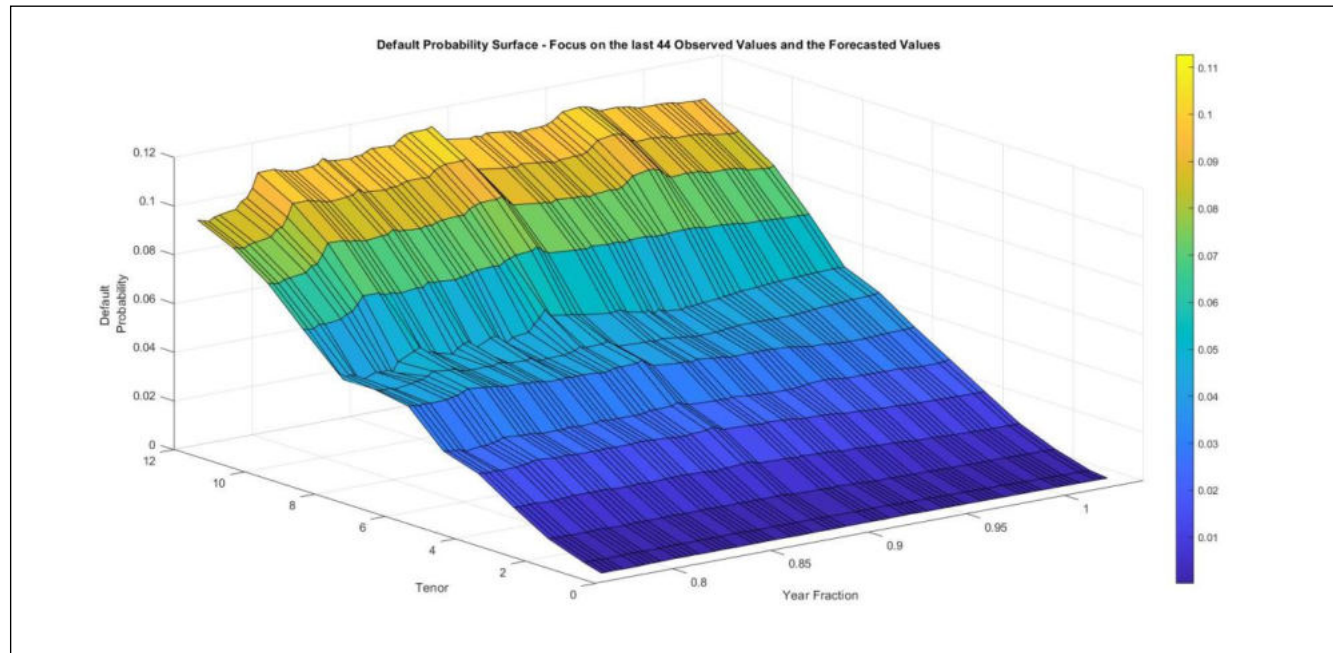
It was thus necessary to regularize the Z-Spreads for the standard market tenors, that is: 6 months, 1 year, 2 years, 3 years and so on up to 12 years. Therefore, for each single day, the synthetic spread corresponding to each standard tenor of the term structure was interpolated across the bonds.

At this point, following this standardization procedure, the problem to be solved becomes like the case of the CDS.

The surfaces shown in Figures 5 and 6 show a summary of the results obtained.



*Figure 5 Surface of the implicit spreads on the analyzed bonds*



*Figure 6 Default Probability surface obtained with the Z-spread method*

#### 4) Current and prospective estimation of the Probability of Default using the KMV method

If the company in question does not have listed Credit Default Swaps or listed corporate bonds on the markets, the KMV model (Kealhofer, Merton and Vasicek) can be used to estimate the probability of default (Bharath and Shumway, 2004).

KMV is a structural model usually implemented in credit risk management and many different versions of this model were successfully implemented (Agosto & Moretto, 2012). Its application is suggested in the estimation of the counterparty risk when market data are not available and the main source of information for the corporate analysis is the balance sheet. As a result, this approach is not suitable for a short-term forecasting ( $T < 1$ ).

The KMV method is based on the equity price and on the balance sheet of the company in question. The assumptions underlying the Merton model can be divided into the following 4 sections:

The Debt is homogeneous and has a time to maturity equal to  $T$ ;

The capital structure of the company in question is given by debt and equity. Consequently, it is true that:  $V_A(t) = D(t) + V_E(t)$ , where:  $V_A(t)$  is the value of the assets at time  $t$ ,  $D(t)$  is the debt to be repaid and  $V_E(t)$  is the value of the company's equity at time  $t$ ;

The model assumes that value of the firm assets follows a Brownian geometric motion of the following type:  $dV_A = \mu_A V_A dt + \sigma_A V_A dW_t$ , where:  $\mu_A$  is the expected instantaneous rate of return,  $\sigma_A$  is the volatility and  $dW_t$  is a Wiener process;

The hypothesis of perfect markets applies: there are no taxes; there are no restrictions on short selling; the market is completely liquid, i.e. investors can buy and sell any asset at market price; sellers and buyers have the same as risk free rate and such interest rate is constant over the reference time span.

Based on the previous assumptions, Robert Cox Merton proposed, in 1974, a model where a company's equity is an option on the assets of the company. Based on the structural model proposed (Merton, 1974), if  $V_A(T) < D$  it is likely, at least in theory, that the company surely defaults at time  $T$ ; in this case the equity at time  $T$  is equal to 0 ( $V_E(T) = 0$ ). Likewise, if  $V_A(T) > D$ , then the company in question should repay its debt at time  $T$  and, in this case, the value of the equity at that same time will be equal to:  $V_E(T) = V_A(T) - D$ . In strictly mathematical terms, the amount of the company's equity at time  $T$  is given by the following pay-off:

$$V_E(T) = \max(V_A(T) - D, 0) \quad (12)$$

Given the previous statement, the analogy with a European-type option is clear: the value of the equity is similar to the payoff of a call option on the value of the assets with a strike price equal to the payment requested on the debt.

Under the assumptions of this model, the traditional Black-Scholes-Merton (BSM) model for option pricing (Black and Scholes, 1973) can be applied and, consequently, the following equations are valid:

$$V_E(t) = V_A(t)\phi(d_1) - \exp(-r(T-t))D\phi(d_2) \quad (13)$$

$$d_1 = \frac{\ln\left(\frac{V_A(t)}{D}\right) + \left(r - \frac{1}{2}\sigma_A^2\right)(T-t)}{\sigma_A\sqrt{T-t}}$$

$$d_2 = d_1 - \sigma_A\sqrt{T-t}$$

Where:

$T$  is the time when the valuation is made;

$\sigma_A$  is the asset volatility;

$r$  is the risk-free rate at time  $T$ ;

$V_A(0) - V_E(0)$  is the value of the debt at time 0;

$\phi(-d_2)$  is the risk neutral probability that the company will default on its debt at maturity  $T$ .

However, in order to calculate the default probability  $\phi(-d_2)$ , a further relationship is necessary since there are two unknown parameters in the above equations, namely the value and the volatility of the assets (respectively:  $V_A$  and  $\sigma_A$ ). The other variables of the model are, on the other hand, directly observable or calculable: the risk-free rate ( $r$ ) can be selected by referring to the Euribor term structure; the value of the debt ( $D$ ) is directly observable from the financial statements of the company in question; the equity value ( $V_E$ ) can be calculated by multiplying the number of shares of the company by their unitary market value; the volatility of the equity ( $\sigma_E$ ) can be calculated using the implied volatility or through a traditional backward looking econometric approach such as the GARCH. Applying Ito's lemma, the stochastic dynamics that regulate the behavior of  $V_E$ :

$$dV_E = \mu_E V_E dt + \sigma_E V_E dW_t \quad (14)$$

can be re-written as:

$$dV_E = \left(\frac{1}{2}\sigma_A^2 V_A^2 \frac{\partial^2 V_E}{\partial V_A^2} + \mu_A V_A \frac{\partial V_E}{\partial V_A} + \frac{\partial V_E}{\partial t}\right) dt + \sigma_A V_A \frac{\partial V_E}{\partial V_A} dW_t \quad (15)$$

By comparing the terms of equations (14) and (15) we can derive the following relationship:

$$\sigma_E V_E = \sigma_A V_A \frac{\partial V_E}{\partial V_A} \quad (16)$$

Following the theory of the Black-Scholes-Merton model, the term  $\frac{\partial V_E}{\partial V_A}$  corresponds to one of the so-called Greeks of a European option, in particular it corresponds to the sensitivity measure defined as Delta ( $\Delta^E$ ), and it is equal to:  $\phi(d_1)$ .

At this point, to obtain the unobserved values of  $V_A$  and  $\sigma_A$ , the following non-linear system of two equations and two unknowns has to be solved:

$$\begin{cases} f_1(V_E, \sigma_E) = V_A \phi(d_1) - \exp(-r(T-t)) D \phi(d_2) - V_E = 0 \\ f_2(V_E, \sigma_E) = \frac{V_A}{V_E} \phi(d_1) \sigma_A - \sigma_E = 0 \end{cases} \quad (17)$$

With  $\frac{\partial f_1}{\partial V_A} = \phi(d_1) > 0$

This is because, just like the Delta in the BSM model framework,  $f_1$  is an increasing function of  $V_A$  and this implies that  $f_1(V_A)$  only has one solution. For the same reason,  $f_2$  also has a unique solution.

Once the previous non-linear system has been solved numerically with Matlab, all the necessary data to calculate the default probability using  $\phi(-d_2)$  are available:

$$DD = d_2 = d_1 - \sigma_A \sqrt{T} = \frac{\ln\left(\frac{V_A}{D}\right) + \left(r - \frac{1}{2}\sigma_A^2\right)(T)}{\sigma_A \sqrt{T}} \quad (18)$$

The term  $d_2$  within this model is often indicated with the term Distance to Default (DD).

Considering now that the assets follow a Brownian geometric motion and, consequently,  $V_A(t)$  is distributed as a log-normal with an expected value at time  $t$  equal to:

$$V_A(t) = V_A \exp\left\{\left(r - \frac{1}{2}\sigma_A^2\right)(T-t) + \sigma_A W_{T-t}\right\} \quad (19)$$

We can state that the default probability,  $PD(T-t)$ , for  $t=0$  can be calculated as follows (Löffler and Posch, 2011):

$$\begin{aligned} PD(T) &= \Pr[V_A(T) < D] = \Pr\left[V_A \exp\left\{\left(r - \frac{1}{2}\sigma_A^2\right)T + \sigma_A W_T\right\} < D\right] = (20) \\ &= \Pr\left[W_T < \frac{\ln\left(\frac{D}{V_A}\right) - \left(r - \frac{\sigma_A^2}{2}T\right)}{\sigma_A}\right] = \Pr\left[Z < \frac{\ln\left(\frac{D}{V_A}\right) - \left(r - \frac{\sigma_A^2}{2}T\right)}{\sigma_A \sqrt{T}}\right] = \\ &= \Pr\left[Z < -\frac{\ln\left(\frac{V_A}{D}\right) + \left(r - \frac{\sigma_A^2}{2}T\right)}{\sigma_A \sqrt{T}}\right] = \Pr[Z < -DD] = \phi(-DD) \end{aligned}$$

#### 4.1) Procedural example

Two operational examples of default probability calculation will be conducted using the KMV method. The first consists in calculating the measure of the default probability at time 0, that is taking as reference data those of the latest available financial statements of one of the main Italian energy companies, which is 31<sup>st</sup> December 2020, at the date of preparation of this paper.

Once all necessary data have been found from Bloomberg®, the issuer's default probability is calculated following the procedure illustrated in the previous section.

The second operational example consists in the calculation of the prospective default probability using different methodologies for forecasting data and for modeling the other parameters required for the calculation of the relevant figures.

The equity value was calculated by multiplying the stock price as of 31<sup>st</sup> December 2020 by the number of company shares on the same date, obtaining the data from Bloomberg®. The share price was 8.548, the number of shares in the company was 3,572,550,000 and, consequently, the equity value was equal to:  $V_E = 8.548 \times 3,572,550,000 = 30,538,157,400$  euros.

As regards the amount of debt, it is always obtained from the balance sheet and here it is equal to 31,704,000,000.

The term structure of the risk-free rate with tenor equal to 6 months (Euribor 6M) was used as the value for the risk-free rate and the value of the zero rate at one year is equal to -0.533%.

The last parameter required in order to set up the system of equations (17) is the volatility of the equity. This was estimated at 53.58% using the close-to-close method based on the daily returns of the share recorded in the past year and annualized with a factor of 252, i.e., the number of trading days within the considered year. (Haug, 2010).

Table 6 displays a summary of the results obtained

Parameters	VE	D	sigma E	r (RISK FREE RATE)	STOCK PRICE	NUM. OUTSTANDING	FIRM VALUE (VA)	ASSET SIGMA (SIGMA A)
Values in Millions of Euro	30,538.16	31,704.00	53.58%	-0.533%	8.548	3,572.55	62,385.93	26.32%
Default Probability	0.00775627%							

Table 6 Estimation of the current default probability using the KMV method

As a second step, the objective is to estimate the prospective default probability. In order to obtain this figure, we need to investigate which is the most reasonable technique to use to determine the behavior of the input data of the one-year model.

Regarding the value of the equity, the number of shares was assumed to be constant during the following year and the prediction of the share price was obtained through NAR. In order to project the value of the debt into the relevant time span, we assumed that it follows a Log-Normal distribution. According to this assumption, the percentage change in the value of the debt over a very short time span is normally distributed. Defining as:

$\mu_D$  the expected value of returns over a year

$\sigma_D$  the volatility of the share price over a year.

The mean and the standard deviation over a time period of  $\Delta t$  are approximately  $\mu_D \Delta t$  and  $\sigma_D \sqrt{\Delta t}$ . Thus, we can state that:

$$\frac{\Delta D}{D} \sim \phi(\mu_D \Delta t, \sigma_D^2 \Delta t) \quad (21)$$

Where:  $\Delta D$  is the change in the value of debt over a period of time  $\Delta t$ , and  $\phi(m, v)$  denotes a normal distribution with mean equal to  $m$  and variance equal to  $v$ .

The BSM model implies that:

$$\begin{aligned} \ln D_T - \ln D_0 &\sim \phi\left[\left(\mu_D - \frac{\sigma_D^2}{2}\right)T, \sigma_D^2 T\right] \\ \ln \frac{D_T}{D_0} &\sim \phi\left[\left(\mu_D - \frac{\sigma_D^2}{2}\right)T, \sigma_D^2 T\right] \\ \ln D_T &\sim \phi\left[\ln D_0 + \left(\mu_D - \frac{\sigma_D^2}{2}\right)T, \sigma_D^2 T\right] \quad (22) \end{aligned}$$

Where  $D_T$  is the value of the debt at a future time  $T$  and  $D_0$  is the value of the debt at time 0. From this, we can conclude that the variable  $\ln D_T$  is normally distributed and consequently,  $D_T$  is log-normally distributed. The mean of  $\ln D_T$  is:  $\ln D_0 + \left(\mu_D - \frac{\sigma_D^2}{2}\right)T$  and the standard deviation is:  $\sigma_D \sqrt{T}$ .

Three different scenarios were constructed: the best scenario with a confidence level of 50%, the average scenario with a confidence level of 75% and the worst scenario with a confidence level of 99%. Table 7 shows different scenarios constructed for the forecasted estimation of the debt.

CONFIDENCE LEVEL	Z	DEBT (LOWER BOUND)	DEBT (UPPER BOUND)
50.00%	0.0100	32629.41	32695.94
60.00%	0.2500	31841.50	33504.99
75.00%	0.6745	30494.18	34985.33
95.00%	1.9600	26752.03	39879.18
99.00%	2.5760	25125.22	42461.28

Table 7 Debt simulation scenarios

To obtain a prospective estimate of the capital volatility ( $\sigma_E$ ), we proceeded using a GARCH (1,1) to model such value. However, before proceeding with the use of this econometric model, a statistical test was preliminarily used, the Ljung-Box Q-Test, to verify the presence of heteroskedasticity in the relevant historical series, i.e. the time series of the stock prices. (Tsay, 2010).

The Ljung-Box Q-Test (Ljung and Box, 1978) is defined as follows

$H_0$ : the data are independently distributed (i.e., the correlations,  $\rho$  in the population from which the sample is taken are 0, so that any observed correlations in the data result from randomness of the sampling process). In mathematical terms  $\rho_1 = \rho_2 = \dots = \rho_n = 0$

$H_a$ : the data are not independently distributed; they exhibit serial correlation.

The test statistics is:

$$Q = N(N+2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{N-k} \quad (23)$$

Where  $N$  is the sample size,  $\hat{\rho}_k$  is the sample autocorrelation at lag  $k$  and  $h$  is the number of lags being tested. Under  $H_0$  the statistic  $Q$  asymptotically follows a  $\chi^2_{(h)}$

The econometric test applied on the residuals of the time series indicates that the evidence is insufficient to reject the null hypothesis of no residual autocorrelation through 20 lags, consequently it is reasonable to use a GARCH type model.



The GARCH (Generalized AutoRegressive Conditional Heteroschedasticity) is a model of generalized auto regressive conditioned heteroskedasticity presented by (Bollerslev, 1986). The more general expression of the GARCH(p,q) model however evaluates the  $\sigma_n^2$  starting from the  $p$  observations of  $u^2$  and from the most current estimates of the variance rate  $q$ . The GARCH (p, q) formula can be generalized as follows:

$$\sigma_n^2 = \gamma V_L + \sum_{i=1}^q \alpha_i u_{n-i}^2 + \sum_{i=1}^p \beta_i \sigma_{n-i}^2 \quad (24)$$

It should be emphasized that, by setting  $p = 0$ , the result obtained is the expression of the ARCH (q). By setting  $\omega = V_L \gamma$  in the GARCH equation, the estimation model can be rewritten as follows:

$$\sigma_n^2 = \omega + \sum_{i=1}^q \alpha_i u_{n-i}^2 + \sum_{i=1}^p \beta_i \sigma_{n-i}^2 \quad (25)$$

The above expression is generally used in order to estimate the parameters since, once  $\omega$ ,  $\alpha$  and  $\beta$  are known, we can calculate  $\gamma$  as a difference:  $\gamma = 1 - \alpha - \beta$ . The Long-Term Variance, on the other hand, is equal to the ratio:  $\frac{\omega}{\gamma}$ .

The equation representing the GARCH (1,1) is the following:

$$\sigma_n^2 = \gamma V_L + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2 \quad (26)$$

Where  $V_L$  is the long-term variance.

A fundamental condition is given by the fact that the sum of the weights must be equal to one:  $\alpha + \beta + \gamma = 1$ . The parameters (1,1) of GARCH (1,1) indicate that the calculation of  $\sigma_n^2$  is focused on the most recent observation of the  $u^2$  and on the closest estimate of the variance rate. It is required that  $\alpha + \beta < 1$  since, otherwise, the weight assigned to the Long-Term Variance would become negative. The model described here recognizes that, over time, the variance tends to converge towards a long-term average level ( $V_L$ ) with an associated weight equal to  $\gamma = 1 - \alpha - \beta$ . The final stage, for the realization of the GARCH (1,1) model, is the estimation of the required parameters starting from the historical price series. The most common method for estimating them is that of the maximum likelihood estimation, MLE (Maximum Likelihood Estimation). The first step for the estimation of the model parameters consists of defining the estimated variance for day  $i$  as:  $v_i = \sigma^2$  and assume that  $u_i$  follows a normal conditional probability distribution. The maximum likelihood function ( $L$ ) to be maximized with respect to the model parameters is given by (Francq and Zakoian, 2010):

$$L = \prod_{i=1}^m \frac{1}{\sqrt{2\pi v_i}} \exp\left(-\frac{u_i^2}{2v_i}\right)$$

Applying the logarithm to the previous equation, the maximum points of the function do not change, but the calculations are simplified:

$$L = \sum_{i=1}^m \left[ -\ln(v_i) - \frac{u_i^2}{v_i} \right] = \sum_{i=1}^m L_i \quad (27)$$

The parameters that allow the maximization of  $L$  were found using a gradient descent algorithm (Byrd *et al.*, 1994) programmed in Matlab around the optimum zone (Figure 7).

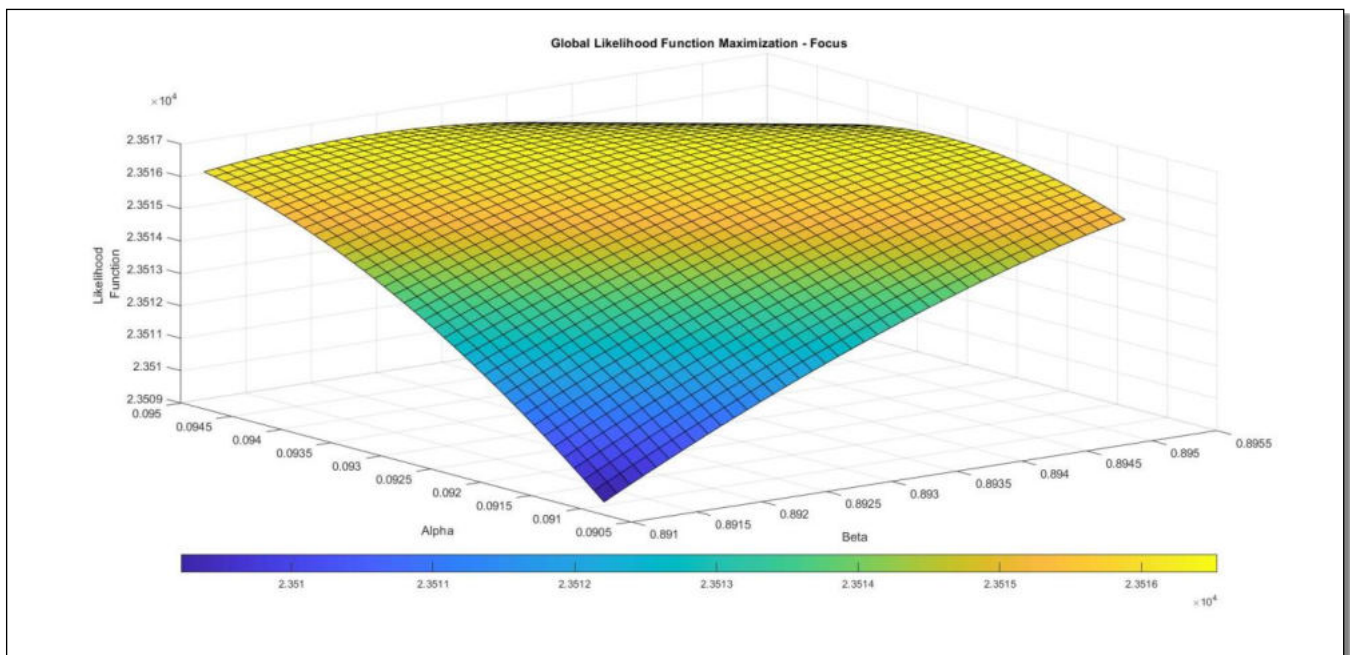


Figure 7 The detail of the surface of the Log-Likelihood Function around the optimum zone

Figure 8 shows the term structure of the volatility calculated from these, in accordance with the equation (Hull, 2015):

$$\sigma(T) = \sqrt{252 \left( V_L + \frac{1 - \exp(-aT)}{aT} [V(0) - V_L] \right)} \quad (28)$$

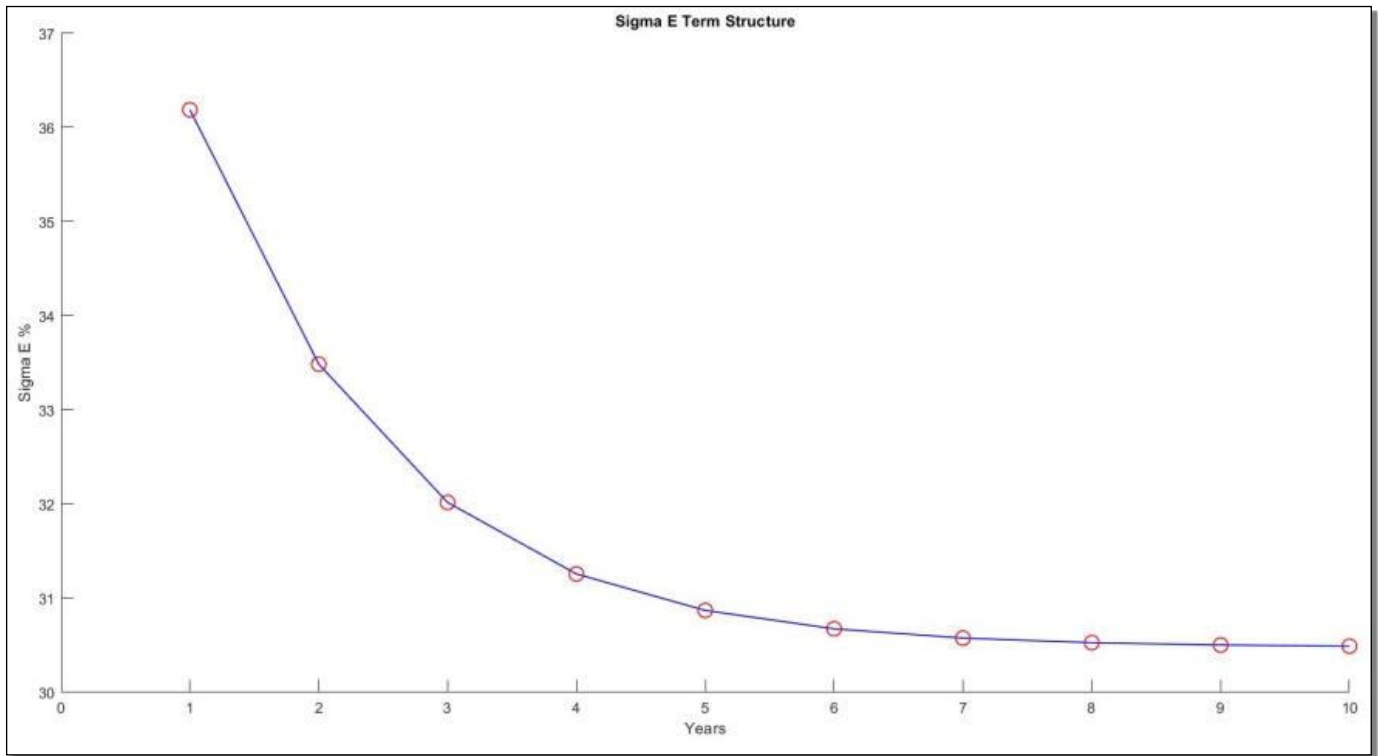


Figure 8 The term structure of the volatility estimated by GARCH (1,1)

As regards the rates used for the calculation of the prospective default probability, they were chosen by calculating the forward rate between 1 year and 2 years ( $F_{1Y,2Y}$ ) and the zero rates ( $r_{0Y,1Y}$  and  $r_{0Y,2Y}$ ) implied by the term structure of the 6-month Euribor at the date of analysis:

$$F_{1Y,2Y} = \frac{(1+r_{0Y,2Y})^2}{(1+r_{0Y,1Y})} - 1 = \frac{(1-0.00524)^2}{(1-0.00533)} - 1 = -0.515\% \quad (29)$$

Once all the future estimates of the parameters required for the application of the KMV model were obtained, they were used for the estimate of the prospective default probability following the same procedure.

Table 8 shows the results obtained in calculating the relevant figure for the three different debt scenarios previously described.

TIME 1 VALUES (31/12/2022) - Values in Millions of Euro			
Parameters	Best Scenario	Average Scenario	Worst Scenario
VE	30,023.71		
D	32,695.94	34,985.33	42,461.28
sigma E	36.19%		
r (RISK FREE RATE)	-0.515%		
STOCK PRICE	8.404		
NUM. OUTSTANDING	3,572.55		
FIRM VALUE (VA)	62,888.24	65,189.41	72,707.77
ASSET SIGMA (SIGMA A)	17.28%	16.67%	14.95%
Default Probability	0.01215340%	0.01476090%	0.02425420%

Table 8 Forecasted default probability using the KMV method

## 5) Applications of the previous methodologies on the companies included in the EuroStoxx 50

The goal of this section is to apply the analyzed methodologies to the companies making up one of the most famous European indices: the Euro Stoxx 50 (SX5E Index) in order to calculate, for each of them, the default probability at time 0 and the same figure as a forecasted value.

For each of the companies included in the reference index, the more appropriate forecasting technique has been selected.

For the calculation of the Probability of Default, the hierarchical principle previously discussed in this report was followed, namely: in the presence of Credit Default Swaps listed on the market, the method used will be that of the CDS; if there are no listed CDSs, but there are corporate bonds listed on the markets, the methodology used will be that of the Z-Spread and finally, in the absence of listed CDS or corporate bonds, the methodology will be that of the KMV.

Tables 9 and 10 show all 50 companies in the index with a progressive ID aimed at identifying them easily later and, in the last column, the forecasting technique that was used to obtain the Default Probability.

In particular, number 1 indicates that the methodology used will be that of the CDS, number 2 indicates that the methodology will be that of the Z-Spread and, finally, number 3 indicates that the methodology will be that of the KMV.

<b>SX5E INDEX</b>				
<b>FORECASTING TECHNIQUE: 1 (CDS METHOD) / 2 (Z-SPREAD METHOD) / 3 (KMV METHOD)</b>				
<b>ID</b>	<b>Ticker (Bloomberg)</b>	<b>Name</b>	<b>Stock Price (30/06/2021)</b>	<b>Forecasting Technique</b>
01	OR FP Equity	L'Oreal SA	375.80	3
02	DG FP Equity	Vinci SA	89.99	1
03	ASML NA Equity	ASML Holding NV	579.40	2
04	SAN SQ Equity	Banco Santander SA	3.22	1
05	PHIA NA Equity	Koninklijke Philips NV	41.79	1
06	TTE FP Equity	TotalEnergies SE	38.16	1
07	AI FP Equity	Air Liquide SA	147.66	1
08	CS FP Equity	AXA SA	21.39	1
09	BNP FP Equity	BNP Paribas SA	52.87	1
10	BN FP Equity	Danone SA	59.37	1
11	EL FP Equity	EssilorLuxottica SA	155.64	3
12	VIV FP Equity	Vivendi SE	28.33	1
13	MC FP Equity	LVMH Moet Hennessy Louis Vuitton SE	661.30	1
14	KER FP Equity	Kering SA	737.00	1
15	AMS SQ Equity	Amadeus IT Group SA	59.32	3
16	SAF FP Equity	Safran SA	116.92	3
17	AD NA Equity	Koninklijke Ahold Delhaize NV	25.07	1
18	IBE SQ Equity	Iberdrola SA	10.28	1
19	INGA NA Equity	ING Groep NV	11.14	1
20	LIN GY Equity	Linde PLC	243.35	2
21	PRX NA Equity	Prosus NV	82.47	3
22	ITX SQ Equity	Industria de Diseno Textil SA	29.71	3
23	KNEBV FH Equity	Kone Oyj	68.80	3
24	FLTR ID Equity	Flutter Entertainment PLC	152.70	3
25	ISP IM Equity	Intesa Sanpaolo SpA	2.33	1

*Table 9 Companies belonging to the Euro Stoxx 50 index at the analysis date - first part*

SX5E INDEX				
FORECASTING TECHNIQUE: 1 (CDS METHOD) / 2 (Z-SPREAD METHOD) / 3 (KMV METHOD)				
ID	Ticker (Bloomberg)	Name	Stock Price (30/06/2021)	Forecasting Technique
26	ENI IM Equity	Eni SpA	10.27	1
27	ENGI FP Equity	Engie SA	11.55	1
28	ABI BB Equity	Anheuser-Busch InBev SA/NV	60.81	1
29	ADYEN NA Equity	Adyen NV	2060.50	3
30	SAN FP Equity	Sanofi	88.36	1
31	ENEL IM Equity	Enel SpA	7.83	1
32	IFX GY Equity	Infineon Technologies AG	33.82	2
33	SU FP Equity	Schneider Electric SE	132.68	3
34	ALV GY Equity	Allianz SE	210.30	1
35	AIR FP Equity	Airbus SE	108.44	1
36	BAYN GY Equity	Bayer AG	51.21	1
37	BMW GY Equity	Bayerische Motoren Werke AG	89.31	1
38	CRH ID Equity	CRH PLC	42.50	2
39	BAS GY Equity	BASF SE	66.44	1
40	SIE GY Equity	Siemens AG	133.62	1
41	VOW3 GY Equity	Volkswagen AG	211.20	1
42	MUV2 GY Equity	Munich Re	230.95	1
43	SAP GY Equity	SAP SE	118.84	2
44	RI FP Equity	Pernod Ricard SA	187.20	1
45	ADS GY Equity	Adidas AG	313.90	3
46	DTE GY Equity	Deutsche Telekom AG	17.81	1
47	DPW GY Equity	Deutsche Post AG	57.36	1
48	DAI GY Equity	Daimler AG	75.30	1
49	DB1 GY Equity	Deutsche Boerse AG	147.20	2
50	VNA GY Equity	Vonovia SE	54.52	2

*Table 10 Companies belonging to the Euro Stoxx 50 index at the analysis date - second part*

In the following three sub-sections, the relevant figure, i.e. the Default Probability, will be calculated for each of the companies included in the index as of 30<sup>th</sup> June 2021.

### 5.1) Use of the CDS method

For each of the individual companies the Default Probability was calculated at time 0 and prospectively. To calculate the relevant forecasted figure, it was necessary, as seen before, to use artificial neural networks to forecast the CDS premiums.

NARs and the network training methodology already described were used. The only difference is the fact that the range in which to look for the best network architecture has been reduced. In the search for the best network, the number of Lags was varied from 1 to 15, the number of neurons in the first layer from 1 to 20, the number of neurons in the second layer from 0 to 20 and the third layer was not considered, thus leaving the number of neurons constant at 0.

The reason for this change in the search for the best network is the aim at reducing the computational time taken by the machine to select the best network.

Table 11 displays the result of applying the CDS method to the SX5E Index companies.

In particular, the following figures are shown: the Default Probability at time 0, the Forecasted Default Probability, and the architecture of the best network used for forecasting.

SX5E INDEX					
METHOD 1 - CDS METHOD					
ID	Ticker (Bloomberg)	Name	Time 0 Default Probability	Forecasted Default Probability	Best Network [Lag; 1st Layer; 2nd Layer; 3rd Layer]
02	DG FP Equity	Vinci SA	0.00141566	0.00181783	[12; 8; 11; 0]
04	SAN SQ Equity	Banco Santander SA	0.00167316	0.00071439	[15; 17; 6; 0]
05	PHIA NA Equity	Koninklijke Philips NV	0.00099950	0.00105345	[6; 13; 6; 0]
06	TTE FP Equity	TotalEnergies SE	0.00141566	0.00131872	[2; 9; 8; 0]
07	AI FP Equity	Air Liquide SA	0.00108275	0.00121097	[11; 8; 6; 0]
08	CS FP Equity	AXA SA	0.00167491	0.00161944	[2; 17; 17; 0]
09	BNP FP Equity	BNP Paribas SA	0.00230348	0.00281530	[4; 9; 0; 0]
10	BN FP Equity	Danone SA	0.00154048	0.00179698	[14; 12; 5; 0]
12	VIV FP Equity	Vivendi SE	0.00309137	0.00407886	[12; 3; 18; 0]
13	MC FP Equity	LVMH Moet Hennessy Louis Vuitton SE	0.00149472	0.00126877	[9; 9; 13; 0]
14	KER FP Equity	Kering SA	0.00140343	0.00400380	[14; 13; 4; 0]
17	AD NA Equity	Koninklijke Ahold Delhaize NV	0.00156544	0.00698189	[14; 8; 6; 0]
18	IBE SQ Equity	Iberdrola SA	0.00204525	0.00185069	[5; 18; 9; 0]
19	INGA NA Equity	ING Groep NV	0.00238968	0.00488713	[10; 20; 19; 0]
25	ISP IM Equity	Intesa Sanpaolo SpA	0.00456957	0.00452180	[5; 6; 12; 0]
26	ENI IM Equity	Eni SpA	0.00248554	0.00190261	[9; 12; 10; 0]
27	ENGI FP Equity	Engie SA	0.00221837	0.00325333	[7; 12; 7; 0]
28	ABI BB Equity	Anheuser-Busch InBev SA/NV	0.00388259	0.00240518	[10; 7; 19; 0]
30	SAN FP Equity	Sanofi	0.00106194	0.00277703	[12; 15; 12; 0]
31	ENEL IM Equity	Enel SpA	0.00218133	0.00243555	[13; 6; 4; 0]
34	ALV GY Equity	Allianz SE	0.00121505	0.00167887	[4; 8; 15; 0]
35	AIR FP Equity	Airbus SE	0.00288598	0.00137127	[3; 13; 17; 0]
36	BAYN GY Equity	Bayer AG	0.00262911	0.00249248	[9; 20; 16; 0]
37	BMW GY Equity	Bayerische Motoren Werke AG	0.00178762	0.00315165	[14; 14; 11; 0]
39	BAS GY Equity	BASF SE	0.00135703	0.00215369	[11; 2; 20; 0]
40	SIE GY Equity	Siemens AG	0.00166906	0.00146033	[6; 13; 5; 0]
41	VOW3 GY Equity	Volkswagen AG	0.00317865	0.00237291	[10; 20; 8; 0]
42	MUV2 GY Equity	Munich Re	0.00132675	0.00200233	[14; 13; 13; 0]
44	RI FP Equity	Pernod Ricard SA	0.00151779	0.00068621	[14; 19; 0; 0]
46	DTE GY Equity	Deutsche Telekom AG	0.00205281	0.00224608	[14; 17; 5; 0]
47	DPW GY Equity	Deutsche Post AG	0.00091625	0.00248906	[13; 9; 18; 0]
48	DAI GY Equity	Daimler AG	0.00237785	0.00044437	[8; 15; 18; 0]

Table 11 Application of the current and prospective version of the CDS method on the SX5E Index

## 5.2) Use of the Z-Spread method

The companies included in the index, for which Credit Default Swaps are not available, but listed corporate bonds are available, are considered here. In this case, as before, the Default Probability was also both calculated at time 0 and forecasted. In order to calculate the forecasted measure, in this operational example it was also necessary to carry out a forecasting of the Z-spreads using the NARs. The procedure followed to search for the best networks is completely similar to the one we already discussed. Table 12 shows the result of the application of the Z-Spread method on the index companies which do not have listed CDSs on the markets but do have listed corporate bonds both at time 0 and prospectively, as well as the architecture of the best network used to the different forecasts.

## 5.3) Use of the KMV method

Lastly, we considered the companies in the index which do not have either listed Credit Default Swaps or listed corporate bonds, therefore the available balance sheet data has been used for calculating the probability of default using the KMV method has been used. To calculate the forecasted figure, it was necessary to use the different methods described in section 4.1 to obtain the requested forecasted values. In particular, the neural networks were used for the forecasting of the different stock prices in order to calculate the forecasted equity value ( $V_E$ ), the Log-Normal distribution was used to estimate the different forecasted debt scenarios and finally a GARCH (1,1) was used to obtain the forecasted value of the capital volatility ( $\sigma_E$ ). Table 13 shows the result of applying the KMV method. In this case, the various Default Probabilities have not been included as the values obtained are extremely low, in line with the checks carried out with the DRSK risk assessment module available on Bloomberg®. The values highlighted in the table are instead the number of shares of the company, the value of the share (at time 0 and forecasted), the value of the equity (at time 0 and forecasted), the value of the debt at time 0 and the 3 different prospective scenarios, the value of the capital volatility ( $\sigma_E$ ) at time zero and the same forecasted value obtained with the GARCH (1,1) (also highlighting the parameters  $\alpha$ ,  $\beta$  and  $\omega$  of the GARCH), the value of the assets ( $V_A$ ) at time 0 and in the three envisaged scenarios and, finally, the value of the asset volatility ( $\sigma_A$ ) at time 0 and in the 3 different scenarios.



SX5E INDEX					
METHOD 2 - Z-SPREAD METHOD					
ID	Ticker (Bloomberg)	Name	Time 0 Default Probability	Forecasted Default Probability	Best Network [Lag; 1st Layer; 2nd Layer; 3rd Layer]
03	ASML NA Equity	ASML Holding NV	0.00177343	0.00592898	[15; 14; 6; 0]
20	LIN GY Equity	Linde PLC	0.00171860	0.00030275	[14; 9; 12; 0]
32	IFX GY Equity	Infineon Technologies AG	0.00363643	0.00187741	[13; 16; 9; 0]
38	CRH ID Equity	CRH PLC	0.00417472	0.00561251	[10; 12; 16; 0]
43	SAP GY Equity	SAP SE	0.00171348	0.00016373	[13; 7; 19; 0]
49	DB1 GY Equity	Deutsche Boerse AG	0.00157778	0.00053382	[15; 20; 3; 0]
50	VNA GY Equity	Vonovia SE	0.00358036	0.00822622	[11; 10; 3; 0]

Table 12 Application of the current and prospective version of the Z-spreads method on the SX5E Index

SX5E INDEX							
METHOD 3 - KMV METHOD (risk-free rate: - 0.432%)							
ID	Shares Outstanding	(Stock Price) [Forecasted S. P.]	(VE <sub>0</sub> ) [VE <sub>1</sub> ]	(D <sub>0</sub> ) [D <sub>1</sub> ] [Best; Medium; Worst]	(σ <sub>E0</sub> ) [σ <sub>E1</sub> ] [α; β; ω]	(VA <sub>0</sub> ) [VA <sub>1</sub> ] [Best; Medium; Worst]	(σ <sub>A0</sub> ) [σ <sub>A1</sub> ] [Best; Medium; Worst]
01	557.67	(375.80) [303.70]	(209,572.386) [169,363.13]	(2,451.60) [2,247.95; 4,148.29; 23,948.26]	(0.1804) [0.2428] {0.0745; 0.9105; 0.0000038}	(212,034.51) [171,620.81; 173,529.38; 192,415.07]	(0.1783) [0.2396; 0.2370; 0.2126]
11	441.75	(155.64) [142.21]	(68,753.97) [62,823.22]	(11,413.00) [14,917.95; 20,053.19; 46,756.78]	(0.1905) [0.2454] {0.0501; 0.9384; 0.0000029}	(80,216.38) [184,345.66; 189,503.14; 216,322.34]	(0.1633) [0.2231; 0.2170; 0.1901]
15	450.50	(59.32) [71.27]	(26,723.66) [32,105.78]	(5,783.00) [6,398.08; 7,333.09; 10,834.22]	(0.2557) [0.2986] {0.8558; 0.8949; 0.0000060}	(32,534.60) [38,531.62; 39,470.68; 42,987.00]	(0.2100) [0.2488; 0.2429; 0.2230]
16	426.26	(116.92) [112.67]	(49,838.32) [48,024.76]	(6,860.00) [7,486.4; 8,297.83; 11,139.12]	(0.3486) [0.3795] {0.0800; 0.9100; 0.0000059}	(56,731.47) [55,543.65; 56,538.59; 59,212.21]	(0.3062) [0.3281; 0.3234; 0.3078]
21	1616.29	(82.47) [87.36]	(133,295.44) [141,197.12]	(6,971.91) [6,579.74; 7,756.06; 12,417.90]	(0.2014) [0.1638] {0.1105; 0.7907; 0.0000096}	(140,297.53) [147,805.35; 148,986.76; 153,668.78]	(0.1913) [0.1565; 0.1552; 0.1505]
22	3114.86	(29.71) [28.63]	(92,542.49) [89,186.65]	(6,089.00) [2,037.47; 7,643.29; 336,008.76]	(0.2322) [0.2859] {0.0670; 0.9140; 0.0000061}	(98,657.85) [91,232.94; 96,863.03; 426,649.88]	(0.2178) [0.2795; 0.2632; 0.0598]
23	518.71	(68.80) [72.39]	(35,687.25) [37,550.80]	(559.20) [584.61; 725.66; 1,346.93]	(0.2122) [0.2626] {0.0630; 0.9141; 0.0000066}	(36,548.87) [38,137.94; 38,279.60; 38,903.56]	(0.2089) [0.2586; 0.2576; 0.2535]
24	174.63	(153.09) [121.78]	(26,734.11) [21,266.17]	(3,834.16) [3,295.26; 7,339.21; 72,573.75]	(0.3014) [0.3128] {0.1004; 0.8407; 0.0000023}	(30,586.8) [24,575.73; 28,637.22; 94,154.56]	(0.2634) [0.2707; 0.2323; 0.0707]
29	30.45	(2,060.50) [1,213.42]	(62,742.23) [36,948.74]	(128.36) [174.93; 243.55; 627.91]	(0.3928) [0.6365] {0.4246; 0.4919; 0.0001371}	(62,871.15) [37,124.43; 37,193.34; 37579.37]	(0.3920) [0.6335; 0.6323; 0.6258]
33	567.07	(132.68) [132.37]	(75,238.85) [75,065.79]	(11,399.00) [10,713.88; 11,507.08; 14,116.47]	(0.2538) [0.3038] {0.0753; 0.9135; 0.0000047}	(86,583.04) [85,826.15; 86,622.79; 89,243.50]	(0.2205) [0.2657; 0.2633; 0.2555]
45	195.07	(313.65) [309.42]	(61,183.71) [60,357.93]	(6,279.00) [6,384.44; 8,540.15; 19,633.66]	(0.3039) [0.2998] {0.0449; 0.9281; 0.0000096}	(67,491.81) [66,770.07; 68,935.13; 80,076.77]	(0.2460) [0.2710; 0.2625; 0.2260]
VALUE OF: SHARES OUTSTANDING - VE - D - VA IN MILLIONS OF EUROS							

Table 13 Application of the KMV method in current and prospective version on the SX5E Index

## 6) Conclusions

The present paper illustrates the application of three different methods for estimating the Default Probability as a measure of the counterparty risk: the CDS method, the Z-Spread method and the KMV method (Kealhofer, Merton and Vasicek).

All the above-mentioned methodologies have been validated both by following the reference scientific literature and by means of the calculation modules made available by one of the main information providers used by professionals in the sector, including the YAS, DRSK and CDSW modules of Bloomberg®.

As already extensively discussed, this figure was estimated both at the current level, using the market data observed at the time of valuation ( $t_0$ ), and as a forecasted measure, using simulations on the model's inputs to this end.

For the perspective values, we have conducted econometric tests (such as absence of autocorrelation errors), checked the goodness of fit on validation sets measured by  $R^2$ , conducted out-of-sample tests on different architectures created by varying the number of the main parameters of the NAR networks (neurons, layers, and lags).

In particular, the paper wants to focus on this specific aspect, providing a set of methodologies if the relevant figure is presented as a forecasted value and not in a current standpoint, in line with the growing attention paid by regulators on the so-called “Through the cycle” measurement.

The use of artificial neural networks plays therefore a primary role in the whole research: the forecasting obtained with the aid of dynamic ANNs proved to be robust in terms of econometric measures and allowed to obtain a robust estimate of the data required for the forecasting.

The above has proved valid, both for the practical examples carried out on one of the main gas and energy suppliers in our country, and for those conducted on the 50 companies included in the Euro Stoxx 50 (SX5E Index).

There is no lack of further insights on this matter. For example, the set of techniques illustrated could be applied to calculate the Default Probability of an issuer not belonging to the European area, but to the US area.

Another interesting potential application could be the construction of a committee machine which, by comparing the dynamic artificial neural networks with the traditional econometric methods aimed at forecasting, could choose, time after time, the most performing solution in terms of goodness of the estimate, thus improving the heterogeneity of forecasting methods (Bagnato *et al.*, 2021).

The authors want to thank the anonymous Referees whose suggestions helped improving this article.

## Bibliography

- [1] Agosto A., Giribone P. G. (2019), Artificial Intelligence in Risk Management: Machine Learning methods in application to market risk, AIFIRM (Italian Association of Financial Industry Risk Management) Conference, Milan, 11th April 2019
- [2] Agosto A., Moretto E. (2012), “Exploiting default probabilities in a structural model with nonconstant barrier”. *Applied Financial Economics*, 22(8), 667-679
- [3] Arbib A.M. (2003), “The Handbook of brain theory and neural networks”. The MIT press
- [4] Bagnato M., Bottasso A., Giribone P. G. (2021). “Implementation of a Commitment Machine for an Adaptive and Robust Expected Shortfall Estimation”. *Frontiers in Artificial Intelligence: A.I. in Finance* (<https://doi.org/10.3389/frai.2021.732805>)
- [5] Beale, M. H., M. T. Hagan, and H. B. Demuth (2014). *Neural Network Toolbox, User’s Guide*. The Mathworks Inc.
- [6] Bharath, S.T. and Shumway, T (2004). “Forecasting Default with the KMV-Merton Model”. Working Paper, The University of Michigan. ([https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=637342](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=637342))
- [7] Black F. and Scholes M. (1973). “The Pricing of Options and Corporate Liabilities”. *Journal of Political Economy*, 81, 637-654
- [8] Bollerslev T. (1986). “Generalized autoregressive conditional heteroscedasticity”. *Journal of Econometrics*, 31(3), 307-327
- [9] Bonini B., Caivano G., Cerchiello P., Giribone P. G. (2019) – “Intelligenza Artificiale: l'applicazione di Machine Learning e Predictive Analytics nel Risk Management” – XIV Position Paper AIFIRM (Associazione Italiana Financial Industry Risk Managers)
- [10] Bottasso A., Giribone P. G., Martorana M. (2019). “Analisi e progettazione di un sistema di misure quantitative per il monitoraggio dei rischi finanziari delle garanzie di origine”. *Risk Management Magazine*, 14(2), 27-45
- [11] Byrd R., Nocedal, J.; Schnabel, R. B. (1994). Representations of Quasi-Newton Matrices and their use in Limited Memory Methods”. *Mathematical Programming*, 63(4), 129–156
- [12] Caligaris, O., P. G. Giribone, and S. Ligato (2015). “Applicazioni delle reti neurali feed-forward per la ricostruzione di superfici di volatilità”. *AIFIRM Magazine* (Italian Association of Financial Industry Risk Managers), 10(1), 4-19
- [13] Chollet F. (2018), “Deep Learning with Python”. Manning Publications
- [14] Cun, Y. L. et al. (1990). “Handwritten digit recognition with a back-propagation network”. In: *Advances in Neural Information Processing Systems*, 396–404
- [15] Decherchi, C. and P. G. Giribone (2020). “Stima prospettica delle misure finanziarie di rischio mediante reti neurali dinamiche: un’applicazione al mercato statunitense”. *Risk Management Magazine*, 15(1), 50–69
- [16] de Simon-Martin M., Bracco S., Rosales-Asensio E., Piazza G., Delfino F., Giribone P. G. (2020). “Electricity Spot Prices Forecasting for MIBEL by using Deep Learning: a comparison between NAR, NARX and LSTM networks” – International Conference on Environment and Electrical Engineering - IEEEIC 2020, Technical Area: Regulation and Electricity Markets
- [17] Fonseca P. G. and Lopez H. D. (2017), “Calibration of Machine Learning Classifiers for Probability of Default Modelling”. *James Finance*, Crowd Process Inc. (<https://doi.org/10.48550/arXiv.1710.08901>)
- [18] Francq C. and Zakoian J. M. (2010). “GARCH Models: Structure, Statistical Inference and Financial Applications”. Wiley
- [19] Freeman J. A. and D. M. Skapura (1992), “*Neural Networks Algorithms, Applications, and Programming Techniques*”. Addison-Wesley, Massachusetts

- [20] Géron, A (2019). "Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow: Concepts, Tools and Techniques to Build Intelligent Systems". O'Reilly
- [21] Giribone P. G. (2020). "Seasonality Modeling through LSTM Network in Inflation-Indexed Swaps" – Data Analytics: The Ninth International Conference on Data Analytics - Special Session: FinTech Risk Management
- [22] Giribone, P. G., Ligato S., and Penone F. (2018). "Combining robust Dynamic Neural Networks with traditional indicators for generating mechanic trading signals", *International Journal of Financial Engineering*, 5(4). (<https://www.worldscientific.com/doi/abs/10.1142/S2424786318500378>)
- [23] Giribone P. G., Ligato S., Raviola P. (2014). "Studio ed implementazione della metodologia Credit Value Adjustment in un framework di pricing automatico" – ASSIOM Forex Newsletter, 25-36
- [24] Haug E. G. (2010). "The complete guide to Option Pricing Formulas". McGraw-Hill
- [25] Haykin, S. (1994). "Neural networks: a comprehensive foundation". Prentice Hall PTR
- [26] Hull J. C. (2015). Options, Futures and other Derivatives. Wiley
- [27] Kolen J. F. and C. S. Kremer (2001). "A field guide to dynamical recurrent networks". Wiley-IEEE Press
- [28] Krizhevsky, A., I. Sutskever, and G. E. Hinton (2012). "ImageNet Classification with Deep Convolutional Neural Networks". In: Proceedings of the 25th International Conference on Neural Information Processing Systems
- [29] Löffler G., Posch P. N. (2011). "Credit risk modeling using Excel and VBA", Wiley Finance
- [30] Ljung G. M. and Box G. E. P. (1978). "On a Measure of a Lack of Fit in Time Series Models". *Biometrika*. 65(2), 297-303
- [31] Merton R. C. (1974). "On the pricing of corporate debt: The risk structure of interest rates". *The Journal of Finance*, 29(2), 449-470
- [32] Principe, J. C. (2000). Artificial Neural Networks – The Electrical Engineering Handbook. CRC press
- [33] Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386–408
- [34] Rojas R. (1996), "Neural Networks – a systematic introduction", Springer Verlag
- [35] Rumelhart, D. E., G. E. Hinton, and R. J. Williams (1986). "Learning representations by back-propagating errors". In: *Nature*, 533–536
- [36] Tsay R. S. (2010). "Analysis of Financial Time Series". Wiley