Executive Summary
Hypothetical Market Data Scenario Generation Using Generative AI

Motivation: Historical data may not be enough to cover every scenario. Synthetic yet realistic data can help us do more robust risk management.

Goal: Creating novel yet realistic synthetic market scenarios for 3-Month TONA Futures contracts.

Generative AI and AI have revolutionized a wide range of industries in recent years. Briefly, generative AI is about learning patterns in data and recreating similar data with same statistical properties. There are patterns in nature, in the surrounding society, and also in financial data which describe how certain factors interact and change in relation to each other. In financial context, this can be co-movements of various instruments in the market. We can design AI/ML algorithms to learn these patterns and generate data that follows these patterns (Fig. 1). In technical language, we call it learning the data probability density function (PDF) and then sampling from it. Both learning the PDF and then sampling from this empirical PDF are non-trivial problems.

Neural network based models are powerful because they are flexible enough to capture complex patterns and non-linearities present in the real data. Non-linearity means that the sum of effects of constituent parts is different from the whole. Individual effects are not additive. Mathematically, a function f is non-linear if $f(x+y) \neq f(x) + f(y)$. The non-linear interactions between constituent parts can lead to synergistic effects completely different from simple linear interactions. Most of the real world data is non-linear and requires non-linear methods to be modelled properly.

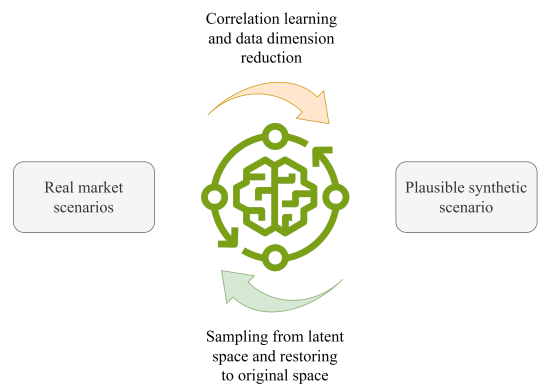


Fig. 1: Generative AI for market scenarios.

In this white paper we focus on modelling the co-movements of the first four 3-Month TONA Futures contracts which have expiries within 1 year. We use a neural network based generative model called "Variational Autoencoder" (VAE). Variational autoencoders non-linearly transform our complicated high-dimensional data and compress it into a low-dimensional latent space. Each data sample gets mapped to a conditional gaussian distribution in VAE latent space. This latent space has lower dimensionality than the data space and keeps the data samples near each other, leading to a smooth and continuous low dimensional representation of the data. Realistic new samples can be obtained either by simple interpolation between any two real data samples or by sampling in the empty regions of this latent space. After modelling the 3-Month TONA futures with a VAE, we generate scenarios by sampling VAE's latent space. We also use the classical Principal Component Analysis (PCA) to create new scenarios for comparison by using a similar methodology (Fig. 2).

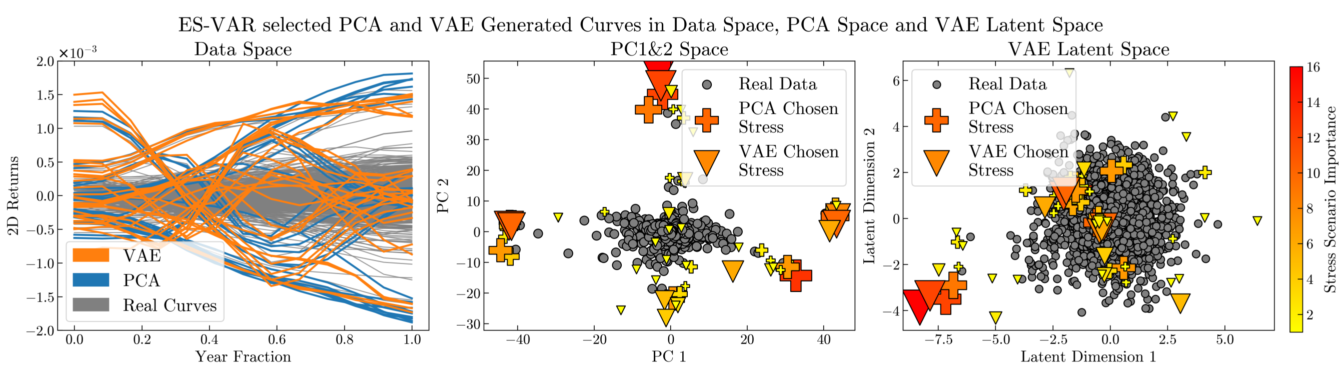


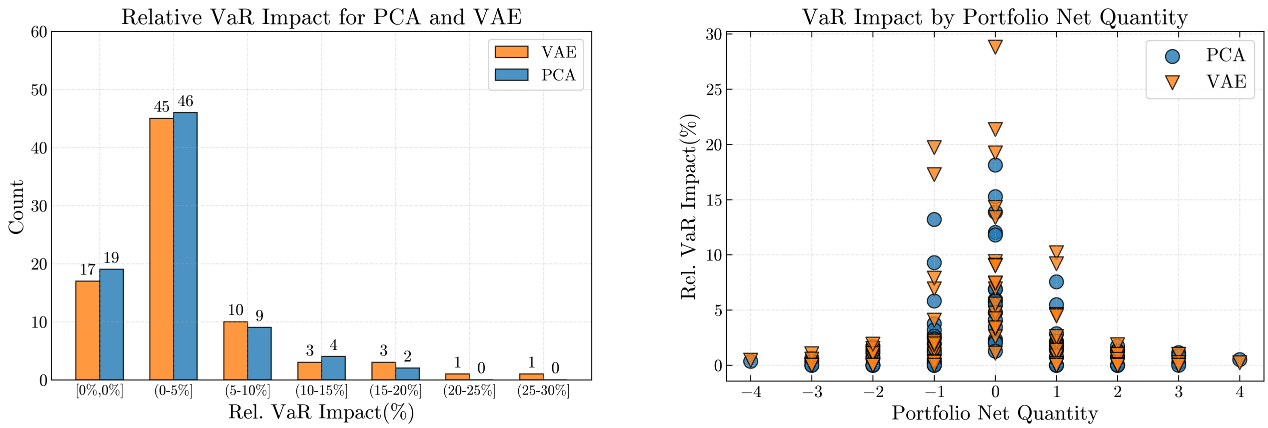
Fig. 2: ES-VaR chosen synthetic VAE and PCA generated curves by farthest point sampling in latent space. All the curves are represented both in PCA and VAE latent space in mid and right figure respectively. The heatmap reflects portfolio-specific stress scenario selection frequency during ES-VaR calculations.

To test the effect of synthetic scenarios on ES-VaR (Expected Shortfall), we construct a simple test by generating hypothetical portfolios with different combinations of 3-Month TONA Futures. Since the scope of the calculation is limited to four specific instruments, we can define a portfolio P_N as vector of quantities for each instrument expiry respectively:

$$P_N = (I_{N1},\ I_{N2}, I_{N3}, I_{N4}) \tag{1}$$

Where I_{N1}, I_{N2}, I_{N3}, I_{N4} are the net position quantities for the first, second, third and fourth closest-to-expiry instrument, respectively. For the purposes of this paper, we created hypothetical portfolios for all -1, 0 and 1 position, resulting in $3^4 - 1 = 80$ portfolio combinations when excluding the empty $(0, 0, 0, 0)$ portfolio.

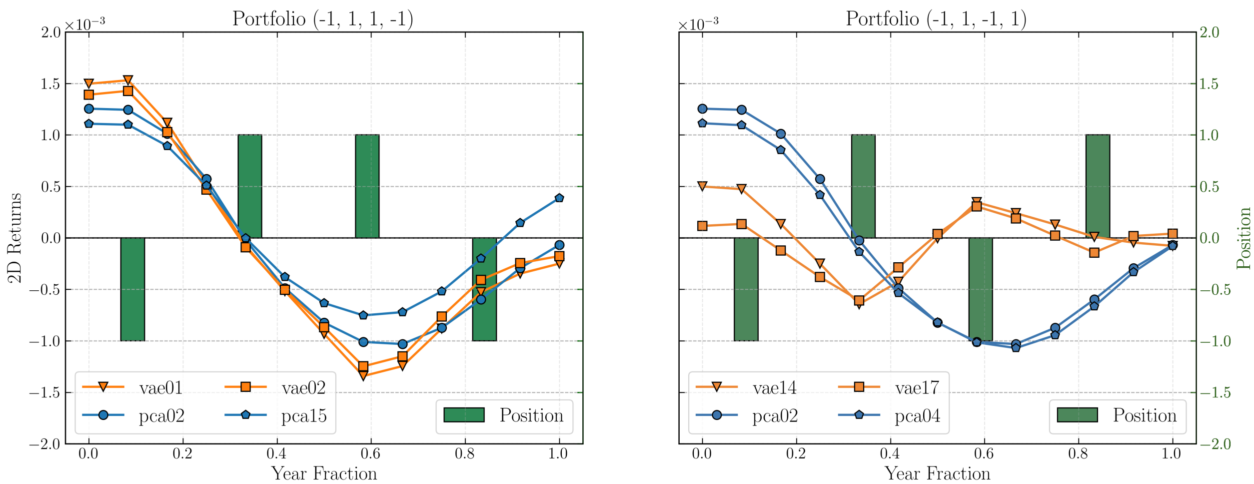
Both VAE and PCA were able to find new scenarios that resulted in a VaR impact in the 0% − 30% range as shown in Fig. 3. We made a similar comparison while also including the training data set of historical scenarios in both calculations, which resulted in VaR impacts in the 0%-10% range.



(a) Overall relative VaR impact for all portfolios.

(b) Relative VaR impact by portfolio net quantity.

Fig. 3: Fig. 3a shows the relative change in ES-VaR after adding VAE/PCA synthetic scenarios as hypothetical stress scenarios. Comparison of portfolio net quantity and VaR impact for both VAE and PCA calculations is shown in Fig. 3b.

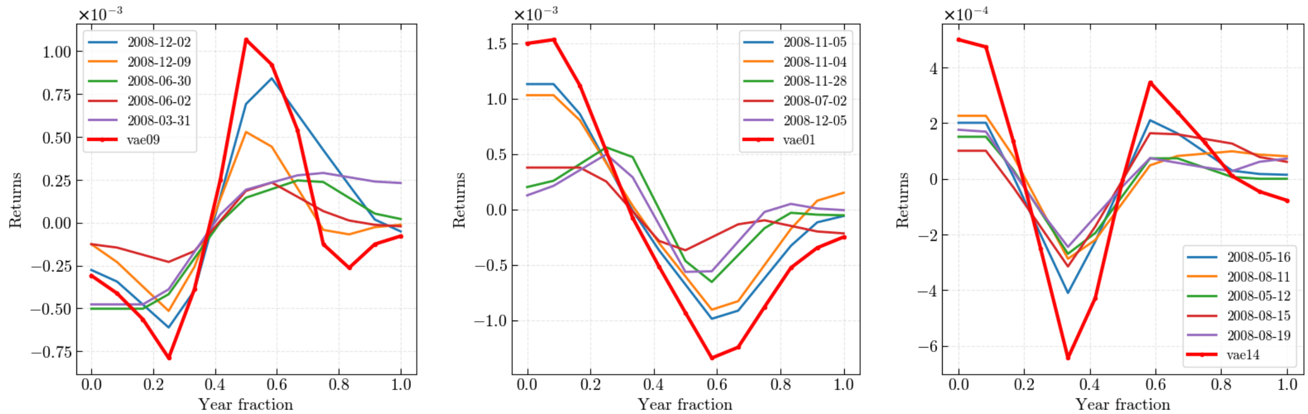


(a) Portfolio (-1, 1, 1, -1) with largest VaR impact.

(b) Portfolio (-1, 1, -1, 1) with different scenario shapes.

Fig. 4: ES-VaR chosen hypothetical scenarios for two portfolios with net quantity of 0 are shown. Portfolio (-1, 1, 1, -1) had the largest VaR impact of all portfolios for both VAE and PCA generated scenarios at 28.8% and 18.1% respectively. Fig. 4b shows the case where VAE and PCA generated scenario selected by ES-VaR have different shape. This portfolio had a VaR impact of 5.5% and 1.3% by VAE and PCA generated scenarios respectively.

When comparing VAE and PCA results, we noticed that VAE had a higher VaR impact (Fig. 3a) due to more extreme highs and lows (Fig. 4a) and having more curve shape variations (Fig. 4b). Both of these are desirable properties when looking for extreme scenarios that can improve our risk coverage. VAE allows us to create new data samples by taking into account local covariance structures in data distribution resulting in synthetic data with greater diversity while PCA only looks at global covariance and misses subtle local nuances in the data distribution. When further looking at the selected scenarios, we could see that they shared similar characteristics and shapes as some of the original training data though as shown in Fig. 5. This is by design but still worth mentioning – we



(a) Synthetic *vae09* sample. **(b)** Synthetic *vae01* sample. **(c)** Synthetic *vae14* sample.

Fig. 5: Selected synthetic VAE scenarios (thick red curve) compared to the closest training data neighbors (thin-colored).

should not expect our introduced sampling methodology to generate scenarios that are completely different from the training data. However, as both approaches were able to output new scenarios that resulted in higher losses than our original scenarios, we can argue that they help us to improve our risk coverage and risk understanding. Nonetheless, adding new synthetic scenarios to VaR should be done with great care to avoid too big or unwanted impact on the calculation output. Also worth mentioning is that within the scope of this paper, we tested a small set of hypothetical portfolios using a single calculation date.

Our original goal was to find a way to sample extreme but plausible scenarios. Plausibility is hard to define in the case of abstract financial data. The challenge, in the case of financial data, is that it is hard for someone to say what data is realistic or not just by looking at it apart from a highly trained expert. Therefore, we can only work with statistical measures and see if our generated data follows similar data distribution as the actual data or not. We see that our generated data does follow contours of the distribution and curves have similar shapes to data curves. Furthermore, many of these scenarios were selected by ES-VaR showing that they are extreme enough to create a significant change in portfolio's value.

However, with any technology there are some drawbacks, and cautions we need to take. Generative AI is based on complex models and architectures, which have complicated mathematics behind them. They are black boxes in two ways:

1. Due to their architectural complexity, their decision-making and predictions are hard to explain.
2. Mathematics governing them is not easily accessible.

This makes AI models powerful but also hard to understand at the same time which can make it difficult for users to fully understand the limitations. Another bottleneck to adoption is the lack of data. AI models require a large amount of high-quality data to be trained. Adding to this, there is no algorithm which can help us decide the best architecture, and various other hyperparameters to train neural network or even how long to train. Most often it is a combination of experimentation, intuition, and some problem specific details. There is indeed a lot of trial and error involved in training AI models.

The linearity of PCA makes it explainable, straightforward to implement and use. However, generating data using PCA transform requires careful sampling which is usually non-trivial. In short synthetic data generation by manually varying principal components does allow us to generate data but requires careful human expert's supervision. On the other hand, VAE converts complex high dimensional data distributions to simple and continuous latent distributions and has sound theoretical framework behind it ensuring its ability as a generative model. The simpler and lower-dimensional latent spaces enable better sampling and are easier to handle computationally due to their lower dimensionality.

In conclusion, we can say that both PCA and VAE-based generative approaches have their own merits and ideally can be used together to generate synthetic data. For example, we can remove the first and second principal components from data which are usually highly correlated with the market, and model the remaining complex dependencies in the data using VAE. Conditional models can also be built to predict the residual principal components' distribution based on the first and second principal components. Furthermore, combining these approaches provides enhanced explainability for VAE or other neural network-based generative models.