

Package CovRegpy: Regularized covariance regression and forecasting in Python

Cole van Jaarsveldt¹ , Gareth W. Peters² , Matthew Ames³ and Mike Chantler¹ 

¹School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, UK; ²Department of Statistics & Applied Probability, University of California, Santa Barbara, CA, USA; and ³ResilientML, Melbourne, Australia
Corresponding author: Cole van Jaarsveldt; Email: cv25@hw.ac.uk

(Received 22 February 2023; revised 18 March 2024; accepted 19 March 2024; first published online 13 May 2024)

Abstract

This paper will outline the functionality available in the **CovRegpy** package which was written for actuarial practitioners, wealth managers, fund managers, and portfolio analysts in the language of Python 3.11. The objective is to develop a new class of covariance regression factor models for covariance forecasting, along with a library of portfolio allocation tools that integrate with this new covariance forecasting framework. The novelty is in two stages: the type of covariance regression model and factor extractions used to construct the covariates used in the covariance regression, along with a powerful portfolio allocation framework for dynamic multi-period asset investment management.

The major contributions of package **CovRegpy** can be found on the GitHub repository for this library in the scripts: `CovRegpy.py`, `CovRegpy_DCC.py`, `CovRegpy_RPP.py`, `CovRegpy_SSA.py`, `CovRegpy_SSD.py`, and `CovRegpy_X11.py`. These six scripts contain implementations of software features including multivariate covariance time series models based on the regularized covariance regression (RCR) framework, dynamic conditional correlation (DCC) framework, risk premia parity (RPP) weighting functions, singular spectrum analysis (SSA), singular spectrum decomposition (SSD), and X11 decomposition framework, respectively.

These techniques can be used sequentially or independently with other techniques to extract implicit factors to use them as covariates in the RCR framework to forecast covariance and correlation structures and finally apply portfolio weighting strategies based on the portfolio risk measures based on forecasted covariance assumptions. Explicit financial factors can be used in the covariance regression framework, implicit factors can be used in the traditional explicit market factor setting, and RPP techniques with long/short equity weighting strategies can be used in traditional covariance assumption frameworks.

We examine, herein, two real-world case studies for actuarial practitioners. The first of these is a modification (demonstrating the regularization of covariance regression) of the original example from Hoff & Niu ((2012). *Statistica Sinica*, 22(2), 729–753) which modeled the covariance and correlative relationship that exists between forced expiratory volume (FEV) and age and FEV and height. We examine this within the context of making probabilistic predictions about mortality rates in patients with chronic obstructive pulmonary disease.

The second case study is a more complete example using this package wherein we present a funded and unfunded UK pension example. The decomposition algorithm isolates high-, mid-, and low-frequency structures from FTSE 100 constituents over 20 years. These are used to forecast the forthcoming quarter's covariance structure to weight the portfolio based on the RPP strategy. These fully funded pensions are compared against the performance of a fully unfunded pension using the FTSE 100 index performance as a proxy.

Keywords: portfolio optimization; regularized covariance regression (RCR); empirical mode decomposition (EMD); singular spectrum analysis (SSA); singular spectrum decomposition (SSD); X11; implicit factors; risk premia parity; risk parity; long short equity

1. Actuarial setting and context

Modern portfolio theory has seen significant developments in the twenty-first century in two separate, but related aspects of portfolio optimization originally proposed by Markowitz (1952):

- development of assumptions regarding statistical features of financial returns and their dynamic covariance–correlation relationships, and
- development of various objectives in pursuit of optimal portfolio weightings for positions in the portfolio based on assumptions made about future forecasts of trend and covariance of asset returns.

Multivariate dynamic covariance time series models can be split into various families of models such as multivariate generalized autoregressive conditional heteroskedasticity (MGARCH), dynamic conditional constant correlation (DCC), and various families of factor models, see a survey on such models in Bauwens *et al.* (2006). These models are generally all endogenous regression models, meaning that they are based on dynamics that arise from a past cross section of asset return residuals and white noise drivers to describe structurally the dynamic evolution of covariance and correlation in the cross section of asset returns. They have been widely adopted in many areas of finance, and there are reasonable software implementations for most MGARCH and DCC time series models in R, Python, and Matlab. Less developed families of multivariate dynamic volatility models for covariance forecasting are the families of covariance regression factor models. These are multivariate generalizations of classical factor models such as capital asset pricing model (CAPM), arbitrage free pricing theory (APT) models, and extended Fama-French 3- and 5-factor models, see a review in Mateus *et al.* (2019). Importantly from the perspective of the software package developed in **CovRegpy**, many of the new classes of dynamic factor models that can be used for covariance forecasting in multivariate asset return time series are not widely available in packages in R or Matlab. It is, therefore, the intention of this work to propose a new library of tools that can implement various new extended families of covariance regression factor models for covariance forecasting, where the emphasis on novelty lies in the choice of factors and new factor construction approaches which are based on nonlinear and nonstationary time series methods, not yet widely adopted in the wealth management space but are proving to be ushering in a new wave of multivariate asset covariance forecasting tools. This paper, therefore, seeks to introduce these tools for actuarial practitioners via a dedicated software library **CovRegpy**.

Before the recent developments of covariance regression methods such as those in Hoff & Niu (2012) and the applications in financial settings using these methods such as those in Ames *et al.* (2017) and (2018), the most common approach to covariance forecasting was using either empirical historical sample estimators as predictions based on the past realized portfolio asset returns covariance estimates (over windows of varying length) used as the best estimate for the forthcoming covariance. Alternatively, for model-based forecasting methods, it is common to utilize methods from multivariate time series such as DCC-MGARCH. It is widely understood that financial markets are often comprised of nonstationary series that exhibit complex dependence structures that vary depending on the current state of the economy, macroeconomic outlooks, and many other factors. DCC-MGARCH, while attempting to explicitly forecast the covariance, results in a square-root law temporarily growing covariance structure – as a result the weights will remain the same at any point in the future. That is, given the DCC-MGARCH forecast 1 day ahead, $\Sigma_{DCC}(1)$, the DCC-MGARCH forecast t days ahead will simply be:

$$\Sigma_{DCC}(t) = \sqrt{t} \times \Sigma_{DCC}(1). \quad (1)$$

Unlike in DCC-GARCH models, covariance regression frameworks such as those developed in Hoff & Niu (2012) can be considered in both non-time series and time series contexts. In general, this method will provide a regression technique for attributing covariance to underlying

factors in a parametric and interpretable manner through the use of a specific type of random-effects linear regression structure that produces a covariance regression model structure that the user can specify explicitly and test. In this work, the focus of the covariance regression structure will be on covariance forecasting in multivariate financial time series settings using a lagged effects model based on extracted factors that will act as implicit factor covariates when making covariance forecasts. Numerous regularizations of the original covariance regression of Hoff & Niu (2012) are presented herein, namely least absolute shrinkage and selection operator (LASSO) regression, ridge regression, elastic-net regression, group-LASSO regression, as well as subgradient descent for optimization solutions when estimating of such models in the **CovRegpy** package. These are all implemented within various speciality Expectation-Maximization (EM) estimation frameworks.

Furthermore, this **CovRegpy** package serves to promote the use of implicit factor extraction techniques and their integration into covariance regression. This is a newly emerging class of feature extraction methods proving useful in forecasting covariance of asset returns, though not yet widely adopted due to a lack of readily available software tools that allow the adoption of these methods by practitioners. We seek to fill this gap with the **CovRegpy** package.

Under these implicit factor covariance regression methods, the subsequent covariance forecasts can then be utilized in portfolio optimization methods built in the proposed **CovRegpy** package, and examples include the risk premia parity (RPP) as well as classical Markowitz mean-variance methods (see Markowitz, 1952). Portfolio optimization methods such as RPP (see Qian, 2005) are interesting modern portfolio allocation methods that complement traditional Markowitz mean-covariance methods. Portfolio weighting strategies such as RPP have become increasingly popular since the subprime mortgage crisis which significantly contributed to the wider-reaching 2007–2008 global financial crisis. In addition to RPP weighting, long short equity weighting rules, introduced in Jacobs & Levy (1993), restrict the cumulative long and short positions within the portfolio, replicating real-world shorting costs.

2. Software contributions and context

Despite many innovations in this space, there does not exist very many public-domain, state-of-the-art packages for the development of these methods for practitioners. The intention of this work is to develop a public domain freely available software toolbox for state-of-the-art dynamic portfolio methods that complement other existing packages such as *cvxPortfolio*, Boyd *et al.* (2016), which focuses on different aspects of portfolio optimization problem compared to the package proposed. The aforementioned package focuses on various dynamic portfolio settings with a wide variety of constraints that can be incorporated, whereas the proposed **CovRegpy** package in this manuscript focuses on building dynamic covariance regression models using advanced regression methods as well as state-of-the-art signal decomposition methods that work in nonstationary settings to extract factors and features to act, for instance, as regression factors.

The state-of-the-art nonstationary and nonlinear time series methods adopted in the package **CovRegpy** to produce the implicit factor extraction models have all been recently developed, with an exception being X11 which was developed in Shiskin *et al.* (1967). The empirical mode decomposition (EMD) used throughout this package is based on a parallel package established for actuarial data analytics in **AdvEMDpy** package which was developed in van Jaarsveldt *et al.* (2023a). EMD was originally developed in Huang *et al.* (1998), Huang (1999), and Huang *et al.* (1999). Singular spectrum analysis (SSA) developed in Hassani (2007) has been added to this package with additions like the Kolmogorov–Smirnov test of the distribution of the errors resulting in Kolmogorov–Smirnov SSA (KS-SSA). Decomposing SSA (D-SSA), decomposes the original algorithm and presents the individual components extracted rather than grouping them as in the original algorithm. Singular spectrum decomposition (SSD), which is a refinement of SSA, was

developed in Bonizzi *et al.* (2014) and has been added to this package. X11 is introduced in Shiskin *et al.* (1967) and further discussed in Sutcliffe (1993) and Doherty (2001). These separate methods included, and sometimes criticized, herein should not exclusively be thought of as competitors as, owing to their nonconstructive nature, and they function very well as complementary algorithms, see van Jaarsveldt *et al.* (2023b).

3. Overview and structure

The structure of this paper replicates the sequence in which the proposed techniques should be applied:

- extract factors using implicit factor extraction techniques,
- use l -lagged model or explicitly forecast implicit factor,
- use factors to forecast covariance using covariance regression, and
- weight portfolio according to risk appetite.

The following overview and structure is summarized with the associated packages and section in Fig. 1. In Section 4, the implicit factor extraction techniques are described with some descriptive figures. In this section, new approaches to these established techniques are discussed in order to increase flexibility for financial applications. These new approaches, such as KS-SSA, D-SSA, L2-SSD, L1-SSD, and ITA-SSD, seek to improve flexibility and solve specific problems encountered when applying the original methods to both synthetic and real-world data. EMD is briefly discussed in Section 4.2 before moving on to the implicit factor extraction techniques introduced in this package, namely X11, SSA, and SSD. In Section 4.3, X11 is discussed. In Sections 4.4 and 4.5, SSA and SSD are discussed, respectively. SSD was originally developed to address some of the shortcomings of SSA. Both of these techniques still have shortcomings as will be noted in Sections 4.4 and 4.5.

In Section 5, the first real-world actuarial case study, regularized covariance regression (RCR) is presented with its numerous available algorithmic variations. In Hoff & Niu (2012), the idea of independently estimating the mean and covariance coefficients is discussed without it being formalized. This independent mean regularized covariance regression (IM-RCR) is formalized in Section 5.3 and in Algorithm 1. This is followed by LASSO RCR, ridge RCR, elastic-net RCR, group-LASSO RCR, and subgradient optimization RCR in Sections 5.4, 5.5, 5.6, 5.7, and 5.8, respectively.

Section 5.1 notes the long-term financial benefits of insurers having more accurate knowledge of the prevalence of chronic obstructive pulmonary disease (COPD) and other early signs of reduced lung capacity, particularly after the SARS-CoV-2 pandemic, regarding policy pricing and the savings of early detection and indicators compared against expensive late-stage treatment and maintenance. Section 5.2 details works examining the diagnostic and mortality forecasting abilities of measures of lung capacity. In Hutchinson (1846) (which cites numerous early works), the diagnostic ability of reduced lung capacity was already shown to be valuable – this is motivated within Section 5 with the intention of pricing insurance policies and eventually constructing individual pricing.

Section 6 presents a more complete case study (using more features of this package) in which the implicit factor extraction techniques are used to isolate different frequency structures from the daily returns of FTSE 100 stocks. These different frequency structures are used to forecast the covariance of the stocks over the forthcoming investment horizon. With these covariance forecasts, the portfolios (funded portions of pensions) are weighted based on the RPP strategy. This uses implicit factor extraction techniques in RCR with RPP weighting. Section 6.1 summarizes some of the shortcomings of either using purely funded or unfunded pensions before the advantages of using hybrid pension schemes to address the demographic changes in developed

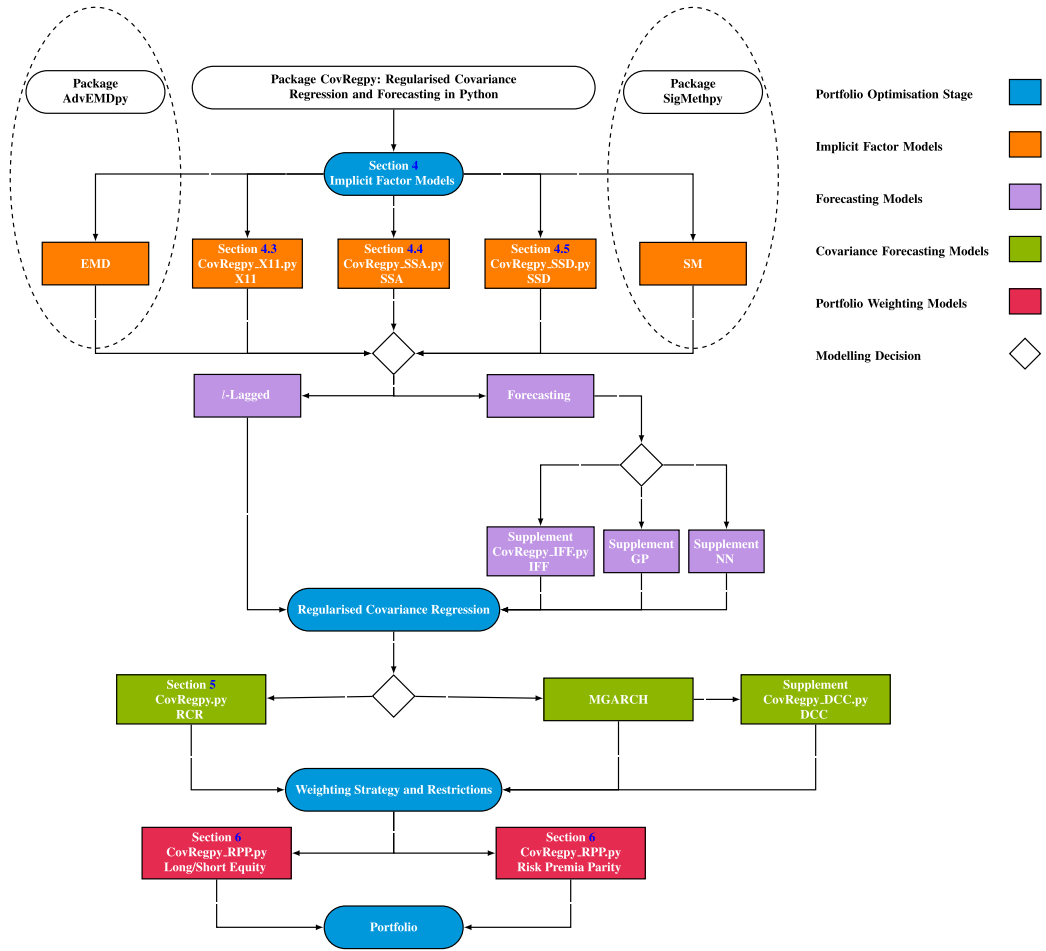


Figure 1. Flow diagram summarizing stages (and providing the respective sub-packages within *CovRegpy*) of constructing leveraged risk premia parity portfolios using implicit factor models in a regularized covariance regression framework with either an *l*-lagged framework or a formal forecasting framework.

nations are given exposition in Section 6.2. RPP portfolio weighting is described in Section 6.3 before the case study constructing hybrid pensions using the modeling framework from Dutta *et al.* (2000) with different covariance forecasts and different risk appetites is presented. This paper is concluded with some closing remarks in Section 7.

4. Implicit factor models

In Smith (1776), it was first postulated among supporting observations that assets have an intrinsic value about which their values fluctuate owing to human uncertainty. This formed the foundational ideas of asset pricing for nearly two centuries. In Sharpe (1964) and Lintner (1965), the CAPM was introduced and despite failing numerous tests, Fama & French (2004), this model did establish formal asset pricing models as a viable field of study with actual testable predictions concerning risk and return. This model is the first of what one can refer to and is referred to herein, as an explicit factor model. It is explicit in the sense that the factor upon which an asset’s price depends is observable or easily inferable.

Contemporaneously with this explicit factor model, Markowitz (1952) and Tobin (1958) developed modern portfolio theory as an explanation of the observable diversification in most if not all, real-world portfolios. This development considered both the returns of assets (above some “risk-free” asset as proposed in the CAPM) and the risk of assets (quantified as the covariation of the various assets). This was necessary and, arguably, vital to the further developments seen in modern economic theory. In Ross (1976), the Arbitrage Pricing Theorem (APT) was developed, which is widely considered the natural successor to the CAPM which has had increased success in its predictive capabilities, see Basu & Chawla (2012).

A further development in this field which arguably combines factor models with traditional portfolio theory is the usage of principal component analysis (PCA) in the construction of principal portfolios by applying the technique to either covariances or correlations of sets of assets – the details of the application thereto are outlined in Jolliffe (1986). Applications of PCA to the financial industry have varied with Darbyshire (2017) constructing interest rate derivative portfolios and Pasini (2017) applying the same framework in the construction of equity portfolios. In Yang (2015), this is further extended to the construction of portfolios of principal portfolios. These principal components are the first implicit factors in the sense that their interpretation is less obvious than CAPM and APT explicit factors.

This section outlines the techniques available in the recently published EMD package, AdvEMDpy, van Jaarsveldt *et al.* (2023a), and the techniques published in this package, CovRegpy. These techniques are in contrast to traditional financial factor models such as principal portfolios as a dimensionality reduction technique and an attributable variance model and the explicit market factor models proposed in Fama & French (1993) and (2015) which relates financial asset prices to several factors – one of which is the difference between the return of a diversified portfolio of small stocks and the return of a diversified portfolio of large stocks. Little research has been conducted into the interpretability of these implicit factors or relating them to easily observable economic factors with a notable exception being van Jaarsveldt *et al.* (2023b) where an implicit Carbon ETF factor is related to the annual cycle of carbon dioxide in the atmosphere.

The explicit factor techniques, while revolutionary, make no attempt to isolate specific time-frequency structures within the asset basket under observation. The techniques discussed in this section either implicitly (EMD and X11) or explicitly (SSA and SSD) isolate structures based on frequency bandwidths and can therefore, to take the model further, be used to construct horizon-specific portfolios based on one’s requirements and the bandwidths of the components. An additive synthetic time series has been generated to demonstrate the various available decomposition techniques in this package. Equation (2) is plotted in Fig. 2 demonstrating a synthetic time series with a noticeable trend and an annual seasonal structure with white noise. With $t \in \{0, 1, 2, \dots, 120\}$, the synthetic time series is

$$x_t = x_{t,\text{trend}} + x_{t,\text{seasonal}} + x_{t,\text{irregular}} \quad (2)$$

with

$$x_{t,\text{trend}} = \frac{1}{100} \times (t - 10) \times (t - 60) \times (t - 110) + 1000, \quad (3)$$

and

$$x_{t,\text{seasonal}} = 100 \times \sin\left(\frac{2\pi}{12}t\right), \quad (4)$$

and

$$x_{t,\text{irregular}} \sim \mathcal{N}(0, 10). \quad (5)$$

Section 4.2 demonstrates the results of the application of EMD to the time series described in Equation (2). We advise the reader to see van Jaarsveldt *et al.* (2023a) for a thorough review of this

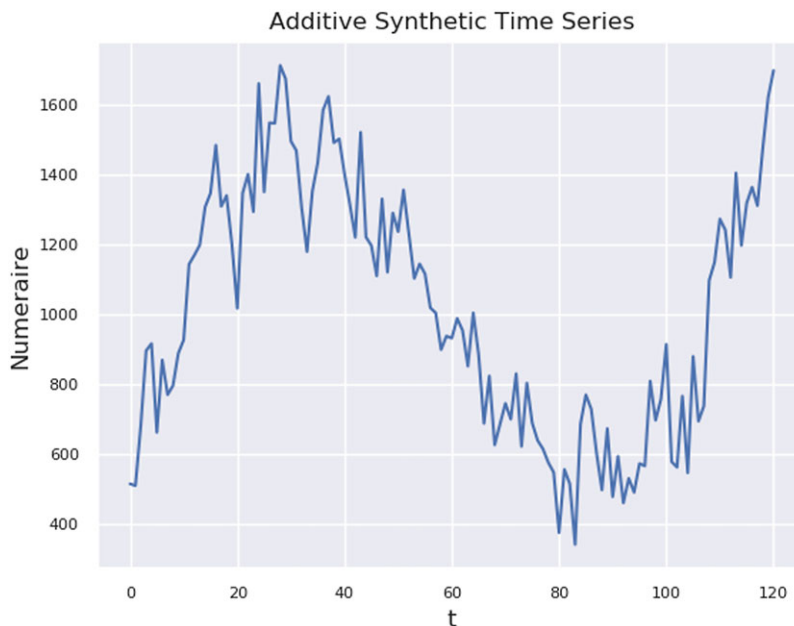


Figure 2. Additive synthetic time series described in Equations (2), (3), (4), and (5), used to demonstrate implicit factor extraction methods in this section which consists of a trend-cycle, seasonal, and irregular component.

algorithm, its algorithmic variations, and the associated software package which also may produce implicit market factors for use in this algorithm – see Fig. 1 for usage context. The application of X11 to this time series is reviewed in Section 4.3 – the results are less favorable than in Section 4.2, but the method is much older and has been shown in van Jaarsveldt *et al.* (2023b) to act favorably as a smoother of EMD to form EMD-X11. The base implementation of the SSA and SSD algorithms to the same synthetic time series is reviewed in Sections 4.4 and 4.5 – the algorithmic extensions to SSA and SSD developed to address some of the shortcomings discovered in this work are addressed in the supplement.

4.1 Additive and multiplicative time series

In Derman & Kani (1994), Dupire (1994), and Rubinstein (1994), it is postulated that the volatility of assets is not only time-inhomogeneous, but it is also a function of the prices of the assets themselves or, by extension, the recent changes in the asset prices – this is inline with modern probability default models that agree that the momentum of the underlying factors is nontrivial in their predictive capabilities. This follows logically as high prices or recent large increases in asset prices often foreshadow periods of heightened uncertainty and, therefore, volatility. These assumptions are not without criticism with Dumas *et al.* (1998), performing an extensive empirical study that concludes that these models are indistinguishable from smoothed versions of the implied volatility model proposed in Black & Scholes (1973).

With this context, it is natural to also explore multiplicative models where seasonality and trend components (and further less-easily named or interpretable components isolable using SSA and SSD within the CovRegpy package and EMD in the AdvEMDpy package) are multiplied onto some underlying trend. The framework in this paper is also well suited for multiplicative models, provided that one uses an appropriate transform such as the natural logarithm as demonstrated in Equation (6):

$$\begin{aligned}
 x_t &= x_{t,\text{trend}} \times x_{t,\text{seasonal}} \times x_{t,\text{irregular}} \\
 \log(x_t) &= \log(x_{t,\text{trend}}) + \log(x_{t,\text{seasonal}}) + \log(x_{t,\text{irregular}}).
 \end{aligned}
 \tag{6}$$

4.2 Empirical mode decomposition

Based on work done in Huang *et al.* (1998), Huang (1999), Huang *et al.* (1999), van Jaarsveldt *et al.* (2023a), and van Jaarsveldt *et al.* (2023b), it can be stated with some confidence that EMD and the various algorithmic variations summarized and presented in van Jaarsveldt *et al.* (2023a) present a class of robust trend estimation and decomposition algorithm that are state-of-the-art when working with general stochastic processes including potential nonstationarity and nonlinearity. The algorithm, Algorithm 1, presented in pseudocode in Appendix A, of Supplement to: Package CovRegpy: Regularized Covariance Regression and Forecasting in Python, was first presented in van Jaarsveldt *et al.* (2023a). The two conditions that each component isolated using EMD, named intrinsic mode functions (IMFs), must satisfy are listed in C1 and C2. An additional, optional (and highly recommended) condition, C3, is referred to as a stopping criterion and is checked after the initial two conditions are checked. This prevents over-sifting and the propagation of errors through successive IMF components with the conditions listed below:

- C1 $\text{abs}\left(\left|\left\{\frac{d\gamma_k(t)}{dt} = 0 : t \in (0, T)\right\}\right| - \left|\left\{\gamma_k(t) = 0 : t \in (0, T)\right\}\right|\right) \leq 1,$
- C2 $\sum_t \text{abs}(\tilde{\gamma}_k^\mu(t)) \leq \epsilon_2,$ and
- C3 $SD_{(p,q)} = \sum_t \left[\frac{|(h_{(p,q-1)}(t) - h_{(p,q)}(t))|^2}{h_{(p,q-1)}^2(t)} \right] < \epsilon_3.$

In C1, $\gamma_k(t)$ is the k^{th} IMF component and with $|\{\cdot\}|$ being the cardinality set. In C2, $\tilde{\gamma}_k^\mu(t)$ is the mean of the k^{th} IMF and in C3, $h_{(p,q)}(t)$ is the q^{th} iteration of the algorithm for the p^{th} IMF. The final result of the sifting presented in Algorithm 1 is

$$x_{\text{IMF}}(t) = \sum_{k=1}^K \gamma_k(t) + r_K(t).
 \tag{7}$$

As mentioned, EMD refers exclusively to the sifting procedure, whereas the Hilbert–Huang transform (HHT) refers to both EMD (the sifting procedure) followed by the Hilbert transform. The Hilbert transform of k^{th} IMF can be seen in Equation (8):

$$\check{\gamma}_k(t) = \mathcal{HT}[\gamma_k(t)] = \frac{1}{\pi} PV \int_{-\infty}^{\infty} \frac{\gamma_k(t^*)}{t - t^*} dt^*,
 \tag{8}$$

with PV being the Cauchy principle value integral. This is an improper integral that assigns a value to an improper integral where there exists a discontinuity. The Hilbert transform in Equation (8) can be expanded to Equation (9):

$$\check{\gamma}_k(t) = \frac{1}{\pi} \lim_{\epsilon \rightarrow 0^+} \left[\int_{t-\frac{1}{\epsilon}}^{t-\epsilon} \frac{\gamma_k(t^*)}{t - t^*} dt^* + \int_{t^*+\epsilon}^{t+\frac{1}{\epsilon}} \frac{\gamma_k(t^*)}{t - t^*} dt^* \right].
 \tag{9}$$

With the Hilbert transform of the time series defined as in Equation (9), the analytical signal can then be defined as in Equation (10):

$$\gamma_k^a(t) = \gamma_k(t) + i\check{\gamma}_k(t) = a(t)e^{i\theta(t)} = a(t)e^{i \int \omega(t) dt},
 \tag{10}$$

with the instantaneous amplitude being defined as in Equation (11):

$$a(t) = \sqrt{\gamma_k(t)^2 + \check{\gamma}_k(t)^2}, \tag{11}$$

and the phase being:

$$\theta(t) = \tan^{-1} \left(\frac{\check{\gamma}_k(t)}{\gamma_k(t)} \right). \tag{12}$$

The instantaneous frequency can then be defined as in Equation (13):

$$\omega(t) = \frac{d\theta(t)}{dt}. \tag{13}$$

4.3 X11

X11 was originally proposed in Shiskin *et al.* (1967). This algorithm is the focus of `CovRegpy_X11.py`. A disadvantage of X11 when compared against EMD is that it is less flexible in that it can only extract three structures namely the trend-cycle component, seasonal component, and the irregular component. Each of these components is individually less flexible than an IMF from EMD. The irregular component would, in most real-world scenarios, contain additional meaningful structural information that is regarded as noise in X11. Many shortcomings of the method are discussed in Sutcliffe (1993) and Doherty (2001).

The Henderson symmetric weights proposed in Henderson (1916), Whittaker (1922), and Henderson (1924) are used in the X11 algorithm. The edges of the time series present problems for the symmetric weighting procedure. Two common solutions are proposed. As in Dagum (1980), (1988), (1996), and Findley *et al.* (1998), the autoregressive integrated moving average (ARIMA) model is used to explicitly forecast the edge of the time series so that the symmetric weights can still be used. An alternative to this without explicitly forecasting the time series involves calculating a set of asymmetric weights. Classical asymmetric weights are calculated in Musgrave (1964b) and (1964a) with the Reproducing Kernel Hilbert Space derivation of these weights being done Bianconcini (2006), Dagum & Bianconcini (2006), and (2008). X11 decomposes the times series as:

$$x(t) = T(t) + S(t) + \epsilon(t), \tag{14}$$

with $T(t)$, $S(t)$, and $\epsilon(t)$ being the trend-cycle component, seasonal component, and irregular component, respectively. The symmetric moving average of order k , $ma_k(\cdot)$, is defined as:

$$ma_k(x(t_h)) = \frac{1}{2(k-1)} x\left(t_{h-\frac{(k-1)}{2}}\right) + \sum_{a=-\frac{(k-2)}{2}}^{\frac{(k-2)}{2}} \frac{1}{k-1} x(t_{h+a}) + \frac{1}{2(k-1)} x\left(t_{h+\frac{(k-1)}{2}}\right), \tag{15}$$

with $x(t_h)$ being the h^{th} time series component. The symmetric seasonal moving average of order $m \times n$, $S_{\{m \times n, c\}}(\cdot)$, is defined as:

$$\begin{aligned} S_{\{m \times n, c\}}(x(t_h)) &= \frac{1}{mn} x\left(t_{h-\frac{c((n-1)+(m-1))}{2}}\right) + \frac{2}{mn} \sum_{b=-\frac{c((n-1)+(m-1)-2)}{2}}^{-\frac{c(n-1)}{2}} x(t_{h+b}) \\ &+ \sum_{b=-\frac{c(n-3)}{2}}^{\frac{c(n-3)}{2}} \frac{1}{n} x(t_{h+b}) + \frac{2}{mn} \sum_{b=\frac{c(n-1)}{2}}^{\frac{c((n-1)+(m-1)-2)}{2}} x(t_{h+b}) \\ &+ \frac{1}{mn} x\left(t_{h+\frac{c((n-1)+(m-1))}{2}}\right), \end{aligned} \tag{16}$$

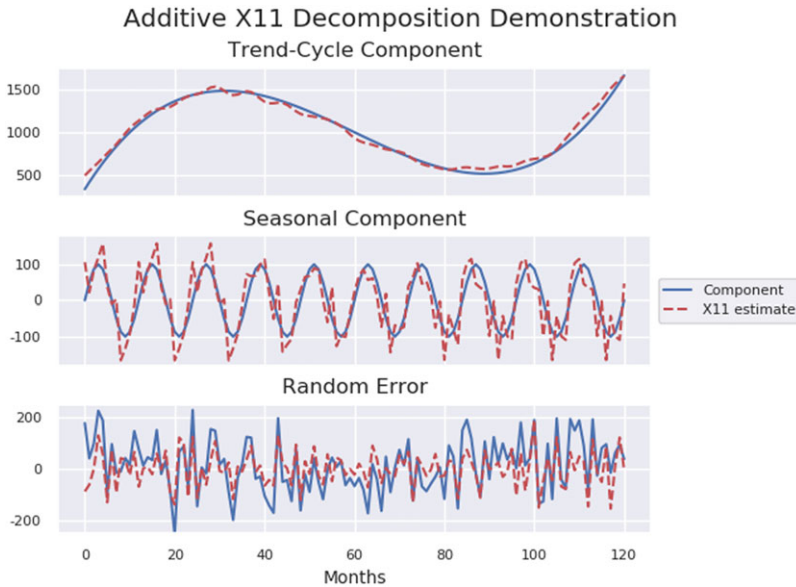


Figure 3. X11 time series decomposition of synthetic additive time series into trend-cycle component, seasonal component, and random error or noise component.

with c being a parameter that depends on both the sampling rate of the time series and the calculated seasonality of the time series. An example is that if the time series is sampled monthly and the seasonality is expected to be quarterly, then $c = 3$. The Henderson moving average of order l , $Hma_l(\cdot)$, is as in Henderson (1916), Whittaker (1922), Henderson (1916), Bianconcini (2006), Dagum & Bianconcini (2006), and Dagum & Bianconcini (2008). An example of the application of the code can be seen below:

```
trend, seasonal, irregular = \
    CovRegpy_X11(time, time_series, seasonality='annual',
                 seasonal_factor='3x3', trend_window_width_1=13,
                 trend_window_width_2=13, trend_window_width_3=13)
```

Fig. 3 demonstrates the result of applying the X11 method in `CovRegpy_X11` to the additive synthetic time series, Fig. 2, described in Section 4 using the parameters as shown in the code above. The trend-cycle component, seasonal component, and random error component are separated from one another, but the noise persists in the system owing to the lack of formal frequency bandwidth or distribution approaches to the decomposition as will be formally demonstrated in Section 4.5. The three components visible in Fig. 3 result from applying the above code snippet to the time series presented in Equation (2) and plotting each component with its corresponding underlying target time series component.

4.4 Singular spectrum analysis

SSA, as presented in Hassani (2007), can be separated into four stages, namely embedding, singular value decomposition (SVD), grouping, and diagonal averaging which will be presented in Sections 4.4.1, 4.4.2, 4.4.3, and 4.4.4, respectively. In Section 9.1.1 of Supplement to: Package `CovRegpy`: Regularized Covariance Regression and Forecasting in Python, D-SSA is introduced and in Section 9.1.2 of “Supplement to: Package `CovRegpy`: Regularized Covariance Regression

and Forecasting in Python,” KS-SSA is introduced. D-SSA and KS-SSA have been developed during this research to modify SSA to develop a more robust decomposition technique. D-SSA seeks to construct a decomposing version of SSA, whereas KS-SSA was developed to optimize and automate the isolation of a trend among white noise – the assumption underpinning this technique is that the errors have a Gaussian distribution.

4.4.1 Embedding in SSA

The first step involves converting the univariate time series into a multivariate time series by embedding lagged increments of the original univariate time series to create matrix \mathbf{X} such that:

$$\mathbf{X} = [\mathbf{X}_1(t), \dots, \mathbf{X}_K(t)], \quad (17)$$

with $\mathbf{X}_j(t) = [x(t_j), \dots, x(t_{j+L-1})]^T$ and $K = T - L + 1$. L is the embedding dimension that determines the structures being isolated and becomes increasingly relevant in the algorithmic variations and methodological extension of SSA, namely SSD (discussed in the next section), for purposes such as removing the initial trend structure.

4.4.2 Singular value decomposition in SSA

This is followed by the SVD of the matrix \mathbf{X} resulting in $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_L)$ and $\mathbf{U} = [\mathbf{U}_1, \dots, \mathbf{U}_L]$ being the vector of eigenvalues in descending order ($\lambda_1 < \dots < \lambda_L$) and their corresponding eigenvectors, respectively. With $\mathbf{V}_j = \mathbf{X}^T \mathbf{U}_j / \sqrt{\lambda_j}$, \mathbf{X} is decomposed as:

$$\mathbf{X} = \sum_{j=1}^L \mathbf{X}_j, \quad (18)$$

where $\mathbf{X}_j = \sqrt{\lambda_j} \mathbf{U}_j \mathbf{V}_j^T$. Equation (18) demonstrates the decomposition of the embedding matrix in Equation (17). In the absence of rounding errors, Equation (18) should precisely reproduce the time series after a reversal of the embedding step.

4.4.3 Grouping in SSA

With I being a subset of indices such that $I = \{i_1, \dots, i_p\} \subset \{1, \dots, L\} = L$, the trend is estimated using a user-selected subset of L , but in practice to estimate the trend, owing to the exponential decay of the eigenvalue, only the first few indices are used:

$$\mathbf{X}_I = \sum_I \mathbf{X}_j. \quad (19)$$

Theoretically, any subset in $\mathcal{P}(L)$ can be used to estimate the trend, but as the eigenvalues monotonically decrease, the amount of variation, and therefore information, in each subsequent \mathbf{X}_j decreases exponentially. It is this step that is adjusted in Section 9.1.1 of “Supplement to: Package CovRegpy: Regularized Covariance Regression and Forecasting in Python” to produce the D-SSA by keeping the components separate before performing the following step on individual components before grouping them based on the user’s requirements such as a pure trend estimate or an evaluation of different cyclical components.

4.4.4 Diagonal averaging in SSA

This final step is, in summary, a reversal of the embedding step. By appropriately defining a vector, such that the lagged embedding of Equation (17) is reversed so that an appropriate averaging can take place. With this vector, $X_I[i]$, defined as in Equation (20):

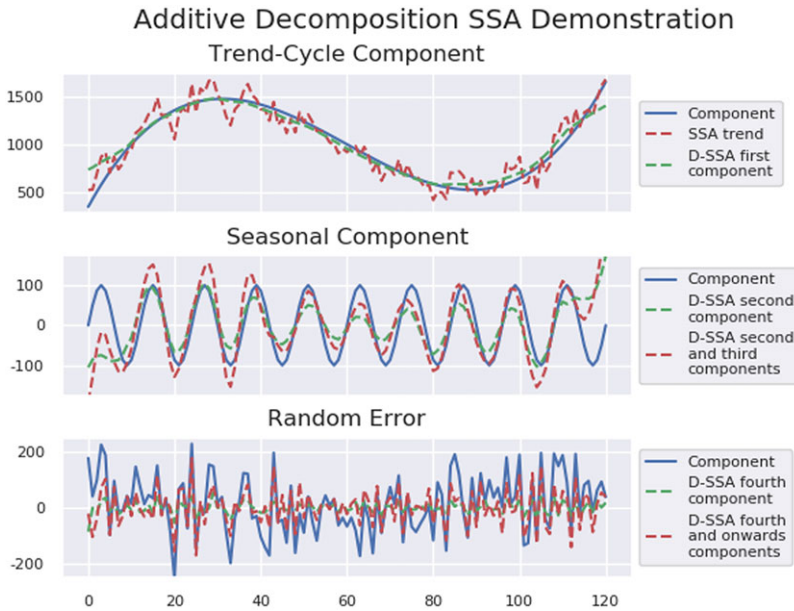


Figure 4. SSA and D-SSA trend estimate and component isolation compared against the corresponding underlying structures.

$$X_I[i] = \sum_{r+c=i} X_I[r, c], \tag{20}$$

with the trend estimation then proceeding by using an appropriate averaging vector to reverse the lagged embedding:

$$x_{SSA}(t) = \sum_{i=1}^{p-1} \frac{(p+1)-i}{p} X_I[i] + \sum_{i=p}^{L-p} \frac{1}{p} X_I[i] + \sum_{i=L-p+1}^L \frac{(L-i)}{p} X_I[i]. \tag{21}$$

In the original SSA, as proposed in Hassani (2007), the only input variables used in the code above are `time_series`, `L`, and `est`. The time series (`time_series`) is required as well as the embedding dimension (`L`) and the grouping factor (`est`). One variation on the original algorithm can be noted as the second output of the algorithm which is `time_series_decomp`. Rather than automatically group the structures, as in `time_series_est`, `time_series_decomp` stores each of the structures sequentially for the user’s convenience.

```
time_series_est, time_series_decomp = \
    CovRegpy_ssa(time_series, L, est=3, plot=False, KS_test=False,
                 plot_KS_test=False, KS_scale_limit=1,
                 figure_plot=False)
```

An example of the D-SSA output (`time_series_decomp`) is plotted in Fig. 4. One is advantaged in this synthetic example in that one has foreknowledge of the underlying structures, but this would not necessarily hold when applied to real-world examples. The various components extracted can be seen against their corresponding underlying components in Fig. 4. The additional inputs in the above code will be discussed in the Supplement to: Package CovRegpy: Regularized Covariance Regression and Forecasting in Python.

4.5 Singular spectrum decomposition

In this section SSD, originally introduced in Bonizzi *et al.* (2014), is discussed and several algorithmic variations are also introduced. This technique is available to the user in `CovRegpy_SSD.py` as well as the algorithmic variations. For a detailed description of the modifications made to SSA to create SSD, see Bonizzi *et al.* (2014). In summation, the major adjustments are the creation and automation of the extraction of a trend described in Section 4.5.1 and the formal extraction of components based on defined narrow bandwidths as described in Section 4.5.2.

4.5.1 Dealing with trended decomposition in the presence of a significant trend

Unlike in SSA, there is a formal test for a significant trend before the remainder of the decomposition algorithm is performed. If the normalized maximum frequency (f_{\max}/F_s) is below some threshold, the time series is deemed to have a significant trend. In Bonizzi *et al.* (2014), the frequency threshold is set to $f_{\text{threshold}} = 10^{-3}$. If this threshold is satisfied, then L is set to $T/3$ to isolate the trend. This follows that done in Vautard *et al.* (1992). If this threshold is not met, the embedding dimension, L , is set to $1.2F_s/f_{\max}$.

4.5.2 Downsampling in singular spectrum decomposition

This process can be summarized as calculating a frequency band which will be used to down-sample the possibly detrended time series. By constructing three Gaussian functions such that:

$$\gamma(f, \theta) = \sum_{i=1}^3 A_i e^{-\frac{(f-\mu_i)^2}{2\sigma_i^2}}, \tag{22}$$

with the μ parameters, and therefore, the centers of the Gaussian functions to be fitted to the power spectral density (PSD), being fixed and defined as:

$$\mu_1 = f_{\max}, \mu_2 = f_2, \mu_3 = \frac{f_{\max} + f_2}{2}, \tag{23}$$

where f_{\max} is the frequency at which the maximum PSD mode is located and f_2 is the frequency at which the second-highest PSD mode is located, with the other parameters being initialized as:

$$\begin{aligned} A_1^{(0)} &= \frac{1}{2} PSD(f_{\max}), & \sigma_1^{(0)} &= f : PSD(f) = \frac{2}{3} PSD(f_{\max}), \\ A_2^{(0)} &= \frac{1}{2} PSD(f_2), & \sigma_2^{(0)} &= f : PSD(f) = \frac{2}{3} PSD(f_2), \\ A_3^{(0)} &= \frac{1}{4} PSD(f_3), & \sigma_3^{(0)} &= 4|f_{\max} - f_2|, \end{aligned} \tag{24}$$

where $\sigma_1^{(0)} = f : PSD(f) = \frac{2}{3} PSD(f_{\max})$ denotes the frequency at which the PSD is equal to two-thirds of the PSD of the maximum PSD mode nearest to the maximum PSD – this allows for probable asymmetric peaks in the PSD.

L1 Singular Spectrum Decomposition:

With μ_1, μ_2 , and μ_3 fixed, Equation (22) can be fitted using LASSO regression as in Equation (25). In practice, L1 SSD can result in a perpetual loop if not managed correctly and should be used with caution.

$$\min_{\theta} ||\gamma(f, \theta) - PSD(f)||_1. \tag{25}$$

L2 Singular Spectrum Decomposition:

With μ_1, μ_2 , and μ_3 fixed, Equation (22) can be fitted using ridge regression as in Equation (26). This is the preferred method as LASSO regression promotes scarcity and in the

event of the θ vector being optimized to be identically zeros, without a counting stopping criterion, would result in an infinite loop:

$$\min_{\theta} \|\gamma(f, \theta) - \text{PSD}(f)\|_2^2. \quad (26)$$

Once Equation (22) has been fitted to the PSD, the optimal frequency band is calculated as $(f_{\max} - \delta f, f_{\max} + \delta f)$ where δf is calculated as in Equation (27):

$$\delta f = 2.5\sigma_1^{\text{opt}}. \quad (27)$$

This frequency band is demonstrated in Fig. 6. Figures, such as Figs. 5 and 6, can be optionally output when running the code below if one has `plot=True` which outputs every incremental PSD fitting process.

4.5.3 Stopping criterion in singular spectrum decomposition

To prevent over-sifting (a word borrowed from EMD, but accurately summarizes the concept), one should introduce a stopping criterion that stops the algorithm after a certain number of components have been isolated. The SSD algorithm results in the following decomposition:

$$x_{\text{SSD}}(t) = \sum_{i=1}^M \tilde{g}_i(t) + v_{M+1}(t), \quad (28)$$

with $\tilde{g}_i(t)$ being the i th component and $v_{M+1}(t)$ being the residual. This stopping criterion stops the algorithm when a certain percentage of the PSD of the time series has been accounted for in the formal components extracted. The algorithm stops when the normalized mean squared error (NMSE), calculated as follows:

$$\text{NMSE}_i = \sum_{t=0}^N \frac{v_{(i+1)}^2(t)}{x^2(t)}, \quad (29)$$

drops below a certain percentage. A value of $\alpha = 0.01$ is recommended in Bonizzi *et al.* (2014) but is adjustable in this package. The time series (`time_series`) is the sole input required for the algorithm with the other additional inputs required being `initial_trend_ratio`, `nmse_threshold`, `plot`, and `debug`.

```
ssd_decomp = CovRegpy_ssd(time_series, initial_trend_ratio=3,
                           nmse_threshold=0.01, plot=False, debug=False)
```

The `nmse_threshold` value has been discussed in this section with `initial_trend_ratio` being discussed in Section 9.2.3 of Supplement to: Package CovRegpy: Regularized Covariance Regression and Forecasting in Python. The `plot=True` input variable results in numerous incremental plots such as Figs. 5 and 6. If `debug=True`, then each incremental value of the calculated NMSE_i will be printed for debugging purposes. The components resulting from SSD being applied to the synthetic time series described in Fig. 2 can be seen in Fig. 7 for comparison with Figs. 3 and 4.

4.6 Implicit factor context with regularized covariance regression

As detailed in Fig. 1, the three implicit factor extraction algorithms described in this section isolate components that may then be used as independent covariates (X) in the RCR framework described in the following section. The implicit factor algorithms are not a necessary step, but it is advised that if covariance forecasting is intended that the components used in Algorithm 1 be smoothed or subjected to some form of preprocessing to ensure the forecasted covariance evolves smoothly over time as in Figs. 8, 9, 10, and 11.

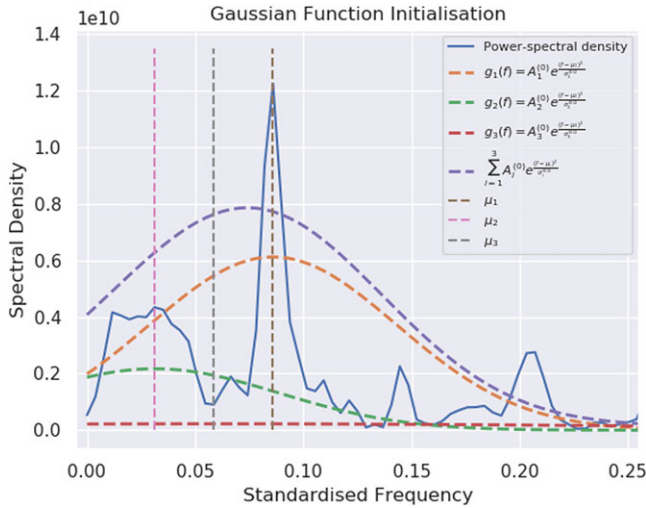


Figure 5. Initialization of Gaussian functions, Equation (22), to be fitted to the PSD of the additive synthetic example presented in Fig. 2 for downsampling in SSD.

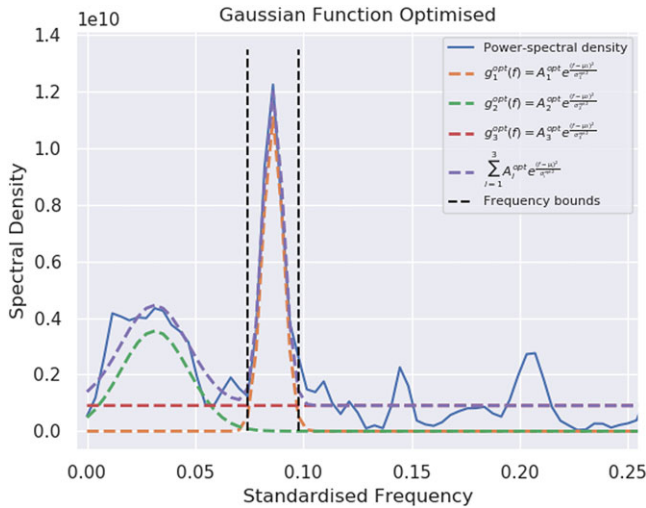


Figure 6. Fitting of Gaussian functions, Equation (22), to the PSD of the additive synthetic example presented in Fig. 2 for downsampling in SSD.

5. Actuarial Case Study 1: Reduction in lung capacity versus age and physiology as an effective early lung disease indicator

In this case study, we introduce the covariance regression model from Hoff & Niu (2012). This model is generalized in this work through the use of the implicit factors as exogenous variables in the covariance regression framework and the regularization of the covariance regression introduced in Sections 5.4, 5.5, 5.6, 5.7, and 5.8. We also formalize the independent mean definition in Section 5.3 wherein users can define the mean of the dependent variable before fitting the covariance regression. We make clear the actuarial applications of this research in Section 5.1. Section 5.2 notes the known diagnostic-assisting strength of forced expiratory volume (FEV) measures as

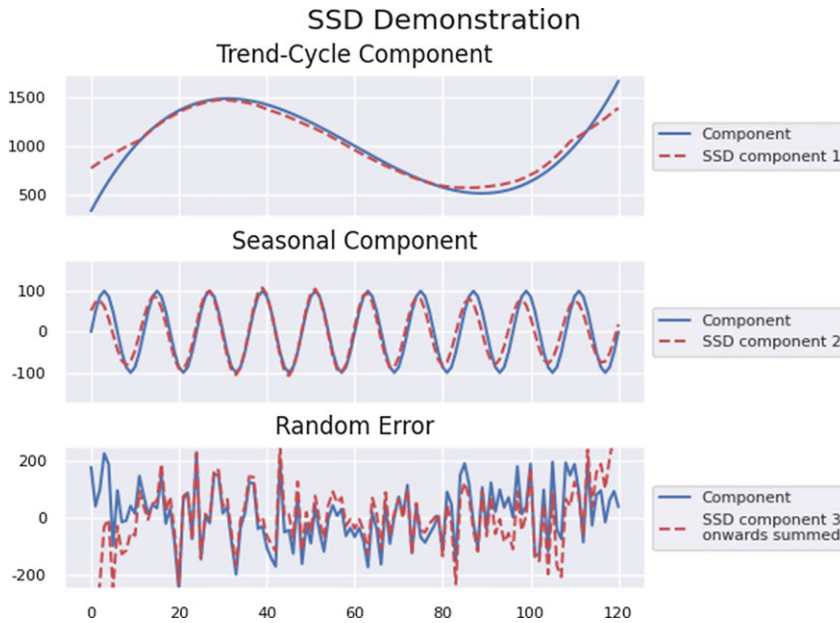


Figure 7. SSD of example time series using adjustable initial trend isolation window with the ubiquitous problem that also permeates EMD analysis known as the edge effect or frequency leakage between trend-cycle component and random error or irregular component.

it pertains to correctly identifying lung obstructions and degeneration as a proxy for underlying conditions. The early diagnosis of common respiratory disorders such as:

- small cell lung cancer;
- non-small cell lung cancers such as:
 - adenocarcinoma;
 - squamous cell cancer;
 - large cell carcinoma;
- emphysema;
- chronic bronchitis; and
- chronic asthma;

would significantly reduce the cost to insurers. Early diagnosis (through cheap tests and indicators such as FEV) would lead to low-cost early treatment or prevention as compared to late-stage diagnosis and expensive treatment and end-of-life therapy. These cost-benefits are discussed in Section 5.1 with the specifics of FEV as an indicator of underlying condition being outlined in Section 5.2. The rest of the section is dedicated to the exposition of RCR with some closing remarks and a brief summary in Section 5.9.

5.1 Early lung disease detection and insurance benefits

In Pyenson *et al.* (2012), the short-term costs and long-term benefits of screening for lung cancer in high-risk populations (aged 50–64 years in the USA) is demonstrated. The further reduced cost of the FEV screening (as an indicator for cancer or further, more-expensive, testing) of at-risk populations in the UK and elsewhere would greatly reduce the strain on national health institutions. The FEV measures examined herein, the research of Cho & Stout-Delgado (2020) which

notes a distinctive linear decrease in lung capacity with age, and our research which puts confidence bounds on the lung capacity as a function of age and size would all assist in cheap, early, and accurate measurement of reduced lung capacity that would reduce the future cost of cancer treatment. The research of Pyenson *et al.* (2012) strongly indicates that cheap and early screening for lung cancer would greatly decrease the future strain of healthcare costs on insurers.

Fitch *et al.* (2011) assesses the prevalence of COPD (in patients without coexisting asthma) in a large group of patients (44,366) with a focus on determining a link between the severity of the condition and the level of adherence to prescribed medication. This work assists in determining the distributional severity of COPD (grouped into mild, moderate, severe, and very severe) among those making claims which allows insurers to make reasonable adjustments to their health plans based on this distribution. In Section 5.2, we make clear the diagnostic value of the FEV measurement for further tests or underlying lung degeneration which would also assist in grouping insurance clients into probabilistic groups based on likely severity of condition and mortality rates applications to policy pricing.

It is noted in Pyenson *et al.* (2014) that lung cancer is, by a wide margin, the greatest cause of cancer deaths worldwide. Low-dose computed tomography (LDCT) is used to confirm the presence of lung cancer, but these tests are expensive and only conducted on populations deemed to be at risk based solely on smoking history and age – these two factors are, admittedly, strongly correlated with lung cancer instances. FEV and other inexpensive lung function tests are effective earlier indicators of reduced lung function for insurers to plan and allocate funds for future LDCT tests and potential treatments. Measures of lung capacity relative to age and size of an individual (which can be further stratified by gender) among a broader grouping of at-risk individuals such as younger populations in polluted cities or with vocations where exposure to hazardous airborne particulates can assist in early detection and further save insurers future health expenditures. The FEV (modeled herein) and other lung capacity tests (such as FEV per second compared to total FEV) are shown in Section 5.2 to be diagnostically valuable indicators of diseased lungs.

Dash & Grimshaw (1993) assess the costs associated with positive cancer diagnoses from a different perspective: Dread Disease contracts. Dread Disease contracts are insurance policies that have become increasingly popular in the UK (Dash & Grimshaw, 1993) and payout predetermined lump sums upon the realization of certain mortal contingencies. As stated above, lung cancer is the largest cause of cancer deaths and any indicators of potential future lung deficiencies (such as FEV discussed herein) that could lead to cancer could significantly decrease the future cost to the underwriters of these Dread Disease contracts. Changes in marketing policies in the UK have led to an increased interest in these contracts.

Efird *et al.* (2014) investigates the stage of lung cancer diagnosis as it relates to racial groups and health insurance policies. It is noted that survival rates among lung cancer diagnoses are intimately linked with the stage at which it is first detected. It is noted that the type of insurance policy did not relate to the stage of the cancer diagnosis. Cheap and diagnostically valuable information such as early indicators of decreased lung capacity (FEV discussed herein) could assist in ameliorating this notably significant disparity between racial groups which would assist with the costs to insurers as a result of late diagnosis and expensive late-stage treatment compared to comparatively cheaper earlier stage treatments.

Yang *et al.* (2019) notes the increasing life expectancies of those living with HIV and AIDS as a result of the provision of antiretrovirals. This decreased HIV-related mortality has had the unintended consequence of increased exposure to chronic illnesses, particularly lung diseases, which are often exacerbated by HIV, AIDS, or the associated antiretrovirals. It is noted in Yang *et al.* (2019) that the presence of HIV or AIDs in an individual's system results symptoms that mimic an increased rate of biological aging which, when taking into account with this research and the known linear degradation in lung capacity with age (Cho & Stout-Delgado, 2020), can be used in the early detection of lung functional decay for the provision, by large private insurers (Yang *et al.*, 2019), of funding for the relevant treatment.

5.2 Forced expiratory volume and mortality

As early as 1846 in Hutchinson (1846) (which cites far older work), the measure of the lung capacity as a proxy for the normal function of the lungs was established as a viable diagnostic tool for determining the presence of lung disease. Specifically, this work focused on the spirometer which measures the rate of air expulsion from the lung as well as the total lung expulsion and uses these measurements (FEV per second and total FEV) against normally distributed values based on an individual's height (as a proxy of size and, therefore, expected lung capacity), gender, and age when making probabilistic predictions (which would be greatly assisted by our research) about an individual's health. These measurements and an individual's expected health incomes would greatly assist in calculating competitive insurance premiums.

Casanova *et al.* (2005) provides evidence that supports the correlative link between decreased FEV and the increased likelihood of death related to COPD which is caused by a number of serious medical conditions such as emphysema and chronic bronchitis. It is noted in this study, Casanova *et al.* (2005), that age is also a contributory factor in the decrease of lung capacity measured by the proxy of lung capacity, FEV. The study of the covariance between FEV and age in this work should be used in conjunction with work such as Dyer (2012) which studies the decreases in lung capacity with age in the absence of confounding medical causes to better predict mortality outcomes.

Further, Dykstra *et al.* (1999) uses the FEV (among numerous other lung capacity measures) as a contributory diagnostic tool when determining whether a patient with a ratio of FEV per second to total FEV of less than 70% has asthma or COPD. By noting the stabilizing trend of FEV versus age in Fig. 8 and the work in Cho & Stout-Delgado (2020) that supports a known approximately linear level of lung deterioration with the age beyond 35 years (approximately 20 mL/year), one can use deviations from this trend as a contributory diagnostic factor. Cho & Stout-Delgado (2020) and our work can also assist in the construction of modified mortality tables for individuals whose lung capacity depreciation rate is above those forecasted and observed.

In the study conducted in French *et al.* (2015), both age and lowered FEV (and the further measure of the ratio of inspiratory capacity to total lung capacity once a lowered ratio of FEV per second to total lung capacity established the presence of COPD) were shown as significant indicators of increased mortality. Specific occupations have increased the risk of lung diseases – Farmer's Lung (caused by the increased inhalation of biologic dust particles) is one such disease. Braun *et al.* (1979) studies the long-term and clinical outcomes of this disease – this research (and other such research) coupled with our research could assist in competitive pricing of life insurance premiums and provide a framework for individual insurance pricing.

Covariance regression was originally proposed in Hoff & Niu (2012). Algorithm 1 defines a modified version of the original algorithm where the local mean of the dependent variables can be estimated independently of the algorithm. This allows for increased flexibility in defining a local mean structure. A rank 1 covariance regression model seeks to calculate Ψ , the base unattributable (or systemic or contemporaneous) covariance, and \mathbf{B} , the matrix of coefficients relating the covariance of \mathbf{Y} to the independent variables \mathbf{X} such that:

$$\text{cov}[\mathbf{Y}|\mathbf{X}] = \Psi + \mathbf{B}\mathbf{X}\mathbf{X}^T\mathbf{B}^T. \quad (30)$$

In the following section, the independent mean-covariance regression model is introduced in translatable pseudocode in Algorithm 1. In this algorithm, Equation (31) is repeated from Hoff & Niu (2012) which assumes the iterative inverses are well behaved. Within this software package, the pseudo-inverse is used to avoid unnecessary errors. The sections that follow this restatement of the algorithm (with a defined pseudocode algorithm and independent mean definition) are discussed with particular focus on variations of the Equation (31) for differing degrees of fitting and variable selection – these algorithmic variations are presented herein for the first time and applied to the same data as in Hoff & Niu (2012) for direct comparison with the original technique both formally and graphically as in Figs. 10 and 11. These algorithmic variations are presented here for the reader to explore and develop – the implicit factor models developed in the previous

Algorithm 1 Covariance Regression

Require: $\mathbf{A}, \mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}$, iterations

Initialize:

(1) $\mathbf{m} \sim \mathcal{MVN}(\mathbf{0}, \mathbb{I}_n)$

(2) $\mathbf{v} = \mathbf{0}_{1 \times n}$

for iter in $[1, \dots, \text{iterations}]$ **do**

construct $\tilde{\mathbf{X}}_2 \in \mathbb{R}^{2n \times r}$ with the i^{th} row being $m_i \mathbf{x}_{i,2}^T$ and the $(n+i)^{\text{th}}$ row being $v_i \mathbf{x}_{i,2}^T$

construct $\tilde{\mathbf{Y}} \in \mathbb{R}^{2n \times p}$ such that $\tilde{\mathbf{Y}} = [(\mathbf{Y} - \mathbf{X}_1 \mathbf{A})^T, \mathbf{0}_{n \times p}^T]^T$

calculate

$$\mathbf{B} = \tilde{\mathbf{Y}}^T \tilde{\mathbf{X}}_2 (\tilde{\mathbf{X}}_2^T \tilde{\mathbf{X}}_2)^{-1} \tag{31}$$

construct $\mathbf{C} = [\mathbf{A}^T, \mathbf{B}^T]^T \in \mathbb{R}^{(q+r) \times p}$

construct $\tilde{\mathbf{X}} \in \mathbb{R}^{2n \times (q+r)}$ with the i^{th} row being $[\mathbf{x}_{i,1}^T, m_i \mathbf{x}_{i,2}^T]$ and the $(n+i)^{\text{th}}$ row being $[\mathbf{0}_q^T, v_i \mathbf{x}_{i,2}^T]$

construct $\tilde{\mathbf{Y}}^* \in \mathbb{R}^{2n \times p}$ such that $\tilde{\mathbf{Y}}^* = [\mathbf{Y}^T, \mathbf{0}_{n \times p}^T]^T$

calculate $\Psi = (\tilde{\mathbf{Y}}^* - \tilde{\mathbf{X}} \mathbf{C})^T (\tilde{\mathbf{Y}}^* - \tilde{\mathbf{X}} \mathbf{C}) / n$

calculate $v_i = (1 + \mathbf{x}_{i,2}^T \mathbf{B}^T \Psi^{-1} \mathbf{B} \mathbf{x}_{i,2})^{-1}$

calculate $m_i = v_i (\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{x}_{i,1}})^T \Psi^{-1} \mathbf{B} \mathbf{x}_{i,2}$

end for

store Ψ, \mathbf{B}

section are ideal candidates for independent variables in the RCR framework owing to (most of) them yielding smoothed factors (compared to discontinuous explicit market factors such as the 10-year bond yield, etc.) for relatively smooth forecasted covariance transitions.

5.3 Independent mean-covariance regression

In Hoff & Niu (2012), covariance regression is introduced along with the formal derivations. In this section, a modified and more robust version of the algorithm originally presented in Hoff & Niu (2012) is presented and detailed in Algorithm 1.

In the absence of some stopping criterion, the modified covariance regression algorithm requires five inputs. Matrices $\mathbf{A} \in \mathbb{R}^{(q \times p)}$ and $\mathbf{X}_1 \in \mathbb{R}^{(n \times q)}$ are the major modification to the original algorithm in that the means of the dependent variable, \mathbf{Y} , can be calculated and optimized independent of the variance calculation. This increased robustness is ideal for variance attributable to different frequency structures which are common in the financial setting. The mean matrix, $\boldsymbol{\mu} \in \mathbb{R}^{(n \times p)}$ is calculated as $\boldsymbol{\mu} = \mathbf{X}_1 \mathbf{A}$ with q being the number of independent variables (implicit or explicit factors) used in the model, p being the number of dependent variables, and n being the number of observations.

With $\mathbf{X}_2 \in \mathbb{R}^{(n \times r)}$ being the matrix of independent variables which can be different to the basis matrix for mean construction, that is, different to \mathbf{X}_1 , with $\mathbf{Y} \in \mathbb{R}^{(n \times p)}$ being the matrix of independent variables, $\mathbf{B} \in \mathbb{R}^{(p \times n)}$ being the matrix of structural covariance coefficients, and with **max_iter** being the maximum allowed algorithmic iterations.

Using Algorithm 1, the data from Hoff & Niu (2012), and the specification of the cubic B-spline knot points as in Hoff & Niu (2012) (knots at 4, 11, 18), one arrives at Equation (32) below

Rank 1 Figure 5 in Hoff and Niu (2012)

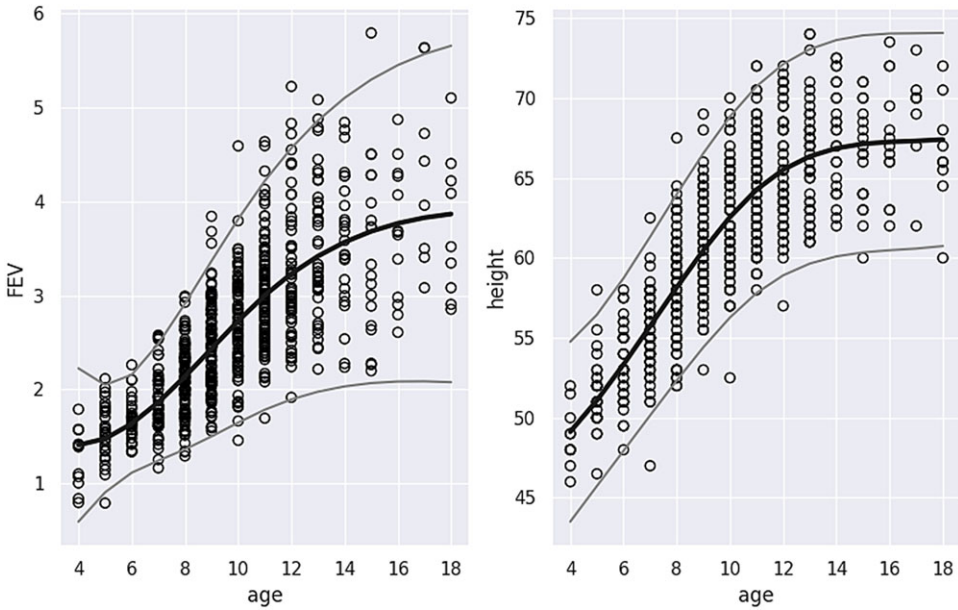


Figure 8. Replication of Fig. 5 from (see p. 24, Hoff & Niu, 2012) plotting FEV data and variance in the left figure and height data and variance in the right figure as a function of age.

with $\mathbf{X}_2 = [\mathbf{1}, \mathbf{age}^{\frac{1}{2}}, \mathbf{age}]^T$. Using these parameters, one can arrive at Figs. 8 and 9 to estimate the covariation of FEV and height as a function of age from Hoff & Niu (2012):

$$\mathbf{B}_{direct} = \begin{bmatrix} -2.60408 & -10.45065 \\ 1.43707 & 6.44579 \\ -0.14644 & -0.82822 \end{bmatrix}. \tag{32}$$

5.4 LASSO covariance regression

Equation (31) is the direct calculation of \mathbf{B} , but in RCR \mathbf{B} can be calculated in several ways. \mathbf{B}_{LASSO} is calculated by minimizing the following:

$$\arg \min_{\mathbf{B}} \left((\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}_2 \mathbf{B}^T)^T (\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}_2 \mathbf{B}^T) + \lambda_1 \|\mathbf{B}\|_1 \right). \tag{33}$$

This form of optimization promotes sparsity as can be seen in Equation (34). In Equation (34), it can be noted that the variances and covariance of FEV and height concerning age are most dependent on age. As a result of the structure of Equation (30), the covariance matrix becomes a function of age squared:

$$\mathbf{B}_{lasso} = \begin{bmatrix} -0.00000 & -0.00000 \\ -0.00000 & -0.00000 \\ -0.04537 & -0.12452 \end{bmatrix}. \tag{34}$$

Rank 1 Figure 6 in Hoff and Niu (2012)

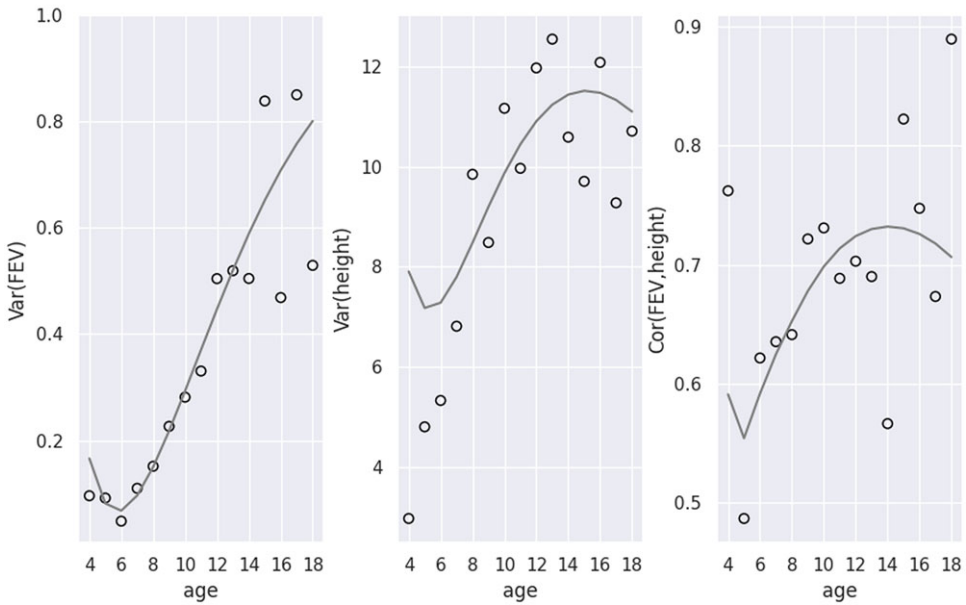


Figure 9. Replication of Fig. 6 from (see p. 25, Hoff & Niu, 2012) plotting FEV data and variance, height data and variance, and FEV and height data and correlation as a function of age.

5.5 Ridge regression covariance regression

Ridge regression penalizes unjustifiably large coefficients in **B** by optimizing the following:

$$\arg \min_{\mathbf{B}} \left((\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}_2 \mathbf{B}^T)^T (\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}_2 \mathbf{B}^T) + \lambda_2 \|\mathbf{B}\|_2^2 \right). \tag{35}$$

In this setting, by manipulating Equation (35), one can observe that ridge regression introduces a bias such that:

$$\arg \min_{\mathbf{B}} \left(\|\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}_2 \mathbf{B}^T + \lambda_2 \mathbf{B}\|_2^2 \right). \tag{36}$$

By referring to Equation (31) and noting the structure of Equation (36), Equation (37) follows as a biased version of the original direct solution where λ_2 can also be seen as a bias term rather than merely a penalty:

$$\mathbf{B} = \tilde{\mathbf{Y}}^T \tilde{\mathbf{X}}_2 \left(\tilde{\mathbf{X}}_2^T \tilde{\mathbf{X}}_2 + \lambda_2 \mathbf{I} \right)^{-1}. \tag{37}$$

By comparing Equation (32) against Equation (38), it can be noted that the coefficients have been greatly reduced as a result of the penalizing or biasing term:

$$\mathbf{B}_{\text{ridge}} = \begin{bmatrix} 0.27839 & 1.31838 \\ 0.07562 & -0.51146 \\ -0.09374 & -0.10844 \end{bmatrix}. \tag{38}$$

5.6 Elastic-net covariance regression

Elastic-net regression is a compromise between LASSO regression and Ridge regression where \mathbf{B} is optimized using the following:

$$\arg \min_{\mathbf{B}} \left((\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}_2 \mathbf{B}^T)^T (\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}_2 \mathbf{B}^T) + \lambda_1 \|\mathbf{B}\|_1 + \lambda_2 \|\mathbf{B}\|_2^2 \right). \tag{39}$$

This can be reparameterized (as is in Python) to:

$$\arg \min_{\mathbf{B}} \left((\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}_2 \mathbf{B}^T)^T (\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}_2 \mathbf{B}^T) + \lambda_1 r_1 \|\mathbf{B}\|_1 + \frac{1}{2} \lambda_1 (1 - r_1) \|\mathbf{B}\|_2^2 \right), \tag{40}$$

with r_1 being the ratio of the L1-norm to the L2-norm. Noting that Equation (39) can be viewed as a combination of Equations (33) and (36), it becomes clear why Equation (41) displays both sparsity and smaller parameters. This is seen as a favorable compromise between the two methods:

$$\mathbf{B}_{\text{elastic-net}} = \begin{bmatrix} 0.00000 & 0.00000 \\ 0.02062 & -0.04304 \\ -0.05310 & -0.12245 \end{bmatrix}. \tag{41}$$

5.7 Group-LASSO covariance regression

The group-LASSO regression promotes a sparse number of groups of parameters and in each group, there will be a sparse set of parameters:

$$\arg \min_{\mathbf{B}} \left((\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}_2 \mathbf{B}^T)^T (\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}_2 \mathbf{B}^T) + \lambda_1 \|\mathbf{B}\|_1 + \lambda_g \sum_{g \in \mathcal{G}} \sqrt{d_g} \|B_g\|_2 \right). \tag{42}$$

In Equation (43), λ_g is the penalty term for the grouped coefficient terms, \mathcal{G} is a partition of the coefficients into subsets, d_g is the cardinality of each of the subsets (larger sets get penalized more than smaller sets), and B_g is the set coefficients within subgroup g . With a grouping of $\{0, 1, 1\}$, that is with the constant being in a group separate from $\mathbf{age}^{\frac{1}{2}}$ and \mathbf{age} , this results in the following parameter estimates:

$$\mathbf{B}_{\text{group-lasso}} = \begin{bmatrix} 0.00823 & -0.00091 \\ 0.00577 & -0.03660 \\ -0.04977 & -0.12262 \end{bmatrix}. \tag{43}$$

5.8 Subgradient covariance regression

Equation (44) can be seen as being identical to Equation (33) in Section 5.4. The differences lie in the method of approaching this solution. Subgradient descent or subgradient optimization allows optimization when functions are continuous, but non-differentiable:

$$\begin{aligned} & \text{minimize } \|\mathbf{B}\|_1 \\ & \text{subject to } \tilde{\mathbf{X}}_2 \mathbf{B}^T = \tilde{\mathbf{Y}}, \end{aligned} \tag{44}$$

To initialize the algorithm, several candidates can be used. In this algorithm, \mathbf{B}_0 is initialized with Equation (31) which is the direct solution in Euclidean Norm Space. The next iteration of the solution, \mathbf{B}_{k+1} , is calculated using the following equation:

$$\mathbf{B}_{k+1} = \mathbf{B}_k - \alpha_k \left(\mathbf{I}_p - \tilde{\mathbf{X}}_2^T (\tilde{\mathbf{X}}_2 \tilde{\mathbf{X}}_2^T)^{-1} \tilde{\mathbf{X}}_2 \text{sign}(\mathbf{B}_k) \right), \tag{45}$$

with $\text{sign}(\cdot)$ being defined as:

$$\text{sign}(x) = \begin{cases} +1, & \text{for } x > 0 \\ 0, & \text{for } x = 0 \\ -1, & \text{for } x < 0 \end{cases}, \quad (46)$$

and with the incremental step size being a member of the square summable, but not summable family of step sizes. In this application, $\alpha_k = \frac{10^{-2}}{k}$ with the formal conditions of the square summable step sizes required to satisfy the following conditions:

$$\alpha_k > 0 \forall k > 0, \sum_{i=0}^{\infty} \alpha_k^2 < \infty, \text{ and } \sum_{i=0}^{\infty} \alpha_k = \infty. \quad (47)$$

With this technique applied to the dataset from Hoff & Niu (2012) using 10 iterations one arrives at:

$$\mathbf{B}_{\text{subgradient}} = \begin{bmatrix} -1.77455 & -4.37165 \\ 1.44572 & 3.53236 \\ -0.28942 & -0.71025 \end{bmatrix}. \quad (48)$$

To estimate the base unattributable (or systemic or contemporaneous) covariance, Ψ , and the coefficients, \mathbf{B} , that relate the attributable (or structural) covariance to the independent variables one can apply the following code:

```
B_est, Psi_est = \
    cov_reg_given_mean(A_est, basis, x, y, iterations, technique,
                       alpha, l1_ratio_or_reg, group_reg,
                       max_iter, groups)
```

Each of these techniques was applied to the original dataset from Hoff & Niu (2012) and the results are plotted in Figs. 10 and 11. In the code above, `A_est` and `basis` refer to the mean coefficients and basis to be used for means of the dependent variable, y , with x being the independent variable. The number of iterations of the core covariance regression algorithm is controlled by the `iterations` parameter. The technique is being used in estimating the \mathbf{B} matrix in each iteration of the total iterations. The options available for `technique` are “direct,” “lasso,” “ridge,” “elastic-net,” “group-lasso,” and “sub-gradient” which correspond to Sections 5.3, 5.4, 5.5, 5.6, 5.7, and 5.8, respectively. The unambiguous penalty term for LASSO, ridge, elastic-net, group-LASSO regression as well as the first value in the subgradient optimization is `alpha`.

The `l1_ratio_or_reg` controls the ratio of the l1-ratio to l2-ratio as in Equation (40) when performing elastic-net regression. The group-LASSO regression group penalty term is controlled by `group_reg`. While possibly misleading when compared against iterations, the `max_iter` parameter controls the number of iterations to be used in the subgradient descent algorithm as in Equation (45). The `groups` variable controls the grouping of the independent variables and the corresponding parameters using a vector of indices such as in Section 5.7. Finally, the `test_lasso` is a Boolean variable that controls whether the `alpha` value for LASSO regression is optimized.

5.9 Closing remarks on lung capacity versus physiology as a cost-saving benefit to insurers

Section 5.1 notes the decrease in costs to insurers from the early detection of a number of degenerative lung diseases. Several of the more well-known conditions that affect human lungs are listed in the introduction to Section 5. The physiological variable that is perceived as the most highly correlated with most of these conditions is a person’s age. This research has provided a framework for the early detection of the presence of a number of degenerative lung diseases using FEV measurement (as outlined in Section 5.2) as well as other relevant variables, besides age. With this

Rank 1 Figure 5 in Hoff and Niu (2012)

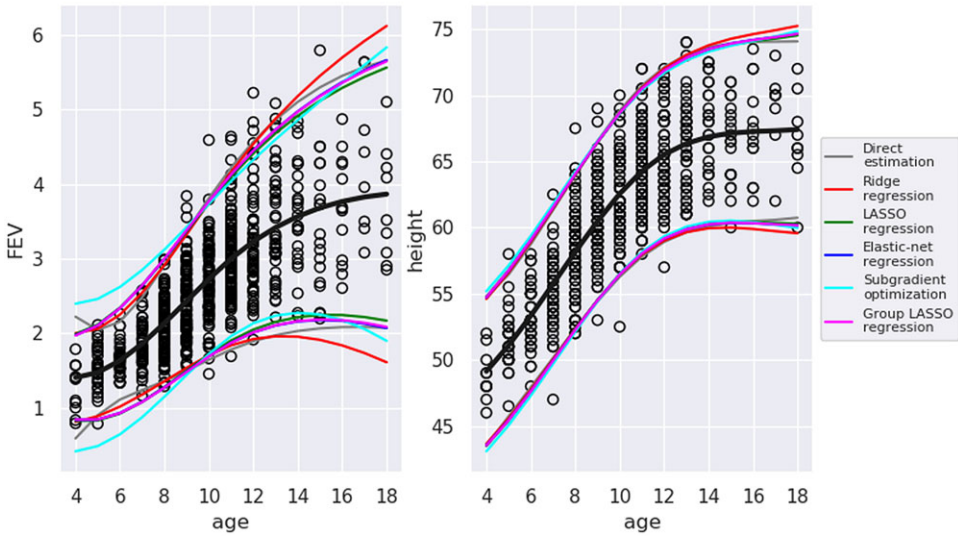


Figure 10. Replication of Fig. 5 from (see p. 24, Hoff & Niu, 2012) plotting FEV data and variance in the left figure and height data and variance in the right figure as a function of age with RCR alternate estimate.

Rank 1 Figure 6 in Hoff and Niu (2012)

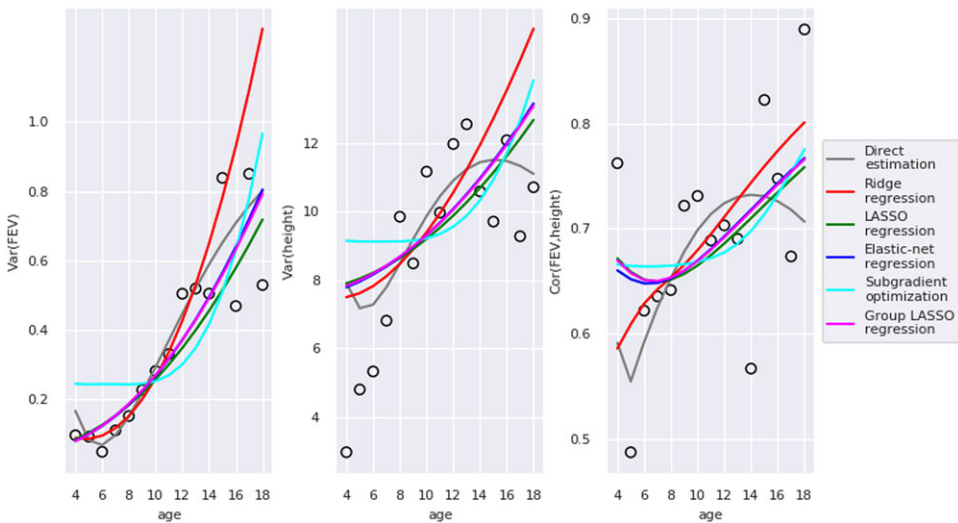


Figure 11. Replication of Fig. 6 from (see p. 25, Hoff & Niu, 2012) plotting FEV data and variance, height data and variance, and FEV and height data and correlation as a function of age with RCR alternate estimate.

framework, one can use any number of lifestyle or physiological variables to assess the presence of outliers in FEV measurements which can lead to early detection and significantly reduce the costs to insurers through comparatively cheaper early-stage treatment compared against late-stage treatment or end-of-life therapy.

Table 1. Summary statistics for quarterly returns of purely funded and unfunded pensions using the quarterly returns of the FTSE 100 as a proxy for unfunded pension schemes

Pension scheme	Funded				Unfunded
	High-Freq	Mid-Freq	Low-Freq	Equal	FTSE 100
Mean	0.0124	0.0144	0.0191	0.0138	0.0094
Variance	0.0088	0.0102	0.0106	0.0078	0.0006
VaR	-0.1762	-0.1485	-0.1189	-0.1254	-0.0073
CVaR	-0.2450	-0.2470	-0.2550	-0.2209	-0.0517
MDD	-2.3770	-2.0044	-2.1559	-2.7839	-2.3596

6. Actuarial Case Study 2: Covariance regression for hybrid funded–unfunded pension portfolios using covariance forecasting and RPP

In this case study, we begin by noting the risks associated with purely unfunded and purely funded pension plans in Section 6.1 – the primary quantitative financial risks are also demonstrated in this case study in Table 1 and Fig. 12 with purely funded pensions having higher, less-certain returns and purely unfunded pensions having significantly lower, more-certain returns. Hybrid pension schemes are motivated with context concerning the problems arising from changing demographics in developed nations in Section 6.2.

The RPP portfolio weighting strategy (included in this package) is introduced in Section 6.3 with it being subsequently used to weight the purely funded pensions based on the different frequency covariance forecasts using the framework given exposition in the previous section, Section 5. This case study and section is concluded in Section 6.4 where the mean-variance utility framework of hybrid pension schemes from Dutta *et al.* (2000) is introduced before we construct hybrid pension schemes using different risk aversion levels and different frequency forecasts of the covariance structure of the assets (in the funded portion) for the upcoming investment horizon.

6.1 Unfunded and funded pension schemes

Chapman *et al.* (2001) notes the lower and more certain returns of bond investment (as a proxy for fully unfunded) versus the higher and less certain returns of equity investment (as a fully funded pension proxy). Chapman *et al.* (2001) goes further in noting that employees, the enterprise, the stakeholders (shareholders), and the government seldom benefit from the same schemes. The stakeholders in the company (and the associated share price) should be taken into account when setting the contribution rates for the unfunded (or defined benefits) portion of the pension as well as the investment strategies of the companies once funds have been attained. Chapman *et al.* (2001) also notes the need for clarity for the employees and company as the funded portion of the pensions are secured versus the relatively unsecured portion of the pension that is unfunded should there be unforeseen market shocks such as the 2008–2009 financial crisis or the SARS-CoV-2 pandemic.

Splinter (2017) notes the broader economic benefits of maintaining an at least partially unfunded pension scheme for state employees in which it provides a stabilizing effect on a nation's economy by providing governments with funds without needing to either raise taxes or cut benefits. There is a downside risk in that the government (and by extensions the employees) then relies on the same assumptions of unfunded pensions in that the economy will continue to grow and the money can be provided or returned at a later date. Splinter (2017) notes that the main contributor to the under-funding of the public employees pensions is the unfunded portions as a result of an increase in state liabilities with an insufficient associated contribution by employees and the government. Whereas, the funded portions properly invested in diversified equity portfolios' returns

were in excess of expectations. Other causes were noted, but the large accumulation of liabilities was seen as the main contributor to less-than-favorable defined benefit pension funding.

Grubbs (1999) notes the difficulty experienced by some private firms in adequately funding defined benefits schemes before adequate legislation and planning was enforced. These failed fully defined benefit plans (or unfunded pensions) relies on inadequate contributions from employees and the companies. These resulted in a number of disastrous instances such as companies canceling all their unfunded or defined benefits schemes or only being able to continue paying a small portion of the retirees at the expense of the future retirees. An extended study on this topic was conducted in Trowbridge (1952).

With the problems with pure funding and unfunded pensions noted above and with the rise of demographic mismatches in developed populations discussed in the next section, Section 6.2, Devolder & de Valeriola (2019) notes that neither funded (referred to as defined contributions or DC) nor unfunded (referred to as defined benefits or DB) are adequate as risk-sharing compromises between generations. This, among other things, motivated the rise of hybrid strategies. The optimization in Devolder & de Valeriola (2019) of the ratio of funded to unfunded schemes is considered from the perspective of minimizing the disparity between retirement living conditions of different generations rather than some risk-returns trade-off of individual investors which is performed in this case study.

Wang & Lu (2019) studies hybrid pensions strategies to address the problem of risk-sharing between generations from a stochastic perspective with adjusted contributions and benefits depending on the performances of the pensions. Using the stochastic optimal control approach, closed-form solutions are derived using both a quadratic loss function and an exponential loss function. Various market parameters are adjusted with the contributions and benefits adjusted accordingly to demonstrate the effectiveness of risk-sharing between different generations when compared against purely funded or unfunded plans. Some simplifying assumptions were made in the analysis (such as assuming mortality rates remained constant over time) to allow the calculation of closed-form solutions.

Further, Blommestein *et al.* (2009) notes that hybrid pension strategies perform better for individual private investors in terms of funding ratios and replacement ratios. The simulations performed by Blommestein *et al.* (2009) show that hybrid strategies perform best in terms of sustainable financial risk-sharing between generations. It is found that in relatively stable environments where over-funding of hybrid strategies takes place, the most effective way of sharing risk is through conditional index-linked investing. It is shown how hybrid plans offer an agreeable compromise between DB plans with no benefit risk to the individual, but risk to the employer, and DC plans where the individual bears the entirety of the risk through investment decisions, inflation, and increased longevity.

Hoevenaars & Ponds (2008) evaluates these hybrid options from the balance sheet perspective as the sum of embedded generational options. This leads to the conclusion that any policy change of an individual inevitably leads to the transfer of risk between generations. It is found, maybe not surprisingly, that shifts in the investment pools of pensions to less risky assets shifts the risks from older generations to younger generations. Whereas, a shift in weightings between funded and unfunded within the hybrid strategy from flexible contributions and fixed benefits to fixed contributions and variable benefits shifts the risk from younger generations to older generations.

6.2 Hybrid funded–unfunded UK pension schemes to address demographic changes

The changing demographics of the more-developed countries in the world (proportionally larger aging working populations than developing countries) are putting significant strain on unfunded or pay-as-you-go (PAYG) pension systems – see Miles (2000). The motivation and evidence for moving away from PAYG pensions is based on developed economies having efficient market

portfolios that significantly exceed GDP growth and funded pensions lead to returns that are more robust to temporal distortions in the availability of working-age people within the population. These reasons, while arguably logical, are not in themselves conclusive for transfers from unfunded to funded pension systems for more developed nations.

Miles (1998) notes that for this shift from unfunded to funded pensions, the population would need more savings. It is also noted that given the aging populations and the unfunded pensions on which they depend, significant leveraged funding would be required by the governments to fund these shifts in policy or there would naturally be some losers when shifting systems owing to the funding needing to come from somewhere. It is, however, noted in Miles (1998) that the UK is further along in this process than most of its mainland Europe counterparts which face significant difficulties that have no doubt been exacerbated by the 2007–2008 financial crisis and SARS-CoV-2 pandemic-induced recessions since its publication. These demographic changes and the resulting difficulties they cause for governments in more developed countries to shift from unfunded to funded pension schemes have given rise to significant literature on the topic such as Disney (1996), Feldstein (1996), Kotlikoff (1996), Mitchell & Zeldes (1996), Roseveare *et al.* (1996), Feldstein & Samwick (1998), and Miles *et al.* (1999).

Sinn (1999) supports the full or partial shift away from PAYG pension systems as a way of addressing the growing demographic-related problems in more developed countries by replacing the loss of labor with capital. These demographic issues (larger ratio of retirees to working-age population, the associated costs of this relative aging population, and the drastically falling birth rates) put strain on PAYG systems in a similar manner as pyramids being dependent on a broader base.

Martell *et al.* (2013) also notes the difficult decision facing governments when funding pensions (by matching personal contributions) using debt increases the scheme's funding ratio which leads to a country's credit rating and global perception decreasing. The majority of the dangers of underfunding pensions for governments are related to the costs associated with an aging population. The costs are increasing as a result of the increasing ratio of retirees to the working population, the increasing life expectancy of retirees, and the increasing healthcare costs associated with an aging population living longer.

6.3 Equal RPP weighting strategy

RPP portfolio weighting strategies, proposed in Maillard *et al.* (2010), have had empirical successes when back-testing on equities over previous financial crises. The RPP weights are calculated as follows:

$$\boldsymbol{\omega}_{\text{med}} = \operatorname{argmin} \sum_{j=1}^N \left(\omega_j \frac{(\boldsymbol{\Sigma}_{\text{med}} \boldsymbol{\omega})_j}{\boldsymbol{\omega}^T \boldsymbol{\Sigma}_{\text{med}} \boldsymbol{\omega}} - b_j \right), \quad (49)$$

with $\boldsymbol{\omega}_{\text{med}}$ being the RPP weighting vector, $\boldsymbol{\Sigma}_{\text{med}}$ being the median of the forecasted covariance, N being the number of assets, and b_j being the relative proportion of risk allocated to asset j . Only the equal RPP portfolio weighting strategy ($b_j = \frac{1}{N}$) is explored in this paper and package, with the other constraints being:

$$\sum_{j=1}^N \omega_j = 1, \text{ and} \quad (50)$$

$$\sum_{j=1}^N \omega_j \mathbf{I}_{\{\omega_j \geq 0\}} \geq k_{\text{long}} \text{ for } k_{\text{long}} \geq 1.3, \text{ and} \quad (51)$$

$$\sum_{j=1}^N \omega_j \mathbf{I}_{\{\omega_j \leq 0\}} \geq k_{\text{short}} \text{ for } k_{\text{short}} \leq -0.3. \tag{52}$$

The final two constraints are referred to as long/short equity constraints that allow a certain level of shorting (k_{short}) to finance further long positions. These two weighting strategies are not exhaustive and are included here for context, clarity of case studies in the package, and completeness.

6.4 Mean-variance utility

We follow the model in Dutta *et al.* (2000) with the stochastic return of the portfolio over time increment $t - 1$ to t being $P_t = w_t R_{f,t} + (1 - w_t) R_{u,t}$ with w_t being the proportion of the portfolio in the diversified RPP portfolio based on either high-, mid-, or low-frequency structures with the stochastic returns of said portfolio being $R_{f,t}$ and $R_{u,t}$ being the stochastic returns of the unfunded portion which we approximate with the quarterly UK GDP. The resulting mean-variance utility function is such that:

$$\mathbb{E}[U(P_t)] = \mathbb{E}[P_t] - \frac{\gamma}{2} \text{Var}(P_t). \tag{53}$$

with γ being the risk aversion level to be discussed later. The resulting expected returns and variance of the portfolio are as follows:

$$\begin{aligned} \mathbb{E}[P_t] &= w_t \mathbb{E}[R_{f,t}] + (1 - w_t) \mathbb{E}[R_{u,t}] \\ &= w_t \mu_{f,t} + (1 - w_t) \mu_{u,t}, \end{aligned} \tag{54}$$

and

$$\text{Var}(P_t) = w_t^2 \sigma_{f,t}^2 + (1 - w_t)^2 \sigma_{u,t}^2 + w_t(1 - w_t) \sigma_{fu,t}, \tag{55}$$

respectively, with the realized returns of the funded portfolio over time increment $t - 1$ to t being $\mu_{f,t}$ (for either the high-, mid-, or low-frequency covariance forecasts), the realized returns of the unfunded portfolio over the same increment being $\mu_{u,t}$, $\sigma_{f,t}^2$ being the variance of the funded portfolio, $\sigma_{u,t}^2$ being the variance of the unfunded portion, and $\sigma_{fu,t}$ being the covariance between the two returns. By taking the derivative of Equation (53) with respect to w_t and setting to zero, it can be shown that:

$$w_t = \frac{\mu_{f,t} - \mu_{u,t} + \gamma(\sigma_{u,t}^2 - \sigma_{fu,t})}{\gamma(\sigma_{f,t}^2 + \sigma_{u,t}^2 - 2\sigma_{fu,t})}. \tag{56}$$

6.5 Different risk aversions in FTSE 100 hybrid pension schemes

In this study, we take the daily logarithmic returns of the constituents of the FTSE 100 from 1 October 2003 until 1 April 2023 (where available) to construct RPP portfolios using the high-, mid-, and low-frequency structures to forecast the covariance of the constituents and weight the portfolios based on these forecasts. The high-, mid-, and low-frequency structures from the daily logarithmic returns isolated from 1 October 2003 until 31 December 2003 are regressed against the daily logarithmic returns for the same time series from 1 January 2004 until 31 March 2004.

The high-, mid-, and low-frequency structures isolated from the daily logarithmic returns isolated from 1 January 2004 until 31 March 2004 are then used to forecast the covariance of

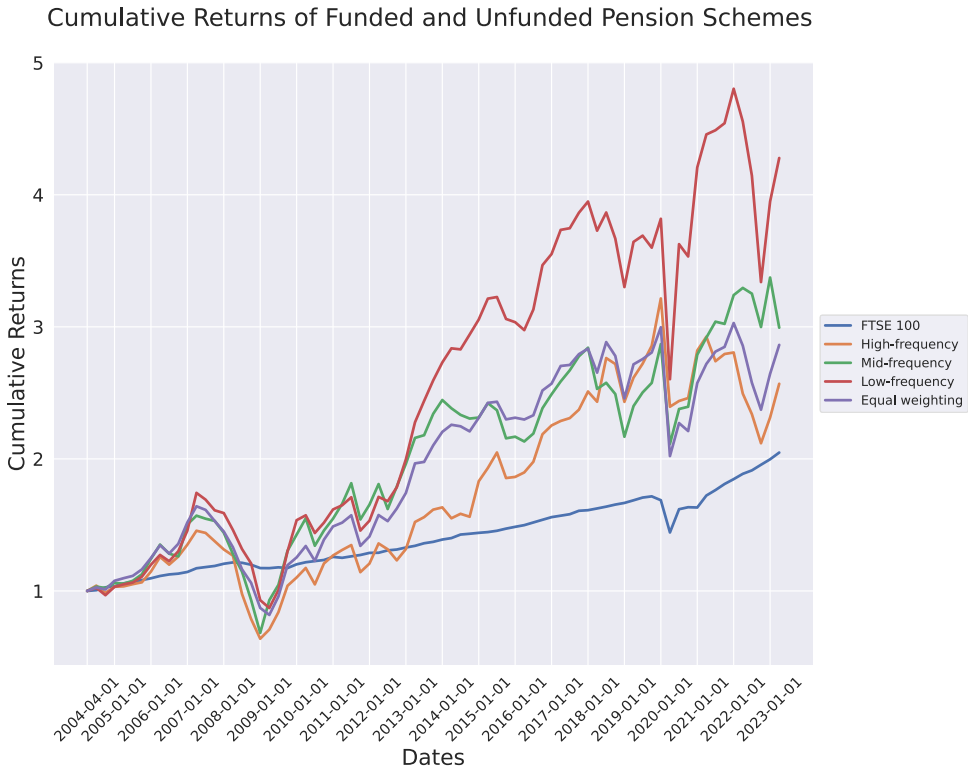


Figure 12. Plot of the cumulative returns of funded and unfunded pensions from 1 April 2004 until 1 April 2023.

the constituents from 1 April 2004 until 30 June 2004. The portfolio is then weighted based on the median of these daily covariance forecasts using the different frequency structures. The case study and resulting returns therefore begin from 1 April 2004 in Fig. 12. This process is repeated until the available GDP proxy of the FTSE 100 quarterly returns is exhausted which is 1 April 2023.

Rather surprisingly (considering the RPP optimization focusing on the minimization of risk), the RPP portfolios weighted based on the high-, mid-, and low-frequency forecasts of the upcoming investment period's covariance structure have higher returns when compared against the FTSE 100 and less desirable variance and value-at-risk (VaR). The cumulative returns of the portfolios, besides the FTSE 100 proxy, are much more susceptible to market-wide shocks such as those seen during the 2007–2008 financial crisis and the SARS-CoV-2 pandemic-induced recession. Based on the variance, the high-frequency content is more accurate when used to forecast the covariance than the mid-frequency and low-frequency content, but the VaR indicates that the low-frequency content is more susceptible to extreme shocks.

Noting the risk aversion level (γ) in Equations (53) and (56), we see that as the level of risk aversion is increased sequentially from 1.2 in Table 2 to 1.4 in Table 3 and, finally, to 1.6 in Table 4 the weightings in the respective funded pension portions of the hybrid schemes decrease. This study indicates that a desirable compromise between the level of acceptable risk and the expected returns of the pensions can be attained by weighting the funded portions of the pensions using RPP based on covariance forecasts and increasing or decreasing the risk aversion level to meet an individual's need regarding the proportion of the pension scheme to be funded or unfunded.

Table 2. Summary statistics for quarterly returns of hybrid pensions schemes using risk aversion level of $\gamma = 1.2$

Pension scheme	Hybrid schemes: $\gamma = 1.2$		
Summary stats	High-freq: $w_t \approx 0.296$	Mid-freq: $w_t \approx 0.429$	Low-freq: $w_t \approx 0.872$
Mean	0.0103	0.0116	0.0179
Variance	0.0014	0.0025	0.0084
VaR	-0.0504	-0.0595	-0.1030
CVaR	-0.1063	-0.1319	-0.2272
MDD	-3.2860	-2.6620	-2.1658

Table 3. Summary statistics for quarterly returns of hybrid pensions schemes using risk aversion level of $\gamma = 1.4$

Pension scheme	Hybrid schemes: $\gamma = 1.4$		
Summary stats	High-freq: $w_t \approx 0.251$	Mid-freq: $w_t \approx 0.362$	Low-freq: $w_t \approx 0.733$
Mean	0.0102	0.0112	0.0165
Variance	0.0012	0.0020	0.0063
VaR	-0.0423	-0.0491	-0.0865
CVaR	-0.0974	-0.1185	-0.1967
MDD	-3.1067	-2.8143	-2.1788

Table 4. Summary statistics for quarterly returns of hybrid pensions schemes using risk aversion level of $\gamma = 1.6$

Pension scheme	Hybrid schemes: $\gamma = 1.6$		
Summary stats	High-freq: $w_t \approx 0.217$	Mid-freq: $w_t \approx 0.312$	Low-freq: $w_t \approx 0.628$
Mean	0.0101	0.0110	0.0155
Variance	0.0010	0.0017	0.0050
VaR	-0.0361	-0.0413	-0.0740
CVaR	-0.0908	-0.1084	-0.1738
MDD	-2.9831	-2.7519	-2.1907

6.6 Closing remarks on hybrid pensions schemes

In Section 6.1, we note works that are critical of both purely funded and purely unfunded pensions schemes. In general (depending on the prevailing economic conditions over the lifetime of the scheme), the funded schemes have higher and less certain returns with the returns of the unfunded schemes being lower and more certain. Numerous works cited in this section allude to the necessity for some desirable compromise between these two schemes. Section 6.2 notes not only that hybrid pension schemes would benefit the holders of these schemes, but they would also address arising demographic issues in the majority of more-developed countries as the shrinking workforces and growing retired populations are putting significant financial strain on traditionally unfunded pension schemes.

We introduce the RPP weighting strategy in Section 6.3 which we use in this case study to allocate portfolio weights. The mean-variance utility methodology of hybrid pension schemes which we use to allocate the ratio of our unfunded portion (using the FTSE 100 as a proxy) and our funded portion (using difference frequency covariance forecasts) is outlined in Section 6.4. The construction of the unfunded proxy and the funded portfolios using different frequency forecasts is given exposition in Section 6.5. We display the results of both purely unfunded and purely funded pension schemes using mean returns as well as various common portfolio risk measures in Table 1.

We note that the funded portfolios have higher returns as well as significantly higher variances, VaRs, and CVaRs. This leads to the hybrid pension scheme results in the following tables. By noting Tables 2, 3, and 4, one can infer that the high-frequency structures used to forecast the forthcoming investment period's covariance structure are more responsive to changes in the covariance structures and therefore require a lower relative weighting (when compared against using mid- and low-frequency structures) in the funded portions of the portfolios to lower the risk measures such as portfolio variance, VaR, and CVar. The mid-frequency structures are less responsive than the high-frequency structures while being more responsive than the low-frequency and constant structures. This forms a hierarchy of performance in both overall returns and the risk measures.

7. Conclusion

In Section 4, several decomposition algorithms are presented. The first of the decomposition algorithms is EMD which had a thorough examination of its algorithmic variations in van Jaarsveldt *et al.* (2023a) and the associated AdvEMDpy package. X11 is presented in Section 4.3 with several algorithmic variations already existing, but not presented herein owing to the scope and the focus of this work being RCR. SSA, originally proposed in Hassani (2007), is presented in Section 4.4. The algorithmic extensions of SSA developed in this package are given exposition in Supplement to: Package CovRegpy: Regularized Covariance Regression and Forecasting in Python – these are D-SSA and KS-SSA. SSD was originally presented in Bonizzi *et al.* (2014) (presented in Section 4.5 herein), and independent trend analyzing SSA (ITA-SSD) is presented in the supplement which can dynamically adjust the isolation and extraction of an initial trend to prevent residual portions of trend obscuring the later decomposition and analysis.

In Section 5, by comparing Figs. 8 and 9 with Figs. 10 and 11, respectively, one can observe and compare the RCR parameter estimation against the original technique proposed in Hoff & Niu (2012). In addition to standard covariance regression, LASSO covariance regression, ridge covariance regression, elastic-net covariance regression, group-LASSO covariance regression, and subgradient covariance regression are also presented herein and are optional extensions in CovRegpy.py for algorithmic flexibility – see Sections 5.4, 5.5, 5.6, 5.7, and 5.8. This RCR algorithm and the algorithmic variations developed in this package are presented within the context of providing actuarial practitioners with a statistical framework for predicting the likelihood of future debilitating lung conditions which can be easily, and cheaply, treated with genetic testing and early-stage treatment compared against late-stage diagnosis and costly medical treatment and end-of-life care – this is discussed in Sections 5.1 and 5.2.

In Section 6, we present a case study which first examines the shortcomings of both purely funded and purely unfunded pensions – see Section 6.1. After addressing the shortcomings of the separate strategies, we note that many of the more-developed countries would not only enrich their citizens by adopting hybrid pension schemes as suitable compromises between the secure returns of unfunded pensions and the higher returns of the funded pensions but would also address the impending drastic demographic changes and their negative effects on the currently prevailing purely unfunded pension schemes with aging populations and shrinking workforces.

We show, using a case study built on equally weighted sector indices of the FTSE 100 over a period of approximately 20 years, that purely funded RPP portfolios using high-, mid-, and low-frequency covariance forecasts outperform unfunded pensions (using the equally weighted FTSE 100 as a proxy) in terms of absolute returns, but the returns are significantly less certain. By constructing hybrid RPP portfolios using different risk aversion levels (see Sections 6.3 and 6.4) for each frequency of covariance forecast, we note the following:

- high-frequency covariance forecasts are more responsive to changes in the underlying inter-relationships between the indices;

- mid-frequency covariance forecasts are less responsive than the high-frequency forecasts, but more responsive than the low-frequency forecasts;
- during periods of strong GDP growth, the ideal strategy favors higher-frequency covariance forecasts with a higher proportion of the cumulative portfolio in the unfunded portion; and
- during periods of economic stagnation (lower or no GDP growth), a higher portion of the pension fund should be directed towards the funded portion which should utilize the lower-frequency covariance forecasts.

This framework has developed a robust approach to address the impending potentially catastrophic demographic changes in the more developed nations, the optimal funding ratio strategies during periods of varying economic growth, and the optimal frequency of structures to use when forecasting covariance for the funded portion of portfolios during different economic climates.

Additional algorithmic extensions and forecasting techniques are presented in Supplement to: Package CovRegpy: Regularized Covariance Regression and Forecasting in Python. One such technique, named IFF (instantaneous frequency forecasting), is presented in the supplement which provides a forecasting technique which uses the changes in a structure's instantaneous amplitude and instantaneous frequency to predict how its temporal structure continues.

Supplementary material. The supplementary material for this article can be found at <http://dx.doi.org/10.1017/S1748499524000101>.

Data Availability Statement. The data, Python code, figures, and other replication materials that support this study are openly available in CovRegpy at:

<https://zenodo.org/doi/10.5281/zenodo.10827714>.

A regularly maintained version catalog of the software package CovRegpy, as well as detailed installation instructions and examples for both experienced and new users, can be found here:

<https://github.com/Cole-vj/CovRegpy>.

Funding statement. There was no external funding.

Competing interests. None.

References

- Ames, M., Bagnarosa, G., Peters, G., & Shevchenko, P. (2018). Understanding the interplay between covariance forecasting factor models and risk-based portfolio allocations in currency carry trades. *Journal of Forecasting*, *37*(8), 8–831. doi: [10.1002/for.2505](https://doi.org/10.1002/for.2505) 2018.
- Ames, M., Bagnarosa, G., Peters, G., Shevchenko, P., & Matsui, T. (2017). Forecasting covariance for optimal carry trade portfolio allocations. In 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp.5910–5914). IEEE. doi:[10.1109/ICASSP.2017.7953290](https://doi.org/10.1109/ICASSP.2017.7953290).
- Basu, D., & Chawla, D. (2012). An empirical test of the arbitrage pricing theory — the case of Indian stock market. *Global Business Review*, *13*, 421–432. doi:[10.1177/097215091201300](https://doi.org/10.1177/097215091201300)
- Bauwens, L., Laurent, S., & Rombouts, J. (2006). Multivariate GARCH models: A survey. *Journal of Applied Econometrics*, *21*(1), 79–109. doi: [10.1002/jae.842](https://doi.org/10.1002/jae.842).
- Bianconcini, S. (2006). Trend-Cycle Estimation in Reproducing Kernel Hilbert Spaces. Ph. D. Dissertation. Department of Statistics, University of Bologna, Bologna.
- Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, *81*(3), 637–654. doi:[10.1086/260062](https://doi.org/10.1086/260062).
- Blommestein, H., Janssen, P., Kortleve, N., & Yermo, J. (2009). Moving beyond the DB vs. DC debate: The appeal of hybrid pension plans. *Rotman International Journal of Pension Management*, *2*(2). doi:[10.2139/ssrn.1493333](https://doi.org/10.2139/ssrn.1493333).

- Bonizzi, P., Karel, J., Meste, O., & Peeters, R. (2014). Singular spectrum decomposition: A new method for time series decomposition. *Advances in Adaptive Data Analysis*, 6(04), 1–34. doi:10.1142/S1793536914500113
- Boyd, S., Busseti, E., Diamond, S., Kahn, R., Koh, K., Nystrup, P., & Speth, J. (2016). Multi-period trading via convex optimization. *Foundations and Trends in Optimization*, 3(1), 1–76. doi: 10.1561/24000000023.
- Braun, S., doPico, G., Tsiatis, A., Horvath, E., Dickie, H., & Rankin, J. (1979). Farmer's lung disease: Long-term clinical and physiologic outcome. *American Review of Respiratory Disease*, 119(2), 185–191. doi:10.1164/arrd.1979.119.2.185.
- Casanova, C., Cote, C., de Torres, J., Aguirre-Jaime, A., Marin, J., Pinto-Plata, V., & Celli, B. (2005). Inspiratory-to-total lung capacity ratio predicts mortality in patients with chronic obstructive pulmonary disease. *American Journal of Respiratory and Critical Care Medicine*, 171(6), 591–597. doi:10.1164/rccm.200407-867OC.
- Chapman, R., Gordon, T., & Speed, C. (2001). Pensions, funding and risk. *British Actuarial Journal*, 7(4), 605–662. doi:10.1017/S1357321700002488.
- Cho, S., & Stout-Delgado, H. (2020). Aging and lung disease. *Annual Review of Physiology*, 82(1), 433–459. doi:10.1146/annurev-physiol-021119-034610.
- Dagum, E. (1980). The X-11-ARIMA Seasonal Adjustment Method. Number Catalogue 12-564E. Statistics Canada, Ottawa, Canada.
- Dagum, E. (1988). The X-11-ARIMA/88 Seasonal Adjustment Method: Foundations and User's Manual. Statistics Canada, Ottawa, Canada.
- Dagum, E. (1996). A new method to reduce unwanted ripples and revisions in trend-cycle estimates from X11ARIMA. *Survey Methodology*, 22(1), 77–83.
- Dagum, E., & Bianconcini, S. (2006). Local polynomial trend-cycle predictors in reproducing kernel hilbert spaces for current economic analysis. In *Anales de economia aplicada* (pp. 1–16).
- Dagum, E., & Bianconcini, S. (2008). The Henderson smoother in reproducing kernel Hilbert space. *Journal of Business & Economic Statistics*, 26(4), 536–545. doi: 10.1198/07350010700000322.
- Darbyshire, J. (2017). *Pricing and trading interest rate derivatives: A practical guide to swaps (revised ed.)*. Aitch & Dee Limited. <https://books.google.co.uk/books?id=gOx3tAEACAAJ>.
- Dash, A., & Grimshaw, D. (1993). Dread disease cover—An actuarial perspective. *Journal of the Staple Inn Actuarial Society*, 33(1), 149–193. doi:10.1017/S2049929900010564.
- Derman, E., & Kani, I. (1994). Riding on a smile - constructing binomial tree models that are consistent with the volatility smile effect. *Risk*, 7(2), 32–39.
- Devolder, P., & de Valeriola, S. (2019). Between DB and DC: Optimal hybrid PAYG pension schemes. *European Actuarial Journal*, 9(2), 463–482. doi: 10.1007/s13385-019-00216-y.
- Disney, R. (1996). *Can we afford to grow older? A perspective on the economics of aging* (1st ed.). MIT Press.
- Doherty, M. (2001). Applications: The surrogate Henderson filters in X-11. *Australian & New Zealand Journal of Statistics*, 43(4), 385–392. doi:10.1111/1467-842X.00187.
- Dumas, B., Fleming, J., & Whaley, R. (1998). Implied volatility functions: Empirical tests. *The Journal of Finance*, 53(6), 2059–2106. doi:10.1111/0022-1082.00083.
- Dupire, B. (1994). Pricing with a smile. *Risk*, 7(1), 18–20.
- Dutta, J., Kapur, S., & Orszag, J. (2000). A portfolio approach to the optimal funding of pensions. *Economics Letters*, 69(2), 201–206. doi:10.1016/S0165-1765(00)00271-8.
- Dyer, C. (2012). The interaction of ageing and lung disease. *Chronic Respiratory Disease*, 9(1), 63–67. doi:10.1177/1479972311433766.
- Dykstra, B., Scanlon, P., Kester, M., Beck, K., & Enright, P. (1999). Lung volumes in 4,774 patients with obstructive lung disease. *Chest*, 115(1), 68–74. doi:10.1378/chest.115.1.68.
- Efird, J., Landrine, H., Shiue, K., O'Neal, W., Podder, T., Rosenman, J., & Biswas, T. (2014). Race, insurance type, and stage of presentation among lung cancer patients. *SpringerPlus*, 3(1), 710–717. doi:10.1186/2193-1801-3-710.
- Fama, E., & French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3–56.
- Fama, E., & French, K. (2004). The capital asset pricing model: Theory and evidence. *Journal of Economic Perspectives*, 18(3), 25–46. doi:10.1257/0895330042162430.
- Fama, E., & French, K. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1–22. doi:10.1016/j.jfineco.2014.10.010.
- Feldstein, M. (1996). The Missing Piece in Policy Analysis: Social Security Reform. Technical Report NBER Working Paper, No. 5413. National Bureau of Economic Research, 1050 Massachusetts Avenue, Cambridge, Massachusetts, USA. <https://doi.org/10.3386/w5413>.
- Feldstein, M., & Samwick, A. (1998). The transition path in privatizing social security. In M. Feldstein (Ed.), *Privatizing social security* (pp. 215–264). University of Chicago Press, <http://www.nber.org/chapters/c6251>
- Findley, D., Monsell, B., Bell, W., Otto, M., & Chen, B. (1998). New capabilities and methods of the X-12-ARIMA seasonal-adjustment program. *Journal of Business & Economic Statistics*, 16(2), 127–152. doi:10.1080/07350015.1998.10524743.

- Fitch, K., Iwasaki, K., Pyenson, B., Plauschinat, C., & Zhang, J. (2011). Variation in adherence with global initiative for chronic obstructive lung disease (GOLD) drug therapy guidelines: A retrospective actuarial claims data analysis. *Current Medical Research and Opinion*, 27(7), 1425–1429. doi:10.1185/03007995.2011.583230.
- French, A., Balfe, D., Mirocha, J., Falk, J., & Mosenifar, Z. (2015). The inspiratory capacity/total lung capacity ratio as a predictor of survival in an emphysematous phenotype of chronic obstructive pulmonary disease. *International Journal of Chronic Obstructive Pulmonary Disease*, 2015, 1305–1312. doi:10.2147/COPD.S76739.
- Grubbs, D. (1999). The public responsibility of actuaries in American pensions. *North American Actuarial Journal*, 3(4), 34–41. doi:10.1080/10920277.1999.10595857.
- Hassani, H. (2007). Singular spectrum analysis: Methodology and comparison. *Journal of Data Science*, 5(2), 239–257. doi:10.6339/JDS.2007.05(2).396.
- Henderson, R. (1916). Note on graduation by adjusted average. *Transactions of the Actuarial Society of America*, 17, 43–48.
- Henderson, R. (1924). A new method of graduation. *Transactions of the Actuarial Society of America*, 25, 29–40.
- Hoevernaars, R., & Ponds, E. (2008). Valuation of intergenerational transfers in funded collective pension schemes. *Insurance: Mathematics and Economics*, 42(2), 578–593. doi:10.1016/j.insmathco.2007.06.007.
- Hoff, P., & Niu, X. (2012). A covariance regression model. *Statistica Sinica*, 22(2), 729–753.
- Huang, N. (1999). Computer Implemented Empirical Mode Decomposition Method, Apparatus and Article of Manufacture. Patent. US Patent 5,983,162.
- Huang, N., Shen, Z., & Long, S. (1999). A new view of nonlinear water waves: The Hilbert spectrum. *Annual Review of Fluid Mechanics*, 31(1), 417–457. doi:10.1146/annurev.fluid.31.1.417.
- Huang, N., Shen, Z., Long, S., Wu, M., Shih, H., Zheng, Q., Yen, N., Tung, C., & Liu, H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 454(1971), 903–995. doi:10.1098/rspa.1998.0193.
- Hutchinson, J. (1846). On the capacity of the lungs, and on the respiratory functions, with a view of establishing a precise and easy method of detecting disease by the spirometer. *Medico-Chirurgical Transactions*, 29(1846), 137–252. doi:10.1177/095952874602900113.
- Jacobs, B., & Levy, K. (1993). Long/Short equity investing - profit from both winners and losers. *Journal of Portfolio Management*, 20(1), 52–63.
- Jolliffe, I. (1986). *Principal component analysis* (2nd ed.). Springer-Verlag.
- Kotlikoff, L. (1996). Privatizing social security at home and abroad. *The American Economic Review*, 86, 368–372. Papers and Proceedings of the 108th Annual Meeting of the American Economic Association, San Francisco, California, USA. <https://www.jstor.org/stable/2118154>.
- Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The Journal of Finance*, 20(4), 587–615.
- Maillard, S., Roncalli, T., & Teiletche, J. (2010). The properties of equally weighted risk contribution portfolios. *The Journal of Portfolio Management*, 36(4), 60–70. doi: 10.3905/jpm.2010.36.4.060.
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91. doi:10.2307/2975974.
- Martell, C., Kioko, S., & Moldogaziev, T. (2013). Impact of unfunded pension obligations on credit quality of state governments. *Public Budgeting & Finance*, 33(3), 24–54. doi:10.1111/j.1540-5850.2013.12013.x.
- Mateus, I., Mateus, C., & Todorovic, N. (2019). Review of new trends in the literature on factor models and mutual fund performance. *International Review of Financial Analysis*, 63, 344–354.
- Miles, D. (1998). The implications of switching from unfunded to funded pension systems. *National Institute Economic Review*, 163, 71–86. doi:10.1177/002795019816300109.
- Miles, D. (2000). Funded and Unfunded Pension Schemes: Risk, Return and Welfare. Technical Report Working Paper, No. 239. Center for Economic Studies and ifo Institute (CESifo), CESifo GmbH, Poschingerstr. 5, Munich, Germany. <http://hdl.handle.net/10419/75591>.
- Miles, D., Timmermann, A., de Haan, J., & Pagano, M. (1999). Risk sharing and transition costs in the reform of pension systems in Europe. *Economic Policy*, 14(29), 253–286. <http://www.jstor.org/stable/1344689>.
- Mitchell, O., & Zeldes, S. (1996). Social Security Privatization: A Structure for Analysis. Technical Report NBER Working Paper, No. 5512. National Bureau of Economic Research, 1050 Massachusetts Avenue, Cambridge, Massachusetts, U.S.A.
- Musgrave, J. (1964a). Alternative Sets of Weights for Proposed X-11 Seasonal Factor Curve Moving Averages. Working paper. U.S. Bureau of the Census, U.S. Department of Commerce, Washington D.C., USA.
- Musgrave, J. (1964b). A Set of End Weights to End All End Weights. Working paper. U.S. Bureau of the Census, U.S. Department of Commerce, Washington D.C., USA.
- Pasini, G. (2017). Principal component analysis for stock portfolio management. *International Journal of Pure and Applied Mathematics*, 115(1), 153–167. doi:10.12732/ijpam.v115i1.12.
- Pyenson, B., Henschke, C., Yankelevitz, D., Yip, R., & Dec, E. (2014). Offering lung cancer screening to high-risk medicare beneficiaries saves lives and is cost-effective: An actuarial analysis. *American Health & Drug Benefits*, 7(5), 272–282.
- Pyenson, B., Sander, M., Jiang, Y., Kahn, H., & Mulshine, J. (2012). An actuarial analysis shows that offering lung cancer screening as an insurance benefit would save lives at relatively low cost. *Health Affairs*, 31(4), 770–779. doi:10.1377/hlthaff.2011.0814.

- Qian, E.** (2005). Risk Parity Portfolios: Efficient Portfolios Through True Diversification. White paper. Panagora Asset Management, Boston, Massachusetts, USA.
- Roseveare, D., Leibfritz, W., Fore, D., & Wurzel, E.** (1996). Ageing Populations, Pension Systems and Government Budgets: Simulations for 20 OECD Countries. Technical Report Working Papers, No. 168. Organisation for Economic Co-operation and Development. <https://doi.org/10.1787/463240307711>.
- Ross, S.** (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, **13**(3), 341–360.
- Rubinstein, M.** (1994). Implied binomial trees. *The Journal of Finance*, **49**(3), 771–818. doi:10.1111/j.1540-6261.1994.tb00079.x.
- Sharpe, W.** (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, **19**(3), 425–442. doi: 10.1111/j.1540-6261.1964.tb02865.x.
- Shiskin, J., Young, A., & Musgrave, J.** (1967). The X-11 Variant of the Census Method II Seasonal Adjustment Program. Technical Report 15. U.S. Department of Commerce, Washington D.C.
- Sinn, H.** (1999). Pension Reform and Demographic Crisis: Why a Funded System is Needed and Why it is Not Needed. Technical Report Working Paper, No. 195. Center for Economic Studies and ifo Institute (CESifo), CESifo GmbH, Poschingerstr. 5, Munich, Germany.
- Smith, A.** (1776). An Inquiry into the Nature and Causes of the Wealth of Nations.
- Splinter, D.** (2017). State pension contributions and fiscal stress. *Journal of Pension Economics & Finance*, **16**(1), 65–80. doi:10.1017/S1474747215000189.
- Sutcliffe, A.** (1993). *X11 time series decomposition and sampling errors*. Australian Bureau of Statistics.
- Tobin, J.** (1958). Liquidity preference as behavior towards risk. *The Review of Economic Studies*, **25**(2), 65–86.
- Trowbridge, C.** (1952). Fundamentals of pension funding. *Transactions of the Society of Actuaries*, **4**(17), 17–43.
- van Jaarsveldt, C., Ames, M., Peters, G., & Chantler, M.** (2023a). Package advEMDpy: Algorithmic variations of empirical mode decomposition in Python. *Annals of Actuarial Science*, **2023**(3), 1–37. doi:10.1017/S1748499523000088.
- van Jaarsveldt, C., Peters, G., Ames, M., & Chantler, M.** (2023b). Tutorial on empirical mode decomposition: Basis decomposition and frequency adaptive graduation in non-stationary time series. *IEEE Access*, **11**, 94442–94478. doi:10.1109/ACCESS.2023.3307628.
- Vautard, R., Yiou, P., & Ghil, M.** (1992). Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. *Physica D: Nonlinear Phenomena*, **58**(1-4), 95–126. doi:10.1016/0167-2789(92)90103-T.
- Wang, S., & Lu, Y.** (2019). Optimal investment strategies and risk-sharing arrangements for a hybrid pension plan. *Insurance: Mathematics and Economics*, **89**, 46–62. doi:10.1016/j.insmatheco.2019.09.005.
- Whittaker, E.** (1922). On a new method of graduation. *Proceedings of the Edinburgh Mathematical Society*, **41**, 63–75. doi:10.1017/S0013091500077853.
- Yang, H., Beymer, M., & Suen, S.** (2019). Chronic disease onset among people living with HIV and AIDS in a large private insurance claims dataset. *Scientific Reports*, **9**(18514), 1–8. doi:10.1038/s41598-019-54969-3.
- Yang, L.** (2015). An Application of Principal Component Analysis to Stock Portfolio Management. Master's Thesis. University of Canterbury, Department of Economics and Finance, 20 Kirkwood Avenue, Upper Riccarton, Christchurch 8041, New Zealand. <https://doi.org/10.26021/5235>.