Predicting Pollution Risk Using Asymmetric GARCH-DCC Models

Passamani Giuliana[1]; Masotti Paola[2]

[1] Department of Economics and Management, University of Trento, Italy.

e-mail address: giuliana.passamani@unitn.it

[2] Department of Economics and Management, University of Trento, Italy.

e-mail address: paola.masotti@unitn.it

Abstract:

Exposure to high levels of pollution represents a major risk factor, as it may produce adverse health effects documented by numerous studies (e.g. Fuller et al. 2022). In order to work towards reaching SDG target 3.9.1, that aims to reduction in illnesses and deaths attributed to ambient air pollution, we need to define a model for measuring and predicting risk associated to exposure. Given that ambient air pollution is the outcome of complex mixtures of air pollutants emitted from various activities, an approximation of their combined effects and of impacts on health is possible if we could assume some form of independence and little correlation between the pollutants. However, there are some limitations in estimating these joint effects given nonlinear interactions among pollutants and their impacts. For the paper purpose, we consider the extension of time varying volatility models for time series data, to dynamic multivariate regression models, in which the diagonal elements of the conditional covariance matrix of the errors are modelled as univariate GARCH models, whereas the off-diagonal elements are modelled as nonlinear dynamic functions of the diagonal terms and of the conditional quasi-correlations. In other words, for measuring and predicting pollution risk we use GARCH-Dynamic Conditional Correlation (DCC) models (Engle, 2002) developed for measuring and hedging financial risk. The data set consists of daily standardized concentrations, over two years, on three pollutants, PM_{10}, NO_2 and O_3, which are interrelated and represent the so-called photochemical pollution factor. The three variables are observed at a single urban monitoring site. Given the non-stationarity in the mean of the observed variables, their stochastic trends are estimated using a smooth-trend unobserved component model and we use these estimated trends to de-trend these variables to make them stationary. As observed, pollutants concentrations show the presence of significant and different GARCH effects. The objective of this paper is to explore whether the use of a multivariate asymmetric GARCH-DCC model can lead to a more accurate risk prediction for air pollution. In particular, we aim to determine how positive shocks to the observed pollutants can increase health risk. Interesting results emerge for particulate matter and ozone, both of which have great effects on human health.

1. Introduction

Governments, international organizations and NGOs continuously emphasize that air pollution poses a serious threat to health and climate all over the world. The health risks of air pollution are particularly severe as WHO shows: air pollution is a risk for all-cause mortality as well as specific diseases. EPA and EEA, the United States and European environmental protection agencies, collaborate with a wide variety of multilateral organizations to protect human health and the environment. Even if the EEA's European Air Quality Index allows people to know the air quality of individual countries, regions and cities in Europe, it reflects the potential impact on health of the single pollutant for which concentrations are poorest due to associated health impacts. They observe up to five key pollutants at each monitoring station, but only the poorest value on the single one determines the index and no relation among different pollutants is taken into account in the index.

We know that correlations could represent critical information for determining the appropriate air quality value when more pollutants show poor values, because their interaction could make a worst quality of air than predicted by the single pollutant, and therefore, a riskier health condition. In

addition, observed volatilities could represent important information given that high variability in the observed values contribute to uncertainty in the determination of the air quality index.
Borrowing from the financial literature, in this paper we focus on a class of Multivariate Generalized AutoRegressive Conditional Heteroskedasticity (MGARCH) models for which we can estimate dynamic conditional correlations. Originally proposed by Engle (2002), this methodological approach has seen a variety of developments and applications for hedging financial risks. Our aim is to study correlations between key gaseous pollutants to identify linked trends as a means of better understanding the impact of pollution on health risk, and of dynamic conditional correlations given the observed concentrations on various pollutants at the urban monitoring site that we consider.

2. Methodology

MGARCH models are dynamic multivariate regression models in which the squared errors in each equation follow an autoregressive-moving-average structure, that is, they assume that the conditional variance of each univariate error at time *t* depends on the squared errors at the previous time plus the lagged conditional variance itself.
The Dynamic Conditional Correlation (DCC) MGARCH model uses a nonlinear combination of univariate GARCH models with time-varying cross-equation weights to model the conditional covariance matrix of the errors.
We give a formal definition of the general multivariate GARCH model and of dynamic conditional correlations to establish notation that facilitates comparisons of models. We can write the general MGARCH model as follows:

$$\boldsymbol{y}_t = \boldsymbol{\Phi}\boldsymbol{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t = \boldsymbol{H}_t^{1/2}\boldsymbol{\upsilon}_t, \tag{1}$$

where $\boldsymbol{y}_t$ is a $(p\times 1)$ vector of stationary observed variables $\boldsymbol{\varepsilon}_t$ is a $(p\times 1)$ vector of errors, $\boldsymbol{\Phi}$ is a $(p\times p)$ matrix of autoregressive parameters and $\boldsymbol{H}_t$ represents the $(p\times p)$ conditional covariance matrix based on information known the previous period, that is:

$$E_{t-1}\left(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t^{'}\right) \equiv \boldsymbol{H}_t. \tag{2}$$

and $\boldsymbol{H}_t^{1/2}$ is the Choleski factor of the time-varying conditional covariance matrix $\boldsymbol{H}_t$. This assumption implies that contemporaneous variances and covariances may be time varying, depending on past information.
MGARCH models differ in the parsimony and flexibility of their specification for the conditional covariance matrix of the errors, $\boldsymbol{H}_t$. As suggested in Engle (2002), in the conditional correlation family of MGARCH models, the diagonal elements of $\boldsymbol{H}_t$ are modeled as univariate asymmetric GARCH models, whereas the off-diagonal elements are modeled as nonlinear functions of the diagonal terms, that is:

$$\boldsymbol{H}_t = \boldsymbol{D}_t^{1/2}\boldsymbol{R}_t\boldsymbol{D}_t^{1/2}, \quad \boldsymbol{D}_t^{1/2} = diag\left\{\sqrt{h_{ii,t}}\right\}, \tag{3}$$

where $\boldsymbol{R}_t$ is a $(p\times p)$ correlation matrix containing the time varying conditional correlations, and $\boldsymbol{D}_t$ is a $(p\times p)$ diagonal properly defined matrix in which each element on the main diagonal evolves according to a univariate asymmetric GARCH model.
$\boldsymbol{R}_t$ is a matrix of conditional quasi-correlations defined as follows:

$$\boldsymbol{R}_t = diag(\boldsymbol{Q}_t)^{-1/2}\boldsymbol{Q}_t\ diag(\boldsymbol{Q}_t)^{-1/2} \tag{4}$$

$$\boldsymbol{Q}_t = (1-\lambda_1-\lambda_2)\bar{\boldsymbol{Q}} + \lambda_1\tilde{\boldsymbol{\varepsilon}}_{t-1}\tilde{\boldsymbol{\varepsilon}}_{t-1}^{'} + \lambda_2\boldsymbol{Q}_{t-1} \tag{5}$$

where $\boldsymbol{Q}_t$ is a symmetric positive definite matrix, $\tilde{\boldsymbol{\varepsilon}}_t = \boldsymbol{D}_t^{-1/2}\boldsymbol{\varepsilon}_t$ is a vector of standardized errors, λ_1 and λ_2 are parameters that govern the dynamics of conditional quasi-correlations, nonnegative and satisfying the restriction $0 \le \lambda_1 + \lambda_2 < 1$. $\bar{\boldsymbol{Q}}$ is the unconditional correlation matrix of the standardized errors $\tilde{\varepsilon}_t$.

The DCC-GARCH model has gained popularity because its parameters can be estimated using a two-stage approach: in the first stage, the parameters of the univariate GARCH models are estimated separately, for each of the *p* variables, and the estimate of the conditional covariance matrix $\boldsymbol{H}_t$ is thus be obtained. In the second stage, residuals transformed by their estimated standard deviations are used to estimate the parameters of the correlation part conditioning on the parameters estimated in the first stage. Matrix $\boldsymbol{H}_t$ could be useful also for discussing volatility spillover effects as in Chang, McAleer and Zou (2017).

3. Empirical results

In the empirical analysis, we consider two years of daily observations, at a single urban monitoring site (Trento – Parco S. Chiara), on the three main pollutants that, according to European Citeair index directives, represent the mandatory pollutants for calculating any background pollution index: particulate matter of dimension less than or equal to 10, nitrogen dioxide and ozone. These make the vector $\boldsymbol{y}_t' = (\mathrm{PM}_{10},\ \mathrm{NO}_2,\ \mathrm{O}_3)_t$.

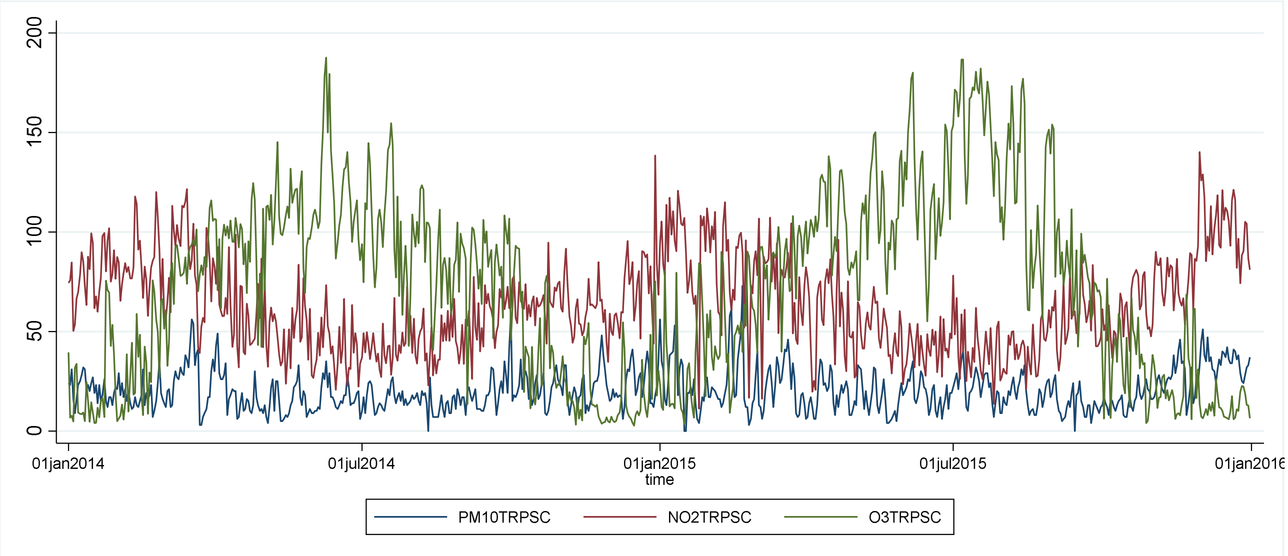

Figure 1: Daily observations, over 2014 and 2015, for the pollutants: PM_{10} (navy), NO_2 (maroon) and O_3 (green), at Trento PSC monitoring site.

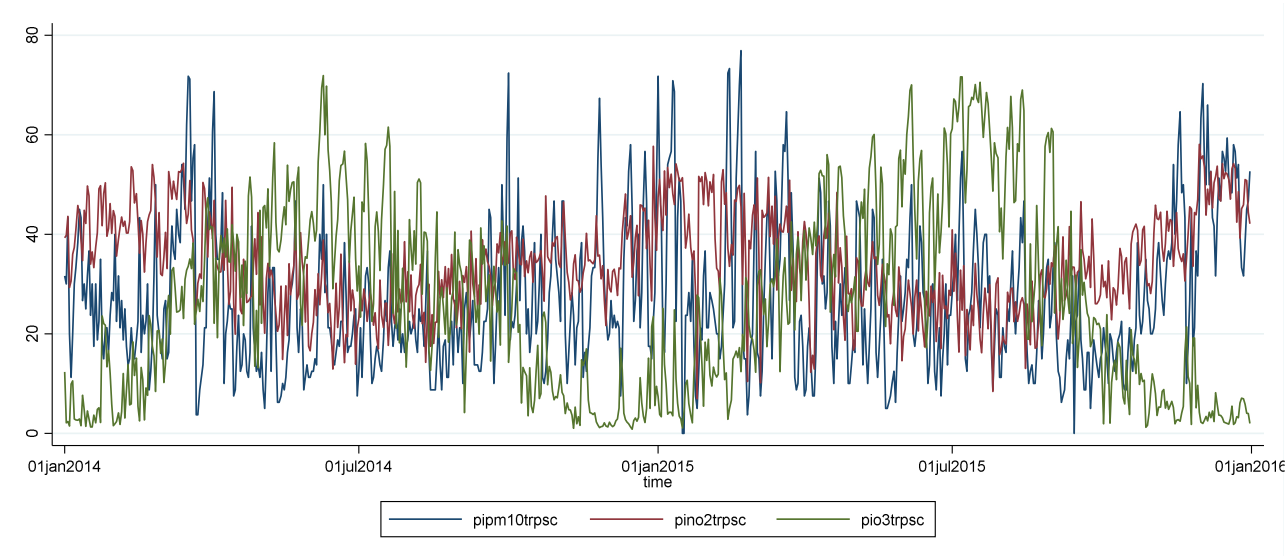

Figure 2: Daily transformed observations, over 2014 and 2015.

As we can observe in Fig. 1, the three variables, measured in terms of micrograms per cubic meter (µg/m3), show some non-stationarity at least in the mean values. Moreover, their measurement range is different due to their gaseous characteristics.
To transform air pollutant concentrations into comparable values in the range [0, 100], we use an algorithm involving piecewise linear functions, as in Murena (2004). We represent the transformed daily observations in Figure 2.
In order to deal with the problem of non-stationarity of the observed variables we estimate an unobserved component model for each time series: the trend component is obtained by applying a single exponential smoothing procedure. The new de-trended variables result from the difference between the observed and the estimated corresponding trend values and are represented in Figure 3. Based on the ADF test the de-trended variables reject the null hypothesis of a unit roots at the 1% level.

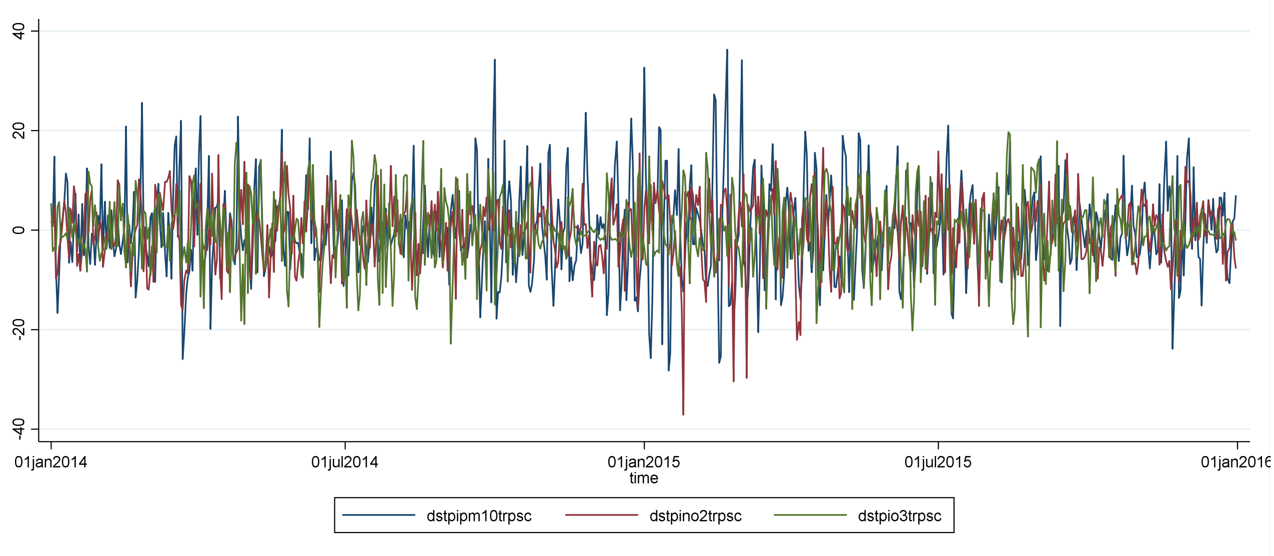

Figure 3: Detrended daily transformed observations, over 2014 and 2015.

The interesting feature of these variables is that they enhance the time varying variability and also some volatility clustering of the pollutants. Therefore, we can analyse their conditional variability using a multivariate model, which estimates their dynamic conditional correlations using an approach that has the flexibility of univariate asymmetric GARCH, but not the complexity of conventional MGARCH (Sadorsky, 2012). These models, introduced in paragraph 2, are naturally estimated in two steps: the different univariate asymmetric GARCH model in the first, and the dynamic conditional correlations in the second. They are estimated by Quasi Maximum Likelihood Estimation (QMLE), assuming that υ_t follow a multivariate *t* distribution.

	Coeff.	St.Err.	*t*-value	p-value	Sig
λ_1	.045	.009	5.15	0	***
λ_2	.939	.014	67.19	0	***
df	8.826	1.365	6.47	0	***

*** *p<.01*, ** *p<.05*, * *p<.1*

Table 1: DCC parameters estimates

From the estimates in Table 1, we can observe that, for the DCC part of the model, the estimated coefficients λ_1 and λ_2 are each positive and statistically significant at 1% level. Their sum is a value that is less than one, though quite close: it means that the dynamic conditional correlations are mean reverting. In particular, the long-term parameter λ_2 is rather high, showing a high persistence in the dynamic correlation. The dynamic conditional correlations

Figure 4 shows time varying conditional correlations from the DCC model for each pair of series: in the top left, the correlations between PM_{10} and NO_2; in the top right, the correlations between NO_2 and O_3; in the bottom left, the correlations between PM_{10} and O_3; in the bottom right, we represent

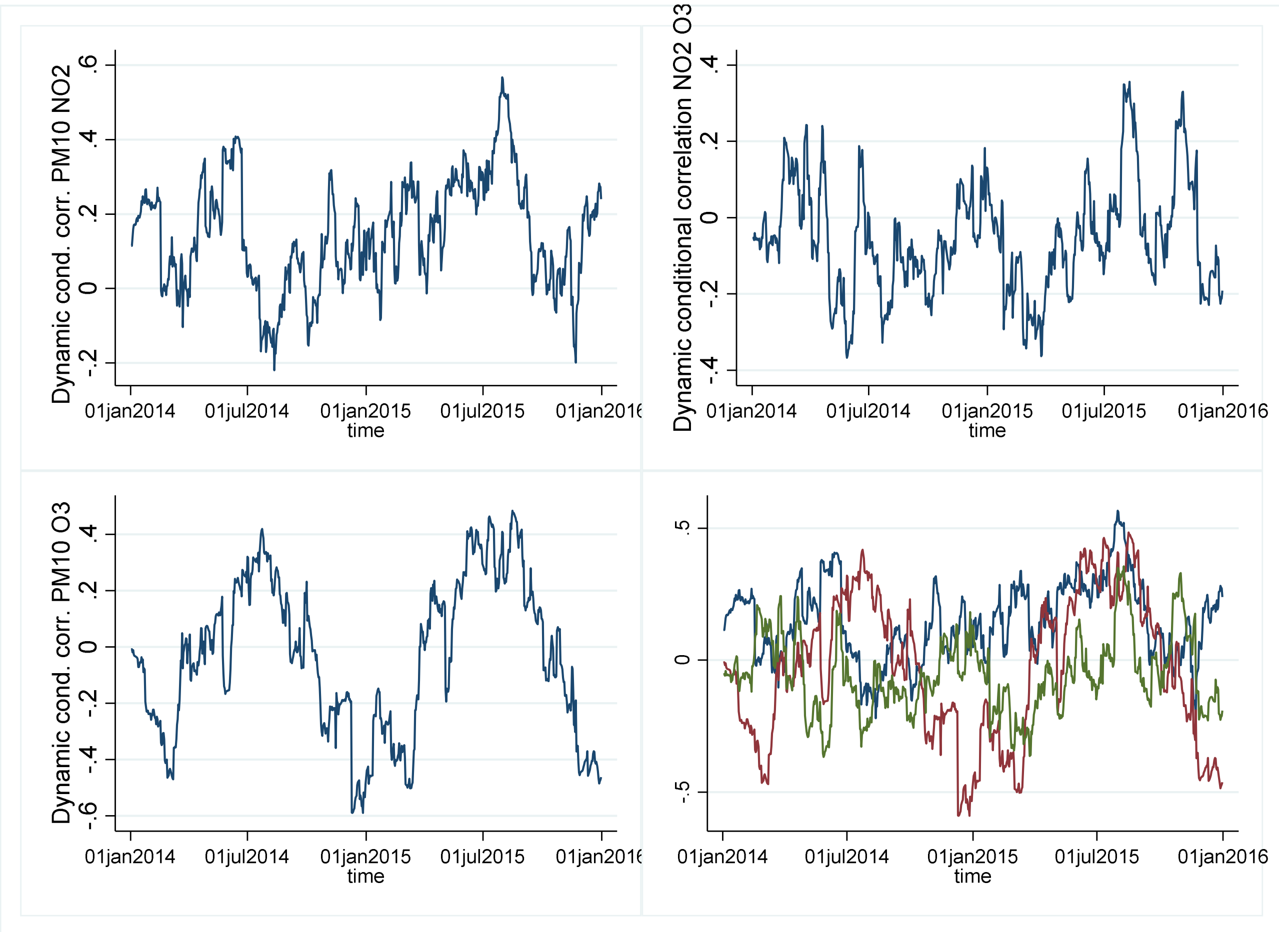

Figure 4: Estimated dynamic conditional correlations between pairs of pollutants.

the three series together. The conditional correlations vary a lot over the period, emphasizing the need to compute the dynamic ones, whereas the unconditional would be not significant.

These graphs show that dynamic conditional correlations reached high positive values in summer 2015, as well as in the first half of 2014. We can observe an upward trend in the correlations between PM_{10} and NO_2 and a clear seasonal pattern in correlations between PM_{10} and O_3. Moreover, there are several peaks in the correlations between NO_2 and O_3.

What should concern us most are the positive high correlations, while the negative correlations are signals that pollution conditions are less worrying.

This is a clear indication that there are spillover effects between pollutants and, when measuring pollution, we should consider them: therefore, we should look at the combined effects between pollutants and not just at the single pollutant value.

Among the results of the procedure adopted for estimating matrix $\boldsymbol{R}_t$, we obtain the estimated matrix $\boldsymbol{H}_t$, which conveys important information on the time varying conditional variances and covariances of the observed pollutants. This information can be used for calculating, for each pollutant, a new time series containing what in finance are called values at risk that are very popular for risk measurement and management.

In order to calculate these values at risk, we take the estimated conditional variances and compute the upper limit of a time varying confidence interval for the observed pollutant. The time series resulting from the computation of these upper limits represent a series of threshold values that can be calculated at appropriate confidence levels. In Figure 5, we represent the observed PM_{10} and the computed corresponding values at risk at 80% confidence level. This approach could then be used for forecasting future risk values for pollutants.

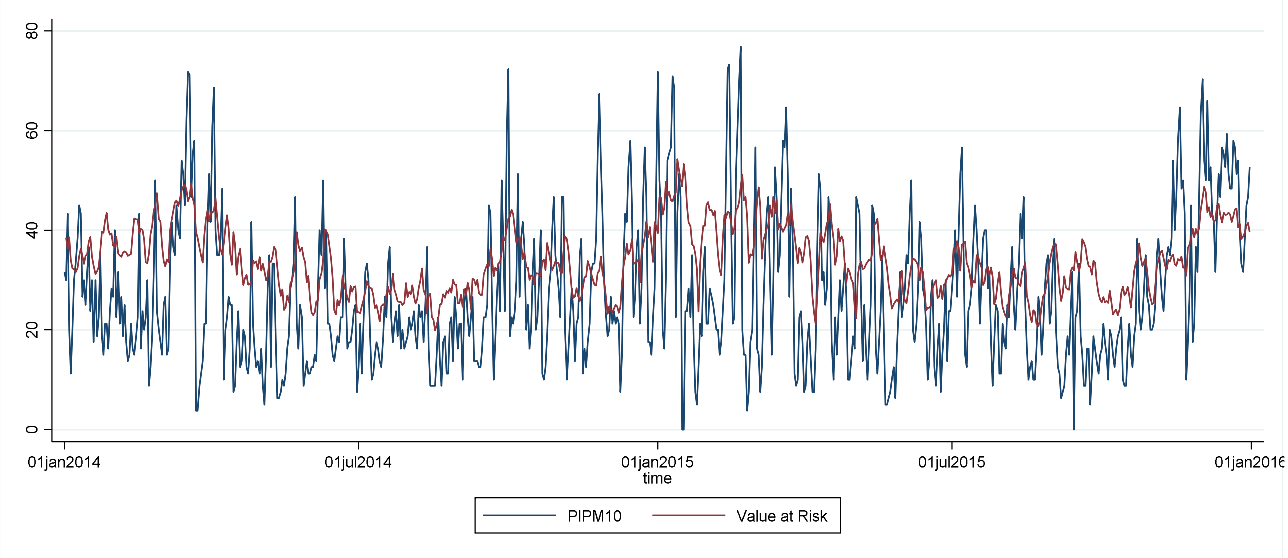

Figure 5: Estimated 80% values at risk for PM_{10}.

4. Conclusions

With the increase in global environmental pollution, it is important to have a better understanding of the volatility characteristics of the pollutants as well as their correlations over time with other pollutants. Our results emphasize the need to compute dynamic conditional covariances and conditional correlations in order to have a deeper knowledge of the pollution phenomenon. Another promising analytical approach that we could borrow from financial methodology, for future research, is a model for computing systemic pollution risk (Adrian and Brunnermeier, 2016), given that pollutants have reached particularly high values at risk.

References

Adrian T., Brunnermeier M. K. (2016). "CoVaR." *American Economic Review*, 106 (7): 1705-41.

Chang C-L, McAleer M., Zuo G. (2017). Volatility Spillovers and Causality of Carbon Emissions, Oil and Coal Spot and Futures for the EU and USA. Sustainability, 9 (10): 1789.

Engle R. F. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* 20 (3): 339–350.

Fuller R., Landrigan P., Balakrishnan K., Bathan G., Bose-O'Reilly S., Brauer M., Caravanos J., Chiles T., Cohen A., Corra L., Cropper M., Ferraro G., Hanna J., Hanrahan D., Hu H., Hunter D., Janata G., Kupka R., Lanphear B., Lichtveld M., Martin K., Mustapha A., Sanchez-Triana E., Sandilya K., Schaefli L., Shaw J., Seddon J., Suk W., Téllez-Rojo M. M., Yan C. (2022). Pollution and health: a progress update. *The Lancet Planetary Health*, 6 (6): e535-e547.

Murena F. (2004), Measuring air quality over large urban areas: development and application of an air pollution index at the urban area of Naples, *Atmospheric Environment*, 38 (36): 6195-6202.

Sadorsky P. (2012). Correlations and volatility spillovers between oil prices and the stock prices of clean energy and technology companies. *Energy Economics*, 34 (1), 248-255.