Journal of Statistical and Econometric Methods, vol. 2, no.3, 2013, 57-73
ISSN: 2051-5057 (print version), 2051-5065(online)
Scienpress Ltd, 2013

EGARCH, GJR-GARCH, TGARCH, AVGARCH, NGARCH, IGARCH and APARCH Models for Pathogens at Marine Recreational Sites

Ghulam Ali[1]

Abstract

The environmental literature lacks the use of volatility based models for environmental stochastic processes. To overcome this deficiency, we use EGARCH, IGARCH, TGARCH, GJR-GARCH, NGARCH, AVGARCH and APARCH models for functional relationships of the pathogen indicators time series for recreational activates at beaches. We use generalized error, Student's t, exponential, normal and normal inverse Gaussian distributions along with their skewed versions to model pathogen indicator time series. Generally speaking, turbidity, rainfall, dew point, river flow and cloud cover are significant variables. EGARCH, TGARCH, NAGARCH and AVGARCH are not radically different from each other in their output. However, TGARCH could be marginally better than the rest of models in capturing response of the pathogen indicator variable. Evidence supports some sign bias effect of the shocks. Dry weather and wet weather conditions of the same magnitude seem to have disproportionate effect on pathogens. Nyblom test shows that the estimated parameters are stable.

Mathematics Subject Classification: 62H12, 62M10, 62P10, 62P12.
Keywords: TGARCH, GJR GARCH, EGARCH, APARCH, Generalized Error Distribution, Student's t, normal, pathogen indicators, Huntington Beach.

1 Introduction

Response variables are not only affected by exogenous variables but also by themselves from their past behavior. On the basis of this theoretical underpinning, autoregressive models have been invented. Box and Jenkins time series modeling is indispensable in

[1]Economist, United States Environmental Protection Agency, Office of Water, Office of Science and Technology, Standards and Health Protection Division, 1200 Pennsylvania Avenue, NW, Washington, DC. 20460, USA. The views expressed herein are those of the author and may not be attributed to the organization of his employment or anyone else associated with it, in any way.

Article Info: *Received :* May 11, 2013. *Revised :* June 9, 2013.
Published online : September 9, 2013

analyzing stochastic processes. Autoregressive and moving average models are used frequently by many disciplines. In environmental science, they have been rarely, if ever, used in reference to the autoregressive variance and the mean of the distribution. What are even rarer in use are the various forms of the specification of the error terms in the estimation of the autoregressive models. Typically, normal distribution is used for the error terms of the stochastic equation; it is a mainstay of the environmental econometrics. The autoregressive framework is used in macroeconomics, such as for money supply, interest rate, price, inflation, exchange rates and gross domestic product. The autoregressive heteroscedastic modeling framework is used in financial economics, such as asset pricing, portfolio selection, option pricing, and hedging and risk management. There are a number of studies in the financial literature about modeling the return on stocks. Usually, in the financial market, upward movements in stock prices are followed by lower volatilities, while negative movements of the same magnitude are followed by much higher volatilities. Engle [1] developed the time varying variance model. Bollerslev [2] extended the model to include the ARMA structure. Since then, a number of studies have adopted the autoregressive conditional heteroscedastic (ARCH) or a generalized autoregressive conditional heteroscedastic (GARCH) framework to explain volatility of the stock market. Bad and good news both seem to increase the volatility of the stock market. Large changes follow the large changes and smaller changes follow the small changes in the stock market. Negative shocks have a much larger effect on stock pricing than positive shocks of the same magnitude. The negative shock has a long lasting impact, causing the stock market to take a long time to recover to the pre-shock level, after only a few days crash. This shows that symmetric distribution or normal distribution is not always a realistic assumption.

Exponential distribution was used by Nelson [3] for the U.S. stock market returns. Hsieh [4], Theodossiou [5] and Koutmos and Theodossiou [6] used it for foreign exchange rates. Akgirary et al [7] applied it for the distribution of prices of precious metal. This shows that the assumption of normal distribution has been relaxed in modeling the effect of volatility. Gallant, Hsieh and Tauchen [8] adopted the non-normal distribution for the financial analysis. Jun Yu [9] and Siourounis [10] also preferred the non-normal distribution. McMillan, Speight and Ap Gwilyn [11] had symmetric and asymmetric densities for the United Kingdom stock market. Fernandez and Steel [12] used the skewed Student's t distribution. Lambert and Laureen [13] used it in the GARCH framework. Bollerslev [14] and Baillie and Bollerslev [15] used the Student's t distribution to model the foreign exchange rate. Harris, Kucukozmen and Yilmaz [16] used the skewed generalized Student's t distribution to capture stylized facts (skewness and leverage effects) of daily returns. Ding, Granger and Engle [17] use the asymmetric power autoregressive conditional heteroscedastic (APARCH) model using Standard and Poor's data.

Negative correlation between the shocks and the return is a salient feature of the stock market. The sign and the magnitude of the shocks have asymmetric effects on returns. Therefore, Glosten, Jagannathan and Runkle (GJR) [18] introduced GARCH with differing effects of negative and positive shocks taking into account the leverage phenomenon. Due to asymmetric effects, skewed distributions are used in modeling stock returns. The above mentioned very brief review shows that financial market analysis has been extended to incorporate various distributions. Fat-tail distributions are used to represent the stylized facts of the stock market. Alberg, Shalit and Yosef [19] showed that the GARCH models with fat-tail distributions are relatively better suited for analyzing returns on stocks.

Fecal indicator bacteria (FIB) density can fluctuate rapidly. Dry weather and wet weather of the same magnitudes may have disproportionate or unequal effects on pathogen indicators. External shocks can dramatically influence the indicator bacteria. Therefore, models that focus on volatility might be useful for these time series. Ali [20] pioneered the use of ARCH and GARCH models for issuing beach advisories for pathogen indicators, in the environmental literature. However, he used the symmetric ARCH and GARCH models. In this article, we relax the symmetry assumption. We use the asymmetric and fat tail distributions because they have an advantage in representing the volatile time series (Alberg, Shalit and Yosef [19]). In addition, the models such as EGARCH, GJR GARCH, AVGARCH, TGARCH and APARCH (asymmetric power autoregressive conditional heteroscedastic models), despite their application for time series, have not been used in the environmental literature. As a result, this study includes them, bridging the gap in the literature. Generalized Error Distribution for GARCH modeling has not been used at all by any study in the environmental literature. Neither have Student-t, exponential or negative inverse Gaussian been used for modeling pathogen indicators with ARCH and GARCH models. This study fills this vacuum in the environmental literature. It incorporates the skewed versions of the distributions, such as the skewed normal and skewed Student's in modeling the pathogen indicator time series.

In sum, we will use the most commonly used models in the financial literature. As was pointed out earlier, we consider EGARCH of Nelson [3], TGARCH of Zakoian [21], APARCH of Ding, Engle and Granger [17], GJR-GARCH of Glosten, Jagannathan and Runkle [18], AVGARCH of Taylor [22] and NGARCH of Higgins and Bera [23] for modeling the pathogen indicators. We will use the various distributions to estimate these models, which might make this study rather unique in the environmental literature.

The rest of the article is structured as follows. The models are presented in section 2, while the data and site description are given in section 3. Section 4 discusses output of the models and section 5 presents the conclusions.

2 Models

The autoregressive model with exogenous variable is expressed as:

$$y_t = \phi(L)y_t + x_t\vec{\beta} + \varepsilon_t. \tag{1}$$

with y and x dependent and exogenous variables respectively, and ε for the error term.

2.1 IGARCH

GARCH models apply both an autoregressive and moving average structure to the variance, σ^2. The integrated GARCH (IGARCH) is specified as

$$\varepsilon_t = \sigma_t z_t;\ \ \sigma^2 = \omega + \sum_{i=1}^{p} \alpha_i {\varepsilon_{t-i}}^2 + \sum_{j=1}^{q} \beta_j\ {\sigma^2}_{t-j}. \tag{2}$$

The sum of coefficients is restricted to 1. The exogenous variable can be easily reflected in the various specifications of GARCH models just by addition of $x_t\vec{\beta}$.

2.2 EGARCH

The exponential GARCH (EGARCH) may generally be specified as

$$\varepsilon_t = \sigma_t z_t;\ \ ln\sigma^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}{}^2 + \sum_{j=1}^{q} \beta_j\ \ ln\sigma^2{}_{t-j}\,. \tag{3}$$

This model differs from the GARCH variance structure because of the log of the variance. The following specification also has been used in the financial literature (Dhamija and Bhalla [24]).

$$\varepsilon_t = \sigma_t z_t;\ \ ln\sigma^2 = \omega + \alpha_i \varepsilon_{t-i}{}^2 + \sum_{j=1}^{q} \lambda_j \ln\left(\sigma^2{}_{t-j}\right) + \sum_{i=1}^{p} \gamma_i \left(\frac{|\varepsilon_{t-i}|}{\sigma_{t-i}} - \sqrt{\frac{2}{\pi}}\right). \tag{4}$$

2.3 AVGARCH

An asymmetric GARCH (AGARCH) is simply

$$\varepsilon_t = \sigma_t z_t;\ \ \sigma^2 = \omega + \sum_{i=1}^{p} \alpha_i |\varepsilon_{t-i} - b|^2 + \sum_{j=1}^{q} \beta_j\ \ \sigma^2{}_{t-j}\,. \tag{5}$$

The absolute value generalized autoregressive conditional heteroscedastic (AVGARCH) model is specified as

$$\varepsilon_t = \sigma_t z_t;\ \ \sigma^2 = \omega + \sum_{i=1}^{p} \alpha_i (|\varepsilon_{t-i} + b| - c(\varepsilon_{t-i} + b))^2 + \sum_{j=1}^{q} \beta_j\ \ \sigma^2{}_{t-j}\,. \tag{6}$$

2.4 GJR GARCH

The GJR GARCH model is represented by the expression

$$\zeta_t{}^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_t{}^2 + \sum_{j=1}^{q} \beta_j\ \ \sigma^2{}_{t-j} + \sum_{i=1}^{p} \gamma_i I_{t-i} \varepsilon^2{}_{t-i}. \tag{7}$$

where:

$$I_{t-i} = \begin{cases} 1 & if\ \varepsilon_{t-i} < 0 \\ 0 & if\, \varepsilon_{t-i} \geq 0. \end{cases}$$

2.5 TGARCH

The threshold GARCH (TGARCH) is similar to the GJR model, different only because of the standard deviation, instead of the variance, in the specification

$$\zeta_t{}^2 = \omega + \sum_{i=1}^{p} \alpha_i (|\varepsilon_{t-i}| - \gamma_i\, \varepsilon_{t-i}) + \sum_{j=1}^{q} \beta_j\ \ \zeta_{t-j}. \tag{8}$$

2.6 GARCH-M

The GARCH in mean (GARCH-M) incorporates the effect of the volatility of the series on the mean. The model is usually represented by the expression

$$x_t = \mu + k\sigma_t + \varepsilon_t; \quad \varepsilon_t = \sigma_t z_t, \text{ and } \sigma^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon^2 + \sum_{j=1}^{q} \beta_j \ \sigma^2{}_{t-j}. \qquad (9)$$

2.7 APARCH

APARCH represents a general class of models that include both ARCH and GARCH models. The Omnibus structure of this model is

$$\varepsilon_t = z_t \sigma_t; \ \sigma^2 = \omega + \sum_{i=1}^{p} \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^{\delta} + \sum_{j=1}^{q} \beta_j \ \sigma^{\delta}{}_{t-j}. \qquad (10)$$

Some of the models can be derived or estimated from this model. For example,

- TS-GARCH of Taylor and Stuart when $\delta = 1$ and $\gamma_i = 0$,
- GJR-GARCH when $\delta = 2$,
- T-ARCH of Zakoian when $\delta = 1$,
- Log ARCH when $\delta = 0$,
- N-ARCH of Higgins and Bera when $\delta = 0$ and $\beta_i = 0$.

3 Site Description

Huntington Beach is located in Ohio, United States. It is a relatively scenic beach on Lake Erie. It is located at latitude 41, 49111 and longitude -81.93472 near the city of Bay Village. It is a public beach situated in an urban setting. Wastewater is discharged east of the beach, into Lake Erie from the Rocky River treatment plant. Stormwater runoff from the beach parking lot is directly discharged by two outfalls in the vicinity of the beach. Sources of FIB are not very well differentiated for the Huntington Beach.

3.1 Data

A number of meteorological and environmental variable time series are available for this beach. These variables include bacterial densities of fecal indicators, wind speed, air temperature, atmospheric pressure, dew point, turbidity, Cuyahoga River flow, Rocky River flow, cloud cover, time of travel of flow and relevant area draining in the vicinity of the beach. However, this study will not consider all of the variables because some of them such as time of travel and acreage along with the Rocky River flow were not found to have any explanatory power by an earlier study (Ali [20]). Atmospheric pressure may not be all that relevant for the FIB densities, hence is not considered in this study. For modeling FIB densities, scientific explanation or scientific rationale of some of the variables is presented by Boehm et al [25]. FIB time series spans from 2006 to 2008, with 225 daily observations of swimming seasons to estimate the models.

4 Discussion of the Model Output

Various distributions such as Generalized Error, Student's t, normal, exponential and normal inverse Gaussian distribution are used for estimation of GARCH/ARCH models in this paper. We use skew normal, skew Student and skew Generalized Error distributions. Generalized Error and Student and their skewed versions have additional shape parameters, which are changed in the estimation of models. Thus, it allows estimation of the various versions of the same model. Although we are using various distributions, we will present results of the best fit only, ignoring the rest. For estimation, we use the R software (see Klien [26], Ghalanos [27] and Wurtz, Chalabi and Luksan [28].

4.1 TGARCH

For the TGARCH specification, a Generalized Error Distribution with skew parameter is assumed. The variables turbidity, rainfall, wind speed, temperature, dew point, and Cuyahoga River flow are included in the model estimation. The coefficients of turbidity, rainfall, temperature, dew point and Cuyahoga river flow are significant, because p-value falls below the critical value. The probability of the null hypothesis of the wind speed exceeds the threshold p-value (0.05). This shows that the coefficient of wind speed is statistically no different from zero. Similarly, alpha, the parameter in the variance equation, is no different from zero. However, the coefficient representing the skewness turns out to be significant.

We dropped wind speed from the equation. The coefficients of turbidity, rainfall, temperature, dew point and the river flow all remain significant. In addition, constant reversed from insignificant to significant in value. We do not find the wind speed to be relevant to the pathogen series. Therefore, we retain these variables in other models. Very interestingly, the variance equation is significant. Not only alpha but also beta is significant. Similarly, the coefficient for mean is significant. Moreover, the skew parameter of the distribution is significant. A skewed distribution seems to fit the data nicely. The output of the model is presented in Table 1.

4.2 Sign Bias Tests

Engle and Ng [29] proposed negative sign test, positive sign bias test, and joint bias test for volatility of the process. The negative shock is supposed to increase the volatility, while the positive shock is supposed to reduce the volatility. The test involves the autoregressive specification estimation of the time series as

$$y_t = \omega + \phi(l)y_t + u_t. \tag{11}$$

The residuals from the estimation are used for the least square equation:

$$u_t{}^2 = c_0 + c_1 s^-{}_t + c_2 s^-{}_{t-1} u_{t-1} + c_3 s^+{}_{t-1} u_{t-1} + v_t. \tag{12}$$

$$\begin{cases} s^{-1}{}_t = 1 & if\ sign\ u_t = -1 \\ s^{-1}{}_t = 0 & if\ sign\ u_t = +1 \end{cases} \text{ and } s^+{}_t = 1 - s^-{}_t.$$

The coefficent c_i $(i = 1, 2, 3)$ is distributed asymmetrically with Student-t. If the coefficient c_i is significant, this means that the positive or negative error terms (external shocks), affect the variance differently to predict the response variable. On the other hand, if the coefficient turns out to be insignificant, it means no bias due to the sign of the disturbances.

The sign bias hypothesis is rejected at the 5 percent significance level, but it is accepted at the 10 percent significance level, because the p-value is 0.09. The test statistic fails to reject the nullity of the negative sign and the joint effect hypothesis. Likewise, positive sign bias is rejected at the 5 percent significance value, but it too is accepted at the 10 percent significance level. This shows that the sign bias test and the positive sign test yield similar results, providing some evidence of the bias. This may suggest that the models that do not take into account bias, may not be able to represent the FIB time series adequately. This could be deemed an important discovery of this analytical framework.

The no-sign bias test failed to reject the no-sign bias effect. The null is not rejected because p-value is 0.10. Similarly, the positive sign bias tests indicate that the coefficient is no different from zero. P-value fails to reject the null of no bias. No negative sign-bias null is rejected at the 10 percent, but supported at the 5 percent significance level. The critical value, 0.056, is almost at the border of the rejection limit. This may show that there is a negative sign bias. On the other hand, the test fails to reject the no-joint-effect null hypothesis, implying no combined effect of the sign and magnitude; p-value (0.17) supports nullity of the hypothesis.

4.3 Nyblom Test

To test the stability or constancy of the parameters we use the Nyblom test. This test assesses the variance of the errors in the parameter. If the parameter is a constant (i.e., no errors), then variance of the error term is zero. If it is not a constant (related to the past values of the parameter) then the error term has variance. The test statistic is based on this logic. In the testing of stability, the test statistic generally does not fall in the rejection region. In our hypothesis testing, the stability of the parameters is generally not rejected. So our estimated parameters are stable (do not shift over time).

4.4 TGARCH

To compare with the above model, an alternative specification is used for FIB data. This model consists of turbidity, rainfall, dew point, river flow, and cloud cover as the exogenous variables. The distribution is the same as above. The model output shows that the coefficients of all the variables are significant. P-values are very low, rejecting the null hypotheses. All the exogenous variables are positively related to the FIB density.

In the variance equation, not only alpha but also beta is significant. Similarly, parameter of shape is significant. The joint biased test indicates significance at the 8 percent, not at the 5 percent critical p-value. The test statistic for negative sign bias test fails to reject the no bias effect of the negative shocks because p-value is 0.35. Similarly, no joint effect hypothesis is not rejected in the joint testing (p-value (0.28)). Moreover, the test fails to reject the no positive sign hypothesis at 5 percent threshold level. However, at the 10 percent level the hypothesis is rejected. P-value is 0.099, which is insignificantly lower than the 10 percent limit. No bias is found at the 5 percent significance level. If the bound is extended to 10 percent, then the bias parameter reverses in significance.

Table 1: TGARCH with GED with Shape Parameter

Variables	Coefficient	Std. Error	t-ratio	p-value	Nyblom test statistics
mu	-5.035	0.081	-62.509	0.000	0.082
inmean	1.093	0.204	5.354	0.000	0.081
mxreg1	0.954	0.0569	16.768	0.000	0.085
mxreg2	0.065	0.011	5.874	0.000	0.075
mxreg3	2.779	0.046	60.445	0.000	0.079
mxreg4	0.179	0.044	4.056	0.000	0.094
mxreg5	0.202	0.053	3.847	0.000	0.186
omega	0.003	NA	NA	NA	NA
alpha	0.062	0.0256	2.421	0.015	0.244
beta	0.928	0.019	48.132	0.000	0.158
shape	1.608	0.259	6.213	0.000	0.088
Information Criteria					
Akaike	1.150	Bayes	1.317	Shibata	1.146
Hannan-Quinn	1.218		Log Likelihood		-118.388
Sign Bias Test		t-value		Prob sig	
Sign Bias		1.7062		0.089	
Negative Sign Bias		0.9443		0.346	
Positive Sign Bias		1.6581		0.099	
Joint Effect		3.8079		0.283	

Dependent variable is log10 *E. coli*; mxreg1... mxreg5 indicate exogenous variables, turbidity, 24- hour rainfall, dew point, Cuyahoga River flow and cloud cover, respectively, in the log10 form. Also note, Nyblom critical value is 2.96 at 5 percent.

4.5 GJR-GARCH

This GJR-GARCH differs from the above TGARCH version because it includes the skewed Generalized Error Distribution. TGARCH did not include the shape parameter. However, the results of the model are no different from the above TGARCH model in terms of number of significant exogenous variables. Moreover, not only alpha and beta, but also skew and shape parameters are significant at the 5 percent critical value.

Sign bias, negative sign bias, and joint effect of the sign and scale are no different from zero. Moreover, the test statistic fails to reject the positive sign bias null of no effect at the 5 percent significance level. If we relax the threshold bound from the 5 percent to 10 percent probability, then the parameter reflecting the positive sign bias is significant because the p-value is 0.065.

The skewed version of the model seems to fit the data very well. The model output is presented in the Table 2.

Table 2: GJR-GARCH with Skewed GED with Shape Parameter

Variables	Estimate	Std. Error	t value	Pr(>|t|)	Nyblom Statistic
mu	-5.211	0.062	-84.655	0.000	0.038
inmean	1.559	0.218	7.157	0.000	0.040
mxreg1	0.974	0.051	18.951	0.000	0.036
mxreg2	0.070	0.010	6.983	0.000	0.038
mxreg4	0.139	0.025	5.563	0.000	0.037
mxreg5	0.159	0.046	3.479	0.001	0.033
omega	0.0004	NA	NA	NA	NA
alpha	0.016	0.006	2.857	0.004	0.034
gamma11	1.000	0.354	2.827	0.000	0.034
beta	0.967	0.006	152.500	0.000	0.034
skew	1.042	0.114	9.110	0.000	0.147
shape	1.691	0.201	8.420	0.000	0.111
Information Criteria					
Akaike	1.118	Bayes	1.315	Shibata	1.112
Hannan-Quinn	1.198		Log Likelihood		-112.768
Sign Bias Test		t-value		Prob sig	
Sign Bias		1.571		0.118	
Negative Sign Bias		0.739		0.461	
Positive Sign Bias		1.854		0.065	
Joint Effect		4.0143		0.260	

Dependent variable is log10 *E. coli*; mxreg1... mxreg5 indicate exogenous variables, turbidity, 24- hour rainfall, dew point, Cuyahoga River flow and cloud cover, respectively, in the log10 form. Also note, Nyblom critical value is 2.96 at 5 percent.

4.6 EGARCH (Skewed normal distribution)

For comparison to the TGARCH and GJR-GARCH above models, a radically different specification is used for the FIB density data. Instead of the Generalized Error Distribution, asymmetric normal distribution is used. The same variables which are used in other models are used for this specification also. Had we used the different set then the comparison would have been difficult. Turbidity, rainfall, dew point, and cloud cover are significant variables. Surprisingly, however, river flow (mxreg4) is an insignificant variable. Similarly, in the variance equation alpha is no different from zero. The model output is presented in Table 3.

The test does not support the negative sign hypothesis at the 5 percent significant level. Similarly, joint effect is not supported at this level of significance. It is interesting that not only the sign-bias null of no effect but also the negative sign null of no negative sign effect are both rejected at 10 percent significant level. However the testing fails to reject the nullity at 5 percent significance level. P-value is 0.06 for sign bias and 0.088 for negative sign bias. This seems to underscore the bias effect.

Table 3: EGARCH with Skewed Normal Distribution

Variables	Coefficient	Std. Error	t-ratio	p-value	Nyblom Statistic
mu	-4.848	0.864	-5.608	0.000	0.054
inmean	1.593	0.356	4.479	0.000	0.059
mxreg1	0.914	0.073	12.538	0.000	0.058
mxreg2	0.069	0.012	5.708	0.000	0.051
mxreg3	2.639	0.471	5.601	0.000	0.052
mxreg4	0.151	0.086	1.766	0.077	0.052
mxreg5	0.193	0.057	3.352	0.001	0.096
omega	0.002	NA	NA	NA	NA
alpha	0.039	0.021	1.843	0.065	0.229
beta	0.955	0.017	55.821	0.000	0.152
skew	1.060	0.108	9.814	0.000	0.191
Information Criteria					
Akaike	1.149	Bayes	1.316	Shibata	1.145
Hannan-Quinn	1.217			Log Likelihood	-118.281
Sign Bias Test		t-value		Prob sig	
Sign Bias		1.889		0.060	
Negative Sign Bias		1.233		0.219	
Positive Sign Bias		1.720		0.087	
Joint Effect		4.625		0.201	

Dependent variable is log10 *E. coli*; mxreg1... mxreg5 indicate exogenous variables, turbidity, 24- hour rainfall, dew point, Cuyahoga River flow and cloud cover, respectively, in the log10 form. Also note, Nyblom critical value is 2.96 at 5 percent.

4.7 EGARCH (normal distribution)

In the model specification, the skewed normal is replaced with a normal distribution making it a regular symmetric GARCH model. The result of the model estimation differs from the EGARCH with the skewed normal distribution. In contrast, all the exogenous variables are significant.

The results of the variance equation are also different. Like the skewed EGARCH alpha, EGARCH alpha is also insignificant. Surprisingly, the sign bias test fails to reject the null hypotheses. So this model does not show any sign bias. The model output is presented in Table 4.

Table 4: EGARCH With Normal Distribution

Variables	Coefficient	Std. Error	t-ratio	p-value	Nyblom Statistic
mu	-4.749	0.863	-5.506	0.000	0.069
inmean	1.103	0.259	4.264	0.000	0.065
mxreg1	0.916	0.072	12.662	0.000	0.078
mxreg2	0.068	0.012	5.726	0.000	0.069
mxreg3	2.668	0.475	5.622	0.000	0.066
mxreg4	0.168	0.086	1.964	0.050	0.079
mxreg5	0.193	0.058	3.339	0.001	0.152
omega	0.002	NA	NA	NA	NA
alpha	0.054	0.033	1.648	0.099	0.286
Information Criteria					
Akaike	1.1495	Bayes	1.301	Shibata	1.146
Hannan-Quinn	1.211		Log Likelihood		-119.322
Sign Bias Test		t-value		Prob sig	
Sign Bias		1.529		0.128	
Negative Sign Bias		0.943		0.347	
Positive Sign Bias		1.555		0.121	
Joint Effect		3.351		0.341	

Dependent variable is log10 *E. coli*; mxreg1... mxreg5 indicate exogenous variables, turbidity, 24- hour rainfall, dew point, Cuyahoga River flow and cloud cover, respectively, in the log10 form. Also note, Nyblom critical value is 2.96 at 5 percent.

4.8 IGARCH (normal)

For comparison with other models, normal distribution, which is usually used in the literature, is used in the IGARCH model. The symmetry assumption is substituted for the asymmetry assumption for modeling the time series. The symmetric GARCH model consists of the same variables which are used in the above model. All the exogenous variables are found to be significant at the 5 percent significant level. P-values are very very small as was the case in other models. Beta of the variance specification is significant. However, alpha of variance turns out to be no different from zero at 5 percent significance.

Surprisingly, the test statistics for sign bias, negative sign, positive sign and joint effect fail to reject the null hypotheses. The p-value reaches 0.13 for the sign bias coefficient, 0.35 for negative sign, 0.12 for the positive sign bias and 0.34 for the joint effect coefficient. This is in contrast to some of the asymmetric models. We do not present this output to save space.

4.9 TGARCH (normal)

An alternative to the TGARCH with the Generalized Error Distribution, TGARCH with normal distribution is used. Shape and skew parameters are not estimated for this specification. The output of the model shows that the coefficients of the exogenous variables are significant. P- values are extremely small for the coefficients. All the variables are positively related to FIB densities.

However, the sign bias tests show rather interesting results. The test statistics for the sign bias shows that the coefficient is insignificant at the 5 percent significance level. However it is significant if we consider the 10 percent significance level because the p-value is 0.07. In contrast, negative sign p-value (0.48) supports the null of no negative sign bias. Similarly, the no joint effect parameter turns out to be no different from zero, the p-value (0.11) shows insignificance. However, in contrast, the test statistic confirms positive sign bias, since the p-value is just 0.02. The t-ratio appears suspect, therefore we do not report the results of this model.

4.10 NAGARCH (normal)

With the presence of the same exogenous variables, the nonlinear asymmetric autoregressive conditional heteroscedastic (NAGARCH) model is applied to the indicator bacterium density data. NAGARCH appears similar to the TGARCH because of the normal distribution. The result do not seems to be radically different from other models discussed above. The output of the NAGARCH is presented in Table 5. Turbidity, rainfall, dew point and cloud cover are significant variables. Nevertheless, river flow turns out to be an insignificant variable, although it was expected to be an important variable for the indicator bacteria.

Negative sign bias and the joint effect null hypotheses are not rejected by the bias tests. Sign bias and negative sign bias tests also fail to reject the null hypotheses, at a 5 percent significant level. Nevertheless, the coefficients are significant at the 10 percent significance level. Thus, there is some evidence that sign of the shocks play an important role in predicting the variable.

Table 5: NAGARCH With Normal Distribution

Variables	Estimate	Std. Error	t value	Pr(>|t|)	Nyblom Statistic
mu	-4.925	0.798	-6.172	0.000	0.048
inmean	1.376	0.204	6.742	0.000	0.050
mxreg1	0.934	0.071	13.161	0.000	0.048
mxreg2	0.070	0.012	5.938	0.000	0.047
mxreg3	2.876	0.444	6.479	0.000	0.048
mxreg4	0.072	0.069	1.049	0.294	0.048
mxreg5	0.151	0.052	2.910	0.004	0.044
omega	0.001	NA	NA	NA	NA
alpha	0.029	0.008	3.249	0.001	0.045
gamma	1.000	0.483	2.072	0.038	0.044
beta	0.937	0.027	34.741	0.000	0.045
Information Criteria					
Akaike	1.112	Bayes	1.279	Shibata	1.108
Hannan-Quinn	1.179		Log Likelihood		-114.097
Sign Bias Test		t-value		prob sig	
Sign Bias		1.715		0.088	
Negative Sign Bias		0.915		0.361	
Positive Sign Bias		1.984		0.048	
Joint Effect		4.800		0.187	

Dependent variable is log10 *E. coli*; mxreg1... mxreg5 indicate exogenous variables, turbidity, 24- hour rainfall, dew point, Cuyahoga River flow and cloud cover, respectively, in the log10 form. Also note, Nyblom critical value is 2.96 at 5%.
We used the skewed version of this distribution also. It was not found to be practical in parametric determination.

4.11 TGARCH (student's t)

The above mentioned forms of the model specifications are not based on the Student's t distribution. For comparison with other models, this distribution is included here in the GARCH specification. As in most of the models discussed above, the exogenous variables are significant under this specification. However, neither alpha nor the shape parameters are different from zero; the p-values do not fall below the threshold level.
No evidence was found for the sign bias using this model specification. As a result we do not find any comparative advantage of this model on other models. The model output is shown in Table 6.

Table 6: TGARCH With Student's t Distribution

Variables	Estimate	Std. Error	t value	Pr(>|t|)	Nyblom Statistic
mu	-4.949	0.870	-5.690	0.000	0.098
inmean	1.102	0.287	3.845	0.000	0.097
mxreg1	0.938	0.071	13.131	0.000	0.101
mxreg2	0.066	0.012	5.553	0.000	0.091
mxreg3	2.760	0.470	5.870	0.000	0.094
mxreg4	0.171	0.082	2.075	0.038	0.110
mxreg5	0.189	0.056	3.373	0.001	0.215
omega	0.004	NA	NA	NA	NA
alpha	0.063	0.040	1.601	0.109	0.141
beta	0.925	0.027	33.860	0.000	0.092
shape	8.470	7.595	1.115	0.265	0.051
		Information Criteria			
Akaike	1.146	Bayes	1.313	Shibata	1.142
Hannan-Quinn	1.214		Log Likelihood		-117.940
Sign Bias Test		t-value		Prob sig	
Sign Bias		1.3864		0.167	
Negative Sign Bias		0.796		0.427	
Positive Sign Bias		1.489		0.138	
Joint Effect		2.881		0.410	

Dependent variable is log10 *E. coli*; mxreg1... mxreg5 indicate exogenous variables, turbidity, 24- hour rainfall, dew point, Cuyahoga River flow and cloud cover, respectively, in the log10 form. Also note, Nyblom critical value is 2.96 at 5 percent.

4.12 AVGARCH

AVGARCH is not different from TGARCH, GJR-GARCH and EGARCH in its output (Table 7) of the significant of the exogenous variables. However, the coefficient gamma11 of the variance equations is not significant. The sign bias test shows insignificance of the bias at 5 percent level. However, the sign bias is almost significant because the p-value is just 0.06. We do not present the result of the use of other distributions because they will crowd the paper causing confusion.

Table 7: AVGARCH With Normal Distribution

Variables	Estimate	Std. Error	t value	Pr(>|t|)	Nyblom Statistic
mu	-3.280	1.005	-3.264	0.001	0.141
inmean	-1.505	0.663	-2.269	0.023	0.124
mxreg1	0.902	0.080	11.323	0.000	0.138
mxreg2	0.066	0.012	5.438	0.000	0.098
mxreg3	2.355	0.537	4.386	0.000	0.131
mxreg4	0.236	0.094	2.505	0.012	0.143
mxreg5	0.207	0.060	3.442	0.001	0.217
omega	0.252	NA	NA	NA	NA
alpha	0.060	0.024	2.537	0.011	0.054
gamma11	1.000	1.249	0.801	0.423	0.058
gamma21	1.000	0.260	3.848	0.000	0.024
beta	0.262	0.035	7.445	0.000	0.100
Information Criteria					
Akaike	1.142	Bayes	1.324	Shibata	1.136
Hannan-Quinn	1.215		Log Likelihood		-116.432
Sign Bias Test		t-value		Prob sig	
Sign Bias		1.911		0.057	
Negative Sign Bias		0.414		0.679	
Positive Sign Bias		1.302		0.194	
Joint Effect		4.398		0.222	

Dependent variable is log10 *E. coli*; mxreg1... mxreg5 indicate exogenous variables, turbidity, 24- hour rainfall, dew point, Cuyahoga River flow and cloud cover, respectively, in the log10 form. Also note, Nyblom critical value is 2.96 at 5 percent.

We also experimented with normal inverse Gaussian distribution version of the GARCH modeling. It was not found to be practical for the estimation of the parameters.

5 Conclusions

Box and Jenkins' framework is infrequently used in the environmental sciences. It is almost essential for stochastic processes. It is used for time series analysis because it incorporates the past behavior of the series to explain the generation of the time series given the exogenous or explanatory variables. It is equally useful in modeling without the explanators. Usually, symmetric distribution is used in the literature on environmental modeling. What is not used is the asymmetric distributions in modeling. It is extensively used in macroeconomics for money, interest rate, inflation, gross domestic product and foreign exchange rate. It is frequently used in financial economics for explaining the behavior of the stock market. However, it is rarely used in environmental economics. What is even rarer is the use of variance and mean structure in the explanation of the dependent variable. This study fills this gap in the literature. Ali [20] introduced the use of ARCH and GARCH models in the environmental literature. However, he used the symmetric assumption in his approach. This study relaxes this assumption. Negative and positive shocks of the same magnitude may have disproportionate or unequal effect on FIB densities. Dry condition shock and wet condition shock of the same magnitude may have differing magnitude of the impact on pathogen indicators. Similarly the negative shocks could have longer-lasting effect than the positive shock. Small fluctuations may follow the smaller fluctuations and larger fluctuation may follow the larger fluctuations. In other words, symmetric assumption may not be sensible. In this paper, we relax the symmetric assumption. In the literature on environmental economics, we did not find any study that had had the Generalized Error Distribution. We bridge this gap in the literature. Since Generalized Error Distribution has wider tails, it could very well be useful to modeling the behavior of the pathogen indicator series. In addition, we use the Student's t distribution. We also investigated the use of other distributions. In this article, we use the Generalized Error, normal and Student's t and normal inverse Gaussian distributions. We also use the skewed version of these distributions. We used the TGARCH, GJR-GARCH, IGARCH EGARCH, APGARCH and AVGARCH models.
We found in general that turbidity, rainfall, dew point, river flow and cloud cover are the significant variables. The variance parameters are found to be significant in most of the model specifications. Similarly, the mean structural parameter is significant. We conclude that there is some evidence of sign bias, even though the model output differed in sign bias test results. We applied Nyblom test for testing the stability of the estimated parameters. We conclude that the value of the parameters did not shift during the sampling period, supporting the invariance assumption of the parameters.

With wider tail distribution, the TGARCH model is reasonable for explaining the data. We did not find TGARCH with normal distribution to have comparative advantage over the TGARCH with Generalized Error Distribution in its explanatory power. Similar to TGARCH, GJR-GARCH is found to provide a reasonable explanation of the pathogen indicator time series. This is so because TGARCH and GJR-GARCH differ very little; they are competitive with each other. We do not find comparative advantage of the AVGARCH over the TGARCH in establishing relationships, although it yielded nearly similar results. EGARCH is not found to be better than the TGARCH in explaining the variations. We did not find relevance of the skewed normal inverse Gaussian distribution in GARCH modeling of the FIB density series. Nor do we find the relevance of the skewed normal distribution in GARCH modeling. Nevertheless, it is instructive to apply not only various distributions and but also various GARCH models for pathogen time series to establish a relationship. The variance and mean structures provide additional explanatory power to the volatility models in establishing functional relationships that would be useful for issuing beach advisories for recreational activities.

ACKNOWLEDGEMENTS: The author is thankful to Elizabeth Southerland, James Pendergast, Thomas Wall and Deborah Nagle for encouragement. In addition, he wishes to thank Lisa E. Faeth for inspiration.

References

[1] R. Engle, Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation, *Econometrica*, **50**(4), (1982), 987-1007.

[2] T. Bollerslev, Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics*, **9**(2), (1986), 121-131.

[3] D.B. Nelson, Conditional heteroscedasticity in asset returns: a new approach, *Econometrica*, **59**(2), (1991), 347-370.

[4] D. Hsieh, Modeling heteroscedasticity in daily exchange rates, *Journal of Business and Economic Statistics*, 7(3), (1989), 307-317.

[5] P. Theodossiou, The Stochastic properties of major Canadian exchange rates, The Financial Review, **29**(2), (1994), 193-221.

[6] G. Kkoutmos and P. Theodossiou, Time-series properties and predictability of Greek exchange rates, *Managerial and Decision Economics*, **15**(2), (1994)159-177.

[7] V. Akgirary, G.C. Booth, J.C. Hatem and C. Mustafa, Conditional dependence in precious metal prices, *The Financial Review*, **26**(3), (1991), 367-86.

[8] R. Gallant, D. Hsieh and G. Tauchen, Estimation of stochastic volatility models with diagnostics, *Journal of Econometrics*, 81(1), (1997), 159-192.

[9] J. Yu, On leverage in a stochastic volatility model, *Journal of Econometrics,* **127**(2), (2005), 165-178.

[10] D. Siourounis, Modeling volatility and testing for efficiency in emerging capital markets: the case of the Athens stock exchange, *Applied Financial Economics*,**12**(1), (2002), 47–55.

[11] D. McMillan, A. Speight and O. Ap Gwilym, Forecasting UK stock market volatility, *Applied Financial Economics*, **10**(4), (2000), 435–48.

[12] C. Fernandez and M. Steel, On Bayesian modeling of fat tails and skewness, *Journal of the American Statistical Association*, **93**(441), (1998), 359–71.

[13] P. Lambert and S. Laurent, Modelling financial time series using GARCH-type models and a skewed student density, (2001), Mimeo, Universite de Liege.

[14] T. Bollerslev, On the correlation structure for the generalized autoregressive heteroscedastic process, *Journal of Time Series Analysis*, **9**(2), (1988), 121-131.

[15] R. Baillie and T. Bollerslev, Common stochastic trends in a system of exchange rates. *Journal of Monetary Economics*, **44** (1), (1989), 167–81.

[16] R. Harris, C. Kucukozmen and F. Yilmaz, Skewness in the conditional distribution of daily equity returns, *Applied Financial Economics*, **14**(3), (2004), 195–202.

[17] Z. Ding, R.F. Engle, and C.W.J. Granger, A long memory property of stock market return and a new model, *Journal of Empirical Finance* **1**(1), (1993), 83-106.

[18] L.R Glosten, R. Jagannathan and D.E. Runkle, Relationship between the expected value and the volatility of the nominal excess return on stocks, *The Journal of Finance,* **48**(5), (1993), 1779-1801.

[19] D. Alberg, H. Shalit and R. Yosef, Estimating stock market volatility using asymmetric GARCH models, *Applied Financial Economics*, **18**(15), (2011), 1201-1208.

[20] G. Ali, ARCH, GARCH and ARMAX models for forecasting pathogen indicators and advisories at marine recreational sites, *Marine Resource Economics*, **26**(3), (2011), 211-224.

[21] J.M. Zakoian, Threshold heteroscedasticity models. *Journal of Economic Dynamics and Control,* **18**(5), (1994), 931–955.

[22] S. Taylor, *Modelling Financial Time Series*, (1986), Wiley, New York.

[23] M.L. Higgins, and A.K. Bera, A class of nonlinear ARCH models, *International Economic Review*, 33 (1), 1992, 137-158.

[24] AK. Dhamija, and VK. Bhalla, Financial time series forecasting: comparison of neutral networks and ARCH models, *International Research Journal of Finance and Management*, **49**(1), (2010), 159-172.

[25] A.B. Boehm, R.L. Whitman, M.B. Nevers, D. Hou, and S.B. Weisberg, Nowcasting recreational water quality, in *Statistical Framework for Water Quality Criteria and Monitoring*, L.J. Wymer, ed., (2007), 179-210, England, John Wiley and Sons, Ltd.

[26] T. Klein, Econometrics II, course material, University of Cambridge, (2011), http://thiloklein.de/Econometrics2.html.

[27] A. Ghalanos, Introduction to the rugarch package, (2012), http://cran.r-project.org/web/packages/rugarch/vignettes/Introduction_to_the_rugarch_package.pdf.

[28] D. Wurtz, Y. Chalabi, and L. Luksan, Parameter estimation of ARMA models with GARCH?APARCH errors; an R and SPlus software implementation, *Journal of Statistical Software* (2008),1-41, http://www.jstatsoft.org.

[29] R.F. Engle, and V.K. Ng, Measuring and testing the impact of news on volatility, *Journal of Finance*, **48**(5), (1993), 1749-1778.