10	N	Nitrate concentration (mg/L N)	0.08	0.1485	0.17	0.0593	74.13
11	P	Phosphorus concentration in lyotic environments (mg/L P)	0.51	0.9648	1.213	0.3706	72.68
12	ST	Total solids(mg/L)	190.9	253.6	218.5	114	59.73
13	Coliforms	Number of coliforms / 100 mL	16,200	31,880	34,140	9,398	58.02
14	DBO	Biochemical oxygen demand (BOD 5 days at 20°C)	2.2	5.835	17.29	0.2965	13.48
15	pH	water pH	7.1	6.879	1.293	0.3706	5.22

Figure 3 shows a cluster analysis of the economic variables. To increase the number of degrees of freedom, the variables that presented the highest CVR% values were selected in each cluster:

Figure 3: Result of cluster analysis of economic variables

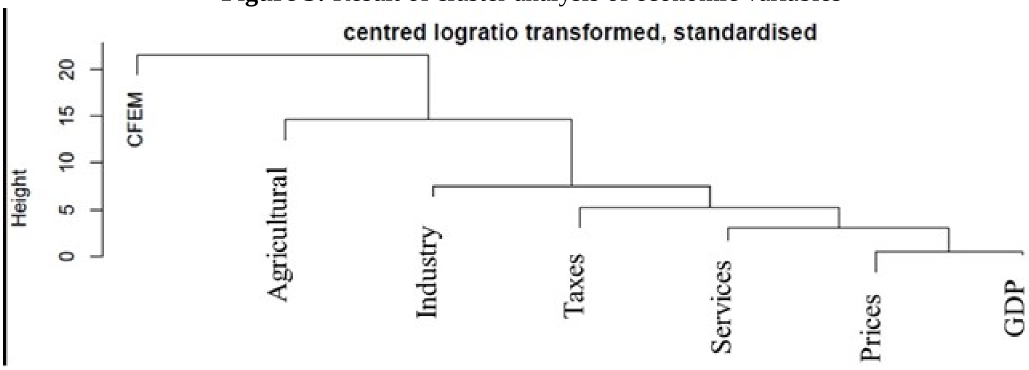


Figure 3: Kuznets Environmental Curve (Source: Based in Ávila and Diniz, 2015.)

The outlier analysis identified the following outliers Juramento (small town), Nova Serrana (intense industrial activity), Pirapora, and Três Marias (intense farming and industrial activities). The clr transformation was then performed, and the multiple regression analyses between turbidity and the selected economic variables were conducted. Turbidity was selected because it was the only environmental variable that presented CVR% greater than 100%. Table 2 presents the regression residues. Symmetry can be observed in relation to the mean zero value, which justifies the use of this model [28].

Table 2 -Multiple regression residues

Waste				
Min	1Q	Median	3Q	Max
-0.00702	-0.00186	-0.00047	0.002186	0.009146

The coefficients of the model are listed in Table 3. The standard error measures the average amount that the coefficient estimates vary from the actual average value of our response variable. The low values observed in Table 3 (<0,5 %) suggest the accuracy of the model. The Pr(>t) acronym found in the model output relates to the probability of observing any value equal or larger than t. A small p-value (<5%) indicates that a relationship between predictor variables and change-related responses is unlikely to exist. Typically, a p-value ≤ 0.05 is a good cut-off point. In the present study, Pr values (> t) were acceptable [28].

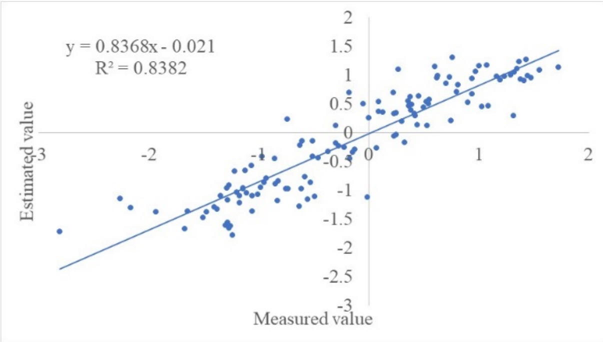
Table 3 - Parameters related to multiple regression model.

	Estimate	Standard error	Pr(>|t|)
Intersession	-0.0028	0.00037	7.10E-12
Agricultural	-1.0028	0.00086	2.00E-16
Industry	-1.0063	0.00157	2.00E-16
Services	-1.0276	0.00379	2.00E-16
Taxes	-0.9094	0.0027	2.00E-16
GDP	-2.0546	0.00463	2.00E-16
CFEM	-1.0002	0.00047	2.00E-16

The residual standard error is a measure of the quality of a linear regression adjustment. Theoretically, every linear model is assumed to contain an error term, €. Due to the presence of this error term, it is not possible to accurately predict the response variable from the predictor. The residual default error is the average value by which the response will deviate from the true regression line. After taking these parameters into account, the obtained value of 0.003652 with 118 degrees of freedom was acceptable [28].

The R^2 method provides a measure of how well the model adjusts to the actual data. It takes the form of a proportion of the variation. R^2 is a measure of the linear relationship between the forecasted variable and the response or target variable. The value of R^2 is always between 0 and 1 (a number close to 0 represents a regression that does not adequately explain the variance in the response variable, and a number close to 1 explains the variance observed in the response variable). In our study, the R^2 value obtained was 0.84. Thus, approximately 84% of the variance found in the response variable can be explained by the predictor variable [28]. The regression result is presented in Fig 4.

Figure 4: Multiple regression result.



The F method coefficient was 1,424 × 10^6 for 6 variables and 118 degrees of freedom. The F value is a good indicator of the existence of the relationship between the predicted and response variables. The further the F value is away from 1, the better the results. However, the F method depends on the number of data points and predicted variables. The p-value was $< 2.2 \times 10^{-16}$. This suggests that the adjustment of the model was good/bad [28].

4. Discussions

Turbidity represents the degree of interference during the passage of light through water, which leads to a blurred appearance of the light [29]. Turbidity is facilitated by the presence of suspended particles such as rock particles, clay, silt, algae, microorganisms, and domestic and industrial sewage, among others. High turbidity reduces the photosynthetic capability of submerged rooted vegetation and algae [15], causing the extinction of biological communities. The municipalities detected as outliers were Juramento, Nova Serrana, Pirapora, and Três Marias (Table 4).

Table 4 - Outliers detected for the variables studied.

City	CFEM	Agricultural	Industry	Services	Taxes	GDP
Juramento	0.01	6,552	1,525	8,7194	1,025	34,481
NovaSerrana	96,388	10,512	738,710	607,431	198,864	1,838,089
Pirapora	6,187	38,621	603,969	537,633	226,285	1,605,428
Três Maria	3,172	68,584	552,606	264,505	201,578	120,7171

In Três Marias and Pirapora, monoculture helped to characterize a scenario of large-scale agricultural activities, with many farms (Ribeiro et al., 2012). The primary activities in these municipalities are agriculture, silting, mining, livestock, diffuse load, urbanization, industrial activities, and forestry [15].

Nova Serrana is known to produce sports shoes; in 2019, the local production was 105 million pairs. The main factors associated with the degradation of water quality in this municipality are the absence of adequate treatment facilities for sewage and industrial effluents, animal slaughter activities, and inadequate soil management practices [31]. Juramento presents one of the lowest GDP in the basin, owing to the absence of significant economic activities.

The variables related to economic activities contributing to the degradation of water quality in the São Francisco River Basin in Minas Gerais were agriculture, industry, services, taxes, GDP, and CFEM. This degradation in water quality is a result of non-compliance with environmental laws, human settlements, real estate growth in marginal areas of water bodies, destruction of forests along the river, and lack of awareness regarding the need for sustainability of natural resources [32].

The presence of an agro-industrial production model in the basin has been causing a chain of environmental problems. Since the 1970s, the basin has seen an accelerated and unbridled expansion of intensive agricultural practices [33]. Agricultural activities have the potential to contaminate water resources because the water supplied to an irrigated area that is not absorbed by the crops returns to rivers by surface runoff, carrying soluble salts, fertilizers, and toxic elements [34] .

As for mining, the São Francisco River Basin accounts for about 20% of the country's official mineral activity, which contrasts with the finding that the mineral sector is one of the major water polluters in the basin [34] ..

The municipalities in the southeast have most of their waterways filled with tailings from mining, industrial processes, and agriculture [35] .

The results report that, in the studied region, the hypothesis of the Kuznets environmental curve is valid and the increase in economic development reflects environmental degradation. Alam et al. (2016) [36], who applied the Kuznets curve hypothesis in Brazil, also validated the Kuznets environmental curve considering the CO_2 emission.

From the economic and environmental perspectives, the development process in the region has not yet improved the living conditions of the residents or environmental quality. Amidst development and the lack of sustainable policies, the regions experience degradation of soil and waterways because of various production processes and general human activities. In this sense, environmental, social, political, and economic conditions are closely associated and may both improve and deteriorate the lives of the population [35] .

5. Conclusions

The aim of this study was to examine the validity of the Kuznets environmental curve hypothesis in the São Francisco River Basin in Minas Gerais. Strong evidence of a correlation between economic variables and those related to water quality was obtained in this study.

As expected, the release of domestic sewage throughout the region, agricultural, mining, and urbanization activities contribute to the bulk of the water pollution during the studied period.

Due to the pressure of development and lack of sustainable policies, the natural resources in the São Francisco River Basin in Minas Gerais face a critical threat of degradation.

In conclusion, we have found the occurrence of the often-hypothesized inverted-U shaped relationship between income and environmental degradation, known as the Environmental Kuznets Curve, in the studied region. Throughout the region, water is contaminated, especially in places with higher income. The data indicate that the municipalities are at lower stages of development. In these stages, the amount and intensity of environmental degradation are related to the impacts of subsistence economic activity and a large amount of biodegradable waste.

Acknowledgements

This work was supported by FAPEMIG (Minas Gerais Research Support Foundation) and PUC Minas (Pontifical Catholic University of Minas Gerais).

References

[1] Brazil, V.M., 2008. The São Francisco River: The physical basis of the national unity of the empire. Revista Mosaico 1, 133–142.

[2] Nasir, M., Ur Rehman, F., 2011. Environmental Kuznets curve for carbon emissions in Pakistan: An empirical investigation. Energy Policy 39, 1857–1864.

[3] Saboori, B., Sulaiman, J., 2013. Environmental degradation, economic growth and energy consumption: Evidence of the environmental Kuznets curve in Malaysia. Energy Policy 60, 892–905.

[4] Shahbaz, M., Lean, H.H., Shabbir, M.S., 2012. Environmental Kuznets Curve hypothesis in Pakistan: Cointegration and Granger causality. J. Renew. Sustain. Energy View 16, 2947–2953. https://doi.org/10.1016/j.energy.2016.12.106

[5] Tutulmaz, O., 2015. Environmental Kuznets curve time series application for Turkey: Why controversial results exist for similar modes? Renew. Sustain. Energy Rev. 50, 73–81.

[6] Cheng, Z., 2016. The spatial correlation and interaction between manufacturing agglomeration and environmental pollution. Ecol. Indic. 61, 1024–1032. https://doi.org/10.1016/j.ecolind.2015.10.060

[7] Dong, K., Sun, R., Jiang, H., Zeng, X., 2018. CO_2 emissions, economic growth, and the environmental Kuznets curve in China: What roles can nuclear energy and renewable energy play? J. Cleaner Prod. 196, 51–63. https://doi.org/10.1016/j.jclepro.2018.05.271

[8] Javid, M., Sharif, F., 2016. Environmental Kuznets curve and financial development in Pakistan. Renew. Sustain. Energy Rev. 54, 406–414. http://doi.org/10.1016/j.rser.2015.10.019 (accessed October 2020).

[9] Menezes, João Paulo Cunha et al. Relação entre padrões de uso e ocupação do solo e qualidade da água em uma bacia hidrográfica urbana. Eng. Sanit. Ambient., Rio de Janeiro, v. 21, n. 3, p. 519-534, set. 2016. http://dx.doi.org/10.1590/S1413-41522016145405 .

[10] Ozturk, I., Mulal, U., 2015. Investigating the Validity of the Environmental Kuznets Curve Hypothesis in Cambodia. Ecol. Indic. 57, 324–330. https://doi.org/10.1016/j.ecolind.2015.05.018

[11] Avila, E.S., Diniz, E.M.S., 2015. Evidence on Kuznets environmental curve and convergence of emissions. Estud. Econ. 45, 97–126. http://dx.doi.org/10.1590/0101-4161201545197ese

[12] BBC. Brumadinho dam collapse in Brazil: Vale mine chief resigns. 2019. https://www.bbc.com/news/business-47432134

[13] Maneta,M.P., Torres, M., Wallender, W.W., Vosti, S., Kirby, M., Bassoi, L.H., Rodrigues, L. N., 2009. Water demand and flows in the São Francisco River Basin (Brazil)

[14] CONAMA. National Council for the Environment, 2008. Resolution. CONAMA. 357/2005. Available online: https://www.gov.br/mma/pt-br (Accessed 10 Aug 2020).

[15] IGAM, Instituto Mineiro de GESTÃO DAS ÁGUAS, 2019. http://www.igam.mg.gov.br/monitoramento-da-qualidade-das-aguas2 (accessed October 2020).

[16] IBGE, 2019. Brazilian Institute of Geography and Statistics. https://www.ibge.gov.br/ (accessed October 2020).

[17] DNPM/ANM. National Department of Mineral Research. https://sistemas.anm.gov.br/scm/site/admin/default.aspx (accessed October 2020).

[18] Sartori, S.D., 2008. *Applications of multivariate analysis techniques in agricultural experiments using R software.* Dissertation (master's degree in Agronomy). Luiz de Queiroz School of Agriculture, Piracicaba, SP.

[19] R Foundation for Statistical Computing, 2019. https://www.r-project.org/foundation/Rfoundation-statutes.pdf (accessed October 2020).

[20] Filzmoser, P., 2013. Statistical analysis for environmental data: Package '**StatDA**'. R Package Version Package version 1.6.5, http://cran.r-project.org/web/packages/StatDA/StatDA.pdf (accessed October 2020).

[21] Todorov, V., 2013. Scalable Robust Estimators with High Breakdown Point. R package version package version 1.3-2. http://cran.r-project.org/web/packages/rrcov/rrcov.pdf (accessed October 2020).

[22] Reimann, C., Filzmoser, P., Garrett, R., Dutter, R., 2008. Statistical Data Analysis Explained: Applied Environmental Statistics with *R*. John Wiley & Sons: Kindle Ed. 335 p.

[23] Katircioğlu, S.T., Taşpinar, N., 2017. Testing the moderating role of financial development in an environmental Kuznets curve: Empirical evidence from Turkey. Renewable and Sustainable Energy Reviews 68, 572–586. http://dx.doi.org/10.1016/j.rser.2016.09.127
[24] Silva, G., Perobelli, F. S., 2018. Sectoral Interconnections and GDP per capita: Is there a direct relationship between both variables? Estud. Econ. 48, 251–282. http://dx.doi.org/10.1590/0101-41614823gdfp
[25] Alam, M.S., Paramati, S.R., 2016. The impact of tourism on income inequality in developing economies: Does Kuznets curve hypothesis exist? Ann. Tourism Res. 61, 111–126. https://doi.org/10.1016/j.annals.2016.09.008
[26] Taghvaee, V.M., Shirazi, J.K., 2014. Analysis of the relationship between economic growth and environmental pollution in Iran (evidence from three sections of land, water and atmosphere) Indian J.Sci.Res. 7, 31-42.
[27] Solarin, S.A., Al-Mulalia, U., Ozturk, I., 2017. Validating the environmental Kuznets curve hypothesis in India and China: The role of hydroelectricity consumption Renewable and Sustainable Energy Reviews 80, 1578–1587. http://dx.doi.org/10.1016/j.rser.2017.07.028
[28] Filzmoser, P., Hron, K., Templ, M. Applied Compositional Data Analysis. 2010. (Springer Series in Statistics) (p. iv). Springer International Publishing. Edição do Kindle. https://doi.org/10.1007/978-3-319-96422-5
[28] R pubs: Multiple Linear Regression R Guide, 2019. https://rpubs.com/bensonsyd/385183 (accessed October 2020).
[28] Rong, Y. (Ed.), 2011. Practical Environmental Statistics and Data Analysis. ILM Publications.
[30] Todorov, V., Filzmoser, P., 2009. An object oriented framework for robust multivariate analysis. J. Stat. Soft. 32, 1–47.
[31] Dantas, M.S., Freitas, L.N., Vani, J.C., Oliveira, J.C., Oliveira, S., M.A. C., 2019. The integrated analysis of the influence of changes in land use and occupation for the identification of water bodies and the most impacted municipalities of the Rio Pará basin -MG. http://abes.locaweb.com.br/XP/XP-EasyArtigos/Site/Uploads/Evento45/TrabalhosCompletosPDF/VI-124.pdf, 30 ABES Congress (accessed October 2020).
[32] Silva, A, P, de S,, Dias, H. C. T., Bastos, R. K. X., Silva, E. Qualidade da água do Reservatório da Usina Hidrelétrica (UHE) de Peti, Minas Gerais. Rev. Árvore [online]. 2009, vol.33, n.6, 1063-1069. http://dx.doi.org/10.1590/S0100-67622009000600009.
[33] Zellhuber, A., Siqueira, R., 2007. RIO São Francisco in the way: Degradation and revitalization. CEAS Notebooks: Critical Journal of Humanities. 227 https://cadernosdoceas.ucsal.br/index.php/cadernosdoceas/article/view/124/104 (accessed October 2020).
[34] Simon, F.R., Faria, M.A., Costa, E.L., Oliveira, P.M., 2009. Collection and Analysis of Environmental Impacts Generated by the Irrigated Perimeter of Jaíba. EPAMIG, Belo Horizonte. 68. http://www.epamig.br/download/tabela-climatica-estacao-jaiba-mg/
[35] Barbosa, M.S.M., Lima, K.S.C., Friede, R.R., Miranda, M.G., 2015. The relationship between poverty and environmental degradation in Brazil from the perspective of the GDP/HDI indicators. Semioses 9, 17–35.
[36] Alam, M.M., Murad, M.W., Noman, A.H.M., Ozturk, I., 2016. Relationships among carbon emissions, economic growth, energy consumption and population growth: Testing Environmental Kuznets Curve hypothesis for Brazil, China, India and Indonesia. Ecol. Indic. 70, 466–479. https://doi.org/10.1016/j.ecolind.2016.06.043
[37] Apergis, N., Ozturk, I., 2015. Testing Environmental Kuznets Curve hypothesis in Asian countries. Ecol. Indic. 52, 16–22. https://doi.org/10.1016/j.ecolind.2014.11.026
[38] Usama, A., Chor, F.T., Ilhan, O., 2015. Estimating the environment Kuznets curve hypothesis: evidence from Latin America and the Caribbean countries. Renewable and sustainable Energy reviews v. 55, 918–924.
[39] Sugiawan, Y., Managi, S., 2016. The environmental Kuznets curve in Indonesia: Exploring the potential of renewable energy. Energy Policy 98, 187–198.
[40] Mehdi, B.; Slim, B.; Ozturkc, I. Testing environmental Kuznets curve hypothesis: The role of renewable and non-renewable energy consumption and trade in OECD countries. Ecological Indicators60(2016)824–831. http://dx.doi.org/10.1016/j.ecolind.2015.08.0311470-160X/
[41] Tingting Li. 2016. Environmental Kuznets Curve in China: New evidence from dynamic panel analysis. Energy Policy 91, 138–147. https://doi.org/10.1016/j.enpol.2016.01.002
[42] Ahmad, N., Du, L., Lu, J., Wang, J., Li, H., Hashmi. M.Z., 2017. Modelling the CO2 emissions and economic growth in Croatia: Is there any environmental Kuznets curve? Energy 123, 164–172.
[43] Wang, S., 2003. Forest cover, environmental quality, and economic well-being are they related. XII Word Foresty Congress, Québec, Canada. http://www.fao.org/3/XII/0597-C5.htm (accessed October 2020).
[44] Taghvaee, V. M.; Shirazi, J. K. (2016) Analysis of the relationship between economic growth and environmental pollution in Iran (evidence from three sections of land, water and atmosphere) MPRA Paper No. 70056. https://mpra.ub.uni-muenchen.de/70056/
[45] Omay, R.E. 2013. The relationship between environment and income: Regression spline approach. Int. J. Energy Econ. Policy 3, 52–61.
[46] Özokcua, S. Özdemir, Ö. Economic growth, energy, and environmental Kuznets curve, Renewable and Sustainable Energy Reviews 72 (2017) 639–647
[47] Tingting Li., Wu, J., Dang, A., Lia, L., Xu, M., 2019. Emission pattern mining based on taxi trajectory data in Beijing. Journal of Cleaner Production 206, 688–700. https://doi.org/10.1016/j.jclepro.2018.09.051

APPENDIX A - SELECTED ARTICLES

No.	Author	Independent variables	Dependent variables	Local	Model*
1	Apergis and ozturk (2015)[37]	CO_2 emission	GDP per capita	Asian countries	II
			Population density		
			Land use		
			Industry		
2	Usama et al. (2015)[38]	CO_2 emission	GDP per capita	Latin America and the Caribbean	I
			Energy consumption per capita		
			Economic development		
3	Dong et al. (2018)[7]	CO_2 emission per capita	GDP per capita	China	I
			Non-renewable energy consumption per capita		
			Renewable energy consumption per capita		
4	Solarin et al. (2017)[27]	CO_2 emission	Consumption of hydroelectric power	China	I
			Urbanisation rate		
			GDP per capita		
5	Alam and Paramati (2016)[25]	Gini	Foreign direct investment	49 developed countries	I
			Economic inequality		
			Inflation		
			Commercial opening		
			Renewable tourism		
			GDP per capita		
6	Katircioğlu and Taşpinar (2017)[23]	CO_2 emission	GDP, Energy consumption	Turkey	Ii
7	Sugiawan and Managi (2016)[39]	CO_2 emission	GDP	Indonesia	Ii
			Renewable energy consumption per capita		
8	Alam et al. (2016)[36]	CO_2 emission	GDP per capita	India, Indonesia, China and Brazil	Ii

			Kg consumed of oil per capita		
			population growth		
9	Mehdi and Slim (2015)[40]	CO_2 emission	GDP	Tunisia	I
			Renewable energy consumption per capita		
			International trade		
10	Javid and Sharif (2016)[8]	CO_2 emission	GDP per capita	Pakistan	I
			Oil consumption		
			Commercial growth		
			International trade		
			Foreign direct investment		
11	Tingting et al. (2016)[41]	Emission liquid effluents	Economic development	28 provinces in China	II
		Solid effluent emission	Energy consumption		
		CO_2 emission	Commercial opening		
			Urbanisation		
	Ahmad et al. (2017)[42]	CO_2 emission	GDP	Croatia	I
13		CO_2 emission	GDP	Cambodia	I
			Total electricity consumption		
			Corruption index		
			Government reliability		
			Government effectiveness		
14	Wang (2013)[43]	Forest cover	Population density	Human Development Report	I
			% of rural population		
			Tropical model		
			GDP per capita		
15	Taghvaee and Shirazi (2014) [44]	CO_2 emission per capita	GDP per capita	IRAN	I
		BOD per capita	Pi		

			Deforestation		
16	Omay (2013)[45]	CO_2 emission	Economic growth	Turkey	Ii
			GDP		
17	Özokcua and Özdemir[46](2017)	CO_2 emission	GDP per capita	26 OECD countries	Ii
				52 emerging countries	
18	Tingting [47](2016)	CO_2 emission	Emission of industrial effluents per capita	China	Ii
			Solid tailings emission per capita		
	*I = Multiple regression, II = Multiple log-linear regression simple and quadratic terms				