MANAGING WATER COMMONS USING MEDIATOR VARIABLES TO BRIDGE THE GAP BETWEEN ENVIRONMENTAL FACTORS AND ANTHROPOGENIC POLLUTION INDICATORS

Davor Mance[1], Diana Mance[2], Darija Vukić-Lušić[3,4]
[1] University of Rijeka, Faculty of Economics, Rijeka (Croatia), e-mail: davor.mance@efri.hr
[2] University of Rijeka, Faculty of Physics, Rijeka (Croatia)
[3] Teaching Institute of Public Health of Primorje Gorski Kotar County, Rijeka (Croatia)
[4] University of Rijeka, Faculty of Medicine, Department of Environmental Health, Rijeka (Croatia)

Abstract – Water is the lifeblood of all life, so population concentrates near potable sweet water sources. People also concentrate near the coastline for economic reasons. Much of the drinking groundwater resources in the Croatian coastal area is stored in karst aquifers. Because of its quick pass-through nature and nonpoint source pollution, its protection is challenging. There have been many monitoring and measurement challenges in the past. For example, there was no empirical confirmation of a conjectured link between precipitation and microbiologic pollution in the monitored coastal areas. Before the use of $\delta^{18}O$ as mediator for the analysis of groundwater dynamics in karst aquifer characterisation, the causal links between precipitation and aquifer- as well as marine pollution were elusive. Data analysis of groundwater dynamics required also some dynamic statistical modelling, as for example dynamic panel data modelling in form of a General Method of Moments with First Differences transformation to control for unobserved time-invariant individual effect heterogeneity in. Static statistical models that include $\delta^{18}O$ values successfully represent the microbial pollution variations within a closed system. We understand this to be a characteristic of a stock pollution. At an open sea location, the results of static microbial pollution modelling have not been as good. Dynamic modelling using first differences of $\delta^{18}O$ values indicate that in these circumstances we deal with flow pollution.
Whether a pollutant is a stock or a flow is not only dependent on the pollutant itself, but mostly on the medium, and its environment. Policies regarded as optimal for stock and flow pollutants are different. The question of stock or flow is of great importance to decide whether the regulatory body should use price or quantitative allocation mechanisms. There are circumstances where we would prefer the one to the other because of political and administrative reasons, but it is ultimately up to the marginal costs and benefits to use the one or the other. When supply and demand curves are flat, it is better to use quota-like quantitative methods for avoiding planning mistakes. For inelastic costs and benefits, a price system of regulation is optimal instead. Whether an ecological system produces elastic or inelastic costs and benefits depends on factors such as pass-through velocity and its sustainability. Stocks are common goods and prefer price mechanisms. Flows are common pool resources and prefer quantitative mechanisms.

Keywords – *economic institutions, water commons management, mediator variable, pollution measurement, stock pollutants, flow pollutants.*

Referee List (DOI 10.36253/fup_referee_list)
FUP Best Practice in Scholarly Publishing (DOI 10.36253/fup_best_practice)

Introduction

Water is the lifeblood of all life, so people congregate near fresh water sources. People congregate near the coast for economic reasons as well. Dinaric karst covers nearly half of Croatia's land area (Fig. 1).

Figure 1 – Boundary of the Dinaric Karst (red dashed line).

A significant portion of Croatia's drinking groundwater resources is stored in karst aquifers. Because of its rapid passage through nature and nonpoint source pollution, karst protection is difficult. For any reasonable design of an institutional mechanism to control for pollution, the rate at which water flows through the system and its accumulation and degradation in the environment through which it travels determines whether a pollutant is a stock or a flow. There are times when we would prefer one to the other for administrative reasons, but it is ultimately up to the marginal costs and benefits to use one or the other. Ecological problems, as addressed in environmental economics and in the design of economic institutional mechanisms, are referred to as stock and flow problems. The rates of pollutant emission and depletion determine the sustainability of the system. According to Elinor Ostrom's classification of common goods into stocks or flows [9] we also classify water media into stocks or flows regarding their capability to retain pollutants. Stocks are better modelled with a static approach in form of an Estimated Generalised Least Squares causal mechanical inferential model of explanation. To account for the unobserved time-invariant variability of individual effects, aquifers with a higher degree of karstification are best modelled using a dynamic inferential causal mechanical method in the form of a General Method of moments with transformation of first differences [6,7]. Panel data statistical methods may also be used for hypothesis testing and coefficient estimation as well as direction of causality testing [4,8].

In this study, we present the results of observations using the $\delta^{18}O$ as a mediator variable between precipitation locations and human pollution at beach sites. Where other measurement and statistical techniques, as well as mediators had failed, have been prohibitively expensive or outright impossible to use, the $\delta^{18}O$, as a mediator variable, has provided statistically significant results.

Materials and Methods

With the help of formal logic, we can neither draw a deductive conclusion going from a description of the cause to the kind of the effect, nor are we able to conclude from a description of the effect about the kind of the cause. It is a common methodological problem to all empirical sciences: natural and social. In empirical sciences, a single consequence may have many different sources of causation. In addition to the problem of causation in natural sciences, which answers the question of how come, social sciences have the problem of grasping the reasons for certain human behaviour (teleologically rational or not), which answers the question of what for. Problems of anthropogenic pollution, its measurement, and ultimately abatement cannot be adequately addressed without resorting to human incentives, and social institutions in the form of rules and norms, as well as other institutions that have evolved to incentivize human behaviour in a certain social context.

In this paper, we firstly try to answer the question of measurement where some other techniques have failed in the past. Secondly, we try to use mediator indicator variables to try to understand the causal mechanics hidden within the black box of karstic aquifers. Finally, we present some unorthodox methods of inferential statistics that enabled us to test causal conjectures regarding the anthropogenic sources of pollution. The concept of causality is essential for scientific explanation. According to the Causal-Mechanical (CM) model of explanation, the fundamental causal mechanisms are causal processes and causal interactions [10, 11]. The most fundamental causal concept, according to Wesley Salmon, is that of a causal process. A causal process is defined by its ability to transmit a mark or its own structure in a spatiotemporally continuous manner. So, for all our practical intents and purposes, a pollutant ending up in an estuary, a spring, or on the beach, may have different sources or causes that may never precisely be identified. According to the CM model of explanation, to be able to claim that something is a source of pollution at a certain beach, one should be able to systematically and spatially continuously track the transmission of the pollutant from the source to the target location, a task that is neither feasible nor economical. This task is usually performed by some form of marker or tracer, usually a liquid of persistent colour, transmitting the colour downhill. At our measurement locations, and by using standard tracers and standard statistical methods, any attempt to find a direct link between precipitation or any other anthropogenic indicator variable gave spurious results [5,6,7]. When a mediator variable was introduced instead in the form of a delta value of the stable isotope ^{18}O, the puzzle finally fell into place: The tracer could serve as a mediator between rainfall in the hinterland and anthropogenic pollution on the beaches. In statistical terms, the

$\delta^{18}O$ is the mediator variable. The $\delta^{18}O$ is not perfect, because there is also the possibility of false correlation due to random mixing of precipitation with different content of stable isotope ^{18}O at all levels as an independent variable, which in the end leads to false positive results. Nevertheless, we consider such results very unlikely, especially if the measurements are repeated many times. According to the basic idea of methodological pluralism, there is as many explanations as there are alternative causes but also as many possible alternative methods to get to the scientific truth. One should pursue as many as possible of these paths. In this paper, we attempt to represent the Causal-Mechanical (CM) model of bacterial contamination in coastal bathing water by using mediator variables to bridge the gap between environmental factors such as total precipitation and anthropogenic pollution indicators such as *E. coli* or enterococci. Prior to the use of stable isotope ^{18}O as a mediator indicator variable for the analysis of groundwater dynamics in karst aquifer characterisation, the causal links between precipitation and aquifer- and marine pollution were elusive. Thus, we cannot overstate the importance of stable isotopes in completing the missing CM links as they stand for proxies of the particular water in the precipitation – groundwater – bathing water process. Another question needed to be answered for an effective governance of water commons is the speed at which water passes through the ground. We used isotopic content of the precipitation and groundwater samples as the primary mediator indicator variable (a naturally occurring tracer) to try to estimate the points of origin of the precipitation and the pass-through velocity of the groundwater. During the 2010 and 2011 bathing seasons, bi-weekly samples were collected from coastal bathing waters in the Kvarner Bay region of Croatia. Sampling sites included Bakar Bay and Kantrida Beach in the city of Rijeka. Four coastal springs were sampled weekly from April 2010 to April 2012. Only groundwater samples that coincided with sampling of marine bathing waters were included in the analysis of this study. The collected groundwater samples were stored in 50-mL double-capped polyethylene bottles until analysed by an isotope ratio mass spectrometer. We used the water equilibration technique to measure stable isotope ratios on a DeltaplusXP (Thermo Finnigan) isotope ratio mass spectrometer with an HDOeq48/24 (IsoCal) equilibration unit and a Dual Inlet system (Thermo Finnigan) as peripherals. The measurements were made against three laboratory standards ranging from -1.58 ‰ to -19.92 ‰ for $\delta^{18}O$, where $\delta^{18}O = R_{sample}/R_{standard} - 1$, with R_{sample} and $R_{standard}$ representing the ratio of ^{18}O and ^{16}O in the sample and the standard respectively. The laboratory standards were calibrated against the primary standards VSMOW2, SLAP2, and GISP. The measurements were normalised and analyzed using the USGS Laboratory Information Management System (LIMS) and the IAEA SiCalib 2.14 program for stable isotopes in water. The precision of measurement was better than 0.1 ‰.

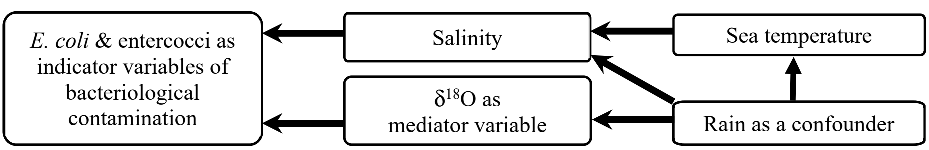

Figure 2 – $\delta^{18}O$ as a mediator variable between precipitation and indicators of anthropogenic bacteriological contamination of marine bathing water.

Panel data, also known as longitudinal data, are cross-sectional units observed over time. The use of both cross-sectional and time-series data allows for more accurate econometric model parameters and generalizations. The Generalized Method of Moments First Differences (GMM FD) is widely used in empirical research [2,3,12] and has numerous advantages over Ordinary Least Squares (OLS) and Estimated Generalized Least Squares (EGLS) Fixed Effects (FE) modelling. The GMM FD is unaffected by distributional assumptions such as heteroskedasticity, it produces the least bias and variance in parameter estimation and eliminates autoregression by explicitly including a time-lagged dependent variable into the model. Firstly, we used a static approach in the form of an EGLS model test to account for heteroscedasticity in the residuals and to incorporate spatiotemporal effects. We then performed dynamic panel data modelling in the form of a General Method of Moments (GMM) with First Differences (FD) transformation to control for unobserved time-invariant individual effect heterogeneity. The GMM estimators are consistent, asymptotically normal, and efficient because they use no information other than that contained in the moment terms. The differencing process eliminates the non-stationarity of the data, accounts for momentum and inertia, and removes the problem of location fixed effects, autocorrelation, and other time-invariant components [3,12]. As a result, Panel GMM FD is an excellent complement for EGLS estimation. Finally, to rule out autoregression as a cause of spurious correlations, we performed the Arellano-Bond test on the residuals [1].

Results and discussion

Our results are divided into two sections: static Estimated Generalized Least Squares (EGLS) Fixed Effects (FE) modelling and dynamic General Method of Moments (GMM) First Differences (FD) modelling. The latter is called dynamic because variables are differenced first to control for non-stationarity and autoregression, idiosyncratic heterogeneities, and heteroskedasticity, all problems in inferential statistics that lead to spurious correlation.

a) The static inferential analysis

First, we present the results of the static EGLS models. Static models explain well the interaction between precipitation, stable oxygen isotopes, and bacterial contamination at closed sites where seawater does not move much (Tables 1 and 2).

Table 1 – Panel EGLS for *E. coli* as a dependent variable at Bakar Bay.

Independent	Coeff.	S.E.	p	R^2
Total precipitation	0.043862	0.003372	<0.001	0.128

Total precipitation has long been used as an instrumental variable for predicting anthropogenic bacterial pollution in Croatian karst areas, since pollution is known to be washed out by rain, and rather quickly, within a few days [5]. However, it is clear from Table 2 that the $\delta^{18}O$ is a better predictor of bacterial pollution than pure information on total precipitation and also than any other indicator we could test. $\delta^{18}O$ is statistically not a determinant, but a mediator variable. Total precipitation is a confounding variable because it plays a role in the transmission of the independent variable (anthropogenic contamination in soil) to the dependent variable (bacterial contamination of coastal bathing water) but also influences the dependent variable (it lowers seawater salinity and decreases its temperature). In a CM model, it is necessary to use a mediator (for example, $\delta^{18}O$) to serve as a tracer in the transmission mechanism when there is no statistically significant relationship between the independent and dependent variables.

Table 2 – Panel EGLS for *E. coli* as a dependent variable at Bakar Bay.

Independent	**Coeff.**	**S.E.**	**p**	**R^2**
$\delta^{18}O$	7.918967	1.354843	<0.001	**0.525**
Salinity	-0.23324	0.048671	<0.001	

The model in Table 2 gives us a much better coefficient of determination (R^2=0.525). Unfortunately, combining the independent variables from Table 1 and Table 2 was not possible due to multicollinearity problems between total precipitation, $\delta^{18}O$, and salinity. Since *E. coli* is less resistant to salinity than enterococci, the salinity variable was statistically significant in the model with the *E. coli* as a dependent variable and thus was included in the model with the correct negative sign, as expected. The same could not be said for enterococci as a dependent variable, as the results did not show statistical significance for salinity at the $p<0.05$ level.

Table 3 – Panel EGLS for enterococci at Bakar Bay.

Independent	**Coeff.**	**S.E.**	**p**	**R^2**
Total precipitation	0.058916	0.005082	<0.001	0.122

Again, total precipitation and $\delta^{18}O$ were measured separately due to multicollinearity issues (Table 3). The main reason why salinity was not included in the model is the relative biological insensitivity of enterococci to salinity compared to *E. coli*, which makes it impossible to draw any statistically significant inferential conclusions regarding causal relations. The $\delta^{18}O$ outperforms all other variables in predicting enterococci in a closed bay environment (Table 4) with the coefficient of determination R^2=0.426.

Table 4 – Panel EGLS for enterococci at Bakar Bay.

Independent	**Coeff.**	**S.E.**	**p**	**R^2**
$\delta^{18}O$	11.88985	1.395156	<0.001	**0.426**

As stated earlier, enterococci are relatively more resistant to salinity and persist longer than *E. coli*. This is more exacerbated in a semi-enclosed bay such as Bakar Bay. We suspect that this is why we could not obtain statistically significant results for the conjecture that salinity reduces bacterial contamination by enterococci in semi-enclosed bays in the short term.

Table 5 – Panel EGLS for enterococci Kantrida (open sea).

Independent	Coeff.	S.E.	p	R^2
Total precipitation	0.195296	0.036986	<0.001	0.059

As shown in Table 5, total precipitation does not result in a large coefficient of determination, especially if one compares it to the model in Table 6.

Table 6 – Panel EGLS for enterococci Kantrida (open sea).

Independent	Coeff.	S.E.	p	R^2
$\delta^{18}O$	22.80015	10.54655	0.033	
Total precipitation	0.090819	0.022678	<0.001	**0.695**
Salinity	-3.36561	0.291998	<0.001	

In the open sea case at Kantrida shown in Table 6, again, the $\delta^{18}O$ oxygen isotope together with salinity and total precipitation in a common static panel EGLS model resulted in a very high coefficient of determination of 0.695.

b) The dynamic inferential analysis

Panel EGLS is a static modelling technique based on non-differenced data prone to give spurious correlations. Static EGLS modelling of bacterial contamination did not give any statistically significant results at open sea locations for short-lived and salinity non-resistant *E. coli.* Thus, it was not possible to get any meaningful statistically significant results for *E. coli,* similar to those for enterococci as shown in tables 3 to 6. This is the reason we had to change our methods to dynamic ones using first-differences of variables.

Table 7 – Panel GMM FD for *E. coli* at Bakar Bay.

Independent	Coeff.	S. E.	p	S.E.	J-stat.	AR(1) p	AR(2) p
$\delta^{18}O$	5.16	0.32	<0.001	7.31	19.14	0.05	0.796
Salinity	-0.46	0.11	<0.001				

Similar to the static EGLS model, the best results for a dynamic panel model GMM FD were obtained with the two variables $\delta^{18}O$ and salinity. In the case of the open ocean at the Kantrida site, we had to use the first difference of the oxygen isotope and precipitation data, which were one day ahead of the isotope sample data, indicating the need for careful dynamic modelling at open ocean sites that are susceptible to various dynamic influences. The most noticeable distinction between Bakar Bay and Kantrida is found in the enterococci modelling, where an autoregressive (-1) variable was required for Bakar Bay, confirming its more static nature (Table 8).

Table 8 – Panel GMM FD for enterococci at Bakar Bay.

Independent	Coeff.	S. E.	p	S.E.	J-stat.	AR(1) p	AR(2) p
enterococci(-1)	-0.49	0.01	<0.001				
$\delta^{18}O$	26.54	4.38	<0.001	19.80	40.34	0.061	0.538
Salinity	-0.26	0.02	<0.001				

Table 8 shows the null-hypothesis of the Arellano-Bond test statistics cannot be rejected at p=0.05 level indicating the constant presence of the enterococci autoregressive component (-1) even after being introduced into the GMM FD model. This is indicative of the static nature of the Bakar Bay but also of the resistance of enterococci. Table 9 shows cumulatively the dynamic panel model for *E. coli* at the open sea location at Kantrida. The models are separated by a straight line within the table.

Table 9 – Panel GMM FD for *E. coli* at Kantrida.

Independent	Coeff.	S. E.	p	J-stat.	AR(1) p	AR(2) p
$\delta^{18}O$	495.75	157.79	0.002	62.62	< 0.001	0.721
Rain day before	12.78	5.11	0.014	68.99	< 0.001	0.670
Salinity	-28.40	6.20	<0.001	57.62	< 0.001	0.684
$\delta^{18}O$	344.36	129.08	0.009			
Rain day before	17.05	7.89	0.033	47.82	< 0.001	0.654
Salinity	-30.49	8.01	<0.001			

The results of the panel GMM FD for *E. coli* at the open sea Kantrida location show the independent variables to be statistically significant at the p<0.05 level both individually as well as collectively in a model, and with correct signs.

Table 10 – Panel GMM FD for enterococci at Kantrida.

Independent	Coeff.	S. E.	p	J-stat.	AR(1) p	AR(2) p
$\Delta(\delta^{18}O)$	156.21	50.11	0.002	72.73	< 0.001	0.968
Rain day before	4.89	1.54	0.002	80.18	< 0.001	0.965
Salinity	-8.50	1.86	<0.001	69.05	< 0.001	0.844

Table 10 shows the variables in a Panel GMM FD for enterococci at Kantrida taken separately. Taken together, the standard error seems to be slightly lower, thus, giving a better model (Table 11).

Table 11 – Panel GMM FD for enterococci at Kantrida.

Independent	Coeff.	S. E.	p	J-stat.	AR(1) p	AR(2) p
$\Delta(\delta^{18}O)$	164.4491	43.11	<0.001			
Rain day before	7.135354	2.00	<0.001	54.71	< 0.001	0.785
Salinity	-8.24550	2.30	<0.001			

The dynamic panel GMM FD results for enterococci at the Kantrida open sea location give the best results when the three independent variables ($\delta^{18}O$, precipitation the day before bacteria measurement, and salinity) are shown together in a common model. We need to emphasize once again, that in a dynamic panel GMM FD model all variables are first differenced, and thereby loose the constant in the process. Although the coefficients were calculated as first moments, that is, the changes in the independent variables are calculated relative to the changes in the dependent variables, the coefficients in Table 7 represent nominal coefficients. The major shortcoming of the method concerns the elimination of the constant and the inability to compare the results to other methods due to the lacking R^2.

Conclusion

The pollutant's status as a stock or a flow is determined not only by the pollutant, but also by the medium and the environment through which it is propagating, as the medium propagating through the environment determines its pass-through velocity. Thus, the same medium, in our case the seawater, may travel slower or faster depending on whether it is a closed bay or open sea. As our panel data analysis of static $\delta^{18}O$ and dynamic $\Delta(\delta^{18}O)$ as well as of the *E. coli* and enterococci bacteria show, there is a statistically significant difference between closed bay and open sea environments. We analysed the precipitation $\delta^{18}O$ data from hinterland locations that might potentially coincide with the locations where anthropogenic pollution originates, and subsequently we analysed $\delta^{18}O$ data springs at several locations near the beaches. Without $\delta^{18}O$ as a mediator variable, no statistically significant results were obtained. With ^{18}O as a naturally occurring tracer as a mediator indicator variable, we were able to obtain acceptable and expectable results. The measures considered optimal for stock and flow pollutants differ. The amount of pollution that renders a water source (a spring) unusable might be at a point where the marginal benefit of an additional unit of pollutant discharge changes very rapidly. In such cases, it is not optimal to regulate prices or impose taxes, but instead to set total allowable quantities of pollutants with tradable permits or quotas. Mixed systems may be optimal in some situations.

Funding

This work was supported by the University of Rijeka under the project numbers: uniri-pr-prirod-19-24, uniri-drustv-18-284-1456 and uniri-biomed-18-292.

References

[1] Arellano, M.; Bond, S. (1991) - *Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations*. Rev. Econ. Stud. 58, 277–297. https://doi.org/10.2307/2297968

[2] Arellano M. (2003) - *Panel Data Econometrics*; Oxford University Press: Oxford, UK.

[3] Baltagi, B. H. (2021) Econometric Analysis of Panel Data, Sixth Edition. Springer Nature, Cham, Switzerland. https://doi.org/10.1007/978-3-030-53953-5

[4] Mance D, Krunić K, Mance D. (2021) - *Protecting Species by Promoting Protected Areas and Human Development—A Panel Analysis*. Sustainability. 13(21):11970. https://doi.org/10.3390/su132111970

[5] Mance, D., Mance, D., Vukić Lušić, D. (2021) - *Environmental and anthropogenic factors affecting coastal bathing water quality: preliminary study for Primorje-Gorski Kotar County (Croatia)*. Pomorstvo: scientific journal of maritime research, 35(1):57-68. https://doi.org/10.31217/p.35.1.7

[6] Mance, D., Mance, D., Vukić Lušić, D. (2018) - *Marine Pollution Differentiation with Stable Isotopes of Groundwater*. Pomorstvo: scientific journal of maritime research, 32 (1), 80-87. https://doi.org/10.31217/p.32.1.10

[7] Mance, D., Mance, D., Vukić Lušić, D. (2018) - *Environmental isotope O-18 in coastal karst spring waters as a possible predictor of marine microbial pollution*. Acta Adriatica, 59(1):3-16. https://doi.org/10.32582/aa.59.1.1

[8] Mance, D.; Vilke, S.; Debelić, B. (2020) - *Sustainable Governance of Coastal Areas and Tourism Impact on Waste Production: Panel Analysis of Croatian Municipalities*. Sustainability, 12(18), 7243; https://doi.org/10.3390/su12187243

[9] Ostrom, E. (1990) Governing the Commons. Cambridge University Press.

[10] Salmon, W. (1971)- *Statistical Explanation and Statistical Relevance*. In *Statistical Explanaton and Statistical Relevance*, ed. Wesley Salmon. University of Pittsburgh Press.

[11] Woodward J. (1989) The causal mechanical model of explanation. Minnesota Studies in the Philosophy of Science 13:359-83.

[12] Wooldridge, J. M. (2010) - *Econometric Analysis of Cross Section and Panel Data*, second edition. MIT Press, Cambridge, Massachusetts.