The Influence of Atmospheric Methane on Human Health

Assistant Cristina Stirbu PhD Student
"Artifex" University of Bucharest

Abstract

In this article we try to determine whether there is a relation between the quantity of athmospheric methane and the number of diseases of circulatory system, at the level of the whole country.We first analyze the time series of the quantity of pollutant between 1990 and 2008 and then we predict the values for the years 2009-2012. In the end we calculate the correlation coefficient between the polution and the number of circulatory system diseases.

Key words: *Time series, trend, seasonal deviation, least square method, correlation coefficient.*

The methane, the main component of natural gas, is used in domestic economy for cooking and central heating, and in the same time is largely consumed in industrial processes, thus a large quantity of this compound is released into atmosphere. Methane is a colourless and odourless gas and also toxic for human organism. Released into atmosphere, it contributes to global warming, in a higher proportion than CO2, although the life time of CH4 is shorter than that of carbon dioxide.

In this article we tried to emphasize the effect that methane has on human body, respectively on how it influences the circulatory system diseases occurrence.

The values measured in tones for the pollutant quantities, between years 1990 and 2008, are presented in table 1. We collected the data from the National Statistical Insitution site, and the calculations were made using Excel 2003.

Regarding data analisys, we operate with time series, and we addopt the following additive model:

$$y = t + S + r \qquad (1)$$

where:

y is a known value of the time series ;
t is the trend ;
S is the seasonal component;
r is the residue ;

The evaluation and interpretation of the three components is made with detailed analysis. Each component will be treated separatelly.

Regarding the trend, this will be evaluated by the least square method, using a linear adjustment.

$$t = ax + b \qquad (2)$$

where the coefficients a and b are calculated by the formulas :

$$a = \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} \qquad (3)$$

and

$$b = \frac{\Sigma y}{n} - b\frac{\Sigma x}{n} \qquad (4)$$

Using the values from table 1, we get :

a=-27187,9973 and b=1628409,5.

To obtain the trend values, we replace the x values of time instants into equation (2) .

Table 1

x (year)	y CH4 (tons)	xy	x^2	t (Trend)	y-t
1	**1931957,22**	1931957,22	1	1601221,464	330735,7564
2	**1695267,42**	3390534,84	4	1574033,466	121233,9537
3	**1532579,05**	4597737,15	9	1546845,469	-14266,41898
4	**1448614,96**	5794459,84	16	1519657,472	-71042,51168
5	**1400681,89**	7003409,45	25	1492469,474	-91787,58439
6	**1460324,24**	8761945,44	36	1465281,477	-4957,237088
7	**1487613,12**	10413291,84	49	1438093,48	49519,64021
8	**1312724,74**	10501797,92	64	1410905,482	-98180,74249
9	**1200511,99**	10804607,91	81	1383717,485	-183205,4952
10	**1168105,07**	11681050,7	100	1356529,488	-188424,4179
11	**1196917,6**	13166093,6	121	1329341,491	-132423,8906
12	**1189067,15**	14268805,8	144	1302153,493	-113086,3433
13	**1227544,31**	15958076,03	169	1274965,496	-47421,186
14	**1277007,78**	17878108,92	196	1247777,499	29230,2813
15	**1249915,59**	18748733,85	225	1220589,501	29326,0886
16	**1260126,89**	20162030,24	256	1193401,504	66725,38589
17	**1275081,88**	21676391,96	289	1166213,507	108868,3732
18	**1235956,5**	22247217	324	1139025,51	96930,99049
19	**1224062,87**	23257194,53	361	1111837,512	112225,3578
190	25774060,27	242243444,2	2470		

We further determine the seasonal component. To this end, we first divide the time interval of 19 years into 5 periods; the first 4 periods of 4 years each, and the last one of three years. We will follow the steps:

- We calculate the y-t values (we subtract the trend from the known y values) for each time instance (for each year);
- We determine the arithmetic mean for each of the five periods ;
- If the sum of all arithmetic means is different from zero, we adjust one or more means for the total to be zero.

As we can see from table 2 in which we made these calculations, the sum of the arithmetic means of the five periods is very close to zero, so it is insignifficant and we can take it as zero.

Table 2

	The seasonal deviation S			
	year1	year2	year3	year4
Period1	330735,7564	121233,9537	-14266,41898	-71042,51168
Period2	-91787,58439	-4957,237088	49519,64021	-98180,74249
Period3	-183205,4952	-188424,4179	-132423,8906	-113086,3433
Period4	-47421,186	29230,2813	29326,0886	66725,38589
Period5	108868,3732	96930,99049	112225,3578	0
Total	117189,864	54013,57053	44380,77702	-215584,2116
Arithmetic means	29297,46601	13503,39263	11095,19425	-53896,05289
Sum	-5,82077E-10			

We further filter the time series using the deviations on each period, see table 3 :

Table 3

period	year	y	S	Filtered values on each season(y-s)
1	1	**1931957,22**	29297,46601	1902659,754
	2	**1695267,42**	13503,39263	1681764,027
	3	**1532579,05**	11095,19425	1521483,856
	4	**1448614,96**	-53896,05289	1502511,013
2	1	**1400681,89**	29297,46601	1371384,424
	2	**1460324,24**	13503,39263	1446820,847
	3	**1487613,12**	11095,19425	1476517,926
	4	**1312724,74**	-53896,05289	1366620,793
3	1	**1200511,99**	29297,46601	1171214,524

period	year	y	S	Filtered values on each season(y-s)
	2	**1168105,07**	13503,39263	1154601,677
	3	**1196917,6**	11095,19425	1185822,406
	4	**1189067,15**	-53896,05289	1242963,203
4	1	**1227544,31**	29297,46601	1198246,844
	2	**1277007,78**	13503,39263	1263504,387
	3	**1249915,59**	11095,19425	1238820,396
	4	**1260126,89**	-53896,05289	1314022,943
5	1	**1275081,88**	29297,46601	1245784,414
	2	**1235956,5**	13503,39263	1222453,107
	3	**1224062,87**	11095,19425	1212967,676

We calculate the rezidual deviation r by the formula r=y-t-S (see table 4).

Table 4

Period	**Year**	y	s	t	y-t-s
1	**1**	**1931957,22**	29297,46601	1601221,464	301438,2904
	2	**1695267,42**	13503,39263	1574033,466	107730,5611
	3	**1532579,05**	11095,19425	1546845,469	-25361,61323
	4	**1448614,96**	-53896,05289	1519657,472	-17146,45879
2	**1**	**1400681,89**	29297,46601	1492469,474	-121085,0504
	2	**1460324,24**	13503,39263	1465281,477	-18460,62972
	3	**1487613,12**	11095,19425	1438093,48	38424,44596
	4	**1312724,74**	-53896,05289	1410905,482	-44284,6896
3	**1**	**1200511,99**	29297,46601	1383717,485	-212502,9612
	2	**1168105,07**	13503,39263	1356529,488	-201927,8105
	3	**1196917,6**	11095,19425	1329341,491	-143519,0848
	4	**1189067,15**	-53896,05289	1302153,493	-59190,29041
4	**1**	**1227544,31**	29297,46601	1274965,496	-76718,65201
	2	**1277007,78**	13503,39263	1247777,499	15726,88867
	3	**1249915,59**	11095,19425	1220589,501	18230,89435
	4	**1260126,89**	-53896,05289	1193401,504	120621,4388
5	**1**	**1275081,88**	29297,46601	1166213,507	79570,90718
	2	**1235956,5**	13503,39263	1139025,51	83427,59786
	3	**1224062,87**	11095,19425	1111837,512	101130,1635
TOTAL		25774060,27	53896,05289	25774060,27	-53896,05289

The residual deviation is small compared to cumulated values of variable y, so we can make prognosis based on presented data. (The data prognosis for a future period can be done only if we have results for at least three past cycles).

The previsioned values for the pollutant in the next 4 years, 2009-2012, are obtained with the aid of the relation :

$$y_{forecast} = t_{forecast} + S \qquad (5)$$

Where :

$y_{forecast}$ *are* the estimated values

$t_{forecast}$ are the estimated values for the trend

S *is* the seasonal component

Table 6

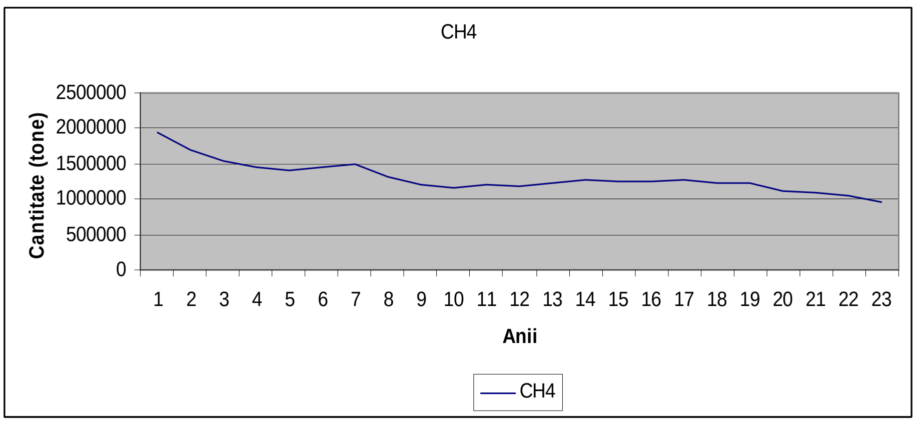
trend t	S	Y	Ciculatory System diseases (thousands cases)
1601221,464	29297,46601	**1931957,22**	
1574033,466	13503,39263	**1695267,42**	
1546845,469	11095,19425	**1532579,05**	
1519657,472	-53896,05289	**1448614,96**	427
1492469,474	29297,46601	**1400681,89**	445
1465281,477	13503,39263	**1460324,24**	459
1438093,48	11095,19425	**1487613,12**	471
1410905,482	-53896,05289	**1312724,74**	486
1383717,485	29297,46601	**1200511,99**	492
1356529,488	13503,39263	**1168105,07**	507
1329341,491	11095,19425	**1196917,6**	543
1302153,493	-53896,05289	**1189067,15**	614
1274965,496	29297,46601	**1227544,31**	646
1247777,499	13503,39263	**1277007,78**	608
1220589,501	11095,19425	**1249915,59**	625
1193401,504	-53896,05289	**1260126,89**	560
1166213,507	29297,46601	**1275081,88**	682
1139025,51	13503,39263	**1235956,5**	609
1111837,512	11095,19425	**1224062,87**	657
1084649,515	29297,46601	1113946,981	682
1057461,518	13503,39263	1084468,303	**713**
1030273,52	11095,19425	1052463,909	**642**
1003085,523	-53896,05289	949189,4701	607

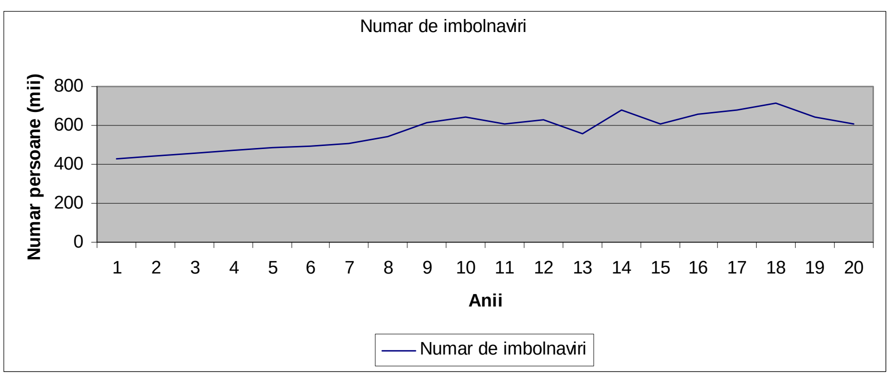
The last 4 values from the trend column were estimated by the "trend" function from Excel, which calculates the forecast values with the aid of least squares method. In the column of y, of the pollutant, the values recorded by year 2008 were kept, and from the year 2009 until year 2012 the values were calculated by formula (5).

The correlation coefficient between the pollutant and the number of diseases, has the value -0,67380708, and it is calculated by formula (6).

We thus found a strong relation between the polution and the number of diseases: although the emission of gas has a slow decrease, the number of sick persons increases. We can visualize that in the graphs below.

$$\rho = \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{n\Sigma x^2 - (\Sigma x)^2}\sqrt{n\Sigma y^2 - (\Sigma y)^2}} \qquad (6)$$





References

Andy Francis " Business Mathematics and Statistics 4th edition", Editura Tehnica 2005;

Emilia Titan "Statistica.Teorie. Aplicatii in sectorul tertiar" , Editura Meteor Press;

www.insse.ro.