The Impact of Level of Solar Activity on Mortality by Cause in Longtime Period

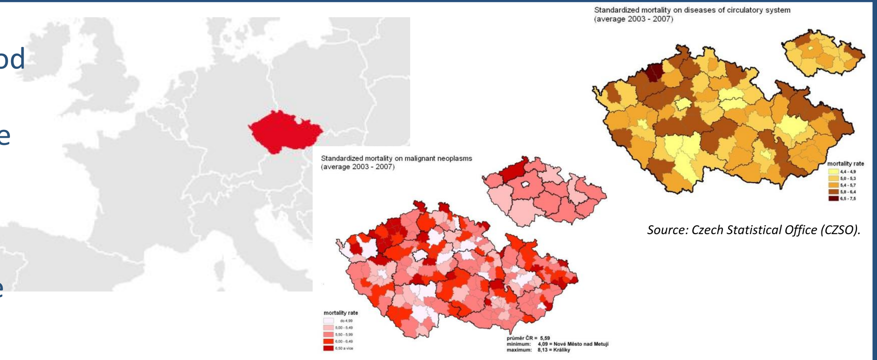
Kateřina Podolská

Departement of Demography and Geodemography, Faculty of Science, Charles University in Prague,
Institute of Atmospheric Physics ASCR, Prague, Czech Republic; kapo@ufa.cas.cz

The aim of this presentation is to show the dependence of the intensity of mortality on the chosen causes of death in the Czech Republic, according to ICD-10, on the solar activity during the increasing and decreasing phases of the solar cycles No.23 and No.24 in the period 1994-2011. We use the methods of multivariate statistical analysis. The typology of time profiles for the causes of death is identified with the help of cluster analysis using time. The solar activity is represented by the indices R, Kp, F10.7 and Dst, and also by the foF2 height of the F2 layer and TEC. There are investigated 67 of individual causes of death according to ICD-10 of groups II.Neoplasms, VI.Diseases of the nervous system, IX.Diseases of the circulatory system and XVII.Congenital malformations, deformations and chromosomal abnormalities. The correlation between the intensity of mortality from cardiovascular disease e.g. I21 (Acute myocardial infarction) and I64 (Brain stroke) and the solar activity parameters is discovered, as well as a stronger dependence on the height of the F2 layer and TEC. It was found congenital defect Q91 (Edwards and Patau syndrom) dependence on these parameters and also cardiovascular congenital defect e.g. Q20, Q21, Q23, Q25, Q26.

DATA AND REGIONAL DELINEATION

The mortality in the Czech Republic (population of about 10 million) in a investigated period of years from 1994 to 2011 has significantly decreased. Registration of deaths by cause of death is very stable in the Czech Republic in long-term. The cardiovascular diseases are the leading causes of death in the CR and in most European populations. The ischemic heart diseases and the cerebrovascular diseases accounted for 70% of all deaths on the cardiovascular causes of death. As reported Bruthans (2011) mortality on cerebrovascular diseases declines further but to other causes of death from this group is still stagnant. The numbers of deaths on main cardiovascular causes of death in the period 1994–2011 are plotted in the Figure 1.



The solar activity indices **Kp** planetary index, **R** (Relative number of sunspots), **AE** (Auroral Electrojet), **F10.7** (Intensity of the Sun radio flux) and **Dst** (Disturbance Storm Time) represent solar activity in models with solar paremetres. Physical parameters describe the ionospheric effects are **foF2** critical frequency of the ionospheric F2 layer and **TEC** (Total Electron Content) and were included to the model with ionospherical parameters. The ionospheric parameters foF2 and TEC were related to the geographic location of the Czech Republic..

The time series of daily aggregated numbers of deaths by cause, men and women together were used. The data were provided by Czech Statistical Office. Whole observed period belongs to time of validity of ICD-10. For detailed analysis were chosen 67 of causes of death of the chapters II. Neoplasms (C43, C44, C69, C70, C71, C72, C73, C81, C84, C85), VI. Diseases of the nervous system (G20, G21, G30, G31, G35, G36, G37, G40, G80, G90), IX. Diseases of the circulatory system (I20, I21, I22, I26, I44, I45, I46, I47, I48, I49, I50, I60, I61, I62, I63, I64, I67, I74) and XVII. Congenital malformations, deformations and chromosomal abnormalities (Q02, Q03, Q04, Q07, Q20–Q26, Q28, Q31, Q33, Q39, Q43, Q44, Q60–Q64, Q79, Q85, Q86, Q87, Q89, Q90, Q91, Q96) of ICD-10 with higher daily numbers of deaths.

The correction in the numbers of died persons were made for elimination of coding effect in 2011 year data (Poppová,2011). The effect of encoding is shown in the Table 1. The decline in mortality by cause, especially for cardiovascular diseases (see Figure 1), in the CR followed the total numbers of deaths in the reference period 1994–2011. The time series of the numbers of deaths on the individual causes of death was for further analysis rectified on the total number of deaths in the CR in a given year.

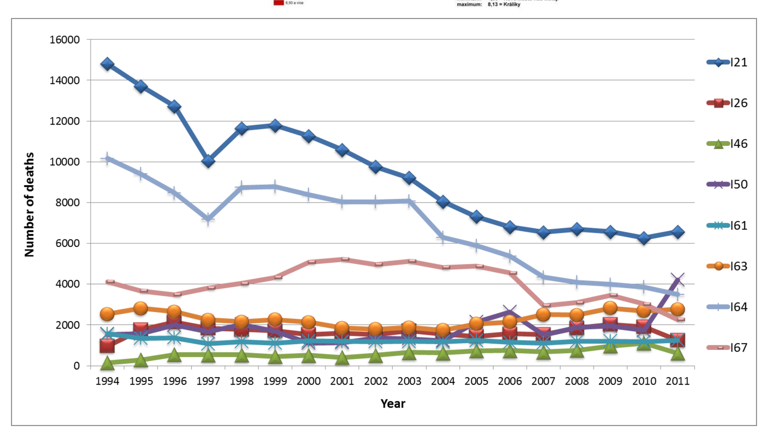


Fig.1: *Numbers of deaths on main causes of diseases of cardiovascular system in the Czech Republic, 1994–2011*

Diagnosis Codes	*Causes of death*	*Coding effect in 2011 (2010=1)*
I21	Acute myocardial infarction	1,08
I26	Pulmonary embolism	0.69
I46	Cardiac arrest	0,4
I50	Heart failure	1,83
I61	Nontraumatic intracerebral hemorrhage	0.74
I63	Cerebral infarction	0.74
I64	Stroke, not specified as haemorrhage or infar	0.74
I67	Other cerebrovascular diseases	0.88

Table 1: *List of causes of death with coding effect*

METODOLOGY

- **The cluster analysis** applied to time series data for study method of mortality on cardiovascular causes of death was used. Cluster analysis, often referred to as segmentation in contexts of large data sets. While segments are often identified using static characteristics, evolving systems may be better described by how processes change over time. In this special case is this method used to investigate the possible influence of abnormal solar events to mortality on selected diseases in a period of years from 1994 to 2011. We looked for similar patterns in time profiles of numbers of deaths on individual causes of death and compare it with timing of the solar cycle and the occurrence of abnormal solar events. Cluster analysis was performed in the SAS 9.2TM software, using procedure FASTCLUS applied to time series. In detail is application of the procedure described in (Corliss, 2012). This computational technique is used to demonstrate how time intervals may be used to identify subgroups of daily numbers of deaths in segmentation in cluster analysis. The object entering into the calculation is realization of a random vector with components *Year* and *Daily number of deaths* described by the equation [1]:

$$X = (Year, Daily\ number\ of\ deaths) \qquad [1]$$

The number of realizations of random vector in the reference period 1994–2011 is 6573. During the calculation data was repeatedly distributed to 1–5 clusters in 10 iterations and the optimal number of clusters is selected. The homogeneity of the groups is measured by cluster mean square deviation. **Cluster centroid** is the median time variables, the random vector component Year in the cluster. Each daily number of deaths from the examined time series is within the computational processes classified into groups with similar characteristics. The time distributions of belonging to groups were compared with the timing phase of the solar cycle.

- **The linear regression model** performed in SAS 9.2TM by procedure REG to verify the number of deaths dependence on solar and ionospheric parameters during the solar cycle No.23 was used. The linear regression model with regressors solar parameters is described by the equation [2]:

$$Y = b_0 + b_1Kp + b_2R + b_3AE + b_4F10.7 + b_5Dst + e_i. \qquad [2]$$

Linear regression model with regressors ionospheric parameters is described by the equation [3]:

$$Y = b_0 + b_1foF2 + b_2TEC + e_i, \qquad [3]$$

where: *Y* is dependent variable of time series of number of death e.g. I21, I64; *Kp, R, AE , F10.7, Dst, foF2, TEC* are explanatory variable of physical parameters, b_i are the regression coefficients, e_i is the error term.

The best model is selected by the value of the adjusted coefficient of determination, which indicates the percentage of variability of the number of deaths by cause of death explained by the regression model for the given data. Variables were included from the model at a significance level of 0.05.

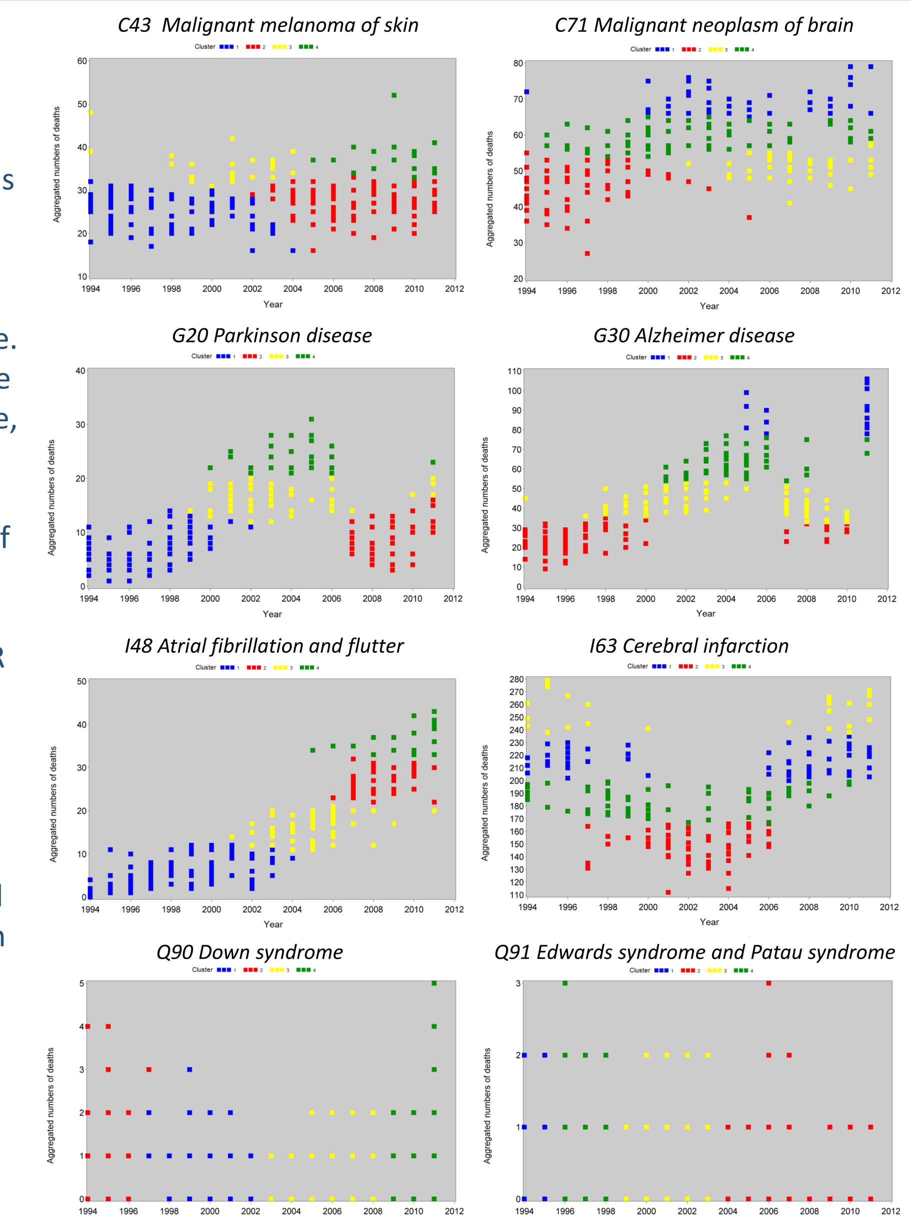
DATA SOURCES

Number of deaths by cause: Czech Statistical Office (CZSO).
Kp index: World Data Center for Geomagnetism, Kyoto University, Japan.
R, AE index, Dst, F10.7: Space Physics Interactive Data Resource, National Geophysical Data Center, Boulder, USA.
Solar cycle evolving: NWRA/CoRA, NorthWest Research Associates, Boulder, USA, Deutsches GeoForschungs Zentrum, Helmholtz-Zentrum, Germany.
foF2: UK Solar System Data Centre, Rutherford Appleton Laboratory, Oxfordshire, GB, ionosonde JR055 Juliusruh/Rugen.
TEC: Institut Géographique National (IGN), France.

TIME PROFILES OF CAUSES OF DEATH TYPOLOGY

Typology of time profiles of numbers of deaths from selected causes of death was identified using time series cluster analysis. The calculation was made separately for each series of the number of deaths from chosen causes of death. Division into clusters revealed that for most causes of death are grouped into four clusters in terms of the daily number of deaths. For a different number of clusters is not classification significant.

II. Neoplasms (C00-D48) most of years centroids occur at the time of maxima and minima of solar cycle. Variability between centroids in this group of diagnoses is large.
VI. Diseases of the nervous system (G00-G99) between the diagnoses are observable differences in the distribution of centroids, especially in the maxima of the solar cycle, the group does not have common features. The reaction of the nervous system to electromagnetic phenomena is differentiated by gender.
IX.Diseases of the circulatory system (I00 – I99) occurs most years cluster centroids if found in the maximum of the solar cycle in 2000–2003 slightly less than the upward phase of the cycle and at the minimum before the start. The graphs in Fig. 4 show location of the calculated clusters centroids in the chart of relative sunspot number R during solar cycle No.23. The clusters centroids are placed at the maximum and the both of minima of the solar cycle. This suggests to changes the nature of the process in years of high solar activity and in the ascending phase of the solar cycle when culminates in the amount of high-energy protons in the solar wind and geomagnetic activity variations. This is consistent with the results of medical studies conducted classical correlation analysis (Cornelissen, 2002) and (Stoupel, 2011), which observed changes in the type of cardiovascular mortality, depending on level of solar activity in the long term.
XVII. Congenital malformations, deformations and chromosomal abnormalities (Q00-Q99) significant incidence of centroids groups in 1995 and somewhat less pronounced in 2010, at a time of the minima between solar cycles No.22/23 and No.23/24. For this group of causes of death can be seen a significant difference between the time profile of the maximum and minimum of the solar cycle.



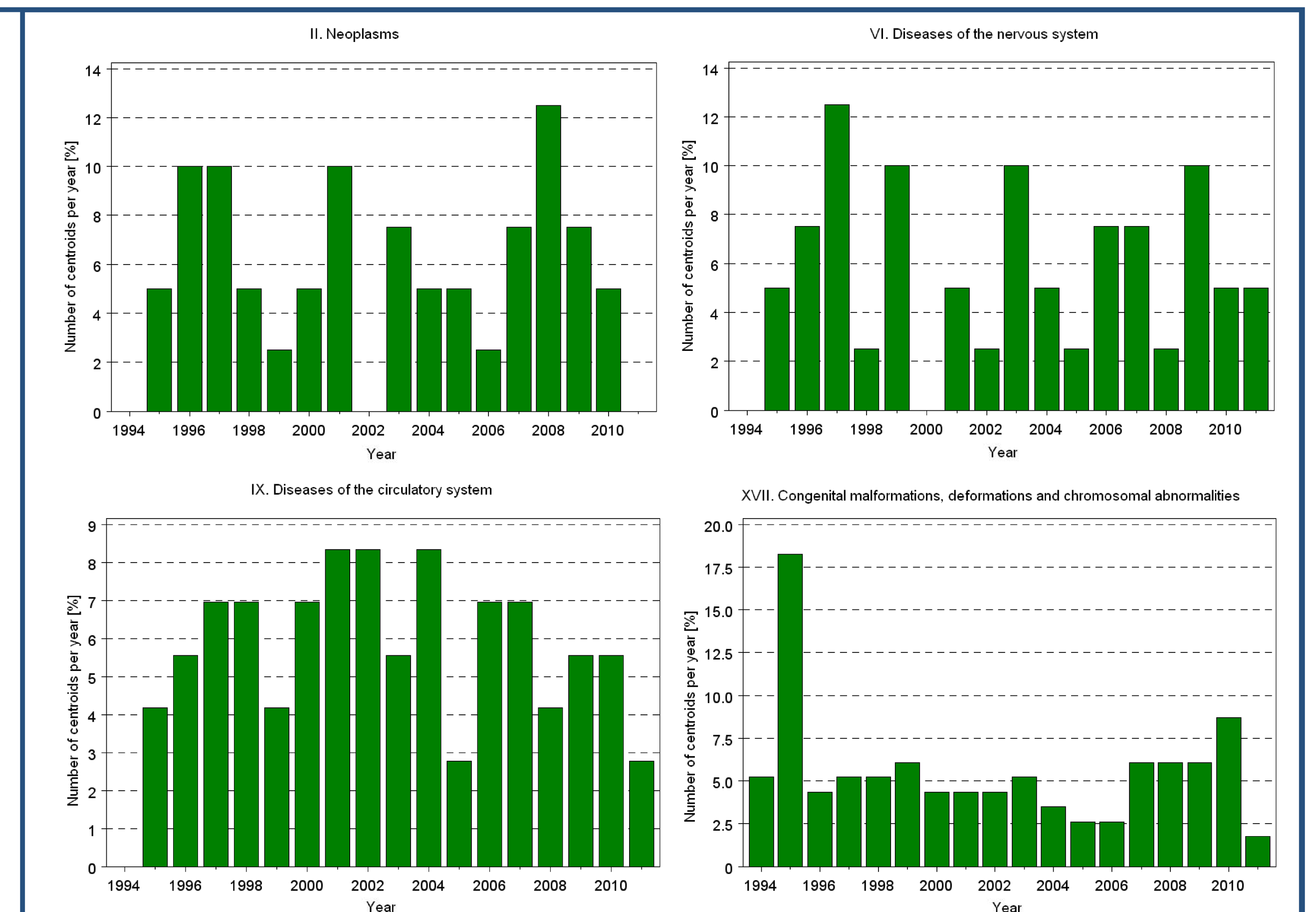


Fig.3: *Histograms of the clusters centroids by ICD-10 group of diagnosis.*
Distribution of clusters in the beginning, end and duration of solar cycle No.23 over a period of years 1994–2011 may reflect changes of occurrence of abnormal solar events. The vertical axis represents number of centroids of the clusters per year, and the horizontal axis is the year of observation. Clusters are differentiated according to the level of mortality and its spread over the reference period.

LINEAR REGRESSION MODELS

Causes of death	b0	Kp	R	AE	F10,7	Dst	R-squared
I21 Acute myocardial infaction	27,60	4,00	0,05	-0,02	-0,04	-0,01	0,06
I45 Other conduction disorders	1,17	0,07	0,00	-0,01	-0,02	-0,03	0,07
I64 Brain stroke	17,68	2,85	0,03	-0,02	-0,02	-0,05	0,08
I67 Other cerebrovascular diseases	8,03	0,74	-0,02	-0,03	0,04	0,02	0,10
Q04 Other congenital malformations of brain	1,13	-0,07				-0,02	0,09
Q20 Congenital malformations of cardiac chambers	1,12		0,01		-0,01	-0,01	0,06
Q21 Congenital malformations of cardiac septa	0,91	0,16	-0,02	-0,01	0,01		0,07
Q23 Congenital malformations of aortic	1,29		0,03		-0,01	0,01	0,08
Q25 Congenital malformations of great arteries	0,84	0,05	-0,02	-0,01	0,02	-0,01	0,06
Q26 Congenital malformations of great veins	1,15		-0,03				0,05
Q87 Other specified congenital malformations	0,97	0,15		-0,01			0,07
Q90 Down syndrome	1,21	-0,07	0,02	0,01	-0,02		0,06
Q91 Edward's and Patau's syndromes	0,98		0,03	-0,03		-0,01	0,45

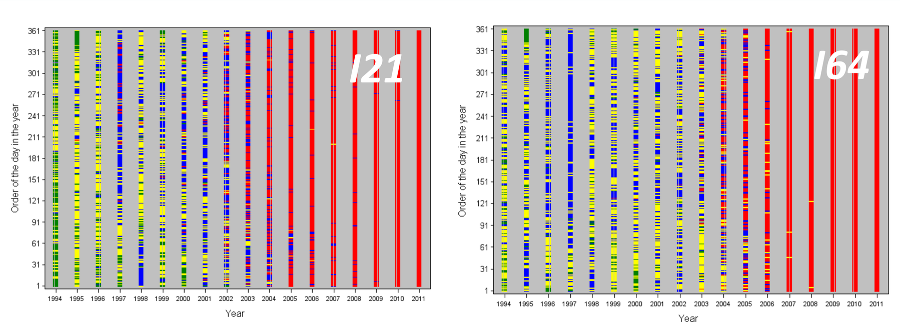
Table 2: *Estimated coefficients of the linear regression model with solar parameters regressors.*

Causes of death	b0	foF2	TEC	R-squared
I21 Acute myocardial infaction	17,04	0,49	0,03	0,25
I46 Cardiac arrest	3,74		-0,01	0,06
I48 Atrial fibrillation and flutter	2,13		-0,01	0,05
I50 Heart failure	8,21	-0,12	-0,01	0,05
I63 Cerebral infarction	8,66	-0,10	-0,01	0,05
I64 Brain stroke	11,74	0,69	0,03	0,20
I67 Other cerebrovascular diseases	9,58	0,40	0,01	0,07
Q89 Other congenital malformations	0,96		0,00	0,05
Q90 Down syndrome	1,50	-0,14	0,00	0,09
Q91 Edward's and Patau's syndromes	1,49	-0,17	0,00	0,15

Table 3: *Estimated coefficients of the linear regression model with ionospheric parameters regressors.*

Summary of results of the linear regression models with solar activity parameters Kp, R, AE, F10.7 and Dst are shown in the Table 2. The regression models with ionospheric parameters foF2 and TEC are shown in the Table 3. Negative coefficient for regressors in the models corresponds to the negative partial correlation. If the explanatory variable is excluded from the model, are not the appropriate parameters in the table above. Verification of conformity of the model with the data was performed using coefficient of determination, which indicates the percentage of variability of the number of deaths by cause of death explained by the regression model for the given data. When were gradually excluded all explanatory variables from the model, model not listed in the table. If for cause of death does not exist statistically significant model is not listed in the tables at all. The residual analysis was performed in the calculation, we calculated 95 % confidence intervals for estimates of the mean value of the dependent variable, 95 % confidence intervals for individual estimates and 95 % confidence intervals for estimates of model parameters. The computed model is consistent with the results of (Stoupel 2002) and (Szczeklik, 1984), where for acute myocardial infarction found negative correlation to shifted indices of solar activity.

The classical linear regression model with regressors of solar parameters explained more than 5 % of the variability in the causes of death Q04, Q20, Q21, Q23, Q25, Q26, Q87, Q90 and 45 % variability for the cause of death Q91 Edwards and Patau syndromes.



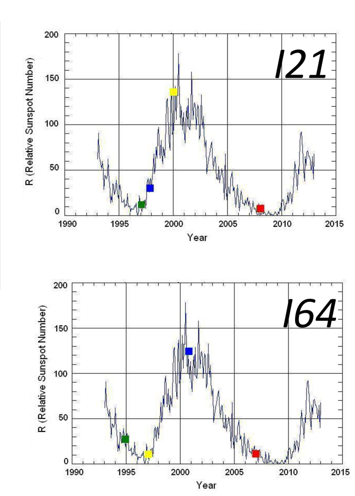


Fig. 4: Time profiles of causes of death I21 and I64
In this plot, the time changes in number of death by day are reflected in the identified clusters. The graph in left displays the result of Acute myocardial infarction I21, very coherent cause of death in time. The figure in right illustrates the result of same analysis for the Brain stroke diagnosis I64. We can observe the subsequent reaction during a solar cycle compared to myocardial infarction. We see that cluster 4 in red looks very homogeneous for this both causes in the period from 2005. The linear regression model proved for these two causes that dependence on solar and ionospheric parameters is weaker after the drop in the number of deaths at a lower level. In right column are plotted clusters centroids on a graph of relative sunspot number R during solar cycle No.23 for I21 Acute myocardial infarction (up) and I64 Brain stroke (down).

CONCLUSIONS

The obtained result is in accordance with the point of view, which was not direct correlation between the number of deaths on Acute myocardial infarction (I21) and Brain stroke (I64) with geomagnetic solar index Ap in the maxima of the solar cycle.

The ionospheric parameters foF2 and TEC explain greater variability for Acute myocardial infarction (I21) and Brain stroke (I64) than the model with solar parameters. They explain 24.8 % of variability for the cause of death Acute myocardial infarction (I21), and 20.5 % for Brain stroke (I64). Thus, the analysis shows that the ionospheric parameters foF2 and TEC may, due to the geographically specific values, better explain the variability in the number of deaths than the solar indices. The cardiovascular diseases thus respond to changes in solar activity and occurrence of abnormal solar events indirectly through a concentration of electrical charge in the environment in which the monitored population lives.

The classical linear regression model with regressors of solar parameters explained more than 5 % of the variability in the causes of death Q04, Q20, Q21, Q23, Q25, Q26, Q87, Q90 and 45 % variability for the cause of death Q91 Edwards and Patau syndromes.

ACKNOWLEDGEMENTS

This study was supported in part by Charles University in Prague, Faculty of Science, Department of Demography and Geodemography and Institute of Atmospheric Physics Academy of Sciences of the Czech Republic.

REFERENCES

Corliss, D.J., „Dynamically Evolving Systems: Cluster Analysis Using Time“, *SAS Global Forum*, 2012, Paper 329-2012, pp. 2-3.
Cornelissen, G. et al. “Non-photic Solar Associations of Heart Rate Variability and Myocardial Infarction.” *Journal of Atmospheric and Solar-Terrestrial Physics 64* (2002): 707.
Bruthans, J. “Pokles úmrtnosti na ICHS a jeho hlavní příčiny.” *Cor Vasa 53* (2011): (4–5).
Poppová, M. „IRIS: jazykově nezávislý software pro kódování příčin smrti - implementace v ČR.“ *Demografie 53,* (2011)
Stoupel, E. “The effect of geomagnetic activity on cardiovascular parameters.” *Biomed. Pharmacother. 56* (2002): 247.
Stoupel, E. et al.. “Twenty years study of solar, geomagnetic, cosmic ray activity links with monthly deaths number” [online]. [cit. 2011-06]. (2011), http://www.SciRP.org/journal/jbise/.
Szczeklik, E.et al. “Solar activity and myocardial infarction” *Cor Vasa 25* (1984): 49-55.