Available online at www.sciencedirect.com
SCIENCE DIRECT®
PERGAMON
Applied Radiation and Isotopes 58 (2003) 697–706
Applied Radiation and Isotopes
www.elsevier.com/locate/apradiso

Application of decision trees to the analysis of soil radon data for earthquake prediction

B. Zmazek*, L. Todorovski, S. Džeroski, J. Vaupotič, I. Kobal

J. Stefan Institute, Ljubljana 1000, Slovenia

Received 7 February 2003; accepted 20 March 2003

Abstract

Different regression methods have been used to predict radon concentration in soil gas on the basis of environmental data, i.e. barometric pressure, soil temperature, air temperature and rainfall. Analyses of the radon data from three stations in the Krško basin, Slovenia, have shown that model trees outperform other regression methods. A model has been built which predicts radon concentration with a correlation of 0.8, provided it is influenced only by the environmental parameters. In periods with seismic activity this correlation is much lower. This decrease in predictive accuracy appears 1–7 days before earthquakes with local magnitude 0.8–3.3.

Keywords: Radon in soil gas; Environmental parameters; Earthquakes; Correlation; Regression trees; Forecasting

1. Introduction

Radon in groundwater was monitored for the first time in Uzbekistan after the Tashkent earthquake (Ulomov and Mavashev, 1971) and it became known that seismogenic processes influence the behaviour of underground fluids (Scholz et al., 1973; Mjachkin et al., 1975). Since then, temporal variations of radon concentration in soil gas and in groundwater have been studied and related to seismic activity in many countries (Teng, 1980; Sultankhodajev, 1984; Wakita, 1979; Ui et al., 1988; Wakita et al., 1988; Singh et al., 1993; Singh and Virk, 1994; Igarashi et al., 1995; Planinić et al., 2000, 2001).

The first radon measurements in Slovenia aimed at predicting earthquakes were made in 1982. In four thermal water sources, radon concentrations were determined weekly, and Cl^-, SO_4^{2-}, hardness and pH, monthly (Zmazek et al., 2000a). This frequency of analyses, however, is not high enough to follow seismic activity properly. So, in 1998, we extended our study from thermal waters (Zmazek et al., 2000b, 2002a, 2002b) to soil gas (Zmazek et al., 2000c, 2002c) and increased the sampling frequency up to once an hour. In both cases, meteorological parameters such as barometric pressure, air temperature, soil temperature and rainfall were also taken into account.

As is the general practice (Yasuoka and Shinogi, 1997; Singh et al., 1999; Virk et al., 2001), anomalies in radon concentration observed in our previous study (Zmazek et al., 2002c) were identified and then related to seismic activity. Some of these anomalies with respect to barometric pressure and air temperature can be seen in Fig. 1. For small earthquakes, it is often impossible to identify an anomaly as resulting solely from seismic activity and not from meteorological or hydrological parameters. Therefore, the implementation of more advanced statistical methods in data evaluation (Di Bello et al., 1998; Cuomo et al., 2000; Biagi et al., 2001; Belayaev, 2001; Negarestani et al., 2001) appears to be essential.

In this paper, regression trees are applied for the first time to earthquake prediction, based on radon in soil gas. Data from our previous study (Zmazek et al., 2002a) are used. Data mining and machine learning methods have been applied to many problems in

*Corresponding author. Tel.: +386-1-477-3580; fax: +386-1-477-3811.
E-mail address: boris.zmazek@ijs.si (B. Zmazek).

0969-8043/03/$ - see front matter
doi:10.1016/S0969-8043(03)00094-0

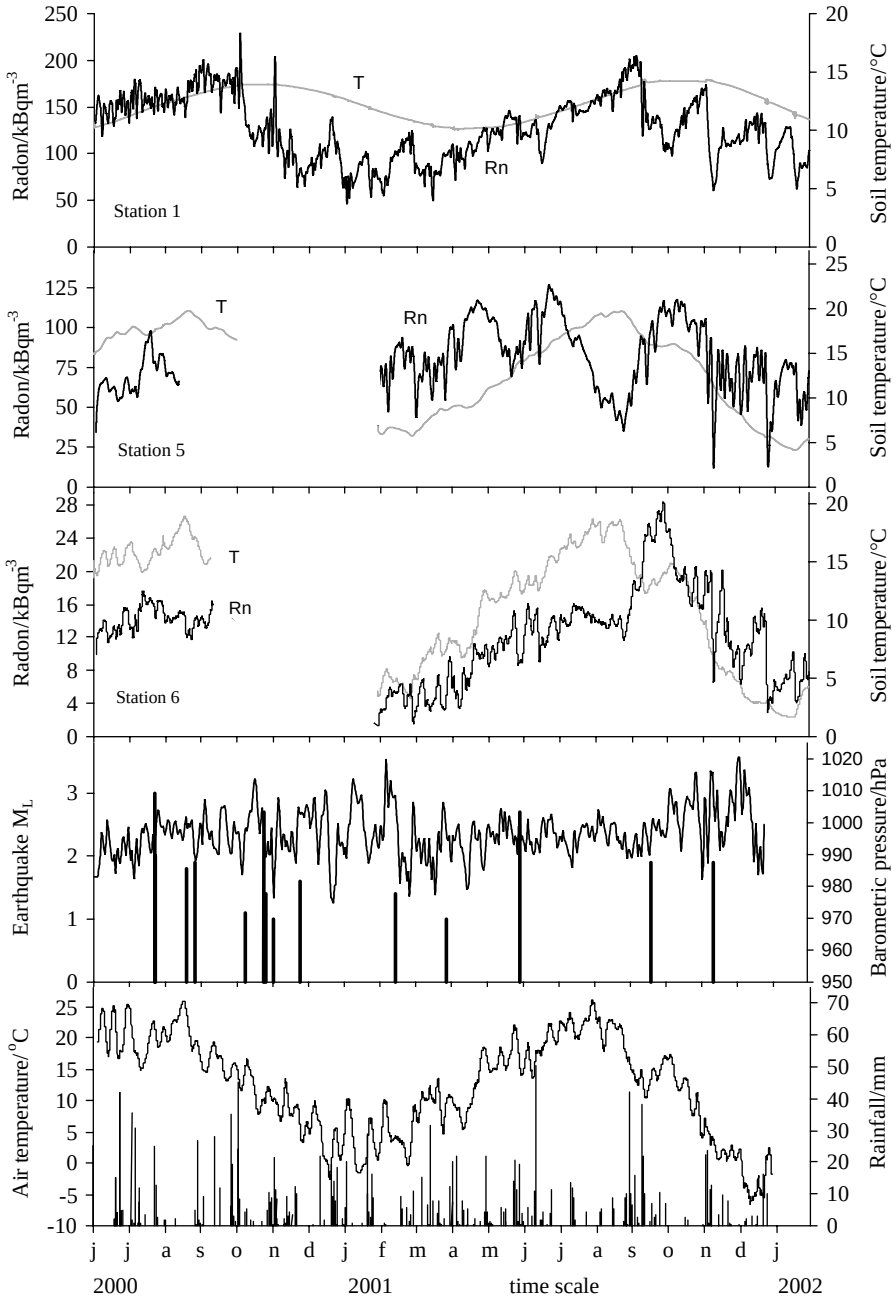

Fig. 1. Radon concentration, 24 hourly averages, recorded at stations 1, 5 and 6 for the period from June 2000 to January 2002. Local seismicity, barometric pressure, soil temperature, air temperature and rainfall are shown for the same period.

environmental sciences; for an overview see (Džeroski, 2002). The tasks we address here are to predict radon concentration from meteorological data, then to test the hypothesis that during seismically active periods the prediction is significantly worse than during seismically inactive periods.

2. Experiment

In boreholes at six stations in the Krško basin, radon concentration in soil gas, barometric pressure and temperature have been measured and recorded once an hour since April 1999 (Zmazek et al., 2002a). For that

purpose barasol probes (MC-450, ALGADE, France) have been used. Boreholes 1 and 4 are located in the active fault zone of the Orlica fault, at a distance of about 4000 m from each other, while the other stations are at distances from 150 to 2500 m on either side of the fault zone. Barasol probes at stations 1 and 2 are fixed at depths of about 5 m, while in other boreholes they are 60–90 cm deep. The borehole wall is protected with a plastic tube and the top is isolated and covered with a plastic cap and soil to reduce hydrometeorological effects on the measurement. For this paper, the data collected from station 1, from the period April 1999–February 2002, and from stations 5 and 6, from the period June 2000–February 2002, were used.

Other meteorological data have been provided by the Office of Meteorology of the Environmental Agency of the Republic of Slovenia, and seismic data by the Office of Seismology of the same Agency. Only earthquakes potentially responsible for strain effects in the investigated area (Dobrovolsky et al., 1979) have been considered. During the period of our measurements at station 1, 21 earthquakes occurred with local magnitude of 0.8–3.3 and at an epicentre distance of 1–31 km, but in the shorter period of measurements at stations 5 and 6, only six earthquakes occurred, with local magnitude of 1.9–3.0 and at an epicentre distance of 1–29 km.

3. Methodology of data analysis

Since radon concentration is a numeric variable, we have approached the task of predicting radon concentration from meteorological data using regression (or function approximation) methods. We used regression trees (Breiman et al., 1984), as implemented with the WEKA data mining suite (Witten and Frank, 1999). For comparison, we also took into consideration two other regression methods, a more traditional statistical method of linear regression (LR) and instance based regression (IB) (Aha and Kibler, 1991), also implemented within the WEKA data mining suite.

Regression trees are a representation for piecewise constants or piecewise linear functions. Like classical regression equations, they predict the value of a dependent variable (called class) from the values of a set of independent variables (called attributes). Data presented in the form of a table can be used to learn or automatically construct a regression tree. In that table, each row (example) has the form (x_1, x_2,…, x_N, y), where x_i are values of the N attributes (e.g., air temperature, barometric pressure, etc.) and y is the value of the class (e.g., radon concentration in soil gas). Unlike classical regression approaches, which find a single equation for a given set of data, regression trees partition the space of examples into axis-parallel rectangles and fit a model to each of these partitions. A regression tree has a test in each inner node that tests the value of a certain attribute and, in each leaf a model for predicting the class. The model can be a linear equation or just a constant. Trees having l inear equations in the leaves are also called model trees (MT).

Given a new example for which the value of the class should be predicted, the tree is interpreted from the root. In each inner node, the prescribed test is performed and, according to the result of the test, the corresponding left or right sub-tree is selected. When the selected node is a leaf then the value of the class for the new example is predicted according to the model in the leaf.

Tree construction proceeds recursively, starting with the entire set of training examples (entire table). At each step, the most discriminating attribute is selected as the root of the subtree and the current training set is split into subsets according to the values of the selected attribute. Technically speaking, the most discriminating discrete attribute or continuous attribute test is the one that most reduces the variance of the values of the class variable. For discrete attributes, a branch of the tree is typically created for each possible value of the attribute. For continuous attributes, a threshold is selected and two branches are created, based on that threshold. The attributes that appear in the training set are considered as thresholds. For the subsets of training examples in each branch, the tree construction algorithm is called recursively. Tree construction stops when the variance of the class values of all examples in a node is small enough (or if some other stopping criterion is satisfied). These nodes are called leaves and are labelled with a model (constant or linear equation) for predicting the class value.

An important mechanism used to prevent trees from over-fitting data is tree pruning. Pruning can be employed during tree construction (pre-pruning) or after the tree has been constructed (post-pruning). Typically, a minimum number of examples in branches can be prescribed for pre-pruning and a confidence level in error estimates expressed in leaves for post-pruning.

A number of systems exist for inducing regression trees from examples, such as CART (Breiman et al., 1984) and M5 (Quinlan, 1992). M5 is one of the best-known programs for regression tree induction. We used the system M5 (Wang and Witten, 1997), a reimplementation of M5 within the WEKA data mining suite (Witten and Frank, 1999). The parameters of M5 were set to their default values, unless stated otherwise.

Both types of regression trees mentioned above were used. Ordinary regression trees predict a constant value in each leaf node. Model trees use a linear regression for prediction in each leaf node. Furthermore, we used different linear regression methods that differ in the way they select and/or eliminate predictive variables. Finally, a single instance based regression method was used with

Table 1
Predictive performance, in terms of correlation coefficient, of different regression methods based on radon concentration in soil gas at three stations in the Krško basis

Method	Station 1	Station 5	Station 6
		r	
MT	0.80	0.81	0.76
RT	0.78	0.68	0.73
LR 1	0.72	0.40	0.66
LR 2	0.72	0.38	0.66
LR 3	0.72	0.40	0.66
LR 4	0.72	0.38	0.66
IB 1	0.53	0.34	0.55
IB 5	0.58	0.52	0.60
IB 10	0.57	0.50	0.56
IB 25	0.58	0.43	0.59
IB 50	0.61	0.44	0.62
IB 99	0.66	0.46	0.64

MT—model trees; RT—regression trees; LR—linear regression four different methods; IB—instance based regression, six different settings for the number of neighbours parameter.

Table 2
Predictive performances, in terms of root mean squared error of different regression methods, for different locations

Method	Station 1	Station 5	Station 6
	RMSE		
MT	20 063	12 769	3508
RT	21 100	16 208	3651
LR 1	22 960	20 048	4019
LR 2	23 008	20 195	4019
LR 3	22 960	20 048	4019
LR 4	23 008	20 195	4019
IB 1	38 100	22 332	4903
IB 5	30 693	18 732	4342
IB 10	29 832	19 076	4489
IB 25	29 298	19 885	4441
IB 50	28 473	20 157	4453
IB 99	27 427	20 830	4532

MT—model trees; RT—regression trees; LR—four different linear regression methods; IB—instance based regression—six different settings for the number of neighbours parameter.

six different settings for the number of nearest neighbouring instances parameter: 1, 5, 10, 25, 50 and 99.

The predictive performance of the regression methods was determined using two different measures. First, the correlation coefficient (r) expresses the level of correlation between the measured and predicted values of radon concentration. Higher values of the correlation coefficient signify a better correlation. Second, the root mean squared error (RMSE) measures the discrepancy between measured and predicted values of radon concentration. Smaller RMSE values indicate lower discrepancies.

In order to estimate the performance of predictors on measurements that were not used for training the predictor, a standard 10-fold cross validation method was applied.

The estimated predictive performance of the different regression methods for the three stations is shown in Table 1 (correlation coefficient) and Table 2 (RMSE). The results show that the model trees outperform all the other regression methods considered, both on average and for each individual station. They perform better in terms of the correlation coefficient and of the root mean squared error. Therefore, for the further evaluations performed to test the hypothesis, we have used model trees to predict the radon concentration.

4. Results and discussion

Experimental data for the period from June 2000 to February 2002 for the three selected stations are shown in Fig. 1. To test the hypothesis about the predictability of radon concentration in periods with and without seismic activities, the following procedure was applied. First the value of the class—daily radon concentration; and the values of attributes—average daily barometric pressure, average daily air temperature, average daily soil temperature, difference between daily soil and daily air temperature, daily amount of rainfall, and difference in daily barometric pressure was selected. Second, this data set was split into two parts. In the first part (labelled SA), data for the periods with seismic activity were included, i.e., periods of 7 days before and after an earthquake. Data for the remaining days were included in the second part, belonging to the seismically non-active periods (labelled non-SA). Then, to evaluate the predictability of radon concentration in the seismically non-active periods, we estimated the performance of model trees on the non-SA data with cross-validation. Furthermore, we induced a model tree on the whole non-SA data and measure its performance on the SA data in order to evaluate the practicability of predicting radon concentration in the SA periods. If our hypothesis is true, the first measured performance should be higher than the second.

The period of 7 days was chosen after inspecting correlation changes between radon concentration and barometric pressure (Zmazek et al., 2002a), and confirmed by an analysis in which the length of the SA periods was varied from 1 to 7 days. Table 3 summarises the correlation coefficients and RMSEs obtained with the MT method for different lengths of the SA period (assuming that radon changes appear 1–7 days before an earthquake). The largest drop in the correlation coefficient was observed between 6 and 7 days before an earthquake.

Table 3
Comparison of predictability of radon concentration based on the assumptions that radon changes appears 1–7 days before earthquake

	Days	Non-SA periods		SA periods		Performance change (%)	
		r	RMSE	r	RMSE	r	RMSE
Station 1	7	0.83	18 719	0.69	23 536	−17.3	25.7
	6	0.84	18 135	0.73	22 551	−13.5	24.4
	5	0.82	18 913	0.76	21 659	−6.9	14.5
	4	0.82	19 052	0.76	21 291	−6.7	11.8
	3	0.80	19 792	0.79	19 975	−1.6	0.9
	2	0.80	19 972	0.78	19 534	−2.5	−2.2
	1	0.80	19 920	0.78	19 197	−2.4	−3.6
Station 5	7	0.81	13 243	0.54	14 910	−33.7	12.6
	6	0.81	13 342	0.53	13 223	−34.4	−0.9
	5	0.81	13 288	0.66	11 208	−18.4	−15.7
	4	0.82	12 720	0.79	8457	−4.1	−33.5
	3	0.81	13 183	0.71	10 295	−11.9	−21.9
	2	0.81	12 990	0.74	12 272	−8.8	−5.5
	1	0.82	12 667	0.73	12 225	−10.8	−3.5
Station 6	7	0.80	3076	0.22	5299	−72.9	72.3
	6	0.79	3185	0.14	5404	−82.6	69.7
	5	0.80	3142	0.18	5321	−77.4	69.4
	4	0.79	3211	0.36	4982	−54.1	55.1
	3	0.77	3415	0.37	4836	−51.3	41.6
	2	0.77	3397	0.25	4884	−67.5	43.8
	1	0.76	3451	0.35	4593	−53.9	33.1

The data for 7 days in Table 3 (bold) clearly confirm the hypothesis: the correlation in the SA periods is much lower than in the non-SA periods, for all three stations. The drop in the correlation coefficient ranged from 0.17 to 0.73. The RMSE in the SA periods is higher than in the non-SA periods. The RMSE increase ranged from 13% to 72%. The confirmation of the hypothesis allows us to predict seismic activity in the following manner. A model tree is built that predicts the concentration of radon in soil gas on the basis of data measured during the non-SA periods. We then follow the discrepancy between the measured values of radon concentration and the values predicted by the model tree. If the discrepancy is low, no seismic activity is anticipated; if it starts to increase, however, an increase in seismic activity may be expected.

From our results, station 6 appears to be the best location for earthquake prediction. However, only six earthquakes occurred in the vicinity of this location during the period of study (June 2000–February 2002). Before one of these earthquakes, a very large difference between the measured and the predicted radon concentrations was observed (Fig. 4). During our measurements at location 1 (April 1999–February 2002), 21 earthquakes occurred, but anomalies occurring before some of the earthquakes at larger distances from the station are not very marked. From a geological point of view, station 6 is the only station on Triassic limestone; the other two are on Miocene limestone. On the other hand, at station 1, the hole is much deeper (5 m) than the holes at stations 5 and 6 (60–90 cm). This could be the reason that certain environmental parameters, such as barometric pressure, have more influence on radon concentration. For this reason, during some seismic events much smaller anomalies in measured radon concentration have been observed.

The main contribution of our work is that we have applied a new method (decision trees) to relate radon data to seismic activity. The results are encouraging. The learned decision trees predict anomalies before and/or during all earthquakes. The radon concentration predicted before an earthquake is mostly lower than the measured radon concentration (the predicted value is higher than the measured on only five occasions). The drop in the correlation coefficient between the two appears on average 10.2 ± 4.8 days before an earthquake. The duration of an anomaly is 13 ± 7 days. The time of appearance of anomalies and their duration and magnitude grow with the magnitude of the earthquake and become smaller as the distance to the earthquake increases. The average difference between the measured and the predicted radon concentrations during these

anomalies is 54±35% of the measured values. On 14 occasions, differences between the measured and the predicted radon concentrations can be observed in the non-SA periods, but these differences of radon concentration, on average 30%, last for less than 3 days on average. The predicted radon concentrations are compared with the measured values for the three measuring stations in Figs. 2–4.

Fig. 2 shows three measurement periods for station 1, with six earthquakes. In Fig. 2(a), two days before the first earthquake, the measured radon concentration first increases, then it starts to decrease a day before the earthquake and returns to the predicted level 3 days after the earthquake. The situation is reversed for the second earthquake on May 20, 1999. Here, 6 days before the earthquake, the measured radon concentration first decreases, and then 2 days before the earthquake it starts to increase, to eventually match the predicted level. In Fig. 2(b), the measured radon concentration increases before the earthquake. Then, at the time of the earthquake, it first suddenly decreases and then abruptly increases to match the predicted value 8 days after the

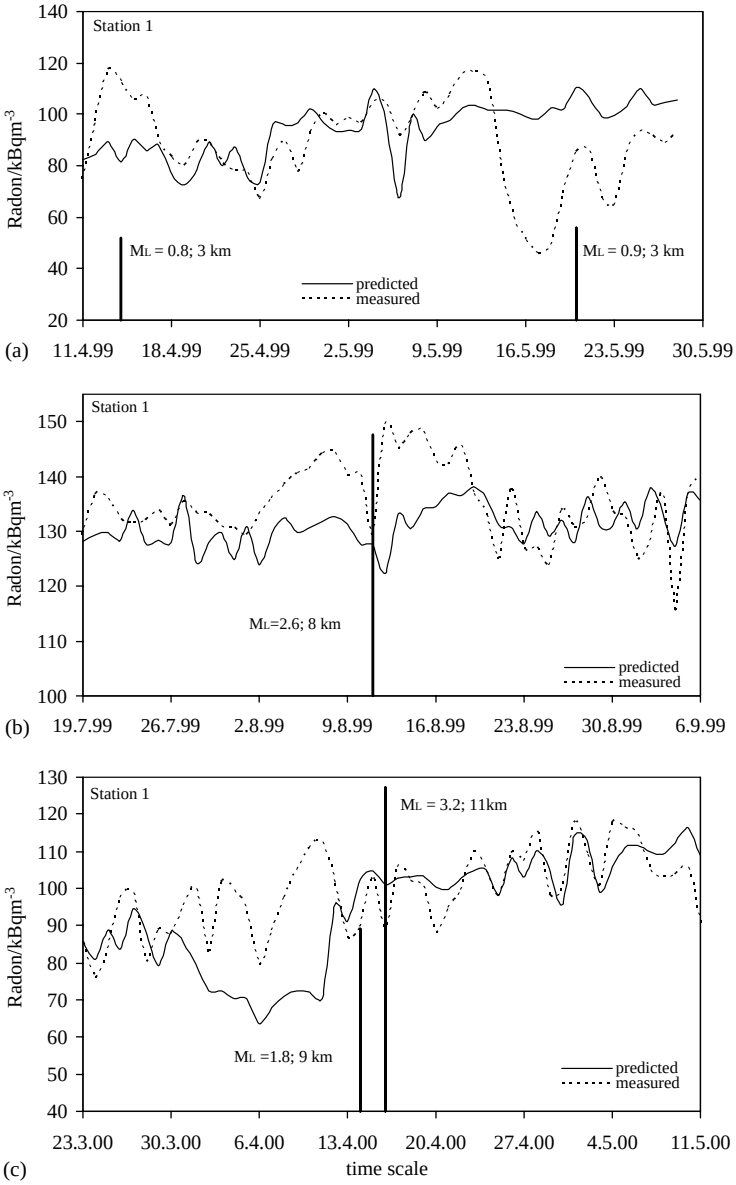

Fig. 2. Predicted and measured radon concentration at station 1 for different periods: (a) from April 11 to May 30, 1999; (b) from July 19 to September 6, 1999; (c) from March 23 to May 11, 2000. Earthquakes with their magnitudes and distances between the station and the epicentres are also shown.

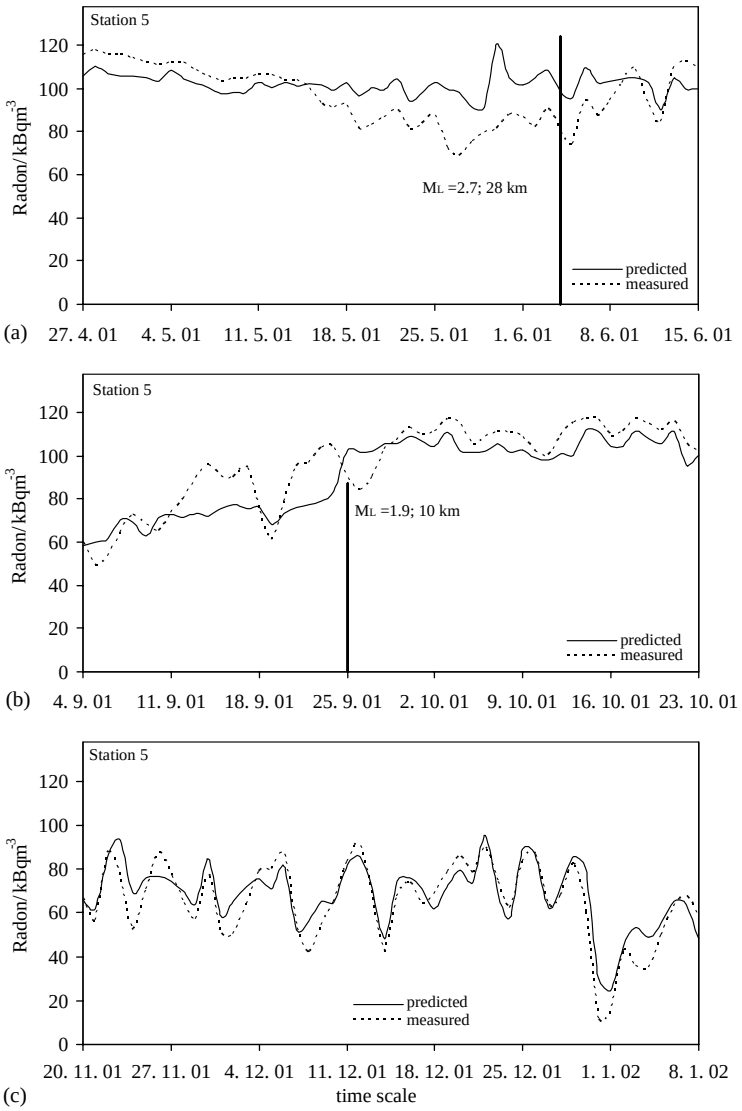

Fig. 3. Predicted and measured radon concentration at Station 5 for different periods: (a) from April 27 to June 15, 2001; (b) from September 4 to October 23, 1999; (c) from November 20, 2001 to January 8, 2002. Earthquakes with their magnitudes and distances between the station and the epicentres are also shown.

earthquake. Two earthquakes in Fig. 2(c) are also preceded by increases in the measured radon concentration. But here, the measured and predicted concentrations start to match before the seismic events.

In Fig. 3, the situation for station 5 is shown. Eighteen days before the earthquake in Fig. 3(a), the measured radon concentration starts to decrease and is below the predicted value until 3 days after the earthquake. In Fig. 3(b), the measured radon concentration before the earthquake is higher than the predicted one, with a decrease 6 days before the earthquake. It is obvious from Fig. 3(c) that the prediction of radon concentration in the non-SA period is good.

Observations at station 6 are presented in Fig. 4. While seismic activity in Figs. 4(a) and (c) is accompanied by an increase in the measured radon concentration, the situation is reversed in Fig. 4(b).

In the above examples, a seismic event was always preceded by a time during which the measured radon concentration was not predicted using regression trees. We were able to detect anomalies related to earthquakes as weak as $M_L = 0.8$. A more extended analysis would be needed in order to explain why at the same measuring station, radon concentration may either increase or decrease before or during an earthquake, or to relate the nature of an anomaly to the local magnitude or

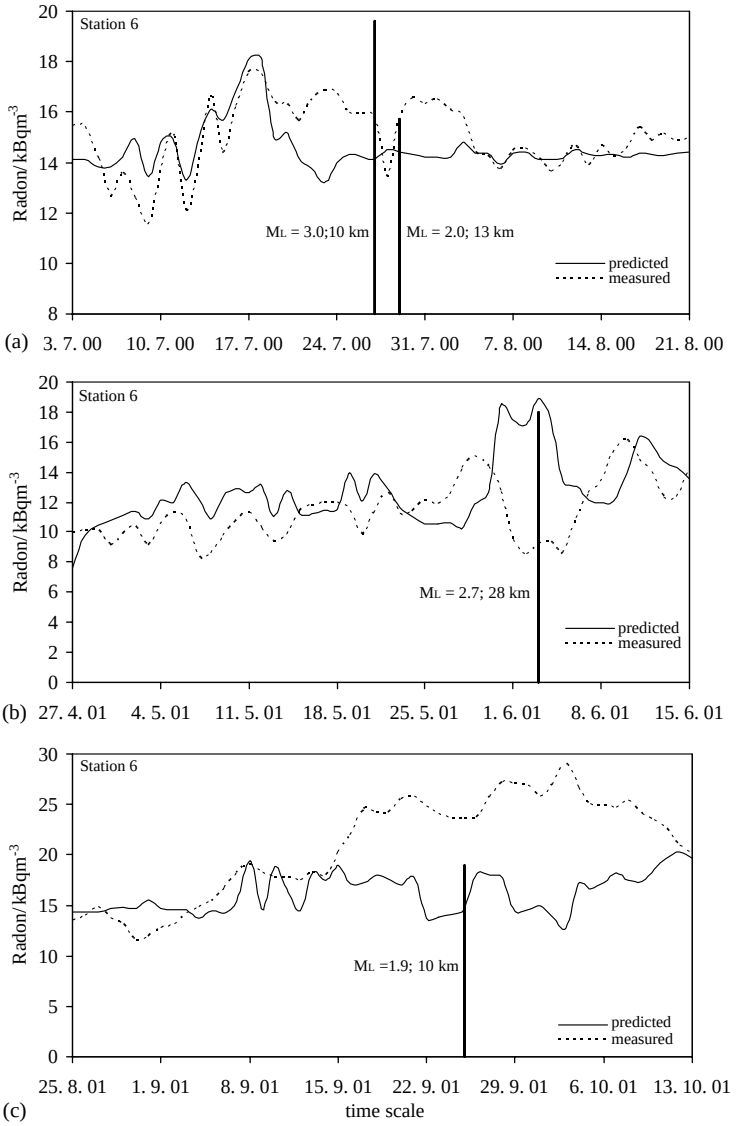

Fig. 4. Predicted and measured radon concentration at station 6 for different periods: (a) from July 3 to August 21, 2000; (b) from April 27 to June 15, 2001; (c) from August 25 to October 13, 2001. Earthquakes with their magnitudes and distances between the station and the epicentres are also shown.

epicentre of an earthquake. This analysis is beyond the scope of this paper.

5. Conclusion

We use the machine learning method of model tree (MT) induction to predict radon concentration in soil gas from measured environmental data, i.e. barometric pressure, air temperature, soil temperature, rainfall, difference between air and soil temperature, daily changes of barometric pressure and rainfall. MTs represent piecewise linear functions. From measured data, MT induction learns piecewise linear functions that predict radon concentration from environmental data. We have compared MTs to two other regression methods: LR and IB regression. The MT method performs best in terms of predictive accuracy.

In the non-SA periods, when the variation of radon concentration with time is affected merely by the environmental parameters and not by seismic activity, radon concentrations have been predicted with correlations over 0.8. If the prediction is significantly worse, an increased seismic activity is expected. This method does not require us to know how the radon concentration is related to environmental data; it discovers this

relationship from measured data. It can also determine the parameters that have the strongest effect on radon concentration. If we measured additional relevant environmental parameters, such as wind and humidity of the soil, we would expect better results for predicting radon concentration in the non-SA periods. Finally, with this method, we can detect radon anomalies resulting from earthquakes with local magnitude lower than 3.3.

Our measurements are still in progress and further analyses will be carried out over longer measuring periods.

References

Aha, D., Kibler, D., 1991. Instance based learning algorithms. Mach. Learning 6, 37–66.

Belayaev, A.A., 2001. Specific Features of radon earthquake precursors. Geochem. Int. 12, 1245–1250.

Biagi, P.F., Ermini, A., Kingsley, S.P., Khatkevich, Y.M., Gordeev, E.I., 2001. Difficulties with interpreting changes in groundwater gas content as earthquake precursors in Kamchatka. Russ. J. Seismol. 5, 487–497.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and Regression Trees. Wadsworth, Belmont.

Cuomo, V., Di Bello, G., Lapenna, V., Piscitelli, S., Telesca, L., Macchiato, M., Serio, C., 2000. Robust statistical methods to discriminate extreme events in geoelectrical precursory signals: implications with earthquake prediction. Nat. Hazard 21, 247–261.

Di Bello, G., Ragosta, M., Heinicke, J., Koch, U., Lapenna, V., Piscitelli, S., Macchiato, M., Martinelli, G., 1998. Time dynamics of background noise in geoelectrical and geochemical signals: an application in a seismic area of Southern Italy. Ill Nuovo Cimento 6, 609–629.

Dobrovolsky, I.P., Zubkov, S.I., Miachkin, V.I., 1979. Estimation of the size of earthquake preparation zones. Pure Appl. Geophys. 117, 1025–1044.

Džeroski, S., 2002. Applications of KDD methods in environmental sciences. In: Kloesgen, W., Zytkow, J. (Eds.), Handbook of Data Mining and Knowledge Discovery. Oxford University Press, Oxford.

Igarashi, G., Saeki, S., Takahata, N., Sumikawa, K., Tasaka, S., Sasaki, Y., Takahashi, M., Sano, Y., 1995. Groundwater radon anomaly before the Kobe earthquake in Japan. Science 269, 60–61.

Mjachkin, V.I., Brace, W.E., Sobolev, G.A., Dieterich, J.H., 1975. Two models for earthquake forerunners. Pure Appl. Geophys. 113, 169–181.

Negarestani, A., Setayeshi, S., Ghannadi-Maragheh, M., Akashe, B., 2001. Layered neural networks based analysis of radon concentration and environmental parameters in earthquake prediction. J. Environ. Radioact. 62, 225–233.

Planinić, J., Radolić, V., Čulo, D., 2000. Searching for an earthquake precursor: temporal variations of radon in soil and water. Fizika B (Zagreb) 9, 75–82.

Planinić, J., Radolić, V., Lazanin, Ž., 2001. Temporal variations of radon in soil related to earthquakes. Appl. Radiat. Isot. 55, 267–272.

Quinlan, J.R., 1992. Learning with continuous classes. In: Proceedings of the Fifth Australian Joint Conference on Artificial Intelligence. World Scientific, Singapore, pp. 343–348.

Scholz, C.H., Sykes, L.R., Agrawal, Y.P., 1973. Earthquake prediction: a physical basis. Science 181, 803–810.

Singh, M., Virk, H.S., 1994. Investigation of radon-222 in soil–gas as an earthquake precursor. Nucl. Geophys. 8, 185–193.

Singh, M., Ramola, R.C., Singh, B., Singh, S., Virk, H.S., 1993. Radon anomalies: correlation with seismic activities in Northern India. In: Proceedings of the Second Workshop on Radon Monitoring in Radioprotection. Environmental and/or Earth Sciences, Trieste, 25 November–6 December, 1991. World Scientific, Singapore, pp. 359–377.

Singh, M., Kumar, M., Jain, R.K., Chatrath, R.P., 1999. Radon in ground water related to seismic events. Radiat. Meas. 30, 465–469.

Sultankhodajev, A.N., 1984. Hydrogeoseismic precursors to earthquakes. In: Earthquake Prediction, UNESCO, Paris, pp. 181–191.

Teng, T.L., 1980. Some recent studies on groundwater radon content as an earthquake precursor. J. Geophys. Res. 85, 3089–3099.

Ui, H., Moriuchi, H., Takemura, Y., Tsuchida, H., Fuji, I., Nakamura, M., 1988. Anomalously high radon discharge from the Atotsugawa fault prior to the Western Nagano Prefecture earthquake (M 6.8) of September 14, 1984. Tectonophysics 152, 147–152.

Ulomov, V.I., Mavashev, B.Z., 1971. Forerunners of the Tashkent earthquake. Izv. Akad. Nauk Uzb. SSR 188–200.

Virk, H.S., Walia, V., Kumar, N., 2001. Helium/radon precursory anomalies of Chamoli earthquake, Garhwal Himalaya, India. J. Geodyn. 31, 201–210.

Wakita, H., 1979. Earthquake prediction by geochemical techniques. Rec. Progr. Nat. Sci. Jpn. 4, 67–75.

Wakita, H., Nakamura, Y., Sano, Y., 1988. Short-term and intermediate-term geochemical precursors. Pure Appl. Geophys. 126, 267–278.

Wang, Y., Witten, I.H., 1997. Induction of model trees for predicting continuous lasses. In Proceedings of the Poster Papers of the European Conference on Machine Learning, University of Economics, Faculty of Informatics and Statistics, Prague.

Witten, I.H., Frank, E., 1999. Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations. Morgan Kaufmann, San Francisco.

Yasuoka, Y., Shinogi, M., 1997. Anomaly in atmospheric radon concentration: a possible precursor of the 1995 Kobe, Japan, earthquake. Health Phys. 72, 759–761.

Zmazek, B., Vaupotič, J., Živčić, M., Premru, U., Kobal, I., 2000a. Radon monitoring for earthquake prediction in Slovenia. Fizika B (Zagreb) 9, 111–118.

Zmazek, B., Vaupotič, J., Živčić, M., Martinelli, G., Italiano, F., Kobal, I., 2000b. Radon, temperature, electrical conductivity and $^{3}He/^{4}He$ measurements in three thermal springs in Slovenia. In: Book of Abstracts: New Aspects of Radiation Measurements, Dosimetry and Spectrometry. The Proceedings of the Second Dresden Symposium on Radiation Protection, September 10–14, 2000.

Zmazek, B., Vaupotič, J., Bidovec, M., Poljak, M., Živčić, M., Pineau, J.F., Kobal, I., 2000c. Radon monitoring in soil gas

tectonic faults in the Krško basin. In: Book of Abstracts: New Aspects of Radiation Measurements, Dosimetry and Spectrometry. The Proceedings of the Second Dresden Symposium on Radiation Protection, September 10–14, 2000.

Zmazek, B., Vaupotič, J., Kobal, I., 2002a. Radon, temperature and electric conductivity in Slovenian thermal waters as potential earthquake precursors. The Proceedings of the First Workshop on Natural Radionuclides in Hydrology and Hydrogeology, Centre Universitaire de Luxembourg, September 4–7.

Zmazek, B., Italiano, F., Živčić, M., Vaupotič, J., Kobal, I., Martinelli, G., 2002b. Geochemical monitoring of thermal waters in Slovenia: relationships to seismic activity. Appl. Radiat. Isot. 57, 919–930.

Zmazek, B., Živčić, M., Vaupoti, J., Bidovec, M., Poljak, M., Kobal, I., 2002c. Soil radon monitoring in the Krško basin, Slovenia. Appl. Radiat. Isot. 56, 649–657.