only published deep boreholes from Ljubljansko barje with more or less complete data sets. There are many more (Mencej, 1990) but data are sparse and no details are available. With many remaining open problems, the publication of other available data and even the drilling of new boreholes is urgently needed.

In 1959, the BV-1 borehole was drilled down to the dolomite basement to a depth of 103.80 m in the area between Notranje gorice and Podpeč, and, in 1962, the BV-2 borehole was drilled south of Črna vas (Fig. 1) to the dolomite basement at 116.80m (Grimšičar & Ocepek, 1967; Pohar, 1978). In the present paper, data of carbonate concentrations (represented as mass ratio of CO_2 or $CaCO_3$ in the sample) from boreholes BV-1 (Grimšičar & Ocepek, 1967) and BV-2 (Pohar, 1978) are analysed based on statistical time series techniques. The article consists of two parts that are equally important. In the first part, an unevenly spaced time series was studied by an interpolation technique based on a weighted influence function. Data of CO_2 or $CaCO_3$ concentrations along the borehole core (borehole depth) in BV-1 and BV-2 can be described as an unevenly spaced time series where time (length) distances between particular data are not equal. These differences have some empirical distributions. For this type of data, several statistical techniques have been developed (e.g. Diggle, 1990), however our calculations were defined by independent consideration and we have developed our own statistical model. Therefore, the method also represents a new contribution to the data reduction of unevenly spaced time series. Based on this model, we have reconstructed an equally spaced data time series which was the basis for the second part of the paper. In the second part, the structure of the time series was explored with an autocorrelation function in the time domain and with a power spectrum in the frequency domain. The periodicity of the data was extracted and the types of noise contained in the signal were analysed. An attempt was made to interpret the results from a sedimentation point of view.

Methods

Data reconstruction

Data for the analyses were obtained by digitization of originally published borehole profiles given in the works of Grimšičar & Ocepek (1967) and Pohar (1978). The diagram showing the CO_2 profile in borehole BV-1 was digitized by scanning all the data points on the borehole diagram (Grimšičar & Ocepek, 1967). In the diagram of the $CaCO_3$ profile in borehole BV-2 (Pohar, 1978), data points were not given. In this case digitization was performed along a line and data points chosen where a discontinuity was observed in the line. These values are probably not exactly the same values as the data obtained in the laboratory by the authors of the original paper. Unfortunately, data especially for borehole BV-2 are not published and are not available to the author. However, because we are not particularly interested in the particular values of the carbonate concentration, but more in the general shape of the curve along the borehole, we believe that data obtained with digitization of published diagrams are suitable for our analyses.

From the published data, it follows that values of the CO_2 ratio in borehole BV-1 (Grimšičar & Ocepek, 1967) and of $CaCO_3$ ratio in BV-2 (Pohar, 1978) are not directly comparable. Unfortunately information about analytical methods for CO_2 and

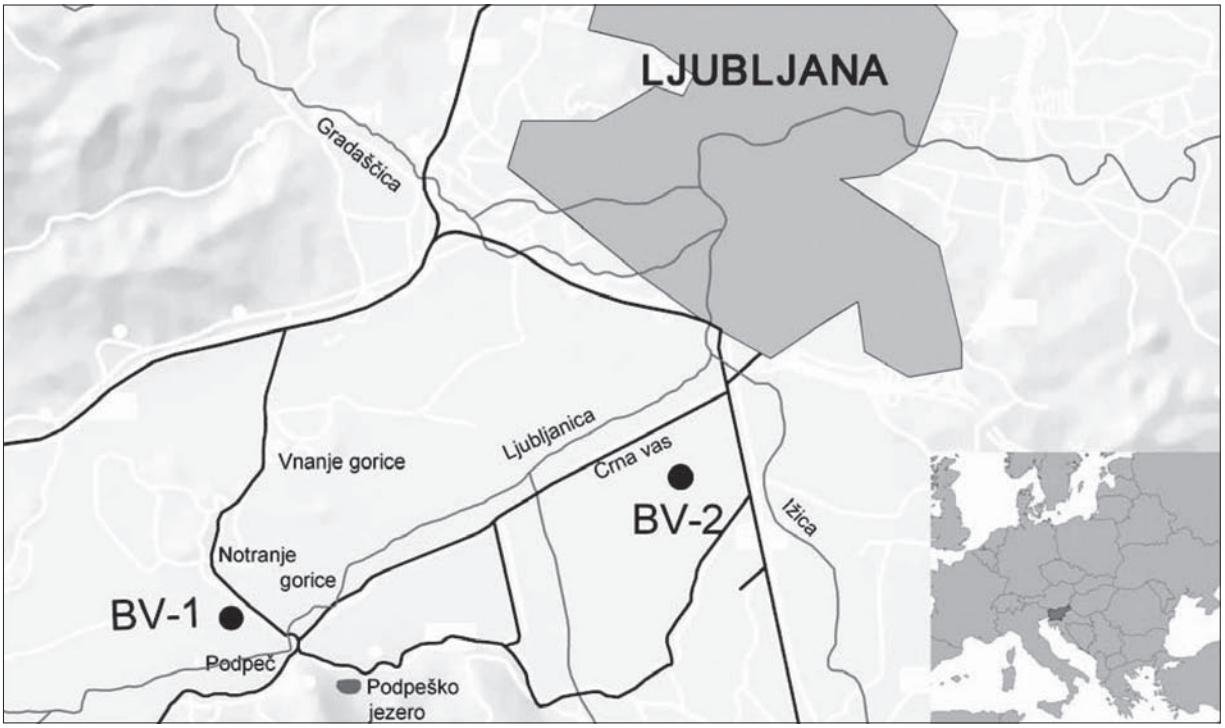

Figure 1. Position of boreholes BV-1and BV-2 at Ljubljansko barje (summarized and adapted according to Grimšičar & Ocepek, 1967)

$CaCO_3$ determination are very sparse. Grimšičar & Ocepek (1967) report a calcimetric analysis with HCl acid in the ratio of 1:2.5 where they probably measured the mass ratio between sample and diluting agent. Readings were controlled for pure calcite and the precision of the reading was 0.1%. No information about the grain size of the sampled sediment is available. Pohar (1978) report that samples were weighed and sieved and grains smaller than 1 mm were used for the $CaCO_3$ determination. How the determination of $CaCO_3$ was performed is not reported.

Because both boreholes BV-1 and BV-2 were drilled within a short period of time and, in both cases, experts from the same institutions were involved and were cooperating on the project, we can suppose that determination of CO_2 and $CaCO_3$ were performed by calcimetry on the same or similar fraction of the sediment sample. Therefore we have to recalculate concentrations of $CaCO_3$ in BV-2 based on molar masses derived from the ratio of CO_2 concentrations. By such a transformation, we obtain data on comparable scales.

Time series analyses

Suppose that $\boldsymbol{X} = \{x_1, x_2, ..., x_i, ..., x_n\}$ is a random vector where values $x_1, x_2, ..., x_i, ..., x_n$ are defined at time coordinates $t = 1,2, ..., n$. If $\Delta t = t_{i-1} - t_i = t_i - t_{i+1}$, the time series is evenly spaced. If $\Delta t = t_{i-1} - t_i \neq t_i - t_{i+1}$ then the time series is uneven and data are not equally spaced. Irregular time series data are common when equally spaced data cannot be obtained owing to limitations of data access, or more often in cases when natural conditions do not allow equally spaced sampling. Such cases are very frequent in stratigraphy. The reader interested in regular time series analyses can find more information elsewhere (e.g Diggle, 1990; Weedon, 2003).

For the regular time series, the τ-th autocovariance coefficient is defined as:

$$g_k = \sum_{i=k+1}^{n} \left(x_i - \bar{x} \right) \left(x_{i-\tau} - \bar{x} \right) / n$$

where the mean value of random vector $\boldsymbol{X}$ is defined as:

$$\bar{x} = \sum_{i}^{n} x_i / n$$

Index τ represents a time lag. The definition of the empirical τ-th autocorrelation coefficient r_τ follows from this result. It represents an estimate of the function $\rho(t)$ of the stochastic process $\boldsymbol{X}$ that gave rise to sample defined as $\boldsymbol{X}$. The τ-th autocorrelation coefficient r_τ is defined as

$$r_\tau = g_\tau / g_0$$

where g_0 is the autocovariance at time lag $\tau = 0$. A diagram of r_τ versus τ is defined as an autocorrelation diagram that defines the structure of the time series in the time domain. Time series analysis can also be performed in the frequency domain, where components of periodic trends inside the time series can be detected by harmonic analysis. We assume that the time series y_t can be modelled as:

$$y_t = \sum_{i=1}^{n/2} \left\{ \alpha_i \cos(\omega_i t) + \beta_i \sin(\omega_i t) \right\} + U(t)$$

where:

- α_i, β_i are parameters to be estimated from the data,
- ω is the frequency of cyclic fluctuations defined as $\omega = 2\pi/p$ where p is the period,
- $U(t)$ is a random component.

We can define:

$$A_i = \sum_{i=1}^{n} y_t \cos(\omega_i t)$$

$$B_i = \sum_{i=1}^{n} y_t \sin(\omega_i t) \qquad \text{and}$$

$$C_i = \left[A_i^2 + B_i^2 \right]$$

where C_i denotes periodogram ordinates. A plot of C_i as ordinate against frequency ω is called the power spectrum of y_t.

Interpolation procedure

From the definitions given above for g_k, it is not possible to calculate the autocorrelation diagram in an unevenly spaced time series. It can be calculated from the variogram (Diggle, 1990) or from a regular time series reconstructed by interpolation techniques from the original unevenly spaced time series. Interpolation of an irregular time series to a regular time series is feasible only if the density of observation points is large enough. From the Parseval theorem, it follows that the smallest frequency and information from it depends on the distance Δt between observations.

Suppose that the random vector $\boldsymbol{X_e}$ is a sample of a stochastic process so that $\Delta t = t_{i-1} - t_i = t_i - t_{i+1}$ is valid. The $\boldsymbol{X_e}$ represents an evenly spaced time series. Suppose also that random vector $\boldsymbol{X_n}$ is a sample of a stochastic process so that $\Delta t = t_{i-1} - t_i \neq t_i - t_{i+1}$. The $\boldsymbol{X_n}$ represents an unevenly spaced time series. Random vectors $\boldsymbol{X_e}$ and $\boldsymbol{X_n}$ correspond to each other if the relation $\boldsymbol{X_e} \mid \Delta t = \boldsymbol{X_n} \mid \Delta t_{min}$ is valid. Interpolation is defined as a procedure where random $\boldsymbol{X_n}$ is transformed into correspondence vector $\boldsymbol{X_e}$ so that $\boldsymbol{X_n} = \{x_1, x_2, ..., x_n\} \rightarrow \boldsymbol{X_e} = \{x_1, {}_1x, ..., {}_nx, x_2, ..., x_n\}$ where ${}_1x, ..., {}_nx$ represent interpolated values for which $\Delta t = t_1 - {}_1t = {}_1t - {}_{i+1}t = {}_nt - t_2 = \Delta t_{min}$ is valid.

From the sedimentation processes, it follows that the properties of adjacent sediments or rocks in the vertical stratigraphical column and therefore in the time series are correlated. In sedimentology, we are very much interested in the value of this correlation and whether some periodicity exists in the data.

In the reconstruction of an evenly spaced time series, an *a priori* dependence model between the data must be proposed. A dependence model based on a weighted influence function W_t defined as

$$W_t = \Delta t_i^{-a}$$

is proposed, where Δt_i denotes the distance between successive data points in an unevenly spaced time series and a is a weighting parameter. If $a>0$, the influence of particular data at t_i reduces with distance. The level of reduction is defined by parameter a. At large a, the influence of x_i diminishes quickly and at small a, it reduces slowly.

Parameter a can be defined based on the expert's judgment. However, it is better to find an independent procedure for selection of the parameter a. There are many possible procedures. For the selection process in our analyses, we have used the maximum entropy principle based on a definition of the entropy E (e.g. Ross, 1989):

$$E = -\sum_i^n p_i \log p_i ,$$

where p_i is the probability of a particular event from stochastic processes $\boldsymbol{X}$ estimated from the frequencies in random vector $\boldsymbol{X_e}$. It was hypothesized that a maximum amount of information is available in the reconstructed time series $\boldsymbol{X_e}$ obtained with parameter a_i when the entropy $E_{\max} | (E_{a1}, E_{a2}, \dots , E_{an})$ is maximum.

Kernel densities

In a classical statistical analysis, the empirical distribution of the data is usually represented by an histogram. Although many approaches for the determination of histogram class width can be found in the literature, they are still biased representations of the real data. Alternative ways for graphically representing the data is the kernel density approach (e.g. Williams, 1997; Reiss & Thomas, 1997). According to this method, the probability density $g_b(x,k(x))$ for particular data is estimated as:

$$g_b(x,k(x)) = \frac{1}{N_b} k\left(\frac{x-x_i}{b}\right),$$

where $k(x)$ is the kernel such that

$$\int k(y)dy = 1$$

and where b is the bandwidth and $b>0$. We have used the Epanechnikov kernel (Reiss & Thomas, 1997):

$$k(x) = \frac{3}{4}(1-x^2) I(-1 \le x \ge 1)$$

In summing up the single terms, one gets the kernel density (Reiss & Thomas, 1997):

$$f_{N,b}(x) = \sum_{i \le N} g_b(x,k(x)) = \frac{1}{Nb} \sum_{i \le N} k\left(\frac{x-x_i}{b}\right)$$

Calculation procedures

After digitization, the data taken from the original publications were analysed by the kernel density approach and basic descriptive statistics were calculated. It was decided that, for interpolation purposes, a minimum distance Δt_{min} would be used for a regular time series reconstruction.

An interpolation was performed based on spreadsheet calculations performed by macros written in Excel. In the first step, all possible differences Δt between unevenly spaced data t_i and data interpolated at the step $\Delta_i t_{min}$ were calculated as $|\Delta_1 t_{min}-t_1|$, $|\Delta_1 t_{min}-t_2|$, … , $|\Delta_1 t_{min}-t_n|$, … , $|\Delta_m t_{min}-t_n|$. Then, each of the $|\Delta_i t_{min}-t_i|^{-a}$ terms was calculated over a chosen interval of $z < a < w$, where z and w are any rational numbers. At each $\Delta_i t_{min}$ and x_i in the original data from the unevenly spaced time series, terms $x_i|\Delta_i t_{min}-t_i|^{-a}$ were calculated. The final values ${}_i x_i$ at $\Delta_i t_{min}$ were obtained as

$${}_i x_i = \frac{\sum_i^n x_i \left|\Delta_i t_{\min} - t_i\right|^{-a}}{\sum_i^n \left|\Delta_i t_{\min} - t_i\right|^{-a}}$$

The ${}_i x_i$ values in fact represent weighted versions of the x_i values in the original data, with the distance between the original data raised to a power.

The maximum entropy was estimated based on a measure p_i of the area between two adjacent x_i, x_{i+1} data points in the particular parameter a reconstructed on the evenly spaced time series.

All autocorrelation functions and power spectra were calculated with the Statistica 6.0 statistical package. Kernel densities were estimated with the program $X_T R_E M_E S$ (Reiss & Thomas, 1997) where automatic bandwidth selection as well as manual bandwidth determination were chosen.

Results

Digitization results of calcimetry data from Grimšičar & Ocepek (1967) are given in Table 1 and results from Pohar (1978) are given in Table 2. Both tables are also represented in the diagrams on Figure 2 where concentrations are expressed in % mass ratio of CO_2. Data obtained form the digitization procedure are well in accordance with the diagrams from the original sources (Grimšičar & Ocepek, 1967; Pohar, 1978). From borehole BV-1, 93 data points were obtained and, from borehole BV-2, 123 data points were obtained. The kernel densities of both data sets are given in Figure 3. The bandwidth was chosen automatically by the program; in the case of BV-1, the bandwidth was 4.87 m and in the case of BV-2 it was 5.11 m. Both densities are very inhomogeneous with several modes. In the data from borehole BV-1, a strong left asymmetry can be observed with a strong mode of around 30 %. Kernel densities for the BV-2 data are more symmetrical, the main mode is around 23 % with two other smaller modes around 13 %

Table 1. Carbonate concentrations from borehole BV-1 expressed as mass ratio of CO_2 in mass % - digitized data from Grimšičar & Ocepek (1967)

Depth [m]	CO_2 [%]	Depth [m]	CO_2 [%]	Depth [m]	CO_2 [%]	Depth [m]	CO_2 [%]	Depth [m]	CO_2 [%]
0.66	8.0	21.96	12.8	41.61	8.8	64.14	11.2	78.43	16.8
1.49	19.2	22.46	15.2	42.44	8.0	64.97	29.6	80.42	5.6
2.31	28.8	23.78	10.4	43.60	9.6	66.30	1.6	81.59	20.8
3.63	28.8	24.61	8.0	44.59	9.6	67.46	34.4	82.75	20.8
4.46	27.2	25.60	4.0	46.07	1.6	68.13	5.6	84.08	0.8
5.61	29.6	26.42	6.4	46.90	3.2	68.29	0.8	84.74	1.6
6.44	24.8	28.07	9.6	47.39	6.4	69.62	28.0	85.24	7.2
7.60	29.6	29.72	6.4	48.22	4.0	71.12	1.6	87.57	15.2
8.42	31.2	30.72	12.0	48.88	8.8	71.28	32.0	88.23	27.2
9.74	28.0	31.71	3.2	50.37	29.6	71.95	28.0	88.73	23.2
10.57	13.6	32.37	9.6	51.85	5.6	72.61	1.6	89.90	27.2
12.39	22.4	33.69	9.6	52.84	14.4	73.44	29.6	90.40	32.0
13.54	3.2	34.51	8.8	53.50	1.6	74.11	1.6	90.89	17.6
14.37	3.2	35.67	4.8	56.66	1.6	74.61	0.8	92.39	11.2
15.19	1.6	36.33	10.4	57.16	1.6	75.44	4.8	93.05	19.2
16.35	2.4	36.83	12.8	58.65	11.2	76.10	0.8	95.21	38.4
17.17	3.2	37.65	9.6	60.48	2.4	76.44	24.0	99.20	38.4
18.50	4.0	38.31	9.6	63.31	25.6	77.60	16.0		
19.65	28.5	40.46	10.4	63.64	19.2	77.93	0.8		

Table 2. Carbonate concentrations from borehole BV-2 expressed as mass ratio of CO_2 in mass % - digitized and recalculated data from Pohar (1978)

Depth [m]	CO_2 [%]	Depth [m]	CO_2 [%]	Depth [m]	CO_2 [%]	Depth [m]	CO_2 [%]	Depth [m]	CO_2 [%]
1.00	34.5	33.75	14.5	60.23	15.0	79.48	1.7	101.38	43.4
2.61	32.9	34.35	21.7	60.43	11.1	79.88	13.9	101.49	42.3
2.71	26.7	35.26	22.8	61.54	10.6	82.08	18.9	101.99	1.7
7.53	36.8	35.56	16.2	62.44	18.9	82.98	19.5	102.39	25.1
15.47	4.5	36.56	27.8	63.04	26.2	83.89	23.9	102.90	37.3
17.38	5.0	37.57	8.9	63.54	21.2	85.49	20.1	103.20	12.3
17.68	30.6	40.58	1.1	63.94	18.9	86.29	24.5	103.70	25.6
18.38	22.3	41.99	41.2	64.64	18.9	86.69	26.7	104.51	30.1
20.39	26.7	42.45	41.8	65.45	18.4	88.10	23.4	105.51	20.1
20.69	26.2	43.49	7.8	65.55	26.2	88.50	25.6	105.72	27.8
21.09	24.5	43.79	38.4	66.25	19.5	88.70	25.6	106.42	35.6
22.40	38.4	44.80	44.0	67.25	1.7	88.90	17.8	106.93	22.3
22.90	40.1	45.00	35.1	70.46	11.1	89.10	39.0	107.33	31.2
23.30	12.3	46.00	3.3	72.36	20.1	89.60	21.7	107.53	32.3
23.71	11.1	47.10	14.5	72.76	11.1	90.00	21.2	108.34	0.0
25.31	16.7	47.71	9.5	73.16	21.2	92.12	21.2	108.74	21.7
25.92	26.7	48.51	8.4	73.56	13.4	92.32	2.2	109.85	29.0
26.22	16.2	48.61	20.1	74.27	23.9	95.34	12.8	110.85	23.4
26.82	10.0	49.71	16.7	75.47	17.8	97.96	32.3	111.06	39.0
27.62	8.9	50.51	22.3	76.17	11.1	98.76	43.4	111.86	1.1
28.53	10.0	51.71	26.7	76.47	13.4	99.77	31.7	113.98	41.2
28.83	33.4	53.42	22.8	76.97	10.0	100.18	43.4	114.38	37.9
29.63	22.8	56.12	40.1	77.57	17.8	100.38	32.9	115.39	40.1
30.54	40.7	57.73	3.3	78.47	12.3	100.98	29.0		
33.35	12.3	59.43	22.3	78.88	14.5	101.18	31.2		

Table 3.
Calcimetry data in CO_2 mass ratio [%] from boreholes BV-1 and BV-2 - descriptive statistics

Descriptive statistics	BV-1	BV-2
Numerus	93	123
Min	0.8	0.0
Max	38.4	44.0
Median	10.0	21.7
Average	13.6	22.1

Table 4.
Distance between data in the vertical column [m] from boreholes BV-1 and BV-2 - descriptive statistics

Descriptive statistics	Distance between data in the vertical column [m]	
	BV-1	BV-2
Numerus	92	122
Min	0.17	0.10
Max	3.99	7.93
Median	0.83	0.70
Average	1.07	0.94

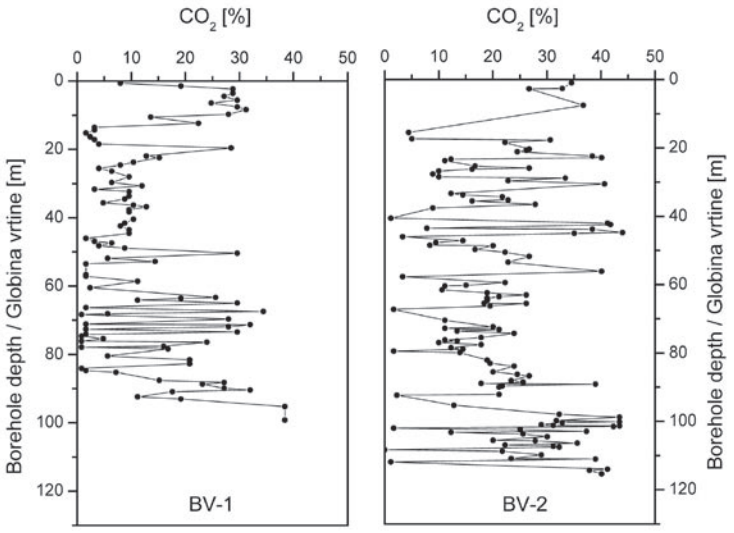

Figure 2. Calcimetry from boreholes BV-1 (Grimšičar & Ocepek, 1967) and BV-2 (Pohar, 1978) expressed as mass ratio of CO_2 in mass % along borehole depth - digitized data

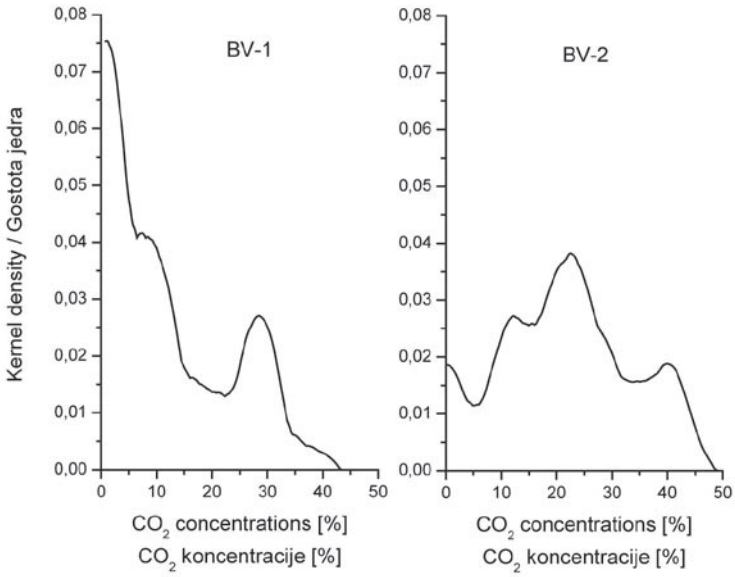

Figure 3. Kernel densities of CO_2 mass ratio concentrations [%] in boreholes BV-1 and BV-2

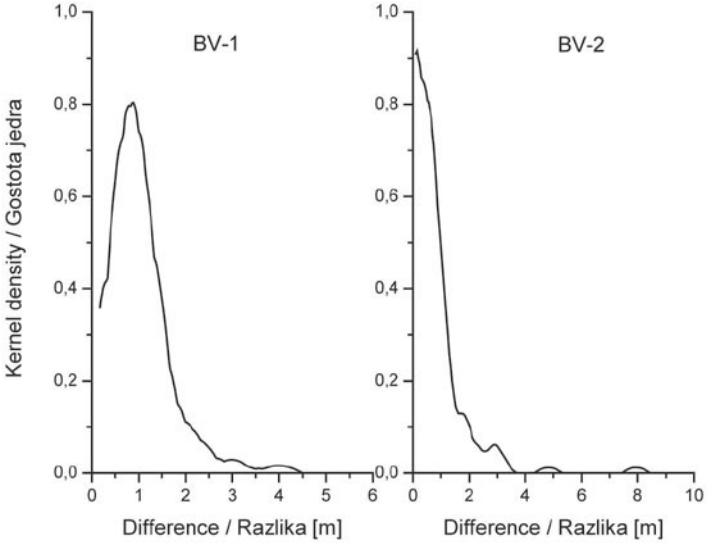

Figure 4. Kernel densities of Δt_i between adjacent data of irregular CO_2 mass ratio time series in boreholes BV-1 and BV-2

and 40 %. The basic statistical characteristics of both data sets are shown in Table 3 where a greater range of CO_2 concentrations can be observed in borehole BV-2 (44.0 %) than in borehole BV-1 (37.6 %). Asymmetry of the data from BV-1 is also reflected in the comparison between the average (10.0 %) and the median (13.6 %). In contrast, the BV-2 values of the average (21.7 %) and median (22.1 %) are very close.

In analyses of unevenly spaced time series, it is crucial to understand the empirical distribution of the time differences Δt_i between adjacent data x_i, x_{i+1} in the series. These differences indicate the sampling interval distribution of the original time series which is supposed to be continuous. Both empirical distributions represented in Figure 4 are again asymmetrical to the left and, from Table 4, it can be seen that the ranges and central tendency measures are different for the two data sets. In the calculation, a bandwidth of 0.5 m was chosen based on a comparison between different representations of the kernel densities. It can be concluded that empirical sampling interval distributions for boreholes BV-1 and BV-2 are not the same and the information content in the original time series reconstruction is not the same.

Reconstructed time series with equally spaced data are given in Figure 5. Both time series were reconstructed based on the minimum interval between adjacent data in the original, unevenly spaced time series. In borehole BV-1, the minimum interval is 0.17 m and in borehole BV-2 the minimum interval is 0.10 m. Therefore the total number of data points in the reconstructed BV-1 time series is 586 and in borehole BV-2 it is 1155. Based on the maximum entropy principle, the parameter a in the weighted influence function was determined for borehole BV-1 as 1.45 and for borehole BV-2 as 1.60.

The autocorrelation function of the reconstructed time series from both boreholes is represented in Figure 6. The autocorrelation function was calculated on the reconstructed time series where the average was subtracted and the linear trend removed. The same procedure was used for the power spectrum calculation. In the diagram, the lag τ in the autocorrelation function is represented as vertical and given as the distance reflecting the depth of the boreholes. We can observe that the autocorrelation function for the data of borehole BV-1 is slowly declining, approaching zero correlation at a lag of 10.0 m. The shape of the autocorrelation function between 0.0 m lag and 10.0 m lag can be described with an exponential decay function $r = 0.97\exp(-0.30\tau)$ with a correlation coefficient for the fitted exponential model of 0.86. The autocorre-

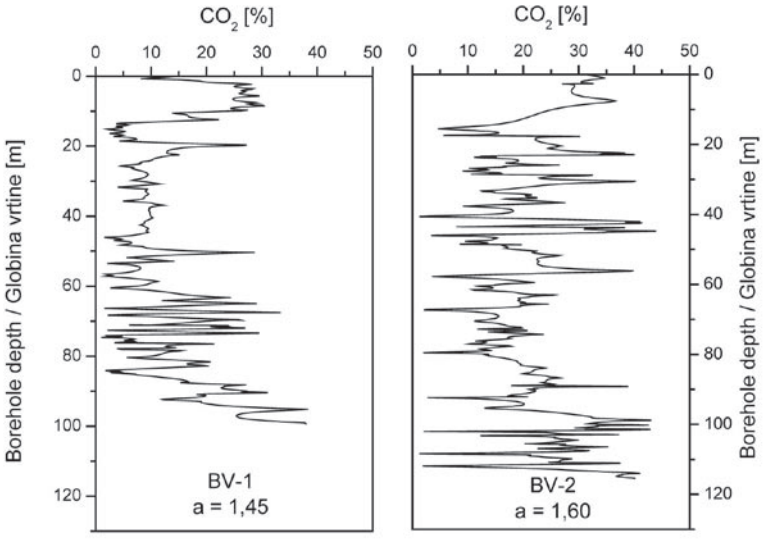

Figure 5. Reconstructed time series of CO_2 mass ratio time series in mass % along the borehole depth of BV-1 and BV-2

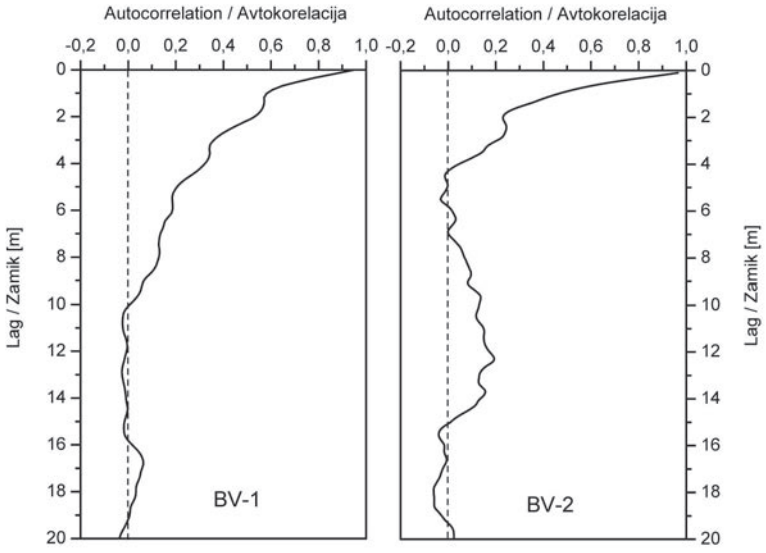

Figure 6. Autocorrelation diagrams of CO_2 mass ratio time series in mass % of BV-1 and BV-2

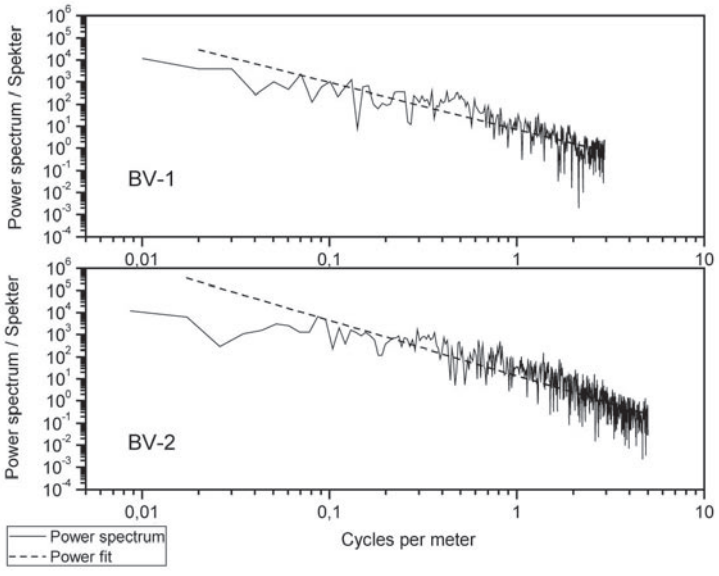

Figure 7. Power spectrum of CO_2 mass ratio time series in mass % of BV-1 and BV-2 with best fit power line

lation function of the reconstructed time series from borehole BV-2 has a different shape from the BV-1 data series, showing a stochastic process with periodicity. The autocorrelation function of the BV-1 data series approaches zero correlation at a lag of 4.2 m, but rises again from a lag of 7.0 m onwards, reaching a correlation coefficient of 0.19 at a lag of 12.3 m, which is significantly different than a zero autocorrelation based on the Box-Jenkins test. This type of autocorrelation function represents periodicity of the BV-2 time series with a period of around 12.3 m.

The structure of reconstructed BV-1 and BV-2 time series in the frequency domain is given by the power spectrum in Figure 7. Both power spectra are rather similar with some important differences reflected and observed in the above autocorrelation functions. In the BV-2 power spectrum, significantly higher values can be observed between frequencies 0.086 m^{-1} and 0.095 m^{-1} showing periodicity with periods between 10 m and 11 m. For both power spectra P, a best fit power model was calculated. For the spectrum from borehole BV-1, the best-fit line is $P = 0.16f^{-2.1}$ with a correlation coefficient of 0.86 and for the spectrum from borehole BV-2, the best-fit line is $P = 0.14f^{-2.5}$ with a correlation coefficient of 0.84. The regression line fits well with values at higher frequencies, while at lower frequencies the regression line is higher than empirical data showing that the complete spectrum is inclined from low to higher frequencies.

Discussion

In our interpretation of calcimetry data from boreholes BV-1 and BV-2 in the Ljubljansko barje, we attempted to reach two goals. The first goal was to reconstruct and prepare a time series for data published in the original papers (Grimšičar & Ocepek, 1967; Pohar, 1978) allowing reliable statistical analyses to be carried out. The second goal was to reveal possible stochastic processes that may have been responsible for generating the two time series. Reconstruction of the stochastic process can be very useful in the study of sedimentary processes as well as other related natural conditions as is the case for climate changes.

In the statistical analysis of unevenly spaced time series, two general approaches can be used. In the first approach, the missing data between existing data points in the time series are interpolated with various mathematical

interpolation functions. In the second approach, the analysis is performed on existing data without interpolation. These procedures are usually based on the assumption that an empirical time series is a sample of a stationary stochastic time series generating process. Such a procedure consists of a variogram calculation and its inversion to an autocorrelation function or spectrum of the process (Diggle, 1990). During our statistical analysis a serious attempt was made to calculate the variogram. Further work in this direction is needed, especially in relation to the empirical sampling distribution reflected in the time differences between adjacent data points.

In our statistical analyses, we have used interpolation, the approach usually adopted in cyclostratigraphy (Weedon, 2003), but the literature rarely reports the exact interpolation procedures. Therefore, we have introduced our own model of weighted influence function which is based on the assumption that the time series has some autocorrelation structure and can be described with some weighted inverse function of parameter a which relates influence to the distance Δt from the original data point. The model of weight W_t was expressed mathematically as $W_t = (\Delta t)^{-a}$. Again, in the selection of the correct and most useful parameter a, several approaches can be followed. We have tried several approaches based on a parametric moment comparison. In such an approach, we compare parametric moments of the original, unevenly spaced time series with the reconstructed, evenly spaced time series, based on the hypothesis that moments in both time series should be similar. However, it was realised that the analysed time series of BV-1 and BV-2 are very dissimilar and other criteria were needed. We have adopted the maximum entropy approach assuming that the maximum amount of information contained in the entropy is also the maximum information required for our analysis. In selection criteria based on the statistical moment comparison, further research is needed as well as better theoretical underpinnings for the maximum entropy approach.

In spite of the fact that statistical tests for the stationarity of the time series were not performed (e.g. Cromwell et al., 1994), it can be seen from their data and from their graphical presentation in Figure 2 that they probably do not represent a sample from a stationary stochastic process. At the same time, data and their time series representation show that time series from boreholes BV-1 and BV-2 are dissimilar and that they do not originate from the same time series generating stochastic process. This conclusion is in accordance with a sedimentological interpretation (Verbič & Horvat, 2009) and with a core age interpretation (Šercelj, 1966). Based on these interpretations, borehole BV-1 goes only to the beginning of Würm and borehole BV-2 goes to Mindel. Sedimentation in the region of borehole BV-1 is more related to a lacrustrine and swampy environment and only occasionally to river sedimentation. In BV-2 there have been more events and the environment has changed more frequently from a lacrustrine to river environment. Our goal here is not to discuss the causes for the relatively large differences between two boreholes in relatively close proximity, but the differences in sedimentation are reflected also in the calcimetry time series and their statistical structure is revealed by our modelling.

From the autocorrelation function and from the spectrum of borehole BV-2 (Figure 6 and Figure 7), we can find some periodicity with a length between 10.0 m and 12.5 m. This periodicity is probably related to climatic fluctuations well known for the Quaternary Period. At this stage of our knowledge, when the sediment depth of borehole BV-2 is not very well related with age, it is not possible to discern what type of periodicity it is. Again, uncovering of this periodicity is reflected in frequent changes in sedimentation environment in borehole BV-2, as detected from the core description (Verbič & Horvat, 2009). Comparing both autocorrelation functions we find that borehole BV-1 has much longer memory effects than borehole BV-2. And the decline of autocorrelation function can be described with an exponential curve. This is typical for a time series with long-term memory processes and is known as the Hurst effect (Beran, 1994).

The spectrum of both boreholes (Figure 7) is rather typical of so-called coloured noise or $1/f$ noise. The time series literature abounds with discussions about different noise types and their causes. Various basic and popular accounts can be found in the literature (e.g. Mandelbrodt, 1982, 1997; Petgen et al., 1992). Especially relevant is the discussion about climatic time series (e.g. Wunsh, 1999) and their proxies. Various noise types can be found in the time series of different natural phenomena. In our discussion about the presence of coloured noise in the BV-1 and BV-2 time series, we just wish to highlight several open questions without going into details. We have left open many of the questions for future research.

If we neglect the periodicity of the BV-2 time series, the shape of the BV-1 and BV-2 spectra in greater part resembles so-called red noise that reflects a random walk process. A regression fit of the power model to a BV-1 reconstructed time series shows that power is very near a value of -2 which is typical for red noise. The parameter for BV-2 is far from the typical red noise value; however it is in the range reported for fractal Brownian motion (e.g. Sithi & Lim, 1995). In both diagrams, the power spectrum at lower frequencies is lower than the power spectrum in the predicted red noise power line. On the left part of the diagrams, values of power spectrum at lower frequencies are relatively flat, declining to the right where the power spectrum at higher frequencies is present. In the BV-1 spectrum, values go flatter to the right than in the BV-2 spectrum. If the spectrum is completely flat, this is the sign of a so-called white noise spectrum that reflects a completely uncorrelated random stochastic process. From the BV-1 and BV-2 spectra, we can anticipate that in both cases at lower frequencies, white noise is predominant and at higher frequencies it transfers to red noise.

Transition from white noise at lower frequencies to red noise at higher frequencies is typical of the so-called Lorentz power spectrum (e.g. Tsallis, 1999).

The identification of the periodicity, red noise and white noise components in the concentrations of $CaCO_3$ based on the CO_2 mass ratio time series in the sedimentation profile at the locations of boreholes BV-1 and BV-2 on the Ljubljansko barje are interesting phenomena that were previously unknown. However, it not only has mathematical meaning from a data reduction point of view, it also has implications for the understanding of sedimentation processes in relation to climatic fluctuations during the Quaternary Period. $CaCO_3$ dynamics varies between the BV-1 and BV-2 boreholes. The sedimentation processes of $CaCO_3$ were more complex in the latter borehole.

Sedimentation processes in the BV-1 borehole can be described with a model consisting of three parts. The first part is completely random and belongs to long-term trends along the whole borehole. The second part is represented by a periodic component with a wavelength between 10.0 m and 12.0 m; this is part of the fluctuations is probably related to climatic fluctuations. The third part is related to long term memory effects at higher frequencies. We can interpret the third part in the way that small scale fluctuations of $CaCO_3$ concentrations have long term memory effects whereas on the shorter time scale, sedimentation conditions of $CaCO_3$ in the sedimentation space are very similar. As we move away in time from a particular time point, conditions become less and less similar; however the similarity declines only very slowly.

The BV-1 borehole time series can be described with a model in two parts. The first part consists of a completely random process as in the case of borehole BV-2. The second part of the model has long term memory effects that are more profound than in borehole BV-2. The similarity between adjacent data also declines very slowly along time. In this case, it is estimated that no similarity exists after an interval of 10.0 m in the sedimentation column.

Conclusions

Based on the published calcimetry data from two boreholes, BV-1 (north of Podpeč) and BV-2 (south of Črna vas), which statistically represent unevenly spaced time series, an auxiliary time series with equally spaced data was reconstructed. Reconstruction was performed with the help of a weighted influence function where the best parameter was estimated with the help of the maximum entropy principle. Reconstructed time series were analysed in the time domain with an autocorrelation function and in the frequency domain with a power spectrum. For both boreholes, our interpretation was that part of the signal consisted of a completely random signal represented by white noise while another part came from a long term memory effect represented by red noise. In the BV-1 borehole, a periodicity component can also be detected, probably reflecting influences from climatic fluctuations in the Quaternary Period during ongoing sedimentation.

Many questions remain open. Several of them relate to the statistical techniques used in the analysis of unevenly spaced time series. They are related to sampling distribution, stationarity of the parent stochastic process, the parameter selection process and the selection of appropriate statistical tests. Many of these questions can be answered with simulated stochastic processes. From a geological point of view, the statistical analysis must be related in greater detail to other geological information in the area of Ljubljansko barje and also to the available information about general conditions in the Quaternary Period. Similar analysis must be performed on other available data sets in the region. Unfortunately the BV-1 and BV-2 boreholes are currently the only boreholes from Ljubljansko barje with comprehensively published data, in spite of the fact that many boreholes have been drilled in the area.

Acknowledgement

The results were obtained through the research programme P1-0020 "Groundwater and geochemistry" supported financially by the Slovene Research Agency – ARRS.

References

Beran, J. 1994: Statistics for Long Memory Procesess. Chapman & Hall/CRC, (Boca Raton): 1–315.

Brenčič, M. 2007: Subsidence rate of Ljubljansko barje in Holocene. Geologija (Ljubljana) 50/2: 455–465.

Cromwell, J.B., Labys, W.C., Terraza, M., 1994: Univariate Tests for Time Series Models. Sage Publications (Thousand Oaks): 1–96.

Diggle, P. 1990: Time series – A Biostatistical Introduction. Oxford Science Publications (Oxford): 1–257.

Grimšičar, A. & Ocepek, V. 1967: Vrtini BV-1 in BV-2 na Ljubljanskem barju. Geologija (Ljubljana) 10: 279–303.

Mandelbrodt, B.B. 1982: Fractal Geometry of Nature. W.H. Freeman (New York): 1–468.

Mandelbrodt, B.B. 1997: Fractals and Scaling in Finance. Springer (New York): 1–551.

Mencej, Z. 1990: Prodni zasip pod jezerskimi sedimenti Ljubljanskega barja. Geologija (Ljubljana) 31/32: 517–548.

Pavšič, J. (ed.) 2008: Ljubljansko barje: neživi svet, rastlinstvo, živalstvo, zgodovina in naravovarstvo. Društvo Slovenska matica (Ljubljana) 1–214.

Petgen, H.O., Jürgens, H., Saupe, D. 1992: Chaos and Fractals. Springer (New York): 1–984.

Pohar, V. 1978: Granulometrična analiza sedimentov z Ljubljanskega barja. Rudarsko metalurški zbornik (Ljubljana) 2/3: 177–186.

Reiss, R.D. & Thomas, M. 1997: Statistical Analysis of Extreme Values. Birkhäuser (Basel): 1–316.

Ross, S. 1989: A First Course in Probability. Macmillan Publishing Group (New York): 1–422.

Sithi, V.M & Lim, S.C. 1995: On the spectra of Riemann-Liouville fractional Brownian motion. Journal of Physics A: Mathematics and General, 28: 2995–3003.

Šercelj, A. 1965: Paleobotanične raziskave in zgodovina Ljubljanskega barja. Geologija (Ljubljana) 8: 5–27.

Šercelj, A. 1966: Pelodne analize pleistocenskih in holocenskih sedimentov Ljubljanskega barja. Razprave SAZU IX/9 (Ljubljana): 431–472.

Sovinc, I. 1965: Nekaj osnovnih geotehničnih značilnosti sedimentov iz vrtine BV-1 med Notranjimi goricami in Podpečjo na Ljubljanskem barju. Geologija (Ljubljana) 8: 28–33.

Tsallis, C. 1999: Nonextensive statistics: Theoretical, experimental and computational evidences and connections. Brazilian Journal of Physics (Sao Paulo) 29: 1–35.

Verbič, T. & Horvat, A. 2009: Geologija Ljubljanskega barja. V: Turk, P., Istenič, J., Knific, T., Nabergoj, T. (ured.) Ljubljanica – kulturna dediščina reke, Narodni muzej Slovenije (Ljubljana): 13–19.

Weedon, G. 2003: Time-Series Analysis and Cyclostratigraphy. Cambridge University Press (Cambridge): 1–259.

Williams, G.P. 1997: Chaos Theory Tamed. Joseph Henry Press (Washington) 1–499.

Wunsh, C. 1999: The interpretation of short climate records, with comments on the north Atlantic and southern oscillations. Bulletin of the American Meteorological Society (Boston) 80: 245–255.