Entomofauna

ZEITSCHRIFT FÜR ENTOMOLOGIE

Band 34, Heft 5: 81-104 ISSN 0250-4413 Ansfelden, 2. Januar 2013

Prediction of Bees (Apoidea, Hymenoptera) Frequency and Distribution Using Best linear Spatial Prediction Techniques (Kriging and Cokriging)

Saman MONFARED, Kavoos KHORSHIDIAN, Alireza MONFARED, Zargham MOHAMMADI & Razeyeh KHODAPARAST

Abstract

1. In this research, data prepared from a field work study on bees (Super family: Apoidea, order: Hymenoptera) in 31 localities of Fars province (Iran), from January 2010 to August 2011 which in summation more than 3000 specimens of Apoidea of 6 families and 171 species were collected.

2. Kriging and Cokriging are techniques of making best linear unbiased estimates of regionalized variables at unsampled locations using hypothesis of stationary and structural properties of the covariance and the initial set of data values. These techniques used for prediction of frequency and geographical distribution of Bees' species.

3. For selecting the best prediction among different methods, method of leave-one-out cross validation with various criteria were used. Also we have presented prediction map for these variables based on the all models.

4. Results shown that the Cokriging method which used parameters of elevation and mean year round temperature auxiliary variables was the best prediction model for predicting of frequency and spatial distribution of Bees. Frequency of bees had positive correlation with elevation and negative correlation with mean year round temperature respectively. However, the ordinary cokriging can present a good prediction based on spatial correlations.

5. Using these methods would make logical predictions of insect frequencies and distributions with low costs in ecological entomology, especially in mountainous regions where sampling might be hard and time consuming.

Keywords: Kriging, Cokriging, Bees frequency, Apoidea, pollinator insects.

Introduction

Most of our crops, fruits, vegetables and other agricultural products are depended on insect pollination (KEARNS & INOUYE 1995). Among pollinator insects, superfamily of Apoidea is the most important ones. Bees are very close to flowering plants which called 'food plants' for gathering nectar and pollen. These two plant products are needed for Bees' survival. In reverse, bees have a unique role in pollination of crops, fruits and greenhouse plants (MICHENER 2007). In recent decade, many species of bees are focused for commercial purpose, e.g. bumblebees which are reared for pollinating of more than 30 vegetables and fruit plant products in greenhouse in many countries (VELTHIUS & DOORN 2006). In Iran, like to other countries, most species of Apoidea are wild and always live in mountainous regions (MONFARED et al 2007). For detecting species number and their distribution, it is needed sampling during their activity period which is always lasting in early spring to end of summer. In this period, plants are in flowering stage. In Iran which has vast mountainous regions, collecting bees' specimens is a much labour study. Bees sampling, necessarily needed costs and human labours. These kind of studies need requirements such as trips to fields, collecting insects by net, making collections, pinning specimens, preserving in good conditions, sending examples abroad to expertises for identification and making database etc. These processes are very time consuming. In this research we studied pollinator bees distribution in Fars province to submit a model of forecasting of their existence in locations and regions based on sampling data.

Since the classical statistical approach not to be able evaluate and modeling patterns of stochastic processes that have spatial variation, recently some Geostatistical techniques have been used that called Kriging and Cokriging. This technique considers a spatial variation function called variogram and covariogram based on spatial observation to analysis and predicting for unsampled data. Based on an accepted ecological fundamental, faunal (animal species live in an area) variations is as a result of floral (vegetation species grown in an area) variations which in turn it's resulted from climate circumstances (SOUTHWOOD & HENDERSON 2000). Such areas with various kinds of climates have been caused a rich floral and faunal diversity. In the other words, topography of each area make variations in weather and climate changes and this make plant variations. Finally, plant species restricts animal species composition in a expand area. As said before, bees are directly depend on flowering plants and these plants grown in an area which has appropriate precipitation and soil materials. So, climate, elevation, soil texture and vegetations are the most important factors involving and controlling their occurrence in a definite region. Based on these explanations we can assume that frequency and distribution of bees in an area is a spatial variable.

Geostatistics is a class of statistical techniques developed to analyze and predict values of a variable distributed in space or time. It begins with a type of autocorrelation analysis called variography or semivariance analysis. The variogram model is used to predict unsampled locations by kriging or conditional simulation, which produces estimates of the variable across the entire spatial or temporal domain. Interpolation based on spatial dependence of samples, was first used by KRIGE (1951) for estimation of the gold content of ore bodies in the mining industry of South Africa. However, classical interpolation procedures were considered inappropriate. They were biased and non-optimal due to the fact that they did not take into account local spatial dependence during estimation. Krige's practical methods were generalized by MATHERON (1962) into the theory of regionalized variables. A regionalized variable $\mathbf{Z(s)}$ is a random variable that takes different values according to its location $\mathbf{s}$ within some region. MATHERON (1962, 1963) showed the Kriging methods are the best linear predictors. There are many published sources for case studies using geostatistics, the books by (DAVID 1977; 1988), (JOURNEL and HUIJBREGTS 1987) and (CLARK 1979) are all devoted to applications in the mining industry. Application of geostatistics in other areas abound; for example in rainfall precipitation (e.g, ORD & REES 1979), soil mapping (e.g, BURGESS 1980), atmospheric science (e.g, THIEBAUX & PEDDER 1987), agriculture (MONFARED et al. 2011), earth quakes (MONFARED et al. 2011). To understand the significance of statistics and diversity of methods and their applications in various fields see CRESSIE (1993), BANERJEE (2004) & SCHABENBERGER 2004).

Material and Methods

Bees Collection

Bees were collected from suburban areas and mountainous regions of Fars province from January 2010 to August 2011. Geographical locations of sampling length, width and height were taking notes by GPS, made by Garmin company series eTrix Hc. Bees were collected by insect net. The locations and bees' frequency are plotted in (Fig. 1). In sampling time, bees were killed in containers in which filled with cork and mixed with ethyl acetate and were pinned by Austrian and Australian insect's 000 to 1 pines. Then these samples have been received labels. Labels had information including, location, collector, date and code of specimens. All samples were identified to genus using keys to identification in the book "The Bees of The World" by MICHENER (2007) in Yasouj University. To identify species, bees were sent to universities and research institutes in the Europe and detected by top expertise scientist on bees' taxonomy of the world (see their names in Acknowledgments). All identified species deposited in "Iranian pollinator Insect Museum", in Dept. of Plant Protection in faculty of agriculture of Yasouj University. All data are presented in (Table. 3).

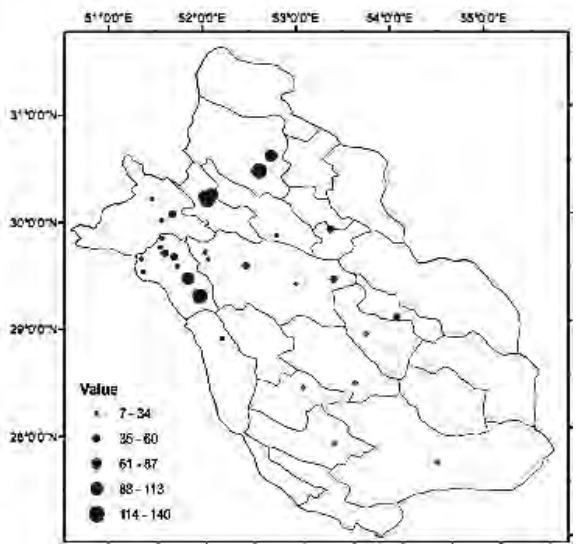

Fig1. Location and frequency of sampled bees in Fars province

Study area

Fars province has an area of about 133,299 km², and covers 8.1 percent of Iran. It is located between 27° 03` and 31° 42` Northern latitude and 50° 30` and 55° 36` Eastern longitude. Its average annual rainfall varies between 100 mm in the Southern parts and more than 400 mm in the Northern parts of the province. Based on information available from Research Centre for Agriculture and Natural Resources of Fars Province, 2011; Fars is one of the most important provinces in agricultural productions. For many years, Fars has had remarkable record in crop production. Mean elevation of Fars Province is 1350 m above the sea level. It's the 3rd largest province in Iran. The majority of the rain producing air masses enters the region from the west and the north-west, yielding relatively high precipitation amounts for those areas. Towards the south and southeast, rainfall is decreases. Furthermore, winter precipitation in the northwest area is in the form of snowfall, but for other areas it is mostly in the form of rain. Highlands of the province can be divided into four regions, the first, the northern and north-western which is composed of interconnected highland, the second located in central region which includes fertile plains among mountains. The third region is western highlands and the fourth region which is located in south and southeast that it is covered with less altitude and wide desert. There are four distinguishable climatic regions commensurate with four mentioned highlands region in the province, the first, covers the northern, the western and the north-western mountainous area, which is significant for its cold and moderate weather, as well as its significant green land. The second region is central, which is significant for its moderate winter with plenty of rain, and warm and dry summer. The third region is located on the south and the south-east of the province, which is

significantly dry and moderate in winter, and very hot in summer, because of its lowland (Fig. 2) (National Geosciences Database of Iran). This province has known as a pole of agriculture production of crops, fruit plants, forest and greenhouses products in the country (Research Centre for Agriculture and Natural Resources of Fars Province). Therefore, pollinators study in this province is at the centre of attentions by biologists.

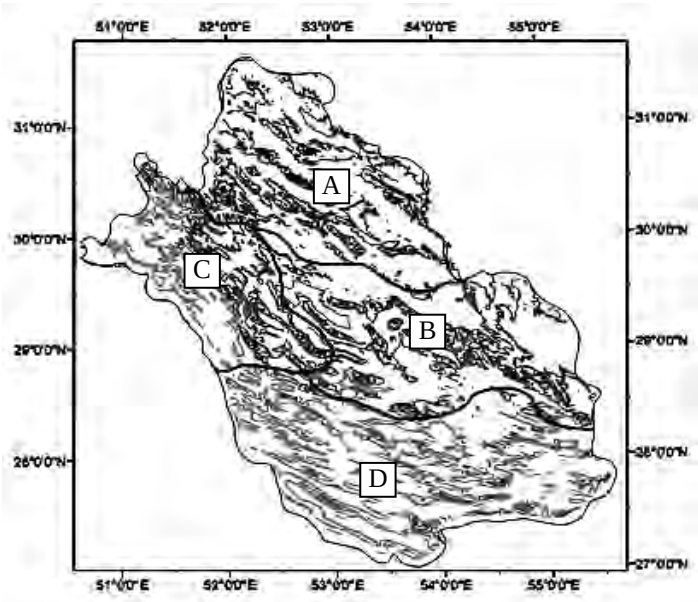

Fig2. Mountainous regions based on topography in Fars province.
Darker lines appear higher elevations.

Spatial linear model

Let $\{\mathbf{Z(s)} : \mathbf{s} \in D\}$ denote a spatial process that is observed at certain points $\{s_i = 1,2,\ldots,n\}$ where D is a fixed subset of $\mathbf{R}^d$ with positive d-dimensional volume. The spatial linear model is defined by

$$\mathbf{Z(s)} = \mathbf{Y(s)} + \boldsymbol{\varepsilon}\mathbf{(s)} \tag{1}$$

A fundamental assumption for geostatistical methods is that any two locations that are a similar distance and direction from each other should have a similar difference squared. This relationship is called stationary. Where $Y(s)$ is the underlying spatial surface and $\varepsilon(s)$ is white-noise measurement error process. Intrinsic stationary is defined if it has the following requirements (2) and (3)

$$E(\mathbf{Z(s+h)} - \mathbf{Z(s)}) = \mathbf{0} \tag{2}$$

$$VAR(\mathbf{Z(s+h)} - \mathbf{Z(s)} = \mathbf{2\gamma(h)} \tag{3}$$

The quantity $\mathbf{2\gamma(h)}$ is known as the variogram and is the crucial parameter of geostatistics. For more details see (MATHERON 1963b) and (CRESSIE 1993, sec 2.3). The classical estimate of the variogram proposed by MATHERON (1962) is

$$2\hat{\gamma}(h) = \frac{1}{N(\mathbf{h})} \sum_{N(\mathbf{h})} (z(s_i) - z(s_j))^2 \tag{4}$$

Where the sum is over $N(\mathbf{h}) \equiv \{i, j : \mathbf{s}_i - \mathbf{s}_j = \mathbf{h}\}$ and $|N(\mathbf{h})|$ is the number of distinct elements of $N(\mathbf{h})$. It is unbiased; however, it is affected by atypical observations due to the $(.)^2$ term in the summand of (4). Variogram analysis consists of the experimental variogram calculated from the data and the variogram model fitted to the data. The experimental variogram is calculated by averaging one-half the difference squared of the z-values over all pairs of observations with the specified separation distance and direction. The variogram model is chosen from a set of mathematical functions that describe spatial relationships. The appropriate model is chosen by matching the shape of the curve of the experimental variogram to the shape of the curve of the mathematical function.

Sample locations separated by distances closer than the range are spatially autocorrelated whereas locations farther apart than the range are not. According to Fig. 3, the value that the semivariogram model attains at the range (the value on the y-axis) is called the sill. The partial sill is the sill minus the nugget. Theoretically, at zero separation distance (lag = 0), the semivariogram value is zero. However, at an infinitesimally small separation distance, the semivariogram often exhibits a nugget effect, which is some value greater than zero.

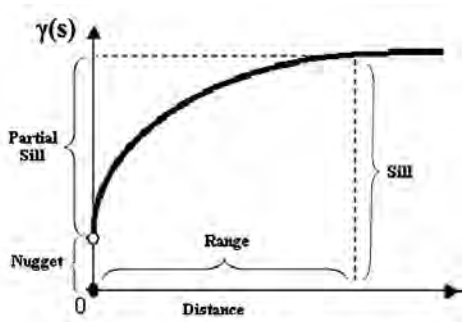

Fig3. Semivariogram, Sill, Partial Sill, Nugget and

Natural phenomena can vary spatially over a range of scales. Variation at micro scales smaller than the sampling distances will appear as part of the nugget effect.

Assume defined $\mathbf{Z(s)}$ process has following trait that

$$E(\mathbf{Z(s)}) = \boldsymbol{\mu} \qquad \text{for all } \mathbf{s} \in D \tag{6}$$

In order to estimate optimal predictors, an additional assumption such as

$$Cov(\mathbf{Z(s+h), Z(s)}) = \mathbf{C(h)} \tag{7}$$

A random function $\mathbf{Z}(.)$ satisfying (6) and (7) is defined to be second-order or week stationary. The function $C(.)$ is called covariogram. Furthermore, if $\mathbf{C(h)}$ and $\mathbf{2\gamma(h)}$ are a function only of $\|\mathbf{h}\|$ then they are called isotropic. The absence of this condition express most correlation between observations in particular direction. Learn more about geostatistical methods that are based on statistical models spatial autocorrelation may

depend only on the distance between two locations, which is called isotropy. However, it is possible that the same autocorrelation value may occur at different distances when considering different directions. Another way to think of this is that things are more alike for longer distances in some directions than in other directions. This directional influence is seen in semivariograms and covariances and is called anisotropy. It is important to look for anisotropy so that if you detect directional differences in the autocorrelation, you can account for them in the semivariogram or covariance models. This in turn has an effect on the geostatistical prediction method.

Kriging and Cokriging Techniques

Kriging belongs to the family of linear least squares estimation algorithms. The aim of kriging is to estimate the value of an unknown real-valued function **Z(s)** at a point $\mathbf{s}_0$ given the values of the function at some other points $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n$. A kriging estimator is said to be linear because the predicted value $\hat{Z}(\mathbf{s}_0)$ is a linear combination that may be written as (8), the weights λ_i are solutions of a system of linear equations which is obtained by assuming that Z, is a sample-path of a random process **Z(s)** and that the error of prediction are given by (9), it is to be minimized in some sense.

$$\hat{Z}(\mathbf{s}_0) = \sum_{i=1}^{n} \lambda_i Z(\mathbf{s}_i) \tag{8}$$

$$\varepsilon(\mathbf{s}) = Z(\mathbf{s}) - \sum_{i=1}^{n} \lambda_i Z(\mathbf{s}_i) \tag{9}$$

Simple kriging is mathematically the simplest, but the least general. It assumes the expectation of the random field to be known, and relies on a covariance function. However, in most applications neither the expectation nor the covariances are known beforehand. The kriging weights of simple kriging have no unbiasedness condition and the interpolation is given by the simple kriging equation system (10), and the simple kriging variance is given by (11) equations.

$$\hat{Z}(\mathbf{s}_0) = \gamma(\mathbf{s} - \mathbf{s}_0)' \mathbf{\Gamma}^{-1} \mathbf{Z(s)} \tag{10}$$

$$\sigma_{sk}^2(\mathbf{s}_0) = \boldsymbol{\gamma}' ((\mathbf{s} - \mathbf{s}_0) \mathbf{\Gamma}^{-1} \boldsymbol{\gamma}(\mathbf{s} - \mathbf{s}_0) \tag{11}$$

The $\boldsymbol{\gamma}$, $\boldsymbol{\lambda}$ and $\mathbf{\Gamma}$ in the above equations defined by

$$\mathbf{\Gamma} = \begin{pmatrix} \gamma(s_1 - s_1) & \cdots & \gamma(s_1 - s_n) \\ \vdots & \ddots & \vdots \\ \gamma(s_n - s_1) & \cdots & \gamma(s_n - s_n) \end{pmatrix}^{-1}$$

$$\boldsymbol{\gamma}(s - s_0) = (\gamma(s_1 - s_0), \ldots, \gamma(s_n - s_0))' \quad ; \quad \boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)'$$

Ordinary kriging is the most commonly used type of kriging. It assumes a constant but unknown mean. The kriging weights of ordinary kriging fulfil the unbiasedness condition $\sum_{i=1}^{n} \lambda_i = 1$, the interpolation are given by (12)**,** the weights and Lagrange coefficient obtained as (13) and (14) respectively. The ordinary kriging variance is given by (15) equations.

$$\hat{Z}(\mathbf{s}_0) = \boldsymbol{\lambda}'\mathbf{Z(s)} = (\boldsymbol{\gamma}_0(\mathbf{s}-\mathbf{s}_0) + \mathbf{1}m)'\Gamma^{-1}\mathbf{Z(s)} \tag{12}$$

$$\boldsymbol{\lambda}' = (\gamma_0(\mathbf{s}-\mathbf{s}_0) + \mathbf{1}\frac{1-\mathbf{1}'\Gamma_0^{-1}\gamma_0(\mathbf{s}-\mathbf{s}_0)}{\mathbf{1}'\Gamma_0\mathbf{1}})'\Gamma_0^{-1} \tag{13}$$

$$m = \frac{1-\mathbf{1}'\Gamma_0^{-1}\gamma_0(\mathbf{s}-\mathbf{s}_0)}{\mathbf{1}'\Gamma_0\mathbf{1}} \tag{14}$$

$$\sigma_{ok}^2(s_0)) = \boldsymbol{\lambda}_m'\boldsymbol{\gamma}_\mathbf{0} \tag{15}$$

The $\boldsymbol{\gamma}_0$, $\boldsymbol{\lambda}_\mathbf{m}$ and Γ_0 in the above equations defined by

$$\boldsymbol{\gamma}_0 = (\gamma(s_1 - s_0), \ldots, \gamma(s_n - s_0), 1)' \text{ ; } \boldsymbol{\lambda}_\mathbf{m} = (\lambda_1, \lambda_2, \ldots, \lambda_n, m)$$

$$\Gamma_0 = \begin{pmatrix} \gamma(s_1 - s_1) & \cdots & \gamma(s_1 - s_n) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma(s_n - s_1) & \cdots & \gamma(s_n - s_n) & 1 \\ 1 & \ldots & 1 & 0 \end{pmatrix}$$

Universal kriging assumes that the mean of process may be expressed by $\mu(s) = \sum_{j=1}^{p} \beta_j X_j(s)$ where $X(\mathbf{s}) = (X_1(\mathbf{s}), X_2(\mathbf{s}), \ldots, X_n(\mathbf{s}))'$ is known function of the spatial coordinate $\mathbf{s}$. Therefore we can write the spatial process by $\mathbf{Z(s)} = \mathbf{X(s)}\boldsymbol{\beta} + \boldsymbol{\varepsilon}\mathbf{(s)}$

Let $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)'$ be the location parameters, deterministic but unknown drift coefficient and $\boldsymbol{\varepsilon}(\mathbf{s})$ is a zero-mean random error process with $\text{var}(\boldsymbol{\varepsilon}(\mathbf{s})) = \sigma^2$ .Let $x(s_0)$ is prediction value of X in location s_0, the kriging weights of universal kriging fulfil the unbiasedness condition $\lambda' X(s) = x'(s_0)$, the interpolation based on covariogram is given by (16)

$$\hat{Z}(s_0) = x'(s_0)\boldsymbol{\beta}_{gls} + \mathbf{c}'(\mathbf{s}, s_0)\Sigma^{-1}(\mathbf{Z(s)} - \boldsymbol{\beta}_{gls}) \tag{16}$$

The weights and regression coefficient obtained as (17) and (18) respectively.

$$\boldsymbol{\lambda} = \Sigma^{-1}\{\mathbf{c}(\mathbf{s}, s_0) - \mathbf{X}(\mathbf{s})(\mathbf{X}'(\mathbf{s})\Sigma^{-1}\mathbf{X}(\mathbf{s}))^{-1}[\mathbf{X}(\mathbf{s})\Sigma^{-1}\mathbf{c}(\mathbf{s}, s_0) - x(s_0)]\} \tag{17}$$

$$\boldsymbol{\beta}_{gls} = (\mathbf{X}'(\mathbf{s})\Sigma^{-1}\mathbf{X}(\mathbf{s}))^{-1}\mathbf{X}'(\mathbf{s})\Sigma^{-1}\mathbf{Z}(\mathbf{s}) \tag{18}$$

$$\sigma_{uk}^2(s_0) = \sigma_0^2 + \mathbf{c}'(\mathbf{s}, s_0)\Sigma^{-1}\mathbf{c}(\mathbf{s}, s_0) + [x'(s_0)(\mathbf{s})]'$$

$$-\mathbf{c}'(\mathbf{s}, s_0)\Sigma^{-1}\mathbf{X}(\mathbf{s})](\mathbf{X}'(\mathbf{s})\Sigma^{-1}\mathbf{X}(\mathbf{s}))^{-1}[x'(s_0) - \mathbf{c}'(\mathbf{s}, s_0)\Sigma^{-1}\mathbf{X} \tag{19}$$

Cokriging is an extension of kriging for prediction of one variable using other variables. The co-variables must have a strong relationship and this relationship must be defined. Use of Cokriging requires the spatial covariance model of each variable and the cross-covariance model of the variables. The method can be quite difficult to do because developing the cross-covariance model is quite complicated. Developing the relationship between the variables can also be complicated.

Let $[\mathbf{Z}_1(\mathbf{s}_i), \ldots, \mathbf{Z}_k(\mathbf{s}_i) : \mathbf{s}_i \in D \subseteq R^k]$ denote k spatial process. It is desired to predict $\mathbf{Z}_i(\mathbf{s}_0)$ based on covariables $\mathbf{Z}_j = [Z(s_1), \ldots, Z(s_n), j = 1,2,\ldots,n]$ $j \neq i$. Let $E(\mathbf{Z(s)}) = \boldsymbol{\mu}$, $\mathbf{s} \in D$ and $Cov(\mathbf{Z(s), Z(u)}) = \mathbf{C(s,u)}$ $\mathbf{s}, \mathbf{u} \in D$ where $\boldsymbol{\mu} = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ldots, \boldsymbol{\mu}_n)'$ and $\mathbf{C(s,u)}$ is a $k \times k$ matrix (not necessarily symmetric). The kriging predictor of $\mathbf{Z}_i(\mathbf{s}_0)$ is a linear combination of all the available variable data values of the k variables

$$\mathbf{Z}_i(\mathbf{s}_0) = \sum_l^n \sum_j^k \lambda_{jl} Z_j(\mathbf{s}_i) \qquad j \neq i \tag{20}$$

Similar to any kriging method, we can obtain the weights and variance of each method. Although the algebra for prediction formula is complicated, the principle of exploiting covariation to improve the mean-square prediction error is the basis of cokriging. See (WOLD 1938, KOLMOGROV 1941, WIENER 1949). Usually the covariance matrix function $\{\mathbf{C(s,u)} : \mathbf{s,u} \in D\}$ has to be estimated from the data. Assuming stationary of this covariance matrix function, that is $\mathbf{C(s,u)} = \mathbf{C}^*(\mathbf{s} - \mathbf{u})$, would allow estimation of $\mathbf{C}^*(.)$ from the available data $\mathbf{Z}(\mathbf{s}_1), \ldots, \mathbf{Z}(\mathbf{s}_n)$. For satisfactory of the cross-variogram for cokriging see (CLARK 1989), for matrix criterion of cokriging see (VERHOEF 2004, MYERS 1982, 1984).

Results

Modelling the covariable

In order to perform cokriging, we must first model the spatial structure of a covariables and their covariance with the target variable. This is called a co-regionalisation. It is an extension of the theory of a single regionalised variable. The covariables for use in cokriging are correlated to the target variable in both feature and geographic space, and

have usually been more densely sampled than the target variable. Candidates for co-variables must have: 1-a feature-space correlation with the target variable; 2-a spatial structure (i.e. be modelled as a regional variable); 3-a spatial covariance with the target variable. There are two main ways to select covariables: 1-theoretically, from knowledge of the spatial process that caused the observed spatial (co-)distribution; 2-empirically, by examining the feature-space correlations (scatter-plots) and then the spatial covariance. In this research we have considered mean year round temperature and altitude variables as covariables. We will compare two possible covariables based on both. 1-Theoretically: Based on accepted ecological fundamental, Faunal (animal species live in an area variations) as a result of floral (vegetation species grown in an area) variations which in turn it's resulted from climate circumstances (SOUTHWOOD & HENDERSON, 2000). Such areas with various kinds of climates have been caused a rich floral and faunal diversity. Therefore, many factors like Temperature, elevation, humidity, sunshine, wind speed and other factors can be effective in the frequency and distribution of bees. According to unavailability of all these information, we have just considered temperature and altitude of sampled species because of their importance. Also, other factors are influenced by these two factors. To illustrate, in regions with high elevation and temperature the wind speed and sunshine are higher than regions with low elevation and temperature respectively. 2- Empirically: based on study of linear correlation and calculate the correlation coefficients of covariables with target variable. According to (Fig.4), bees frequency decrease with increasing temperature and it increase with increasing Altitude. The correlation coefficients of temperature and Altitude with bees frequency are (-0.77, 0.61) respectively.

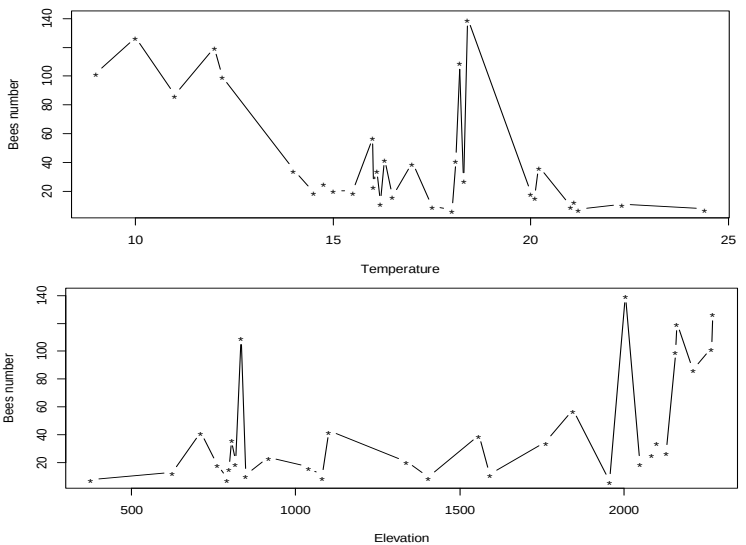

Fig4. Changes of bees frequency with decreasing temperature and increasing Elevation

According to calculated correlation coefficients we expect that temperature has more effect than Altitude. But, we don't consider these variables separately to predict bees frequency. We perform cokriging with both of them that it can help to gain best prediction with consider of interaction between temperature and Altitude with bees frequency.

Normality Test

In order to normality test of observations, we used Shapiro-Francia (SHAPIRO 1972) test, if its normality is rejected, one of the simplest ways to extend the Gaussian model is to assume that the model holds after applying a transformation to the original data. For response variables, a useful class of transformations is the Box-Cox family (Box and COX 1964), it is defined by

$$Z^* = \begin{cases} (Z^{\lambda} - 1)/Z & : \lambda \neq 0 \\ \log Z & \lambda = 0 \end{cases} \tag{21}$$

The Shapiro-Francia test used for bees frequency observations. The p-value= 0.0001 for this test reject hypothesis of normality. Therefore, the Box-Cox transformation with optimal λ =-0.14 applied for observations. Histogram of observation for original data and transformed data are presented in (Fig. 5)

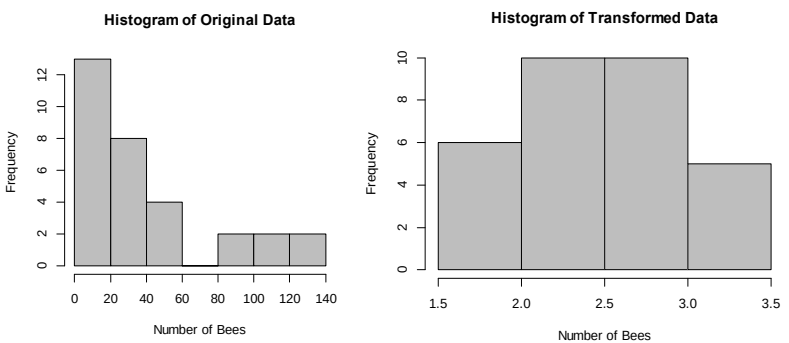

Fig5. Histogram of Bees frequency before and after Box-Cox transformation

Test of anisotropy

According to concepts in section 2, in order to evaluate of anisotropy we applied the directional sample variogram in four main direction (0, 45, 90,135) degrees (Fig. 6).

Directional variogram in three directions (0, 90, 45) degree have same model but in 45 degree direction model is different. Therefore, we can accept isotropy because it is same in three directions.

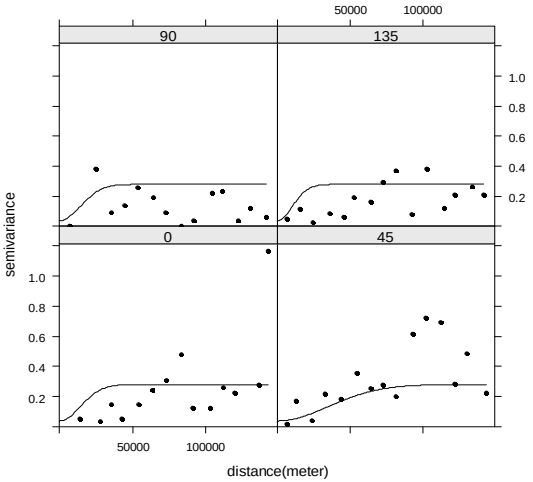

Fig6. Directional sample variogram (dot) and plotted model (curve line) in four directions (0 is North, 90 is East)

Variography

In order to choose best variogram model we obtained the best variogram and cross-variogram based on (4) for all variables with minimum of errors based on general procedure that has presented in (Fig. 7).

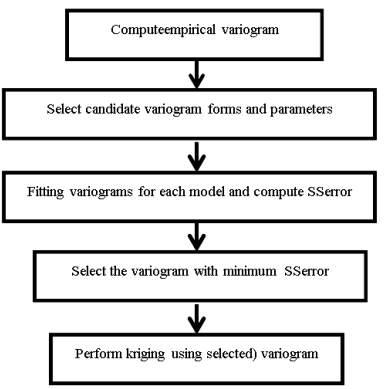

Fig7. General Procedure of selecting variogram

Results shown that for bees frequency variable for simple and ordinary kriging, Gaussian models have the lowest sum of square error (SSerror) with value (1.292119e-10) and parameters (Partial Sill= 0.2249, Nugget= 0.0367, Range= 45006.24) is the best fitted model. This model is plotted in (Fig. 8(a)). This model shows the variance between observations distance. increases with increasing Therefore, frequency of bees in closer regions has a high correlation and this correlation decrease with increasing distance. To perform universal kriging we assumed the mean of processes is a spatial variable of coordinates. In this case Gaussian models have the lowest sum of square error (SSerror) with value (4.98596e-11) and parameter (Partial Sill= 0.10368, Nugget= 0.0298, Range= 19815.81) is the best fitted model. This model is plotted in (Fig. 8(b)). This model shows that the frequency of bees have a high spatial correlation in distance less than 19 km.

In order to perform cokriging the first stage is modelling the co-regionalisations. We have to fit models to both the direct and cross-variograms simultaneously, and these models must lead to a positive definite cokriging system. The easiest way to ensure this is to fit a linear model of co-regionalisations all models (direct and cross) have the same shape and range, but may have different partial sills and nuggets. For universal cokriging we assumed that means of all processes is a spatial variable of coordinates. Parameters of best models are presented in (Table. 1) Parameters of best model for simple and ordinary cokriging are presented in (Table. 1) these models are plotted in (Fig. 9) and (Fig. 10).

According to fitted variogram both model without external drift (simple and ordinary) and with external drift (universal) have the same range 78006.24, it shows that the spatial correlation of bees frequency, temperature and altitude simultaneously occur at most in a 78 km distance. Therefore, the highest spatial correlation between observations occurs at a distance less than 78 km.

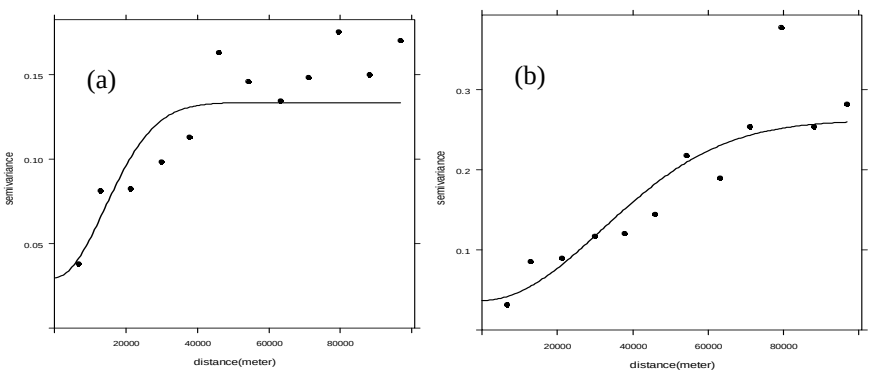

Fig8. Best fitted model for Bees frequency in simple and ordinary kriging (a), universal kriging (b)

		Simple and Ordinary		Universal	
Variogram	**Model**	**Partial Sill**	**Range**	**Partial Sill**	**Range**
Bee[1]	Nug	0.047	0	0.057	0
Bee[2]	Gau	0.33	78006.24	0.155	78006.24
Altitude[1]	Nug	23723	2372	2075	0
Altitude[2]	Gau	641280	78006.24	419310	78006.24
Temperat[1]	Nug	1.51	0	1.28	0
Temperature[2]	Gau	18.6	78006.24	7.03	78006.24
Bee.Altitude[1]	Nug	-5.9	0	-4.31	0
Bee.Altitude[2]	Gau	271	78006.24	109.34	78006.24
Bee.Temperature[1]	Nug	-0.048	0	-0.063	0
Bee.Temperature[2]	Gau	-1.85	78006.24	-0.46	78006.24
Altitude.Temperature[1]	Nug	-71.4	0	-40.7	0
Altitude.Temperature[1]	Gau	-2747	78006.24	-1260	78006.24

Table1. Parameters of best cross-variogram model for Simple,Ordinary and Universal cokriging

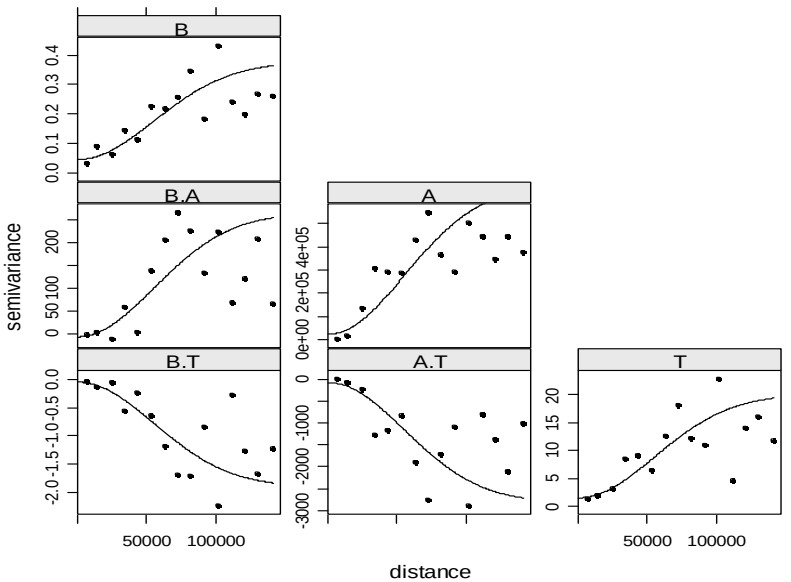

Fig9. The best fitted model on empirical cross-variogram in Simple and Ordinary. Cokriging

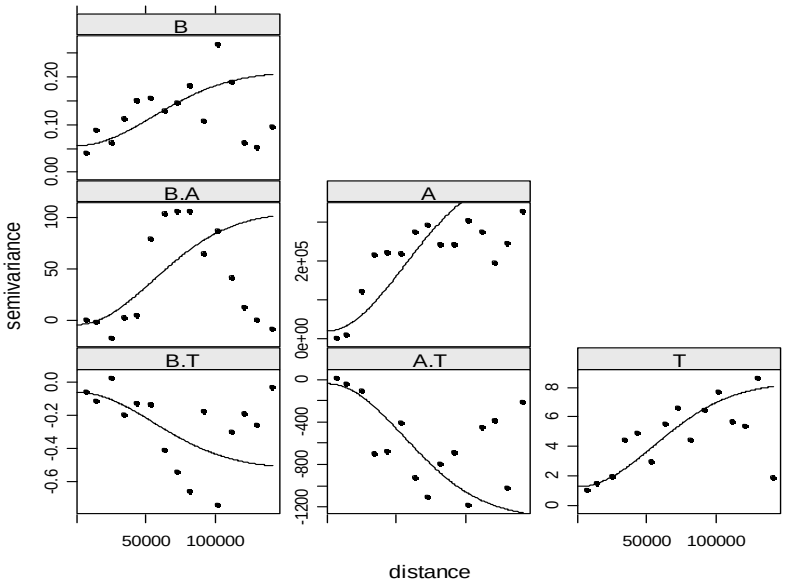

Fig10. The best fitted model on empirical cross-variogram in Universal Cokriging

Prediction and cross validation

Usually, cross-validation techniques are used to check the performance of the kriging procedure (COOPER & ISTOK 1988). Particularly, sample values $Z(\mathbf{s}_0)$ were deleted from the dataset one at a time and then the kriging process were carried out with their remaining sample values, in order to estimate the value of $Z(\mathbf{s}_0)$ at the location of the deleted sample, this procedure called leave-one-out cross validation. Three different scores were used: the mean error (ME), the mean square error, and the standardized mean square error SMSE defined as follows:

$$ME = \frac{1}{N}\sum_{i=1}^{N}(z^*(s_i) - z(s_i)) \tag{21}$$

$$MSE = \frac{\sum_{i=1}^{N}(z^*(s_i) - z(s_i))^2}{N} \tag{22}$$

$$SMSE = \sum_{i=1}^{N}\left(\frac{(z^*(s_i) - z(s_i))^2}{N \times \sigma_k^2(s_i)}\right) \tag{23}$$

According to selected variograms and mentioned concepts in section 2, The bees' frequency and distribution calculated for each kriging method (simple, ordinary, universal) and cokriging (simple, ordinary, universal) to produce the final results. Predictions for bees frequency and distribution performed for all region of study area (Fars province) in each method and leave-one-out cross validation for all method is applied to choose the best prediction method of kriging and cokriging, results are presented in (Table. 2).

Table2. Cross validation result for all methods of kriging and cokriging

Method	ME	MSE	SMSE
Simple kriging	-3.63	28.33	1.2
Ordinary kriging	-2.55	28.47	1.01
Universal kriging	-0.45	26.15	1.54
Simple cokriging	2.59	25	1.5
Ordinary cokriging	-1.53	24.2	1.01
Universal cokriging	1.35	24.38	1.4

According to results of cross validation the ordinary cokriging has a minimum mean and root-mean square error and its value of standardized mean square error is very close to 1 that it shows we don't have over or under estimate in this method. Finally according to best cross validated model the prediction map (Fig.11) and surface plot (Fig.11) for bees frequency of study area are presented.

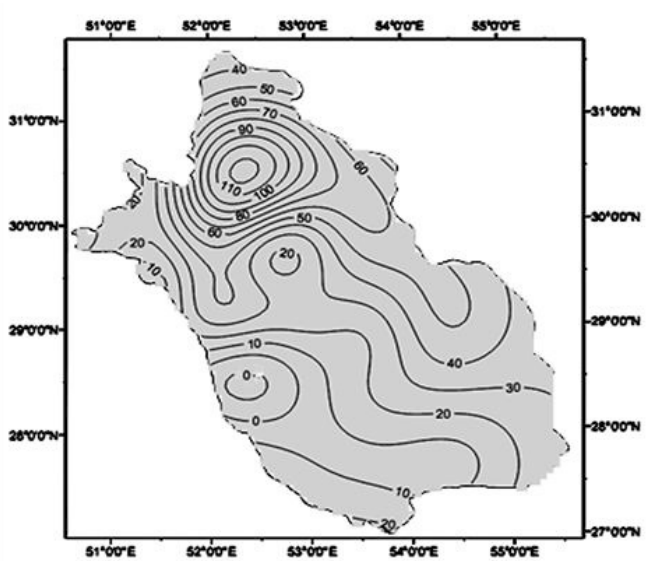

Fig 11. Prediction map for bees frequency in Fars province.

Discussion

Predictions based on ordinary cokriging, indicated that bees frequency had negative spatial correlation with temperature and positive spatial correlation with altitude. Observations shown that region A (Fig. 2) had the most bees' number which in turn had the most altitude and the least temperature. Then after regions C, B and D respectively were in turn. MONFARED et al (2008) which studied bumblebees (Apidae, *Bombus* spp.) in Alborz Mountains in Northwest of Iran, could collect these bees in elevation ranges of 70-2800m. Based on their results, most Iranian bumblebees occurred in an altitudinal range between 1500-2500m. The maximum number of species they could find was at 2000-2500m. Therefore they assumed that bumblebees habitat selection was depends on the altitude (MONFARED et al. 2008). They also found a sharp increase in the species richness peaking at 2500m after which it declines again.

Whereas the irregular changes in bees frequency with decreasing temperature and increasing altitude show that just increasing and decreasing elevation and temperature alone does not guarantee an increase in the number of bees (Fig. 3). But it depends on consideration of some factors such as temperature, elevation and other climate factors simultaneously. Bees need other requirements to survive their coloney like proper soil texture to make their nest either for their colony or pass the winter period, various food plant species to exploit nectar and pollen etc. Therefore, some factors like proper nest place which bees have to borrow ground to make their nest, winter nests and food plant which bees are depended on would affect frequency of bees' populations in an area other than two main mentioned factors of climate and elevation. However, the ordinary cokriging with consideration elevation and temperature simultaneously present a good prediction model based on spatial correlation between selected covariables. According to Mountainous regions based on topography in Fars province (Fig. 2) and prediction map based on ordinary cokriging (Fig. 10), the northern and north-western (region A) has interconnected highland with maximum altitude and the lowest temperature in Fars province has the most Bess frequency. The other hand the region (D) has the lowest bees' frequency with ranges 0-20 number which is located in south and southeast and this region has the less altitude and wide desert. Another point that it should be considered is that the bees' frequency does not necessarily increase with decreasing temperature and increasing altitude, because decreasing temperature near 0°C might make some difficulties for bees. Collectively this method would made predictions of bees as well as other insects' distributions in localities reliable and easy for biologists.

Acknowledgment

The authors are very thankful from expertise Scientists in bees taxonomy Drs. Andreas Müller, Maximilian Schwarz, Christophe Praz, Andreas Werner Ebmer, Alain Pauly, Stephan Risch, Michael Terzo, Erwin Scheuchl and Ardeshear Ariana for their valued helps in specimens identification. We also thank to Dr. Alireza Nematollahi for his advices and Dr. A. Dehdari for making facilities in the Yasouj University in this study.

Table3. Observations with their related coordinates and locations.

Longitude	Latitude	Elvations	Mean Temp.	Bees freq.	Locations
724591.20	3314618.73	1643	16	58	Arsenjan(Peerebasapha)
664398.37	3391017.32	2266	10	102	Eqlid
652492.09	3375473.02	2270	10	127	Eqlid
791486.33	3222963.66	1641	14	35	Estahban(Sahrayeserishk)
615214.23	3199954.38	776	18	7	Firuzabad (Farashband)
557144.39	3289368.70	711	18	42	Kazerun (Dadin)
553887.95	3304675.15	917	16	24	Kazerun (Ghaemeih)
580437.78	3262701.79	834	18	41	Kazerun (Ghalae seied)
552237.21	3294977.07	761	20	19	Hasanabadeabkhonar
569835.27	3275392.67	798	20	16	Kazerun (Kacekan)
533284.60	3283199.39	849	22	11	Kazerun (Kamarej)
534975.80	3269481.82	376	24	8	Kazerun (Konartakhteh)
566461.74	3285191.96	805	20	37	Kazerun (Seifabad)
727898.50	3262032.69	1762	16	35	Kherameh.
670460.36	3307627.95	1591	16	12	Marvdasht.
564217.87	3330263.90	1101	16	43	Nurabad (Ghandil)
553422.84	3323067.22	1040	16	17	Nurabad (Khomezar)
604205.54	3350337.55	2210	10	87	Sepidan.
599993.36	3345022.44	2161	12	120	Sepidan (Bahrghan)
597174.68	3348780.60	2157	12	100	Sepidan (Bahrghan)
639266.96	3276246.60	1558	16	40	Shiraz (Baghejannat)
597926.32	3289776.30	2049	14	20	Shiraz (ChehelCheshmeh)
600932.92	3283011.48	2084	14	26	Shiraz (Dashte arzhan)
689867.41	3257422.80	2128	18	28	Shiraz (Zangeneh)
760786.36	3204935.40	1336	15	21	Fasa
696955.50	3148049.90	1082	21	10	Ghirokarzin (Rikan)
728705.57	3090172.20	625	21	13	Khonj (Hanganooyeh)
832885.46	3070328.45	792	21	8	Lar
749841.67	3153355.20	1404	16	10	Jahrom
543751.34	3345928.02	817	15	20	Nurabad (Zirdu)
592519.65	3244132.35	864	18	46	Kazerun (Bidzard)

Contribution of authors

Project design: Saman Monfared (Msc student of Mathematical statistics) & Dr. Alireza Monfared (Assistant professor in Agricultural Entomology).

Data collections: Razeyeh Khodaparast (MSc. Student in Entomology) & Dr. Alireza Monfared.

Analysis: Saman Monfared, Dr. Kavoos Khorshidian, Dr. Zargham Mohammadi & Dr. Alireza Monfared.

Paper writing: Saman Monfared & Dr. Alireza Monfared, Dr. Kavoos Khorshidian, Dr. Zargham Mohammadi & Razeyeh Khodaparast.

References

BANERJEE S., CARLIN B.P. & A.E. GELFAND (2004): Hierarchical Modeling and Analysis for Spatial Data, Boca Raton. – Chapman and Hall.CRC.

BOX G.E.P. & D.R. COX (1964): An analysis of transformations. – JRSS B **26**: 211-246.

BURGESS T. M. (1980): Optimal interpolation and isarithmic mapping of soil properties. I. hesemi-variogram and punctual kriging. – Journal of Soil Science **31**: 315-331.

CLARK I. (1979): Practical Geostatistics. – Applied Science Publishers, 129 pp.

CHORTI A. & D.T. HRISTOPULOS (2008): Non-parametric Identification of Anisotropic (Elliptic) Correlations in Spatially Distributed Data Sets, IEEE. – Transactions on Signal Processing **56** (10): 4738-4751.

CRESSIE N.A.C. (1993): Statistics for Spatial Data. – New York, John Wiley and Sons.

DAVID M. (1977): Geostatistical Ore Reserve Estimation. – 2nd edition, Elsevier, 364 pp.

DAVID M. (1988): Handbook of Applied Advanced Geostatistical Ore Reserve Estimation. – Elsevier, 216pp.

JOURNEL A.G. & C.J. HUIJBREGTS (1978): Mining Geostatistics. – London: Academic Press.

KOLMOGOROV A. (1941): Interpolation and extrapolation of stationary random se quences. – Izv. Akad. Nauk SSSR. Ser. Mat **5**: 3-14.

MATHERON G. (1962): Traite de geostatistique applique, Tome I. – Memoires du Bureau de Recherches Geologiques et Minieres. No. 14. Editions Technip (Paris).

MATHERON G. (1963): Principles of geostatistics. – Economic Geology **58**: 1246-1266.

MICHENER C.D. (2007): The bees of the world. – The John Hopkins University Press, Baltimore (MD)/USA.

MONFARED A., TALEBI G. & P.H WILLIAMS (2007): A Survey of the Localities and Food Plants of the Bumblebees of Iran. – Entomologia Generalis **30** (4): 283-299.

MONFARED A., TALEBI A.A., TAHMASBI G. EBRAHIMI E. & J. BIESMEIJER (2008): Bumblebee (Hymenoptera: Apidae) diversity and abundance in the Iranian Alborz Mountains. – Zoology in the Middle East **44**: 67-74.

MONFARED S., ZAMANI A., ISMAILI S. & K. KHORSHIDIAN (2011): Spatial Prediction Analysis of Magnitude and Focal Depth of Earthquakes in Fars Province of Iran Using Kriging and Cokriging. – 1st Student Conference in Earth Science; Iran. [In Persian9.

MONFARED S., ZAMANI A., ISMAILI S. & K. KHORSHIDIAN (2011): A new application of Geostatistical Techniques in Predicting Seismo-tectonic Stress Trends in Fars Province of Iran. – 15th Symposium of Iranian Geological Society Iran. [In Persian].

MYERS D. (1982): Matrix formulation of co-kriging. Math. Geology **14**: 249-257.

MYERS D.E. (1984): Cokriging: New Developments. Geostatistics for Natural Resource Characterization. In: VERLY G. et al. (eds), D. Reidel Pub. Co., Dordrecht, 295-305. National Geoscience Database of Iran, 2011. http://ngdir.ir/.

ORD J.K.M.R. (1979): Spatial processes: recent developments with applications to hydrology, pp. 95-118. – In: LLOYD E.H., O'DONNELL T. & J.C. WILKINSON (eds), The mathematics of hydrology and water resources. Academic Press, London.

PETRAKIS Em. & D.T. HRISTOPULOS (2009): A non-parametric test of statistical isotropy for Differentiable Spatial RandomFields in Two Dimensions. Work in progress.

PETRAKIS E.a.H., D.T. (2009): A non-parametric test of statistical isotropy for Differentiable Spatial Random Fields in Two Dimensions. – Work in progress. Research Center for Agriculture and Natural Resources of Fars Province, 2011. http://fars.areo.ir.

KEARNS C.A. & D.W INOUYE (1995): Techniques for pollination biologists. – University press of Colorado, Niwot, Colorado, P583.

SHAPIRO S.S. & R.S. FRANCIS (1972): An Approximate Analysis of Variance Test for Normality. – American Statistical Association **67**: 215-216.

SCHABENBERGER O. (2004): Statistical Methods for Spatial Data Analysis. – Boca Raton, Chapman and Hall – CRC.

SOUTHWOOD T.R.E. & P.A. HENDERSON (2000): Ecological Methods. – Blackwell Science Ltd, Oxford, P573.

THIEBAUX H.J. & M.A. PEDDER (1987): Spatial Objective Analysis with Applications in Atmospheric Science London. – Academic Press, 299pp.

WIENER N. (1949): Extrapolation, Interpolation, and Smoothing of Stationary Time Series. – New York, Wiley.

VELTHUIS H.H.W. & A. van DOORN (2006): A century of advances in bumble bee domestication and the economic and environmental aspects of its commercialization for pollination. – Apidologie **37**: 421-451.

VER HOEF J.M., CRESSIE N.A.C. & R.P BARRY (2004): Flexible spatial models based on the fast Fourier transform (FFT) for cokriging. – Computnl Graph. Statist **13**, 265-282.

WOLD H. (1938): A Study in the Analysis of Stationary Time-Series. – Uppsala, Sweden, Almquist & Wiksell.

Author's addresses:

Saman Monfared
Kavoos Khorshidian
Department of Statistics, College of Sciences, Shiraz Universitry, Shiraz, Iran.

Zargham Mohammadi
Department of Earth Sience, College of Sciences,Shiraz Universitry, Shiraz, Iran.

Alireza Monfared[1]
Razeyeh Khodaparast
Plant Protection Group, Faculty of Agriculture, Yasouj University, Yasouj, Iran.

[1]Correspondence author: Dr. Alireza Monfared, Plant Protection Group, Faculty of Agriculture, Yasouj University, Yasouj, Iran. Email:alirezamonfared1@yahoo.com or amonfared@mail.yu.ac.ir

Buchbesprechungen

NENTWIG, W. (Hrsg.): **Unheimliche Eroberer.** Invasive Pflanzen und Tiere in Europa. – Haupt Verlag, Bern, 2011. 251 S.

Mehr oder weniger unbemerkt wandern jährlich Pflanzen- und Tierarten bei uns (Europa) ein; sie werden mit dem Fachbegriff Neobioten bezeichnet. Viele dieser Arten sind schon vor so langer Zeit eingewandert, dass sie von weiten Teilen der Bevölkerung gar nicht mehr als "Neubürger" betrachtet werden (z. B. Regenbogenforelle). Exakterweise gilt hier das Stichdatum 1492; alles was nach diesem Jahr in Europa von selbst einwanderte oder vom Menschen eingeschleppt wurde, ist ein Neophyt oder Neozoon. Beeinträchtigen diese Einwanderer die einheimische Biodiversität, richten wirtschaftliche Schäden an und/oder schädigen die menschliche Gesundheit, spricht man von invasiven Arten - um diese "unheimlichen" Eroberer geht es in diesem Buch. 24 Arten wurden vom Herausgeber, dem bekannten Ökologen Wolfgang Nentwig, ausgewählt und werden von entsprechenden Spezialisten vorgestellt. Diese Arten sind: Knötericharten, Riesen-Bärenklau, Drüsiges Springkraut, Opuntien, Robinien, Götterbäume, Tigermücke, Maiswurzelbohrer, Asiatischer Marienkäfer, Kartoffelkäfer; Argentinische Ameise, Pazifische Auster, Wollhandkrabbe, Amerikanischer Sumpfkrebs, Varroa-Milbe, Kanadagans, Amerikanischer Nerz, Nutria, Marderhund, Bisamratte, Waschbär, Halsbandsittich und Schmuckschildkröten.

Dieser Bild- und Textband übernimmt die wichtige und lobenserte Aufgabe, die Leser über diesen Sachverhalt aufzuklären. Anhand hervorragender Farbfotos, exakten Verbreitungskarten und gut lesbaren Texten werden Ursachen und Zusammenhänge leicht zugänglich und treffend dargestellt. Und, was fast noch wichtiger ist, es werden Auswege und Alternativen aufgezeigt, wie man aus dieser Krise wieder herauskommen könnte.

Man kann diesem hervorragenden Werk nur eine weite Verbreitung und entsprechende Aufmerksamkeit wünschen.

R. Gerstmeier

KAPPELER P.M. & D.P. WATTS (eds): **Long-term field studies of primates.** – Springer Verlag, Heidelberg, 2012. 460 S.

Langzeitstudien sind das Salz in der ökologischen Suppe, sei es "reine" Ökologie, Verhaltensökologie oder verhaltensbiologische Studien. Ihre Wichtigkeit und Außerordentlichkeit kann nur bejaht und befürwortet werden, die Erkenntnisse, welche man aus derartigen Studien ziehen kann, sind unschätzbar und stellen nicht selten sporadische Kurzzeitstudien in Frage.

In diesem Konsens spielen Primaten logischerweise eine besonders wichtige Rolle; es sind unsere nächsten Verwandten, bei vielen Arten lassen sich die Tiere individuell erkennen und es ist absolut lohnenswert, mehrere Generationen zu erforschen.

Begonnen hat dies alles mit den Untersuchungen Imamishi's in den späten 1940er Jahren an japanischen Makaken (*Macaca fuscata*). Erst später (ab 1960) folgten die Untersuchungen über Schimpansen (Goodall, Nishida, Reynolds), Bonobos (Furuichi), Paviane (Altmann) und andere.

Die Herausgeber konnten einige dieser hochrangigen Primatologen für die Beiträge in diesem Band gewinnen. Die einzelnen Artikel sind geographisch (Madagaskar, Amerika, Asien und Afrika) gegliedert und beinhalten Studien u.a. über Lemuren, Sifakas, Kapuzineraffen, Weißhandgibbons, Steppenpavianen, Diademmeerkatzen, Schimpansen und Bonobos. Sie zeigen Aufbau, Organsiation und Durchführung solcher Studien, stellen bestimmte Spezialthemen heraus und zeigen den input auf lokale Naturschutzaktivitäten. - Eine abschließende, zusammenfassende Diskussion liefert Tim Clutton-Brock.

Ein außergewöhnliches, wichtiges und damit sehr empfehlenswertes Buch für alle Ökologen, die sich "Gedanken" über Langzeitstudien machen und speziell natürlich für Primatologen.

R. Gerstmeier

Druck, Eigentümer, Herausgeber, Verleger und für den Inhalt verantwortlich:
Maximilian SCHWARZ, Konsulent für Wissenschaft der Oberösterreichischen Landesregierung, Eibenweg 6, A-4052 Ansfelden, E-Mail: maximilian.schwarz@liwest.at.

Redaktion: Erich DILLER, ZSM, Münchhausenstraße 21, D-81247 München;
Roland GERSTMEIER, Lehrstuhl f. Tierökologie, H.-C.-v.-Carlowitz-Pl. 2, D-85350 Freising
Fritz GUSENLEITNER, Lungitzerstr. 51, A-4222 St. Georgen/Gusen;
Wolfgang SPEIDEL, MWM, Tengstraße 33, D-80796 München;
Thomas WITT, Tengstraße 33, D-80796 München.

Adresse: Entomofauna, Redaktion und Schriftentausch c/o Museum Witt, Tengstr. 33, 80796 München, Deutschland, E-Mail: thomas@witt-thomas.com; Entomofauna, Redaktion c/o Fritz Gusenleitner, Lungitzerstr. 51, 4222 St. Georgen/Gusen, Austria, E-Mail: f.gusenleitner@landesmuseum.at