Tatra Mt. Math. Publ. **17** (1998), 251–264

TATRA MOUNTAINS
Mathematical Publications

CONTINUOUS SEASON MODELS

SLAVOMÍR STANO

ABSTRACT. The paper deals with seasonal models with two seasons. There is improper to use a simple changepoint to divide the seasons in some seasonal models. The aim of this paper is to recommend to use a change period instead of the changepoint. To do this we propose and justify the use of a "generalization" of the standard linear seasonal models. An algorithm for estimating of parameters in the proposed model is presented. Some examples and simulations for a comparison are discussed.

1. Linear model with constraints on the parameters as one possible seasonal model

Consider the following model:

$$\begin{aligned} y(t) &= \sum_{k=1}^{m} \theta_k f_k(t) + \varepsilon(t) \qquad \text{for} \quad t \in T_1\,, \\ y(t) &= \sum_{k=1}^{m} \gamma_k f_k(t) + \varepsilon(t) \qquad \text{for} \quad t \in T_2\,, \end{aligned} \tag{M1}$$

where $f_k(t)$ are known smooth functions, $\boldsymbol{\theta}' = (\theta_1, \ldots, \theta_m)$ and $\boldsymbol{\gamma}' = (\gamma_1, \ldots, \gamma_m)$ are unknown parameters from any parametric space $\boldsymbol{\Omega}$ with dimension m, $\varepsilon(t)$ is the Gauss white noise with unknown dispersion, and T_1, T_2 are periods of time (so called seasons), often considered as the intervals like $T_1 = \langle t_U, t_W)$ and $T_2 = \langle t_W, t_V \rangle$.

Further assume that the mean value of the process $y(t)$ is a p-times continuously differentiable function during both seasons and in the point of the change, too.

These models are used to model data describing two seasons with the same modeling functions, but the different season parameters, and, moreover, this

2000 Mathematics Subject Classification: 62J05, 62M10.
Keywords: changepoint, linear and nonlinear regression.

Supported by the VEGA grant of the Slovak Grant Agency No. 1/4196/97.

change between the seasons is continuous. Briefly, we can talk about a model of the altering season parameters without discontinuity in the mean. These models try to describe a system, which had stabilized functions describable by linear model, but after some change its characteristics (or parameters) were altered.

Methods based on linear model with constraints on parameters are used to estimate the parameters $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$. Assume that the observations are taken in times $t_1, \ldots, t_{n_1}, \ldots, t_{n_1+n_2}$ such that $\{t_1, \ldots, t_{n_1}\} \subset T_1$ and $\{t_{n_1+1}, \ldots, t_{n_1+n_2}\} \subset T_2$. Denote $\boldsymbol{y}' = \big(y(t_1), \ldots, y(t_{n_1+n_2})\big)$ and $\boldsymbol{\varepsilon}' = \big(\varepsilon(t_1), \ldots, \varepsilon(t_{n_1+n_2})\big)$, so the model (M1) may be rewritten to the form

$$\boldsymbol{y} = \begin{pmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{\theta} \\ \boldsymbol{\gamma} \end{pmatrix} + \boldsymbol{\varepsilon}\,, \tag{M1'}$$

where

$$\mathbf{F} = \begin{pmatrix} \boldsymbol{f}(t_1) \\ \vdots \\ \boldsymbol{f}(t_{n_1}) \end{pmatrix} \qquad \mathbf{G} = \begin{pmatrix} \boldsymbol{f}(t_{n_1+1}) \\ \vdots \\ \boldsymbol{f}(t_{n_1+n_2}) \end{pmatrix}\,,$$

and $\boldsymbol{f}(t) = \big(f_1(t), \ldots, f_m(t)\big)$. Constraints on the parameters can be written as

$$\mathbf{C} \cdot \begin{pmatrix} \boldsymbol{\theta} \\ \boldsymbol{\gamma} \end{pmatrix} = \mathbf{0}\,,$$

where

$$\mathbf{C} = \begin{pmatrix} \boldsymbol{f}(t_W) & -\boldsymbol{f}(t_W) \\ \frac{\partial \boldsymbol{f}(t_W)}{\partial t} & -\frac{\partial \boldsymbol{f}(t_W)}{\partial t} \\ \vdots & \vdots \\ \frac{\partial^p \boldsymbol{f}(t_W)}{\partial t^p} & -\frac{\partial^p \dot{\boldsymbol{f}}(t_W)}{\partial t^p} \end{pmatrix}\,.$$

The LSE-estimation of $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$ can be obtained as

$$\begin{pmatrix} \widehat{\boldsymbol{\theta}} \\ \widehat{\boldsymbol{\gamma}} \end{pmatrix} = \arg \min_{(\boldsymbol{\theta},\boldsymbol{\gamma}) \in \mathbf{Q}} \big((\boldsymbol{y}_1 - \mathbf{F} \cdot \boldsymbol{\theta})'(\boldsymbol{y}_1 - \mathbf{F} \cdot \boldsymbol{\theta}) + (\boldsymbol{y}_2 - \mathbf{G} \cdot \boldsymbol{\gamma})'(\boldsymbol{y}_2 - \mathbf{G} \cdot \boldsymbol{\gamma})\big)\,,$$

where

$$\mathbf{Q} = \left\{ (\boldsymbol{\theta}, \boldsymbol{\gamma}) \in \boldsymbol{\Omega}^2,\ \mathbf{C} \cdot \begin{pmatrix} \boldsymbol{\theta} \\ \boldsymbol{\gamma} \end{pmatrix} = \mathbf{0} \right\}\,,$$

$$\boldsymbol{y}_1' = \big(y(t_1), \ldots, y(t_{n_1})\big) \qquad \boldsymbol{y}_2' = \big(y(t_{n_1+1}), \ldots, y(t_{n_1+n_2})\big)\,.$$

Solution of this minimalization problem is

$$\begin{pmatrix} \widehat{\boldsymbol{\theta}} \\ \widehat{\boldsymbol{\gamma}} \end{pmatrix} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{y} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'(\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}')^{-1}\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{y}\,.$$

For details on this subject see [1].

EXAMPLE 1. Consider modelling of temperatures in any given locality. Further assume that the simple sinusoidal model is sufficient for modeling temperature in a stabilized ecosystem. After some changes at the end of a year with an influence on temperature (setting on a coal powerplant, big refinery or a dam), the characteristics of this function can be changed. Assume that this change corresponds to a change of the parameters. Therefore the following model can be appropriate:

$$y(t) = \theta_1 . \sin\left(\tfrac{2\pi t}{365}\right) + \theta_2 . \cos\left(\tfrac{2\pi t}{365}\right) + \theta_3 + \varepsilon(t)\,, \qquad t \in \langle 0, 365\rangle\,,$$
$$y(t) = \gamma_1 . \sin\left(\tfrac{2\pi t}{365}\right) + \gamma_2 . \cos\left(\tfrac{2\pi t}{365}\right) + \gamma_3 + \varepsilon(t)\,, \qquad t \in (365, 730\rangle\,.$$

With respect to the assumption that the function corresponding to the mean of the temperature is at least twice continuously differentiable, the parameters must satisfy the following conditions:

$$\theta_1 \cdot \sin(2\pi) + \theta_2 \cdot \cos(2\pi) + \theta_3 = \gamma_1 \cdot \sin(2\pi) + \gamma_2 \cdot \cos(2\pi) + \gamma_3\,,$$

$$\theta_1 \cdot \tfrac{2\pi}{365} \cdot \cos(2\pi) - \theta_2 \cdot \tfrac{2\pi}{365} \cdot \sin(2\pi) = \gamma_1 \cdot \tfrac{2\pi}{365} \cdot \cos(2\pi) - \gamma_2 \cdot \tfrac{2\pi}{365} \cdot \sin(2\pi)\,,$$

$$-\theta_1 \cdot \left(\tfrac{2\pi}{365}\right)^2 \cdot \sin(2\pi) - \theta_2 \cdot \left(\tfrac{2\pi}{365}\right)^2 \cdot \cos(2\pi)$$
$$= -\gamma_1 \cdot \left(\tfrac{2\pi}{365}\right)^2 \cdot \sin(2\pi) = -\gamma_2 \cdot \left(\tfrac{2\pi}{365}\right)^2 \cdot \cos(2\pi)\,,$$

hence

$$\theta_2 + \theta_3 = \gamma_2 + \gamma_3\,,$$
$$\theta_1 = \gamma_1\,,$$
$$\theta_2 = \gamma_2\,,$$

which is equivalent to $\boldsymbol{\theta} = \boldsymbol{\gamma}$, and so the model is reduced to an inadequate model as if no change happens.

The first disadvantage of the modelling with constraints on parameters (as the previous example shows) is a strong influence of the order p on the parametric space $\mathbf{Q}$. Dimension of this space is strongly limited by p, concretely $\dim \mathbf{Q} = 2m - p$. As the parameter p grows, the space $\mathbf{Q}$ shrinks and hence causes bad characteristic of estimators $\widehat{\boldsymbol{\theta}}$ and $\widehat{\boldsymbol{\gamma}}$, e.g., big variance of these estimators. So it is strongly recommended to ensure that $2m \gg p$.

The second drawback of this method is the fact that the seasonal alteration usually lasts for a significant period of time until the system reaches a stabilized state, which can be described by a linear model again.

Our main aim is to propose a different approach suitable for seasonal data, which tries to avoid the disadvantages mentioned above.

2. Alternative seasonal model

The suggested model is based on replacing the changepoint t_W by a period of time. Thus, the observation period (both seasons T_1 and T_2) consists of three different periods with the same modelling functions f_k, but different parameters. More precisely, we consider the following periods of time:

1. P_1 –period before the change, where the mean of the observations is the same as in the first season T_1 in the model (M1), and $P_1 \subset T_1$,
2. P_{xy} –period of the change, where the parameters of the mean depend on the time in this period,
3. P_2 –period after the change, where the mean of the observations is the same as in the second season T_2 in the model (M1), and $P_2 \subset T_2$

Consider the period of the change as an interval, so $P_{xy} = \langle t_x, t_y \rangle$, where $t_x \in T_1$ is the beginning of the change period and $t_y \in T_2$ is the end of the change period, therefore the relations between all periods can be written as

$$P_1 \cup P_{xy} \cup P_2 = \langle t_U, t_x) \cup \langle t_x, t_y \rangle \cup (t_y, t_V \rangle = \langle t_U, t_W) \cup \langle t_W, t_V \rangle = T_1 \cup T_2 \,.$$

Using this notation, we can write the proposed model

$$\begin{aligned} y(t) &= \sum_{k=1}^{m} \theta_k f_k(t) + \varepsilon(t)\,, && t \in P_1\,, \\ y(t) &= \sum_{k=1}^{m} \psi_k(t) \cdot f_k(t) + \varepsilon(t)\,, && t \in P_{xy}\,, \\ y(t) &= \sum_{k=1}^{m} \gamma_k f_k(t) + \varepsilon(t)\,, && t \in P_2\,, \end{aligned} \tag{M2}$$

where $\psi(t)$ is a monotonous and continuous function describing the change from $\boldsymbol{\theta}$ to $\boldsymbol{\gamma}$. The simplest one can be obtained by interpolation from $\boldsymbol{\theta}$ to $\boldsymbol{\gamma}$. Hence, for $k = 1, \ldots, m$,

$$\psi_k(t) = \left(1 - \alpha\left(\frac{t - t_x}{t_y - t_x}\right)\right) * \theta_k + \alpha\left(\frac{t - t_x}{t_y - t_x}\right) * \gamma_k$$

or simply

$$\psi_k(t) = \theta_k + \alpha\left(\frac{t - t_x}{t_y - t_x}\right) * (\gamma_k - \theta_k)\,,$$

where $\alpha(t)$ is any increasing continuous function with $\alpha(0) = 0$ and $\alpha(1) = 1$. The function $\alpha(t)$ is the so-called change function.

Just defined model (M2) can be regarded as a "generalization" of (M1), because (M1) is only a limit case of (M2). It is easy to show it by setting a sequence of increasing continuous functions converging to Heaviside function instead of the change function $\alpha(t)$.

In the case when we require only simple continuity ($p = 0$), any obvious change function is suitable. The identic function is the simplest one.

Generally, in the case when the mean of the observed process should have higher order of continuousness ($p > 0$), the change function must be at least p times differentiable and all first p derivatives of this function in 0 and 1 must be zero (of course, they are only one-sided derivatives).This condition is sufficient for higher order of continuousness.

For the sake of brevity denote by $\alpha^*(t)$ change function transformed to the interval $\langle t_x, t_y \rangle$,

$$\alpha^*(t) = \alpha\left(\frac{t - t_x}{t_y - t_x}\right).$$

A modified logistic function $\alpha_{R,S}(t)$ described below is one of good change functions. It is the simple logistic function transformed onto the interval $\langle 0, 1 \rangle$. The function $\cot(t)$ can be used as a suitable transformation, thus for $S > 0$

$$\begin{aligned} \alpha_{R,S}(t) &= \frac{1}{1 + e^{\left(R + S \cdot \cot(\pi t)\right)}} \qquad \text{for} \quad t \in (0,1)\,, \\ \text{and} \quad \alpha_{R,S}(0) &= 0\,, \qquad \alpha_{R,S}(1) = 1\,. \end{aligned} \tag{2.1}$$

The smooth continuousness (it means p can be arbitrarily large) is the main argument in favour of the use of this function. Of course, there are many similar suitable functions but we would like to recommend this function because of its simplicity and interpretability of its parameters R and S. The solution of $\alpha_{R,S}(t) = 0.5$ is

$$t_H = \frac{\operatorname{arccot}(-R/S)}{\pi}\,.$$

Let us denote this time as a halftime of the change function $\alpha(t)$. Therefore the parameter R can be interpreted as the parameter of the shift of the halftime and parameter S can be called the parameter of the speed of the change around the halftime. On Fig. 1 there is a graphical comparison of a few change functions for different parameters.

Some other details on the logistic function can be found in [2].

3. Estimation of the parameters in the proposed model

Proposed model (M2) can be rewritten in the form

$$\begin{aligned} y(t) &= \boldsymbol{f}(t)\cdot\boldsymbol{\theta}+\varepsilon(t)\,, && t\in P_1\,,\\ y(t) &= \boldsymbol{f}(t)\cdot\Big(\big(1-\alpha^*(t)\big)\cdot\boldsymbol{\theta}+\alpha^*(t)\cdot\boldsymbol{\gamma}\Big)+\varepsilon(t)\,, && t\in P_{xy}\,,\\ y(t) &= \boldsymbol{f}(t)\cdot\boldsymbol{\gamma}+\varepsilon(t)\,, && t\in P_2\,. \end{aligned}$$

The middle equation can be rewritten in another form

$$y(t)=\boldsymbol{f}(t)\cdot\big(\boldsymbol{\theta}+\alpha^*(t)\cdot(\boldsymbol{\gamma}-\boldsymbol{\theta})\big)+\varepsilon(t)\,,\qquad t\in P_{xy}\,. \tag{3.1}$$

The mentioned model can be also expressed as the matrix form

$$\boldsymbol{y}=\begin{pmatrix}\mathbf{F}_1 & \mathbf{0}\\ (\mathbf{I}-\mathbf{A})\mathbf{H} & \mathbf{A}\mathbf{H}\\ \mathbf{0} & \mathbf{G}_2\end{pmatrix}\cdot\begin{pmatrix}\boldsymbol{\theta}\\ \boldsymbol{\gamma}\end{pmatrix}+\boldsymbol{\varepsilon}\,, \tag{M2'}$$

where $\mathbf{A}=\operatorname{diag}\big(\alpha^*(t_{n_x}),\dots,\alpha^*(t_{n_y})\big)$ and

$$\mathbf{F}=\begin{pmatrix}\mathbf{F}_1\\ \mathbf{F}_2\end{pmatrix},\qquad \mathbf{G}=\begin{pmatrix}\mathbf{G}_1\\ \mathbf{G}_2\end{pmatrix},\qquad \mathbf{H}=\begin{pmatrix}\mathbf{F}_2\\ \mathbf{G}_1\,,\end{pmatrix},$$

$$\mathbf{F}_1=\begin{pmatrix}\boldsymbol{f}(t_1)\\ \vdots\\ \boldsymbol{f}(t_{n_x-1})\end{pmatrix},\qquad \mathbf{F}_2=\begin{pmatrix}\boldsymbol{f}(t_{n_x})\\ \vdots\\ \boldsymbol{f}(t_{n_1})\end{pmatrix},$$

$$\mathbf{G}_1=\begin{pmatrix}\boldsymbol{f}(t_{n_1+1})\\ \vdots\\ \boldsymbol{f}(t_{n_y})\end{pmatrix},\qquad \mathbf{G}_2=\begin{pmatrix}\boldsymbol{f}(t_{n_y+1})\\ \vdots\\ \boldsymbol{f}(t_{n_1+n_2})\end{pmatrix},$$

and $\{t_{n_x},t_{n_x},\dots,t_{n_y}\}\subset(t_x,t_y)$ are times of observations which were taken from inside of the change period P_{xy}.

Before estimating of parameters $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$ we must distinguish four cases:

1a. The borders of the change period as well as the parameters of the change function are known.

1b. The borders of the change period are known but the parameters of the change function are not.

2a. The borders of the change period are not known but the parameters of the change function are known.

2b. We know neither the borders nor the parameters.

The first case (1a) is the simplest one. It is a linear model, which is generally very well investigated so it is not necessary to mention it any more.

In this paper we will deal with the case (1b) only. Cases (2a) and (2b) are not subjects of this paper.

In the case of well known borders, the parametric space is enlarged only by the parameters $\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots, \phi_l)'$. These parameters are parameters of the change function $\alpha(t)$, so from now we will denote change function by $\alpha_{\boldsymbol{\phi}}(t)$. For the logistic function (2.1): $l = 2$ and $\boldsymbol{\phi} = (R, S)$.

Denote the observations before the change as $\boldsymbol{y}_1$, the observations from the change period as $\boldsymbol{y}_{12}$, and the observations after the change as $\boldsymbol{y}_2$. Likewise, denote errors in mentioned periods by $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_{12}, \boldsymbol{\varepsilon}_2$, respectively. Moreover, again

$$\alpha^*_{\boldsymbol{\phi}}(t) = \alpha_{\boldsymbol{\phi}}\left(\frac{t - t_x}{t_y - t_x}\right).$$

In the most cases the change function is not expressible as a linear function of $\boldsymbol{\phi}$ (e.g., function (2.1)). Moreover, obviously, after introducing of the new parameter $\boldsymbol{\phi}$, the model (M2') becomes nonlinear.

$$\boldsymbol{y} = \begin{pmatrix} \mathbf{F}_1 & \mathbf{0} \\ (\mathbf{I} - \mathbf{A}(\boldsymbol{\phi}))\mathbf{H} & \mathbf{A}(\boldsymbol{\phi})\mathbf{H} \\ \mathbf{0} & \mathbf{G}_2 \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{\theta} \\ \boldsymbol{\gamma} \end{pmatrix} + \boldsymbol{\varepsilon}, \qquad \text{(M2")}$$

where $\mathbf{A}(\boldsymbol{\phi}) = \operatorname{diag}\left(\alpha^*_{\boldsymbol{\phi}}(t_{n_x}), \ldots, \alpha^*_{\boldsymbol{\phi}}(t_{n_y})\right)$

The parameters $\boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\phi}$ in the model (M2") are not known. The Marquardt gradient method or Gauss-Newton method are used to estimate the unknown parameters in the nonlinear models. Both these methods are iteration methods. More details about them can be found in [3].

The derivation of the nonlinear part of the model can be derived from (3.1):

$$\frac{\partial\left(\boldsymbol{f}(t) \cdot \left(\boldsymbol{\theta} + \alpha^*_{\boldsymbol{\phi}}(t) \cdot (\boldsymbol{\gamma} - \boldsymbol{\theta})\right)\right)}{\partial \boldsymbol{\phi}'} = \boldsymbol{f}(t) \cdot (\boldsymbol{\gamma} - \boldsymbol{\theta}) \cdot \boldsymbol{\nabla}\alpha^*(t, \boldsymbol{\phi}),$$

where

$$\boldsymbol{\nabla}\alpha^*(t, \boldsymbol{\phi}) = \left(\frac{\partial \alpha^*_{\boldsymbol{\phi}}(t)}{\partial \boldsymbol{\phi}_1}, \ldots, \frac{\partial \alpha^*_{\boldsymbol{\phi}}(t)}{\partial \boldsymbol{\phi}_l}\right).$$

Thus, the model (M2") after the linearization can be written as

$$\begin{pmatrix} \boldsymbol{y}_1 - \boldsymbol{y}^0_1 \\ \boldsymbol{y}_{12} - \boldsymbol{y}^0_{12} \\ \boldsymbol{y}_2 - \boldsymbol{y}^0_2 \end{pmatrix} = \begin{pmatrix} \mathbf{F}_1 & \mathbf{0} & \mathbf{0} \\ (\mathbf{I} - \mathbf{A}(\boldsymbol{\phi}_\mathbf{0}))\mathbf{H} & \mathbf{A}(\boldsymbol{\phi}_\mathbf{0})\mathbf{H} & \mathbf{D}_\mathbf{0} \\ \mathbf{0} & \mathbf{G}_2 & \mathbf{0} \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{\theta} - \boldsymbol{\theta}_\mathbf{0} \\ \boldsymbol{\gamma} - \boldsymbol{\gamma}_\mathbf{0} \\ \boldsymbol{\phi} - \boldsymbol{\phi}_\mathbf{0} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_{12} \\ \boldsymbol{\varepsilon}_2 \end{pmatrix}, \qquad \text{(M3)}$$

or, after a small simplification,

$$\begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_{12} + \mathbf{D_0}\boldsymbol{\phi_0} \\ \boldsymbol{y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{F}_1 & \mathbf{0} & \mathbf{0} \\ (\mathbf{I} - \mathbf{A}(\boldsymbol{\phi_0}))\mathbf{H} & \mathbf{A}(\boldsymbol{\phi_0})\mathbf{H} & \mathbf{D_0} \\ \mathbf{0} & \mathbf{G}_2 & \mathbf{0} \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{\theta} \\ \boldsymbol{\gamma} \\ \boldsymbol{\phi} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_{12} \\ \boldsymbol{\varepsilon}_2 \end{pmatrix}, \quad \text{(M3')}$$

where

$$\mathbf{D_0} = \mathbf{D}(\boldsymbol{\theta_0}, \boldsymbol{\gamma_0}, \boldsymbol{\phi_0}) = \begin{pmatrix} \boldsymbol{f}(t_{n_x}) \cdot (\boldsymbol{\gamma_0} - \boldsymbol{\theta_0}) \cdot \boldsymbol{\nabla}\alpha^*(t_{n_x}, \boldsymbol{\phi}) \\ \vdots \\ \boldsymbol{f}(t_{n_y}) \cdot (\boldsymbol{\gamma_0} - \boldsymbol{\theta_0}) \cdot \boldsymbol{\nabla}\alpha^*(t_{n_y}, \boldsymbol{\phi}) \end{pmatrix},$$

$$\begin{aligned} \boldsymbol{y}_1^0 &= \mathbf{F}_1\boldsymbol{\theta_0}\,, \\ \boldsymbol{y}_{12}^0 &= (\mathbf{I} - \mathbf{A}(\boldsymbol{\phi_0}))\mathbf{H}\boldsymbol{\theta_0} + \mathbf{A}(\boldsymbol{\phi_0})\mathbf{H}\boldsymbol{\gamma_0}\,, \\ \boldsymbol{y}_2^0 &= \mathbf{G}_2\boldsymbol{\gamma_0}\,. \end{aligned}$$

The main problem in nonlinear regression is the convergence of the iterations. The speed of the convergence strongly depends on the values of the parameters for the first iteration. It turns out that we can choose a reasonable starting values of $\boldsymbol{\theta}, \boldsymbol{\gamma}$ as follows

$$\widehat{\boldsymbol{\theta}}_{\mathbf{0}} = (\mathbf{F}_1'\mathbf{F}_1)^{-1}\mathbf{F}_1'\boldsymbol{y}_1\,, \qquad \widehat{\boldsymbol{\gamma}}_{\mathbf{0}} = (\mathbf{G}_2'\mathbf{G}_2)^{-1}\mathbf{G}_2'\boldsymbol{y}_2\,.$$

On the other hand, starting values of the parameter $\boldsymbol{\phi}$ can not be obtained so easily. These parameters can be separated for the model (M2"), of course only under the condition that we know $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$.

Temporary, let us assume that the parameters $\boldsymbol{\theta}$, $\boldsymbol{\gamma}$ are known and $\boldsymbol{\theta} \neq \boldsymbol{\gamma}$. We can transform equation (3.1) to the form

$$\frac{y(t) - \boldsymbol{f}(t) \cdot \boldsymbol{\theta}}{\boldsymbol{f}(t) \cdot (\boldsymbol{\gamma} - \boldsymbol{\theta})} = \alpha^*_{\boldsymbol{\phi}}(t) + \frac{\varepsilon(t)}{\boldsymbol{f}(t) \cdot (\boldsymbol{\gamma} - \boldsymbol{\theta})}\,,$$

which we can regard as another nonlinear model for the values of the change function, obtained from observations $\boldsymbol{y}_{12}$. If we denote these values as $z(t)$

$$z(t) = \frac{y(t) - \boldsymbol{f}(t) \cdot \boldsymbol{\theta}}{\boldsymbol{f}(t) \cdot (\boldsymbol{\gamma} - \boldsymbol{\theta})}$$

the mean of these values as $\alpha^*_{\boldsymbol{\phi}}(t)$, and the errors of these observations as $\xi(t)$

$$\xi(t) = \frac{\varepsilon(t)}{\boldsymbol{f}(t) \cdot (\boldsymbol{\gamma} - \boldsymbol{\theta})}$$

we can write this model for $t \in P_{xy}$ as

$$z(t) = \alpha^*_{\boldsymbol{\phi}}(t) + \xi(t)\,. \qquad \text{(M4)}$$

We should remark that while the $\mathrm{Cov}(\boldsymbol{\varepsilon}) = \sigma^2 \cdot \mathbf{I}$, the errors $\boldsymbol{\xi} = \mathbf{K}^{-1} \cdot \boldsymbol{\varepsilon}_{12}$ have the covariance matrix $\sigma^2 \cdot \mathbf{K}^{-2}$, where $\mathbf{K} = \mathrm{diag}(k_1, \ldots, k_q)$ and $\mathbf{k} = \mathbf{H}(\boldsymbol{\gamma} - \boldsymbol{\theta})$, hence it gives us the possibility to use weighted estimators with weights given by the covariance matrix $\sigma^2 \cdot \mathbf{K}^{-2}$. Of course, we assume that $\mathbf{K}$ is a regular matrix, what is equivalent with the condition that all k_i are nonzero. Otherwise we will remove the observations corresponding to zeroes.

The model (M4) can be written as

$$\mathbf{K}^{-1} \cdot (\boldsymbol{y}_{12} - \mathbf{H}\boldsymbol{\theta}) = \mathbf{A} \cdot \mathbf{j} + \mathbf{K}^{-1} \cdot \boldsymbol{\varepsilon}_{12} \,,$$

where $\boldsymbol{j} = (1, \ldots, 1)'$.

This separation can be successfully performed only with known parameters $\boldsymbol{\gamma}, \boldsymbol{\theta}$, but in the case of unknown parameters (as we have), we can replace them by $\widehat{\boldsymbol{\theta}}_{\mathbf{0}}$ and $\widehat{\boldsymbol{\gamma}}_{\mathbf{0}}$. Some properties (e.g., unbiased estimation) of these linear estimators justificate this replacement in order to obtain $\widehat{\boldsymbol{\phi}}_{\mathbf{0}}$.

Generally, this model is also nonlinear, so the mentioned methods of the nonlinear regression can be applied here. Therefore, a selection of an appropriate starting values is the first step of these methods. The chosen values $\widehat{\boldsymbol{\phi}}_{\boldsymbol{z}}$ will be used as the first iteration in (M4), and the last estimation will be used as $\widehat{\boldsymbol{\phi}}_{\mathbf{0}}$ in the first iteration in the model (M3).

In the rest of the paper we want to propose one nonconventional method for choosing of $\widehat{\boldsymbol{\phi}}_{\boldsymbol{z}}$, when the logistic function (2.1) is used as the change function. The methods based on the means of groups of data are often used in a practice. Let us consider the model (M4) with $\alpha_{\boldsymbol{\phi}}(t) = \alpha_{R,S}(t)$ and observations $z(t)$ for $t = t_{n_x}, \ldots, t_{n_y}$. Denote by $\overline{t}_1$ the mean of times and by $\overline{z}_1$ the mean of the observations from the first half of the change period. Analogically, let $\overline{t}_2$ and $\overline{z}_2$ denote respective values of the second half of P_{xy}. $\widehat{\boldsymbol{\phi}}_{\mathbf{z}}$ is an estimation $(\widehat{R}, \widehat{S})$ such that $\alpha_{\widehat{R},\widehat{S}}(\overline{t}_1) = \overline{z}_1$ and $\alpha_{\widehat{R},\widehat{S}}(\overline{t}_2) = \overline{z}_2$. Usually these equations give us the unique solution, but in some nontrivial cases, the values of $\overline{z}_1$ or $\overline{z}_2$ exceed their expected limitation — the interval $(0, 1)$. The conventional solution of this case is to remove observations which caused this fact, but in practice an unconventional solution based on the truncation of these observation can be used. The truncation can give better results when the first method would remove too many observations and the resulting estimation would have too large dispersion.

4. Simulations and graphs

Twenty graphs of the function (2.1) for various R and S are presented on Fig. 1. The steepest one is the graph for $[R, S] = [4, 6]$ and the shallowest one is the graph for $[R, S] = [0.2, 0.3]$. The graphs between them are graphs

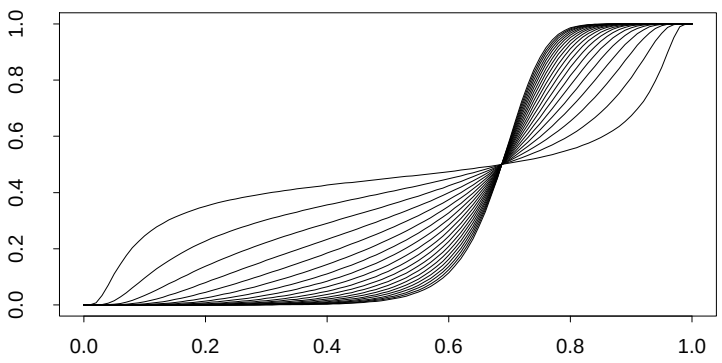

FIG. 1. Change functions from $(0.2, 0.3)$ to (4.6)

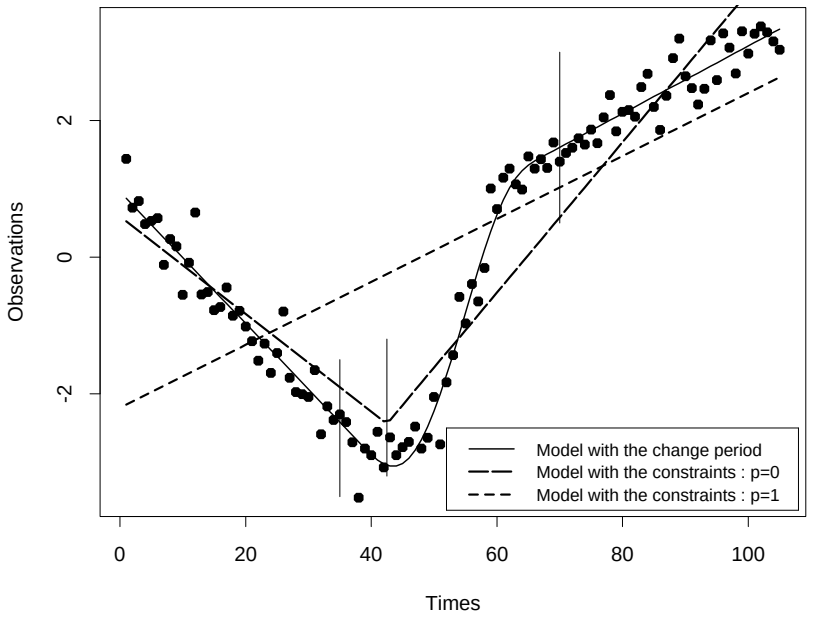

FIG. 2. Comparison of season models

for $[R, S] \in \{[0.2i, 0.3i] : i = 1, \ldots, 20\}$. Since they have the same ratio of the parameters $R/S = 2/3$, all of them have the same halftime, therefore each of them crosses the point $[\operatorname{arccot}(-2/3)/\pi, 0.5]$.

Two models were used for simulation studies of properties of the proposed model.

The first model is a model of simple linear trends for two years. The observations are taken weekly. At the end of August some change happened and it took over 8 months while the observed process was stabilized again. So the model is

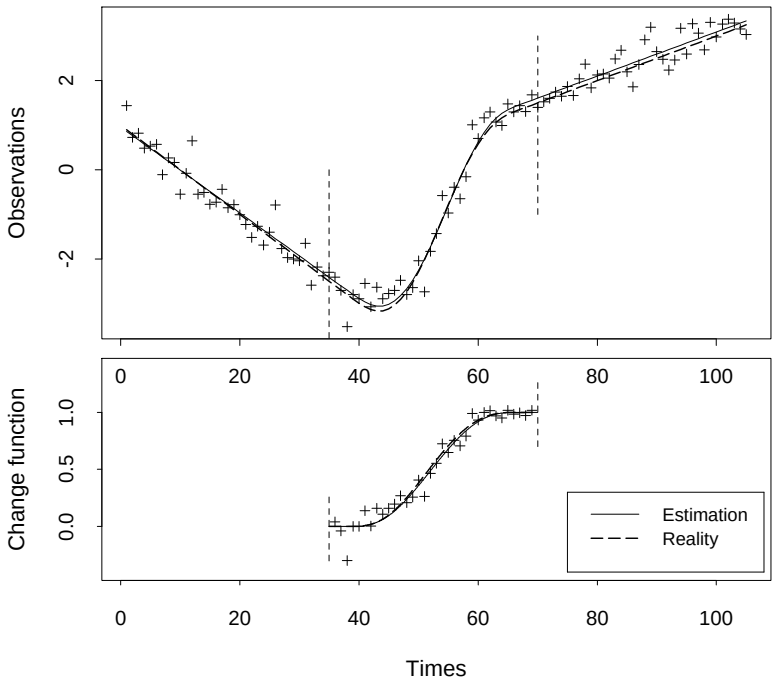

FIG. 3. Model with the change period

(the units of time are weeks)

$$\begin{aligned} y(t) &= \theta_1 + \theta_2 t + \varepsilon(t)\,, && t \in \langle 0, 35\rangle, \\ y(t) &= \psi_1(t) + \psi_2(t)\cdot t + \varepsilon(t)\,, && t \in \langle 35, 70\rangle, \\ y(t) &= \gamma_1 + \gamma_2 t + \varepsilon(t)\,, && t \in \langle 70, 105\rangle\,. \end{aligned}$$

We demand a smooth continuousness of the mean of this process. There are compared. On Fig. 2, the results of application of the model M1 to this process (with required $p \in \{0, 1\}$ and a changepoint between years) are compared with the results given by the proposed model (interpolations with the change function $\alpha_{R,S}$ are used). A plot of the change function is appended too (on Fig. 3). 200 simulations were made. Estimations of means and variances of the mentioned estimated parameters from these simulations are presented in the following rows.

Reality:	1	$-0,1$	-2	0.05	-0.2	3	$\sigma^2 = 1$
	θ_1	θ_2	γ_1	γ_2	R	S	
mean	0.9956	-0.0996	-1.9452	0.0496	-0.2243	3.8570	taken from the 1–200
variance	0.0897	0.0002	1.0363	0.0001	0.3823	8.2662	taken from the 1–200

The same procedures as in previous model are made for the second model like in the example, but the change lasts 3 years and observations are taken

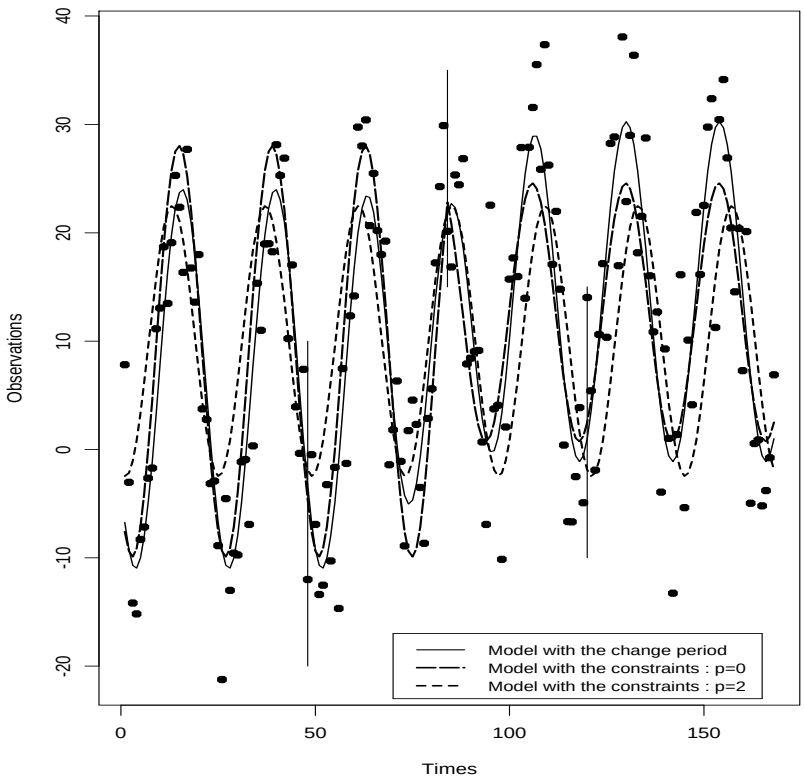

FIG. 4. Comparison of season models

twice per month for 7 years. Hence the model is (the units of time are halves of months)

$$\begin{aligned} y(t) &= \theta_1 + \theta_2 \cdot \sin\left(\tfrac{2\pi t}{24}\right) + \theta_3 \cdot \cos\left(\tfrac{2\pi t}{24}\right) + \varepsilon(t)\,, && t \in \langle 0, 48\rangle\,, \\ y(t) &= \psi_1(t) + \psi_2(t) \cdot \sin\left(\tfrac{2\pi t}{24}\right) + \psi_3(t) \cdot \cos\left(\tfrac{2\pi t}{24}\right) + \varepsilon(t)\,, && t \in \langle 48, 120\rangle\,, \\ y(t) &= \gamma_1 + \gamma_2 \cdot \sin\left(\tfrac{2\pi t}{24}\right) + \gamma_3 \cdot \cos\left(\tfrac{2\pi t}{24}\right) + \varepsilon(t)\,, && t \in \langle 120, 168\rangle\,. \end{aligned}$$

Results of the 200 simulations of mentioned procedures for this model are presented in the following rows and the last simulation in Figures 4 and 5, too.

Reality:	6	−14	−10	14	8	−15	0	2	$\sigma^2 = 36$
	θ_1	θ_2	θ_3	γ_1	γ_2	γ_3	R	S	
mean	6.0458	−14.0210	−9.8646	14.0736	8.1029	−15.0893	−0.2054	3.0058	1–200
variance	0.8783	1.9380	2.1052	0.9466	2.3227	1.8070	1.0361	20.4182	1–200

These simulations indicate that the estimators of the season parameters $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$ are very close to unbiased estimator, but the estimators of the parameters of the change function have a bias and also rather big dispersion. Big dispersion

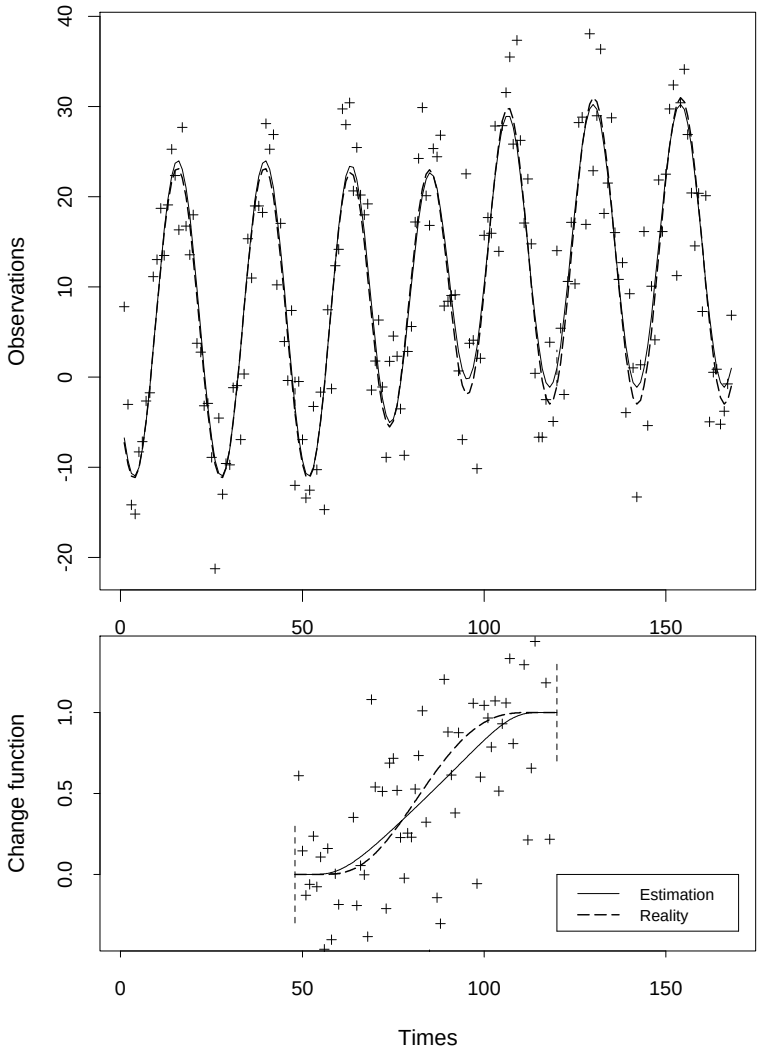

FIG. 5. Model with the change period

of ϕ can be caused by low identifiability of the change function, as it is obvious from Fig. 1. A profil of this function changes very slowly in spite of big changes of ϕ.

REFERENCES

[1] KUBÁČEK, L.: *Základy teórie odhadu*, Veda, Bratislava, 1983.

[2] RATKOWSKY, D. A.: *Handbook of Nonlinear Regression Models*, Marcel Dekker, Inc., New York, 1990.

[3] PÁZMAN, A.: *Nonlinear Statistical Models*, Kluwer, Dodrecht, 1993.

Received May 4, 1998

Department of Probability
and Mathematical Statistics
Faculty of Mathematics and Physics
Comenius University
Mlynská dolina
SK–842 15 Bratislava
SLOVAKIA
E-mail: stano@fmph.uniba.sk