ANNALI DI GEOFISICA, VOL. XL, N. 5, October 1997

Integer variables estimation problems: the Bayesian approach

Fernando Sansò and Giovanna Venuti
DIIAR, Politecnico di Milano, Italy

Abstract
In geodesy as well as in geophysics there are a number of examples where the unknown parameters are partly constrained to be integer numbers, while other parameters have a continuous range of possible values. In all such situations the ordinary least square principle, with integer variates fixed to the most probable integer value, can lead to paradoxical results, due to the strong non-linearity of the manifold of admissible values. On the contrary an overall estimation procedure assigning the posterior distribution to all variables, discrete and continuous, conditional to the observed quantities, like the so-called Bayesian approach, has the advantage of weighting correctly the possible errors in choosing different sets of integer values, thus providing a more realistic and stable estimate even of the continuous parameters. In this paper, after a short recall of the basics of Bayesian theory in section 2, we present the natural Bayesian solution to the problem of assessing the estimable signal from noisy observations in section 3 and the Bayesian solution to cycle slips detection and repair for a stream of GPS measurements in section 4. An elementary synthetic example is discussed in section 3 to illustrate the theory presented and more elaborate, though synthetic, examples are discussed in section 4 where realistic streams of GPS observations, with cycle slips, are simulated and then back processed.

Key words *Bayesian theory – prior and posterior probability – integer and continuous variables*

1. Introduction

In geodesy as well as in geophysics there are a number of examples where the unknown parameters are partly constrained to be integer numbers, while other parameters have a continuous range of possible values.

Mailing address: Dr. Giovanna Venuti, DIIAR, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy; e-mail: giove@ipmtf4.topo.polimi.it

(1) Of course a much more complex model could be established with unknown parameters for earth orientation in space, unknown parameters for orbit corrections, unknown parameters for clocks modelling and for ionospheric effects; most of these however can be either determined by ancillary observations or eliminated by suitable differencing the observation equations.

Physically this situation is realised for instance when we observe phases of waves (*e.g.*, electromagnetic waves), which are related to the travelled path, reduced by an integer number of wavelength.

This happens for instance in GPS theory where satellite to receiver observation equations include at least(1) geodetic coordinates of the receiver, tropospheric parameters and so-called integer bias, *i.e.*, the integer number of wavelengths when the satellite signal is locked to the receiver as well as possible cycle slips where the lock is lost and this integer number has a sudden change from one epoch to the next.

The same is true for instance in SAR interferometry where integer numbers are defined on subsets of the interferogram, so that a discontinuity is realised between one zone and the next, and these have to be determined together with the horizontal position and height of the reflectors, *i.e.*, the digital terrain model.

Even more generally, it is possible to model in this way one of the basic problems of signal theory, *i.e.*, the determination of the information content of a certain flow of measurement about a signal expressed on a certain fixed basis of a Hilbert space, by considering the «variable» number of determinable coefficients on this basis as an unknown.

In all such situations the ordinary least square principle, with integer variates fixed to the most probable integer value, can lead to paradoxical results, due to the strong non-linearity of the manifold of admissible values; this has happened for instance in geodetic GPS theory where for a long time the accuracy of the estimation of the unknown coordinates has been grossly overestimated (Betti *et al.*, 1993).

On the contrary an overall estimation procedure assigning the posterior distribution to all variables, discrete and continuous, conditional to the observed quantities, like the so-called Bayesian approach, has the advantage of weighting correctly the possible errors in choosing different sets of integer values, thus providing a more realistic and stable estimate even of the continuous parameters.

In this paper, after a short recall of the basics of Bayesian theory in section 2, we present the natural Bayesian solution to the problem of assessing the estimable signal from noisy observations in section 3 and the Bayesian solution to cycle slips detection and repair for a stream of GPS measurements in section 4.

An elementary synthetic example is discussed in section 3 to illustrate the theory presented and more elaborate, though synthetic, examples are discussed in section 4 where realistic streams of GPS observations, with cycle slips, are simulated and then back processed.

2. Recalls of Bayesian estimation theory

Bayes' theory of elementary probability theory is as follows:

Theorem: given a set Ω with a probability distribution P, given a partition of Ω in non-overlapping measurable sets D_i,

$$\bigcup_{i=1}^{n} D_i = \Omega \qquad D_i \cap D_j = 0 \qquad i \neq j,$$

and given an event A, the following holds:

$$P(D_i \mid A) = \frac{P(A \mid D_i) P(D_i)}{\sum_k P(A \mid D_k) P(D_k)}. \tag{2.1}$$

This formula has the following nice interpretation: if the only knowledge we have about our stochastic system is P, for any A, of $\{D_i\}$ we can indeed compute $P(D_i)$, which are considered the prior information on D_i as well as $P(A \mid D_i)$, namely the probability of sampling in A given D_i; now if we come to know that A is true for our system (additional information or observation), formula (2.1) allows us to compute the posterior probabilities $P(D_i \mid A)$.

When $\{D_i\}$ is just the partition associated with a discrete random variable

$$D_i \equiv \{\omega;\ x(\omega) = x_i \quad i = 1, 2, \ldots, n\} \tag{2.2}$$

and A is also an event associated with an «observable» discrete r.v. Y

$$A \equiv \{\omega;\ Y(\omega) = y_k\} \tag{2.3}$$

we write (2.1) as

$$P(X = x_i \mid Y = y_k) = \\ = \frac{P(Y = y_k \mid X = x_i) P(X = x_i)}{\sum_j P(Y = y_k \mid X = x_j) P(X = x_j)}. \tag{2.4}$$

Formula (2.4) says that observing the value y_k for Y modifies the prior distribution of X, $P(X = x_i)$ into the posterior distribution $P(X = x_i \mid Y = y_k)$.

Let us notice that if $X = x_i$ is thought of as a fixed value of the parameter x, $P(Y = y_k \mid X = x_i)$ can be considered as the probabilistic model of the observation of Y, given $X = x_i$, *i.e.*, the likelihood of the observations given the value of the parameter; for this reason we will also call

$$P(Y = y_k \mid X = x_i) = L(y_k \mid x_i). \quad (2.5)$$

The transition from (2.5) to the case of continuous variables (X, Y), is straightforward, namely, in terms of probability densities one gets

$$p(x \mid y) = \frac{L(y \mid x) p_0(x)}{\int L(y \mid x) p_0(x)\, dx} \quad (2.6)$$

where X and Y can be scalar as well as vector variates and the integral ranges over the whole parameter space.

In (2.6) $p_0(x)$ is the prior distribution of the parameter X, while $p(x \mid y)$ is the posterior distribution, *i.e.*, $p(x \mid y) = f_{X \mid Y}(x \mid y)$, $f_{X \mid Y}(x \mid y)$ indicating the conditional probability density.

All the information on X after observing Y is now contained in $p(x \mid y)$ and of course if we would like to determine representative values we could for instance compute $\hat{x}_B = E_X\{X \mid Y = y\}$ interpreted as a Bayesian estimator, while the variance associated with the distribution $p(x \mid y)$ can be interpreted as the estimation error of $\hat{x}_B$.

An interesting feature of the Bayesian scheme is its ability to represent mixed models where the parameters can be divided into two groups, one of continuous variables X, the other of discrete variables K (where X and K can be both vectors).

If the prior knowledge of X and K is such that we can assume them *a priori* independent

$$p_0(x, k) = p_0(x) p_k \quad (2.7)$$

then (2.6) becomes

$$p(x, k \mid y) = \frac{L(y \mid x, k) p_0(x) p_k}{\sum_j \left[\int L(y \mid x, j) p_0(x)\, dx\right] p_j} \quad (2.8)$$

all the information on X, K is now contained in the function $p(x, k \mid y)$ which we underline to be simultaneously a probability density with respect to X and a probability with respect to K.

If we want to concentrate more on one or the other variable, we have just to switch to the marginal distributions, namely

$$p(x \mid y) = \sum_k p(x, k \mid y) \quad (2.9)$$

$$p(k \mid y) = \int p(x, k \mid y)\, dx. \quad (2.10)$$

Naturally willing to analyze the posterior distribution of K, searching for a representative value, it would be awkward to use the average of $p(k \mid y)$ because this is in general a non-integer value and therefore it has no clear interpretation in terms of variable K. On the contrary, in this case a MAP criterion, *i.e.*, looking for $\bar{k}$ maximizing $p(k \mid y)$, is more suitable and understandable; we stress, however, that only in case a very large probability is concentrated on a single value $\bar{k}$, it is reasonable to take it as representative for K. In this case moreover $1 - p(\bar{k} \mid y)$ represents just the error commetted by fixing $K = \bar{k}$.

3. Maximum information from noisy signal observations

The following situation is standard in signal theory. A «signal» $u(t)$, $t \in R^n$ is considered as a smooth function belonging to some Hilbert space H, endowed with some (normalized but not necessary orthogonal) Ritz basis $\{\varphi_k(t)\}$ such that

$$u(t) = \sum_{k=1}^{\infty} u_k \varphi_k(t), \quad (\| \varphi_k \|_H = 1) \quad (3.1)$$

where u_k are unknown parameters.

We assume to observe the quantities

$$y_i = u(t_i) + v_i \qquad t_i \in R^n, \quad i = 1, 2, \ldots, m \quad (3.2)$$

with v_i independent, equal variance noises. The problem of estimating u_k from y_i is indeed always undetermined, however due to the presence of the noise and to the fact that necessarily

$$\lim_{k\to\infty} u_k = 0, \quad \lim_{N\to\infty} \left\| \sum_{k=N}^{+\infty} u_k \varphi_k \right\| = 0, \qquad (3.3)$$

one way to reduce the problem to an overdetermined form is to assume that the relevant part of the signal is well described by a finite series

$$u(t) = \sum_{k=1}^{N} u_k \varphi_k(t) \; (N \le m) \qquad (3.4)$$

for a suitable value of N.

The choice of the best N, *i.e.*, the highest estimable degree of the coefficients compatible with the observations, represents the maximum information we can draw from $\{y_i\}$ on u, relative to the base $\{\varphi_k\}$.

This problem could be treated in different ways, particularly with the least squares formalism and the related testing procedures, although it finds a very natural formulation in terms of Bayes' theory. So we will consider as unknown parameters of the problem the discrete variable N and the vector of continuous variables $\boldsymbol{X}_{(N)}$

$$\boldsymbol{X}_{(N)} = (u_1, u_2, \ldots, u_N)^{+}. \qquad (3.5)$$

To apply the Bayesian formula we need first of all the priors of $\boldsymbol{X}_N$, N; to make it simple we assume these variables to be *a priori* independent and to have uniform priors(2), *i.e.*

$p_0(\boldsymbol{x}_N) = \text{const}$

$$p_n = \begin{cases} \dfrac{1}{\bar{N}} & n = 1, 2, \ldots, \bar{N} \\ 0 & n > \bar{N} \end{cases} \qquad (3.6)$$

with $\bar{N}$ sufficiently large as to be sure that $P(N > \bar{N})$ be negligible. We observe that in any way as maximum value we have $\bar{N} = m$; in the contrary case the measurements are not sufficient to identify the signal.

Then we need $(\boldsymbol{y} \mid \boldsymbol{x}_N, N)$, which according to the hypotheses done on $\boldsymbol{\nu}$ is given by

$$\boldsymbol{Y} \approx N[\boldsymbol{y}, \sigma_\nu^2 I] \qquad (3.7)$$

$$\boldsymbol{y} = A_N \boldsymbol{x}_N$$

$$A_N = \begin{bmatrix} \varphi_1(t_1) & \ldots & \varphi_N(t_1) \\ \ldots & \ldots & \ldots \\ \varphi_1(t_m) & \ldots & \varphi_N(t_m) \end{bmatrix} (N < m);$$

therefore, we can write

$$L(\boldsymbol{y} \mid \boldsymbol{x}_N, N) = \frac{1}{(2\pi)^{m/2} \sigma_\nu^m} \exp\left[-\frac{1}{2\sigma_\nu^2} \left| \boldsymbol{Y} - A_N \boldsymbol{x}_N \right|^2 \right]. \qquad (3.8)$$

If at this point we realize that we would prefer to consider σ_ν as well as a parameter to be estimated from the observations, we should also introduce a prior hypothesis on the σ_ν variable, assuming for instance that σ_ν is independent of $\boldsymbol{x}_{(N)}$, N and that

$$p_0(\sigma_\nu) = \frac{\text{const}}{\sigma_\nu}. \qquad (3.9)$$

This choice is considered in Bayes' theory as

(2) Since formula (2.8) is homogeneous of degree zero in $p_0(x)$, we can accept for such priors also improper distributions like the uniform distribution in R^n.

non-informative for variables ranging on [0, ∞) (Box Tiao, 1992).

Formula (2.8) then writes

$$p(N, \boldsymbol{x}_N, \sigma_v \mid \boldsymbol{Y}) = \frac{\frac{1}{(2\pi)^{m/2}\sigma_v^m}\exp\left[-\frac{1}{2\sigma_v^2}\mid \boldsymbol{Y}-A_N\boldsymbol{x}_N\mid^2\right]\frac{1}{\bar{N}\sigma_v}}{\frac{1}{(2\pi)^{m/2}}\sum_{n=1}^{\bar{N}}\left(\int_0^{+\infty}\frac{d\sigma_v}{\sigma_v^{m+1}}\int_{R^N}dx_N\exp\left[-\frac{1}{2\sigma_v^2}\mid \boldsymbol{Y}-A_N\boldsymbol{x}_N\mid^2\right]\right)\frac{1}{\bar{N}}}. \tag{3.10}$$

Willing to concentrate on the variable N, we can further write, after some simplifications,

$$p(N \mid \boldsymbol{y}) = \frac{\int_0^{+\infty}\frac{d\sigma_v}{\sigma_v^{m+1}}\int_{R^N}dx_N\exp\left[-\frac{1}{2\sigma_v^2}\mid \boldsymbol{Y}-A_N\boldsymbol{x}_N\mid^2\right]}{\sum_{n=1}^{\bar{N}}\left(\int_0^{+\infty}\frac{d\sigma_v}{\sigma_v^{m+1}}\int_{R^N}dx_N\exp\left[-\frac{1}{2\sigma_v^2}\mid \boldsymbol{Y}-A_N\boldsymbol{x}_N\mid^2\right]\right)}. \tag{3.11}$$

The result of the integration can be expressed as

$$p(N \mid \boldsymbol{y}) = \frac{B_N}{\sum_{n=1}^{\bar{N}} B_n}$$

$$B_n = \frac{(2\pi)^{n/2}\,\Gamma\left(\frac{m-n}{2}\right)}{\sqrt{\det A_n^+A_n}\cdot 2\left(\frac{m-n}{2}\right)^{\frac{m-n}{2}}\hat{\sigma}_{0n}^{m-n}} \tag{3.12}$$

$$\hat{\sigma}_{0n}^2 = \frac{\mid \boldsymbol{y} - A_n\,(A_n^+A_n)^{-1}A_n^+\,\boldsymbol{y}\mid^2}{m-n}$$

and Γ is Euler's function.

As an example we have simulated 10 equispaced observations on the function (signal)

$$u(\vartheta) = 2\cos 2\vartheta + \sin\vartheta$$
$$\vartheta_i = -\pi + i\,\frac{2\pi}{10} \qquad i = 1, 2, \ldots, 10 \tag{3.13}$$

to which a white noise v_i with zero average and $\sigma_v = 0.5$ has been added;

$$y_i = u(\vartheta_i) + v_i. \tag{3.14}$$

Disregarding the Bayesian approach, one could think of interpolating by least squares our function with an increasing value of N; due to the well known orthogonality relations of Discrete Fourier Transform, the coefficients estimated from the model

$$y_i = a_0 + \sum_{i=1}^{N}[a_i\cos(i\vartheta) + b_i\sin(i\vartheta)] + v_i \tag{3.15}$$

stay the same even changing N.

The behaviour of the $u(\vartheta)$ and the estimated (a_i, b_i) are shown in figs. 1 and 2a,b.

As one can see, indeed the coefficients $\hat{a}_2$ and $\hat{b}_1$ are larger than the others, however not to such an extent as to make it obvious (just by looking) which is good and which is not.

Moreover, if we compute $\hat{\sigma}_0^2$ after the adjustment for n = 1, 2, 3, 4 we find the values plotted in fig. 3; here too we see a clear jump at n = 2, but after that it is not obvious that n = 3 or larger should not be accepted.

By using the approach summarized in formula (3.12) however, one obtains the following values for $P(N = n \mid \boldsymbol{y})$:

n	1	2	3	4
$P(N = n \mid \boldsymbol{y})$	0.0004	0.9978	0.0000	0.0020

this corresponds to having chosen $\bar{N} = 4$ in (3.12).

As one can see the distribution P_n points at the value $N = 2$, with an error probability P_e

$$P_e \cong 0.22\%$$

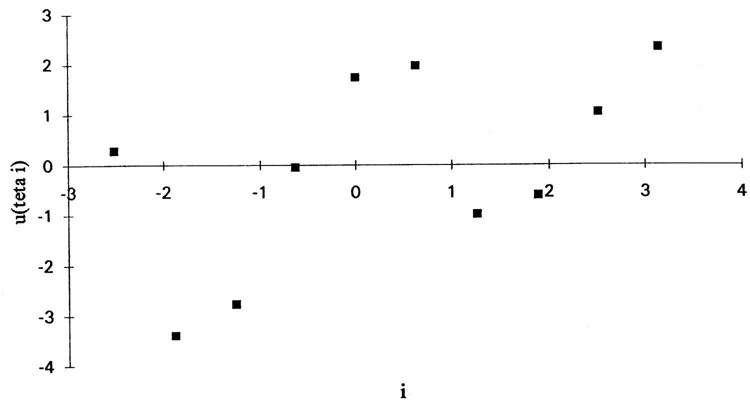

Fig. 1. Equispaced values of the function $u(\vartheta) = 2 \cos 2\vartheta + \sin \vartheta$.

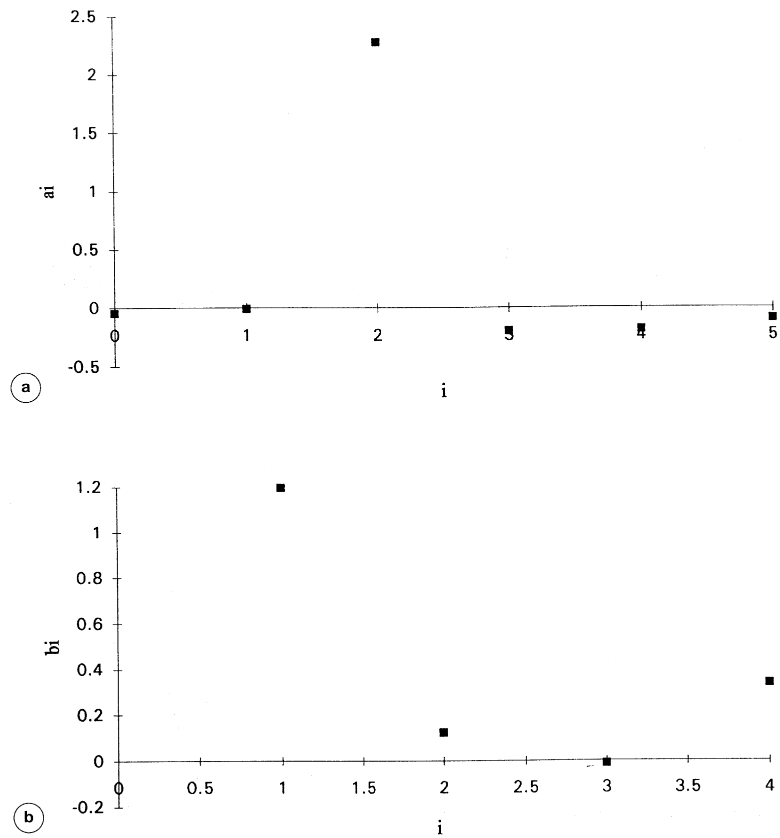

Fig. 2a,b. a) Estimated values of coefficients a_i of Discrete Fourier Transform of $u(\vartheta)$; b) estimated values of coefficients b_i of Discrete Fourier Transform of $u(\vartheta)$.

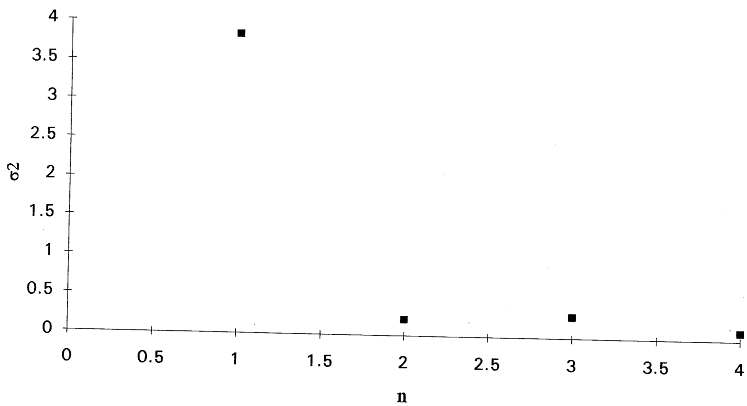

Fig. 3. $\hat{\sigma}_0^2$ for different values of n.

when fixing N at the value 2. So this seems to be a nice confirmation of the theory, although extensive testing of the approach should be performed in the future.

4. Cycle slips detection and identification in GPS phases measurement

In GPS theory the set of code and phase observation equations for a single receiver and a single satellite (Marana, 1994), reads

$$\begin{aligned} R_1(t_k) &= \rho_k + I + v_1 \\ \varphi_1(t_k) &= \rho_k - I + \Lambda_1 N_1 + \eta_1 \\ R_2(t_k) &= \rho_k + \alpha I + v_2 \\ \varphi_2(t_k) &= \rho_k - \alpha I + \Lambda_2 N_2 + \eta_2, \end{aligned} \tag{4.1}$$

where

$R_i(t_k)$ = code readings at time t_k on $L_i (i = 1, 2)$;
$\varphi_i(t_k)$ = phase readings at time t_k on $L_i (i = 1, 2)$;
$\rho_k = \rho(t_k)$ = satellite-receiver range at time t_k;
I = ionospheric correction for L_1;

$$\alpha = \left(\frac{\Lambda_2}{\Lambda_1}\right)^2 \cong 1.65;$$

v = code noise, σ_v = 20 cm;
η = phase noise, σ_η = 0.5 cm;
Λ_i = wavelengths of L_i;
N_i = initial integer ambiguity for φ_i.

From eq. (4.1) a single time solution can be derived, namely

$$\begin{aligned} \rho_k &= \frac{\alpha R_1(t_k) - R_2(t_k)}{\alpha - 1} \\ (I)_k &= \frac{R_2(t_k) - R_1(t_k)}{\alpha - 1} \\ (N_1)_k &= \frac{1}{\Lambda_1}\varphi_1(t_k) - \frac{1}{\Lambda_1}\frac{(1+\alpha)R_1(t_k) - 2R_2(t_k)}{\alpha - 1} \\ (N_2)_k &= \frac{1}{\Lambda_2}\varphi_2(t_k) - \frac{1}{\Lambda_2}\frac{2\alpha R_1(t_k) - (1+\alpha)R_2(t_k)}{\alpha - 1}. \end{aligned} \tag{4.2}$$

Considering the equations for the integer biases only, one can say that through (4.2) one has a stream of values of the type

$$(N)_k = N + \omega_k \qquad k = 1, 2, \ldots, m \tag{4.3}$$

with ω_k a white noise constant variance. This model at least holds as far as the phase observations in the receiver are locked to the incoming wave so that the initial ambiguity is the same for all the observations.

However if lock is lost at sometime $t_{\bar{k}} = \bar{k}\Delta$ (Δ = observating leg) the corresponding phase φ has a jump of height $h\Lambda$. So that the eq. (4.3) has to be substituted by the new model

$$N(t_k) = N + h\vartheta_{\bar{k}}(t_k) + \omega_k$$
$$\vartheta_{\bar{k}}(t_k) = \begin{cases} 1 & k \geq \bar{k} \\ 0 & k < \bar{k} \end{cases}. \tag{4.4}$$

We note that the same model can be written as well

$$y_k = N(t_k) = N\vartheta_0(t_k) + h\vartheta_{\bar{k}}(t_k) + \omega_k$$

or in vector form

$$\boldsymbol{y} = N\boldsymbol{\vartheta}_0 + h\boldsymbol{\vartheta}_{\bar{k}} + \boldsymbol{\omega}. \tag{4.5}$$

Here we have as a matter of fact three discrete parameters, N, $\bar{k}$, h; however for the sake of simplicity of the subsequent computations, we have considered N as continuous variable, that is not wrong but it corresponds to accept a weakening of the model because we do not exploit fully information on this variable.

So we call $x = N$, and we assume that its prior is uniform on R; as for $\bar{k}$ we also assume this variable to be uniformly distributed on the integers between 0 and m (number of epochs); concerning h (the height of the jump) we have considered again a uniform distribution on the integers over a reasonable interval, for instance 0 to 5 since in the simulations we are going to illustrate, we have always taken positive slips smaller than 4. Moreover all the priors have been taken as independent.

Under these hypotheses the posterior distribution of x, $\bar{k}$, h is

$$p(x, \bar{k}, h \mid \boldsymbol{y}) = \frac{L(\boldsymbol{y} \mid x, \bar{k}, h)\, p_0(\bar{k})\, p_0(h)}{p(\boldsymbol{y})} \tag{4.6}$$

where

$$L(\boldsymbol{y} \mid x, \bar{k}, h) =$$

$$\frac{1}{(2\pi)^{m/2}\sigma_\omega^m} \exp\left[-\frac{1}{2\sigma_\omega^2} |\boldsymbol{y} - x\boldsymbol{\vartheta}_0 - h\boldsymbol{\vartheta}_{\bar{k}}|^2\right]$$

$$p(\boldsymbol{y}) =$$

$$\sum_{\bar{k}} \sum_h \left[\int dx\, p_0(x)\, L(\boldsymbol{y} \mid x, \bar{k}, h)\right] p_0(\bar{k})\, p_0(h).$$

Willing to concentrate the attention on the possible epoch of the jump, $\bar{k}$, one can then derive

$$p(\bar{k} \mid \boldsymbol{y}) = \sum_h \int dx\, p(x, \bar{k}, h \mid \boldsymbol{y});$$

this can be explicitly calculated giving as a result (Sansò and Venuti, 1995)

$$p(\bar{k} \mid \boldsymbol{y}) = \frac{\sum_h \exp\left[-\frac{1}{2\sigma_\omega^2} |U_{\bar{k},h}|^2\right]}{\sum_{\bar{k}} \sum_h \exp\left[-\frac{1}{2\sigma_\omega^2} |U_{\bar{k},h}|^2\right]} \tag{4.7}$$

where

$$U_{\bar{k},h} = \boldsymbol{y} - \hat{x}\,\boldsymbol{\vartheta}_0 - h\boldsymbol{\vartheta}_{\bar{k}} \tag{4.8}$$

$$\hat{x} = \frac{1}{m} \sum_{i=1}^{m} y_1 - h\frac{m - \bar{k} + 1}{m} \; (^3) \tag{4.9}$$

Remark: it is important to stress here that, as we did in section 3, it would be possible to include σ_ω in the Bayesian scheme, averaging on its posterior distribution; however, experience has proved that the formulas so derived, though exact, suffer from a numerical instabil-

(3) This value is just the least squares estimate of x (floating) when $\bar{k}$, h are fixed.

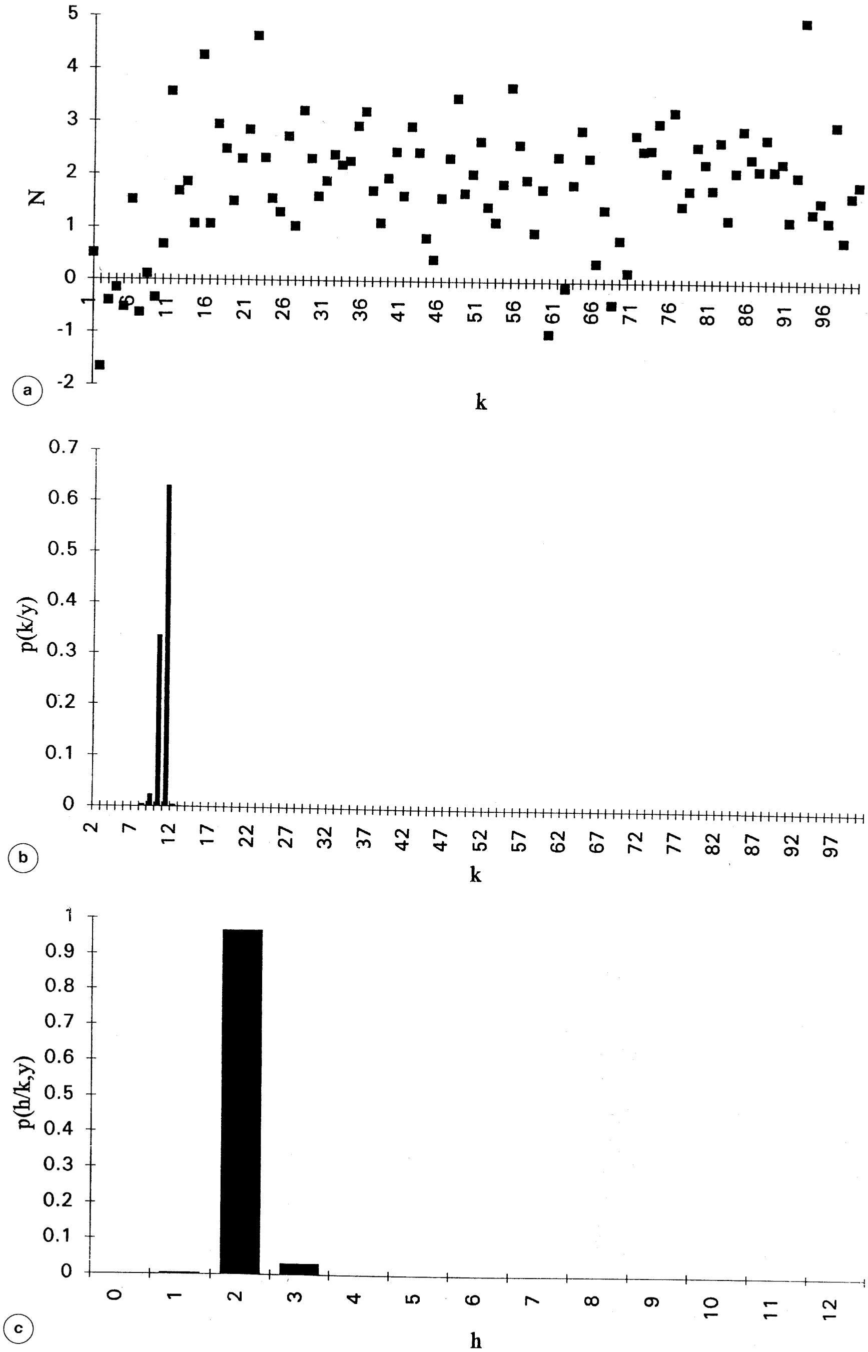

Fig. 4a-c. a) Simulated values of initial integer ambiguity N ($\bar{k} = 11$, $h = 2$); b) posterior distribution of the cycle slip epoch k; c) posterior distribution of cycle slip height h (for $k = 11$).

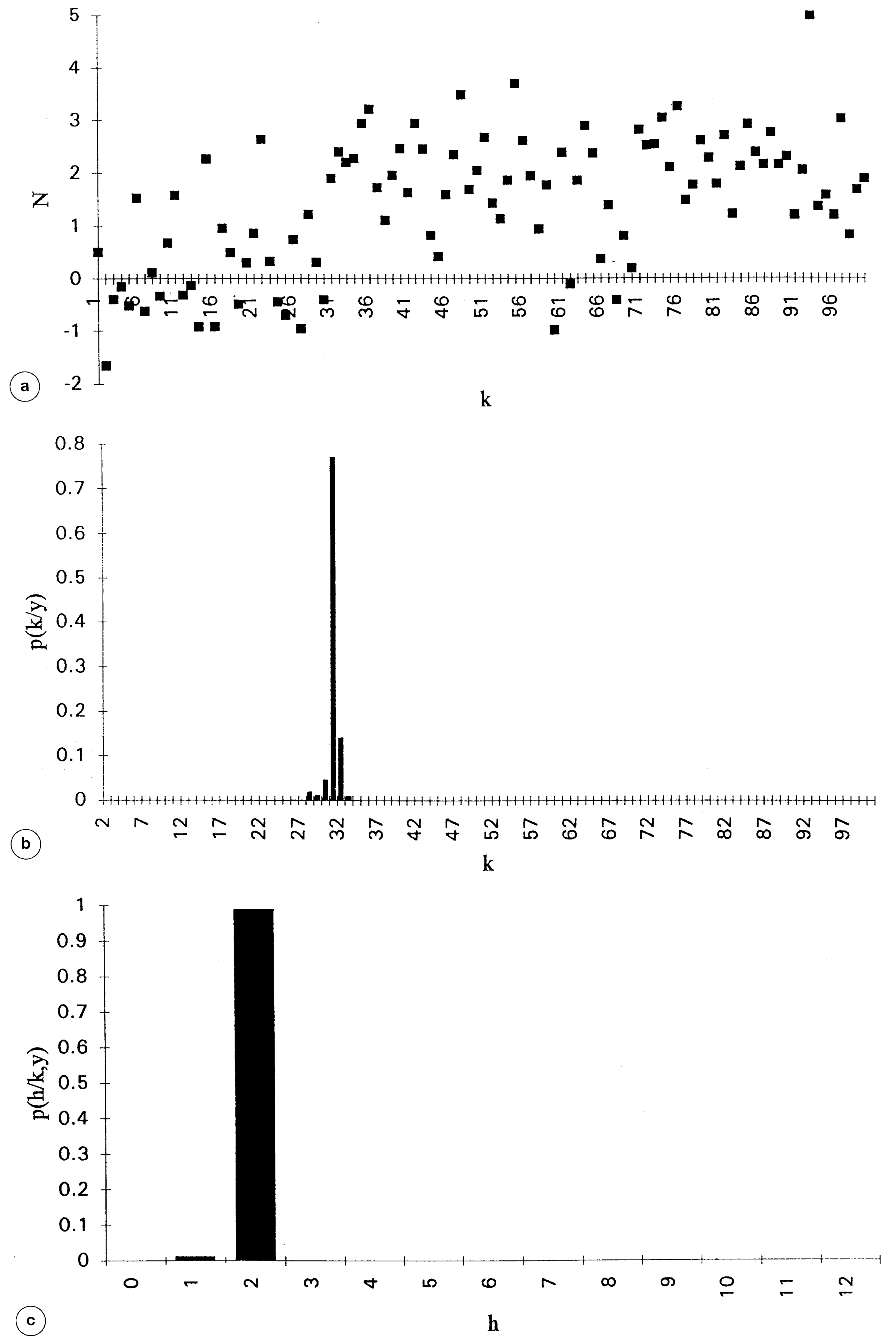

Fig. 5a-c. a) Simulated values of initial integer ambiguity N ($\bar{k}$ = 30, h = 2); b) posterior distribution of the cycle slip epoch k; c) posterior distribution of cycle slip height h (for k = 30).

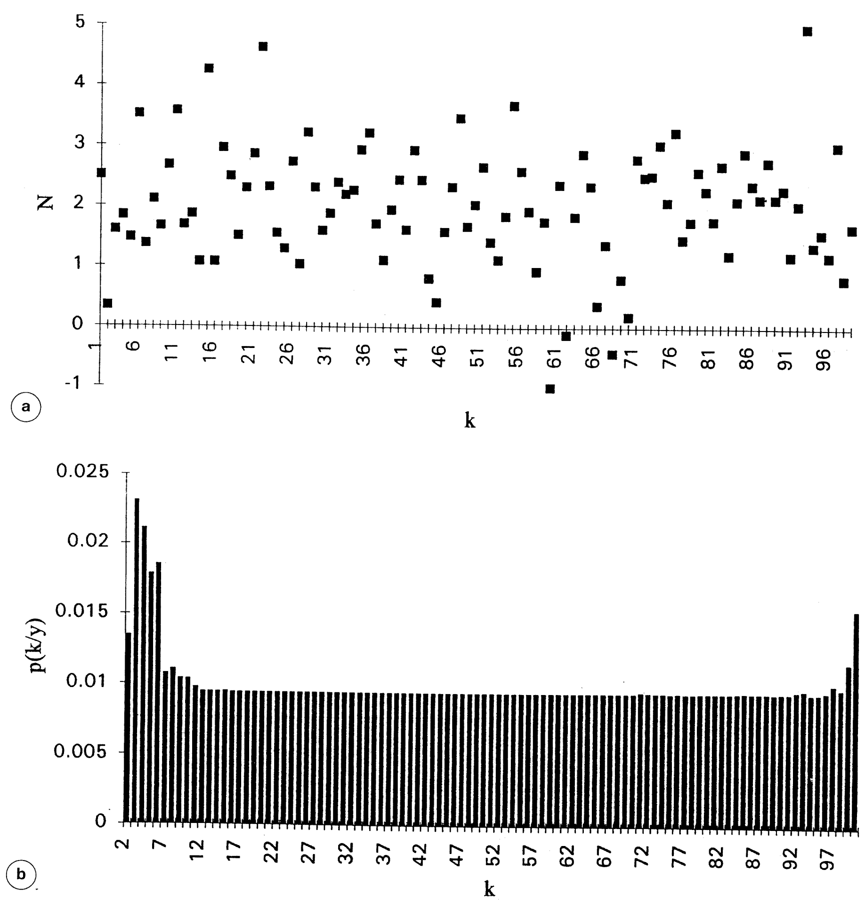

Fig. 6a,b. a) Simulated values of initial integer ambiguity N (no discontinuity); b) posterior distribution of the cycle slip epoch k.

ity. To avoid this we will present examples where σ_ω is fixed as, according to our tests, the result does not depend critically on σ_ω, as far as we fix it at a value smaller than the real one.

So, for instance, in our simulations we will add to the model (4.5) a noise with $\sigma_\omega = 1$ and will perform all the computations with $\sigma_\omega = 0.5$.

Once $\bar{k}$ has been identified one can think of studying the distribution of the variable h; naturally in this case it would look strange to average the probability for the values of h on different jump epochs, so we decided that in this case what is relevant to the problem is the distribution

$$p(h\,|\,\bar{k}, y) = \frac{\exp\left[-\frac{1}{2\sigma_\omega^2}\,|\,U_{\bar{k},h}\,|^2\right]}{\sum_h \exp\left[-\frac{1}{2\sigma_\omega^2}\,|\,U_{\bar{k},h}\,|^2\right]} \qquad (4.10)$$

where, contrary to the case of formulas (4.7)-(4.9), here $\bar{k}$ is fixed to the value previously detected for the jump time.

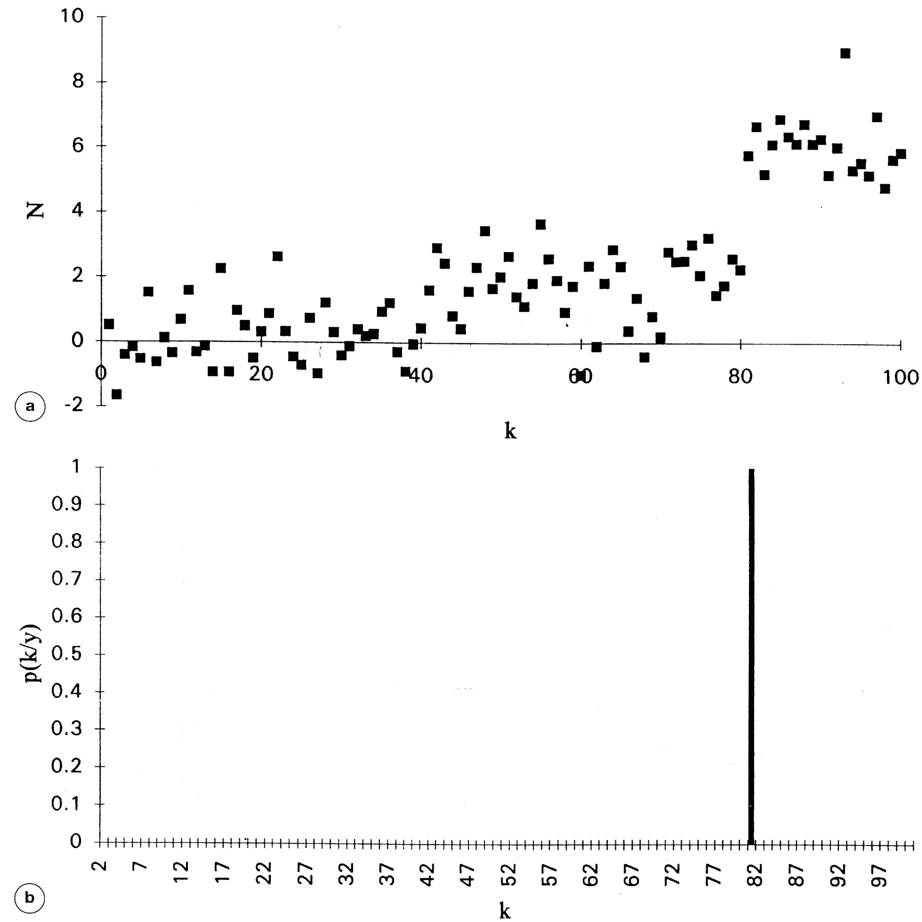

Fig. 7a,b. a) Simulated values of initial integer ambiguity N ($\bar{k} = 41$, $h = 2$; $\bar{k} = 81$, $h = 4$); b) posterior distribution of the cycle slip epoch k.

Results coming from processing simulated data are shown in figs. 4a-c and 5a-c, where jumps of height $\bar{h} = 2$ have been introduced, respectively, at $\bar{k} = 11$ and $\bar{k} = 30$.

As we see in both cases we achieve a correct estimation of the jump epoch as well as of the height; neverthless, in the first case the identification of $\bar{k}$ is more doubtful due to border effects, as is evidenced by the non-smaller error probability.

A case with no discontinuities introduced into the data is shown in fig. 6a,b where the almost uniform distribution of $\bar{k}$ is a clear indication that no jump is found.

At this point we wanted to understand how to treat a more general model with more than one slip. Of course also in this case one can apply the general theory, arriving however at a rather complicated computational scheme; while in the single jump example one has to compute probabilities for each possible jumping epoch (namely m probabilities), in the multiple slips case one has to test all the combinations of the possible discontinuity epochs. So, in order to

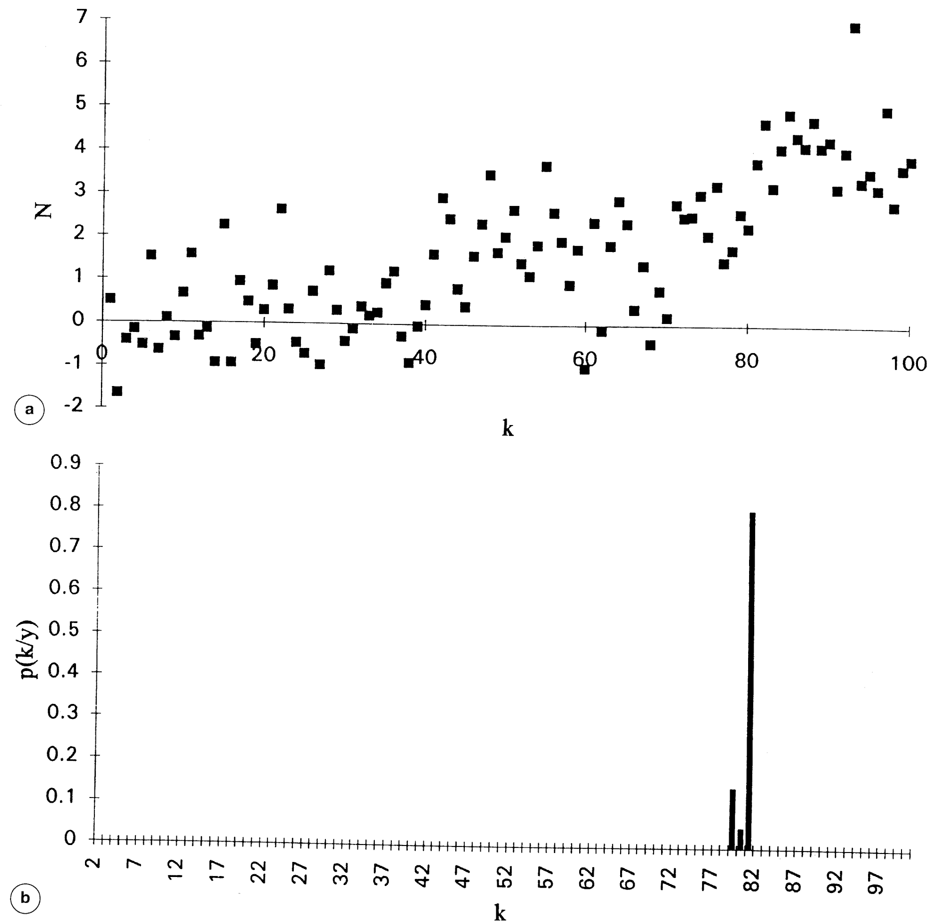

Fig. 8a,b. a) Simulated values of initial integer ambiguity N ($\bar{k} = 41$, $h = 2$; $\bar{k} = 81$, $h = 2$); b) posterior distribution of the cycle slip epoch k.

reduce the number of calculations and simplify the procedures, we decided to test the actual one-jump model to see whether it is possible to implement a sequential detection of the jumps.

Of course once a jump has been found, the data can be split into pieces and the procedure can be repeated.

In the fig. 7a we plot a simulated example with jumps at $\bar{k} = 41$ and $\bar{k} = 81$ of amplitude $h = 2$ and $h = 4$, respectively.

The noise level is always $\sigma_\omega = 1$ and in the computation it has been fixed at $\sigma_\omega = 0.5$ as explained in a previous remark. As one can see (fig. 7b), the method clearly identifies the jump at $\bar{k} = 81$ so that, after this step all the subsequent analysis runs smoothly as in the single jump case.

In fig. 8a,b we have a similar situation, with the amplitude of the two discontinuities equal to $h = 2$ in both cases $\bar{k} = 41$ and $\bar{k} = 81$. Also in this case the method identifies one discontinuity in the right place $\bar{k} = 81$, although indicating a larger uncertainty through a greater estimation error.

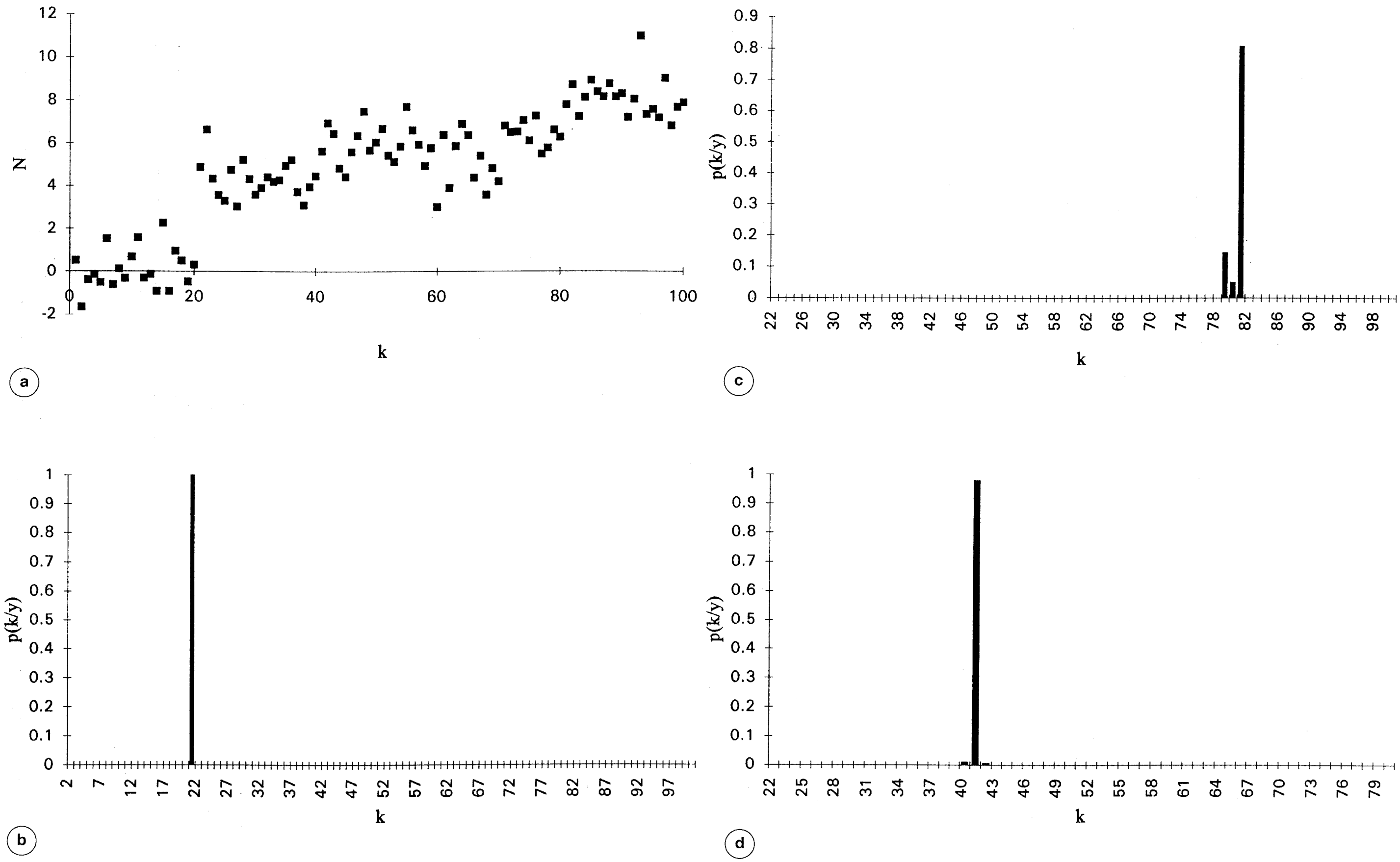

Fig. 9a-d. a) Simulated values of initial integer ambiguity $N(\bar{k} = 21, h = 4; \bar{k} = 41, h = 2; \bar{k} = 81, h = 4)$; b) posterior distribution of the cycle slip epoch k ($k = 0 \div 100$); c) posterior distribution of the cycle slip epoch k ($k = 22 \div 100$); *d* posterior distribution of the cycle slip epoch k ($k = 22 \div 80$).

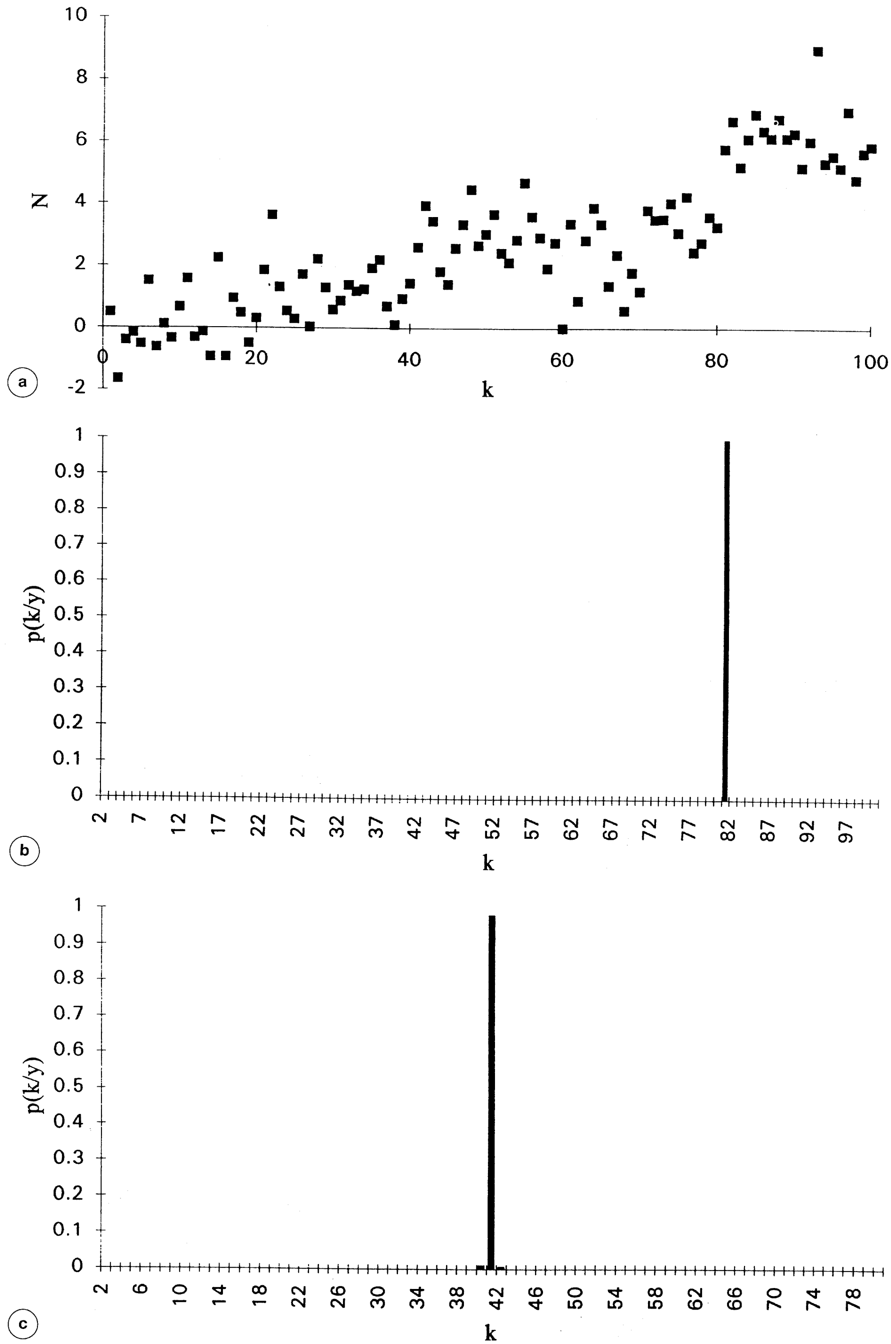

Fig. 10a-c. a) Simulated values of initial integer ambiguity $N(\bar{k} = 21, h = 1; \bar{k} = 41, h = 2; \bar{k} = 81, h = 3)$; b) posterior distribution of the cycle slip epoch k ($k = 0 \div 100$); c) posterior distribution of the cycle slip epoch k ($k = 0.80$).

To gain insight into this procedure we tried to push it to the analysis of data with three jumps. In fig. 9a-d a three jump data set with

$\bar{k}$	21	41	81
h	4	2	2

is presented. The processing, as expected, first identifies the largest discontinuity at $\bar{k} = 21$ and then the other two in sequence $\bar{k} = 81$, $\bar{k} = 41$.

Another example with

$\bar{k}$	21	41	81
h	1	2	3

has been processed and presented in fig. 10a-d

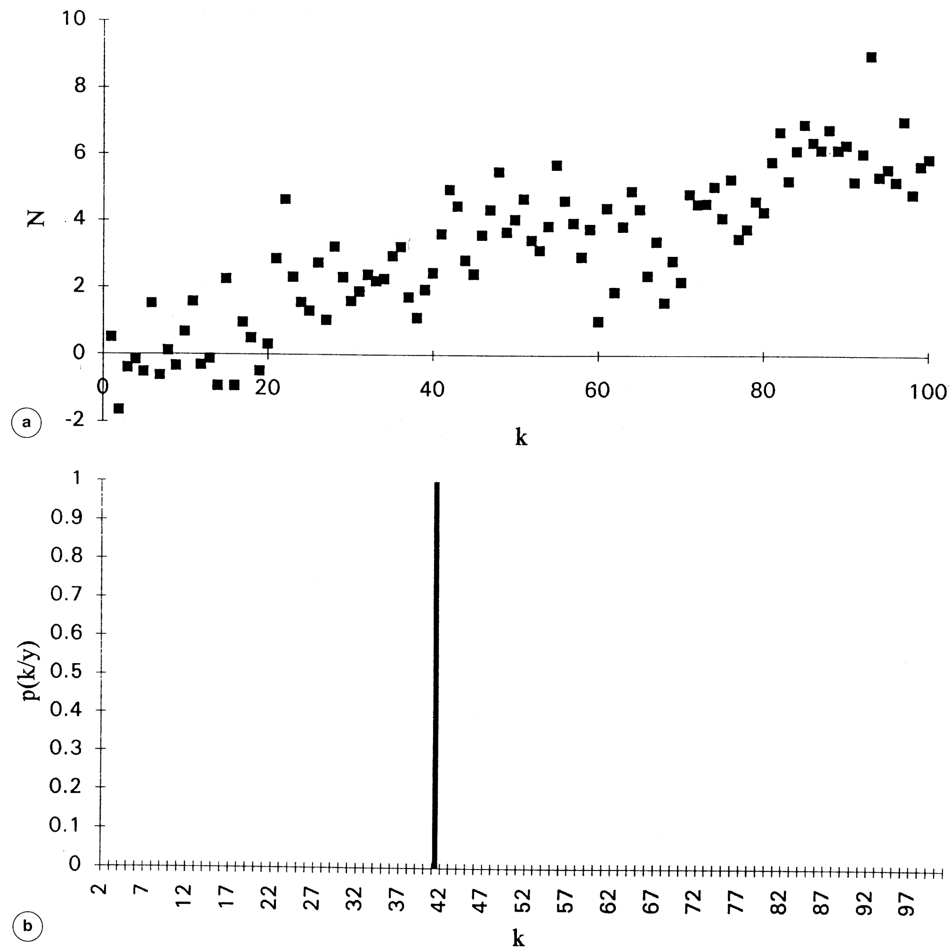

Fig. 11a,b. a) Simulated values of initial integer ambiguity N ($\bar{k} = 21$, $h = 2$; $\bar{k} = 41$, $h = 2$; $\bar{k} = 81$, $h = 2$); b) posterior distribution of the cycle slip epoch k.

where the three discontinuities are detected in the predictable order; here the last distribution of $\bar{k}$ presents a certain spread of probability, though always identifying the right $\bar{k} = 21$ with a larger probability.

As the last example we have tried with three jumps of the same amplitude, (see fig. 11a,b), namely

$\bar{k}$	21	41	81
h	2	2	2

After the processing the time $\bar{k} = 41$ is clearly identified, so that a splitting of the data into two subsets clearly reduces the problem to a single jump case, already discussed.

As a conclusion for the present example apart from some problems which still have to be analyzed more deeply, like the σ_0^2 problem, we feel that the Bayesian approach has proved to be quite reliable and stable in identifying the discontinuity epochs as well as their amplitude; in all cases, even in those where a larger error probability was present, the answer of the method was correct.

Our impression is that this procedure can be adopted quite advantageously to detect cycle slips in GPS data. Naturally our results have to be read on a proper scale; if the noise is increased, as it is almost certain for realistic GPS observations, the identifiable jumps must also be larger to obtain results as good as those presented here.

5. Conclusions

The examples proposed in the paper show that the Bayesian approach, with its ability to treat together continuous and discrete variables, is ideal to discuss and solve different estimation problems where discrete variables are present. The method works reliably in cases where different approaches could be used as well as in cases which are not treated by the ordinary statistical literature (Betti *et al.*, 1993). This is the strength of the method for instance in discussing the GPS data analysis for cycle slip detection. This case, in fact, cannot be reduced to the alternative of hypotheses where the manifold of the admissible values are contained (nested) one into the other.

Many more exciting problems could then be analyzed by the same approach and we expect this to happen in the future.

REFERENCES

BETTI, B., M. CRESPI and F. SANSÒ (1993): A geometric illustration of ambiguity resolution in GPS theory and a Bayesian approach, *Manuscripta Geodaetica*, **18** (5), 317-330.

BETTI, B., M. CRESPI, F. SANSÒ and D. SGUERSO (1993): Discriminant analysis to test not-nested Hypotheses, *Geodetic Theory Today*, edited by F. SANSÒ (Springer), *IAG Symposium*, No. 114, 259-271.

BOX, TIAO (1992): *Bayesian Inference in Statistical Analysis* (Wiley).

MARANA, B. (1994): GPS signals preprocessing, in *Atti del 13° Convegno GNGTS*, 269-280.

SANSÒ, F. and G. VENUTI (1995): Ricerca di discontinuità in segnali immersi nel rumore, in *Atti del 14° Convegno GNGTS*, 229-238.