

PAPER • OPEN ACCESS

Approximate Bayesian Computation in the estimation of the parameters of the Forbush decrease model

To cite this article: A. Wawrzynczak and P. Kopka 2017 *J. Phys.: Conf. Ser.* **936** 012088

View the [article online](#) for updates and enhancements.

You may also like

- [Evidence of Energy and Charge Sign Dependence of the Recovery Time for the 2006 December Forbush Event Measured by the PAMELA Experiment](#)
R. Munini, M. Boezio, A. Bruno et al.
- [The Presence of Turbulent and Ordered Local Structure within the ICME Shock-sheath and Its Contribution to Forbush Decrease](#)
Zubair Shaikh, Anil Raghav and Ankush Bhaskar
- [Quantitative understanding of Forbush decrease drivers based on shock-only and CME-only models using global signature of February 14, 1978 event](#)
Anil Raghav, Ankush Bhaskar, Ajay Lotekar et al.



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Approximate Bayesian Computation in the estimation of the parameters of the Forbush decrease model

A. Wawrzynczak^{1,2}, P. Kopka²

¹Institute of Computer Sciences, Siedlce University, Poland,

²National Centre for Nuclear Research, Poland.

E-mail: awawrzynczak@uph.edu.pl, piotr.kopka@ncbj.gov.pl

Abstract. Realistic modeling of the complicated phenomena as Forbush decrease of the galactic cosmic ray intensity is a quite challenging task. One aspect is a numerical solution of the Fokker-Planck equation in five-dimensional space (three spatial variables, the time and particles energy). The second difficulty arises from a lack of detailed knowledge about the spatial and time profiles of the parameters responsible for the creation of the Forbush decrease. Among these parameters, the central role plays a diffusion coefficient. Assessment of the correctness of the proposed model can be done only by comparison of the model output with the experimental observations of the galactic cosmic ray intensity. We apply the Approximate Bayesian Computation (ABC) methodology to match the Forbush decrease model to experimental data. The ABC method is becoming increasingly exploited for dynamic complex problems in which the likelihood function is costly to compute. The main idea of all ABC methods is to accept samples as an approximate posterior draw if its associated modeled data are close enough to the observed one. In this paper, we present application of the Sequential Monte Carlo Approximate Bayesian Computation algorithm scanning the space of the diffusion coefficient parameters. The proposed algorithm is adopted to create the model of the Forbush decrease observed by the neutron monitors at the Earth in March 2002. The model of the Forbush decrease is based on the stochastic approach to the solution of the Fokker-Planck equation.

1. Introduction and background information

The Forbush decrease (Fd) of galactic cosmic ray (GCR) intensity are identified as a short period decrease (5-12 days) in the stream of the GCR particles reaching the Earth [1]. The Fds occur as an outcome the substantial disturbances in the interplanetary space occurring due to the powerful coronal mass ejecta and solar flares on the Sun. It was shown [2, 3, 4, 5] that Fd of GCR intensity observed by super neutron monitors can provide beneficial information about the structure of the interplanetary magnetic field (IMF) turbulence. In connection with this, the theoretical modeling plays an important role to explain the information obtained from Fds observed by neutron monitors. In general modeling of the GCR transport in the heliosphere is a quite challenging task due to complicated interaction of four main processes: convection by the solar wind, diffusion on irregularities of interplanetary magnetic field (IMF), particle drifts in the non-uniform magnetic field and adiabatic cooling (e.g. [6]). The interplay of these processes represents the Parker transport equation (PTE) [7] being the second order parabolic type partial



differential equation.

To compose the complete model of the Fd is complex problem relating with the complex structure of the disturbed vicinity of the interplanetary space responsible for the Fd of the GCR intensity. To take into account the simultaneous changes of the solar wind velocity, structure of the heliospheric neutral sheet (HNS), regular and turbulence components of the IMF in the restricted disturbed vicinity of the interplanetary space is, basically impossible.

In papers [2, 3, 4, 5] was shown that during the Fd is observed the increase of the exponent ν of the power spectral density (PSD) of the components of the IMF ($PSD \propto f^{-\nu}$, where f is the frequency). From the other side according to the quasi-linear theory [8, 9] the diffusion coefficient K is proportional to the rigidity R of the GCR particles as, $K \propto R^\gamma$, where $\gamma = 2 - \nu$. Thus, during Fd we observe a decrease of the diffusion coefficient due to the increase if the IMF turbulence. However, direct insertion of the exponent ν calculated based on the experimental IMF data is not possible. One reason is that in the model we should consider the spatial distribution of the exponent ν while the IMF data cover only one point of space where the spacecraft recording the IMF is placed. The second is that we observe the increase of ν mainly for the IMF components perpendicular to the field line i.e B_y and B_z , while in the model we should include the overall ν . The third problem is that we cannot calculate the value of ν with arbitrary resolution. The reason is the requirement of employing relatively long data series to attain the values of the exponent ν in low-frequency range $f \in < 10^{-6}, 10^{-5} >$ responsible for the scattering of the GCR particles to which neutron monitor and ground meson telescopes respond. This usually was done (e.g. [2, 3]) based on the comparison of the single data series with length predetermined by the duration of the Fd; i.e. there were calculated the PSDs for three periods, before, during and after the Fd.

In this paper we employ the Approximate Bayesian Computation (ABC) method to compose the model of the Fd best fitted to the Fd recorded by the neutron monitors in the period of 18 March-2 April 2002. (Fig. 1). We assume that Fd occurs due to decrease of the diffusion coefficient K in the vicinity of space where the Fd is observed. The parameter γ responsible for the K decrease will be fitted by the ABC method.

2. Model of the Forbush decrease

We model the transport of GCR in the heliosphere by the stochastic approach. In this approach the individual particle motion is described as a Markov stochastic process, and the system evolves probabilistically. Consequently the Parker equation is brought into a form of the backward Fokker-Planck equation of the form:

$$\frac{\partial f}{\partial t} = A_1 \frac{\partial^2 f}{\partial r^2} + A_2 \frac{\partial^2 f}{\partial \theta^2} + A_3 \frac{\partial^2 f}{\partial \varphi^2} + A_4 \frac{\partial^2 f}{\partial r \partial \theta} + A_5 \frac{\partial^2 f}{\partial r \partial \varphi} + A_6 \frac{\partial^2 f}{\partial \theta \partial \varphi} + A_7 \frac{\partial f}{\partial r} + A_8 \frac{\partial f}{\partial \theta} + A_9 \frac{\partial f}{\partial \varphi} + A_{10} \frac{\partial f}{\partial R} \quad (1)$$

where $f = f(\vec{r}, R, t)$ is an omnidirectional distribution function of three spatial coordinates $\vec{r} = r(r, \theta, \varphi)$, particles rigidity R and time t ; The coefficients A_1, \dots, A_{10} are dependent on the solar wind velocity \vec{U} , the drift velocity components \vec{v}_d , and the 3D anisotropic diffusion tensor K_{ij} of GCR particles [11].

Based on the Ito formulation the corresponding set of the stochastic differential equations (SDE) is solved (e.g., [12]). In the case of the GCR transport the set of SDEs in three dimensional heliocentric spherical space have a form

$$\begin{aligned} dr &= A_7 \cdot dt + [B \cdot dW]_r \\ d\theta &= A_8 \cdot dt + [B \cdot dW]_\theta \\ d\varphi &= A_9 \cdot dt + [B \cdot dW]_\varphi \\ dR &= A_{10} \cdot dt, \end{aligned} \quad (2)$$

where \vec{r} is the trajectory of individual pseudoparticle in the phase space and dW_i is the Wiener process. All coefficients and the details of the numerical solution of the set Eq. 2 are given in [13, 14, 15].

In the proposed model we assume that Fd takes place due to established corotating heliolongitudinal disturbances in the interplanetary space. In this region the IMF turbulence is increased resulting in a gradual diminishing of the diffusion at the Earth orbit, causing larger scattering of the GCR particles, and in effect, fewer GCR particles reach the Earth (e.g., [3, 5]). We simulate this process by the gradual decrease and then the increase of the diffusion coefficient at the Earth orbit with respect the heliolongitudes. The diffusion coefficient K_{II} of cosmic ray particles has a form:

$$\begin{aligned} K_{II} &= K_0 \cdot K(r) \cdot K(R, \gamma) \\ K(r) &= 1 + 0.5 \cdot (r/1AU) \\ K(R, \gamma) &= R^\gamma \\ \gamma &= \begin{cases} \alpha + \beta \sin(\varphi - \zeta); & \text{for } (r < 30AU) \text{ and } (90^\circ \leq \varphi \leq 270^\circ), \\ \alpha; & \text{otherwise} \end{cases} \end{aligned} \quad (3)$$

The parameters α , β and ζ values are selected by the ABC method so that the Fd model matches the observational amplitude of Fd in March 2002 (see Fig. 1).

3. Approximate Bayesian Computation algorithm setup

Let λ be a parameter vector, given the prior distribution $\pi(\lambda)$. The goal of Bayesian inference is to approximate the posterior distribution, $\pi(\lambda|d_{obs}) \propto \pi(d_{obs}|\lambda)\pi(\lambda)$, where $\pi(d_{obs}|\lambda)$ is the likelihood of λ given the observed data d_{obs} . The main idea of ABC methods is to accept λ as an approximate posterior draw if its associate (modeled) data d is close enough to the observed data d_{obs} . Accepted parameters are a sample from $\pi(\lambda|\rho(d, d_{obs}) < \epsilon)$ where the $\rho(d, d_{obs})$ is the chosen measure of discrepancy, and ϵ is a threshold defining 'closeness margin'. If ϵ is sufficiently small then the distribution $\pi(\lambda|\rho(d, d_{obs}) < \epsilon)$ will be a good approximation for the posterior distribution $\pi(\lambda|d_{obs})$. It is often difficult to define an adequate distance function $\rho(d, d_{obs})$ between the simulated and observed data. So, in many cases it is replaced with a distance defined by summary statistics, $\rho(S(d), S(d_{obs}))$. However, as here we consider the Fd amplitudes in specific time points, we can compare those data directly without a use of summary statistics.

In ABC methods, Sequential Monte Carlo (SMC) is used to automatically, sequentially 'clean' approximation of posterior distribution to be used to generate proposals for further steps. In ABC SMC method, the set of samples with weights, called particles sampled from the prior distribution $\pi(\lambda^1)$, are propagated through a sequence of intermediate posterior distributions $\pi(\lambda^t|\rho(d, d_{obs}) < \epsilon_t)$, $t = 1, \dots, T$, until it represents a sample from the target distribution, $\pi(\lambda^T|\rho(d, d_{obs}) < \epsilon_T)$. These methods aim to generate draws from $p(\lambda^t|\rho(d, d_{obs}) < \epsilon_t)$, at each of a series of sequential steps t , where ϵ_t define a series of thresholds. In [16] the authors proposed strategies called ABC SMC with Adaptive Weights (ABC SMC AW). In the first step algorithm 1 initializes the threshold schedule $\epsilon_1 > \epsilon_2 > \dots > \epsilon_T$. Then N samples are simulated based on the predefined a priori distribution $\pi(\lambda^1)$ and the corresponding acceptance condition $\rho(d, d_{obs}) < \epsilon_1$. Next, the initial uniform weights are calculated. Samples, denoted by a tilde are drawn from the previous generation with probabilities w_j^{t-1} . Using perturbation kernel $K_{\lambda, t}(\lambda_i^t | \tilde{\lambda}_i)$ new 'fresh' samples λ_i^t are obtained, with the veracity of the condition $\rho(d, d_{obs}) < \epsilon_t$. The weights are calculated according to the formula in stage (11); in stage (12) the weights are normalized and the time step is increased $t = t + 1$. This procedure is repeated until $t \leq T$.

In this paper the scanned parameters vector λ is

$$\lambda \equiv (\alpha, \beta, \zeta). \quad (4)$$

with the following priori distribution on particular parameters:

$$\begin{aligned} \alpha &\sim U(0.7, 1.5), \\ \pi(\lambda^1) &\equiv \begin{cases} \beta \sim U(0.1, 0.5) \\ \zeta \sim U(70, 100) \end{cases} \end{aligned}$$

To estimate the fit of the Fd model to the observations as a distance measure we used normalized approximation error between all the data. This measure is called Fractional Bias (FB) is used to indicate

Algorithm 1 ABC SMC AW

```

1. Initialize threshold schedule  $\epsilon_1 > \epsilon_2 > \dots > \epsilon_T$ 
2. Set  $t = 1$ 
for  $i = 1$  to  $N$  do
  3. Simulate  $\lambda_i^t \sim \pi(\lambda^t)$  and  $d \sim \pi(d|\lambda_i^t)$ 
  4. Until  $\rho(d, d_{obs}) < \epsilon_t$ 
  5. Set  $w_i^t = \frac{1}{N}$ 
end for
for  $t = 2$  to  $T$  do
  for  $i = 1$  to  $N$  do
    8. Pick  $\tilde{\lambda}_i$  from the set  $\{\lambda_j^{t-1}\}_{1 \leq j \leq N}$ 
      with probabilities  $\{v_j^{t-1}\}_{1 \leq j \leq N}$ 
    9. Draw  $\lambda_i^t \sim K_{\lambda,t}(\lambda_i^t|\tilde{\lambda}_i)$  and  $d \sim \pi(d|\lambda_i^t)$ 
    10. Until  $\rho(d, d_{obs}) < \epsilon_t$ 
    11. Compute new weights as
      
$$w_i^t \propto \frac{\pi(\lambda_i^t)}{\sum_j v_j^{t-1} K_{\lambda,t}(\lambda_i^t|\lambda_j^{(t-1)})}$$

    12. Normalize weights  $w_i^t$  for  $i = 1, \dots, N$ 
  end for
end for

```

a bias towards underprediction or overprediction of data by the model. We have slightly modified the classical FB measure $\rho(d, d_{obs})$ as:

$$\rho(d, d_{obs}) = \frac{1}{days} \sum_{i=1}^{days} \frac{|d_i - \hat{d}_i|}{|d_i + \hat{d}_i|}, \quad (5)$$

under additional definition, that $\frac{|d_i - \hat{d}_i|}{|d_i + \hat{d}_i|} = 0$ when $d_i = 0$ and $\hat{d}_i = 0$.

The most commonly used adaptive scheme for threshold choice is based on the quantile of the empirical distribution of the distances between the simulated data and observations (e.g. [18]). This method determines ϵ_t in the t time-iteration by sorting the measure $\{\rho(d_{\theta_i}^{t-1}, d_{obs})\}_{1 \leq i \leq N}$ and setting ϵ_t such that predetermined percent α_t of the simulated data fulfill the condition $\rho(d, d_{obs}) < \epsilon_t$. The basic problem with this approach is that we are not able to estimate the value of the measure $\rho(d, d_{obs})$ based on the values obtained in iteration $t - 1$. In this paper, we propose an efficient adaptive mechanism for setting the α_t . The value of the quantile is judged based on of the previous stage $t - 1$, but with three major improvements:

- The distribution of $\{\rho_{\theta_i}^{t-1}\}_{i=1, \dots, N}$ measure of samples set is not discrete. To keep the procedure flexible and support the interval $[0, 1]$ the best choice for continuous distributions seems to be the family of *Beta* distributions.
- The set $\{\rho_{\theta_i}^{t-1}\}_{i=1, \dots, N}$ is extended by adding the values of $\rho(d, d_{obs})$ calculated in step t regardless whether the samples have been accepted or not.
- We propose to define the value of α_t to be dependent on the ratio between a number of currently accepted samples (AS) and the number of all hitherto generated samples (AGS) i.e.: $\alpha_t = 1 - AS/AGS$.

We chose transition kernel $K_{\lambda,t}(\lambda_i^t|\tilde{\lambda}_i)$ to be a three dimensional Gaussian kernel (e.g. [17]). The set of samples from $t - 1$ is used to construct a kernel $K_{\lambda,t}(\lambda_i^t|\tilde{\lambda}_i)$, which will describe the probability mass. Additionally only samples keeping the parameter γ in a physically acceptable interval i.e., $0 < \gamma < 2$ are accepted.

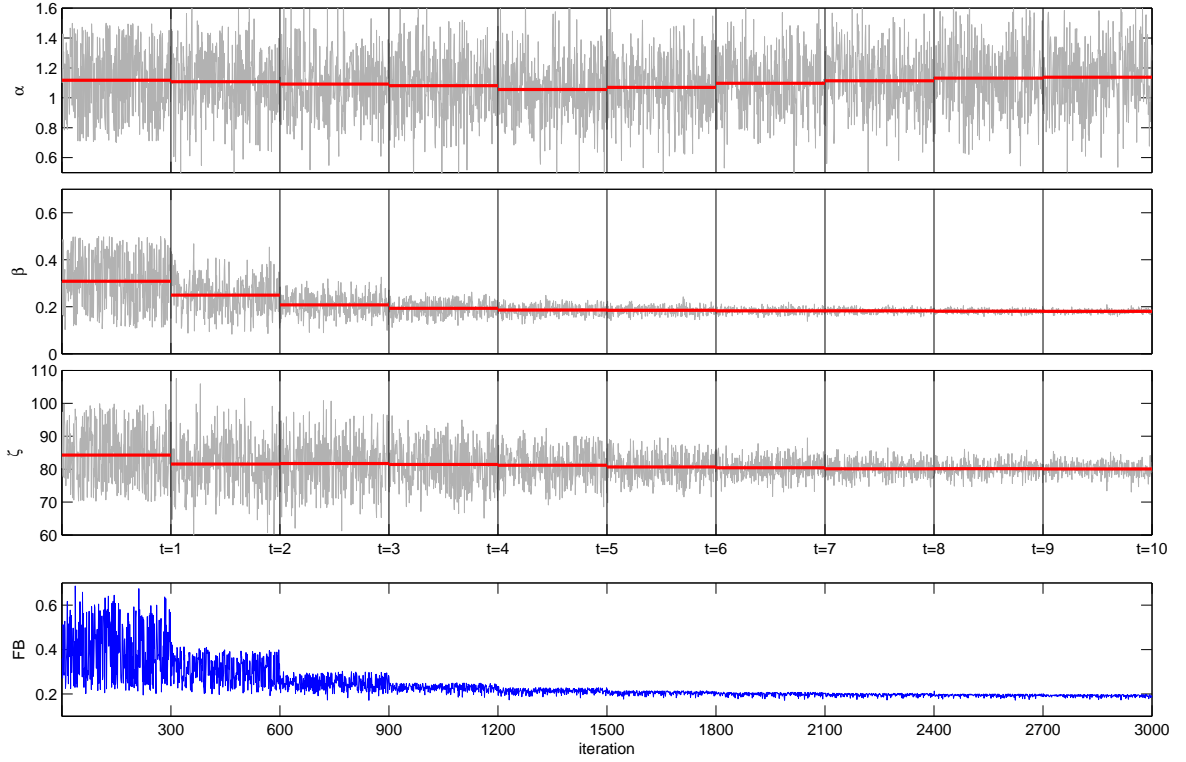


Figure 1. The trace plots for all searched parameters $\lambda \equiv (\alpha, \beta, \zeta)$ in subsequent iterations. Red line marks the mean value within 300 iterations; gray color represents simple trace plot of all samples. The bottom panel presents the FB value for all the accepted samples shown in top three panels.

4. The simulations results and discussion

We have employed the ABC SMC with Adaptive Weights method described in section 3 to find the diffusion coefficient parameters $\lambda \equiv (\alpha, \beta, \zeta)$ matching the Fd observed in March 2002 (Fig. 3). The code for the numerical realization of the described algorithm is realized in Julia environment [19]. The model of GCR stochastic transport in the heliosphere described in section 2 is easy to parallelize versus the number of simulated pseudoparticles using Distributed Arrays Package. This approach fallouts from the assumption that any random process is independent of the other realization, accordingly each pseudoparticles trajectory is independent on another. We performed the simulations applying the ABC SMC AW method setup described in section 3. Each calculation of the FB (Eq. 5) value requires one run of the model of stochastic transport of GCR in the heliosphere represented by the SDEs (Eq. 2). In a single run, the trajectories of $N = 3000$ pseudoparticles were traced backward in time in the spherical heliocentric coordinate system. The pseudoparticles with the rigidity $R=10$ GV were initialized in the point representing the Earth's orbit and traced backward in time until crossing the heliosphere boundary assumed at 100 AU. Details of the numerical solution of Eqs.2 are given in [13].

Fig. 1 illustrates the trace plots for all searched parameters in subsequent stages of the algorithm. The red line marks the mean value within 300 iterations. One can see, that with iterations the samples are more focused on the average value. This is the best visible in the case of β and ζ parameters. These parameters influence the most the amplitude of Fd. From the other side, we can see that the estimation of the parameter λ is not so unequivocal. The variance of this parameter is the largest. The bottom panel in Fig.1 presents the values of FB for accepted samples throughout the scanning procedure. One can see, that the FB value changes significantly in subsequent stages. In the initial iterations of given time, we observe the meaningful change of FB value. This is connected with adapting the threshold to

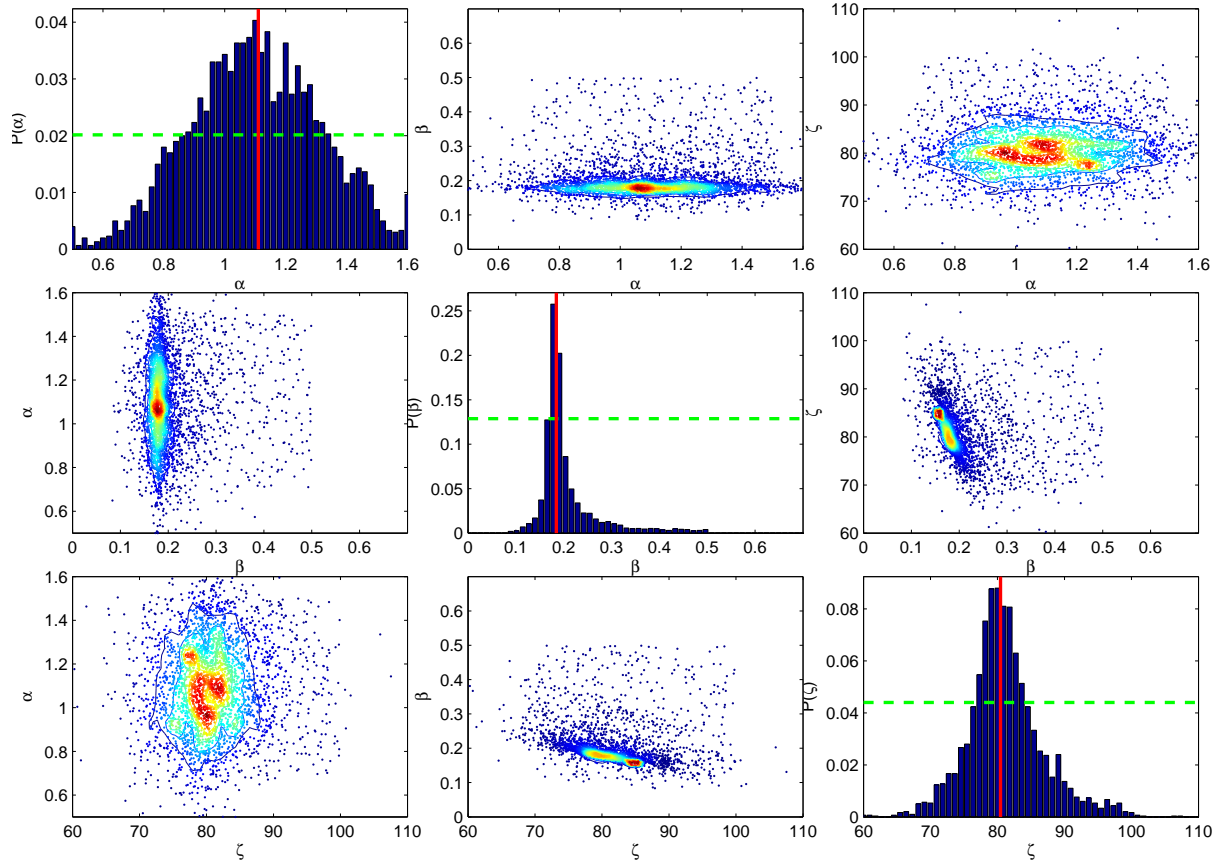


Figure 2. The bivariate and marginal posterior distributions for all searched parameters $\lambda \equiv (\alpha, \beta, \zeta)$. Probability density colors the plot, the more red regions, the higher is its probability. The red vertical lines in diagonal plots mark the value of the parameter with the higher probability. The green horizontal line cuts the value 50% of maximal probability.

the knowledge about the Fd amplitude, i.e., the value ϵ_t is modified according to the Acceptance Ratio (AR). Starting from 1500 iteration, no significant improvement in FB value is observed. In result, ABC SMC AW algorithm accepts only samples with FB value lower than one estimated in previous stages. This reveals the strength of the proposed procedure, which keeps the Acceptance Ratio at the appropriate level in all stages.

The summarized results of reconstruction procedure are presented in Fig.2 as a trellis plot. Recall that a color pattern reflected in the Fig.2 corresponds to the empirical 2D probability distribution of all parameter combinations. The colored contour lines are enveloping higher probability of the joint posterior distributions. The diagonal plots are marginal empirical posterior distributions of the parameters (α, β, ζ) .

The estimated as most probable values of parameters are highlighted with a vertical red line in diagonal subplots. The most probable values are: $P(\alpha = 1.11 \pm 0.01) = 0.0404$, $P(\beta = 0.1845 \pm 0.00635) = 0.2577$, $P(\zeta = 80.4545 \pm 0.4545) = 0.0880$. The green horizontal line cuts the half of the maximal probability distribution for each parameter, which we consider as a high probability region. Thus, as the values with high probability for each considered parameter we consider the interval $\alpha \in < 0.87, 1.35 >$, $\beta \in < 0.1718, 0.1973 >$ and $\zeta \in < 76.8182, 84.0909 >$.

We have incorporated the estimated parameters of diffusion coefficient into the stochastic model of the Fd. The expected amplitudes of the GCR intensity for the rigidity of 10 GV in comparison with the profile of the daily GCR intensity recorded by Apatity and Moscow neutron monitor during 18 March - 2 April presents Fig. 3. The shaded region denotes the amplitudes of the Fd obtained for the border

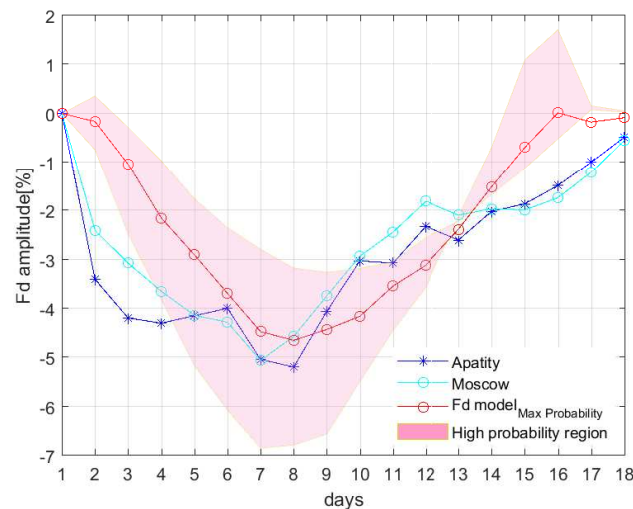


Figure 3. Temporal changes of the observed daily amplitudes of Fd in 18-March - 2 April 2002 for Apatity and Moscow neutron monitors and expected amplitude modeled with use of the parameters vector λ denoted by ABC SMC AW algorithm as the most probable. The shady region marks the expected amplitude for the minimal/maximal parameters indicated by ABC SMC AW with at least 50% of the maximal probability.

values of the high probability region of parameters. One can see that proposed model reconstructs the temporal changes of the GCR intensity observed in March 2002. However, in some periods, the model results do not coincide one to one with the experiment. The reasons are the assumed simplifications. The discrepancy results from the assumption that the Fd is modeled only by the change in diffusion coefficient, while the solar wind velocity was assumed as constant. However, one can see that proposed ABC SMC AW method in conjunction with the stochastic model of GCR transport was able to confirm that as the main cause of the Fd observed in March 2002 we can consider the decrease of the exponent γ causing the change of the diffusion coefficient. These results coincide with the experimental data for Fd in March 2002 presented in [2].

The obtained results confirm that the ABC methodology can be successfully applied to judge about the spatial and temporal changes of the parameters responsible for GCR modulation in the heliosphere. The future work will include testing other FB measures and summary statistics. The energy dependence of the diffusion coefficient and the spatial dependence of the drift velocity will also be fitted.

5. Acknowledgments

This work is supported by The Polish National Science Centre grant awarded by decision number DEC-2012/07/D/ST6/02488. Calculations were performed at the Interdisciplinary Centre for Mathematical and Computational Modelling (ICM) at Warsaw University within the computational grant no. G66-19. We thank the principal investigators of Apatity and Moscow neutron monitor for the ability to use their data.

6. References

- [1] Forbush S E 1954 *J. Geophys. Res.* **59** 525
- [2] Wawrzynczak A and Alania M V 2008 *Adv. Space Res.* **41** 325-334
- [3] Wawrzynczak A and Alania M V 2010 *Adv. Space Res.* **45** 622-631
- [4] Alania M V and Wawrzynczak A 2012 *Adv. Space Res.* **50** 6 725-730
- [5] Wawrzynczak A and Alania M V 2015 *J. of Phys.: Conference Series* **632** 1742-6596
- [6] Moraal H 2013 *Space Science Reviews* **176** 299-319
- [7] Parker E N 1965 *Planetary and Space Science* **13** 9-49

- [8] Jokipii J R 1971 *Rev. Geophys. Space Phys.* **9** 21
- [9] Toptygin I N 1985 *Cosmic rays in interplanetary magnetic fields*, Reidel Publishing Company, Dordrecht
- [10] Potgieter M S 2013 *Space Science Reviews* **176** 165-176
- [11] Alania M V 2002 *Acta Physica Pol. B* **33** 4 1149-1166
- [12] Gardiner C W 2009 *Handbook of stochastic methods. For physics, chemistry and the natural sciences* Springer Series in Synergetics
- [13] Wawrzynczak A, Modzelewska R and Gil A 2015 *J. of Phys.: Conference Series* **574** 012078
- [14] Wawrzynczak A, Modzelewska R and Kluczek M 2015 *J. of Phys.: Conference Series* **633** 012058
- [15] Wawrzynczak A and Modzelewska R *J. of Phys.: Conference Series* **738** 1
- [16] Bonassi F V, West M et al. 2015 *Bayesian Analysis* **10** 1 171187.
- [17] Filippi S, et.al. 2013 *Statistical applications in genetics and molecular biology* **12** 1 87107.
- [18] Beaumont M A, Zhang W and Balding D J 2002 *Genetics* **162** 4 20252035.
- [19] <https://julialang.org/>