Erratum: Statistical validation of event predictors: A comparative study based on the field of seizure prediction [Phys. Rev. E 83, 066704 (2011)]

Hinnerk Feldwisch-Drentrup,* Andreas Schulze-Bonhage, Jens Timmer, and Björn Schelter

(Received 30 August 2011; published 20 October 2011)

DOI: 10.1103/PhysRevE.84.049902 PACS number(s): 05.10.−a, 07.05.Tp, 87.19.xm, 91.30.pd, 99.10.Cd

In this paper we have compared methods for the statistical validation of prediction performances. We present corrected figures of both the empirical size (Fig. 1) and the statistical power (Fig. 2). The results and conclusions of our original paper remain qualitatively unchanged. Quantitative changes regard in particular the alarm time surrogate (ATS), for which a lower empirical size is observed in comparison to the original results (Fig. 1), notably also for larger r_{FA}. In parallel to the decreased empirical size, the power of the ATS is accordingly decreased (cf. Fig. 2). For seizure times surrogates with random offset (STS) or with bootstrap sampling (BST), the effective size is decreased, in particular for the case of a normalization parameter $\gamma_0 = 0.1/\mathrm{h}$ [Fig. 1(h)], for which the influence of false predictions on the performance measure Π_{M} is increased compared to the default value of $\gamma_0 = 1/\mathrm{h}$. Here it complies to the chosen significance level for the STS. Still an empirical size larger than the chosen significance level, $\alpha = 5\%$ is observed for the ATS and BST for short simulation durations including few events and alarms. An invalid effective size is observed for all numerical methods for γ-distributed interseizure intervals with shape parameter $k = 0.2$ [Fig. 1(f)], resembling clustered events.

*hinnerk.feldwisch@bcf.uni-freiburg.de

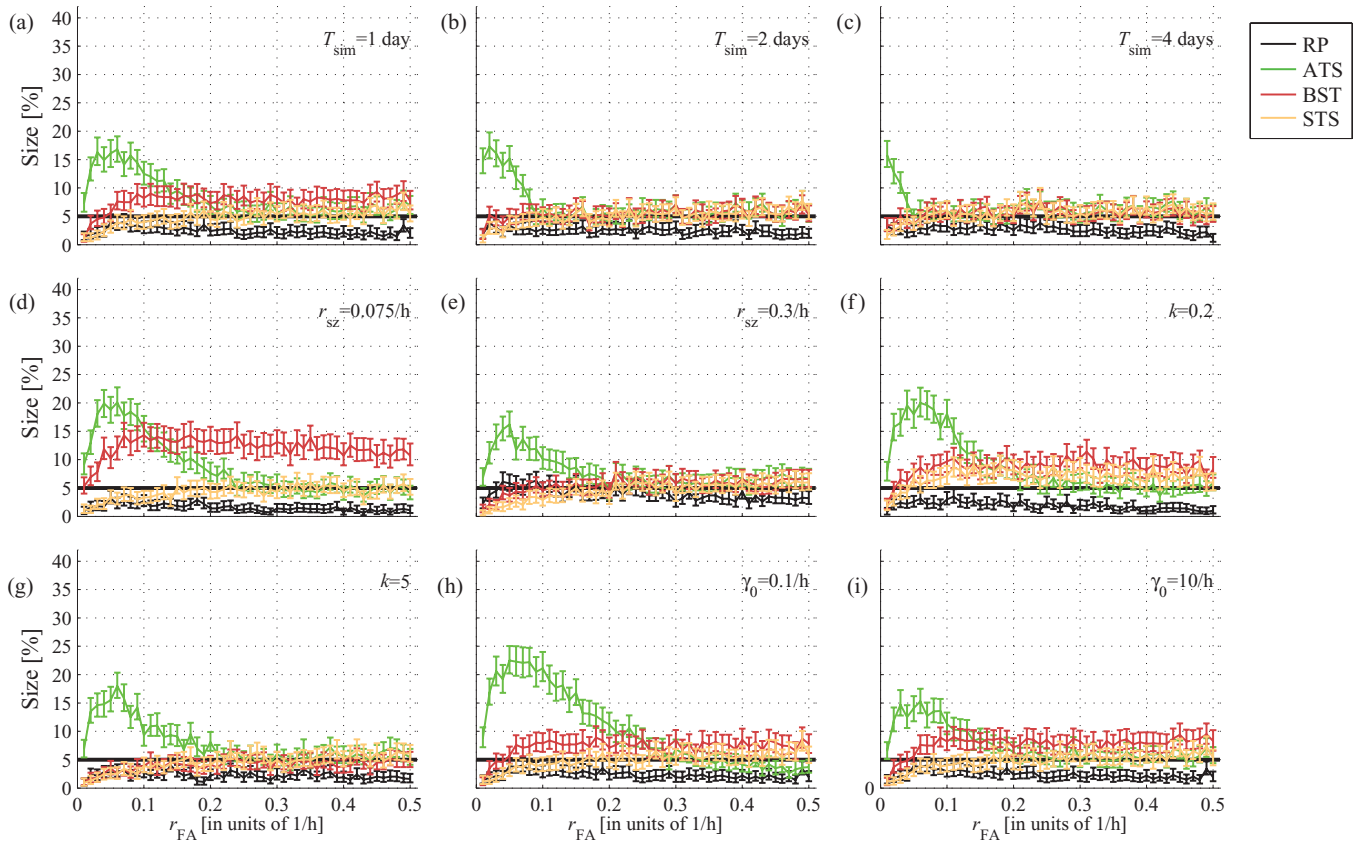

FIG. 1. (Color online) Empirical size of the analytical random predictor (RP, black), the alarm time surrogates (ATS, green or medium gray), and seizure time surrogates with random offset (STS, orange or light gray) and with bootstrap resampling (BST, red or dark gray), depending on the rate of false alarms r_{FA}, for $\alpha = 5\,\%$ (black horizontal lines). For exponentially distributed interseizure intervals, the results are shown for varying simulation durations and fixed seizure rate $r_{\mathrm{sz}} = 0.15/\mathrm{h}$ and $\gamma_0 = 1/\mathrm{h}$ in (a)–(c), and for varying r_{sz} and fixed simulation duration $T_{\mathrm{sim}} = 1$ day and $\gamma_0 = 1/\mathrm{h}$ in (d) and (e). In (f) and (g), the empirical size is shown for γ-distributed interseizure intervals with shape parameter $k = 0.2$ and $k = 5$ for $r_{\mathrm{sz}} = 0.15/\mathrm{h}$ and $T_{\mathrm{sim}} = 1$ day. In (h) and (i), the normalization factor γ_0 of the performance measure Π_{M} was varied, again for $r_{\mathrm{sz}} = 0.15/\mathrm{h}$ and $T_{\mathrm{sim}} = 1$ day, for $k = 1$. Based on 1000 simulation instances, 95% confidence intervals are given.

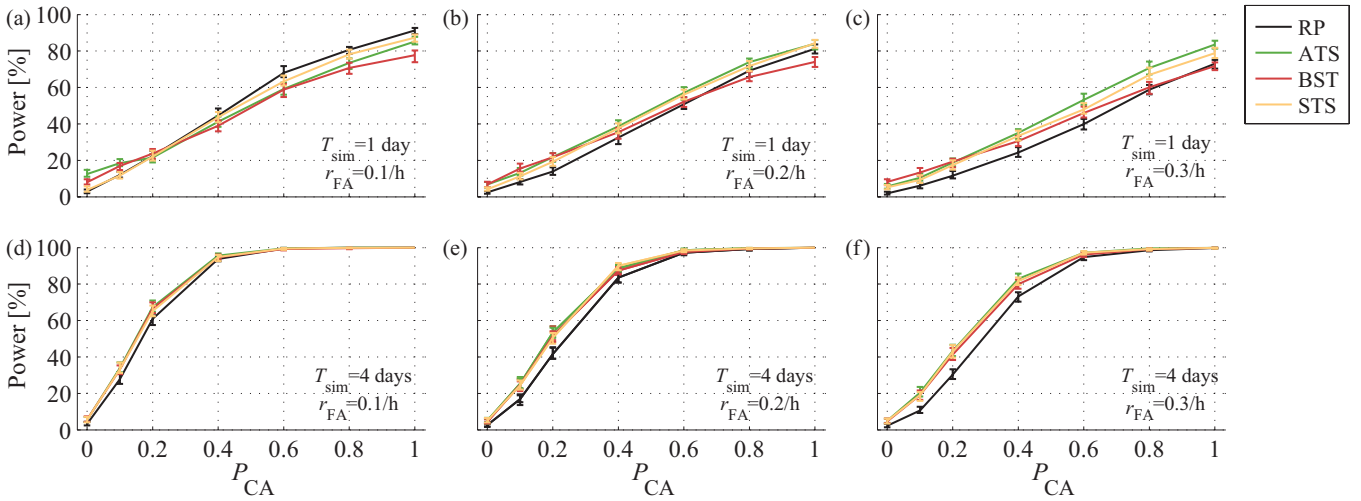

FIG. 2. (Color online) Statistical power of the analytical random predictor (RP, black), the alarm time surrogates (ATS, green or medium gray), seizure time surrogates with random offset (STS, orange or light gray), and with bootstrap resampling (BST, red or dark gray), depending on the probability of predictive alarms P_{CA}. Simulation durations of 1 day in (a)–(c) and 4 days in (d)–(f) are shown for exemplary rates of false alarms r_{FA} and $\gamma_0 = 1/\text{h}$. For the case $P_{CA} = 0$, the power is equivalent to the empirical size (cf. Fig. 1). 95% confidence intervals are based on 1000 simulation instances.

We would like to thank Dr. Ralph Andrzejak for constructive discussions on this matter.