

Using Monte Carlo simulation to propagate processing parameter uncertainty to the statistical analyses of biomechanical trajectories

Todd C. Pataky, Ph.D.

Department of Human Health Sciences

Kyoto University Graduate School of Medicine

Keywords: biomechanics, motion analysis, statistical parametric mapping, stochastic parameterization, probabilistic analysis, sensitivity analysis

Word counts:

Abstract: 192 (max 200)

Body: 2993 (max 3000, including Abstract and captions)

Corresponding author:

Todd Pataky, Associate Professor

Kyoto University Graduate School of Medicine

Department of Human Health Sciences

53 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto, Japan 606-8507

pataky.todd.2m@kyoto-u.ac.jp

Abstract

Biomechanical trajectories are often routed through a chain of processing steps prior to statistical analysis. Since changes in processing parameter values can affect these trajectories, care is required when choosing data processing specifics. The purpose of this Research Note was to demonstrate how data processing parameter uncertainty can be propagated to statistical inferences regarding biomechanical trajectories. As an example application, the correlation between foot contact duration and vertical ground reaction force (VGRF) during constant-speed treadmill walking was considered. Uncertainty was modeled using plausible-range uniform distributions in three data processing steps, and Monte Carlo (MC) simulation was used to construct probabilistic representations of both individual VGRF measurements and the ultimate statistical results. While an initial, plausible set of parameter values yielded a significant correlation between contact duration and late-stance VGRF, MC simulations revealed strong sensitivity, with “significance” being reached in fewer than 40% of simulations, with relatively little net effect of parameter uncertainty magnitude. These results indicate that propagating processing parameter uncertainty to statistical results promotes a cautious, nuanced and robust view of observed effects. By extension, MC simulations may yield greater interpretive consistency across studies involving data processing uncertainties.

1. Introduction

Many human movement variables like joint angles and joint moments cannot be directly measured and are estimated based on external measurements like reflective markers and external forces. This estimation consists of several data processing steps including filtering, inverse dynamics, interpolation, etc., and each processing procedure is generally parameterized, requiring specific parameter values to execute. For example, lowpass filtering generally requires a specific cutoff frequency value. Selection of specific parameter values is usually justified based on community consensus or literature convention (e.g. Wu et al. 2002). However, it is known that multiple, justifiable processing choices can qualitatively affect the ultimately estimated human movement

variables and even their clinical interpretations (Devaprakash et al., 2016). Due to a plethora of commercial and in-house hardware and software systems, experimental tasks, and literature justification possibilities, one is unlikely to know --- in an *a priori* sense --- the specific processing parameter values that ought to be used in a specific measurement scenario.

Various algorithmic approaches have emerged in the literature to deal with the processing parameter uncertainty. One approach is numerical optimization, wherein optimal processing parameter values are sought to minimize the error between measurement and mechanical expectation (e.g. van den Boggert & Koning, 1996). The main limitation of optimization approaches is that they settle on single parameter values, and therefore do not consider how uncertainty in those processing parameters can translate into uncertainty regarding final, reported results.

Another approach is to regard results-relevant parameters as uncertain, and to conduct a Monte Carlo analysis which repeatedly draws random values from a plausible range of parameters. This general procedure has been applied extensively in the modeling literature (e.g. Hughes & An 1997, Valero Cuevas et al. 2003, Ackland et al. 2012). These studies demonstrate how uncertainty in model parameters like initial kinematic postures and muscle moment arms propagate to uncertainty in final results. Plausible value ranges for these biomechanical model parameters can generally be determined through *in vitro* and/or *in vivo* experimentation. Although not generally appearing in the literature, Monte Carlo simulations can also handle uncertainty in *non-measurable* parameters like filter cutoffs, movement start/stop velocity thresholds, etc. Variation in these parameters can also affect the final results, but generally they are treated as fixed in the literature.

A recent study has demonstrated that it is possible to explicitly define distributional models of parameter uncertainty, based on known or assumed measurement error, and then to use Bayesian inference to propagate that error to the estimate of biomechanical quantities like joint angle trajectories (Pohl et al., 2022). The main advantage of this approach is that uncertainty is retained in estimated quantities, and this uncertainty can be propagated to arbitrary statistical inferences like two-sample comparisons. The main disadvantage is that there is a substantial gap between those methods and routine practice: there is no software package aimed at general researchers for conducting these analyses, so adapting these methods to arbitrary experimental scenarios may be

difficult for many researchers. Additionally, the methods have not yet been developed for the analysis of 3D quantities like joint angles, so cannot yet be employed in routine practice. There is therefore a niche for a practical alternative.

The primary aim of this paper was to demonstrate how processing parameter uncertainty in a relatively minimal set of processing parameters can be propagated to final statistical results. This short paper couples a simple, Monte Carlo simulation to trajectory-level statistical inference (Friston et al., 2007) to demonstrate both this error propagation, and how the results can be interpreted in terms of sensitivity.

As an example application, the correlation between step duration and vertical ground reaction forces (VGRF) during treadmill walking is considered below. This example was chosen because VGRF can be readily and accurately measured, and requires just minimal and relatively well-accepted processing; if the value of the proposed probabilistic approach is clear for this relatively simple and accurate VGRF case, then it should also be clear for less accurately measured / estimated quantities like joint moments and joint angles.

2. Methods

2.1 Dataset

A single participant's right-limb VGRF data (31 years, 74 kg, 173 cm) from McClymont et al. (2016) were reanalyzed. The participant walked barefoot on a Zebris FDM-THM treadmill for 30 s at 1.1 m/s, yielding 25 steps per limb. Consideration of both a single subject and a minimal processing parameter scheme (see below) were considered adequate because the goal was not to make scientific inferences regarding VFRG biomechanics; instead the primary goal was to demonstrate how processing parameter uncertainty can propagate to final results.

2.2 Data analysis: Baseline scenario with fixed parameters

Data were processed using three sequential procedures: filter, segment, and register (Table 1).

Filtering consisted of a 4th-order lowpass Butterworth filter with a cutoff frequency of 20 Hz.

Segmentation used a threshold force of 10 N to define stance phase and yielded stance-phase VGRF trajectories with variable duration (Fig.1a). Registration involved linear interpolation across stance phase at $n=101$ equally-spaced time nodes (Fig.1b).

Statistical parametric mapping (Friston 2017) was used to test the null hypothesis (H_0) of no correlation between step duration and the VGRF trajectory. The correlation coefficient r at time q maps to the t -statistic according to:

$$t(q) = r(q) \sqrt{(n-2)/(1-r(q)^2)}$$

The two-tailed critical t -value (t_c) corresponding to a Type I error rate of $\alpha=0.05$ was calculated using random field theory (Pataky 2016). H_0 was rejected if the maximum absolute $t(q)$ exceeded t_c .

Table 1: Data processing and Monte Carlo (MC) simulation summary.

Simulation scenarios

Category	Ordinal Variance	Variance Label	Scenario Name	Filter Cutoff (Hz)	Segment Threshold (N)	Register Nodes	n_{MC}
All-Fixed	0	Zero	Baseline	20	10	100	0
All-Random	1	Small	AR-1	U (19,21)	U (9,11)	UI (99,101)	1
	2		AR-2	U (18,22)	U (8,13)	UI (95,105)	1
	3	Moderate	AR-3	U (15,25)	U (7,15)	UI (90,110)	1
	4		AR-4	U (10,30)	U (6,20)	UI (75,125)	1
	5	Large	AR-5	U (5,35)	U (5,25)	UI (50,150)	1

				Filter	Segment	Register	
Category	Ordinal Variance	Variance Label	Scenario Name	Cutoff (Hz)	Threshold (N)	Nodes	nMC
One-Fixed	3	Moderate	Cutoff-Fixed	[5, 7, 9, ... 33, 35]	(Moderate)	(Moderate)	16
	3	Moderate	Threshold-Fixed	(Moderate)	[5, 7, 9, ... 23, 25]	(Moderate)	11
	3	Moderate	Nodes-Fixed	(Moderate)	(Moderate)	[50,60,70, ... 140, 150]	11

U = Uniform distribution

UI = Uniform (integer) distribution

nMC = number of 10,000-iteration MC simulations

2.3 Monte Carlo simulations

A set of Monte Carlo (MC) simulations was conducted wherein the three aforementioned processing parameters (cutoff, threshold, and nodes) were varied randomly according to the scenarios and distributions specified in Table 1. Each MC simulation consisted of $N=10,000$ iterations, where each iteration consisted of (i) a random drawing of parameter values, then (ii) processing and statistical analysis. A total of 43 MC simulations were conducted. Each simulation produced N t-value trajectories, N critical thresholds, and thus N null hypothesis (H_0) rejection decisions. The H_0 rejection rate was used to summarize each MC simulation. Changes in H_0 rejection rates across MC simulations were qualitatively compared.

*** Insert Figure 1 about here ***

3. Results

Fixed-parameter analysis yielded a moderately strong correlation between contact duration and VGRF (Fig.1c-d), with maximum correlation at time=74% ($r=0.619$) (Fig.1c). This correlation was sufficiently strong to reject H_0 at $\alpha=0.05$ (Fig.1c).

The “moderate” variance MC simulation scenario (see Table 1) yielded an H_0 rejection frequency of $p=0.369$ (Fig.2c). The locally maximum H_0 rejection rate was $p=0.231$ at time=76% (Fig.2b,e). Thus H_0 was rejected occasionally when t-values in the time=90-100% region were large (Fig.2e). H_0 rejection rates were highly variable for $N < 100$ (i.e., the first 100 simulation iterations), and stabilized for $N > 1000$ (Fig.2f).

Increasing parameter uncertainty from “small” to “large” (see Table 1) yielded qualitatively more uncertainty in the t-value trajectories (Fig.3a-c), critical t-values (Fig.3a-f) and maximum absolute t-values (Fig.3d-f). Nevertheless this increased variance did not have a systematic effect on either the maximum local variance (time=76%, Fig.3g) or the overall H_0 rejection rate (Fig.3h). Fixing one processing parameter and randomly varying the other two revealed that individual parameters can affect the H_0 rejection rate in different and highly nonlinear ways (Fig.3i).

*** Insert Figure 2 about here ***

*** Insert Figure 3 about here ***

4. Discussion

The “baseline” fixed-parameter result (Fig.1) suggests a significant positive correlation between contact duration and VGRF. The “moderate” variance MC simulation results (Fig.2) contradict that

conclusion and suggest that the Fig.1 results are merely a single case from a distribution of feasible results. Systematic MC simulations suggest that increasing the uncertainty in modeled processing parameters can substantially affected H0 rejection rates in some cases (Fig.3i) but can also have negligible effects in others (Fig.3g-h). More generally, the MC approach indicates that fixed-parameter results may be insufficient; robust statistical conclusions may require consideration of ranges of feasible parameter values.

A minor but notable result was that the correlation between contact duration and VGRF was positive (Fig.1c-d), in apparent contrast to gait studies which generally report *negative* correlations (e.g. Rosenbaum et al., 2004). This apparent disagreement is partly resolved by the fact that the treadmill controlled average walking speed, and thus that contact duration variability was relatively small, ranging only 30 ms across the 25 recorded steps (Fig.1d); walking speed studies generally consider a much broader range of walking speeds and thus contact durations. The current study did not aim to make scientific conclusions regarding VGRF, so no further consideration is given to the biomechanical meaning of these results.

The Fig.2-3 results represent a relatively straightforward sensitivity analysis, wherein MC simulation iterations are completely independent. Other approaches to the problem of processing parameter uncertainty also exist. For example, a Bayesian model could be constructed (similar to Serrien et al., 2019) wherein both VGRF and modeled processing parameter uncertainty are used to estimate the true linear regression slope. Similar to the current MC analysis, Bayesian inference embeds both unknown parameters of interest (i.e., regression slope) and nuisance parameters (i.e., processing parameters) within a statistical model and treats the unknowns as random variables. The *a priori* uncertainty in these parameters is described using prior distributions and this distribution is updated via Bayes' rule using observed data to formulate posterior distributions which are subsequently used to make statistical inferences. While robust, constructing models and executing Bayesian analysis is technically more complex than MC simulation, so Bayesian inference may be unnecessary if one's main aim is simply to scrutinize one's primary conclusions.

A problem common to MC simulation and Bayesian inference is the selected parameter distribution. In MC simulation this distribution is static throughout analyses, but in Bayesian

inference the initially specified prior distribution is updated iteratively using Bayes' rule as described above. The sensitivity of Bayesian inference to the prior distribution has been recently considered for kinematics analyses (Pohl et al. 2022). Regardless of the inference approach, a researcher must choose a starting distribution. As a first approximation, an apparently feasible distribution can be derived from the range of values used in the literature, as approximately embodied in Table 1. Since MC results can be affected by distribution particulars (Fig.3i) one may need to conduct multiple MC simulations with systematic distribution changes to ensure the robustness of one's conclusions. A more robust approach to distributional assumptions would likely require substantial research effort to determine what processing algorithms, what parameters, and what parameter ranges may be appropriate for particular data modalities with particular signal and noise characteristics. At this stage the only justifiable, practical recommendation is: *if you wish to claim that your results are insensitive to processing parameter value changes, then select the widest possible parameter range to substantiate that claim.*

The main benefit of an MC sensitivity approach is that it encourages a cautious and nuanced interpretation of observed effects. It achieves this using parameter uncertainty to transform binary null hypothesis rejection decisions (Fig.1c) into a continuous probability space (Fig.2b-e) wherein the apparent importance of a specific effect can substantially diminish. In the specific VGRF case considered in this study, the sensitivity results in Figs.2-3 suggest that there is insufficient evidence to support a conclusion of an effect, and this in turn implies that a potentially erroneous biomechanical interpretation is avoided.

The proposed approach also promotes reporting consistency in that parameter value ranges are more easily justified than are specific parameter values. Imagine two similar studies whose only data processing difference is that one uses a filter cutoff of 20 Hz and one uses 25 Hz. These studies may justify those choices by citing consensus, previous research, or neither, and these justifications or lack thereof may be regarded as satisfactory. The results of the current study suggest that apparently minor processing differences like these can substantially affect statistical conclusions, and thus that a more robust approach would be is to consider ranges of parameter values.

This paper considered a relatively small set of common data processing algorithms and a relatively small effect (Fig.1c-d). Real data processing can involve considerably more complex calculations (e.g., inverse dynamics) and thus results can depend on a much larger set of unknown parameters (Valero-Cuevas et al. 2003, Ackland et al. 2012). It is conceivable that uncertainty in this larger set of parameters could broaden the range of expected thresholds (like in Fig.3a-c), and therefore could require that considerably larger effects are also interpreted with caution.

In summary, this study has shown how uncertainty in data processing parameters can be propagated to statistical inferences regarding biomechanical trajectories, and how to robustly interpret classical hypothesis testing results in terms of feasible outcomes. The approach is relatively easy to implement through Monte Carlo simulation, and amounts to a systematic sensitivity analysis that could improve reporting robustness.

Acknowledgments

Thank you to Juliet McClymont for your wonderful willingness to share, and to Russ Savage and Robin Crompton for your unending support. Thank you to Andy Pohl and team for inspiring this article. Thank you most of all to Vladimir Zatsiorsky: you gave me a chance and launched my career, and I am forever indebted to you.

References

Ackland, D. C., Lin, Y. C., & Pandy, M. G. (2012). Sensitivity of model predictions of muscle function to changes in moment arms and muscle–tendon properties: a Monte-Carlo analysis. *Journal of Biomechanics*, 45(8), 1463-1471.

Devaprakash, D., Weir, G. J., Dunne, J. J., Alderson, J. A., & Donnelly, C. J. (2016). The influence of digital filter type, amplitude normalisation method, and co-contraction algorithm on clinically

relevant surface electromyography data during clinical movement assessments. *Journal of Electromyography and Kinesiology*, 31, 126-135.

Friston, K. J., Ashburner, J. T., Kiebel, S. J., & Nichols, T. E., Penny, W. D. (Eds.). (2007). *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Elsevier.

Hughes, R. E., & An, K. N. (1997). Monte Carlo simulation of a planar shoulder model. *Medical and Biological Engineering and Computing*, 35(5), 544-548.

McClymont, J., Pataky, T. C., Crompton, R. H., Savage, R., & Bates, K. T. (2016). The nature of functional variability in plantar pressure during a range of controlled walking speeds. *Royal Society Open Science*, 3(8), 160369.

Pataky, T. C. (2016). RFT1D: Smooth one-dimensional random field upcrossing probabilities in Python. *Journal of Statistical Software*, 71(7), 1-22.

Pohl, J.A., Schofield, M. R., Ferber, R. (2022). A generalised smoothing approach for continuous, planar, inverse kinematics problems. *Journal of Biomechanics*, in review.

Serrien, B., Goossens, M., & Baeyens, J. P. (2019). Statistical parametric mapping of biomechanical one-dimensional data with Bayesian inference. *International Biomechanics*, 6(1), 9-18.

Valero-Cuevas, F. J., Johanson, M. E., & Towles, J. D. (2003). Towards a realistic biomechanical model of the thumb: the choice of kinematic description may be more critical than the solution method or the variability/uncertainty of musculoskeletal parameters. *Journal of Biomechanics*, 36(7), 1019-1030.

van den Bogert, A. J., & Koning, J. J. (1996). On optimal filtering for inverse dynamics analysis. In *Proceedings of the IXth Biennial Conference of the Canadian Society for Biomechanics* (pp. 214–215). Vancouver, Canada.

Wu, G., Siegler, S., Allard, P., et al. (2002). ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion—part I: ankle, hip, and spine. *Journal of Biomechanics*, 35(4), 543-548.

Figures

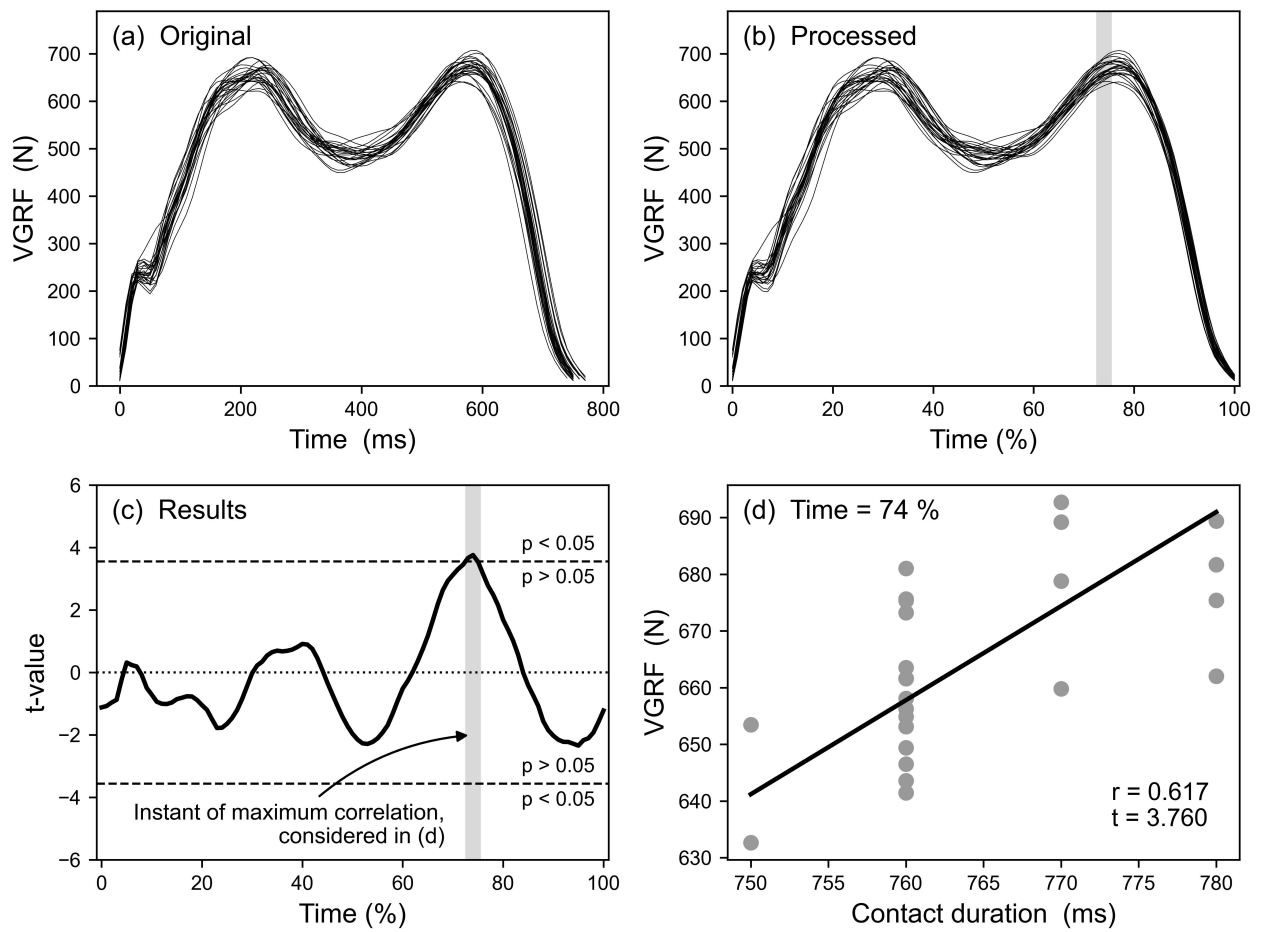


Figure 1: Results for the "Baseline" scenario (see Table 1). VGRF = Vertical ground reaction force. (a) Original data, note that durations are variable. (b) Processed data. (c) Linear regression results, positive/negative t-values indicate positive/negative correlations. (d) Scatterplot at instant of maximal correlation.

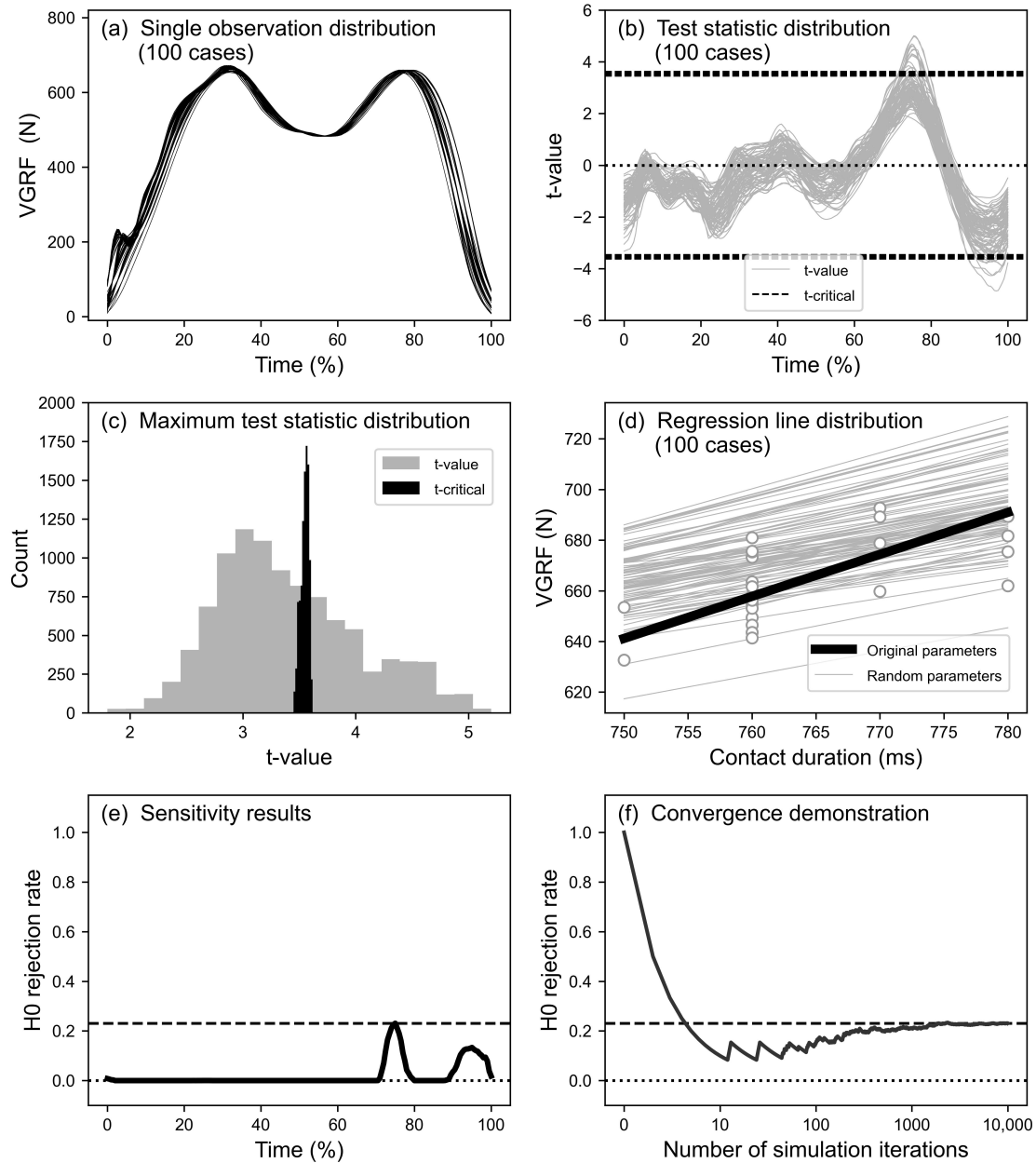


Figure 2: Monte Carlo (MC) simulation results, "Moderate" variance scenario (see Table 1). For visual interpretability, panels (a), (b) and (d) depict only 100 cases. (a) Due to modeled processing parameter uncertainty, each observation also possesses uncertainty. (b) Hypothesis testing results; both t-values and critical thresholds embody uncertainty. (c) Histogram of maximum t-value across all 10,000 MC simulation iterations. (d) Depiction of linear regression variance. (e) Null hypothesis (H0) rejection sensitivity results; the horizontal dashed line highlights the maximum H0 rejection rate. (f) Demonstration of H0 rejection rate convergence as the number of MC simulation iterations increases.

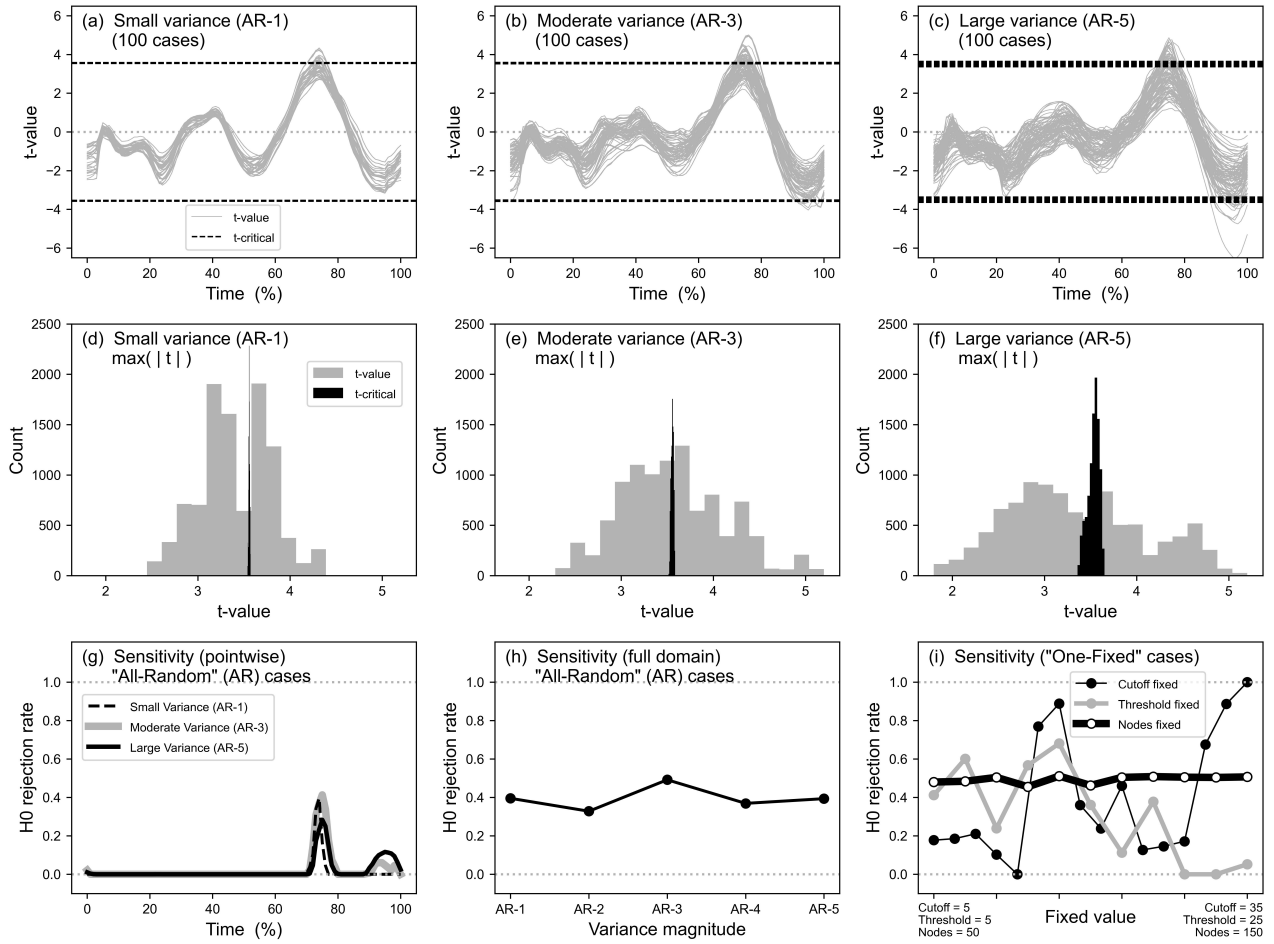


Figure 3: Summary of all Monte Carlo (MC) simulation results. (a-c) Test statistic uncertainty for the “Small” (AR-1), “Moderate” (AR-3) and “Large” (AR-5) variance cases (see Table 1); for visual interpretability only 100 cases are depicted. (d-f) Histograms of the domain-wise maximum t-value across the 10,000 MC simulation iterations. (g) Pointwise null hypothesis (h_0) rejection sensitivity results for the AR-1, AR-3 and AR-5 scenarios. (h) Full domain H_0 rejection sensitivity across the AR-1 to AR-5 scenarios. (i) Full domain H_0 rejection sensitivity for the three “One-fixed” scenarios; each point summarizes 10,000 MC simulation iterations where two parameters varied randomly and the fixed parameter remained constant.