Supplementary Information

Electrochemical tomography as a nondestructive technique to study localized corrosion of metals

M.C. van Ede[1], C.J. Earls[2], A. Fichtner[3], U. Angst[1]*

[1]Institute for Building Materials, ETH Zurich, 8093, Zurich, Switzerland
[2]School of Civil & Environmental Engineering, Cornell University, Ithaca, NY, 14850, United States
[3] Institute of Geophysics, ETH Zurich, 8092, Zurich, Switzerland
*Corresponding author. E-mail: ueli.angst@ifb.baug.ethz.ch

1 Supplementary Note: Convergence diagnostics

Following Brooks and Gelman[1], a number of m Markov chains are generated, each with a different starting point, randomly selected from the prior range. From these chains the burn-in samples are discarded and one is left with chains of length n. The between-sequence variance B/n and the within-sequence variance W can then be computed:

$$B/n = \frac{1}{m-1}\sum_{j=1}^{m}(\overline{\varphi}_{j.} - \overline{\varphi}_{..})^2 \tag{1}$$

$$W = \frac{1}{m(n-1)}\sum_{j=1}^{m}\sum_{t=1}^{n}(\varphi_{jt} - \overline{\varphi}_{j.})^2 \tag{2}$$

Here φ_{jt} is the tth sample of chain j, $\overline{\varphi}_{j.}$ the mean of all samples in chain j and $\overline{\varphi}_{..}$ the total mean of all samples in each chain.

Using these variances, we estimate the true variance σ^2 of the sampled distribution using a weighted average of B and W:

$$\hat{\sigma}_+^2 = \frac{n-1}{n}W + \frac{B}{n} \tag{3}$$

From this one can compute the potential scale reduction factor $\hat{R}$, which is an estimation of the variance ratio between pooled and within-chain inference.

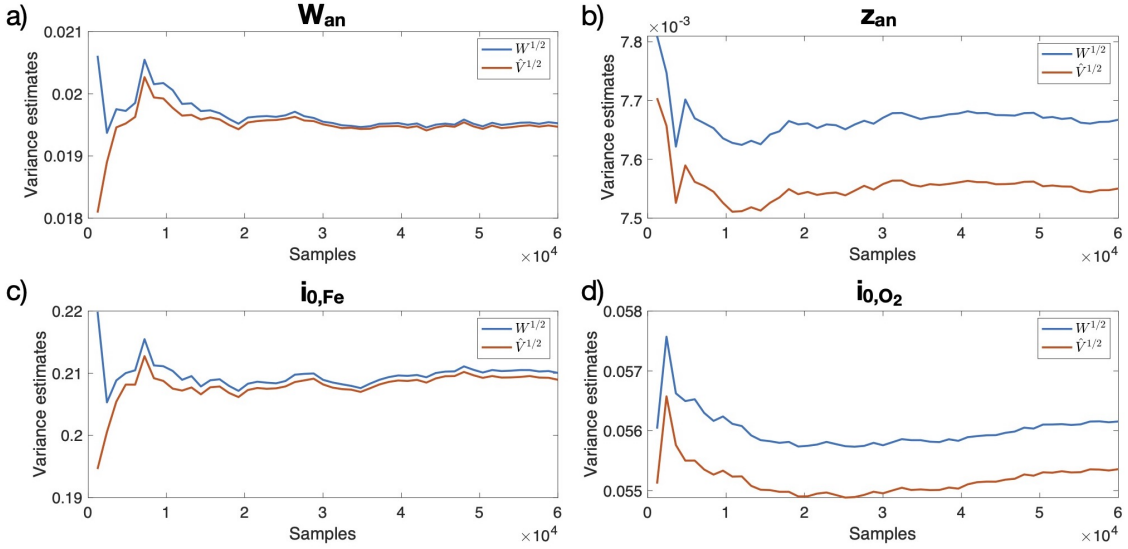
$$\hat{R} = \frac{m+1}{m}\frac{\hat{\sigma}_+^2}{W} - \frac{n-1}{mn} \tag{4}$$

When $\hat{R}$ is close to 1, we most likely reached convergence and the Markov chains closely describe the target distribution.

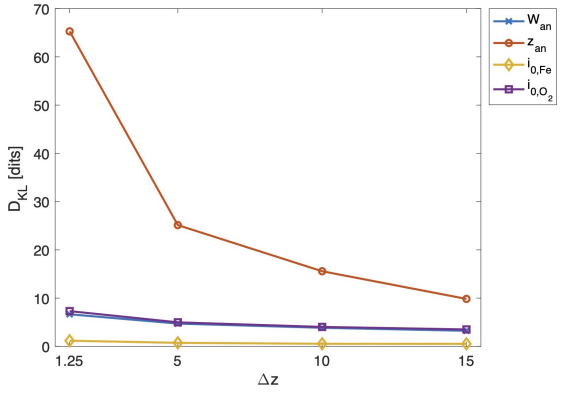
Brooks and Gelman[1] proposed to divide the chains into k batches of length b, where k should typically be around 20, or more for reasonable output. This enables monitoring of convergence as a function of the batches. Brooks and Gelman[1] state three conditions which should hold at convergence: The within-sequence variance, $W(k)$, and the mixture-of-sequences variance, $\hat{V}(k) = \hat{\sigma}_+^2 + \frac{B}{mn}$, should stabilize and $\hat{R}(k)$ should approach 1.

We generate 5 Markov chains with a burn-in of 5000 and 60,000 samples, starting each at a different randomly selected set of initial model parameters, within the bounds set by the priors (+/- 2*σ). The chains where divided in k=50 batches. Supplementary figure 1 shows, in addition to figure 5a, the evolution of $W(k)$ and $\hat{V}(k)$ over these batches. In around 30,000 iterations both $W(k)$ and $\hat{V}(k)$ seem to have stabilized.

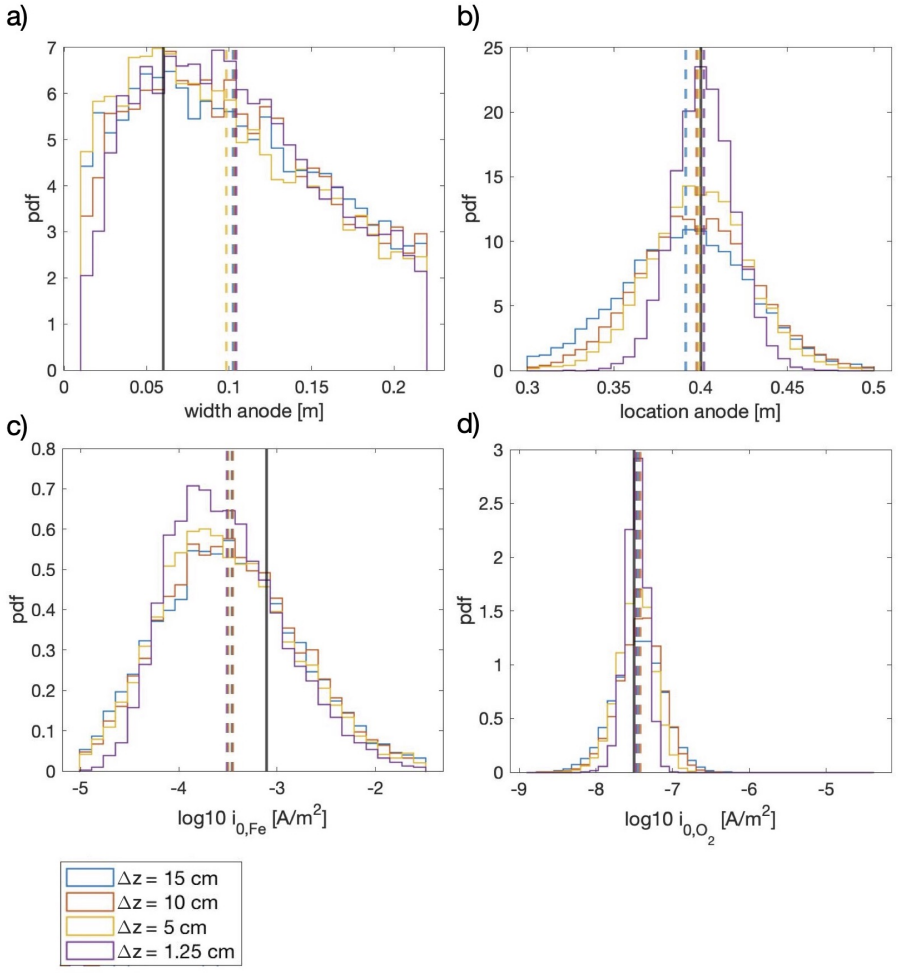
2 Supplementary Figures



Supplementary Figure 1: The evolution of $W(k)$ & $\hat{V}(k)$ for the four different model parameters: a) the width of the anode W_{an} b) the location of the anode z_{an} c) the exchange current density of iron $log10(i_{0,Fe})$ d) the exchange current density of oxygen $log10(i_{0,O_2})$. In around 30,000 samples $W^{1/2}(k)$ & $\hat{V}^{1/2}(k)$ seem to have stabilized for all parameters.



Supplementary Figure 2: The information gain for different spacing between data points (Δz) at the concrete surface, for the four model parameters: W_{an}, z_{an}, $log10(i_{0,Fe})$ and $log10(i_{0,O_2})$.



Supplementary Figure 3: MCMC results for different spacing between data points (Δz) at the concrete surface, and thus different length of the data vector $\vec{\phi}^{obs}$, for model parameters a) W_{an}, b) z_{an}, c) $log10(i_{0,Fe})$, d) $log10(i_{0,O_2})$. The vertical black lines represent the known model parameters.

References

[1] Brooks, S. P. & Gelman, A. General methods for monitoring convergence of iterative simulations. *Journal of computational and graphical statistics* **7**, 434–455 (1998).