Supplementary materials for ‘Openness and computational reproducibility in plant pathology: where do we stand and a way forward’

Adam H. Sparks1,2,✉, Emerson M. Del Ponte3, Kaique S. Alves3, Zachary S. L. Foster4, and Niklaus J. Grünwald5

1 Department of Primary Industries and Regional Development, Farming Systems Innovation, Perth WA 6000, Australia  
2 University of Southern Queensland, Centre for Crop Health, Toowoomba Qld 4350, Australia  
3 Departmento de Fitopatologia, Universidade Federal de Viçosa, Brazil  
4 Department of Botany and Plant Pathology, Oregon State University, Corvallis OR 97331, USA  
5 Horticultural Crops Research Unit, USDA Agricultural Research Service, Corvallis OR 97330, USA

✉ Corresponding author: [Adam H. Sparks <[Adam.Sparks@dpird.wa.gov.au](mailto:Adam.Sparks@dpird.wa.gov.au)>](mailto:Adam.Sparks@dpird.wa.gov.au)

# Supplementary Materials

## Supplementary Tables

Table 1: Full description of model fit for the effect of journal title on code availability. We fitted a Bayesian logistic mixed model (estimated using MCMC sampling with 4 chains of 10000 iterations and a warmup of 5000) to predict comp\_mthds\_avail with abbreviation (formula: comp\_mthds\_avail ~ abbreviation). The model included assignee as random effect (formula: ~1 | assignee). Priors over parameters were set as normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00) and student\_t (location = 0.00, scale = 2.50) distributions.

| Parameter | Median | CI | CI Low | CI High | pd | Rhat | ESS | Fit |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Intercept[1] | 0.17 | 0.95 | -1.39 | 1.71 | 0.59 | 1.00 | 20,885.20 |  |
| Intercept[2] | 0.56 | 0.95 | -0.98 | 2.09 | 0.77 | 1.00 | 22,977.13 |  |
| AustralasPlantPath | -0.12 | 0.95 | -2.03 | 1.65 | 0.55 | 1.00 | 30,384.50 |  |
| CanJPlantPathol | -0.18 | 0.95 | -2.06 | 1.57 | 0.58 | 1.00 | 31,898.52 |  |
| CropProt | -0.22 | 0.95 | -2.02 | 1.44 | 0.60 | 1.00 | 30,761.34 |  |
| EurJPlantPathol | -0.17 | 0.95 | -2.02 | 1.54 | 0.57 | 1.00 | 28,146.82 |  |
| ForestPathol | -0.20 | 0.95 | -2.00 | 1.50 | 0.58 | 1.00 | 31,905.00 |  |
| JPhytopathol | -0.33 | 0.95 | -2.08 | 1.30 | 0.65 | 1.00 | 31,231.12 |  |
| JPlantPathol | -0.21 | 0.95 | -2.05 | 1.48 | 0.59 | 1.00 | 29,063.96 |  |
| MolPlantMicroIn | 0.49 | 0.95 | -1.24 | 2.06 | 0.72 | 1.00 | 30,218.34 |  |
| MolPlantPathol | -0.26 | 0.95 | -2.07 | 1.39 | 0.61 | 1.00 | 31,148.37 |  |
| Nematology | -0.20 | 0.95 | -2.05 | 1.50 | 0.59 | 1.00 | 31,651.84 |  |
| PhysiolMolPlantP | -0.21 | 0.95 | -2.03 | 1.46 | 0.60 | 1.00 | 30,455.31 |  |
| Phytoparasitica | -0.24 | 0.95 | -2.10 | 1.44 | 0.61 | 1.00 | 30,266.81 |  |
| PhytopatholMediterr | -0.19 | 0.95 | -2.10 | 1.54 | 0.58 | 1.00 | 29,854.90 |  |
| PlantDis | -0.17 | 0.95 | -2.05 | 1.54 | 0.57 | 1.00 | 26,753.21 |  |
| PlantHealthProgress | -0.15 | 0.95 | -2.02 | 1.61 | 0.56 | 1.00 | 30,382.01 |  |
| PlantPathol | -0.24 | 0.95 | -2.08 | 1.43 | 0.61 | 1.00 | 32,634.40 |  |
| RevMexFitopatol | -0.22 | 0.95 | -2.08 | 1.42 | 0.60 | 1.00 | 29,507.33 |  |
| TropPlantPathol | 0.66 | 0.95 | -1.16 | 2.34 | 0.77 | 1.00 | 29,671.83 |  |
| VirolJ | -0.13 | 0.95 | -1.98 | 1.60 | 0.56 | 1.00 | 31,229.60 |  |
| ELPD |  |  |  |  |  |  |  | -34.10 |
| LOOIC |  |  |  |  |  |  |  | 68.21 |
| WAIC |  |  |  |  |  |  |  | 66.58 |

Table 2: Full description of model fit for the effect of journal title on data availability. We fitted a Bayesian logistic mixed model (estimated using MCMC sampling with 4 chains of 10000 iterations and a warmup of 5000) to predict comp\_mthds\_avail with abbreviation (formula: comp\_mthds\_avail ~ abbreviation). The model included assignee as random effect (formula: ~1 | assignee). Priors over parameters were set as normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00) and student\_t (location = 0.00, scale = 2.50) distributions.

| Parameter | Median | CI | CI Low | CI High | pd | Rhat | ESS | Fit |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Intercept[1] | 0.39 | 0.95 | -0.96 | 1.69 | 0.72 | 1.00 | 4,719.32 |  |
| Intercept[2] | 0.67 | 0.95 | -0.66 | 1.96 | 0.83 | 1.00 | 4,808.63 |  |
| Intercept[3] | 1.14 | 0.95 | -0.19 | 2.43 | 0.95 | 1.00 | 4,868.56 |  |
| AustralasPlantPath | 0.62 | 0.95 | -0.71 | 1.84 | 0.83 | 1.00 | 22,512.63 |  |
| CanJPlantPathol | 0.30 | 0.95 | -0.83 | 1.32 | 0.70 | 1.00 | 20,869.52 |  |
| CropProt | -1.23 | 0.95 | -2.73 | 0.02 | 0.97 | 1.00 | 20,199.84 |  |
| EurJPlantPathol | 0.08 | 0.95 | -1.11 | 1.15 | 0.55 | 1.00 | 21,647.41 |  |
| ForestPathol | -0.09 | 0.95 | -1.26 | 0.95 | 0.57 | 1.00 | 21,567.86 |  |
| JPhytopathol | -0.25 | 0.95 | -1.22 | 0.62 | 0.71 | 1.00 | 18,003.18 |  |
| JPlantPathol | 0.15 | 0.95 | -0.94 | 1.14 | 0.62 | 1.00 | 20,336.25 |  |
| MolPlantMicroIn | 0.54 | 0.95 | -0.43 | 1.47 | 0.87 | 1.00 | 19,561.11 |  |
| MolPlantPathol | 0.94 | 0.95 | 0.10 | 1.75 | 0.99 | 1.00 | 16,326.91 |  |
| Nematology | -0.19 | 0.95 | -1.51 | 0.95 | 0.62 | 1.00 | 22,360.61 |  |
| PhysiolMolPlantP | 0.59 | 0.95 | -0.34 | 1.47 | 0.89 | 1.00 | 17,765.30 |  |
| Phytoparasitica | -0.47 | 0.95 | -1.72 | 0.62 | 0.79 | 1.00 | 21,343.22 |  |
| PhytopatholMediterr | 1.68 | 0.95 | 0.78 | 2.57 | 1.00 | 1.00 | 17,609.88 |  |
| PlantDis | -1.26 | 0.95 | -2.74 | -0.02 | 0.98 | 1.00 | 22,282.14 |  |
| PlantHealthProgress | -0.59 | 0.95 | -2.01 | 0.64 | 0.82 | 1.00 | 21,371.07 |  |
| PlantPathol | -0.08 | 0.95 | -1.08 | 0.83 | 0.57 | 1.00 | 19,017.40 |  |
| RevMexFitopatol | -1.15 | 0.95 | -2.64 | 0.12 | 0.96 | 1.00 | 24,184.64 |  |
| TropPlantPathol | 0.35 | 0.95 | -0.76 | 1.39 | 0.74 | 1.00 | 20,363.83 |  |
| VirolJ | 0.87 | 0.95 | -0.03 | 1.73 | 0.97 | 1.00 | 17,099.91 |  |
| ELPD |  |  |  |  |  |  |  | -284.23 |
| LOOIC |  |  |  |  |  |  |  | 568.45 |
| WAIC |  |  |  |  |  |  |  | 568.32 |

Table 3: Full description of model fit for the effect of year of publication on code availability. We fitted a Bayesian logistic mixed model (estimated using MCMC sampling with 4 chains of 10000 iterations and a warmup of 5000) to predict comp\_mthds\_avail with year (formula: comp\_mthds\_avail ~ year). The model included abbreviation and assignee as random effects (formula: list(~1 | abbreviation, ~1 | assignee)). Priors over parameters were set as normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), student\_t (location = 0.00, scale = 2.50) and student\_t (location = 0.00, scale = 2.50) distributions.

| Parameter | Median | CI | CI Low | CI High | pd | Rhat | ESS | Fit |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Intercept[1] | 398.63 | 0.95 | -252.91 | 1,160.62 | 0.88 | 1.00 | 19,651.43 |  |
| Intercept[2] | 398.98 | 0.95 | -252.28 | 1,161.30 | 0.88 | 1.00 | 19,650.40 |  |
|  | 0.20 | 0.95 | -0.13 | 0.58 | 0.88 | 1.00 | 19,642.86 |  |
| ELPD |  |  |  |  |  |  |  | -31.43 |
| LOOIC |  |  |  |  |  |  |  | 62.86 |
| WAIC |  |  |  |  |  |  |  | 60.93 |

Table 4: Full description of model fit for the effect of year of publication on data availability. We fitted a Bayesian logistic mixed model (estimated using MCMC sampling with 4 chains of 10000 iterations and a warmup of 5000) to predict comp\_mthds\_avail with year (formula: comp\_mthds\_avail ~ year). The model included abbreviation and assignee as random effects (formula: list(~1 | abbreviation, ~1 | assignee)). Priors over parameters were set as normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), student\_t (location = 0.00, scale = 2.50) and student\_t (location = 0.00, scale = 2.50) distributions.

| Parameter | Median | CI | CI Low | CI High | pd | Rhat | ESS | Fit |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Intercept[1] | 166.73 | 0.95 | -16.03 | 355.77 | 0.96 | 1.00 | 21,806.65 |  |
| Intercept[2] | 167.02 | 0.95 | -15.75 | 356.08 | 0.96 | 1.00 | 21,807.78 |  |
| Intercept[3] | 167.49 | 0.95 | -15.23 | 356.46 | 0.96 | 1.00 | 21,806.72 |  |
|  | 0.08 | 0.95 | -0.01 | 0.18 | 0.96 | 1.00 | 21,691.79 |  |
| ELPD |  |  |  |  |  |  |  | -284.84 |
| LOOIC |  |  |  |  |  |  |  | 569.68 |
| WAIC |  |  |  |  |  |  |  | 569.48 |

Table 5: Full description of model fit for the effect of five-year impact factor on code availability. We fitted a Bayesian logistic mixed model (estimated using MCMC sampling with 4 chains of 10000 iterations and a warmup of 5000) to predict comp\_mthds\_avail with year (formula: comp\_mthds\_avail ~ year). The model included abbreviation and assignee as random effects (formula: list(~1 | abbreviation, ~1 | assignee)). Priors over parameters were set as normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), student\_t (location = 0.00, scale = 2.50) and student\_t (location = 0.00, scale = 2.50) distributions.

| Parameter | Median | CI | CI Low | CI High | pd | Rhat | ESS | Fit |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Intercept[1] | 1.54 | 0.95 | -0.45 | 3.54 | 0.94 | 1.00 | 13,682.75 |  |
| Intercept[2] | 1.93 | 0.95 | -0.08 | 3.95 | 0.97 | 1.00 | 14,594.99 |  |
| IF\_5year | 0.45 | 0.95 | -0.05 | 1.04 | 0.96 | 1.00 | 13,935.20 |  |
| ELPD |  |  |  |  |  |  |  | -32.04 |
| LOOIC |  |  |  |  |  |  |  | 64.08 |
| WAIC |  |  |  |  |  |  |  | 62.86 |

Table 6: Full description of model fit for the effect of five year impact factor on data availability. We fitted a Bayesian logistic mixed model (estimated using MCMC sampling with 4 chains of 10000 iterations and a warmup of 5000) to predict comp\_mthds\_avail with year (formula: comp\_mthds\_avail ~ year). The model included abbreviation and assignee as random effects (formula: list(~1 | abbreviation, ~1 | assignee)). Priors over parameters were set as normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), normal (mean = 0.00, SD = 1.00), student\_t (location = 0.00, scale = 2.50) and student\_t (location = 0.00, scale = 2.50) distributions.

| Parameter | Median | CI | CI Low | CI High | pd | Rhat | ESS | Fit |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Intercept[1] | 0.86 | 0.95 | -0.45 | 2.03 | 0.90 | 1.00 | 4,068.28 |  |
| Intercept[2] | 1.12 | 0.95 | -0.20 | 2.28 | 0.95 | 1.00 | 4,109.16 |  |
| Intercept[3] | 1.55 | 0.95 | 0.23 | 2.73 | 0.99 | 1.00 | 4,188.23 |  |
| IF\_5year | 0.15 | 0.95 | -0.00 | 0.30 | 0.97 | 1.00 | 14,822.60 |  |
| ELPD |  |  |  |  |  |  |  | -295.85 |
| LOOIC |  |  |  |  |  |  |  | 591.71 |
| WAIC |  |  |  |  |  |  |  | 591.70 |

## Supplementary Figures

![Figure 1: Criteria scores for 450 articles computational materials and data availability for each of the five evaluators. Each article was evaluated on a 0 to 3 scale for computational materials (Code) and raw data availability (Data) by one of five evaluators.](data:application/eps;base64,)

Figure 1: Criteria scores for 450 articles computational materials and data availability for each of the five evaluators. Each article was evaluated on a 0 to 3 scale for computational materials (Code) and raw data availability (Data) by one of five evaluators.

![Figure 2: Equivalence test for a model testing the effect of the publishing journal on articles’ supporting code availability.](data:application/eps;base64,)

Figure 2: Equivalence test for a model testing the effect of the publishing journal on articles’ supporting code availability.

![Figure 3: Equivalence test for a model testing the effect of the publishing journal on articles’ supporting data availability.](data:application/eps;base64,)

Figure 3: Equivalence test for a model testing the effect of the publishing journal on articles’ supporting data availability.

![Figure 4: Equivalence test for a model testing the effect of the year of publication title on articles’ supporting code availability.](data:application/eps;base64,)

Figure 4: Equivalence test for a model testing the effect of the year of publication title on articles’ supporting code availability.

![Figure 5: Equivalence test for a model testing the effect of the year of publication on articles’ supporting data availability.](data:application/eps;base64,)

Figure 5: Equivalence test for a model testing the effect of the year of publication on articles’ supporting data availability.

![Figure 6: Equivalence test for a model testing the effect of the publishing journal’s five-year impact factor on articles’ supporting code availability.](data:application/eps;base64,)

Figure 6: Equivalence test for a model testing the effect of the publishing journal’s five-year impact factor on articles’ supporting code availability.

![Figure 7: Equivalence test for a model testing the effect of the year of publication on articles’ supporting data availability.](data:application/eps;base64,)

Figure 7: Equivalence test for a model testing the effect of the year of publication on articles’ supporting data availability.