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

Adam H. Sparks1,2,✉, Emerson M. Del Ponte3, Kaique S. Alves3, Zachary S. L. Foster4, and Niklaus J. Grünwald4

1 Department of Primary Industries and Regional Development, 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 Horticultural Crops Disease and Pest Management 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 S1: Full description of the model that was fit to scoring data, which were used to evaluate the effect of journal title on code availability for 450 papers published in 21 plant pathology journals or plant pathology focused articles from other specialized journals. Scoring for 'Code Availability' was scored 0-3 where, '0' was 'Not available or not mentioned in the publication'; '1' was 'Available upon request to the author; '2' was 'Online, but inconvenient or non-permanent (e.g., login needed, paywall, FTP server, personal lab website that may disappear, or may have already disappeared)'; and '3' was 'Freely available online to anonymous users for foreseeable future (e.g., archived using Zenodo, dataverse or university library or some other proper archiving system)'; 'NA' indicates that no code was created to conduct the work that was detectable. We fit 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 | Effect Size |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| AustralasPlantPath | -0.13 | 0.95 | -2.02 | 1.67 | 0.55 | 1.00 | 32,338.99 | very small |
| CanJPlantPathol | -0.18 | 0.95 | -2.08 | 1.56 | 0.58 | 1.00 | 28,515.96 | very small |
| CropProt | -0.22 | 0.95 | -2.06 | 1.48 | 0.60 | 1.00 | 30,612.06 | very small |
| EurJPlantPathol | -0.19 | 0.95 | -2.04 | 1.52 | 0.58 | 1.00 | 31,234.25 | very small |
| ForestPathol | -0.19 | 0.95 | -2.06 | 1.52 | 0.58 | 1.00 | 31,202.06 | very small |
| JPhytopathol | -0.31 | 0.95 | -2.07 | 1.28 | 0.64 | 1.00 | 32,113.03 | very small |
| JPlantPathol | -0.21 | 0.95 | -2.07 | 1.49 | 0.59 | 1.00 | 31,482.07 | very small |
| MolPlantMicroIn | 0.49 | 0.95 | -1.26 | 2.05 | 0.72 | 1.00 | 30,985.04 | very small |
| MolPlantPathol | -0.26 | 0.95 | -2.06 | 1.39 | 0.61 | 1.00 | 30,645.93 | very small |
| Nematology | -0.20 | 0.95 | -2.05 | 1.52 | 0.59 | 1.00 | 30,839.64 | very small |
| PhysiolMolPlantP | -0.22 | 0.95 | -2.07 | 1.45 | 0.60 | 1.00 | 30,942.22 | very small |
| Phytoparasitica | -0.24 | 0.95 | -2.04 | 1.41 | 0.61 | 1.00 | 30,868.29 | very small |
| PhytopatholMediterr | -0.19 | 0.95 | -2.04 | 1.54 | 0.58 | 1.00 | 32,950.47 | very small |
| PlantDis | -0.19 | 0.95 | -2.06 | 1.54 | 0.58 | 1.00 | 32,160.84 | very small |
| PlantHealthProgress | -0.15 | 0.95 | -2.02 | 1.58 | 0.57 | 1.00 | 30,622.46 | very small |
| PlantPathol | -0.24 | 0.95 | -2.06 | 1.45 | 0.60 | 1.00 | 33,692.04 | very small |
| RevMexFitopatol | -0.22 | 0.95 | -2.09 | 1.48 | 0.60 | 1.00 | 31,375.00 | very small |
| TropPlantPathol | 0.66 | 0.95 | -1.13 | 2.27 | 0.77 | 1.00 | 30,524.87 | small |
| VirolJ | -0.14 | 0.95 | -2.03 | 1.59 | 0.56 | 1.00 | 31,373.40 | very small |

Table S2: Full description of the model that was fit to scoring data, which were used to evaluate the effect of journal title on data availability for 450 papers published in 21 plant pathology journals or plant pathology focused articles from other specialized journals. Scoring for 'Data Availability' was scored 0-3 where, '0' was 'Not available or not mentioned in the publication'; '1' was 'Available upon request to the author; '2' was 'Online, but inconvenient or non-permanent (e.g., login needed, paywall, FTP server, personal lab website that may disappear, or may have already disappeared)'; and '3' was 'Freely available online to anonymous users for foreseeable future (e.g., archived using Zenodo, dataverse or university library or some other proper archiving system)'; 'NA' indicates that no data were generated, e.g., a methods paper. We fit 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 | Effect Size |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| AustralasPlantPath | 0.61 | 0.95 | -0.69 | 1.81 | 0.83 | 1.00 | 24,307.12 | small |
| CanJPlantPathol | 0.30 | 0.95 | -0.85 | 1.33 | 0.70 | 1.00 | 22,129.24 | very small |
| CropProt | -1.24 | 0.95 | -2.75 | 0.01 | 0.97 | 1.00 | 22,755.47 | small |
| EurJPlantPathol | 0.07 | 0.95 | -1.14 | 1.15 | 0.55 | 1.00 | 19,980.53 | very small |
| ForestPathol | -0.10 | 0.95 | -1.28 | 0.95 | 0.57 | 1.00 | 20,503.42 | very small |
| JPhytopathol | -0.26 | 0.95 | -1.24 | 0.62 | 0.71 | 1.00 | 18,975.15 | very small |
| JPlantPathol | 0.15 | 0.95 | -0.97 | 1.14 | 0.61 | 1.00 | 20,369.20 | very small |
| MolPlantMicroIn | 0.54 | 0.95 | -0.45 | 1.46 | 0.86 | 1.00 | 18,290.43 | small |
| MolPlantPathol | 0.92 | 0.95 | 0.10 | 1.73 | 0.98 | 1.00 | 17,388.41 | small |
| Nematology | -0.19 | 0.95 | -1.49 | 0.94 | 0.62 | 1.00 | 23,521.11 | very small |
| PhysiolMolPlantP | 0.58 | 0.95 | -0.38 | 1.47 | 0.89 | 1.00 | 19,046.97 | small |
| Phytoparasitica | -0.47 | 0.95 | -1.70 | 0.59 | 0.80 | 1.00 | 22,297.86 | very small |
| PhytopatholMediterr | 1.67 | 0.95 | 0.77 | 2.57 | 1.00 | 1.00 | 19,527.70 | medium |
| PlantDis | -1.26 | 0.95 | -2.74 | -0.03 | 0.98 | 1.00 | 24,518.10 | medium |
| PlantHealthProgress | -0.59 | 0.95 | -2.01 | 0.64 | 0.82 | 1.00 | 22,227.55 | small |
| PlantPathol | -0.09 | 0.95 | -1.09 | 0.82 | 0.57 | 1.00 | 19,747.43 | very small |
| RevMexFitopatol | -1.14 | 0.95 | -2.61 | 0.13 | 0.96 | 1.00 | 25,620.08 | small |
| TropPlantPathol | 0.36 | 0.95 | -0.77 | 1.38 | 0.74 | 1.00 | 22,568.90 | very small |
| VirolJ | 0.87 | 0.95 | -0.02 | 1.73 | 0.97 | 1.00 | 19,005.73 | small |

## Supplementary Figure

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. Scoring for ‘Code Availability’ was scored 0-3 where, ‘0’ was ‘Not available or not mentioned in the publication’; ‘1’ was ‘Available upon request to the author; ’2’ was ‘Online, but inconvenient or non-permanent (e.g., login needed, paywall, FTP server, personal lab website that may disappear, or may have already disappeared)’; and ‘3’ was ‘Freely available online to anonymous users for foreseeable future (e.g., archived using Zenodo, dataverse or university library or some other proper archiving system)’; ‘NA’ indicates that no code was created to conduct the work that was detectable. And the scoring for ‘Data Availability’ was scored 0-3 where, ‘0’ was ‘Not available or not mentioned in the publication’; ‘1’ was ‘Available upon request to the author; ’2’ was ‘Online, but inconvenient or non-permanent (e.g., login needed, paywall, FTP server, personal lab website that may disappear, or may have already disappeared)’; and ‘3’ was ‘Freely available online to anonymous users for foreseeable future (e.g., archived using Zenodo, dataverse or university library or some other proper archiving system)’; ‘NA’ indicates that no data were generated, e.g., a methods paper.

Figure S1: 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. Scoring for ‘Code Availability’ was scored 0-3 where, ‘0’ was ‘Not available or not mentioned in the publication’; ‘1’ was ‘Available upon request to the author; ’2’ was ‘Online, but inconvenient or non-permanent (e.g., login needed, paywall, FTP server, personal lab website that may disappear, or may have already disappeared)’; and ‘3’ was ‘Freely available online to anonymous users for foreseeable future (e.g., archived using Zenodo, dataverse or university library or some other proper archiving system)’; ‘NA’ indicates that no code was created to conduct the work that was detectable. And the scoring for ‘Data Availability’ was scored 0-3 where, ‘0’ was ‘Not available or not mentioned in the publication’; ‘1’ was ‘Available upon request to the author; ’2’ was ‘Online, but inconvenient or non-permanent (e.g., login needed, paywall, FTP server, personal lab website that may disappear, or may have already disappeared)’; and ‘3’ was ‘Freely available online to anonymous users for foreseeable future (e.g., archived using Zenodo, dataverse or university library or some other proper archiving system)’; ‘NA’ indicates that no data were generated, e.g., a methods paper.