APPENDICES TO
Bayesian hierarchical modelling: incorporating spatial information in water resources assessment and accounting

Numbering of all figures and equations below continues from the main text of Chiu and Lehmann (2011).

APPENDIX A

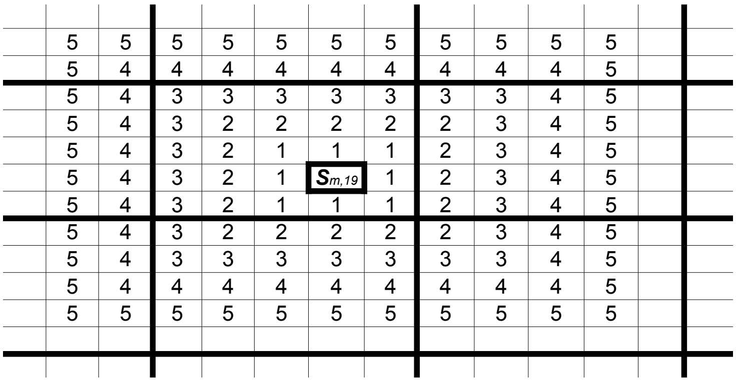

Figure 7. Spatial dependence structure for $s_{m,19}$ (the 19th AWAP pixel inside the mth AMSR-E pixel). Bold lines outside of $s_{m,19}$ delimit AMSR-E pixels. An AWAP pixel showing integer q is a neighbour along the qth nearest rectangular border surrounding $s_{m,19}$ (rectangular adjacency); such a pixel exhibits spatial dependence with $s_{m,19}$, for q=1, ..., 5. Thus, s_{mr} (r=1, ..., 25) is assumed to exhibit non-trivial spatial dependence with an adjacent AMSR-E pixel, regardless of the location of s_{mr} inside an AMSR-E pixel.

We impose a "5th order" CAR structure on $\boldsymbol{\phi}$. Specifically, writing

$$\boldsymbol{\phi} = \{\phi_{mr}\} = (\phi_{1,1}, \ldots, \phi_{1,25},\ \phi_{2,1}, \ldots, \phi_{2,25}, \ldots, \phi_{M,1}, \ldots, \phi_{M,25})' \equiv (\phi_1, \phi_2, \ldots, \phi_{25M})'$$

and $\boldsymbol{\phi}_{-i}$ to denote $\boldsymbol{\phi}$ with ϕ_i removed, we assume

$$\phi_i \left| \boldsymbol{\phi}_{-i}, \tau^2 \right. \sim \mathrm{N}\left(\frac{1}{w_{i+}} \sum_{i' \neq i} w_{ii'} \phi_{i'}, \frac{\tau^2}{w_{i+}} \right) \tag{7}$$

where the dependence structure of $\boldsymbol{\phi}$ (see Fig. 7) is imposed by

$$w_{ii'} = \begin{cases} e^{-(q-1)} & \text{if } s_{i'} \text{ is a neighbour of } s_i \text{ along its } q\text{th nearest rectangular border},\ q = 1, \ldots, 5 \\ 0 & \text{otherwise} \end{cases}$$

with w_{i+}=$\sum_{i'=1}^{25M} w_{ii'}$. Note that w_{ii}=0 for all i, and $w_{ii'}$ decays exponentially over the dependence neighborhood of s_i. Taking $\sum_i \phi_i$=0, the log-likelihood of $\boldsymbol{\phi}$ given τ^2 is (Banerjee *et al.*, 2004)

$$\begin{aligned} \log f(\boldsymbol{\phi}|\tau^2) &= -\frac{1}{2} \log \left| \tau^2 (\mathbb{D} - \mathbb{W})^{-1} \right| - \frac{1}{2\tau^2} \boldsymbol{\phi}' (\mathbb{D} - \mathbb{W}) \boldsymbol{\phi} + \text{constant} \\ &= -\frac{25M - 1}{2} \log \tau^2 - \frac{1}{2\tau^2} \boldsymbol{\phi}' (\mathbb{D} - \mathbb{W}) \boldsymbol{\phi} + \text{constant} \end{aligned} \tag{8}$$

where $\mathbb{W}$=$[w_{ii'}]$ is the symmetric adjacency matrix corresponding to Fig. 7 and $\mathbb{D}$=diag$\{w_{1+}, \ldots, w_{25M+}\}$. Note that $\mathbb{D} - \mathbb{W}$ has rank $25M-1$.

APPENDIX B

Let $\boldsymbol{v}^{\text{obs}}$ be the vector of $\{v_m\}$ where observed, and similarly for $\boldsymbol{g}^{\text{obs}}$ and $\boldsymbol{p}^{\text{obs}}$. Also let $\boldsymbol{v}^{\text{mis}}$ be the vector of $\{v_m\}$ where unobserved, and similarly for $\boldsymbol{p}^{\text{mis}}$ and for $\boldsymbol{g}^{\text{mis}}$ (g^{mis}_{mrk} exists only if $K_{mr}\geq 1$). Then, the posterior distribution for our model parameters is

$$\begin{aligned} & f\left(\boldsymbol{g}^{\text{mis}}, \boldsymbol{v}^{\text{mis}}, \boldsymbol{p}^{\text{mis}}, \boldsymbol{\psi}, \boldsymbol{\phi}, \boldsymbol{\Omega} \middle| \, \boldsymbol{g}^{\text{obs}}, \boldsymbol{v}^{\text{obs}}, \boldsymbol{p}^{\text{obs}}\right) \\ \propto \; & f(\boldsymbol{g}, \boldsymbol{v}, \boldsymbol{p}, \boldsymbol{\psi}, \boldsymbol{\phi}, \boldsymbol{\Omega}) \\ = \; & f\left(\boldsymbol{g} | \, \boldsymbol{\psi}, \{\sigma^2_m\}\right) \, f\left(\boldsymbol{v} | \, \alpha_1, \boldsymbol{\psi}, \sigma^2_\delta\right) \, f\left(\boldsymbol{\psi} | \, \boldsymbol{\beta}, \boldsymbol{p}, \sigma^2_\eta\right) \, f\left(\boldsymbol{p} | \, \boldsymbol{\phi}, \boldsymbol{\gamma}, \sigma^2_\zeta\right) \, f\left(\boldsymbol{\phi} | \, \tau^2\right) \, f(\boldsymbol{\Omega}) \end{aligned} \tag{9}$$

where $f(a|b)$ denotes the conditional distribution of a given b, and $f(\boldsymbol{\Omega})$ is the prior from (6) for $\boldsymbol{\Omega} = [\alpha_1, \boldsymbol{\beta}, \boldsymbol{\gamma}, \{\sigma^2_m\}, \sigma^2_\delta, \sigma^2_\eta, \sigma^2_\zeta, \tau^2]'$.

Specifically, for the priors of $\alpha_1, \boldsymbol{\beta}$, and $\boldsymbol{\gamma}$, we ensure reasonable diffuseness by taking a=0.01 in (6). This reflects the fact that we have very little idea of the inherent variability of these parameters. For the same reason, we take a_1=a_2=1 for the priors of σ^2_ζ and τ^2. In contrast, we wish to impose some constraints on the priors of $\sigma^2_m, \sigma^2_\delta$, and σ^2_η. The reason is as follows. Linearity is reasonable for (1) and (2) which involve measurement devices, and also for (3) which is supported by exploratory data analyses. Yet, very large values of the error term in (1)–(3) sampled while running the MCMC could produce negative values of $g^{\text{mis}}, v^{\text{mis}}$, and ψ. (This is not of practical concern here for p in (4) due to its spatial smoothness and lack of missing data.) Thus, we take a naive empirical approach to set a_1 and a_2 for these three variance parameters (see Appendix C).

With values of a, a_1, and a_2 pre-specified, below we work out the full conditional distributions of all model parameters based on (9).

First, let $\mathcal{M}_g$={(m,r,k): $m\in\mathcal{S}$ and g_{mrk} is unobserved}, $\mathcal{M}_v$={m: v_m is unobserved}, and $\mathcal{M}_p$={(m,r): p_{mr} is unobserved}. Then, from (1), (2), and (4), we have

$$f\left(\boldsymbol{g}^{\text{mis}} \middle| \, \boldsymbol{.}\right) = \prod_{\substack{m=1 \\ (m,r,k)\in\mathcal{M}_g}}^{M} \prod_{r=1}^{25} \prod_{k=1}^{K_{mr}} \text{N}\left(\psi_{mr}, \sigma^2_m\right) , \tag{10}$$

$$f\left(\boldsymbol{v}^{\text{mis}} \middle| \, \boldsymbol{.}\right) = \prod_{m\in\mathcal{M}_v} \text{N}\left(\alpha_1 \overline{\psi}_m, \sigma^2_\delta\right) , \tag{11}$$

$$f\left(\boldsymbol{p}^{\text{mis}} \middle| \, \boldsymbol{.}\right) = \prod_{\substack{m \\ (m,r)\in\mathcal{M}_p}} \prod_{r} \text{N}\left(h_{\boldsymbol{\gamma}}(\boldsymbol{x}_{mr}) + \phi_{mr}, \sigma^2_\zeta\right) \tag{12}$$

Next, from (3) and (9), we have

$$\begin{aligned} & f(\boldsymbol{\psi}|\boldsymbol{.}) \propto f\left(\boldsymbol{g} | \, \boldsymbol{\psi}, \{\sigma^2_m\}\right) \, f\left(\boldsymbol{v} | \, \alpha_1, \boldsymbol{\psi}, \sigma^2_\delta\right) \, f\left(\boldsymbol{\psi} | \, \boldsymbol{\beta}, \boldsymbol{p}, \sigma^2_\eta\right) \\ \Longrightarrow \quad & \log f(\boldsymbol{\psi}|\boldsymbol{.}) = -\frac{1}{2} \sum_{m=1}^{M} \left\{ \sum_{r=1}^{25} \sum_{k=1}^{K_{mr}} \frac{(g_{mrk} - \psi_{mr})^2}{\sigma^2_m} + \frac{(v_m - (\alpha_1/25)\sum_{r=1}^{25} \psi_{mr})^2}{\sigma^2_\delta} + \right. \\ & \left. \sum_{r=1}^{25} \frac{(\psi_{mr} - \beta_0 - \beta_1 p_{mr})^2}{\sigma^2_\eta} \right\} + \text{constant}. \end{aligned}$$

Due to conjugacy and after some algebra, we have

$$\psi_{mr}|\boldsymbol{.} \sim \text{N}\left(\frac{c_{1mr}}{c_{2mr}}, \frac{1}{c_{2mr}}\right) \tag{13}$$

where

$$c_{1mr} = \frac{\alpha_1}{25\sigma^2_\delta} \left(v_m - \frac{\alpha_1}{25} \sum_{r'\neq r} \psi_{mr'} \right) + \frac{1}{\sigma^2_m} \sum_k g_{mrk} + \frac{\beta_0 + \beta_1 p_{mr}}{\sigma^2_\eta} ,$$

$$c_{2mr} = \frac{\alpha_1^2}{25^2\sigma_\delta^2} + \frac{K_{mr}}{\sigma_m^2} + \frac{1}{\sigma_\eta^2}\,.$$

Note that for fixed m, the elements of $\{\psi_{mr}\}$ are dependent, but the sets $\{\psi_{mr}\}$ and $\{\psi_{m'r}\}$ are independent for $m{\neq}m'$.

Next, from (7) and (9), and reindexing (m, r) as i, we have

$$\begin{aligned} & f\left(\phi_i|\,\boldsymbol{.}\right) \propto f\left(p_i\left|\phi_i, \boldsymbol{\gamma}, \sigma_\zeta^2\right.\right)\, f\left(\phi_i|\,\boldsymbol{\phi}_{-i}, \tau^2\right) \\ \implies \quad & \log f\left(\phi_i|\,\boldsymbol{.}\right) = -\frac{1}{2}\left\{\frac{(p_i - h_{\boldsymbol{\gamma}}(\boldsymbol{x}_i) - \phi_i)^2}{\sigma_\zeta^2} + \frac{[\phi_i - (1/w_{i+})\sum_{i'} w_{ii'}\phi_{i'}]^2}{(\tau^2/w_{i+})}\right\} + \text{constant.} \end{aligned}$$

Due to conjugacy and after some algebra, we have

$$\phi_i|\boldsymbol{.} \sim \mathrm{N}\left(\frac{c_{1i}}{c_{2i}}, \frac{1}{c_{2i}}\right) \tag{14}$$

where

$$c_{1i} = \frac{p_i - h_{\boldsymbol{\gamma}}(\boldsymbol{x}_i)}{\sigma_\zeta^2} + \frac{1}{\tau^2}\sum_{i'} w_{ii'}\phi_{i'}\,, \qquad c_{2i} = \frac{1}{\sigma_\zeta^2} + \frac{w_{i+}}{\tau^2}\,.$$

Similarly, from (6) and (9), we have

$$\alpha_1|\boldsymbol{.} \sim \mathrm{N}\left(\frac{c_1}{c_2}, \frac{1}{c_2}\right) \tag{15}$$

where $c_1 = (1/\sigma_\delta^2)\sum_m v_m\overline{\psi}_m + 3\sqrt{a}$ and $c_2 = (1/\sigma_\delta^2)\sum_m \overline{\psi}_m^2 + a$.

Also from (6) and (9) we have

$$\begin{aligned} & f(\boldsymbol{\beta}|\boldsymbol{.}) \propto f\left(\boldsymbol{\psi}|\boldsymbol{\beta}, \boldsymbol{p}, \sigma_\eta^2\right)\, f(\boldsymbol{\beta}) \\ \implies \quad & \log f(\boldsymbol{\beta}|\boldsymbol{.}) = -\frac{1}{2}\left\{(\boldsymbol{\psi} - \mathbb{P}\boldsymbol{\beta})'\left(\sigma_\eta^{-2}\mathbb{I}_{25M}\right)(\boldsymbol{\psi} - \mathbb{P}\boldsymbol{\beta}) + (\boldsymbol{\beta} - \boldsymbol{\mu})'\left(a\mathbb{I}_2\right)(\boldsymbol{\beta} - \boldsymbol{\mu})\right\} + \text{constant} \\ & \text{where} \quad \mathbb{P} = \begin{bmatrix}\mathbb{P}_1 \\ \vdots \\ \mathbb{P}_M\end{bmatrix}, \quad \mathbb{P}_m = \begin{bmatrix}1 & p_{m,1} \\ \vdots & \vdots \\ 1 & p_{m,25}\end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix}0 \\ 3a^{-1/2}\end{bmatrix}, \end{aligned}$$

and $\mathbb{I}_n$ is the $n{\times}n$ identity matrix. Due to conjugacy and after some algebra, we have

$$\boldsymbol{\beta}|\boldsymbol{.} \sim \mathrm{BVN}\left(\mathbb{C}_4^{-1}\boldsymbol{c}_3, \mathbb{C}_4^{-1}\right) \tag{16}$$

where $\boldsymbol{c}_3 = (1/\sigma_\eta^2)\mathbb{P}'\boldsymbol{\psi} + a\boldsymbol{\mu}$ and $\mathbb{C}_4 = (1/\sigma_\eta^2)\mathbb{P}'\mathbb{P}{+}a\mathbb{I}_2$.

Similarly, we have

$$f(\boldsymbol{\gamma}|\boldsymbol{.}) \propto f\left(\boldsymbol{p}|\,\boldsymbol{\phi}, \boldsymbol{\gamma}, \sigma_\zeta^2\right)\, f(\boldsymbol{\gamma}) \quad \implies \quad \boldsymbol{\gamma}|\boldsymbol{.} \sim \mathrm{MVN}\left(\mathbb{C}_6^{-1}\boldsymbol{c}_5, \mathbb{C}_6^{-1}\right) \tag{17}$$

where $\boldsymbol{c}_5{=}(1/\sigma_\zeta^2)\mathbb{H}'(\boldsymbol{p}{-}\boldsymbol{\phi})$, $\mathbb{C}_6{=}(1/\sigma_\zeta^2)\mathbb{H}'\mathbb{H}{+}a\mathbb{I}_5$,

$$\mathbb{H} = \begin{bmatrix}\mathbb{H}_1 \\ \vdots \\ \mathbb{H}_M\end{bmatrix}, \quad \mathbb{H}_m = \begin{bmatrix}\boldsymbol{h}'_{m,1} \\ \vdots \\ \boldsymbol{h}'_{m,25}\end{bmatrix}, \quad \boldsymbol{h}_{mr} = \begin{bmatrix}x_{mr1} \\ x_{mr2} \\ x_{mr1}x_{mr2} \\ x_{mr1}^2 \\ x_{mr2}^2\end{bmatrix}.$$

Next, from (6), (8), and (9), we have

$$f\left(\tau^2\middle|\bullet\right) \propto f\left(\phi\middle|\tau^2\right)\, f\left(\tau^2\right)$$

$$\implies \quad \log f\left(\tau^2\middle|\bullet\right) = -\left(a_1 + \frac{25M-1}{2} + 1\right)\log\tau^2 - \frac{1}{\tau^2}\left\{a_2 + \frac{\phi'(\mathbb{D}-\mathbb{W})\phi}{2}\right\} .$$

Due to conjugacy, we have

$$\tau^2\big|\bullet \sim \mathrm{IG}\left(a_1 + \frac{25M-1}{2}, a_2 + \frac{\phi'(\mathbb{D}-\mathbb{W})\phi}{2}\right) . \tag{18}$$

Similarly, writing $\boldsymbol{g}_{mr} = [g_{mr1}, \ldots, g_{mrK_{mr}}]'$, we have

$$f\left(\sigma_m^2\middle|\bullet\right) \propto f\left(\boldsymbol{g}_{m,1}, \ldots, \boldsymbol{g}_{m,25}\middle|\psi_{mr}, \sigma_m^2\right)\, f\left(\sigma_m^2\right)$$

$$\implies \quad \sigma_m^2\big|\bullet \sim \mathrm{IG}\left(a_1 + \frac{\sum_{r=1}^{25} K_{mr}}{2}, a_2 + \frac{\sum_{r=1}^{25}\sum_{k=1}^{K_{mr}}\left(g_{mrk} - \psi_{mr}\right)^2}{2}\right) , \tag{19}$$

$$f\left(\sigma_\delta^2\middle|\bullet\right) \propto f\left(\boldsymbol{v}\middle|\alpha_1, \boldsymbol{\psi}, \sigma_\delta^2\right)\, f\left(\sigma_\delta^2\right)$$

$$\implies \quad \sigma_\delta^2\big|\bullet \sim \mathrm{IG}\left(a_1 + \frac{M}{2}, a_2 + \frac{\sum_m\left(v_m - \alpha_1\overline{\psi}_m\right)^2}{2}\right) , \tag{20}$$

$$f\left(\sigma_\eta^2\middle|\bullet\right) \propto f\left(\boldsymbol{\psi}\middle|\boldsymbol{\beta}, \boldsymbol{p}, \sigma_\eta^2\right)\, f\left(\sigma_\eta^2\right)$$

$$\implies \quad \sigma_\eta^2\big|\bullet \sim \mathrm{IG}\left(a_1 + \frac{25M}{2}, a_2 + \frac{\sum_m\sum_r\left(\psi_{mr} - \beta_0 - \beta_1 p_{mr}\right)^2}{2}\right) , \tag{21}$$

$$f\left(\sigma_\zeta^2\middle|\bullet\right) \propto f\left(\boldsymbol{p}\middle|\boldsymbol{\phi}, \boldsymbol{\gamma}, \sigma_\zeta^2\right)\, f\left(\sigma_\zeta^2\right)$$

$$\implies \quad \sigma_\zeta^2\big|\bullet \sim \mathrm{IG}\left(a_1 + \frac{25M}{2}, a_2 + \frac{\sum_m\sum_r\left(p_{mr} - \boldsymbol{h}'_{mr}\boldsymbol{\gamma} - \phi_{mr}\right)^2}{2}\right) . \tag{22}$$

However, note from Appendix C that the values of a_1 and a_2 are not identical for (19), (20), and (21).

Finally, MCMC approximation of (9) is via Gibbs sampling using (10)–(22), with constraint $\sum_m\sum_r \phi_{mr}$=0 to ensure propriety of (8).

APPENDIX C

To impose some constraints on an IG(a_1, a_2) prior, first note that it is right-skewed, with mean $a_2/(a_1-1)$ for a_1>1 and mode $a_2/(a_1+1)$. Arbitrarily, suppose we restrict the skewness such that the mean is less than twice the size of the mode, say, 3/2 times. This leads to a_1=5 and a_2=6×(mode).

The mode can be set to a reaonsable naive empirical estimate of the variance parameter, as follows. (As our research progresses, we intend to investigate the sensitivity of the Bayesian inference to the following values of a_1 and a_2.)

1. Mode for σ_m^2. We can take the mode to be the square of the largest documented instrumentation error for the ground probes inside B_m, available from Young *et al.* (2008). Specifically, for so-called old stations, we take the mode as (2.5 %volume)^2, and so-called new stations, (3.3 %volume)^2. For B_m containing both old and new stations, we take the larger value to be conservative.

2. Mode for σ_δ^2. From (1), we know $\overline{g}_m \approx \overline{\psi}_m$, where $\overline{g}_m = \sum_r\sum_k g_{mrk}/\sum_r K_{mr}$. Thus, naively, we have

   $$v_m = \alpha_1\left(\overline{g}_m + \text{error}\right) + \delta_m$$

   where "error" is naively assumed to be approximately $\overline{\varepsilon}_m$ defined similarly as $\overline{g}_m$. Then, the overall mean-square-error associated with a linear regression of v_m on $\overline{g}_m$ is approximately

   $$\mathrm{MSE} \approx \frac{\alpha_1^2\sigma_m^2}{\sum_r K_{mr}} + \left(\text{mode for } \sigma_\delta^2\right) . \tag{23}$$

The values of the MSE and α_1 in (23) can be approximated from the actual linear regression of v_m on $\overline{g}_m$, and when solving for the mode of σ_δ^2, σ_m can be replaced by 2.5 and $K_{m+}{\equiv}\sum_r K_{mr}$ by $\max_m K_{m+}$ to allow a more conservative solution.

However, on any given day there are often about only 20 pairs of $(\overline{g}_m, v_m)$ on which to approximate (23). Alternatively, invariance of the model statements over time assumed by Chiu (2011) would allow us to fix m and perform the regression over days from late 2002 to 2009. We perform this regression individually for each m for which B_m contains ground stations. We then take the largest of the m values of the mode for σ_δ^2 to determine a_2. We take this very naive approach in handling the batches (over m) of temporal paired data since the approximation here is merely for the purpose of determining a reasonable value of a_2 in the inverse-gamma prior.

3. Mode for σ_η^2. Substitute (3) into (1) to see that

$$\overline{g}_{mr} = \beta_0 + \beta_1 p_{mr} + \eta_{mr} + \overline{\varepsilon}_{mr} \tag{24}$$

where $\overline{g}_{mr} = \sum_k g_{mrk}/K_{mr}$ and similarly for $\overline{\varepsilon}_{mr}$. Thus, the mean-square-error from a linear regression of $\overline{g}_{mr}$ on p_{mr} approximates the overall error in (24). That is,

$$\text{MSE} \approx \left(\text{mode of } \sigma_\eta^2\right) + \frac{\sigma_m^2}{K_{mr}}\,. \tag{25}$$

As we do for determining the mode of σ_δ^2, here we also fix (m, r) and perform a regression on all available daily values of $(p_{mr}, \overline{g}_{mr})$ to approximate the mode of σ_η^2. We then take the largest of these K_{m+} approximated values to determine a_2.

REFERENCES

Banerjee, S., B.P. Carlin, and A.E. Gelfand (2004). *Hierarchical Modeling and Analysis for Spatial Data*. Chapman and Hall/CRC.

Chiu, G.S. (2011). *Statistical Model-Based Alternatives to Integrate and Evaluate Soil Moisture Products*. Technical Report #CMIS2011-EP11722. CSIRO Mathematics, Informatics and Statistics.

Chiu, G.S. and E.A. Lehmann (2011). Bayesian hierarchical modelling: incorporating spatial information in water resources assessment and accounting. *Proceedings of the 2011 International Congress on Modelling and Simulation* (MODSIM 2011).

Young, R., J. Walker, N. Yeoh, A. Smith, K. Ellett, O. Merlin, and A. Western (2008). *Soil Moisture and Meteorological Observations From the Murrumbidgee Catchment.* Department of Civil and Environmental Engineering, The University of Melbourne.