Supplement of Adv. Stat. Clim. Meteorol. Oceanogr., 3, 1–16, 2017
http://www.adv-stat-clim-meteorol-oceanogr.net/3/1/2017/
doi:10.5194/ascmo-3-1-2017-supplement

Supplement of

Reconstruction of spatio-temporal temperature from sparse historical records using robust probabilistic principal component regression

John Tipton et al.

Correspondence to: John Tipton (jtipton25@gmail.com)

S Supplementary Material

S.1 Marginal distribution

Conditional on the unobserved latent principal components, we can write the robust Student's-t pPCR likelihood as

$$
\begin{aligned}
&\left[y_{it}\middle|\mathbf{Z}_i,\boldsymbol{\beta}_t,v_{it}^2\right]\left[\mathbf{X}_i|\mathbf{Z}_i,\sigma\right]\left[\mathbf{Z}_i\right]\\
&\qquad\propto \exp\left\{-\frac{(y_{it}-\mathbf{Z}_i\boldsymbol{\beta}_t)^2}{2v_{it}^2}-\frac{\left(\mathbf{X}_i-\hat{\mathbf{K}}\mathbf{Z}_i\right)'\left(\mathbf{X}_i-\hat{\mathbf{K}}\mathbf{Z}_i\right)}{2\sigma^2}-\frac{\mathbf{Z}_i'\mathbf{Z}_i}{2}\right\}\\
&\qquad\propto \exp\left\{-\frac{1}{2}\mathbf{Z}_i'\left(\frac{\boldsymbol{\beta}_t\boldsymbol{\beta}_t'}{v_{it}^2}+\frac{\hat{\mathbf{K}}'\hat{\mathbf{K}}}{\sigma^2}+\mathbf{I}\right)\mathbf{Z}_i+\mathbf{Z}_i'\left(\frac{\boldsymbol{\beta}_t y_{it}}{v_{it}^2}+\frac{\hat{\mathbf{K}}'\mathbf{X}_i}{\sigma^2}\right)\right\}\\
&\qquad\quad\times\exp\left\{-\frac{y_{it}^2}{2v_{it}^2}-\frac{\mathbf{X}_i'\mathbf{X}_i}{2\sigma^2}\right\}\\
&\qquad\propto \exp\left\{-\frac{1}{2}\left(\mathbf{Z}_i-\boldsymbol{\mu}_{Z_{it}}\right)'\boldsymbol{\Sigma}_{Z_{it}}^{-1}\left(\mathbf{Z}_i-\boldsymbol{\mu}_{Z_{it}}\right)-\frac{y_{it}^2}{2v_{it}^2}-\frac{\mathbf{X}_i'\mathbf{X}_i}{2\sigma^2}+\frac{\boldsymbol{\mu}_{Z_{it}}'\boldsymbol{\Sigma}_{Z_{it}}^{-1}\boldsymbol{\mu}_{Z_{it}}}{2}\right\}
\end{aligned}
$$

where $\boldsymbol{\Sigma}_{Z_{it}}=\left(\frac{\boldsymbol{\beta}_t\boldsymbol{\beta}_t'}{v_{it}^2}+\frac{\hat{\mathbf{K}}'\hat{\mathbf{K}}}{\sigma^2}+\mathbf{I}\right)^{-1}$ and $\boldsymbol{\mu}_{Z_{it}}=\boldsymbol{\Sigma}_{Z_{it}}\left(\frac{\boldsymbol{\beta}_t y_{it}}{v_{it}^2}+\frac{\hat{\mathbf{K}}'\mathbf{X}_i}{\sigma^2}\right)$. Integrating the above equation with respect to $\mathbf{Z}_i$ we get

$$
\begin{aligned}
\left[y_{it}\middle|\boldsymbol{\beta}_t,v_{it}^2,\sigma\right]&\propto\exp\left\{-\frac{y_{it}^2}{2v_{it}^2}+\frac{\boldsymbol{\mu}_{Z_{it}}'\boldsymbol{\Sigma}_{Z_{it}}^{-1}\boldsymbol{\mu}_{Z_{it}}}{2}\right\}\\
&\propto\exp\left\{-\frac{y_{it}^2}{2v_{it}^2}+\frac{\left(\boldsymbol{\Sigma}_{Z_{it}}\left(\frac{y_{it}\boldsymbol{\beta}_t}{v_{it}^2}+\frac{\hat{\mathbf{K}}'\mathbf{X}_i}{\sigma^2}\right)\right)'\boldsymbol{\Sigma}_{Z_{it}}^{-1}\left(\boldsymbol{\Sigma}_{Z_{it}}\left(\frac{y_{it}\boldsymbol{\beta}_t}{v_{it}^2}+\frac{\hat{\mathbf{K}}'\mathbf{X}_i}{\sigma^2}\right)\right)}{2}\right\}\\
&\propto\exp\left\{-\frac{1}{2}y_{it}^2\left(\frac{1}{v_{it}^2}-\frac{\boldsymbol{\beta}_t'\boldsymbol{\Sigma}_{Z_{it}}\boldsymbol{\beta}_t}{(v_{it}^2)^2}\right)+y_{it}\left(\frac{\boldsymbol{\beta}_t'\boldsymbol{\Sigma}_{Z_{it}}\hat{\mathbf{K}}'\mathbf{X}_i}{v_{it}^2\sigma^2}\right)\right\}
\end{aligned}
$$

which is $\mathrm{N}\left(\mu_{y_{it}},\sigma_{y_{it}}^2\right)$ with

$$
\begin{aligned}
\sigma_{y_{it}}^2&=\left(\frac{1}{v_{it}^2}-\frac{\boldsymbol{\beta}_t'\boldsymbol{\Sigma}_{Z_{it}}\boldsymbol{\beta}_t}{(v_{it}^2)^2}\right)^{-1}\\
\mu_{y_{it}}&=\left(\frac{1}{v_{it}^2}-\frac{\boldsymbol{\beta}_t'\boldsymbol{\Sigma}_{Z_{it}}\boldsymbol{\beta}_t}{(v_{it}^2)^2}\right)^{-1}\frac{\boldsymbol{\beta}_t'\boldsymbol{\Sigma}_{Z_{it}}\hat{\mathbf{K}}'\mathbf{X}_i}{v_{it}^2\sigma^2}
\end{aligned}
$$

To further simplify this equation, we use the Morrison-Woodbury matrix inversion formula

$$(\mathbf{A}+\mathbf{UDV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}\left(\mathbf{D}^{-1}+\mathbf{VA}^{-1}\mathbf{U}\right)^{-1}\mathbf{VA}^{-1}.$$

Then, defining $\mathbf{A} = v_{it}^2\mathbf{I}$, $\mathbf{U} = \boldsymbol{\beta}_t'$, $\mathbf{V} = \boldsymbol{\beta}_t$, and $\mathbf{D} = \left(\frac{\hat{\mathbf{K}}'\hat{\mathbf{K}}}{\sigma^2}+\mathbf{I}\right)^{-1}$,

$$\begin{aligned}
\sigma_{y_{it}}^2 &= \left(\frac{\mathbf{I}}{v_{it}^2} - \frac{\boldsymbol{\beta}_t'\boldsymbol{\Sigma}_{Z_{it}}\boldsymbol{\beta}_t}{(v_{it}^2)^2}\right)^{-1} \\
&= \left(\frac{\mathbf{I}}{v_{it}^2} - \frac{\boldsymbol{\beta}_t'}{v_{it}^2}\left(\frac{\boldsymbol{\beta}_t\boldsymbol{\beta}_t'}{v_{it}^2} + \frac{\hat{\mathbf{K}}'\hat{\mathbf{K}}}{\sigma^2}+\mathbf{I}\right)^{-1}\frac{\boldsymbol{\beta}_t}{v_{it}^2}\right)^{-1} \\
&= v_{it}^2 + \boldsymbol{\beta}_t'\left(\frac{\hat{\mathbf{K}}'\hat{\mathbf{K}}}{\sigma^2}+\mathbf{I}\right)^{-1}\boldsymbol{\beta}_t \\
&= v_{it}^2 + \boldsymbol{\beta}_t'\mathbf{M}_p^{-1}\boldsymbol{\beta}_t
\end{aligned}$$

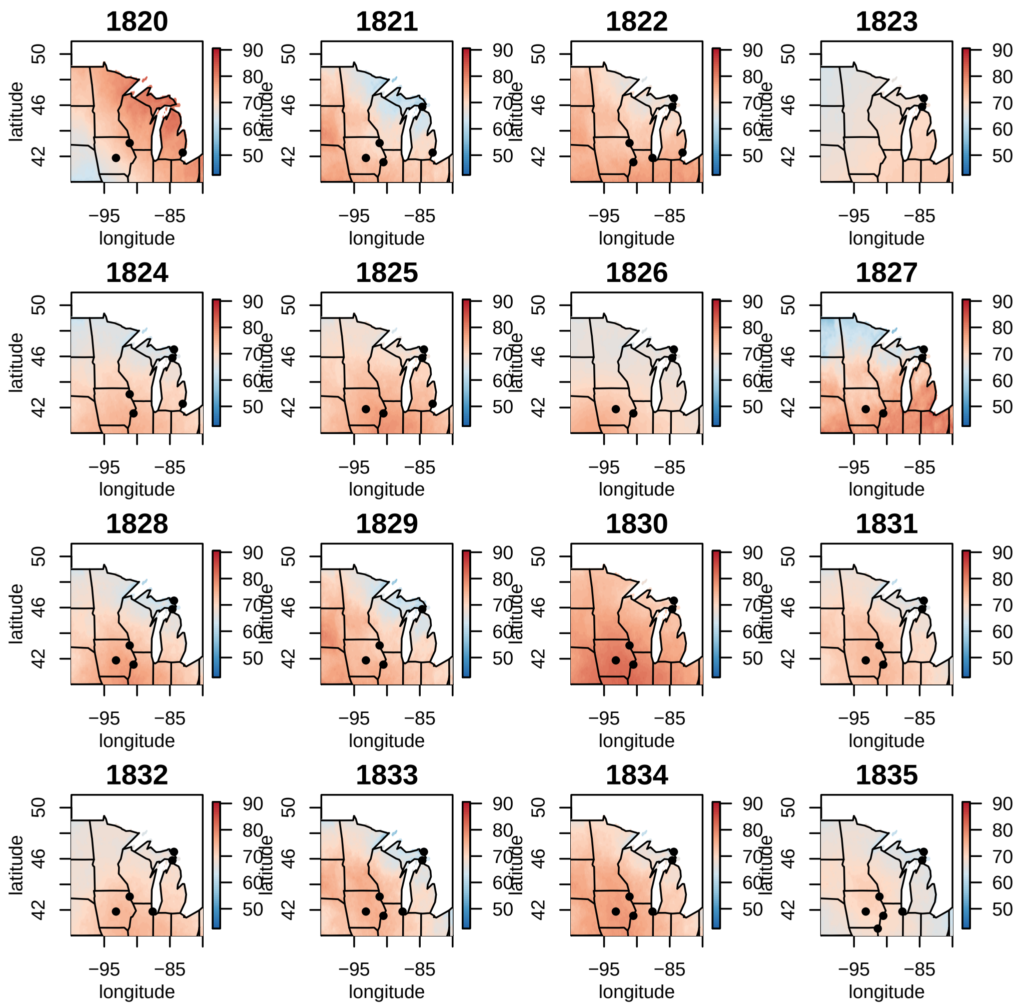
with $\mathbf{M}_p = \frac{\hat{\mathbf{K}}'\hat{\mathbf{K}}}{\sigma^2}+\mathbf{I}$.

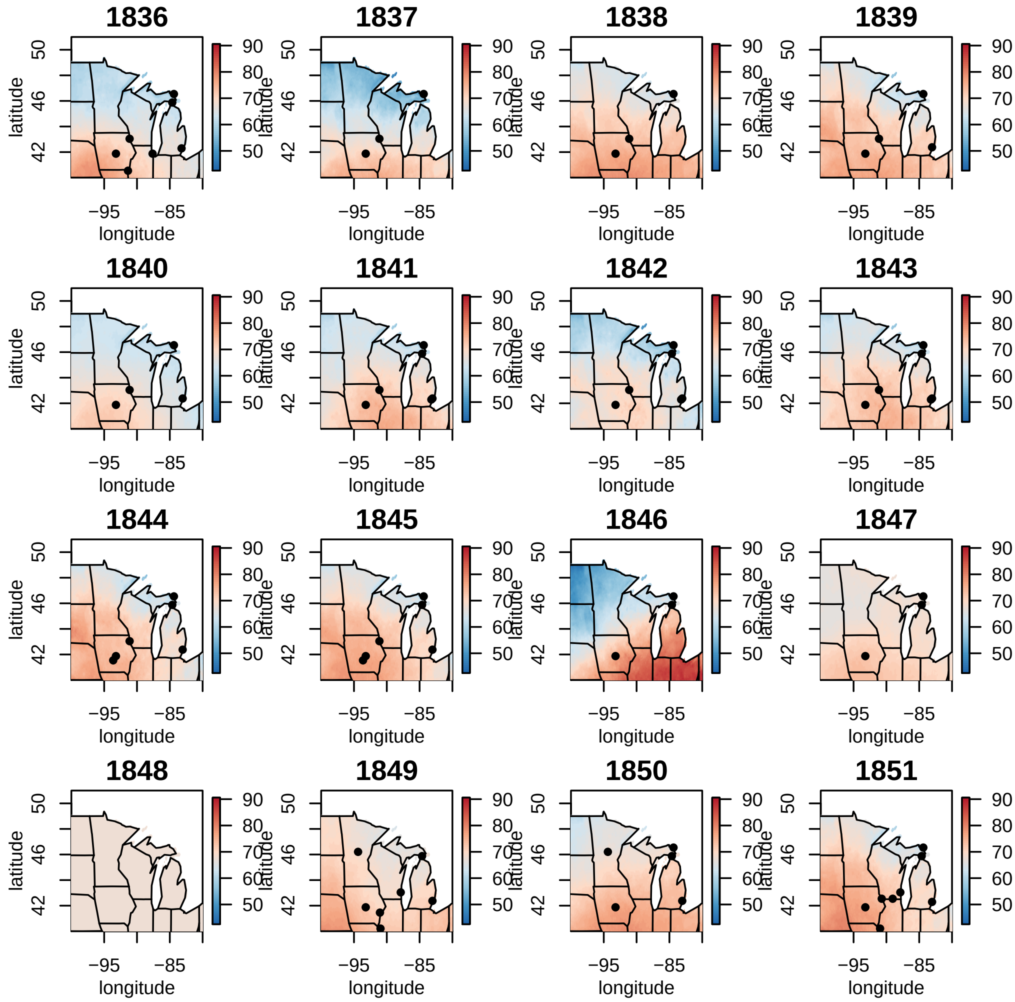
Then, letting $\mathbf{B}_t = \mathbf{M}_p^{-1/2}\boldsymbol{\beta}_t$ the conditional mean $\mu_{y_{it}}$ is

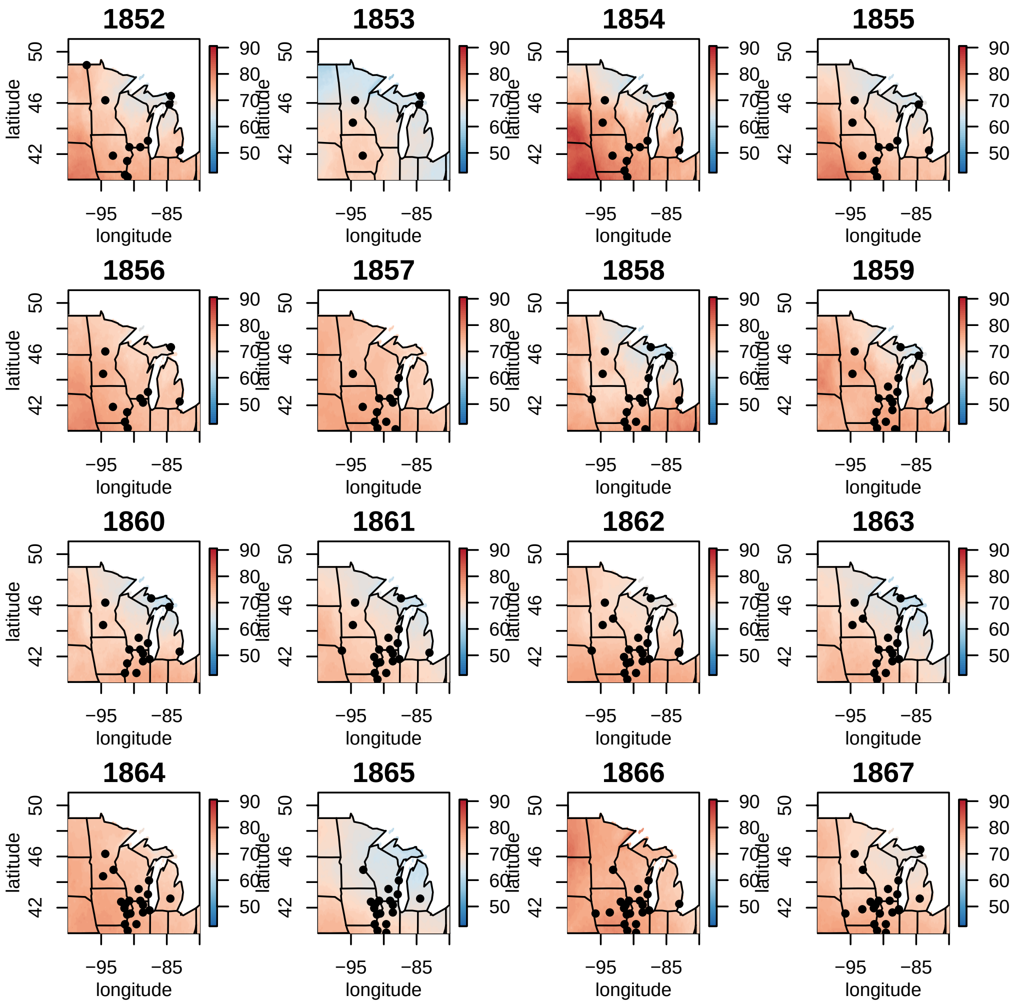
$$\begin{aligned}
\mu_{y_{it}} &= \left(\frac{1}{v_{it}^2} - \frac{\boldsymbol{\beta}_t'\boldsymbol{\Sigma}_{Z_{it}}\boldsymbol{\beta}_t}{(v_{it}^2)^2}\right)^{-1}\left(\frac{\boldsymbol{\beta}_t'\boldsymbol{\Sigma}_{Z_{it}}\hat{\mathbf{K}}'\mathbf{X}_i}{v_{it}^2\sigma^2}\right) \\
&= \left(v_{it}^2 + \boldsymbol{\beta}_t'\mathbf{M}_p^{-1}\boldsymbol{\beta}_t\right)\left(\frac{\boldsymbol{\beta}_t'}{v_{it}^2}\left(\frac{\boldsymbol{\beta}_t\boldsymbol{\beta}_t'}{v_{it}^2} + \frac{\hat{\mathbf{K}}'\hat{\mathbf{K}}}{\sigma^2}+\mathbf{I}\right)^{-1}\frac{\hat{\mathbf{K}}'\mathbf{X}_i}{\sigma^2}\right) \\
&= \left(v_{it}^2 + \mathbf{B}_t'\mathbf{B}_t\right)\left(\frac{\boldsymbol{\beta}_t'}{v_{it}^2}\left(\frac{\boldsymbol{\beta}_t\boldsymbol{\beta}_t'}{v_{it}^2} + \mathbf{M}_p\right)^{-1}\frac{\hat{\mathbf{K}}'\mathbf{X}_i}{\sigma^2}\right) \\
&= \left(v_{it}^2 + \mathbf{B}_t'\mathbf{B}_t\right)\left(\frac{\boldsymbol{\beta}_t'}{v_{it}^2}\left(\mathbf{M}_p^{\frac{1}{2}}\left(\frac{\mathbf{M}_p^{-\frac{1}{2}}\boldsymbol{\beta}_t\boldsymbol{\beta}_t'\mathbf{M}_p^{-\frac{1}{2}}}{v_{it}^2}+\mathbf{I}\right)\mathbf{M}_p^{\frac{1}{2}}\right)^{-1}\frac{\hat{\mathbf{K}}'\mathbf{X}_i}{\sigma^2}\right) \\
&= \left(v_{it}^2 + \mathbf{B}_t'\mathbf{B}_t\right)\frac{\mathbf{B}_t'}{v_{it}^2}\left(\frac{\mathbf{B}_t\mathbf{B}_t'}{v_{it}^2}+\mathbf{I}\right)^{-1}\mathbf{M}_p^{-\frac{1}{2}}\frac{\hat{\mathbf{K}}'\mathbf{X}_i}{\sigma^2} \\
&= \left(\mathbf{B}_t' + \frac{\mathbf{B}_t'\mathbf{B}_t\mathbf{B}_t'}{v_{it}^2}\right)\left(\frac{\mathbf{B}_t\mathbf{B}_t'}{v_{it}^2}+\mathbf{I}\right)^{-1}\mathbf{M}_p^{-\frac{1}{2}}\frac{\hat{\mathbf{K}}'\mathbf{X}_i}{\sigma^2} \\
&= \mathbf{B}_t'\left(\mathbf{I} + \frac{\mathbf{B}_t\mathbf{B}_t'}{v_{it}^2}\right)\left(\frac{\mathbf{B}_t\mathbf{B}_t'}{v_{it}^2}+\mathbf{I}\right)^{-1}\mathbf{M}_p^{-\frac{1}{2}}\frac{\hat{\mathbf{K}}'\mathbf{X}_i}{\sigma^2} \\
&= \mathbf{B}_t'\mathbf{M}_p^{-\frac{1}{2}}\frac{\hat{\mathbf{K}}'\mathbf{X}_i}{\sigma^2} \\
&= \frac{\boldsymbol{\beta}_t'\mathbf{M}_p^{-1}\hat{\mathbf{K}}'\mathbf{X}_i}{\sigma^2}
\end{aligned}$$

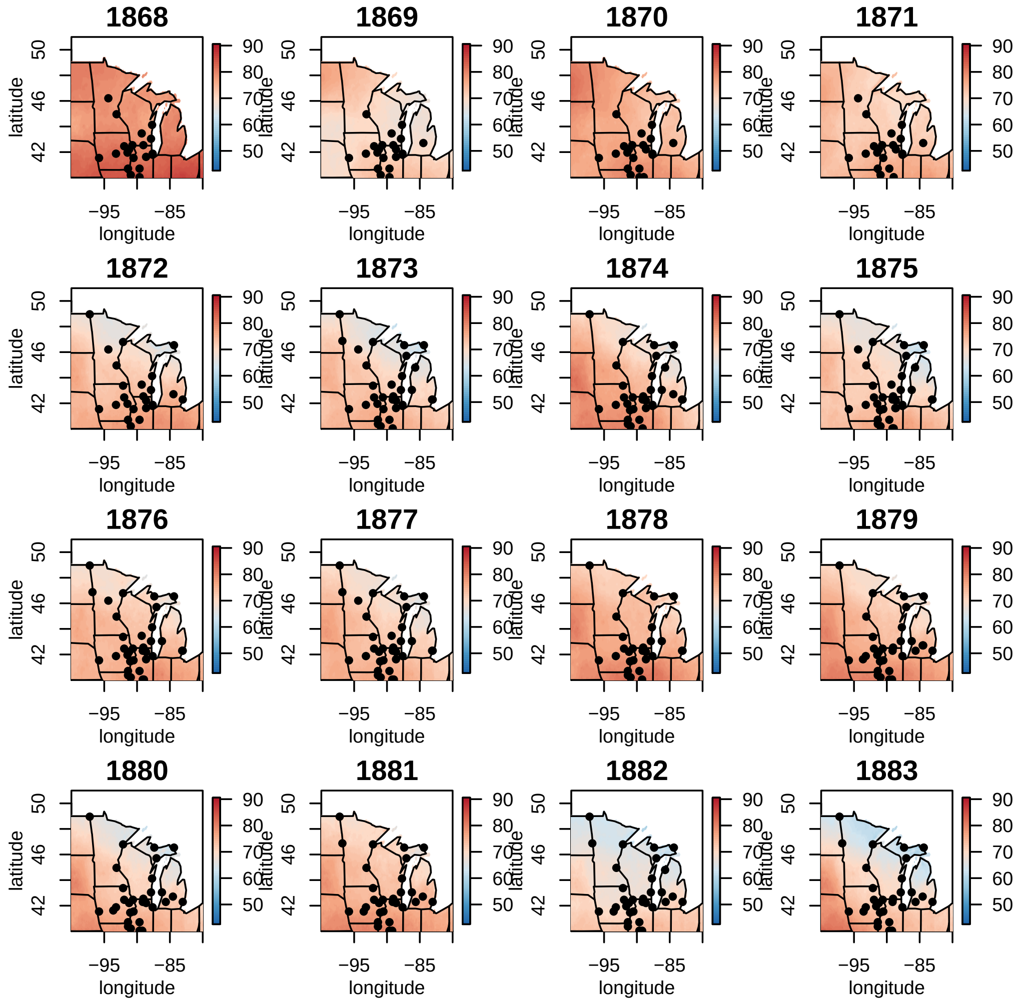
$$y_{it}|\boldsymbol{\beta}_t, v_{it}^2, \sigma \sim \mathrm{N}\left(\frac{\mathbf{X}_i'\hat{\mathbf{K}}\mathbf{M}_p^{-1}\boldsymbol{\beta}_t}{\sigma^2}, v_{it}^2 + \boldsymbol{\beta}_t'\mathbf{M}_p^{-1}\boldsymbol{\beta}_t\right).$$

S.2 Robust PCR Model Posterior Mean July Temperature

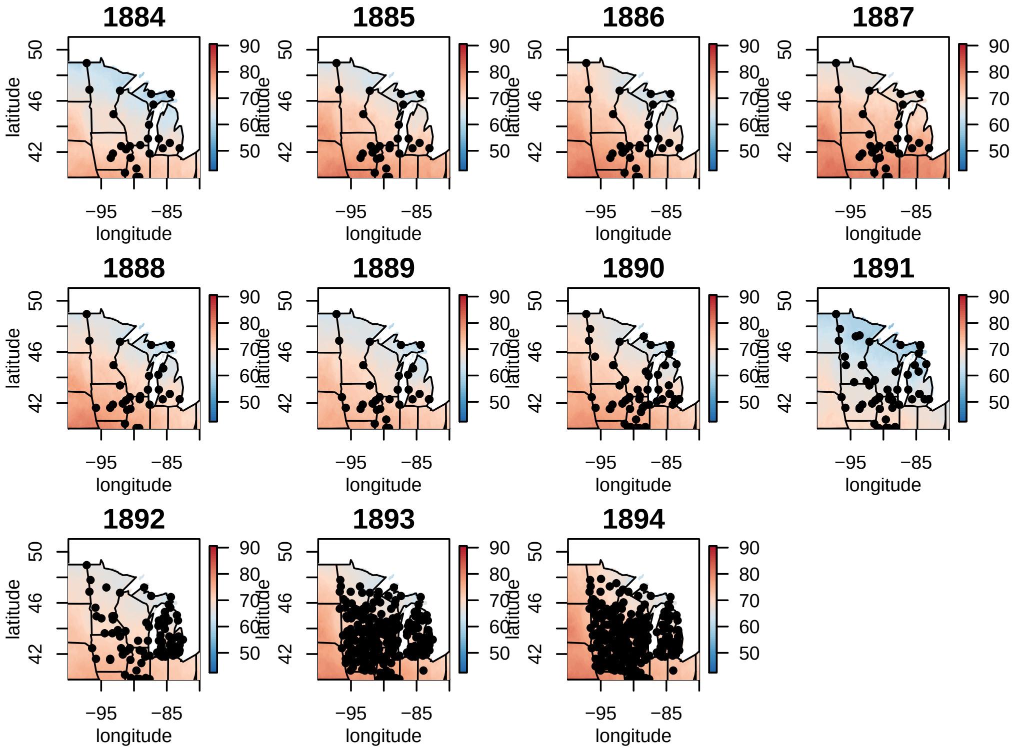



1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
latitude
longitude
42
46
50
−95
−85
50
60
70
80
90

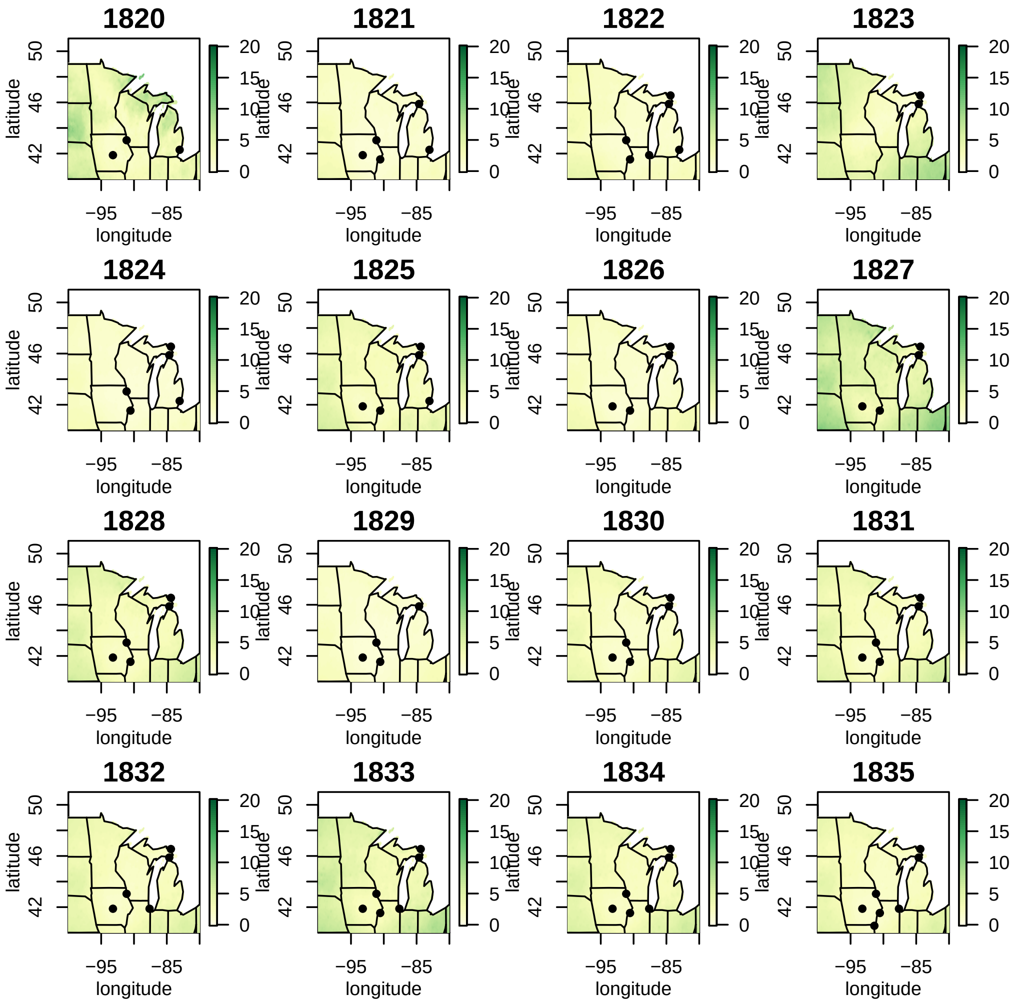

1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
latitude
longitude
50
46
42
−95
−85
90
80
70
60
50

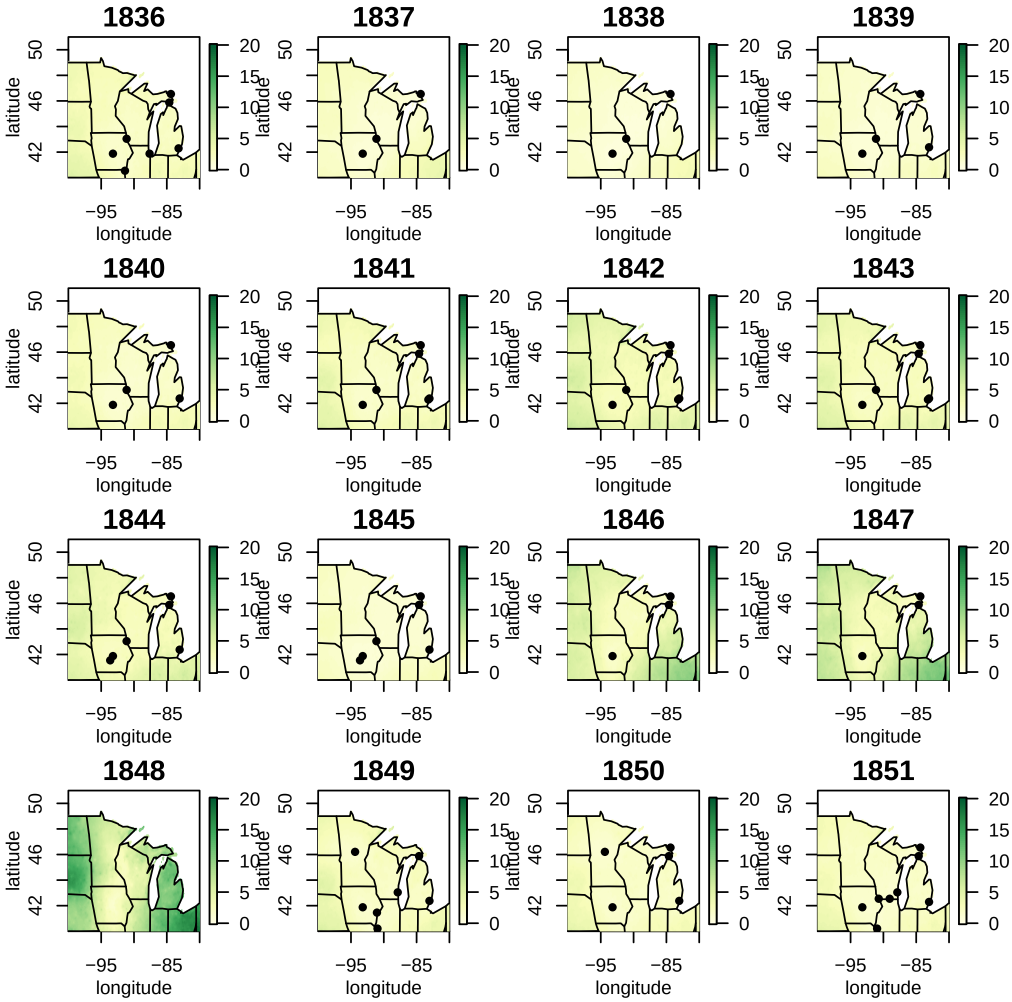


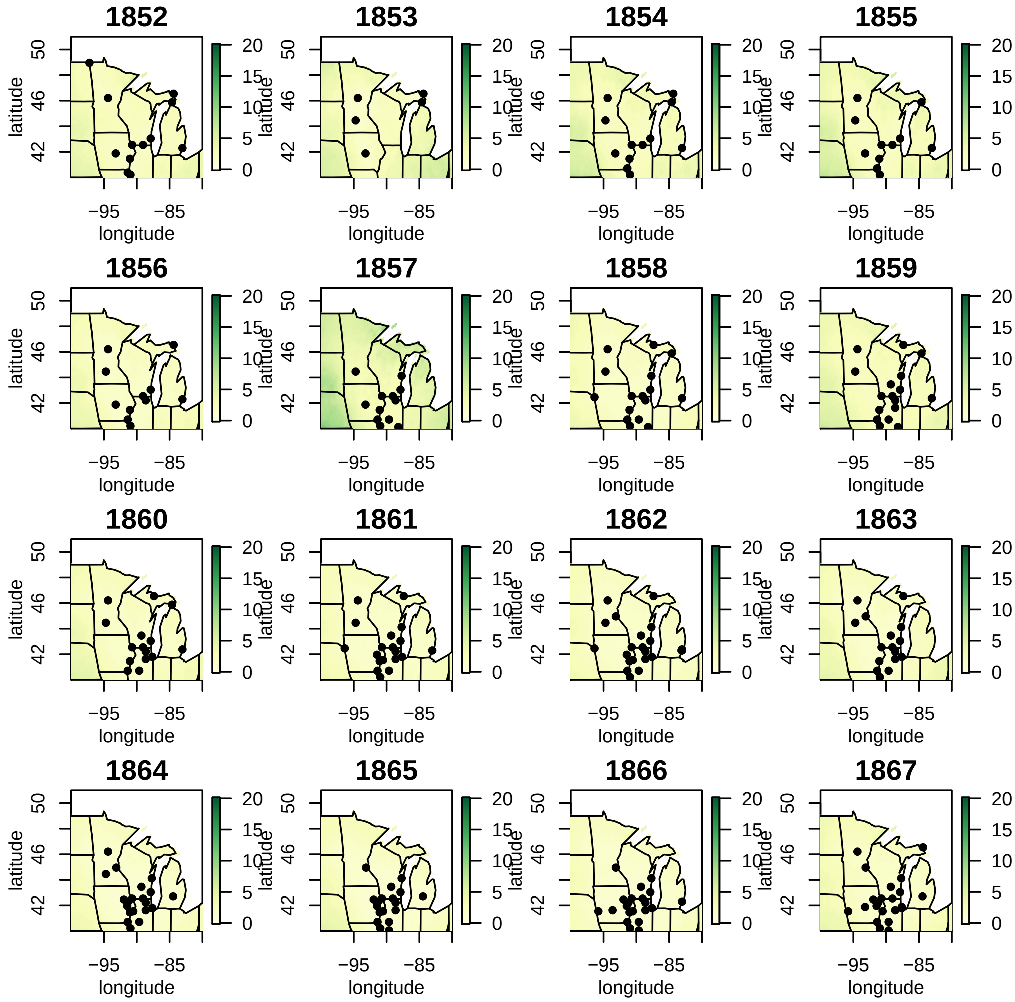
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
latitude
longitude

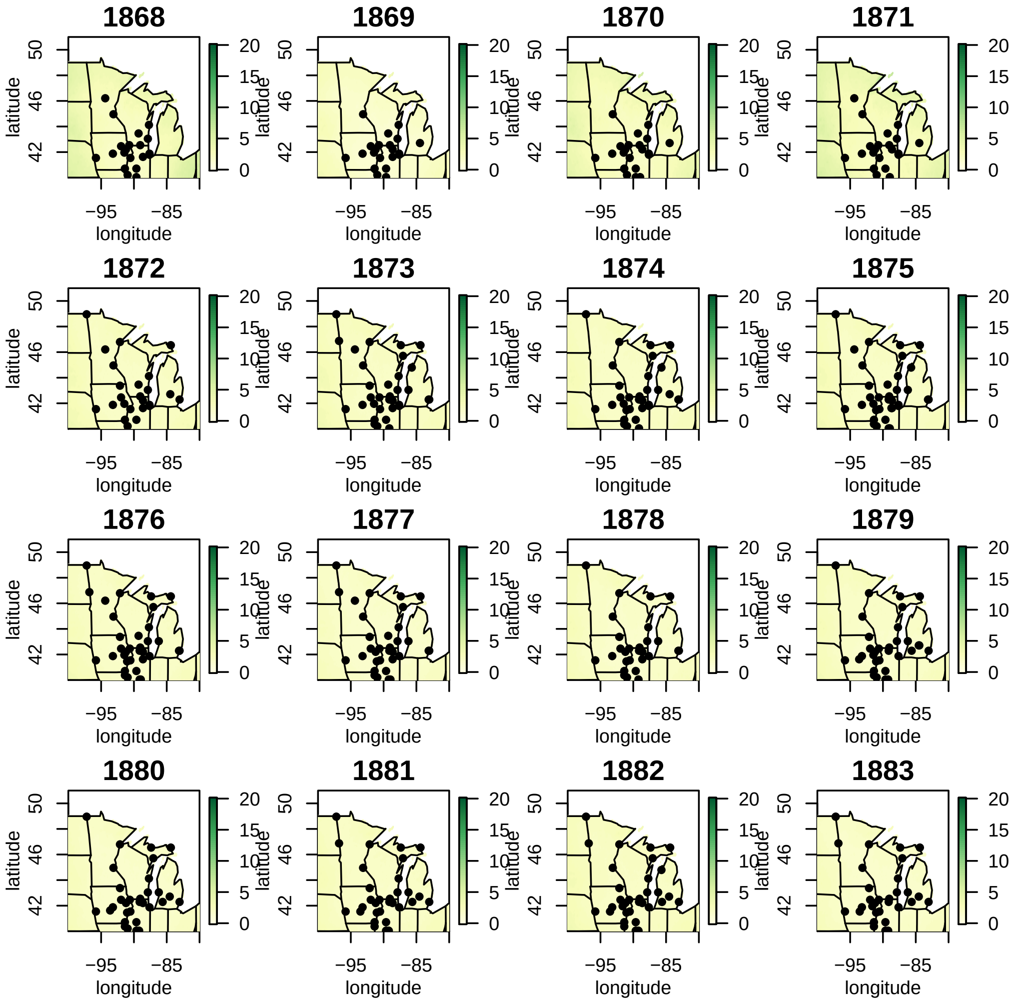

1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
latitude
longitude
50
46
42
−95
−85
90
80
70
60
50

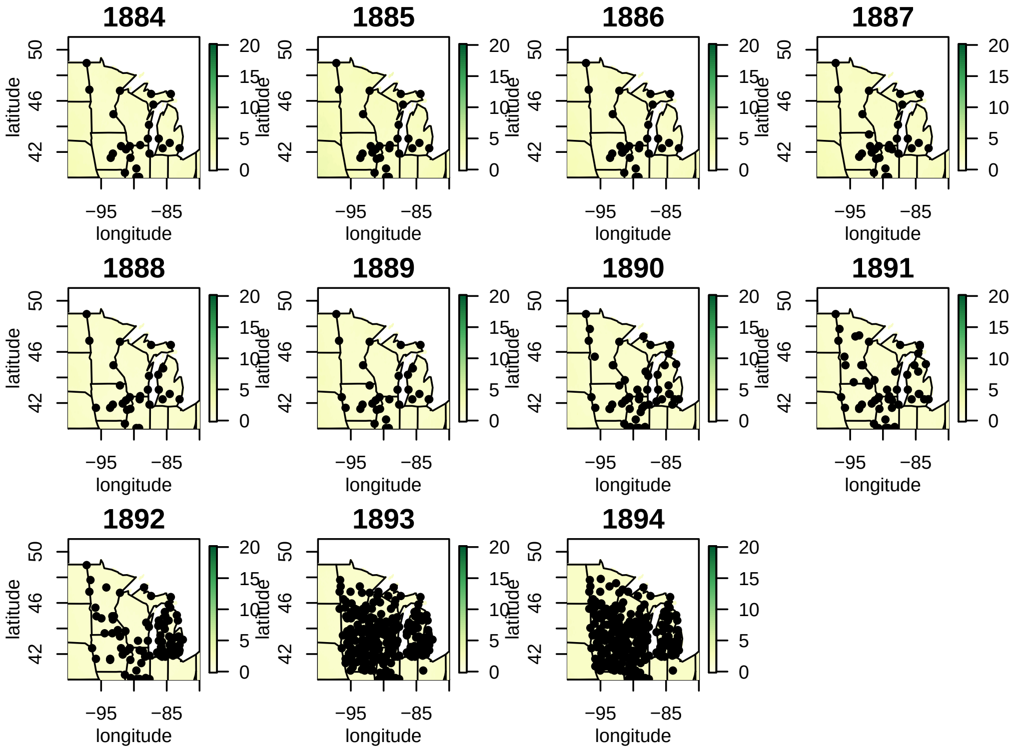
S.3 Robust PCR Model July Temperature Posterior Standard Deviations



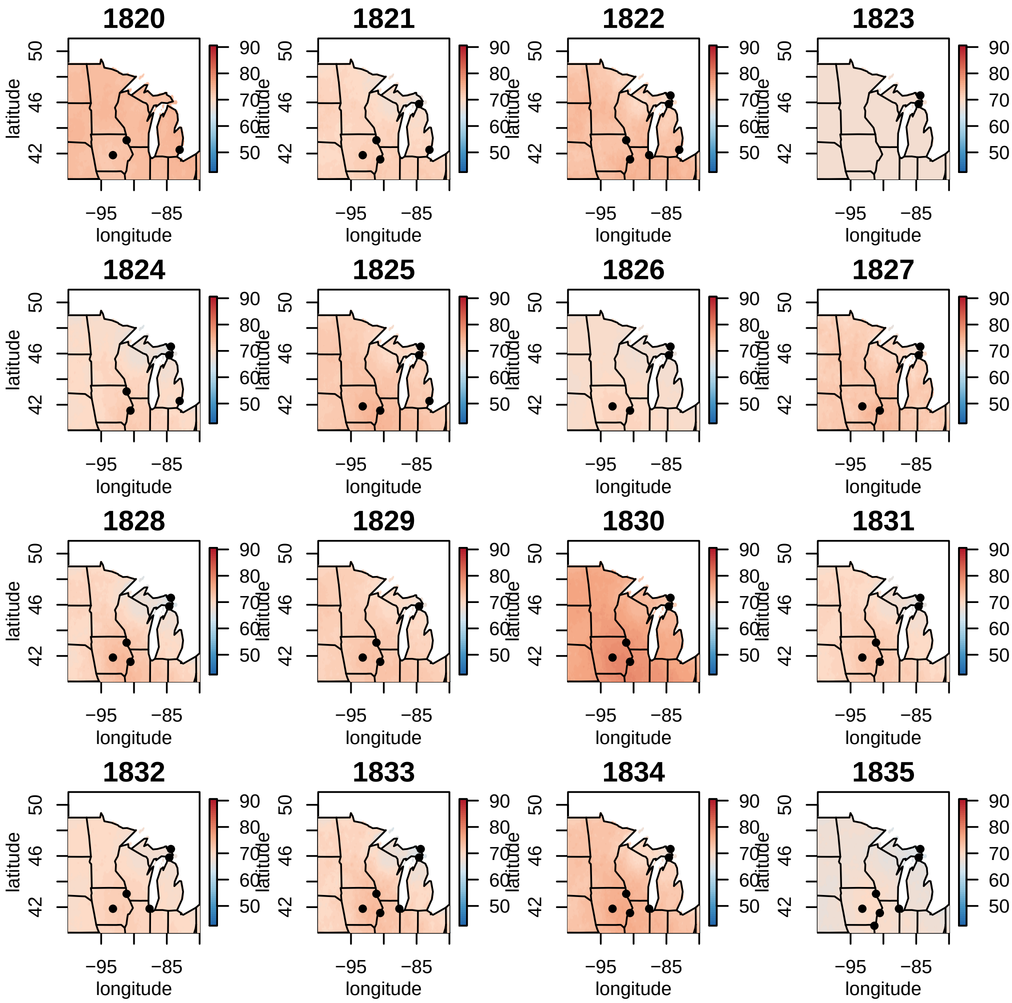

1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
latitude
longitude
50
46
42
−95
−85
20
15
10
5
0

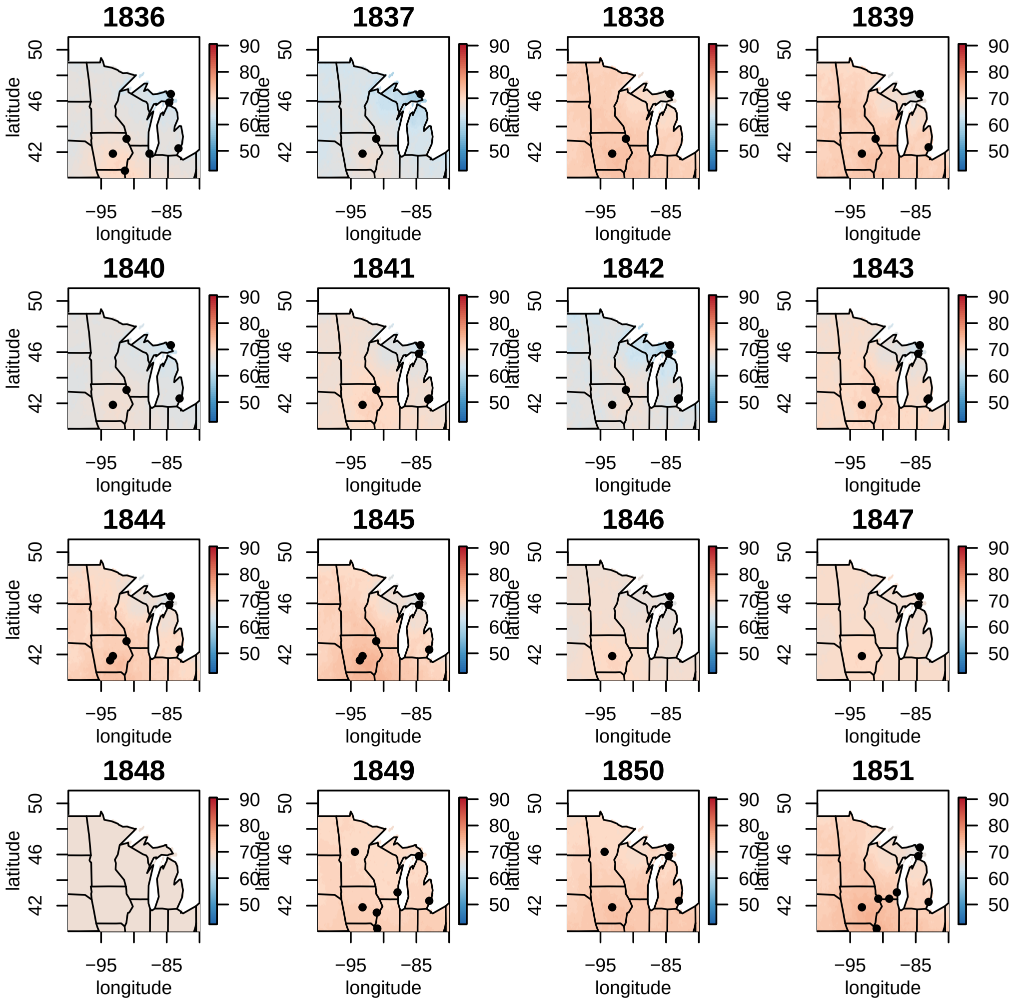

1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
latitude
longitude
50
46
42
−95
−85
20
15
10
5
0

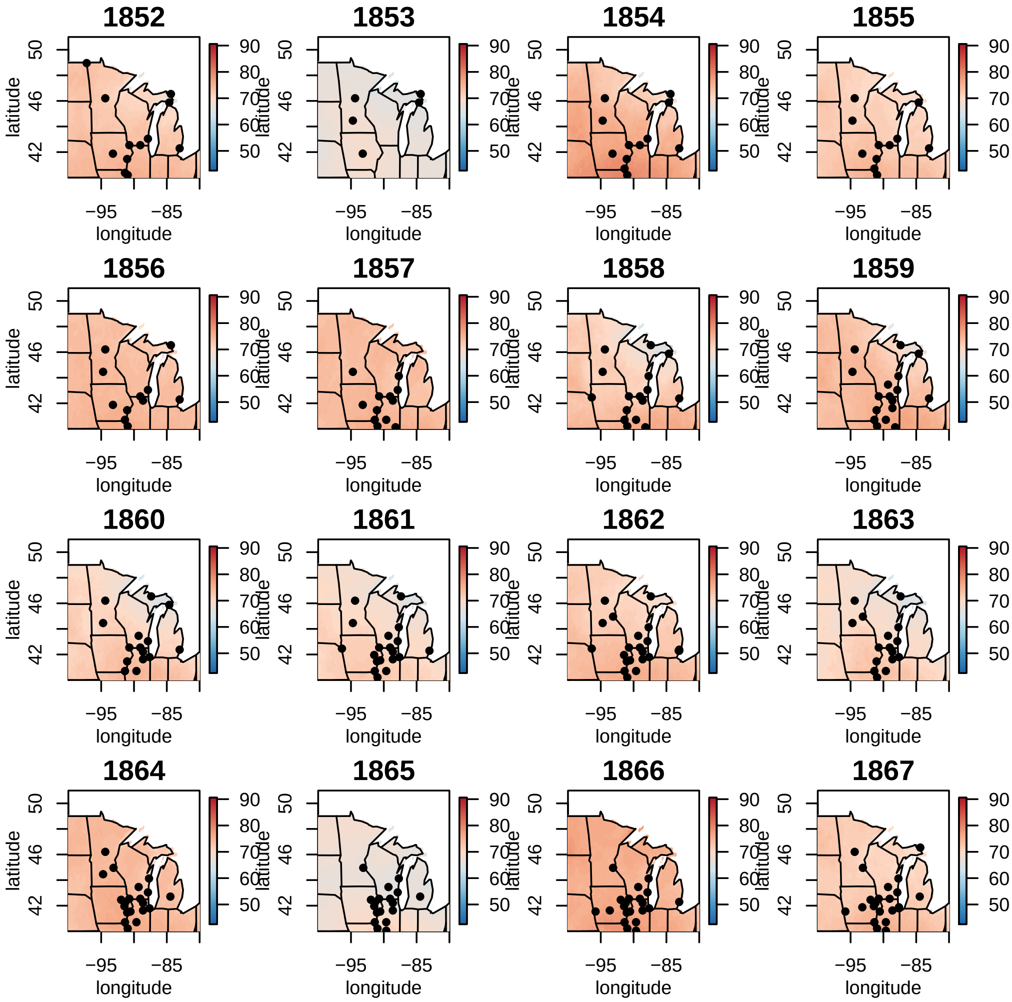

1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
latitude
50
46
42
longitude
−95
−85
20
15
10
5
0

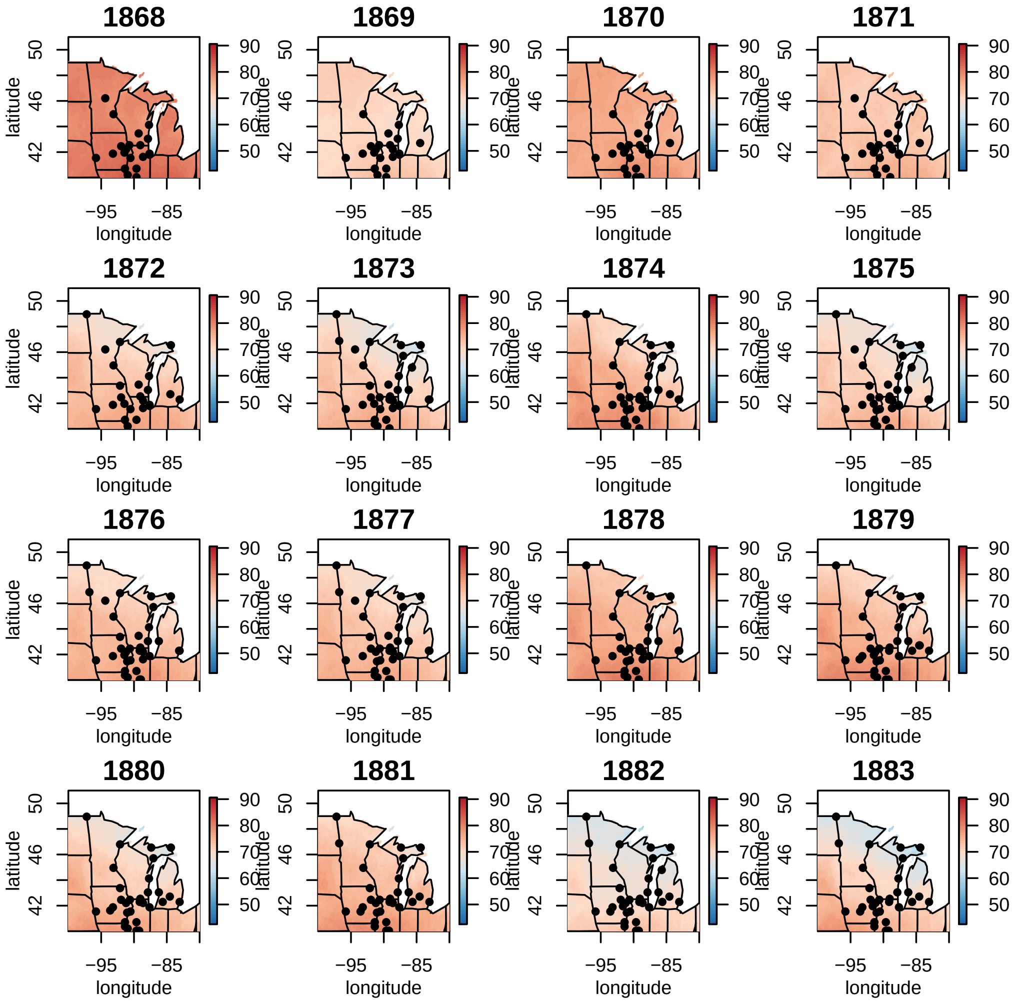

1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
latitude
longitude
50
46
42
−95
−85
20
15
10
5
0

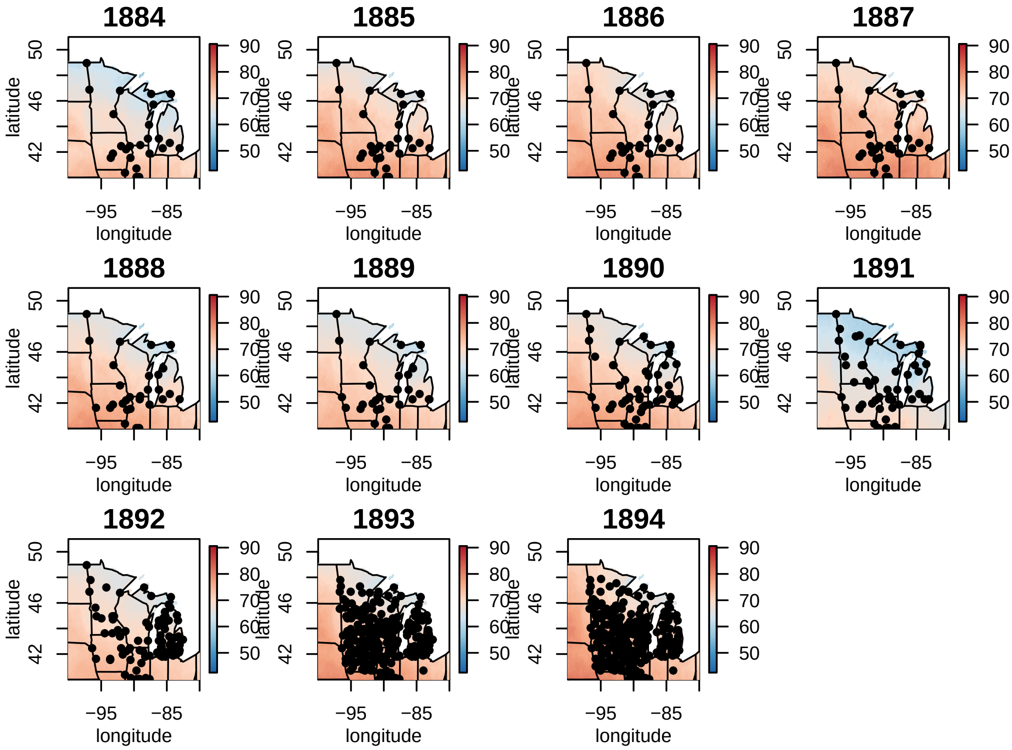
S.4 Robust Probabilistic PCR Model Posterior Mean July Temperature



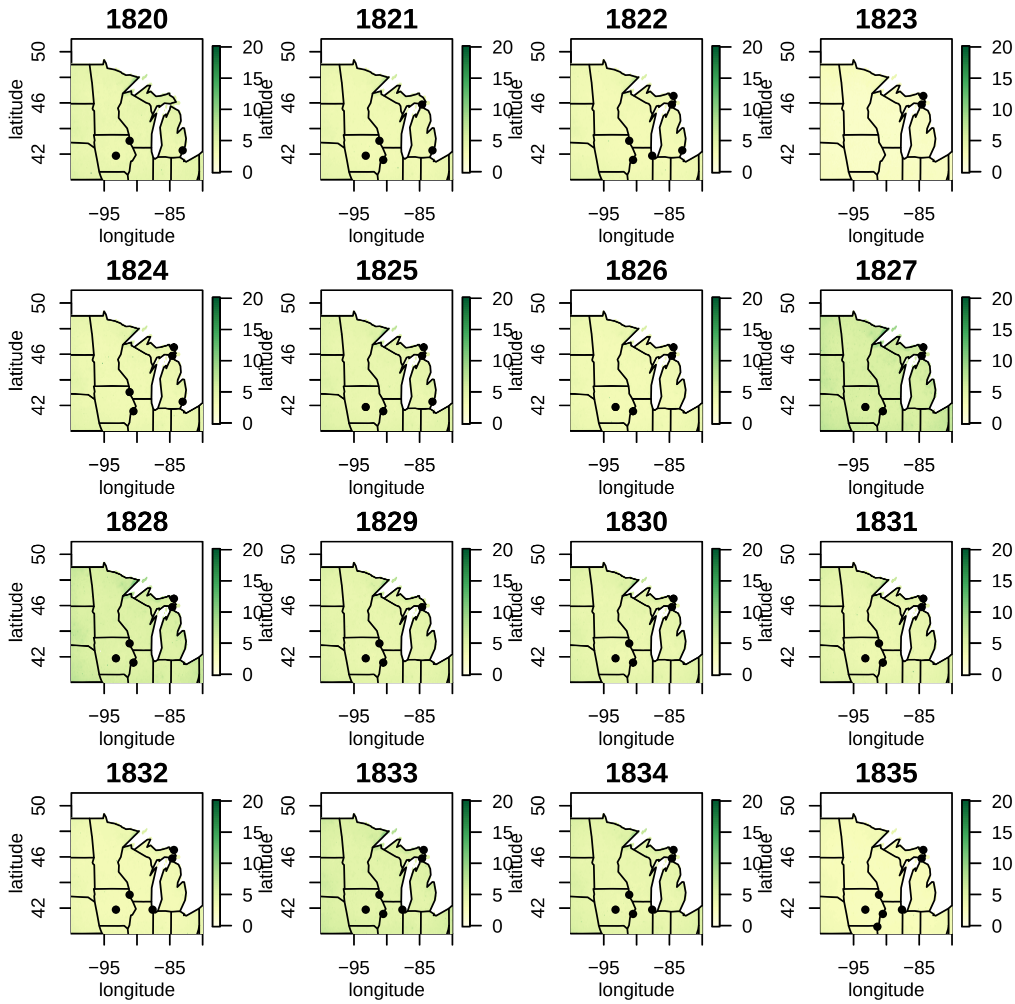

1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
latitude
longitude
50
46
42
−95
−85
90
80
70
60
50

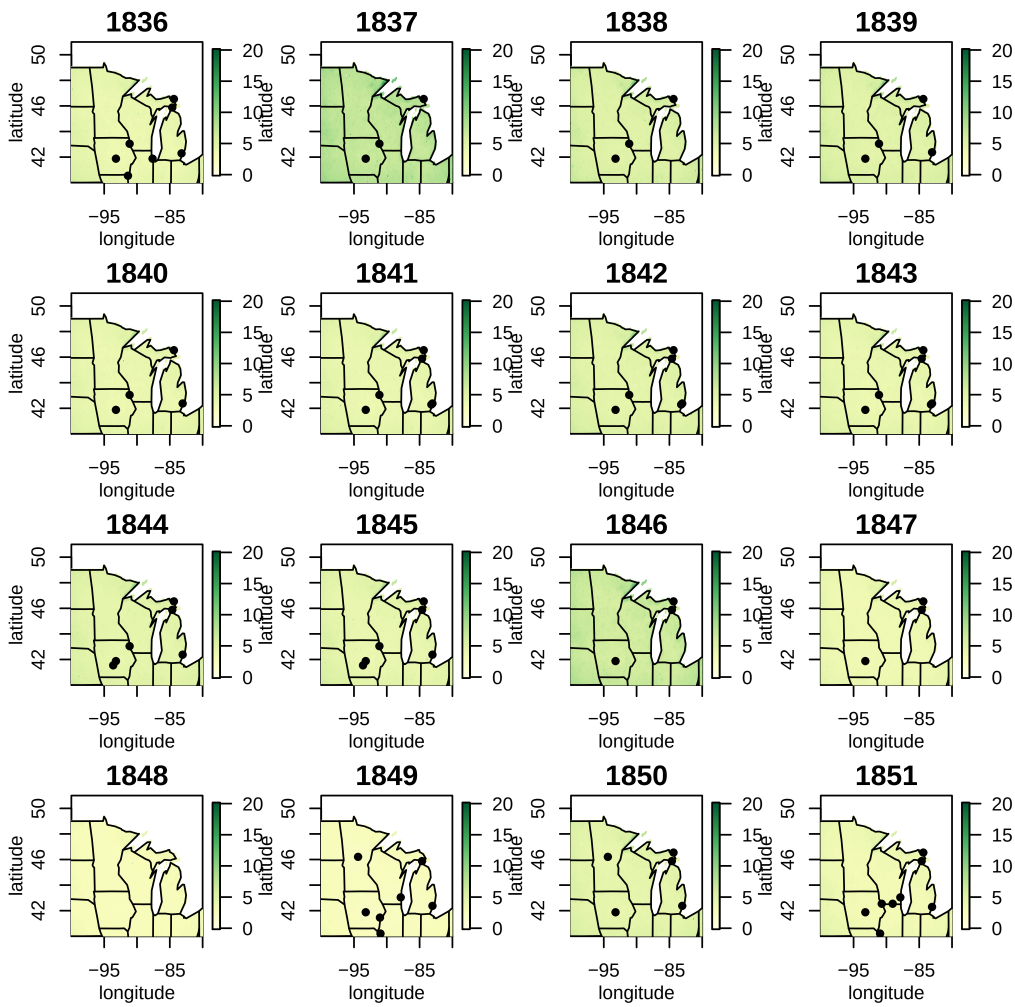

1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
latitude
longitude
50
46
42
−95
−85
90
80
70
60
50

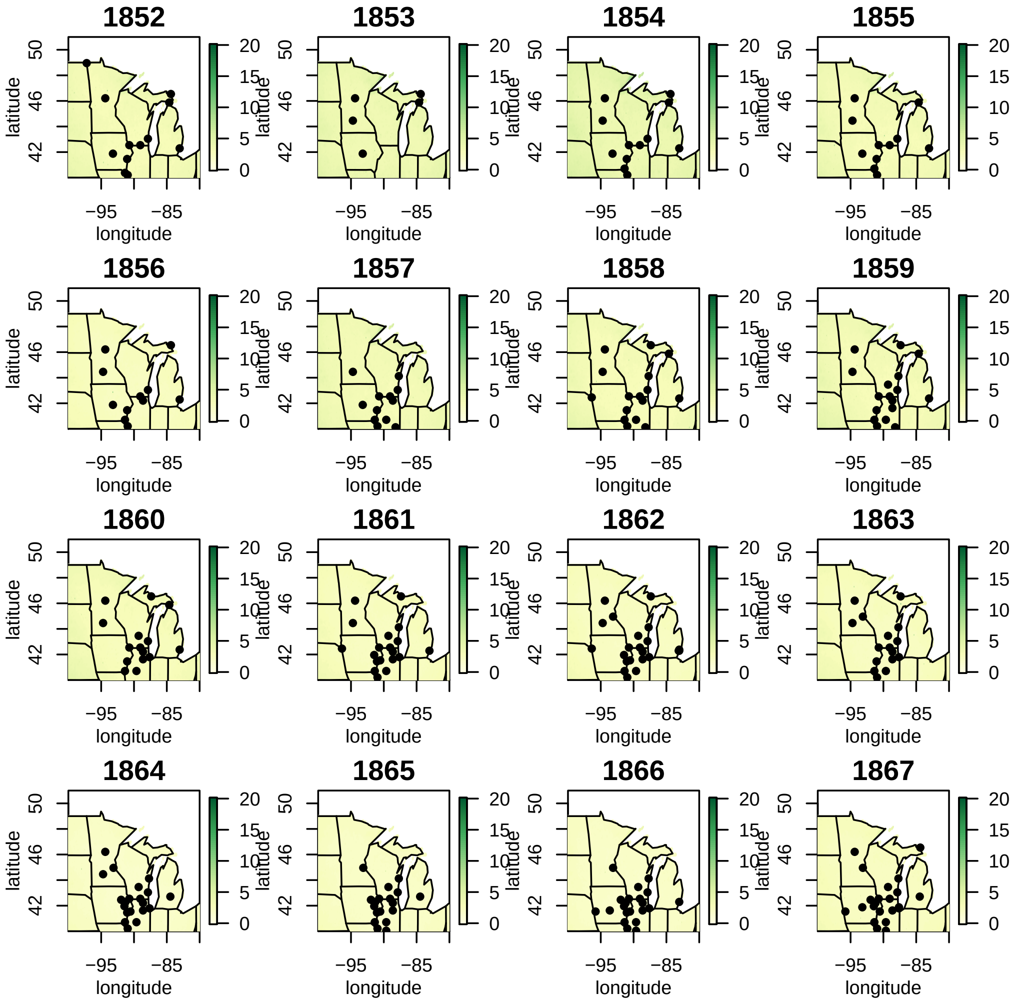

1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
latitude
longitude
50
46
42
−95
−85
90
80
70
60
50


1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
latitude
longitude
50
46
42
−95
−85
90
80
70
60
50

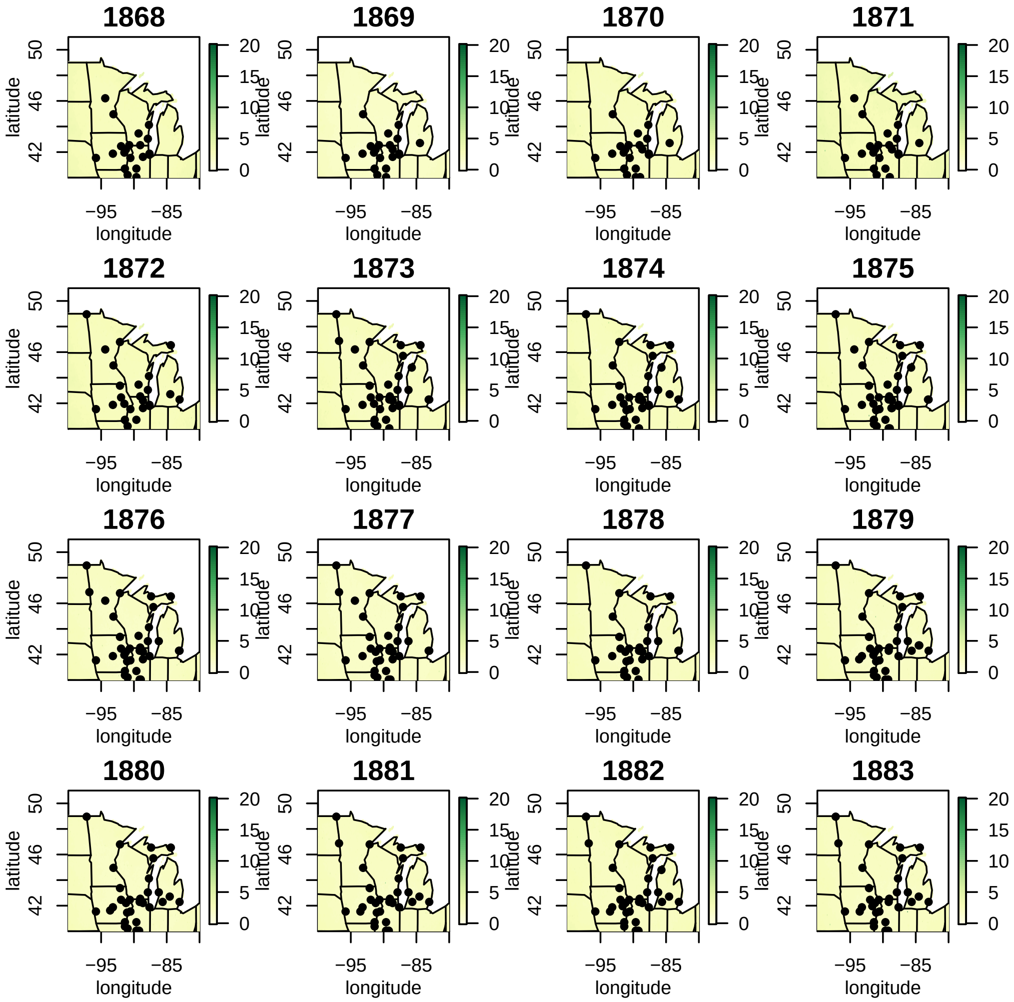
S.5 Robust Probabilistic PCR Model July Temperature Posterior Standard Deviations

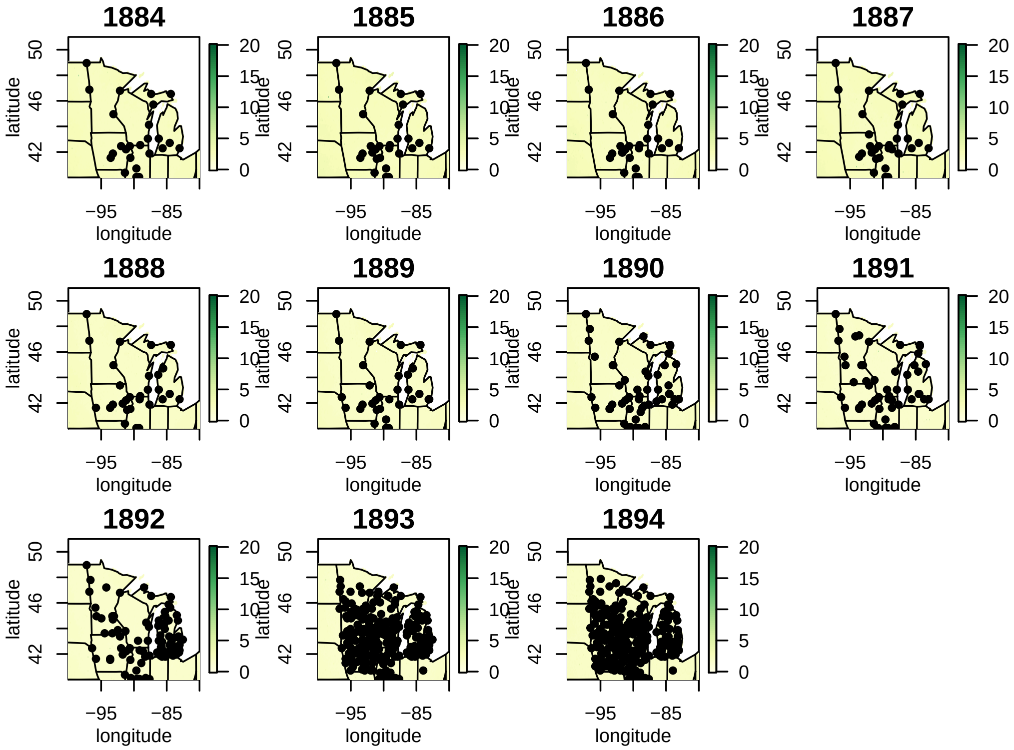



1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
latitude
longitude
50
46
42
−95
−85
20
15
10
5
0


1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
latitude
longitude
50
46
42
−95
−85
20
15
10
5
0

60


1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
latitude
longitude
50
46
42
−95
−85
20
15
10
5
0


1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
latitude
longitude
50
46
42
−95
−85
0
5
10
15
20