Marco Ferreira (PF&W 11.1)

- Latent factor models with one factor
- MCMC computations
- Latent *k*-factor models
- Choice of number of factors
- Example: troposphere temperature

Marco Ferreira (PF&W 11.1)

- Latent factor models with one factor
- MCMC computations
- Latent k-factor models
- Choice of number of factors
- Example: troposphere temperature

Marco Ferreira (PF&W 11.1)

- Latent factor models with one factor
- MCMC computations
- Latent k-factor models
- Choice of number of factors
- Example: troposphere temperature

Marco Ferreira (PF&W 11.1)

- Latent factor models with one factor
- MCMC computations
- Latent k-factor models
- Choice of number of factors
- Example: troposphere temperature

Marco Ferreira (PF&W 11.1)

- Latent factor models with one factor
- MCMC computations
- Latent *k*-factor models
- Choice of number of factors
- Example: troposphere temperature

Working example

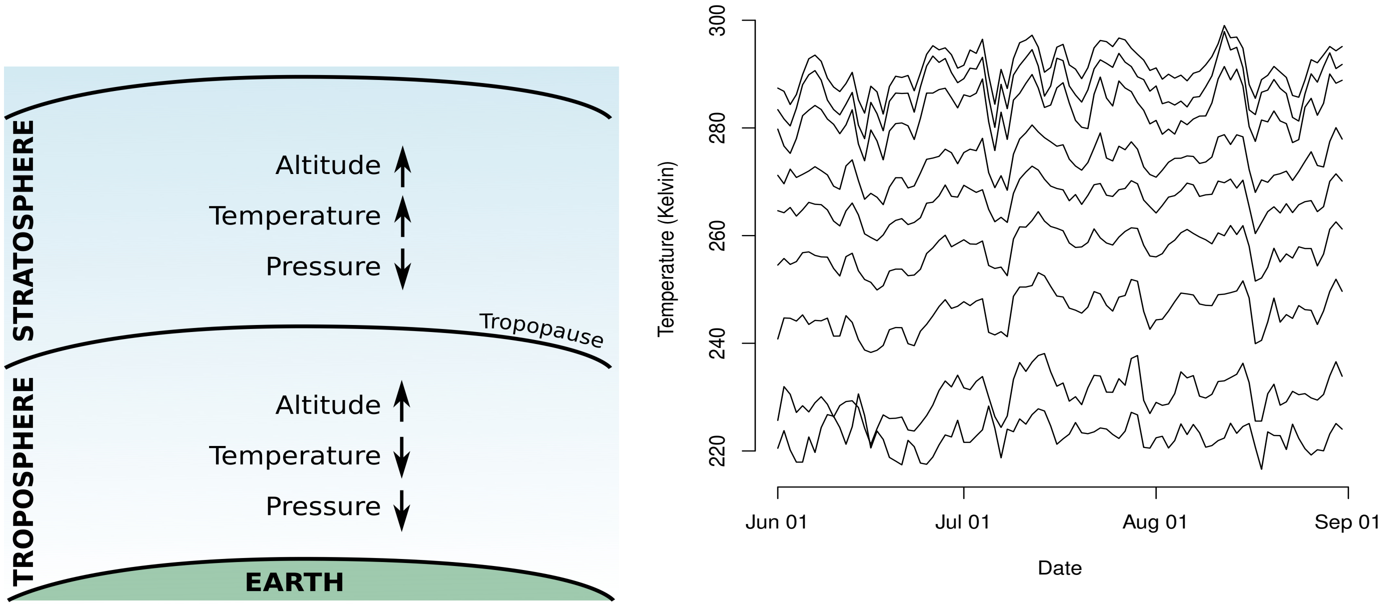


Figure: Daily average troposphere temperatures at longitude -100^o and latitude 55^o (over Canada) from June 1 to August 31, 2015, at different altitudes. Observation vector of dimension $r = 9$. $n = 92$ time points.

Latent 1-factor model

Observation equation: $\mathbf{y}_i = \mathbf{B}x_i + \mathbf{v}_i,$

- $\mathbf{y}_i$: ith observation of r-dimensional vector of interest.
- $\mathbf{B} = (b_1, \ldots, b_r)'$: r-dimensional vector of factor loadings.
- x_i: ith realization of one-dimensional latent factor process.
- $x_1, \ldots, x_n$ i.i.d. $N(0, H)$.
- $\mathbf{v}_i$: r-dimensional error vector for the ith observation.
- $\mathbf{v}_1, \ldots, \mathbf{v}_n$ i.i.d. $N(\mathbf{0}, \mathbf{V})$, with $\mathbf{V} = diag(\sigma_1^2, \ldots, \sigma_r^2)$.
- $\sigma_1^2, \ldots, \sigma_r^2$: idiosyncratic variances.

Latent 1-factor model

Observation equation: $\mathbf{y}_i = \mathbf{B}x_i + \mathbf{v}_i,$

- $\mathbf{y}_i$: ith observation of r-dimensional vector of interest.
- $\mathbf{B} = (b_1, \ldots, b_r)'$: r-dimensional vector of factor loadings.
- x_i: ith realization of one-dimensional latent factor process.
- $x_1, \ldots, x_n$ i.i.d. $N(0, H)$.
- $\mathbf{v}_i$: r-dimensional error vector for the ith observation.
- $\mathbf{v}_1, \ldots, \mathbf{v}_n$ i.i.d. $N(\mathbf{0}, \mathbf{V})$, with $\mathbf{V} = diag(\sigma_1^2, \ldots, \sigma_r^2)$.
- $\sigma_1^2, \ldots, \sigma_r^2$: idiosyncratic variances.

Latent 1-factor model

Observation equation: $\mathbf{y}_i = \mathbf{B}x_i + \mathbf{v}_i,$

- $\mathbf{y}_i$: ith observation of r-dimensional vector of interest.
- $\mathbf{B} = (b_1, \ldots, b_r)'$: r-dimensional vector of factor loadings.
- x_i: ith realization of one-dimensional latent factor process.
- $x_1, \ldots, x_n$ i.i.d. $N(0, H)$.
- $\mathbf{v}_i$: r-dimensional error vector for the ith observation.
- $\mathbf{v}_1, \ldots, \mathbf{v}_n$ i.i.d. $N(\mathbf{0}, \mathbf{V})$, with $\mathbf{V} = diag(\sigma_1^2, \ldots, \sigma_r^2)$.
- $\sigma_1^2, \ldots, \sigma_r^2$: idiosyncratic variances.

Latent 1-factor model

Observation equation: $\mathbf{y}_i = \mathbf{B}x_i + \mathbf{v}_i,$

- $\mathbf{y}_i$: ith observation of r-dimensional vector of interest.
- $\mathbf{B} = (b_1, \ldots, b_r)'$: r-dimensional vector of factor loadings.
- x_i: ith realization of one-dimensional latent factor process.
- $x_1, \ldots, x_n$ i.i.d. $N(0, H)$.
- $\mathbf{v}_i$: r-dimensional error vector for the ith observation.
- $\mathbf{v}_1, \ldots, \mathbf{v}_n$ i.i.d. $N(\mathbf{0}, \mathbf{V})$, with $\mathbf{V} = diag(\sigma_1^2, \ldots, \sigma_r^2)$.
- $\sigma_1^2, \ldots, \sigma_r^2$: idiosyncratic variances.

Latent 1-factor model

$$\text{Observation equation:} \quad \mathbf{y}_i = \mathbf{B}x_i + \mathbf{v}_i,$$

- $\mathbf{y}_i$: ith observation of r-dimensional vector of interest.
- $\mathbf{B} = (b_1, \ldots, b_r)'$: r-dimensional vector of factor loadings.
- x_i: ith realization of one-dimensional latent factor process.
- $x_1, \ldots, x_n$ i.i.d. $N(0, H)$.
- $\mathbf{v}_i$: r-dimensional error vector for the ith observation.
- $\mathbf{v}_1, \ldots, \mathbf{v}_n$ i.i.d. $N(\mathbf{0}, \mathbf{V})$, with $\mathbf{V} = diag(\sigma_1^2, \ldots, \sigma_r^2)$.
- $\sigma_1^2, \ldots, \sigma_r^2$: idiosyncratic variances.

- If $x_1, \ldots, x_n$ were known, then the latent 1-factor model would be a multivariate regression model.
- In that case, conditional on x_i, the second-order moments would be $var(y_{ij}|x_i) = \sigma_j^2$ and $cov(y_{ij}, y_{ik}|x_i) = 0$.
- With known $x_1, \ldots, x_n$, the factor loadings $b_1, \ldots, b_r$ would be the regression coefficients that measure the impact of the univariate factor x_i on each element of the vector $\mathbf{y}_i$.
- However, for statistical factor models, $x_1, \ldots, x_n$ are not observable.

- If $x_1, \ldots, x_n$ were known, then the latent 1-factor model would be a multivariate regression model.
- In that case, conditional on x_i, the second-order moments would be $var(y_{ij}|x_i) = \sigma_j^2$ and $cov(y_{ij}, y_{ik}|x_i) = 0$.
- With known $x_1, \ldots, x_n$, the factor loadings $b_1, \ldots, b_r$ would be the regression coefficients that measure the impact of the univariate factor x_i on each element of the vector $\mathbf{y}_i$.
- However, for statistical factor models, $x_1, \ldots, x_n$ are not observable.

- If $x_1, \ldots, x_n$ were known, then the latent 1-factor model would be a multivariate regression model.
- In that case, conditional on x_i, the second-order moments would be $var(y_{ij}|x_i) = \sigma_j^2$ and $cov(y_{ij}, y_{ik}|x_i) = 0$.
- With known $x_1, \ldots, x_n$, the factor loadings $b_1, \ldots, b_r$ would be the regression coefficients that measure the impact of the univariate factor x_i on each element of the vector $\mathbf{y}_i$.
- However, for statistical factor models, $x_1, \ldots, x_n$ are not observable.

- If $x_1, \ldots, x_n$ were known, then the latent 1-factor model would be a multivariate regression model.
- In that case, conditional on x_i, the second-order moments would be $var(y_{ij}|x_i) = \sigma_j^2$ and $cov(y_{ij}, y_{ik}|x_i) = 0$.
- With known $x_1, \ldots, x_n$, the factor loadings $b_1, \ldots, b_r$ would be the regression coefficients that measure the impact of the univariate factor x_i on each element of the vector $\mathbf{y}_i$.
- However, for statistical factor models, $x_1, \ldots, x_n$ are not observable.

- Thus, to obtain the first- and second-order moments of the observation vector $\mathbf{y}_i$ we have to integrate out x_i.
- By doing that, we find that the expected value of $\mathbf{y}_i$ is $E(\mathbf{y}_i) = \mathbf{0}$ (we usually center the variables of interest before using these factor models).
- In addition, the covariance matrix of $\mathbf{y}_i$ is $Cov(\mathbf{y}_i) = H\mathbf{B}\mathbf{B}' + \mathbf{V}$.

- Thus, to obtain the first- and second-order moments of the observation vector $\mathbf{y}_i$ we have to integrate out x_i.
- By doing that, we find that the expected value of $\mathbf{y}_i$ is $E(\mathbf{y}_i) = \mathbf{0}$ (we usually center the variables of interest before using these factor models).
- In addition, the covariance matrix of $\mathbf{y}_i$ is $Cov(\mathbf{y}_i) = H\mathbf{B}\mathbf{B}' + \mathbf{V}$.

- Thus, to obtain the first- and second-order moments of the observation vector $\mathbf{y}_i$ we have to integrate out x_i.
- By doing that, we find that the expected value of $\mathbf{y}_i$ is $E(\mathbf{y}_i) = \mathbf{0}$ (we usually center the variables of interest before using these factor models).
- In addition, the covariance matrix of $\mathbf{y}_i$ is $Cov(\mathbf{y}_i) = H\mathbf{B}\mathbf{B}' + \mathbf{V}$.

- Hence, this factor model with one factor induces a very specific covariance structure with $var(y_{ij}) = Hb_j^2 + \sigma_j^2$ and $cov(y_{ij}, y_{ik}) = Hb_j b_k$.
- Therefore, the factor loadings vector $\mathbf{B} = (b_1, \ldots, b_r)$ plays a critical role in the covariance structure of $\mathbf{y}_i$, with larger b_j implying larger variances and covariances.

- Hence, this factor model with one factor induces a very specific covariance structure with $var(y_{ij}) = Hb_j^2 + \sigma_j^2$ and $cov(y_{ij}, y_{ik}) = Hb_j b_k$.
- Therefore, the factor loadings vector $\mathbf{B} = (b_1, \ldots, b_r)$ plays a critical role in the covariance structure of $\mathbf{y}_i$, with larger b_j implying larger variances and covariances.

- Alternatively, we can write the 1-factor model in matrix notation

$$\mathbf{Y} = \mathbf{X}\mathbf{B}' + \mathbf{v},$$

- $\mathbf{Y} = (\mathbf{y}_1', \ldots, \mathbf{y}_n')'$ is the n by r matrix of observations,
- $\mathbf{X} = (x_1, \ldots, x_n)'$ is the n dimensional vector of latent factors,
- $\mathbf{B} = (b_1, \ldots, b_r)'$ is the r-dimensional vector of factor loadings.
- $\mathbf{v} = (\mathbf{v}_1', \ldots, \mathbf{v}_n')'$ is the n by r matrix of errors.

- Alternatively, we can write the 1-factor model in matrix notation
$$\mathbf{Y} = \mathbf{X}\mathbf{B}' + \mathbf{v},$$
- $\mathbf{Y} = (\mathbf{y}_1', \ldots, \mathbf{y}_n')'$ is the n by r matrix of observations,
- $\mathbf{X} = (x_1, \ldots, x_n)'$ is the n dimensional vector of latent factors,
- $\mathbf{B} = (b_1, \ldots, b_r)'$ is the r-dimensional vector of factor loadings.
- $\mathbf{v} = (\mathbf{v}_1', \ldots, \mathbf{v}_n')'$ is the n by r matrix of errors.

- Alternatively, we can write the 1-factor model in matrix notation
$$\mathbf{Y} = \mathbf{X}\mathbf{B}' + \mathbf{v},$$
- $\mathbf{Y} = (\mathbf{y}_1', ..., \mathbf{y}_n')'$ is the n by r matrix of observations,
- $\mathbf{X} = (x_1, \ldots, x_n)'$ is the n dimensional vector of latent factors,
- $\mathbf{B} = (b_1, \ldots, b_r)'$ is the r-dimensional vector of factor loadings.
- $\mathbf{v} = (\mathbf{v}_1', ..., \mathbf{v}_n')'$ is the n by r matrix of errors.

- Alternatively, we can write the 1-factor model in matrix notation
$$\mathbf{Y} = \mathbf{X}\mathbf{B}' + \mathbf{v},$$
- $\mathbf{Y} = (\mathbf{y}_1', ..., \mathbf{y}_n')'$ is the n by r matrix of observations,
- $\mathbf{X} = (x_1, \ldots, x_n)'$ is the n dimensional vector of latent factors,
- $\mathbf{B} = (b_1, \ldots, b_r)'$ is the r-dimensional vector of factor loadings.
- $\mathbf{v} = (\mathbf{v}_1', ..., \mathbf{v}_n')'$ is the n by r matrix of errors.

- Alternatively, we can write the 1-factor model in matrix notation

$$\mathbf{Y} = \mathbf{X}\mathbf{B}' + \mathbf{v},$$

- $\mathbf{Y} = (\mathbf{y}_1', ..., \mathbf{y}_n')'$ is the n by r matrix of observations,
- $\mathbf{X} = (x_1, \ldots, x_n)'$ is the n dimensional vector of latent factors,
- $\mathbf{B} = (b_1, \ldots, b_r)'$ is the r-dimensional vector of factor loadings.
- $\mathbf{v} = (\mathbf{v}_1', ..., \mathbf{v}_n')'$ is the n by r matrix of errors.

Identifiability

- Identifiability is an important issue in factor models.
- Note that if we define $\mathbf{B}^* = a\mathbf{B}$ and $x_i^* = x_i/a$ for a scalar $a \neq 0$, then $\mathbf{y}_i = \mathbf{B}^* x_i^* + \mathbf{v}_i$ is exactly the same model for the observations as $\mathbf{y}_i = \mathbf{B} x_i + \mathbf{v}_i$.
- There are many ways to resolve this identifiability problem.
- One simple and intuitive way that we use here is to set the first element of $\mathbf{B}$ to be equal to 1, that is, $b_1 = 1$.

Identifiability

- Identifiability is an important issue in factor models.
- Note that if we define $\mathbf{B}^* = a\mathbf{B}$ and $x_i^* = x_i/a$ for a scalar $a \neq 0$, then $\mathbf{y}_i = \mathbf{B}^* x_i^* + \mathbf{v}_i$ is exactly the same model for the observations as $\mathbf{y}_i = \mathbf{B} x_i + \mathbf{v}_i$.
- There are many ways to resolve this identifiability problem.
- One simple and intuitive way that we use here is to set the first element of $\mathbf{B}$ to be equal to 1, that is, $b_1 = 1$.

Identifiability

- Identifiability is an important issue in factor models.
- Note that if we define $\mathbf{B}^* = a\mathbf{B}$ and $x_i^* = x_i/a$ for a scalar $a \neq 0$, then $\mathbf{y}_i = \mathbf{B}^* x_i^* + \mathbf{v}_i$ is exactly the same model for the observations as $\mathbf{y}_i = \mathbf{B} x_i + \mathbf{v}_i$.
- There are many ways to resolve this identifiability problem.
- One simple and intuitive way that we use here is to set the first element of $\mathbf{B}$ to be equal to 1, that is, $b_1 = 1$.

Identifiability

- Identifiability is an important issue in factor models.
- Note that if we define $\mathbf{B}^* = a\mathbf{B}$ and $x_i^* = x_i/a$ for a scalar $a \neq 0$, then $\mathbf{y}_i = \mathbf{B}^* x_i^* + \mathbf{v}_i$ is exactly the same model for the observations as $\mathbf{y}_i = \mathbf{B} x_i + \mathbf{v}_i$.
- There are many ways to resolve this identifiability problem.
- One simple and intuitive way that we use here is to set the first element of $\mathbf{B}$ to be equal to 1, that is, $b_1 = 1$.

Priors

- For the factor loadings $(b_2, \ldots, b_r)$ we consider independent priors $N(0, \tau^2)$.
- For the latent factor variance H we assume $IG(n_H/2, n_H s_H^2/2)$.
- For the idiosyncratic variances $\sigma_1^2, \ldots, \sigma_r^2$ we assume independent priors $IG(n_\sigma/2, n_\sigma s_\sigma^2/2)$.
- Finally, for the factor loadings variance τ^2 we assume the prior $IG(n_\tau/2, n_\tau s_\tau^2/2)$.

Priors

- For the factor loadings $(b_2, \ldots, b_r)$ we consider independent priors $N(0, \tau^2)$.
- For the latent factor variance H we assume $IG(n_H/2, n_H s_H^2/2)$.
- For the idiosyncratic variances $\sigma_1^2, \ldots, \sigma_r^2$ we assume independent priors $IG(n_\sigma/2, n_\sigma s_\sigma^2/2)$.
- Finally, for the factor loadings variance τ^2 we assume the prior $IG(n_\tau/2, n_\tau s_\tau^2/2)$.

Priors

- For the factor loadings $(b_2, \ldots, b_r)$ we consider independent priors $N(0, \tau^2)$.
- For the latent factor variance H we assume $IG(n_H/2, n_H s_H^2/2)$.
- For the idiosyncratic variances $\sigma_1^2, \ldots, \sigma_r^2$ we assume independent priors $IG(n_\sigma/2, n_\sigma s_\sigma^2/2)$.
- Finally, for the factor loadings variance τ^2 we assume the prior $IG(n_\tau/2, n_\tau s_\tau^2/2)$.

Priors

- For the factor loadings $(b_2, \ldots, b_r)$ we consider independent priors $N(0, \tau^2)$.
- For the latent factor variance H we assume $IG(n_H/2, n_H s_H^2/2)$.
- For the idiosyncratic variances $\sigma_1^2, \ldots, \sigma_r^2$ we assume independent priors $IG(n_\sigma/2, n_\sigma s_\sigma^2/2)$.
- Finally, for the factor loadings variance τ^2 we assume the prior $IG(n_\tau/2, n_\tau s_\tau^2/2)$.

Full conditional distributions

- Latent factor x_i :

$$N\left(\frac{\sum_{j=1}^{r}\sigma_j^{-2}b_jy_{ij}}{H^{-1}+\sum_{j=1}^{r}\sigma_j^{-2}b_j^2},\frac{1}{H^{-1}+\sum_{j=1}^{r}\sigma_j^{-2}b_j^2}\right).$$

- Factor loading b_j:

$$N\left(\frac{\sum_{i=1}^{n}x_iy_{ij}}{\sigma_j^2\tau^{-2}+\sum_{i=1}^{n}x_i^2},\frac{\sigma_j^2}{\sigma_j^2\tau^{-2}+\sum_{i=1}^{n}x_i^2}\right).$$

Full conditional distributions

- Latent factor x_i :

$$N\left(\frac{\sum_{j=1}^{r}\sigma_j^{-2}b_jy_{ij}}{H^{-1}+\sum_{j=1}^{r}\sigma_j^{-2}b_j^2},\frac{1}{H^{-1}+\sum_{j=1}^{r}\sigma_j^{-2}b_j^2}\right).$$

- Factor loading b_j:

$$N\left(\frac{\sum_{i=1}^{n}x_iy_{ij}}{\sigma_j^2\tau^{-2}+\sum_{i=1}^{n}x_i^2},\frac{\sigma_j^2}{\sigma_j^2\tau^{-2}+\sum_{i=1}^{n}x_i^2}\right).$$

- Idiosyncratic variance σ_j^2:

$$IG\left(\frac{n_\sigma + n}{2}, \frac{n_\sigma s_\sigma^2 + \sum_{i=1}^{n}(y_{ij} - b_j x_i)^2}{2}\right).$$

- Factor loadings variance τ^2 :

$$IG\left(\frac{n_\tau + r - 1}{2}, \frac{n_\tau s_\tau^2 + \sum_{j=2}^{r} b_j^2}{2}\right).$$

- Latent factor variance H :

$$IG\left(\frac{n_H + n}{2}, \frac{n_H s_H^2 + \sum_{i=1}^{n} x_i^2}{2}\right).$$

- Idiosyncratic variance σ_j^2:

$$IG\left(\frac{n_\sigma + n}{2}, \frac{n_\sigma s_\sigma^2 + \sum_{i=1}^n (y_{ij} - b_j x_i)^2}{2}\right).$$

- Factor loadings variance τ^2 :

$$IG\left(\frac{n_\tau + r - 1}{2}, \frac{n_\tau s_\tau^2 + \sum_{j=2}^r b_j{}^2}{2}\right).$$

- Latent factor variance H :

$$IG\left(\frac{n_H + n}{2}, \frac{n_H s_H^2 + \sum_{i=1}^n x_i{}^2}{2}\right).$$

- Idiosyncratic variance σ_j^2:

$$IG\left(\frac{n_\sigma + n}{2}, \frac{n_\sigma s_\sigma^2 + \sum_{i=1}^{n}(y_{ij} - b_j x_i)^2}{2}\right).$$

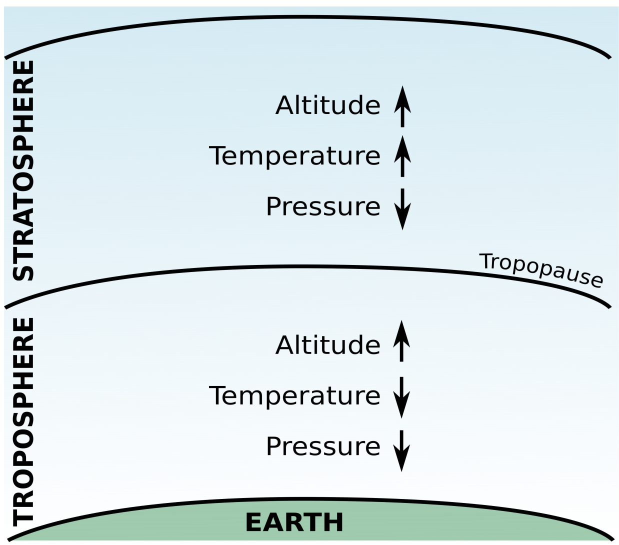
- Factor loadings variance τ^2 :

$$IG\left(\frac{n_\tau + r - 1}{2}, \frac{n_\tau s_\tau^2 + \sum_{j=2}^{r} {b_j}^2}{2}\right).$$

- Latent factor variance H :

$$IG\left(\frac{n_H + n}{2}, \frac{n_H s_H^2 + \sum_{i=1}^{n} {x_i}^2}{2}\right).$$

Back to working example



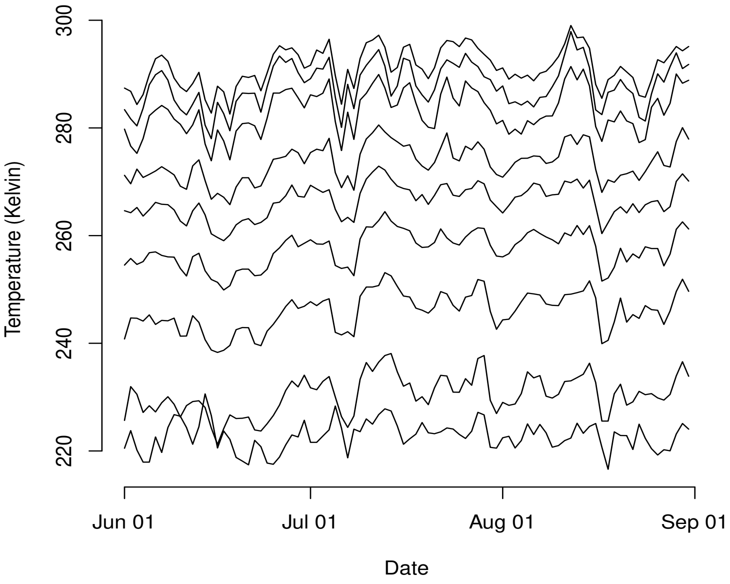
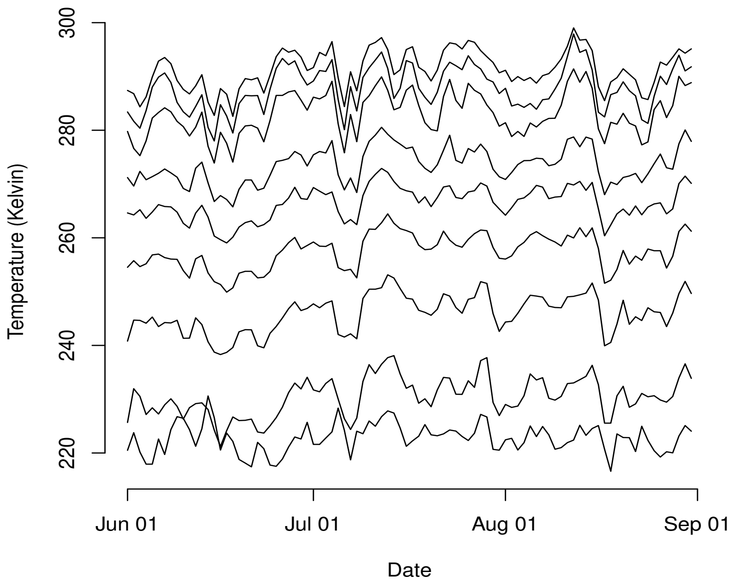
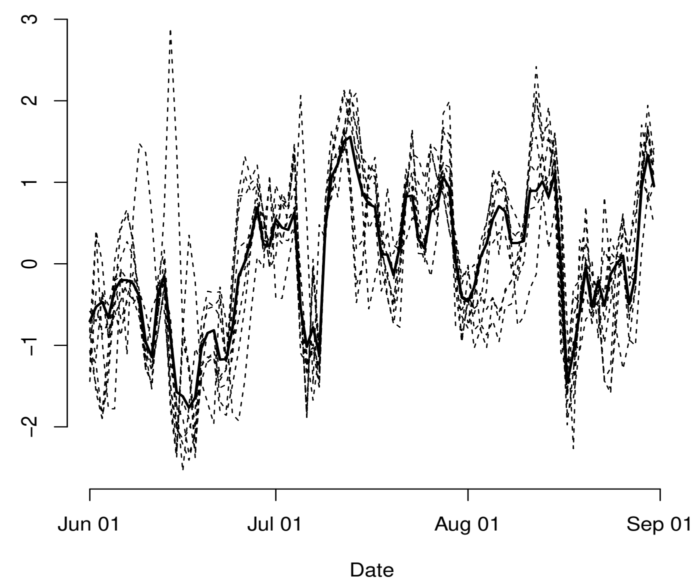


Figure: Daily average troposphere temperatures at longitude -100^o and latitude 55^o (over Canada) from June 1 to August 31, 2015, at different altitudes. Observation vector of dimension $r = 9$. $n = 92$ time points.



(a)



(b)

Figure: (a) Daily average troposphere temperatures. (b) Standardized temperature series (dashed lines) and estimated common factor (bold solid line).

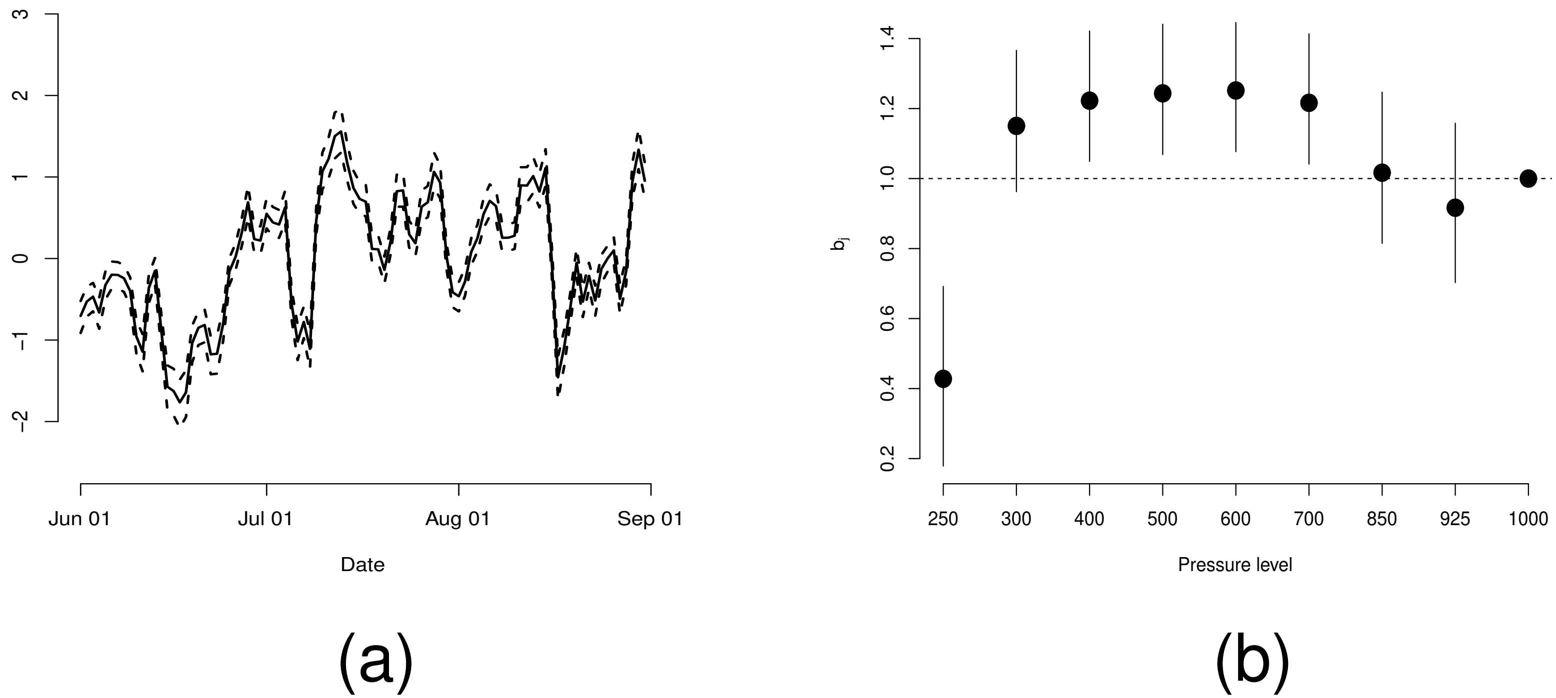


Figure: (a) Common factor ($x_{1:T}$) posterior mean (solid line) and 95% credible intervals (dashed lines); (b) Posterior means (circles) and 95% credible intervals (vertical lines) of factor loadings b_j for each pressure level.

Latent k-factor model

Observation equation: $\mathbf{y}_i = \mathbf{B}\mathbf{x}_i + \mathbf{v}_i.$

- $\mathbf{y}_i$: ith observation of the r-dimensional vector of interest.
- $\mathbf{B}$: $r \times k$ matrix of factor loadings.
- $\mathbf{x}_i$: ith realization of k-dimensional latent factor process.
- $\mathbf{x}_1, \ldots, \mathbf{x}_n$ i.i.d. $N(\mathbf{0}, \mathbf{H})$ with $\mathbf{H} = diag(h_1, \ldots, h_k)$.
- $\mathbf{v}_i$: r-dimensional error vector for the ith observation.
- $\mathbf{v}_1, \ldots, \mathbf{v}_n$ i.i.d. $N(\mathbf{0}, \mathbf{V})$, with $\mathbf{V} = diag(\sigma_1^2, \ldots, \sigma_r^2)$.
- $\sigma_1^2, \ldots, \sigma_r^2$: idiosyncratic variances.

Latent k-factor model

Observation equation: $\mathbf{y}_i = \mathbf{B}\mathbf{x}_i + \mathbf{v}_i$.

- $\mathbf{y}_i$: ith observation of the r-dimensional vector of interest.
- $\mathbf{B}$: $r \times k$ matrix of factor loadings.
- $\mathbf{x}_i$: ith realization of k-dimensional latent factor process.
- $\mathbf{x}_1, \ldots, \mathbf{x}_n$ i.i.d. $N(\mathbf{0}, \mathbf{H})$ with $\mathbf{H} = diag(h_1, \ldots, h_k)$.
- $\mathbf{v}_i$: r-dimensional error vector for the ith observation.
- $\mathbf{v}_1, \ldots, \mathbf{v}_n$ i.i.d. $N(\mathbf{0}, \mathbf{V})$, with $\mathbf{V} = diag(\sigma_1^2, \ldots, \sigma_r^2)$.
- $\sigma_1^2, \ldots, \sigma_r^2$: idiosyncratic variances.

Latent k-factor model

Observation equation: $\mathbf{y}_i = \mathbf{B}\mathbf{x}_i + \mathbf{v}_i.$

- $\mathbf{y}_i$: ith observation of the r-dimensional vector of interest.
- $\mathbf{B}$: $r \times k$ matrix of factor loadings.
- $\mathbf{x}_i$: ith realization of k-dimensional latent factor process.
- $\mathbf{x}_1, \ldots, \mathbf{x}_n$ i.i.d. $N(\mathbf{0}, \mathbf{H})$ with $\mathbf{H} = diag(h_1, \ldots, h_k)$.
- $\mathbf{v}_i$: r-dimensional error vector for the ith observation.
- $\mathbf{v}_1, \ldots, \mathbf{v}_n$ i.i.d. $N(\mathbf{0}, \mathbf{V})$, with $\mathbf{V} = diag(\sigma_1^2, \ldots, \sigma_r^2)$.
- $\sigma_1^2, \ldots, \sigma_r^2$: idiosyncratic variances.

Latent k-factor model

Observation equation: $\mathbf{y}_i = \mathbf{B}\mathbf{x}_i + \mathbf{v}_i.$

- $\mathbf{y}_i$: ith observation of the r-dimensional vector of interest.
- $\mathbf{B}$: $r \times k$ matrix of factor loadings.
- $\mathbf{x}_i$: ith realization of k-dimensional latent factor process.
- $\mathbf{x}_1, \ldots, \mathbf{x}_n$ i.i.d. $N(\mathbf{0}, \mathbf{H})$ with $\mathbf{H} = diag(h_1, \ldots, h_k)$.
- $\mathbf{v}_i$: r-dimensional error vector for the ith observation.
- $\mathbf{v}_1, \ldots, \mathbf{v}_n$ i.i.d. $N(\mathbf{0}, \mathbf{V})$, with $\mathbf{V} = diag(\sigma_1^2, \ldots, \sigma_r^2)$.
- $\sigma_1^2, \ldots, \sigma_r^2$: idiosyncratic variances.

Latent k-factor model

$$\text{Observation equation:} \quad \mathbf{y}_i = \mathbf{B}\mathbf{x}_i + \mathbf{v}_i.$$

- $\mathbf{y}_i$: ith observation of the r-dimensional vector of interest.
- $\mathbf{B}$: $r \times k$ matrix of factor loadings.
- $\mathbf{x}_i$: ith realization of k-dimensional latent factor process.
- $\mathbf{x}_1, \ldots, \mathbf{x}_n$ i.i.d. $N(\mathbf{0}, \mathbf{H})$ with $\mathbf{H} = diag(h_1, \ldots, h_k)$.
- $\mathbf{v}_i$: r-dimensional error vector for the ith observation.
- $\mathbf{v}_1, \ldots, \mathbf{v}_n$ i.i.d. $N(\mathbf{0}, \mathbf{V})$, with $\mathbf{V} = diag(\sigma_1^2, \ldots, \sigma_r^2)$.
- $\sigma_1^2, \ldots, \sigma_r^2$: idiosyncratic variances.

Latent k-factor model

Observation equation: $\mathbf{y}_i = \mathbf{B}\mathbf{x}_i + \mathbf{v}_i.$

- $\mathbf{y}_i$: ith observation of the r-dimensional vector of interest.
- $\mathbf{B}$: $r \times k$ matrix of factor loadings.
- $\mathbf{x}_i$: ith realization of k-dimensional latent factor process.
- $\mathbf{x}_1, \ldots, \mathbf{x}_n$ i.i.d. $N(\mathbf{0}, \mathbf{H})$ with $\mathbf{H} = diag(h_1, \ldots, h_k)$.
- $\mathbf{v}_i$: r-dimensional error vector for the ith observation.
- $\mathbf{v}_1, \ldots, \mathbf{v}_n$ i.i.d. $N(\mathbf{0}, \mathbf{V})$, with $\mathbf{V} = diag(\sigma_1^2, \ldots, \sigma_r^2)$.
- $\sigma_1^2, \ldots, \sigma_r^2$: idiosyncratic variances.

Identifiability

Hierarchical structural constraint on factor loadings matrix **B**:

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ b_{21} & 1 & 0 & \cdots & 0 \\ b_{31} & b_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{k1} & b_{k2} & b_{k3} & \cdots & 1 \\ b_{k+1,1} & b_{k+1,2} & b_{k+1,3} & \cdots & b_{k+1,k} \\ \vdots & \vdots & \vdots & & \vdots \\ b_{r1} & b_{r2} & b_{r3} & \cdots & b_{rk} \end{pmatrix}$$

B is a full rank block lower triangular matrix with diagonal elements equal to 1.

- Hierarchical structural constraint on $\mathbf{B}$ has important implications for the interpretation of the factors.
- Order of the variables in $\mathbf{y}_i$ defines the factors.
- First variable is equal to first factor plus noise, second variable is equal to b_{21} times first factor plus second factor plus noise, etc.
- Order may influence interpretation of the factors.

- Hierarchical structural constraint on $\mathbf{B}$ has important implications for the interpretation of the factors.
- Order of the variables in $\mathbf{y}_i$ defines the factors.
- First variable is equal to first factor plus noise, second variable is equal to b_{21} times first factor plus second factor plus noise, etc.
- Order may influence interpretation of the factors.

- Hierarchical structural constraint on $\mathbf{B}$ has important implications for the interpretation of the factors.
- Order of the variables in $\mathbf{y}_i$ defines the factors.
- First variable is equal to first factor plus noise, second variable is equal to b_{21} times first factor plus second factor plus noise, etc.
- Order may influence interpretation of the factors.

- Hierarchical structural constraint on $\mathbf{B}$ has important implications for the interpretation of the factors.
- Order of the variables in $\mathbf{y}_i$ defines the factors.
- First variable is equal to first factor plus noise, second variable is equal to b_{21} times first factor plus second factor plus noise, etc.
- Order may influence interpretation of the factors.

- Order may influence choice of number of factors k (Aguilar and West, 2000; Lopes and West, 2004).
- However, for a fixed number of factors k, the ordering has no influence either on the covariance matrix of $\mathbf{y}_i$ or on forecasts.
- Order of variables in $\mathbf{y}_i$ is a modeling decision.
- If possible, order variables based on prior knowledge of their relative importance and possible relationships.

- Order may influence choice of number of factors k (Aguilar and West, 2000; Lopes and West, 2004).
- However, for a fixed number of factors k, the ordering has no influence either on the covariance matrix of $\mathbf{y}_i$ or on forecasts.
- Order of variables in $\mathbf{y}_i$ is a modeling decision.
- If possible, order variables based on prior knowledge of their relative importance and possible relationships.

- Order may influence choice of number of factors k (Aguilar and West, 2000; Lopes and West, 2004).
- However, for a fixed number of factors k, the ordering has no influence either on the covariance matrix of $\mathbf{y}_i$ or on forecasts.
- Order of variables in $\mathbf{y}_i$ is a modeling decision.
- If possible, order variables based on prior knowledge of their relative importance and possible relationships.

- Order may influence choice of number of factors k (Aguilar and West, 2000; Lopes and West, 2004).
- However, for a fixed number of factors k, the ordering has no influence either on the covariance matrix of $\mathbf{y}_i$ or on forecasts.
- Order of variables in $\mathbf{y}_i$ is a modeling decision.
- If possible, order variables based on prior knowledge of their relative importance and possible relationships.

Bayesian analysis

- Conditionally conjugate priors.
- MCMC to explore posterior distribution.
- Marginal data density to choose number of factors (Laplace-Metropolis approximation).

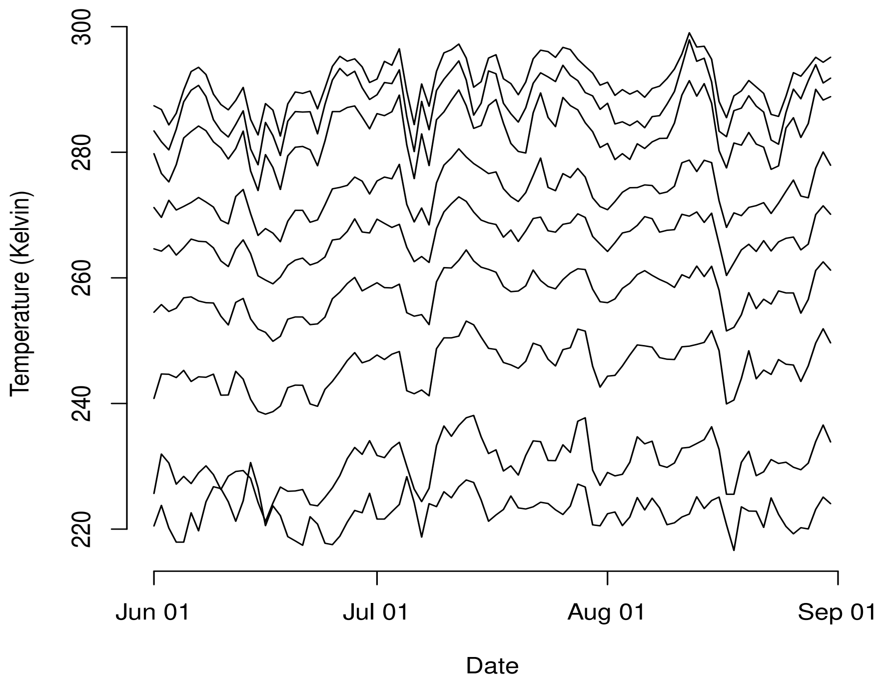
Bayesian analysis

- Conditionally conjugate priors.
- MCMC to explore posterior distribution.
- Marginal data density to choose number of factors (Laplace-Metropolis approximation).

Bayesian analysis

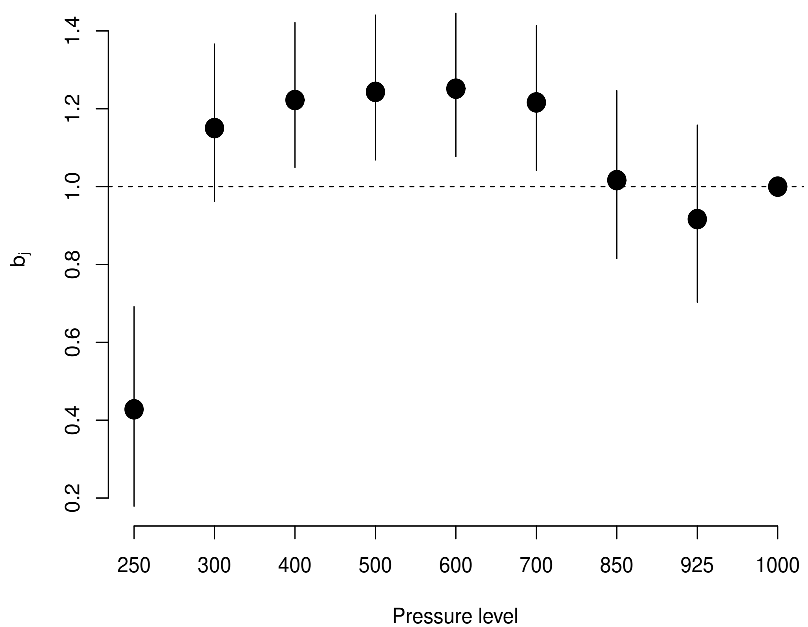
- Conditionally conjugate priors.
- MCMC to explore posterior distribution.
- Marginal data density to choose number of factors (Laplace-Metropolis approximation).

Latent k-factor model for troposphere temperature



Deciding the order of the variables

Recall the factor loadings b_j for the 1-factor model:



- In 1-factor model, the largest estimated factor loading is for altitude pressure level of 600 millibars.
- Temperature at that altitude seems to behave similarly to temperatures at other troposphere mid-altitudes.
- First position in $\mathbf{y}_i$: standardized temperature at altitude pressure level of 600 millibars.

- In 1-factor model, the largest estimated factor loading is for altitude pressure level of 600 millibars.
- Temperature at that altitude seems to behave similarly to temperatures at other troposphere mid-altitudes.
- First position in $\mathbf{y}_i$: standardized temperature at altitude pressure level of 600 millibars.

- In 1-factor model, the largest estimated factor loading is for altitude pressure level of 600 millibars.
- Temperature at that altitude seems to behave similarly to temperatures at other troposphere mid-altitudes.
- First position in $\mathbf{y}_i$: standardized temperature at altitude pressure level of 600 millibars.

- Second position: standardized temperature at altitude pressure level of 1000 millibars.
- Third position: standardized temperature at altitude pressure level of 250 millibars, which is close to the tropopause and seems to behave differently from the temperatures at lower altitudes.
- The order for the remaining variables is the original order from lower to upper altitudes.

- Second position: standardized temperature at altitude pressure level of 1000 millibars.
- Third position: standardized temperature at altitude pressure level of 250 millibars, which is close to the tropopause and seems to behave differently from the temperatures at lower altitudes.
- The order for the remaining variables is the original order from lower to upper altitudes.

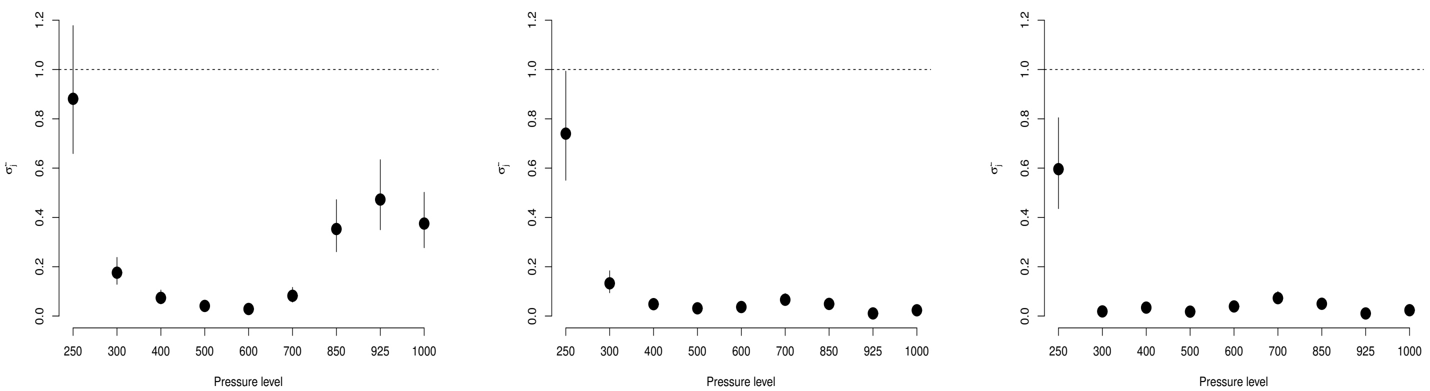
- Second position: standardized temperature at altitude pressure level of 1000 millibars.
- Third position: standardized temperature at altitude pressure level of 250 millibars, which is close to the tropopause and seems to behave differently from the temperatures at lower altitudes.
- The order for the remaining variables is the original order from lower to upper altitudes.

Choice of the number of factors

Table: Troposphere temperature. Logarithm of marginal data density for factor models with distinct number of factors.

	Number of factors				
	1	2	3	4	5
Log marginal density	89.7	319.2	326.3	310.0	300.6

Choice of the number of factors



(a) (b) (c)

Figure: Troposphere temperature – k-factor models. Idiosyncratic variances σ_j^2 based on factor models with (a) 1, (b) 2, and (c) 3 factors.

Analysis with 3-factor model

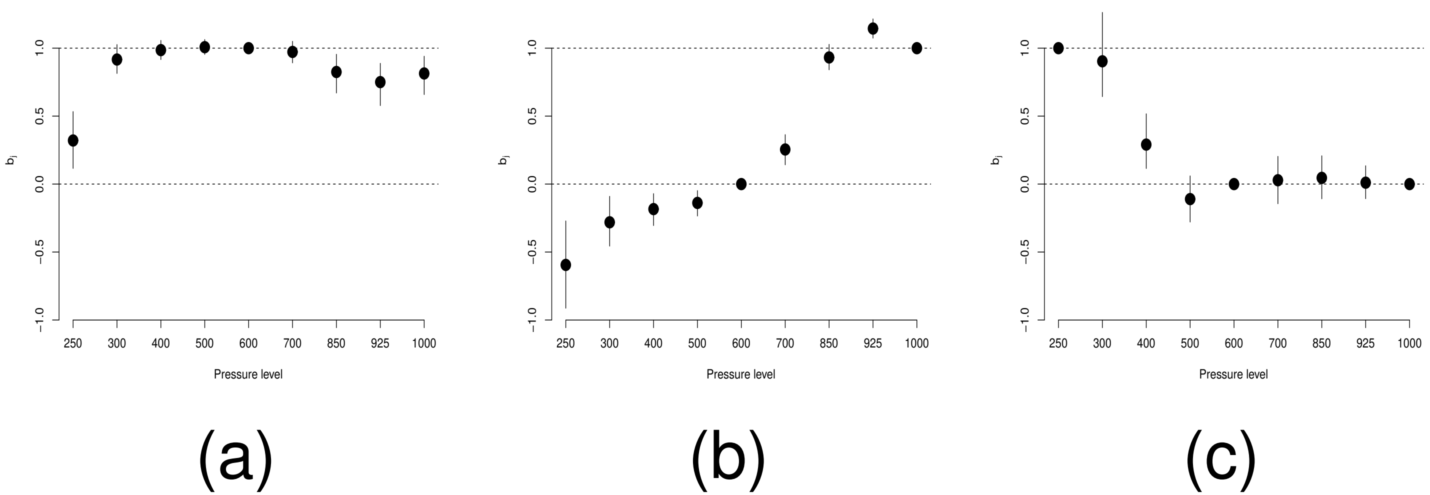
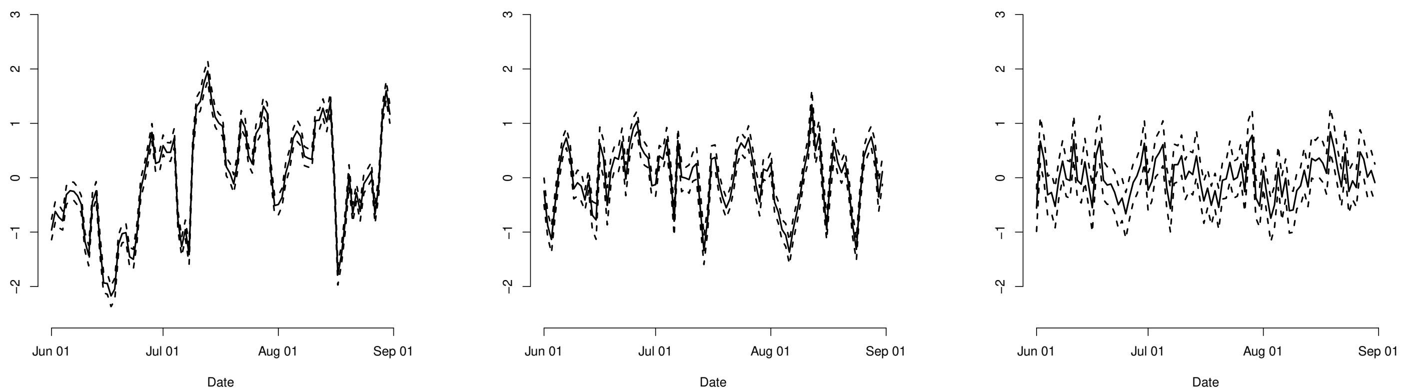


Figure: Troposphere temperature – 3-factor model. Factor loadings for (a) first factor, (b) second factor, and (c) third factor.

Analysis with 3-factor model



(a) (b) (c)

Figure: Troposphere temperature – 3-factor model. Time series plots of common factors posterior means (solid line) and 95% credible intervals for (a) first factor, (b) second factor, and (c) third factor.

Marco Ferreira (PF&W 11.2-11.4)

- Dynamic latent factor models
- Example: time series troposphere temperature
- Dynamic spatiotemporal factor model
- Example: spatiotemporal troposphere temperature

Marco Ferreira (PF&W 11.2-11.4)

- Dynamic latent factor models
- Example: time series troposphere temperature
- Dynamic spatiotemporal factor model
- Example: spatiotemporal troposphere temperature

Marco Ferreira (PF&W 11.2-11.4)

- Dynamic latent factor models
- Example: time series troposphere temperature
- Dynamic spatiotemporal factor model
- Example: spatiotemporal troposphere temperature

Marco Ferreira (PF&W 11.2-11.4)

- Dynamic latent factor models
- Example: time series troposphere temperature
- Dynamic spatiotemporal factor model
- Example: spatiotemporal troposphere temperature

Dynamic latent factor model

Putting together latent factor models and DLMs:

$$\begin{aligned}\mathbf{y}_t &= \mathbf{F}_t'\theta_t + \mathbf{B}_t\mathbf{x}_t + \nu_t, \quad \nu_t \sim N(\mathbf{0}, \mathbf{V}_t),\\ \theta_t &= \mathbf{G}_t\theta_{t-1} + \omega_t, \quad \omega_t \sim N(\mathbf{0}, \mathbf{W}_t),\end{aligned}$$

- $\mathbf{x}_t = (x_{t,1}\ldots, x_{t,k})'$, a k-vector dynamic latent factor process,
- the $r \times k$ dynamic factor loadings matrix $\mathbf{B}_t$ provides coefficients that map factors to observations.

Dynamic latent factor model

Putting together latent factor models and DLMs:

$$\begin{aligned}\mathbf{y}_t &= \mathbf{F}_t'\theta_t + \mathbf{B}_t\mathbf{x}_t + \nu_t, \quad \nu_t \sim N(\mathbf{0}, \mathbf{V}_t),\\ \theta_t &= \mathbf{G}_t\theta_{t-1} + \omega_t, \quad \omega_t \sim N(\mathbf{0}, \mathbf{W}_t),\end{aligned}$$

- $\mathbf{x}_t = (x_{t,1}\ldots, x_{t,k})'$, a k-vector dynamic latent factor process,
- the $r \times k$ dynamic factor loadings matrix $\mathbf{B}_t$ provides coefficients that map factors to observations.

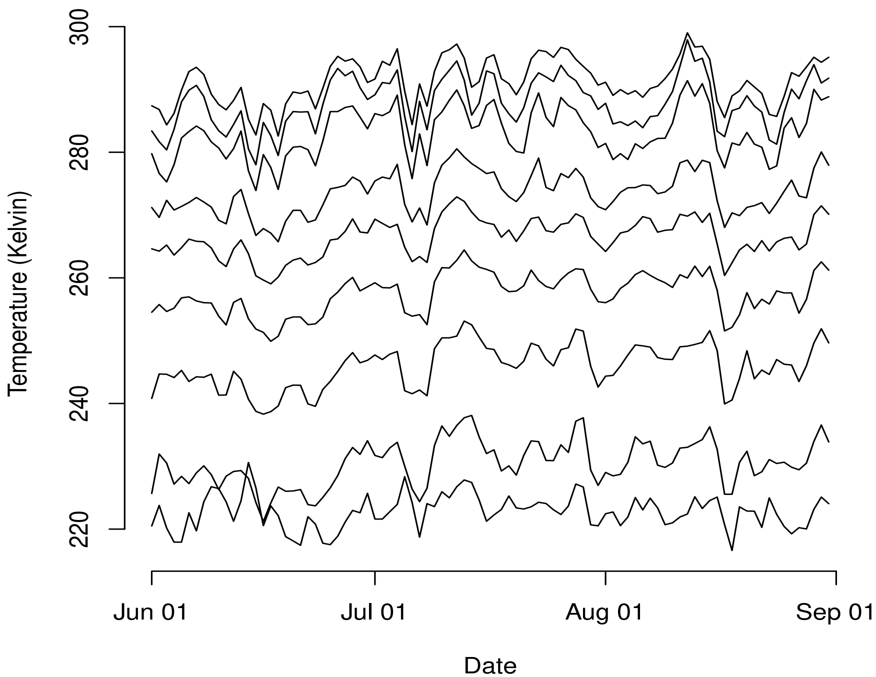
Dynamic latent factor model

Putting together latent factor models and DLMs:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{F}_t'\theta_t + \mathbf{B}_t\mathbf{x}_t + \nu_t, \quad \nu_t \sim N(\mathbf{0}, \mathbf{V}_t), \\ \theta_t &= \mathbf{G}_t\theta_{t-1} + \omega_t, \quad \omega_t \sim N(\mathbf{0}, \mathbf{W}_t), \end{aligned}$$

- $\mathbf{x}_t = (x_{t,1} \ldots, x_{t,k})'$, a k-vector dynamic latent factor process,
- the $r \times k$ dynamic factor loadings matrix $\mathbf{B}_t$ provides coefficients that map factors to observations.

Dynamic 3-factor model: troposphere temperature



- Exploratory analysis of estimated common factors in the static 3-factor model indicates that each of the three common factors may be better modeled with autoregressive processes.
- Autoregressive processes of order 2 for the first two common factors.
- Autoregressive process of order 1 for the third common factor.

- Exploratory analysis of estimated common factors in the static 3-factor model indicates that each of the three common factors may be better modeled with autoregressive processes.
- Autoregressive processes of order 2 for the first two common factors.
- Autoregressive process of order 1 for the third common factor.

- Exploratory analysis of estimated common factors in the static 3-factor model indicates that each of the three common factors may be better modeled with autoregressive processes.
- Autoregressive processes of order 2 for the first two common factors.
- Autoregressive process of order 1 for the third common factor.

Dynamic factor model for troposphere temperature

$$\begin{aligned}
\mathbf{y}_t &= \mathbf{B}\mathbf{x}_t + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim N(\mathbf{0}, \mathbf{V}),\\
\mathbf{x}_t &= \mathbf{F}'\boldsymbol{\theta}_t,\\
\boldsymbol{\theta}_t &= \mathbf{G}\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim N(\mathbf{0}, \mathbf{W}).
\end{aligned}$$

- $\mathbf{B}$ fixed through time,
- $\mathbf{x}_t$ is a k-dimensional vector of common factors with lth element x_{tl} following an AR(p_l) process.

Dynamic factor model for troposphere temperature

$$\begin{aligned} \mathbf{y}_t &= \mathbf{B}\mathbf{x}_t + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim N(\mathbf{0}, \mathbf{V}), \\ \mathbf{x}_t &= \mathbf{F}'\boldsymbol{\theta}_t, \\ \boldsymbol{\theta}_t &= \mathbf{G}\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim N(\mathbf{0}, \mathbf{W}). \end{aligned}$$

- $\mathbf{B}$ fixed through time,
- $\mathbf{x}_t$ is a k-dimensional vector of common factors with lth element x_{tl} following an AR(p_l) process.

Dynamic factor model for troposphere temperature

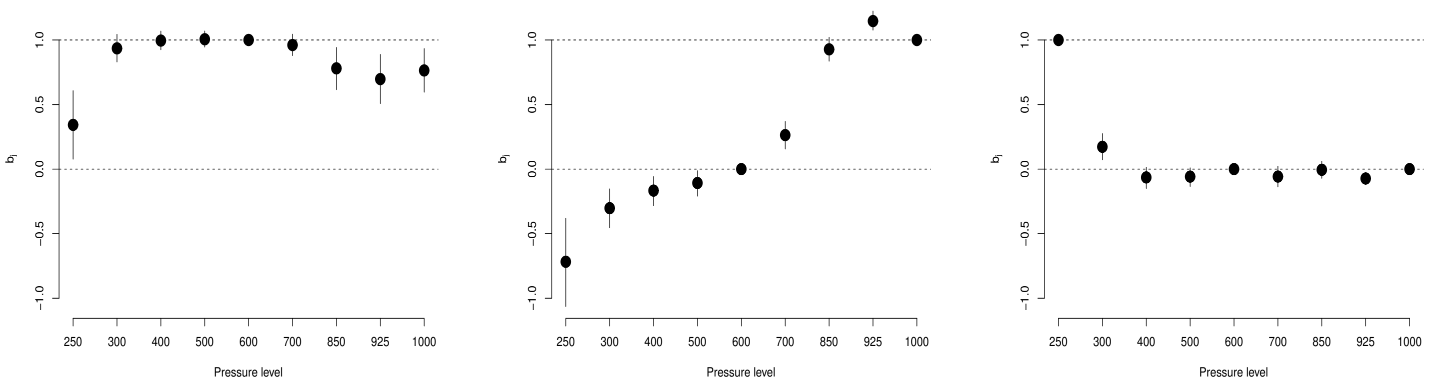
$$\begin{aligned}
\mathbf{y}_t &= \mathbf{B}\mathbf{x}_t + \nu_t, \quad \nu_t \sim N(\mathbf{0}, \mathbf{V}), \\
\mathbf{x}_t &= \mathbf{F}'\theta_t, \\
\theta_t &= \mathbf{G}\theta_{t-1} + \omega_t, \quad \omega_t \sim N(\mathbf{0}, \mathbf{W}).
\end{aligned}$$

- $\mathbf{B}$ fixed through time,
- $\mathbf{x}_t$ is a k-dimensional vector of common factors with lth element x_{tl} following an AR(p_l) process.

- Matrices **F**, **G**, and **W** are block-diagonal matrices with lth block representing in state-space form the AR(p_l) process for x_{tl}.

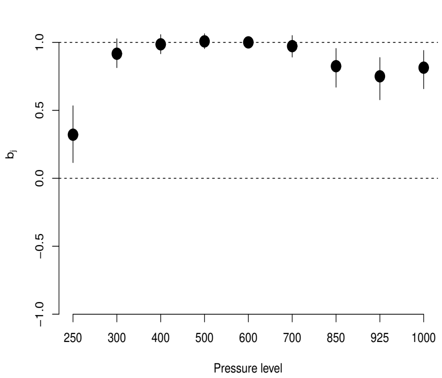
- $\mathbf{V} = diag(\sigma_1^2, \ldots, \sigma_r^2)$ where $\sigma_1^2, \ldots, \sigma_r^2$ are idiosyncratic variances.

Dynamic 3-factor model: Factor loadings

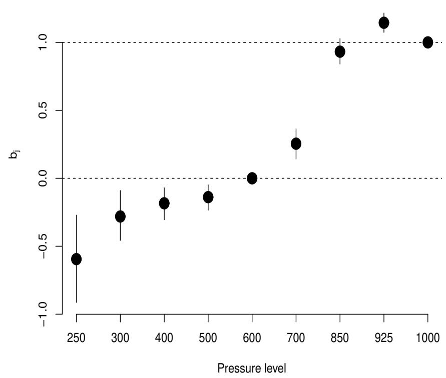


1st factor 2nd factor 3rd factor

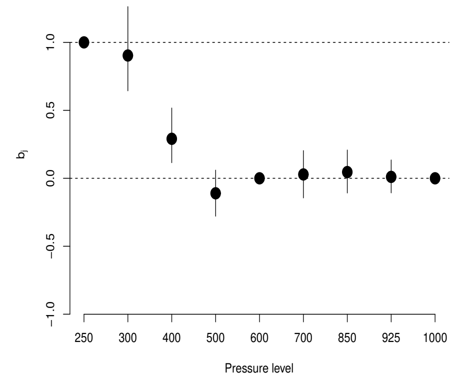
Compare: static 3-factor model loadings



1st factor

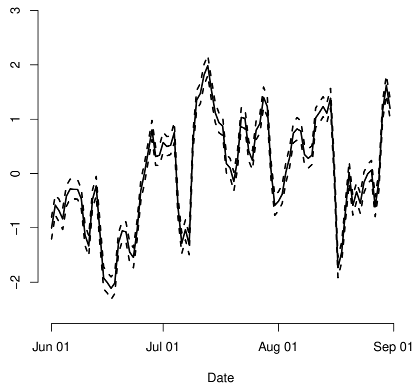


2nd factor

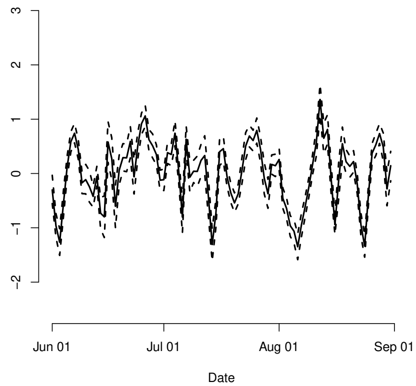


3rd factor

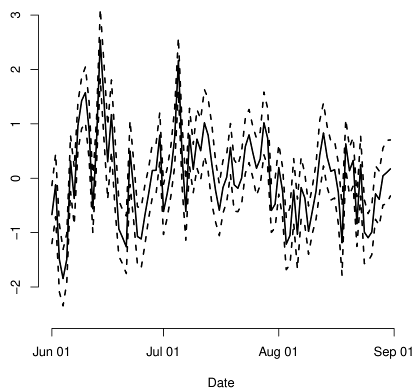
Dynamic 3-factor model: Estimated latent factors



1st factor

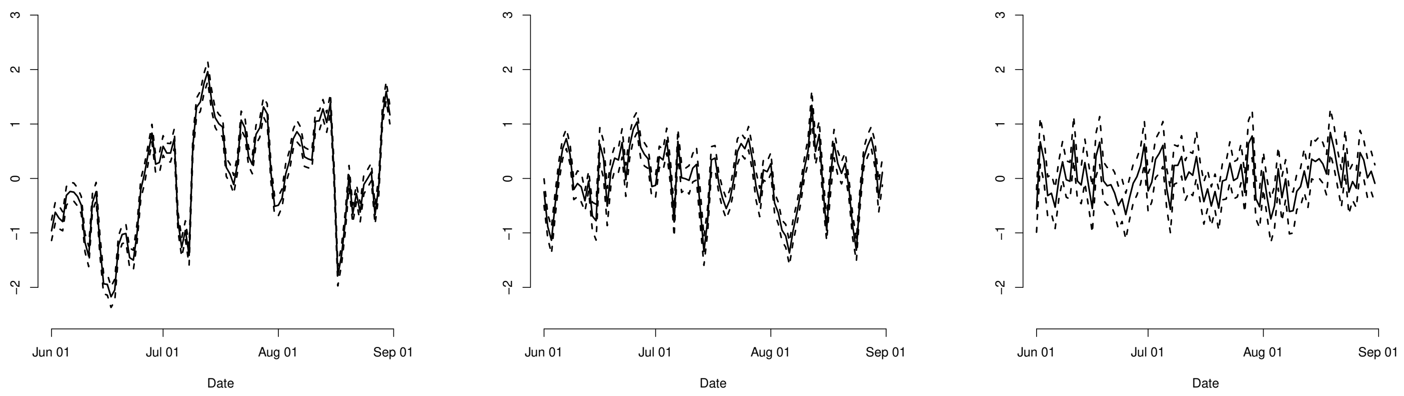


2nd factor



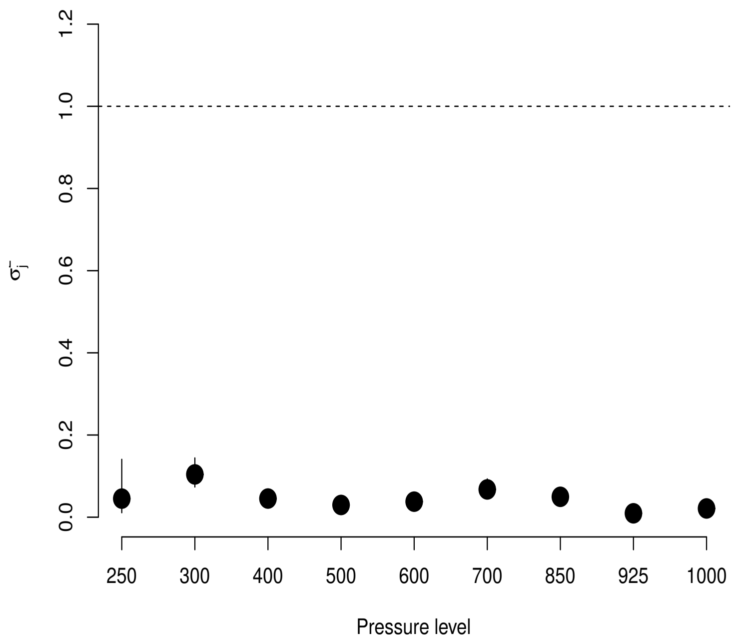
3rd factor

Compare: latent factors for static 3-factor model

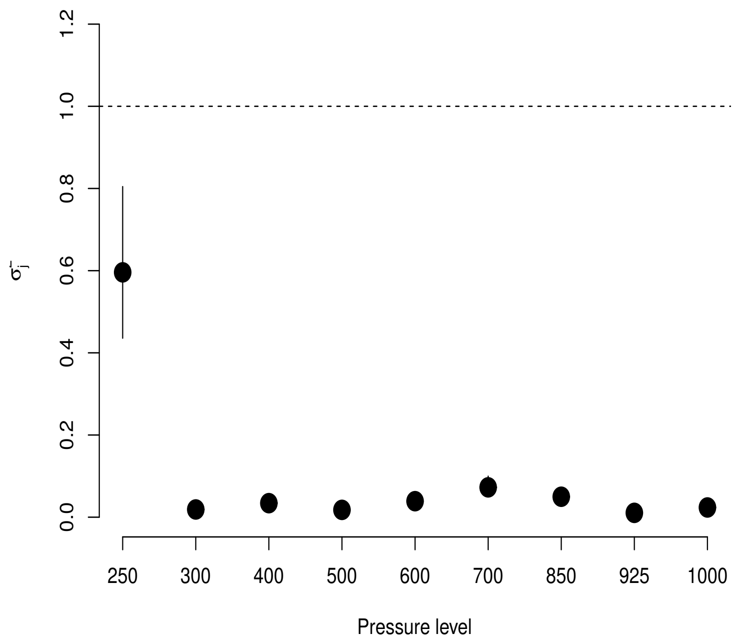


1st factor 2nd factor 3rd factor

Compare: Idiosyncratic variances

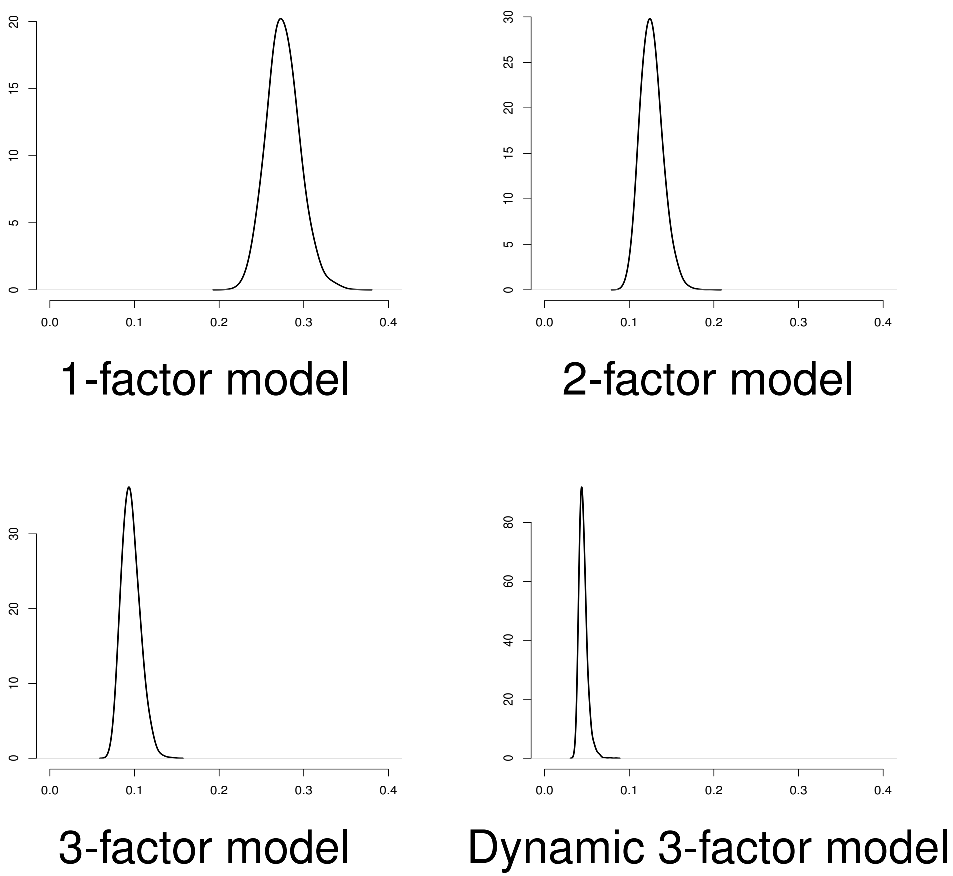


Dynamic 3-factor model

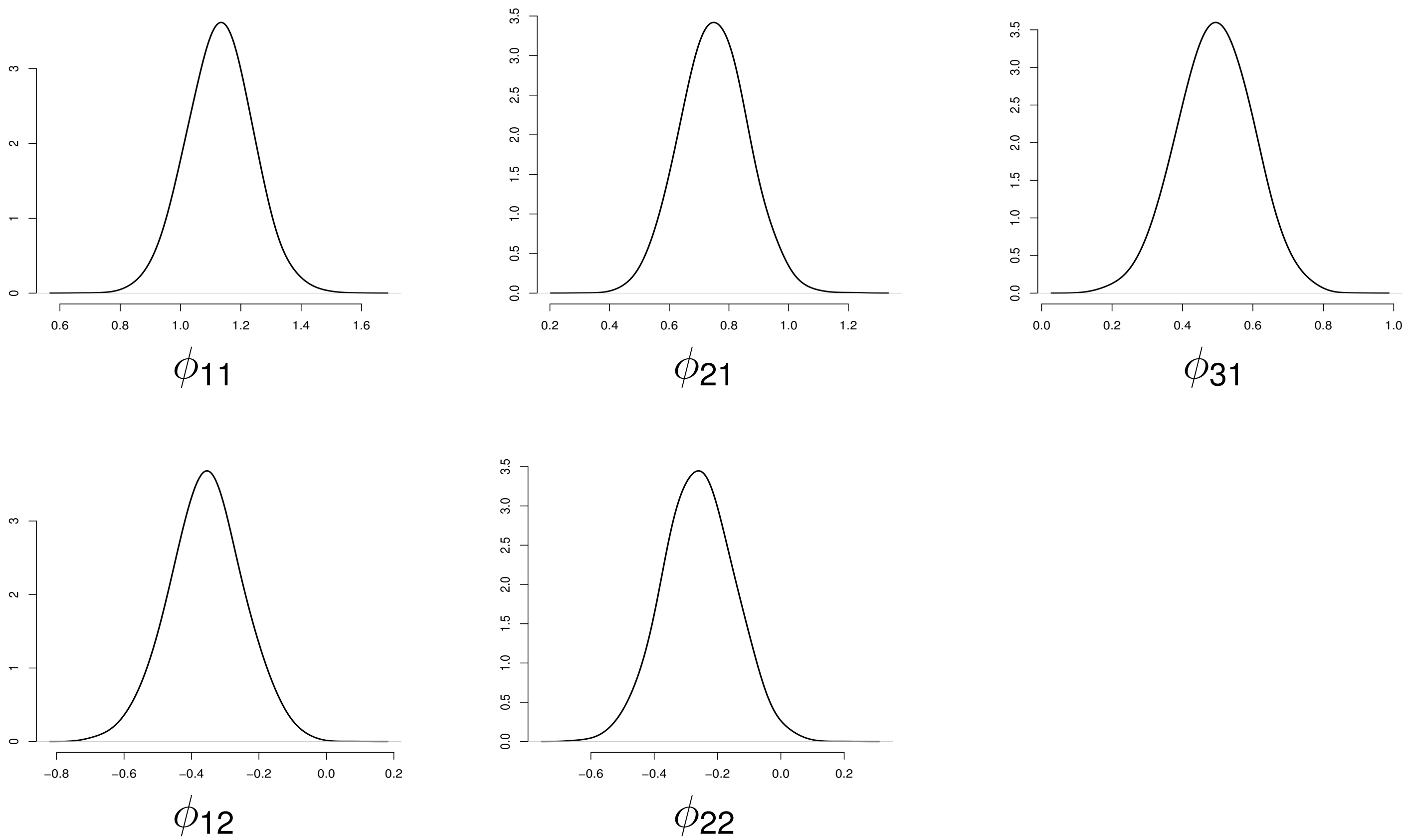


Static 3-factor model

Proportion of unexplained variance



Dynamic 3-factor model: autoregressive coefficients for common factors



- Posterior means: $\widehat{\phi}_{11} = 1.13$, $\widehat{\phi}_{12} = -0.36$, $\widehat{\phi}_{21} = 0.75$, $\widehat{\phi}_{22} = -0.26$, and $\widehat{\phi}_{31} = 0.49$.

- Roots of the characteristic polynomials for the first and second common factors are complex conjugates, implying that these common factors have stochastic cycles.

- Posterior means: $\widehat{\phi}_{11} = 1.13$, $\widehat{\phi}_{12} = -0.36$, $\widehat{\phi}_{21} = 0.75$, $\widehat{\phi}_{22} = -0.26$, and $\widehat{\phi}_{31} = 0.49$.
- Roots of the characteristic polynomials for the first and second common factors are complex conjugates, implying that these common factors have stochastic cycles.

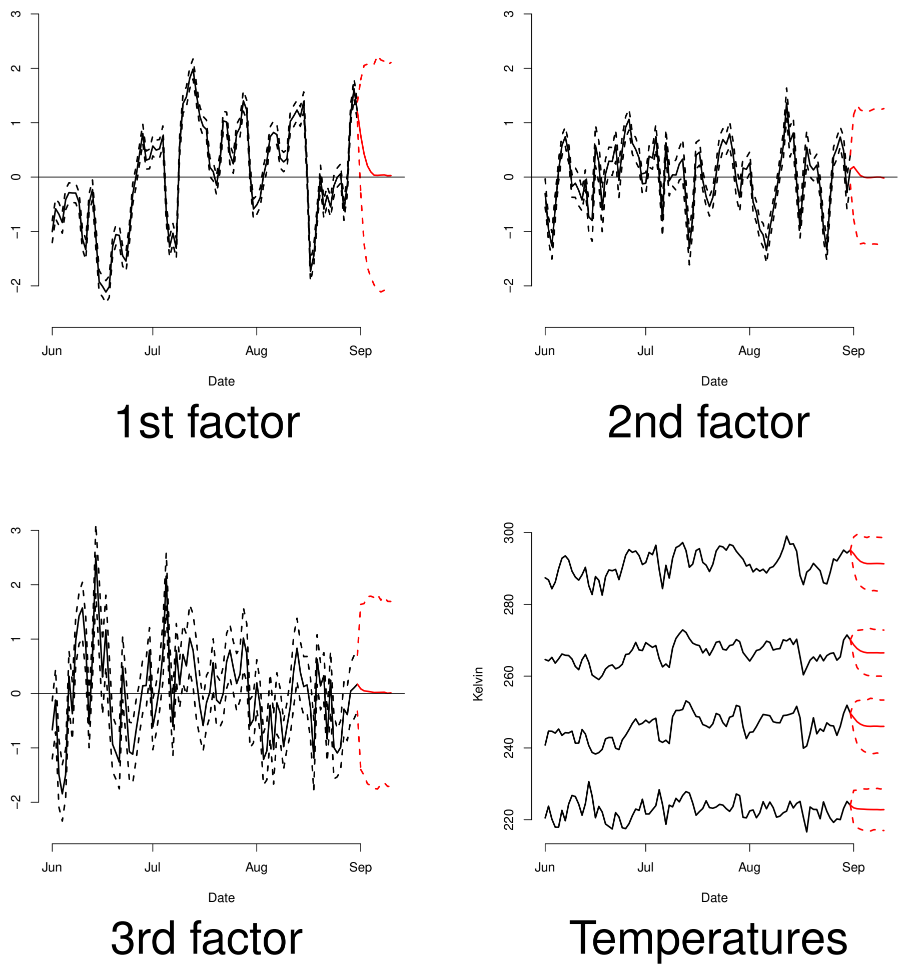
- First common factor: estimated period is 19.3 days; estimated modulus of reciprocal roots is 0.59.
- Second common factor: estimated period is 8.4 days; estimated modulus of reciprocal roots is 0.51.

- First common factor: estimated period is 19.3 days; estimated modulus of reciprocal roots is 0.59.
- Second common factor: estimated period is 8.4 days; estimated modulus of reciprocal roots is 0.51.

- Estimated moduli of the reciprocal roots for the three dynamic common factors are all less than 0.60.
- Therefore, the forecast functions for each of these three common factors decay to zero fairly fast.

- Estimated moduli of the reciprocal roots for the three dynamic common factors are all less than 0.60.
- Therefore, the forecast functions for each of these three common factors decay to zero fairly fast.

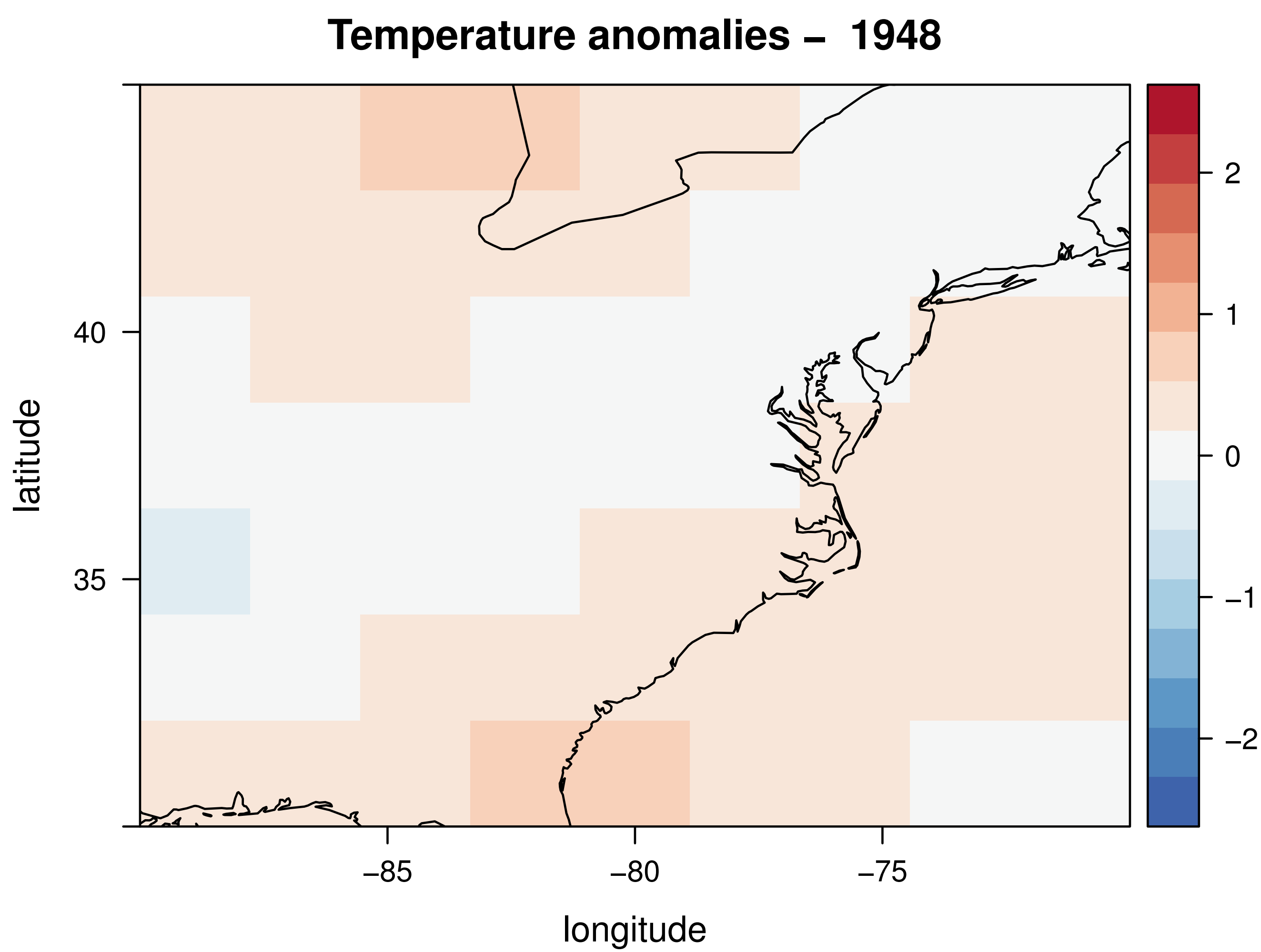
Dynamic 3-factor model: forecasts



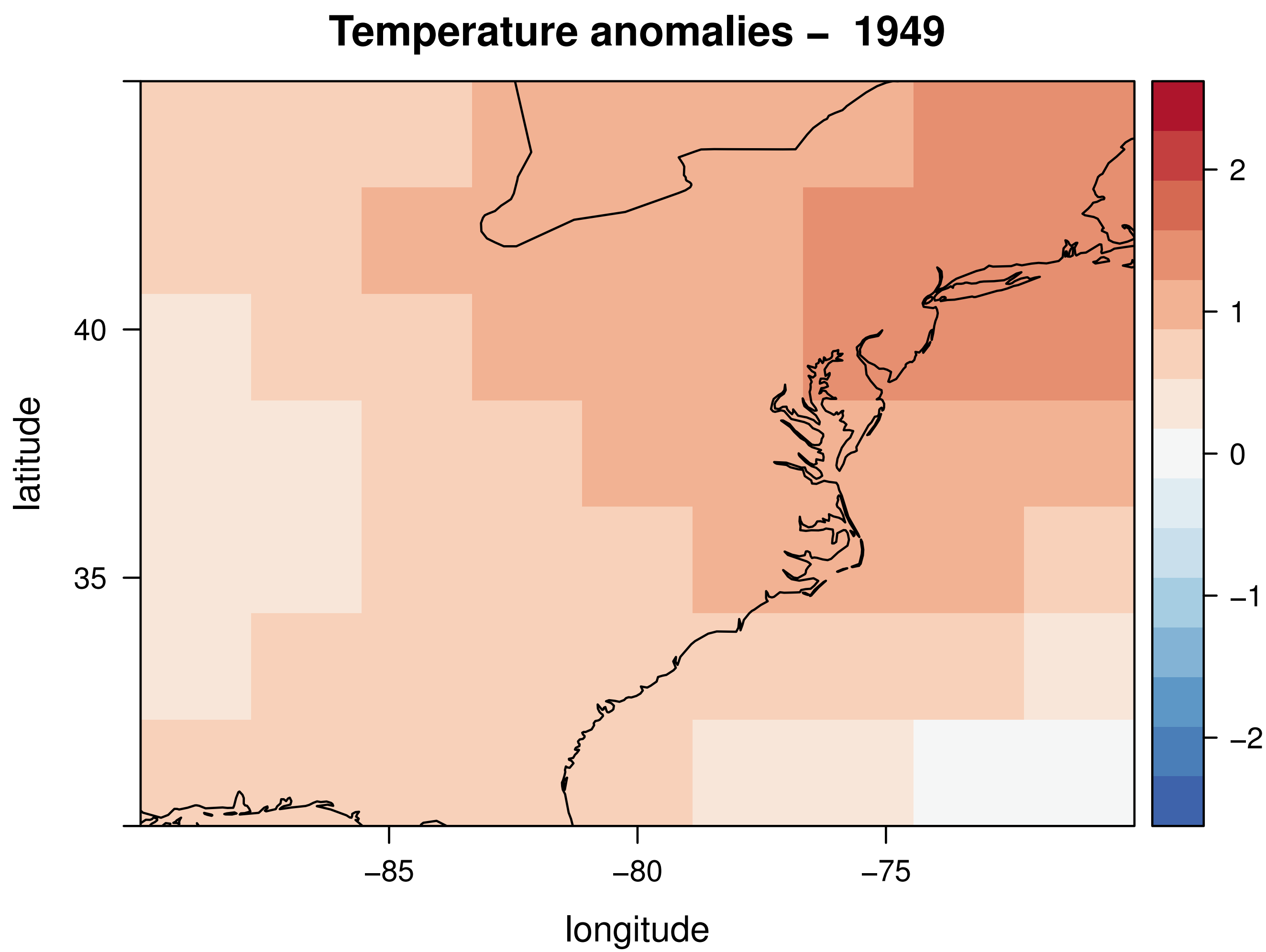
Working example: Temperature Anomalies in Eastern USA

- Temperature anomalies (Kelvin) in Eastern USA.
- Pressure level 1000 Psi.
- From 1948 to 2016.
- NCEP Reanalysis dataset.

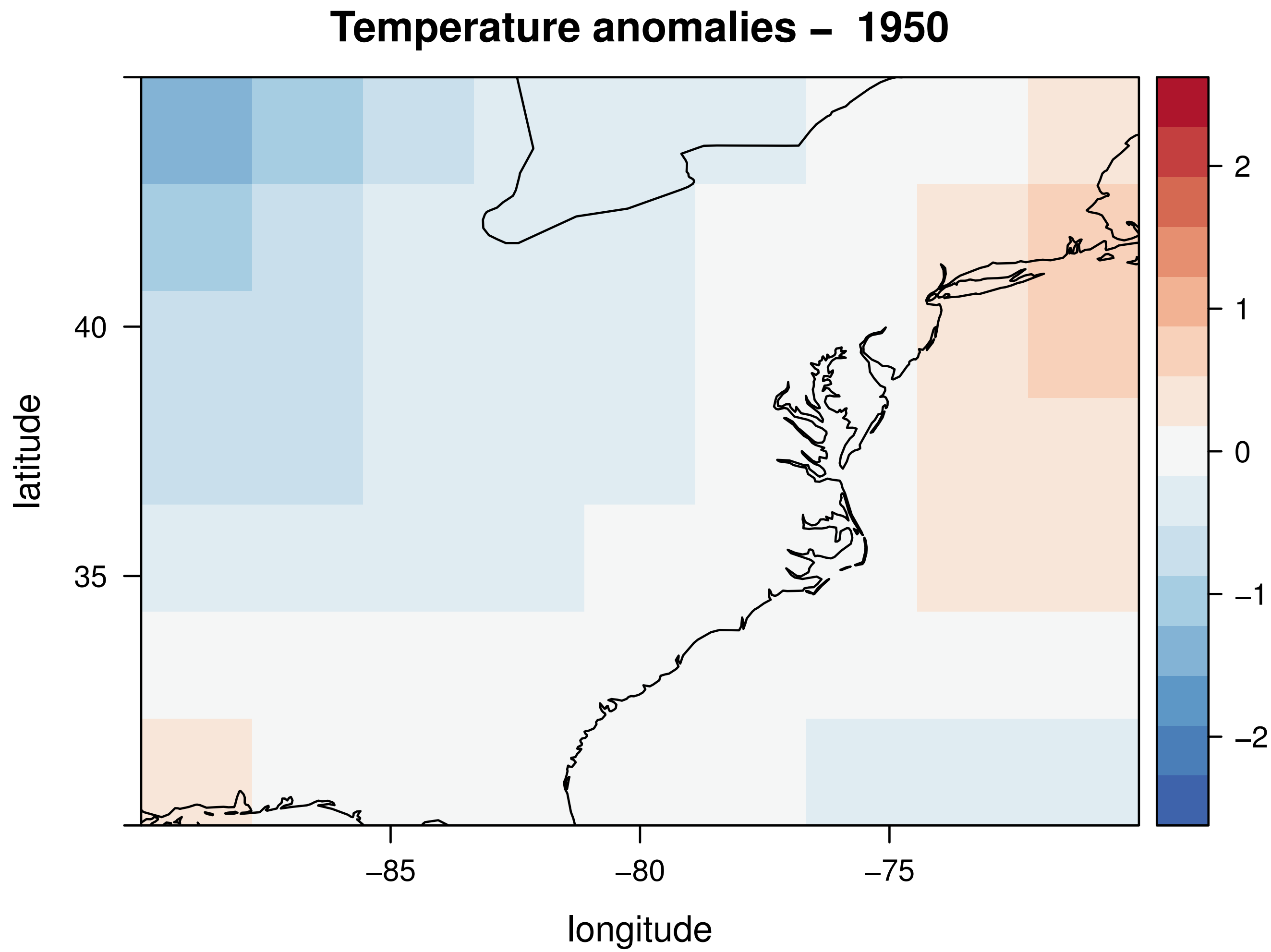
Temperature Anomalies in Eastern USA



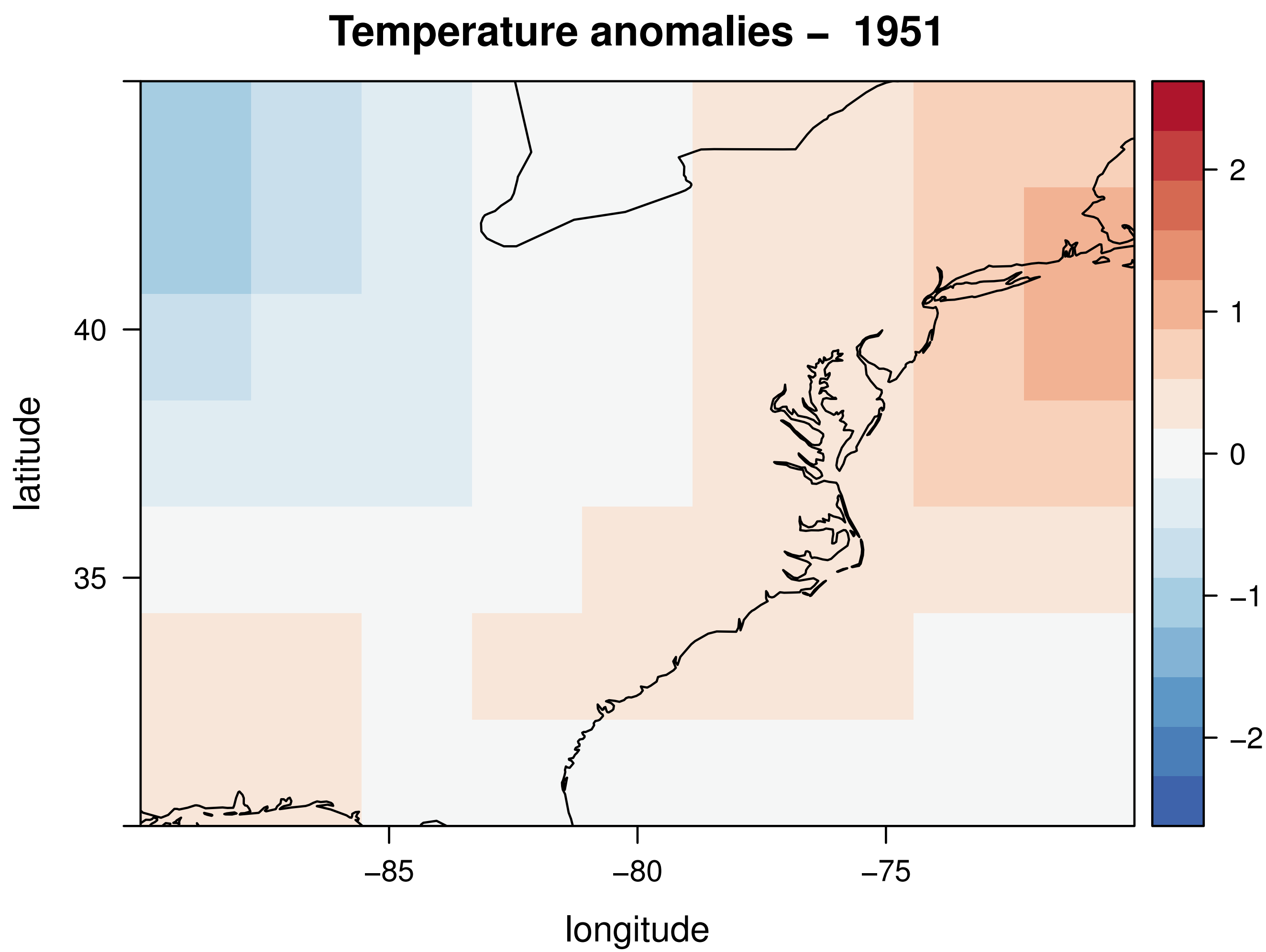
Temperature Anomalies in Eastern USA



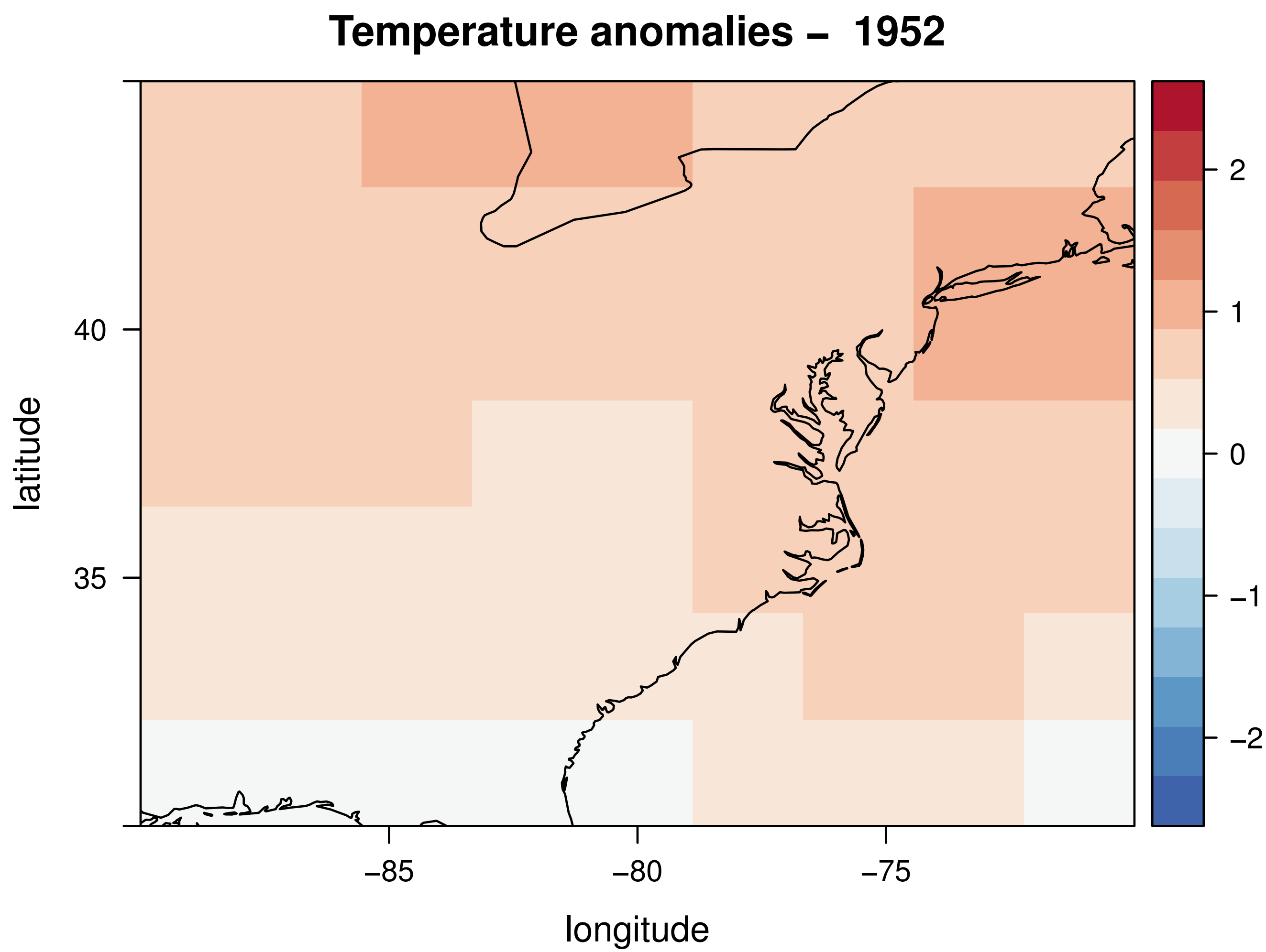
Temperature Anomalies in Eastern USA



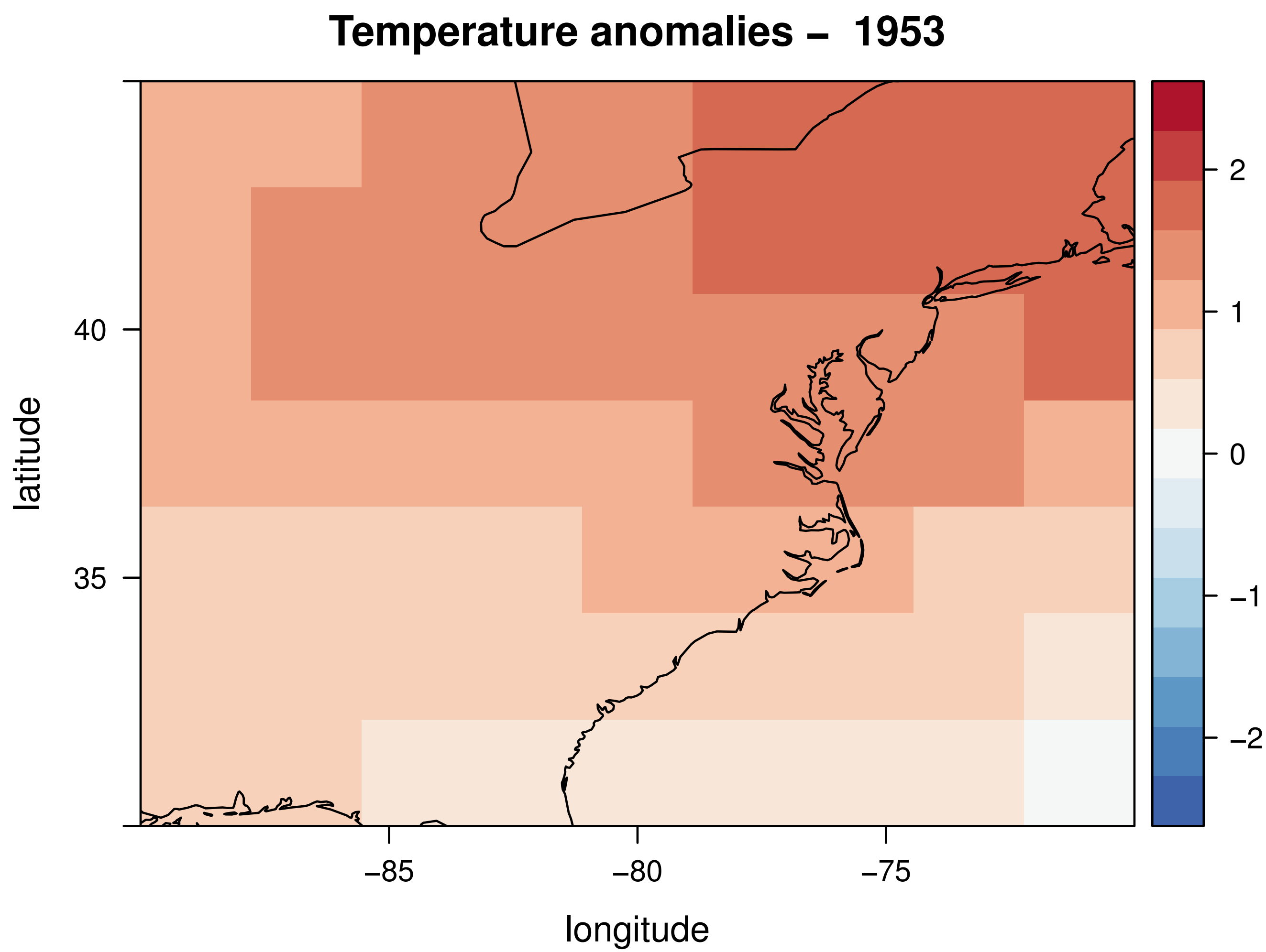
Temperature Anomalies in Eastern USA



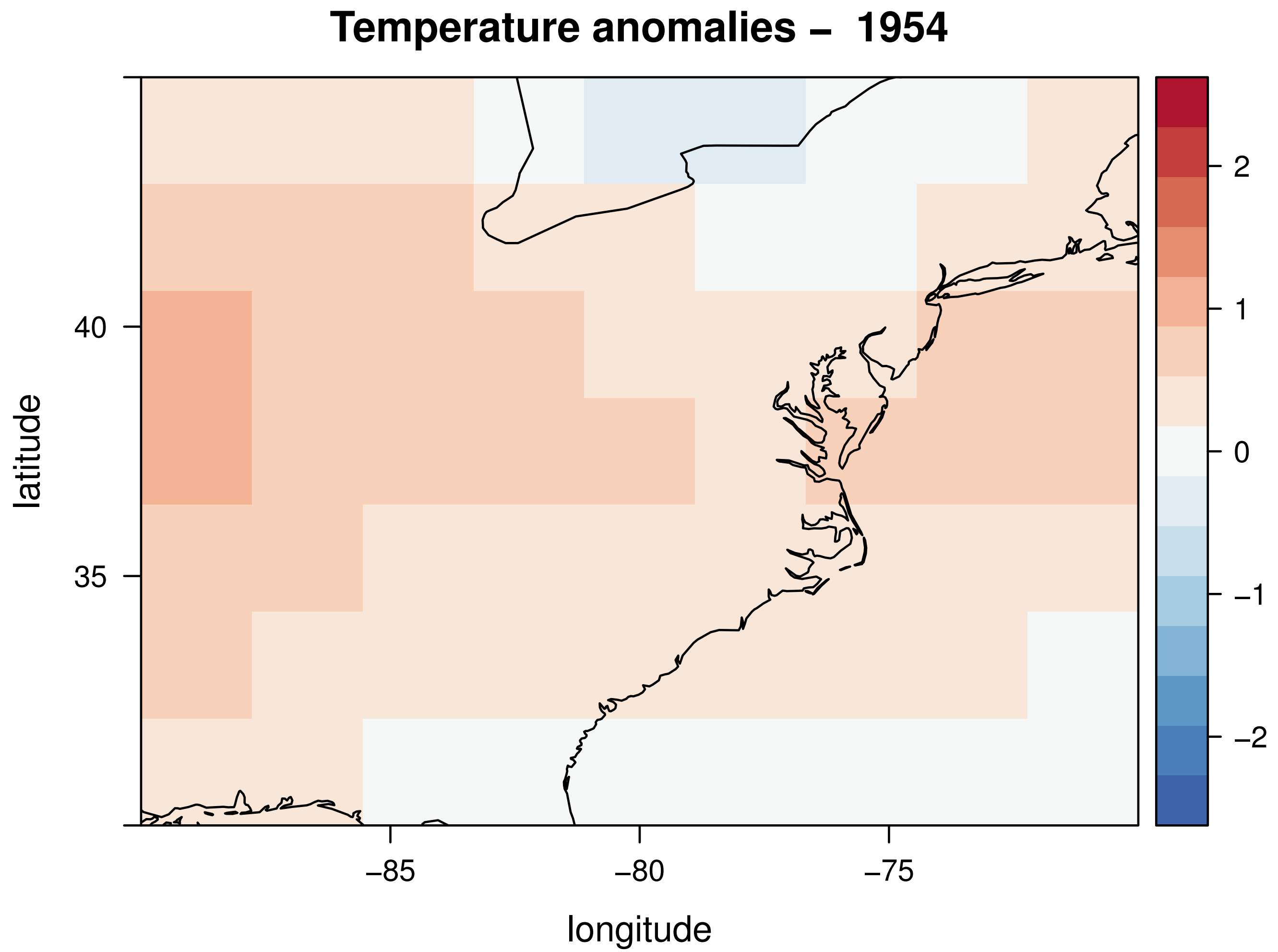
Temperature Anomalies in Eastern USA



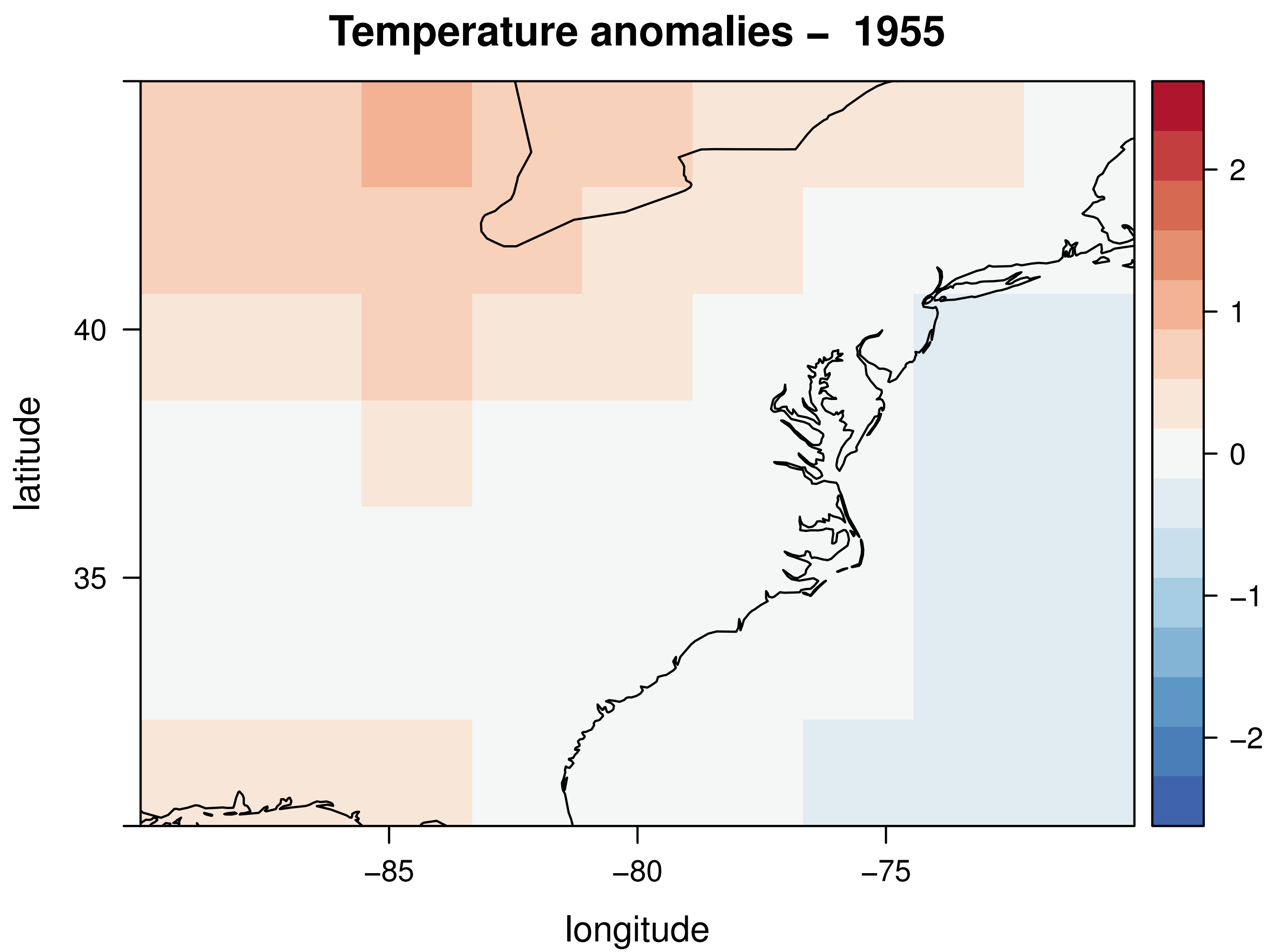
Temperature Anomalies in Eastern USA



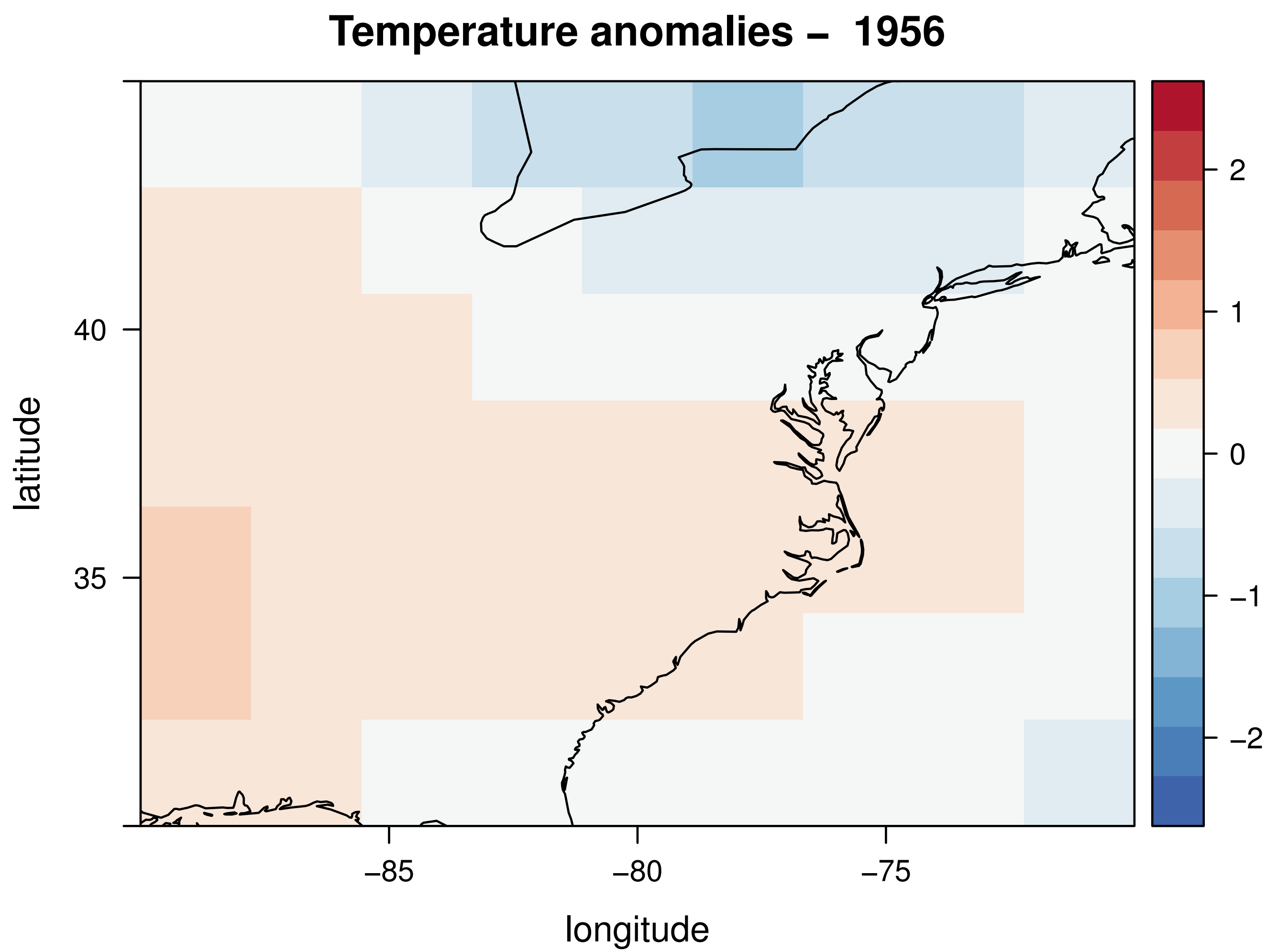
Temperature Anomalies in Eastern USA



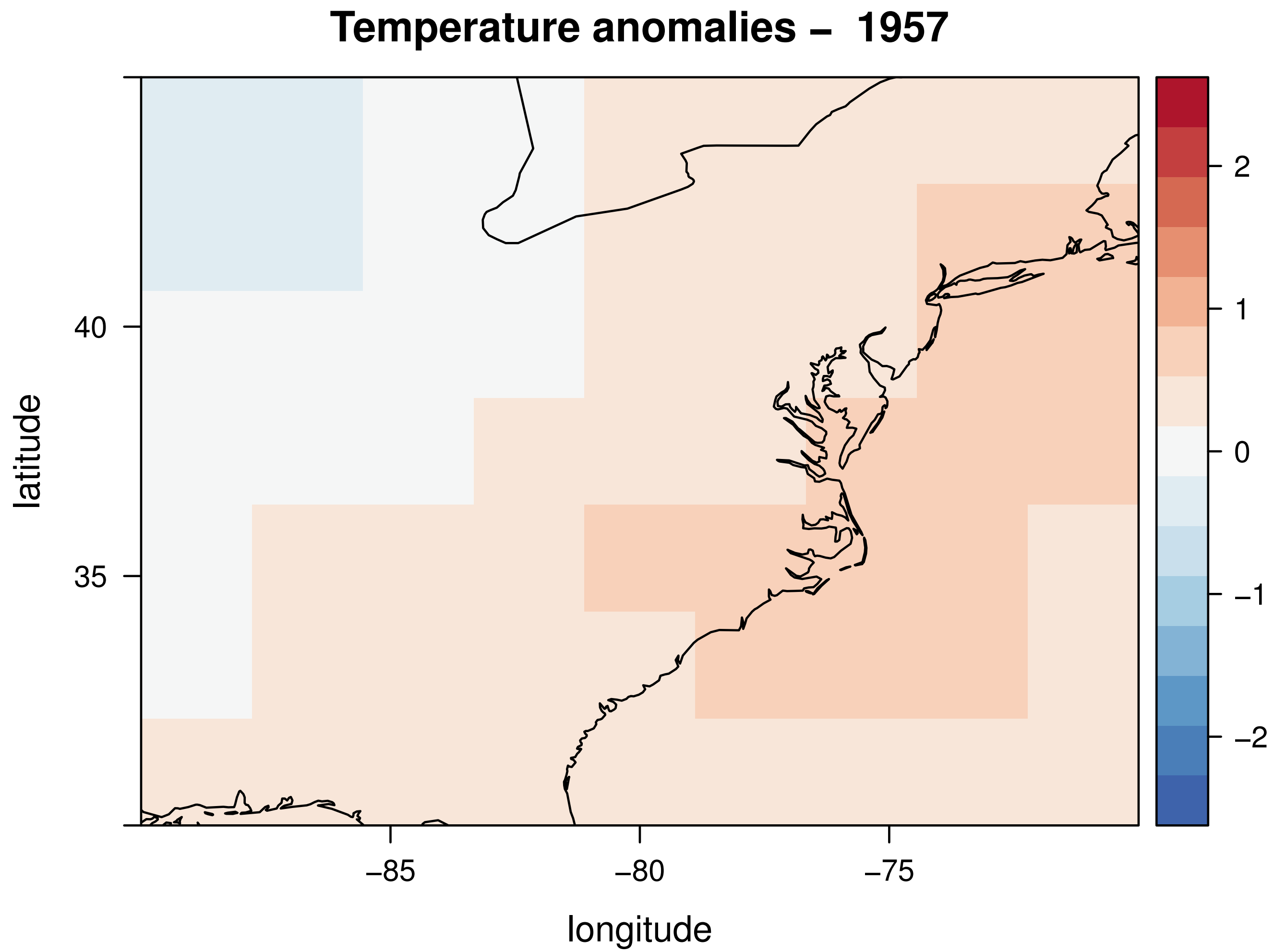
Temperature Anomalies in Eastern USA



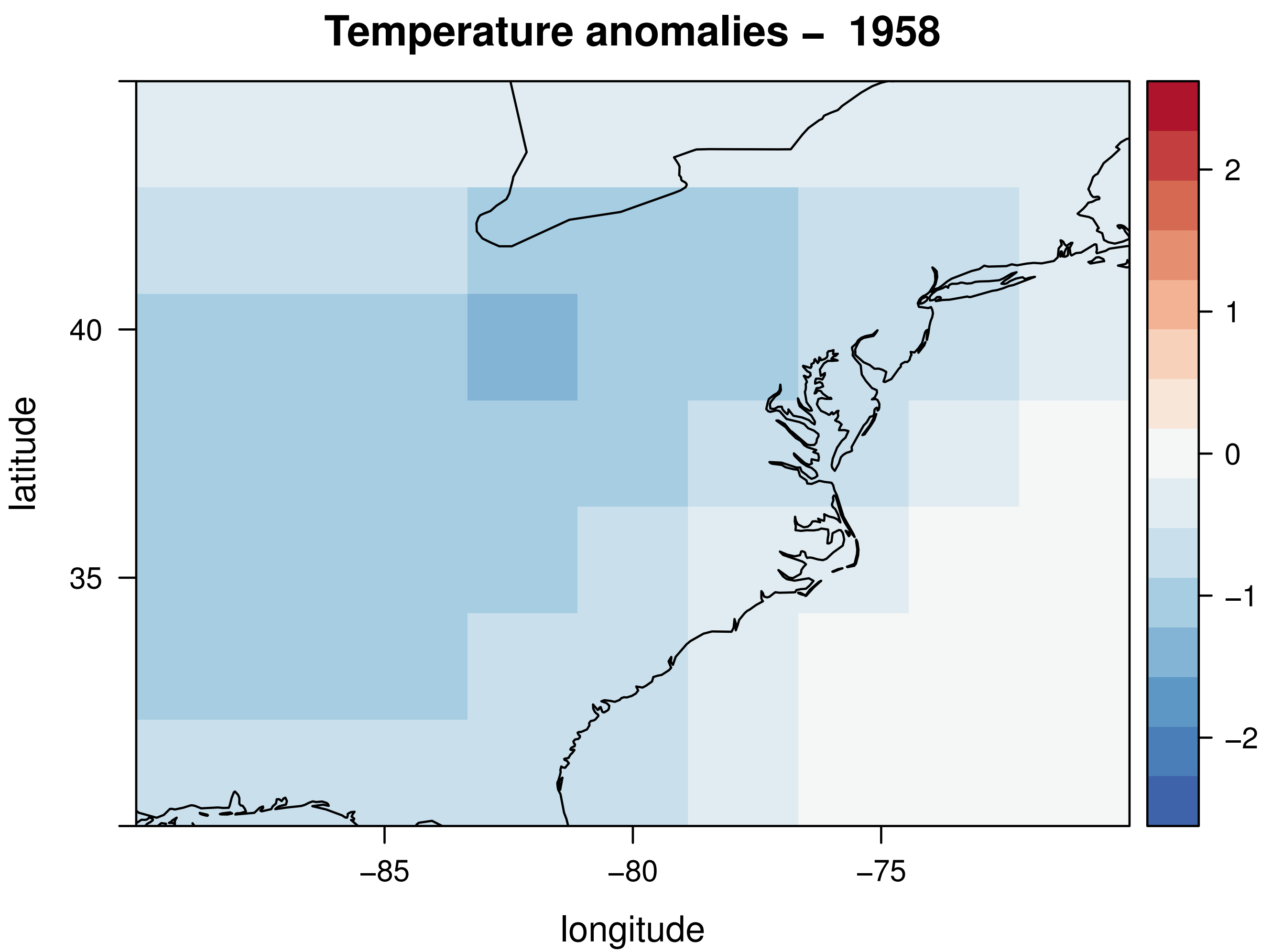
Temperature Anomalies in Eastern USA



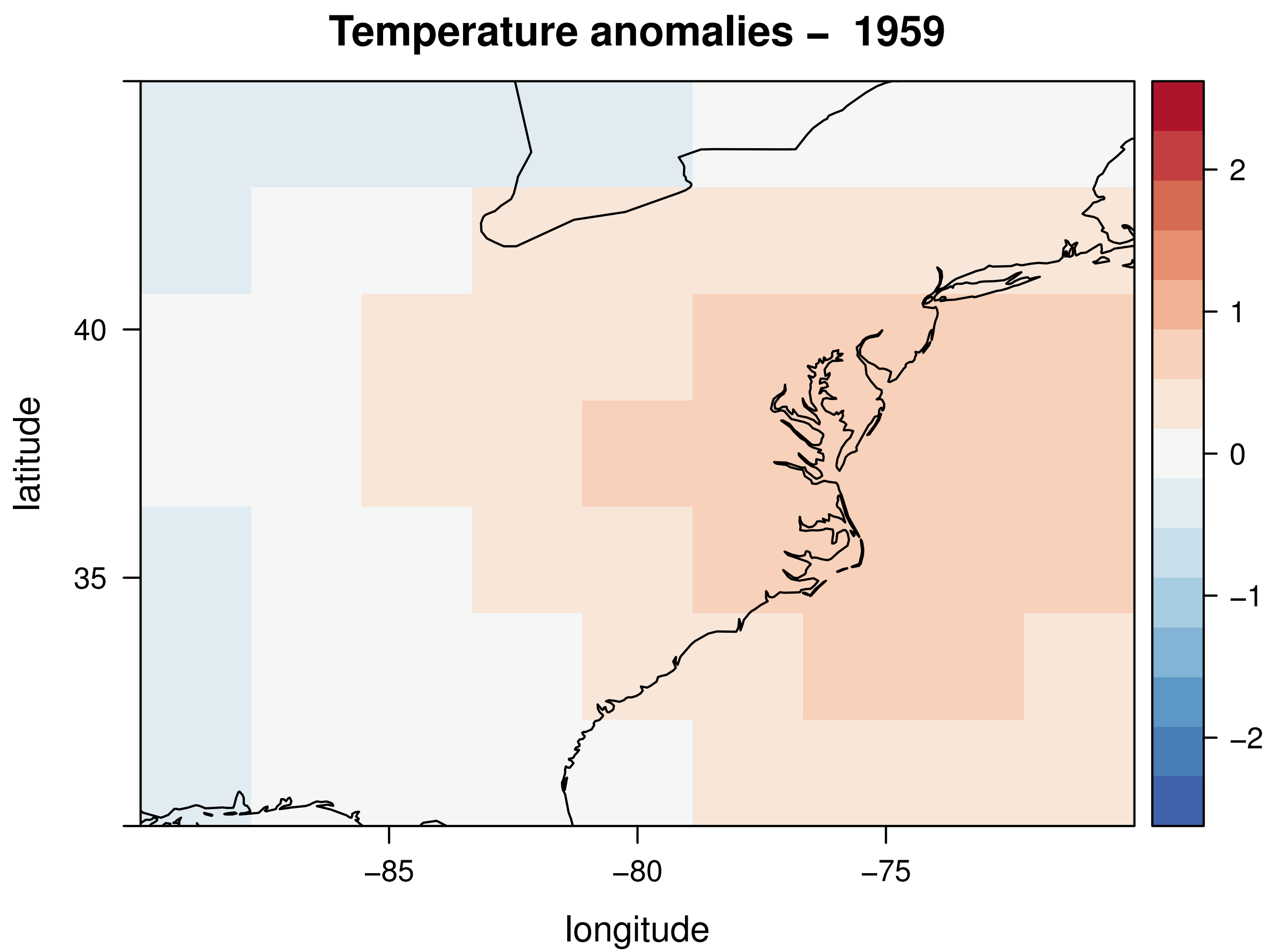
Temperature Anomalies in Eastern USA



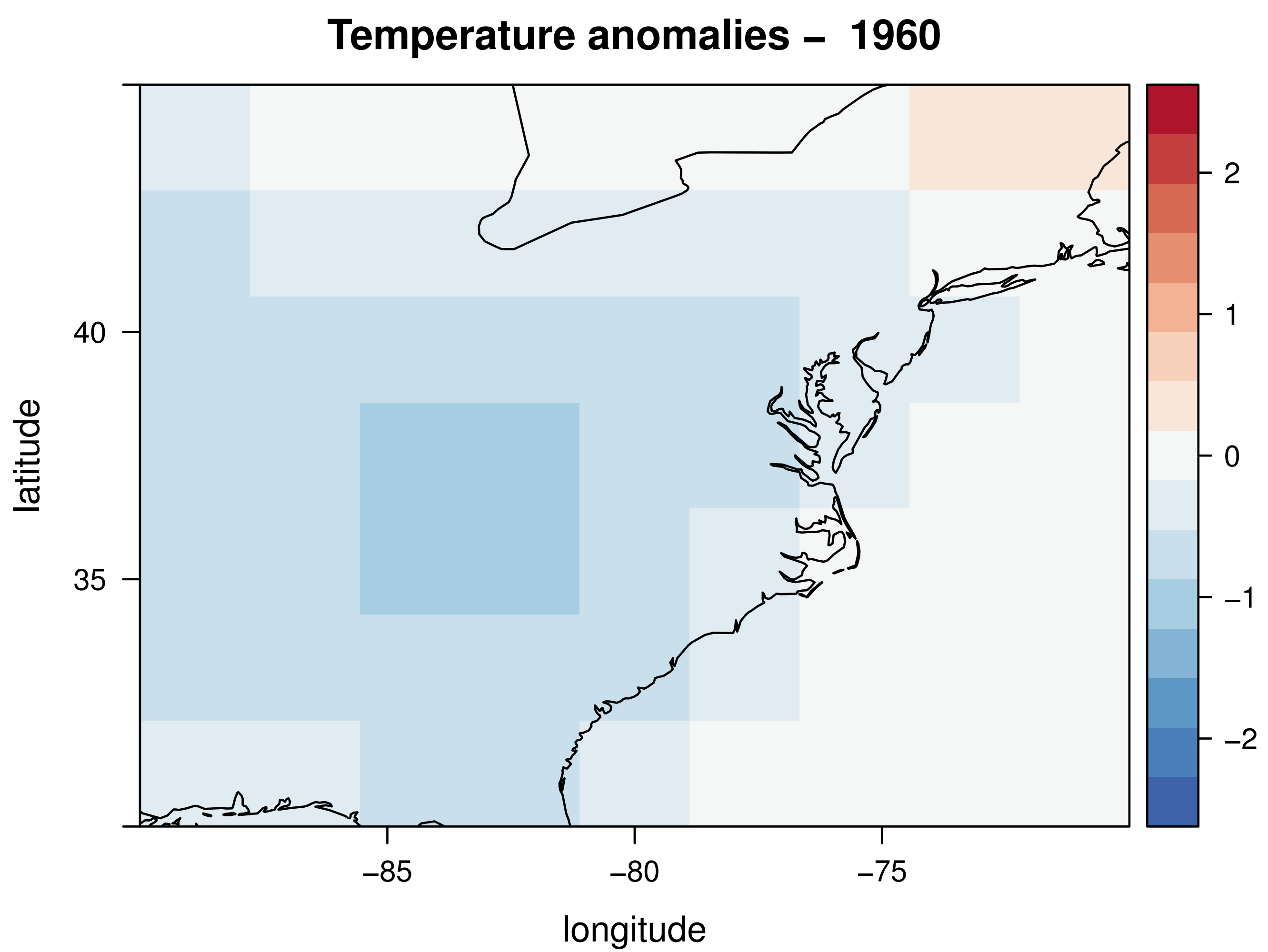
Temperature Anomalies in Eastern USA



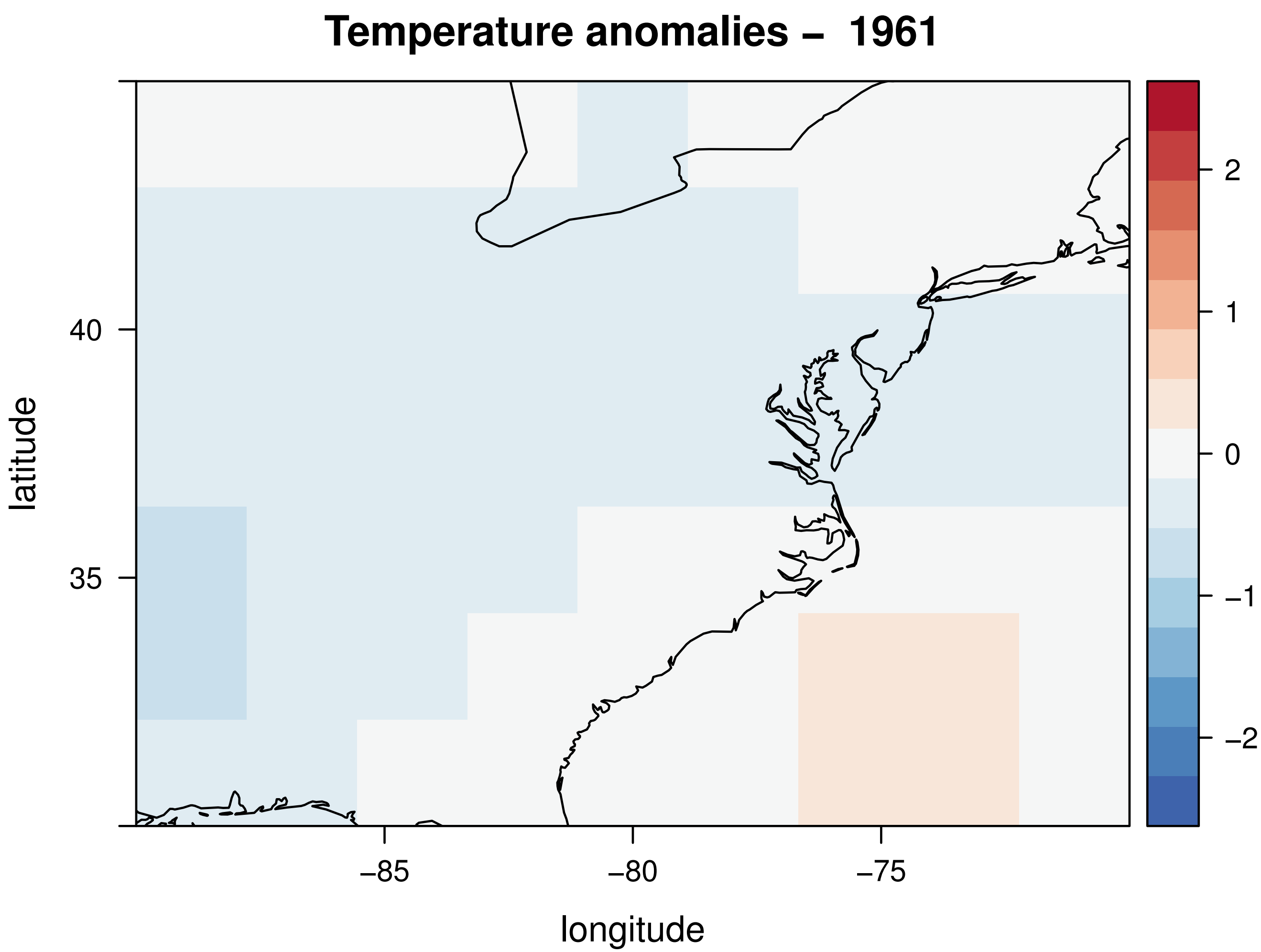
Temperature Anomalies in Eastern USA



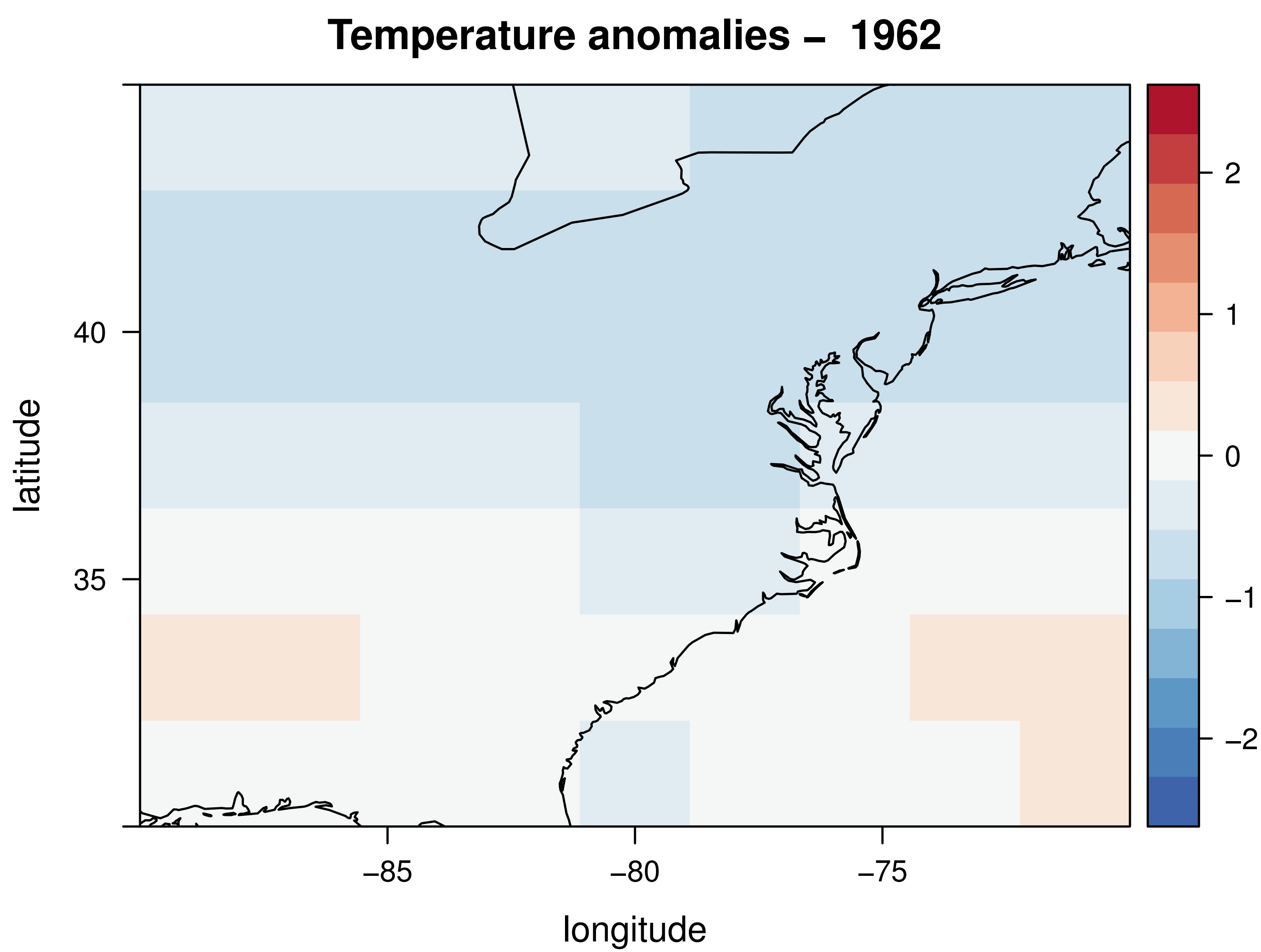
Temperature Anomalies in Eastern USA



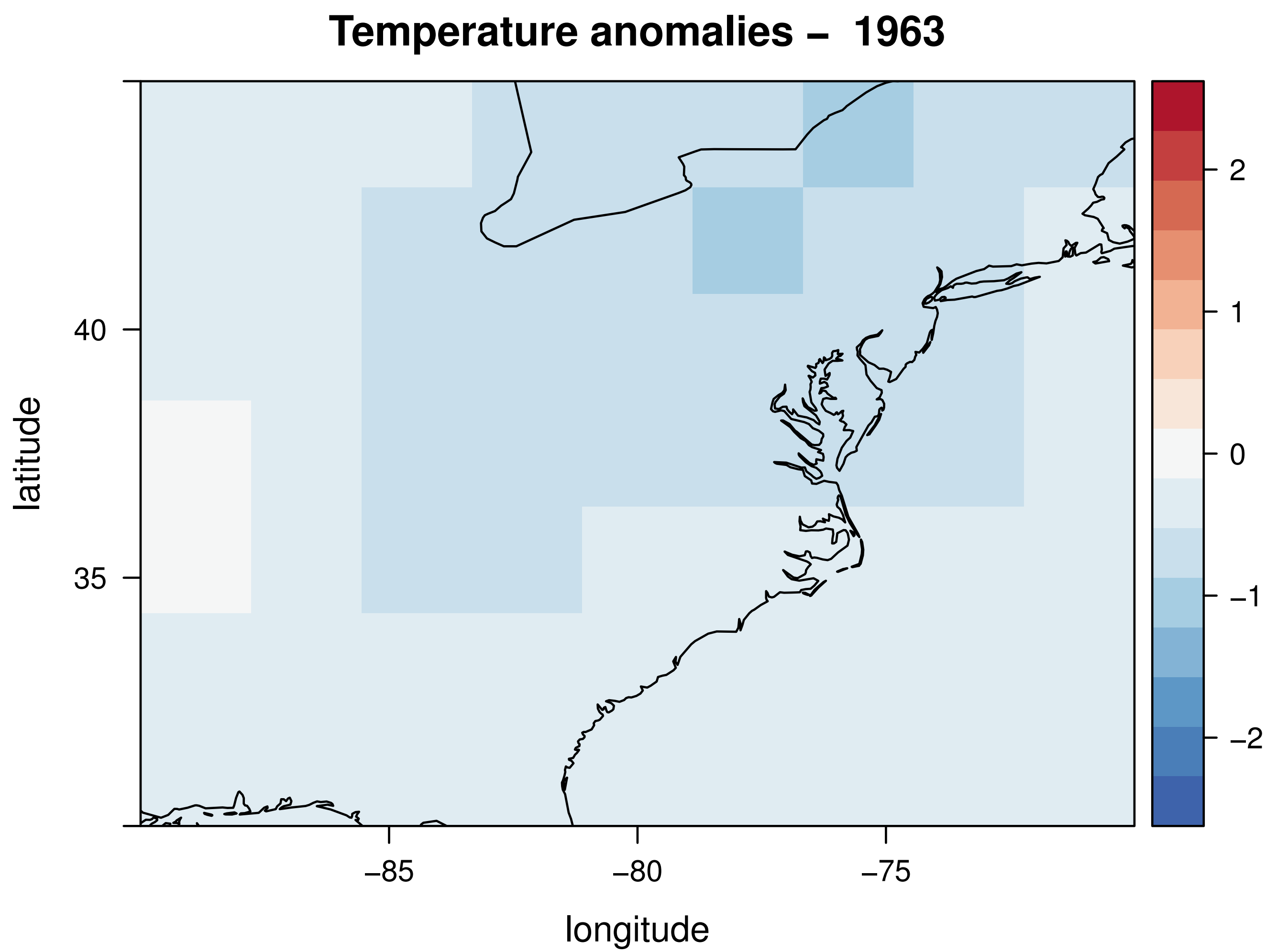
Temperature Anomalies in Eastern USA



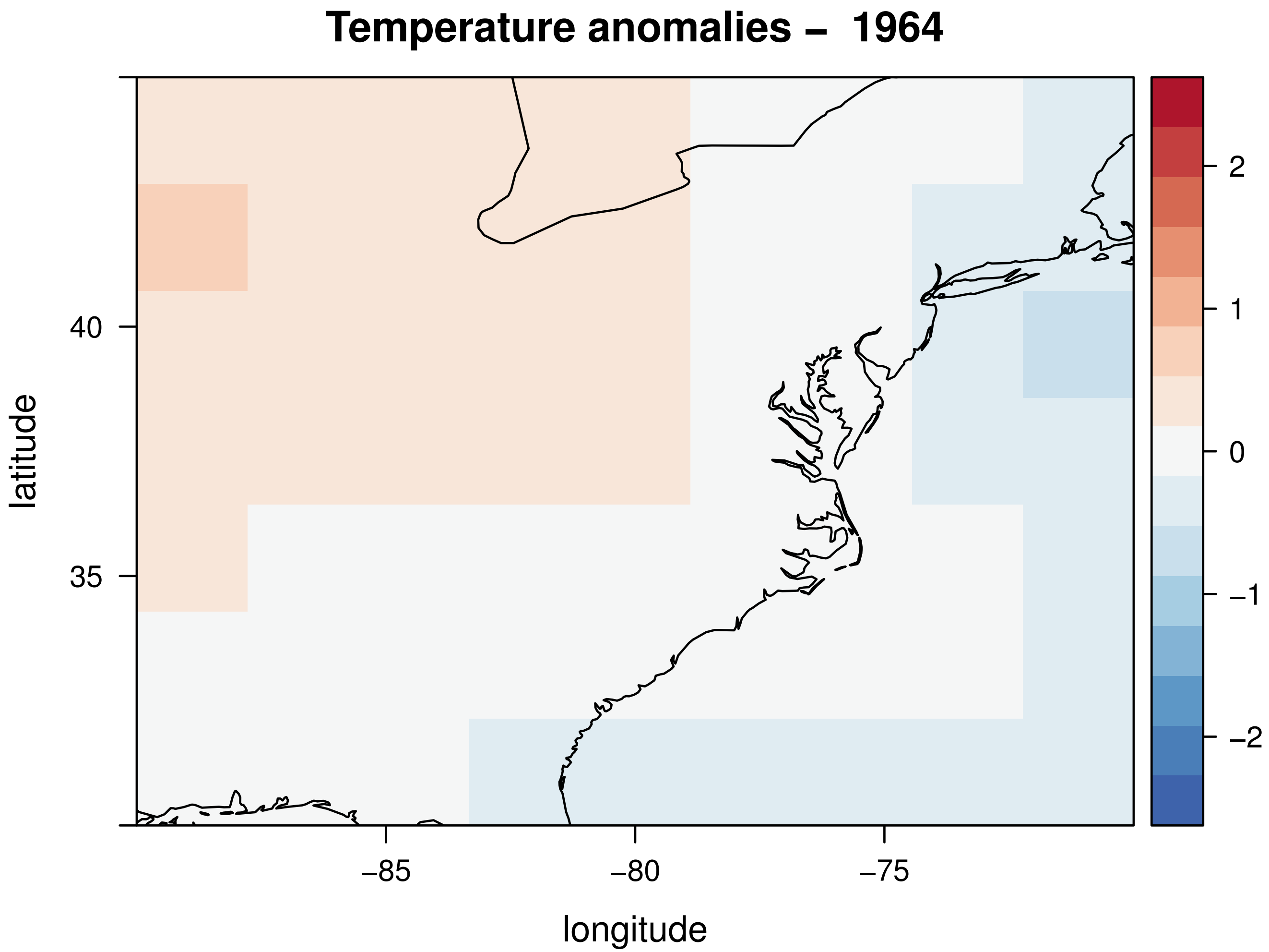
Temperature Anomalies in Eastern USA



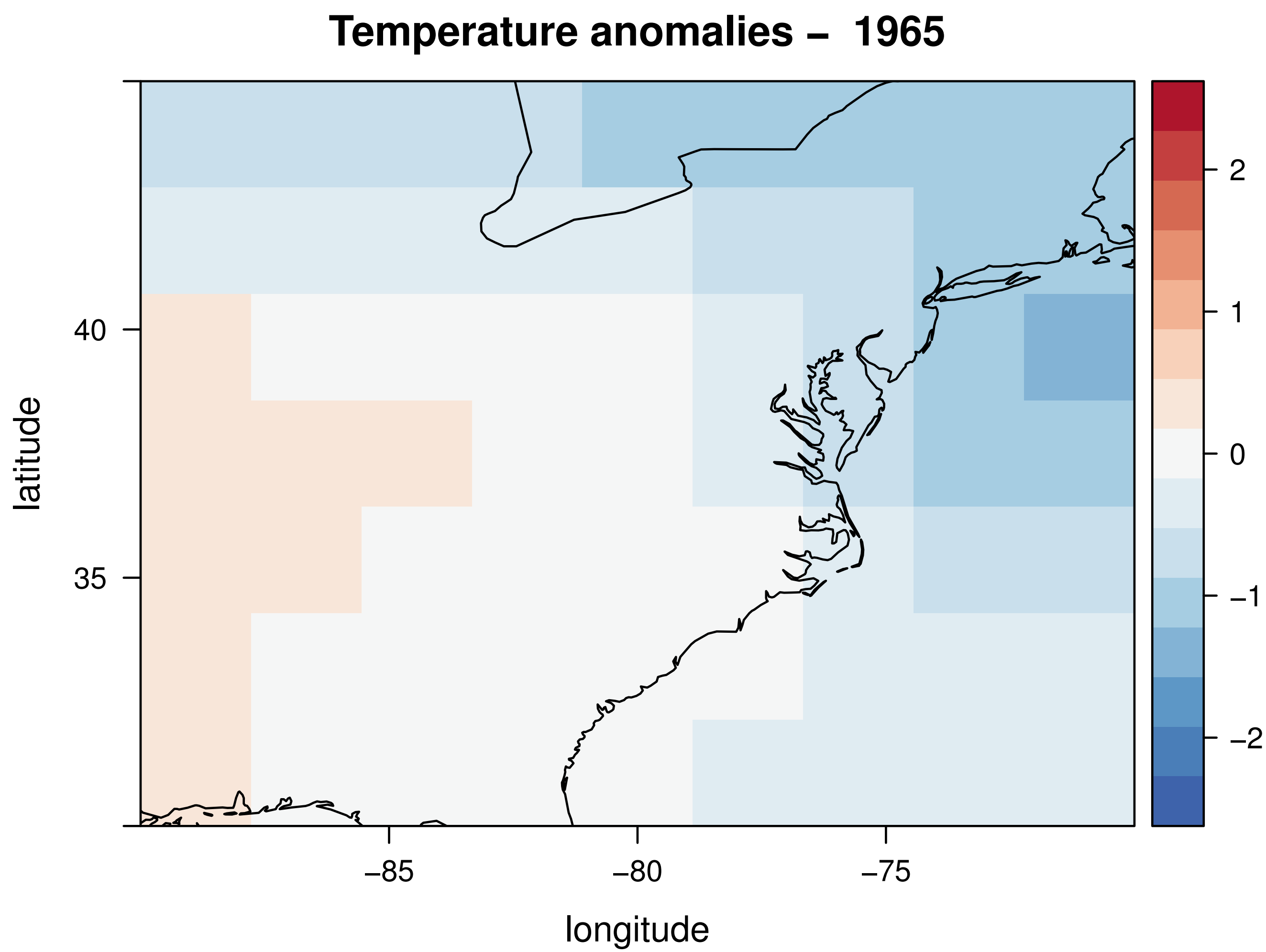
Temperature Anomalies in Eastern USA



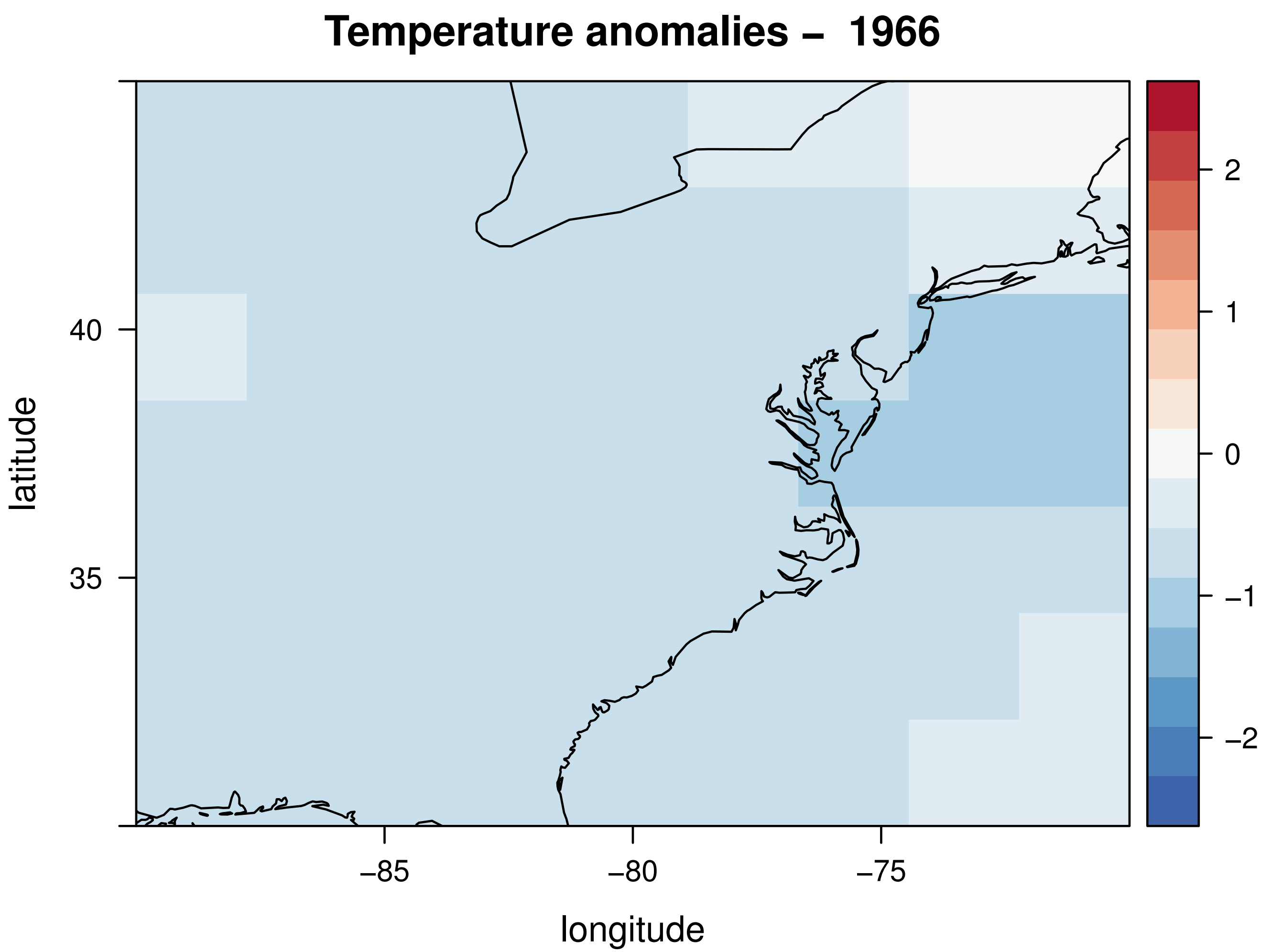
Temperature Anomalies in Eastern USA



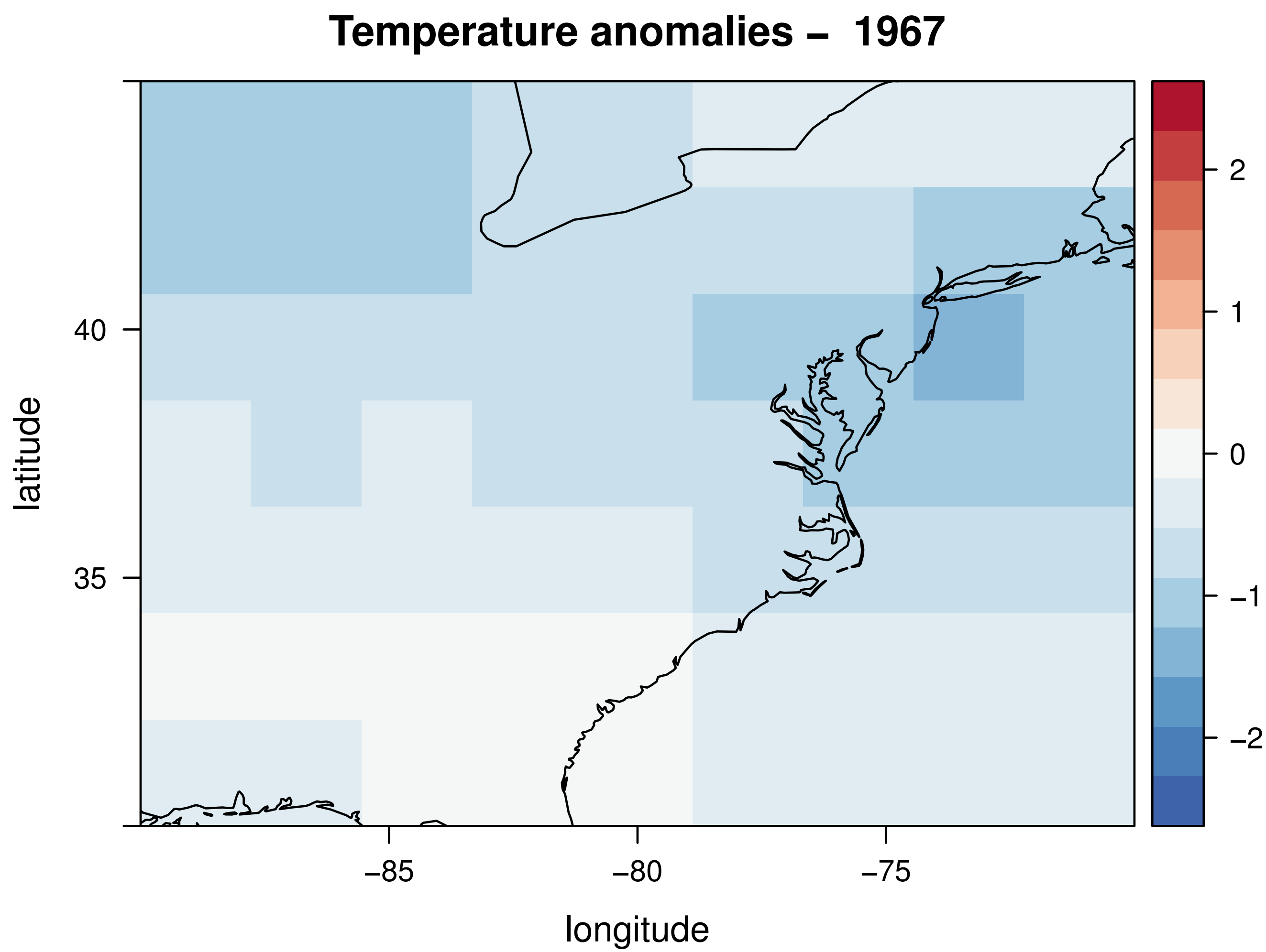
Temperature Anomalies in Eastern USA



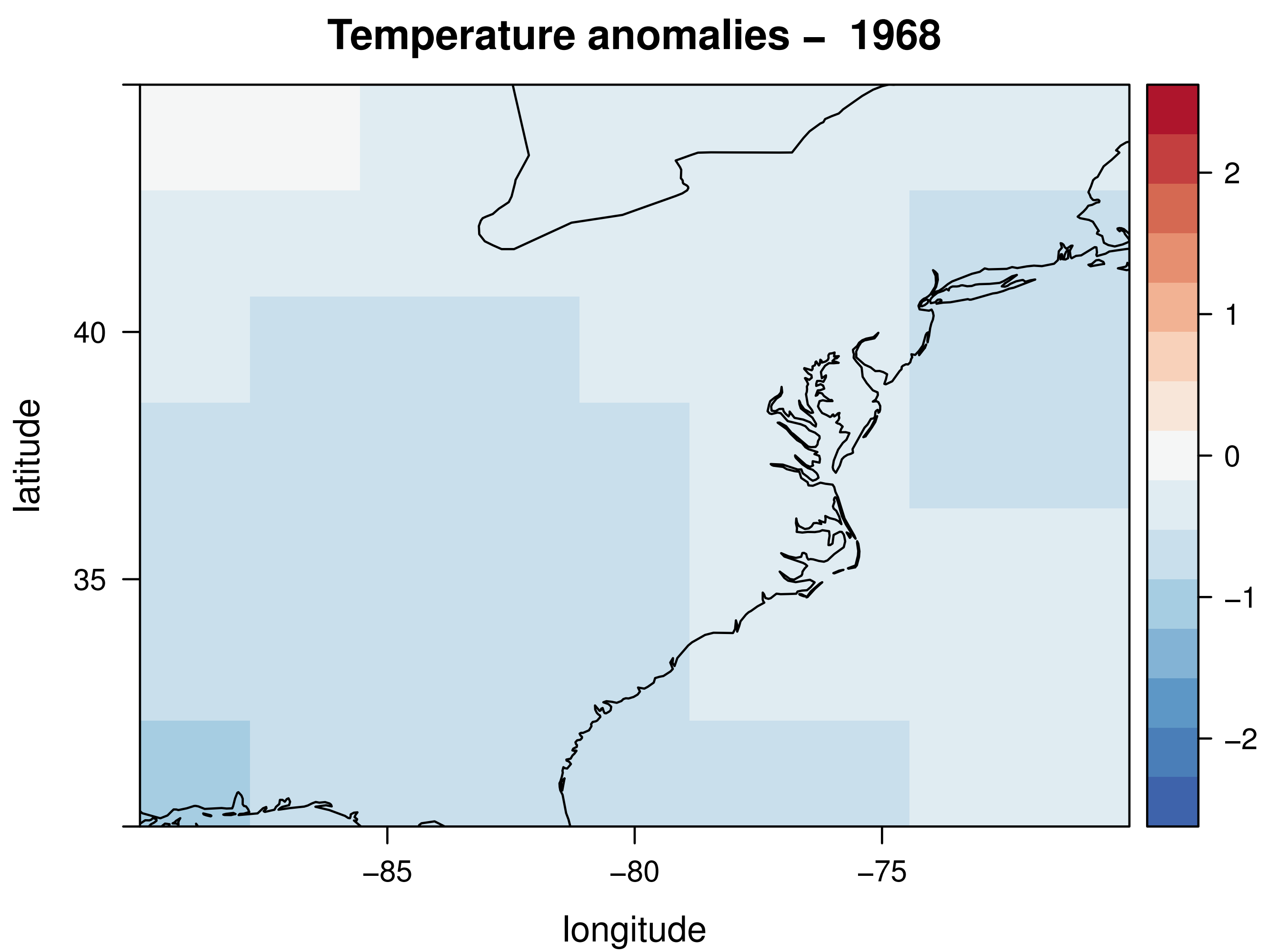
Temperature Anomalies in Eastern USA



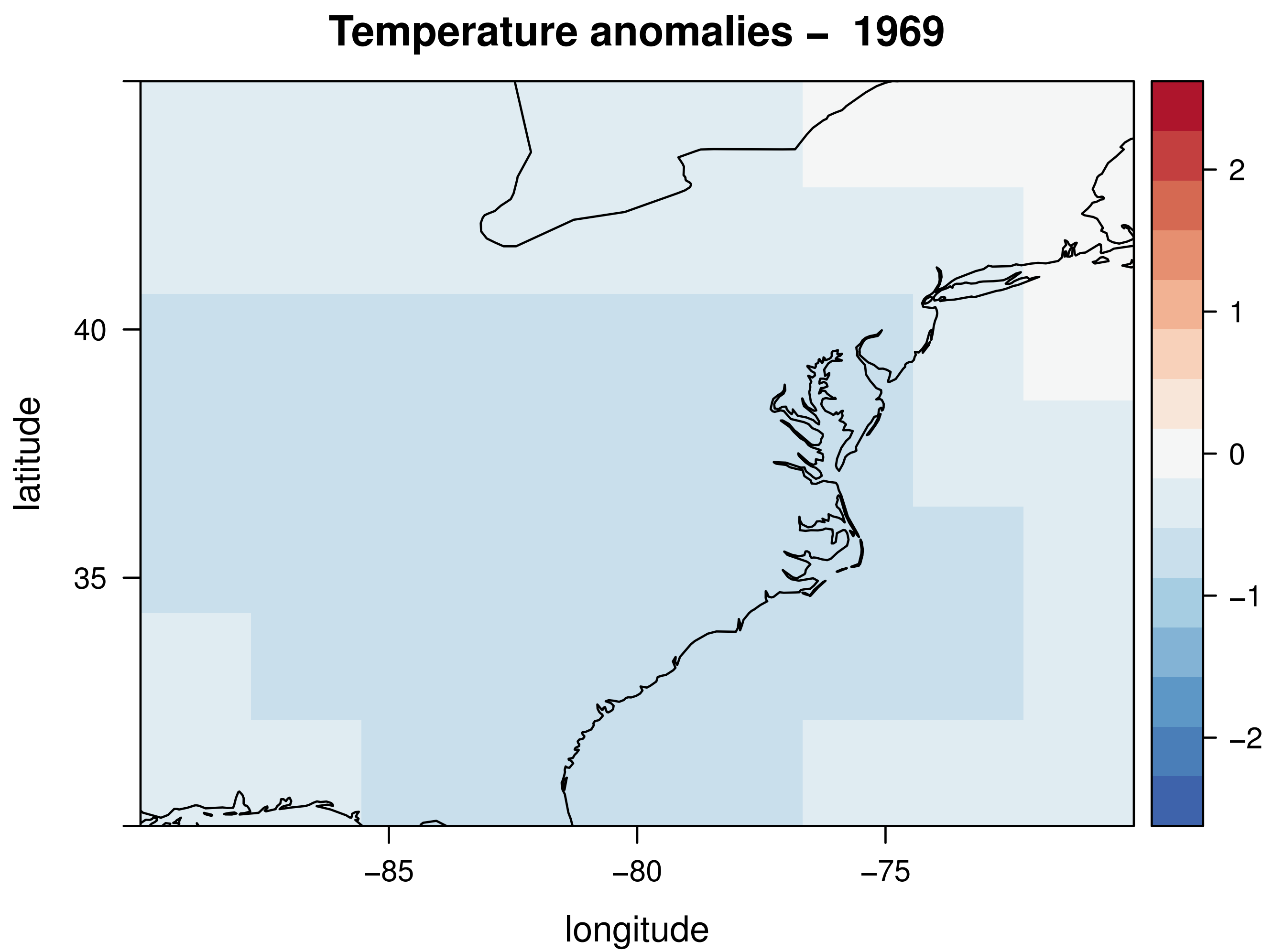
Temperature Anomalies in Eastern USA



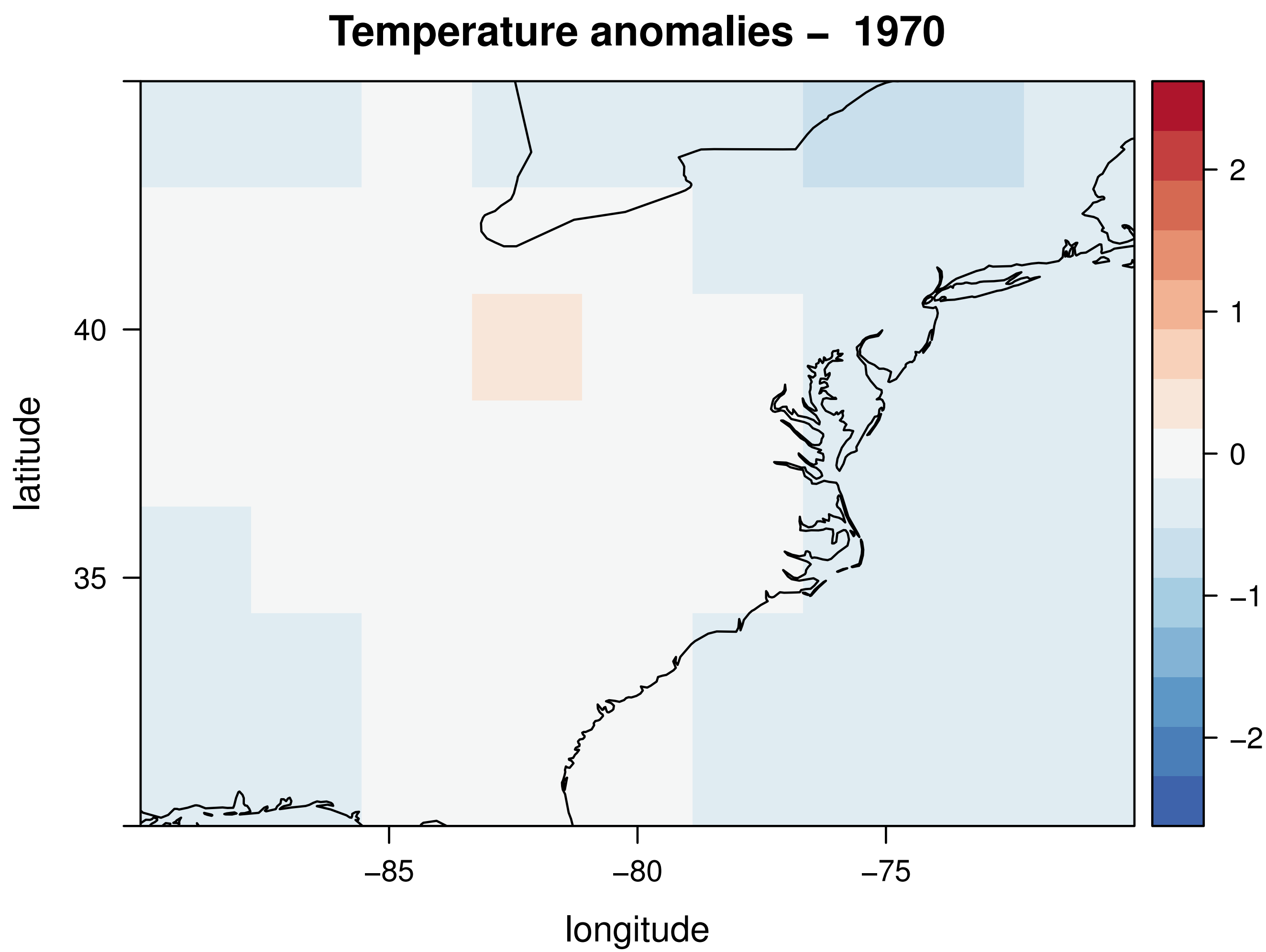
Temperature Anomalies in Eastern USA



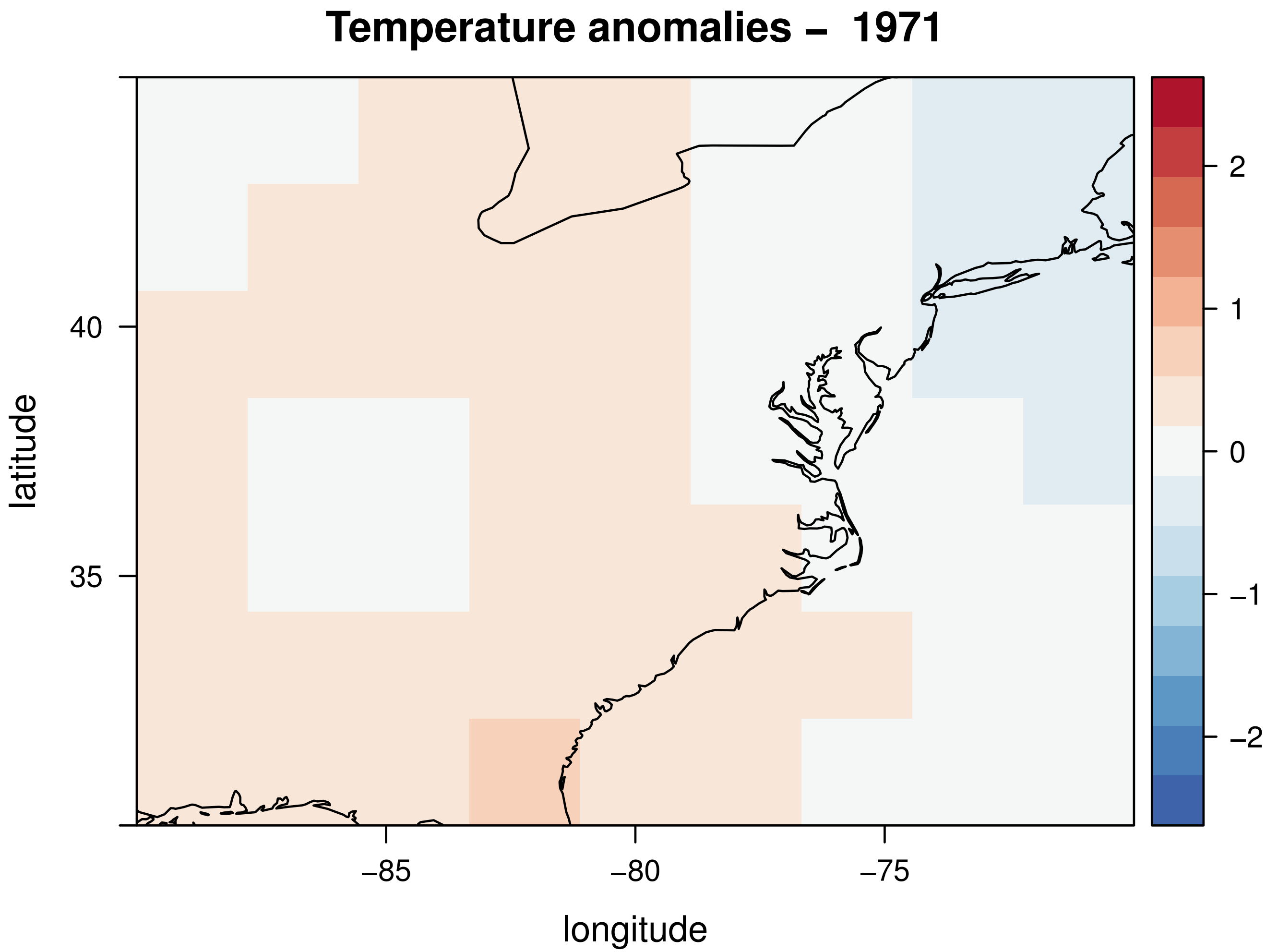
Temperature Anomalies in Eastern USA



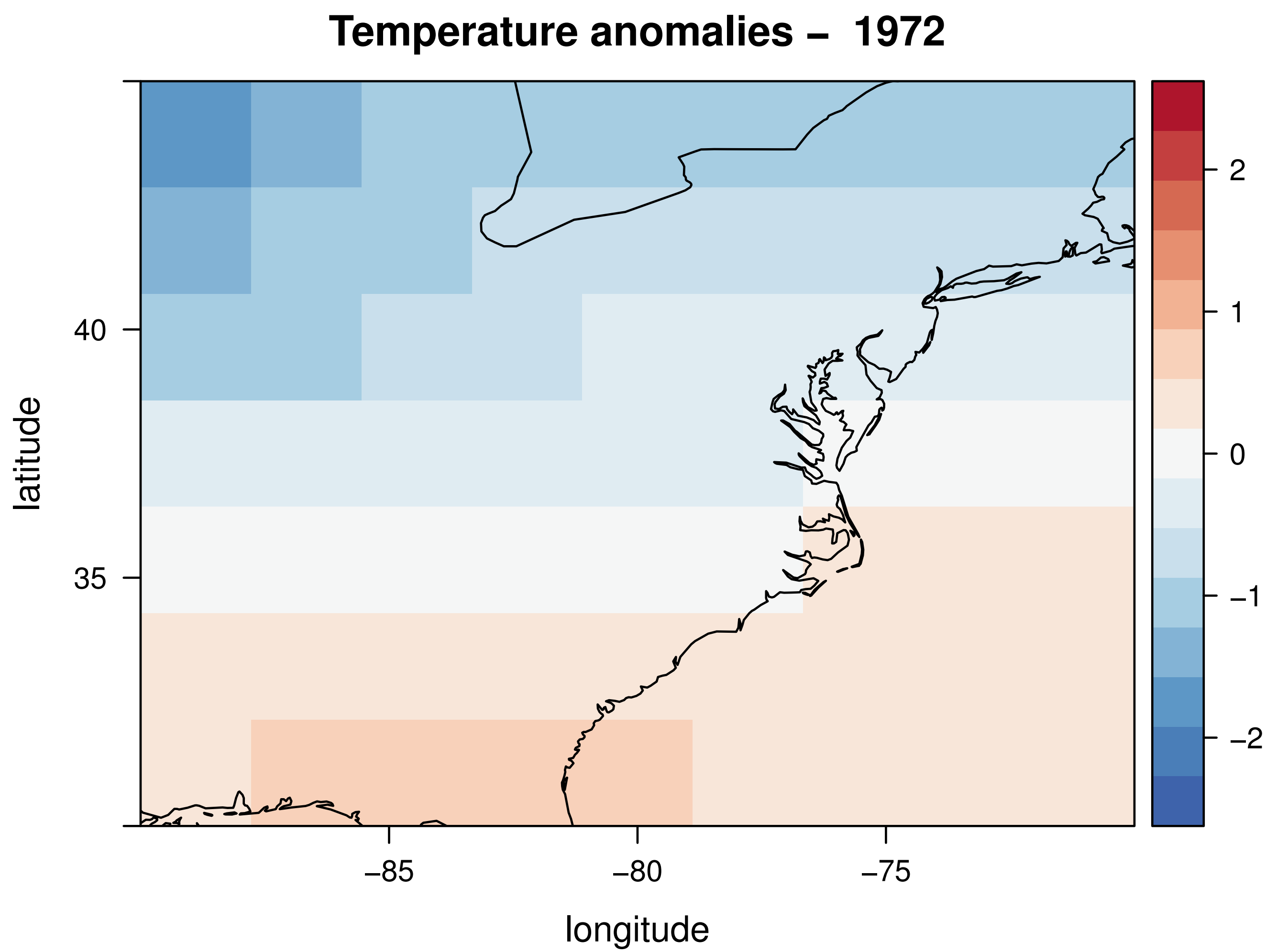
Temperature Anomalies in Eastern USA



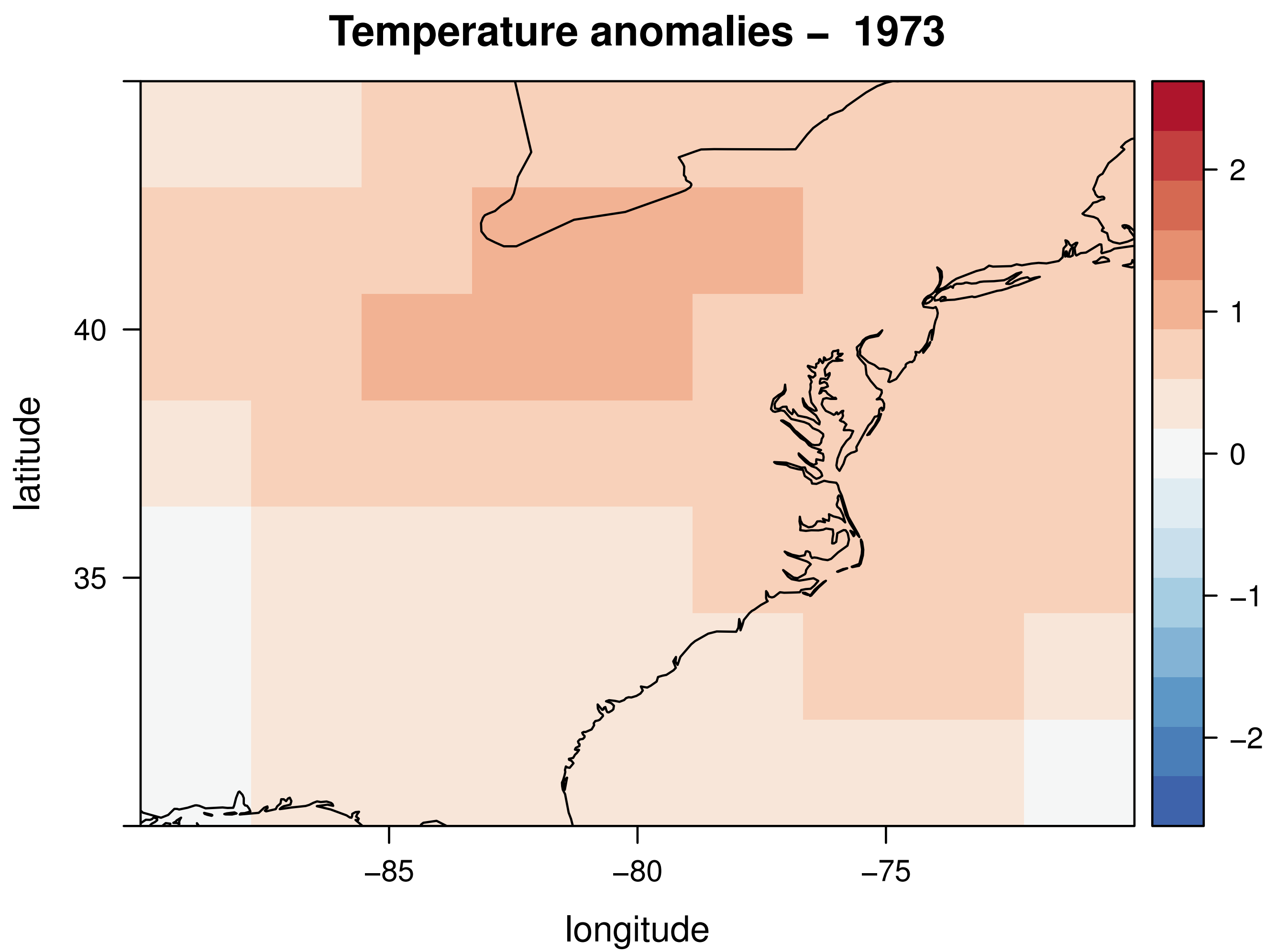
Temperature Anomalies in Eastern USA



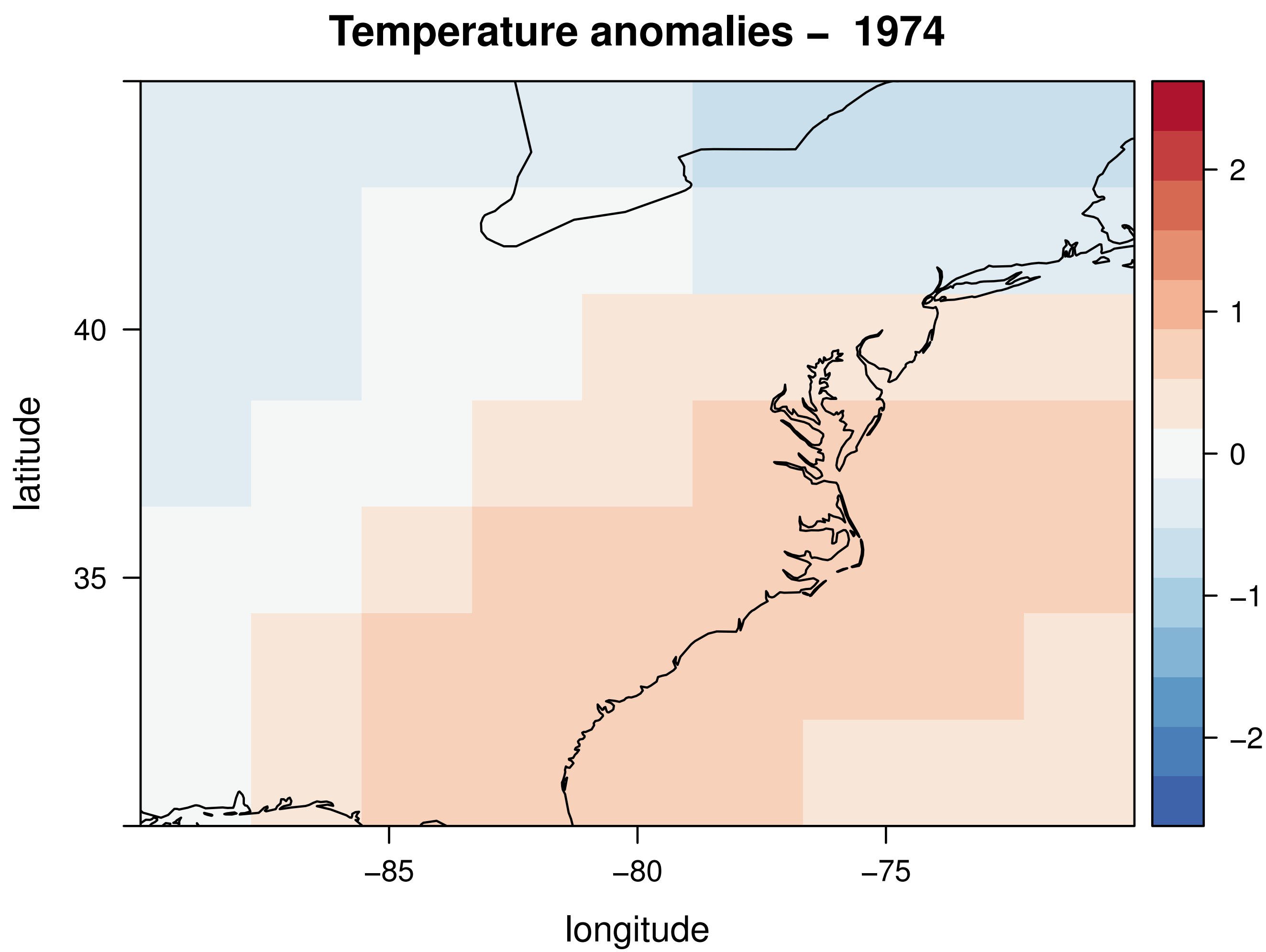
Temperature Anomalies in Eastern USA



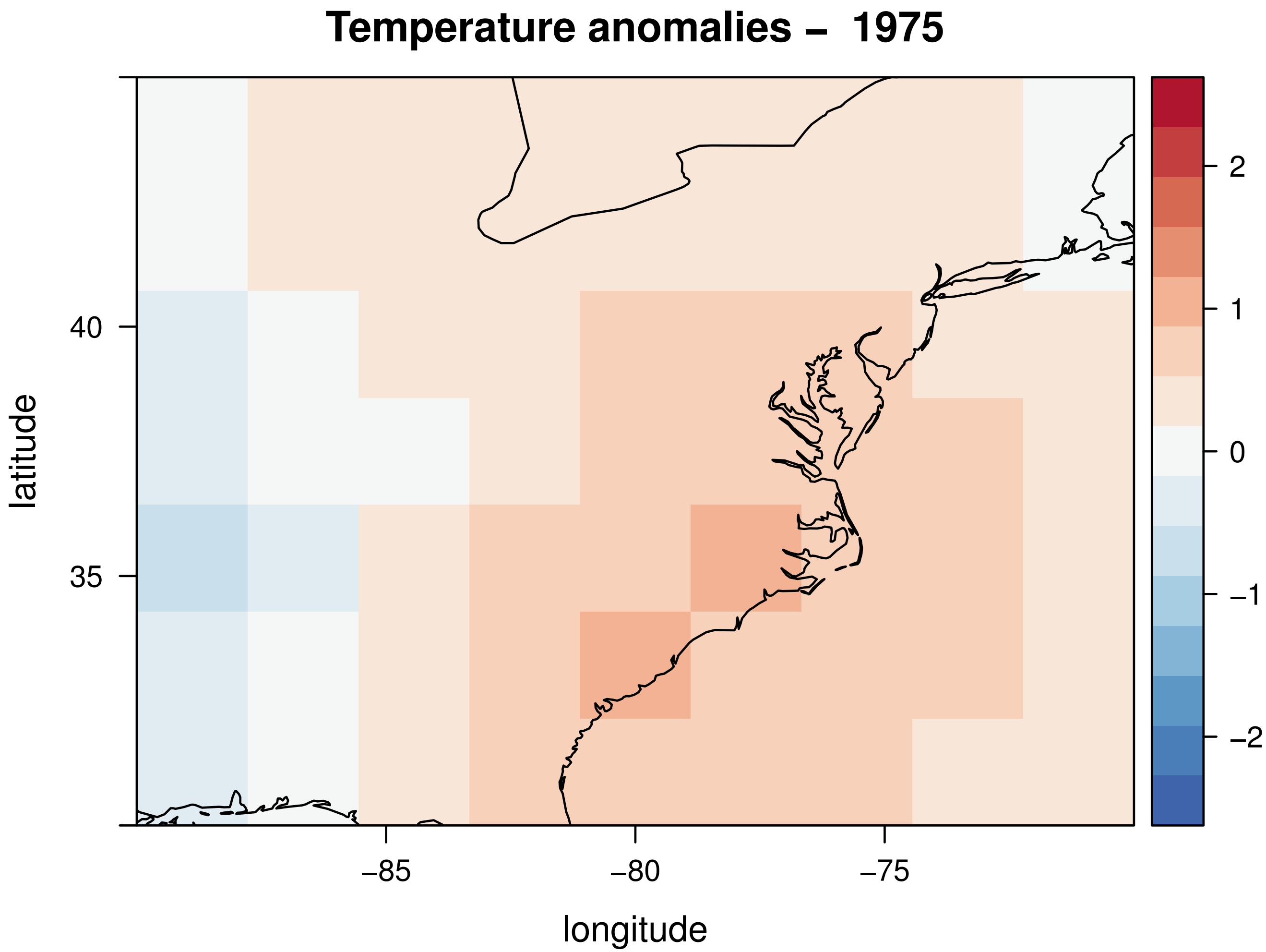
Temperature Anomalies in Eastern USA



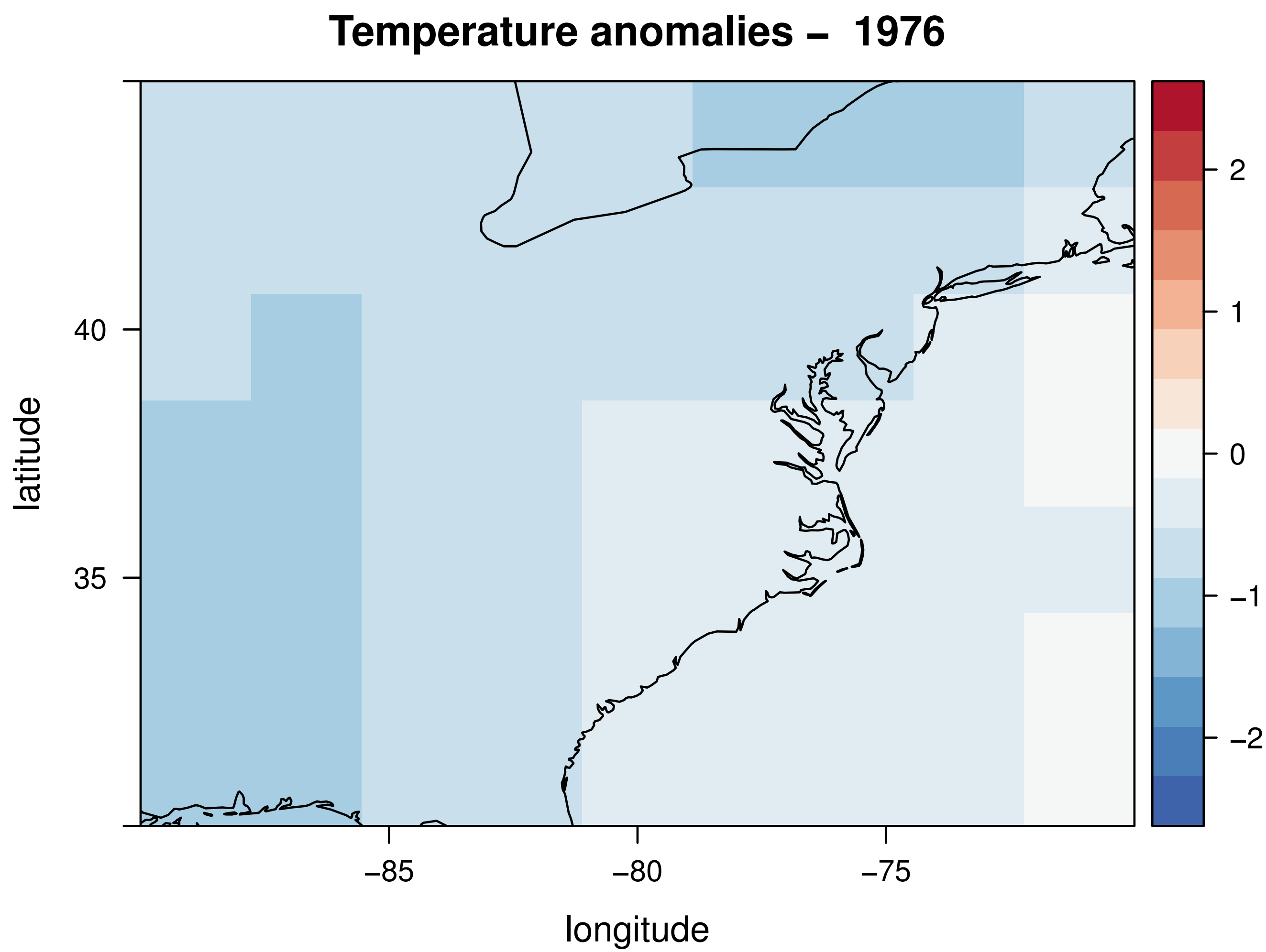
Temperature Anomalies in Eastern USA



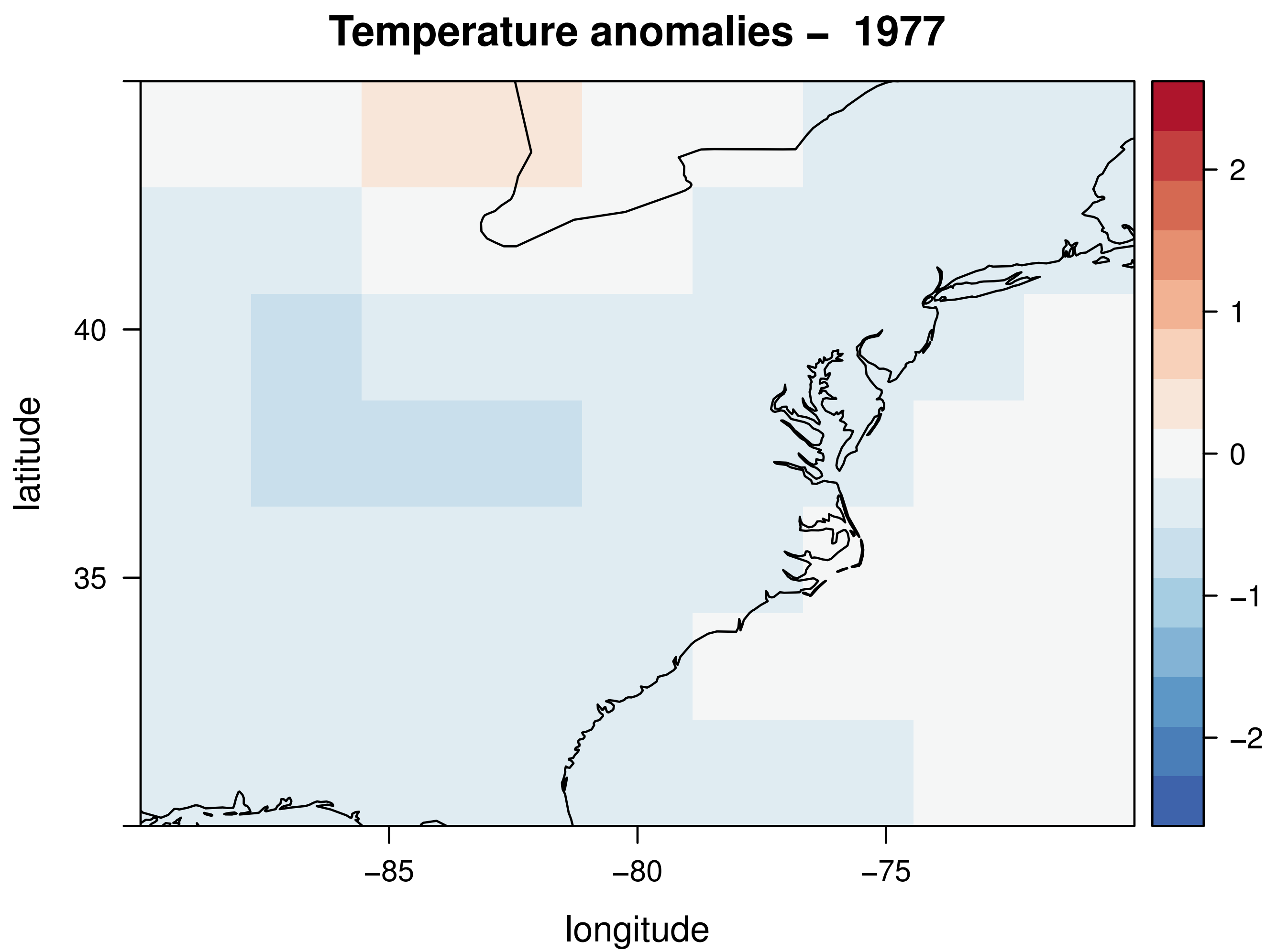
Temperature Anomalies in Eastern USA



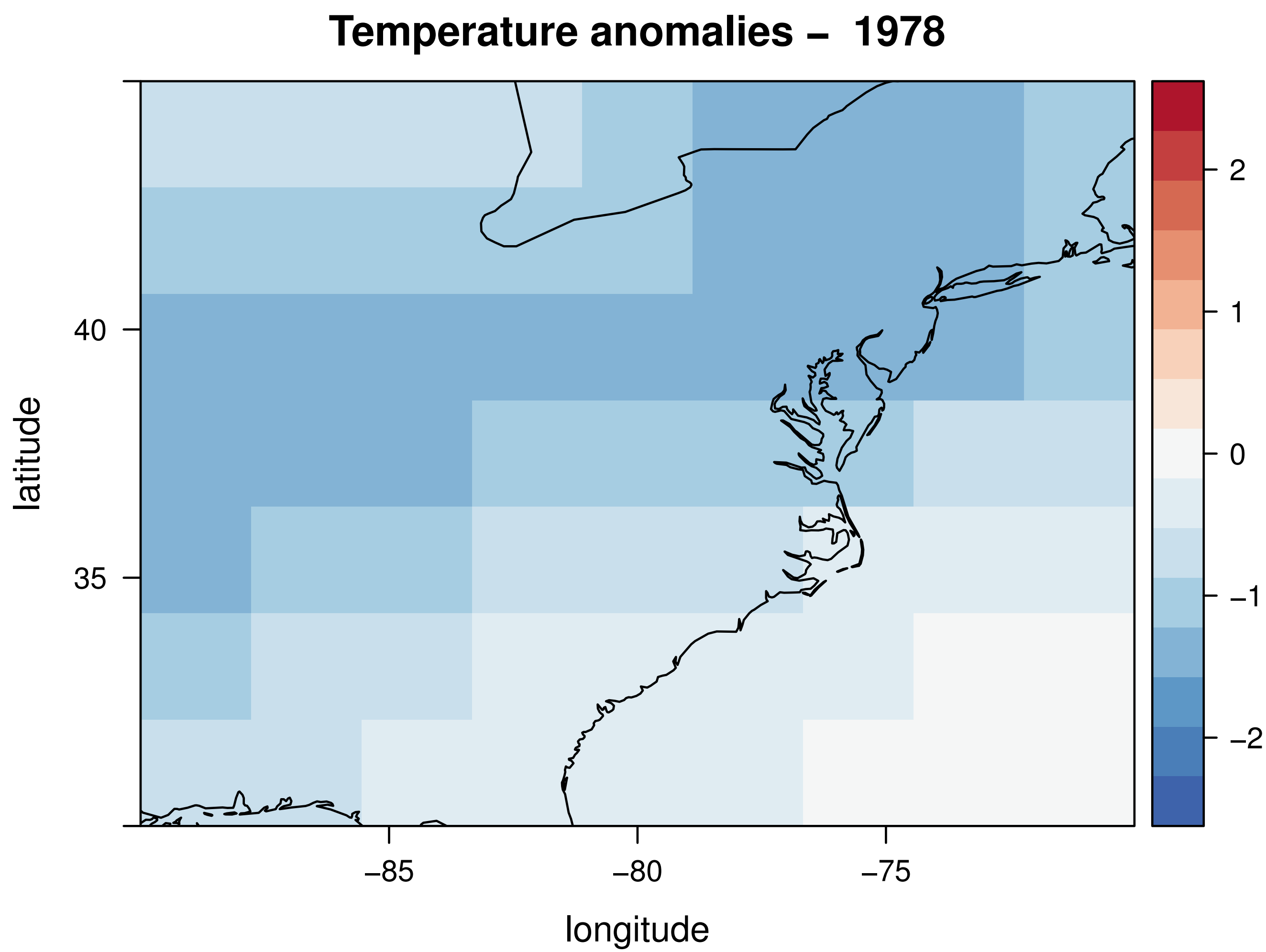
Temperature Anomalies in Eastern USA



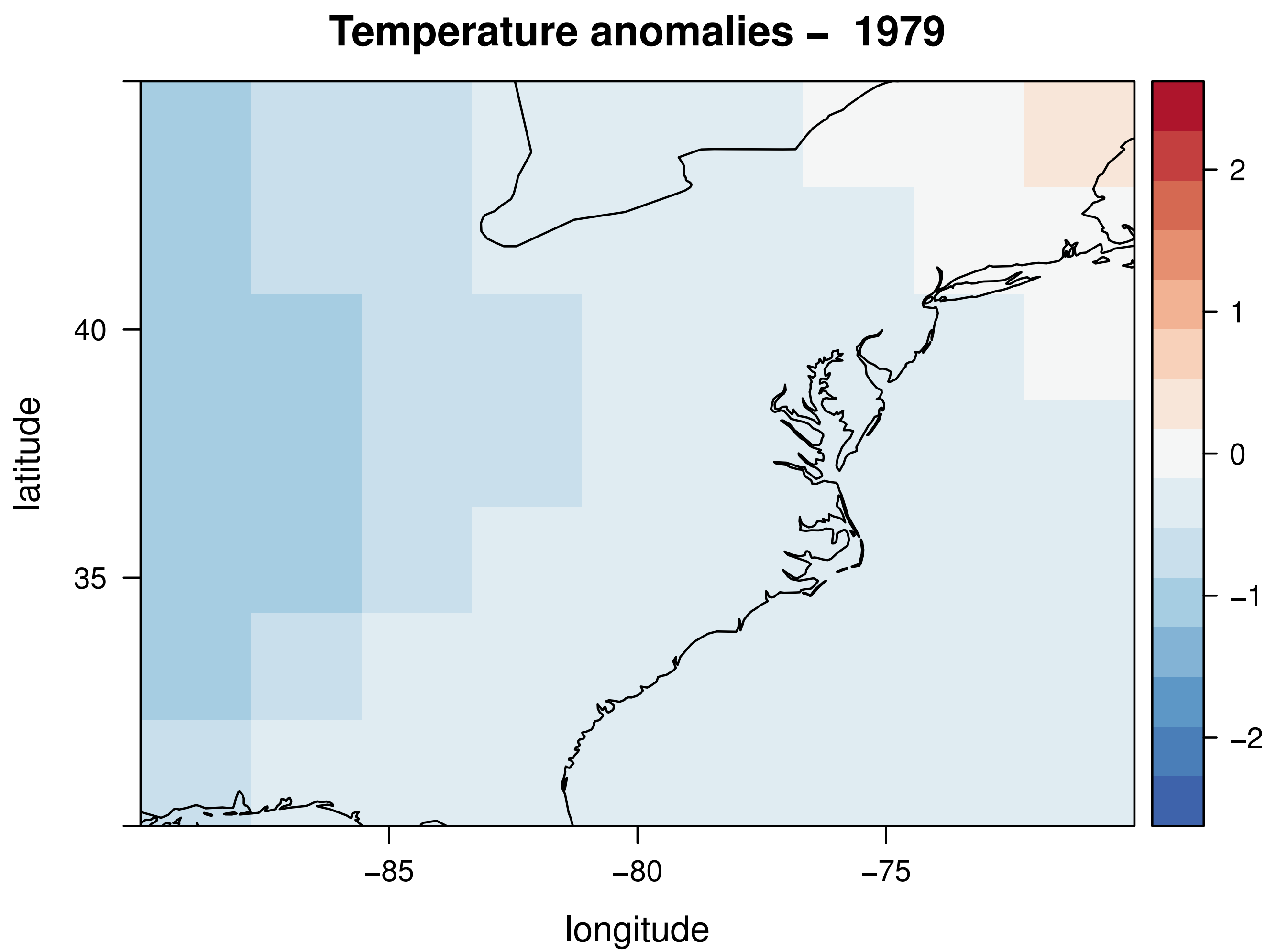
Temperature Anomalies in Eastern USA



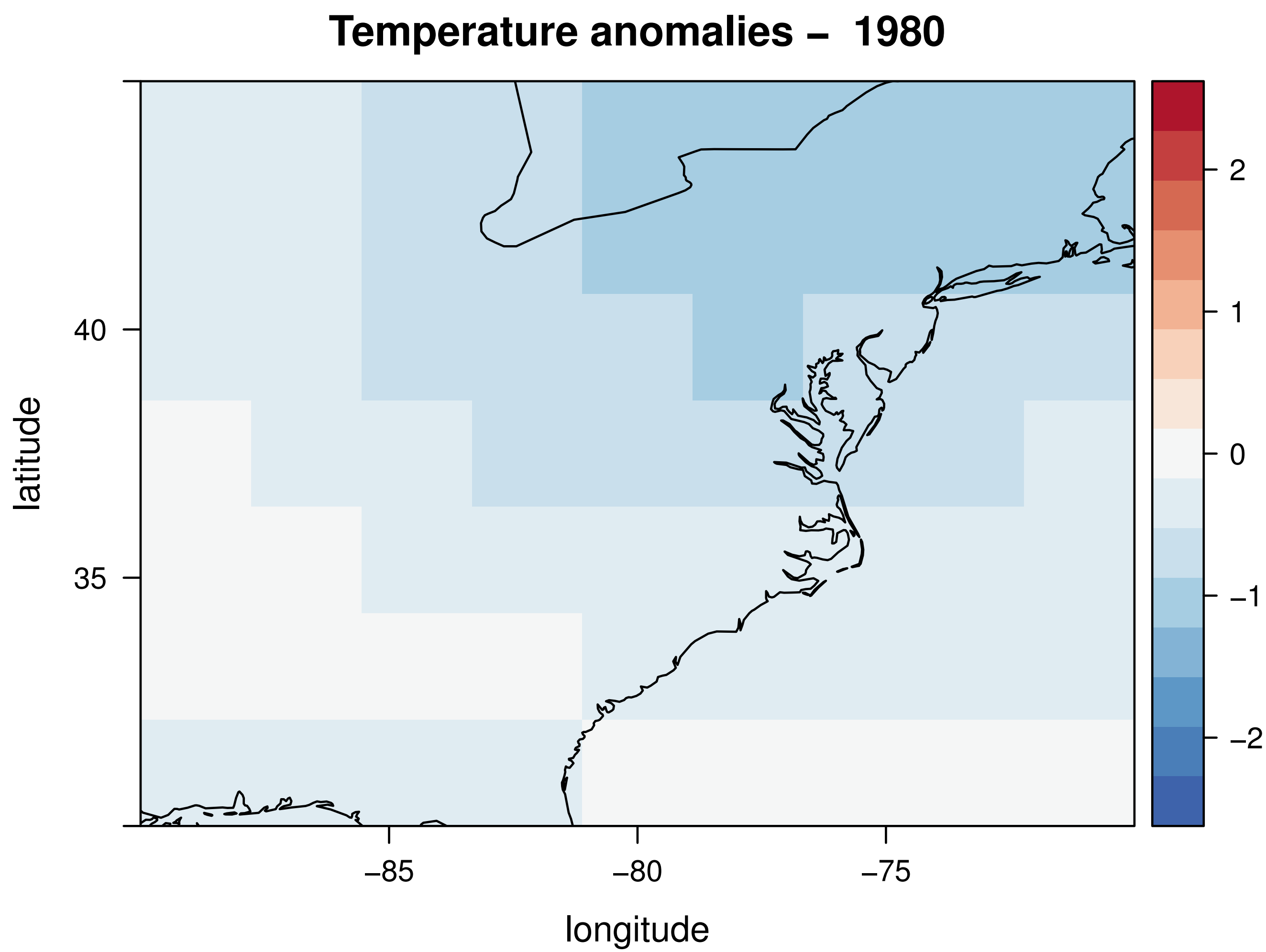
Temperature Anomalies in Eastern USA



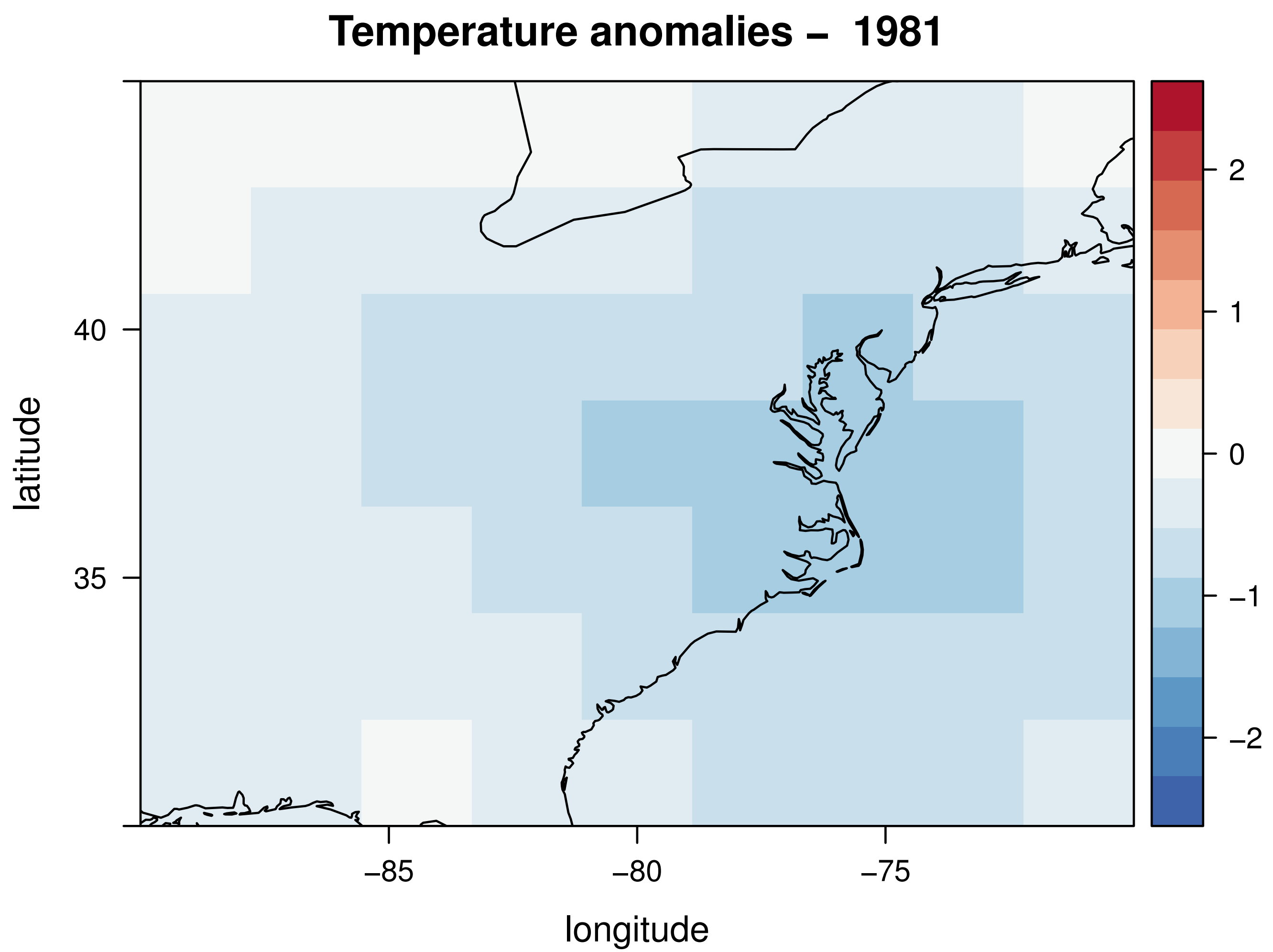
Temperature Anomalies in Eastern USA



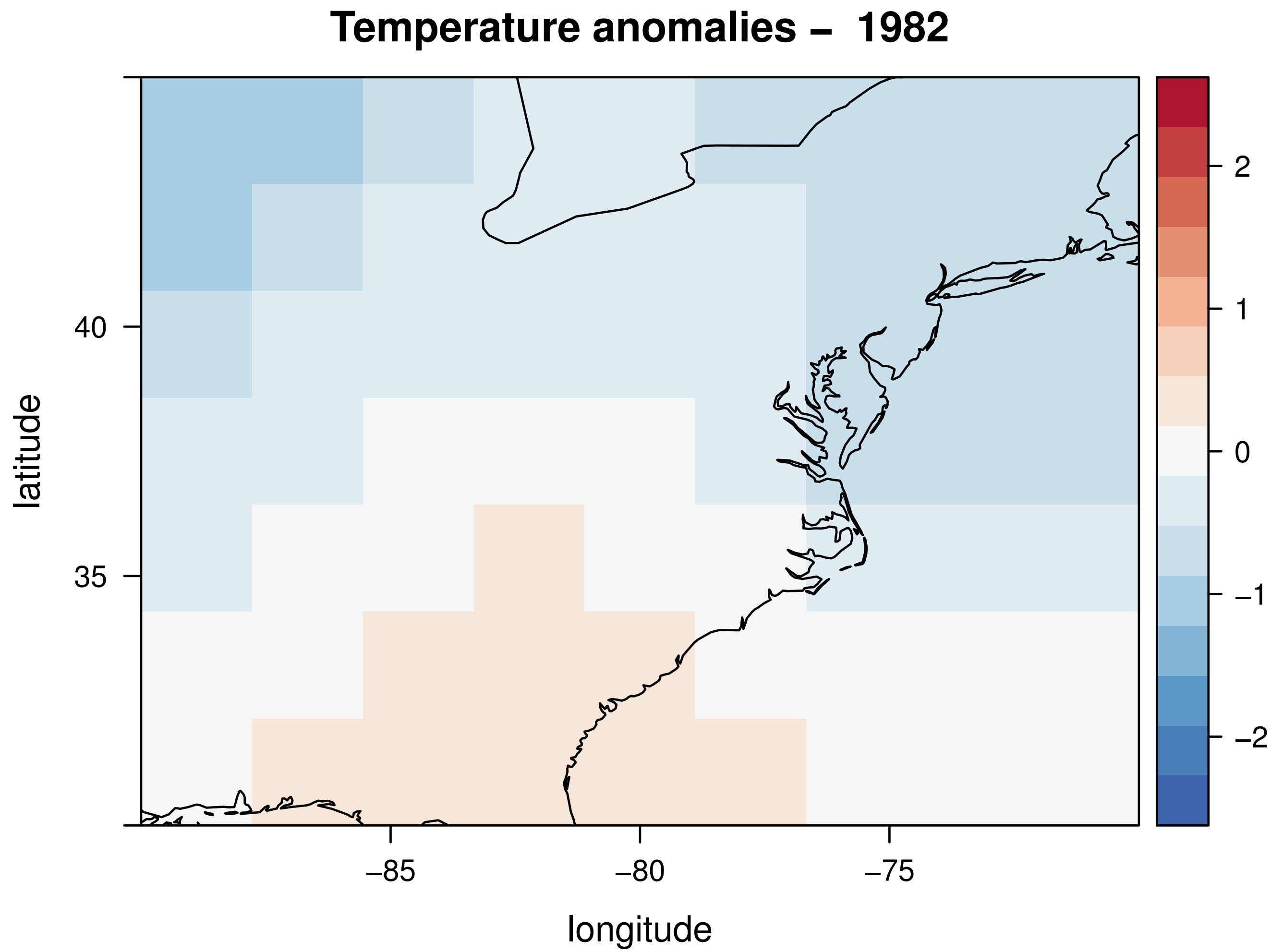
Temperature Anomalies in Eastern USA



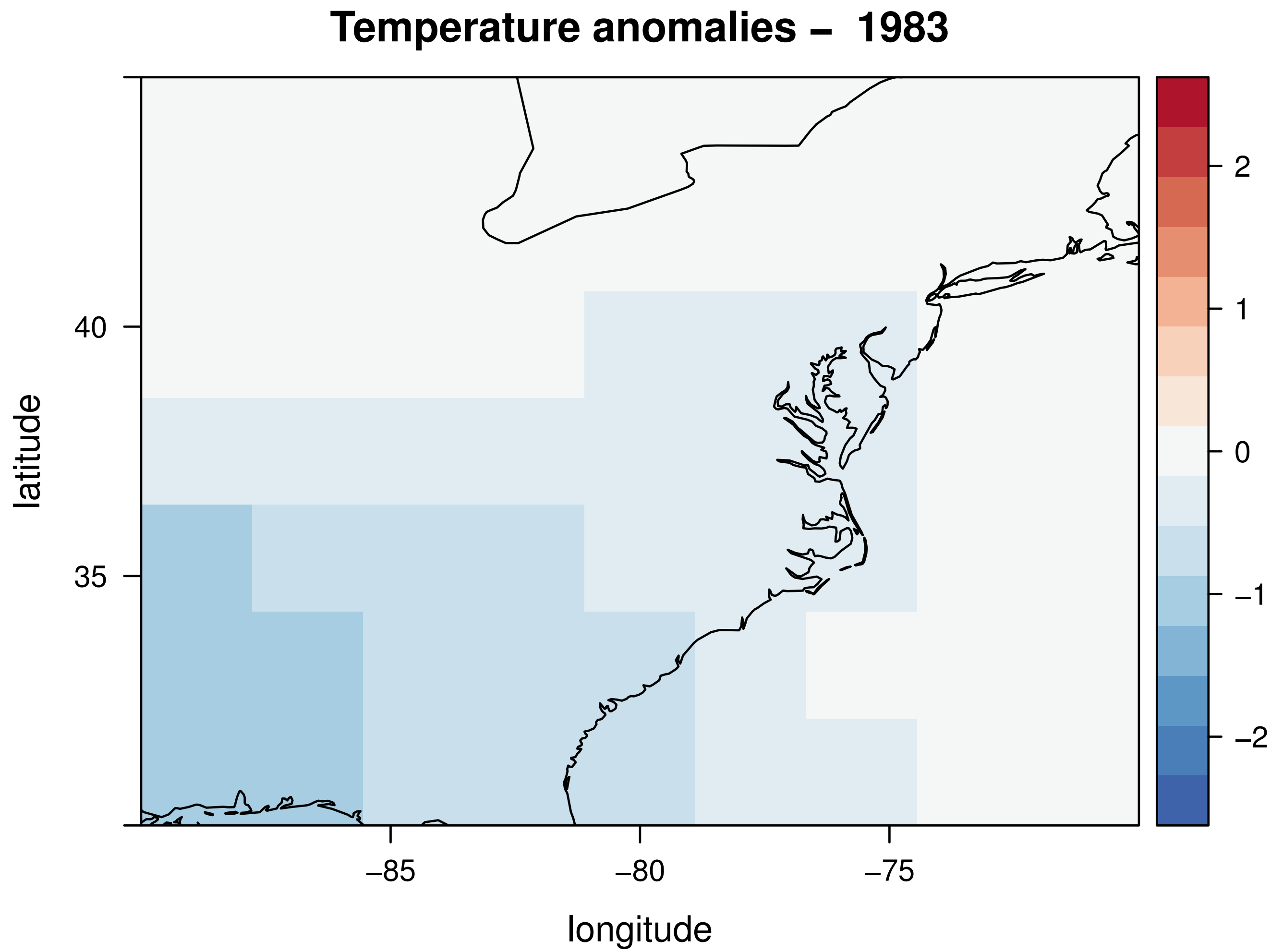
Temperature Anomalies in Eastern USA



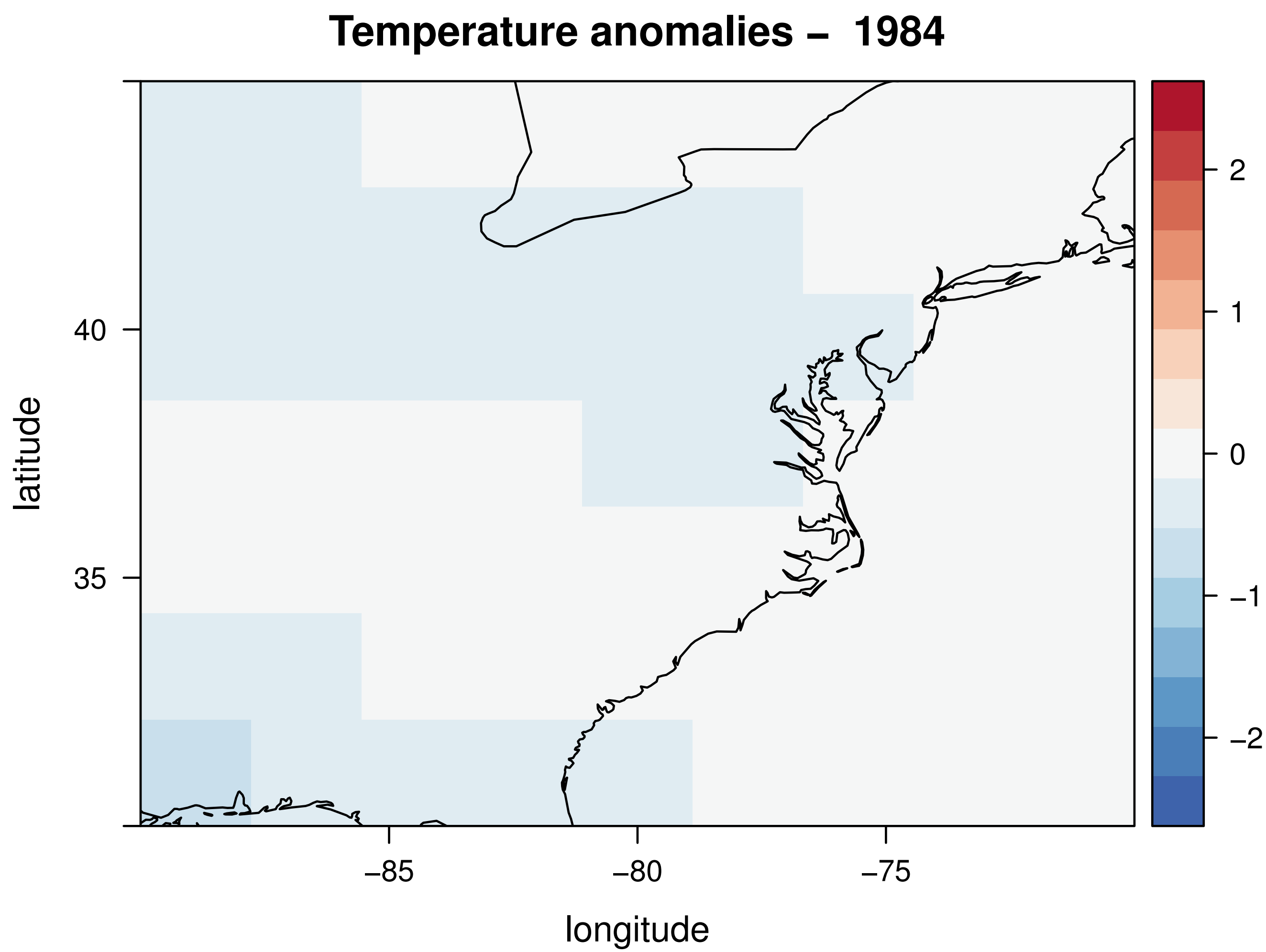
Temperature Anomalies in Eastern USA



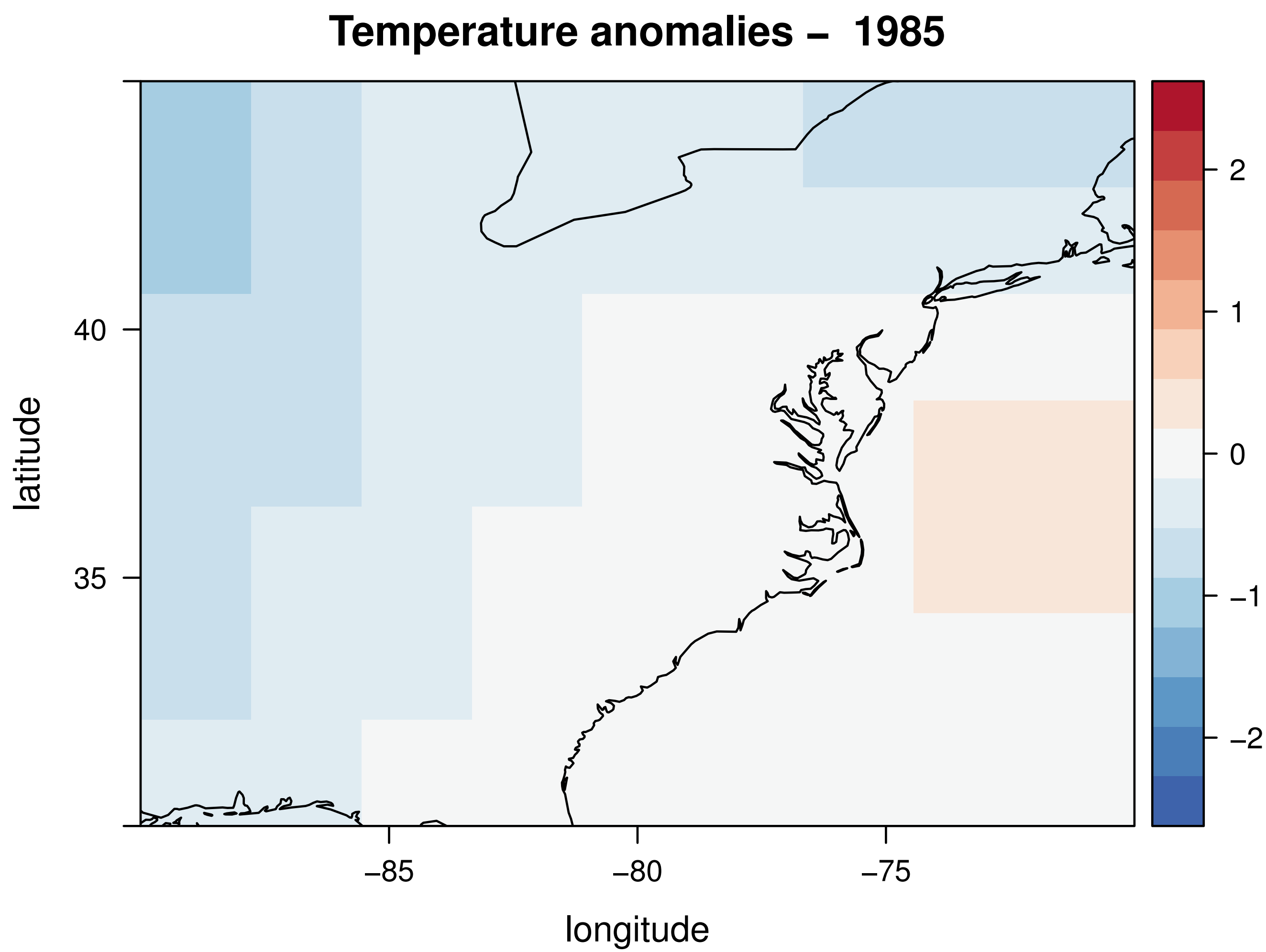
Temperature Anomalies in Eastern USA



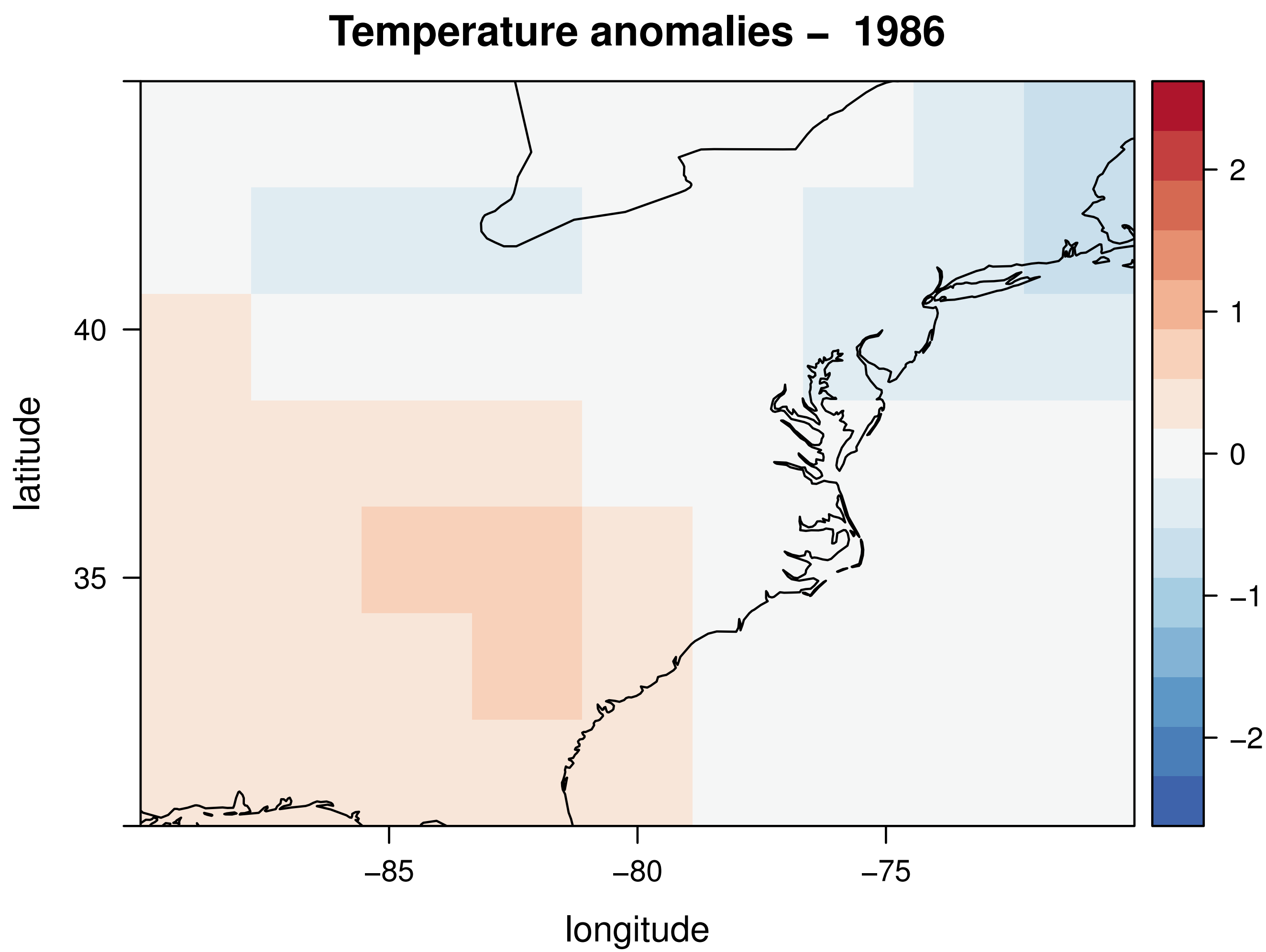
Temperature Anomalies in Eastern USA



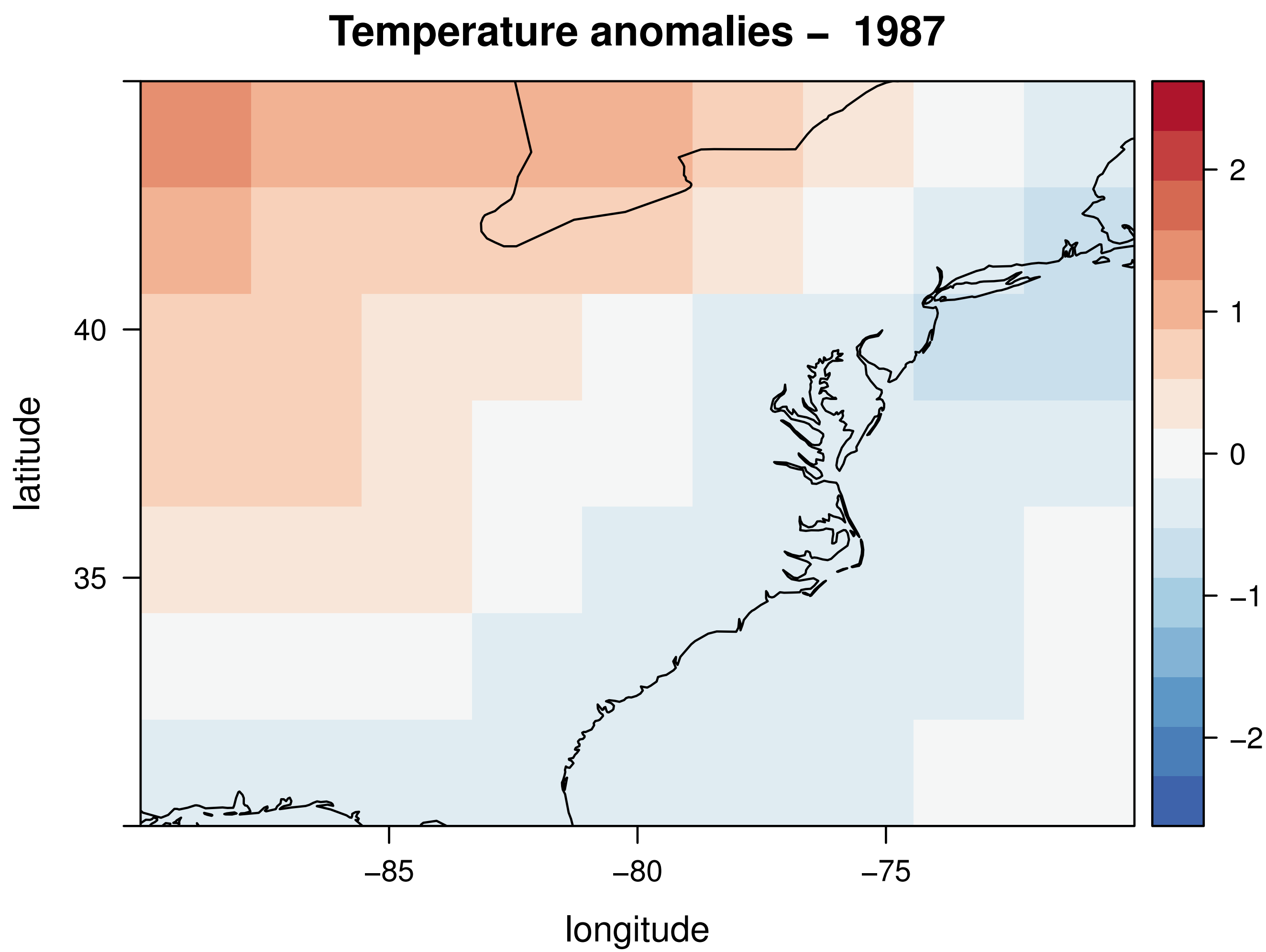
Temperature Anomalies in Eastern USA



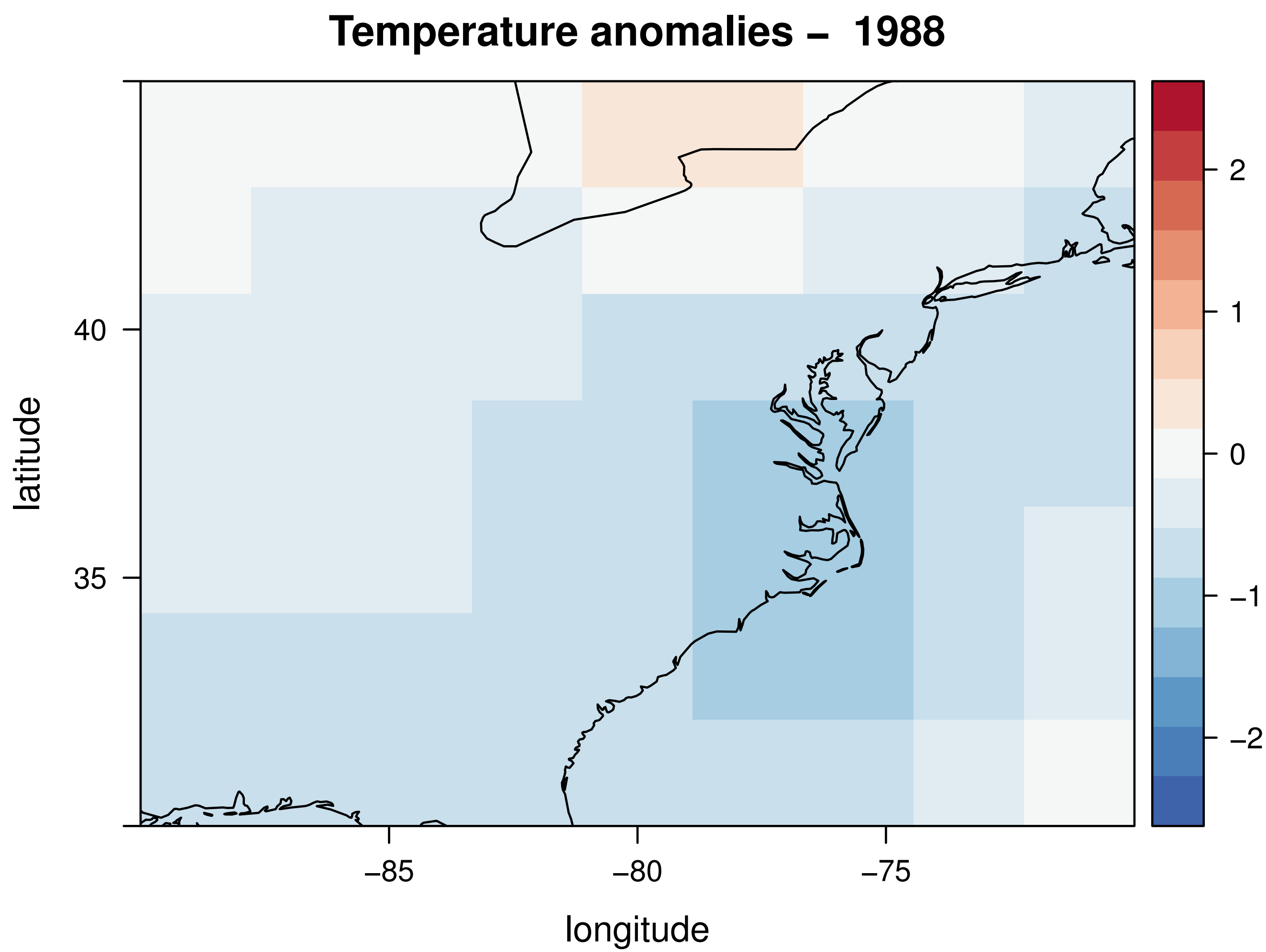
Temperature Anomalies in Eastern USA



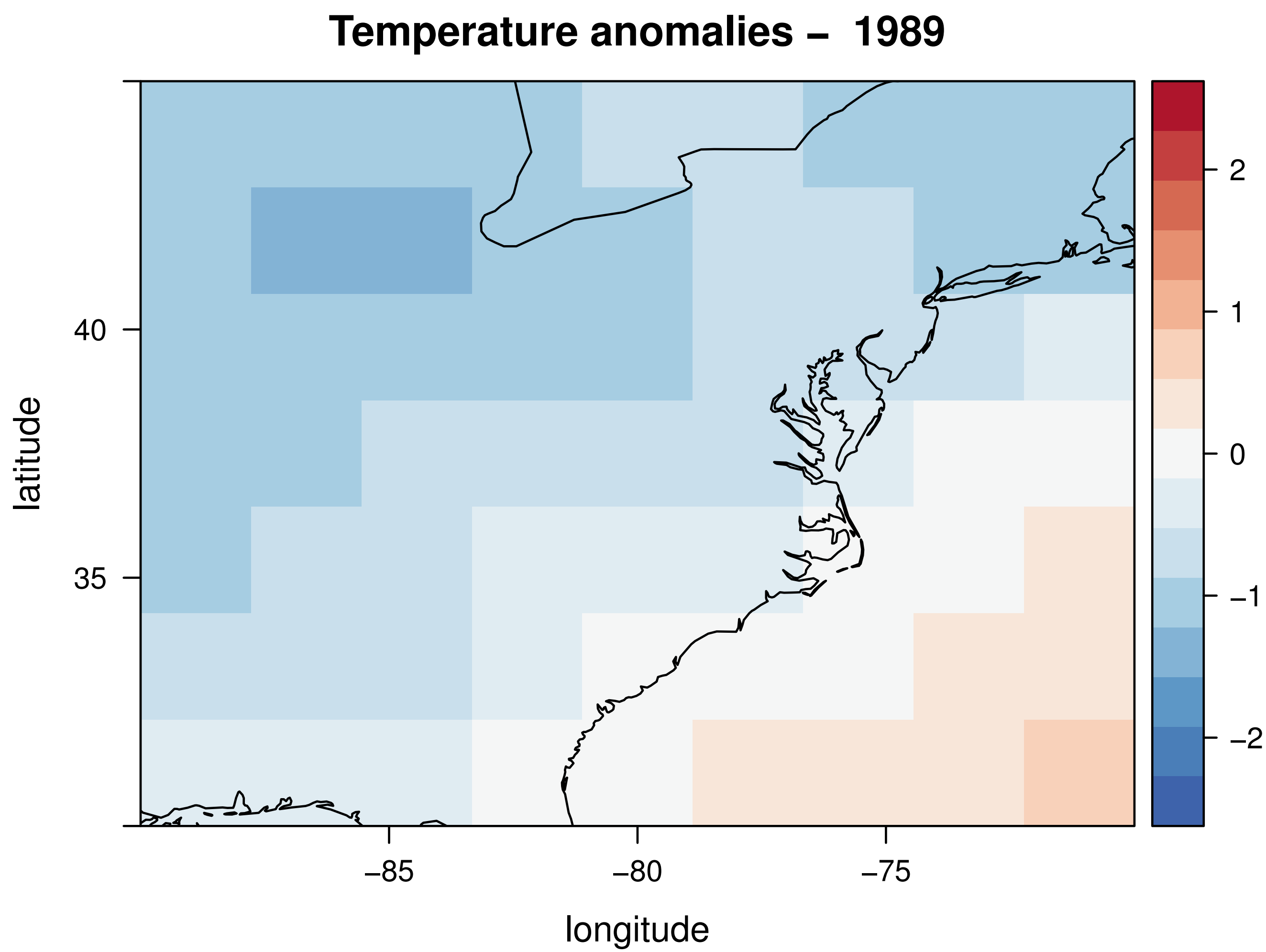
Temperature Anomalies in Eastern USA



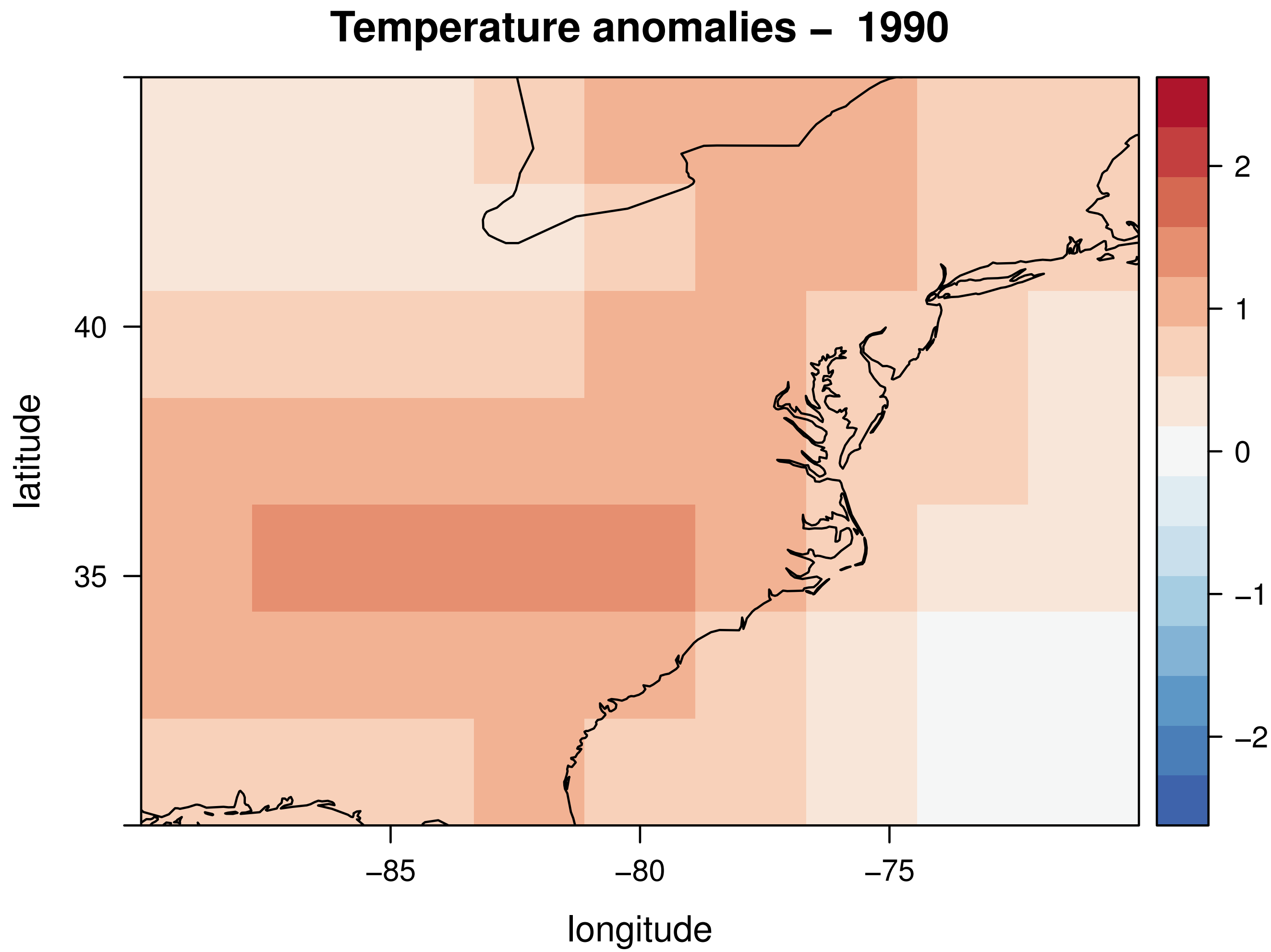
Temperature Anomalies in Eastern USA



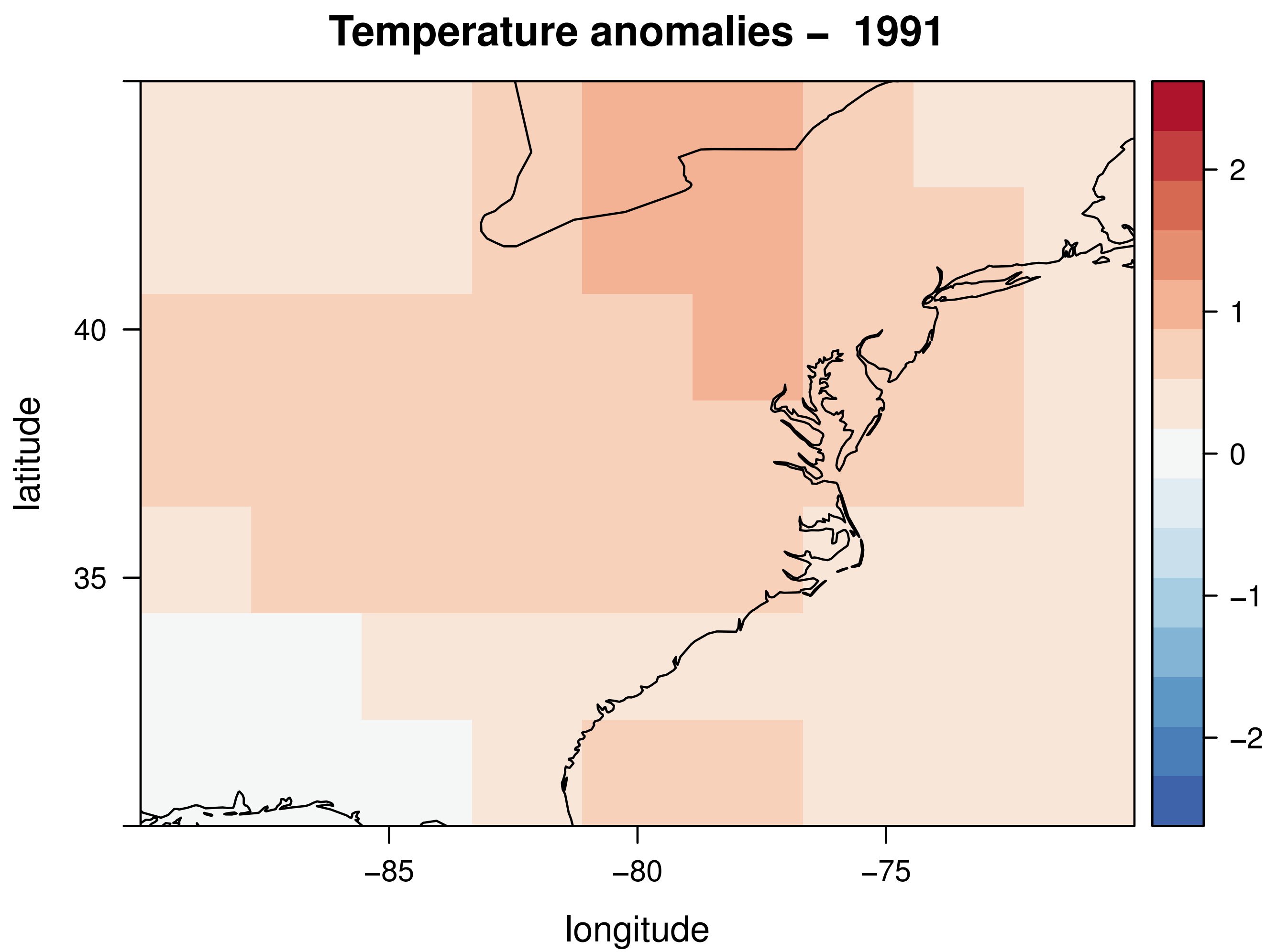
Temperature Anomalies in Eastern USA



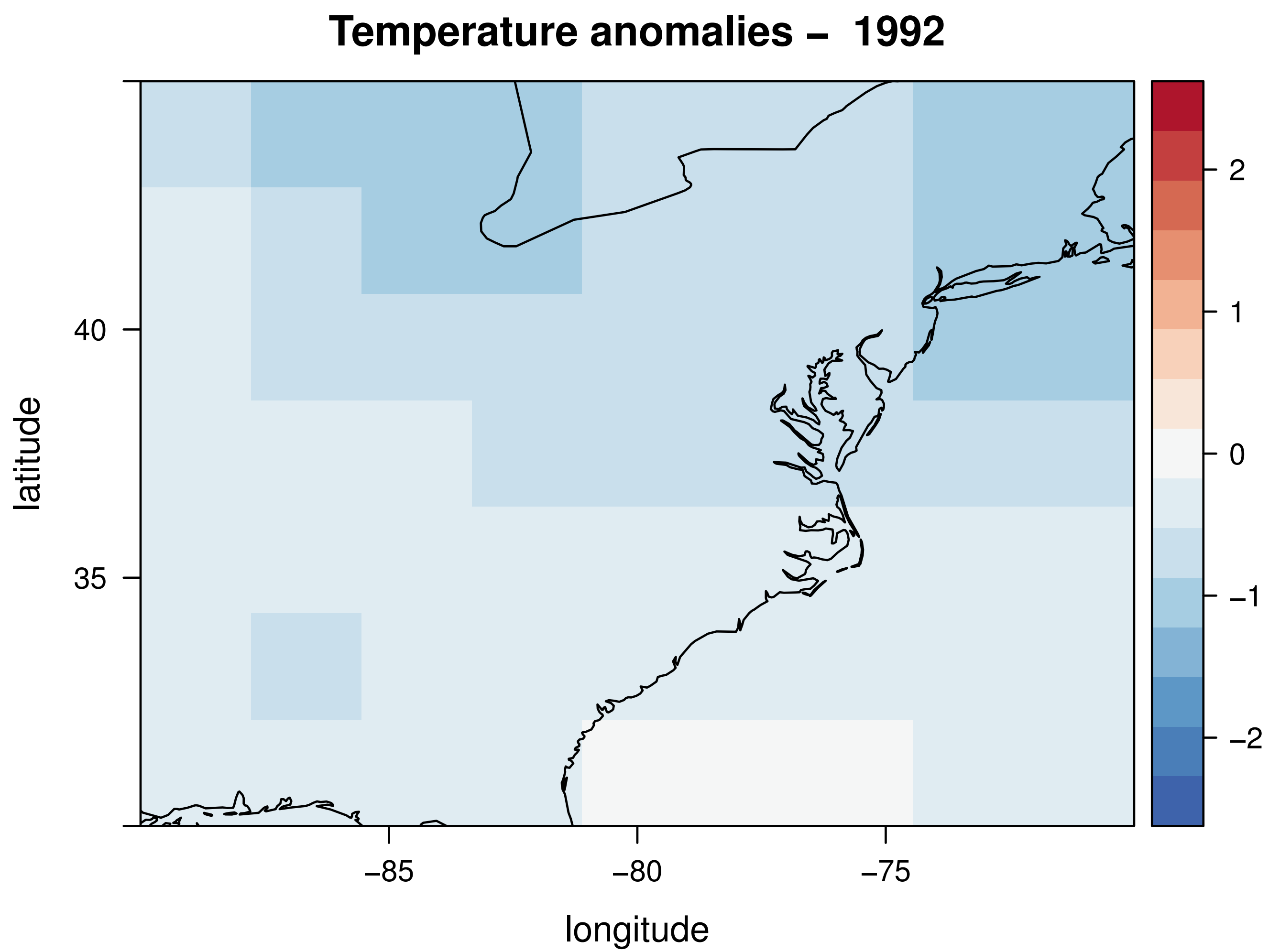
Temperature Anomalies in Eastern USA



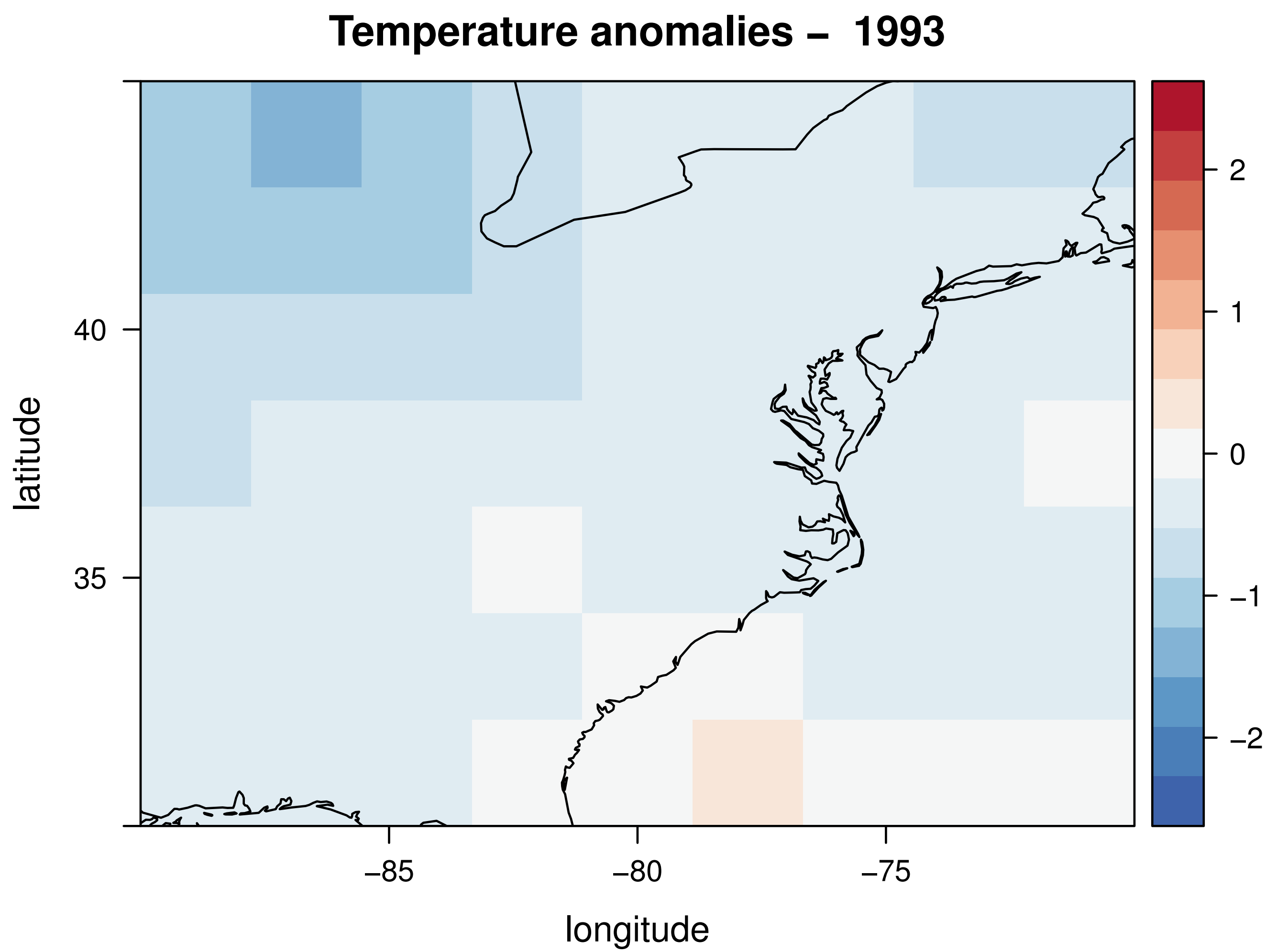
Temperature Anomalies in Eastern USA



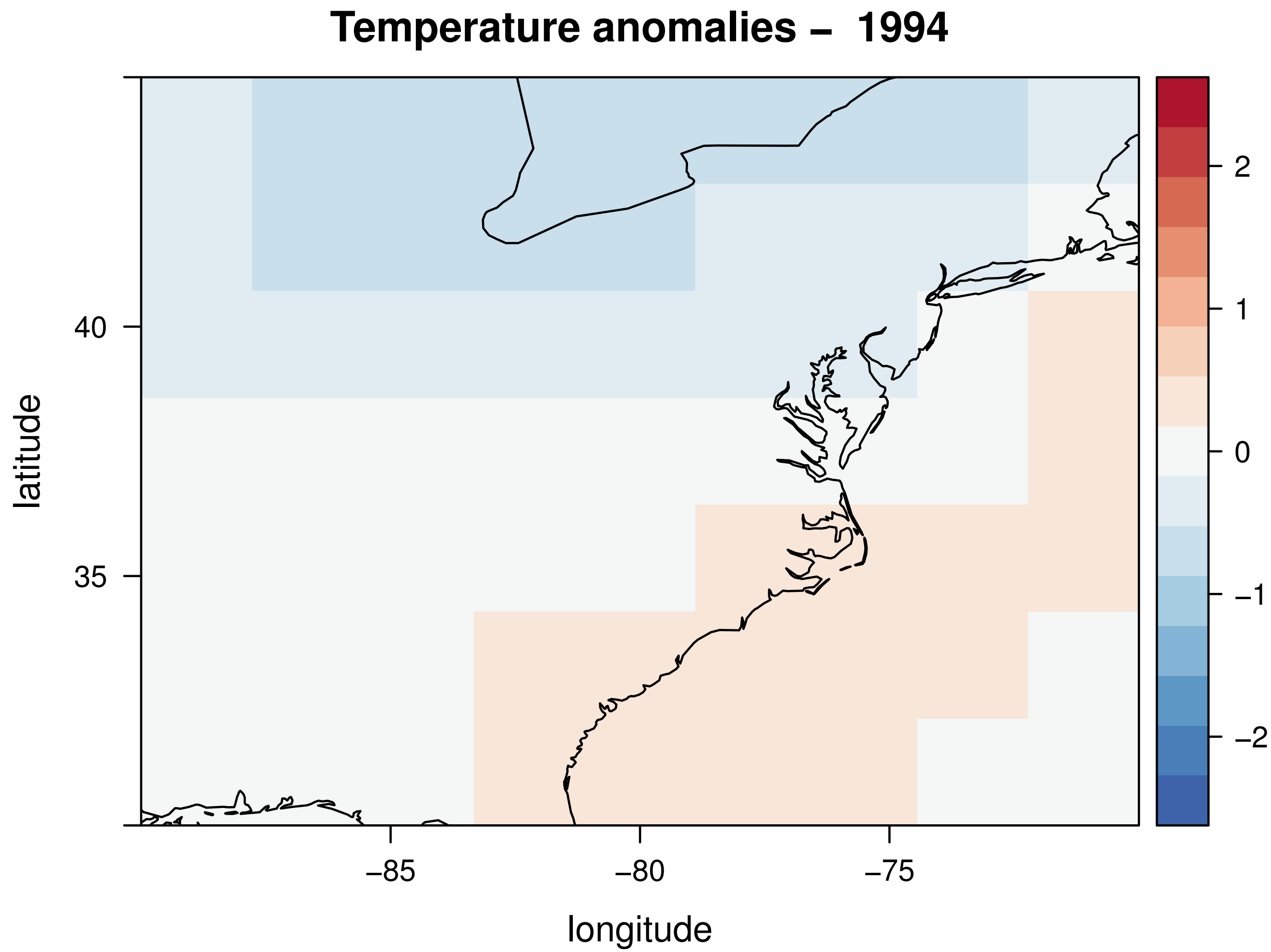
Temperature Anomalies in Eastern USA



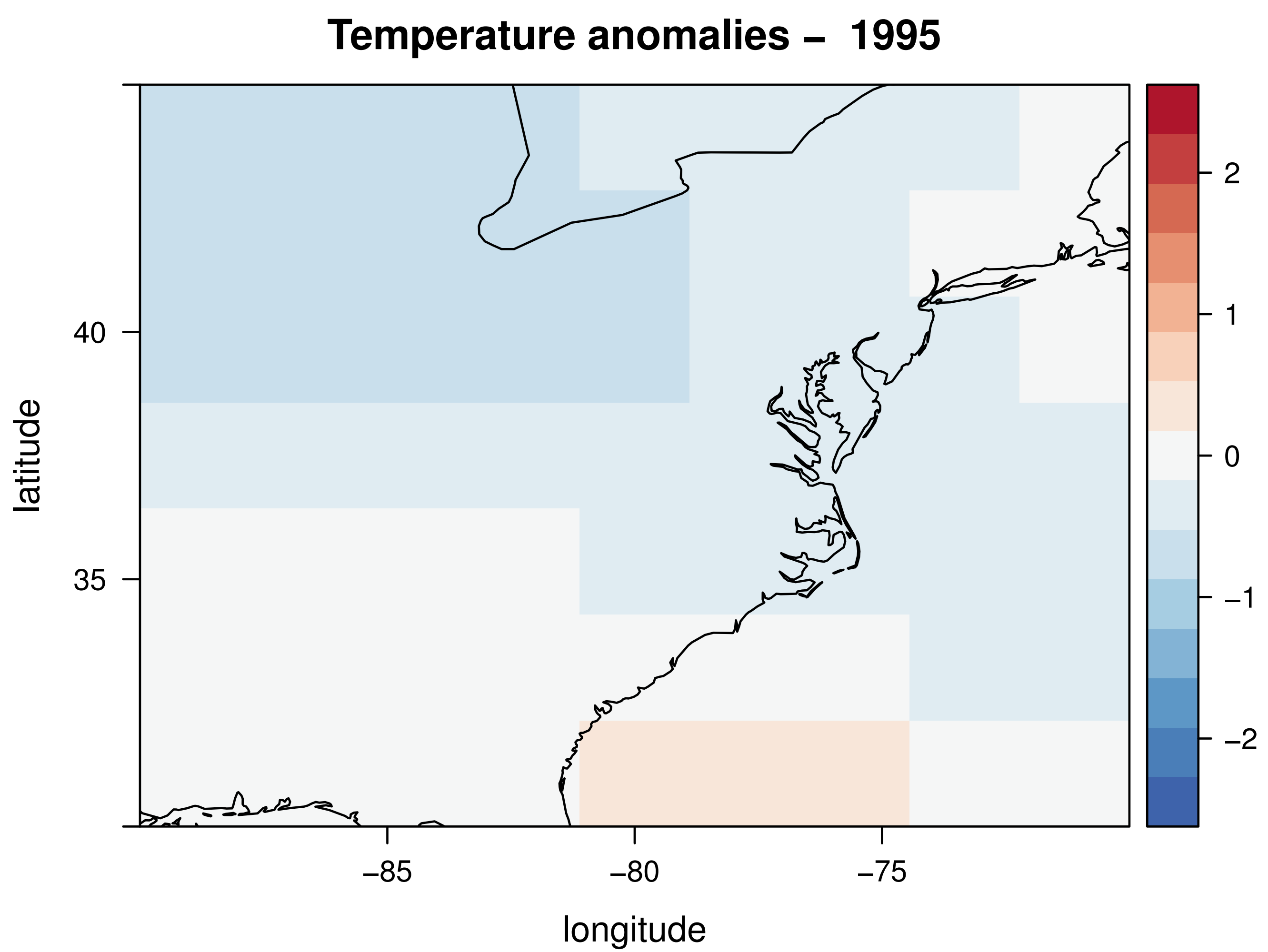
Temperature Anomalies in Eastern USA



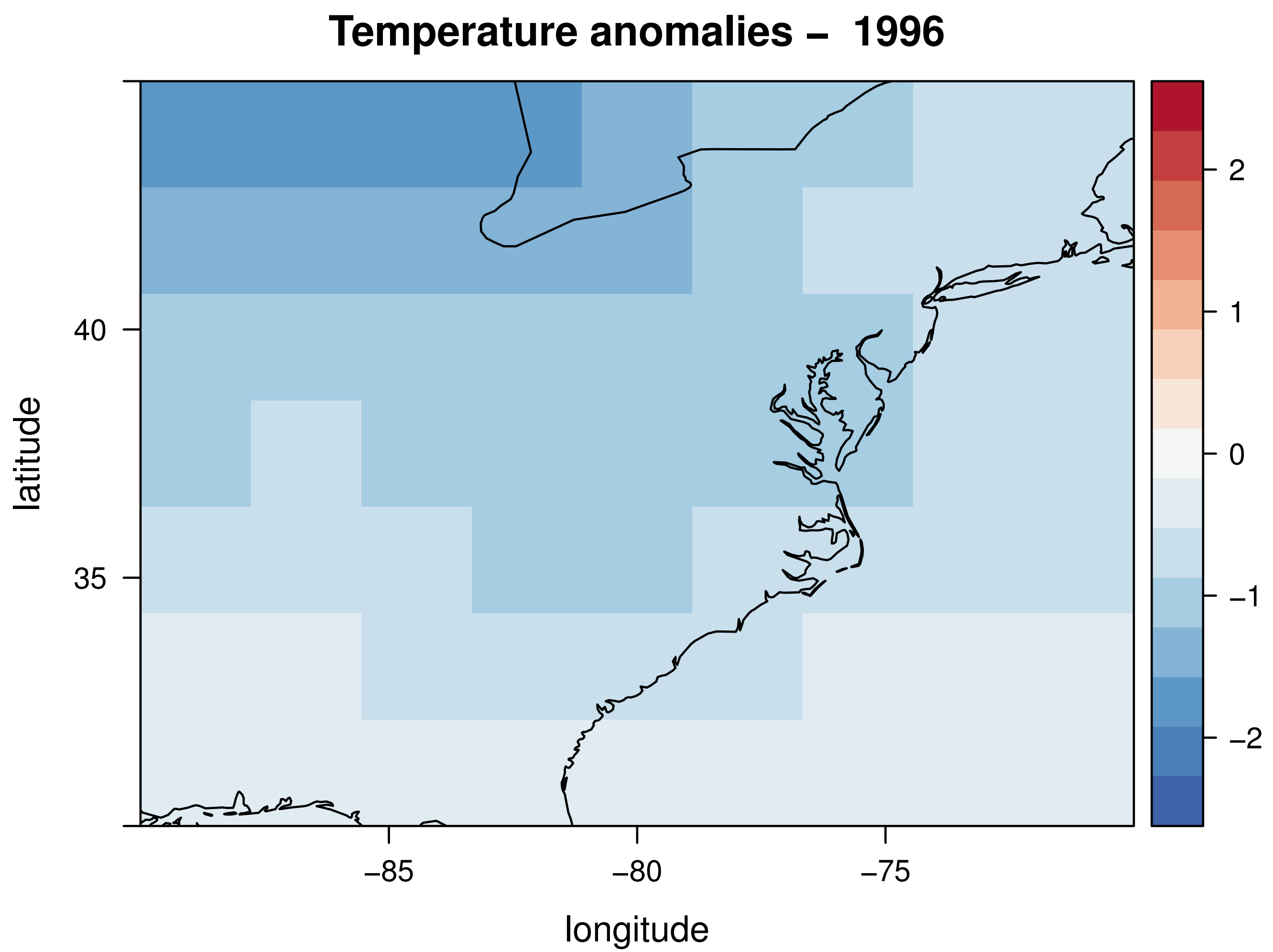
Temperature Anomalies in Eastern USA



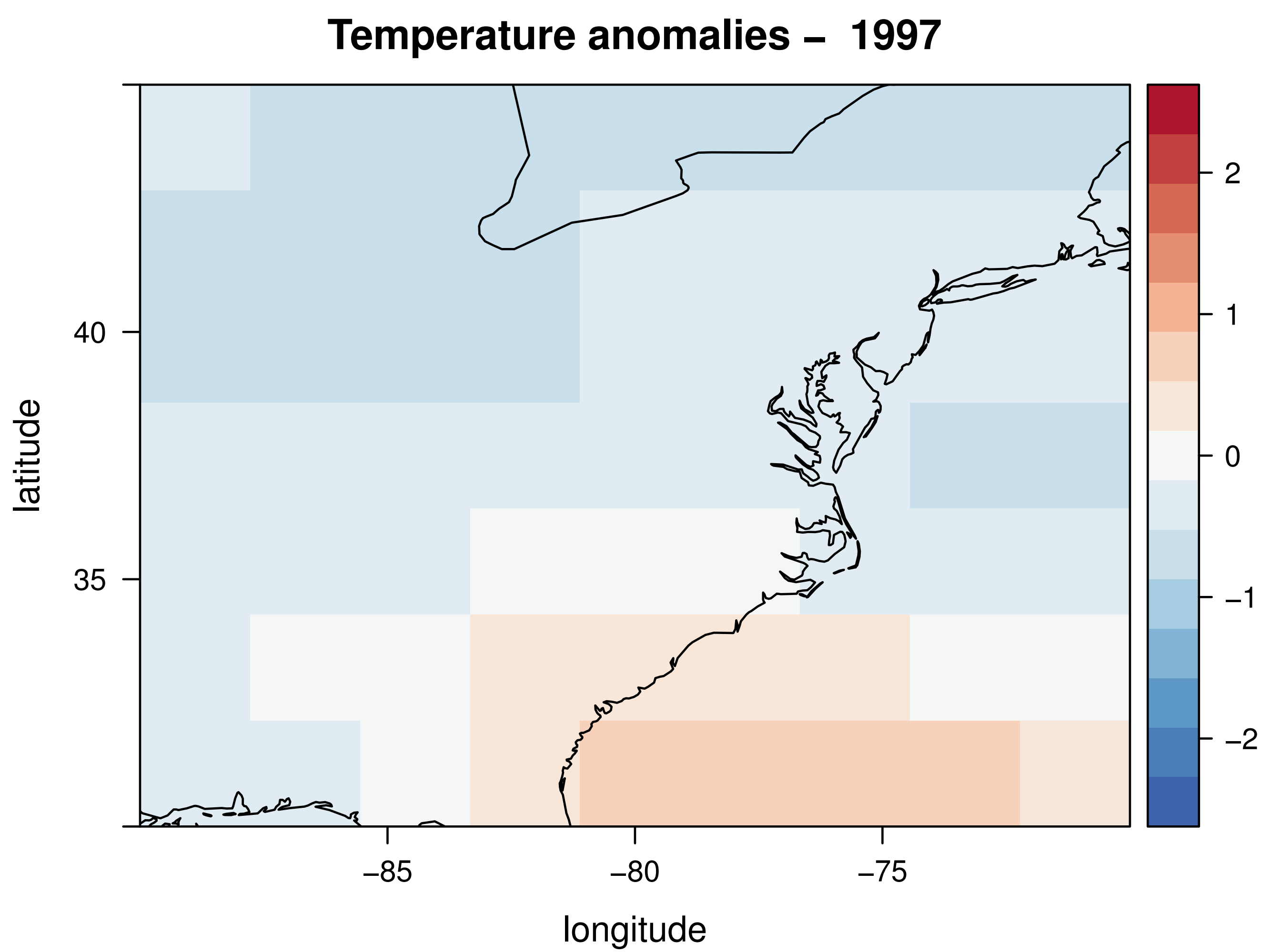
Temperature Anomalies in Eastern USA



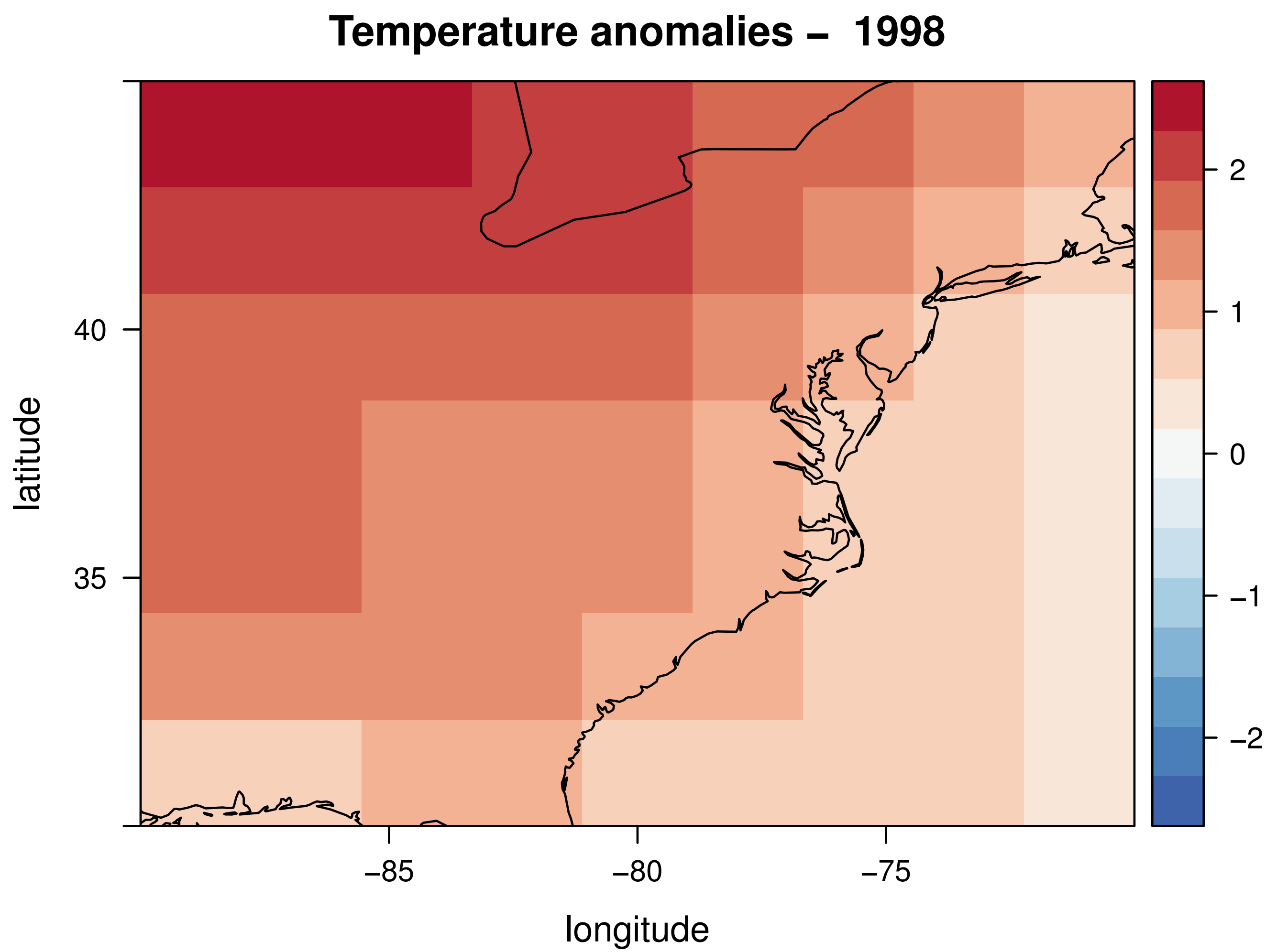
Temperature Anomalies in Eastern USA



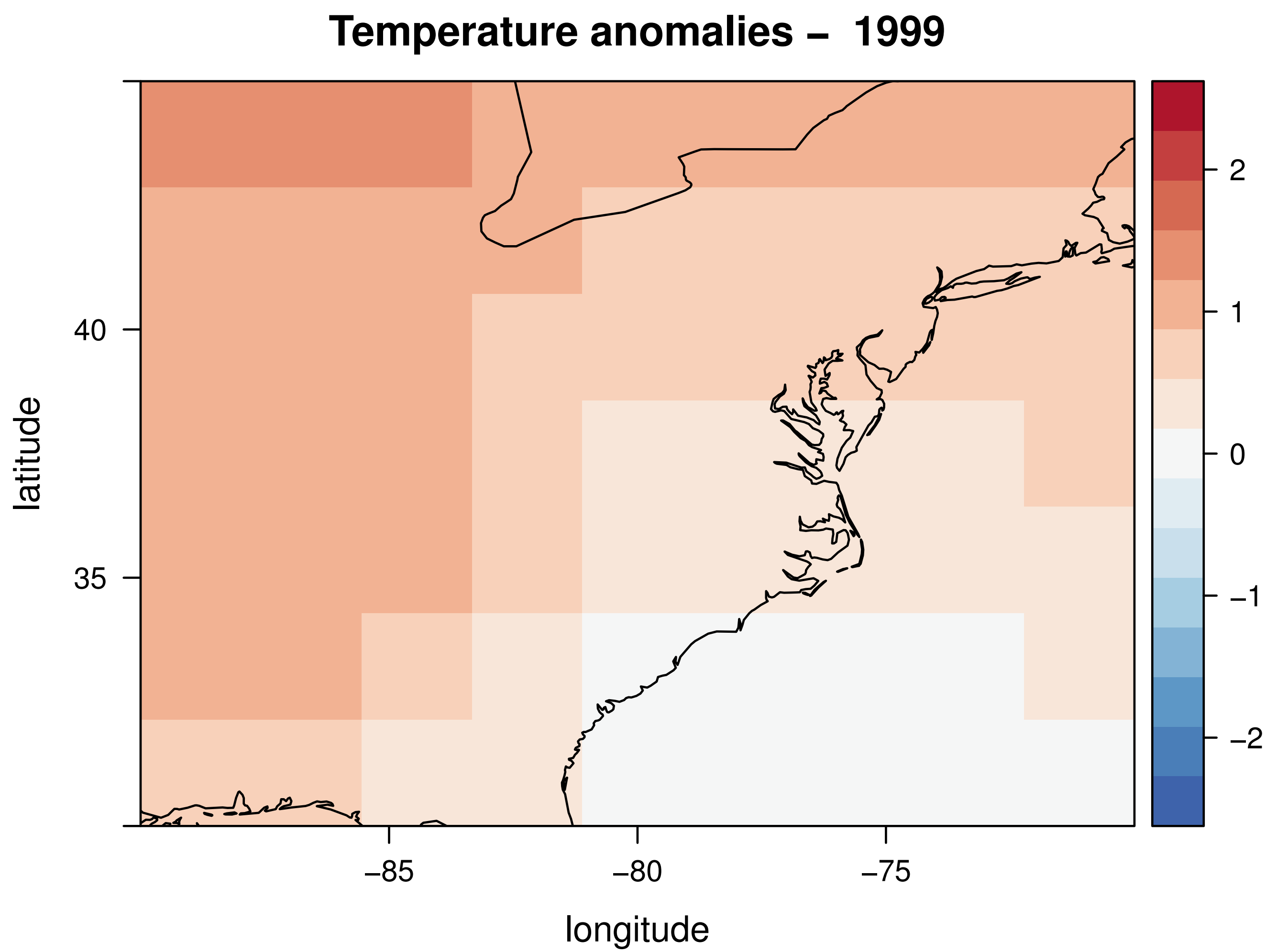
Temperature Anomalies in Eastern USA



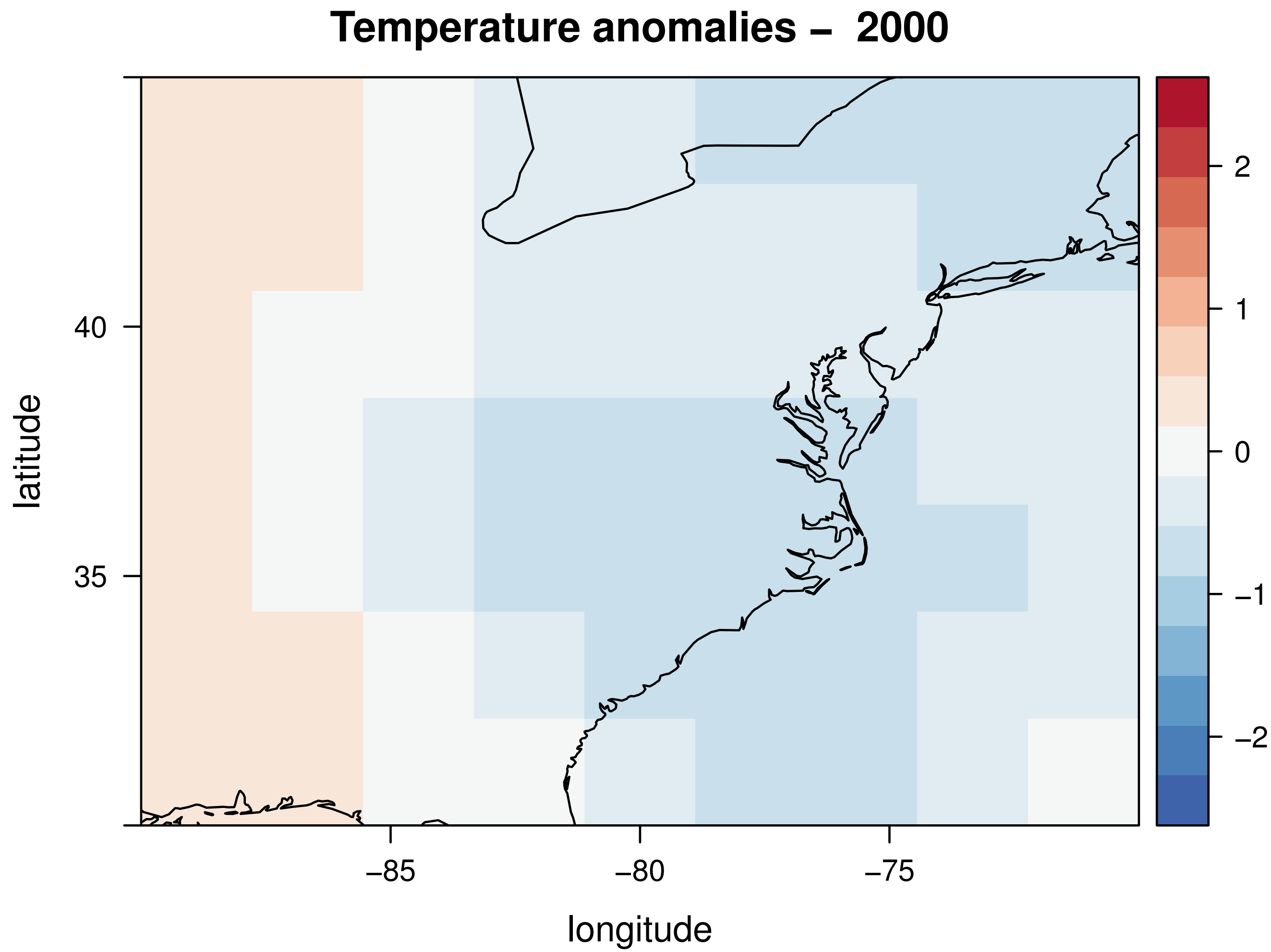
Temperature Anomalies in Eastern USA



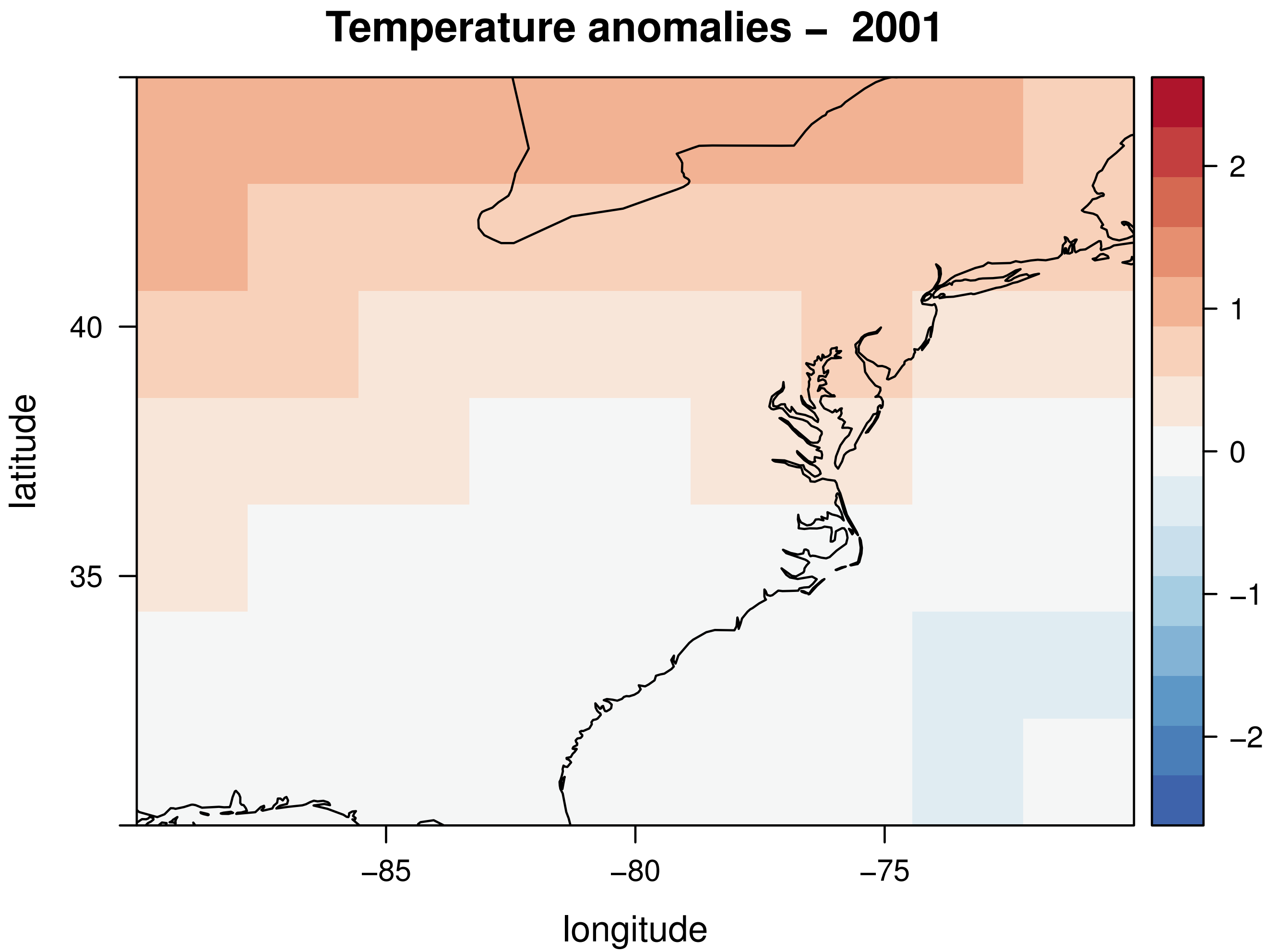
Temperature Anomalies in Eastern USA



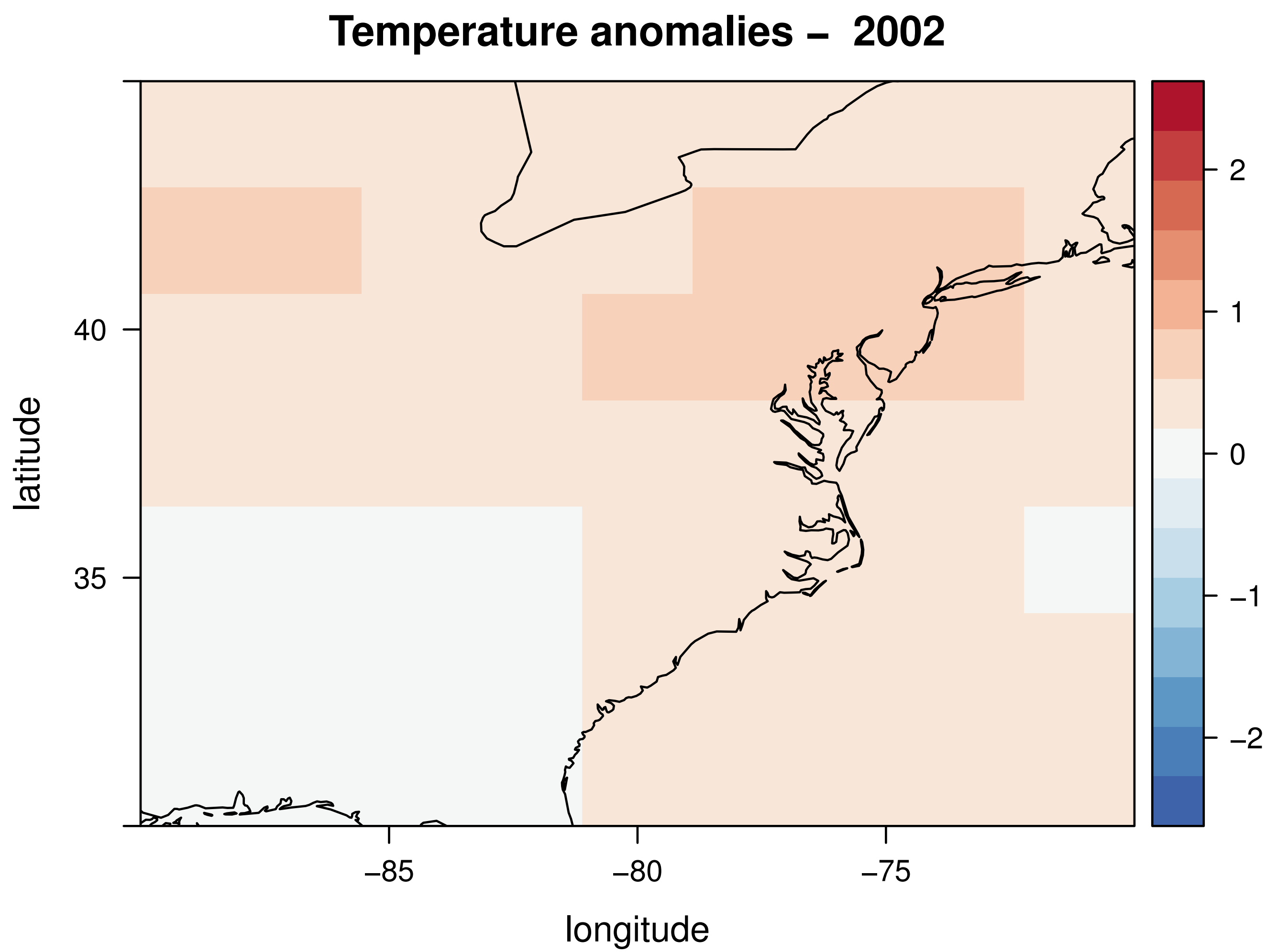
Temperature Anomalies in Eastern USA



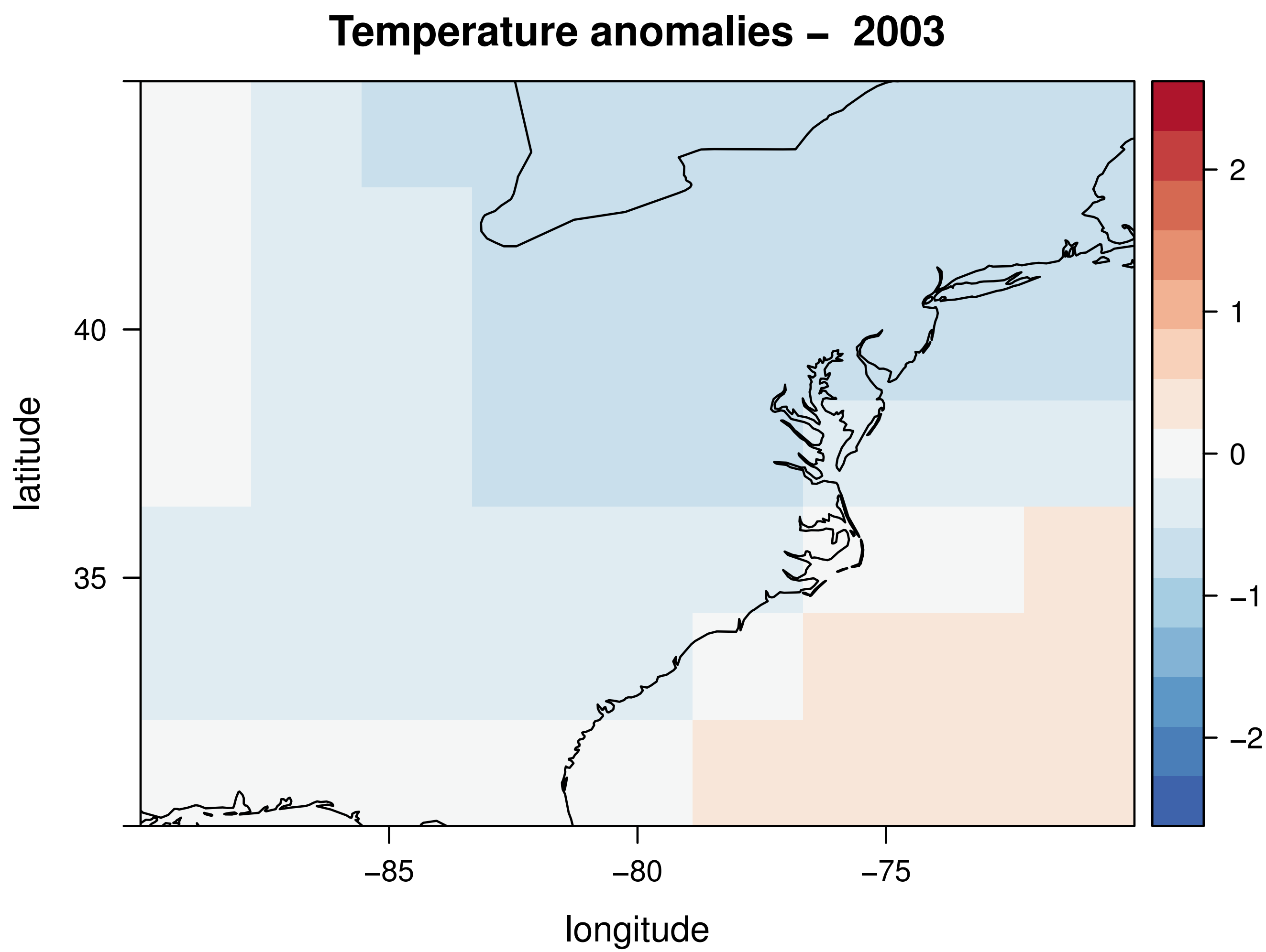
Temperature Anomalies in Eastern USA



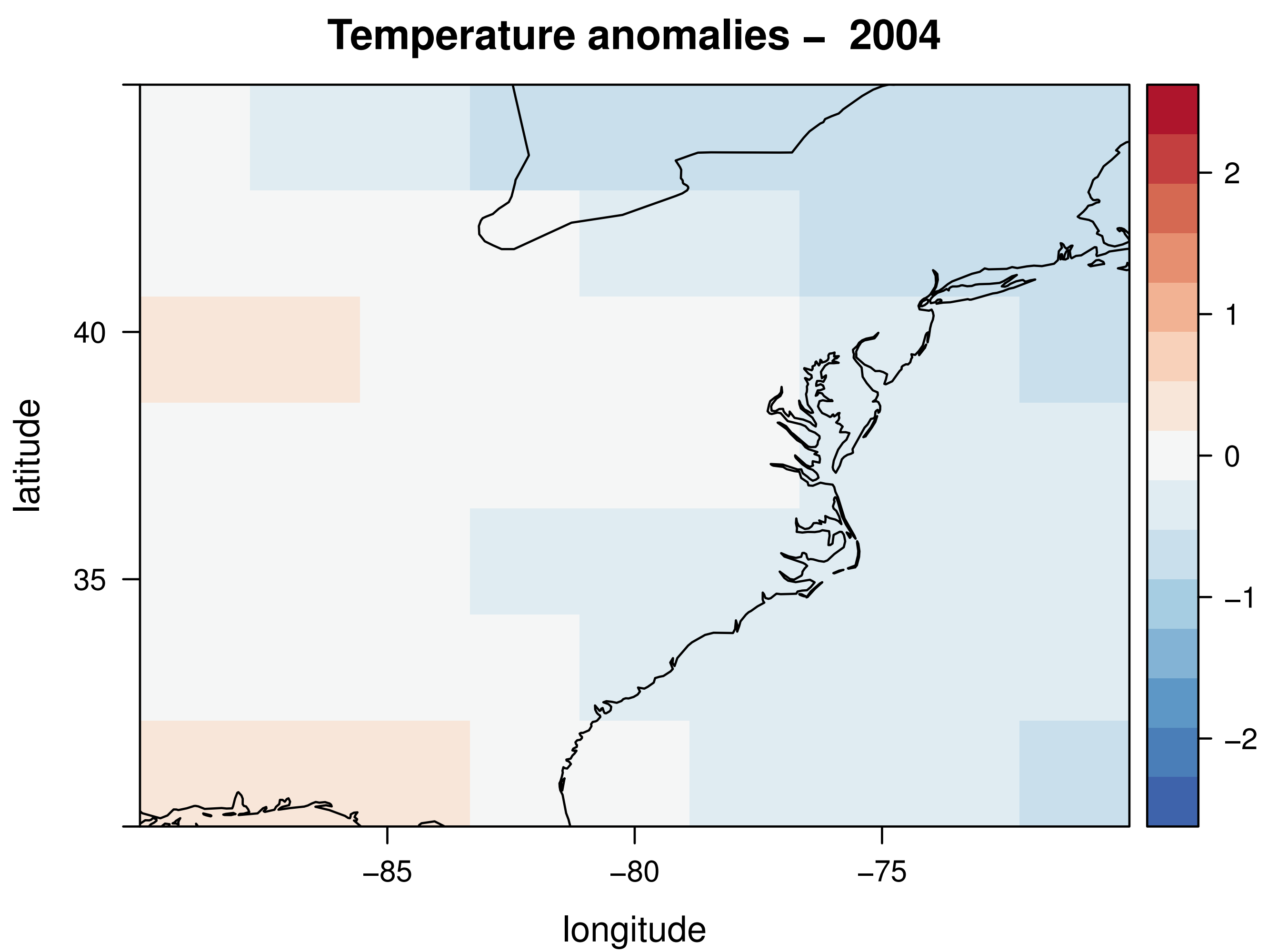
Temperature Anomalies in Eastern USA



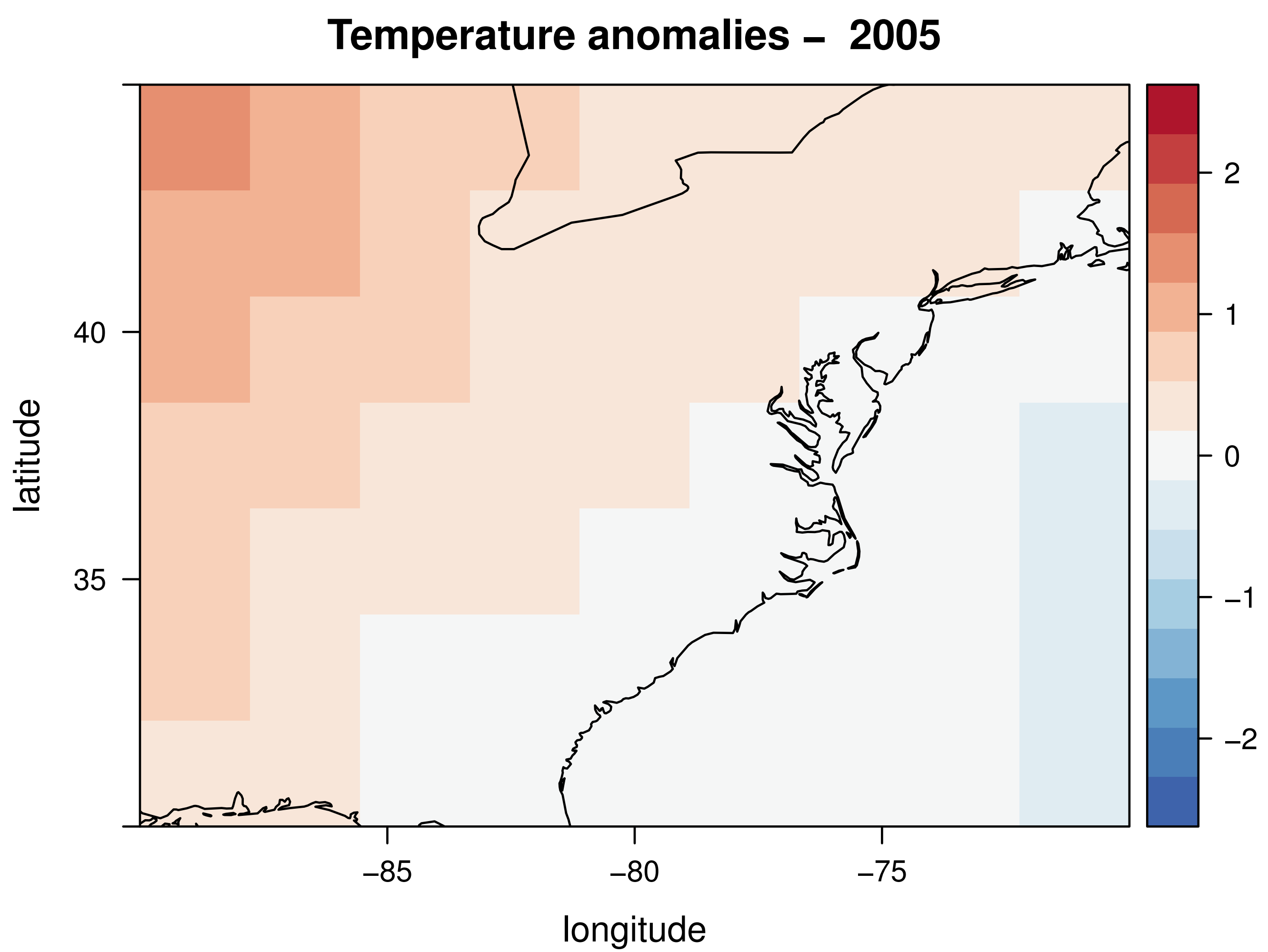
Temperature Anomalies in Eastern USA



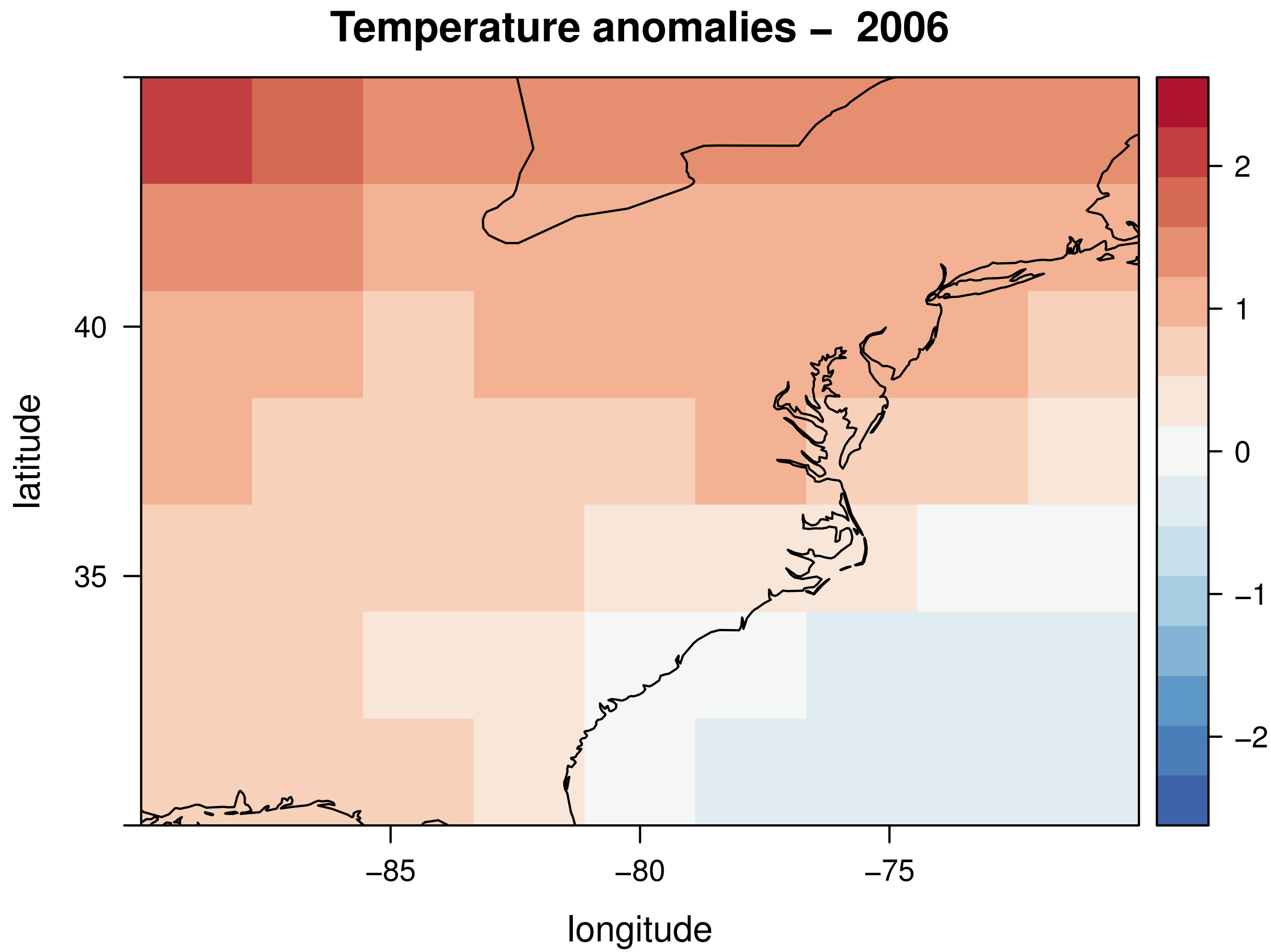
Temperature Anomalies in Eastern USA



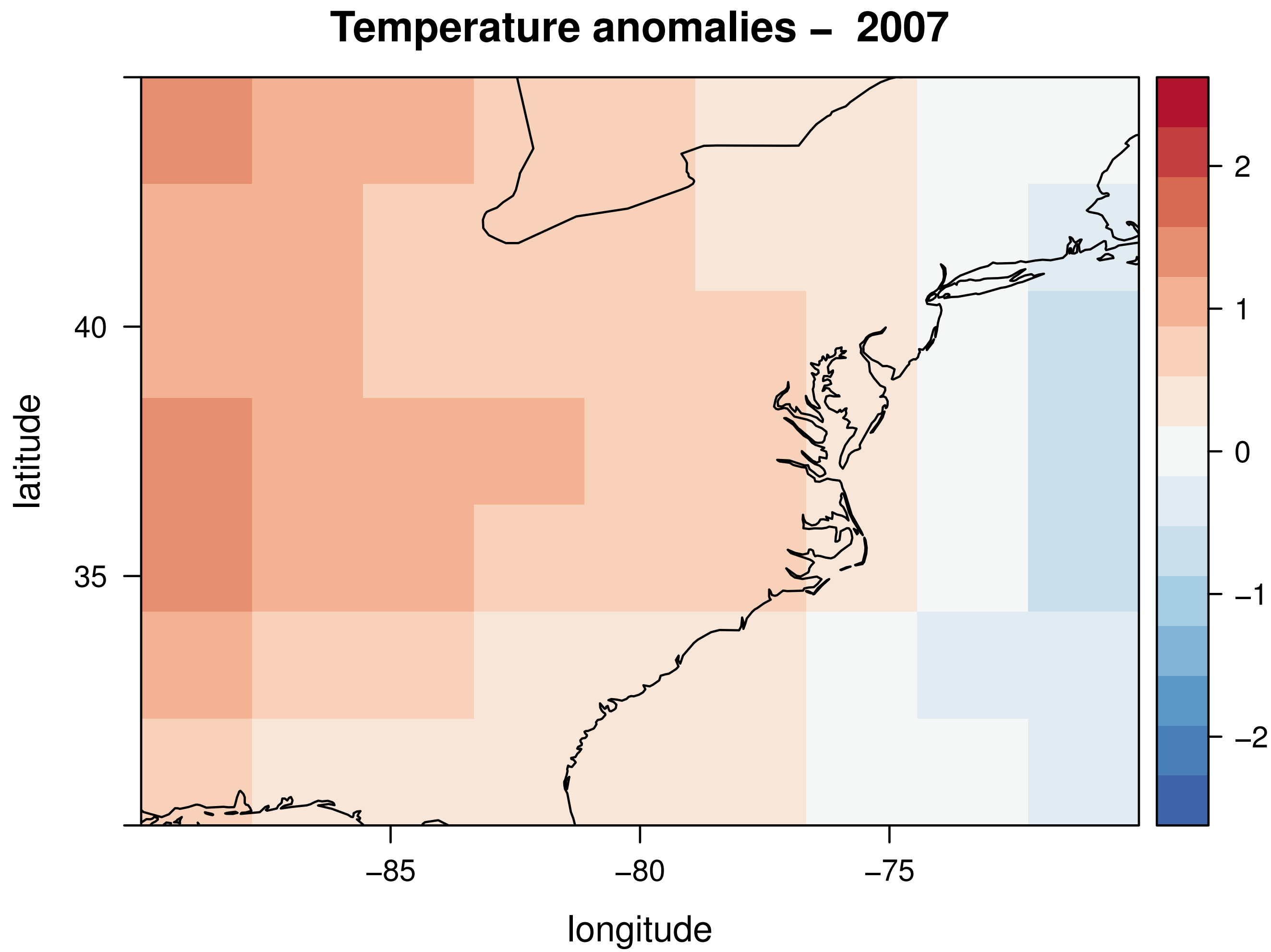
Temperature Anomalies in Eastern USA



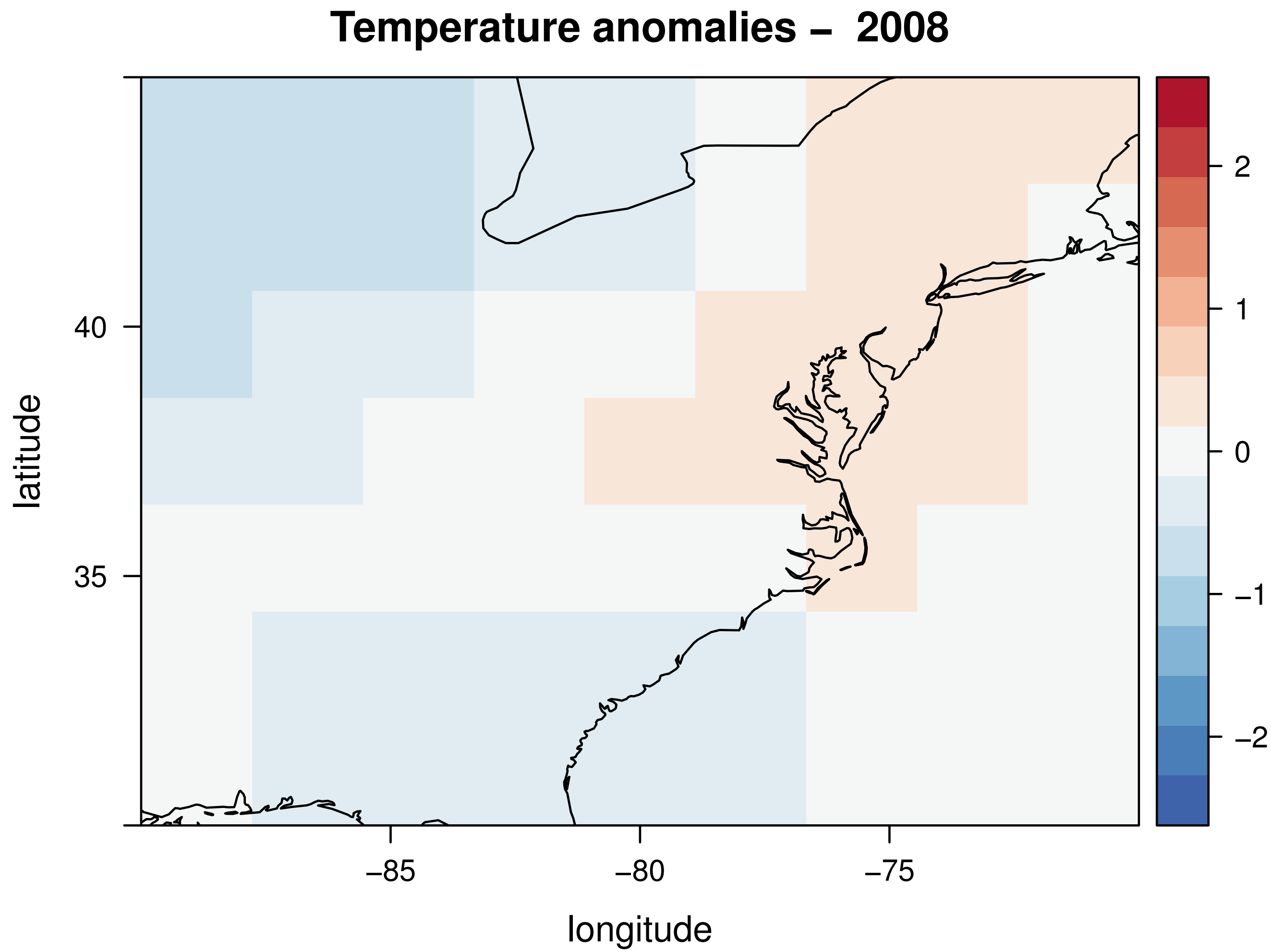
Temperature Anomalies in Eastern USA



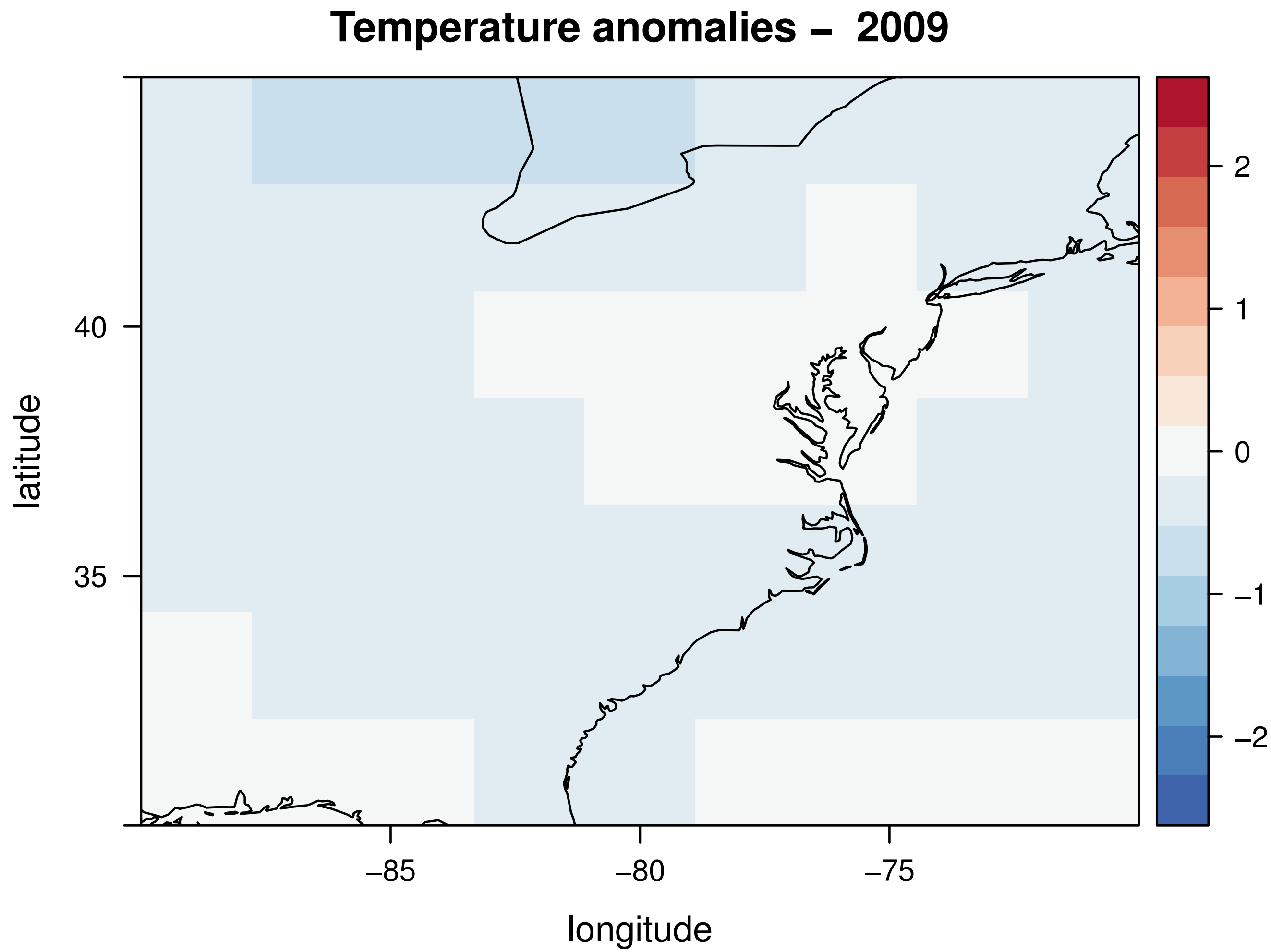
Temperature Anomalies in Eastern USA



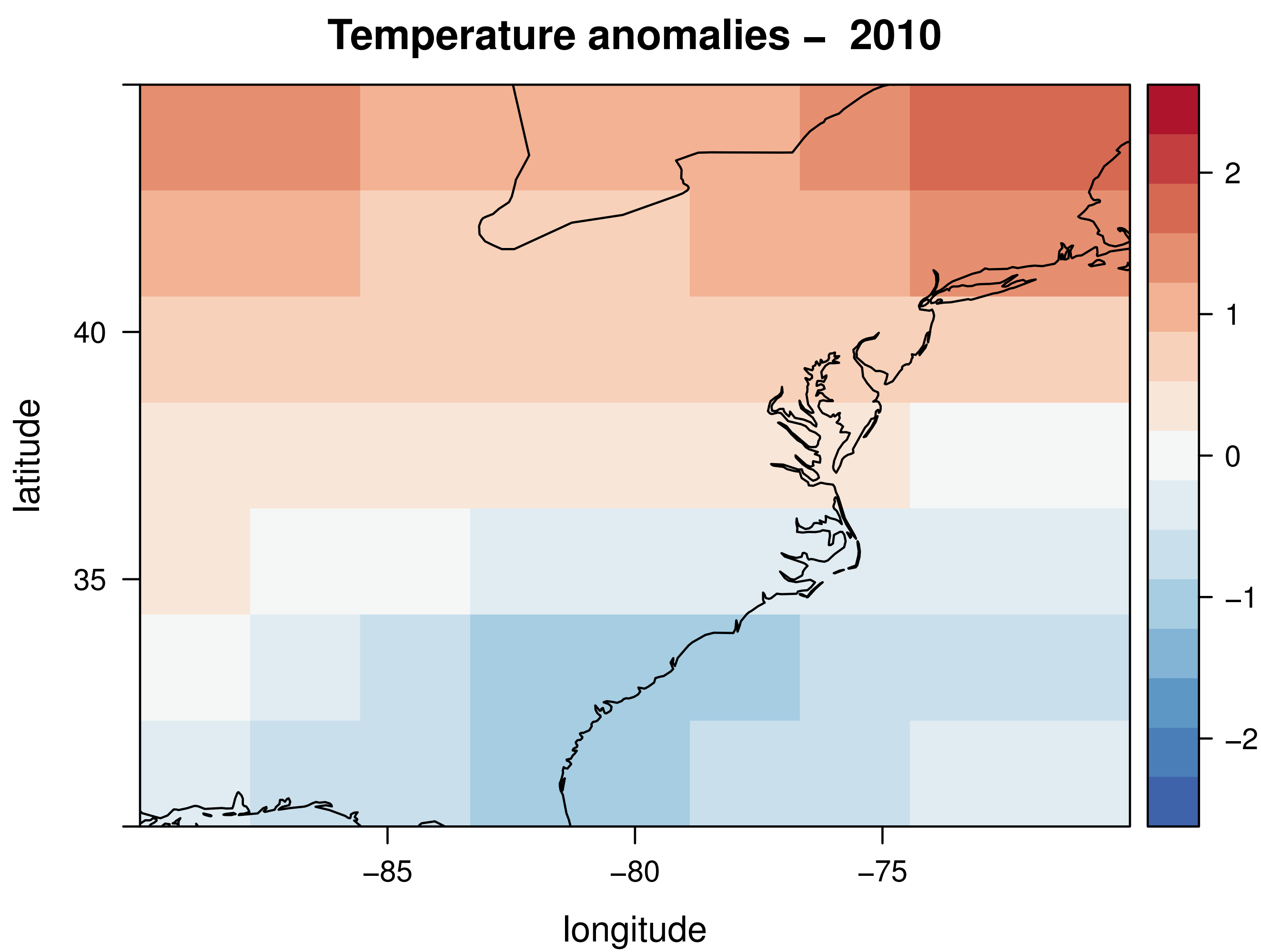
Temperature Anomalies in Eastern USA



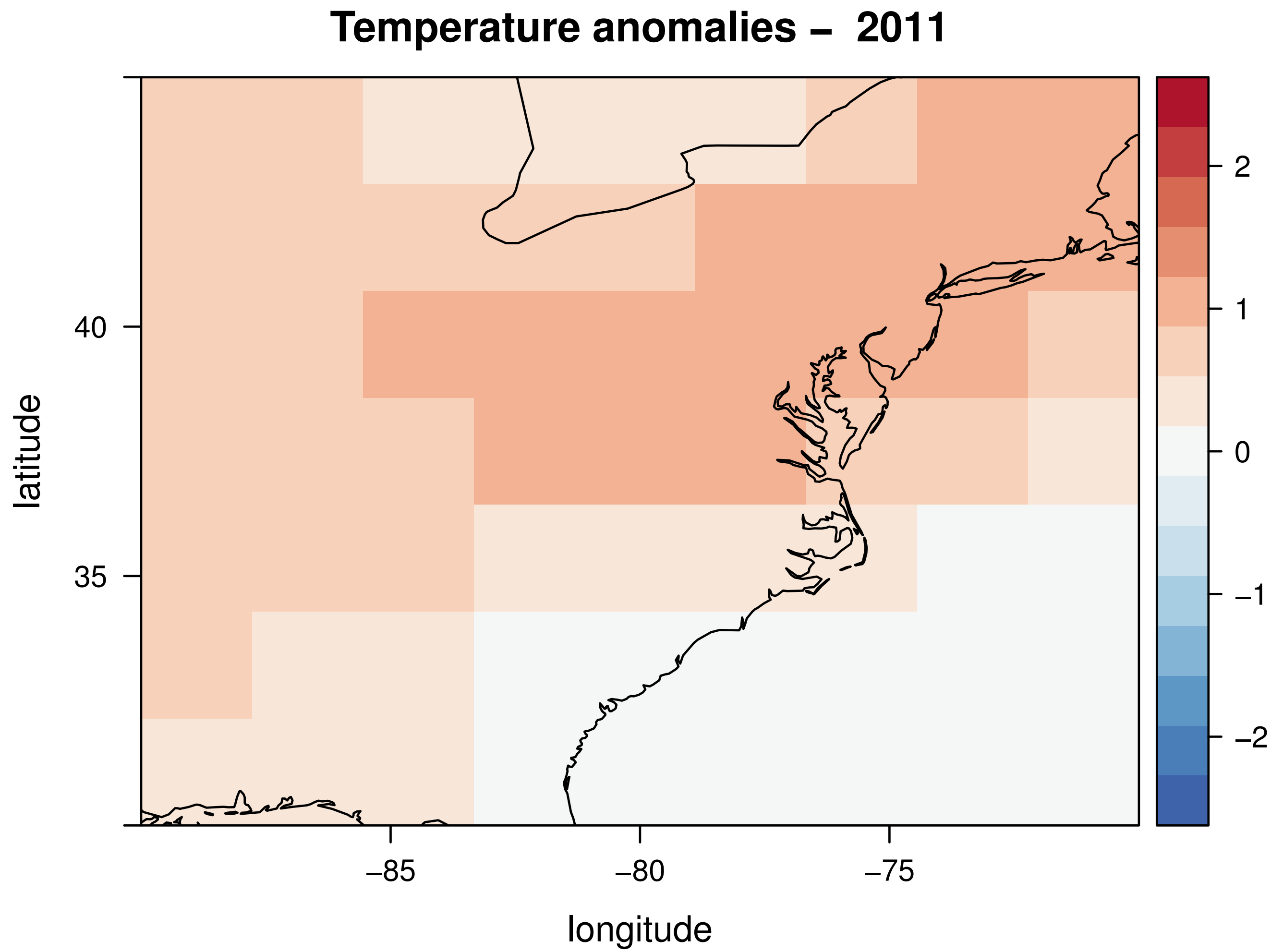
Temperature Anomalies in Eastern USA



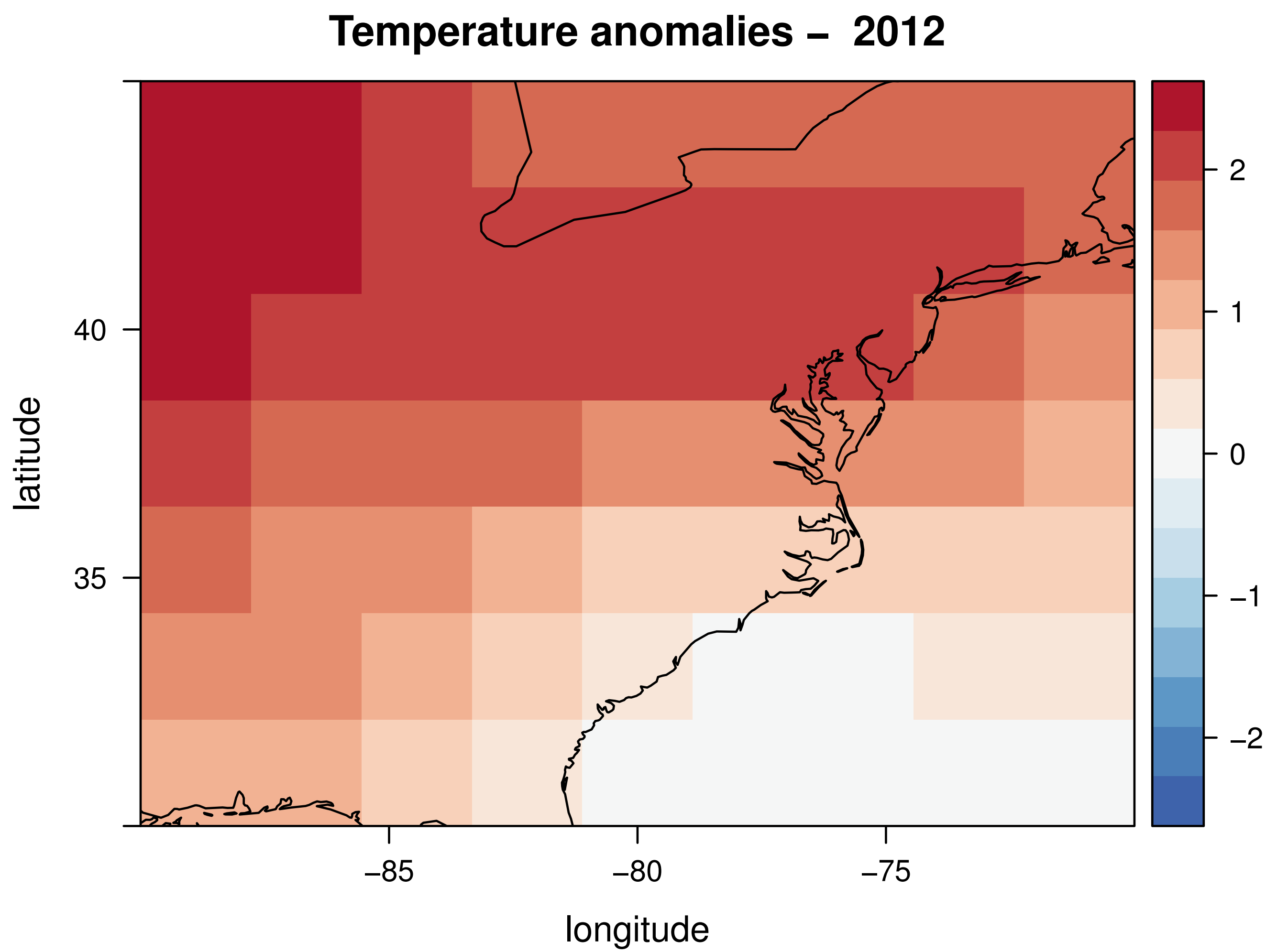
Temperature Anomalies in Eastern USA



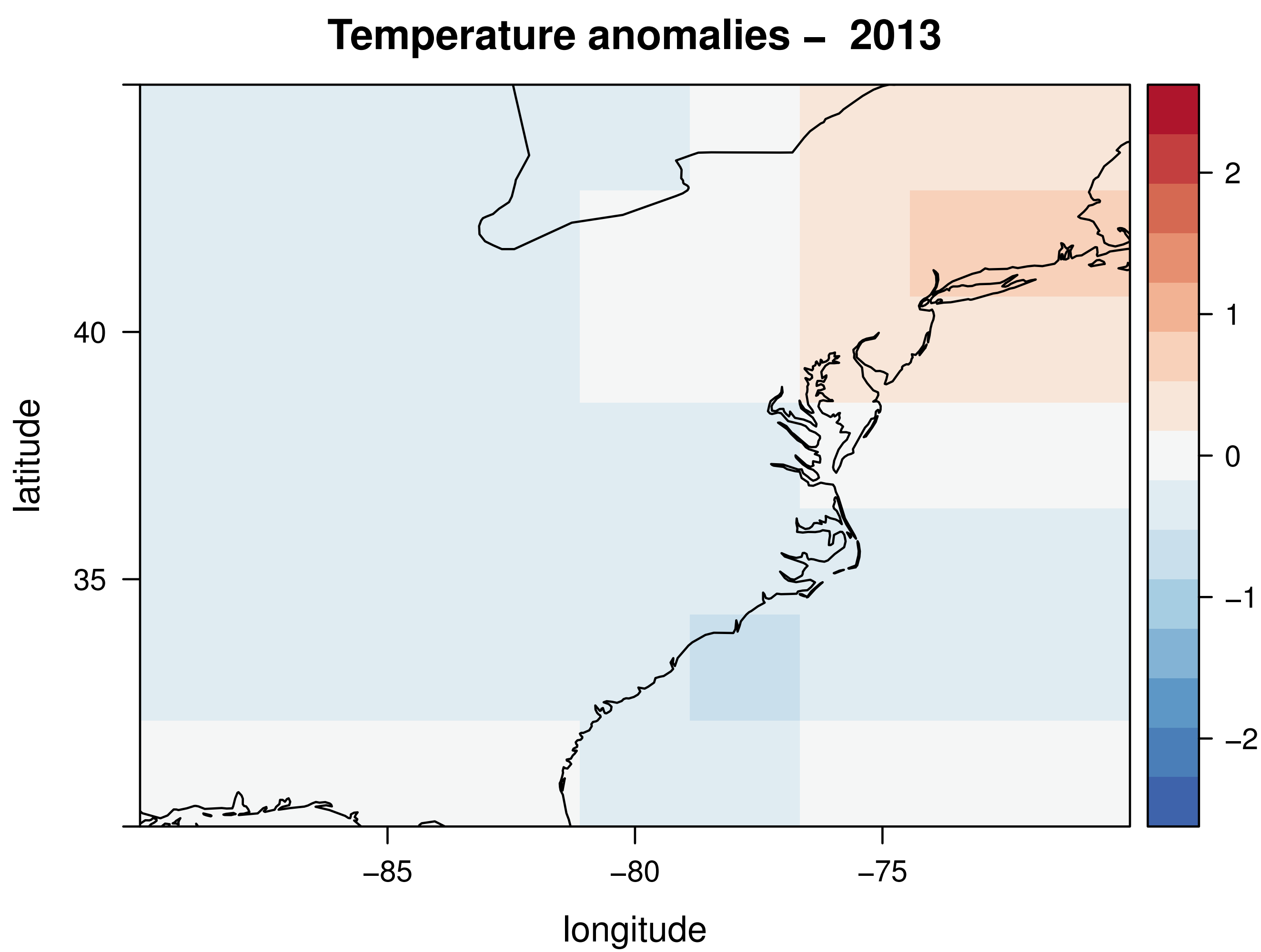
Temperature Anomalies in Eastern USA



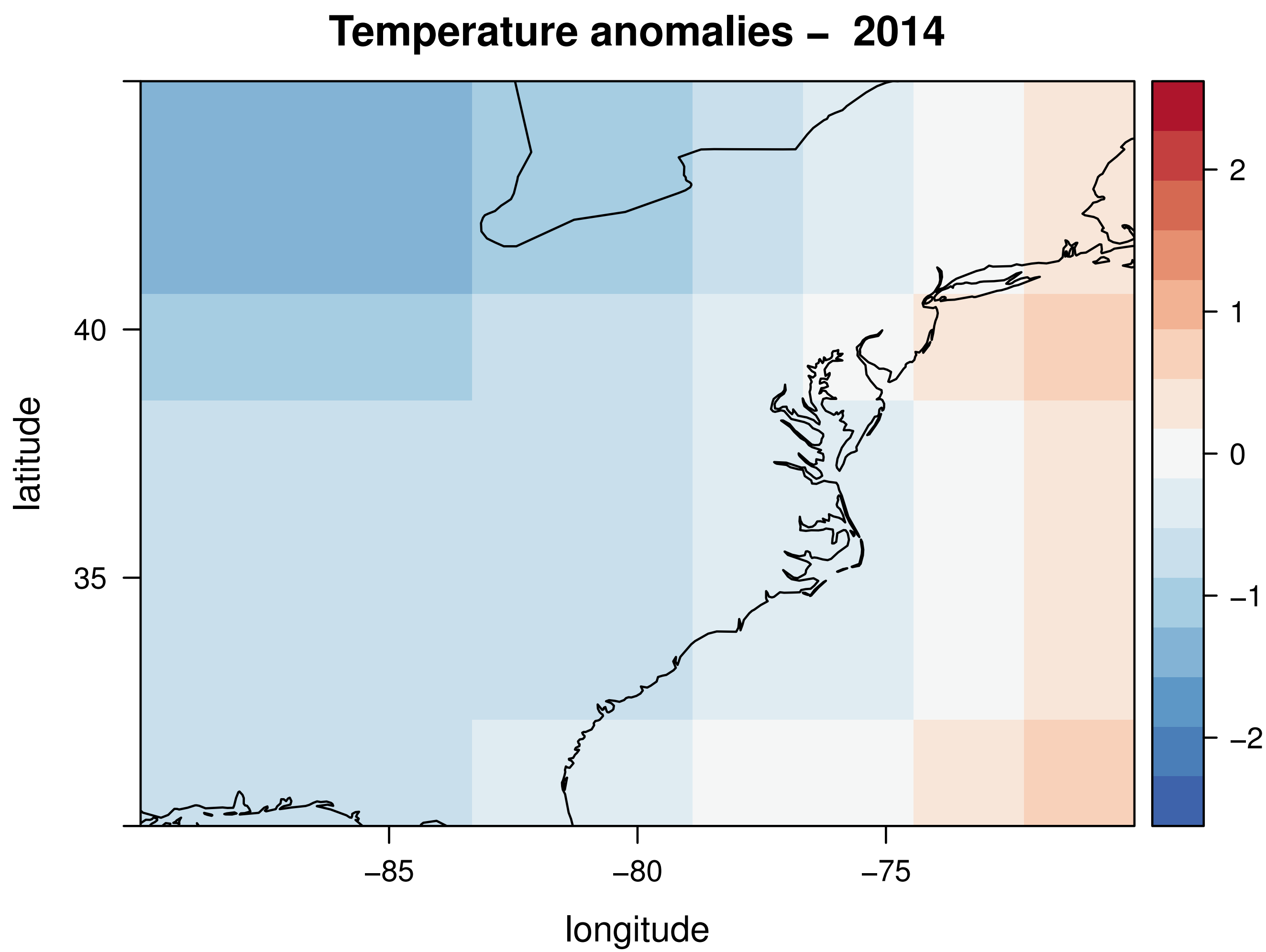
Temperature Anomalies in Eastern USA



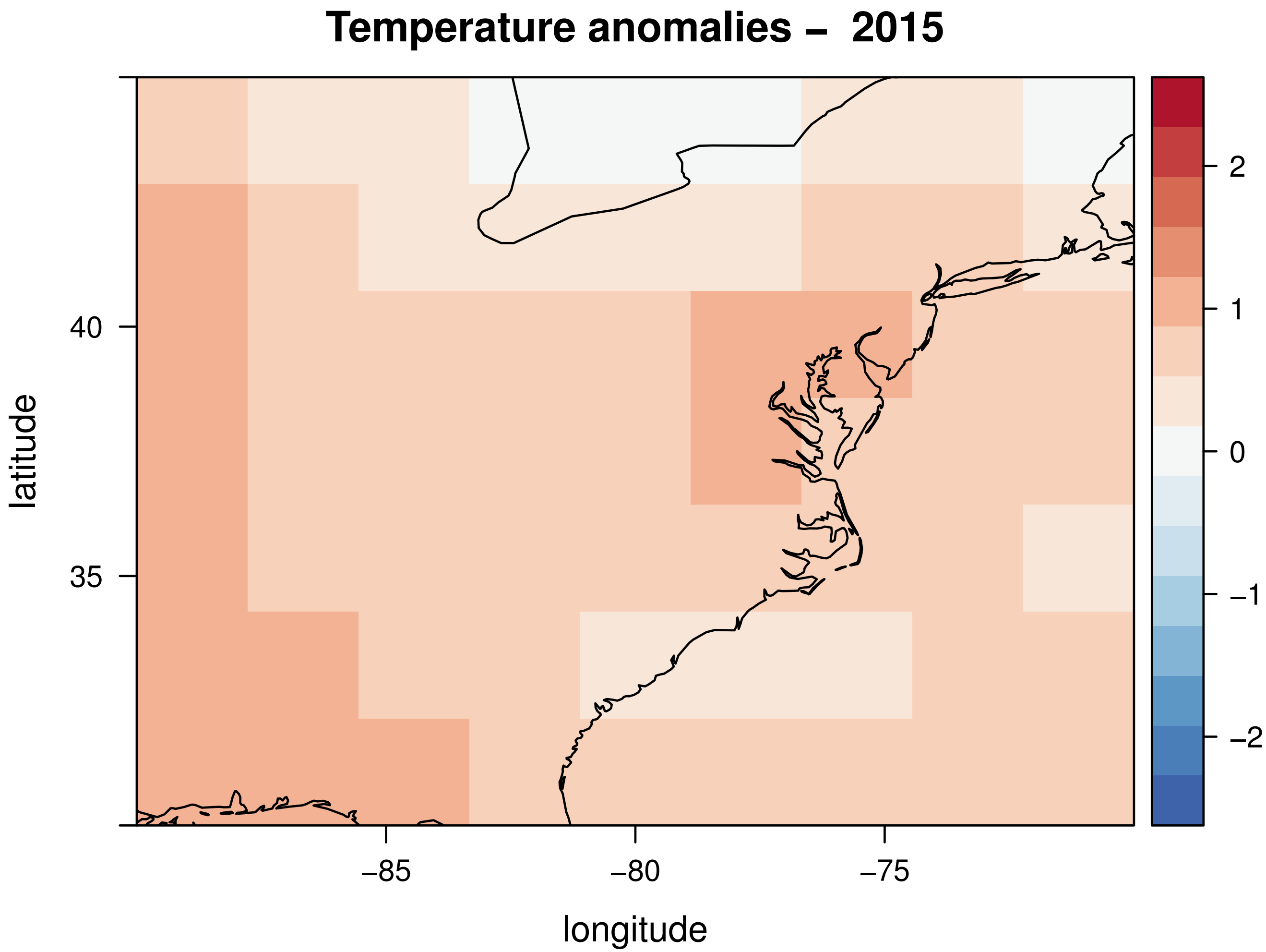
Temperature Anomalies in Eastern USA



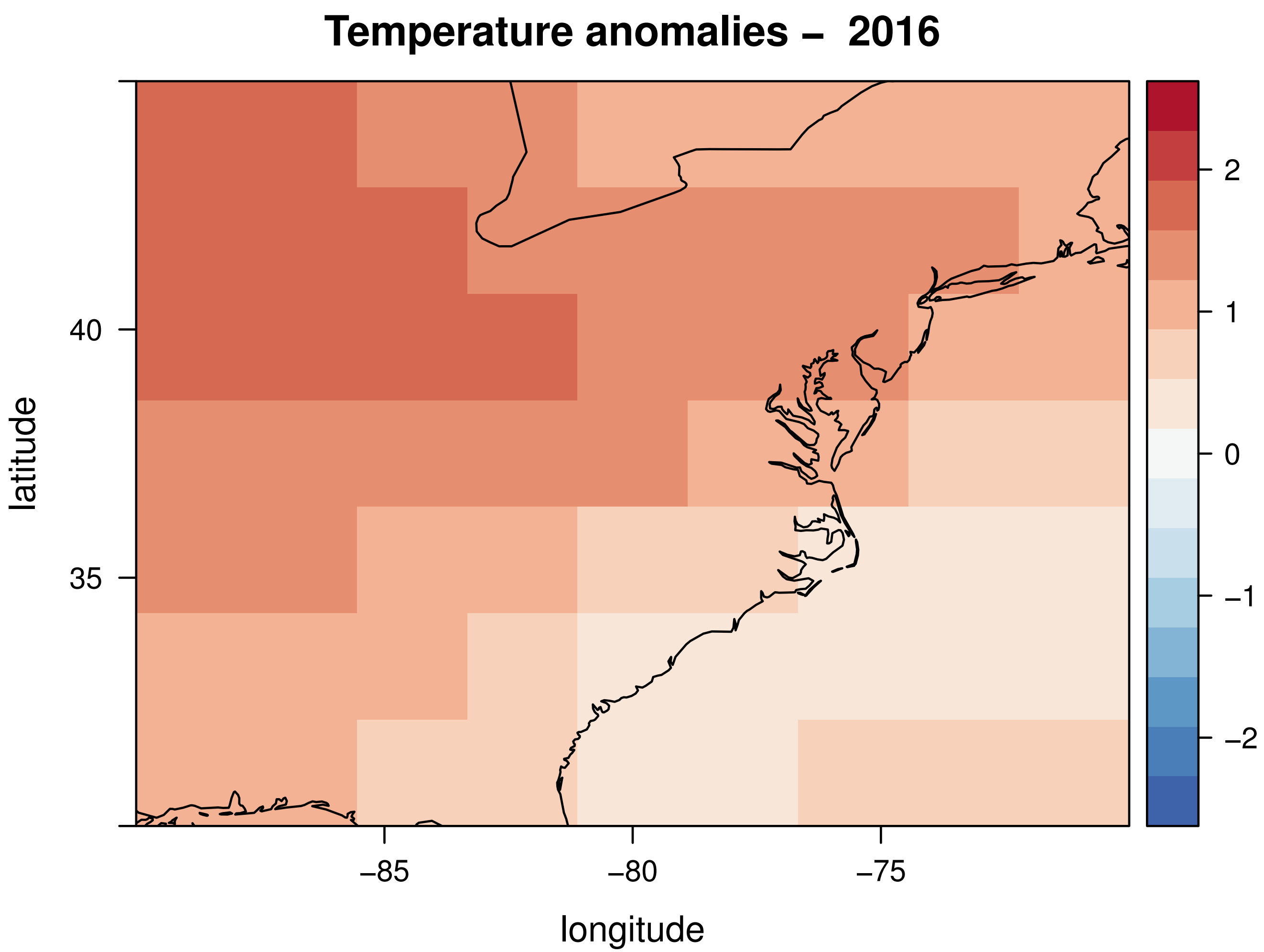
Temperature Anomalies in Eastern USA



Temperature Anomalies in Eastern USA



Temperature Anomalies in Eastern USA



Dynamic spatiotemporal factor model for temperature anomalies

- $\mathbf{y}_t$ = vector of observed temperatures in year t;
- Factor loadings vary spatially.
- Elements of each column of factor loadings matrix $\mathbf{B}$ follow a spatial Gaussian process;
- $\mathbf{x}_t$ is a k-dimensional vector of common factors.
- lth element x_{tl} follows an AR(p_l) process.

Dynamic spatiotemporal factor model for temperature anomalies

- $\mathbf{y}_t$ = vector of observed temperatures in year t;
- Factor loadings vary spatially.
- Elements of each column of factor loadings matrix $\mathbf{B}$ follow a spatial Gaussian process;
- $\mathbf{x}_t$ is a k-dimensional vector of common factors.
- lth element x_{tl} follows an AR(p_l) process.

Dynamic spatiotemporal factor model for temperature anomalies

- $\mathbf{y}_t$ = vector of observed temperatures in year t;
- Factor loadings vary spatially.
- Elements of each column of factor loadings matrix $\mathbf{B}$ follow a spatial Gaussian process;
- $\mathbf{x}_t$ is a k-dimensional vector of common factors.
- lth element x_{tl} follows an AR(p_l) process.

Dynamic spatiotemporal factor model for temperature anomalies

$$\begin{aligned}
\mathbf{y}_t &= \mathbf{B}\mathbf{x}_t + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim N(\mathbf{0}, \mathbf{V}), \\
\mathbf{x}_t &= \mathbf{F}'\theta_t, \\
\theta_t &= \mathbf{G}\theta_{t-1} + \omega_t, \quad \omega_t \sim N(\mathbf{0}, \mathbf{W}).
\end{aligned}$$

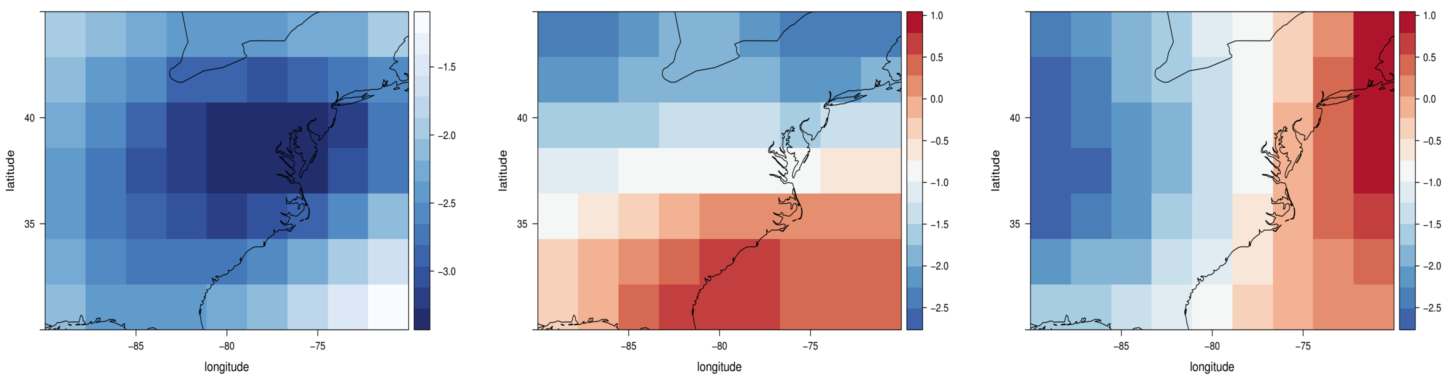
- $b_{.l} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_l)$,
- $\{\boldsymbol{\Sigma}_l\}_{ij} = \tau_l^2 \rho_l(||s_i - s_j||; \lambda_l)$.

Dynamic spatiotemporal factor model for temperature anomalies

$$
\begin{aligned}
\mathbf{y}_t &= \mathbf{B}\mathbf{x}_t + \boldsymbol{\nu}_t, \quad \boldsymbol{\nu}_t \sim N(\mathbf{0}, \mathbf{V}), \\
\mathbf{x}_t &= \mathbf{F}'\boldsymbol{\theta}_t, \\
\boldsymbol{\theta}_t &= \mathbf{G}\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim N(\mathbf{0}, \mathbf{W}).
\end{aligned}
$$

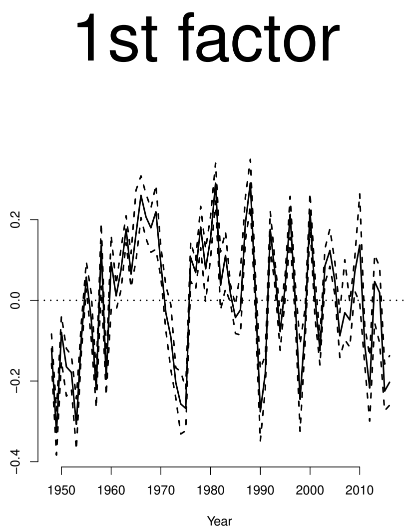
- $b_{.l} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_l)$,
- $\{\boldsymbol{\Sigma}_l\}_{ij} = \tau_l^2 \rho_l(||s_i - s_j||; \boldsymbol{\lambda}_l)$.

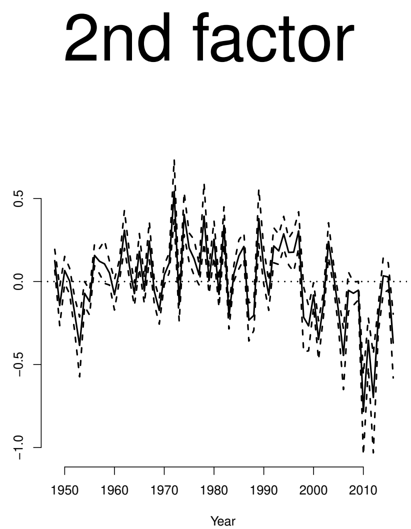
Dynamic spatiotemporal 3-factor model: Factor loadings

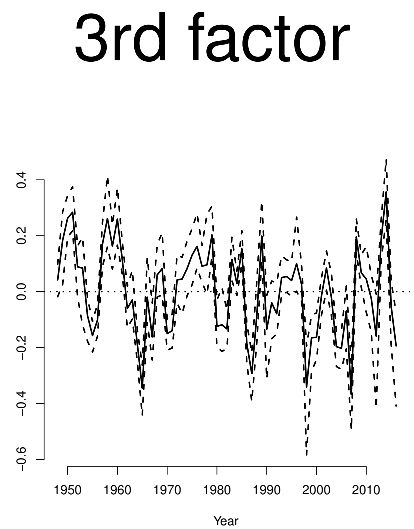


1st factor 2nd factor 3rd factor

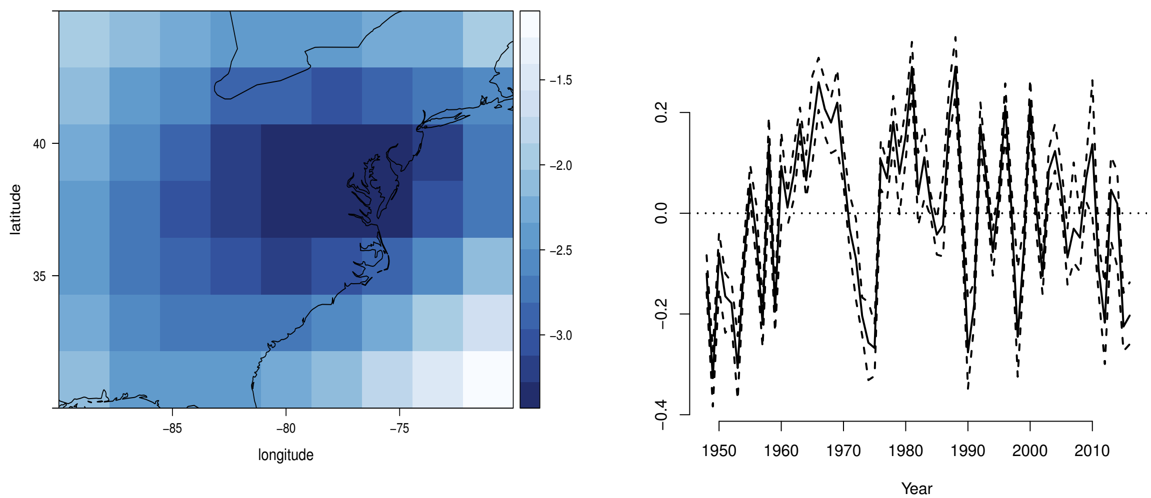
Common factors



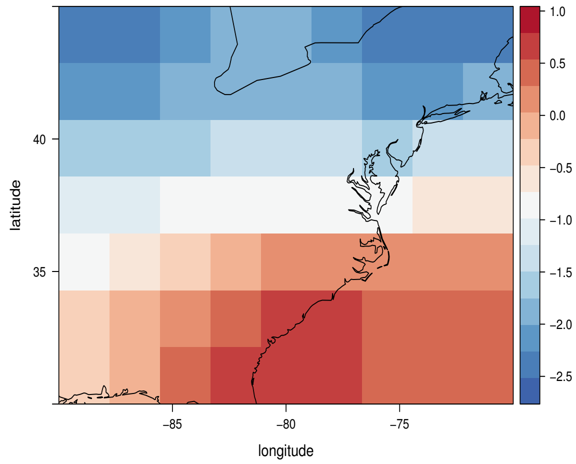


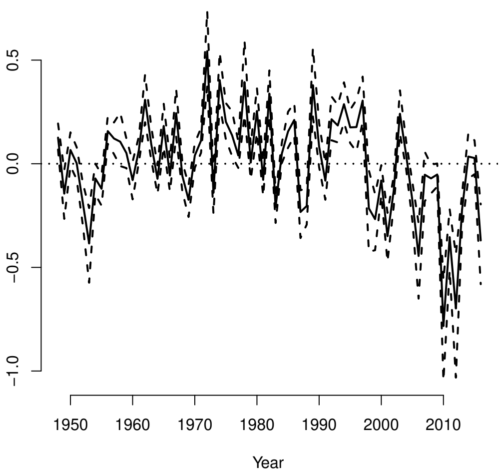


1st common factor and factor loadings

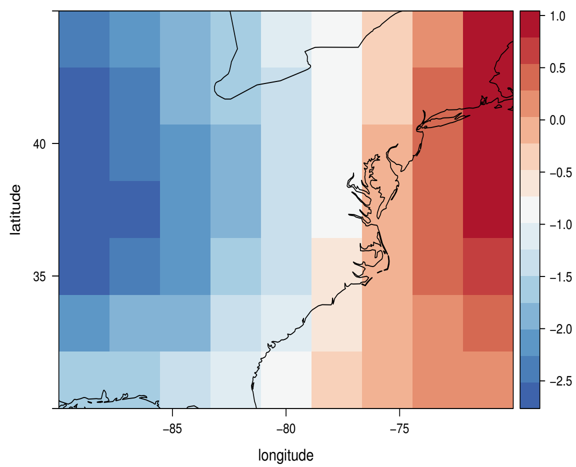


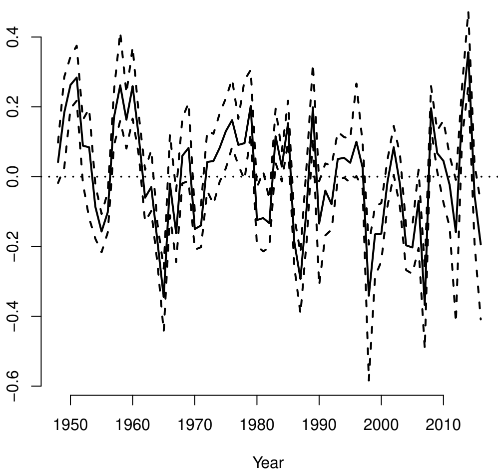
2nd common factor and factor loadings





3rd common factor and factor loadings

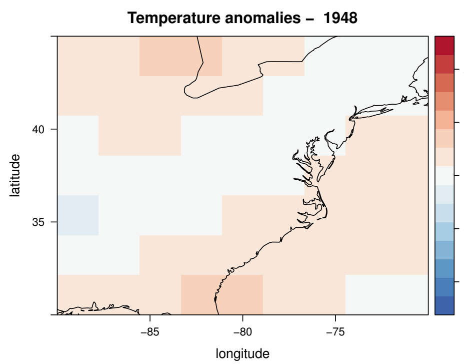


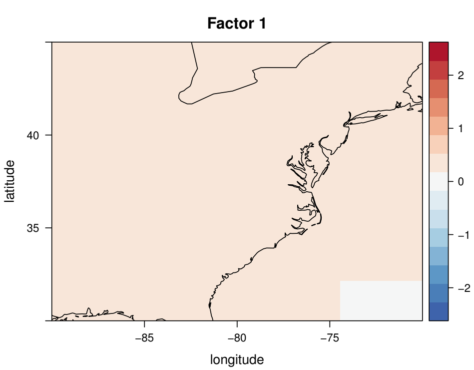


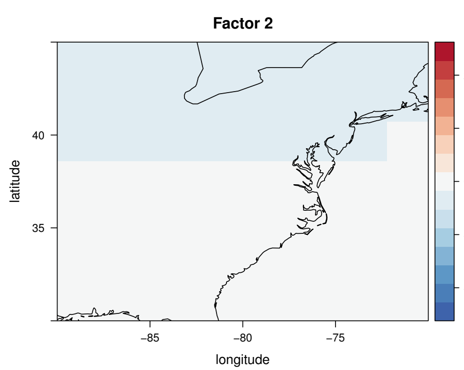
Factor model for temperature anomalies

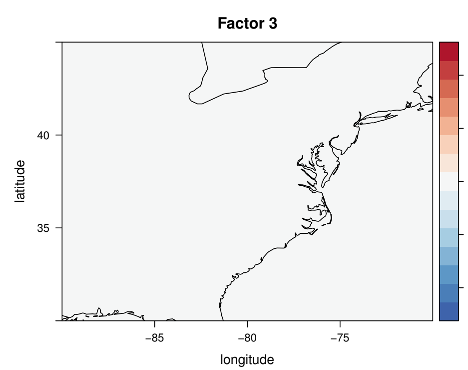
$$\mathbf{y}_t = \mathbf{x}_{t1} b_{.1} + \mathbf{x}_{t2} b_{.2} + \mathbf{x}_{t3} b_{.3} + \nu_t$$

Factor contributions in 1948

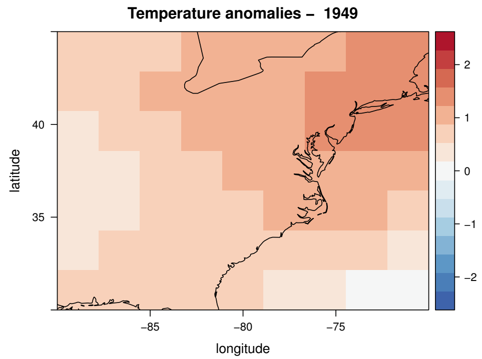


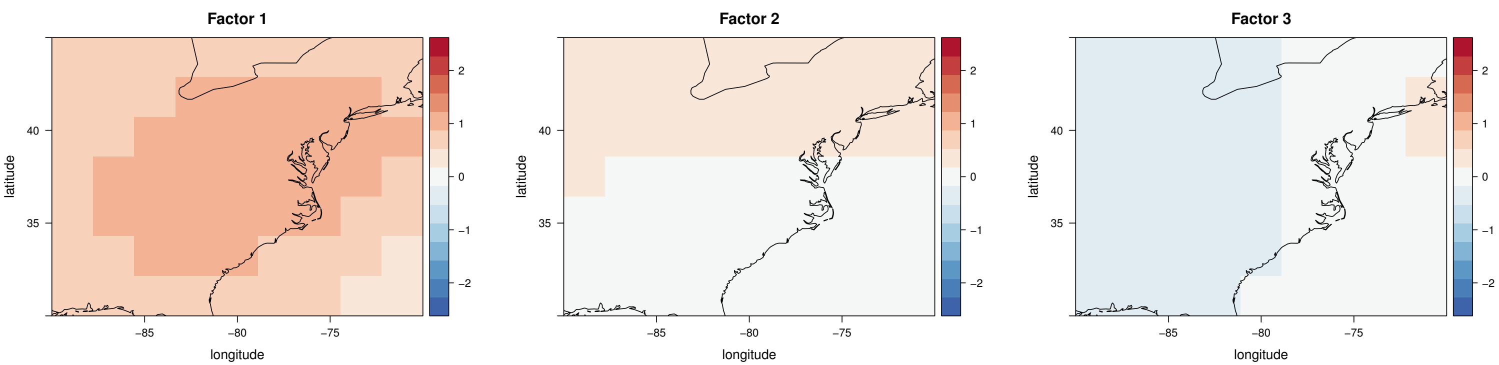




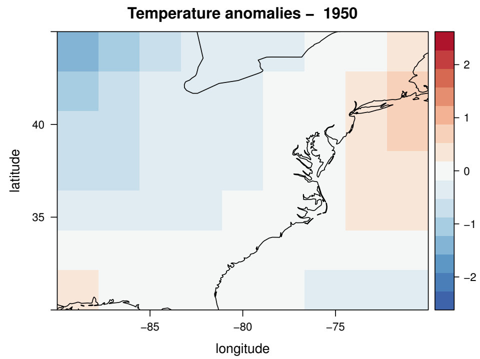


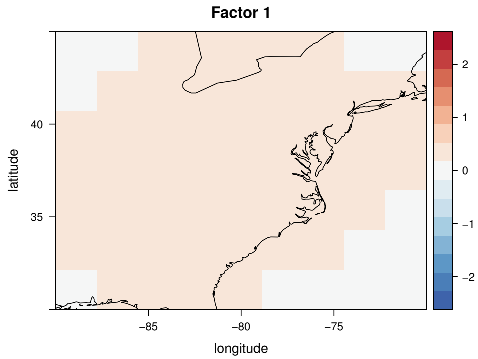
Factor contributions in 1949

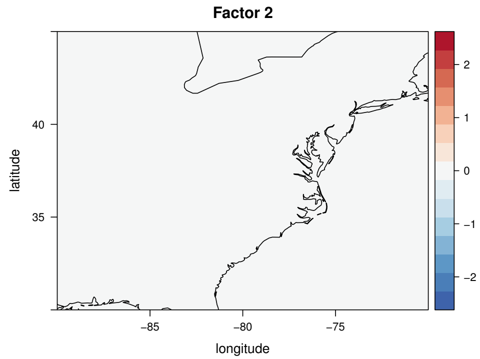


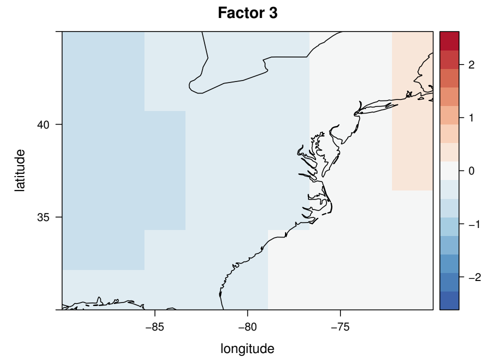


Factor contributions in 1950

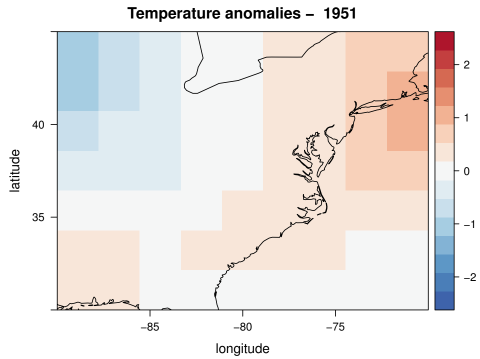


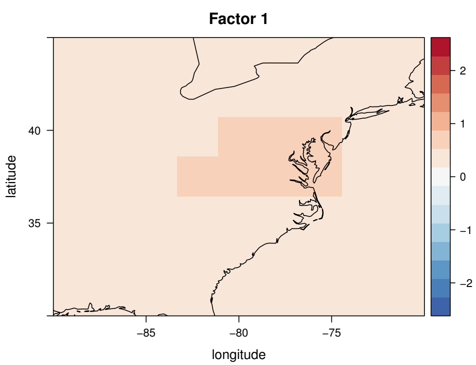


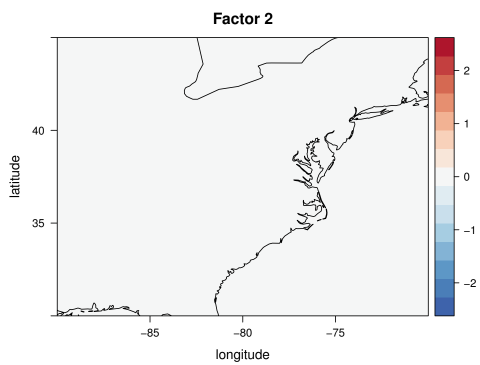


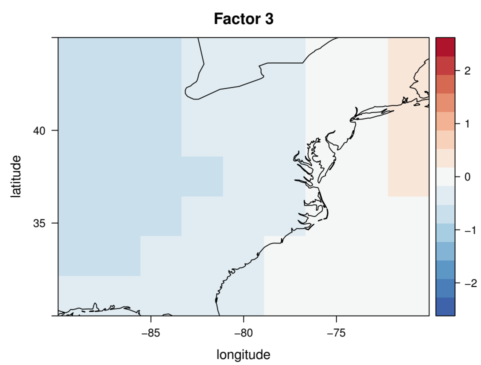


Factor contributions in 1951

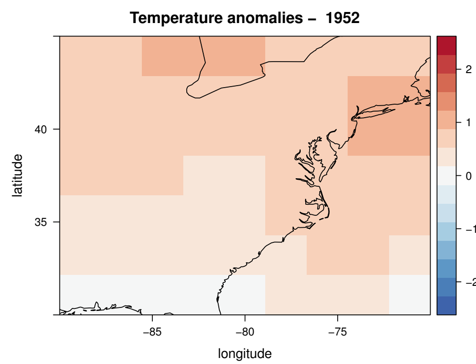


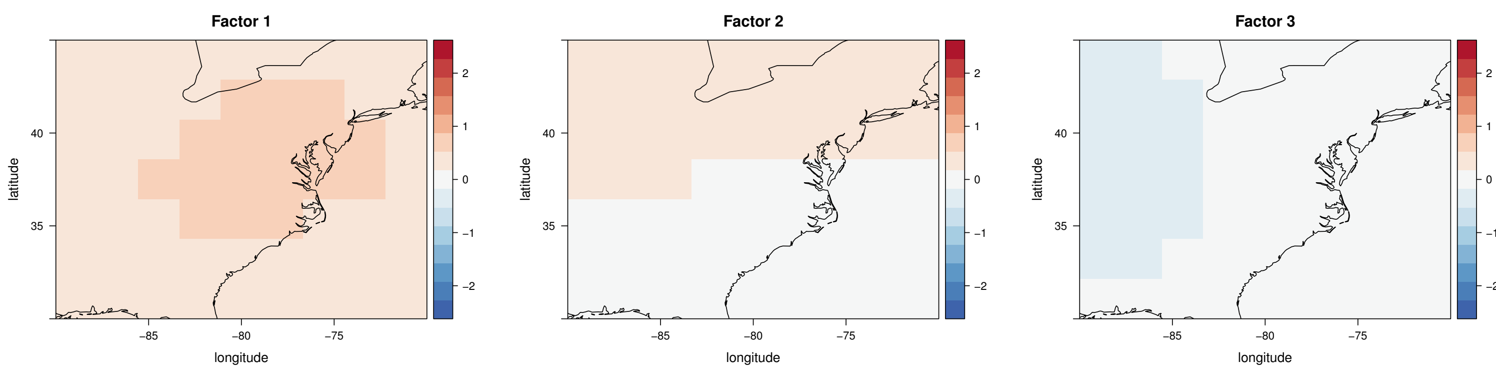




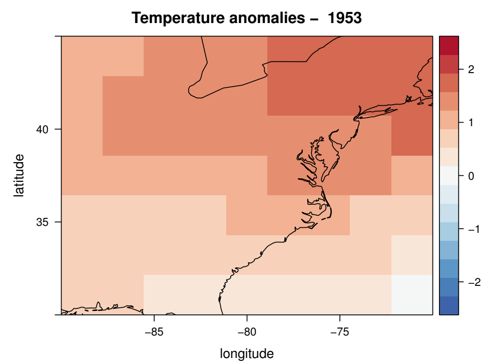


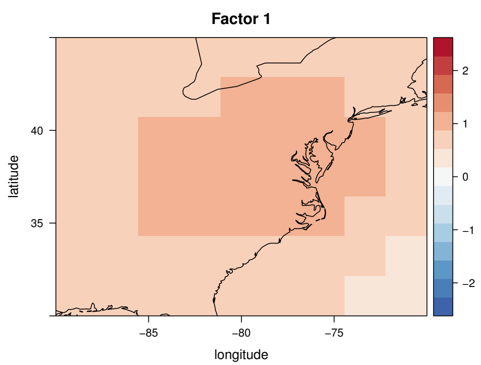
Factor contributions in 1952

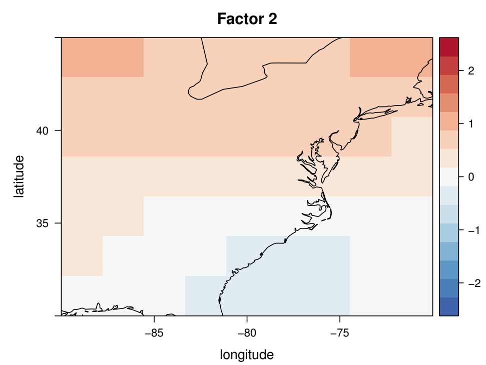


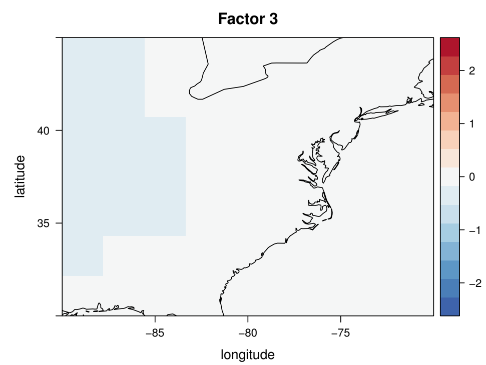


Factor contributions in 1953

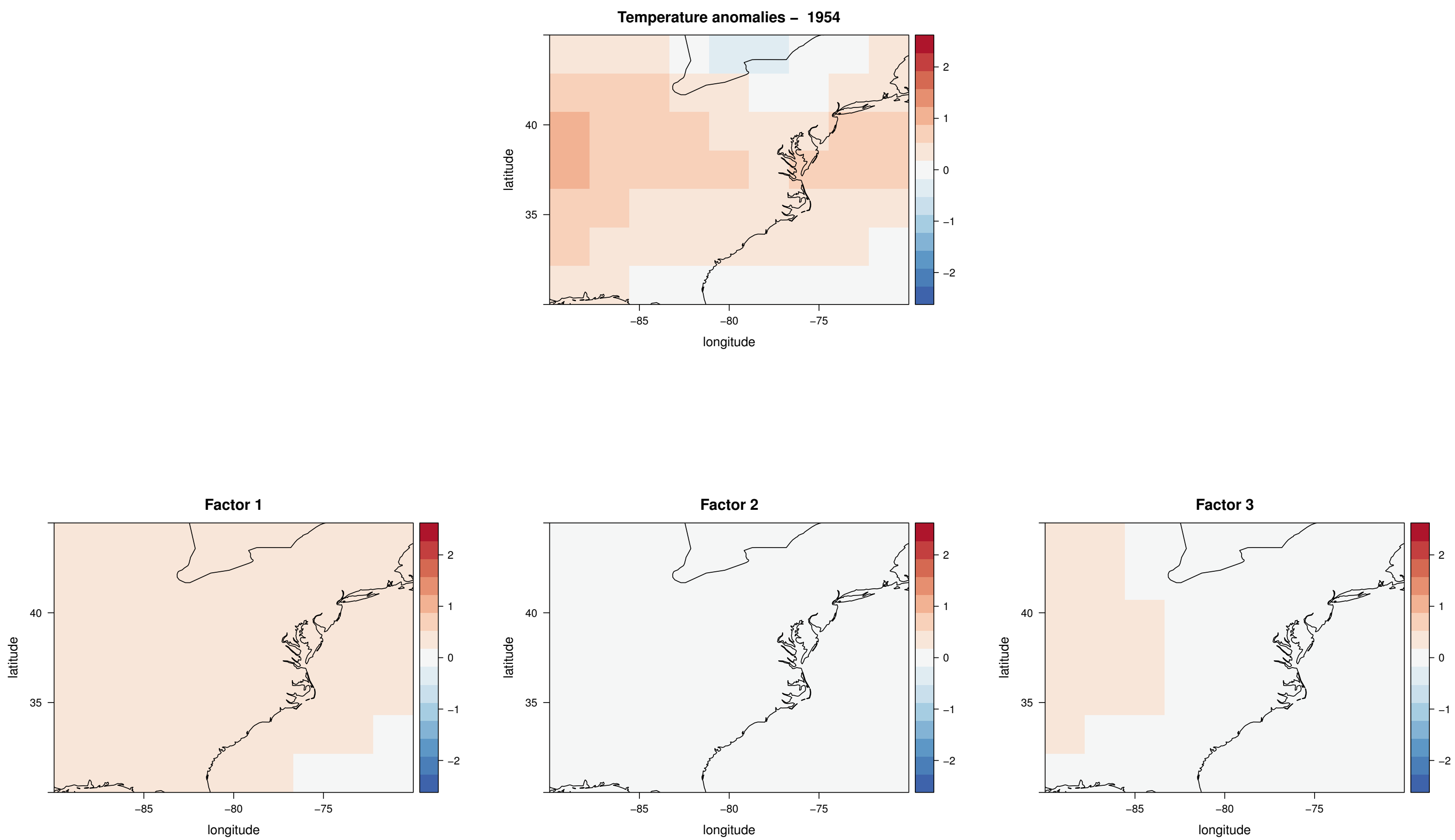




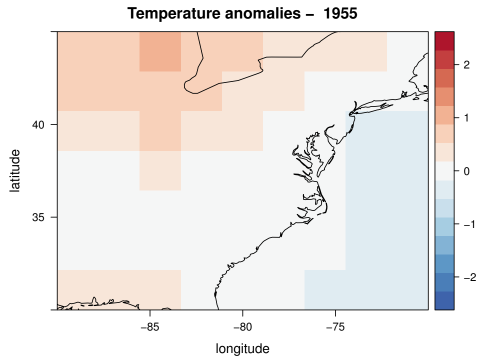


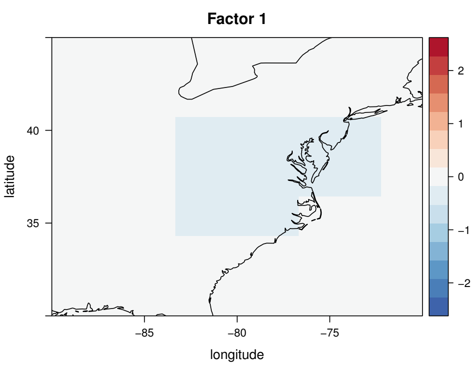


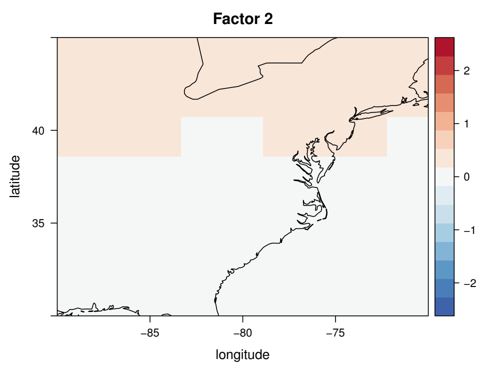
Factor contributions in 1954

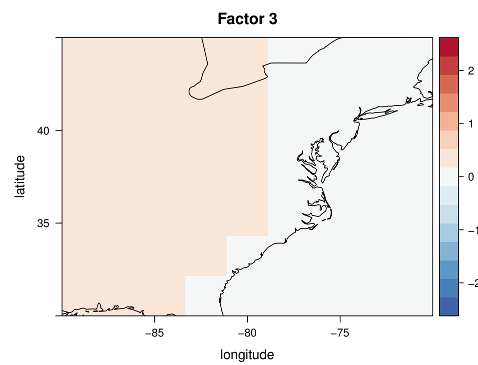


Factor contributions in 1955

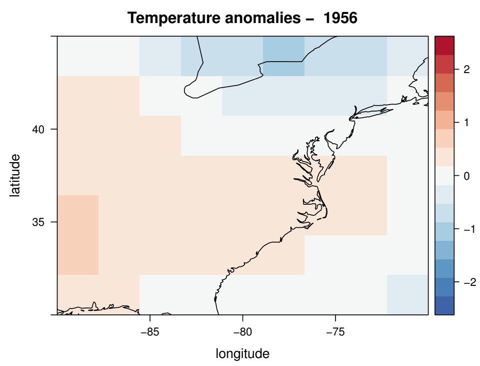


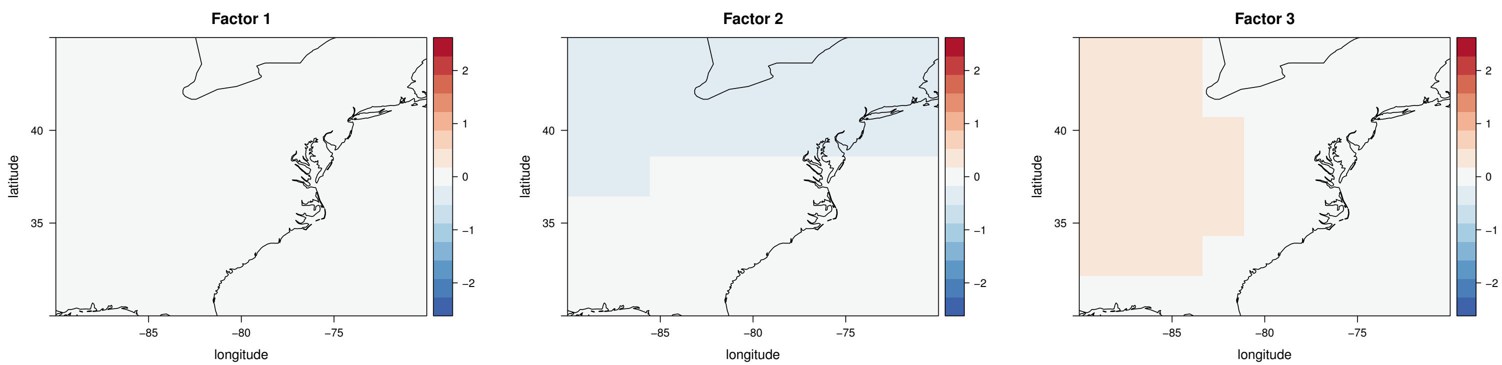




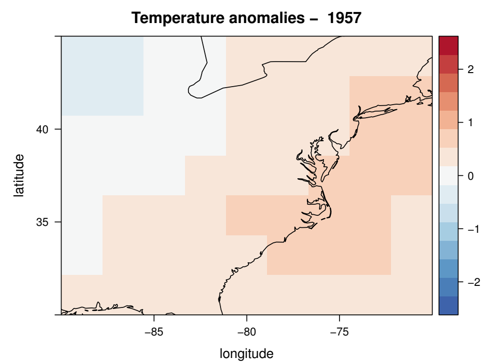


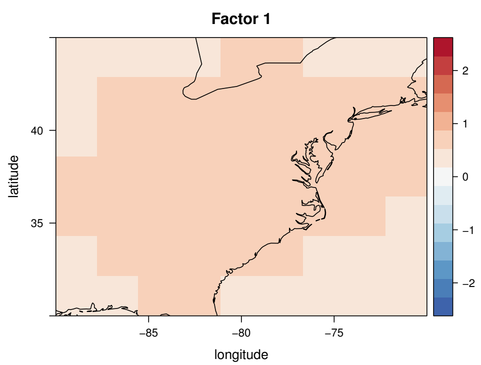
Factor contributions in 1956

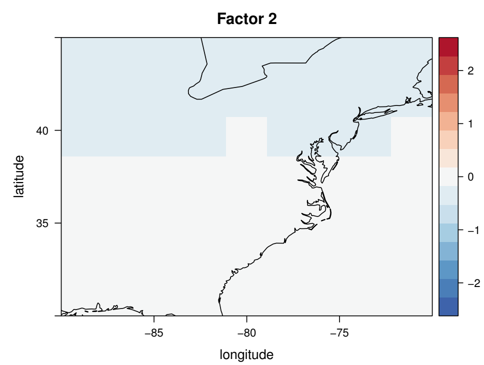


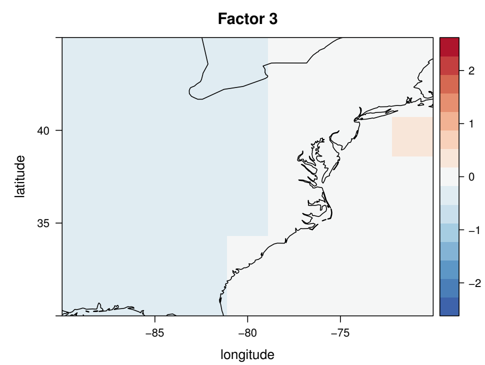


Factor contributions in 1957

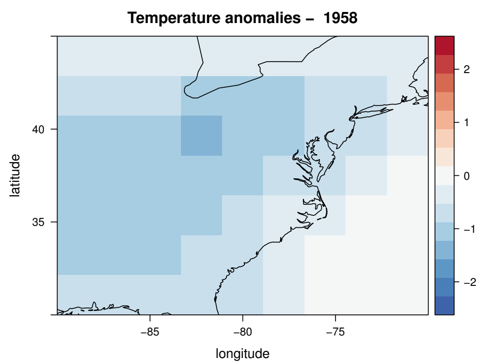


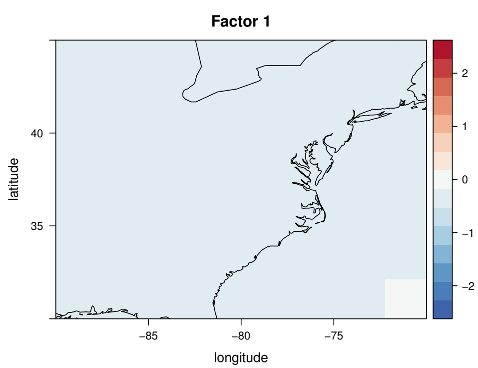


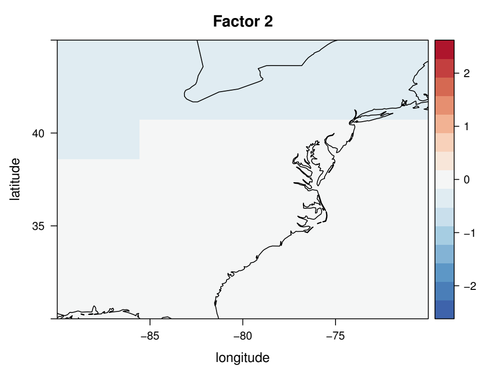


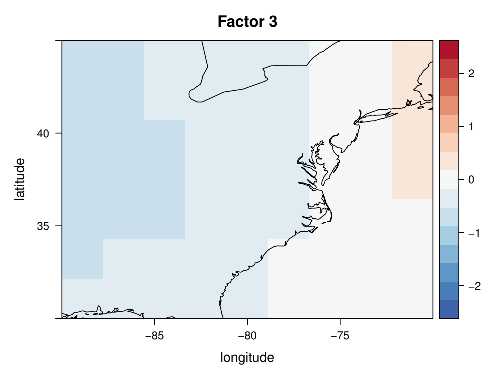


Factor contributions in 1958

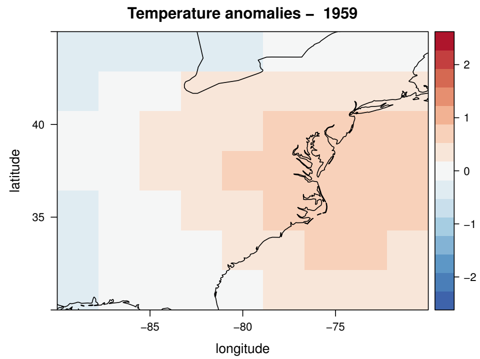


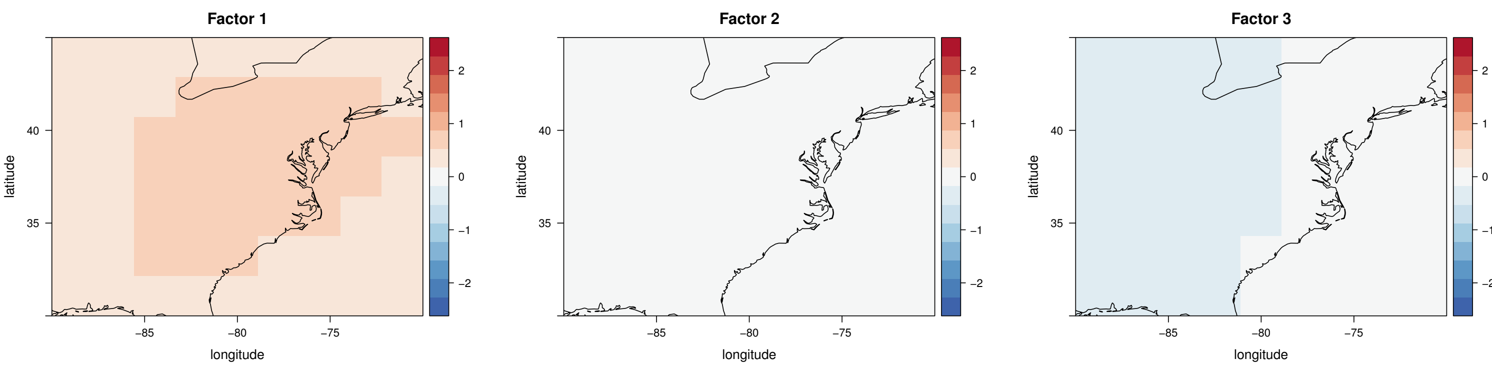




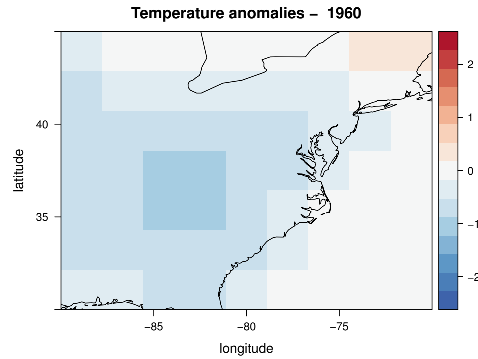


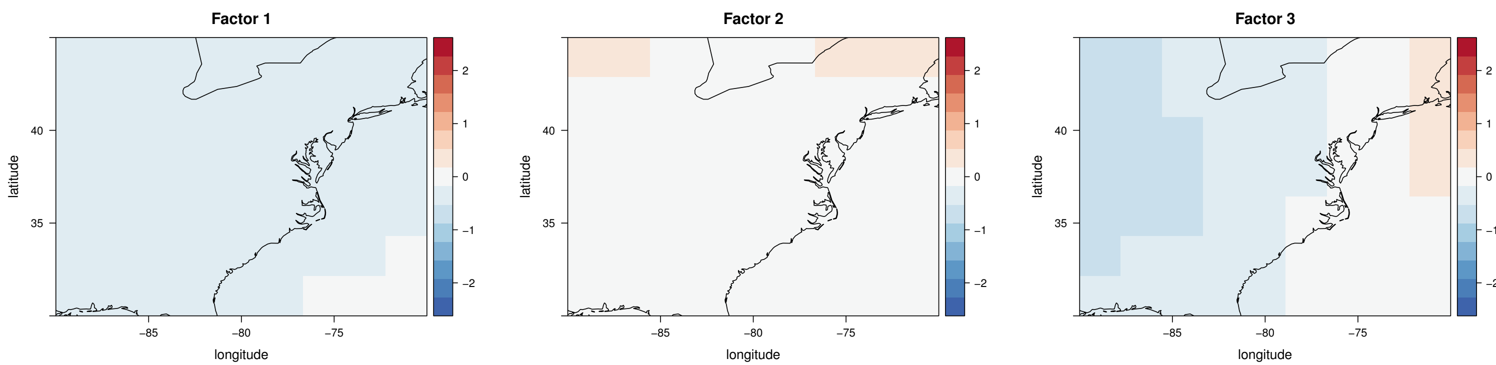
Factor contributions in 1959



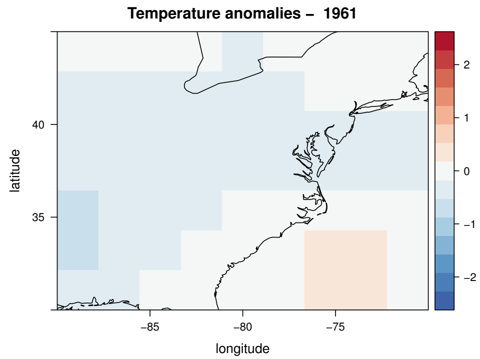


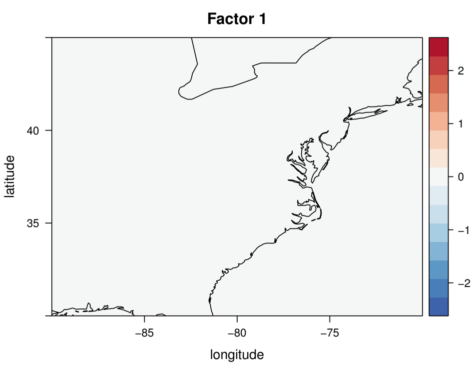
Factor contributions in 1960

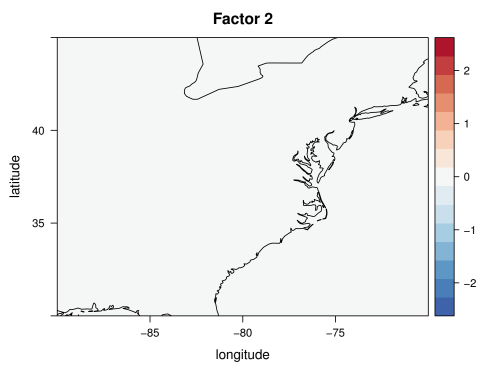


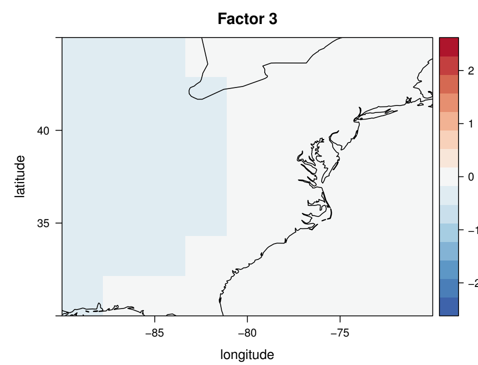


Factor contributions in 1961

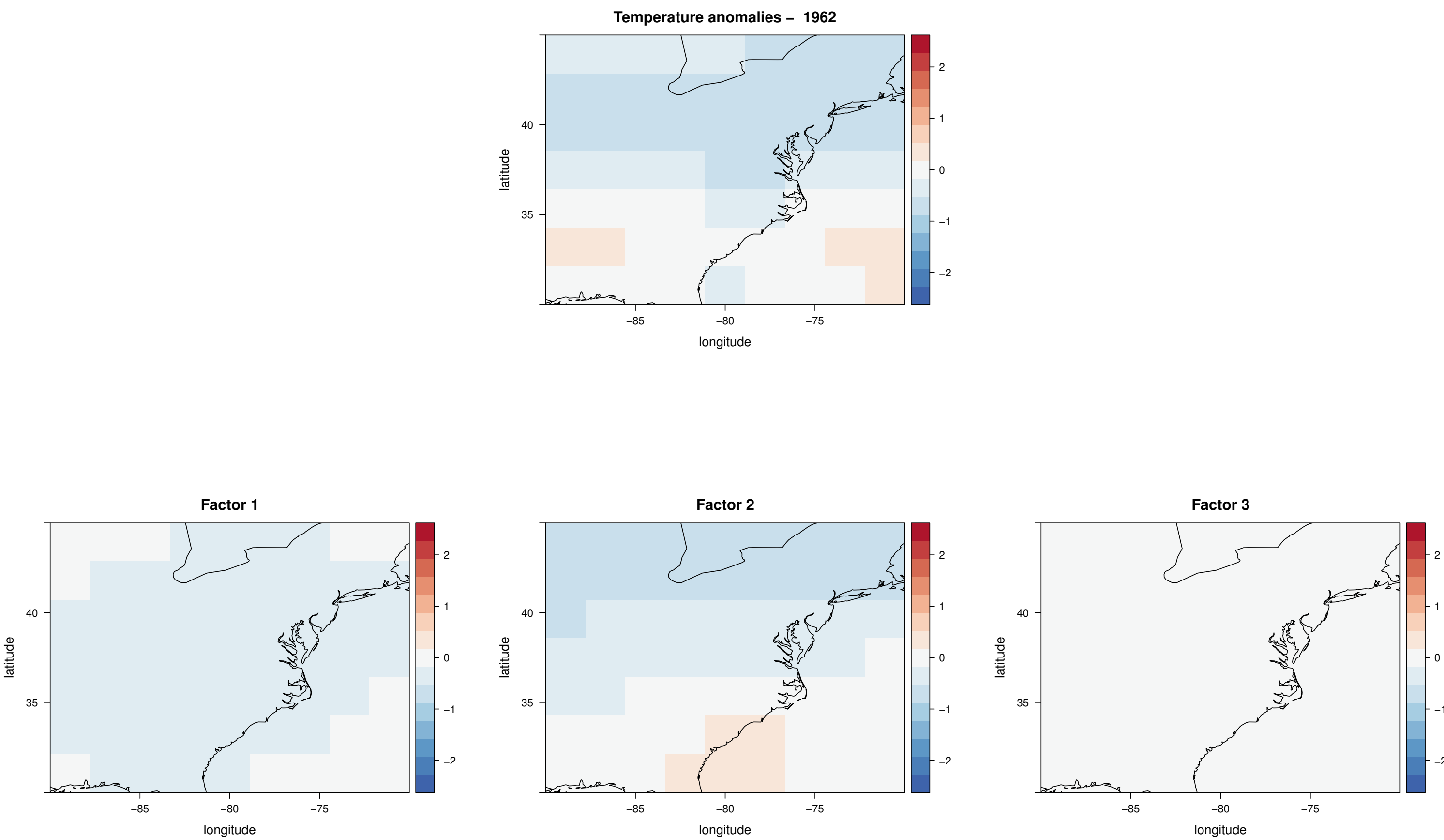




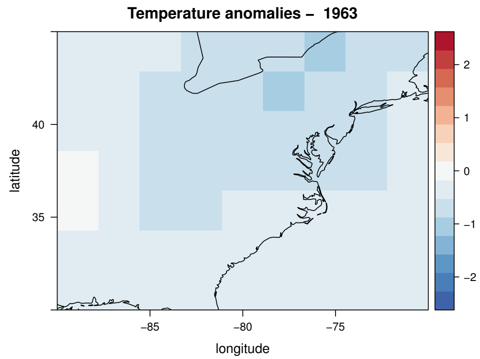


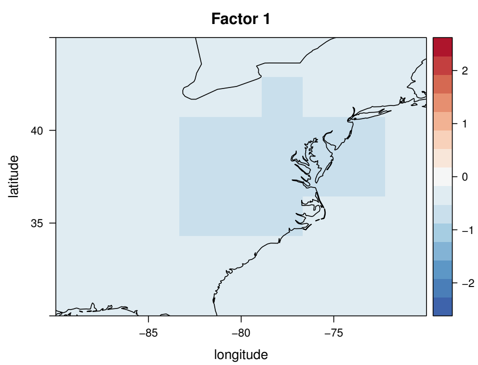


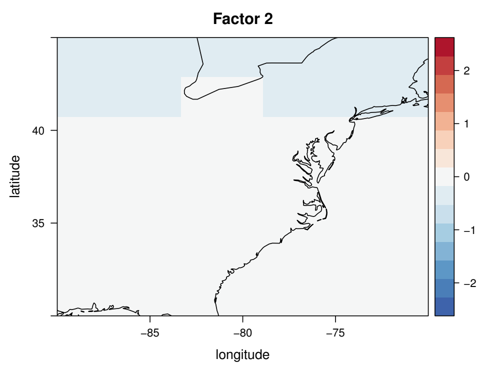
Factor contributions in 1962

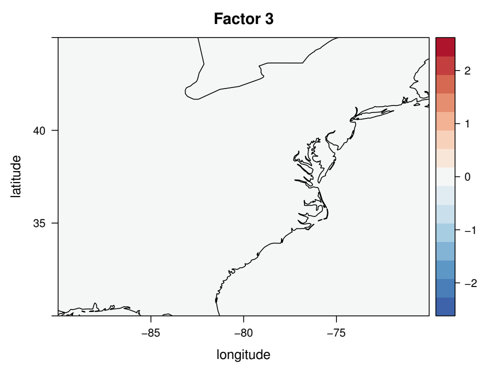


Factor contributions in 1963

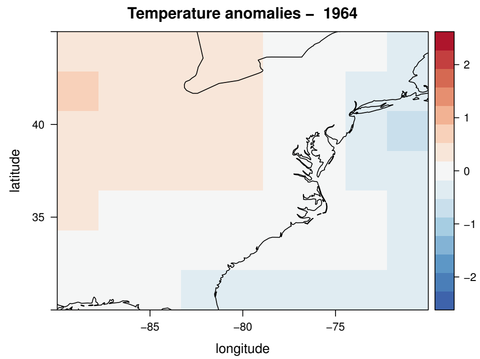


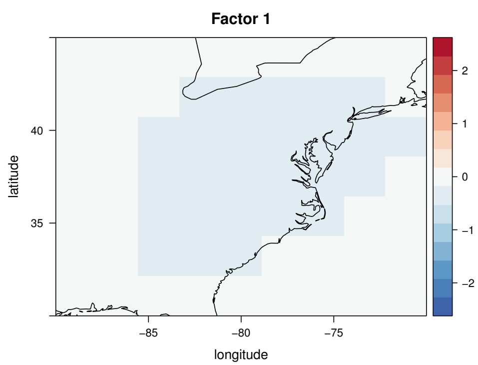


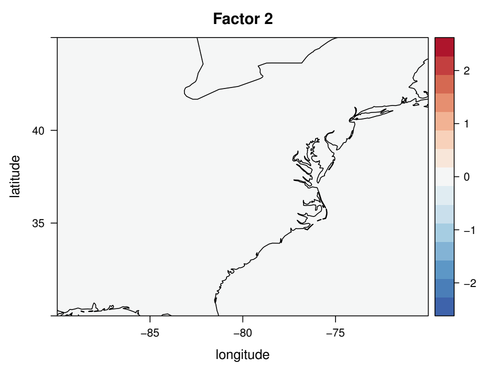


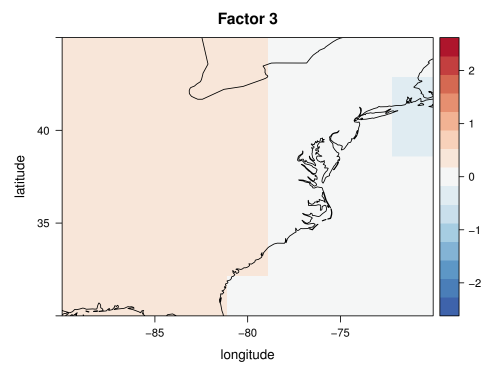


Factor contributions in 1964

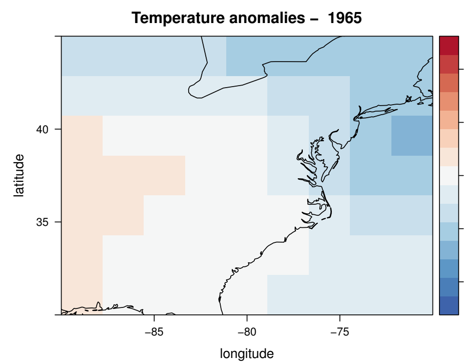


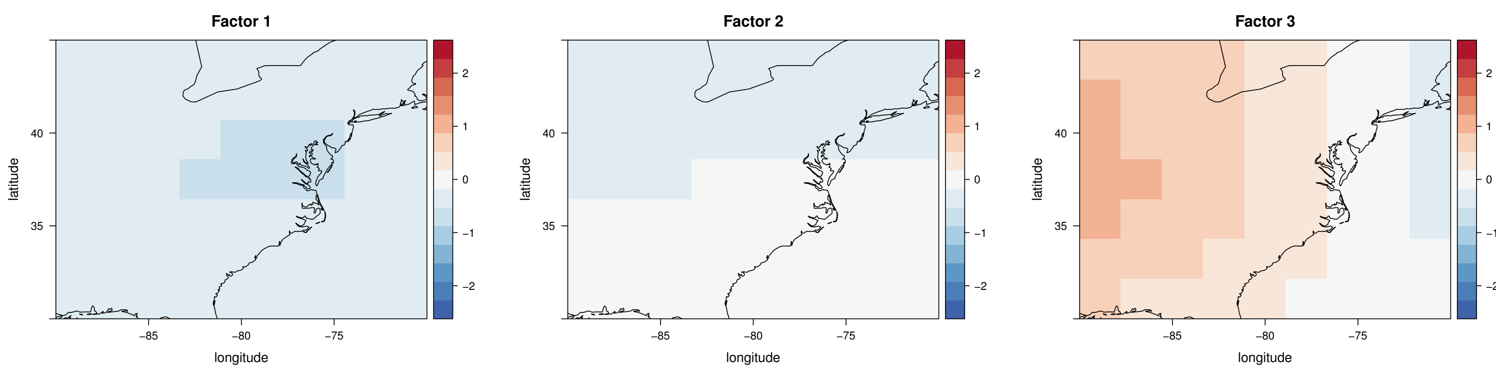




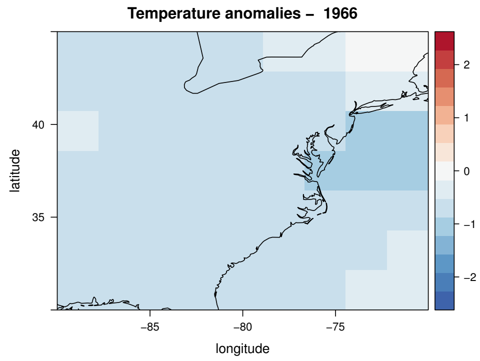


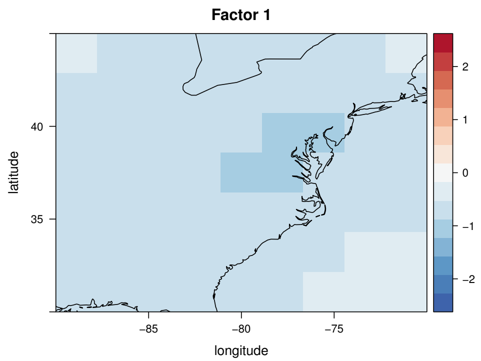
Factor contributions in 1965

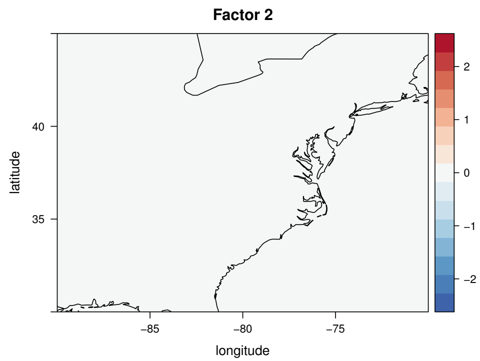


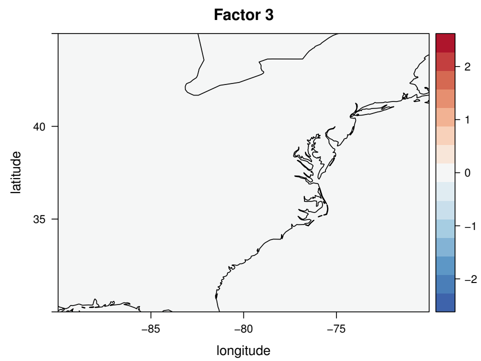


Factor contributions in 1966

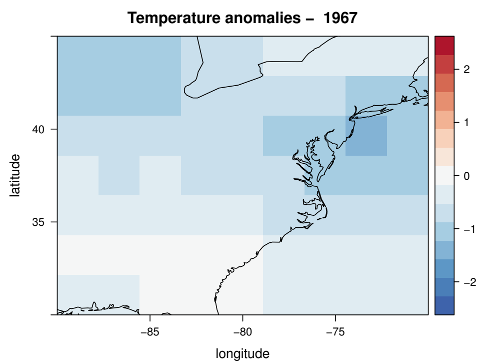


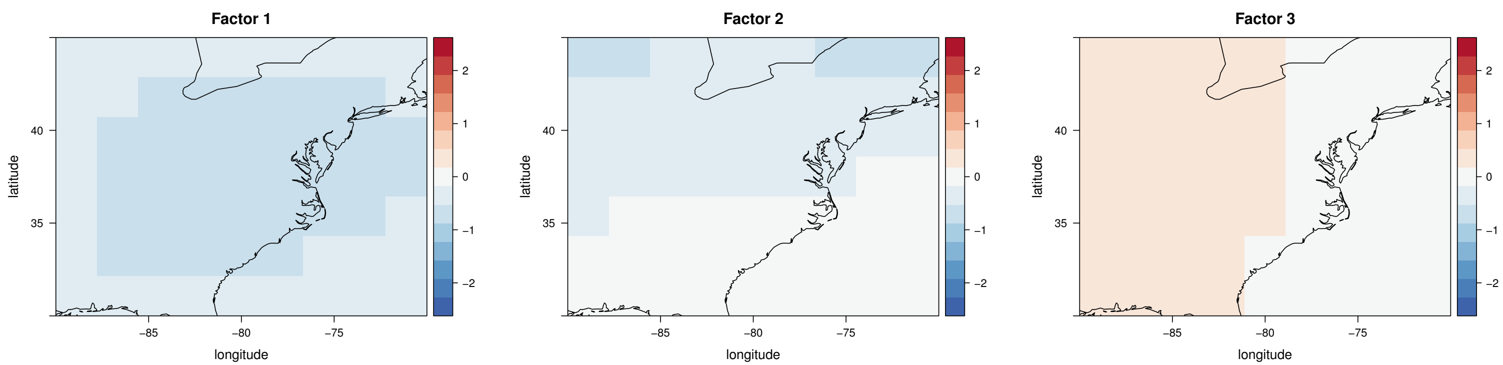




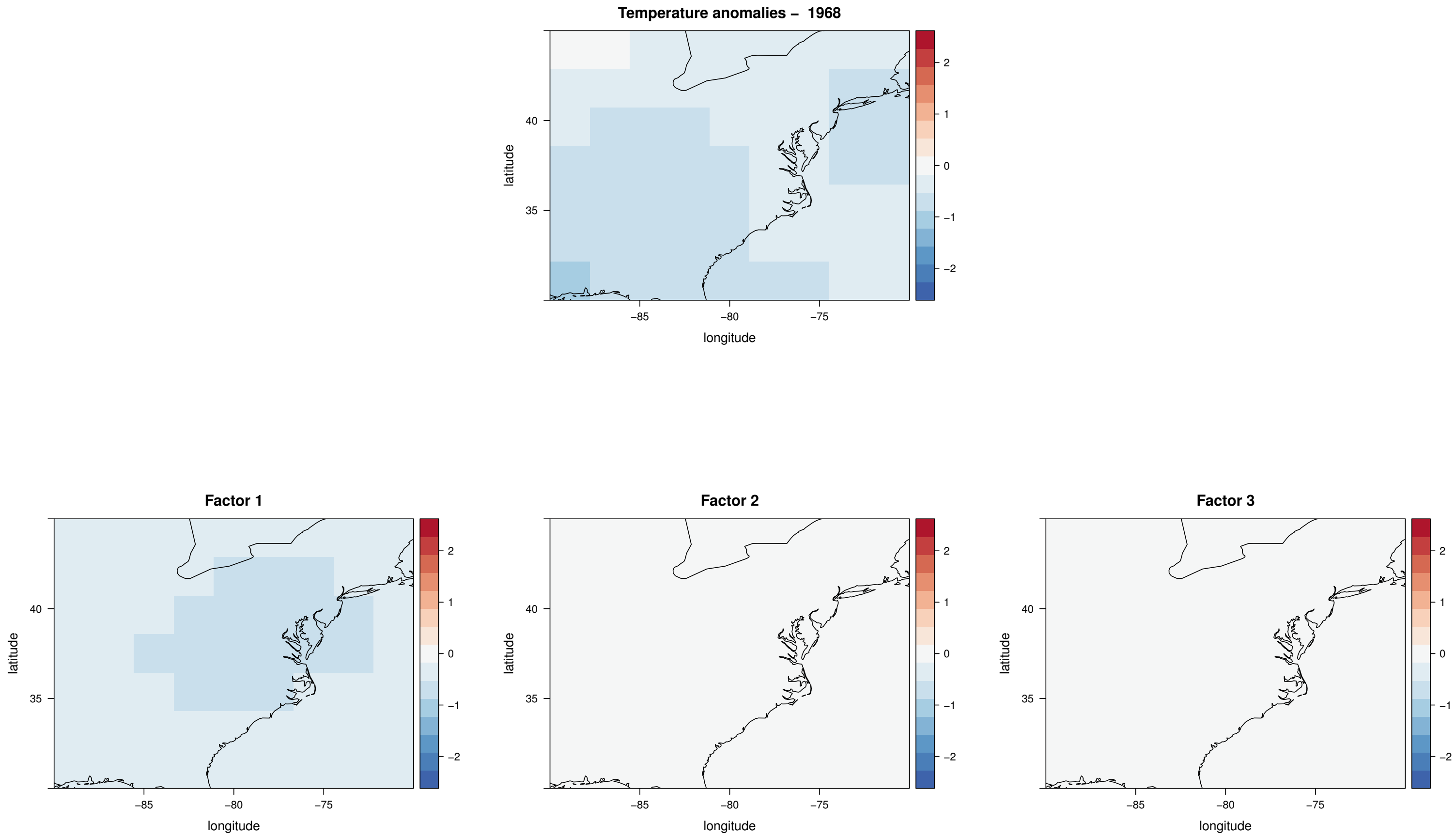


Factor contributions in 1967

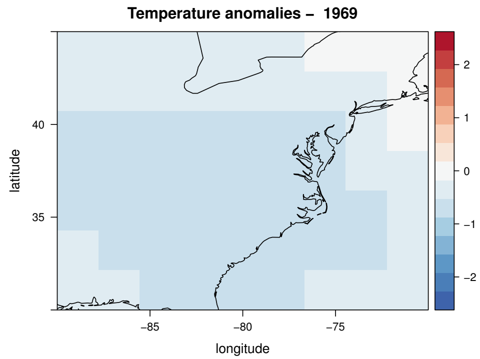


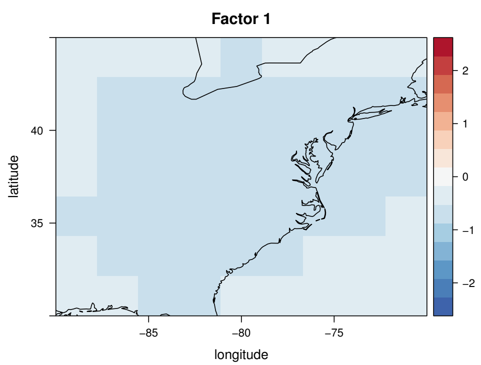


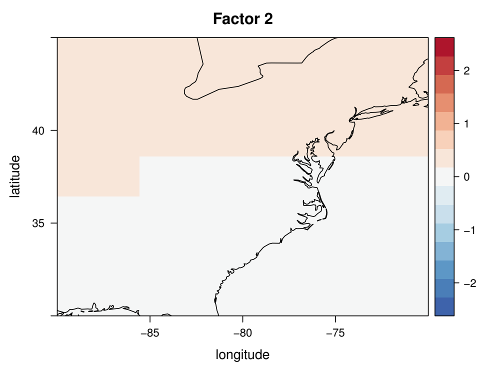
Factor contributions in 1968

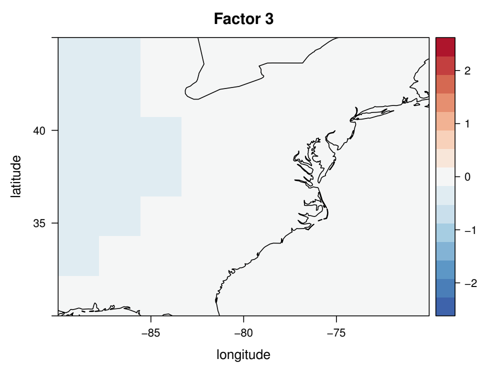


Factor contributions in 1969

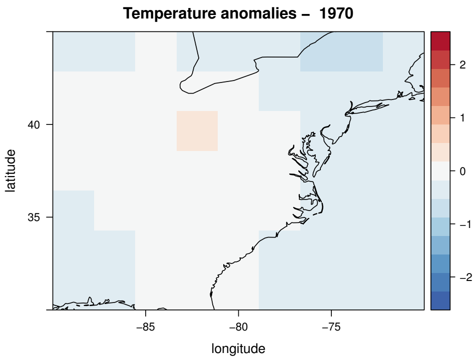


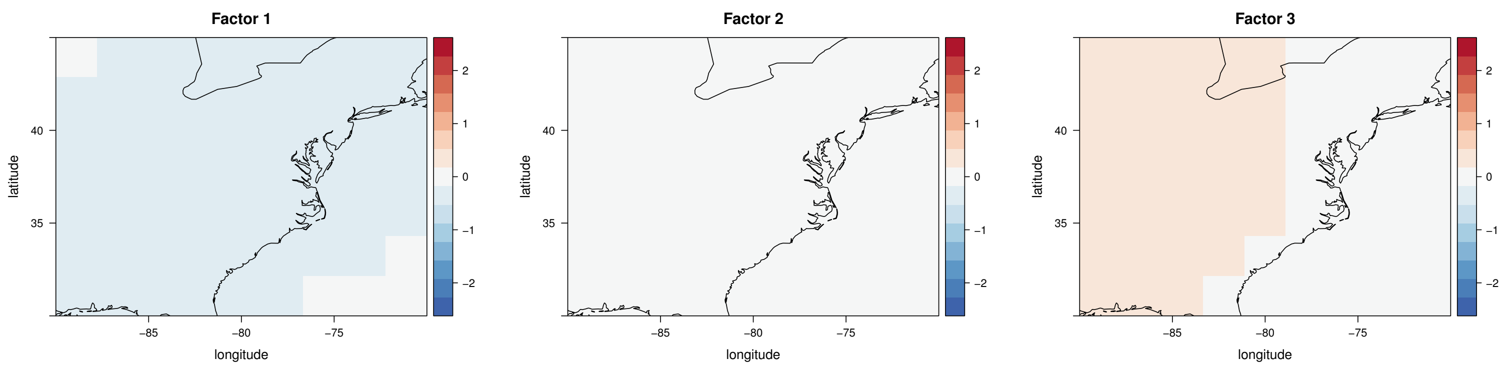




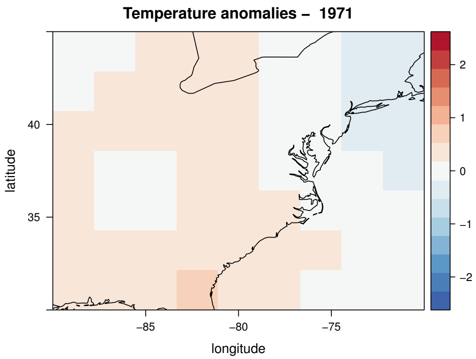


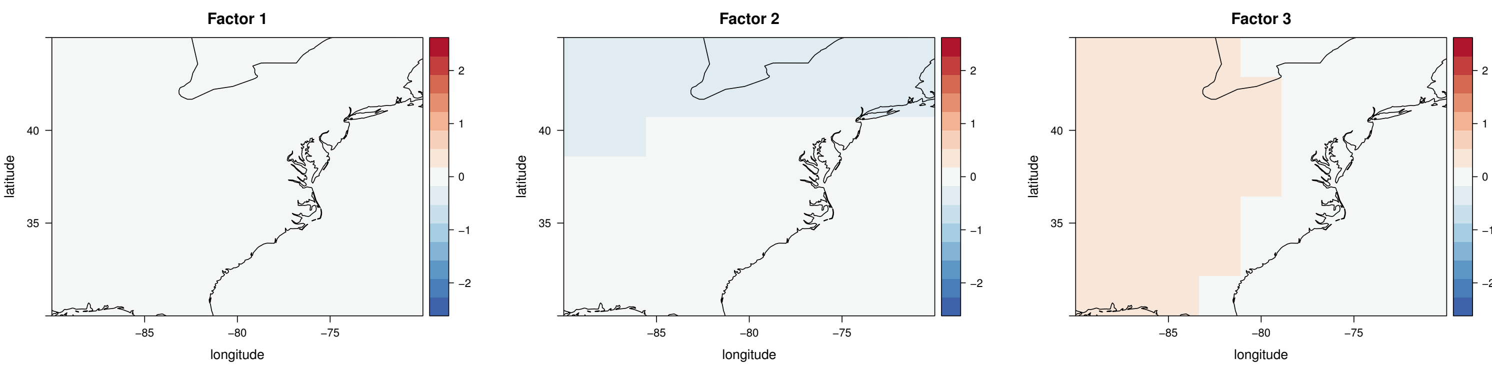
Factor contributions in 1970



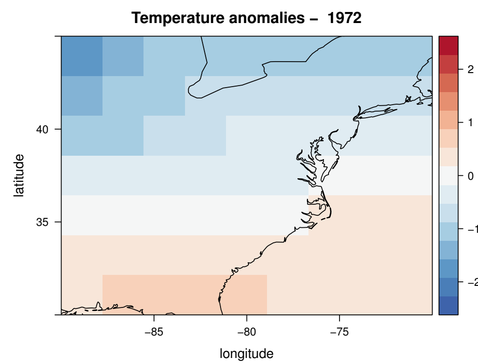


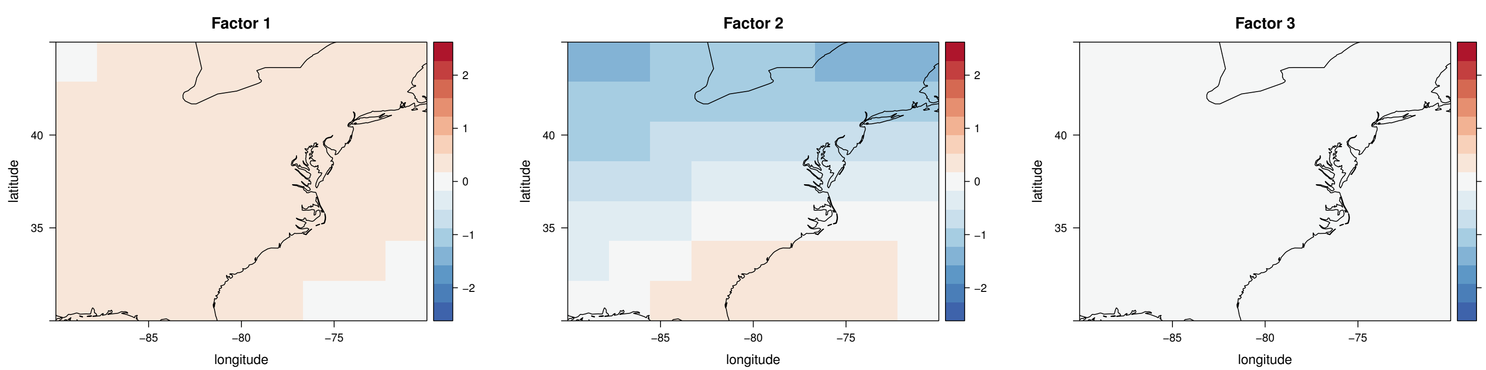
Factor contributions in 1971



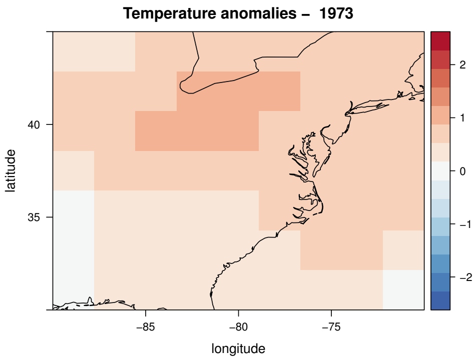


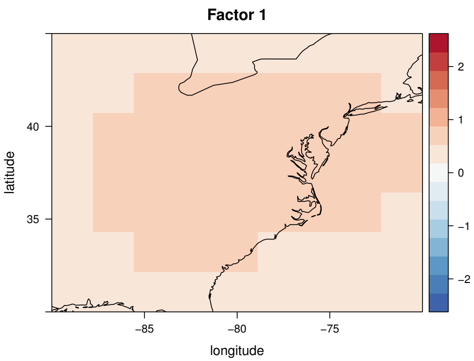
Factor contributions in 1972

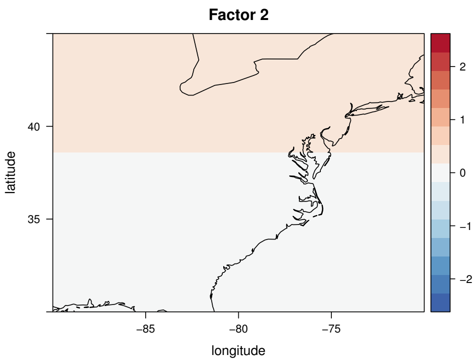


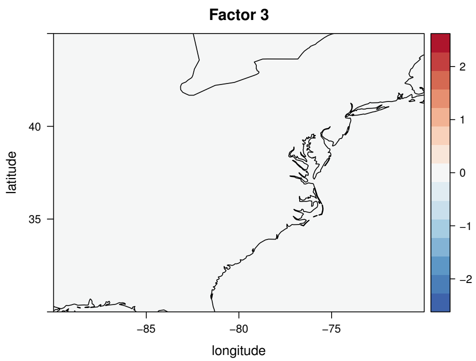


Factor contributions in 1973

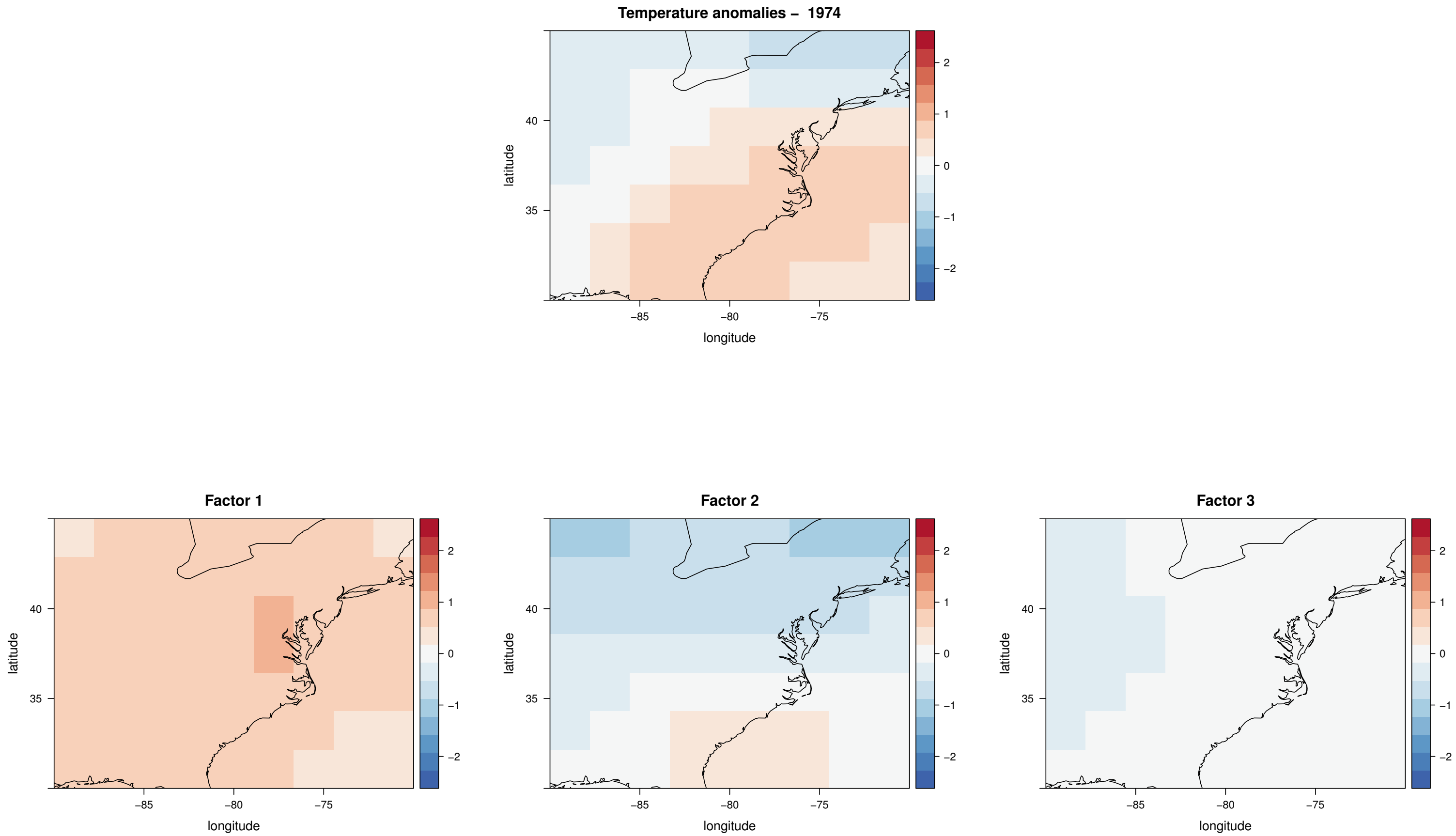




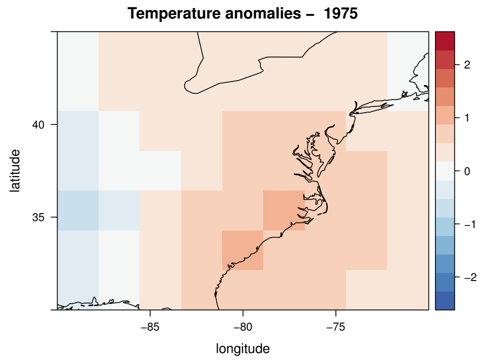


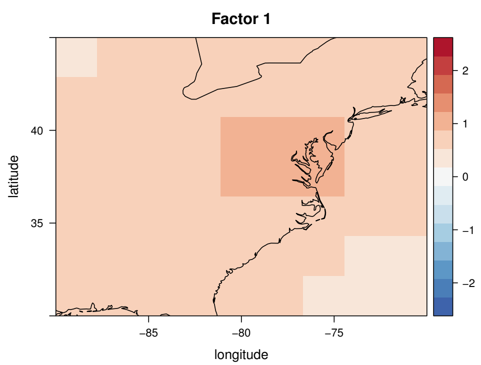


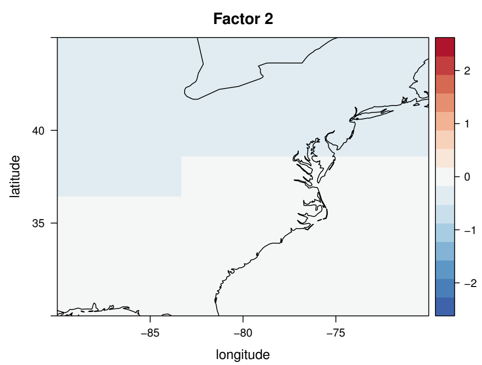
Factor contributions in 1974

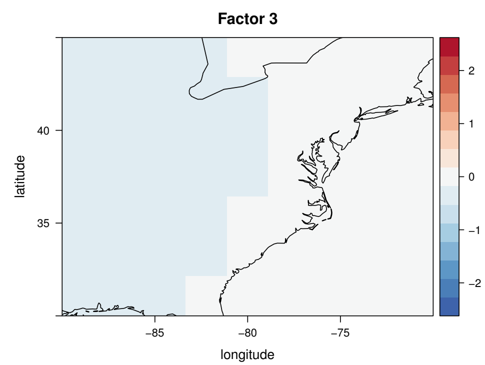


Factor contributions in 1975

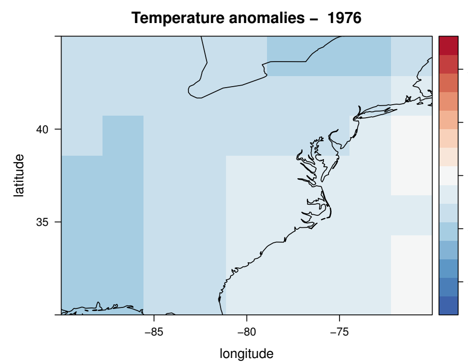


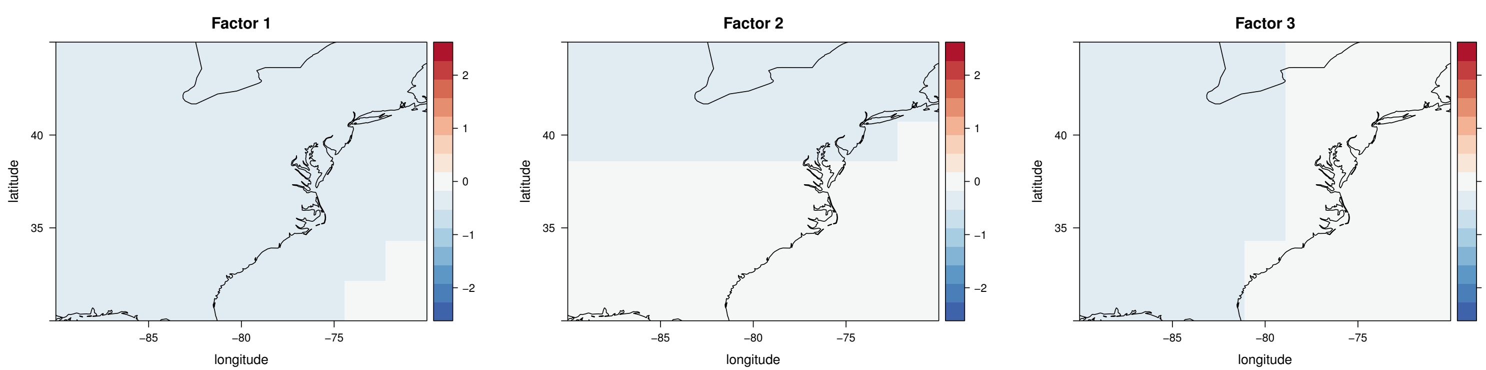




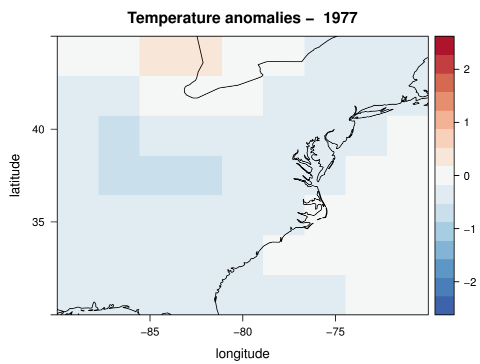


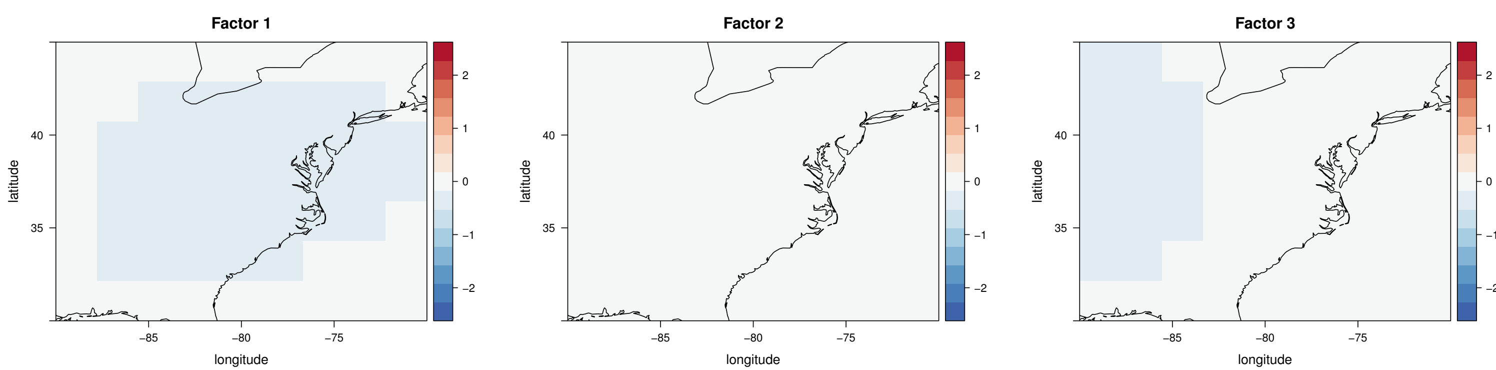
Factor contributions in 1976



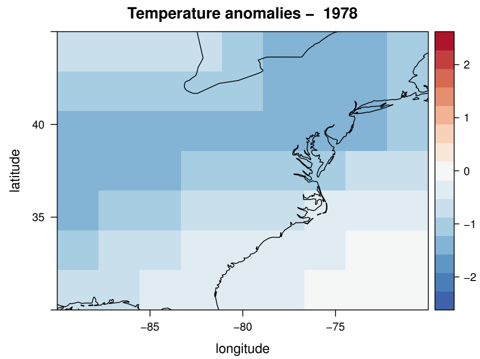


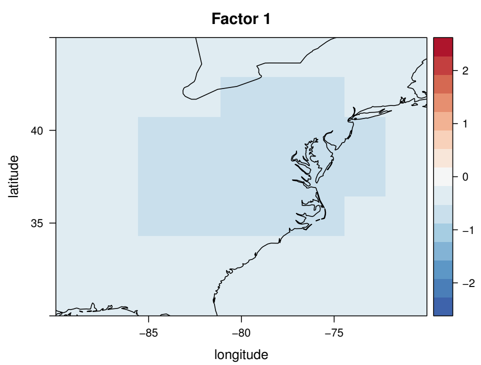
Factor contributions in 1977

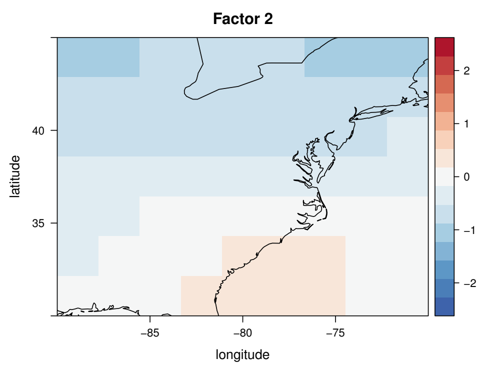


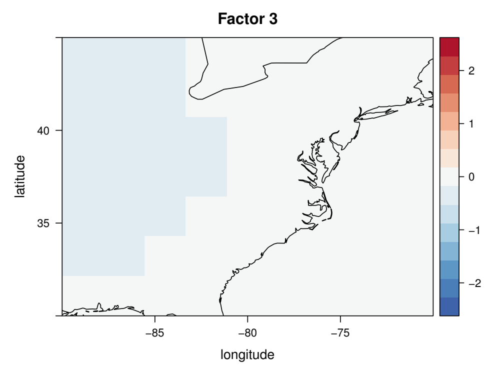


Factor contributions in 1978

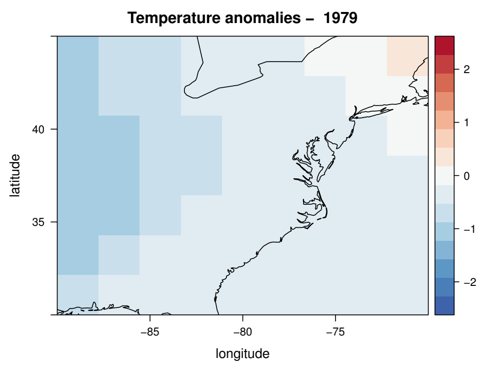


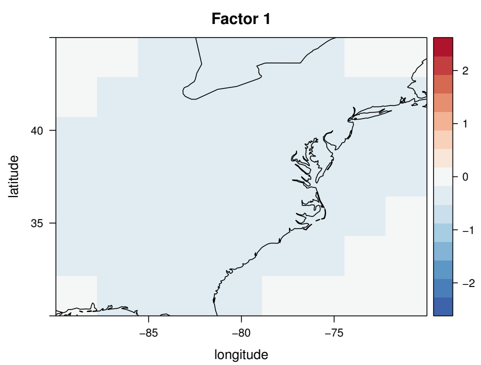


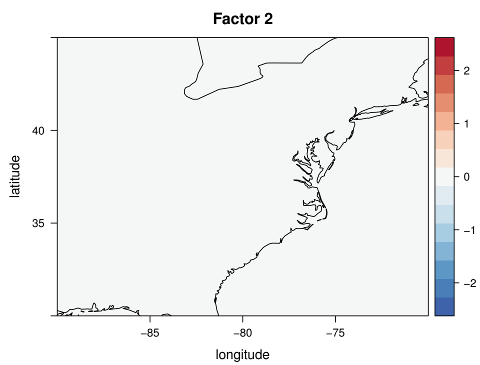


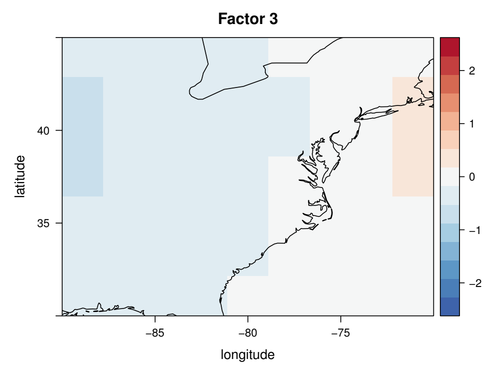


Factor contributions in 1979

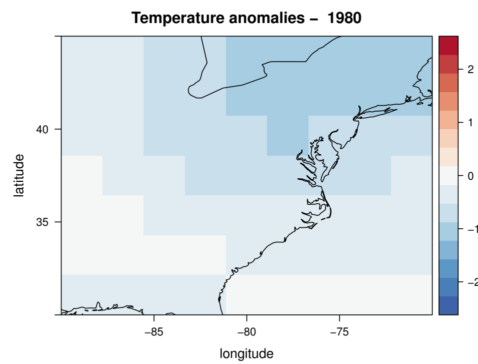


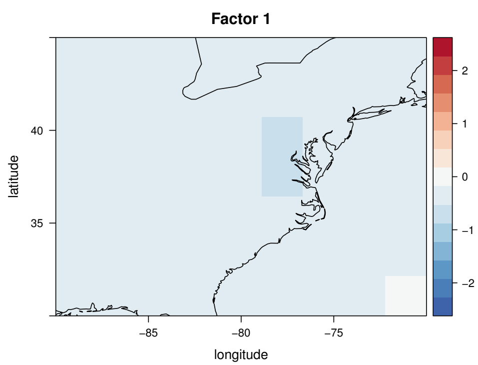


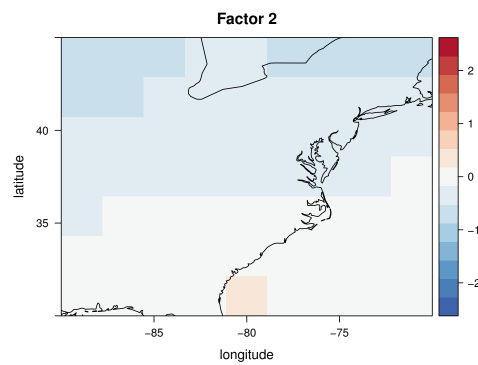


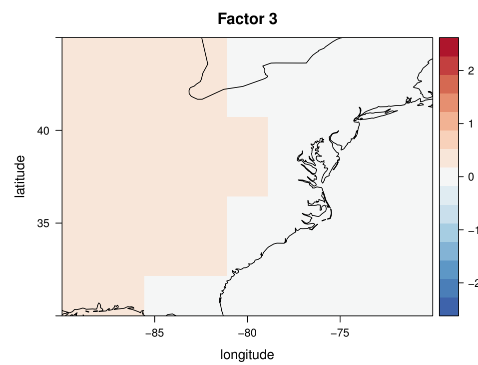


Factor contributions in 1980

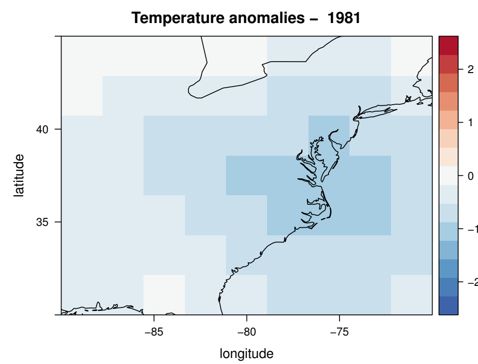


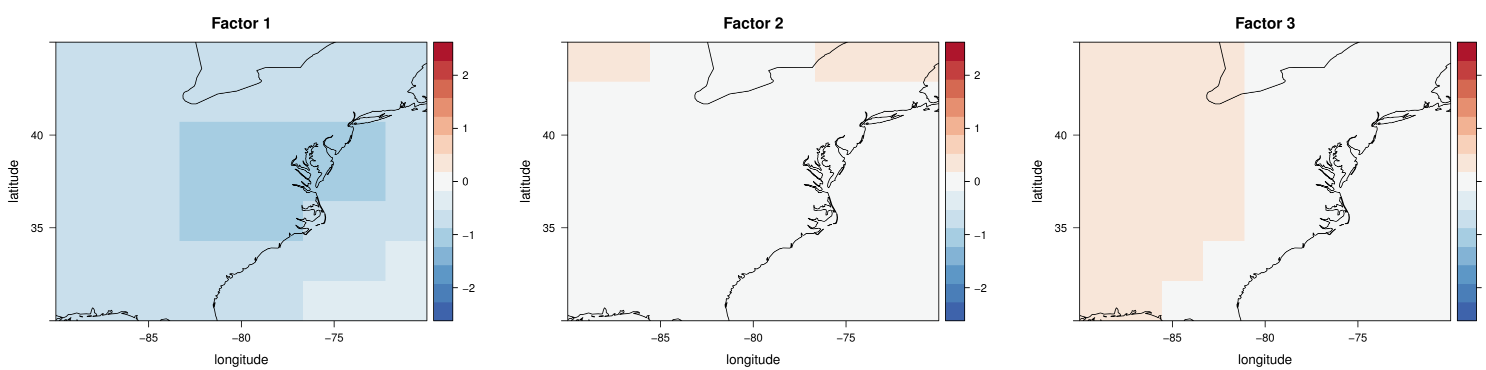




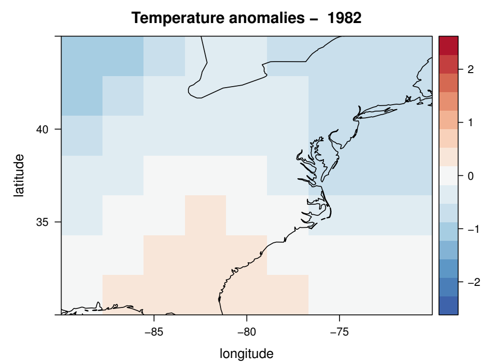


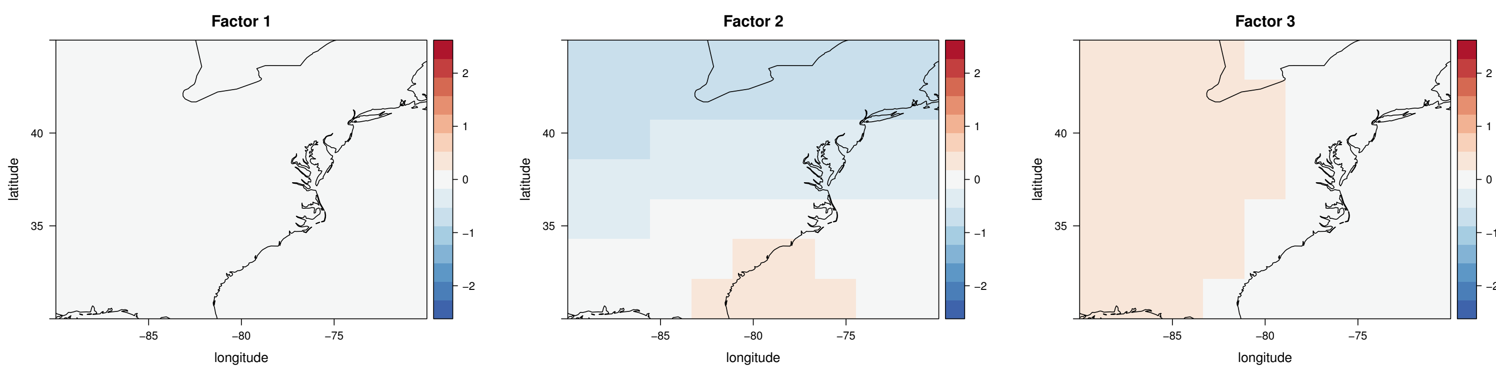
Factor contributions in 1981



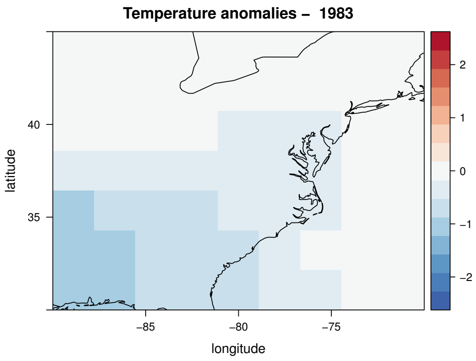


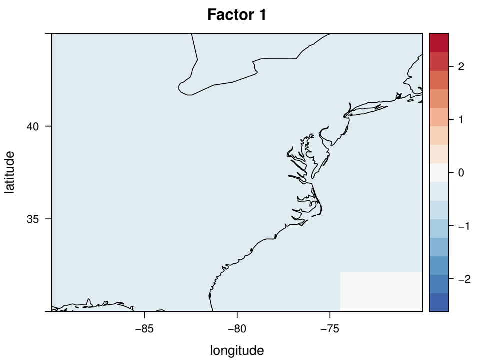
Factor contributions in 1982

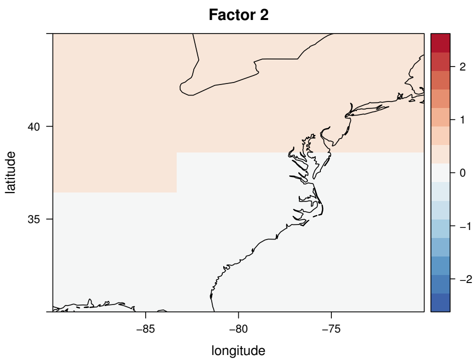


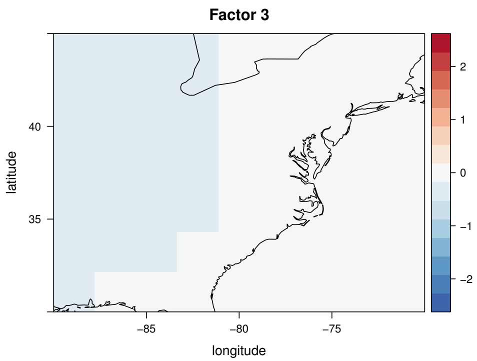


Factor contributions in 1983

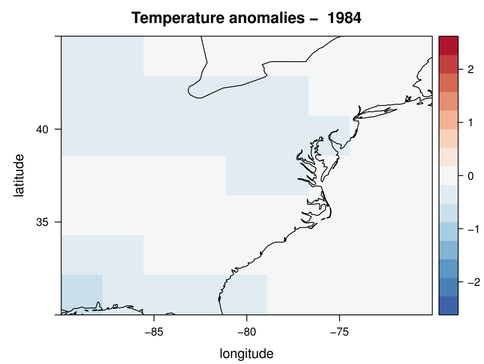


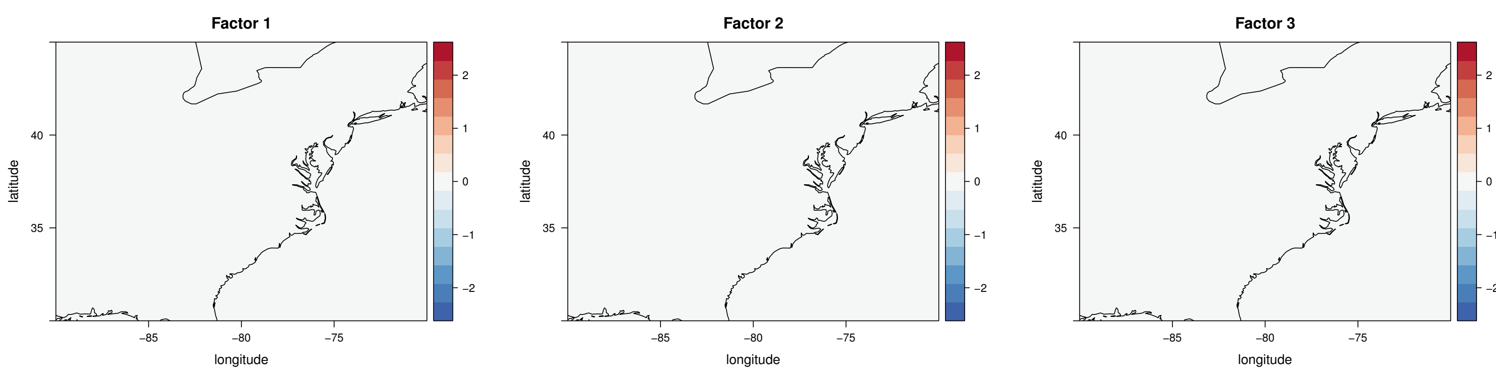




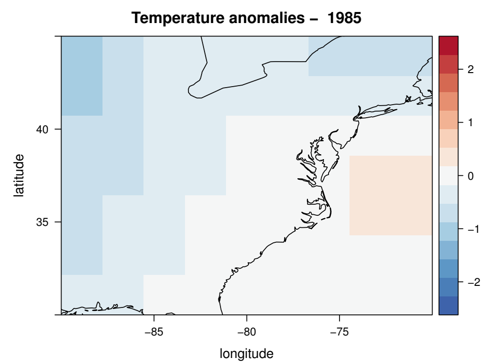


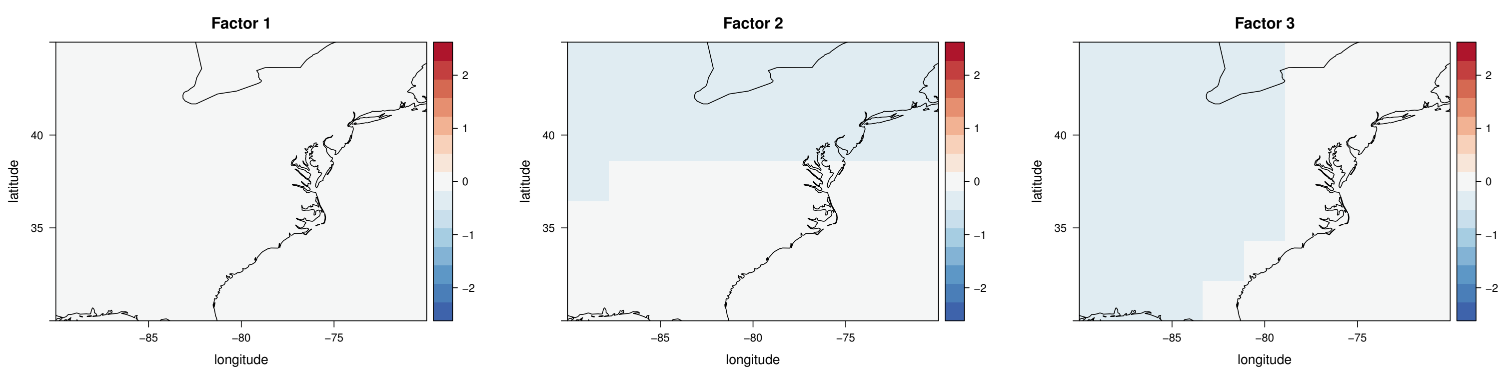
Factor contributions in 1984



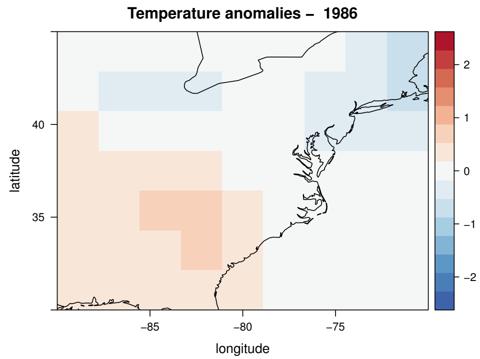


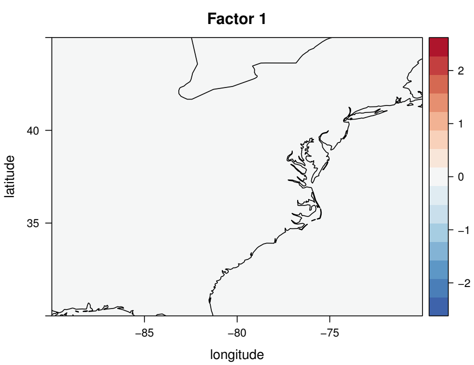
Factor contributions in 1985

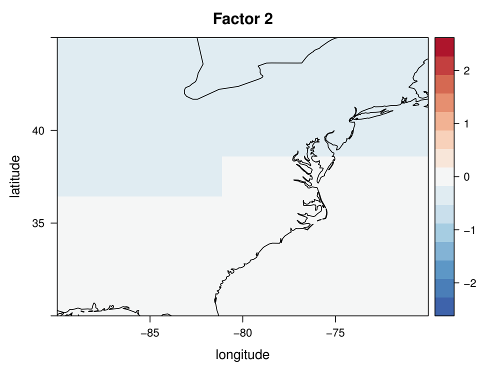


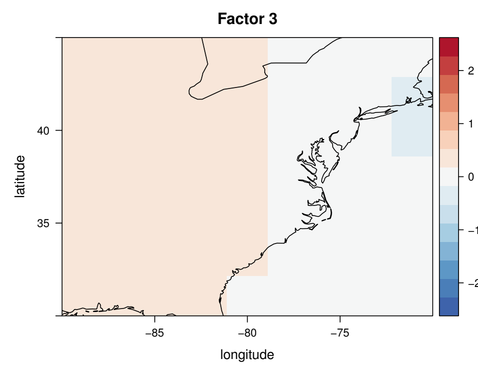


Factor contributions in 1986

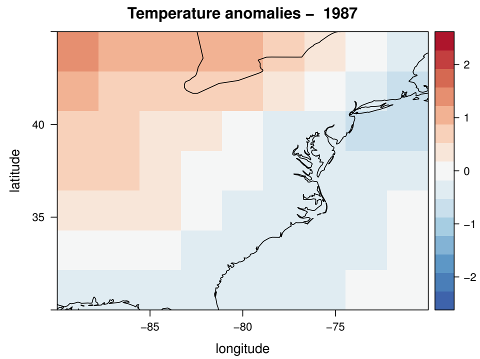


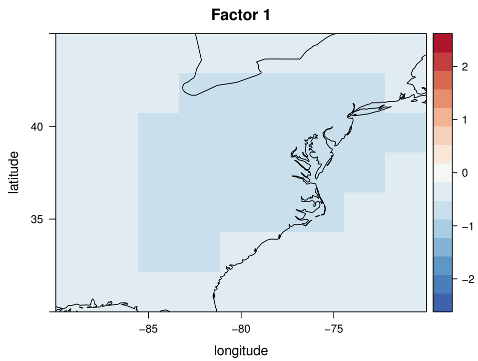


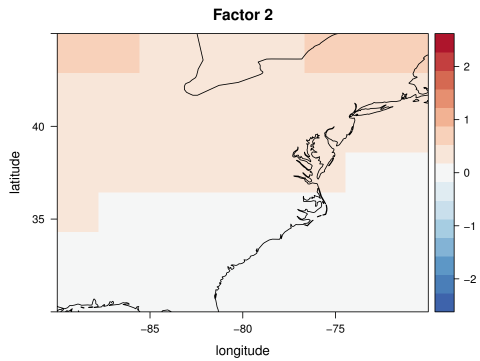


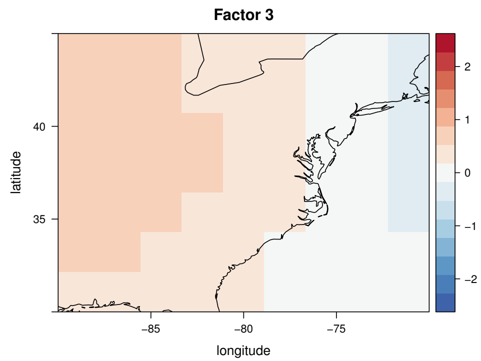


Factor contributions in 1987

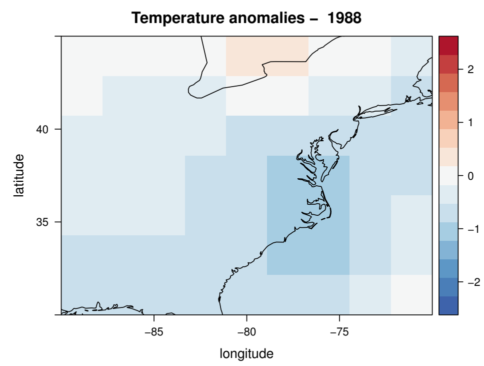


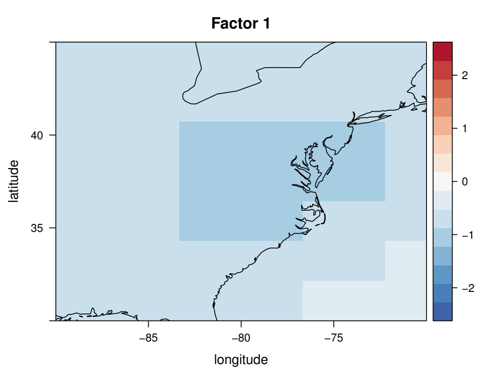


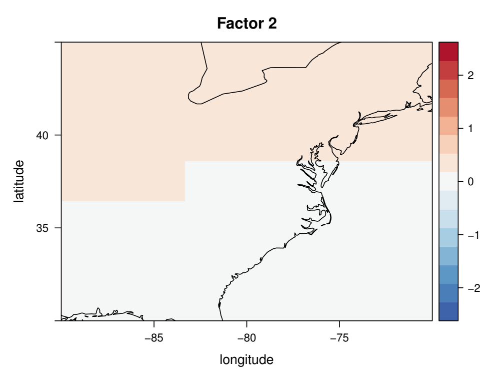


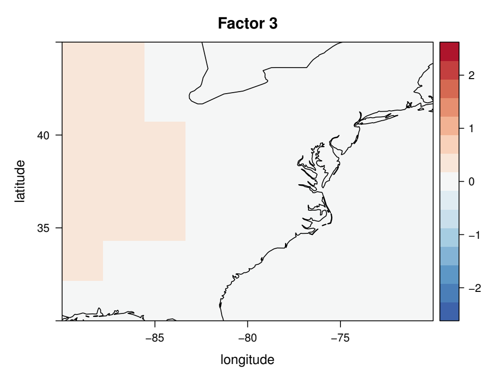


Factor contributions in 1988

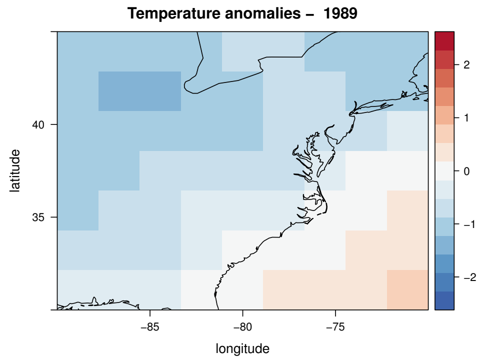


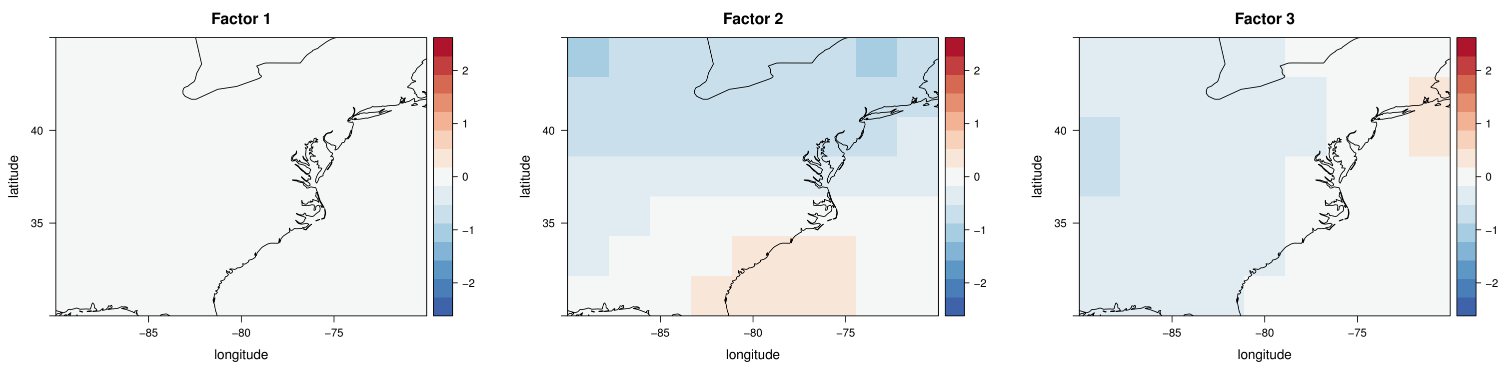




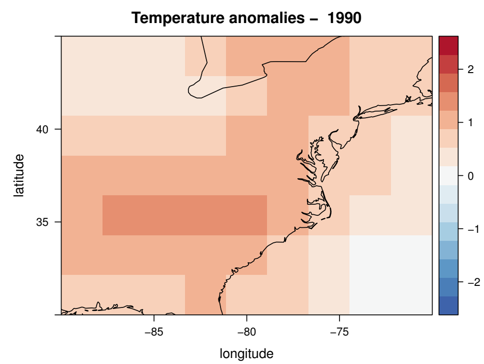


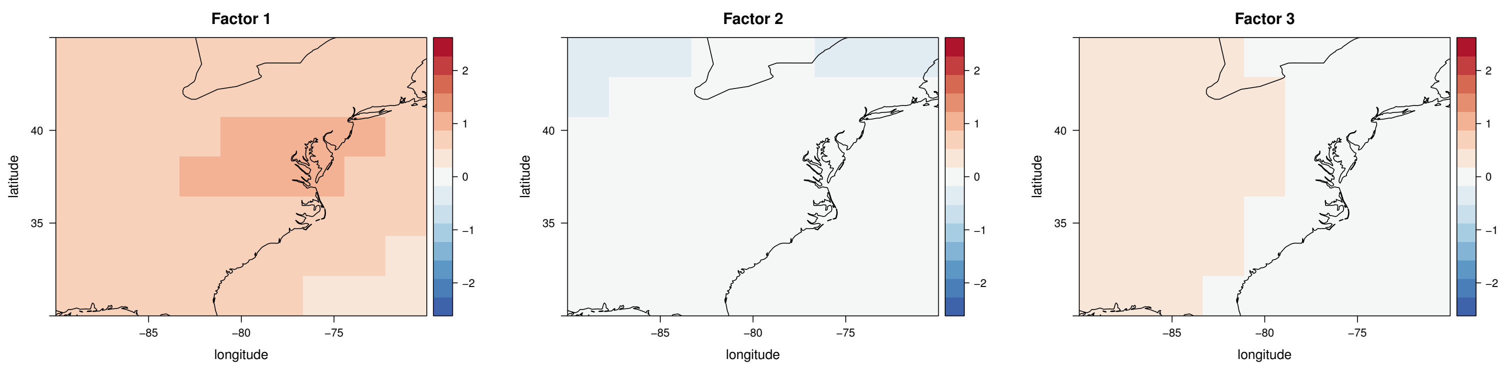
Factor contributions in 1989



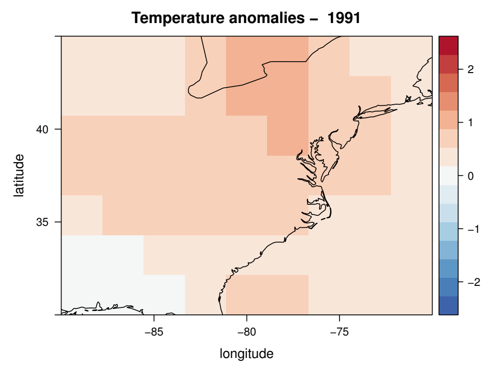


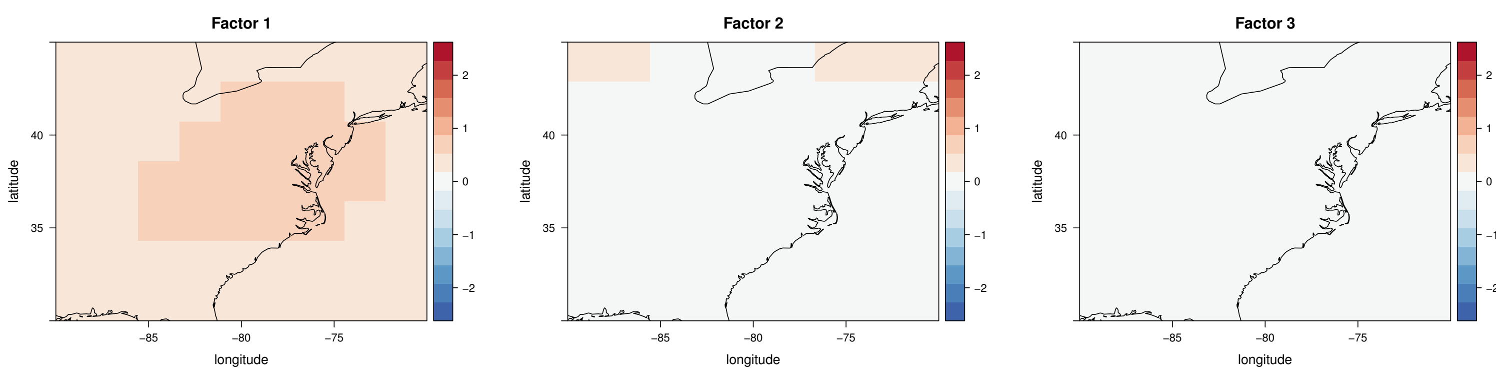
Factor contributions in 1990



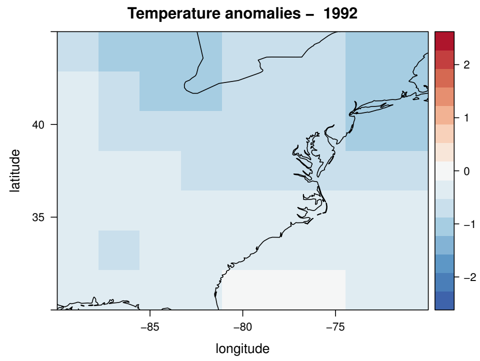


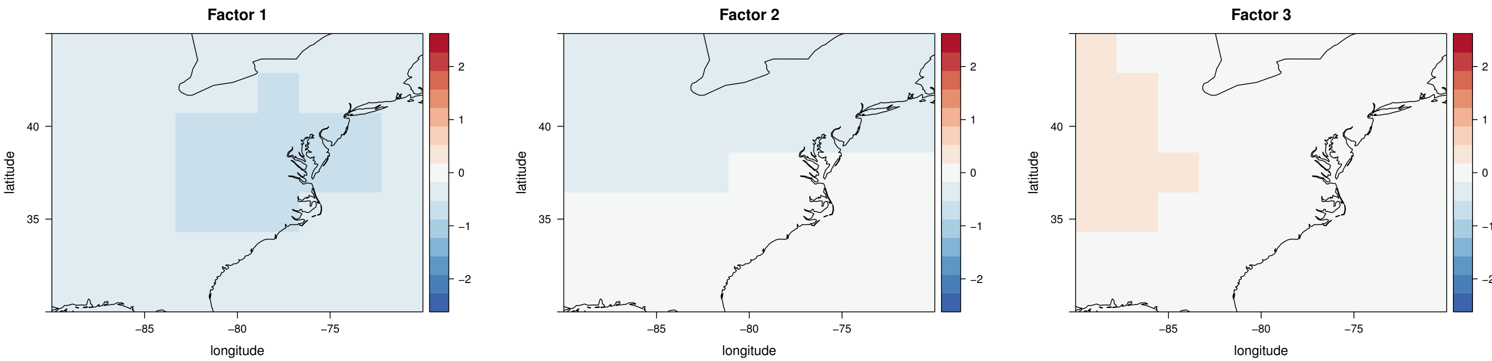
Factor contributions in 1991



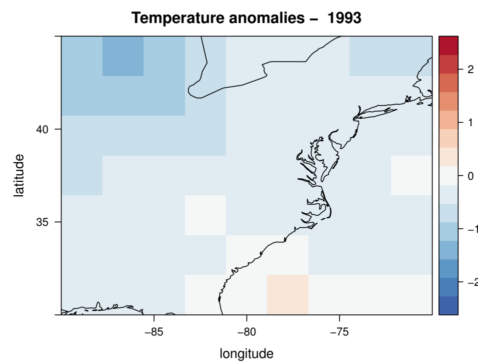


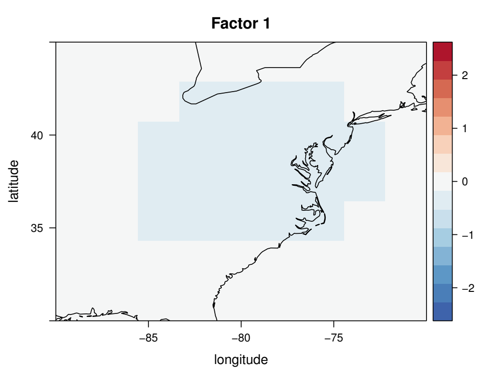
Factor contributions in 1992

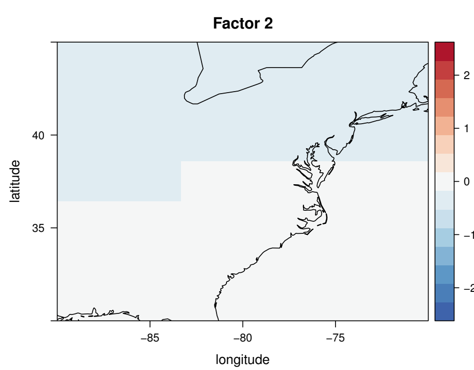


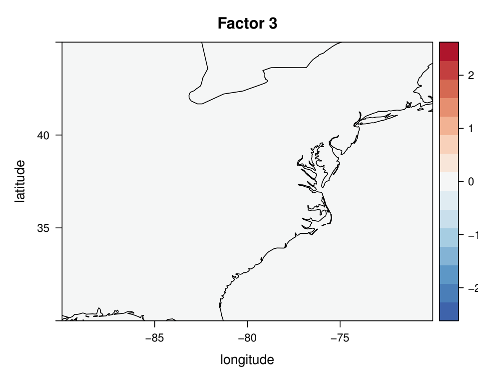


Factor contributions in 1993

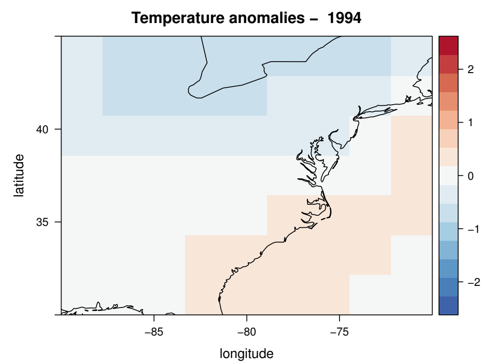


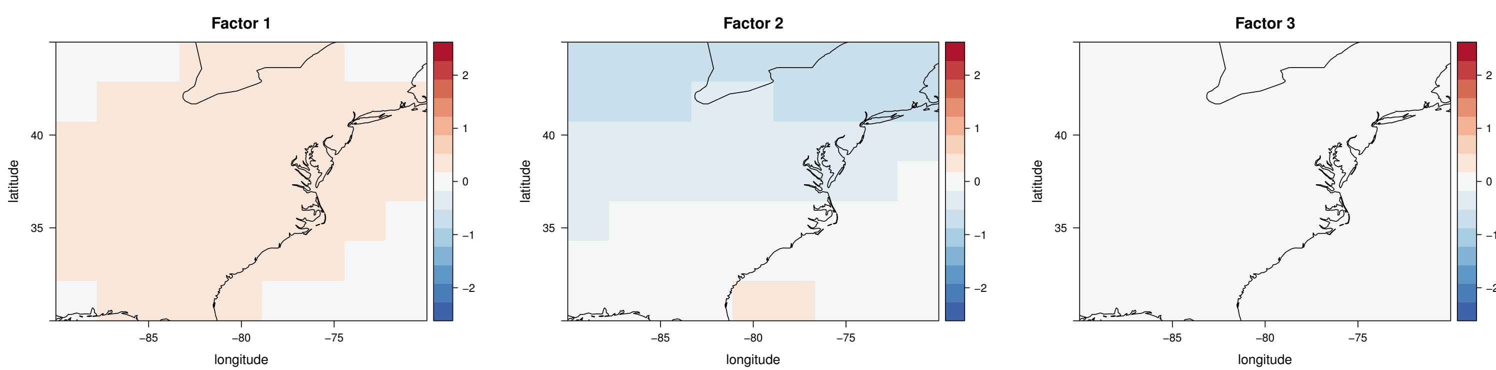




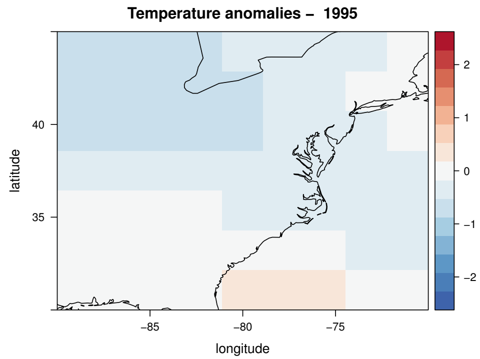


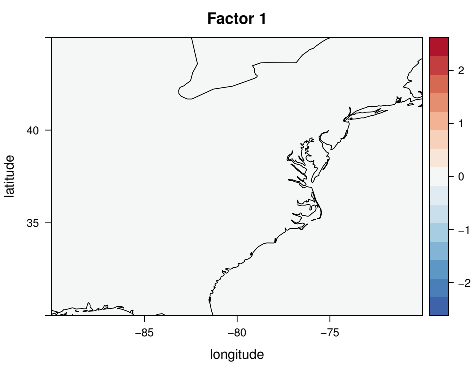
Factor contributions in 1994

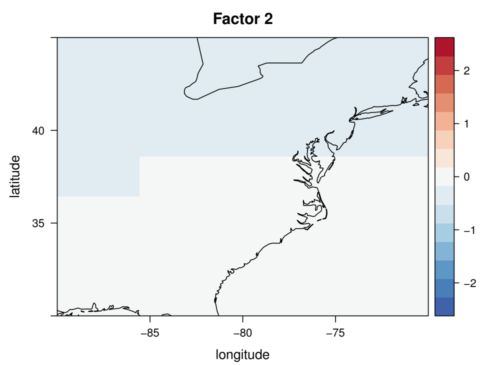


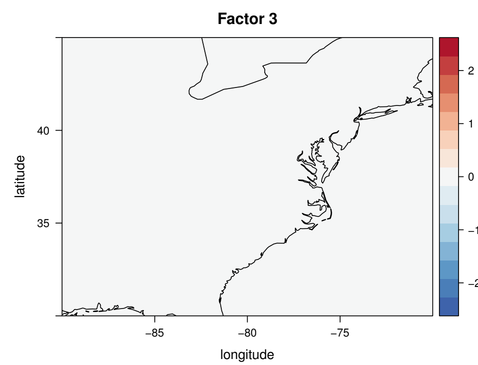


Factor contributions in 1995

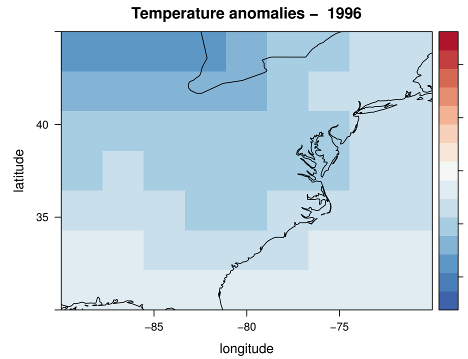


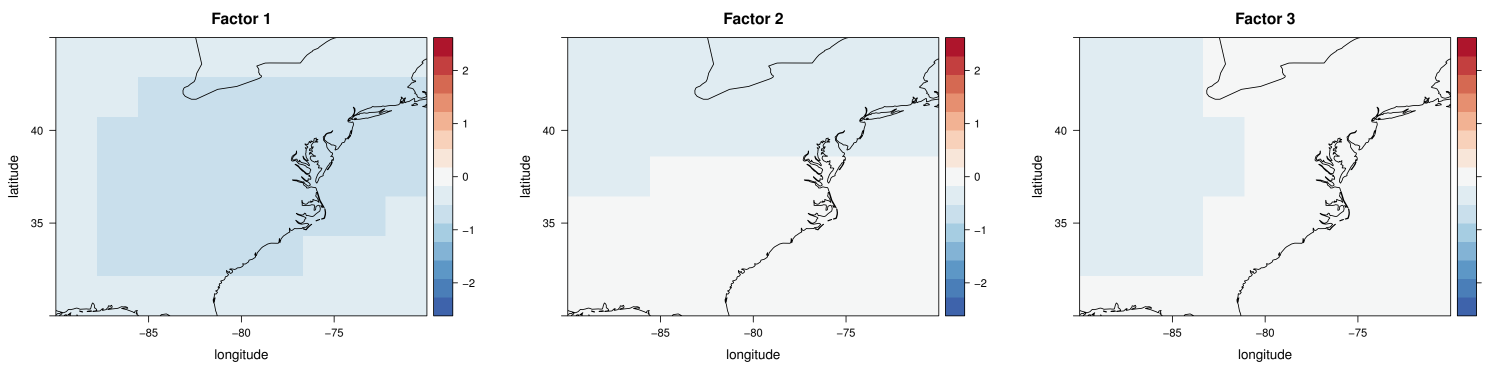




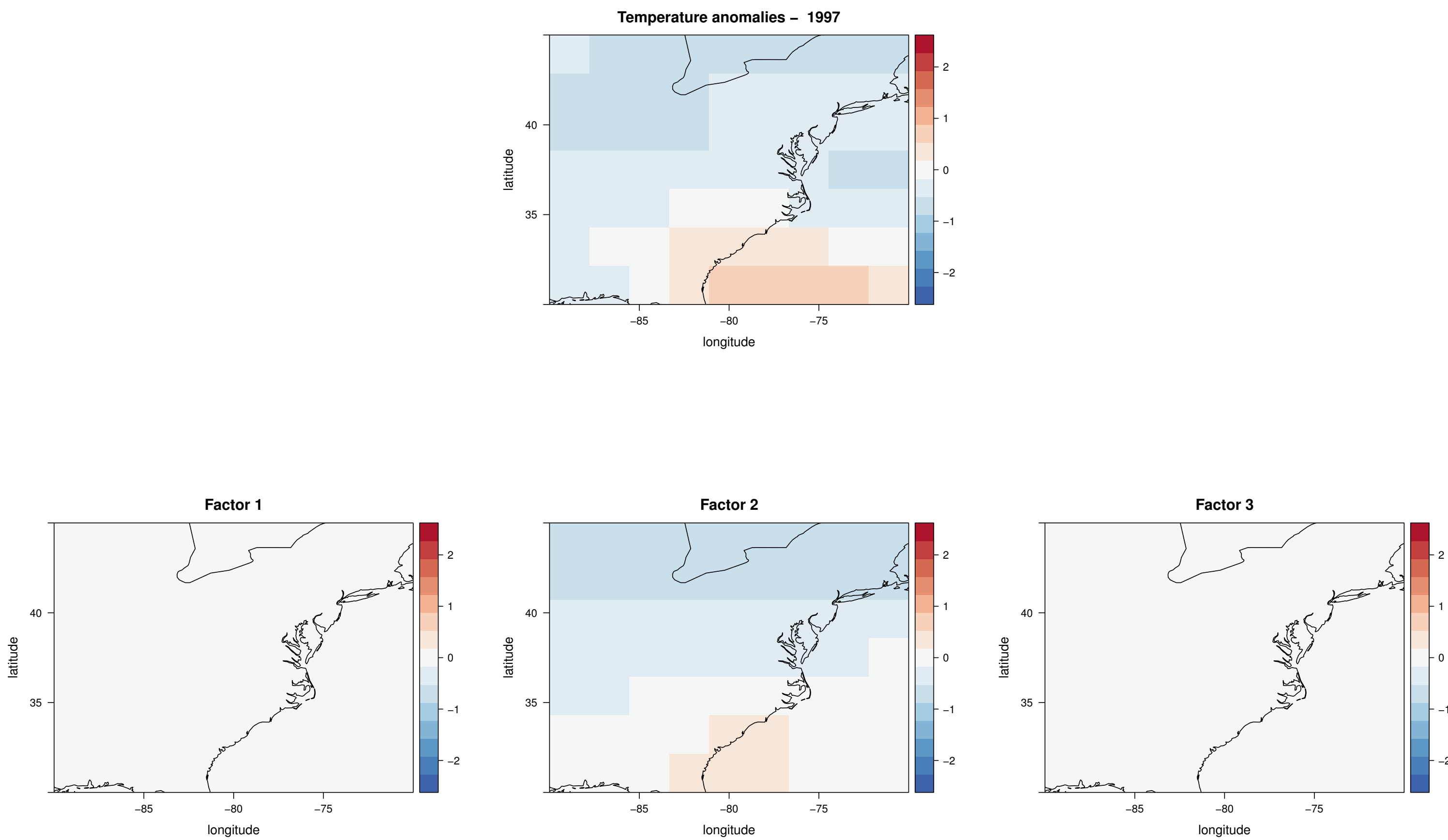


Factor contributions in 1996

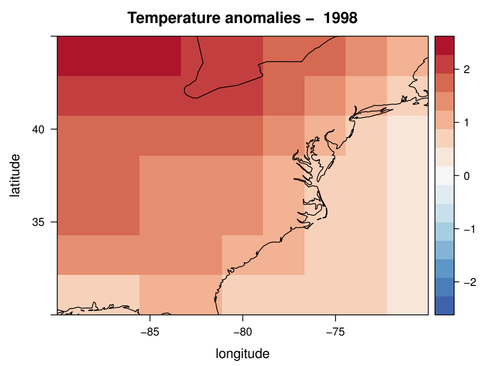


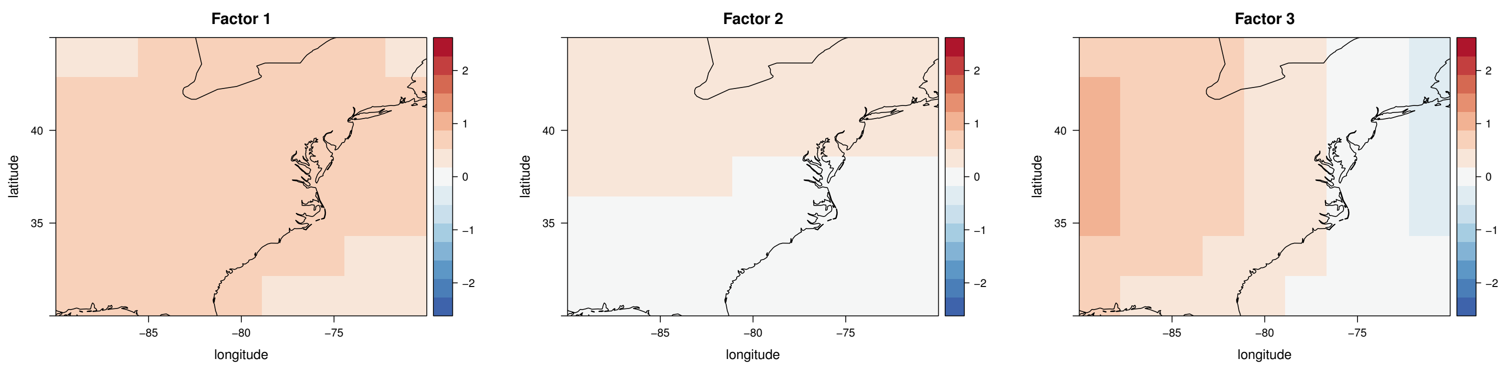


Factor contributions in 1997

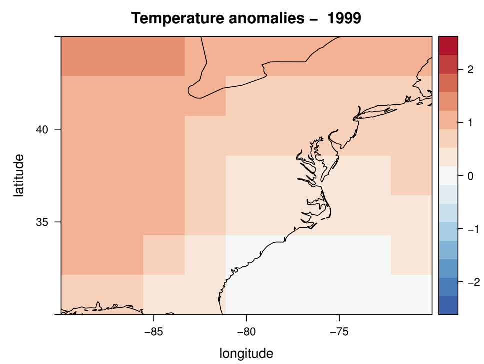


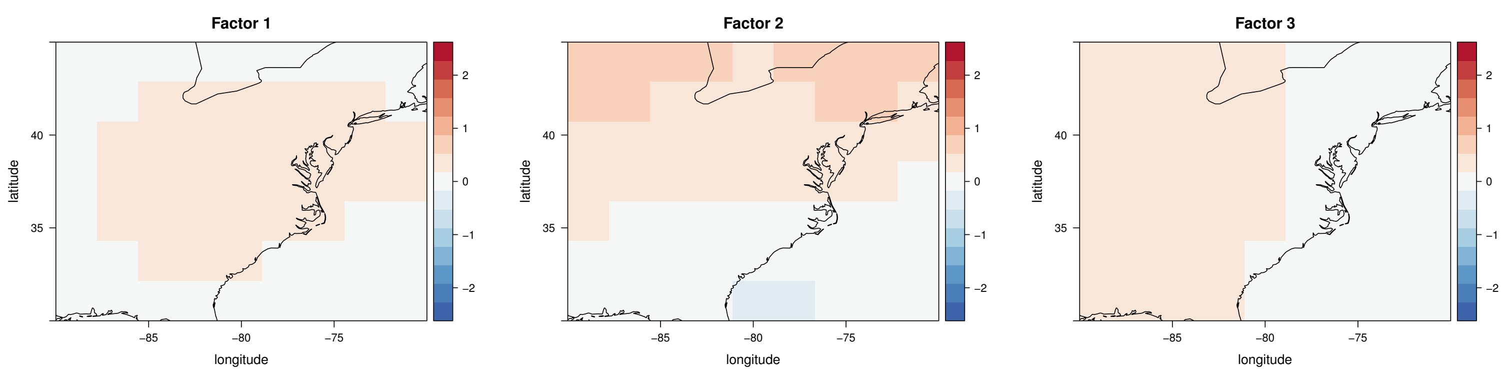
Factor contributions in 1998



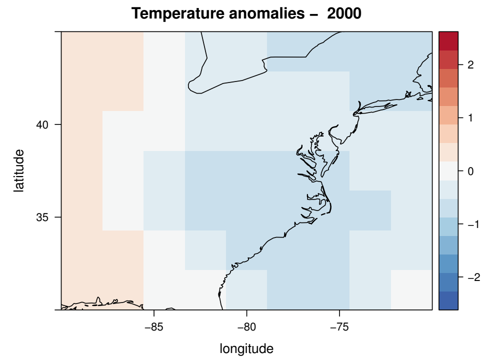


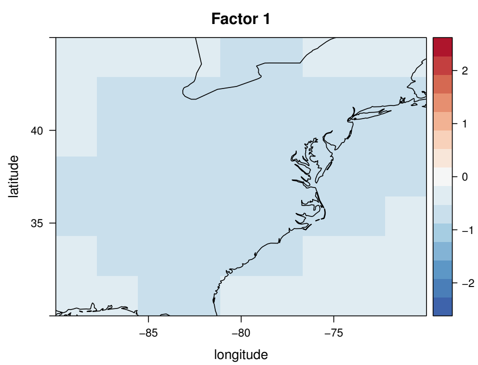
Factor contributions in 1999

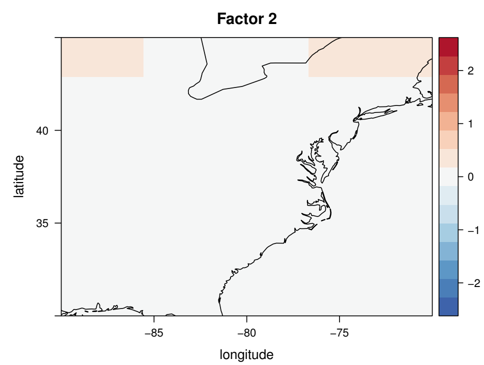


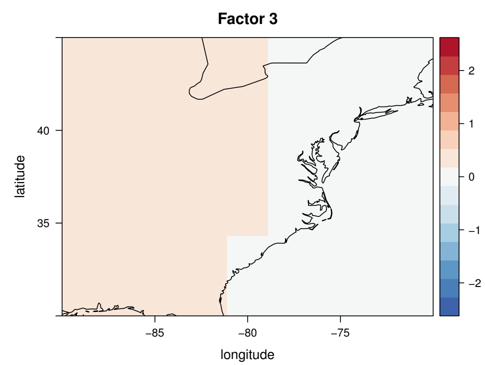


Factor contributions in 2000

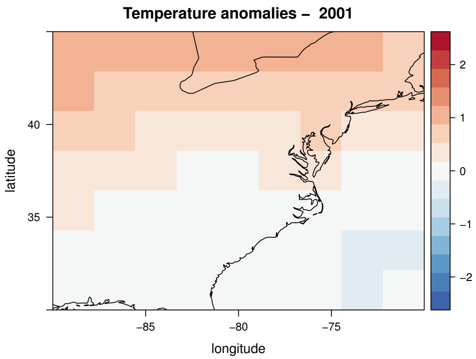


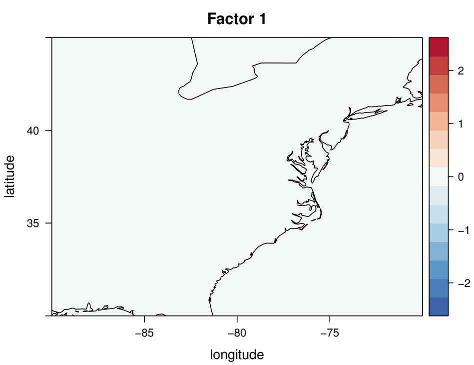


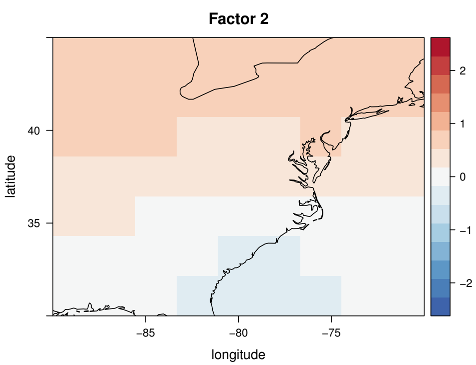


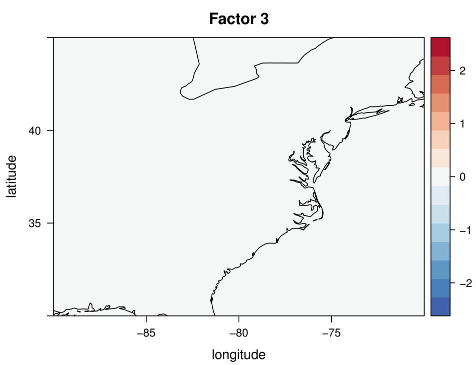


Factor contributions in 2001

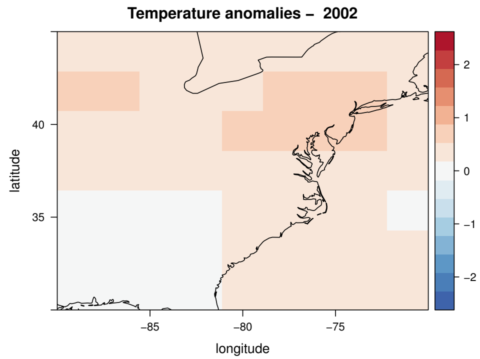


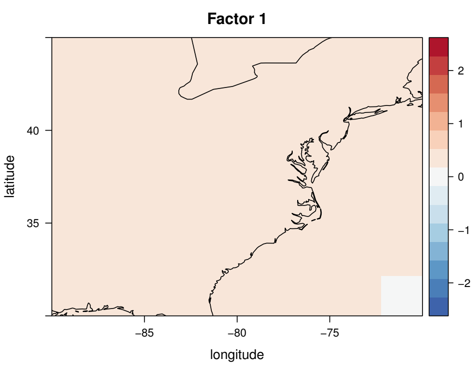


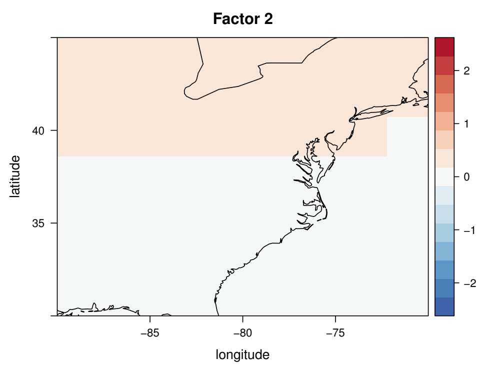


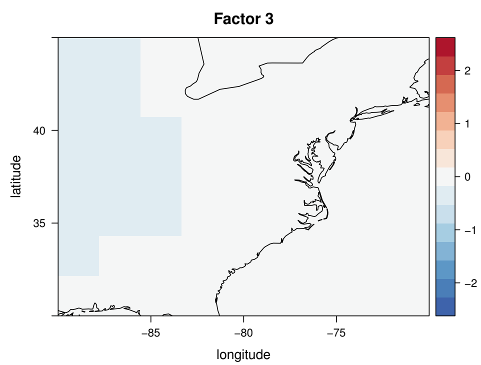


Factor contributions in 2002

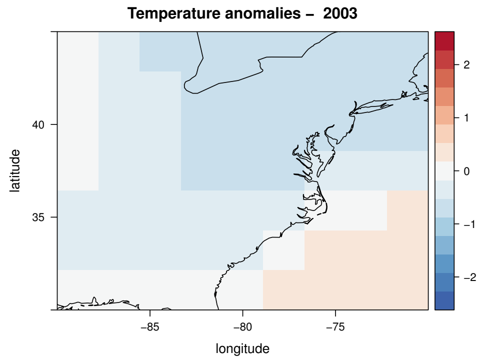


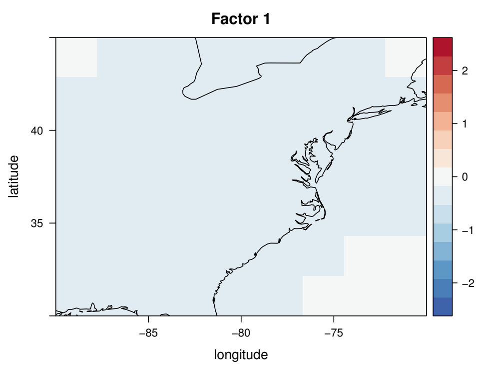


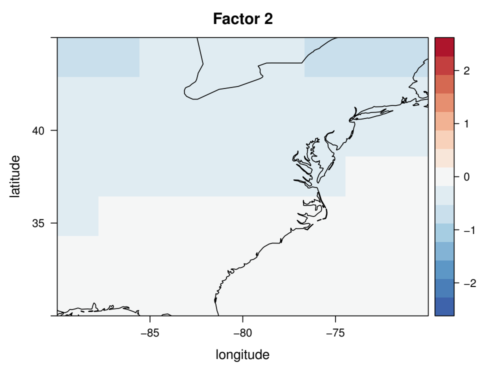


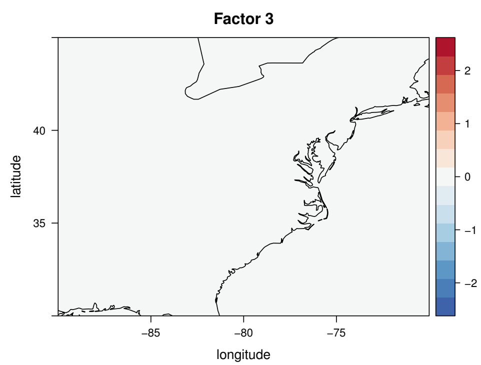


Factor contributions in 2003

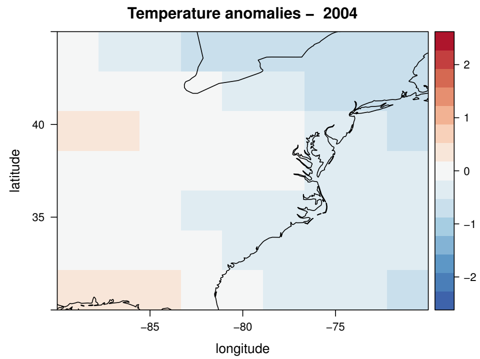


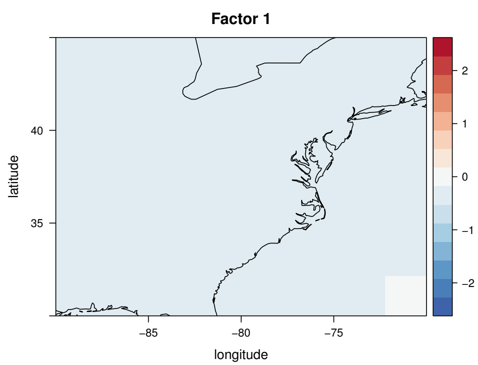


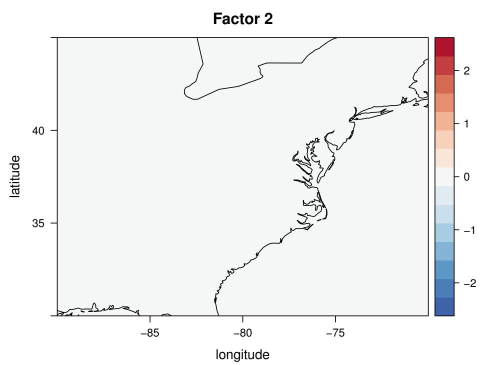


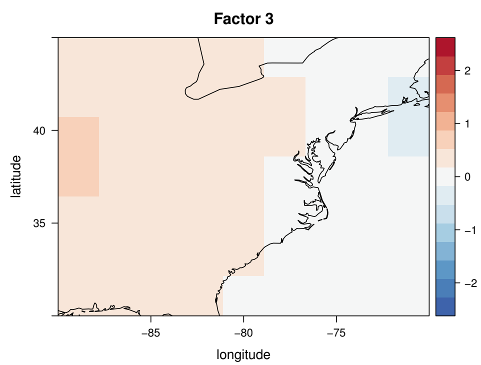


Factor contributions in 2004

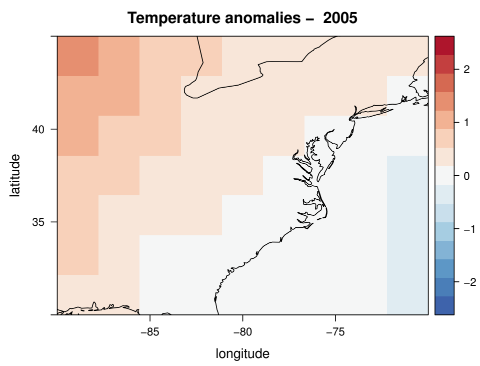


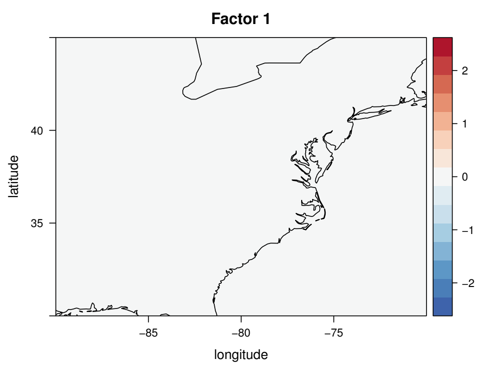


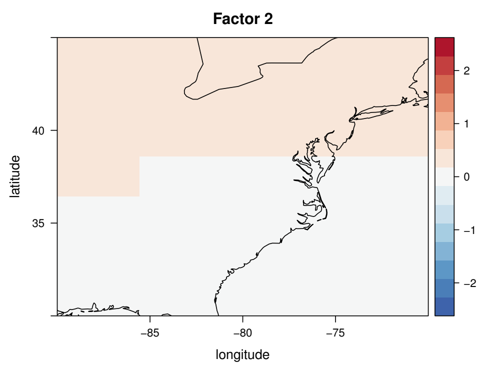


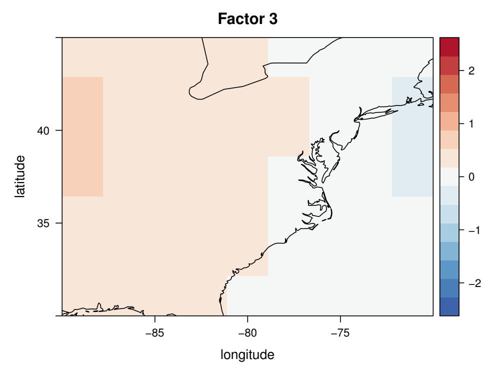


Factor contributions in 2005

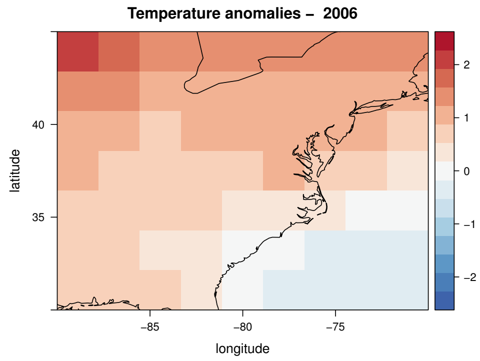


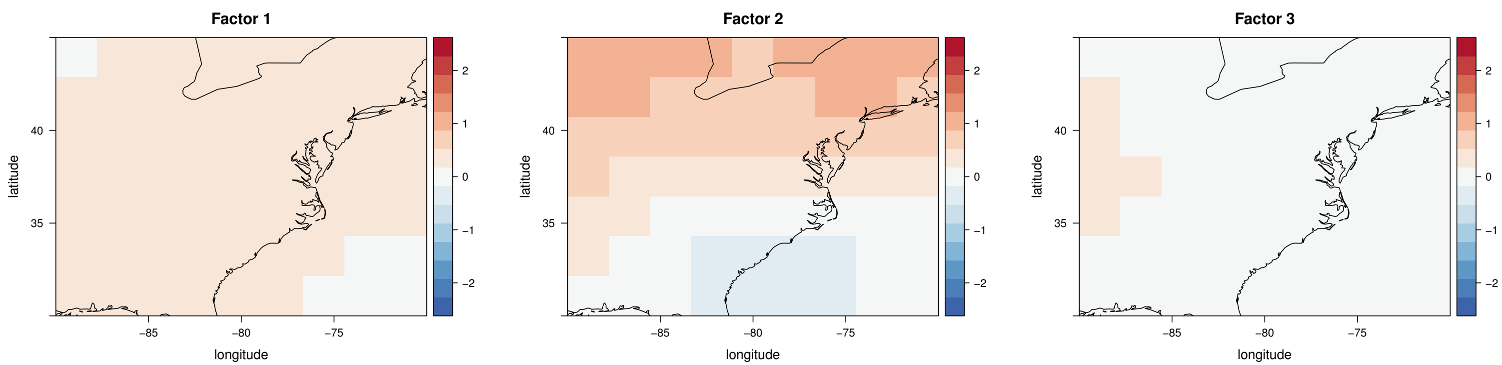




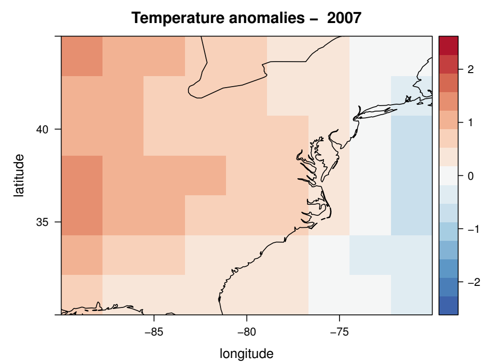


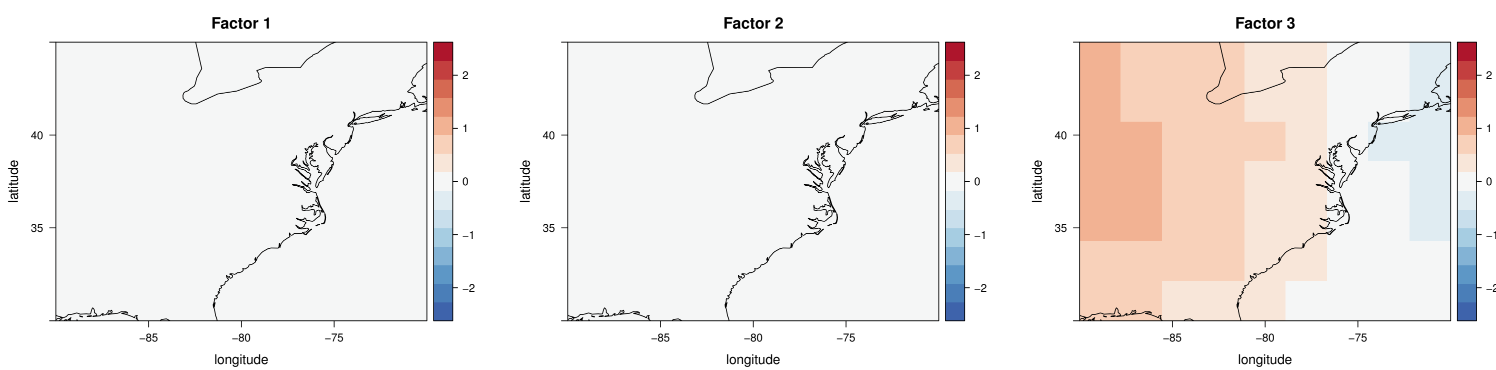
Factor contributions in 2006



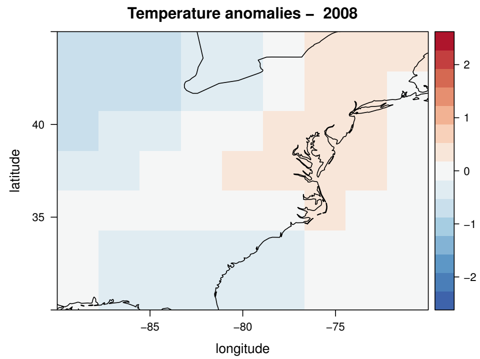


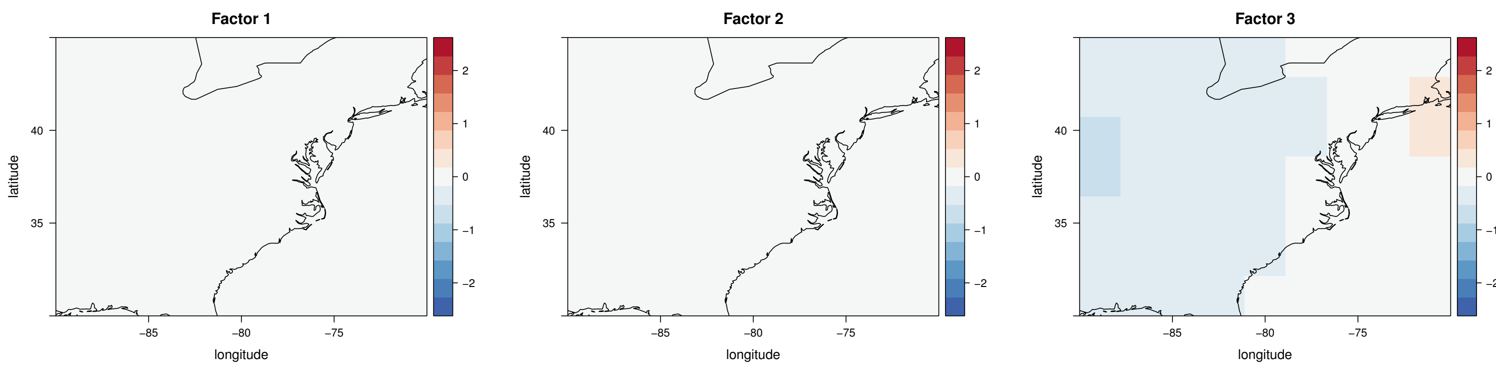
Factor contributions in 2007



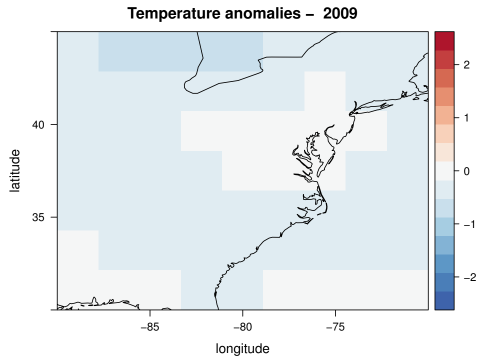


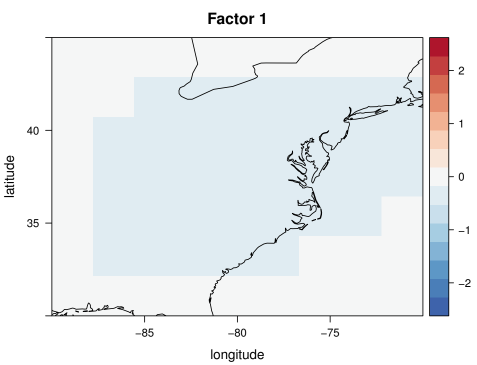
Factor contributions in 2008

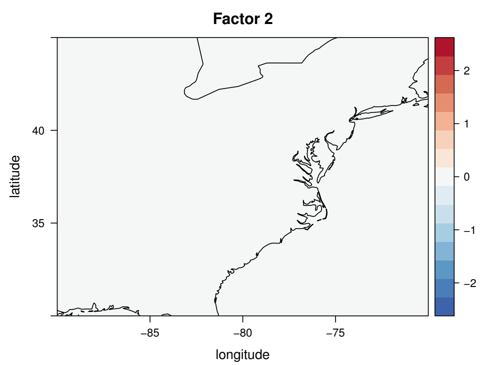


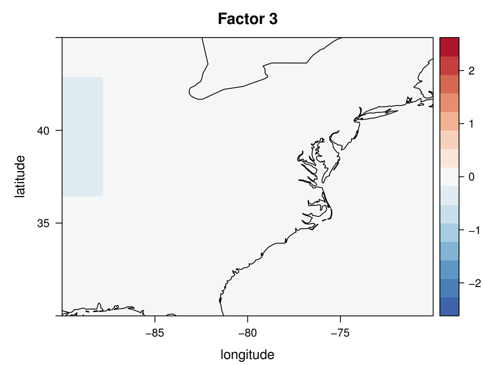


Factor contributions in 2009

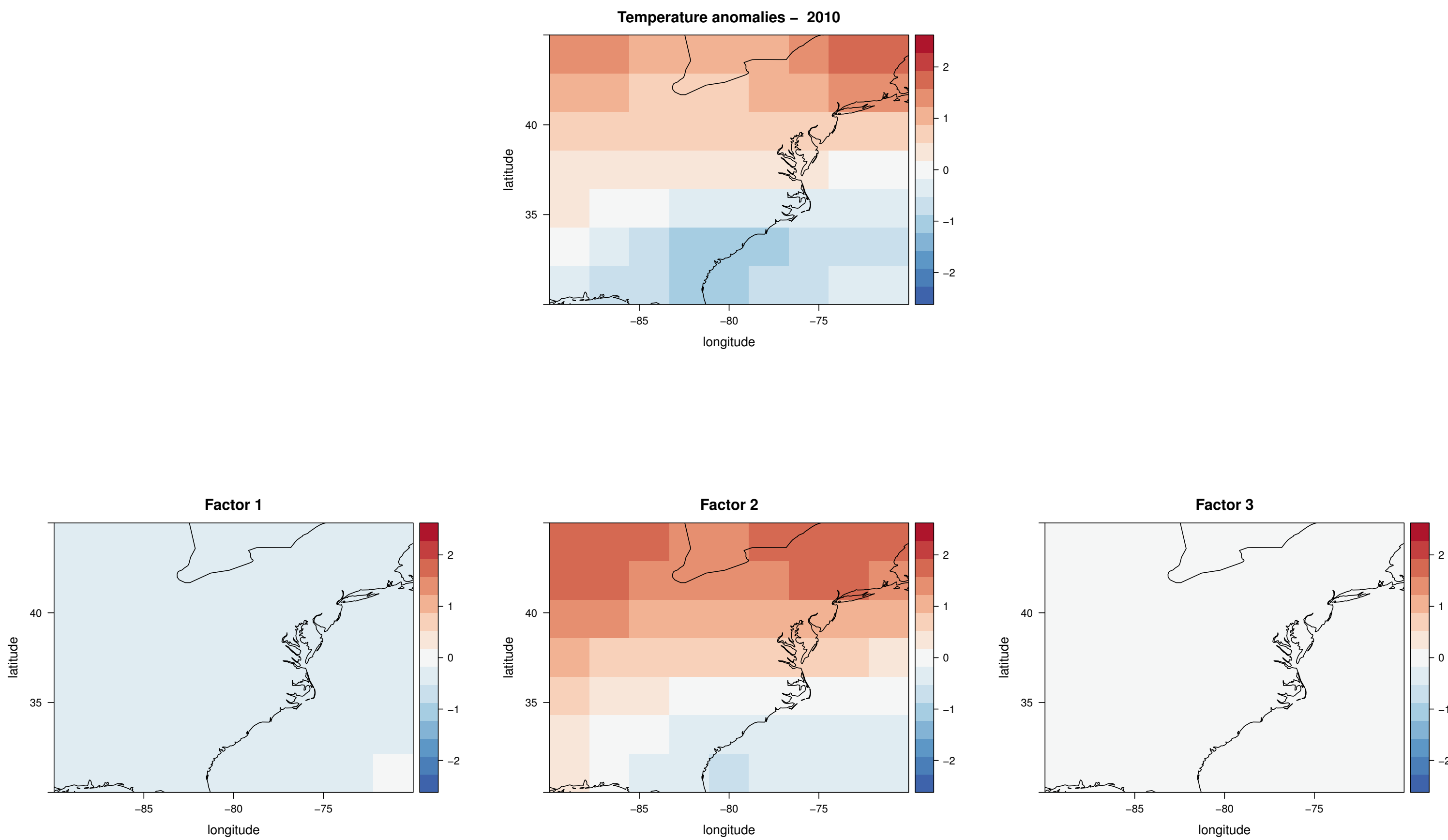




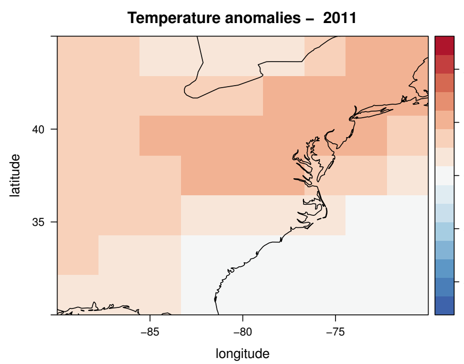


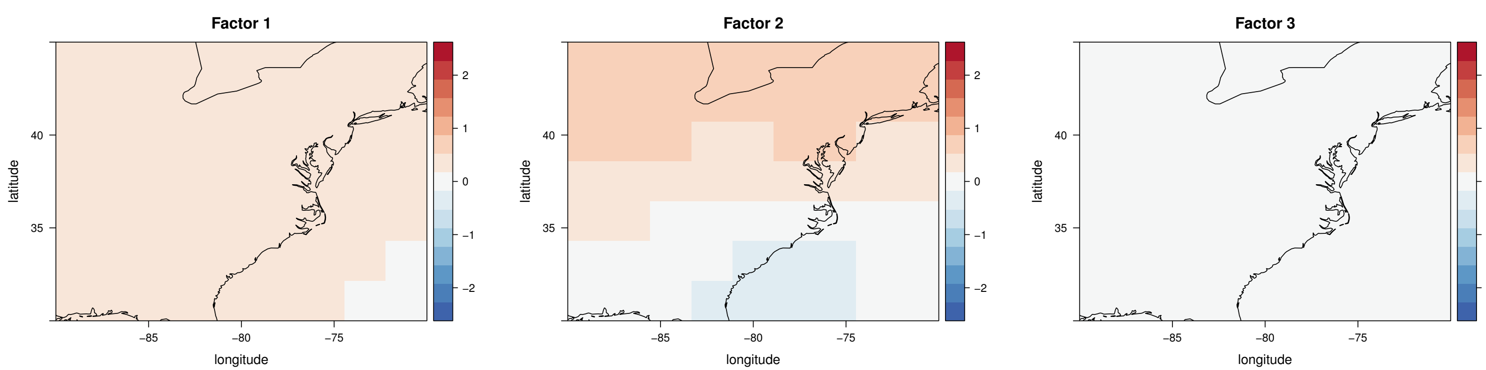


Factor contributions in 2010

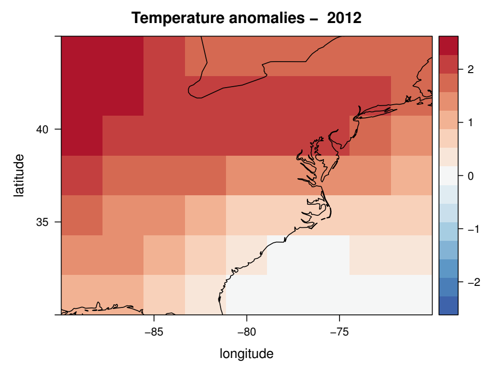


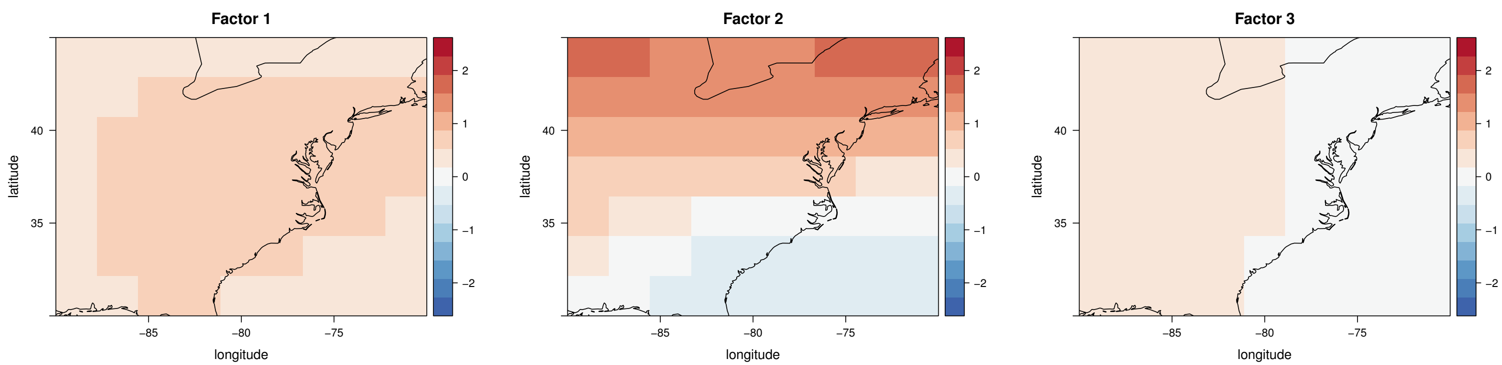
Factor contributions in 2011



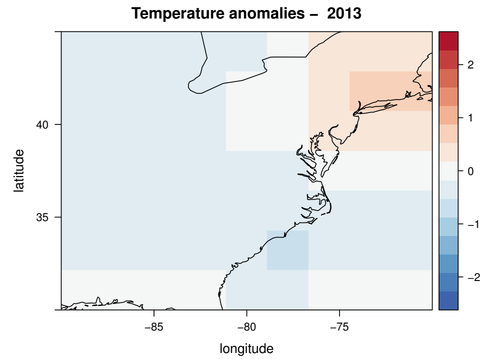


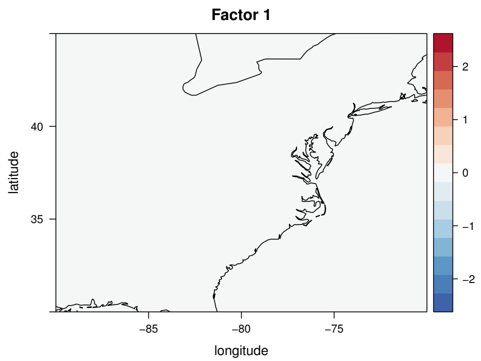
Factor contributions in 2012

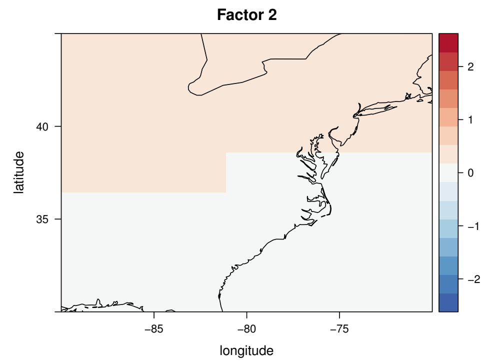


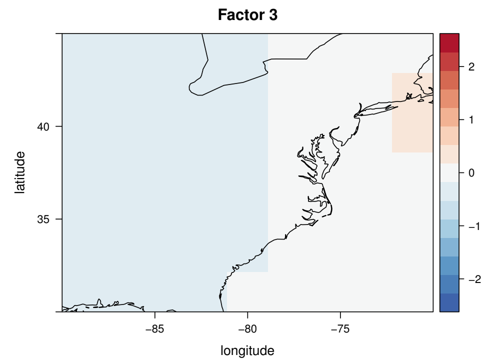


Factor contributions in 2013

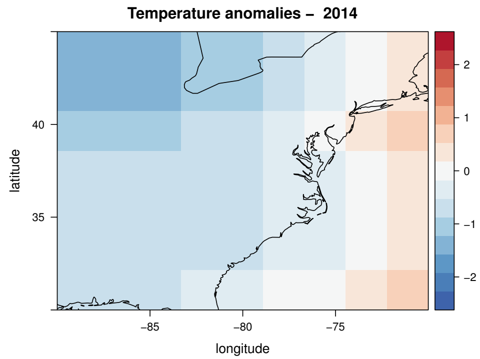


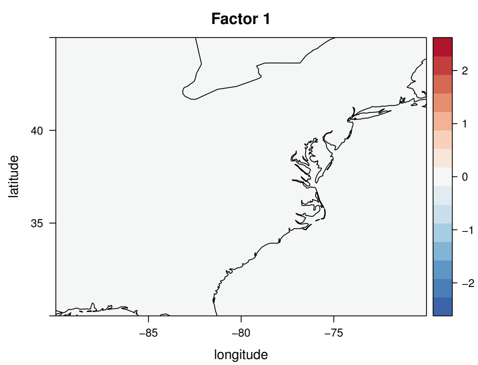


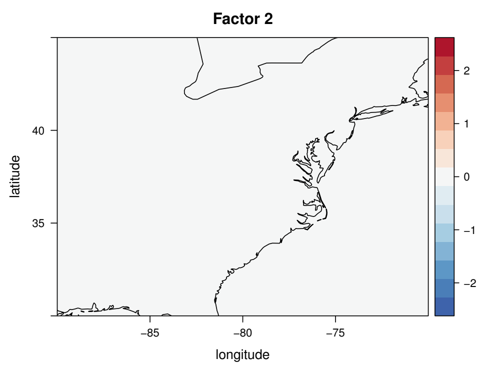


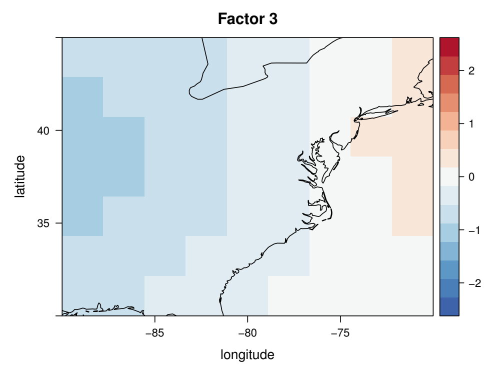


Factor contributions in 2014

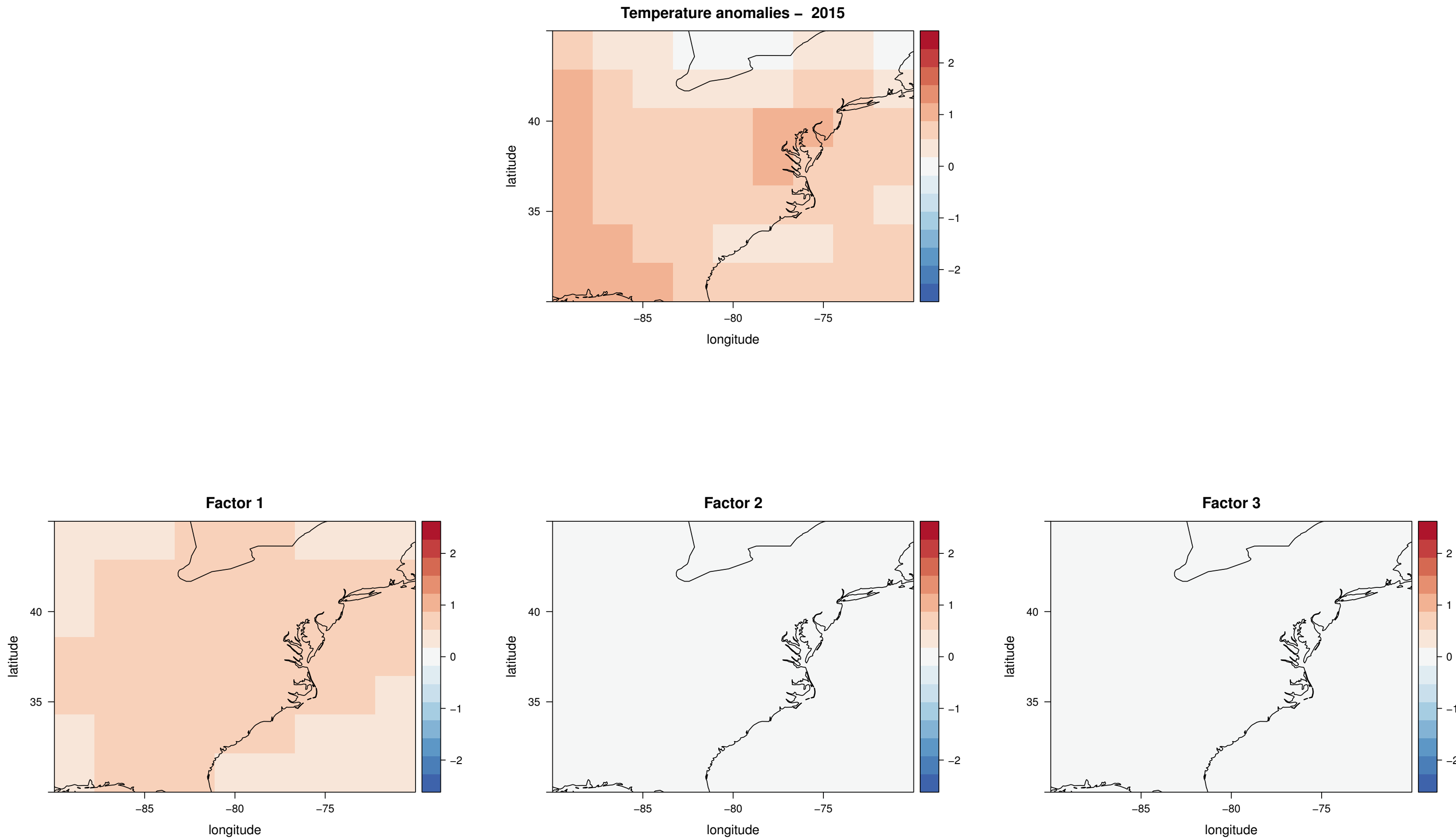




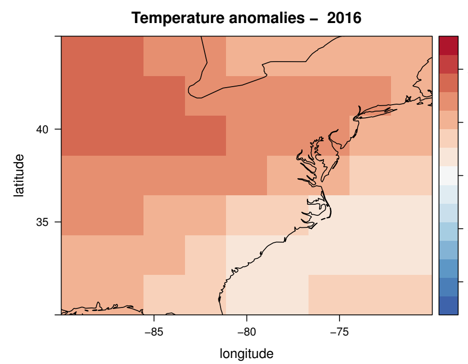


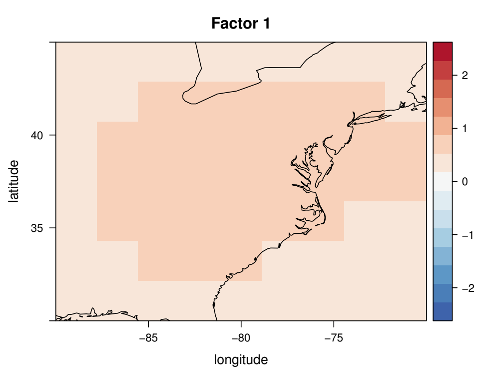


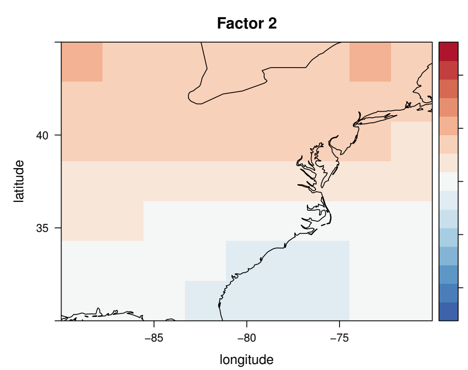
Factor contributions in 2015

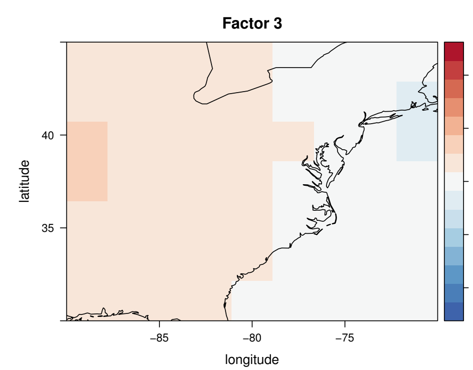


Factor contributions in 2016

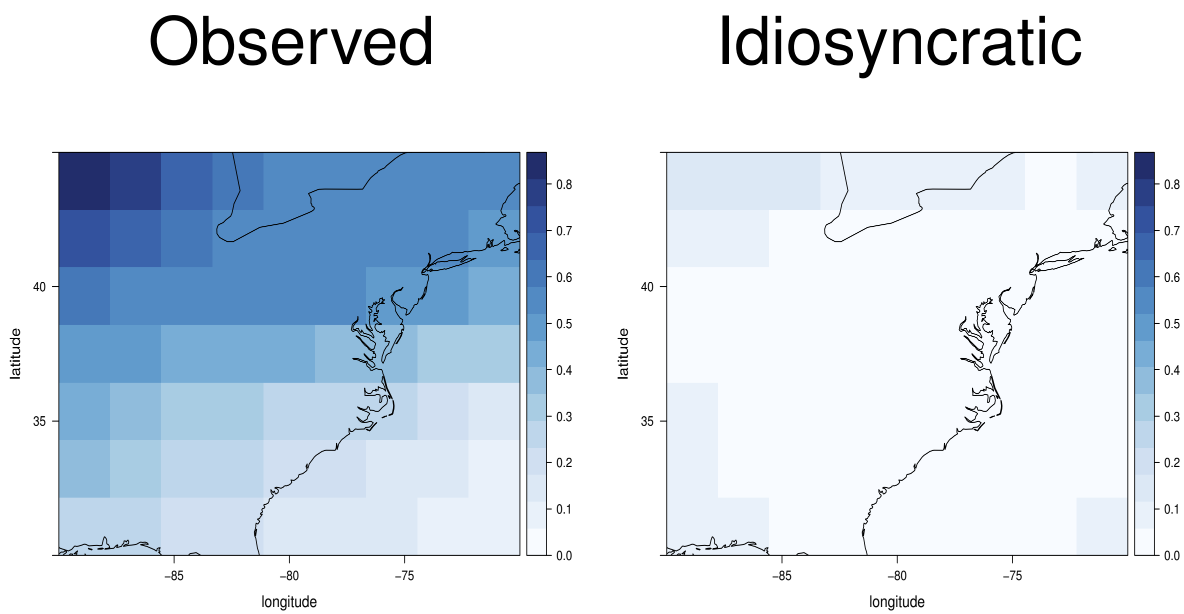








Observed and idiosyncratic variances



- Model explains about 90% of the total variability.

Wrapping up Sessions 4 and 5

- Models for multivariate time series;
- Shared structure/dependence through latent factors;
- Dynamic latent factor models (environmental application);
- Dynamic spatiotemporal factor models (environmental application).

Wrapping up Sessions 4 and 5

- Models for multivariate time series;
- Shared structure/dependence through latent factors;
- Dynamic latent factor models (environmental application);
- Dynamic spatiotemporal factor models (environmental application).

Wrapping up Sessions 4 and 5

- Models for multivariate time series;
- Shared structure/dependence through latent factors;
- Dynamic latent factor models (environmental application);
- Dynamic spatiotemporal factor models (environmental application).

Wrapping up Sessions 4 and 5

- Models for multivariate time series;
- Shared structure/dependence through latent factors;
- Dynamic latent factor models (environmental application);
- Dynamic spatiotemporal factor models (environmental application).