Problem 1. When derived from first principles, the analysis equation for a linear observation operator $\mathbf{H}$ is

$$\bar{\mathbf{x}}^a = \mathbf{P}^a \left[(\mathbf{P}^b)^{-1} \bar{\mathbf{x}}^b + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{y}^o \right], \tag{1}$$

$$\mathbf{P}^a = \left[(\mathbf{P}^b)^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \right]^{-1}, \tag{2}$$

Show that this equation can be equivalently written in the computationally more advantageous form

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{K}(\mathbf{y}^o - \mathbf{H}\bar{\mathbf{x}}^b), \tag{3}$$

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^b, \tag{4}$$

where

$$\mathbf{K} = \mathbf{P}^a \mathbf{H}^T \mathbf{R}^{-1} \tag{5}$$

is the “Kalman gain”

Hint. When the inverse $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$ of two arbitrary matrices $\mathbf{A}$ and $\mathbf{B}$ exist

$$(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}. \tag{6}$$

Solution 1. Taking the inverse of (2), then rearranging the terms in the resulting equation yields

$$(\mathbf{P}^b)^{-1} = (\mathbf{P}^a)^{-1} - \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}. \tag{7}$$

Substituting the left side of (7) for $(\mathbf{P}^b)^{-1}$ in (1) and using (5) yields(4).
To show the equivalence of (2) and (4) we first take the inverse of (2) again, but this time we substitute $(\mathbf{P}^a)^{-1}\mathbf{K}$ for $\mathbf{H}^T\mathbf{R}$ on the left side, which yields

$$(\mathbf{P}^a)^{-1} = [(\mathbf{P}^b)^{-1} + (\mathbf{P}^a)^{-1}\mathbf{K}\mathbf{H}]. \tag{8}$$

From (8)

$$(\mathbf{P}^a)^{-1} = (\mathbf{P}^b)^{-1}(\mathbf{I} - \mathbf{K}\mathbf{H})^{-1}. \tag{9}$$

First applying (6) to the left side of (9) and then taking the inverse of the resulting equation yields (4).

Problem 2. While equation (4) provides important insight into the role he Kalman gain plays in reducing the uncertainty from $\mathbf{P}^b$) to $\mathbf{P}^a$, it cannot be used in its present form to calculate $\mathbf{P}^a$) in (1), because the definition (5) of $\mathbf{K}$ itself includes $\mathbf{P}^a$. For this purpose we use the equation

$$\mathbf{P}^a = (\mathbf{I} + \mathbf{P}^b\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{P}^b. \tag{10}$$

Show that this equation can be derived either directly from (2) or from (4) and (5).

Problem 3. Assume that we have two locations where we want to estimate the wind speed. At those two locations the background value of the wind speed is 8 ms^{-1} and 12 ms^{-1}, while the observed wind speed is 9 ms^{-1} and 14 ms^{-1}, respectively. What are the analyzed values of of the wind speed and the expected standard deviation of the errors at the two locations for the following values of the background and observational errors:

a) The standard deviation of the background error and the observation error is 1 ms^{-1} independently of the location and the errors between the different locations are uncorrelated?

b) The standard deviation of the background error and the observation error is 1 ms^{-1}

independently of the location, the observation errors are uncorrelated, but the covariance between the background errors at the two locations is 0.5?

c) The standard deviation of the background error and the observation error is 1 ms^{-1} independently of the location, the background errors at the two locations are uncorrelated with each other, but the covariance between the two observations is 0.5?

d) What happens when we increase the correlation to 0.9 in b) and c)?

Reminder. The inverse of a 2-by-2 matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{11}$$

is

$$\mathbf{A}^{-1} = \frac{1}{|\,\mathbf{A}\,|} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}, \tag{12}$$

where the determinant $|\,\mathbf{A}\,| = a_{11}a_{22} - a_{12}a_{21}$.

Solution 3.

a) The analyzed values are 8.5 ms^{-1} and 13 ms^{-1}, while the standard error is $\sqrt{0.5}$ ms^{-1} = 0.71 ms^{-1} at both locations.

b) The analyzed values are 8.56 ms^{-1} and 12.92 ms^{-1}, while the standard error is $\sqrt{21/45}$ ms^{-1} = 0.68 ms^{-1} at both locations.

c) The analyzed values are 8.07 ms^{-1} and 12.76 ms^{-1}, while the standard error is $\sqrt{21/45}$ ms^{-1} = 0.68 ms^{-1} at both locations.

Problem 4. Assume that the background is the same as in Problem 2, but only one of the two observations is available, the one that measured 9 ms^{-1}. What are the analyzed values of the wind speed and the expected standard deviation of the errors at the two locations for the following values of the background and observational errors:

a) The standard deviation of the background error and the observation error is 1 ms^{-1} independently of the location and the errors between the different locations are uncorrelated?

b) The standard deviation of the background error and the observation error is 1 ms^{-1} independently of the location, the observation errors are uncorrelated, but the covariance between the background errors at the two locations is 0.5?

c) How does the result of b) change if the covariance between the background errors at the two locations is increased to 0.9?

Hint. For this configuration

$$\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{13}$$

Solution 4.

a) The analyzed values are 8.5 ms^{-1} and 12 ms^{-1}, while the standard errors are $\sqrt{0.5}$ ms^{-1} = 0.71 ms^{-1} and 1 ms^{-1}.

b) The analyzed values are 8.53 ms^{-1} and 12.45 ms^{-1}, while the standard errors are 0.71 ms^{-1} and 0.94 ms^{-1}.

c) The analyzed values are 8.44 ms^{-1} and 12.51 ms^{-1}, while the standard errors are 0.71 ms^{-1} and 0.77 ms^{-1}.