The $PM_{10-2.5}$ DQO Model

This is a companion document to the report titled*: Sensitivity of the PM10-2.5 Data Quality Objective to Spatially Related Uncertainties* provided to the Clean Air Scientific Committee (CASAC) for the meeting scheduled for September 21-22, 2005.

1.0 THE SIMULATION MODEL

The simulation model used for establishing the $PM_{10-2.5}$ DQOs and equivalency method requirements is fairly complex. It was developed as a generalization of the model used for the $PM_{2.5}$ DQO simulation model. The basic model was first implemented in MathCad and then reprogrammed in C^{++} with corresponding functions to increase the speed of execution. The C^{++} functions have been successfully tested against the MathCad functions. Consequently, it suffices to describe the MathCad functions that are much easier to read. The MathCad version is described below.

2.0 CONCEPTUAL MODELS

There are several layers to the conceptual model. The $PM_{10-2.5}$ and $PM_{2.5}$ components are viewed separately. Since the $PM_{2.5}$ data generation is simpler and motivates the rest, it is described first. Within the code, however, it is easier to generate it last.

2.1 The $PM_{2.5}$ Conceptual Model

The $PM_{2.5}$ component is thought of as a spatially uniform (within the scale of the monitor) process with a sinusoidal mean having one period per year. The basic sine curve has its modulus and vertical shift related so that models vary multiplicatively. This is a very important point for the implementation, as all the parameters have a multiplicative interpretation. This means that a site is associated with a fixed ratio between the highest point on the sign curve and its lowest point. Consequently, the sine wave is parameterized as:

$$A \cdot \left(1 + \frac{R-1}{R+1} \sin\left(\frac{2\pi \cdot day}{365}\right)\right)$$

where *A* is a multiplicative factor, *R* is the seasonal ratio, and *day* is the number of days into the simulation period (day ranges from 1 to 3x365). Since three full years are modeled, there is no need for a phase shift in the $PM_{2.5}$ component.

From this mean the true day-to-day concentrations are generated as lognormally distributed random deviations from the mean with a constant coefficient of variation (CV). These deviations are allowed to have autocorrelation that is modeled as the exponential of a simple autoregressive process.

2.2 The $PM_{10-2.5}$ Conceptual Model

The coarse portion of the PM is simulated as a random field that varies with correlation both spatially and temporally. The spatial domain is simulated on a grid that is supposed to represent a neighborhood scale monitor with grid points located every 2 km in a square grid that is 8 km on a side. The temporal variation also includes a seasonal pattern throughout the three-year simulation periods. For each grid point, separate three-year design values are computed. Then the grid's design values (one for the annual standard and one for the daily standard) are computed by averaging over the grid-point specific design values. Finally, this design value is compared to the design value for the simulated observations corresponding to the center point. This whole process is repeated 1,000 times to determine the probability that the observed-design value correctly predicts whether or not the grid's true-design value is above the standard.

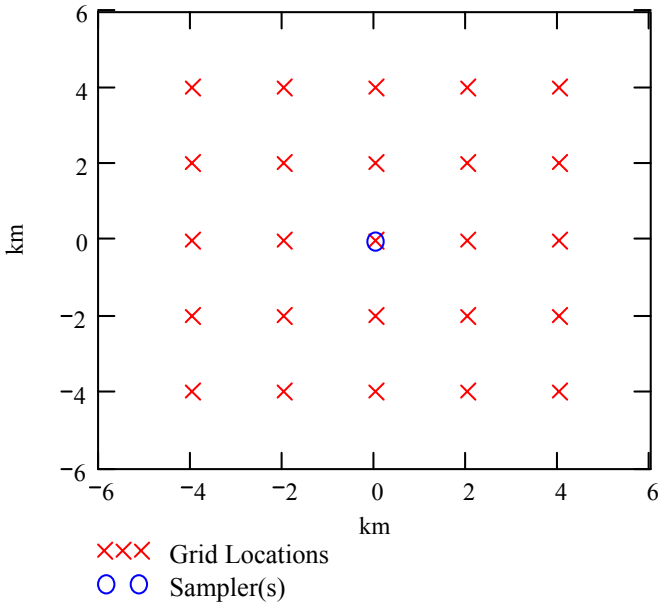

Figure A-1. Simulation grid points.

The temporal pattern for the coarse fraction is also based on random deviations from a sinusoidal pattern. The pattern can have either one or two periods per year and can be phase shifted from the fine PM portion. The random deviations can have autocorrelation and be correlated with the random deviations of the fine fraction.

3.0 PROGRAMMING DETAILS

The details are made somewhat complex by the need to pre-specify the various correlations, variances, and other model parameters. To achieve the desired results, the simulation is done with a series of functions/modules with separate tasks. Each of these is described below.

Notes:

1. The indexing always starts with 0.
2. The parameter vector is referred to as parms.
3. Within the code, $A^{<n>}$ refers to the n^{th} column of the matrix *A*.
4. The code shown is the MathCad code that was used to develop the model. The C^{++} code is tested against the MathCad code. To accommodate that testing, the format of the final output has been modified from what is currently shown.

3.1 Scaled Design Value Distribution

Input: The vector *parms* of parameter values.

```
Gen_scaled_dist(parms) := | Seed(1233547)
                          | bchol ← BChol(parms)
                          | for k ∈ 0.. reps − 1
                          |     in ← Gen_truth(bchol, parms)
                          |     X ← submatrix(in, 0, rows(in) − 2, 0, cols(in) − 1)
                          |     tm ← in_{rows(in)−1, 0}
                          |     Y ← Gen_obs(X, parms)
                          |     Z^⟨k⟩ ← Cal_pobs(Y, parms, tm)
                          | Z^T
```

This function forms the main simulation loop. For each repetition of the loop, a three-year set of scaled true concentrations are generated for each point in the grid. The output of the Gen_truth, however, only contains the simulation values for the center grid point and a summary appended to the end. The next steps in the code separate the simulation values and summary for use later in the code. The fourth step in the loop generates the observations corresponding to the scaled true concentrations. The final step in the loop summarizes the observed-design values for this three-year simulation. The entire loop is repeated as often as is practical and is controlled by the value of the global variable *reps*.

The output of the function is a collection of scaled design values that all correspond to having the mean over the grid points of the design values of the true concentrations equal to 1. The scaling is such that multiplying the simulated observed-design values by 20 results in a set of simulated observed-design values. These values all correspond to examples with simulated true concentrations such that the mean of the true-design values over the grid points is exactly 20. This is true for both the daily standard design values and the annual standard design values. (Note that 20 is *not* the value of the long-term mean for the hypothetical site.)

3.2 The Cholesky Decomposition Matrix

Input: *parms*, the vector of simulation parameters.

$$
\begin{aligned}
&\text{BChol(parms)} := \left|\begin{array}{l}
\text{if } \text{parms}_{22} > 0 \\
\quad \left|\begin{array}{l}
\text{for } i \in 0..4 \\
\quad \text{for } j \in 0..4 \\
\quad\quad \text{for } s \in 0..4 \\
\quad\quad\quad \text{for } t \in 0..4 \\
\quad\quad\quad\quad x_{5\cdot i+j,\,5\cdot s+t} \leftarrow \text{parms}_{22}\cdot\exp\left[-\left|\begin{pmatrix} i-s \\ j-t \end{pmatrix}\right|\cdot\frac{2}{\text{parms}_{23}}\right] \\
\text{out} \leftarrow \text{cholesky}(x)
\end{array}\right. \\
\text{out}_{24,24} \leftarrow 0 \quad \text{if } \text{parms}_{22} = 0 \\
\text{out}
\end{array}\right.
\end{aligned}
$$

Note:

The matrix multiplication of the Cholesky decomposition of a matrix **A** and a vector/matrix of values independently drawn from a standard normal distribution, N(1,0), results in a vector/matrix with the elements that have a variance-covariance matrix equal to **A**. We use this to create the user-specified spatial correlation structure. Specifically, this function creates the Cholesky decomposition of the spatial correlation matrix, except when the sill is set to 0. When the sill is set to zero, the matrix is all zeros. In either case, a separate temporal pattern is added to reflect the fact that the temporal variance is larger than the spatial variance. An exponential spatial correlation is assumed. The factor of 2 in the exponential is to reflect the fact that the grid points are 2 km apart.

3.3 The Coarse Mass Deviations Generator

Input: The matrix *bchol* from the function BChol and *parms*, the vector of simulation parameters.

$$\text{Gen\_CM\_deviations}(bchol, parms) := \left| \begin{array}{l} tmp^{\langle 0 \rangle} \leftarrow \text{rnorm}(26, 0, 1) \\ tmp^{\langle 0 \rangle} \leftarrow tmp^{\langle 0 \rangle} - \text{mean}(tmp) \\ \text{for } i \in 1..365 \cdot 3 + 10 \\ \quad tmp^{\langle i \rangle} \leftarrow \text{rnorm}(26, 0, 1) \cdot \sqrt{1 - \left(parms_{13}\right)^2} + tmp^{\langle i-1 \rangle} \cdot parms_{13} \\ tmp2 \leftarrow bchol \cdot \text{submatrix}(tmp, 0, 24, 11, 3 \cdot 365 + 10) \\ \text{for } i \in 0..365 \cdot 3 - 1 \\ \quad tmp2^{\langle i \rangle} \leftarrow tmp2^{\langle i \rangle} + \sqrt{\left(parms_{16}\right)^2 - parms_{22}} \cdot tmp_{25, i+10} \\ tmp2^{T} \end{array} \right.$$

Notes:

This function creates the (multiplicative) temporal deviations of the long-term trend for the $PM_{10\text{-}2.5}$ component. The first loop creates an autocorrelated temporal sequence for all of the grid points and an extra sequence. The first 25 sequences are then given the desired spatial correlation structure by multiplying by the Cholesky decomposition. However, the spatial variation can be less than temporal variance. This is fixed by adding a spatially uniform temporal sequence to all of the points in the grid based on the "extra" sequence generated in the first loop. Since everything has the same temporal autocorrelation, then the end result has the A1 structure ($X_{n+1} = \alpha X_n + Y_{n+1}$, where the Ys are independent with an appropriate variance).

Also, note that there is a very short "burn in." The temporal sequences are generated for a slightly longer period than needed by the rest of the program. This is needed to adjust for the fact that the A1 structure does not hold initially.

3.4 The Fine Mass Deviations Generator

Input: The matrix *CM_dev* from the function Gen_CM_deviations, and *parms*, the vector of simulation parameters.

$$
\begin{aligned}
&\text{Gen\_FM\_deviations}(CM\_dev, parms) := \\
&\quad x \leftarrow parms_{14} \cdot \frac{parms_{15}}{parms_{16}} \cdot \frac{\left(1 - parms_{12} \cdot parms_{13}\right)}{\left[1 - \left(parms_{13}\right)^2 \cdot \left(parms_{14}\right)^2\right]} \\
&\quad a \leftarrow \left(parms_{14}\right)^2 \\
&\quad b \leftarrow \left(parms_{12}\right)^2 \\
&\quad c \leftarrow \left(parms_{13}\right)^2 \cdot \left(parms_{14}\right)^2 \\
&\quad d \leftarrow 2 \cdot \left(parms_{14}\right)^2 \cdot parms_{13} \cdot parms_{12} \\
&\quad y \leftarrow \sqrt{\frac{1 - a - b - c + d}{1 - \left(parms_{13}\right)^2 \cdot \left(parms_{14}\right)^2}} \\
&\quad k \leftarrow \frac{\left[parms_{12} - \left(parms_{14}\right)^2 \cdot parms_{13}\right]}{1 - \left(parms_{13}\right)^2 \cdot \left(parms_{14}\right)^2} \\
&\quad tmp \leftarrow rnorm\left(3 \cdot 365, 0, parms_{15}\right) \\
&\quad tmp_0 \leftarrow CM\_dev_{0,12} \cdot parms_{14} \cdot \frac{parms_{15}}{parms_{16}} + tmp_0 \cdot \left[\sqrt{1 - \left(parms_{14}\right)^2}\right] \\
&\quad \text{for } i \in 1 .. 3 \cdot 365 - 1 \\
&\qquad tmp_i \leftarrow x \cdot CM\_dev_{i,12} + k \cdot tmp_{i-1} + y \cdot tmp_i \\
&\quad tmp
\end{aligned}
$$

Note:

As mentioned earlier, it turns out to be easier to generate the fine mass deviations after generating the coarse mass deviations. The reason is that, in general, we want sequences that have a distinct autocorrelations and are correlated with each other either spatially or just through time. This function takes as an input the coarse mass deviations (although it only uses the center point). The constants *x*, *k*, and *y* are adjusted to yield (approximately) the desired correlations.

3.5 The True Concentrations Generator

Input: The matrix *bchol*, from the function BChol, and *parms*, the vector of simulation parameters.

$$
\begin{aligned}
&\text{Gen\_truth}(bchol, parms) := \\
&\quad X \leftarrow \text{Gen\_CM\_deviations}(bchol, parms) \\
&\quad z \leftarrow \text{Gen\_FM\_deviations}(X, parms) \\
&\quad f25 \leftarrow e^{\frac{-\left[\sqrt{\ln\left[(parms_{15})^2+1\right]}\right]^2}{2}} \\
&\quad f10 \leftarrow e^{\frac{-\left[\sqrt{\ln\left[(parms_{16})^2+1\right]}\right]^2}{2}} \\
&\quad B25 \leftarrow \frac{-1 + parms_{18}}{1 + parms_{18}} \\
&\quad B10 \leftarrow \frac{-1 + parms_{19}}{1 + parms_{19}} \\
&\quad \text{for } j \in 0..3\cdot 365 - 1 \\
&\quad\quad Truth25_j \leftarrow \left(1 + B25\cdot\sin\left(\frac{j\cdot 2\cdot\pi}{365}\right)\right)\cdot\exp\left(z_j\right)\cdot f25 \\
&\quad\quad TruthC^{\langle j\rangle} \leftarrow \left[\left(1 + B10\cdot\sin\left(\frac{j\cdot 2\cdot\pi\cdot parms_{21}}{365} + \frac{parms_{20}}{12}\cdot 2\cdot\pi\right)\right)\cdot\exp\left[\left(X^T\right)^{\langle j\rangle}\right]\cdot f10\right] \\
&\quad trans\_tC \leftarrow TruthC^T \\
&\quad \text{for } i \in 0..24 \\
&\quad\quad tmp1_i \leftarrow \text{Cal\_pt}\left(trans\_tC^{\langle i\rangle}, parms\right) \\
&\quad tm \leftarrow \text{mean}(tmp1) \\
&\quad out \leftarrow \text{augment}\left(Truth25\cdot parms_{17}, trans\_tC^{\langle 12\rangle} + parms_{17}\cdot Truth25\right)\cdot\frac{1}{tm} \\
&\quad out_{\text{rows}(out),0} \leftarrow \frac{\text{mean}(trans\_tC)}{tm} \\
&\quad out
\end{aligned}
$$

Notes:

This function calls the previous two functions and combines them to create the sequences of true concentrations with the appropriate long-term seasonal trends, except that

everything is scaled to have a daily true-design value of 1. The output includes the annual true-design value corresponding to the scaled sequences.

Note that the output only includes the temporal sequences for the center grid and a summary value.

3.6 The Observed Values Generator

Input: The matrix of (scaled) true concentrations from the function Gen_truth and *parms*, the vector of simulation parameters.

$$
\begin{aligned}
&\text{Gen\_obs}(\text{truth}, \text{parms}) := \\
&\quad v1_0 \leftarrow 0 \\
&\quad v2_0 \leftarrow 0 \\
&\quad \text{for } ii \in 0..2 \\
&\quad\quad \text{for } j \in 0..3 \\
&\quad\quad\quad \text{min} \leftarrow \text{floor}\left(\frac{j \cdot 365}{4 \cdot \text{parms}_5}\right) \cdot \text{parms}_5 + ii \cdot 365 \\
&\quad\quad\quad \text{max} \leftarrow \text{floor}\left[\frac{(j+1) \cdot 365}{4 \cdot \text{parms}_5}\right] \cdot \text{parms}_5 + ii \cdot 365 \\
&\quad\quad\quad \text{num\_sampled25} \leftarrow \text{ceil}\left[\frac{(\text{max} - \text{min})}{\text{parms}_5} \cdot \text{parms}_6\right] \\
&\quad\quad\quad \text{num\_sampled10} \leftarrow \text{ceil}\left[\frac{(\text{max} - \text{min})}{\text{parms}_5} \cdot \text{parms}_7\right] \\
&\quad\quad\quad v1 \leftarrow \text{stack}\left[v1, \text{parms}_5 \cdot \text{Select}\left[\frac{(\text{max} - \text{min})}{\text{parms}_5}, \text{num\_sampled25}\right] + \text{min}\right] \\
&\quad\quad\quad v2 \leftarrow \text{stack}\left[v2, \text{parms}_5 \cdot \text{Select}\left[\frac{(\text{max} - \text{min})}{\text{parms}_5}, \text{num\_sampled10}\right] + \text{min}\right] \\
&\quad k \leftarrow 0 \\
&\quad \text{for } ii \in 1..\text{rows}(v1) - 1 \\
&\quad\quad \text{for } jj \in 1..\text{rows}(v2) - 1 \\
&\quad\quad\quad \text{if } v1_{ii} = v2_{jj} \\
&\quad\quad\quad\quad v_k \leftarrow v1_{ii} \\
&\quad\quad\quad\quad k \leftarrow k + 1 \\
&\quad \text{total\_sampled} \leftarrow \text{rows}(v) \\
&\quad e \leftarrow \text{augment}(\text{rnorm}(\text{total\_sampled}, 0, 1), \text{rnorm}(\text{total\_sampled}, 0, 1)) \\
&\quad \text{for } i \in 0..\text{total\_sampled} - 1 \\
&\quad\quad obs_{i,1} \leftarrow \left[\left(\text{truth}_{v_i,1}\right) \cdot \left(1 + \text{parms}_{11}\right) \cdot \left(1 + e_{i,0} \cdot \text{parms}_9\right)\right] - \text{truth}_{v_i,0} \cdot \left(1 - \text{parms}_{10}\right) \cdot \left(1 + e_{i,1} \cdot \text{parms}_8\right) \\
&\quad\quad obs_{i,0} \leftarrow \left[\text{truth}_{v_i,1} \cdot \left(1 - \text{parms}_{11}\right) \cdot \left(1 + e_{i,0} \cdot \text{parms}_9\right)\right] - \text{truth}_{v_i,0} \cdot \left(1 + \text{parms}_{10}\right) \cdot \left(1 + e_{i,1} \cdot \text{parms}_8\right) \\
&\quad\quad obs_{i,0} \leftarrow 0 \text{ if } obs_{i,0} < 0 \\
&\quad\quad obs_{i,1} \leftarrow 0 \text{ if } obs_{i,1} < 0 \\
&\quad \text{augment}(obs, v)
\end{aligned}
$$

Note:

This function takes in the true concentrations for the center grid, selects the days throughout the 3-year period that will have corresponding sample values, and then simulates the observed values. Two versions of the observed values are generated: one corresponding to the maximum positive bias and one corresponding to the most extreme negative bias. The last column gives the number of days into the simulation to which the observed values correspond.

3.7 The Design Values Calculator for the Observations

Input: The matrix *X* of (scaled) observed concentrations from the function Gen_obs, *parms*, the vector of simulation parameters, and the scaling factor, *tm*, for switching between the daily and annual standard.

$$
\begin{aligned}
&\text{Cal\_pobs}(X, parms, tm) := \\
&\quad j \leftarrow 0 \\
&\quad k \leftarrow 0 \\
&\quad \text{for } i \in 0..\text{rows}(X) - 1 \\
&\quad\quad j \leftarrow j + \left(X_{i,2} < 365\right) \\
&\quad\quad k \leftarrow k + \left(X_{i,2} < 365 \cdot 2\right) \\
&\quad r_0 \leftarrow \text{floor}\left(j \cdot parms_4\right) \\
&\quad r_1 \leftarrow \text{floor}\left[(k - j) \cdot parms_4\right] \\
&\quad r_2 \leftarrow \text{floor}\left[(\text{rows}(X) - k) \cdot parms_4\right] \\
&\quad \text{for } i \in 0..1 \\
&\quad\quad p_{0,i} \leftarrow \text{sort}(\text{submatrix}(X, 0, j - 1, i, i))_{(r_0)} \\
&\quad\quad am_{0,i} \leftarrow \text{mean}(\text{submatrix}(X, 0, j - 1, i, i)) \\
&\quad \text{for } i \in 0..1 \\
&\quad\quad p_{1,i} \leftarrow \text{sort}(\text{submatrix}(X, j, k - 1, i, i))_{(r_1)} \\
&\quad\quad am_{1,i} \leftarrow \text{mean}(\text{submatrix}(X, j, k - 1, i, i)) \\
&\quad \text{for } i \in 0..1 \\
&\quad\quad p_{2,i} \leftarrow \text{sort}(\text{submatrix}(X, k, \text{rows}(X) - 1, i, i))_{(r_2)} \\
&\quad\quad am_{2,i} \leftarrow \text{mean}(\text{submatrix}(X, k, \text{rows}(X) - 1, i, i)) \\
&\quad \text{for } i \in 0..1 \\
&\quad\quad m_i \leftarrow \text{mean}\left(p^{\langle i \rangle}\right) \\
&\quad\quad m_{i+2} \leftarrow \text{mean}\left(am^{\langle i \rangle}\right) \cdot \frac{1}{tm} \\
&\quad m
\end{aligned}
$$

Note:

Since the number of days with both PM_{10} and $PM_{2.5}$ data can vary, the function first determines to which year the data values correspond. Then, as with real data, annual values are evaluated for each of the three years. Finally, these are averaged over the three years to obtain the two scaled design values for the simulation for the positive bias case and two scaled design values for the negative bias case.

3.8 The Percentile Standard Calculator for the Truth

Input: A vector *X* of (scaled) true concentrations and the vector *parms*, of the input parameters.

$$\text{Cal\_pt}(X, \text{parms}) := \left| \begin{array}{l} r \leftarrow \text{floor}\left(\text{floor}\left(\frac{\text{rows}(X)}{3}\right)\cdot \text{parms}_4\right) \\ \text{for } i \in 0..2 \\ \quad p_i \leftarrow \text{sort}\left[\text{submatrix}\left[X, i\cdot\text{floor}\left(\frac{\text{rows}(X)}{3}\right), (i+1)\cdot\text{floor}\left(\frac{\text{rows}(X)}{3}\right) - 1, 0, 0\right]\right]_r \\ \text{mean}(p) \end{array} \right.$$

Note:

This function is less complicated than the version for the observations since the data are always complete with the same number of data points per year. Moreover, this completeness property implies that the annual design value is just the mean of all of the values and is not computed by this function.

3.9 Random Selection of Sample Days

Input: Integers *n* and *k*.

$$\text{Select}(n, k) := \left| \begin{array}{l} j \leftarrow 0 \\ \text{for } i \in 0..n-1 \\ \quad \text{if } [\text{rnd}(1)\cdot(n-i) < (k-j)] \\ \quad\quad \left| \begin{array}{l} v_j \leftarrow i \\ j \leftarrow j+1 \end{array} \right. \\ v \end{array} \right.$$

Note:

This function randomly selects *k* integers between 0 and n-1. This is an auxiliary function used in the generation of the observations to randomly pick the days that will have PM_{10} data and which will have $PM_{2.5}$ data (the selections are made independent of each other).

3.10 Decision Performance Calculation

Input: A scalar *x* equal to the desired observed-design value target, an indicator *std_ind* to indicate whether the daily or annual standard is being considered, the matrix output *tt* from the function Gen_scaled_dist, and the vector *parms* of input parameters.

$$
\begin{aligned}
&\text{P\_obdv\_ge\_stnd}(x, \text{std\_ind}, tt, parms) := \\
&\quad \text{if } \text{std\_ind} = 0 \\
&\quad\quad \text{std} \leftarrow parms_2 \\
&\quad\quad \text{round} \leftarrow 0.5 \\
&\quad \text{otherwise} \\
&\quad\quad \text{std} \leftarrow parms_3 \\
&\quad\quad \text{round} \leftarrow 0.05 \\
&\quad mm \leftarrow tt \cdot x \\
&\quad rr_0 \leftarrow 0 \\
&\quad rr_1 \leftarrow 0 \\
&\quad \text{for } k \in 0..\text{reps} - 1 \\
&\quad\quad rr_0 \leftarrow \left[mm_{k,\,0+\text{std\_ind}} > (\text{std} + \text{round}) \right] + rr_0 \\
&\quad\quad rr_1 \leftarrow \left[mm_{k,\,1+\text{std\_ind}} > (\text{std} + \text{round}) \right] + rr_1 \\
&\quad rr \leftarrow rr \cdot \frac{1}{\text{reps}} \\
&\quad rr
\end{aligned}
$$

Note:

This function computes the probability that the observed-design value exceeds the standard given that the design value of the true concentrations is equal to *x*.

3.11 Gray Zone and Performance Curve Computation

Input: The vector *parms* of input parameters.

```
Gray_zone(parms) := | tt ← Gen_scaled_dist (parms)
                    | out_day ← Gray_zone_day(parms, tt)
                    | out_ann ← Gray_zone_ann(parms, tt)
                    | out ← augment(out_day , out_ann )
                    | out
```

Note:

In addition to computing the gray zones, this function computes the values for plotting the performance curves.

The final output consists of a matrix with 102 rows and 6 columns. The last row has the gray zones for the daily and annual standard. The other rows contain the values for plotting the decision performance curves. The first two columns are the probabilities of observing design values greater than the daily standard for the positive and negative bias cases. The third column consists of the daily true-design values for the grid. The fourth and fifth columns are the probabilities of observing design values greater than the annual standard for the positive and negative bias cases. The sixth column consists of the annual true-design values for the grid.

3.11.1 Gray Zone and Performance Curve Computation for the Daily Standard

Input: The vector *parms* of input parameters and the output of the Gen_scaled_dist function.

$$
\begin{aligned}
&\text{Gray\_zone\_day}(\text{parms}, \text{tt}) := \\
&\quad \text{min2} \leftarrow \min\left(\text{tt}^{\langle 0 \rangle}\right) \\
&\quad \text{max2} \leftarrow \max\left(\text{tt}^{\langle 1 \rangle}\right) \\
&\quad \text{start\_day} \leftarrow \text{floor}\left(\frac{\text{parms}_2 - .0}{\text{max2}}\right) \\
&\quad \text{end\_day} \leftarrow \text{ceil}\left(\frac{\text{parms}_2 + .75}{\text{min2} + .001}\right) \\
&\quad \text{delta2} \leftarrow \frac{\text{end\_day} - \text{start\_day}}{100} \\
&\quad \text{for } k \in 0..100 \\
&\quad\quad \text{ttest}^{\langle k \rangle} \leftarrow \text{P\_obdv\_ge\_stnd}(\text{start\_day} + k \cdot \text{delta2}, 0, \text{tt}, \text{parms}) \\
&\quad\quad v_k \leftarrow \text{start\_day} + k \cdot \text{delta2} \\
&\quad L \leftarrow 0 \\
&\quad u \leftarrow 0 \\
&\quad \text{for } i \in 0..\text{cols}(\text{ttest}) - 2 \\
&\quad\quad L \leftarrow i \text{ if } \text{ttest}_{1,i} < \text{parms}_0 \\
&\quad\quad u \leftarrow i \text{ if } \text{ttest}_{0,i} < \left(1 - \text{parms}_1\right) \\
&\quad \text{out\_day} \leftarrow (\text{start\_day} \quad \text{end\_day}) \text{ if } \left(\text{ttest}_{1,L+1} - \text{ttest}_{1,L}\right) \cdot \left(\text{ttest}_{0,u+1} - \text{ttest}_{0,u}\right) = 0 \\
&\quad \text{out\_day} \leftarrow \text{start\_day} + \left[L + \left(\frac{\text{parms}_0 - \text{ttest}_{1,L}}{\text{ttest}_{1,L+1} - \text{ttest}_{1,L}}\right) \quad u + \left[\frac{\left(1 - \text{parms}_1\right) - \text{ttest}_{0,u}}{\text{ttest}_{0,u+1} - \text{ttest}_{0,u}}\right]\right] \cdot \text{delta2} \quad \text{otherwise} \\
&\quad \text{plot\_data} \leftarrow \text{augment}\left(\text{ttest}^{T}, v\right) \\
&\quad \text{out\_day}_{0,2} \leftarrow 0 \\
&\quad \text{stack}(\text{plot\_data}, \text{out\_day})
\end{aligned}
$$

3.11.2 Gray Zone and Performance Curve Computation for the Annual Standard

Input: The vector *parms* of input parameters and the output of the Gen_scaled_dist function.

$$
\begin{aligned}
&\text{Gray\_zone\_ann}(parms, tt) := \\
&\quad min1 \leftarrow \min\left(tt^{\langle 2 \rangle}\right) \\
&\quad max1 \leftarrow \max\left(tt^{\langle 3 \rangle}\right) \\
&\quad start\_ann \leftarrow \text{floor}\left(\frac{parms_3 - .0}{max1}\right) \\
&\quad end\_ann \leftarrow \text{ceil}\left(\frac{parms_3 + .5}{min1 + .001}\right) \\
&\quad delta \leftarrow \frac{end\_ann - start\_ann}{100} \\
&\quad \text{for } k \in 0..100 \\
&\quad\quad test^{\langle k \rangle} \leftarrow P\_obdv\_ge\_stnd\ (start\_ann + k \cdot delta, 2, tt, parms) \\
&\quad\quad v_k \leftarrow start\_ann + k \cdot delta \\
&\quad L \leftarrow 0 \\
&\quad u \leftarrow 0 \\
&\quad \text{for } i \in 0..\text{cols}(test) - 2 \\
&\quad\quad L \leftarrow i \text{ if } test_{1,i} < parms_0 \\
&\quad\quad u \leftarrow i \text{ if } test_{0,i} < \left(1 - parms_1\right) \\
&\quad out\_ann \leftarrow (start\_ann \quad end\_ann) \text{ if } \left(test_{1,L+1} - test_{1,L}\right) \cdot \left(test_{0,u+1} - test_{0,u}\right) = 0 \\
&\quad out\_ann \leftarrow start\_ann + \left[L + \left(\frac{parms_0 - test_{1,L}}{test_{1,L+1} - test_{1,L}}\right) \quad u + \left[\frac{\left(1 - parms_1\right) - test_{0,u}}{test_{0,u+1} - test_{0,u}}\right]\right] \cdot delta \quad \text{otherwise} \\
&\quad out\_ann_{0,2} \leftarrow 0 \\
&\quad plot\_data \leftarrow \text{augment}\left(test^T, v\right) \\
&\quad \text{stack}(plot\_data, out\_ann)
\end{aligned}
$$