DEFINITIONS and STATISTICAL TESTS

Building-physical definitions

R is the *thermal resistance* (surface-to-surface) defined as the difference between the two surface temperatures in steady state divided by the density of heat flow rate, in °Cm^2/W

C_i is the *internal thermal capacity* per unit area of the wall defined as the amount of heat that goes into the wall per m^2 as the result of a change from one steady state situation to another by increasing the internal surface temperature with 1°*C*, in $Wh/°Cm^2$

C_e is the *external thermal capacity* per unit area of the wall defined as the amount of heat which goes into the wall per m^2 as the result of a change from one steady state situation to another by increasing the external surface temperature with 1°*C*, in $Wh/°Cm^2$

C is the *thermal capacity* per unit area of the wall defined as the sum $C_i + C_e$, in $Wh/°Cm^2$

Mathematical definitions

The wall may be considered as a constant parameter linear system with the density of heat flow rate at the internal surface as the output variable and the two surface temperatures as input variables. In continuous time the relationship may be written as

$$q_i(t) = \int_0^\infty w_1(t) q_i(t-t)dt - \int_0^\infty w_2(t) q_e(t-t)dt \qquad (2)$$

where $w_1(t)$ and $w_2(t)$ are weighting functions.

The thermal resistance *R* may defined by:

$$\frac{1}{R} = \int_0^\infty w_1(t)dt = \int_0^\infty w_2(t)dt \qquad (3)$$

In the case of a 'symmetric' wall, the heat flow density at the external surface may be written as

$$q_e(t) = \int_0^\infty w_2(t) q_i(t-t)dt - \int_0^\infty w_1(t) q_e(t-t)dt \qquad (4)$$

The thermal capacity *C* may in this case be defined by the following equation

$$C = C_i + C_e = 2\int_{t=0}^{\infty} \int_{t=0}^{t} \{w_1(t) - w_2(t)\} dt dt \qquad (5)$$

STATISTICAL TESTS

To obtain a correct evaluation the organisers have defined the parameters and tests as given below. The following tests are considered:

the **t-test**, which is used for test of unbiased estimates, i.e. to test for correct mean value of the estimates.

the **F-test**, which is used to test whether the variance of the parameter estimates as provided by the estimation method is equal to the empirical variance between stochastic independent runs.

the c^2 **test**, which is used to test whether the ratio between the observed deviation and the estimated standard deviation of the parameter estimates is reasonable.

x_r : the real value x_r is the value that the organisers have used to generate the data

x_e : the estimated value x_e is a point estimate of x_r given by the participant

s_e^2 : the variance of the estimate, also given by the participant

The empirical variance of N point estimates is defined by:

$$s_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_{e,i} - \bar{x}_e\right)^2 \tag{6}$$

where $x_{e,i}$ is the point estimate in data series *i* and $\bar{x}_e$ is the empirical mean, i.e.

$$\bar{x}_e = \frac{1}{N}\sum_{i=1}^{N}\left(x_{e,i}\right) \tag{7}$$

N corresponds to the total number of data series.
The mean of the estimated variance of the parameter estimates equals :

$$\bar{s}^2 = \frac{1}{N}\sum_{i=1}^{N} s_{e,i}^2 \tag{8}$$

with $s_{e,i}^2$ defined as the variance of the estimate in data series *i*

t-test
In order to test whether the estimates are unbiased (have the correct mean value) the t-test quantity is defined:

$$z_t = \frac{\left|\bar{x}_e - x_r\right|}{s_x/\sqrt{N}} = \frac{\sqrt{\left(\bar{x}_e - x_r\right)^2}}{\sqrt{\frac{1}{N(N-1)}\sum_{i=1}^{N}\left(x_{e,i} - \bar{x}_e\right)^2}} \tag{9}$$

It can be stated that $z_t \in t(N-1)$. This means that the critical region for testing on level a is $\{z > t(N-1)_{1-a/2}\}$. Which results that for N=20 the following critical regions are obtained:

a = 20%	z >1.328
a = 10%	z >1.729
a = 5%	z >2.093
a = 1%	z >2.861
a = 0.1%	z >3.883

F-test

In order to test whether the variance of the parameters provided by the estimation tool is equal to the empirical variance the F-test quantity is defined. The F-statistic is given by:

$$z_f = \frac{s_x^2}{\bar{s}^2} = \frac{\frac{1}{N-1}\sum_{i=1}^{N}\left(x_{e,i} - \bar{x}_e\right)^2}{\frac{1}{N}\sum_{i=1}^{N} s_{e,i}^2} \tag{10}$$

Since the number of observations used in the calculation of $s_{e,i}^2$ is large, the situation exists that $\{ z_F \in F(N-1,\infty) \}$. This means that the critical region for testing on level a is $\{ z < F(N-1,\infty)_{a/2}$ V $z > F(N-1,\infty)_{1-a/2} \}$.

This means that for N=20 the following critical regions are obtained:

a = 20%	z <0.613 V z >1.432
a = 10%	z <0.532 V z >1.586
a = 5%	z <0.469 V z >1.729
a = 1%	z <0.372 V z >2.031
a = 0.1%	z <0.258 V z >2.420

The c^2 test

For cases were only one data series is available, another test quantity is necessary. The c^2 test is used to test for a reasonable ratio between the deviation and the estimated standard deviation of the parameters and hence to obtain a quantity to evaluate the estimated standard deviation.

The c^2 test quantity as defined by:

$$z_{c^2} = \frac{(x_e - x_r)^2}{s_e^2} \tag{11}$$

Since the number of observations in each test is large, the approximation is used that $z_{c^2} \in c^2(1)$. The critical region for testing on level a is $\{ z < c^2(1)_{a/2}$ V $z > c^2(1)_{1-a/2} \}$.

For level a = 1% and 20% the critical regions are:

a =20%	z <0.016 V z >2.706
a =1%	z <0.001 V z >7.879

Prediction evaluation

L : the time period of prediction measured in data points

$y_{r,t}$: the real value at time t

$y_{e,t}$: the predicted value at time t

The accuracy of the predicted data series can be calculated by three statistical measures; the RMSE, the CV and the NMBE. The RMSE, the root mean square error, can be calculated from :

$$RMSE = \sqrt{\frac{\sum_{t=1}^{L}\left(y_{e,t} - y_{r,t}\right)^2}{L}} \qquad (12)$$

The CV, the coefficient of variation, is a dimensionless measure of the RMS error:

$$CV = \frac{RMSE}{\bar{y}_r} \qquad (13)$$

where $\bar{y}_r$ is the average of the L data points.

The NRMS, the normalised root mean square error is :

$$NRMSE = \frac{RMSE}{\sqrt{s_r^2}} \qquad (14)$$

where s_r^2 is the empirical variance of the L data points

Finally the NMBE, the normalised mean bias error, is a dimensionless estimate of the bias of the prediction :

$$NMBE = \frac{\sum_{t=1}^{L}\left(y_{e,t} - y_{r,t}\right)}{\sqrt{s_r^2}} \qquad (15)$$

In general the rule is that the smaller these values are, the better that the prediction has performed and approximates the real data series. A rule of thumb is: a NRMSE value less than 50% performs reasonable. For the NMBE a value of 8% can be applied for a reasonable result.