LCC sample treatment saturation statistical evaluation

Test sample saturation (%) sorted by treatments:

AD = Air-dried
H_{100} = High humidity environment for one week
M_5 = submerged in water for 5 minutes then high humidity environment for 1 week
D_1 = submerged in water 1 day then high humidity environment for 1 week
W_1 = submerged in water 1 week then high humidity environment for 1 week

AD	H_{100}	M_5	D_1	W_1
2.955	4.449	11.907	15.986	17.531
3.091	4.441	9.783	15.504	18.374
3.067	4.275	10.281	15.480	20.368
3.122	4.227	9.679	15.802	19.438
3.620	4.981	10.444	17.123	18.652
3.716	4.647	10.055	17.533	16.078
3.541	4.967	9.884	15.493	15.842
3.455	4.996	9.860	15.764	17.084
3.449	5.033	9.889	17.973	18.536
3.329	5.100	9.389	15.708	18.283
4.014	5.288	9.142	15.832	17.781
3.825	5.112	9.293	17.547	18.503

Sample means:

$$\mu_{AD} := \text{mean}(AD) = 3.432$$

$$\mu_{H100} := \text{mean}(H_{100}) = 4.793$$

$$\mu_{M5} := \text{mean}(M_5) = 9.967$$

$$\mu_{D1} := \text{mean}(D_1) = 16.312$$

$$\mu_{W1} := \text{mean}(W_1) = 18.039$$

Population standard deviations:

$\sigma_{AD} := \text{stdev}(AD) = 0.317$

$\sigma_{H100} := \text{stdev}(H_{100}) = 0.349$

$\sigma_{M5} := \text{stdev}(M_5) = 0.690$

$\sigma_{D1} := \text{stdev}(D_1) = 0.900$

$\sigma_{W1} := \text{stdev}(W_1) = 1.235$

Histogram data:

$n_{bins} := 7$

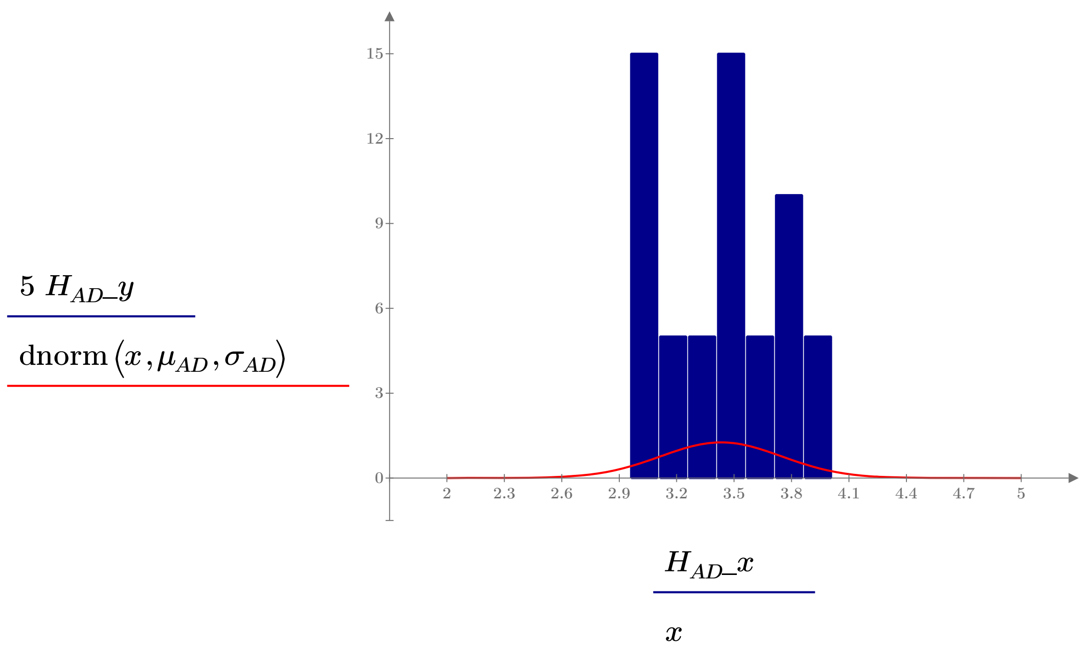

Figure 1. Histogram and normal probability distribution plot of saturation for data set AD.

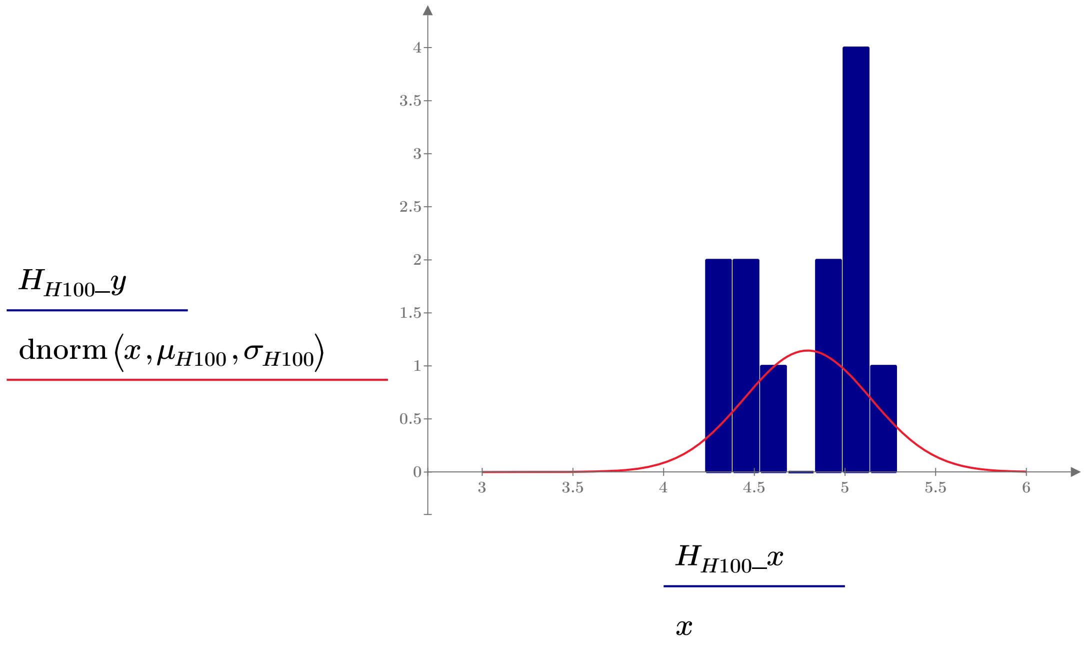

Figure 2. Histogram and normal probability distribution plot of saturation for data set H_{100}.

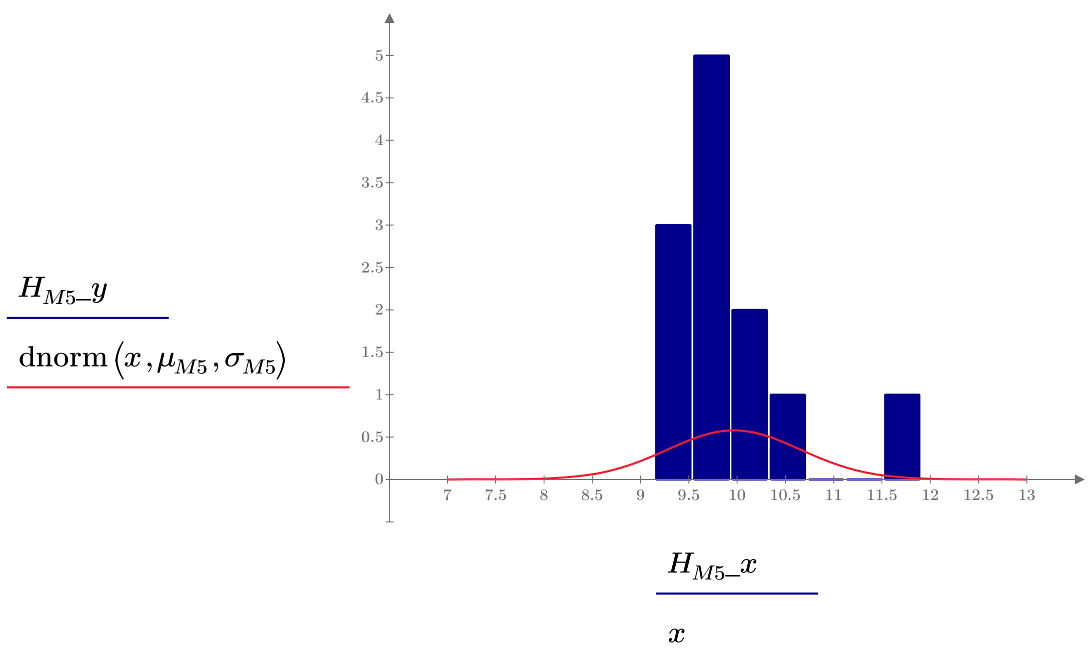

Figure 3. Histogram and normal probability distribution plot of saturation for data set M_5.

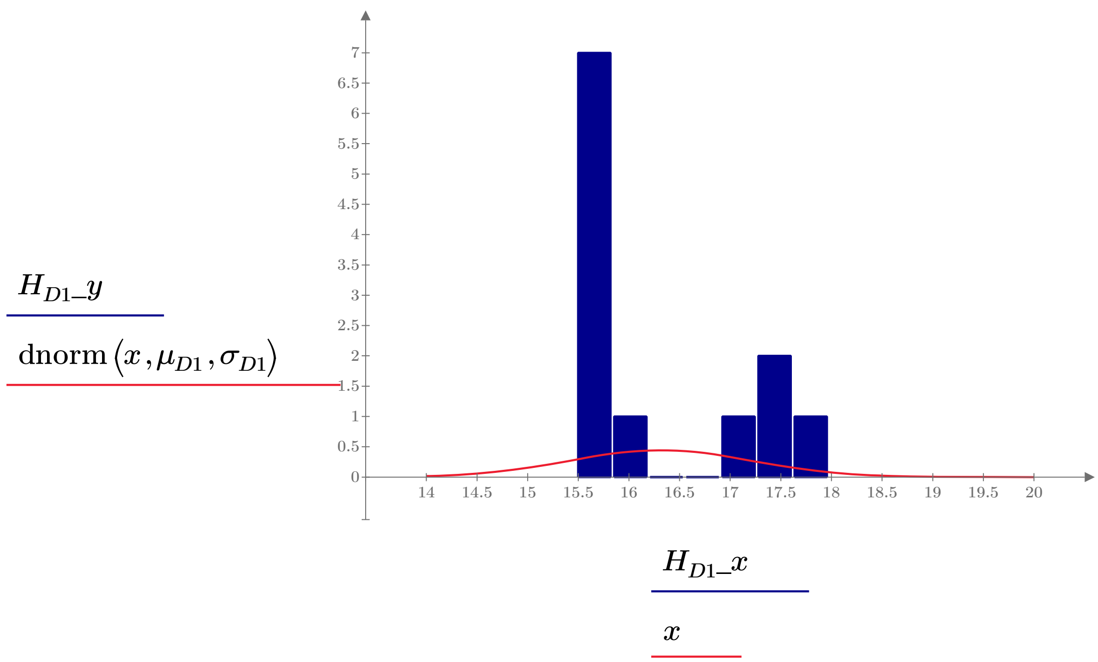

Figure 4. Histogram and normal probability distribution plot of saturation for data set D_1.

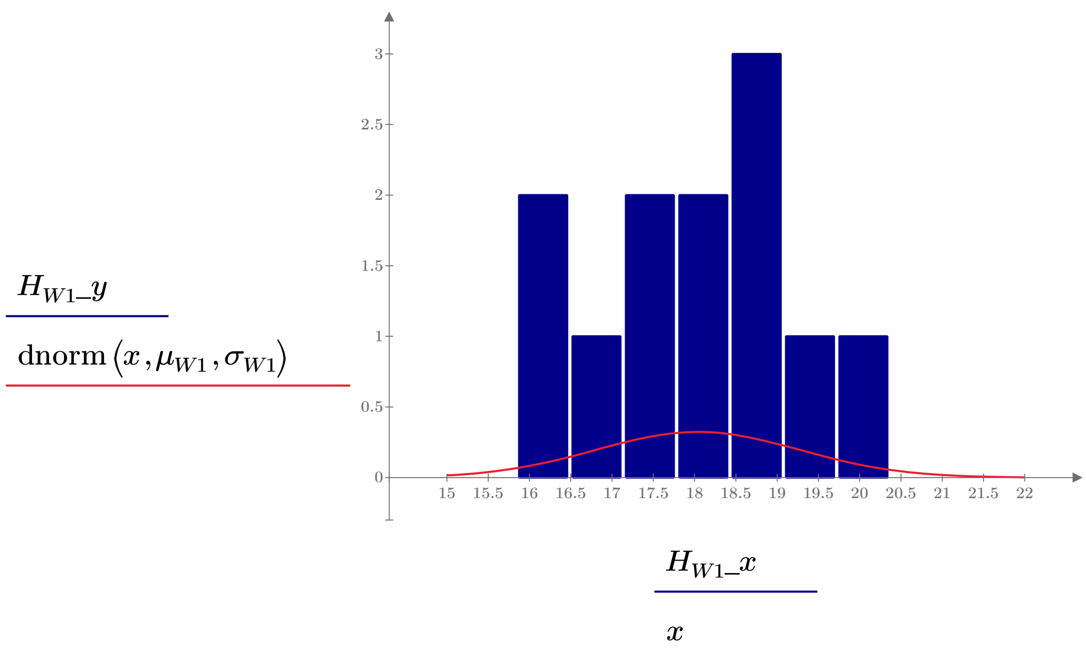

Figure 5. Histogram and normal probability distribution plot of saturation for data set W_1.

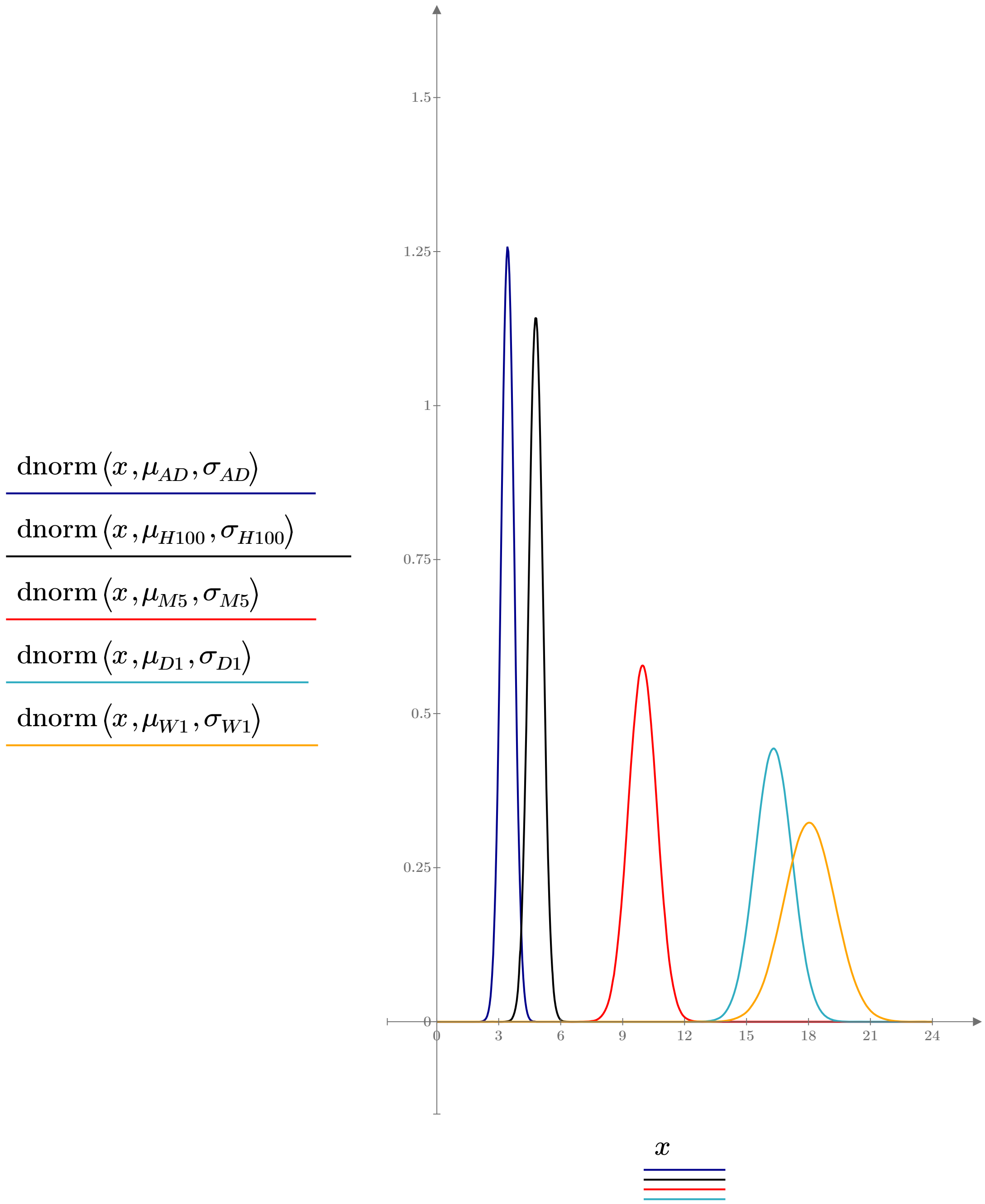

Figure 6. Composite normal probability distribution plots of saturation for all data sets. Note the variability increases with the degree of saturation.

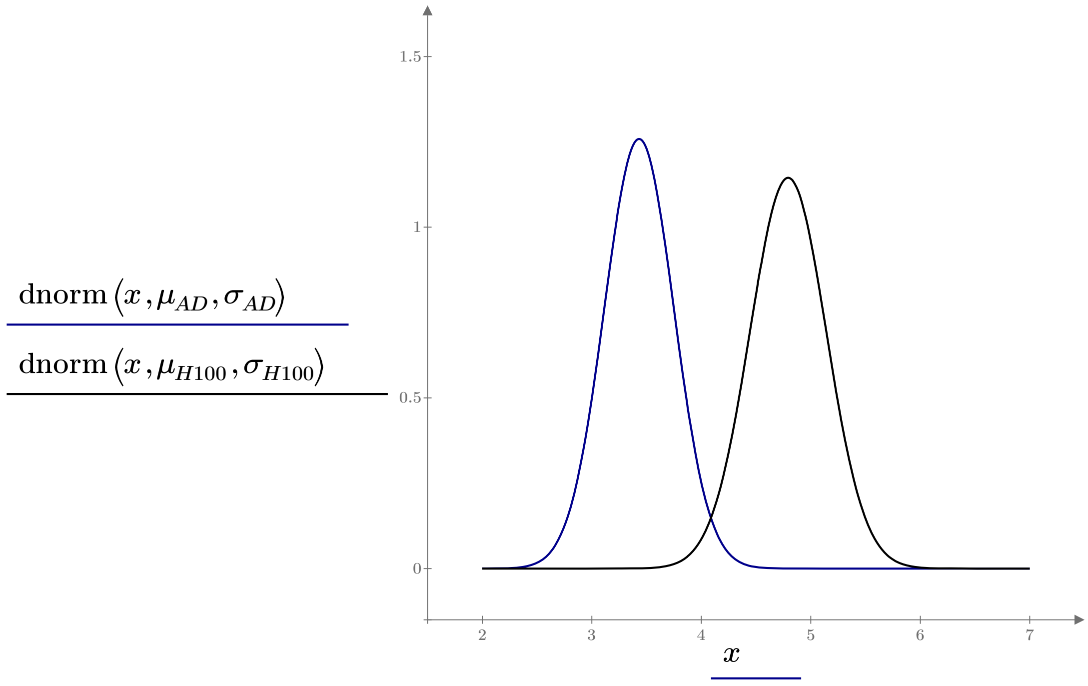

Figure 7. Comparison of the normal distribution plots of AD and H_{100}.

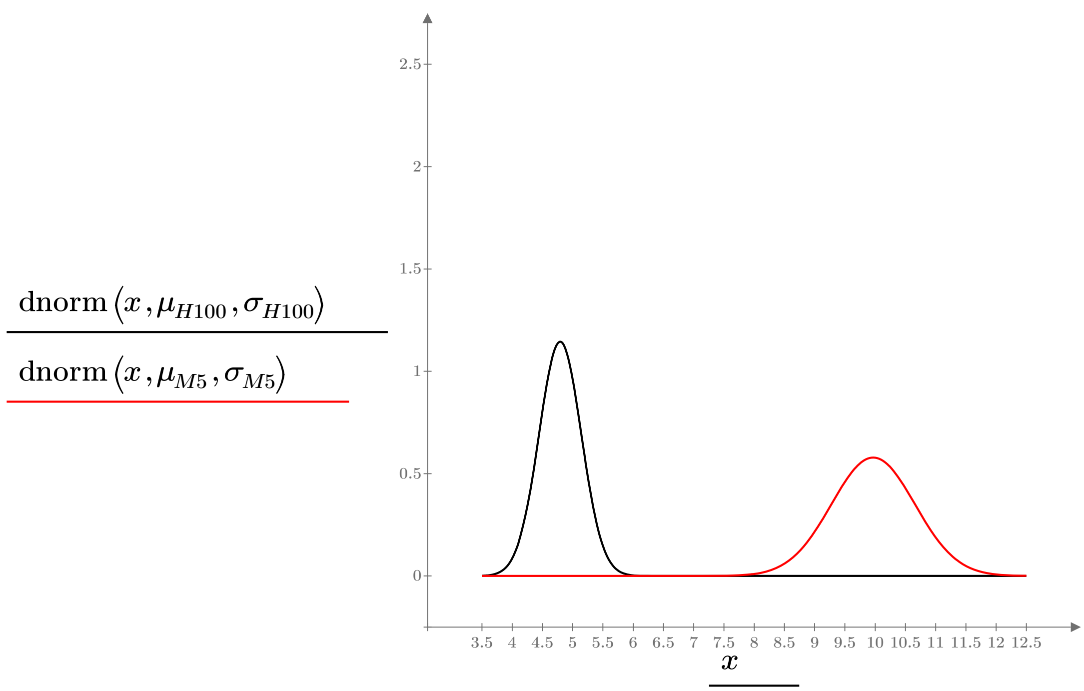

Figure 8. Comparison of the normal distribution plots of H_{100} and M_5.

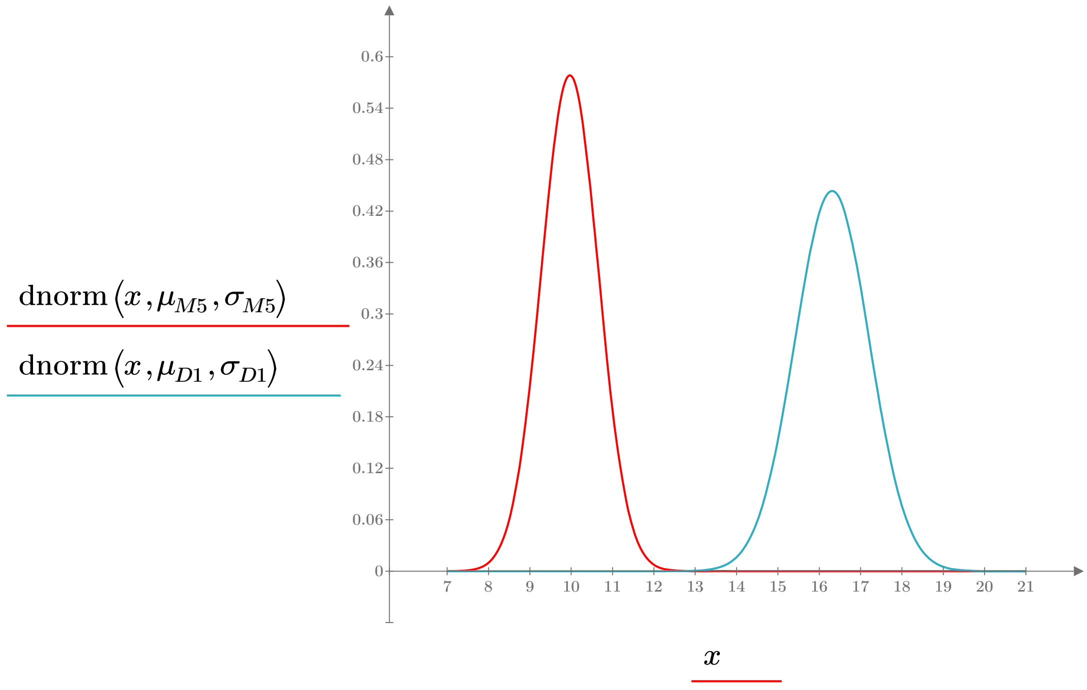

Figure 9. Comparison of the normal distribution plots of M_5 and D_1.

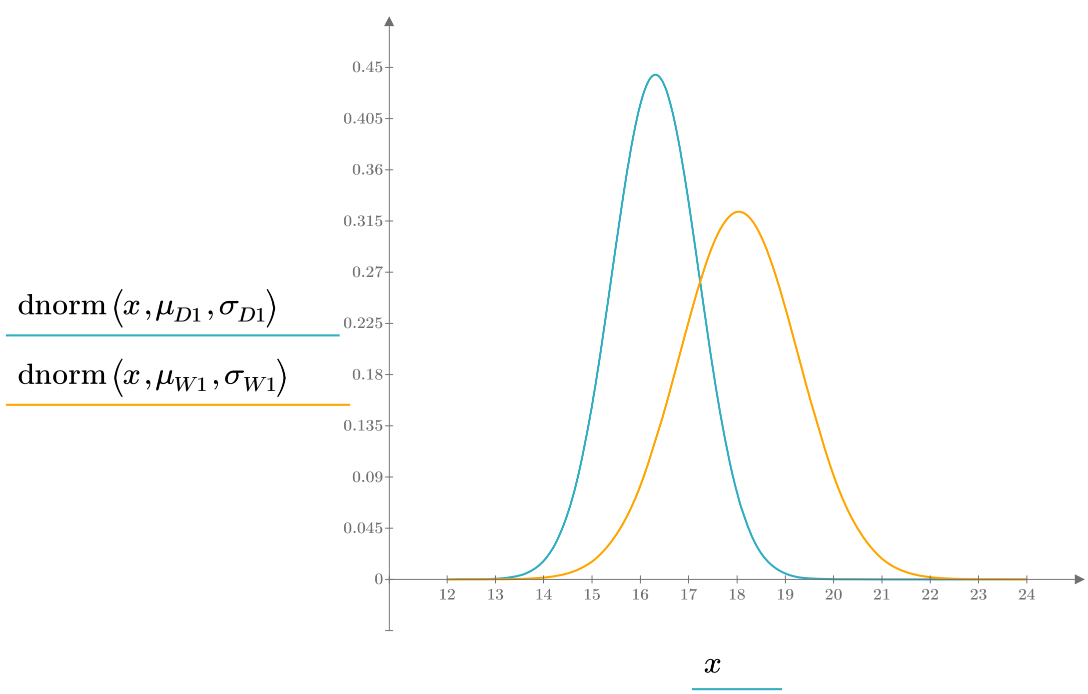

Figure 10. Comparison of the normal distribution plots of D_1 and W_1. Additional analysis is warranted to determine if there is a significant difference between the means of the two sets.

Box plot data:

$$BP_{AD} := \text{boxplot}(AD) = \begin{bmatrix} 3.099 \\ 3.452 \\ 3.692 \\ 2.955 \\ 4.014 \end{bmatrix}$$

25% quartile
50% quartile
75% quartile
Minimum of data set
Maximum of the data set

$$BP_{H100} := \text{boxplot}(H_{100}) = \begin{bmatrix} 4.443 \\ 4.974 \\ 5.083 \\ 4.227 \\ 5.288 \end{bmatrix}$$

$$BP_{M5} := \text{boxplot}(M_5) = \begin{bmatrix} 9.462 \\ 9.872 \\ 10.225 \\ 9.142 \\ 10.444 \\ 11.907 \end{bmatrix}$$

$$BP_{D1} := \text{boxplot}(D_1) = \begin{bmatrix} 15.555 \\ 15.817 \\ 17.431 \\ 15.48 \\ 17.973 \end{bmatrix}$$

$$BP_{W1} := \text{boxplot}(W_1) = \begin{bmatrix} 17.196 \\ 18.329 \\ 18.623 \\ 15.842 \\ 20.368 \end{bmatrix}$$

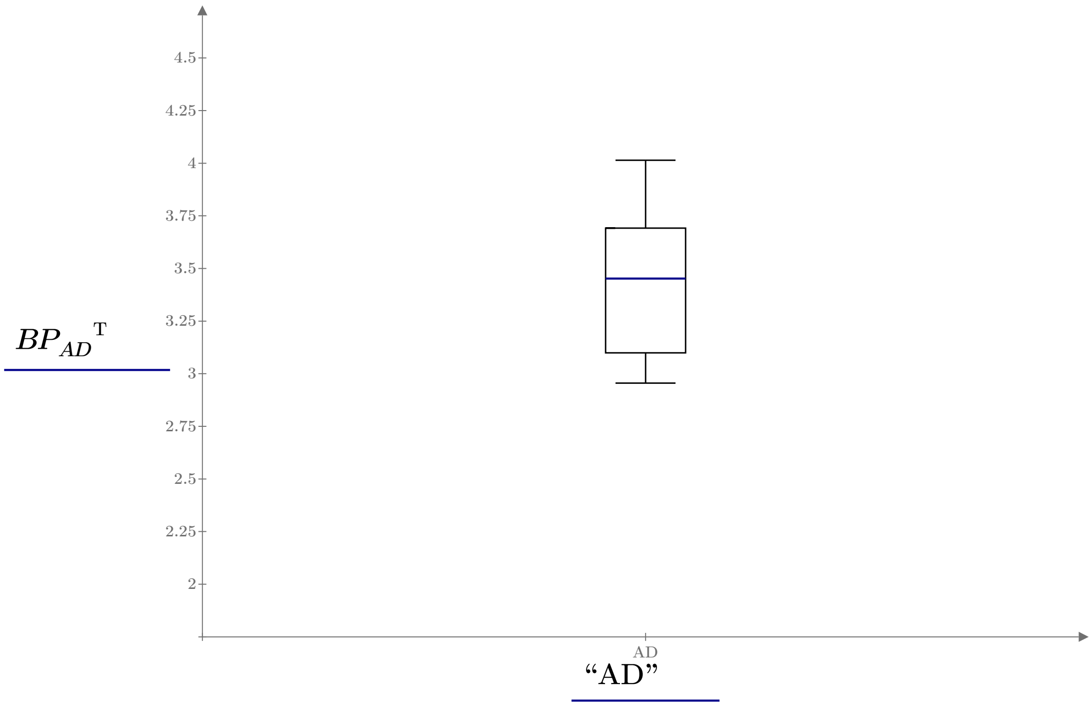

Figure 10. Box plot of the saturation for data set AD.

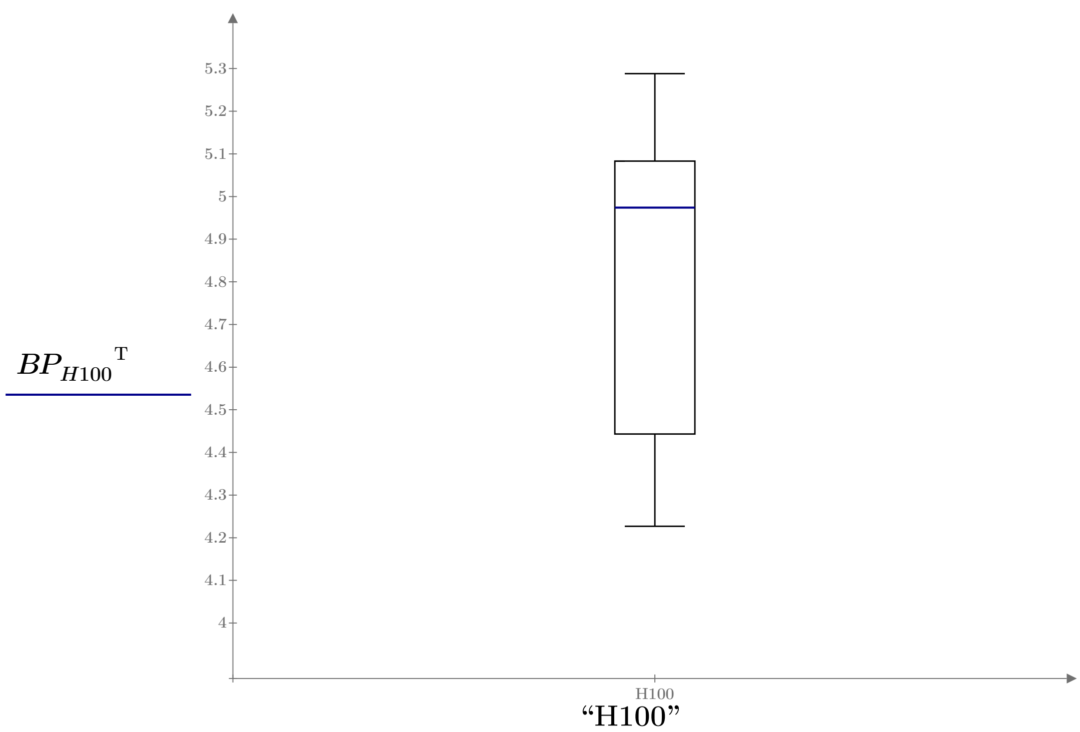

Figure 11. Box plot of the saturation for data set H_{100}.

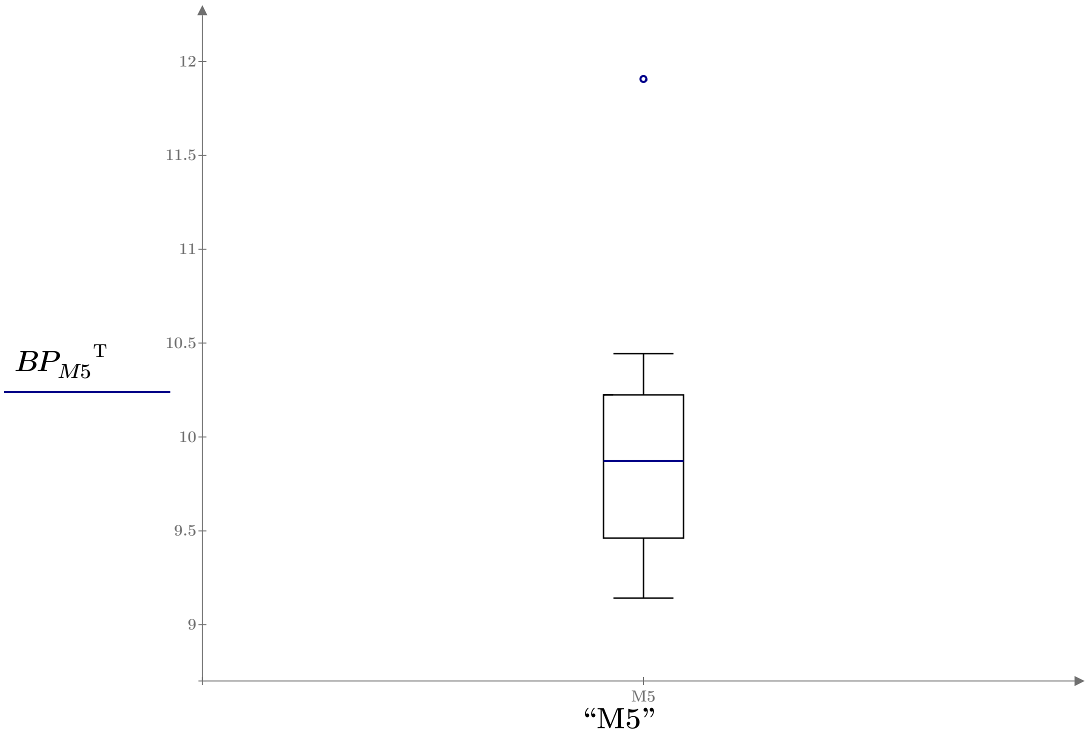

Figure 12. Box plot of the saturation for data set M_5.

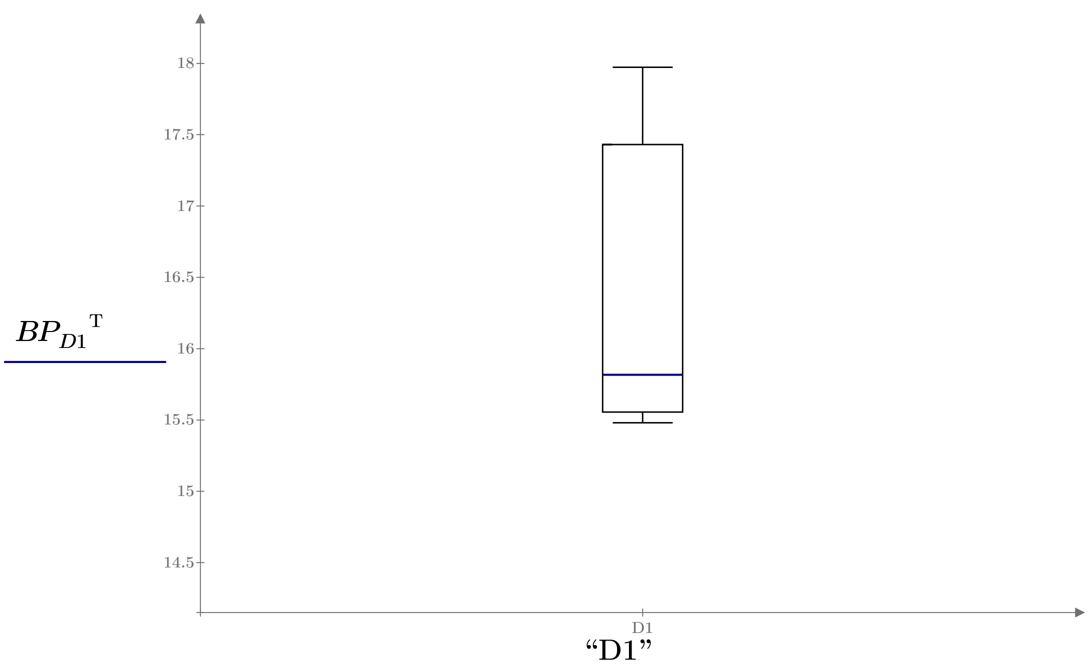

Figure 13. Box plot of the saturation for data set D_1.

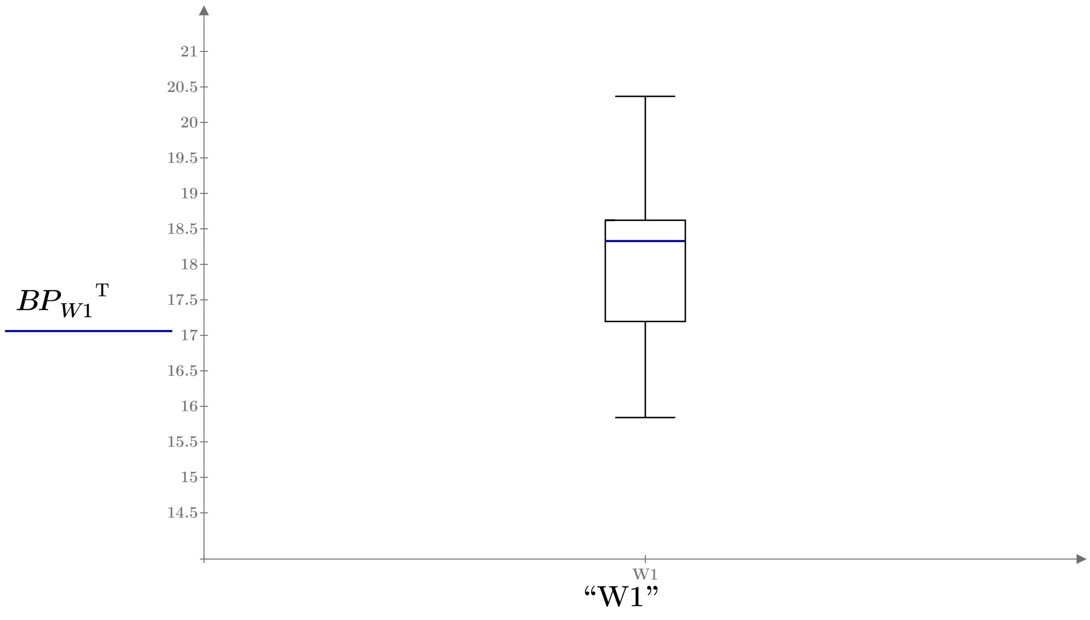

Figure 14. Box plot of the saturation for data set W_1.

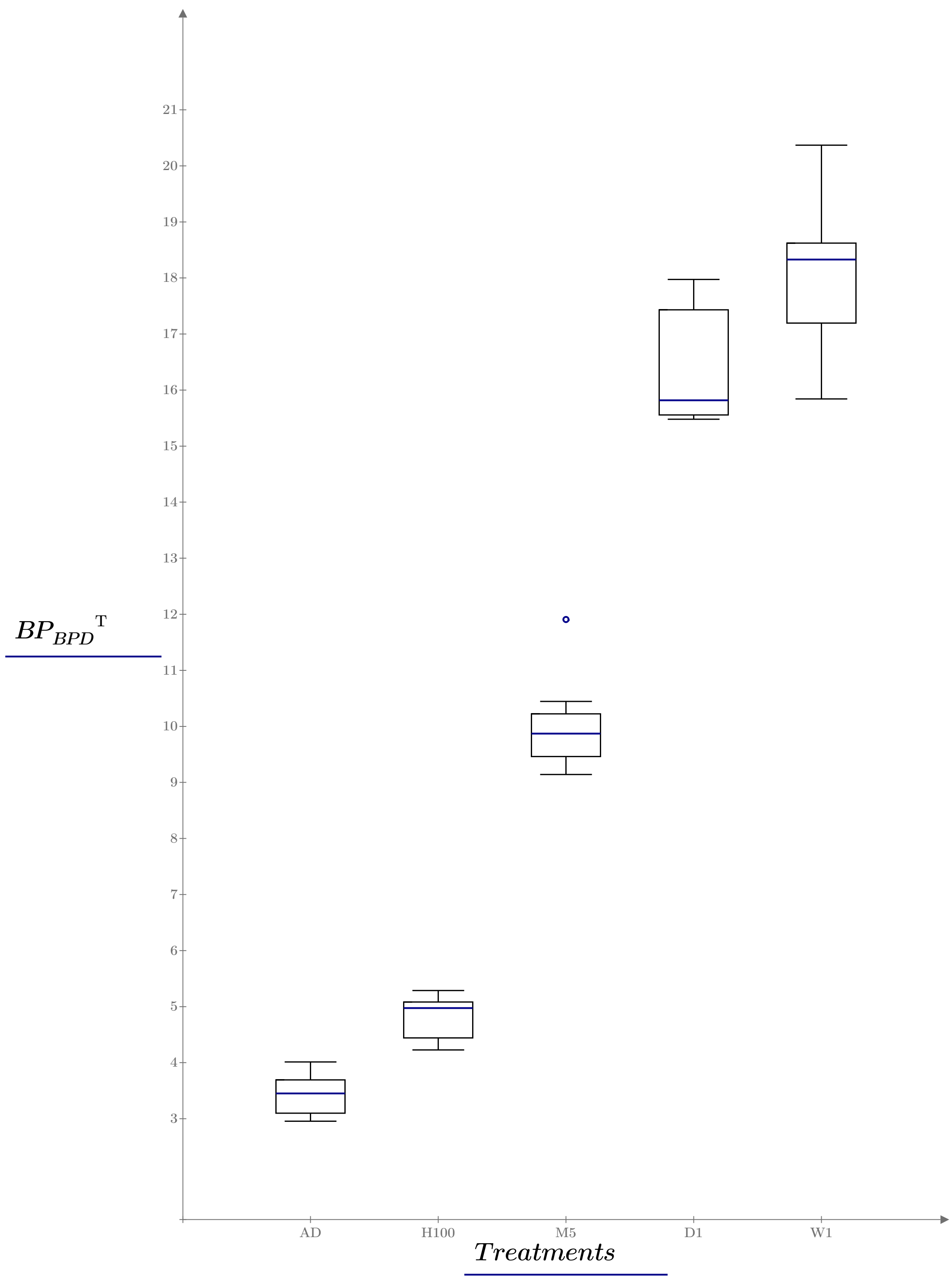

Figure 15. Comparison box plot of the saturation for all data sets.

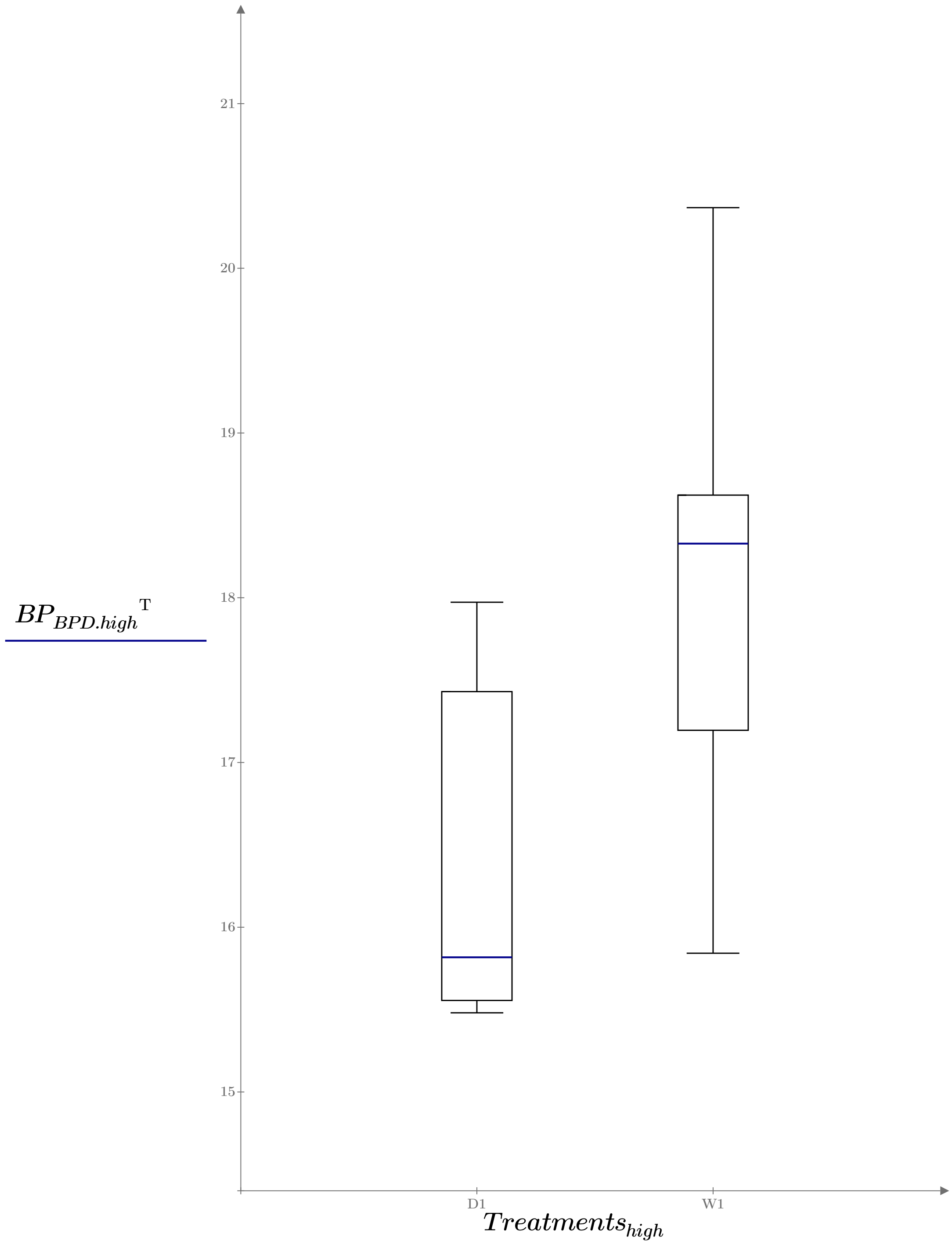

Figure 16. Comparison box plot of the saturation for data sets D_1 and W_1. Additional analysis is warranted to determine if there is a significant difference between the means of the two sets.

Upon inspection of the histograms, probability distribution curves and the box plots, further inspection is warranted to evaluate whether or not there is a difference between group D_1 and group W_1.

Hypothesis testing (F-test):

Since it is evident that the variance increases with an increase in saturation due to the treatments, an initial F-test is warranted to test for equal variance in the two data sets to decide on a pooled-variance t-test or an independent variance t-test is appropriate:

Null and alternate hypothesis:

H0: There is no statistically significant difference between the variance of the two data sets.
H1: There is a statistically significant difference between the variance of the two data sets.

$\alpha := 0.05$ Significance level

$DF_{D1} := \text{length}(D_1) - 1 = 11$ Degrees of freedom

$DF_{W1} := \text{length}(W_1) - 1 = 11$ Degrees of freedom

$F := \text{Ftest}(D_1, W_1) = \begin{bmatrix} 1.885 \\ 0.308 \end{bmatrix}$ F-test statistic
Probability the statistic is larger

$crit_F := \text{qF}(1 - \alpha, DF_{D1}, DF_{W1}) = 2.818$

$F_0 > crit_F = 0$ Hypothesis test, reject H0, accept H1

Since there is statistical evidence of a difference between the variance between the two data sets, an independent variance t-test is warranted.

<u>Hypothesis testing (t-test, independent variance):</u>

Null and alternate hypothesis:

H0: There is no significant difference between the two samples.
H1: m1 < m2 (means)

$n1 := \mathrm{length}\left(D_1\right) = 12$

$n2 := \mathrm{length}\left(W_1\right) = 12$

No. samples for each data set

$m1 := \mathrm{mean}\left(D_1\right) = 16.312$

$m2 := \mathrm{mean}\left(W_1\right) = 18.039$

Sample Means

$s1 := \mathrm{stdev}\left(D_1\right) \cdot \sqrt{\dfrac{n1}{n1-1}} = 0.94$

$s2 := \mathrm{stdev}\left(W_1\right) \cdot \sqrt{\dfrac{n1}{n1-1}} = 1.29$

Sample standard deviations

$\nu := n1 + n2 - 2 = 22$ Degrees of freedom

$w1 := \dfrac{s1^2}{n1} = 0.074$ Weight factor 1

$w2 := \dfrac{s2^2}{n2} = 0.139$ Weight factor 2

$t := \dfrac{|m1 - m2|}{\sqrt{\dfrac{s1^2}{n1} + \dfrac{s2^2}{n2}}} = 3.749$ Test statistic (t-value)

$crit_{D1} := \left|\mathrm{qt}\left(\dfrac{\alpha}{2}, n1 - 1\right)\right| = 2.201$ Critical value, D_1

$crit_{W1} := \left|\mathrm{qt}\left(\dfrac{\alpha}{2}, n2 - 1\right)\right| = 2.201$ Critical value, W_1

$$crit_t := \frac{w1 \cdot crit_{D1} + w2 \cdot crit_{W1}}{w1 + w2} = 2.201$$ Weighted critical value.

$t > crit_t = 1$ Hypothesis test, reject H0, accept H1

Since the t-value is greater than the critical value, we reject the null hypothesis, H0, and accept the alternate hypothesis, H1.

There *is* statistical evidence that m1 < m2.

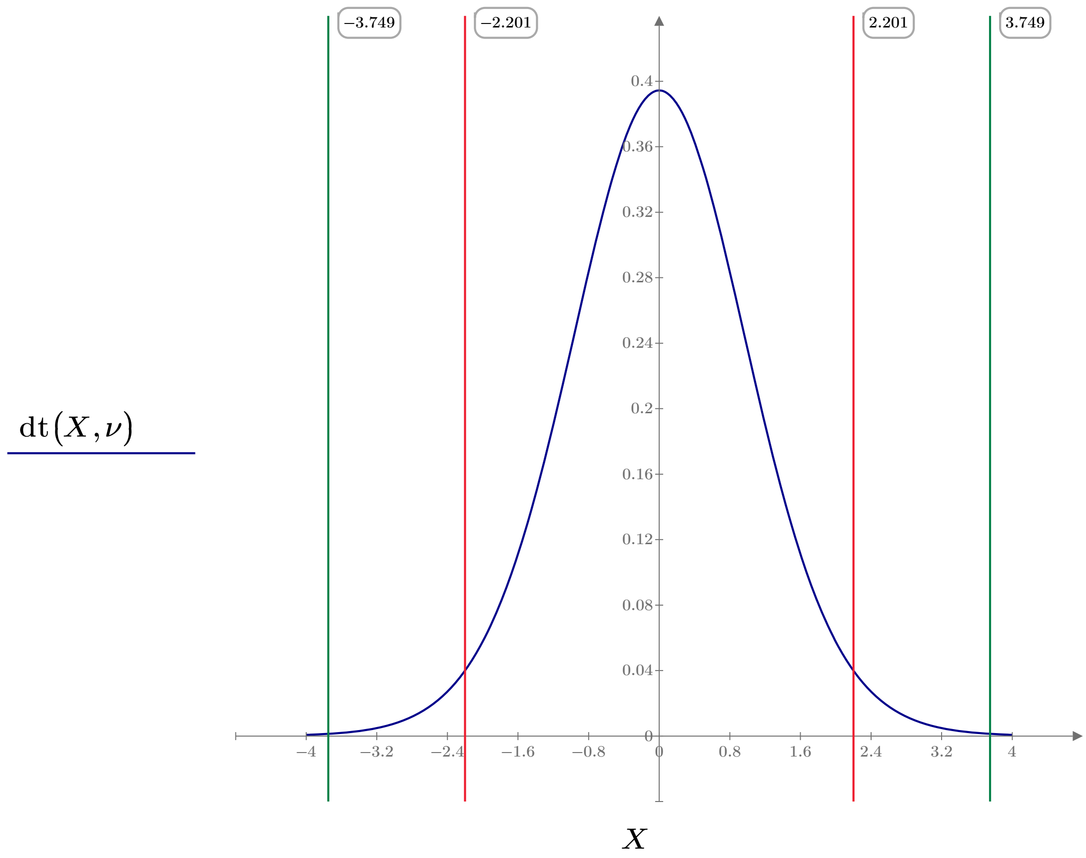

Figure 17. Plot of the critical values and t-statistic values for a t-distribution. Since the t-statistic falls outside of the 95% confidence area, Ho is rejected and H1 is accepted, therefore therefore there is statistical evidence that m1 < m2 and there is a difference in the means for D_1 and W_1 data sets.