Questionable mirrored S-curves used in literature on crop yield relations with soil salinity to determine salt tolerance of crops

R.J. Oosterbaan, February 2021. On www.waterlog.info public domain

Abstract
In some articles published in journals peculiar (mirrored) S-curve relations of crop yield and soil salinity are present. They are used to detect the salt tolerance of crops. The weaknesses are, among other (1) scarce data, (2) majority of data with high salinity and therefore low yields, uninteresting for agriculture (3) many cases where a straight line would be statistically preferable instead of an S-curve, (4) absence of replications despite the fact that it is well known that the crop yield at a certain soil salinity value is usually quite variable, (5) Anova tables and F-tests are absent, (6) the yields are expressed as a fraction of the maximum yield so that the absolute values of the yield are unknown while it is a fact that with generally low yield levels the salt tolerance of crops is higher than with generally high yield levels, (7) the salt tolerance levels are very low, below a soil salinity (expressed in electrical conductivity of the saturation extract of a soil sample, ECe) of 2 dS/m. Hence the crops discussed are extremely en exceptionally sensitive, (8) there are cases in which the S-curve is *not* useful, (9) the determination of the parameters is not explained, (10) in many cases the fit of the S-curve to the data is so precise that the impression arises that the data have been manipulated.
In this paper examples found in literature are given together with their references and internet specifications. These examples are critically reviewed and all together the cover all the weak points 1 to 8.

Contents

1. Introduction

In the analysis of crop yield data (Y) versus soil salinity (X) to obtain a salt tolerance index of the crop, the Maas-Hoffman (MH) model has been frequently used. The model shows a horizontal line at the lower salinity values at level Y=Yo, indicating that here the yield is constant with increasing salinity. At a certain X value (called the breakpoint, threshold or tolerance index) and beyond, the line shows a negative slope, meaning that the yield declines at increasing X values and the salinity is apparently damaging (see for example *figure 5* in *section 4*).

There have been attempts to use mirrored S-curves instead of the breakpoint (threshold, tolerance index) model, even though the mirrored S-curves do not reveal a definite breakpoint, so here it is difficult to define a precise tolerance index. Yet, the S-curves have been brought in connection to salt tolerance. Steppuhn et. al. [**Ref.1**] and van Straten et. al. [**Ref. 2**] have used different ways to select, with the help of the S-curve, a well defined tolerance level, that is to say a value of the soil salinity below which the crop withstands the salinity, while at higher values the crop yield is negatively affected.

In literature, the efforts to fit S-curves to available data sets have given results that are far from perfect, as summarized in the abstract of this paper.

In addition to the two first references mentioned, similar efforts were made by van Genuchten and Gupta [**Ref. 4**] and van Genuchten and Hoffman [**Ref.5**]

The cases reviewed and criticized in this paper stem from the above 4 references.

2. Steppuhn data

Steppuhn et al. apply four different S-curve methods, but they use only one dataset obtained from controlled experiments with the Biggar spring wheat crop, so that it is not evident that the methods have a general applicability, valid under many different circumstances and for different crops.

Their methods are shown in *figure 1* below. It is a copy from their publication. Four mathematical methods have been used: a) the Weibull response function, b) the bi-exponential resonse function, c) the (modified) Gompertz response function, and d) the (modified) discount response function. The differences in the outcomes of the 4 methods are negligible.

It is a disadvantage of the Steppuhn publication [**Ref. 1**] that only one data set has been used, so that there is no information on the general applicability of the 4 methods. In addition, the presented data lack replications. As the yield at a certain salinity may show considerable variation (see for example *figure 5* in *section 4*), replications would be needed to assess that variation. Without replications, the dataset is incomplete and the outcomes may have arisen by chance, because replications may have led to different outcomes. Form that point of view, the validity of the methods have not been proven.

Further, there are only two data points with Y=1, suggesting that the critical X value is about 2 dS/m. This would make wheat a crop very sensitive to soil salinity, while many data in literature show wheat to be quite salt tolerant [**Ref. 6**].

In continuation: in the Steppuhn publication it is not explained how the parameters of the S-curve equations have been determined.

Finally, the relative large amount of data at the tail end, though they help to construct the S-curve, are uninteresting for agriculture and unimportant for the determination of the critical value of the salinity (the salt tolerance index) beyond which the yield is negatively affected by the soil salinity, because here the yields are very low.

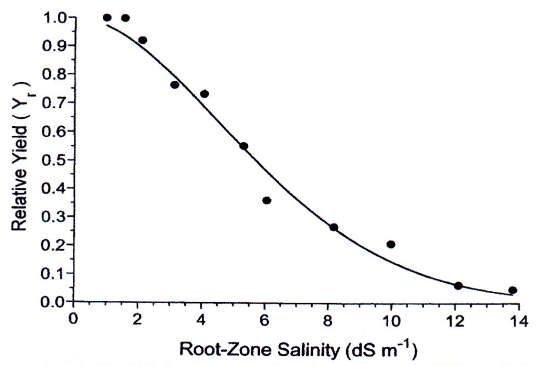


Fig. 2. Modified Weibull response function, $Y_r = \exp[a(C^b)]$, applied to the Biggar spring wheat data.

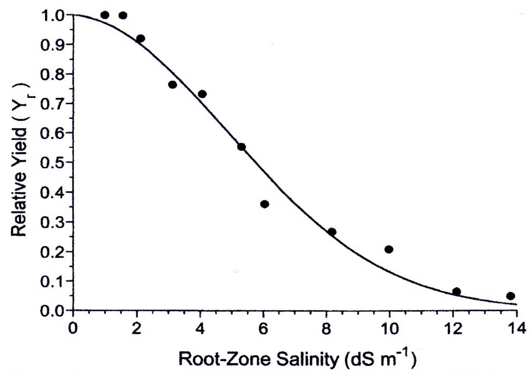


Fig. 3. Bi-exponential response function, $Y_r = \exp[aC - b(C^2)]$, applied to the Biggar spring wheat data.

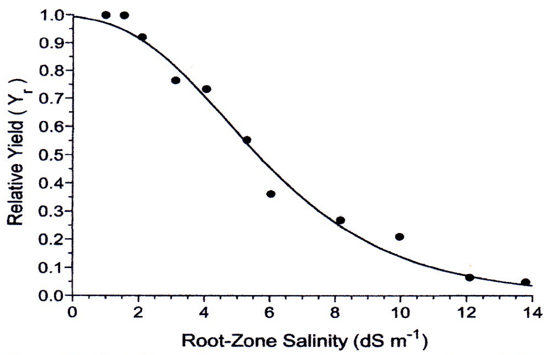


Fig. 4. Modified Gompertz response function, $Y_r = 1 - \exp[a \exp(bC)]$, applied to the Biggar spring wheat data.

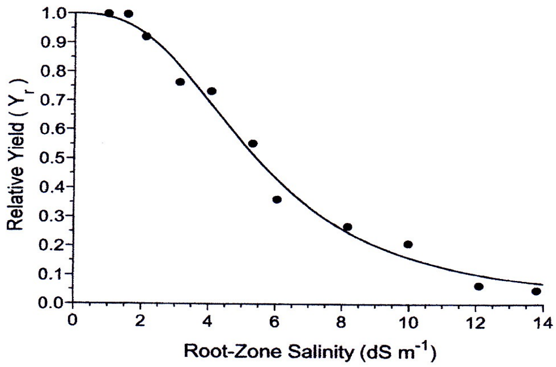


Fig. 6. Modified discount response function, $Y_r = 1/[1 + (C/C_{50})^{\exp(sC_{50})}]$, applied to the Biggar spring wheat data.

***Figure 1**. S-curve methods used bij Steppuhn et al. [**Ref. 1**] on Biggar spring wheat data. The differences in the outcomes of the 4 methods are negligible.*

NOTE. There are many more probability distributions that can be use for S-curves, for example Gumbel (Fig. 2, made with SegRegA [**Ref. 7**].

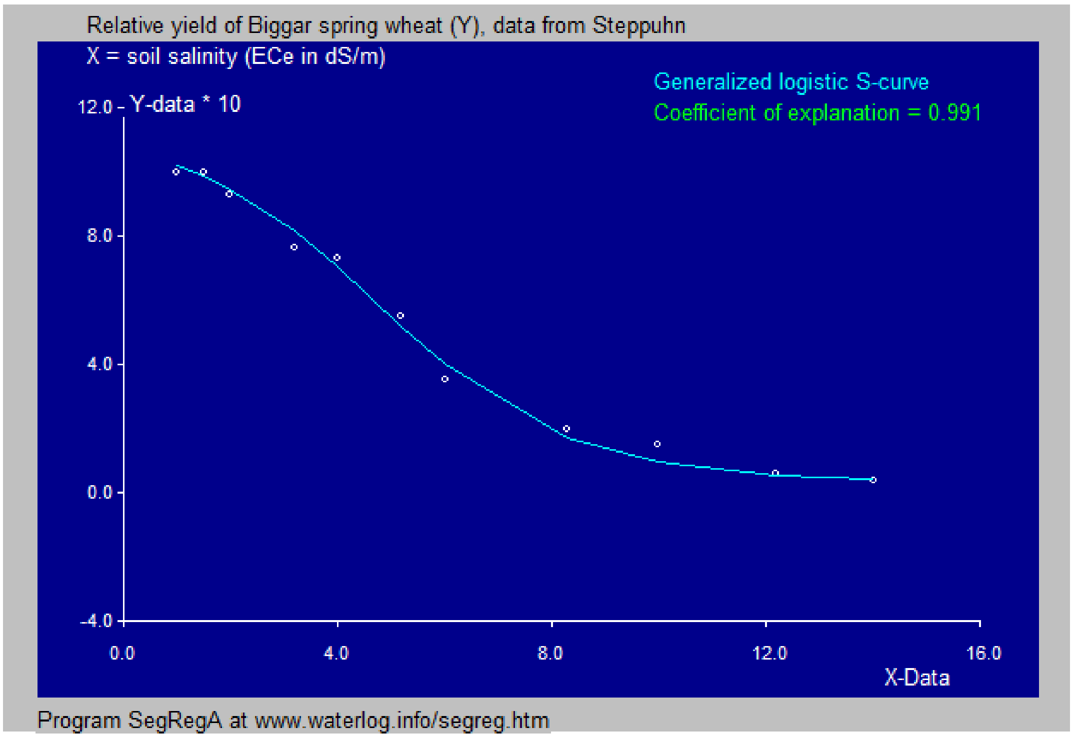


Figure 2. *Biggar spring wheat data of Steppuhn (**figure 1**) used in SegRegA [**Ref.** 7] to obtain another similar S-curve based on a probability distribution, in this case the logistic distribution. The coefficient of explanation (R^2) is given.*

3. Van Genuchten - Gupta data

Van Genuchten and Gupta use the equation $Y = Ym / [1 + (X/X_{50})^P]$, where Ym is the theoretically maximum value of Y (the yield), X_{50} is the X (soil salinity) value where Y is 50% of Ym, and P is a parameter. This will be called the vGG model.

In their article there are two crops whose yield is related to soil salinity: barley and meadow.

The results of the analysis for barley is shown in *figure 3.* This figure reveals weaknesses as discussed before: (1) scarce data, (2) majority of data with high salinity and therefore low yields, not important for agriculture, (3) absence of replications, which would be replications would be needed to assess the variation. Without replications, the data set is incomplete and the outcomes may have arisen by chance, because replications may have led to different outcomes. From that point of view, the validity of the methods have not been proven. (4) The yields are expressed as a fraction of the maximum yield so that the absolute values of the yield are unknown while it is a fact that with generally low yield levels the salt tolerance of crops is higher than with generally high yield levels, (5) the salt tolerance level (threshold, breakpoint, salt tolerance index) is very low, below a soil salinity (expressed in electrical conductivity of the saturation extract of a soil sample, ECe) of even 1 dS/m. Hence the barley discussed is extremely, even exceptionally, sensitive, while barley is well know for its high salt tolerance [**Ref. 6**]

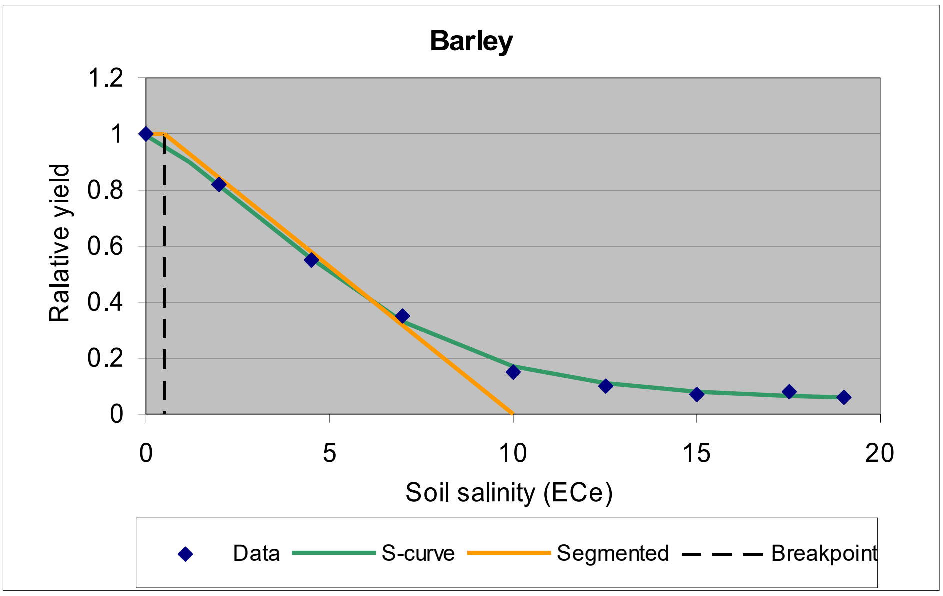


Figure 3. *Relative yield of barley versus soil salinity (ECe in dS/m). The vGG S-curve is shown in green color. The orange line signifies a segmented linear regression with breakpoint (threshold, tolerance index) at X=0.05 dS/m (dotted line), as determined by vGG with a view to the Maas-Hoffman model (see the introduction, section 1).*

The correspondence between the S-curve and the observations is amazingly high. This may give rise to the question: have the data been manipulated?

Figure 4 depicts the yield of meadow versus soil salinity. The S-curve according to the vGG model (in red color) is peculiar, as explained in the subscript. In fact, it is not justified, and it would be better to stick to the straight line (green color). The meadow crop has no breakpoint according to the Maas-Hoffman (MH) model and the crop production declines immediately at increasing soil salinity, even at a low level. Hence, the tolerance to salinity is zero. This is probably due to the absence of data between soil salinity ECe = 0 dS/m and ECe = 12 dS/m. Here we have an example of experimentation with mainly very high salinity values. (see weakness 2 in the abstract). Also all the other weaknesses are apparent.

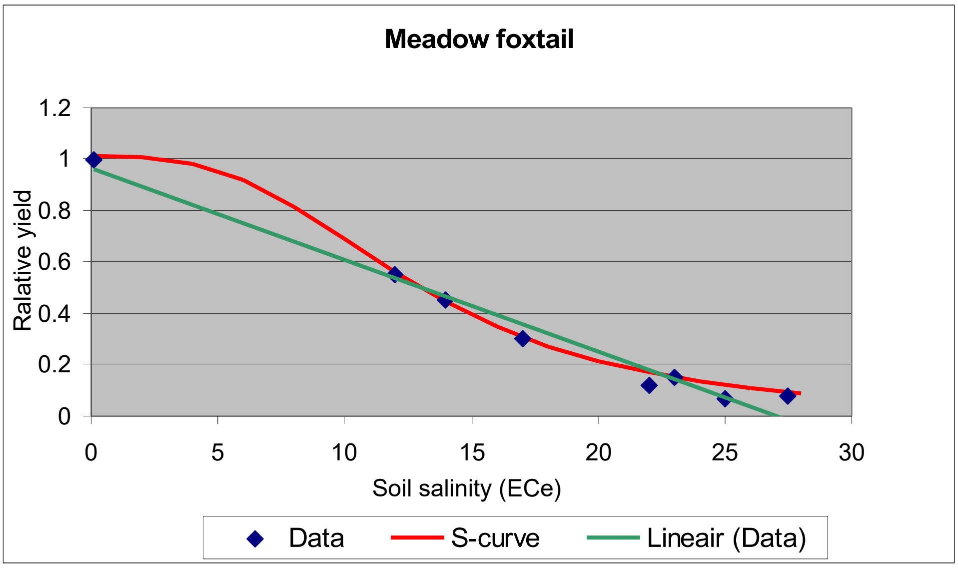


Figure 4. *Relative yield of meadow versus soil salinity (ECe in dS/m). The red line shows the fitted S-curve, while the green line represent a simple linear regression. In the range between X = 1 dS/m to X = 12 dS/m there are no data. The large deviation of the S-curve from the straight line, therefore, has no experimental justification. Beyond X = 12 dS/m, the differences between the S-curve and the straight line are negligibly small. The use of the S-curve in this case is not justified.*

4. Van Straten data

Van Straten et al. describe experiments in test fields. The experimental fields were irrigated during the growing season with large amounts of water of different salinity which was held constant during the growing season. Hence, in each test block the soil salinity was more or less constant during the season and it corresponded to the salinity of the irrigation water in that block. There were 4 replications each year. Many different crops were treated this way.

In the van Straten article [**Ref. 2**], only the data of the potato variety Achilles (t/ha) were dealt with. The soil salinity was expressed in electrical conductivity of the soil moisture (ECe, dS/m). The data were obtained repeatedly at varying ECe levels and during 5 years (2012-2016). In the article, the majority of the examples are given for Achilles 2014.

The Achilles data appear to be useless [**Ref. 3**], as is explained in continuation.

Using the Maas-Hoffman (MH) model (see section 1, introduction), it is found that the breakpoint (threshold) is lower as the Yo value is higher (see table 1). In fact, at low yield levels, the threshold values are not representative for the variety Achilles. At yields above 50 t/ha (which is the normal range of the yield in the Netherlands), the threshold is zero. At yields around 30 t/ha farmers in the Netherlands would not be able to make a living, so data with low Ym yield levels should not be used at all.

In the van Straten article **[Ref. 2**] this phenomenon is not mentioned and this important feature is disregarded [**Ref. 3**].

Table 1. *Annual values of the maximum level Ym of the yield (Y) in the MH model (see section 1, introduction) and the breakpoint (threshold) value of the soil salinity (X) for the potato variety Achilles. In year 2013 and 2014 no threshold could be detected (it is zero), the crop yield in these years immediately reduced at increasing soil salinity from the start, so that the salt tolerance is zero and the crop is extremely sensitive.*

Year	Yield Ym	Threshold (breakpoint)
2012	33 (low !)	5
2013	51	0 (zero !)
2014	27 (low !)	5
2015	61	0 (zero !)
2016	35 (low !)	4

If the yields would have been expressed in relative yield, the relation between overall yield level and threshold value would not have been found, but the above table proves that the absolute values need to be considered. In section 3 (Van Genuchten - Gupta data) only relative crop productions have been used, which could be of a great disadvantage.

The figures below (copied from the article) illustrate this principle for Achilles 2014: left the MH model, right the vGG (S-curve) model. The 90% yield in the vGG model corresponds to ECe=6 dS/m. Strange enough, its confidence interval is not shown. Note that the Yo yield in these figures is only 27 t/ha! This yield is not representative for potato in agriculture in The Netherlands. In fact this case should not have been used.

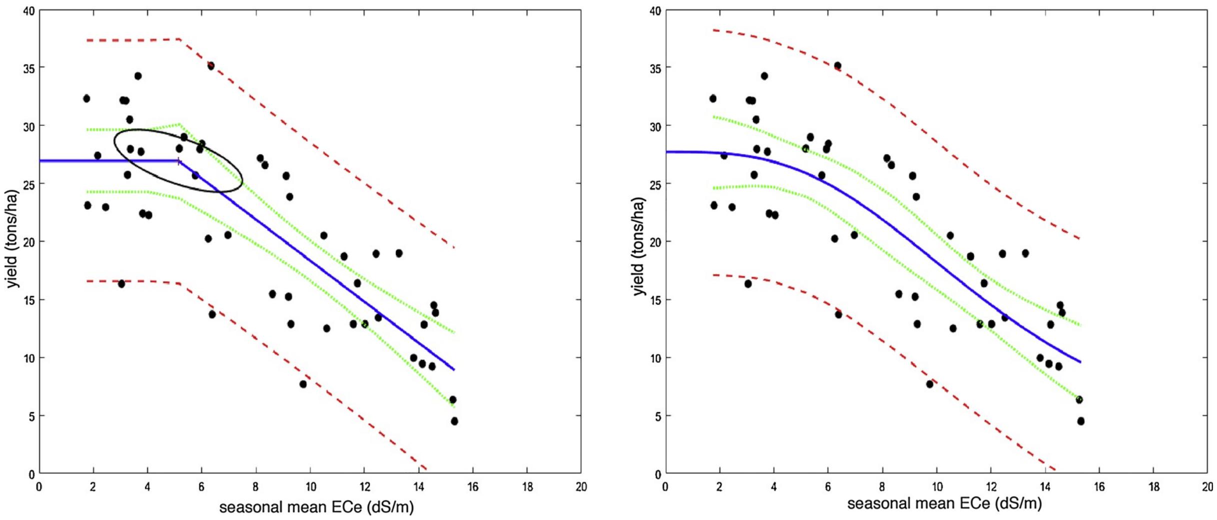


Figure 5. *Van Straten data of the yield of potato Achilles 2014 versus soil salinity (ECe in dS/m). The general yield level is low, reason why the MH model is applicable as there is a breakpoint (threshold, tolerance index), see left figure. Owing to the replications, the variance is large. The right hand figure shows the vGG S-curve model (see section 3).*

The authors claim that the ECe at the 90% yield level in the S-curve model constitutes an “improved methodology” to define the salt tolerance index of crops. However, in the left hand figure they indicate a statistical confidence interval of the threshold as an ellipse, while in the right hand figure there is no such interval for the 90% point, even though in their article this point is strongly promoted as an index.

Given the large variations, it is debatable that the authors do not use analysis of variance (ANOVA) to test whether the models are significantly different compared to a simple linear regression. For the vGG model (Achilles 2014) the ANOVA table looks as follows (table 2):

Table 2*. Analysis of variance (Anova) of the vGG model to test statistically the improvement of the vGG model with respect to a simple linear regression in the Achilles 2014 case. The F-test provides a significance level of only 48%, meaning that there is 52% chance that the vGG model has arisen by coincidence and that the model does significantly improve the straightforward linear regression.*

Sum of squares of deviations		Degrees of freedom	Variance	Fisher's F-test	Probability (significance) (%)
Total	44.200	47	0.940		
Explained by lin. regress.	25.300	1	25.300	F(1,46)= 61.577	99.9 %
Remaining Unexplained	18.900	46	0.411		
Extra explained by vGG model	0.533	2	0.271	F(2,44)= 0.666	48.1 %
Remaining Unexplained	18.375	44	0.408		

The following figure (6) was made with the PartReg model [**Ref. 6**] to explain why in the case of Achilles 2013, with a relatively high yield level, the MH model did not help to find a critical salinity level (threshold, tolerance index), see *table 1*.

Van Straten et al. [**Ref. 2**] determine the parameters of the MH model so that the sum of the differences between the observed and modeled Y values over the entire domain is minimized. The PartReg model, however does a partial regression to detect only the horizontal stretches in a data set *(figure 6*). PartReg finds that the Achilles 2013 data possess a long horizontal tail, reason why in the MH model the slope of the line beyond the breakpoint is flat and drawn to the left, so that it intersects the Y axis at X=0. The horizontal part of the MH model is thus not found. On the other hand, the PartReg method does find this part, and in *figure 6* it can be seen that there is a breakpoint at about X = 6 dS/m. This can be used as a salt tolerance index. In this case, an S-curve is certainly not required as it does not help to determine a salt tolerance index. The large number of high salinity values at the tail end is agriculturally not interesting.

Finally, the methods employed by Steppuhn et al. (*Section 2*) were disregarded, including their salinity tolerance index based on the vGG model.

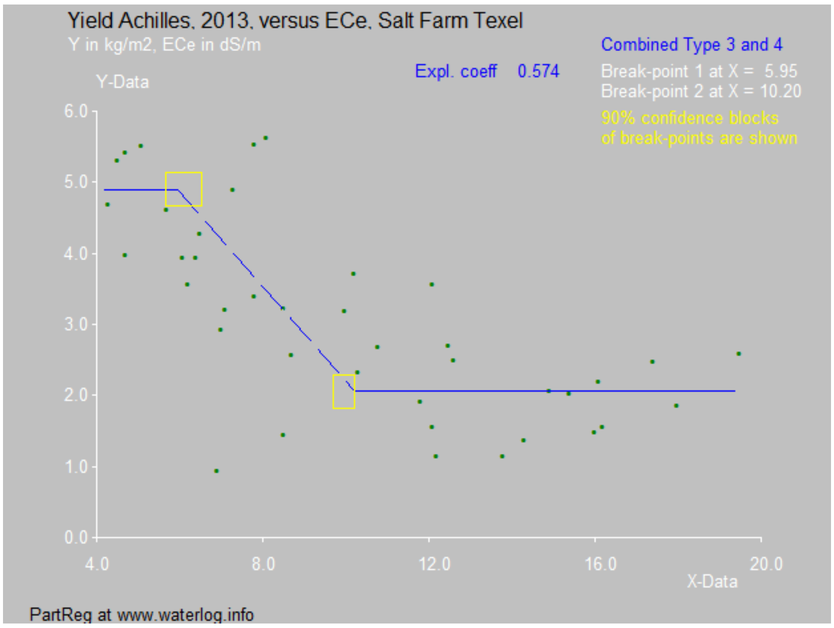


Figure 6. *Result of the PartReg program applied to the yield and salinity data of Achilles 2013. PartReg detected two horizontal stretches giving rise to a Z-type relation. The data between the two break points have been left untreated.*

The PartReg method produces a Z-type yield-salinity relation (*figure 6*). The long tail in figure 6 is inviting to use the vGG model to apply an S-type model, which has similarity with the Z-type. The S-type, however does not produce breakpoint and is less useful to determine critical X values and formulate a salt tolerance index. In PartReg, contrary to the MH model and the S-curve, the large tail end does not influence the salt tolerance index.

More examples of PartReg applications can be found in [**Ref. 9**].

4. Van Genuchten-Hoffman data

Van Genuchten and Hoffman, in their article, show crop yield relations with soil salinity for the crops: Brome grass, Tomato, Perennial Rye, and Tall fescue. The corresponding details are depicted hereunder in figures 7 to 10. Comments are written in the subscripts of the figures.

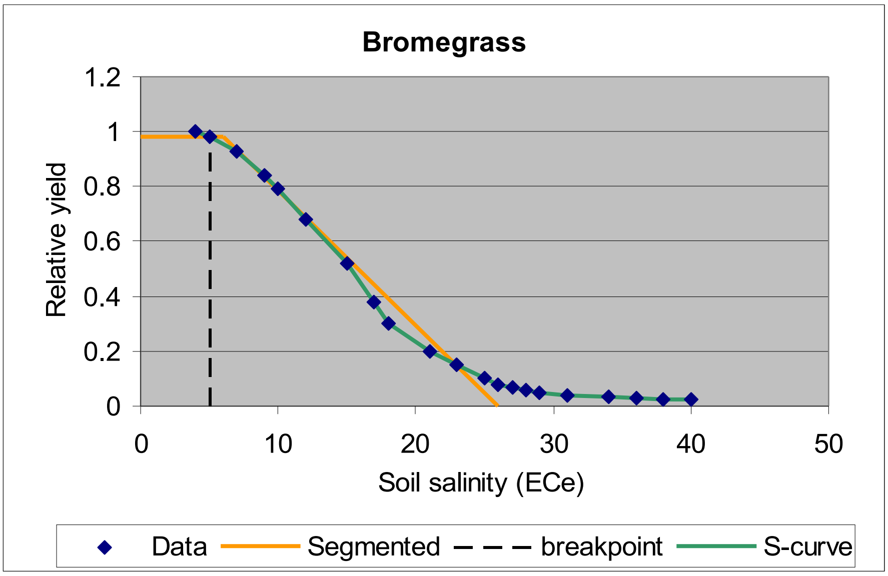


Figure 7. *Relative yield (Y) of bromegrass in versus soil salinity (X). The disadvantage of plotting the relative yield has been discussed before (under table in section 3). The fit of the S-curve (in green color) to the data is amazing and brings up the idea the data have been manipulated, as in* ***figure 3***. *The authors claim that there is a breakpoint in the MH model (orange color) at ECe of about 5 dS/m. This statement is based on only two observations of a relative yield near the value Y= and therefore statistically doubtful.*

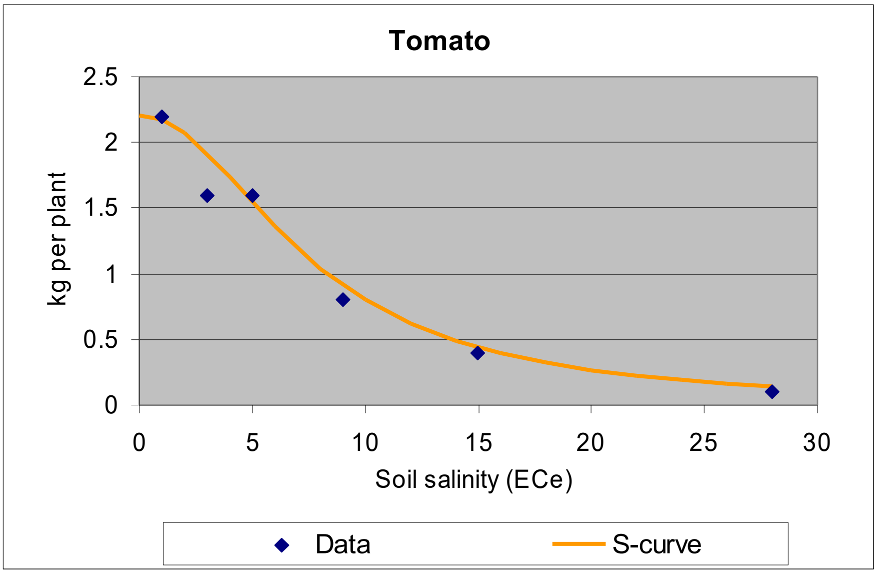


Figure 8. *Yield of tomato in kg/plant (Y) in relation to soil salinity (X). The number of data at low X values is small so that a tolerance index cannot be found. The shape of the S-curve (in orange color) depends greatly on the higher X values of the data. In this case, the sensitivity of tomato to soil salinity is high as the descent of the curve starts immediately at X values grater than ECe = 1 dS/m. This may not be realistic. A further disadvantage is that the are no replications as discussed in section 2 (Steppuhn data).*

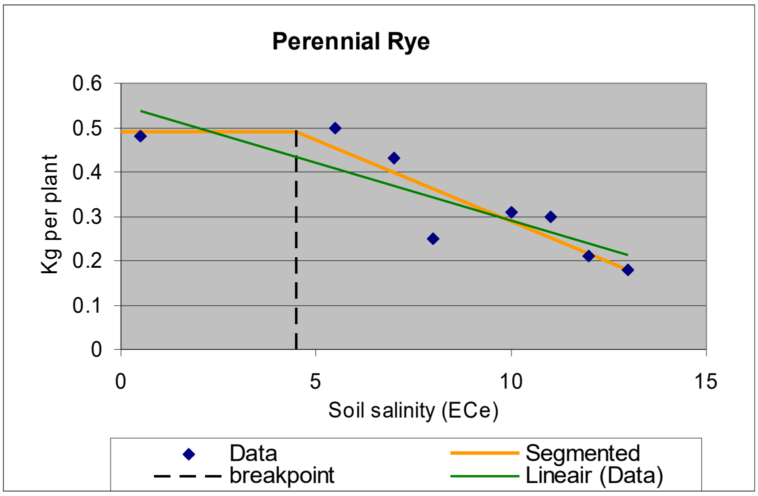


Figure 9. *Relation between yield of perennial rye in kg/plant (Y) and soil salinity (ECe in dS/m, X). An S-curve could not be found, which illustrates that S-curves are not always applicable (item 8 of the weaknesses mentioned in the abstract). The authors have introduced the MH model (drawn in orange color) and claim that the threshold value is close to X = 5. This conclusion is based on only two observations. Statistically an analysis of variance would be needed to decide if the MH model is significantly better than the straight line found by linear regression (in green color). The lack of replications will make it difficult to decide on this matter.*

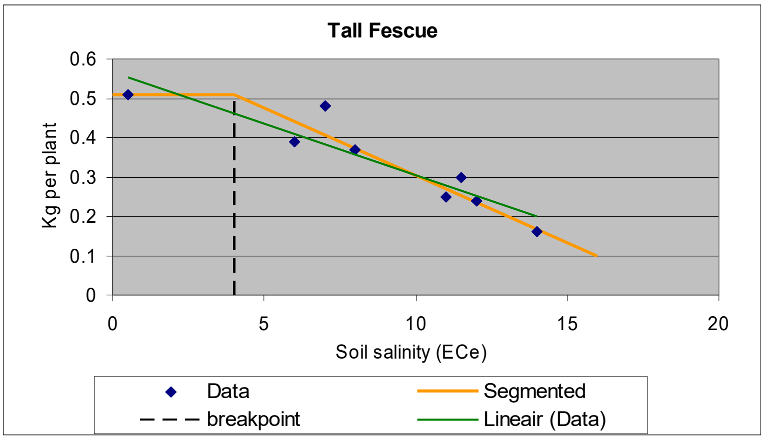


Figure 10. *Relation between yield of tall fescue in kg/plant (Y) and soil salinity (ECe in dS/m, X). The comments made under **figure 9** are equally relevant here. The determination of the breakpoint (at X between 4 and 5) as a tolerance index is questionable as the straight line (in green color) seems to be statistically the most appropriate solution here.*

5. Mousavi data

The following image stems from an article written by Hesam Mousavi et al. 2022 [**Ref. 8**].

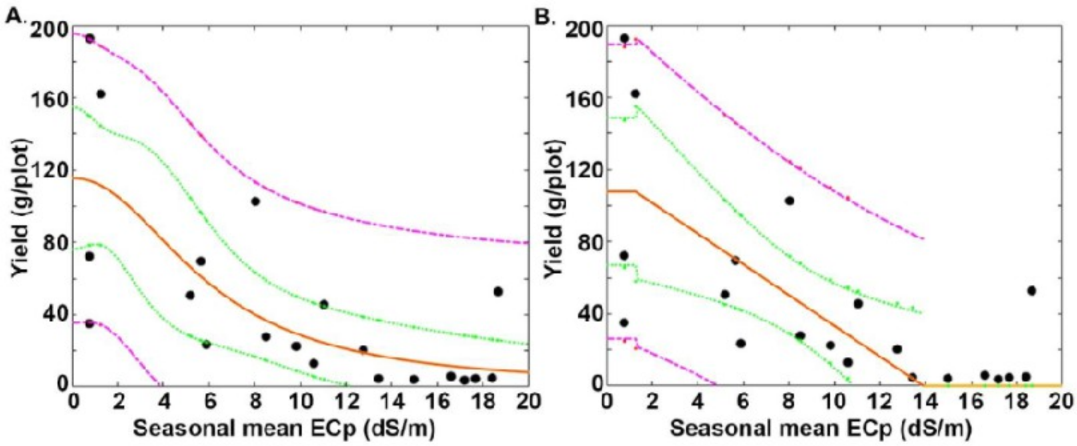
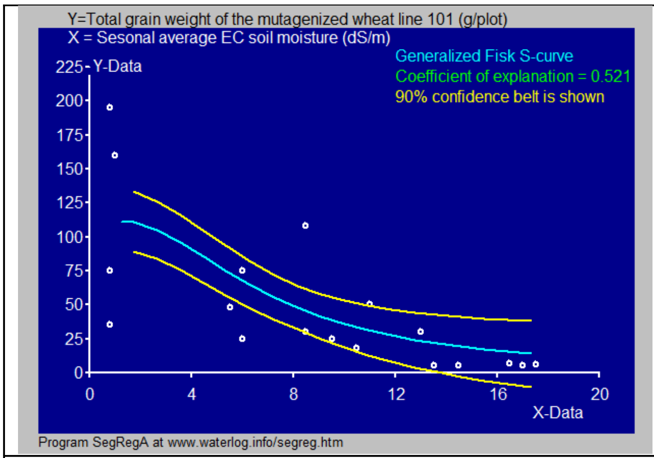
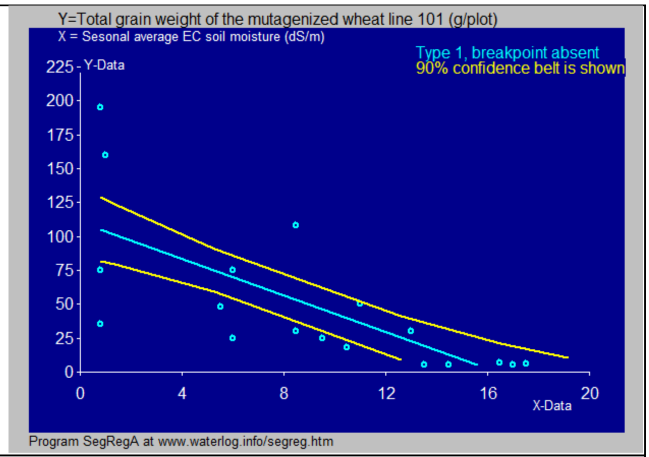


Figure 5. Total grain weight (TGW) of the mutagenized line 101 as a function of salinity obtained using the van Genuchten–Hoffman (A) and Maas–Hoffman (B) models. Dots represent measurement points and the orange lines show the least square best fits. The dotted green lines represent upper and lower simultaneous prediction error bounds, i.e., the (approximately 95% CI) uncertainty bounds for a new experiment with the same number of data points. The magenta dash-dot lines represent upper and lower non-simultaneous prediction error bounds, i.e., the approximately 95% CI range when a single measurement is done.

With a plot size of 10 m^2 the yield levels are very low.

Using SegRegA **[Ref. 7]** the van Genuchten-Hoffman model (left side) and the Maas-Hoffman model (right side) look as follows:

<table>
<tr>
<td>
</td>
<td>
</td>
</tr>
<tr>
<td>The van Genuchten-HoffMan model is based on the Fisk probability distribution.</td>
<td>For the Maas-Hoffman model the break point does not exist because it is less than 1 dS/m</td>
</tr>
</table>

The data with EC values beyond 13 dS/m are useless as they are practically zero.

The ANOVA (ANalysis Of VAriance) table made in SegRegA for the van Genuchten-HoffMan model (a mirrored S-curve) is as shown below.

From the table it follows that the van Genuchten-Hoffman model is highly insignificant and should not be used.

Variance Analysis, ANOVA table, Regression Type: mirrored S-curve

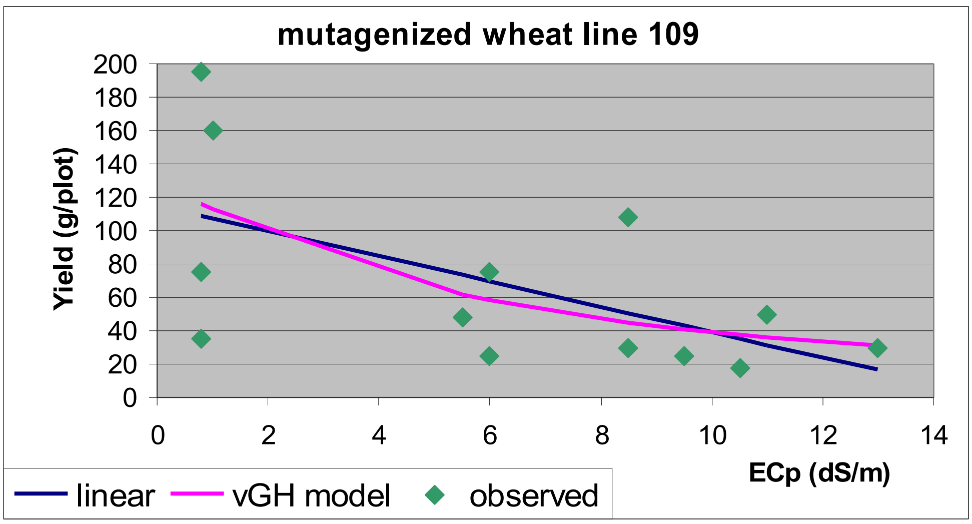
Sum[(Y-Av.Y)sq.] = 50900.000 (total sum of squares of deviations)

Total nr. of data = 18
Degrees of freedom = 17

Sum of squares of deviations	Degrees of freedom	Variance	F-Test	Probability/ Significance	
explained by lin. regr. 25800.000	1	F(1,16) = 25800.000	16.446	99.9 %	
remaining unexplained 25100.000	16	1568.750			
extra expl. by S-curve regr. 736.869	2	F(2,15) = 368.434	0.227	20.0 %	**#)**
remaining unexplained 24363.131	15	1624.209			
total expl. by S-curve regr. 26536.869	3	F(3,14) = 8845.623	5.083	98.6 %	

#) there is 80.0 % chance that the mirrored S-curve is invalid

The difference between the mirrored S-curve (vGH model) and the straight line is negligibly small (see next figure), certainly when compared with the scatter of the data. The yield variation for ECp<2 is enormous.



The vGH model and linear regression of the Mousavi data after cut-off of the *useless near zero yield values when ECp>13 dS/m*

7. Conclusion

The weaknesses of the S-curve techniques mentioned in the abstract have been amply illustrated in the previous sections.

As there were no explanations about the determination of the parameters of the S-curves, for more insight one could read the paper "Free software for the determination of positive and inverted S-curves for the response function of influential treatments or conditions with examples of crop yield versus soil salinity and depth of the water table" [**Ref. 9**].

7. References

[Ref. 1] H. Steppuhn et al. , 2005, CROP ECOLOGY, MANAGEMENT & QUALITY; *Root-Zone Salinity: I. Selecting a Product–Yield Index and Response Function for Crop Tolerance* . In: Crop Sci. 45:209–220 (2005), Crop Science Society of America. On line:
https://www.researchgate.net/publication/43257218_Root-zone_salinity_I_Selecting_a_product-yield_index_and_response_function_for_crop_tolerance
or:
https://pubag.nal.usda.gov/download/3381/PDF

[Ref. 2] G. van Straten et al., 2019. *An improved methodology to evaluate crop salt tolerance from field trials.* In: ***Agricultural Water Management***, ***Volume*** 1, March 2019, Pages 375-387. On line:
https://www.sciencedirect.com/science/article/pii/S0378377418310370

[Ref. 3] Methods to evaluate crop salt tolerance from field trials, a critical review of the Salt Farm Texel article entitled: "An improved methodology to evaluate crop salt tolerance from field trials", which gives no improvement at all, to the contrary.
On line:
https://www.researchgate.net/publication/349038823_Methods_to_evaluate_crop_salt_tolerance_from_field_trials_a_critical_review_of_the_Salt_Farm_Texel_article_entitled_An_improved_methodology_to_evaluate_crop_salt_tolerance_from_field_trials_which_give
or:
https://www.waterlog.info/pdf/SaltMethods.pdf

[Ref. 4] M. Th. Van Genuchten and S.K. Gupta, 1993. *Reassessment of the Crop Tolerance Response Function*. Journal of the Indian Society of Soil Science, Vol. 41, No. 4, pp 730-737 (1993). On line: https://www.ars.usda.gov/ARSUserFiles/20360500/pdf_pubs/P1295.pdf?origin=publication_detail

[**Ref. 5**] Martinus Th. Van Genuchten and G.J Hoffman, 1984. *Analysis of crop salt tolerance data*. On line:
https://www.researchgate.net/publication/238185339_Analysis_of_crop_salt_tolerance_data

[Ref. 6] PartReg, free software for detection of horizontal segments in a (Y,X) data set.

On line: https://www.waterlog.info/partreg.htm

[Ref. 7] SegRegA, free software for segmented and curved regressions, including S-curves. On line: www.waterlog.info/segreg.htm

[Ref. 8] Hesam Mousavi et al. 2022. Effects of Increasing Salinity by Drip Irrigation on Total GrainWeight Show High Yield Potential of Putative Salt-Tolerant Mutagenized Wheat Lines. In Sustainability 2022, 14, 5061. https://doi.org/10.3390/su14095061

[Ref. 9] R. J. Oosterbaan. (2018) *Crop Tolerance to Soil Salinity, Statistical Analysis of Data Measured in Farm Lands*. International Journal of Agricultural Science, **3**, 57-66.
On line: https://www.waterlog.info/pdf/AgrJournal.pdf
or:
https://www.researchgate.net/publication/332466260_CROP_TOLERANCE_TO_SOIL_SALINITY_STATISTICAL_ANALYSIS_OF_DATA_MEASURED_IN_FARM_LANDS