Escalas diagramáticas

Diseño y validación

Juan Pablo Edwards Molina
Juan Andrés Paredes
Bruno Pugliese

Evaluación visual de enfermedades

Tipo de variables

Continuas

min max

Discretas

min | | | | | | | | | | | | | | | | | max

Categóricas

Conteos

Severidad
Altura
Rendimiento

Binaria
Sana o enferma
Incidencia
Prevalencia

Ordinal cualitativo
Clases de severidad
Ej (Sana < clorotica < necrotica)

n esclerocios / capítulo
n nematodes / g raíz
lesions / cm2

Modelos Lineales LM

Modelos lineales generalizados aditivos (GAM)

Modelos lineales generalizados - GLM

What is **Severity?**

Conceptually and operationally

Table 1 The Horsfall and Barratt interval scale showing the disease severity ranges, midpoints and interval sizes [58]

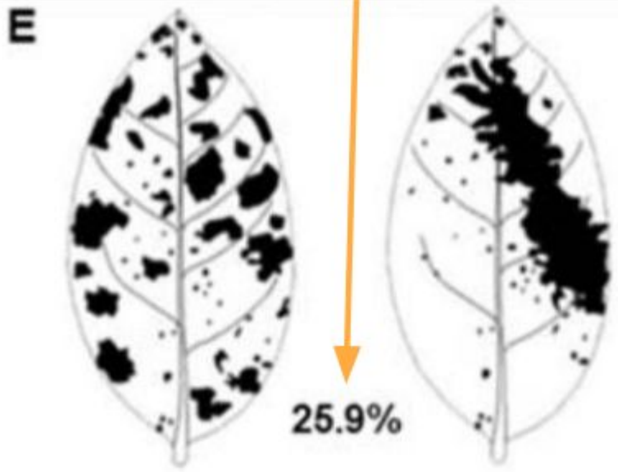
H–B category	Disease severity range (%)	Midpoint	Interval size
1	0	0	0
2	0^+–3	1.5	3
3	3^+–6	4.5	3
4	6^+–12	9.0	6
5	12^+–25	18.5	13
6	25^+–50	37.5	25
7	50^+–75	62.5	25
8	75^+–88	81.5	13
9	88^+–94	91.0	6
10	94^+–97	95.5	3
11	97^+–100	98.5	3
12	100	100	0

Ordinal

Class
(% interval)

Ordinal
(midpoint %)

Ratio (%)



Standard area diagrams - SAD
Escalas diagramáticas estandarizadas

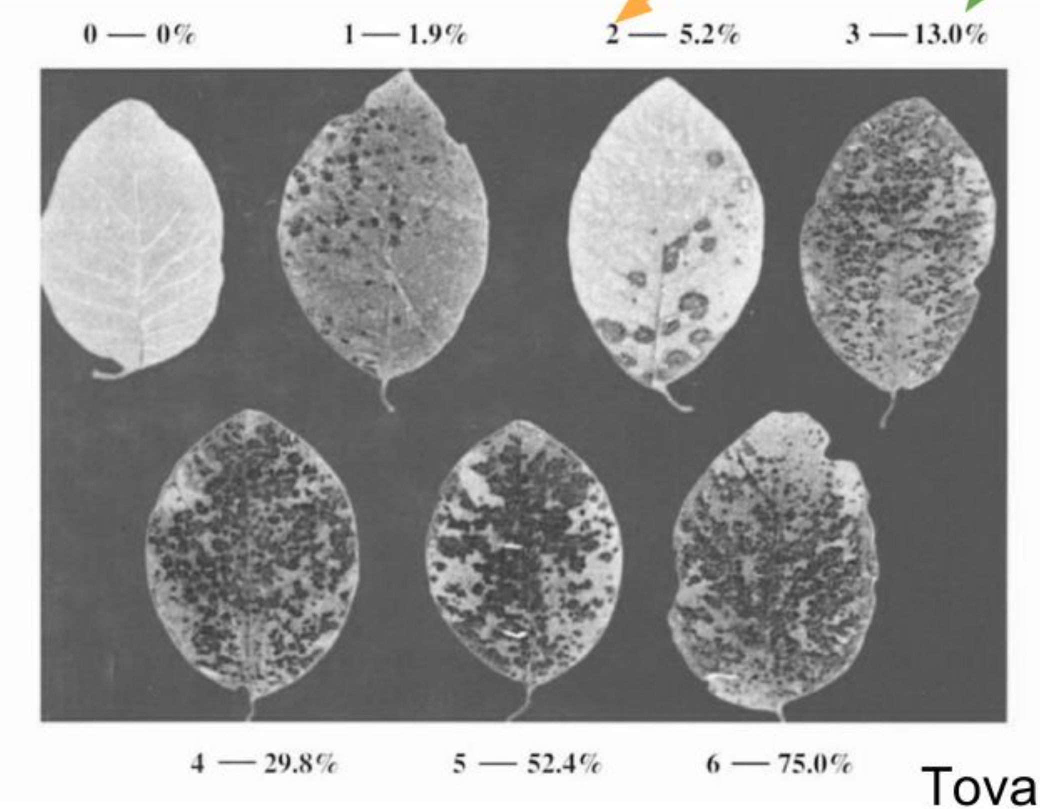
0,6% 2,0% 7,0% 18,0% 42,0% 78,5%

Godoy et al (2006)

Hybrid system

Ordinal (0-6)

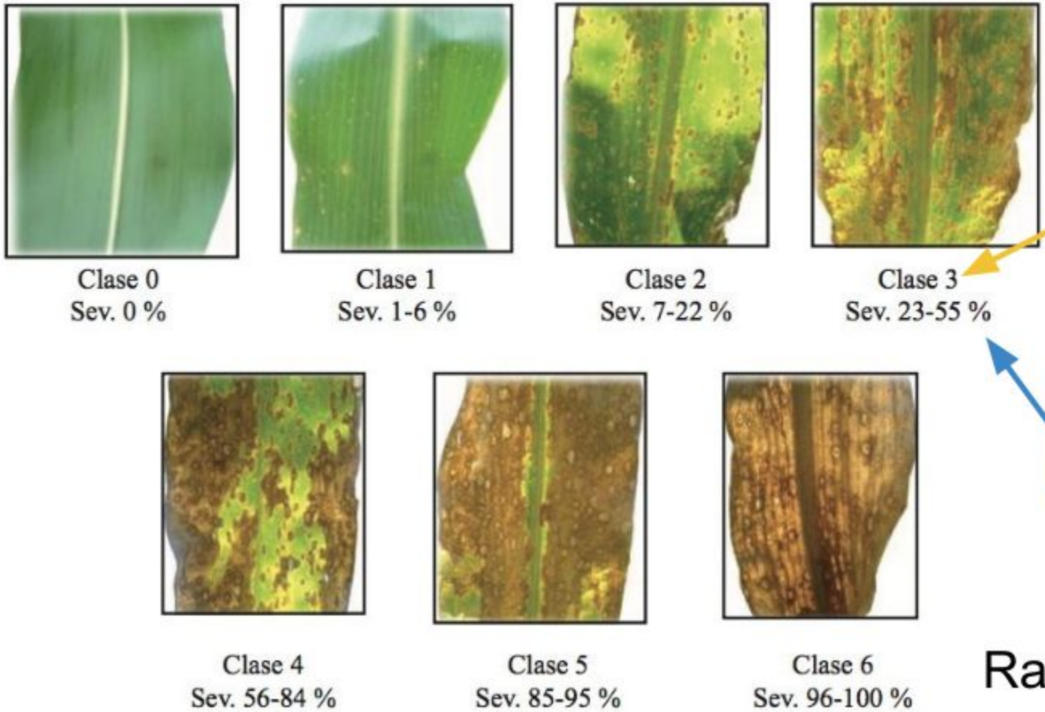
Ratio (%)



Tovar-soto et al (2002)

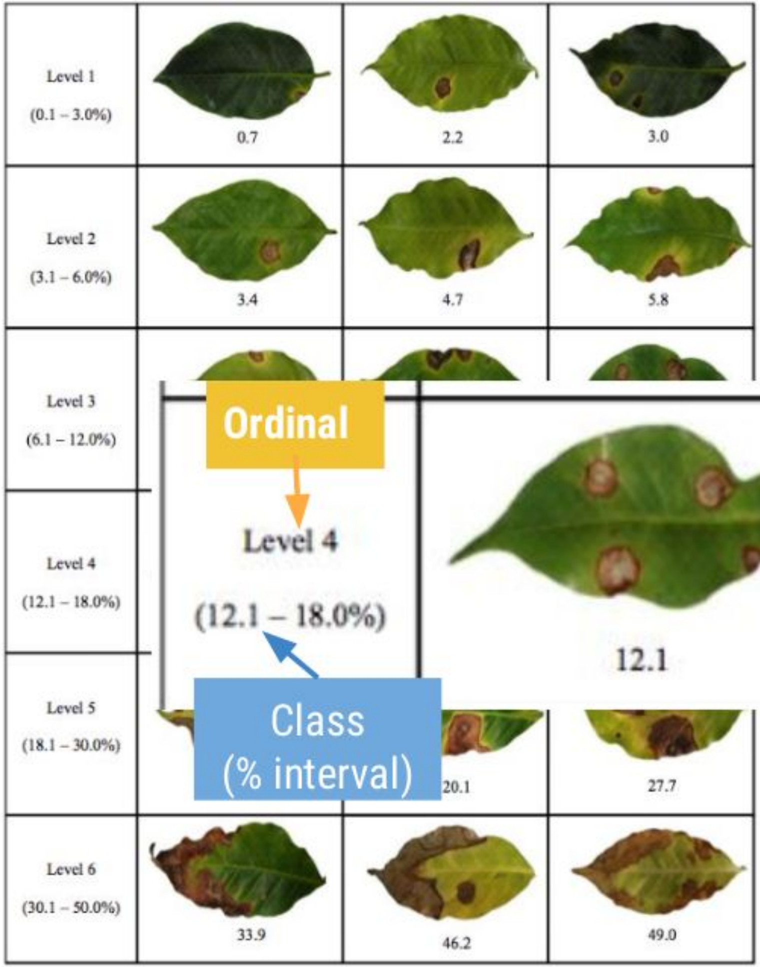
Hybrid system

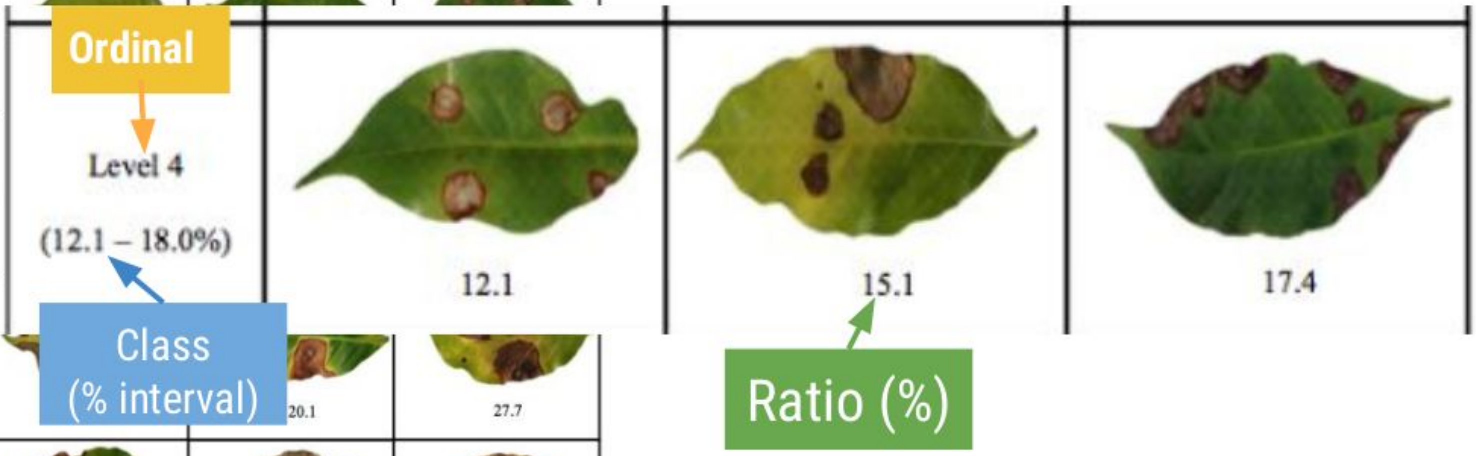
Scores and severity interval (class)



RamosandIslas (2015)

multiple systems:



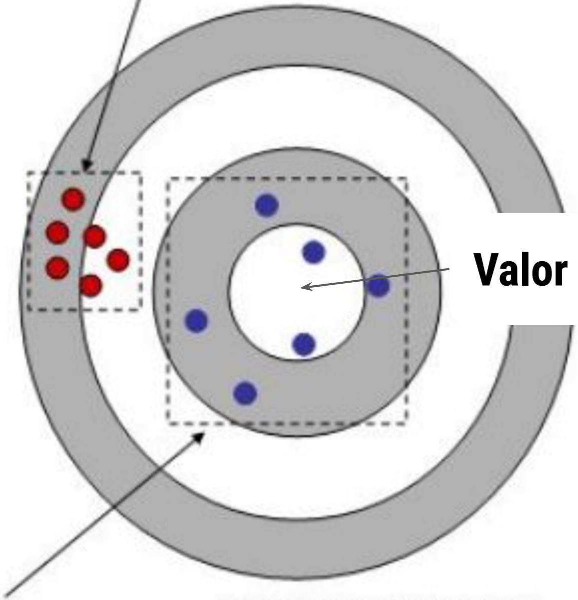


Belan et al. (2014)

Calidad de las estimaciones

Precisión

Variabilidad asociada a las estimaciones de una muestra
Medida de dispersión, +Preciso = bajo SD



Valor real (“standard gold”)

“Accuracy”

Exactitud

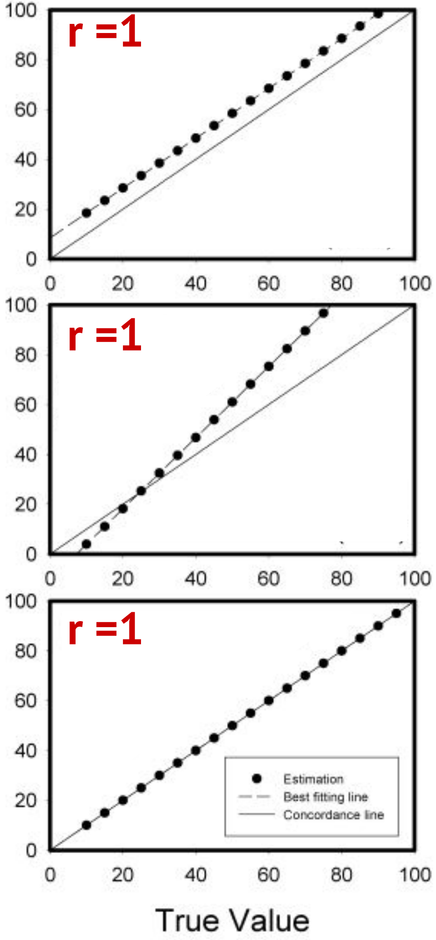
Grado de proximidad de los valores estimados vs reales
“Medida de posición”

Precisión Total

Cambio de posición

Cambio de escala
<y posición>

Exacto



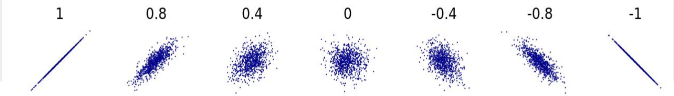
***r* Insensible a cambios de Posición o de escala!**

Sesgo sistemático en todos los valores reales
<siempre arriba en 10 unidades>

Sesgo variable con el aumento de los valores reales

Concordancia absoluta
Todos los puntos sobre la diagonal 1:1

Coeficiente de correlacion de Pearson r <precisión>
Estrechez de los datos sobre su línea de mejor ajuste de los datos



Madden, 2007

Coeficiente de correlación de concordancia de Lin (Lin's CCC)
Varía entre -1 y 1 (concordancia total); 0 es ausencia de concordancia

Coeficiente de Pearson
<Precisión>

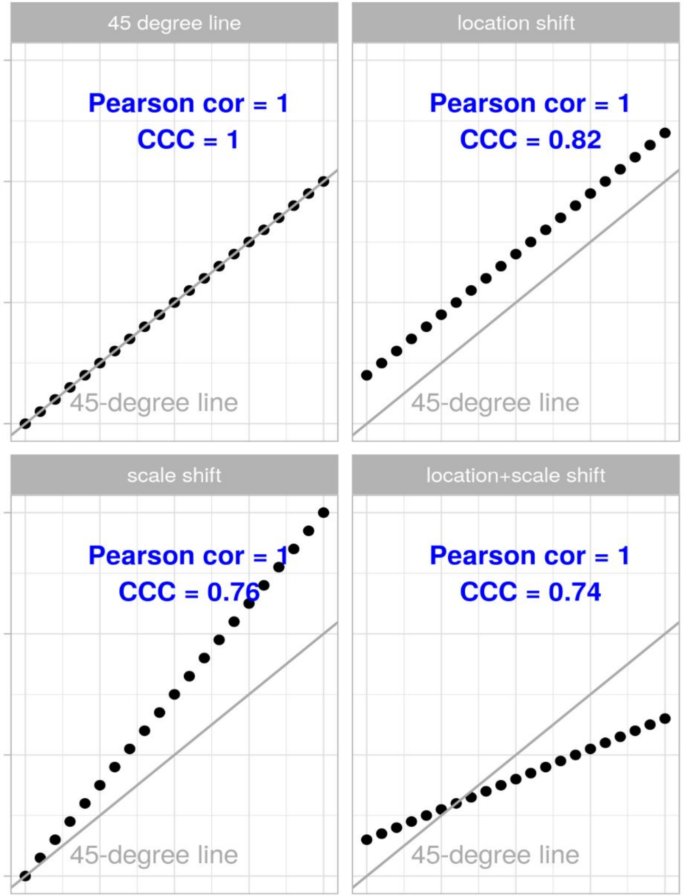
CCC = *r Cb*

Coeficiente de sesgo *<Exactitud>*
Cuán lejos está la reg. de los puntos de la 45°

$$C = \frac{2}{v + \frac{1}{v} + u^2}, v = \frac{\sigma_1}{\sigma_2}, u = \frac{(\mu_1 - \mu_2)^2}{\sigma_1 \sigma_2}$$

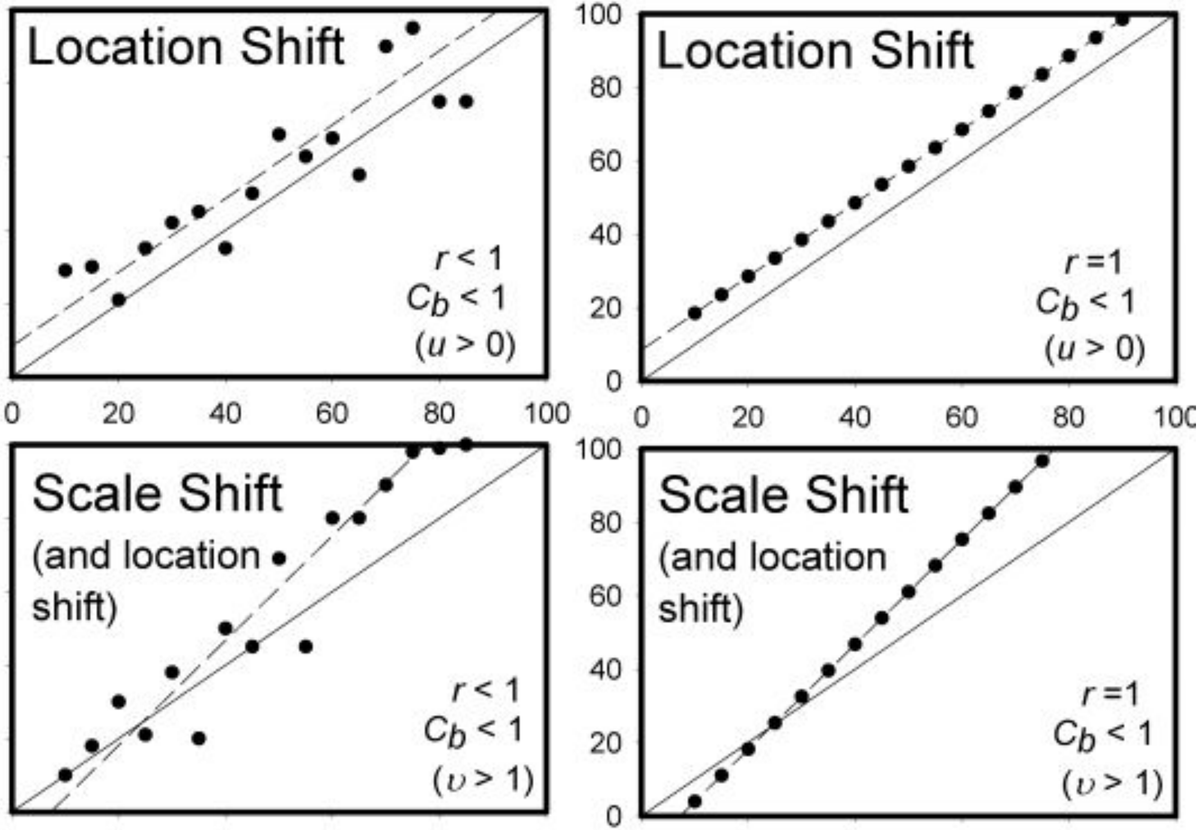
Sesgo de Escala

Sesgo de Posición



https://www.alexejgossmann.com/ccc/

Valores estimados

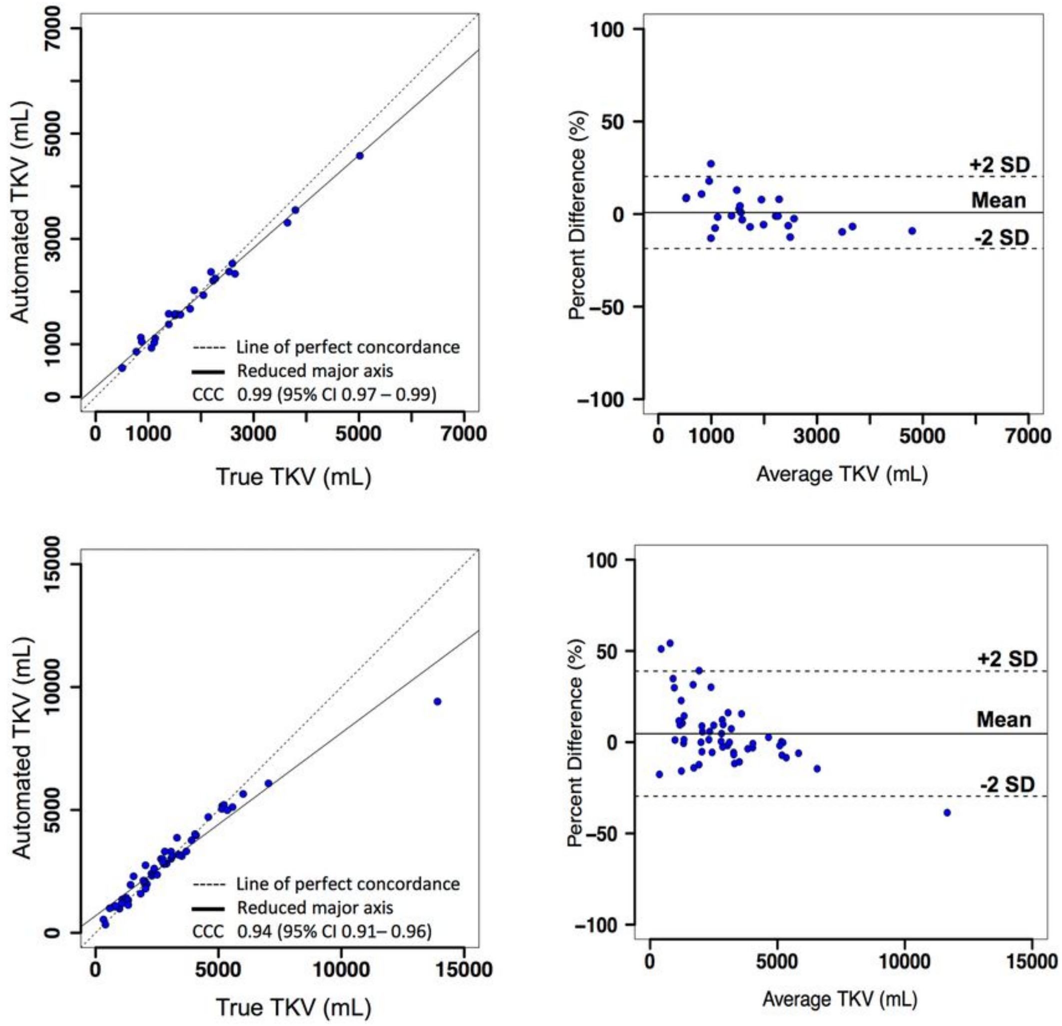


Cambio de posición
u
<difieren las medias>

Cambio de escala
y posición
υ
<difieren las varianzas>

Valores reales

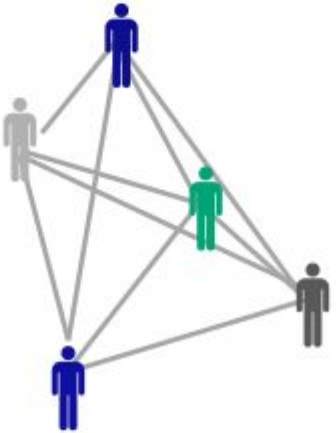
Madden, 2007



DOI:10.1038/s41598-017-01779-0

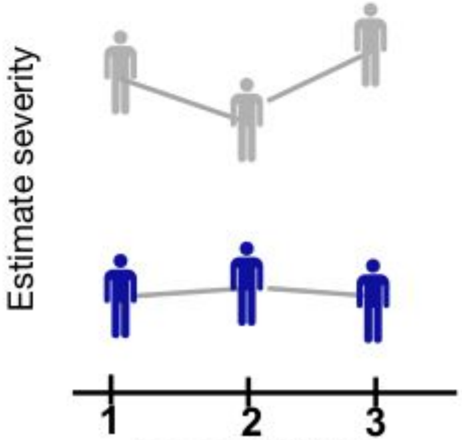
Calidad de las estimaciones - Cont.

Confiabilidad inter-evaluador



Reproducibilidad

Confiabilidad intra-evaluador



Tiempo

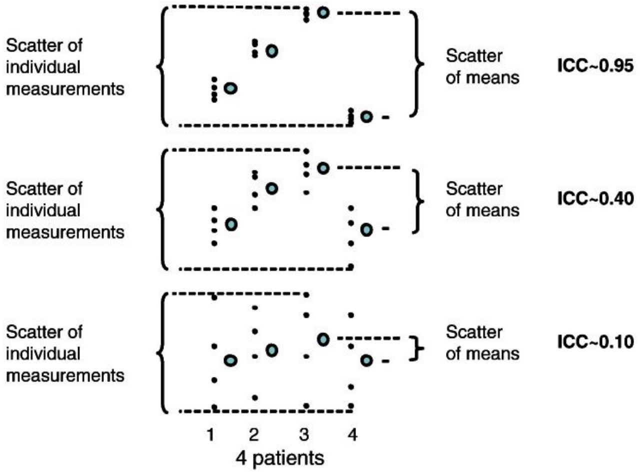
Repetibilidad

Bock et al. (2016)

Coeficiente de correlación intraclase ICC

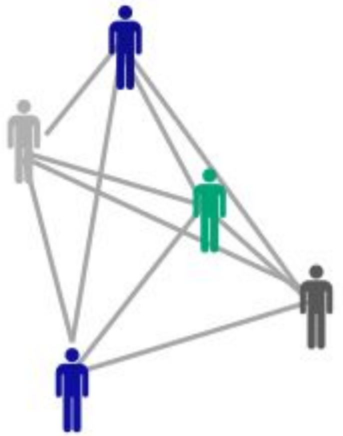
Estadística descriptiva utilizada cuando se realizan mediciones cuantitativas en unidades que están organizadas en grupos

Describe cuán fuertemente las unidades en el mismo grupo se parecen entre sí.



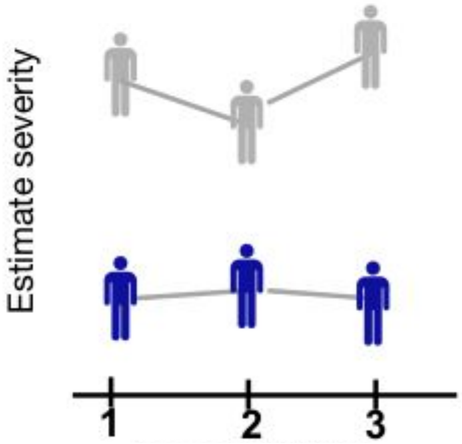
Calidad de las estimaciones - Cont.

Confiabilidad inter-evaluador



Reproducibilidad

Confiabilidad intra-evaluador



Tiempo

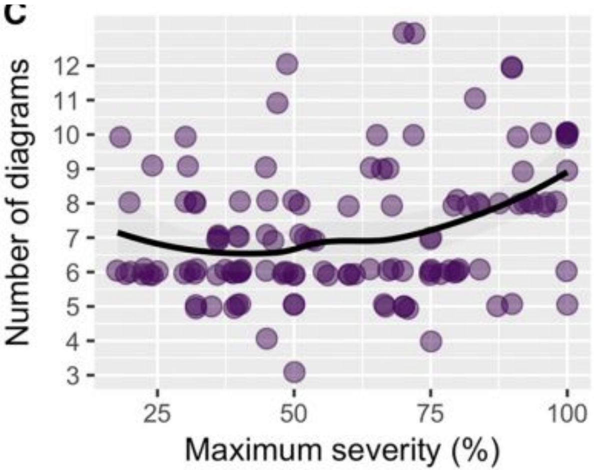
Repetibilidad

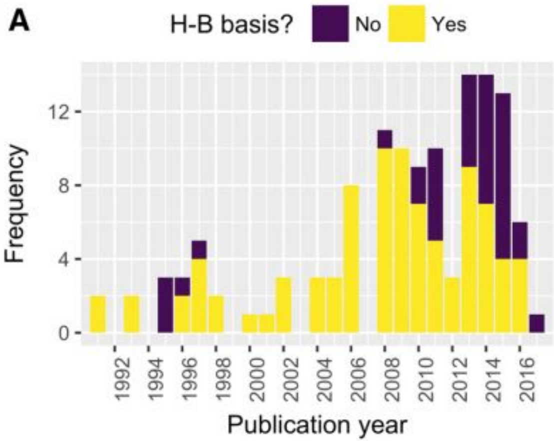
CCC general <OCCC, Overal...>

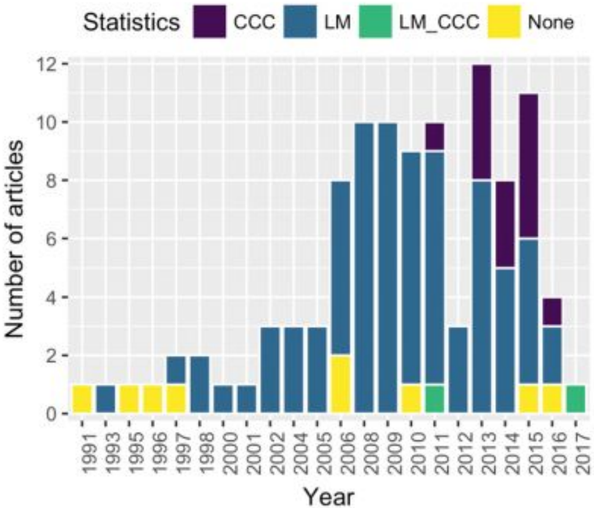
Lin’s CCC es un índice de concordancia en el contexto de comparar dos observadores fijos. Para utilizar múltiples observadores en un estudio que involucre a un gran número de sujetos, es necesario evaluar el acuerdo entre estos múltiples observadores. OCCC permite evaluar la concordancia entre múltiples observadores fijos.

Standard Area Diagrams for Aiding Severity Estimation: Scientometrics, Pathosystems, and Methodological Trends in the Last 25 Years

Emerson M. Del Ponte,† Sarah J. Pethybridge, Clive H. Bock, Sami J. Michereff, Franklin J. Machado, and Piérri Spolti







BOX 1

Best practices for conducting and reporting studies on the validation of standard area diagrams (SAD) for aiding visual assessment of plant disease severity

- State and justify clearly the need and importance for designing SAD for a specific disease.
- Test hypothesis (see knowledge gaps in the discussion) related to SAD design and evaluation that may affect gains in accuracy and reliability of the severity estimates when using the aid.
- Sample a minimum number (e.g., $n = 100$) of specimens from natural epidemics representing the range of disease severity and typical symptoms observed. Do not use damaged, blemished specimens or those infected with other diseases. If the symptoms are induced by artificial inoculation, make clear limitations of use of the tool for assessing severity in field samples.
- Use standard image analysis software optimized to recognize and discriminate disease symptoms from healthy areas to calculate actual severity. If new tools are used or developed, provide evidence that the measurements are comparable with standard software.
- When designing the illustrations for the SAD set, ensure that the individual diagrams are prepared realistically, whether line drawn, actual photos, or computer generated. It is critical that the diagram allows the rater to discriminate easily between the diseased and healthy areas. The diagrams should preserve the characteristics of the symptoms for the range of severity found in the field.
- It is not established how many diagrams are required in a SAD set. We suggest that the number of diagrams should be no less than 6 and no more than 10, distributed approximately linearly, and spaced no more than 15% apart. Additional diagrams (±2) should be included between 0 and 10% severity, considering the tendency of raters to overestimate severity in this range, especially when symptoms consist of numerous small lesions.
- For validation, select at least 50 specimens representing the full range of actual severity and symptom patterns, because there may be variable patterns of symptoms for similar severity values. Provide data on the age, position, and size of the specimens, which could further affect the utility of the SAD in some situations.
- When selecting raters for validation, make sure they do not have previous experience in using the SAD under evaluation. Provide standard instructions on how to recognize the symptoms of the disease and how to assess severity, first without and then with the SAD. It is not clear how many raters are required but we suggest a minimum of 15 raters selected randomly.
- Ideally repeat the assessment in time, with a 1- or 2-week interval, both without and with the aid, using the same set of raters in order to evaluate the effect of training and experience on gains in accuracy. One or more assessments could be made with a different set of randomly selected raters and the data analyzed separately to check consistency of the results obtained.
- Both pre- and posttest experiment conditions should be the same to avoid any impact of distraction on accuracy of estimates during the tests.
- To evaluate the effect of SAD on accuracy components, analyze the data, preferably using concordance analysis methods, to fully explore which component is affected and to gain insight into the ramification of errors. Linear regression should not be used as the sole method but it could be complementary for comparison with previous literature. Inferential methods should be used for testing hypotheses related to gain in accuracy and reliability. If parametric tests are used, make sure to check that the assumptions are not violated. Alternatively, nonparametric bootstrapping should be used when the conditions for parametric tests are not met. Reliability (intrarater or interrater) analysis should also be performed using concordance or intraclass correlation methods.
- When preparing the report, we encourage the use of reproducible research practices, including the availability of the raw data and the computer codes used for the analysis, which can be provided as supplementary materials and ideally hosted in a public repository.

Estudio de caso

[Franceschi, 2020](https://osf.io/3zmv8/)
https://osf.io/3zmv8/

Received: 8 October 2019 | Accepted: 1 January 2020
DOI: 10.1111/ppa.13148

ORIGINAL ARTICLE

Plant Pathology WILEY

A new standard area diagram set for assessment of severity of soybean rust improves accuracy of estimates and optimizes resource use

Vinicius T. Franceschi[1] | Kaique S. Alves[2] | Sergio M. Mazaro[1] | Cláudia V. Godoy[3] | Henrique S. S. Duarte[4] | Emerson M. Del Ponte[2]

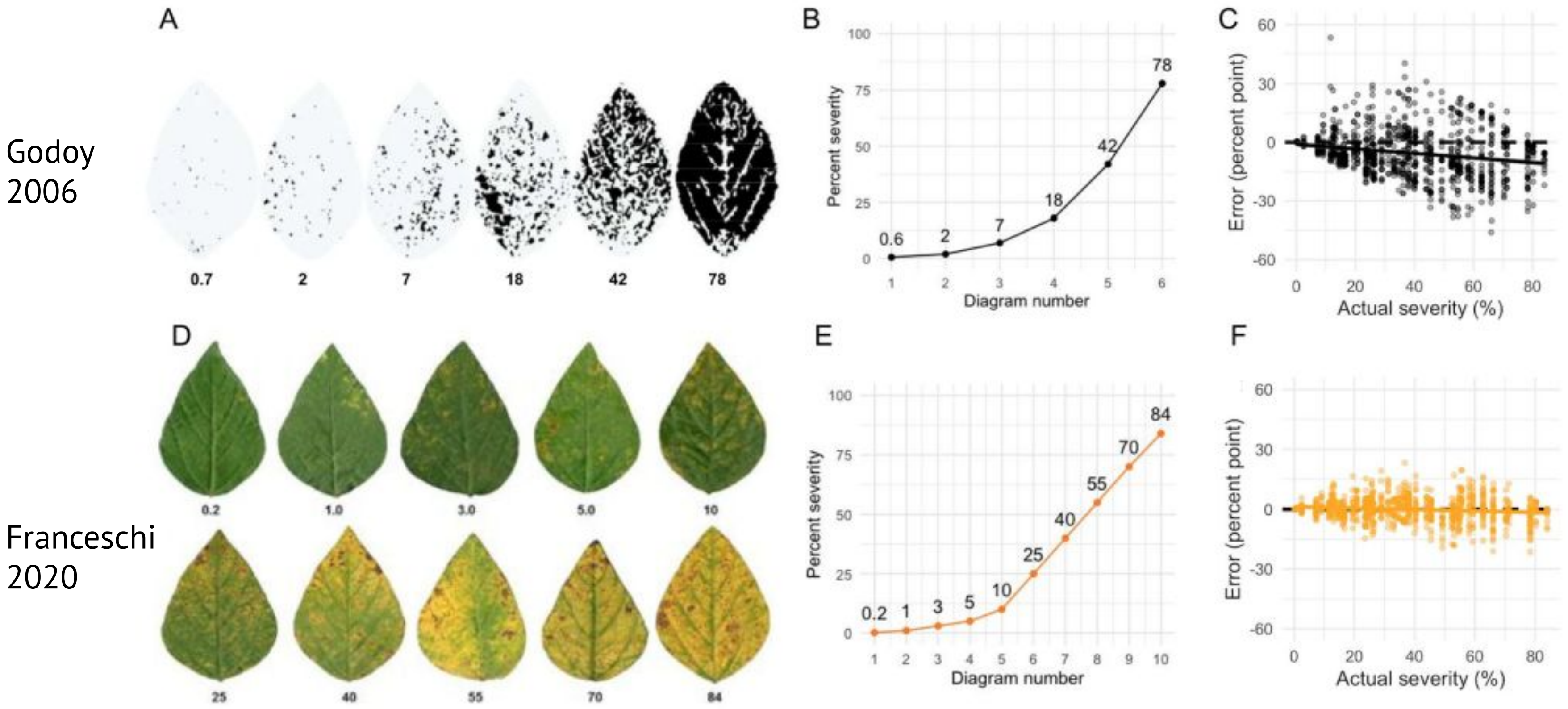
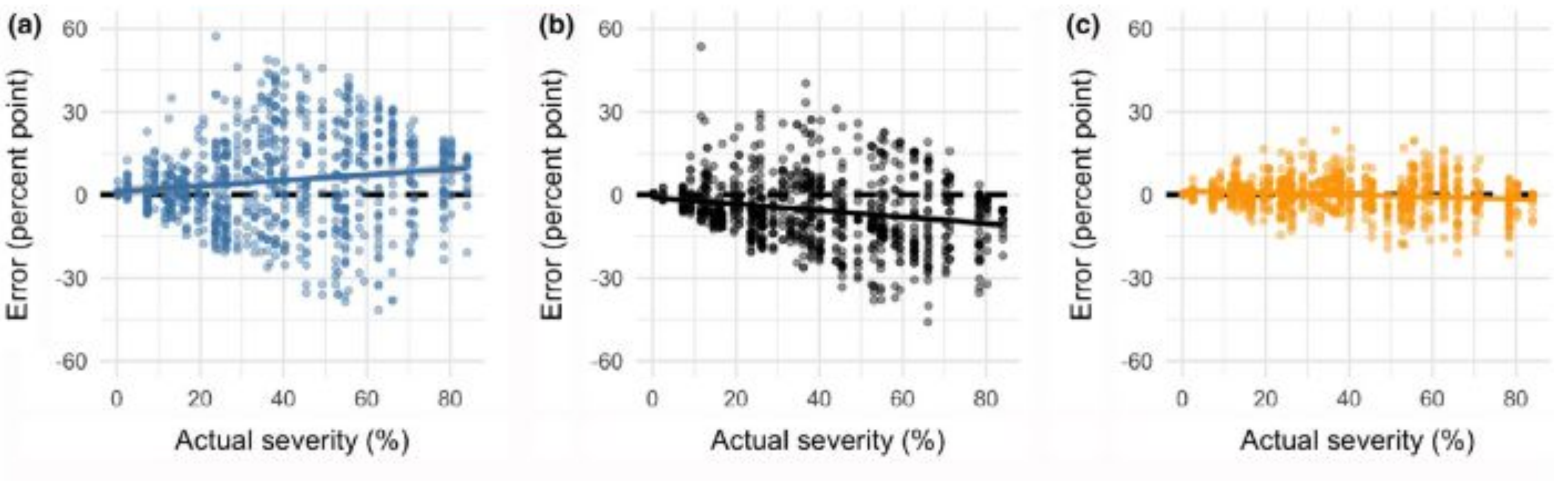


Fig. 4 Standard area diagrams (SADs) to estimate severity of rust (*Phakopsora pachyrhizi*) on soybean (*Glycine max*) leaves. **A** The original SADs (Godoy et al. 2006) **B** the relationship between the illustated SAD severity and diagram number for the original SAD **C** the absolute errors of estimates when using the original SADs **D** the newly developed and validated SADs (Franceschi et al. 2020) that is a tool for more accurate estimates of rust severity **E** the relationship between the illustated SAD severity and diagram number for the newly developed SAD **F** the absolute errors of estimates when using the newly developed SADs. The numbers under each leaf represent actual percentage leaf area showing symptoms (necrosis and chlorosis)

Sin Auxilio | Godoy et al., 2006 | Franceschi et al, 2020



Franceschi, 2020

Method	N	u^{a}	υ^{b}	C_b[c]	r^{d}	ρ_c[e]
New SAD	18	0.01 a	0.99 a	0.99 a	0.96 a	0.96 a
Old SAD	19	−0.26 b	0.98 a	0.92 b	0.77 b	0.83 b
Unaided New	18	0.19 c	1.23 b	0.92 b	0.77 b	0.82 b
Unaided Old	19	0.21 c	1.23 b	0.90 b	0.72 b	0.80 b

Note: Means followed by the same letter in the column are not significantly different (Tukey's HSD, 5% level).

[a]Location shift (u, 0 = no bias relative to the concordance line).

[b]Scale shift (υ, 1 = no bias relative to the concordance line).

[c]Bias correction factor (C_b) measures how far the best fitted line deviates from 45° and is a measure of accuracy.

[d]Correlation coefficient as a measure of precision (r).

[e]Lin's concordance correlation coefficient (LCC), that combines both measures of precision (r) and accuracy (C_b) to measure overall accuracy (agreement) with the true value.

TABLE 3 Measures of inter-rater reliability of severity estimates by 37 inexperienced raters during two assessments unaided, or with the use of a new standard area diagram set (New SAD) or an Old SAD (Godoy *et al.*, 2006) as an aid to assessment of disease severity

Method	Intra-class correlation coefficient (ICC) ρ [95% CI][a]	Overall concordance correlation (OCC)[b]
New SAD	0.94 [0.92–0.96]	0.940
Unaided New	0.83 [0.77–0.88]	0.759
Old SAD	0.81 [0.74–0.87]	0.736
Unaided Old	0.83 [0.77–0.89]	0.746

[a]Calculated with decisions of ICC model described elsewhere (Schwanck and Del Ponte, 2014).

[b]Overall agreement statistics based on Lin (1989) and Barnhart *et al.* (2002) to evaluate agreement among multiple observers.

Franceschi, 2020