Proceedings World Geothermal Congress 2020+1
Reykjavik, Iceland, April - October 2021

# Statistical Validation Method for Soil Thermal Response Test Comparison Between Vertical and Horizontal Exchangers

Valentín A. Stefanini

Av Lafinur 301, CP 5700 San Luis, Capital - Argentina

Stefanini824@gmail.com

**Keywords:** Statistical comparison method, thermal response test

## ABSTRACT

This work was carried out in order to put forth a statistical comparison between two calibration lines of Thermal Response Test (TRT) in situ. The calibration lines were substantially different in design, time and place. A statistical method was used to compare straight-line slopes based on variance homogeneity analysis.

In order to determine whether the variances were homogeneous or not, the F-Test was performed then the T-Test evaluating similarity between slopes. The F-test, determined that the slopes were not homogeneous. Afterwards, the T-test, for non-homogeneous slopes, with 95%, 97.5% and 99% probability, determines that these slopes were different; but the T-test, with a probability of 99.75%, established similarity between both slopes.

This established that this method in the particular horizontal exchanger to obtain an effective soil heat transfer coefficient is worthwhile.

## 1. INTRODUCTION

This work aim is to verify the horizontal exchanger (TRT) obtained results compared with a vertical ones using a mathematical model consisting of the heating transfer process with "λeff" and "Rb" as adjustment variables developed by, Ingersoll, L.R. and Plass, H.J. (1948), Carslaw and Jaeger (1959) and put into Eklöf C, Gehlin S. (1996) thesis were taken as references.

A horizontal exchanger, built in San Luis (Argentina), was tested using the Kelvin Infinite Line Source model (ILS) was used to determine the effective soil thermal conductivity coefficient "λeff" and the soil thermal resistance "Rb" with the slope method.

The reference line corresponds to a TRT realized at Norwest National University (UNNe) belonging to a vertical standard exchanger which was compared to the line obtained by a San Luis National University (UNSL) particular horizontal exchanger.

The horizontal exchanger is schematized in Figure 1, it was adapted to an existing excavation which had been made for another purpose, where the horizontal exchanger was put in the middle of a 15 cm thickness cement layer. The lower part of the layer was on the subsoil and the upper part contained a water pool, so its characteristics were different and asymmetric.

## 2. DEVELOPING

### 2.1. Why it should be compared two slopes?

Comparison between slopes as result of different tests in time, place and configuration enables us to validate one with respect to a standard reference.

### 2.2. The validation method allows us:

- To study a method's reproducibility (Compare straight-line slopes obtained from different moments)
- To study a method's robustness (Compare straight-line slopes obtained in different conditions)
- To prove whether the sample has matrix effect (Comparing the obtained straight-line slope with a pattern one)

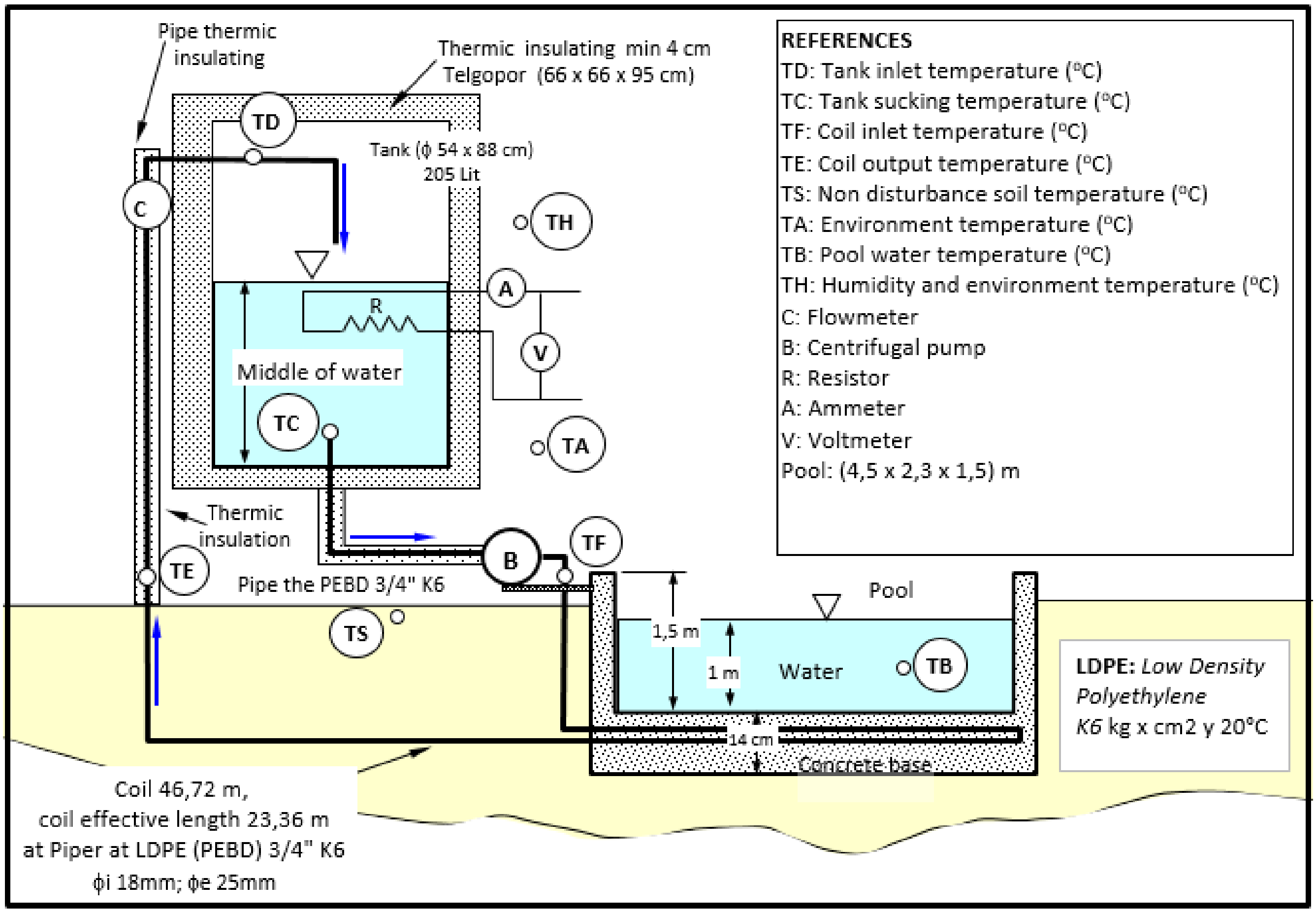


**Figure 1. Scheme of the measuring device that was set in San Luis – Argentina.**

## 3. METHODOLOGY

A statistical method was applied to compare the straight-line slope calibration, whose method was based on variance homogeneity analysis.

The first step, an F-Test was employed to determine variances homogeneity, once it was finished, on a second step a T-Test was performed to verify similarity of the straight-line slope.

## 4. PROCEDURE

Straight-line slopes belong to two different soil TRT, the first from Corrientes National University and the second comes from San Luis National University. The first study was developed for a vertical exchanger in which a vertical standard mathematical model to process data was used. This model was used to process data from the test realized in the San Luis University horizontal exchanger.

### 4.1. Starting Conditions

The test must meet several conditions and then slopes can be compared for obtaining reliable statistical results:
1. - Provide data base for both tests to be compared.
2. - Verify that independent variable units be equal (x axis), the time sample as well.
3. - The starting point and ending point for both tests must be the same.
4. - Data amount for both tests must be equal.
5. – As of the precedent items, it is possible to obtain the regression lines and mean quadratic (squared) error for both tests.
6. – The parallelism of both slopes must be done first by visual comparison.
7. - Mean quadratic (squared) errors must be similar.

### 4.2. Suggestions

Considering that the two data bases allow us to determine regression lines on the independent variable, in this case “time”. Both tests must have the same time control base (10 minute).
Taking care that the parallelism be as mentioned above in order to be efficient.
If mean squared errors are not similar then the slopes cannot be compared.
Time criterion, established by the mathematical model, rejects lower time data.
The time criterion used in San Luis test was nine (9) hours.
In this analysis, two slopes were compared with starting point from 9 to 72 hours.

### 4.3. TRT from Corrientes

It starts at time zero to time 426600 seconds (118.5 hours) with data collection per minute.

### 4.4. TRT from San Luis

It starts at time zero until the time 259200 seconds (72 hours) with data collection every 10 minutes.

### 4.5. Criterion unification

Regarding the previous points, both regression lines were compared between 32400 and 259200 seconds with time criterion and a 10 minutes period corresponding to TRT SAN LUIS, and data from TRT CORRIENTES were considered in the same period in order to compare them.

### 4.6. Corrientes TRT regression line

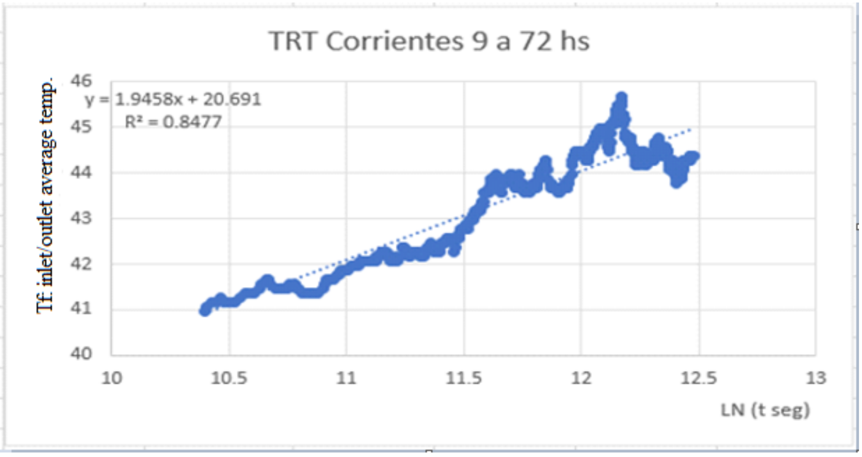


**Figure 2: Corrientes TRT graph with time criterion from 9 to 72 hs.**

### 4.7. San Luis TRT regression line

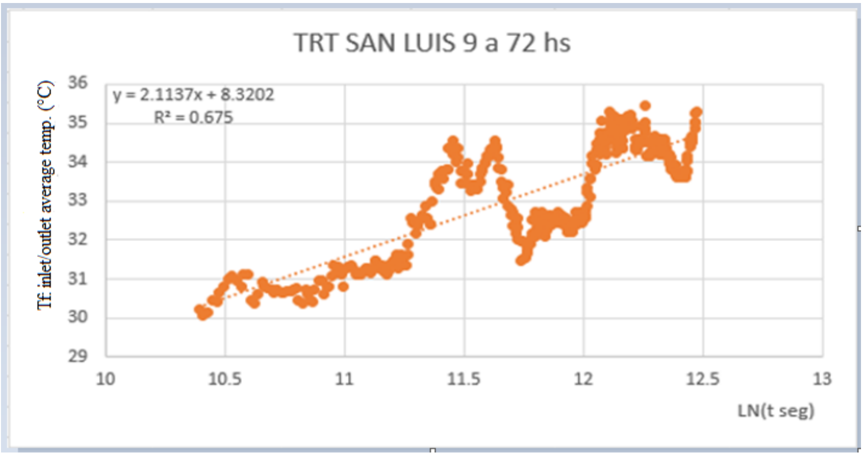


**Figure 3: San Luis TRT Graph with time criterion of 9 to 72 hs.**

## 5. DATA REQUIRED TO PUT FORTH IN TEST

Spreadsheet table showed for both tests.

### 5.1 Data Corrientes

Corrientes TRT statistical data

| | | | |
|---|---|---|---|
| Slope | 1,945774992 | y-intercept | 20,69127688 |
| Sb: Desv. Slope | 0,042469074 | Sa: Desv. Intercept | 0,500035226 |
| R^2 | 0,847746375 | Sxy: Desv. Line | 0,450208327 |
| Freedom Degree | 377 | ∑{[Y-Y(est.)]^2} | 76,41320177 |

**Table 1: Corrientes TRT statistical data**

Corrientes TRT regression line equation

$$Y(est.Corr) = [(1{,}95 \pm 0{,}042) * X + (20{,}7\ \pm 0{,}5)] \quad (1)$$

**5.2 Data San Luis**

San Luis TRT statistical data

| | | | |
|---|---|---|---|
| Slope | 2,11367936 | y-Intercept | 8,32021568 |
| Sb: Desv. Slope | 0,07553806 | Sa: Desv. Intercept | 0,8893928 |
| R^2 | 0,67499235 | Sxy: Desv. Line | 0,80076768 |
| Freedom degree | 377 | ∑{[Y-Y(est.)]^2} | 241,743285 |

**Table 2: San Luis TRT statistical data**

San Luis TRT regression line equation

$$Y(est.SL) = [(2{,}11 \pm 0{,}075) * X + (8{,}32\ \pm 0{,}88)] \quad (2)$$

**5.3 Data Summary**

Slope for both lines:

| | |
|---|---|
| Bcorr= 1,95 | Bsl= 2,11 |

**Table 3: Slopes**

Standard Deviation for Slopes

| | |
|---|---|
| $S_{Bcorr} = 0{,}042$ | $S_{Bsl} = 0{,}075$ |

**Table 4: Standard Deviation for slopes.**

Slope Variances

| | |
|---|---|
| $S^2_{\mathrm{Bcorr}} = 0{,}0018$ | $S^2_{\mathrm{Bsl}} = 0{,}0057$ |

**Table 5: Slope Variances**

Amount of numbers for each line:

| | |
|---|---|
| Ncorr = 379 datas | Nsl = 379 datas |

**Table 6: Amount of data to consider**

Note:
The amount for San Luis was considered with 379 data that were taken every 10 minutes beginning from 9 hours, when the test started, until the 72 hours proposed.

## 6. STATISTICAL APPLICATION

### 6.1. Statistic tested at 95% of probability

6.1.1 Statistic calculation of "F value"

The variance homogeneity F-test for both slopes will be confirmed.
Statistical calculation of "F value" which will be named Fcalc:

$$F_{calc.} = \frac{San\ Luis\ Variance}{Corrientes\ Variance} \quad (3)$$

$$F_{calc.} = \frac{S^2_{\mathrm{Bsl}}}{S^2_{\mathrm{Bcorr}}}$$

$$F_{calc.} = \frac{0{,}0057}{0{,}0018} = 3{,}16$$

Note
Fcalc. must be always bigger than 1 (one), that will determine which variance to put forth to the numerator and which one to the denominator.
In this case San Luis variance is bigger than Corrientes one.

Freedom degree calculation for numerator and denominator

$$U_{num/den} = N - 2 \quad (3)$$

Numerator Freedom Degree

$$U_{numerator} = N_{sl} - 2 = 379 - 2 = 377$$
$$U_{sl} = 377$$

Denominator Freedom Degree

$$U_{denominator} = N_{corr} - 2 = 379 - 2 = 377$$

$$U_{corr} = 377$$

Tabbed F value determination, Ftab. 95%, Usl, Ucorr
Ftab value at 95% of probability has 377 freedom degree for numerator and 377 freedom degrees for denominator. See table "F"

Ftab at 95%,377,377

"x axis" SL," y axis" Corr
With software "R" is obtained value of table F.95,377,377

$$F_{.95,377,377} = 1.184$$

$$F_{calc.}(3,16) > F_{tab.}(1,184)$$

It means that variances are NOT HOMOGENEOUS

Note:
Otherwise where Fcalc is MINOR than Ftab, then variances ARE HOMOGENEOUS.

Deduce if slopes are statistically similar with a T-test for NOT HOMOGENEOUS VARIANCES AT 95%.

<u>6.1.2. Statistical calculation of "T value"</u>
The T test is done as follows

Statistical calculation for tcalc.

$$t_{calc.} = \frac{|B_{sl.} - B_{corr.}|}{\sqrt{S_{sl.}^2 + S_{corr.}^2}} \quad (4)$$

$$t_{calc.} = \frac{|2{,}11 - 1{,}95|}{\sqrt{0{,}0057 + 0{,}0018}} = 2{,}73$$
$$t_{calc.} = 2{,}73$$

Statistical calculation for t"calc.

$$t''_{calc.} = \frac{t_{sl} * S_{\mathrm{Bsl}}^2 + t_{corr} * S_{\mathrm{Bcorr}}^2}{S_{\mathrm{Bsl}}^2 + S_{\mathrm{Bcorr}}^2} \quad (5)$$

Where: tsl and tcorr are taken from the table.
Calculation for ttab.,sl

$$t_{tab.95\%,Nsl.-2gl.}$$

San Luis freedom degree calculation

$$Fredom\ Degree = N_{sl.} - 2$$

$$Fredom\ Degree = 379 - 2 = 377$$

With software "R" is obtained the table value for $t_{.95,377}$ for San Luis

$$t_{.95,377} \cong 1{,}645$$

Calculation for ttab.,corr.

$$t_{tab.95\%,Ncorr.-2gl.}$$

Corrientes freedom degree calculation

$$\text{Fredom Degree} = N_{corr.} - 2$$

$$Fredom\ Degree = 379 - 2 = 377$$

From "t" table or from software "R" we get a probability of 95% and with a freedom degree of 377 for the Corrientes value $t_{.95,377} \cong 1{,}645$
Statistical calculation for t"calc..

$$t''_{calc.} = \frac{1,645 * 0,0057 + 1,645 * 0,0018}{0,0057 + 0,0018} = 1,62$$

$$t''_{calc.} = 1,62$$

$$t_{calc.} = 2,73$$

$$t_{calc.} = 2,73 > t''_{calc.} = 1,62$$

It means that slopes are different for 95%

**6.2. Statistic tested at 97.5% of probability**

For a 97.5% probability, the F test determines that the variances continue being non-homogeneous.

Deduce if slopes are statistically similar with a T-test for NOT HOMOGENEOUS VARIANCES AT 97.5%.

ttab for San Luis

$$t_{.975,377} = 1.966276$$

ttab for Corrientes

$$t_{.975,377} = 1.966276$$

$$t''_{calc.} = \frac{1.966 * 0,0059 + 1.966 * 0,0018}{0,0059 + 0,0018} = 1,97$$

$$t_{calc.} = 2,73 > t''_{calc.} = 1,97$$

Slopes are statistically different for 97.5% with variances NOT HOMOGENEOUS.

**6.3. Statistic tested at 99% of probability**

For a 99% probability, the F test determines that the variances continue being non-homogeneous.

Deduce if slopes are statistically similar with a T-test for NOT HOMOGENEOUS VARIANCES AT 99%.

ttab for San Luis

$$t_{.99,377} = 2,33628$$

ttab for Corrientes

$$t_{.99,377} = 2,33628$$

$$t''_{calc.} = \frac{2,336 * 0,0059 + 2,336 * 0,0018}{0,0059 + 0,0018} = 2,34$$

$$t_{calc.} = 2,73 > t''_{calc.} = 2,34$$

It means that slopes are different for 99% and variances are not homogeneous.

**6.4. Statistic tested at 99.75% of probability**

Calculation for $F_{tab.}$ at 99.75% with 377 freedom degrees in numerator and denominator respectively are:
$F_{.99.75,377,377}$= 1.33

$$F_{calc.}(3,16) > F_{tab.}(1,33)$$

It means that variances still are NOT HOMOGENEOUS

Deduce if slopes are statistically similar with a T-test for NOT HOMOGENEOUS VARIANCES AT 99.75%.

ttab for San Luis

$$t_{.99.75,377} = 2,823$$

ttab for Corrientes

$$t_{.99.75,377} = 2,823$$

$$t''_{calc.} = \frac{2{,}823 * 0{,}0059 + 2{,}823 * 0{,}0018}{0{,}0059 + 0{,}0018} = 2{,}82$$

$$t_{calc.} = 2{,}73 < t''_{calc.} = 2{,}82$$

Slopes are statistically similar at 99.75% and the variances are not homogeneous.

## 7. CONCLUSIONS

The slopes for the test realized in San Luis province are statistically similar to the slope in Corrientes test at 99.75% certainty, but is not the case for 95%, 97.5% and 99% with parameter unification and taking the time criterion of San Luis.

Therefore, we can assert that the method used in the particular horizontal exchanger of San Luis which was put forth the Kelvin Infinite Line source (ILS) to determine the subsoil effective thermal conductivity coefficient λeff, and the subsoil thermal resistance Rb with slope method were worthwhile.

With this new perspective to compare the TRT, generated by the author, in an excavation realized with other objectives, it is possible to fit a geothermal heat exchanger with different designs adapted to the place.

Later I propose that, in this way with an easy, fast and economical TRT you can determine the subsoil characteristics and thereafter you can make the proposed comparison which would validate the mentioned parameters.

If that validation is correct the geothermal entrepreneurship will be designed based on real values.

This method would ensure an efficient design which would help you not to over-size or sub-size the equipment by means of saving money, therefore, giving a great prestige to the low enthalpy geothermal energy (Stefanini V. 2018).

In conclusion, we can affirm the method's reproducibility and robustness for the specified parameters.

## REFERENCES

[1] Carslaw H. S., Jaeger J. C., Conduction of Heat in Solids (Oxford science publications, 1964)

[2] Busso A., Reuss M., Müller J., Sograri N. (2000), Almacenamiento Subterráneo de Energía Térmica: Resultado de un ensayo de respuestas térmicas del subsuelo. Avance en Energías Renovables y Medio Ambiente, Vol. 4, Nº 1.

[3] Busso A., Reuss M., (2000), Ensayo de Respuestas Térmicas: Método de Evaluación de la Pendiente vs Ajuste Con Dos Parámetros Variables. Dpto. de Física – Facultad de Ciencias Exactas y Naturales, UNNE – Corrientes Argentina.

[4] Eklöf C, Gehlin S. (1996). A Mobile Equipment for Thermal Response Test. Master of science thesis 1996:198E. Luleå University of Technology. Sweden.

[5] Gehlin, S. and Nordell, B. 1997. "Thermal Response Test - a Mobile Equipment for Determining Thermal Resistance of Borehole." Proc. Megastock '97 Sapporo, pp. 103-108

[7] Ingersoll, L.R. and Plass, H.J. 1948. "Theory of the ground pipe heat source for the heat pump." Heating, Piping & Air Conditioning 20/7, pp. 119-122

[8] John E. Freud, Ronald E. Walpole. Estadística Matemática con aplicaciones. Printice Hall Hispanoamericana. 1990. pp 494-498.

[9] Miler y Freud, Probabilidad y Estadística para Ingenieros, Prentice Hall. 1997. pp 356-360.

[10] Meseguer Lloret S., Ensayos estadísticos para comparar las pendientes de 2 rectas de calibrado, 2017.

[11] The R Project for Statistical Computing. www.r-project.org

[12] Estadística Básica para Ingenieros, Luis María Dicovskiy Riobóo. Editorial Uni Norte. 2012.

## NOMENCLATURE

TRT: Thermal Response Test
Bcorr and Bsl.: Corrientes and San Luis city slope
$S_{Bcorr}$ and $S_{Bsl}$: Slope Standard Deviation
$S^2_{\text{Bcorr}}$ and $S^2_{\text{Bsl}}$: Slope Variances
Ncorr and Nsl : Amount of numbers for each line
$F_{calc.}$: "F-value" Statistic calculation
$U_{num/den}$ : For numerator and denominator freedom degree calculation
Ftab at 95%,377,377: value at 95% of probability with 377 freedom degree for numerator and 377 freedom degrees for denominator. See table "F"
$t_{calc.}$: Statistic calculation for $t_{calc.}$
t"calc.: Statistic calculation for t"calc
$t_{tab.95\%,Nsl.-2gl.}$ : tabulated value of "t", for a probability of 95% and for a certain number of degrees of freedom for each regression line.