

EL ROL DEL ANÁLISIS FACTORIAL EN LA EVALUACIÓN DE ENSAYOS DE RESPUESTA TÉRMICA DE SUELO - SERPENTÍN VERTICAL/HORIZONTAL CYPRUS/CORRIENTES/SAN LUIS

V. Stefanini, M. Cesarini, D. Nazario, R. Guaycochea

Proyecto de Investigación: “Sistemas de climatización a partir de la geotermia de baja entalpía, utilizando el subsuelo como fuente de energía almacenada, para economizar otras energías sin contaminación.”

Universidad Nacional de San Luis - Facultad de Ciencias Económico, Jurídica y Sociales
(FCEJS - UNSL) - CeReCoN (Centro Regional de Investigación y Desarrollo en Computación y Neuroingeniería) – Area de Infometría, Lexicometría e Ingeniería del Conocimiento – Laboratorio ReAVi (Realidad aumentada y Virtual) – Universidad Tecnológica Nacional FRM
Tel: 266-4667741, 266-4008419

E-mail: stefanini28@hotmail.com.ar, matilde.cesari@frm.utn.edu.ar, ronioguaycochea@gmail.com

RESUMEN: A partir de los resultados que se obtuvieron de un Test de Respuestas Térmicas (TRT), en el cual se determinaron los valores de la conductividad térmica efectiva del subsuelo λ_{eff} , y la resistencia térmica R_b , correspondiente a un sistema horizontal de intercambiador particular de calor, realizado en la provincia de San Luis-Argentina. Este trabajo se realizó con el objetivo de validar los parámetros establecidos del ensayo propuesto, se aplicó un Análisis Factorial de Componentes Principales (ACP), entre tres TRT, sustancialmente diferentes en diseño, lugar y tiempo, de serpentines verticales vs Horizontales. Así mismo se estudió cada serie en 7 rangos, a través de un Análisis Factorial de Correspondencias Múltiple (AFCM). El ACP aplicado, permitió demostrar que la correspondencia entre los valores obtenidos por los tres TRT estudiados es superior al 99 %, por lo cual estadísticamente, el método usado para obtener el coeficiente de transferencia de calor efectivo del suelo es válido para el intercambiador horizontal propuesto. A su vez el análisis no lineal del AFCM demostró que los tres test tienen la misma variación en el tiempo, a pesar de las diferentes magnitudes en las diferencias de temperaturas observadas.

Palabras clave: estadística, análisis de componentes principales, análisis de correspondencias, test de respuesta térmica.

INTRODUCCIÓN

El objetivo del presente trabajo es poder validar los resultados obtenidos de un Test de Respuestas Térmicas in situ (TRT) de un intercambiador particular horizontal, respecto de dos intercambiadores estándar verticales, aplicando un modelo matemático desarrollado para estos dos últimos.

En la Universidad Nacional San Luis (UNSL), Argentina, se realizó un intercambiador particular horizontal, y se lo ensayó aplicando el modelo de la fuente de línea infinita de Kelvin (LSM) (Line Source Model)¹, y a partir de éste se determina el coeficiente de conductividad efectivo de subsuelo λ_{eff} , y la resistencia térmica de subsuelo R_b , por el método de la pendiente.

Suscitando una controversia respecto si los resultados obtenidos en dicho ensayo son fiables, por lo cual se consideró en establecer si hay una correlación entre el TRT particular de San Luis, con los TRT estándar (vertical y simétricos) correspondientes a la UNNE, Corrientes, Argentina y la Universidad de Cyprus.

¹ Ingersoll L.R. y Plass H.J. (1948). Theory of the ground pipe heat source for the heat pump

Para ello se aplica una de las técnicas estadísticas multivariantes: Análisis de Componente Principales (ACP), que como lo indicaron Sánchez, López, Cruz, Gutiérrez, Posada, Zamora, Torrijos, & Osorio Lama (2016)², tiene como propósito reducir el número de dimensiones en un conjunto de variables, tratando de conservar la mayor información posible. Este análisis ACP reduce las dimensiones de los datos multivariados a dos o tres componentes principales que se pueden visualizar gráficamente con una pérdida mínima de información³.

El ACP fue propuesto por Pearson (1901)⁴ a modo de solución para algunos problemas que eran de interés para la biometría de la época. En dependencia del campo de aplicación, las componentes principales suelen ser conocidas también como la transformada discreta de KarhunenLoève (análisis de señales), funciones bases ortogonales empíricas (meteorología e investigaciones atmosféricas), índices semánticos latentes (recuperación de datos). Actualmente constituye una herramienta esencial para el análisis de datos multivariados y la reducción de la dimensión; de hecho, prácticamente la totalidad de los libros de texto sobre análisis de datos multivariados cubren el tema. El análisis de correspondencias simples AFCS, se utiliza a menudo en la representación de datos que se pueden presentar en forma de tablas de contingencia de dos variables nominales u ordinales. Otras utilidades implican el tratamiento de tablas de proximidad o distancia entre elementos. La extensión del análisis de correspondencias simples al caso de varias variables nominales (tablas de contingencia multidimensionales) se denomina Análisis de Correspondencias Múltiples ACM⁵, y utiliza los mismos principios generales que la técnica anterior. El ACM, a diferencia de las formas usuales de representación de datos, permite visualizar de manera simultánea todas las categorías de una tabla, poniendo de relieve cuáles tienen mayor o menor probabilidad de ocurrir al mismo tiempo⁶.

La idea fundamental del ACP y AFCS, es encontrar una secuencia de vectores ortogonales que expliquen de la forma más eficiente la varianza de las observaciones, es decir, reducir la dimensión del conjunto de datos conservando tanto como sea posible la variación presente en los mismos.

Para aplicar el ACM sobre tramos o rangos de las series, primero es necesario particionar cada serie, para lo cual un método basado en la varianza y/o inercia, dentro de los métodos multivariados en el algoritmo óptimo de Fisher. El algoritmo de Fisher está inspirado en la programación dinámica, en términos generales, la programación dinámica requiere optimizar un criterio numérico aditivo w que dependa de m pasos. Pese a que el algoritmo de Fisher tiene la limitación de producir una clasificación óptima sobre datos univariados, tiene gran importancia como método previo a otras técnicas de particionamiento para datos multivariados, tales como los métodos de segmentación y los métodos de discriminación, tomando la variable cuantitativa como la variable a explicar y las demás como variables explicativas⁷.

El análisis de componentes principales propuesto se realizó a partir de los datos correspondientes a los TRT verticales estándar, realizado en la Universidad Nacional del Noreste (UNNe), el realizado en la Universidad de Cyprus y el TRT realizado en la Universidad Nacional de San Luis (UNSL), esquematizado en la (Fig. 1), el mismo fue adecuado a una excavación ya existente para otro fin, quedando empotrado en el medio de un bloque de cemento de 15 cm. de espesor. La parte inferior del bloque está sobre el subsuelo y la parte superior contiene agua de una pileta, por ello es que sus características son diferentes.

Para profundizar el análisis y estudiar las relaciones no lineales se particionó cada serie, según la varianza y se obtuvieron 17 rangos que se analizaron a través del ACM.

²Sánchez López, A., Cruz, Gutiérrez V., Posada, Zamora, M. A., Torrijos, M. T., y Osorio, Lama, M. A. (2016). Estudio del análisis de componentes principales. *Research in Computing Science*, 9-19

³Kassambara A. (2017). *Practical Guide to Principal Component Methods in R*.

⁴Pearson K. (1901). On lines and planes of closest fit to systems of points in space. *Phil. Mag.* 2(6), pp. 559-572

⁵Benzecri, J.P. (1992). *Correspondence Analysis Handbook*. CRC Press,

⁶Langrand C. y Pinzón L.M. (2009). *Análisis de datos. Métodos y ejemplos*. Escuela Colombiana de Ingeniería Julio Garabito, Bogotá, Colombia. ISBN: 9789588060903, 388 páginas.

⁷Espinoza, José. (2012). Partición óptima: el algoritmo de Fisher. *Revista de Matemática: Teoría y Aplicaciones*. 4. 77. 10.15517/rmta.v4i1.142.

Esquema del dispositivo

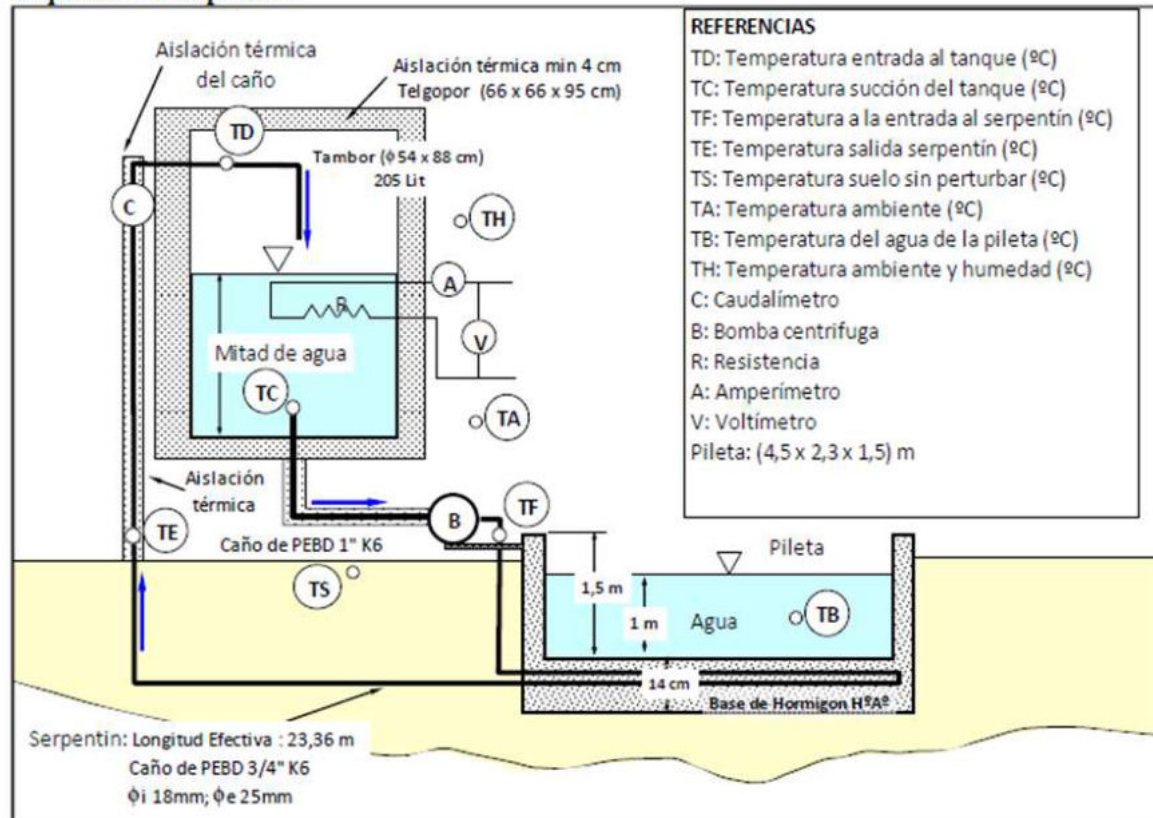


Fig. 1 Esquema del dispositivo de medición montado en la Universidad Nacional San Luis – Argentina

Las bombas de calor geotérmicas (GSHPs) (ground source heat pumps) son sistemas eficientes para la calefacción y refrigeración de edificios, ampliamente usados en Europa y EE. UU. Estos consisten en acoplar una unidad de bomba de calor a un grupo de intercambiadores de calor de pozo (BHEs) (bore hole heat exchanges).

Las mejoras que se pueden realizar a estos equipamientos de bomba de calor geotérmico están centradas en la optimización del mismo y en la reducción de los costos de instalación; Como estos costos de instalación pueden ser elevados, antes de poner en marcha esta construcción se necesita realizar el cálculo y la simulación para una mayor optimización de las instalaciones a construir. Para ello se deben conocer dos parámetros que afectan la eficiencia del depósito de calor, los cuales se deben determinar entre el fluido portador de calor, y la pared del pozo, ellos son: la conductividad térmica efectiva del subsuelo λ_{eff} , y la resistencia térmica de subsuelo R_b .

Teniendo en cuenta que tanto la conductividad térmica del suelo (λ_{eff}), como la resistencia del subsuelo (R_b) deben ser valores pequeños para asegurar la mejor transferencia de calor entre el disipador y el medio, (material de la cañería, y el subsuelo).

Estos valores se determinaron mediante un Test de Respuesta Térmica (TRT) "in situ", que consiste en hacer circular un fluido caloportador a través de un intercambiador de calor geotérmico. En el caso de estudio dos verticales convencionales, uno de la UNNe, Corrientes argentina, el otro de la Universidad de Cyprus donde se estableció si hay, o no, una correlación con el particular de la UNSL.

METODOLOGÍA

Este estudio se realizó en tres fases: en primer lugar, la organización de la tabla de datos, seguidamente la aplicación de las técnicas estadísticas descriptivas multivariadas, y, en tercer lugar, los análisis de los resultados obtenidos.

Organización de la tabla de datos

Entre las condiciones que deben cumplir los ensayos para poder realizar un ACP, se debe: utilizar las mismas unidades de tiempo, medidas en los mismos periodos, mismo punto de partida y finalización de los ensayos

- TRT de Corrientes. el mismo comenzó en el tiempo cero hasta el tiempo 426600 segundos, (118,5 horas), con una toma de datos por minuto.
- TRT de Cyprus. El mismo se efectuó para un intercambiador enterrado de forma vertical de 50 metros de profundidad, hecho de tubo de polietileno con 32 mm diámetro externo y 25 mm de diámetro interno, realizado en Julio del 2005 y por un tiempo de 240 horas.
- TRT de San Luis. comenzó en el tiempo cero hasta el tiempo 259200 segundos, (72 horas), con una toma de datos cada 10 minutos.

Por lo expuesto en este caso para realizar la comparación se aplicó el criterio de tiempo de San Luis el cual resultó de 9 horas. Por lo cual se compararon dos pendientes cuyo punto de inicio es el correspondiente a las 9 horas hasta el punto de culminación del ensayo de San Luis el cual fue a las 72 hs.

En un principio se comparó las dos rectas de regresión entre el tiempo 32400 y 259200 segundos correspondientes al TRT San Luis, con criterio de tiempo y en un periodo de 10 minutos, por lo cual se tomaron del TRT Corrientes y del TRT Cyprus los datos en los mismos periodos de tiempo.

Para la aplicación de las técnicas de análisis multivariante de datos, fue necesario tabular los datos en diversas tablas de distinta naturaleza (cualitativa y cuantitativa). En columna se colocó atributos numéricos medidos y atributos cualitativos discretizados en rangos, para lo cual se utilizó método de partición univariada de Fisher, clases óptimas⁸. La tabla de datos contiene en la primera columna Ln de horas, en decimal, desde que comenzó el test hasta su finalización 72 horas posteriores. La segunda, tercera y cuarta columna corresponde a la diferencia entre la temperatura de entrada y la de salida del intercambiador para el TRT San Luis, Corrientes y Cyprus, respectivamente. La columna de tiempo se transformó a 20 rangos de intervalos iguales de 5,4 minutos.

Las diferencias para los tres TRT son discretizadas en 7 grupos, cada uno con diferentes intervalos según la variación de la serie.

Aplicación De Las Técnicas Estadísticas Descriptivas Multivariadas

Para el análisis de una tabla de datos haciendo uso de métodos multivariados, se siguió la estrategia de realizar un ACP y para la interpretación - validación se utilizan las técnicas DESCO (Descripción de Variables Continuas) con el valor de prueba, "Valores Test"⁹ que son índices descriptivos construidos siguiendo la metodología de pruebas de hipótesis y la prueba exacta de Fisher¹⁰¹¹. Así mismo se aplicó el ACM a partir de los grupos discretizados con las tres series TRT.

Para el análisis factorial y de validación estadística se utilizó el paquete computacional FactoMineR¹² en R¹³, mientras que para representar gráficamente en 3D el espacio de las componentes principales se

⁸ Fisher, W.D. (1958). On grouping for maximum homogeneity. Journal American Statistics Association, N° 53, pag.789-798,

⁹Lebart, L., Morineau, A. & Piron, M. (1995), Statistique exploratoire multidimensionnelle Dunod, Paris

¹⁰Bertaut, M. B., & i Marsal, J. V. Manual de introducción a los métodos factoriales y clasificación con SPAD. Universidad Politécnica de Catalunya y Universidad Autónoma de Barcelona

¹¹ Pardo, C. E., & Del Campo, P. C. (2011). Combination of Factorial Methods and Cluster Analysis in R: The Package FactoClass. Revista Colombiana de Estadística, 30(2), 231-245

¹²Lê, S., Josse, J. & Husson, F. (2008), 'FactoMineR: an R package for multivariate analysis', J. of Stat. Soft. 25(1), 1-18.

¹³ R Development Core Team (2009), R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria. *<http://www.R-project.org>

utilizó el paquete R scatterplot3d¹⁴, La discretización de tiempo y diferencias de temperatura se realizó a través de la librería classint de R.

RESULTADOS

Resultados análisis descriptivo de las series de tiempo

A partir de la comparación de las dos rectas de regresión (Fig. 2 y Fig. 3) entre el tiempo 32400 y 259200 segundos correspondientes al TRT San Luis y TRT Corrientes, en un periodo de 10 minutos; se obtiene dos ecuaciones de la recta, con R2 de 0.85 y 0.67 respectivamente.

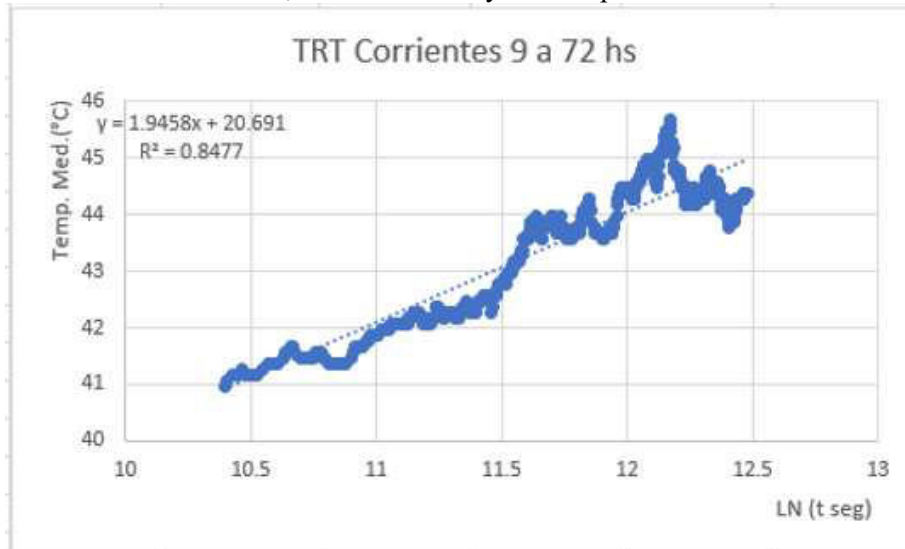


Fig. 2 Recta de regresión lineal del TRT Corrientes con criterio de tiempo 9 a 72 hs.

La ecuación de estas rectas no permite modelar con precisión la variación en el tiempo, por el bajo coeficiente de determinación, mucho menos comparar ambas rectas. En estas curvas se incluyen los errores de medición y aleatorios.

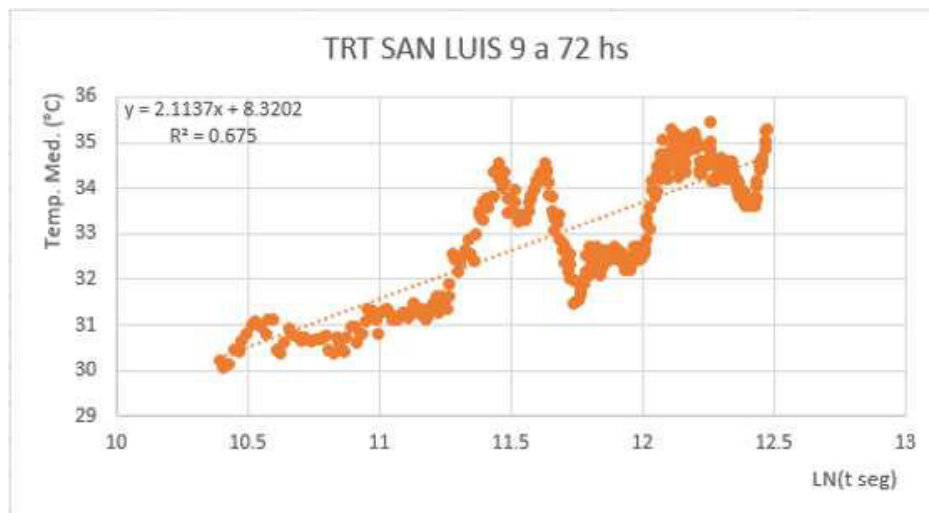


Fig. 3 Recta de regresión lineal del TRT San Luis con criterio de tiempo 9 a 72 hs.

Resultados del análisis multivariado

El resultado de aplicar el ACP se presenta en la (Fig. 4 y Fig. 5). Las 2 primeras componentes principales explican el 100 %, minimizando los errores de medición y aleatorios.

¹⁴Ligges, U. & M'achler, M. (2003), 'Scatterplot3d—an R package for Visualizing Multivariate Data', J. of Stat. Soft. 8(11), 1–20

En la Figura 4, se observa el plano de las variables, es decir la relación lineal entre las tres series de tiempo. Las tres están altamente correlacionadas (coeficiente de determinación 1, igual al coseno del ángulo entre sí), ahora se puede apreciar que varían de igual manera en el tiempo.

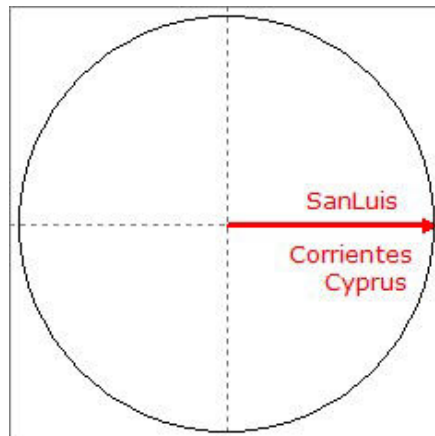


Fig. 4 Plano de las variables. Las tres series están totalmente correlacionadas

En la (Fig. 5), se observa el plano de las observaciones, es decir los valores distribuidos según las tres series, en los 20 tiempos y se colorea los grupos los 7 intervalos de las diferencias en San Luis.

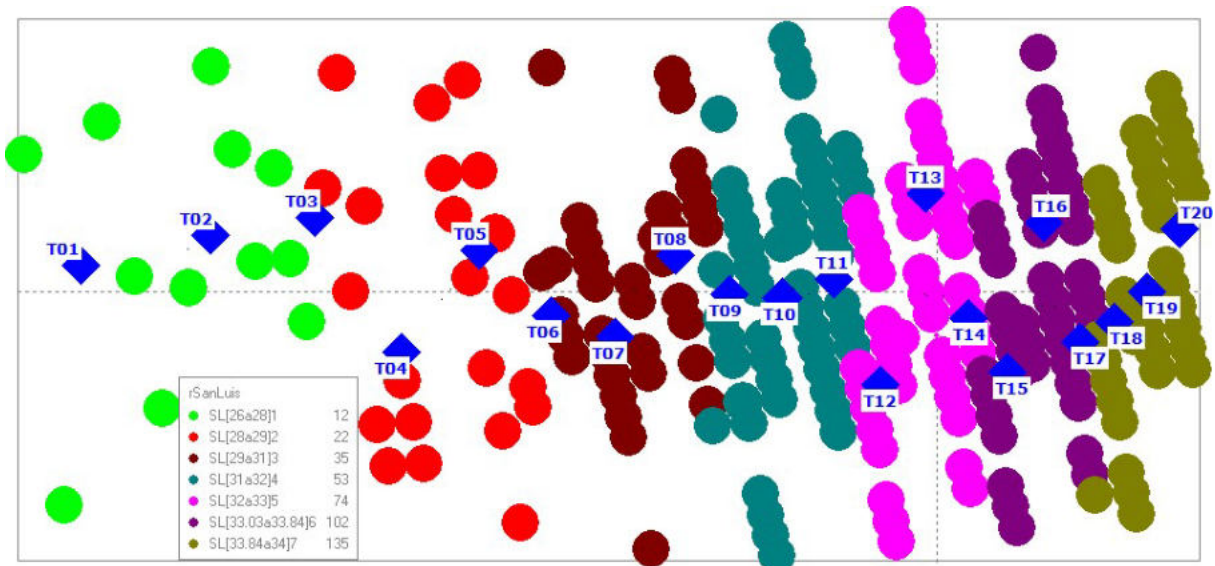
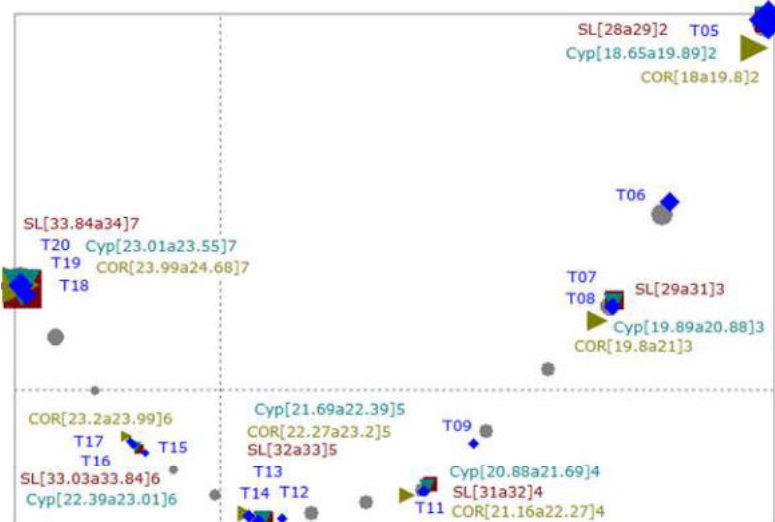


Fig. 5. Plano de las observaciones formado por las componentes 1 y 2. (Tx, tiempo x)

El análisis de los 7 intervalos discretizados para cada serie, nos permitió observar que, si bien las tres variables están totalmente correlacionadas, lo cual significan que tienen la misma variación en el tiempo, presentan distintas magnitudes en las diferencias de temperaturas, lo cual se observa en los resultados del DESCO y el ACM (Fig. 6). Dado que el primer intervalo son pocas observaciones, no contribuyen en el análisis por lo cual se excluye del mismo. El tamaño de representación en el plano es consistente con los niveles de significación estadística, a mayor tamaño, mayor valor de test.

Rangos Característicos	LN Hora Decimal	Valor-Test	
Cyp[18.65a19.89]2	1.71	-9.20	***
SL[28a29]2	1.71	-9.20	***
COR[18a19.8]2	1.74	-9.81	***
SL[29a31]3	2.24	-11.40	***
Cyp[19.89a20.88]3	2.24	-11.50	***
COR[19.8a21]3	2.29	-11.36	***
SL[31a32]4	2.81	-7.71	***
Cyp[20.88a21.69]4	2.82	-7.59	***
COR[21.16a22.27]4	2.88	-7.23	***
SL[32a33]5	3.30	-2.45	**
Cyp[21.69a22.39]5	3.31	-2.35	**
COR[22.27a23.2]5	3.37	-1.59	NS
LNHoraDecimal	3.48		***
SL[33.03a33.84]6	3.73	4.43	***
Cyp[22.39a23.01]6	3.74	4.63	***
COR[23.2a23.99]6	3.78	5.11	***
SL[33.84a34]7	4.12	13.51	***
Cyp[23.01a23.55]7	4.12	13.38	***
COR[23.99a24.68]7	4.13	13.06	***



- *** Para valores de Test superiores a 2.58, $p < 0.01$, muy significativo (99% certeza).
- ** Para valores de Test superiores a 1.96, $p < 0.05$, significativo (95% certeza).
- * Para valores de Test superiores a 1.65, $p < 0.10$, poco significativo (90% certeza).
- NS Para valores de Test inferiores a 1.65, $p > 0.10$, no significativo.

Figura 6. Plano Factorial del AFCM, Validación estadística. Sin considerar el primer intervalo

En el intervalo 5 de las tres series no se observa alta significación, esto podría ser debido a pocas observaciones estudiadas en este intervalo.

CONCLUSIONES

En primer lugar, se concluye a partir de los resultados, que el método utilizado para la comparación de series de tiempo es apropiado, dado que al transformar en variables latentes las originales, la interferencia de los errores de medición y aleatorios no se producen. Ahora las ecuaciones de las variables latentes obtenidas del análisis factorial tienen un coeficiente de terminación muy alto. Entonces se puede validar también el método factorial, para la comparación de la pendiente de dos rectas de calibrado, basado en el análisis de la homogeneidad de las varianzas.

Utilizando este ACP, se puede afirmar que los datos obtenidos en el intercambiador particular propuesto en la Universidad de San Luis (UNSL), con respecto al intercambiador de la Universidad Nacional de Corrientes y al de la Universidad de Cyprus, son válidos en más del 99% de certeza.

Por lo tanto, se puede afirmar que el método usado en el intercambiador horizontal particular de San Luis aplicando el modelo de la fuente de línea infinita de Kelvin (LSM), determinando el coeficiente de conductividad efectivo de subsuelo λ_{eff} , y la resistencia térmica de subsuelo R_b , por el método de la pendiente es válido, y describe el modelo planteado.

El ACM nos permitió comparar las series, por intervalos, independiente de las magnitudes de cada TRT. Con la aplicación de los métodos factoriales y de la comparación estadística de los TRT, se abren un sinnúmero de caminos respecto a la determinación de los parámetros de subsuelo ya que en una excavación realizada con otros objetivos se puede colocar un intercambiador de calor geotérmico de varios diseños amoldado al lugar y con un TRT de realización rápida y económica se determina las características del subsuelo.

A partir de aquí se realiza la comparación estadística y un análisis factorial que validaría dichos parámetros. Si esta validación es correcta, se diseña el emprendimiento geotérmico sobre valores reales, lo cual aseguraría un diseño muy eficiente, ya que no se estaría a la deriva de un sobre o sub dimensionamiento del mismo, traducido en dinero y sobre todo en prestigio de la geotermia de baja entalpía.

REFERENCIAS

- Benzecri. J. P. (1992). *Correspondence Analysis Handbook*. CRC Press.
- Benzecri. J. P. (1992). *Correspondence Analysis Handbook*. CRC Press.
- Bertaut M. B. y Marsal, J. V. (1999). *Manual de introducción a los métodos factoriales y clasificación con SPAD*. Universidad Politécnica de Catalunya y Universidad Autónoma de Barcelona
- Busso A. y 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.
- Busso A., Reuss M., Müller J. y 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, N° 1).
- Carslaw H. S. y Jaeger J. C., (1964). *Conduction of Heat in Solids*. Oxford science publications.
- Eklöf C. y Gehlin S. (1996). *A Mobile Equipment for Thermal Response Test*. Master of science thesis 1996:198E. Luleå University of Technology. Sweden.
- Espinoza, J. (2012). *Partición óptima: el algoritmo de Fisher*. *Revista de Matemática: Teoría y Aplicaciones*. 4. 77. 10.15517/rmta.v4i1.142.
- Fisher W.D. (1958). *On grouping for maximum homogeneity*. *Journal American Statistics Association*, N° 53, pag.789-798.
- Gehlin S. y Nordell, B. (1997). “*Thermal Response Test - a Mobile Equipment for Determining Thermal Resistance of Borehole*.” *Proc. Megastock*. (97 Sapporo, pp. 103-108).
- Lê S.Josse, J. y Husson, F. (2008), *FactoMineR: an R package for multivariate analysis*, *J. of Stat. Soft.* 25(1), 1–18.
- Langrand C. y Pinzón L.M. (2009). *Análisis de datos. Métodos y ejemplos*. Escuela Colombiana de Ingeniería Julio Garabito, Bogotá, Colombia. ISBN: 9789588060903. 388 páginas.

THE ROLE OF FACTORIAL ANALYSIS IN THE EVALUATION OF SOIL THERMAL RESPONSE TESTS - VERTICAL SERPENTINE / HORIZONTAL CYPRUS/CORRIENTES/ SAN LUIS.

ABSTRACT: In order to validate the established parameters proposed for a Thermal Responses test, a Principal Component Factor Analysis (PCA) has been applied, between three in situ Thermal Response Tests (TRT), substantially different in design, place and time, of Vertical vs. Horizontal coils. Likewise, each series was studied in 7 ranges, through a Multiple Correspondence Factor Analysis (MCA). The applied ACP, allowed to demonstrate that the correspondence between the values obtained by the three TRT studied is greater than 99%, for which statistically, the method used to obtain the effective heat transfer coefficient of the soil is valid for the proposed horizontal exchanger. . In turn, the nonlinear analysis of the AFCM showed that the three tests have the same variation over time despite the different magnitudes in the observed temperature differences.

Keywords: statistics, principal component analysis, correspondence analysis, thermal response test.