Geothermal Energy

RESEARCH Open Access

GaS_GeoT: A computer program for an effective use of newly improved gas geothermometers in predicting reliable geothermal reservoir temperatures

A. Acevedo-Anicasio[1], E. Santoyo[2*], D. Pérez-Zárate[3], Kailasa Pandarinath[2], M. Guevara[2] and L. Díaz-González[4]

*Correspondence: esg@ier.unam.mx
[2] Instituto de Energías Renovables, Universidad Nacional Autónoma de México, Priv. Xochicalco S/N, Temixco, 62580 Morelos, México
Full list of author information is available at the end of the article

Abstract

A geochemometric study based on a multi-criteria decision analysis was applied, for the first time, for the optimal evaluation and selection of artificial neural networks, and the prediction of geothermal reservoir temperatures. Eight new gas geothermometers ($GasG_1$ to $GasG_8$) were derived from this study. For an effective and practical application of these geothermometers, a new computer program GaS_GeoT was developed. The prediction efficiency of the new geothermometers was compared with temperature estimates inferred from twenty-five existing geothermometers using gas-phase compositions of fluids from liquid- (LIQDR) and vapour-dominated (VAPDR) reservoirs. After applying evaluation statistical metrics (DIFF%, RMSE, MAE, MAPE, and the Theil's U test) to the temperature estimates obtained by using all the geothermometers, the following inferences were accomplished: (1) the new eight gas geothermometers ($GasG_1$ to $GasG_8$) provided reliable and systematic temperature estimates with performance wise occupying the first eight positions for LIQDR; (2) the $GasG_3$ and $GasG_1$ geothermometers exhibited consistency as the best predictor models by occupying the first two positions over all the geothermometers for VAPDR; (3) the $GasG_3$ geothermometer exhibited a wider applicability, and a better prediction efficiency over all geothermometers in terms of a large number of samples used (up to 96% and 85% for LIQDR and VAPDR, respectively), and showed the smallest differences between predicted and measured temperatures in VAPDR and LIQDR; and lastly (4) for the VAPDR, the existing geothermometers ND84c, A98c, and ND98b sometimes showed a better prediction than some of the new gas geothermometers, except for $GasG_3$ and $GasG_1$. These results indicate that the new gas geothermometers may have the potential to become one of the most preferred tools for the estimation of the reservoir temperatures in geothermal systems.

Keywords: Geothermal energy, Renewable energy, Geochemometrics, Fluid geochemistry, Artificial intelligence

Introduction

Geothermal energy has emerged as a clean alternative source of renewable energy for electric power generation and other direct uses (Wu and Li 2020). The enormous amount of energy stored in geothermal systems makes them an important renewable and sustainable energy source (Nieva et al. 2018; Gutiérrez-Negrín et al. 2020). Among the geothermal systems today exploited, hydrothermal systems stand out for their storage capability of hot fluids, which are used for the generation of electricity (Thien et al. 2015). The chemical composition of liquid and steam (gas) phases of geothermal fluids provides useful information on hydrogeological processes, thermal and recharge conditions of reservoirs, and underground flow patterns (Nicholson 1993). Within these applications, the reliable estimation of reservoir temperatures is a crucial task to evaluate the energy potential of geothermal resources (Gutiérrez-Negrín 2019). To carry out this task, several chemical geothermometers have been proposed for the prediction of deep equilibrium temperatures in geothermal systems (Guo et al. 2017). Chemical geothermometers are low-cost tools used for predicting reservoir temperatures in the early exploration and exploitation stages (Yan-guang et al. 2017). Solute geothermometers are mostly recommended for the prediction of reservoir temperatures in liquid-dominated reservoirs, LIQDR (Verma et al. 2008), whereas gas geothermometers are predominantly suggested for the calculation of reservoir temperatures in vapour-dominated (VAPDR) reservoirs (García-López et al. 2014).

The gas chemistry of geothermal fluids has been applied not only for the estimation of geothermal reservoir temperatures but also to elucidate water–rock interaction processes (Pang, 2001). The direct escape from deep magmatic sources, and the occurrence of gas–gas and gas–mineral reactions may explain the consumption and production of gases in these systems (Minissale et al. 1997).

The equilibria attainment among gas species is considered as a fundamental thermodynamic basis for gas geothermometry. This physicochemical process has been widely studied for the prediction of reservoir temperatures in LIQDR and VAPDR using the composition of vapour (or gas) samples collected from wells and fumaroles. To our knowledge, Ellis (1957) was the first geochemist to point out that gases in magmatic steam might be used to estimate reservoir temperatures. The first gas geothermometer is attributed to Tonani (1973) who found some relationships among geothermometric models, and the composition of geothermal gases.

From these studies, numerous gas geothermometers have been proposed based upon the analysis of gas-phase compositions, and the thermodynamic study of gas–gas and gas–mineral equilibria reactions. A comprehensive review on gas geothermometers has not been published yet in indexed journals, although some efforts have been conducted in compiling some of the geothermometers most commonly used (e.g. Henley et al. 1985; Nicholson 1993; Powell 2000; Powell and Cumming 2010). In this work, an update compilation of the most commonly used gas geothermometers is reported in Table 1. Most of these geothermometers may be roughly grouped as follows:

- Simple equations calibrated with databases of gas-phase compositions of fluids collected from geothermal wells;

Table 1 Updated compilation of gas geothermometers reported in the literature for determining geothermal reservoir temperatures

No	Gas geothermometer	Acronym	Gas-Mineral equilibria reactions	Gas concentration units	Geothermal reservoir-type[a]	Calculation method	Temperature interval (°C)	References
1	CO_2-H_2S-CH_4-H_2	DP80	$CaSO_4 + FeS_2 + 3H_2O + CO_2 \leftrightarrow CaCO_3 + 1/3Fe_3O_4 + 3H_2S + 7/3O_2$ $C + CO_2 + H_2 \leftrightarrow 2CH_4 + 2H_2O$	% volume	LIQDR and VAPDR	Analytical	145–300	D'Amore and Panichi (1980)
2	H_2O-CO_2-CH_4-H_2	G80a	$CO_2 + 4H_2 \leftrightarrow CH_4 + 2H_2O$ $2NH_3 \leftrightarrow N_2 + 3H_2$	mmol/mol	LIQDR	Numerical[b]	100–340	Giggenbach (1980)
3	NH_3-H_2-N_2	G80b	$CO_2 + 4H_2 \leftrightarrow CH_4 + 2H_2O$ $2NH_3 \leftrightarrow N_2 + 3H_2$	mmol/mol	LIQDR	Numerical[b]	100–340	Giggenbach (1980)
4	CO_2-CH_4-H_2	ND84a	$CO_2 + 4H_2 \leftrightarrow CH_4 + 2H_2O$	Mole fraction	LIQDR	Analytical	150–350	Nehring and D'Amore (1984)
5	CO_2-H_2	ND84b	$H_2 + 1/2O_2 \leftrightarrow H_2O$ $C + O_2 \leftrightarrow CO_2$	Mole fraction	LIQDR	Analytical	150–350	Nehring and D'Amore (1984)
6	CO_2-H_2S	ND84c	$3FeS_2 + 2H_2 + 4H_2O \leftrightarrow Fe_3O4 + 6H_2S$ $C + O_2 \leftrightarrow CO_2$	Mole fraction	LIQDR	Analytical	150–350	Nehring and D'Amore (1984)
7	CO_2	AG85a	$2clinozoisite + 2calcite + 3quartz + 2H_2O \leftrightarrow 3prehnite + 2CO_2$	mmol/kg	LIQDR and VAPDR	Analytical	200–300	Arnórsson and Gunnlaugsson (1985)
8	CO_2-H_2	AG85b	$2clinozoisite + 2calcite + 3quartz + 2H_2O \leftrightarrow 3prehnite + 2CO_2$ $4pyrrhotite + 2prehnite + 2H_2O \leftrightarrow 2epidote + 2pyrite + 3H_2$	mmol/kg	LIQDR and VAPDR	Analytical	200–300	Arnórsson and Gunnlaugsson (1985)
9	CO_2-H_2	AG85c	$2clinozoisite + 2calcite + 3quartz + 2H_2O \leftrightarrow 3prehnite + 2CO_2$ $4pyrrhotite + 2prehnite + 2H_2O \leftrightarrow 2epidote + 2pyrite + 3H_2$	mmol/kg	LIQDR and VAPDR	Analytical	200–300	Arnórsson and Gunnlaugsson (1985)
10	H_2	AG85d	$4pyrrhotite + 2prehnite + 2H_2O \leftrightarrow 2epidote + 2pyrite + 3H_2$	mmol/kg	LIQDR and VAPDR	Analytical	200–300	Arnórsson and Gunnlaugsson (1985)
11	H_2	AG85e	$4pyrrhotite + 2prehnite + 2H_2O \leftrightarrow 2epidote + 2pyrite + 3H_2$	mmol/kg	LIQDR and VAPDR	Analytical	200–300	Arnórsson and Gunnlaugsson (1985)
12	H_2S	AG85f	$pyrite + pyrrhotite + 2prehnite + 2H_2O \leftrightarrow 2epidote + 3H_2S$	mmol/kg	LIQDR and VAPDR	Analytical	200–300	Arnórsson and Gunnlaugsson (1985)
13	H_2S	AG85g	$pyrite + pyrrhotite + 2prehnite + 2H_2O \leftrightarrow 2epidote + 3H_2S$	mmol/kg	LIQDR and VAPDR	Analytical	200–300	Arnórsson and Gunnlaugsson (1985)
14	H_2S-H_2	AG85h	$pyrite + pyrrhotite + 2prehnite + 2H_2O \leftrightarrow 2epidote + 3H_2S$ $4pyrrhotite + 2prehnite + 2H_2O \leftrightarrow 2epidote + 2pyrite + 3H_2$	mmol/kg	LIQDR and VAPDR	Analytical	200–300	Arnórsson and Gunnlaugsson (1985)
15	CO_2-CH_4-CO	B85	$CO_2 + H_2 \leftrightarrow CO + H_2O$ $CH_4 + 3CO_2 \leftrightarrow 4CO + 2H_2O$	Mole fraction	LIQDR and VAPDR	Numerical[b]	n.r	Bertrami et al. (1985)
16	CO_2-N_2	A87a	n.a	mmol/kg	LIQDR	Analytical	200–300	Arnórsson (1987)
17	CO_2-N_2	A87b	n.a	mmol/kg	LIQDR	Analytical	200–300	Arnórsson (1987)

Table 1 (continued)

No	Gas geothermometer	Acronym	Gas-Mineral equilibria reactions	Gas concentration units	Geothermal reservoir-type[a]	Calculation method	Temperature interval (°C)	References
18	CO_2-H_2S-CH_4-H_2-CO	SD89	$CO_2 + 4H_2 \leftrightarrow CH_4 + 2H_2O$ $H_2 + 1/2O_2 \leftrightarrow H_2O$ $H_2 + 1/2S_2 \leftrightarrow H_2S$ $H_2 + CO_2 \leftrightarrow CO + H_2O$ $1/3Fe_3O_4 + S_2 \leftrightarrow FeS_2 + 2/3O_2$	% mol	VAPDR	Numerical[b]	140–370	Saracco and D'Amore (1989)
19	CO_2-CH_4	G91a	$3CO_2 + CH_4 \leftrightarrow 4CO + 2H_2O$	mmol/mol	LIQDR	Analytical	100–350	Giggenbach (1991)
20	H_2-Ar	G91b	$2Fe_3O_4 + H_2O \leftrightarrow 3Fe_2O_3 + H_2 + O_2$	mmol/mol	LIQDR	Analytical[b]	100–350	Giggenbach (1991)
21	CO_2-Ar	G91c	n.a	mmol/mol	LIQDR	Analytical[b]	100–350	Giggenbach (1991)
22	H_2-Ar	GG92a	$FeS_2 + FeO + 2H_2O \leftrightarrow Fe_2O_3 + 2H_2S$	mmol/mol	LIQDR	Analytical[b]	100–350	Giggenbach and Glover (1992)
23	CO_2-CO	GG92b	$FeS_2 + FeO + 2H_2O \leftrightarrow Fe_2O_3 + 2H_2S$	mmol/mol	LIQDR	Analytical[b]	100–350	Giggenbach and Glover (1992)
24	CO_2-H_2	K95a	Empirical geothermometer. Regression approach	% volume	LIQDR	Analytical	n.r	Koga et al. (1995)
25	CH_4-H_2	K95b	Empirical geothermometer. Regression approach	% volume	LIQDR	Analytical	n.r	Koga et al. (1995)
26	CO_2-H_2S	S96a	Empirical geothermometer. Multidimensional approach	% volume	LIQDR	Analytical	n.r	Supranto et al. (1996)
27	CO_2-H_2S-CH_4	S96b	Empirical geothermometer. Multidimensional approach	% volume	LIQDR	Analytical	n.r	Supranto et al. (1996)
28	CO_2	A98a	2clinozoisite + 2calcite + 3quartz + $2H_2O \leftrightarrow$ 3prehnite + $2CO_2$	mmol/kg	LIQDR	Analytical	> 230	Arnórsson et al. (1998)
29	CO_2	A98b	2clinozoisite + 2calcite + 3quartz + $2H_2O \leftrightarrow$ 3prehnite + $2CO_2$	mmol/kg	LIQDR	Analytical	> 230	Arnórsson et al. (1998)
30	H_2S	A98c	pyrite + pyrrhotite + 2prehnite + $2H_2O \leftrightarrow$ 2epidote + $3H_2S$	mmol/kg	LIQDR	Analytical	> 150	Arnórsson et al. (1998)
31	H_2	A98d	4pyrrhotite + 2prehnite + $2H_2O \leftrightarrow$ 2epidote + 2pyrite + $3H_2$	mmol/kg	LIQDR	Analytical	> 150	Arnórsson et al. (1998)
32	CO_2-N_2	A98e	n.r	mmol/kg	LIQDR	Analytical	n.r	Arnórsson et al. (1998)
33	H_2S-Ar	A98f	n.r	mmol/kg	LIQDR	Analytical[b]	n.r	Arnórsson et al. (1998)
34	H_2-Ar	A98g	n.r	mmol/kg	LIQDR	Analytical[b]	n.r	Arnórsson et al. (1998)
35	H_2S	B06	$2Fe_3O_4 + 12H_2S + CO_2 \leftrightarrow 6FeS_2 + 10H_2O + CH_4$	% mol	LIQDR	Analytical	200–300	Blamey (2006)
36	CO_2-H_2-Ar	PC10a	$2Fe_3O_4 + H_2O \leftrightarrow 3Fe_2O_3 + H_2 + O_2$ $2H_2O + C \leftrightarrow 2H_2 + CO_2$	% mol	LIQDR and VAPDR	Grid[b]	100–350	Powell and Cumming (2010)
37	CO_2-CO-CH_4	PC10b	$CacO_3$ + K-mica $\leftrightarrow$ $CaAl_2$-silciate + K-feldspar + CO_2 $CO_2 + H_2O \leftrightarrow CO + H_2O$ $3CO_2 + CH_4 \leftrightarrow 4CO + 2H_2O$	% mol	Liquid—and VAPDR	Grid[b]	100–350	Powell and Cumming (2010)

Table 1 (continued)

No	Gas geothermometer	Acronym	Gas-Mineral equilibria reactions	Gas concentration units	Geothermal reservoir-type[a]	Calculation method	Temperature interval (°C)	References
38	CO_2-CH_4-H_2-H_2S	PC10c	$CO_2 + 4H_2 \leftrightarrow CH_4 + 2H_2O$ $3FeS_2 + 2H_2 + 4H_2O \leftrightarrow Fe_3O_4 + 6H_2S$ $FeS_2 + H_2 \leftrightarrow FeS + H_2S$	% mol	LIQDR and VAPDR	Grid[b]	100–350	Powell and Cumming (2010)
39	CO_2-CH_4-H_2	PC10d	$CO_2 + 4H_2 \leftrightarrow CH_4 + 2H_2O$ $CaCO_3 + K\text{-mica} \leftrightarrow CaAl_2\text{-silicate} + K\text{-feldspar} + CO_2$	% mol	LIQDR and VAPDR	Grid[b]	100–350	Powell and Cumming (2010)
40	CO_2-CH_4-H_2-H_2S	PC10e	$CO_2 + 4H_2 \leftrightarrow CH_4 + 2H_2O$ $3FeS_2 + 2H_2 + 4H_2O \leftrightarrow Fe_3O_4 + 6H_2S$ $FeS_2 + H_2 \leftrightarrow FeS + H_2S$	% mol	LIQDR and VAPDR	Grid[b]	200–350	Powell and Cumming (2010)
41	CO_2-H_2S-H_2	B16a	$CO_2 + 4H_2 \leftrightarrow CH_4 + 2H_2O$ $H_2 + 2H_2O + 3/2FeS_2 \leftrightarrow 3H_2S + 1/2Fe_3O_4$	% volume	LIQDR	Numerical[b]	125–350	Barragán et al. (2016)
42	CO_2-H_2S-H_2	B16b	$CO_2 + 4H_2 \leftrightarrow CH_4 + 2H_2O$ $5/4H_2 + 3/4Fe_2O_3 + 3/2FeS_2 + 7/4H_2O \leftrightarrow 3H_2S + Fe_3O_4$	% volume	LIQDR	Numerical[b]	150–350	Barragán et al. (2016)
43	CO_2-H_2S-H_2	B16c	$CO_2 + 4H_2 \leftrightarrow CH_4 + 2H_2O$ $H_2 + FeS_2 \leftrightarrow H_2S + FeS$	% volume	LIQDR	Numerical[b]	100–350	Barragán et al. (2016)

[a] Type of geothermal reservoir used to calibrate the gas geothermometers

[b] Gas geothermometers that were not considered in this work

n.r. = not reported; n.a. = not available. Analytical method refers to a gas geothermometer that provides a direct mathematical function for the calculation of the reservoir temperatures, i.e., *BHT = f(gas concentration)*; whereas Numerical method refers to a more complex function that correlate multiple variables (i.e., gas concentration, pressure, the fraction of water, temperature, among others); Grid geothermometric method refers to a gas geothermometer that provides a complex grid-numerical algorithm for the estimation of temperatures and other key parameters of the reservoir (i.e., steam fraction, distribution coefficients, and the steam/gas ratio)

Table 2 A listing of software reported in the technical literature for computing geothermal reservoir temperatures

No	Open-source computer programs	Computational platform/ language	Number of gas geothermometers	Working temperature interval (°C)	Gas concentration units	Geothermal reservoir-type[a]	References
1	CO2B	Fortran	1	140–370	% mol	VAPDR	Saracco and D'Amore (1989)
2	Spreadsheet application	Excel	5	100–350	% mol	LIQDR and VAPDR	Powell and Cumming (2010)
3	GasGeo	Visual Basic	21	100–350	mmol/kg; mole fraction; mmol/mol; %mol; %vol	LIQDR and VAPDR	Pandarinath et al. (2011)
4	EQUILGAS	Visual Fortran	3	100–350	% vol	LIQDR	Barragán et al. (2016)
5	GasGeoPlus	Fortran	21	100–350	mmol/kg; mole fraction; mmol/mol; %mol; %vol	LIQDR and VAPDR	Pérez-Zárate et al. (2019)
6	SYS_ GASCHEM	Perl	21	100–350	mmol/kg; mole fraction; mmol/mol; %mol; %vol	LIQDR and VAPDR	García-Manduano (2019)
7	GaS_GeoT	Java	8	170–374	mmol/mol	LIQDR and VAPDR	This work

[a] Type of geothermal reservoir used to calibrate the gas geothermometer

- Complex grid-numerical geothermometers that calculate reservoir temperatures and other key parameters (e.g. the steam excess, the distribution coefficients of gases between liquid and steam, etc.); and
- Geothermometric equations derived from fluid–rock interaction experiments.

Although these geothermometers have been proposed for the prediction of geothermal temperatures, their generalized application has been limited by the following issues:

1. The significant statistical differences among the temperature estimates predicted by a group of gas geothermometers;
2. The scarcity of the gas-phase compositions at lower temperature levels (between 90 and 150 °C) which has hindered the proposal of new improved equations;
3. The limited intervals of the gas composition and temperature for the application of the geothermometric equations;
4. The scarcity of geochemometric studies for the evaluation of the temperature estimates and their uncertainties; and
5. The lack of practical computer programs to estimate reservoir temperatures by using a wide variety of concentration units, and the complicated calculations involved in some equations.

To address the referred issues (1–3), geochemical databases with representative compositions of gas and/or steam phases from LIQDR and VAPDR are required.

Considering the complex nature of the gas–mineral equilibria processes, it is also necessary to explore new regression tools based on artificial intelligence techniques for calibrating multivariate gas geothermometers, and so, to predict reservoir temperatures with a better accuracy. Among these techniques, the artificial neural networks (ANN) have been used in solving multivariate problems in Earth sciences (Poulton, 2001). For example, in geothermal studies, the use of ANN has been applied for (i) the development of Na–K geothermometers (Díaz-González et al. 2008; and Serpen et al. 2009); (ii) the prediction of mass and heat transport in geothermal wells (Bassam et al. 2010; Álvarez del Castillo et al. 2012; Porkhial et al. 2015); and (iii) the optimization of geothermal power plants (Arslan and Yetik 2011), among others.

One of the latest ANN applications reported for geothermal studies was conducted by Pérez-Zárate et al. (2019) who performed a preliminary study to evaluate ANNs for the prediction of geothermal temperatures using gas-phase compositions from which the present research work has been comprehensively completed.

To evaluate the uncertainties of gas geothermometers (the issue referred as 4), a limited number of geochemometric studies have been reported (e.g. D'Amore and Panichi 1980; Arnórsson et al. 2006; García-López et al. 2014). These studies stated that the prediction efficiency of gas geothermometers is affected by several error sources, such as gas sampling errors, analytical errors, coefficient errors, and the total propagated errors, associated with the calculation of temperatures (e.g. Kacandes and Grandstaff 1989; García-López et al. 2014).

The availability of new computer programs to calculate reservoir temperatures (the issue referred as 5) still constitutes a current necessity for the geothermometric studies. Computer programs to apply solute geothermometers in estimating temperatures are widely available for the study of geothermal systems (e.g. Verma et al. 2008; Spycher et al. 2016), whereas for the gas geothermometry, these programs are rarely shared in the literature. A short listing of open-source programs commonly used in gas geothermometry is included in Table 2. To fulfil the constraints (1–5), the development of new improved gas geothermometers and computer programs are still claimed by the geothermal industry.

To address a reliable prediction of geothermal reservoir temperatures, new improved gas geothermometers and a practical computer program called GaS_GeoT have been developed in this work. GaS_GeoT was calibrated for the reliable prediction of reservoir temperatures in LIQDR and VAPDR for a temperature interval between 170 and 374 °C.

To evaluate the prediction efficiency of the new gas geothermometers in geothermal wells, an updated worldwide geochemical database containing gas-phase compositions of fluids were compiled. The prediction efficiency of these geothermometers was compared against twenty-five existing gas geothermometers (listed in Table 3). Details of this geochemometric study are also outlined in this work.

Work methodology

A comprehensive computational methodology was developed to achieve the following research objectives (Fig. 1a, b): (i) to evaluate ANNs based on a Multi-Criteria Decision Analysis (MCDA) for selecting optimal prediction models that enable new improved gas

Table 3 Existing gas geothermometers reported in the geothermal literature which were used in the present work

No.	Acronym	Temperature function (°C)	Gas concentration units	References
1	DP80	If $\%CO_2 \geq 75$ $\frac{24775}{2log\left(\frac{CH_4}{CO_2}\right)-6log\left(\frac{H_2}{CO_2}\right)-3log\left(\frac{H_2S}{CO_2}\right)+36.05} - 273.15$ If $\%CO_2 < 75$ $\frac{24775}{2log\left(\frac{CH_4}{CO_2}\right)-6log\left(\frac{H_2}{CO_2}\right)-3log\left(\frac{H_2S}{CO_2}\right)-7log(0.1)+36.05} - 273.15$	% volume	D'Amore and Panichi (1980)
2	ND84a	$\frac{\sqrt{(5.5697-R)+(R-5.5697)^2+1056.34368}}{0.03936} - 273.15$ $R = log\left(X_{CH_4}\right) - 4log\left(X_{H_2}\right) - log\left(X_{CO_2}\right)$	Mole fraction	Nehring and D'Amore (1984)
3	ND84b	$\frac{\sqrt{(S+1.89499)+(S+1.89499)^2+128.90864}}{0.01432} - 273.15$ $S = log\left(X_{H_2}\right) + \frac{1}{2}log\left(X_{CO_2}\right)$	Mole fraction	Nehring and D'Amore (1984)
4	ND84c	$5.78356V^2 + 110.212V + 662.814$ $V = log\left(X_{H_2S}\right) + \frac{1}{6}log\left(X_{CO_2}\right)$	Mole fraction	Nehring and D'Amore (1984)
5	AG85a	$-44.1 + 269.25\left[log(CO_2)\right] - 76.88\left[log(CO_2)\right]^2 + 9.52\left[log(CO_2)\right]^3$	mmol/kg	Arnórsson and Gunnlaugsson (1985)
6	AG85b	$341.7 - 28.57log\left(\frac{CO_2}{H_2}\right)$	mmol/kg	Arnórsson and Gunnlaugsson (1985)
7	AG85c	$311.7 - 66.72log\left(\frac{CO_2}{H_2}\right)$	mmol/kg	Arnórsson and Gunnlaugsson (1985)
8	AG85d	$277.2 + 20.99log(H_2)$	mmol/kg	Arnórsson and Gunnlaugsson (1985)
9	AG85e	$212.2 + 38.59log(H_2)$	mmol/kg	Arnórsson and Gunnlaugsson (1985)
10	AG85f	$246.7 + 44log(H_2S)$	mmol/kg	Arnórsson and Gunnlaugsson (1985)
11	AG85g	$173.2 + 65.04log(H_2S)$	mmol/kg	Arnórsson and Gunnlaugsson (1985)
12	AG85h	$304.1 - 39.48log\left(\frac{H_2S}{H_2}\right)$	mmol/kg	Arnórsson and Gunnlaugsson (1985)
13	A87a	$148.5 + 64.35log\left(\frac{CO_2}{N_2}\right) + 5.239\left[log\left(\frac{CO_2}{N_2}\right)\right]^2 - 1.832\left[log\left(\frac{CO_2}{N_2}\right)\right]^3$	mmol/kg	Arnórsson (1987)
14	A87b	$135.9 + 63.14log\left(\frac{CO_2}{N_2}\right) + 6.241\left[log\left(\frac{CO_2}{N_2}\right)\right]^2 - 1.813\left[log\left(\frac{CO_2}{N_2}\right)\right]^3$	mmol/kg	Arnórsson (1987)
15	G91a	$\frac{4625}{10.4+log\left(\frac{CH_4}{CO_2}\right)} - 273.15$	mmol/mol	Giggenbach (1991)

Table 3 (continued)

No.	Acronym	Temperature function (°C)	Gas concentration units	References
16	K95a	$352.45 - 53.36log\left(\frac{CO_2}{H_2}\right)$	% volume	Koga et al. (1995)
17	K95b	$228.95 - 24.06\left(\frac{CH_4}{H_2}\right)$	% volume	Koga et al. (1995)
18	S96a	$\frac{738.3091}{(\%gas^{0.0008})\left(\frac{\%CO_2}{\%H_2S}\right)^{0.1037}} - 273.15$	% volume	Supranto et al. (1996)
19	S96b	$\frac{1138.9501(\%Gas)^{0.0037}(\%H_2S)^{0.1076}}{(\%CO_2^{0.2019})(\%CH_4)^{0.0056}} - 273.15$	% volume	Supranto et al. (1996)
20	A98a	$121.8 + 72.012log(CO_2) + 11.068\left[log(CO_2)\right]^2 - 4.724\left[log(CO_2)\right]^3$	mmol/kg	Arnórsson et al. (1998)
21	A98b	$129.2 + 88.299\log(CO_2) + 21.946\left[\log(CO_2)\right]^2 - 10.103\left[log(CO_2)\right]^3$	mmol/kg	Arnórsson et al. (1998)
22	A98c	$177.6 + 66.152log(H_2S) + 4.811\left[log(H_2S)\right]^2$	mmol/kg	Arnórsson et al. (1998)
23	A98d	$227.1 + 56.168log(H_2) + 5.836\left[log(H_2)\right]^2 + 6.630\left[log(H_2)\right]^3$	mmol/kg	Arnórsson et al. (1998)
24	A98e	$173.2 + 48.751log\left(\frac{CO_2}{N_2}\right) + 7.599\left[log\left(\frac{CO_2}{N_2}\right)\right]^2 + 1.793\left[log\left(\frac{CO_2}{N_2}\right)\right]^3$	mmol/kg	Arnórsson et al. (1998)
25	B06	For H_2S concentrations ≥ 0.00175 $75.85log(H_2S) + 458.6$ For H_2S concentrations < 0.00175 $64.68log(H_2S) + 427.8$	%mol	Blamey (2006)

All the gas geothermometers use concentration units in wet-basis, except DP80, G91a, S96a, S96b, and B06 which use dry-basis

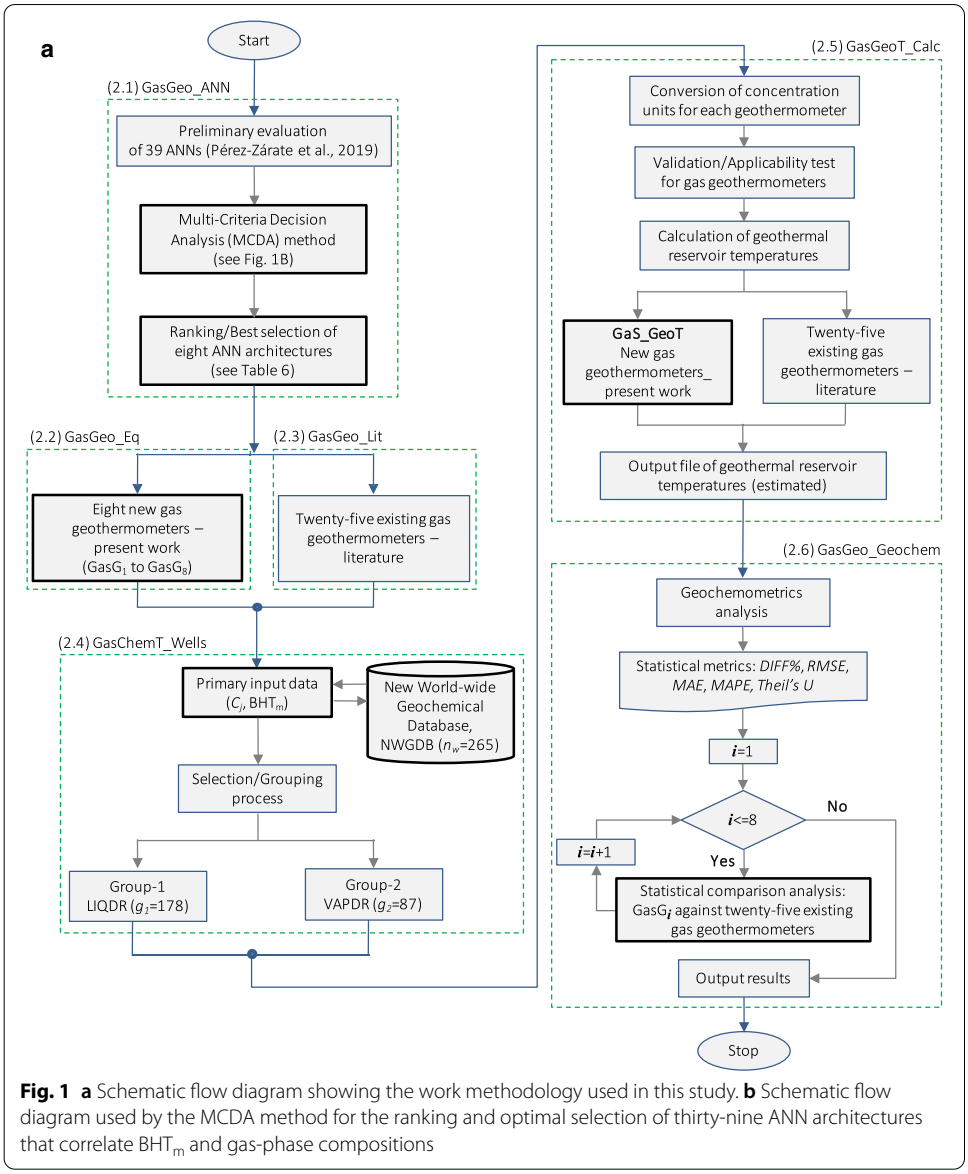

Fig. 1 **a** Schematic flow diagram showing the work methodology used in this study. **b** Schematic flow diagram used by the MCDA method for the ranking and optimal selection of thirty-nine ANN architectures that correlate BHT_m and gas-phase compositions

geothermometers to be developed; (ii) to create a new computer program for the effective use of the new gas geothermometers; (iii) to evaluate the prediction efficiency of the new gas geothermometers using the gas-phase compositions from well samples collected in LIQDR and VAPDR; and (iv) to compare the temperature estimates inferred from new gas geothermometers with those temperatures predicted by some existing geothermometers. In order to address these research goals, six computational modules were structured:

GasGeo_ANN: To carry out the optimal evaluation and selection of ANN architectures by using a novel application of the MCDA method;
GasGeo_Eq: To describe the new gas geothermometer equations developed and their applicability conditions;

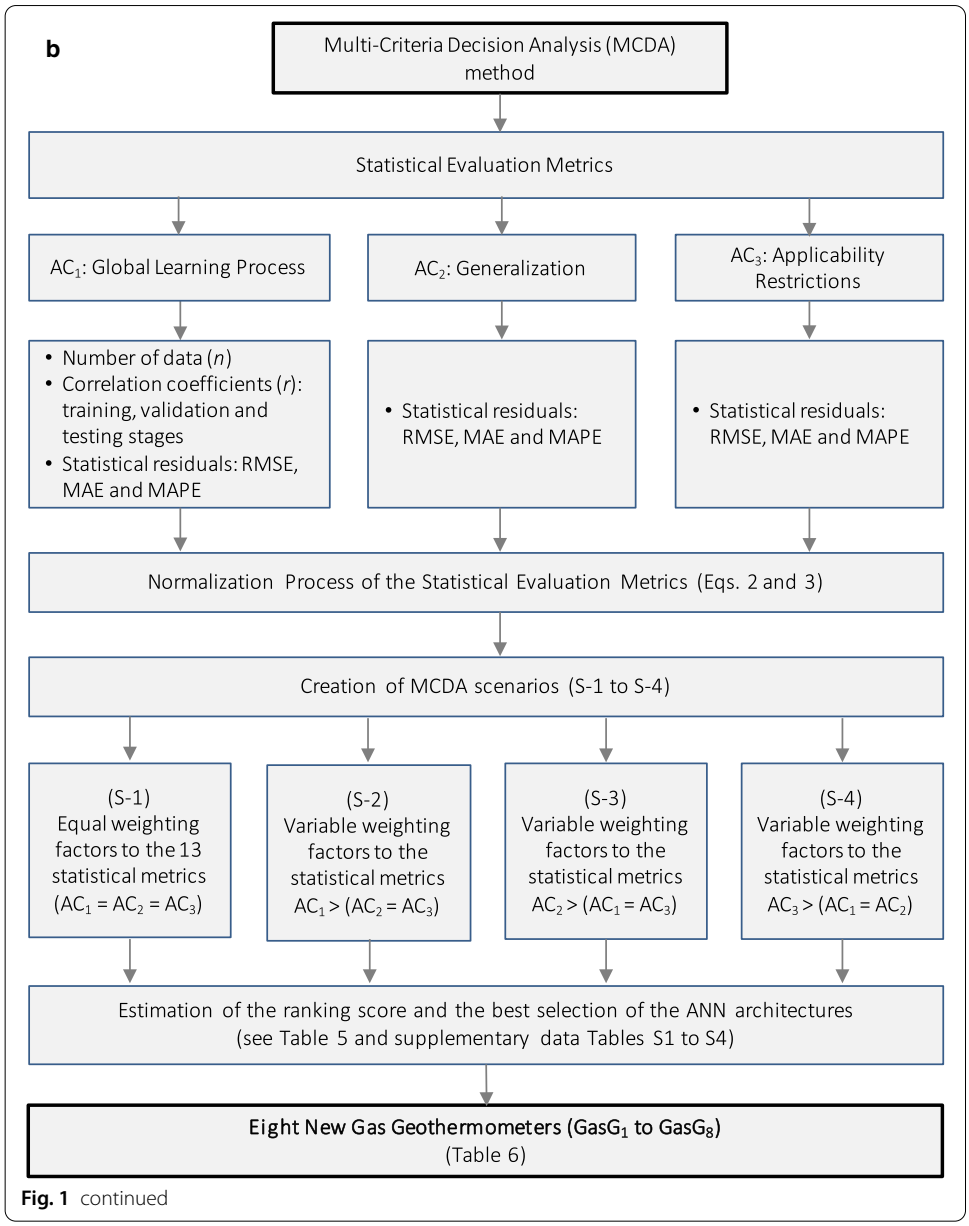

Fig. 1 continued

GasGeo_Lit: To select twenty-five existing gas geothermometers for carrying out statistical comparison analyses;

GasChemT_Wells: To compile a new Worldwide Geochemical Database (NWGDB) containing gas-phase compositions, and bottom-hole temperatures (BHT_m) measured in geothermal wells;

GasGeoT_Calc: To describe the calculation of geothermal reservoir temperatures by using the new and existing gas geothermometers, and the operation of the new computer program (GaS_GeoT); and

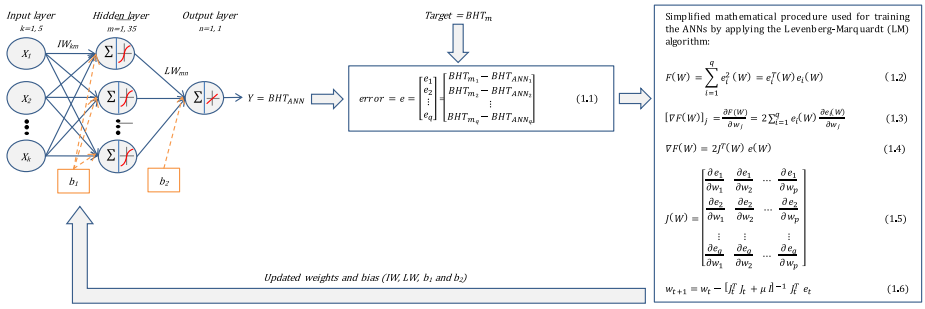

Fig. 2 Schematic diagram showing the conceptual model of a multilayer perceptron with three-layers, and a summary of the fundamental equations used by applying the Levenberg–Marquardt algorithm: Eq. (1.1) error vector; Eq. (1.2) cost function, where W is the vector of all the weights and biases and q is the total number of samples or patters; Eq. (1.3) where j is the jth element of the gradient; Eq. (1.4) the gradient in matrix form; Eq. (1.5) Jacobian matrix, where p is the total number of neurons; Eq. (1.6) iterative update of weights using the LM algorithm, where w_{t+1} is the next weight, w_t is the current weight, μ is the learning rate factor, and I is the identity matrix

GasGeo_Geochem: To perform a comprehensive geochemometric analysis among the temperature estimates (predicted from all the geothermometers) and the bottom-hole temperature measurements (BHT_m).

A summary of these computational modules is described as follows:

GasGeo_ANN

A previous evaluation study of ANNs to predict geothermal reservoir temperatures was preliminarily conducted by Pérez-Zárate et al. (2019), which constitutes the foundational basis of the present research work. 455 ANN architectures were originally trained to predict reservoir temperatures using three Worldwide Geochemical Subdatabases (WG_SubDB_1: $q_1 = 527$; WG_SubDB_2: $q_2 = 498$; and WG_SubDB_3: $q_3 = 97$) containing gas-phase (CO_2, H_2S, CH_4, and H_2) compositions of geothermal fluids as input variables. According to Fig. 2, a multilayer perceptron model was used for the design of these ANNs using the conventional matrix notation (Eq. 1) for the determination of the ANN output (or target):

$$Y = \mathrm{BHT_{ANN}} = LW \cdot f[IW \cdot X + b_1] + b_2, \tag{1}$$

where X is the input variable, and IW and b_1 represent the matrix and the vector of coefficients between the input and hidden layers. LW and b_2 are the layer weighting coefficients and the vector of coefficients between the hidden and output layers, and f corresponds to an activation function for the hidden layer.

A full description of the multilayer perceptron model is reported by Pérez-Zárate et al. (2019), and schematically summarized in Fig. 2 by listing the fundamental equations used for adjusting weighting and bias coefficients in the learning process of each ANN layer (referred as Eqs. 1.1–1.6). Such a perceptron model was solved by creating several Matlab numerical scripts that were previously reported (https://github.com/ANNGroup/GasG-Scripts-ANNs.git).

For the learning process of ANNs, the well-known Levenberg–Marquardt algorithm (LM), the hyperbolic tangent function, and the linear function were applied. All the ANNs were characterized by an input layer, one hidden layer, and an output layer. For the input and output layers, the data were normalized between -1 and 1, whereas for the hidden layer, the number of neurons varied from 1 to 35. Input data sets were randomly divided into training (80%), validation (10%), and testing (10%). From this study, thirty-nine ANNs were preliminarily reported as the most acceptable prediction models to correlate multivariate relationships among gas-phase compositions and BHT data *(*referred as ANN-1 to ANN-39 in Pérez-Zárate et al. 2019*).*

By using traditional evaluation metrics based on small differences and acceptable correlation coefficients between measured (BHT_m) and simulated (BHT_{ANN}) temperatures, six ANNs were proposed as the '*better prediction models*' (cited as ANN-12, ANN-13, ANN-22, ANN-25, ANN-33, and ANN-38 in the same paper). Although acceptable results were obtained from these earlier prediction models, some applicability problems were later identified in two ANNs that used a small data set for the learning process (WG_SubDB_3: $q_3 = 97$, which was used for the ANN-33 and ANN-38 models). The low representativeness of this sub-data set was lastly reflected on very limited intervals of applicability.

A generalization problem was therefore identified in the final development stage of new gas geothermometers, which may affect the future application of these tools in geothermometric studies for the geothermal prospection and exploitation. To correct these problems, a new evaluation methodology based on a novel MCDA method was applied for the optimal selection of the most reliable ANNs among the thirty-nine architectures previously proposed (Fig. 1b). The MCDA method was used, for the first time in the ANN literature, to comprehensively evaluate the efficiency of these ANN prediction models by utilizing the evaluation results derived from the following:

I. The analysis Case 1 (AC_1) described by the global learning processes of three sub-databases (WG_SubDB_1, WG_SubDB_2, and WG_SubDB_3) applying the thirty-nine ANNs (pre-selected) and the following statistical metrics: the number of data used by each sub-database ($q_1 = 527$; $q_2 = 498$ and $q_3 = 97$), the correlation coefficients (r) obtained between measured (BHT_m) and simulated (BHT_{ANN}) temperatures for the training, validation, and testing stages, and the residuals calculated between measured and predicted temperatures (RMSE, MAE, and MAPE);
II. The analysis Case 2 (AC_2) described by the application of the thirty-nine ANNs to the largest sub-database ($q_1 = 527$), as a generalization case, using the same statistical residuals between measured and predicted temperatures (RMSE, MAE, and MAPE); and
III. The analysis Case 3 (AC_3) described by the application of the thirty-nine ANNs to the same sub-database ($q_1 = 527$) but strongly restricted by the applicability conditions of each ANN prediction model, using the same residuals between measured and predicted temperatures.

Table 4 A summary of the thirty-nine ANNs (proposed by Pérez-Zárate et al., 2019) for determining reservoir temperatures

ANN	Input variables					ANN architecture	Statistical evaluation metrics											
							Original learning process[a]						Full data[b]			After applicability conditions[c]		
	1	2	3	4	5		n data	(*r*) training	(*r*) validation	(*r*) testing	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
WG_SubDB$_1$: q_1=527																		
1	ln(CO_2)					[1-32-1]	527	0.69	0.58	0.62	34.83	27.08	10.29	34.83	27.08	10.29	34.83	27.08	10.29
2	ln(H_2S)					[1-3-1]	527	0.67	0.59	0.67	35.45	27.58	10.46	35.45	27.58	10.46	35.45	27.58	10.46
3	ln(CO_2/H_2)					[1-13-1]	527	0.68	0.64	0.55	35.47	27.85	10.31	35.47	27.85	10.31	35.47	27.85	10.31
4	ln(H_2S/H_2)					[1-26-1]	527	0.48	0.35	0.36	42.06	32.65	12.27	42.06	32.65	12.27	42.06	32.65	12.27
5	ln(CH_4/CO_2)	ln(H_2S/H_2)				[2-28-1]	527	0.69	0.57	0.31	38.21	27.94	10.44	38.21	27.94	10.44	36.68	27.36	10.25
6	H_2S/CO_2	CH_4/CO_2	H_2/CO_2			[3-12-1]	527	0.80	0.74	0.74	29.26	20.98	7.99	29.26	20.98	7.99	29.12	20.88	7.94
7	ln(H_2S/CO_2)	ln(CH_4/CO_2)	ln(H_2/CO_2)			[3-18-1]	527	0.85	0.80	0.64	26.84	18.39	6.87	26.84	18.39	6.87	26.46	18.17	6.78
8	ln(CO_2/H_2S)	ln(CH_4/H_2S)	ln(H_2/H_2S)			[3-9-1]	527	0.83	0.80	0.61	28.22	19.45	7.31	28.22	19.45	7.31	28.22	19.45	7.31
9	ln(CO_2/CH_4)	ln(H_2S/CH_4)	ln(H_2/CH_4)			[3-15-1]	527	0.85	0.82	0.63	26.45	18.94	7.14	26.45	18.94	7.14	26.43	18.91	7.12
10	ln(CO_2/H_2)	ln(H_2S/H_2)	ln(CH_4/H_2)			[3-9-1]	527	0.83	0.79	0.70	27.24	19.36	7.27	27.24	19.36	7.27	27.24	19.36	7.27
11	ilr_1	ilr_2	ilr_3			[3-9-1]	527	0.83	0.82	0.59	27.94	19.93	7.40	27.94	19.93	7.40	27.92	19.89	7.38
12	ln(CO_2)	ln(H_2S)	ln(CH_4)	ln(H_2)		[4-13-1]	527	0.86	0.78	0.75	25.77	17.60	6.66	25.76	17.60	6.66	25.74	17.57	6.64
13	ln(H_2S/CO_2)	ln(CH_4/CO_2)	ln(H_2/CO_2)	ln(H_2S)	ln(H_2S/H_2)	[5-9-1]	527	0.85	0.79	0.72	26.50	18.32	6.92	26.50	18.32	6.92	26.50	18.32	6.92

Table 4 (continued)

ANN	Input variables					ANN architecture	Statistical evaluation metrics											
							Original learning process[a]						Full data[b]			After applicability conditions[c]		
	1	2	3	4	5		n data	(*r*) training	(*r*) validation	(*r*) testing	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
WG_SubDB$_2$: q_2=498																			
14	$ln(CO_2)$					[1-33-1]	498	0.71	0.66	0.71	33.68	26.39	10.10	40.13	30.11	11.37	34.40	26.82	10.19
15	$ln(H_2S)$					[1-34-1]	498	0.72	0.63	0.63	33.85	25.61	9.68	35.03	26.64	9.96	35.01	26.52	9.94
16	$ln(CO_2/H_2)$					[1-16-1]	498	0.64	0.63	0.73	36.06	28.76	10.69	37.05	29.63	10.93	36.68	29.17	10.75
17	$ln(H_2S/H_2)$					[1-25-1]	498	0.48	0.46	0.38	41.94	32.99	12.41	42.30	33.26	12.43	42.38	33.33	12.48
18	$ln(CH_4/CO_2)$	$ln(H_2S/H_2)$				[2-34-1]	498	0.61	0.56	0.54	38.29	28.22	10.62	39.54	29.38	10.96	38.62	28.46	10.64
19	H_2S/CO_2	CH_4/CO_2	H_2/CO_2			[3-27-1]	498	0.82	0.80	0.80	27.74	19.94	7.64	146.81	41.14	14.80	28.59	20.39	7.74
20	$ln(H_2S/CO_2)$	$ln(CH_4/CO_2)$	$ln(H_2/CO_2)$			[3-18-1]	498	0.83	0.79	0.83	26.85	19.00	7.27	29.22	20.66	7.78	27.46	19.30	7.29
21	$ln(CO_2/H_2S)$	$ln(CH_4/H_2S)$	$ln(H_2/H_2S)$			[3-6-1]	498	0.80	0.75	0.75	29.13	21.26	8.11	30.64	22.25	8.41	29.90	21.72	8.21
22	$ln(CO_2/CH_4)$	$ln(H_2S/CH_4)$	$ln(H_2/CH_4)$			[3-34-1]	498	0.88	0.80	0.79	24.91	16.60	6.33	48.58	22.76	8.38	23.76	16.33	6.12
23	$ln(CO_2/H_2)$	$ln(H_2S/H_2)$	$ln(CH_4/H_2)$			[3-15-1]	498	0.85	0.77	0.77	26.01	18.19	6.84	31.94	20.69	7.68	26.54	18.47	6.87
24	ilr_1	ilr_2	ilr_3			[3-14-1]	498	0.85	0.76	0.77	26.44	18.45	7.00	35.17	21.90	8.12	27.16	18.89	7.10
25	$ln(CO_2)$	$ln(H_2S)$	$ln(CH_4)$	$ln(H_2)$		[4-10-1]	498	0.86	0.80	0.79	24.93	17.20	6.47	59.96	22.64	8.55	25.80	17.66	6.58
26	$ln(H_2S/CO_2)$	$ln(CH_4/CO_2)$	$ln(H_2/CO_2)$	$ln(H_2S)$	$ln(H_2S/H_2)$	[5-5-1]	498	0.81	0.78	0.82	27.86	20.55	7.81	323.58	69.74	24.67	28.34	20.86	7.87
WG_SubDB$_3$: q_3=97																			
27	$ln(CO_2)$					[1-10-1]	97	0.77	0.73	0.83	23.95	17.37	6.31	126.01	71.75	25.33	44.39	36.45	12.60

Table 4 (continued)

ANN	Input variables					ANN architecture	Statistical evaluation metrics											
							Original learning process[a]						Full data[b]			After applicability conditions[c]		
	1	2	3	4	5		n data	(*r*) training	(*r*) validation	(*r*) testing	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
28	$\ln(H_2S)$					[1-14-1]	97	0.81	0.83	0.82	21.62	15.60	5.81	57.35	39.24	14.34	41.09	32.97	12.02
29	$\ln(CO_2/H_2)$					[1-4-1]	97	0.56	0.59	0.51	30.51	22.72	8.25	49.59	41.80	14.66	50.63	42.93	14.62
30	$\ln(H_2S/H_2)$					[1-8-1]	97	0.71	0.71	0.68	25.98	19.71	7.01	52.31	40.83	15.11	49.80	38.88	14.12
31	$\ln(CH_4/CO_2)$	$\ln(H_2S/H_2)$				[2-9-1]	97	0.86	0.82	0.73	20.05	13.47	4.80	107.09	61.29	22.34	52.88	38.80	13.19
32	H_2S/CO_2	CH_4/CO_2	H_2/CO_2			[3-8-1]	97	0.90	0.87	0.86	16.88	10.63	3.86	172.56	67.79	24.89	36.21	26.83	9.45
33	$\ln(H_2S/CO_2)$	$\ln(CH_4/CO_2)$	$\ln(H_2/CO_2)$			[3-8-1]	97	0.95	0.92	0.91	12.56	8.64	3.08	93.54	57.17	20.69	44.15	31.54	11.23
34	$\ln(CO_2/H_2S)$	$\ln(CH_4/H_2S)$	$\ln(H_2/H_2S)$			[3-8-1]	97	0.92	0.90	0.90	14.83	10.94	3.88	63.15	42.59	16.23	40.54	29.64	10.57
35	$\ln(CO_2/CH_4)$	$\ln(H_2S/CH_4)$	$\ln(H_2/CH_4)$			[3-11-1]	97	0.94	0.92	0.86	14.18	9.64	3.50	124.97	65.67	23.30	50.39	37.10	12.78
36	$\ln(CO_2/H_2)$	$\ln(H_2S/H_2)$	$\ln(CH_4/H_2)$			[3-7-1]	97	0.92	0.91	0.79	15.87	11.45	4.14	88.95	53.02	19.17	45.02	33.73	11.70
37	ilr_1	ilr_2	ilr_3			[3-9-1]	97	0.92	0.93	0.82	16.08	10.74	3.78	63.24	40.40	15.03	35.66	25.94	9.08
38	$\ln(CO_2)$	$\ln(H_2S)$	$\ln(CH_4)$	$\ln(H_2)$		[4-8-1]	97	0.94	0.92	0.92	14.56	10.33	3.68	102.77	64.26	22.86	43.31	32.94	11.29
39	$\ln(H_2S/CO_2)$	$\ln(CH_4/CO_2)$	$\ln(H_2/CO_2)$	$\ln(H_2S)$	$\ln(H_2S/H_2)$	[5-6-1]	97	0.92	0.90	0.88	15.37	10.75	3.84	66.45	44.07	15.72	39.57	29.27	10.21

[a] The global learning process was reported in Pérez-Zárate et al. (2019): n data (number of samples used in each sub-database); (*r*) training, validation and testing (the correlation coefficient calculated between the bottom-hole temperatures (measured) and the temperature estimates (predicted) by ANNs; RMSE, MAE and MAPE (residuals calculated between the bottom-hole temperatures and temperature estimates by ANNs)

[b] Full data: RMSE, MAE and MAPE (residuals calculated using the largest representative sub-database: $q_1 = 527$)

[c] After applicability conditions: RMSE, MAE and MAPE (residuals calculated using the samples that fulfil the applicability conditions of the ANN architectures (reported in Table 8)

To apply the MCDA, it was necessary to normalize the resulting thirteen statistical metrics (Table 4). The normalization was performed either to minimize or maximize the evaluation metrics using the methodology proposed by Dincer and Acar (2015).

For example, in the particular case of the residual errors (RMSE, MAE, and MAPE), the best efficiency of the ANN prediction models would be given when these metrics achieve lower values, whereas for the linear correlation coefficient (r), higher values would be desirable. For the statistical residual (SR_i), or the minimization obtained during the global learning and application processes of the ANNs, normalized data were calculated by using the RMSE, MAE, and MAPE results and the following equation:

$$\text{Normalized Data}_i = \frac{SR_{\max} - SR_i}{SR_{\max} - SR_{\min}} \times 10 \tag{2}$$

whereas for the maximization, the normalized data was determined by using the following equation:

$$\text{Normalized Data}_i = \frac{SP_i - SP_{\min}}{SP_{\max} - SP_{\min}} \times 10, \tag{3}$$

where SP_i are the statistical regression parameters: n (the data number used by each database) and r, the correlation coefficients, which were obtained from the training, validation, and testing of the ANNs.

The MCDA is suggested as an optimized method for the integral evaluation of various statistical metrics among predictor models using different scenarios to achieve a specific objective function (Adem and Geneletti 2018). To apply the MCDA, the Multi-Attribute Value Theory (MAVT) algorithm was used (Santoyo-Castelazo 2011; Santoyo-Castelazo et al. 2011; Santoyo-Castelazo and Azapagic 2014; Estévez et al. 2018). This algorithm required the determination of partial value functions, and the estimation of weighting factors for each evaluation metric. The global value function $V(s)$ which represents the total score to be obtained by any ANN architecture in each scenario s is calculated as follows:

$$V(s) = \sum_{i=1}^{E} w_i u(s), \tag{4}$$

where w_i and $u(s)$ are the weighting factors used by the evaluation metric, and the value function which provides the metric efficiency for each ANN architecture and scenario, respectively; and E is the total number of the evaluation metrics.

Wang et al. (2009) suggested the use of variable weighting factors by considering the importance of the scenarios assumed on the predictor model efficiency. To apply the MCDA method, Santoyo-Castelazo and Azapagic (2014) suggested the use of a sensitivity analysis to determine the ranking score of prediction models based on different scenarios that enable indicators (or statistical metrics) to be varied under certain weighting criterion. In this work, the sensitivity analysis was applied for evaluating the prediction efficiency of each ANN model in the calculation of reservoir temperatures using four different scenarios (S), where thirteen statistical metrics were comprehensively analysed by assigning the following weighting criteria:

- (S-1) Equal weighting factors assigned to the statistical metrics computed for the three analysis cases already postulated: AC_1 (global learning processes); AC_2 (the generalization case); and AC_3 (the applicability restrictions imposed by each ANN model), i.e. $AC_1 = AC_2 = AC_3$, Fig. 1b; see Additional file 1: Table S1.
- (S-2) Greater weighting factors assigned to the seven statistical metrics computed for the AC_1 (global learning processes) in comparison with equal weighting values assumed for the metrics of the AC_2 and AC_3, i.e. $AC_1 > (AC_2 = AC_3)$, Fig. 1b; see Additional file 1: Table S2.
- (S-3) Greater weighting factors assigned to the three statistical metrics computed for the AC_2 (the generalization case) in comparison with equal weighting values assumed for the metrics of the AC_1 and AC_3, i.e. $AC_2 > (AC_1 = AC_3)$, Fig. 1b; see Additional file 1: Table S3.
- (S-4) Greater weighting factors assigned to the three statistical metrics computed from the AC_3 (the restricted case of the applicability conditions) in comparison with equal weighting values assumed for the metrics of the AC_1 and AC_2, i.e. $AC_3 > (AC_1 = AC_2)$, Fig. 1b; see Additional file 1: Table S4.

After applying the MCDA (by means of the algorithm represented in Fig. 1b), the best ANN architectures were optimally selected and used for the development of the new gas geothermometers. A summarized description of the ranking score obtained by each ANN together with the optimal selection results are reported in Table 5. A full report containing all MCDA calculations obtained for the thirty-nine ANNs is reported in Additional file 1: Tables S1 to S4.

GasGeo_Eq

This module was created to describe the new gas geothermometer equations inferred from the best ANNs (from here referred as $GasG_1$ to $GasG_8$), and their applicability conditions. Table 6 summarizes the eight ANNs that were optimally selected for the development of the new gas geothermometer equations. The number of neurons used for the three layers of the ANNs (input, hidden, and output) are included. The input variables used by each ANN are also reported, including their relative contribution (in %) obtained from the sensitivity analysis (Garson 1991). Weighting and bias coefficients of the optimal ANNs are reported in Table 7. These coefficients were used for the development of each gas geothermometer ($GasG_i$, $i = 1$ to 8) by using the following general equation (Eq. 5):

$$\mathrm{BHT}_{\mathrm{GasG_i}}\left(^{\circ}C\right) = \left\{ \left[\frac{\left(\sum_{j=1}^{m} \left(LW_{(m,n)} \cdot \left(\frac{2}{1+\exp\left(-2\cdot\left(\sum_k IW_{(m,k)} \cdot X_{(k)} + b_{1(m)}\right)\right)} - 1 \right) \right) + b_{2(n)} \right) + 1}{2} \right] \cdot \alpha \right\} + 170, \tag{5}$$

where X is the input variable; IW and b_1 represent the associated coefficients to the input–hidden layers of the ANN; and LW and b_2 are the coefficients for the hidden–output layers. The subscripts k, m, and n refer to input, hidden, and output neurons, respectively, whereas α represents a normalization factor equal to 214 for the $GasG_1$, $GasG_2$, $GasG_6$, and $GasG_8$, and 193 for the remaining geothermometers ($GasG_3$, $GasG_4$, $GasG_5$, and $GasG_7$).

Table 5 Ranking score and optimal selection of the eight ANN architectures using the MCDA method

Ranking/ best selection	Neural network model	ANN architecture	Score				Score average
			Scenario 1	Scenario 2	Scenario 3	Scenario 4	
1	ANN-22[a]	[3-34-1]	8.45	8.12	8.58	8.74	8.47
2	ANN-12[a]	[4-13-1]	8.38	7.98	8.68	8.57	8.40
3	ANN-25[a]	[4-10-1]	8.23	7.94	8.37	8.45	8.25
4	ANN-13[a]	[5-9-1]	8.18	7.78	8.50	8.36	8.20
5	ANN-23[a]	[3-15-1]	8.12	7.76	8.39	8.31	8.15
6	ANN-7[a]	[3-18-1]	8.11	7.68	8.45	8.32	8.14
7	ANN-9[a]	[3-15-1]	8.04	7.63	8.37	8.21	8.06
8	ANN-20[a]	[3-18-1]	8.02	7.68	8.31	8.15	8.04
9	ANN-24	[3-14-1]	7.97	7.62	8.23	8.14	7.99
10	ANN-10	[3-9-1]	7.95	7.55	8.29	8.10	7.97
11	ANN-8	[3-9-1]	7.76	7.33	8.13	7.92	7.78
12	ANN-11	[3-9-1]	7.74	7.32	8.11	7.90	7.77
13	ANN-6	[3-12-1]	7.49	7.09	7.87	7.60	7.51
14	ANN-21	[3-6-1]	7.33	6.96	7.71	7.42	7.36
15	ANN-37	[3-9-1]	7.22	7.44	7.12	7.04	7.21
16	ANN-19	[3-27-1]	6.89	6.86	6.68	7.16	6.90
17	ANN-39	[5-6-1]	6.86	7.22	6.75	6.50	6.83
18	ANN-34	[3-8-1]	6.84	7.24	6.75	6.43	6.82
19	ANN-33	[3-8-1]	6.46	7.19	6.05	5.95	6.41
20	ANN-38	[4-8-1]	6.06	6.82	5.55	5.61	6.01
21	ANN-15	[1-34-1]	5.95	5.53	6.47	5.95	5.97
22	ANN-32	[3-8-1]	5.98	6.58	5.24	5.97	5.94
23	ANN-36	[3-7-1]	5.96	6.54	5.76	5.44	5.93
24	ANN-14	[1-33-1]	5.87	5.52	6.31	5.87	5.89
25	ANN-28	[1-14-1]	5.88	6.06	6.08	5.45	5.87
26	ANN-1	[1-32-1]	5.69	5.23	6.25	5.70	5.72
27	ANN-2	[1-3-1]	5.62	5.18	6.18	5.61	5.65
28	ANN-3	[1-13-1]	5.56	5.10	6.13	5.56	5.59
29	ANN-26	[5-5-1]	5.56	5.95	4.56	6.06	5.53
30	ANN-16	[1-16-1]	5.42	5.02	5.97	5.36	5.44
31	ANN-35	[3-11-1]	5.49	6.44	5.02	4.79	5.43
32	ANN-5	[2-28-1]	5.06	4.47	5.71	5.15	5.10
33	ANN-18	[2-34-1]	4.98	4.47	5.61	4.98	5.01
34	ANN-31	[2-9-1]	4.67	5.37	4.47	4.00	4.63
35	ANN-30	[1-8-1]	4.42	4.57	4.86	3.79	4.41
36	ANN-27	[1-10-1]	4.20	4.74	3.83	3.89	4.16
37	ANN-17	[1-25-1]	3.51	2.90	4.31	3.46	3.54
38	ANN-4	[1-26-1]	3.47	2.81	4.30	3.47	3.51
39	ANN-29	[1-4-1]	3.30	3.23	3.96	2.73	3.31

[a] Neural network models that represent the new gas geothermometer equations proposed in this work

Considering the complexity of the Eq. (5), it is pertinent to assume that a future application of the eight geothermometric equations could be discouraged by users when compared with simple equations already proposed from the existing gas geothermometers (Table 3). To overcome this limitation, a computer program (GaS_GeoT) codified

Table 6 New improved gas geothermometers proposed in this research work

Neural network model	New gas geothermometer nomenclature ($GasG_i$)	Input variables					Number of inputs neurons	Number of hidden neurons	Number of outputs neurons	ANN architecture
		1	2	3	4	5				
ANN-7	$GasG_1$	$\ln(H_2S/CO_2)$ 46.32%	$\ln(CH_4/CO_2)$ 24.84%	$\ln(H_2/CO_2)$ 28.84%			3	18	1	[3–18-1]
ANN-9	$GasG_2$	$\ln(CO_2/CH_4)$ 42%	$\ln(H_2S/CH_4)$ 31.02%	$\ln(H_2/CH_4)$ 26.97%			3	15	1	[3–15-1]
ANN-20	$GasG_3$	$\ln(H_2S/CO_2)$ 39.12%	$\ln(CH_4/CO_2)$ 37.68%	$\ln(H_2/CO_2)$ 23.2%			3	18	1	[3–18-1]
ANN-22	$GasG_4$	$\ln(CO_2/CH_4)$ 33.05%	$\ln(H_2S/CH_4)$ 34.51%	$\ln(H_2/CH_4)$ 32.44%			3	34	1	[3–34-1]
ANN-23	$GasG_5$	$\ln(CO_2/H_2)$ 24.6%	$\ln(H_2S/H_2)$ 50.2%	$\ln(CH_4/H_2)$ 25.21%			3	15	1	[3–15-1]
ANN12	$GasG_6$	$\ln(CO_2)$ 26.18%	$\ln(H_2S)$ 30.19%	$\ln(CH_4)$ 15.37%	$\ln(H_2)$ 28.26%		4	13	1	[4–13-1]
ANN-25	$GasG_7$	$\ln(CO_2)$ 27.83%	$\ln(H_2S)$ 12.81%	$\ln(CH_4)$ 25.87%	$\ln(H_2)$ 33.48%		4	10	1	[4–10-1]
ANN-13	$GasG_8$	$\ln(H_2S/CO_2)$ 25.13%	$\ln(CH_4/CO_2)$ 5.11%	$\ln(H_2/CO_2)$ 10.29%	$\ln(H_2S)$ 42.86%	$\ln(H_2S/H_2)$ 16.61%	5	9	1	[5-9-1]

According to the Garson's method (Garson 1991), the relative contribution for each input variable was determined in subsequent calculations of the geothermal reservoir temperatures. This value appears in percentage units below each input variable

Table 7 Weighting and bias coefficients of the best ANN models (new gas geothermometers: $GasG_1$ to $GasG_8$)

Neuron	$IW(_{m,k})$					$b1(_{m,1})$	$LW(_{m,n})$	$b2(_n)$
	Input variable 1	Input variable 2	Input variable 3	Input variable 4	Input variable 5	Bias	Weight	Bias
$GasG_1$ ($k=3$, $m=18$, $n=1$)								
1	−10.4966	−1.1378	0.8463			9.5785	−1.7731	−1.0356
2	22.0807	−12.9076	19.0550			−10.6989	−0.4489	
3	−14.8833	4.3913	−0.1612			8.5887	4.2903	
4	−51.9879	−0.3416	5.9872			19.3727	−0.2385	
5	16.7261	−4.8153	−0.0926			−9.4315	4.2683	
6	4.3149	3.8523	−5.4022			−0.5974	0.8856	
7	10.5391	−20.9557	33.7107			2.5871	−0.2241	
8	−2.0890	1.1073	20.2039			7.9015	0.3611	
9	0.5418	1.3637	−2.3029			0.0489	−3.3014	
10	−5.6519	0.9427	4.8160			1.3659	−10.5466	
11	1.1402	−3.2332	2.6054			−4.1995	1.5774	
12	5.2141	−0.5295	−4.0756			−1.3200	−7.5987	
13	3.6485	−19.4886	−0.8257			12.7539	5.0275	
14	−4.4414	0.0312	0.4327			4.3220	6.0399	
15	−3.5400	−0.7734	3.5304			1.4321	−2.7057	
16	−4.0591	19.9584	0.6652			−13.0292	5.0045	
17	5.5351	−1.0741	−5.2122			−1.3846	−4.8823	
18	2.7430	1.2762	0.3413			1.2202	−0.9442	
$GasG_2$ ($k=3$, $m=15$, $n=1$)								
1	−74.7200	−29.3826	71.2961			−9.1189	0.1740	−0.4764
2	−51.6547	21.8776	5.1177			13.0007	−0.3961	
3	−0.5704	−8.5959	11.2749			1.0424	−11.5175	
4	−2.8170	−5.8255	2.3507			6.9290	0.1195	
5	138.9491	−164.2573	14.7326			−2.4335	−0.2194	
6	10.0731	10.0604	3.4971			−0.7834	−10.0376	
7	8.2489	8.4971	2.0266			−0.4685	20.6834	
8	−6.9094	4.3475	2.2922			0.6182	0.5344	
9	−0.6145	−8.8942	11.4466			0.9874	11.3964	
10	−6.5749	−7.2235	−1.1089			0.2658	10.8816	
11	−109.4214	3.9406	18.2203			−19.5511	−0.2835	
12	68.6847	−37.4115	−78.8355			0.2217	0.1522	
13	−7.8109	1.3878	−6.8636			−9.5630	−14.2420	
14	8.0752	−1.1217	7.0651			10.1010	−13.8272	
15	−20.7201	10.9331	−12.9707			−8.4363	−0.2044	
$GasG_3$ ($k=3$, $m=18$, $n=1$)								
1	0.2605	−2.8324	0.7405			−4.6395	0.0965	−0.6741
2	−4.1589	5.9679	0.2775			5.0844	0.4591	
3	1.2094	13.6531	3.8001			−6.2841	0.1207	
4	−18.5398	0.6195	0.5262			10.0722	−0.5122	
5	2.7557	−5.8167	2.0639			−6.4099	0.0843	
6	−2.5005	−0.3730	−2.0823			−0.7033	−2.6769	
7	−8.7356	3.4340	1.9566			5.5936	−0.4428	
8	−3.9373	−0.2286	3.8324			1.0464	−0.6335	
9	2.1539	0.4060	3.8254			1.0421	−3.2267	
10	0.5414	12.1314	0.1324			−0.6745	0.5703	

Table 7 (continued)

Neuron	$IW_{(m,k)}$					$b1_{(m,1)}$	$LW_{(m,n)}$	$b2_{(n)}$
	Input variable 1	Input variable 2	Input variable 3	Input variable 4	Input variable 5	Bias	Weight	Bias
11	2.4237	9.6744	0.8198			−1.0205	−0.4172	
12	2.1196	0.1036	−0.6293			−1.2896	−1.8327	
13	−3.1159	−1.4763	0.4641			−5.7187	−0.1605	
14	0.8259	1.6177	9.0616			1.9310	1.2339	
15	3.6085	−2.0623	−2.2694			5.3824	−0.1270	
16	−13.5662	1.6842	−2.0641			3.1308	−0.3221	
17	1.0129	4.8087	0.1825			4.8000	−0.2732	
18	4.7110	6.9895	−4.6717			−2.6020	−0.2882	
$GasG_4$ ($k=3, m=34, n=1$)								
1	0.8413	1.0481	4.5492			−4.3604	0.5430	0.7175
2	1.6852	−1.1809	4.0247			−4.2626	−0.6184	
3	−3.1846	3.0345	0.6129			4.2989	0.4547	
4	−2.4474	−2.2656	3.1111			3.6910	−0.1940	
5	−1.7082	4.3553	1.0333			3.4071	−2.2762	
6	−4.8797	6.0548	0.1265			2.8825	1.9591	
7	0.4419	4.7801	0.6585			−1.5392	2.1423	
8	1.9559	3.0935	−4.7555			−1.7757	3.0085	
9	−0.3619	−5.2932	6.4612			1.1353	3.5850	
10	1.5416	−1.9525	4.7684			−2.3951	2.4549	
11	−4.0310	−2.8368	8.6461			2.0696	−3.2083	
12	1.3875	2.8019	2.1562			−0.2151	−2.1271	
13	−6.2492	−0.8381	4.9963			−0.5368	2.3720	
14	−5.4812	−2.5508	0.2280			1.4242	2.2546	
15	1.2885	3.7342	−3.5705			−1.3129	−3.2678	
16	0.6120	6.8702	1.6604			0.4849	5.1290	
17	−8.0473	7.0992	3.3462			−0.6557	6.2486	
18	−6.5842	4.8933	4.9639			−0.5687	−5.3591	
19	2.0680	−2.0030	3.2922			0.5423	1.8140	
20	−2.7406	9.8262	8.5937			2.1876	0.6365	
21	7.7279	2.7843	−10.4992			0.5845	1.5671	
22	−10.3807	11.2299	1.2403			−0.8739	−3.2182	
23	−2.5572	−7.6695	−0.8698			−1.1968	2.8984	
24	6.7225	−3.9473	2.0077			2.1102	−4.8149	
25	12.5036	3.5952	−5.8924			2.4635	0.5574	
26	1.1006	−0.6048	−4.1736			2.9012	−0.9523	
27	−3.2534	5.5943	−2.5071			−0.9127	2.2162	
28	6.1765	2.4106	−1.4982			5.4936	7.1628	
29	−8.6020	4.5660	−3.5287			−3.1652	−3.1148	
30	2.2915	−3.0145	−2.4441			3.5085	0.0843	
31	−5.0789	−1.0713	−4.5243			−7.5724	−3.3251	
32	−1.0222	2.7998	−5.5370			−3.9064	−2.1729	
33	−10.9591	−4.7211	4.3928			−8.6022	3.8019	
34	6.5221	1.8689	2.3093			8.0946	−5.0940	
$GasG_5$ ($k=3, m=15, n=1$)								
1	4.6610	9.3354	−2.9977			−10.1278	1.2400	−1.9220
2	−2.0011	3.1019	4.6559			1.5880	2.0416	
3	−1.3652	−6.8953	0.2085			−4.3337	−3.5757	
4	2.2086	6.9970	−0.7490			4.9651	−3.5151	

Table 7 (continued)

Neuron	IW($_{m,k}$)					b1($_{m,1}$)	LW($_{m,n}$)	b2($_n$)
	Input variable 1	Input variable 2	Input variable 3	Input variable 4	Input variable 5	Bias	Weight	Bias
5	15.0925	12.5673	−1.6365			−2.8091	−0.6210	
6	−0.2917	−4.4285	−4.8459			−2.2553	1.9381	
7	−1.4554	−3.8719	0.6722			0.5524	−3.7513	
8	14.6274	−18.1442	1.1720			2.4447	−0.3234	
9	−4.1866	2.2442	6.8280			−2.1283	0.5290	
10	1.2033	−3.7615	−7.3399			1.4214	0.4381	
11	−0.9754	−5.1881	0.8466			0.5475	2.5457	
12	−4.1395	−8.2389	14.8899			−5.2272	−0.1067	
13	4.3414	15.7990	7.7251			−8.5051	−0.5169	
14	3.0032	0.5427	0.7922			2.7043	2.2648	
15	−0.2435	3.8179	0.3571			−2.9697	−1.0737	
$GasG_6$ ($k=4$, $m=13$, $n=1$)								
1	−7.0424	1.0592	0.5872	−4.9261		−0.9909	0.6853	−1.4850
2	−7.7923	0.7611	2.4586	11.5728		3.6960	1.0193	
3	−4.0733	−1.6240	1.7513	−3.4984		−0.9385	−0.7838	
4	−2.3369	−9.6032	0.3973	−0.2256		4.4306	−0.6460	
5	4.6053	3.2060	−6.2870	2.5629		0.3976	−1.5127	
6	4.7627	2.8469	−6.2893	3.3543		0.4991	1.4433	
7	−0.7276	5.4561	1.5209	4.0939		2.8770	3.4748	
8	−6.3221	5.2401	−0.0758	−1.8561		−3.3356	−0.7253	
9	0.0618	3.3627	0.2201	2.8292		−5.0749	−3.2557	
10	1.3088	1.8037	0.9426	6.5565		2.9480	−2.2063	
11	5.6895	3.8142	−3.7349	0.3184		3.0027	−0.6407	
12	3.5135	−3.4591	2.5028	−4.2444		5.3923	−1.8108	
13	−1.2860	−6.6278	−2.7831	−3.4112		−2.6695	2.0116	
$GasG_7$ ($k=4$, $m=10$, $n=1$)								
1	7.9032	−0.6015	0.7257	−0.8062		−4.9130	−0.4283	−0.6240
2	−2.9711	−3.8961	−0.3714	5.0930		2.1358	−0.4423	
3	0.6463	−1.2954	0.1709	7.7610		−0.2102	6.7284	
4	4.6109	−0.1063	−18.6311	−1.7928		−3.7415	−4.8884	
5	−8.3235	2.7237	−6.0207	−4.6909		−1.0212	0.7313	
6	5.7018	−4.4436	7.5091	−5.5924		7.9518	7.8252	
7	−5.6660	3.7079	−8.5863	7.8219		−8.8848	7.7244	
8	−0.5289	1.7097	−0.2783	−8.7773		0.0500	6.2932	
9	5.7362	−0.4683	−22.2945	−1.7934		−4.4083	4.7232	
10	8.8670	−2.9926	−0.6986	8.4312		6.8119	0.9991	
$GasG_8$ ($k=5$, $m=9$, $n=1$)								
1	22.4842	−33.9535	−29.5633	15.8975	48.5322	20.3047	0.1346	−1.3893
2	3.4291	0.5420	1.0172	−4.9756	4.5659	0.9671	1.6130	
3	−14.4388	3.7895	−7.2506	79.5516	5.1701	−33.4122	0.2263	
4	−43.2597	−1.2331	−24.8234	50.0978	−9.2286	−14.9195	−0.6028	
5	−2.2463	0.2702	0.2465	2.9529	1.1975	−1.8629	−19.6929	
6	48.8397	−3.6155	19.3550	−55.8381	20.4814	10.5299	−0.2055	
7	−2.9234	−0.2702	−0.2371	5.0167	−2.1591	−0.3359	9.1898	
8	−3.2910	−0.3027	−0.4428	6.2711	−2.9274	−0.4075	−7.1448	
9	1.7139	−0.3069	1.3356	−4.0765	−0.0239	2.9979	−17.7178	

The input variables of each new gas geothermometer are reported in Table 6. The subscripts *k*, *m* and *n* refer to input, hidden and output neurons, respectively

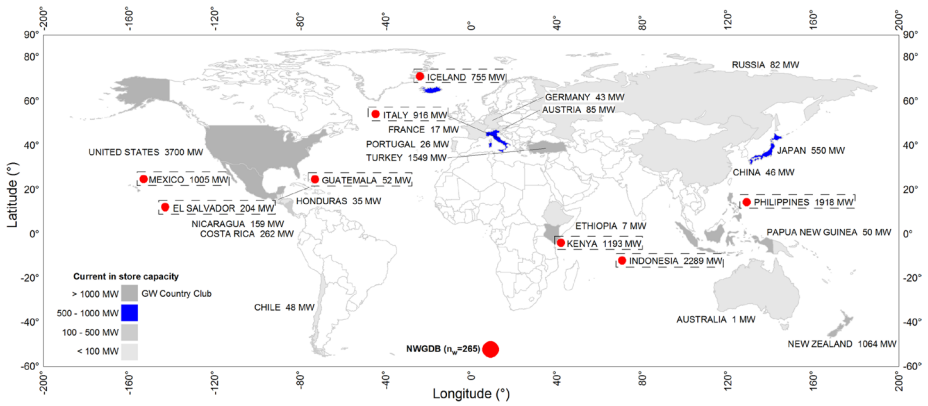

Fig. 3 Location of the worldwide geothermal fields used for the creation of the new Geochemical Database (NWGDB). The updated world geothermal power generation is also shown

in Java object-oriented programming was created. This program was developed for the effective use of the new gas geothermometer equations for the calculation of geothermal temperatures from gas-phase compositions (CO_2, H_2S, CH_4, and H_2). GaS_GeoT and a quick user manual are available in the public server (https://github.com/ANNGroup/GaS_GeoT.git).

The applicability conditions of the new gas geothermometers ($GasG_1$ to $GasG_8$) are reported in Table 8. These limits are the min–max values of concentration and temperature used during the ANN learning processes. The gas concentrations must be given in mmol/mol units (dry-basis), whereas the temperature in °C.

The input variables for the $GasG_1$–$GasG_5$ and $GasG_8$ are given by the natural logarithm values (dimensionless) because these equations use gas ratios, whereas for the $GasG_6$ and $GasG_7$, the input variables are directly given in mmol/mol units.

GasGeo_Lit

This module was created to carry out a comparison of prediction efficiencies among the new ($GasG_1$ to $GasG_8$) and existing gas geothermometers using their temperature estimates. Twenty-five existing gas geothermometers have been included in Table 3. The fundamental criteria for selecting these geothermometers were as follows: (i) the determination of temperatures as a direct function of gas concentrations; and (ii) the use of gas concentrations (CO_2, H_2S, CH_4, H_2, and N_2) most commonly applied in geothermometric studies.

GasChemT_Wells

This module was created for compiling the NWGDB ($n_w = 265$) with data that were not used in the ANN learning process. NWGDB contains compositions (in mmol/mol, dry-basis) of major (CO_2 and H_2S) and trace (H_2, N_2, CH_4) geothermal gases,

and BHT measurements (in °C) logged in wells from thirteen world geothermal fields (Fig. 3): (1) Berlin geothermal field (BGF), El Salvador (compiled from: Renderos 2002; n = 41: collected from a LIQDR); (2) Zunil geothermal field (ZGF), Guatemala (Giggenbach et al., 1992; n = 4: LIQDR); (3) Krafla geothermal field (KGF), Iceland (Stefánsson 2017; n = 8: LIQDR); (4) Kamojang geothermal field (KaGF), Indonesia (Laksminingpuri and Martinus 2013; n = 10: VAPDR); (5) Sibayak geothermal field (SGF), Indonesia (Abidin et al. 2005; n = 9: LIQDR); (6) Amiata geothermal field (AGF), Italy (Chiodini and Marini 1998; n = 4: VAPDR); (7) Larderello geothermal field (LGF), Italy (Chiodini and Marini 1998; n = 30: VAPDR); (8) Olkaria geothermal field (OGF), Kenya (Karingithi et al. 2010; Wamalwa 2015; n = 29: 13 data collected from a LIQDR and 16 from a VAPDR); (9) Cerro Prieto geothermal field (CPGF), Mexico (Nieva et al. 1982; Nehring and D'Amore 1984; n = 9: LIQDR); (10) Las Tres Virgenes geothermal field (LTVGF), Mexico (Verma et al. 2006; n = 4: LIQDR); (11) Los Azufres geothermal field (LAGF), Mexico (Nieva et al. 1985 and 1987; Santoyo et al., 1991; Arellano et al., 2005; Barragán et al., 2005, 2008, 2012; n = 36: 21 data collected from a LIQDR and 15 from a VAPDR; (12) Los Humeros geothermal field (LHGF), Mexico (López-Mendiola and Munguía, 1989; Tello et al., 2000; Arellano et al. 2003, 2015; n = 72: 60 data compiled from a LIQDR and 12 from a VAPDR); and (13) Palinpinon geothermal field (PGF), Philippines (D'Amore et al.1993; n = 9: LIQDR).

To perform the analysis of prediction efficiency in new and existing gas geothermometers, the input data compiled in NWGDB were classified in two major groups: Group-1, LIQDR ($g_1 = 178$), and Group-2, VAPDR ($g_2 = 87$). A complete version of the NWGDB is reported in Additional file 1: Table S5.

GasGeoT_Calc

This module was created to describe the new computer program GaS_GeoT and the effective use of the new gas geothermometers ($GasG_1$ to $GasG_8$). The user interface and a calculation routine of geothermal reservoir temperatures by using GaS_GeoT are also described. Before running GaS_GeoT and to apply the twenty-five geothermometer equations (Table 3), the concentration units of the gas-phase compositions and the applicability conditions of each gas geothermometer were verified (Table 8).

To proceed with the calculation of the reservoir temperatures, all gas geothermometers were executed (see Fig. 1a). As one of the goals of this work was focused on the effective use of the new gas geothermometers ($GasG_1$ to $GasG_8$), the calculation process used by GaS_GeoT is schematically referred in five major sections of Fig. 4 and summarized as follows.

To open or upload the input data through the *File* option (menu 4.1 in Fig. 4), an Excel spreadsheet (with a filename extension: *xlsx*) is required. This file must contain eight columns: Column 1, Sample ID (numerical data: integer); Columns 2–4: Geothermal Well, Geothermal Field, and Country information, respectively (in alphanumerical format); and the Columns 5–8, to enter the gas-phase compositions of the gases CO_2, H_2S, CH_4, and H_2 in mmol/mol (dry-basis), respectively (numerical data in floating format).

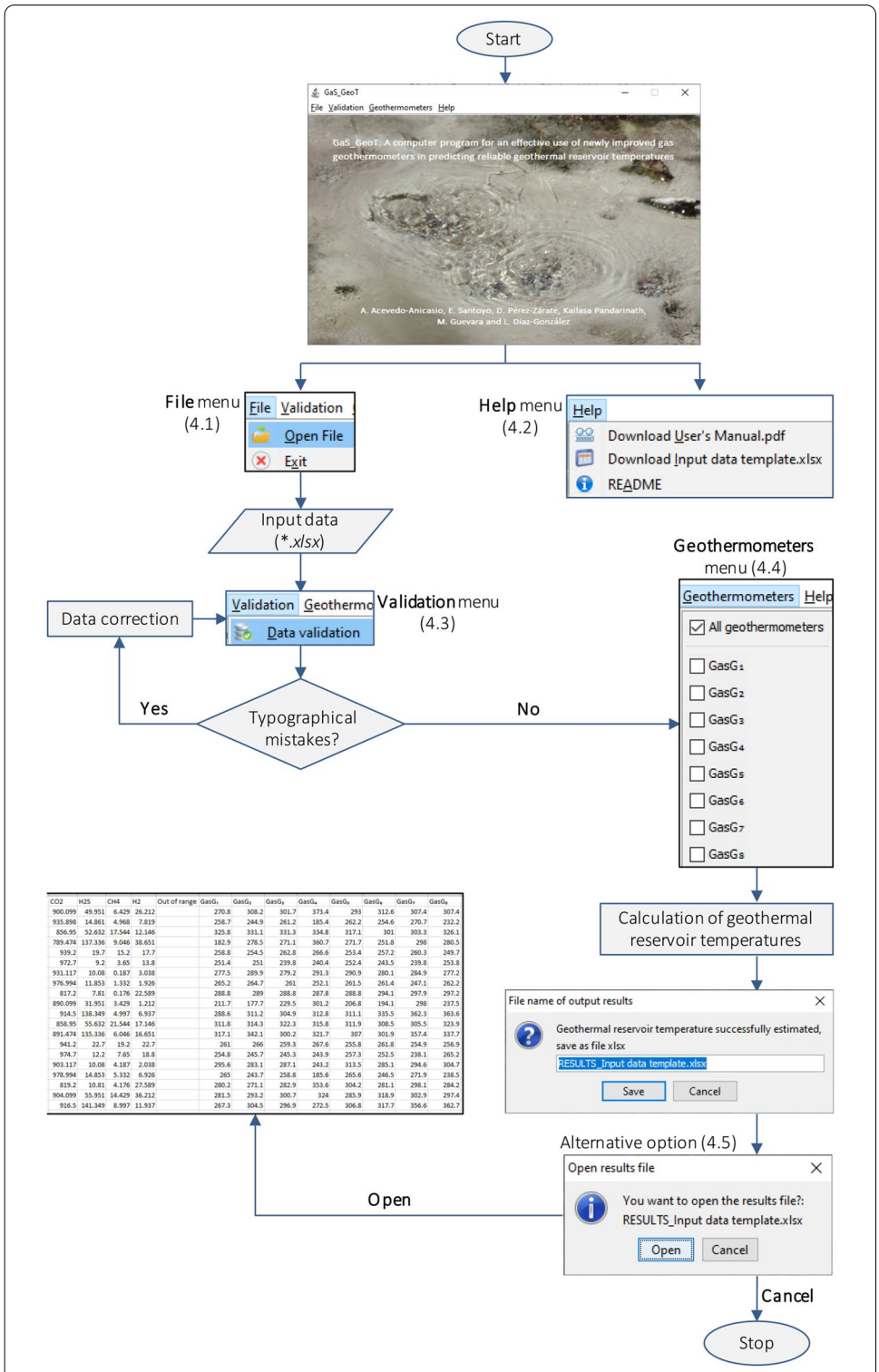

Fig. 4 Schematic flow diagram and menu screens that show the execution of the computer program GaS_GeoT for the estimation of geothermal reservoir temperatures

In the *Help* option (menu 4.2 in Fig. 4), a template file with some input data examples is available as a useful query. A User's Manual is also included in the *Help* option of the program, and the public server (https://github.com/ANNGroup/GaS_GeoT.git).

Two additional menu options appear in the main screen of GaS_GeoT: *Validation* and *Geothermometers*. The *Validation* menu option performs an input data validation for checking the existence of typographical mistakes in the CO_2, H_2S, CH_4, and H_2 compositions (menu 4.3 in Fig. 4). The *Geothermometers* menu option invokes a computer subroutine which will make either the total or partial selection of the gas geothermometers ($GasG_1$ to $GasG_8$) by checking their corresponding boxes (menu 4.4 in Fig. 4). After selecting the geothermometers, the calculation of temperatures is performed, and the GaS_GeoT will prompt the user to provide an output file name to print a report with the gas-phase compositions, and the temperature estimates calculated by the new gas geothermometers. An alternative output option is also requested by GaS_GeoT for displaying the output results on screen (option 4.5 in Fig. 4).

After applying GaS_GeoT and twenty-five existing geothermometers to the gas-phase composition of NWGDB (LIQDR $g_1 = 178$ and VAPDR $g_2 = 87$), an output file with all the temperature estimates is generated both to carry out the geochemometric evaluation, and to estimate the prediction efficiencies of the thirty-three gas geothermometers (Fig. 1a). A complete version of the output file is reported in Additional file 1: Tables S6 and S7, which contain the results obtained from the new and existing gas geothermometers, respectively.

GasGeo_Geochem

This module was developed to perform the geochemometric analyses using the temperature estimates obtained from new and existing gas geothermometers. To carry out these analyses, the following statistical metrics were applied: (1) Percent Difference, DIFF%; (2) Root Mean Square Error, RMSE; (3) Mean Absolute Error, MAE; (4) Mean Absolute Percentage Error, MAPE; and (5) the Difference Coefficient Ratio or statistical Theil's U test. As these metrics require the knowledge of actual bottom-hole temperatures, most of these were used as accuracy measures. The calculation equations used by these metrics are summarized in Table 9. RMSE, MAE, and MAPE metrics are interpreted as statistical residuals obtained between the bottom-hole temperatures (BHT_m) and the temperature estimates predicted by any gas geothermometer, whereas the metrics DIFF% and Theil's U require a previous analysis. For example, when the DIFF% value is positive, the calculated temperature is greater than the BHT_m of the respective geothermal well (which means an overestimation) and vice versa. Differences (DIFF%) either positive or negative $\leq 20\%$ are assumed as acceptable estimates by considering the total propagated errors that are commonly quantified in some solute geothermometers (Verma and Santoyo 1997; García-López et al. 2014).

On the other hand, the statistical Theil's U is recommended for evaluating the prediction efficiency among several predictor models (Theil 1961). This metric was therefore used to evaluate the efficiency of the new geothermometers ($GasG_1$ to $GasG_8$) among other existing gas geothermometers. If the calculated values for the

Theil's coefficient ratios are less than 1, it means that the errors obtained by the new geothermometers are lower than that those obtained from the existing geothermometers and vice versa (Álvarez del Castillo et al. 2012).

Results and discussion

After applying thirteen evaluation metrics and the MCDA, the thirty-nine ANNs were ranked (Table 5). Based on these optimization results, eight new gas geothermometers were successfully developed (Table 6): (1) $GasG_1$ [$\ln(H_2S/CO_2)$, $\ln(CH_4/CO_2)$, $\ln(H_2/CO_2)$]; (2) $GasG_2$ [$\ln(CO_2/CH_4)$, $\ln(H_2S/CH_4)$, $\ln(H_2/CH_4)$]; (3) $GasG_3$ [$\ln(H_2S/CO_2)$, $\ln(CH_4/CO_2)$, $\ln(H_2/CO_2)$]; (4) $GasG_4$ [$\ln(CO_2/CH_4)$, $\ln(H_2S/CH_4)$, $\ln(H_2/CH_4)$]; (5) $GasG_5$ [$\ln(CO_2/H_2)$, $\ln(H_2S/H_2)$, $\ln(CH_4/H_2)$]; (6) $GasG_6$ [$\ln(CO_2)$, $\ln(H_2S)$, $\ln(CH_4)$, $\ln(H_2)$]; (7) $GasG_7$ [$\ln(CO_2)$, $\ln(H_2S)$, $\ln(CH_4)$, $\ln(H_2)$]; and (8) $GasG_8$ [$\ln(H_2S/CO_2)$, $\ln(CH_4/CO_2)$, $\ln(H_2/CO_2)$, $\ln(H_2S)$, $\ln(H_2S/H_2)$].

By considering the gas-phase compositions and the BHT_m measurements compiled in the NWGDB ($n_w = 265$), thirteen geothermal fields of the world (Berlin, Zunil, Krafla, Kamojang, Sibayak, Amiata, Larderello, Olkaria, Cerro Prieto, Las Tres Virgenes, Los Azufres, Los Humeros, and Palinpinon) were used for the geochemometric evaluation (see Additional file 1: Table S5). The geothermal reservoir temperatures estimated by applying the new gas geothermometers (GaS_GeoT) along with twenty-five existing geothermometers are reported in Additional file 1: Tables S6 and S7, respectively.

The prediction efficiency for all the gas geothermometers to determine reservoir temperatures was compared with measured BHT_m values using the evaluation metrics (DIFF%, RMSE, MAE, MAPE, and Theil's U). A summary of the prediction efficiency results is reported in Table 10. A full version of this statistical comparison is also reported in Additional file 1: Tables S8 to S13.

Results for wells located in LIQDR

The reservoir temperatures estimated for the gas phase compositions from LIQDR using the new gas geothermometers ($GasG_1$ to $GasG_8$) and twenty-five existing gas geothermometers were statistically compared with the corresponding BHTs, and the obtained results are presented as rounded-off values in the following sections:

Analysis of the DIFF% metric

The percent difference (DIFF%) calculated from the temperature estimates by all the thirty-three gas geothermometers (new and existing) and BHTs shows better efficiencies by the newly gas geothermometers ($GasG_1$ to $GasG_8$) with 91 to 96% of the predicted temperatures falling within the limits of acceptance (DIFF% $\pm$ 20%; Table 10). This behaviour is clearly observed when the temperature estimates are plotted against the measured BHTs (Fig. 5a–d). However, only four existing geothermometers from the literature (ND84c, AG85b, AG85d, and AG85f) predicted 85 to 88% of the temperatures in the same acceptable limits (Table 10 and Fig. 6a–d). About 55–63% of the estimated reservoir temperatures are overestimated (with DIFF% $\leq$ 20%), 30–34% of

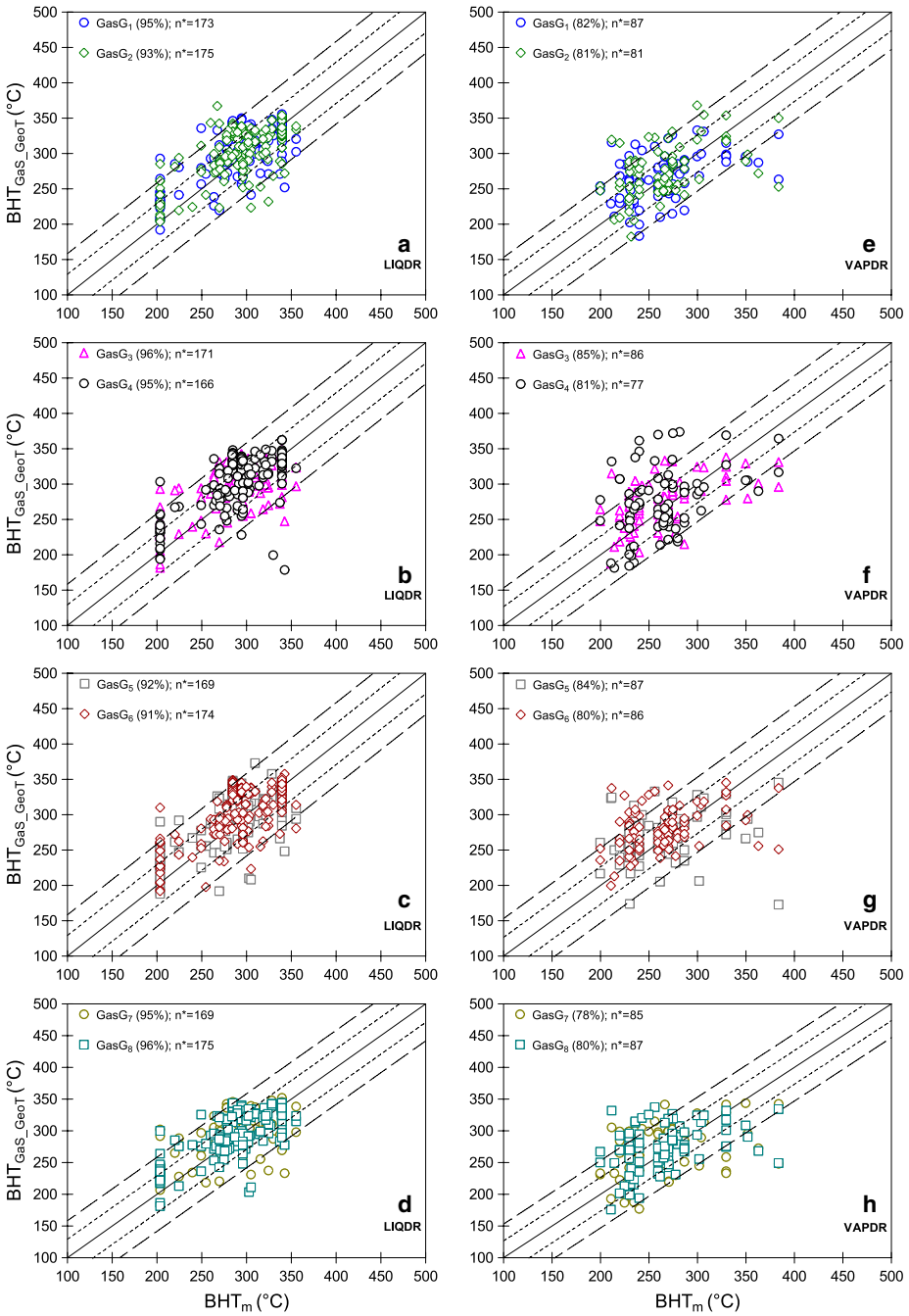

Fig. 5 Results of a comparative analysis between measured (BHT_m) and temperature estimates calculated for the new gas geothermometers (BHT_{GaS_GeoT}). The plots **a**–**d** show the temperature results predicted by the new geothermometers for LIQDR wells, whereas the plots **e**–**h**] show the results for VAPDR. n* is the number of samples that fulfil the applicability conditions of each new gas geothermometer (Table 8). Dotted and dashed lines represent a hypothetical variability of the ± 10% and ± 20% with respect to the BHT mean value, respectively

the temperatures are underestimated (with DIFF% ≤ 20%), and 5–12% are equal (with DIFF% ≤ 1%), by the eight new gas geothermometers ($GasG_1$ to $GasG_8$), when compared to the BHTs, respectively (Table 10).

Table 8 Applicability conditions defined for the eight new gas geothermometers developed

New gas geothermometers ($GasG_i$)	Gas concentrations (mmol/mol dry-basis)				Bottom-hole temperature °C
	CO_2	H_2S	CH_4	H_2	BHT_m
	Concentration intervals [Min, Max]				[Min, Max]
$GasG_1$, $GasG_2$, $GasG_6$, and $GasG_8$	[311, 3206]	[0.1032, 357]	[0.000059, 57]	[0.1084, 477]	[170, 374]
$GasG_3$, $GasG_4$, $GasG_5$ and $GasG_7$	[699, 994]	[0.1032, 197]	[0.000059, 57]	[0.1084, 128]	[170, 374]
	Input variable 1	Input variable 2	Input variable 3	Input variable 4	Input variable 5
$GasG_1$	[− 9.17, − 0.21]	[− 16.56, − 2.67]	[− 9.07, 0.43]	–	–
$GasG_2$	[2.67, 16.56]	[− 0.82, 13.95]	[− 2.99, 11.13]	–	–
$GasG_3$	[− 9.17, − 1.39]	[− 16.56, − 2.67]	[− 9.07, − 1.77]	–	–
$GasG_4$	[2.67, 16.56]	[− 0.82, 13.95]	[− 2.99, 11.13]	–	–
$GasG_5$	[1.77, 9.07]	[− 2.90, 6.14]	[− 11.13, 2.99]	–	–
$GasG_6$	[5.74, 8.07]	[− 2.27, 5.88]	[− 9.73, 4.04]	[− 2.22, 6.17]	–
$GasG_7$	[6.55, 6.90]	[− 2.27, 5.28]	[− 9.73, 4.04]	[− 2.22, 4.85]	–
$GasG_8$	[− 9.17, − 0.21]	[− 16.56, − 2.67]	[− 9.07, 0.43]	[− 2.27, 5.88]	[− 2.90, 6.14]

The input variables of each new gas geothermometer are reported in Table 6

Table 9 Statistical metrics used for evaluating the prediction efficiency of gas geothermometers (new and existing)

Statistical metric	Calculation equation	Equation number	Reference
Percent difference (DIFF%)	$DIFF\% = \left[\left(\frac{BHT_{CALC(i)} - BHT_{m(i)}}{BHT_{m(i)}}\right) \times 100\right]$	(6)	García-López et al. (2014)
Root mean square error (RMSE)	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(BHT_{m(i)} - BHT_{CALC(i)}\right)^2}$	(7)	Willmott et al. (2009)
Mean absolute error (MAE)	$MAE = \frac{1}{n}\sum_{i=1}^{n}\left\lvert BHT_{m(i)} - BHT_{CALC(i)}\right\rvert$	(8)	Wang and Lu (2018)
Mean absolute percentage error (MAPE)	$MAPE = \left(\frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{BHT_{m(i)} - BHT_{CALC(i)}}{BHT_{m(i)}}\right\rvert\right) \times 100$	(9)	Li and Shi (2010)
Difference coefficient (Theil's U)	$Theil'sU = \frac{\sqrt{\sum_{i=1}^{n}\left(BHT_{m(i)} - BHT_{GaS\_GeoT(i)}\right)^2}}{\sqrt{\sum_{i=1}^{n}\left(BHT_{m(i)} - BHT_{GasGeo\_Lit(i)}\right)^2}}$	(10)	Álvarez del Castillo et al. (2012)

$BHT_{m(i)}$ is the bottom-hole temperature measured in a geothermal well; $BHT_{CALC(i)}$ is the temperature calculated by any gas geothermometer; $BHT_{GaS\_GeoT(i)}$ is the temperature calculated by the new gas geothermometers ($GasG_1$ to $GasG_8$); $BHT_{GasGeo\_Lit(i)}$ is the temperature calculated by the existing gas geothermometers (Table 3), and n is the total number of gas samples

Arnórsson and Gunnlaugsson (1985) reported reservoir temperatures in some geothermal fields where the gas geothermometers may yield both under- and overestimates for reservoir temperatures. Powell (2000) suggested that temperature underestimates may be due to the addition of un-equilibrated biogenic methane, whereas Barragán et al. (2000) stated that an overestimation of temperatures may derive from an excess of gases in total discharge of the wells.

Based on the DIFF% metric, the resulting ranking between 1st and 12th positions for the gas geothermometers under evaluation is occupied by $GasG_8$, $GasG_3$, $GasG_1$, $GasG_7$, $GasG_4$, $GasG_2$, $GasG_5$, $GasG_6$, AG85d, AG85b, ND84c, and AG85f, respectively.

Analysis of RMSE, MAE, and MAPE metrics

The RMSE values of the geothermometers vary between 32 and 136 (see Additional file 1: Table S9). The lowest values of RMSE ranging from 32 to 35 were obtained for the eight new gas geothermometers: $GasG_1$ to $GasG_8$ (Table 10). This systematic behaviour of RMSE replicates the better efficiency provided by the new gas geothermometers in comparison to those existing geothermometers, in predicting reservoir temperatures comparable to BHT_m, which is clearly observed in Fig. 7a.

Moreover, all new gas geothermometers were characterized by the lowest values of MAE (ranging from 25 to 27), which also demonstrated that their temperature estimates have a better accuracy (Fig. 7b). With respect to the MAPE metric, small temperature differences between the temperature estimates and BHTs were also provided by the new gas geothermometers, which varied from 9 to 10 (Table 10). As MAPE represents an average value of absolute percentage errors, the lowest values calculated by the new gas geothermometers suggest a small better efficiency in predicting the reservoir temperatures, which is observed in Fig. 7c.

Based on the RMSE values, the resulting ranking between 1st and 12th positions for all the gas geothermometers was $GasG_3$, $GasG_8$, $GasG_5$, $GasG_6$, $GasG_7$, $GasG_1$, $GasG_2$,

Table 10 Summary of statistical results obtained from the geochemometric evaluation study of gas geothermometers

Evaluation metrics	Geothermal reservoir-type[a]	Gas Geothermometers (new and existing[b]); NWGDB $n_w = 265$ data (LIQDR: $g_1 = 178$; VAPDR: $g_2 = 87$)											
Percent difference		$GasG_1$	$GasG_2$	$GasG_3$	$GasG_4$	$GasG_5$	$GasG_6$	$GasG_7$	$GasG_8$	ND84c	AG85b	AG85d	AG85f
		Number of Samples (%—applicability)											
Over-estimation (DIFF% ≤ 20%)	LIQDR	103(60)	111(63)	99(58)	105(63)	93(55)	101(58)	97(58)	101(58)	70(39)	57(32)	55(31)	67(38)
	VAPDR	38(44)	42(52)	43(50)	37(48)	47(54)	51(59)	48(57)	42(48)	27(31)	62(71)	67(77)	69(79)
Under-estimation (DIFF% ≤ 20%)	LIQDR	56(32)	56(32)	52(30)	50(30)	57(34)	54(31)	53(31)	56(32)	94(53)	104(58)	117(66)	88(49)
	VAPDR	43(49)	32(39)	38(44)	36(47)	34(39)	29(34)	30(35)	39(45)	53(61)	20(23)	16(18)	13(15)
Equal to BHT (DIFF% ≤ 1%)	LIQDR	14(8)	8(5)	20(12)	11(7)	19(11)	19(11)	19(11)	18(10)	14(8)	17(10)	6(3)	23(13)
	VAPDR	6(7)	7(9)	5(6)	4(5)	6(7)	6(7)	7(8)	6(7)	7(8)	5(6)	4(5)	5(6)
Ranking[c]		1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th
DIFF% (± 20%)[d]	LIQDR	$GasG_8$	$GasG_3$	$GasG_1$	$GasG_7$	$GasG_4$	$GasG_2$	$GasG_5$	$GasG_6$	AG85d	AG85b	ND84c	AG85f
		168(96)	164(96)	164(95)	160(95)	157(95)	163(93)	156(92)	159(91)	156(88)	154(87)	152(85)	152(85)
	VAPDR	A98c	$GasG_3$	$GasG_5$	$GasG_1$	*ND84c*	$GasG_2$	$GasG_4$	$GasG_8$	ND84b	$GasG_6$	AG85f	$GasG_7$
		74(85)	73(85)	73(84)	71(82)	71(82)	66(82)	62(81)	70(81)	70(81)	69(80)	69(79)	66(78)
RMSE (°C)	LIQDR	$GasG_3$	$GasG_8$	$GasG_5$	$GasG_6$	$GasG_7$	$GasG_1$	$GasG_2$	$GasG_4$	AG85b	AG85d	AG85f	ND84c
		31.6	31.8	32.6	33.2	33.4	33.5	33.9	34.7	36.9	38.7	38.8	45.2
	VAPDR	$GasG_3$	$GasG_1$	ND84c	A98c	$GasG_2$	ND84b	AG85f	$GasG_6$	$GasG_8$	$GasG_7$	$GasG_5$	$GasG_4$
		35.6	38.0	38.9	39.8	39.8	40.1	41.0	41.9	42.1	43.4	44.8	49.0
MAE (°C)	LIQDR	$GasG_8$	$GasG_5$	$GasG_3$	$GasG_7$	$GasG_4$	$GasG_1$	$GasG_6$	$GasG_2$	AG85f	AG85b	AG85d	ND84c
		25.2	25.2	25.3	25.8	26.3	26.5	26.6	27.3	27.7	28.3	30.6	33.9
	VAPDR	$GasG_3$	$GasG_1$	$GasG_2$	$GasG_6$	$GasG_5$	A98c	$GasG_8$	ND84c	ND84b	$GasG_7$	AG85f	$GasG_4$
		27.6	29.3	30.2	30.4	31.3	31	32.1	32.7	33.8	33.8	34.0	39.2
MAPE (%)	LIQDR	$GasG_5$	$GasG_3$	$GasG_8$	$GasG_7$	$GasG_4$	$GasG_1$	$GasG_6$	$GasG_2$	AG85b	AG85f	AG85d	ND84c
		9.1	9.1	9.1	9.3	9.4	9.5	9.6	9.8	10.0	10.6	11.3	12.6
	VAPDR	$GasG_3$	$GasG_1$	$GasG_2$	$GasG_6$	$GasG_5$	A98c	$GasG_8$	ND84c	$GasG_7$	AG85g	ND84b	$GasG_4$
		10.6	11.2	11.5	11.8	11.9	11.9	12.4	12.7	12.8	12.9	13.1	15.4
Theil's U test	LIQDR	$GasG_3$	$GasG_8$	$GasG_5$	$GasG_7$	$GasG_6$	$GasG_1$	$GasG_4$	$GasG_2$	AG85b	AG85d	AG85f	ND84c
	VAPDR	$GasG_3$	$GasG_1$	ND84c	A98c	ND84b	AG85f	$GasG_2$	$GasG_6$	$GasG_8$	$GasG_7$	$GasG_5$	$GasG_4$

[a] Type of geothermal reservoir used to calibrate the gas geothermometers

[b] For the existing gas geothermometers (literature), we have selected only the best four geothermometers

[c] The full ranking can be checked in the Additional file 1: Tables S8 to S13

[d] Within the limits of acceptance. The new gas geothermometers were typed in italic font

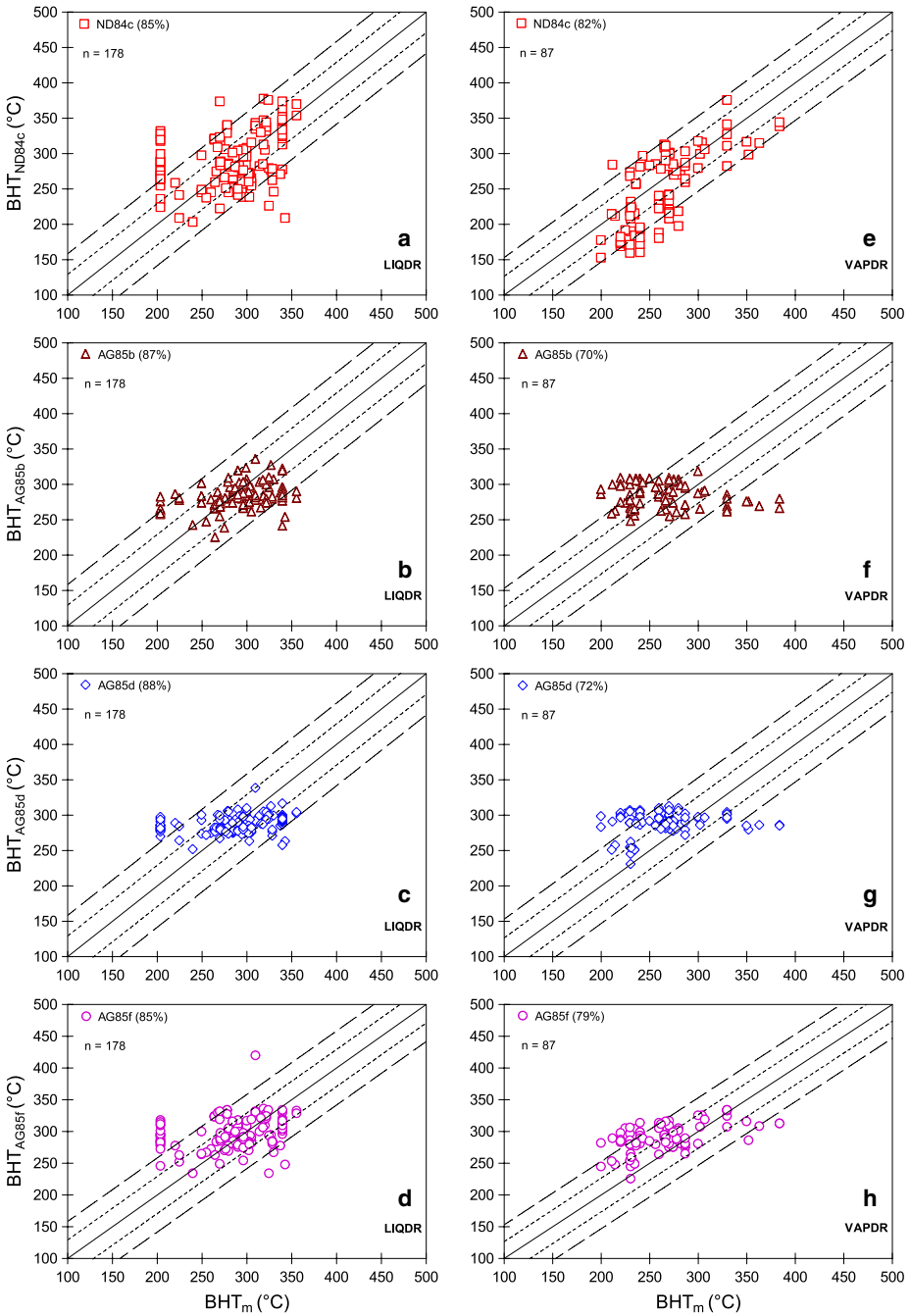

Fig. 6 Results of a comparative analysis between measured (BHT_m) and temperature estimates calculated by four existing gas geothermometers (BHT_{GasGeo_Lit}). The plots **a**–**d** show results of the ND84c, AG85b, Ag85d, and AG85f geothermometers obtained for LIQDR wells, whereas the plots **e**–**h** represent the results of the same geothermometers for VAPDR wells. Dotted and dashed lines represent a hypothetical variability of the ± 10% and ± 20% with respect to the BHT mean value, respectively

$GasG_4$, AG85b, AG85d, AG85f, and ND84c, respectively, whereas for the MAE and MAPE metrics, the first five positions corresponded to the new gas geothermometers with an order slightly different but consistent as the best geothermometric tools: $GasG_8$,

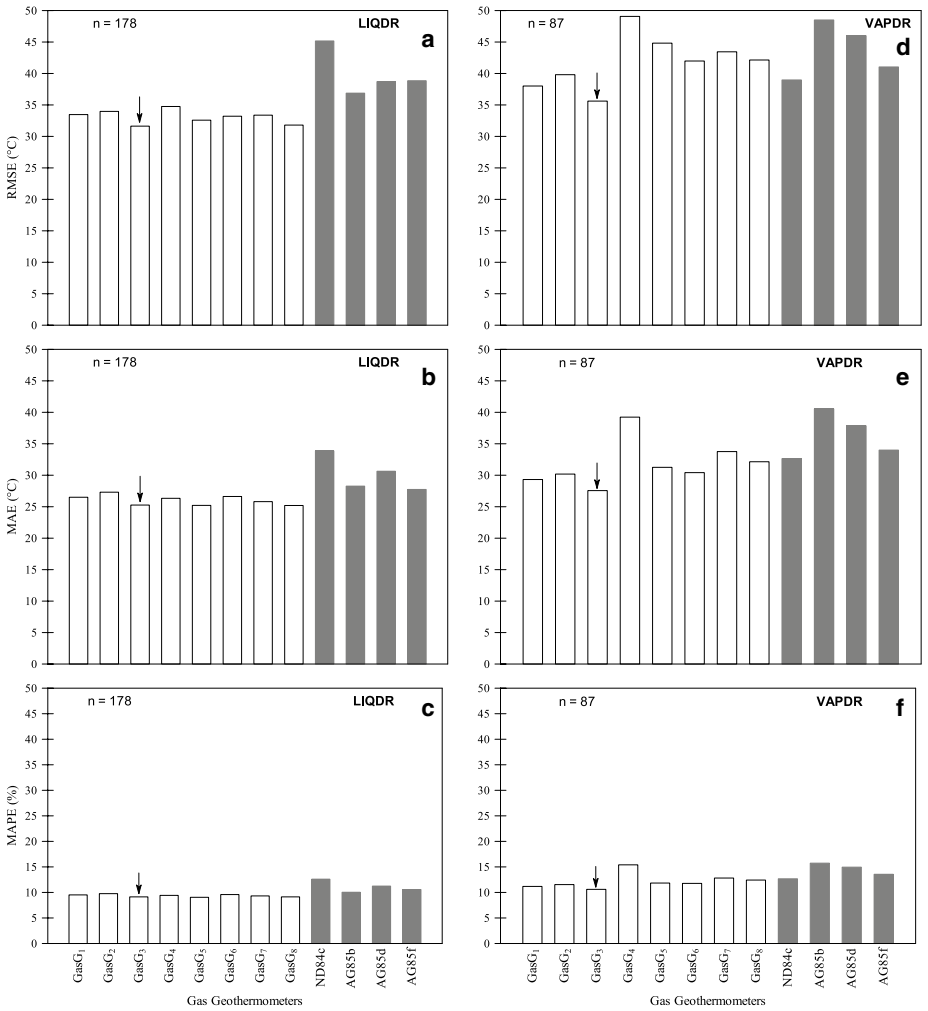

Fig. 7 Results of the statistical metrics used to evaluate the prediction efficiency of the eight new gas geothermometers against four existing geothermometers of the literature. The **a**–**c** plots show the residual results obtained for LIQDR wells (RMSE, MAE, and MAPE), whereas the **d**–**f** plots show the same residuals for VAPDR wells

$GasG_5$, $GasG_3$, $GasG_7$, and $GasG_4$; and $GasG_5$, $GasG_3$, $GasG_8$, $GasG_7$, and $GasG_4$, respectively (Table 10).

Analysis of the statistical Theil's U test (LIQDR)

After analysing the Theil's U results, the following inferences were accomplished: (1) the $GasG_3$ geothermometer may be considered with confidence as the best predictor model among all other geothermometers because the Theil´s U values were systematically lower than 1; (2) it was also observed that the new geothermometers ($GasG_1$ to $GasG_8$) systematically provided the lowest errors in comparison with twenty-five existing geothermometers (Table 10); and (3) after comparing the prediction efficiency among the new gas geothermometers, the $GasG_3$ exhibited Theil´s U lower values than 1 in comparison with those obtained for the remaining geothermometers ($GasG_1$–$GasG_2$ and $GasG_4$–$GasG_8$). A full description of the LIQDR results obtained for all the

gas geothermometers under evaluation using the statistical evaluation metrics (DIFF%, RMSE, MAE, MAPE, and Theil´s U) is reported in Additional file 1: Tables S8 to S10.

Results for wells located in VAPDR

The prediction efficiency of the reservoir temperatures calculated in VAPDR samples using the thirty-three gas geothermometers (new and existing) was also compared with the actual BHT measurements.

Analysis of the DIFF% metric

Based on the percent difference (DIFF%) values calculated by the new gas geothermometers, a good prediction efficiency was observed with 78 to 85% of the predicted temperatures falling within the limits of acceptance (DIFF% $\pm$ 20%), as it is observed in Fig. 5e–h. Out of the total twenty-five gas geothermometers, four geothermometers (ND84b, ND84c, AG85f, and A98c) from the literature have shown statistical differences comparable to these new geothermometers with 79 to 85% (Table 10, and Fig. 6e–h). About 44–59% of the estimated reservoir temperatures are overestimated (with DIFF% $\leq$ 20%), 34–49% of the temperatures are underestimated (with DIFF% $\leq$ 20%), and 5–9% are equal (with DIFF% $\leq$ 1%), by the eight new gas geothermometers ($GasG_1$ to $GasG_8$), when compared to the BHTs, respectively (Table 10).

According to these results, the new ranking between 1st and 12th positions for the gas geothermometers under evaluation was given by A98c, $GasG_3$, $GasG_5$, $GasG_1$, ND84c, $GasG_2$, $GasG_4$, $GasG_8$, ND84b, $GasG_6$, AG85f, and $GasG_7$, respectively. In this ranking, the prediction efficiency was actually very close between the geothermometers A98c and $GasG_3$, differing in only two decimals of the applicability percentage (i.e. 85.1% for A98c, and 84.9% for $GasG_3$; both percentages rounded-off as 85%). With these results, the gas geothermometer $GasG_3$ systematically shows a high prediction efficiency similar to those results obtained for LIQDR systems.

Analysis of the RMSE, MAE, and MAPE metrics

The RMSE values calculated in VAPDR samples for all geothermometers showed a wider variability between 36 and 181 in comparison with those values estimated for LIQDR samples (see Additional file 1: Table S12). For the new gas geothermometers ($GasG_1$ to $GasG_1$), seven out of the eight (except $GasG_4$) predict low values of RMSE ranging from 36 to 45 (Table 10 and Fig. 7d), whereas for the existing gas geothermometers, a wider interval of RMSE was calculated (from 39 to 181).

In relation to the MAE metric, seven out of the eight new geothermometers were characterized by the lowest values ranging from 28 to 34, which also demonstrated that the temperature estimates predicted by the seven new geothermometers (except $GasG_4$) are more accurate, as it is shown in Fig. 7e. Regarding the differences calculated in the MAPE values of the new gas geothermometers varied from 11 to 15 (Table 10 and Fig. 7f); which also suggest that seven out of eight geothermometers (except $GasG_4$) provided a slight better efficiency.

Based on the integrated calculations obtained for all the metrics (RMSE, MAE, and MAPE), it was found that seven out of the eight new gas geothermometers show a better prediction efficiency to determine the reservoir temperatures in VAPDR gas samples.

The resulting ranking between 1st and 12th positions according to RMSE was given by $GasG_3$, $GasG_1$ ND84c, A98c, $GasG_2$, ND84b, AG85f, $GasG_6$, $GasG_8$, $GasG_7$, $GasG_5$, and $GasG_4$, respectively, whereas for the MAE and MAPE metrics, the five positions corresponded to the new gas geothermometers: $GasG_3$, $GasG_1$, $GasG_2$, $GasG_6$, and $GasG_5$.

Analysis of the statistical Theil's U test (VAPDR)

After applying the Theil's U results (Table 10), two interpretations are inferred: (1) it was systematically observed that the $GasG_3$ predictor model (Theil's U values lower than 1) is the most reliable and accurate gas geothermometer over the rest of the new geothermometers ($GasG_1$–$GasG_2$ and $GasG_4$–$GasG_8$); and (2) it was found that two ($GasG_1$ and $GasG_3$) out of the eight new geothermometers exhibited lower errors in estimating reservoir temperatures in comparison with those estimates predicted by the existing geothermometers.

A full description of the results obtained from the use of the thirty-three gas geothermometers in VAPDR systems is reported in Additional file 1: Tables S11 to S13.

As a final remark of this research work, it was demonstrated, for the first time, the effectiveness of the MCDA optimization method for ranking and selecting the most reliable ANN architectures to predict a dependent variable (BHT_m) as a function of multiple independent variables (gas-phase compositions: CO_2, H_2S, CH_4, and H_2). We consider that this new evaluation proposal is actually innovating the evaluation methods commonly used in ANNs for evaluating their training, validation, and test stages, which typically rely on the simple use of the linear correlation coefficients (r) obtained between measured and simulated data.

With the optimal selection of the best predictor models used to correlate BHT_m and gas-phase compositions, eight new improved gas geothermometers ($GasG_1$ to $GasG_8$) were successfully developed. Most of these new geothermometers provide reliable estimations of geothermal reservoir temperatures. We have also demonstrated that the prediction efficiency of these new geothermometric tools (mainly $GasG_1$ and $GasG_3$ geothermometers) exceeds the efficiency of those existing gas geothermometers available in the literature.

Conclusions

Eight new improved gas geothermometers ($GasG_1$ to $GasG_8$) based on an optimized selection of artificial neural networks by using the MCDA method were successfully developed for the reliable prediction of the geothermal reservoir temperatures. For an effective and practical use of these geothermometers, a new computer program GaS_GeoT was successfully developed. The evaluation of the efficiency of the new improved gas geothermometers in predicting the reservoir temperatures was successfully demonstrated for geothermal wells of LIQDR and VAPDR systems.

The new gas geothermometers ($GasG_1$ to $GasG_8$) provided the best prediction efficiencies for geothermal wells from LIQDR, whereas two out the eight ($GasG_1$ and $GasG_3$) demonstrated their best efficiency in predicting reservoir temperatures for VAPDR. Among the new gas geothermometers, the best geothermometric tool for predicting reservoir temperatures in LIQDR and VAPDR systems is the $GasG_3$

geothermometer which uses the gas ratios: $\ln(H_2S/CO_2)$, $\ln(CH_4/CO_2)$, and $\ln(H_2/CO_2)$.

Out of the total twenty-five existing gas geothermometers, the most consistent geothermometers to predict reservoir temperatures comparable to those values inferred from the new geothermometers were as follows: (1) the ND84c, AG85b, AG85d, and AG85f geothermometers for LIQDR (which use CO_2-H_2S, CO_2-H_2, H_2, and H_2S concentrations); (2) the ND84b, ND84c, AG85f, and A98c geothermometers for VAPDR (which use CO_2-H_2, CO_2-H_2S, H_2S, and H_2S concentrations).

Taking in consideration the higher prediction efficiencies observed in predicting reservoir temperatures, the new gas geothermometers and the GaS_GeoT program may have the potential to become one of the most preferred geothermometric tools for the reliable estimation of the reservoir temperatures in geothermal prospection and exploitation.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40517-020-00182-9.

Additional file 1. The four statistical evaluation scenarios created with the MCDA methodology for ranking and selection of thirty-nine ANNs. Details of the NWGDB compiled from world-wide geothermal fields, and used for the development of new gas geothermometers are also included, together with deep temperatures determined by using gas geothermometers (new and existing). **Table S1.** MCDA Scenario 1 and ranking scores estimated for thirty-nine ANNs using statistical evaluation metrics. **Table S2.** MCDA Scenario 2 and ranking scores estimated for thirty-nine ANNs using the statistical evaluation metrics. **Table S3.** MCDA Scenario 3 and ranking scores estimated for thirty-nine ANNs using the statistical evaluation metrics. **Table S4.** MCDA Scenario 4 and ranking scores estimated for thirty-nine ANNs using the statistical evaluation metrics. **Table S5.** Chemical composition of gases from thirteen geothermal fields compiled from the literature: NWGDB ($n_w = 265$). **Table S6.** Temperature estimates predicted by the new gas geothermometers ($GasG_1$ to $GasG_8$) using the NWGDB ($n_w = 265$). **Table S7.** Temperature estimates predicted by twenty-five existing gas geothermometers using the gas-phase compositions compiled in NWGDB ($n_w = 265$). **Table S8.** Comparison among reservoir temperatures estimated by the new gas geothermometers and BHT_m of LIQDR systems ($g_1 = 178$). **Table S9.** Residual errors calculated for the thirty-three gas geothermometers using the gas compositions of LIQDR systems ($g_1 = 178$). **Table S10.** Statistical Theil's U results for evaluating the prediction efficiency of the gas geothermometers using LIQDR systems. **Table S11.** Comparison among reservoir temperatures estimated by the new gas geothermometers and BHTm of VAPDR systems ($g_2 = 87$). **Table S12.** Residual errors calculated for the thirty-three gas geothermometers using the gas compositions of VAPDR systems ($g_2 = 87$). **Table S13.** Statistical Theil's U results for evaluating the prediction efficiency of the gas geothermometers using VAPDR systems.

Abbreviations

LIQDR: Liquid-dominated reservoir; VAPDR: Vapour-dominated reservoir; ANN: Artificial neural networks; MCDA: Multi-criteria decision analysis; NWGDB: New Worldwide Geochemical Database; BHT_m: Bottom-hole temperatures (measured); BHT_{ANN}: Simulated ANN temperatures; WG_SubDB_1: Worldwide geochemical sub-database ($q_1 = 527$); WG_SubDB_2: Worldwide geochemical sub-databases ($q_2 = 498$); WG_SubDB_3: Worldwide geochemical sub-databases ($q_3 = 97$); LM: Levenberg–Marquardt algorithm; AC_1: Analysis case 1; AC_2: Analysis case 2; AC_3: Analysis case 3; MAVT: Multi-attribute value theory; S-1: MCDA scenario 1; S-2: MCDA scenario 2; S-3: MCDA scenario 3; S-4: MCDA scenario 4; BGF: Berlin geothermal field; ZGF: Zunil geothermal field; KGF: Krafla geothermal field; KaGF: Kamojang geothermal field; SGF: Sibayak geothermal field; AGF: Amiata geothermal field; LGF: Larderello geothermal field; OGF: Olkaria geothermal field; CPGF: Cerro Prieto geothermal field; LTVGF: Las Tres Virgenes geothermal field; LAGF: Los Azufres geothermal field; LHGF: Los Humeros geothermal field; PGF: Palinpinon geothermal field; DIFF%: Percent difference; RMSE: Root mean square error; MAE: Mean absolute error; MAPE: Mean absolute percentage error.

Acknowledgements

The authors also express their gratitude to E.O. García-Mandujano of IER-UNAM for her valuable comments and assistance in the literature compilation. The first author also wants to thank CONACyT for the scholarship awarded to carry out this research work as part of his PhD studies. The second author also wants to thank to the Physics and Mathematics Department of the Universidad Iberoamericana (Mexico) for the Professorship support in a sabbatical leave program carried out on 2018–19, time period where the present research work was started.

Authors' contributions

AAA contributed to design of ANNs and computer code programming; ES is the Leader of the research, and involved in design of computational methodology, fluid geochemistry, and geochemometrics; DPZ contributed to gas geothermometry analysis and unit conversions; KP is involved in gas geothermometry and mineral–gas assemblages;

MG created the geochemical databases; LDG calculated statistical metrics. All authors read and approved the final manuscript.

Funding
This work was partially funded by the P09 CeMIE-Geo research project (207032 CONACyT-SENER).

Availability of data and materials
The data used in this work are included in both the article and additional material. GaS_GeoT computer program was written in Java programming language, and is available for downloading from the following public server or repository (https://github.com/ANNGroup/GaS_GeoT.git), which may be directly accessed by the users. A User's Manual for running GaS_GeoT is also included, together with a README file, and the open-source code license (MIT). Hardware specifications to install and to run effectively the GaS_GeoT program are also described in the User's Manual.

Competing interests
The authors declare that they have no competing interests.

Author details
[1] Doctorado en Ciencias, Centro de Investigación en Ciencias-IICBA, Universidad Autónoma del Estado de Morelos, Av. Universidad 1001, Col. Chamilpa, 62209 Morelos, México. [2] Instituto de Energías Renovables, Universidad Nacional Autónoma de México, Priv. Xochicalco S/N, Temixco, 62580 Morelos, México. [3] CONACYT-Instituto de Geofísica, Universidad Nacional Autónoma de México, Circuito interior s/n, Coyoacán, 04510 Ciudad de México, México. [4] Centro de Investigación en Ciencias-IICBA, Universidad Autónoma del Estado de Morelos, Av. Universidad 1001, Col. Chamilpa, 62209 Morelos, México.

Received: 15 September 2020 Accepted: 26 November 2020
Published online: 18 January 2021

References

Abidin Z, Alip D, Nenneng L, Ristin PI, Fauzi A. Environmental isotopes of geothermal fluids in Sibayak geothermal field. In: Use of Isotope Techniques to Trace the Origin of Acidic Fluids in geothermal systems, IAEA Press, Austria. 2005;37–60. https://inis.iaea.org/search/search.aspx?orig_q=RN:36065710. Accessed 19 Nov 2020.

Adem EB, Geneletti D. Multi-criteria decision analysis for nature conservation: a review of 20 years of applications. Methods Ecol Evol. 2018;9:42–53.

ÁlvarezdelCastillo A, Santoyo E, García-Valladares O. A new empirical void fraction correlation inferred from artificial neural networks for modeling two phase flow in geothermal wells. Comput Geosci. 2012;41:25–39.

Arellano VM, García A, Barragán RM, Izquierdo G, Aragón A, Nieva D. An updated conceptual model of the Los Humeros geothermal reservoir (Mexico). J Volcanol Geotherm Res. 2003;124:67–88.

Arellano VM, Torres MA, Barragán RM. Thermodynamic evolution of the Los Azufres, Mexico, geothermal reservoir from 1982 to 2002. Geothermics. 2005;34:592–616.

Arellano VM, Barragán RM, Ramírez M, López S, Paredes A, Aragón A, Tovar R. The response to exploitation of the Los Humeros (México) geothermal reservoir. In Proceedings of the World Geothermal Congress 2015; Melbourne, Australia. 2015. https://www.geothermal-energy.org/pdf/IGAstandard/WGC/2015/14029.pdf. Accessed 19 Nov 2020

Arnórsson S, Gunnlaugsson E. New gas geothermometers for geothermal exploration—calibration and application. Geochim Cosmochim Acta. 1985;49:1307–25.

Arnórsson S. Gas chemistry of the Krísuvík geothermal field, Iceland, with special reference to evaluation of steam condensation in upflow zones. Jökull. 1987;37:31–48. https://jokulljournal.is/21-39/1987/031.pdf. Accessed 19 Nov 2020

Arnórsson S, Fridriksson T, Gunnarsson I. Gas chemistry of the Krafla geothermal field, Iceland. In: Intl Symp Water-Rock Interaction, Auckland, New Zealand. 1998;613–616. https://www.tib.eu/en/search/id/BLCP:CN023965846/Gas-chemistry-of-the-Krafla-Geothermal-Field-Iceland?cHash=ea876fde0d9fec307ca89a09bd0ce5ca. Accessed 20 Nov 2020

Arnórsson S, Bjarnason JÖ, Giroud N, Gunnarsson I, Stefánsson A. Sampling and analysis of geothermal fluids. Geofluids. 2006;6:203–16. https://doi.org/10.1111/j.1468-8123.2006.00147.x.

Arslan O, Yetik O. ANN based optimization of supercritical ORC-Binary geothermal power plant: Simav case study. Appl Therm Eng. 2011;31:3922–8.

Barragán RM, Arellano VM, Nieva D, Portugal E, García A, Aragón A, Tovar R, Torres-Alvarado I. Gas geochemistry of the Los Humeros geothermal field, México. In: Proceedings World Geothermal Congress 2000; Kyushu-Tohoku, Japan. 2000. https://www.geothermal-energy.org/pdf/IGAstandard/WGC/2000/R0130.PDF. Accessed 19 Nov 2020

Barragán RM, Gómez VA, Portugal E, Sandoval F, Segovia N. Gas geochemistry for the Los Azufres (Michoacán) geothermal reservoir, México. Ann Geophys. 2005;48:145–157. https://www.annalsofgeophysics.eu/index.php/annals/article/viewFile/3189/3234. Accessed 19 Nov 2020

Barragán RM, Arellano VM, Armenta M, Aguado R. Cambios químicos en fluidos de pozos del campo geotérmico de Los Humeros: Evidencia de recarga profunda. Geotermia. 2008;21:11–20. http://pubs.geothermal-library.org/lib/journals/Geotermia-Vol21-2.pdf. Accessed 20 Nov 2020

Barragán ReyesRM., Arellano Gómez VM, Mendoza A, Reyes L. Variación de la composición del vapor en pozos del campo geotérmico de Los Azufres, México, por efecto de la reinyección. Geotermia. 2012;25(1), 3–9. https://biblat.unam.mx/hevila/Geotermia/2012/vol25/no1/1.pdf. Accessed 20 Nov 2020

Barragán RM, Núñez J, Arellano VM, Nieva D. EQUILGAS: Program to estimate temperatures and in situ two-phase conditions in geothermal reservoirs using three combined FT-HSH gas equilibria models. Comput Geosci. 2016;88:1–8.

Bertrami R, Cioni R, Corazza E, D'Amore F, Marini L. Carbon monoxide in geothermal gases. Reservoir temperature calculations at Larderello (Italy). Geother Res Council Trans. 1985;9:299–303. http://pubs.geothermal-library.org/lib/grc/1001282.pdf. Accessed 20 Nov 2020

Blamey NJ. H_2S concentrations in geothermal and hydrothermal fluids—a new gas geothermometer. In: Proceedings thirty-first workshop on geothermal reservoir engineering, Stanford University. Stanford, California, USA. 2006;403–407. https://pangea.stanford.edu/ERE/pdf/IGAstandard/SGW/2006/blamey.pdf. Accessed 20 Nov 2020

Chamorro CR, Mondéjar ME, Ramos R, Segovia JJ, Martín MC, Villamañán MA. World geothermal power production status: Energy, environmental and economic study of high enthalpy technologies. Energy. 2012;42:10–8.

Chiodini G, Marini L. Hydrothermal gas equilibria: the H_2O-H_2-CO_2-CO-CH_4 system. Geochim Cosmochim Acta. 1998;62:2673–87.

D'Amore F, Panichi C. Evaluation of deep temperatures of hydrothermal systems by a new gas geothermometer. Geochim Cosmochim Acta. 1980;44:549–56.

D'Amore F, Ramos-Candelaria, MN, Seastres JrJ, Ruaya JR, Nuti S. Applications of gas chemistry in evaluating physical processes in the Southern Negros (Palinpinon) geothermal field, Philippines. Geothermics. 1993;22:535–553.

Díaz-González L, Santoyo E, Reyes J. Tres nuevos geotermómetros mejorados de Na/K usando herramientas computacionales y geoquimiométricas: aplicación a la predicción de temperaturas de sistemas geotérmicos. Rev Mex Cienc Geol. 2008;25:465–482. http://www.scielo.org.mx/pdf/rmcg/v25n3/v25n3a7.pdf. Accessed 20 Nov 2020

Dincer I, Acar C. A review on clean energy solutions for better sustainability. Int J Energy Res. 2015;39:585–606.

Ellis AJ. Chemical equilibrium in magmatic gases. Am J Sci. 1957;255:416–31.

Estévez RA, Alamos FH, Walshe T, Gelcich S. Accounting for uncertainty in value judgements when applying multi-attribute value theory. Environ Model Assess. 2018;23:87–97.

García-López CG, Pandarinath K, Santoyo E. Solute and gas geothermometry of geothermal wells: a geochemometrics study for evaluating the effectiveness of geothermometers to predict deep reservoir temperatures. Int Geol Rev. 2014;56:2015–49.

García-Mandujano EO. SYS_GASCHEM: Information system based on web technologies for the processing of geochemical databases and the determination of the temperature of geothermal systems. Universidad Politécnica del Estado de Morelos. 2019;137. http://www.cie.unam.mx/cgi-bin/CemieGeo/sys_gaschem.pl. Accessed 20 Nov 2020

Garson DG. Interpreting neural network connection weights. AI Expert. 1991;6:46–51.

Giggenbach WF. Geothermal gas equilibria. Geochim Cosmochim Acta. 1980;44:2021–32.

Giggenbach WF. Chemical techniques in geothermal exploration. In: D'Amore F. (Ed) Applications of geochemistry in geothermal reservoir development: series of technical guides on the use of geothermal energy, UNITAR/UNDP Centre Press, Rome. 1991;119–142. https://www.studocu.com/cl/document/universidad-catolica-del-norte/introduccion-a-la-geoquimica/otros/giggenbach-1991/5193236/view. Accessed 20 Nov 2020

Giggenbach WF, Glover RB. Tectonic regime and major processes governing the chemistry of water and gas discharges from the Rotorua geothermal field. New Zealand Geothermics. 1992;21:121–40.

Guo Q, Pang Z, Wang Y, Tian J. Fluid geochemistry and geothermometry applications of the Kangding high-temperature geothermal system in eastern Himalayas. Appl Geochem. 2017;81:63–75.

Gutiérrez-Negrín LC. Current status of geothermal-electric production in Mexico. Environ Earth Sci. 2019;249:1–11. https://doi.org/10.1088/1755-1315/249/1/012017/pdf.

Gutiérrez-Negrín LC, Canchola Félix I, Romo-Jones JM, Quijano-León JL. Geothermal energy in Mexico: update and perspectives. In Proceedings, Proceedings World Geothermal Congress 2020; Reykjavik, Iceland. 2020. https://www.researchgate.net/profile/Luis_Gutierrez-Negrin/publication/343111483_Geothermal_energy_in_Mexico_update_and_perspectives/links/5f17406292851cd5fa3a0275/Geothermal-energy-in-Mexico-update-and-perspectives.pdf. Accessed 20 Nov 2020

Henley RW, Truesdell AH, Barton PB, Whitney JA. Fluid-mineral equilibria in hydrothermal systems. Reviews in economic geology, vol 1, Society of Economic Geologists Inc. 1985;1:258. https://www.segweb.org/store_info/REV/REV-01-Additional-Product-Info.pdf. Accessed 20 Nov 2020

Kacandes GH, Grandstaff DE. Differences between geothermal and experimentally derived fluids: how well do hydrothermal experiments model the composition of geothermal reservoir fluids? Geochim Cosmochim Acta. 1989;53:343–58.

Karingithi CW, Arnórsson S, Grönvold K. Processes controlling aquifer fluid compositions in the Olkaria geothermal system. Kenya J Volcanol Geotherm Res. 2010;196:57–76.

Koga A, Kita I, Hikino E, Nitta K, Taguchi S. New gas geothermometers using CO_2/H_2 and CH_4/H_2 ratios. J Geother Res Soc of Japan. 1995;17:201–211. https://www.jstage.jst.go.jp/article/grsj1979/17/3/17_3_201/_pdf. Accessed 20 Nov 2020

Laksminingpuri N, Martinus A. Studi Kandungan Dan Temperatur Gas Panas Bumi Kamojang Dengan Diagram Grid. Beta Gamma Tahun. 2013;4:69–79. http://jurnal.batan.go.id/index.php/BetaGamma/article/download/1502/1431. Accessed 20 Nov 2020

Li G, Shi J. On comparing three artificial neural networks for wind speed forecasting. Appl Energy. 2010;87:2313–20.

López-Mendiola JM, Munguía F. Evidencias geoquímicas del fenómeno de ebullición en el campo de Los Humeros. Geotermia. 1989;5:89–106. https://colecciondigital.cemiegeo.org/xmlui/handle/123456789/2062. Accessed 20 Nov 2020

Minissale A, Evans WC, Magro G, Vaselli O. Multiple source components in gas manifestations from north-central Italy. Chem Geol. 1997;142:175–92.

Moya D, Aldás C, Kaparaju P. Geothermal energy: Power plant technology and direct heat applications. Renew Sustain Energy Rev. 2018;94:889–901.

Nehring N, D'Amore F. Gas chemistry and thermometry of the Cerro Prieto, Mexico, geothermal field. Geothermics. 1984;13:75–89.

Nicholson K. Geothermal fluids: Chemistry and Exploration Techniques. Berlín, Alemania, Springer-Verlag. 1993;263. https://www.springer.com/gp/book/9783642778469. Accessed 20 Nov 2020
Nieva D, Fausto J, González J, Garibaldi F. Flow of vapor into the production zone of Cerro Prieto I wells. In: Proceedings Fourth Symposium on the Cerro Prieto Geothermal Field, Baja California, México. 1982;2:455–461. https://www.osti.gov/servlets/purl/7369515. Accessed 20 Nov 2020
Nieva D, Gonzales J, Garfias A. Evidence of two extreme flow regimes operating in the production zone of different wells from Los Azufres. In: Proceedings Tenth Workshop on Geothermal Reservoir Engineering, Stanford University. Stanford, California, USA. 1985;233–240. https://www.osti.gov/servlets/purl/892539. Accessed 20 Nov 2020
Nieva D, Verma M, Santoyo E, Barragan RM, Portugal E. Chemical and isotopic evidence of steam upflow and partial condensation in Los Azufres reservoir. In: Proceedings Twelfth Workshop on Geothermal Reservoir Engineering, Stanford University. Stanford, California, USA. 1987;253–260. https://www.osti.gov/servlets/purl/888553/. Accessed 20 Nov 2020
Nieva D, Barragán RM, Arellano V. Geochemistry of Hydrothermal Systems. In: Bronicki L, editor. Power stations using locally available energy sources. Encyclopedia of sustainability science and technology series. New York: Springer; 2018.
Pandarinath K, Pérez-Barrera J, Pérez-Orozco JP. GasGeo—software to estimate the reservoir temperatures of geothermal systems using gas geothermometers: In: Proceedings XXI National Geochemistry Congress 2011, Actas INAGEQ 17. 2011.
Pang Z. Isotope and chemical geothermometry and its applications. Sci China Technol Sci. 2001;44:6–20. https://doi.org/10.1007/BF02916784.pdf.
Pérez-Zárate D, Santoyo E, Acevedo-Anicasio A, Díaz-González L, García-López C. Evaluation of artificial neural networks for the prediction of deep reservoir temperatures using the gas-phase composition of geothermal fluids. Comput Geosci. 2019;129:49–68.
Porkhial S, Salehpour M, Ashraf H, Jamali A. Modeling and prediction of geothermal reservoir temperature behavior using evolutionary design of neural networks. Geothermics. 2015;53:320–7.
Poulton MM. Computational Neural Networks for Geophysical Data Processing. Pergamon Press, Amsterdam. 2001;30:335. https://www.elsevier.com/books/computational-neural-networks-for-geophysical-data-processing/poulton/978-0-08-043986-0. Accessed 20 Nov 2020
Powell T. A review of exploration gas geothermometry. In: Proceedings 25th Workshop on Geothermal Reservoir Engineering, Stanford University. Stanford, California, USA. 2000;9. https://pangea.stanford.edu/ERE/pdf/IGAstandard/SGW/2000/Powell.pdf. Accessed 20 Nov 2020
Powell T, Cumming W. Spreadsheets for geothermal water and gas geochemistry. In: Proceedings 35th Workshop on Geothermal Reservoir Engineering. Stanford University. Stanford, California, USA. 2010;10. https://pangea.stanford.edu/ERE/pdf/IGAstandard/SGW/2010/powell.pdf. Accessed 20 Nov 2020
Renderos RE. Chemical characterization of the thermal fluid discharge from well production tests in the Berlin geothermal field, El Salvador: geothermal training program. The United Nations University. 2002;2:205–231. https://orkustofnun.is/gogn/unu-gtp-report/UNU-GTP-2002-12.pdf. Accessed 20 Nov 2020
Santoyo E, Verma SP, Nieva D, Portugal E. Variability in the gas phase composition of fluids discharged from Los Azufres geothermal field. Mexico J Volcanol Geotherm Res. 1991;47:161–81.
Santoyo-Castelazo E. Sustainability assessment of electricity options for Mexico: current situation and future scenarios. Ph. D. Dissertation, The University of Manchester, United Kingdom, 2011:286. https://www.research.manchester.ac.uk/portal/files/54515414/FULL_TEXT.PDF. Accessed 20 Nov 2020
Santoyo-Castelazo E, Gujba H, Azapagic A. Life cycle assessment of electricity generation in Mexico. Energy. 2011;36:1488–99.
Santoyo-Castelazo E, Azapagic A. Sustainability assessment of energy systems: integrating environmental, economic and social aspects. J Clean Prod. 2014;80:119–38.
Saracco L, D'Amore F. CO2B: a computer program for applying a gas geothermometer to geothermal systems. Comput Geosci. 1989;15:1053–65.
Serpen G, Palabiyik Y, Serpen U. An artificial neural network model for Na/K geothermometer. In: Proceedings 34th Workshop on Geothermal Reservoir Engineering, Stanford University. Stanford, California, USA. 2009;12. https://pangea.stanford.edu/ERE/pdf/IGAstandard/SGW/2009/serpen.pdf. Accessed 20 Nov 2020
Spycher N, Peiffer L, Finsterle S, Sonnenthal E. GeoT User's Guide, A Computer Program for Multicomponent Geothermometry and Geochemical Speciation, Version 2.1. 2016. https://escholarship.org/content/qt8hs3b99h/qt8hs3b99h.pdf. Accessed 20 Nov 2020
Stefánsson A. Gas chemistry of Icelandic thermal fluids. J Volcanol Geotherm Res. 2017;346:81–94.
Supranto S, Budianto T, Djoko W, Idrus A. Proposed empirical gas geothermometer using multidimensional approach. In: proceedings twenty first workshop on geothermal reservoir engineering, Stanford University. Stanford, California, USA. 1996;195–199. https://pangea.stanford.edu/ERE/pdf/IGAstandard/SGW/1996/Supranto.pdf. Accessed 20 Nov 2020
Tello E, Verma MP, Tovar R. Origin of acidity in the Los Humeros, Mexico, geothermal reservoir. In: Proceedings World Geothermal Congress 2000; Japan. 2000. https://www.geothermal-energy.org/pdf/IGAstandard/WGC/2000/R0081.PDF. Accessed 20 Nov 2020
Theil H. Economic Forecasts and Policy. North-Holland Press. 1961;567.
Thien BM, Kosakowski G, Kulik DA. Differential alteration of basaltic lava flows and hyaloclastites in Icelandic hydrothermal systems. Geotherm Energy. 2015;3:1–32.
Tonani F. Equilibria that control the hydrogen content of geothermal gases. Bull Volcanol. 1973;44:547–64.
Verma SP, Santoyo E. New improved equations for Na/K, Na/Li and SiO2 geothermometers by outlier detection and rejection. J Volcanol Geotherm Res. 1997;70:9–23.
Verma SP, Pandarinath K, Santoyo E, González-Partida E, Torres-Alvarado IS, Tello-Hinojosa E. Fluid chemistry and temperatures prior to exploitation at the Las Tres Vírgenes geothermal field. Mexico Geothermics. 2006;35:156–80.

Verma SP, Pandarinath K, Santoyo E. SolGeo: A new computer program for solute geothermometers and its application to Mexican geothermal fields. Geothermics. 2008;37:597–621.
Wamalwa RN. Evaluation of factors controlling the concentration of non-condensible gases and their possible impact on the performance of wells in Olkaria, Kenya. Geothermal Training Programme, United Nations University, Reykjavik, Iceland. 2015;787–808. https://orkustofnun.is/gogn/unu-gtp-report/UNU-GTP-2015-34.pdf. Accessed 20 Nov 2020
Wang JJ, Jing YY, Zhang CF, Zhao JH. Review on multi-criteria decision analysis aid in sustainable energy decision-making. Renew Sustain Energy Rev. 2009;13:2263–78.
Wang W, Lu Y. Analysis of the mean absolute error (MAE) and the root mean square error (RMSE) in assessing rounding model. Mater Sci Eng. 2018;324:1–10. https://doi.org/10.1088/1757-899X/324/1/012049/pdf.
Willmott CJ, Matsuura K, Robeson SM. Ambiguities inherent in sums-of-squares-based error statistics. Atmos Environ. 2009;43:749–52.
Wu Y, Li P. The potential of coupled carbon storage and geothermal extraction in a CO_2 enhanced geothermal system: a review. Geotherm Energy. 2020;8:1–28.
Yan-guang L, Bing L, Chuan L, Xi Z, Gui-ling W. Reconstruction of deep fluid chemical constituents for estimation of geothermal reservoir temperature using chemical geothermometers. J Groundw Sci Eng. 2017;5:173–181. http://gwse.iheg.org.cn/article/id/271. Accessed 20 Nov 2020

Publisher's Note