ORIGINAL PAPER

Testing solution models for phase equilibrium (forward) modeling of partial melting experiments

Félix Gervais[1] · Pierre-Henri Trapy[2]

Abstract

This study compares four sets of solution (or activity-composition) models using two internally consistent thermodynamic datasets for calculating isochemical phase diagram sections of six partial melting experiments covering a wide range of metasedimentary bulk compositions. Compared parameters are: (1) biotite-breakdown temperatures; (2) parageneses of Fe–Mg phases and Ti-oxides; (3) proportion of biotite, garnet, liquid and plagioclase; (4) Mg# of biotite, garnet and liquid and the anorthite content of plagioclase. Results reveal significant differences between sets that are mainly related to the construction of the biotite solution, with the model of Tajčmanová et al. (2009) yielding better results for the majority of investigated parameters. Owing to the success of the investigated set of solution models at reproducing the proxy that constitutes partial melting experiments, but also at reproducing natural observations as published elsewhere, it is suggested that this biotite model should be used in the Ti–Mn–Na–Ca–K–Fe–Mg–Al–Si–H–O system for phase equilibrium (forward) modeling of metasediments.

Introduction

Phase equilibrium, forward modeling (PEFM) ("forward" to distinguish it from the multiequilbrium approach that is an inverse modeling; see Powell and Holland 2008) has become an essential tool to unravel the evolution of metamorphic rocks (see reviews by Holland and Powell 2011 and Lanari and Duesterhoeft 2019). It has been used to interpret the evolution of migmatites (Indares et al. 2008; White and Powell 2002, White et al. 2005; Yakymchuk and Brown 2014); to reconstruct Pressure–Temperature (P–T) paths that were then linked to U-Th–Pb geochronology of accessory phases to reconstruct P–T-time paths used to test regional-scale

Communicated by Mark S Ghiorso.

✉ Félix Gervais
   felix.gervais@polymtl.ca

1 Civil, Geological and Mining Engineering Department, Polytechnique Montréal, Montréal, Canada

2 Rueil-Malmaison, France

tectonic models (Lang and Gilotti 2015; Larson et al. 2013; Dumond et al. 2015; Gervais and Crowley 2017); to predict directly accessory phases petrogenesis by incorporating their thermodynamics properties into the calculation (Spear and Pyle 2010; Spear 2010). One of main strengths of the method is that it allows a direct comparison between the observed minerals modes and composition to those predicted by equilibrium thermodynamics for a specific bulk composition and a recent improvement further allows to quantify this comparison (Duesterhoeft and Lanari 2020). Currently, three main softwares are used for such calculations: Thermocalc (Powell et al. 1998); Theriak-Domino (de Capitani and Petrakakis 2010) and Perple_X (Connolly and Petrini 2002; Connolly 2005). Out of the three softwares, the latter two allow the user to choose the desired solution models.

Although solution models (or activity-composition relations) have been identified as a major source of errors (White et al. 2011) there is generally no justification provided for the choice of a set of solution models (apart from the fact that they use internally consistent database). A few studies have evaluated a given thermodynamic database (see Lanari and Duesterhoeft 2019 for a discussion on what constitutes a database) against partial melting experiments with results broadly consistent with observations (Johnson et al. 2008; Grant 2009;

White et al. 2011; García-Arias 2020), but few have compared solution models or complete database to each other (except for Tajčmanová et al. 2009 and Tropper et al. 2002). With the release of the thermodynamic database by Thermocalc composed of a revised internally consistent dataset (Holland and Powell 2011) and a set of compatible solution (activity-composition) models (White et al. 2014) the hopes were high that PEFM would achieve even better results than before, but results of two studies rather suggested that the older database (dataset tcds55 with associated solution models) reproduces natural observations, while the newer database does not (Guevara and Caddick 2016; Kendrick and Indares 2018), and one other study indicated that the new database reproduced well the paragenesis, but not the composition of clinopyroxene and amphibolite in several mafic granulites (Forshaw et al. 2019). In the absence of rigorous test, users cannot justify their choice of thermodynamic datasets and solution models to conduct PEFM. It is, therefore, crucial to evaluate different sets of solution models and thermodynamic datasets available and test whether some perform better than others for modeling different bulk compositions. Although we are aware of the advantages and pitfalls of PEFM and partial melting experiments (see White et al. 2011 and Sect. 2.3 below), we consider that comparing the two methods is one of the best way to conduct our test, another alternative would be to use the Bingo-Antidote software described in Duesterhoeft and Lanari (2020).

This study compares four sets of solution models aiming at reproducing six partial melting experiments of metapelite and aluminous greywacke by PEFM using four different sets of solutions models and two internally consistent thermodynamic datasets. Parameters investigated include: (1) proportion of major phases; (2) proportion of titanium oxides; (3) composition of major phases. Our results reveal significant discrepancies between the sets and identify two that outperforms the others.

Methodology

Experiments

Six partial melting experiments were chosen (Fig. 1; Table 1) such that on AFM, AFC, AKF they cover most of the range of composition compiled on the MetPetDB database (Spear et al. 2009) for metapelitic and greywacke compositions. Experiments HQ36 (Patiño-Douce and Johnston 1991) and HP60 (Pickering and Johnston 1998) are close to the most common bulk composition of metapelites in the database, experiments MS (Patiño-Douce and Harris 1998) and NBS (Stevens et al. 1997) represent aluminous metapelites, whereas experiments NB (Stevens et al. 1997) and CEPV (Montel and Vielzeuf 1997; Vielzeuf and Montel 1994) represents metagreywackes (Fig. 1). All experiments contained MnO except HP60 and all experiment contained between 2.3 and 2.8 wt% TiO_2, except MS that contained 3.23 wt% TiO_2 (Table 1). Experiment NBS consisted of 13 runs from 750 to 1000 °C at 500 and from 800 to 1000 °C at 1000 MPa; experiment MS consisted of 11 runs from
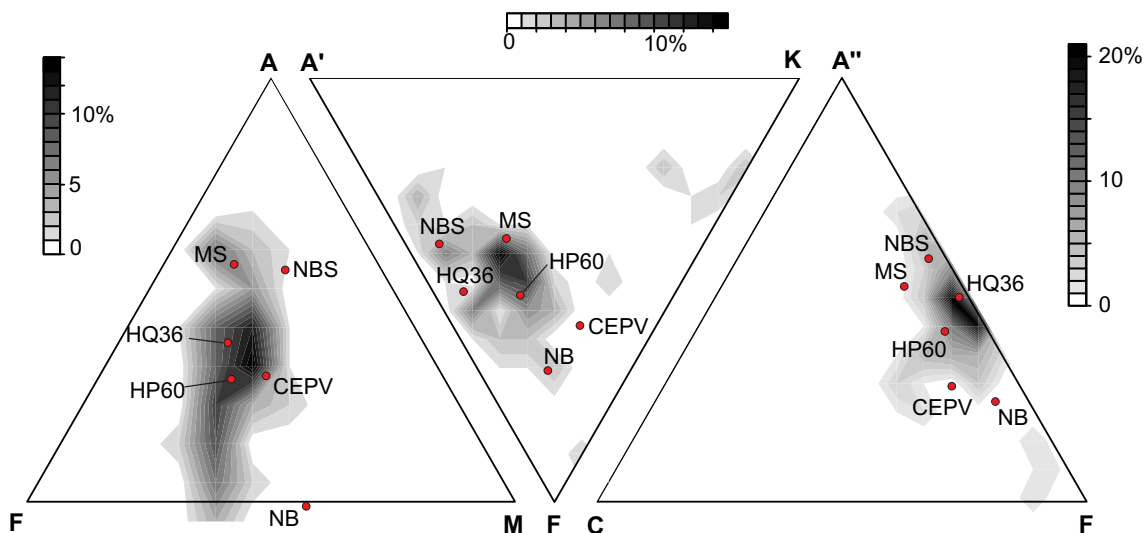

Fig. 1 Bulk composition of starting material for the six fluid-absent partial melting experiments investigated in this contribution: HQ36 (Patiño-Douce and Johnston, 1991); NBS and NB (Stevens et al., 1997); MS (Patiño-Douce and Harris, 1998); HP60 (Pickering and Johnston, 1998); CEPV (Vielzeuf and Montel, 1994; Montel and Vielzeuf, 1997). Density plots were constructed with GCDkit 3.0 from the composition of 129 samples extracted from the MetPetDB database (Spear et al., 2009) with the query: protolith types = metapelite or metagreywacke. AFM: A = Al_2O_3–K_2O, F = FeO, M = MgO. A'KF: A' = Al_2O_3 + Fe_2O_3–(K_2O + Na_2O + CaO),   K = K_2O, F = FeO + MnO + MgO. A''CF: A'' = Al_2O_3 + Fe_2O_3–(K_2O + Na_2O), C = CaO; F = FeO + MnO + MgO. See text for details

Table 1 Key chemical chracteristics of the six partial melting experiments investigated herein along with the biotite-breakdown temperatures

		Bulk composition			Biotite composition		
	Experiment	Mn wt%	Mg#	O fugacity	TiO$_2$ wt%	F wt%	Bt-out ($^{\circ}$C)
Metapelites	HP60	–	39	Below or at QFM	2.81	0.28	>950
	MS	0.13	33	QFM -QFM2	3.23	0.2	>900
	HQ36	0.17	36	QFM	2.68	0.31	975–1000
	NBS	0.01	57	NNO-QFM	2.29	<0.01	900–950
Metagreywackes	NB	0.01	58	NNO-QFM	2.29	<0.01	900–950
	CEPV	0.06	49	0.19 wt% Fe$_2$O$_3$	2.80	n.a	980–1000

750 to 900 °C at 600 and from 800 to 900 °C at 1000 MPa; experiment HQ36 consisted of 19 runs from 825 to 1075 °C at 700 and 1000 MPa; experiment HP60 consisted of 6 runs from 812 to 950 °C at 1000 and from 800 to 1000 °C at 1000 MPa; experiment CEPV consisted of 24 runs from 805 to 875 °C at 300, from 809 to 898 °C at 500 MPa, from 855 to 1040 °C at 800 MPa, from 855 to 1040 °C at 800 MPa and from 803 to 1000 °C at 1000 MPa; experiment NB consisted of 13 runs from 800 to 1000 °C at 500 and from 800 to 1000 °C at 1000 MPa. Out of the six experiments, only CEPV was used to calibrate a solution model and it was used to calibrate the three Ti-biotite models investigated herein. The six studies did not report the same data, which limited possible comparisons as described in the result section.

Phase-equilibria (forward) modeling

PEFM was conducted with the program Perple_X (version 6.7.3) in the Ti–Mn-Na–Ca-K-Fe–Mg–Al–Si–H-O system. Oxygen fugacity was incorporated in the calculations either as a buffer by creating a new thermodynamic entity (qfm=2 mt+3 q-3 fa) in the datasets and fixing O$_2$ as a saturated component or as the thermodynamic component O$_2$ for the CEPV experiment that reported the Fe$_2$O$_3$ composition. The first set investigated, referred herein as VAR, consists of various solution models identified by the authors as performing best by trial and errors. It uses the same models as HP04 for melt, staurolite, spinel, cordierite and orthopyroxene, but includes the ternary feldspar of Fuhrman and Lindsley (1988), the Mn-bearing garnet model of White et al. (2005), the white mica model of Smye et al., (2010), the ideal ilmenite model slightly modified to include W^{ilm,pnt} = 2200 (Wu and Zhao 2006), and the biotite model of Tajčmanová et al. (2009). The second set investigated, referred herein as set TCHP04, also includes the biotite model of Tajčmanová et al. (2009) and Mn in garnet (without interaction parameters for spessartine), but all the other solution models are the same as the third set. This third set, referred herein as set HP04, consists of the solution models provided by White et al. (2007), which do not include MnO. The fourth set

investigated, referred herein as set HP11, consist of the solution models provided by White et al. (2014). The first three sets use the internally consistent thermodynamic dataset of Powell et al. (1998; updated in 2004 as tcds55), whereas the last set uses the tcds62 dataset of Holland and Powell (2011).

Strategy of investigation

Experiments and PEFM both have advantages and inconvenient. It is difficult to test solution models by modeling natural rocks, although the Bingo-Antidote software (Lanari and Duesterhoeft 2019; Duesterhoeft and Lanari 2020) now provides a more robust way to test the quality of a given calculation. In theory, experiments should provide "answers" toward which PEFM should tend. In practice, however, there will never be a perfect fit because both methods have pitfalls and uncertainties that are difficult to quantify. White et al. (2011) and García-Arias, (2020) provided an extensive description of the various sources of errors and pitfalls of each methods and we will only mention some of them here. For experiments, attainment of equilibrium and the oxygen fugacity in the capsule are two important factors that are difficult to ascertain (see Douce and Beard 1994). For PEFM, some important chemical components are omitted. For example, the fluor content of biotite significantly controls its stability, but cannot be modeled. There are also significant uncertainties in the thermodynamic formulation of each solution model (see Lanari and Duesterhoeft 2019 for an extensive discussion) that combines in PEFM, rendering quantification of uncertainties very difficult. Nevertheless, our working hypothesis is that well-formulated sets of solution models should systematically yield results closer to that of experiments.

Four sets of solid solutions were chosen for this study (see Table 2 for a complete description). The main research group on PEFM is responsible for the development of the *Thermocalc* software and the production of internally consistent thermodynamic datasets (now hosted at http://hpxeosandthermocalc.org/). Thermocalc proposes different sets of solution models depending on the bulk composition investigated.

Table 2 Solution models used for phase equilibria (forward) modeling

	Garnet	Biotite	feldspar	Ilmenite	Cordierite	Liq, Opx
HP11 + Mn	W	W	W	W	W	W
HP04	WPH	WPH	WPH	WPH	ideal	WPH
TCHP04 + Mn	WPH	TCC	WPH	WPH	ideal	WPH
VAR + Mn	WPPH	TCC	feldspar (F&L)	ideal $W^{ilm,pnt} = 2200$	ideal	WPH

W: White et al. JMG 2014

WPH: Thermocalc 2015

WPPH: White et al., 2005

TCC: Tajčmanová et al. 2009

F&L: Fuhrman and Lindsley 1988

For metasediments, the proposed sets are HP04 and HP11 that accompany the older and newer versions of the internally consistent thermodynamic datasets tcds55 (Powell et al. (1998; updated in 2004) and tcds62 (Holland and Powell 2011), respectively. Note that solution models developed for the tcds62 dataset cannot be used with the tcds55 dataset, and vice versa. Because our informal work with the Ti-biotite model of Tajčmanová et al. (2009) appeared to significantly influence results of PEFM, we decided to specifically test this model. For this sake, we designed a set similar to HP04 (TCHP04), but replacing the Ti-biotite model of White et al. (2007) with that of Tajčmanová et al. (2009). The VAR set includes various solution models, including the Tajčmanová et al. (2009) biotite model, that appear to yield better results based on our informal experience and success at modeling natural rocks (Larson et al. 2013; Gervais and Crowley 2017; Perrot et al. 2020; Soucy La Roche et al. 2019). Although Thermocalc suggests excluding Mn in PEFM with the tcds55 dataset, we decided to include it in the TCHP04 set. We tried a set not including Mn, but it yielded redundant results. We, therefore decided to include Mn for a more robust comparison between sets VAR, TCHP04 and HP11 in a system as close as possible to natural compositions (Ti–Mn-Ca–Na–K-Fe–Mg–Al–Si–H-O). Finally, it is worth mentioning that the large majority of data presented herein were acquired before the publication of the HGP database (Holland et al. 2018). Nevertheless, we consider that it would not change significantly our results because this database yielded very similar results compared the HP11 sets for average metapelitic composition (Fig. 12 in Holland et al. 2018), a conclusion that we have confirmed for the MS bulk composition by PEFM (results not presented in this contribution).

Several parameters are used herein to compare the four sets of solution models. We first calculated isochemical phase diagram sections (IPDS) for all six partial melting experiments using the four different sets of solution models, producing a total of 24 IPDS. IPDS are difficult to compare (see Figs. 2 and 3), hence we extracted data from them for each set for all experimental runs. Paragenesis, phase proportions and chemical compositions were, therefore, extracted from a total of 344 pressure–temperature data points and compared with results yielded by partial melting experiments of metasediments covering the entire field of natural compositions (Fig. 1) and P–T conditions varying between 500 and 1000 MPa and 750 to 1050 °C, depending of experimental conditions (see Sect. 2.1 above and supplementary files). The temperature at which biotite disappears is of uttermost importance for testing the validity of solution models because its breakdown is the main melting reaction in metasediments (see reviews by Clemens 2006 and Brown 2010) and it influences the stability of other ferromagnesian minerals. Figure 4 compares experimental and calculated biotite-breakdown temperatures. Parageneses of ferromagnesian minerals is another very important parameter because their proportion and chemical composition are regularly used for isopleth thermobarometry. Figure 5a presents the number of times a given set has reproduced the Fe–Mg experimental paragenesis and Fig. 5b shows the number of times one of the Fe–Mg mineral (we included muscovite in this list) caused the failure of the calculation (either as predicted to be there, while it was not observed or vice versa). Another important test for the effectiveness of the different sets of solution models is the comparison between predicted and experimental proportion for major phases. Proportions and Mg# of liquid, garnet, biotite are reported in the six investigated experiments (NB, NBS, MS, HQ36, HP60 and CEPV), whereas proportion and the anorthite content of plagioclase are reported in two (MS, CEPV). Only data points reproducing the experimental paragenesis of the main phases are included. Sets that best reproduced the experimental proportion and composition for a given data point is shown as larger symbol on Figs. 6 and 7, respectively, and the number of times this occurred for each set and each phase is shown on Fig. 8. Boxplots of the absolute differences between calculated and experimental proportions abs(calc.–exp) and composition are presented for each set on Fig. 9. Another way of comparing
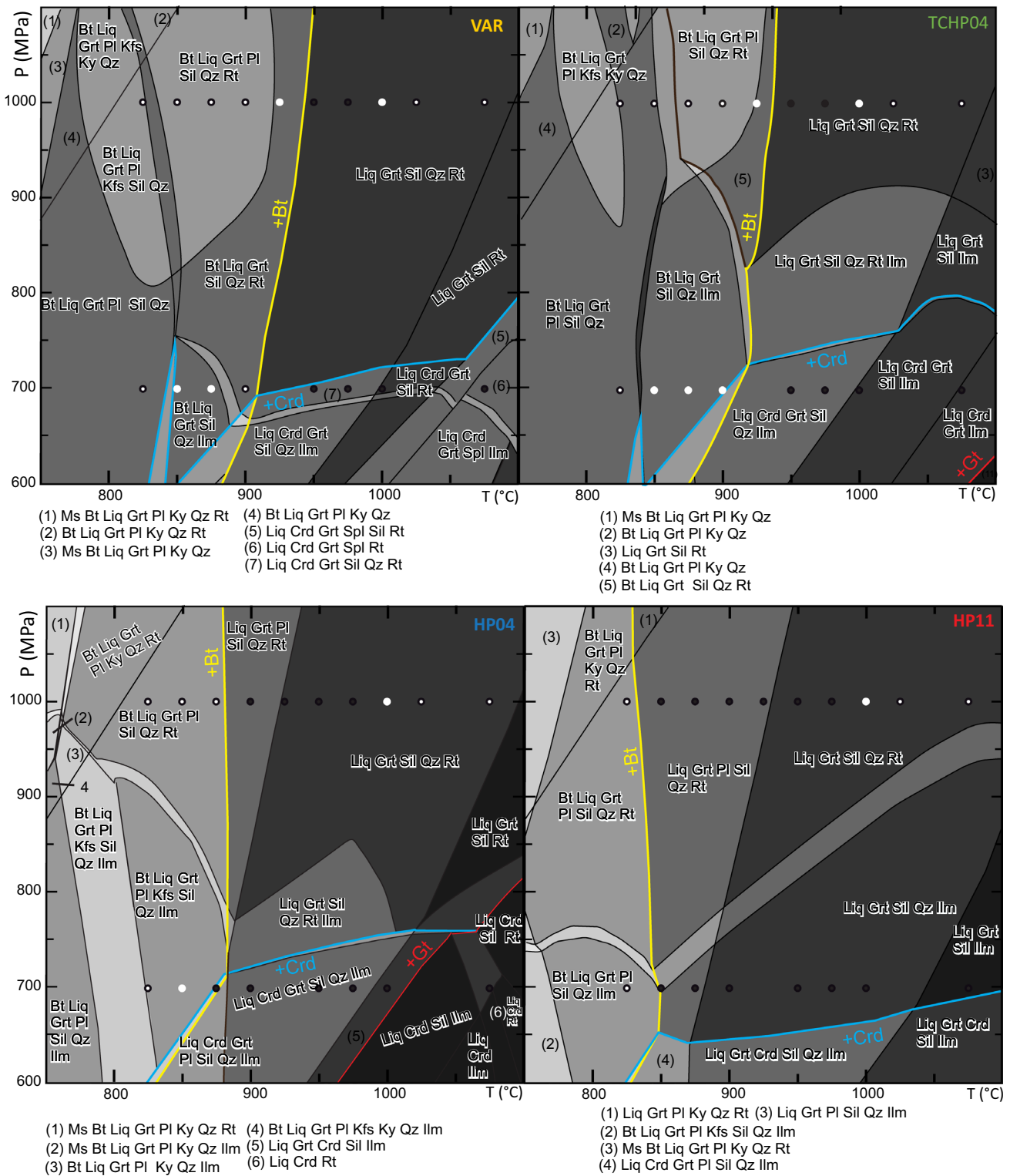
Fig. 2 Isochemical phase diagram sections calculated with the four sets of solution models for the pelitic bulk composition HQ36 (Patiño-Douce and Johnston 1991). Each point represents one experimental run with white circles indicating calculations reproducing experimental paragenesis of ferromagnesian and Ti-oxides minerals, white circles with black contour indicating calculations reproducing experimental paragenesis of ferromagnesian, but not of Ti-oxides minerals, and black circles indicating failure of calculations at reproducing experimental paragenesis of ferromagnesian and of Ti-oxides. Yellow, blue, and red lines show disappearance of biotite, cordierite and garnet, respectively, with the +mineral name on the side of appearance
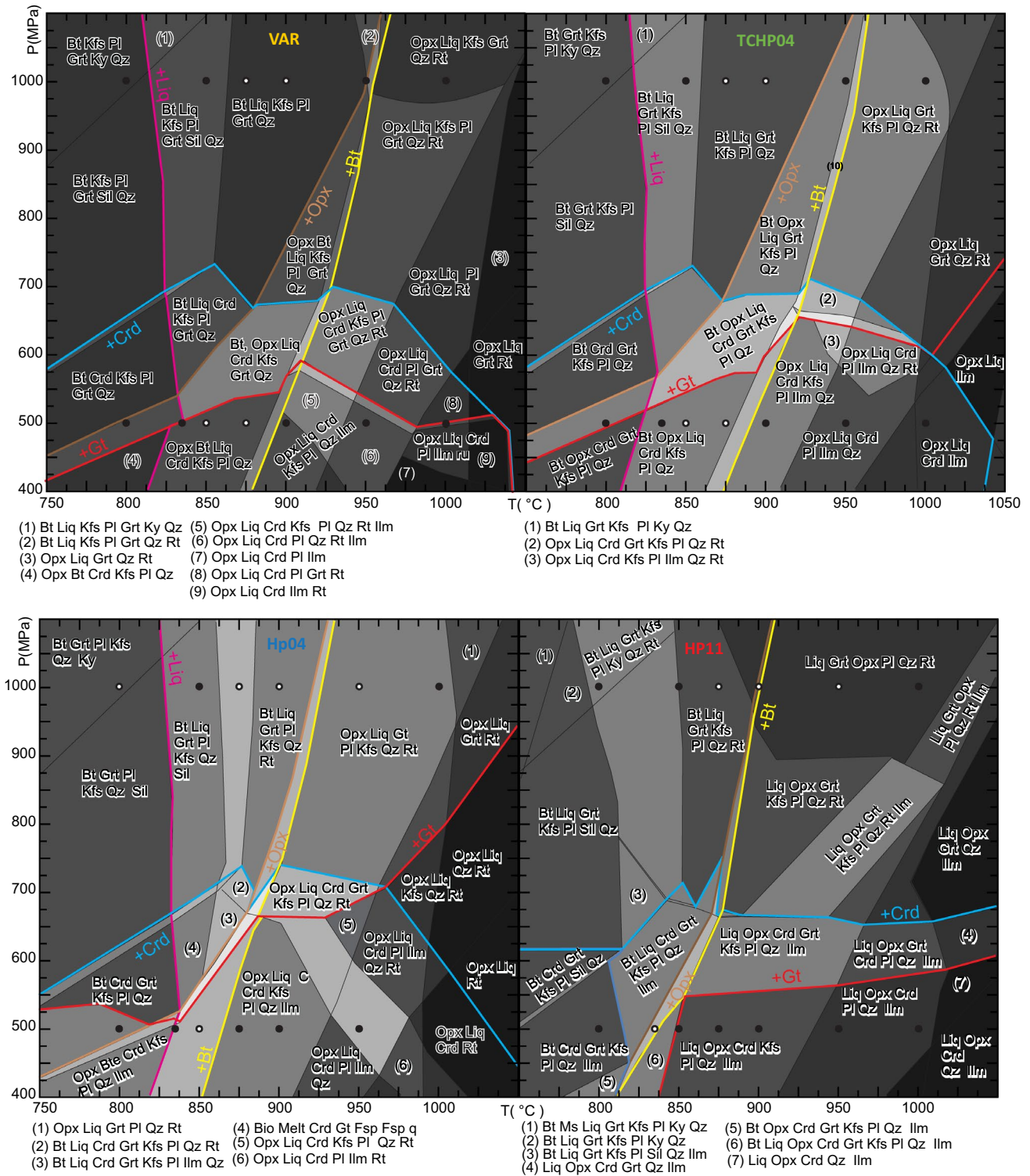
VAR

(1) Bt Liq Kfs Pl Grt Ky Qz
(2) Bt Liq Kfs Pl Grt Qz Rt
(3) Opx Liq Grt Qz Rt
(4) Opx Bt Crd Kfs Pl Qz

(5) Opx Liq Crd Kfs Pl Qz Rt Ilm
(6) Opx Liq Crd Pl Qz Rt Ilm
(7) Opx Liq Crd Pl Ilm
(8) Opx Liq Crd Pl Grt Rt
(9) Opx Liq Crd Ilm Rt

TCHP04

(1) Bt Liq Grt Kfs Pl Ky Qz
(2) Opx Liq Crd Grt Kfs Pl Qz Rt
(3) Opx Liq Crd Kfs Pl Ilm Qz Rt

Hp04

(1) Opx Liq Grt Pl Qz Rt
(2) Bt Liq Crd Grt Kfs Pl Qz Rt
(3) Bt Liq Crd Grt Kfs Pl Ilm Qz
(4) Bio Melt Crd Gt Fsp Fsp q
(5) Opx Liq Crd Kfs Pl Qz Rt
(6) Opx Liq Crd Pl Ilm Rt

HP11

(1) Bt Ms Liq Grt Kfs Pl Ky Qz
(2) Bt Liq Grt Kfs Pl Ky Qz
(3) Bt Liq Grt Kfs Pl Sil Qz Ilm
(4) Liq Opx Crd Grt Qz Ilm
(5) Bt Opx Crd Grt Kfs Pl Qz Ilm
(6) Bt Liq Opx Crd Grt Kfs Pl Qz Ilm
(7) Liq Opx Crd Qz Ilm

Fig. 3 Isochemical phase diagram sections calculated with the four sets of solution models for the grewacke bulk composition NB (Stevens et al., 1997). Each point represents one experimental run with white circles indicating calculations reproducing experimental paragenesis of ferromagnesian and Ti-oxides minerals, white circles with black contour indicating calculations reproducing experimental paragenesis of ferromagnesian but not of Ti-oxides minerals, and black circles indicating failure of calculations at reproducing experimental paragenesis of ferromagnesian and of Ti-oxides. Yellow, blue, red and pale brown lines show disappearance of biotite, cordierite and garnet, and orthopyroxene, respectively, with the +mineral name on the side of appearance
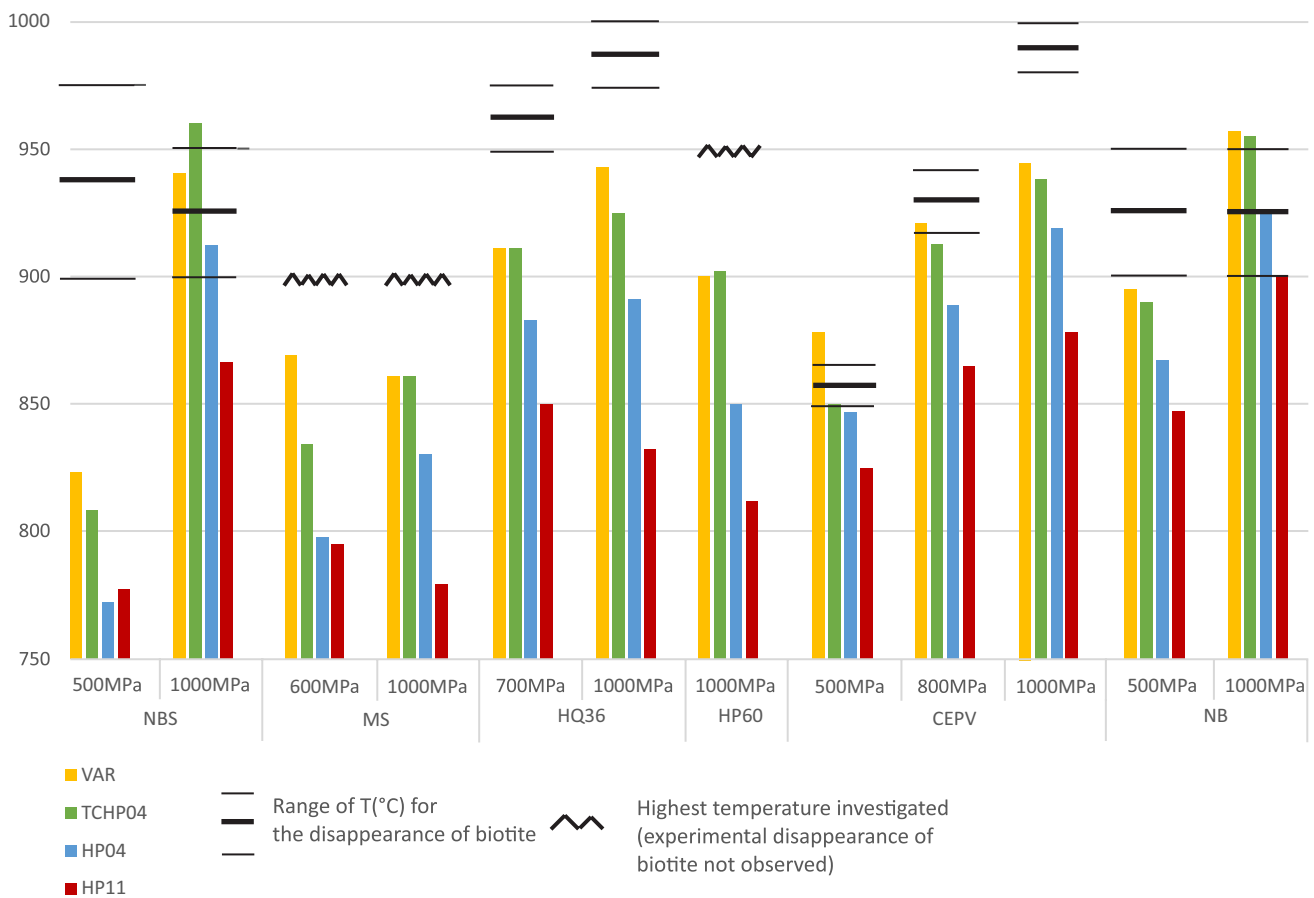
Fig. 4 Calculated vs experimental temperatures at given pressures coincident with the disappearance of biotite for the four sets of solution models. VAR, TCHP04, HP04 and HP11 are the names given to the four sets of solution models investigated in this contribution and described in Table 2

mineral compositions is to investigate trends yielded by experiments. The HQ36 experiment yielded compositional trends for garnet end members as well as for biotite Mg# and TiO$_2$ (Fig. 10). Because Ti-oxides (rutile and ilmenite) parageneses constitute important clues for reconstructing the P–T evolution of metasediments, it is crucial to investigate whether PEFM is able to yield meaningful results. For this, the three experiments that yielded traceable reactions in a P–T diagram were selected (Fig. 11). Below, sets are generally compared by listing values in order of success at reproducing experimental results (from best to worst).

Results

Isochemical phase diagram sections (IPDS)

Figures 2 and 3 present IPDS calculated with the four different sets for one metapelitic (HQ36) and one greywacke (NB) compositions. Obvious differences include the stability field of cordierite (blue lines), garnet (red lines) and

biotite (yellow lines). Although there are obvious differences in topologies, the details are difficult to see on such diagrams. It is, therefore, important to extract specific properties from each diagram to thoroughly investigate the discrepancies between sets.

Biotite-out temperature

Figure 4 presents the upper stability of biotite at each pressure conditions in the six experiments compared with results of PEFM for the four sets investigated. Out of the 12 investigated conditions, calculations with sets VAR, TCHP04, HP04 and HP11 yield a biotite-out temperature within or higher than that observed in the experiments four, three, two and zero times, respectively. A striking observation is that sets VAR and TCHP04 systematically yield higher biotite-out temperatures than the HP's sets, with HP04 generally yielding higher temperatures than HP11.

Paragenesis

Figure 5a presents the number of calculations reproducing the experimental paragenesis for ferromagnesian minerals (biotite, garnet, cordierite, orthopyroxene with the addition of muscovite) of a given experimental run for each set. The first important result is that only 15–37% of the runs are reproduced. The experimental paragenesis for ferromagnesian minerals is correctly modeled for the majority (> 50%) of runs only in experiments HP60 and HQ36 with calculations using sets VAR and TCHP04. Combining all points from all experiments, sets VAR (35%) and TCHP04

(37%) reproduce more parageneses than sets HP04 (28%) and HP11 (15%).

Figure 5b shows a compilation of the numbers of times, for each set of solution models, a calculation failed to reproduce the experimental paragenesis because a given phase was predicted but was not observed or vice versa. These histograms must be interpreted with caution because some minerals are more common than others (e.g., garnet and biotite). Hence, despite that cordierite is responsible for a similar number of failed calculations as garnet (~ 15 for all sets), the former is a much more problematic solution model because it is not expected in as many investigated points as the latter. The orthopyroxene model of the HP11 set performs better

Fig. 6 Calculated vs experimental proportion for biotite (**a**), garnet (**b**) liquid (**c**) and plagioclase (**d**) expressed in volume%. Each data point compares results of a calculation against that of the corresponding experimental run. Solution model sets are color-coded, whereas experiments are represented by symbols. The set that best reproduce the experimental result in a given experimental run is shown as a larger symbol
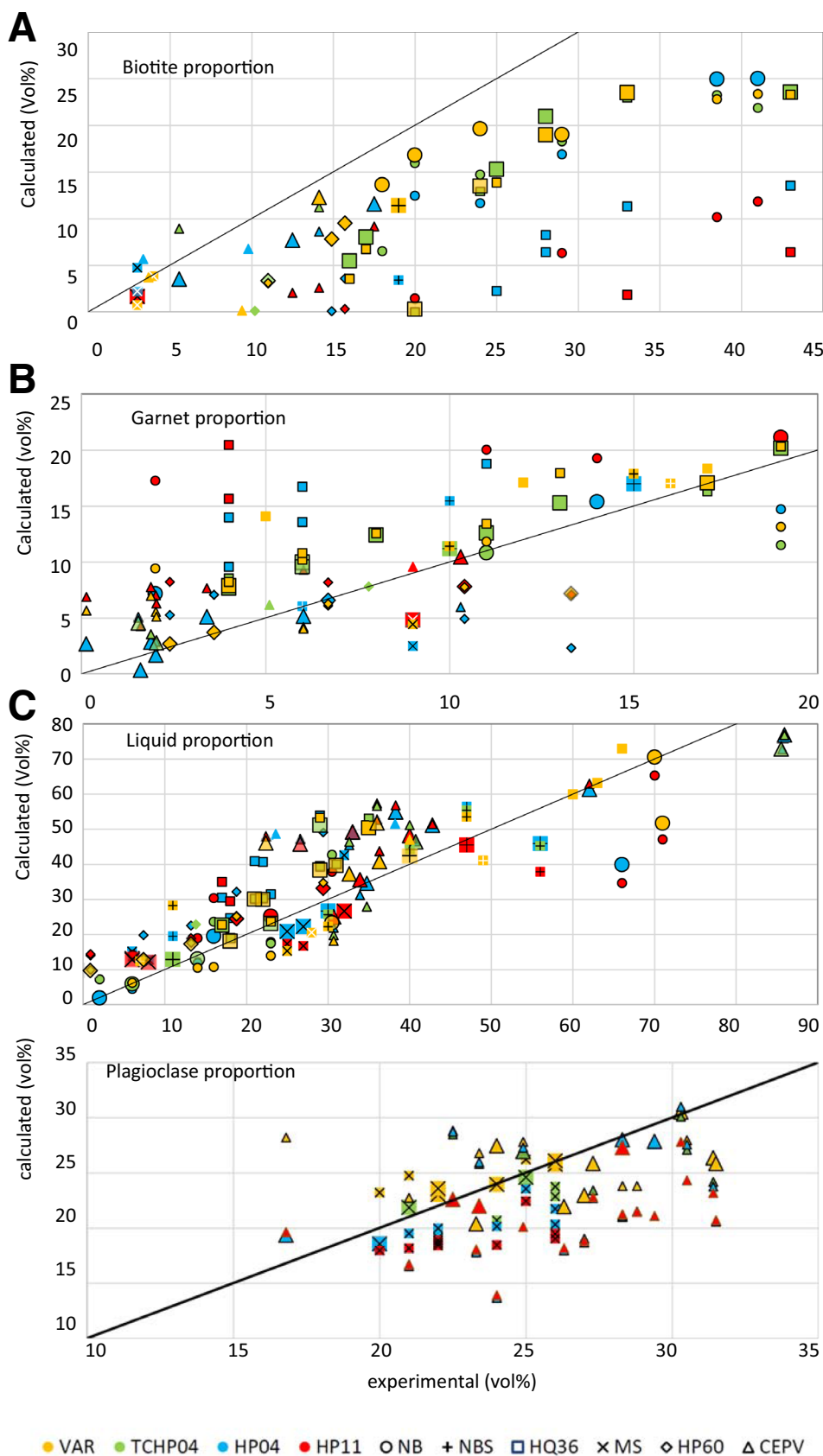
Fig. 7 Calculated vs experimental Mg# (100*Mg/Mg+Fe in mol) for biotite (**a**), garnet (**b**) liquid (**c**) and 100*X$_{An}$ of plagioclase (**d**). Each data point compares results of a calculation against that of the corresponding experimental run. Solution model sets are colorcoded, whereas experiments are represented by symbols. The set that best reproduce the experimental result in a given experimental runs is shown as a larger symbol
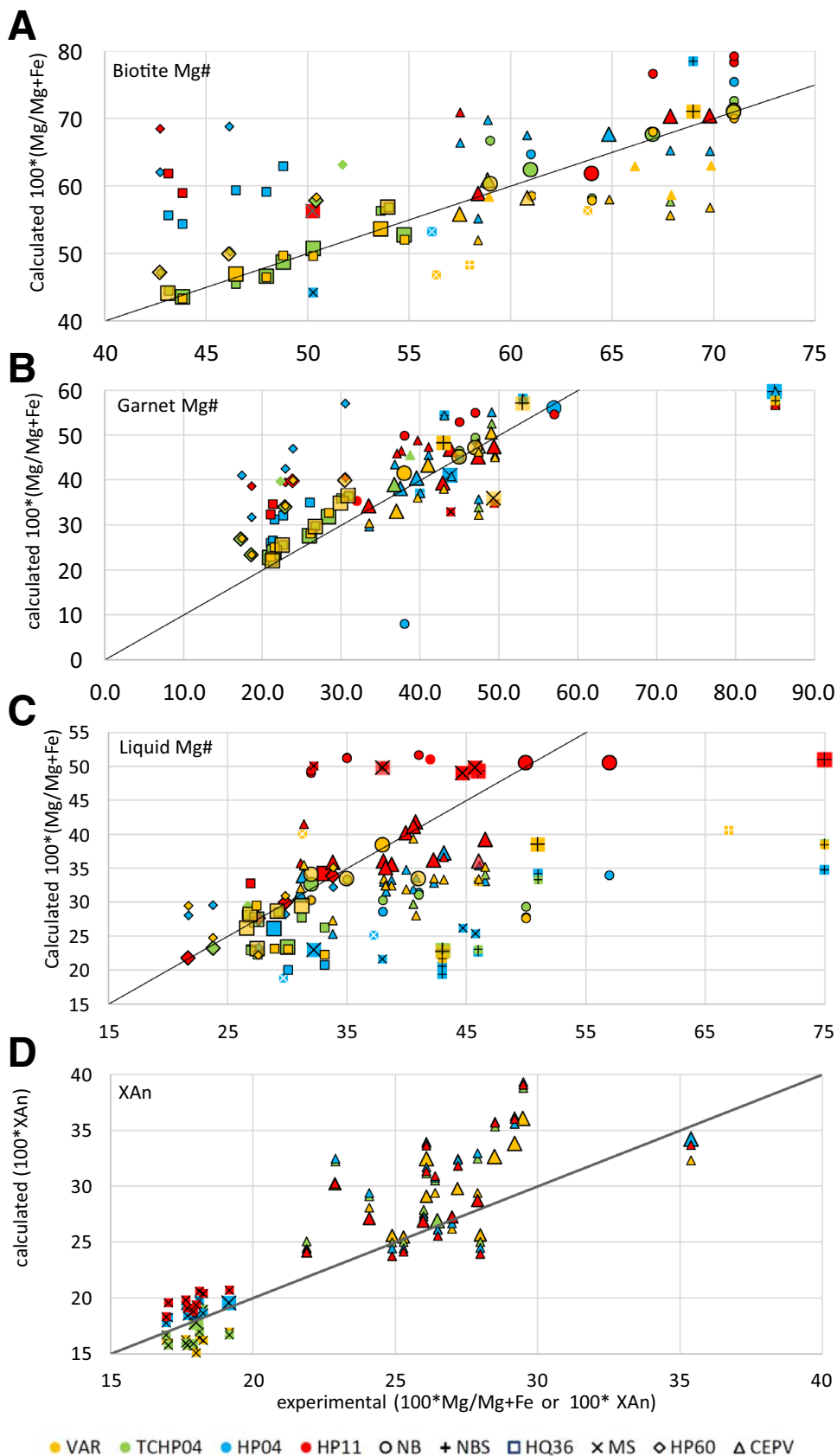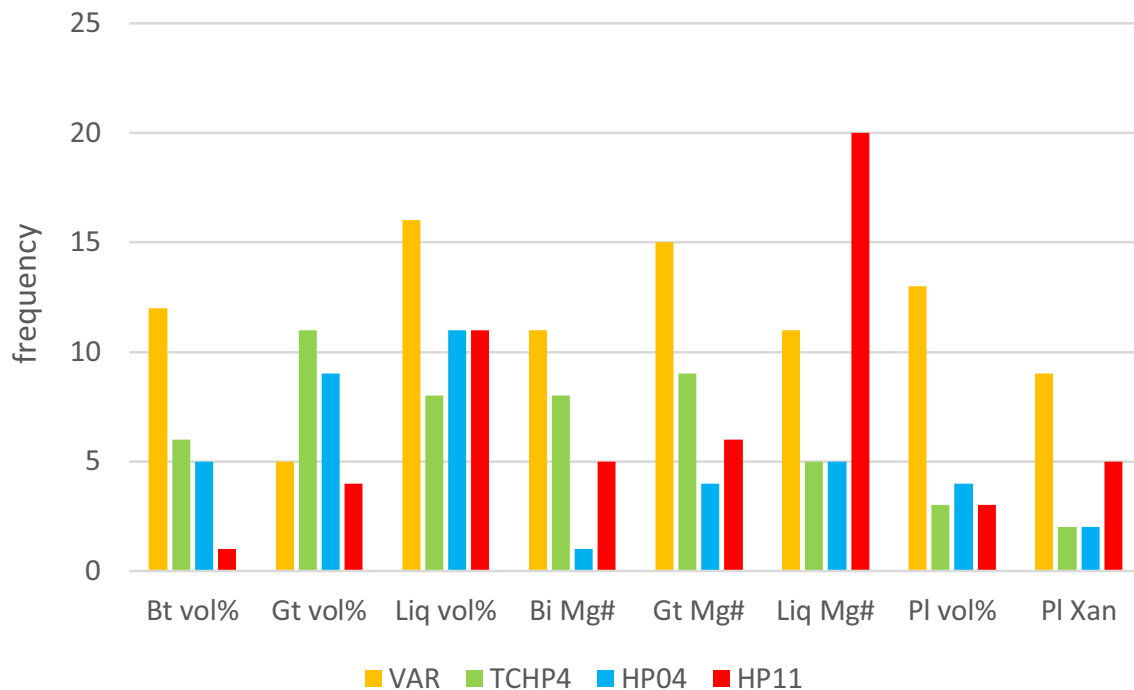
Fig. 8 Frequency plot of each sets of solution models that best reproduce each phase proportion and their composition (Mg# or X_{An} in experimental runs). Shown as larger symbols on Figs. 6 and 7

the other sets, with VAR being the worst. The only garnet model that does not incorporate Mn, HP04, is the most problematic. Moreover, it is clear that the biotite of the two HP sets are the most problematic solution models with 34 and 47 failed calculations, for sets HP04 and HP11, respectively, vs 21 and 18 for sets Var and TCHP04, respectively.

Proportion of major phases

General conclusions can be derived from graphs of calculated vs experimental proportion of phases and mineral compositions. Almost all calculations underestimate biotite proportion (Fig. 6a), whereas garnet proportions were overestimated for a majority of them (Fig. 6b). Data for liquid proportion plot above and below the 1: 1 line (Fig. 6c). Mg# (Mg/Fe + Mg mol%) of biotite plot evenly above and below the 1: 1 line (Fig. 8a), whereas that of garnet generally plot slightly above the line (Fig. 8b). Mg# for liquid generally plot below the 1:1 line except for results from the HP11 set that plots generally above (Fig. 8c). For plagioclase, calculations tend to slightly underestimate its proportion and overestimate the anorthite content (Fig. 9a and b). Apart perhaps for biotite proportion (Fig. 6a), there is no relationship between absolute differences (calculated-experimental) and experimental values suggesting that relative measures (e.g., calculated-experimental/experimental) would not constitute adequate parameters of comparison.

Systematic difference between sets are observed for biotite proportion. On the plot of calculated vs experimental vol%, sets VAR and TCHP04 yield values closer to that of experiments than sets HP04 and HP11 (Fig. 6a). Sets VAR, TCHP04, HP04 and HP11 reproduced best experimental runs 12, 6, 5 and 1 times, respectively (Fig. 8). On boxplots of absolute difference between calculated and experimental vol% (Fig. 9a), sets VAR and TCHP04 yield lower median values (8.8, 10.1, respectively) than sets HP04 (12.6) and HP11 (19.3), and the former two sets have upper whiskers that are lower than the upper quartile of set HP04, which also has upper whiskers lower than the upper quartile value of set HP11. The lower quartile and whiskers of sets VAR are also lower than all other sets. These results demonstrate that sets VAR and TCHP04, which use the Ti-biotite model of Tajčmanová et al. (2009), better reproduce experimental biotite proportion. This exerts significant control on all results presented herein.

Systematic difference between sets are also observed for garnet proportion, although it is not as clear as for biotite proportion. Sets TCHP04, HP04 VAR and HP11 reproduced best experimental runs 11, 9, 5 and 4 times, respectively (Fig. 8). On boxplots of absolute difference between calculated and experimental vol% (Fig. 9b), set TCHP04 yield a lower median value (1.6,) than the other sets (3.5–4.3), and has upper whiskers that are lower than the upper quartile of set HP04, which also has upper whiskers lower than the upper quartile value of set HP11. It is striking that the

Fig. 9 Boxplots compiling absolute differences between calculations and experiments of all data points (i.e., for all experiments) for biotite (**a**), garnet (**b**), liquid (**c**) and plagioclase (**d**) proportions (in vol.%), as well as for biotite (**e**), garnet (**f**), liquid (**g**) 100*Mg# and plagioclase 100*X_{An} (**h**). Lines and X inside the boxplots are the median and average values, respectively. The box itself represents the interquartile range (25-75 percentile) and whiskers shows the maximum and minimum values. Data points yielding anomalous results for all sets were discarded. See text for details
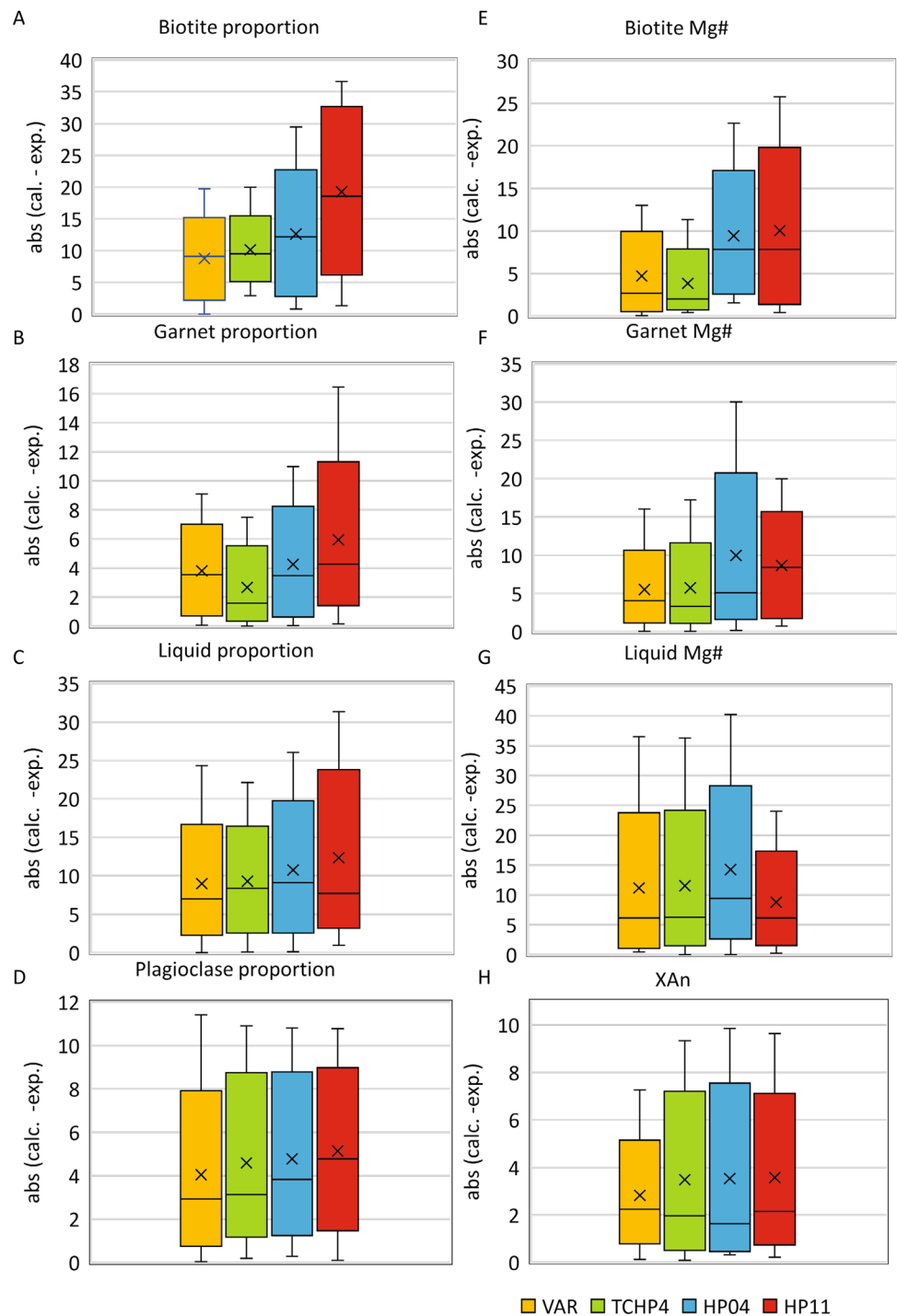
incorporation of Mn (albeit with no margules parameters) and the use of a different biotite model increase the ability of the HP04 garnet model (used in set TCHP04) at better reproducing experimental results.

No systematic differences between sets are observed for liquid proportion. The VAR set best reproduces experimental results for more runs (16) than the other sets (8–11; Fig. 8), but on boxplots of absolute difference between calculated and experimental vol%, it yields similar results as for sets TCHP04 and HP04, whereas set HP11 yields similar median value, but higher upper quartile and whisker values as the other sets (Fig. 9c).

For plagioclase, set VAR reproduce more experimental runs than the other sets (13 vs 3–4; Fig. 8), but on boxplots of absolute calculated–experimental vol% all
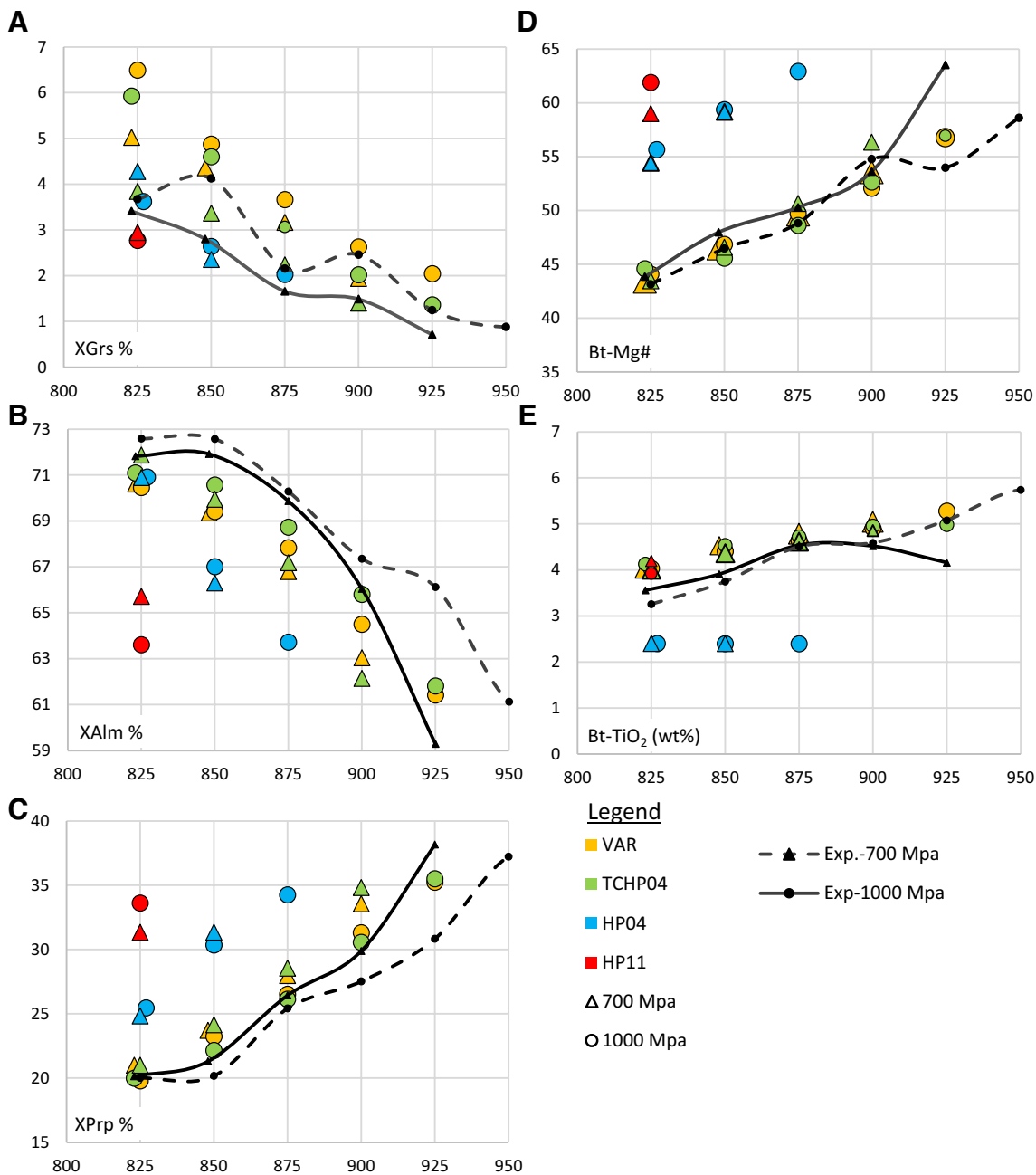
Fig. 10 Calculated (colored symbols) and experimental (lines) garnet and biotite composition for HQ36, the only experiment that yielded compositional trends with temperature. Triangles and circles are for experiments and calculations at 700 and 1000 MPa, respectively. **a**, **b**, **c** The grossular, almandine and pyrope content of garnet expressed in % ($100*X_{Alm}$). **d** Depicts Mg# ($100*Mg/Mg + Fe$ in mole) in biotite. **e** The TiO_2 (in wt%) of biotite. See text for details

sets reproduce equally well the experiment with a pooled median value around 4 (Fig. 9d).

Composition of major phases

Systematic difference between sets are observed for biotite Mg#. Sets VAR, TCHP04, HP11 and HP04 reproduce best experimental runs 11, 8, 5 and 1 times, respectively (Fig. 8).

On boxplots of absolute difference between calculated and experimental Mg# (Fig. 9e), sets VAR and TCHP04 yield lower median values (2–2.7) than sets HP04 and HP11 (~7.8), and the former two sets have upper whiskers that are lower by > 7 than the upper quartile values of sets HP04 and HP11. The lower quartile and whiskers of sets VAR and HP04 are also lower than all other sets. These results demonstrate that sets VAR and TCHP04, which use the Ti-biotite

Fig. 11 Frequency plot of calculations that reproduce the paragenesis of Ti-oxide minerals by the four sets of solution models for all runs of each experiment investigated in this study. For reference, gray bars behind the colored histograms represent the total number of data points calculated for each experiment. See text for details
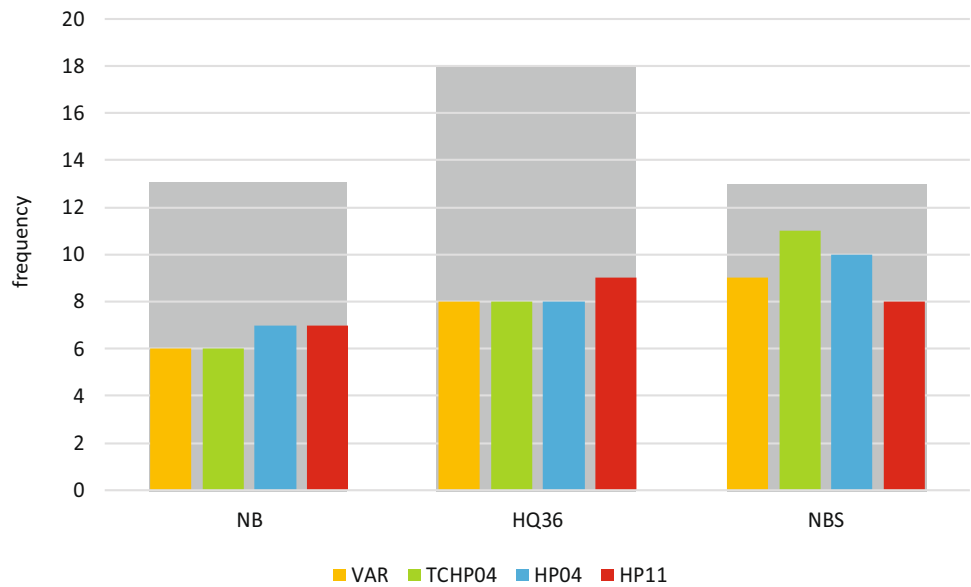
model of Tajčmanová et al. (2009), better reproduce experimental biotite Mg#.

Systematic difference between sets are also observed for garnet Mg#, although it is not as clear as for biotite Mg#. Sets VAR, TCHP04, HP11 and HP04 reproduce best experimental runs 15, 9, 6 and 4 times, respectively (Fig. 8). On boxplots of absolute difference between calculated and experimental Mg# (Fig. 9f), sets VAR, TCHP04 and HP04 yield lower median value (1.6,) than set HP11 (3.5–4.3), whereas sets VAR, TCHP04 and HP11 have upper whiskers that are lower than the upper quartile value of set HP04 (16–17 vs 20–21, respectively). Sets VAR and TCHP04 thus better reproduce garnet Mg# than the HP's sets.

Systematic difference between sets are observed for liquid Mg#. Sets HP11, VAR, TCHP04 and HP04 reproduce best experimental runs 20, 11, 5 and 5 times, respectively (Fig. 8). On boxplots of absolute difference between calculated and experimental Mg# (Fig. 9g), sets VAR, TCHP04, HP11 yield a lower median value (~ 6.1) than set HP04 (9.4). Set HP11 has upper whiskers that are lower than the upper quartile values of sets VAR, HP04 and HP11. Mg# of liquid is the only parameter investigated in this contribution for which set HP11 reproduces better experimental results than the other sets.

Systematic difference between sets are observed for the anorthite content of plagioclase. Sets VAR, HP11, TCHP04 and HP04 reproduce best experimental runs 9, 5, 2 and 2 times, respectively (Fig. 8). On boxplots of absolute difference between calculated and experimental X_{An} (Fig. 9h), all sets have a similar median values (1.6 to 2.3), but set VAR has an upper whisker value lower than the upper quartile values of the other sets. The plagioclase model of set VAR, constructed from Fuhrman and Lindsley (1988), thus better reproduces experimental compositions than the HP's sets.

Another way of comparing mineral compositions is to investigate compositional change with temperature in the experiments. The HQ36 experiment yielded compositional trends for garnet end members as well as for biotite Mg# and TiO$_2$. At 800 and 1000 MPa, X_{Prp} and Bt-Mg# increase with increasing temperature, whereas X_{Alm} and X_{Grs} decreases and Bt-TiO$_2$ remains relatively flat (lines on Fig. 10). There are not as many data points for the HP's sets because only calculations that reproduce the paragenesis of the main phase are considered. In fact, there is only one data point for set HP11, which precludes any comparison with the other sets. Sets VAR, TCHP04 and HP04all follow the experimental trends. The large majority of data points calculated with sets VAR and TCHP04 further overlap experimental results, whereas set HP04 yield results that systematically diverge from experimental runs for all compositions (Fig. 10b, c, d, e), except for grossular (Fig. 10a).

Titanium oxides paragenesis and reactions

There is no difference between sets for the frequency of successful calculations at reproducing experimental parageneses for Ti-oxides. All sets reproduce ~ 50% of the observed parageneses for the NB and HQ36 experiment and 60–70% for the NBS experiment (Fig. 11).

Three experiments yielded results with coherent reactions for titanium oxides (Fig. 12). The NBS experiment yielded an ilmenite-in reaction with a positive slope from low to high P and an ilmenite→rutile reaction at high P. NB yielded similar reactions except that the ilmenite-in has a negative slope. Finally, HQ36 appears to have a V-shaped ilmenite-rutile reaction at high P, although its exact location is not clearly defined.
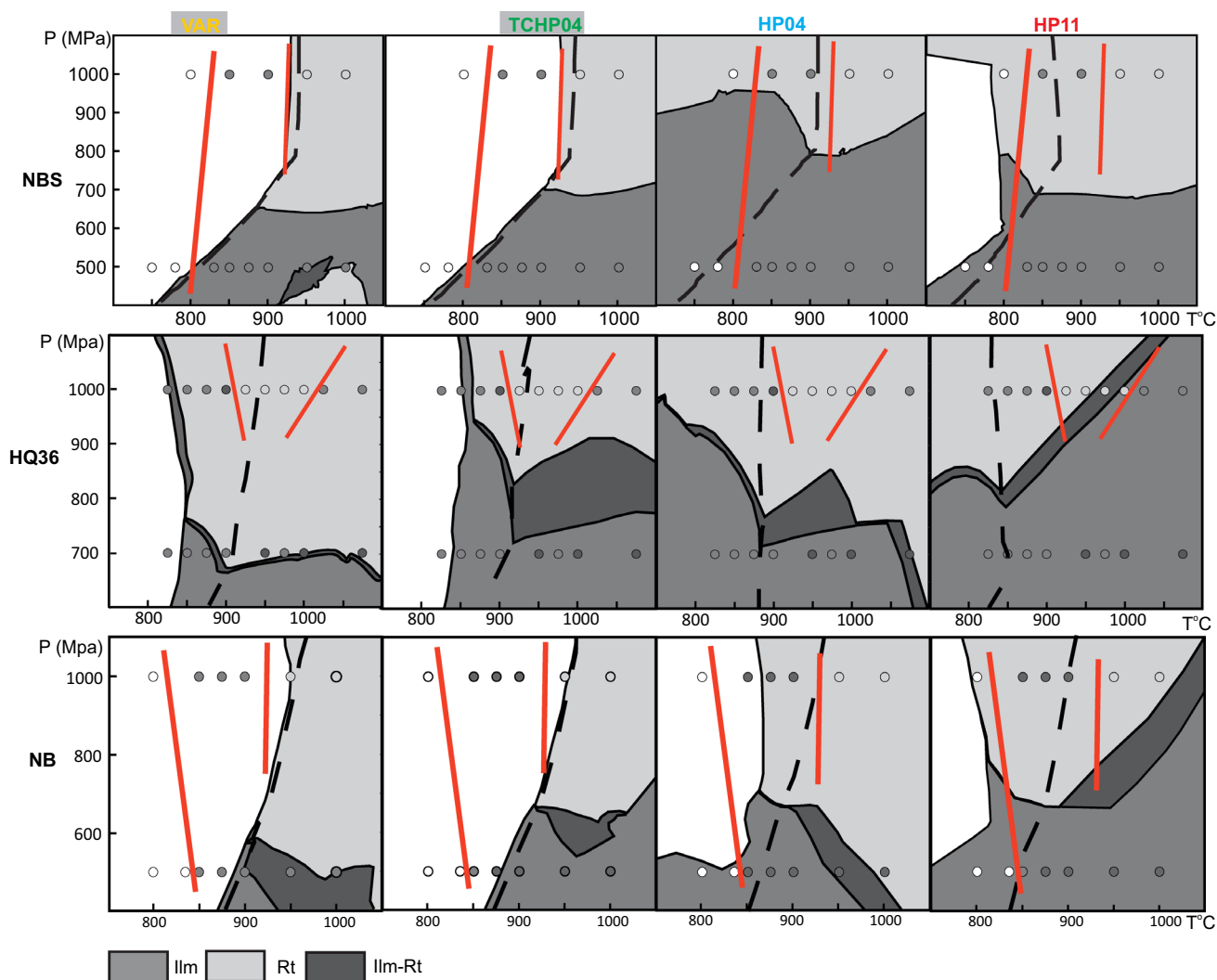
Fig. 12 Isochemical phase diagram sections showing Ti-oxides parageneses calculated with the four sets of solution models for the experiments that yielded coherent reactions (indicated as red lines). Circles represent experimental runs. Circles and background fill are color-coded for the Ti-oxides observed in experiments or calculated, respectively. Stippled black lines represent biotite-breakdown reaction. See text for details

Systematic differences between sets are observed for reactions involving titanium oxides. For the NBS experiment, sets VAR and TCHP04 reproduce the low-P ilmenite-in and the high-P rutile-in reactions, but ilmenite is not predicted to be stable at a high enough P compared to that observed in the experiment. In contrast, HP04 and HP11 do not reproduce any of the reactions, but HP04 does predict ilmenite stability at a higher P than the other three sets. For the NB experiment, sets VAR and TCHP04 predict the rutile-in reaction and an ilmenite-in reaction at a T ~ 50 °C higher than its observed appearance. Set HP04 predict the ilmenite-in reaction within 30 °C, albeit at a slightly higher P, and a rutile-in reaction at high-P at a T ~ 40 °C lower than its observed appearance. Set HP11 yield results diverging from that of the experiments. Interestingly, for the HQ36 experiment, all

sets predict a V-shaped boundary for the rutile-in reaction. Set TCHP04 predicts the negatively sloping ilmenite→rutile reaction within ~ 50 °C, while the positively sloping rutile → ilmenite reaction is predicted at 100–200 MPa lower P and ~ 50 °C lower T, whereas set HP11 predicts the positively sloping rutile → ilmenite reaction within 50 °C. All the other sets yield reactions that are far off the experimentally derived ones. It is interesting to note that for all three experiments, set TCHP04 yield results that are more similar to set VAR than to set HP04 despite that it uses the same ilmenite solution model as the latter. Because the main difference between sets TCHP04 and set HP04 is the biotite solution model, it appears that the choice of this model has a stronger influence on the stability field of Ti-oxides than the choice of the ilmenite solution model.

Discussion

As discussed by White et al. (2011), comparing results of partial melting experiments to that of phase equilibria modeling is fraught with many uncertainties that are difficult to ascertain. In any given comparison, it is almost impossible to determine if an observed discrepancy between the two methods is related to experimental problems, such as failure to reach equilibrium or poorly constrained or variable oxidation state; or if it is related to modeling problems, such as uncertainties in the thermodynamic dataset or with solution models. However, by comparing the same parameters (mineral proportion, composition and specific metamorphic reactions) between six experiments spanning a wide range of bulk compositions and results of PEFM using four different sets of solution models, we consider that we have largely circumvented the above-mentioned problems. Not only that this analysis allows us to determine if one (or more) set of solution models better reproduces experiments than the others (see below), but it could also provide an independent evaluation of partial melting experiments. For example, boxplots of calculated-experimental garnet Mg# values (Fig. 9e) show that the large majority of calculations are within 15 (absolute value). On the calculated vs experimental values graph (Fig. 7b), this is reflected by data points plotting near the 1:1 line. There is one exception, however, for points of all sets that plot at the extreme right of the graph (at experimental Mg# = 85), which corresponds to the run at 1000 MPa/1000 °C of the NBS experiment). The inconsistency of these data points compared to all other results strongly suggests it is an experimental anomaly, We therefore discarded this data point in the boxplots of Fig. 8e. Because only two other data points were discarded, we consider that the vast majority of experimental runs yield good results and can be used in our comparison with results of PEFM. Furthermore, these results suggest that PEFM using different sets of thermodynamic datasets and solid solutions could be used as an independent test for judging the validity and repeatability of experimental runs.

This contribution indicates that some sets systematically better reproduce experimental results than others. Set VAR and set TCHP04 systematically yield modeled biotite-breakdown temperature that are closer to those observed in experiments compared with set HP04 (higher by 20–100 °C), which, in turn, generally yields slightly higher temperatures than set HP11 (Fig. 4). The same two sets also reproduce the observed Fe–Mg parageneses (Fig. 5a) for more experimental runs than set HP04 (by 8–9%) and set HP11 (by 20–22%). Figure 8 shows that set VAR outnumbered the other sets for the number of times it best reproduces a given experimental run (shown

as larger symbols on Figs. 6 and 7) for biotite, liquid and plagioclase proportion and for biotite, garnet and plagioclase composition, although set TCHP04 yields better results more frequently than the other sets for garnet proportion. On boxplots pooling all the difference between calculated and experimental values (absolute), sets VAR and TCHP04 clearly yield better results than the HP's sets for biotite and garnet proportion (Fig. 9a and b) as well as for biotite and garnet Mg# (Fig. 9e and f). Set VAR also yields better results for the anorthite content of plagioclase (Fig. 9h). Set HP11 yields a better result than the other sets only for the liquid Mg# (both for the frequency of best reproduction, Fig. 8, and on boxplot of Fig. 9g). Perhaps that more attention was given to this parameter in the calibration of the solution model. Although the four sets reproduce the same proportion of Ti-oxides parageneses observed in experiments HQ36, NBS and NB, only sets VAR and TCHP04 reproduce the HP ilmenite → rutile and the LP ilmenite-in reactions observed in the latter two experiments (Fig. 11). Consequently, results of this study clearly indicate that sets VAR and TCHP04 are preferable than sets HP04 and HP11 for PEFM of metasediments for a wide variety of investigated parameters.

Several clues suggest that one solution model is mainly responsible for the observed discrepancies between sets. The striking differences for temperatures of biotite-breakdown observed between solution model sets using the Tajčmanová et al. (2009) biotite model and those that do not (setVAR/ TCHP04 vs HP04/HP11), hint that this model has a profound influence on PEFM of metasediments (Fig. 4). This hypothesis is supported by the compilation of minerals causing problems in the modeled Fe–Mg paragenesis (Fig. 5b) showing that biotite was the cause of failure of calculations at reproducing experimental paragenesis more frequently than any other phases. Finally, the fact that biotite is the only solution model distinguishing sets TCHP04 and HP04 confirms that it is the main cause of differences between sets. Because biotite is a major carrier of Fe, Mg and Ti and that it is the main phase involved in partial melting of metasediments, its proportion in calculations would necessarily have a significant impact on the proportion and composition of garnet, liquid and Ti-oxides, as observed herein.

Several factors could explain the discrepancies between biotite models documented herein. Tajčmanová et al. (2009) significantly modified the HP04 biotite model, from which the HP11 was also built upon. They first reduced the enthalpy of the ordering reaction (obi) from $\Delta H^0_{order} = -2$ kJmol^{-1} (Holland and Powell 2006) to $\Delta H^0_{order} = -6.8$ kJmol^{-1}. This had the effect of lowering the octahedral Al-content of biotite that resulted in a better match with natural data and an increase in predicted biotite proportion. Based on crystallographic studies, they also changed the Ti

substitution in biotite from ordering on the M1, as in the HP04 and HP11 sets, to the M2 octahedral site. This has the net effect of increasing the entropy of mixing, and thus lead again to an increase in biotite proportion. In parametrizing the model, Tajčmanová et al. (2009) determined by regression that the non-ideal mixing parameter involving the Ti endmember was null, whereas it is positive for the HP's sets. This also has the net effect of increasing the predicted biotite proportion. Consequently, as predicted by the thermodynamic formulation of the model, PEFM of partial melting of sediments resulted in significant increase in the calculated biotite proportion by using the Tajčmanová et al. (2009) biotite model (Fig. 6a), which has a cascading effect on other parameters (i.e., proportion and Mg# of liquid and garnet).

Conclusion

- This contribution tested four sets of solution models using two thermodynamic datasets (tcds55 and tcds62; Powell et al. 1998 and Holland and Powell 2011, respectively) for calculating isochemical phase diagram sections (Figs. 2 and 3) of six partial melting experiments covering a wide range of metasedimentary bulk compositions (Table 1 and Fig. 1).
- Sets using the biotite model of Tajčmanová et al. (2009) and the datasets tcds55 better reproduce several experimentally derived parameters, notably: (1) biotite-breakdown temperatures (Fig. 4); (2) frequency of success at reproducing parageneses of the main phases (Fig. 5a); (3) proportion of biotite and garnet (Figs. 8, 9a, b); (4) more frequent success at reproducing experimental proportion of liquid and plagioclase (Fig. 8); (5) composition of biotite, garnet and plagioclase (Figs. 8, 9e, f, h); (6) reactions involving Ti-oxides (Fig. 12).
- The solution model causing the discrepancies between sets is the biotite model with the model of Tajčmanová et al. (2009) being preferable.
- Results presented herein and the excellent results obtained on natural rocks using this biotite model (Larson et al. 2013; Gervais and Crowley 2017; Perrot et al. 2020; Soucy La Roche et al. 2019), strongly argue in favor of using this biotite model for phase equilibrium forward modeling of migmatitic metasediments.

References

Brown M (2010) Melting of the continental crust during orogenesis: the thermal, rheological, and compositional consequences of melt transport from lower to upper continental crust. Can J Earth Sci 47:655–694

Clemens JD (2006) Melting of the continental crust: fluid regimes, melting reactions, and source-rock fertility. In: Brown M, Rushmer T (eds) Evolution and differentiation of the continental crust. Cambridge University Press, Cambridge, pp 297–331

Connolly JA (2005) Computation of phase equilibria by linear programming: a tool for geodynamic modeling and its application to subduction zone decarbonation. Earth and Planetary Science Letters 236(1–2):524–541

Connolly JAD, Petrini K (2002) An automated strategy for calculation of phase diagram sections and retrieval of rock properties as a function of physical conditions. J Metamorph Geol 20(7):697–708

de Capitani C, Petrakakis K (2010) The computation of equilibrium assemblage diagrams with Theriak/Domino software. Am Miner 95(7):1006–1016

Douce AEP, Beard JS (1994) H2O loss from hydrous melts during fluid-absent piston cylinder experiments. Am Mineral J Earth Planet Mater 79(5–6):585–588

Duesterhoeft E, Lanari P (2020) Iterative thermodynamic modelling—Part 1: A theoretical scoring technique and a computer program (Bingo-Antidote). J Metamorph Geol 38(5):527–551

Dumond G, Goncalves P, Williams ML, Jercinovic MJ (2015) Monazite as a monitor of melting, garnet growth and feldspar recrystallization in continental lower crust. J Metamorph Geol 33(7):735–762

Forshaw JB, Waters DJ, Pattison DR, Palin RM, Gopon P (2019) A comparison of observed and thermodynamically predicted phase equilibria and mineral compositions in mafic granulites. J Metamorph Geol 37(2):153–179

Fuhrman ML, Lindsley DH (1988) Ternary-feldspar modeling and thermometry. Am Miner 73(3–4):201–215

García-Arias M (2020) Consistency of the activity–composition models of Holland, Green, and Powell (2018) with experiments on natural and synthetic compositions: a comparative study. J Metamorph Geol 00:1–18

Gervais F, Crowley JL (2017) Prograde and near-peak zircon growth in a migmatitic pelitic schist of the southeastern Canadian Cordillera. Lithos 282:65–81

Grant JA (2009) Thermocalc and experimental modelling of melting of pelite, Morton Pass, Wyoming. J Metamorph Geol 27:571–578

Guevara VE, Caddick MJ (2016) Shooting at a moving target: phase equilibria modelling of high-temperature metamorphism. J Metamorph Geol 34(3):209–235

Holland TJB, Powell R (2006) Mineral activity–composition relations and petrological calculations involving cation equipartition in multisite minerals: a logical inconsistency. J Metamorph Geol 24(9):851–861

Holland TJB, Powell R (2011) An improved and extended internally consistent thermodynamic dataset for phases of petrological interest, involving a new equation of state for solids. J Metamorph Geol 29(3):333–383

Holland TJ, Green EC, Powell R (2018) Melting of peridotites through to granites: a simple thermodynamic model in the system KNCFMASHTOCr. J Petrol 59(5):881–900

Indares A, White RW, Powell R (2008) Phase equilibria modelling of kyanite-bearing anatectic paragneisses from the central Grenville Province. J Metamorph Geol 26(8):815–836

Johnson TE, White RW, Powell R (2008) Partial melting of metagreywacke: a calculated mineral equilibria study. J Metamorph Geol 26:837–853

Kelsey DE, Powell R (2011) Progress in linking accessory mineral growth and breakdown to major mineral evolution in metamorphic rocks: a thermodynamic approach in the Na2O-CaO-K2O-FeO–MgO–Al$_2$O$_3$–SiO$_2$–H$_2$O–TiO$_2$–ZrO$_2$ system. J Metamorph Geol 29(1):151–166

Kendrick J, Indares A (2018) The reaction history of kyanite in high-P aluminous granulites. J Metamorph Geol 36(2):125–146

Lanari P, Duesterhoeft E (2019) Modeling metamorphic rocks using equilibrium thermodynamics and internally consistent databases: past achievements, problems and perspectives. J Petrol 60(1):19–56

Lang HM, Gilotti JA (2015) Modeling the exhumation path of partially melted ultrahigh-pressure metapelites, North-East Greenland Caledonides. Lithos 226:131–146

Larson KP, Gervais F, Kellett DA (2013) A P–T–t–D discontinuity in east-central Nepal: implications for the evolution of the Himalayan mid-crust. Lithos 179:275–292

Montel JM, Vielzeuf D (1997) Partial melting of metagreywackes, Part II. Compositions of minerals and melts. Contrib Miner Petrol 128(2–3):176–196

Patiño-Douce AE, Harris N (1998) Experimental Constraints on Himalayan Anatexis. J Petrol 39(4):689–710

Patiño-Douce AE, Johnston AD (1991) Phase equilibria and melt productivity in the pelitic system: implications for the origin of peraluminous granitoids and aluminous granulites. Contrib Miner Petrol 107(2):202–218

Pattison DRM, Tinkham DK (2009) Interplay between equilibrium and kinetics in prograde metamorphism of pelites: an example from the Nelson aureole, British Columbia. J Metamorph Geol 27(4):249–279

Perrot M, Tremblay A, Ruffet G, Labrousse L, Gervais F, Caroir F (2020) Diachronic metamorphic and structural evolution of the Connecticut Valley-Gaspé trough, Northern Appalachians. J Metamorph Geol 38(1):3–27

Pickering JM, Johnston DA (1998) Fluid-absent melting behavior of a two-mica metapelite: experimental constraints on the origin of Black Hills granite. J Petrol 39(10):1787–1804

Powell R, Holland TJB (2008) On thermobarometry. J Metamorph Geol 26(2):155–179

Powell R, Holland TJB, Worley BA (1998) Calculating phase diagrams involving solid solutions via non-linear equations, with examples using THERMOCALC. J Metamorph Geol 16(4):577–588

Smye AJ, Greenwood LV, Holland TJB (2010) Garnet–chloritoid–kyanite assemblages: eclogite facies indicators of subduction constraints in orogenic belts. J Metamorph Geol 28(7):753–768

Soucy La Roche R, Godin L, Cottle JM, Kellett DA (2019) Tectono-metamorphic evolution of the tip of the Himalayan metamorphic core in the Jajarkot klippe, west Nepal. J Metamorph Geol 37:239–269

Spear FS (2010) Monazite–allanite phase relations in metapelites. Chem Geol 279(1–2):55–62

Spear FS, Pyle JM (2010) Theoretical modeling of monazite growth in a low-Ca metapelite. Chem Geol 273(1–2):111–119

Spear FS, Hallett B, Pyle JM, Adalı S, Szymanski BK, Waters A, Linder Z, Pearce SO, Fyffe M, Goldfarb D, Glickenhouse N (2009) MetPetDB: A database for metamorphic geochemistry. Geochem Geophys Geosyst 10(12)

Stevens G, Clemens JD, Droop GTR (1997) Melt production during granulite-facies anatexis: experimental data from "primitive" metasedimentary protoliths. Contrib Miner Petrol 128(4):352–370

Tajčmanová L, Connolly JAD, Cesare B (2009) A thermodynamic model for titanium and ferric iron solution in biotite. J Metamorph Geol 27(2):153–165

Tropper P, Hauzenberger C (2015) How well do pseudosection calculations reproduce simple experiments using natural rocks: an example from high-P high-T granulites of the Bohemian Massif. Austrian J Earth Sci 108:123–138

Tropper P, Konzett J, Bertoldi C, Cooke RA, Finger F (2002) How close can we resemble nature? Experimental investigations on high-P dehydration melting of biotite and the comparison to high-P/high-T granulites. Mineral Petrol 134:17–32

Vielzeuf D, Montel JM (1994) Partial melting of metagreywackes. Part I. Fluid-absent experiments and phase relationships. Contribut Mineral Petrol 117(4):375–393.Yakymchuck and Brown, 2013

White RW, Powell R (2002) Melt loss and the preservation of granulite facies mineral assemblages. J Metamorph Geol 20(7):621–632

White RW, Pomroy NE, Powell R (2005) An in situ metatexite–diatexite transition in upper amphibolite facies rocks from Broken Hill. Australia. Journal of Metamorphic Geology 23(7):579–602

White RW, Powell R, Holland TJB (2007) Progress relating to calculation of partial melting equilibria for metapelites. J Metamorph Geol 25(5):511–527

White RW, Stevens G, Johnson TE (2011) Is the Crucible Reproducible? Reconciling Melting Experiments with Thermodynamic Calculations. Elements 7(4):241–246

White RW, Powell R, Johnson TE (2014) The effect of Mn on mineral stability in metapelites revisited: new a–x relations for manganese-bearing minerals. J Metamorph Geol 32(8):809–828

Wu CM, Zhao GC (2006) The applicability of the GRIPS geobarometry in metapelitic assemblages. J Metamorph Geol 24(4):297–307

Yakymchuk C, Brown M (2014) Behaviour of zircon and monazite during crustal melting. Journal of the Geological Society 171(4):465–479