Hydrothermal heat flow near the Main Central Thrust, central Nepal Himalaya

L.A. Derry*, M.J. Evansb, R. Darlingc, C. France-Lanordd

a Cornell University, Dept. of Earth and Atmospheric Sciences, Ithaca, NY, USA
b Dept. of Chemistry, Wheaton College, Norton, MA, USA
c Dept. of Geology, SUNY Cortland, Cortland, NY, USA
d Centre des Recherches Petrographiques et Geochimiques, CNRS, Nancy, France

**A R T I C L E  I N F O**

Article history:
Received 18 February 2009
Received in revised form 12 June 2009
Accepted 15 June 2009
Available online 3 August 2009

Editor: T.M. Harrison

Keywords:
heat flow
hydrothermal
Himalaya
hot springs
thermochronometry

**A B S T R A C T**

In central Nepal hot springs are common in a broad zone where deeply incised river gorges cross the MCT along steep stream reaches. The chemistry of the hydrothermal fluids is distinct from that of the rivers, enabling the use of chemical mass balance to estimate the hydrothermal flux. Spring exit temperatures are 25–70 °C. We combine mass balance estimates of spring flux with observed T to estimate hydrothermal power dissipation of springs at the surface of 500 MW across the Narayani basin. Estimates of reservoir T and a simple conductive cooling model indicate an additional 1260 MW is transferred to the uppermost crust by fluid flow, for a total heat transfer of 1760±953 MW. Fluid inclusions from post-ductile quartz veins yield T = 301±6 °C, P = 1056±110 bars, and imply a geothermal gradient 75±7.8 °C km−1. Simple models of thermal balance suggest that hydrothermal circulation is a major mechanism of heat loss from this region of steep geothermal gradient. A 1-D model of rock advection and radiogenic heating indicates that coupled erosion-rock uplift is the major source for heat in the upper crust, and is consistent with the observed magnitude of hydrothermal heat transfer. On a more local scale, a simple porous media flow model appears to predict observed T-depth relationships in the zone of active meteoric flow, and implies T ≈ 100 °C at depths only on the order of 100 m. These very shallow isotherms have the potential to influence low T thermochronometers.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Rapid uplift and exhumation in mountain belts has been shown to create steep geothermal gradients that can drive hydrothermal circulation (Koons and Craw, 1991; Jenkin et al., 1994; Koons et al., 1998). Hydrothermal circulation can be an important term in the thermal budget of an orogen (Ingebritsen et al., 1989; Forster and Smith, 1989; Manga, 1998). Hot springs occur in all of the major river valleys of central Nepal, but their contribution to the thermal budget of the Himalayan front is poorly known. They are primarily found in two kinds of tectonic–geomorphic settings. A number of hot springs are located in major graben structures, such as in the Thakola or Mustang graben that contains the upper reaches of the Kali Gandaki river at altitudes of 3500–4000 m. The second common setting for hot springs is along steep reaches of major streams in deeply incised valleys, generally near 1000–1500 m altitude. Only large, deeply incised valleys in this section of Nepal have significant geothermal activity. Many of the springs are located near the intersection of the Main Central Thrust (MCT) with the major valleys (Le Fort and Jest, 1974; Bhattacharai, 1980). The zones of geothermal activity tend to be associated with over-steepened reaches of the rivers, which are believed to reflect active tectonic processes (Seeber and Gornitz, 1983; Harrison et al., 1997; Catlos et al., 2001; Lave and Avouac, 2001; Wobus, et al. 2003; Wobus et al., 2006). This topographic break is roughly coincident with the Main Central Thrust across much of the study areas, but lies to the south of the MCT in the Trisuli valley where the surface trace of the MCT turns sharply north. Hot springs are rare or absent in smaller, less deeply incised valleys associated with lower order streams in the region. The association of the springs with first order topographic and structural features of the Himalayan front suggests a tectonic–geomorphic control on their occurrence. Models suggest that the zone of steep topographic gradients along the front is the locus of rapid erosion and rock advection toward the surface, resulting in steepened geothermal gradients (Koons and Craw, 1991; Beaumont et al., 2001).

The chemistry of the hot spring fluids in Nepal creates readily detectable geochemical anomalies in Himalayan rivers (Evans et al., 2001; 2004). Here we use those anomalies as part of a geochemical mass balance that permits an estimate of the hydrothermal heat flux along the Himalayan front in central Nepal. We combine these estimates of the hot water flux to the surface with fluid inclusion data that constrain the local geothermal gradient to make an overall estimate of the heat budget in this area of active uplift and incision.
1. Study area and spring locations

The study area includes eight major rivers that combine to form the Narayani river system of central Nepal (Fig. 1). Hot springs are found in all of these river valleys. Springs occur as clusters, with one or more seeps and pools located within 10 to 100 m of each other, typically along river banks, and discharge from both bedrock and fluvial deposits. Springs in this study are hosted in both gneisses of the Lesser Himalaya (LH) and phyllites and schists of the High Himalayan Crystalline formation (HHC).

2. Geothermal fluid compositions and temperatures

Spring fluids sampled for this study have measured exit temperatures ranging from 25 to 70 °C. The chemistry of central Nepal hot springs has been studied by several authors (Bhattarai, 1980; Kotarba, et al. 1981; Grabczak and Kotarba, 1985; Kotarba, 1986; Evans, et al. 2001; Evans and Derry, 2002; Evans, 2003; Becker et al., 2008). The spring fluids are sodium–potassium bicarbonate waters, supersaturated with carbon dioxide. Application of standard Si, and Ca-corrected alkali chemical geothermometers (Fournier, 1981) to spring chemistry produces estimates for equilibrium reservoir temperatures from 60–125 °C. These results are consistent with those of other workers (Kotarba, 1985; Becker et al., 2008), however these estimates represent minima because unavoidable mixing with surface waters commonly biases the temperature down. Equilibrium modeling of Ge/Si systematics yielded estimated reservoir temperatures of 200 ± 50 °C (Evans and Derry, 2002). Simple scaling relationships (Turcotte and Schubert, 2002) that suggest \( T_{res} > 2 \times T_{exit} \) in the absence of dilution by cooler groundwaters imply \( T_{res} \geq 100 \) to \( \geq 140 \) °C. Comparison with explored geothermal systems along strike in India also suggest \( T_{res} > 150 \) °C (Shankar et al., 1991). We take 150° ± 25 °C as our estimate for the hydrothermal reservoir temperature, while recognizing that there may be significant along-strike variation.

2.1. Isotopic constraints on water–rock interaction and recharge zone

Oxygen and hydrogen isotope data from the springs (Grabczak and Kotarba, 1985; Evans et al., 2008) indicate that the hydrothermal waters are meteoric in origin. The data define a line (\( \delta D = 8.16 \times \delta ^{18}O + 10.7 \)) which is very close to the world meteoric water line and to a regional correlation observed in the Kali Gandaki system (Garzione, et al. 2000). The absence of significant oxygen enrichment relative to deuterium indicate that water–rock interaction and recrystallization are limited. This is consistent with reservoir temperatures below 200 °C and may indicate relatively short residence time of the water in the hydrothermal system.

Grabczak and Kotarba (1985) also reported tritium contents for fourteen springs sampled in 1977 and 1980 from the Kali Gandaki, Seti Khola and Trisuli geothermal systems, which provide an additional constraint on the transit time of the spring water. They found tritium contents with a range of 7–45 TU (tritium units) for MCT springs and 60–68 TU for springs in the Thakola graben. As they note, the interpretation of these data in terms of a transit time for the meteoric water is non-unique. In the simplest case of a single component, values at the higher end of the range (>60 TU) would correspond to a time scale on the order of one half life of \(^3\)H (12.3 yr). The lower values (7–45 TU) could indicate a two component mixing of young meteoric water and an older “tritium-free” water equilibrated with atmospheric \(^3\)H prior to the bomb era (i.e. before ca. 1952). In both cases the presence of detectable tritium implies the presence of a component of the fluid with isolation times from the atmosphere no longer than 1–2 half lives of \(^3\)H (12–24 yr).

The \( \delta D \) data provide an additional constraint on the recharge zone of the meteoric water in the spring fluids. The deuterium content of stream waters in the Himalaya follows a simple empirical relationship with the mean elevation of the catchment zone upstream of the sampling point (Garzione et al., 2000; Gajurel et al., 2006). At a given sampling elevation, the \( \delta D \) values of the central Nepal hot springs are higher than the streams. The relatively heavy hydrogen in the springs

![Fig. 1. Geologic map of Narayani basin, central Nepal. Position of major hot spring systems marked with stars. Sample locations yielding key fluid inclusion data labeled.](image-url)
implies that the mean elevation of the recharge zone of the springs is significantly lower than the mean elevation of the stream catchment above the sampling point. That is to say that the springs recharge from a zone that is much more local than the catchment as a whole.

Strontium isotope data support local derivation of the hot spring fluids (Evans et al., 2004). There are large contrasts in $^{87}\text{Sr}/^{86}\text{Sr}$ between the three main lithologic units (Lesser Himalaya (LH), High Himalayan Crystalline series (HHC), and Tethyan Sedimentary Series (TSS)) in the basins. In each case the $^{87}\text{Sr}/^{86}\text{Sr}$ of the hot spring fluids is similar to the host lithology where the spring exits, although in some cases the map distance to other potential sources of Sr is only a few km.

The available chemical, $\delta^{18}\text{O}$, $\delta^3$T and $^{87}\text{Sr}/^{86}\text{Sr}$ data on the Nepalese springs indicate that the spring fluids are locally derived meteoric water, have not been penetrated deeply into the crust nor had extended residence times at high temperatures, and have reacted with lithologies similar to those observed at the spring exits. Consequently, the heat carried by the geothermal fluids must have been extracted from relatively shallow crustal levels within a few to 10 km of the discharge point.

2.2. Geochemical mass balance for spring flux

We have developed a chemical mass balance technique for estimating the flux of hydrothermal waters into the major river basins of central Nepal (Evans et al., 2004). This approach takes advantage of distinct and reliable differences between the chemistry of conservative components in the hot spring fluids and in the tributaries that form the river discharge. Hydrothermal fluids in a wide range of settings, including the Himalaya, have high germanium–silicon ratios ($\text{Ge}/\text{Si}$) from 30 to 1100 µmol/mol which contrast sharply with local unpolluted rivers yet reported. These high values can be estinated from a mixing model for $\text{Ge}/\text{Si}$ ratios in flu- ids adequately modeled by mixing tributary water with low silica fluids (Evans and Derry, 2002). Downstream from (3) and by taking the average hot spring values for each basin, estimated the minimum $P$–$T$ conditions of fluid immiscibility (and trapping) at 280–310 °C and 700 ± 300 bars based on total homogenization temperatures and the position of the $\text{H}_2\text{O}$–$\text{CO}_2$ solvus for the NaCl-free system. He also established a maximum pressure of 1600 bars for fluid trapping based on a maximum density isochore of late stage $\text{CO}_2$ fluid inclusions. Since this work, improvements in modeling of the $\text{H}_2\text{O}$–$\text{CO}_2$–NaCl system (Duan et al., 1995) allow more accurate estimates of trapping pressure (Darling and Bassett, 2002).

$\text{H}_2\text{O}$–$\text{CO}_2$–NaCl fluid inclusions in quartz samples from seven locations in the Kali Gandaki and Trisuli valleys were studied. The fluid inclusions are characterized by three phases at room temperature (25 °C), an aqueous phase containing $\text{H}_2\text{O}$ and dissolved NaCl, liquid $\text{CO}_2$, and vapor $\text{CO}_2$. The $\text{CO}_2$ component occupies anywhere from 40 to 70 volume % of the inclusions (Fig. 3) at room temperature. At high

Table 1

<table>
<thead>
<tr>
<th>River</th>
<th>Mean annual HS discharge (m³ s⁻¹)</th>
<th>±</th>
<th>Thermal power, MW</th>
<th>±</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myagdi</td>
<td>0.36</td>
<td>0.22</td>
<td>60</td>
<td>37</td>
</tr>
<tr>
<td>Kali</td>
<td>1.25</td>
<td>1.32</td>
<td>211</td>
<td>221</td>
</tr>
<tr>
<td>Modii</td>
<td>0.33</td>
<td>0.15</td>
<td>55</td>
<td>24</td>
</tr>
<tr>
<td>Seti</td>
<td>0.05</td>
<td>0.03</td>
<td>8.5</td>
<td>4.6</td>
</tr>
<tr>
<td>Maryandy</td>
<td>0.29</td>
<td>0.26</td>
<td>49</td>
<td>43</td>
</tr>
<tr>
<td>Bhardi</td>
<td>0.04</td>
<td>0.02</td>
<td>6.5</td>
<td>3.9</td>
</tr>
<tr>
<td>Trisuli</td>
<td>0.66</td>
<td>0.68</td>
<td>111</td>
<td>114</td>
</tr>
<tr>
<td>Langtang</td>
<td>0.01</td>
<td>0.003</td>
<td>1.2</td>
<td>0.6</td>
</tr>
<tr>
<td>SUM</td>
<td>2.98</td>
<td>1.57</td>
<td>501</td>
<td>257</td>
</tr>
</tbody>
</table>

Thermal power calculated from Eq. (3). Uncertainty on integrated fluxes calculated with standard error propagation techniques (Bevington, 1969).

3. Fluid inclusion constraints on the geothermal gradient

Fluid inclusions in quartz veins from central Nepal have been studied and interpreted by Pêcher (1979), Sauniac and Touret (1983), Craw (1990), and Boullier et al. (1991). These studies describe a variety of inclusion types (e.g. $\text{H}_2\text{O}$, $\text{H}_2\text{O}–\text{NaCl}$, $\text{H}_2\text{O}$–$\text{CO}_2$, $\text{H}_2\text{O}$–$\text{CO}_2$–NaCl immiscibility in Nepal quartz veins and estimated the $\text{H}_2\text{O}$–$\text{CO}_2$ solvus for the NaCl-free system. He also established a maximum pressure of 1600 bars for fluid trapping based on a maximum density isochore of late stage $\text{CO}_2$ fluid inclusions. Since this work, improvements in modeling of the $\text{H}_2\text{O}$–$\text{CO}_2$–NaCl system (Duan et al., 1995) allow more accurate estimates of trapping pressure (Darling and Bassett, 2002).

$\text{H}_2\text{O}$–$\text{CO}_2$–NaCl fluid inclusions in quartz samples from seven locations in the Kali Gandaki and Trisuli valleys were studied. The fluid inclusions are characterized by three phases at room temperature (25 °C), an aqueous phase containing $\text{H}_2\text{O}$ and dissolved NaCl, liquid $\text{CO}_2$, and vapor $\text{CO}_2$. The $\text{CO}_2$ component occupies anywhere from 40 to 70 volume % of the inclusions (Fig. 3) at room temperature. At high

Fig. 2. Cross-cutting quartz vein in LH phyllites near Syabru Bensi, Nepal.
temperature (>200 °C), many of these inclusions homogenize by shrinking of the CO2 phase (i.e. homogenization to liquid). However, some homogenize by expansion of the CO2 phase (i.e. homogenization to vapor), suggesting the trapping of immiscible H2O–CO2–NaCl fluids.

We noted that fluid inclusions within a single fluid inclusion array (FIA) generally displayed one type of homogenization behavior, that is, all the fluid inclusions homogenize to liquid or all the fluid inclusions homogenize to vapor. This could be interpreted as the trapping of an immiscible fluid, however, one cannot rule out the possibility that the fluids were trapped at different times and under different PT conditions. Consequently, we focused our study on identifying FIA’s that showed both types of homogenization behavior. Such FIA’s are the best indicator of trapping of immiscible fluids. Of the seven locations studied, only three showed FIA’s with direct evidence of fluid immiscibility (GA-10; MLB-33, MLB-89, Fig. 1). Fig. 3 shows one such FIA in sample MLB-89 (from the Modi Khola valley, Annapurna Himal).

Fluid inclusions in quartz samples from the three locations were analyzed on a U.S.G.S gas-flow heating and cooling stage (Werre et al., 1979). The stage was calibrated with synthetic fluid inclusions and the microthermometric data are believed accurate to +0.1 °C for low temperature measurements (<40 °C) and ±2.5 °C for high temperature measurements (>200 °C).

Measured CO2 melting temperatures fall within 0.1 °C of pure CO2 (−56.6 °C) for only two of the three locations studied (MLB-33 and MLB-89). CO2 melting temperatures from the immiscible FIA in sample GA-10 averaged −57.0 °C indicating the presence of another gas phase. Raman spectroscopy of fluid inclusions from sample GA-10 revealed the presence of CH4, but no CH4 was detected in fluid inclusions from samples MLB-33 and MLB-89. CO2 was confirmed in all inclusions by Raman spectroscopy.

Table 2 shows measured microthermometric data and calculated parameters for FIA’s in samples MLB-33 and MLB-89 only. Final clathrate melting temperatures range from +8.5 to +9.3 °C and indicate salinities of 3.02 to 1.44 wt.% NaCl relative to H2O (Darling, 1991). CO2 homogenized to liquid at temperatures ranging from +25.9 to +30.4 °C indicating moderate CO2 densities (Ely et al., 1989) as modified by (Sterner and Bodnar, 1991). Total fluid inclusion homogenization temperatures varied depending on the mode of homogenization. Those that homogenized to liquid did so between +295.1 and +306.6 °C, whereas those homogenizing to vapor did so between +310.0 and +315.5 °C (Table 2). In theory, the trapping of immiscible fluids on either side of the H2O–CO2–NaCl solvus should result in identical total homogenization temperatures. However, because H2O preferentially “wets” the walls of all inclusions, it occurs in excess in those inclusions that homogenize to the vapor state, and consequently results in slightly higher total homogenization temperatures (Roedder and Bodnar, 1980).

It is, therefore, total homogenization temperatures recorded from inclusions that homogenize to the liquid state that are believed to be reflective of actual temperatures of trapping. Table 2 shows calculated pressures of homogenization, inclusion compositions, and inclusion molar volumes for only those inclusions that homogenized to the liquid state. Because these are immiscible fluids, the total homogenization temperatures and calculated homogenization pressures are interpreted to be the temperatures and pressures of trapping (Roedder, 1984).

The methods used in calculating the parameters in Table 2 are described in Darling and Bassett (2002). In brief, the microthermometric data are used in combination with an equation of state (EOS) for the H2O–CO2–NaCl system (Duan et al., 1995) to calculate the fluid composition, pressure (on the solvus) at homogenization, and the molar volume of the liquid phase. Fluid inclusions that homogenized to liquid have compositions averaging XH2O = 0.624, XCO2 = 0.371, XNaCl = 0.005 and molar volumes averaging 37.95 cm3/mol. Calculated pressures of homogenization range from 936 to 1160 bars, and average 1056 bars. Although the amount of data in Table 2 is small, the inclusions demonstrate strong evidence of fluid immiscibility and have compositions that can be modeled in the ternary H2O–CO2–NaCl system.

Our inferred temperature range of trapping (295.1 to 306.6 °C) is similar to that determined by Craw (1998) (i.e. 280° to 310 °C) and our inferred pressure range of trapping (936 to 1160 bars) falls between Craw’s maximum of 1600 bars and minimum of 700 ± 300 bars. Our calculated immiscibility pressures are about 350 bars higher than estimated because we account for the strong effect of NaCl on the position of the H2O–CO2 solvus (Takenouchi and Kennedy, 1965; Gehrig et al., 1986).

Some uncertainty with the equation of state in Duan et al. (1995) has been recently noted, particularly at temperatures between 300 and 400 °C (Schmidt and Bodnar, 2000; Hu et al., 2007). It is difficult to assess the quality of the EOS model in this application because there is little or no PVTx and solvus data in the ternary H2O–CO2–NaCl system at salinities between 0 and 6 wt.% NaCl (relative to H2O) and the location of the solvus varies widely as a function of pressure and NaCl content. However, a minimum trapping pressure of 600–700 bars (at 300 °C) can be inferred based on the location of the NaCl-free H2O–CO2 solvus at XClO2 values of 0.2 to 0.4 (Schmidt and Bodnar, 2000). Given the present lack of experimental data in ternary system between 0 and 6 wt.% NaCl, and the models derived from it, the calculated parameters listed in Table 2 should be viewed as only a good approximation.

Inferred temperatures and pressures of trapping average 301 ± 6 °C and 1056 ± 110 bars. Assuming lithostatic load, the trapping pressures correspond to a depth of 3.9 km ± 0.4 km. For a mean surface temperature of 10 °C we obtain a geothermal gradient of 75 ± 78 Ckm−1 over the upper 3.9 km of the crust (assuming a steady state geotherm from the time of fluid trapping to today). Despite the uncertainties, our pressure estimates indicate substantially elevated heat flow at the time of trapping, consistent with a young orogen. Furthermore, our PT determinations for undeformed quartz veins are consistent with the clockwise retrograde metamorphic path modeled by Beaumont et al. (2001). Our estimates are also similar to those recently determined for fluid inclusions from the eastern Himalayan syntaxis (Craw et al., 2005).

Stable isotope analysis of fluid inclusions from these same samples extracted by crushing shows that the trapped fluids are non-meteoric and are consistent with derivation from an ascending metamorphic fluid (Evans et al., 2008). The depth of trapping (3.9 km) thus indicates a likely lower limit on the depth of circulation of meteoric water.

4. Heat transfer by meteoric hydrothermal flow

Meteoric hydrothermal circulation transfers heat from the high temperature reservoir to the surface. Some of this heat is lost to rocks of the uppermost crust by conductive cooling, as the fluids cool from a reservoir Teq ≈ 150 °C to an exit Texit = 40 to 70 °C. The rest is lost as the exiting fluids mix with cool surface water (Tsurf ≈ 10 °C). Adiabatic...
temperature losses from fluid ascent to the surface may be estimated by

\[
\frac{dT}{dz} = \frac{\alpha_v \cdot g \cdot T}{C_p}
\]

where \(\alpha_v\) = coefficient of thermal expansion = \(6 \times 10^{-4} \degree \text{C}^{-1}\) at \(T = 100 \degree \text{C}\) (see Table 3 for list of symbols). For a circulation depth < 3 km such losses are small (<2 °C) and may be ignored.

Conductive cooling of the fluids during ascent through the crust can be constrained from the initial fluid reservoir T and the exit T at \(T_{exit} = 60 \pm 70 \degree \text{C}\). As discussed above, we note that \(T_{exit}\) may well be >150 °C for some systems. Combining these values with the observed hydrothermal water flux:

\[
\Delta H_{\text{cond}} = \int_{T_{\text{in}}}^{T_{\text{exit}}} C_p dT = 3.2 \times 10^7 \text{J} \cdot \text{kg}^{-1}
\]

For an integrated geothermal water flux \(j_{\text{HW}}\) across the Narayani basin of \(3 \pm 1.5 \text{ m}^3 \text{s}^{-1}\) (Evans et al., 2004) this term yields 1260 MW (±716), which is transferred to the shallow crust from the geothermal reservoir.

\[ \Delta P_{\text{surf}} = Q_{\text{HW}} \cdot \int_{T_{\text{in}}}^{T_{\text{exit}}} C_p dT = 5.0 \times 10^8 \text{W} \]

for \(T_{\text{surf}} = 10 \degree \text{C}\). The computed heat flux into the individual basins studied ranges from 1 to 211 MW, for a total heat flux across the Narayani system of 501 ± 257 MW (Table 1). This value represents only heat advected to the surface by the hydrothermal circulation. The two river systems with the largest thermal power output are the Kali Gandaki and the Trisuli. These two valleys contain the highest temperature springs we have observed \((T_{\text{exit}} = 60–70 \degree \text{C})\). Both rivers also pass through extensional graben systems with active hydrothermal systems before entering the MCT zone. In the case of the Trisuli, most of the hydrothermal activity appears to be located near the MCT at Syabru Bensi but the upstream region has been little explored for geothermal purposes and the extent of hydrothermal activity in the graben structure is not well known. In the Kali Gandaki, significant hot spring systems occur in the Thakola graben, but a major locus of geothermal activity is at Tatopani near the MCT. We cannot at present quantitatively separate the thermal contribution from hydrothermal systems associated with the graben from those associated with the MCT zone, but in both cases the springs in the MCT zone appear to carry the majority of the heat output.

The net thermal power transfer across the Narayani basin is \(\approx 1760 \pm 953 \text{ MW}\) of which 500 MW is direct loss to the surface. The remainder heats the shallow crust, steepens the local geothermal gradient, and is lost by conduction to the surface or possibly by diffuse fluid flow into groundwater that we do not observe. The two largest sources of uncertainty are the calculated spring fluxes (±50%) and the reservoir temperatures (±25 °C). We have used standard error propagation methods (Blevington, 1989) in assessing their combined uncertainties of ≤ 60%. There are inherent limitations on the precision of our mass balance approach, and it is likely that significant improvement would require either intensive long term monitoring and/or drilling to obtain independent estimates on flow rates and down-hole temperatures. The latter would be particularly helpful constraining the extent of lateral transport of fluids and heat. We here use a simple 1-D model, while recognizing that this can only yield a first approximation of fluid heat transfer in a complex 3-D system.

The along-strike region defined by the Kali Gandaki to Trisuli rivers is ca. 200 km in length. Following Seeber and Görnitz (1983), Lave and Auvouic (2001) propose a 40–50 km wide swath of rapid uplift near
the Himalayan front, beginning roughly coincident with the surface trace of the MCT. As noted above, the stable isotope data suggest derivation of the heat-dissipating meteoric fluids from an even more limited source area. As a first order estimate, we assume that the thermal power carried by the hydrothermal fluids is extracted from a region 200 km × 45 km = 9000 km². For a thermal power of 1760 MW, this region yields an areally averaged heat flow \( \phi_{\text{basal}} = 196 \text{ mW m}^{-2} \). Clearly this calculation is sensitive to the somewhat arbitrarily chosen surface area, but is quite similar to that estimated for the regional heat flow near the Puga geothermal field in India (Harinarayana et al., 2006). The calculations are designed to illustrate the scale of relevant processes rather than provide precise estimates of thermal transfer.

4.1. Geothermal gradient and heat flow

We estimate the thermal structure of the upper crust using the \( P, T \) data at the surface and at 3.9 km from the fluid inclusions. We assume that \( T = f(z) \) has an exponential form:

\[
T(z) = \left( T_{\text{basal}} - T_{\text{surf}} \right) e^{\left( -\frac{z}{h} \right)}
\]

(5)

where \( T_{\text{basal}} \) is taken to be 500 °C (Cattin and Avouac, 2000). We estimate the scale depth \( h \) to be near 4.3 km by fitting this function to the data (Fig. 4). Differentiating to obtain the heat flow:

\[
\phi = -k \left( \frac{\partial T}{\partial z} \right) = -k \frac{\left( T_{\text{basal}} - T_{\text{surf}} \right)}{h} e^{\left( -\frac{z}{h} \right)}
\]

(6)

where \( k = 2 \text{ W K}^{-1} \text{ m}^{-1} \) is the thermal conductivity. At the surface \( \phi = 228 \text{ mW m}^{-2} \). This figure is close to the estimated hydrothermal heat transfer of 196 mW m\(^{-2}\) but the uncertainties are such that this apparent close agreement between the heat flow estimated from the thermal profile and from measurement of the hot spring flux may be in part fortuitous.

4.2. Sources of heat for hydrothermal systems

Geothermal activity is widespread along the Himalayan front, although there appear to have been few if any systematic attempts to document the extent of such activity on an orogen-wide scale. An early publication catalogs >200 springs in colonial India (and thus excludes Nepal) distributed along the Himalayan front (Oldham, 1883). There is no evidence of magmatic activity associated with these hydrothermal systems.

The location of the central Nepal springs with respect to topographic and structural features provides some clues to the possible mechanisms for driving hydrothermal flow. The central Nepal hot springs are primarily confined to major cross-cutting valleys in zones of high relief; at or near the Main Central Thrust (MCT) shear zone. A hypothesis for the zone of high hydrothermal heat flux is that rapid uplift and incision of rocks in major N–S valleys that cut the MCT leads to the creation of a dynamically maintained steep thermal gradient. In this scenario rocks from mid-crustal depths are brought to shallow levels by coupled uplift-erosion processes, and provide a heat source to drive meteoric hydrothermal circulation. One prediction of this model is that the springs should be located in areas of rapid incision. Lave and Avouac (2001) have calculated river incision rates for four rivers in the region of interest using a modified stream power model. There is a good correspondence between hot spring locations and calculated incision rate along those major rivers (Kali Gandaki, Marsyangdi, Buri Gandaki, Trisuli). A second observation is that the spring locations are, in most cases, within 10 km of the mapped trace of the MCT. It is possible that fracturing associated with the MCT serves as a convenient structural conduit for fluid flow, and/or that interaction with HHC rocks at depth is an important source of heat. It may simply be that high rates of rock uplift in this zone brings heat to the near sub-surface rapidly. It is worth noting that the MCT zone is complex, often several kilometers wide at the surface. We do not have a detailed picture of the structure of the HHC zone at depth, especially in the vicinity of a proposed crustal ramp. Finally, we note that the calculated incision rates for the Buri Gandaki are the lowest of the major rivers in the area (Lave and Avouac, 2001) and the Buri has the fewest known springs and lowest estimated heat flow.

4.3. Heat sources

We can ask what heat production is necessary to support the observed hydrothermal heat flow. We take the depth of meteoric circulation to not exceed 3 km, thus the fluids extract their heat from the upper 3 km of the crust yielding:

\[
\phi = \frac{196 \text{ mW m}^{-2}}{3000 \text{ m}} = 65 \mu \text{W m}^{-2}
\]

(7)

Radiogenic heat production is \( \Phi_{\text{rad}} \approx 3.2 \mu \text{W m}^{-3} \) based on analysis of heat producing elements in river sands from our study area. Two samples of river sand from the Chepe khola basin within the HHC yielded an average composition of K\(_2\)O = 2.2 wt %, Th = 12.4 ppm, and U = 2.2 ppm. The required heat production greatly exceeds the possible radiogenic contribution irrespective of the details of our assumptions about the depth of fluid circulation. The balance of the “heat production” required to support the observed heat loss may come from diffusion of heat from below 3 km and from tectonic advection of hot rock to the surface. The latter may be written in terms of the uplift/erosion rate and the geothermal gradient as:

\[
\frac{\partial T}{\partial t} = \frac{C_P}{V} \frac{\partial V}{\partial t} = \frac{1}{V} \frac{\partial H}{\partial t} = \Phi_{\text{adv}}
\]

(8)

where \( C_P = 2.7 \times 10^4 \text{ J K}^{-1} \text{ m}^{-3} \) is the heat capacity of rock. With an average geothermal gradient estimated from the fluid inclusions:

\[
\frac{\partial T}{\partial z} = 75 \text{ °C km}^{-1}
\]

(9)

and an uplift-erosion (rock advection) rate from (Lave and Avouac, 2001) of

\[
\frac{\partial z}{\partial t} = 5 \text{ mm yr}^{-1}
\]

(10)
the “heat production” produced by advection of material into the upper 3 km region of the crust accessible to meteoric circulation is:

$$\phi_{\text{adv}} = \frac{1}{V} \frac{\partial h}{\partial t} \approx 32 \mu W m^{-3}$$

(11)

The ratio of “heat production” provided by radioactive decay and crustal advection to that necessary to support the observed fluid heat flow is:

$$\frac{\phi_{\text{fluid}}}{\phi_{\text{adv}}} = \frac{32 + 3.2 \mu W m^{-3}}{65 \mu W m^{-3}} = 0.54$$

(12)

While this is clearly a very simplified calculation it suggests that advection of hot rock by coupled uplift/erosion can supply a significant fraction of the observed heat. The conductive heat flow estimated at the base of the fluid circulation zone \((z=3 \text{ km})\) is 113 mW m\(^{-2}\) (Fig. 4), sufficient to supply the remaining heat input into the portion of the upper crust sampled by the meteoric fluids. We conclude that the observed fluid heat transfer is consistent with the steep geothermal gradient implied by the fluid inclusion data and with the tectonic advection of heat to the shallow crust. The hydrothermal circulation is an efficient means of mining and then dissipating heat at the surface provided by these processes.

4.4. Shallow thermal structure in zones of active fluid flow

The high heat flow and steep geothermal gradient we observe can impact the interpretation of data from low-temperature thermochronometers, such as U-Th-He dates in apatite and zircon. Given that the nominal closure temperatures of U-Th-He in apatite and zircon are ca. 70 °C and 110 °C respectively, and that we observe hot spring exit temperatures approaching 70 °C, the influence of fluid circulation on such ages clearly can be significant. Individual hot spring systems are geologically ephemeral. Recent U series dating of travertine from a geothermal system in southern Tibet implies that individual springs may function on the order of 10 kyr (Zentmyer et al., 2008). This spring is located in a graben north of the steep river reaches, protecting it from erosion. Travertine deposits along the main river valleys are subject to rapid erosion and removal, and so do not accumulate as do the ones located in the northern grabens. It is very likely that spring locations migrate up and down the valleys on a 10 kyr time scale as hydrothermal precipitation cements fluid conduits in one location but stress release and fracturing develop new pathways in another. Thus while the area impacted by fluid flow at any one time may be modest, it is probable that most of the geomorphologically active steep stream reaches have experienced hot fluid flow in the recent geological past. Explored geothermal systems in the NW Indian Himalaya indicate the scale of the surface area impacted by geothermal flow. Geothermal exploration wells near the Puga geothermal system, Ladakh, which is similar in many respects to the Nepali geothermal systems, show strongly elevated subsurface temperatures \((T>100 °C \text{ at } 70 \text{ m depth})\) along a 3–4 km reach of the Puga river valley (Shankar et al., 1991).

We modified a simple 1-D model of steady fluid flow in a geothermally heated porous medium from Turcotte and Schubert (2002) to examine the subsurface temperature distribution in a zone of active fluid flow. Given that we lack subsurface data from the active geothermal systems in Nepal the simplicity of this model appears to outweigh its limitations for our present purposes. For a 1-D case the temperature may be expressed as:

$$T(z) = T_{\text{res}} -(T_{\text{res}}-T_{\text{surf}}) \exp\left(-\frac{\rho c_p \alpha (K \Delta T)}{\lambda_m} z\right)$$

(13)

where the fluid density \(\rho = 1000 \text{ kg m}^{-3}\), \(\lambda_m\) is the volumetrically averaged thermal conductivity of the medium \(\approx 3.35 \text{ W m}^{-1} \text{K}^{-1}\), and \(K\) is the hydraulic conductivity. Whipp and Ehlers (2007) estimated the hydraulic conductivity for the central Nepal region using a thermokinematic model with an embedded fluid flow model. They calibrated a value for \(K\) to satisfy the observed surface hydrothermal heat flux derived from the data for the Marsyandi basin in (Evans et al., 2004), using data and a procedure essentially the same as outlined above. They found a best-fit value for \(K \approx 5 \times 10^{-7} \text{ m s}^{-1}\). We calculated the subsurface temperature distribution from Eq. (4) using \(K = 5 \times 10^{-7} \text{ m s}^{-1}\) (Fig. 5). With this value of \(K\) the temperature gradient is steep, and although the calculated fluid temperatures do not reach the two phase curve, the results suggest that boiling, particularly of CO\(_2\)-rich fluids, might take place in localized zones of rapid fluid flow. However there is no evidence of boiling in the Nepal hot spring fluid chemistry. With a \(K = 1 \times 10^{-7} \text{ m s}^{-1}\) the fluid cooling path is less steep in the final few hundred meters and the fluids are unlikely to boil. This lower value of \(K\) is within the range of uncertainty given by Whipp and Ehlers (2007) while still well above their threshold for fluid advection-dominated heat flow of \(10^{-7} \text{ m s}^{-1}\), and may be a better estimate for the “average” hydraulic conductivity near the MCT in central Nepal.

Along the fluid cooling path shown in Fig. 5 the fluids rise nearly isothermally to above 500 m, and in this respect it differs significantly from that obtained by Becker et al. (2008) using a pipe flow conductive cooling model. They found a much more gradual path, with reservoir temperatures of 125 °C below 2000 m, cooling to 40–60 °C below 500 m. The overall shape of their cooling path is concave-up. Neither of these features resemble most explored geothermal systems, and they are unlike temperature-depth profiles from wells drilled into geothermal systems in northwest India that in many respects are similar to those of central Nepal (Shankar et al., 1991; Chandrasekharan, 2000). The upper boundary of the porous media model is the ambient \(T_{\text{surf}}\) used in the pipe flow model, but this small difference cannot explain the fundamental difference in behavior. We suggest that a porous media model is a much more realistic representation of the fluid temperatures at depth. The inability of the pipe flow model of Becker et al. (2008) to produce realistic temperature profiles has implications for their calculation of CO\(_2\) solubility and degassing behavior, since these quantities are strong functions of the temperature profiles.
With our range of estimated hydraulic conductivities calculated fluid temperatures reach 100 °C within 100–200 m of the surface, as observed at Puga. These very steep temperature gradients are only valid for the zone of active fluid flow, which as noted above may be a zone with a length scale (\( \theta \)) of 1 km. As the geothermal activity migrates along the valley axis a significant reach will have been impacted by fluid flow and consequent steep sub-surface temperature gradients. It appears likely that low temperature thermochromometers can be impacted by hydrothermal fluid flow in this environment, as originally suggested by (Copeland et al., 1991).

5. Conclusions and implications

Geothermal systems along the Himalayan front in central Nepal deliver 500 ± 257 MW of thermal power to the surface in hot springs. Another 1260 ± 716 MW is transferred by fluid flow to the uppermost crust via conductive cooling of geothermal fluids, for a total of 1760 ± 973 MW. Along 200 km of range front length and across a 45 km wide active zone this heat transfer implies an areally averaged heat flow of ca. 196 mW m⁻². Fluid inclusion data define P-T conditions of trapping of a CO₂-rich metamorphic fluid at 1056 ± 10 bars and 301 ± 6 °C, implying a mean geothermal gradient of 75 °C km⁻¹. A simple exponential temperature model indicates surface heat flow \( \approx 228 \text{ mW m}^{-2} \), of which \( \approx 85\% \) is transferred by fluid circulation. Radiogenic heat sources in the upper crust can only support a small fraction (\(<5\%) of the fluid-driven heat transfer. Simple calculations suggest that advection of hot rock from depth can support another 50%, while conductive cooling with the steepened geotherm can support the rest. Thus within the uncertainties of the available data the thermal balance of the zone of geothermal activity could be considered to be in a kind of steady state between tectonically advected heat inputs and fluid-adveective and conductive losses.

Models that consider erosion as a driving mechanism of Himalayan tectonics share several characteristics. Rapid denudation along the Himalayan front should be approximately balanced by the advection of crust into the zone of rapid erosion — otherwise either the range front must prograde to the south, or retreat to the north. Progradation or retreat would be uneven, as rivers with higher incision rates should be more deeply embedded in the range front than rivers with lower incision rates. As pointed out by Lave and Avouac (2001) this is not the case, implying a dynamic equilibrium between crustal inputs and erosional losses. Such a “steady state” (sensu lato) balance between crustal advection and erosion along the Himalayan front has been proposed by several groups (Avouac and Burov, 1996; Willett, 1999; Beaumont et al., 2001). Channel flow type models realize stabilizing feedback by the displacement of the brittle-ductile transition toward the surface in the zone of rapid denudation. This weakened zone of crust becomes the preferred flow path for deformation, and hot rocks move rapidly toward the surface, maintaining the shallow isotherms. These shallow isotherms drive fluid flow, as proposed by Koons and Craw (1991). The majority of Himalayan hot springs are found along major incised valleys, in zones of steep river reaches and rapid fluvial downcutting. Most are located within or near the MCT zone and are associated with strong gradients in range-front topography and river profiles. The position of the springs is consistent with the view that the heat is supplied to the meteoric system by tectonic advection of hot rock to the sub-surface. Here we show that the heat transfer by this fluid flow is substantial and efficiently extracts a large fraction of the tectonically advected heat.

Geothermal activity results in very steep temperature gradients in the zone of active fluid flow, with temperatures exceeding 100 °C in the upper 50–200 m of the crust. These high temperatures at such shallow levels have obvious implications for the interpretation of low-T thermochromometer data, particularly when sampled near river valleys. It is likely that at least some of the young cooling ages reported from Himalayan samples reflect perturbation by hot fluids and not merely conductive cooling.

To our knowledge, this geothermically based heat flow estimate is the first direct estimate of heat loss from the Himalayan front, and suggests that additional studies of geothermal systems can provide important geophysical constraints on the shallow thermal structure of active orogens.

Acknowledgements

We wish to thank two anonymous reviewers who made very valuable comments that improved the manuscript. We thank members of the Canadian Institute for Advanced Research ESEP Program and Jerome Lavé for insightful discussions about Himalayan tectonics, erosion and heat flow. Field and lab work for this project was supported by U.S. NSF–EAR grants 0087671 and 0090116 and by the French INSU program PNSE.

References


Ehlers, J.F., Haynes, W.M., Bain, B.C., 1989. Isotopic (p. Vm) measurements on CO₂ and on (0.982CO₂ + 0.018 N₂) from 250 to 330 at pressures to 35 Mpa. J. Chem. Thermodyn. 21, 879–894.


