Hydrothermal circulation and the thermal structure of shallow subduction zones

Robert N. Harris¹, Glenn A. Spinelli², and Andrew T. Fisher³

¹College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, 104 CEOAS Administration Building, Corvallis, Oregon 97330, USA
²Department of Earth and Environmental Science, New Mexico Institute of Mining and Technology, 801 Leroy Place, Socorro, New Mexico 87801, USA
³Department of Earth and Planetary Sciences, University of California, Santa Cruz, 1156 High Street, Earth and Planetary Sciences, Room A232, Santa Cruz, California 95064, USA

ABSTRACT

Hydrothermal circulation within oceanic basement can have a profound influence on temperatures in the upper crust, including those close to the subduction thrust and in the overlying plate. Heat flow evidence for hydrothermal circulation in the volcanic basement of incoming plates includes: (1) values that are well below conductive predictions due to the advection of heat into the ocean, and (2) variability about conductive predictions that cannot be explained by variations in seafloor relief or thermal conductivity. In this review we summarize evidence for hydrothermal circulation in subducting oceanic basement from the Nankai, Costa Rica, south-central Chile, Haida Gwaii, and Cascadia margins and explore its influence on plate boundary temperatures. Models of these systems using a high Nusselt number proxy for hydrothermal circulation are used to illustrate the influence of this process on seafloor observations and thermal conditions at depth. We show that at these subduction zones, patterns of seafloor heat flow are best explained by thermal models that include the influence of hydrothermal circulation.

INTRODUCTION

A significant emphasis over the past two decades of subduction zone research has been on understanding transitions in frictional sliding behavior along the plate boundary (e.g., Hyndman, 2007; Lay et al., 2012). Temperature along plate boundary faults has emerged as an important parameter because the thermal state of this interface influences its rheology, including the updip and downdip limits of seismicity, which can vary as a result of thermally driven rock-fluid reactions and differences in friction (Hyndman and Wang, 1993; Oleskevich et al., 1999; Moore and Saffer, 2001). Measuring temperature at the plate interface, especially at depths of the seismogenic zone, is impractical in most settings because of the time requirements, expense, and technical difficulty of drilling deep enough to reach this part of an active subduction zone. Instead temperatures at depth are estimated through a combination of surface heat flow measurements and modeling.

A primary influence on subduction zone temperature is the thermal state of the incoming plate (e.g., Dumitrut, 1981; Molnar and England, 1985). In broad terms the thermal state of the incoming lithosphere is a function of age that is largely controlled by conductive cooling (e.g., Davis and Lister, 1974; Parsons and Sclater, 1977). In detail however, hydrothermal circulation and other environmental processes such as sedimentation can modify the thermal state of the incoming igneous crust sufficiently so as to have a significant impact on plate boundary temperatures. Therefore, understanding the nature and impact of hydrothermal circulation on the incoming lithosphere temperature structure is a prerequisite for accurate thermal models of the shallow subduction zone. In thermal models, a portion of the effect of hydrothermal circulation can be represented by modifying the geothermal gradient (geotherm) within the incoming plate, which serves as either the seaward boundary condition for steady-state models or the initial condition for transient models.

In this review we focus on the influence hydrothermal circulation within incoming and subducting oceanic basement has on thermal conditions of the shallow subduction zone. Many of the parameters important to hydrothermal circulation during subduction are also of interest in studies of ridge-flank hydrothermal systems more generally, including crustal permeability, basement relief, and sediment thickness and continuity (e.g., Davis et al., 1989, 1999; Fisher et al., 2003a; Hutnak et al., 2007). It is worth noting that the subduction zones receiving much attention over the past two decades have tended to be subducting young igneous basement (e.g., Nankai, Costa Rica, central Chile, Cascadia), in part because warm and buoyant conditions make these systems logistically easier to characterize and sample through ship-based studies and drilling. Similar to ridge flanks, youthfulness of subducting oceanic basement facilitates hydrothermal circulation because of high permeability, large thermal gradients, and thin or incomplete sediment cover.

We start with a general discussion of the shallow thermal regime during subduction, and how this thermal regime can be perturbed by hydrothermal circulation in the upper oceanic basement. We present and discuss examples of thermal modeling studies from several convergent margins that illustrate key processes. We summarize these results, and review how thermal models of subduction can incorporate characteristics of hydrothermal circulation to assess where this process occurs.

THERMAL MODELS OF THE SHALLOW SUBDUCTION ZONE

Figure 1 shows a generic thermal model and the important parameters influencing surface heat flow and temperature along the subduction thrust. We
Figure 1. Subduction zone thermal model. (A) Generic thermal model of a subduction zone showing important parameters for thermal modeling (circ. = circulation). Although not depicted, thermal physical rock properties are also important. For the thermal regime of the shallow subduction zone and plate boundary fault, the least-well-known parameters are the seaward geotherm and magnitude of shear heating on the plate interface. (B) Uncertainties in seaward geotherms based on GDH1 model (Stein and Stein, 1992) for 10 Ma (red) and 100 Ma (blue) oceanic crust. Lines show mean and shaded areas show one standard deviation. (C) Heat flow across forearc for 10 (red) and 100 Ma (blue) crust. Shaded area shows uncertainty due to uncertainty in geotherm. Magenta and cyan lines show modeled heat flow resulting from insulated hydrothermal circulation in subducting oceanic crust. (D) Temperature along the plate interface for 10 (red) and 100 Ma (blue) crust. Shaded area shows uncertainty due to uncertainty in geotherm. Magenta and cyan lines show modeled heat flow resulting from insulated hydrothermal circulation in subducting oceanic crust.
use the algorithm of Wang et al. (1995) that solves the heat transfer equation assuming steady-state conditions, as applied to a two-dimensional trench normal profile at a subduction zone,

$$\nabla \cdot (\lambda \nabla T) - \rho cv \cdot \nabla T + A = 0, \quad (1)$$

where $\lambda$ is thermal conductivity, $T$ is temperature, $\rho c$ is volumetric heat capacity, $v$ is the convergence rate of the subducting plate, and $A$ is heat production. Mantle corner flow is incorporated by assigning an isoviscous rheology to the mantle wedge and coupling it to the subducting slab (Peacock and Wang, 1999).

The kinematic parameters, the convergence rate and slab dip, control how quickly cold crust and mantle material are advected downward. The convergence rate is usually well determined from geodetic measurements or models of plate motion, and the dip of the subduction thrust is usually well known from active and/or passive seismic experiments. The thermal field is influenced by thermophysical rock properties, the most important of which are the thermal conductivity and heat production. These properties can be determined by dredging or drilling to collect representative material, or estimated from catalogues of measurements coupled with knowledge of site lithology and structure. Our emphasis is on shallow subduction thrusts, an area that is largely unaffected by the mantle wedge flow field, a process that mainly influences thermal conditions landward of sites considered in the present review (e.g., Currie et al., 2004; Wada and Wang, 2009).

Less well determined parameters in these systems are the geotherm at the seaward edge of the model, and the effective coefficient of friction that controls the magnitude of shear heating across the main thrust and along secondary fault systems. In the absence of heat flow data from the incoming plate, the geotherm is often assumed to follow a conductive lithospheric reference model such as GDH1 (Stein and Stein, 1992). The geotherm governs the magnitude of the incoming heat flow and the heat content of the plate; this heat is advected downdip with the subducting plate and governs the thermal budget of the shallow subduction zone. As shown by case examples herein, uncertainties in the geotherm, including property variations in and out of the plane of two-dimensional simulations, can lead to significant uncertainties in the state of the shallow thermal regime.

Shear heating, $Q_s$, along the plate interface is a function of the effective coefficient of friction, $\mu'$, and can be expressed as (e.g., Wang et al., 1995),

$$Q_s = \tau v = \mu' \sigma_n \frac{v}{w}, \quad (2)$$

$$Q_l = \tau \dot{e} \quad (3)$$

for the brittle and ductile regimes, respectively, where $\tau$ is the shear stress, $\dot{e}$ is the strain rate, $\sigma_n$ is the normal stress across the fault, $v$ is the sliding velocity, and $w$ is the width over which shearing takes place. Evidence from thermal modeling, borehole measurements, and geomechanical laboratory tests suggests that the effective coefficient of friction along shallow subduction thrusts often has a low value ($\mu' < 0.2$) compared to a higher value ($\mu' \sim 0.6$) inferred from early laboratory experiments (Byerlee, 1978; Di Toro et al., 2004; Fulton et al., 2013; Gao and Wang, 2014; Wang and He, 2008). Numerical models show that frictional heating has a large effect on fault zone temperatures but may have a small effect on seafloor heat flow, particularly in the presence of hydrothermal circulation that can redistribute or extract excess heat (e.g., Rotman and Spinelli, 2014; Hass and Harris, 2016). As a result, observations of seafloor heat flow can provide weak constraints on frictional processes at depth.

### THERMAL REGIME OF THE OCEANIC CRUST AND LITHOSPHERE

#### Conductive Models

Seafloor bathymetry and heat flow furnish the primary evidence for the thermal state of oceanic lithosphere. Under conditions of isostatic equilibrium, large-scale bathymetry provides an estimate of the heat content of the plate and an integrated temperature structure through the lithosphere. These data have been used to develop plate-scale cooling models of the lithosphere (e.g., Davis and Lister, 1974; Parsons and Sclater, 1977; Stein and Stein, 1992; Hastenrath, 2013). Heat flow data are highly sensitive to the shallow thermal gradient and can be used to aid in the estimation of lithospheric physical parameters, especially in old seafloor where heat transfer is thought to be dominated by conduction. However, heat flow data can present challenges for determination of lithospheric physical and cooling parameters, particularly for younger basements, because local and shallow processes, including environmental disturbances associated with sedimentation and mass wasting, changes in bottom water temperatures, and hydrothermal circulation in the underlying basaltic basement can perturb the data.

Figure 2 summarizes marine bathymetric and heat flow data used to construct a cooling-plate reference model for the lithosphere, GDH1 (Stein and Stein, 1992). These data are from the North Pacific and northwest Atlantic, and are generally considered to be broadly representative of the global data set. Other analyses have led to somewhat different fitting parameters, but follow similar trends. Superimposed on the data are cooling plate model predictions and their uncertainty based on the standard deviation of the basal plate temperature, $T_w$, and lithospheric thickness, $L$, for conductively cooling lithosphere. These plots illustrate how the bathymetry and heat flow change in response to cooling. Seafloor depth increases from the ridge, near 2.5 km at zero age, to ~5.6 km at 160–170 Ma. For ages younger than ca. 80 Ma, the depth increases linearly with the square root of age as predicted by simple boundary layer cooling (Parker and Oldenburg, 1973; Davis and Lister, 1977). For older seafloor, depth approaches an asymptotic value of ~5.65 km (Stein and Stein, 1992). Heat flow also generally decreases with increasing age, but the mean of observations for a given age is generally less than the lithospheric prediction to ages of ca. 65 Ma (Stein and Stein, 1994). The global heat flow data set combined with well-established models of plate cooling provides a robust estimate...
of the total heat content in the lithosphere. Less well constrained by these models is the shallow geotherm that is susceptible to shallow environmental processes, particularly fluid flow.

To illustrate how uncertainties in conductive model predictions map to uncertainties in temperatures along the subduction thrust, we show the uncertainty in the geotherm (Fig. 1B) due to the standard deviation in the basal temperature ($T_m = 1450 \pm 250^\circ C$) and plate thickness ($L = 95 \pm 15$ km) as predicted by GDH1 (Stein and Stein, 1992). We show plate models having ages of 10 and 100 Ma, which bracket a large number of subduction zones and show key differences between young and old plates. We use these geotherms to construct thermal models of a generic subduction zone (Fig. 1C). The decrease in heat flow from the deformation front to ~20 km landward is dictated by the convergence rate, slab geometry, and landward thickening of the margin. The heat flow uncertainty in these calculations is based on uncertainty in the geotherm and decreases with distance from the boundary condition. In the generic thermal model, this leads to an uncertainty in surface heat flow of ~15 and ~5 mW m$^{-2}$ and uncertainties in plate boundary temperatures of ~90 °C and 40 °C at seismogenic depths, for the 10 and 100 Ma models, respectively (Fig. 1D). These differences are resolvable with high-quality heat flow data on the incoming plate and margins. The uncertainties, especially for young plates, likely represent minimum values for three reasons. First, these models represent a global average and specific regions need not conform to the global mean. Second, in young oceanic crust, heat flow values can be substantially (>1σ) lower than the mean (i.e., departures from the global mean are greater than modeled here). Third, these models are highly idealized, neglecting local variability in system properties and processes occurring out of the plane of the simulations.

Hydrothermal Circulation

Figure 2 shows that heat flow on young oceanic lithosphere is below conductive model predictions and its uncertainty based on established cooling models. This discrepancy is a measure of the advective heat loss of the oceanic crust (e.g., Wolery and Sleep, 1976; Stein and Stein, 1994). The discrepancy comes about because the vast majority of marine heat flow measurements rely on inserting a probe into sediment to measure conductive heat loss from the seafloor. In areas where there is considerable volcanic rock exposure and heat flow has been measured in sediment ponds between outcrops, heat flow values are typically low because a significant fraction of lithospheric heat is advected laterally by fluids and discharged through areas of basement exposure. Advective cooling of the volcanic crust lowers the thermal gradient in the overlying sediments, biasing even randomly distributed heat flow observations toward lower values. This sampling bias is exacerbated because conventional heat flow measurements cannot be made in areas of exposed basement where fluid discharge normally occurs. In a few cases where researchers have targeted areas of thin sediment cover adjacent to basement highs in regions where there is evidence for advective heat extraction, elevated heat flow values close to rock outcrops have been interpreted to be indicative of rapid fluid discharge (e.g., Davis et al., 1992; Villinger et al., 2002; Fisher et al., 2003a, 2003b).

In addition to this bias, there is a relatively large standard deviation associated with the heat flow discrepancy within young seafloor (Fig. 2). Fluid flow can homogenize temperatures within the basaltic aquifer such that observed
heat flow is proportional to sediment thickness, leading to considerable local variability (e.g., Davis et al., 1989; Davis and Becker, 2004; Langseth et al., 1992). Studies show that when the heat flow variability is larger than can be attributed to heat refraction or sedimentation effects due to variations in seafloor relief or thermal conductivity, the variability can be confidently attributed to hydrothermal circulation (e.g., Davis et al., 1992; Von Herzen, 2004; Fisher and Von Herzen, 2005). Thus hydrothermal circulation can result in two primary seafloor heat flow signatures, one or both of which may be found at individual sites: measured values that are systematically below the lithospheric prediction, and measured values that have high local variability.

Regional and local modeling of fluid flow based on detailed heat flow surveys provides insight into hydrothermal circulation. Modeling studies constrained by seafloor heat flow data point to the importance of permeability, sediment thickness, completeness of sediment cover, and basement relief on the vigor, pattern, and degree of ventilation of hydrothermal circulation (Sclater et al., 1974, 1976; Lister, 1972; Anderson and Hobart, 1976; Davis and Lister, 1977; Anderson et al., 1977; Davis et al., 1989, 1999; Fisher et al., 2003b).

The style of hydrothermal circulation is primarily dictated by the permeability structure of the basaltic oceanic crust and overlying sediment. The upper igneous crust is composed of extrusive pillow basalts, flows, and breccia that can have very high permeability ($\sim 10^{-6} - 10^{-9}$ m$^2$) because of interconnected porosity largely resulting from coarse clast size, fractures, lava flow contacts, and faults (Fisher, 1998; Becker and Davis, 2004; Fisher et al., 2014). Below this extrusive basalt aquifer, permeability measurements in sheeted dikes are low ($10^{-12} - 10^{-15}$ m$^2$) (Becker and Davis, 2004). Where the basement aquifer is not exposed to the ocean, it is capped by sediment typically many orders of magnitude less permeable ($10^{-16} - 10^{-19}$ m$^2$) that depends on lithology and porosity (Spinelli et al., 2004). For example, siliceous oozes have relatively high permeability and fine-grained turbidites have relatively low permeability. For layered sediment, the effective vertical permeability is dominated by the low-permeability layers, and interbedded high-permeability layers have little influence on vertical flow (e.g., Giambalvo et al., 2000).

Resistance to flow can be characterized by hydraulic impedance, a function of permeability and thickness. In marine sediments hydraulic impedance can vary over five orders of magnitude for sediment thicknesses of 10–100 m, increasing greatly as sediment thickens and permeability decreases due to porosity loss. Because the buoyancy forces driving fluid flow is limited to a few hundred kilopascals, there is typically little flow through sediments extracting significant quantities of heat where sediment is greater than a few tens of meters thick (Spinelli et al., 2004). Slow seepage through marine sediments is geochemically detectable at rates of millimeters per year, but much higher flow rates are required to allow advective heat extraction from the volcanic ocean crust. Thus sediment plays a key role in insulating the oceanic basement and limiting the advective loss of lithospheric heat.

Where the upper basaltic crust is exposed at the seafloor, fluid flow can bypass sediments and advect heat directly from the volcanic aquifer to the ocean. The strongest evidence for this process is regional heat flow measurements that are systematically below lithospheric predictions of conductive cooling on a regional basis. However, local heat flow values near areas of discharge can be larger than lithospheric predictions (e.g., Langseth and Herman, 1981; Fisher and Wheat, 2010). This style of hydrothermal circulation is referred to as ventilated.

As sediment thickness increases and there are fewer pathways between the crustal aquifer and overlying ocean, hydrothermal advection of lithospheric heat becomes restricted, decreasing the discrepancy between observed and predicted heat flow. In this case, hydrothermal circulation may be manifested as local variability in surface heat flow, due to the pattern of convecting cells and/or the homogeneity of upper basement temperatures in areas of basement relief. Regional and local studies show that spatial variability in heat flow can be related to the roughness of the upper basement as the sediment-base ment interface approaches an isothermal state (Davis et al., 1989; Fisher and Becker, 1995; Wang et al., 1997). This style of hydrothermal circulation is referred to as insulated. Davis et al. (1989) documented insulated circulation in the igneous crust of the Juan de Fuca Ridge flank where circulation is vigorous enough to homogenize the sediment-base ment interface temperature over lateral distances of kilometers.

The concepts summarized here can be used to estimate the geotherm forming the seaward boundary and initial conditions for thermal models of subduction. Key parameters in determining the shallow crustal geotherm are the depth and thickness of the aquifer, the degree of thermal homogenization within the aquifer, heat input below the aquifer, and the thermal properties of the rock-fluid system. These same parameters are important for estimating the structure of the geotherm below the aquifer, but also include the history and nature of hydrothermal circulation (Spinelli and Harris, 2011).

### Hydrothermal Circulation in Subducting Oceanic Crust

Spinelli and Wang (2008) argued that hydrothermal circulation within the basaltic basement aquifer should not be expected to stop at the deformation front of a subduction zone, but likely continues as long as there is sufficient permeability in volcanic crust. Following Davis et al. (1997), Spinelli and Wang (2008) applied a computationally efficient way of exploring hydrothermal circulation within the subducting basaltic aquifer using a Nusselt number approximation. The Nusselt number, Nu, is the ratio of the total heat transport to heat that would be transferred by conduction alone,

$$Nu = q_t/\lambda \Gamma,$$  \hspace{1cm} (4)

where \(q_t\) is the total heat transport, \(\lambda\) is the thermal conductivity, and \(\Gamma\) is the conductive thermal gradient. In this approach the aquifer is given an effective thermal conductivity that is increased by a factor of \(Nu\), simulating the thermal impact of efficient fluid mixing. The vigor of hydrothermal circulation in an idealized porous medium is described by the Rayleigh number, \(Ra\),

The strongest evidence for this process is regional heat flow measurements that are systematically below lithospheric predictions of conductive cooling on a regional basis. However, local heat flow values near areas of discharge can be larger than lithospheric predictions (e.g., Langseth and Herman, 1981; Fisher and Wheat, 2010). This style of hydrothermal circulation is referred to as ventilated.

As sediment thickness increases and there are fewer pathways between the crustal aquifer and overlying ocean, hydrothermal advection of lithospheric heat becomes restricted, decreasing the discrepancy between observed and predicted heat flow. In this case, hydrothermal circulation may be manifested as local variability in surface heat flow, due to the pattern of convecting cells and/or the homogeneity of upper basement temperatures in areas of basement relief. Regional and local studies show that spatial variability in heat flow can be related to the roughness of the upper basement as the sediment-base ment interface approaches an isothermal state (Davis et al., 1989; Fisher and Becker, 1995; Wang et al., 1997). This style of hydrothermal circulation is referred to as insulated. Davis et al. (1989) documented insulated circulation in the igneous crust of the Juan de Fuca Ridge flank where circulation is vigorous enough to homogenize the sediment-base ment interface temperature over lateral distances of kilometers.
Research Paper

\[ Ra = \frac{\alpha g k_p q h^4}{\mu \kappa} \]  

(5)

where \( \alpha \) is the coefficient of thermal expansion, \( g \) is gravitational acceleration (9.81 m/s\(^2\)), \( k \) is permeability, \( \rho_f \) is the fluid density, \( h \) is aquifer thickness, \( \mu \) is viscosity, \( q \) is the vertical conductive heat flow at the base of the aquifer, and \( \kappa \) is thermal diffusivity (\( \kappa = \lambda / \rho c_p \)), where \( \rho_c \) is the specific heat of the aquifer. In practice, \( \kappa \) is set a priori (or modified as part of a modeling study to achieve a particular result). \( h \) is specified in the model geometry, and \( \alpha, \rho_f, \mu, \) and \( \kappa \) are calculated as functions of pressure and temperature (Harvey et al., 1997). Application of Ra calculations to assess stability in hydrothermal systems necessarily simplify geological conditions. Aquifer permeability can be parameterized as a function of depth, \( z \) (Spinelli and Wang, 2008),

\[ \log(k) = -9 - 5.5(z - 0.5Z_{aq}) \]  

(6)

where \( Z_{aq} \) is the depth of the aquifer before subduction. \( Nu \) is computed using an empirical relationship with Ra (Wang, 2004; Spinelli and Wang, 2008),

\[ Nu = 0.08Ra^{0.85} \]  

(7)

Because there is a nonlinear feedback between temperature and Nu, temperature is computed using a multistep method. The iteration begins using a simulation with \( Nu = 1 \) and Ra is calculated for each aquifer element in the model and a new Nu is determined from Equation 7. Then the thermal conductivity is increased in each aquifer element by a factor of Nu, the simulation is rerun, Ra and Nu are recalculated until temperature changes are less than 1 °C between iterations. Values of Nu typically range from 1 to 1000. Nu values of 1000 are large enough to homogenize temperatures within the aquifer; larger values have only a small additional impact on the thermal structure (e.g., Wang et al., 2015).

Simulations are run for either insulated or ventilated conditions. Example geotherms are shown in Figure 3. The conductive geotherm (red line) shows a change in trend at the sediment-basement interface reflecting the change in thermal conductivity. The geotherm reflecting insulated conditions (green line, Fig. 3) has the same thermal gradient through the sediments, but an isothermal aquifer. The geotherm through the sediments sets the temperature of the aquifer. Below the aquifer the geotherm is based on conductive cooling models with a net heat deficit reflecting the increased heat transfer into the aquifer relative to conductive conditions. A ventilated geotherm (blue line, Fig. 3) using a thermal gradient through the sediment of 10 °C/km is also shown. Like the insulated case, the temperature of the aquifer is based on the thermal gradient through the sediments, and below the aquifer the geotherm is based on conductive cooling models. For both ventilated and insulated circulation, the geotherm through the aquifer can be adjusted to mimic the degree of thermal homogenization; in the ventilated case, the geotherm through sediments can be adjusted to indicate the efficiency of heat extraction from the volcanic crustal aquifer. With insulated models all heat transfer from the aquifer to the ocean is along the conductive thermal gradient through the sediments.

![Figure 3. Geotherms for defining the seaward boundary condition (bsf—below seafloor). The conductive geotherm (red) corresponds to a conductive plate cooling model appropriate for the age of the crust. In this figure the age of the oceanic basement is 24 Ma, such that the basal heat flow is 102 mW m\(^{-2}\). The change in gradient through the 400 m of sediment is due to the contrast in thermal conductivity between the sediment (1.1 W m\(^{-1}\) K\(^{-1}\)) and the crust (2.9 W m\(^{-1}\) K\(^{-1}\)). The intermediate geotherm (green) has the same gradient through the sediments as the plate cooling geotherm but includes an isothermal aquifer 0.5 km thick. This geotherm is coupled to the plate cooling geotherm at depth and is used for insulated models. The cool geotherm (blue) is parameterized with a low thermal gradient (10 °C km\(^{-1}\)) appropriate for observed ventilated conditions through the sediments, an isothermal aquifer, and plate cooling at depth. This geotherm is used for ventilated models.](https://pubs.geoscienceworld.org/gsa/geosphere/article-pdf/13/5/1425/3995579/1425.pdf)

Thermal models simulating ventilated conditions set the geotherm at the deformation front. In these simulations all heat transported from depth along the subducting aquifer is removed so that the incoming plate is not warmed seaward of the deformation front. In contrast, thermal models simulating insulated conditions set the geotherm well seaward (150 km or more) of the deformation front, and the incoming plate is allowed to warm seaward of the deformation front. Hybrid cases where the crust transitions from ventilated to insulated conditions can be simulated by modeling ventilated hydrothermal circulation to a time corresponding to a distance in front of the deformation front and then imposing insulated conditions at that point. The pastel lines in Figure 1C show the impact of hydrothermal circulation on heat flow across the trench and forearc corresponding to ventilated and
insulated hydrothermal circulation. Two conspicuous features include higher heat flow at the trench than can be explained with conductive cooling models (anomalous magnitudes ~50–100 mW m⁻²), and a more abrupt landward decrease in heat flow (by ~50–100 mW m⁻² over ~10–20 km) than can be explained with reasonable variations in the convergence rate and slab dip. In these models, the high effective thermal conductivity transports heat from the deeper, warmer portion of the aquifer to the shallower, cooler portion of the aquifer at and seaward of the trench.

Hydrothermal circulation within subducting oceanic crust is manifested by a change in temperature along the plate boundary relative to models without fluid flow (Fig. 1D). In general temperatures are increased near the deformation front and decreased landward, consistent with the transport of heat from depth to the trench region. The landward distance affected by hydrothermal circulation is a function of how the aquifer permeability decreases with distance from the deformation front, a parameter that is not well known. If the updip and downdip extent of the seismogenic zone is between ~100 °C and 350 °C, then the impact of fluid flow is to increase the width of the seismogenic zone. This increase in width has implications for seismic and tsunami hazard analysis, patterns of interseismic locking, and the progress of diagenetic and/or metamorphic reactions that, in turn, affect the distribution of fluid sources from subducting material.

### CASE STUDIES

Heat flow surveys designed to understand hydrothermal circulation in oceanic crust approaching and entering subduction zones are rare; instead, scattered heat flow measurements are often combined to make inferences about the thermal state and hydrothermal circulation of the incoming oceanic crust. Similar to ridge flank studies, hydrothermal circulation in subduction zones can be understood in terms of insulated and ventilated conditions. Case studies chosen for this review are either regions that are relatively data rich or have data that help to highlight important aspects of hydrothermal circulation.

**Nankai**

The Nankai Trough (Fig. 4) has been the site of numerous heat flow surveys and thermal studies (e.g., Yamano et al., 1992, 2003; Kinoshita et al., 2008; see review in Harris et al., 2013), although none has specifically targeted hydrothermal circulation on the incoming plate. Nevertheless, the regional environment is conducive to hydrothermal circulation as indicated by the heat flow data. Here, the Shikoku Basin (Philippine Sea plate) is subducting beneath the Eurasian plate at a rate of 80 mm/yr. The Shikoku Basin basement varies in age between ca. 14 and 30 Ma. Well seaward of the trough, sediment cover is discontinuous but can be as thick as 250 m in basement depressions (Higuchi et al., 2007). As the Philippine Sea plate approaches the Nankai Trough, the sediment thickness increases. Nankai Trough sediment consists of turbidite fill overlain by hemipelagic sediment (Ike et al., 2008). The sediment thickness in the trench is typically 200–600 m greater than the basement relief (Ike et al., 2008). The increase in sediment thickness (such that it exceeds basement relief) as the plate approaches the trough could facilitate a switch from ventilated to insulated hydrothermal circulation.

Harris et al. (2013) assessed the likelihood of hydrothermal circulation within Shikoku Basin basement by computing the heat flow fraction defined as the ratio of observed heat flow to that predicted by conductive cooling models. They found that the heat flow fraction is 0.8 with a standard deviation of 0.4 and is remarkably similar to the global heat flow fraction for this age of crust. The low mean and large standard deviation of fractional heat flow is characteristic of ventilated hydrothermal circulation and consistent with its young age and exposed basement.

Along the Nankai Trough, where sediment is thicker and more continuous, values of heat flow are 20% higher than predicted by conductive cooling models based on basement age (Yamano et al., 2003). The incoming sediment is overlain by coarse turbidites that may support lateral flow of fluids, but if the flow is parallel to isotherms then little heat is transported. The high mean fractional heat flow suggests an additional source of heat relative to conductive predictions. Locally, the Kinan seamounts show evidence for volcanism that persisted to as late as 78 Ma with many samples in the 10–8 Ma range (Sato et al., 2002; Ishizuka et al., 2009). Although this late-stage volcanism may add heat in the vicinity of the Kinan Seamount Chain, it was not enough to overprint the observed magnetic lineations and does not explain the regionally higher than predicted heat flow along the Nankai Trough (Yamano et al., 2003).

Figure 4 shows heat flow along three drilling transects illustrating features discussed in previous compilations (Yamano et al., 1984; Kinoshita and Yamano, 1986; Yamano et al., 1992, 2003; Harris et al., 2013): (1) relatively large variability between closely spaced heat flow measurements, (2) heat flow higher than conductive predictions within and just seaward of the Nankai Trough, and (3) a decrease in heat flow just landward of the deformation front. These features are best expressed in the Muroto transect where heat flow just seaward (<25 km) of the deformation front has a value of ~200 mW m⁻², much higher than the predicted value for 15 m.y. old lithosphere, 130 ± 22 mW m⁻². These high values decline steeply landward to ~60 mW m⁻² over a distance of 65 km. The landward decrease in heat flow is less abrupt along the Kumano transect, where the crust is 20 Ma and the predicted heat flow is 100 ± 20 mW m⁻². Here heat flow values decline from ~140 mW m⁻² 30 km seaward of the deformation front to values of ~60 mW m⁻² 30 km landward of the deformation front, a distance of 60 km. Sparse data from the Ashizuri transect (Fig. 4C) preclude a definitive assessment of heat flow variations, but the general pattern is consistent with those at the other two transects.

In Spinnell and Wang (2008) a number of potential sources for the elevated heat flow near the deformation front and the steep decrease landward of this area at the Muroto transect were considered, the most likely being hydrothermal circulation within the subducting oceanic aquifer. They applied a model...
Figure 4. Heat flow data from the Nan-kai Trough (modified from Harris et al., 2013) (smts.—seamounts). All distances are relative to the deformation front with landward distance being positive. Inset shows heat flow data color and location of thermal models. Marine probe (circles), boreholes (stars), and BSR (bottom simulating reflector; small circles) are color coded by heat flow. On land, circles show borehole values of heat flow. (A–C) Heat flow data and thermal models along the Muroto (Spinelli and Wang, 2008), Kumano (Spinelli and Harris, 2011), and Ashizuri transects (Harris et al., 2013). Thermophysical rock properties and other model details are given in the original studies. The permeability evolution of the subducting aquifer is based on Equation 6. Solid and dashed black lines show thermal models with and without fluid flow, respectively. (D–F) Temperature along the subduction thrust for Muroto, Kumano, and Ashizuri.
of insulated hydrothermal circulation in the incoming crust and interpreted that fluid flow in the subducting oceanic crust leads to temperatures that are generally 25 °C higher near the toe of the margin wedge, and 50–100 °C lower near the down-dip limit of the seismogenic zone, relative to models without fluid flow. Both the Kumano and Azshizuri transects are simulated using similar style of models (Spinelli and Harris, 2011; Harris et al., 2013). Collectively, these studies suggest that there is a wider seismogenic zone than would be present if there were not hydrothermal circulation in the subducting plate, an interpretation that is generally consistent with geodetic estimates of the geometry or the seismogenic zone (Sagiya and Thatcher, 1999; Baba et al., 2002, 2006; Ito et al., 2007).

**Costa Rica**

The region offshore the Nicoya Peninsula was surveyed specifically to understand the thermal state of the oceanic crust just prior to subduction (Fisher et al., 2003a; Hutnak et al., 2007, 2008). A significant along-strike thermal transition allows comparisons between models of hydrothermal circulation in subducting crust.

Offshore of western Costa Rica, the 24–16 Ma Cocos plate is subducting under the Caribbean plate at a rate of ~87 mm/yr (Fig. 5). Here the lithology consists of hemipelagic mud overlying pelagic carbonate (Spinelli and Underwood, 2004). The bathymetry and sediment thickness of the incoming Cocos plate varies substantially along strike and is a function of the plate's origin and subsequent history. The Cocos plate was formed at two ridges, the fast spreading East Pacific Rise (EPR) and the intermediate-spreading Cocos-Nazca spreading center (CNS). The EPR and CNS derived crust is juxtaposed across a plate suture that is a combination of a triple junction trace and a fracture zone (Barckhausen et al., 2001). After the plate formed it was intruded by Galapagos magmatism that generated the Cocos Ridge and many seamounts that contribute to large variations in bathymetric relief. Sediment on the EPR crust, where it is not disrupted by seamounts, is generally 350–450 m thick (Spinelli and Underwood, 2004); sediment is thinner on CNS crust, and thins toward the Cocos Ridge. Exposed basement rock on the Cocos plate is more common on CNS crust than on EPR crust.

Heat flow offshore of Costa Rica, both on the incoming plate and margin, is well characterized and was summarized in Hutnak et al. (2007) and Harris et al., 2008a. Heat flow data come from a variety of sources including older data (black circles). Red circles are from a regional study of the incoming plate (Hutnak et al., 2007), and data crossing the margin are shown as white circles (Harris et al., 2010a). Ocean Drilling Program and Integrated Ocean Drilling Program heat flow values are shown as yellow stars. EPR—East Pacific Rise; CNS—Cocos-Nazca spreading center (CNS). The EPR and CNS derived crust is juxtaposed across a plate suture that is a combination of a triple junction trace and a fracture zone (Barckhausen et al., 2001). After the plate formed it was intruded by Galapagos magmatism that generated the Cocos Ridge and many seamounts that contribute to large variations in bathymetric relief. Sediment on the EPR crust, where it is not disrupted by seamounts, is generally 350–450 m thick (Spinelli and Underwood, 2004); sediment is thinner on CNS crust, and thins toward the Cocos Ridge. Exposed basement rock on the Cocos plate is more common on CNS crust than on EPR crust.

Figure 5. Heat flow along Costa Rica margin. (A) Data come from a variety of sources including older data (black circles). Red circles are from a regional study of the incoming plate (Hutnak et al., 2007), and data crossing the margin are shown as white circles (Harris et al., 2010a). Ocean Drilling Program and Integrated Ocean Drilling Program heat flow values are shown as yellow stars. EPR—East Pacific Rise; CNS—Cocos-Nazca spreading center (CNS). The EPR and CNS derived crust is juxtaposed across a plate suture that is a combination of a triple junction trace and a fracture zone (Barckhausen et al., 2001). After the plate formed it was intruded by Galapagos magmatism that generated the Cocos Ridge and many seamounts that contribute to large variations in bathymetric relief. Sediment on the EPR crust, where it is not disrupted by seamounts, is generally 350–450 m thick (Spinelli and Underwood, 2004); sediment is thinner on CNS crust, and thins toward the Cocos Ridge. Exposed basement rock on the Cocos plate is more common on CNS crust than on EPR crust.
al. (2010a), respectively. These measurements come from a variety of sources, including probe and downhole measurements and estimates from bottom simulating reflectors. Values seaward of the deformation front are used to parameterize the geothermal of thermal models and values across the margin are used to assess model performance. We discuss each set of values in turn.

Early measurements (Von Herzen and Uyeda, 1963; Vacquier et al., 1967) were made without the benefit of modern bathymetric or seismic reflection records and tend to be widely spaced (Fig. 5). These measurements indicated that EPR-generated igneous crust of the Cocos plate is associated with very low heat flow relative to the global mean for crust of the same age. In contrast, the mean of early heat flow values on CNS-generated igneous crust is close to the conductive prediction, but above the observed global mean for crust of this age. These values exhibit large variability indicative of fluid flow with a mean and standard deviation of 110 and 60 mW m\(^{-2}\), respectively (Fig. 5).

In 2000 and 2001, heat flow studies specifically designed to investigate the nature of the cool EPR crust and its transition with warm CNS crust were completed (Fisher et al., 2003a; Hutnak et al., 2007). All heat flow measurements are closely spaced and collocated with seismic reflection lines, providing an environmental context for interpreting the data. Heat flow values on CNS crust have a mean of 103 ± 74 mW m\(^{-2}\) corresponding to a fractional mode of 0.2, suggesting efficient advective extraction of geothermal heat, and probably requiring very high basement permeability at the updip limit of seismicity. Processes that do not include advective heat transport in the basalt aquifer. On the EPR side of the thermal transition, heat flow values are significantly lower than conductive predictions (Fig. 6). Heat flow increases seaward of the deformation front, showing the impact of bringing heat from depth. On the CNS side of the thermal transition, heat flow increases seaward to values greater than can be accounted for with a conductive model.

These calculations indicate significant differences in the depth to the updip limit of seismicity on either side of the thermal transition (Newman et al., 2002). Histograms of hypocenters (Kyriakopoulos et al., 2015) along the subduction thrust indicate that the updip limit of seismicity is ~10 km deeper on the cold side of the transition than the warm side (Fig. 6). We use the thermal models to assign a temperature to each hypocenter and find that for both sides the updip limit of seismicity corresponds to ~100 °C. Because the sediment and fault properties are likely similar over this short along-strike section of the Middle America Trench, these results are consistent with the idea that thermally controlled processes likely affect the updip limit of seismicity. Processes that may affect frictional behavior of plate interface material and/or the pore fluid pressure at temperatures ~100 °C include the opal to quartz and smectite to illite transitions (e.g., Moore and Saffer, 2001; Spinelli et al., 2006).

**South-Central Chile**

A third example of hydrothermal circulation from subducting oceanic crust comes from the south-central Chile Trench. Four heat flow transects across this region were analyzed (Rotman and Spinelli, 2014; Fig. 7). Offshore south-central Chile, the Nazca plate is subducting under the South American plate at an approximately uniform rate of 60 mm/yr (Fig. 7). Sediments seaward of the trench comprise mainly siliceous red clay, ≤200 m thick, and exposed base-
Figure 6. Thermal models from Costa Rica across the thermal transition (Harris et al., 2010b) combined with new hypocenter relocations (Kyriakopoulos et al., 2015). (A) Heat flow (HF) data and thermal models along seismic profile BGR99–44. Solid and dashed lines show thermal models with and without fluid flow, respectively. (B) Thermal model showing hypocenters (red circles) along plate boundary. EPR—East Pacific Rise. Temp—temperature. (C, D) Histograms showing depth and temperature distribution of plate boundary hypocenters. (E) Heat flow data and thermal models along seismic profile SO81–8. (F) Thermal model showing hypocenters (red circles) along plate boundary. CNS—Cocos-Nazca spreading center. (G, H) Histograms showing depth and temperature distribution of plate boundary hypocenters. Thermophysical rock properties and other model details are given in Harris et al. (2010b). The evolution of permeability in the subducting aquifer is based on Equation 6 (see text).
Figure 7. Results from south-central Chile. ODP—Ocean Drilling Program. (A) Tectonic map showing location of trench (barbed line) between the subducting Nazca plate and overriding South America plate. Arrows show the convergence direction and rate. Black lines show location of thermal models and age of crust at trench. (B–E) Measured and modeled heat flow along the profiles. Bottom simulating reflector (BSR) measurements are from Grevemeyer et al. (2005), and the shallow probe measurements are from Cande et al. (1987), Flueh and Grevemeyer (2005), Grevemeyer et al. (2005), and Contreras-Reyes et al. (2007). Heat flow peaks seaward of the deformation front show the modeled change from ventilated (seaward) to insulated (landward) conditions (gray and black curves). (F–I) The impact of hydrothermal circulation and frictional heating on plate boundary temperatures. Thermophysical rock properties and other model details are given in Rotman and Spinelli (2014). Temp Diff—temperature difference.
ment is common. Turbidites are common within and near the trench, and the sediment thickness generally increases to the south, from <2.0 to ~2.4 km. For this study, it is important that crustal age at the deformation front decreases from ca. 1.2 Ma at 45°S, close to where an actively spreading ridge is subducting, to 31.5 Ma at 36°S. Conductive models of crustal cooling predict that heat flow should decrease from ~450 to 90 mW m⁻² over this age range, and global marine heat flow data show an even more abrupt decrease with age (Fig. 2D).

Although heat flow data near the south-central Chile Trench are limited, values from regions where the sediment thickness is <2 km are generally 70% less than conductive predictions (based on crustal age), and there is high spatial variability, with measurements that are 6%–148% of conductive models (Rotman and Spinelli, 2014). Contreras-Reyes et al. (2007) reported heat flow values associated with exposed basement varying between 150 and 7 mW m⁻² over a distance of <10 km. Observations of seafloor heat flow that tend to be suppressed well below lithospheric cooling predictions are characteristic of ventilated hydrothermal circulation, and high variability can be explained by a combination of local convection and proximity to sites where fluids enter or exit the volcanic crust. Other evidence for widespread hydrothermal circulation includes variable heat flow over plate-bending normal faults (Grevemeyer et al., 2005), and lower than expected seismic velocities at depth that might result from hydration of the oceanic crust and uppermost mantle (Contreras-Reyes et al., 2007).

Similar to examples from Nankai and Costa Rica, heat flow data across the Chile Trench are better replicated with thermal models incorporating the effects of ventilated fluid flow than those not incorporating fluid flow (Fig. 7). All models use the same permeability structure (Equation 6), but vary the position of the transition from ventilated to insulated hydrothermal circulation as a function of sediment thickness and basement relief. Rotman and Spinelli (2014) gave more weight to lower values of heat flow on the margin, arguing that high and scattered values likely reflected areas of focused sediment dewatering; they also investigated tradeoffs between hydrothermal circulation and frictional heating, applying a coefficient of friction μ' = 0.03, consistent with mechanical studies of subduction thrusts (Wang and He, 2008). These simulations showed that friction could account for ≤7 mW m⁻² of surface heat flow, illustrating that these measurements are relatively insensitive to frictional heating at depth.

Differences in plate temperature for these models illustrate a number of key points (Fig. 7). First, the efficacy of hydrothermal circulation in cooling plate boundary temperatures depends in part on the increase in temperature with distance from the deformation front (ΔT/ΔL) in the absence of fluid flow (Rotman and Spinelli, 2014). The age of the subducting oceanic basement has a strong control on this ratio. Second, although modest frictional heating does not have a large impact on surface heat flow, it does have a strong influence on plate boundary temperatures. Because of the dependence on the normal stress, maximum frictional heating increases with depth to the brittle-ductile transition (at the 350 °C isotherm in these models), which deepens with increasing basement age. Models combining hydrothermal circulation and frictional heating show tradeoffs in the plate boundary temperature that depend on age, and illustrate the difficulty of resolving the coefficient of friction in the presence of hydrothermal circulation: hydrothermal circulation can both extract and redistribute excess (frictional) heat.

Models with no hydrothermal circulation show that the distance along the plate boundary between the 100 °C and 350 °C isotherms should increase with incoming plate age. Instead, it was found (Rotman and Spinelli, 2014) that hydrothermal circulation proportionately cools younger and warmer crust, homogenizing thermal conditions such that the distance between the 150 °C and 350 °C isotherm is relatively constant along strike.

Haida Gwaii

Wang et al. (2015) explored the implications of hydrothermal circulation offshore western Canada at Haida Gwaii. Here, the 8–6 Ma Pacific plate is converging obliquely and underthrusting the North American plate at a rate of ~20 mm/yr (Fig. 8). The incoming plate is covered by ~1.0–1.5 km of sediment that promotes conditions for insulating hydrothermal circulation. Although there are seamounts well offshore, there are no known areas of basement exposure on the incoming plate near the Queen Charlotte fault.

Heat flow measurements on the incoming plate in this region are scarce, but near the margin tend to indicate values ~250 mW m⁻², ~75 mW m⁻² higher than conductive predictions (Hyndman et al., 1982). Landward of the Queen Charlotte fault observed heat flow drops to ~75 mW m⁻² landward of the Queen Charlotte fault (Yorath and Hyndman, 1983). This decrease in heat flow is more rapid than can be explained with the plate geometry, convergence rate, and landward thickening of the margin. Wang et al. (2015) showed that the data can be fit well with models that incorporate hydrothermal circulation using a high-Nu proxy. Because of the lack of heat flow data on the incoming plate, Wang et al. (2015) estimated the geotherm using GDH1. Figure 8B shows heat flow models corresponding to Nu = 1 (no fluid flow), and Nu = 1000 (vigorous flow). The model that represents the thermal influence of vigorous hydrothermal circulation clearly provides a better fit to the data, allowing prediction of temperatures along the subduction thrust (Fig. 8C). The intersection of the Queen Charlotte fault with the subduction thrust occurs at a temperature of ~350 °C, ~50 °C cooler than the no-fluid-flow model.

The implication of these data and models is that, even though the incoming plate is well sealed by a thick accumulation of sediment, heat advected from depth within the subducting crust can substantially increase the surface heat flow near the deformation front. This process is important for accurately predicting the distribution of temperatures along the subduction thrust, and thus the behavior of the fault.

Cascadia

One area where thick sediment cover may mask the heat flow signal from insulated circulation is along the Cascadia margin of Washington, Oregon, and...
northern California, USA (Cozzens and Spinelli, 2012). Here, the 9–7 Ma Juan de Fuca plate subducts beneath the North American plate at ~3–4 cm/yr (Fig. 9). Pleistocene glaciation has resulted in a laterally continuous and thick section of sediment that extends to near the active spreading centers. Sediment thickness along the northern deformation front is as much as 3–5 km (Davis and Hyndman, 1989), approximately three times that observed at Haida Gwaii.

The eastern flank of the Juan de Fuca Ridge, in the FlankFlux experiment area (Fig. 9), has been the site of numerous heat flow studies designed to understand hydrothermal circulation (e.g., Davis et al., 1992, 1999; Fisher et al., 2003b; Spinelli and Fisher, 2004; Hutnak et al., 2006). Among other findings, these studies show that hydrothermal circulation is ventilated close to the ridge, but the circulation system becomes insulated and heat flow approaches conductive predictions at ~35 km from the ridge at an age of 1.3 Ma. Variability in heat flow across sedimented regions is inversely proportional to sediment thickness, implying that the sediment-basement interface is nearly isothermal (Davis et al., 1997).

Heat flow measurements outside of the FlankFlux area are sparse and the apparent agreement between heat flow data and conductive cooling models at the eastern end of the FlankFlux area are used as arguments for conductive conditions at the deformation front (Hyndman and Wang, 1993). However, given the young age of the basement, it was argued (Cozzens and Spinelli, 2012) that insulated hydrothermal circulation might be expected, and the impact of hydrothermal circulation along this margin was explored. They found that the thick sediment cover at the deformation front results in a broad low-amplitude surface heat flow anomaly that may be difficult to discern from conductive predictions (Fig. 9). However, Cozzens and Spinelli (2012) argued for the presence of hydrothermal circulation on the basis of correlating pressure-temperature conditions with the seismically estimated location of the basalt to eclogite transition within the subducting slab at depths of 40–45 km inferred from analysis of seismic velocities (Bostock et al., 2002; Rondenay et al., 2008; Abers et al., 2009). In this setting, models incorporating hydrothermal circulation are in better agreement with the basalt to eclogite transition than models without fluid flow (Cozzens and Spinelli, 2012).

Heat flow measurements seaward of the deformation front can be used to test for the presence of hydrothermal circulation. With the recognition that insulated hydrothermal circulation leads to an inverse relationship between sediment thickness and heat flow (e.g., Davis et al., 1989; Fisher and Becker, 1995; Wang et al., 1997), a “topo test” was proposed (Harris and Chapman, 2004) to explore for the presence of hydrothermal circulation. In this test, closely spaced heat flow measurements are made over seismically imaged basement topography with continuous, but variable thickness, sediment covering the basement. If advection is sufficiently vigorous to create a nearly isothermal condition at the top of the permeable igneous basement, then this state will be manifested by heat flow that is inversely proportional to sediment thickness. High heat flow will occur over buried topographic highs; low heat flow will occur over thickly buried topographic lows. If hydrothermal circulation does not occur, or if it is
insufficiently vigorous to thermally homogenize the upper crust, the heat flow through the sediment covering the rough topography will be constant except for relatively small thermal refraction and sedimentation effects (e.g., Fisher and Von Herzen, 2005). The relationship between heat flow variability and basement relief can be used to make estimates of aquifer permeability that can then be fed into thermal models of the shallow subduction zone. Such a test can be applied to the Cascadia subduction zone (and other thickly sedimented margins) near the trench in a location with appropriate basement topography.

**DISCUSSION**

The studies discussed in this review show evidence consistent with persistent hydrothermal circulation within subducting oceanic crust. Many of these studies benefit from modern heat flow survey techniques that include high-precision navigation and closely spaced measurements (<1 km) on both the incoming plate and across the deformation front. Closely spaced measurements overcome issues associated with aliasing and help to constrain the extent and significance of heat flow variability. Swath mapping and seismic reflection data give information on the measuring environment and sediment thickness, allowing downward continuation of seafloor data into the volcanic basement. These survey techniques are especially important for distinguishing between deviations in thermal conditions associated with hydrothermal circulation, conductive refraction, and/or sedimentation, which provide rationale explanations for heat flow variability and help to discriminate between competing thermal models.

Relative to conductive predictions, insulated hydrothermal circulation seaward of the trench leads to high values of heat flow at the deformation front and a steeper landward decline in heat flow than can be accounted for on the basis of the plate convergence rate, slab dip, and/or landward thickening of the margin. Insulated hydrothermal circulation is most likely to occur where the sediment cover is greater than a few hundred meters thick and regionally continuous at a length scale of 100 km or more.

In contrast, ventilated hydrothermal circulation is characterized by heat flow that is significantly lower than predicted with conductive models. Ventilated hydrothermal circulation is more likely to occur where sediment is thin, sedimentation rates are low, and basement rocks are exposed. These
conditions may occur at island arcs (where sediment supplies are limited) or where sediments being shed from larger continental landmasses are deposited upslope of the trench (e.g., Vannucchi et al., 2016). Ventilated hydrothermal circulation can generate large differences in subduction thrust temperatures, relative to no flow conditions, due to depression of the geotherm relative to conductive conditions (Fig. 3). If the hydrothermal circulation is ventilated, heat being transported updip along the subducting aquifer may subsequently discharge to the ocean where the aquifer is exposed at the seafloor (possibly as far landward as the deformation front), such that the high heat flow in the trench and the steep landward decrease in heat flow characteristic of insulated circulation are not present.

Both insulated and ventilated models of hydrothermal circulation in the subducting oceanic basement show the greatest impact on seafloor heat flow within ~20 km landward and seaward of the deformation front. To better understand the potential roles of hydrothermal circulation in this setting, closely spaced heat flow data from both sides of the deformation front are needed. As seen with examples from the Nankai Trough and the south-central Chile Trench, hydrothermal circulation may transition from ventilated to insulated conditions seaward of the deformation front. Where a system shifts from ventilated to insulated conditions (e.g., where sediment accumulation on a plate approaching a trench is sufficient to isolate the basement aquifer from the ocean), heat flow may display a landward increase. Heat flow data on the incoming plate are essential to identify and understand this transition.

Deviations in seafloor heat flow 20–100 km landward of the deformation front are both less apparent in observational studies and more difficult to use to distinguish between underthrust plates dominated by conductive versus advective thermal processes. Data from these areas are subject to numerous environmental perturbations, including sediment slumping and changes in the temperature of bottom water; the latter becomes especially common in oceanic depths <2 km. In addition, the relatively thick wedge of sediment commonly found above the underthrust plate tends to obscure spatial variations in thermal properties at and below the subduction thrust, because of both conductive refraction and fluid flow within the wedge.

In some cases hydrothermal circulation in the oceanic crust may occur without producing obvious anomalies in seafloor heat flow. These cases include weak ventilated circulation (resulting in a small fraction of advective heat extraction), and cases of insulated convection where the sediment cover is thick enough to attenuate local variations at depth, as may be happening at Cascadia. These environments may still lead to significant perturbations of temperature on the plate interface relative to conductive predictions.

Stein (2003) used the global heat flow data set to assess whether heat flow associated with trenches show a bias toward low values that could be linked to enhanced heat extraction from the oceanic lithosphere by hydrothermal circulation. The core hypothesis of that study is that plate-bending normal faults, induced as the lithosphere traverses the outer rise, enhance ventilated hydrothermal circulation relative to oceanic lithosphere of the same age away from trenches (and not affected by outer rise faulting). In this scenario, the normal faults are envisioned to act as high-permeability conduits facilitating advective heat loss from the lithosphere at focused locations (and not commonly observed by typical heat flow observations). If this were to occur, it would tend to reduce temperatures in the basement, such that heat flow observations made in seafloor sediments overlying the aquifer are anomalously low for the plate age.

An alternative hypothesis posits that plate-bending normal faults could increase the depth extent of hydrothermal circulation within the basement rocks (i.e., increase the aquifer thickness) without developing conduits through the overlying sediment with permeability sufficient for substantial heat advection to the seafloor (i.e., the system maintains insulated hydrothermal circulation; Kawada et al., 2014; Yamano et al., 2014). In this second scenario, the plate-bending faults facilitate delivery of heat from deeper in the lithosphere to the upper basement aquifer; aquifer temperatures increase, increasing the conductive heat flow through the overlying sediment. This mechanism is proposed as an explanation for a broad (~150 km wide) low-amplitude (~20 mW m⁻²) heat flow anomaly observed across the outer rise seaward of the Japan Trench (Kawada et al., 2014; Yamano et al., 2014). For the global data set, Stein (2003) observed no significant difference between heat flow data near trenches and away from trenches on lithosphere of the same age. Thus, if plate-bending normal faults enhance hydrothermal circulation in either of the two ways presented here, the results of Stein (2003) indicate that neither the ventilated scenario nor the insulated scenario dominate the global data set. Testing these hypotheses for effects of plate-bending normal faults on seafloor heat flow may require detailed studies of closely spaced heat flow measurements that are collocated with complementary data, and bottom surveys to locate potential venting sites where advective heat extraction from the plate is hypothesized to occur.

In spite of the progress made over the past decade on understanding temperature along the subduction thrust, there is a lack of knowledge for some key parameters. A fundamental uncertainty in these models is the permeability structure of the igneous crust and how the permeability evolves with depth. The evolution of permeability with depth is poorly known but is likely a function of pressure, temperature, and diagenetic and/or metamorphic effects. To date, the surface heat flow distribution near the deformation front has proved to be a somewhat blunt instrument in constraining the details of permeability evolution within the subducting aquifer (e.g., Spinelli and Wang, 2008). To better define this permeability distribution and/or evolution may require (or at least, benefit from) a multipronged approach. For example, surface heat flow has been used in conjunction with the seismically inferred location of slab alteration to constrain the vigor and extent of hydrothermal circulation in Nankai (Spinelli and Wang, 2008) and Cascadia (Cozzens and Spinelli, 2012). Future studies may benefit from also including geochemical constraints on the sources of fluids within the basement aquifer (e.g., Solomon et al., 2009).

Distinguishing between the impacts of hydrothermal circulation on surface heat flow from other processes remains a challenge. At margins accreting sediment these processes include the sediment thickening and de-watering due to porosity loss as sediment is accreted to the margin (Wang...
et al., 1993). Sediment thickening can decrease margin heat flow while sediment dewatering can increase it. The impact on surface heat flow for these processes is likely to be smaller than the impact of hydrothermal circulation. These processes act over similar length scales, and the lack of data regarding these processes makes it difficult to separate these effects from hydrothermal circulation in the subducting aquifer.

Another key process influencing the thermal regime of the shallow subduction zone is frictional heating along the plate interface. Temperature along the subduction thrust is extremely sensitive to the coefficient of friction, whereas surface heat flow is much less sensitive to this parameter. Several studies have explored the impact of frictional heating and hydrothermal circulation on surface heat flow (e.g., Rotman and Spinelli, 2014; Hass and Harris, 2016; Wang et al., 2015). These studies are consistent in finding that hydrothermal circulation and frictional heat have opposing effects on plate interface temperatures and surface heat flow above the subduction thrust. In general, hydrothermal circulation decreases temperatures on the subducting thrust. The magnitude of cooling on the subduction thrust generally increases with depth until the aquifer loses sufficient permeability. In contrast, frictional heating increases plate boundary temperatures. Because frictional heating is proportional to the normal stress (Equation 2), warming increases with depth until ~350 °C, where the rheology changes from brittle to ductile. Figure 7 shows that these processes work over a similar length scale, but as the age of the incoming basement ages and the brittle-ductile transition deepens, peak frictional heating shifts landward. Studies evaluating both hydrothermal circulation and frictional heating indicate that frictional heat can be efficiently extracted from the system by hydrothermal circulation. As a result, in order to use heat flow observations above a subduction thrust to constrain the amount of frictional heating on the fault, independent knowledge of the vigor of hydrothermal circulation or the lack of hydrothermal circulation is required. In Hass and Harris (2016) the root mean square misfit between thermal models and heat flow data at Costa Rica were evaluated as a function of the initial permeability and coefficient of frictional heating; for Costa Rica, it was found that the initial permeability is relatively well resolved and that the coefficient of friction is less well resolved.

Abundant evidence for high pore fluid pressure (i.e., pressure greater than hydrostatic) along the subduction thrust seems at odds with a high-permeability aquifer immediately below it. Sediments between the basaltic aquifer and subduction thrust must be relatively impermeable to maintain high pore pressure along the fault, but little is known about the permeability or continuity of this seal.

**CONCLUSIONS**

At subduction zones, the thermal regime of the incoming plate has a large influence on temperatures on and near the subduction thrust. This study summarizes thermal evidence for hydrothermal circulation in subducting oceanic crust. Heat flow data from Nankai, Costa Rica, south-central Chile, and Haida Gwaii are best fit by models incorporating hydrothermal circulation in subducting oceanic crust. This result is consistent with the young basement ages (35–8 Ma) of the incoming igneous crust and results from global heat flow analysis showing a significant heat flow deficit for crustal ages younger than 65 Ma (e.g., Stein and Stein, 1992) interpreted in terms of advective fluid flow.

In ridge flank settings, buoyancy driven hydrothermal circulation is pervasive in the basaltic basement aquifer. Where sediment is thin and discontinuous and basement is exposed, the oceanic crust tends to be well ventilated. Surface heat flow associated with subducting oceanic crust that is well ventilated just seaward of the deformation front is characterized by low magnitudes relative to conductive predictions and large variability. Surface heat flow just landward of the deformation front may increase as the crustal aquifer becomes insulated. Subducting crust that includes thick and continuous sediment cover may also host vigorous fluid flow. In this environment, surface heat flow values may approach conductive predictions on a regional basis, but variability is larger than can be explained with conductive refraction. Near the deformation front, heat flow may be much higher than conductive predictions and decrease landward at a greater rate than can be explained by plate convergence and geometry and landward thickening of the margin. If hydrothermal circulation is present within the incoming crust, it is likely to continue with subduction as long as permeability remains sufficiently high to allow convection to occur.

The presence of hydrothermal circulation within subducting oceanic crust has a large influence on plate boundary temperatures. In general, hydrothermal circulation leads to plate boundary temperatures that are cooler than those where fluid flow does not occur. The magnitude of cooling depends on the permeability structure of the incoming plate and the evolution of permeability with depth and time. Resolving complex relationships between subduction processes, the permeability structure in the ocean crust, and the dynamics of hydrothermal circulation remains an interdisciplinary frontier.

**ACKNOWLEDGMENTS**

Harris was supported by National Science Foundation (NSF) grants OCE 1458211 and OCE 1355878. Spinelli was supported by NSF grant OCE 1515167. Fisher was supported by NSF grants OIA-0939564, OCE-1260408, and OCE-1355870. We thank Guest Associate Editor Laura Wallace and reviewers Carol Stein and Michael Underwood for providing constructive comments that improved this study. This is C-DEBI (Center for Dark Energy Biosphere Investigations) contribution 373.

**REFERENCES CITED**


Baba, T., Cummins, P.R., Hori, T., and Kaneda, Y., 2006, High precision slip distribution of the 1944 Tonga-Kermadec earthquake inferred from tsunami waveforms: Possible slip on a splay fault: Tectono-


Contreras-Reyes, E., Greveveymer, I., Flesch, E.R., Scherwath, M., and Heesemann, M., 2007, Alter-


