Surface processes and tectonics: Forcing of continental subduction and deep processes

E. Burov*, G. Toussaint

University Paris VI, Case 129, 4 Place Jussieu, 75252, Paris, France

Received 13 November 2006; accepted 7 February 2007
Available online 6 April 2007

Abstract

It is now well accepted that surface processes provide a critical feedback on the surface tectonic deformation, whatever it is, orogenic building or basin evolution. However, the idea that the influence of these processes may go below the crustal levels, is less common. In this preliminary study, we use coupled thermo-mechanical numerical models to investigate the possible influence of surface processes on the styles of continental collision, in particular, continental subduction. For that, we further exploit the recent successful model of continental subduction of the early stages of India–Asia collision by Toussaint et al. [Toussaint G., Burov, E., and J.-P. Avouac, Tectonic evolution of a continental collision zone: a thermo-mechanical numerical model, Tectonics, 23, TC6003, doi:10.1029/2003TC001604, 2004b]. On the example of India–Asia-like settings, we show that not only the surface topography but also the total amount of subduction may largely vary as function of denudation rate (controlled by the coefficient of erosion, k). Erosion provides a dynamic discharge of the hanging wall of the major thrust zone, whereas the sedimentation increases loading on the footwall and this helps down-thrusting of the lower plate. Both processes reduce the resistance of the major thrust and subduction channel to subduction. However, very strong or very slow erosion/sedimentation enhance the possibility of plate coupling and promote whole-scale thickening or buckling. The maximal amount of subduction is thus achieved for some intermediate values of erosion rates when the tectonic uplift rate is fine-balanced by the denudation rate. In our case the optimal balance is reached for the values of k on the order of 3000 m²/yr. We then extended our model beyond the conditions of India–Asia collision, in terms of the tested range of k and convergence rates. The experiments suggest that for provided settings, both extra slow (k<50–100 m²/yr) and extra rapid erosion (k>6000–8000 m²/yr) limit, by up to 50%, the total amount of subduction, if not totally prevent it. The model demonstrates the large capability of surface processes to adopt to different deformation styles: the orogenic building and subduction successfully develop (subduction number, S>0.5) in the range of k between 500 m²/yr and 6000 m²/yr at convergence rates ranging from 1 cm/yr to 6 cm/yr. Within this range, some peculiar features of orogenic style such as the geometry of the accretion prism, amount of upper crustal subduction, horizontal progression of the mountain range/thrust fault and the amount of exhumation of metamorphic facies are sometimes quite different. We conclude that surface processes may control deep, mantle level tectonic evolution.

© 2007 Elsevier B.V. All rights reserved.

Keywords: surface processes and tectonics; mountain building; collision tectonics; rheology; numerical modelling; erosion

1. Introduction

A number of workers have emphasized the importance of dynamic feedbacks between surface processes and tectonic evolution (e.g., Molnar and England, 1990;
Masek et al., 1994a; Kooi and Beaumont, 1994, 1996; Avouac and Burov, 1996; Snyder et al., 2000; Molnar, 2001; Basile and Allemend, 2002; Garcia-Castellanos, 2002; Garcia-Castellanos et al., 2002, 2003; Simpson and Schlunegger, 2003; Persson et al., 2004; Cloetingh et al., 2004). Indeed, surface processes modify the topography and redistribute tectonically very significant volumes of sedimentary material, acting as vertical, or normal loads, over large horizontal distances. This may result in dynamic loading and unloading of the underlying crust and mantle lithosphere (e.g., Willett, 1999). At the same time, tectonically induced topographic contrasts are absolutely required to set up erosion and sedimentation processes. Consequently, tectonics could be a forcing factor of surface processes, whereas the surface processes could be a forcing factor of tectonics, which evokes a positive feedback.

Continental maintain belts, such as, for example, Himalaya or Tien Shan (Central Asia, Fig. 1), are characterized by topography elevations persistently growing over tens of millions years. However, the fact that gravitational potential energy per unit surface $g \rho h^2$ scales as $h^2$ implies that a thrust belt should grow more easily in width than in height (Molnar and Lyon-Caen, 1988, $h$ is the mean topography elevation above sea level, $p$ is density and $g$ is acceleration due to gravity). Indeed, a range results from thrusting on faults that cut through the upper crust and root into the lower crust. Uplift of the range implies an increase in the vertical stress acting on the fault. This acts to oppose further frictional sliding on the fault, inhibiting further thrusting. A new fault will then form farther away from the range front leading to widening of the range. For example, 5 km vertical displacement on the fault (average height of Himalaya) would be associated with frictional stress along the fault plain of up to 300 MPa (0.6–0.8 of $\rho gh$ assuming typical friction angle of $30^\circ$–$40^\circ$ and absence of slab pull on the lower plate). This

Fig. 1. Actively growing intercontinental belts and plateaus: an example showing a schematic map of India–Eurasia collision with its main features such as the Himalayan mountain belt and Tibetan plateau. The topography peaks to 8800 m in Himalaya. Modified from Avouac and Tapponnier (1993).
stress is too high compared to the strength of the crustal lithosphere and to the average estimates of the mean intraplate stresses (50–100 MPa). Consequently, the subduction may only occur if the frictional stress is reduced by a factor of 3–5. Such reduction may be produced by some intrinsic strength softening mechanisms resulting in reduction of the friction angle from 33° to 5–10°. However, such mechanisms are not very well known. The other, more natural way to reduce friction on the fault would be to reduce normal stress acting on the fault plain. In ocean subduction settings, this normal stress is reduced since the lower plate subducts under its own weight and only weakly interacts with the upper plate. In continents, the subducting plate has far much higher floatability and the interaction with the upper plate may be ultimately stronger. In this situation, it may be suggested that the major thrust faults and the growth of the surface topography can be localized in space and time by dynamic discharge of the fault surfaces and of the bulk of the supporting lithosphere. This discharge can be provided by surface processes (Fig. 2A) that continuously remove new elevating topography and transport the resulting sediment to the flexural foreland basin provoking its additional subsidence (e.g.,

![Diagram](image)

**Fig. 2.** A. Conceptual model of interaction between surface and subsurface processes in continental subduction settings. Large empty arrows correspond to the directions of material flow. Large filled arrows correspond to major stresses acting on the crust and lithosphere. Black arrows correspond to velocities. $\sigma_n$ is normal stress, $\tau_f$ is mean frictional stress in the major thrust/subduction channel. Erosion of the surface topography results in discharge of the hanging wall whereas filling of the foreland basin leads to increased loading of the footwall. Thus flexes the lower plate down thus enlarging the subduction channel and reducing coupling between the upper and lower plate. B. Typical rheology profiles for continental lithosphere: (1) geotherms that yield YSEs shown in the middle and on the right; (2) yield stress envelope (YSE) for quartz-rich upper and lower crust and olivine mantle; (3) yield stress envelope (YSE) for quartz-rich upper crust, diabase lower crust, olivine mantle. EET — equivalent elastic thickness of the lithosphere computed for each of YSEs. Profile a-1 corresponds to the YSE used for the lower plate in our numerical experiments. Same profile but assuming a 10 km thicker crust was used for the upper plate (this maintains the upper plate afloat due to higher floatability of its crust). In this study, we do not try to reproduce formation of the Tibetan plateau. Otherwise the rheology profile for the upper plate had to be b-1, according to the present day estimates of EET.
Loading of the flexural basin compensates the upward restoring force due to the flexure and isostatic compensation of the lower plate, and helps to force this plate down (Burov et al., 1990). Sedimentary filling thus may decrease the normal stress exerted on the thrust fault. Selective removal of elevated topography loads favours compensatory uplift of the deep material towards the zones that are most affected by erosion, and helps to localize orogenic building in one area (e.g., Avouac and Burov, 1996), as well as to maintain sliding on a single fault rather than on several faults. Foreland basin typically accumulate several km of sediments, which is equivalent to 10–150 MPa of normal stress increment on the footwall of the fault, or to vertical force increment on the order of $10^{12}$–$10^{13}$ N per unit length, which is just the same order as typical estimates for slab pull forces (e.g., Bott, 1993). Erosion produces a comparable dynamic unloading of the hanging wall. For example, the estimated rock uplift rates in Tien Shan (Central Asia) suggest that without erosion this range (Avouac and Burov, 1996) would be 10 to 15 km high, instead of the actual 3–5 km. Such a range would exert normal stresses $gph$ of up to 400 MPa on the supporting basement (assuming density of the rocks $\sim 2700$ kg m$^{-3}$), or a net vertical force, $f_{h}$, of $10^{14}$ N per unit length. This value is 10 to 100 times higher than typically estimated horizontal tectonic forces, $f_{h}$ (on the order of $10^{11}$ to $10^{12}$ N per unit length; Bott, 1993), that normally should balance the vertical forces. The resulting Argand numbers ($Ar = f_{h}/f_{l}$) would be on the order of 0.1 or 0.01, which implies physical impossibility of mountain growth (because $Ar$ should be $\geq 1$ for stable growth). Erosion reduces $f_{h}$ and thus allows for mountain growth when $f_{h}$ is on the order of $f_{l}$.

Of course, despite of the importance of erosion and sedimentation, mountain growth primarily depends on the mechanism of accommodation of tectonic shortening: subduction (simple shear), collision (pure shear), folding etc. Yet, these deep mechanisms were always thought to be independent of surface processes. We suggest that principal modes of continental deformation such as subduction may be favoured (or disfavoured) by the way the surface processes distribute vertical loads acting on the upper and lower plate. As mentioned above, erosion discharges the hanging wall of the major thrust faults while sedimentation provides extra loading on the footwall and the flexural foreland basin (sedimentary accretion, Fig. 2A). Consequently, surface processes may limit normal stresses and friction on the surface of the major thrust faults. By this bias they may ease the subduction process that needs overall low resistance of the thrust fault to sliding. As suggested in a number of studies (e.g., Toussaint et al., 2004a), the initial phases of continental convergence may be accommodated by subduction of one continental plate below another. Subduction of the entire lithosphere may not probably continue during an important period of time ($>10$ Myr) since the continental plates have a positive mean floatability due to their thick low density crust. Continental subduction may therefore continue only if the “continental slab” is mechanically strong enough to be pushed–pulled by tectonic or slab pull forces, and, of course, if these forces are sufficient to overcome those due to positive floatability of the slab and viscous resistance of the surrounding asthenosphere. For given rheology, the strength of the slab at depth depends on the ratio of the thermal diffusion rate to the subduction rate (thermal softening), plate bending (flexural softening) and a number of other factors. For example, thermal softening results in overall weakening of the slab after several Myr (Toussaint et al., 2004a). Starting from this point subduction is impossible and plates accommodate shortening by pure shear or folding (Toussaint et al., 2004a). In the numerical models of collision/subduction, one can trace the amount of subduction (subduction length, $S$) and compare it with the total amount of shortening on the borders, $\Delta x$. The subduction number $S$, which is the ratio of these two values, may be used to characterize the deformation mode:

$$S = s_{1}/\Delta x$$  \hspace{1cm} (1)

When $S=1$, all shortening is likely to be accommodated by subduction, which refers to full subduction mode. In case when $0.5 < S < 1$, pure shear or other deformation mechanisms participate in accommodation of shortening. When $S<0.5$ subduction is no more a leading mechanism of shortening. Finally, when $S>1$, one deals with full subduction plus a certain degree of “unstable” subduction. This refers to the cases of high $s_{1}$ ($>300$ km) when a large portion of the subducted slab is reheated by the surrounding hot asthenosphere. As a result, the deep portion of the slab mechanically weakens and can be additionally stretched by gravity forces (slab pull). In general, $S>1$ corresponds to the initial stages of slab break-off. $S>1$ is also often associated with the development of Rayleigh–Taylor instabilities in the weakened part of the slab.

In this paper, we test the influence of surface processes on deep sub-crustal tectonics, and thus on “$S$” number, using as reference scenario our previous successful model of India–Asia collision (Toussaint et al., 2004a)
et al., 2004b). This model has demonstrated the possibility of continental subduction of India beneath Tibet during the first stages (10–15 Myr) of the collision. The experiments made by Toussaint et al. (2004b) tested only the case of fine balance between the surface and subsurface rates, using the “stable” values predicted in (Avouac and Burov, 1996). We here consider a large range of erosion and convergence rates to test the sensitivity of deep tectonic evolution to surface processes. The particular interest of testing the model for the conditions of the India–Himalaya–Tibet collision refers to the fact that this zone, characterized both by intensive convergence (Patriat and Achache, 1984) and erosion (e.g., Hurtrez et al., 1999), belongs to the same geodynamic framework of India–Eurasia collision as the Tien Shan range, which growth was shown to be strongly controlled by balance between surface and subsurface processes (Avouac and Burov, 1996).

2. Interplays between surface and tectonic processes

2.1. Tectonic forcing on surface processes

Surface topography elevations are required to set up erosion and sedimentation processes. Following Ahnert (1970), and Pinet and Souriau (1988), Summerfield and Hulton (1994) have compiled rates of denudation at the scale of major river basins. These studies indicate that denudation is primarily influenced by basin topography so that rates of denudation appear to be systematically high in areas of active tectonic uplift. Common values of mean denudation rates in such areas would be of the order of a few 0.1 mm/yr to about 1 mm/yr at the scale of large drainage basins. Such rates are generally consistent with estimates derived from balancing sediment volumes over geological periods of time (Leeder, 1991; Summerfield and Hulton, 1994). Thermochronologic studies indicate, however, local values as great as 1 mm/yr (Molnar and England, 1990; Leeder, 1991). The discrepancy between local and basin-averaged estimates is due to the fact that tectonic uplift is probably distributed in brief pulses over localized domains within a drainage basin (Copeland and Harrison, 1990). Pinet and Souriau (1988) have demonstrated that denudation leads to an exponential decay of the topography of a range with a characteristic time constant of the order of 2.5 m.y. Thus, in the absence of a strong tectonic feedback, common values of denudation rates should lead to the disappearance of major mountain belts such as the Alpes, Himalaya or Tien Shan in a few million years. Since these ranges persist for tens of millions of years, the presence of some strong tectonic feedback appears to be inevitable.

2.2. Coupling between denudation and tectonic uplift due to isostasy

Many recent models have investigated coupling between isostatic reaction and surface processes (e.g., Beaumont, 1981; Beaumont et al., 1992; Kooi and Beaumont, 1994; Willett, 1999; Snyder et al., 2000; Basile and Allemand, 2002; Garcia-Castellanos, 2002; Garcia-Castellanos et al., 2002, 2003; Simpson and Schlunegger, 2003; Persson et al., 2004; Castelltort and Simpson, 2006). Redistribution of surface loads by erosion and sedimentation must induce tectonic deformation to maintain isostatic balance. Vertical uplift is expected to partly compensate unloading in the area subjected to denudation while subsidence should occur in response to loading by sedimentation. This feedback mechanism may lead to some coupling between denudation and tectonic uplift (e.g., Ahnert, 1970). A first consequence is that the time needed to erode a topographic relief must take into account removal of the topographic relief and of the crustal root. If local isostasy is assumed and if horizontal strains are neglected, denudation is dynamically compensated by uplift and the characteristic time of decay of the topography would then be of the order of 10 m.y. (Leeder, 1991). In addition, it has been argued that a positive feedback may arise (Molnar and England, 1990; Masek et al., 1994b; Avouac and Burov, 1996; Burov and Cloetingh, 1997). If the slopes of valley steepen during river incision, isostatic readjustment following denudation in a mountain range may result in a net uplift of the higher summits in spite of the average lowering of reliefs. Alternatively regional compensation due to the elasticity of the lithosphere might lead to the uplift of the eroded edge of a plateau. Erosion might therefore induce some uplift of topographic summits leading in turn to enhanced erosion. The uplift of the Himalayan belt during the last few million years may have resulted from such a coupling rather than from thrusting at the Himalayan front (Burbank, 1992; Burbank and Vergés, 1994).

However, while the peaks might reach higher elevations following isostatic adjustment, the net effect of erosion is crustal thinning. Thus, pure isostatic models cannot explain the growth of mountains over long time periods. Moreover, the strongest feedback between erosion and isostatic reaction would be obtained for local isostasy. It will be mitigated in case of regional compensation and become negligible for
lithospheres which equivalent elastic thickness exceeds 60 km. This is a reason why more efficient feedback mechanisms should also take place in collisional settings. These mechanisms may refer to lower crustal flow, that provides rapid feedback between surface loading and deep deformation (Avouac and Burov, 1996), and to the above mentioned mechanisms of flexural discharging of the major thrust faults that reduce frictional resistance and enable long-lasting subduction.

2.3. Coupling between surface processes and horizontal strains

As mentioned in the previous sections, small lateral variations of the crustal thickness should drive horizontal flow in the lower crust. Some studies have already pointed out to the importance of such process in continental tectonics (e.g., Lobkovsky, 1988; Lobkovsky and Kerchman, 1991; Avouac and Burov, 1996; Burov and Cloetingh, 1997). For example, Kruse et al. (1991) have shown that horizontal flow in the lower crust may have regulated isostatic equilibrium during extension in the Basin and Range. The lower crust would have been extruded from under the high topography during that process. Following Westaway (1994) we will call this sense of flow “outward”. On the other hand, Gregory and Chase (1994) inferred “inward” flow, toward the crustal root, during the Laramide orogeny of the Frontal Range, Colorado. The characteristic time associated with flow in the lower crust induced by the topography of a range a few thousands of m high, a few hundreds of km wide, is in the order of a few m.y. The characteristic times of erosional decay of the topography of a range and of lateral collapse of a crustal root are thus of the same order of magnitude. Since both processes are driven by topographic slopes, some coupling must arise. Although it is not often pointed out, it has long been recognized that this kind of process might play a major role in elevation changes within continents (see Westaway, 1994 for an review of historical development of these ideas). Westaway (1994) made a case for such a coupling, with inward flow, in the context of extensional tectonics in western Turkey. He proposed that sediment loading in the sedimentary basins would have driven flow toward the uplifted area. This kind of process was first modelled by King and Ellis (1990) who modelled crustal extension using a thin elastic plate (upper crust) overlying an inviscid fluid (lower crust).

We proposed (Avouac and Burov, 1996) that this kind of coupling might also appear in compressional context. The lithosphere in the region of the range is weakened, since (1) the crust is thick and hot, and because (2) bending of the lithosphere beneath the mountain load tends to reduce its flexural strength (Burov and Diament, 1992, 1995; Ranalli, 1995). Higher strain rates in the area below the range should then be expected. A low viscosity channel in the lower crust beneath the high topography might therefore allow lateral flow. In the absence of horizontal shortening and erosion, the lower crust below the range would be extruded laterally as discussed by Bird (1991) and Gratton (1989). If erosion takes place, a regime may be established in which horizontal shortening would be preferentially accommodated either by crustal thickening in the area below the range or by down-thrusting of the lower plate along the major thrust faults-bordering high topography (Fig. 2A):

a) Surface processes remove material from the range and feed the adjacent flexural basins inducing isostatic imbalance.

b) This imbalance produces a temporary excess of normal stress below the foreland basins and stress deficit below the range favouring flow in the lower crust towards the crustal root. The range uplifts reducing flexural loading on the major thrust fault, while the basins subside, forcing the hanging wall down. Both these processes tend to widen the subduction channel, reduce friction on the fault and, consequently, favour subduction.

Ultimately this coupled regime might lead to some dynamic equilibrium in which the amount of material removed by erosion would balance the material supplied to the range by subsurface deformation.

Apart of the direct mechanical effect of erosion/sedimentation (loading/unloading) on the lithosphere, it also has a very important thermal, and, by proxy, mechanical consequences, because the removal and accumulation of sedimentary matter modifies surface heat flux and thermal conditions in the upper crust (e.g., England and Richardson, 1977). Accumulation of sediments in the forelands leads to (1) cooling of the accretion wedge at a short term, in case of rapid advection/filling (initial stages of collision when the convergence rate is highest); (2) heating of the accretion wedge at long term in case of slow advection, when collision rate slows down, due to heat screening (sediments have low thermal conductivity) and the abundance of heat producing radiogenic elements in the sedimentary matter. Furthermore, penetration of the mechanically weak sediment in the subduction channel...
should serve as lubrication and may enhance the conditions for subduction processes.

2.4. Coupling of surface processes and tectonic input/reaction in full scale mechanical models: major stages

A number of earlier modelling studies (e.g., Beaumont, 1981; Beaumont et al., 1992; Willett, 1999) have investigated various relationships between erosion and tectonic processes. However, tectonic reaction was not fully accounted as most of these models have exploited semi-kinematic formulations for the crust or the mantle lithosphere. These models thus cannot be used to study the influence of erosion on deep tectonic processes such as subduction. One of the first full-scale parametric semi-analytical models was developed by Avouac and Burov (1996) in order to validate the coupled regime between surface and subsurface processes. This model accounted for: (1) surface processes, (2) the effect of topographic loads and variations of crustal thickness on the mechanical behaviour of the lithosphere, (3) ductile flow in the lower crust, (4) depth and strain dependent rheology of the lithosphere. This semi-analytical model has a number of limitations in terms of model geometry and its inability to account for some key deformation modes such as formation of major thrust faults and subduction. For this reason, we go further by introducing an unconstrained fully coupled numerical thermo-mechanical model of continental collision/subduction similar to that used by (Burov et al., 2001; Toussaint et al., 2004a,b). This model takes into account more realistic (than in the previous studies) geometry of the convergent plates, accounts for large strains and brittle–elastic–ductile rheology including localized brittle (faulting) and ductile deformation.

3. Surface processes modelling: principles and numerical implementation

3.1. Basic models of surface processes

A growing amount of field and experimental studies have investigated and validated various forms of long and short range erosion and sedimentary transport laws and models (Ahnert, 1970; Beaumont, 1981; Ashmore, 1982; Pinet and Souriau, 1988; Burbank, 1992; Beaumont et al., 1992; Burbank and Vergès, 1994; Densmore et al., 1997, 1998; Mizutani, 1998; Beaumont et al., 2000; Davy and Crave, 2000; Crave et al., 2000; Lague et al., 2000; Lavé and Avouac, 2001; Molnar, 2001; Crave and Davy, 2001; Lague et al., 2003).

3.1.1. Short range erosion

A simple two-dimensional law may be used to simulate erosion and sedimentation at the scale of a mountain range. The evolution of a landscape results from the combination of weathering processes that prepare solid rock for erosion, and transportation by hillslope and stream processes (see Carson and Kirkby, 1972 for a review). Although many factors, depending on the lithologies and on climate (e.g., Fournier, 1960; Nash, 1980), may control this evolution, quite simple mathematical models describing the geometrical evolution of the morphology at the small scale have been proposed and tested successfully (e.g., Kirkby, 1971, 1986; Smith and Bretherton, 1972; Luke, 1972, 1974; Chorley et al., 1984; Kirkby et al., 1993). For example, the two-dimensional evolution of a scarp-like landform can be modelled assuming that the rate of downslope transport of debris, \( q \), is proportional to the local slope, \( \nabla h \) (Culling, 1960, 1965; Hanks et al., 1984; Avouac, 1993; Kooi and Beaumont, 1994, 1996; Braun and Sambridge, 1997):

\[
q = -k \nabla h,
\]

where \( k \) is the mass diffusivity coefficient, expressed in units of area per time [e.g., m\(^2\)/yr]. Assuming conservation of matter along a 2-D section and no tectonic deformation, \( h \) must obey:

\[
\frac{dh}{dt} = -\nabla q.
\]

With constant \( k \), Eqs. (1) and (2) lead to the linear diffusion equation:

\[
\frac{dh}{dt} = k \nabla^2 h.
\]

We consider the values for \( k \) varying from 50 to 10\(^3\) m\(^2\)/yr that yield denudation rates of the order of a few 0.01 mm/yr to 1 mm/yr for a 500 km wide range with a few thousand meters of relief. This model of surface processes holds only for particular conditions. The regolith must form more rapidly than it is removed by surface transport and slopes must not exceed the frictional angle of the material. Even for scarps formed in loose alluvium some complications arise when high scarps are considered. Scarps with height typically in excess of about 10 m in arid climatic zones, tend to have systematically sharper curvatures at crest than at base (e.g., Andrews and Bucknam, 1987). Gravity-driven erosion processes such as hillslope landsliding impose strong limitations on the applicability of the diffusion equation since these processes are rather slope — then curvature dependent, which basically requires to introduce slope and height dependent terms in the Eq. (4a). At a larger
scale, hillslope and stream processes interact and the sediment transport then depends non-linearly on the slope and on other factors such as the slope gradient, the area drained above the point, the distance from the water divide, so that the simple 2-D linear diffusion does not apply in general (e.g., Gossman, 1976). In spite of these limitations, we have chosen to stick to a linear diffusion law to model erosion in the upland. Although this model does not accurately mimic the spatial distribution of denudation in the mountain range, it leads to a sediment yield that is roughly proportional to the mean elevation of the basin relative to that point (a rough approximation to the sediment yield resulting from a change of elevation \( h \) over a horizontal distance \( d \) is \( k \times h / d \)) and therefore accounts for the apparent correlation between elevation and denudation rates (Ahnert, 1970; Pinet and Souriau, 1988; Summerfield and Hulton, 1994). We did not apply the diffusion model to the whole system, however. We felt that we should take into account the major discontinuity in surface processes that occurs at the mountain front/thrust fault. As a river emerges into the adjacent basin its gradient is sharply reduced and deposition occurs. The streams shift from side to side and build up alluvial fans and tend to form a broad gently sloping pediment at the base of the mountain range. In addition, a lateral drainage often develops along the foothills of mountain ranges. The Ganges along the Himalayan foothills, the Parana along the Andes or the Tarim along the Tien Shan are good examples. Altogether the formation of the pediment and lateral drainage tend to maintain gentle slopes in the foreland. There is therefore a sharp contrast between river incision that maintains a rugged topography with steep slopes in the mountain range and widespread deposition of alluvium in the foreland. This discontinuity of processes must be considered to model the sharp break-in-slope at the mountain front that is generally observed on topographic profiles across mountain belts. In order to simulate this major change in surface processes, sedimentation in the lowland is modelled assuming flat deposition by fluvial network: we assume conservation of matter along the section and the sediment at the mountain front is distributed in order to maintain a flat horizontal topography in the foreland. We arbitrarily set the change from diffusional erosion to sedimentation (“flat deposition”) at a differential elevation of 500 m, which is, however, representative for the transition from highlands to forelands.

Finally, it is noteworthy that the erosion law can be also considered in a non-linear form:

\[
dh/dt = k^*(x, h, \nabla h) \nabla^2 h
\]

(4b)

where \( k^*(x, h, \nabla h) = k(x)(\nabla h)^n \) (e.g., Gossman, 1976; Andrews and Bucknam, 1987). The phenomenological Eq. (4b) differs from one obtained assuming a non-linear diffusion coefficient in the Eq. (2). In that case conservation of mass would lead to an additional term \( \nabla k^* \nabla h \). The existing non-linear erosion laws are not limited to the Eq. (4b) (e.g., Newman, 1983; Newman and Turcotte, 1990), which only presents the simplest way to account for dependence of erodibility on the morphology. In this study we test only first order effects of erosion on the collision processes, for which the Eq. (4a) is sufficient.

3.1.2. Long-range surface processes

The long range surface processes are associated with fluvial transport, i.e. with river incision, slope geometry, character of sediment matter and conditions for deposition (Flint, 1973, 1974; Shepard and Schumm, 1974; Hirano, 1975; Schumm et al., 1987; Seidl and Dietrich, 1992; Govers, 1992a,b; Hairine and Rose, 1992; Howard et al., 1994; Sklar and Dietrich, 1998; Howard, 1998; Smith, 1998; Davy and Crave, 2000; Snyder et al., 2000; Sklar and Dietrich, 2001; Crave and Davy, 2001; Snyder, 2001; Hancock and Willgoose, 2001; Simpson, 2004). The characteristic laws for this range are different as the these mechanisms are dependent on the incision and transport capacity of the fluvial network, local slope and type of sediment. Deep steep rivers can carry sediment to longer distances as it can be caught in turbulent flow layer. Shallow rivers would deposit sediment rapidly resulting in rapid river blockage and frequent change of the direction of the fluvial network. There is also a strong dependence of transport capacity on the grain size and climate episodicity (e.g., Davy and Crave, 2000). The long-range fluvial models were used with success by Kooi and Beaumont (1994, 1996), Garcia-Castellanos (2002), Garcia-Castellanos et al. (2002, 2003), Persson et al. (2004). The cumulative material flow, \( q_{fe} \), due to the fluvial transport can be presented, in most simple form, as:

\[
q_{fe} = -K_r q_t dh/dl
\]

(4c)

where \( q_t \) is the river discharge, \( K_r \) is non-dimensional transport coefficient and \( dh/dl \) is the slope in the direction of the river drainage with \( l \) being the distance along the transporting channel. The diffusion Eq. (4a), except if it is not strongly non-linear (4b), provides symmetrical, basically over-smoothed shapes whereas the fluvial transport Eq. (4c) may result in more realistic asymmetric behaviours, because, locally, the direction of each bi-furcation of the fluvial network is affected.
by negligibly small factors, even though the overall direction of the flow is controlled by the regional slope of topography. Any important change in the regional slope of topography, such as at the transition from tectonically built steep highlands to flat sedimentary built forelands, may result, at some moment, in a drastic change of the direction of the fluvial network, which may choose a principally new stream direction orthogonal to that of the highland network (as it is the case of the Ganges river, for example). This happens when the sedimentary basin is filled to a point that the inclination of its surface in the direction of tectonic convergence becomes less important than that in some other direction (basically in the direction of the boundary between the steep highlands and flat lowlands). In this study we use a highly simplified form of

<table>
<thead>
<tr>
<th>Variable</th>
<th>Values and units</th>
<th>Definition</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_{xx}$, $\tau_{xy}$, $\tau_{yy}$</td>
<td>Pa, MPa</td>
<td>Viscous stress components</td>
<td></td>
</tr>
<tr>
<td>$\sigma_{xx}$, $\sigma_{xy}$, $\sigma_{yy}$</td>
<td>Pa, MPa</td>
<td>Full stress components</td>
<td>$\sigma = \tau - P I$, $\sigma_{xy} = \tau_{xy}$ etc.</td>
</tr>
<tr>
<td>$P$</td>
<td>Pa, MPa</td>
<td>Pressure</td>
<td></td>
</tr>
<tr>
<td>$v$</td>
<td>m/s, mm/yr</td>
<td>Total velocity vector</td>
<td></td>
</tr>
<tr>
<td>$u$</td>
<td>m/s, mm/yr</td>
<td>Horizontal velocity</td>
<td></td>
</tr>
<tr>
<td>$v$</td>
<td>m/s, mm/yr</td>
<td>Vertical velocity</td>
<td></td>
</tr>
<tr>
<td>$\mu$</td>
<td>Pa s</td>
<td>Effective viscosity</td>
<td>$10^{19}$ to $10^{25}$ Pa s</td>
</tr>
<tr>
<td>$k$</td>
<td>m$^3$/yr</td>
<td>Coefficient of erosion</td>
<td>~ Mass diffusivity</td>
</tr>
<tr>
<td>$dh$</td>
<td>m, km</td>
<td>Topographic uplift</td>
<td>Or subsidence</td>
</tr>
<tr>
<td>$du$</td>
<td>m, km</td>
<td>Tectonic uplift</td>
<td>Do not mix with $u$</td>
</tr>
<tr>
<td>$de$</td>
<td>m, km</td>
<td>Erosion</td>
<td>Or sedimentation</td>
</tr>
<tr>
<td>$\varepsilon$</td>
<td></td>
<td>Strain</td>
<td></td>
</tr>
<tr>
<td>$\dot{\varepsilon}$</td>
<td>s$^{-1}$</td>
<td>Effective strain rate</td>
<td>$\dot{\varepsilon} = (1/2 \varepsilon_{ij} \varepsilon_{ij})^{1/2}$</td>
</tr>
<tr>
<td>$E$</td>
<td>$8 \times 10^{10}$ N/m$^2$</td>
<td>Young’s modulus</td>
<td></td>
</tr>
<tr>
<td>$\nu$</td>
<td>0.25</td>
<td>Poisson’s ratio</td>
<td>In the semi-analytical model</td>
</tr>
<tr>
<td>$\lambda$, $\mu_e$</td>
<td>N/m$^2$</td>
<td>Lamè’s constants</td>
<td>In the semi-analytical model</td>
</tr>
<tr>
<td>$A^*$</td>
<td>Pa$^{-n}$s$^{-1}$</td>
<td>Material constant</td>
<td>Power law</td>
</tr>
<tr>
<td>$n$</td>
<td>3 to 5</td>
<td>Stress exponent</td>
<td>Power law</td>
</tr>
<tr>
<td>$H^*$</td>
<td>kJ mol$^{-1}$</td>
<td>Activation enthalpy</td>
<td>Power law</td>
</tr>
<tr>
<td>$R$</td>
<td>8.314 J/mol K</td>
<td>Gas constant</td>
<td>Power law</td>
</tr>
<tr>
<td>$T$</td>
<td>°C, K</td>
<td>Temperature</td>
<td></td>
</tr>
<tr>
<td>$\gamma(v)$</td>
<td>Pa/m, MPa/km</td>
<td>Depth gradient of yield stress</td>
<td>$\gamma(v) \propto d (e) / dt$</td>
</tr>
<tr>
<td>$w$</td>
<td>m, km</td>
<td>Plate deflection</td>
<td>~ Deflection of mantle lithosphere</td>
</tr>
<tr>
<td>$T_e$</td>
<td>m, km</td>
<td>Effective elastic thickness</td>
<td>~ Instant integrated strength</td>
</tr>
<tr>
<td>$\rho_{c1}$</td>
<td>2650 kg/m$^3$</td>
<td>Density</td>
<td>Of upper crust</td>
</tr>
<tr>
<td>$\rho_{c2}$</td>
<td>2900 kg/m$^3$</td>
<td>Density</td>
<td>Of lower crust</td>
</tr>
<tr>
<td>$\rho_m$</td>
<td>3330 kg/m$^3$</td>
<td>Density</td>
<td>Of mantle</td>
</tr>
<tr>
<td>$g$</td>
<td>9.8 m/s$^2$</td>
<td>Acceleration due to gravity</td>
<td></td>
</tr>
<tr>
<td>$T_b$</td>
<td>°C</td>
<td>Temperature</td>
<td></td>
</tr>
<tr>
<td>$t_a$</td>
<td>My</td>
<td>Thermal age</td>
<td>≤ Geological age</td>
</tr>
<tr>
<td>$a$</td>
<td>250 km</td>
<td>Thermal thickness of the lithosphere</td>
<td></td>
</tr>
<tr>
<td>$T_m$</td>
<td>1330 °C</td>
<td>$T$ at depth $a$</td>
<td>Base of lithosphere</td>
</tr>
<tr>
<td>$\chi$</td>
<td>m$^2$ s$^{-1}$</td>
<td>Thermal diffusivity</td>
<td>$\chi = k / \rho C_p$</td>
</tr>
<tr>
<td>$\chi_{c1}$</td>
<td>$8.3 \times 10^{-2}$ m$^2$ s$^{-1}$</td>
<td>Thermal diffusivity</td>
<td>Upper crust</td>
</tr>
<tr>
<td>$\chi_{c2}$</td>
<td>$6.7 \times 10^{-2}$ m$^2$ s$^{-1}$</td>
<td>Thermal diffusivity</td>
<td>Lower crust</td>
</tr>
<tr>
<td>$\chi_m$</td>
<td>$8.75 \times 10^{-1}$ m$^2$ s$^{-1}$</td>
<td>Thermal diffusivity</td>
<td>Mantle</td>
</tr>
<tr>
<td>$k_{c1}$</td>
<td>2.5 W m$^{-1}$ K$^{-1}$</td>
<td>Thermal conductivity</td>
<td>Upper crust</td>
</tr>
<tr>
<td>$k_{c2}$</td>
<td>2 W m$^{-1}$ K$^{-1}$</td>
<td>Thermal conductivity</td>
<td>Lower crust</td>
</tr>
<tr>
<td>$k_m$</td>
<td>3.5 W m$^{-1}$ K$^{-1}$</td>
<td>Thermal conductivity</td>
<td>Mantle</td>
</tr>
<tr>
<td>$h_r$</td>
<td>10 km</td>
<td>Decay scale of radiogenic heat production</td>
<td>Upper crust</td>
</tr>
<tr>
<td>$H_s$</td>
<td>$9.5 \times 10^{-10}$ W kg$^{-1}$</td>
<td>Surface radiogenic heat production rate per unit mass</td>
<td>Upper crust</td>
</tr>
<tr>
<td>$H_{2e}C_v^{c2-1}$</td>
<td>$1.7 \times 10^{-15}$ K s$^{-1}$</td>
<td>Radiogenic heat</td>
<td>Lower crust</td>
</tr>
<tr>
<td>$\rho C_p$</td>
<td>J (m$^3$ K)$^{-1}$</td>
<td>Density × specific heat</td>
<td></td>
</tr>
</tbody>
</table>

(continued on next page)
the Eq. (4c), which refers to flat deposition in the downlands (see previous section).

4. Structure and rheology of model lithosphere

4.1. Rheology

We consider same lithological and mechanical rheological layering of the brittle–elastic–ductile lithosphere as in (Toussaint et al., 2004b) (Appendix A, Table 1). For the model demonstrated here, four lithological layers were defined: the upper crust, the lower crust, the mantle and asthenosphere (Fig. 2A). Each layer has specific properties (rheology, density, mechanical and thermal constants) that are given in Table 1. We assume no compositional changes due to deformation or cooling. The lithological boundary between the upper and lower crust is initially found at depth of 20 km. The bottom of the mantle lithosphere is limited by the 1330 °C isotherm at a depth of about 250 km. We use the same rheology for the lithospheric mantle and the asthenosphere. In the model, the lower plate has a normal crustal thickness of 37 km. The crustal of the upper plate is 10 km thicker, which insures its positive buoyancy.

The elastic, plastic and ductile constitutive laws form the yield stress envelope (YSE, Fig. 2B). Most of the upper crust remains elastic or brittle–elastic (depth interval between approx. 5 and 15–20 km). The lower crust may become ductile below 20 km. Depending on the geotherm and strain rates, first 30 to 70 km of the mantle lithosphere remains effectively elastic. The lithosphere becomes weaker when submitted to increasing horizontal forces or flexural stresses and when the crust gets thicker.

4.2. Thermal model

A thermal model is equivalent to that used of (Toussaint et al., 2004b). The initial geotherm is computed according to a half-space heat transfer model (for details see Burov et al., 1993; Burov and Diament, 1995):

$$\dot{T} + uT' + \nu T' = H_d + H_t + \nu \Omega$$

(5)

where primes mean spatial differentiation by respective coordinate. The thermal diffusivity parameter $\chi_t$ equals
to $\chi_{c1}$, $\chi_{c2}$, $\Omega_m$ depending on the lithological layer (see Table 1 for all parameters). $H_r = \chi_{c1} k_{c1} \rho_c H_c \exp(-y h_r^{-1})$ is the radiogenic heat. $H_r$ equals to constant heat generation $H_{c2} C_{c2}^{-1}$ in the lower crust and to zero in the mantle. $H_d$ refers to heat generation due to mechanical dissipation (e.g. frictional heating). The adiabatic temperature gradient in the asthenosphere, $\Omega_a$, is 0.3 °C/km (Turcotte and Schubert, 1982).

The boundary and initial conditions are: $T(x, 0, t_a) = 0$ °C (temperature at the upper surface = const at time $t_a$, where $t_a$ is the thermal age); $T(x, a, t) = T_m = 1350$ °C ($a \approx 250$ km is the depth to the thermal bottom, or thermal thickness of the lithosphere); $T(x, y, 0) = T_m$ (homogeneous temperature distribution at the beginning). $t_a$ is defined as the age of the last large-scale thermal event determined from geological data. We initially impose an adiabatic temperature gradient in the interval between the base of the lithosphere ($a$) and the bottom of the model ($y = 600$ km). The temperature at depth of 600 km is set to 2000 °C.

5. Implementation of coupled model

We use a fully coupled (mechanical behaviour–surface processes–heat transport) numerical models, which combine brittle–elastic–ductile rheology and account for large strains, fault localization and erosion/sedimentation mechanisms (Fig. 3).

We have extended the Para(o)voz code (Poliakov et al., 1993, Appendix A) based on the FLAC (Fast Lagrangian Analysis of Continua) algorithm (Cundall, 1989). This “2.5 D” explicit time-marching, large-strain Lagrangian algorithm locally solves Newtonian equations of motion in continuum mechanics approximation and updates them in large-strain mode. The solution of these equations is coupled with those of constitutive and heat-transfer equations. Parovoz v9 is thus a fully thermally coupled code that also handles explicit elastic–ductile–plastic rheologies, free-surface boundary condition, full metamorphic changes, and surface processes (erosion and sedimentation). The Lagrangian numerical mesh, which periodically becomes distorted and thus needs remeshing, is doubled by a denser passive marker grid allowing to interpolate grid values, specifically stresses, with minimal losses during remeshing.

We test continental collision assuming same initial scenario of (Toussaint et al., 2004b, Fig. 3), in which (1) the rapidly subducting oceanic slab first entrains a very small part of a cold continental “slab” (we do not impose continental subduction at the beginning), and (2) the initial convergence rate equals to or is smaller than the rate of the preceding oceanic subduction.

Two-sided initial closing rate, $u$, of $2 \times 3$ cm/yr during the first 5–10 My and coefficient of erosion $k$ of 3000 m$^2$/yr were used following (Toussaint et al., 2004b) for the India–Asia collision test. In this paper,
Fig. 4 (continued).

1. First 0 to 5 Myr: similar initial development

continental subduction and formation a small accretion prism mainly of lower crustal material

2. After 5 Myr, significant differences appear

whole-scale subduction in all cases

topographie de 20km

S=0.8
7Ma, 440km

S=1.0
7.5Ma, 470km

exhumation l. crust sediments

S=1.1
7.3Ma, 460km

S=1.12
10.46Ma, 660km

3. After 8 Myr, evolution scenarios become different

experiment ends up before 8 Myr, as whole-scale buckling with excessively high unrealistic topography develops

topography becomes too high due to buckling and cannot be compensated by erosion

after 500 km of subduction, plate coupling occurs and the system enters into pure shear collision mode with low topography
Fig. 4. A. Coupled numerical models of India–Eurasia type of collision as a function of the coefficient of erosion: \( k = 3000 \text{ m}^2/\text{yr} \) (left) and \( k = 6000 \text{ m}^2/\text{yr} \) (right), assuming horizontal convergence rate \( u \) of \( 3 \times 2 \text{ cm/yr} \). These experiments were performed in collaboration with G. Toussaint using a numerical setup (Fig. 3) identical to Toussaint et al. (2004b). The numerical method is identical to that of Burov et al. (2001), Toussaint et al. (2004a,b). Sub-vertical stripes associated with little arrows point to the position of the passive marker initially “drawn” across the middle of the foreland basin (Fig. 3). Displacement of this marker indicates the amount of subduction. \( \Delta x \) is amount of shortening. Different brittle–elastic–ductile rheologies are used for sediment, upper crust, lower crust, mantle lithosphere and the asthenosphere (Table 1b). B. Summary of the experimental results for different values of the coefficient of erosion, \( k \) (in m\(^2/\text{yr}\)), and for a fixed convergence rate \( u \) of \( 3 \times 2 \text{ cm/yr} \) (see caption to Fig. 4A for details). C. Zoom to the central part of Fig. 4B also showing distribution of the effective shear stress and effective “viscosity” (ratio of shear stress to strain rate) for the experiment with \( k = 3000 \text{ m}^2/\text{yr} \) for \( \Delta x = 460 \text{ km} \). Note that in case of \( k = 3000 \text{ m}^2/\text{yr} \), the amount of subduction (\( S \) number) is highest, and a larger part of sediment and upper crust is buried in the subduction channel, compared to the case of \( k = 500 \text{ m}^2/\text{yr} \) and \( k = 6000 \text{ m}^2/\text{yr} \). Orange — upper crust, yellow — lower crust, blue — mantle lithosphere, gray — asthenosphere, black — sediment (upper part) or HP metamorphic rock (eclogite). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
we greatly extended the range of convergence rates by testing the values as low as $2 \times 1$ cm/yr, and by varying $k$ from 50 to 11,000 m$^2$/yr. The rate ($2 \times 3$ cm/yr) chosen for the reference India–Asia collision test is relatively high, compared to the values for some other thrust belts. However, it is smaller than the average historical convergence rate between India and Asia ($2 \times 4$ to $2 \times 5$ cm/yr) during the first 10 m.y. (Patriat and Achache, 1984). The model and the model setup (Fig. 3, bottom) are identical to those described in detail in Toussaint et al. (2004b). For this reason, we send the interested reader to this study (see also Appendix A).

We use commonly inferred crustal structure and rheology parameters derived from rock mechanics (see previous section, Table 1; Burov et al., 2001). The thermo-mechanical part of the model is coupled with surface process model Eqs. (4a) (4b) (4c). On each type step the geometry of the free surface is modified by erosion and deposition. The surface area affected by sediment deposition change its material properties according to those prescribed for sedimentary matter (Table 1). The initial geotherm is derived from the common half-space model (e.g., Parson and Sclater, 1977) as discussed in section “Thermal model” and in Appendix A.

The universal controlling variable parameter of all continental experiments is the initial geotherm (Fig. 2B), or thermotectonic age (Turcotte and Schubert, 1982), identified with the Moho temperature $T_{\text{moho}}$. The geotherm or age define major mechanical properties of the system, e.g., the rheological strength profile (Fig. 2B). The second major parameter is the composition of the lower crust, which, together with the geotherm, controls the degree of crust–mantle coupling. In this study, we do not vary crustal and thermal structure but adopt same parameters as (Toussain et al., 2004a,b, Table 1).

The experiments of this study present a logical extension of the India–Asia collision model by (Toussain et al., 2004a,b). Indeed, the way the Central Asia has absorbed indentation of India may well reflect the sensitivity of the tectonic deformation to surface processes. Present kinematics of active deformation in Central Asia corroborates a highly heterogeneous distribution of strain. The 5 cm/yr convergence between India and stable Eurasia is absorbed by lateral extrusion of Tibet and crustal thickening, with crustal thickening accounting for about 3 cm/yr of shortening. About 2 cm/yr would be absorbed in the Himalayas and 1 cm/yr in the Tien Shan. The indentation of India into Eurasia has thus induced localized strain below two relatively narrow zones of active orogenic processes while minor deformation has been distributed elsewhere. Our point is that, as in our numerical experiments, surface processes might be partly responsible for this highly heterogeneous distribution of deformation that has been maintained over several millions or tens of millions years. First active thrusting along the Himalaya and in the Tien Shan range to the north may have been sustained during most of the Cenozoic period thanks to continuous erosion. Second, the broad zone of thickened crust in Central Asia has resulted in part from the redistribution of the sediments eroded from the localized growing reliefs. While the Indian upper crust would have been delaminated and brought to the surface by erosion, the Indian lower crust would have flowed below Tibet. Surface processes might therefore have facilitated down-thrusting of Indian lower crust below Tibet.

However, in this paper, we do not attempt to convincingly reproduce the India–Asia collision. Our primary goal is to study dynamic interactions between the surface and subsurface processes in a rapid convergence zone that has accommodated an important amount of shortening (600–700 km, according to DeCelles et al., 2002). The term “India–Asia” collision is thus used primarily because the parameters of our model are compatible with the parameters characterizing the early stages of this collision in a number of important points:

1. The convergence rate (e.g., Toussaint et al., 2004b).
2. The Rheological profile used for the lower plates yields Effective Elastic Thickness ($\sim 65–80$ km) that matches the observed one (e.g., Watts and Burov, 2003; Burov and Watts, 2006).
3. Thermal state of the lower plate (thermotectonic age $>400–800$ Ma, approximately $=\text{thermotectonic age of the Indian Plate (e.g., Watts and Burov, 2003)}$).
4. Initial crustal thickness $=37$ km (approximately the same as for the Indian plate (e.g., Watts and Burov, 2003)).
5. Rheological structure (Watts and Burov, 2003) likely fits the structure of the Indian lithosphere.
6. Total estimated amount of subduction (600–660 km, comparable with 600–700 km estimated by DeCelles et al. (2002)).

6. Experiments and discussion

In all numerical experiments, there is no initially predefined localized topography; the topography forms and evolves in time as a result of deformation and surface processes. We run series of tests for various coefficients
of erosion ($k=50–11,000 \text{ m}^2/\text{yr}$) for a fixed convergence rate of $2 \times 3 \text{ cm/yr}$. We then test the model sensitivity to variations in the convergence rate.

The numerical experiments (Figs. 4–6) confirm the initial idea that surface processes result in dynamic tectonically-coupled unloading of the lithosphere below.

---

**Topography evolution**

**Fig. 5. Smoothed surface profiles generated in the experiments of Fig. 4B (India–Eurasia type of collision).** $k$ is coefficient of erosion. Topography is unrealistically high for small $k=50 \text{ m}^2/\text{yr}$, and too low for high $k>1000 \text{ m}^2/\text{yr}$; the range $500 \text{ m}^2/\text{yr} < k < 3000 \text{ m}^2/\text{yr}$ corresponds to localized mountain growth. Note complex behaviour of the topography in case of $k=3000 \text{ m}^2/\text{yr}$: the mountain range and the major thrust tend to migrate towards the subducting plate. Small arrows show the position of the passive marker.
Fig. 6. Summary of the numerical experiments testing model sensitivity to erosion coefficient, $k$, and the convergence rate, $u$. $S$ is “subduction number”, that is the ratio of the subduction length, $s$, (total displacement of the passive marker) to the amount of shortening ($\Delta x$). $S = 1$ refers to “full subduction”, i.e. to the case where all shortening was accommodated through subduction. $S > 1$ indicates cases when the slab-pull or other gravity instabilities stretch the subducting plate so that the subduction length exceeds the amount of shortening. In particular, $S > 1$ corresponds to the initial stages of slab break-off. Numbers below the plate correspond to the attained subduction length.
the thrust belt, whereas the deposition of the eroded matter in the foreland basin results in additional subsidence. Fig. 4A shows experiments for \( k = 3000 \text{ m}^2/\text{yr} \) (case considered in Toussaint et al., 2004b) and \( k = 6000 \text{ m}^2/\text{yr} \). It can be seen that even if subduction develops in both cases, the amount of subduction (\( S \) number) is 20% higher in the case \( k = 3000 \text{ m}^2/\text{yr} \) than in the reference case \( k = 6000 \text{ m}^2/\text{yr} \). The amount of subducted crustal and sedimentary material (black color in Fig. 4A), is also significantly higher in case of \( k = 3000 \text{ m}^2/\text{yr} \), which practically means higher amount of material that reaches HP–HT metamorphic conditions. Moreover, in the case of \( k = 6000 \text{ m}^2/\text{yr} \), subduction slows down when the total amount of shortening reaches \( \Delta x = 400 \text{ km} \), and stalls after \( \Delta x = 580 \text{ km} \), whereas it continues in case \( k = 3000 \text{ m}^2/\text{yr} \) until \( \Delta x \) reaches 660 km (we stop experiments when the slab touches one of the boundaries of the model). Surprisingly, the size of the accretion prism is also much larger for \( k = 3000 \text{ m}^2/\text{yr} \) than for \( k = 6000 \text{ m}^2/\text{yr} \), so that the subduction front progrades towards the subducting plate, just as in the case of the major thrust fault in Himalaya, which was initially active far to the north of the current main front thrust. These differences between the two cases may be explained by larger spatial range of distribution of sedimentary matter in case of higher erosion coefficient, which results in smaller deposition in the foreland basin. The foreland basin remains relatively under-loaded and exerts less forcing on the down-thrusting plate.

Toussaint et al. (2004b) tested the possibility of subduction of the Indian plate beneath the Himalaya and Tibet at early stages of collision (first 15 My). This study used by default “stable” values of the coefficient of erosion \((3000\pm1000 \text{ m}^2/\text{yr})\) derived from predictions of the semi-analytical model of (Avouac and Burov, 1996). Toussaint et al. (2004b) thus did not test the sensitivities of the collision processes to variations in the erosion/sedimentation rate. In the new experiments of this study we fill the gap by testing this model for a large range of \( k \), from 50 \( \text{ m}^2/\text{yr} \) to 11000 \( \text{ m}^2/\text{yr} \) (Fig. 4B, C). These experiments show that, similarly to the cases of very high \( k \), very small values of \( k \) also lead to under-loading of the forelands and, more importantly, to insufficient removal of normal charges associated with the growing mountain topography. This leads to increase of normal stress exerted on the hanging wall of the thrust zone and thus to increase of friction on the thrust fault. In case of low \( k \), the amount of sediment buried to metamorphic depths is small, and exhumation of lower crustal material is not observed (Fig. 4C). Conversely, for erosion coefficients between 500 and 6000 \( \text{ m}^2/\text{yr} \), strong feedback between tectonic and surface processes is rapidly established and regulates the processes of mountain building during a long period of time (in the experiments \( >15 \text{ m.y.} \)) corresponding to a very large amount of shortening (up to 660 km). As can be seen, the amount of subduction of the upper crust and sedimentary material is highest for \( k = 3000 \text{ m}^2/\text{yr} \) (Fig. 4C). The exhumation of the lower crustal material is also observed for \( k = 3000 \text{ m}^2/\text{yr} \) but not for smaller \( k = 500 \text{ m}^2/\text{yr} \) nor for higher \( k = 6000 \text{ m}^2/\text{yr} \). Within the range \( 500 \text{ m}^2/\text{yr} \leq k \leq 6000 \text{ m}^2/\text{yr} \), the erosion–sedimentation prevents the mountain range from reaching gravitationally unstable geometries, and reduces the resistance of the major fault zone. These experiments demonstrate that the feedback between surface and tectonic processes may allow for development of long lasting continental subduction. The experiments show that once the major thrust zone is initialized, the tectonically balanced surface processes help to keep major thrust working. In the absence of a sufficient feedback between surface and subsurface processes, the major thrust fault is soon blocked, the upper plate couples with the lower plate, and the system evolution switches from simple shear subduction to pure shear collision or folding. However, topography cannot infinitely grow even in the strong “feedback” mode: as soon as the range grows to some critical size, it cannot be supported anymore by the underlying lithosphere, due to the limited strength of the constituting rocks, and ends up by gravitational collapse. The load of over-grown topography increases friction on the major thrust and results in narrowing and closure of the subduction channel. As a consequence, mechanical coupling between the lower and upper plate increases, which promotes pure shear thickening or large-scale plate buckling instead of subduction. On the other hand, extra-rapid removal of topography in case of \( k > 6000 \text{ m}^2/\text{yr} \) prevents formation of crustal roots, whereas the sedimentary matter is distributed over larger distances than for smaller \( k \) resulting in lesser volume of sediment deposited in the foreland basin. This all finally results in coupling between the upper and lower plate. The experiments (Figs. 4B and 5) demonstrate that, depending on the intensity of the surface processes, horizontal compression of continental lithosphere can lead either to strain localization below a growing range and continental subduction, or to distributed thickening or buckling/folding. These experiments indicate that subduction does not develop and homogeneous thickening or folding occurs when erosion is either too strong \((k > 6000 \text{ m}^2/\text{yr})\) or too weak \((k < 500 \text{ m}^2/\text{yr})\). In case of a dynamic balance between surface and subsurface
processes ($k=500–3000 \text{ m}^2/\text{yr}$, close to the predictions of the semi-analytical model, Fig. 6), erosion/sedimentation resulted in long-term localization of the major thrust fault that keeps working during 10 My. In the same time, in the experiments with $k=500–1000 \text{ m}^2/\text{yr}$ (moderate feedback between surface and subsurface processes), the major thrust fault and topography were almost stationary (Figs. 4B and 5). In case of a stronger feedback ($k=2000–5000 \text{ m}^2/\text{yr}$) the range and the thrust fault migrated horizontally in the direction of the lower plate (“India”). This basically happened when both the mountain range and the foreland basin reached some critical size. In this case, the “initial” range and major thrust fault were abandoned after about 500 km of subduction, and a new thrust fault, foreland basin and range were formed “to the south” (i.e. in the direction towards the subducting plate) of the initial location.

In the experiments, we used only very weak cohesion softening (stress drop $<20$ MPa, see Appendix A), which is taken into account because it represents a real property of most rocks. Yet, this softening affects only the first $1–2$ km of the upper crust (confining pressures less than 50 MPa). The numerical experiments thus confirm the idea that intercontinental orogenies could arise from coupling between surface/climatic and tectonic processes, and that orogeny–subduction systems may develop without specific help of other sources of strong large-scale strain localization. Given the differences in the problem setting, the results of the numerical experiments are in good agreement with the semi-analytical predictions of Avouac and Burov (1996) that predict mountain growth for $k$ on the order of $3000–10000 \text{ m}^2/\text{yr}$ for strain rates on the order of $0.5 \times 10^{-16} \text{ s}^{-1} - 10^{-15} \text{ s}^{-1}$. The numerical experiments, however, predict somewhat smaller values of $k$ than the semi-analytical experiments. This can be explained by the difference in the convergence mode attained in the numerical experiments (mainly simple shear subduction) and in the analytical models (pure shear). For the same convergence rate, subduction resulted in smaller tectonic uplift rates than pure shear collision. Consequently, “stable” erosion rates and $k$ values are smaller for subduction than for collision. For the fast convergence rate assumed for these experiments ($2 \times 3 \text{ cm/yr}$), stable subduction was observed in the range $50 < k < 6000 \text{ m}^2/\text{yr}$ for total amounts of shortening $\Delta x < 500$ km. For higher amounts of shortening ($500 \text{ km} < \Delta x < 660$ km), subduction continued only in the range of $k$ $500 < k < 4000 \text{ m}^2/\text{yr}$.

Finally, we have conducted a series of experiments to test model sensitivity both to the variations of the shortening rate $u$ and of the erosion coefficient, $k$. Summary of these experiments are shown in Figs. 6 and 7, as function of shortening rate $u$, $k$ and $S$ (“subduction number”). For all shortening rates, subduction does not develop for $k > 10000 \text{ m}^2/\text{yr}$ and $k < 50 \text{ m}^2/\text{yr}$. It also does not develop in the domain $k > 6000 \text{ m}^2/\text{yr}$ and $u < 2 \times 2 \text{ cm/yr}$ and $k > 500 \text{ m}^2/\text{yr}$ and $u < 2 \times 1.75 \text{ cm/yr}$. For $500 \text{ m}^2/\text{yr} < k < 3000 \text{ m}^2/\text{yr}$, subduction develops in the entire range of tested convergence rates. These experiments confirm the results of the experiments described in the previous sections but span to a wider range of the convergence rates. In addition, Fig. 7 suggests an appearance of local maximum on the $S–k$ curves for convergence rates $>2 \times 1.75 \text{ cm/yr}$. This maximum corresponds to $k$ values on the order of $3000 \text{ m}^2/\text{yr}$. It can be explained by additional downward stretching of the slab due to R–T instability.

7. Conclusions

We investigated the interactions between the surface and deep subsurface processes in continental collision settings for a case of relatively fast convergence rate and strong subducting plate that corresponds to India–Asia collision. We conclude that not only topography but also deep evolution of continental collision zones is sensitive to surface processes and, consequently, to climate. In some cases, the surface processes may be a forcing factor not only of the orogenic evolution but also of the entire
collision style. In particular, we suggest that the development of continental subduction at early stages of collision may be conditioned, among other factors, by the presence of an efficient feedback between surface processes and tectonic deformation.

We have shown that for the particular settings of India–Asia collision, continental subduction may develop in case of strong coupling between surface and subsurface processes. The most “favourable” “subduction” range of mass diffusivity corresponds to 500 m²/yr<k<3000 m²/yr. This parameter range corresponds to subduction numbers S>0.8 and also includes the range that favours the development of stable topography in pure shear collision defined in (Avouac and Burov, 1996). Within this range, some peculiar features of orogenic style such as the geometry of the accretion prism, the amount of subduction of the upper crust and sediments, horizontal progression of the mountain range/thrust fault and the amount of exhumation of metamorphic facies may be quite different (e.g., Fig. 4B, C). We demonstrated that both, excess topography (small k) and under-filled flexural basins (small or very high k) increase the resistance of the major thrust zone to down-thrusting of the lower plate.

The results of this study do not apply to India–Asia collision only, because we tested a large range of erosion coefficients and convergence rates that cover the initial conditions for most collision zones. Actually, the only feature of our model, that may be regarded as specific of the India–Asia collision, is the assumption of a strong lithosphere for the subducting plate. However, it is not excluded that the initial integrated strength of the lithosphere that subducted, for example, below the Alpes (Europe), was higher than the present-day integrated strength estimated from the Equivalent Elastic Thickness. Indeed, slow convergence rates in the Alpine settings could result in progressive thermal weakening of the lithosphere during the post-subduction period.

We finally conclude that surface processes must be taken into account in the interpretation and modelling of long-term deformation of continental lithosphere. Conversely, the mechanical response of the lithosphere must be accounted for when large-scale topographic features are interpreted and modelled in terms of geomorphologic processes. The models of surface process are most realistic if treated in two dimensions in horizontal plane, while most of the current mechanical models are two dimensional in the vertical cross-section. Hence, at least for this reason, a next generation of 3D tectonically realistic thermo-mechanical models is needed to account for dynamic feedbacks between tectonic and surface processes. With that, new explanations of evolution of tectonically active systems and surface topography can be provided.

Acknowledgements

We thank C. Faccenna and the anonymous reviewer for highly constructive comments.

Appendix A. Numerical algorithm for the full thermo-mechanical model

This mixed finite-element volume/finite difference code Parovoz (Poliakov et al., 1993) is based on the FLAC technique (Cundall, 1989). It solves simultaneously Newtonian dynamic equations of motion (A1), in a Lagrangian formulation, coupled with visco-elasto-plastic constitutive equations (A2), heat transport equations (A3) and state equation (A4) (see Appendix A (Burov and Guillou-Frottier, 1999; Burov et al., 2001; Le Pourhiet et al., 2004) for details concerning numerical implementation).

\[
\rho \frac{q^2}{\partial t^2} - \text{div}\sigma - \rho g = 0 \quad (A1)
\]

\[
\frac{D\sigma}{Dt} = F(\sigma, l, v, \nabla v, ... T...) \quad (A2)
\]

\[
\rho C_p \frac{\partial T}{\partial t} + v \nabla T - k_c \text{div}(\nabla T) - H - frac{\sigma_l \frac{\partial e_{II}}{\partial t}}{\partial t} = 0 \quad (A3)
\]

assuming adiabatic temperature dependency for density and Boussinesq approximation for thermal body forces:

\[
\rho = \rho_0 (1 - \alpha \Delta T) \quad (A4)
\]

Here I, v, \(\sigma\), g, and \(k_c\) are the respective terms for displacement, velocity, stress, acceleration due to body forces and thermal conductivity. The brackets in Eq. (A1) specify conditional use of the related term: in quasi-static mode, the inertia is dumped using inertial multiplier (Cundall, 1989). The terms \(t, \rho, C_p, T, H, \alpha, frac{\partial e_{II}}{\partial t}\) designate respectively time, density, specific heat, temperature, internal heat production, thermal expansion coefficient and shear heating term moderated by experimentally defined \(frac\) multiplier (\(frac\) was set to 0 in our experiments). The terms \(\partial e_{II}/\partial t\), \(D\sigma/Dt\), F are a time derivative, an objective (Jaumann) stress time derivative and a functional, respectively. In the Lagrangian framework, the incremental displacements are added to the grid coordinates allowing the
mesh to move and deform with the material. This enables solution of large-strain problems locally using small-strain formulation: on each time step the solution is obtained in local coordinates, which are then updated in the large strain mode. Volume/density changes due to phase transitions are accounted via application of equivalent stresses to affected material elements.

Solution of Eq. (A1) provides velocities at mesh points used for computation of element strains and of heat advection \( \nabla T \). These strains are used in Eq. (A2) to calculate element stresses, and the equivalent forces are used to compute velocities for the next time step.

All rheological terms are implemented explicitly. The rheology model is serial viscous–elastic–plastic (Table 1). The plastic term is given by explicit Mohr–Coulomb plasticity (non-associative with zero dilatency) assuming linear Navier–Coulomb criterion. We imply internal friction angle \( \phi \) of 30° and maximal cohesion \( S \) of 20 MPa, which fit best the experimental Byerlee’s law of rock failure (Byerlee, 1978):

\[
\tau = S + \sigma_n \tan \phi
\]  
(A5)

where \( \tau \) is the shear stress and \( \sigma_n \) is the normal stress. Linear cohesion softening is used for better localization of plastic deformation \( e_p(S(e_p)) = S_0 \min(0, 1 - e_p/e_{p0}) \) where \( e_{p0} = 0.01 \).

The ductile–viscous term is represented by power law with material parameters (Table 1) that correspond to the properties of four lithological layers: upper crust (quartz), middle-lower crust (quartz-diorite), mantle and asthenosphere (olivine):

\[
\mu_{\text{eff}} = \hat{e}^{(1-n)/n} (A^*)^{-1/n} \exp(H/nRT)
\]  
(A6)

where \( \hat{e} = \sqrt{I_2(\hat{\varepsilon}_{ij})} = \left( \frac{1}{2} \frac{\partial \hat{e}_{ij}}{\partial t} \right)^{1/2} \) is the effective strain rate and \( A^* = \frac{1}{2} A_3^{(n+1)/2} \) is the material constant, \( H \) is the activation enthalpy, \( R \) is the gas constant, and \( n \) is the power law exponent (Table 1b). The elastic parameters (Table 1a) correspond to commonly inferred values from Turcotte and Schubert (1982). At small differential stresses the rocks behave elastically. In terms of principal components, the relationship between the stress tensor, \( \sigma \), and the strain tensor, \( \varepsilon \), can be written (Table 1):

\[
\sigma_{ij} = 2 \mu \varepsilon_{ij} + \delta_{ij} \lambda \varepsilon_{ii}
\]  
(A7)

Surface processes are taken into account by diffusing Eq. (A8) the topographic elevation \( h \) of the free surface along \( x \) using conventional Culling erosion model (Culling, 1960, 1965) with a diffusion coefficient \( k \).

\[
\frac{\partial^2 h}{\partial t^2} = k \frac{\partial^2 h}{\partial x^2}
\]  
(A8)

This simple model is well suited to simulate fan deltas, which can be taken as a reasonably good analogue of typical foreland basin deposits. This model is not well adapted to model slope dependent long-range sedimentation, yet, it accounts for some most important properties of surface processes such as dependency of the erosion/sedimentation rate on the roughness of the relief (surface curvature).

Para(o)voz allows for large displacements and strains in particular owing to an automatic remeshing procedure, which is implemented each time the mesh becomes too distorted to produce accurate results. The remeshing criterion is imposed by a critical angle of grid elements. This angle is set to 10° to reduce frequency of remeshing and thus limits the associated numerical diffusion. The numerical diffusion was effectively constrained by implementation of the passive marker algorithm. This algorithm traces passively moving particles that are evenly distributed in the initial grid. This allows for accurate recovering of stress, phase and other parameter fields after each remeshing. Para(o)voz has been already tested on a number of geodynamical problems for subduction/collision context (Burov et al., 2001; Toussaint et al., 2004a,b).

References


