While the mechanisms of oceanic convergent processes – subduction - are well understood and successfully modeled, at least in major details[1,2], the mechanisms of continental convergence and the associated processes such as mountain building and high-pressure rock exhumation are still matter of lively discussions. Actually, in the continental realm, subduction, or simple shear shortening, presents one of at least four competing mechanisms of shortening (simple shear, pure shear, folding and Rayleigh-Taylor (RT) instabilities).

In continents; subduction indeed faces more physical obstacles than in the oceans[3,4]. First of all, the continental plates are positively buoyant owing to their light, thick crust. Second, mechanisms of upper-lower plate decoupling are there less clear than in the oceans (where lubrication of the subduction channel is warranted by serpentisation). Third, continental plates have a stratified multilayer rheology that results in complicated behaviors associated with possible mechanical detachments between the layers and lateral flow of the ductile material within the plates. Forth, continental convergence rates are slow (3–10 times slower than in the oceans). As the result, the Peclet numbers of the system are small, and the continental “slab” may be dominated by conductive, not by advective, heat transfer. In this case, hot temperatures propagate into the slab by thermal conduction from the asthenosphere as fast as the cold temperatures are advected with the sinking slab. Such a “slab” may undergo thermal readjustment leading to strong mechanical weakening well before it reaches any considerable depth. Hence, far-field push or pull forces will result in pure shear deformation or in de-blobbing due to RT instabilities rather than in simple shear subduction sensu stricto.

As mentioned, the alternative collision scenarios are related to accommodation of tectonic shortening by various mechanisms: (1) pure-shear thickening, (2) folding, and (3) gravitational (Raleigh-Taylor [RT]) instabilities in thickened, negatively buoyant lithosphere. The RT instabilities lead to sinking of subvertical sections of the lithosphere into the asthenosphere (“unstable subduction”). Superimposed scenarios are also quite possible: for instance, “megabuckles” created by lithospheric folding can localize and evolve into subduction-like zones or result in development of RT instabilities. RT instabilities can also occur at greater depth in slowly subducting slab. Despite these complexities, a number of geologic and geophysical observations point to the possibility of stable continental subduction. Among these observations the first rank is occupied by the data on exhumation of Ultra-High-Pressure metamorphic rocks (UHP) that represent former continental crustal units first buried at important depth and then brought back to the surface. These rocks record pressures as high as 3-5 GPa suggesting burial depths of 100-150 km. Subduction is by far the only plausible mechanism that can bring crustal rock to such great depth. Yet, while the exhumed UHP rocks can be considered as indicators of subduction processes, the mechanisms responsible for their return to the surface remain largely enigmatic. Indeed, it has been shown that accretion prism mechanisms that are well applicable to low and middle pressure exhumation, should not work at depths exceeding 40 km. Hence, specific multi-stage exhumation mechanisms are required to explain exhumation of UHP rocks. Consequently, to be matched with observations, successful model of continental collision should also reproduce characteristic P-T-t paths of the exhumed material. A number of obstacles to continental subduction may be bypassed at the initial phases of continental convergence. For example, the effect of the positive buoyancy of the lithosphere can be circumvented if low-density crust early separates from the mantle lithosphere or if it undergoes metamorphic changes leading to increase of its density. Geodynamic data also suggest that...
during the first million years of transition from fast ocean-continent subduction to slower continent-continent collision, convergence rates are considerably higher than at later stages. If true, then the continental “slab” may remain cold and strong enough to maintain subduction for a few million years. This minimal condition for subduction is usually defined by the Péclet number, $Pe$, equal to the ratio of the advection rate to the diffusion rate. $Pe$ must be $>20–30$ for subduction, which suggests that subduction is unlikely for convergence rates below $1–1.5$ cm/yr ($Pe < 20$), but is not impossible for convergence rates $>1.5–2.5$ cm/yr ($Pe = 30–70$). Many additional conditions must be also satisfied to enable stable subduction: other modes of deformation such as RT instabilities, folding, and pure-shear thickening should develop slower than in the simple-shear mode (subduction); the subduction channel should maintain low resistance to shear, which needs either specific mechanisms of softening, or continuous supply of weak material, fluids, shear heating or metamorphic reactions. The multitude of factors influencing continental collision justifies a numerical modeling approach (Fig. 1). Some of the existing numerical models are limited by the assumption of simplified viscous-plastic rheologies and by use of a fixed-displacement (or velocity) upper-boundary condition, instead of free-surface boundary condition. The use of fixed upper-boundary condition forces stable subduction, attenuates pure shear, cancels folding and does not allow for realistic topography evolution. Consequently, successful collision model[2-6] should be free of a number of constrains, otherwise typical for many existing geodynamic models. In particular it should (1) allow for all modes of small and large-strain deformation to occur; (2) account for “realistic” stratified viscous-elastic-plastic rheology and thermal evolution; (3) handle free surface boundary condition and account for surface processes; (4) incorporate phase changes (density and rheology); (5) eventually account for fluids, latent and shear heating and partial melting. Such models are expected to handle multi-physical processes and hence be able to reproduce (and be constrained by) various kinds of observations, from structural geology, topography, sedimentary and geodesy data, to petrology, gravity, thermal data and seismic tomography.

We here analyze major mechanisms of shortening of continental lithosphere (subduction, pure shear collision, folding, Rayleigh-Taylor instabilities), and, by proxy, of HP-UHP exhumation. We show examples of thermo-dynamically coupled thermo-mechanical numerical models that were successful in reproduction of crucial features of continental convergent processes. The “popular” numerical techniques used for collision models refer to finite-element, hybrid (finite difference-finite element + passive markers) Flac type techniques or staggered grid finite-differences + marker-in-cell techniques[2-6]. The most elaborated models account for free surface, viscous-elastic-plastic rheology and use marker techniques to trace P-T-t-z paths of metamorphic facies that are compared with petrology data. The phase changes are implemented via coupling thermo-mechanical codes to thermo-dynamic algorithms based on Gibbs free energy minimization (e.g. Perple-X by J.Conolly). The thermo-dynamic models also yield estimation of the amount of internal heating and of fluids that can be released or absorbed during phase changes. Fluid circulation is accounted in different ways, from simple computation of hydration fronts assuming vertical fluid migration at given rate to more elaborated porous flow models. The experiments suggest that continental subduction occurs in case of relatively strong mantle lithospheres (temperature at Moho depth, $T_m < 550^\circ$C) and high initial convergence rates ($> 1.5$-5 cm/yr). Depending on the lower-crustal rheology (strong or weak), either the whole (upper and lower) crust or only the lower crust can be involved in subduction. In case of weak metamorphic rheologies, phase changes improve chances for stable subduction. In general, exhumation of UHP-HP rocks to the surface is favored if crustal rheological profile is characterized by two internal ductile decollement levels (between the upper and lower or intermediate crust and the lower crust and mantle lithosphere). Pure shear collision is dominant when temperature at Moho depth, $T_m > 550^\circ$C or convergence rates are lower than 1.5-3 cm/yr. Large-scale folding is favored in case of $T_m = 500 – 650^\circ$C and is more effective in case of mechanical coupling between crust and mantle (e.g., strong diabase lower crust). Gravitational R-T instabilities overcome other mechanisms for very high values of $T_m (>800^\circ$C) and lead to the development of subvertical mantle or crustal “de-blobbing” (or “down-sagging”). The experiments show that active subduction channel is characterized by nearly lithostatic pressure conditions. Large-scale zones of tectonic overpressure may be built outside the channel but do not affect the exhumed rocks. Overpressure is also built inside
the channel during a short period of its closure when subduction stalls and the deformation mode switches from simple shear to pure shear or folding. We suggest that most continental orogenic belts could have started their formation from continental subduction. This evokes a multi-level mechanism of exhumation of UHP rocks to the surface, where deep units are either dragged up or rapidly ascend by Stokes mechanism, to the depths on the order of 40 km. The final phases of their ascend to the surface are controlled by slow accretionary prism mechanism and surface erosion.

The above results can be obtained only by “tuning” surface erosion in a way that it yields erosion rates comparable to the rates of rock uplift due to tectonic processes. It is indeed essential that during experiments, the model surface topography grows or remains stable having statistically same morphology and heights as in nature. The experiments show that any significant deviation from surface-subsurface balance results in smaller amount of subduction and disappearance of localized growth of topography. For example, for Himalayan collision, the optimal erosion coefficients allowing for total 800 km of subduction are on the order of 2000-3000 m²/y for shortening rates on the order of 5 – 6 cm/y and grid cell resolution of 5-10 km[8]. Increase of this coefficient by a factor of 2 results in 2 times smaller amount of subduction. Same effect would have a decrease of the coefficient of erosion by a factor of 3-4. Such a strong coupling between surface and tectonic processes is specifically important for rapid convergence (> 2 cm/y). For slowly converging regions (< 1.5 cm/y) the effect of surface processes on deep deformation is much smaller and almost negligible for the rates on the order of < 5 mm/yr. Hence it can be also concluded that continental convergence is largely controlled by the degree of coupling between the surface and subsurface processes, evoking strong feedbacks between climate-dependent erosion and tectonics.

References
[1] Sobolev S.V., A.Y. Babeyko (2005), What drives orogeny in the Andes?, Geology; v. 33; no. 8; p. 617-620; DOI: 10.1130/G21557AR.1
Figure 1: Typical continental convergence model setup and major controlling factors. $l_s$ is the amount of pure shear subduction to be compared with the total amount of shortening $\Delta x$. 
Figure 2: Examples of continental collision models and some of multi-physical data constraints such as structural, petrology and tomography data. After Yamato et al. [6] (bottom, slow early Alpine collision/subduction), Burov and Yamato [1] (middle, fast early Indian-type collision/subduction), Jolivet et al. [7] (top right), Gerya [2] and Faccenda et al. [4] (top, left).