The frequency response of dynamic friction: Enhanced rate-and-state models

CrossMark

A. Cabboi [a,*], T. Putelat [b], J. Woodhouse [a]

[a] Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK
[b] Department of Engineering Mathematics, University of Bristol, Queen's Building, University Walk, Bristol BS8 1TR, UK

ARTICLE INFO

ABSTRACT

The prediction and control of friction-induced vibration requires a sufficiently accurate constitutive law for dynamic friction at the sliding interface: for linearised stability analysis, this requirement takes the form of a frictional frequency response function. Systematic measurements of this frictional frequency response function are presented for small samples of nylon and polycarbonate sliding against a glass disc. Previous efforts to explain such measurements from a theoretical model have failed, but an enhanced rate-and-state model is presented which is shown to match the measurements remarkably well. The tested parameter space covers a range of normal forces (10–50 N), of sliding speeds (1–10 mm/s) and frequencies (100–2000 Hz). The key new ingredient in the model is the inclusion of contact stiffness to take into account elastic deformations near the interface. A systematic methodology is presented to discriminate among possible variants of the model, and then to identify the model parameter values.

1. Introduction

The fields of vibration engineering and earthquake mechanics both provide problems in dynamics in which interfacial friction plays a key role. To make progress with modelling the phenomena in either field requires an understanding of dynamic friction, and in particular requires a reliable constitutive law. A number of such "laws" have been proposed, based on a range of experimental approaches to characterising the friction force (or stress) at an interface exhibiting sliding or stick-slip motion (for general reviews, see for example Woodhouse et al., 2015; Sheng, 2008; Baumberger and Caroli, 2006; Marone, 1998). However, there is no consensus in the literature, especially when it comes to problems involving relatively high frequency vibration: vehicle brake squeal, for example, which is a friction-driven vibration typically occurring in the kilohertz range. Recently a novel measurement method for dynamic friction was described (Wang and Woodhouse, 2011), which extends into this kilohertz range. When the results were compared with existing models from the literature, none were found to agree (Woodhouse and Wang, 2011).

The new measurement method was motivated by the needs of applications like brake squeal, in which a key question is to determine when the state of steady sliding between brake liner and brake disc is stable, and when it is unstable to some kind of oscillatory disturbance. Such an instability is one route by which undesirable brake noise can arise. It is not the only possible route, but it is the one most extensively studied in the past because it can be approached via linearised analysis. In a

* Corresponding author.
  E-mail addresses: ac932@cam.ac.uk (A. Cabboi), jw12@cam.ac.uk (J. Woodhouse).

similar way, a route to study earthquake processes is also based on mapping out the frictional stability regime, since it has been widely recognised that stick-slip frictional instability, rather than fracture, plays a primary role in triggering earthquakes (Brace and Byerlee, 1966; Marone, 1998; Scholz, 1998; Bizzarri, 2014).

For any linearised stability analysis applied to mechanical systems under steady sliding conditions, it is easy to see what form the constitutive information about the frictional interface must take. In the early stage of a growing instability, the steady sliding speed is perturbed by a small oscillatory disturbance. If that disturbance is sinusoidal, and if linear theory is applicable, then the evoked variation in friction force must also be sinusoidal at the same frequency, with some amplitude and phase relative to the speed perturbation. If the sliding speed and tangential force are modulated around their steady-state values v_0 and F_0 according to

$$v \approx v_0 - v' e^{i\omega t} \quad \text{and} \quad F \approx F_0 + F' e^{i\omega t}, \tag{1}$$

it must be possible to write

$$F'(\omega) \approx \beta(\omega) v'(\omega) \tag{2}$$

where $\beta(\omega)$ is in general a complex number, which may vary with frequency: it is a kind of linearised frequency response function. Of course, β may also depend on other variables such as the sliding speed v_0 and the normal load N_0, as will be explored later in this study. The negative sign in Eq. (1) is chosen for consistency with earlier work (Woodhouse and Wang, 2011).

It is sometimes assumed that the friction force depends only on the instantaneous sliding speed, as in the "Stribeck model" (see for example Sheng, 2008). In that case β would be simply the slope of the force–velocity curve at the operating point v_0: it will be a real number, independent of frequency, for given values of v_0 and N_0. However, if the frictional force perturbation is influenced by other state variables then β is likely to be complex, with an amplitude and phase that may vary with frequency in a non-trivial and informative way, as has been shown in the preliminary measurements described by Wang and Woodhouse (2011). The impedance-like function $\beta(\omega)$ encapsulates information about friction needed by any linearised analysis to detect instability initiation of the steady sliding state.

It may not be the whole story, though. In Butlin and Woodhouse (2013) a first glimpse has been provided of how the prediction of friction-induced vibration might be improved by including similar frequency-response information relating to normal force and velocity fluctuations. A general linearised contact model would then involve a 2×2 matrix of which $\beta(\omega)$ is one entry. However, in the present work attention is restricted to $\beta(\omega)$ alone.

The recent published measurements of $\beta(\omega)$ by Wang and Woodhouse (2011) clearly revealed a complex, frequency-varying result: some further examples will be shown in Section 2, and the test apparatus will be described. The results showed puzzling features, and before getting into detail it is useful to summarise some general issues. The most striking of these concerns the role played by the normal load. Most published frictional models assume the Amontons–Coulomb "law" that for a dry sliding contact the friction force F_0 is proportional to the normal load N_0 so that $F_0 = \mu N_0$, in terms of the familiar idea of a coefficient of friction μ. In the new measurements it was found that the dynamic component of friction, captured in the frequency response function described above, showed a non-trivial dependence on N_0. Rather than a simple proportional relationship, the measured $\beta(\omega)$ showed different patterns of variation with frequency for different values of N_0. However, simultaneously with the dynamic measurements the mean value F_0 was also measured, and this *was* found to follow the Amontons–Coulomb proportionality relation.

The second striking qualitative feature observed in the earlier measurements concerned repeatability. A series of tests was run, in which the imposed mean sliding speed was varied. Starting from a relatively slow speed, successive runs used progressively higher speeds up to a maximum value, then the sequence was followed in reverse so that the final run used the same slow speed as the first. The entire sequence of tests took several hours. It was found that the frequency response function capturing the dynamic friction force was accurately reproduced between the first and last runs, at the same slow speed. However, the mean friction force, or equivalently the steady sliding coefficient of friction μ_{ss}, was quite different at the end. For one particular sequence of runs the friction force was shown by Wang and Woodhouse (2011) to increase monotonically, with the final value exceeded the initial value by more than a factor of 2. A possible explanation is that the mean friction force is sensitive to temperature, and of course the temperature of the apparatus will have gone up during the hours of testing with continual sliding. But then it is striking and unexpected to find that the dynamic component of friction, represented in $\beta(\omega)$, seems to be quite insensitive to this temperature rise.

In a real application such as the design of vehicle brakes, manufacturers may devote a lot of effort to measurements of μ_{ss}, to see how it varies with sliding speed, environmental conditions, wear state of the brake lining material and so on. They may describe such a measurement programme as a "full tribological characterisation". However, important as μ_{ss} may be in determining how well the brake will stop a moving vehicle, the results just mentioned suggest very strongly that if these brake manufacturers also want to predict and control the occurrence of squeal they will need to augment their testing programme with dynamic measurements along the general lines of those to be described here, since the dynamic behaviour is qualitatively different and is unlikely to be captured usefully by steady-sliding measurements alone.

The structure of the paper is as follows. Section 2 provides information about the experimental rig with which the frictional frequency response $\beta(\omega)$ is measured. The testing procedure is described in sufficient detail to give confidence in the measurements and their reproducibility, and also to provide enough information to allow independent checking by
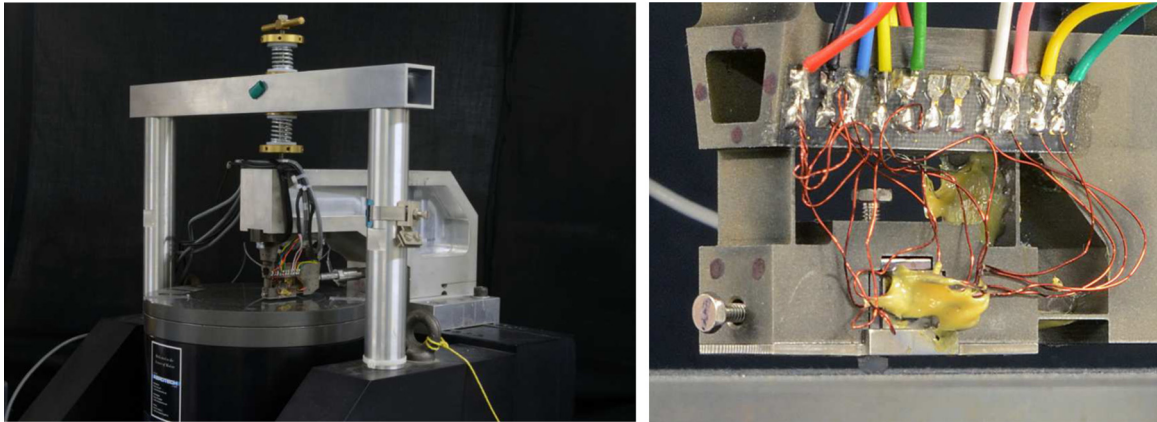
Fig. 1. The measurement rig: global view and dynamometer unit.

other researchers. In Section 3, theoretical material relating to friction modelling is presented. In particular, a range of alternative enhanced rate-and-state models is introduced, with special emphasis on the rôle and importance of the contact stiffness. Section 4 presents detailed measurements of the frictional frequency response for various sliding speeds and normal forces, and uses them for testing and discriminating these friction models. A systematic optimisation methodology is used to identify the parameter values of the various candidate models. Finally, Section 5 gives discussion and summarises the main conclusions, along with possible future research directions.

2. Experimental work

2.1. Description of the measurement rig

The test rig, based on a pin-on-disc configuration and extensively described in Wang and Woodhouse (2011), is shown in Fig. 1. Its key components are shown schematically in Fig. 2. The disc is arranged in a horizontal plane, carried by a direct-drive rotary stage. Brought into contact with this disc is a hemispherical pin sample of the material to be tested, attached to a dynamometer unit. The pin sample is glued to the bottom part of a small inner block that is inserted into a larger outer block, held by the dynamometer unit through three thin members. The dynamometer is in turn supported by an aluminium arm through two flexures to allow small movement in the direction parallel to sliding. To accommodate any small vertical movements from non-flatness of the disc and to control the normal force, a spring system connected on a gantry is screwed on the top of the supporting arm.

As already described in the introduction, the target measurement is a linearised frequency response function, so a piezoelectric actuator is used to provide a small and controllable dynamic perturbation v' on the steady sliding speed v_0 induced by the rotating disc. In order to excite and measure over a broad frequency band, the imposed fluctuation consists of band-limited random noise, whose maximum amplitude needs always to be much smaller than the actual rotation speed of the disc so that the assumption of linearised theory remains valid. As described earlier (Wang and Woodhouse, 2011), the validity of the linearisation assumption can be directly checked by performing a series of runs at different excitation amplitudes. For all measurements to be presented here, $\beta(\omega)$ showed excellent agreement in amplitude and phase within the excited frequency range across a wide range of excitation amplitudes, and also exhibited high values of the coherence function calculated as part of the measurement process.
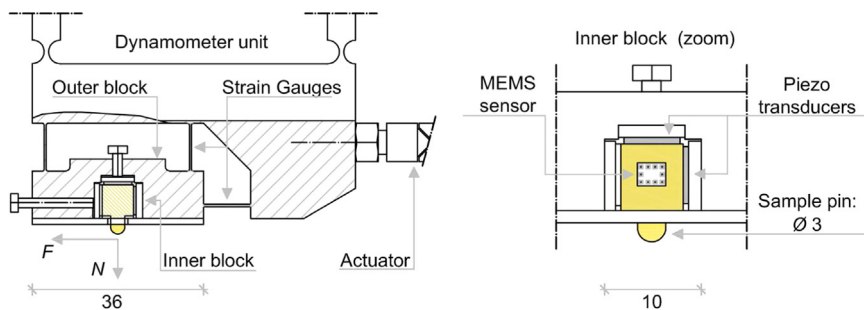

Fig. 2. The dynamometer unit, zoom of the inner block and installed sensor package. All dimensions are in mm. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

During a test run, forces and velocities in the tangential (frictional) and normal direction are measured, F and N in Fig. 2. In order to do this, the dynamometer unit is provided with a package of sensors (see Fig. 2 for accompanying diagram):

(a) Strain gauges are mounted on two of the thin members that connect the outer block to the dynamometer, which allow the DC and AC components of the normal and tangential forces to be measured.
(b) Piezoelectric force transducers are installed on the top and on the side of the inner block, since the AC components provided by the strain gauges exhibit poor noise performance. The inner block is held by two screws, which provide pre-load to these two piezoelectric force gauges. The top of the inner block is dynamically isolated by a small piece of synthetic rubber sheet so that the dynamic friction force is transferred effectively to the sensing element without complications from friction at the base of the block, given that the applied normal load has to be carried through this block. In the tangential direction the inner block, piezoelectric element and outer block are all in direct contact, so that no relative motion is possible in the frequency range relevant to the measurement.
(c) A 3-axis MEMS sensor (Willow KXTC-2050) is glued to the side of the inner block, to monitor the motion of the pin (yellow area in Fig. 2).

Standard signal processing methods (see for example McConnel and Varoto, 2008) are then used to construct the required frequency response function: the averaged transfer function computed using the dynamic sliding velocity as input and the measured force fluctuation as output (see Section 2.2.1). Concurrently, during the dynamic testing procedure, the mean DC values of the forces measured by the strain gauges are also calculated over the entire duration of each run.

Before performing any measurements the installed sensors needed to be subjected to several stages of calibration. Some of these have already been described in Wang and Woodhouse (2011), but all are briefly mentioned below for completeness:

1. Synchronisation of data acquisition. The DAQ device used (NI DAQ 6023E) is based on a multiplexed architecture, so there is an inter-channel time delay made up of a settling time (De Silva, 2007) plus the time needed by the A/D converter to perform a conversion. In order to identify and compensate the time delay, several otherwise identical measurements were performed after interchanging the order of channels. For each test, transfer functions were estimated using the signal measured by the first channel as input and the quantities measured by the other channels as outputs. The required time delays can be identified by estimating the phase deviation between similar transfer functions obtained with a different channel layout. Once the time delay had been identified (5 μs), it could be compensated in subsequent measurement runs.
2. Static calibration and decoupling of the strain gauges and the piezoelectric force transducers. Because of the geometrical details of the dynamometer and the location of the strain gauges and the piezoelectric force sensors, the normal and tangential force measurements are not entirely uncoupled. For example, a tangential force will evoke a significant response in the strain gauge aligned along the vertical direction, and conversely a normal force causes a small response in the tangential gauge. These coupling terms need to be identified, to allow the decoupling of the orthogonal force components by using a linear transformation based on a 2×2 matrix of influence coefficients. The strain gauges were calibrated using the procedure described in Wang and Woodhouse (2011), while the calibration of the piezoelectric force transducers was performed by gently pulling loops of thin copper wire (connected to the bottom of the inner block) in the horizontal and vertical directions until they broke. By knowing the ultimate load of the wire and by reading the sign and voltage of each force jump due to the wire rupture, it is straightforward to estimate the 2×2 influence matrix of the piezoelectric sensors.
3. Calibration and compensation of the MEMS device. The circuit of the triaxial MEMS sensor includes first-order passive low-pass filters with a cut-off frequency of 50 Hz, which act approximately as an integrator in the time domain. Hence above 50 Hz, the MEMS device (whose actual measurement is acceleration) starts to provide an approximation to velocity as output data. However, the precise characteristics of the device and its filters need to be taken into account. Furthermore, MEMS devices exhibit a certain amount of cross-sensitivity between the axes due to manufacturing errors, so that the resulting FRFs can be significantly contaminated by coupling terms due to the non-orthogonality between axes (McConnel and Varoto, 2008). The response was checked by employing a laser-Doppler vibrometer, and estimating the transfer function using as input the velocity obtained by the laser vibrometer and as output the signal provided by the MEMS device in the same direction. Amplitude and phase corrections were found to be necessary to optimise the accuracy of the velocity measurement. In addition, in order to decouple the measurements along the two orthogonal axes a simple test rig was built, in which a symmetrical cantilever beam was used to provide vibration in a well-controlled direction. The MEMS device was attached to the end of the beam, in a manner that allowed it to be rotated. A miniature impulse hammer was then used to excite the cantilever, and measure the transfer functions along the two relevant axes of the MEMS device. In order to identify the cross-sensitivity between the two directions, the test was repeated after rotating the sensor by small angles. The goal was to determine the angle that provides a "null measurement" for each measurement axis, from which it is straightforward to construct a 2×2 rotation/decoupling matrix in order to compensate the effects induced by manufacturing errors.
4. Mass compensation of the inertial forces induced by the outer and inner block masses. The procedure is based on measuring force sensor outputs with the pin not in contact with the disc, as described in Wang and Woodhouse (2011). Since a new, although rather similar, dynamometer is used here, slightly different masses were found.
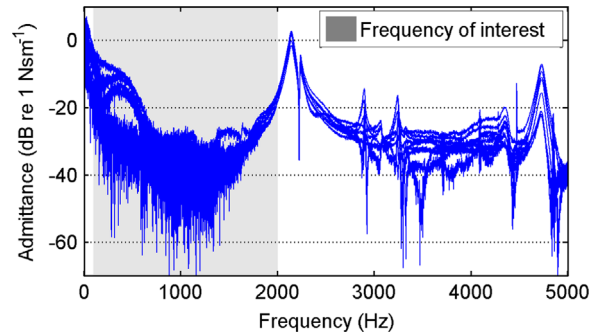
Fig. 3. Frequency response of the dynamometer block along the sliding direction. The main resonance peak indicates the need to tailor the amplitude of the induced random exitation for the dynamic friction measurements.

The final step necessary before dynamic measurements could safely be carried out was to take account of the first mode of the dynamometer unit involving significant motion along the sliding direction. Vibration measurements were carried out using a miniature impulse hammer to excite the front face of the dynamometer unit at nine different points. By using the dynamic response of the block along the tangential direction, the respective averaged transfer functions were estimated (Fig. 3). This admittance (mobility) plot shows that the first significant mode in the tangential direction appears at 2141 Hz, with a Q factor of 43. The modal amplitudes revealed a mode shape which is approximately a rigid body motion of the dynamometer block moving on its flexures in the tangential direction, against the stiffness of the piezo actuator.

By fitting the modal parameters of a single-degree-of-freedom oscillator to this resonance at 2.14 kHz, the excitation noise signal could be shaped using the inverse of the theoretical frequency response of this oscillator. This shaped noise is used, via a digital-to-analogue converter, to drive the amplifier that powers the piezo actuator. In this way an approximately flat spectrum of the induced velocity perturbation is produced, avoiding any dangerous resonant oscillation that might damage the rig. In principle this allows measurements to be made up to and beyond the first resonance frequency: repeatable results have been obtained up to 4 kHz. However, results at higher frequencies are rendered unclear by other rig resonances, and for the purposes of this paper results will only be shown up to 2 kHz.

2.2. Experimental measurements

2.2.1. Dynamic friction test procedure

The aim of the measurements to be described here was to check the results reported in the previous work (Wang and Woodhouse, 2011) using two different kinds of polymer—nylon and polycarbonate—and then to collect a more complete and organised set of data on these materials to allow systematic discrimination between different candidate theoretical models to be described in Section 3. In all cases the disc material is glass: a new glass disc was mounted on the rig for the purpose of these tests.

Before starting the description of the test procedure, it is worth underlining that the measurements done so far, and presented in the current paper, are mainly proofs of concept of the experimental technique and the subsequent model-fitting process. There is probably little direct technological interest in the friction of nylon or polycarbonate against glass, but the methodology presented here could be extended to the characterisation of any frictional interface. The information gained could then be used for a linearised stability analysis for friction-induced vibration with some hope of success, since the correct material information would be, for the first time, included in the modelling.

The initial step of the testing procedure is to make sure that the glass surface of the disc is clean and dry, and to balance the strain gauge bridges to ensure correct DC measurements. As already mentioned, the dynamic friction measurement relies on an averaged transfer function, estimated in the cases to be presented from 120 samples each of length 1 s. Six channels of data were collected, using a sampling frequency of 20 kHz: the velocity signals, and the forces from the piezo transducers and strain gauges; each physical parameter is measured along the tangential and normal direction. Hence, 5 averaged transfer functions, and 5 corresponding coherence functions, were estimated using the velocity along the frictional direction as input signal. The mean values of the friction force F_0 and normal force N_0, the ratio of which provides the steady-state coefficient of friction μ_{ss} for each test run, are also calculated.

For each material combination, several sequences of runs were performed using different velocities and normal forces. A typical sequence of tests consisted of fixing a desired value of normal load, and progressively changing the sliding speeds starting at 1 mm/s and going in steps up to 40 mm/s, and then back down again to 1 mm/s. Before starting a sequence of tests, the disc was run at 20 mm/s for about 30 min, which seemed to give a good compromise between adequate running-in of the contact conditions and avoiding lengthy measurements that might raise concerns about the wear of the pin sample. Once a measurement sequence started, any drift in the balance of the strain gauge bridges was monitored after every three runs of testing by recording 10 s of signal from the strain gauges with the pin lifted temporarily out of contact with the disc. An entire sequence involves about two hours of sliding contact.
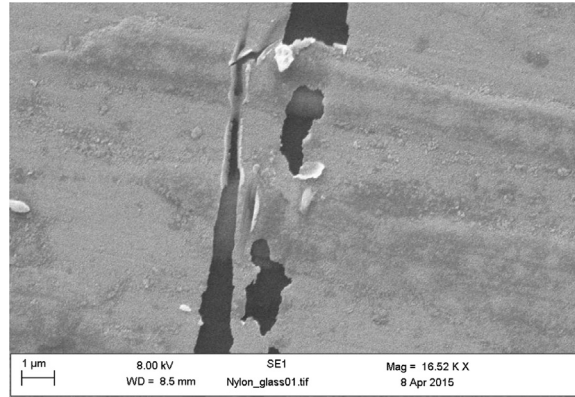
Fig. 4. SEM scan of polymer deposit along the sliding track. The obvious vertical scars result from a small scratch on the surface, which seems to have lifted the polymer transfer film away from the underlying glass.

It is worth anticipating the reassuring conclusion that the results obtained from these new tests are in good agreement in terms of phase and amplitude variation with those obtained a few years ago using a slightly different dynamometer head, with different sensors and different pin geometry (Wang and Woodhouse, 2011).

2.2.2. Sample results: nylon on glass

For the material combination of nylon pin against glass disc, a full data matrix of dynamic responses was collected using a range of sliding speeds and normal forces. The sliding speeds were accurately controlled, but the normal force was determined in each case by hand adjustment of the spring system so that precise pre-determined values were hard to achieve. The test values were roughly 10, 20, 30, 40 and 50 N, but the exact value was always determined as part of the measurement process.

In the light of results obtained in Wang and Woodhouse (2011), it is worth discussing first the estimated steady-state coefficient of friction as a function of sliding speed and normal force. This kind of measurement is very commonly performed in connection with industrial applications, and commercial testing machines are made for that purpose. Whether or not the variation with steady sliding speed is sufficient to predict dynamic variation of the friction force (as is claimed by the "Stribeck law", for example), unambiguous measurement of μ_{ss} itself can be very problematic under certain conditions. For example in Wang and Woodhouse (2011), some results were shown from a series of runs with progressively increasing and then decreasing sliding speed: these results clearly showed a rising friction force throughout the testing period, leading to unrepeatable results for the same sliding speed. The coefficient of friction rose by more than a factor of 2, from 0.2 up to 0.45. In order to shed some light on this behaviour, the old glass disc was removed and a small section was cut out of it. The piece of glass was then examined in a scanning electron microscope (SEM). A representative micrograph (Fig. 4) clearly highlights the presence of a polymeric transfer film on the glass surface: a scratch on the surface while cutting the glass appears to have lifted the film in places. Such a transfer film inevitably affects the tribological behaviour of the material combination. For example, transfer layer formation might easily lead to self-mated sliding (Blau, 2009) that produces a subsequent rise in the coefficient of friction due to the thermal effects caused by higher sliding speeds.

As already mentioned, for the measurements to be reported here a new glass disc was mounted on the rig, and a data matrix of dynamic responses was collected. Fig. 5a shows the measured μ_{ss} against sliding speed for different normal forces (with speed increasing and then decreasing through the same set of values for each force), and conversely the friction force
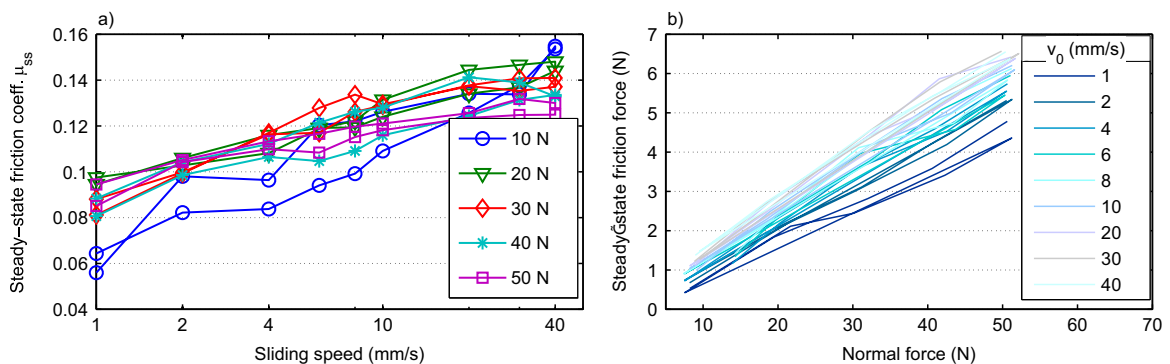

Fig. 5. Nylon on glass: (a) steady-state coefficient of friction, μ_{ss}, vs. sliding speed; (b) measured frictional force against normal force.
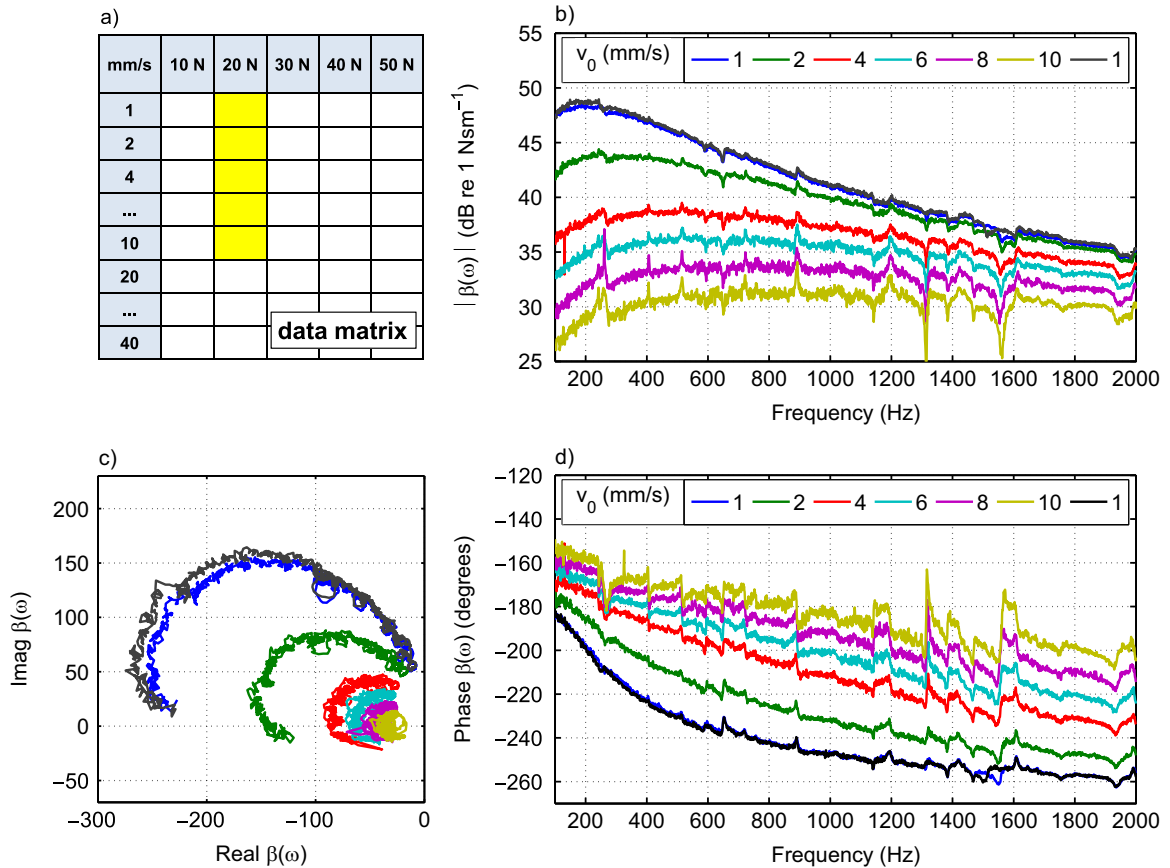
Fig. 6. Nylon on glass: typical examples of frictional frequency response function $\beta(\omega)$ for a fixed normal force $N=20$ N and different sliding speeds. In (b), (c) and (d) the repeatability of the measurement is highlighted by the close agreement of the blue and black lines. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

against the normal force plotted for different values of sliding speed (Fig. 5b). If compared with the results shown in Wang and Woodhouse (2011), the following observations can be made:

1. The friction coefficients, μ_{ss}, exhibit lower values and a more limited variation, ranging between 0.06 and 0.16.
2. Contrary to what was shown in Wang and Woodhouse (2011) no increase of μ_{ss} was observed for the runs performed by decreasing the disc speed from 40 mm/s to 1 mm/s.
3. The figure on the left suggests a logarithmic variation of μ_{ss} as a function of velocity.
4. The figure on the right highlights that Coulomb's law seems to be satisfied fairly well, despite slight scatter of the μ_{ss} values. One possible reason for this scatter is that the five run sequences were performed over five days (one per day). The precise test conditions may have varied slightly from one day to the next, for example because of humidity variations.

The first two comments are connected: the lower friction coefficients may well be caused by the absence of an effective polymeric transfer film on the new glass disc, which in turn reduced thermal effects that might trigger an increase of μ_{ss}, persisting even at lower disc speeds. This highlights the strong sensitivity of the μ_{ss} measurement to the testing conditions of the tribosystem, reinforcing the suggestion that the use of the steady-state friction curve for friction-induced vibration analysis is likely to be unreliable for the materials tested here. As already remarked, it was shown in Wang and Woodhouse (2011) and will be confirmed in the current work that for the material combinations tested, the dynamic quantity $\beta(\omega)$ seems to be much more consistent than μ_{ss}. This should be good news for linear stability analysis: the correct quantity to use in any such analysis is in some ways easier to measure than the traditional but incorrect quantity.

The response data for the nylon-glass combination was next used to estimate the frictional frequency response function for the matrix of different velocities and normal forces. Some examples of the resulting $\beta(\omega)$ are shown in Fig. 6 for a selection of speeds with a fixed normal force, and in Fig. 7 for a set of normal forces with a fixed sliding speed. It was reassuring to find that all these $\beta(\omega)$ measurements exhibited a coherence function close to unity over the tested frequency range. In Fig. 6 the magnitude (plotted on a dB scale), the phase and the Nyquist plot of $\beta(\omega)$ all clearly demonstrate the remarkable level of repeatability between the initial and final runs at 1 mm/s, especially considering the noise limits of this
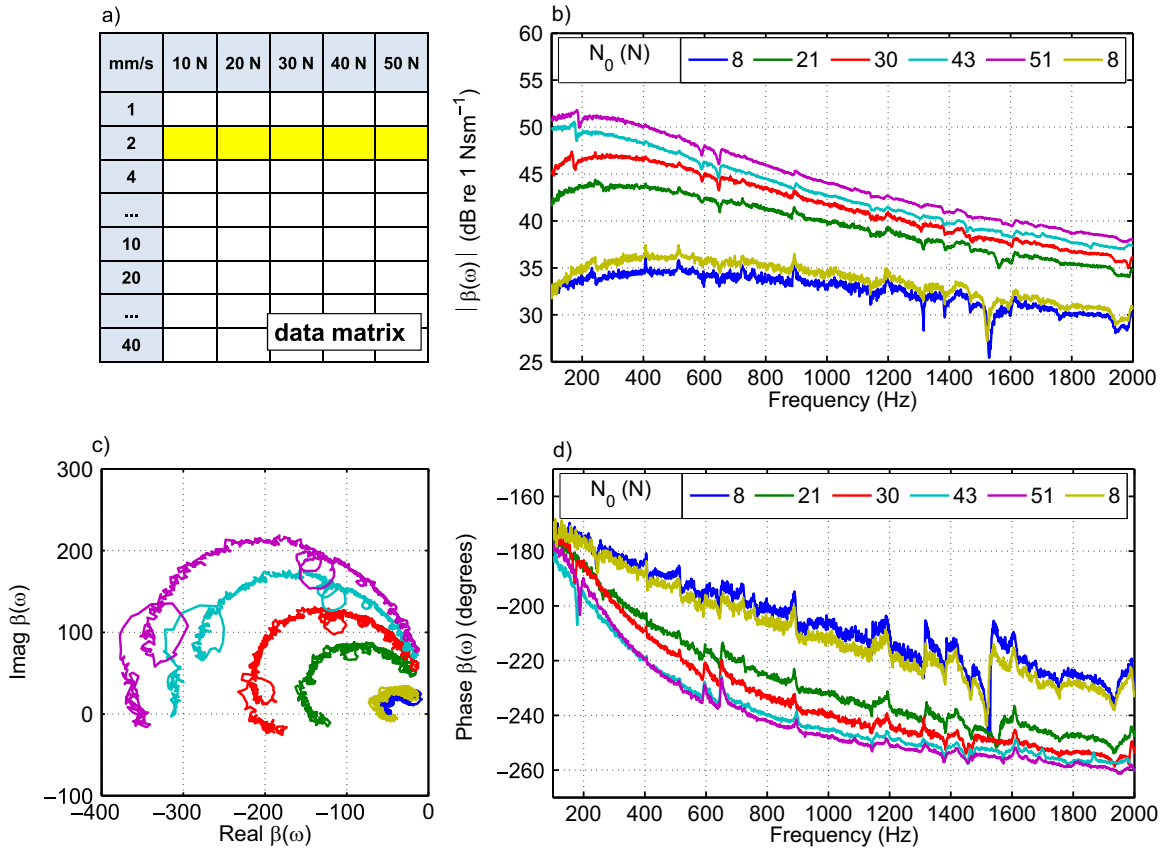
Fig. 7. Nylon on glass: typical examples of frictional frequency response function $\beta(\omega)$ for a fixed sliding speed $v_0 = 2$ mm/s and different normal forces. In (b), (c) and (d) the repeatability of $\beta(\omega)$ is highlighted by the close agreement of the two measurements at 8 N.

type of measurement. Many features of $\beta(\omega)$, such as the variation with frequency and sliding speed, confirm results shown earlier in Wang and Woodhouse (2011). Some narrow peaks are visible in the results: these are attributable to artefacts of the rig. These peaks, which have the character of antiresonances since $\beta(\omega)$ is an impedance-like quantity, are emphasised at higher sliding speeds when the magnitude of the results reduces.

The Nyquist plot in Fig. 6 highlights a further suggestive feature that was already mentioned in Woodhouse and Wang (2011): the frequency band under analysis is characterised by a roughly circular shape in the complex plane, which becomes clearer and less noisy as the sliding speed decreases. As the sliding speed increases the circular segments rotate in a counterclockwise direction. The low frequency end of the circles ends up in the lower half-plane with a negative imaginary part, a fact previously shown to be relevant when comparing with the predictions of certain theoretical models.

A further demonstration of repeatability of the $\beta(\omega)$ measurement can be seen in Fig. 7. In this case, by picking a sliding speed and choosing tests performed at different normal forces it is possible to observe how the measured dynamic frictional curve changes by varying the normal force. The set of normal forces is covered twice, because of the increasing/decreasing sliding speed runs. The results shown here for normal forces of 7.6 N and 8.2 N are very similar to each other. In contrast to the variation with sliding speed shown in Fig. 6, here an increase of normal force is seen to cause an increase of the magnitude of $\beta(\omega)$. Further discussion of the results shown in these figures is deferred until Section 4: once some candidate theoretical models have been introduced, the full range of $\beta(\omega)$ measurements from the data matrix will be used to discriminate between models, to estimate parameter values for them, and finally to judge whether any of the proposed models give a sufficiently close fit to the observed behaviour.

2.2.3. Sample results: polycarbonate on glass

To check that the behaviour seen in the measurements is not somehow special to the particular combination of nylon against glass, it is useful to show some results obtained with a different material combination. A pin sample of polycarbonate was fitted to the test rig, keeping the same glass disc and choosing a different sliding path to avoid any pre-existing transfer film of nylon. The intention was to perform a matrix of tests using the same set of sliding speeds and normal forces described before. However, by the time runs had been performed with the full range of sliding speeds for normal forces around 10 and 20 N, the polycarbonate pin started to release visible wear debris and the contact surface of the pin became larger. This evolution of the contact conditions caused significant changes in the dynamic response of the
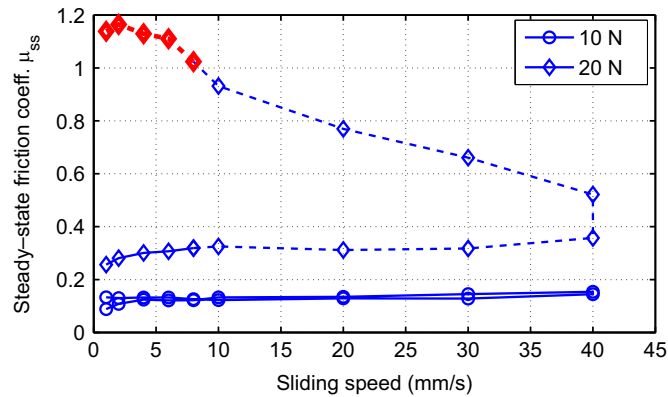
Fig. 8. Polycarbonate on glass: variation of the steady-state coefficient of friction, μ_{ss}, as sliding speed is first increased and then decreased, for two different values of normal force; the red portion indicated squeal initiation. The lack of repeatability of μ_{ss} at the higher normal force is probably due to the effects of temperature and wear of the polycarbonate pin, changing the surface properties. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

frictional interface, and led to self-excited squeal in the test rig. Because of that, no further meaningful tests were possible using the same pin and path on the disc. Fortunately, the limited range of data collected was already enough to show some important features in common with the nylon results.

Fig. 8 shows that the evolution of the contact conditions strongly affected the value of the steady-state friction coefficient, probably because of build-up of a transfer film. The first test series, performed at 10 N, shows a relatively small variation of μ_{ss} in the range 0.1–0.15, following more or less the same track as the sliding speed increased and then decreased although not returning to the initial low value at the lowest speed. The second test series at 20 N shows very different behaviour. Even at the lowest speed, the coefficient of friction was significantly higher than with the smaller normal load. As speed increased the coefficient of friction rose gradually, but it then jumped upwards between the two runs at the highest speed, and continued to rise as the speed was decreased: behaviour reminiscent of that reported before (Wang and Woodhouse, 2011). It seems likely that there was an interaction between temperature, transfer film generation, and friction: higher friction causes additional heating, which softens the polymer leading to more wear and material transfer, leading in turn perhaps to greater real area of contact and even higher friction. Once the coefficient of friction reached 1 with the last few slow-speed runs, the whole dynamometer block started exhibiting self-excited vibration. This violates the assumptions behind the linearised frequency response measurement, so it is no surprise that the corresponding dynamic measurements of $\beta(\omega)$ gave very odd-looking results (not reproduced here) under these conditions.

Fortunately, at the lower normal force of 10 N the measurements of $\beta(\omega)$ gave clear results: Fig. 9 provides some examples. The general trends in both amplitude and phase look qualitatively very similar to those for nylon seen in Fig. 6. The frictional frequency response shows remarkable consistency between the first and last runs of the test sequence, as was seen with the nylon results, despite a slight increase of μ_{ss} between the two runs. The tentative conclusion is that the build-up of a transfer film has a more pronounced effect on μ_{ss} than on $\beta(\omega)$. For both nylon and polycarbonate the results support the suggestion by Wang and Woodhouse (2011) that the novel measurement method is able to provide a "fingerprint" of the

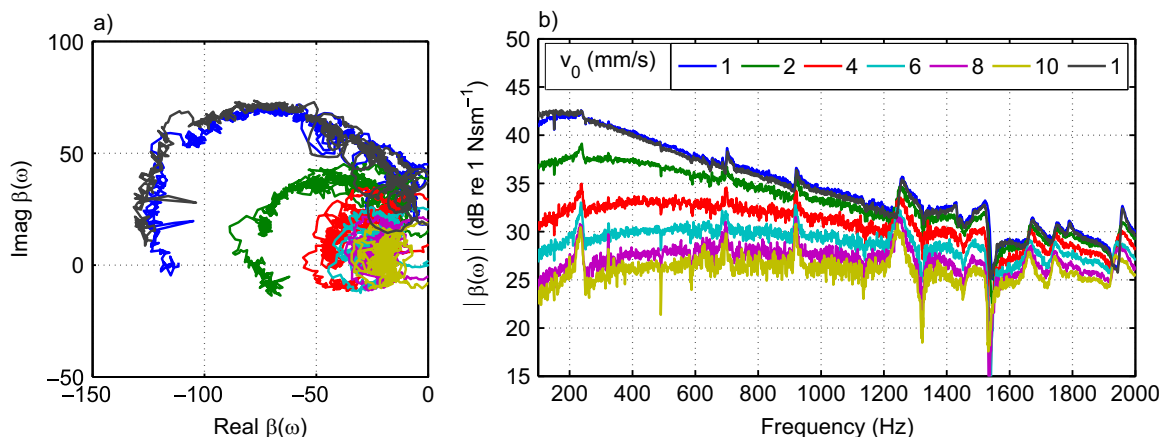**Fig. 9.** Polycarbonate on glass: Nyquist and magnitude plot of $\beta(\omega)$ for a fixed normal force $N = 10$ N and different sliding speeds. In (b), (c) and (d) the repeatability of the measurement is highlighted by the close agreement of the blue and black lines. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

sliding interface between the tested materials. However, the physical mechanisms behind the observed behaviour of $\beta(\omega)$ still need to be unraveled, and a satisfactory theoretical model formulated. This is the task of the remainder of this paper.

3. Friction models

3.1. Overview

As mentioned in the introduction, none of the constitutive models of friction explored in Woodhouse and Wang (2011) were found to be compatible with the dynamic friction measurements. Stribeck-type models based on a velocity dependent steady-state friction curve alone were a priori excluded, since they cannot give a complex or frequency-varying value of $\beta(\omega)$. The original comparison of Woodhouse and Wang (2011) was focused on the simplest of the rate-and-state models and a rate-temperature model. Both models can predict a circular shape for the Nyquist plots, but they are confined to the lower or upper half of the complex plane, whereas results such as those shown in Figs. 6 and 7 show progressive rotation of the circular segments by changing either the sliding speed v_0 or the normal force N, eventually crossing the real axis into the other half-plane. More strikingly, none of the models investigated in Woodhouse and Wang (2011) gave any useful clue to a possible mechanism for the non-trivial influence of the normal force.

Enhancements to the classical rate-and-state models will be suggested shortly, introducing additional physical ingredients and leading to a very encouraging fit to the experimental results. Before considering in detail the different modelling assumptions, a brief discussion of the background to rate-and-state (RS) models is presented. This family of models was originally motivated by measurements by Dieterich (1979) on rock samples, which showed that in response to a sudden velocity jump the variation of the friction coefficient is characterised by a stress/force relaxation over some slip distance or characteristic timescale, before a new steady state is reached. This observation clearly called for the introduction of some kind of additional internal state variable(s) into the friction model, with an evolution law capturing the relaxation process.

The original realisation of RS models, introduced in the context of earthquake mechanics, is known as the Dieterich–Ruina law (Dieterich, 1979; Ruina, 1980, 1983) and is defined by a friction force F of the form

$$F = N\mu = N[\mu_* + a\ln(v/V_*) + b\ln(\phi/\phi_*)], \quad \phi_* = L/V_*. \tag{3}$$

Coulomb's law is assumed here (N denoting the constant normal force), but it could be relaxed if necessary, allowing a nonlinear relation between the friction force and the normal force, as some experiments have suggested (Aronov et al., 1984; Shimamoto, 1986; Linker and Dieterich, 1992; Bureau et al., 2006; Scheibert et al., 2008). Eq. (3) is usually accompanied by one or other of the ad hoc empirical state evolution laws

$$d\phi/dt = 1 - v\phi/L, \quad \text{referred to as the Dieterich ageing law, or} \tag{4}$$

$$d\phi/dt = -(v\phi/L)\ln(v\phi/L), \quad \text{referred to as the Ruina slip law.} \tag{5}$$

In expressions (3)–(5), v denotes the slip rate of a frictional interface, subscripts "*" denote chosen reference values of the relevant variables, and a and b are dimensionless model parameters quantifying the rate-and-state deviation from classical Coulomb friction. The variable ϕ represents an internal state that can be thought as a measure of the resistance to slip of a frictional interface. It could have various possible physical interpretations (Putelat et al., 2011), but it is most commonly described as *the average lifetime of a population of interfacial asperity contacts* (Dieterich, 1979; Baumberger and Caroli, 2006). Parameter L is the so-called "memory length" of the frictional interface. Within the geophysical community, L is often assumed to be a material constant independent of the sliding velocity v. Such a characteristic length scale could define the slip distance required for the complete rejuvenation of the asperity contact population (Dieterich, 1979). The original experiments of Dieterich, followed by Ruina, Baumberger and many others (e.g. Ruina, 1980, 1983; Heslot et al., 1994; Marone, 1998; Persson, 2000; Scholz, 2002) gave a memory length that was typically of the order of a few microns. Historically, the first demonstration of the importance of such a memory effect in sliding friction was given by Rabinowicz (1951, 1957). Further evidence of the fact that friction depends not only on the instantaneous sliding velocity but also on the sliding history was later provided by Dieterich (1979), and confirmed by others for a wide range of materials (e.g. Dieterich and Kilgore, 1994; Marone, 1998; Baumberger and Caroli, 2006; Persson, 2000).

The Dieterich–Ruina laws (3)–(5) are phenomenological in nature and represent an attempt to close the gap between static and kinetic friction (Berthoud et al., 1999; Baumberger et al., 1999; Rice et al., 2001; Persson et al., 2003; Putelat and Dawes, 2015). The framework of RS models allows the unification of three important experimental observations: logarithmic time variation of static friction, the sliding memory effect already mentioned, and logarithmic velocity dependence of the steady-state friction force, that may be written as

$$F_{ss} = N\mu_{ss}(v_0) = N[\mu_* + (a-b)\ln(v_0/V_*)]. \tag{6}$$

The value μ_* is chosen to be the reference value of μ_{ss} at velocity V_*. Eq. (6) is deduced from assuming that, in steady-state sliding, the interfacial state variable tends to the value

$$\phi_0 = \phi_{ss}(v_0) = L/v_0, \tag{7}$$

which satisfies either of the state evolution laws (4) or (5).

Regardless of the exact mathematical expression for the frictional force (3) and the state evolution law (4) and (5), RS models can be more generally defined by a pair of equations of the form (see e.g. Rice et al., 2001)

$$F = f(N, v, \phi), \quad d\phi/dt = -g(N, v, \phi). \tag{8}$$

The state evolution Eq. (8)$_2$ could be vectorial if more than one internal variable is necessary (Ruina, 1983), but that possibility is disregarded in this paper for simplicity.

As stated above, RS models are phenomenological: the model coefficients and the functional forms are chosen to fit experimental results, without necessarily making explicit any relation between the physical properties of the tribosystem and the state variables or model parameters. Indeed, there is still no well-accepted and definite microphysical derivation based on first principles of mechanics. Nevertheless, constitutive laws of this general type have achieved a significant consensus, extending now beyond the scientific community of earthquake mechanics, and to date they seem to give the best way to mimic a variety of observed frictional behaviour ranging from microscale to large scale mechanical systems. Admittedly, this success is partly due to successive refinements of the state evolution law proposed over the last three decades (Ruina, 1983; Gu et al., 1984; Weeks, 1993; Heslot et al., 1994; Baumberger and Caroli, 2006; Putelat and Dawes, 2015) in order to fit different experimental datasets. Assessing which model has the largest application range is clearly one of the most challenging problems in this field, since large sets of different experimental data are required to perform any systematic discrimination among the models. Unfortunately, relevant literature is quite sparse.

That is precisely the task for the remainder of this paper: to move a significant step towards the long-term goal of identifying a general constitutive law able to give satisfactory predictions for a wide range of friction tests, based initially on the extensive data matrix for the nylon/glass experiments described in Section 2.2.2. Two major novelties of that test, compared to the bulk of previous literature, are the high sliding speed, and the frequency bandwidth extending into the kilohertz range. Both aspects are of direct interest in structural vibration problems such as brake squeal, but in the context of the broader question of model development and validation the main significance is that they provide an extension of empirical data into new regions of parameter space, and thus allow new aspects of candidate theories to be tested.

3.2. Frequency response function of classical rate-and-state models

The first step is to linearise (8) in order to deduce the corresponding frictional frequency response function $\beta(\omega)$. In the steady state condition, so that $v = v_0$, the equilibrium interfacial state (Eq. (7)) leads to $g(v_0, \phi_{ss}) = 0$, which means the time derivative of g is zero. Concurrently, the friction force curve is given by $F_{ss} = f(v_0, \phi_{ss}(v_0), N) := f_{ss}(v_0, N)$ and its slope by $f'_{ss} = df_{ss}(v_0)/dv_0$.

The experiments involve a small speed fluctuation v' to the sliding speed of the disc v_0 such that $|v'| \ll v_0$, so that the interfacial slip rate can be expressed as $v = v_0 - v'e^{i\omega t}$. Linearisation of (8) then gives the corresponding force fluctuation amplitude

$$F' = -f_{,v}v' + f_{,\phi}\phi', \quad \text{with } d\phi'/dt = g_{,v}v' - g_{,\phi}\phi' \tag{9}$$

where, for example, $f_{,v}$ denotes the partial derivative of the function f with respect to v, evaluated at the steady-state operating point. Taking the Fourier transform and solving the state evolution law in the frequency domain then gives

$$F'(\omega) = -f_{,v}v'(\omega) + f_{,\phi}\phi'(\omega), \quad \text{with} \quad \phi'(\omega) = g_{,v}v'(\omega)/(g_{,\phi} + i\omega). \tag{10}$$

Accordingly the frequency response function $\beta(\omega)$ of any basic RS friction model (8) is given by

$$\beta_{rs}(\omega) = -\frac{g_{,\phi}f'_{ss} + i\omega f_{,v}}{g_{,\phi} + i\omega}, \tag{11}$$

where the slope of the steady-state curve $f'_{ss} = f_{,v} - f_{,\phi}g_{,v}/g_{,\phi}$ follows from the differentiation chain rule. For example, the frequency response function for both variants of the Dieterich–Ruina law (3)–(5) is

$$\beta_{dr} = -\frac{(a-b)N/L + i\omega aN/v_0}{v_0/L + i\omega} \tag{12}$$

since $f'_{ss} = (a-b)N/v_0$, $f_{,v} = aN/v_0$ and $g_{,\phi} = v_0/L$.

As discussed in Woodhouse and Wang (2011), expression (11) is a complex bilinear transformation with real coefficients, satisfying the Hermitian symmetry condition $\beta(-\omega) = \beta^*(\omega) \; \forall \; \omega \in \mathbb{R}$, where $*$ denotes the complex conjugate. This condition ensures that β_{rs} corresponds to a real-valued impulse response. Separating the real and imaginary parts of (11) leads to the well-known result that such a bilinear transformation maps the real half-line $\omega > 0$ into a semicircle: it lies in the lower half of the complex ω-plane for a velocity weakening rate-and-state law (i.e. $f'_{ss} < 0$), and in the upper half-plane for a velocity-strengthening law (i.e. $f'_{ss} > 0$). This immediately shows that no rate-and-state law (8) can match measurements

like those shown in Figs. 6 and 7: the Nyquist plots showed approximately circular traces, but the centres of the circles clearly do not lie on the real axis, and in some cases the circles straddle the real axis rather than being confined to one half-plane. Furthermore, since Coulomb's law has been assumed the dependence on normal force is simple proportionality.

3.3. The role of contact stiffness

A clue about how to proceed has been given by Bureau et al. (2000): if a mechanical system exhibits high frequency oscillations with small amplitudes, elastic deformations near the interface may need to be taken into account. It was shown by Bureau et al. (2000) that in order to predict the amplitude of the first two harmonic components of the measured frictional shear response of a sliding mass, excited by a small modulation of the normal load with a given amplitude and a frequency of 120 Hz, it was necessary to compensate the slip rate for an interfacial shear stiffness. This suggests that it might be interesting to explore how to incorporate interfacial contact stiffness into an RS friction model: perhaps this might help explain the measurements of $\beta(\omega)$.

The concept of "contact stiffness" is relevant to both normal and tangential motion near a region of contact. The story begins with the classical Hertz theory (see for example Johnson, 1985), developed in the late 19th century. Hertz found a closed-form solution for the linear elastic deformation near the contact when an ellipsoidal solid is loaded against a half-plane, or equivalently when two crossed cylinders are loaded together. Given a normal force N, the linear elastic displacement is given by

$$\delta = \left(\frac{9N^2}{16\, RE^2} \right)^{\frac{1}{3}},$$

(13)

where the parameters E and R are composite values formed, respectively, from the elastic moduli and radii of curvature of the two surfaces in contact (see Johnson, 1985). It can be seen that the elastic displacement δ varies as the 2/3 power of the applied load N. The linearised contact stiffness is found by differentiating δ with respect to N, evaluated at the nominal load, and taking its inverse (Johnson, 1985). This suggests that the contact stiffness can be modelled by using a power law of the normal force N:

$$\frac{dN}{d\delta} = k_{b,i} \equiv \kappa N^{\alpha}$$

(14)

where for simple Hertzian contact, the exponent $\alpha = 1/3$. Similar results can be obtained for the tangential elastic response when a shear stress is applied across the contact: the normal and tangential contact stiffnesses are generally of the same order of magnitude (Johnson, 1985).

However, this is by no means the whole story. It has been well known since Bowden and Tabor (1950) that the real area of contact between two sliding surfaces is usually far smaller than the apparent area of contact. The discrepancy is caused by surface roughness: true contact occurs only at the tips of interacting asperities. Since friction and interfacial stiffness processes take place around these asperity tips, the starting point for most tribological analysis is to characterise the surface roughness. Hertzian contact theory gives the behaviour around a single asperity tip, and Greenwood and Williamson (1966) incorporated this theory into a statistical model, usually referred to as the GW model. Combined with other assumptions, they modelled a rough surface with an ensemble of spherical asperities, with a Gaussian distribution of heights. The GW model predicts that the real area of contact will be proportional to the normal force: contact areas at individual asperities grow more slowly than this by Hertzian theory, but the area is augmented by additional asperities coming into contact. This proportionality is believed to be the explanation of Coulomb's law. GW theory also predicts that contact stiffness (normal or tangential) is approximately proportional to the normal load, so that α would be close to unity (Adams and Nosonovsky, 2000).

Over the years, many refined and extended versions of the GW model have been proposed, which preserve the linear relationship between real contact area and normal load (for example Bush et al., 1975; McCool, 1986; Persson, 2001). However, these models were all developed by considering only small normal loads. For higher load variations the linear relationship breaks down and deviations from Coulomb's law must be taken into account: eventually the "seizure load" is reached, where the asperities are squashed flat and the real area of contact becomes equal to the apparent area (Johnson, 1985). Such variations have an influence on the contact stiffness.

In the current work, the power law assumption (14) will be investigated in Section 4, based on the measured data. The exponent α cannot be predicted a priori, it could range between 0 and 1: a simple linear spring gives $\alpha=0$, Hertz gives 1/3, simple rough-surface theory gives 1. An attempt is made in Section 4 to estimate values of α using the extensive results from the glass/nylon dynamic friction tests.

It is worth mentioning that there is a familiar type of friction model which includes some allowance for contact stiffness, typified by the LuGre model (deWit et al., 1995). It is important to note that this class of model is *not* relevant to the present study. Models of this type were developed in the context of position control, where the overall state of a frictional contact may be described as sticking or rolling, but where part of the contact footprint exhibits micro-slip (see for example Johnson, 1985). Under those circumstances, the "stiffness" and "frictional" parts of the contact stress distribution act in parallel. By contrast, the measurements under discussion here were obtained under conditions of full sliding at all times. As will now be
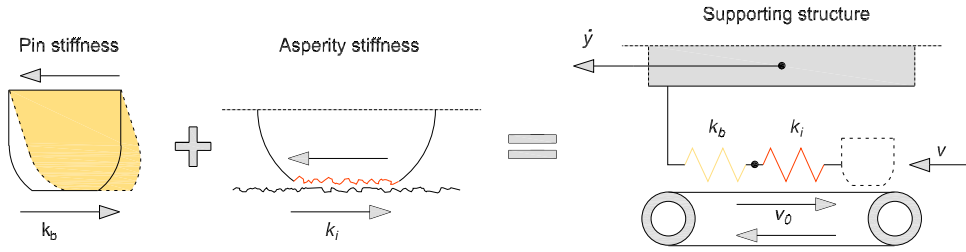
Fig. 10. Schematic stiffness modelling of the pin-interface compliance.

explained, under those conditions the "stiffness" and "frictional" components act in series, leading to significantly different dynamic response.

3.4. Frequency response function of enhanced rate-and-state models

The way that contact stiffness may influence the dynamic friction measurements is illustrated schematically in Fig. 10. In the measurement rig the sensors measuring force and velocity are placed as near to the contact zone as possible, but of course they cannot be exactly on the interface. The measured velocity, on the inner block (see Fig. 2), may differ slightly from the actual dynamic perturbation of slip rate on the interface because of two effects: the tangential contact stiffness determined by deformation of the asperities, as just discussed, and bulk shear deformation of the pin sample. These are represented in Fig. 10 by two springs in series: a stiffness k_i for the asperities, and another stiffness k_b for the bulk deformation of the pin. It is important to keep in mind that the goal of this paper is to find a model that matches the measurements, including any necessary imperfections in instrumentation: philosophical debates about whether a friction model "should" include these stiffness effects are not relevant here.

In Bureau et al. (2000) the rate-dependent term of an RS model (the equivalent of the function f in Eq. (8)) was modified by introducing a stiffness parameter analogous to k_i to characterise the shear deformation of the asperities within the interfacial contact region. To generalise this kind of model for further application requires some care: it is not immediately clear whether the two stiffnesses k_i and k_b play separate roles in any way, and it is also not clear whether some allowance for contact stiffness should also be included in the function g, in the evolution law for the state variable. These will be treated as empirical questions: models will be formulated representing different possibilities, and they will all be tested against the set of measurement data to see which perform best.

A natural way to introduce an interfacial shear stiffness within rate-and-state friction models is suggested by the sketch in Fig. 10: use an additive decomposition of the strain rate between elastic and "frictional" contributions as is done for a "Maxwell material". This suggests decomposing the slip rate into a sum

$$v = v_e + v_f, \tag{15}$$

where the elastic contribution to the slip rate is simply proportional to the stress rate, so that $v_e = \dot{F}/k_i$, while the explicit frictional contribution is expected to be a nonlinear function of the friction force $v_f(F)$. Then inverting this latter would lead to a Maxwell combined rate-and-state constitutive model of friction force having the form

$$F = f(v - \dot{F}/k_i, \phi), \quad \dot{\phi} = -g(v, \phi). \tag{16}$$

If in addition there is some bulk shear deformation between the pin-on-disc contact point and the measurement point of the velocity, the interfacial slip rate is given by

$$v = v_0 - \dot{y} - \dot{F}/k_b, \tag{17}$$

where $\dot{y}$ denotes the velocity as measured. Combining the last two expressions leads to the effective rate-and-state friction model

$$F = f(v_0 - \dot{y} - \dot{F}/k_t, \phi), \quad \dot{\phi} = -g(v_0 - \dot{y} - \dot{F}/k_b, \phi) \tag{18}$$

where the composite tangential stiffness k_t is given by the resultant of the two springs in series:

$$1/k_t = 1/k_b + 1/k_i. \tag{19}$$

The corresponding frequency response function then reads

$$\beta_c(\omega) = \frac{\beta_{rs}}{1 + i\omega f_{,v}/k_i - i\omega\beta_{rs}/k_b}. \tag{20}$$

where β_{rs} is given by (11). Eq. (20) will be called the "compliant pin and interface model" (CPI).

Two further models can be obtained by taking two different limits: $k_b \to \infty$ or $k_i \to \infty$. The first limit only accounts for the interface compliance due to the asperities, assuming that the pin is rigid enough to identify the velocity of the measurement point with the velocity of the contact point. It yields

$$\beta_i(\omega) = \frac{\beta_{rs}(\omega)}{1 + i\omega f_{,v}/k_i},$$
(21)

which will be referred to as the CI (compliant interface) model. The second limit neglects the elastic contribution of asperity deformation, but keeps the bulk compliance of the pin. The corresponding frequency response function is

$$\beta_b(\omega) = \frac{\beta_{rs}(\omega)}{1 - i\omega\beta_{rs}(\omega)/k_b}.$$
(22)

and will be referred to as the CP (compliant pin) model.

Exploring the limiting behaviour of the proposed models gives some useful information about their prediction performance. For example, expressions (20)–(22) all lead to similar asymptotic behaviour at low and high frequency. As would be expected from the definition of the frictional frequency response, the low frequency limit simply tracks the slope of the steady-sliding friction curve

$$\beta(0) = -f'_{ss}.$$
(23)

The low- frequency end of the β curves therefore always lies on the real axis in the Nyquist representation. For classical RS models, f'_{ss} depends on the difference $(a - b)$: for velocity weakening of the friction curve $(a < b)$, the limit of the Nyquist plot is located on the positive side of the real axis; for velocity strengthening $(a > b)$ it appears on the negative side. The value of the slope coefficient $(a - b)/v$ for the steady-state friction curve given by classical Dieterich–Ruina models (6) can be obtained by extrapolating the graph of $\text{Re}(|\beta|/N)$ as $\omega \to 0$ in a log–log plot. In general, such information would also show any departure from Coulomb's law for data obtained with different normal loads N if the results were not to collapse onto a single curve. Expressions (20)–(22) show that the different possible elastic corrections of rate-and-state models can be disregarded for frequencies such that

$$\omega \ll k_t/f_{,v},$$
(24)

or in other words, when the frictional Deborah number

$$\text{De} \equiv \omega/(k_t/f_{,v}) \ll 1.$$
(25)

At high frequency, the frequency response function is independent on the friction law as

$$\beta(\omega) \sim -k_t/(i\omega):$$
(26)

the asymptotic force fluctuation is purely elastic, with stiffness k_t (becoming k_i or k_b in the two limiting models). Again, a log–log plot of $\text{Im}(\beta)$ can be very instructive: it should show a straight line of slope -1, from which the stiffness can be deduced directly. If the stiffness is normal load dependent, as suggested by Eq. (14), that should emerge directly from the plot.

These asymptotic limits can be illustrated using the experimental data shown previously in Figs. 6 and 7. Fig. 11 shows log–log plots of the real (Fig. 11a) and imaginary (Fig. 11b) parts of $\beta(\omega)$, for different normal loads and for different sliding speeds, respectively. When the frequency goes to zero the real parts of the different measurements (all obtained at the same sliding speed) roughly collapse to a single line, broadly as just discussed. Somewhat less convincingly, when the frequency
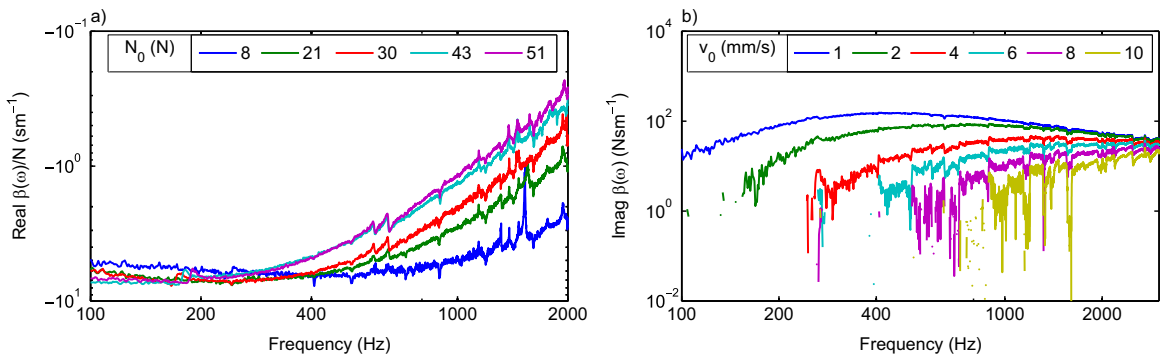

Fig. 11. Illustration of asymptotic limits for the experimental $\beta(\omega)$ curves of nylon-glass: (a) $\text{Re}(\beta(\omega))$ measured at different normal loads with a fixed sliding speed of 2 mm/s; (b) $\text{Im}(\beta(\omega))$ measured at different sliding speeds with a fixed normal load of 20 N.

tends to infinity the imaginary plot perhaps shows the different $\beta(\omega)$ functions converging asymptotically towards inclined straight lines, but in the frequency range shown here they do not all appear to have the same slope.

3.5. Enhanced non-monotonic Dieterich–Ruina laws

One limitation of the extended RS models presented so far is the assumption that the steady-state velocity dependence of the frictional force is monotonic. Although the current experimental observations showed monotonic velocity-strengthening behaviour over the speed range studied, other experimental results in the literature suggest that many tribological systems exhibit non-monotonic variation of the friction coefficient when studied over a wider range of speeds. This implies the possibility of crossovers between velocity-weakening and velocity-strengthening frictional force regimes, at least for some material combinations and conditions. In the simplest form, the frictional force decreases for a low-speed regime and increases when operating with higher velocities. However, in a more complicated case the steady-state friction μ_{ss} might show a "spinodal" (i.e. N-shaped) dependence: velocity-strengthening branches at low and high velocities and velocity-weakening behaviour over a range of intermediate speeds (Putelat and Dawes, 2015).

As shown in Putelat et al. (2010), a straightforward way to reproduce the simplest case of non-monotonic behaviour of the friction force is to introduce a small constant c in Eq. (3), leading to

$$F = N\mu = N[\mu_* + a \ln(v/V_*) + b \ln(c + \phi/\phi_*)]. \tag{27}$$

The parameter c is introduced in order to give a residual strength to the interface at very high slip rates, when the interfacial state is supposed to have no influence on the friction force since $\phi \approx 0$. The effect is to produce logarithmic velocity-strengthening at high velocity (Putelat et al., 2010). It may be mentioned that the constant c is related to the concept of transition or cut-off velocity V_c introduced by Weeks (1993) in order to simulate high-velocity strengthening: $V_c = V_*/c$.

Another common feature of traditional RS models, which may be challenged if a more general form is wanted, regards the interpretation of the memory length-scale L in the state evolution law. As described before, the Dieterich–Ruina law was originally formulated on the assumption that the interfacial state relaxes on a characteristic length-scale L, which is the distance over which the interface must slip before the effect of past variations in velocity on the current asperity population is lost. Such a characteristic length-scale would perhaps be independent of sliding velocity, and indeed it is often supposed to be a sort of "universal constant" in the fault mechanics community. This viewpoint has been called into question by Putelat et al., 2011; Putelat and Dawes, 2015, who introduced a relaxation time-scale t_ϕ rather than the length-scale L.

Of course, t_ϕ could be assumed to be simply a constant time, or it could be velocity dependent. For instance, a possible choice to fit the pattern provided by the Dieterich–Ruina laws is $t_\phi \equiv L/v$, in which case L is indeed the classical constant memory length. Following the argument of Putelat et al. (2011), it is possible to formulate a rather generic state evolution law as follows:

$$\dot{\phi} = -\frac{\phi - \phi_{ss}(v, \ldots)}{t_\phi(v, \ldots)} \tag{28}$$

where the relaxation timescale t_ϕ and the equilibrium state ϕ_{ss} are functions of the instantaneous interfacial slip rate v, and possibly of other variables like temperature or normal stress, a case which is disregarded here for simplicity. Then, the classical Dieterich–Ruina model (3) and (4) with a friction curve $\mu_{ss} \approx \mu_* + (a - b)\ln(v_0/V_*)$ can be recovered by assuming

$$\phi_{ss}(v) \equiv L/v. \tag{29}$$

From the experimental point of view, the challenge is then to measure the dependence on velocity, and possibly on other state variables, of the unknown functions ϕ_{ss} and t_ϕ in order to determine (28). A possible experimental protocol using the stick-slip boundary location within the velocity-stiffness parameter plane of a spring-block system was proposed in Putelat et al. (2011) in the special case where t_ϕ is proportional to ϕ_{ss}. In Section 4 of this paper, it will be shown how to determine t_ϕ in the case where ϕ_{ss} follows (29). For this, it is useful to rewrite the functional form of $\beta(\omega)$ derived from (28), by realizing that $g_{,\phi} = 1/t_\phi$, as

$$\beta_{rs}(\omega) = -\frac{f'_{ss} + i\omega t_\phi f_{,v}}{1 + i\omega t_\phi}. \tag{30}$$

Combining (27)–(29) and substituting $c = V_*/V_c$ based on the concept of cut-off velocity yields

$$\beta_{rs}(\omega) = -aN/v_0 + \frac{bN/v_0}{1 + v_0/V_c} \times \frac{1}{1 + i\omega t_\phi}. \tag{31}$$

Table 1
Enhanced monotonic rate-and-state models.

Model	Rate-and-state model $\beta_{rs}(\omega)$				Elastic contribution		
	$f'_{ss} = \frac{(a-b)N}{v_0}$	$f_{,v} = \frac{aN}{v_0}$	$g_{,\phi} = \frac{v_0}{L}$	$g_{,\phi} = \frac{1}{t_\phi}$	$\frac{1}{1+i\omega f_{,v}/k_i}$	$\frac{1}{1-i\omega\beta_{rs}/k_b}$	$\frac{1}{1+i\omega f_{,v}/k_i - i\omega\beta_{rs}/k_b}$
CI1	×	×	×		×		
CI2	×	×		×	×		
CP1	×	×	×			×	
CP2	×	×		×		×	
CPI1	×	×	×				×
CPI2	×	×		×			×

4. Comparisons with measurements

4.1. Candidate models

Several candidate models have been proposed, and for the sake of clarity it is useful to present a short summary to orient the reader in preparation for a systematic attempt to discriminate between them based on the experimental results. The sets of monotonic and non-monotonic RS models to be considered are listed in Tables 1 and 2, respectively. Each model appears in two variants, flagged by a final digit 1 or 2, depending on whether L or t_ϕ is used as the fitting variable. This distinction will make a difference when multiple datasets are to be fitted with a single model. The list of parameters to be fitted is as follows: for monotonic RS models with $c=0$ (Table 1) there is $k_{b,i}$, a, b and t_ϕ or L; for non-monotonic RS models with $c \neq 0$ (Table 2) there is the same list plus the cut-off velocity V_c. For both subgroups of models the stiffness parameter $k_{b,i}$ may be expressed as a power-law function of the normal load (14), in which case $\kappa_{b,i}$ and $\alpha_{b,i}$ are further two fitting parameters.

It is worth noting a general point about all these models. Because of the power-law function, the normal force N appears differently in the expressions for the "friction" and "stiffness" terms. Therefore, non-trivial variation of $\beta(\omega)$ with N can be expected to arise quite naturally: whether it is the right type of variation to match the measurements can now be explored.

4.2. Model fitting methodology

Any individual measurement of $\beta(\omega)$ can be best-fitted to a chosen model from the candidate set by minimising the cost function

$$\epsilon_{j,i} = \sum_{\omega=\omega_{min}}^{\omega_{max}} |\beta_{j,i}^{exp}(\omega) - \beta_{j,i}^*(\omega, \Phi_{j,i})|^2 \tag{32}$$

which is simply the sum over all frequencies of the squared difference between the experimental data $\beta^{exp}(\omega)$ for normal force j and sliding speed i and the corresponding theoretical prediction $\beta^*(\omega, \Phi)$. The vector Φ contains the fitting parameters. The minimisation will be performed here either using an unconstrained non-linear optimisation algorithm, specifically the Matlab routine "*fminsearch*" (which implements the Nelder–Mead method, well-known for its simplicity and low storage needs), or, when necessary, a constrained non-linear algorithm such as the one provided by the Matlab routine "*fmincon*". The latter is a gradient-based optimiser.

Remarkably, the result is that an entirely satisfactory fit can be found for any individual test and any chosen model from the set, in the sense that the disparity between model and measurement can always be made to lie within the noise bounds of the measurement. Examples will be shown shortly. There are two consequences of this successful fitting. First, it suggests that the inclusion of contact stiffness into the model, in one way or another, is indeed a promising way forward. But second, it shows that no single measurement gives enough information to discriminate between the candidate models: it will be necessary to make use of the full data matrix. This requires some care over the design of a fitting methodology.

Table 2
Enhanced non-monotonic rate-and-state models.

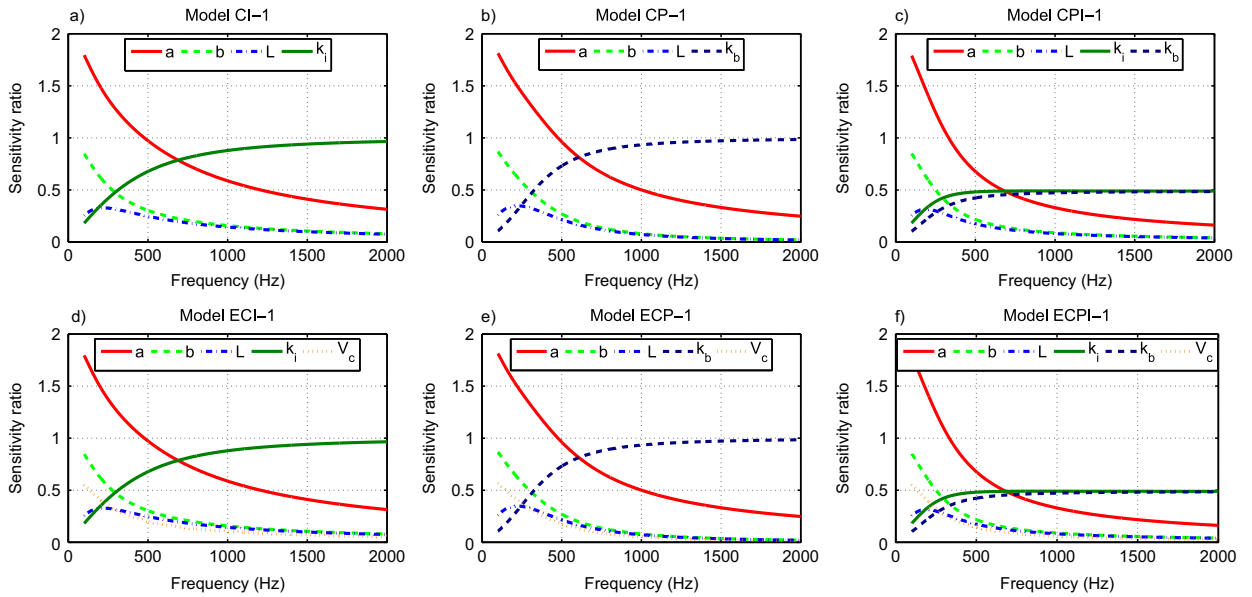| Model	Rate-and-state model $\beta_{rs}(\omega)$				Elastic contribution		
	$f'_{ss} = \frac{aN}{v_0} - \frac{bN}{v_0(1+v_0/V_c)}$	$f_{,v} = \frac{aN}{v_0}$	$g_{,\phi} = \frac{v_0}{L}$	$g_{,\phi} = \frac{1}{t_\phi}$	$\frac{1}{1+i\omega f_{,v}/k_i}$	$\frac{1}{1-i\omega\beta_{rs}/k_b}$	$\frac{1}{1+i\omega f_{,v}/k_i - i\omega\beta_{rs}/k_b}$
ECI1	×	×	×		×		
ECI2	×	×		×	×		
ECP1	×	×	×			×	
ECP2	×	×		×		×	
ECPI1	×	×	×				×
ECPI2	×	×		×			×

Fig. 12. Sensitivity analysis of monotonic and non-monotonic enhanced RS models.

As a first step, it is helpful to show a sensitivity analysis of the different fitting parameters. This highlights the relative importance of each model parameter, and provides useful information to interpret the fitting results. The metric used is the relative sensitivity of $\beta(\omega)$ to a given parameter X_i according to RS model j, as a function of frequency:

$$SR_{i,j}(\omega) = \frac{\partial \beta_j(\omega)}{\partial X_i} \frac{X_i}{\beta_j(\omega)}.$$

(33)

To compute this metric, each single variable was independently perturbed by a small amount, in this case +5%. The reference values of the model parameters were chosen by fitting to a particular nylon/glass experimental curve obtained with a normal force $N_0 \approx 30$ N and a speed $v_0 = 2$ mm/s. Results are shown in Fig. 12. It is apparent that the pattern of sensitivities shares common features over the different models. At high frequencies, the stiffnesses $k_{i,b}$ have the most pronounced effect on $\beta(\omega)$. At low frequencies, the friction parameters a, b and V_c dominate. It also appears that $\beta(\omega)$ weakly depends on the partial derivative of the state variable at steady-state sliding, $g_{,\phi}$.

It is useful to give a preliminary view of how the most significant parameters, namely a, b and the stiffness $k_{i,b}$, influence the general form of the frictional frequency response function. Since all the models are affected in a broadly similar way, only the first model CI1 is used for illustrative purposes. Fig. 13 shows how the Nyquist plot of $\beta(\omega)$ varies when the reference values of a, b and k_i are increased by 10%, 20%, 30%, 40% and 50%. The three parameters induce different kinds of change in the shape of the plots, all of which can be recognised as ingredients of the experimental results.
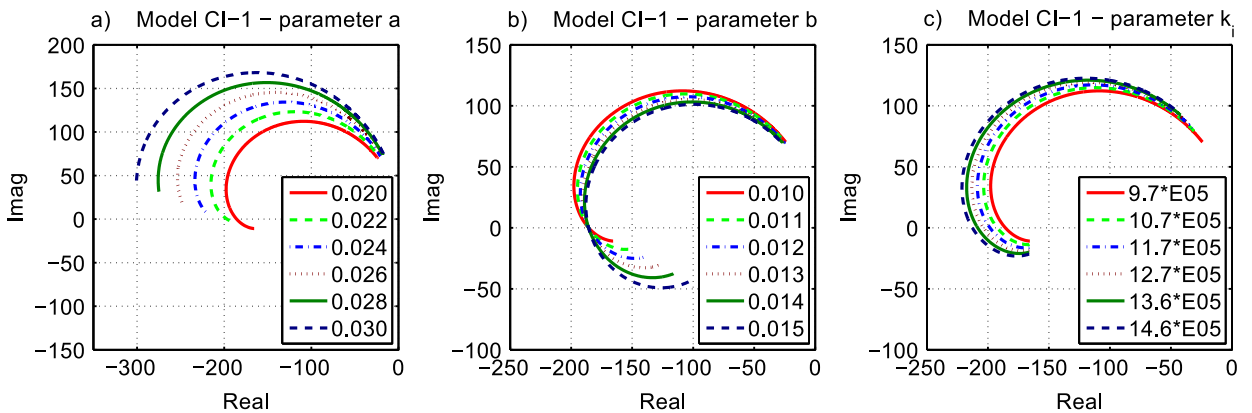

Fig. 13. Variations of the Nyquist plot of model CI1 induced by changing the parameters a, b and k_i.

4.3. Results for nylon/glass

4.3.1. Normal force variation and contact stiffness identification

Detailed results will now be shown for the measurements using a nylon pin and glass disc, since this represents the most complete set of data for different normal forces and sliding speeds. For the purpose of model fitting and discrimination, only the measurements for the sliding speed range 1–10 mm/s are used. The reason for this restriction to lower speeds is simply that the frictional frequency response function estimated for higher speeds becomes noisier and more prone to show rig artefacts. Some stages of the fitting methodology involve simultaneous fitting of multiple experimental results, in which case the appropriate cost function is obtained by summing the relevant versions of Eq. (32).

The first stage is to examine the contact stiffness and see whether it does indeed follow a power-law function as suggested in (14). If so, it is intuitively clear that the exponent α has to be treated somewhat differently from the other fitting parameters, because it cannot be fitted to any single test; it requires data at multiple normal loads. At the same time, a second question can be addressed: is it possible to discriminate between the different types of stiffness, i.e. interfacial, bulk or both? Two fitting approaches are explored to address these questions:

1. The first approach consists in fitting each single $\beta(\omega)$ measurement by using the effective stiffness $k_{b,i}$ as a fitting parameter rather than the power-law expression. By following this route, the fitting will provide a set of stiffness values for different sliding speeds and normal forces. A log–log plot of the identified stiffness against the normal force, for each different sliding speed, can then reveal directly whether a power law is appropriate.
2. The second approach assumes the power-law dependence of stiffness on normal force, and works by running the optimisation row-wise along the data matrix: at each single speed, each candidate model is fitted to a set of five $\beta(\omega)$ measurements collected under different normal forces, to estimate values of the scale factor $\kappa_{b,i}$ and the exponent $\alpha_{b,i}$ in Eq. (14).

First, it is useful to investigate whether or not the measured data makes it possible to discriminate between the two different types of stiffness, or if it is preferable to use a model based on only one stiffness component (i.e. k_i or k_b). This discussion represents a first step in the process of discrimination among the models listed in Tables 1 and 2. Four out of the 12 candidate models listed there involve two separate stiffnesses. For the sake of brevity, results are shown only for model CPI2: similar behaviour was seen for models CPI1, ECPI1 and ECPI2.

Fig. 14a and b shows the pattern of variation of the cost functions extracted from the two approaches described above. The results in Fig. 14a were calculated by taking the identified values of a, b and t_ϕ referred to a particular $\beta(\omega)$ curve having a sliding speed of 1 mm/s and a normal force around 20 N. The two stiffness values were varied over ranges covering the identified minimum point, and the cost function calculated for each combination. The red cross in Fig. 14a marks the identified minimum, which can be seen to lie in a narrow valley in this parameter subspace. The contours of the cost function follow an approximately hyperbolic shape, and furthermore these can be seen to follow reasonably closely the contour lines of k_t, overlaid on the plot. This immediately demonstrates that the optimisation procedure can only be expected to identify k_t, rather than to provide any reliable distinction between the interfacial and bulk compliances. Similar
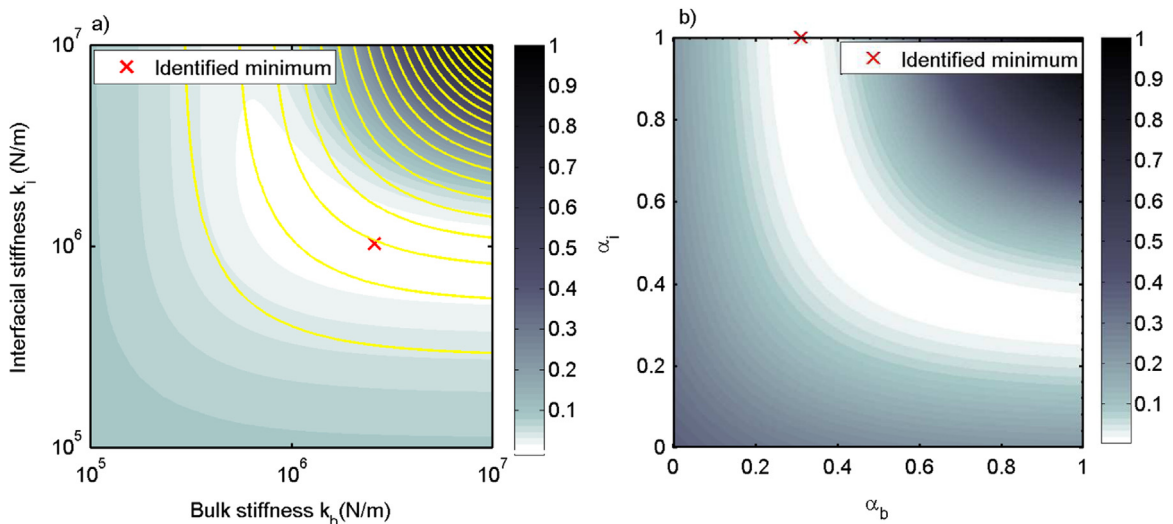

Fig. 14. Shaded contours of the cost function computed by using model CPI-2: (a) minimum search by changing both stiffness values, k_i and k_b, referred to $\beta(\omega)$ measured with a normal force of 20 N and a sliding speed of 1 mm/s; the yellow lines show contours of constant k_t in this parameter plane; (b) minimum search by changing both α parameters, referred to five $\beta(\omega)$ curves measured for different normal forces and a sliding speed of 2 mm/s. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)
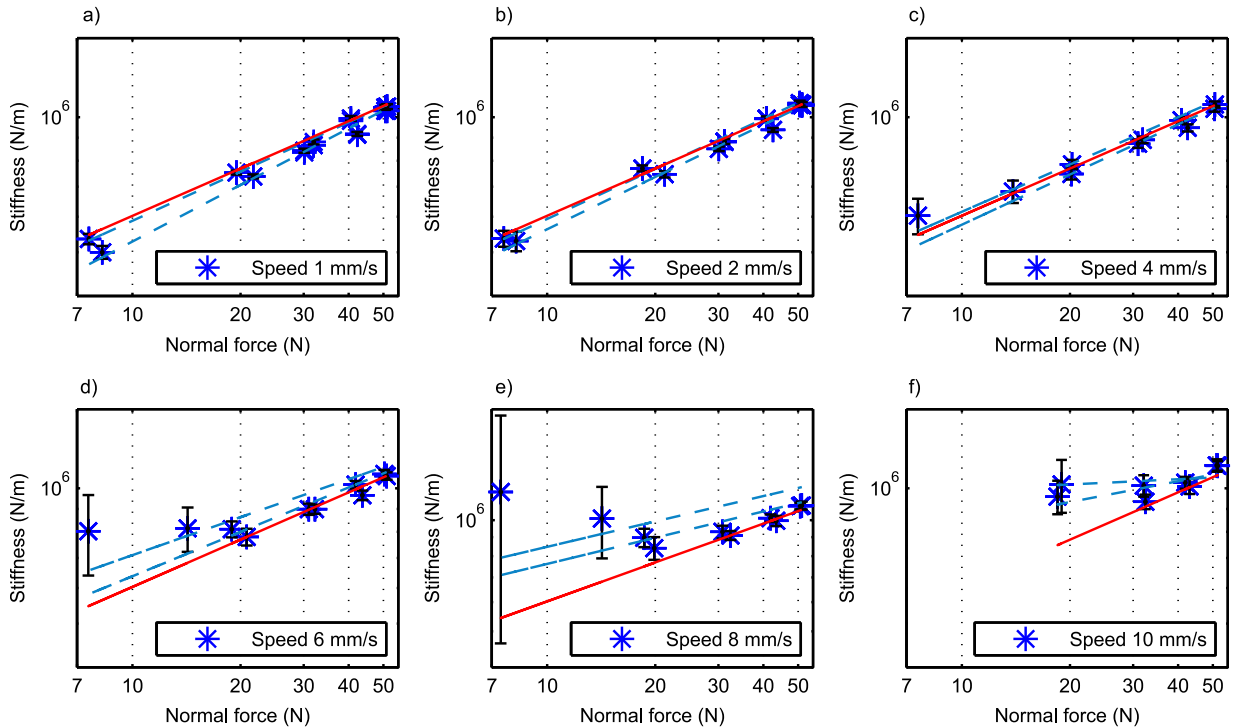
Fig. 15. Log–log plots of the identified stiffness values against the normal force for each sliding speed. Points are plotted for every model, but the values are so similar that they can scarcely be distinguished. The dashed blue straight lines define the power-law obtained from the row-wise fitting; note that there are two rows for each sliding speed, during the ascending and descending sequence of runs. The solid red straight line indicates the power-law identified from the global fit (see text). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

behaviour can be observed in Fig. 14b, calculated using the identified values of a, b, t_ϕ, κ_i and κ_b using five $\beta(\omega)$ measurements with different normal forces, all having a sliding speed of 2 mm/s. Again it is clear that the minimum lies in a narrow valley, so that no reliable distinction can be drawn between the two different α parameters.

These results strongly suggest that the CPI and ECPI models cannot be reliably fitted by using the current experimental data for model parameter identification. In addition, Fig. 14a has shown that the cost function varies very little along the k_t contour lines. This means that if a single stiffness parameter is used, as in the CI and CP models, the fitted value tends to represent the combined elastic contribution rather than the specific shear compliance of the interface or the bulk stiffness of the pin, even though in the model development separate notations k_i and k_b were used. From here on, the fitted contact stiffness will be regarded as giving a value of k_t rather than trying to make any distinction between an interfacial and bulk stiffness. Whether or not the tangential contact stiffness k_t needs to be introduced in the state evolution law, as described in model CP, will be discussed later.

Having excluded models CPI and ECPI from consideration, the models that remain all feature a single contact stiffness. As mentioned above, the first approach consisted in fitting a value of this stiffness to each $\beta(\omega)$ curve independently. Fig. 15 shows these identified values on a log–log plot against the normal load, separately for each sliding speed. Values are plotted for all the fitted models, but they cluster tightly in the plots: this shows that regardless of the type of model used, the fitting process leads to similar values of stiffness for each $\beta(\omega)$ measurement.

It is immediately clear that for lower sliding speeds the points lie close to a straight line in each case, suggesting that power-law dependence on normal load is indeed a good approximation to the observed behaviour. At higher sliding speeds, though, the results deviate from straight-line behaviour at lower values of normal force. To interpret this deviation, it is necessary to take account of the uncertainty of each identified stiffness parameter. This uncertainty was estimated by running the optimisation for contact stiffness again, artificially forcing it to converge to a value resulting in an error 5% bigger than the minimum found before. The change of the identified stiffness was used to plot error bars around the points in Fig. 15. It can be seen that the values which deviate most strongly from the straight-line trends also have the largest error bars: this is mainly a result of the noisier nature of the associated experimental data.

Each subplot contains three straight lines, whose slopes and intercepts give the exponents α and the stiffness scale factors κ for the corresponding power law. The solid red line has a fixed slope equal to 0.349, corresponding to the value obtained from a global fit to the entire matrix of fitted stiffness values. Interestingly, this value of α is close to that predicted for an ideal Hertzian contact. The dashed blue lines have a different slope for each sliding speed. These different slopes reflect the identified values of α obtained by performing the second optimisation approach. For low sliding speeds, the dashed and solid lines agree quite closely. For the higher speeds, the dashed lines represent a kind of compromise fit to data
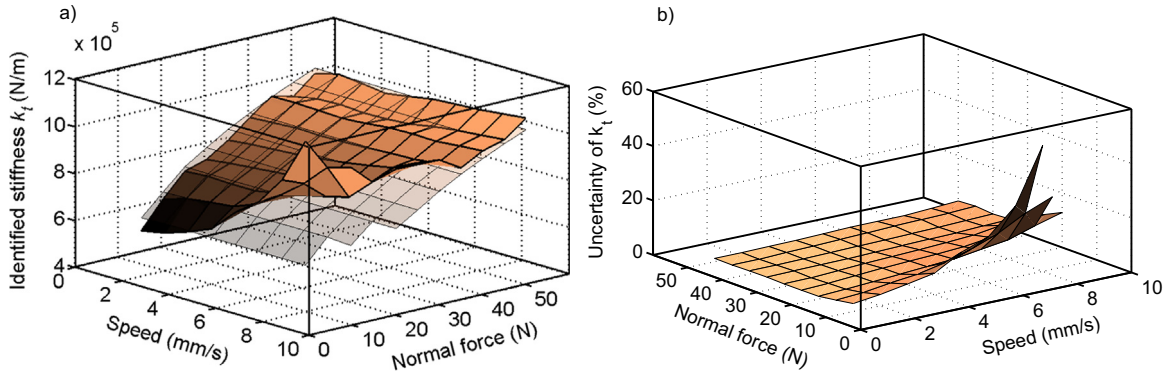
Fig. 16. (a) Identified contact stiffness values; (b) relative uncertainty of these values, expressed as a percentage. The smooth shadowed surface in (a) represents the global fit (see text). Note that different view angles have been used for the two plots, for maximum clarity.

that does not really follow a power law variation. Taken at face value, these dashed lines would suggest that α should be a function of sliding speed. However, as has already been pointed out, the pattern of error estimates in the fitted values suggests that such an interpretation should be treated with caution. Notice that even for these higher sliding speeds, the solid line continues to fit well for the data points with low error.

Fig. 16a collects the fitted stiffnesses into a 3D surface plot against sliding speed and normal force. It also includes, as the shadowed surface, the global fit with $\kappa_t = 2.7 \times 10^5$ and $\alpha_t = 0.349$. Fig. 16b shows the error estimates in a similar format, expressed as a percentage of the base value. These plots show very clearly that the global fit works well over most of the parameter space, but that the results deviate in one corner where speed is high and force is low. The errors increase dramatically in the same corner. However, it is by no means clear that these rising errors justify ignoring the anomalous behaviour entirely: the stiffness results do not simply become more random in this corner: there is a systematic tendency to increase. This observation suggests that further work is called for, but the issue will not be pursued here because the present set of measurements cannot give a clear resolution. For the purposes of fitting other parameters in the following Sections, the global fit to contact stiffness represented by the shadowed surface in Fig. 16a is used.

4.3.2. Full-matrix fitting and model parameter identification

The next stage is to use the entire data matrix to explore the relatively subtle differences between the remaining proposed models, and decide if there is a clear best choice among them. Now that the contact stiffness has been characterised, the remained parameters can be estimated by fitting each measured response independently. Each of these fitted models can then be used to predict the entire data matrix, and the cost function obtained by summing all the versions of Eq. (32) for the full matrix can be calculated. The results can be used to address some key questions:

1. Is it preferable to model the relaxation of the internal states using a fixed time scale t_ϕ or a fixed length scale L?
2. Are the (slightly more complicated) non-monotonic RS models necessary to fit the current experimental data?
3. Should the contact stiffness be introduced in the state evolution law as well as in the friction law?

The first question can be tackled simply by plotting the parameter values for L and t_ϕ obtained from fitting each $\beta(\omega)$ measurement. As shown in Fig. 17a, the memory length L appears to be linearly dependent on the sliding speed v_0, regardless of the model. Conversely, the fitted relaxation time-scale t_ϕ seems to be rather constant (Fig. 17b), apparently
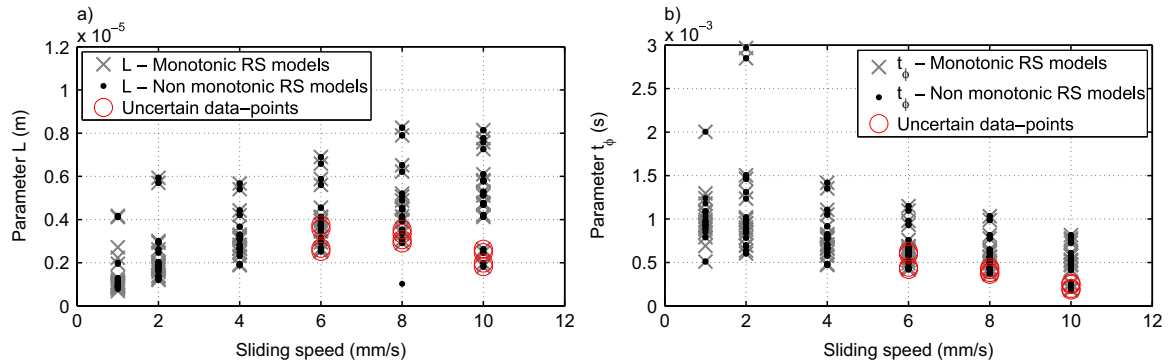

Fig. 17. Identification of parameter (a) L and (b) t_ϕ for the enhanced RS models (Nylon pin on glass disc). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

independent of the velocity. In both plots the scatter of the results is mostly due to the presence of noise in the experimental data. A few outliers can be seen at low speeds, which are linked to two "noisier" $\beta(\omega)$ measurements obtained at 10 and 40 N. The red circles mark the data-points related to cases that showed a significant deviation from the global fit of contact stiffness (i.e. $\beta(\omega)$ measured around 10 N and sliding speeds higher than 4 mm/s). These results clearly suggest that, for the present material combination, the relaxation process of the internal state variable can be well modelled using a fixed time-scale. The median value of t_ϕ is 0.67 ms, and the sensitivity analysis in Fig. 12 suggests that the quality of the prediction would not be much changed by varying that value within the range 0.5–1.5 ms seen in Fig. 17b.

This conclusion challenges the assumption of a constant memory length, mentioned in Section 3.5 and commonly used in geophysical problems. It is not suggested here that the geophysical models are wrong: their materials are different (crystalline rather than polymeric) and the sliding speeds used in their experiments were much lower then the ones used in the current work. However, it might be interesting to test such materials in a similar experiment to the one described here, since the present methodology offers an alternative perspective on this question.

To begin addressing the second listed question, a useful consistency check between model behaviour and the observed experimental data can be performed by inspecting the fitted parameters a, b and V_c. Since the measured steady friction force exhibits a velocity strengthening nature, the analysis from earlier Sections gives rise to the following expectations:

- For the monotonic RS models, $a > b$.
- For the non-monotonic RS models, $a < b$ since the presence of V_c produces a sign change.

It should be noted that these two alternatives may have physical significance: geologists have observed cases for which the sign of $a - b$, i.e. the slope of the friction curve $\mu_{ss}(v_0)$, can change as a function of temperature (Brace, 1972; Stesky et al., 1974; Scholz, 1998). Such a change can strongly influence the spatio-temporal complexity of frictional slip along faults and consequently the earthquake and stick-slip dynamics (Rice, 1993). Changes of monotonicity of the friction curve have also been observed in a range of other materials (e.g. Grosch, 1963; Shimamoto, 1986; Kilgore et al., 1993; Weeks, 1993; Heslot et al., 1994; Estrin and Brechet, 1996; Tsutsumi and Shimamoto, 1997; Bar-Sinai et al., 2014).

The histograms in Fig. 18 confirm that both types of model give rise to the expected combination of model parameters a and b. For the monotonic models, the parameters a and b show relatively tight distributions around mean values of 0.02 and 0.009, respectively. The non-monotonic models yield a similarly tight distribution for a, but the distribution for b, while giving consistently higher values than that for a, is much more broad and ill-defined. It should be mentioned that the "occurrence" numbers for the non-monotonic models are higher because more optimisation runs were performed, using many different starting values of the model parameter V_c. This proved to be necessary because, in contrast to what was found for the other parameters, a change of the starting value for the cut-off velocity often led to a different result.

In light of this it is interesting to observe what range of V_c values the different non-monotonic models suggest, bearing in mind that the direct measurements of mean friction force do not apparently exhibit any transition from velocity-weakening to velocity-strengthening in the sliding speed range studied. The distribution of the fitted parameter V_c for all the non-monotonic RS models is shown in Fig. 19. The distribution is concentrated at low sliding speeds, $v_0 < 2$ mm/s, with the peak around speed 0.8 mm/s.

As a physical check on these fitting results some additional sliding contact runs were performed at a fixed normal force around 30 N, at low sliding speeds down to a minimum value of 0.1 mm/s. For these low-speed runs only the steady-state coefficient of friction μ_{ss} was measured: for disc speeds lower than 1 mm/s, the present rig configuration makes it hard to avoid reversals in the velocity of the pin during dynamic testing, preventing a meaningful measurement of $\beta(\omega)$. The results are plotted in Fig. 19. The values are quite scattered, but the estimated μ_{ss} appears to change from velocity-weakening behaviour at speeds lower than about 0.4 mm/s to velocity-strengthening behaviour at higher speeds. This experimental evidence should only be regarded as a preliminary consistency check, but the approximate agreement is very encouraging. In order to validate the non-monotonic RS models in a more definitive way, it is likely that different types of dynamic
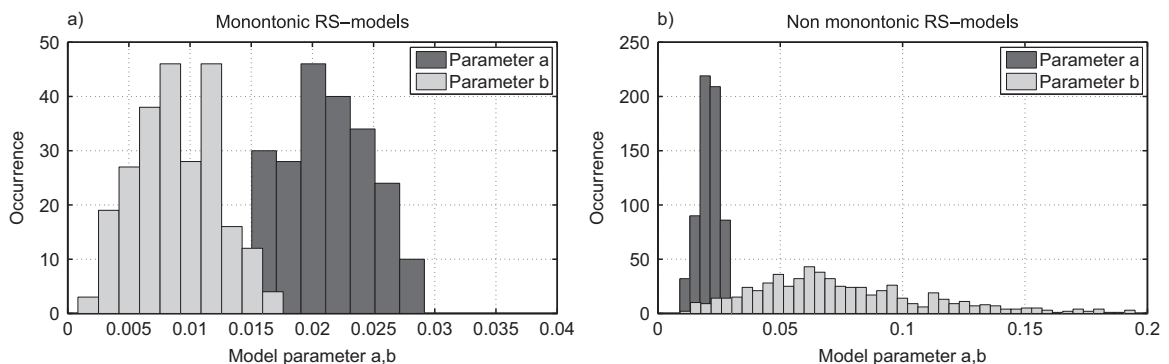

Fig. 18. Histograms of parameters a and b for (a) monotonic and (b) non-monotonic RS models (Nylon pin on glass disc).
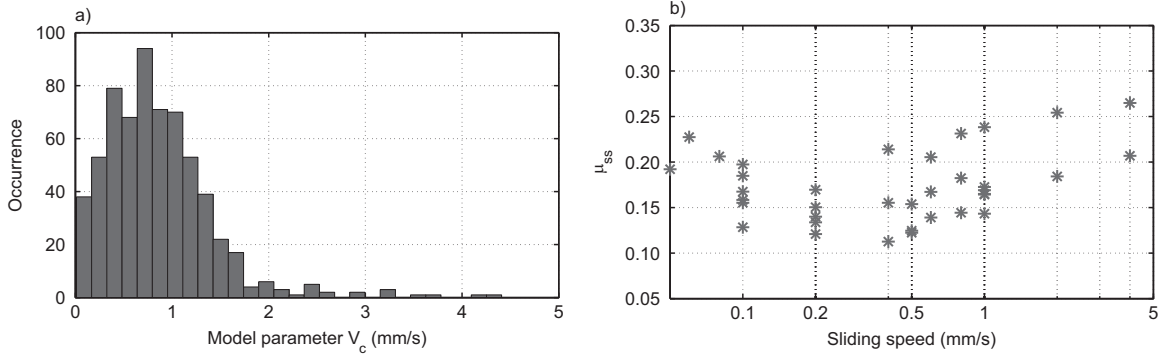
Fig. 19. Distribution of (a) parameter V_c and (b) experimental measurements of the steady state curve $\mu_{ss}(V)$ focused for lower speeds and obtained with a normal force of 30 N (Nylon pin on glass disc).

friction measurements will be necessary. It can be noted that the sensitivity analysis in Fig. 12 suggests that a change of the cut-off velocity V_c affects the prediction behaviour only slightly, predominantly at low frequencies.

The second of the listed questions can now be addressed. It has already been seen that the fitted model parameters a and b are more tightly distributed around mean values for the monotonic RS models than for the non-monotonic models. A more general comparison of the two families of models can now be carried out, as mentioned earlier. The eight remaining models have been independently fitted to the 60 $\beta(\omega)$ curves separately, and simulations run for each fit in order to predict the whole data matrix. For every model it is then possible to estimate the normalised Schwartz Bayesian Information Criterion (SBIC):

$$\text{SBIC} = \ln(\hat{\delta}_{err}^2) + (p/n)\ln(n) \tag{34}$$

in which $\hat{\delta}_{err}^2$ is the error variance between the model predictions and the whole experimental data matrix, n is the number of observations used to compute the error vector and p is the number of fitted parameters. The SBIC quantifies the goodness of fit per simulation (Wei, 2006). The last term in (34) is a penalty function that takes into account the number of fitting parameters. Models with low values of SBIC should be preferred to other models. Fig. 20 summarises some statistical information about the estimated SBIC indices, produced using the Matlab function "*boxplot*". The ends of the rectangles show the 25th and 75th percentiles of the distribution, and the central lines indicate the median. The vertical dashed lines are used by "*boxplot*" to identify outliers, indicated by the isolated dots.

In terms of the medians, it is immediately evident that all the monotonic RS models (the first four) exhibit a better prediction performance than the non-monotonic models. This seems to give a clear answer to the second question raised above. The results also suggest an answer to the third question: based on the prediction results, model CI2 stands out as the most suitable model for the current set of experimental data. Summarising, model CI2 is a monotonic model, that includes the contact stiffness only in the friction term of the RS model (not in the state evolution law).

The final challenge is to see if this "best" model can really give acceptable predictions for the whole data matrix, capturing the dependence on normal force and sliding speed. The distributions of identified model parameters a, b and t_ϕ shown in Figs. 17b and 18a give a promising clue. All three parameters show a confined and almost uniform distribution
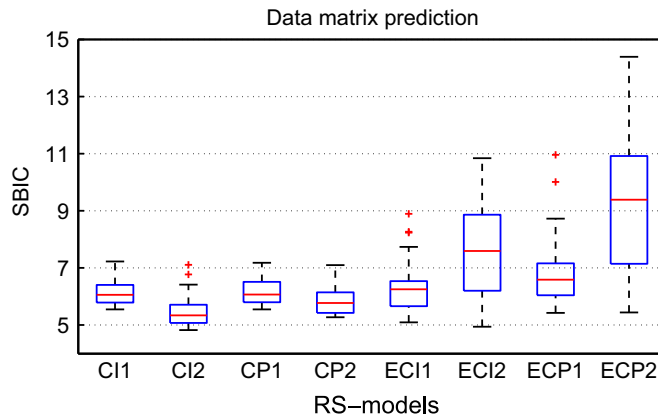

Fig. 20. Statistics of the different SBIC indices for each enhanced RS model (Nylon pin on glass disc). The central horizontal line indicates the median of the different SBIC values, the rectangle box defines the 25th and 75th percentiles of the data distribution and the outliers are marked as red dots. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)
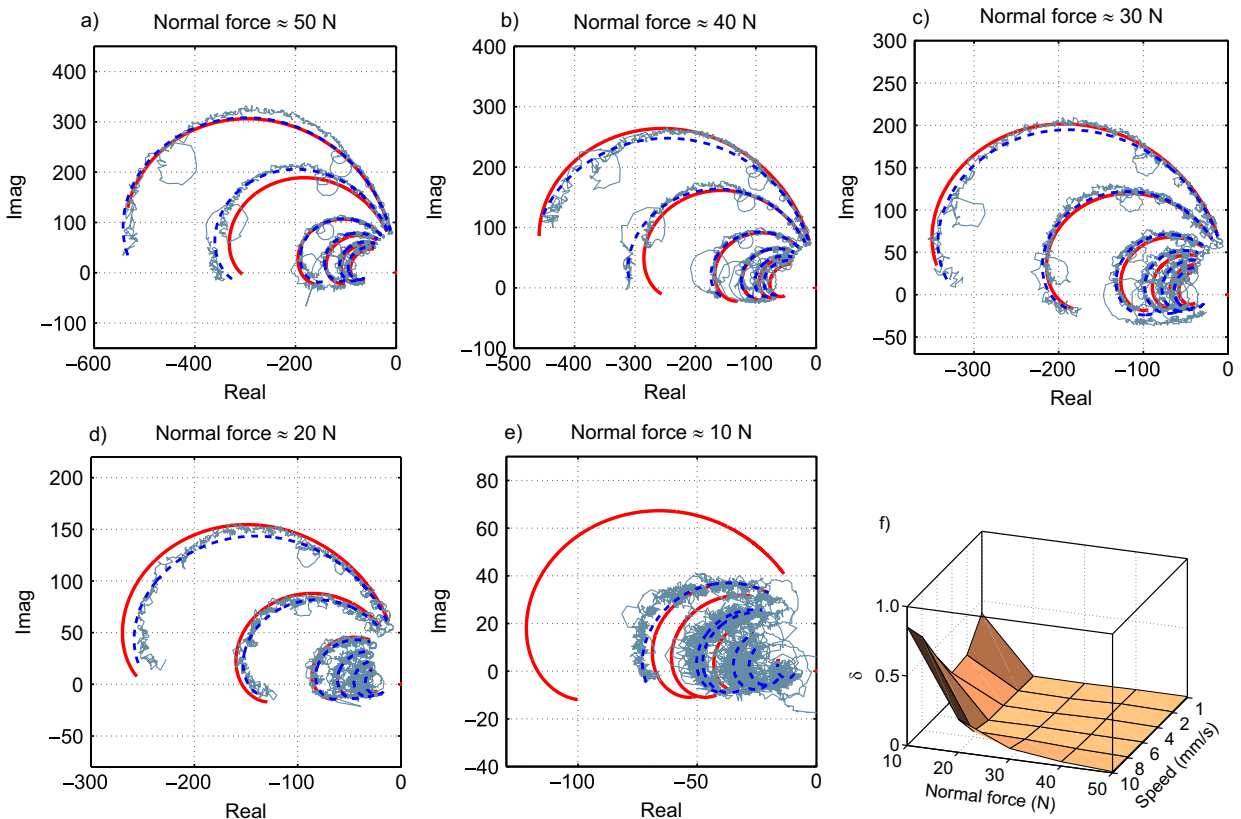
Fig. 21. Nylon pin on glass disc. Fitting performance of model CI2 vs. whole data matrix of $\beta(\omega)$: the solid red lines within the Nyquist plots indicate the global model prediction performed by fitting the whole data matrix; the dashed blue lines indicate the single curve fitting; the error plot (f) indicates the normalised prediction error (δ) for each curve of $\beta(\omega)$. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

around their mean values, and in addition no clear correlation was found between these parameters and the sliding speed or normal force. This suggests that a single set of such parameters might be good enough, at least for this specific material combination. In this context, previous studies on RS models by geologists (Scholz, 1998) have already suggested that the parameters a and b are likely to have a material-dependent nature.

A global fit of model CI2 was carried out, using as experimental reference the upper part of the data matrix, which corresponds to the $\beta(\omega)$ measurements done by increasing the sliding speed. This produced the following set of model parameters: $a=0.0196$, $b=0.0094$ and $t_\phi = 0.00075$. These are reassuringly close to the mean values estimated before. Each case can then be simulated using the appropriate measured sliding speeds and normal forces. Fig. 21 shows the quality of the resulting predictions for five subsets of data, clustered for similar normal forces. Each Nyquist plot contains the experimental data, a red solid curve that shows the prediction of the global model, and a dashed blue line that corresponds to the individual best fit of each experimental curve.

It can be seen that in most cases the global fit gives remarkable agreement with the measured $\beta(\omega)$, and usually lies very near the individual fit to that particular measurement. The model seems to catch the changes produced by varying the normal force and the velocity of the sliding contact quite naturally: the amplitude of the Nyquist segments increases as the normal force increases from 10 to 50 N, and for higher sliding speeds the Nyquist circles rotate somewhat in the counterclockwise sense.

In order to quantify the accuracy of prediction, a normalised version of Eq. (32) was used. The resulting values are plotted over the parameter space of the measurements in Fig. 21(f). It can be seen that most of the experimental data is well reproduced by the global model CI2. However, it was not possible to get a satisfactory match for the subset of $\beta(\omega)$ curves obtained with a fixed normal force at 10 N, as indeed can be seen in Fig. 21(e). This departure from otherwise excellent agreement is not currently understood. That data set was noisier, and it has also been shown earlier that the same subset of data strongly deviates from the global power-law fit for contact stiffness (see Fig. 15). However, this may not be the full explanation for the deviation between prediction and measurement here.
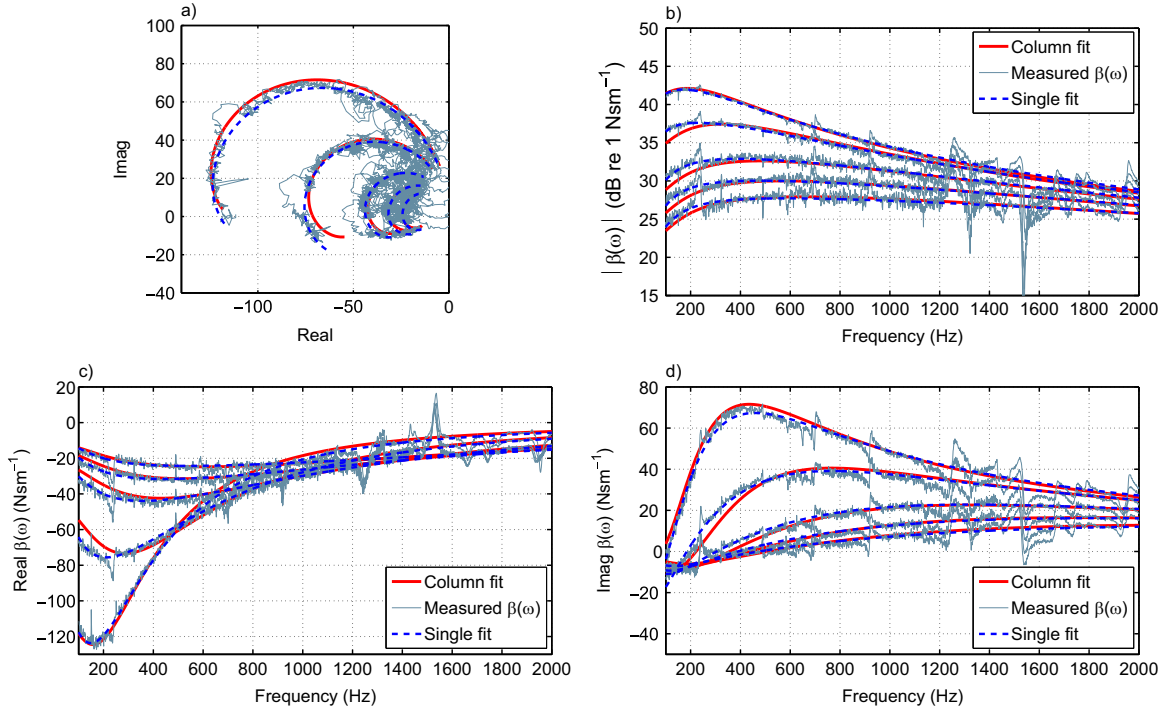
Fig. 22. Polycarbonate on glass disc. Fitting performance of model CI2 vs. experimental $\beta(\omega)$ with a normal force around 10 N and increasing sliding speed (1 2 4 6 8) mm/s: the solid red lines indicate the model prediction performed by combined fitting of the subset of data; the dashed blue lines indicate the single curve fitting. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

4.4. Results for polycarbonate/glass

The proposed model has been shown to work well for the particular material combination of nylon against glass. This combination was chosen principally on pragmatic grounds; readily available materials, easy to use in this particular rig, and not showing much tendency for self-excited "squeal" under the conditions tested. But it cannot be claimed that this combination has any very great technological significance in terms of friction testing. In order to start the process of exploring the applicability of the model to other materials and conditions, some preliminary results can be shown for a different combination, a polycarbonate pin on a glass disc.

As already mentioned in Section 2.2.3, because of the more pronounced wear process which characterised the polycarbonate pin and the consequent appearance of self-excited squeal, only two series of runs were performed having fixed normal forces of 10 N and 20 N, respectively. Given the limited nature of the available data, only the most representative fitting results obtained using model CI2 are illustrated here, although all the proposed models were also investigated. However, it was not possible to follow exactly the fitting methodology described above, because the contact stiffness could not be fully characterised with only two subsets of measured data. Therefore, the adopted fitting parameters are: a, b, t_ϕ and k_t. The experimental curves were first fitted independently, and then combined fittings for each series of data were performed.

Fig. 22 shows the results obtained for the first subset of experimental curves with different sliding speeds (1–2–4–6–8 mm/s) and a constant normal force around 10 N. In a similar format to Fig. 21, the dashed blue line indicates the single fit of each $\beta(\omega)$ curve and the solid red line shows the single fit to the combined subset of data. The results show behaviour that is, in general terms, very similar to what was seen for nylon and glass. Again, remarkable agreement between the model and the experimental curves is seen. The different mechanisms triggered by a change of sliding speed, which are a decrease in amplitude and a slight counterclockwise rotation of the Nyquist circles as velocity increases, seem to be well reproduced by the proposed model. The parameter values obtained by performing a single fit for the entire subset of data are: $a=0.021$, $b=0.012$, $t_\phi = 0.00064$ s and $k_t = 3.4 \times 10^5$ N/m. The first three are quite close to the ones identified for the Nylon pin.

The second subset of data, obtained with normal force 20 N, was more problematic, as described earlier: the mean friction force grew during the measurement series, and by the end the rig began to squeal spontaneously. In the light of this, it was a pleasant surprise to find that the dynamic measurements could still be matched reasonably well by the same model, although in order to achieve this a somewhat different set of parameter values was necessary: $a=0.029$, $b=0.016$, $t_\phi = 0.0012$ s and $k_t = 5.2 \times 10^5$ N/m. This seems to be another example of the robustness and stability of $\beta(\omega)$ measurements, compared to the more familiar "twitchiness" of steady friction measurements. The results are generally similar to those shown in Fig. 22 and are not reproduced here. The most likely reasons for the changed parameter values involve the

pronounced wear process during the second test series, and a probable difference in the nature of the film transferred to the glass. There is also the fact that the contact stiffness could not be fully characterised. At this stage, nothing further can usefully be said about this material combination apart from underlining again the goodness of fit to the proposed model.

5. Discussion and conclusions

In earlier work (Wang and Woodhouse, 2011) it was argued that to characterise a frictional interface in a form appropriate to the prediction of self-excited vibration such a brake squeal, a novel measurement was required. Specifically, if linearised theory is to be used to predict the threshold of instability, and associated unstable frequencies, then a linearised description of sliding friction is needed. It takes the form of a frequency response function, which encapsulates the amplitude and phase of the force perturbation in response to a small fluctuation in sliding speed at a given frequency. Some preliminary measurements of this frictional frequency response were shown: they revealed that it is indeed a complex and frequency-varying function.

In a second paper (Woodhouse and Wang, 2011), attempts were made to compare these measurements with the predictions of some models for dynamic friction taken from the literature. None of the models tested were able to give even qualitative agreement with the measurements, and this left a puzzle. The work presented here proposes a solution to this puzzle. A model has been presented which, at least for two particular material combinations involving a polymer in contact with a glass disc, gives very satisfactory agreement with measurements. This model is based on the family of "rate and state" models, developed initially in the context of rock mechanics but later applied to many other materials; the crucial extension of these models for the present purpose involves taking contact stiffness into account. Contact stiffness is known to be important in other contexts such as frictional dampers and the position control of mechanical actuators, but there is very little existing literature in which it is studied in connection with rate-and-state models for sliding friction (Berthoud and Baumberger, 1998; Bureau et al., 2000; Bouchbinder et al., 2011).

A brief overview was presented of a general framework for rate-and-state models, within which a number of commonly used models can be placed. The physical background for contact stiffness was discussed, and a number of related models were proposed to allow contact stiffness to be incorporated into the rate-and-state framework. A methodology was then developed to make systematic use of a new body of experimental measurements to test these models, to discriminate between them, and finally to identify the best candidate model from the set. This chosen model was shown to match the experimental results well, including catching in a natural way the variation with imposed changes in sliding speed and normal force. Both these parameters had previously been shown to make non-trivial changes to the magnitude and functional form of the frictional frequency response.

To allow these model investigations, a more extensive set of experimental measurements was needed than those published in the earlier study. For the case of a nylon pin in contact with a glass disc, a full matrix of data was collected with a chosen set of sliding speeds and normal forces. Similar measurements were attempted with a polycarbonate pin against the glass disc, but in this case only a limited number of tests were possible before significant changes took place in the contacting surfaces: probably because of thermal effects, the mean coefficient of friction rose to very high levels, after which the rig exhibited spontaneous squeal-like vibration. Nevertheless, such measurement as could be made were found to agree well with the same dynamic model, with parameter values quite similar to the case of nylon.

The difficulties encountered with the polycarbonate measurements serve to highlight an important general observation. Friction measurement is notoriously "twitchy": it is hard to get repeatable results, and apparently small changes (for example in surface contamination) can have large effects. The growth in mean friction for the polycarbonate was a fairly typical example of the difficulties encountered in such measurements. Strikingly, the dynamic frictional frequency response function seem to be far more robust and repeatable. This has now been shown experimentally in several different ways. This robustness could be rather good news for the prospects of squeal prediction. At least provisionally, it appears that the quantity required for squeal prediction is quite reproducible and consistent, to a degree that is quite unfamiliar to those accustomed to tribological measurement.

There is a corollary to this observation. Far and away the most common tribological tests, as applied to such things as brake lining materials, involve measuring the friction force during steady sliding. This is the context of the familiar "twitchiness" of measurements. Speed, normal load, and environmental and material factors may be varied between tests, but all such measurements are fundamentally at zero frequency in the language of the frequency response function investigated here. Furthermore, those involved in designing braking systems often try to infer the dynamic properties needed for squeal prediction from such zero-frequency tests. The results presented here demonstrate that this approach is fundamentally flawed. Squeal commonly occurs at frequencies of hundreds or thousands of Hz, and successful prediction requires as input the interface properties appropriate to that frequency range. What has been shown in the earlier work and again here is that any connection between friction behaviour at high frequency and at zero frequency is a best tenuous. This is not merely a question of slight differences of parameter values: there seems to be a rather deep difference in the underlying physics between the two cases, leading to the strikingly better repeatability of the high-frequency behaviour.

The frictional frequency response, as measured here, thus has direct technological utility. However, it has also been shown here that it gives a powerful tool for scientific investigation of friction. The measurement give a new and rich source of information about the underlying constitutive model governing frictional behaviour. That information has been used here

to discriminate successfully between proposed models which differ only in relatively subtle ways. Of course, it may turn out that when other material combinations are tested they will not all follow the particular model which has worked well for the tests here, but the methodology set out here could be applied again in such cases to track down what additional ingredients an improved model might require.

The particular model presented here highlights the importance of contact stiffness to the high-frequency dynamics of an interface. However, it should be noted that the physical interpretation of this stiffness needs some care. In the experimental rig, the sensors for force and velocity are placed as close as possible to the sliding interface, but inevitably they cannot be exactly on the interface. The "contact stiffness" deduced by fitting to the measured results includes any physical effects that are relevant to the signals from those sensors. As well as "true" contact effects from, for example, the shear stiffness of the interacting surface asperities, there may be a contribution from bulk stiffness of whatever structure lies between the contact zone and the sensors. It might in principle be possible to discriminate between such contributions directly from measured data, but it has been shown that for the particular tests reported here no such discrimination is justified: the sensitivity to the relevant parameters is simply too low. Instead, an "Occam's razor" argument has been used, to select the most parsimonious model that is consistent with all the measurements within the limits set by their intrinsic noise.

Several directions of future work are suggested by the results discussed here. At a purely experimental level, there would be obvious benefits to testing a wide variety of materials in the experimental rig. This would help to map out how widely the particular model proposed here can be applied. In order to do such tests, it will probably be necessary to enhance the apparatus and procedures to get round the difficulties encountered when polycarbonate was tested: this measurement relies on remaining in the linearised regime, and if self-excited vibration is spontaneously generated by the rig, that assumption is immediately invalidated.

Another line of future investigation would involve developing the analysis and measurement methodology from localised frictional contact to the more important case of extended contact. This would be directly relevant to vehicle braking systems, and would also relate to tests frequently performed in experimental seismology, by the use of an array of strain gauges and accelerometers placed near a sliding interface.

If it does turn out that this model, or a close relative, has wide applicability, that raises further possibilities. If one were confident ahead of time of the form of the underlying constitutive model, the goal of testing would become the more restricted one of estimating parameter values. It is possible that a less elaborate measurement method could be developed to satisfy that goal: perhaps something that is more readily implemented using existing commercial tribological test equipment. Equally, there would be obvious interest in developing ways to incorporate this model into commercial software for squeal prediction and related vibration calculations: there seems to be no reason in principle why this model should not be implemented in a Finite Element context.

Acknowledgements

References

Adams, G., Nosonovsky, M., 2000. Contact modeling—forces. Tribol. Int. 33 (5), 431–442.
Aronov, V., D'Souza, A., Shareef, I., 1984. Interactions among friction, wear, and system stiffness—part 1: effect of normal load and system stiffness. Trans. ASME J. Tribol. 106 (1), 54–58.
Bar-Sinai, Y., Spatschek, R., Brener, E., Bouchbinder, E., 2014. On the velocity-strengthening behavior of dry friction. J. Geophys. Res. 119, 1738–1748.
Baumberger, T., Berthoud, P., Caroli, C., 1999. Physical analysis of the state- and rate-dependent friction law: dynamic friction. Phys. Rev. B 60 (6), 3928–3939.
Baumberger, T., Caroli, C., 2006. Solid friction from stick-slip down to pinning and aging. Adv. Phys. 55 (3–4), 279–348.
Berthoud, P., Baumberger, T., 1998. Shear stiffness of a solid-solid multicontact interface. Proc. R. Soc. Lond. A: Math. Phys. Eng. Sci. 454, 1615–1634.
Berthoud, P., Baumberger, T., G'Sell, S., Hiver, J.-M., 1999. Physical analysis of the state- and rate-dependent friction law: static friction. Phys. Rev. B 59 (22), 14313–14327.
Bizzarri, A., 2014. The mechanics of seismic faulting: recent advances and open issues. Riv. Del Nuovo Cim. 37, 181–271.
Blau, P.J., 2009. Friction Science and Tecnology: from Concepts to Applications. CRC Press, Taylor and Francis, Boca Raton, USA.
Bouchbinder, E., Brener, E., Barel, I., Urbakh, M., 2011. Slow cracklike dynamics at the onset of frictional sliding. Phys. Rev. Lett. 107 (November), 235501.
Bowden, F., Tabor, D., 1950. The Friction and Lubrication of Solids. Oxford University Press, New York, USA.
Brace, W., 1972. Laboratory studies of stick-slip and their application to earthquakes. Tectonophysics 14, 189–200.
Brace, W.F., Byerlee, J.D., 1966. Stick slip as a mechanism for earthquakes. Science 153, 990–992.
Bureau, L., Baumberger, T., Caroli, C., 2000. Shear response of a frictional interface to a normal load modulation. Phys. Rev. E 62, 6810–6820.
Bureau, L., Baumberger, T., Caroli, C., 2006. Non-Amontons behavior of friction in single contacts. Eur. Phys. J. E 19 (2), 163–169.
Bush, A., Gibson, R., Thomas, T., 1975. The elastic contact of a rough surface. Wear 35 (1), 87–111.
Butlin, T., Woodhouse, J., 2013. Friction-induced vibration: model development and comparison with large-scale experimental tests. J. Sound Vib. 332 (21), 5302–5321.
De Silva, C.W., 2007. Vibration: Fundamentals and Practice, 2nd edition. CRC Press, Taylor and Francis, Boca Raton, USA.
deWit, C., Olsson, H., Astrom, K., Lischinsky, P., 1995. A new model for control of systems with friction. IEEE Trans. Autom. Control 40 (3), 419–425.

Dieterich, J.H., 1979. Modeling rock friction: 1. Experimental results and constitutive equations. J. Geophys. Res. 84, 2161–2168.

Dieterich, J.H., Kilgore, B.D., 1994. Direct observation of frictional contacts: new insights for state-dependent properties. Pure Appl. Geophys. 143, 283–302.

Estrin, Y., Brechet, Y., 1996. On a model of frictional sliding. Pure Appl. Geophys. 147 (4), 745–762.

Greenwood, J.A., Williamson, J.B.P., 1966. Contact of nominally flat surfaces. Proc. R. Soc. Lond. A: Math. Phys. Eng. Sci. 295 (1442), 300–319.

Grosch, K., 1963. The relation between the frictional and visco-elastic properties of rubber. Proc. R. Soc. Lond. A 274, 21–39.

Gu, J.C., Rice, J.R., Ruina, A.L., Tse, S.T., 1984. Slip motion and stability of a single degree of freedom elastic system with rate and state dependent friction. J. Mech. Phys. Solids 32, 167–196.

Heslot, F., Baumberger, T., Perrin, B., Caroli, B., Caroli, C., 1994. Creep, stick-slip, and dry-friction dynamics: experiments and a heuristic model. Phys. Rev. E 49, 4973–4988.

Johnson, K.L., 1985. Contact Mechanics. Cambridge University Press, Cambridge, UK.

Kilgore, B.D., Blanpied, M.L., Dieterich, J.H., 1993. Velocity dependent friction of granite over a wide range of conditions. Geophys. Res. Lett. 20 (10), 903–906.

Linker, M., Dieterich, J., 1992. Effects of variable normal stress on rock friction: observations and constitutive equations. Journal. Geophys. Res. 97 (B4), 4932–4940.

Marone, C., 1998. Laboratory-derived friction laws and their application to seismic faulting. Annu. Rev. Earth Sci. 26, 643–696.

McConnel, K.G., Varoto, P.S., 2008. Vibration Testing: Theory and Practice, 2nd edition. Wiley, Hoboken, USA.

McCool, J.I., 1986. Comparison of models for the contact of rough surfaces. Wear 107 (1), 37–60.

Persson, B., 2001. Elastoplastic contact between randomly rough surfaces. Phys. Rev. Lett. 87, 116101.

Persson, B., Albohr, O., Mancosu, F., Peveri, V., Samoilov, V., Sivebaek, I., 2003. On the nature of the static friction, kinetic friction and creep. Wear 254, 835–851.

Persson, B.N.J., 2000. Sliding Friction: Physical Principles and Applications, 2nd edition. Nanoscience and Technology. Springer-Verlag, Berlin, Germany.

Putelat, T., Dawes, J.H.P., 2015. Steady and transient sliding under rate-and-state friction. J. Mech. Phys. Solids 78, 70–93.

Putelat, T., Dawes, J.H.P., Willis, J.R., 2010. Regimes of frictional sliding of a spring-block system. J. Mech. Phys. Solids 58 (1), 27–53.

Putelat, T., Dawes, J.H.P., Willis, J.R., 2011. On the microphysical foundations of rate-and-state friction. J. Mech. Phys. Solids 59 (5), 1062–1075.

Rabinowicz, E., 1951. The nature of the static and kinetic coefficients of friction. J. Appl. Phys. 22 (11), 1373–1379.

Rabinowicz, E., 1957. The intrinsic variables affecting the stick-slip process. Proc. Phys. Soc. (London) 71, 668–675.

Rice, J., 1993. Spatio-temporal complexity of slip on a fault. J. Geophys. Res.: Solid Earth 98, 9885–9907.

Rice, J.R., Lapusta, N., Ranjith, K., 2001. Rate and state dependent friction and the stability of sliding between elastically deformable solids. J. Mech. Phys. Solids 49, 1865–198.

Ruina, A., 1980. Friction Laws and Instabilities: a Quasistatic Analysis of Some Dry Frictional Behavior (Ph.D. thesis), Brown University, Providence, USA.

Ruina, A.L., 1983. Slip instability and state variable friction laws. J. Geophys. Res. 88, 10,359–10,370.

Scheibert, J., Prevost, A., Frelat, J., Rey, P., Debrégeas, G., 2008. Experimental evidence of non-Amontons behaviour at a multi-contact interface. EPL (Eur. Lett.) 83, 34003.

Scholz, C.H., 1998. Earthquakes and friction laws. Nature 391, 37–41.

Scholz, C.H., 2002. The Mechanics of Earthquakes and Faulting, 2nd edition. Cambridge University Press, Cambridge, UK.

Sheng, G., 2008. Friction-induced Vibrations and Sound: Principles and Applications. CRC Press, Taylor and Francis, Boca Raton, USA.

Shimamoto, T., 1986. Transition between frictional slip and ductile flow for halite shear zones at room temperature. Science 231, 711–714.

Stesky, R., Brace, W., Riley, D., Robin, P.-Y., 1974. Friction in faulted rock at high temperature and pressure. Tectonophysics 23, 177–203.

Tsutsumi, A., Shimamoto, T., 1997. High-velocity frictional properties of gabbro. Geophys. Res. Lett. 24, 699–702.

Wang, S., Woodhouse, J., 2011. The frequency response of dynamic friction: a new view of sliding interfaces. J. Mech. Phys. Solids 59 (5), 1020–1036.

Weeks, J.D., 1993. Constitutive laws for high-velocity frictional sliding and their influence on stress drop during unstable slip. J. Geophys. Res. : Solid Earth 98 (B10), 17637–17648.

Wei, W.W.S., 2006. Time Series Analysis: Univariate and Multivariate Methods, 2nd edition. Addison Wesley, Boston, USA.

Woodhouse, J., Putelat, T., MacKay, A., 2015. Are there reliable models of friction? Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci. 373, 2051.

Woodhouse, J., Wang, S., 2011. The frequency response of dynamic friction: model comparisons. J. Mech. Phys. Solids 59 (11), 2294–2306.