INVESTIGATION OF CONTACT OF METROLOGICAL PROBE TIPS WITH ROUGH ENGINEERING SURFACES

by

Ismail M. R. I. Najjar

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Engineering

University of Warwick
School of Engineering

November 2002

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENT

I would like to take this opportunity to express my sincere thanks to my supervisor, Prof. D. G. Chetwynd, for his support and encouragement during the period of the research and his suggestions and comments on this piece of work, and also for his help in correcting the thesis.

Many thanks go to the people in the Centre for Nanotechnology and Microengineering, specially Dave Robinson and Steve Wallace, for their general help to progress my experimental work.

I would also like to express my gratitude to the Ministry of Higher Education in Saudi Arabia for the financial support of my study.

DECLARATION

This thesis is presented in accordance with the regulations for the degree of Doctor of Philosophy at the University of Warwick. The thesis has been composed and written by myself based on the research undertaken by myself. This research has not been submitted in any previous application for a higher degree. All sources of information used are specifically acknowledged at the relevant points in the text.

I. Najjar

SUMMARY

This investigation explores the behaviour of contact interaction between probe tips of indicators (displacement gauges) and rough engineering surfaces, in order to gain new insights into uncertainty in industrial precision gauging. The conditions of interest are contact of a mm-scale hard sphere at 0.1 to 5 N force with metallic surfaces, $R_q < 2\,\mu m$. The motivation arises from the growing demand of modern industries for higher accuracies in machining processes and for finer inspection capabilities. In the near future, such a demand is likely to require careful analyses of the measurement error budget of these instruments with a greater consideration of the potential errors induced by the complex behaviour of contact, which is currently assumed to be negligible.

A preliminary experimental study was carried out on an existing test-rig of limited accuracy. It raised several concerns related to the effects of surface roughness and material on the resulting contact deflection at the load regimes of indicators' probes. It stimulated the need for more reliable data. So, a new test-rig of advanced capabilities was designed to perform a comprehensive study of a wider range of contact parameters to simulate real situations of industrial gauging processes. This advanced study confirms the consistent effects of roughness, material and probe tip size not only on the surface deflections but also on the probe rotations. Roughness variability across the same surface caused some inconsistencies in the deflection behaviour and repeatability, and surface contamination gave additional unsystematic effects on these relations. Re-establishing contact on the same position of a contaminated surface appeared to contribute errors of the order of those of a clean surface.

Seeking a quick and easy tool for predicting displacement errors contributing to measurement uncertainty, approximations are proposed in two contact models. The contact conditions of probe tips are unlike those commonly studied by tribologists. The experimental results show that both approximations can provide conservative estimates of surface deformation at loads up to 2 N and roughness below 0.5 µm.

Based on the results of this investigation, systematic errors of contact can exceed the 1 µm level with uncertainty up to at least 0.4 µm in many measurement processes with precision indicators. Such figures reveal that the contact cannot any longer be considered a negligible source of errors, and precision metrology must account for these induced errors in the error budget of displacement gauges.

CHAPTER 1

INTRODUCTION

1.1 DEVELOPMENT OF PRECISION MANUFACTURING

Manufacturing to higher precision is a development that has been gathering momentum over the last two-hundred years and accelerating over the last two decades in terms of research, development, product innovation, and investment [1]. It has been driven by the increasing demand for much higher performance of products, higher reliability, longer wear/fatigue life, and greater miniaturization and packing densities. Eliminating "fitting", promoting "assembly", and improving interchangeability of components are other general motives for such development. In addition, this demand embraces the improvement of quality control processes through higher machine accuracy capabilities and, hence, reducing scrap, rework, and conventional inspection procedures [2]. The most widely used categories for the "manufacture with higher precision" are: precision engineering, micro-engineering, and nanotechnology. *Precision engineering* may be defined as manufacturing to tolerances smaller than one part in 10^4 or perhaps one part in 10^5, whereas *micro-engineering* is where the physical dimensions of the component or features are small, namely in the order of 1 μm. *Nanotechnology* is a term coined by Norio Taniguchi in 1974 to describe the manufacture to finishes and tolerances in the nanometre region (usually quoted as less than 100 nm).

The historical root of precision engineering could be said to be horology, the development of chronometers and watches, and, of course, optics (e.g., the manufacture

of mirrors and lenses for telescopes and microscopes). Major contributions were made to the development of high precision machine tools and instruments in the late 1800s and early 1900s by ruling engines for the manufacture of scales, reticules, and spectrographic diffraction gratings. Today, ultra precision machine tools under computer control using single-point or multi-point diamond grinding wheels can position the tool relative to the workpiece to a resolution and positioning accuracy in the order of 1 nm [3]. However, publications such as Evans [4] provide more details about the history and development of precision engineering industries.

1.2 TANIGUCHI'S HISTORICAL PROGRESS OF MACHINING ACCURACY

The historical development of the accuracy of material processing over the past century has been traced by Taniguchi from various sources on processing technologies. The result was the plot shown in Figure 1.1 which shows the machine tools, processing devices, and measuring instruments/inspection devices together with the corresponding positioning accuracies and measuring resolutions. This plot appeared in different forms and in several of his publications such as [5 and 6]. Based on this survey, Taniguchi extrapolated the specifications from existing and past machine tools, such as lathes and grinders, to the new generation of machine tools. He concluded quite correctly that in the late 1980s and 1990s accuracies between 0.1 μm and 1 nm would be needed to cater for industries' needs [7]. He also concluded that the attainable processing accuracy could be expected to reach the order of nanometres early in the twenty-first century. From such estimations, he aimed to strongly emphasize the urgent need to develop nanotechnology to improve processing accuracies for many recent products.

The predictions postulated by Taniguchi have been updated by several workers in

his field; one of them was his noted colleague, Pat McKeown, who modified the extrapolations up to the year 2000 as illustrated in Figure 1.2 [3]. McKeown also provided an update on the machine tools and equipment accuracies in his new plot. When comparing this updated plot with Figure 1.1, it seems that Taniguchi's predictions between 1980 and 2000 were also correct.

Whitehouse stated in [7] that it emerged that the only way to achieve Taniguchi's results was to incorporate very sophisticated instrumentation and metrology into the design of the machines. He illustrated that nanotechnology expanded several disciplines within each science and nanotechnology in, for example, engineering encompasses metrology, instrumentation, function, and manufacture. Moreover, Taniguchi pointed out in [6] that it is evident (from his survey and from the tolerances or functionally allowable errors of various kinds of products and intelligent devices with ultra-high precision and ultra-fine structure reported in several publications such as [5, 6, and 8]) that nanotechnology includes not only extra-high precision processing but also measuring and positioning technologies with sub-nanometre resolution and scattering error, respectively. Accordingly, nanotechnology requires the development of integrated systems of materials processing, dimensional measurement, and positional control technologies which achieve ultra-high precision with nanometre accuracies, sub-nanometre resolution and in-process feedback and feed-forward networks. Hence, from these two authorities, it is very apparent that the ever-increasing precision of manufacturing processing necessitated an ever-increasing development in metrology together with the other disciplines needed for the progress of the whole technology. It is, then, conceivable that tracing the development of accuracies (plus resolution, precision, etc.) of metrology instruments over the past century would lead to trends similar in shape to those introduced by Taniguchi for the development of machining accuracies. Extrapolation of these new trends could also be employed in order to obtain

future estimations that may be useful in determining the potential for continued use of many conventional measurement/inspection tools and techniques.

1.3 THE FUTURE OF CONTACT-GAUGING OF SURFACES WITH INDICATORS

History gives us every reason to expect greater demands for precision during the next decade. Since the 1940s, typical machining tolerances have become about 50% tighter every ten years, continually shifting the accepted meaning of "high precision" – from 0.05 mm in the 1940s to 0.02 mm in 1950s, all the way to 1 μm by 2000. Some metal cutting industries already are working to 0.5 μm tolerances, and sub-micrometre tolerances may be considered normal before long [9]. Moreover, depending on Taniguchi's extrapolations of the achievable machining accuracies shown in Figure 1.1, normal machining processes will be capable of producing components of tolerances down to 0.1 μm by about year 2030. Accordingly, a development in precision inspection tools conventionally used in precision machine-shops and quality control checkpoints (e.g., co-ordinate measuring machines and indicators) is confidently expected in order to fulfil such increasing requirements. There is a widely-used guideline that the process capability at each step of a traceability (calibration) chain should be ten times better than the previous step. Although, in practice, technological and economic constraints often lead to agreement to use a lower figure, there may be demands for uncertainties of around 100 nm, or smaller, from such instruments within the next fifteen to twenty years. Well before that, the major individual contributory factors to the uncertainty budget will need to be controlled at the 100 nm level.

Contact measurement and inspection remain very popular processes for the conventional dimensional control of the production of mechanical components. They

provide appreciably robust, cheaper single-point control, and more tolerant of some types of contamination with reasonable accuracies compared to the non-contact dimensional inspection processes. So, it seems very unlikely that the demands for them will show a notable decrease, at least over the coming decades. An increasing demand for improving the accuracy performance of contact inspection instruments seems inevitable. This implies that advanced mechanical and electronic systems are to be implemented in their designs.

In recent years, many of the manufacturers of contact probe instruments have produced some sophisticated models with advanced design technologies and improved measuring accuracy and precision in order to cope with the growing needs of industry. To give a typical example, digital electronic indicators appeared as a result of the limited measuring performance of the mechanical dial indicators at the level of high precise measurements. In regard to this, the authors in [9] (who are managers at Federal Products Co., a leading USA manufacturer of precision dimensional gauging products) stated that most digital indicators already offer 1 μm resolution, and 0.5 μm resolution is readily available. In contrast, few mechanical dial indicators resolve finer than 1 μm. Digital gauging amplifiers offer 0.1 μm resolution, and some transducers are capable of such accuracy. They added that improvements in transducer technology will allow gauges to combine high resolution accuracy with long measurement range. Digital indicators with resolution of 1 μm and ranges of 12.7 mm or 25.4 mm are now on the market, and the authors expect these ranges to double or triple and resolution to improve to 0.5 μm or less in simple indicators over the next few years.

The authors also pointed out [9] that as dimensional tolerances become tighter, surface finish and geometric variations represent a larger proportion of the total allowable part variation, and their measurement becomes increasingly important. Engineers are learning more about how these variations influence functionality, and

tolerances are being specified more frequently. They added that as production requirements demand more complex measurements – especially surface finish and geometry – on the shop floor, gauge makers will respond with deceptively simple but actually sophisticated gauges that do more work for the user. Increased data storage and processing power will be built into gauges, and more gauging functions will be built into computers. Computer-aided gauges will guide procedures and setups, establish datum and compensate or account for geometric and environmental variation. They concluded that the future of "traditional" contact-style dimensional gauging will be largely a process of "digitalizing" and computerizing many tasks that currently are performed with mechanical instruments. As this occurs, some measurements will become more precise. Many routine inspection tasks will become easier and/or quicker to perform, while some highly complex tasks will become routine. Gauges will be integrated increasingly into feedback-controlled manufacturing processes and into company-wide networks. Overall, gauging will play an even more important role than it does today in quality-oriented product development and manufacturing.

This simple example reveals in general terms the future of contact-gauging processes. It emphasizes the essential development of conventional contact-measuring instruments to accommodate the growth in precision of many technological industries: electronic indicators are the coming era of contact-inspection tools at the level of high precision machine shops and QC checkpoints. Moreover, the authors gave a clue to the importance of surface finish in the contact-gauging processes as smaller dimensional errors are allowed, also addressed in an earlier article [10] on new trends in gauging. It stated that a new quality issue has arisen as dimensional tolerances have decreased: factors of part geometry and surface roughness, once so subtle that they could previously be ignored, are becoming increasingly important. So, it seems from these two articles that instrument manufacturers are already concerned about the significant

effect of surface quality on the errors contribution to the contact-measurement with displacement gauges. Manufacturers rarely provide information beyond the features of their products; they sometimes assist users through general guidelines for the optimum option to accomplish the required measurements within the design accuracies of the selected instruments. Users' full awareness of this effect is questionable as manufacturers' recommendations do not usually go that deeply into details.

1.4 SCOPE OF THESIS

Given the increasing demand for tighter tolerances, modern manufacturing industries increasingly need to control workpiece dimensions or positions to micrometre tolerances. In-process, in-line, or off-line inspection adjacent to machine tools is then needed. The factory-floor environment is poor for the precision metrology needed to maintain production control. Hence, there is a preference for robust displacement gauging systems that make quite heavy mechanical contact with the surfaces of the workpieces. Such gauges might be touch-trigger probes (as used on CMM's) or linear dial and digital indicators with side-acting or plunger probes. Typical indicators have contact tips of hard spheres of few millimetres diameter held against the surface by a spring-generated force in the Newton range. Touch-trigger probes have smaller static forces but greater dynamic effects.

At the expected measurement capability of around 100 nm under normal operating conditions within the next one or two decades, a greater understanding of the contact behaviour of probe tips onto engineering surfaces will be essential. The physical contact interaction is far from simple; such surfaces have some roughness and generally contaminated by dust and machining debris particles and oil and aqueous films. Therefore, it is doubtful whether it can any longer be assumed that the contact

itself is a negligible error source either for repeatability or for systematic effects. Currently, there is a little relevant published data: the contact regime is unlike those generally of interest to tribologists.

Consequently, this study has been carried out to explore the nature and magnitude of uncertainties in precision measurements with indicators resulting from the contact conditions on engineering surfaces. Different possible effects of the parameters involved in the contact interaction were investigated in an attempt to estimate the gauging precision beyond which such uncertainties must be taken into consideration in the error budget of displacement contact-gauges. Obviously, this requires a detailed analysis of the behaviour of surface deflections resulting from gauging forces of indicators and, especially, of their repeatability under the typical variability of contact effects that might be encountered in normal inspection processes. This was achieved through an experimental programme of two stages.

Before reporting on the experimental programme, a concise theoretical survey of some common mechanical dial and digital electronic gauge indicators is introduced in Chapter 2. Potential deficiencies in the performance of these linear displacement gauges are also discussed. This chapter also reports on tests carried out on some of the precise indicators available in the Warwick Precision Engineering Laboratory.

Chapter 3 gives an overview of the mathematical modelling of contact between rough surfaces. It discusses several factors which have been of concern to many tribologists about the formulation of these models. Two approaches were approximated and their results were compared with experimental ones in order to suggest a simplified theoretical model that reasonably represents the real contact configuration between the indicator's probe tip of and the rough surface. The idea was to explore whether useful accuracy could be obtained from a formula sufficiently simple that it could be applied by practicing metrologists, who will not be well-versed in contact theory.

A preliminary experimental investigation was performed by renovating and using an existing test-rig of limited capabilities in order to provide initial data that would guide the design of the more detailed investigations. Chapter 4 demonstrates the full characterisation of this test-rig and discusses the results of those initial experiments.

A special-purpose test-rig was then designed and manufactured for evaluating the uncertainties in the position ("height") of rough surfaces reported by contact probes under conditions typical of industrial gauging. Its design incorporates features to address the most serious experimental uncertainties encountered in the preliminary study and so has much improved accuracy, with a sensitivity that allows contact height repeatability to be compared to around 10 nm. The design principles, characterisation, and specifications of this new test-rig are given in Chapter 5.

Chapter 6 presents the second stage of the experimental programme which was accomplished on the advanced test-rig. Comprehensive analyses of surface deflection behaviour under the variation of contact parameters are discussed in this chapter. Two types of repeatability were studied; same-point and different-point repeated contacts. In practice, the former type could represent the case when the probe is released after initial contact(s) and then brought again into contact on nominally the same point on the surface for measurement. The majority of indicator probes apply less force at their outward movement and some users might prefer to minimize contact force at the instant of recording by adjusting the probe after initial discovery of the contact condition. The latter type could represent a case when several measurements are made to establish a feature (such as a plane) or determine a height difference. However, it is also the principal means for studying the expected uncertainty of single point measurements.

Chapter 7 includes the final discussions and the main conclusions of the research. The main contributions to knowledge and recommendations for future work are also discussed in this chapter.

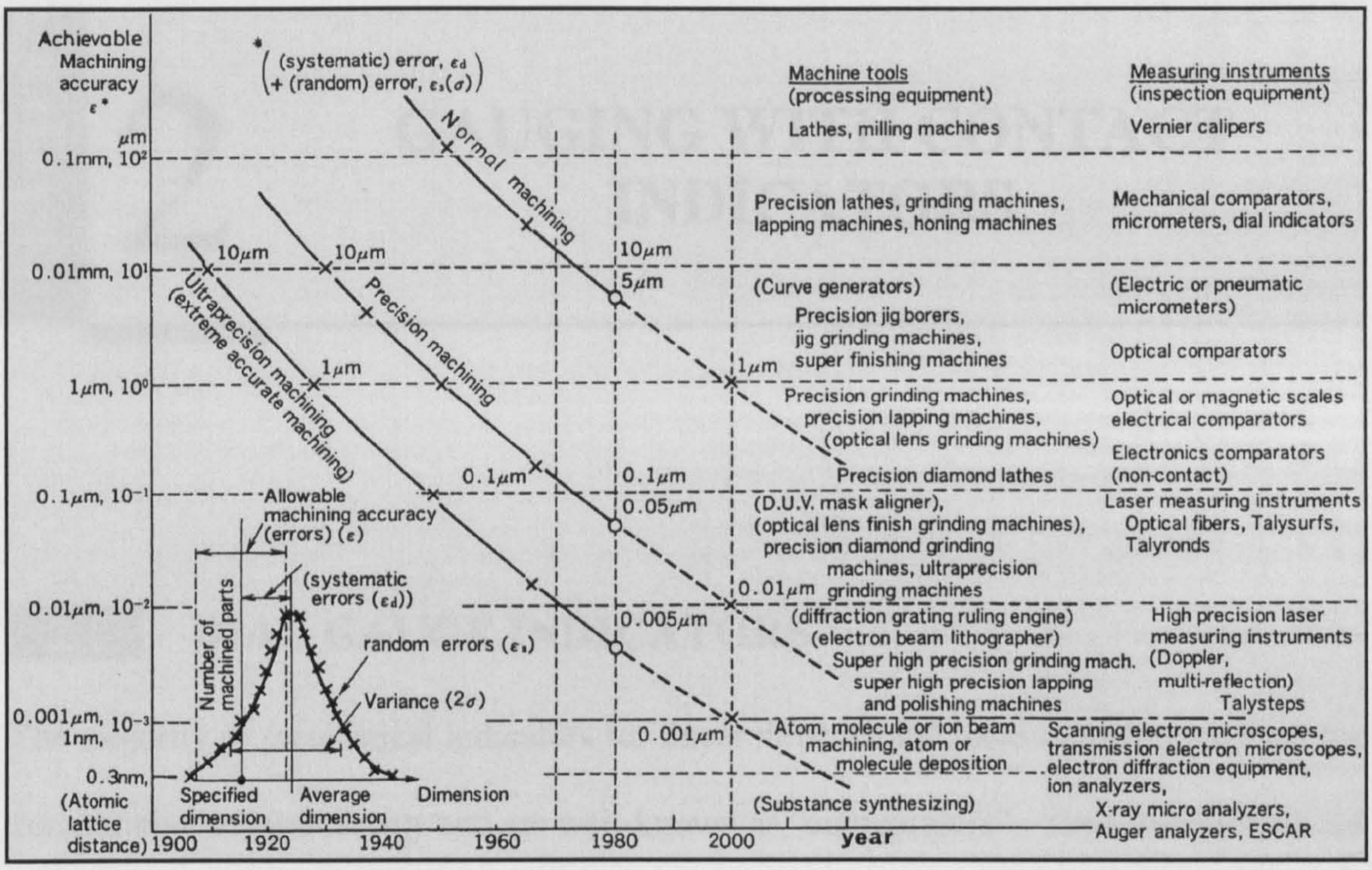


Figure 1.1 Taniguchi's plot of the historical progress of achievable machining accuracy over the past century [5].

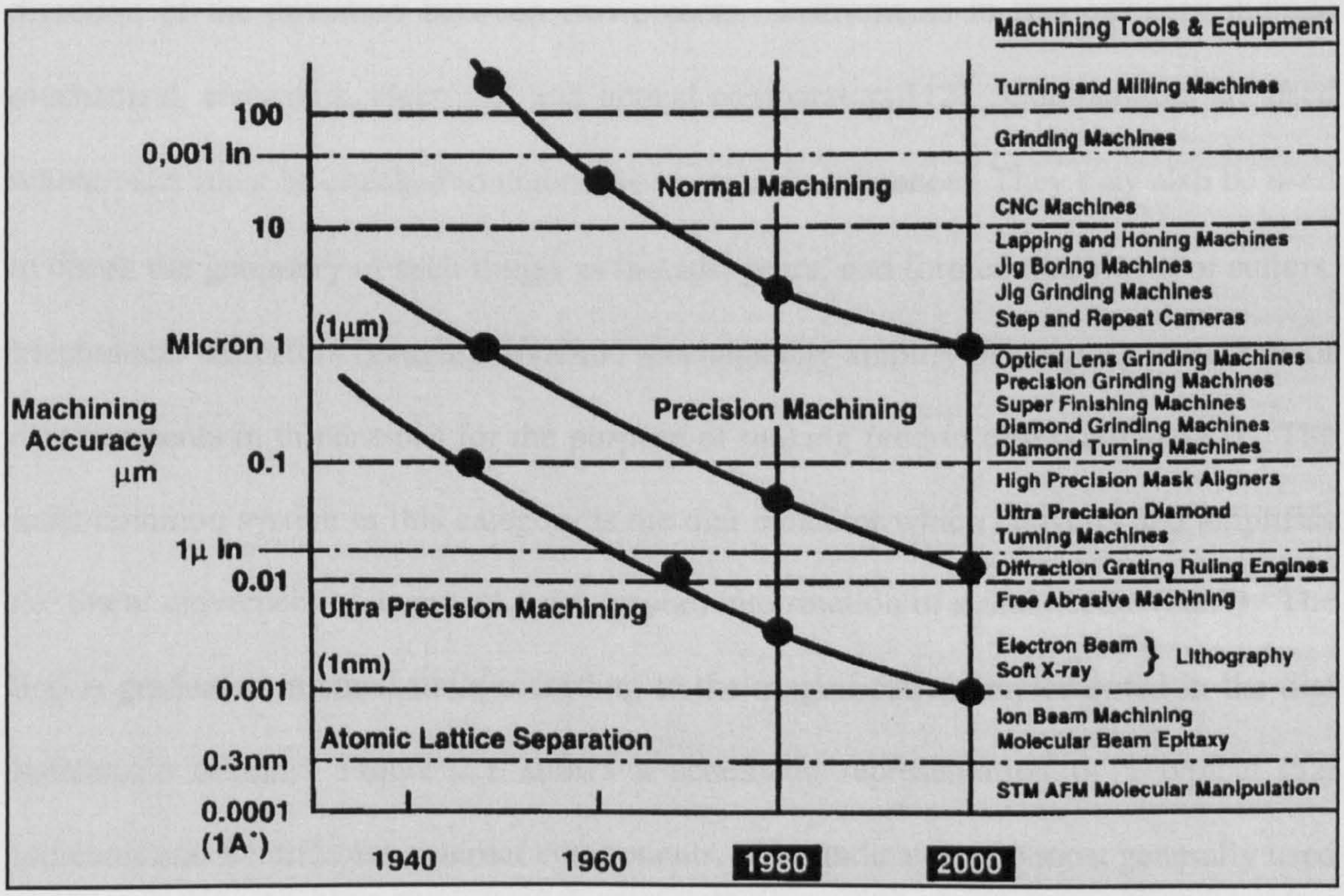


Figure 1.2 McKeown's update (after Taniguchi) of the development of achievable machining accuracy over the last sixty years [3].

CHAPTER 2

GAUGING WITH CONTACT INDICATORS

2.1 DIAL GAUGE INDICATORS

The majority of mechanical indicators for linear dimensional measurements are used for comparative measurements and are well-known as "comparators". Such instruments are used to make dimensional comparisons between two objects, such as a workpiece and a reference surface [11]. They are usually not capable of providing an absolute measurement of the quantity of interest; instead, they measure the magnitude and direction of the deviation between two objects. Instruments in this category include mechanical, electronic, electrical, and optical comparators [12]. Comparators are used where parts must be checked to determine acceptable tolerance. They may also be used to check the geometry of such things as threads, gears, and formed machine tool cutters. Mechanical indicators (gauging systems) mechanically amplify or magnify variations or displacements in dimensions for the purpose of making precise observations [13]. The most common system in this category is the dial indicator which converts and amplifies the linear movement of a contact point (probe) into rotation of a dial needle (hand). The dial is graduated in small units according to the magnification implemented in the dial indicator's design. Figure 2.1 shows a schematic representation of a typical dial indicator and its different external components. Dial indicators are most generally used because their magnification accuracy meets the large majority of requirements. Magnification is obtained by means of a gear-train.

In operation, a sensitive contact or point is attached to a rack gear. A train of three to five gears, depending on the magnification desired, magnifies and transmits the movement of the contact to the pinion gear on which the indicator hand is mounted with a hairspring and take-up gear to eliminate backlash. Figure 2.2 illustrates some internal components of a dial indicator. Some indicators are designed with a spring-loaded link at some part of the gear-train to absorb the impact of a sudden shock and protect the indicator against damage. The pressure of the sensitive contact is established by a pullback spring which exerts a downward pressure on the rack spindle, thus keeping the contact constantly in touch with the workpiece. There must always be sufficient pressure to assure positive contact and to overcome the internal friction of the indicator itself. There is normally a sliding seal to prevent entrainment of dirt which tends to introduce significant friction forces.

The test indicator, a typical one shown in Figure 2.3, is a widely used type of dial indicator. It is frequently quite small so that it can be used to indicate in locations inaccessible to other indicators. The magnification is obtained by gears and levers. Its spindle or tip can be swivelled to any desired position. Test indicators are usually equipped with a movement reversing lever which means that the indicator can be actuated by pressure from either side of the tip, so the instrument need not be turned around. They, like dial indicators, have a rotating bezel for zero setting. But their contact tips are usually smaller in diameter and interchangeable with the levers. Their dial faces are generally balanced-graduated.

The dial indicators are perhaps the most widely used instrument for moderately precise measurement [14]. They provide a quick, economical, reliable method of dimensional measurement and have been around since the 1800s and are expected to find extensive, continued application well through this millennium [15]. In many industrial countries, dial indicator specifications for design features, geometry, and

performance have been regulated through general recommended guidelines (or standards), such as the British Standards BS 907, for the dial gauges [16], and BS 2795, for the dial test indicators [17]. On the other hand, dial indicator manufacturers have established standard specifications, known as American Gauge Design (AGD) specifications, which also assist the users with reliable bases for selection criteria in order to achieve optimum measurement performance and better indicator utilisation. Standard series in these specifications vary greatly in size, amplification ratio, mounting facilities, and precision. However, due to the importance of gauging tools accuracy and productivity, other more specific guidelines for this selection have also been offered by some common commercial sources of dial indicators, like Mahr-*FEDERAL* [18] and Mitutoyo [19], and published by several authors such as [15, 20, 21 and 22].

2.2 ELECTRONIC GAUGE INDICATORS

Digital indicators also measure linear dimensions, but work different than dial indicators [15]. Figure 2.4 shows a typical standard digital indicator. Also, they typically cost more than dial indicators, and the advantages and disadvantages of each need to be weighed when choosing between them. Most people consider reading a number on a digital indicator easier than counting graduations on a dial indicator. Also, unlike the dial indicator, which must be purchased as either an inch or metric tool, an electronic indicator converts between the two measurement systems. The strength of the dial indicator is its ability to instantly show the high or the low point of its readings. Experienced users frequently prefer the dial indicator because they can easily see where the needle stops and then reverses direction. Many feel that it is easier to do sweep measurement with a dial than with a digital indicator. The motion of the indicating needle is a visual aid for sweeping surfaces. On the other hand, the inattentive operator

of the dial indicator commonly fails to notice that the needle has made a complete revolution of the dial and returned to zero [23]. The part may be grossly out of tolerance, but to this operator, the needle appears to be exactly where it should be. Most also agree that it is more difficult to view numbers and quickly determine which was the lowest (or highest) number. However, the greatest advantage of the digital indicators is their ability to output data automatically for statistical process control. They may also provide much better accuracy than the dial indicators in the long run [21].

Digital indicators are a member of the electronic gauges family: measuring and gauging instruments based on transducers capable of converting a linear displacement into an electrical signal (changes of electric current or voltage). The electrical signal is then amplified and transformed into a suitable data format such as a digital readout. The types of transducers used in electronic gauges include linear variable differential transformer (LVDT), strain gauges, inductance bridge, variable capacitors, and piezoelectric crystals; also some with optical gratings. The transducer is contained in a gauging head designed for the application.

One of the early designs of electronic gauging systems based on the use of the LVDT was described in [24]. A basic single-channel electronic measuring system consists of a transducer (gauge head), signal conditioning and processing circuits, and an appropriate readout. The transducer element is a LVDT, an electromechanical device that produces an electrical output proportional to the displacement of a separate movable core. By mechanically coupling the transformer core to a simple spring-biased probe shaft supported on precision bearings, displacement of the gauge head is converted to an electrical signal that is proportional to probe movement. There are no cams, linkages, contacts, wiping elements or similar moving parts to affect the inherent accuracy, repeatability and infinite resolution of the LVDT.

In recent sophisticated designs of these indicators, such as series 542 linear

gauges from Mitutoyo [25], the spindle, which carries the contact point from one side and attached to an ultra-precise optical scale from the other side, travels through a precise linear ball bearing cylinder. Near the contact point, the spindle is also attached into a sort of diaphragm for dust-, splash-, and oil-proofing. A transmission-type photoelectric detecting unit senses the movements of the graduated scale and transmits electrical signals to circuitry for signal processing. The output data for this gauge are in a digital format which can be processed further on external devices such as digital displays and data acquisition systems.

Such an example of an advanced digital indicator shows the rapid growth of the applications of electronic gauges in recent years. They are gradually replacing many of the conventional measuring and gauging devices. Advantages of electronic gauges include (a) good sensitivity, accuracy, precision, repeatability, and speed of response; (b) ability to sense very small dimensions, down to 0.025 μm; (c) ease of operation; (d) reduced human error; (e) display of electrical signal in various formats; and (f) capability to be interfaced with computer systems for data processing. Moreover, they are applied in absolute measurement in addition to their role as comparative measuring instruments, such as modern micrometers and graduated callipers which are available with electronic devices that display a digital readout of the measurement.

2.3 POTENTIAL ERRORS IN INDICATORS

Indicator gauges are delicate instruments and the amount of satisfactory service that they will give depends to a very large extent on the way they are used [26]. In general, the higher the magnification the more delicate they become. In dial indicators, as in any mechanical device, friction, dimensional tolerances between parts and wear may cause eventual loss of sensitivity and accuracy [27]. In addition, motion of sensing element

leads to problems with hysteresis, non-linearity and temperature variation.

In normal operation, potential sources of error from the internal mechanism of the dial indicator include: (a) the gear teeth mesh with some clearance, causing backlash and lagging response, particularly when wear progresses, (b) the effect of imperfect gear form, or play, results in cumulative error, which is usually proportional to the length of the gauging travel, (c) a clearance of the measuring spindle in its guides is needed to prevent binding and to provide an unimpeded operation movement; this clearance can result in positional variations of the rack in relation to the meshing pinion, (d) play in the pivot bearings due to initial inaccuracies or wear affects the precise meshing of the gear teeth, causing measuring errors [28]. Hence, such errors mainly cause two functional deficiencies, which are (1) lower accuracy if it is used over the whole range of travel and (2) slight variations between forward-moving-spindle (or upward) and backward-moving-spindle (or downward) readings. The first shortcoming can be avoided by keeping the difference between the two readings to be compared as small as possible. While the second one is probably not possible to correct, such variation in the readings is not serious for most current applications and the indicator can still be used. The electronic indicator may also suffer from the second deficiency which could be as a result of the error source in (c) above. Another primary source of errors is the temperature variation. Most electronic indicators will operate properly at temperatures between 10 °C and 45 °C; however, this temperature range is sometimes violated in shop settings where indicator gauges are used [28].

2.3.1 Effects of Spring Rate and Friction

The force of the pullback spring, which is needed to keep the contact tip constantly in touch with the workpiece, is not uniform along the range of the spindle travel, as with the behaviour of any spring mechanism. The spring force is designed not only to

maintain contact but also to overcome the total friction between the different moving components (e.g. at gears' joints) and between the dirt-extracting elements (wiping seals) and the spindle. Thus, the variation of the resultant force exerted on the spindle is not linear with its displacement, as will be seen in the next section. Additionally, such force causes bending and compressive strains in the metrology loop of which the joints, gear-train, spring, and spindle are parts.

The determination of the magnitude and variability of these strains and their contribution to the error budget of the indicator seems to attract little interest from researchers or industry, probably, because the calibration processes have been thought sufficient for providing the required final measuring accuracy of the indicator. There has been some recent interest in advanced techniques for the calibration of dial indicators, such as [29 and 30] which highlighted new insights into the sources of error of these indicators. A setup has been constructed [29] that comprises a laser interferometer for measuring the displacement of the dial indicator rod and a special non-contact angular transducer for reading the dial indicator hand position. It is able to compute displacement errors for forward and backward movements, displacement error amplitude, fidelity error, and hysteresis error of the dial indicator. Automatic machine vision-based systems have been used [30] to extend the calibration of the dial indicators by checking several hundreds of points on its scale in order to get a complete picture of the errors. A method has been described for the calculation of the different components of measurement uncertainty of a 0.01 mm graduated dial indicator.

2.3.2 Thermal Effects

Another very important source of mechanical errors in such indicators is the thermal effects. Most engineering metals expand at a rate in the range 11 to 22 µm/m/°C, for steel and aluminium, respectively [31]. So, typical workshop temperature variations (of

±5 °C or more) can have a significant effect on the accuracy of these instruments, particularly if the rate of change is significant (greater than 3 °C/hour). To give an indication of this effect, for a steel spindle of an indicator of 100 mm long (with a typical thermal expansion coefficient of 12×10^{-6} °C^{-1}) the change in length for a temperature change of 5 °C is about 6 µm [32].

In the same way, the thermal expansion in gears diameters and their centre distances can be determined. These expansions (and probably more from other components) interact with each other to result a variable backlash error with temperature and, in turn, affect the indicator's measuring accuracy. Therefore, continuous inspection tasks using high-accuracy indicators are supposed to take place in a temperature controlled environment to minimise these thermal effects. It is worth noting here that temperature variations affect workpieces, as well. Thus, it is important to allow parts to be measured to come to thermal equilibrium in the measurement room before they are inspected.

2.3.3 Effect of Interchanging Contact Tips

Manufacturers of dial and electronic gauge indicators commonly offer the feature of interchangeable contact tips on them, for purposes such as maintenance, application, etc. The contact tip is usually mounted at the free end of the indicator's spindle by means of a screw thread. This means that an additional source of uncertainty is created in the metrology loop. In other words, the compliance of the tip mount (screw threads interaction) is also expected to contribute displacement errors to the indicator readings. These errors could introduce a gross adverse effect if, for example, an unstable mounting (insufficient tightening) of the contact tip is accidentally attained. To highlight the importance of such effect on measurement, Chapter 4 briefly reports some experiments on the effect of contact probe tightening on the repeatability of

measurement. It is considered one of the genuine characteristics inherent in indicators as the above sources of error.

The electronic indicator also suffers from all of the above sources of error because of the mechanical system that couples the electronic transducer to the workpiece. Although, this mechanical system is of shorter span within the metrology loop than that of the dial indicator, errors due to strain, friction, and thermal effects are still inherent in it. Moreover, the electronic transducer itself has inherent errors such as linearity, electronic noise, and digitisation; and can also be vulnerable to environmental noise and thermal effects. With the exception of non-linearity, these potential errors of a well-designed transducer are likely to have considerably smaller effects on measurement compared to those of the mechanical system.

2.4 PROBING FORCES OF INDICATORS

The spring-generated force, that pushes the indicator's probe rod towards the maximum extension, can show a significant variation along the range of the rod travel because of the spring rate effect. As the rod moves, this force resists the friction from the gears mechanism and its joints, bearings, and wiper. This causes the output force (measuring force) at the indicator's rod to vary (non-linearly) as the rod displacement increases. Thus, the magnitude of the force of contact exerted on the workpiece depends on the position of the rod. In addition, it depends primarily on the last direction of movement of the rod before accomplishing the contact with the surface. These effects on the measuring force will be described in detail in the following subsections.

2.4.1 Measuring Force Data Provided by Manufacturers

The variety of needs of indicators' users has lead manufacturers not only to produce a

variety of indicators of different accuracies and travel ranges, but of different measuring forces, as well. Gauging highly compliant surfaces needs indicators of low measuring forces in order to attain the required accuracy and stability of measurement. Most of the manufacturers of dial gauge indicators provide approximate, brief figures about the magnitude of the average or maximum measuring forces in their product publications. A few of them give more attention to this issue and show, especially, the users of the electronic gauge indicators further details, such as the measuring forces in different attitudes of operation (upward, downward, and horizontal) and, sometimes, the hysteresis of either the force or displacement. Tables 2.1 and 2.2 give typical values of the measuring forces of some models of dial and electronic gauge indicators, respectively, which are listed in product catalogues by some common manufacturers.

Indicators designed with pullback spring mechanisms, such as all of the dial gauge and most of the electronic gauge indicators, have the normal spring rate effect which introduces a distinctive and systematic variations of spring response with displacement. So, the probe rod of these indicators usually presents its maximum gauging force when it is nearly in full retraction, i.e., around the end of its travel inside the gauge enclosure. Since, the rod movements in dial indicators is governed by the rack and gear-train movements in addition to the friction and spring forces (generating a sort of "self-locking" mechanism), the Earth's gravitational forces are usually considered of negligible effect on the gauging forces. Dial indicators are commonly used in any direction (vertically, horizontally, upside-down, etc.) and their manufacturers do not associate their measuring forces with any specific attitude of operation. They normally provide, in this case, one value only for these forces, see Table 2.1, which is sometimes unclear whether it represents the average or the maximum gauging force.

Some manufacturers specify fuller information for only those types of electronic

indicators which are meant in their designs to offer the reduced gauging pressure feature. The probe rods of electronic indicators are generally ballbush- or slide-guided and secured only by much lower magnitudes of pullback spring and friction forces compared to those of the dial ones. The weight of the probe rod, and its associated components that move with it, tend to either aid or oppose the spring force depending on the orientation of the indicator gauge. So, when such indicators are in vertical-downward (the surface is below the probe) use, the resultant measuring force will be in its maximum regime. Obviously, the opposite would be true if they are in vertical-upward (i.e. upside-down) use. If they are used horizontally, these forces are in their average regimes, see Table 2.2. Thus, generally speaking, the gauging force of the spring-type electronic indicators are not only affected by the spring rates, but also influenced by the operating orientation of the indicator gauge.

As mentioned earlier, the friction forces from the gear-train mechanism and its joints (in dial indicators), bearings, and wiper affect the spring rate behaviour. This results in uncertainties in the "varying" gauging force along the travel range of the probe rod. Moreover, the interaction between the rod and the wiper is, in practice, sensitive to the direction of movement. The wiper is a kind of dust seal used to repel dirt from one side and, hence, needs to exert more forces on the moving member in one direction. The wiper does this as the rod retracts inside the gauge to repel dirt attached to the rod surface. As the rod extends outside the gauge, the wiper responds to the outwards movement that forces it to expand and, hence, exerts a lower force. As a result, the rod produces greater forces along the inwards direction than along the outwards direction, and a force hysteresis could be developed over one complete stroke of its travel. Thus, this friction-derived hysteresis leads to an uncertainty that could reach, as will be seen later, more than one-third of the maximum gauging force in some commonly-used precision dial indicators. Even if the sliding force at the seal is

consistent, there will still be a hysteresis effect in the contact force. At the instant the probe ceases to move, the friction will be in the opposite direction to that movement. The situation would then be as illustrated in Figure 2.5 if the spring force itself repeated exactly with position. It is unlikely that the friction force will reduce to zero once movement stops due to the static force. If lower contact forces are needed, the last direction of movement of the rod, before establishing contact, should be outwards.

2.4.2 Force Measurement on Real Indicators

To demonstrate the typical behaviour of the indicator's measuring force along the travel range of its rod, three linear high-accuracy dial and electronic gauge indicators were tested. These are: a John Bull (British Indicators Ltd.) dial indicator (model: 2U, travel: 12.7 mm, accuracy: unknown, resolution: 2 µm, and contact force: unknown), a Mitutoyo Digimatic series 543 electronic indicator (model: ID-F125E, travel: 25.4 mm, accuracy: ±3 µm, resolution: 1 µm, and contact force: 1.8 N or less) [33], and a Heidenhain digital length gauge (model: MT30, see Table 2.2 for its main specifications). All three were taken from the Warwick Precision Engineering Research Laboratory, where they had been used regularly and so were regarded as typical of 'working' instruments rather than being as-new ones. For convenience in this context, these indicators were given the codenames: G1, G2, and G3, respectively. The force measurement of the probe rod was accomplished with the aid of a Hounsfield Tensile/Compressive Test Equipment (model: H1KS) which has accuracy of ±25 mN and ±10 µm in force and displacement measurement [34], respectively, which is regularly calibrated to meet the industrial standards. Each indicator was mounted on the top fixed head on the stand of the Hounsfield system, as shown in Figure 2.6. The force transducer was clamped on the motorised head (crosshead) on this stand and was, then, pushed and pulled during measurement against the indicator's probe rod using the

system's controls. Additionally, the displacement readings of the Hounsfield system crosshead were also collected to be compared with those shown by the indicators.

Figures 2.7, 2.8, and 2.9 illustrate the hysteresis of gauging force measured along a complete return travel of the probe rods of the G1, G2, and G3 gauge indicators, respectively. The systematic effect of the spring rate was clearly shown by all of these indicators in both directions of travel of their probe rods, in addition to the distinctive effect of the wiper friction. Moreover, relatively minor variations of this force, mainly due to the combined effect of the stick-slip friction forces from both the bearing and wiper, were also noticed superimposed on the general trends.

The G1 indicator, which is the dial one, showed a maximum uncertainty in its contact force of around 2 N, as illustrated in Figure 2.7. Although there is no information available about the manufacturer's measuring force of this indicator, a rival one from Mitutoyo (refer to Table 2.1), for instance, has a measuring force of about 55% higher than its maximum one. The figure shows an increasing displacement error along the upwards travel of its probe rod, which reached to more than 230 µm at the end of the travel range where the maximum force regime is nominally located. At the onset of downwards travel of the rod and during the rapid decrease in its force (due to the change of magnitude and direction of the wiper friction force), this displacement error decreased back (because of reversing the rod motion) to its initial magnitude. Along the downwards travel of the rod, the differences between the two displacement readings were below 55 µm. While there is some evidence of a calibration divergence between the gauge and the Hounsfield and there may also be a cyclic error in the gauge, the magnitude of the errors correlates with increasing force. The different manufacturing tolerances and wear processes within the gear-train may have influenced this increase (or decrease) of such error along the travel range. The higher forces may directly introduce errors by straining the gauge mechanism and may also cause higher wear.

As expected, the G2 and G3 indicators showed much lower uncertainty in their contact forces than the G1 indicator, since they are electronic ones. The maximum magnitudes of their gauging force were observed to be within the values provided by their manufacturers, with maximum uncertainties of about 0.9 N and 0.5 N, respectively. The rod displacement error was also noticed on these two indicators, but with much lower maximum magnitudes and at different locations in the hysteresis loop. Along the upwards movement of the rod, the G2 indicator produced its maximum displacement error of nearly 120 µm between 3 and 6 mm of rod displacement: the region has higher and more variable contact forces than normal, as shown in Figure 2.8, presumably due to high friction forces. The G3 indicator revealed approximately 90 µm maximum displacement error within the 4 to 8 mm region which has the highest contact forces, as shown in Figure 2.9. Similarly to the G1 indicator, at the start of downwards movement of the rod and the large change in its force, this displacement error decreased and, along the downwards movement of the rod, the maximum differences between the displacement readings did not exceed 75 µm and 45 µm on the G2 and G3 indicators, respectively. When considering gauging over only a limited region of travel in this direction, the relative deviations in displacements are obviously far less than the maximum ones. This should easily be recognised through the calibration processes that should be carried out regularly on any indicator to certify its serviceability. However, a very detailed correction scheme would be needed to compensate directly for an error varying as these do. Furthermore, the effect of varying force on a compliance surface has not been taken into account.

2.5 REMARKS

In many modern designs of precision indicators, the wiping seal is replaced by an

external corrugated rubber cover that conceals the rod surface in order to reduce the force uncertainty due to the friction-derived hysteresis. In some designs of electronic indicators of high accuracy, the gauge head is also equipped with a motorised or a pneumatically-controlled plunger to eliminate the spring rate effects, as well. However, effects of the force contributed by such stretching cover to the resultant gauging forces could also be of notable magnitudes. In addition, the measuring (or approach) speed of the contact probe, towards the surface to be gauged using indicators with controlled probe rods, could be a new significant parameter, since it may lead to additional plastic deflections on the surface due to the impact.

The resultant interaction between the different inherent forces in the indicator mechanism is becoming increasingly important in evaluating the error budget of the complete gauging process. When high precision comparative measurements are to be achieved on surfaces of different characteristics, the contact force uncertainties of the indicators used obviously generate significant uncertainties through the deformation at these surfaces which, consequently, affect the final accuracy of such measurements. To give exaggerated estimates for such deformations on perfectly smooth surfaces using the Hertzian contact analyses, a 3 mm diameter probe tip with an average gauging force of 2 N could cause indentation depths of 0.5 µm and 0.8 µm on steel and aluminium surfaces, respectively. These deflection values will at least be doubled on such surfaces if a moderate roughness of just around 0.2 µm is considered on them, as will be discussed in the following chapters. Hence, a contact force uncertainty of 0.5 N could additionally lead to variation in deflections in the micrometre range on such surfaces.

Among the experienced users of indicators, it might be well-known that, in high precision measurements, probe rods are recommended to be brought gently into contact with surfaces during their downwards (return) travel in order to avoid the region of the high gauging forces and backlash errors. In practice, using indicators with small ranges

of travel (2 mm or below, for instance), which are usually chosen for such purposes, makes it difficult to guarantee that adjustment in the recommended way occurs in every location of measurement. So, the maximum uncertainty in contact force (due to the spring rate) is still of a potential effect in gauging with these indicators, in addition to the uncertainty in displacement. This is also possible in absolute measurements with digital indicators, such as the G2 and G3, where both probe rod directions of movement could be involved. In all of the three indicators tested, long ranges of travel were generally observed to contribute more displacement errors to the measurements because of the wider chance for friction forces to show a larger variation.

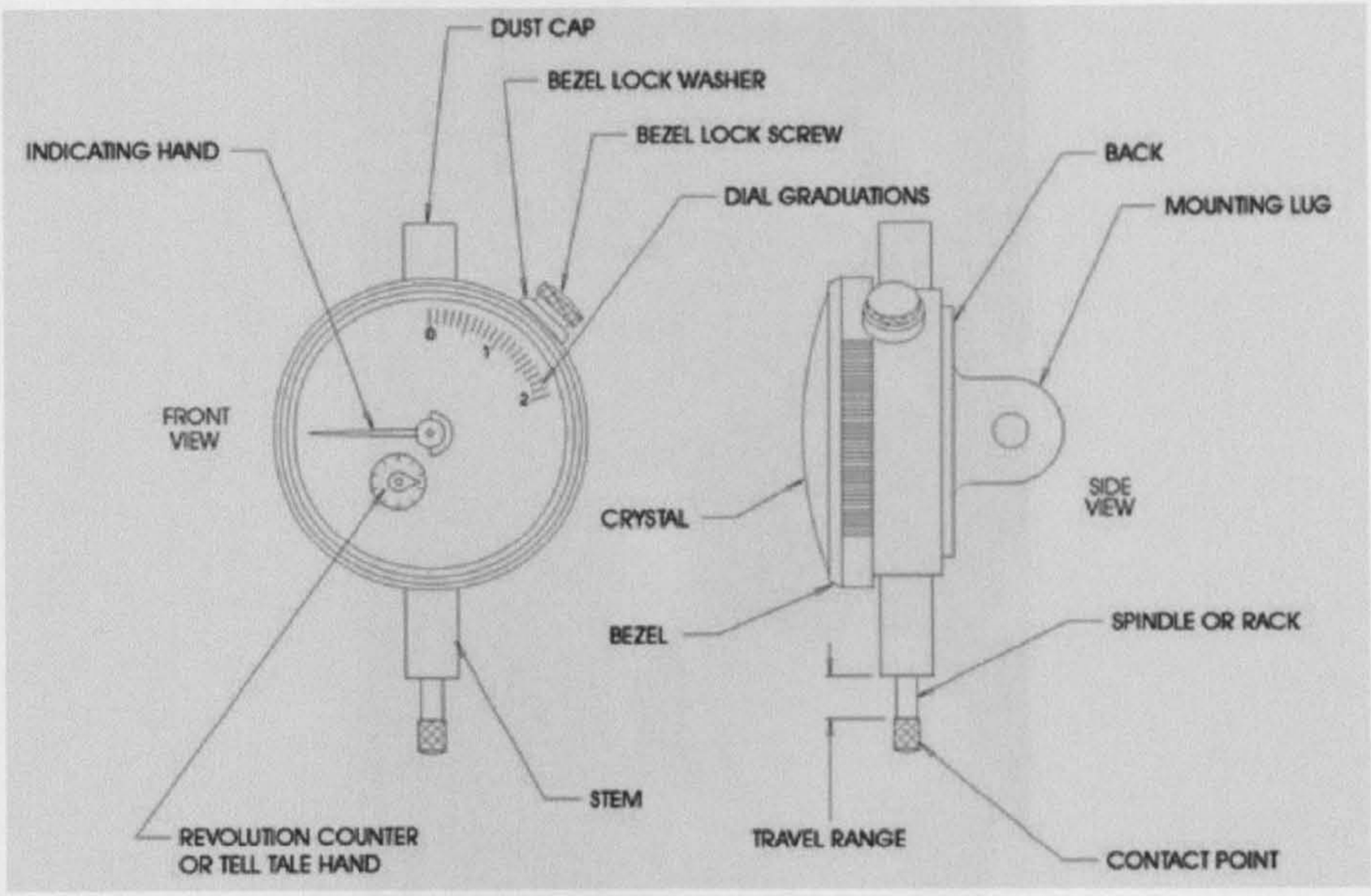


Figure 2.1 A schematic representation of a typical dial indicator and its different external parts [35].

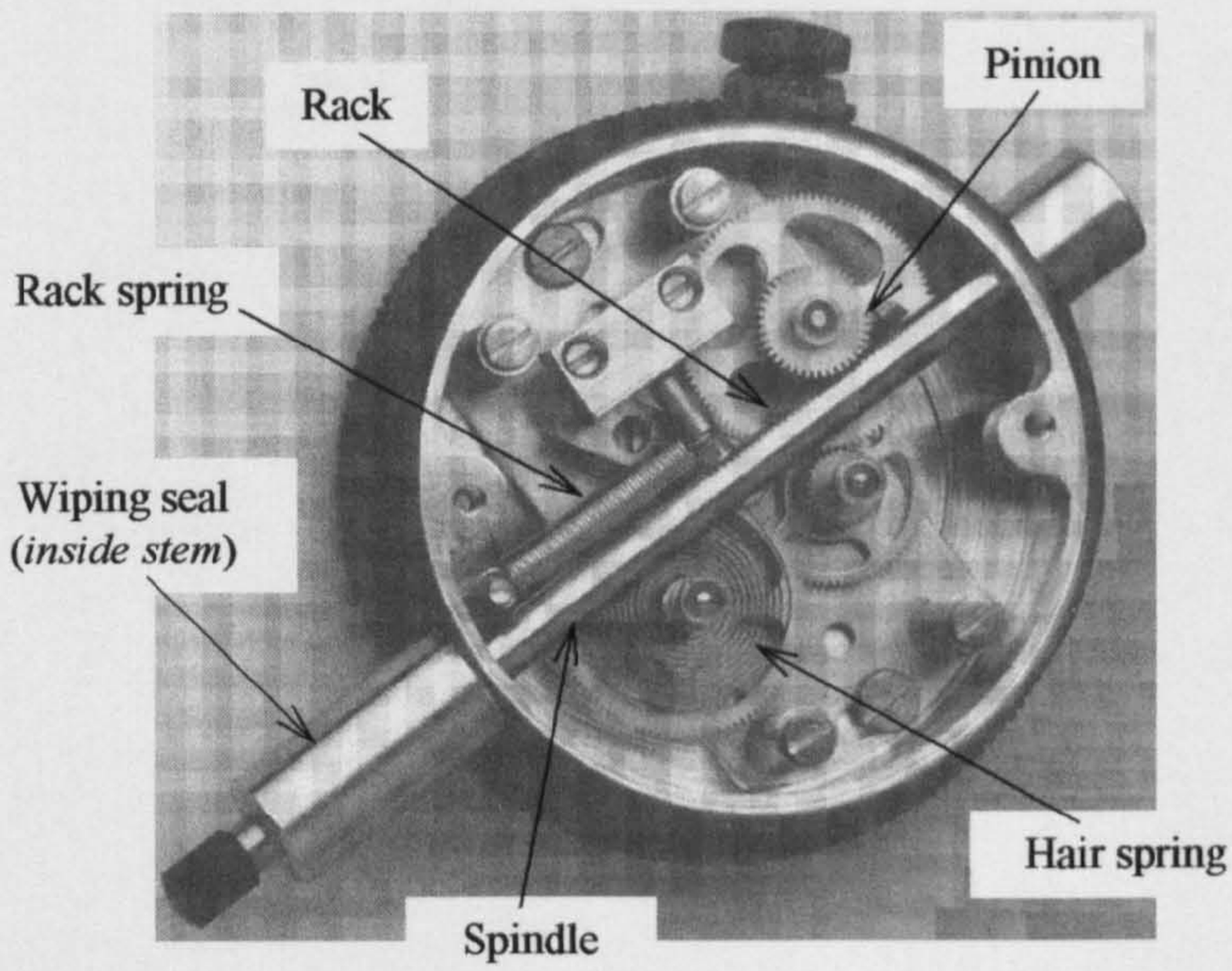


Figure 2.2 Rear view of a dial indicator showing some of its internal components [11].

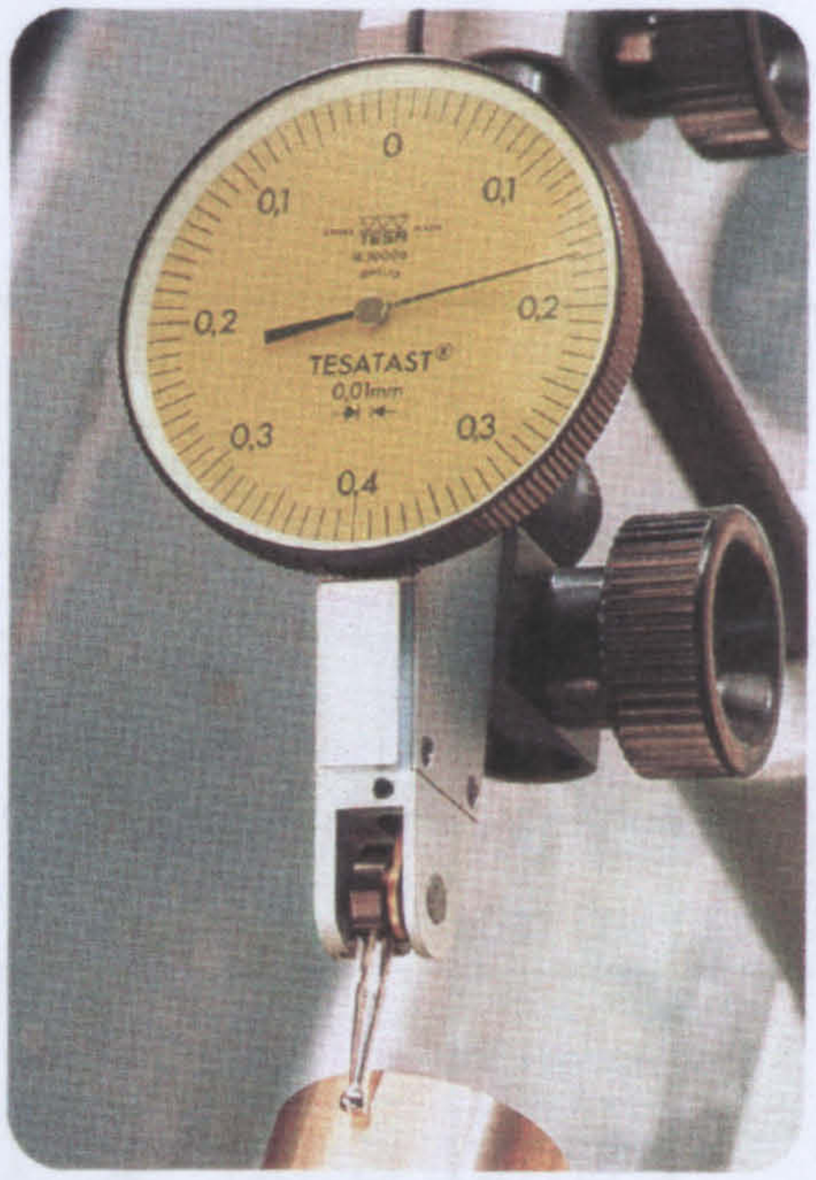


Figure 2.3 A typical standard dial test indicator, TESATAST lever-type dial test indicator [36].

Figure 2.4 A typical standard digital gauge indicator, FEDERAL μMaxμm digital indicator [37].

	Gauge Model No.	Travel (mm)	**Measuring Force (N)**	Graduations (mm)	Accuracy (μm)	Hysteresis (N)
Mitutoyo [38] Series 1 and 2 Metric Dial Indicators	1900F-60	0.1	1.5	0.001	2	N/A
	2109F	1			3	
	2113F	2			5	
	1125E-10	3.5	1.4	0.005	12	
	2045FE	5		0.01	10	
	2046FE-80	10	5		12	
TESA [36] Precision Dial Gauges	40	1	2	0.001	5	0.4 - 0.5
	40	3		0.01	12	
	57	10	2.2		17	
Mahr [39] Precision Dial Indicators	803B	0.4	1.1	0.01	9	N/A
	810SB	0.8				
	803A	3			12	
	810A	10			17	
	810V	40	0.75 – 1.5		30	

Table 2.1 Typical measuring forces of some models of dial gauge indicators from three well-known manufacturers.

	Gauge Model No.	Travel (mm)	Measuring Force (N) *Direction of use (Operating attitude)* *vertically downwards*	*vertically upwards*	*horizontally*	Resolution (µm)	Accuracy (µm)	Hysteresis (µm)
Mitutoyo [40] Linear Gauge Heads & Digimatic Indicators	542-204	5	0.65	0.55	0.6	1	2	N/A
	542-222	10	0.8	0.7	0.75			
	542-131		1.2	1.1	1	0.5		
	542-132	25	4.6	4.3	4			
	542-133	50	5.7	5.3	4.9		2.5	
	Series 543	12.7 - 50.8	0.9 - 2.3			1 - 10	3 - 40	
TESA [41] Electronic Linear Measuring Probes	GT43	2.1	0.4			0.1 - 1	0.01	0.015
	GT212	3.2	1.2				0.015	0.02
	GT21	4.3	0.63				0.01	
	FMS100	5.8	2				0.5	0.5
	GT61	10.3	0.9				0.05	0.05
Mahr FEDERAL [37&39] Electronic Indicators	µMaxµm	2	0.85 - 1.13 (at centre of spindle travel)			0.5 - 1	2 - 5	N/A
	µMaxµm XL	12.7 - 25.4					5	
	1083 Millitast	12.5	0.7 - 0.95			1	5	
	1085 Millitast	25	0.6 - 1.1					
	1082	50	2 - 3			10	20	
		100	2.3 - 4					
HEIDENHAIN [42] Digital Length Gauge Heads	MT10	10	0.3 - 0.8	0.1 - 0.6	0.2 - 0.7	1	1	N/A
	MT30	30	0.5 - 0.9	0.2 - 0.6	0.3 - 0.8			
	MT60m	60	1.75	0.75	1.25			

Table 2.2 Typical measuring forces of some models of electronic gauge indicators from four well-known manufacturers.

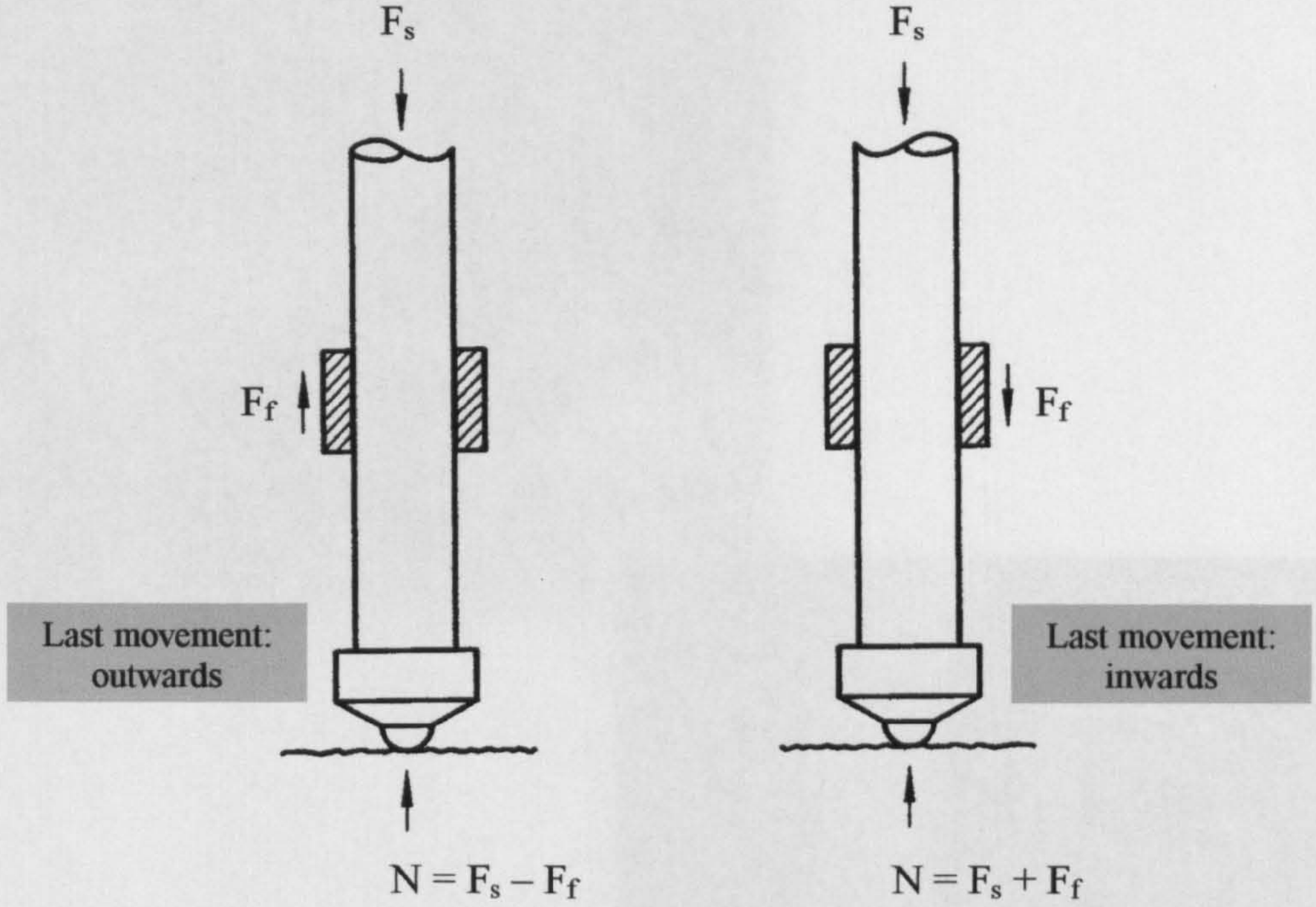


Figure 2.5 The effect of friction force from the wiping seal on the gauging force, N, which is governed by the direction of last movement of the indicator's probe rod.

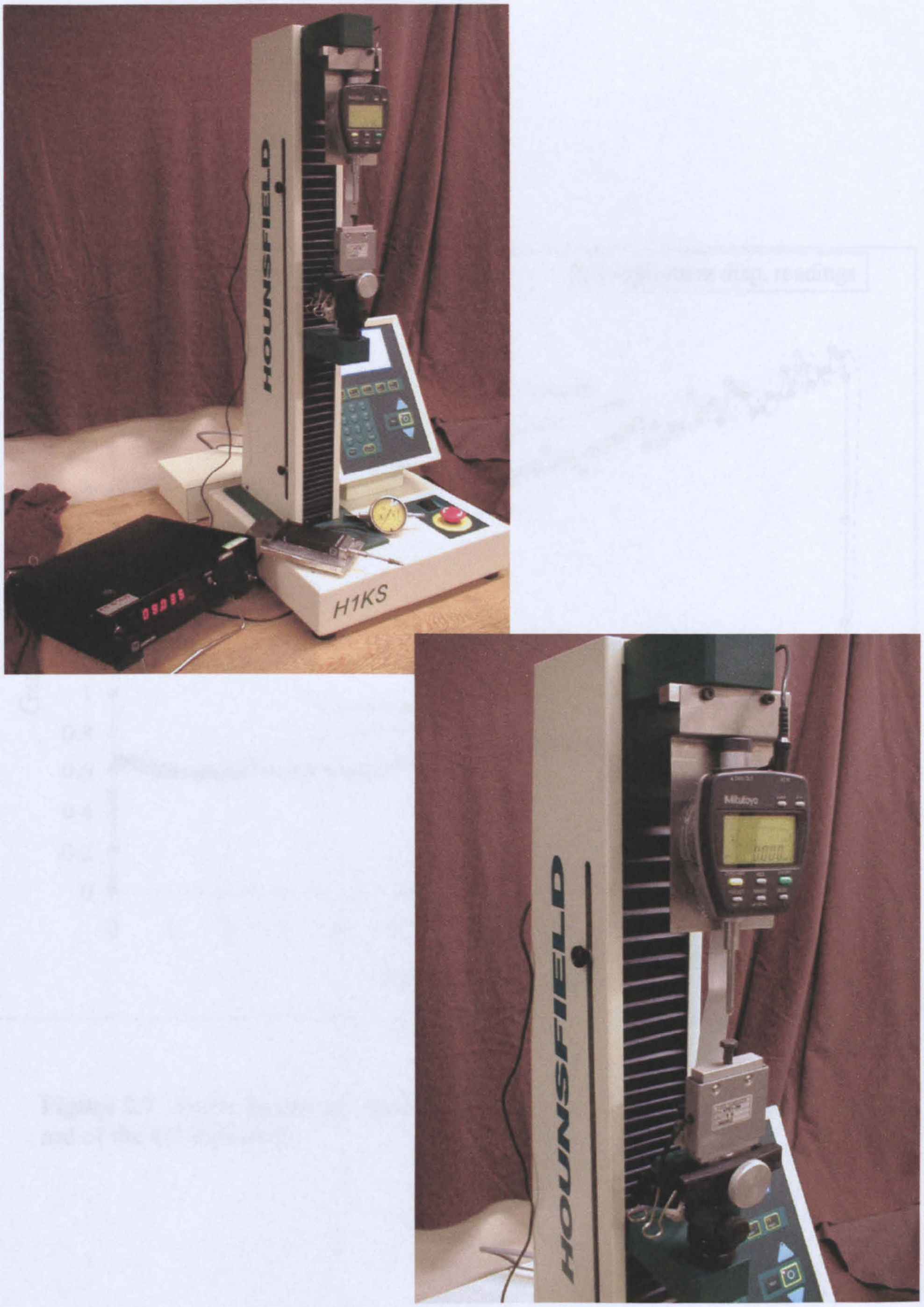


Figure 2.6 Photographs of the setup used for measuring the gauging force of the three available dial and digital gauge indicators using the Hounsfield Test Equipment.

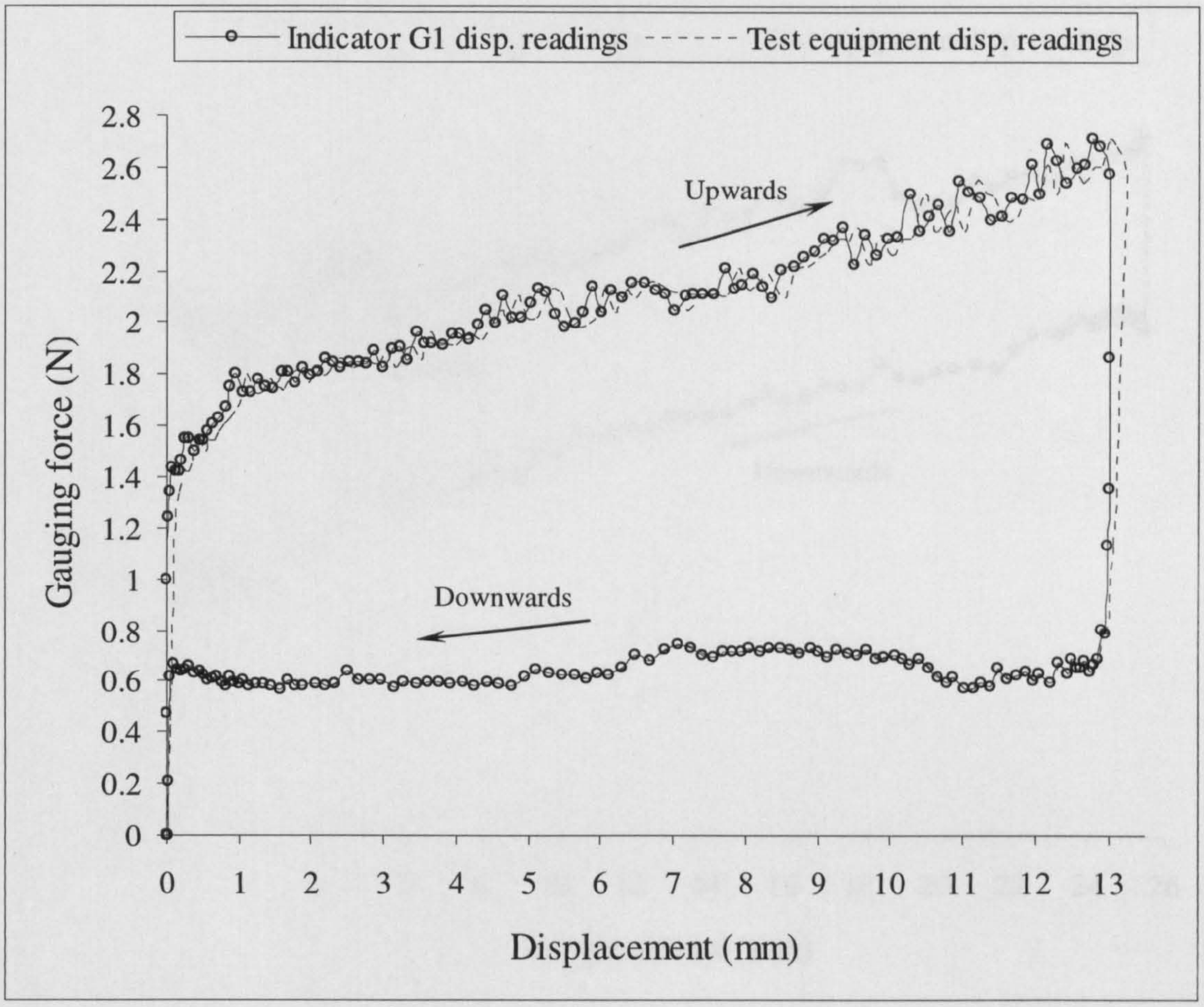


Figure 2.7 Force hysteresis measured along the travel range of the probe rod of the G1 indicator.

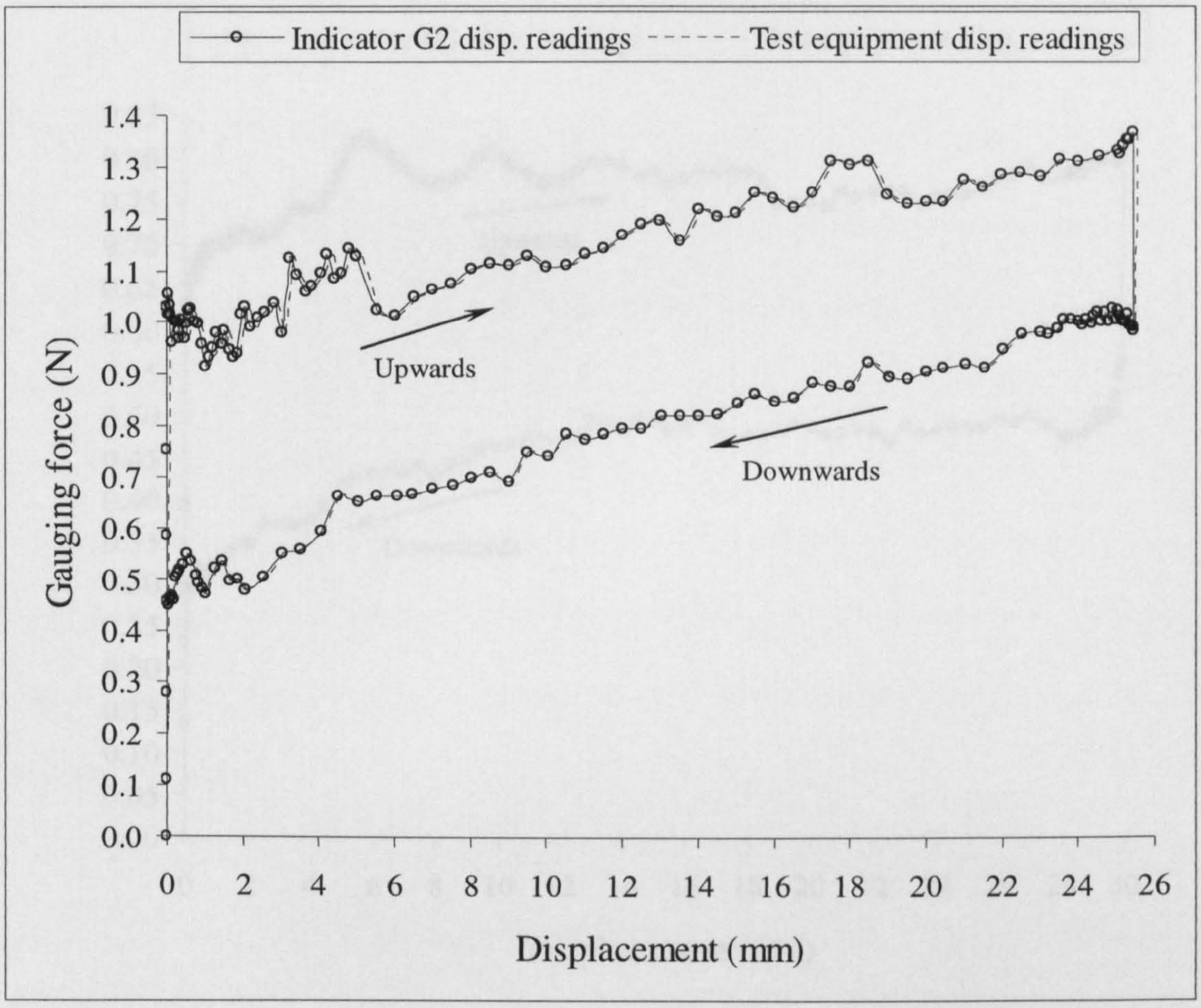


Figure 2.8 Force hysteresis measured along the travel range of the probe rod of the G2 indicator.

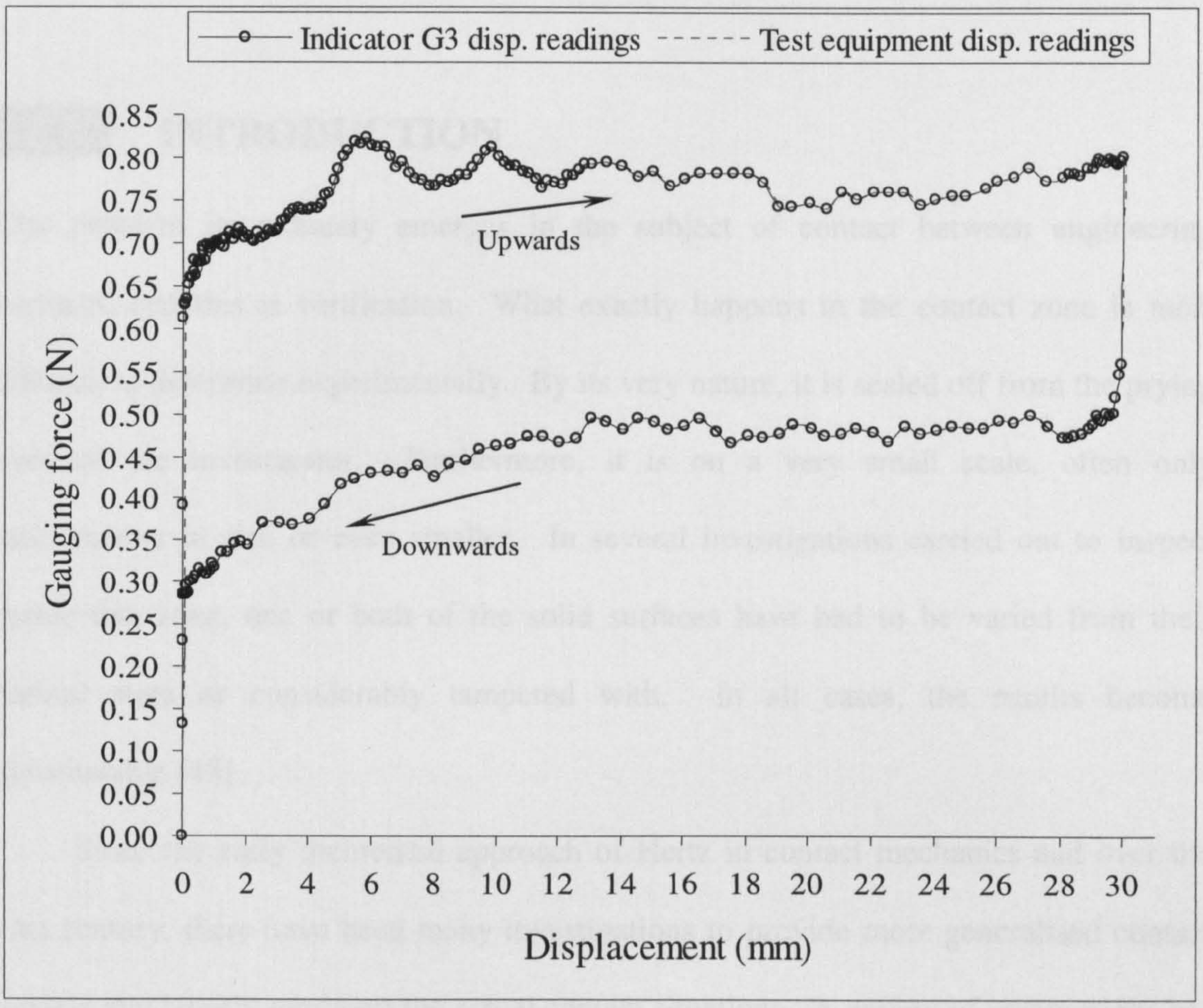


Figure 2.9 Force hysteresis measured along the travel range of the probe rod of the G3 indicator.

CHAPTER 3

THEORETICAL APPROACHES FOR DETERMINING CONTACT BEHAVIOUR

3.1 INTRODUCTION

One problem immediately emerges in the subject of contact between engineering surfaces, and this is verification. What exactly happens in the contact zone is most difficult to determine experimentally. By its very nature, it is sealed off from the prying eyes of the investigator. Furthermore, it is on a very small scale, often only micrometers in size or even smaller. In several investigations carried out to inspect inside this zone, one or both of the solid surfaces have had to be varied from their natural state or considerably tampered with. In all cases, the results become questionable [43].

Since the early theoretical approach of Hertz in contact mechanics and over the past century, there have been many investigations to provide more generalised contact models that closely represent the actual contact situations. A variety of assumptions has been inevitably adopted in the analysis of the different contact configurations of the real applications in order to simplify the formulations and, consequently, expedite the convergence to the solution of these contact problems. So, such contact models of rough surfaces generally offer predictions that are useful in, for example, designing improved functional performance only at very specific contact conditions of interest. In this study, the use of these models has been customised to the contact situations which are believed to be exist during the contact gauging of rough engineering surfaces, towards estimating the associated errors contributed to the measurement.

A complete review on the models of contact between rough surfaces is obviously beyond the scope of the thesis. Comprehensive coverage is generally within the extents of many specialised textbooks such as [43 and 44]. This chapter gives only an overview of the main concerns and contributions that are of close relevance to the thesis interest. It also describes the approximations made in two common contact models towards suggesting simplified theoretical approaches that could be used for estimating contact uncertainties included in the measurement with indicators.

3.2 HERTZ THEORY OF ELASTIC CONTACT

The modelling of stresses and deflections arising from the contact between two elastic solids has practical applications in hardness testing, wear and impact damage of engineering materials, and the design of gear teeth and ball and roller bearings [45]. The first satisfactory analysis relating the induced stresses and strains at the contact of two elastic solids was initially developed by Heinrich Hertz. He was studying Newton's optical interference fringes in the gap between two glass lenses and was concerned at the possible influence of elastic deformation of the surfaces of the lenses due to the contact pressure between them. His theory aroused considerable interest when it was first published in 1882 and has stood the test of time [44]. The history and analysis details of the theory were extensively presented by K. L. Johnson in [46], who collated much practical and theoretical work relating the induced stresses and strains under a wide variety of load conditions and presented, later, the subject of contact mechanics [44] in a mathematically rigorous and concise manner [47].

Hertz idealised his theory by formulating the following assumptions in order to facilitate the analysis:

(i) The contacting surfaces are continuous. This means that each surface is

topographically smooth, which implies the absence or disregard of small surface irregularities that would lead to a discontinuous contact or highly local variations in contact pressure.

(ii) The radius of the contact circle is small compared with the dimension of each body (non-conforming bodies). This allows the indentation stresses and strains to be considered independently of those arising from the geometry, method of attachment, and boundaries of each contacting solid.

(iii) The radius of the contact circle is small compared with the radii of curvature of the contacting bodies. With this assumption, each surface may be treated as an elastic half-space where the methods for solving boundary-value problems can be readily used.

(iv) The contacting bodies are in frictionless contact. That is, only a normal pressure is transmitted between them.

The contact between a rigid indenter and a flat extensive specimen is of a particular interest. The shape of the indenter may be spherical, conical, a cylindrical flat punch, or even take the form of a uniformly applied pressure. All such phenomena are usually referred to as "Hertzian contact". The most well-known scenario is the contact between a rigid sphere and a flat surface, as given in Figure 3.1, where the top view of the contact area is circular. Hertz found that the radius of the circle of contact, a, is related to the indenter load, P, the indenter radius, R, and the elastic properties of the two materials by

$$a = \sqrt[3]{\frac{3PR}{4E^*}} \tag{3.1}$$

where E^* is the elastic modulus of contact given by

$$E^* = \frac{E_1 E_2}{E_1\left(1-\nu_2^2\right)+E_2\left(1-\nu_1^2\right)} \tag{3.2}$$

where E_1 and E_2 are the Young's modulus, and ν_1 and ν_2 are the Poisson's ratios of the two surface materials. The mutual approach of distant points in the two materials is given by

$$\delta = \frac{a^2}{R} = \sqrt[3]{\frac{9P^2}{16RE^{*2}}} \tag{3.3}$$

He also found that the maximum contact pressure on the surface of the specimen occurs at the centre of the contact circle and is given by

$$p_0 = \frac{3P}{2\pi a^2} = \sqrt[3]{\frac{6PE^{*2}}{\pi^3 R^2}} \tag{3.4}$$

The mean contact pressure is given by the indenter load divided by the contact area, which is two-thirds of the maximum contact pressure

$$p_m = \frac{P}{\pi a^2} \tag{3.5}$$

The maximum tensile stress in the specimen happens at the edge of the contact circle at the surface, i.e. at $r = a$ and $z = 0$, and is given by

$$\sigma_r = \frac{(1-2\nu)p_0}{3} \tag{3.6}$$

and the maximum shear stress appears at the centre of the contact circle, $r = 0$, but below the surface, $z = 0.48a$, and is given by

$$\tau_1 = 0.31 p_0 \tag{3.7}$$

The analyses of other shapes of indenter have been demonstrated, concisely, in several publications such as [48] and, comprehensively, in more recent ones such as [49].

Due to the boundary conditions required in this theory, its direct implementation, without enhancing its classical approximations to the real contact conditions, became somewhat rare. For instance, great caution must be taken when applying its results to materials of low modulus, such as rubber, where it is easy to produce deformations which exceed the restriction to small contact strains. Another more contentious issue in

the theory is the assumption of smooth surfaces which has received a vast amount of investigations during the last four decades. The consideration of friction in contact has also been recognised in other investigations since the contact pressure is not usually perpendicular to the interface which is not necessarily planar, and in addition it is necessary to account for the changes in the boundary forces arising from the deformation that they produce. However, these considerations and others are briefly discussed within the following section in order to highlight most of the different approaches of concern to researchers towards a better understanding of the mystery of contact between engineering surfaces.

The results of Hertz theory can still be used to provide simple and quick estimates for the "ideal" stresses and strains in many elementary contact problems. The theory has perhaps been used beyond its limits and the Hertz formulae now form the basis of an ever-widening range of methods for testing and determining the properties of materials. Because of this, it is still attracting some interest, such as [50 and 51], for generalising its implementation to wider "ideal" contact conditions.

3.3 CONTACT OF ROUGH SURFACES

In reality, continuous interfaces between real contacting components are extremely rare due to the inevitable topographical roughness inherent on their man-made or natural surfaces. The contact between such solid surfaces is generally discontinuous and the real area of contact is a small fraction of the nominal contact area [44]. If the roughness is high, the area of intimate surface contact is usually much less than the apparent contact area [52]. The real contact pressure at the true contact spots will inevitably be much higher than the pressure given by the Hertz theory applied to the contact as a whole; in between the true contact spots, the pressure must fall to zero [46]. The

number of these spots is determined by the fine detail of the surface texture that must reveal the density of peaks or summits and the distribution of their heights and shapes, which also determine the possibility of interaction. It is, however, rarely that the behaviour between real surfaces in contact could be experimentally determinable since the contact region is almost totally inaccessible and very small relative to the component size. Many attempts have been made to measure directly the number of contacts, but most have been relatively unsuccessful. The usual method has been to resort to estimating the contacts either from pre-contact measurement of the mating surfaces with some degree of extrapolation or by computer simulation. Both techniques are better than nothing but not ideal [43].

In the dynamic function of surfaces, the influence of their roughness on contact behaviour is of a great importance in many tribological situations. Apart from causing high contact stresses, the surface roughness is crucial with respect to the stiffness, wear, adhesion, friction, lubrication, and thermal and electrical resistance properties of the contact. Since there are cost and performance considerations when smooth surfaces are manufactured, it is important for tribologists to understand how surface roughness affects the contact conditions and to be able to predict an acceptable degree of roughness for specific contact situations [53]. The contact properties between gear teeth, roller and race way, cam and follower, and shaft and journal are examples of common problems which frequently attract many tribologists towards more specific and better compromises. Such investigations are beyond the concern of this thesis since the contact between the probe tip of the indicator and the surface is nominally static.

Models for the contact of rough surfaces are numerous and they have been refined over many years. However, generally the main objective of different investigations that considered roughness within the contact condition between surfaces has been the determination of the actual area of contact. This has been obtained by: (i) experimental

techniques, (ii) stochastic approaches, and (iii) numerical models. In the experimental approach, several methods have been reported: thermal, electrical, optical direct observation, internal reflection, Nomarski interferometry, Neutrographical method, computer simulation, and relocation profilometry. Refer, for example, to [54] for the contributors of these methods.

Numerical models have been developed over the last two decades and used on rough surface contact problems. This approach provides a deterministic solution to rough surface contact stresses and areas. It does not require an asperity model and makes no probabilistic assumptions such as the distribution of asperity heights, slopes, curvatures. The technique takes full account of the interaction of deformation from all the contact points and highlights the deformable contact geometry of rough surfaces under loading. It provides useful information on the number of contacts, their sizes and distributions, and the spacing between contacts [54]. On the other hand, the approach requires detailed information about the surface profile (geometry), which is usually obtained from either the direct measurement of real surfaces or computer simulation. This limits the utilisation of such an approach to contact problems where comprehensive analyses are needed. However, the approach can be regarded as an exhaustive method when quick, computationally feasible, and convenient estimations are frequently required.

Stochastic models have also been developed to describe the relationships between load, mean separation, and real contact area in static contact [55]. They are usually known as the "asperity models", since the rough surface is represented by a collection of asperities of prescribed shape that are scattered over a reference plane. Such contact models are numerous and the assumptions common to all of them are: (1) that the summits can be approximated as quadratic surfaces so that Hertz theory governs the relation between load, displacement, and contact area at each individual micro contact;

(2) that the statistical properties of the surface have some mathematically defined form, often a Gaussian distribution of asperity heights and curvatures; and (3) that the surfaces deform elastically. Adopting such conditions was essential in order to bound the behaviour of the contact interaction and to provide less complicated analyses.

The work of Greenwood and Williamson [56] (GW) is an early and leading example of such models which provided a considerable simplified understanding of the rough contact problem. In this model, it is assumed that the heights of asperities have a Gaussian distribution and that the contacting asperities deform elastically according to Hertz theory. It is also assumed that these asperities do not merge during the deformation. Moreover, the contact was considered normal between a flat surface and an equivalent rough surface comprising uniformly-distributed (over the projected in-plane surface area) spherical asperities of a constant radius.

A large number of contributions in the contact of rough surfaces were based on the GW work in order to relax most of its controversial assumptions and provide more realistic approaches. For instance, the assumption that the asperities do not interact may be violated for large contact loads which would result in an increase of the ratio of true contact area to apparent area. Another shortcoming of the GW model is the predicted area of contact. From the statistical point of view, even for very small contact loads there will be some contacting asperities that have reached their elastic limit and deformed plastically. As the load is increased the number of these plastic contacts also increases, yet the corresponding area of contact in the GW model is calculated from the Hertz theory for elastic deformation, neglecting volume conservation in the plastically deformed regions [57].

3.3.1 Effect of Plasticity

Most engineering surface finishing processes create asperity geometries, particularly on

the small scale of roughness, which will undergo some level of plastic deformation even at the lightest of loads [58]. This is evident from the application of plasticity index criteria, such as that of GW [56], and from more recent numerical elastic contact solutions, such as in [59], to rough surface profile data [60]. However, the basic plastic contact model, known as the "profilometric model", was introduced in 1933 by Abbott and Firestone. It assumed that the area of contact is the geometrical intersection of the flat with the original profile of the rough surface [57]. Unlike the theories based on asperity model, the theories based on profilometry attempt to start from the real situation, and the simplifying assumptions are made during the solution [61].

The investigation in [62] made notable progress in understanding the nature of plastic contact and is based on the results presented in [63], which revealed that when incremental deformation of the asperity tips occurs the incremental plastic volume reappears where there is no contact. And in the case where the surface is constrained, this effect may inhibit asperities collapsing with the application of the load, which was called "asperity persistence" and has the effect of making the surface appear to be harder than it is [43]. In reality, this phenomena is created by the interface friction, the strain hardening, and the subsurface interaction of plastically deforming material from neighbouring asperity deformations [60]. It has been concluded that under heavily loaded contact conditions the precise form of the surface texture has no practical effect on the persistence, but the plastic asperity interaction does cause higher contact pressures [64].

The above two assumptions of the plastic contact model may be valid for very heavily loaded contacts when neighbouring asperities merge to form large contact patches. For lightly and moderately loaded contacts in which there is a large percent of elastically deformed asperities and also an appreciable percent of asperities beyond their elastic limit, the assumptions made in these two models appear to be unrealistic.

Although, the boundary between elastic and plastic deformation of rough surfaces has been considered in GW itself and by several workers such as [65, 66, 67, and 68], the fact of volume conservation of plastically deformed asperities was not considered. However, it has been considered by the authors of [57] in their alternative elastic-plastic asperity model for moderately loaded contacts.

3.3.2 Effect of Anisotropy

Most commonly-machined metal surfaces have orientation or a lay with respect to the direction(s) of motion of the cutting tools relative to the workpieces. Such surfaces usually have highly directional properties which are likely to affect their functional behaviour [69]. So, the anisotropic roughness must be considered in the analysis of their contact problems. However, when a random and Gaussian roughness is also considered, the analysis of the asperity model will become extremely complicated since, at least, seven parameters are required to define the topography of surfaces [70]. Nayak [62] also approached this fact by considering plastic surfaces in three dimensions in which a lay was present and using the "random process model" which had been developed in 1957 by Longuet-Higgins. Bush *et al* also utilised this model to analyse the static contact of isotropically rough surfaces in elastic [71] and plastic [66] modes of deformation, and solely considered the parameter of anisotropy for a simplified and extensive analysis [72].

The characterisation of surfaces in three dimensions, obviously, increases the time and effort of computation compared to the asperity-based methods, and introduces some new topics such as anisotropy [69]. However, when the numerical techniques are used, the need for considering the anisotropy in the contact analysis of rough surfaces became satisfied since the different features of the digitised (or computer-generated) surface have been fully taken into account in this analysis. The choice of the mechanical state

of contact, which largely represents the demanded contact condition, remains rather crucial as it could significantly affect the analysis. Investigators, who used these techniques, were also observed adopting different modes of contact deformation. For example, the authors of [53, 55, 59, and 73] assumed a completely elastic contact mode and used computer-generated surface profiles to examine their proposed numerical analyses; but in [58], digitised measurements from real rough surfaces have been used instead. The elastic-plastic contact mode is considered in several numerical techniques such as in [70 and 74] where simulated topographies of the rough surfaces are used, but the analyses in [54 and 75] have been developed to deal with the data from both the real and the simulated surface topographies towards a more generalised numerical model.

3.3.3 Effect of Adhesion

The adhesion phenomenon is another parameter which is observed, attracting some concerns when dealing with the force analysis between surfaces in contact. It is basically a force that operates at the contact interface due to the atomic nature of the surfaces. It is a result of the actions of attraction and repulsion between atoms or molecules in the two surfaces, which greatly depend on the separation between them. It was first realised when the contact of a glass ball on a soft and smooth rubber was found not to be in agreement with the Hertz theory: a finite contact area was observed under zero load, and a finite force was needed for separation [76].

Johnson *et al* [77] discussed the influence of surface energy on the contact between elastic solids and, accordingly, they modified the Hertz theory (JKR theory) and verified their analysis experimentally by contacting a glass sphere with flats of rubber and gelatine. Using a thermodynamic approach, Derjaguin *et al* [78] presented a completely different theory (DMT theory) which took into account the reciprocal influence of the contact deformation and molecular attraction of a ball and a plane, in

combination with the Hertz theory. These two theories appeared at first to be contradictory, which led to a sharp debate, until it was agreed that the JKR theory is valid for soft solids with a large surface energy and radius, and that the DMT theory is valid for hard solids with a low surface energy and small radius. However, through many recent publications, such as [79 and 80], the mechanics of adhesion between homogeneous elastic solids having spherical surfaces has become well-established. The different models of adhesion have found several practical applications where a guide, such as in [81], is generally useful for the selection of the appropriate model to use, depending on the conditions (the size and elasticity of the spheres and the load to which they are subjected).

Considering the contact of rough surfaces, it is, however, a common experience that no adhesion force is observed when two metals are pushed together. Investigators, like Bowden and Tabor, suggested two reasons why normal adhesion is not detected. Firstly, the surfaces are covered with adsorbed films even when the surfaces have been carefully cleaned in the atmosphere. Secondly, which is more important, the adhesion junctions formed between the lower asperities on loading are elbowed apart on unloading by the elastic compression between the higher asperities. In this way, adhesion developed at the points of real contact is progressively broken down. For materials with high modulus, the compression is more significant than the adhesion and, therefore, adhesion is not detected [82].

3.3.4 Effect of Friction

The interfacial friction induced during the contact deformation is another parameter which could affect the contact stresses and strains at the contact interface of the two bodies. Considering a normal and non-conforming contact between any two real surfaces, it plays a part only if the elastic constants of the two materials are different

[44]. The mutual contact pressure produces tangential displacements at the interface as well as normal compression. If the materials of the two solids are dissimilar, the tangential displacements will, in general, be different so that slip will take place. Such slip will be opposed by friction and may, to some extent, be prevented. Therefore, a central region might be expected where the surfaces stick together and also regions of slip towards the edge of the contact. If the coefficient of limiting friction were sufficiently high, slip might be prevented entirely [44]. However, this sounds more convincing for heavily loaded static contacts where relatively large deformations at both surfaces could readily stimulate such process.

Friction is usually considered to be a significant parameter when a relative movement (dynamic contact) is present between the contacting surfaces. For instance, refer to the classification of surface interaction in contact and tribology problems given in Tables 7.1 and 7.2 in [43], in Figure 3 in [83], and in Figure 6 in [84]. Obviously, the study of this parameter in tribology arises from the demand for improvement in the functional performance of surfaces (in sliding and rolling contacts) which is governed by the efficiency of lubrication. Tribologists regard the friction force as the combined effect of adhesion, deformation, and ploughing; in which the relative contribution of these components depends on the environmental conditions of the sliding interface and materials used [82]. Furthermore, the process of wear is inevitably discussed in conjunction with the friction since it also depends on the surface finish and the mode of deformation. However, several mechanisms and models of friction (at the asperity scale) and their associated analyses have been suggested for greater understanding and better prediction of the real contact situations, which are beyond the extent of the thesis.

3.3.5 Other Concerns

In addition to the consideration of the previous parameters in the contact analysis of

rough surfaces, other factors of concern were found by several investigators. Such concerns arose from the frequent needs to match the analytic solutions with the actual contact conditions and to provide more general and reliable contact models. The investigations, which are briefly described below, are only few examples.

The effects of elastic conformity within rough surface contacts has been studied in [85] to define the spatial size of asperities which are likely to deform elastically or plastically, assuming a Hertzian asperity contact. A parameter of conformity was introduced in this study and was demonstrated to identify and explain some practical problems in common applications of rolling contacts.

The contact between non-Gaussian isotropic random surfaces has been numerically analysed in [86] since most of the common machining processes produce surfaces with asperity heights of non-Gaussian distributions. The analysis highlighted the effect of such a consideration on the functional contact of surfaces and its tribological characteristics which could offer promise for the design of interface roughness to provide low friction and wear.

In [87], the effect of the mutual influence of asperities on the Hertzian micro-contacts has been considered in the numerical analysis of the elastic contact of ideal rough surfaces, where asperities are assumed to be spherical with the same vertical height and spatial distance. The location of the reference plane, for the Hertzian asperity contacts between rough surfaces, has been proved to be varying with the contact load, and not as it is usually assumed to be in the mean line position.

3.4 CONTACT OF GAUGE PROBE TIPS ON ROUGH SURFACES

The aim in this section is to suggest a simplified theoretical approach that reasonably

represents the real contact configuration between the spherical tip of the gauge (indicator) probe rod and the surface under gauging, based on the experimental results presented in this thesis. Such a theoretical model could be employed to obtain a reliable prediction about those errors (due to surface deformations) which are certainly included in the measurement with contact indicators. For such a frequently-encountered contact problem in precision machine shops and quality checkpoints, the model should also be easy to use (without a great deal of complicated mathematical formulae) as well as time-saving (without involving in exhaustive experimental procedures to describe the prerequisite topography of surfaces being gauged). It is, then, obvious to restrict the concerns to the common analysis of contact of a smooth ball on a flat rough surface. Furthermore, the asperity (stochastic) models seems, to a large extent, more reasonable for use in this case than the numerical (profilometric) models, since the former ones require much less details about the surface and, hence, provide a considerably easier and faster convergence to the required contact solution.

As reported in the previous chapter, almost all of the precise models of contact indicators have gauging forces of less than 3 N. At such a regime of contact force, which is much below those regimes of interest by the contact mechanics researcher and tribologists, the contact between the indicator's probe tip and the rough surface can be regarded, with confidence, as a lightly loaded contact in which the main mode of deformation of asperities that bear the contact force is elastic. The effect of plastic deformation of a small percentage of these asperities on the suggested analysis is insignificant and, therefore, the parameter of plasticity will not be considered. Moreover, as the anisotropy is an inevitable characteristic inherent in the topography of any machined surface and depends greatly on the type of machining process used, it seems impractical to be taken into account for all the surfaces (which are probably produced by different processes) under gauging. As pointed out earlier, the

consideration of such a parameter could lead to an extremely complicated analysis of the asperity model. So, since a simplified analysis is the objective here, this parameter will also not be considered.

In regard to the contact adhesion, it seems very reasonable to follow the common experience, mentioned before, and assume that no adhesion force is taking place between the indicator's probe tip and the rough surface. This is because of the same two reasons noted previously, i.e., the difficulty of achieving perfect cleanliness on any real surface and the unavoidable presence of the roughness. In addition, it is sensible to assume that the interfacial friction due to deformations in the static contact between surfaces of dissimilar materials is of negligible effect. This is because the gauging forces of indicators are not adequately large to cause a considerable plastic deformation of asperities that encourages the tangential slips to occur, besides assuming that the indicator's probe tip has a perfectly smooth surface.

Although, contact indicators are usually available with precisely polished probe tips made from hardened steel (such as the contact tips used in the experimental investigations), it is conceivable that, for such a contact problem, the probe tip produces a relatively trivial deformation compared with that of the rough surface which probably belongs to a steel component. In addition to the low contact forces involved, the high continuity of the probe tip surface (compared to that of the rough surface under gauging) is another condition that is believed of benefit in preserving the rigidity of the probe tip during the contact. The consideration of rigid probe tips allows Equation 3.2 of the elastic modulus of contact to be reduced to

$$E^* = \frac{E}{1-\nu^2} \tag{3.8}$$

where E and ν are the material properties of the rough surface. As a result, the theoretical contact deformation becomes relatively smaller. This assumption will be

rigorously tested together with the suggested contact solution against some of the experimental results observed.

Two asperity-based models for the contact between a smooth spherical surface and a flat rough one have been adapted, as demonstrated in the next two subsections. They are: the Greenwood and Tripp (GT) model [88], and the Kagami, Yamada, and Hatazawa (KYH) model [89]. These models have been chosen for their simplicity, as they considered none of the parameters, discussed in the previous section as of little relevance, in their analysis. Additionally, they were introduced in two different periods in the history of the development of contact analyses of rough surfaces: the GT model directly followed the early important contribution of GW, while the KYH model followed a progress of nearly two decades. Hence, two different analyses of distinctive backgrounds have been used for the purpose of comparison with the experimental results of the thesis, seeking a simplified analysis for estimating the displacement errors due to the contact interaction which are included in the measurement when using contact indicators.

3.4.1 The GT Model

In this basic asperity-based model, two main assumptions have been made by the investigators to derive the relation connecting the separation between the mean surfaces with the pressure created by compressing the asperities. These are: (1) the asperity heights follow the normal (Gaussian) distribution about the mean surface, and (2) the tops of the asperities are spherical, all with the same radius, and that they deform elastically according to the laws of Hertz. The investigators pointed out that the essential point is that the behaviour of rough surfaces is determined primarily by the statistical distribution of asperity heights and secondarily by their mode of deformation. Although, it is clear that some of the higher asperities must suffer plastic deformation,

they found that pure plastic deformation leads to very similar results, and it is plausible that a combination of modes would introduce no new features. Their derivation resulted in the relation of the dimensionless pressure distribution, $p^*(r)$, with the dimensionless separation, u^*, which is

$$p^*(\rho) = \mu F_{3/2}(u^*) \tag{3.9}$$

where ρ is the dimensionless radial distance from the centre of the contact and equals $r/\sqrt{2R\sigma}$, $\mu = \frac{8}{3}\eta\sigma\sqrt{2R\beta}$, and $F_{3/2}(u^*)$ is given by

$$F_{3/2}(u^*) = \frac{1}{\sqrt{2\pi}} \int_{u^*}^{\infty} (s - u^*)^{3/2} \exp\left(-\tfrac{1}{2}s^2\right)\, ds \tag{3.10}$$

Here, r is the radial distance, R is the radius of the sphere, σ is the standard deviation of the height distribution and can be regarded as the roughness measure, η is the spatial density of asperities, and β is the asperity radius. The dimensionless separation, u^*, is given by

$$u^* = d^* + \rho^2 + w^*(\rho) - w^*(0) \tag{3.11}$$

where $d^* = d/\sigma$ is the dimensionless minimum separation between the mean lines of the two surfaces, and $w^*(\rho)$ is the dimensionless bulk displacement of the nominal surfaces at ρ from the centre and determined from Equations 7a and 7b in [88].

The investigators indicated that the movement of the centre of the sphere caused by an increment in the load is the change in $[w^*(0) - d^*]$; at light loads the movement is due to the squashing of asperities (the d term), while at heavier loads this becomes negligible and the bulk elastic deformation (the w term) dominates. Hence, it is safe enough to assume that, at the low regime of gauging forces exerted by the contact indicators, the bulk elastic deformations of the probe tip and the subsurface of the component being measured are of negligible effect. This assumption reduces Equation 3.11 to

$$u^* = d^* + \rho^2 \tag{3.12}$$

and Equation 3.9 to

$$p^*(\rho) = \mu F_{3/2}\left(d^* + \rho^2\right) \tag{3.13}$$

Furthermore, since the movement of the centre of the sphere is basically equal to the maximum deformation of asperities, which appears at the centre of the contact area (i.e., at $r = 0 \Rightarrow \rho = 0$), Equation 3.12 above becomes even more simple, i.e., $u^* = d^*$ (meaning that the minimum separation is located at the centre of the contact area).

The dimensionless total load, T, has been defined as

$$T = \int_0^\infty 2\pi\rho p^*(\rho)\, d\rho \tag{3.14}$$

so that the actual load, P, equals $\frac{1}{2}\sigma E^* \sqrt{2R\sigma} T$. Solving this equation for T will give another form for Equation 3.14

$$T = 2P / \sigma E^* \sqrt{2R\sigma} \tag{3.15}$$

Comparing Equations 3.14 and 3.15,

$$2\pi \int_0^\infty \rho p^*(\rho)\, d\rho - \frac{2P}{\sigma E^* \sqrt{2R\sigma}} = 0 \tag{3.16}$$

with $p^*(\rho)$ determined by substituting Equation 3.10 into Equation 3.13. The only unknown is d* included in the pressure distribution function $p^*(\rho)$. The ρ term will obviously be substituted after solving the integral of the total load. However, by using any zero-finding numerical method, Equation 3.16 can be solved iteratively.

Following recommendations in texts on numerical algorithms, the Newton-Raphson method was judged appropriate for use here and the algorithm demonstrated in [90] was followed to write the iterative routine of calculations. The routine was developed in the formats of the MathCAD 2000 Professional software which was used on a Pentium III personal computer of a processor speed of 850 MHz. Appendix A1 shows this simple routine with a sample of results. With any reasonable initial guess for

d^*, convergence to the exact value of d^* (that satisfies Equation 3.16 with an accuracy of 10^{-12}) can be attained within fifteen seconds and in less than ten iterations on this computer. Then, the maximum deformation of asperities is computed by multiplying the converged value of d^* by the roughness σ (to obtain the non-dimensionless minimum separation), and subtracting the result from 2σ. It is assumed here that the first contact with the highest asperities happens at a distance of 2σ above their mean line (where there is about 2.5% of the surface material) and, hence, the deformation was calculated starting from that level.

3.4.2 The KYH Model

The investigators here assumed a different statistical distribution for the asperity heights about the mean surface, but used the same assumption for the summits shape of asperities. They did not follow the Hertzian laws in treating the elastic deformation of asperities. Instead, they considered a mixed asperity contact theory in which asperities with displacements below and above a critical value are deformed elastically and plastically, respectively. They also analysed contact asperities under pure elastic or pure plastic deformation modes, independently. Some differences between the three analytical results have been found and confirmed by their experimental investigations.

The model is based on the comparison between an assumed ideal pressure distribution (for a smooth contact) and another pressure distribution that considers the asperity contacts [91]. In other words, the latter pressure distribution must match the former one; this matching is done iteratively using the numerical methods. The investigators selected the following ideal pressure distribution

$$q_I(r) = q_0\left[\exp\left\{-\left(\frac{r}{b}\right)^2\right\} - \exp\left\{-\left(\frac{a}{b}\right)^2\right\}\right] \tag{3.17}$$

where q_0 is the maximum pressure at the centre of the contact area, r is the radial

distance from the centre of the contact area, a is the radius of the circular contact area, and b is a parameter related to the shape of the distribution. The contact load, P, could be obtained from

$$P = \int_0^a r\, q_1(r)\, dr \tag{3.18}$$

The results of this integral could be used to solve for q_0 to give

$$q_0 = \frac{P \exp\left\{(a/b)^2\right\}}{\pi b^2 \left[\exp\left\{(a/b)^2\right\} - (a/b)^2 - 1\right]} \tag{3.19}$$

The investigators defined the compliance α between the sphere and the rough plate at a distance r from the centre of the contact area as

$$\alpha = \frac{r^2}{2R} + u(r) + w(r) \tag{3.20}$$

where R is the radius of the smooth sphere, $u(r)$ is the asperity deformation, and $w(r)$ is the elastic deformation of the sphere and the subsurface of the plate. If, as assumed in the GT model, the bulk deformations are of insignificant magnitudes at the low loads of the contact indicators, the $w(r)$ term could be eliminated from the previous equation. Moreover, when $r = a$, $u(r)$ becomes zero (i.e., $u(a) = 0$) and, then, $\alpha = a^2/2R$. Substituting this result back in Equation 3.20 and solving for $u(r)$, after eliminating $w(r)$, will give

$$u(r) = \frac{a^2 - r^2}{2R} \tag{3.21}$$

For the elastic contact case, the investigators followed the same derivation, as used in the GT model, to obtain the pressure distribution due to the asperity contacts which is given by

$$q_2(r) = \tfrac{4}{3} E' \eta \beta^{1/2} \int_0^{u(r)} \{u(r) - z\}^{3/2} f(z)\, dz \tag{3.22}$$

where E' is the elastic contact modulus given by Equation 3.2, η is the density of

asperities, and β is the mean radius of curvature of the asperity peaks. The distribution of the asperity heights of the rough plate, $f(z)$, is assumed by KYH to be the Rayleigh distribution which is defined in [92] as

$$f(z) = 2\varepsilon z \exp\left(-\varepsilon z^2\right) \tag{3.23}$$

where ε is equal to $(4-\pi)/4\sigma^2$; σ is, as before, the standard deviation of the asperity heights. By substituting Equations 3.21 and 3.23 into Equation 3.22, the radius of contact, a, will be the only unknown in the $q_2(r)$ distribution which, as mentioned earlier, should be identical to the ideal one, $q_1(r)$ (where both a and b are unknown).

Hence, the two parameters, a and b, must be determined such that these two distributions match each other. The investigators repeated the calculations, by the Newton-Raphson method and using several values for a and b, until $q_1(r)$ became equal to $q_2(r)$ at the two radial distances, $r = 0$ and $r = \frac{1}{2}a$. The same procedure was followed here to implement the suggested approximations in this model. So, the two equations, which are to be solved simultaneously to find these two unknowns, could be given as

$$q_1(0) - q_2(0) = 0 \tag{3.24a}$$

$$q_1\left(\tfrac{1}{2}a\right) - q_2\left(\tfrac{1}{2}a\right) = 0 \tag{3.24b}$$

The modified Newton-Raphson method for two variables was used, and the algorithm demonstrated in [93] was followed to write the iterative routine of calculations in the formats of the MathCAD 2000 Professional software. This routine is shown in Appendix A2 with a sample of results. A sensible initial guess for a and b could lead the calculation to converge to the solution of Equations 3.24 (with an accuracy of 10^{-12}) within ten seconds and in less than ten iterations on the personal computer used before.

It worth to noting here that when dealing with the exact solution of the full KYH model, the calculations time for just three iterations was observed to be nearly eight

minutes, with a reduced accuracy of 10^{-6}. This was when both of the initial guesses were chosen very carefully; from experience, close to the expected solution to within ±10%. Guesses greater than this limit could result in the convergence in a period of time of, at least, fifteen minutes through six iterations only using the pc mentioned before. But, in most of the occasions, such arbitrary guesses could lead to divergence after spending a considerable amount of time iterating. This difficulty in computation is obviously a result of the double evaluation of the bulk deformations of surfaces; when not eliminating the $w(r)$ term from Equation 3.20, $u(r)$ in Equation 3.21 becomes $\left(\left(a^2 - r^2\right)/2R\right) + w(a) - w(r)$. Additionally, the derivation of $w(r)$, which is based on the theories of elastic deformations between a sphere and a flat described in [94], involves a double integral with respect to the two coordinates used to define the circular contact area. As it is misprinted [91] in Equation 4 in reference [89], the correct form is

$$w(r) = \frac{2q_0}{\pi E'} \int_0^{\pi} \int_0^{r\cos\varphi + \sqrt{a^2 - r^2 \sin^2\varphi}} \exp\left(-\frac{r^2 + s^2 - 2rs\cos\varphi}{b^2}\right) - \exp\left(-\frac{a^2}{b^2}\right) ds\, d\varphi \tag{3.25}$$

3.4.3 Comparison between the Modified GT and KYH Models

The two calculation routines written for the approximations suggested in the previous section were employed to establish a comparison between them at different material and surface parameters. Three material properties were used for this purpose: mild steel (E = 210 GPa and ν = 0.29), copper (E = 117 GPa and ν = 0.34), and aluminium (E = 71 GPa and ν = 0.34) [95]. The experimentally determined values of the three surface parameters (β or R, η or n_0, and σ) shown in Table 1 in reference [89] were also used assuming that such parameters are producible using each of these materials. The first four surfaces (of σ of 0.08μm, 0.18μm, 0.46μm, and 1.45μm) were used to observe the roughness effect in each model. The assumption of a rigid smooth sphere (probe tip)

discussed earlier, in which E^* could be regarded as given in Equation 3.8, was taken into consideration here.

Figure 3.2 shows plots of the asperities deformation at the centre of the contact area computed by each model approximation under the four roughness values and a sphere diameter of 5 mm for the mild steel material. The corresponding Hertzian deformation curve (for a rigid indenter) is also shown in this figure. It is very obvious from this figure that as the roughness parameter σ decreases, these calculated deformations by both approximations also decrease and approach the Hertzian deformations of the ideal smooth surfaces. The difference between the deformation curves of the two approximated models tends to be more significant as the roughness parameter increases. The GT approximation showed an instability at the highest roughness value and at contact loads below, say, 2.0 N, since its estimated deformations went sharply down to below zero deformation at positive contact loads. It seems that the assumption of negligible bulk deformations did not work well with the GT analysis as with the KYH analysis. From the iterative solution of the GT approximation at this region of high roughness and very low loads, it was observed that Equation 3.16 gives values of d^* ($= d/\sigma$) larger than 1. This means that d (the minimum separation from the asperities mean line) becomes near to 2σ, and the asperities deformation becomes implausibly very small compared to that obtained by the KYH approximation. As the roughness decreases and/or the load increases, d^* decreases to below 1; the GT results become more stable, and the deformation tends to be more reasonable and closer to that computed by the KYH approximation. However, this instability is responsible for the significant difference between the two approximations.

To examine the discrepancy between the two approximated models on the other two materials, the change of the percentage difference between the two computed deformations with the contact load is illustrated in Figure 3.3 for the three materials and

a sphere diameter of 5 mm. As shown in this figure, the difference decreases systematically with the decrease of E^* at every roughness value. This decrease became almost insignificant at the lowest roughness value. The percentage difference descended asymptotically as the contact load increased; i.e., the difference in the estimated deformations between the two approximated models remained nearly constant at high contact loads and at the given surface roughnesses and materials. A systematic decrease in this percentage is also expected with the decrease of the radius of the sphere. Figure 3.4 illustrates this fact with three different diameters for the sphere and under the four roughness values on the aluminium material. Similarly, the decrease becomes nearly negligible at low roughness values. The percentage difference descends asymptotically with the contact load under the surface properties used.

However, it is believed that this difference becomes negligible at very high contact loads (the consideration of bulk deformations is then necessary), regardless of the magnitude of surface roughness. This is since the pressure distribution suggested in each model has been demonstrated, by its investigators, to be nearly identical to the Hertzian pressure distribution at heavily loaded contacts. As quoted before, the effect of roughness on the contact pressure becomes almost unimportant.

3.4.4 Comparison with the Experimental Results

In order to determine how well these approximated models of contact represent the actual situations of concern in this study, some of the data computed from these two models were compared with appropriate experimental results selected from Chapter 6. The average resultant surface deflections obtained from the different-point repeated contact tests on the *St1*, *Co2*, and *Al3* surfaces, shown in Table 6.3, with a contact tip of 5 mm diameter are each plotted in Figure 3.5 together with the surface deformation curves calculated by each approximation. With a contact tip of 3 mm diameter, the

same type of deflections observed on the *S1C*, *A1C*, and *A2C* surfaces, given in Table 6.9, are also plotted in Figure 3.6 in the same way. The roughness value of each of these six surfaces is near to one of the four roughness parameters used for the comparison in the last subsection.

From all of the six plots in these two figures, it is clear that the surface deformations observed experimentally are, in general, within or below the estimated ones computed by the two approximated models. This justifies the assumption that there are no bulk deformations contributed by both the spherical contact tip and the specimen's subsurface at such low regimes of contact force. Obviously, the original analysis in both models suggests higher magnitudes of deformation than the approximated analysis due to the consideration of these bulk deformations.

It is also clear that both approximations are generally overestimating the actual deformations at contact forces below 0.5 N. As shown, this is observed with all the three materials and the two probe tip sizes used within the range of roughness magnitudes given. The trend of surface deformations seems likely linear with contact loads and of a smaller slope than those of the estimated trends. In Figure 3.5a, the instability of the GT approximation at this high roughness value, mentioned before, on mild steel might have aided an apparent good agreement with the experimental data, but it failed to estimate any deformations below 0.2 N load. With the 3 mm tip and the aluminium, in Figure 3.6c, the reduced effect of this instability allowed the GT approximation to overestimate again the real deformations below the 0.5 N load. But beyond this load, the experimental deformations showed a good agreement with the predictions as illustrated in Figures 3.6a and 3.6b. The observed deformation at the highest load used on the *A2C* surface, shown in Figure 3.6c, seems somewhat higher than expected, but the trend of deformations on this surface is fairly unlike that observed on the others. So, it is believed that this agreement will carry on to, at least,

the 2 N or 3 N range.

It worth recalling from Chapter 6 that the initial contact dead-load adjusted during the tests reported in Table 6.3 was around 0.2 N, while it was below 0.05 N during those reported in Table 6.9. This means that the experimental average deformations plotted in Figure 3.5 are lower than the actual asperity deformations by that amount of deformation caused initially by that initial load. This also explains why some of these deformations are very near and even below the Hertzian predictions, as shown in Figures 3.5b and 3.5c. Hence, the agreement of these deformations with the approximations, above the 0.5 N load, is also expected. The observed average deformations plotted in Figure 3.6 are, however, closer to the actual ones due to the lower initial dead-load adjustment.

It also worth pointing out here that the other two surface parameters (β and η), which were deliberately overlooked in this comparison, have some effects. The surface deformations obviously decrease as either one of them increases. Each of the selected real surfaces for this comparison was assumed to have similar magnitudes of these two parameters as the calculated one of the nearest roughness value. At much higher regimes of contact load than of interest here, these effects have been noticed by GT to be greater than at low regimes of load. Moreover, they are minor compared with the effect of σ, which led GT to conclude that the assumption of equal spherical asperities randomly located is not critical [88]. Hence, the comparison established above is valid to a large extent.

To conclude based on the available results, the approximations made in the GT and KYH models could provide very conservative predictions below 0.5 N and at all the roughnesses, materials, and probe tip sizes used in this investigation. Above this load level, the accuracy of their predictions could vary depending on the contact conditions involved. Since the difference between the approximations favours the GT or KYH

ones at different loads, it might be considered rigorous to compute both and to use the larger result as the conservative estimate. Although it has been recorded by Greenwood *et al* [96] that the experimental results of KYH did not correlate well with their findings and showed unknown greater discrepancy from the Hertz theory (at contact loads higher than of interest here), the approximation in KYH model was noticed providing better estimates than the approximation in GT model at high roughness magnitudes and within the regime of loads of concern. If only one model is used, the approximate KYH is recommended for contact gauge error estimation.

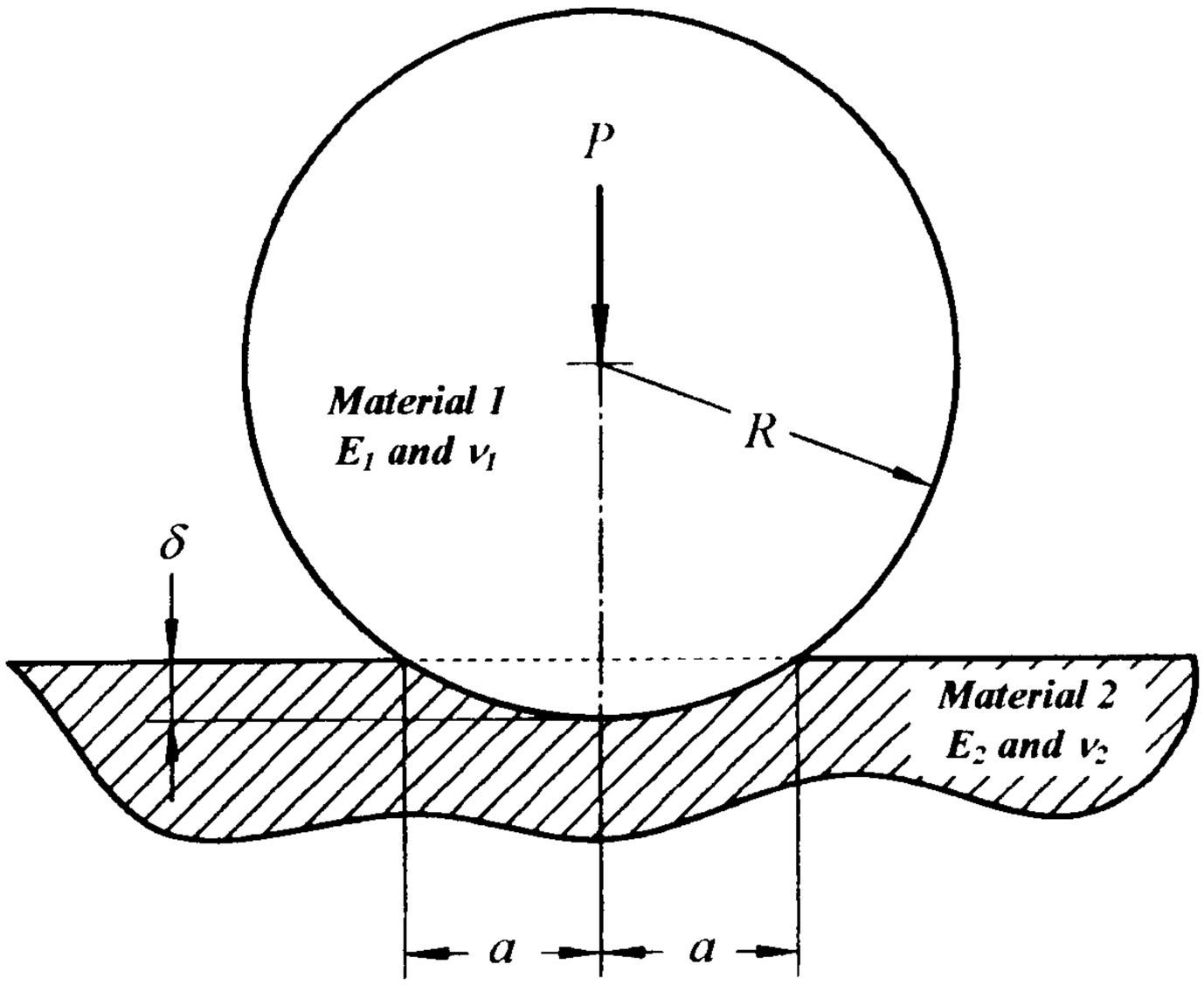


Figure 3.1 The normal contact between a sphere and a nominally flat smooth surface.

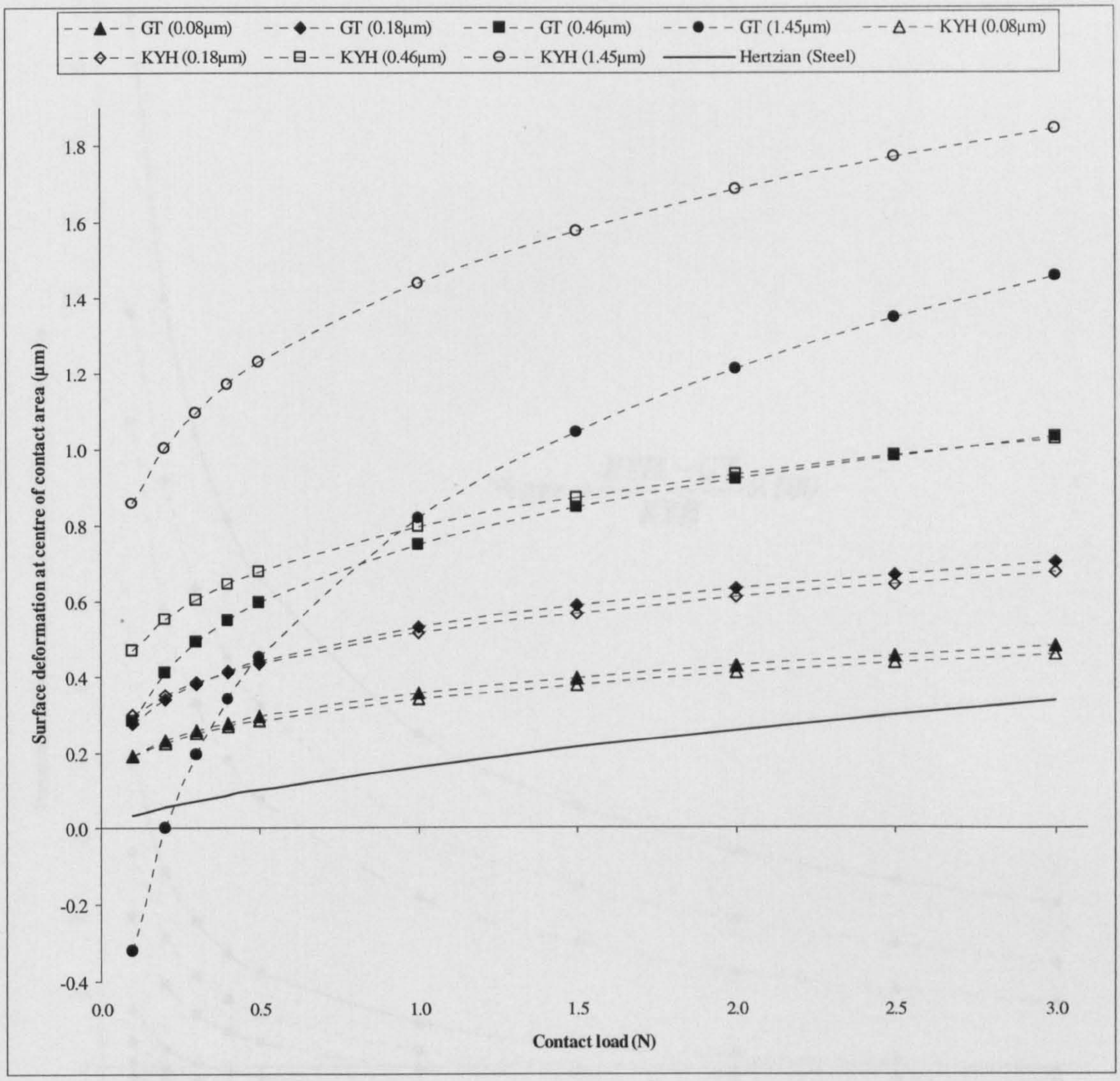
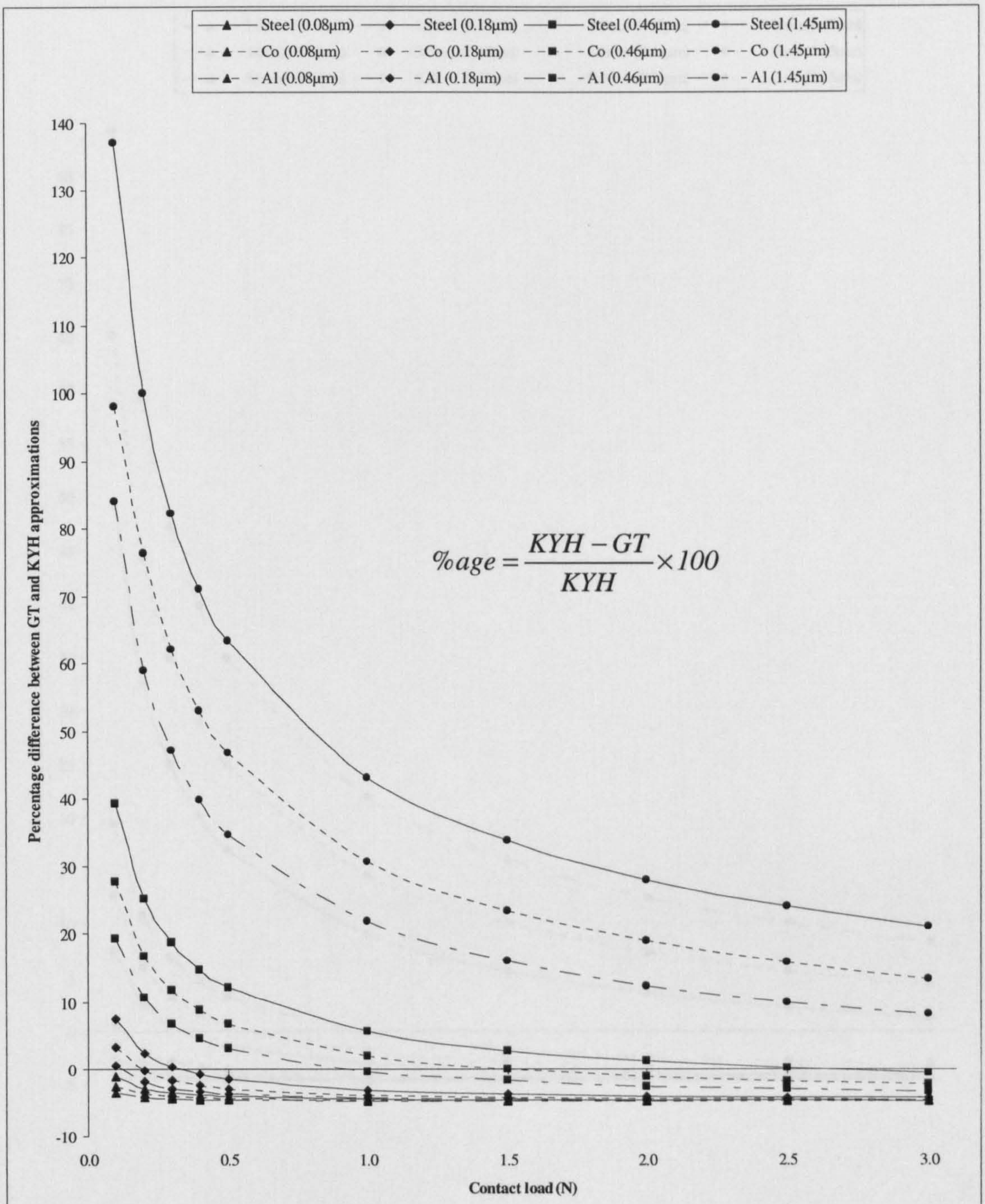
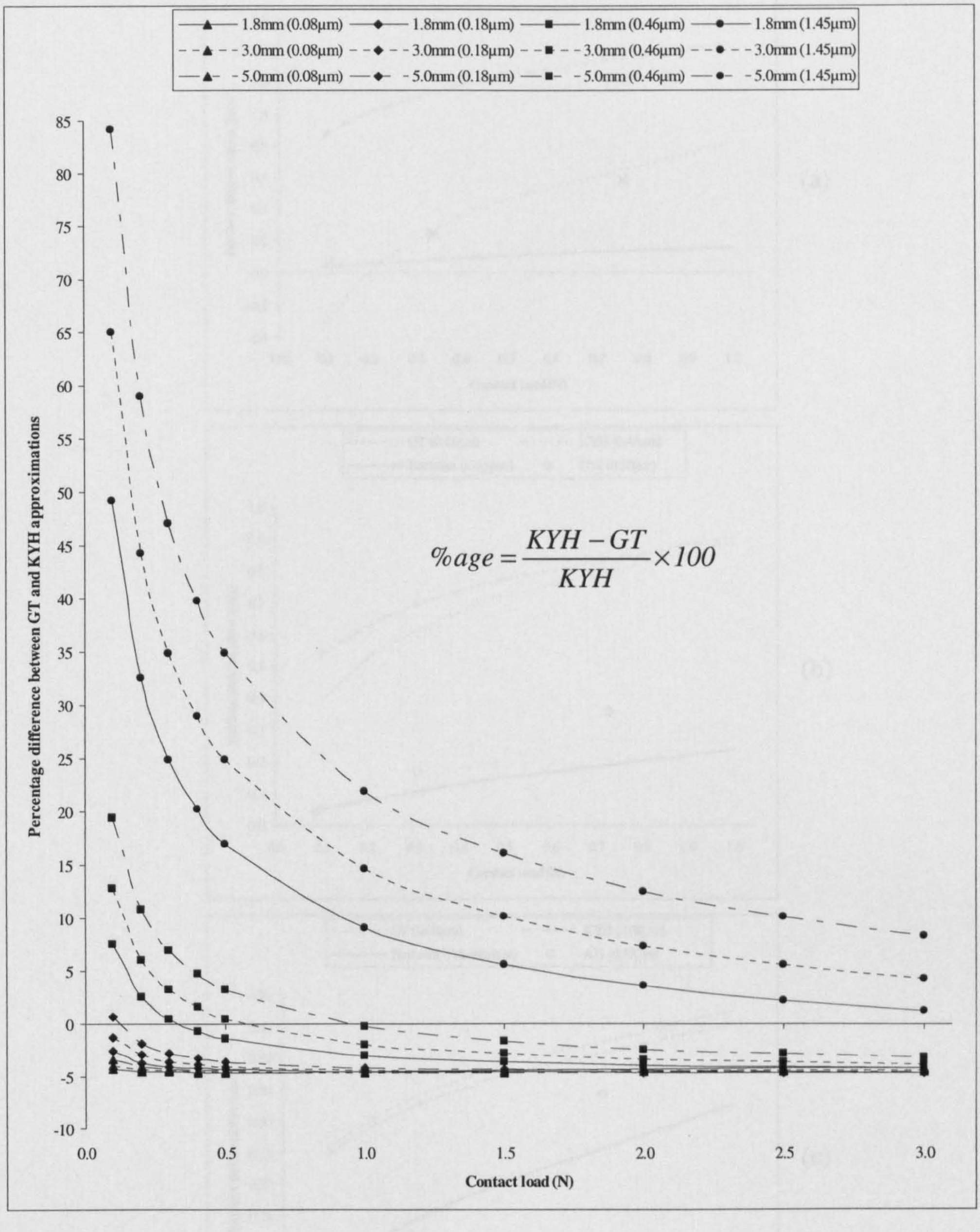


Figure 3.2 Comparison of surface deformations at the centre of contact area calculated by the two approximated models, under four roughness values, a probe tip of 5 mm diameter, and the "mild steel" material.



$$\%age = \frac{KYH - GT}{KYH} \times 100$$

Figure 3.3 Change of percentage difference between the GT and KYH approximations under four roughness values, three materials, and a sphere of 5 mm diameter.



$$\%age = \frac{KYH - GT}{KYH} \times 100$$

Figure 3.4 Change of percentage difference between the GT and KYH approximations under four roughness values, a sphere of three sizes, and the "aluminium" material.

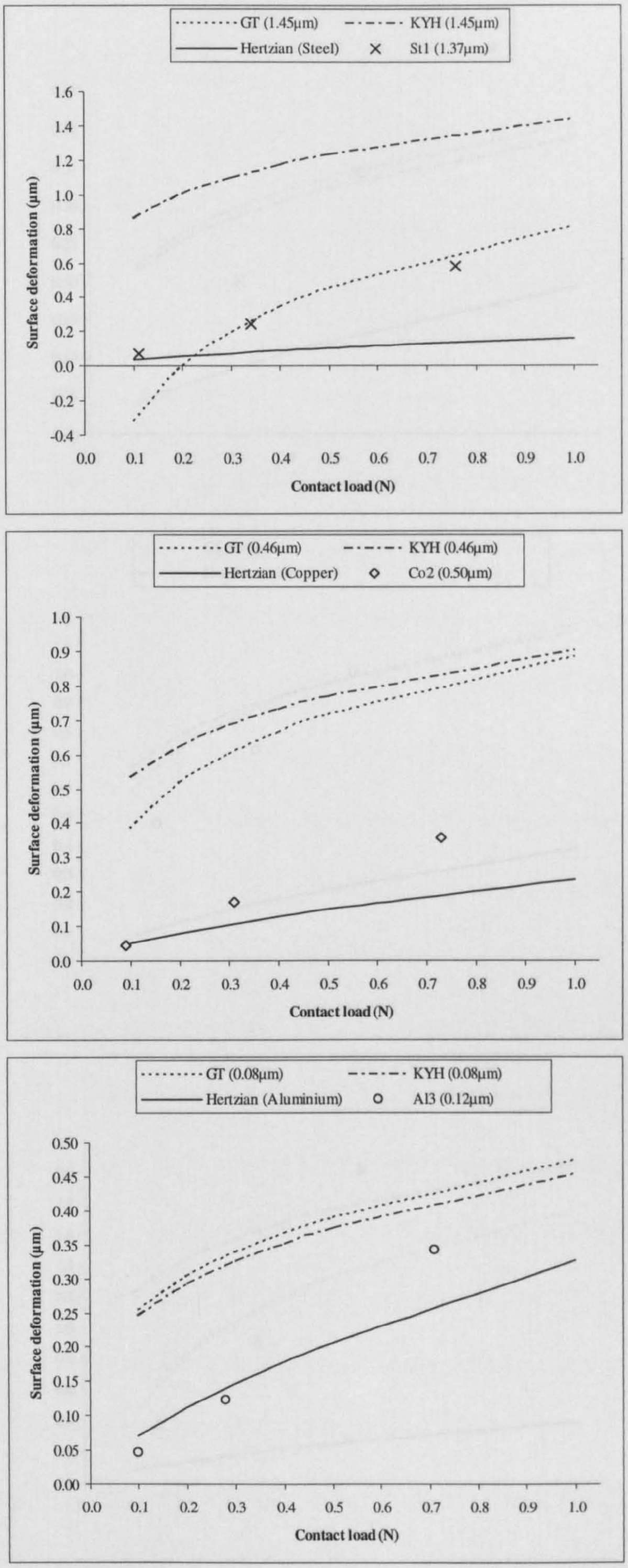
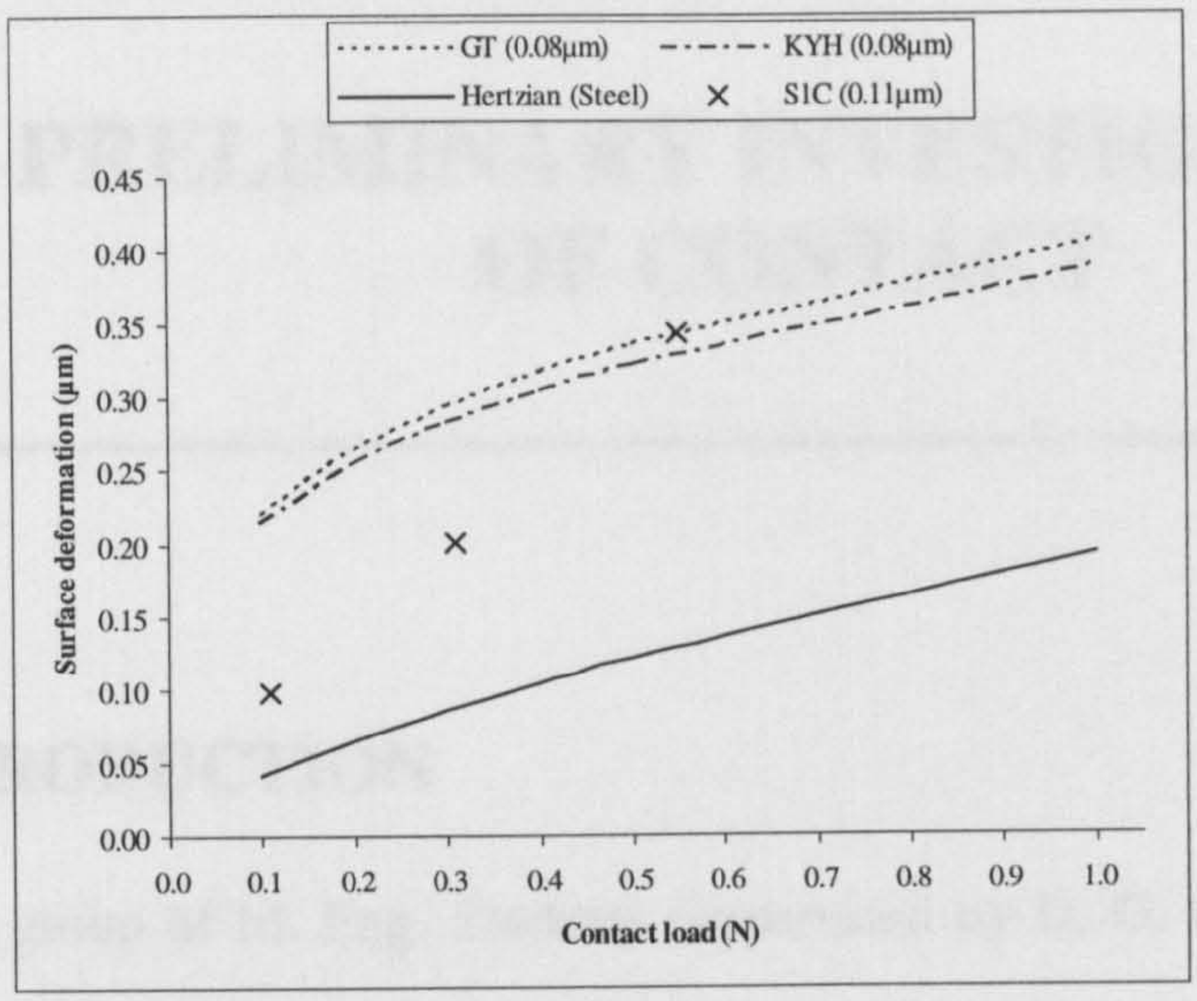
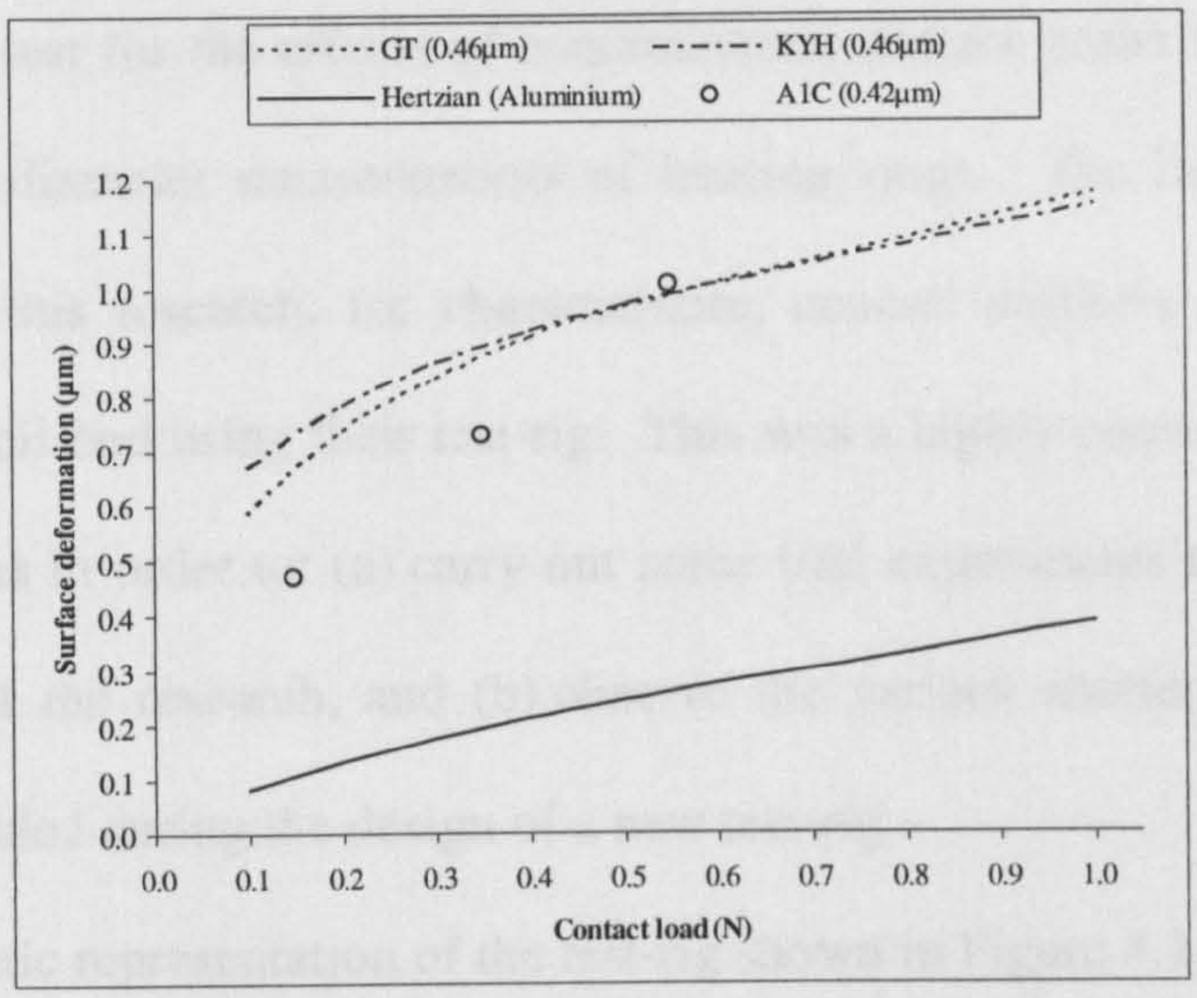


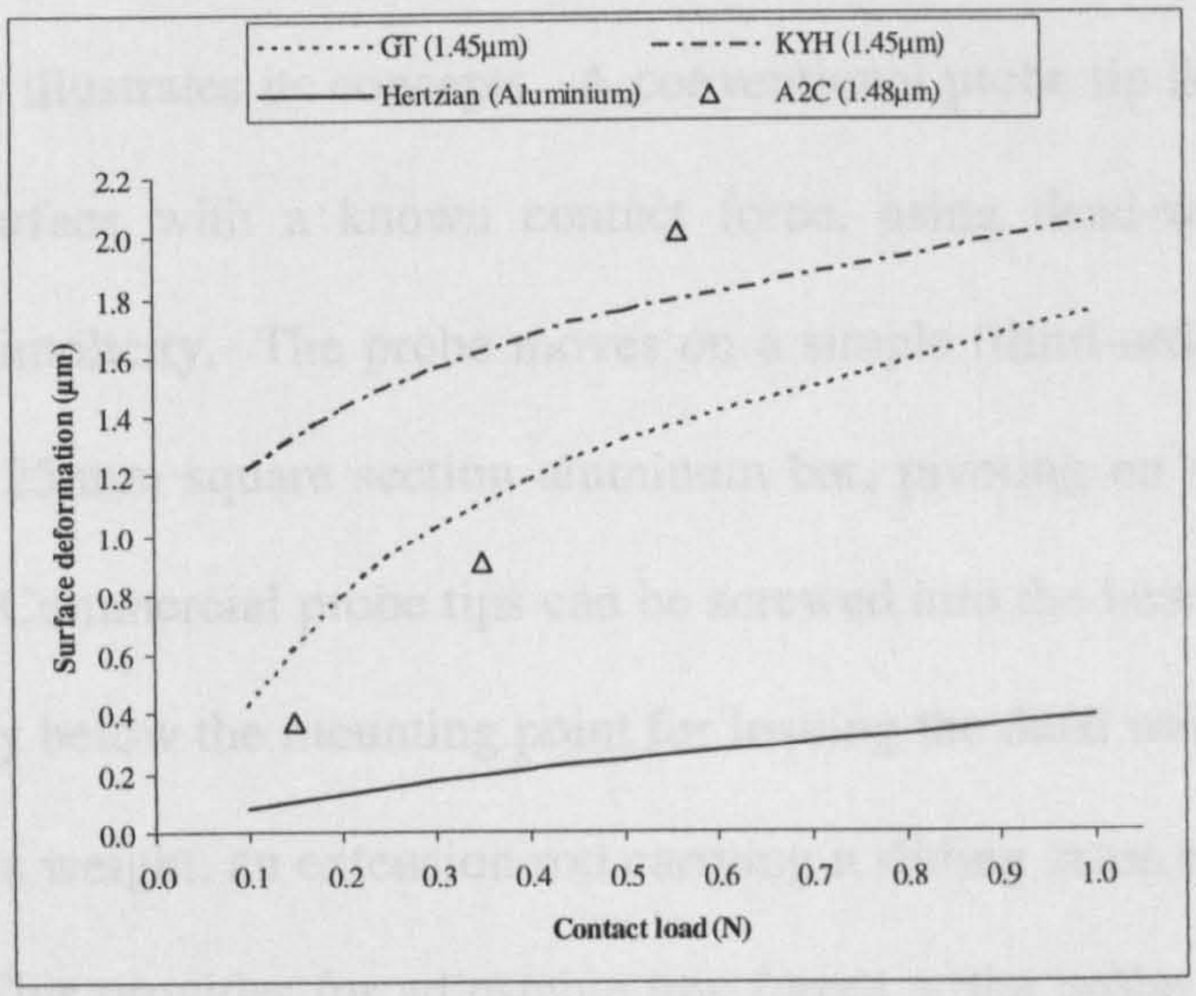
Figure 3.5 Comparison between the calculated surface deformations and the observed ones on (a) *St1*, (b) *Co2*, and (c) *Al3* surfaces with a contact tip of 5 mm diameter.



(a)



(b)



(c)

Figure 3.6 Comparison between the calculated surface deformations and the observed ones on (a) *S1C*, (b) *A1C*, and (c) *A2C* surfaces with a contact tip of 3 mm diameter.

CHAPTER 4

PRELIMINARY INVESTIGATIONS OF CONTACT

4.1 INTRODUCTION

In spring 1998 a group of M. Eng. Students (supervised by D. G. Chetwynd) built a simple test-rig to test for the effects of contaminants, contact probe types, and contact force relevant to diameter measurements of bearing rings. The initial experimental investigations in this research, for characterizing contact stiffness and repeatability, have been accomplished using their test-rig. This was a highly convenient way to start such investigations in order to: (a) carry out some trial experiments to make certain of the basic ideas of the research, and (b) observe the various shortcomings and errors sources to be avoided during the design of a new test-rig.

The schematic representation of the test-rig shown in Figure 4.1 is a little different to their design but illustrates its concepts. A conventional probe tip is pressed against a horizontal test surface with a known contact force, using dead-weight loading for consistency and simplicity. The probe moves on a simple (third-order) lever, 175 mm long, made from 25 mm square section aluminum bar, pivoting on a ball bearing and short steel shaft. Commercial probe tips can be screwed into the beam 90 mm from the pivot, immediately below the mounting point for loading the dead weights. To counter-balance the lever's weight, an extension rod carrying a sliding brass cylinder is fitted to the lever's tail. This provides for adjustable low forces at the probe tip for finding the initial contact on the specimen's surface.

The rest position of the tip relative to the test surface is measured by a capacitive

micrometry displacement gauge. One electrode is mounted below the lever, glued to the lower end of a vertically-sliding rod used for adjusting readings that is centered 130 mm from the pivot. The other is fixed rigidly to the same base as the pivot and specimen. The electrodes are cut from a copper-clad glass-fiber printed circuit board into squares of 15 mm side. The copper cladding was left entirely covering the one fitted to the lever, while it was partly removed to leave an area of 10 mm square in the middle of the lower one. Hence, the gauge has a nominal area of 100 mm^2 and separation of around 100 μm and is used with the electronics from Queensgate Instruments for commercial gauges.

In general, the overall design of this rig combines cheapness with simplicity in manufacturing. It also provides a clear and easy access for probe tip and specimen cleaning and replacements. The use of leverage mechanism created a better resolution for the capacitive gauge, since the genuine movements at the probe tip are amplified at the gauge by a factor of 13/9; while the pivot error effects are restricted, when the probe tip acts as a fulcrum, since the dead-weight load acts in-line with the probe tip.

4.2 MODIFICATION

After the first few tests made on the original test-rig towards its initial characterisation, there was clear need of modifications in order to get reliable measurements. The stability and repeatability of the capacitive gauge readings were fairly poor. Moreover, the rig was designed to cope with a very small range of size of specimens, which did not offer good flexibility in using specimens of different thickness without readjusting the gauge electrodes. The following easily-made modifications were therefore applied, giving the form shown in Figure 4.1.

- The ball bearing set installed at the pivot has been replaced with a more precise

set, for a better pivot accuracy; hence, a new steel shaft has been installed.

- The dead-weight loading point has been redesigned to look as in Figure 4.1, allowing a better concentration of the loads applied on the probe tip.
- A specimens' block, screwed to the base plate, was removed to reduce additional systematic errors due to its presence and to permit specimens of larger thickness to be tested.
- The gauge electrodes have been re-wired with shielded cables, since their original wiring conditions were affecting the consistency of readings.

It should be noted that these modifications were proposed to alter specific features on the rig that were producing highly significant errors in the resultant measurements. Yet, there are some inherent errors which could not be eliminated, as discussed in the following sections.

4.3 CHARACTERISATION

Clearly, the modifications applied on the rig are expected to present it with new error conditions. In other words, the installation of the new precise set of ball bearings at the pivot and re-wiring the capacitive gauge electrodes must have affected the accuracy performance of both the pivot and the gauge. Thus, the relative movements of the gauge electrodes have to be re-calibrated and the pivot errors have to be reviewed.

4.3.1 Calibration of Capacitive Gauge

The vertical movements of the probe tip were obtained by the use of gauge-blocks combinations available in the laboratory. Adjustments were made to the rod of the gauge upper electrode in order to read an output voltage just around the lower value of the expected range of linearity, approximately, –3.5 V. This was done in conjunction

with the use of a suitable sized ground steel block underneath the first combination of gauge-blocks. The gauge-blocks were then arranged to provide the probe tip equal displacement increments of 0.01 mm each, until the capacitive gauge showed a reading around +3.5 V. As a result, forty readings were collected and analysed; the best linearity of the data was found to be between −1 and +2 V, see Figure 4.2.

From the equation given in this figure for the best-fit straight line, which represents the most linear range of data, the calibration value of the capacitive gauge can be calculated to be around 65 nm/mV. So, every 1 mV difference at the capacitive gauge display corresponds to a change of 65 nm in the vertical position of the tip.

4.3.2 Evaluation of Pivot's Errors

Pivot rotation and stiffness are the dominant sources of error that cause some uncertainties in the measurements obtained using this test-rig. These errors cannot, in practice, be avoided completely, but they could be fairly estimated analytically and then eliminated from the end results. Hence, it is reasonable to assume that they were being included during the calibration of capacitive gauge.

The rolling elements in the two ball bearings appeared to cause stability errors for the pivot. Consequently, these errors will be reflected on the capacitive gauge readings. This has been clearly observed when repeated measurements, of the same nominal contact point on a smooth steel surface, were performed by lifting, replacing, and loading the probe tip. Three sets of data were taken (each consists of 25 readings) at four different lifts and using a 3 N load on the probe tip. The sample standard deviation σ_{n-1} for each data record is then computed and plotted as shown in Figure 4.3.

To a certain extent, the pivot errors caused by the bearing rolling elements tend to increase with the probe lift increase. The last three probe lifts in this figure showed nearly the same average standard deviation, which is about 2.5 times higher than the one

showed by the 2 mm probe lift. Obviously, a large probe lift allows the balls in the two bearings to have a new random arrangement; hence, a new different pivot error is created and transmitted to the capacitive gauge. A weaker repeatability of the readings is more likely to occur at higher lifts. Small probe lifts tend to obstruct the chance of further ball re-arrangements and a tighter scatter of readings could be noticed from the capacitive gauge.

More tests have been carried out for a complete evaluation of the pivot errors performance, which seems to affect the repeatability of readings. Independent repeated measurements have been done on each side of the pivot in order to quantify the clearance of each bearing along the vertical directions. These measurements used a dismounted probe from a Taylor-Hobson Talysurf-5 instrument (TS5) as an independent displacement sensor to pick up the local vertical displacements at the left and right bearings, separately, while repeatedly rotating and replacing the lever, and then pulling the pivot upward by a 7 N force. A load of 3 N was kept on top of the contact tip to ensure its action as a fulcrum.

The results of fifteen rotate and pull tests showed a fairly poor repeatability of vertical displacements at both bearings, with a mean value and full-range (95% confidence) scatter of about 27 μm and ±1.5 μm, respectively, at the left bearing; and around 13 μm and ±1.0 μm at the right bearing. The left bearing seemed having a dominant role in contributing the final resultant pivot errors, since it produced over double the displacements that produced by the right one. However, the bearing could be defective or have been over-stressed during its installation.

One should note that these measured displacements are inclusive of the shaft elastic deflections (bending), since the two bearings are not located at the fixed ends of the shaft. These could be easily computed using simple force analyses and then eliminated. Of course, having done so will not influence the repeatability of the bearing

errors, but it would give more accurate estimates for them.

In fact, using this test-rig in the way described in the above test represents a case, which could never be attained during normal use. It is just as if a load of about 16 N was being applied at the capacitive gauge location after rotating the lever and replacing it back, considering negligible lever deflections. The aim for doing such tests on this test-rig was to uncover the pivot errors in their extreme values and to use them for comparison purposes.

Measurements of the pivot resultant vertical displacement have been recorded, when repeatedly lifting and replacing the lever and then applying the 7 N load on the contact tip, which it represents the normal way of using the rig. The TS5 probe has been repeatedly placed on top of the lever and immediately above the pivot's shaft and centered between the two bearings. Fifteen readings of such displacements have been collected and they were widely dispersed, nearly ±26 nm about their mean value of around 24 nm, but dramatically low compared to those obtained from the previous test. While in principle such small displacements are within the capability of the TS5 probe, in practice the uncertainty will be high and the numbers are best considered as order of magnitude values. Clearly, this test showed a gigantic difference between the full errors of the pivot and those which transmitted from the pivot to the capacitive gauge during the actual use of the test-rig. However, the maximum displacement, which has been noted from these readings, was about 50 nm. This includes a lever-reduced fraction of 4/9 of the tip deflection since the measuring point is above the pivot line, suggesting that the effect of real pivot displacement at the capacitive gauge will be below 25 nm under the normal operational conditions. As before, these resultant displacements also contained the shaft deflections, which are expected to be much smaller, since the load was not being directly applied on the pivot.

Another type of test was run on the pivot to evaluate its effect on the capacitive

gauge readings, by repeatedly applying the loads on the pivot after lifting and replacing the lever. It is another abnormal way of using this test-rig, but useful to give a complete picture about all the possible errors that could emerge from the pivot. Simultaneously, twenty readings from both of the capacitive gauge and the TS5 instrument were taken for a repeated load of 3 N placed on top of the pivot and centered between the shaft ends. The TS5 probe was kept on top of the lever and as near as possible to the gauge location; and a 0.5 N load was maintained continuously on the probe tip. Both the gauge and TS5 readings showed nearly the same mean value of 195 nm and 184 nm, respectively, for the upward displacements at the gauge location due to loading the pivot. In general, they also showed good repeatability with σ_{n-1} of approximately 11 nm for the TS5 readings and about four times this deviation for the gauge reading. Nearly, the same standard deviation (of about 13 nm) also resulted from TS5 readings for a repeated load on the pivot of 5 N, using the same setup configuration described above. The mean displacement at the gauge location was 309 nm, which is consistent with the ratio between the two loads (i.e., $3/5 \approx 184/309$). Hence, the contact probe was acting as a fulcrum and the bearings' shaft at the pivot was responding to loads by bending.

This behavior was closely examined through gradually increasing the load on the bearings from 0.5 to 7.5 N and recording the displacements at the gauge location, while maintaining a 3 N force on the contact probe. Figure 4.4 illustrates the effect of load increase on top of the pivot on the resultant displacement sensed by the TS5 probe near the capacitive gauge. The change in this displacement is fairly linear with the change of load, as given by the results of correlation for the best fitted straight line ($R^2 \approx 0.99$), since the shaft was bending elastically (straining) as the load was increasing (stressing). Obviously, the stress-strain relationship is linear within the elastic region. Deviations from linearity are believed to be due to other sources of errors such as clearance of the ball-bearings and specimen indentation (where a different part of each load applied on

the pivot would appear at the contact probe due to unavoidable eccentricity in loading).

4.3.3 Checking of Deflections of Specimen's Base

It is clear from the design of the test-rig that the loop of contact force and the metrology loop of contact deflection are coincident between where the contact probe joins the lever and where the specimen meets the base plate. So, as in conventional hardness testing, the deflection of the specimen relative to the base is included in the measurement. Since the contact load is spread at this interface over a much larger area than at the indenter, such deflection is normally assumed to be negligible. However, this assumption has been confirmed, on several specimens within the loading range to be used, by examining the deflection of some points on the top surface of the specimen a few millimeters from the contact area.

For this purpose, the dismounted TS5 probe was mounted on the base plate and its stylus was placed at random points on the top surface of a steel specimen, 5 to 8 mm from the centre of the contact probe. At this distance, deflection of the surface from the contact stress should be negligible, while Abbe effects will remain small. The deflection error generally increased as the probe force was increased to 7 N; the interface (mounting) stiffness at three different locations around the contact probe was between 0.6 and 1.0 GN/m. Contact tests on the same specimen showed typical values of stiffness between 4.4 and 5.6 MN/m. This means that the deflection of specimen's base is more than 100 times stiffer than those of typical test contacts.

4.4 ESTABLISHING PARAMETERS

Despite the still limited performance of the modified test-rig, it was used to check the magnitude of some major contact-related parameters that should be taken into

consideration when using contact-probe gauges in precision measurements. The tests performed for this purpose considered the repeatability of contact under different surface roughness, material, and surface condition. The behaviour of surface deflection with contact load was also checked under these different parameters. Other tests examined the influence of the probe tip size and tightening on this repeatability. These tests showed clear relations that are significant at the given errors status of the test-rig.

4.4.1 Load–Surface Deflection Tests

Four specimens of an identical size of 50×20×8 mm were prepared for these tests; two specimens are from mild steel and the other two are from aluminium. One of the two 50×20 mm surfaces of each specimen was ground such that there were two different surface roughnesses on each of the two materials. The average R_q roughnesses measured on the steel specimens, at 0.8 mm cut-off, were 0.78 μm and 0.21 μm; and on the aluminium specimens were 1.08 μm and 0.20 μm. These specimens were tested on top of a ground steel block of a suitable height that was screwed to the base plate of the test-rig in order to allow the probe tip to contact the surface of the specimen within the calibration range of the capacitive gauge. A screw-in steel Mitutoyo probe tip of a 5 mm diameter was used in all of these tests.

At the beginning of each test, an initial probe load of about 0.05 N was set on a randomly selected point on the surface being tested by adjusting the counter-balance weight of the lever. Dead loads were then placed on the contact probe (according to the sequence: apply load→ release it → increase to next load → apply it). Throughout the test, the probe was kept in contact with the surface. Hence, the dismounted TS5 transducer was used, for a better accuracy of deflection readings, mounted on the lever as near as possible to the capacitive gauge. Changes in voltage readings from this transducer were monitored and acquired by a PC using a customized routine written in

LabVIEW software and a data acquisition card from National Instruments Inc. The 12-bit analogue to digital converter was sampled at 4 ms intervals.

Figure 4.5 illustrates the plots of the load-deflection test for all of the four specimens, together with the Hertzian load-approach distance curves for steel and aluminium smooth surfaces. In general, all the plots nearly have shown the same trend as the ideal (Hertzian) one, specially, those plots of the surfaces of the lower roughness values. The consistent relation between surface deflection and both surface roughness and material is quite clear from the distinct difference between each plot, which can be easily recognized throughout the whole loading range used in these tests. The surface with the higher roughness of each material showed a greater deflection and, so, a bigger deviation from the corresponding ideal deflection of its material.

Over most of the test range it is reasonable to take a linear approximation of surface deflection with contact load. The graph for the smooth surfaces of steel and aluminium have a ratio of slopes of about 0.62, which is near to that of the ideal (Hertzian) curves, i.e., 0.65. This ideal ratio of slopes, obtained from the Hertzian load-deflection curves of steel and aluminium, is basically the ratio of the E^* values (based on steel indenter) of these materials raised to the power 2/3, as given by the Equation 3.3 of the Hertzian approach distance δ discussed in Chapter 3. But, as the roughness of the surfaces (of different E) being compared goes higher, the ratio of slopes of their load-deflection curves tends to be less than the ideal one. A ratio of about 0.47 was found between the slopes of the plots of the rough surfaces of steel and aluminium, although their R_q values are not as similar as those of the smooth ones. However, this observation will be examined in further advanced investigations.

Likewise, comparing the slopes from the rough and smooth surfaces gave nearly the same relation, within both materials, to the roughness ratio of the two surfaces. A ratio of slopes of about 0.80 was found between the two steel surfaces which is nearly

three times the ratio of their roughness values. Similarly, the two aluminium surfaces showed a ratio of around 0.60 which is also about three times their roughness ratio. Alternatively, the relation between these two ratios can be written in a more accurate empirical form as

$$R_s \approx 2R_r^{2/3} \tag{4.1}$$

where R_s is the ratio of slopes of deflection curves and R_r is the ratio of roughness values. Although the Hertzian approach ignores the surface roughness effects, the exponent of R_r in the above relation is the same as that for the P and E^* terms in Equation 3.3 for the approach distance. However, from the available data, such relations could not be proved to be valid for all sorts of surfaces and roughness regimes. The mathematical models presented in the previous chapter suggested more surface-related parameters that can reasonably affect the deformation of surface asperities. Detailed checks on the validity of these relations will be carried out on the new test-rig.

4.4.2 Repeatability of Contact Tests

Two types of repeatability tests were considered here, i.e. repeating the contact on: (a) nominally the same point on the surface and (b) different randomly selected points across the surface. The former type deals with one set of features on the rough surface; hence, a difference in contact deformation behaviour is expected between earlier and later contacts during the test. The latter type deals with several sets of "intact" features; hence, the deformation behaviour is expected to depend, to a large extent, on the variability of features across the rough surface.

Surface asperities might be reduced in height by elastic-plastic distortion, if the contact caused no work-hardening. A repeat contact on the same point might then show pretty well the same distortion, but starting from the reduced height of the plastic distortion of the first one. Thus, a consistent drift downwards of the apparent contact

point could occur. However, if we assume the radius of local Hertz contact is dominated by the sharp curvature of the higher asperities under light loading, then we might expect the plastic flow to increase the radius quite rapidly so that succeeding contacts do not deflect so much.

It is expected that the repeatability of contact at nominally the same point on any engineering surface will be affected by the different parameters of contact. Moreover, with repeated contacts with the same load, the deformation process of surface asperities on that specific point of surface depends not only on these parameters alone but on the number of such successive contacts as well. In other words, as it will be discussed later in Chapter 6, the plastic deformation of asperities is not fully achieved by the first contact, but continues in the following ones. Clearly, this behaviour can adversely influence the measurements (and even the concept) of such repeatability.

Two specimens, one mild steel and one aluminium, with an average R_q of around 0.30 μm and 0.57 μm respectively, were used to perform these tests. Each one was tested at two different contact loads, 0.5 N and 1.0 N, at two points randomly selected on the surface of each specimen. Thirty repeated contacts were made with each. The dead-loads were kept on the loading point on the lever, and the contact probe was lifted less than 2 mm and gently released. For each contact, the voltage reading of the capacitive gauge was directly taken from the multimeter display.

During each of the four tests, the voltage readings from successive contacts showed a descending trend that varied in its magnitude from one test to another. As noted earlier, this trend indicates that the plastic deformation of surface asperities was being gradually increased from repeated contact. Moreover, in all tests, the descending trend was starting at high rates in the first few contacts, then, ending at low rates in later contacts. The rate of this descent was found to be different in each test, and the repeatability of these contacts must be able to provide better indications about this.

Figure 4.6 illustrates the general behaviour, showing the sequence on the 0.5 N steel test. At 0.5 N load, a scatter within about ±0.14 µm and ±0.30 µm, at 95% confidence, was found on the steel and aluminium surfaces, respectively, and ±0.09 µm and ±0.25 µm at 1.0 N load. These values indicate that repeated contacts with a higher load can result in a lesser scatter of probe position readings, since the majority of plastic deformation of asperities can be attained faster. Hence, it might be expected that a softer surface material also shows a better repeatability of probe position than a harder one, provided that both surfaces have the same sort of roughness and are tested with identical loads. Similarly, a smoother surface is expected to give a better repeatability of probe position than a less smooth one. These results and expectations will be examined in more detail in Chapter 6, through some tests on the advanced test-rig.

The repeatability of deformation of surface asperities on different points across the surface has been briefly examined on the steel specimen of 0.78 µm R_q with a 2 N load. An initial contact load of about 0.05 N had first been adjusted to bring the probe gently into contact with the surface and form a reference position for measuring the deformation, then the 2 N load was applied. For a better accuracy, the TS5 probe was used and its stylus was remounted as near as possible to the capacitive gauge after each replacement of the contact probe on the surface. The contact probe was lifted by not more than 2 mm from the surface in order to move the specimen for testing the next point. Fifteen points were randomly sampled within a 20 mm square area on this surface. The mean deformation was about 0.68 µm with a standard deviation of 75 nm. The mean indicates a systematic error on position measurement that could in principle be compensated by calibration. The scatter indicates an effect of roughness variation over the entire surface, which gives an uncertainty of around ±0.15 µm at 95% confidence. It is believed that this uncertainty can also be affected by the different contact parameters, and it will be closely studied using the new test-rig.

4.4.3 Effect of Surface Conditions

The condition of surface cleanliness must surely play a role in contributing additional uncertainties to the contact probe position on the surface, beyond those that are contributed by surface roughness. How significant this role might be in practice is not clear. The presence of surface contamination by particles of dust and surface debris within the contact area will definitely cause poorer repeatability of contact height. Random false heights will result as particles get trapped. Contamination by oil and aqueous films might also directly affect the height detected. On a rough surface, the probe tip will generally contact the summits of asperities; off-centre high asperities might cause sliding friction with the tip as it settles into position and might force it to rotate. Contamination films might alter such behaviour, for example, acting as a lubricant and causing greater variation in the settled position. The effect on contact repeatability might well depend on the amount of roughness.

In order to check these effects using the available test-rig, repeated contact tests were performed at nominally the same position on the steel surface of 0.30 μm R_q under different successive surface conditions. A 0.5 N load was kept on the contact probe which was lifted, after taking the capacitive gauge reading, by about 2 mm and then returned gently onto the surface for the next contact. Thirty height (voltage) readings were collected for repeated contacts on a clean condition. Then, a thin layer of thick oil was applied on the surface and another thirty readings were taken. Afterwards, the surface was re-cleaned with isopropanol and re-tested as before. The standard deviation of the probe position under the first clean condition was around 47 nm, reduced to 43 nm under the oil condition and to 33 nm under the second clean one. This suggests that a slight surface smoothing (or asperities deformation, as discussed in the previous section) was taking place throughout each test and was responsible for the consistent change in these deviations rather than the changed surface conditions. Another less

plausible explanation: oil could reduce scatter either by a lubricant effect or by "filling" valleys so that probe is less likely to settle fully into them. However, re-cleaning should then remove such effects. A removal of debris from the nominally clean original surface might have happened.

For a closer look at these effects, two steel surfaces of different roughness have been tested in a similar way as given above, but using separate points for each surface condition on each of the two surfaces. Figure 4.7 shows the deviation of readings of the probe position at successive contacts from their mean value under clean and oiled conditions on 0.34 μm and 1.55 μm R_q steel surfaces. Clearly, the rough surface, Figure 4.7b, gave a higher standard deviation (about 245 nm and 456 nm for clean and oiled conditions, respectively) than the smooth surface, Figure 4.7a (89 nm and 229 nm). Hence, these results support the idea, given earlier in this section, that the effect of surface conditions can contribute additional uncertainties of contact beyond those that are contributed by the surface roughness; oil caused about 86% increase in uncertainty on the rough surface and over 157% increase on the smooth surface. This may also lead to the conclusion that the contact on smooth surfaces is more vulnerable to their cleanliness conditions due to, for example, the narrow spacing between the surface asperities which offer better possibilities for trapping oil films and dust particles at the site of contact. The other contact parameters, such as load, material, tip radius, etc., are also expected to have influences on this increase in contact uncertainties due to poor cleanliness conditions. These will be investigated in more detail within Chapter 6.

4.4.4 Effect of Probe Tip Size

From the ideal mathematical contact analysis of Hertz and all the other contact models such as those presented in the previous chapter, the size of indenter controls the amount of deformation as it changes the size of the contact area. Hence, tip radius is considered

one of the contact parameters that can affect the repeatability of contact. A bigger probe tip causes a larger area of contact and deforms more surface asperities, which might imply a wider scatter for such deformations between repeated contacts.

This effect was checked on the rough steel surface of 1.55 µm R_q with two Mitutoyo steel probes of 3 mm and 5 mm contact tip diameters. Repeatedly contacting nominally the same point on this surface with each probe, but separate point for each probe, gave the deviations plotted in Figure 4.8. These were measured by the capacitive gauge using a 0.5 N load on a probe that was being lifted not more than 2 mm from the surface. At 95% confidence, the scatter was about ±490 nm with the 5 mm probe and ±200 nm with the 3 mm one. The larger scatter of the probe positions readings on the bigger contact area is believed due to the rate of achieving full plastic deformation of asperities throughout the repeated contacts, which was slower in this case, although these deformations were smaller. As shown in this figure, the deviations of the 5 mm probe position readings have, in general, a descending trend which caused a lower repeatability for them than those of the 3 mm probe position. Hence, as noted in Subsection 4.4.2, the slower in attaining the full plastic deformation the larger is the uncertainty in probe position. Moreover, the 3 mm tip shows slight downward trend but the 5 mm tip has sudden and large change in behaviour. This might be an artefact of the test-rig (e.g., dirt and/or clearance in bearings) or a typical behaviour at the surface (e.g., fracture of an important asperity). So, these scatters may not represent a genuine effect of the tip size. However, for better evaluations of the effects of this contact parameter, more tests will be carried out on the advanced test-rig.

4.4.5 Effect of Probe Tightening

The condition of probe tightening is considered neither as one of the contact parameters nor one of the contact gauge (indicator) characteristics illustrated in Chapter 2. Yet, it is

an artefact capable of a significant error contribution, as far as the whole performance of the gauging process is concerned. It was thought to be worth examining here, since the available test-rig offers good possibilities for manipulating the contact probe without greatly disturbing its accuracy. However, considering the normal operating condition of indicators, it looks odd to raise an issue that the contact probe has not been well secured to the indicator's rod: it is more likely to happen when no great attention has been given during assembling or maintaining the indicator itself.

Three different tightening positions for the contact probe were used: "tight" is the normal operating one where a tool was used to obtain it; "loose" where the probe was screwed by hand only until its shoulder just touched the bottom surface of the lever; and "very loose" where the probe shoulder was just not touching that surface. The 5 mm probe was used in these tests with a 0.5 N load on a smooth steel surface of a fixed specimen (so isolating any errors of the base surface). Under each of the tightening conditions above, three contact repeatability tests were performed. In each, the probe was lifted not more than 2 mm from the surface and replaced, repeatedly, for thirty times at nominally the same position on the surface. Figure 4.9 shows a plot for the standard deviations of all tests at each probe tightening.

The scatters of the probe position with a "tight" probe, shown in this figure, were basically due to the probe lifts. They are within the same range as those given earlier in Figure 4.3 for the 2 mm probe lifts. When a "loose" probe as described above was used, deviations were almost doubled which indicates how severely such inadequate tightening can affect the probe position repeatabilities. When the probe was in a "very loose" condition, these deviations were at least eight times higher than normal and greatly varying from test to another. Obviously, this is due to "backlash" between the probe screw and its threaded hole, which is an undetermined behaviour. The above ratios of increase in deviations are not necessary true for every probe and tightening

condition. However, such tests highlight all the possible sources of uncertainty that could adversely affect the gauging performance of contact indicators.

4.5 PRELIMINARY CONCLUSIONS

Various tests have been carried out to characterize the original, existing test-rig, aimed at comprehending all the possible errors which may interfere with the measurements taken, although some of these errors seem less likely to be transmitted under the normal operational conditions of the rig. At this preliminary stage of the research, information about the overall accuracy performance of this rig is strongly needed not only for the design of a new test-rig of advanced performance but also for identifying the principle relations for a more comprehensive study.

Throughout the characterization, the measurement noise which combines vibration, electronic noise and digitization effects was typically below 12 nm peak to valley; hence, the useful resolution was below 15 nm. Thus, it is believed that, for an individual contact test, tip displacement can be measured to within ±5%, although with noise errors rising rapidly at deflections much below 100 nm. In addition, a typical repeatability error of around 80 nm must be included if the probe is lifted by up to 2 mm, and an active uncertainty from pivot errors of less than 25 nm if a 7 N load is used. So, great caution is needed when interpreting variations of less than 120 nm.

Load-deflection curves for those surfaces tested on this test-rig have shown consistent variabilities with surface material and roughness, which are to be examined in further investigations. The repeatability of the probe vertical position at nominally the same point on the surface appears to be altered by the rate of achieving the full plastic deformation of asperities through the repeated contacts which, in turn, is believed to be governed by the different contact parameters. The presence of oil on the

surface has given additional uncertainties to the probe position, with their magnitude dependent on the roughness. No clear evidence was found that such surface condition has affected the mean elastic deformation of asperities, which is normally shown by the last repeated contacts in every test. Successive contacts randomly sampled across the surface have given uncertainty due to variations of local roughness, which seem to correlate with roughness. Again more tests are needed.

The test-rig has proved its capability to reveal some hints about these basic relations. Further advanced investigations need more reliable data than can be realized with its limited accuracy performance. Additionally, its design configuration does not offer the required flexibility for testing specimens of different form factors without disturbing the operating linearity of the capacitive gauge. In other words, changing the specimen height requires readjusting the gap between the electrodes, which, in turn, sets the gauge to a different linear behaviour, as the upper electrode rotates with the lever and its adjustment is not genuinely perpendicular to the lower electrode. In addition, due to the tight fit between the electrode-adjusting rod and its hole, this adjustment (sliding the upper electrode vertically) cannot be easily done without rotating this rod, which creates another effect on the gauge linearity. However, the optimum linear response of the gauge could be obtained when the specimen size makes the lever nearly parallel to the base plate. The same arguments apply when changing the probe tip size.

The need for constructing a new test-rig, with a high-precision performance and enhanced design capabilities, has been strongly justified. The interim conclusions presented in this chapter will be rigorously tested and these investigations will be extended to a wider set of conditions that represent real probing environments.

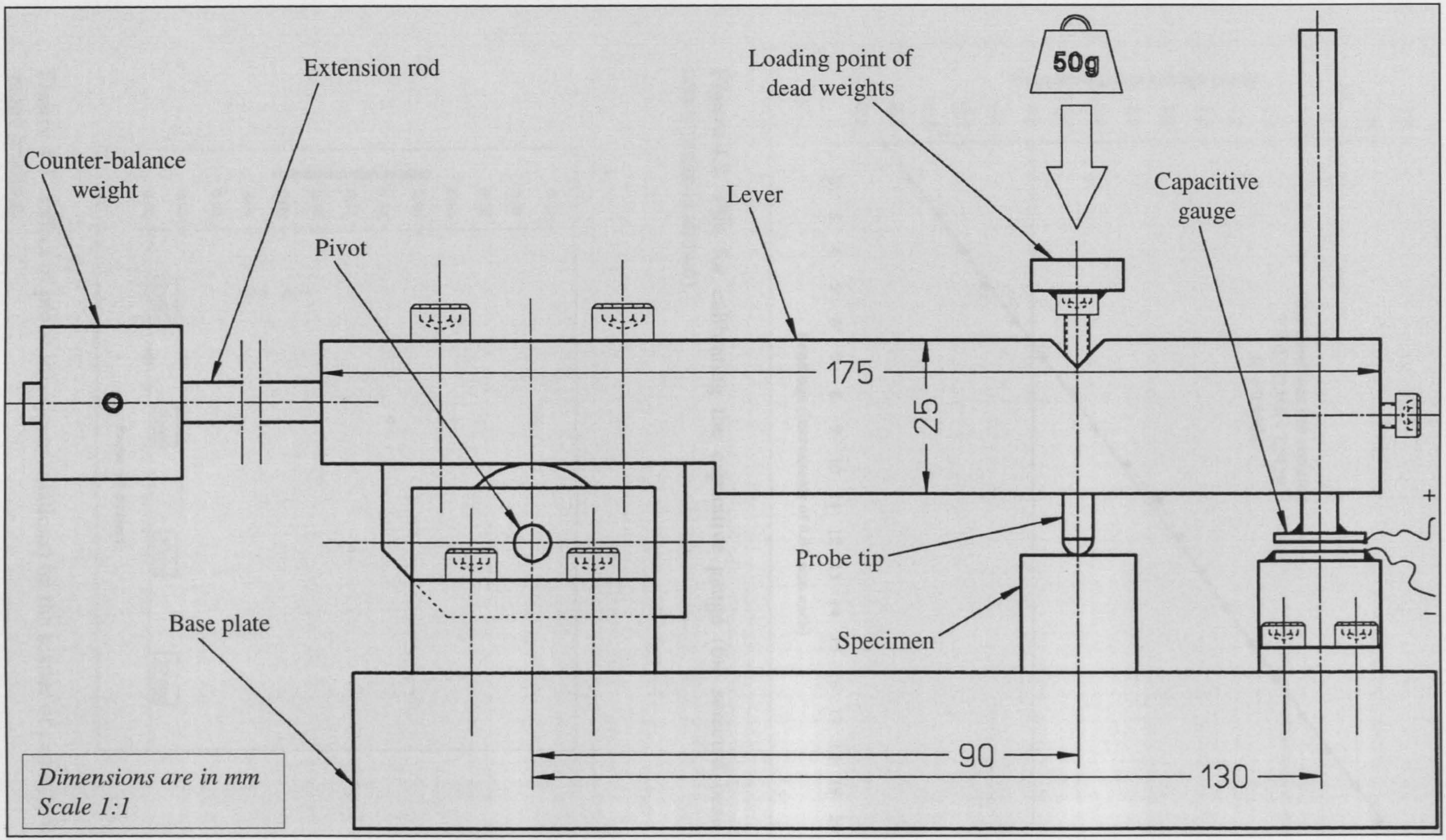


Figure 4.1 Schematic representation of the existing simple test-rig (after modifications).

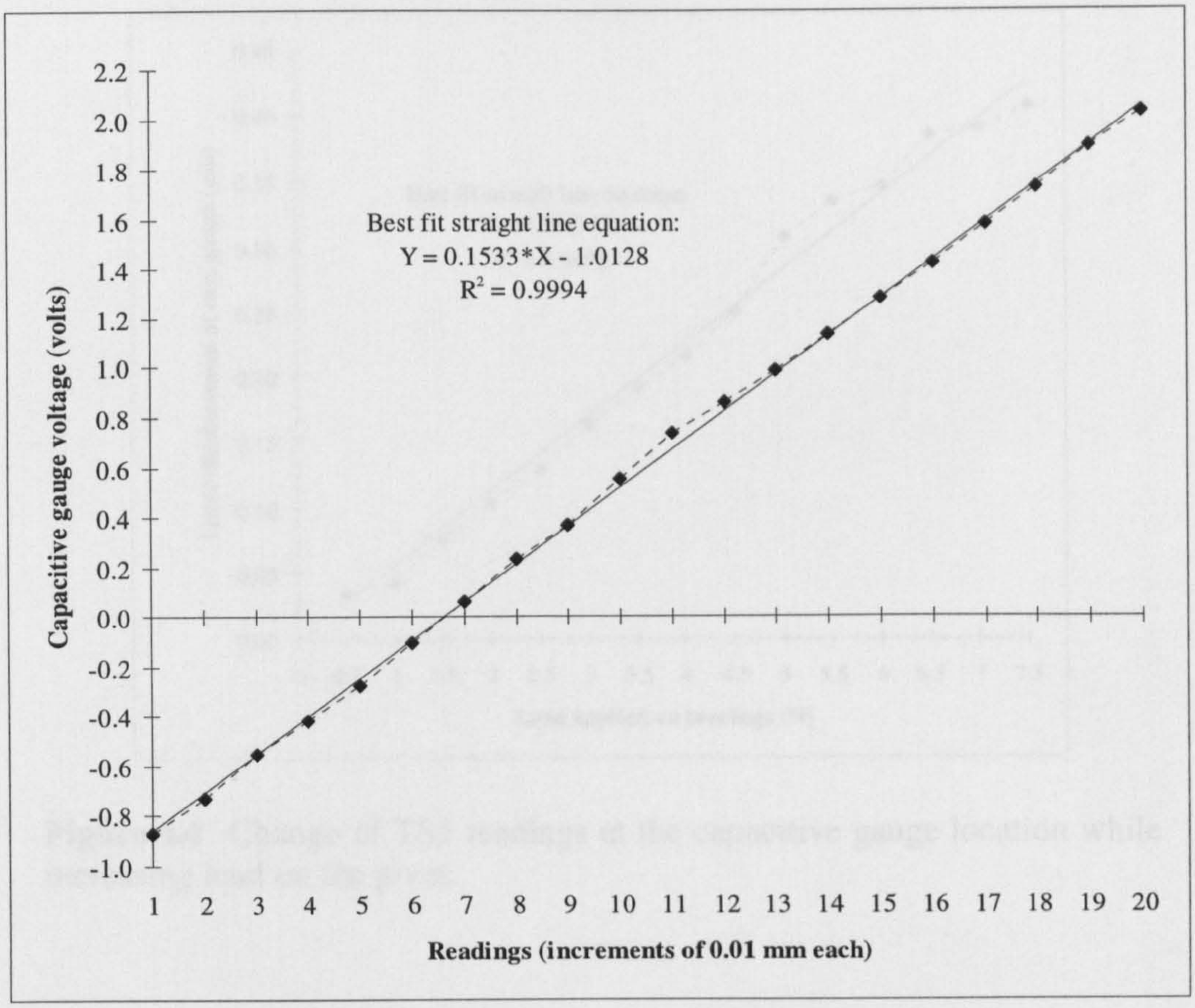


Figure 4.2 Plot for calibrating the capacitive gauge (the selected best-linear range is shown).

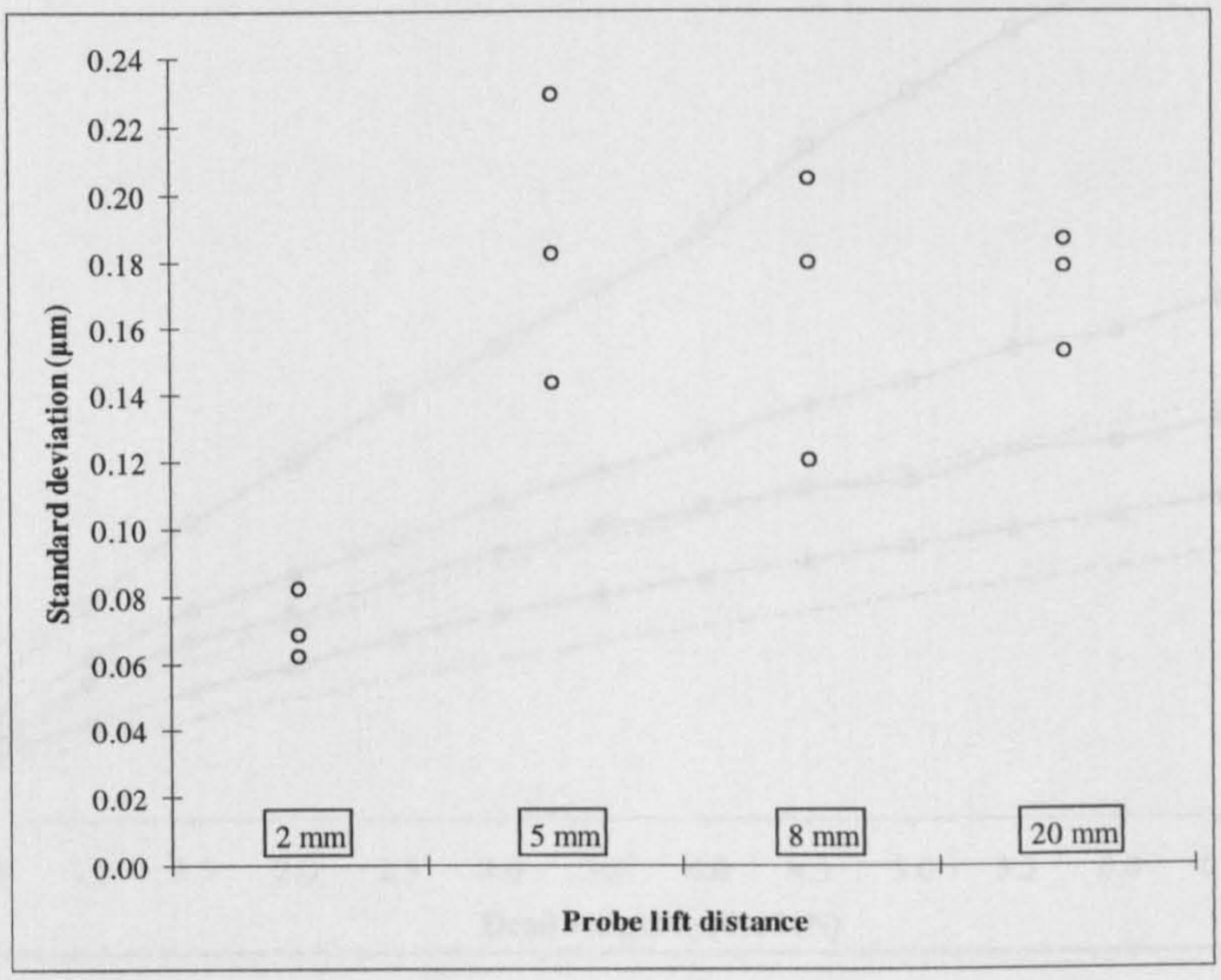


Figure 4.3 Effect of probe lifts (lever rotations) on the scatter of capacitive gauge readings.

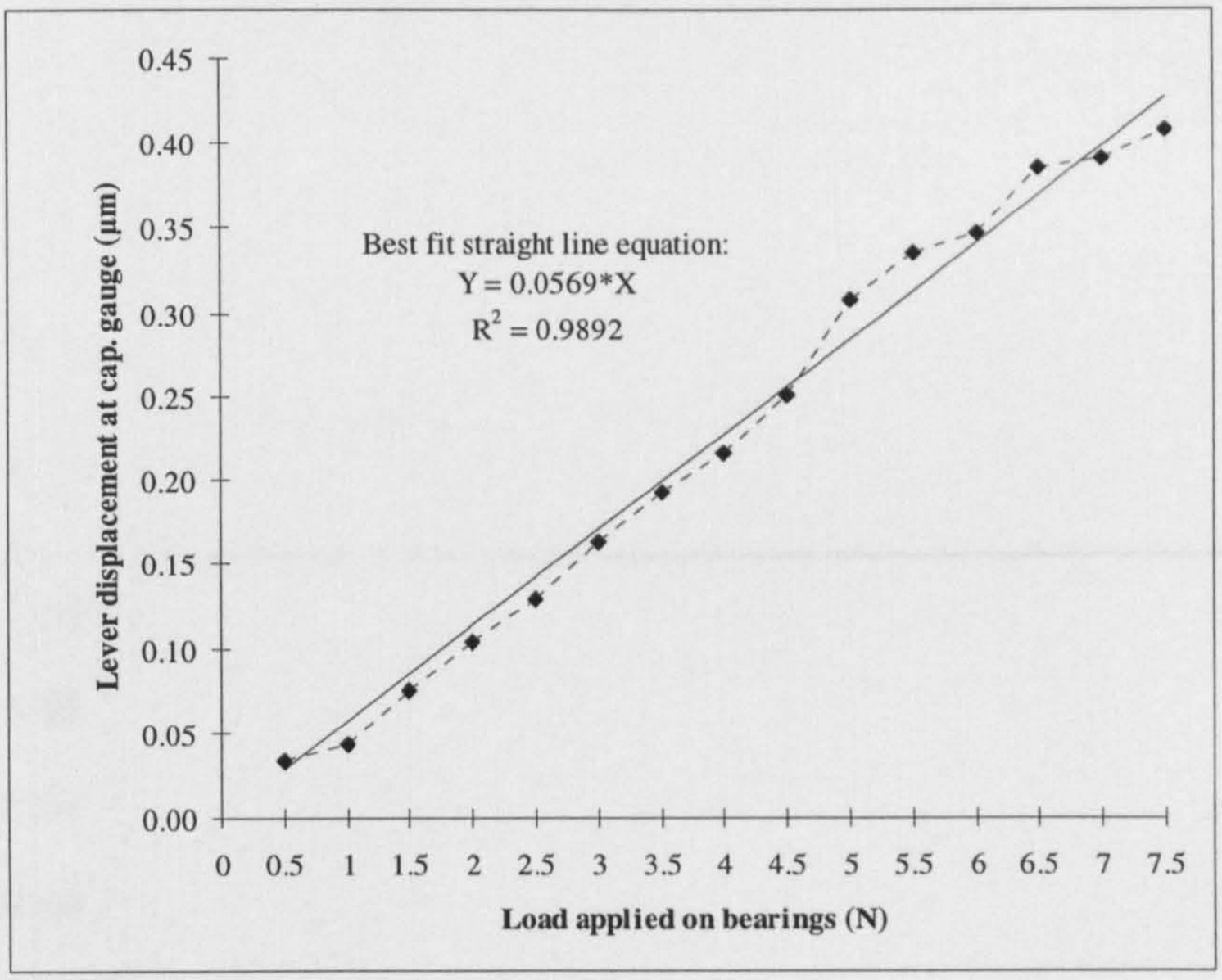


Figure 4.4 Change of TS5 readings at the capacitive gauge location while increasing load on the pivot.

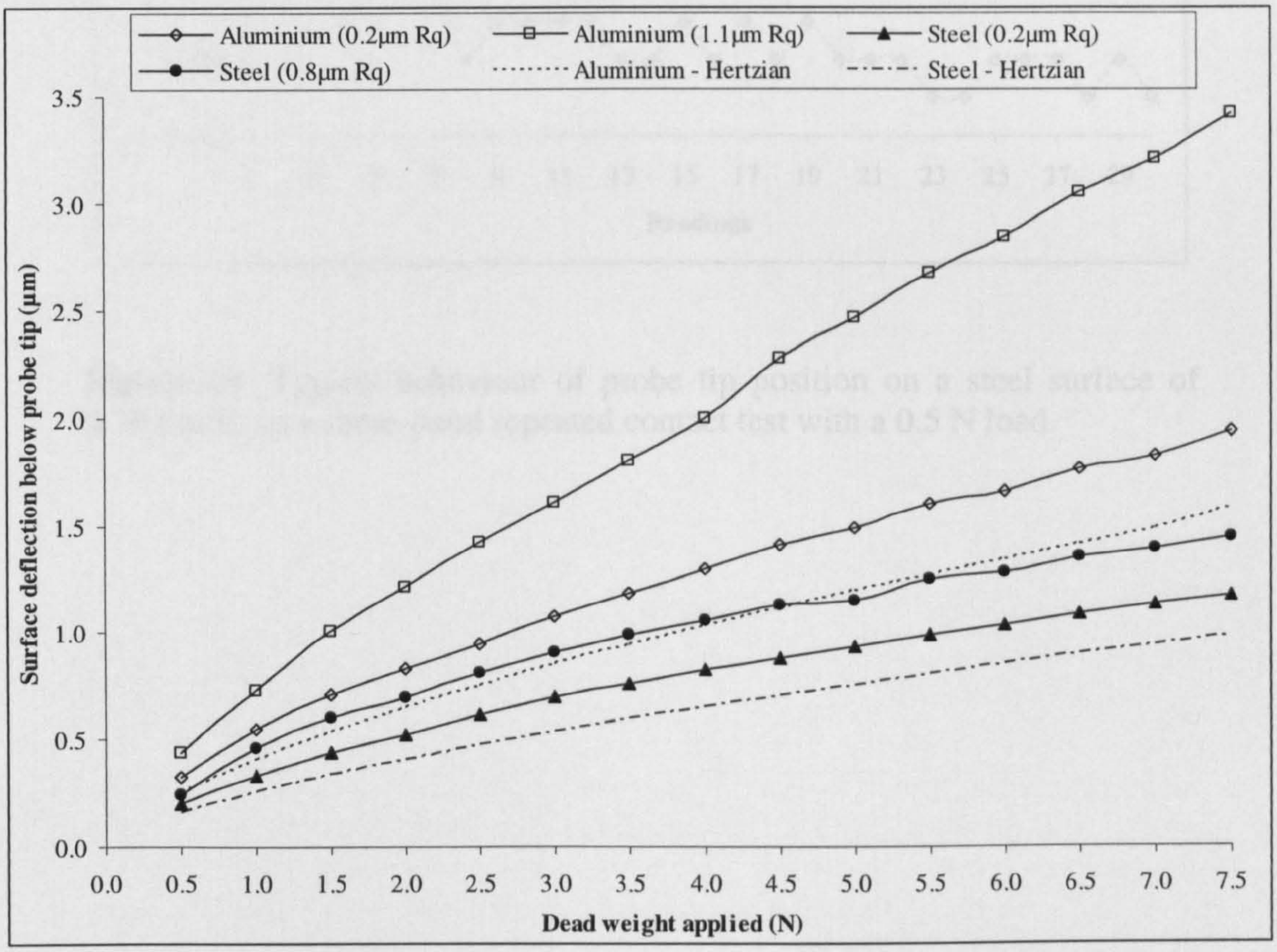


Figure 4.5 Load-deflection plot of four steel and aluminium specimens with the corresponding Hertzian curve.

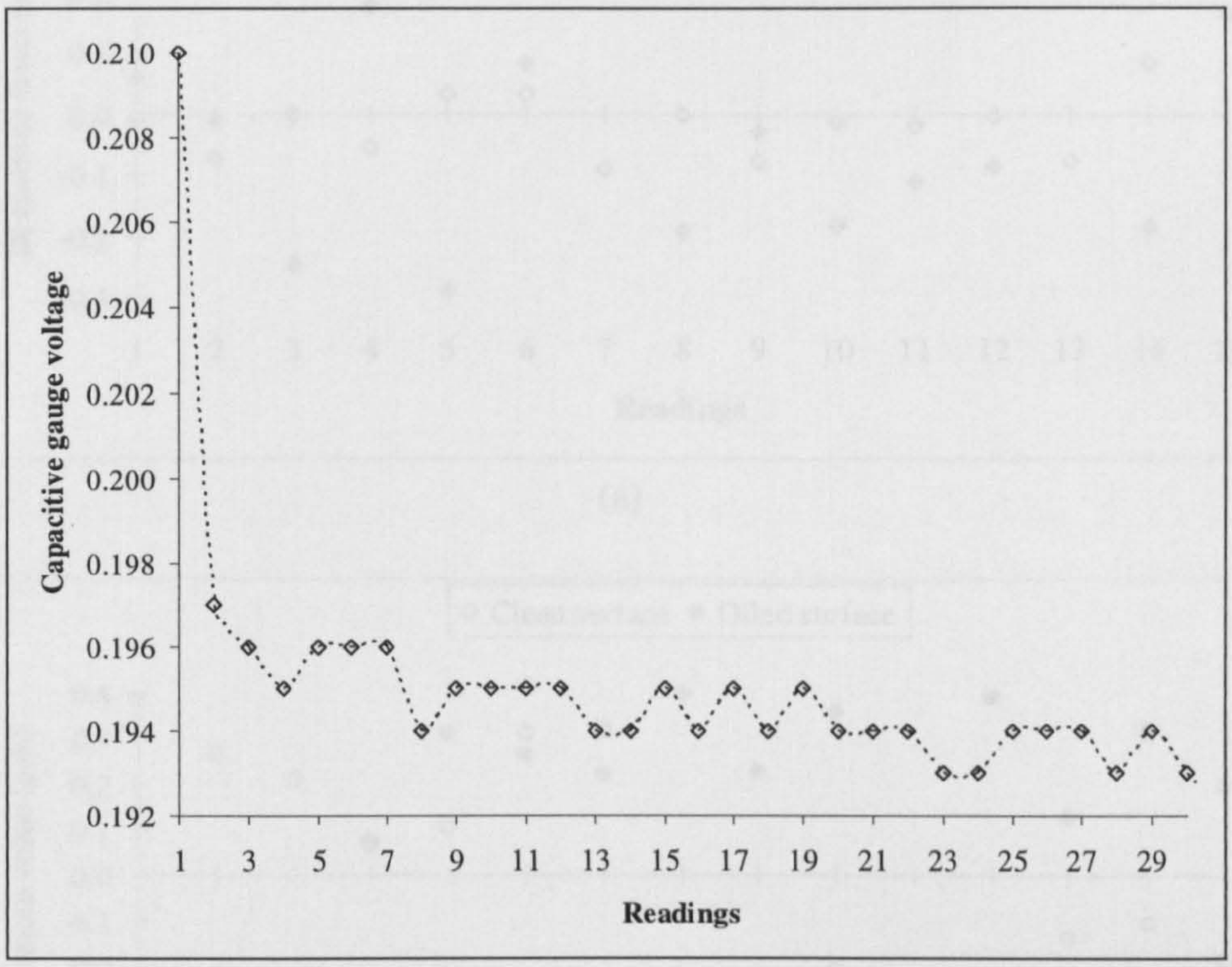
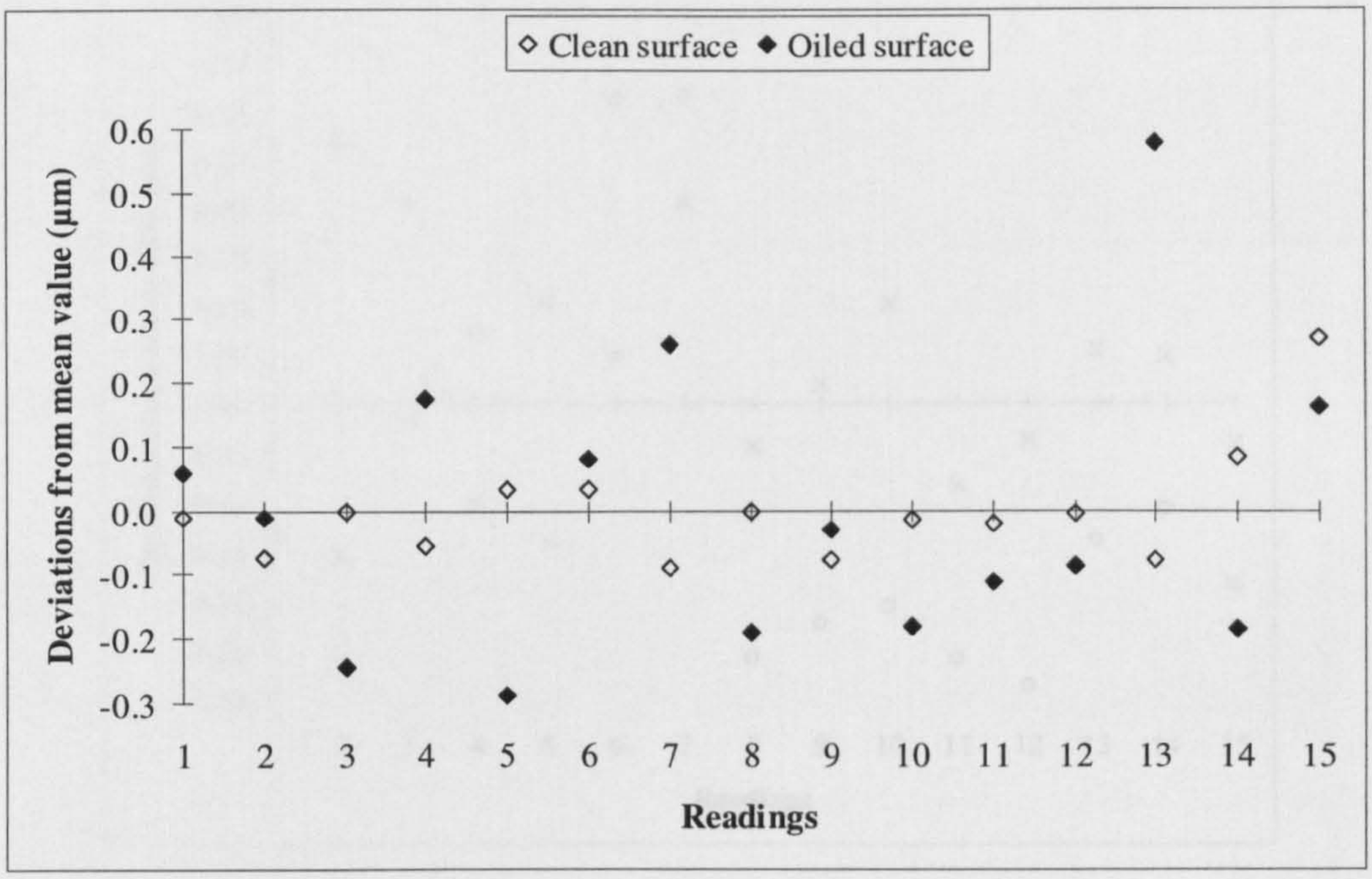
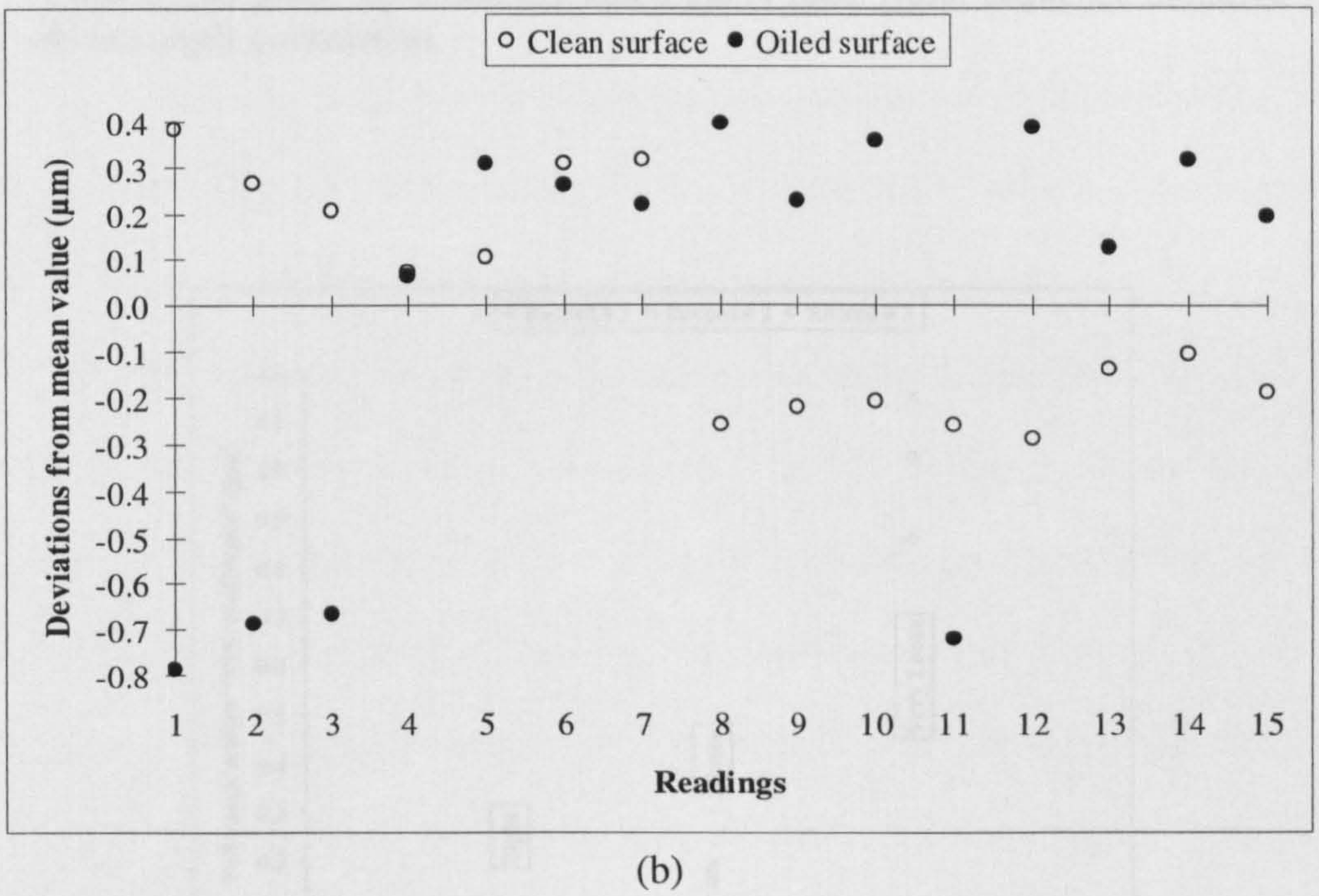


Figure 4.6 Typical behaviour of probe tip position on a steel surface of 0.30 μm R_q in a same-point repeated contact test with a 0.5 N load.



(a)



(b)

Figure 4.7 Repeatability of contact height on different points on steel surfaces of different cleanliness conditions and roughness of (a) 0.34 µm R_q and (b) 1.55 µm R_q, using a 0.5 N load on a probe tip of 5 mm diameter. *Note: the sequence number does not imply correlation of points.*

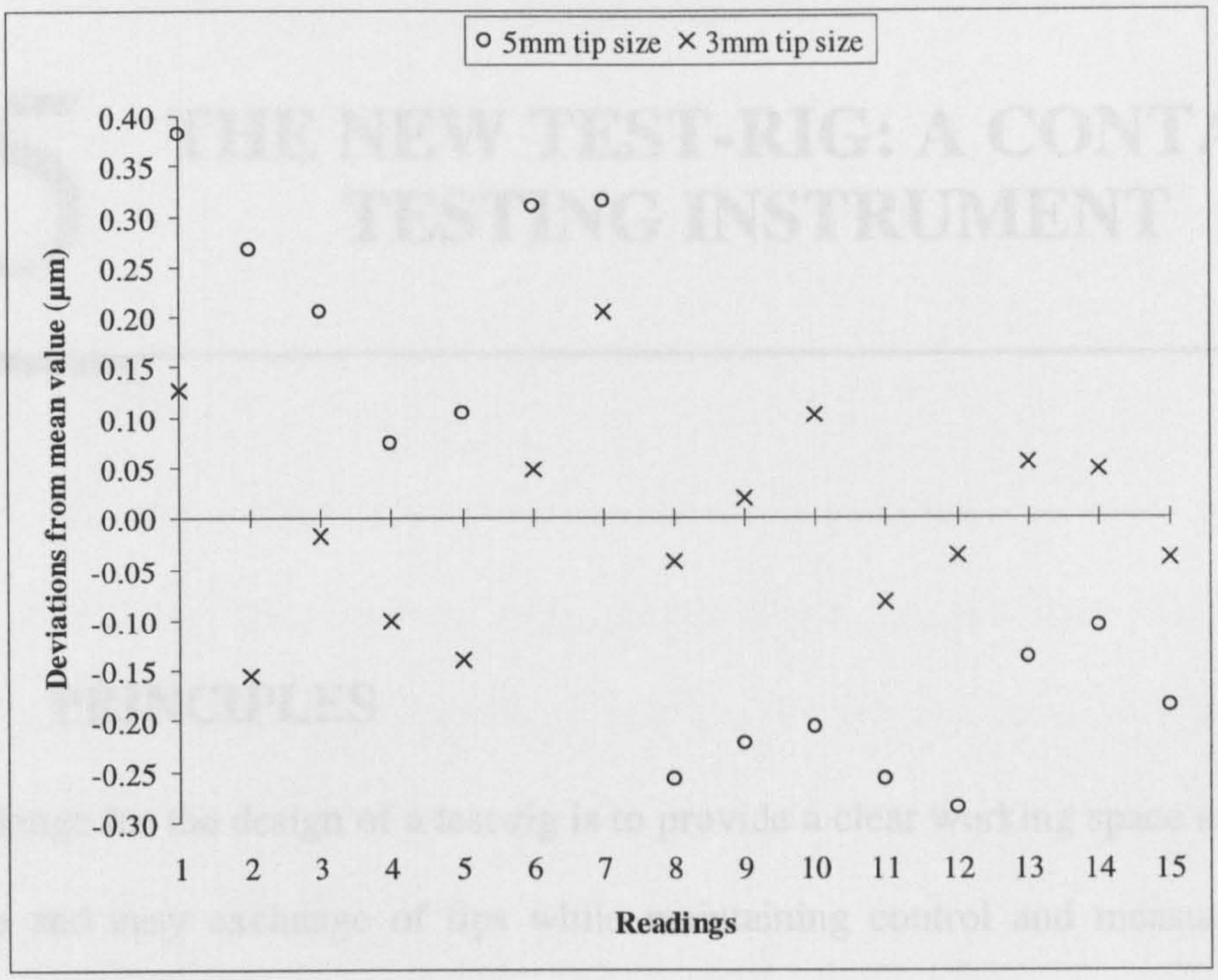


Figure 4.8 Repeatability of contact on a steel surface of 1.55 µm R_q, using 5 and 3 mm probe tip diameters and a 0.5 N load. *Note: sequence numbers do not imply correlation.*

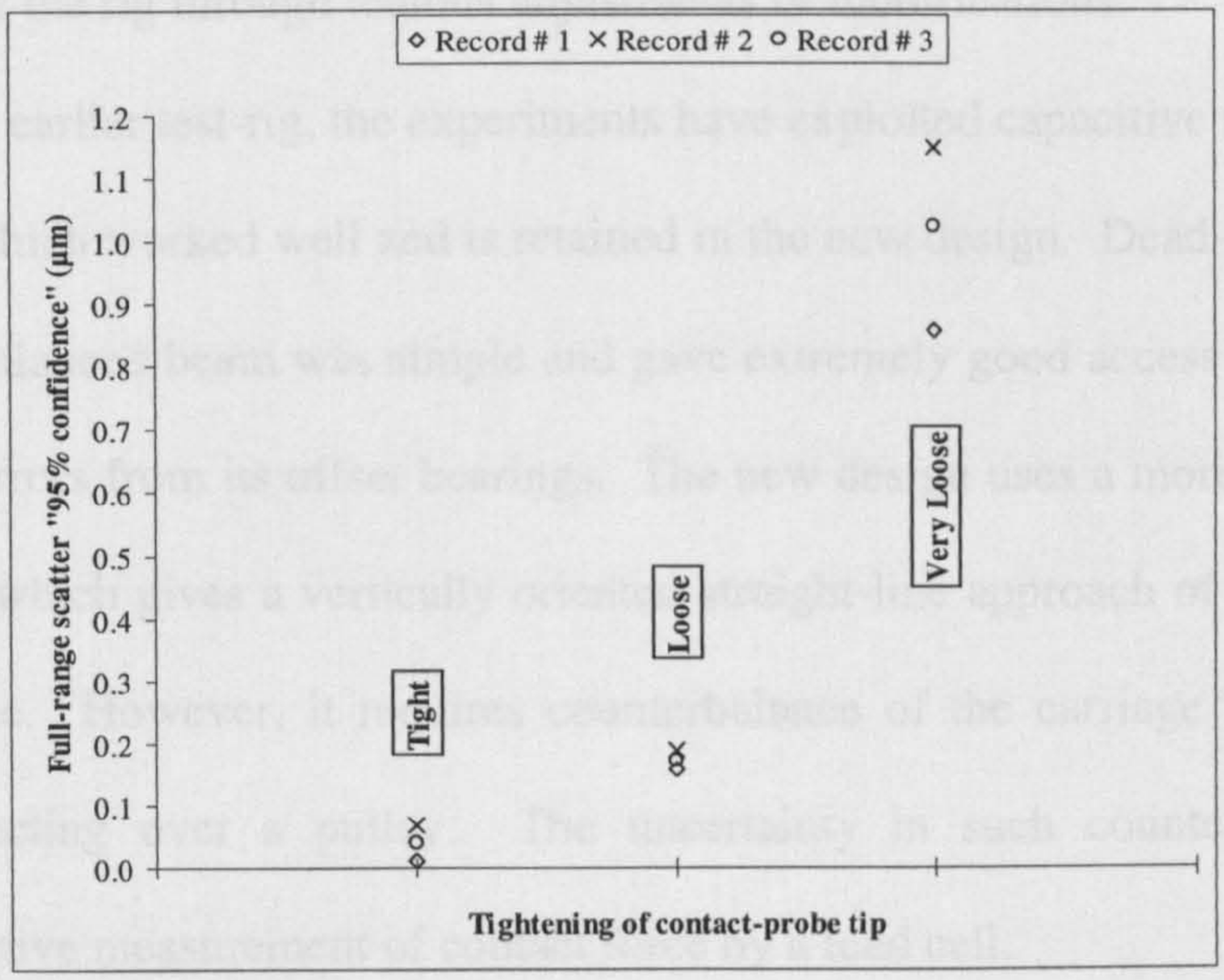


Figure 4.9 Effect of probe tip tightening on the repeatability of contact on a fixed steel surface of 0.34 µm R_q roughness, using a 0.5 N load on a probe tip of 5 mm diameter.

CHAPTER 5 THE NEW TEST-RIG: A CONTACT TESTING INSTRUMENT

5.1 PRINCIPLES

The challenge for the design of a test-rig is to provide a clear working space around the probe tip and easy exchange of tips while maintaining control and measurement of contact forces to below newton precision and of displacements to around 25 nm. It must be possible to lift the probe tip a short distance from a test surface and return it with high precision to the same lateral position on that surface. It should be capable of testing a wide range of specimen sizes without significantly affecting the measuring performance of the rig through tedious adjustments or modifications.

With the earlier test-rig, the experiments have exploited capacitive micrometry for the gauging, which worked well and is retained in the new design. Dead-weight loading via a counterbalanced beam was simple and gave extremely good access but introduced repeatability errors from its offset bearings. The new design uses a more robust guided linear motion which gives a vertically oriented straight-line approach of the probe onto the test surface. However, it requires counterbalance of the carriage by a spring or dead-weight acting over a pulley. The uncertainty in such counterbalances then necessitates active measurement of contact force by a load cell.

The test-rig operating principles have much in common with precision indentation testers. In particular, it shares with them the inherent requirement to impose a significant and variable force at a point of precision positional (or displacement)

measurement. It is essential that the path of this force is separated from that of the metrology loop to the largest practical extent.

5.2 DESIGN

Figure 5.1 illustrates a schematic representation of the main features of the new probe repeatability test-rig; front view is shown in this figure. A complete set of detailed assembly and part drawings, which were prepared for the machining workshops, are attached in Appendix B. In addition, Figure 5.2 shows two photographic views (front and rear) for this test-rig and its attachments.

The test-rig basically comprises a 250×150×20 mm flat steel base-plate that provides a working stage onto which specimen surfaces can be placed. The probe holder and its associated instrumentation are mounted from a vertically moving carriage perpendicular to a 160×50 mm bridge structure secured to the base-plate. The basic idea is that the body of the carriage slides freely on high precision cross-roller thrust bearing slides from Schneeberger attached to the two steel uprights of the bridge. An upward force generated by weights acting over pulleys balances the weight of the carriage; although the issue of pulleys friction is likely in such a system, it offers a nearly constant counterbalance force along the carriage travel in addition to an easy control on this force in comparison with the spring system. Extra weight can then be added on the carriage to provide the nominal downward contact force of the probe onto the specimen surface.

The carriage (90×30 mm) shown in Figure 5.3a incorporates a horizontal leaf-spring flexure element cut from a copper plate of 1 mm thick (over-constrained system, built in at both ends) which carries the probe on its centreline. It provides for a degree of smooth guided motion at the level where the conventional bearings may exhibit stick-

slip behaviour. It can also be used to impart smoothly increasing loads by driving the carriage downwards on its main bearings once the probe has made contact with the surface. In either mode, the counterbalance force applied by the string is totally within the force path and outside the metrology loop, so it does not introduce errors from cross-talk. To ensure consistent running of the main bearings, one side of the carriage is mounted to a small copper flexure element, see the image of Figure 5.3b, with high stiffness vertically and moderate stiffness horizontally. This allows some movement to compensate for the inevitable residual errors in the alignment of the bearings, while keeping the line of action of the probe closely registered to one datum.

Immediately below the horizontal leaf-spring there is a commercial load-cell to monitor the actual resultant contact forces. It is needed for two reasons. First, the effectiveness of the counterbalances and variations on the main rolling bearings may produce small, but potentially significant, variations in force for the same nominal dead-load applied. Second, it provides a means of detecting initial contact of the probe with the surface in a consistent way: "contact" occurs when a pre-specified reaction force appears at the probe tip.

Below the load-cell is the steel probe holder which passes through, and is bonded to, the centre of a square glass plate of 38 mm sides and 3 mm thickness. The lower surface of this plate carries evaporated gold electrodes that form part of the displacement gauging. The holder then passes through a clearance hole in a second glass plate of $50\times44\times3$ mm, the top surface of which has gold electrodes that form capacitive micrometers with those on the other plate, as shown in the photo of Figure 5.3c. The lower glass plate is mounted directly to the base-plate through a slotted steel holder which is carried by a non-rotating micrometer (lower centre in Figure 5.2b) that provides height adjustment to set the nominal capacitor gap for each combination of probe tip and specimen. The upper plate has a single annular electrode,

acting as a common ground for four equal annular segments on the other plate. In this way, sum and difference readings allow the z-motion and error rotations in two axes to be monitored. In practice, electrodes are usually run as pairs in parallel, monitoring rotation only in the less stiff torsional axis of the rig.

It worth mentioning here that the assembly and adjustment of these plates were accomplished in the final stage of the whole assembly process of the test-rig. First, the upper glass plate was centred onto the probe holder, levelled horizontally, and glued to this holder using Araldite epoxy adhesive. Then, the lower glass plate was inserted in the clearance slot of its holder, adjusted and spring-clamped to the upper one, and glued. Of course, the clamps were released later when the glue had cured. Although, some relaxation is likely on release, this procedure reduces parallelism and rotation errors between the two plates which, in turn, affect the sensitivities of the capacitive gauges.

Commercial probe tips are screwed into the end of the probe holder. They project through the lower glass plate by only a few millimetres in order to reduce the metrology loop dimensions and aid thermal and mechanical stability. The lower electrodes plate is approximately 44 mm across, leading to relatively poor access for other instrumentation close to the probe tip although there is a large unimpeded region for moving specimens relative to the tip. An extensive experimental study of the rig with an offset electrode to give easy access to the probe tip from one side demonstrated that the restriction was unavoidable. The minor error rotations could not be guaranteed insignificant at the 20 nm vertical sensitivity level even with Abbe offsets of only 25 mm.

A means of clamping the intended specimens ($20\times20\times6$ mm) to the base-plate was added in the final stage of the test-rig design in order to reduce the effect of their bottom surfaces on the measurement. To avoid direct clamping, each specimen was first glued to a micrometre slide that will then be clamped to the base-plate; more details on this will be given in Section 6.2. A U-shape holder (not drawn in Figure 5.1, but

shown at lower centre in Figure 5.2a) was designed for this purpose and loosely attached to the base-plate from rear side of the test-rig. So, it can be released (and raised) to place, move, or remove the specimen with its slide and tightened by hand from the front side with the aid of a wings nut, a screw projected from the base-plate, and a clamping plate. However, such a clamping system can be adjusted, modified, or eliminated for any future needs.

The two attachments of the test-rig shown in Figure 5.2 are a stand that carries a fine micrometre for applying the dead-weights on the carriage; and another stand that carries the pulleys and the counterbalance weights for balancing the carriage weight. Both the micrometre and pulleys were meant to be separated from the rig body so as not to disturb the probe contact with the surface (hence, the measurement) while rotating the micrometre or changing the dead-weights.

5.3 OPERATION

A block diagram representation of the complete test setup and its associated instrumentation is given in Figure 5.4. As shown, two power supply units (one for the capacitive gauges and one for the load-cell) and three voltmeters (to monitor the changes in the output voltage of the two capacitive gauges and the load-cell) were used. It was necessary to use a low-noise preamplifier to filter the output signal of the load-cell; a low-pass mode was selected with a cut-off frequency of 100 Hz. The two oscillator gauges are the electronics modules of Queensgate Instruments Nanosensor modified for external reference capacitor; a 10 pF capacitor (open to air with `Zerodur` body) was used as a common reference. A channel distribution box was used to link the three signals to the 12-bit data acquisition PC card from National Instruments Inc. The LabVIEW routine used in the previous investigations was also utilized here to acquire

the measurement data needed.

In operation, the specimen is placed on base-plate (with its glass slide underneath the two arms of the U-shape holder) and moved with respect to the probe tip to the required point on the surface to be tested. After adjusting the proper height of the U-shape holder from the base-plate, the slide is then clamped to the base-plate by pressing the arms end by tightening the wing nut on the clamping plate. The probe tip is brought onto contact with the surface by displacing the carriage downwards using its micrometre; the counterbalance force should be adjusted to be sensibly higher than the carriage weight in order to keep the carriage moving with the micrometre spindle movements. A great care is then needed when establishing the initial contact with the required force from the first attempt; repeating the same small force in series of tests would be difficult as it depends greatly on the operator's hand. Alternatively, a more convenient way is followed: the counterbalance force is adjusted to be sensibly less than the carriage weight so that the carriage is just able to move downwards very slowly. This movement is then controlled by hand to have a gentle initial contact with the surface. Here, the resulting initial force obviously depends on the pre-adjustment of weights rather than the sense of the operator; hence, a better repeatability of the initial force would be expected in repeated contact tests.

Once the initial contact is established, the gap between the two electrodes of the capacitive gauges is adjusted by using the lower electrode fine micrometre in order to attain the calibrated linear range of the gauges. At this stage, readings from the capacitive gauges and the load-cell are taken as reference or "no-load" readings. The dead-weight required for testing can then be hung on the dead-weight fine micrometre which is provided to assist in applying it very gently on the carriage. As the dead-weight is applied, readings are taken again to compute the resulting force and displacement encountered by the contact probe. Obviously, in order to release the

specimen being measured for testing another point or replacing another specimen, the dead-weight should be released first to raise the carriage easily by either using its micrometre or adding a reasonable weight to the counterbalance weights so that it is able to move upwards slowly.

5.4 CHARACTERISATION AND CALIBRATION

Evaluating the inherent errors of the test-rig mainly focussed on the accuracy of the carriage motion, as it greatly affects the accuracy of measurements and, hence, the general functional performance of the rig. The calibration factor of the capacitive gauges and their useful operating range had also to be determined.

The work in this task has used 5 mm diameter steel probe tips and the configuration with two capacitive gauges only, designated for convenience "front" and "rear", and the view shown in Figure 5.1 is from the "front" in this context. This configuration allows the detection of tilts that may occur at the onset of contact in the less stiff torsional axis of the test-rig.

5.4.1 Sensitivity of Capacitive Gauges

The two gauges were directly calibrated against a Queensgate Instruments Digital Piezo-Translator that offers nanometre precision and traceability to laser interferometers. It was fixed on the test-rig in the same way as a specimen. Its datum surface was glued to a glass slide and positioned directly below the contact probe and the U-shape holder was used to clamp the slide to the base plate. The counterbalance weight that acts over the pulleys was reduced in order to allow the probe to rigidly touch the translator top surface (with about 2 N force) and, hence, follow its movements.

Accordingly, the lower glass plate of the gauges was adjusted. The translator was controlled to produce increments of displacement to the contact probe and the changes in the output voltage of the gauges were recorded.

The small differences in electrode area and the unavoidable parallelism error between the two glass plates lead to different sensitivities for the two gauges. Tracing the behaviour of these sensitivities from the lowest possible output voltage (around 4 V) to about 10 V showed a nearly systematic decrease of sensitivities. For convenience, the range between 6 and 8 V was selected to be the operating range of the gauges. Figure 5.5 shows the change of sensitivities at the start and at the end of this range. The sensitivities of the front and rear gauges have changed by 2.5 nm and 2 nm, respectively. The difference in their readings was varying between 1.4 and 1.8 nm, respectively, as the separation between the two plates increased by approximately 30 μm. A maximum standard deviation of around 0.4 nm was noticed on the sensitivities readings collected. However, these readings were averaged and the sensitivities were taken as 0.016 and 0.015 μm/mV, respectively.

5.4.2 Evaluation of Carriage Rotations

Since the accuracy of motion of the carriage is closely related to the metrology loop of the test-rig and greatly affects the accuracy of vertical positions of the contact probe, it was essential that it was fully characterised. It was plausible to assume that as the carriage moves vertically it must suffer from rotations due to the assembly error of the two slides (each one has a different perpendicularity to the base plate), as well as, the inherent bearing errors. An additional effect on this movement could arise from pulling the carriage from one side only. Along the carriage vertical path, such rotations may cause greater variations to the sensitivities of the gauges than the inherent ones evaluated in the previous subsection.

A Hilger & Watts Simultaneous Two Directional Microptic Auto-collimator was used to measure these carriage rotations about the two horizontal axes (x and y) of the test-rig. The auto-collimator was mounted vertically on the column (freed from the gearbox) of an old Talysurf 4 instrument and the test-rig was placed below it on the worktable, such that the objective was viewing the carriage from the top. A silicon wafer, providing a reasonable flat and reflective surface, was attached to the top of the upper glass plate of the capacitive gauges and adjusted to face the auto-collimator objective at the side of the carriage. Accordingly, the whole auto-collimator instrument was levelled and adjusted to get the best focus for both the vertical and horizontal lines of the reflected target image at the eyepiece. Extra weight was added to the counterbalance weights in order to control the movements of the carriage by using its own micrometer.

The carriage was repeatedly allowed to rise and lowered by a distance of around 15 mm and the resultant shifts in the image lines were monitored. During its upwards movement, increasing clockwise and counter-clockwise rotations were observed about the x- and y-axes of the rig, respectively, i.e. tilting to the left side (as viewed from front) and to the rear side at the same time. These directions of rotations were, of course, reversing as the carriage was in its downward movement. After several repeated movements, the maximum rotations about the x- and y-axis were determined as approximately 11 and 16 seconds of arc, respectively. Both rotations increased steadily with the carriage height. However, within the operating voltage range of the gauges where a total displacement of around 30 µm could be attained, these rotations values were less by at least 500 times, considering that their trends were linear. Moreover, they even were less as all the deflections measured have not exceeded the 8 µm regime on all the surfaces used in this research. Hence, it is acceptable to assume that the systematic carriage rotations due to its movement were transmitting no significant

effects to the sensitivities of the capacitive gauges. The inevitable small random fluctuations in orientation at any given height will almost certainly be larger. These latter effects appear within the overall repeatability of the carriage positioning.

5.4.3 Evaluation of Probe Tip Repeatability

Probe tip repeatability was examined through series of contacts on nominally the same point on a steel gauge block. To avoid (or minimize) any errors from the test-rig that might be contributed to the measurement as a result of lifting the probe, the contact was maintained while repeatedly modulating the force (repeatedly releasing and applying the dead-weight on the carriage). Two tests were performed on the smooth hard surface at two different dead-weights, 0.95 N and 1.45 N. Figure 5.6 shows plots of the load-cell force readings and the capacitive gauges displacement readings of the successive contact on the gauge block surface at both tests. The high deflections at the first application of the dead-weights have been briefly discussed in the previous chapter and will be extensively discussed in the next chapter. The tests, however, showed consistent vertical displacements at the capacitive gauges with load, subject to a slight random variation with standard deviations of contact height well below 10 nm (excluding the first displacement readings).

But, as shown in Figure 5.6, the load-cell reading also suffered variations. Ideally, in such a type of repeatability tests, the load-cell should show the same voltage difference, as long as the same dead-weight was repeatedly applied on the carriage. But, as shown in Figure 5.6, that difference decreased, mainly at the second and the third applications of load. Then, it became nearly constant at the later applications of load. A major justification for this behaviour of force may be valid here. The applied dead-weight is basically reacted by the tip (desired) and by (undesired) carriage friction, in addition to the counterweight pulleys friction. Friction always opposes motion, so on

the first loading, the carriage tends to move down and so the tip force would be less than the dead-weight. On unloading, if the carriage moves up (due to the spring-back action or elastic recovery from the surface), extra load might be expected on the tip, roughly the size of the "missing" load previously. This would result a significant hysteretic effect. On the next loading, the starting height of the tip on the surface became lower than that before the first loading (obviously due to the plastic deformation), and the starting tip force became higher than the initial load; so the relative changes in the probe force and displacement would be less than those caused by the first loading. This behaviour seems to continue over the next few loadings and relax at the latter ones as the plastic deformation becomes smaller and the reactions to the same dead-weight become nearly similar.

So, the bearing performance of the slides is responsible for this behaviour of the load-cell, since it was caused by a lack of response to surface movements by the carriage. The stick-slip behaviour (static friction) of the rolling elements in these slides (which is a normal characteristic of conventional roller bearings) must have opposed the motions and "locked" the carriage over these very small movements which were observed well below 0.5 μm at the 1.45 N on the gauge block surface used. As it will be seen later, this response of the load-cell was also noticed in all similar tests performed on all surfaces used in this research. Although this bearing behaviour cannot be practically compensated in the present configuration of the test-rig, it could be reasonably avoided by considering only the magnitude of first relative load reading, in addition to the variation of such load over the later applications of weight within the test, which is directly related to the repeatability of the surface deflections. However, it has been observed from similar tests demonstrated in the next chapter that this behaviour of the slides had no significant effect on the deflections measured since the overall load on surface was nearly constant during each test.

5.4.4 Operating Noise of Measurement

The noise of measurement of the load-cell and capacitive gauges were monitored during this characterisation and throughout the experimental work accomplished on this test-rig. As stated earlier, this noise combines the laboratory floor vibrations, electronic noise, and digitisation effects. A standard deviation of about 135 mN was observed within the load-cell readings, and standard deviations of around 32 nm and 35 nm were found within the readings of the front and rear gauges, respectively. Hence, the useful resolutions of the test-rig were well below 0.3 N and 70 nm in single load and displacement measurements, respectively. However, these noise figures were collected during working daytimes in a laboratory noise environment beyond the standards of industrial metrology labs. Furthermore, the noise was distributed across a frequency band and so averaging, either by multiple readings or by low-pass filtering, can provide improved resolution, when required.

5.5 CONCLUSIONS

A new enhanced apparatus for studying the contact behaviour of precision industrial gauges (indicators) has been constructed. In its design, some ideas from a previous test-rig have been retained and new design requirements have been realized. Its characterizations have revealed a few error sources of measurable magnitude but they should have insignificant effect on the accuracy of intended measurements. The main features and specifications of this new test-rig are summarized below. Some design "weaknesses" are also mentioned here for possible future improvements.

5.5.1 Test-Rig Features

Workspace

A clear working space around the contact probe and an easy exchange of tips and specimens are provided on a flat steel base-plate.

Specimens size	The large workspace allows specimens of up to 45 mm in height to be tested.
Probe motion	A smooth guided straight-line motion of the probe tip through a secured bridge structure attached to two precise cross-roller bearing slides.
Displacement and rotations measurement	The probe tip vertical position and rotations about the horizontal axes are monitored by a glass capacitive micrometre which consists of four evaporated gold electrodes in equal annular segments.
Load measurement	The use of a commercial load-cell allows the detection of any potentially significant variations in force and provides consistent means of detecting the initial contact.
Dead-weight loading	Dead-weight loading provides a consistent repeatability of force on the surface and better stability of the probe tip.
Metrology loop	Counterbalance of the carriage and dead-weight loads are totally within the force path and outside the metrology loop, so no error transmitted from cross-talk.

5.5.2 Test-Rig Specifications

Sensitivity of capacitive gauges and optimum operating range	16.5 nm/mV (front) and 15.0 nm/mV (rear), 6–8 V (≈ 27 μm change in gap)
Make, model, sensitivity, and range of load-cell	Entran International, ELH-TC400, 16.4 mV/lb. (≈ 3.6 mV/N), 4 lb. (≈ 18 N)
Specimens size range	≈ 25–45 mm (height) × max. 60 mm (length) × max. 30 mm (width) – with manipulations of specimen
Dimensions (L×H×W)	250×180×150 mm (excluding carriage micrometer length and dead-weights and pulleys stands)
Noise resolutions at the current lab environment	Below: 70 nm (displacement) and 0.3 N (load)
Make, type, model, and accuracy of slides	Schneeberger, cross-roller thrust bearing slide, ND 2-80.50, better than ±1 μm over 50 mm
Make, type, model, and accuracy of micrometers	Mitutoyo, non-rotating spindle micrometer head, 153-202, ±3 μm direct reading
Make of contact tips recommended and used	Mitutoyo contact points – Steel tips of diameters of: 5 mm (No. 101118) and 3 mm (No. 120045)
Other instruments used for capacitive gauges	Queensgate Instruments Nanosensor electronics modules modified for external reference capacitor. 10 pF reference open to air with Zerodur body

5.5.3 "Weaknesses" in Test-Rig Design

Throughout the characterisation and operation of this test-rig, a few design "weaknesses" were recognised and noted below. Although, they were considered within this investigation to have insignificant effects on measurements, Chapter 7 will also discuss them as recommendations for future improvements and better overall performance of the test-rig.

- The noise figures, which have been measured earlier in the existing laboratory environment, were high compared to those deviations noticed on the surface of a gauge block. Isolation of the rear surfaces of the electrodes glass plates is believed to be significant in reducing electromagnetic interference from environment and different parts of the rig.
- As shown before, the carriage rotation about the y-axis was increasing as the carriage was moving upward (tilting off the pulling side of the carriage) which is believed due to, mainly, the offset counterbalancing of its weight. This suggests that there should be a balanced pull from both its ends, which could be practically attained by some modifications.
- The slides' bearing performance was observed not only to change the magnitude of the initial load within the same-point repeated contact test, as discussed earlier, but also to affect the adjustment of this load prior to any single contact test. A better control of such load could reduce uncertainties of surface deflections due to initial load variations which could be clearly observed on the different-point repeated contact tests.

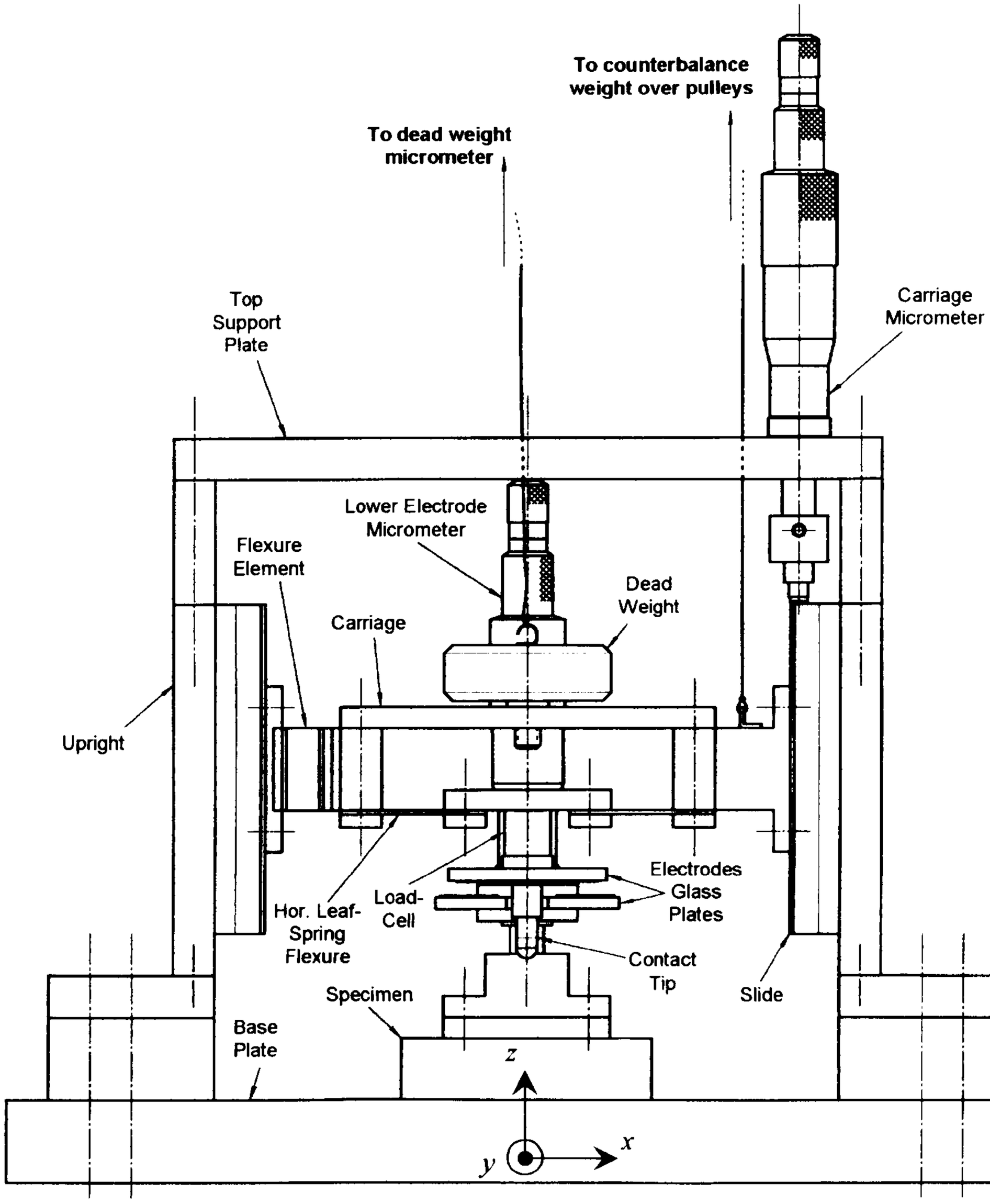
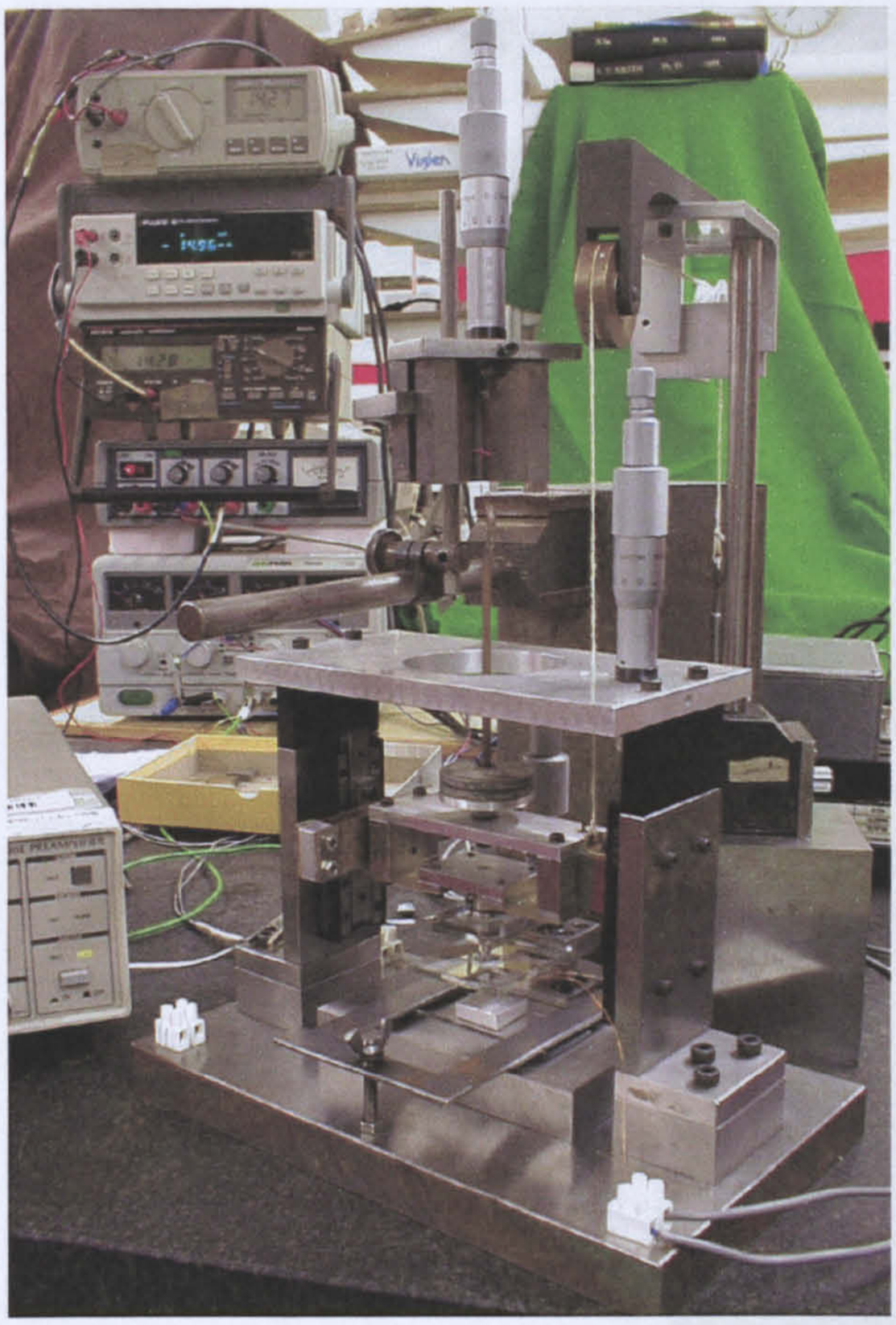
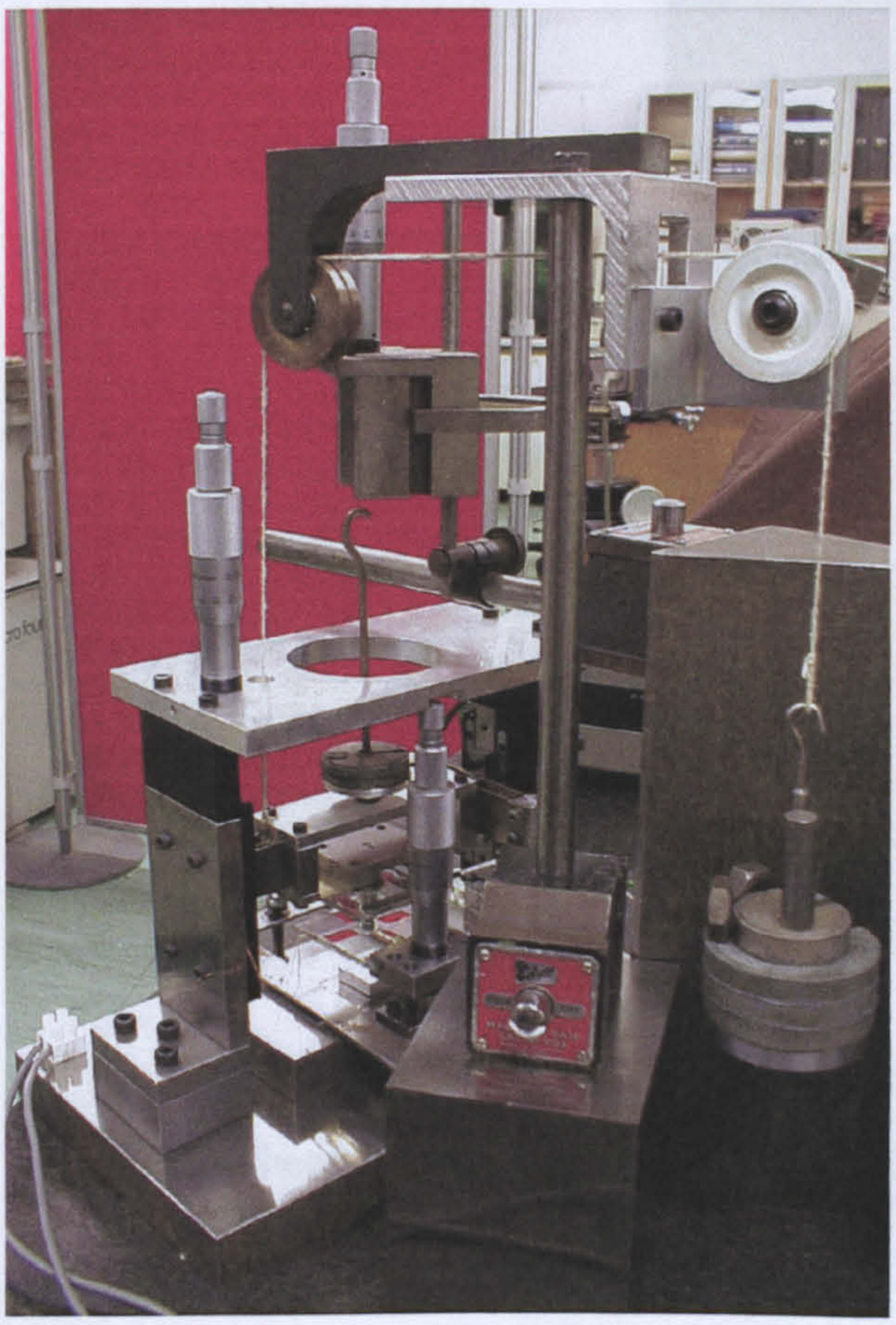


Figure 5.1 Schematic representation of front view of the new probe repeatability test-rig.

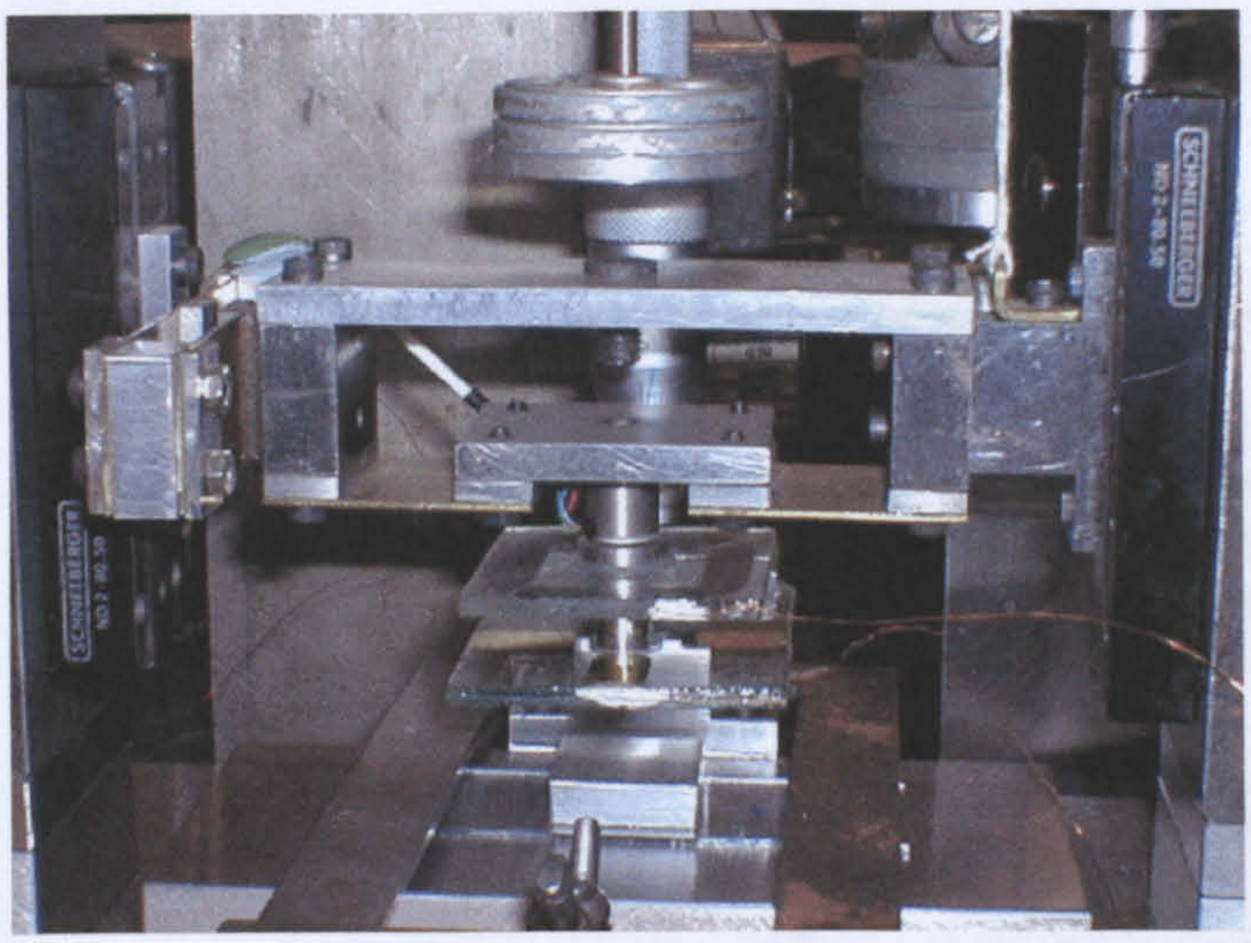


(a)

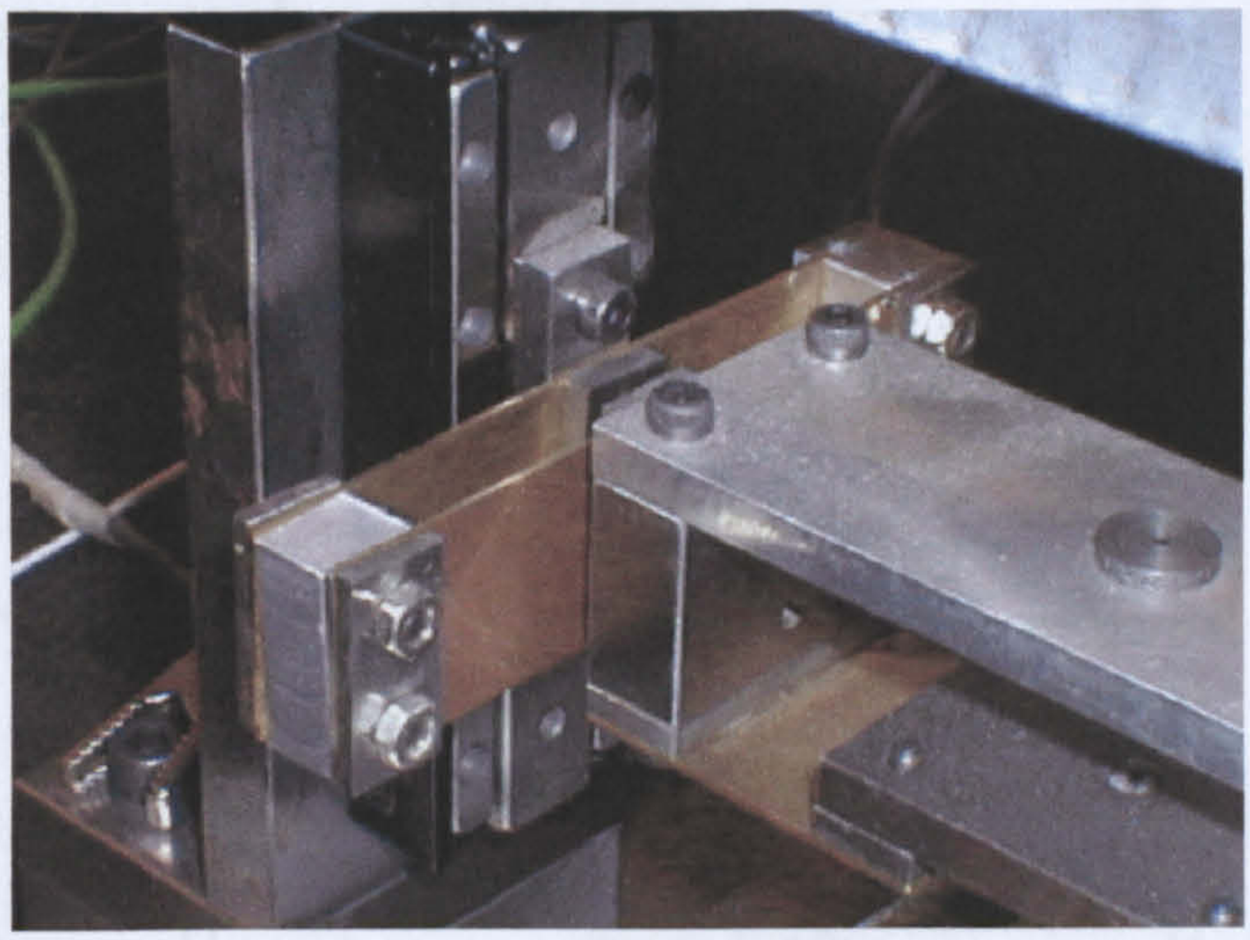


(b)

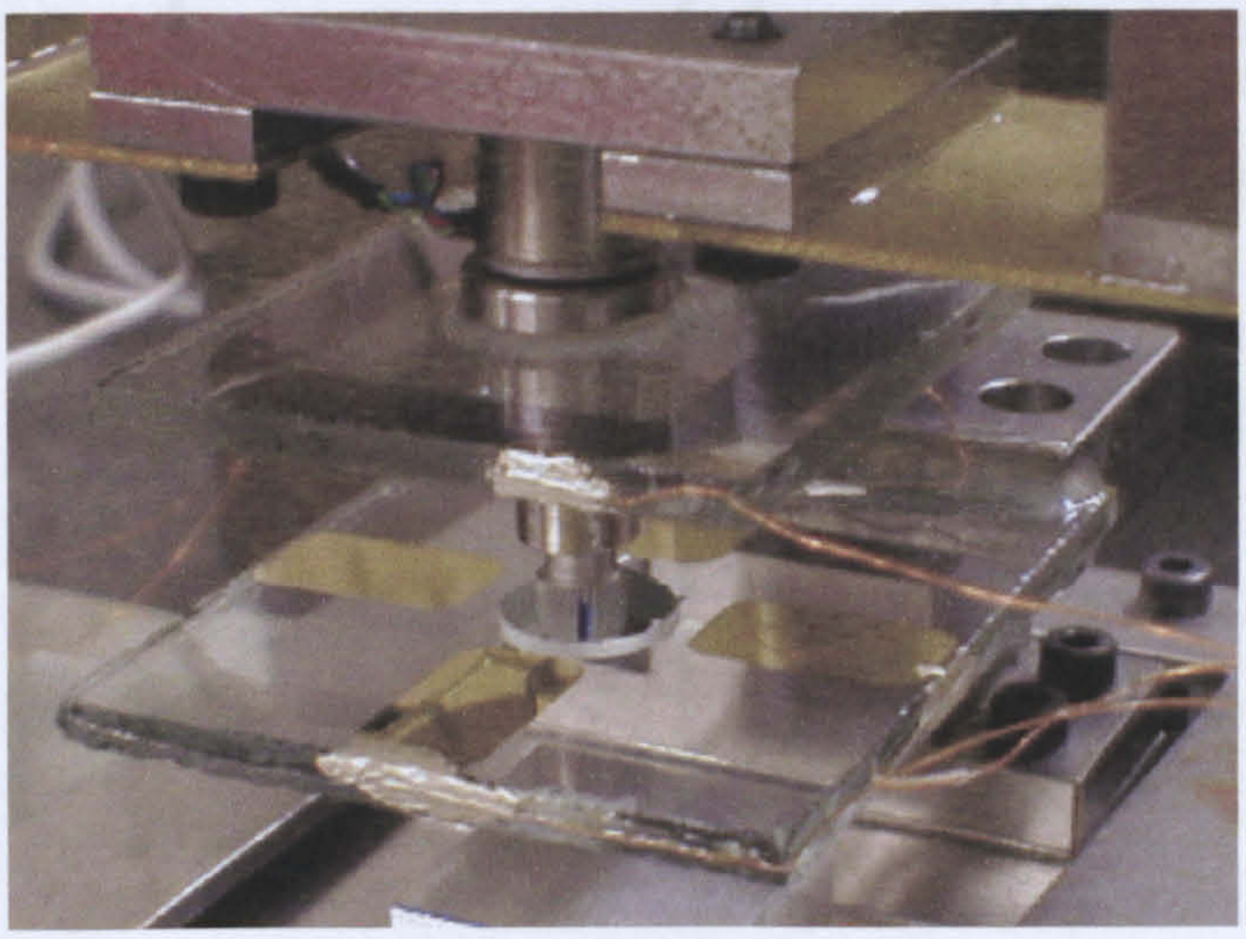
Figure 5.2 Images of the new test-rig with its associated instrumentation, from the (a) front and (b) rear sides.



(a)



(b)



(c)

Figure 5.3 Clips of some main parts of the new test-rig: (a) the carriage, (b) the horizontal leaf-spring flexure, and (c) the glass plates of the gauge electrodes and the load-cell.

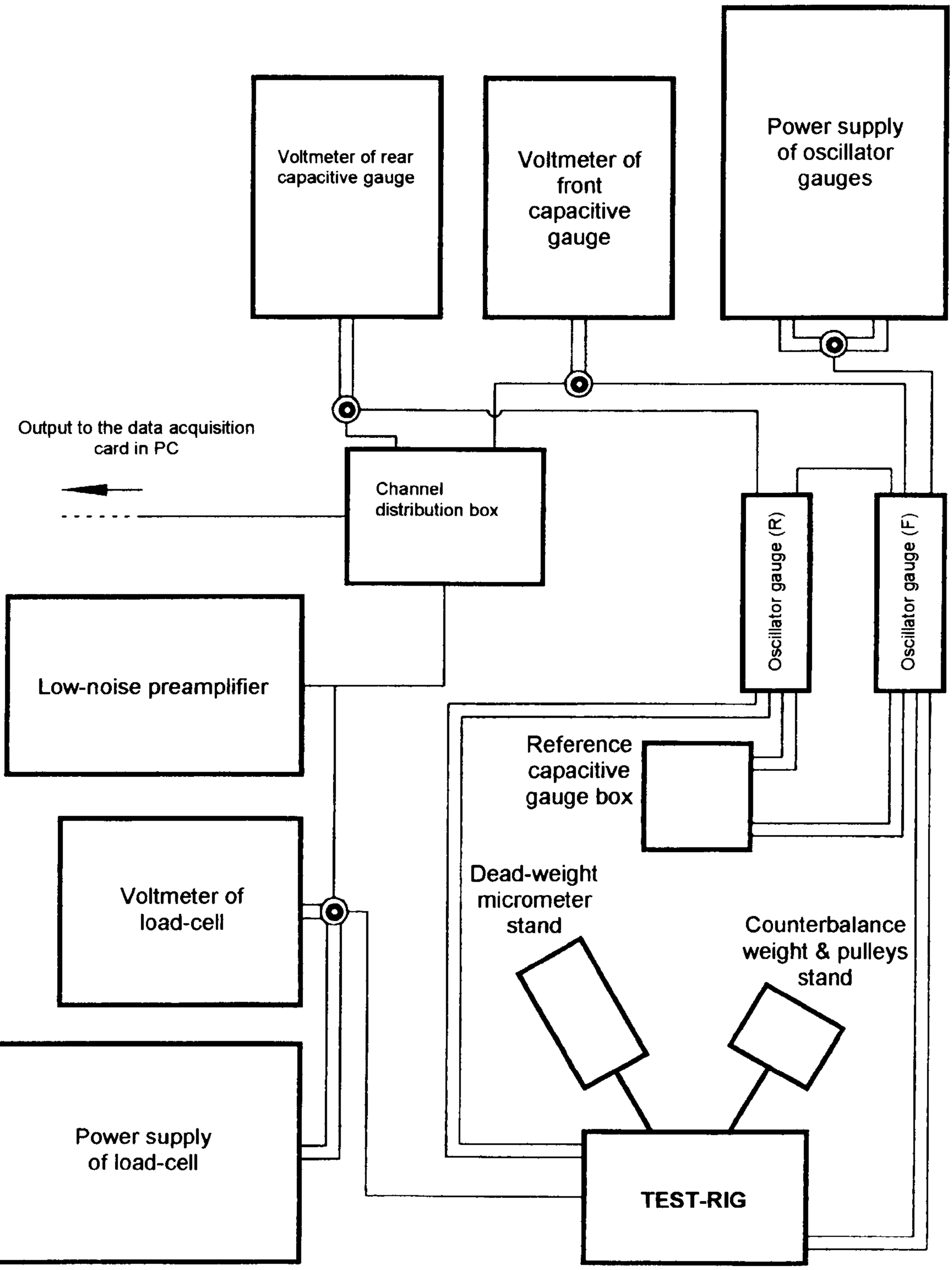


Figure 5.4 A block diagram representation of the complete test setup and its associated instrumentation.

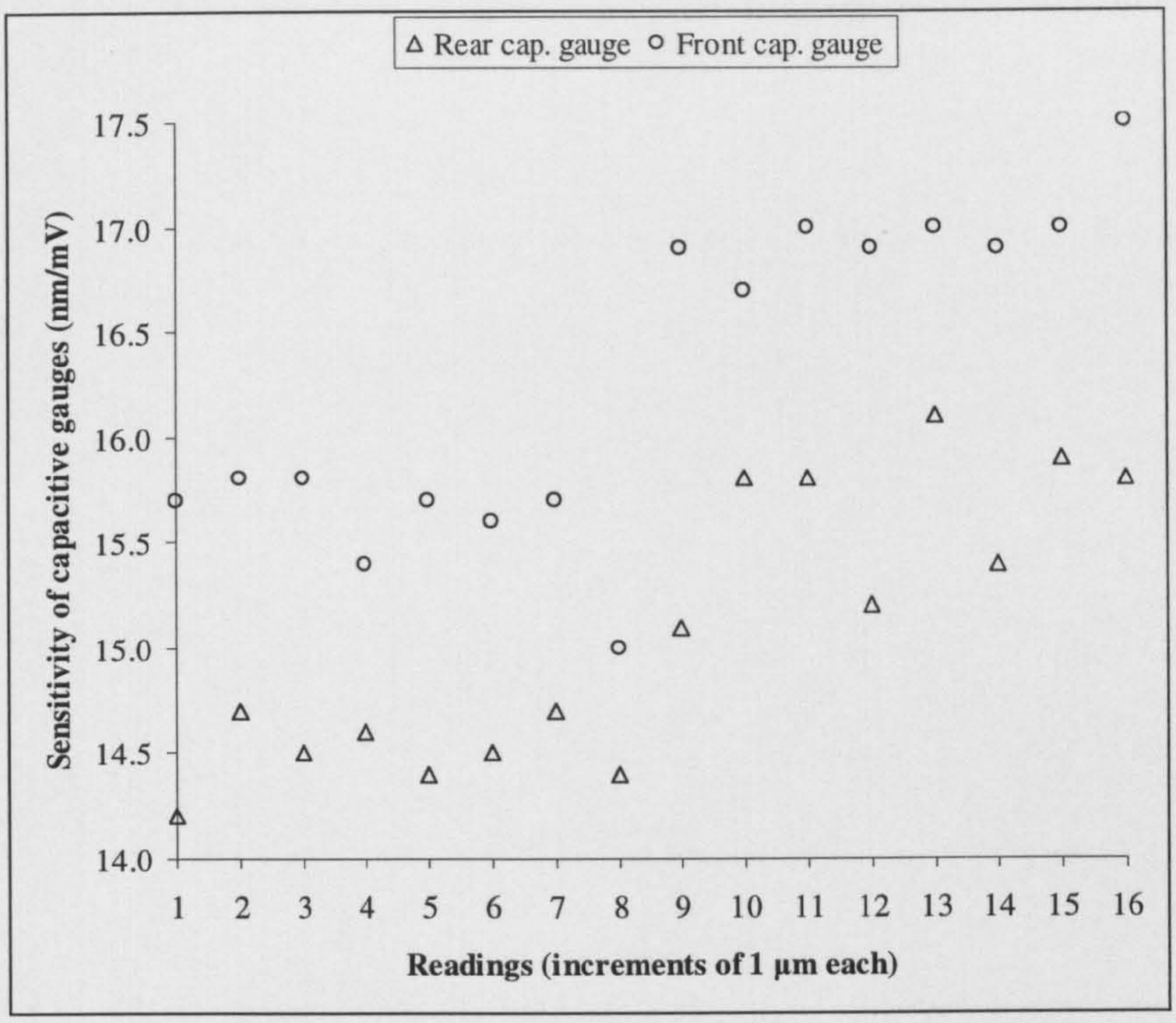


Figure 5.5 Change of sensitivity of the two capacitive gauges with displacements from the DPT gauge. The first eight readings were taken at the start of the selected operating range of voltages, and the last eight readings were taken at the end of this range.

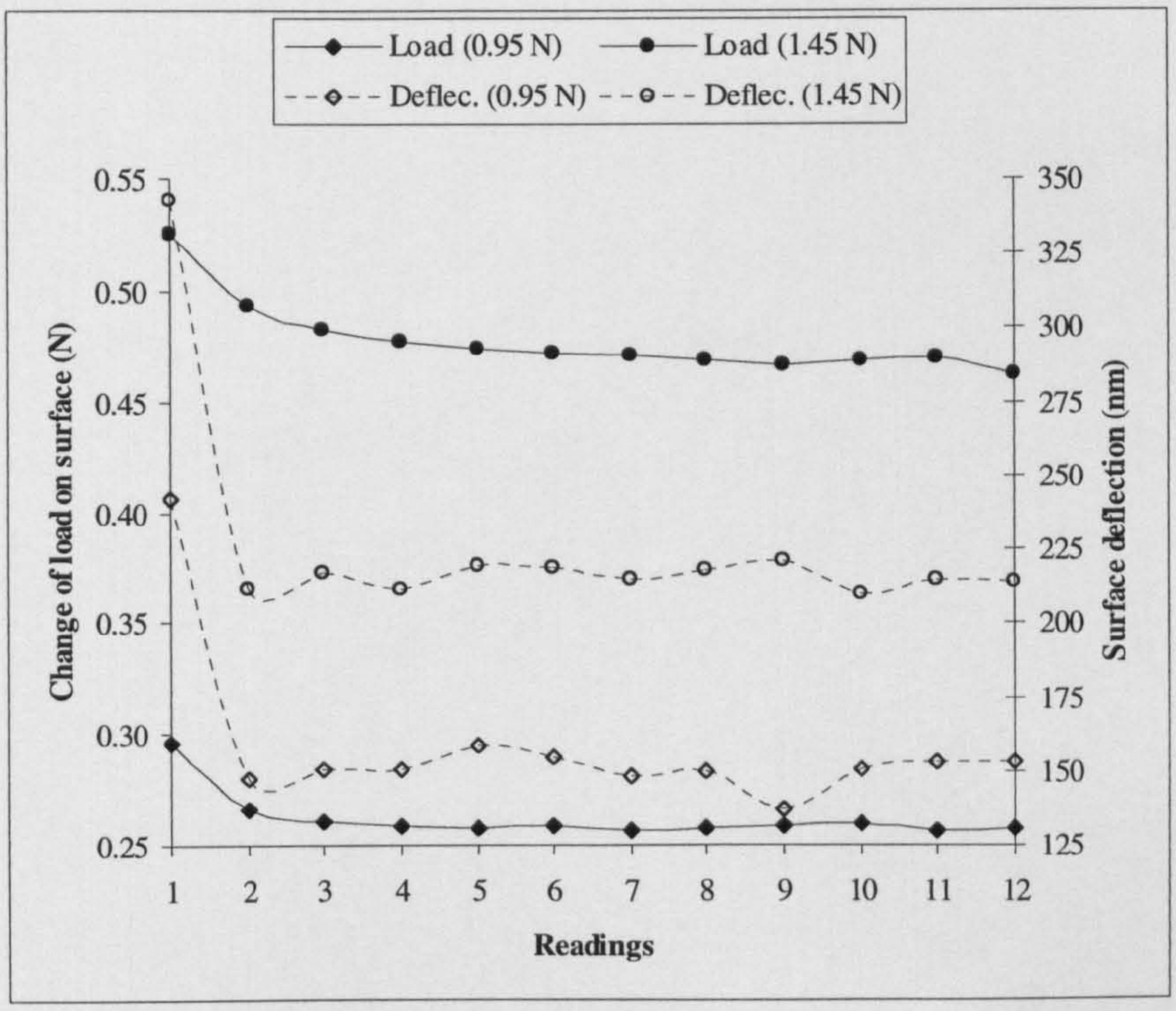


Figure 5.6 Behaviours of contact load and deflection on two same-point repeated contact tests on a surface of a gauge block.

CHAPTER 6 ADVANCED INVESTIGATIONS ON PARAMETERS AFFECTING CONTACT REPEATABILITY

6.1 INTRODUCTION

The initial investigations accomplished on the old test-rig, which have been discussed in Chapter 4, revealed a substantial clue to some effects from those parameters involved at the contact interface. From the various brief tests performed on this test-rig, the surface roughness was the chief source of variation in the deflection of the probe tip. Successive contacts sampled at the same location on the surface, as well as those randomly sampled across it, showed a hint of other contact parameters that could also affect their repeatability. Tests of repeated deflections on contaminated surfaces did not provide enough information about the effect of surface conditions on the mean deflections. Similarly, the change in contact tip size gave no clear evidence of its effect on such deflections, but it revealed a vague discrepancy in repeatability. However, the limited accuracy of this test-rig was the main reason for these concise and quick tests, and it stimulated the need for a careful investigation with more reliable results.

This chapter discusses the advanced experimental study accomplished on the new test-rig (discussed in the previous chapter) which was constructed to eliminate the most serious of those deficiencies encountered with the old test-rig and to offer a more precise environment for rigorous testing. This study includes a wider range of parameters to simulate the real situations of gauging with contact indicators and to establish a better understanding of the importance of errors contributed to these measurements due to the contact uncertainty. In these new experiments, improved

methods will be followed for testing and evaluating repeated contacts since the new test-rig facilitates more controlled test procedures with considerably better accuracies.

Surface deflections will be comprehensively investigated against a larger variety of surfaces of different roughnesses, materials, and cleanliness under different loads (within the regime of interest) and contact tip sizes. Repeatability of the two types of successive deflections (i.e., same- and different-point) will be the main concern in this investigation, and Sections 6.3 and 6.4 will be dedicated to each type, respectively. Rotations of the contact tip will be introduced in this study, as another important parameter that could exist as a result of the contact interaction during actual measurements. Tip rotation is also an important diagnostic of the test-rig performance, especially since the gauging has a small Abbe offset (see Chapter 5). The analysis of deflections behaviour under incremental loading will be re-examined here through more extensive tests, and better governing relations will be suggested in Section 6.5. The effect of surface conditions on the two contact repeatabilities will be rigorously checked on clean and contaminated surfaces (of real contamination: left without cleaning in the test laboratory environment after machining) with two contact probes of different tip diameter, and discussed in Section 6.6. A further discussion will be raised in Section 6.7 regarding the effect of the contact tip size on the surface deflections and their repeatability.

6.2 PREPARATION OF SPECIMENS

Nine square blocks (specimens), 20 mm in length and width and 6 mm in thickness, were manufactured nominally identically from three different materials: mild steel, copper, and aluminium. The two square faces of each specimen were prepared, separately, by grinding to reduce the parallelism error between them. This was,

however, done at the same time for the three specimens of each material to get a similar sort of roughness quality on all surfaces of the same material. Then, one of these two square surfaces was either ground (once more) or sandpapered, independently, to obtain the required roughness quality for testing. The three surfaces of each material were given three different roughness qualities. All surfaces were frequently cleaned, prior to every contact test, by hand using Isopropanol and cleaning tissues.

The Form Talysurf instrument of Taylor-Hobson was used to measure the root-mean-square roughness parameter, R_q, of these prepared surfaces at 0.8 mm cut-off. Multiple scans were taken across each surface and the average R_q value was calculated and is listed in Table 6.1 for all the surfaces. The first column of this table gives qualitative descriptions, for the roughness quality of three groups of surfaces of different regimes of roughness, which will be used in this context. The codenames, shown in the second column, were arbitrary chosen, for simplicity, and assigned to each surface. They will also be used throughout the discussion.

An initial set of tests on these specimens showed a considerable amount of inconsistent results, which made the search for any correlations a difficult task. This discrepancy in the results was traced to the interaction between the bottom surface of the specimen and the surface of the base plate of the test-rig. Obviously, this interaction is mainly caused by the roughness of both surfaces, which transmit additional "unwanted" displacements to the measurements. Moreover, the inevitable presence of dirt and surface debris particles at the interface may also cause a sort of unstable mounting of the specimen on the base plate which, in turn, may lead the specimen to rock as the probe tip contacts its top surface. Commonly in metrology and indentation, such effects are neglected on the grounds that the much larger nominal area of the base means that pressures, and deflections, are small. This assumption is inadequate at the sensitivity available here. Therefore, to overcome this effect, every specimen was glued

(using Araldite epoxy adhesive) to a glass microscope slide, as shown in Figure 6.1. Accordingly, the test-rig was modified to include the U-shape holder and the clamping plate in order to secure both end of the glass slide to the base plate in the manner illustrated in Figure 5.2a. The set of specimens given in Table 6.1 was used only to carry out the same-point and different-point repeated contact tests and the load-surface deflection tests which will be discussed in the following three sections. Later, another set of specimens was manufactured, to investigate more contact parameters, and these are described in Section 6.6.

6.3 REPEATABILITY OF SAME-POINT CONTACT

This type of repeatability was introduced in Chapter 4, which reported brief tests only on steel surfaces of different roughness values. Using the new test-rig, it has been carefully investigated through tests with different combinations of load, material, and surface roughness. During each test, the probe vertical position on the surface was repeatedly measured after applying the same dead-weight on the test-rig carriage using nominally the same contact point on the surface. Twelve probe position readings were collected from each of the two capacitive gauges in each test and five tests were performed on each specimen with five different loads. The five points for contact on the specimen's surface were randomly selected within a 1 cm^2 area around its centre. Readings from the load-cell were also collected in order to monitor its behaviour during the test as well as to observe the change in the contact loads. Hence, a total of forty-five tests have been analysed in this investigation and Appendix C1 provides full details of the readings and their plots for each of these tests.

At the beginning of each test, an initial dead-load of about 0.2 N was applied on

the contact probe to keep it in a firm contact to the surface throughout the test. Hence, the probe was not moved off the surface between the repeated contacts, unlike what had been done in similar tests using the previous test-rig. The initial load was adjusted by using a slightly smaller counter-weight over the pulleys than the weight of the carriage. However, the difference between the carriage weight and the counter-weight was adequate to allow the contact tip to move down freely towards the specimen's surface. This movement was carefully controlled to establish the first contact on the surface as gently as possible, guided by hand. When the tip was believed to have been bumped against the surface, another location for contact was chosen and the test restarted. Any sort of impact by the contact tip on a certain point on the surface must produce some degree of damage to surface asperities, which makes that point unsuitable for testing.

After carefully applying the required weight for the particular test on the carriage, the change in the voltage reading of each gauge was recorded as representing the local displacement of the contact tip due to the change in contact load. Hence, the resultant displacement (surface deflection) at the tip location will be the average of the two readings, since both gauges are located at the same distance from the tip. However, the difference between these values (divided by the centre distance between the two gauges) gives the relative resultant rotation of the tip about the x-axis of the test-rig, refer to Figure 5.1. Similarly, the change in voltage reading given by the load-cell represents the absolute increase in load beyond the adjusted initial load. Then, the local compliance of the contact is the local displacement divided by this change in force.

All tests in this investigation have shown, without exception, a high surface deflection value at the first application of any weight, regardless of how big or small that weight was. Succeeding weight applications gave much lower deflection values. The difference of their average from the first one depends mainly on the change of contact load and, probably, on the specimen's material and its surface roughness as

well. This will be discussed further later in this chapter. A similar phenomenon has been observed by Chetwynd and Davis [97] under the different contact regimes of stylus topography on surfaces that were heavily contaminated with oil.

As explained earlier in Chapter 4, this pattern of behaviour seems reasonable since the first application of the dead-weight is likely to cause almost all of the plastic damage of the highest asperities that are in contact with the tip, in addition to the elastic deformations. Releasing that dead-weight will allow these asperities to recover nearly all of their elastic deformation, pushing the tip upwards from the surface. At this stage, the tip will be at the starting (reference) position for the next application of the same dead-weight, but in a lower position (and closer to the asperities bases) than before. Therefore, the compliance values for the following applications of the dead-weight will nominally be lower. The deformations are probably nearly elastic in later contacts since they tend to show the same repeatability.

In most of the tests, a significant reduction in the deflection values was noticed over, at most, the next three or four applications of weight after the first one, as shown by the plots given in Appendix C1. Obviously, this behaviour adversely affects the repeatability of contact in the first few cycles. It is believed that further plastic deformation was taking place after each of these applications of weight in addition to the elastic deformation. It could be that there was not enough time during the test to allow for the full plastic deformation to develop after the first application of weight and, hence, it was being completed in the later cycles. However, the contact is maintained for 30 seconds or more and, so time dependent behaviour would not be greatly expected in metals at room temperature. An alternative explanation would be that elastic recovery alters the shape of asperities such that further plastic deformation can occur in the following cycle. This settling behaviour occurs independently of the specimens' material and the contact load; as observed from all the tests, there is no evidence of a

direct relation between this behaviour and these two main contact-affecting parameters. Moreover, the surface roughness of specimens did not affect this behaviour. It appeared, without a definite trend, across all roughness values used within each material and each load. As a result of this and the initial contact force variation discussed in the previous chapter, the compliance of contact showed a similar behaviour in all the tests; sharply decreasing trend at the earliest few cycles and, then, nearly settling to a certain value at the latest ones.

In almost all of the tests that showed the decreasing trend of the deflection values the last six readings in each test showed little effect of such behaviour. Thus, position repeatability and rotation of the probe tip, due to the repeated contact on nominally the same point on the surface, have been studied here by considering only these six readings in each of the forty-five tests performed in this investigation. Although this number of data looks somewhat small for statistical reliability, the variation of any further readings was expected to be within the same range that has been observed. Drift and noise from temperature variation, floor vibration, and the instrumentation will be the major limits on the repeatability of further readings. It is recognised that the larger deflections caused by the first loading cycles will be the appropriate values for some metrology error budgets, which will be discussed in the next section.

Table 6.2 summarizes the results of the last six readings from all of the same-point repeated contact tests performed on the steel, copper, and aluminium specimens. The second column of this table provides both the dead-weight, which was repeatedly applied on the carriage, and the actual average change in contact load measured by the load-cell in the last six readings of each test. This measured change in load was smaller because of the random bearing friction behaviour of the slides and the pulleys, which has been explained before in Subsection 5.4.3. The resultant deflection at the contact tip location represents, as mentioned before, the mean value of the two average vertical

displacements read by the front and rear capacitive gauges. As discussed in the previous chapter, the tip is located midway between the two gauges. The rotations (or tilts) of the tip are given in micro-radians and obtained by dividing the difference between these two average displacements by the centre distance between the gauges, which is 28 mm. A measure for the scatter of deflection readings at the tip area is taken as the 95% confidence interval which is $\pm 2\sigma$, where σ is the standard deviation of these readings. The average genuine compliance of contact for each test is calculated by dividing the average resultant change in the probe tip position by the average change in the relative contact force on the surface.

These resultant deflection values of the three surfaces are plotted together for each of the three materials as shown in Figure 6.2. The corresponding Hertzian deformation curves (for a steel indenter) are also plotted in this figure for the purpose of comparison. The plots show a generally consistent increase in the resultant "elastic" deflection with the contact load on each surface tested. This is no more than expected behaviour from elastic or plastic theories. However, if the tip covers many asperities, smaller "elastic" deformations will be expected as contact loads rise to a level where most of the asperities' damage is plastic, at loads approaching the end of the elastic-plastic deformation range. To a large extent, this could explain the fall of the resultant deflection value on the *St3* surface when the 0.75 N load was repeatedly applied. Another fall was also observed at the next higher load (0.95 N), as shown in Figure 6.2a. Similarly, this was also noted when the 0.95 N load was used on the *Co3* surface, see Figure 6.2b, and when the 1.15 N load was used on the *Al1* surface, see Figure 6.2c. But, both the *St3* and *Co3* surfaces showed an increase afterwards at, co-incidentally, the same repeated load of 1.15 N. However, this suggests that some other parameters must also be influencing this "anomalous" behaviour.

It is well-known that any engineering surface manufactured by regular machine-

tool operations is expected to show a random distribution of asperities of dissimilar height, size, and shape, regardless of the surface finish quality achieved. The commonly used surface roughness measurement methods, which use stylus mechanisms, are regarded as two-dimensional measuring techniques and are not able to provide a complete evaluation (three-dimensional characterisation) of roughness variations across the whole surface. This is also true even by taking a succession of parallel tracks, in order to cover the area of surface in concern, which is a non-conventional way of surface mapping that involves many problems as discussed comprehensively in [98]. The different surface roughness parameters obviously represent only the local features at those areas of the measured surface where the stylus tip has come near.

Surface function is largely manifest in three dimensions. For instance, tribological and sealing (functional contact) qualities are related to the three-dimensional nature of the surface. Hence, the analysis of a surface based often on a single profile is misleading and always limited in scope. There are many problems with assessing a profile taken from a surface, some largely statistical. The profile may fail to take in the major features of a surface by missing predominant summits and valleys. Features such as ridges and valleys are indistinguishable from peaks and pits when it comes to calculating height parameters. However, the two-dimensional analysis can often be used as a process monitor, and it is able to provide a simple indication whether or not the topography is changing [99].

Therefore, the comprehensive assessment of surfaces is of a great importance for predicting their contact-related performance, yet the deformation (or wear) process of their asperities due to any type of contact remains a random process. In other words, it is possible to determine precisely neither the number and geometry of those asperities from the surface, which will suffer the contact load, nor the way in which these

asperities will deform and their final geometry after deformation. Nevertheless, we may conclude in general from those major mathematical models reviewed in Chapter 3 that the larger the number of asperities in contact, the smaller is their deformation. Such numbers turn out, in practice, to be questionable, at least, because of the roughness variations all over the surface. So, regardless of the mean roughness value, it is normal to observe variations in responses to the loads over the same surface. Therefore, at load regimes of surface gauging, both elastic and plastic deformations measured on the surface depend, to a significant extent, on the set (group) of asperities which happen to be present within the contact area and on their particular behaviour in resisting the load. So, a case could exist where a certain load on a closely grouped set of asperities in the contact area causes less deformation than a rather lower load on a different less-condensed set of asperities. Encountering such an asperity group after some initial deformation of others could well explain the observations that do not follow the other normal trends shown in Figure 6.2.

This anomalous behaviour could also explain the odd order of the plots in Figure 6.2 with respect to the change in roughness. The usual behaviour is, however, that surfaces of low roughness should nominally show smaller deflections than those of high roughness, at the same E^* values and repeated loads. The *St3* surface gave, in general, the highest "elastic" deflections, as shown in Figure 6.2a, while the other two surfaces, *St1* and *St2*, showed no significant difference between their deflections that correlates to the roughness difference between them. The plot of the *Co1* surface in Figure 6.2b showed a consistent location of its plot (as it is the highest one) relative to the other two plots, but the *Co3* surface did not reveal the lowest deflections except at the largest two loads used. The plots of the three aluminium surfaces in Figure 6.2c show nearly the same odd location relative to each other: the *Al3* surface gave the lowest deflections, but the *Al1* surface did not give the highest deflections. So, such inconsistencies between

deflections in all these groups of surfaces are genuinely because of the random variations of roughness across the same surface. In addition, due to this effect it appears from these results that no clear systematic relation could be observed between the surface roughness and these average "elastic" deflections; the general effect of roughness on such deflections was, however, very obvious.

Comparing these deflections with the Hertzian ones for smooth surfaces, an appreciable amount of the observed deflections on each material are unexpectedly below the curve of the ideal case. All these deflections are expected to be higher than those computed by the Hertzian relation because of roughness. The anomalous behaviour might again be responsible for such inconsistency. On the other hand, the Hertzian analysis comprises of several ideal assumptions, as discussed in Chapter 3, which are impossible to be encountered in practice. This could be another source for this inconsistency. In Section 6.5, there will be a more emphasis on those aspects that should be taken into consideration when generating such comparisons.

Rearranging the plots of the elastic deflections, Figure 6.3 classifies these plots according to the roughness of the surfaces. The plots of the roughest surfaces, in Figure 6.3a, shows a consistent location relative to the E^* variation, but note that the *Al1* surface is of a rather higher roughness than the other two surfaces, *St1* and *Co1*. The elastic deflections of the *St2* and *Co2* surfaces, plotted in Figure 6.3b, are nearly the same, while those of the *Al2* surface are the highest, as expected and observed between the roughest surfaces before. Figure 6.3c shows a reverse order for the plots of the rough surfaces against the E^* variation, most likely, due to the relatively large differences in roughness between these surfaces becoming a more dominant effect than the E^* variation. The *St3* surface which has the highest roughness shows, in general, the highest elastic deflections and the *Al3* surface which has the lowest roughness shows nearly the lowest deflections, with some exceptions at the two largest loads because of,

it is believed, the anomalous behaviour discussed earlier. However, from these data, it is also noticeable that the roughest surfaces are able to reveal the effect of E^* on the elastic deflections clearer than the others. Similarly, no clearly relevant ratios could be drawn from the E^* variation and the trends of these deflections with loads, yet the general effect is reasonably realized.

It worth mentioning here that, in addition to the roughness variation across the same engineering surface, E^* variability is also expected as a result of the variation of material properties. In practice, perfect homogeneity of the surface bulk material is unachievable due to, for example, the inevitable presence of impurities. Moreover, due to the applied machining forces in material-cutting processes, a strained and metallurgically altered material structure is usually formed at the final exposed layer of the surface: its thickness is measured in tenths of a millimetre [100]. Figure 6.4 shows a schematic of this layer for a turned surface. The mechanical properties of such a layer are obviously different from those of the bulk material. This figure also shows the oxide and adsorbate layers which will exist in most environments. All these different layers are expected to affect greatly the material elasticity and, hence, the consistency of contact deflection across the same rough surface.

Furthermore, as the contact tip progresses in deforming surface asperities once the load is applied, a strain-hardening effect is also expected at those asperities in contact. This is in addition to the increase of the real contact area because of the new contacting asperities and the deformation of those already in contact. In the exaggerated section through a contact region shown in Figure 6.5, asperities 1, 3 and 5 have larger contact spots than before loading due to their plastic distortion which could cause, it is believed, more straining of them. Such an effect could inhibit further deformation from taking place at the same application of load. However, asperities 2, 4 and 6 will be in contact with the tip if extra load is applied, and they then will exhibit the same behaviour.

Obviously, the number of contacting asperities (roughness) plays here a significant role as well. So, the effect of roughness and E^* variabilities across the same surface are incorporated into the (consistent or inconsistent) resultant variation of surface deflection using the same contact load.

Considering the scatter values of these repeated elastic deflections given in Table 6.2, a narrower scatter range is broadly noticed as the repeated load increased on the surface. Three of the nine surfaces examined have not resulted in the highest scatters at the lowest load used as the others, but shown this at the next lowest load, i.e. at 0.55 N. Such high scatter values of the low elastic deflections is likely to be due to the effects of measurement noise, since most of these deflections were low enough to be significantly vulnerable to the noise figures mentioned earlier in Subsection 5.4.4. At the highest repeated load used, almost all of the surfaces have shown the highest deflections with scatter ranges within around ±12 nm which is approximately half of that shown by the lowest deflections. However, it is believed that the operating noise of the test is the main effect on the repeatability of those elastic deflections resulted from the same-point on the surface, as there are no certain correlations found between the variation of the deflections scatter and either the surface roughness or the E^* value.

The rotations of the contact probe in the less-stiff torsional axis of the test-rig, i.e. about the x-axis, were calculated, as mentioned before, by dividing the difference between the two displacements sensed by the two capacitive gauges by their mean spacing. The rotation values shown in Table 6.2 were determined based on the average of the last six displacements collected in each test and, obviously, relative to the original orientations of the probe tip before applying the loads. For convenience, the displacements of the front gauge were subtracted from those of the rear one; hence, the tilts of the probe tip to the rear side (clockwise about x-axis) were considered as positive rotations. Negative values of this rotation represent counter-clockwise probe tip

rotations about the *x*-axis, i.e. tilts toward the front side of the rig. Tip rotation may occur from two sources: imperfection in the probe guidance system and the lateral force components resulting from off-centre contacts with asperities. On nearly every surface there was a general increase in the probe tilt to either side of the rig as the repeated load increased. This seems reasonable behaviour since higher loads nominally mean higher surface deflections and, in turn, more possibilities for the probe tip to interact with the surface asperities and rotate accordingly. Exceptions to this behaviour were mainly noticed on the *Co1* and *Al2* surfaces. Complete consistency would not be expected on random process with small number of events. However, these are the same surfaces for which the random combination of asperities in the contact area and their deformation process was discussed before. These surfaces nevertheless showed a consistent increase in deflections with the loads used. Hence, it could be concluded that for such type of repeated contacts, rotations of probe tip are not absolutely guaranteed to increase with the repeated loads on every rough surface. The random nature of the asperities group contacted has forced the tip to comply with the final arbitrary geometry of these asperities and also to show a random rotation. Consequently, given the available data, neither the surface roughness nor E^* variations have clear effects on these computed contact tip rotations.

Theoretically, from the linear relation between load and deflection within the elastic regime, the "elastic" compliance of contact on any point on the surface should remain constant regardless of the load magnitudes applied. In this investigation, most of the surfaces tested at the different repeated loads, separately, have shown fairly major variations in their average relative compliance of contact. In other words, such compliance was significantly inconstant across the same surface as the repeated load increased, as given in Table 6.2. For instance, the *St3* and *Co3* surfaces (which have resulted in irregular deflection behaviour with load) showed a sharply decreasing trend

as the load increased; the *Co1* surface gave similar behaviour but with relatively lower descent rate. The *Al1* and *Al2* surfaces showed no definite compliance trends but relatively large variations. The rest of the surfaces gave more stable compliance values but with at least one inconsistent value, such as the 0.75 N load on the *St1* and *St2* surfaces. Apart from the contact force uncertainty caused by the slides' bearing performance, the anomalous behaviours of "elastic" deflections, which were discussed earlier, was the main reason of such variations in the elastic compliance. Hence, for these contact conditions, neither the surface roughness nor its E^* value have shown clear effects on such compliance. While surfaces of high roughness magnitudes and low E^* values might be expected to show greater contact compliances, practice is much less predictable.

6.4 REPEATABILITY OF DIFFERENT-POINT CONTACT

Using the previous test-rig, the repeatability of deformation of surface asperities across the surface was briefly examined on only one steel surface of 0.78 µm R_q with a 2 N load, as described in Subsection 4.4.2. Subsequently, it has been studied in more detail using the advanced test-rig through several tests on the same specimens used earlier. In each test, the vertical position of the contact probe was repeatedly measured on different points on the surface, after applying the same dead-weight on the test-rig carriage. Seven probe position readings were collected from each capacitive gauge in every test and three tests were performed on each specimen with three different loads. The twenty-one points for contact on the specimen's surface were randomly chosen within a 1 cm^2 area around its centre. The load-cell readings were also collected during the test to compute the change in contact forces. Thus, a total of twenty-seven tests have been

analysed in this investigation and Appendix C2 provides full details of the readings and their plots for each of these tests.

At all points chosen on surfaces for contact, the probe was first brought into contact, gently, with the same initial dead-load of 0.2 N, approximately, which was adjusted by the counter-weight over the pulleys. Then, the required weight was carefully applied on the carriage and the change in the voltage readings of the two capacitive gauges and load-cell were recorded. The applied weight was removed, afterwards, and the probe was pulled few millimetres above the specimen's surface. The glass slide that holds the specimen was, then, released (from the pressure of the U-shape holder), randomly moved (for gauging another point on the surface), and clamped again. From this test procedure, it can be noticed that all the readings collected in this type of test are similar in nature to the first contact reading, which has been encountered earlier in every same-point repeated contact test.

Table 6.3 summarizes all the results of the twenty-seven different-point repeated contact tests performed on the steel, copper, and aluminium surfaces that have been used in the previous tests. The second column of this table also compares the repeated dead-weight applied on the carriage to the average real change in the contact load on the specimen's surface. Likewise, this measured change in load was, in general, smaller than the applied one due to the bearing friction of the slides and the pulleys. The four parameters, shown in the next columns of this table, were calculated as given in the previous section; the standard deviation σ of the resultant surface deflections at the contact tip is shown here instead of the 95% confidence scatter range. The last column gives the general compliance value of each surface, which is basically the slope of the linear least-squares best-fit straight line of the three deflection values that passes through the origin point.

Figure 6.6 presents the plots of the average resultant deflection values, with those

of the three surfaces of the same material are shown together for comparison. The corresponding Hertzian deformation curves (for a steel indenter) are also plotted in this figure for the purpose of comparison. In general, a consistent approximately linear increase of such deflections with the change in contact load could easily be observed from these plots. This linearity is as would be expected from the basic relation between stress and strain within the elastic regime. Another consistent increase might be expected with the surface roughness, since, as mentioned before, surfaces of high roughness values (low number of asperities in contact) might show greater deflections than those of low roughness values (high number of asperities in contact). The three copper surfaces clearly revealed this relation, as shown in Figure 6.6b, with distinctive locations of their plots relative to the roughness variation. The *St1* surface also gave high deflections, but the *St3* surface generally resulted in fairly high deflections compared to those of the *St2* surface, as shown in Figure 6.6a, while the opposite behaviour should be true. This could also be observed, easily, from the slope values of the deflections' best-fit straight lines of these two surfaces, refer to Table 6.3, where the *St3* surface was shown to have a slightly higher slope (compliance) of its deflections trend. The slope values of deflections trends and Figure 6.6c also show that the *Al2* surface has a higher compliance than the *Al1* surface by more than 10% of magnitude because of, obviously, its higher deflections. The anomalies with respect to the simple expectations are quite similar to those reported in Section 6.3. It is conceivable that the higher than expected compliance values derive from bottom surface effects whereby some bending of the glass slide occurs under loading since it is only clamped to the sides of the loading region. However, the consistent behaviour in different test set-ups suggests it is a real surface behaviour.

As observed on those deflections discussed in the previous section, there is also a significant amount of these average deflections on each material which is surprisingly

below the Hertzian curve. Due to roughness, all these deflections were expected to be higher than those of the ideal case of smooth surfaces. One source of such inconsistency might be the roughness variability across the same surface. As seen in Chapter 3, the Hertzian analysis embraces some ideal assumptions which are impossible to be fulfilled in practice. However, this could be another source for this inconsistency. For more convenience, the discussion about the comparison between these experimental values and the Hertzian ones was shifted to the next section.

Figure 6.7 rearranges these deflection plots in terms of the roughness quality (or description) used in this context. In general, the increase of the surface deflections with the decrease of the E^* value was observed. In Figure 6.7a, the plots of the roughest surfaces show fairly distinctive locations relative to the E^* variation. Clearly, the general compliance values of these surfaces, given in Table 6.3, also show such relationships. Similarly, the rougher surfaces also revealed this relation, as shown in Figure 6.7b, although, the *Al2* surface gave deflections higher than expected at the two largest dead-weights used. The *St3* surface (which also gave deflections higher than expected) was also higher in deflections (see Figure 6.7c) and, consequently, in compliance than the *Co3* surface. Although the *Co3* surface has nearly 0.1 μm roughness lower than the *St3* surface, the opposite behaviour should have resulted since, it is believed, the E^* difference is more dominant than this roughness difference.

From these available results, there is no clear systematic relation between the surface roughness and the different-point repeated contact deflections. This seems a reasonable conclusion since the theoretical models of contact additionally consider the effect of other surface parameters, as presented in Chapter 3. It is also since, as believed, the variations of roughness and E^* across the surface itself makes each of the repeated contacts happen at a group of asperities of unique characteristics and, hence, causes a different deflection on the surface. However, the effect of the roughness, in

general, on the surface compliance was strongly noted.

According to the Hertzian analyses presented in Chapter 3, the contact approach distance (surface deflection), given by Equation 3.3, is inversely proportional to 2/3 of power of the contact modulus, E^*. This relation was reflected fairly closely in the available results. The ratios of compliance of the *St1* to *Co1* surfaces and of the *St2* to *Co2* surfaces are 0.98 and 0.80, respectively, while the ratio of $(E^*)^{2/3}$ of copper to steel is about 0.82. A compliance ratio of around 0.70 was found between the *Co1* and *Al1* surfaces, which is nearly identical to that one between the *Co3* and *Al3* surfaces, while the ratio of $(E^*)^{2/3}$, in this case, is about 0.79. The *St1* and *Al1* surfaces showed a closer compliance ratio (of around 0.68) to their $(E^*)^{2/3}$ ratio, which is about 0.65. However, such differences between the two compared ratios above were probably due to, mainly, the difference in roughness between the two surfaces considered. The variation of deformation (due to the variations of roughness and E^*) across the same surface must also have played a significant role in stimulating these differences since it eventually affects the surface compliance.

The standard deviation of the different-point repeated contact deflections also showed an increasing trend with both the contact load and surface roughness on, generally, all of the surfaces tested, as shown in Figure 6.8. This increase in the deviations seems obviously related to the increase of the magnitudes of deflection. Such trends were almost linear on those surfaces of roughness values lower than 0.6 μm. They appeared to be curves of reducing slope on the roughest surfaces, namely, the *St1* surface in Figure 6.8a and the *Co1* surface in Figure 6.8b. It is believed that these roughness-related deviations are unlikely to show much greater (probably similar) magnitudes on the same rough surface as the loads increase: the larger the deflections the smaller the effects of roughness. The deflections observed on both the *Al1* and *Al2* surfaces revealed, accidentally, an unexpected standard deviation (as shown in Figure

6.8c) when the second load, i.e. 0.35 N, was used, but the general increase with contact loads was also found.

Figure 6.9 also shows no clear clues for the effect of E^* variation on such deviations – unlike the deflections – since, in each of the three figures, the plots did not show a consistent arrangement with respect to the surface material. This seems a reasonable result because surfaces of dissimilar materials, and within the same range of roughness and tested under the same repeated load, should nominally show similar sort of deviations of contact deflections. The variations of roughness and E^* across the surface itself are solely responsible for such deviations. From the theoretical contact models, if the surface was assumed to be of a uniform distribution of asperities, the deflection caused by a certain load should nominally be the same at each point on that surface and, consequently, the deviations will nominally be zero.

The same idea could be pointed out here as in Subsection 4.4.2 before, where a similar test was performed on the previous test-rig, that the mean deflection represents a systematic error on position measurement that could in principle be compensated by calibration, and the deviation represents the effect of roughness variation across the surface. The theoretical models of the contact of rough surfaces could readily be used to estimate the surface deflection. The repeatability of contact over the entire surface depends not only on the magnitude of its roughness but on the degree of its real anisotropy, as well. However, it is expected that the numerical methods of contact could be able to provide an estimate for such repeatability if they are used with the actual three-dimensional map of the surface.

As in the same-point repeated contact tests, the average relative rotations of the contact probe about the x-axis of the test-rig, given in Table 6.3, showed a general increase, in either of the two directions (front or rear), with the contact load on all the surfaces. This is because of the increasing interaction of the probe tip with those

surface asperities within the contact area. However, when the second dead-weight (i.e., 0.35 N) was used, all the roughest surfaces accidentally showed a magnitude of rotation inconsistent with the current trend of rotation on the surface. For instance, on the *St1* surface, the average rotation of 0.4 μrad, observed when the second dead-weight was repeatedly applied, is not in between the other two average rotations of 1.4 and 1.8 μrad noticed at the first and the third repeated dead-weights, respectively. The same words could be said about the *Co1* and *Al1* surfaces. As pointed out before, such behaviour seems expected due to the random distribution of asperities across the surface (anisotropy), which has been discussed in the previous section. This behaviour also confirms that such rotations are genuinely because of the contact interaction, since it appeared strongly on the roughest surfaces and not due to errors inherent in the test-rig or due to tilts in the specimens.

Moreover, this odd dispersion in the probe tip rotations on the roughest surfaces may give a clue to the roughness effect on these rotations. To investigate this in more detail, the raw values of rotations of all the repeated contacts performed are plotted in Figure 6.10. From this plot, the effect of contact load is clear. It is also evident that those surfaces of the least roughness (*St3*, *Co3*, and *Al3*) showed, in general, the least scatter in these rotations. The roughest surfaces revealed, as expected, the opposite at nearly each dead-weight used. This effect seems plausible since the asperities on the rough surface are more spaced (bigger gaps between them) than on the smooth one. This obviously means that a smaller number of asperities sustain the contact load and, hence, there are more possibilities for relatively larger random probe tip rotations to take place (less constrained probe tip).

Since the surface of the soft material could also produce, at the same contact load and roughness, a larger interaction with the probe tip than that of the hard material, then, it can be expected that the rotations of the probe tip are also affected by the E^* value of

the surface material. This effect is somewhat evident from Figure 6.10. For instance, the *Al2* surface gave a greater scatter in the rotation values than the *St2* surface at all the dead-weights used. The same behaviour could be easily seen between the *Co1* and *St1* surfaces, and between the *Al3* and *Co3* surfaces at the 0.35 N dead-weight.

6.5 LOAD-SURFACE DEFLECTION TESTS

The behaviour of surface deflections induced due to the incremental loading on a single point on the surface has been briefly investigated, as described in Subsection 4.4.1, using the previous test-rig. With the newly-designed test-rig and the same nine specimens listed in Table 6.1, this behaviour has been studied in much more detail, as discussed in this section. A point on each of the nine surfaces was randomly selected and incrementally loaded, up to around 3 N, using in each case the same steel probe tip of 5 mm diameter. A dead-load of about 0.2 N was first adjusted by the counter-weight over the pulleys to get the probe to an initial rigid contact with the surface. Then, the dead-weights were repeatedly increased and placed on the carriage according to the same sequence followed in the previous tests. So, the probe was not released from the surface throughout the test. After every application of the new dead-weight, the surface deflection was computed based on the voltage readings collected from the two capacitive gauges. Fifteen deflection values were eventually obtained for each surface. The load-cell voltage readings were also used to compute the changes in the contact force. Full information on all these readings is available in Appendix C3 which provides the raw readings and their initial analyses for all the tests.

Figure 6.11 illustrates the behaviour of surface deflection with the incremental load observed at each of the nine points tested on all the surfaces; arranged according to the surface material. For the purpose of comparison, the ideal (Hertzian) curve of the

approach distance is also plotted for each of the three materials. From this figure, the effect of the surface roughness on this behaviour is very obvious, as noticed from those similar tests described in Chapter 4. Since most of these behaviours showed a more linear trend than the ideal one, the slope of the least-squares best-fit straight line (which has an intercept at zero deflection) of each of these deflection curves is computed and also given in Figure 6.11. This result is somewhat like the previous one (which, in general, revealed deflection curves of a similar trend to the ideal one, as shown in Figure 4.5) when considering the same range of loading, i.e., up to 3.5 N. Hence, this observation may lead us to conclude that this deformation of asperities at such low contact forces is most likely to follow the common linear stress-strain trend of the elastic compression rather than the Hertzian trend. This linear relation between the deflections and the contact loads was also observed at the different-point repeated contact tests described in the previous section.

On the other hand, Figure 6.11 also shows that a considerable number of these deflection curves are unexpectedly below the corresponding Hertzian curve, especially those of the smoothest surfaces. This is unlike the earlier results shown in Figure 4.5, under the same range of loads. But, the four specimens used in the early tests were not fixed to the base of the previous test-rig, which means that their bottom surface interaction with the base surface may have interfered with the measurement of deflections. The load-surface deflection curves shown in Figure 6.11 are more genuine. This fact leads to the conclusion that the behaviour of the incremental deflection (obtained in the way described here) of a certain point on a rough surface is not necessarily larger in magnitude than the Hertzian behaviour because of the roughness effect. Moreover, the comparison between these two behaviours is questionable since they are of dissimilar natures. Tests of deforming incrementally a certain point on the surface are similar in nature to, for example, the common tension and compression tests

in which the continuous loading causes a successive deformation. The plastic deformation of asperities caused by a certain load (history of deformation at the point of contact) affects the next deformation caused by the increase in this load. The Hertzian relation does not possess the same nature, since the calculated deformations at the different loads are not related to each other (in effect not for the same point of contact). When considering the contact models of rough surfaces such as those presented in Chapter 3, this problem becomes more evident since the computation of each deformation value on their load-deflection curves is based on the full statistical height of asperities (i.e., each deformation value is for a different point on the surface). Hence, unless the history of plastic deformation at the point of contact is theoretically considered, there will be no true comparison between the theoretical and experimental behaviours of the incremental deflection of rough surfaces.

Another potential source of such unexpected discrepancies between the Hertzian relation and these incremental deflection curves could be the friction. While the ideal relation assumes "perfectly frictionless materials" and "single point of contact" which could be approached under some loading conditions, the effects of friction are likely more important in incrementally loaded contacts. "Stick-slip" actions could be generated between the off-centre asperities (near and at the edge of contact area) and the tip surface as the indentation continues. This means that the larger the number of these asperities (smoother surface) the greater the effects of friction. However, such actions could significantly hinder further deformation of asperities and result in lower deflections than the ideal case.

Another important consideration with this type of experimental deflection, which has been pointed out before, is that the amount of elastic and plastic deformations depends primarily on the specific nature (e.g., height distribution, shape and geometry, local material properties, etc.) of the asperity group at the area of contact. The

behaviour of their successive deformations is also expected to depend on their nature. Hence, the linearity of deformation (surface compliance) could show a discrepancy between points on the same real rough surface due to the roughness and E^* variabilities. This effect may cause an additional difficulty in the comparison with the theoretical behaviour. However, to a great extent, it may explain why the point tested on the *St2* surface showed a smaller compliance value than that of the point tested on the *St3* surface, as given in Figure 6.11a, while it should show a larger value because of the high roughness of its surface. The same could be said about the two compliance values shown by the points tested on the *Al1* and *Al2* surfaces in Figure 6.11c. The difference between the compliance values of the two points tested on the *Co2* and *Co3* surfaces might be expected to be clearer than is shown in Figure 6.11b, as the difference between the roughnesses of these two surfaces is quite large.

The previous attempt discussed in Chapter 4 to relate the change in the surface compliance, caused by incremental loading on a single point on two surfaces of similar material, to the change in their roughness led to the empirical relation Equation 4.1 (which compared the ratio of deformation slopes to ratio of roughnesses) was based on two data points only. Here, using the nine different ratios that could be computed for each compliance and roughness, this relation was more rigorously checked. Figure 6.12 shows a plot of these nine data points which were denoted by the ● shape. For comparison, the old two data points were also plotted and denoted by the ✱ shape. The data point, denoted by the ✖ shape in this figure has a coordinate of (1, 1) and represents a nominal case of two separate surfaces of similar material and roughness resulting in nearly identical slopes of their deformation curve.

From all of these plotted data points, the linear relation between the two ratios (as observed previously) seems inapplicable here. Moreover, neither the zero roughness ratio nor the zero compliance ratio is practically obtainable, and the suggested

governing relation must, to some extent, reflect this fact. Therefore, a logarithmic curve fitting was found more suitable and all the data points were used to compute the best fit logarithmic trend line shown also in Figure 6.12 with its equation. From the correlation coefficient or residual R^2, it seems that this new relation is rather weak as a result of the relatively high deviations of some data points from the fitted curve. Recalling that some of the compliance values (such as of the *Al2* surface) given in Figure 6.11c were somewhat inconsistent with the roughness difference between the surfaces due to the variabilities of roughness and E^* on each of them, this equation of the best fit curve becomes more acceptable. If, however, the highly scattered data points are ignored, the equation tends to be closer to the form

$$R_s = \tfrac{1}{5}\ln(R_r) + 1 \tag{6.1}$$

where R_s and R_r are the ratios of deflection slopes (compliances) and roughnesses of the two surfaces, respectively. Specifically, leaving out the two data points near the coordinates (0.2, 0.5) and (0.4, 1.1) would enhance the correlation coefficient to 0.80. When also leaving out the data point (1, 1) which represents the ideal case, the resulting equation of the fitted curve would remain near to the above form (with coefficients: 0.22 and 1.06); the correlation coefficient in this case would nearly be the same, 0.77.

In regard to the effect of surface material, Figure 6.13 rearranges the load-surface deformation curves in terms of their roughness description used in the discussion. This effect is very apparent in each of the three groups, although with the distinctive roughness differences between the surfaces in comparison. To study such effect in more details, Table 6.4 compares the Hertzian deformation ratio between the different materials (based on a steel probe tip) with the compliance ratio between the different surfaces examined. The ratio of elasticity modulus of the materials is also shown in this table to provide another comparison with the case of linear deformation.

It seems obvious that the ratio of deformation (under the same contact load) on

two surfaces of identical roughness but dissimilar materials should nominally equal that of another two surfaces of the same dissimilar materials but of higher (or lower) identical roughness. This is fairly clear from the ratios of compliance between the steel and copper surfaces (which have the closest roughnesses to each other); a compliance ratio of about 0.9 was found between each pair of surfaces of nearly equal roughness. However, there was no correlation in the ratios of compliance between the steel and aluminium surfaces or between the copper and aluminium surfaces. The dispersion of these ratios is believed to be due to the relatively large roughness differences between each two surfaces involved, in addition to the odd compliance values caused by the variability of roughness on each of the surfaces. Given that the ratio for steel to copper pairs is not so far from unity, the similarity of individual ratios of steel and copper to aluminium gives more evidence for this interpretation.

However, these observed ratios of compliance showed no clear relation to the Hertzian E^* ratios; some of them are higher than the Hertzian ones, while the others are lower. This behaviour becomes obvious, when recalling that the observed load-deflection curves have not originally followed the Hertzian trends or showed a specific location with respect to them. The ratios of Hertzian deformation, computed in Table 6.4, sound very convincing between perfectly smooth surfaces, but when dealing with rough surfaces, this comparison turns out to be as difficult as encountered here. As stated before, the non-uniform distribution of asperities across a rough surface (anisotropy) and the E^* variability play an important role in affecting the amount of surface deformation. Therefore, it seems believable that compliance ratios of such surfaces could not be compared with the ideal deformation ratios of Hertz.

Although surface deflections measured here possessed a good linear relation with the increasing contact forces, the ratio of linear "elastic" deformations (the ratio of modulus of elasticity computed in Table 6.4) did not show a distinct relation to the

experimental ratios of surface compliance. Since the slope of the deformation trend of the asperity group in contact depends primarily on its random nature (material property, number and geometry of asperities), this slope becomes unpredictable if another group of asperities is tested. It is not easy to describe any systematic differences between the two ratios, as experienced here.

6.6 EFFECT OF SURFACE CONDITIONS

To study the effect of this important parameter on the repeatability of contact more rigorously than what has been discussed in Chapter 4, a new set of specimens was prepared. A total of eight specimens (square blocks the same sizes, 20x20x6 mm^3, as those of the previous set) were manufactured: four from mild steel and four from aluminium. As described in Section 6.2, the two square faces of each specimen were ground separately, with one of them given the required roughness quality for testing. The specimens were ground in this way in pairs in order to obtain nominally the same roughness on two of the same material. Two different roughness qualities were obtained on the four specimens of each material (each roughness on two specimens). During grinding these surfaces, an ordinary machining coolant was used in order to have on them the usual surface conditions of machining.

One specimen only of each pair was cleaned by hand using isopropanol and cleaning tissues. So, for each of the four roughness qualities obtained, there were two surfaces of different conditions; clean and unclean. The four clean specimens were then taken to the Form Talysurf instrument to measure the root-mean-square roughness parameter, R_q, of their prepared surfaces. At 0.8 mm cut-off length, multiple traces were taken across each surface and the average Rq value was computed and is listed in Table 6.5. For convenience, codenames were also assigned to each of these surfaces,

and will be used throughout this work.

The unclean surfaces were flicked gently by hand directly after machining to get rid of excess coolant drops, and left to dry in the machine shop environment. Later, a very thin layer of oil and dust was observed of different extent and location on each surface. Obviously, these surfaces were unsuitable for roughness measurement; they were assumed to be of the same roughness qualities as the clean ones. As shown in Table 6.5, they were given different codenames as their topographies include features unlike of those of the clean surfaces.

As with the previous set of specimens, each of these clean and unclean specimens was glued to a glass microscope slide, as shown in Figure 6.1, using Araldite epoxy adhesive. This is, as noted before, to avoid the effect of the interaction between the bottom surface of the specimen and the surface of the base plate of the test-rig.

6.6.1 Repeatability of Same-Point Contact

The effect of surface cleanliness was investigated on repeatability of contact at the same point on the surface using two contact tips of different sizes and the set of specimens described before. Three tests were performed on each surface with three different dead-weights, and twelve readings were collected for the probe vertical position in each test from each of two capacitive gauges. Hence, using each of the two contact tips, a total of twenty-four tests have been analysed in this investigation. Appendix C4 gives full details of their raw readings and preliminary plots. The scenario of testing and data analysis is similar to that described in Section 6.3. But in this group of tests, the initial dead-load adjusted on the specimens was much smaller; around 0.03 N.

The phenomenon of the relatively high magnitude of surface deformation at the first application of dead-weight was also inevitable in each of these tests, as shown in the initial plots illustrated in Appendix C4. Obviously, this phenomenon was already

discussed through the analysis of the different-point repeated contact tests given in Section 6.4, and will be analysed again in the next subsection with the consideration of the effect of surface conditions. The pattern of behaviour of deformations at the succeeding applications of dead-weight during each test was observed to be as discussed in Section 6.3. Again, the last six readings only from each test were considered in the analysis of these results.

Table 6.6 shows the statistics of these sets of six readings for all the twenty-four tests performed on the eight specimens using the contact tip of 5 mm diameter. The computation of each parameter shown in the first row of this table was done in the same manner described earlier for Table 6.2. The average resultant "elastic" deflections for the eight clean and unclean surfaces were plotted in Figure 6.14 for each material separately against the change in contact load. The consistent increase of these "elastic" deflections with the increase of contact load is in general very clear from these plots on both materials and at both surface conditions. This increase was continuous and almost linear on all the clean surfaces, unlike what was observed on some of the surfaces illustrated in Figure 6.2, which showed first an increase and, then, an unexpected decrease in these deflections in later tests (at different locations on the surface) with higher repeated loads, because of the roughness variations across the same surface.

Both the clean and unclean surfaces showed the consistent effect of surface roughness on these "elastic" deflections on both materials, as shown in Figure 6.14. This effect was more evident in this group of tests than in the previous one discussed in Section 6.3, which revealed that the roughness and E^* variabilities of surfaces might also inhibit this effect and could result in deflections on surfaces of higher roughness being lower than on those of lower roughness at any arbitrary change of contact loads. Furthermore, the effect of surface conditions on these deflections seems to be greater on the smooth surfaces than on the rough ones. The discrepancy between the plots of S/C

and *S1D* surfaces in Figure 6.14a is generally larger than that between the plots of *S2C* and *S2D* surfaces. Dust and machining debris particles and oil layers tend to reside in the large spaces between the higher asperities on the rough surface. Hence, they are expected to share the interaction with the asperities against the contact tip force and, probably, get supported by the asperities. On the smooth surface, these contaminants are expected to be less supported by the smaller asperities and more vulnerable to the interaction than these asperities since they are more exposed to the contact. However, this behaviour was not greatly confirmed between the plots of aluminium surfaces in Figure 6.14b, which suggest that the surface material is an affecting parameter here.

The unclean surfaces showed, in general, an increase in such deflections at nearly all the dead-loads used. The dust particles and oil layers remaining on these surfaces obviously add features to their topographies, of random shapes and distribution all over the surface, probably statistically similar to the surface asperities. This might cause the surface to function as if at a higher roughness than its actual one. Thus, it is reasonable to assume that such features added to the topography must have interfered in the contact interface and possibly caused the interaction to take place at the summits of these particles (in addition to the summits of the highest asperities), as exaggeratedly sketched in Figure 6.15a. The supporting asperities to such particles from below and the oil pockets between these particles must also be involved in such interactions. These supporting asperities (and probably others in the contact zone) must show a deformation as well as the highest ones. The support of such asperities does not prevent these dust and debris particles from showing magnitudes of deformation generally larger than the asperities themselves because of their greater freedom. In addition, movements and rotations of these particles are also expected to take place at the incidence of contact or change of load. The presence of oil could induce (or inhibit) such actions. So, the increase in the "elastic" deflections on the unclean surfaces is readily justified.

The complexity of the contact on "dirty" surfaces is easily appreciated from all these interacting effects. Factors more unsystematic than the surface roughness come into action at the contact interface. Clearly, the theoretical models of contact could not be employed in such situations. The surface topography is difficult to describe to the contact model, and the resultant effect of interaction is difficult to predict. Since perfect cleanliness is impossible (at least, under normal operating environments of QC checkpoints), the effect of surface conditions is however present, but usually considered to be insignificant.

Therefore, it is expected that the contact deflections on the unclean surfaces could show additional inconsistencies with the other parameters of contact. This was observed strongly from the "elastic" deflections of the *S1D* surface plotted in Figure 6.14a. Changing the contact load repeatedly with about 0.25 N at the same point on this surface gave a deflection magnitude of more than twice that expected, and even nearly double that seen at the next higher change of load. It was also observed on the *A2D* surface, Figure 6.14b, which showed a value of deflection at the second repeated change in load of about 50% higher than expected (compared to a linear trend of "elastic" deformations as nearly shown by most of the other surfaces).

On the other hand, the *S2D* surface revealed, in general, relatively low magnitudes of deformation compared to those revealed by the *S2C* surface, Figure 6.14a. Such behaviour is similar to some of the deformations trends illustrated in Figure 6.2. This suggests that the combination of surface asperities, dust particles, and oil layers do not always lead to an increase in this type of deflection. During each test, the first few applications of load on a particular point on the unclean surface might have created a combination of plastically deformed asperities and particles that was unable to show a further "elastic" deformation of magnitudes larger than those shown by the clean surface. The deformed dust particles might have reduced the gaps between the

asperities, as illustrated in Figure 6.15b, with the aid of the deformed asperities (and probably the oil, as well), which finally formed a set of strongly supported features of a low "elastic" deformation at the same repeated load. Another explanation to this behaviour is a basic one: the distribution of dust particles might have been very low at that particular point on the surface; hence, the *S2D* surface showed its second elastic deflection in the same regime of that of the *S2C* surface.

Table 6.7 gives results of the last six readings of the twenty-four same-point repeated contact tests, performed on the eight specimens used before, with a contact tip of 3 mm diameter. The average resultant deflections shown in the third column of this table are plotted in Figure 6.16 for each of the two materials. The consistent nearly linear increase of these "elastic" deflections with the repeated contact load is observed on all the surfaces. The effect of roughness is apparent, as well, at both conditions of surfaces. Similarly, the deflections between the two steel surfaces of the low roughness, *S1C* and *S1D* in Figure 6.16a, showed greater deviations than those between the other two steel surfaces of the high roughness, *S2C* and *S2D*. This higher effect of surface conditions on the smoother surfaces was not observed between the deflections of the aluminium surfaces in Figure 6.16b, which supports the previous observation from Figure 6.14b and shows that the E^* value has an additional effect beside the roughness.

The behaviour of these "elastic" deflections on the unclean surfaces is different from what was observed when the 5 mm contact tip was used. Excluding the *S1D* surface, all the other three surfaces showed in general smaller deflections than their rival clean ones, in particular, at the second and the third repeated changes of load. At the first change of load, they showed relatively larger "elastic" deflections than the clean surfaces. This suggests that the dust particles and oil layers might have obstructed such deflections with the small contact tip to a greater extent than with the large one, although the small tip interacts with fewer asperities and particles since its expected

area of contact is smaller at the same given contact load. The suggested idea of "narrow gaps between asperities and particles due to the first few cycles of load", which was discussed before and illustrated in Figure 6.15b, might also be valid here. Section 6.7 will be dedicated to a detailed comparison between the results observed with the two sizes of contact tip used.

To check the effect of E^* value in the presence of the effect of surface conditions, the "elastic" deformations caused by the two contact tips on the *S2C* and *S2D* surfaces were re-plotted in Figure 6.17 together with those of the *A1C* and *A1D* surfaces; the difference between the two roughness magnitudes (about 0.1 μm) is considered to be minor for the sake of comparison. In general, the effect of E^* value from Figure 6.17a is still evident on the unclean surfaces, although the clean ones failed to show it. The difference in roughness might have been more dominant than the difference between E^* values on these two clean surfaces. But using the 3 mm contact tip, the two clean surfaces showed the effect of E^* value as shown in Figure 6.17b, in spite of that difference in roughness. Similarly, the two unclean surfaces also revealed this effect. Thus, it seems from these results that the effect of material is more visible on the unclean surfaces (prepared in the way described before), and as the "elastic" deformations become generally greater.

With the previous set of specimens of the clean surfaces, the repeated "elastic" deflections revealed distinct quantitative relations with neither the surface roughness nor the contact modulus, E^*, yet the general effect of these two parameters was reasonably clear. The situation on the clean surfaces of this set of specimens is similar with both the sizes of contact tip used. In addition, the unclean surfaces revealed broadly such an effect, as expected, without clear systematic relations between these deflections. The additional randomly-distributed features on the rough surface could lead to additional variations in such deflections of a more unpredictable nature, as observed here.

In regard to the deviation of these repeated deflections, Table 6.6 and Table 6.7 give (in the fourth column of each) the standard deviation values observed in all the tests performed with the 5 mm and 3 mm contact tips, respectively. At both surface conditions, most of the surfaces showed the highest deviation at either the lowest two repeated changes in contact load. Few of them showed this at the highest repeated change in load which generally caused the highest "elastic" deflection on every surface. As believed from the previous similar results, the noise level of measurement during each test is a very significant parameter for the repeatability of such deflections. However, from Table 6.6, these highest deflections showed deviations between 2.7 and 8.6 nm on the clean surfaces, and between 5.2 and 12.5 nm on the unclean ones. And from Table 6.7, these deviations were between 2.9 and 5.4 nm on the clean surfaces, while on the unclean ones they were between 3.2 and 7.7 nm. Thus, it seems from these values that the surface cleanliness has some influences on such deviations at magnitudes of "elastic" deflection significantly above the noise floor. Although these differences are very minor compared to the deflections, they appear plausible since, as mentioned earlier, movements and rotations of the additional features on the rough surface might take place at the incidents of repeated change in load, which could cause additional variations. This is unlike the surface roughness and material parameters which showed no clear effects on these deviations within the available data.

Rotations of the contact tip revealed a consistent increase to either side of the test-rig with the repeated contact loads on some of the clean and unclean surfaces, Tables 6.6 and 6.7. Most of the surfaces that showed inconsistent variations of these rotations were unclean; namely, the *S1D* and *A2D* surfaces with the 5 mm tip, and the *S1D*, *S2D*, and *A1D* surfaces with the 3 mm tip. Hence, the surface conditions are also able to affect the tip rotations. This supports the previous conclusion that the increase of such rotations with the repeated loads is not always expected due to the random distribution

of surface asperities across the same clean surface. A consistent increase is even more unexpected to occur with the additional random distribution of dust particles and oil layers across the same unclean surface. Such conditions on the rough surface are likely to force the contact tip to comply with the resultant arbitrary deformed asperities and particles and to show a more random rotation across the unclean surface than across the clean one. As a result, all the surfaces showed no clear effects of both the surface roughness and its material on these tip rotations.

The average compliance of contact was also observed to decrease with the increase of the change in repeated load on all the surfaces when the 5 mm contact tip was used, as shown in Table 6.6. But with the 3 mm contact tip, some of the surfaces, such as the *S1C* and *S1D* in Table 6.7, showed similar compliance values at each change in repeated load. This is a genuine behaviour because of, as noted earlier, the linear relation between loads and deflections within the elastic regime. Only the *A2C* surface revealed, unexpectedly, an increasing trend of compliance with the changes in repeated load. The decreasing magnitude of compliance with the increase of repeated load was, however, the common behaviour. Although a larger repeated load usually caused larger elastic deflections, when normalised to this load, such deflections could be relatively smaller. It seems reasonable that surface asperities and dust particles deform elastically at the low repeated loads in greater magnitudes per unit load than at the high ones, since they are less supported by each other and the spaces between them are higher and, hence, they are able to show a more elastic deformation.

The previous set of tests performed on clean surfaces have shown no clear correlation of either the surface roughness or its E^* value to these elastic compliance values. This is unlike the observation on the current set of clean surfaces. Although, the material effect is unclear when comparing the compliance values in Table 6.6 of *S2C* surface with those of *A1C* surface, it is apparent in Table 6.7 at all the repeated

loads used. The roughness effect was revealed by these surfaces with each contact tip and at nearly every repeated load. On the unclean surfaces, the behaviour was rather different. The effect of E^* value is evident between the *S2D* and *A1D* surfaces with both contact tips. On the aluminium surfaces, the effect of roughness on the compliance values was more obvious than on the mild steel surfaces, which could be attributed to the material effect involved. When the 3 mm tip was used, the *S2D* surface was in general more compliant than the *S1D* surface, but not when the 5 mm tip was used. This strong effect of contact tip size will be discussed in Section 6.7.

6.6.2 Repeatability of Different-Point Contact

The repeatability of contact at different points on the rough surface was examined further in the presence of the additional effect of surface cleanliness on the same surfaces used in the previous set of tests. The two contact tips and the three contact loads discussed in Subsection 6.6.1 were used again to generate the same number of tests, during each of which twelve readings were collected for the probe vertical position from each of the two capacitive gauges. A total of forty-eight tests have been analysed in this investigation and Appendix C5 gives full details of their raw readings and preliminary plots. The procedure of testing and data analysis is similar to that described previously in Section 6.4. The initial dead-load adjusted on each specimen prior to every test was also around 0.03 N.

Tables 6.8 and 6.9 provide a summary of results observed when the 5 mm and 3 mm contact tips were used, respectively. Figures 6.18 and 6.19 illustrate the behaviours of the average different-point repeated deflection with each of these two tips. The consistent approximately linear behaviour of this deflection with the contact load on the clean surfaces could also be seen in all the plots in these two figures. The effect of surface roughness on the behaviour is very obvious, as well. The unclean

surfaces also revealed a fairly consistent increase in such deflections with loads, but with some odd values: the first deflection values on the *S1D* and *A1D* surfaces in Figure 6.18 and the second one on the *A2D* surface in Figure 6.19b. Such inconsistent (higher than expected) average deflections are believed to be due to the random nature of the dust particles and oil layers on the unclean rough surface, as well as to its asperities. The combination of dirt and roughness is expected to increase the statistical variation and, over the number of tests, may occasionally result in an average deflection inconsistent with the general trend. However, the roughness effect is also very clear on these unclean surfaces for both materials and with both contact tips.

Regarding the material effect, as was done before, the average deflections of the surfaces of the closest roughness from each material are re-plotted together in Figure 6.20 for both contact tips. The *S2C* surface revealed somewhat higher magnitudes of deflection than the *A1C* surface with the 5 mm tip, as shown in Figure 6.20a, while these two surfaces showed nearly the same magnitudes with the 3 mm tip, as shown in Figure 6.20b. No effect of E^* value is apparent from these deflections, possibly because of the roughness difference (the *S2C* surface is about 0.1 μm rougher than the *A1C* surface) which might have been more effective on these deflections than the E^* value. Unexpectedly, the other two unclean surfaces also did not reveal this effect with either contact tip, although the same-point repeated deflections observed on them clearly showed it; as discussed in the previous subsection. This suggests that when deforming a new set of asperities the surface cleanliness can be a more dominant effect than the E^* effect; the surface roughness may enhance this effect by supporting contamination particles. The effect of material could become more important when a further "elastic" deformation is performed on that plastically deformed set of asperities and dust particles of the surface.

To explain this behaviour in more detail, Figure 6.21 illustrates typical output

voltages acquired from the two capacitive gauges during testing a point on the *S2C* and *S2D* surfaces with the 0.35 N dead-weight. The behaviour is different on each of these surfaces: in addition to the inevitable large difference in the deflection, there was also a large difference in the contact tip rotation (the difference between the front and rear voltage levels after loading compared to that before loading is much smaller on the *S2C* surface than on the *S2D* surface). Another more important discrepancy is shown inside the dotted circle in Figure 6.21b. These disturbances in the vertical motion of the contact tip during the deformation of surface asperities were not detected on the clean surface as shown in Figure 6.21a. Such behaviour was encountered on most of the points tested on the unclean surfaces, while it was not seen on any of the clean ones. It is believed to be due principally to the presence of the dust particles and oil layers which might have suffered rearrangement (rotation, displacement, fall, etc.) with respect to each other, throughout the indentation of the unclean surface. Such actions could possibly cause resistance to the tip force which, in turn, could create these disturbances during the vertical movement of the tip. The geometry of asperities (or, say, the surface roughness) plays an important role in aiding these actions: the higher the roughness the greater these actions. It is believed that most of the difference in deflections between the clean and unclean surfaces is because of the rearrangements of the dust particles and oil layers. Hence, the E^* value seems to be of a minor effect in such circumstances.

As observed in Section 6.4, the standard deviation of this type of repeated deflections also showed, in general, an increasing trend with both the contact load and surface roughness on all the clean surfaces and with both contact tips, as shown in Figures 6.22 and 6.23. This increase is basically related to the consistent increase in deflections with these two parameters. It is also observed to be nearly linear with the load on the smoothest surfaces, namely, the *S1C*, *S2C*, and *A1C* surfaces which are all of roughnesses below 0.6 μm. On the roughest surface, *A2C*, the behaviour was

inconsistent using both contact tips: standard deviations increased and, then, decreased at the higher load used. This supports the previous belief that the increasing trend of such roughness-related deviations with the contact load is not always continuous on every surface roughness. At high loads and deflections, the effect of roughness variation across the surface becomes less significant.

On the unclean surfaces, the behaviour of deflections repeatability against the contact load is rather similar, with some inconsistencies. All these surfaces showed, as expectations, higher standard deviations than the clean ones, except the *A1D* surface which revealed lower deviations than the *A1C* surface with the 3 mm tip, as shown in Figure 6.23b. The increasing trend of these deviations with the contact load was also shown by all except the *A2D* surface which gave a decreasing trend, as shown in the same figure. Moreover, the rougher unclean surfaces generally had higher standard deviations than the smoother ones, except the *S1D* and *S2D* surfaces which revealed the opposite behaviour, as shown in Figure 6.23a. Thus, more consistent behaviour was observed with the 5 mm tip than with the 3 mm one, which suggests that the tip size has an important effect on the repeatability of deflections on these unclean surfaces.

Figure 6.24 combines the trends of standard deviation of the repeated deflections of the *S2C*, *S2D*, *A1C*, and *A1D* surfaces in order to examine the material effect on these deviations. On both of the clean surfaces, these deviations were of similar magnitudes at every contact load and tip used, with no significant variations due to the change of E^* value. This confirms the previous result encountered in Section 6.4, which stated that the surfaces of dissimilar materials having the same regime of roughness and tested under the same repeated load, are likely to show similar standard deviations of deflections. On the other two unclean surfaces, it seems that the effect of the roughness difference (assumed to be minor) became more dominant; deflections of the rougher surface, *S2D*, showed much greater standard deviations than those of the *A1D* surface.

However, the E^* value is not expected to have much effect on the repeatability of deflections on the unclean surfaces. The available data for these deflections failed to show such an effect, as discussed earlier.

Recalling the results given in Tables 6.8 and 6.9, the average rotations of the contact tip about the x-axis of the test-rig also tended to increase, in either of the two directions (front or rear), with the contact loads on all of the eight surfaces. This is due to the increase in the contact tip interaction with those surface asperities, as well as with dust particles and oil layers (on the unclean surfaces), within the contact area. This increase was greater and more consistent with the 3 mm tip than with the 5 mm tip: an observation which proves the effect of tip size on these rotations, as will be discussed later. As noticed earlier, the roughest clean surfaces also caused generally greater variations in the tip rotation than the smoothest ones. For a clearer view, Figure 6.25 illustrates the raw values of tip rotations computed for all the repeated contacts made on all the surfaces with both contact tips. The roughness effect is very obvious from this figure: the *S2C* and *A2C* surfaces showed a higher scatter of the repeated rotations, especially of the 5 mm tip, than the *S1C* and *A1C* surfaces, respectively. It is also very clear that the unclean surfaces caused larger average rotations to the contact tips used, and revealed larger scatters of these rotations. The roughness effect is also evident, broadly, from the average tip rotation values of these surfaces, as seen on Figure 6.25. The statistics of asperity heights are non-stationary (local areas have different statistics) and become more so as roughness increases. Rougher surfaces will have more extreme peaks, but likely to be at larger average spacing, so it is more likely that probe tip sees significant net side-thrust as it comes into contact.

The earlier tests of this type suggested that such rotations can be affected by the change of E^* value, since the surface from the softer material could produce a larger interaction with the contact tip. The current tests show no evidence for this when

comparing the average rotations of the *S2C* surface with those of the *A1C* surface. But in Figure 6.25, the *A1C* surface shows a slightly larger scatter of rotations than the *S2C* surface at nearly every load used. This could be a genuine effect from the E^* difference between them, as the effect of roughness difference would cause the opposite result if it was the dominant effect. However, the unclean versions of these two surfaces showed the roughness difference effect on the scatter of rotations instead of the material effect; the *S2D* surface caused the highest scatter since it has rather higher roughness than the *A1D* surface.

In regard to the computed values of the average compliance of contact provided in Tables 6.8 and 6.9, the clean surfaces showed no particular trend with the contact loads – similar to those computed in Table 6.3. This seems to be a normal behaviour since the deflections were observed to be nearly linear with the repeated loads applied and, hence, similar magnitudes of compliance should nominally result. The variability of compliance is basically due to the variabilities of surface roughness and E^* across the same surface. In general, the roughest clean surfaces also gave greater compliance values than the smoothest ones, simply, because of their higher deflections. With both contact tips, the average compliance decreased as the repeated loads increased on all the unclean surfaces as observed, also, on the average "elastic" compliance of the same-point repeated contact tests. It again suggests that the surface asperities, dust particles, and oil layers rearrange themselves and deform at the low repeated loads in greater magnitudes per unit load than at the high ones. The roughest unclean surfaces also revealed larger "decreasing" compliance values than the smoothest ones, once more due to their larger deflections. Similarly to the deflections, the effect of E^* value on the surface compliance was not seen between the *S2C*, *A1C*, *S2D* and *A1D* surfaces; the slight difference in roughness might have been of a more serious effect, instead.

To conclude in general terms about this parameter (i.e., the surface conditions)

based on the available data, machining debris, dust particles and oil layers remaining randomly across the rough surface were observed to be of decisive effect on the contact deflections, in addition to the roughness effect. Same- and different-point repeated deflections on such surfaces were also observed to be affected by the presence of these features and showed significantly greater deviations. Likewise, the tilts of contact tips used appeared to be generally larger in magnitude and scatter on these unclean surfaces than on the clean ones. The effect of surface material on the deflections of unclean surfaces seemed to be significant when deforming the same point (that already has been deformed using the same load), but it was trivial when deforming a new point on the unclean surface.

The random geometry and distribution (like those of the surface asperities) of these additional features in addition to their freedom could simply lead to further arbitrary behaviour from all the other parameters involved in the contact interaction. Anisotropy of the clean rough surface seemed to be the main source of variabilities which are largely stimulated by the presence of such features. Although the above qualitative analysis of the different behaviours is plausible, the quantitative assessment obviously seems greatly complicated. Even when assuming that a rigorous evaluation of the topography of an unclean surface has been achieved in a way or another, the actions of rearrangement and deformation of these features (at the onset of contact) and their interaction with the deforming asperities remain problematic behaviours. Thus, predictions provided by the mathematical contact analyses become worthless under those unsystematic severe effects from such a parameter.

6.7 EFFECT OF CONTACT TIP SIZE

The effect of this parameter on the repeatability of contact has been briefly examined

using the old test-rig on only one rough mild steel surface. As discussed earlier in Subsection 4.4.4, same-point repeated contact tests have been performed with each of the two contact probes (of 5 mm and 3 mm tip diameters) and showed that the bigger probe tip resulted in a larger uncertainty of the "elastic" deflections. This observation will be checked within the results of the experimental investigation discussed in the previous section which will be rearranged in this section to also emphasize such effect on both types of repeatability. Obviously, the involvement of effects from other parameters, such as the E^* value, will be highlighted as before.

6.7.1 Repeatability of Same-Point Contact

Figures 6.26 and 6.27 combine the average "elastic" deflections plots made by the two contact tips together for the mild steel and aluminium surfaces, respectively. On the clean surfaces, the effect of the smaller tip is very obvious; this tip caused generally higher "elastic" deflections than the larger tip, specially, at the highest load used. This is simply an ordinary behaviour from what the Hertzian analysis revealed. To have a more detailed comparison, the ratio of approach distance (surface deflection) between the 5 mm and 3 mm contact probes, at similar loads and on similar materials, is about 0.84 according to the Hertz theory. Assuming linear trends of "elastic" deflection with loads, the *S1C* and *S2C* surfaces showed ratios of slopes very close to the Hertzian ratio; 0.88 and 0.87, respectively. On the contrary, the *A1C* and *A2C* surfaces showed ratios of slopes of 0.59 and 0.53, respectively. This could mainly be referred to the inevitable effect of roughness variability across the same surface.

The effect of roughness on these deflections was nearly similar with each probe tip. In other words, the ratio of slopes of the deflection trends between the *S1C* and *S2C* surfaces with the 5 mm tip is about 0.64, while it is around 0.63 with the 3 mm tip. Between the aluminium clean surfaces, this ratio is about 0.53 and 0.48 with the 5 mm

and 3 mm tips, respectively. This seems a reasonable result since it agrees with the previous empirical relation (suggested between the roughness and compliance ratios) which was discussed earlier in Section 6.5. Although the method of applying the dead-loads and the resulting deflections in these two types of test are different, but the same thinking is still applicable here: the same roughness ratio could generally give the same compliance ratio, regardless of the tip size. Unexpectedly, when substituting the roughness ratio between the two mild steel surfaces (i.e., about 0.21) into Equation 6.1 before, it would give a compliance ratio of around 0.68 which is near to the observed values. Substituting the roughness ratio between the two aluminium surfaces (i.e., 0.29) would give a compliance ratio of 0.75 which is far from the observed values.

On the unclean surfaces, refer to Figures 6.26b and 6.27b, the effect of tip size is also clear. The "elastic" deflections made with the 5 mm tip on these surfaces are somewhat higher than those made with the 3 mm tip (unlike the behaviour observed on the clean ones); namely, the first and second deflections observed on the *S1D* and *A2D* surfaces and the first deflection on the *S2D* surface. This seems a reasonable behaviour as the interaction of the 5 mm tip with the dust particles and oil layers is greater. Additionally, it seems plausible that the small tip tends to stimulate the actions of rearrangement of these features in more proportions than the large tip. In other words, the displacements (skidding away from the tip centre) of these loose features as the small tip gets its initial contact with them could be greater due to its high inclination. Then, the final settled contact (after the first cycle of load that causes further of these actions and the major plastic deformations) on the unclean surface could be not only on a smaller area of contact but on a more reduced amount of such features. On the other hand, this behaviour is not always true as the random nature of surface asperities and these features also plays a dominant effect in controlling the amount of deflections. The later higher loads on all the three surfaces mentioned above with the 5 mm tip caused

lower deflections than with the 3 mm tip, which reflect the normal behaviour.

The effect of roughness on the repeated "elastic" deflections was clearer on the aluminium unclean surfaces with each probe tip than on the mild steel ones. The *A2D* surface showed higher deflections with both tips than the *A1D* surface, while the *S2D* surface failed to reveal a clear consistent difference in deflections from the *S1D* surface with both tips. This might suggest that the E^* effect could have supported the roughness effect to show greater deflections on the roughest aluminium surfaces. Such inconsistent behaviour could also be referred, yet again, to the greater arbitrary distribution of the dust particles and oil layers than that of the surface asperities which might have been of larger effects than the roughness.

Figure 6.28 shows another attempt to highlight the effect of E^* value on these deflections of the clean and unclean surfaces: combines the deflection plots of *S2C*, *S2D*, *A1C*, and *A1D* surfaces (which have the closest roughnesses) under both contact tips. On the clean ones, the deflections caused by the 3 mm tip revealed this effect, as shown in Figure 6.28a, but it was not seen on those made with the other tip. The greater interaction of the larger tip with the surface asperities (of the slightly rougher surface, *S2C*) could have been of a more decisive effect than the E^* effect. In Figure 6.28b, both unclean surfaces also showed, in general, such an effect using both the contact tips. So, the elasticity of surface material is broadly of an effect that depends largely on the other parameters involved in the interaction.

In regard to the effect of probe tip size on the deviations of these repeated deflections given in Tables 6.6 and 6.7, such deflections generally showed lesser scatters with the 3 mm tip – like what was observed in those similar tests performed on the old test-rig. To illustrate this fact clearly, in all the tests carried out with the 3 mm tip the deviations of the last six repeated deflections were below the 10 nm level, while in eight of the twenty-four tests performed with the 5 mm tip they were either very near

to or above this level. To a large extent, this could be an evidence of such an effect and could greatly support the thinking that the bigger the contact area (bigger tip size) the larger the interaction with the asperities (and those additional features remained); hence, greater deviations are expected.

On the other hand, small contact tips are expected to be more vulnerable to rotations during their interaction with all these features than large ones. This is since the small area of contact provides a sort of lower stability (lesser number of contact points) to the tip against tilting than the large area which offers less chances for the tip to balance itself on several highest features on the surface. From the average rotation values given Tables 6.6 and 6.7, such an effect of tip size is fairly clear. Four surfaces (namely, *S2C*, *A1D*, *A2C*, and *A2D*) caused, in general, higher rotations to the 3 mm tip than to the 5 mm tip at every change in load made. The *S2D* and *A1C* surfaces also caused higher rotations to the 3 mm tip when the 0.15 N dead-weight was repeatedly applied on them. However, the greater interaction of the large tip and the random distribution of features on the surface are the limiting factors of this effect. The other higher magnitudes of rotation observed with the 5 mm tip could greatly be because of some influences from these two important factors.

When comparing the average compliance values computed in Table 6.6 with those computed in Table 6.7, the effect of contact tip size would not be consistent. The average "elastic" deflections observed with the 3 mm tip were generally higher, but the compliances shown by almost all the surfaces with the 5 mm tip were surprisingly greater when the lowest dead-load was applied. This is simply due to that the values of the resultant change in contact load were lower when the 5 mm tip was used; hence, the corresponding compliances appeared greater. When the highest dead-load was applied, the difference in these values between the two tips became less important against the high deflections caused by the 3 mm tip; hence, all the surfaces appeared to be, as

expectations, less compliant under the 5 mm tip than under the 3 mm tip. Obviously, when normalising all the deflections to the magnitudes of the used dead-weights instead of those of the measured changes in load, all the resulting compliances at each tip would be in such consistency.

6.7.2 Repeatability of Different-Point Contact

The average values of the repeated deflections at different points on the surface given in Tables 6.8 and 6.9 are re-plotted for each material as shown in Figures 6.29 and 6.30 in order to examine the effect of contact tip size on these deflections. On the clean surfaces, the high deflections observed by the 3 mm tip (compared to those made with the 5 mm tip) are broadly clear, specially, on the *S2C* surface. At the first and second changes in load on all the other three surfaces, higher deflections were also observed with the 3 mm tip. Thus, most of the tests performed on the clean surfaces revealed the normal effect of tip size at the contact interaction. But, the inconsistent behaviours shown by the *A2C* surface at first and second loads and by the *S1C* and *A1C* surfaces at the third load (gave larger deflections with the 5 mm tip) could be, as believed, due to the roughness and E^* variations across the same surface. The large interactions of the large tip could create more possibilities to deal with asperity groups that are able to deform in greater proportions than others, which could significantly increase the average magnitude of deflections at any arbitrary repeated load. Obviously, the opposite situation could be true, as well, with the small tip. The systematic effect of tip size on these deflections was not observed because of such inconsistencies.

It is also very obvious from Figures 6.29a and 6.30a that the rougher surface from each material showed higher deflections than the other surface with both tips: the usual effect of roughness at the contact interaction. This effect is not similar at each contact tip and on both materials; yet again, because of those dominant effects from the

roughness variability across the surface on such type of deflections.

Figures 6.29b and 6.30b show somewhat more inconsistent behaviours on the unclean surfaces, but the general effect of tip size is also apparent. Eight of the twelve tests performed on these surfaces resulted in higher average deflections with the 3 mm tip than those resulted from the rival tests performed with the 5 mm tip. Such behaviour of these deflections is the opposite of that of the average "elastic" deflections shown by these surfaces (discussed in the previous subsection). The idea of "the large interactions of the big tip with asperities, dust, and oil and the less vulnerability of the small tip to these features" is still plausible here and it might justify the inconsistent behaviours resulted from the other four tests. This is in addition to the greater variability of these features across the unclean surface than across the clean one. The effect of roughness is also very clear on these surfaces since the two rougher surfaces revealed higher deflections than the other two surfaces with both tips and at all the loads. This effect is also of different proportions at each contact tip and on both materials.

Figure 6.31 rearranges the plots of average deflections of *S2C*, *S2D*, *A1C*, and *A1D* surfaces to examine the effect of E^* value under each contact tip. From this figure and as discussed in Subsection 6.6.2, all these surfaces did not reflect this effect. The two steel surfaces (which are slightly rougher than the aluminium ones) showed, in general, the highest deflections with both tips. This small difference in roughness is believed to be responsible for such behaviour. So, it seems that at the process of deforming a new set of features on these unclean surfaces, the slight extra roughness on the steel surfaces could be more effective when combined with the contamination particles effect than the extra elasticity of the aluminium surfaces.

Likewise, the standard deviations of these repeated deflections are re-plotted as shown in Figures 6.32 and 6.33 to check the effect of contact tip size on them for each surface condition and material. In general, the repeatability of such deflections on the

clean surfaces was observed to be better with the 3 mm tip than with the 5 mm one, although the standard deviations were broadly higher on the roughest surfaces than on the others. Figures 6.32a and 6.33a show that in eight of the twelve tests performed with the 3 mm tip, the standard deviations are lower than those resulted from the rival tests carried out with the 5 mm tip. Such behaviour could be attributed to the large interaction of the largest tip with the surface asperities, which could significantly increase the variation of deflections. The 3 mm contact tip also revealed the same fact on the unclean surfaces as shown in Figures 6.32b and 6.33b. It also caused deflections of low repeatability in eight tests compared to those made with the 5 mm tip in rival tests. The roughness effect on this repeatability was nearly obvious at each tip on these surfaces with inevitable more inconsistencies than on the clean surfaces.

From the previous observations and from Figure 6.34, the change of E^* value appeared to be of a clear effect at neither the two contact tips. This seems a plausible result since the deviation of deflections at different locations on the rough surface is mostly associated with the roughness variability across it. The slightly rougher steel surfaces showed, in general, the usual effect of roughness at both tips and gave repeated deflections of higher deviations than of those given by the aluminium ones. It worth pointing out once more here that, regardless of the tip size and surface material used, the increase of the standard deviation of repeated deflections with load is believed to be not absolute, especially at high loads where the effect of roughness (and probably the surface cleanliness) becomes less significant. This might justify the decrease of slope of the deviation trends when the highest load was repeatedly used on some surfaces from both materials and with both contact tips.

Comparing the rotations caused to the two contact tips during their interaction with all the different tested surfaces, the average values of the repeated rotations given Tables 6.8 and 6.9 show a clue to the tip size effect. As noted in the previous

subsection, the increase in the average rotation with the contact load was more consistent with the 3 mm tip than with the 5 mm tip. In addition, this average was, in general, greater in magnitude when the 3 mm tip was used; especially at the highest two changes in load on all the surfaces. This is also evident from the plots of the repeated rotations themselves shown in Figure 6.25: the departure of points from the zero line is more visible in Figure 6.25b than in Figure 6.25a. Additionally, the scatter of these data points is rather less in the former figure than in the latter one. This could be referred to the lower stability of the small tip on surface (less interaction with the surface features as a result of the smaller contact area) than the large tip.

Since the repeated deflections caused by the 3 mm tip are generally higher than those caused by the 5 mm tip, it is obviously expected that the compliance of surfaces will also be higher. Referring again to Tables 6.8 and 6.9, the average compliance values reflected this behaviour in most of the tests. Some inconsistent average deflections simply gave compliances inconsistent with such behaviour. As stated before, the roughness and E^* variabilities across the same surface and the large interactions of the large tip are believed to be the major cause of these inconsistencies.

To conclude in general terms from the present data, the size of contact tip was observed to have a considerable effect at the contact interaction. Whether testing a new or an already deformed group of asperities on the surface, the resulting deflections tended to reveal this effect in a consistent manner. Moreover, the repeatability of each of these two types of deflections was also noticed to be significantly influenced. The small contact tip seemed of a less stability against tilts than the large tip due to the smaller area of contact. This effect showed a variation with respect to the two crucial parameters involved (i.e., the roughness and cleanliness variabilities across the same surface) and revealed some inconsistent results. Therefore, it had no clear systematic relationship to all of the above factors concerned.

	Codename	Material	Average R_q (μm)
Specimens with **roughest** surfaces	*St1*	Mild steel	1.374
	Co1	Copper	1.324
	Al1	Aluminium	1.884
Specimens with **rougher** surfaces	*St2*	Mild steel	0.589
	Co2	Copper	0.497
	Al2	Aluminium	0.705
Specimens with **rough** surfaces	*St3*	Mild steel	0.339
	Co3	Copper	0.240
	Al3	Aluminium	0.117

Table 6.1 List of specimens used in the same-point and different-point repeated contact and the load-surface compliance tests.

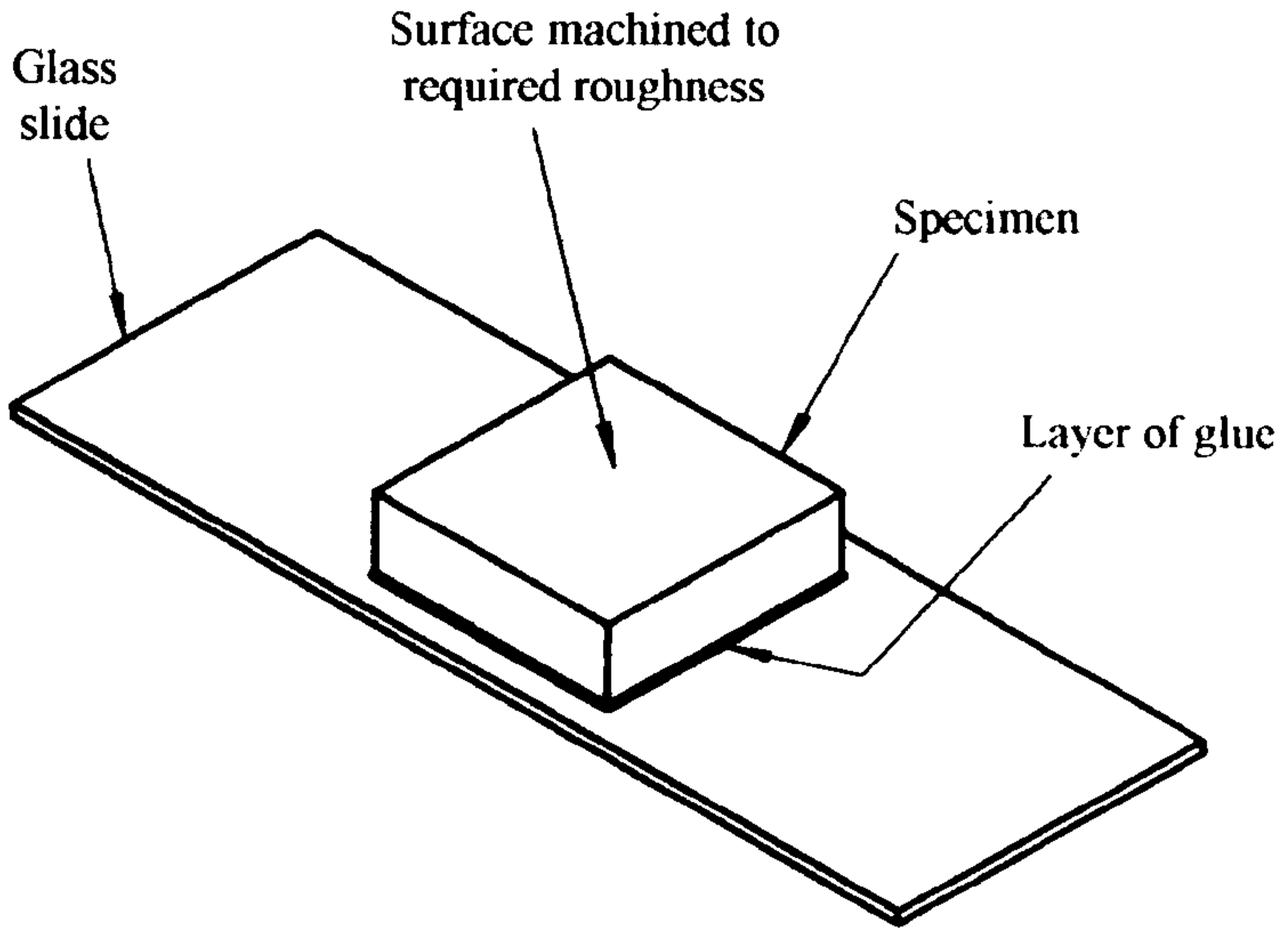


Figure 6.1 Schematic representation of mounting for specimens used to reduce bottom surface effects.

	*Dead-weight applied / **Change in contact load*** (N)	Average of repeated "elastic" deflections (nm)	95% confidence scatter range of repeated deflections (nm)	Average of repeated rotations (tilts) of contact tip (μradian)	Average compliance of contact (nm/N)
St1	*0.35 / **0.26***	79.1	± 17.0	0.3	309.8
	*0.55 / **0.44***	141.3	± 12.0	1.7	322.9
	*0.75 / **0.65***	161.6	± 18.8	1.4	249.8
	*0.95 / **0.93***	308.9	± 19.8	4.4	331.9
	*1.15 / **1.11***	357.7	± 10.6	4.5	323.3
St2	*0.35 / **0.20***	70.7	± 10.2	0.4	355.6
	*0.55 / **0.34***	116.5	± 12.2	2.5	342.0
	*0.75 / **0.56***	122.9	± 4.4	3.3	221.3
	*0.95 / **0.76***	230.3	± 3.4	3.3	303.5
	*1.15 / **1.00***	325.8	± 9.8	4.2	324.4
St3	*0.35 / **0.29***	248.6	± 20.6	- 1.7	851.3
	*0.55 / **0.49***	316.6	± 12.4	3.7	648.1
	*0.75 / **0.71***	299.6	± 7.8	4.5	423.9
	*0.95 / **0.87***	280.9	± 8.6	4.7	322.6
	*1.15 / **1.12***	355.3	± 8.6	2.7	317.3
Co1	*0.35 / **0.22***	148.6	± 9.6	- 0.9	662.6
	*0.55 / **0.41***	261.0	± 16.2	- 2.9	632.6
	*0.75 / **0.67***	412.3	± 8.2	- 0.9	616.9
	*0.95 / **0.84***	521.0	± 5.0	- 0.3	618.0
	*1.15 / **1.06***	637.6	± 6.0	7.6	598.8
Co2	*0.35 / **0.21***	61.9	± 15.2	0.8	300.8
	*0.55 / **0.37***	114.5	± 10.8	0.8	309.8
	*0.75 / **0.58***	170.9	± 15.2	3.3	294.9
	*0.95 / **0.78***	193.1	± 4.2	3.1	248.2
	*1.15 / **1.02***	363.6	± 9.4	5.1	354.8
Co3	*0.35 / **0.30***	128.6	± 7.2	0.6	431.2
	*0.55 / **0.48***	211.0	± 11.6	- 0.1	435.8
	*0.75 / **0.70***	290.8	± 11.4	0.3	415.9
	*0.95 / **0.91***	232.2	± 15.6	1.5	256.4
	*1.15 / **1.12***	280.2	± 10.0	3.7	250.2
Al1	*0.35 / **0.26***	209.5	± 17.6	- 0.2	792.1
	*0.55 / **0.46***	296.7	± 11.2	- 1.2	644.3
	*0.75 / **0.67***	484.8	± 14.6	- 1.2	721.5
	*0.95 / **0.89***	681.5	± 14.4	- 5.0	769.9
	*1.15 / **1.05***	665.2	± 9.4	- 2.5	631.3
Al2	*0.35 / **0.19***	176.9	± 24.8	- 0.4	924.4
	*0.55 / **0.34***	348.5	± 9.2	- 1.8	1014.7
	*0.75 / **0.54***	496.7	± 9.8	0.0	919.7
	*0.95 / **0.77***	820.9	± 13.2	0.7	1059.6
	*1.15 / **1.01***	1158.5	± 16.0	- 4.9	1145.6
Al3	*0.35 / **0.22***	74.0	± 10.0	1.4	341.8
	*0.55 / **0.37***	134.7	± 16.8	0.9	361.5
	*0.75 / **0.60***	227.5	± 14.8	1.9	381.9
	*0.95 / **0.80***	296.5	± 12.0	3.3	372.8
	*1.15 / **0.97***	336.8	± 12.2	2.5	348.4

Table 6.2 Results of the last six readings of the same-point repeated contact tests on steel, copper, and aluminium specimens of surfaces of different roughness.

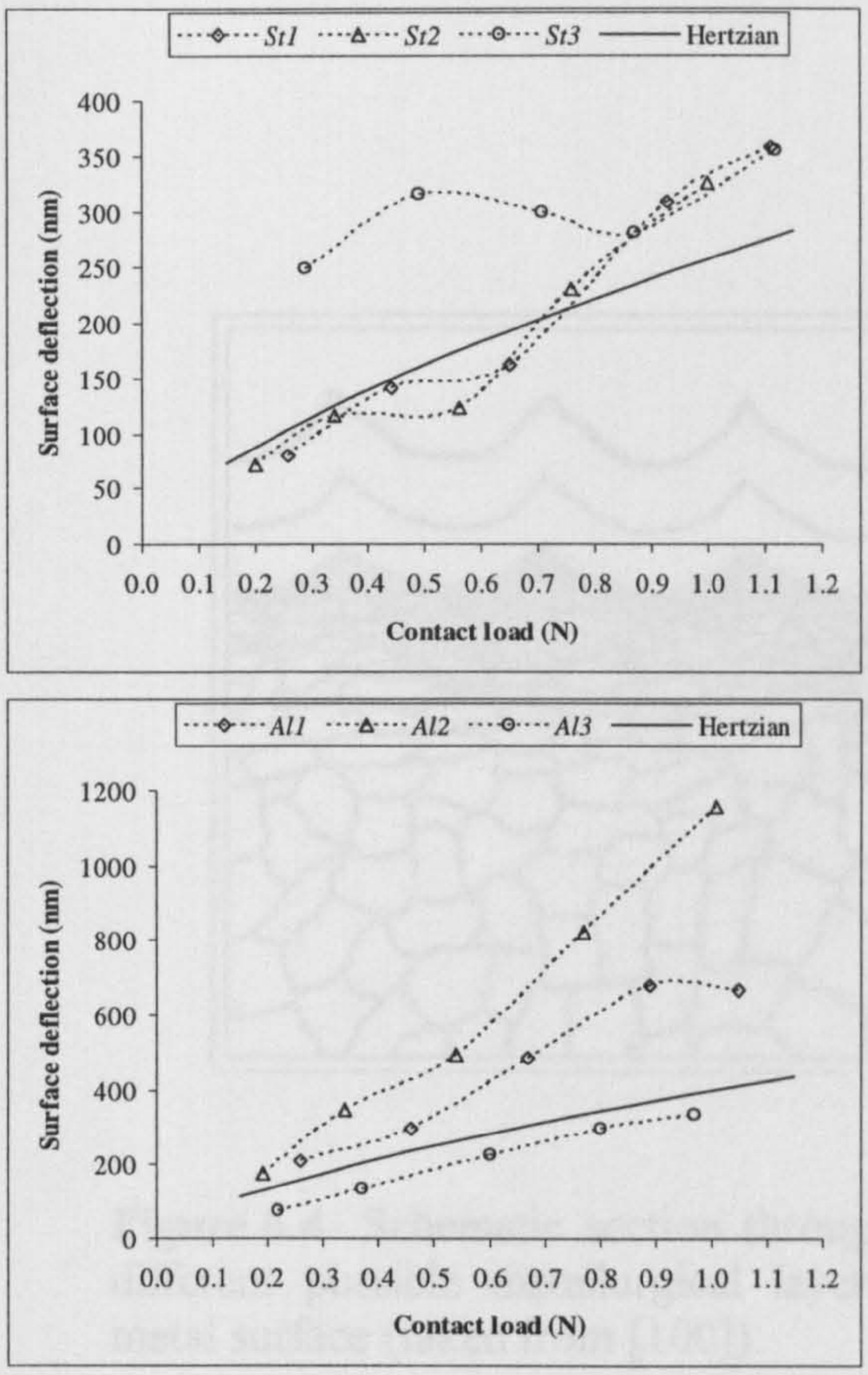


Figure 6.2 Plots of the average "elastic" deflections of same-point repeated contact tests on (a) steel, (b) copper, and (c) aluminium surfaces of different roughness.

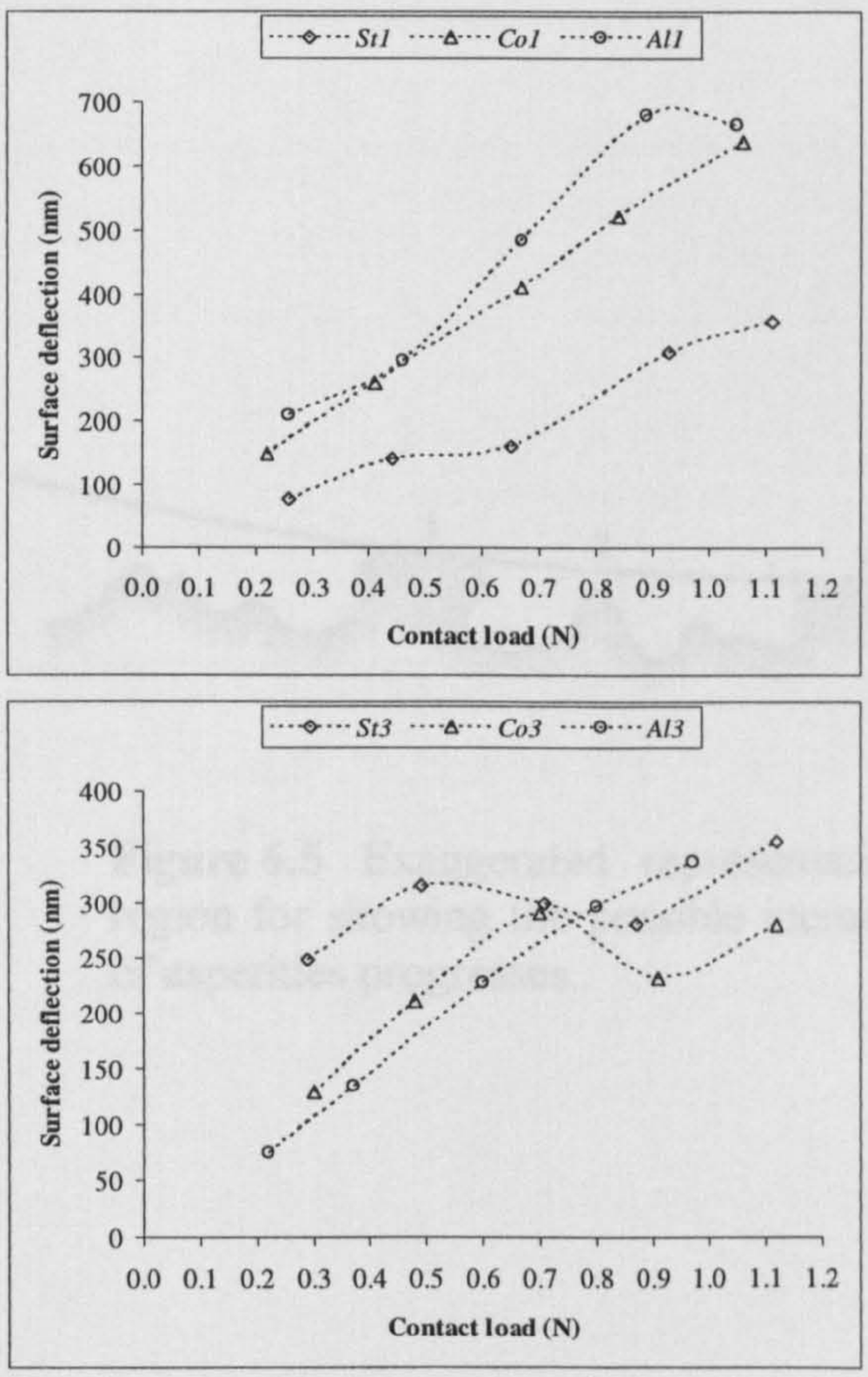


Figure 6.3 Comparison of the average "elastic" deflections of same-point repeated contact tests on the three specimens' materials of (a) roughest, (b) rougher, and (c) rough surfaces.

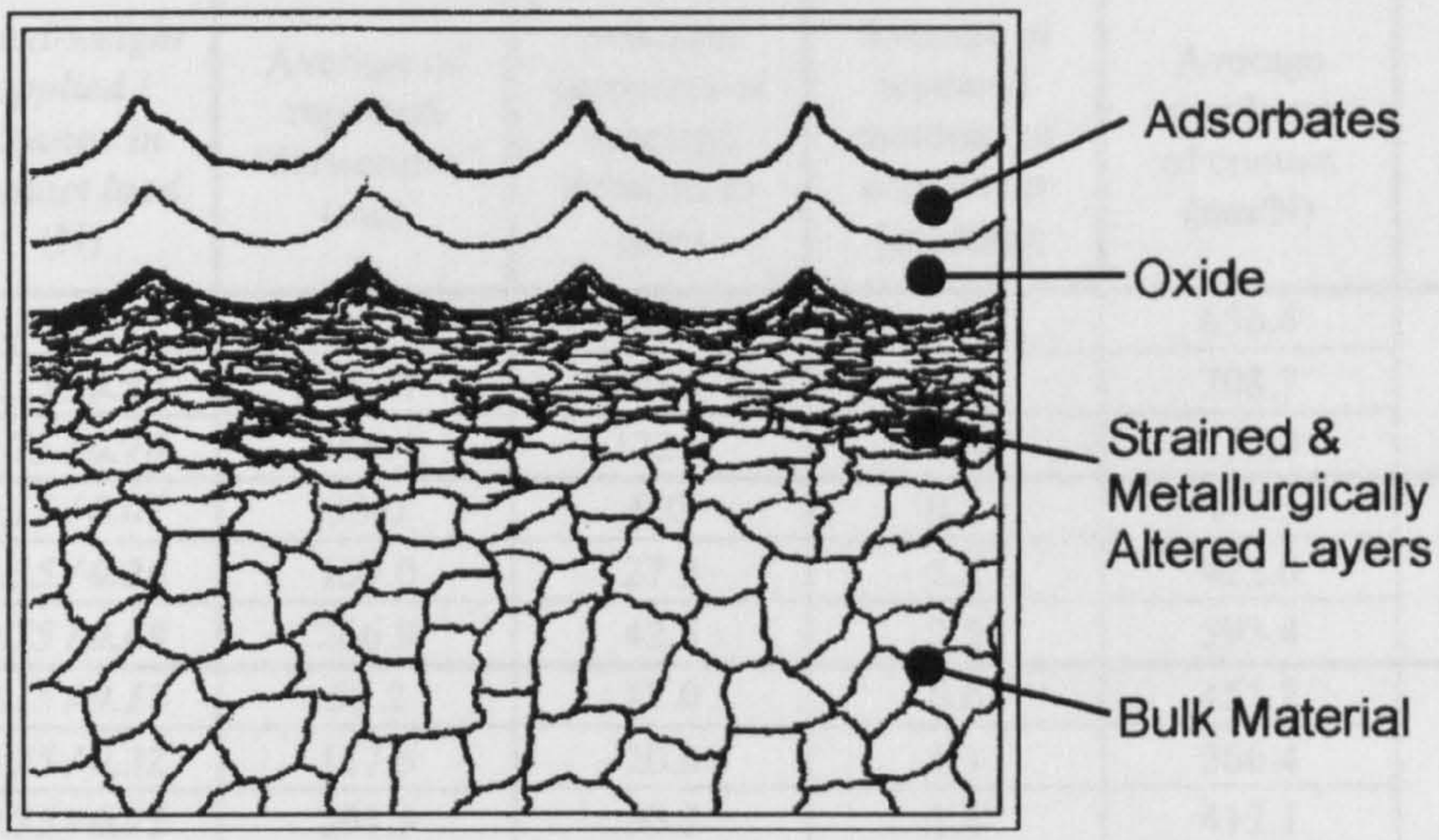


Figure 6.4 Schematic section through a turned surface for illustrating the different possible metallurgical layers that could exist on any machined metal surface (taken from [100]).

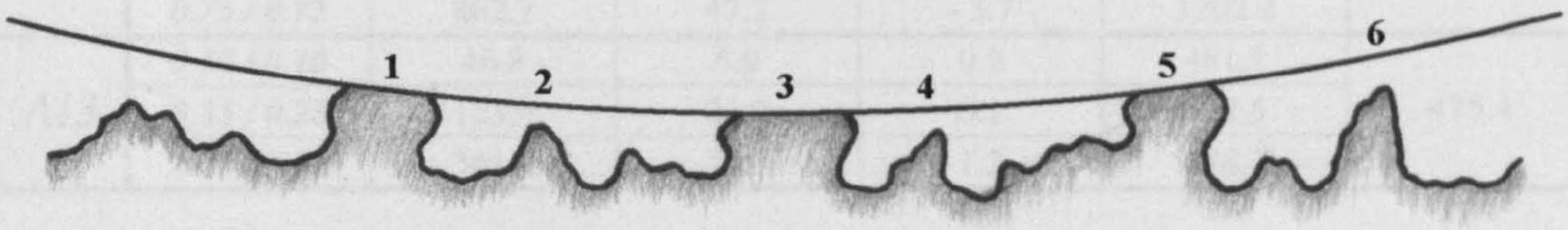


Figure 6.5 Exaggerated representation of a section through a contact region for showing the possible increase in contact area as the deformation of asperities progresses.

	*Dead-weight applied / **Change in contact load*** (N)	Average of repeated deflections (nm)	Standard deviation of repeated deflections (nm)	Average of repeated rotations of contact tip (μradian)	Average compliance of contact (nm/N)	Slope of best-fit straight line (nm/N)
St1	*0.15 / **0.11***	70.5	10.6	1.4	656.4	758.8
	*0.35 / **0.34***	242.8	70.6	0.4	708.7	
	*0.75 / **0.76***	585.4	132.0	1.8	772.0	
St2	*0.15 / **0.07***	33.0	9.0	0.2	447.5	396.6
	*0.35 / **0.26***	109.0	27.9	1.2	422.0	
	*0.75 / **0.68***	266.9	42.5	2.5	393.4	
St3	*0.15 / **0.12***	55.2	12.0	0.6	453.2	404.3
	*0.35 / **0.32***	117.5	20.8	1.1	366.4	
	*0.75 / **0.71***	291.3	50.5	1.6	412.1	
Co1	*0.15 / **0.13***	177.4	37.8	1.0	1383.0	775.2
	*0.35 / **0.32***	306.0	104.3	- 0.3	942.8	
	*0.75 / **0.74***	535.1	168.4	0.1	727.3	
Co2	*0.15 / **0.09***	44.1	9.0	0.6	471.9	497.3
	*0.35 / **0.31***	167.9	21.4	1.2	544.5	
	*0.75 / **0.73***	357.3	55.7	2.7	488.2	
Co3	*0.15 / **0.12***	48.2	12.3	- 0.2	402.8	332.6
	*0.35 / **0.32***	117.2	12.4	0.0	371.9	
	*0.75 / **0.75***	243.5	23.7	1.8	325.8	
Al1	*0.15 / **0.10***	112.9	20.8	0.1	1102.2	1106.0
	*0.35 / **0.32***	338.7	20.9	- 0.4	1071.4	
	*0.75 / **0.73***	813.7	104.6	0.3	1107.7	
Al2	*0.15 / **0.08***	95.0	18.7	- 0.1	1154.4	1251.2
	*0.35 / **0.30***	468.4	65.2	- 1.8	1562.0	
	*0.75 / **0.72***	862.7	47.2	- 3.7	1202.2	
Al3	*0.15 / **0.10***	46.8	5.9	0.2	481.1	475.4
	*0.35 / **0.28***	123.0	21.2	0.2	432.5	
	*0.75 / **0.71***	341.6	29.9	1.2	478.7	

Table 6.3 Results of the different-point repeated contact tests on steel, copper, and aluminium specimens of surfaces of different roughness.

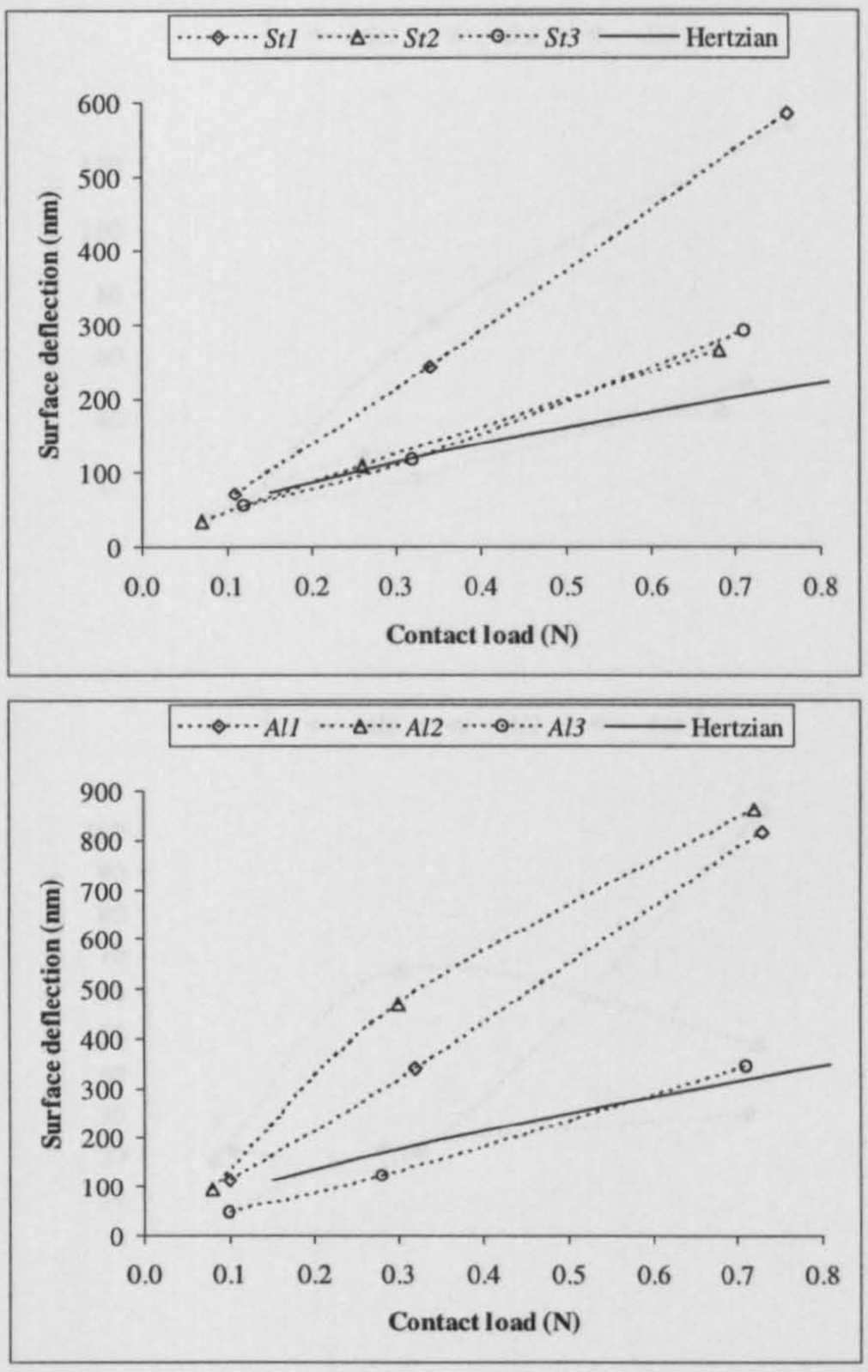


Figure 6.6 Plots of the average deflections of the different-point repeated contact tests on (a) steel, (b) copper, and (c) aluminium surfaces of different roughness.

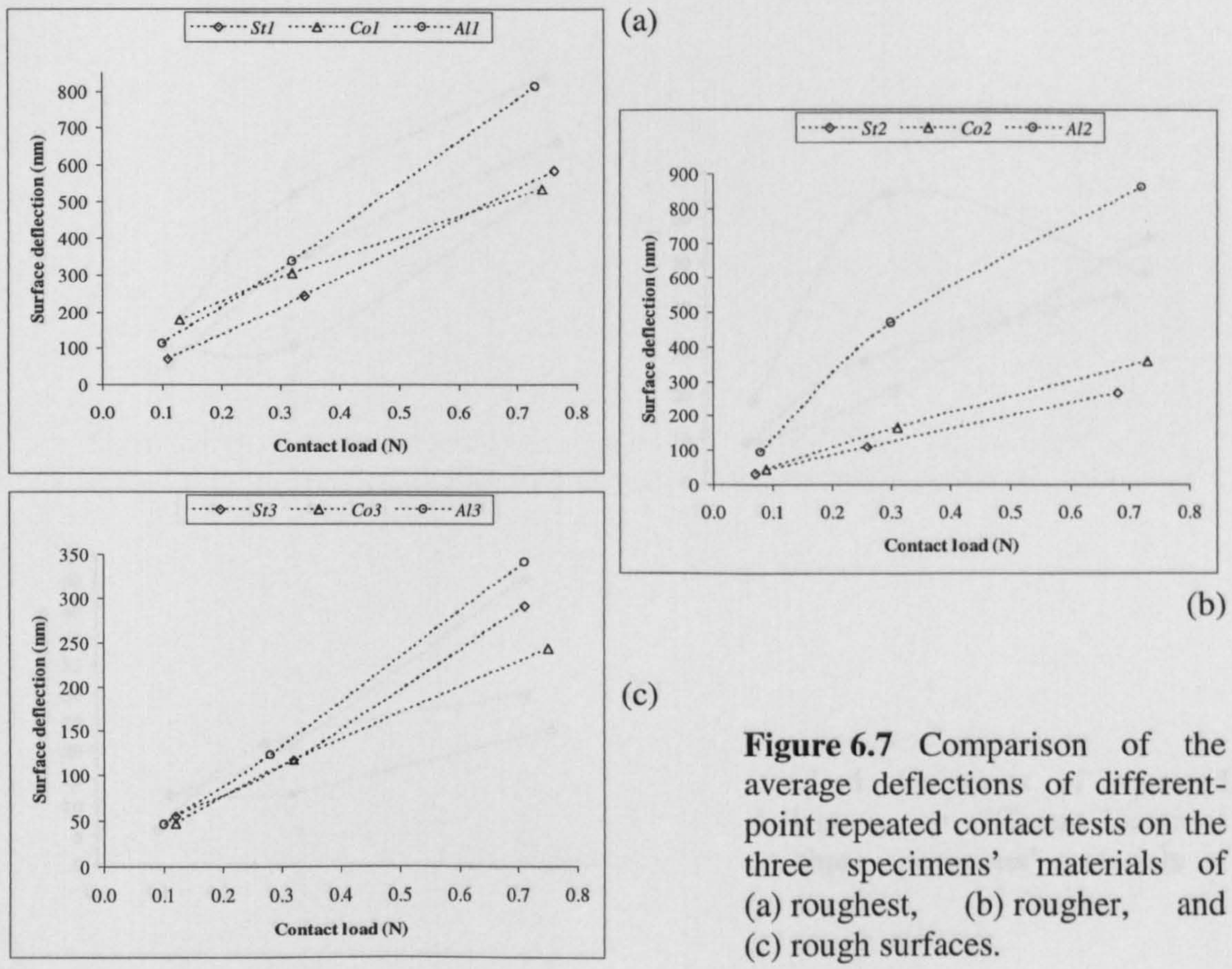
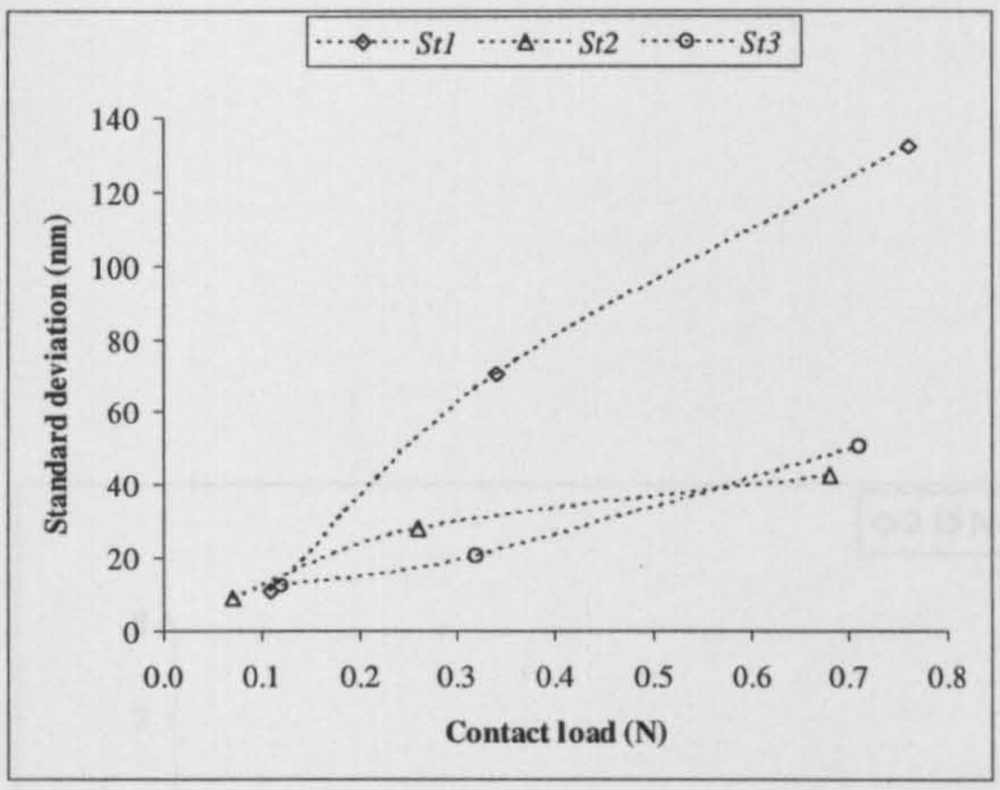
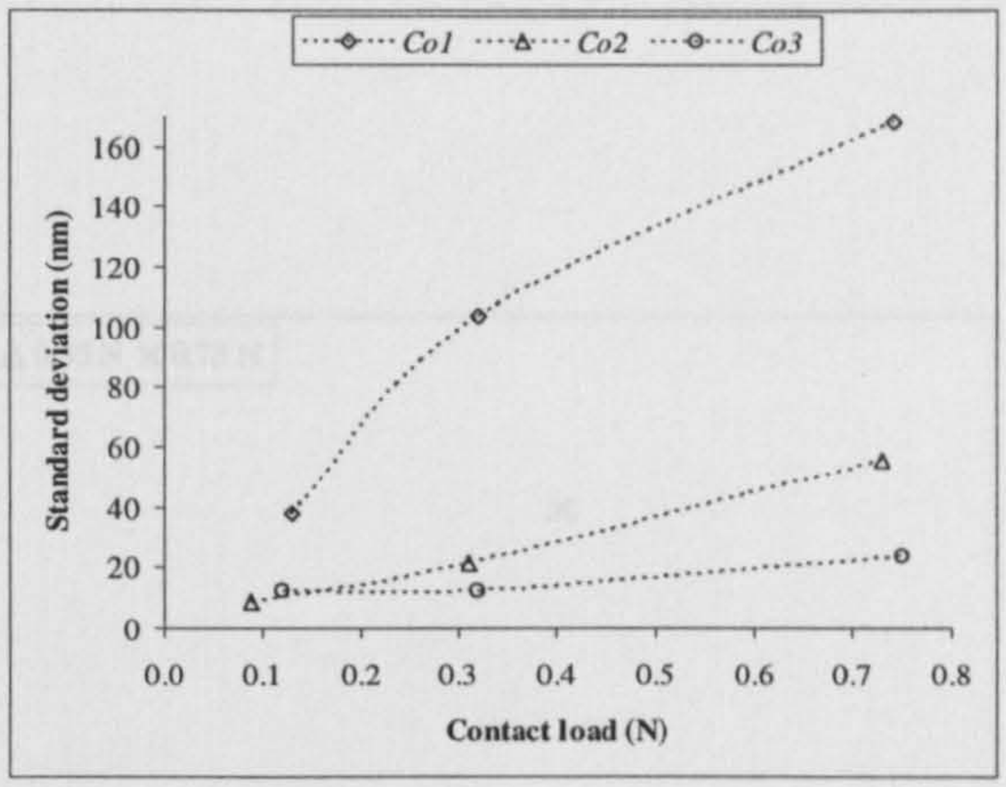


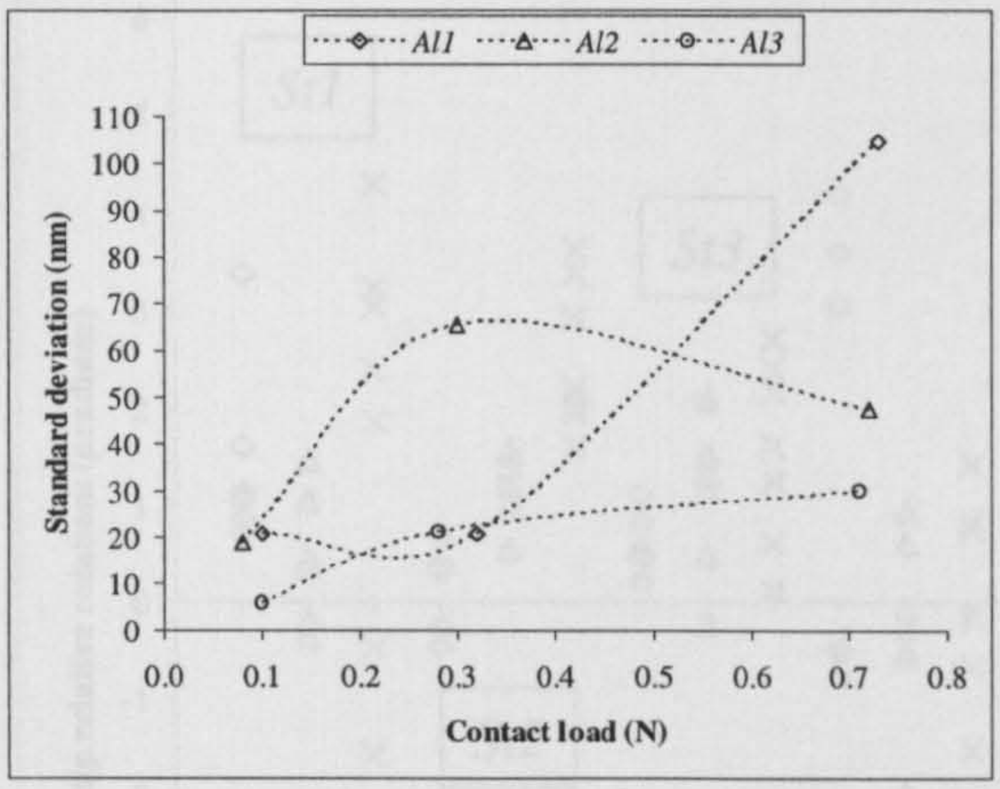
Figure 6.7 Comparison of the average deflections of different-point repeated contact tests on the three specimens' materials of (a) roughest, (b) rougher, and (c) rough surfaces.



(a)

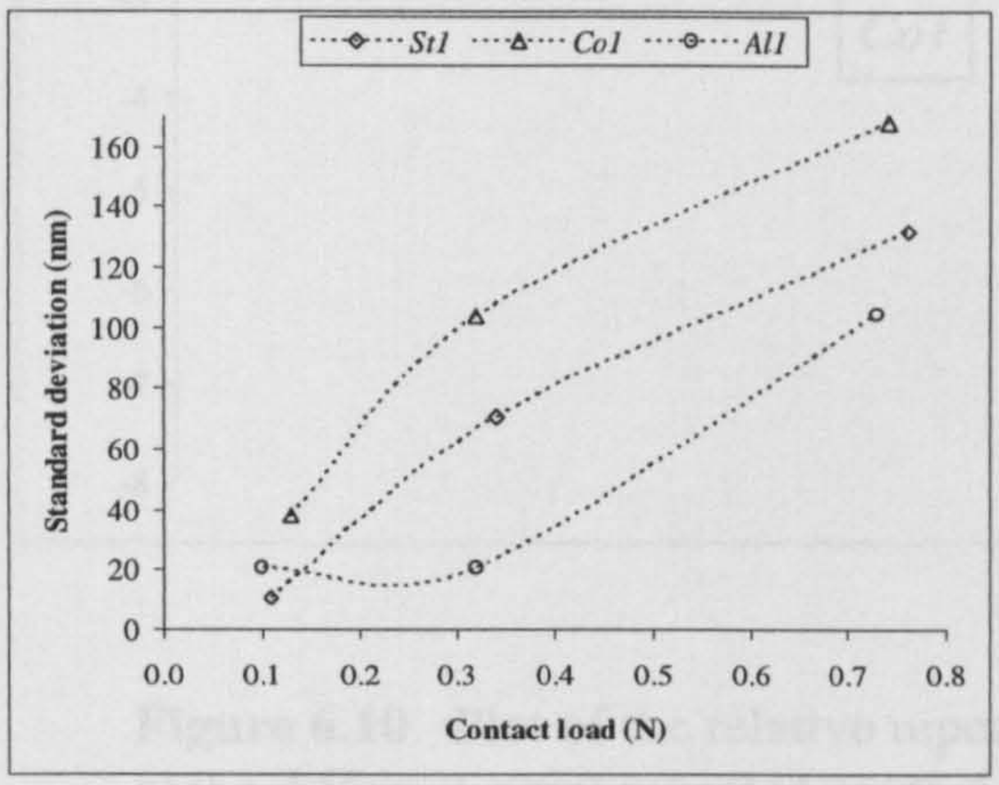


(b)

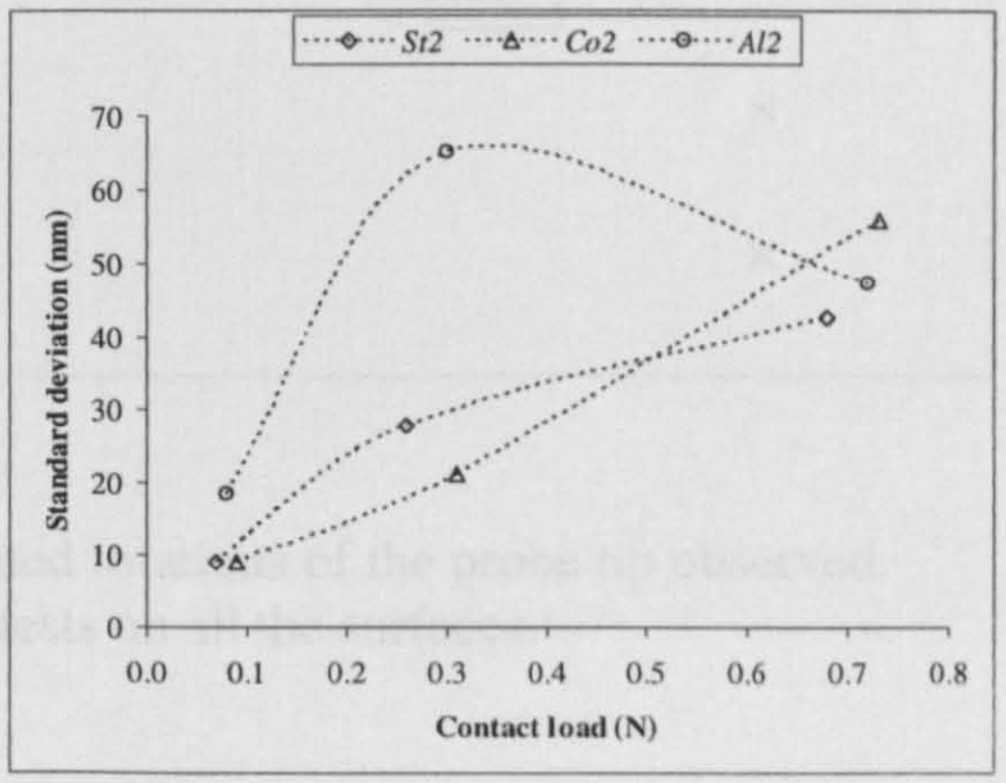


(c)

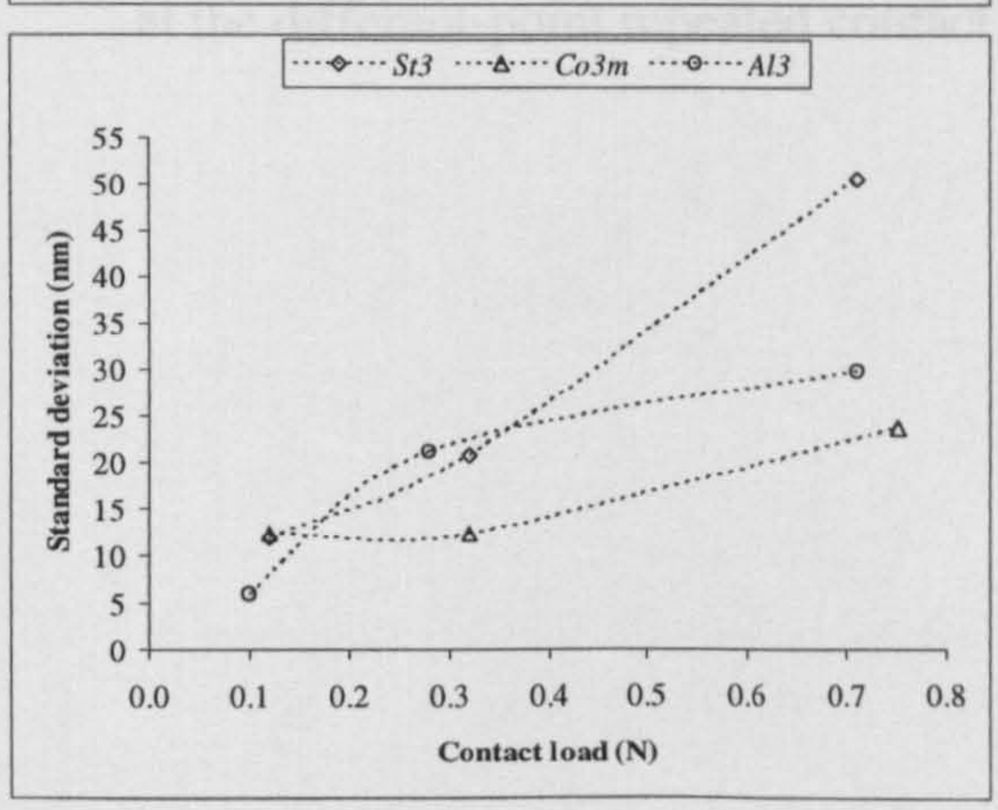
Figure 6.8 Plots of the standard deviations of repeated deflections on different locations on (a) steel, (b) copper, and (c) aluminium surfaces of different roughness.



(a)



(b)



(c)

Figure 6.9 Comparison of the standard deviations of repeated deflections on different locations on three specimens' materials of (a) roughest, (b) rougher, and (c) rough surfaces.

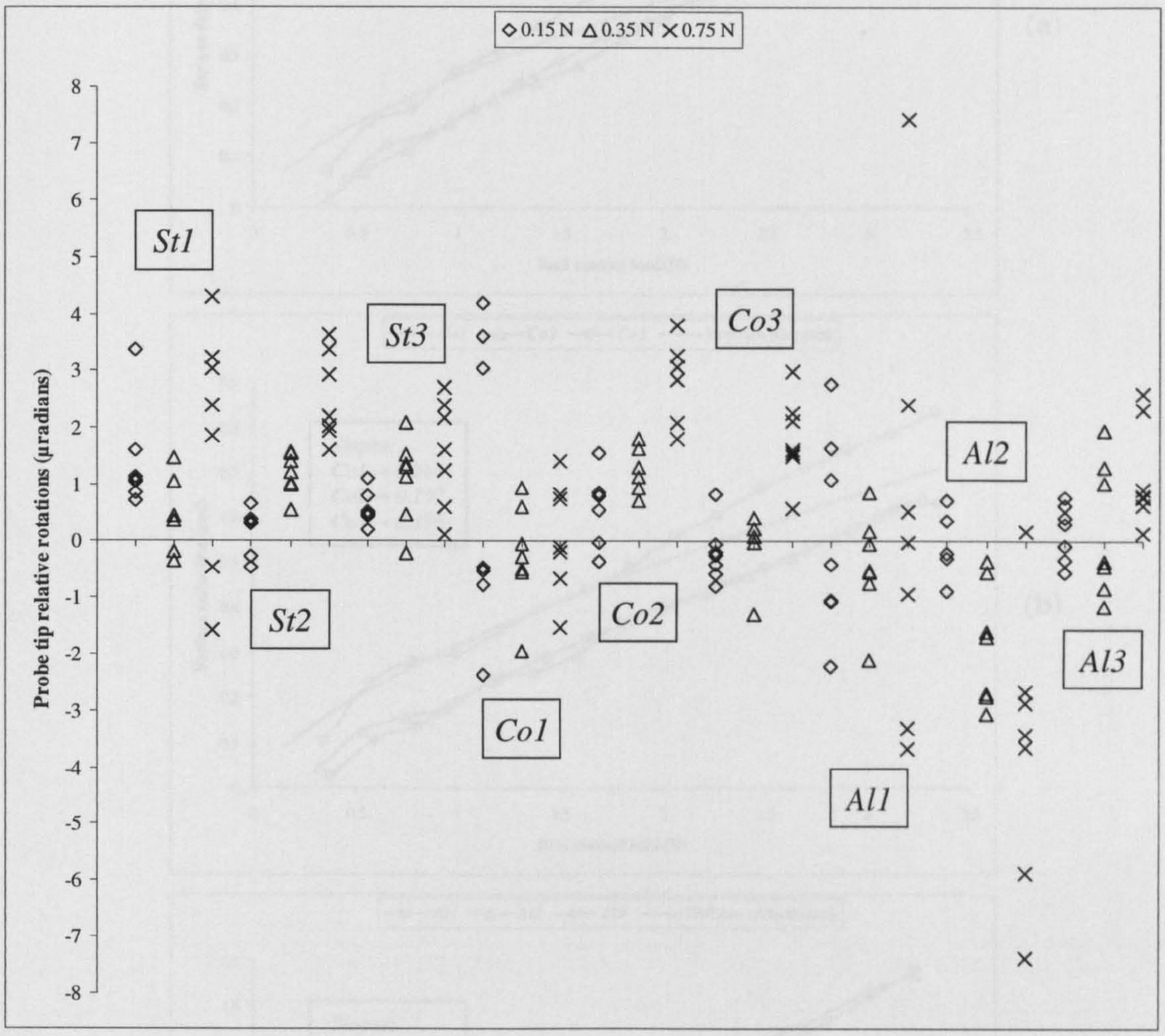
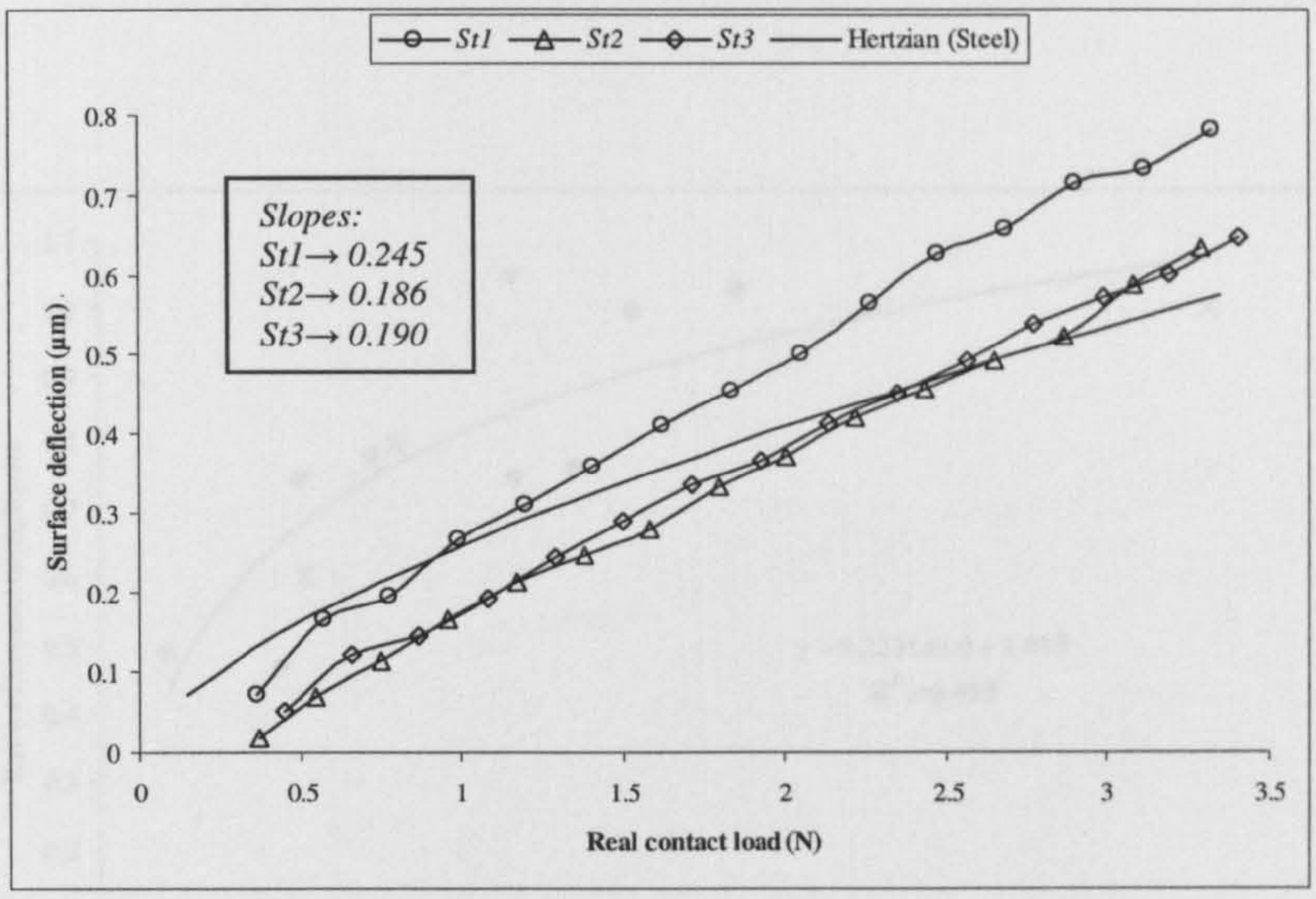
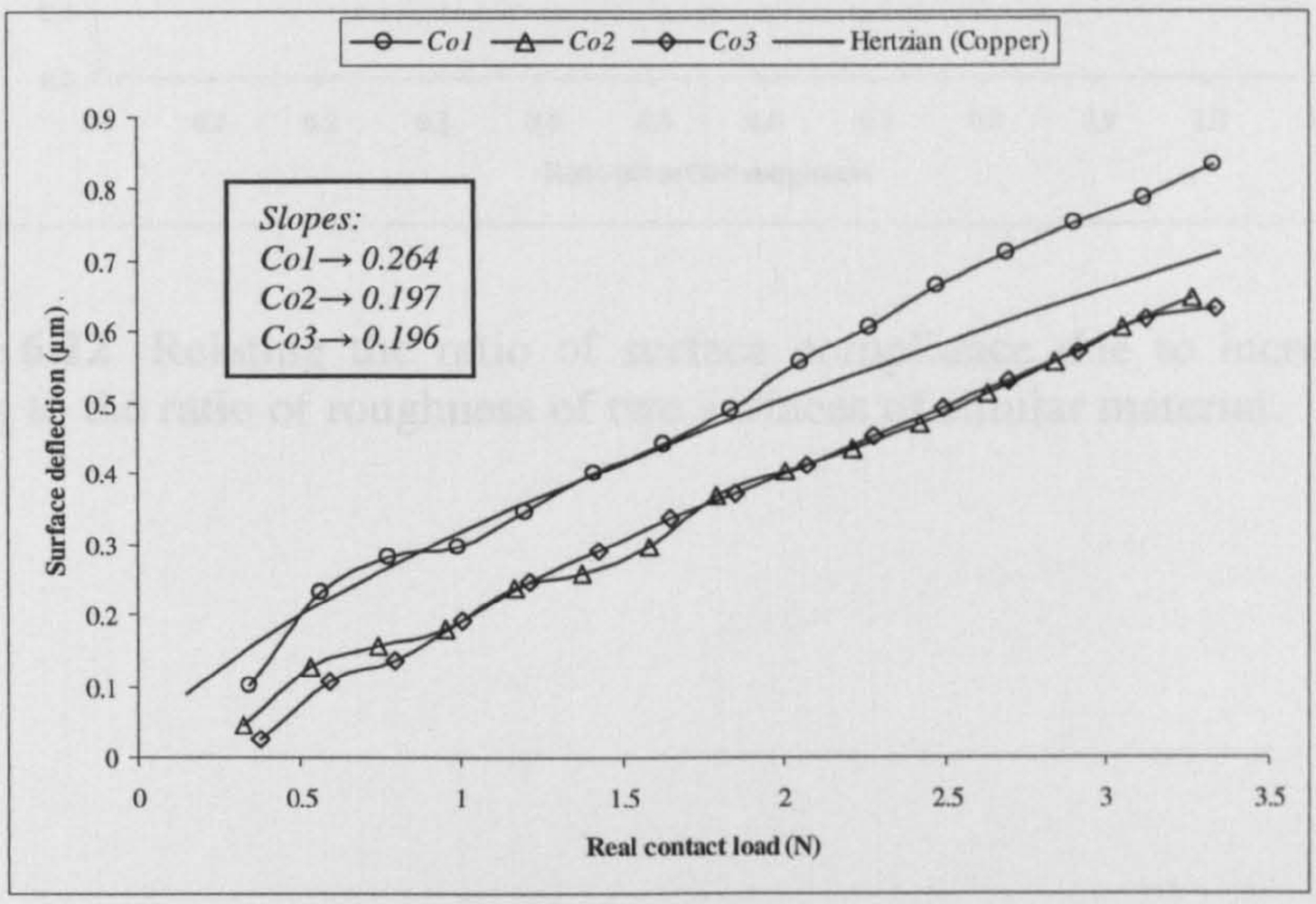


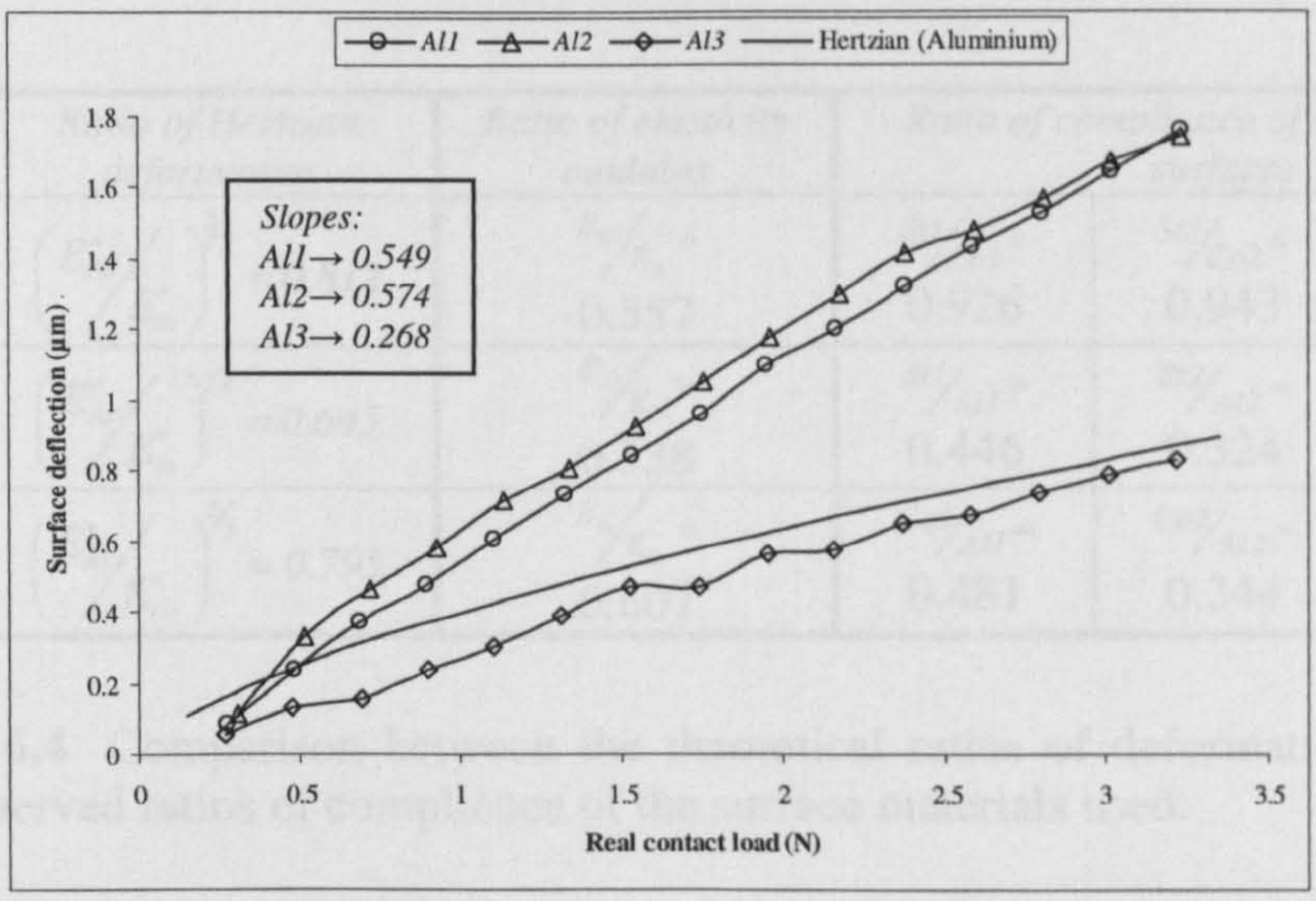
Figure 6.10 Plot of the relative repeated rotations of the probe tip observed at the different-point repeated contact tests on all the surfaces.



(a)



(b)



(c)

Figure 6.11 Plots of surface deflections resulting from the incremental loading on a point on (a) steel, (b) copper, and (c) aluminium surfaces of different roughness.

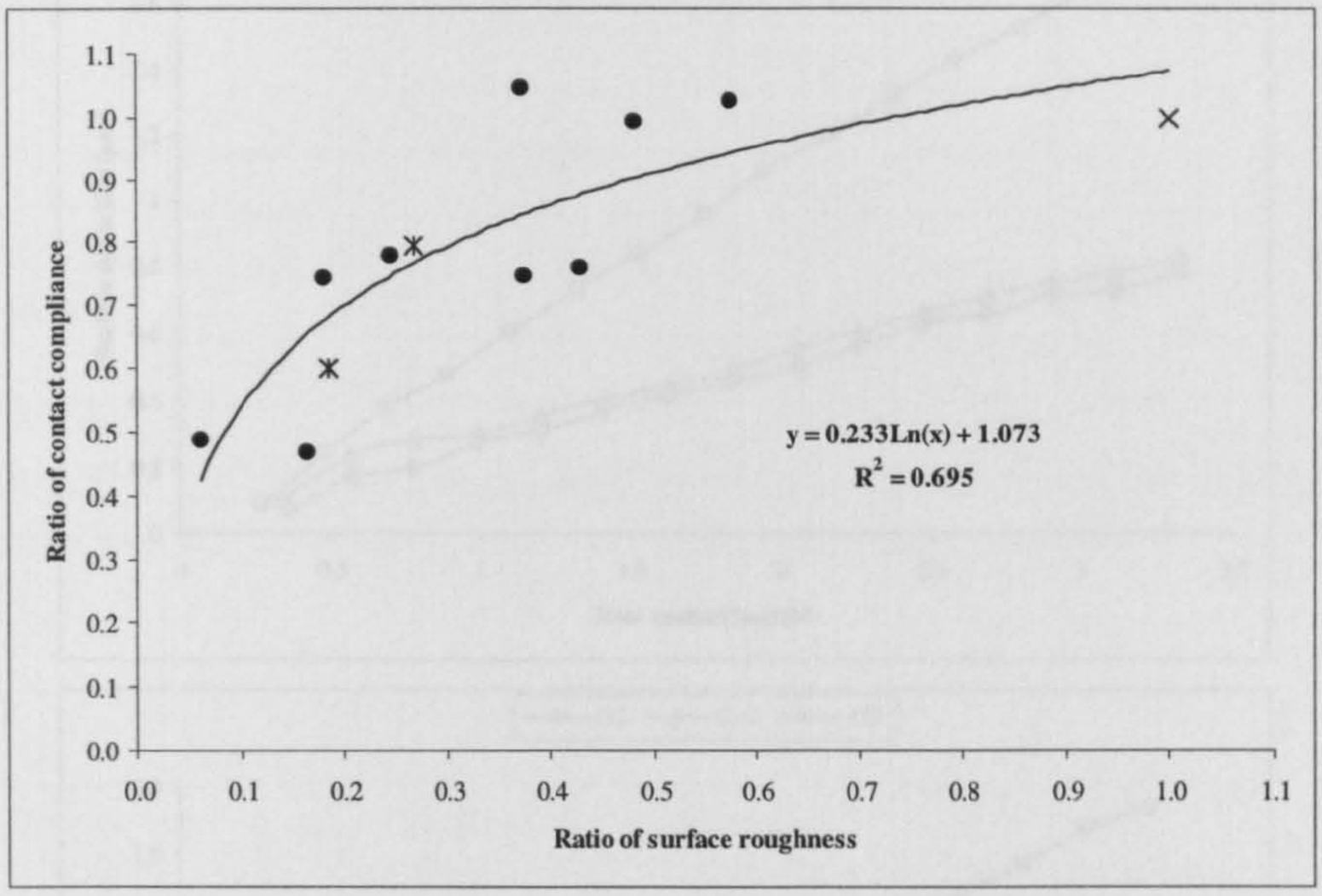


Figure 6.12 Relating the ratio of surface compliance due to incremental loading to the ratio of roughness of two surfaces of similar material.

	Ratio of Hertzian deformation	*Ratio of elasticity modulus*	*Ratio of compliance of the examined surfaces*		
$Steel/Co$	$\left(E^*_{Co}/E^*_{St}\right)^{2/3} = 0.812$	$E_{Co}/E_{St} =$ 0.557	$St1/Co1 =$ 0.926	$St2/Co2 =$ 0.943	$St3/Co3 =$ 0.971
$Steel/Al$	$\left(E^*_{Al}/E^*_{St}\right)^{2/3} = 0.645$	$E_{Al}/E_{St} =$ 0.338	$St1/Al1 =$ 0.446	$St2/Al2 =$ 0.324	$St3/Al3 =$ 0.712
Co/Al	$\left(E^*_{Al}/E^*_{Co}\right)^{2/3} = 0.795$	$E_{Al}/E_{Co} =$ 0.607	$Co1/Al1 =$ 0.481	$Co2/Al2 =$ 0.344	$Co3/Al3 =$ 0.733

Table 6.4 Comparison between the theoretical ratios of deformation and the observed ratios of compliance of the surface materials used.

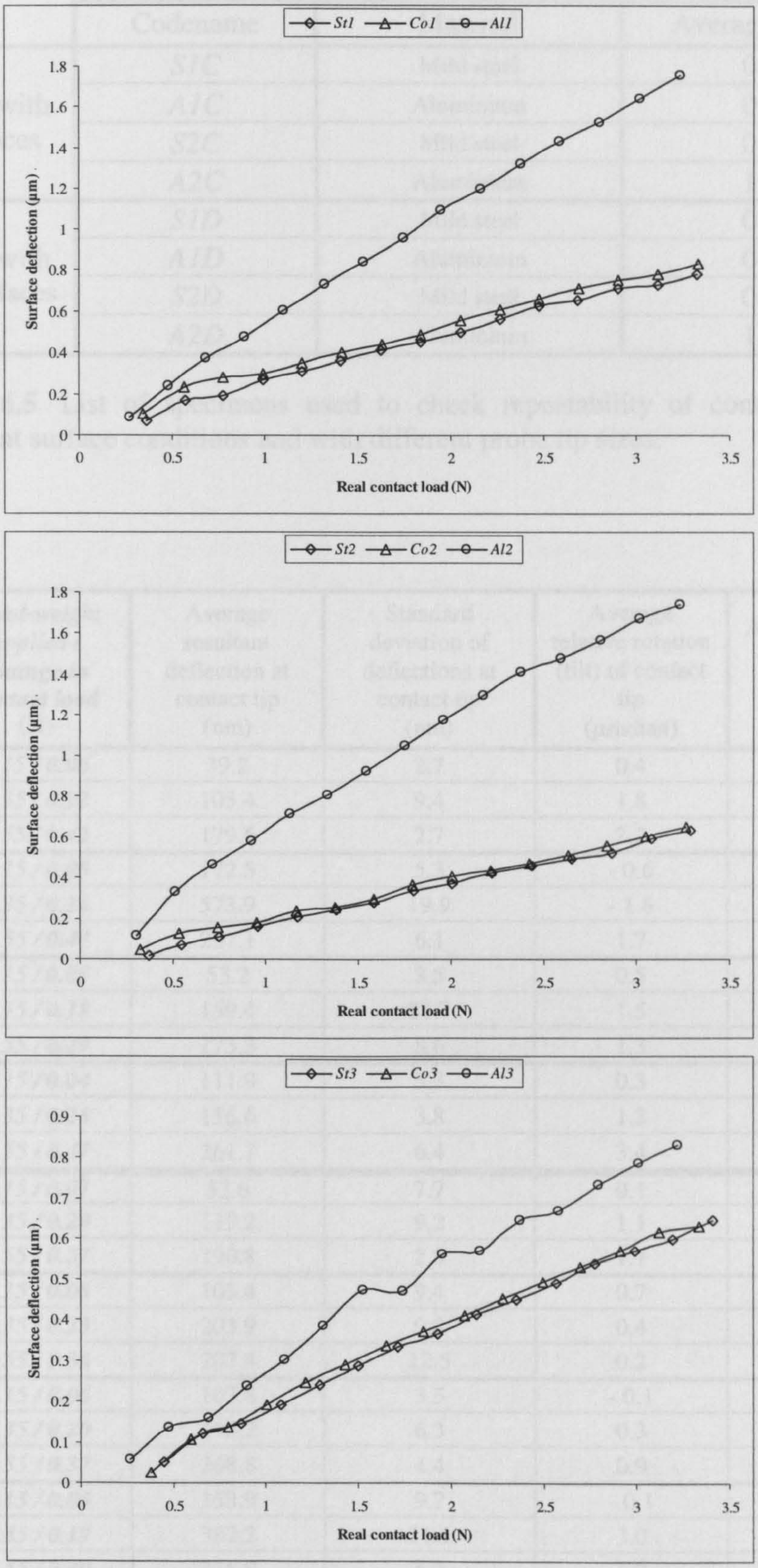


Figure 6.13 Comparison of surface deflections due to the incremental loading between the three specimens' materials of (a) roughest, (b) rougher, and (c) rough surfaces.

	Codename	Material	Average R_q (μm)
Specimens with **clean** surfaces	*S1C*	Mild steel	0.110
	A1C	Aluminium	0.424
	S2C	Mild steel	0.533
	A2C	Aluminium	1.480
Specimens with **unclean** surfaces	*S1D*	Mild steel	0.110
	A1D	Aluminium	0.424
	S2D	Mild steel	0.533
	A2D	Aluminium	1.480

Table 6.5 List of specimens used to check repeatability of contact on different surface conditions and with different probe tip sizes.

	Dead-weight applied / ***Change in contact load*** (N)	Average resultant deflection at contact tip (nm)	Standard deviation of deflections at contact tip (nm)	Average relative rotation (tilt) of contact tip (μradian)	Average relative compliance of contact (nm/N)
S1C	*0.15* / ***0.06***	39.2	2.7	0.4	617.0
	0.35 / ***0.22***	105.4	9.4	1.8	470.8
	0.55 / ***0.42***	179.6	2.7	2.2	424.7
S1D	*0.15* / ***0.05***	172.5	5.3	- 0.6	3166.3
	0.35 / ***0.26***	573.9	19.9	- 1.6	2173.3
	0.55 / ***0.47***	297.1	6.1	1.7	630.7
S2C	*0.15* / ***0.05***	53.2	8.5	0.5	1142.7
	0.35 / ***0.18***	159.4	21.7	1.5	897.0
	0.55 / ***0.29***	173.3	8.6	1.3	607.1
S2D	*0.15* / ***0.04***	111.9	5.3	0.3	2656.2
	0.35 / ***0.26***	156.6	3.8	1.3	606.7
	0.55 / ***0.47***	261.7	6.4	3.4	557.1
A1C	*0.15* / ***0.07***	53.6	7.7	0.1	819.9
	0.35 / ***0.20***	119.2	9.2	1.1	612.0
	0.55 / ***0.37***	190.8	2.5	1.7	519.8
A1D	*0.15* / ***0.05***	105.4	9.4	0.7	2266.7
	0.35 / ***0.23***	203.9	9.3	0.4	902.3
	0.55 / ***0.36***	297.4	12.5	0.2	818.8
A2C	*0.15* / ***0.06***	107.8	3.5	- 0.1	1805.8
	0.35 / ***0.20***	193.2	6.3	0.3	993.1
	0.55 / ***0.37***	368.8	4.4	0.9	990.7
A2D	*0.15* / ***0.05***	153.9	9.7	- 0.1	3202.4
	0.35 / ***0.19***	382.2	11.1	3.0	2062.2
	0.55 / ***0.40***	415.9	5.2	1.9	1041.4

Table 6.6 Results of the last six readings of same-point repeated contact tests on steel and aluminium specimens of clean and unclean surfaces using a contact tip of 5 mm diameter.

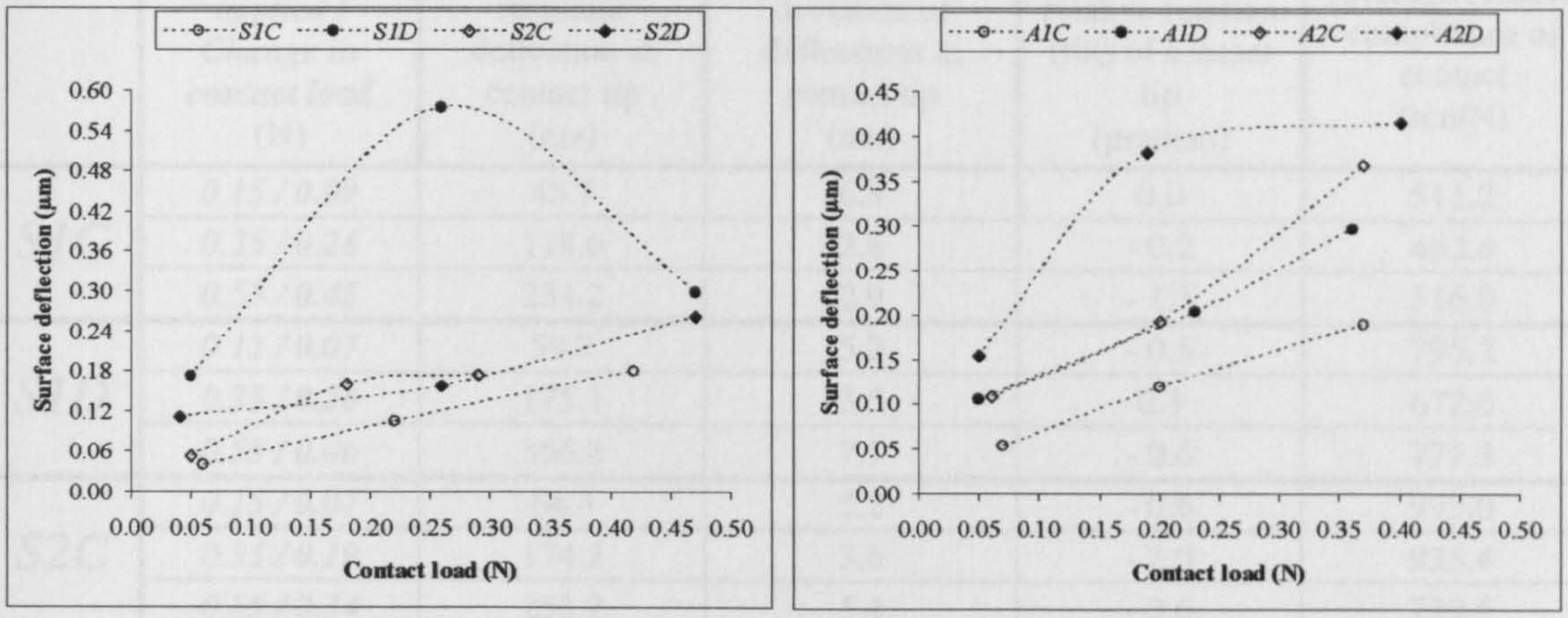


Figure 6.14 Plots of "elastic" deflections of same-point repeated contact tests with a contact tip of 5 mm diameter on clean and unclean specimens of (a) mild steel and (b) aluminium surfaces of different roughness.

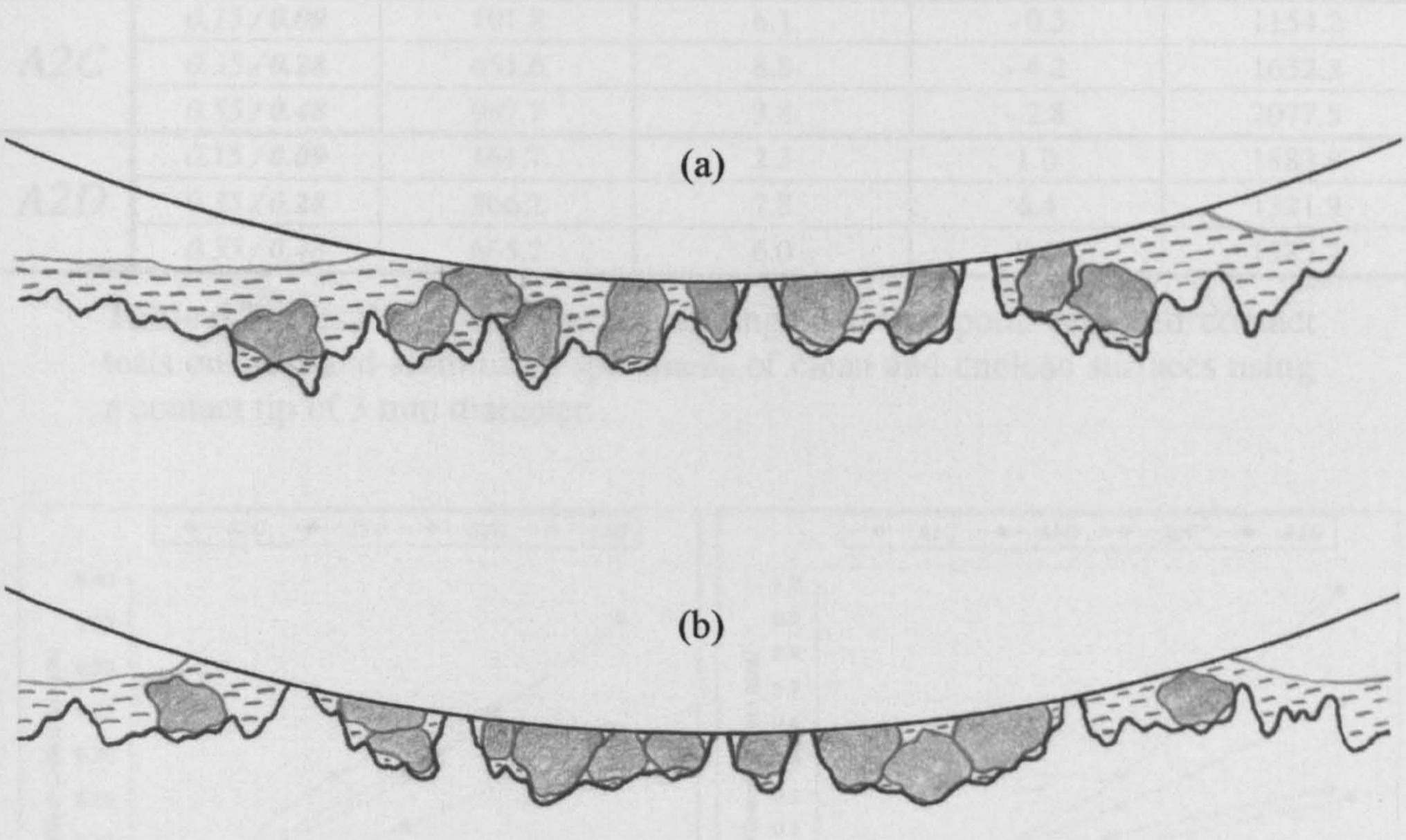
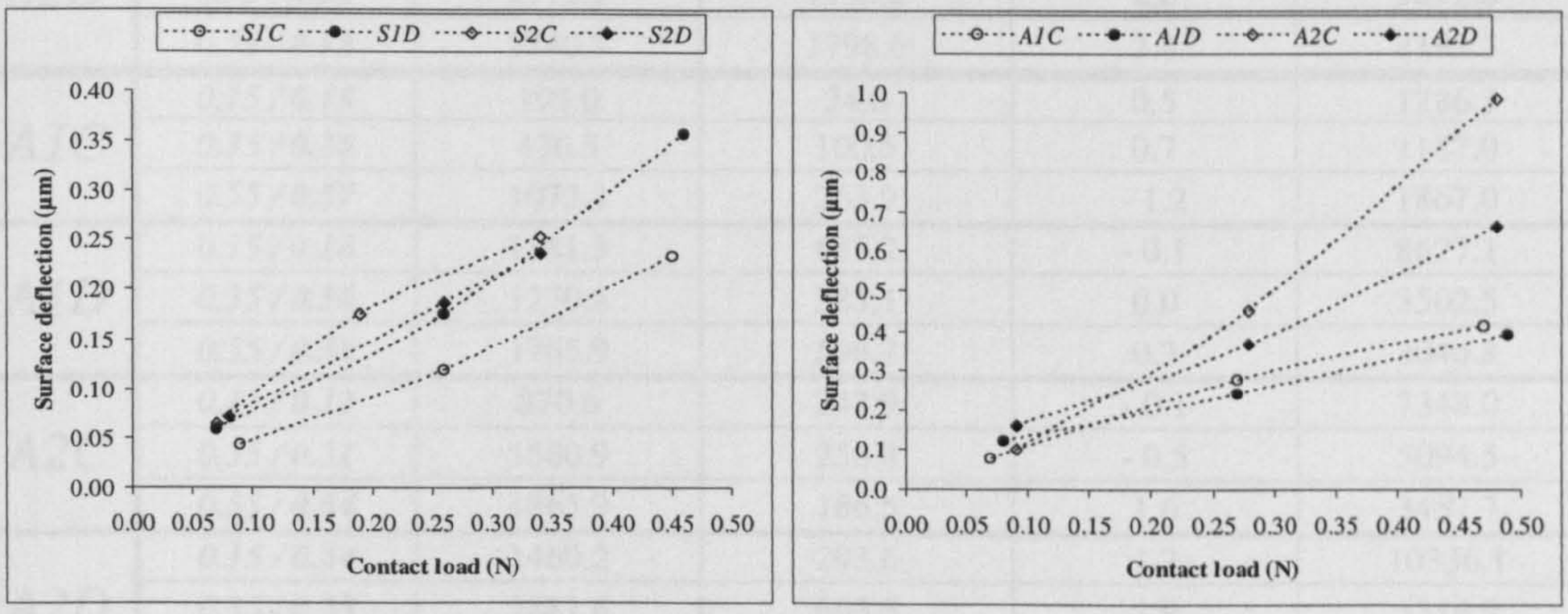


Figure 6.15 Exaggerated representation of two possible contact conditions on unclean rough surfaces of (a) low and (b) high density of asperities and dust particles. At the latter condition, further "elastic" deflections at the same repeated load (beyond the initial plastic deformation) are believed to be less than at the former condition.

	*Dead-weight applied / **Change in contact load*** (N)	Average resultant deflection at contact tip (nm)	Standard deviation of deflections at contact tip (nm)	Average relative rotation (tilt) of contact tip (μradian)	Average relative compliance of contact (nm/N)
S1C	*0.15 / **0.09***	43.7	6.8	0.0	511.2
	*0.35 / **0.26***	118.6	2.8	- 0.2	462.4
	*0.55 / **0.45***	234.2	2.9	- 1.3	516.0
S1D	*0.15 / **0.07***	59.2	5.2	- 0.5	795.2
	*0.35 / **0.26***	175.1	3.5	0.3	672.6
	*0.55 / **0.46***	356.8	7.7	- 0.6	771.3
S2C	*0.15 / **0.07***	64.3	5.2	- 0.6	995.0
	*0.35 / **0.19***	174.2	3.6	- 2.0	935.4
	*0.55 / **0.34***	253.2	5.4	- 3.6	739.5
S2D	*0.15 / **0.08***	70.9	4.9	- 0.4	851.3
	*0.35 / **0.26***	186.1	6.4	0.0	704.7
	*0.55 / **0.34***	236.0	3.7	- 0.2	696.3
A1C	*0.15 / **0.07***	78.4	4.7	0.4	1112.9
	*0.35 / **0.27***	277.6	2.3	- 0.5	1045.9
	*0.55 / **0.47***	411.6	3.1	- 0.2	868.4
A1D	*0.15 / **0.08***	122.1	2.9	- 1.4	1462.7
	*0.35 / **0.27***	241.5	6.5	- 0.8	908.4
	*0.55 / **0.49***	390.1	3.2	- 1.2	795.5
A2C	*0.15 / **0.09***	101.8	6.1	- 0.3	1154.2
	*0.35 / **0.28***	451.6	8.8	- 4.2	1632.3
	*0.55 / **0.48***	987.7	3.8	- 2.8	2077.5
A2D	*0.15 / **0.09***	161.7	2.3	1.0	1883.8
	*0.35 / **0.28***	366.2	7.8	6.4	1321.9
	*0.55 / **0.48***	665.2	6.0	8.6	1385.7

Table 6.7 Results of the last six readings of same-point repeated contact tests on steel and aluminium specimens of clean and unclean surfaces using a contact tip of 3 mm diameter.



(a) (b)

Figure 6.16 Plots of "elastic" deflections of same-point repeated contact tests with a contact tip of 3 mm diameter on clean and unclean specimens of (a) mild steel and (b) aluminium surfaces of different roughness.

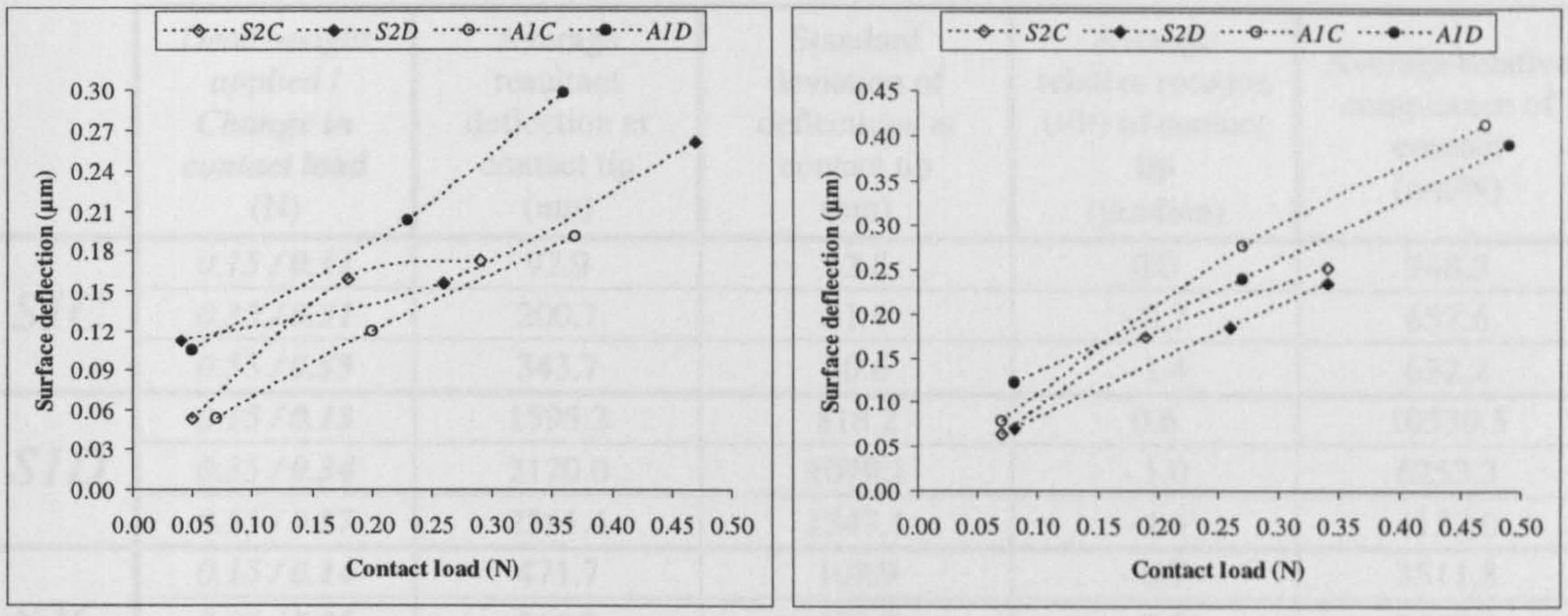


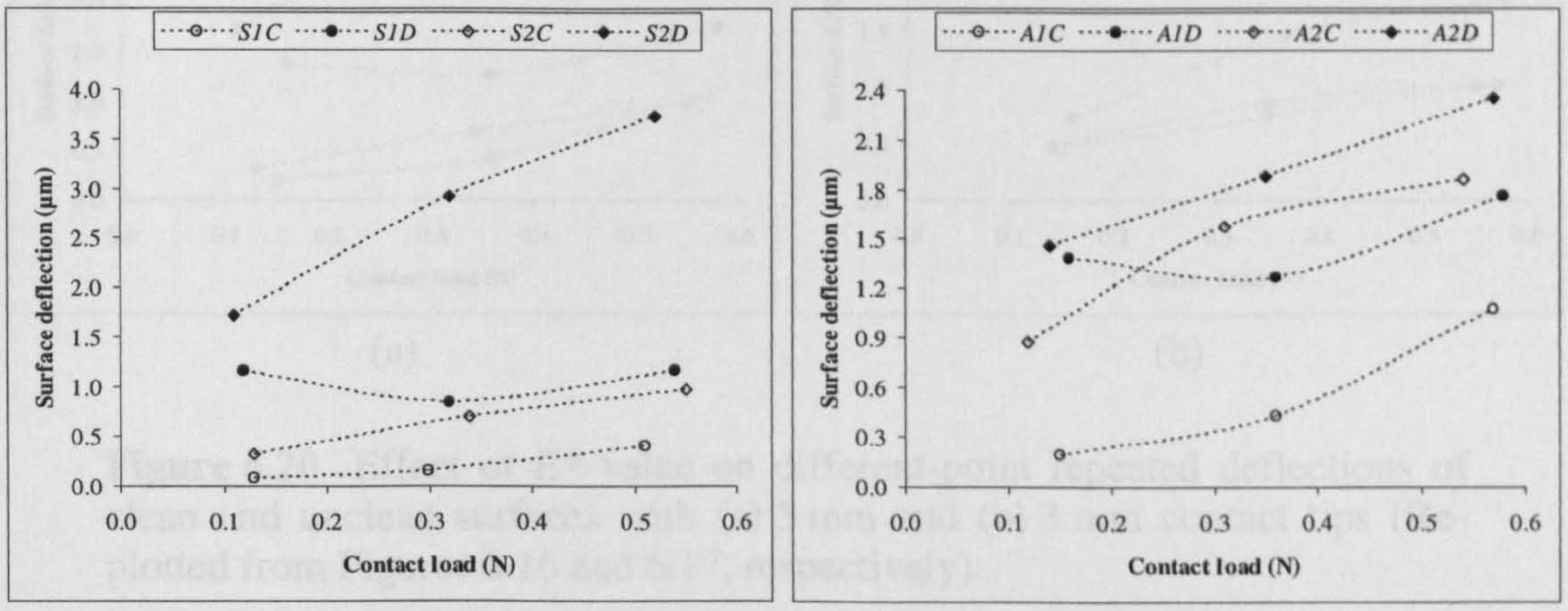
Figure 6.17 Effect of E^* value on same-point repeated deflections of clean and unclean surfaces with (a) 5 mm and (b) 3 mm contact tips (Re-plotted from Figures 6.12 and 6.14, respectively).

	Dead-weight applied / ***Change in contact load*** (N)	Average resultant deflection at contact tip (nm)	Standard deviation of deflections at contact tip (nm)	Average relative rotation (tilt) of contact tip (µradian)	Average relative compliance of contact (nm/N)
S1C	*0.15* / ***0.13***	78.7	5.6	0.2	607.4
	0.35 / ***0.30***	155.2	25.2	0.4	512.0
	0.55 / ***0.51***	411.7	41.6	0.7	816.6
S1D	*0.15* / ***0.12***	1157.3	592.6	- 0.3	10279.5
	0.35 / ***0.32***	836.1	678.6	0.9	2628.2
	0.55 / ***0.54***	1162.2	1149.8	0.0	2154.0
S2C	*0.15* / ***0.13***	326.2	70.9	0.0	2663.3
	0.35 / ***0.34***	702.7	104.5	- 0.8	2048.5
	0.55 / ***0.55***	960.6	123.4	0.2	1746.7
S2D	*0.15* / ***0.11***	1715.9	669.9	1.2	16038.6
	0.35 / ***0.32***	2919.1	1133.2	2.6	9072.0
	0.55 / ***0.52***	3740.3	1798.6	2.4	7192.3
A1C	*0.15* / ***0.15***	193.0	24.6	0.5	1286.3
	0.35 / ***0.36***	430.3	100.5	0.7	1187.0
	0.55 / ***0.57***	1072.1	263.9	- 1.2	1867.0
A1D	*0.15* / ***0.16***	1381.3	683.2	- 0.1	8677.1
	0.35 / ***0.36***	1270.8	385.1	0.0	3502.5
	0.55 / ***0.58***	1765.9	596.7	0.3	3046.8
A2C	*0.15* / ***0.12***	870.6	243.9	- 0.1	7348.0
	0.35 / ***0.31***	1580.9	256.8	- 0.5	5094.5
	0.55 / ***0.54***	1865.9	186.5	1.6	3457.7
A2D	*0.15* / ***0.14***	1460.2	293.6	1.2	10336.1
	0.35 / ***0.35***	1881.6	603.5	1.9	5314.9
	0.55 / ***0.57***	2359.7	676.5	1.8	4155.3

Table 6.8 Results of the different-point repeated contact tests on steel and aluminium specimens of clean and unclean surfaces using the 5 mm tip.

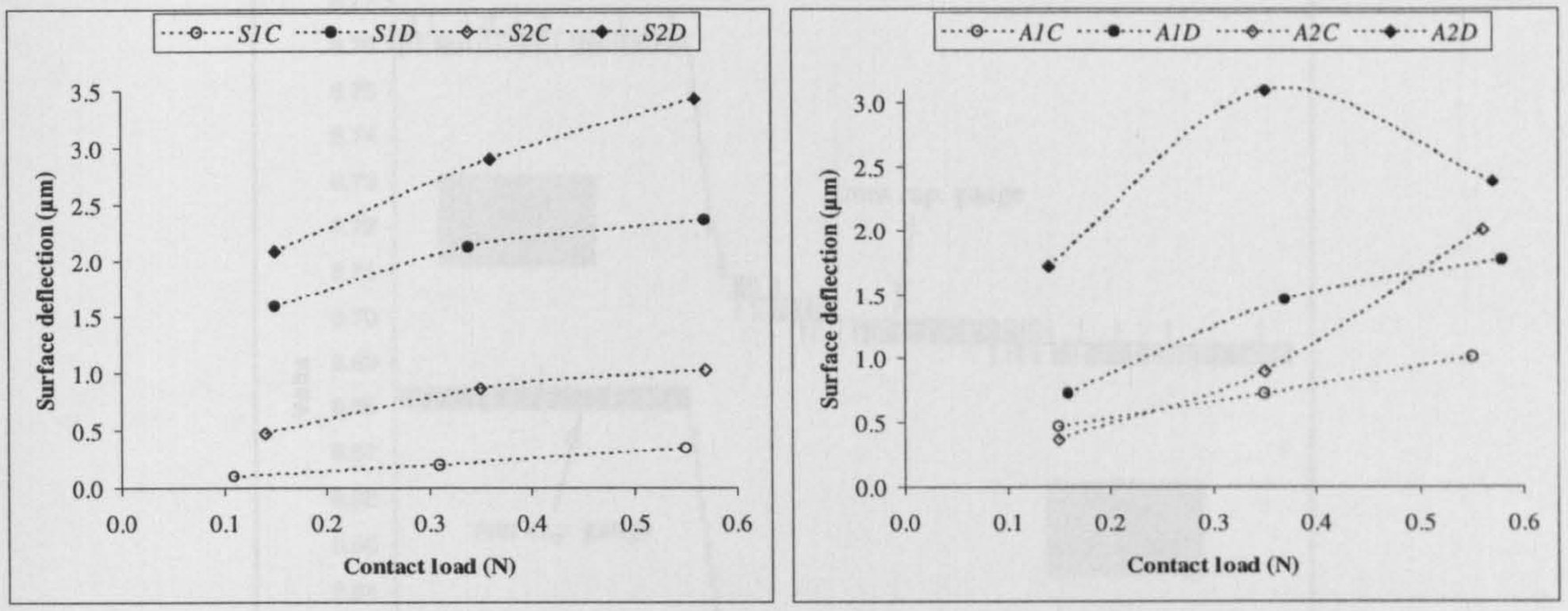
	*Dead-weight applied / **Change in contact load*** (N)	Average resultant deflection at contact tip (nm)	Standard deviation of deflections at contact tip (nm)	Average relative rotation (tilt) of contact tip (μradian)	Average relative compliance of contact (nm/N)
S1C	*0.15 / **0.11***	97.9	12.8	0.0	948.3
	*0.35 / **0.31***	200.7	21.4	- 0.7	657.6
	*0.55 / **0.55***	343.7	30.6	- 1.4	632.2
S1D	*0.15 / **0.15***	1595.2	818.2	0.6	10530.5
	*0.35 / **0.34***	2120.0	1099.1	- 1.0	6253.3
	*0.55 / **0.57***	2361.4	1347.1	- 4.6	4153.0
S2C	*0.15 / **0.14***	471.7	108.9	- 0.7	3511.8
	*0.35 / **0.35***	864.7	128.9	- 2.6	2455.0
	*0.55 / **0.57***	1032.0	101.7	- 5.2	1811.6
S2D	*0.15 / **0.15***	2077.9	652.3	- 1.4	13489.6
	*0.35 / **0.36***	2906.2	842.2	- 0.1	8180.2
	*0.55 / **0.56***	3428.0	1314.7	4.9	6084.2
A1C	*0.15 / **0.15***	469.9	49.2	0.7	3257.1
	*0.35 / **0.35***	730.1	75.7	1.7	2099.8
	*0.55 / **0.55***	1006.8	135.9	1.7	1830.2
A1D	*0.15 / **0.16***	726.7	8.3	0.3	4700.2
	*0.35 / **0.37***	1462.7	38.8	- 1.4	4005.4
	*0.55 / **0.58***	1780.6	74.2	- 2.7	3098.3
A2C	*0.15 / **0.15***	373.8	49.9	- 0.6	2507.0
	*0.35 / **0.35***	903.2	188.7	- 2.5	2577.0
	*0.55 / **0.56***	2018.0	158.8	- 5.9	3583.8
A2D	*0.15 / **0.14***	1719.8	499.3	0.1	11967.6
	*0.35 / **0.35***	3095.1	353.8	3.6	8876.4
	*0.55 / **0.57***	2398.7	230.3	12.9	4249.1

Table 6.9 Results of the different-point repeated contact tests on steel and aluminium specimens of clean and unclean surfaces using the 3 mm tip.



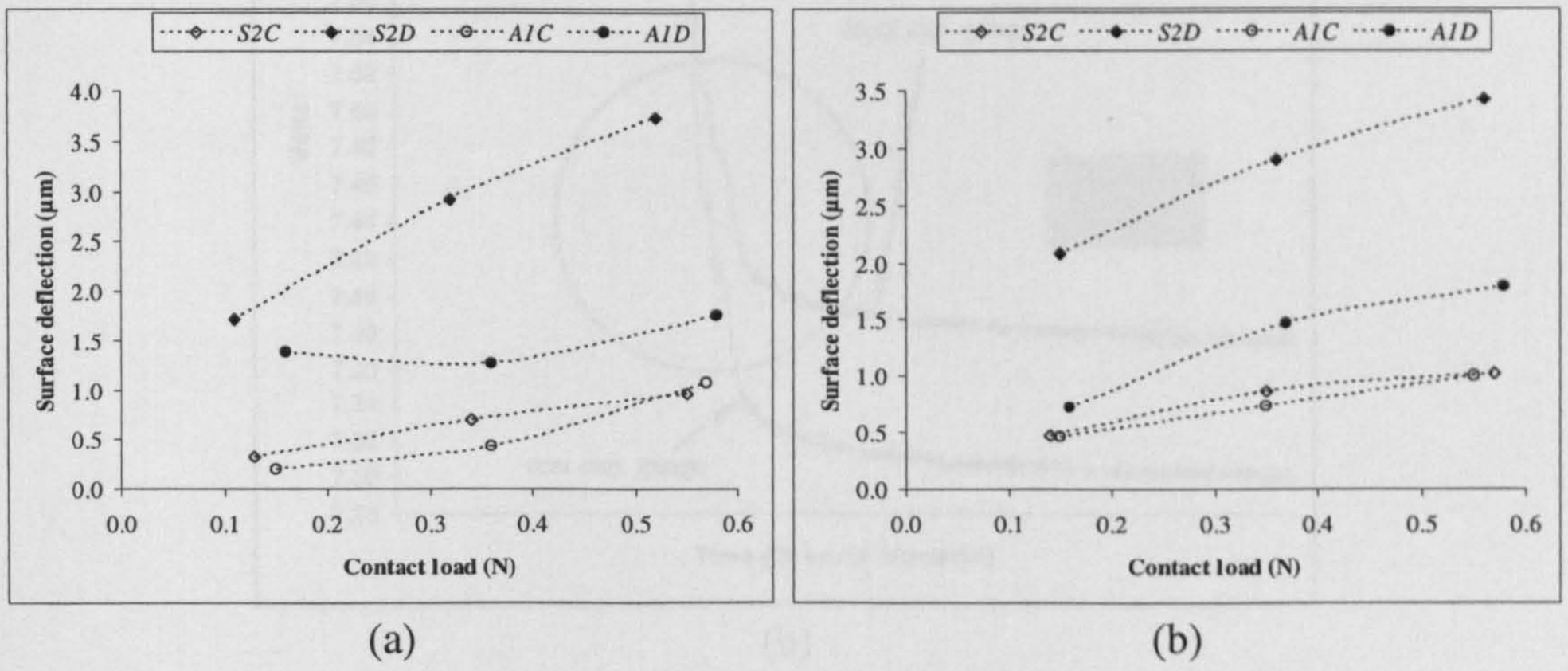
(a) (b)

Figure 6.18 Plots of average deflections of different-point repeated contact tests with a contact tip of 5 mm diameter on clean and unclean specimens of (a) mild steel and (b) aluminium surfaces of different roughness.



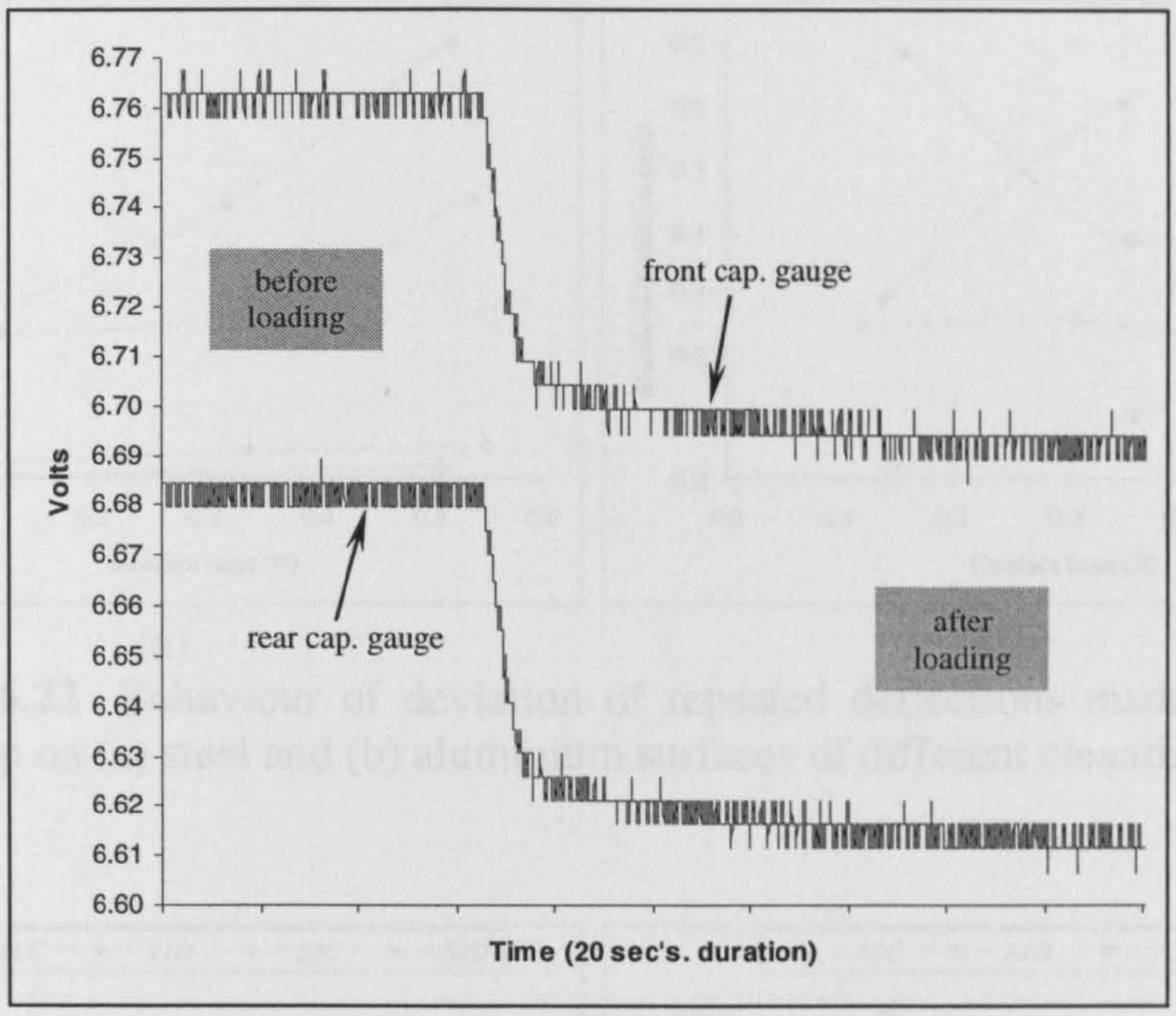
(a) (b)

Figure 6.19 Plots of average deflections of different-point repeated contact tests with a contact tip of 3 mm diameter on clean and unclean specimens of (a) mild steel and (b) aluminium surfaces of different roughness.

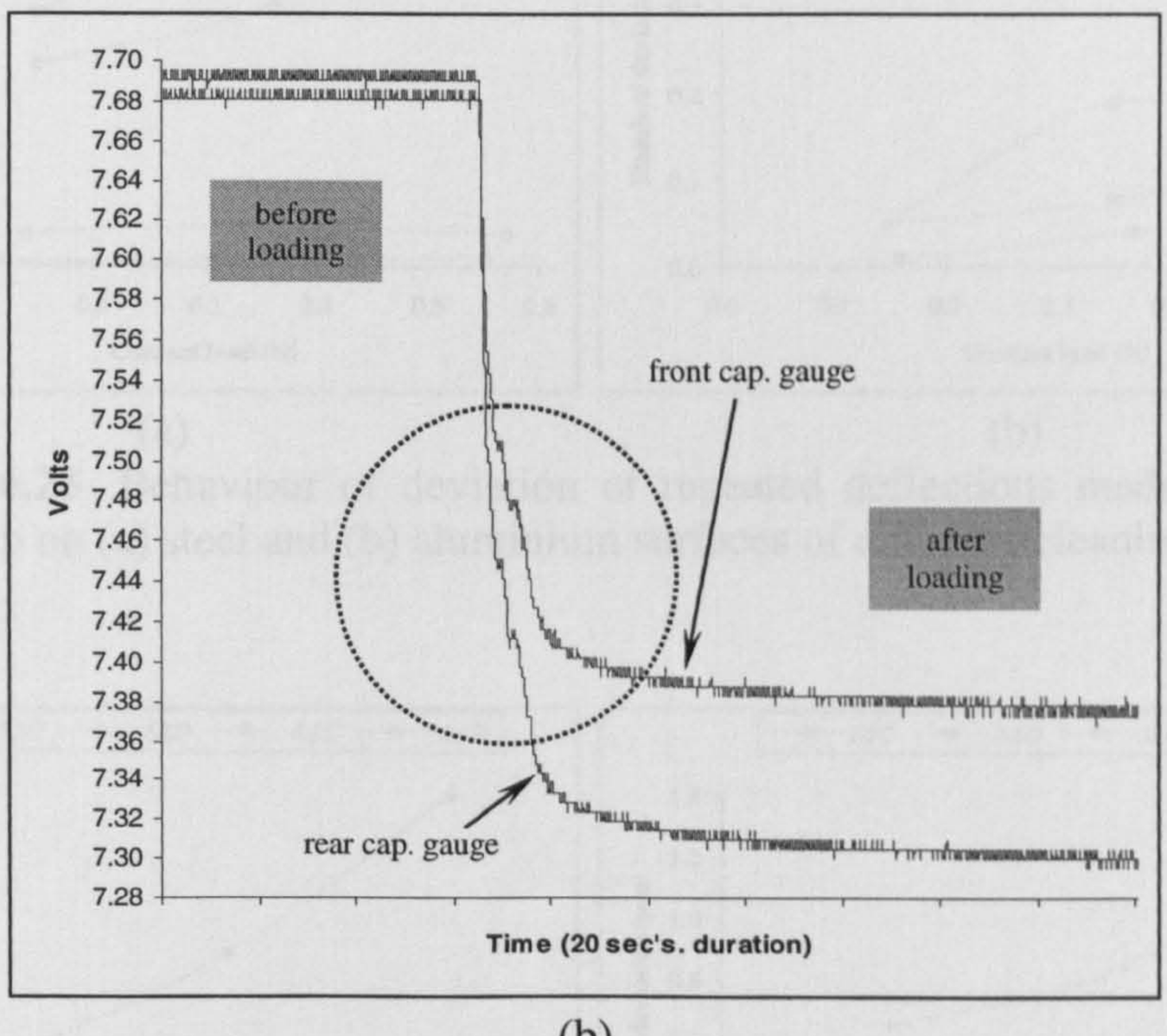


(a) (b)

Figure 6.20 Effect of E^* value on different-point repeated deflections of clean and unclean surfaces with (a) 5 mm and (b) 3 mm contact tips (Re-plotted from Figures 6.16 and 6.17, respectively).

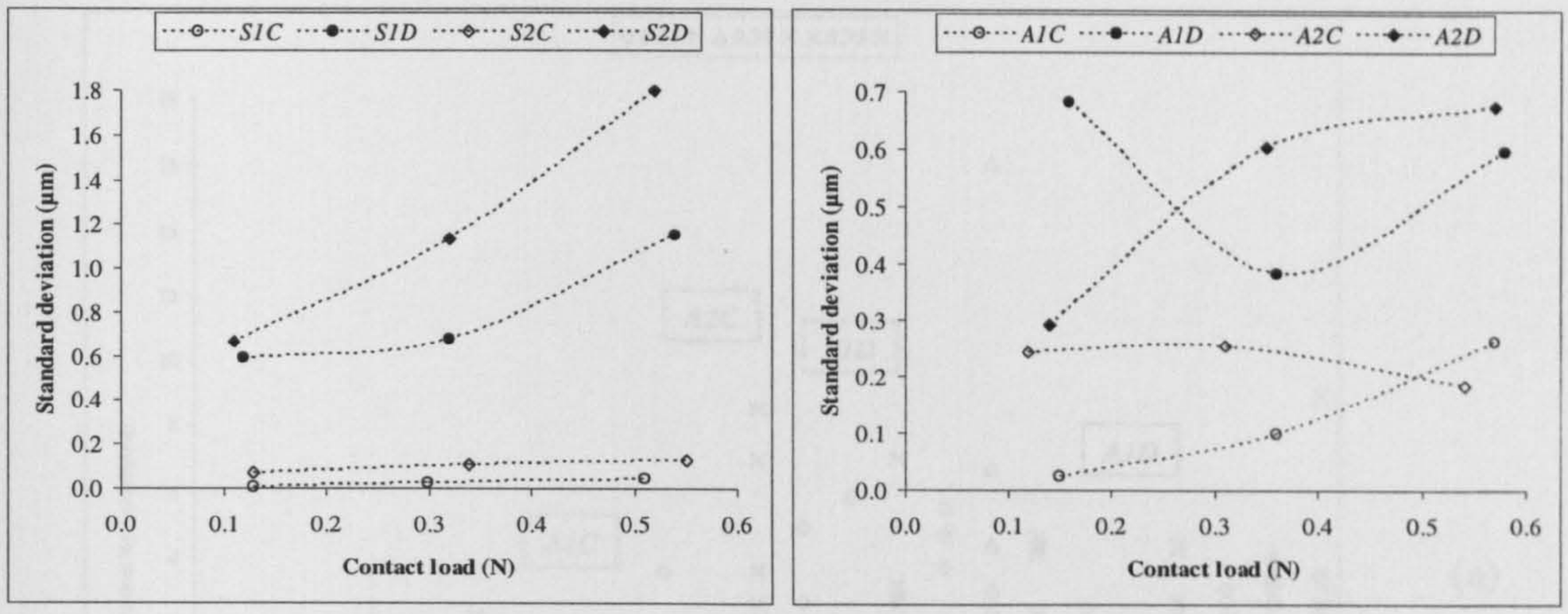


(a)



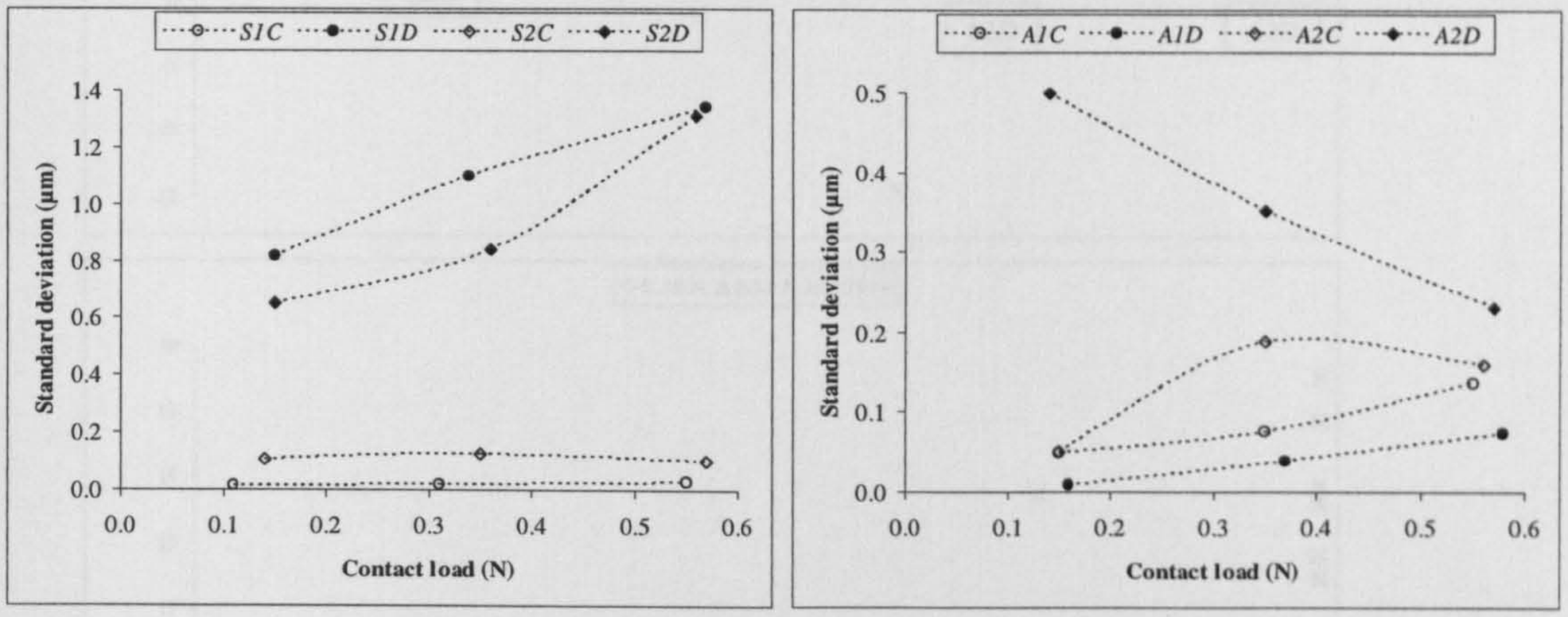
(b)

Figure 6.21 Typical behaviour of the output voltage from the two capacitive gauges during testing a point on the (a) *S1C* and (b) *S1D* surfaces with a dead-weight of 0.35 N.



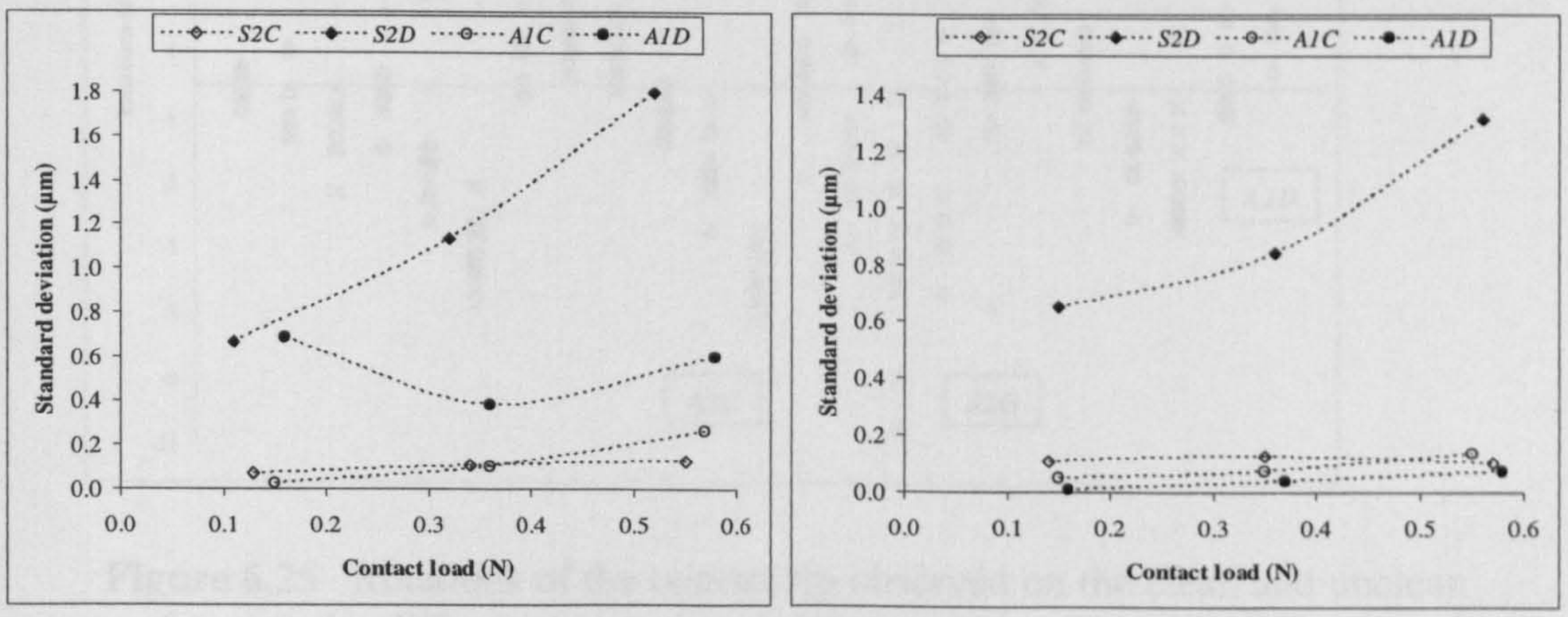
(a) (b)

Figure 6.22 Behaviour of deviation of repeated deflections made with a 5 mm tip on (a) steel and (b) aluminium surfaces of different cleanliness.



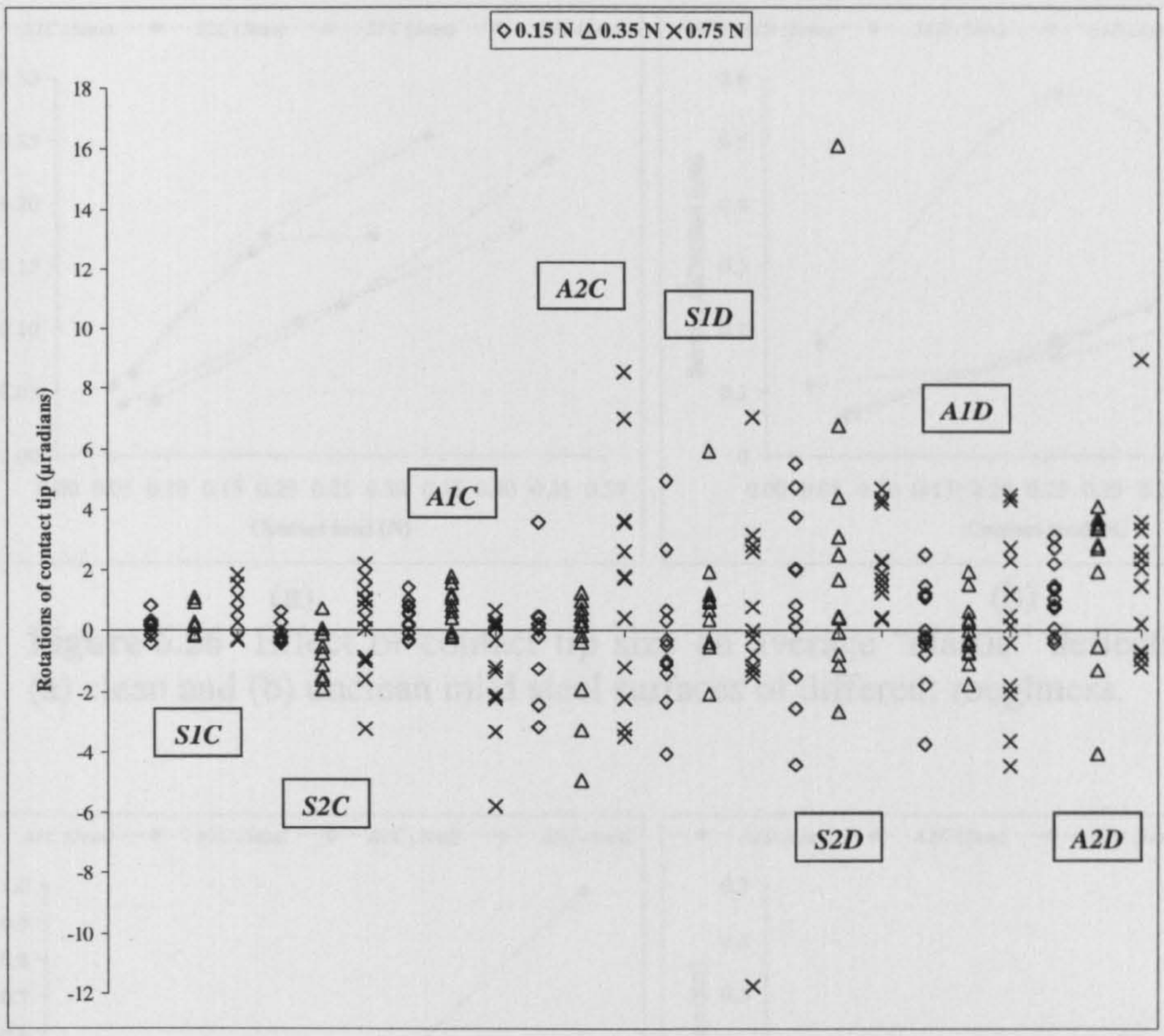
(a) (b)

Figure 6.23 Behaviour of deviation of repeated deflections made with a 3 mm tip on (a) steel and (b) aluminium surfaces of different cleanliness.

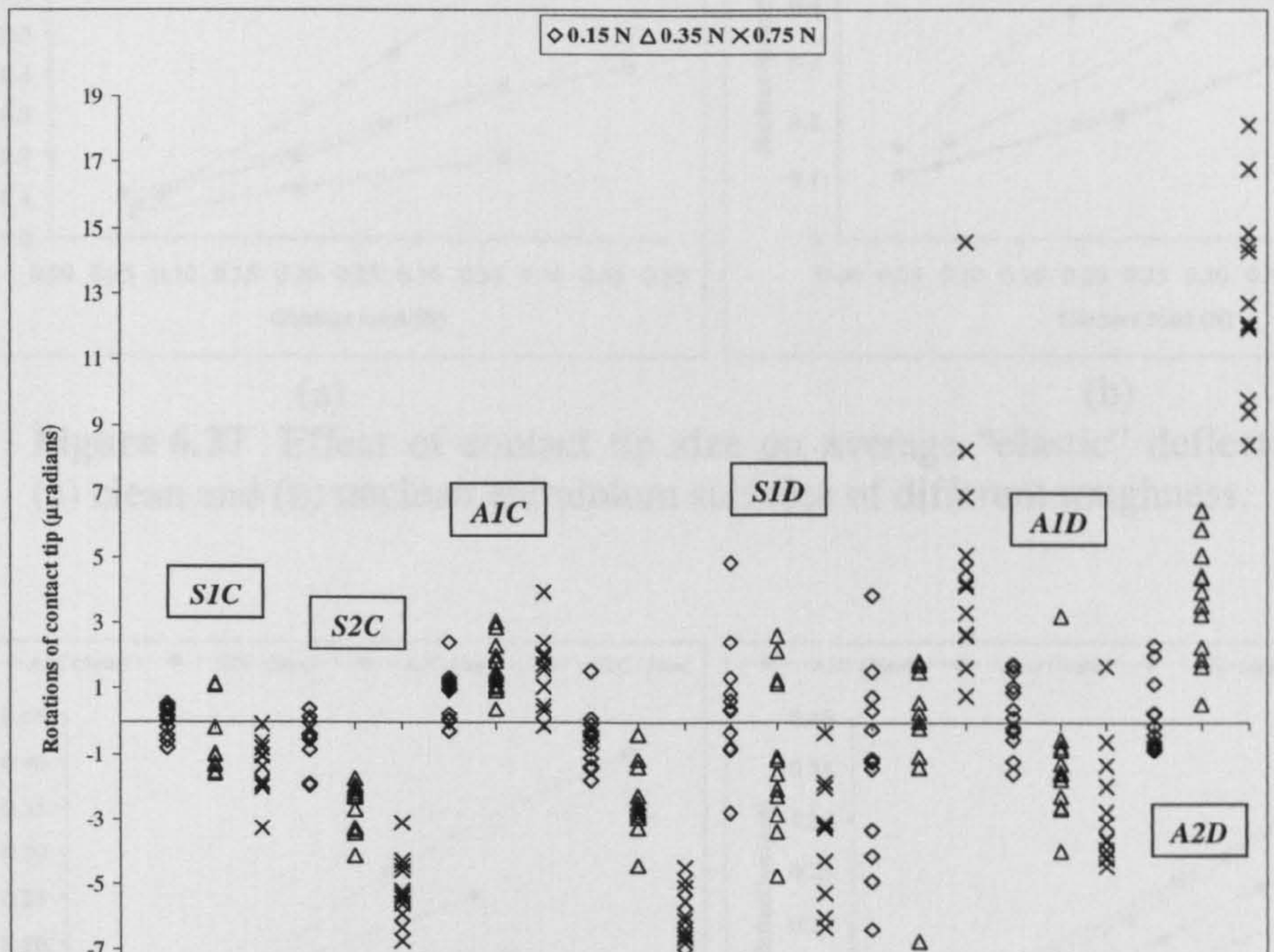


(a) (b)

Figure 6.24 Effect of E^* value on deviation of repeated deflections made with (a) 5 mm and (b) 3 mm tips on surfaces of different cleanliness.

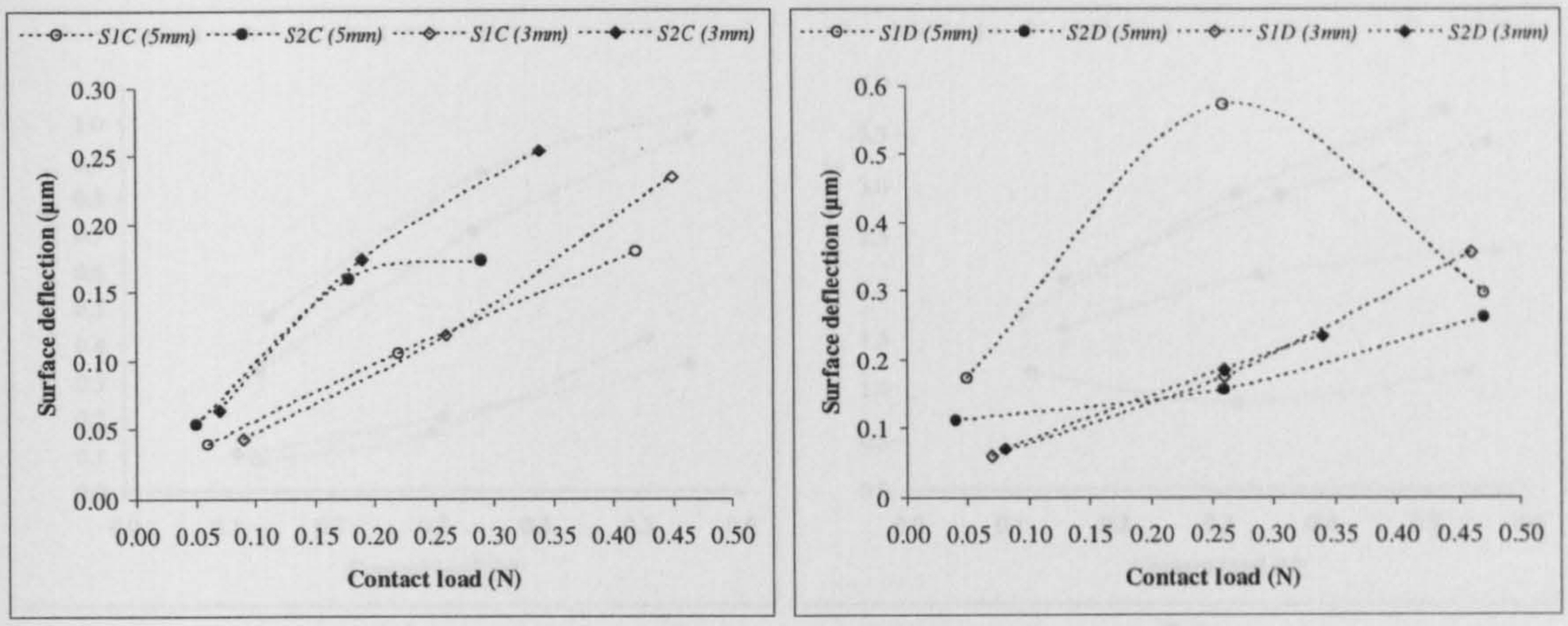


(a)



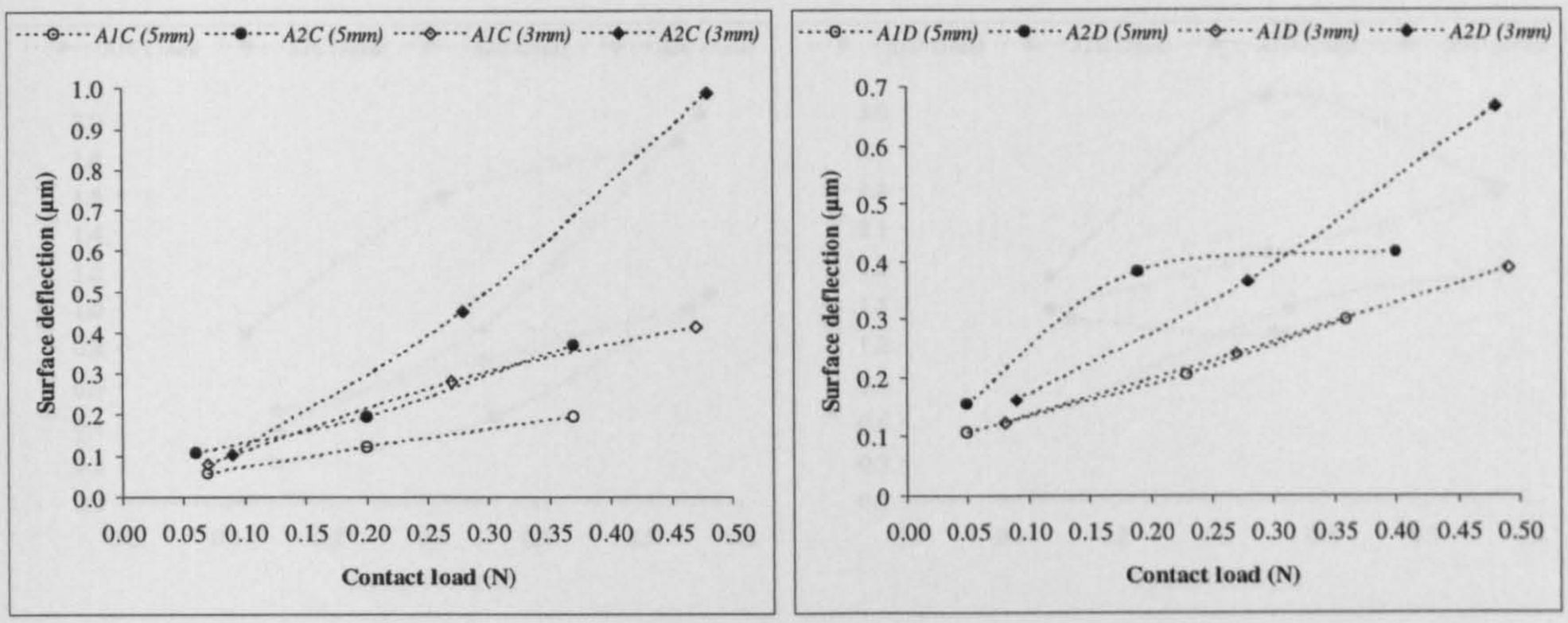
(b)

Figure 6.25 Rotations of the contact tip observed on the clean and unclean surfaces at the different-point repeated contact tests with the (a) 5 mm and (b) 3 mm contact tips.



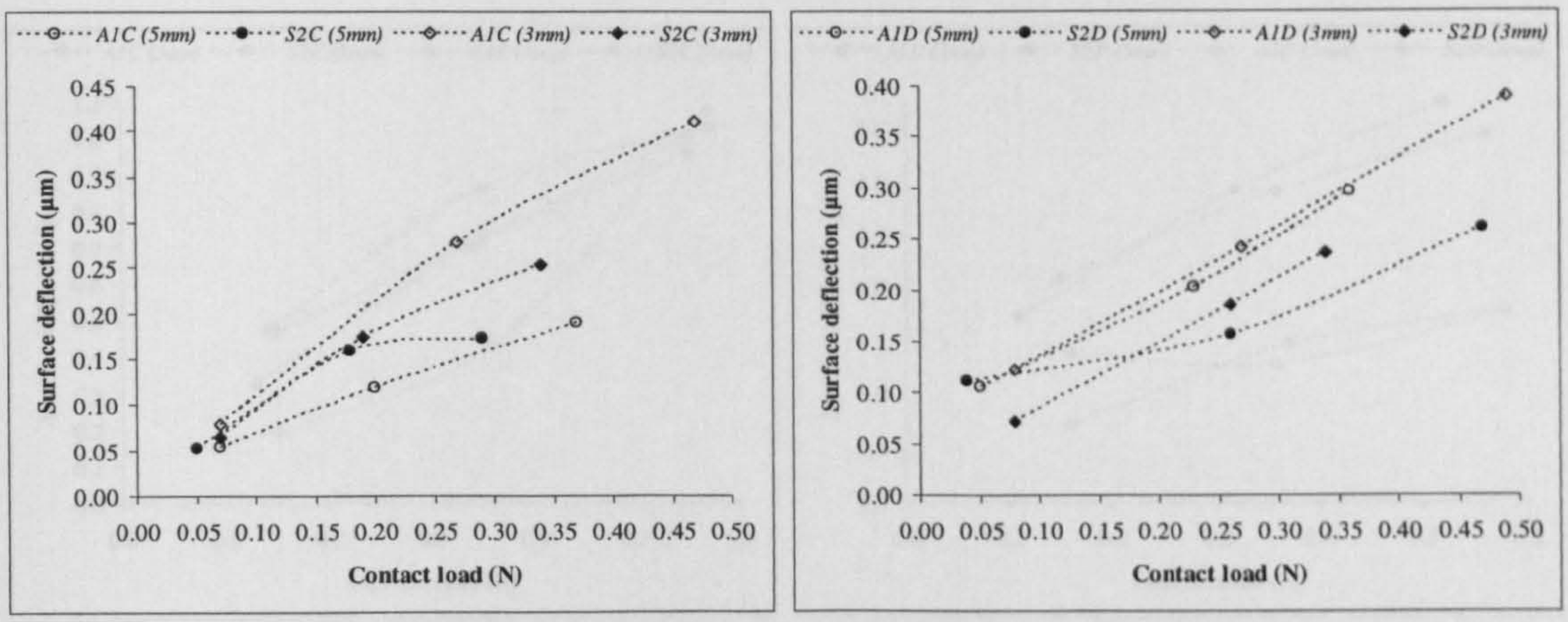
(a) (b)

Figure 6.26 Effect of contact tip size on average "elastic" deflections of (a) clean and (b) unclean mild steel surfaces of different roughness.



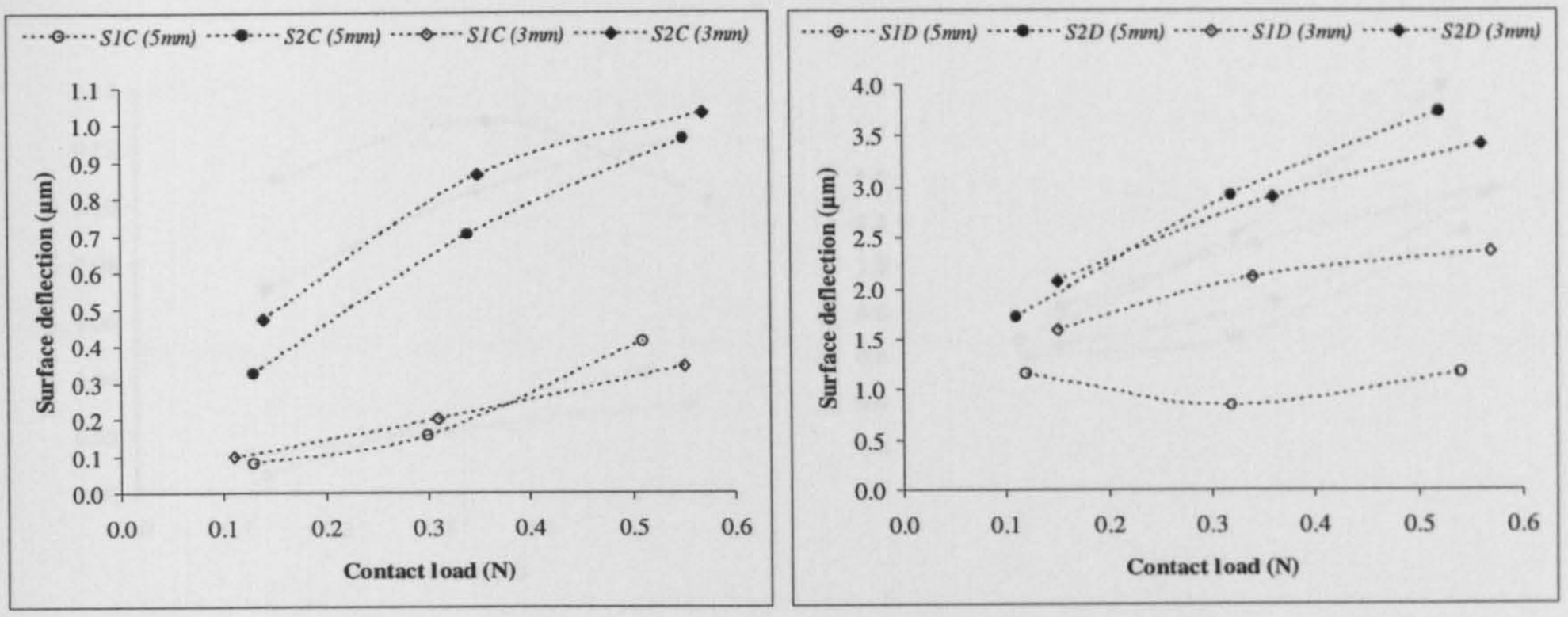
(a) (b)

Figure 6.27 Effect of contact tip size on average "elastic" deflections of (a) clean and (b) unclean aluminium surfaces of different roughness.



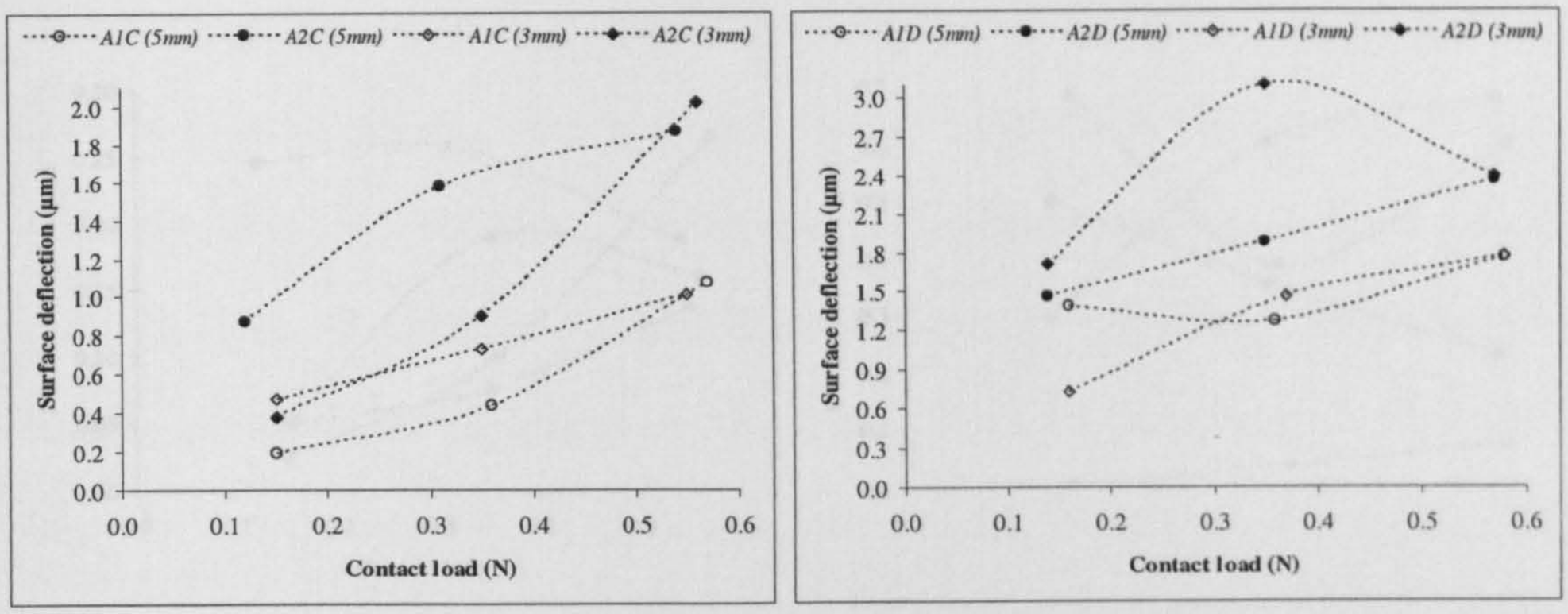
(a) (b)

Figure 6.28 Effect of contact tip size on average "elastic" deflections of (a) clean and (b) unclean surfaces of different E^* value.



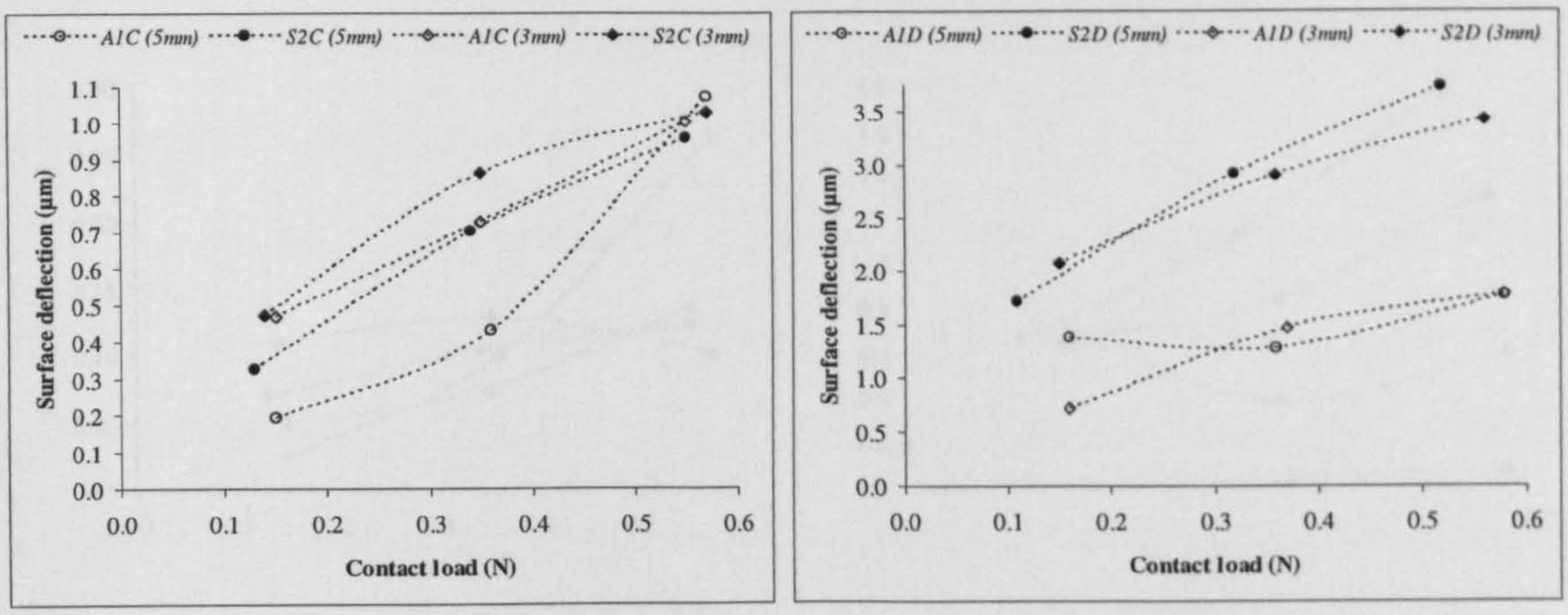
(a) (b)

Figure 6.29 Effect of contact tip size on average deflections of (a) clean and (b) unclean mild steel surfaces of different roughness.



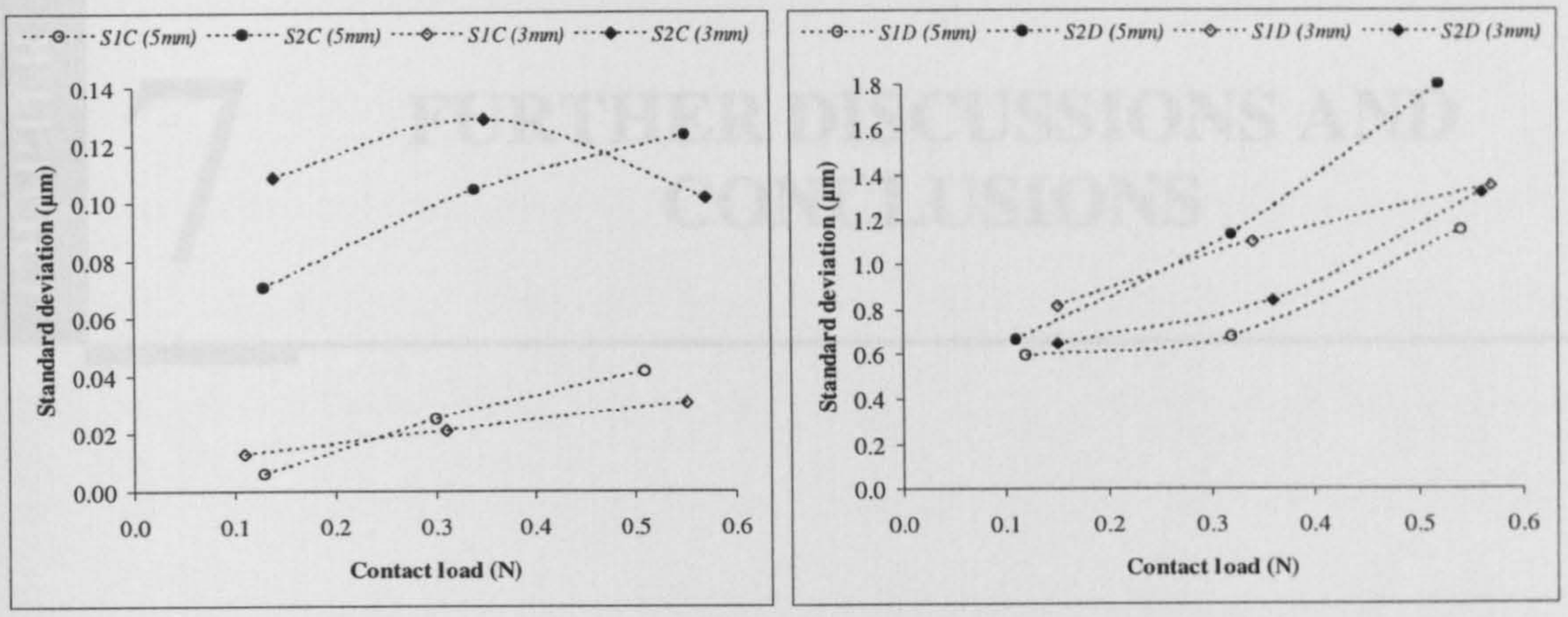
(a) (b)

Figure 6.30 Effect of contact tip size on average deflections of (a) clean and (b) unclean aluminium surfaces of different roughness.



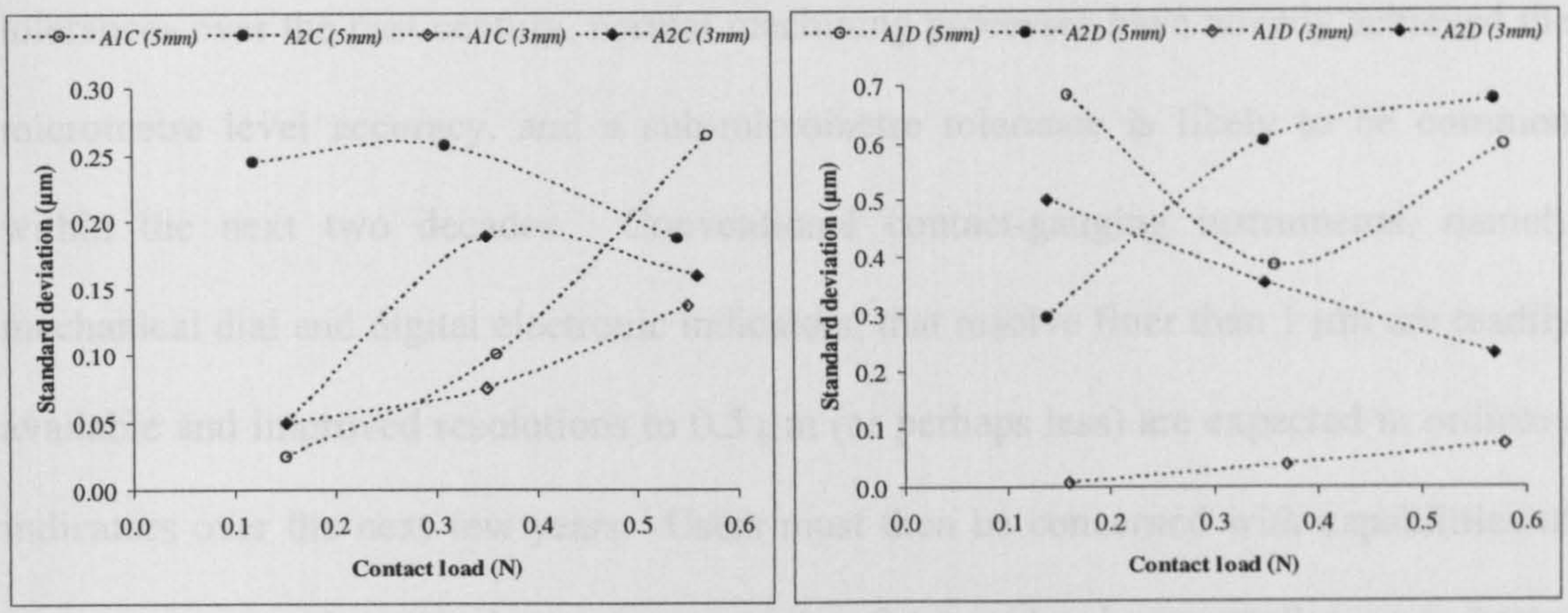
(a) (b)

Figure 6.31 Effect of contact tip size on average deflections of (a) clean and (b) unclean surfaces of different E^* value.



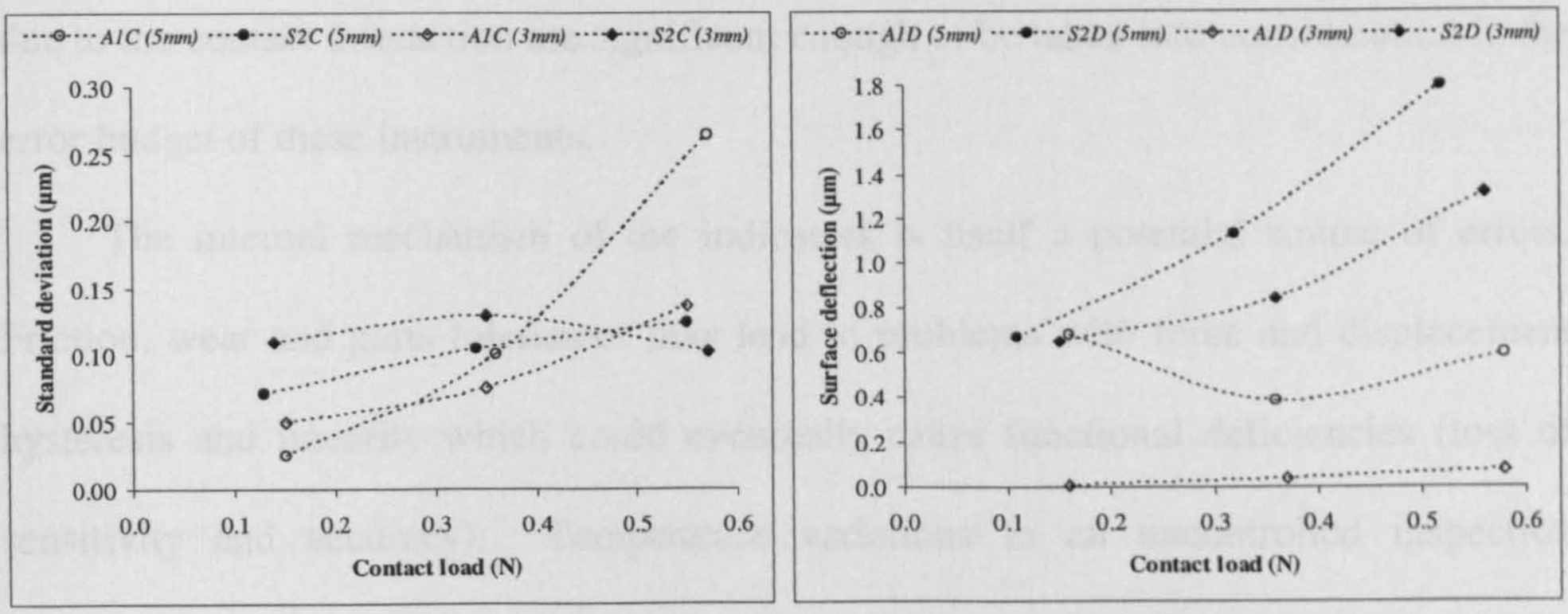
(a) (b)

Figure 6.32 Effect of contact tip size on deviations of repeated deflections of (a) clean and (b) unclean mild steel surfaces of different roughness.



(a) (b)

Figure 6.33 Effect of contact tip size on deviations of repeated deflections of (a) clean and (b) unclean aluminium surfaces of different roughness.



(a) (b)

Figure 6.34 Effect of contact tip size on deviations of repeated deflections of (a) clean and (b) unclean surfaces of different E^* value.

CHAPTER 7

FURTHER DISCUSSIONS AND CONCLUSIONS

7.1 GENERAL REMARKS

Given the history of the ever-increasing demand of modern industries for tighter tolerances over the past century, normal machining processes have already achieved the micrometre level accuracy, and a sub-micrometre tolerance is likely to be common within the next two decades. Conventional contact-gauging instruments, namely mechanical dial and digital electronic indicators, that resolve finer than 1 μm are readily available and improved resolutions to 0.5 μm (or perhaps less) are expected in ordinary indicators over the next few years. Users must then be concerned with capabilities of measurement with such instruments at the 0.1 μm level or smaller. A better understanding of the complex contact behaviour of probe tips onto engineering surfaces is needed at such a level of concern. This investigation clearly proved that uncertainties due to the contact interaction are significant enough to be taken into consideration in the error budget of these instruments.

The internal mechanism of the indicators is itself a potential source of errors. Friction, wear and parts tolerances may lead to problems with force and displacement hysteresis and linearity which could eventually cause functional deficiencies (loss of sensitivity and accuracy). Temperature variations in an uncontrolled inspection environment could also cause a serious effect not only on the indicator performance but also on the workpiece dimensional stability. Tests performed on three available high-

accuracy mechanical dial and digital electronic indicators revealed systematic effects of spring rate in both directions of the probe rod movement, beside a substantial force difference between these two directions due to the friction effect of the wiping seal. Stick-slip friction forces of the rod bearing and wiper have been observed to cause additional demonstrable variations to the gauging force. Moreover, a displacement error has been detected at, broadly, the maximum force regime of the hysteresis loop. Although the most accurate one of the three gauges tested (resolution of 1 μm) showed the smallest force uncertainties (about 0.5 N in both directions and 0.25 N in each direction), the experimental investigation of this research confirmed that these force uncertainties are capable of generating contact uncertainties well above the 100 nm level on most common conditions of engineering surfaces. Such systematic and scattering errors are expected to contribute significantly to the gauge error budget, but their analysis has received little attention from researchers compared with the exploration of advanced calibration techniques.

Through reviewing the published data of some common brands of indicators, most manufacturers have been observed not to pay much attention to providing detailed force characteristics for their best (precise) models. A few of them provide only the extreme magnitudes of the gauging force along the probe rod travel, and the others provide just a value of which it is unknown whether it represents the mean, maximum or minimum value and whether it is at the forward or backward movements of the rod. Force and displacement hysteresis can be much to the user to utilize the optimum performance of the gauge and to guide future calibration processes. In addition, the general guidelines provided by few manufacturers to users (for selecting the proper gauge for the measurement to be done) do not strongly emphasize the measuring force as a criterion for selection. Users' needs may vary beyond the time of purchase and, due to cost issues the same gauges might be used on surfaces of different properties from what was

originally planned. Based on this study and similar future ones, such guides could also include recommendations on the permissible limits of surface material and finish for each gauge model beyond which uncertainties from the contact interaction can be significant (compared to the gauge accuracy, for instance) in the error budget, which should then be evaluated and compensated.

Since what exactly happens in the contact zone between real rough surfaces is almost totally inaccessible and dramatically small relative to the component size, experimental investigations for determining the resulting behaviour of interaction between the different interacting processes (e.g., deformation, friction, wear, etc.) become questionable. So, the theoretical approaches developed to evaluate the resulting contact properties also remain, to a certain extent, doubtful as investigators try to idealize the contact behaviour through several assumptions and to avoid complicated solutions. The vast amount of these experimental and theoretical studies has been carried out by tribologists to inform future designs for better functional performance of surfaces and enhanced reliability of mechanical parts.

The numerical analyses of contact between rough surfaces provide more reliable estimates for the resulting contact behaviour than the stochastic analyses since they start from the actual topographies of surfaces; assumptions are employed in later stages to govern this behaviour under the required loading conditions. The stochastic analyses start with probabilistic assumptions for the asperity model and also employ further assumptions for the contact behaviour; hence, they are fully based on postulation. But they offer quick, computationally feasible, and convenient estimations at the level of frequent use; thus, they can readily be utilized in, for example, simple precision QC checkpoints to predict the systematic errors due to the contact between indicators' probe tips and surfaces being gauged. In both techniques, the mode on which the contact interaction takes place requires a model to be assumed (e.g., plasticity model, friction

model, etc.) that depends basically on the contact case being analysed. The contact interaction between the spherical probe tips of indicators and rough surfaces is much simpler than those which are usually considered in tribology and contact mechanics problems. In comparison with the tribological operating regimes, the contact is likely to be elastic (gauging forces are well below levels expected to cause bulk plastic yield) and frictionless (no relative movement, but the interfacial friction during deformation may exist and must be assumed of negligible effect). Anisotropy of the rough surface under measurement could also be considered an insignificant parameter since the resulting area of contact is very small (lightly loaded contacts). Even if it is analytically possible to generate a model of anisotropy for each component surface to be measured and to integrate it into the contact asperity model adopted, for an individual contact, the uncertainty from this source is still there because of the difficulty of determining exactly where the contact will take place (on which local features of an asperity group). Adhesion is unlikely to appear in such an interaction due to the roughness and imperfect cleanliness of surfaces.

This simple operating regime of contact interaction between the indicators' probe tips and components' surfaces has here led to propose approximations in two simple asperity contact models, seeking a quick and easy tool for predicting those displacement errors induced from such an interaction and contributing to measurement uncertainty. The GT (Greenwood and Tripp [88]) and KYH (Kagami, Yamada and Hatazawa [89]) models, of which each has been established based on different backgrounds, were simplified to meet the probable contact conditions of the case in concern here and compared with the experimental results. The simplification suggested in each model was the rigidity of the spherical tip (no bulk deformations from indenter since small loads are of interest here), and so only the deformation of asperities was considered. Accordingly, a simple E^* (effective Young's modulus) relation was used. The modified

models showed large discrepancies (outside ±5%) between each other at high roughness values (above 0.2 μm) and low contact loads (below 0.5 N) and at the three types of materials and contact tips used. This is as a result of a mathematical instability observed in the approximated GT model. Thus, at roughness values below 0.2 μm (a finish quality of many engineering surfaces in precision applications), the difference between the two approximations is minor within the load regimes of interest.

In comparison with the experimental results, the assumption of a rigid probe tip was justified. To a large extent, the two approximations are capable of providing conservative estimates of the surface deformation at contact loads in the range up to 2 N (a gauging force level within which most precise indicators are designed) and surface roughness below 0.2 μm. The larger results from both analyses could be used for more rigorous predictions. However, a clearer picture about such limits of use is expected if the actual roughness parameters of the surfaces tested are used in the two modified models. Systematic errors generated from the initial deformation accomplished prior to every contact have also caused some inconsistencies in the experimental results. Absolute agreement is not to be expected, also, because of the impractical assumptions made within such theoretical models.

7.2 EXPERIMENTAL RESULTS

The preliminary experimental investigation carried out on an existing test-rig raised several important issues related to the contact interaction between spherical probe tips and rough engineering surfaces at low loads. In general terms, it has showed the consistent effects of surface roughness and material on the surface deflections resulting from the interaction. Results of the deflections repeatability tests were somewhat questionable because of the different sources of error inherent in this test-rig, although

they have shown some consistent relations. The investigation has given some clues to the effects of surface cleanliness and probe tip size on the repeatabilities of deflection, but it did not reveal these effects on the deflections themselves. The test-rig overall evaluation of errors has revealed a displacement uncertainty for an individual contact test of up to 120 nm or more. However, throughout the characterisation of this test-rig and performing the experiments on it, its limited design configuration and accuracy performance have been kept in mind. This stage of the experimental programme provided adequate data needed to guide the design of the new test-rig for more reliable and comprehensive investigations.

The design and manufacture of the new test-rig of advanced accuracy and capabilities was undertaken for a detail study of the contact behaviour tips typical of precision industrial gauges (indicators). This test-rig offers greater flexibility for testing a variety of specimens of different form factors with different contact tip sizes, without altering its operating accuracies (unlike what had been encountered with the previous test-rig). It uses a more robust guided linear motion of commercial precision roller bearing slides which gives a vertically oriented straight-line approach of the contact probe onto the test surface and a linear coaxial movement of the measuring capacitive electrodes towards each other. It also incorporates active measurement of contact forces by a load-cell, easy adjustment of the useful measuring range of the capacitive gauges by a fine micrometre, and indirect stress-free clamping of specimens by gluing on glass slides. Its characterisation has shown a few possible sources of error of measurable magnitude but, in general, they have been observed to be of insignificant effect on the accuracy of measurements accomplished. The existing laboratory environment has restricted the useful resolution of single displacement measurement to be around 70 nm, although the sensitivity of the capacitive gauges is well below 20 nm.

The second stage of the experimental research has been fully accomplished on the

new test-rig, and has comprised a wider range of contact parameters to simulate real situations of industrial gauging processes. New procedures for evaluating the uncertainties of contact have been followed with improved control and accuracy of testing. Surface deflections resulting from the contact interaction have been carefully examined on surfaces of different roughnesses, material, and cleanliness using different contact loads (within the actual regime of gauging forces) and probe tip sizes. Same-point and different-point uncertainties of contact have been rigorously checked as well, against this variety of contact parameters.

Generally speaking, the investigation has confirmed obvious relations: the increase of deflection magnitudes with the increase of load or roughness or with the decrease of material elasticity or probe tip size. Surface contamination has significant unsystematic effects on these relations; especially, on the last two parameters (they dominate the effect of these two parameters). Same-point contact repeatability was affected by the change of load, probe tip size, and cleanliness; surface roughness and material have shown no clear effects on this repeatability. Different-point contact repeatability was influenced by all of these parameters. The roughness variability (variation of local asperities height and distribution) across the same surface was observed to cause some distinct inconsistencies in the deflection and repeatability results; consequently, no clear correlations have been found from the available data. Probe tip rotation generally increased with the increase of load and the decrease of probe tip size in the same-point repeated contacts, and was not observed to be affected by the surface roughness, material, or contamination. Although there were large inconsistent effects from the roughness variability on the probe tip rotations, in the different-point repeated contacts, the rotations have shown a plausible response to each of these effects.

Based on the experimental results of the research, gauging systematic errors due

to the contact interaction between indicators' probe tips of 5 mm diameter and steel and aluminium clean surfaces of roughness values below 0.5 μm are expected to approach 0.5 μm and 0.8 μm, respectively, at gauging forces only below 0.5 N. With contact tips of 3 mm diameter (which most indicators' probes are usually equipped with), such errors can reach 0.7 μm and 1.0 μm on these two materials. At these contact conditions, scattering errors due to, mainly, the surface roughness variability have been seen within 0.2 μm. As the two approximations would estimate, surface deflections can approach 1.5 μm on these surfaces with gauging forces near 2.0 N. When re-contacting the surface at nominally the same point for measurement, the transmitted systematic errors will be largely less; a minimum reduction of around 60% has been noted on the four deflection magnitudes quoted above. Obviously, these errors will be greater if rougher and/or softer surfaces are to be gauged or smaller probe tips are used.

The contact interaction errors and uncertainties on the unclean surfaces are much worst and unpredictable as a result of the random distribution of the loose dirt particles across the rough surface. Tests on surfaces left without cleaning after machining have shown considerably larger inconsistencies in the relations of deflection with the other contact parameters, in addition to significantly smaller repeatabilities. Changes in roughness and/or probe tip size have shown obvious effects on these surfaces, unlike the change in E^* which has been of unclear effect. However, these tests have resulted in deflections highly scattered within about 4.0 μm with the 3 mm probe tip and loads below 0.5 N on unclean surfaces of roughnesses below 0.5 μm. Re-establishing the contact with the same force on nominally the same point on the unclean surface is expected to contribute much lower errors to measurements, as the initial contact force are likely to "clear" the contact area by rearranging the loose features and deforming them together with highest surface asperities. The tests have revealed deflections of higher magnitudes but very near to those observed on the clean surfaces.

The above magnitudes of systematic and scattering errors of contact suggest that caution is needed when precision gauging processes are considered, as they are within the range of operating accuracy of many current precise indicators. Accordingly, we can recommend that systematic errors must be evaluated (or estimated from theoretical models) and compensated form the measurement, if a micrometre level accuracy or below is of interest in gauging engineering surfaces. Evaluating or estimating scattering errors is problematic, as they depend greatly on the inevitable roughness variability created by the complex interaction of parameters during machining the surface. It can also be recommended to repeat the contact of probe tip with the surface before taking the indicator reading, since the contributed errors will be dramatically minimized. Such a method of measurement may necessitate bringing the indicator probe tip (while the gauge head is fixed) onto final contact with the surface in the outward movement of the probe, which would also provide reduced gauging forces and systematic errors. However, the experimental investigations have proved that this method will also be effective in minimizing such errors on the unclean surfaces to near the magnitudes of those on the clean surfaces.

The preliminary recommendation from this work is that users unwilling to attempt a detailed, specific analysis should allow a type B uncertainty of, at least, 1 μm (95%) if systematic compression is not compensated and a further 0.4 μm (95%) type A uncertainty to account for roughness induced variability.

7.3 RECOMMENDATIONS FOR FUTURE WORK

The work discussed in this thesis represents an attempt to highlight new insights into uncertainty in industrial precision gauging, as it concludes that the contact cannot any

longer be assumed as a negligible source of errors. Greater emphases on this topic are expected in the near future due to the increasing demand for tighter tolerances of components and higher accuracies of measurements. Comprehensive investigations are then needed with a wider range of contact parameters and a better overall performance of test setups than what has been dealt with in this work. However, this investigation stimulates further work that could be carried out using even more advanced test capabilities and improved evaluation and prediction of errors. Some general ideas of relevance are noted below.

The performance limitations recognized at the characterization and operation of the new test-rig could be evaluated and minimized in order to enhance the existing operating accuracy; refer to Subsection 5.5.3. These limitations are much related to the test-rig design; hence, improvements could reflect a significant difference in performance. Shielding the rear surfaces of the glass plates of the capacitive gauge electrodes could greatly reduce environmental noise and offer better measurement resolutions; an appropriate coating could be used for this purpose. A balanced pull of the carriage weight can be recommended to minimize rotations along the travel range; although such errors are of negligible effect over the very small carriage movements required during the test, achieving this modification could provide more stability and repeatability to the probe approach everywhere along the carriage travel. As the stick-slip friction of the slides' bearing affects the amount of initial load applied which could lead to different systematic errors from one contact to another, a further modification can also be recommended here to offer better control of such a load.

Such studies are likely to be in need to establish a widely acceptable definition of "first contact" and its detection.

To a large extent, the results of this work appear encouraging enough to set up similar investigations on the contact errors of touch-trigger probes of co-ordinate

measuring machines. Although the gauging forces applied by these probes are generally below those applied by the indicators' probes, the main concern in the measurement with such instruments is the potential dynamic effect. Tips of these probes usually approach the workpiece surface at significant speeds, which would lead to concern that the impact of the tip on the surface could induce significant errors. This measurement will require a rather different test-rig to that used in the present work.

Based on the concerns highlighted in this investigation, rigorous evaluations of theoretical contact models could be carried out in conjunction with extensive experimental investigations in order to provide industry with information, perhaps through standard calibration tables (or charts), for compensating both systematic and scattering errors from contact.

To define the accuracy of a position measurement, there must be a model that proposes where the ideal position would be. In the case of a rough surface, tests such as those reported here can reveal the random uncertainty in multiple readings but do not address the question of where the tip rests on average compared to the "real" position of the surface. In fact, there is no consensus on what level should be used for assessment: the mean (50% material) level might be logical geometrically, but the height of the functional surface could well be higher (e.g. many mechanical contacts) or lower (e.g. some electrical phenomena). Further theoretical and experimental studies are urgently required to establish widely acceptable definitions, and detection methods, for terms such as "effective surface position" and "first contact".

With such additional work should come enough high quality data to provide guidelines on uncertainty of practical use to the international, industrial metrology community.

REFERENCES

[1] McKeown, P. A. (1996), "From micro- to nano-machining – towards the nanometre era", *Sensor Review* **16** (2) 4-10

[2] McKeown, P. A. (1987), "The role of precision engineering in manufacturing of the future", *Annals of the CIRP* **36** (2) 495-501

[3] McKeown, P. A. (2000), "The ultra precision technologies – from micro to nanotechnologies: a glimpse of the future", in: *Proc. of the euspen International Seminar - Warwick University - 18th/19th Sept. 2000*, ISBN 1861940637, 9-37

[4] Evans, C. J. (1989), *Precision Engineering: An Evolutionary View*, Cranfield Press, Cranfield

[5] Taniguchi, N. (1983), "Current status in, and future trends of, ultraprecision machining and ultrafine materials processing", *Annals of the CIRP* **32** (2) 573-582

[6] Taniguchi, N. (Ed.) (1996), *Nanotechnology – Integrated Processing Systems for Ultra-precision and Ultra-fine Products*, Oxford University Press, Oxford

[7] Whitehouse, D. J. (1991), "Nanotechnology instrumentation", *Measurement and Control* **24** 37-46

[8] Taniguchi, N. (1985), "Atomic bit machining by energy beam processes", *Precision Engineering* **7** (3) 145-146

[9] Christy, D., Schuetz, G., Tabenkin, A., and Tullar, P. (Apr. 1998), "The future of gaging", *Quality Digest*

[10] Schuetz, G. (Apr. 1996), "Gages: the hand tolls of quality", *Quality Digest*

[11] Groover, M. P. (1996), *Fundamentals of Modern Manufacturing: Materials, Processes, and Systems*, Prentice-Hall Inc., New Jersey

[12] Kibbe, R. R., Neely, J. E., Meyer, R. O., and White, W. T. (1995), *Machine Tool Practices*, Prentice-Hall Inc., New Jersey

[13] Dallas, D. B. (1976), *Tool and Manufacturing Engineers Handbook*, Society of Manufacturing Engineers, Dearborn

[14] Reid, D. T. (1991), *Fundamentals of Tool Design*, Society of Manufacturing Engineers, Dearborn

[15] Pike, S. (Feb. 1998), "Metrology toolbox: Choosing and using dial indicators", *Quality*

[16] British Standard 907:1965, *Specification for Dial Gauges for Linear Measurement*, BSI, London

[17] British Standard 2795:1981, *Specification for Dial Test Indicators (lever type) for Linear Measurement*, BSI, London

[18] Webpage: http://www.fedgage.com/selecting_guide.htm (*11/11/2002*)

[19] Webpage: http://www.mitcat.com/r02-3/index.htm (*11/11/2002*)

[20] Scriff, J. (1993), "Update on hand-held, dimensional, and variable gauges", *Quality Progress* **26** (4) 87-89

[21] Minner, J. F. (Jun. 1996), "Get inside your indicator needs", *Quality*

[22] Stewart, T. (Apr. 1998), "Selecting the right gage", *Quality Digest*

[23] Koppelmann, D. (Sep. 1996), "GageGuide: Digital or dial indicators?", *Quality Digest*

[24] Anonymous (1979), "Electronic gaging systems simplify complex measurements", *Tooling & Production* **45** (3) 96-98

[25] Webpage: http://www.mitcat.com/r04-8/index_3.htm (*11/11/2002*)

[26] National Physical Laboratory (1967), *Notes on Applied Science No. 5: Gauge Making and Measuring*, H. M. S. O., London

[27] Heske, W. A. and Johnson A. W. (1969), "Digital pressure indicator", *Instruments & Control Systems* **42** (12)
[28] Farago, F. T. and Curtis, M. A. (1994), *Handbook of Dimensional Metrology*, Industrial Press Inc., New York
[29] Sporea, D. G. and Miron, N. (1996), "Dial indicators checking-up by laser interferometry", *Review of Scientific Instruments* **67** 612-614
[30] Hemming, B. and Lehto, H. (2002), "Calibration of dial indicators using machine vision", *Measurement Science & Technology* **13** 45-49
[31] Slocum, A. H. (1992), *Precision Machine Design*, Prentice-Hall Inc., New Jersey
[32] Flack, D. (2001), *Measurement Good Practice Guide No. 43: CMM Probing*, National Physical Laboratory, Middlesex
[33] User's Manual: "ID-F125/150 Digimatic Indicator", No. 99MAH001B, Mitutoyo Corporation, Japan
[34] System Reference Manual: Houndsfield Test Equipment, Issue 02, 1997, HTE, England
[35] Webpage: http://www.dialindicator.com/diagram.html (*11/11/2002*)
[36] Catalogue: "Instruments and systems for quality assurance", Issue 1986, TESA SA, Switzerland
[37] Website: http://www.fedgage.com (*11/11/2002*)
[38] Catalogue: "Measurement technology", No. U397, 1997, Mitutoyo (UK) Ltd.
[39] Webpage: http://www.mahr.com/en/content/products/handtools/indicator/overview.html (*11/11/2002*)
[40] Website: http://www.mitcat.com (*11/11/2002*)
[41] Webpage: http://www.rosebank-eng.com.au/tesa_index.htm (*11/11/2002*)
[42] Catalogue: "Heidenhain-Metro digital length gauges", 1981, Dr. Johannes Heidenhain GmbH, Germany
[43] Whitehouse, D. J. (1994), *Handbook of Surface Metrology*, Institute of Physics Publishing Ltd., Bristol
[44] Johnson, K. L. (1985), *Contact Mechanics*, University Press, Cambridge
[45] Fischer-Cripps, A. C. (1999), "The Hertzian contact surface", *Journal of Materials Science* **34** 129-137
[46] Johnson, K. L. (1982), "One hundred years of Hertz contact", *Proceedings of the Institution of Mechanical Engineers* **196** 363-378
[47] Smith, S. T. and Chetwynd, D. G. (1994), *Foundation of Ultraprecision Mechanical Design*, Gordon and Breach Science Publishers S. A., Switzerland
[48] Sackfield, A. and Hills, D. A. (1983), "Some useful results in the classical Hertz contact problem", *Journal of Strain Analysis* **18** 101-105
[49] Fischer-Cripps, A. C. (2000), *Introduction to Contact Mechanics*, Springer-Verlag Inc., New York
[50] Yoffe, E. H. (1984), "Modified Hertz theory for spherical indentation", *Philosophical Magazine A* **50** 813-828
[51] Greenwood, J. A. (1997), "Analysis of elliptical Hertzian contacts", *Tribology International* **30** 235-237
[52] Poon, C. Y. and Sayles, R. S. (1994), "Contact analysis of a smooth ball on an anisotropic rough surface", *Journal of Tribology – Trans. of the ASME* **116** 850-859
[53] Björklund, S. (2001), "The influence of surface roughness in elliptical contacts", *Tribology International* **34** 841-845

[54] Poon, C. Y. and Sayles, R. S. (1994), "Numerical contact model of a smooth ball on an anisotropic rough surface", *Journal of Tribology – Trans. of the ASME* **116** 194-201
[55] Liang, X. and Linqing, Z. (1991), "A numerical model for the elastic contact of three-dimensional real rough surfaces", *Wear* **148** 91-100
[56] Greenwood, J. A. and Williamson, J. B. P. (1966), "Contact of nominally flat surfaces", *Proceedings of the Royal Society* (London) **A295** 300-319
[57] Chang, W. R., Etsion, I., and Bogy, D. B. (1987), "An elastic-plastic model for the contact of rough surfaces", *Journal of Tribology – Trans. of the ASME* **109** 257-263
[58] Sayles, R. S. (1996), "Basic principles of rough surface contact analysis using numerical methods", *Tribology International* **29** 639-650
[59] Webster, M. N. and Sayles, R. S. (1986), "A numerical model for the elastic frictionless contact of real rough surfaces", *Journal of Tribology – Trans. of the ASME* **108** 314-320
[60] Lee-Prudhoe, I., Sayles, R. S., and Kaderic, A. (1999), "Investigations into asperity persistence in heavily loaded contacts", *Journal of Tribology – Trans. of the ASME* **121** 441-448
[61] Greenwood, J. A. (1967), "The area of contact between rough surfaces and flats", *Journal of Lubrication Technology – Trans. of the ASME* **?** 81-91
[62] Nayak, P. R. (1973), "Random process model of rough surfaces in plastic contact", *Wear* **26** 305-333
[63] Pullen, J. and Williamson, J. B. P. (1972), "On the plastic contact of rough surfaces", *Proceedings of the Royal Society* (London) **A327** 159-173
[64] Childs, T. H. C. (1977), "The persistence of roughness between surfaces in static contact", *Proceedings of the Royal Society* (London) **A353** 35-53
[65] Whitehouse, D. J. and Archard, J. F. (1970), "The properties of random surfaces of significance in their contact", *Proceedings of the Royal Society* (London) **A316** 97-121
[66] Bush, A. W., Gibson, R. D., and Keogh, G. P. (1976), "The limit of elastic deformation in the contact of rough surfaces", *Mech. Res. Comm.* **3** 169-174
[67] Francis, H. A. (1977), "Application of spherical indentation mechanics to reversible and irreversible contact between rough surfaces", *Wear* **45** 221-269
[68] Ishigaki, H., Kawaguchi, I., and Mizuta, S. (1979) "A simple estimation of the elastic-plastic deformation of contacting asperities", *Wear* **54** 157-164
[69] Thomas, T. R. (1999), *Rough Surfaces*, Imperial College Press, London
[70] So, H. and Liu, D. C. (1991), "An elastic-plastic model for the contact of anisotropic rough surfaces", *Wear* **146** 201-218
[71] Bush, A. W., Gibson, R. D., and Thomas, T. R. (1975), "The elastic contact of a rough surface", *Wear* **35** 87-111
[72] Bush, A. W., Gibson, R. D., and Keogh, G. P. (1979), "Strongly anisotropic rough surfaces", *Journal of Lubrication Technology – Trans. of the ASME* **101** 15-20
[73] Yu, M. M. H. and Bhushan, B. (1996), "Contact analysis of three-dimensional rough surfaces under frictionless and frictional contact", *Wear* **200** 265-280
[74] Yan, W. and Komvopoulos, K. (1998), "Contact analysis of elastic-plastic fractal surfaces", *Journal of Applied Physics* **84** 3617-3624
[75] Chang, L. and Gao, Y. (1999), "A simple numerical method for contact analysis of rough surfaces", *Journal of Tribology – Trans. of the ASME* **121** 425-432
[76] Maugis, D. (1996), "On the contact and adhesion of rough surfaces", *Journal of Adhesion Science and Technology* **10** 161-175

[77] Johnson, K. L., Kendall, K., and Roberts, A. D. (1971), "Surface energy and the contact of elastic solids", *Proc. Roy. Soc. Lond.* **A324** 301-313
[78] Derjaguin, B. V., Muller, V. M., and Toporov, Y. P. (1975), "Effect of contact deformation on the adhesion of particles", *Journal of Colloid and Interface Science* **53** 314-326
[79] Johnson, K. L. (1998), "Mechanics of adhesion", *Tribology International* **31** 413-418
[80] Maugis, D. (2000), *Contact, Adhesion and Rupture of Elastic Solids*, Springer-Verlag Inc., New York
[81] Johnson, K. L. and Greenwood, J. A. (1997), "An adhesion map for the contact of elastic spheres", *Journal of Colloid and Interface Science* **192** 326-333
[82] Poon, C. Y. and Sayles, R. S. (1992), "The classification of rough surface contacts in relation to tribology", *Journal of Applied Physics D* **25** A249-A256
[83] Whitehouse, D. J. (1998), "Surfaces: an essential link in nanotechnology", *Nanotechnology* **9** 113-117
[84] Whitehouse, D. J. (1997), "Surface metrology" – Review Article, *Measurement Science and Technology* **8** 955-972
[85] Sayles, R. S., deSilva, G. M. S., Leather, J. A., Anderson, J. C., and Macpherson, P. B. (1981), "Elastic conformity in Hertzian contacts", *Tribology International* **?** 315-322
[86] Chilamakuri, S. K. and Bhushan, B. (1998), "Contact analysis of non-Gaussian random surfaces", *Proc. Instn. Mech. Engrs. J* **212** 19-32
[87] Yongsheng, L., Guiping, Y., Yan, H., and Linqing, Z. (1996), "The characteristics of elastically contacting ideal rough surfaces", *Journal of Tribology – Trans. of the ASME* **118** 90-97
[88] Greenwood, J. A. and Tripp, J. H. (1967), "The elastic contact of rough spheres", *Journal of Applied Mechanics – Trans. of the ASME* **34** 153-159
[89] Kagami, J., Yamada, K., and Hatazawa, T. (1983), "Contact between a sphere and rough plates", *Wear* **87** 93-105
[90] Burden, R. L. and Faires, J. D. (1997), *Numerical Analysis*, Brooks/Cole Publishing Co., Pacific Grove
[91] Communication (Nov. 2000): Prof. J. Kagami (first author of ref. 89 above)
[92] Barnes, J. W. (1994), *Statistical Analysis for Engineers and Scientists – A Computer-based Approach*, McGraw-Hill Inc., New York
[93] Chapra, S. C. and Canale, R. P. (1998), *Numerical Methods for Engineers with Programming and Software Applications*, WCB/McGraw-Hill, Boston
[94] Timoshenko, S. and Goodier, J. N. (1951), *Theory of Elasticity*, McGraw-Hill, New York
[95] Cartwright, A. J. (Ed.) (1998), *Engineering Data Book*, University of Warwick, Coventry
[96] Greenwood, J. A., Johnson, K. L., and Matsubara, E. (1984), "A surface roughness parameter in Hertz contact", *Wear* **100** 47-57
[97] Chetwynd, D. G. and Davis, L. A. J. (1996), "Stylus drag forces on lubricated surfaces", in: *Proc. ASPE* **14** 141-145
[98] Whitehouse, D. J. (1987), "Surface metrology instrumentation", *J. of Phys. E: Sci. Instrum.* **20** 1145-1155
[99] Mainsah, E., Greenwood, J. A., and Chetwynd, D. J. (Eds.) (2001), *Metrology and Properties of Engineering Surfaces*, Kluwer Academic Publishers, Boston
[100] Griffiths, B. J. (2001), *Manufacturing Surface Technology – Surface Integrity & Functional Performance*, Penton Press, London

APPENDIX A

ROUTINES FOR SOLVING APPROXIMATIONS SUGGESTED IN TWO CONTACT MODELS WRITTEN IN "MATHCAD® 2000 PROFESSIONAL" FORMATS

A1
THE GT CONTACT MODEL

Constants:

Contact load (N): $P := 3$

Radius of sphere (mm): $B := 2.5$

Young's Modulus of sphere material (N/mm^2): $E_1 := 2.10 \cdot 10^5$

Poisson's Ratio of sphere material: $\nu_1 := 0.292$

Young's Modulus of rough surface material (N/mm^2): $E_2 := 2.10 \cdot 10^5$

Poisson's Ratio of rough surface material: $\nu_2 := 0.292$

Standard deviation of asperity heights of rough surface (mm): $\sigma := 1.450 \cdot 10^{-3}$

Number of asperities of rough surface per mm^2: $\eta := 1430$

Radius of asperity summit of rough surface (mm): $\beta := 28.0 \cdot 10^{-3}$

Modulus of contact (N/mm^2), assuming a rigid smooth sphere:

$$E := \frac{E_2}{1 - \nu_2^{\,2}} \qquad E = 2.296 \times 10^5$$

Governing equations:

$$\mu := \frac{8}{3} \cdot \eta \cdot \sigma \cdot \sqrt{2 \cdot B \cdot \beta} \qquad u_s(\rho, d_s) := d_s + \rho^2$$

$$p_s(\rho, d_s) := \frac{\mu}{\sqrt{2 \cdot \pi}} \cdot \int_{u_s(\rho, d_s)}^{\infty} \left(s - u_s(\rho, d_s)\right)^{\frac{3}{2}} \cdot e^{-\frac{s^2}{2}}\, ds$$

$$T_1(d_s) := 2 \cdot \pi \cdot \int_0^{\infty} \rho \cdot p_s(\rho, d_s)\, d\rho \qquad T_2 := \frac{2 \cdot P}{\sigma \cdot E \cdot \sqrt{2 \cdot B \cdot \sigma}}$$

$$f(d_s) := T_1(d_s) - T_2$$

A1 Continued

Iterative solution:

$Tol := 1 \cdot 10^{-12}$ $\quad$ $Maxiter := 50$

$$Ans(d_{si}) := \left| \begin{array}{l} Iter \leftarrow 1 \\ \text{while } Iter \le Maxiter \\ \quad \left| \begin{array}{l} d_s \leftarrow d_{si} - \dfrac{f(d_{si})}{\dfrac{d}{dd_{si}} f(d_{si})} \\ \text{break if } |d_s - d_{si}| < Tol \\ d_{si} \leftarrow d_s \\ Iter \leftarrow Iter + 1 \end{array} \right. \\ \begin{pmatrix} d_s \\ Iter \end{pmatrix} \end{array} \right.$$

Initial guess: $\quad$ $M := Ans(0)$

ANSWER:

$$M = \begin{pmatrix} 0.996 \\ 7 \end{pmatrix}$$

Dimensionless separation from the mean-line of the rough surface :

$d_s := M \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $\quad$ $d_s = 0.996$

Separation from the mean-line of the rough surface (mm):

$d := d_s \cdot \sigma$ $\quad$ $d = 1.444 \times 10^{-3}$

Asperity deformation at the centre of contact area (mm):

$A_d := 2 \cdot \sigma - d$ $\quad$ $A_d = 1.456 \times 10^{-3}$

A2
THE KYH CONTACT MODEL

Constants:

Contact load (N): $P := 3$

Radius of sphere (mm): $B := 2.5$

Young's Modulus of sphere material (N/mm^2): $E_1 := 2.10 \cdot 10^5$

Poisson's Ratio of sphere material: $\nu_1 := 0.292$

Young's Modulus of rough surface material (N/mm^2): $E_2 := 2.10 \cdot 10^5$

Poisson's Ratio of rough surface material: $\nu_2 := 0.292$

Standard deviation of asperity heights of rough surface (mm): $\sigma := 1.450 \cdot 10^{-3}$

Number of asperities of rough surface per mm^2: $\eta := 1430$

Radius of asperity summit of rough surface (mm): $\beta := 28.0 \cdot 10^{-3}$

Modulus of contact (N/mm^2), assuming a rigid smooth sphere:

$$E := \frac{E_2}{1 - \nu_2^2} \qquad E = 2.296 \times 10^5$$

Governing equations:

$$q_0(a,b) := \frac{P \cdot \exp\left[\left(\frac{a}{b}\right)^2\right]}{\pi \cdot b^2 \cdot \left[\exp\left[\left(\frac{a}{b}\right)^2\right] - \left(\frac{a}{b}\right)^2 - 1\right]}$$

$$q_1(r,a,b) := q_0(a,b) \cdot \left[\exp\left[-\left(\frac{r}{b}\right)^2\right] - \exp\left[-\left(\frac{a}{b}\right)^2\right]\right]$$

$$u(r,a) := \frac{a^2 - r^2}{2 \cdot B} \qquad \alpha := \frac{4 - \pi}{4 \cdot \sigma^2}$$

$$q_2(r,a) := \frac{4}{3} \cdot E \cdot \eta \cdot \beta^{\frac{1}{2}} \cdot \int_0^{u(r,a)} (u(r,a) - z)^{\frac{3}{2}} \cdot \left(2 \cdot \alpha \cdot z \cdot \exp\left(-\alpha \cdot z^2\right)\right) dz$$

$$f_1(a,b) := q_1(0,a,b) - q_2(0,a) \qquad f_2(a,b) := q_1\left(\frac{a}{2},a,b\right) - q_2\left(\frac{a}{2},a\right)$$

$$R(a,b) := \begin{pmatrix} a \\ b \end{pmatrix} - \begin{pmatrix} \frac{d}{da} f_1(a,b) & \frac{d}{db} f_1(a,b) \\ \frac{d}{da} f_2(a,b) & \frac{d}{db} f_2(a,b) \end{pmatrix}^{-1} \cdot \begin{pmatrix} f_1(a,b) \\ f_2(a,b) \end{pmatrix}$$

A2 Continued

Iterative solution:

$Tol := 1 \cdot 10^{-12}$ $\qquad$ $Maxiter := 50$

$Sol(a_i, b_i) :=$ | Iter ← 1
| while Iter ≤ Maxiter
| | $a_{New} \leftarrow R(a_i, b_i) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
| | $b_{New} \leftarrow R(a_i, b_i) \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$
| | break if $|a_{New} - a_i| < Tol \wedge |b_{New} - b_i| < Tol$
| | $a_i \leftarrow a_{New}$
| | $b_i \leftarrow b_{New}$
| | Iter ← Iter + 1
| $\begin{pmatrix} a_{New} \\ b_{New} \\ Iter \end{pmatrix}$

Initial guesses:

$Q := Sol(0.05, 0.02)$

ANSWERS:

$$Q = \begin{pmatrix} 0.096 \\ 0.050 \\ 8.000 \end{pmatrix}$$

$$a := Q \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$a = 0.096$

$$b := Q \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$b = 0.050$

Pressure distributions:

$q_0(a, b) = 435.590$

$q_1(0, a, b) = 425.164$ $\qquad$ $q_2(0, a) = 425.164$

$q_1\left(\frac{a}{2}, a, b\right) = 160.907$ $\qquad$ $q_2\left(\frac{a}{2}, a\right) = 160.907$

Asperity deformation at the centre of contact area (mm): $\qquad$ $r := 0$

$u(r, a) = 1.846 \times 10^{-3}$

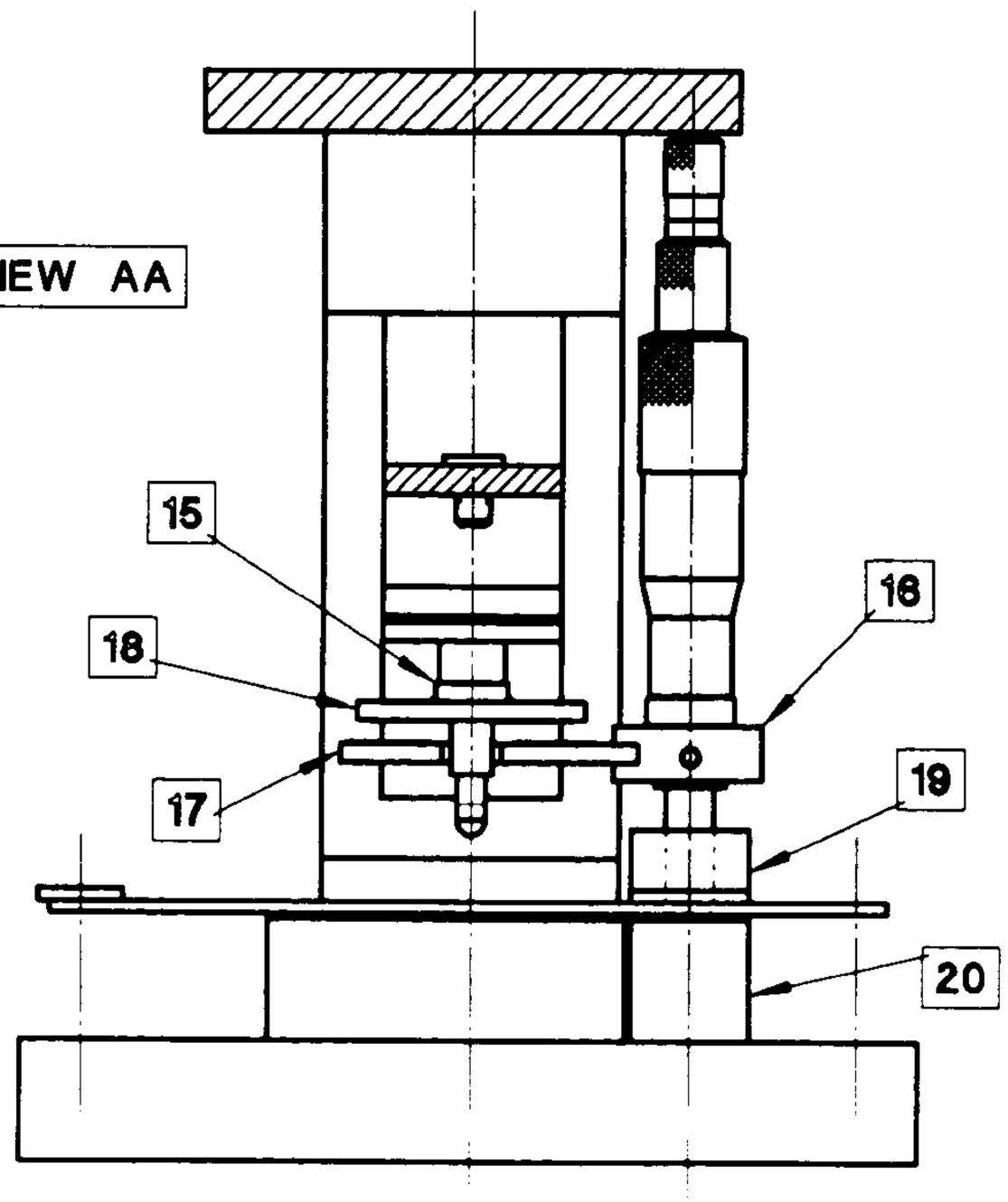
Compliance at the centre of contact area (mm):

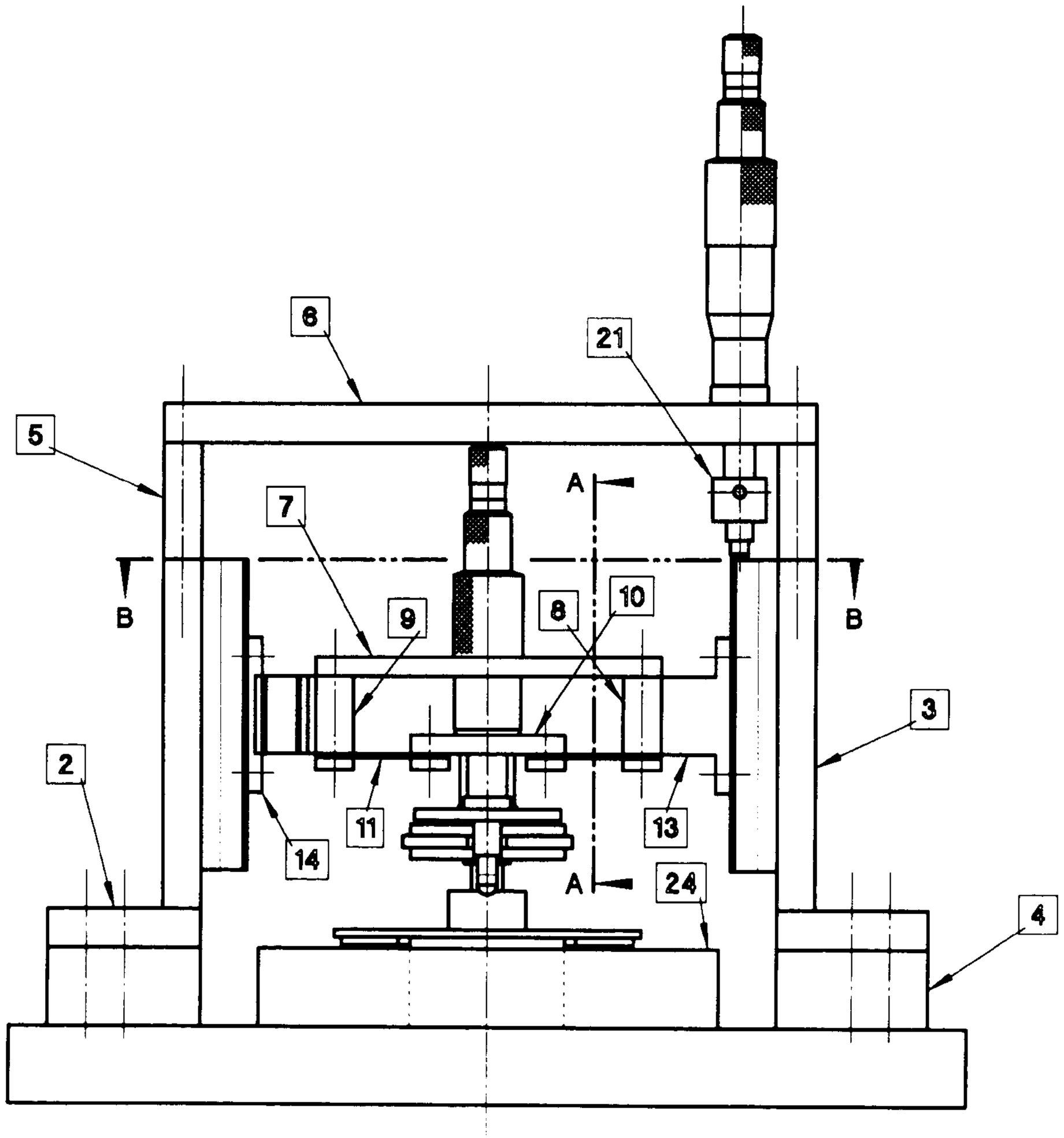
$\alpha := \frac{r^2}{2 \cdot B} + u(r, a)$ $\qquad$ $\alpha = 1.846 \times 10^{-3}$

APPENDIX B

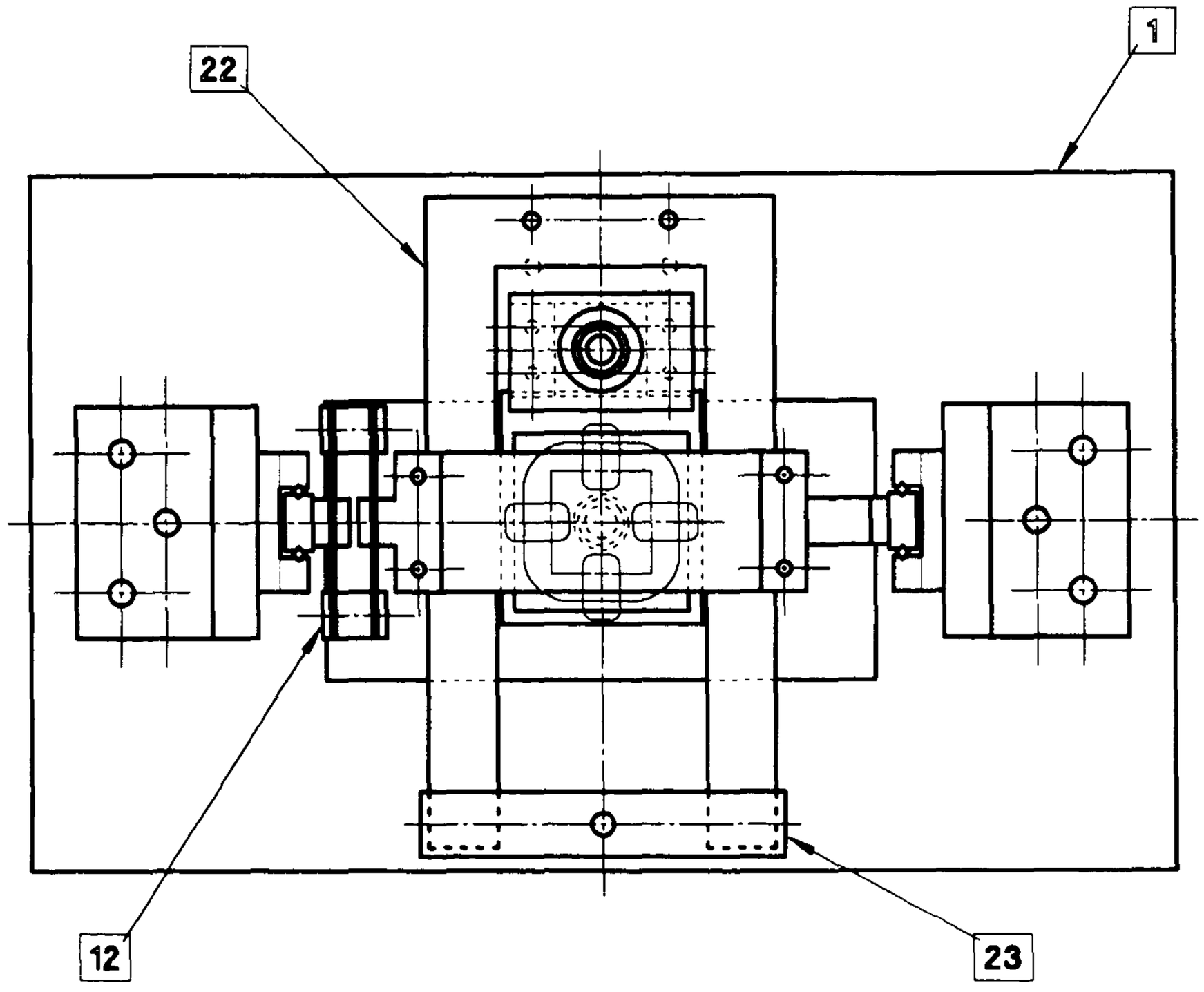
ASSEMBLY & PART WORKSHOP DRAWINGS OF THE NEW TEST-RIG

VIEW AA





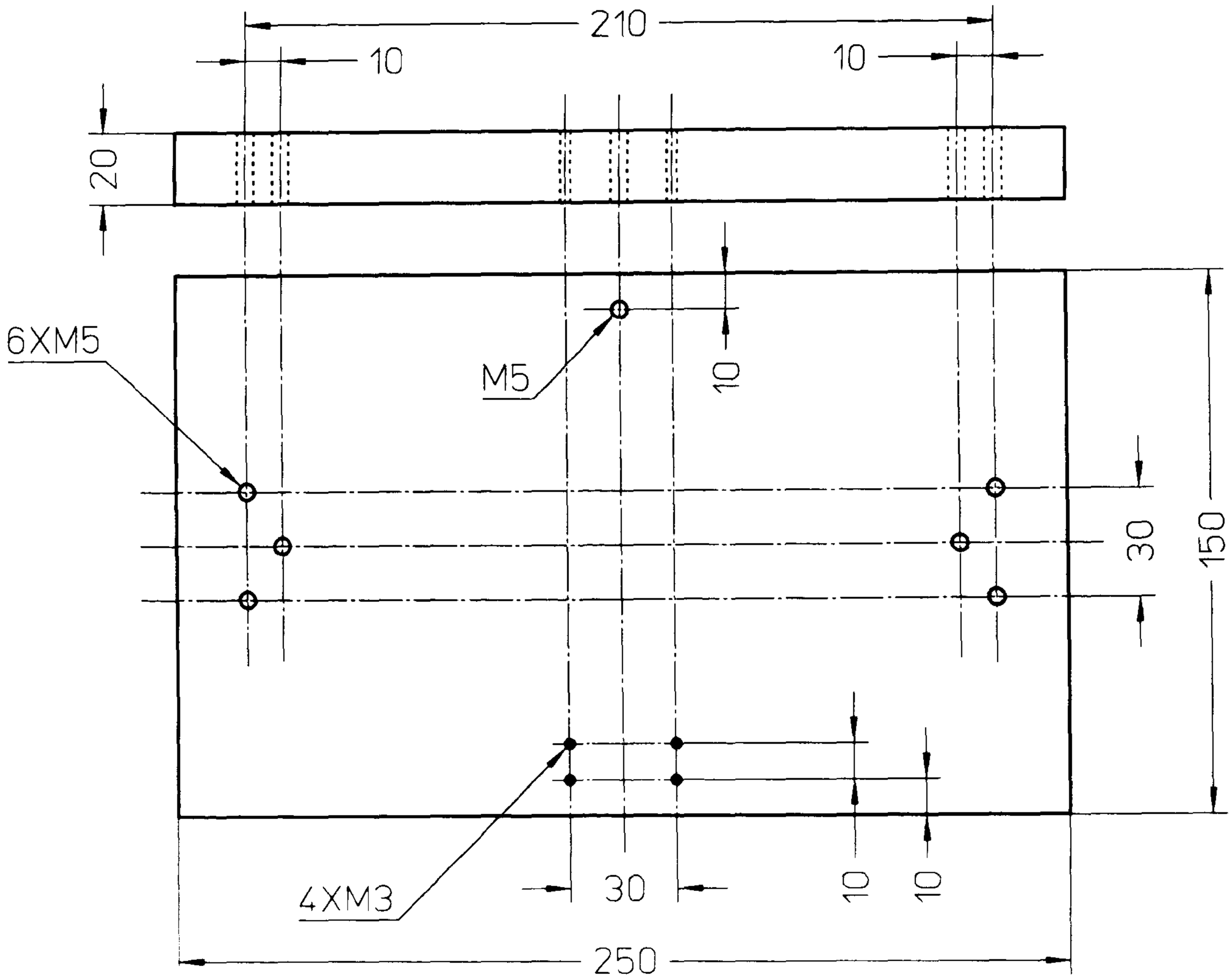
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
OO	THE NEW TEST-RIG	---	1	1:2	mm	30.06.2000	1 Of 2



VIEW BB

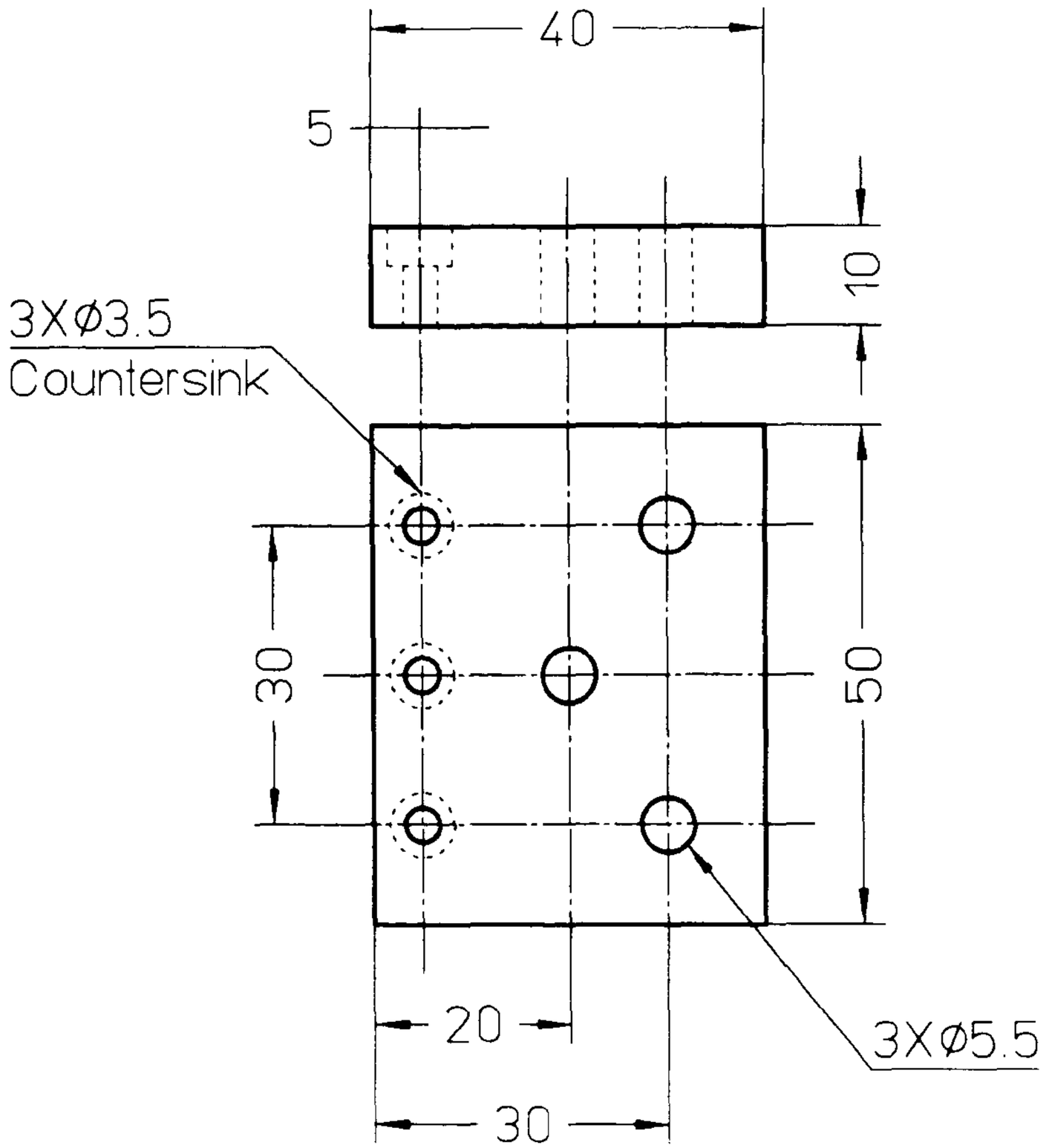
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
OO	THE NEW TEST-RIG	---	1	1:2	mm	30.06.2000	2 Of 2

2107

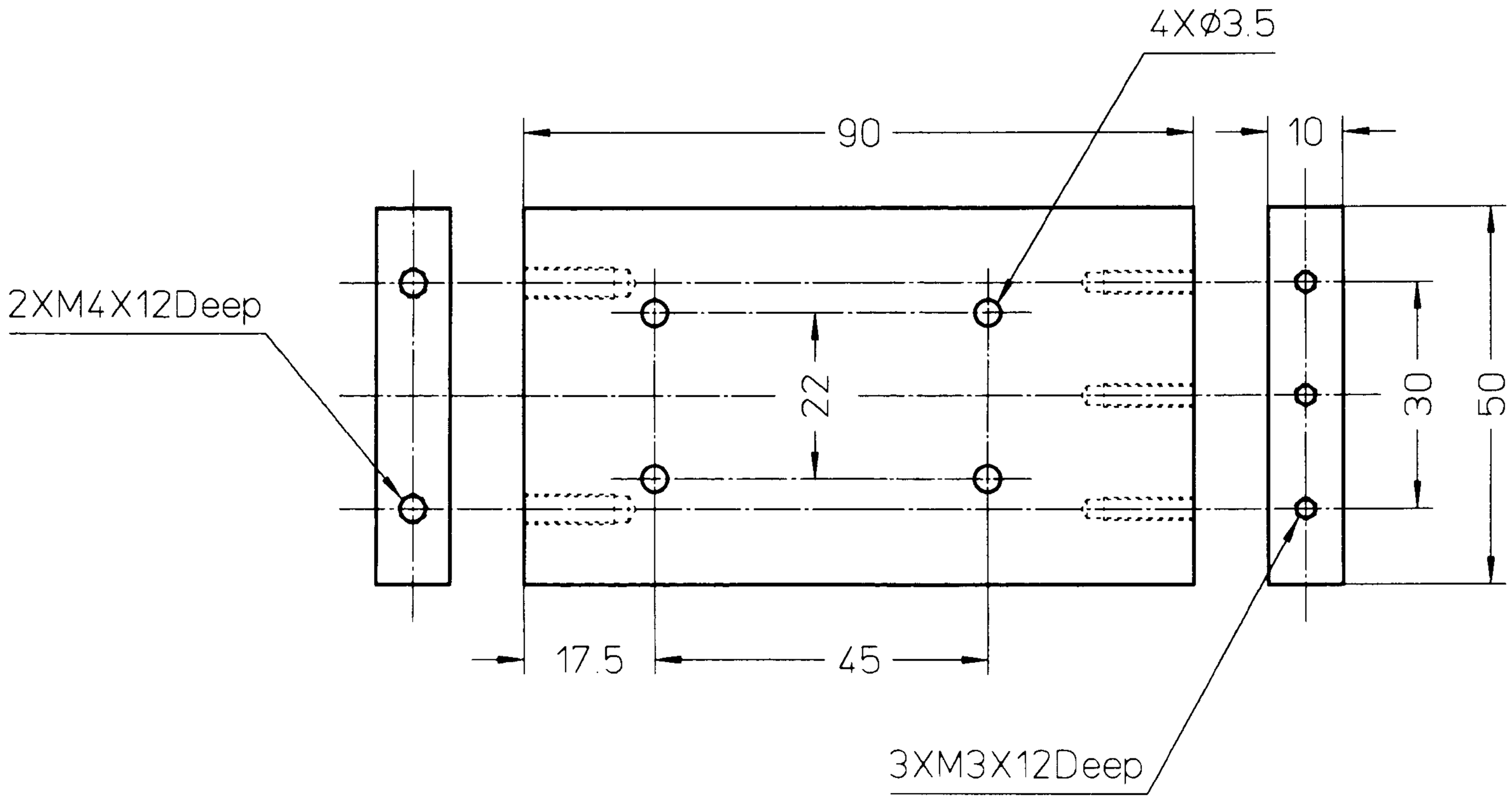


Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
01	Base-Plate	Steel	1	1:2	mm	30.06.2000	1 Of 1

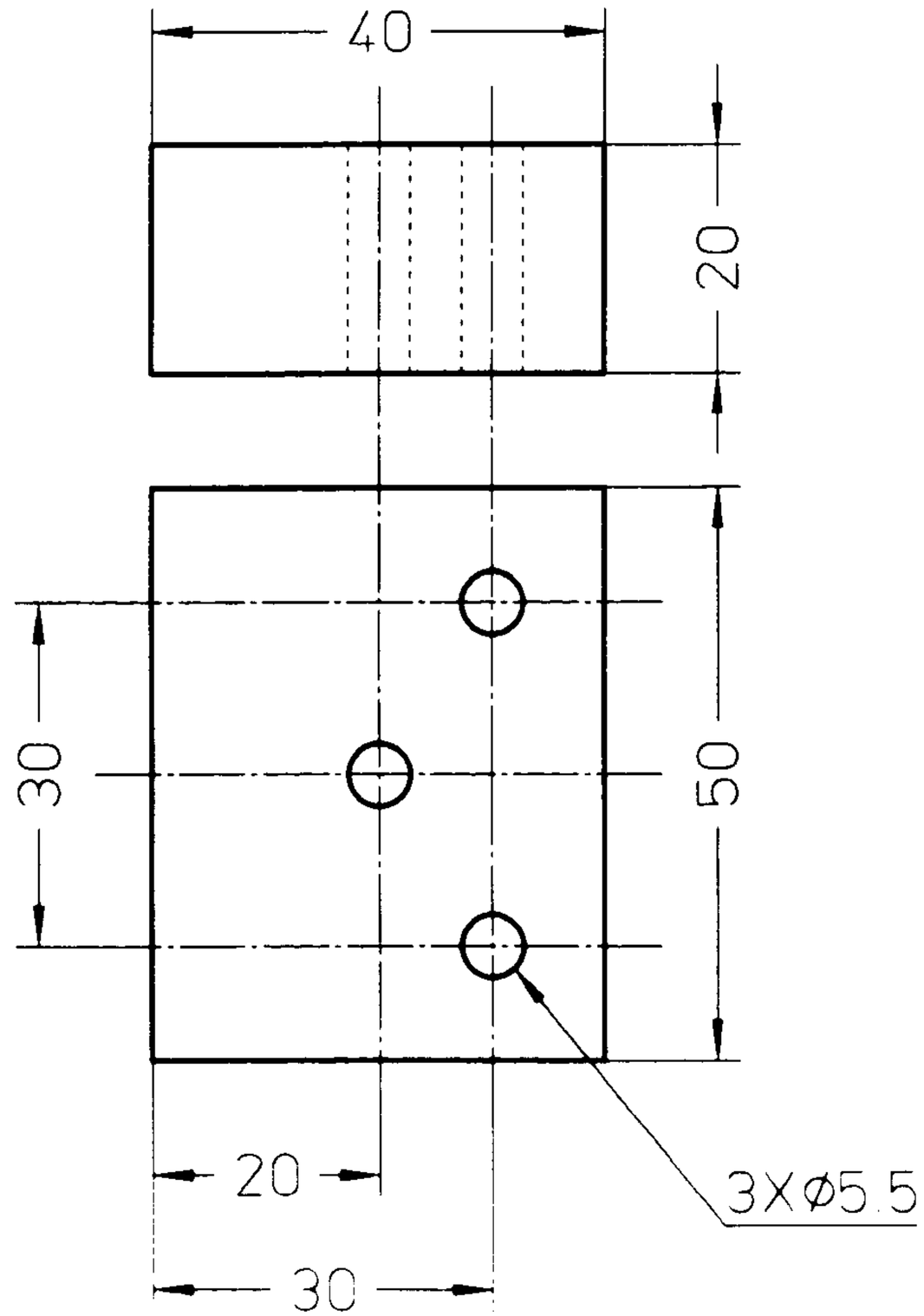
210



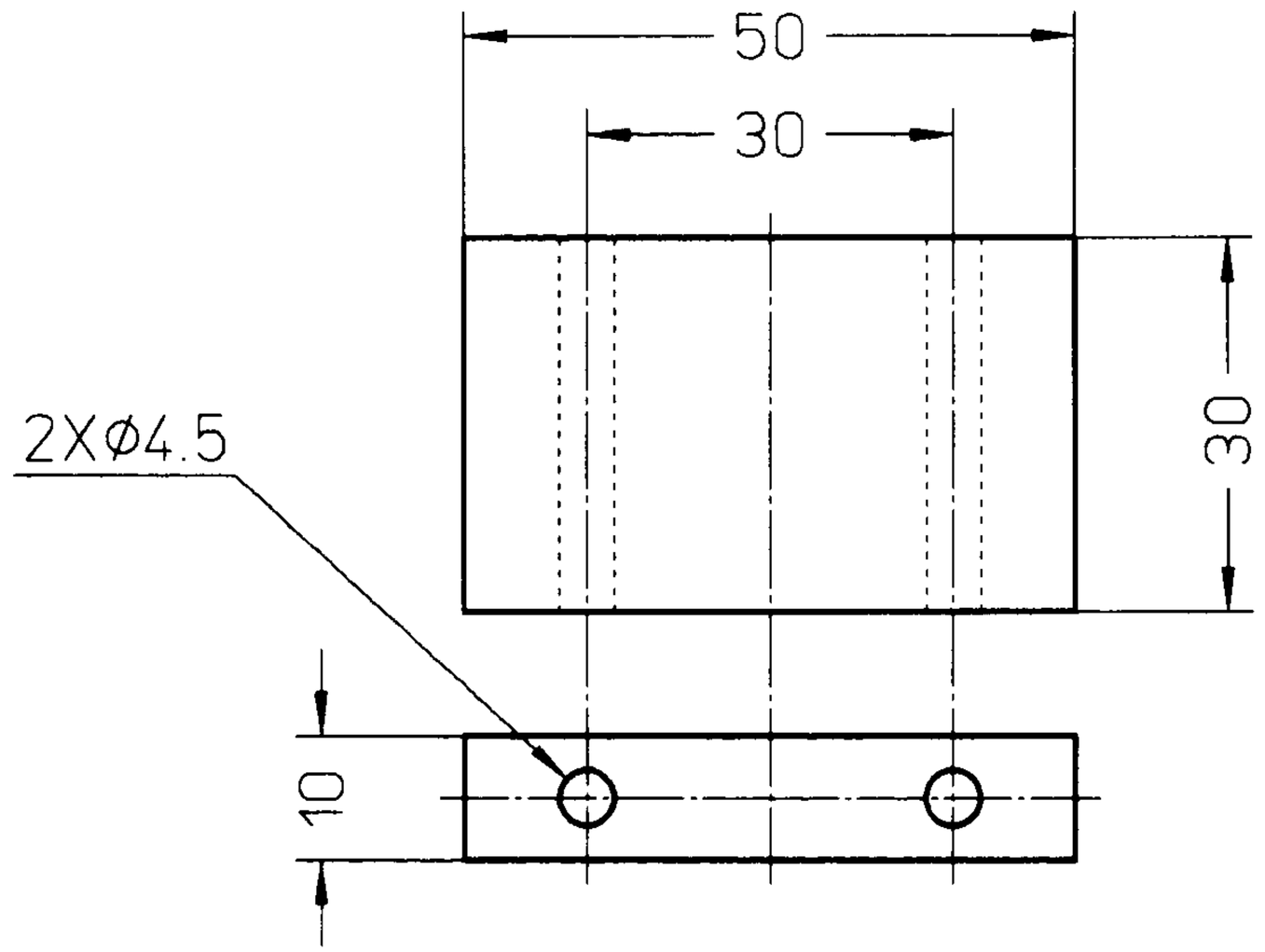
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
02	Upright Base	Steel	2	1:1	mm	30.06.2000	1 Of 1



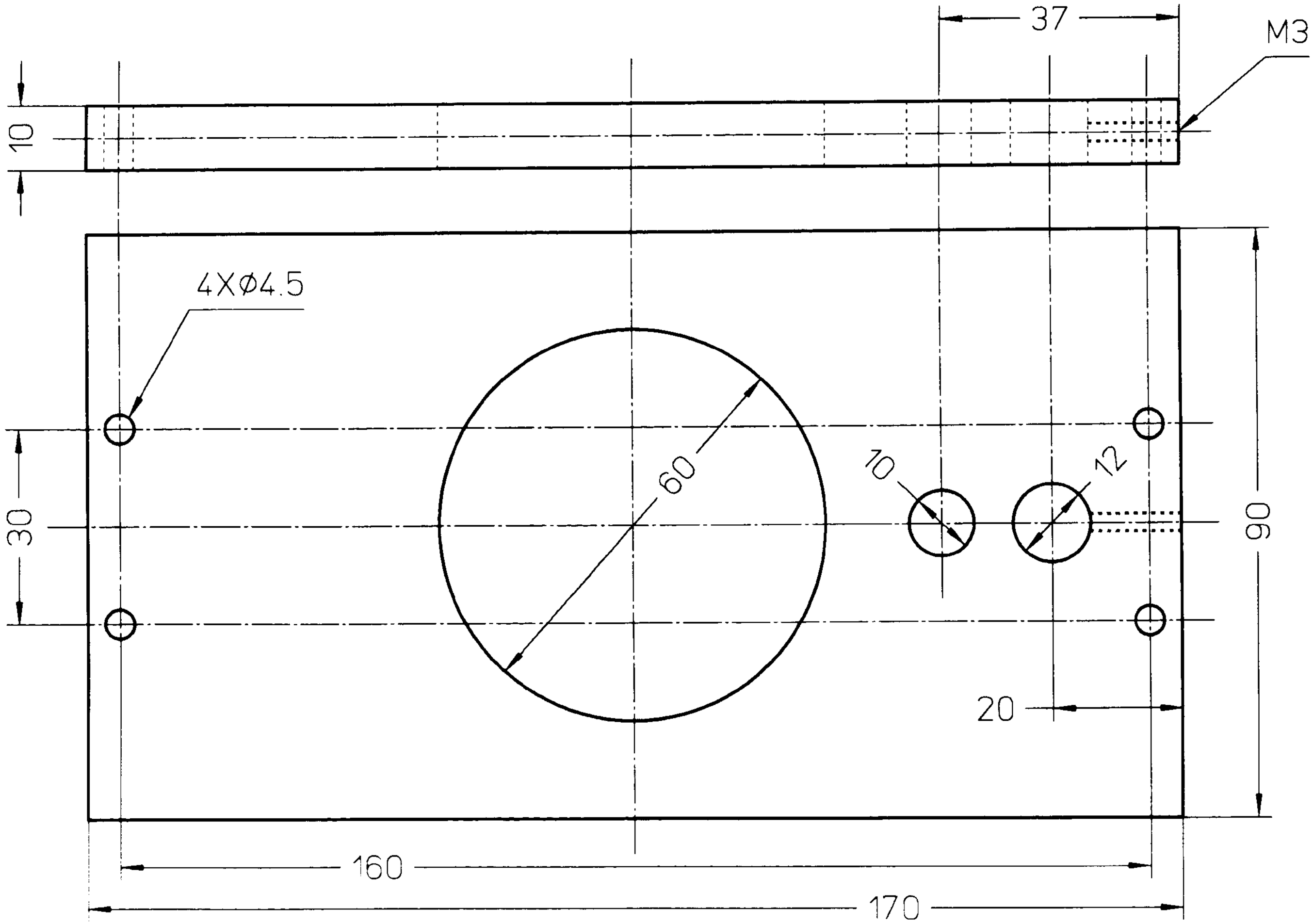
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
03	Upright	Steel	2	1:1	mm	30.06.2000	1 Of 1



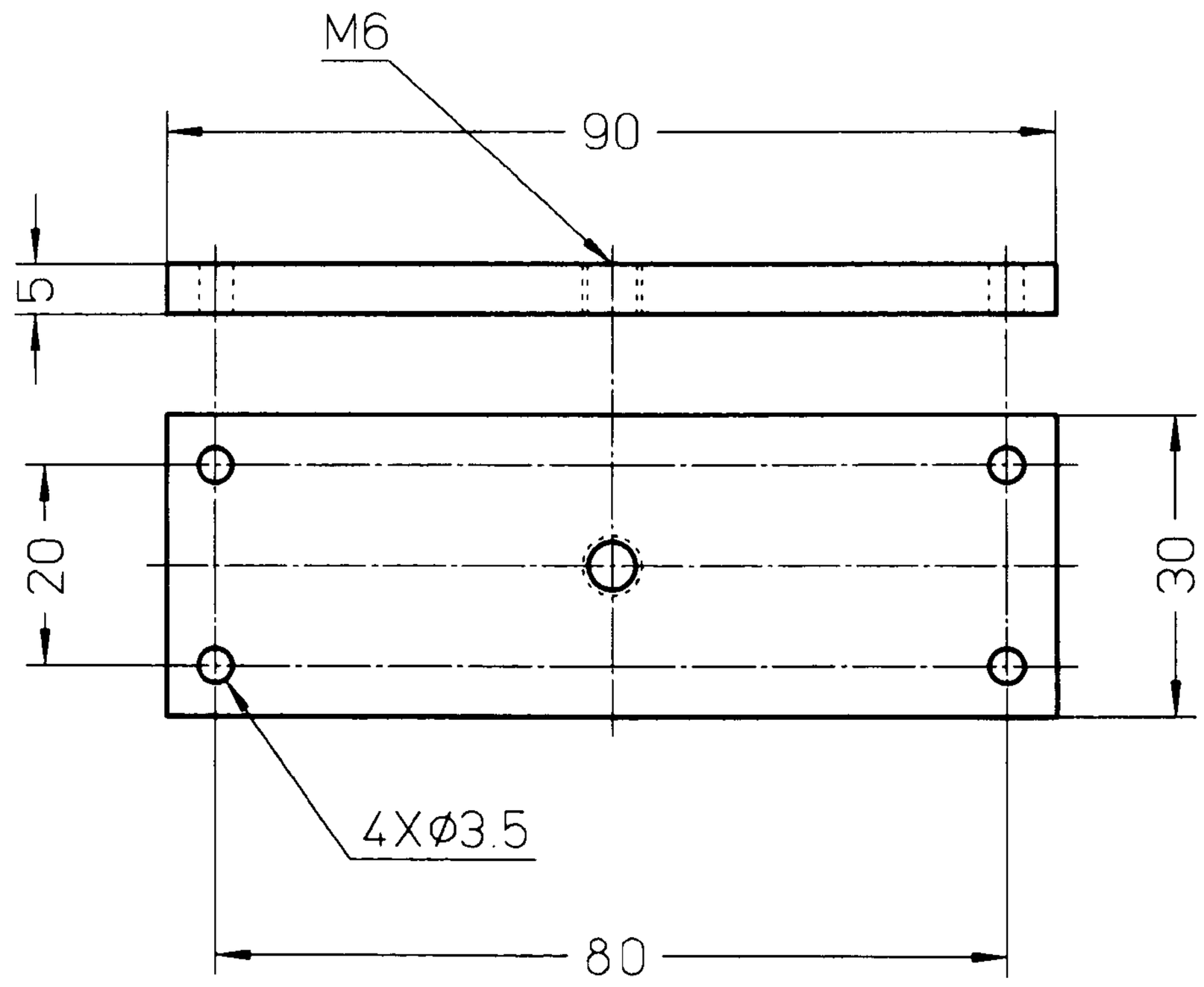
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
04	Upright Bottom Spacer	Steel	2	1:1	mm	30.06.2000	1 Of 1



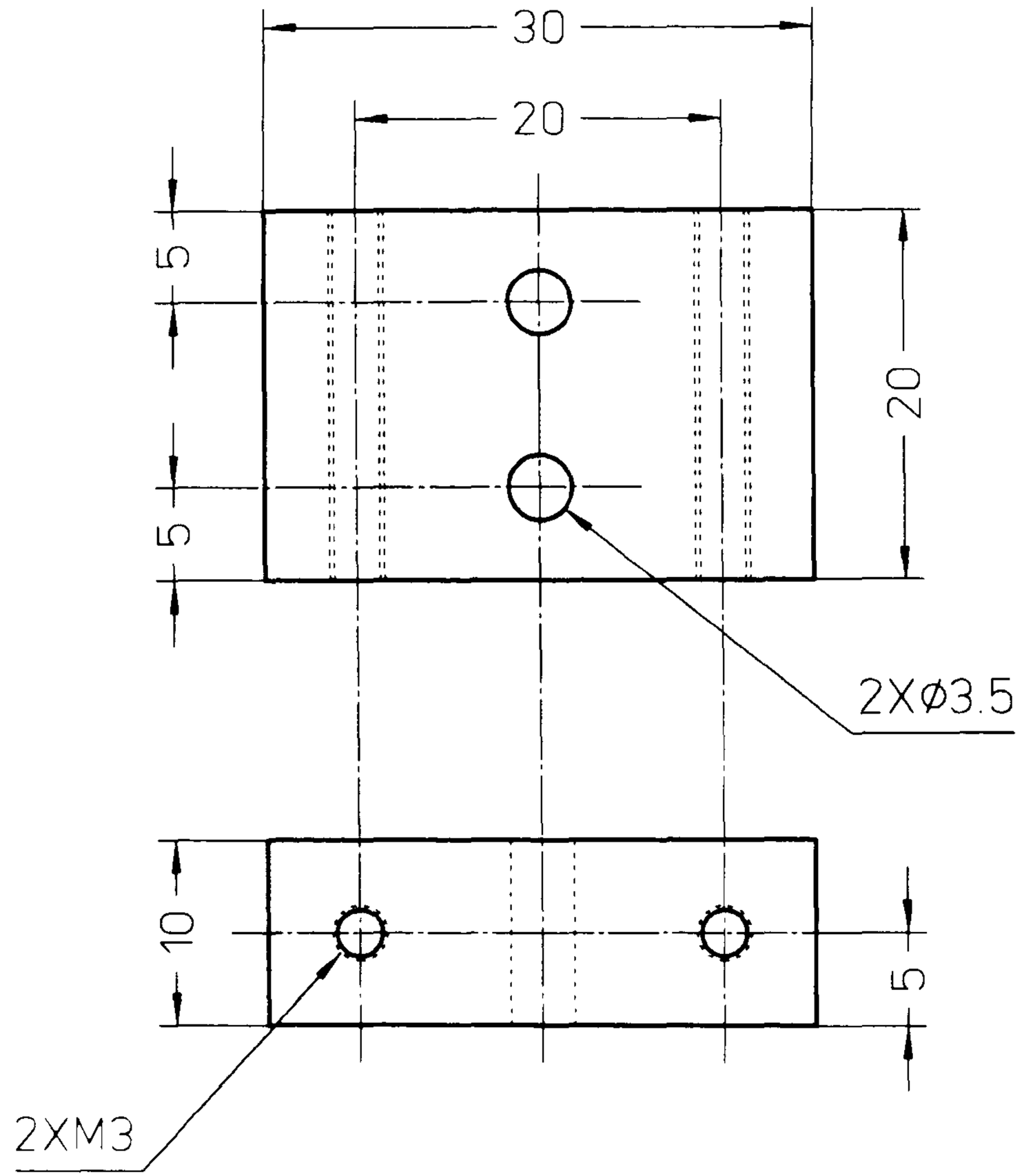
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
05	Upright Top Spacer	Aluminium	2	1:1	mm	30.06.2000	1 Of 1



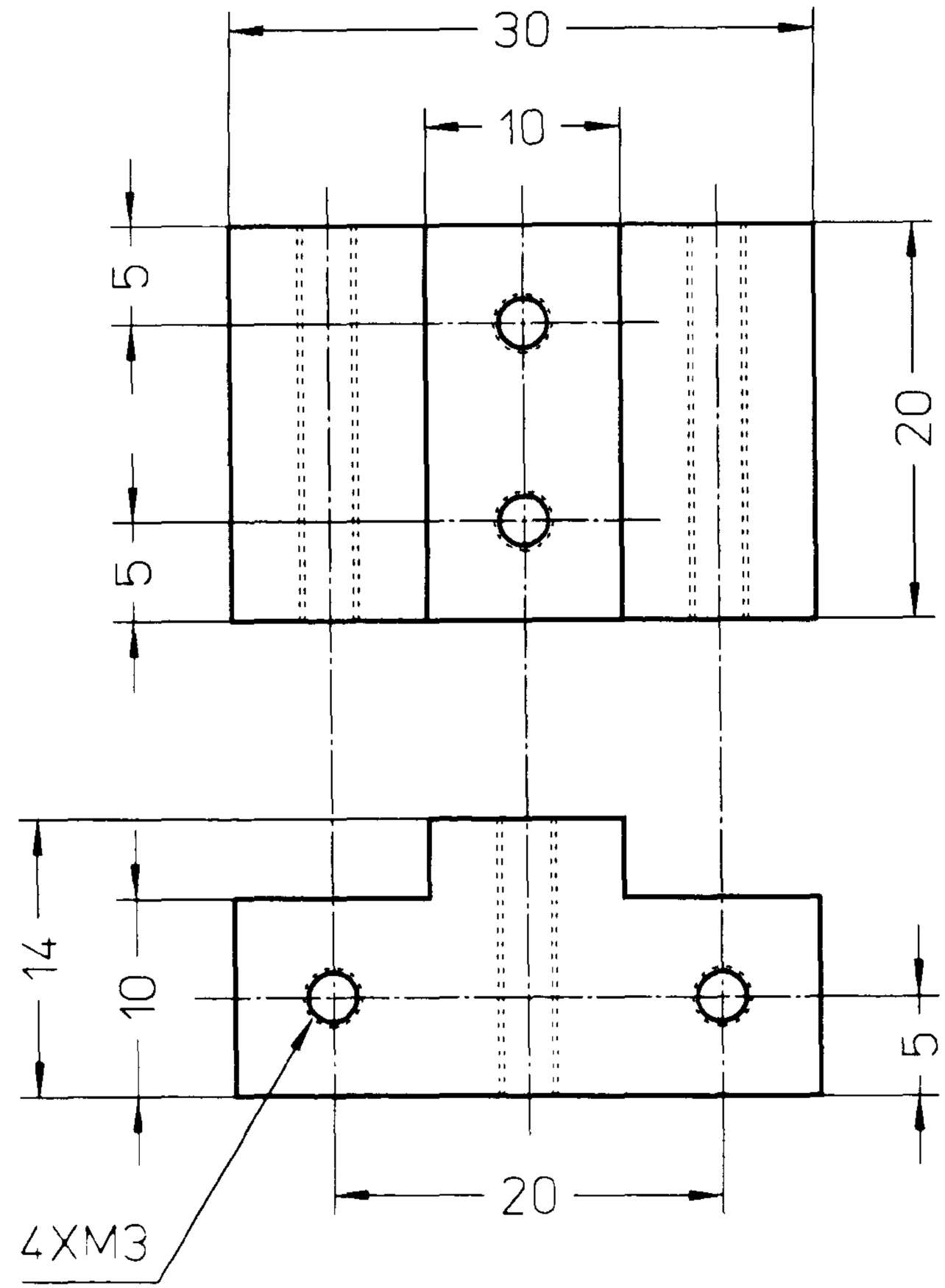
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
06	Top Support Plate	Aluminium	1	1:1	mm	30.06.2000	1 Of 1



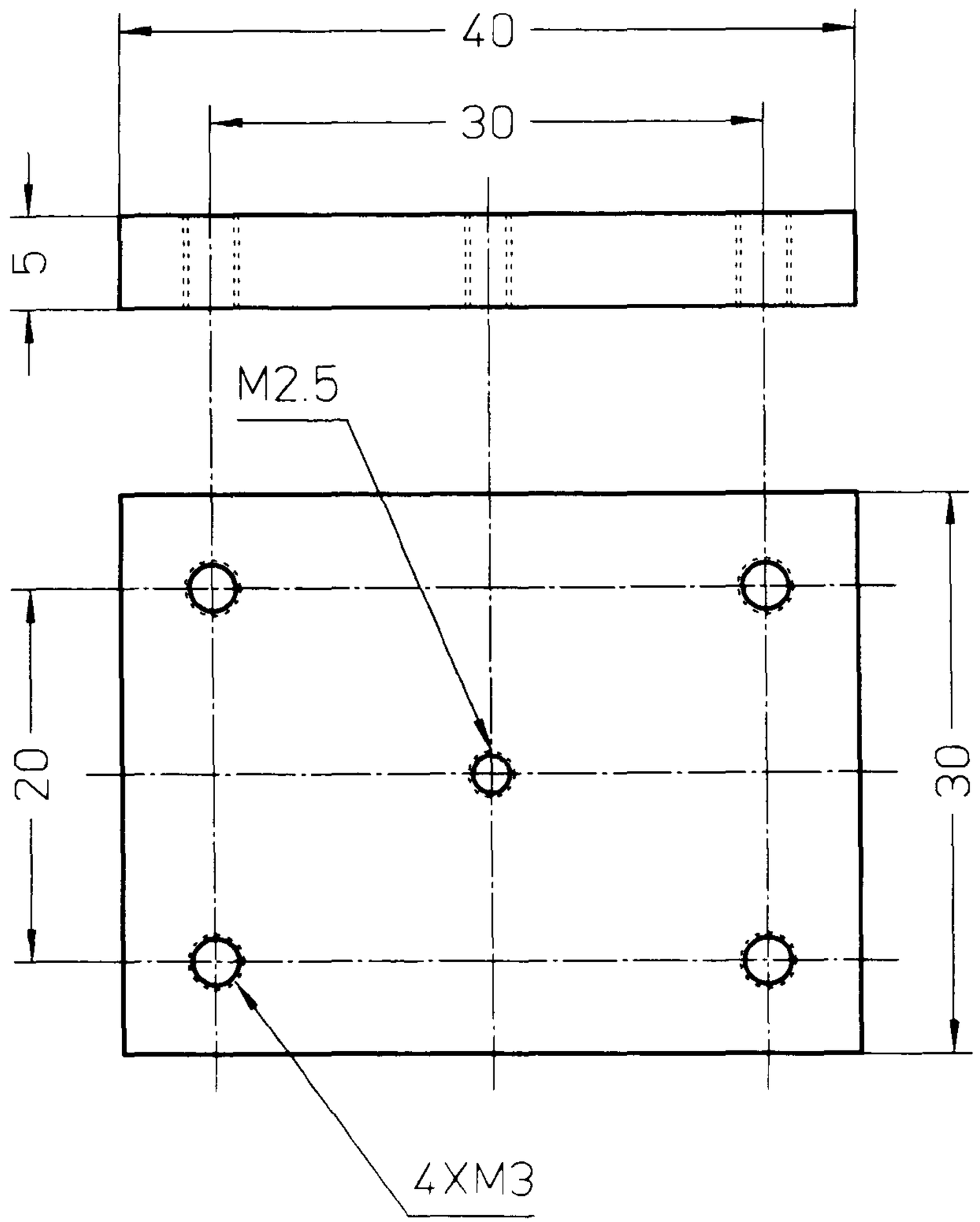
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
07	Carriage Top Plate	Steel	1	1:1	mm	30.06.2000	1 Of 1



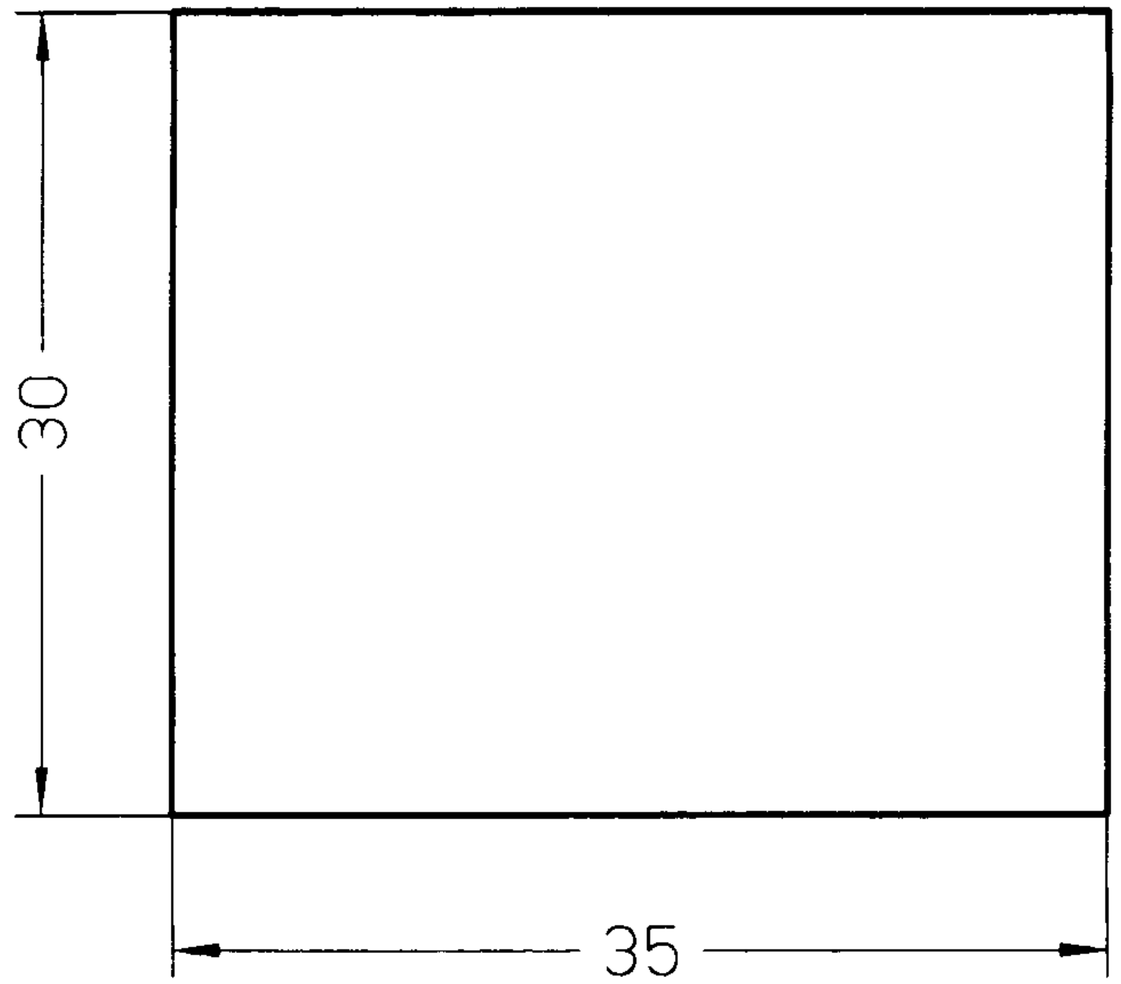
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
08	Carriage Right Plate	Steel	1	2:1	mm	30.06.2000	1 Of 1

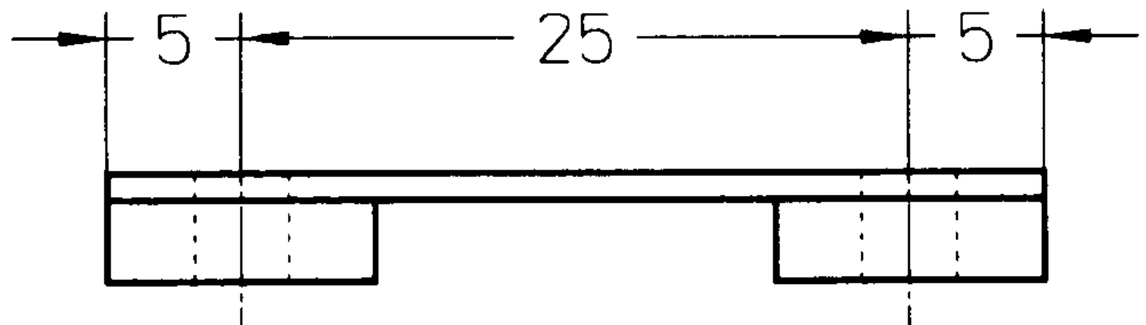


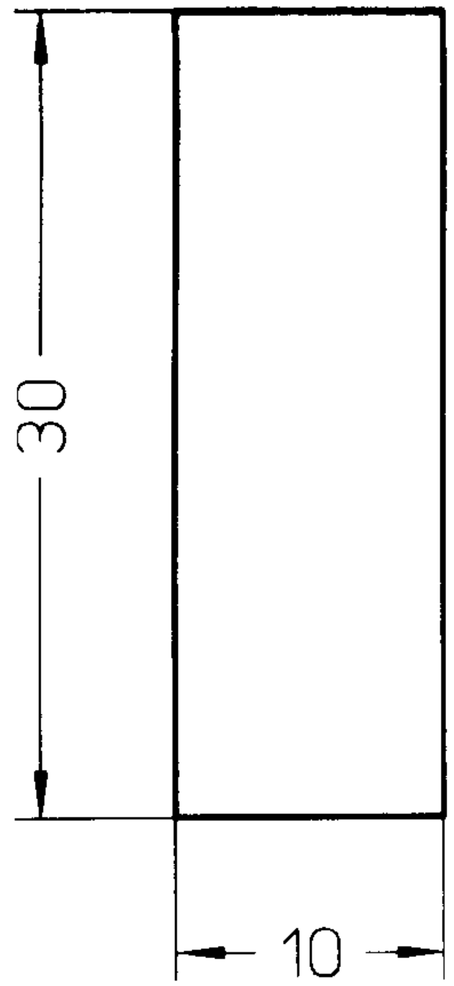
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
09	Carriage Left Plate	Steel	1	2:1	mm	30.06.2000	1 Of 1



Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
10	Carriage Bottom Plate	Steel	1	2:1	mm	30.06.2000	1 Of 1

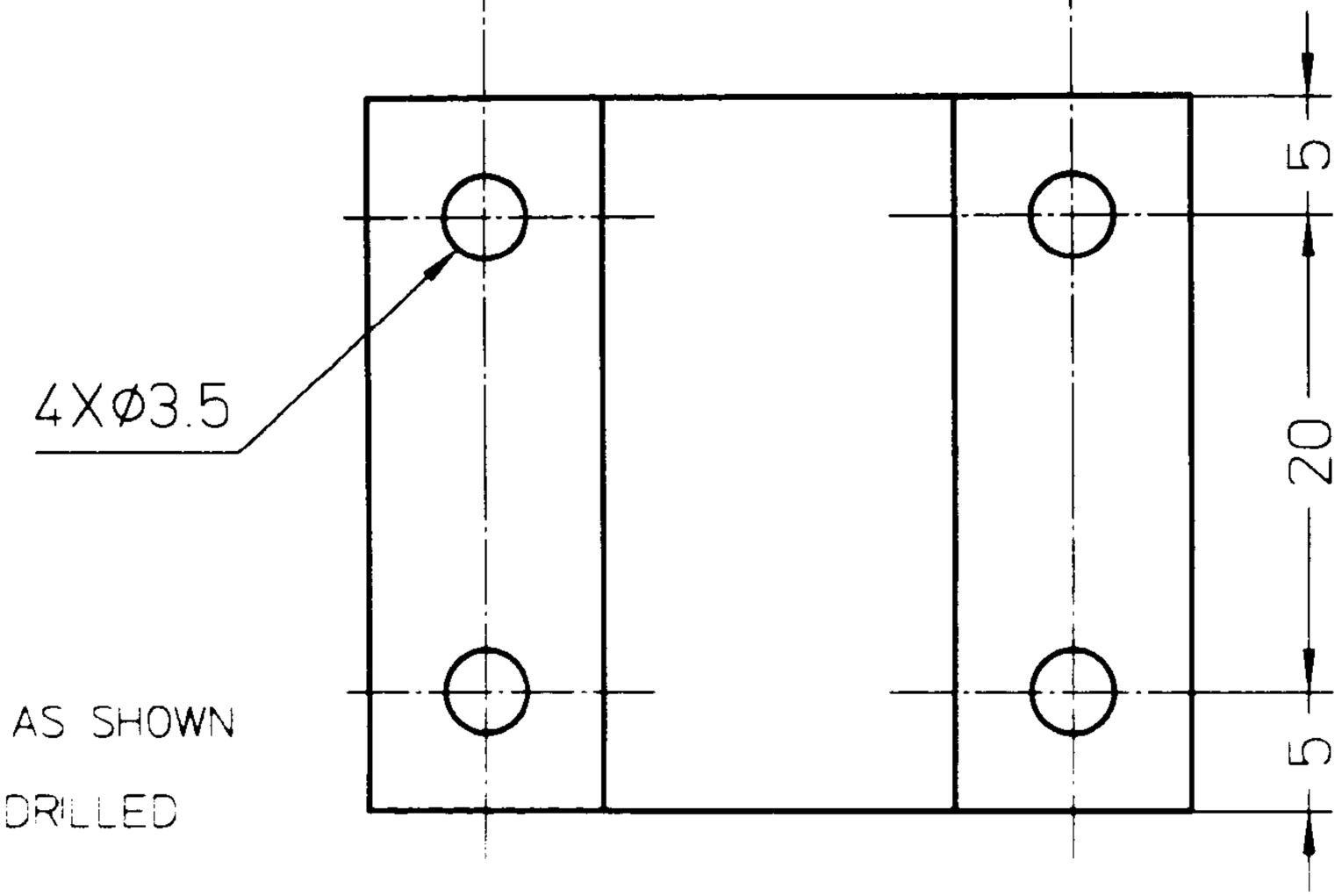






11-02
FLEXTURE
1 MM STEEL SHEET
(1X)

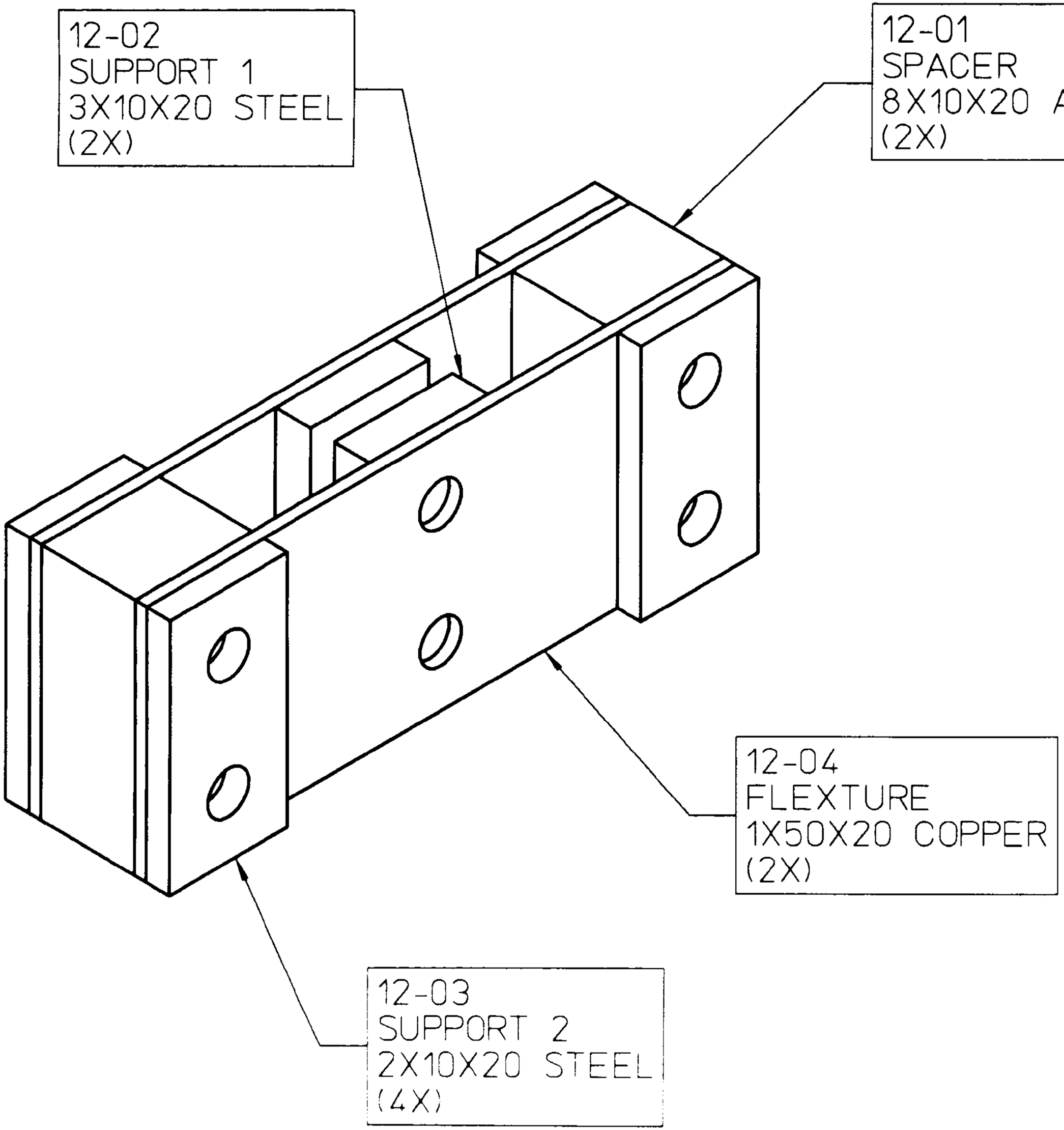
11-01
SUPPORT
STEEL
(2X)



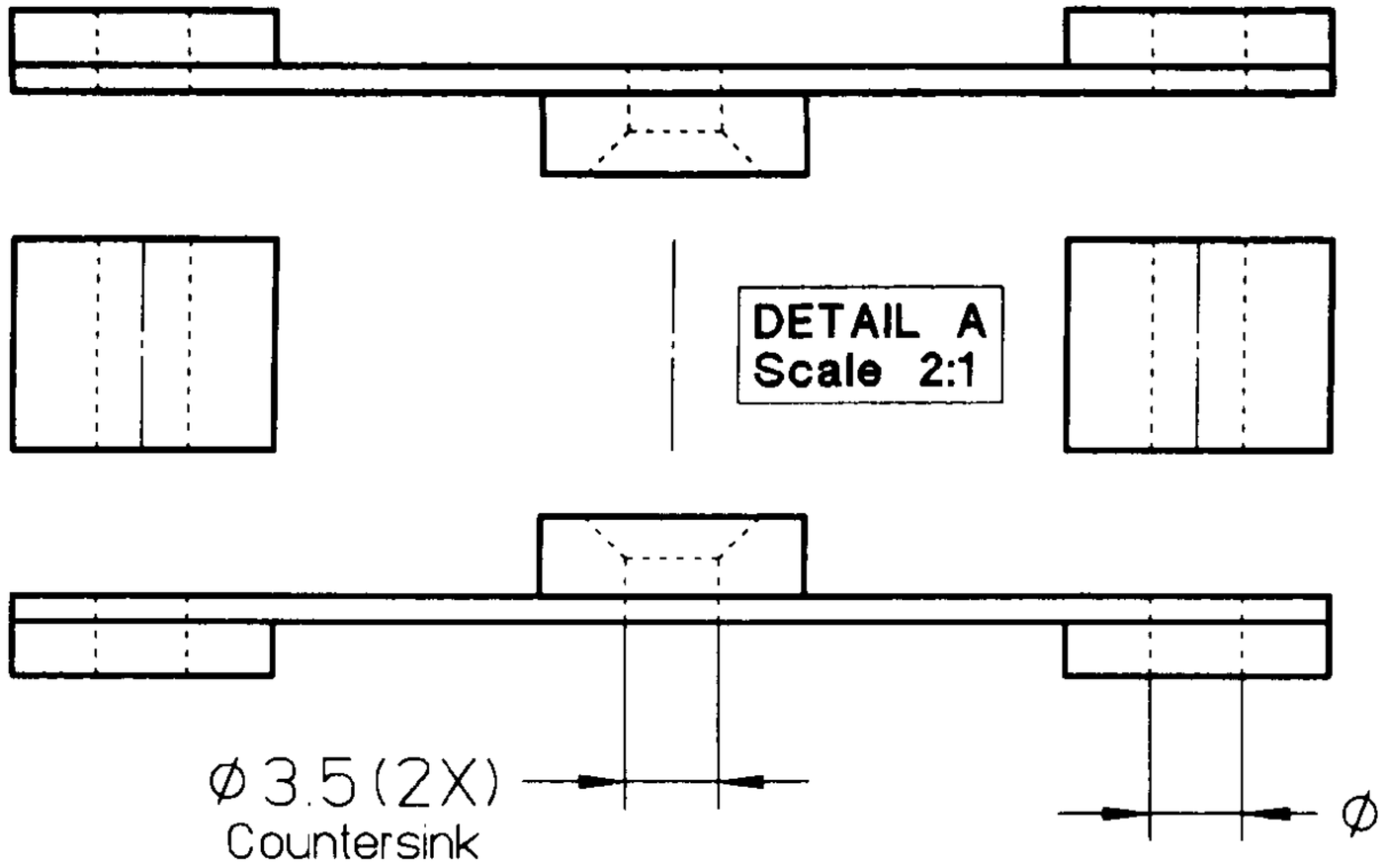
ASSEMBLY VIEWS

NOTES
1 PARTS CAN BE ASSEMBLED BY GLUE AS SHOWN IN THE ASSEMBLY VIEWS
2 THEN, FOUR HOLES OF Ø3.5 CAN BE DRILLED ACCORDING TO DIMENSIONS GIVEN

Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
11	**Horizontal Flexture Assembly**	**"as indicated"**	**2**	**2:1**	**mm**	**30.06.2000**	**1 Of 1**



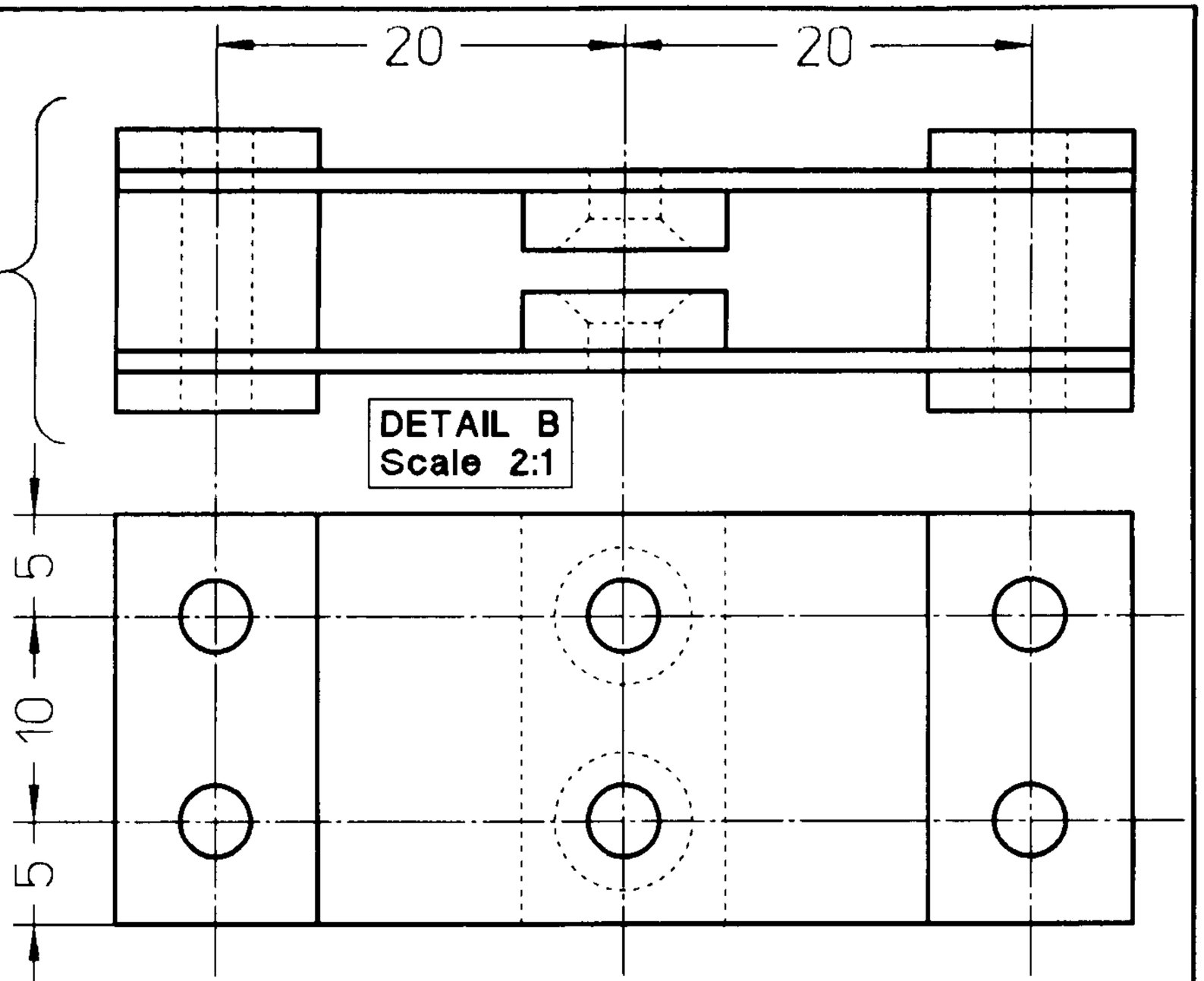
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
12	Slide Flexture Assembly	"as indicated"	1	2:1	mm	30.06.2000	1 Of 2



DETAIL A
Scale 2:1

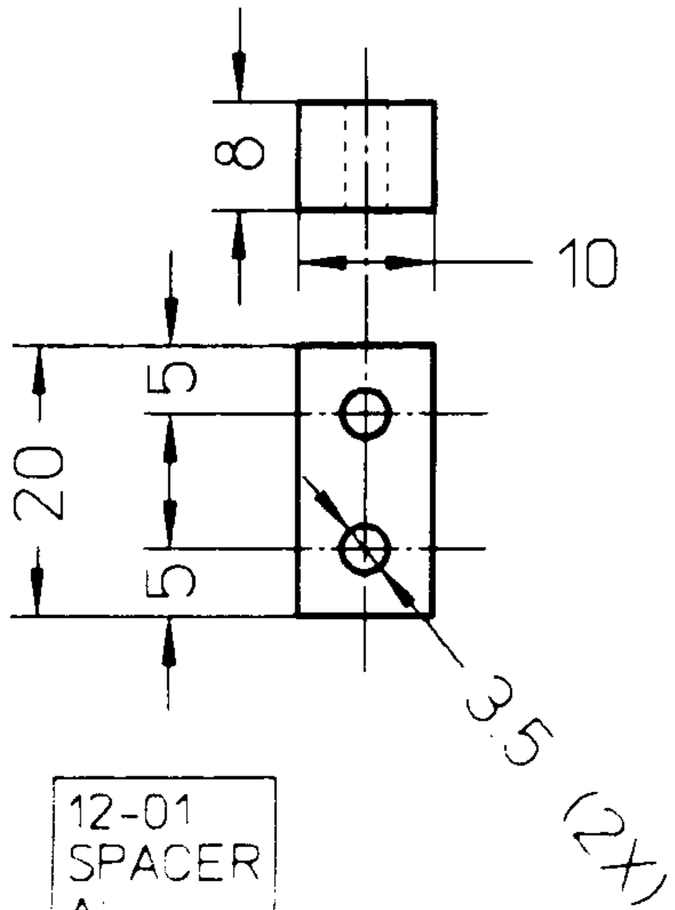
Ø3.5(2X)
Countersink

Ø3.5(4X)

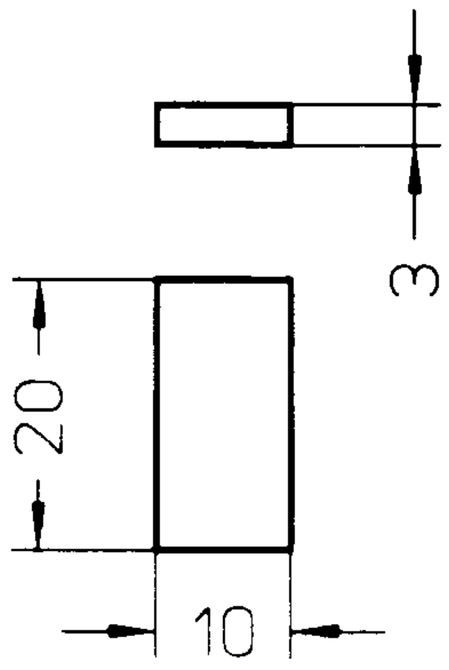


DETAIL B
Scale 2:1

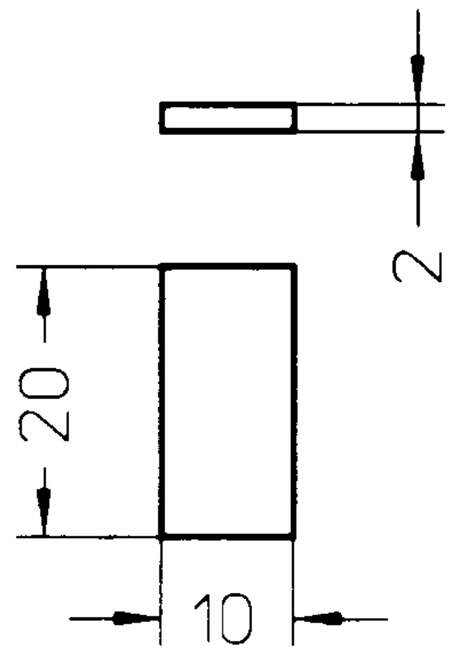
NOTES:
1. THIS ASSEMBLY CAN BE PRE-ASSEMBLED BY GLUE, AS SHOWN IN DETAIL A, TO HAVE TWO JOINED SIDES.
2. THEN, 6 HOLES OF Ø3.5 CAN BE DRILLED ON EACH SIDE AT THE INDICATED POSITIONS IN DETAIL B. THE TWO HOLES AT THE CENTER ARE COUNTERSINK HOLES.



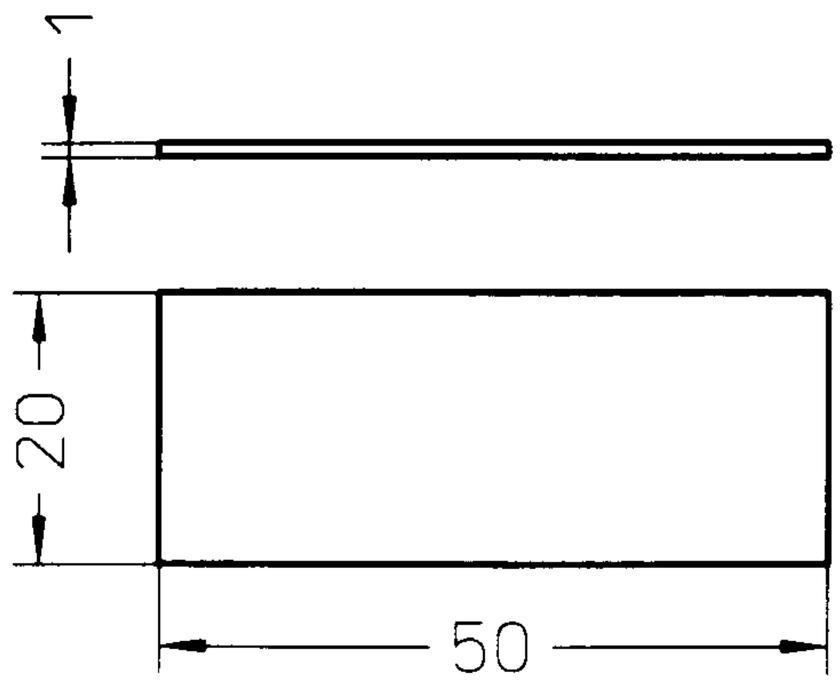
12-01
SPACER
AL
(2X)



12-02
SUPPORT 1
STEEL
(2X)

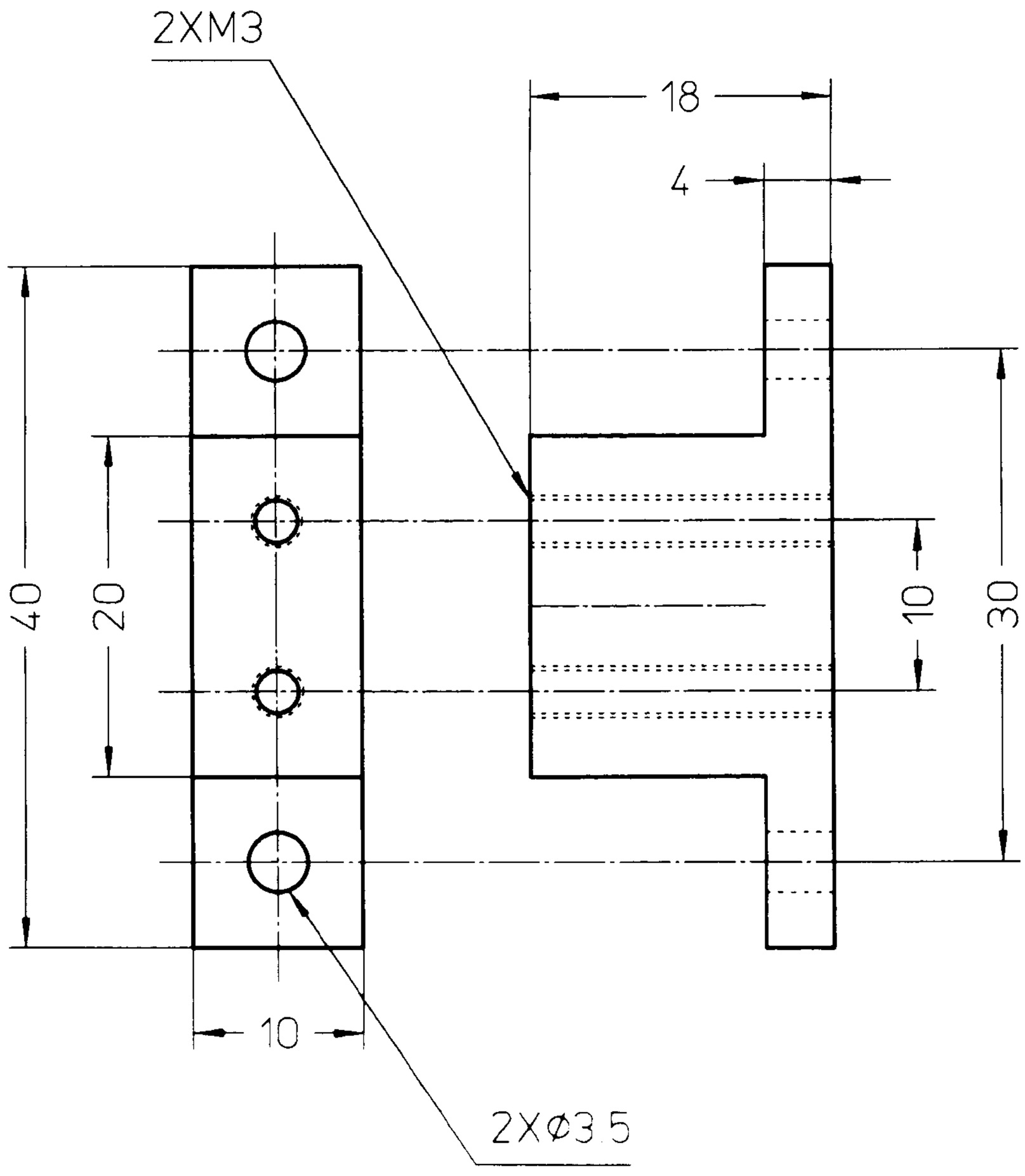


12-03
SUPPORT 2
STEEL
(4X)

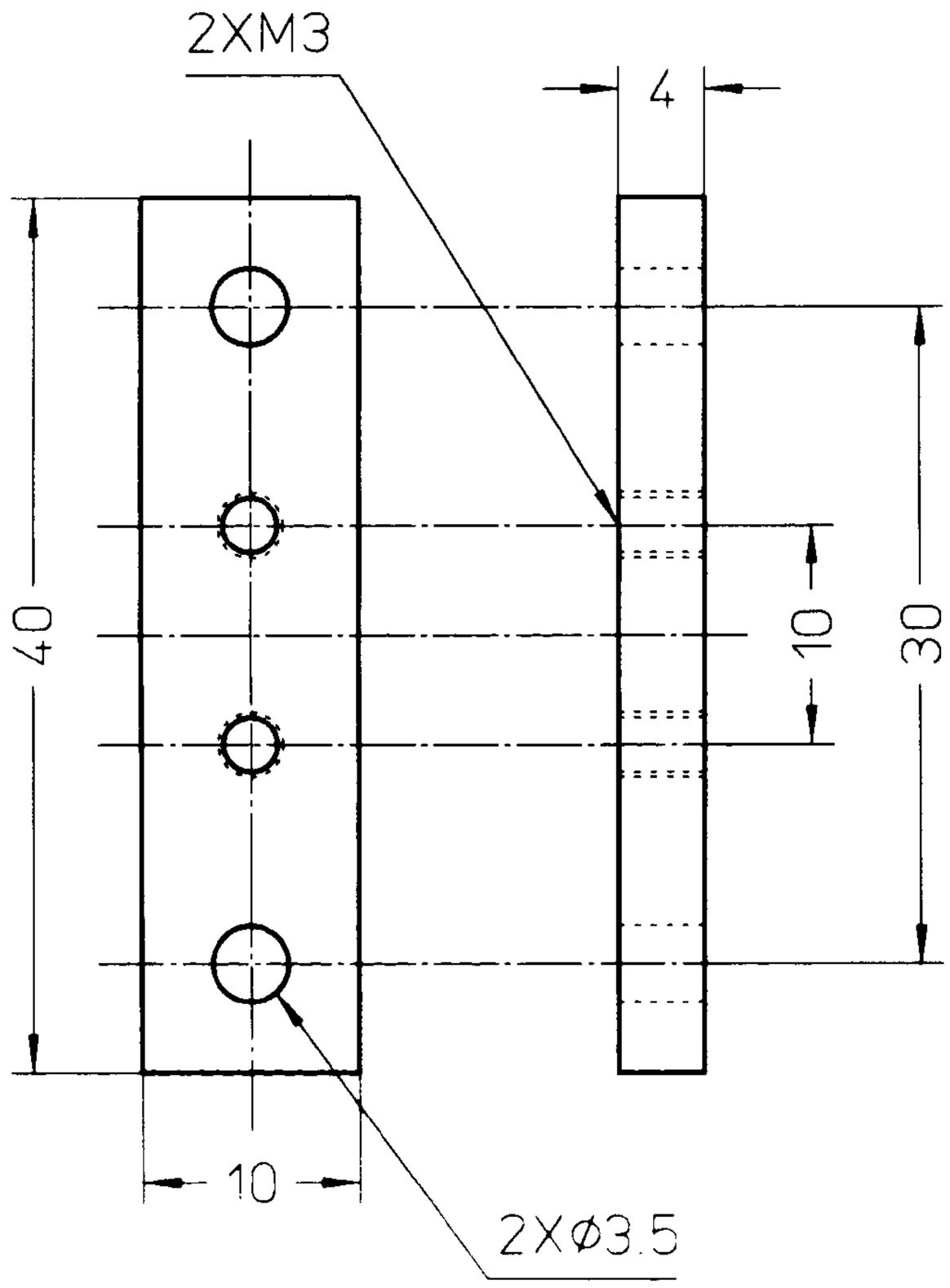


12-04
FLEXTURE
1 MM COPPER SHEET
(2X)

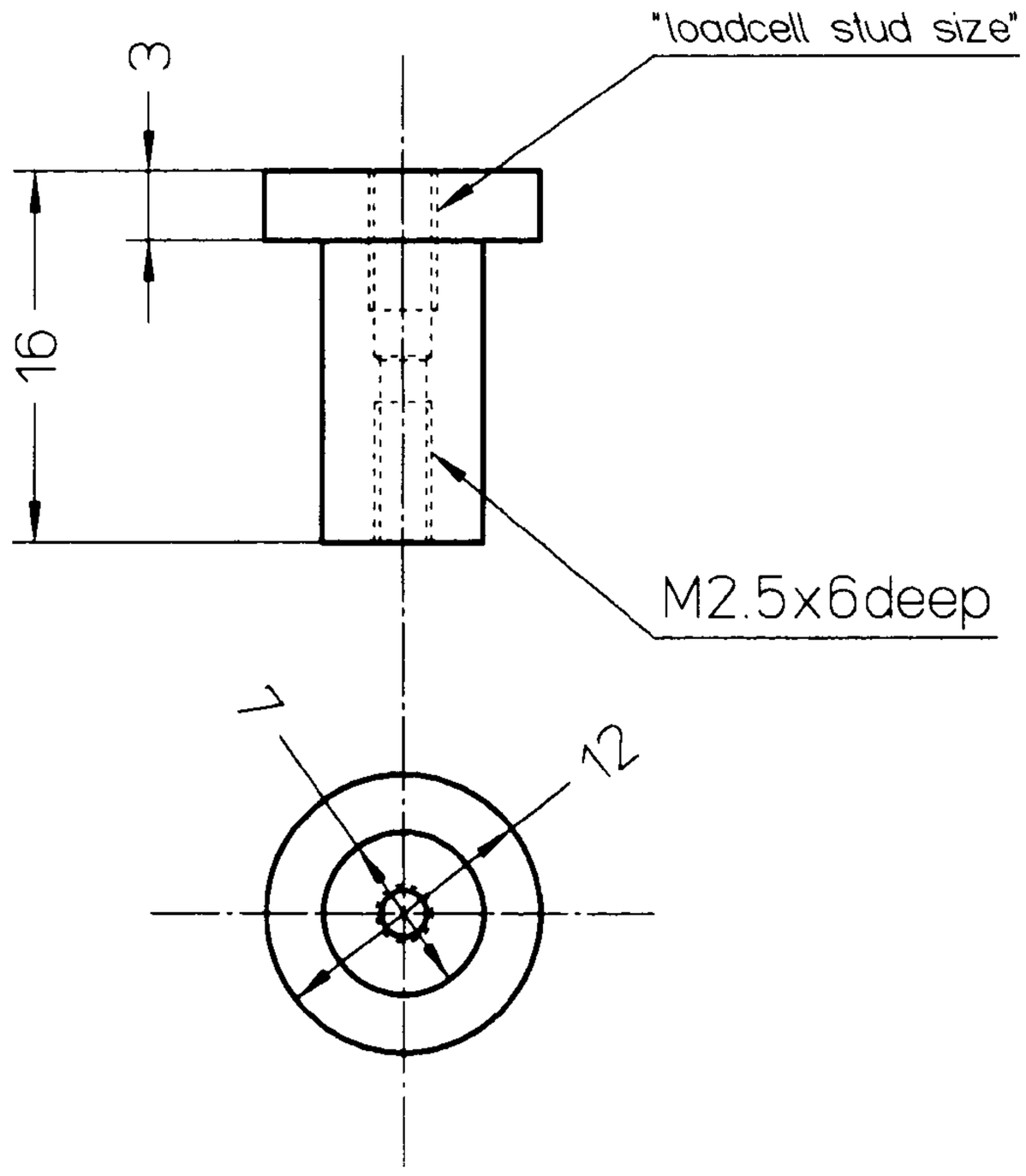
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
12	Slide Flexture Assembly	"as indicated"	1	1:1	mm	30.06.2000	2 Of 2



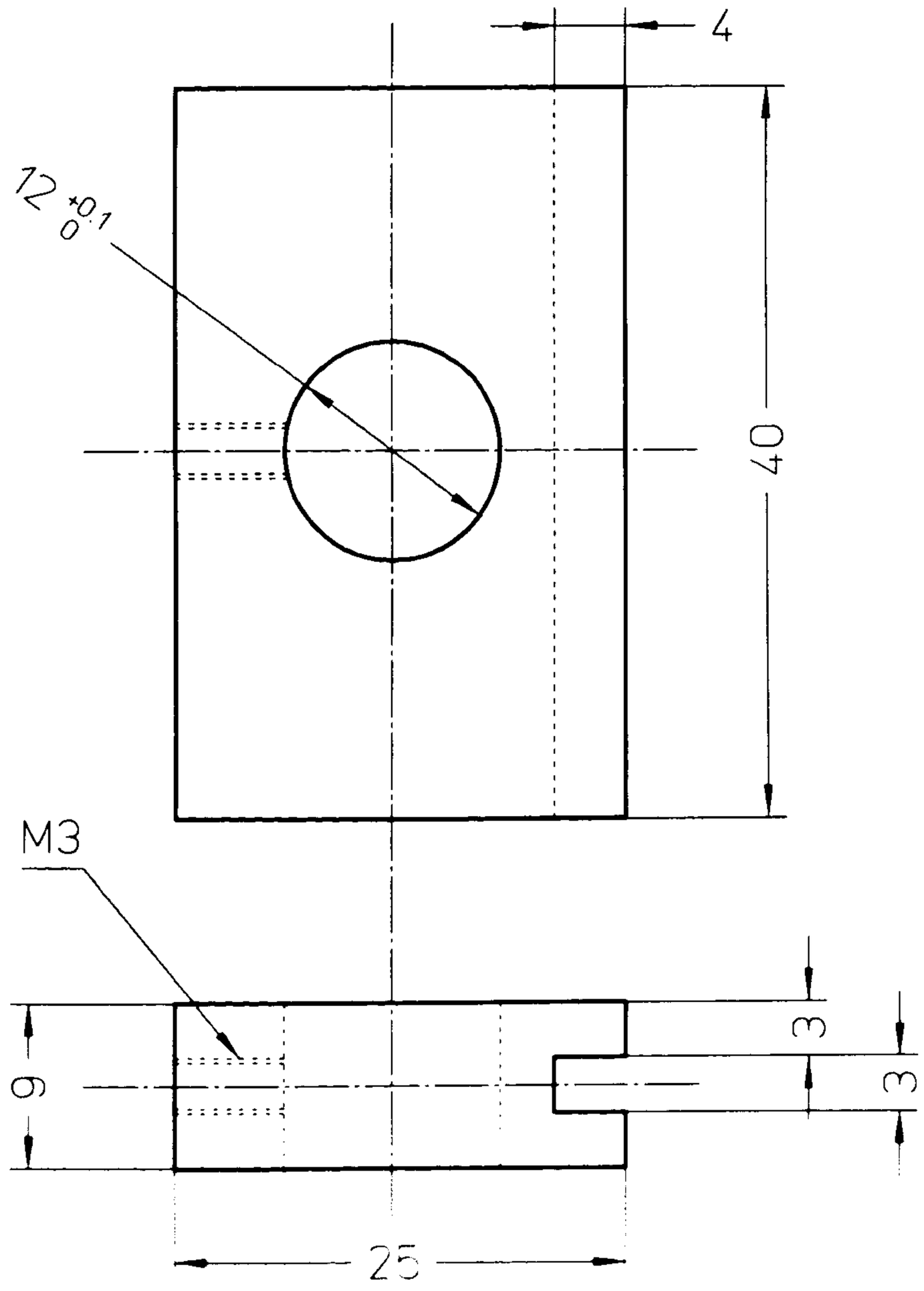
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
13	Carriage Right Holder	Steel	1	2:1	mm	30.06.2000	1 Of 1



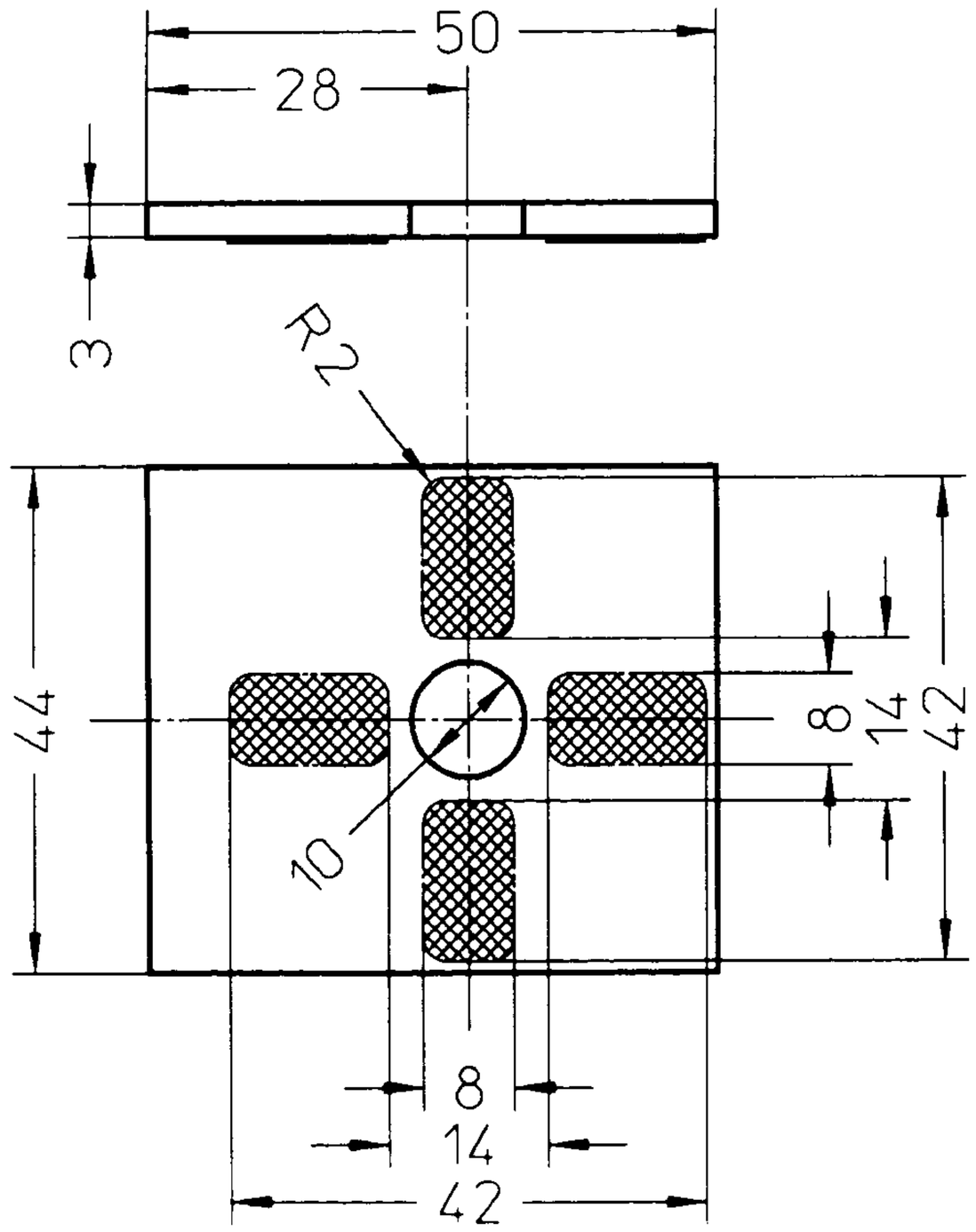
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
14	**Carriage Left Holder**	**Steel**	**1**	**2:1**	**mm**	**30.06.2000**	**1 Of 1**



Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
15	Probe Holder	Steel	1	2:1	mm	30.06.2000	1 Of 1

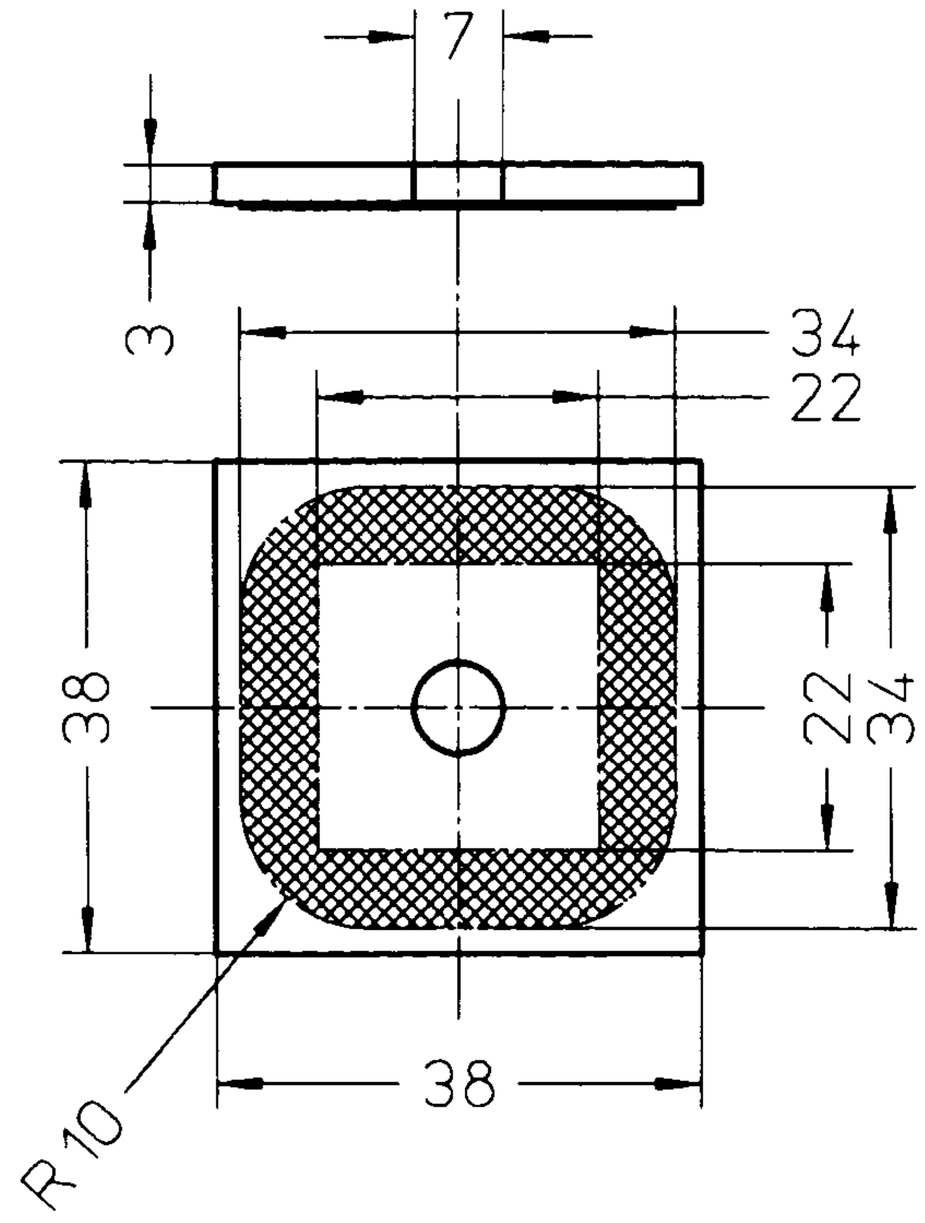


Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
16	Lower Electrode Holder	Steel	1	2:1	mm	30.06.2000	1 Of 1



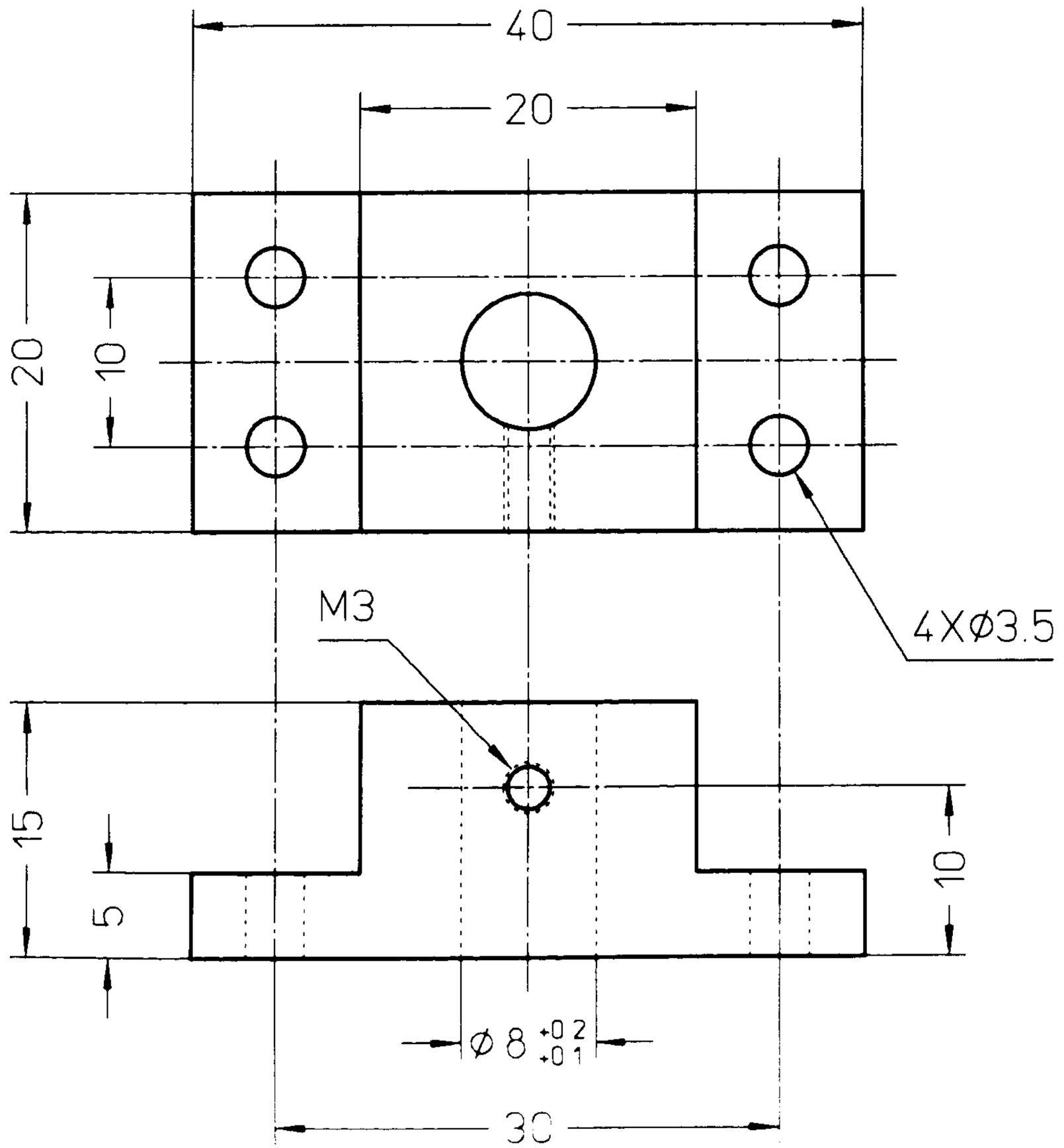
Note: Shaded areas shall be used for gold coating

Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
17	Lower Electrode	Glass plate (3mm thick)	1	1:1	mm	30.06.2000	1 Of 1

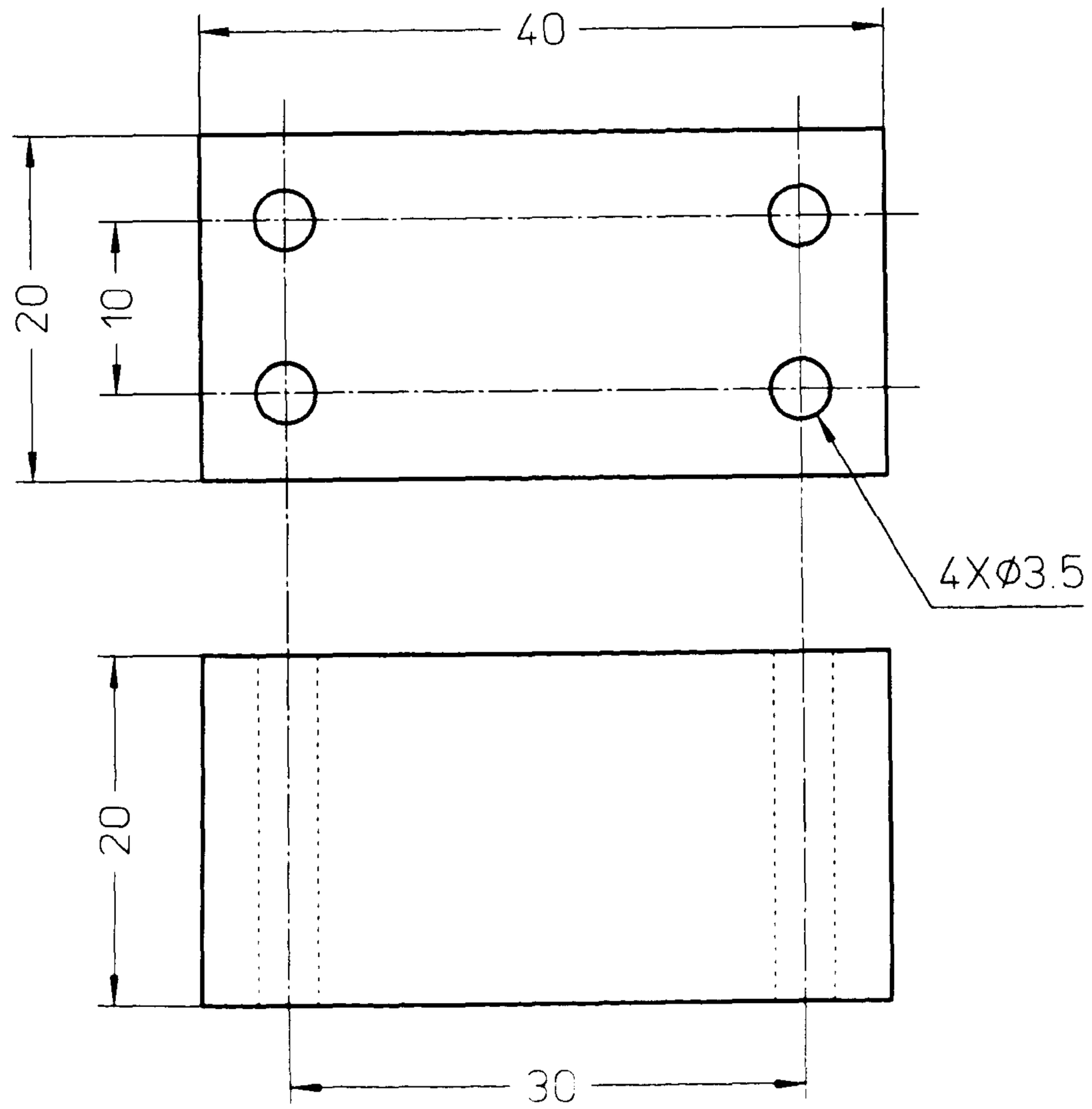


Note: Shaded areas shall be used for gold coating

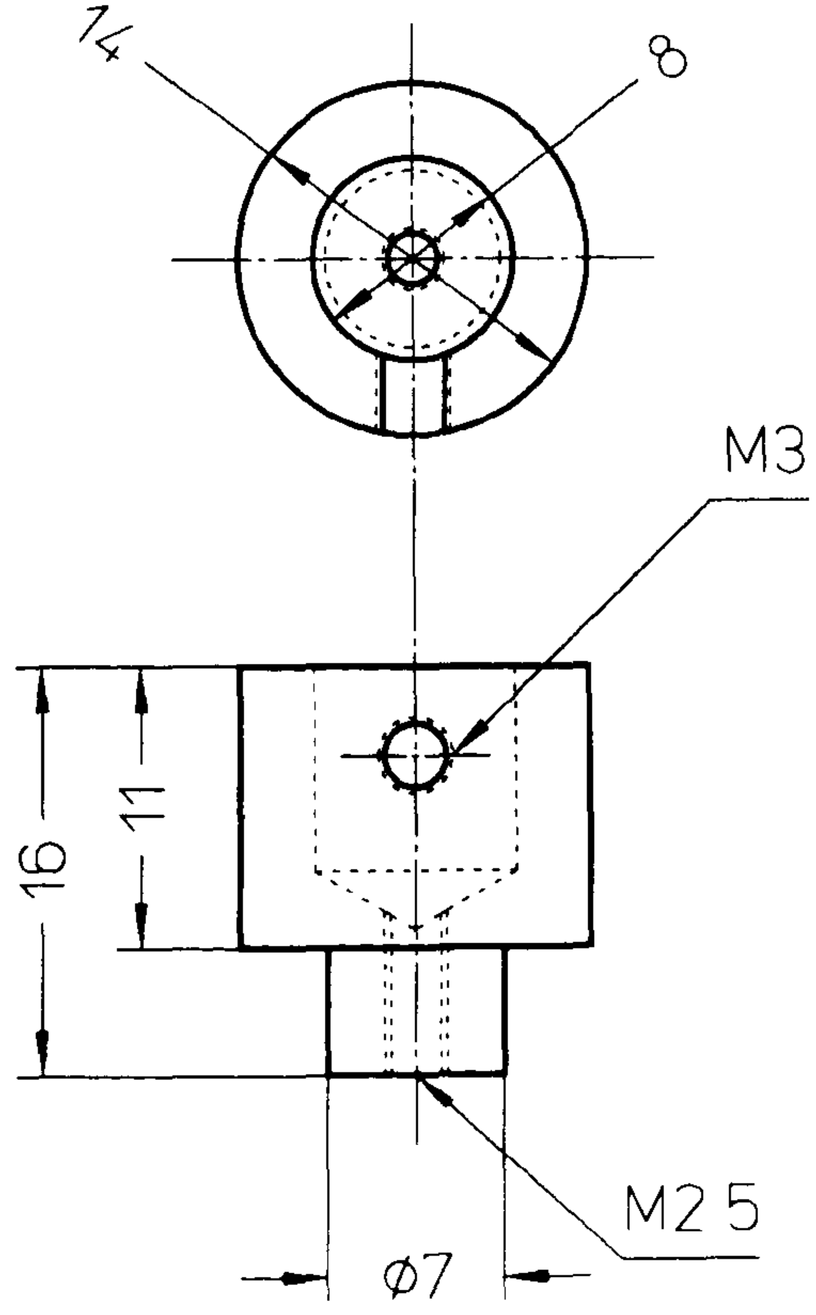
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
18	**Upper Electrode**	**Glass plate (3mm thick)**	1	1:1	mm	30.06.2000	1 Of 1



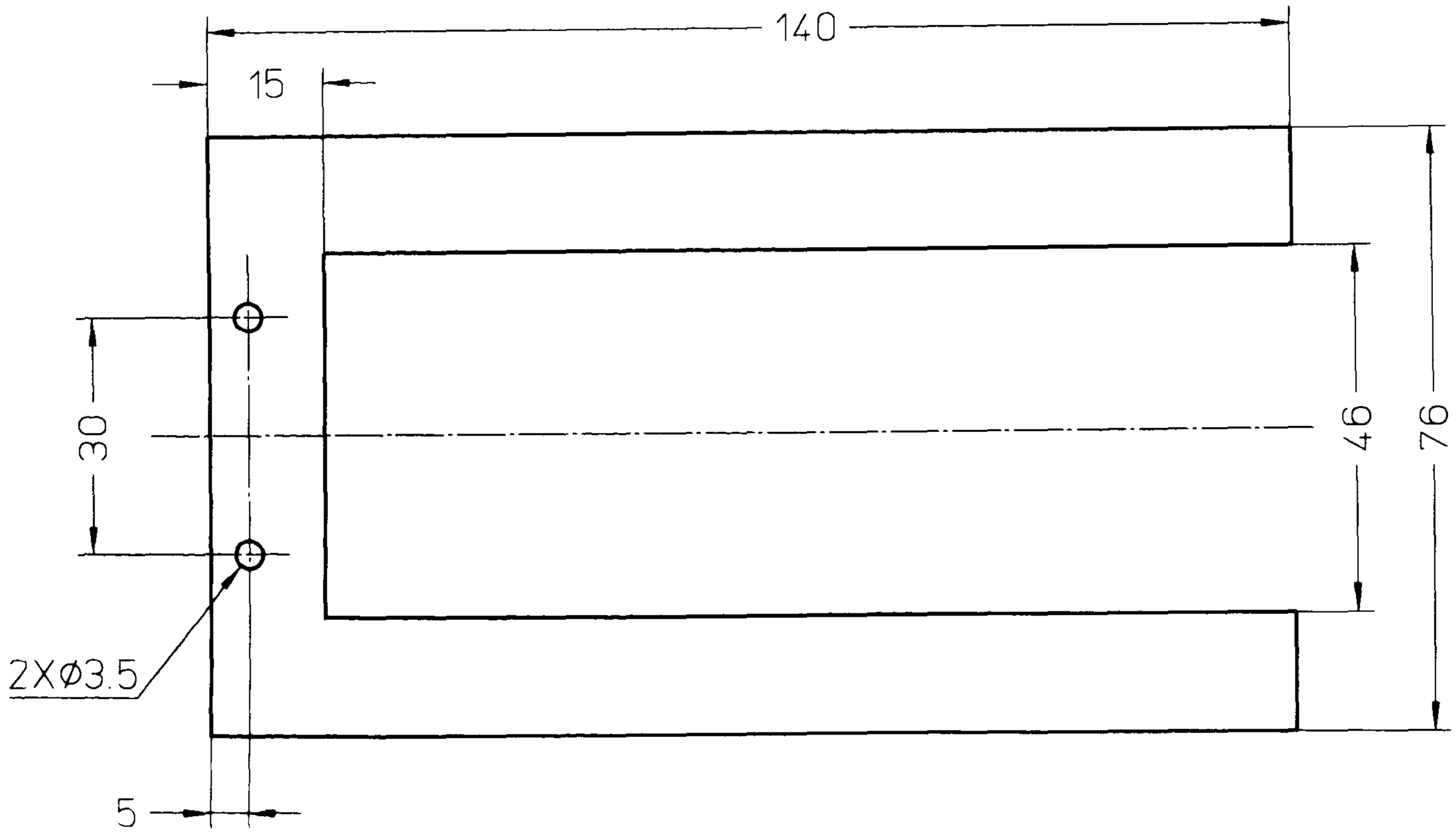
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
19	Bracket, Gauge Micrometre	Steel	1	2:1	mm	30.06.2000	1 Of 1



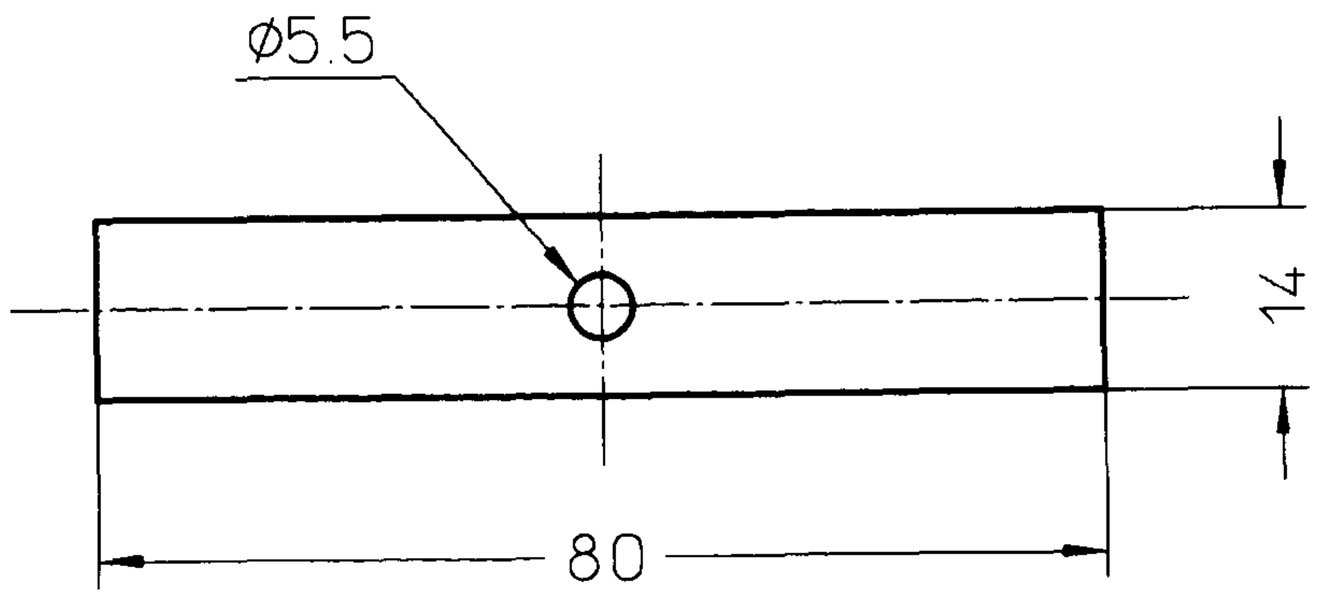
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
20	Spacer, Gauge Micrometre	Steel	1	2:1	mm	30.06.2000	1 Of 1



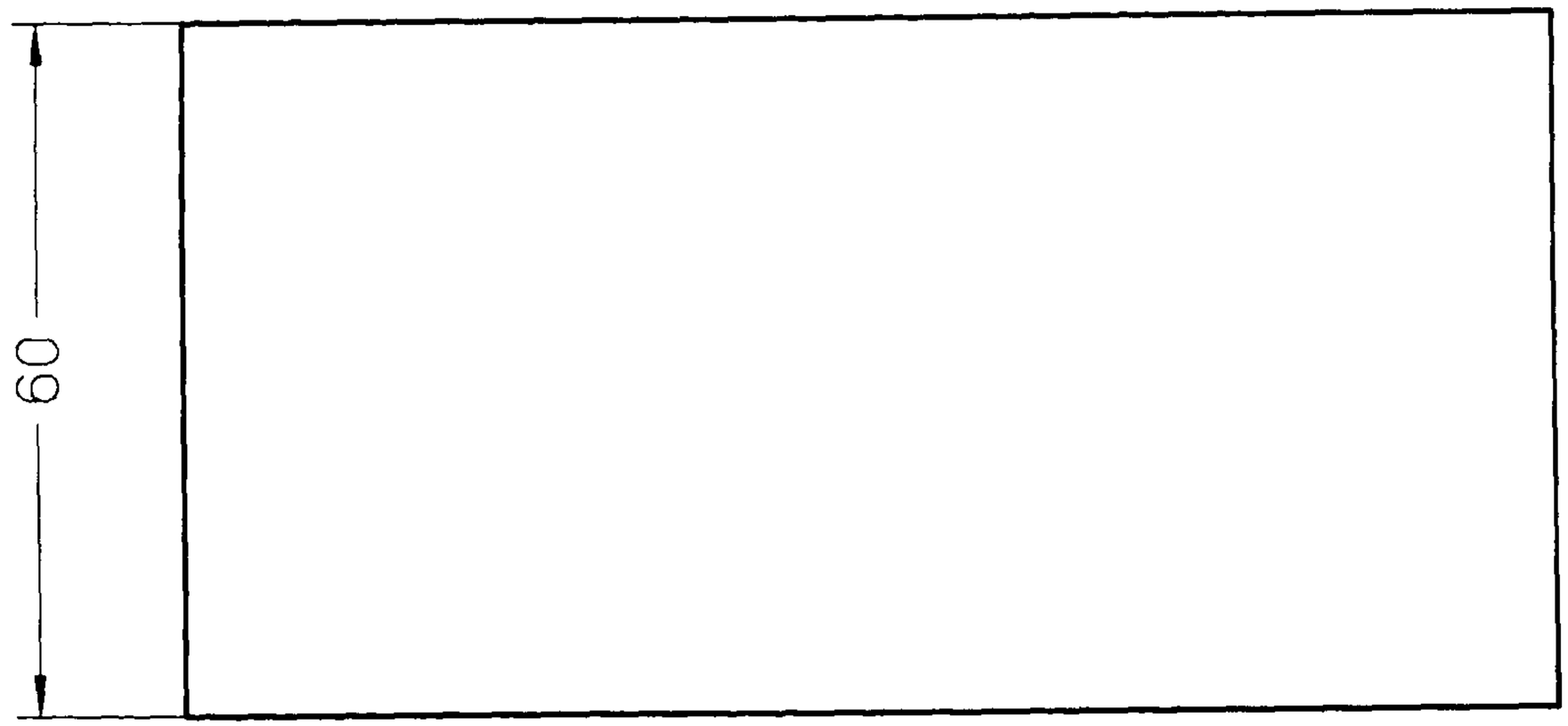
Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
21	Adapter, Slide Micrometre	Aluminium	1	2:1	mm	30.06.2000	1 Of 1

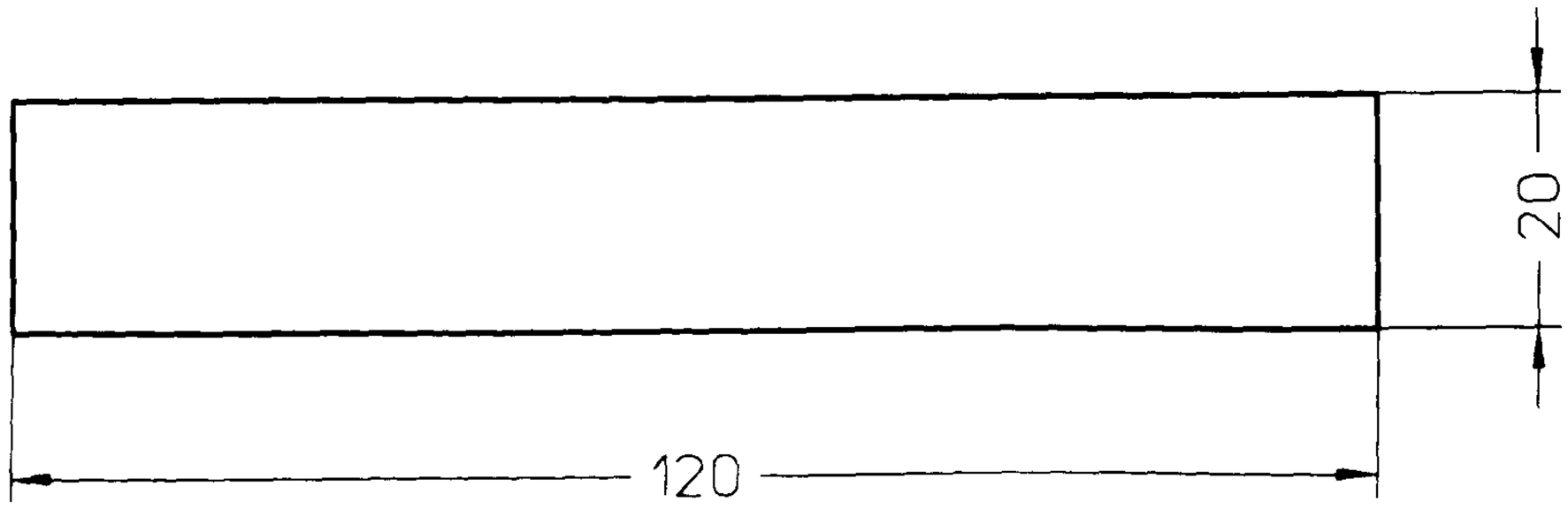


Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
22	U-Shape Holder	Steel plate (2mm thick)	1	1:1	mm	30.06.2000	1 Of 1



Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
23	Clamping Plate	Steel plate (2mm thick)	1	1:1	mm	30.06.2000	1 Of 1





Drawing No	Part Name	Material	No Of	Scale	Dimensions in	Date	Page
24	Specimen Test Plate	Steel	1	1:1	mm	30.06.2000	1 Of 1

APPENDIX C

RAW READINGS & INITIAL ANALYSES OF TESTS PERFORMED ON THE NEW TEST-RIG

these pages of data are available as electronic files in PDF formats on the CD-ROM attached below

C1	for Section 6.3	**samepont.pdf**
C2	for Section 6.4	**diffpont.pdf**
C3	for Section 6.5	**loadsurf.pdf**
C4	for Section 6.6	**sp_cl_un.pdf**
C5	for Section 6.7	**dp_cl_un.pdf**

Note:
To view the above PDF files, the Adobe® Acrobat Reader software (version 5.0) must be available on the PC. Run the "ar500enu.exe" file included on the CD-ROM below in order to install this software or download a newer version from www.adobe.com on the internet.

CD-ROM