Working Group VIII.33
"Reliability of Nondestructive Diagnostic Systems"

BAM
Bundesanstalt für Materialforschung und -prüfung

FIELD USER GUIDE

CEN WORKSHOP AGREEMENT 14747:2003

Berlin August 2008

Dieter Guelle, Mate Gaal

Foreword

The document has been written as a standard like document for supporting persons working in the field. It explains the necessary requirement for tests answering the question if the tested metal detector suits the situation of their local conditions.

The development of this Document has benefited mainly from:

- The individual interest and support of N. Mulliner UNMAS
- The working group VIII.33 BAM under the leadership of Ch. Müller
- The knowledge and experiences of M. Gaal (BAM) predominantly formulating the chapters connected with maximum detection depth and the reliability test but as well involved in the other parts
- M. Rosenthal (BAM) and P-T. Wilrich (FU-Berlin) designing the software and making it usable for the humanitarian demining community.

Persons working with this document had close contact with person working in R&D institutions, with experience of metal detector testing and development, demining engineers and demining Non Governmental Organisations using metal detectors.

This document tried to represent the current state of the art. However, the contents could be reviewed after a period of implementation in order to input more refined information.

Comments or suggestions from the users of this Field User Guide are welcome and should be addressed to the ITEP Secretariat (secretariat@itep.ws) which will be the administrator of the document.

Dieter Guelle

Content

1 Introduction

Metal detectors are still an essential part of the toolkit of a humanitarian deminer and added discrimination functions have increased their value. Even in combination with ground penetrating radar (GPR) the metal detector is the primary sensor. Metal detection is one of the few “non-contact” methods available to search for mines in most of the areas of the world where humanitarian mine clearance operations are carried out.

Metal detector trials performed in recent years [13][15][6][7][9] demonstrated the importance of having the CWA 14747 [1] for executing metal detector tests in a standardised way. Seminars for its implementation showed the complexity of the included tests and the limits for most of the organisations to carry out all described tests. However, even standardised tests did not solve the problem that a detector performance cannot be predicted in specific local conditions. Demining organisations have to carry out tests that identify the most suitable metal detector for their specific environmental conditions, the main threat to the deminers, and other local or organisational requirements. Some advice is given for the pre-selection of metal detectors for a test.

CWA 14747 Part 2 deals with the main factor influencing detector performance, i.e. the electromagnetic properties of the ground. The relevant information concerning the soil characterisation is included in Annex B.

The reduction from 30 tests in the CWA 14747 to three key tests for selecting a metal detector for field user is the main content. This part will assist in the design of the reliability (blind) test, data collection, and evaluation.

In formulating the standardized test procedures for this document, extensive use has been made of the test procedures developed by the CWA 14747 and followed during the projects “Reliability Model for Test and Evaluation of Metal Detectors” [13][15] carried out by the German Federal Institute for Materials Research and Testing (BAM) in a close cooperation with the EC Joint Research Centre (JRC) and the STEMD Project [7][9]

The authors took into account the experiences of the STEMD Project, the BAM trials and the requirements of the field users. The document includes the detailed explanation of the key tests, the reliability (blind) test, maximum detection depth, and sensitivity profile to a target. Those key tests should be carried out for defining the most suitable metal detector under local conditions as a minimum request.
A special chapter will deal with new features of metal detectors that had not been available when the CWA 14747:2003 was confirmed. They concern the discrimination between metals, magnetic and non magnetic metals, and a learn function of targets (object response).

The International Test and Evaluation Program for Humanitarian Demining (ITEP) supported that BAM improves and provides the humanitarian demining community with a field user version of the basic CWA 14747. Support has also been given by CEN BT/WG 126, by the United Nations Mine Action Service (UNMAS), which is responsible for International Mine Action Standards (IMAS and by the Geneva International Centre for Humanitarian Demining (GICHD). Close co-operation has been maintained with GICHD and UMAS, with the aim of possibly including it in the IMAS system at a later stage.

Previous standardization work on demining testing [3] has also been useful in the preparation of the documents given in the Normative References for the CWA 14747. Further documentation had been published, which will help the user in their efforts to evaluate MD: the “Metal Detector Handbook for HD” [4] and the “Metal Detector Trials Test Results and their Interpretation”[5], both published by the JRC in 2003 and 2005 respectively. Concerning the design of reliability tests and maximum detection depth measurements, the doctorate thesis “Trial Design for Testing and Evaluation in Humanitarian Mine Clearance [6] 2007 can be used for detailed scientific approach.

The document is mainly intended to be used by MACs and mine clearance organisations as well as R&D laboratories, manufacturers, operators of test and evaluation facilities, needing to procure metal detectors. Metal detector operators in the field may use it too.

The order of the testing follows a logic that begins with the planning of the tests, the execution and data gathering and ends with data evaluation and reporting. The included tests are designed to be executed in field conditions, for which targets are buried in soil. The conditions for such tests are less controlled and the limitations in accuracy are defined. The necessary equipment for the test execution will be described together with the data collection requirements, data collection and their evaluation.

The users of this document, who wish to, or are able to perform more tests than specified in here should use the CWA 14747 Part 1. Most of the tests are described for lab and field execution. A short description is given in Annex A.

Manufacturers may form one group of users of this document. As well as performing tests according to this document, manufacturers can also help others in their testing by provision of information on their product. Such requirements to the manufacturer can be found in Annex C.

Users of this document who wish to conduct a trial of various metal detectors using the tests specified, may also wish to conduct a pre-trial assessment to exclude detectors at the beginning that clearly do not meet their requirements. Such a pre-trial assessment could include one or more of the tests specified in the CWA 14747, with acceptance levels set by the users according to their own needs.

2 Scope

This document provides guidelines, principles and procedures for the basic testing and evaluation of metal detectors under field conditions and is reduced in detail to three key tests, reliability (blind) test, maximum detection depth, and sensitivity profile (footprint) to a target.

NOTE This document is to be used mainly by Mine Action Centres and mine clearance organisations. Test and R&D organisations and other entities purchasing metal detectors may use this user guide too. It is intended that the users will select the appropriate portions of this document.

This document can be used for the evaluation of all hand-held types of metal detectors for use in humanitarian demining (HD). It is intended to be used for "commercial off-the-shelf" (COTS) detectors, but many of the tests specified within it could be applied to instruments under development and foreseen for HD.

3 Normative References

This document incorporates by dated or undated reference, provisions from other publications. The references are cited in the appropriate places in the text and the publications are listed hereafter. For dated references, subsequent amendments to or revisions of any of these publications apply to this guide only when incorporated in it by amendment or revision. For undated references the latest edition of the publication referred to applies. These references are also given in the Bibliography at the end of the document. If one needs the original sources or references they are given in CWA Part 1 and in the bibliography of the document.

4 Symbols and abbreviations

BAM
Bundesanstalt für Materialforschung und –prüfung, German Federal Institute for Materials Research and Testing

EMC
Electromagnetic compatibility. Considerations of the emission of electromagnetic fields and radiation from equipment, or the immunity of the equipment to such electromagnetic fields and radiation.

GICHD
Geneva International Centre for Humanitarian Demining
COTS
Commercial off-the-shelf. A fully-developed product available on the market. Not a technology demonstrator or prototype

IMAS
International Mine Action Standard

ITEP
International Test and Evaluation Program for Humanitarian Demining

ITOP
International Test Operations Procedure, a testing standard agreed between French, German, UK and US government defence establishments.

MAC
Mine Action Centre

NGO
Non-Governmental (non-profit) Organization, such as a charitable aid and development organization

R&D
Research and development.

5 Terms and definitions

For the purposes of this document, the following terms and definitions apply. Definitions follow Draft IMAS 03.40 or IMAS 04.10 or other references if terms are defined therein.

alarm indication
A signal to warn of the detection of a metal object. The indication can be visual and/or auditory. A positive alarm indication is repeatable under the same conditions and is not intermittent.

NOTE For metal detectors with auto-zero or other functions giving a dynamic mode response, relative motion between target and detector is required to give an alarm indication.

alarm indicator
The device used to generate the alarm indication, often an acoustic device giving a characteristic sound.

blind test
A test in which the detector operator does not know details of the location or the depth or nature of the target(s) being sought.

controlled laboratory tests
Tests performed in conditions where the external factors that may affect a detector are controlled. For example constant or controlled temperature, movement of detector using motorized and even automated scanning mechanisms to ensure control of detector position and sweep speed.

demining
Activities which lead to the removal of mine and UXO hazards, including clearance.

detection
The discovery or finding of a metallic object. The operator is made aware of the detection of a metallic object by means of a true alarm indication on an alarm indicator.

detection halo
The circle around the actual location of a test target, within which an alarm indication is considered a true indication of detection when performing blind detection tests.

detection reliability
Detection reliability is the degree to which the metal detector is capable of achieving its purpose, which is to have maximum capability for giving true alarm indications without producing false alarm indications.

dynamic mode
Some detectors use an auto-zero or high-pass filtering mechanism so that unchanging signals do not cause an alarm indication. This is a way of reducing noise signals from magnetic soils, for example. Only changes in signals produce an alarm indication. In this document, this is defined as a dynamic mode of operation.

electrical conductivity
The ease with which an electrical current flows in a medium. Measured in Siemens per metre (S/m).

false alarm indication
Alarm indication not caused by the presence of a metal object.

field tests
Tests to determine the performance of a metal detector in conditions that are close to real operating conditions.

forward direction
The direction perpendicular to the sweep direction for a metal detector in normal use. Typically the direction in which the operator faces while using the detector.

ground compensation
An operating function of a metal detector, designed to reduce or eliminate alarm indications from noisy soil while maintaining its detection capability for metal.

heterogeneous electromagnetic properties
A medium whose electrical conductivity and/or magnetic susceptibility are not the same at all points within the volume of the medium interrogated by a metal detector and therefore whose effect on a metal detector can vary with location is defined as heterogeneous within this document.

homogeneous electromagnetic properties
A medium whose electrical conductivity and magnetic susceptibility are the same at all points within the volume of the medium interrogated by a metal detector and therefore whose effect on a metal detector does not vary with location is defined as homogenous within this document.

immunity (to an electromagnetic disturbance)
The ability of a device, equipment or system to perform without degradation in the presence of an electromagnetic disturbance (see [14]).

in-air tests
Tests to determine the characteristics and performance of a metal detector, without the influence of **soil**.

in-soil tests
Tests to determine the characteristics and performance of a metal detector for targets buried in soil.

less-controlled tests
Tests performed in conditions without the same degree of control over test variables as in controlled laboratory tests. Such tests may be performed indoors without temperature control, or outdoors. The detector may be manipulated manually, but usually with the assistance of jigs and timing devices to control position and sweep speed.

magnetic susceptibility
The degree to which a medium becomes magnetized in an applied magnetic field. Measured in dimensionless SI $10x10^{-5}$ units.

maximum detection height
The maximum height above a test target at which a metal detector at given settings produces a true alarm indication due to that target, this term is preferred to be used for in-air detection.

maximum detection depth
The maximum depth above a test target at which a metal detector at given settings produces a true alarm indication due to that target in soil, this term is preferred to be used for in-soil detection.

metal detector
A device which uses the principles of electromagnetic induction to reveal the presence of metal in its vicinity.

mine
munition designed to be placed under, on or near the ground or other surface area and to be exploded by the presence, proximity or contact of a person or a vehicle.

noisy soil

Soil that by its composition and/or layering or structure, reduces the performance of metal detectors, to the extent that the operator’s task is made more difficult. This performance reduction is most likely to be a reduced sensitivity to metal and/or producing signals that are not easily distinguished from signals from metal.

open test (non-blind test)
A test in which the detector operator knows details of the location or the depth or nature of the target(s) being sought.

realistic test target
A test target designed to simulate the geometry and material properties of mines or of the metal components contained in mines. Realistic targets also include real mines, mines without explosive or otherwise rendered safe.

sensitivity
The sensitivity of a metal detector is the measure of its capability to detect metal objects. A detector having a high sensitivity can detect small metal objects at a given distance that may be undetected by one having a low sensitivity. The sensitivity of all detectors reduces with distance from the detector sensor head. In many detectors the sensitivity may be adjusted. Within this document, sensitivity is measured in terms of the maximum height of the detector head above a given metal test target at which the target can be detected. Sensitivity may also be expressed as the minimum target (in terms of size, shape and material) that can be detected at a given height
above target.

sensitivity profile (footprint)
The sensitivity profile of a metal detector is a plot of the variation of detection sensitivity with location beneath the detector sensor head along one axis of the sensor head or in two dimensions. The location and extent of the area of maximum sensitivity is of particular interest when specifying the maximum step between detector sweeps to ensure full coverage of an area.

sensor head
The part of the metal detector (usually a flat coil arrangement) that generates and receives alternating magnetic fields in order to detect metal objects.

sensor plane
The plane of the detector sensor head (typically a coil) that is kept parallel to the ground in normal operation.

static mode
Some detectors do not perform auto-zero or high-pass filtering on their output. When an alarm indication is given due to the proximity of a piece of metal, it continues for as long as the metal is there, even if the detector is held motionless. In this document, this is defined as static mode of operation

soil
The medium in which mines may be buried in the ground

sweep direction
The direction in which the sensor head of a metal detector is moved over the ground in normal operation. This is typically a side-to-side (transverse) direction in the plane of the sensor head, when the detector is held in its normal position.

test
Determination of one or more characteristics (of a metal detector) according to a procedure

test lane
A metal detector test area (usually long and narrow) that mimics the lanes into which minefields are divided for clearance operations.

test target
An object that is used to test the detection performance of the metal detector. This is a metallic item that can be intended to mimic the response of a mine or mine component, or it can be a simple metal object to be used in sensitivity measurement.

trial

A series of tests organized in a systematic manner, the results of which lead to an overall evaluation of a component, of equipment or of a system.

true alarm indication
Alarm indication caused by the presence of a metal object.

unexploded ordnance (UXO)
Explosive ordnance that has been primed, fuzed, armed or otherwise prepared for use or used. It may have been fired, dropped, launched or projected yet remains unexploded, either through malfunction or design or for any other reason.

6. General testing principles

6.1 Factors influencing metal detector performance

The most important characteristic of metal detector performance is the capability for detection of metal objects at distance. The smaller the metal objects and the further it is from the detector sensor head, the lower is the probability that the object will be detected. The relationship between the size, shape, orientation, and material of a metal object and the distance at which it can be detected is therefore used to determine detector performance.

Within this document, detection capability is measured in terms of the maximum height for establishing the reliability of selective capabilities in air and maximum depth in soil of the detector head above a given metal test target at which the target can be detected. By implication, detection capability may also be expressed as the minimum target (in terms of size, shape and material) that can be detected at a given height/depth above target (and in a given orientation).

Maximizing the detection capability of a detector in order to detect small or distant objects is however not the only consideration for optimizing detection performance in demining. All detectors suffer a certain amount of unwanted noise: for example from external electromagnetic fields, from the electronics of the instrument itself and due to the electromagnetic properties of the soil over which the detector is used. Evaluation of detector performance therefore takes these effects into account, particularly the latter. The capability for detecting metal targets at a controlled depth in soil is the key basis on which comparisons can be made between detectors.

There are several factors influencing the sensitivity of the detector essentially to the extent that the detector will not perform at all. Most of these factors are mentioned in the CWA 14747:2003 Part 1. To each of these factors there is a corresponding test described in that standard. These factors are:

- sweep speed,
- time after switching on the detector (test of sensitivity drift),
- search head orientation,
- shaft extension,
- moisture on the search head,
- temperature extremes,
- temperature shock,
- battery life,
- soil or other medium,
- target properties.

Another factor has a proven an important influence: the operator [6]. This influence is always present and it has to be taken into account in planning all tests.

Mines are the object of most clearance operations. Their metal content (mass), type of metal, and their shape will influence their detectability. The detectability of the metal part depends on the surface of the metal object. For instance, a ball-like shape is most difficult to detect in comparison to any other shape with the same mass, speaking about parts used in mines [5].

The Human Factor influence can be best measured in a blind test. Operators rarely achieve the same results using the same detector in well designed blind tests. The differences between the operators are usually not significant if the operators are skilled and currently active deminers. It is not known which other personal qualities influence the performance of the operator. The operator’s abilities (experience, technical knowledge, health, mental situation etc.) together with the environment will have influence on the results.

6.2 Purpose of the specified tests from the CWA 14747:2003

IMAS 03.40 gives categories of trials in which testing and evaluation of mine action equipment may take place. Of the four categories given, this guide applies to the "consumer report" and "acceptance trials" categories:

– A "consumer report" trial aims to test equipment against standard general tests, so that the results are of general interest to metal detector users.
– An "acceptance trial" aims to test equipment against specific requirements of a customer, in order to make purchasing decisions, for example.

This guide was established primarily because of a need for “acceptance trial” aims. It is more designed to be used to test against specific requirements or specific local conditions. This means that a metal detector can be tested in conditions similar to those that are likely to be encountered during operation. By their nature, these field tests tend to be less controlled, although appropriate measures are taken in an effort to ensure that the tests specified form a valid basis for comparison of detector performance. These tests may also be used for collecting data to a “consumer report” if the targets are of common interest to a region or standard targets.

The previous section presented most of the factors having an influence on the detector performance. Some of these factors have a more important influence than the others. These are the target (e.g. mine), the soil (or other medium), the setup and the operator. Their influence is more relevant because they always affect the performance of metal detectors, while the other predictor variables in most realistic situations are constant most of the time. This has been experimentally confirmed in two recent trials [6] [9]. If the test goal is to compare metal detectors and their ability to detect a certain target, then the soil, the setup and the operator have to be varied in the experiment. If they are not varied, they would introduce systematic errors and lead to wrong conclusions.

For the Field User Guide the following tests had been selected as key tests:

– Maximum detection depth in soil (Chapter 8)[18]
– Reliability (blind) test (Chapter 9) [20]
– Sensitivity profile (footprint) (Chapter 10)

The **maximum detection depth in soil** allows, in a very short time, the establishment of the detection capability of the detector to specified targets or threat in the area/region. The influences of the soil, setup and the operator are included in the maximum detection depth measurements.

The **reliability test** includes most of the factors influencing on the detector performance:

– the detector with its intrinsic capabilities,
– the environment, including the rules for clearance operations, and
– the human being.

A Reliability Test is a blind test, meaning that the operators do not know the positions of the targets. The maximum detection depth measurements are open tests, meaning that the operators know the positions of the targets. Reliability tests have the advantage that they are the closest one to realistic mine-field situations. They include most of the influencing factors on the detector performance. The Human Factor influence can best be measured in a blind test. Beside this, the false alarm rate can be measured only in a reliability test. There is a disadvantage compared to maximum detection depth measurements and that is that it requires much more time.

The knowledge about the **sensitivity profile**[1]has operational importance for the establishment of the safe search head advance to the main threat in the clearance operation. Following this knowledge, the deminer will not miss targets when searching for mines in a clearance operation in a proper way.

With the results of the key tests contained in this document, and following the established rules for mine clearance, an optimum of safety for the deminer and end-user of the land can be achieved.

6.3 Test of new technical metal detector features

[1] Sensitivity profile – also called footprint, sensitivity area or sensitivity cone.

Since 2003 new technical features appeared on the market for metal detectors involved in humanitarian demining. One feature has already been used for a long time for treasure hunting but only now first attempts are going on to use such equipment in the area of humanitarian demining. The technical background is that the response of magnetic metals is quite different to non magnetic metals. This different response can easily be displayed on a screen and creates a certain capability for discrimination between such metals. Some manufacturer have concluded that this kind of discrimination is reliable enough to be used in humanitarian demining. But they are also warning about the limits of such a system, that's why most of the manufacturers of such equipment do not sell them for humanitarian demining.

The other promising new function is the capability of a metal detectors is to possess a response learning function (RLF). In this case, the detector needs an original target and will save the signal response in its memory for comparing the response of metal parts to the saved sample.

Unfortunately, most of the mines have not only one type of metal, but a composition of different metal parts. The aim of testing those features is to make it clear to the user that there are limits and in which case the feature is limited. The second feature has the advantage of being independent, which kind of metal had been used in mines. Additionally it includes the real response excluding the aging[2] of mines i.e. rust and other processes that reduce the response to the magnetic field of the detector.

6.4 Guidance for field evaluation and assessment

6.4.1 General

The tests specified in this document are intended for use in the evaluation of metal detectors in field tests, for selection of detectors prior to procurement. This testing fits into the "acceptance trial" category as defined by IMAS 03.40. It also may be used for establishing general understanding of detector performance when standard or common targets are used for the comparison of the detector performance.

The results of laboratory tests or another field situation, remote from the required clearance site, provide an essential background for the pre-selection of suitable detectors for a particular user to test. The particular local conditions where humanitarian mine clearance is required will however create specific requirements – perhaps more demanding than those in previous trials. Field evaluation and assessment of metal detectors is therefore normally required before any final selection of a metal-detector suitable for a specific area/region.

Field evaluation and assessment includes conducting tests so that the end-user is able to evaluate the relevance of the results in any particular local conditions. In this way, the field testing may be seen as "confidence testing" but also provides the opportunity to achieve more than that.

Field-tests are intended to:
- assess detector performance in the potential user's field conditions;
- provide users with the opportunity to assess ergonomic preferences and ease of actual use;
- build confidence in the ability of any selected detector to meet the user's performance needs;
- allow evaluation of the field-relevance of the earlier test results;
- extend existing knowledge of detection capability for recorded detector/target/soil combinations.

6.4.2 Pre-selection of detectors

When an acceptance trial is to be performed with the aim of selecting the best detector or detectors to meet particular user needs, previous results from laboratory tests and any field evaluations made by other user groups may be used to create a list of the detectors to be tested. The main pre-selection criterion will often be the capability for detection of relevant targets in appropriate conditions (for example, similar soil properties).

[2] Rust and other factors influence the electromagnetic response of the primary field of a metal detector. In general this process reduces the response of mines, make them more difficult to find with a detector.

6.4.3 Repeatability of field testing

Field assessments are performed to determine how detectors perform in specific local conditions. Nevertheless, these tests need to be done in a common, repeatable way so that the results are useful more generally. Results may then be used by other groups for comparison or by the same group at a later date. Following the testing and reporting requirements of this document ensures that this comparison is possible.

6.4.4 Content of field testing

The tests described in here are the recommended minimum amount of tests for the detector evaluation and operational needs. All tests to be performed in a trial for field assessment are defined in the CWA 14747 chapter 5.1(ANNEX A) [33], in the "field" column of Table 1. The key tests described, also included in this list, could normally be performed in any trial, depending on available time and resources.

6.5 *Test geometry*

In the tests specified, the convention for describing the geometric configuration of the detector, target and their relative movement shall be based upon that for a detector in normal field use. The convention shall be that the target be fixed and the detector swept from side to side (transversely) above the target, with the plane of the detector sensor head kept horizontal – parallel to the soil. The use of terms in the test specifications shall be understood as applying to this conventional configuration.

If the design of a detector means that its normal mode of use is different to the convention above, the tests shall be changed accordingly. For example if a detector is swept forwards and backwards in normal use, then it shall be used in this way in the tests.

For detection tests in soil, the detector sweep height is measured between the underside of the sensor and the soil level. The target depth is measured between the soil level and the top of the target. The total height of the sensor above the target is therefore the sweep height plus the target depth. Mostly it is impossible to determine or to keep a constant sweep height above the soil surface. In minefields the deminers try to keep the search head as close as possible to the surface. Therefore, the detection depth shall be measured from surface level to the top (highest point) of the target.

The above established convention may be changed if tests and other investigation does not allow moving the search head or are carried out where a target distance to an object in soil should be established using the CWA Part 2 mentioned assumptions about signal symmetry and decoupling of target and soil response.

Chapter 9 of CWA 14747 Part 2[3] Effect of soil on field performance.

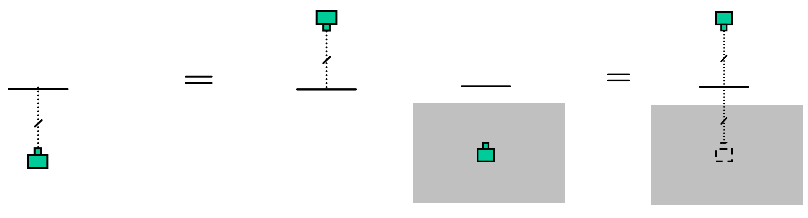


Figure 1 and 2

For measuring the sensitivity to an object or sensitivity area of a metal detector the assumption is that the magnetic field of the metal detector is symmetric to both sides of the detector search head, so measurements should have the same result on both sides. Therefore measurements can be done in air if the detector is set up to the soil and does not lose the contact during the process as shown in Figure 2

The second assumption is that the response of the soil and the target decouple. If soil electrical conductivity and magnetic susceptibility are low enough and the soil and target responses decouple. This means that the total response is the sum of:

- the response of the target in air and
- the response of the soil in absence of target.

The practical value is considerable and can make tests much faster as before. The empirical confirmation should be done for confirming the assumptions. Further the tests can be carried out in the way that the target will be moved and the search head stay fixed at his place.

[3] The figures 1 and 2 are taken over from the CWA 14747 Part 2

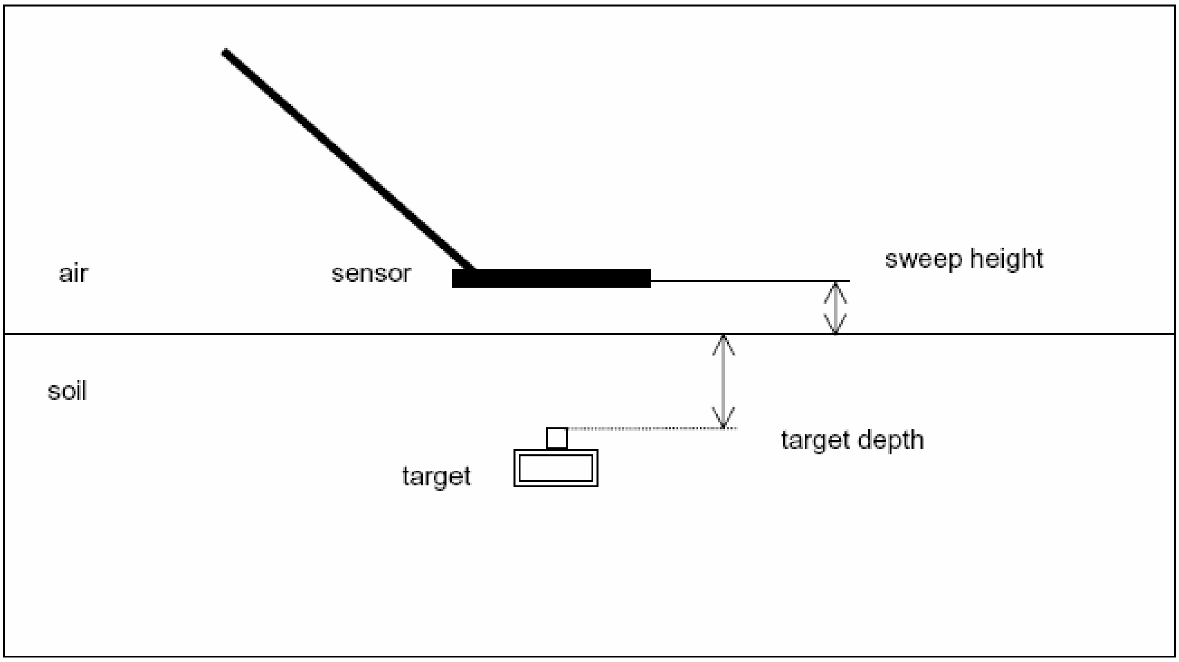


Figure 1: Sweep height and determination of target depth, which is equal to detection depth.

6.6 *Requirements to operators and training*

The operator has influence on the detector performance and should fulfil certain general conditions before he can be used in field trials.

The testing operators shall be capable of using the metal detectors as intended by the manufacturer and they shall be representative of the operators that would use the detectors in the field. The operators shall be currently active deminers, since they represent the persons who will actually perform clearance operations. They should have at least 6 months of continuous work in minefields directly before the trial.

Operators shall be trained on the proper use of each metal detector that they are to operate during the test. They shall have adequate time to become familiar with the operation of the metal detector and to attune their ears to the sounds of the device. At least two days of training for each detector model shall be performed.

Depending on the test and its objectives, other conditions may apply but they shall be as such mentioned in the test description.

6.7 *Criterion for determining detection*

This document applies to hand held metal detectors of the types used in humanitarian mine action. Most of these detectors are designed to be easy to use in difficult conditions, by operators without a high level of technical education. The detection of metal is therefore usually indicated by a simple acoustic alarm. In the field it is up to the trained operator to interpret whether the sound he hears is an alarm indication. As this is the actual situation in the field, the tests specified in this document requiring detection of a target are also based on recognition of the alarm by the operator.

Many of the tests specified in this document require the determination of the conditions in which the detector is just detecting a target. At such low signal levels, the alarm indication may well be intermittent. Some judgement is therefore required to determine whether a target has or has not been detected, for example when determining the maximum distance from a target at which detection is possible.

The person performing the tests shall first become familiar with alarm indication for detection of each detector being tested.

The criterion for detection as obtained from the manual or otherwise from the manufacturer shall be used for guidance within the tests specified in this document, insofar as it is consistent with the criterion below.

The criterion for detection shall be that the detector gives a consistent, non-intermittent alarm indication, repeatable under the same conditions and audible to a person with normal hearing (for acoustic alarms). The signal should be audible in both sweep directions five times in a row.

Other than audible signals shall be clearly defined in the manufacturers' manual for the particular interface to the operator.

6.8 *Measurement of Soil properties*

Where possible the magnetic susceptibility and the electrical conductivity of the soil(s) in the test area(s) shall be measured in situ with calibrated measurement instruments, together with the moisture content in the top layer of the soil, ie. to a depth of 0,2 m. The measurement instruments used shall be reported (brand, type and settings, if applicable). These measurements shall be made at a number of locations on the test area(s) (at least one measurement in every 0,25 m^2 of the test area). Spatial variations shall be reported. The measurements shall be repeated after any occurrence that may change the soil properties, e.g. rain. The magnetic susceptibility shall be given in dimensionless SI units, where possible the frequency dependency of the magnetic susceptibility should be established, i.e. are the measurements different if they are done with different frequencies. The electrical conductivity shall be measured in S/m and the moisture content as a percentage by volume together with the depth of the measurement(s).
Soil response to electromagnetic waves of metal detectors can also be measured by a metal detector. The preferred type is a static detector with continuous sensitivity calibrated to a repeatable sensitivity, by the help of the test piece or other defined objects. Then the detector should be brought down to the ground and the height when the detector starts to signal gives a measure that shall be recorded in centimetre. With an increase of height the difficulties of the detector will increase. This type of measurements includes the influence of the magnetic susceptibility on the metal detectors.

NOTE the detector shall not be using ground compensation.

Simple description of soil properties can be taken over from the CWA Part 2[4].

6.9 *Test targets*

Most manufacturers provide a metal test target to check the proper functioning of their detector. However the test targets provided are not standardized. Some manufacturers use a metal ball, some a pin, others an aluminium plate. It is well known that the distance at which a metal object can be detected by a metal detector depends on the object's size, shape, material and orientation - among other parameters. Thus the selection of a suitable common set of targets is very important for the purpose of comparing performances of various detectors and results of tests conducted at different times and by different agencies. As a general rule for a consumer report the main threat in clearance operations of the country/region should be used for establishing the suited detectors for the clearance operations.

Within this document, test targets are specified for each test. The targets used may include the following categories:

1) Simple geometric targets.

2) Targets designed to simulate metal components of mines.

3) Test targets designed to simulate a whole mine (generic or specific).

4) Targets that are metal components of real mines or UXO

5) Targets that are real mines or UXO.

---

[4] The work on this standard for soil characterisation for metal detector and ground penetrating radar (GPR) performance evaluation is in progress. It will be finished in spring 2008 under the lead of RMA Belgium.

Within this document the most important distinction is between the first type and all the others. The first types are not "realistic" insofar as they are not intended to give a response like that of a mine. Simple geometric shapes can however be used for families of parametric test targets with the same shape but different size. Parametric test targets can be used to measure the minimum detectable target size as a function of detector height above target – larger targets being detected at greater height.

The other types of target may all be used in a common way. The maximum height at which the target is detected can be measured, or it can be recorded whether an alarm indication occurs at a given detector height above target or at a given depth in soil.

The details of targets recommended to be used are found in Annex C.

6.10 *Requirements for recording of test results*

All relevant data relating to a test shall be recorded. This may include the following, depending on the test being performed:

1) Detector make, model, serial number and operating program (software) version if known.
2) Detector settings (as appropriate).
3) Detector operator.
4) Date, start time and end time of test.
5) Target details, including dimensions, materials and drawings where appropriate.
6) Environmental conditions; temperature, humidity, meteorological conditions (as appropriate)
7) Soil type, condition and properties (see Annex B).
8) Target burial depths (top of target to soil surface).

7 Tests included in CWA 14747 2003 (general approach)

The CWA 14747 Part 1 describes the conditions for testing in controlled and less controlled conditions, establishes the accuracy of data collection and the reporting in general and for every test. A summary of CWA Part 1 is at Annex A.

8 Maximum detection depth

8.1 General principle

This test is a modification of the fixed-depth detection test described in Section 8.4 of the CWA 14747:2003. The objective of this test is to determine the detection capability of the metal detector for buried test targets at fixed depths in a given soil. The detection capability is expressed as the largest depth at which the target can be detected in that soil. The test is an open test, i.e. the target positions are clearly marked and known to the detector operators. This test is primarily designed for targets like inert mines or mine surrogates to be used in a field trial. For smaller targets like ITOP inserts, spheres, other small targets the tests described in Section 8.3 of the CWA 14747:2003 are appropriate, for reducing a void effect. They should be repeated with several operators and repeated setup.

Repeating the measurements with several operators is essential for this test, because the maximum detection depth has a large variability caused by the operator, the setup and other sources of experimental error [6] [9]. Only repeated measurements can give some information about the experimental error, which is present in all experiments. The least number of operators shall be four (4), but it is recommended to use as many as possible. Repeated measurements after a new setup and with the same operators may also be used. They are strongly recommended if the number of operators is small.

8.2 Test area, test targets

The requirements for the test area are defined in Section 6.7. Electromagnetic and other soil properties influencing the detector performance shall be measured as defined in CWA 14747 Part 2.

The targets to be used for this test are real disarmed mines or simulated mines.

The targets shall be placed at depths in steps of 1 cm. The positions of the targets shall be marked on the surface above the targets with non-metallic markers (e.g. plastic discs). Adjacent targets shall be separated by a minimum of 0,5 m. Targets shall be buried at least 0,5 m from the edge of the test area to ensure that no interference is caused by the edge. If targets are used that are not rotationally symmetrical, the orientation of these targets when they are buried shall be recorded and reported. The depth of the targets is defined as the distance from the soil surface to the top of the targets when buried.

The locations, depths and orientations of the targets shall be measured whenever they are dug up for removal from the test area and the measured locations, depths and orientations shall be recorded. As well as the exact target locations, the extent of the detection halo as specified in Annex F[48] shall be indicated (e.g. with markers of another colour or with a disc-shaped marker under the actual target marker).

8.3 Measurement procedure

The detector shall be set up according to the procedure described in the user manual of the manufacturer. If applicable, the settings used shall be recorded. The detector shall be moved manually, as close as possible to the ground, as it would be used in clearance operations.[5] The operator shall decide whether an alarm indication occurs when the detector head moves over a target marker. The operator shall also report alarm indications not occurring above a target marker. Any such false indications shall be investigated. If found to be due to the soil (rather than an extraneous metal piece), the detector shall be re-adjusted where possible, to ensure that the soil being used does not trigger the alarm. The test shall then be restarted.

For each target, record the depth and "detected" or "not detected" as appropriate. If an alarm indication occurs that can not be related with one of the known targets, i.e. occur outside the detection halo of the targets, these indications shall be recorded with their location for further investigation.

This measurement procedure shall be repeated with several operators. The experiment shall be designed in a way that the order of execution of measurements is randomised, regarding the operators and regarding the detectors. The reason for randomisation is to avoid any systematic influences related to time, for example tiredness. An example of an experiment with such a design is presented and discussed in Annex G [49].

8.4 Test results and reporting

In the previous section a single pass of an operator over the complete set of depths for one target was called “measurement”. The measurements with all operators on the same target are called “test”. There are two possible outcomes of a maximum detection depth measurement:

[5] The requirement of the CWA 14747:2003 that the detector search head be swept over the test area at a height of 30 mm is not necessary.

1. All targets up to a certain depth are detected and targets buried deeper than that depth are not detected.
2. Some deeper buried targets are detected, while some shallower buried targets are not detected.

If the measurement outcome is of the first kind, the data analysis is relatively simple. The largest depth at which the target has been detected in a single measurement shall be called maximum detection depth of that measurement. If the achieved results are normally distributed, no measurement gaps between the depths at least the following quantities shall be reported:
1. the average of all measurement results for that detector, here marked d_m
2. the standard deviation of all measurement results for that detector, here marked s
3. the number of measurements, here marked n

The average maximum detection depth d_m increased by 5mm is an estimate of the target depth at which the probability of detection is ½. This conclusion about the probability of detection is valid only for the conditions present in the test. This means that the actual probability of detection in blind test is most likely to be lower, because the operators would not know the positions of the targets. The increase by 5mm is necessary as a correction of the systematic error caused by the 1-cm size of the step between the depths. The standard deviation s of the maximum detection depth is a measure of the stability of the detector's performance.

If the measurement outcome is of the second kind, with some deeper buried targets detected, while some shallower targets not detected, the data evaluation is more complex. In that case it is recommended to create POD curves using the generalised linear model described in Annex J [55].

9 Detection reliability tests

9.1. *General principles*

Reliability tests are blind tests defined in Section 8.5 of the CWA 14747:2003. The objective of detection reliability tests is to evaluate the detection reliability of the mine detection device when used by an operator who does not know the location of the targets. Detection reliability is the degree to which the mine detection device is capable of achieving its purpose, which is to detect all targets of the identified threat with a minimum number of false alarms. Detection reliability tests include the influence of most factors that affect the performance of detectors, including a large part of human factor influences. They are also the only tests that can evaluate the ability of detectors to deal with false alarms. Detection reliability tests are typically performed in or near to an area to be cleared of mines, in representative local soil and with targets representative of the local mine threat. Such tests may be part of either "consumer report" trials or "acceptance trials".

In a reliability test, targets are placed in metal free lanes at positions not known to detector operators. While searching, the operators mark the places of indications and, later, test personnel measure and record the positions of the markers. A target is considered to have been detected when a marker is dropped within a prescribed radius around the true target location. The area defined with this radius is called a detection halo and it is defined in Annex F.

The operators shall use coloured markers to mark the locations of alarms. The positions of their indications shall be measured and recorded. They shall be compared with the actual positions of test targets and the detection halo radius. If a marker lies within the detection halo, the target is detected and the indication is called a true positive indication. Markers outside the halo radii of test targets are counted as false alarms and called false positive indications. An optimised device is one that maximizes the number of true alarm indications and minimizes the number of false alarm indications.

9.2. *Lane preparation*

A site shall be chosen for the tests at which:
- the native soil at the test site is representative of an area to be cleared of mines (the test site may be within that area), or
- soil representative of the region requiring mine clearance has been placed in test lanes.

Note that if soil is moved, its response to a metal detector will not necessarily be the same as it was prior to being disturbed. Check it with the detector before and after the movement. The soil in which it is most difficult to detect mines found in the mine-affected area should be used where possible. Either the test site should be set up in an area of that soil, or that soil should be transported and used in a test lane.

The vegetation over lanes shall be removed or cut short (so as not to impede sweeping as close as possible to the ground[6]). Any metal objects on the soil surface shall be removed. Following this, using a metal detector, any buried metal objects shall be removed. This clearing procedure shall be repeated after the tests to verify consistency in results, and to ensure that no metal objects have been inadvertently introduced in the lanes during the test. In some cases, during this clearing procedure, detection signals may come from the soil itself and be mistaken for metal. Therefore, some judgement will have to be exercised in determining how much effort is spent on metal object removal to minimize alarm indications from sources other than the placed targets. The goal is a test site free of metal objects other than the targets placed intentionally, so that detection of mines can be assessed.

If soil is brought in from another area for construction of the test site, the soil should be compacted so that it resembles as far as practicable the state found in an actual demining area – although this does not guarantee to give the same response of a metal detector as at the original place.

The test areas used for reliability tests shall comply with the requirements of Section 8.1.2 of CWA 14747:2003 as a minimum. They include the following general requirements:

- the minimum test area should be 1x1m and the soil 0.5m deep
- no other metal that may influence on the detector and create a signal than the investigated metal shall be present
- no ground water should be able to leak into the test area if it is not a part of the test
- a calibration area with the size of the minimum test area shall be provided for detector setup and for every soil type used in the test
- the general evaluation of other electromagnetic emission in the area shall be carried out.

However, in order to perform such testing effectively a large area is needed, which should be divided into well defined test lanes. Corner locations of the test lanes can be marked with non-metallic corner posts or pegs. These stationary markers serve to define the test lane boundaries and are used as the reference points for measuring actual target locations and locations at which target detections are declared. Lane width should be 1 m and depth at least 0,5 m. Two stripes each at least 25 cm wide shall be established on both sides of the lane. The lane and the two stripes shall comply with the requirements of Section 8.1.2 of CWA 14747:2003 mentioned above.[7] The width and depth of the lanes are so specified to ensure there is no effect from the indigenous soil of the test site (if different from the soil in the test lane) on the metal detectors during the test.

A training area shall be established where operators can practise using the metal detectors assigned to them for the test (as specified in 8.1.2 of CWA 14747:2003). In addition there shall be a similar area containing test targets of the same type and construction as used in the reliability test lanes. The training area should use soil that is similar to the soil used in the test lanes. This training area should be located away from the test lanes to minimize the possibility of operators picking up clues as to where mines are buried in the test lanes to be used for blind tests.

An accurate method is required to measure and record the placement of targets and to record the location of detections declared during the field trials. A laser-based total station survey system is ideal for this purpose. Three or four reference benchmarks outside the lanes should be measured immediately before and after each set of lane corner, target, or detection-marker measurements to confirm data integrity and provide some chance of data recovery due to operator error. If such a survey system is not available, sufficiently long non-stretch tape measures should be provided to measure these positions relative to the reference point. These tape measures are also required to facilitate the laying out of the test site. They shall not cause audio signals of the tested detectors.

---

[6] Annex B.5 of CWA 14747:2003 specifies: "Normal maximum sweep height above the surface of the ground shall be 30 mm." The specification of the maximum sweep height is redundant, because it is not possible to monitor the sweep height during a reliability test.

[7] In the CWA 14747:2003, the lanes are defined to be 1,5 to 2 m wide and the targets are placed in the 1-m stripe in the middle of the lane. The formulation provided in this document is simpler and corresponds to the meanwhile established use of the term "test lane".

Once the test targets are laid in the lanes, all visual clues that might indicate a presence of a target shall be removed. Sweeping or lightly raking the test lane surface can remove such clues. If tests extend over a period of days, lanes shall be inspected daily to remove clues that may result from settling, rain, and other causes. Even with these safeguards, telltale clues can still occur. Therefore, operators shall be told that the test is to measure detection reliability of the detector (and is not a test of individual operator performance) and therefore to ignore any visual evidence of target location.

Electromagnetic and other soil properties influencing metal detector performance shall be measured as defined in CWA 14747 Part 2[8].

9.3. *Placement of targets*

The choice of test targets depends on the objectives of the test and shall follow the description in 6.7 in this document.

Test targets shall be placed to random locations within the lane, with their entire halo inside the lane. The test objects and their depths shall be intermixed randomly down the length of the lane. Separation of targets in the test lane shall be large enough to ensure that the device under test cannot give an alarm indication due to the response from two targets at the same time. The separation shall be at least 0.5 m. This condition limits the number of targets in a lane. The formation of patterns in the target positions shall be avoided. In each lane there shall be at least one area 1 m long without any targets.

The orientation of buried targets shall be as normally laid – horizontal with activation device uppermost.

It would be pointless to bury the targets for a blind reliability test to a depth at which they cannot be detected in an open test, in which the position of the target is visible. The test targets of a certain type should be placed to depths between zero and the expected average maximum detection depth for that target type measured with the most sensitive detector in the test and established in repeated measurements.[9] In most cases, repeated measurements of maximum detection depth are not available prior to trial; in that case the average maximum detection depth needs to be estimated from the available knowledge.

The choice of target depths depends on the analysis that will be performed after the test. The simpler option is to bury the targets at one or more of the depths specified in CWA 14747:2003 (flush with the surface, 5cm, 10cm, 15cm and 20cm), depending on local clearance depth requirements, the mine types of interest and the results of in-soil maximum detection depth measurements.[10] Several pieces of the same mine type would be buried to the same depth. The results are evaluated for each depth separately. The second option is to bury the targets to depths in smaller steps, for example, 1, 2, 3, 4 ... cm, only one or two targets to each depth. The corresponding data evaluation is described in Annex K [55].

As targets are buried, their location shall be accurately noted. The location shall be expressed as down-lane and cross-lane measurements from the end and sides of the test lane (using the corner posts or pegs as datum points). The measurement of each target actual location, its type, and its depth, shall be recorded for each test lane. Access to this information shall be restricted to personnel managing the test and shall not be disclosed to any of the operators participating in the test.

9.4. *Operators*

The operators shall be chosen in accordance with the requirements of section 6.5. Most importantly, the operators shall be experienced well trained persons currently active as deminers.

An essential part of the detector operation is interpretation of alarm indications by the operator, variable human factors are introduced. In order to minimize the effect of a single operator on the results, at least three

[8] The work on this standard for soil characterisation for metal detector and ground penetrating radar (GPR) performance evaluation is in progress. It is lead by the Royal Military Academy (RMA) in Brussels, Belgium.
[9] CWA 14747:2003 prescribes: "Testing with targets at depths of 50 and 100 mm more than those derived from the in-air maximum detection height should ensure that the depth range used spans the maximum detection depth." However, it is clearly better to conform to the maximum detection depth measurements, if available and if they include repeated measurements.
[10] CWA 14747:2003 specifies: "For each target type buried, at least seven (7) identical targets shall be buried at the same depth. At least 28 targets shall be used in each test soil." This specification is not necessary.

(3) operators shall use each detector type. The recommended number of operators is at least six (6). If possible, all of the operators should operate each detector in each test lane. This may be difficult to achieve in practice, because all operators would need training on all detectors. If that would not be feasible, groups of operators should be created for each detector. It should be ensured that these groups have similar characteristics (age, demining experience). If necessary, a qualification performance test can be made so that the actual performance of an operator in that specific test site can also be estimated and controlled.

The trainer shall conduct a quality assessment of trained operators' skill level. If the trainer determines that the operator is not qualified to operate their equipment, based on evidence of inadequate skill level in the training lanes, then he/she will recommend that this person not participate in the testing. This assessment shall be conducted as a clearly described, objective and repeatable testing procedure, which shall be noted by the trial organiser. If an operator does not satisfy the quality assessment criteria, but he nevertheless participates in the test, based on a decision by the test manager, the overall test results shall be processed with and without this operator's data included.

During the testing, the operators shall not feel time pressure. The working hours and the breaks shall be similar to those in clearance operations. Some other elements of the local standard operating procedures and the local demining practice may be applied to improve the concentration of the operators and to make their work similar to their daily routine. The most important element may be the presence of a section leader performing quality assurance.

NOTE: Deminers in any given demining operation are frequently familiar with specific metal detectors. This familiarity could affect the results of a test. Therefore, the make and model of the metal detector they have been using shall be noted and mentioned in the report.

9.5. *Test procedures*

At the beginning of each test, the metal detector shall be set up according to the manufacturers instructions as given in the user manual. The metal detector shall be adjusted for the maximum sensitivity attainable on the soil of the test lane. It shall be set up to the soil if necessary. The detection tests shall be blind; that is to say the operator shall not know the location of targets laid in the lane prior to searching for them with the metal detector. The operator shall be supervised at all times by personnel managing the testing. Each metal detector shall be used in its normal operating mode, sweeping manually in a transverse direction at a speed that ensures optimum detection capability (see Sections 6.4.2 and 6.4.3 of CWA 14747:2003 for more detail). The operator shall move the metal detector forward along the test lane between sweeps. The distance in the forward direction between successive sweeps shall be small enough to ensure that the sensitivity area of the metal detector covers all of the ground in accordance with the target depth. The appropriate sweep overlap may be determined by measuring the sensitivity profile according to Chapter 9 [20].

The operators shall attempt to detect all targets in the test lane. The operators shall use markers to mark the locations of alarms.

The direction in which any single operator moves down a lane shall be alternated (bottom to top, top to bottom) with each pass. This will make it harder for the operator to memorize locations where he/she has previously detected a target. Blind tests shall not be observed by other operators for the same reason. The only observers shall be the test personnel.

Each complete pass of an operator with a metal detector through a lane is called a run. After each run, test personnel shall measure the location of each marker, record the measurements, and retrieve the markers. After the test, the recorded marker locations shall be compared to the actual test target locations to determine the test results. This process shall be repeated for each metal detector in each soil type. If possible, the test shall be performed with at least two (2) samples of the same detector model. When the test is over and as the targets are removed, the identity and position of the removed targets shall be verified and compared to the original location measurements. Any additional objects found in the lanes shall be accurately noted. Should there be any discrepancy, the test results shall be corrected as appropriate.

If locations in the test lane give persistent false indications, such locations shall be investigated to ensure that there is no object that has been inadvertently left there. If such an object is found, all reported indications of prior tests that correspond to this object (within detection halo) shall be omitted from analysis of the results.

9.6. *Design of experiment*

The trial organiser needs to decide how to combine the detectors, operators, and lanes in the trial. Unfortunately it is not possible to give simple prescriptions applicable to every experimental problem. Each trial is different: both the available resources and the experimental goal vary from trial to trial. The experimenter should be familiar with the principles of experimental design and approach every trial as a new experimental problem. However, some general recommendations can be made.

In an ideal trial all metal detectors would be tested by all operators, in order to reduce the influence of the individual differences between the operators. However, the time required for training would than be very long. It is therefore acceptable that each metal detector model is used by other operators.

The design of experiment should be a fractional factorial design based on a Latin square or a Graeco-Latin square. Variable “start” shall be introduced to indicate the execution order of the runs. Run is a single complete pass of an operator with a detector through a lane. A start is a group of runs executed at the same time. Examples for such designs are provided in Annex G. A single Latin square or a Graeco-Latin square should be repeated with permuted (meaning: switched, mixed) operators or detectors or both, so that eventually all detectors will have been used by all operators trained for that detector in all soil types, if the scope of the trial allows that. The number of soils and target types should be as small as possible, so that the number of targets in each target-soil combination can be as large as possible. The number of operators should be as large as possible, since operator can be understood as a nuisance factor, in the sense that the experimenter is not interested in the results of any particular operator. If some earlier measurements imply that there is very little variation between specimens (i.e. copies) of the same metal detector model, than only one specimen may be used in a reliability test, or more specimens treated as identical.

The number of necessary repetitions depends on the least difference between PODs of metal detectors which needs to be discovered in the test. This least difference that needs to be discovered is here called a critical difference d. If the difference between the PODs of two detectors is larger than the critical difference d, the two detectors will have significantly different PODs. For example, $d = 0.2$ means that all differences between two detectors larger than 0.2 need to be seen as statistically significant. This also means that the differences smaller than 0.2 would be not statistically significant and it could not be said if they are caused by actual differences or they are a result of pure chance. The critical difference d is a measure of the required precision of the whole test: the smaller the d is the more repetitions are needed. The number of repetitions is expressed with the number of opportunities to detect a target. It is recommended that the number of opportunities to detect a target is larger than $(2/d)^2$. The rationale for this recommendation with a detailed example is given in Annex J. For example, if two detectors need to be compared in one soil, the number of opportunities to detect the target in that soil will be the number of runs with any of these detectors in that soil multiplied with the number of targets in the same soil.

9.7. *Test results and data evaluation*

The output variables of a reliability test are probability of detection (POD) and false alarm rate (FAR). The estimated probability of detection for a particular combination of an operator, detector and a lane is the ratio of the number of detected targets and the total number of opportunities to detect a target. This number is here denoted n.[11] The estimated false alarm rate is defined as the number of false alarms counted on an area divided by the size of that area, or the average number of false alarms per square metre. The area is calculated as the area of the test lane minus the area of all detection halos. If we assume a binomial distribution for the number of true positive indications, we can find the 95% confidence limits for the probability of detection. Similarly, if we assume Poisson distribution for the false alarms, we can construct the confidence limits for the false alarm rate [10][11][12]. They are described in Annex K [52].

As described in Section 9.5, the positions of detections shall be indicated with markers. The positions of the markers shall be compared with the actual positions of the test targets. The criterion for detection is the size of the so called detection halo, which is defined in Annex F. If a marker lies within the halo radius of a test target, that target is detected and the indication is called a true positive indication. Markers outside the halo radii are counted as false alarms, also called false positive indications. True positive (TP) and false positive (FP) indications shall be counted. The numbers of TP and FP indications for the metal detector are here marked x_{MD} and y_{MD} respectively.

[11] For example, if there are 15 test targets in a lane and 4 runs are analysed, the total number of opportunities to detect a test target is 4 times 15, that is n = 60.

	number of indications
number of TP	x_{MD}
number of FP	y_{MD}

For establishing the estimated POD, the above mentioned numbers x_{MD} and x_R shall be divided by the total number of opportunities to detect a test target, here marked *n*. For establishing the estimated FAR, the numbers of false negative indications shall be divided by the searched area.

$$POD_{MD} = \frac{x_{MD}}{n} \text{ and } FAR_{MD} = \frac{y_{MD}}{NA} \text{ ,}$$

where *N* is the number of repeated scans over an area of size *A*. The POD and the FAR shall be combined in an ROC (receiver operating characteristic) diagram, so that the FAR is on the abscissa and POD on the ordinate axis, with the corresponding 95% confidence limits. The calculation of 95% confidence limits is described in Annex I. For a particular choice of lanes, targets and operators, the result of each detector shall be represented by a point on an ROC diagram. The best achievable result is a point in the upper left corner of the diagram, where POD = 1 and FAR = 0.

The other kind of diagrams used in reporting shall be POD curves. POD curves are curves of POD in dependence on target depth. For creating POD curves, the best choice is to bury the test targets to depths in smaller steps, for example, 1, 2, 3, 4 ... cm, only one or two targets to each depth. A possible way of creating POD curves is described in Annex I.

A simpler analysis of POD in dependence on depth is possible if targets are buried to a smaller number of depths, for example, flush with the surface, 5, 10 and 15 cm deep. Detections for each depth are simply counted separately and expressed as POD with the appropriate confidence limits (as explained in Annex I), in a diagram of POD versus depth [13][6]. The confidence limits are calculated with the assumption that the POD is binomially distributed. Such a method has the advantage that it needs fewer calculations and that it does not depend on any assumptions about the relationship between the POD and the depth; the only assumption is that the detections on a certain depth are binomially distributed. Its disadvantage is that it produces larger than necessary confidence intervals and that it cannot give information about the depths not present in the test.

10 Detector sensitivity profile

The detailed description for establishment of the detector sensitivity profile (footprint, sensitivity area or cone are other used names) is in two versions field and lab conditions given in the CWA 14747 : 2003 chapter 6.7. In both cases special equipment is necessary for establishing it.

The objective of this test is to determine the sensitivity profile of a detector. The sensitivity profile is the variation of sensitivity with location in a plane parallel to the sensor head. The sensitivity profile depends on all factors defining the sensitivity of a detector to the objects that should be detected. Their metal content (mass), type of metal, and their shape will influence their detectability. The detectability of the metal part depends on the surface of the metal object. For instance, a ball-like shape is more difficult to detect in comparison to any other shape with the same mass that may be a part of a mine, more details in [6.1]. The soil may also influence on the size of the sensitivity area if soil compensation has to be used. The form of the sensitivity area is cone-like but rounded at the lower edge.

Whether or not a target is detected can depend on what part of the search head is swept over it. It is very important to characterize this aspect of sensitivity variation in order to assess the field performance of a detector. This data can be used to define the overlap needed between consecutive scans to ensure complete coverage at he required sensitivity. If a detector is capable of detecting the targets being sought at the clearance depth only when the target is directly under the central part of the sensor head, then the detector sweeps need to be highly overlapping.

10.1.1 Testing principle

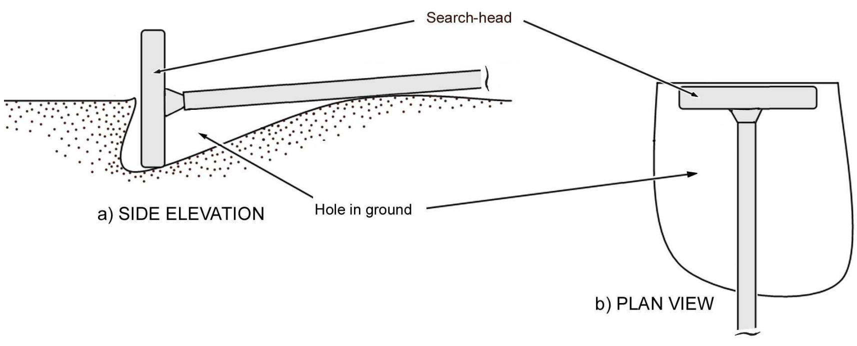
The test described below is to examine the metal detector’s capability to detect objects of interest in clearance operations and in field environment. It is difficult or impossible to control all of the variables that may affect a detector. A detector’s ability to detect objects in air does not directly indicate its ability to detect objects buried in the ground. The performance in the field can be defined in reasonable accuracy by the described procedures.

This test is principally intended for use in field like environments.

10.1.2 Test procedure

Set up the detector to maximum sensitivity as defined in the user manual of the manufacturer. Check if the detector signals if kept close to the ground. If not, you are on ground that is neutral to the metal detectors magnetic field. If YES, you shall use the ground compensation function so that the detector will not create a signal from the ground without metal is present.

A place free of vegetation with a sloping hole must be prepared. The hole should be deep enough for the search-head to be presented horizontally to the vertical wall of the hole as shown in the drawing below. Where available a plastic, wooden board can be placed on the soil surface for accurate marking of the sensitivity area.



Now the target/mine should be moved (on its side), parallel to the centre of the bottom of the search-head. Start close to the search-head and move the target away in 20mm increments. Mark the place where the detector-signal becomes uncertain or ends. This is the maximum detection depth. Now, mark the “centre-line” on the ground/board from the centre of the detector head to the maximum depth. Move the target towards this from one side, marking the place where the detector signal begins. Repeat this at 20mm intervals, moving the target from each side towards the centre-line. Make a mark each time the detector starts to signal. The marks you make show the limit of the sensitivity area. Turn the detector search-head 90° and repeat. You mus t turn the detector head to allow for different search-head geometry. If the form of the search head is a circle you do not need a repetition.

The outline you have marked on the ground/board will be close to a cone with a round edge at maximum detection depth. It will be specific to the target – so the mine with the lowest metal content expected in the area should be used. To show how the cone varies according to the metallic signature of the target, repeat this again using another target/mine. With large targets such as UXO, the cone will extend beyond the sides of the search-head.

Several deminers shall collaborate to do this test; it should be 3 - 4 times repeated with each deminer taking a turn to move the target and to define the place where the signal starts/stops. Slight differences will occur but the general shape will not change. The average of the results will be accepted as the sensitivity profile of this target in the specific soil.

NOTE The double "D" detector can only be used in a position that both half's of the D cover the same amount of the ground otherwise the detector will permanent signal.

10.2 *Target choice*

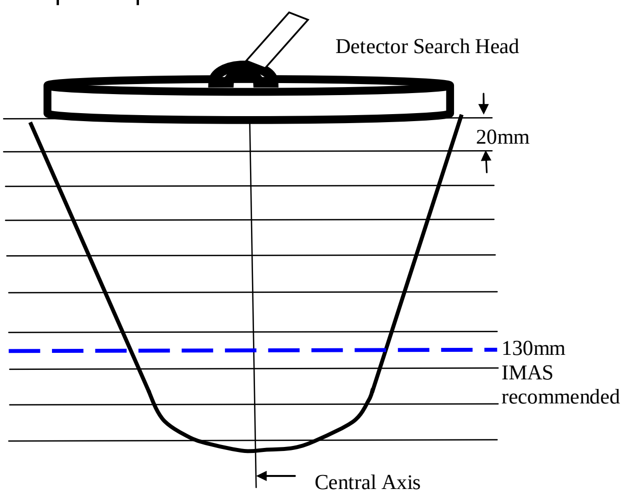
Depending on the local conditions and mine threat the choice of the target/mine should be done under following general rules:
The target

- has been involved in mine accidents during clearance operations
- is difficult to detect up to the established clearance depth in the country (minimum metal mine)
- has changed during the time of being placed in the soil (aging).

These rules are equal in their importance and combinations are possible too. The aging process is connected with the metal content of the mines where rust of ferrous parts can significantly change (reduce) the electromagnetic response of such target.

10.3 *Data evaluation and reporting*

The Sensitivity area shall be measured in length and width, the length along the central axis and the width of the profile parallel to the search head bottom in 20mm increments. This is important to know for the establishment of the search advance of the metal detector in clearance operations. The picture clearly shows that at maximum detection depth just a small radius will be searched and easily a mine can be missed. At the IMAS recommended standard clearance depth about 2/3 of the search head will only be covered by the sensitivity profile in the example of the graph.

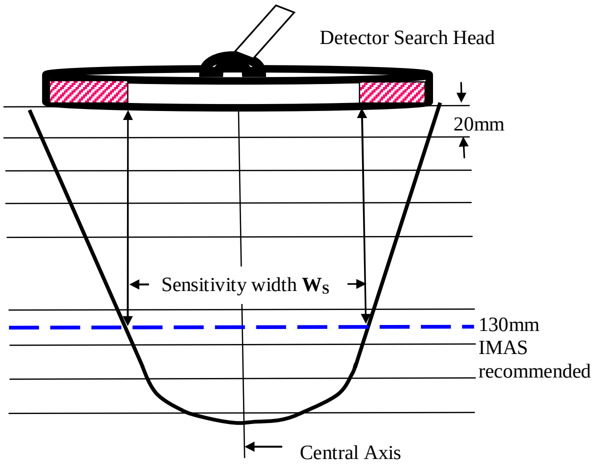


An advance of the full search head means that an area of 1/3 of the search length/diameter will not be searched at the established clearance depth.

10.4 *Defining the search advance*

The search advance of the metal detector shall only be defined by the sensitivity area from the target with the smallest width in the established national or international clearance depth. With this it can be avoided that operators due to their search procedures may miss a mine.
The metal detector width (diameter, length, W_{MD}) minus the measured width at clearance standard depth (W_S) is the basic result for establishing the advance. The result divided by 2 shows the part of the outer search head diameter that will not fully cover with the sensitivity area the established clearance depth. In the picture below the red marked area covers only partially the area to the clearance depth. It demonstrates how the sensitivity profile of the detector search head covers only a certain area at the required depth, not the full search head width. Inside the sensitivity profile targets will be detected outside not. Below the white part of the

search head the sensitivity profile will provide permanent coverage and detection of the full clearance depth.



The advance of the detector search head shall be not more than ¾ of W_S length for having an overlap of the fully covered area.

Example:
W_{MD} = 300mm
W_S = 200mm

MD Advance = $\dfrac{200 \times 3}{4} = 150$

With the advance of 150mm the deminer covers with each movement of the detector permanent the necessary clearance depth. Additionally the 50mm overlap is a tolerance that provides the sureness that the whole clearance area is searched with the optimum sensitivity area.

10.5 *Use of the symmetric and decoupling function*[12]

For users with the possibility to create additional equipment the below given ideas can establish the distances and a profile by using the approach below.

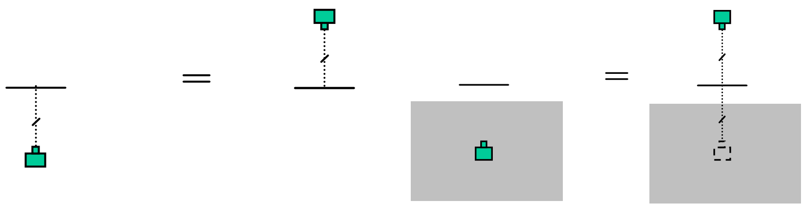


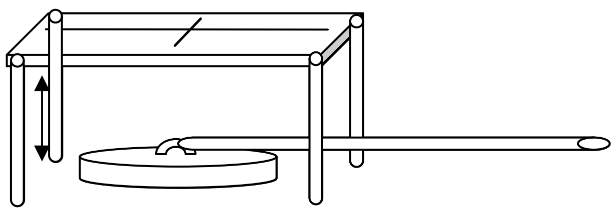
Figure 1 and 2

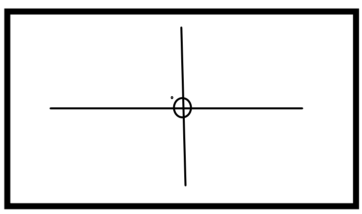
For measuring the sensitivity to an object or sensitivity area of a metal detector the assumption is that the magnetic field of the metal detector is symmetric to both sides of the detector search head, so measurements should have the same result on both sides. Therefore measurements can be done in air if the detector is set up to the soil and does not lose the contact during the process as shown in Figure 2

The second assumption is that the response of the soil and the target decouple. If soil electrical conductivity and magnetic susceptibility are low enough and the soil and target responses decouple. This means that the total response is the sum of:

- the response of the target in air and
- the response of the soil in absence of target.

The practical value is considerable and can make tests much faster as before. The empirical confirmation should be done for confirming the assumptions. Further the tests can be carried out in the way that the target will be moved and the search head stay fixed at its place.
For this purpose polycarbonate (other non-metallic material) can be used as the working platform and for establishing the profile in different height.





[12] The content is taken from the CWA 14747 Part 2 and adapted to the purpose

The detector stays at its place. The centre of the search head as well as the X and y-axis should be marked on the used plate. The height of the plate above the search head should be reduced by 20mm increments to establish the sensitivity area by moving the target along the both axis from outside over the centre to the other side. The position should be the opposite (mirrored) to in soil. Places where the signal starts or stops are to be marked for recording the result including the height above the search head.. If wanted and if the search head has not a circular form other directions may be added for a more accurate sensitivity area under (in this case above) the search head.

11 Evaluation of selective metal detector features

11.1 *General*

Since the publication of the CWA 14747 for metal detectors in 2003, new technical features are proposed to be used in the area of humanitarian demining. Those features are of interest for every clearance operation because they concern selective capabilities of metal detectors that may reduce the false alarm rate during the clearance process.

11.1.1 Description

The selective capabilities are based on the different response of metals received by the detector's antenna. These differences are visible on a screen and give an orientation which metal is dominating, a non-magnetic metal or a ferromagnetic metal. In most cases the mines have in their mechanical part for the initiation of explosive different metal in different shape. Mines with only one metal may be used too and in such cases already a selective approach may significantly reduce the false alarm rate in certain regions/minefields.
The other feature is a learning function where the detector determines discrimination based on conductivity and ferrous content of a target. The response will be a mixture signal of metals characterising just one mine. This also can be used to discriminate between a mine and other metal. When such a detector learns a target it will ignore other targets that have different conductivity and ferrous content.

Note other selective capabilities may be developed not similar to the above described. The structure and the approach of this test may be used in such cases too.

11.2 *Discrimination between ferromagnetic - other metals, and learning function*

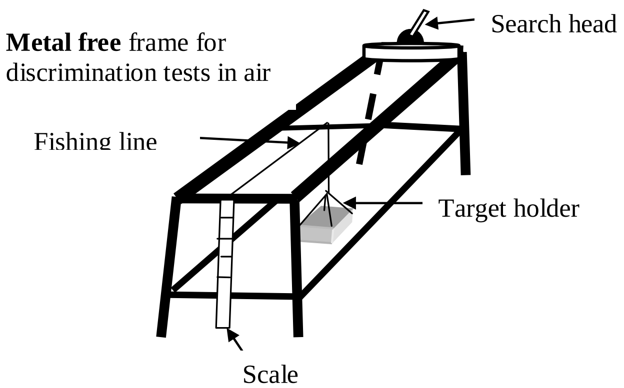
11.2.1 Testing principle

The test described below is to examine the metal detector's capability to discriminate objects of interest in clearance operations and in field environment. The main interest is to establish the reliability of the discrimination capability over the full range of detection. The performance for tests in air shall be confirmed in soil and in field conditions. The discrimination (selection) can be done by excluding something or by selecting something.

11.2.2 Test procedure

The place for carrying out the test shall exclude any electromagnetic interference, like telecommunications, power transmission and other objects that may create signals for the metal detector. This test should be carried out in air as well as in different soil for establishing the reliability of discrimination of the metal detector.

The detector shall be setup as determined in the user manual. A check shall first be conducted using the manufacturer's test piece (if supplied) to confirm detection at the specified distance to ensure proper operation. Then the detector shall be set-up in the suitable mode – the selective mode in which the detection capability is to be evaluated. The sensitivity and other settings (where applicable) shall be recorded. Then the detector shall be passed over a target at the constant sweep. The detector height above the target shall be varied until the maximum height that produces an alarm indication is determined. Some detectors have a distinctive alarm/no alarm transition within 10 mm while other detectors have a 20 to 30 mm range where it may be difficult to determine whether detection has occurred or not. A practical increment with which to vary detector height above target is



therefore 30 mm to establish the limits of detection. The height increment from the last clear target signal to the max detection depth shall be 10mm.

The target to be used shall be placed on the target holder or in soil in the correct orientation, adjusted to the position of the detector sensor head to ensure that the required sweep path is followed and that the sensor height above the target is maintained.

The location of maximum sensitivity to a target varies with the design of the sensor head. Chapter [10] gives details on how to establish the full sensitivity profile of a footprint. A simple manual trial at the required height is however sufficient to find the maximum sensitivity location for the purpose of the discrimination test.

The indication on the screen demonstrates the received signal from the coil. The form and direction of the signal should be carefully registered and any change with the growing distance described in the protocol or by pictures.

Depending on the amount of available targets or their simulants the maximum detection depth approach for in soil tests shall be used for the measurements in soil. For estimating the maximum detection depth measurements in air can be used.

The execution of the test in different soils will give the possibility to evaluate the influence of the ground compensation on the sensitivity but also the influence on the response signal which may be influenced by the mineralisation of the soil too. A change of the target orientation to the search head and its influence on the discrimination capability shall be evaluated too.

The possibility that the target may be hidden by stronger signals from other metal pieces shall be a part of the evaluation. Similar to the test for evaluating the discrimination of adjacent targets should be included into this test.

11.2.3 Target choice

The target choice is for the discrimination test the most sensitive part and should express the local requirements. Mines and metal found in clearance operations are to be selected for receiving an overview about the metal detector's capabilities but also the limits of such a possibility. In most cases the electromagnetic response is defined by a signal consisting of multiple metal, amount and shape i.e. a combination of metal reacting to the magnet field of the detector. Where possible, for the test typical metal parts should be selected by similar shape and surface. Ideal objects are balls that response from all directions in a similar way. For field use geometrical forms with equal measurements will fulfil the requirements too.

The mass and surface of metal may also create quite different signals being from the same metal as in mines used. The surface of the used targets should achieve the size that is covered by the mines found in local conditions including metal anti tank mines and also minimum metal content mines.

The approach for selecting or discrimination shall go in two directions:

- Which way of discrimination provides the biggest effect?
    - The selection of the mines representing the threat of the area, or
    - The selection of clutter that may be typical and exclude the mines.

<u>Example:</u> In most cases the mines laid in minefields are established by survey and known before the clearance operations starts. Minefields in certain regions include only AP mines with aluminium percussion caps as the only metal part. In this case it makes sense to look for aluminium in such an area and exclude other metals.

The test shall include the basic metals used in mines, iron, copper and alloys of it, aluminium, lead but also the original mine (rendered safe) or mine parts used in the configuration of the mine.

11.2.4 Data collection

Detectors with selective capabilities have a screen where the deminer (operator) can visually receive the information about the target response. There are scales showing the response and giving information about the responding metal or metal combination. A huge metal amount may create a signal in another sector of the screen as it may be for a small amount of the same type of metal. This information has to be taken out of the user manual from the manufacturer.

All changes of the acoustical and visual signals demonstrating the selected target/ section shall be recorded for the tested depth in the test protocol ANNEX E.2 and comments shall be added if there are differences to the signal before or other changes not described by the manufacturer. The detector should confirm the detection five times in a row for every depth by moving over the target in both directions. It is important to repeat this test with different persons and especially at depth or with targets where the signal is not clear. The number of repetitions shall achieve an amount of attempts that may be statistically evaluated. Where such cases appear at least 10 attempts should be done.

11.2.5 Data evaluation and reporting

The evaluation of the data shall be done as described in [8.4] for maximum detection depth.

A single pass of an operator over the complete set depths for one target is called “measurement”. The measurements with all operators on the same target are called “test”. There are two possible outcomes of a discrimination detection test:

1. All targets up to a certain depth are identified by both signals and targets buried deeper than that depth are not identified.
2. Some deeper buried targets are identified, while some shallower buried targets are not identified.

For easier treatment of the data it is recommended to evaluate the discrimination ability of the acoustical and the visual signal as single measurements and results in two columns.

If the measurement outcome is of the first kind, the data analysis is relatively simple. The largest depth at which the target has been identified in a single measurement shall be called maximum identification depth of that measurement. If the achieved results are normally distributed, at least the following quantities shall be reported:

3. the average of all measurement results for that detector, here marked d_m
4. the standard deviation of all measurement results for that detector, here marked s
5. the number of measurements, here marked n

The average maximum identification depth d_m increased by 5mm is equal to the target depth at which the probability of detection is ½. The increase by 5mm is necessary as a correction of the systematic error caused by the 1-cm size of the step between the depths in the last phase for establishing the maximum identification depth. The standard deviation s of the maximum identification depth is a measure of the stability of the detector’s performance.

If the measurement outcome is of the second kind, with some deeper buried targets detected, while some shallower targets not detected, the data evaluation is more complex. In that case it is recommended to create POD curves using the generalised linear model described in Annex K [55].

It is of highest importance to define the cases and depths where the reliability of identification achieves the level below 10%, in dependence on the target depth, its size, the influence of adjacent targets, soil, and other factors influencing on detector performance.

For a more complex reliability test the proposed and from the ITEP ExCom endorsed reliability test for dual-sensors may be used.

A Annex: CWA 14747:2003 Tests, their purpose and intended result (trial type; consumer report trial (CRT) or acceptance trial (AT),

The main content and the information that may be gathered by executing the tests under the described conditions will be summarised below. 30 tests are included in that document with a description of their execution.

The chapters of the CWA for metal detectors include:
- Chapters 1 to 5 – the general conditions for testing and reporting,
- Chapter 6 – Measuring the maximum detection height, tests in air to different targets, determining the detector sensitivity are to a target[13],
- Chapter 7 – Immunity to environmental conditions and their influence,
- Chapter 8 – Detection capability for targets buried in soil (including the reliability (blind) trial),
- Chapter 9 – Operational performance characteristics (pinpointing, adjacent targets etc.), and
- Chapter 10 – Evaluation of ergonomic and operational aspects (shock & bump, drop tests, interchangeability of parts).

The annexes give details for soil characterisation, test targets and reporting forms.

The guide uses the same approach as the basic CWA and is focused on the execution of the selected key tests under field conditions.

A.1 *Tests included into the CWA 147472003*

A.1.1 Measuring the maximum detection height (CWA Chap.6 / 9 tests)

The tests are foreseen as general tests for establishing detector parameters necessary for further testing under lab conditions i.e. optimum sweep speed, or the capability of the detector to be set up in the same way achieving the same sensitivity and to keep this sensitivity during a certain time period. For international comparison the sensitivity tests are described to standard targets/balls consisting of different materials or other specific targets.
The test 6.7.1 and 6.7.2 of the CWA are of interest because the sensitivity profile (footprint) is one of the key tests in more detail explained above in Chapter 10 for field conditions. It has operational importance due to its result. It allows determining the search head advance to a target buried in soil.

A.1.2 Immunity to environment and operational conditions (CWA Chap.7 / 6 tests)

The tests provide for the evaluation of the detectors' reaction to environmental conditions like moisture on the search head, to temperature changes, and the effect of electromagnetic waves on detector performance. Another test includes the detectors' sensitivity during the battery life-time.

Lab tests have shown that these factors have little influence on modern detector performance and the information about those factors shall be given in the in detector manual by the manufacturers. If somebody wants to check those data, the tests are in detail laid out in CWA 2003 Chapter 7.

A.1.3 Detection capability for Targets buried in soil (CWA Chap. 8 / 5 tests)

The tests include the main factor influencing on detector performance the soil. Its characterisation will be given in Annex B. In this document, only factors that may be evaluated in the field will be included. For information that is more detailed the CWA 2003 Part 2 shall be consulted.

The tests provide information about the detector capability for detecting targets buried in soil. Different approaches are possible to establish maximum detection depth in soil and to evaluate the results. The tests are still the only way to establish the detector performance to a target to a specific soil.

[13] The sensitivity area is in publications also named footprint, sensitivity cone.

Due to their importance two of the tests included in this chapter belong to the key tests that shall be carried out as the minimum to define the detector performance under field conditions – the reliability (blind) test and the establishment of the maximum detection depth to a target. The reliability test is the most complex test and allows evaluating all factors influencing on detector performance. The maximum detection depth is a precondition for the reliability test and demonstrates the capability of the detector to find a target. Both tests and their planning and execution will be described in detail later.

A.1.4 Operational performance characteristics (CWA Chap. 9 / 7 tests)

The tests include the evaluation of the detectors' capability in combination with the deminer. The accurate pinpointing of a target, the determination of the shape as well as the resolution of adjacent targets is a part of operational requirements. Other evaluation includes the influence of salt water, power lines and other specific media. The work close to large linear targets is another part (metal fences, railways, buildings etc.) important in field conditions and may be tested too.
In a very specific case the mutual interference of detectors shall be known and is important – a mine accident or other affect taking the deminer out of order in a mine field and his detector is still working.

All the operational tests shall be used for evaluating the capabilities of the deminer in the detector use and internal organisational standards may be established for such purpose.

A.1.5 Evaluation of ergonomics and operational aspects (CWA Chap. 10 / 3 tests)

The ergonomics belong to applied science of equipment design intended to maximize productivity by reducing operator fatigue and discomfort. It should remove barriers to quality, productivity, and safe human performance by fitting products, tasks and environments to people. The operator should not experience discomfort but feel support from the used tool. In the case of the metal detector, it means the design, the weight, ease of use including the interface to the operator, fault signals, and transport cases are to be assessed. The manufacturers of metal detectors undertook to the end of nineties/beginning of the new century immense efforts in this direction. This resulted in the one piece detector with different ways to communicate with the operator. The trend is that the user will be able to adjust the detector to specific environmental conditions in the area of employment. Those assessments have never been an object of a trial but in a form of questionnaire information about those facts were everywhere included. The factors which contribute first hand to the POD are the signal cleanness and weight.[5]

Table 1 — Categories of testing

Clause	Test	Testing category							
		CRT	AT	open	blind	lab	field	air	soil
6	*Detection capability testing in-air*								
6.3.3 & 6.4.1	**General test –** **Measuring the maximum detection height**	●	●	●		●	●	●	
6.4.2	**Sweep speed – mechanized movement**	●	○	●		●		●	
6.4.3	**Sweep speed – manual movement**	●	○	●			●	●	
6.4.4	**Repeatability of sensitivity on set-up**	●	○	●		●	○	●	
6.4.5	**Sensitivity drift**	●	○	●		●	○	●	
6.5.2	**Minimum target detection curves for steel balls**	●	●	●		●	●	●	
6.5.3	Minimum target detection curves for other metals	●	○	●		●	○	●	
6.6	**Detection capability for specific targets**	●	●	●		●	●	●	
6.7.1	Sensitivity profile (footprint) measurement - Method 1	●	○	●		●		●	
6.7.2	Sensitivity profile (footprint) measurement – Method 2	●	○	●		●	○	●	
7	*Immunity to environment and operational conditions*								
7.2	Sensor head orientation and shaft extension	●	○	●		●	○	●	
7.3	Moisture on sensor head	●	○	●		●	○	●	
7.4	Temperature extremes	●	○	●		●	○	●	
7.5	Temperature shock	●	○	●		●	○	●	
7.6	Sensitivity during battery life	●	●	●		●	○	●	
7.7	Effect of EM/RF interference	●	○	●		●		●	
8	*Detection capability for targets buried in soil*								
8.2	**Minimum detectable target as a function of depth**	●	●	●		●	○		●
8.3	**Detection capability for specific targets in soil**	●	●	●		●	○		●

Clause	Test	Testing category							
		CRT	AT	open	blind	lab	field	air	soil
8.4	**Fixed-depth detection test**	●	●	●		○	●		●
8.5	**Detection reliability tests**	●	●		●		●		●
8.6	Additional detection reliability testing	●	●		●		●		●
9	*Operational performance characteristics*								
9.2	**Target location accuracy**	○	●		●	○	●	●	○
9.3	Shape determination of targets	○	●		●	○	●	○	●
9.4	**Resolution of adjacent targets**	○	●		●	○	●		●
9.5	The influence of specific media on detection		●	●		○	●		●
9.6	Detection near large linear metal objects	○	●		●	○	●		●
9.7	Effect of specific electromagnetic interference sources		●	●			●	●	
9.8	Mutual interference between detectors	○	●	●		○	●	●	
10	*Evaluation of ergonomic and operational aspects*								
10.1.1	Shock and bump tests	●	○	●		●		●	
10.1.2	Drop tests	●	●	●			●	●	
10.3	Interchangeability of parts	○	●	●		○	●	●	

A.2 General description and grouping for field use

Name		Objective and content	Preferred to apply – Locality: Lab	Field	In air	In soil	Type of test: Consumer	Acceptance	Blind	Open	Info from manufacturers (*)	Remarks
IN AIR TESTS		Stability/Drift of sensitivity										
		• After set-up	☒	☐	☒	☐	☒	☐	☐	☒	☐	
		• During operation	☒	☐	☒	☐	☒	☐	☐	☒	☒	
		Optimal sweep speed	☒	☐	☒	☐	☒	☐	☐	☒	☐	
		Maximum detection height										
		• Standard targets	☒	☒	☒	☐	☒	☐	☐	☒	☒	
		• Different metals	☒	☒	☒	☐	☒	☐	☐	☒	☒	
		• Specific targets	☒	☒	☒	☐	☒	☐	☐	☒	☐	
		Sensitivity profile (footprint)	☒	☐	☒	☐	☒	☐	☐	☒	☐	Defined standard test target
		Miscellaneous may be included here	☒	☐	☒	☐	☒	☐	☐	☒	☐	As consumer test
USER-SPECIFIED TESTS		Effect of sensor head orientation	☒	☒	☒	☒	☒	☐	☐	☒	☒	
		Moisture on sensor head	☒	☐	☒	☐	☐	☒	☐	☒	☐	
		Temperature extremes/shock	☒	☐	☒	☐	☒	☐	☐	☒	☒	
		Effect on EM/RF interference	☒	☒	☒	☒	☒	☐	☐	☒	☒	
		Sensitivity during battery life	☒	☐	☒	☐	☐	☒	☐	☒	☒	Time before alarm signal
		Shock and bump test	☒	☐	☒	☐	☐	☒	☐	☒	☐	
		Drop test	☒	☒	☒	☒	☐	☒	☐	☒	☒	
		Mutual interference of detectors	☒	☒	☒	☒	☒	☐	☐	☒	☒	With other detectors
		Interchangeability of parts	☒	☐	☒	☐	☐	☒	☐	☒	☒	Where possible
IN SOIL AND FIELD TESTS	In soil	Detection depth in different soils	☐	☒	☐	☒	☐	☐	☐	☒	☐	Measurements of soil
		• Standard targets	☐	☒	☐	☒	☐	☐	☐	☒	☐	As in air
		• Different metals	☐	☒	☐	☒	☐	☐	☐	☒	☐	As in air
		• Specific targets	☐	☒	☐	☒	☐	☐	☒	☐	☐	Chosen by end-user
		Reliability tests	☐	☒	☐	☒	☐	☐	☒	☐	☐	
	Miscellaneous	Locating accuracy (pinpointing)	☒	☒	☐	☐	☐	☒	☒	☐	☒	Not above mines
		Shape determination of targets	☒	☒	☐	☐	☐	☒	☒	☐	☐	Point, linear, polygon
		Resolution of adjacent targets	☒	☒	☐	☐	☐	☒	☒	☒	☐	AP and AT mines
		Influence of specific media	☒	☒	☐	☐	☒	☐	☐	☒	☐	
		Detection near large linear metal	☒	☒	☐	☐	☒	☐	☒	☒	☐	Railway, fence
		Effect of EM/RF interference	☒	☒	☐	☐	☒	☐	☐	☒	☒	Power lines, radio
		Mutual interference of detectors	☒	☒	☐	☐	☒	☐	☐	☒	☒	Recovery test

B Annex: Soil characterisation CWA 14747 Part 2

The influencing electromagnetic soil properties include:

- the magnetic susceptibility and the electrical conductivity their frequency dependence and spatial distribution,
- both properties determine the ground response to a metal detector in connection with other soil properties.

In extreme situations they may cause a metal detector to be unusable or force the use of ground compensation that may reduce the detector's sensitivity. Those situations had been described before and the CWA Part 2[1] gives indicative values influencing on detector performance and its level characterisation.

Table 1: Values for continuous wave detectors

Soil effect class (Defined in CWA 14747 Part 2 Annex A.1)	Indicative values of magnetic susceptibility 10^{-5} SI
Neutral	Below 50
Moderate	50 to 500
Severe	500 to 2 000
Very severe	Above 2 000
NOTE These values are indicative only because the effect of a soil on a metal detector depends on the metal detector, and the measured values of susceptibility depend on the measuring instrument.	

The type (pulse or continuous wave) of detector and manufacturer influence the degree of weight on the performance. For pulse detectors especially the frequency dependence may have severe impact on the performance.

Table 2: Values for pulse detectors

Soil effect class (Defined in CWA 14747 Part 2 Annex A.1)	Indicative values of frequency variation of magnetic susceptibility (465 Hz and 4 650 Hz) 10^{-5} SI
Neutral	Below 5
Moderate	5 to 15
Severe	15 to 25
Very severe	Above 25
NOTE These values are indicative only because the effect of a soil on a metal detector depends on the metal detector, and the measured values of susceptibility depend on the measuring instrument.	

The electromagnetic soil properties can be measured by different tools and the values may change. The use of the (a) metal detector as measuring instrument is another method to characterise the electromagnetic properties of the soil. The ground reference height is one way. It is described in the metal detector Handbook. This empirical measure indicates the level of soil influence by the height where a calibrated detector starts to signal due to the soil response. For this purpose static detectors with a continuous change of sensitivity are preferred.

There is another soil property that can influence on metal detectors the saline water content in soil. The soil water content in general may be neutral but in connection with salt or other chemical compounds ionic parts for the transport of electrical charges may be available. This situation (tides on a beach, flooded areas) can create problems for the detectors too.

C Annex: Test Targets

C.1 Parametric test targets

To measure the detection capability of a metal detector as it varies with target-to-detector height requires a set of simple geometric targets (see 5.6) of different size. This makes it possible to plot the maximum detection eight against target size. The targets used shall be spherical metal balls. The acceptable tolerance on the steel ball sphericity shall be ±0,1mm.

Balls with diameters in the range 3 mm to 30 mm shall be used.

Balls of the following materials shall be used to evaluate the relative sensitivity of a detector to different metals:
1) High carbon, low alloy steel, also known as chrome steel (designations; UNS G52986, AISI 52100, UNI 100 Cr 6, DIN 1.3505) shall be used for the basic detection capability curve. Balls of this material are easily available all around the world and so it is straightforward to measure detection curves in field conditions. This is a ferromagnetic steel. Tests have shown [13] that balls of different ferromagnetic steels give practically indistinguishable sensitivity curves, so the curves are not sensitive to small variations in material properties. Many mine parts are made from ferromagnetic steel (e.g. striker pins).

2) Austenitic stainless steel (designations; UNS S31600, AISI 316). This metal has a low electrical conductivity and is essentially non-magnetic (effective magnetic permeability is approximately 1). Small parts of such metals are often difficult to detect.

3) Aluminium. Many mine parts, such as detonator tubes are made from aluminium. This metal is therefore important to be able to detect. Additional materials of interest may be added for purposes of comparison. For example, lead and copper are found in some landmines. In every case the material specification shall be recorded.

C.2 Standard targets that simulate metal components of mines ("ITOPs")

Test targets shall be used that have the metal content typical of a class of mine. Such targets shall be selected from the set of test targets listed below. These targets are known as the "ITOP inserts" since they were conceived for an International Test Operations Procedure project as the metal content of larger simulant mines. The targets are defined in [11]. The targets are metal parts set in silicone within a plastic cylinder. The metal parts sit on the base of the cylinder. The cylinders were designed to be inserted into the bottom of a simulant mine and are marked with an engraved lid, which is therefore on the underside of the cylinder in normal use. When these ITOP inserts are used therefore, the engraved lid shall always be on the underside and the sensor height above the target and the target depth shall be measured to the top of the cylinder in this orientation (See Figure 11).

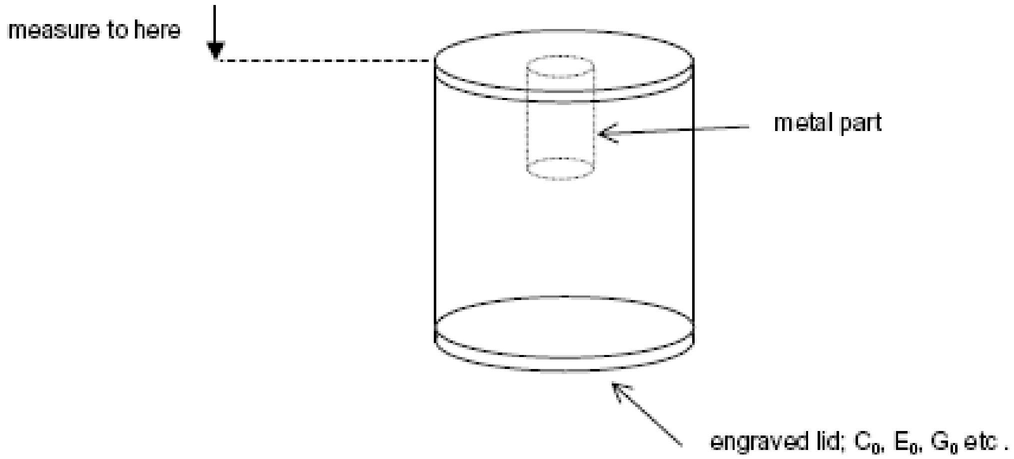


Figure 11 — Geometry and correct orientation of "ITOP insert" test target

ITOP Insert	Level of detectability	Metal Content
C_0	very difficult	carbon steel ball; diameter 3,2 mm
E_0	very difficult	vertical carbon steel pin; length 7 mm, diameter 1,6 mm
G_0	hard to detect	vertical copper tube; length 12,7 mm, outer diameter 3,2 mm, wall thickness 0,4 mm
I_0	hard to detect	vertical aluminium tube; length 12,7 mm, outer diameter 4,8 mm, wall thickness 0,38 mm
K_0	moderately difficult	Two (2) parts: 1. vertical carbon steel pin; length 7 mm, diameter 1,6 mm 2. vertical aluminium tube; length 12,7 mm, outer diameter 6,35 mm, wall thickness 0,38 mm
M_0	moderately difficult	vertical aluminium tube; length 38 mm, outer diameter 6,35 mm, wall thickness 0,38 mm
O_0	easiest to detect	Four (4) parts: 1. vertical carbon steel pin; length 14 mm, diameter 1,6 mm, 2. vertical aluminium tube; length 38 mm, outer diameter 6,35 mm, wall thickness 0,38 mm 3. vertical carbon steel spring; length 25,4 mm, outer diameter 8,7 mm, wire diameter 1,0 mm 4. carbon steel ball; diameter 6,35 mm

C.3 Specific mine test targets

The response of metal detectors to specific mines that are a threat in a particular area often needs to be easured. In addition, the response to particular mines that are difficult to detect is of interest. In such cases the response of a detector to an actual mine that has been made safe is often measured. There are difficulties with this approach. It may be difficult or impossible to make the mine safe enough (to be transported freely, for example), while preserving the metallic components in the position they would be in the armed state. A target designed to have the same metallic content, in configuration and properties of the components, may therefore be used instead.

If the test target is, or is intended to simulate a specific type of mine, the location, orientation and electromagnetic properties of the metallic components within it shall accurately represent those of the metallic components within that specific mine in its normal armed state. The ability of any given detector to detect real mines of a particular type may vary, since the metals used in the mine may not always be the same. In addition, corrosion of metal parts in mines that have been in the ground for several years will also make them harder to detect. The use of test targets to predict the ability of a detector to detect a real mine must always take this into account. For example, if a steel component in a mine is known to rust badly when left in the ground for several years, the signal from this component is likely to reduce. Modifying or even removing the component from a surrogate target to give the worst-case situation can simulate the response to the mine with a corroded component.

When a test target represents the metal component(s) from a particular type of mine, the metal piece(s) shall be mounted in an insulating holder or other means to reproduce the correct location and orientation of the component(s) from the top of the mine when in its normal armed state. The equivalent location of the top of the mine shall be clearly marked on any such holder or mount. The target-to-detector distance for detection shall then be measured to the marked “top of mine” point. It is the metal content that is the most important aspect of targets for testing metal detectors. However, in some noisy soils with some detectors, the volume of soil displaced by the mine may have an effect on the detector response. Therefore a target intended to replicate the response of an actual mine shall also have a similar overall shape and volume. If the test target is a generic simulation of a class or size of mine, the overall volume shall be typical of such mines.

C.4 Targets for shape determination

The following targets are to be used to test the capability of the detector to discriminate between point-like, linear and planar targets.

1) Steel ball as specified in the parametric test target set B.1 above, with diameter 5 to 10 mm.

2) Steel rod with circular, square or other cross-section. Largest dimension of the cross-section shall be 10 mm maximum. Length of rod shall be 50 mm. The rod shall be placed in the ground with its axis horizontal.

3) Steel disc of diameter 50 mm and minimum thickness 2 mm. The plane of the disc shall be placed horizontally in the ground.

B.5 Target depth and separation and sweep height

Since detection tests on targets buried at fixed depths in soil are performed, it is necessary to define standard target burial depths. The target burial depths shall be; 0 mm (surface flush), 50 mm, 100 mm, 150 mm and 200mm. IMAS 09.10, "Clearance requirements" [10] gives a guidance clearance depth of 130 mm. There may also be a requirement therefore, to test detection capability on targets at a depth of 130 mm. The burial depth is defined as the vertical depth of the top of the test object beneath the soil surface. In blind detection tests, burial flush with the surface leaves the target visible. In this case the top of the target shall be obscured either by a thin layer of soil - not greater than 10 mm thick, or by some other means. Minimum separation between adjacent targets shall be 0,5 m.

If the detection performance of the metal detector under test is to be determined for large, deep-buried metal targets such as metal anti-tank mines, the corresponding test targets shall also be buried at 300 mm below the surface. These different depths are chosen for testing metal detectors to evaluate different requirements for detection performance. Most anti-personnel mines are laid to be flush with, or just below the surface. After time they may end up further below the surface. Also the metallic content of the mine may be towards its bottom. Many metal detector detection capabilities vary rapidly at such short range, hence the choice of the additional 50 mm and 100mm testing depth. At ranges of the order of the coil radius, the relative performance of different detector coil configurations can change; in other words, a detector that is relatively sensitive at short range, may be relatively insensitive at long range. This is the reason for the 200 mm burial depth. Normal maximum sweep height above the surface of the ground shall be 30 mm. Sensor height above target shall be based on the above burial depths and sweep height. For example a sweep height of 30 mm combined with a target burial depth of 50mm is equivalent (in neutral soil) to a 80mm sensor height above target. Standard sensor heights above target therefore become 30 mm, 80 mm, 130 mm and 230 mm.

B.6 Criterion for determining true/false indication in blind test assessment

During blind tests, the locations at which operators report alarm indications shall be recorded. To determine whether these indications can be considered to be from the intended targets (i.e. they are true indications) or are false indications, the distance of the indication location from actual target locations shall be used. If the indication location is within a certain distance of a target location it shall be considered a true detection indication. If the indication is further than this from the target location, it shall be considered a false detection indication. The circle around the target location whose radius is defined by this maximum distance is known as the detection halo. The radius of the detection halo shall be half of the maximum horizontal extent of the metal components in the target plus 100mm.

D Annex: Information from manufacturers

Information from metal detector manufacturers that assists the test and evaluation process of their products for use in humanitarian demining:

1) Comprehensive operation manual.

2) Dimensions and mass of detector in operation.

3) Explanation of the operation mode of the detector; e.g. continuous wave or pulsed, frequencies used/ pulse bandwidth, dynamic/static response, differential sensor response, ground compensation functions.

4) Recommended sweep orientation, sweep pattern and sweep speed. Effect of different sensor head orientations.

5) Explanation of alarm operation, including criterion recommended for detection/ no detection decision.

6) Method to be used for locating ("pinpointing") objects that have been detected and the location accuracy attainable.

7) Sensitivity profile of the sensor head and recommended overlap between successive sweeps (for standard target depths) when searching.

8) Target or targets used for testing.

9) Detection performance as measured by manufacturer in air in terms of maximum detection height of specified target. Detection performance as measured by manufacturer in terms of maximum detection depth of a specified target in a specified soil.

10) Specification of limits of the properties of soils on which the detector will function.

11) Recommended battery type and typical operating life with one set of fresh batteries of this type.

12) If the detector circuitry has a back-up battery – the expected battery lifetime, with the detector in regular use or in storage.

13) Guidance on the extent to which parts or modules may be exchanged from one detector to another for field maintenance (see 10.3).

14) Compliance with environmental immunity tests, according to the tests specified in sections 7.3, 7.4 and 7.5. Compliance with other standards, for example for immunity to high temperature, low temperature, temperature shock, solar radiation, low pressure – in operation or in storage/transit conditions (details to be supplied).
Possible test standards; IEC 60721 series, IEC 60068-2 series.

15) Compliance with electromagnetic interference immunity tests according to section 7.7, or according to other EMC standards (test details to be supplied).
Possible test standards;
ENV 50204:1995, Electromagnetic Compatibility (EMC) - radiated electromagnetic field from digital radio telephones - immunity test
EN 55022: 1994+A1: 1995+A2: 1997, Information technology equipment – Radio disturbance characteristics – Limits and methods of measurement (class B)
EN 55011: 1998, Limits and methods of measurement of radio disturbance characteristics of industrial, scientific and medical (ISM) radio-frequency equipment
IEC 61000-4-2 (see section 2)
IEC 61000-4-3 (see section 2)
IEC 61000-4-8 (see section 2)
IEC 61000-6-2: 1999, Electromagnetic compatibility (EMC), Part 6-2: Generic standards – Immunity for industrial environments

IEC 61000-6-4: 2001, Electromagnetic compatibility (EMC), Part 6-4: Generic standards - Emission standard for industrial environments

16) Any guidance on the safe working distance to avoid mutual interference between detectors; of same type or different types.

17) Compliance with robustness and durability tests, according to section 10.1 or according to other standards (bump, shock, vibration) – in operation or in storage/transit conditions (test details to be supplied).
Possible test standards; IEC 60068-2 series

18) Compliance with other standards for equipment testing for conditions not covered by this CWA (test details to be supplied).
Possible standards;
IEC 60529: 1976, Specification for classification of degrees of protection provided by enclosures
(IP 66, IP 67 water and dust protection)

If the user needs other information he should contact the concerned manufacture and ask him to provide the necessary information., hints which information in his particular case are necessary he can take out of the ANNEX A of this document.

E Annex: Test forms

E.1 Manual measurement of sensitivity profile

Equipment details:			
Detector model and serial no: Time of "power on":	Detector setting: Target:	Sweep speed: Operator:	Date: Temperature (°C):
Time of day	**Sweep position (mm)**	**Max detection height (mm)**	**Comments**

E.2 Maximum detection depth of targets in soil/air (balls/ other specific targets)

Equipment details:			
Soil description:	Magnetic susceptibility[1]:	Electrical conductivity (S/m):	Humidity (% by vol):
Detector model and serial no: Time of "power on":	Detector setting: Sweep height:	Auto/manual sweep speed: Operator:	Date: Temperature (°C):
Time of day	**Target details** ***e.g.*** **ball diameter (mm)**	**Max detection depth (mm)**	**Comments**

E.3 Detection of fixed-depth specific targets in soil

Equipment details:			
Soil description:	Magnetic susceptibility[2]:	Electrical conductivity (S/m):	Humidity (% by vol):
Detector model and serial no: Time of "power on":	Detector setting: Sweep height:	Auto/manual sweep speed: Operator:	Date: Temperature (°C):
Time of day	**Target details**	**Burial depth (mm)**	**Detection (Y/N)**

[2] If possible, give frequency-dependent susceptibility also

F Annex: Detection halo

During blind tests, the locations at which operators report alarm indications shall be recorded. To determine whether these indications can be considered to be from the intended targets (i.e. they are true indications) or are false indications, the distance of the indication location from actual target locations shall be used. If the indication location is within a certain distance of a target location it shall be considered a true detection indication. If the indication is further than this from the target location, it shall be considered a false detection indication. If there are several indications within the halo, all accept one of them shall be ignored. The circle around the target location whose radius is defined by this maximum distance is known as the detection halo. The radius of the detection halo shall be half of the maximum horizontal extent of the target plus 100mm.

G Annex: Example of maximum detection depth measurements, design of experiment

There are four operators labelled A, B, C and D and four detectors labelled alpha, beta, gamma and delta. A possible design of experiment is given in Table 1. The order of execution is marked with the variable “start”. Each detector is used by each operator.

Table 1. Possible design of experiment for maximum detection depth measurements.

start	operator	detector
1	C	beta
2	D	alpha
3	A	delta
4	B	gamma
5	A	alpha
6	B	beta
7	C	gamma
8	D	delta
9	A	gamma
10	B	delta
11	C	alpha
12	D	beta
13	C	delta
14	D	gamma
15	A	beta
16	B	alpha

Any systematic influence related to time, for example tiredness, is “distributed” to all detectors and all persons equally. The order of execution is similar to random. Random order of execution would also be a good solution, but not the most practical one. With a random order it could easily happen that the same person appears in two starts in a row. If that is avoided, like in Table 1, than measurements can be performed faster, because each person can prepare for his/her start while the previous start is still running [6].

If these measurements are performed once, each detector will be used by each person exactly once and the total number of measurements with each detector will be four. The measurements can be repeated to increase the sample size and to enable an analysis of the differences between operators.

H Annex: Design of experiment for a reliability test based on a Latin square

Table 2 is an example of a 4x4 Latin square. Table 3 is an example of a 4x4 Graeco-Latin square. These squares are recommended as the basis for a trial design if four detectors are to be tested. Detectors are here marked with letters A, B, C and D. In Table 3, the letters mark the operators and the Greek letters mark the detectors. For other number of detectors, other Latin and Graeco-Latin squares should be used. Latin squares and Graeco-Latin squares are described in [10][6] and their use in metal detector tests in [13] and [6]. Graeco-Latin squares should be used if all operators are trained for all devices. In that case, crossover design should be applied [6]. With crossover design, it is possible to achieve that the operators work with fewer detector models at a time. This is strongly recommended, since it reduces stress on the operators. An example of such a design is provided in Table 4. In the first week, operators A and B are trained for detectors alpha and beta, while the other two operators use the other two detectors. In the second week they switch the detectors. The main disadvantage of such a design is that the training has to be performed in two rounds, what introduces additional costs.

Table 2. Example of a 4x4 Latin square.

	start 1	start 2	start 3	start 4
lane 1	A	B	C	D
lane 2	B	A	D	C
lane 3	C	D	A	B
lane 4	D	C	B	A

Table 3. Example of a 4x4 Graeco-Latin square.

	start 1	start 2	start 3	start 4
lane 1	A alpha	B beta	C delta	D gamma
lane 2	B gamma	A delta	D beta	C alpha
lane 3	C beta	D alpha	A gamma	B delta
lane 4	D delta	C gamma	B alpha	A beta

Table 4. Example of a design based on a Graeco-Latin square and crossover design.

Week 1:

	start 1	start 2	start 3	start 4
lane 1	A alpha	C delta	B beta	D gamma
lane 2	C gamma	A beta	D delta	B alpha
lane 3	B beta	D gamma	A alpha	C delta
lane 4	D delta	B alpha	C gamma	A beta

Week 2:

	start 1	start 2	start 3	start 4
lane 1	C alpha	A delta	D beta	B gamma
lane 2	A gamma	C beta	B delta	D alpha
lane 3	D beta	B gamma	C alpha	A delta
lane 4	B delta	D alpha	A gamma	C beta

Table 5 is an example of an experimental design based on a 3x3 Latin square. The variable “start” marks the order of execution. The experimental problem is to test three detector models, two specimens each, in six lanes. The letters A-L denote the operators, while the numbers denote the detectors. The first number (1, 2 and 3) denotes the detector model, the second number (1 and 2) the specimen (i.e. the copy) of the same model. Persons A B, C and D are trained to work with detector 1; persons E, F, G and H with detector 3; and persons I, J, K and L with detector 3. The detector models form 3x3 Latin squares; four of these squares are combined to form a round. After both rounds are executed, each detector will have been operated in each lane by all four persons trained for that detector. The major disadvantage of this design is the small number of persons operating each detector. In the case that there are large differences between the operators, they will have a large influence on the detector comparisons. The best would be to repeat rounds 1 and 2 with other four persons, but the time often does not allow that. In that case, it should be taken care that all four-person groups are similar in those properties that might influence the results, most importantly in their demining experience.

Table 5. Example of an experimental design based on a 3x3 Latin square.

		start 1		start 2		start 3		start 4		start 5		start 6	
round 1	lane 1	A	1-1	G	2-2	J	3-1	D	1-2	E	2-1	K	3-2
	lane 2	E	2-1	K	3-2	B	1-1	H	2-2	I	3-1	C	1-2
	lane 3	I	3-1	C	1-2	F	2-1	L	3-2	A	1-1	G	2-2
	lane 4	B	1-2	H	2-1	I	3-2	C	1-1	F	2-2	L	3-1
	lane 5	F	2-2	L	3-1	A	1-2	G	2-1	J	3-2	D	1-1
	lane 6	J	3-2	D	1-1	E	2-2	K	3-1	B	1-2	H	2-1

		start 1		start 2		start 3		start 4		start 5		start 6	
round 2	lane 1	B	1-1	H	2-2	I	3-1	C	1-2	F	2-1	L	3-2
	lane 2	F	2-1	L	3-2	A	1-1	G	2-2	J	3-1	D	1-2
	lane 3	J	3-1	D	1-2	E	2-1	K	3-2	B	1-1	H	2-2
	lane 4	A	1-2	G	2-1	J	3-2	D	1-1	E	2-2	K	3-1
	lane 5	E	2-2	K	3-1	B	1-2	H	2-1	I	3-2	C	1-1
	lane 6	I	3-2	C	1-1	F	2-2	L	3-1	A	1-2	G	2-1

I Annex: Confidence limits for the POD and the FAR

The upper and the lower confidence limits of the POD and those of the FAR can be computed with the help of many commercially available computer programmes or read from statistical tables. First the calculation of confidence limits for the POD is elaborated.

The number of opportunities to detect a target is marked *n*, and the number of detections *x*. The estimated POD is $\hat{POD} = x/n$. The number of detections is assumed to be binomially distributed with the parameter POD. The two-sided 1−α confidence limits [11] are

$$\begin{aligned} POD_{lower} &= \frac{x}{x + (n - x + 1)F_{1-\alpha/2,\, f_1,\, f_2}} \quad \text{with } f_1 = 2(n - x + 1),\ f_2 = 2x \\ POD_{upper} &= \frac{(x+1)F_{1-\alpha/2,\, f_1,\, f_2}}{n - x + (x+1)F_{1-\alpha/2,\, f_1,\, f_2}} \quad \text{with } f_1 = 2(x+1),\ f_2 = 2(n - x) \end{aligned} \tag{I.1}$$

The quantities $F_{1-\alpha/2,\, f_1,\, f_2}$ are F-quantiles (also called percentage points) of the F distribution. The difference between the upper and the lower confidence limit is called a confidence interval. Binomial confidence intervals are also called Clopper-Pearson intervals. In the special case when *x* = 0 the two-sided confidence limits are

$$\begin{aligned} POD_{lower} &= 0 \\ POD_{upper} &= 1 - \sqrt[n]{\alpha/2} \end{aligned} \tag{I.2}$$

When *x* = 0, than

$$\begin{aligned} POD_{lower} &= \sqrt[n]{\alpha/2} \\ POD_{upper} &= 1 \end{aligned} \tag{I.3}$$

The usual choice is α = 0,05, that is, 95% confidence limits.

The confidence limits for the FAR are described in the following lines. The number of false alarms in a single run follows a Poisson distribution. A variable which is a sum of Poisson distributed variables also follows a Poisson distribution. Consequently, the total number of false alarms *y* in *N* repeated scans over an area of size *A* also follows a Poisson distribution. The estimated false alarm rate is *y*/(*NA*). The two-sided confidence limits are [12]

$$\begin{aligned} FAR_{lower} &= \frac{1}{2NA}\chi^2_{\alpha/2,\, f} \quad \text{with } f = 2y \\ FAR_{upper} &= \frac{1}{2NA}\chi^2_{1-\alpha/2,\, f} \quad \text{with } f = 2(y+1) \end{aligned} \tag{I.4}$$

where $\chi^2_{\alpha/2,\, f}$ and $\chi^2_{1-\alpha/2,\, f}$ are called quantiles or probability points of the χ^2-distribution. In the special case when y=0, the confidence limits are

$$\begin{aligned} FAR_{lower} &= 0 \\ FAR_{upper} &= \ln(2/\alpha) \end{aligned} \tag{I.5}$$

Equations from (I.1) to (I.5) shall be used to calculate the 95% confidence limits for the POD and the FAR.

Before the widespread use of computers some approximative procedures had been developed. However, today their use seems to be hardly justified, since even the most common spreadsheet programmes can deal with the F-distribution and χ^2-distribution, functions necessary for computing the confidence limits of a binomial and a Poisson distributed variable. However, two extreme approximations can be helpful for a quick preliminary assessment of the size of the confidence intervals.

New abbreviations are introduced: p = POD and $q = 1 - p$. It has been proposed [10] that a normal approximation of a binomial distribution can be used if $n > 5$ and

$$\frac{\left|\sqrt{\frac{p}{q}} - \sqrt{\frac{q}{p}}\right|}{\sqrt{n}} = \frac{|p-q|}{\sqrt{npq}} < 0{,}3 \qquad \text{(I.6)}$$

This condition means that n is sufficiently large and p = POD is sufficiently far from 1 and 0. For the confidence level $1 - \alpha = 95\%$ and with some additional approximations [11] the following relation holds:

$$POD_{upper/lower} = p \pm 2\sqrt{\frac{pq}{n-1}} \qquad \text{(I.7)}$$

If p is close to 0,5, further approximation is possible:

$$POD_{upper/lower} = p \pm \frac{1}{\sqrt{n}} \qquad \text{(I.8)}$$

When y exceeds 15, the Poisson distribution can be approximated by a normal distribution. Setting the variance of that normal distribution to be equal to the variance of the Poisson distribution, we get approximate 95% confidence limits for the FAR:

$$FAR_{upper/lower} = F\hat{A}R \pm 2\sqrt{\frac{F\hat{A}R}{NA}} \qquad \text{(I.9)}$$

where FÂR = y/NA is the estimated FAR.

Because of their simplicity, equations (I.7), (I.8) and (I.9) can be helpful in the preparation phase of the experiment, for example to determine the sample size. Normally equations from (I.1) to (I.5) shall be used in trial reports, since F-quantiles and χ^2-quantiles can be easily calculated with most spreadsheet programmes or read from tables.

J Annex: Number of repetitions

The number of necessary repetitions depends on the least difference between PODs of metal detectors which needs to be discovered in the test. This least difference that needs to be discovered is here called a critical difference *d*. If the difference between the PODs of two detectors is larger than the critical difference *d*, the two detectors will have significantly different PODs. This *d* is a measure of the required precision of the whole test. The difference between PODs of two metal detectors can be roughly estimated as statistically significant if the confidence intervals of two PODs do not overlap. Confidence interval is an interval between the confidence limits. It is shown in Annex H that the 95% confidence limits for the case of POD close to 0.5 can be approximated as

$$POD_{upper/lower} = p \pm \frac{1}{\sqrt{n}}$$

where p is the estimated POD and n is the number of opportunities to detect a target. To detect differences $POD_1 - POD_2$ larger than a critical difference *d*, the following relation needs to be satisfied:

$$2\frac{1}{\sqrt{n}} < d \qquad \text{(J.1)}$$

This follows directly from the condition that the confidence intervals do not overlap. Written in a different form, this equation gives the approximate necessary number of opportunities to detect a target:

$$n > \left(\frac{2}{d}\right)^2 \qquad \text{(J.2)}$$

For example: The test is described with Table 3 in Annex H (Graeco-Latin Square). The goal is to compare detectors in lane 1 and to detect any differences larger than $d = 0.2$. There are 20 targets in each lane. It is expected that the results will be close to POD = 0.5. The number of opportunities to detect the target needs to be larger than $(2/d)^2$ for each detector, which is $(2/0.2)^2 = 100$. Because the number of targets in the lane is 20, the experiment described with Table 3 needs to be repeated at least five times, to get $n = 5 \cdot 20 = 100$. The repeated measurements shall not be performed according to exactly the same table, but with permuted operators, in a way that eventually all operators use each detector in each lane. Permuting the operators means switching their places in the design, for example, by replacing operator A with B, B with C, C with D and D with A. This so called circular permutation can be performed the necessary number of times. In this particular example 4 permutations may be even better than 5, because the differences between the persons would have an undesirable influence on the results observed in each lane separately if 5 permutations would be performed.

This example clearly illustrates that a universally valid prescription for the design of experiment and the number of repetitions does not exist. Each test is a new challenge and requires a unique design of experiment. The test organiser needs to be familiar with the principles of experimental design to make a valid design.

Only the differences in POD were discussed above. For the differences in FAR, a similar procedure should be applied, using Equation I.9 in Annex I.

K Annex: POD curves

The detection of a target is modelled as a Bernoulli experiment where the binary random variable Y takes its value y = 1 (“detected”) with the probability p and its value y = 0 (“not detected”) with the probability $1-p$. The parameter p is specific for each choice of operator-detector-lane combination and it depends on the influence variables characterising that treatment. We cannot relate p linearly with the influence variables, since it is limited to $0 \le p \le 1$. Therefore, the parameter p of the Bernoulli distribution is transformed into the parameter η:

$$\eta = \ln \frac{p}{1-p} \qquad \text{(D.1)}$$

This transformation is called logistic (or logit) transformation and the inverse function

$$p = \frac{1}{1+e^{-\eta}} \qquad \text{(D.2)}$$

is called the logistic function. It is a monotonically increasing S-shaped curve starting with $p(-\infty) = 0$ and ending with $p(\infty) = 1$. The parameter η, which is between $-\infty$ and $+\infty$, is linearly related to the influence variables:

$$\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_m x_m = \sum_{j=0}^{m} \beta_j x_j \qquad \text{(D.3)}$$

where one of the x_j’s stands for the depth of the target and the other x_j’s are indicator variables with values 1 or 0 indicating the presence of a particular level of a qualitative factor. This model is called a generalised linear model [14]. The unknown parameters β_j of the generalised linear model are estimated by maximum likelihood estimation. The result is a curve of POD versus target depth for each combination of other factor levels. The calculation of confidence bounds to POD curves is described in [15].

Bibliography

[1] CWA 14747 CEN Workshop Agreement, June 2003; ICS 95.020, Ref N° CWA 14747:2003 (E); European Committee for Standardisation, Management Centre: rue de Stassart, 36 B-1050 Brussels

[2] CWA 14747 Part 2 CEN Workshop Agreement, June 2003; European Committee for Standardisation, Management Centre: rue de Stassart, 36 B-1050 Brussels

[3] **Normative References of the CWA 14747**: IMAS 04.10, Glossary of mine action terms and abbreviations, First Edition, 01 October 2001, UNMAS, New York;IMAS 03.40, Test and evaluation of mine action equipment, Draft First edition, 01 October 2001, UNMAS, New York; EN 61000-6-1:1997, Electromagnetic compatibility (EMC), Part 6-1: Generic standards – Immunity for residential, commercial and light-industrial environments. (IEC 61000-6-1:1997, modified);IEC 60068-2-27:1987 Basic environmental testing procedures. Part 2: tests – test Ea and guidance: shock; IEC 60068-2-29:1987 Basic environmental testing procedures. Part 2: tests – test Eb and guidance: bump; IEC 61000-4-2:1995 + A1:1998 + A2:2000, Electromagnetic compatibility (EMC), Part 4-2: Testing and measurement techniques – Electrostatic discharge immunity test. (=EN 61000-4-2:1995 + A1:1998 + A2:2001), IEC 61000-4-3:2002, Electromagnetic compatibility (EMC), Part 4-3: Testing and measurement techniques – Radiated, radio-frequency, electromagnetic field immunity test. (=EN 61000-4-3:2002);IEC 61000-4-8:1993 + A1:2000, Electromagnetic compatibility (EMC), Part 4-8: Testing and measurement techniques – Power frequency magnetic field immunity test. (=EN 61000-4-8:1993 + A1:2001)

[4] D. Guelle, A. Smith, A. Lewis, T. Bloodworth; "METAL DETECTOR HANDBOOK FOR HUMANITARIAN DEMINING" ISBN 92-894-6236-1, © Commission of European Communities, 2003

[5] D. Guelle, A. Lewis, P. Ripka "**METAL DETECTOR TRIALS** DETECTOR TEST RESULTS AND THEIR INTERPRETATION", EUR 22534 EN**;** EUR - Scientific and Technical Research series; ISSN 1018-5593

[6] M. Gaal, "Trial Design for Testing and Evaluation in Humanitarian Mine Clearance", PhD thesis, Brandenburg Technical University Cottbus, Germany, July 2007. Available at http://se6.kobv.de:8000/btu/volltexte/2007/349/ and at http://www.itep.ws/

[7] STEMD Project Laos D.M. Guelle, A. M. Lewis, M. A. Pike, A. Carruthers, S. M. Bowen; Systematic Test and Evaluation of Metal Detectors (STEMD). Interim Report Field Trial Laos, 27th September - 5th November 2004 www.itep.ws

[8] STEMD Project Mozambique D. M. Guelle, A. M. Lewis, M. A. Pike, C. Craill; Interim Report Field Trial Mozambique, 12th April - 5th May 2005Guelle at all; Lab tests JRC Ispra A M Lewis at all; search STEMD www.itep.ws

[9] STEMD Project Croatia D. Guelle, M. Gaal, M. Bertovic, C. Mueller, M. Scharmach, M. Pavlovic, "South-East Europe Interim Report Field Trial Croatia (STEMD-continuation)", March 2007. Available at http://www.itep.ws/

[10] G. E. P. Box, W. G. Hunter, J. S. Hunter, Statistics for Experimenters, John Wiley & Sons, 1978.

[11] K. M. Simonson, Statistical Considerations in Designing Tests of Mine Detection Systems: I — Measures Related to the Probability of Detection, technical report, Sandia National Laboratories, August 1998. Available at http://www.itep.ws/.

[12] K. M. Simonson, Statistical Considerations in Designing Tests of Mine Detection Systems: II — Measures Related to the False Alarm Rate, technical report, Sandia National Laboratories, August 1998. Available at http://www.itep.ws/.

[13] C. Mueller, M. Gaal, M. Scharmach, U. Ewert, A. Lewis, T. Bloodworth, P.-T. Wilrich, and D. Guelle, "Reliability Model for Test and Evaluation of Metal Detectors". Final report of the ITEP project 2.1.1.2, Federal Institute for Materials Research and Testing (BAM), Berlin, Germany, Sep 2004. Available at http://www.itep.ws/.

[14] StatSoft Inc. "Electronic Statistics Textbook", Tulsa, OK, USA, 2007. http://www.statsoft.com/textbook/stathome.html/

[15] C. Mueller et al, Reliability Model for Test and Evaluation of Metal Detectors – Part 2. Final report of the ITEP project 2.1.1.8, Federal Institute for Materials Research and Testing (BAM), Berlin, Germany. Soon available at http://www.itep.ws/.

[1] CWA 14747-2 Draft version 2 (under discussion in CEN WS7); the final document will be available in spring 2008 (www.cenorm.be and www.itep.ws)