VALIDITY AND ANALYTICAL ROBUSTNESS OF THE OLSEN SOIL P TEST AND OTHER AGRONOMIC SOIL P TESTS USED IN NORTHERN EUROPE

GITTE HOLTON RUBÆK (EDITOR)

DCA REPORT NO. 071 • DECEMBER 2015

AARHUS
UNIVERSITY
DCA - DANISH CENTRE FOR FOOD AND AGRICULTURE

Validity and analytical robustness of the Olsen soil P test and other agronomic soil p tests used in the northern Europe

Supplementary information and clarifications (October 2019)

In an effort to ensure that this report complies with Aarhus University's guidelines for transparency and open declaration of external cooperation, the following supplementary information and clarifications have been prepared in collaboration between the researcher (s) and the faculty management at Science and Technology:

The preface mentions that a draft version of the updated analytical protocol for the Olsen soil P test presented in Appendix 1 was sent for commenting to the three laboratories (Agrolab, Ok lab and Eurofins-Steins). The update of the protocol was requested to improve the robustness of the method and at the same time ensure that the method in principle remained the same. The laboratories commented on correctness of the updated protocol and on whether they would foresee substantial difficulties implementing the suggested changes in their laboratories.

This input from the laboratories was essential to ensure that the updated protocol was feasible not only in one research laboratory, but also in large-scale commercial laboratories. Based on the input from the laboratories the literature review and data analyses presented by researchers (chapter 3-5) the board made their recommendations (chapter 2) and decided how key issues in the analytical protocol in appendix 1 was handled in the final version.

Therefore, the recommendations in chapter 2 and the final version of the protocol in appendix 1 expresses what the entire board agreed upon (consensus) and the author and board member Gitte Rubæk was the one who put this into writing.

VALIDITY AND ANALYTICAL ROBUSTNESS OF THE OLSEN SOIL P TEST AND OTHER AGRONOMIC SOIL P TESTS USED IN NORTHERN EUROPE

GITTE HOLTON RUBÆK (EDITOR)

DCA REPORT NR. 071 · DECEMBER 2015

AARHUS
UNIVERSITY
DCA - DANISH CENTRE FOR FOOD AND AGRICULTURE

Gitte Holton Rubæk (Editor)

Aarhus University
Department of Agroecology
Blichers Allé 20
PO Box 50
DK-8830 Tjele

VALIDITY AND ANALYTICAL ROBUSTNESS OF THE OLSEN SOIL P TEST AND OTHER AGRONOMIC SOIL P TESTS USED IN NORTHERN EUROPE

Series:	DCA report
No.:	071
Authors:	Gitte Holton Rubæk (Editor)
Publisher:	DCA - Danish Centre for Food and Agriculture, Blichers Allé 20, PO box 50, DK-8830 Tjele. Tlf. 8715 1248, e-mail: dca@au.dk Web: www.dca.au.dk
Commissioned by:	The Danish Ministry of Environment and Food, Environmental Protection Agency
Photo:	Colourbox
Print:	www.digisource.dk
Year of issue:	2015
	Copying permitted with proper citing of source
ISBN:	978-87-93398-18-4
ISSN:	2245-1684

Reports can be freely downloaded from www.dca.au.dk

Scientific report
The reports contain mainly the final reportings of research projects, scientific reviews, knowledge syntheses, commissioned work for authorities, technical assessments, guidelines, etc.

Preface

Analyses of soil P status with soil P tests have for for many years formed the cornerstone for recommendations on how to fertilise agricultural soils. Recommendations are typically based on the economic balance between the cost of the fertiliser P and the yield depressions you can expect if soil P status is limiting crop production. If these principles are followed, soils will be fertilised to obtain and maintain a soil P status supporting optimal plant production, and P fertilisation beyond that will not take place. In areas dominated by intensive animal production, livestock manure is the main source of plant nutrients. In such areas P is often added in surplus year after year due to the unfavourable N-P ratio in livestock manures. As a result, areas exist where soil P levels are considerably higher than required by the crops. Fortunately the restrictions on livestock density on agricultural land, which for decades have been a key regulatory parameter in Danish legislation ("Harmonireglerne"), have set upper limits for the yearly P additions. Danish soils have therefore not received the extremely high P doses as known from, for example, regions in the Netherlands.

The increasing awareness of the role of soil P as a contributor to surface water eutrophication together with the renewed focus on phosphate rock as a valuable non-renewable resource has put emphasis on the way we utilise P in soil, fertilisers, manure and waste products. To ensure and improve optimal utilisation of P in soil, a valid, precise and reliable soil P test method is crucial, which becomes even more important when a soil P test designed and primarily used for advisory purposes is engrafted in the rules and regulations where it typically is used for defining limits for how much phosphorus can be applied to a given field.

Olsen P (in Denmark known as "Ptallet" or "fosfortallet") was selected as the "official" soil P test method in Denmark in 1987. The method was chosen in consensus by ministries and research institutions as the best and most universal method for estimating soil P status on agricultural soils based on literature reviews and investigations on Danish soils. In Denmark we therefore have almost thirty years of experience with this soil P test method and a comprehensive database with test results.

It has long been recognised that results for reference soils analysed in proficiency test programmes in commercial soil laboratories vary too much and apparently systematically between labs and over time. It is therefore clear that initiatives leading to better soil P tests in Denmark with high laboratory precision and valid information on soil P status for farmers, researchers and authorities are highly needed.

Moreover an increasing body of evidence seems to indicate that the Olsen-P method too frequently does not reflect the P availability to plants in soil, which leads to erroneous predictions of fertiliser P

requirements. New alternatives to the Olsen P methods now exist and these are discussed in the present report with special focus on the DGT (Diffusive Gradients in Thin Films) method.

This report presents an update and elaboration on the recommendations on how to improve soil P testing and the quality control of soil analyses in Denmark given in Rubæk and Sørensen (2011). It is commissioned by The Danish Ministry of Environment and Food, Environmental Protection Agency, who also funded the work. The work has been supported and supervised by an advisory board consisting of:

- Søren Husted, Department of Plant and Environmental Sciences (PLEN), KU Science.
- Leif Knudsen, SEGES
- Esben Jensen, Agrolab
- Hans Estrup Andersen, Department of Bioscience, Aarhus University
- Jørgen Eriksen, Department of Agroecology, Aarhus University
- Henriette Hossy, Nikolaj Ludvigsen, The Environmental Protection Agency (Board observers)
- Anders Nemming, The Danish Agrifish Agency (Board observer)

The Advisory board met on four occasions (23/9-2014, 6/2-2015, 23/4-2015 & 17/9-2015) to discuss the objectives, progress and outcome of the work and agree upon recommendations. I wish to thank the board for their support and constructive input to this report.

A draft version of the analytical protocol presented in Appendix 1 was sent for commenting to the three laboratories (Agrolab, OK lab and Eurofins-Steins), who currently all participate in the voluntary proficiency test programme arranged by SEGES. I am grateful for the thorough and prompt responses from all three laboratories. The comments from the laboratories were discussed at the final meeting of the advisory board and formed a significant input to the recommendations and to the protocol presented in Appendix 1.

During the work, several other persons, laboratories, and research institutions were consulted and I am grateful for the help I have been given by:

- Dr. Maria Kreimeyer and Dr. Markus Rupprecht, Agrolab, Germany
- Dr. Wim Chardon, Alterra, Wageningen UR, the Netherlands
- Winnie van Vark, Wepal, Waageningen University, the Netherlands
- Dr. L. Blake and Dr. M.M.A Blake-Kalff, Hill Court Farm Research Ltd, UK
- Prof. Tore Krogstad, Norwegian University of Life Sciences, Norway
- See Mei Ngo, Lisbeth Hartzell, and Martin Frandsen, Eurofins Sweden and Denmark, respectively
- Lene Skovmose og Mette Sahl Haferbier, Department of Agroecology, Aarhus University, Denmark

Gitte H. Rubæk,
December 2015

Table of contents

1. Background

Gitte H. Rubæk, Department of Agroecology, Aarhus University

There is a long tradition of using results from laboratory analyses of agricultural soils as basic input to fertilisation recommendations. Soil phosphorus tests are one important group of such analyses. A wide range of soil P test methods exists worldwide, and typically within a country there is consensus on which soil P test method is to be used for given soil types and purposes. However, if you cross borders there is no consensus on either which soil P test to use or on the fertiliser recommendations it leads to even when the same soil P test method is used (Jordan-Meille et al., 2012). This causes much confusion and difficulty when trying to compare soil phosphorus status and phosphorus fertilisation recommendations among countries. Even though this is a well-known drawback, most countries are reluctant to change soil test methods. That is mainly because: (1) each country/region has chosen a method they trust to suit their dominant agricultural soils; (2) local documentation exists from field experiments for the threshold values of the P tests used in recommendation schemes, and (3) there is a considerable knowledge base and familiarity with the results of the old method. Changing to another method requires significant efforts and resources to establish the sufficient knowledge base, familiarity, new thresholds and recommendation schemes for the new method.

In Denmark, we changed from an extraction with dilute sulphuric acid (fosforsyretallet, Bondorff, 1950) to the Olsen P soil test (Olsen et al., 1954) in 1987. This change was based on scientific evidence that the Olsen P was more reliable in its prediction of plant-available soil P (e.g. Olsen et al., 1954; Sibbesen, 1983, Nielsen, 1979 and 1981). Soils sampled in 1987 from farmed fields were analysed with both the new and the old method and in the following approx five years both analytical methods were used on a steadily decreasing number of the samples (Leif Knudsen, personal communication). The version of the Olsen soil P test used in Denmark is in principle the modification described by Banderis et al. (1976). A formal Danish description of Danish soil P test (Ptallet) and other soil analyses was last updated in 1994 by the Danish Ministry of Agriculture (Plantedirektoratet, 1994). An ISO-Standardised description of the original Olsen P was published in 1994 (ISO 11263, 1994).

Other applications of soil P tests

Today soil P tests are increasingly used for other purposes than just fertiliser recommendations, for example for estimating the amount of P stored in agricultural soils and the associated risk of losing P to the environment (Heckrath et al., 2009), and Soil P testing is becoming an issue in relation to regulation of manure and fertiliser application. In Denmark the Commission on Nature and Agriculture has, for example, recently suggested a more coherent regulation of the use of phosphorus in Danish Agriculture

(Natur og Landbrugskommissionen, 2013). This has been followed up by ideas of including norms for P application in future regulations. From a plant nutritional perspective, norms for P application should ideally take soil legacy P, i.e. fertiliser P accumulated in soils estimated by a soil P test, into account plus the expected removal of P by the crop (Poulsen and Rubæk, 2005). Such a system aims for a balanced fertilisation where soil P levels are adequate, while soils with a low soil P status would be allowed a norm corresponding to the off-take plus a little more and soils with a high soil P status would be allowed a lower amount of P fertilisation than what is expected to be removed by the crop. This corresponds to the way P fertilisation recommendations typically work (Jordan-Meille et al., 2012).

Since soil P test methods are increasingly used for a range of purposes, it becomes even more crucial for all stakeholders that the P test used is valid, robust, and reliable.

Quality assurance of soil analyses

The quality analytical work in soil laboratories is typically assured by accreditation of the labs, in some cases even by authorisation and the use of well-described standardised analytical protocols. The analytical performance of the laboratories is then checked and compared on a regular basis in proficiency testing programmes where all participating labs analyse a number of reference soils and report their results to the institute arranging the test programme (e.g. Rubæk and Sørensen, 2011). A report is then prepared where the performance of the labs is evaluated.

In Denmark we had a national authorisation system with a public laboratory supervising labs and proficiency testing until 2003, when the supervising laboratory closed and the system was abandoned. In replacement, the Danish Knowledge Centre for Agriculture (now SEGES) organised a a proficiency test programme, which is offered to the laboratories carrying out soil tests based on the Danish method descriptions (Plantedirektoratet, 1994) for the most common soil analyses, and all labortories carrying out these soil analyses have chosen to participate in the programme. Results of these tests are published yearly at www.landbrugsinfo.dk (Videncentret for Landbrug, 2014). These ring test programs and the former programmes arranged before 2003 show very clearly that the number of laboratories offering soil analyses according the Danish method descriptions has declined dramatically (seven labs in 1997 down to three labs from 2010 onwards). The number of labs is now far below the minimum required for a classical proficiency testing programme. To compensate for this, SEGES developed a test system which includes more soil samples than the classical proficiency test programmes, and they furthermore use the subsamples of the same soils repeatedly. This strategy has the advantage that it allows comparison of results obtained at different times and years. It is therefore well documented that for the P test (Ptallet), there are problems asthere is a very large variation between results obtained for different test campaigns

on the same reference soil samples, which has also been documented earlier on an older data set. These observations have critically lowered the credibility of the P test results in Denmark.

Objectives and method

In 2009, The Knowledge Centre for Agriculture (now SEGES), University of Copenhagen and Aarhus University raised these problems with the Ministry for Food, Agriculture and Fisheries and the Ministry of Environment, requesting initiatives to: (1) assure the quality of soil analyses in Denmark, (2) revise and update the method description, and (3) ensure that soil testing in Denmark is carried out with uniform, well-described updated methods supported by the main Danish stakeholders using analytical data on agricultural soils. Based on this the ministries initiated a process leading to a synthesis report on the quality and applications of soil analyses in Denmark (Rubæk and Sørensen, 2011) and recommendations from the advisory board supervising the writing process (Kristensen et al., 2011).

The objective of the present report is therefore, first of all, to update and elaborate the recommendations on soil P testing given in Rubæk and Sørensen (2011). The present report includes:

- Updated and elaborated recommendations regarding soil P testing in Denmark (Chapter 2).
- A brief summary of earlier conclusions regarding the validity of some existing standard soil P tests used in our neighbouring countries for agronomic and environmental recommendations and regulations, including a brief update on the work with the new DGT-method at University of Copenhagen and a description of the main differences in existing Olsen P method variations (Chapter 3).
- A presentation of the dominant, well-established, routine P test methods used in Denmark and neighbouring countries and of the proficiency test programmes, which include these soil P tests (Chapter 4).
- A description of how the robustness of the Olsen P method can be improved by correction or calibration (Chapter 5).
- An updated method description for Ptallet/Olsen P (Appendix 1).

2. Recommendations

Gitte H. Rubæk, Department of Agroecology, Aarhus University

The recommendations below are based on the work presented in the subsequent chapters of this report and on the discussions conducted by the advisory board. Table 2.1 attempts to summarise the evaluation of the different soil P analyses used in Denmark's neighbouring countries and Table 2.2 gives an overview of the key comments from the laboratories on a draft of the preliminary protocol in Appendix 1 and a brief summary of the rationale forthe recommendations.

In agreement with the advisory board, the following is recommended:

1. *A permanent advisory board or forum for stakeholders regarding analyses of agricultural soil (ministries, research institutes, farmers and farmers' advisory service, laboratories) is established. The remit for this board should be to: (1) oversee the quality of the soil P test and other soil analyses; (2) suggest improvements for soil test methods and quality control on soil testing, and (3) suggest and supervisetiming and changes to the soil P test method and other soil analyses.*
2. *Currently we recommend that soil P tests in Denmark should be carried out according to an updated analytical protocol for the bicarbonate-extractable soil P test (in Denmark known as "Ptallet" and internationally often referred to as Olsen P). A preliminary protocol for the method is presented in Appendix 1.*
3. *The protocol in Appendix 1 should be revised and finalised when the pending issues specified in Table 2.2 have been clarified.*
4. *Soil P tests for regulatory purposes in Denmark should be carried out in accredited laboratories that participate in sufficiently comprehensive international proficiency testing programmes, e.g. WEPAL or BIPEA and in a test programme similar to that organised by SEGES.*
5. *A portfolio of reference soils corresponding to those presently maintained by SEGES should be made available for laboratories carrying out P tests in Denmark.*
6. *Laboratories should be obliged to document their results and analytical error pertaining to the measurements of bicarbonate-extractable P according to the updated method description.*
7. *Whether to implement a correction to the measured bicarbonate-extractable P based on calibration for at least four standard/reference soils as described in Chapter 5 should be decided after further discussions in the advisory board/stakeholder forum.*
8. *The bicarbonate-extractable P method should be replaced with a more valid method (i.e. a method that is more robust and assesses equally well the need for P fertilisation on all major agricultural soils in Denmark), when such a method is available, cost-effective and ready for implementation as a routine soil P test.*

The main reasons for the recommendations of choice of method are:
(1) The Olsen P method has shortcomings and does not work equally and well on all soil types. Therefore a change to a method for which there is clear scientific evidence and that it is more valid would be advantageous when such a method is ready for implementation. The DGT technique, which is currently being evaluated in Denmark, may turn out to be such a future alternative to the Olsen P method.
(2) The Olsen P method/Ptallet is well-established as the standard procedure for assessing plant availability of P in Denmark and many other countries and is considered to be more valid for agronomic purposes for Danish conditions than other routine soil P tests used in our neighbouring countries;
(3) Comprehensive experience and a valuable collection of regional and historical data exist based on the Olsen P method.

The main reasons for suggesting a preliminary and not a final protocol for bicarbonate extraction in Appendix 1 are: (1) To allow the protocol to be tested in practice before making it final; (2) to avoid enforcement of changes in the protocol at short notice that can be difficult and costly for the laboratories and without proper documentation for the effect of the changes and for potential alternatives; (3) to leave time for key elements in the protocol to be tested before implementing them. For further details see also Table 2.2.

Table 2.1. A rough summary of our evaluation of the validity for agronomic and environmental purposes of some soil P test methods and their applicability for routine soil testing. Three stars indicate "good" validity, two stars "fair" and one star "weak". A question mark indicates that our estimate, if given, is not based on sufficient evidence to form a judgement.

P test method	**The Danish P tal/Olsen P**	**Updated Danish Ptal as suggested in this report including correction**	**CAL-P**	**AL-P**	**DL**	**Pw**	**DGT**
Method description	Plantedirek-toratet, 1994/ISO 11263:1994	This report	Schüller, 1969	Egner et al., 1960	Egner and Reihm, 1955	Sis-sing, 1971	In pro-gress
Validity as a guideline for fertilizer recommendations (Danish conditions)	**	**	*	*	*	**?	***?

Validity as a partial risk indicator for dissolved P through surface run-off and leaching (Danish conditions)	**	**	*?	*?	*?	***	***?
Validity for estimating how much P is accumulated in an agricultural soil	**	**	**	***	***	*	*
Detailed method description available (ISO or similar)	Yes	Yes	Yes	Yes	Yes	Yes	Not yet
Ease of usein routine soil lab	*1)	*?1)	***	*2)	** 3)	*?4)	?4)
Cost-effectiveness in the lab (relative cost)8)	1.1	1.4	1	2	1	2.55)	?
Robustness	*	***?	***?6)	**?6)	**?6)	?	?
Feasibility of including in ring tests	***	***7)	**	**	**	*	Not yet

1) (-) Due to lack of chemical equilibrium, all outer conditions like temperature, shaking conditions, time span between shaking and separation of soil and solution have a large influence on the results. Eliminating gaseous CO_2 from the extract is tedious and time-consuming.
2) (-) Four hours' shaking time; lactic acid must be stored 48 h at 100 °C before usage. Concentration of all components must be determined.
3) (-) Extraction solution is not stable.
   (+) K, P and Mg measurement in the same extract.
4) (-) Analysis takes several days.
5) High costs due to multiple days' handling.
6) Question mark added due to limited information on robustness.
7) The ring test would be on the uncorrected result of the analysis.
8) Estimates of relative costs provided by Maria Kreimeier, Agrolab. A key issue making the otherwise very simple protocol for water extraction more expensive is the overall processing time, which for this analysis is more than 24 hours.

Table 2.2. Key elements in the protocol for bicarbonate extraction identified by the board, which need to be further specified, the comments from the laboratories on the draft protocol and the reasoning by the board before giving their recommendation.

Element in the existing protocols for bicarbonate extraction identified by the board as not sufficiently well described or in need of an update	**Comments from the laboratories**	**The board's rationale for its final recommendation**
Drying temperature of soil prior to analysis. The existing Danish method description stipulates drying at 50-60 ^{o}C, while the ISO-standard stipulates 40 ^{o}C. The board prefers a drying temperature of 40 ^{o}C because of its reduced impact on the soil.	The laboratories currently dry at 50-60 ^{o}C and state that it will be costly and difficult for them to implement drying at 40 ^{o}C, because the drying process will take longerwhich is not compatible with the drying capacity for the high number of soil samples processed daily.	The board decided to keep the drying temperature at 50-60^{o}C for this preliminary version of the protocol. However, the aim is to reduce the temperature in the final version, after having documented and quantified the importance of this change in the protocol.
Amount of soil and dimensions of extraction containers. The dimensions of the container used for extraction should be specified in relation to the amount of soil and extraction solution. The board is in favour of allowing less soil for each extraction since this is common practice in research labs as it eases centrifugation and reduces the amount of chemicals needed.	One laboratory questions the decision to allow as little as 1 g of soil per analysis, as this might increase the variability of the result.	The board sticks to their first suggestion to allow smaller amounts of soil per analysis, but states in the protocol that variability might increase with smaller amounts of soil per analysis and that 5 g would be preferred for routine analyses.
Shaking method, type and speed of rotation. The board is in favour of end-over-end shaking because it is standard procedure in most soil labs and low speed because it minimizes disaggregation during shaking.	Some laboratories question the importance of this.	The board sticks to their first suggestion.
Temperature thoughout the extraction procedure. It is crucial that the temperature is kept at the specified level throughout the analysis until soil and solute have been separated. The board finds that is is reasonable to aim at the same temperature as used in the ISO-standard (ISO 11263, 1994) as that ensures direct compatibility with this standard and with proficiency test programmes for this method.	Some laboratories state that keeping a lower temperature is challenging, especially in summer.	The board sticks to their first suggestion.

Time spent on extraction and handling of samples. Clear and narrow limits have to be specified for how much the extraction time can deviate and how much time can be allowed for handling samples after extraction .	The laboratories gave information on how fast a set of samples could be handled at present, and stated that the initially suggested time for handling of samples is unrealistic in their laboratory procedures.	The board has now specified the acceptable time limit for handling samples after extraction, which is less strict than our first suggestion but expected to be realistic in routine laboratories.
Method for separating soil and solute after extraction. Method for soil and solute separation should be specified. Basically the board prefers separation by centrifugation at a fixed temperature similar to the extraction temperature, because theyexpect this procedure to lead to the most well-defined separation.	It became clear that the laboratories all use different separation methods: Classical filtration, filtration under pressure and centrifugation. One lab had compared filtration and centrifugation, with a surprising result. The laboratories stated that it would be costly and time-consuming to change method.	Due to very different practices at the labs, the limited documentation of the importance of the different separation methods and the difficulties and costs related to a change, the board decided to allow all existing separation methods in the preliminary protocol, but with the aim of selecting one method in the final protocol based on a new thorough comparison of the separation methods.

Laboratory facilities for high throughput analyses of bicarbonate extactable P at Agrolab, Germany.

3. Soil P test methods and their validity for agronomic and environmental purposes

Gitte H. Rubæk, Department of Agroecology, Aarhus University
Simon Mundus and Søren Husted, Department of Plant and Environmental Sciences, KU Science

Only a minor part of the total soil P is present in the soil solution and therefore immediately available to the crop. However a large proportion of the total soil P is reversibly bound to the soil constituents or temporarily immobilised in soil biota. This P is constantly being exchanged with the solution where it replenishes the concentration, when the solution is depleted by crop P uptake, and retains fertiliser P entering the soil solution in high concentrations, i.e. buffers the soil solution P concentration. This buffering capacity takes place through numerous and complex chemical equilibrium processes and biological mobilisation and immobilisation, and it governs the potential plant-availability of soil phosphorus and crop P nutrition and therefore typically depends more on the soil P status than on the amount of fertiliser P added prior to a growing season. Soil properties such as mineralogy, texture, pH, organic matter content define the soil type and soil chemical properties, and soil P buffering capacity is therefore highly dependent on soil type (Frossard et al., 2000).

Due to the complexity and nonlinerarity of the processes governing P binding and release in a soil, it is in principle impossible to fully describe the potential availability to crops of P in a soil by one single number. That is nevertheless what nearly all soil P tests are aiming at. They estimate soil P status with one single number in mg/kg soil from a simple extraction. This number is used as an indicator for potential availability of P to plants in the soil and included in fertiliser recommendation schemes (Jordan-Meille et al., 2012).

Because of not only the different soil types, but also the traditions in different countries and regions, many different soil P tests in exists (Rubæk et al., 2011, Jordan Meille et al., 2012, Beegle, 2005, Sibbesen and Sharpley, 1997, Tunney et al., 1997). Some of the methods used in Northern Europe are shown in table 2. These methods are fairly simple chemical extractions. When used on single soil types with different P fertilisation histories, they typically measure well how much P is accumulated in the soil and it is mostly possible deduct reasonable relations between yield responses to P fertilization and soil P test level . Acid extractants are typically preferred in areas with acid soils, while the bicarbonate extraction is typically preferred in countries with weakly acid, neutral to alkaline soils.

Table 3.1 Examples of P extraction methods used for P fertiliser recommendations in European countries.

Method	**Countries**	**Reference**
Ammonium lactate, P-AL (ammonium lactate/acetic acid, pH 3.75)	Belgium, the Netherlands, Norway, Sweden, Germany	Egner et al., 1960
Double lactate, P-DL (calcium lactate, hydrochloric acid, pH 3.7)	Belgium, Germany	Egner and Reihm, 1955
Olsen (sodium bicarbonate, pH 8.5)	Denmark, Italy, France, England, Wales, Northern Ireland	Olsen et al., 1954
Morgan (sodium acetate, pH 4.8)	Ireland	Morgan, 1941
CAL (calcium lactate/calcium acetate, acetic acid, pH 4.1)	Austria, Belgium, Germany	Schüller, 1969

Due to the complex processes governing soil P retention and release and the variable conditions facing a growing plant, the single number provided by any soil test as an estimate of the plant-available soil P can only serve as a rough estimate or index for how well the soil potentially can supply P to a crop. However, a good soil P test should extract from exactly the same pools as the plants do (e.g. Mason et al., 2013, Six et al., 2012; Olsen et al., 1954). The methods extracting a large proportion of soil total P typically fail in that respect because they also extract P which is not immediately available to plants. Soil P tests therefore typically have to be calibrated and interpreted differently on different soil types (Sibbesen and Sharpley, 1997). Furthermore the actual amount of P available to the growing crop can deviate substantially from the potential availability depending on the actual growing conditions, including situations where subsoil contributes substantially to crop nutrition, while only the topsoil is included in the soil testing (Rubæk et al., 2013, Tóth et al., 2014). Clear relations between soil P status and responses to P application in crop growth under field conditions can therefore be difficult to establish and require comprehensive field experiments including several sites and experimental years.

The proportion of total soil P extracted varies considerably for the soil tests in table 2. The amount decreases in the following order: Ammonium lactate (AL), DL, CAL, Olsen, Pw (Neyroud and Lischer, 2003). When the soil P test was changed in Denmark in 1987, it was from a method extracting large proportions of total inorganic soil P (Bondorff, 1950, Rubæk and Sibbesen, 2000) to the Olsen soil P test which extracts much less P and which had shown much stronger relations to plant P uptake than other methods (Sibbesen, 1983). It has since become clear that on Danish soils with temporary, high water tables (lowland soils) the Olsen P method is not recommendable. Such soils typically support good yields in spite of low Olsen P values and have been shown to sometimes contain very large amounts of total P (Knudsen et al., 2011).

Other results indicate that Olsen P for Danish conditions would benefit from a recalibration towards somewhat higher threshold values for optimal crop production on some sandy soils (Rubæk, 1999; Rubæk and Sibbesen, 2000) where a threshold of around 30 mg P/kg soil would appear more appropriate instead of 20 mg P/kg on one soil and even higher on another sandy soil. The same study confirms that the threshold commonly used in Denmark of 20 mg P/kg is reasonable on the more clayey Danish soils.

The DGT-method and other alternatives

Several alternative methods have been developed, which more precisely estimate the plant-availability of soil P compared with the Olsen P method and other routine soil P test methods. One group of such methods is the isotopic labelling techniques (e.g. Morel et al., 2002 and Schneider and Morel , 2000). Such techniques are not suitable for routine purposes, but are frequently used in research. Another group is the so-called infinite sink extraction methods, where either anion-exchange resin beads or membranes or ion-oxide-impregnated filter papers with large binding capacities for P are used as extractants in a mixture of soil and water (Sibbesen et al., 1983; Menon et al., 1990; Hosseinpur and Sinegani, 2009). In Brazil an ion exchange resin method has been used for routine purposes for many years (van Raij, 1998) and a resin method is also listed as an alternative method in the British fertiliser recommendation manual (MAFF, 1986). In spite of their qualities, this type of infinite sink methods has not been used for routine purposes, with the few mentioned exceptions. This reluctance to change method is most probably related to the comprehensive work needed to validate a new method under practical conditions and to the work required to implement new methods in routine soil laboratories.

The Diffusive Gradient in Thin film (DGT) technique was developed in the 1990s at Lancaster University, initially as a tool to measure metal pollution in aquatic environments. Around the millennium it moved into soil pollution and in the last 5-8 years it has been recognised as a powerful tool in agronomic soil testing. The DGT unit consists of a small plastic holder, around 4 cm in diameter. It holds together a binding gel containing Fe-oxide, a diffusive gel and a protective filter (Fig. 1). The DGT unit is deployed in a water-saturated soil and left for normally 24 hours. Nutrients, in this case P, will diffuse from the soil solution towards the binding gel where it is adsorbed. Hence, the concentration at the binding gel will remain zero so the amount of P adsorbed over the 24 hours is determined by the diffusion gradient and the re-supply from the solid soil phase. This relies on the same theoretical principles as P uptake by a plant root where a P depletion zone in the rhizosphere is driving P diffusion towards the root and induces resupply from the solid phase. The advantages of the method over common extraction techniques are clearly demonstrated and described in Degryse et al. (2009), Mason et al. (2010) and Mundus et al., (2013).

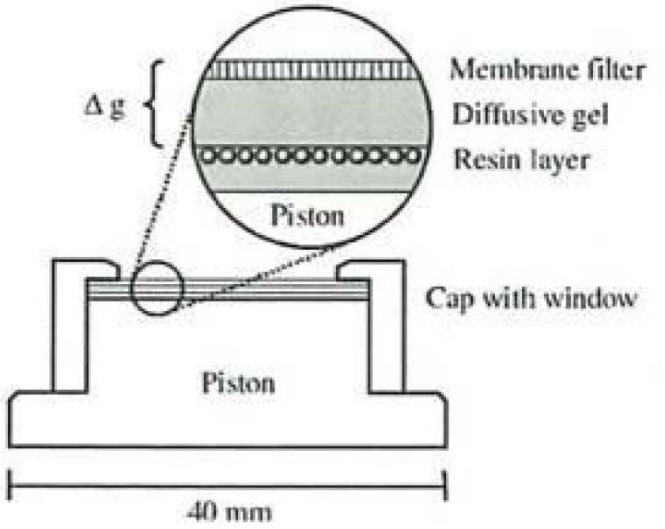


Figure 1. Cross section of the DGT device showing a membrane filter on top of a diffusive gel with a known diffusion coefficient. Below these are the resin gel with resin which will bind the nutrients (Dahlqvist et al., 2002).

The first work using DGT to predict plant-availability in agricultural soils was done by Menzies et al. (2005). They showed that P uptake by tomato was well predicted using DGT while the Colwell-P extraction, a modification of the Olsen-P method used in Australia, provided a poor prediction. In a later study, Mason et al. (2010) tested the DGT technique on a range of Australian field trials and found that it accurately predicted yield of wheat while extraction methods gave poor correlations. Since then, several studies have shown that this method estimates plant P-availability in soil well and is superior to most other routine soil P tests including the Olsen P method (Six et al., 2012; Six et al., 2014; Tandy et al., 2011). Mason et al. (2013) showed by isotopic labelling that DGT measures P from the same pools that supply growing plants, while other methods including the Olsen P also to varying degrees extract P from other pools. In Australia the method has recently been tested and compared to routine soil P tests including Olsen P on a comprehensive dataset including 164 field experiments with P fertilisation from 1968 and until today (Speirs et al., 2013), which again showed the potential of the DGT method. In this study, the Olsen P (and Colwell P) also performed reasonably. It is also worth noticing, that "outliers" in this study seem to be more pronounced for the Olsen P than for the DGT-P. This is most probably because the Olsen P method is much more soil type specific than the DGT method. Speirs et al. (2013) to some extent addresses this in their discussion where they demonstrate that carbonate-rich soils in their study seem to behave differently from other soils with the Colwell test. It should also be noted that soil measurements in Speirs et al. (2013) are carried out on stored soil samples which may have influenced the results. Speirs et al. (2013) conclude that the DGT method has potential and should be further developed, and that it is vital to do further work comparing the methods on underrepresented soil types and soil P levels. Their study furthermore clearly demonstrates the large variations in results which are common and unavoidable in field experiments.

A Danish GUDP project is presently testing the DGT method as a soil P test and comparing it to routine soil P tests in pot trials as well as in field experiments across Scandinavia. Nine field esperiments were

carried out in 2013 and 15 in 2014. Sixteen of the trials were carried out in Denmark and the rest in Sweden, Norway and Finland. Each trial was cropped with spring barley with or without P fertiliser in the amount of 30 kg ha^{-1}. Only four out of the 24 trials responded to P fertiliser by producing higher yields, but 16 of the 24 trials showed transient deficiency where there was a clear fertiliser effect early in the season. For the DGT measurements four soils were above the critical threshold established in Australia. None of these four soils responded to P fertiliser addition, neither in the early season, nor at maturity. It was concluded that above the Australian threshold no deficiency problems arose. However, below the threshold transient deficiencies were often found and whether or not these manifested into effects on yield depended on the actual growing conditions at each site (soil temperature and water content). For the Olsen-P method, it was found that two of the responsive soils had low Olsen-P values of 1.3 and 2.0, but the other two had values above the threshold (4.1 and 5.3 mg P/100 g soil), where responses was not be expected. It was therefore concluded that the Olsen-P method in these soils did not reflect plant P availability well.

Validity of agronomic soil P tests for evaluation of risk of P losses

Methods more specifically designed to address environmental purposes exist. Such methods typically aim at determining the P sorption capacity, the degree of P saturation and the P that can be released to water (Pw) (Schoumans, 2015; van der Zee et al., 1990; Sissingh, 1971). These methods are generally more time-consuming in the lab and there is only limited data available at the scale and the geographical scale needed for land managers. Therefore agronomic soil P tests are frequently used in tools predicting the risk of P losses to surface waters, because they are cheap and because there is already a large knowledge base with field-scale observations (e.g. Heckrath et al., 2009). Typically the test already used in a country/region is used as a proxy for all kinds of P losses and for accumulation of fertiliser P in soils (legacy P). However, it is probably reasonable to assume that the agronomic soil P test methods, which extract a relatively large proportion of total inorganic P will be better predictors of the legacy P and for P lost in particulate forms (because they measure a large proportion of the soil P) than methods only extracting a minor fraction of the total soil P. In contrast, the methods that only measure a minor part of the soil P and the less strongly bound P may generally perform better when it comes to prediction of dissolved P losses. It is therefore very probable, but remains to be further elucidated, that e.g. the DGT method can give reasonable information on the risk of losing especially dissolved inorganic P from the plough layer, while the relation to the loss of P in particulate forms is expected to be weaker. Olsen P has been shown to predict P losses through leaching reasonably well (Heckrath et al., 1995, Glæsner et al., 2011; Kjærgaard et al., 2010) and it is also used as a proxy for soil total P in relation to particulate P losses in the Danish P index (Heckrath et al., 2009). The acid extraction methods, which extract relative large proportions of soil total P will most probably do quite well as proxies for soil total P, while their relation to dissolved P losses are expected to be weaker.

4. Performance of selected soil P tests in proficiency test programmes

Gitte H. Rubæk, Department of Agroecology, Aarhus University

This report focuses on the major soil P tests used for routine soil analyses in Denmark and neighbouring countries, i.e. the Olsen P method and its Danish version "Ptallet", P-CAL, P-AL, Pw (water-extractable P). Information on the proficiency test programmes used for these methods in Denmark and bordering countries and listed in table 3 along with key information on these programmes.

While gathering this information, it became clear that there are important differences among the proficiency test programmes: Some programs run more frequently at many labs with few soil samples at each test round, while others use more soils in each round, but are less frequent and/or include fewer labs. Also the way the results are reported differs between the programs. In some countries participation and passing of certain criteria in the test program are required by authorities or the agricultural agencies. The SEGES program differs from the others by including the largest number of soils per test round and very few labs. When few labs are taking part, the more traditional inter-lab comparison becomes weak. To compensate for this, the SEGES program also includes a comparison to results obtained on the same reference soils in earlier test rounds, which is unique, and this program is the only one which can analyse consistency of test results over time and therefore allows examination of other aspects of the analytical performance and certainty in soil laboratories (see, for example, Chapter 5).

Reference soils for used in proficiency test programme arranged by Seges.

Table 4.1 Proficiency test programs for routine soil tests in Denmark and neighbouring countries.

Method	Organisation in charge	No. of labs in the most recent round	No. of soils analysed in each round	Frequency of test rounds per year	Comments
Olsen P (NF ISO 12263)	WEPAL P.O. Box 8005 NL-6700 EC Wageningen the Netherlands Info.Wepal@wur.nl http://www.wepal.nl/	38	4	4	The number of labs participating varies a lot between rounds as it is a voluntary test for most of the participating labs and they therefore sometimes skip a test round
	BIPEA CAP 18 - 189 rue d'Aubervilliers F-75018 PARIS FRANCE contact@bipea.org http://www.bipea.org/	28	1	10	
Olsen P the British version (MAFF, 1986))	WEPAL	12	4	4	The number of participating labs is relatively constant, because participation is required by the Department for Environment, Food and Rural Affairs, UK
Ptallet (DK)	SEGES Agro Food Park 15, 8200 Aarhus N http://www.seges.dk/Seges.htm	3	10	3	Compares results to averages for standard soils obtained in previous years
P-CAL (Germany)	BIOANALYTIK Weihenstephan (LÜRV-A BODEN) Zentralinstitut für Ernährung und Lebensmittelforschung, ZIEL TECHNISCHE UNIVERSITÄT MÜNCHEN Alte Akademie 10 85354 Freising, Germany	104	2	1	

	Att: Dr. Ludwig Nätscher http://www.bioanalytik-weihenstephan.de				
	Bayerische Landesanstalt für Landwirtschaft Abteilung AQU http://www.lfl.bayern.de/zentrale_analytik/030223/index.php	13	2	1	
P-AL	Ringtesten & Erkenningen VITO NV | Boeretang 200 | 2400 Mol, Belgium Att: Siegfried Hofman	7	1		Some Swedish soil labs also use this test program
	BIPEA	4	1	10	Possible since September 2014
	Norway, Landbruksdepartementet, who has appointedProfessor Tore Krogstad, Norwegian University of Life Sciences to be in charge of a national proficiency test programme)	7	6	1	Lab-performance has to be approved (by a ministry agency) before results can be used in mandatory fertiliser planning tools Some Swedish soil labs also use this test program
	WEPAL	8	4	4	
Pw	WEPAL	2	4	4	
DL-P	BIOANALYTIK Weihenstephan (LÛRV -A Boden)	37	2	1	Together with CAL

How the Olsen soil P test and its modifications differ

A thorough method description is an indispensable part of qualified analytical work and when comparing performance between labs, it is essential that the laboratories refer to the same well-defined analytical protocol. The original soil P test method using a 0.5 M sodium bicarbonate solution at pH 8.5 for extraction was first published by Olsen et al. in 1954. It is now used as a routine soil test in many countries world-wide, frequently under the nickname “Olsen P”. Apart from the original publication, several laboratory protocols describing how to carry out variants of this analysis are available (e.g. ISO 11263,

1994, MAFF, 1986, Plantedirektoratet, 1994; Sparks, 1996). Key elements of the original description and some important international protocols of the existing "official" Danish variant and of the updated Danish variant described in Appendix A are listed in table 4.

Table 4.2 Analytical details of the original Olsen soil P test and four widely used protocols describing this test

Analytical detail	**Olsen et al., 1954**	**Methods of Soil analysis (Sparks,1996)**	**ISO 11263 (1994)**	**Plante-direktoratet (1994)**	**MAFF (1986)**	**Updtated Preliminary method. Appendix A.**
Amount of soil	5 g	2.0 g	5.00 g dried at maximum 40°C	5 g dried at 50-60 °C	5 ml	1.00 to 5.00 g dried at max. 50-60° C.
Extracting solution	0.5 M NaHCO3 adjusted to pH 8.5.	0.5 M NaHCO3 adjusted to pH 8.5.	0.5 M NaHCO3 adjusted to pH 8.5.	0.5 M NaHCO3 adjusted to pH 8.5.	0.5 M NaHCO3 adjusted to pH 8.5.	0.5 M NaHCO3 adjusted to pH 8.5.
Amount of extracting solution	100 ml	40 mL	100 ml	100 ml	100 ml	Ensure soil to solution ratio of 1 to 20
Volume (and type) of flask	Not specified	125 ml Erlemeyer flasks	250 ml	250 ml	175 ml	Ensure soil weight to bottle volume of 1:50
Shaking time	30 minutes	30 minutes	30 minutes	30 minutes	30 minutes	30 minutes
Temperature	Not specified	Not specified	20 +/- 1 °C	22 °C +/- 1 °C	20 °C +/- 1 °C	20 °C +/- 1 °C. Thoughout the entire procedure
Shaking method/speed /distances	Only stated that the rate of shaking should be constant	Not specified	Shake to prevent settling of soil, otherwise not specified	Rotating, end over end, shaking machine	(Grifffin bottle shaker), 275 strokes per minute, length of travel 25 mm	Rotating end over end, speed 20 rounds pr minute
Filtration	Filtration through Whatman no. 40 or other suitable paper	0.45 µm membrane filter or Whatman no.42 filter paper	Filtration through phosphorus free paper	Filtration through phosphate free filter paper	Immediate filtration Whatman nr 2 filter paper	Separation by centrifugation at 1800 g for 5 minutes at 20°C or filtration initiated within 15 minutes after end of extraction
Handling of colour/organic material	One teaspoon of activated carbon black	Half a spoonful of activated carbon to each extraction flask	1 g of activated carbon added to each extraction flask	Polyacrylamide added to extraction solution	Polyacrylamide added to extraction solution	Polyacrylamide added to extraction solution
P detection	Colorimetric using the Dickman and Bray method ($SnCl_2$/molybdate reagent)	Colorimetric determination ascorbic acid/ammonium molybate reagent	Colorimetric determination using a sulfo-molybdate reagent	Colorimetric determina-tion, ascorbic acid/ammo-nium molybate reagent	Colorimetric determina-tion, ascorbic acid/ammonium molybate reagent	Colorimetric determination, ascorbic acid/ammonium molybate reagent

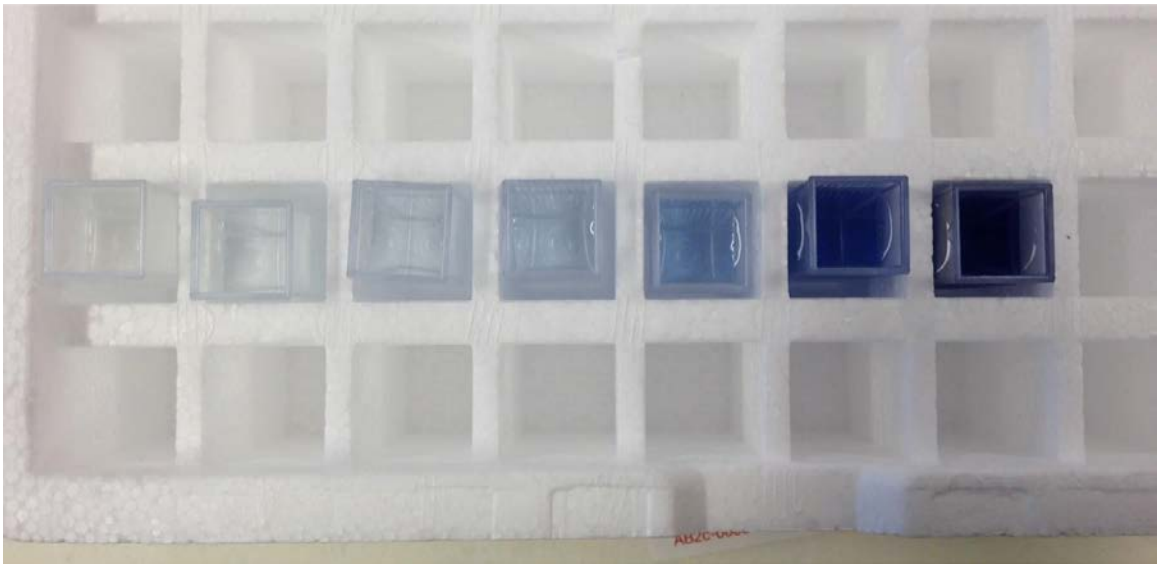
The five protocols for bicarbonate extraction of soil P listed in table 3 are identical when it comes to the extracting solution and duration of extraction, but for details in the protocols there are deviations: The original method does not specify extraction temperature, size of extraction flasks and shaking method and speed, while this is addressed in several of the newer protocols. All versions are based on soil weight,

except for the British which specifies a soil volume, and use identical soil-to-solution ratios. The handling of dissolved coloured organic substances in the extract are done by adding activated carbon in three methods, while the Danish and the British version relies on the addition of polyacrylamide (Banderis et al., 1976). Phosphorus concentrations in the extract are in all methods determined by colorimetry which measures the colour intensity of a phosphorus-molybdate complex, but the exact method differs somewhat, especially when it comes to choice of reducing agent for the colour development.

Even though the identified differences among the protocols are minor and the methods in principle are identical, there will most probably be small and systematic differences in the obtained results from each of the methods, especially if operating with large differences in temperature and shaking intensity. An updated method description should therefore specify such details precisely. Additionally there are a number of difficulties and pitfalls which are often faced when analysing soil for bicarbonate-extractable P. Many of these issues are not dealt with in the old protocols listed in table 3. Below is a list of key issues that needs to be properly addressed in an updated analytical protocol and in all labs performing this analysis:

- Soil pretreatment, especially temperature for drying is most probably important for the result. Drying procedure should therefore be similar for all laboratories.
- Since shaking intensity and method affect the results, these should be kept constant and in line with the protocol used.
- It is important to keep the extraction temperature within the designated limits throughout the lab work. I.e. the extracting solution should have the designated temperature before the extraction starts.
- Extraction time has to be precise. I.e. separation of soil and extractant should take place immediately after the 30 minutes extraction time.
- The method used for separting soil and solute after extraction may influence the result. It is therefore important that separation procedure is defined and properly described in the analytical protocol.
- The gaseous CO_2 which develops after adding acid to the extract should be released carefully. Otherwise small air bubbles in the extract can form during colour development and disturb the measurement.
- Foaming is often experienced during acidification. This should be handled without losing extract.
- The P detection method in the extract is important, because it is the molybdate reactive P in the extract which makes up the Olsen P not the total P in the extract. Therefore detection methods like ICP may lead to overestimation of Olsen P because it measures the total P.
- In automated flow systems for acidification, degasification and subsequent addition of the reagents needed for colour development it is important to assure full removal of analyte in the flow system between samples (Maria Kreimeyer, Agrolab, personal communication).

In Appendix I, an updated protocol of the Danish version of the Olsen P method is presented. In this protocol the description of the procedure is modernised and more practical details regarding the extraction (temperature, shaking intensity, etc.) are specified in order to eliminate small but persistent differences that otherwise may occur. We have also listed the difficulties/pitfalls that should be taken into account when setting up the method in a laboratory and we have written the description in English to make it easier to implement it in labs outside Denmark, too.



A standard curve of the blue color developed by the ascorbic acid/ammonium molybate reagent for spectrophotometric determination of the P concentration in the bicarbonate extracts.

5. Reduction of systematic variation between laboratories and time of analysis using correction of results

Kristian Kristensen and Gitte H. Rubæk - Department of Agroecology, Aarhus University

Introduction

It has previously been documented that the Ptal measurements on subsamples of standard soils vary significantly and systematically between the laboratories and the time at which samples are submitted for analysis. I.e. the results obtained for the same soil sample depend on which laboratory you choose and the time you submit your sample (Rubæk et al, 2011; Videncentret for Landbrug, Planteinfo 2014). A consequence of this is that the uncertainty with an average of several measurements of one or more soils, e.g. a set of standard soils, will be larger when the samples are submitted to different laboratories and/or at different times than when they are submitted to the same laboratory at the same time and then analysed in the same run (Appendices 4, 5 and 6 of Rubæk and Sørensen, 2011). In other words: The difference needed between two analytical results to make the difference statistically significant may become unreasonably large when lab and/or time of analysis differ, and this hampers our ability to detect, for example, when the soil P status has declined or increased significantly due to too little or too much P input over some years.

Inclusion of one or more standard soils in each run of a soil analysis is a standard procedure in most laboratories. The inclusion of standard soils with a known test value allows the checking for analytical problems in each run. Typically a range for the test result is defined for the standard soils and if the result for the standard soil falls within this range, the run is accepted; if not, all analyses have to be repeated. This is a common procedure in most analytical work.

Inclusion of standard soils with well-known “true” test results in each test run also allows the use of these samples for corrections of minor deviations among test runs within the lab if that is necessary. This can be further extended if the same set of standard soils and the same “true values” for these are used at different laboratories, where systematic variations related to both time of analysis and laboratory can then be corrected. In the best of both worlds such corrections are not necessary because by careful work in the laboratory and detailed and precise protocols for the methods, it should be possible to minimise such systematic deviations in the test results. But in some cases, like for the Danish Ptal, it has so far not been possible to reduce this systematic error sufficiently. In the following we therefore examine different ways to carry out corrections on the Ptal analyses, with the objective to identify the most suitable correction method in case the problems with systematic variation on the Ptal analysis persist even with an update of the method description.

The dataset and the tested correction methods

Rubæk et al. (2011) already showed that is possible to reduce this unwanted systematic bias by adjusting the actual measurements according to simultaneous measurements of well-known standard soils analysed in the same batch, but that study was on a limited dataset which only allowed examination of very simple correction procedures for a few labs and years. For this report we have therefore expanded the investigation on the Ptal measurements to include three strategies on data obtained in the ring tests carried out by the "Knowledge Centre for Agriculture"/SEGES between 2008 and 2013. During these five years subsamples of 10 different standard soils were sent for analysis to three commercial laboratories three times in the period between October and February for each of the seasons 2008/2009, 2009/2010, 2010/2011, 2011/2012 and 2012/2013. In 2008/2009 only a subset of the soils was submitted and we therefore omitted data from this season in the present analysis. The mean Pt values together with their minimums and maximums are shown for each soil in table 5.1. A graphical presentation of the data is shown in figure 5.1.

Table 5.1 Mean, minimum, and maximum Pt values for the 10 soils in the ring-test. The last column shows the role of each soil in the investigation.

Soil identification	Mean	Minimum	Maximum	Used as
Dansk Standard 1	3.4	1.8	4.5	*Submitted*
Foulum 99 Have	8.6	6.4	10.4	Standard 1
Foulum Hvede	5.8	4.6	8.1	*Submitted*
Jens K Mark	3.9	2.6	5.2	*Submitted*
Liselund	2.4	1.9	3.4	Standard 2
Lolland 2000	6.2	5.1	7.6	Standard 3
Roum 2 1996.08	3.2	2.5	4.3	*Submitted*
Roum 3	4.9	3.7	6.2	*Submitted*
Troestrup 1995	4.2	3.4	5.1	Standard 4
Troestrup 1996	4.2	1.8	5.2	*Submitted*

Figure 5.1 Plot of the resulting Pt values for each laboratory and for each of 10 soils submitted 12 times during 4 years.

For testing the three methods of correction we considered four of the 10 soils as standard soils, while the remaining six soils were considered as "normal" soils *submitted* for analyses. The four soils used as standard soils were chosen to cover the range of Pt values in the soils to be adjusted.

We have tested and evaluated three different correction approaches:

a) Adjust all results from the run by the difference between the actual values of the standard soils and the "true" values of the standard soils (here called additive adjustment).

b) Adjust all results from the run by the quotient between the actual values of the standard soils and the "true" values of the standard soils (here called multiplicative adjustment).

c) Adjust all results from the run by using a "standard curve" obtained by regression of the actual values of the standard soils against the "true" values of the standard soils (here called calibration).

The true value of the standard soils would most often be based on the mean of many analyses of each standard soil carried out over a reasonable time and/or at relevant laboratories.

It should be noted that no adjustment can be expected to be exact as an adjustment also introduces some uncertainty. Therefore, adjustments that introduce more uncertainty than they remove will not be beneficial.

For methods a) and b), each of the submitted soils was adjusted using each of the standard soils. For method c) each of the submitted soils was adjusted using a calibration curve based on the four standard soils. In all cases, the mean of each standard soil was used as the "true" value. The effect of adjusting was then evaluated by comparing the standard error on the difference between two samples for different simulated conditions (see tables 5.3 and 5.4). The standard error on the difference was calculated from variance components which were estimated using two different models for the submitted soils:

- A mixed model used for data from each laboratory and each adjustment method where year, time within year and residual were included as random effects
- A mixed model used for all data and each adjustment method, wherelaboratory, year, laboratory by year, time within laboratory and time and residual were included as random effects.

The effect of the *submitted* soils was included as a fixed effect in both analyses.

For further details on the analyses and the adjustment methods, see Appendix 3.

Results

The standard errors on the difference between two soils are shown in tables 5.3 and 5.4. For commercial laboratory 1 the standard error on the difference between two samples submitted in different years was reduced from 0.97 to 0.61 if an additive adjustment using standard soil 4 (Troestrup 1995 with an average Pt of 4.2) was applied, whereas the standard error was only reduced to 0.83 if the additive adjustment using standard soil 1 (Foulum 99 Have with an average Pt of 8.6) was applied. For commercial laboratory 3 none of the applied methods reduced the standard error on the difference between two samples, and in fact some adjustment methods increased the standard errors on the differences between two soils. For commercial laboratory 2 the size of the reduction/increase of the standard error was somewhere between that of commercial laboratories 1 and 3. The reason for the difference between laboratories is most likely related to the origin of variance at the lab: For commercial laboratory 1 a relatively high part of the variation (63%) occurred between time and years, whereas for commercial laboratory 3 only a relative

small part of the variation (22%) occurred between time and years. This means that the adjustment using a single standard soil adds more noise than is removed by the adjustment, if only a small part of the total variance occurs between time and year. In addition, in a few cases there was a tendency for the relation between the "true" values and actual recorded values to be non-linear for commercial laboratory 3 (figure 5.2).

Table 5.2 Absolute and relative variance components for each laboratory

	Variance components (Pt^2)				**Relative variance components (%)**		
Laboratory	**Years**	**Time:years**	**Residual**	**Total**	**Years**	**Time:years**	**Residual**
Commercial 1	0.071	0.228	0.175	0.474	15	48	37
Commercial 2	0.032	0.021	0.066	0.119	27	18	55
Commercial 3	0.000	0.028	0.099	0.127	0	22	78

From table 5.3 it can also be seen that additive adjustment was better than multiplicative adjustment if the standard soil had a relatively low Pt (i.e. standard soil 2), while an multiplicative adjustment was better than additive adjustment if the standard soil had a relatively high Pt (i.e. standard soil 1 and 3). For standard soil 4, with a mean Pt value of 4.2, the additive and multiplicative adjustment had approximately the same effect. In addition, using standard soil 4 for the adjustment reduced the standard error on the difference for commercial laboratories 1 and 2 by the largest amount, and only increased the standard error on the difference by a small value for commercial laboratory 3.

The calibration method had approximately the same effect as adjustment using standard soil 4, but also for this method there are both benefits and drawbacks: The main drawback is that four standard soils are required instead of just one; the benefit is that a correction based on a calibration curve offers more scope for evaluating the quality of the run (and the adjustment) and thus also for detecting and handling dubious results for certain standard soils. Such matters can be evaluated by just looking at the calibration curve (figure 5.3) or a measure for goodness of fit, e.g. by using the coefficient of correlation. For the 36 curves used here, the coefficient of correlation in this dataset varied between 0.979 and 0.999.

Table 5.3 Standard error on difference between two measurements at each laboratory for raw data and adjusted values using additive or relative adjustment based on one standard soil or a calibration based on four standard soils. (The unit for Ptallet/Olsen P is "mg P/100 g soil".)

Laboratory	**Submission time**	**Recorded values**	**Adjusted values for each standard soil[a]**								**Calibrated values**
			a1	**b1**	**a2**	**b2**	**a3**	**b3**	**a4**	**b4**	
Commercial 1	Different year	0.97	.83	.67	.70	.73	.82	.77	.61	.62	0.60
	Same year	0.90	.76	.67	.70	.73	.82	.77	.61	.62	0.60
	Same time	0.59	.59	.56	.59	.53	.59	.54	.59	.58	0.59
Commercial 2	Different year	0.49	.70	.45	.46	.57	.53	.45	.44	.45	0.41
	Same year	0.42	.70	.45	.38	.53	.53	.44	.42	.43	0.39
	Same time	0.36	.36	.37	.36	.41	.36	.37	.36	.38	0.36
Commercial 3	Different year	0.50	1.5	.96	.56	.75	.77	.69	.52	.54	0.60
	Same year	0.50	1.3	.81	.56	.75	.77	.69	.52	.54	0.58
	Same time	0.45	.45	.46	.45	.48	.45	.47	.45	.47	0.49
Mean value of standard soil			8.59		2.41		6.25		4.16		

a) Standard soils numbered 1 to 4 (see table 5.1) and adjustment method (a=additive adjustment, b=multiplicative adjustment)

On average, similar reductions on the standard error of differences were obtained by the two correction methods (adjustment by use of standard soil 4 and calibration): The average standard error of difference when the samples were analysed at different laboratories was reduced by 22% to 25% (two top lines of table 5.4) and by 19% to 23% if the samples were sent to the same commercial laboratory at different times (or years) (see lines 3 and 4 in table 5.4).

From both tables 5.3 and 5.4, it can be seen that the effect of adjustment had only a very limited effect if the samples were sent to the same laboratory at the same time – in fact it can be shown theoretically that an additive adjustment cannot change the uncertainty in such a case.

Table 5.4 Average standard error on difference between two measurements at each Laboratory for raw data and adjusted values using a calibration based on four standard soils. (Unit for Ptallet/Olsen P is “mg P/100 g soil”).

Submission time	**Recorded values**	**Adjusted values for each standard soil[a]**								**Calibrated values**
		a1	b1	a2	b2	a3	b3	a4	b4	
Different lab different year	0.77	1.1	.78	.59	.77	.82	.71	.57	.59	.56
Different lab same year	0.73	1.1	.78	.58	.77	.81	.69	.57	.59	.55
Same lab different year	0.70	1.1	.73	.59	.70	.73	.66	.54	.55	.56
Same lab different time	0.66	1.0	.69	.59	.70	.73	.66	.54	.55	.54
Same lab same time	0.50	.50	.49	.50	.50	.50	.48	.50	.50	.51
Mean value of reference soil		8.59		2.41		6.25		4.16		

[a)] Standard soils numbered 1 to 4 (see table 1) and adjustment method (a=additive adjustment, b=multiplicative adjustment)

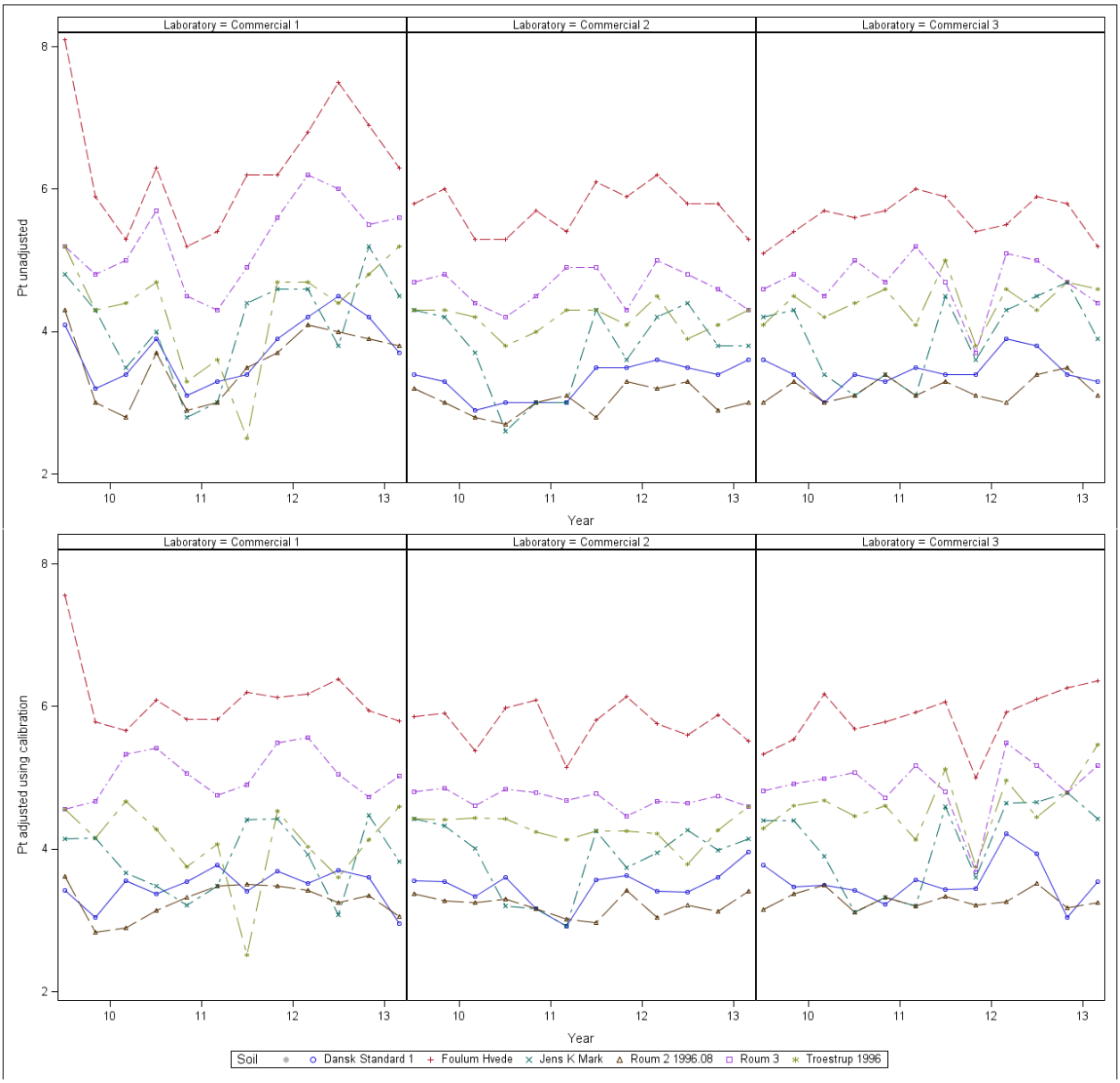


Figure 5.2 Visual comparison of Pt before and after adjustment using the calibration method (only the six soils used as *submitted* soils are shown)

The effect of calibration on the analyses used in this study is show in figure 5.2. The Pt values are clearly more equal across laboratories after calibration. Especially for Commercial Laboratory 1 the values are clearly more in line with those at the other two laboratories, and also the variation over time is clearly reduced for this laboratory. The very high value for the "Foulum Hede" soil at the first sampling time remains high, and could be due to an erroneous measurement of exactly this sample at this time. For the other two laboratories the calibration only modified the variation over time moderately.

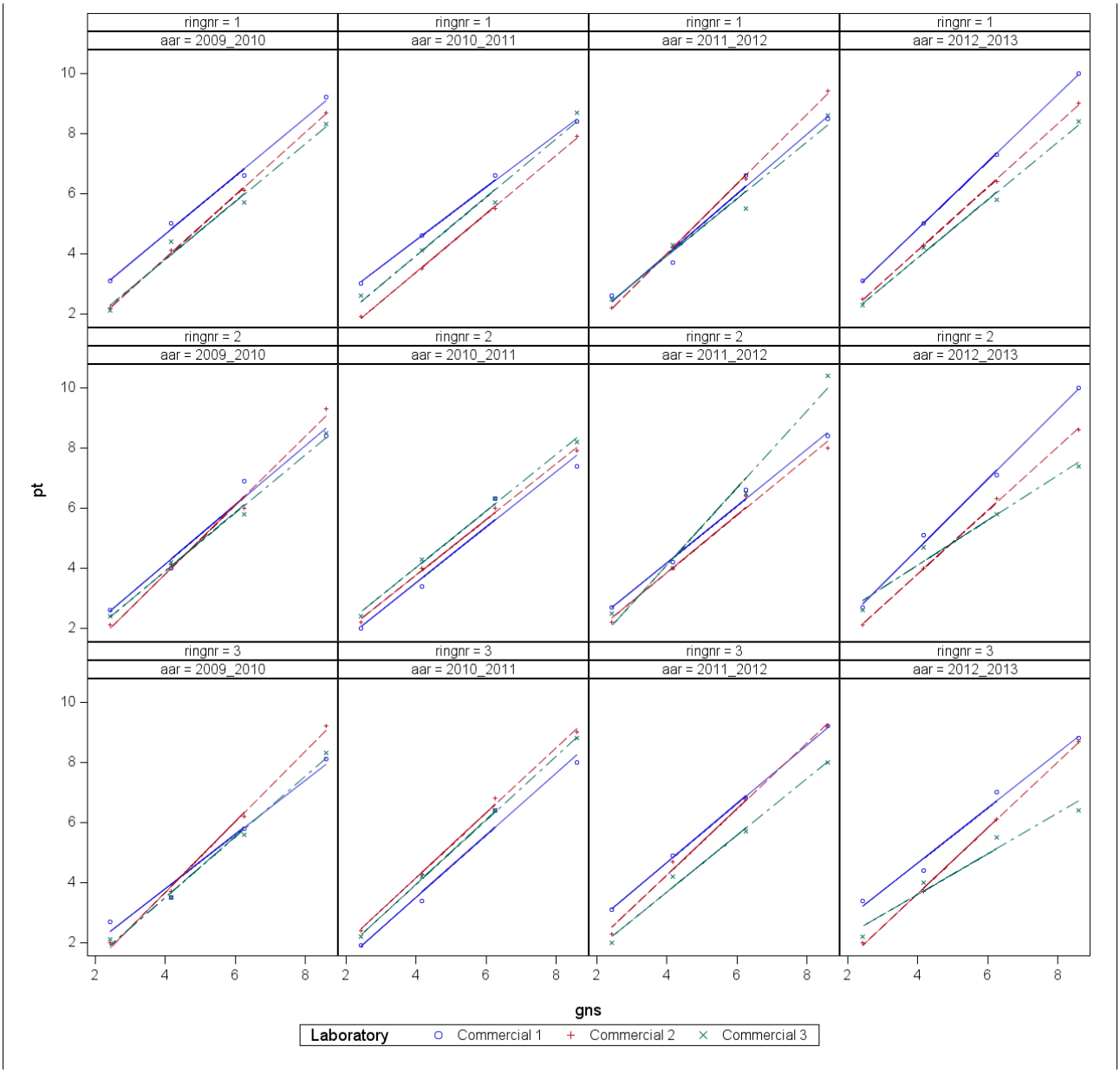


Figure 5.3 Applied calibration curves for each of the three commercial laboratories, years and ring tests (averages for each soil are used as "true" values on the horizontal axis)

Similar analyses of older datasets from the research laboratories in Foulum (Rubæk et al., 2011 and Appendix 6 in Rubæk and Sørensen, 2011) are in accordance with the results presented here. In the former study only two standard soils were available in the dataset ("Danish standard 1" and "Liselund"), which, in contrast, included numerous analytical runs (205 runs at "Centrallaboratoriet" during 1999 and 2003 and

77 runs at “Institut for Jordbrugsproduktion” during 2004 and 2008). Using “Danish standard 1” as the standard soil and “Liselund” as the *Submitted* soil and vice versa showed that the standard error on average could be reduced by 29% at “Centrallaboratoriet” and by 19% at “Institut for Jordbrugsproduktion”. The calibration method could not be evaluated due to the limited number of soils included. For further details on this study see Appendix 6 in Rubæk and Sørensen (2011).

How much certainty can be gained on the difference between the means of several samples?

A core question for whether to implement a correction procedure on the absolute measurements or not is how much certainty can be gained in practice on the average of, for example, two sets of 10 or 40 samples from e.g. one farmer’s 10 or 40 fields that are sent for analyses at different times. This can also be formulated this way: How much correction with calibration to reduce the difference between average test results of 10 or 40 soil samples analysed at two randomly chosen laboratories would be required to make the difference statistically significant at the 95% confidence interval? Unfortunately, our data do not allow us to estimate this properly for the suggested calibration method. In Appendix IV we have made an estimation based on additive correction by an average of four standard soils. Here the standard error on the difference was on average reduced by 20-25% if the two set samples were submitted to different laboratories in different years. I.e. the averages of 10 samples analysed at different years should differ by more than 1.22 mg P/100 soil for the difference to be significant at the 95% level before correction and by more than 0.96 mg P/100 g soil after correction.

Even though we cannot estimate whether the reduction would be the same for the calibration method, or bigger or smaller, we have reasons to believe that it will not deviate much from the above-mentioned calculation if a correction based on the calibration method is used (see Appendix 4).

Stability of standard soils

Rubæk et al. (2011) also showed that Pt values for a standard soil decrease with time, especially in the early years, and increase with temperature at “Centrallaboratoriet”. At “Institut for Jordbrugsproduktion” there were no such significant differences but the same tendencies of much smaller magnitudes were seen. The decrease of Pt at “Centrallaboratoriet” was largest in the beginning of the period (1999 to 2001) when the soils were more recently sampled and dried. This indicates that the soils have to be stored for some time under dry and constant climate conditions before they are stable and can be used as standard soils. This has also been observed in other countries (Dr L. Blake and Dr M.M.A Blake-Kalff, Hill Court Farm Research Ltd, personal communication). Also Castro and Torrent (1993) and Bramley et al. (1992) have shown that dry and constant storage conditions are important for storing soil samples.

Conclusions

By using the same "true" values of stable standard soils for correction in all laboratories, the three suggested adjustment methods could level out some "systematic" variations between laboratories and time of analyses of the reported results: However, it was also clear that correction increased the overall uncertainty slightly at the laboratory having a small "systematic" variation for the soils over time compared with the size of their analytical error (i.e. having a relatively large residual error). It is therefore important that laboratories carry out their analyses with small residual errors, which agrees well with overall aims of good analytical work. It is also important to note that correction is only worthwhile if there is a high risk of "systematic" variation. Detection of "systematic" variation requires a proficiency testing programme where identical soil samples are repeated year after year, much like the system offered by SEGES (Videncentret for Landbrug, 2014), and to our knowledge no other soil P test has been scrutinised as thoroughly for its robustness over time and between labs as is the case for the Danish Olsen P test.

Adjustments using standard soil 4 (both multiplicative and additive) or the calibration method furthermore reduced the standard error on the difference between two samples submitted to different laboratories and/or at different times to approximately to the same extent. For some laboratories these methods reduced the standard error considerably, whereas at other laboratories the reduction was small or slightly negative.

The calibration method has one important advantage over the two other adjustment methods: It allows a check of the validity of the calibration curve (e.g. by looking at the graph and the coefficient of correlation), making a more solid foundation for the decision on whether to discard a whole analytical run.
To avoid systematic variations between laboratory used and time of analyses, which has repeatedly been observed for the Danish Ptal, we therefore recommend correction of results of the Ptal analyses to be calibrated against 4 standard soils covering the range of Olsen P values from *ca.* 1 to *ca.* 8. The soils used should be identical for all laboratories in order to assure a stable overall level of the Pt values used by farmers, consultancies, authorities and researchers.

6. References

Banderis, A, Barter, DD & Henderson, K, 1976. The use of polyacrylamide to replace carbon in the determination of ”Olsen’s” extractable phosphate in soil. Journal of soil Science 27:71-74.

Beegle, D, 2005. Assessing soil phosphorus for crop production by soil testing. In: (Eds: JT. Sims & A. Sharpley). Phosphorus: Agriculture and the Environment. Page 123-143.

Bondorff, KA, 1950. Studier over jordens fosforsyreindhold. VI. Jordfosforsyrens opløselighed i fortyndes svovlsyre. Tidsskrift for Planteavl. 53:336-342.

Bramley, RGW, Barrow, NJ & Shaw, TC, 1992. The reaction between phosphate and dry soil. I. The effect of time, temperature and dryness. Journal of Soil Science, 43:749-758.

Castro, B &Torrent, J, 1993. Phosphate availability in soils at water activities below one. Coommun. Soil Sci. Plant Anal. 24:2085-2092.

Dahlqvist, R, Zhang, H, Ingri, J & Davison, W, 2002. Performance of the diffusive gradients in thin films technique for measuring Ca and Mg in freshwater. ANALYTICA CHIMICA ACTA, 460:247-256 .

Degryse, F, Smolders, E, Zhang, H & Davison, W, 2009. Predicting availability of mineral elements to plants with the DGT technique: a review of experimental data and interpretation by modelling. ENVIRONMENTAL CHEMISTRY, 6:198-218.

Egner, E & Reihm, H, 1955. Die Dobbellaktatmethode. In: Thon, R., Hermann, R. & Knikemann, E. (eds.) Die Untersuchung von Boden Verbandes Deutscher Landwirtschaaftlicher-Untersuchungs- un Forschungsanstalten, Methodenbuch I. Nuemann Verlag, Radebeul and Berlin.

Egner, MT, Riehm, H & Domingo, WR, 1960. Untersuchungen über die chemishe boden-analyse als grundlage fur die beurteilung des nährsoffzustandes der böden. II. Chemische extraktionsmethoden zur phosphor und kaliumbestimmung. Kungl. Lantbrukshoegskolans Annaler, 26, 199–215.

Frossard, E, Condron, LM, Oberson, A, Sinaj, S & Fardeau, JC, 2000. Processes governing phosphorus availability in temperate soils and their relevance to phosphorus losses to water. J. Environ. Qual. 29:15-23.

Glæsner, N, Kjaergaard, C, Rubæk, GH & Magid, J, 2011.Relation between soil P test values and mobilization of dissolved and particulate P from the plough layer of typical Danish soils from a long-term field experiment with applied P fertilizers. Soil Use and management 29: 297-305.

Heckrath, G, Andersen, HE, Rubæk, G, Kronvang, B, Kjærgaard, C & Hoffmann, CC, 2009. Et web-baseret P-indeks som miljøplanlægningsredskab: del 1. Vand og Jord nr. 2: p44-48.

Heckrath, G, Brookes, P, Poulton, P & Goulding, KWT, 1995. Phosphorus leaching from soils containing different P concentrations in the Broadbalk experiment. J. Environ. Qual. 24: 904-910.

Hosseinpur, AR & Sinegani, AS, 2009. Correlation of Iron Oxide-Impregnated Paper Method with Selected Soil Phosphorus Tests and Alfalfa Indices. Soil Science and Plant Analysis, 40, 7-8; 1183-1190

ISO 11263, 1994. Soil quality – Determination of phosphorus – Spectrometric determination of phosphorus soluble in sodium hydrogen carbonate solution. International Organization for Standardization, Geneve, Switzerland. 5 pages.

Jordan-Meille, L, Rubæk, GH, Ehlert, PAI, Genot, V, Hifman, G, Goulding, K, Recknagel, KJ, Provolo, G & Barraclough, P, 2012. An overview of fertiliser-P recommendations in Europe: soil testing, calibration and fertiliser recommendations. Soil Use and Management 28, 419–435.

Kjærgaard, C., Rubæk, GH & Heckrath, G, 2010. Mobilisering og tab af fosfor fra pløjelaget og effekt af dyrkningsmetoder. I Plantekongres 2010: Sammendrag af indlæg, Aarhus Universitet, Det Jordbrugsvidenskabelige Fakultet s. 263-264.

Knudsen, L, Kristensen, K, Heckrath, G & Kjaergaard C, 2011. Udtagning af jordbundsprøver, opbevaring,og forberedelse til analyse. In: GH. Rubæk and P. Sørensen (eds.) Jordanalyser – Kvalitet og anvendelse. DCA Rapport nr. 002 Aarhus Universitet. DCA Nationalt center for Fødevarer og Jordbrug. p. 17-32.

Kristensen, ES, Rubæk, GH, Borggaard, OK, Mathiesen, GB & Ernstsen, V, 2011. Udvalgets anbefalinger. In: GH. Rubæk and P. Sørensen (eds.) Jordanalyser – Kvalitet og anvendelse. DCA Rapport nr. 002 Aarhus Universitet. DCA Nationalt center for Fødevarer og Jordbrug. p. 97-100.

MAFF, 1986. The Analysis of Agricultural materials. Third edition. Reference book 427. Ministry of agriculture food and fisheries. London, Her Majesty's Stationery Office. UK. Page 183-186

Mason, S, McNiell, A, Mc.Laughlin, MJ & Zhang, H, 2010. Prediction of wheat response t an application of phosphorus under field conditions using diffusive gradients in thin-films (DGT) and extraction methods. Plant Soil 337:243-258.

Mason, SD, McLaughlin, MJ, Johnston, C & McNeill, A, 2013. Soil test measures of available P (Colwell, resin and DGT) compared with plant P uptake using isotope dilution. PLANT AND SOIL, 373: 711-722

Menzies, NW., Kusumo, B., & Moody, PW. 2005, Assessment of P availability in heavily fertilized soils using diffusive gradient in thin films (DGT) technique. Plant and Soil 269: 1-9.

Menon, RG , Chien, SH, Hammond, LL & Arora BR, 1990. Sorption of phosphorus by the Iron Oxide-Impregnated filter paper embedded in soils. Plant and Soil, 126:,287-294

Morel, C, Schneider, A, Plenet, D & Fardeau, J C, 2002. Transfer of phosphorus in soil-plant and soil-solution systems using isotopic labelling and dilution methods. Nuclear techniques in integrated plant nutrient, water and soil management. Proceedings of an International Symposium held in Vienna, Austria, 16-20 October 2000, Pages: 49-56

Morgan, MF, 1941. Chemical soil diagnosis by the universal soil testing system. Connecticut Agricultural Experimental Station Bulletin 450, Connecticut, USA.

Mundus, S, Carstensen, A & Husted, S, 2013. Testing for plant available phosphorus in soils. IFS, International Fertiliser Society, Proceeding 738. Page 1-24

Natur og Landbrugskommissionen, 2013. Natur og Landbrug – en ny start. 124 pages (in Danish)(www.naturoglandbrug.dk)

Neyroud, JA & Licher, P, 2003. Do different methods used to estimate soil phosphorus availability across Europe give comparable results? J. Plant Nutr. Soil Sci. 166:422.431.

Nielsen, JD, 1979. Comparison of some analytical methods for evaluation of the P-fertility in soils Tidsskrift for Planteavl 83,485-491 (in Danish with English summary).

Nielsen, JD, 1981. Comparison of some analytical methods for evaluation of the P-fertility in some soils. Tidsskrift for Planteavl 85,31-38 (in Danish with English summary).

Olsen, SR, Cole, CV, Watanabe, FS & Dean, LA, 1954. Estimation of available phosphorus in soils by extraction with sodium bicarbonate. Circular 939, United States Department of Agriculture, Washington DC.

Plantedirektoratet, 1994. Fælles arbejdsmetoder for jordbundsanalyser (NK Sørensen & A. Bülow-Olsen red.) Plantedirektoratet, Landbrugsministeriet, Denmark.

Poulsen HD. & Rubæk GH. 2005. Fosfor i dansk landbrug. Aarhus Universitet, Det Jordbrugsvidenskabelige Fakultet. 211 p.

Rubæk, GH & Sibbesen, E, 2000. Long-term phosphorus fertilization – Effects on crop yield and soil phosphorus status. DIAS report no. 31 Plant production. Danish Institute of Agricultural Sciences, Ministry of Food, Agriculture and Fisheries.

Rubæk, GH & Sørensen, P (eds.), 2011. Jordanalyser – kvalitet og anvendelse. DCA rapport Nr 002. December 2011. 129 pages.

Rubæk, GH, 1999. “Soil phosphorus dynamics. Effects of land use, fertilisation and liming”. Ph.D. thesis. Danish Institute of Agricultural Sciences, Foulum and The Royal Veterinary and Agricultural University, Copenhagen

Rubæk, GH, Kristensen, K & Knudsen, L, 2011. Analyser for tilgængeligt fosfor i jord. In: Rubæk, G. H & Sørensen, P (eds.) 2011. Jordanalyser – kvalitet og anvendelse. DCA rapport Nr 002. December 2011. page 53-71.

Rubæk, GH, Kristensen, K, Olesen, SE, Østergaard, HS & Heckrath, GJ, 2013. Phosphorus accumulation and spatial distribution in agricultural soils in Denmark. Geoderma, 209-210: 241-250.

Schneider, A & Morel, C, 2000 Relationship between the isotopically exchangeable and resin-extractable phosphate of deficient to heavily fertilized soil. European Journal of Soil Science, 51, 709-715

Schoumans, OF, 2015. Phosphorus leaching from soils: process description, risk assessment and mitigation. PhD. Thesis, Wageningen University, Wageningen, NL. 261 pages.

Schüller, H, 1969. Die CAL-Methode, eine neue Mathode zurBestimmung des pflanzenverfu¨ gbaren Phosphats im Boden. Journal of Plant Nutrition and Soil Science, 123, 48–63.

Sibbesen, E & Sharpley, A, 1997. Setting and justifying upper critical limits for phosphorus in soils. I: Tunney, H, Carton, OT & Brookes, PC (red.) Phosphorus Loss From Soil to water, CABI Publishing, Wallingford, UK. s. 151-176.

Sibbesen, E, 1983. Phosphate soil tests and their suitablilty to assess the phosphate status of soil. J. Sci. Food Agric. 34:1368-1374.

Sissingh, HA, 1971. Analytical technique of the Pw method, used for the assessment of the soil phosphate status of arable soils in the Netherlands. Plant and Soil, 34, 483–486.

Six , L., Pypers, P., Degryse, F., Smolders, E. & Merkx, R. 2012. The performance of DGT versus conventional soil phosphorus tests in tropical soils – an isotope dilution study. Plant and Soil 359:267-279.

Six, L., Smolders, E. &Merckx, R. 2014. Testing phosphorus availability for maize with DGT in weathered soils amended with organic materials. Plant and Soil 316:177-192 DL (ed.). 1996. Methods of Soil Analysis part 3 - Chemical Methods. Soil Science Society of America/American Society of Agriculture. Madison, Wisconsin, USA.

Speirs, SD, Scott, BJ, Moody, PW & Mason, SD, 2013. Soil phosphorus tests II: A comparison of soil test-crop response relationships for different soil tests and wheat. CROP & PASTURE SCIENCE, 64: 469-479.

Tandy, S, Mundus, S, Yngvesson, J, de Bang, TC, Lombi, E, Schjoerring, JK & Husted, S, 2011. The use of DGT for prediction of plant available copper, zinc and phosphorus in agricultural soils.PLANT AND SOIL 346:67-180

Tóth, G, Guicharnaud, RA, Tóth, B & Hermann, T, 2014. Phosphorus levels in croplands of the European Union with implications for P fertilizer use. European Journal of Agronomy 55: 42–52.

Tunney, H, Breeuwsma, A, Withers, PJA & Ehlert, PAI, 1997. Phosphorus fertilizer strategies: Present and Future. In: Tunney, H, Carton, OT & Brookes, PC (red.) Phosphorus Loss from Soil to water, CABI Publishing, Wallingford, UK. p 177-203.

van der Zee, SEATM, van Riemsdijk, WH & de Haan, FAM, 1990. Het protocol fosfaatverzadigde gronden, Wageningen, Landbouwuniversiteit. Vakgroep Bodemkunde en Plantevoeding, The Netherlands.

van Raij, B, 1998. Bioavailable tests: Alternatives to standard soil extractions. Commun. in Soil Sci. Plant Ana. 29:1553-1570.

Videncentret for Landbrug, 2014. Resultater af kontrol med analysekvalitet af jordbunds-, husdyrgødnings-, og planteanalyser 2013/2014. 29 pages (in Danish). (https://www.landbrugsinfo.dk/planteavl/jordbund/jordbundsproever/Sider/saa-gode-er-laboratorierne-til-at-analysere-proever_pl_po_14_225.aspx)

Appendix 1. The sodium bicarbonate extraction method for testing soil P status - an updated description of the Danish "Ptal"

Gitte H. Rubæk, Department of Agroecology, Aarhus University

1.0 Principle

Phosphorus is extracted from soil with a sodium bicarbonate solution at pH 8.5 for exactly 30 minutes at 20 °C ± 1 °C, after which soil and solution are immediately separated. In the clear filtrate, the concentration of the blue phosphomolybdate complex is measured by spectrophotometry after adding sulphuric acid, ascorbic acid and ammonium molybdate reagent to the extract.

This method extracts only a modest proportion of soil total P and can therefore be very sensitive to small deviations in extraction time and temperature and intensity of shaking. Temperature should therefore be kept at 20 °C ± 1 °C from initiation of the extraction until soil and solute is separated. The bicarbonate extractant can produce coloured soil extracts, which may precipitate upon acidification of the extract during the colorimetric determination of P. These problems are handled by addition of polyacrylamide to the extracting solution as described by Banderis et al. (1976).

2.0 Apparatus

- Rotating shaking apparatus "end-over-end", shaking intensity 20 ± 2 rounds per minute.
- Scale for measureing 1-5 g with two decimal places.
- Acid-washed bottles and lids and glassware (material suitable for soil and not retaining P).
- Spectrophotometer or similar for determination of light absorbance at wavelength 880 nm.

3.0 Reagents

All reagents shall be analytical grade and water should be purified (Resistivity at 25 °C of maximum 18.2 MΩ·cm).

3.1. 4M sodium hydroxide solution. Dissolve 160.0 g sodium hydroxide (NaOH) pellets in 700 ml water. Cool and dilute to 1000 ml with water. Store the solution in an inert and hermetically sealed bottle.

3.2 Polyacrylamide solution. Polyacrylamide (Granular powder MW over 5.000.000, BDH Laboratory supplies prod. no. 297883N or similar) *ca.* 0.05% water solution. Dissolve 0.10 g polyacrylamide in 200 ml water. Note that it takes several hours to dissolve the polyacrylamide.

3.3 Extracting solution. Dissolve 210 g of sodium hydrogen carbonate ($NaHCO_3$) in 4500 ml water. Add 25 ml of the polyacrylamide solution (3.2). Adjust the pH to 8.50 ± 0.02 with the 4.0 M sodium hydroxide solution (3.1). Add water to 5000 ml volume. The solution should be prepared and sealed within 10-15 minutes. If stored air-tight, it can be kept for weeks. However pH should be controlled daily and a new solution should be prepared if pH deviates from 8.50 ± 0.04.

3.4. 4M Sulphuric acid: In a fume hood, pour *ca*. 350 ml of water into >1000 ml container, add 110.0 ml concentrated (95-97%) sulphuric acid (H_2SO_4) while stirring, cool to room temperature and add up to 500 ml volume.

3.5. *0.1M Sulphuric acid*. Dilute 4.0 M sulphuric acid (3.4) 40 times with water, by adding 25 ml 4.0 M sulphuric acid to *ca*. 900 ml water and fill up to 1000 ml volume with water.

3.6 Ammonium molybdate potassium antimonyl tartrate solution (Sulfomolybdic reagent)

a. Dissolve 13.0 g ammonium heptamolybdate-tetrahydrate ($(NH_4)_6Mo_7O_{24} \cdot 4H_2O$) in 100 ml water
b. Dissolve 0.35 g potassium antimonyl tartrate ($K(SbO)C_4H_4O_6 \cdot 0.5\ H_2O$) in 100 ml water
c. In a fume hood add approx. 120.0 concentrated sulphuric acid (95-97%) into *ca*. 170 ml water while stirring and cool to room temperature. Mix solution “a” and ”b” into the diluted sulphuric acid and fill up to 500 ml with water. Keep reagent cool and protect against sunlight.

3.7 Ascorbic acid solution

Dissolve 5.00 g ascorbic acid ($C_6H_8O_6$) in water and dilute to 100 ml volume.

3.8 Stock solution, 200 mg P/l. Dissolve 0.87775 g dried potassium dihydrogen phosphate (KH_2PO_4) in 1000 ml volume of 0.1 M H_2SO_4 (3.5).

3.9 Standard solutions

Prepare standard solutions with concentrations of PO_4-P ranging from 0 to 8 ppm as suggested in table A4 by appropriate dilution of the stock solution with the extracting solution (3.3).

Table A4: Concentrations of P in standard curve solutions and the amount of stock solution (3.8) to transfer to 100 volumes to obtain these concentrations.

PO4-P concentration (mg/l)	**Amount of stock solution (3.8) (µl) to dilute with extracting solution (3.3) to 100 ml volume**
0	0
0.1	50
0.2	100
0.5	250
1.0	500
3.0	1500
5.0	2500
8.0	4000

4.0 Procedure

4.1. Extraction

Weigh between 1.00 and 5.00 g dried at max. 50-60 ^{o}C, sieved (<2.0 mm) and well-mixed soil into a 50-250 ml flask or container. Ensure soil weight to container volume ratio of 1:50! Add extraction solution corresponding to a soil to solution ratio of 1:20 with a temperature of 20 ^{o}C ± 1 ^{o}C (3.3). Close flasks immediately and mount them on the shaker for exactly 30 minutes at 20 ^{o}C ± 1 ^{o}C. Within maximum 15 minutes after shaking has ended, start separation of soil and solute by either centrifugation of samples at minimum 1800 g for 5 minutes at 20 ^{o}C ± 1 ^{o}C or by filtration. When separated is carried out by filtration, the first milliliters of filtrate should be discarded. Prepare blanks following the same procedure, but excluding soil.

4.2. Measurement

Transfer 1 ml of extract to a beaker large enough to handle foaming and bubbles upon acidification (25 ml Erlenmeyer flasks work well. Handling in racks makes work easier). Add 9 ml of water and 125 µl 4.0 M H_2SO_4 (3.4). Swing flask and leave for CO_2 evolvement and foaming to cease. Then add 400 µl ascorbic acid solution and swing. Add 400 µl of the sulfomolybdic solution (3.6) and swing.

A standard curve is produced by transferring 1 ml of each standard solution and adding water, acid and reagents the same way as to the samples.

Flasks are left for 10-15 minutes at room temperature for colour development to complete. The blue colour is typically stable for up to 24 hours. The colour intensity of the samples and standards are measured on a spectrophotometer at 880 nm. Use the zero standard for setting zero. A path length of 1 cm of the cuvette is appropriate for most measurements, but at concentrations of less than 0.25 mg/l the path should be 4 cm or more. Make sure that bubbles of CO_2 do not obstruct the measurement.

If blanks do not produce zero absorbance or very close to (less than 0.004), the analysis should be repeated. A thorough check for contamination of reagents, bottles and glassware can be necessary.

If the soil extract is highly coloured, it should be tested if this colour absorbs light at 880 nm and if it does corrections for this absorbance will be necessary.

Automated procedures for measurements are accepted, as long as they rely on the above described principle of measuring the intensity of the blue colour developed after addition of the above mentioned reagents.

5.0 Calculations

Carry out a linear regression of measured absorbance of standard solutions against their known concentrations of P according to this equation:

$Abs_{st} = \alpha * C_{Pst}$

Where:

Abs_{st} is the measured absorbance for each standard solution,

C_{PSt} is the known P concentration in each standard solution,

and α is the constant derived from the regression line crossing the origin.

P concentration in the soil extracts can then be calculated as:

$P_{cons\_extract} = (Abs_{sample} - Abs_{blank})/\alpha$

The amount of bicarbonate-extractable P in mg P kg^{-1} dry soil can then be calculated as:

P extracted = $P_{cons\_extract}*20$

Detection area is 2 to 160 mg Olsen P kg^{-1} soil.

If the result is requested as the Danish Ptal, the result should be divided by 10 and the unit is then mg P extracted per 100 g of soil.

6.0 Repeatability

Reference soils should be included in every analytical run.The standard deviation of independent measurements on the reference soils measured at different times in the same laboratory with the same equipment should be less than 10% of the measured value or less than 2 mg P kg^{-1} soil.

7.0 Test report

A test report shall contain the following:

a. A reference to this method description

b. All information necessary for complete identification of the sample

c. Results of the determination in whole numbers in milligram per kilogram calculated on the basis of dried soil (dried at max. 50-60 ^{o}C)

d. Any details of operations not specified in this method description as well as any other factors, which may have affected the results.

8.0 Comments

This method description is an update of the former Danish method description (Plantedirektoratet, 1994) where key details of the procedure are described in more detail. It corresponds in major aspects to the ISO 11263:1994 and to the original method description by Olsen et al. (1954).

9.0 References

Banderis, A, Barter, DD & Henderson, K, 1976. The use of polyacrylamide to replace carbon in the determination of ”Olsen’s” extractable phosphate in soil. Journal of soil Science 27:71-74.

ISO 11263, 1994. Soil quality – Determination of phosphorus – Spectrometric determination of phosphorus soluble in sodium hydrogen carbonate solution. International Organization for Standardization, Geneve, Switzerland. Fem sider.

Olsen, SR, Cole, CV, Watanabe, FS & Dean, LA, 1954. Estimation of available phosphorus in soils by extraction with sodium bicarbonate. Circular 939, United States Department of Agriculture, Washington DC.

Plantedirektoratet, 1994. Fælles arbejdsmetoder for jordbundsanalyser (NK Sørensen & A. Bülow-Olsen red.) Plantedirektoratet, Landbrugsministeriet

Appendix 2. Draft for a protocol on how to correct Olsen P test results based on simultaneous analysis of standards

Kristian Kristensen, Department of Agroecology, Aarhus University

The purpose of correcting the soil is to ensure that the results can be compared across laboratories and times of analyses. However, in order to achieve this, it is important that the same method of correction is used in all laboratories and that the correction is based on the same soil samples. Here the procedure for such a correction based on calibration is described.

The calibration should be based on four standard soils and their "true" Pt-values, which should be obtained by triplicate analysis repeated at least five times over a period of at least one year according to the method description given in Appendix 1, preferably at two different laboratories.

These four standard soils should be included in each run. The four soils should be placed randomly in the sequence of all soils in the run.

When the results from a run are available, the data from the standard soils are regressed on the mean values from each standard soil, i.e. doing a regression analysis on the following eight observations:

Standard soil	Recorded values from the run	True values of the soils
1	Y_1	X_1
2	Y_2	X_2
3	Y_3	X_3
4	Y_4	X_4

This results in an equation: $Y_i \approx \hat{\alpha} + \hat{\beta} X_i$ for the calibration line where Y_i and X_i are respectively the recorded and true value of standard soil number *i*. $\hat{\alpha}$ and $\hat{\beta}$ are the estimated intercept and slope of the line.

In order to assess the quality of this equation (and the results from the run), a plot will show how well the equation fits the data and whether there are any serious deviations (see the following example).

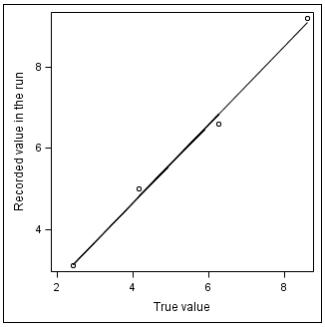


Figure A1 Plot of recorded values against true values together with their calibration curve. The coefficient of correlation is 0.997. The equation is Y=0.804+0.965X.

In addition, it should be tested whether the coefficient of correlation is larger than 0.970. If serious-/systematic deviations are found or the coefficient of correlation is less than 0.970, the results should not be used and new analyses of the standard and submitted soils should be carried out.

If no serious/systematic deviations are found and the coefficient of correlation is larger than 0.970, each of the submitted soils is calibrated using the following equation:

$C_j = (R_j - \hat{\alpha}) / \hat{\beta}$

where

C_j is the calibrated value for submitted soil j in the run

R_j is the recorded value for submitted soil j in the run

$\hat{\alpha}$ and $\hat{\beta}$ are the estimated intercept and slope of the calibration curve for the run

Both the recorded and calibrated values are reported to the submitter of the soils.

Appendix 3. Detailed description of applied statistical analysis and adjustment method

Kristian Kristensen, Department of Agroecology, Aarhus University

Data from the ring tests were analysed using a linear mixed model (see e.g. McCulloch and Searle, 2001) with the soil as a fixed effect. The data were first analysed separately for each laboratory. Here there were three random effects: Year, Time and Residual.

In the analyses for all laboratories, the following random effects were included: Laboratory, Year, Year:time, Laboratory:year, Laboratory:Year:Time and the residual variance.

Mathematically the models may be written as follows:

For data from each laboratory separately:

$$Y_{syt} = \mu + \alpha_s + B_y + C_{yt} + F_{syt} \quad \text{(model 1)}$$

For all data:

$$Y_{slyt} = \mu + \alpha_s + A_l + B_y + C_{yt} + D_{ly} + E_{lyt} + F_{slyt} \quad \text{(model 2)}$$

where

Y_{slyt} is the recorded value for soil s from laboratory l at time t in year y

μ and α_s is the intercept and fixed effect of soil s

$A_l, B_y, C_{yt}, D_{ly}, E_{lyt}$ and F_{slyt} are random effects that are assumed to be normally distributed with means zero and constant variance

The variance components from the models were used to calculate the variance between two samples submitted to different laboratories at the same or different times in the same or different years.

In order to see if the variance between two submitted samples could be reduced by including one or more soils as standards at each batch, two different methods were applied:

Adjustment using additive or multiplicative adjustment

Using one of the standard soils as a reference soil, i.e. to adjust the *submitted* samples for variability between laboratories and/or submission times, the following adjustments were examined:

Additive adjustment: $Z^{A}_{syt} = Y_{syt} - Y_{ryt} + R_r$

and

Multiplicative adjustment: $Z^{M}_{syt} = (Y_{syt} / Y_{ryt}) \times R_r$

where

Z^{A}_{syt} and Z^{M}_{syt} are the adjusted values of soil s at time t in year y

Y_{syt} is the recorded values of submitted soil s at time t in year y

Y_{ryt} is the the recorded value of standard soil r at time t in year y

R_r is the true value of standard soil r (here the average over all submitted samples of soil r was used)

The variables Z^{A}_{syt} and Z^{M}_{syt} were then analysed in the same mixed models as the recorded values for *submitted* soils and the estimated variance components from this analysis were used to calculate the variance between two samples in order to compare these variances with the ones calculated for the recorded values.

The calculations were performed using four different soils as standard soils covering the range of soils in the ring tests. The applied standard soils were not included in the mixed models.

Using a calibration curve based on four standard soils

Four of the 10 soils were selected as references. The standard soils were selected in such a way that they covered the range of Olsen P; the means were 2.4, 4.2, 6.2 and 8.6, respectively. These values are in the following treated as the true values of the soils. The remaining six soils (used as *submitted* soils) had mean values of 3.2, 3.4, 3.9, 4.2, 4.9 and 5.8, respectively.

For each laboratory and time, a calibration curve was estimated by fitting the best straight line to the actual response of these four standard soils against their means so that we for each laboratory and time have the following relation:

$P_{rlyt} = \alpha_{lyt} + \beta_{lyt} \times M_r + E_{rlyt}$

where

P_{rlyt} is the recorded amount of available P on the actual day (time t in year y) for standard soil r in laboratory l

M_r is the true value of standard soil r (here their mean value were used)

α_{lyt} and β_{lyt} are specific values for describing the calibration curve for the actual laboratory and time (time t in year y) in laboratory l - and were estimated by linear regression

Then this relation was used to calibrate the values on each day:

$C_{slyt} = (P_{slyt} - \hat{\alpha}_{lyt}) / \hat{\beta}_{lyt}$

where

C_{slyt} is the calibrated value for soil s at time t in year y for laboratory l

P_{slyt} is the recorded value for soil s at time t in year y for laboratory l

$\hat{\alpha}_{lyt}$ and $\hat{\beta}_{lyt}$ are the estimated intercept and slope of the calibration curve for time t in year y for laboratory l

Reference

McCulloch, C.E. & Searle, S.R. 2001. Generalized, Linear, and Mixed Models. John Wiley & Sons, Inc. 325 pp.

Appendix 4. Effect of adjustment of mean of 1, 10 or 40 samples submitted at the same time using all four standard soils

Kristian Kristensen, Department of Agroecology, Aarhus University

The effect of calibration on the mean if more samples were submitted to different laboratories at the same or different times could not be calculated from the variance components (because of the correlation between the individual adjusted values), but only if a given number of samples for each soil were sent in a ring test and analysed together with the four standard soils. However, it was possible to calculate the effect of additive adjustment (using the variance components) for the standard error on the difference between two means. Therefore, this was done for a mean of 10 and 40 samples using an adjustment based on the mean of the four standard soils. The variance for this situation is given by the following formula:

$Var(\overline{Y}_{slyt}) = \sigma^2_{EY} / n$

$Var(\overline{Z}^{A_4}_{slyt}) = \sigma^2_{EZ} / n + \sigma^2_{EZ} / 4$ The term, $\sigma^2_{EZ} / 4$,was included in order to take into acount the correlation between the adjusted values used i the same mean

where $\overline{Y}_{slyt}$ and $\overline{Z}^{A_4}_{syt}$ are the mean of unadjusted and adjusted Pt-value for soil *s* submitted to laboratory *l* at time *t* in year *y*

σ^2_{EY} and σ^2_{EZ} are the residual variance for the unadjusted and adjusted Pt-values, respectively

n is the number of replicates (here 10 or 40).

The results are shown in Table A2 and A3.

Table A2 Approximate standard error on difference between two rounds of measurements at each laboratory for raw data and adjustment based on four standard soils assuming that either 10 or 40 individual samples were analysed each time. (The unit for Ptallet/Olsen P is "mg P/100 g soil".) The calculations are performed using the correlation that would apply if an additive adjustment using four standard soils had been used.

Laboratory	**Submission time**	**For recorded values**			**For adjusted values**		
		1 sample	10 samples	40 samples	1 sample	10 samples	40 samples
Commercial 1	Different year	0.97	0.80	0.78	0.61	0.38	0.35
	Same year	0.90	0.70	0.68	0.61	0.38	0.35
	Same time	0.59	0.19	0.09	0.59	0.19	0.09
Commercial 2	Different year	0.49	0.35	0.33	0.43	0.31	0.29
	Same year	0.42	0.23	0.21	0.43	0.31	0.29
	Same time	0.36	0.12	0.06	0.36	0.12	0.06
Commercial 3	Different year	0.50	0.28	0.25	0.72	0.62	0.61
	Same year	0.50	0.28	0.25	0.70	0.60	0.58
	Same time	0.45	0.14	0.07	0.45	0.14	0.07

For commercial laboratory 1 the standard error on the difference between the two means was reduced by approximately 50% if the two sets of samples were submitted at different times. For commercial laboratory 2 the reduction was less (about 10%) if the two sets of samples were submitted in different years, whereas the effect of adjustment was negative if the two sets of samples were submitted at different times in the same year. For commercial laboratory 3 the effect of adjustment was always negative (Table A2).

On average (over all laboratories) the standard error on the difference was reduced by 20-25% if the two sets of samples were submitted to different laboratories in different years. If the two sets of samples were submitted to different laboratories in the same year or to the same laboratory in different years the average reduction was less (ca. 12-15%) and if the two set samples were submitted to the same laboratories in the same year (or at the same time) the average reduction was very close to 0 (Table A3).

Table A3. The average standard error on the difference between two measurements at each Laboratory for raw data and calibration adjustment on four standard soils where either 10 or 40 samples are analysed each time and where the individual samples are not necessarily the same in each round of analyses (the unit for Ptallet/Olsen P is "mg P/100 g soil"). The calculations are performed using the relevant correlation if an additive adjustment using four standard soils had been used.

Submission time	**For recorded values**			**For adjusted values**		
	1 sample	10 samples	40 samples	1 sample	10 samples	40 samples
Different lab/different year	0.77	0.61	0.60	0.62	0.48	0.46
Different lab/same year	0.73	0.55	0.54	0.62	0.48	0.46
Same lab/different year	0.70	0.52	0.51	0.61	0.46	0.44
Same lab/different time	0.66	0.46	0.44	0.61	0.46	0.44
Same lab/same time	0.50	0.16	0.08	0.50	0.16	0.08

The columns for "1 sample" may be compared to the values in the results shown when using adjustment with just 1 standard soil and to values shown when using calibration (in table 5.3 and 5.4 in section 5 of the main report).

The available data did not allow us to estimate whether the reduction would be the same for the calibration method, and not even whether it would be bigger or smaller than for the above-mentioned example. However, there are reasons to expect that it will not deviate a lot from the above-mentioned calculation These reasons are:

- The uncertainty on a single sample (tables 5.3 and 5.4 in the main report) and using additive adjustment using an average of four standard soils (Appendix IV) are comparable.
- There are also elements pulling in opposite directions which could cause the uncertainty to be either bigger or smaller than in the above-mentioned example; these elements could counterbalance each other.

DCA - National Centre for Food and Agriculture is the entrance to research in food and agriculture at Aarhus University (AU). The main tasks of the centre are knowledge exchange, advisory service and interaction with authorities, organisations and businesses.

The centre coordinates knowledge exchange and advice with regard to the departments that are heavily involved in food and agricultural science. They are:

Department of Animal Science
Department of Food Science
Department of Agroecology
Department of Engineering
Department of Molecular Biology and Genetics

DCA can also involve other units at AU that carry out research in the relevant areas.

SUMMARY

Analyses of soil P status with soil P tests have for many years formed the cornerstone for recommendations on how to fertilise agricultural soils. The increasing awareness of the role of soil P as a contributor to surface water eutrophication together with the renewed focus on phosphate rock as a valuable non-renewable resource has put emphasis on the way we utilise P in soils, fertilisers, manure and waste products. To ensure this, a valid, precise and reliable soil P test method is crucial. Olsen P (in Denmark known as “Ptallet” or “fosfortallet”) is the official soil P test method, but it has long been recognised that Olsen P results vary too much and apparently systematically between labs and over time. Moreover an increasing body of evidence indicates that the Olsen-P method too frequently does not reflect the P availability to plants in soil. It is therefore clear that initiatives leading to better soil P tests in Denmark with high laboratory precision and valid information on soil P status for farmers, researchers and authorities are highly needed. This report gives recommendations on how to improve soil P testing in Denmark and a draft for an updated protocol for the Olsen P analysis is presented.