

Nordsyn ecodesign task 1 report

Improvements of the
EN16510 test standard

Contents

1. Background	3
2. Task description – Nordsyn project part 1	5
3. Test philosophies compensating for lack of reproducibility	6
Technical assessment of existing relevant test parameters and conditions (a) – including suggestions for modified and improved test proceedings (b)	7
3.1 Start and end definitions of the cycle	7
3.2 Determination of flue gas temperature	9
3.3 Order and number of burn cycles	9
3.4 Reset minimum cycle duration to 45 minutes	10
3.5 Avoid any intermediate pauses between combustion cycles	11
3.6 Part load heat output testing	12
3.7 Standardize the test fuel	13
4. Assessment of practical feasibility and added costs for manufacturers	15
5. Estimation of the achievable tolerance improvement by adopting these recommendations	19
6. Summary	21
Annexes	22
About this publication	23

This publication is also available online in a web-accessible version at:
<https://pub.norden.org/nordicenergyresearch2025-06>

1. Background

This project was initiated by the Energy Agencies in Denmark, Norway and Sweden aimed at identifying and addressing elements of the EN16510-1:2022 standard that currently leads to poor reproducibility of test results, when type testing Solid Fuel Local Space Heaters (SFLSH or 'stove' for simplicity) to CE-labelling. There is special attention on the batch operated log wood appliances (over pellets) as the performance characteristics of batch operated stoves are more likely to be influenced by different manual operational habits. However, some of the recommendations like the suction pyrometer and CO₂ end of test criterion are technology neutral and can be used equally for testing of pellets stoves.

Here we use the term *reproducibility* to describe variability in test results, fully aware that there are two aspects involved, *repeatability* when conducting a series of tests within the same test lab and *reproducibility* when re-testing the same or an equivalent SFLSH at another test lab. This situation can occur in case of a market surveillance action, a round robin test (where the same stove is circulated among several test labs) or a proficiency test (where test labs are requested to conduct tests for benchmarking purposes and report results to a well-defined test protocol issued by the proficiency test operator).

The review of the EN16510 family of standards was initiated by CEN/TC295 in 2012 with purpose of replacing the former product specific room heater standards in the EN132XX series. This review was mainly justified by needs of upgrading the room heater standards following the Construction Products Regulation – CPR (Regulation (EU) no 305/2011) in 2013 repealing the former Construction Products Directive – CPD (Directive 89/106/EEC) from 1988.

The CPR repealing the CPD posed requirements on industry to declare the performance of their construction products with a Declaration of Performance (DoP) involving several Basic Work Requirements (BWR) and essential characteristics. The lack of these formats in the former CPD posed an increasing problem for industry, as national or regional building authorities began requesting updated Declarations of Performance.

Before that, the EN132XX series of standards replaced national room heater standards starting with the EN13240 standard for room heaters issued in 2001 and harmonized to the CPR in 2004 to allow CE-labelling of the room heaters.

During the 10-year drafting period, CEN and the testing organisation introduced several patches on the EN132XX series of standard, by means of decisions and interpretations made by the Group of Notified Bodies network, sector group 03 working group 02 (GNB SG03-WG2), which is a CPR framework under DG GROW.

However, in several cases, an outdated test option was failed to be removed, when a new option was introduced, which over time blurred the testing methodology, leaving several paths through the test for disadvantage of the reproducibility.

This project is about identifying and correcting these deficiencies, and further to suggesting improvements of the test, based on extensive experience with type testing among the partners DTI, RISE and SINTEF.

According to the task description DTI was lead on this task 1, however the initial draft was duly adapted according to multiple valuable contributions from RISE and SINTEF. Thanks for your commitment.

2. Task description – Nordsyn project part 1

Project I: Modification and improvements of test standard, EN16510

- a. Technical assessment of existing relevant test parameters, measurements methods and test conditions.
- b. Suggestions for modified and improved version of EN16510 test standard
- c. Assessment of practical feasibility and cost estimation for manufacturers
- d. Estimation (cautious) of the potentially achieved tolerance percentage.

The full project description is attached as [annex 1](#).

3. Test philosophies compensating for lack of reproducibility

Batch operated stoves are unique in the way that the resulting performance relies on not only physical properties of the stove, and the fuel used, but also the skills and abilities of the person operating the stove. This contrasts with pellets stoves and -boilers, which have a high degree of automation of the operation and use a standardized pellets fuel type, leaving only limited space for an element of the randomness caused by fuelling and operating habits.

Reproducibility as discussed is in this context, is defined as attainment of identical or close to identical test results, when a stove for some reason is tested at another test lab, different from the one that carried out the initial type test. Such cases include re-testing commissioned by the manufacturer, audit testing commissioned by a labelling scheme operator, round robin measurements commissioned by a label scheme operator, and market surveillance measurements commissioned by a national or EU market surveillance authority (MSA).

From a tradition standardization perspective, lack of reproducibility of test results can be compensated by using more elaborated and unambiguous testing approaches. This can be achieved by narrowing down several test methods, parameters and peripheral conditions to a minimum and setting standardized conditions where a whole range of unpredictable operational conditions occur. Another useful approach is to increase the number of repetitions to further exclude random impacts on the test result.

In the next chapter 'Potential for improving the current test method' we address the most influential dual options and other misconduct forming the basis for poor reproducibility of test results.

According to the task description of this Nordsyn task 1 project, reproducibility ranks over Real-Life operation. The partners SINTEF and RISE have most diligently contributed with several good proposals, some of them targeting better approximation of the testing method to Real-Life operation. However, at the expense of reproducibility, these initiatives are discarded in this evaluation, but they are listed in [annex 3](#) for the sake of order, and as inspirations for any possible future refinement of the European testing method for batch operated wood stoves.

Technical assessment of existing relevant test parameters and conditions (a) – including suggestions for modified and improved test proceedings (b)

This chapter aims to identify elements of the EN16510 test procedure that holds potential of improvement, with the objective of improving reproducibility of test results. Several of the listed issues, have over time been driven away from the starting point. Solving these issues will improve reproducibility of test results. It is worth to bear in mind that test labs can come under some pressure from industry, to produce test reports with 'good results', not to mention the never-ending arms race between industry and policy makers, which in other sectors have caused flawed tests over time such as the well-known diesel gate event – or the many defective hip implants (Cohen, Deborah 2012) – in Denmark further aggravated by the Boneloc scandal.

3.1 Start and end definitions of the cycle

Tests are accomplished having the stove mounted on a platform scale in the so-called test trihedron. It is an assembly of walls, a floor and a ceiling fitted with thermocouples to enable determination of safe clearance to combustible material next to the stove.

The test is a warm-to-warm test, meaning that the test is not begun until the stove has reached a thermal equilibrium and there has been build up an adequate bed of embers to safeguard rapid ignition of the next fuel batch entered. This is called the *basic firebed*. When the test begins, the mass of the basic firebed is recorded and the burn cycle shall proceed until the fuel has been fully combusted and the flames are out, leaving only the same mass of embers and glowing char coal as from the beginning – the *basic firebed*. The platform scale is used to identify this value.

According to fuel analysis, the ash remainder of hardwood without bark is approximately 0,3% of the mass, slightly more with bark. For a typical fuel load at 1,4 kg hardwood, the ashes amount to 7 grams based on 0,5% ash content. Consequently, one needs to adjust the basic firebed by the build-up of ashes from each refuelling, but ideally no more than 0,5% or typically 5–10 grams for each burn cycle.

Over time this principle has succumbed in some labs. Many are building up the firebed by as much as 50 grams per refuelling and under EN16510 even as much as 100 grams. This praxis inevitably led to fuelling on flames and transferring of unburnt wood fuel from one burn cycle to the subsequent, rendering the heat output determination invalid and reporting of artificially high efficiency numbers.

Tests are accomplished having the stove mounted on a platform scale in the so-called test trihedron. It is an assembly of walls, a floor and a ceiling fitted with thermocouples to enable determination of safe clearance to combustible material next to the stove.

It is therefore recommended to define the basic firebed as an amount of char and glowing embers, but without visible flames, as 15–25% of the mass of the fuel load. And further define that the firebed shall not be raised by more than the mass of the ashes produced by combustion one batch of fuel, which is typically 0,3–0,5% of the fuel mass. As an example, given the fuel load is defined by the manufacturer as 1,40 kg, the test shall begin on a basic firebed in the range of 210 to 350 grams and the build-up mass from one burn cycle to the next shall be maximum 7 grams. For simplicity it is suggested to agree on incremental masses of 10 grams for fuel loads at or below 2 kg (this will cover 90% of alle tests), 15 grams for fuel loads between 2–3 kg and 20 grams for fuel loads higher than 3 kg.

There are a few appliance types that are not well suited for use the platform scale as criterion for refuelling time. They are limited to appliances with a high mass close to or above the maximum capacity of the platform scale (500 kg) and appliance types with a water circuit for indirect heat dissipation to a water based central heating system.

Those types of appliances usually suffer from *thermal lift forces* when warming up and the opposite when cooling down. This thermal lift force can affect the approximation toward the basic firebed to such a degree, that the platform scale reading becomes useless as a criterion. To overcome that problem, an alternative refuelling criterion by CO₂ was introduced years back. However, there are many indications that the CO₂-criteria is chosen by some test labs, even for ordinary appliance types that does not qualify for it.

In case thermal lift forces is recognized as a problem, it is possible to counter its effects in terms of lack of precise mass determination, by carrying out the so-called **hot tare operation**. Hot tare is accomplished by first heating the stove to a thermal equilibrium. Then remove all embers, tare the platform scale and put the embers back in place again in rapid succession. This operation will balance any thermal lift forces.

This diversion leads to uneven testing practices among the test labs, and it hampers reproduction of test results. It is also a problem for market surveillance measurements, if a stove is tested according to one criterion and checked according to another. The benefits in terms of aligning the tests to one uniform mass-based refuelling criterion far outweighs any disadvantages for the rare stove special stoves type (boiler stoves amounts no more that 1% of the total test volume).

It is there recommended to remove the CO₂-option from the standard, except for boiler stoves and for stoves with a mass of over 500 kg. Should thermal lift be a problem for stoves with a mass less than 500 kg, making the platform scale not trustworthy, test labs are encouraged to solve the problem by performing the 'hot tare' operation which will balance any thermal lift.

3.2 Determination of flue gas temperature

Ever since the national test standards and certainly from the very beginning of the EN13240 standard from 2001, the flue gas temperature was determined with the so-called suction pyrometer. It is a metal tube placed across the measurement section, through which a small proportion of the fumes is extracted with a pump, thus heating a thermocouple placed inside the tube to the flue temperature. A suction pyrometer is only working properly when the flow through the pyrometer is high enough. The difference in not having enough flow can lead to up to 50 °C lower temperature and thus affecting the efficiency by several points. To avoid this element of uncertainty an alternative way of measuring the flue gas temperature with shielded thermocouple placed in the centre of the measurement section was introduced years back, but without ruling out use of the suction pyrometer.

Again, an undue diversion that leads to different practises throughout the test labs in Europe. As flue gas temperature is decisive for the energy efficiency together with the CO₂ percentage, it leads to poor reproducibility and reporting of a flawed energy efficiency figure.

It is therefore recommended to remove the option of using a suction pyrometer and rely alone on determination of the flue gas temperature by means of shielded thermocouples, placed in the centre of the measurement section.

3.3 Order and number of burn cycles

The dominating operating regime of wood stoves is called intermittent combustion. It means frequent refuelling of the stove with a limited amount of wood over the use period of an evening or a day. The opposite regime is called continuous combustion, where a substantial fuel load is entered at once and left to burn down unattended over an evening or a night. While intermittent combustion is the dominating regime in Europe, continuous combustion is the normal operating regime in the United States, including unattended over-night burning at a minimum air valve setting.

Back in EN13240 it took three burn cycles to compose a test result. However, one of them could be de-selected by the test lab, so in the end test results were only based on two combustion cycles.

Although clearly intended to match the intermittent combustion regime, it was nowhere in the standard stated that the three burn cycles had to be in direct succession of each other. This loophole has been used by some test labs to compile tests from several odd combustion cycles, perhaps even from different days. By the current EN16510 standard this loophole was closed, but only partially. Currently a test result can consist of two consecutive combustion cycles and one odd one.

It will be only natural to request combustion cycles in direct continuation or even better increase the number of cycles from three to four or even five combustion cycles in succession. This will lead to more robust and reproducible test results. Incidentally, the UK always requested five repetitions for their national particle tests under the Clean Air Act.

It is therefore recommended to request strictly consecutive combustion cycles with no pause or slip in-between and at the same time to increase the number of burn cycles from 3 to 5.

3.4 Reset minimum cycle duration to 45 minutes

There was no other justification for reducing the minimum cycle duration from formerly 45 minutes to now 40 minutes, other than better enabling 'cutting the tail'.

The time reduction shall be seen in context with allowing to artificially raise the firebed not only by 50 grams but as much as 100 grams under EN16510. In combination the two measures have degraded the former finite cycle definition to a sort of designer test which can end at almost any time or firebed, as long as the 3 combustion cycles of a test are kept within 10% variation by heat output.

This extended manoeuvre space of 5 minutes times 100 grams is illustrated by the red rectangle in the figure below. It's harmful for reproducibility and it hampers effective market surveillance measurements as well.

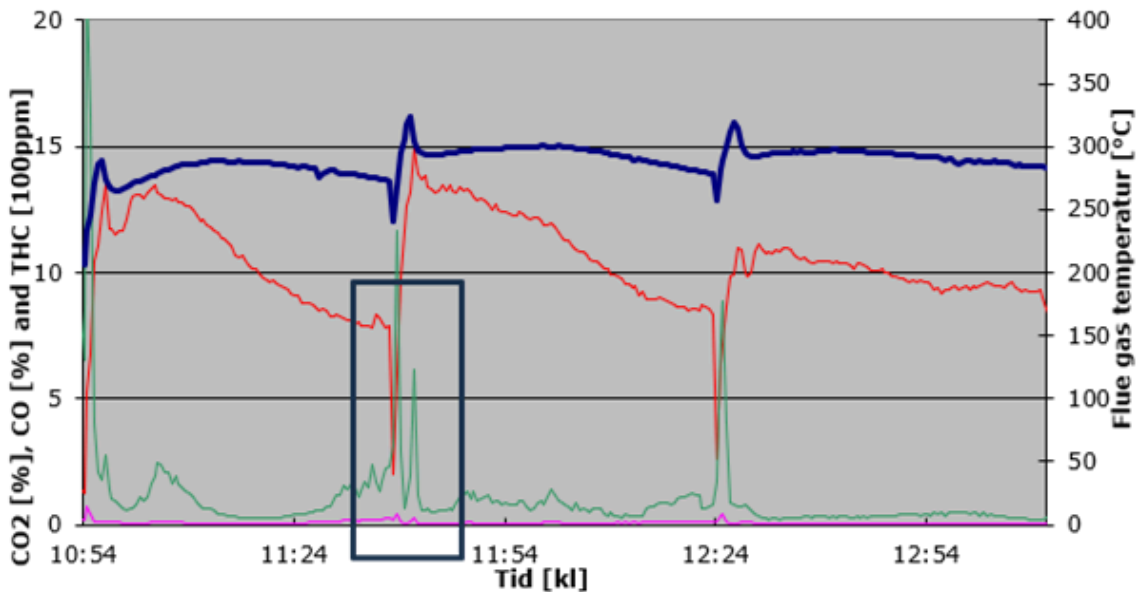


Figure 1 Typical course of output testing

It is therefore recommended to reset the cycle duration back to 45 minutes as a minimum. In daily use, refuelling too late is a much bigger problem than refuelling too soon. Ideally, the refuelling interval during the test should correspond to the refuelling interval in daily use, otherwise there will be a strong bias towards higher emissions during field use. Therefore, a refuelling interval of 45 minutes is far preferable to 40 minutes.

3.5 Avoid any intermediate pauses between combustion cycles

Since the combustion regime is **intermittent batch combustion**, it follows that the combustion cycles must be consecutive; that is, each subsequent cycle must begin immediately after completion of the previous combustion cycle. There must be no delay or idle period between cycles, and in particular, sampling of pollutants must not be paused between combustion cycles for any reason.

Recently, Germany has proposed introducing an intermediate pause of up to two minutes, during which sampling of all the pollutants is stopped. The reasoning for this is presented as time needed to replace the PM filter prior to the next combustion cycle, but this does not justify pausing CO, OGC and NO_x measure-

ments in the same time PM is paused. However, the filter can and should preferably be replaced during the last couple of minutes of a combustion cycle. Changing the filter during the last minutes of a cycle is possible without affecting the results reported, as any, however tiny reduction of particles collected by skipping the last minute or two is offset by a proportional reduction in the sampled air volume, resulting in the same reported emission in mg/m³.

Alternatively, if the EN-PME is more difficult to operate than assumed, it is possible to apply the former Heater Filter sampling method occupying only 33 minutes of a burn cycle, thus leaving ample time to replace the filter before the next combustion cycle begins. Again, this can be done without affecting the PM results reported, as the EN-PME validation project (CEN/TC295-WG5 documents N282 and N315) has most conveniently established that the Heater Filter and the EN-PME is reporting the same order of magnitude of fine dust emissions. The Heated Filter method is just much simpler to operate than the EN-PME.

Allowing an undefined break of 0–2 minutes will only lead to greater variability and reduced reproducibility. If introduction of the EN-PME method introduces new problems not anticipated, problems must be solved by altering the EN-PME method, not by altering the EN16510 standard.

It is therefore recommended to reject the proposal on a break of the test sequence – or revert to the former HF sampling method for determination of emission of fine dust (PM), which has proven its reliability in more than 20 years.

3.6 Part load heat output testing

Testing at part heat load means adding an operation condition frequently occurring in daily use of stoves, where homeowners try to limit heat release from the stove by turning down the air controls. Turning down the air supply or using less wood will inevitably move the working point away from the ideal centre point at Lambda close to 1 (which by large corresponds to the Nominal heat output condition) in direction of *rich combustion* where the emissions are inevitably higher.

However, if stoves with manual air controls are effectively exempt from the market by ecodesign regulation, the relevance of testing stoves at part load is limited. Part load operation in the sense of turning down the air controls is only relevant for stoves having manual air controls, while an ACC stove does pursue its setpoint(s), regardless it is set to operate at a certain fluegas temperature, room temperature or a specific Lambda figure.

In case automated air controls should become compulsory in the next ecodesign regulation, there is no justification for requesting testing at part load. Instead, it

might be considered whether it is possible to include in the test, ways of varying heat output from low up to high heat output, as permitted by the ACC technology. Appliances with automatic combustion control (ACC) are developed with multiple heat output settings, typically ranging from low to high. While ACC can improve combustion performance, this improvement is more effective when verified through type testing at different wood fuel load, reflecting the conditions that occur during actual use.

There is currently an ACC test under development by CEN/TC295 WG1, by the task group TGE. This work might suggest ways of varying the heat output rate for testing purposes.

It is therefore recommended to reserve testing at part load to manually controlled batch operated wood stoves exclusively, until in due course of time, the WG1 TGE task group has developed an ACC testing method.

3.7 Standardize the test fuel

Varying shape and dimensions, including how the fuel is chopped, placed and stacked in the firebox, has an influence on the combustion quality and hence for the test results. In other parts of the world there is a preference for using standardized test fuels to homogenize the tests, for benefit of reduced variability, however still using hardwood for the output tests. This standardized fuel load, which is inspired by NS3058 or US EPA Cribwood testing, is composed by pieces of lumber which is assembled as a 'fuel crib' of variable length and numbers to match the target fuel load mass, but with fixed spacing.

This type of test fuel load can be adopted regardless of maintaining the current approach where the fuel mass is defined by the manufacturer or upgrading to an approach like that of the NS3058 or US EPA with the fuel mass calculated as function of the firebox volume. The latter approach however having more implications, as it is going to affect also the refuelling interval and the rated heat output value too. It is a slightly more intrusive intervention that needs to be considered in more detail. But transforming the current fuel load, as defined by the manufacturer from log wood to standardized fuel is straightforward.

There is reason to assume that the current 50x50 mm cross dimension might be too tiny for an output tests supposed to last 45 minutes. There are indications that cross dimensions 80x80 mm or 90x90 mm are better suited for this purpose. The nature of this issue should be explored in more details by a literature review or by experimental measurements, before selecting the final dimension. A study from TFZ was published at the 20th European Biomass Conference in 2012 involving different fuel dimension and number of 'logs' per refuelling, could serve as a starting point for further investigations.

Use of standardized test fuel will benefit repeatability of combustion cycles within the same test lab and as well reproducibility when testing the stove at another test lab. Further to the physical dimensions, the fuel moisture and the wood species is open for discussion. Currently birch, beach and hornbeam are accepted as test fuels and the current fuel moisture interval is 12–20% on wet basis. Over time concerns have been flagged by test labs in the southern part of Europe that these wood species are not commonly available in their part of Europe. They would like oak added as an extra fuel option.

It is therefore recommended to standardize the test fuel by replacing natural wood logs with a standardized fuel crib but still made of hardwood. We are open for discussions on requesting also the fuel mass calculated as function of the firebox volume, taking into due account any implications on refuelling interval and heat output determination, as in most cases, the calculated fuel mass is major to the one defined by the manufacturer.

4. Assessment of practical feasibility and added costs for manufacturers

Treating the initiatives one by one we will here evaluate feasibility and any effects on testing costs.

Increase the basic firebed by ashes build up only

This initiative has no impact on feasibility or cost. It only implies that the combustion cycle duration is extended by 3–5 minutes typically.

Skip the CO₂ option as end of test criterion

This initiative has no impact on feasibility or costs. It only implies that the combustion cycle duration is altered a little. Some combustion cycles will last a few minutes longer, some a few minutes shorter. This is due to the unpredictable drop of the CO₂ value as the fire goes out.

Skip the suction pyrometer option (flue temp)

This initiative has no or little impact on feasibility and costs. Some test labs might keep the suction pyrometer as an implement to extraction of gasses for CO and CO₂ analysis, some might switch to using a heated probe and sampling line (same principle as for THC and NO_x).

Increase number of burn cycles from 3 to 5

This initiative has no or little impact on feasibility and costs, despite extending the effective testing time. Test labs are assigning a full day for an output test, where they are able to complete 7–8 combustion burn cycles, assuming the average cycle duration is 45 minutes and the first hour or two is spent on ignition and pre-heating the stove to a thermal equilibrium. So, the time is spent anyway, and it is no problem to accomplish the minimum combustion cycles in sequence in one day. If the stove burns uneven or it is hard to control due to poor preparation for the test, it might take an extra day to frame the output test, but this is frequently occurring already today.

Some manufacturers have a full in-house test lab and skilled technicians, that allow them to trim the stove and its operation to an optimum before undertaking the test. Other manufacturers have only limited in-house test facilities, and they will typically spend more time at the test lab to accomplish a satisfactory test result. In the end it is a trade-off between investing in elaborated in house test capacity and save time and money at the test lab – or having no or little internal test capacity and spend more time and money at the test lab. Going from 3 to 5 burn cycles does not change this fundamental trade-off. If the stove is well trimmed by an experienced technician, it is no big deal producing 5 good combustion cycles in sequence.

Reset minimum cycle duration to 45 min

This initiative has no or little impact on feasibility and costs. Obviously, the combustion cycle duration is going to be extended by a couple of minutes. This is however not going to affect the yield of a days of work in the test lab, in terms of number of combustion cycles accomplished.

Remove inter-cycle break (and revert to HF)

The first part of removing the break is of paramount importance on the test quality. It can be accomplished regardless of keeping the EN-PME dust sampling principle or skipping the EN-PME and reverting to the former Heated Filter sampling principle. There are no added cost costs nor any adverse feasibility effect of removing the break (in many test labs, the pause has been introduced yet as it is not yet formalized in the EN16510 standard). Reverting to the Heated Filter (HF) sampling principle represents an upside, as the HF principle is significantly easier to operate – it can be handled by one staff member alone and it offers ample time to replace the filter before commencement of the subsequent combustion cycle. Emission wise there is no difference as demonstrated by the WG5 validation project referred above proving parity between EN-PME and HF.

Include part load testing

If ACC controls (Automated Combustion Controls) become compulsory there is no justification for introducing part load testing at the time being. CEN has not yet drafted a test method for ACC stoves; the work has just begun. Part load testing can currently be reserved to manually controlled stoves alone, where it is well justified and there is a method documented in the test standard. Introducing part load testing almost double the time consumption at the test lab for the output test, and consequently testing costs are increased, yet not doubled as the temperature safety test is not affected by part load. There can also be an effect on

feasibility, if the stove is not well trimmed for part load operation, as discussed above. But most important, part load is counterproductive to test reproducibility. Part load is further away from the ideal working point combustion wise and it carries a wider range of results than measurements at nominal heat output.

Standardize the test fuel

This initiative has no impact on feasibility or costs. The fuel costs are going to be the same and so is the time consumption at the test lab. There might be a potential minor upside, if the relative standard deviation is minor than when using log wood. This could minimize needs for retests.

Summing up the effects of these initiatives, setting part load testing aside, and considering upsides and downsides, the effects in terms of feasibility as well costs are close to neutral. Only part load is undoubtedly a cost driver. Assuming part load is reserved to manually controlled stoves alone, this bias could also urge manufacturers of manually controlled stoves to move on and develop ACC systems faster than otherwise planned.

Other improvement initiatives suggested by the partners worth considering – the effect however is not easily quantified.

- Enable better correlation between how the appliance is fired during the Initial Type Test (ITT) and how the firing description is written in the instruction manual of the appliance. Introduce the Quick User Guide (QUG).
- Allow only optimizing the air control setting during the first three minutes (e.g. with a stopwatch and a caliper) if it is clearly stated in the instruction manual under start up guidance.
- Have better clarification in the test standard to which degree a manufacturer may witness a type test - or take active part in some of the operations during the test. Consider requesting any guidance to the test lab documented by a set of testing instructions written or as a video or a combination of both.
- Enforce compulsory participation of NoBos in interlaboratory comparisons tests among the members of the Group of Notified bodies network SG03-WG2 in fixed 3-year intervals. And always when a new standard replaces the former one. There is already a position paper from GNB-CPR/GNB-AG stressing the importance of using interlaboratory comparative tests as an essential part for ensuring even performance of the Notified Bodies. It is further stressed by ISO17025 clause 7.7.2 that Notified Bodies are obliged to take part in interlaboratory tests.

So, all in all the framework is in place. We just need to make sure it happens. This might be supported by closer follow-up by the notifying authorities appointing the NoBos, and/or the national accreditation bodies monitoring the operation of the NoBos. The paper from GNB is attached as [annex 2](#).

Such compulsory interlaboratory comparisons will have a significant positive impact on reproducibility of tests, and it could also work as a strong QA tool, strengthening the impact of Market Surveillance measurements. In combination, the two measures will be very effective – one at the front end and the other on the output of the test.

5. Estimation of the achievable tolerance improvement by adopting these recommendations

There is limited research or literature addressing the repeatability of test results from wood stoves, or the relationship between individual factors and overall repeatability. The French ecolabelling scheme, the Flamme Verte does organize round robin tests on a regular basis, where the same stove is circulated among their affiliated test labs. Taking OGC as one of the more volatile parameters the test results reported range from 29 to 122 mg/m³ with 92 mg/m³ for the declared values (a factor of x³ in spread).

SINTEF has reported relative standard deviation figures of the spread of test results when comparing the European method and the Norway (FFDT) method, suggesting that the Relative Standard Deviation (RDS) is somehow minor by the FFDT method. For a test at nominal heat output, sampling PM by the DIN Plus method, the RDS is 41% for the European method going down to 27% when sampling by the Norway method. A rough improvement of one third.

From the Danish ecodesign market surveillance activities in 2022 we saw poor reproducibility by the European test method, more so on the indicators of incomplete combustion than on efficiency and PM. Out of 15 stoves tested 11 failed on one or more parameters during the first unattended testing round. At a second testing round with presence of the manufacturer, we saw an improvement of reproducibility ranging from a few % for efficiency to 37% on OGC.

Increasing the number of consecutive combustion cycles from three to five will add significantly to improving reproducibility. By increasing the number of combustion cycles to five, the impact of any lucky punch cycle is balanced by the full set of five, yielding a more robust test result which is closer to the true values. It is the same basic idea behind any of the current fixed load cycle tests, the BeReal, the Blue Angel and the EU Real LIFE protocol.

Trying to aggregate the improvement potential of the recommendations is not an easy task since there is not much documentation available. But we can make some experience-based estimates which can at least indicate the combined effect of adopting these recommendations, bearing in mind that this is not science (see table below). **This is just to say that the aggregated result indicating a relative improvement potential of 31%, might as well be 25% or 35%.** Regardless, the figures demonstrate that there is a significant improvement potential to be harvested, some of it even in the form of low hanging fruits.

Initiative	Attributed relative improvement potential (%) of overall repeatability
Better definitions of the burn cycle end time in terms of basic firebed incrementation by ash build up alone	2%
Skip the option for use of CO2 as start and end criterion	3%
Skip the suction pyrometer option for flue gas temp	3%
Increase the number of consecutive burn cycles from 3 to 5	10%
Reset minimum cycle duration back to 45 minutes	3%
Remove inter-cycle break or revert to Heated Filter	3%
Including of part load testing	- *
Standardizing the test fuel	4%
Compound effect	31,3%

* The effect of including part load is not evaluated. There is likely a negative effect on reproducibility. Part loads just offset the test results, but it will have a positive effect on approximating the test results to live field operation.

6. Summary

In this report there is listed a few initiatives that can improve the reproducibility of tests of wood stoves. The initiatives can be selected and implemented individually, or the full set can be implanted which is recommended but the author. If implemented as a set, roughly 30% improvement of repeatability can be anticipated.

These improvements will improve the test, without placing undue burden on the manufacturers. Time and cost wise these improvements are rendered neutral. Implementation will benefit testing costs overall by making redundant several days of repeated tests to reach the optimum test result. As a side effect these initiatives can also increase the yield of market surveillance measurements, which in turn can streamline market surveillance measurements and reduce the costs associated.

Annexes

[Annex 1: The task description from Nordsyn](#)

[Annex 2: GNB paper on interlaboratory measurements](#)

[Annex 3: List of initiatives not considered here, mainly contributing to better approximation of the test to Real Life operating conditions](#)

About this publication

Nordsyn ecodesign – task 1 report

Jes Sig Andersen

© Nordic Energy Research 2025

<http://dx.doi.org/10.6027/NER2025-06>

Cover photo: Juho S/Unsplash

Published: September 2025

Nordic Energy Research

Nordic Energy Research is the Nordic institution for joint energy research and research-based policy development, under the Nordic Council of Ministers. www.nordicenergy.org