External Scientific Report

APPROVED: 28 November 2023
doi: 10.2903/sp.efsa.2023.EN-8482

Assessing the confidence in pest freedom gained in the past pine wood nematode surveys

Salla Hannunen[1], Juha Tuomola[1], Mariela Marinova-Todorova[1], Niklas Björklund[2], Johanna, Boberg[2], Aino-Maija Alanko[1], Daniel Flø[3], Birger Ilau[4], Mart Kinkar[4], Gintaras Lapinskas[5], Adas Vaicekauskas[5], Loreta Valatkevičienė[5], Liisa Vihervuori[1] and Micael Wendell[3]
[1]Finnish Food Authority, [2]Swedish University of Agricultural Sciences, [3]Norwegian Scientific Committee for Food and Environment, [4]Estonian Agriculture and Food Board, [5]State Plant Service under the Ministry of Agriculture of the Republic of Lithuania

Abstract

All Member States of the European Union must conduct annual surveys of pine wood nematode (PWN) to ensure its timely detection. However, the statistical confidence of these surveys is rarely assessed. To facilitate such assessments, we developed two easy-to-use web applications: NoBaSURV-PWN for assessment of the statistical confidence of past PWN surveys, and NoBa Land Cover Retriever for retrieving the land cover data needed in the assessments. This report explains how the statistical confidence of past PWN surveys can be assessed with NoBaSURV-PWN. In addition, the report presents the assessments done with the NoBaSURV-PWN application for Estonia, Finland, Lithuania, Norway, and Sweden. The technical details of the developed applications are presented, and some matters that have general relevance for statistical assessment and planning of quarantine pest surveys are discussed. Also, the capacity building activities done in the project are described and their impact is evaluated. The assessments for the five Nordic-Baltic countries show that, in most of the countries, PWN surveys have been extensive enough to provide evidence for facilitating trade with a rather high confidence. Yet, the surveys have clearly not been extensive enough to ensure detection of PWN invasions at such an early stage that they could be eradicated.

Key words: survey, statistical confidence, quarantine pest, priority pest, pine wood nematode, capacity building

Question number: EFSA-Q-2023-00752
Correspondence: salla.hannunen@foodauthority.fi (Coordinator)

Disclaimer: The present document has been produced and adopted by the bodies identified above as author(s). In accordance with Article 36 of Regulation (EC) No 178/2002, this task has been carried out exclusively by the author(s) in the context of a grant agreement between the European Food Safety Authority and the author(s). The present document is published complying with the transparency principle to which the Authority is subject. It cannot be considered as an output adopted by the Authority. The European Food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors.

Acknowledgements: We would like to thank the plant health risk managers of Estonia, Finland, Lithuania, Norway, and Sweden for defining the design prevalences used in the assessments, and the EFSA plant health monitoring, assessment, partnership and procurement teams for supporting the project, namely by José Cortiñas Abrahantes, Tomasz Kaluzski, Olaf Mosbach-Schulz and Sybren Vos.

Suggested citation: Hannunen S, Tuomola J, Marinova-Todorova M, Björklund N, Boberg J, Alanko A-M, Flø D, Ilau B, Kinkar M, Lapinskas G, Vaicekauskas A, Valatkevičienė L, Vihervuori L and Wendell M, 2023. Assessing the confidence in pest freedom gained in the past pine wood nematode surveys. EFSA supporting publication 2023:EN-8482. 68 pp. doi:10.2903/sp.efsa.2023.EN-8482

ISSN: 2397-8325

Table of contents

1 Introduction

Pine wood nematode (PWN, *Bursaphelenchus xylophilus*) is the causal agent of pine wilt disease, which, under suitable conditions, can lead to mass mortality of susceptible pine trees (e.g., Futai 2013). In the European Union (EU), PWN is a quarantine pest, whose introduction into and spread within the Union is prohibited (EU 2016). Furthermore, PWN is classified as a priority pest which means that all Member States must conduct surveys annually to ensure, as far as possible, its timely detection, with a high degree of confidence (EU 2016, European Commission 2019).

Although PWN has been surveyed annually in the EU Member States since the year 2000, the confidence in pest freedom achieved in the surveys is rarely assessed (but see Hannunen and Tuomola 2020a). Consequently, it is not known if the surveys achieve the required high degree of confidence of timely detection. Furthermore, since the confidence in pest freedom accumulated in the past surveys is not utilised, the single year surveys must start from zero every year, which can make achieving the required high degree of confidence very resource consuming (Hannunen and Tuomola 2020a).

To facilitate easy assessment of statistical confidence of PWN surveys done in the past, we developed two web applications specifically for that purpose. The applications are called NoBaSURV-PWN (Hannunen and Tuomola 2023) and NoBa Land Cover Retriever (NoBa LCR, Tuomola et al. 2023b). The first is tailored for assessing the statistical confidence of past PWN surveys, and the latter for retrieving the land cover data needed in the assessments from the CORINE Land Cover (CLC) database (EEA 2022).

The main purpose of this report is to explain how to assess the statistical confidence of PWN surveys done in the past using the NoBaSURV-PWN application (section 2) and to present the assessment methodologies and results for five Nordic-Baltic countries, namely Estonia, Finland, Lithuania, Norway, and Sweden (section 3). In addition, we present the technical details of the developed applications (Annexes A and B), discuss some insights we gained in this project that have general relevance for statistical assessment and planning of quarantine pest surveys (section 4), as well as describe the capacity building activities done in the project and assess the impact of the project (Annex E).

Information on the biology of PWN and guidance on its surveillance is available in, e.g., the EPPO standard PM9/1(6) (EPPO 2018) and EFSA's pest survey card on *B. xylophilus* (EFSA 2020a). General guidelines on statistically sound and risk-based surveys of plant pests are available, e.g., in EFSA (2020b).

1.1 Grant details

This grant was awarded by EFSA to: Finnish Food Authority as the leader of the consortium of project partners from Estonian Agriculture and Food Board, the Norwegian Scientific Committee for Food and Environment, State Plant Service under the Ministry of Agriculture of the Republic of Lithuania, and Swedish University of Agricultural Sciences.

Beneficiary: Finnish Food Authority

Grant title: Assessing the confidence in pest freedom gained in the past pine wood nematode surveys

Grant number: GP/EFSA/ENCO/2020/03 - Partnering grants

2 NoBaSURV-PWN application

NoBaSURV-PWN is a web application tailored for assessing the statistical confidence of the official detection surveys of PWN done in the past. It can be used to assess 1) the confidence of each year's survey separately, and 2) the confidence accumulated in all years' surveys. In this report, the first will be referred to as "the sensitivity of annual surveys", and the latter as "the probability of pest freedom after the last annual survey".

In NoBaSURV-PWN, a survey is assumed to be composed of inspections that each target an area occupied by PWN host plants (i.e., epidemiological units), and in each inspection, one or more wood and/or *Monochamus* samples (i.e., samples of inspection units) are collected. (For a full explanation of epidemiological units and inspection units, see the glossary.)

The country, whose PWN surveys are assessed, is divided into smaller regions, that may be, e.g., counties or communes, and the statistical confidence of surveys is assessed separately for each region, and the entire country. Ideally, the regions should be such that the probability of PWN infestation is homogenous within each region, or if the risk based-survey design option 2 is used, within the risk areas of each region and within the baseline areas of each region (see section 2.4). Also, the inspection sites (i.e., places where wood is sampled or *Monochamus* beetles are trapped) should have been selected randomly within each region, or if the risk based-survey design option 2 is used, within the risk areas of each region and within the baseline areas of each region.

NoBaSURV-PWN is available at http://nobasurv-pwn.rahtiapp.fi/ and its source code has been published at EFSA's Knowledge Junction repository (https://zenodo.org/record/7766617) under the GNU General Public License version 3 (https://www.gnu.org/licenses/gpl-3.0.html).

This section gives instructions on how to set the parameter values and how to prepare the data needed to run an assessment with NoBaSURV-PWN. Technical details of the app are presented in Annex A.

2.1 Components of the survey

Components of a survey are survey entities that differ in the target population (e.g., host plant species), the inspection unit (e.g., host plants vs. vectors), or the detection method (e.g., visual examination vs. laboratory testing). In NoBaSURV-PWN, survey entities for which different design prevalences (see section 2.3) or method sensitivities (see section 2.5) are used, must be treated as separate survey components.

In NoBaSURV-PWN, the survey can have, at maximum, the following three components

1) wood sampling 1,
2) wood sampling 2, and
3) trapping of *Monochamus* adults.

Wood sampling should be divided into components 1 and 2 if sampling has targeted wood materials that clearly differ in their probability of harbouring PWN, e.g., if samples have been collected both from healthy wood that is unsuitable for *Monochamus* breeding and from wood that has signs of *Monochamus* breeding. However, to be able to use two wood sampling components, the data on survey activities (see section 2.8) must be available separately for both components.

2.2 The aim of the survey

NoBaSURV-PWN is tailored for assessing the statistical confidence of detection surveys, i.e., surveys conducted to determine the presence or absence of a pest species (IPPC Secretariat 2022). The more specific aim of the survey is defined by first, selecting the aim of the survey (see below) and then, by setting the design prevalences of the survey (section 2.3).

In NoBaSURV-PWN, the aim of the surveys can be

1) to provide evidence of pest freedom for trading partners to justify import requirements related to PWN and to facilitate export to countries with corresponding requirements. These are referred to as **trade facilitation surveys**.
2) to detect possible PWN invasions early enough to enable successful eradication. These are referred to as **early detection surveys**.

For trade facilitation surveys, it is assumed that PWN infestation would be randomly distributed throughout the country, while for early detection surveys, PWN infestation is assumed to be restricted to one region.

2.3 Design prevalences

Proving that a pest is absent from an area is not possible unless all material that can harbour the pest (e.g., all host plants and vector individuals) is tested with a perfect test. Therefore, we need to set design prevalence, i.e., the prevalence that the survey is designed to detect in the event that the pest is present in the survey area.

In NoBaSURV-PWN, design prevalences must be set for two spatial levels, namely inspection site level and areal level. Areal level design prevalence is set differently for pest freedom and early detection surveys. For the trade facilitation survey, the user sets the country level design prevalence (see section 2.3.3), and for the early detection survey, the maximum acceptable area of a PWN infestation at detection (see section 2.3.4). From these, the app calculates areal level design prevalences separately for each region.

Setting design prevalences is a risk management decision and therefore, the design prevalences of official quarantine pest surveys should be decided by the risk managers. Preferably, design prevalences should be set before the surveys are done. Yet, if that has not been done, design prevalences must be set a posteriori to enable an assessment of the probability of pest freedom achieved with surveys done in the past.

2.3.1 Biologically plausible design prevalences

All design prevalences should be biologically plausible. This means that they should be such that PWN could reach them if it was present in the area.

Biological plausibility of design prevalences can be best ensured by using information about observed prevalences of the same or similar (enough) species in similar (enough) conditions. The idea is that the observed prevalence of the same, or similar proxy species (under similar conditions), can be used to define the upper limit of a biologically plausible design prevalence.

In areas where PWN is not expected to cause symptoms, *Bursaphelenchus mucronatus* can be a suitable proxy species for setting biologically plausible design prevalences for PWN. This is because, it is both closely related and ecologically similar to PWN, e.g., it lives in the same host species, does not cause symptoms, and is vectored by the same *Monochamus* species as PWN. Furthermore, in several countries, the presence of *B. mucronatus* in the samples collected in the PWN surveys is analysed and thus, its prevalence in the survey components can be estimated (see Annex C).

2.3.2 Inspection site level design prevalences

Inspection site level design prevalences set the minimum prevalence of PWN in inspection units that an inspection is aiming to detect. In NoBaSURV-PWN, the inspection site level design prevalences must be set separately for each component of the survey. This is because, if PWN was present, its prevalence in the different components would likely differ, and ideally, the design prevalences of the different components should correspond to a similar PWN infestation level.

For either survey aim (i.e., trade facilitation and early detection survey), inspection site level design prevalences should not be higher than the prevalence that an established PWN population would likely have in the considered conditions (Figure 1). In areas where *B. mucronatus* is considered a suitable proxy for defining the design prevalences, inspection site level design prevalences should not be higher than the observed prevalences of *B. mucronatus*.

For the early detection survey, inspection site level design prevalences should preferably correspond to a PWN population that is still growing (Figure 1). In areas where *B. mucronatus* is considered a suitable proxy for defining the design prevalences, inspection site level design prevalences should preferably be lower than the observed prevalences of *B. mucronatus*.

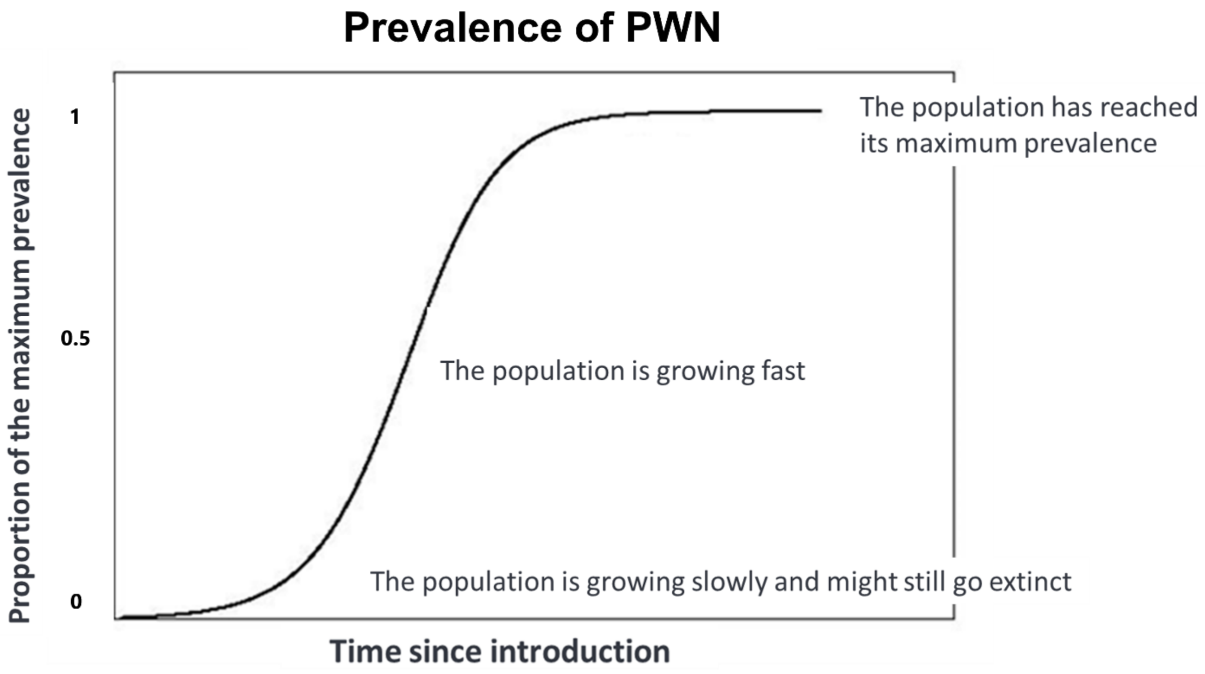

Figure 1: A schematic presentation of how the prevalence of PWN is expected to evolve over time after it is introduced to an area.

Sometimes inspection site level design prevalence can be set using information about the prevalence of *B. mucronatus* in the samples collected in the PWN surveys whose confidence is being assessed. In such cases, inspection site level design prevalence may be set relative to the apparent prevalence of *B. mucronatus*, without knowledge of its true prevalence. In this approach, if the design prevalence is set, e.g., equal to the apparent prevalence of *B. mucronatus*, it indicates that the aim is to detect a PWN prevalence that is equal to the prevalence of *B. mucronatus*. Similarly, if it is set equal to, e.g., 50% of the apparent prevalence of *B. mucronatus*, it indicates that the aim is to detect a PWN prevalence that is equal to 50% the prevalence of *B. mucronatus*. This approach is appropriate if method sensitivity and specificity can be assumed to be the same for both species, which is possible if both species have been analysed from the same samples with the same methods. In such a case, a given apparent prevalence is likely to result in a similar true prevalence (and vice versa) for both species.

If inspection site level design prevalences are set based on other prevalence estimates of a proxy species, a couple of issues should be considered.

1) Is the used prevalence estimate true or apparent prevalence? If the method sensitivity is not reported and accounted for in the study, the prevalence reported is apparent prevalence. In such cases, method sensitivity and specificity should be estimated to get an estimate of the true prevalence of the proxy species which can then be used for setting the design prevalence.

2) Is the used prevalence estimate from the same type of material that was sampled in the considered survey component? For example, if logging residues were sampled, prevalence estimates in trees cannot be used as such to set the design prevalence.
3) Is the used prevalence estimate from an area where control measures may have affected the prevalence of the (proxy) species? If yes, the prevalence of the pest, in uncontrolled conditions, may rise higher than the used prevalence estimate.
4) Are the samples in the survey whose statistical confidence is being assessed composed of material collected from several inspection units (e.g., wood from several wood objects)? If yes, either a) the design prevalence should be set based on the prevalence in the samples, not the prevalence in inspection units, or b) the number of samples should be adjusted so that each inspection unit from which material was collected is considered as a separate sample. An estimate of the prevalence in inspection units can be converted to an estimate of the prevalence in the samples as

   $$ps = 1-(1-pi)^n \quad (1)$$

   where *pi* = the prevalence in inspection units, *n* = the number of inspection units from which material was taken to one sample.

2.3.3 Country level design prevalence

For trade facilitation surveys, the user must set the country level design prevalence, which sets the minimum proportion of infested area (in which the pest prevalence is at least at the inspection site level design prevalence) of the total area with PWN host plants that the survey is aiming to detect. It can be based on requirements of trading partners, political considerations, such as national priorities, the biology of the pest, e.g., species specific Allee effects, and availability of resources. Therefore, the choice of country level design prevalence should preferably be carefully considered for each individual pest.

According to the Commission implementing regulation (EU) 2020/1231 (European Commission 2020a) 1% is typically used as design prevalence for detection surveys. However, unlike NoBaSURV-PWN, the regulation does not differentiate between detection surveys carried out to provide evidence for trading partners (trade facilitation surveys) and those done to facilitate eradication of invasions (early detection surveys).

2.3.4 Maximum acceptable area of PWN infestation at detection

For early detection surveys, the user must set the maximum acceptable area of PWN infestation at detection in square kilometres, from which the app calculates the areal level design prevalence for each region separately. These set the minimum proportion of infested area (in which the pest prevalence is at least at the inspection site level design prevalence) of the total area with PWN host plants that the survey is aiming to detect.

The maximum acceptable area of PWN infestation at detection should represent an infestation that can still be eradicated. It is up to the risk managers that are responsible for the surveys to decide that size since the larger the area is the less likely it is that an eradication attempt will be successful. When defining the size, at least, the availability of labour and machinery, accessibility of terrain and the density of host trees should be considered.

The maximum area from which eradication could, in practise, be attempted can be used as a reference value to set the upper limit for the maximum acceptable area of infestation at detection. It is limited by the financial and physical resources available for the delimiting surveys and other eradication measures (e.g., the availability of personnel and machinery). For Finland, it has been estimated that harvesters would be available for taking measures on an infestation of, at maximum, 598 km^2 between July 15 and May 1 (Hannunen and Tuomola unpublished). For Norway, it has been estimated that the area of forest that could be removed between August 1 and May 1 is 426 km^2 assuming that all the harvesters in the country were available for the measures (Økland et al. 2010). Note that the time and resources needed for delimiting the infestation were not considered in either estimate.

EU emergency measures for PWN (European Commission 2012) require that Member States must attempt eradication if the diameter of the infested area is less than 20 km. If such an infested area was circular, it would be, at maximum, 314 km^2. However, in practise, it would be smaller since not all of it would contain suitable host plants for PWN. Again, although Member States are required to attempt eradication from such a large area, the likelihood of successful eradication depends on the local conditions and the resources available for the measures.

2.4 Risk-based survey design options

A risk factor is a biotic or abiotic factor that increases the probability of infestation by the pest in the area of interest. A risk factor must be such that the target population can be divided into, at least, two risk factor levels, that can be characterised by their relative risk (EFSA 2020). (Note that the severity of impact of an infestation is not considered in "relative risk". Instead, it is based only on the probability of PWN infestation.)

In NoBaSURV-PWN, the risk factor considered is human activity related to international trade that increases the probability of PWN entry to the country, such as logistic centres, harbours etc., and the target population can be divided into risk factor levels, which differ in their relative risk, in two different ways called risk-based survey design options 1 and 2 (Table 1).

Table 1: The options for risk-based survey design available in NoBaSURV-PWN.

	Risk-based survey design option 1	Risk-based surveys design option 2
Risk factor	Human activity related to international trade that increases the probability of PWN entry to the country	Human activity related to international trade that increases the probability of PWN entry to the country
Risk factor levels	Regions	Risk area (close to entry sites) and baseline area (further away from entry sites)
Relative risk of the risk factor levels	Defined based on the proportion of the area or number of entry sites in the region of the total area or number of entry sites in the country	Defined based on the distance from the closest entry site using 1) the proportion of PWN population in risk vs. baseline areas after one flight season of vectors predicted by the '*2Dt*' dispersal location kernel parametrised for *Monochamus galloprovincialis* with the estimates provided by Etxebeste et al. (2016), 2) the proportion of PWN population in risk vs. baseline areas after one or more flight seasons of vectors predicted by some other spread/population dynamics model, or 3) some other process identified by the user, meaning that the user must manually define the relative risk to be used.

In risk-based survey design option 1, regions are assumed to differ in relative risk, and their relative risk is assumed to depend on the relative area of entry sites or the relative number of entry sites (i.e., sites where the probability of pest entry is elevated) in the region. For details on the data needed for this option, see section 2.8.

In risk-based survey design option 2, areas with target population that are within a user defined radius from entry sites are considered "risk areas", and other areas with target population are considered "baseline areas". This option can be used

- if the radius of risk areas was defined before the surveys were initiated and data on the number of inspected sites and the number of samples is available separately for risk and baseline areas, or
- if the GPS coordinates of the inspected sites are available and thus, the number of inspected sites and the number of samples in risk vs. baseline areas can be calculated for any radius of risk areas.

Note that if the radius of risk areas is defined retrospectively, it should be such that the inspected sites provide a representative sample of the risk factor levels in which inspections were done. At minimum, this means that if all the inspections were done within, e.g., 2 km radius from entry sites, the radius of risk areas should not be much more than 2 km.

In risk-based survey design option 2, the user can select one of three alternative ways to define the relative risk of risk vs. baseline areas. The first two assume that the probability of PWN infestation at different distances form entry sites depends on the dispersal behaviour of PWN vectors. The third allows defining the relative risk based on whichever process identified by the user.

The first alternative is to define the relative risk of risk vs. baseline areas based on the proportion of PWN population in risk vs. baseline areas after one flight season of its vectors predicted by the '*2Dt*' dispersal location kernel parameterised for *Monochamus galloprovincialis* with the estimates provided by Etxebeste et al. (2016). In this alternative, the user defines the radius of risk areas, and the app calculates the rest.

The second alternative is to define the relative risk of risk vs. baseline areas based on the proportion of PWN population in risk vs. baseline areas after one or more flight seasons of its vectors predicted by a spread/population dynamics model selected by the user. In this alternative, the user must obtain the predicted proportion of the PWN population in risk areas her-/himself and input that to the app.

If the relative risks are defined based on the distribution of the pest population after one flight season of vectors, only a suitable dispersal kernel is needed. Yet, if more than one flight seasons are considered, a model that covers both PWN population dynamics and the annual probability of new PWN invasions is needed. Ideally, the number of flight seasons considered should be linked to the duration of the invasion threat of the pest. In other words, if the pest has been posing a threat to the considered area for, e.g., 20 years, the number of flight seasons considered should be 20.

When using dispersal kernels to calculate the predicted proportion of PWN population within different radiuses from entry sites, it is important to recognise if the used kernel is a 1- or 2-dimensional dispersal location kernel or a dispersal distance kernel, which are always 1-dimensional. (For more information see, e.g., Nathan et al. 2012). The proportion of the population within different radiuses can be obtained from 1-D kernels with 1-D integration, while if a 2-D kernel is used, 2-D integration is needed. Alternatively, 2-D kernels can be converted into 1-D dispersal distance kernels by multiplying with $2\pi r$. 1-D dispersal kernels can be integrated, e.g., with the integrate function of the 'stats' package of R (R Core Team 2022).

The third alternative is to define the relative risk of risk vs. baseline areas based on some other user defined factor. In this alternative, the user simply defines the relative risk in the risk areas, per unit area (e.g., per km^2). Note that the size of the risk areas should not be considered when defining their relative risk since the area is accounted for by the app. The relative risk in baseline areas is set to 1, and thus the relative risk in risk areas should be greater than 1.

2.5 Method sensitivity

Method sensitivity is the probability that a sample that contains material from one or more infested inspection units tests positive. It is the product of sampling effectiveness and diagnostic sensitivity.

Sampling effectiveness is the probability of selecting infested parts from an infested sampling unit(s). For the wood sampling components, it depends on the distribution and abundance of PWN in trees and other wood objects suitable for sampling, and on the distribution and number of points from which wood material is taken to the sample.

In NoBaSURV-PWN, sampling effectiveness for the *Monochamus* trapping component can be assumed to be close to one. This is because in NoBaSURV-PWN, the inspection unit is *Monochamus* adult, and normally whole beetles are taken as samples. [Note that this approach differs from that in EFSA (2020a), where inspection unit is the area covered by one trap where at least one host plant is present.] However, the used trapping procedure may affect sampling effectiveness if the conditions are such that the nematodes may disembark their vectors.

Diagnostic sensitivity is the probability that a truly positive sample will test positive in the laboratory analysis. For PWN, it depends on the sensitivity of the extraction and identification methods, estimates for which may be found from diagnostic standards or other literature (see, e.g., EPPO 2013). Note that if the sampling technique, transport, or storage conditions of the samples are not ideal, they may also affect diagnostic sensitivity.

If inspection site level design prevalence is set relative to the prevalence of *B. mucronatus* observed in the samples analysed in the PWN surveys whose statistical confidence is being assessed (see 2.3.2), setting the value for the field “method sensitivity” differs from the normal case. In this special case, 1 should be inserted in that field (since that appropriately represents the case in which the inspection site level design prevalence is set relative to the prevalence of a proxy species observed in the same samples.)

2.6 Initial prior probability of freedom

Initial prior probability of freedom is the probability that the pest was absent from the country before the first annual survey, or to be exact the probability that defines the confidence with which the prevalence of the pest is expected to be below the design prevalence(s) before the first survey.

Initial prior probability of freedom can be estimated, e.g., with expert knowledge elicitation (for detailed guidance, see EFSA 2014). In the estimation, different sources of information may be considered, such as data on interceptions and trade volumes and, if PWN is expected to cause symptoms in the considered area, general surveillance of forest pests.

If no information is available about the presence/absence of PWN before the surveys were started, the initial prior probability of freedom can be set to 0.5 which indicates that PWN is as likely to be present as it is to be absent. Note that although this is often called an uninformative prior, it means that the calculations will be based on the assumption that the probability of absence and presence is 0.5. The choice of value for the initial prior probability of freedom can affect the probability of pest freedom for several years, especially if the sensitivity of annual surveys is low (for further discussion, see section 4.1).

The initial prior probability of freedom should be in line with the mean time between invasions (section 2.7). If the mean time between invasions is assumed to have been

Republic of Estonia Agriculture and Food Board; Finnish Food Authority; SLU Swedish University of Agricultural Sciences; VKM Vitenskapskomiteen for mat og miljø

constant, the consistency between the prior probability of freedom and the mean time between invasions can be ensured by defining the prior probability of freedom from the Poisson distribution as

$$PriorPfree = e^{-\frac{y}{Tinv}} \tag{2}$$

where *y* = the number of years during which invasions of the pest are considered to have been possible before the first survey, and *Tinv* = the mean time between invasions (years) during this time.

2.7 Mean time between invasions

In NoBaSURV-PWN, mean time between invasions is the mean time between events in which PWN enters the country, is transferred to a suitable host plant, and manages to infest it.

Mean time between invasions (*Tinv*) can be obtained from the annual probability of invasion (*Pinv*) as

$$Tinv = -\frac{1}{\ln(1 - Pinv)} \tag{3}$$

If estimates about the mean time between invasions (or the probability of invasion) are not available, a wide range of values can be considered. Still, the studied range should be in line with the initial prior probability of freedom, unless reason exists to assume that the probability of invasion was different before the surveys were initiated (see section 2.6).

2.8 Data needed in the assessment

Table 2: A summary of the data needed for the assessments using the two risk-based survey design options available in NoBaSURV-PWN. All data is needed separately for each survey year and administrative region considered, and separate files are needed for all components of the survey.

	Risk-based survey design option 1	**Risk-based survey design option 2**
Data on survey activities	• The number of inspected sites • The number of samples	• The number of inspected sites in risk areas • The number of inspected sites in baseline areas • The number of samples in risk areas • The number of samples in baseline areas
Data on landcover	• The area (km^2) or number of entry sites • The area of the target population (km^2)	• The area (km^2) or number of entry sites • The area of the target population in risk areas (km^2) • The area of the target population in baseline areas (km^2)

2.8.1 The number of inspected sites

For the wood sampling components, the number of inspected sites is the number of sites from which wood samples were collected. Note that only data on sampling of domestic wood should be included. For the *Monochamus* trapping component, the number of inspected sites is the number of trapping sites.

Data on the number of inspected sites must be provided in separate csv files for each component of the survey, and if the risk-based survey design option 2 is used, data must be provided in separate csv files for risk and baseline areas. Detailed instructions on how to prepare the csv files are given in section 2.8.5.

2.8.2 The number of samples

For the wood sampling components, the number of samples is the number of wood samples analysed per inspected site or per region depending on the availability of information. Note that only data on sampling of domestic wood should be included. For the *Monochamus* trapping component, the number of samples is the number of *Monochamus* adults analysed per region.

If the number of wood samples analysed per inspected site is always the same, it is enough to specify it once in the app in the field "Fixed number per inspected site". If the number of wood samples varies between sites, the total number of samples per region must be provided in a csv file. The number of *Monochamus* adults per region must always be provided in a csv file.

Data on the number of samples must be provided in separate csv files for each component of the survey (except if the number of wood samples per inspected site is always the same). In addition, if the risk-based survey design option 2 is used, data must be provided in separate csv files for risk and baseline areas (except if the number of wood samples per inspected site is always the same). Detailed instructions on how to prepare the csv files are given in section 2.8.5.

2.8.3 The area of entry sites or number or of entry sites

In NoBaSURV-PWN, entry sites refer to sites in which the probability of PWN entry to the country is assumed to be elevated, such as ports and industrial areas (but the types of sites that are used as entry sites may vary between countries).

Depending on the definition of entry sites and availability of data, either the area of entry sites or the number of entry sites can be used. For most types of entry sites, area is the more appropriate measure. Yet, if, e.g., sawmills and logistics centres are considered as entry sites, data on their area may be unavailable, in which case the number of entry sites can be used instead.

For many types of PWN entry sites, e.g., ports and industrial areas, data can be obtained from the CORINE Land Cover (CLC) database (EEA 2022). For Estonia, Finland, Lithuania, Norway, and Sweden CLC data can be easily retrieved with NoBa Land Cover Retriever (NoBa LCR, Tuomola et al. 2023b), which is a web application specifically tailored for this purpose (for more information on NoBa LCR, see Annex B).

Data on the area or number of entry sites must be provided in a csv file. If area is used, it must be provided in square kilometres. Detailed instructions on how to prepare the csv file are given in section 2.8.5.

2.8.4 The area of the target population

In NoBaSURV-PWN, the area of the target population refers to the area with PWN host plants in the country or regions to which the results of the survey will be generalised to.

The area of the target population can be obtained from the CORINE Land Cover (CLC) database (EEA 2022) or from other sources. For Estonia, Finland, Lithuania, Norway, and Sweden, CLC data can be easily retrieved with NoBa LCR (Tuomola et al. 2023b), which is a web application specifically tailored for this purpose (for more information on NoBa LCR, see Annex B).

Data on the area of the target population must be provided in square kilometres. If the risk-based survey design option 2 is used, data must be provided in separate csv files for risk and baseline areas. Detailed instructions on how to prepare the csv files are given in section 2.8.5.

2.8.5 Preparing the files to be uploaded

Data must be uploaded to NoBaSURV-PWN as comma separated csv files in which

- Data for regions is in columns, and data for years is in rows.
- Data is provided for at least two regions and two years.
- The first row has the names of the regions, and the first column indicates the years covered.
- The names of the regions are written without special characters.
- The regions are in the same order in all the files.
- The years are in ascending order in all the files.
- Every year between the first and the last is included in all the files, even if the survey was not done in all years.
- When the number of inspected sites or the number of samples is zero, that is indicated with 0.
- Data is given separately for all years, even if the area of entry sites or target population was the same for some (or all) of the years.
- Point is used as a decimal separator.

The files can be compiled with Excel (Figure 2), but they must be saved as comma separated csv files (Figure 3). Note that in Excel, there are several options for saving csv files, of which the option “CSV (comma delimited)” should be used.

	A	B	C	D	E	F	G	H
1	Year	Uusimaa	Varsinais-Suomi	Satakunta	Hame	Pirkanmaa	Kaakkois-Suomi	Etela-Savo
2	2000	100	75	75	50	50	100	50
3	2001	4	3	0	0	0	8	1
4	2002	5	3	0	0	0	0	0
5	2003	37	30	50	0	0	72	0
6	2004	57	50	50	0	0	98	0
7	2005	61	50	50	0	0	95	0
8	2006	65	51	40	0	0	27	0
9	2007	55	29	46	25	25	23	54
10	2008	52	45	50	25	25	33	52
11	2009	45	50	50	25	25	40	53
12	2010	45	45	42	19	20	36	46
13	2011	45	57	43	18	20	40	46

Figure 2: An example of how the data to be saved as a csv file should look like when compiled in a Microsoft® Excel© spreadsheet.

```
Year,Uusimaa,Varsinais-Suomi,Satakunta,Hame,Pirkanmaa,Kaakkois-Suomi,Etela-Savo
2000,100,75,75,50,50,100,50
2001,4,3,0,0,0,8,1
2002,5,3,0,0,0,0,0
2003,37,30,50,0,0,72,0
2004,57,50,50,0,0,98,0
2005,61,50,50,0,0,95,0
2006,65,51,40,0,0,27,0
2007,55,29,46,25,25,23,54
2008,52,45,50,25,25,33,52
2009,45,50,50,25,25,40,53
2010,45,45,42,19,20,36,46
2011,45,57,43,18,20,40,46
```

Figure 3: An example of how the csv files to be uploaded should look like when viewed with, e.g., Microsoft® Notepad©.

2.9 Interpretation of the results

NoBaSURV-PWN calculates

- the statistical confidence of each annual survey separately, and
- the statistical confidence accumulated in all the annual surveys.

The first is measured as sensitivity (which is a synonym to confidence level) of annual surveys and the latter, as probability of pest freedom after the last annual survey.

The exact definitions of sensitivity and probability of freedom differ but for practical purposes, it is enough to known that they both indicate the statistical confidence level of the surveys. In other words, they indicate, how certain we can be that the pest's prevalence is lower than the considered design prevalence, if it was not found in the survey(s). The exact probabilistic definitions of the terms are given below in 2.9.1 and 2.9.2.

2.9.1 The sensitivity of annual surveys

The sensitivity of annual survey is the probability with which PWN would have been detected in the area (that may be the entire country or a single region) in a given year if it was present there

- in an area that corresponds to the areal level design prevalence
- in a proportion of inspection units that is equal to the inspection site level design prevalence of the respective survey component.

If the pest is present at lower prevalence(s) than the design prevalence(s), the probability of detecting it is lower than the sensitivity of the survey. And vice versa, if the pest is present at higher prevalence(s) than the design prevalence(s), the probability of detecting it is higher.

2.9.2 The probability of pest freedom after the last annual survey

The probability of pest freedom is the probability that, if PWN was not detected in any of its annual surveys, it is not present in the area (that may be the entire country or a single region)

- in a proportion of the area that is larger than the areal level design prevalence
- in a proportion of inspection units that is greater than the inspection site level design prevalence of the respective survey component.

In NoBaSURV-PWN, the probability of pest freedom after the last annual survey is plotted against the considered rage of mean time between PWN invasions. The probability of pest freedom always increases with the mean time between invasions but, if a wide enough range of mean times between invasions is considered, the probability tends to level out at some point. Therefore, if the mean time between invasions is not known, it may be enough to consider whether it is likely to be below or above this levelling-out point.

The probability of pest freedom after the last annual survey must be interpreted with caution! This is because the initial prior probability of freedom can have a major impact on the probability of pest freedom even if a seemingly uninformative initial prior probability of freedom (0.5) is used. For further discussion, see section 4.1.

3 The confidence in pest freedom gained in past PWN surveys in five Nordic-Baltic countries

The confidence in pest freedom gained in the PWN surveys conducted in Estonia, Finland, Lithuania, Norway and Sweden was assessed using country specific parameter values and data described below.

3.1 The PWN survey in the five Nordic-Baltic countries

The PWN surveys in the Nordic-Baltic countries were conducted based on the EPPO standard PM 9/1 (EPPO 2018), EU pine wood nematode survey protocol (European Commission 2009) and country specific survey guidelines from respective national plant protection organisation.

Since the PWN is not expected to cause symptoms in host trees in the Nordic-Baltic region (Gruffudd et al. 2016, Tuomola et al. 2021), the surveys have mainly focused on sampling of wood objects suitable for breeding of the vector, i.e., *Monochamus* spp., and sampling of adult *Monochamus* spp. Samples were taken from standing trees, wood or logging residues with signs of *Monochamus* activity and also in some countries from trees that were dead or dying for unknown reasons and from asymptomatic trees. In the assessment, inspection unit of the wood sampling component(s) was defined as “the type of wood objects from which wood samples were collected” in the respective country (e.g., fallen branches, logging residuals). Inspection unit was not “a tree”, because the type of material from which samples were collected would be infested via oviposition of vector beetles only after it has been detached form the tree. (Which means that if PWN was found in such a wood object, it would not indicate that the tree from which it originates was infested.)

Wood samples were obtained from wood of host trees, mainly from *Pinus sylvestris* and *Picea abies*. These tree species constitute a major part of coniferous forests in the region. Adults of *Monochamus* spp. were captured using traps with pheromones. All *Monochamus* spp. present in the respective countries are expected to be able to vector the PWN and were targeted in the survey.

In all countries, the sampling was focused to the areas surrounding locations with elevated likelihood of PWN entry. The samples were analysed for the presence of PWN according to EPPO diagnostic standard PM 7/4(3) (EPPO 2013), and the samples were also analysed for the presence of the native and closely related species *B. mucronatus*.

3.2 The components of the survey and the type of survey

In all countries, both wood sampling and trapping of *Monochamus* adults were included in the PWN survey. In Estonia, Finland, Norway and Sweden only wood suitable for *Monochamus* breeding (only with signs of *Monochamus* activity, or also without, depending on the country) was sampled and thus two survey components were included in the analysis. In Lithuania, also healthy wood, unsuitable for *Monochamus* breeding was sampled and thus three survey components were included in the analysis.

In all countries, the PWN survey aimed to both facilitate trade and to enable early detection of the PWN. Thus, assessments were done for both survey aims (i.e., trade facilitation and early detection) for all countries.

3.3 Design prevalences

The design prevalences were obtained from the risk managers of the respective country. The values set by the risk managers are presented in Table 3 and were justified as follows.

3.3.1 Design prevalences for trade facilitation surveys

In all countries, the inspection site level design prevalences for trade facilitation surveys were set equal to the estimated apparent prevalence of the native *Bursaphelenchus* species, *B. mucronatus*, in the sampled wood and in the trapped *Monochamus* adults in the respective country (Table 3 and Annex C). *B. mucronatus* was considered to be a suitable proxy species for setting the design prevalences since it is closely related and ecologically similar to PWN. For trade facilitation surveys, it was assessed to be acceptable to detect the pest when the population has reached its maximum expected prevalence at the inspection site, and this was considered to be represented by the prevalence of *B. mucronatus* in the samples.

In all countries, the inspection site level design prevalences were set using information about the apparent prevalences of *B. mucronatus* observed in the samples analysed in the PWN survey whose confidence is being assessed (see section 2.3.2). Since it was considered reasonable to assume that method sensitivity and specificity were the same for both species, 1 was inserted to the field "method sensitivity" for all survey components (see section 2.5).

The selected country level design prevalences for the trade facilitation survey differed between countries as follows:

- In Estonia, the country level design prevalences for the trade facilitation survey was set as high as 5% until otherwise required by trade partners. It was considered reasonable to have a clearly higher design prevalence than in the early detection survey as the aim of the trade facilitation survey is not to detect the pest early enough for successful eradication, but to provide evidence of pest freedom to trade partners.
- In Finland, the country level design prevalence for trade facilitation survey was set to 2%.
- In Lithuania, the country level design prevalence for the trade facilitation survey was set to 1% which is a common choice for the design prevalence (European Commission 2020a).
- In Norway, the country level design prevalence for trade facilitation survey was set to 1% referring to (EU) 2020/1231 (European Commission 2020a).
- In Sweden, the country level design prevalence for trade facilitation survey was set to 1%, which reflects a balance between the size of an infected area that can be accepted and the resources available to conduct the survey. A design prevalence for detection survey is typically set to a value of 1% (European Commission 2020a). This is for example implemented in the requirement for *Xylella fastidiosa* where it is stated that the survey shall be able to detect a level of presence of infected plants of 1% (European Commission 2020b).

Table 3: The design prevalences used in the assessments in the five Nordic-Baltic countries. Note that the design prevalences were set by the risk managers of the respective country solely for the purpose of this project and hence different values might be used in other cases.

Design prevalences	Estonia	Finland	Lithuania	Norway	Sweden
Trade facilitation survey					
Inspection site level design prevalences					
Wood sampling 1	0.078	0.123	0.053	0.009	0.082
Wood sampling 2	-	-	0.015	-	-
Trapping of Monochamus	0.14	0.157	0.036	0.017	0.071
Country level design prevalence	0.05	0.02	0.01	0.01	0.01
Early detection survey					
Inspection site level design prevalences					
Wood sampling 1	0.039	0.062	0.027	0.0045	0.041
Wood sampling 2	-	-	0.008	-	-
Trapping of Monochamus	0.07	0.079	0.018	0.0085	0.036
Maximum acceptable area of infestation at the time of detection, km^2	75	30	60	426	314

3.3.2 Design prevalences for early detection surveys

In all countries, the inspection site level design prevalence for early detection surveys was set to 50% of the estimated apparent prevalences of *B. mucronatus*. This value was expected to represent a prevalence of a population that is most likely in an exponential growing phase and would therefore aim for detecting the pest population at an earlier stage of invasion. This was assumed to represent the aim of the survey (to detect PWN so early that eradication is feasible) more adequately. It would be beneficial to detect PWN before the exponential phase, but this has to be balanced by the availability of resources.

In all countries, the inspection site level design prevalences were set using information about the apparent prevalences of *B. mucronatus* observed in the samples analysed in the in the PWN survey whose confidence is being assessed (see section 2.3.2). Since it was considered reasonable to assume that method sensitivity and specificity were the same for both species, 1 was inserted to the field "method sensitivity" for all survey components (see 2.5).

The selected maximum acceptable area of infestation at the time of detection differed between countries as follows:

- In Estonia, the maximum acceptable area of PWN infestation at detection was set to 75 km^2. It was considered that the early detection survey should aim to detect PWN before it has spread to an area of the size defined in the EU PWN emergency measures (European Commission 2012) to be beyond the obligation to attempt eradication. In the EU PWN emergency measures (European Commission 2012), it is stated that Member States may decide to contain PWN, instead of eradicating it, in cases where the diameter of the infested zone exceeds 20 km. For Estonia it would mean that when the size of an infestation is larger than 78.5 km^2, the measures would allow not to attempt eradication but rather switch to containment measures [a circular area with 20 km diameter is 314 km^2, which in case of Estonia would include roughly 25% coniferous forest (approximately 50% of the whole territory of Estonia is covered with forests, and about 50% are coniferous forests)]. An area of 75 km^2 of coniferous forest should be well below the maximum area that could be removed by Estonian harvesters (logging capacity) and therefore eradication could be attempted. Although this estimation is not based on any calculations but on the comparison of estimations done by Finland and Norway (Kukkonen et al. 2011, Økland et al. 2010).
- In Finland, the maximum acceptable area of PWN infestation at detection, was set to 30 km^2. It was estimated based on the Finnish contingency plan for pine wood nematode (Kukkonen et al. 2011) and the experiences gained in the recent eradication process of *Anoplophora glabripennis* from Finland (EPPO 2021).
- In Lithuania, the maximum acceptable area of PWN infestation at detection was set to 60 km^2. The decision was made considering the population size of conifers and the forestry data on the amount of coniferous trees cut per year in Lithuania. Clear-cutting a bigger PWN infestation size would probably be unacceptable both in terms of resources and damage to environment, especially if PWN symptoms do not appear in Lithuanian climate.
- In Norway, the maximum acceptable area of PWN infestation at detection was set to 426 km^2 following the estimated capacity of harvesters available in Norway provided by Økland et al. (2010).
- In Sweden, the maximum acceptable area of PWN infestation at detection was set to 314 km^2. Many different factors will affect the size of the area from which eradication could be attempted. Eradication measures involving removal and destruction of the entire host trees at a regional scale involves enormous practical and logistic challenges. The larger the infested area the more challenging the eradication will be and less likely to be successful. For the purposes of the analysis in this project, the estimation of the maximum size is therefore based on the requirement formulated in the EU emergency measures for PWN (European Commission 2012) where it is stated that member states are allowed to refrain from attempting eradication if the diameter of the infested area is more than 20 km (= circular area of 314 km^2). Thus, this is the largest area of PWN infestation from which the Member States are required to attempt eradication (assuming that the whole area is covered with host trees).

3.4 Options for a risk-based survey design

The risk-based survey design option selected by the different countries varied due to the design of the survey and the availability of data (Table 4).

Risk-based survey design option 1 was used when assessing the PWN surveys of Finland and Lithuania as described below:

- In Finland, sampling was instructed to be done in risk areas, which were defined as pine forests within a 5 km radius from entry sites. However, because the exact coordinates of the inspection sites were not available, the confidence in pest freedom was assessed considering that the relative risk of each administrative region (i.e., each Centre for Economic Development, Transport and the Environment of Finland) depends on the area of entry sites in the region.
- In Lithuania, the probability of PWN infestation was assumed to be elevated around entry sites (i.e., in risk areas). Samples were taken from regular forests and from risk areas. The exact radius from entry sites was not defined. Coordinates of each survey site were not registered. Hence, the confidence in pest freedom was assessed considering the relative risk of each administrative region in Lithuania. The relative risk was calculated based on the number of entry sites in the region.

Risk-based survey design option 2 was used when assessing the PWN surveys of Estonia, Norway and Sweden as described below:

- In Estonian, both trapping of *Monochamus* adults and wood sampling were mainly done in risk areas. Risk areas were defined to be within a 5 km radius around entry sites. Coordinates of the *Monochamus* trapping sites were available for all the survey years, and coordinates of the wood sampling sites for 2010–2021. For 2002–2009 the proportion of wood samples in risk areas and baseline area was estimated for each administrative region based on the proportion of wood samples in risk areas and baseline areas in 2010–2021.
- In Norway, the radius of the risk areas was defined as 50 km (Sundheim et al. 2010), and all sampling was done within these risk areas.
- In Sweden, sampling was done in risk areas surrounding entry sites and the coordinates of inspection sites were available. Sampling was only done in suitable clear cuts as close as possible to entry sites, but no specific distance was defined for the risk area surrounding the entry sites. Instead, for the assessments the radius of the risk areas was defined based on empirical data and dispersal curves provided in Etxebeste et al. (2016). Based on a fitted '*2Dt*' distance dispersal kernel, 99% of dispersing *Monochamus* beetles would be found within a radius of 3.5 km. The data were collected for *M. galloprovincialis* in a landscape with continuous pine forests, which should be comparable to the conditions in Sweden and the '*2Dt*' model forecasted the widest tail representing the most conservative approach. It is, however, acknowledged that the considered radius only takes into account dispersal of one generation of beetles and that no effect of overlapping risk areas was considered.

Table 4: The selected options for risk-based survey design used in the assessments for the five Nordic-Baltic countries.

Risk-based survey design option	Estonia	Finland	Lithuania	Norway	Sweden
Option 1		x	x		
Option 2	x			x	x
In option 2, the relative risk of risk areas was based on					
The '*2Dt*' dispersal location kernel of *M. galloprovincialis* from Etxebeste et al. (2016)	x			x	x
Some other dispersal kernel of PWN vectors					
Something else					
Radius of risk areas (km)	5			50	3.5

3.5 Initial prior probability of freedom and mean time between invasions

The initial prior probability of freedom (i.e., the probability of pest freedom before the first official annual survey included in the assessment) was set to 0.5 for both the trade facilitation survey and the early detection survey in all countries. All assessments were done with mean times between invasions ranging from 2 to 50 years.

3.6 Data used in the assessments

Data from the PWN surveys conducted in the different countries were provided by Estonian Agriculture and Food Board, Finnish Food Authority, the Norwegian Food Safety Authority, State Plant Service under the Ministry of Agriculture of the Republic of Lithuania and the Swedish Board of Agriculture. All data used in the assessments are available as csv files in Zenodo in EFSA's Knowledge Junction repository (Tuomola et al. 2023a).

The availability of PWN survey data suitable for the assessments varied between countries, and the time periods covered in the assessments are presented in Table 5.

Table 5: Years included in the assessments.

Component of the survey	Estonia	Finland	Lithuania	Norway	Sweden
Wood sampling 1	2002-2021	2000-2021	2015-2021	2000-2019	2013-2021
Wood sampling 2	-	-	2015-2021	-	-
Trapping of *Monochamus*	2012-2021	2012-2021	2020-2021	2012-2019	2013-2021

In Estonia, Finland, Lithuania and Sweden each inspection site was represented by one wood sample while in Norway the number of wood samples collected per inspection site varied and the information was input to the app as a separate csv file.

Entry sites were defined differently for the PWN survey in the different countries (Table 6). The distribution and area of entry sites per region was retrieved from the CORINE Land Cover (CLC) data [version 2020_20u1 (CLC2018) in a 100 m GeoTIFF format (EEA 2022)] or from other sources (Table 6) using NoBa LCR (Annex B) except for Estonia.

In all the countries, the target population of the survey was defined as PWN host plants in the whole country. CLC data was used for the calculations, but the CLC classes selected to represent the target population differed slightly between countries (Table 6). The areas of the target population in risk areas and baseline areas were calculated for all the countries except for Estonia using the NoBa LCR application. For Estonia, the areas of entry sites and target populations were calculated using ArcGIS Pro® software.

Table 6: The definition of entry sites and target population, and the sources of land cover data used to calculate their areas or numbers.

	Estonia	Finland	Lithuania	Norway	Sweden
Entry sites					
Definition	Harbours, railways, highways, border points, wood processing companies, WPM production and storage sites	Harbours, industrial areas, landfills	Wood industry, WPM production, storage and trade facilities, border inspection posts, harbour	Ten entry sites defined (mainly wood handling facilities, such as sawmills, pulp factories and large timber yards)	Urban areas, including locations that handle goods and wood packaging material (e.g., industrial areas, harbours) and processing imported wood (e.g., pulp factories, sawmills).
Sources of data	Port areas: EEA (2022), railways and highways: Estonian Topographic Database, Land Board (2021); border points and wood processing companies: EAFB (2022, personal communication)	EEA (2022)	SPSMoA (2022, personal communication)	VKM (2022, personal communication)	EEA (2022)
CLC classes used	123: Port areas	121: Industrial or commercial units; 123: Port areas; 132: Dump sites	-	-	111: Continuous urban fabric; 112: Discontinuous urban fabric; 121: Industrial or commercial units; 122: Road and rail networks and associated land; 123: Port areas; 124: Airports; 133: Constructions sites
Target population					
Definition	Area with PWN host plants				
Sources of data	EEA (2022)	EEA (2022)	EEA (2022)	EEA (2022)	EEA (2022)
CLC classes used	312: Coniferous forests; 313: Mixed forests; 324: Transitional woodland-shrub	312: Coniferous forests; 313: Mixed forests; 324: Transitional woodland-shrub	312: Coniferous forests; 313: Mixed forests; 324: Transitional woodland-shrub	312: Coniferous forests; 313: Mixed forests; 324: Transitional woodland-shrub	312: Coniferous forests; 313: Mixed forests; 324: Transitional woodland-shrub; 334: Burnt areas

3.7 Results of the assessments

PWN was not found in any of the wood or *Monochamus* samples collected and analysed in any of the countries. The sensitivity of annual surveys varied between countries, years and survey type (Figure 4, Annex D). As expected, for all countries, the sensitivity was higher for trade facilitation surveys than for early detection surveys. The probability of pest freedom after the last annual survey is presented by country and survey type in Figure 5.

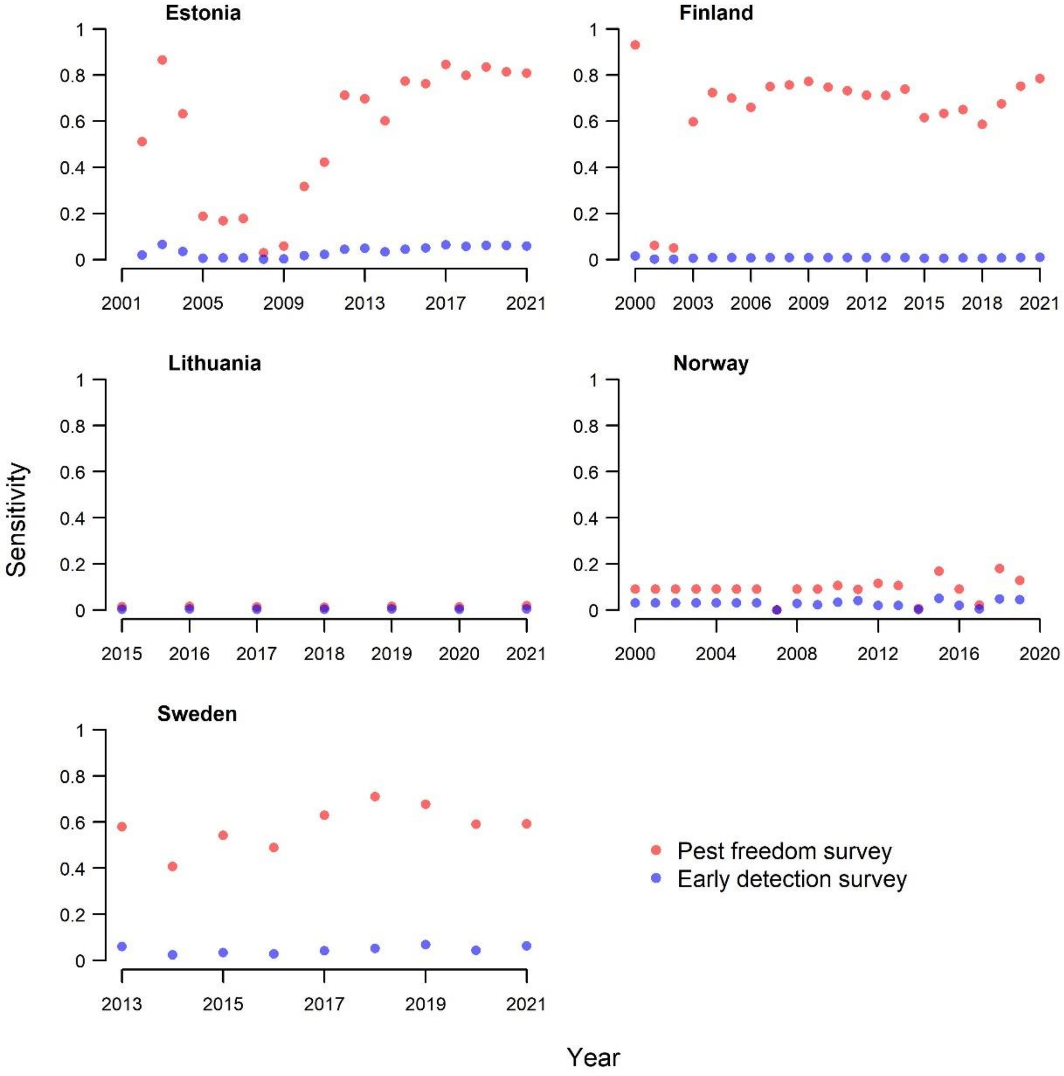

Figure 4: Sensitivity of the PWN surveys for individual years for each country and survey aim. The data is also provided in Annex D.

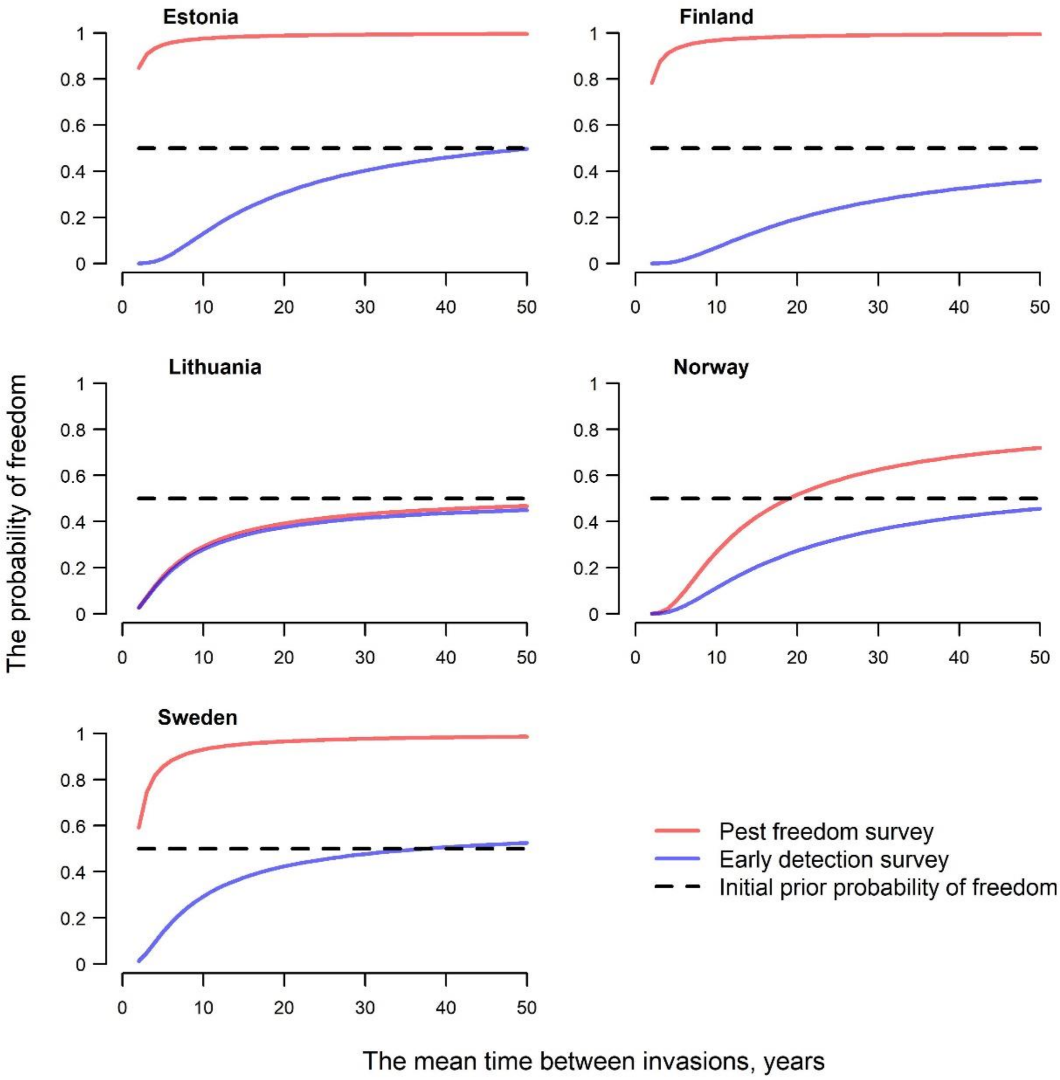

Figure 5: The probability of pest freedom after the last annual survey for a range of mean times between invasions, for the different countries and survey aims. The data is also provided in Annex D.

In Estonia, Finland and Sweden the sensitivity of the trade facilitation surveys was relatively high in many years and when analysed together, the probability of pest freedom achieved by 2021 was very high (close to 1), unless the considered mean time between invasion was very short. This supports the conclusion that PWN is not established in these countries. For Lithuania and Norway, the sensitivity of the trade facilitation surveys was low in all years (<0.18) and although the multiannual analysis resulted in much higher confidence, the achieved probability of pest freedom was still rather low even when the considered mean time between invasions was high (<0.5 for Lithuania and <0.8 for Norway), suggesting that the surveys may not be comprehensive enough to ensure detection of an established population of PWN with high confidence.

The surveys conducted were not extensive enough to ensure early detection of PWN invasions in any of the countries. The sensitivity of the early detection surveys was low for all years in all countries (0.0004–0.068). A higher confidence was obtained through the multiannual analysis but still the probability of pest freedom after the last annual survey was low throughout the studied range of mean time between PWN invasions for all countries (<0.53). Even in Sweden, where the maximum acceptable area of PWN infestation at detection, was set very large (314 km^2), the confidence barely reached above 50% and only if the mean time between PWN invasions was about 40 years.

Further, it should be noted that a large infested area (such as 314 km^2) may not be feasible to eradicate and aiming to detect a smaller infested area would increase the likelihood of succeeding with the eradication. However, in order to detect a small infested area with high confidence would require a much more intensive survey effort than currently performed.

The results of the assessments clearly show that, if the mean time between invasions is not very short, a higher confidence is obtained when analysing survey data from many years together (but see section 4.1). The mean time between invasions in EU-countries through trade is not likely to be short since, although the PWN was introduced to Europe already at the end of the 20th century, it is still restricted to Portugal (with the exception of some isolated outbreaks in the neighbouring country Spain) and no new introductions into EU from non-EU countries have been reported (EPPO 2023).

It should be highlighted that the results of the surveys presented here reflect the design prevalences set in the different countries, and therefore, e.g., if lower country level design prevalences would have been set, the conclusion could have been that none of the countries performed comprehensive enough surveys to ensure detection of an established population of PWN with high confidence. Furthermore, the prevalence of the native *B. mucronatus* in the samples collected in the surveys differed greatly between countries and therefore the inspection site level design prevalences (that were set using information about the estimated apparent prevalences of *B. mucronatus*) were very low in some of the countries. Consequently, the results presented in this report for the different countries cannot be directly compared. However, comparisons using the same design prevalences for all the countries can be made by any interested party since all the data needed to run the assessments has been published online (Tuomola et al. 2023a).

4 Discussion

The statistical methods used in NoBaSURV-PWN are, for the most parts, the same as used by EFSA (2012, 2020b) and referred to in the EU plant health legislation (European Commission 2020a, 2020b). Although these methods can therefore be considered established, their plant health applications are relatively few (but see, e.g., Dominiak et al. 2011, Kean et al. 2015). Hence, the field is relatively young and thus, there should still be room for learning, discussion, and further development.

In this project, we identified some issues that are worth highlighting, or could benefit from clarification or further development (see sections 4.1–4.5). We assessed the sensitivity achieved in past surveys, but the same issues apply when calculating the number of samples needed to achieve a given sensitivity when planning future surveys.

4.1 The probability of pest freedom gained in several years of surveys must be interpreted with caution

The idea of accumulating confidence in pest freedom over time with repeated surveys is tempting since it could potentially save a considerable amount of plant health resources. However, if the sensitivity of annual surveys is low, the results for “the probability of pest freedom after the last annual survey” can be misleading or simply incorrect.

The probability of pest freedom gained in repeated surveys must be interpreted with caution because the initial prior probability of freedom has a major impact on the results. And, if the sensitivity of annual surveys is low, this effect is clear also after several years of surveys (Figure 6 and Hannunen and Tuomola 2020a). Therefore, e.g., if the initial prior probability of freedom has been overestimated, the probability of pest freedom will be overestimated too. This is true even if a seemingly uninformative initial prior probability of freedom of 0.5 is used, as is often recommended.

The results for “the probability of pest freedom after the last annual survey” will be more solid if the initial prior probability of freedom is assessed based on sound evidence, or if the sensitivity of the first annual survey(s) is high enough. Note however, that even then, the probability of pest freedom may be lower for long survey programs than for shorter ones. This is because the impact of the initial prior probability of freedom, and that of the first annual survey(s) with high sensitivity, on the probability of pest freedom, diminishes over time (Figure 6).

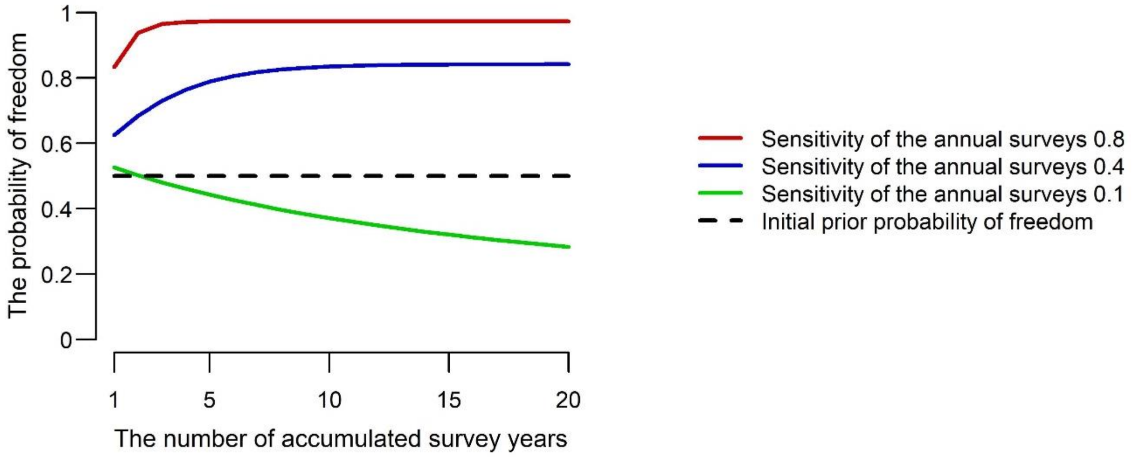

Figure 6: An illustration of the relationship between the number of survey years and the probability of pest freedom after the last survey for different sensitivities of the annual surveys, given an initial prior probability of freedom of 0.5 and a mean time between invasions of 10 years.

4.2 It is essential to know if the survey should aim to ensure eradication of invasions

When planning a survey, it must be clear if its aim is to ensure that invasions are detected so early that the pest can be eradicated. This is because, at the early stages of invasion, pest populations are highly aggregated, and this aggregation must be accounted for when calculating the number of samples needed to achieve a given sensitivity. Otherwise, the number of samples needed can be seriously underestimated.

Furthermore, it must be known if the aim of the survey is to detect invasions so early that there is only one cluster of infested plants such that the distribution of the pest is limited, e.g., to one administrative region. This too affects how the number of samples should be calculated, and if ignored, the number of samples needed will be underestimated.

Ideally, the number of samples needed in early detection surveys should be calculated using beta-binomial based sampling (Venette et al. 2002). This approach requires that the degree of spatial aggregation of the pest population is known, which often presents a problem. However, for detection surveys that aim to facilitate eradication, the expected degree of aggregation could be defined by assuming that there is only one outbreak, and that its size is equal to the maximum acceptable area of infestation at detection.

Another way to account for aggregation of the pest population, at least to some extent, is to use stratified sampling. In it, the target population is divided into subpopulations (strata) that are sampled separately. This approach is recommended by EFSA (EFSA 2012, 2020b) and applied in NoBaSURV-PWN. Furthermore, in NoBaSURV-PWN, the sensitivity of the survey at national level is calculated so that, for early detection surveys, the pest population is expected to be limited to one region while all other regions are assumed to be free from the pest (see Annex A).

4.3 Area based sampling from finite population is problematic

In NoBaSURV-PWN, sensitivity of surveys is assessed using a 2-step approach presented by Martin et al. (2007) and applied to a plant health case by EFSA (2020c) and by Hannunen and Tuomola (2020a, 2020b). At the 1st step, the sensitivity of inspecting one epidemiological unit (i.e., inspection sensitivity) is calculated, and at the 2nd step, the sensitivity of the survey is calculated using inspection sensitivity as method sensitivity (Table 7 and Annex A). Respectively, the sample size needed is estimated by first estimating the number of inspection units (e.g., host plants) that need to be sampled within each epidemiological unit (e.g., single hectare) and then by estimating the number of epidemiological units that need to be inspected in the entire area to achieve the desired confidence level for detecting the pest above the design prevalence (EFSA 2020c).

Table 7: Summary of the 2-step approach and its application in NoBaSURV-PWN. Note that population size is not needed when the sensitivity of inspections and the sensitivity of the survey are calculated using the Poisson distribution.

	1st step	**2nd step**
Population size	The number of inspection units per epidemiological unit (*pop*)	The number of epidemiological units in the region or country (*POP*)
Design prevalence	Inspection site level design prevalence (*dp*)	Areal level design prevalence (*DP*)
Design prevalence set as	The proportion of infested inspection units	The proportion of infested epidemiological units
Sampled units	Inspection units	Epidemiological units
Number of samples	The number of inspection units analysed per epidemiological unit (*n*)	The number of epidemiological units inspected in the region or country (*N*)
Method sensitivity	Sampling effectiveness × analytical sensitivity (*MSe*)	Sensitivity of inspections (*ISe*)
Measure of confidence	Sensitivity of inspections (*ISe*)	Sensitivity of the survey (*SSe*)
Calculation of confidence in NoBaSURV-PWN	Using the Poisson distribution as $ISe = 1 - e^{-n \cdot dp \cdot MSe}$	Using the Poisson distribution as $SSe = 1 - e^{-N \cdot DP \cdot ISe}$

In the aforementioned plant health applications of the 2-step approach, the sensitivity of the survey / the sample size is calculated either using hypergeometric or binomial distribution (assuming sampling from a finite or infinite population respectively). However, using hypergeometric distribution in the 2-step approach is problematic in situations where the size of the epidemiological units cannot be defined objectively. This is because the size of epidemiological units affects the results. To demonstrate this, we present an example calculation in Figure 7.

For small epidemiological unit sizes (red background in Figure 7 b), the sensitivity of the survey decreases steeply with increasing size of epidemiological units. This is because within this range, the proportion of sampled inspection units per epidemiological unit is so large (>0.1) that the population size of the 1st step has an impact on the sensitivity of inspections, but the proportion of inspected epidemiological units is so small (<0.1) that the population size of the 2nd step does not have an impact on the sensitivity of the survey.

For large epidemiological unit sizes (blue background in Figure 7 b), the sensitivity of the survey increases gently with increasing size of epidemiological units. Within this range, the proportion of sampled inspection units per epidemiological units is so small (<0.1) that the population size of the 1st step does not have an impact on the sensitivity of inspections, but the proportion of inspected epidemiological units is so large (>0.1) that the population size of the 2nd step has an impact on the sensitivity of the survey.

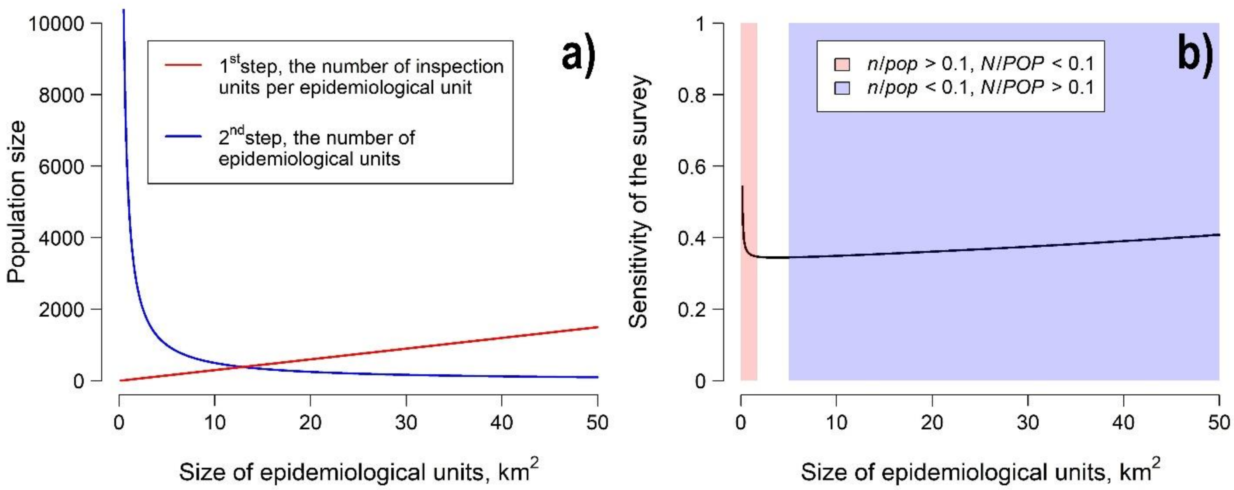

Figure 7: If hypergeometric distribution is used, the size of epidemiological units affects a) the population size at the 1st and 2nd steps of the 2-step approach (*pop* and *POP* respectively), and b) the sensitivity of the survey. In the example, the total area of the target population = 5000 km^2, the density of inspection units = 30 per km^2, design prevalence for the 1st step (*dp*) = 0.1, design prevalence for the 2nd step (*DP*) = 0.01, the number of inspection units sampled per epidemiological unit (*n*) = 5, and the number of epidemiological units inspected in the survey (*N*) = 100.

Since continuous forest environments cannot be divided into similar-sized epidemiological units objectively, in NoBaSURV-PWN, both the sensitivity of inspections and the sensitivity of the survey are assessed using the Poisson distribution (see Annex A).

4.4 Optimising 1st and 2nd step design prevalences could save resources

When the 2-step approach is used, design prevalences must be set separately for each step. The design prevalence of each step must be biologically plausible, and it must reflect the aim of the survey. However, if the aim is “to detect, with a given confidence, a given level of presence of infested plants”, this can be achieved by several different combinations of 1st and 2nd step design prevalences.

For example, the requirement imposed on *Xylella fastidiosa* surveys (European Commission 2020b), i.e., “the survey design and sampling scheme used shall be able to identify within the Member State concerned, with at least 80% of confidence, a level of presence of infected plants of 1%” can be fulfilled, e.g., using 10% as design prevalences for both steps, or using 20% as design prevalences for the 1st step and 5% as design prevalences for the 2nd step. (Note that if 1% is used as design prevalence for both steps, the survey will detect, with at least 80% of confidence, a level of presence of infected plants of 0.01%, which is a much lower prevalence than required.)

Since the cost of achieving the required confidence level with different combinations of the 1st and 2nd step design prevalences likely differ, the survey design could be optimised to minimise the total cost of the survey, keeping in mind, of course, that the design prevalences must be biologically plausible. Tools to facilitate such optimisation could potentially save resources and improve the effectiveness of surveys.

Note also that whenever the 2-step approach is used, the conclusions of the survey should be reported so that the design prevalences of both steps are presented and the interpretation of the results is explained considering both steps.

4.5 Anyone can build on NoBaSURV-PWN and NoBa LCR

The source codes of NoBaSURV-PWN and NoBa LCR have been published with an open-source software licence in EFSA's Knowledge Junction repository (Hannunen and Tuomola 2023, Tuomola et al. 2023b) to promote further development of easy-to-use applications needed in statistical assessment and planning of surveys.

NoBaSURV-PWN can, in principle, be used also for assessments of past surveys of other quarantine pests, not only PWN. However, the terminology used in the app has been tailored for PWN and thus, using it for other pests requires some effort. Therefore, a truly pest generic app for assessing the statistical confidence accumulated in past surveys might be useful and could easily be developed based on the source code of NoBaSURV-PWN.

NoBa LCR can be used for retrieving the land cover data needed when planning surveys for many quarantine pests, especially those that inhabit coniferous, broadleaved and/or mixed forests. Hence, adding the CORINE land cover data of other EU countries to NoBa LCR, or an app developed based on it, could be useful.

5 References

Dominiak BC, Gott K, McIver D, Grant T, Gillespie PS, Worsley P, Clift A and Sergeant ESG, 2011. Scenario tree risk analysis of zero detections and the eradication of yellow crazy ant ('*Anoplolepis gracilipes*' (Smith)), in New South Wales, Australia. Plant Protection Quarterly, 26, 124–129. https://search.informit.com.au/documentSummary;dn=593474062968814;res=IELHSS

Futai K, 2013. Pine wood nematode, *Bursaphelenchus xylophilus*. Annual Review of Phytopathology, 51, 61–83. https://doi.org/10.1146/annurev-phyto-081211-172910

EFSA (European Food Safety Authority), 2012. A framework to substantiate absence of disease: the risk based estimate of system sensitivity tool (RiBESS) using data collated according to the EFSA Standard Sample Description - An example on *Echinococcus multilocularis*. Supporting Publications, 2012 9(12):EN-366, 1–44. https://doi.org/10.2903/sp.efsa.2012.EN-366

EFSA (European Food Safety Authority), 2014. Guidance on expert knowledge elicitation in food and feed safety risk assessment. EFSA Journal 2014, 12(6):3734, 278 pp. https://doi.org/10.2903/j.EFSA.2014.3734

EFSA (European Food Safety Authority), Ciubotaru RM, Cortiñas Abrahantes J, Oyedele J, Parnell S, Schrader G, Zancanaro G and Vos S, 2018. Work-plan and methodology for EFSA to develop plant pest survey guidelines for EU Member States. EFSA supporting publication 2018, 15(3):EN - 1399, 1–36. https://doi.org/10.2903/sp.efsa.2018.EN-1399

EFSA (European Food Safety Authority), Baker R, Gilioli G, Behring C, Candiani D, Gogin A, Kaluski T, Kinkar M, Mosbach-Schulz O, Neri FM, Preti S, Rosace MC, Siligato R, Stancanelli G and Tramontini S, 2019. *Bursaphelenchus xylophilus* – Pest report and datasheet to support ranking of EU candidate priority pests [Data set]. Zenodo. https://doi.org/10.5281/zenodo.2788667

EFSA (European Food Safety Authority) 2020a. Pest survey card on *Bursaphelenchus xylophilus*. Plant pest survey cards gallery. https://efsa.maps.arcgis.com/apps/MapJournal/index.html?appid=4a10774369904c8c9c0c9d7b41ed1f5f

EFSA (European Food Safety Authority), Lázaro E, Parnell S, Vicent Civera A, Schans J, Schenk M, Cortiñas Abrahantes J, Zancanaro G and Vos S, 2020b. General guidelines for statistically sound and risk-based surveys of plant pests. EFSA supporting publication, 2020: 17(9): EN-1919, 65 pp. https://doi.org/10.2903/sp.efsa.2020.EN-1919

EFSA (European Food Safety Authority), Lázaro E, Parnell S, Civera AV, Schans J, Schenk M, Schrader G, Abrahantes JC, Zancanaro G and Vos S, 2020c. Guidelines for statistically sound and risk-based surveys of *Xylella fastidiosa*. EFSA Supporting publication, 2020:EN-1873. 76 pp. https://doi.org/10.2903/sp.efsa.2020.EN-1873

EPPO (European and Mediterranean Plant Protection Organization), 2013. PM 7/4 (3) *Bursaphelenchus xylophilus*. EPPO Bulletin, 43, 105–118. https://doi.org/10.1111/epp.12024

EPPO (European and Mediterranean Plant Protection Organization), 2018. PM 9/1(6) *Bursaphelenchus xylophilus* and its vectors: procedures for official control. EPPO Bulletin 48(3), 503–515. https://doi.org/10.1111/epp.12505

EPPO (European and Mediterranean Plant Protection Organization), 2021. Eradication of *Anoplophora glabripennis* in Finland. EPPO Reporting Service no. 02-2021 Num. article: 2021/032. https://gd.eppo.int/reporting/article-6971

EPPO (2023) *Bursaphelenchus xylophilus*. EPPO datasheets on pests recommended for regulation. https://gd.eppo.int/taxon/BURSXY/datasheet (Accessed 31 Marc 2023)

Estonian Topographic Database, Land Board, 2021. https://geoportaal.maaamet.ee/eng/Spatial-Data/Estonian-Topographic-Database-p305.html (Accessed 22 December 2021)

Etxebeste I, Sanchez-Husillos E, Álvarez G, Mas i Gisbert H, and Pajares J, 2016. Dispersal of *Monochamus galloprovincialis* (Col.: Cerambycidae) as recorded by mark–release–recapture using pheromone traps. Journal of Applied Entomology, 140, 485–499. https://doi.org/10.1111/jen.12278

European Commission, 2012. Commission implementing decision of 26 September 2012 on emergency measures to prevent the spread within the Union of *Bursaphelenchus xylophilus* (Steiner et Buhrer) Nickle et al. (the pine wood nematode). Official Journal of the European Union L 266, 2.10.2012, p. 42. http://data.europa.eu/eli/dec_impl/2012/535/oj

European Commission, 2019. Commission delegated regulation (EU) 2019/1702 of 1 August 2019 supplementing Regulation (EU) 2016/2031 of the European Parliament and of the Council by establishing the list of priority pests. Official Journal of the European Union, L 260, 19.10.2019, 8–10. http://data.europa.eu/eli/reg_del/2019/1702/oj

European Commission, 2020a. Commission implementing regulation (EU) 2020/1231 of 27 August 2020 on the format and instructions for the annual reports on the results of the surveys and on the format of the multiannual survey programmes and the practical arrangements, respectively provided for in Articles 22 and 23 of Regulation (EU) 2016/2031 of the European Parliament and the Council. Official Journal of the European Union L 280, 28.8.2020, 1–17. https://eur-lex.europa.eu/eli/reg_impl/2020/1231/oj

European Commission, 2020b. Commission implementing regulation (EU) 2020/1201 of 14 August 2020 as regards measures to prevent the introduction into and the spread within the Union of *Xylella fastidiosa* (Wells et al.). Official Journal of the European Union L 269, 17.8.2020, 2–39. http://data.europa.eu/eli/reg_impl/2020/1201/oj

EEA (European Environment Agency), 2022. Copernicus, Land Monitoring Service, CLC 2018. In: Copernicus. https://land.copernicus.eu/pan-european/corine-land-cover/clc2018

European Union, 2016. Regulation (EU) 2016/2031 of the European Parliament of the Council of 26 October 2016 on protective measures against pests of plants, amending Regulations (EU) No 228/2013, (EU) No 652/2014 and (EU) No 1143/2014 of the European Parliament and of the Council and repealing Council Directives 69/464/EEC, 74/647/EEC, 93/85/EEC, 98/57/EC, 2000/29/EC, 2006/91/EC and 2007/33/EC. Official Journal of the European Union, L 317, 23.11.2016, 4–104. http://data.europa.eu/eli/reg/2016/2031/oj

Gruffudd HR, Jenkins TAR and Evans HF, 2016. Using an evapo-transpiration model (ETpN) to predict the risk and expression of symptoms of pine wilt disease (PWD) across Europe. Biological Invasions, 18, 2823–2840. https://doi.org/10.1007/s10530-016-1173-7

Hannunen S and Tuomola J, 2020a. Assessing the probability of freedom from pine wood nematode based on 19 years of surveys. NeoBiota, 58, 75–106. https://doi.org/10.3897/neobiota.58.38313

Hannunen S and Tuomola J, 2020b. FinnSURV-Assess PWN - A tool for assessing the confidence in pest freedom gained in official pine wood nematode surveys (1.0). Finnish Food Authority, Helsinki, Finland. The source code available at Zenodo https://doi.org/10.5281/zenodo.3842358

Hannunen S and Tuomola J, 2023. NoBaSURV-PWN - A tool for assessing the confidence in pest freedom gained in official pine wood nematode surveys. Finnish Food Authority, Helsinki, Finland. The app available at https://nobasurv-pwn.rahtiapp.fi/, and the source code at Zenodo https://doi.org/10.5281/zenodo.7766617

IPPC Secretariat, 2022. Glossary of phytosanitary terms. International Standard for Phytosanitary Measures No. 5. Rome. FAO on behalf of the Secretariat of the International Plant Protection Convention. https://www.fao.org/3/mc891e/mc891e.pdf

Kean JM, Burnip GM and Pathan A, 2015. Detection survey design for decision making during biosecurity incursions. In: Jarrad F, Low-Choy S and Mengersen K (Eds) Biosecurity surveillance Quantitative approaches. CABI (USA), 238–250.

Kukkonen H, Pouttu A, Tomminen J, Lilja P, Ronni P, Niemi K, Sahin T, Soukainen J, Kainulainen A, Islander A, Lampela H and Hannunen S, 2011. [Contingency plan for the pine wood nematode] in Finnish. Finnish Food Safety Authority Evira. Evira publications 7/2011. ISBN 978-952-225-089-6.

Nathan R, Klein E, Robledo-Anuncio JJ and Revilla E, 2012. Dispersal kernels: review. Chapter 15. Dispersal Ecology and Evolution, in Clobert J, Baguette M, Benton TG and Bullock JM (eds), Dispersal Ecology and Evolution, Oxford University Press, UK.

R Core Team, 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/

Sundheim L, Økland B, Magnusson C, Solberg B, Rafoss T, 2010. Pest risk assessment of the pine wood nematode (PWN) *Bursaphelenchus xylophilus* in Norway – Part 2. Opinion of the Plant Health Panel of the Scientific Committee for Food Safety 31.08.2010, VKM Report 2010: 24, 08/906-6_final, 21 pp.

https://vkm.no/download/18.2994e95b15cc545071682c61/1500471103388/718f593d02.pdf

Tuomola J, Gruffudd H, Ruosteenoja K and Hannunen S, 2021. Could pine wood nematode (*Bursaphelenchus xylophilus*) cause pine wilt disease or even establish inside healthy trees in Finland now—or ever? Forests, 12, 1679. https://doi.org/10.3390/f12121679

Tuomola J, Alanko A-M, Björklund N, Boberg J, Flø D, Ilau B, Kinkar M, Marinova-Todorova M, Vaicekauskas A, Valatkevičienė L, Vihervuori L, Wendell M and Hannunen S, 2023a. Data on the Nordic-Baltic pine wood nematode surveys [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7793987

Tuomola J, Marinova-Todorova M and Hannunen S, 2023b. NoBa Land Cover Retriever - A tool for retrieving land cover data needed in statistical assessment and planning of quarantine pest surveys. Finnish Food Authority, Helsinki, Finland. The app available at http://www.noba-lcr.rahtiapp.fi/, and the source code at Zenodo http://doi.org/10.5281/zenodo.7560539

Økland B, Skarpaas O, Schroeder M, Magnusson C, Lindelöw Å and Thunes K, 2010. Is eradication of the pinewood nematode (*Bursaphelenchus xylophilus*) likely? An evaluation of current contingency plans. Risk Analysis, 30, 1424–1439. https://doi.org/10.1111/j.1539-6924.2010.01431.x

Venette R, Moon R and Hutchison W, 2001. Strategies and statistics of sampling for rare individuals. Annual Review of Entomology, 47, 143-174. https://doi.org/10.1146/annurev.ento.47.091201.145147

Glossary and Abbreviations

Glossary of terms used in this report.

Apparent prevalence	The proportion of analysed units testing positive.
Areal level design prevalence	Design prevalence for the 2nd step of the 2-step approach (see Table 7 in section 4.3).
Baseline area	In a risk-based survey design, the area with the lowest relative risk.
Components of a survey	Survey entities that differ in the target population (e.g., host plant species), the inspection unit (e.g., wood vs. vectors), or the detection method (e.g., visual examination vs. laboratory testing).
Confidence	A measure of reliability of the survey procedure. In this project, it is quantified as the sensitivity of annual surveys and the probability of pest freedom after the last annual survey.
Design prevalence	Roughly, design prevalence determines the minimum prevalence that the survey is aimed to detect, and sensitivity determines the probability with which the survey is expected to succeed in this aim. If the pest prevalence is equal to (or greater than) the design prevalence, at least one infested individual will be detected in the survey, with the probability equal to (or greater than) the sensitivity of the survey.
Detection survey	Survey conducted to determine the presence or absence of pests (IPPC Secretariat 2022).
Diagnostic sensitivity	The probability that a truly positive sample will test positive.
Early detection survey	A detection survey that aims to detect possible pest invasions early enough to facilitate successful eradication.
Entry (of a pest)	Movement of a pest into an area where it is not yet present, or present but not widely distributed and being officially controlled (IPPC Secretariat 2022)
Entry site	A site where the probability of pest entry (to the country) is elevated. For PWN, e.g., harbours, industrial areas and landfills.
Epidemiological unit	Subdivision of the target population that contains inspection units. Each epidemiological unit should be homogenous with respect to the biotic and abiotic factors that are relevant for the pest such that, if the pest was present in an epidemiological unit, it could be randomly distributed within the inspection units.
Establishment (of a pest)	Perpetuation, for the foreseeable future, of a pest within an area after entry (IPPC Secretariat 2022)
Expert knowledge elicitation	A systematic, documented and reviewable process to retrieve expert judgements from a group of experts in the form of a probability distribution (EFSA 2014).
Initial prior probability of freedom	The probability that the prevalence of the pest is below the design prevalence before the first survey.

Inspection site	A place where wood is sampled or *Monochamus* beetles are trapped.
Inspection site level design prevalence	Design prevalence for the 1st step of the 2-step approach (see Table 7 in section 4.3).
Inspection unit	The plants, plant parts or vectors that are scrutinised to detect the pests, i.e., units within the epidemiological units that could host the pests and on which the pest diagnosis takes place (EFSA 2018).
Introduction (of a pest)	The entry of a pest resulting in its establishment (IPPC Secretariat 2022)
Invasion	In this report, invasion refers to an instance where a pest enters an area, manages to transfer to a suitable host plant and infest it (at least for a short while), but that does not necessarily result in the establishment of the pest (for the foreseeable future) in the area.
Method sensitivity	The probability that a truly positive inspection unit tests positive. It is the product of sampling effectiveness and diagnostic sensitivity.
Prevalence	The proportion of infested inspection units (e.g., host plants or vectors).
Prior probability of (pest) freedom	The probability that the prevalence of the pest is below the design prevalence(s) before an annual survey.
Probability of (pest) freedom	The probability that the prevalence of the pest is below the design prevalence(s) if the pest is not detected in the survey(s).
Relative risk	The ratio of the probability of infestation in one group to the probability of infestation in another group. (Note that the severity of impact of infestation is not considered.)
Risk area	Area where the probability of pest infestation is elevated, normally around entry sites.
Risk-based survey design	A survey design in which the target population is divided into subpopulations that differ in their relative risk, and the survey efforts are divided among those subpopulations.
Risk factor	A biotic or an abiotic factor that affects the probability of infestation by the pest.
Sampling effectiveness	The probability with which a sample taken from an infested inspection unit is infested.
Sensitivity	The probability with which the pest is detected if its prevalence in the target population is equal to the design prevalence.
Sensitivity of inspection (inspection sensitivity)	The probability with which the pest is detected in an inspection if its prevalence in inspection units is equal to the inspection site level design prevalence. (This can also be called the confidence level of the inspection.)
Sensitivity of survey	The probability with which the pest is detected in a survey if its prevalences are equal to inspection site and areal level design prevalences. (This can also be called the confidence level of the survey.)
Survey	An official procedure conducted over a defined period to determine the presence or absence of pests, or the boundaries or

	characteristics of a pest population, in an area, place of production or production site (IPPC Secretariat 2022).
Target population	The population (of, e.g., host plants and vectors) to which the results of the survey will be generalised to.
Trade facilitation survey	A detection survey that aims to provide evidence for pest freedom to justify import requirements related to the pest and to facilitate export to countries with corresponding requirements.
True prevalence	The actual proportion of the infested units in the population.
Wood object	Tree, log, detached branch, logging residual, or any piece of wood.

Abbreviations used in this report.

CLC	CORINE Land Cover
EAFB	Estonian Agriculture and Food Board
EEA	European Environment Agency
EFSA	European Food Safety Authority
EPPO	European and Mediterranean Plant Protection Organization
EU	European union
NoBa LCR	NoBa Land Cover Retriever
PWN	Pine wood nematode
SLU	Swedish University of Agricultural Sciences
SPSMoA	State Plant Service under the Ministry of Agriculture of the Republic of Lithuania
VKM	Norwegian Scientific Committee for Food and Environment

Annex A – Technical details of the NoBaSURV-PWN application

A.1. Introduction

NoBaSURV-PWN is a tool for assessing the confidence in pest freedom gained in official pine wood nematode (PWN, *Bursaphelenchus xylophilus*) surveys done in the past (Hannunen and Tuomola 2023). It can be used to assess both the sensitivity of annual surveys and the probability of pest freedom after the last annual survey.

In NoBaSURV-PWN, the assessment of statistical confidence of surveys is based on the principles developed by Cannon (2002) and Martin et al. (2007), i.e., the same principles that are applied in EFSA's risk-based estimate of system sensitivity tool RiBESS (EFSA 2012).

NoBaSURV-PWN is written with R version 4.2.1 (R Core Team 2022), its package 'shiny' (1.7.2) (Chang et al. 2022) and several other packages. The app is available at https://nobasurv-pwn.rahtiapp.fi/, and its source code is published at https://zenodo.org/record/7766617 under the GNU General Public License version 3 (https://www.gnu.org/licenses/gpl-3.0.html).

A.2. Options for risk-based surveys design

In NoBaSURV-PWN, the target population can be divided into subpopulations (i.e., risk factor levels) that are assumed to differ in the probability of PWN infestation (i.e., relative risk) in two different ways. In this document, these are called risk-based survey design options 1 and 2. Note that the word risk-based here refers only to the probability of PWN infestation, meaning that the severity of impact of PWN infestation is not considered when the relative risk of risk factor levels is defined.

In risk-based survey design option 1, administrative regions (or other geographical regions for which data is available separately) are assumed to differ in their relative risk, which is assumed to depend on the relative area, or number of, entry sites (i.e., areas with elevated probability of PWN entry) in the region as

$$RISK_j = \frac{E_j}{\sum_{j=1}^{J} E_j} \qquad (1)$$

where j denotes the region, J = the total number of regions, and E = the area or number of entry sites.

In risk-based survey design option 2, areas with target population that are within a user defined radius (r) from entry sites are considered "risk areas", whereas other areas with target population are considered "baseline areas".

The user can select one of three alternative ways to estimate the relative risk of risk vs. baseline areas. The first two assume that the probability of PWN infestation at different distances from entry sites depends on the dispersal behaviour of PWN vectors. The third allows defining the relative risk based on whichever user identified process.

<u>The first alternative</u> is to define the relative risk based on the '*2Dt*' dispersal location kernel parametrised for *Monochamus galloprovincialis* with the estimates provided by Etxebeste et al. (2016), i.e.

$$k_L(r) = \frac{p}{\pi \cdot u \cdot \left[1 + \frac{r^2}{u}\right]^{p+1}} \tag{2}$$

where r = the radius of risk areas (m), the shape parameter p = 0.804, and the scale parameter u = 39 760.1 m^2.

The two-dimensional dispersal location kernel (k_L, eq. 2) is converted into one-dimensional dispersal distance kernel (k_D) by multiplying it with $2\pi r$, which is then integrated numerically with respect to r with the 'integrate' function of R Stats Package version 4.2.1 (R Core Team 2022) to get the proportion of *M. galloprovincialis* that is predicted to be within risk areas after their first flight season (PWN_{risk}). This proportion of the vector population in risk vs. baseline areas is assumed to reflect the proportion of PWN population in the respective areas, and hence also the relative risk of the risk vs. baseline areas.

The relative risk of the baseline areas ($RISK_{baseline}$) is set to 1, and the relative risk of risk areas is calculated as

$$RISK_{risk} = \frac{PWN_{risk}}{PWN_{baseline}} \tag{3}$$

where PWN_{risk} = the predicted proportion of the PWN population in the risk areas, and $PWN_{baseline}$ = the predicted proportion of the PWN population in the baseline areas.

<u>The second alternative</u> is to define the relative risk of risk vs. baseline areas based on some other dispersal kernel of PWN vectors or population dynamics model of PWN. In this option, the user must define the proportion of the PWN population in risk areas. The relative risk of risk areas is then calculated as explained above for the first alternative.

<u>The third alternative</u> is to define the relative risk of risk vs. baseline areas based on some other user defined factor. In this option, the user simply defines the relative risk of the risk areas ($RISK_{risk}$) directly.

A.3. Adjusting the areal level design prevalences

For trade facilitation survey, the user sets "the country level design prevalence", which is used to calculate the effective probability of infestation for the different risk factors levels as

$$EPIpf_i = DP \cdot \frac{RISK_i}{\sum_{i=1}^{I}\left(PropPop_i \cdot RISK_i\right)} \quad (4)$$

where *i* denotes the risk factor level, *I* = the total number of risk factors levels, *DP* = the country level design prevalence, *RISK* = the relative risk of the risk factor level, and *PropPop* = the proportion of the target population in the risk factor level *i* of the total area of the target population in the country.

For early detection survey, the user sets "the maximum acceptable area of infestation at detection", which is used to calculate the regional level design prevalences as

$$DPr_j = \frac{MaxInfSize_j}{Pop_j} \quad (5)$$

where *j* denotes the region, *MaxInfSize* = maximum acceptable area of PWN infestation at detection, and *Pop* = the area of the target population in the region. Areal level design prevalences are calculated separately for each region because the total area of the target population may differ between regions. For the risk-based survey design option 2, *DPr* is further converted into effective probability of infestation in risk and baseline areas separately for each region as

$$EPIed_{j,i} = DPr_j \cdot \frac{RISK_i}{\sum_{i=1}^{I}\left(PropPop_{j,i} \cdot RISK_i\right)} \quad (6)$$

where *i* denotes the risk factor level, *I* = the total number of risk factors levels, *DPr* = the region level design prevalence, *RISK* = the relative risk of the risk factors level, and *PropPop* = the proportion of the target population in the risk factor level *i* of the total area of the target population in the region *j*.

A.4. The 2-step approach for assessing the sensitivity of annual surveys

To calculate the sensitivity of annual surveys, NoBaSURV-PWN employs a 2-step approach that is used, e.g., in EFSA (2020). In this approach, each inspection is assumed to target one epidemiological unit (i.e., area with host plants), and in each inspection, a given number of inspection units (i.e., wood objects and/or *Monochamus* adults) is sampled.

At the 1st step of the 2-step approach, the sensitivity of inspections (*ISe*) is calculated as explained below in section A.5. For this step, design prevalence (i.e., inspection site level design prevalence) is expressed as the proportion of infested inspection units (i.e., wood objects or *Monochamus* adults).

Then at the 2nd step, the sensitivity of the annual surveys is assessed using the sensitivity of inspections (*ISe*) as method sensitivity, as explained below in section A.6. For this step, design prevalence (i.e., areal level design prevalence) is expressed as the proportion of infested epidemiological units (i.e., area with PWN host plants).

A.5. Sensitivity of inspections

Sensitivity of inspections (*ISe*) is the probability that the pest will be detected in an inspected epidemiological unit when it is present in it at a prevalence equal to the inspection site level design prevalence. It is calculated separately for each region and survey component (i.e., wood objects or *Monochamus* adults) based on the Poisson probability distribution as

$$ISe_{j,c} = 1 - e^{-n_{j,c} \cdot MSe_c \cdot dp_c} \quad (7)$$

where *j* denotes the region, *c* denotes the survey component, *n* = the number of inspection units sampled per epidemiological unit, *dp* = the inspection site level design prevalence, and *MSe* = method sensitivity, i.e., the probability that the pest is detected in the sample, given that it was present in the inspection unit(s) from which the sample was taken.

A.6. Sensitivity of annual surveys

The sensitivity of annual surveys is the probability that the pest will be detected in the survey if it is present in the target population at a prevalence equal to the areal level design prevalence. It is calculated first at the regional level and then at the national level.

A.6.1. At regional level

First, the sensitivity of each survey component is calculated separately for each risk factor level based on the Poisson probability distribution as

$$GSe_{j,i,c} = 1 - e^{-N_{j,i,c} \cdot ISe_{j,c} \cdot adjDP_{j,i}} \quad (8)$$

where *j* denotes the region, *i* denotes the risk factor level, *c* denotes the survey component, *N* = the number of inspected epidemiological units, *ISe* = the sensitivity of inspections, and *adjDP* = the adjusted design prevalence. For trade facilitation surveys, *adjDP* = *EPIpf* (i.e., the effective probability of infestation, see eq. 4). For early detection surveys risk-based survey design option 1, *adjDP* = *DPr* (i.e., the regional level design prevalence, see eq. 5) and for risk-based survey design option 2, *adjDP* = *EPIed* (i.e., the effective probability of infestation, see eq. 6).

Then, the sensitivity of each risk factor level is calculated as the complement of the probability that, if the pest is present in the risk factor level at or above the adjusted design prevalence, it is not detected in any component of the survey as

$$GSe_{j,i} = 1 - \prod_{c=1}^{C} \left(1 - GSe_{j,i,c}\right) \quad (9)$$

where C = the total number of components in the survey.

Finally, the overall sensitivity of the survey for each region is obtained as the complement of the probability that, if the pest is present in the region at or above the design prevalence, it is not detected in any risk factor level as

$$GSe_{j} = 1 - \prod_{i=1}^{I} \left(1 - GSe_{j,i}\right) \quad (10)$$

where I = the total number of risk factor levels. Note that in the risk-based survey design option 1, there is only one risk factor level per region, and therefore, this last step is not needed.

A.6.2. At national level

The sensitivity of the annual survey at the national level is obtained by combining the sensitivities of the annual survey in the different regions.

For the trade facilitation survey, PWN infestation is assumed to be distributed randomly throughout the country. Therefore, the sensitivity of the survey is obtained as the complement of the probability that, if PWN is present in the country, it is not detected in any of the regions as

$$SSe = 1 - \prod_{j=1}^{J} (1 - GSe_{j}) \quad (11)$$

where j denotes the region, J = the total number of regions, and GSe_j = the overall sensitivity of the survey in region j (eq. 10).

For the early detection survey, PWN infestation is assumed to be limited to one region while all other regions are assumed to be free from the pest. Therefore, the sensitivity of the survey for the entire country (SSe) is calculated as the probability of correctly detecting the pest in the survey given that it is present in one region as

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (12)$$

where A denotes the event “PWN is detected”, and B denotes the event ”PWN is present in one region”.

When the probability that PWN is present in the region j is denoted by p_j and the total number of regions by J, the probability that the pest is present in one region is

$$P(B) = \sum_{j=1}^{J} p_{j} \quad (13)$$

based on disjoint events. (The first infestation can only be in one region, hence disjoint events). The probability that PWN is present in the country and detected is likewise the sum

$$P(A \cap B) = \sum_{j=1}^{J} GSe_j \cdot p_j \qquad (14)$$

where GSe_j = the sensitivity of the survey in region j.

The probability of PWN infestation in the region j can be expressed relative to the probability of PWN infestation in the region where it is the lowest (i.e., baseline region) as

$$p_j = q \cdot R_j \qquad (15)$$

where q = the probability of infestation in the baseline region, and R_j = the risk of infestation in the region j relative to the risk of the baseline region. (Note that here the word risk refers only to the probability of infestation and the impact of the infestation is not considered.) For the baseline region, $R = 1$, and for other regions $R_j = p_j/q > 1$.

By inserting equation 4 to equation 2 and 3 and then those to equation 1 we get

$$P(A|B) = \frac{\sum_{j=1}^{J} GSe_j \cdot q \cdot R_j}{\sum_{j=1}^{J} q \cdot R_j} \qquad (16)$$

From this, the unknown baseline probability of infestation q cancels to get

$$P(A|B) = \frac{\sum_{j=1}^{J} GSe_j \cdot R_j}{\sum_{j=1}^{J} R_j} \qquad (17)$$

which is used to calculate the sensitivity of the early detection surveys for the entire country (SSe).

The relative risk of infestation in the regions (R_j) is approximated assuming that the probability of infestation in the region (p_j) depends linearly on the area or number of entry sites in the region (E_j) as $p_j = \theta \cdot E_j$, where $\theta < 1/\max\{E_j \ldots E_J\}$ to ensure that $p_j < 1$.

Since $R_j = p_j/q$, where $q = \min\{q_j \ldots q_J\}$ (see eq. 15), its approximation using the area or number of entry sites is

$$R_j \approx \frac{\theta \cdot E_j}{\theta \cdot \min\{E_j \ldots E_J\}} \qquad (18)$$

from which the unknown factor θ (which defines the relationship between the area or number of entry sites and the probability of infestation) cancels to get

$$R_j \approx \frac{E_j}{\min\{E_j ... E_J\}} \tag{19}$$

A.7. The probability of pest freedom after the last annual survey

The probability of pest freedom is the probability that the prevalence of the pest is below the design prevalence given that the pest has not been detected in the surveys. It is estimated for each region and for the entire country in a stepwise manner by progressively updating the estimate with evidence gained in the annual surveys using Bayes' theorem according to Martin et al. (2007) as

$$Pfree_{t,j} = \frac{PriorPfree_{t,j}}{PriorPfree_{t,j} + \left[(1 - PriorPfree_{t,j}) \cdot (1 - Se_{t,j})\right]} \tag{20}$$

where *j* denotes the area considered (that may be a region or an entire country), *t* = time, *PriorPfree* = the prior probability of pest freedom, and *Se* = the sensitivity of the annual survey. For the regions *Se* = *GSe* (eq. 10), and at country level *Se* = *SSe* (eq. 11 or 17 depending on the survey type).

The initial prior probability of freedom, (i.e., the prior probability of freedom for the first time-step) is defined by the app user, and the same value is used for all the regions and for the entire country.

For all the other time steps, the prior probability of freedom is calculated as the complement of the probability that a) the prevalence of the pest would be above the design prevalence although it was not detected in the previous survey, or b) the pest was introduced to the area after the previous survey, according to Martin et al. (2007), as

$$PriorPfree_{t,j} = 1 - \left[\left(1 - Pfree_{t-1,j}\right) + Pinv_j - \left(1 - Pfree_{t-1,j}\right) \cdot Pinv_j\right] \tag{21}$$

where *Pinv* = the probability of PWN invasion to the considered area after the survey that was conducted at time *t*—1.

The probability of at least one invasion to the country per time step (i.e., year) is calculated using the Poisson distribution as the complement of the probability of no invasions as

$$Pinv_{COUNTRY} = 1 - e^{(-\frac{1}{Tinv})} \tag{22}$$

where *Tinv* = the mean time between PWN invasions defined by the app user and *e* is the mathematical constant (approximately equal to 2.71828). From this, the probability of invasion to the different regions is calculated as

$$Pinv_j = Pinv_{COUNTRY} \cdot RP_j \tag{23}$$

where RP_j = the ratio of the probability of pest entry to region *j* to the probability of PWN entry to the entire country normalised between 0 and 1 (eq. 13).

A.8. Symbols and their definitions

A	The event "PWN is detected"
adjDP	The adjusted design prevalence
B	The event "PWN is present in one region"
c	The component of the survey (i.e., wood sampling component 1 and 2, or trapping of *Monochamus* adults)
C	The total number of components in the survey
dp	Inspection site level design prevalence (i.e., design prevalence for the 1st step of the 2-step approach) expressed as the proportion of infested inspection units (i.e., wood objects or *Monochamus* adults)
DP	National level design prevalence (i.e., design prevalence for the 2nd step of the 2-step approach) expressed as the proportion of infested epidemiological units (i.e., area with PWN host plants) in the country
DPr	Regional level design prevalence (design prevalence for the 2nd step of the 2-step approach) expressed as the proportion of infested epidemiological units (i.e., area with PWN host plants) in the region
E	The area or number of entry sites
EPIed	The effective probability of infestation for early detection surveys
EPIpf	The effective probability of infestation for trade facilitation surveys
GSe	The sensitivity of the survey at regional level (this can also be called the confidence level of surveys)
i	Risk factor level
ISe	The sensitivity of inspections (this can also be called the confidence level of inspections)
j	Region
J	The total number of regions
MaxInfSize	The maximum acceptable area of PWN infestation at detection
MSe	Method sensitivity, i.e., the probability that the pest is detected in the sample, given that it was present in the inspection unit(s) from which the sample was taken
N	The number of epidemiological units inspected (the sample size for the 2nd step of the 2-step approach)
n	The number of inspection units sampled per epidemiological unit (the sample size for the 1st step of the 2-step approach)
p	The probability that PWN is present in the considered area

Pfree	The probability of pest freedom
Pinv	The probability of PWN invasion to the considered area between the previous and the following survey
$Pinv_{COUNTRY}$	The probability of PWN invasion to the country
$PWN_{baseline}$	The predicted proportion of the PWN population in baseline areas (in case of PWN invasion)
PWN_{risk}	The predicted proportion of the PWN population in risk areas (in case of PWN invasion)
PriorPfree	Prior probability of pest freedom
PropPop	The proportion of the target population
q	The probability of infestation in the baseline region, i.e., in the region where it is the lowest
R	The ratio of the probability of infestation in a region to the probability of infestation in the baseline region (i.e., in the region where it is the lowest).
RISK	The relative risk of a risk factor level
RP	The ratio of probability of PWN entry to a region to the probability of PWN entry to the entire country normalised between 0 and 1
SSe	The sensitivity of the survey at national level (this can also be called the confidence level of the survey)
t	Time
T_{inv}	The mean time between invasions
θ	The factor that defines the relationship between the area or number of entry sites and the probability of infestation in a region

A.9. References

Cannon RM, 2002. Demonstrating disease freedom - Combining confidence levels. Preventive Veterinary Medicine, 52, 227–249. https://doi.org/10.1016/S0167-5877(01)00262-8

Chang W, Cheng J, Allaire J, Sievert C, Schloerke B, Xie Y, Allen J, McPherson J, Dipert A and Borges B, 2022. shiny: Web Application Framework for R. R package version 1.7.2, https://CRAN.R-project.org/package=shiny

EFSA (European Food Safety Authority), 2012. A framework to substantiate absence of disease: the risk based estimate of system sensitivity tool (RiBESS) using data collated according to the EFSA Standard Sample Description - An example on *Echinococcus multilocularis*. Supporting Publications, 2012 9(12):EN-366, 1–44. https://doi.org/10.2903/sp.efsa.2012.EN-366

EFSA (European Food Safety Authority), Lázaro E, Parnell S, Civera AV, Schans J, Schenk M, Schrader G, Abrahantes JC, Zancanaro G and Vos S, 2020. Guidelines for statistically sound and risk-based surveys of *Xylella fastidiosa*. EFSA Supporting publication, 2020:EN-1873. 76 pp. https://doi.org/10.2903/sp.efsa.2020.EN-1873

Etxebeste I, Sanchez-Husillos E, Álvarez G, Mas i Gisbert H and Pajares J, 2016. Dispersal of *Monochamus galloprovincialis* (Col.: Cerambycidae) as recorded by mark–release–recapture using pheromone traps. Journal of Applied Entomology, 140, 485–499. https://doi.org/10.1111/jen.12278

Hannunen S and Tuomola J, 2023. NoBaSURV-PWN – A tool for assessing the confidence in pest freedom gained in official pine wood nematode surveys. Finnish Food Authority, Helsinki, Finland. The app is available at https://nobasurv-pwn.rahtiapp.fi/, and the source code at Zenodo https://doi.org/10.5281/zenodo.7766617

Martin PAJ, Cameron AR and Greiner M, 2007. Demonstrating freedom from disease using multiple complex data sources 1: A new methodology based on scenario trees. Preventive Veterinary Medicine, 79, 71–97. https://doi.org/10.1016/j.prevetmed.2006.09.008

R Core Team, 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/

Annex B – NoBa Land Cover Retriever application

B.1. Introduction

The NoBa Land Cover Retriever (NoBa LCR) is a web application for retrieving CORINE Land Cover (CLC) data (EEA 2022) needed in the statistical assessment and planning of quarantine pest surveys. The countries currently included in NoBa LCR are Estonia, Finland, Lithuania, Norway and Sweden.

NoBa LCR has been tailored for retrieving the data needed for analysing and planning risk-based surveys in which a) the relative risk of each administrative region is calculated based on the area or number of entry sites in the region (i.e., risk-based survey design option 1 of the NoBaSURV-PWN application), or b) the target population is divided into risk areas that are close to entry sites and baseline areas that are further away from entry sites (i.e., risk-based survey design option 2 of NoBaSURV-PWN).

NoBa LCR can be used to retrieve:

- The area of entry sites per region
- The number of entry sites per region
- The total area of the target population per region
- The area of the target population within a user-defined radius from entry sites (risk areas) and outside it (baseline areas) per region

The results can be explored on an interactive map and downloaded in the following formats:

- The summary table of the results (csv)
- The retrieved data in a format needed for NoBaSURV-PWN (csv)
- The retrieved data in GIS formats (geoTIFF or shapefile)
- The parameter values used for retrieving the data (txt)

NoBa LCR is freely available at https://noba-lcr.rahtiapp.fi/ and its source code has been published at EFSA's Knowledge Junction repository (Tuomola et al. 2023) under the GNU General Public License version 3 (https://www.gnu.org/licenses/gpl-3.0.html).

B.2. Programming language

NoBa LCR is written with R version 4.2.1 (R Core Team 2022) and its package 'shiny' (Chang et al. 2022). R packages 'raster' (Hijmans 2022), 'sf' (Pebesma 2018), 'sp' (Pebesma and Bivand 2005, Bivand et al. 2013) and 'rgdal' (Bivand et al. 2022) are used for retrieving and analysing the GIS data. R package 'leaflet' (Cheng et al. 2022) is used to create an interactive map for visual exploration of the results.

The following R packages are used to finalize the user experience of the application: 'shinythemes' (Chang 2021), 'shinyhelper' (Mason-Thom 2019), 'shinybusy' (Meyer and Perrier 2022), 'shinyWidgets' (Perrier et al. 2022), 'tidyverse' (Wickham et al. 2019) and 'zip' (Csárdi et al. 2021).

Further details on the R operations used in the NoBa LCR can be found in the source code of the app (Tuomola et al. 2023).

B.3. GIS data used by the application

B.3.1. CORINE Land Cover (CLC) data

CORINE Land Cover (CLC) is the European land cover database of the Copernicus Land Monitoring Service (CLMS) (EEA 2022). CLMS is implemented by the European Environment Agency (EEA) and the European Commission DG Joint Research Centre.

NoBa LCR application uses the CLC 2018, version 2020_20u1 dataset (CLC2018) in a 100m GeoTIFF format (EEA 2022).

The CLC 2018 data includes 44 classes of land cover, but in NoBa LCR only 31 land cover classes, that were considered potentially relevant in the statistical assessment and planning of quarantine pest surveys, are included.

B.3.2. User's own data on entry sites

NoBa LCR enables the addition of user's own data on entry sites. The added data can be used instead or together with the CLC data. The added data should represent point locations of the entry sites as WGS84 coordinates and should be uploaded to the application as a csv file.

B.3.3. Data on administrative regions

For Lithuania, Norway and Sweden, data on counties was derived from the GADM database of Global Administrative Areas (GADM 2020). For Estonia, data on counties was derived from the Estonian Geoportal (Estonian Land Board 2021). For Finland, data on the Centres for Economic Development, Transport and the Environment of Finland was derived from Statistics Finland (Statistics Finland 2022). All data on administrative regions was derived in a shapefile format.

B.4. Spatial resolution

NoBa LCR uses 100 m resolution for spatial operations with two exceptions.

In the case where the target population is divided into risk areas that are close to entry sites and baseline areas that are further away from entry sites, the resolution of the CLC data on entry sites is first aggregated into 1000 m resolution, then converted into spatial vector point data, and finally used to deploy the risk areas and the baseline areas. Aggregating the resolution speeds up the deployment of the risk areas and the baseline areas significantly. However, it also impairs the accuracy of the analysis but fortunately only by a few percent at most.

For the interactive maps, the CLC data is aggregated to allow smooth display of the maps. For Estonia and Lithuania, the CLC data is aggregated to 300 m, and for Finland, Norway and Sweden, to 400 m resolution.

B.5. References

Bivand RS, Pebesma E and Gomez-Rubio V, 2013. Applied spatial data analysis with R, Second edition. Springer, NY. https://asdar-book.org/

Bivand R, Keitt T and Rowlingson B, 2022. rgdal: Bindings for the 'geospatial' data abstraction library. R package version 1.5–32, https://CRAN.R-project.org/package=rgdal

Chang W, 2021. shinythemes: Themes for shiny. R package version 1.2.0, https://CRAN.R-project.org/package=shinythemes

Chang W, Cheng J, Allaire J, Sievert C, Schloerke B, Xie Y, Allen J, McPherson J, Dipert A and Borges B, 2022. shiny: Web application framework for R. R package version 1.7.2. https://CRAN.R-project.org/package=shiny

Cheng J, Karambelkar B and Xie Y, 2022. leaflet: Create interactive web maps with the JavaScript 'leaflet' library. R package version 2.1.1. https://CRAN.R-project.org/package=leaflet

Csárdi G, Podgórski K and Geldreich R, 2021. zip: Cross-platform 'zip' compression. R package version 2.2.0. https://CRAN.R-project.org/package=zip

GADM, 2020. GADM database of global administrative areas, version 4.1. www.gadm.org (Accessed 14 September 2022)

EEA (European Environment Agency), 2022. Copernicus, Land Monitoring Service, CLC 2018. In: Copernicus. https://land.copernicus.eu/pan-european/corine-land-cover/clc2018

Estonian Land Board, 2021. Administrative and settlement units, Estonian Land Board. https://geoportaal.maaamet.ee/eng/Spatial-Data/Administrative-and-Settlement-Division-p312.html (Accessed 1 December 2021)

Hijmans R, 2022. raster: Geographic data analysis and modeling. R package version 3.5–29. https://CRAN.R-project.org/package=raster

Mason-Thom C, 2019. shinyhelper: Easily add markdown help files to 'shiny' app elements. R package version 0.3.2. https://CRAN.R-project.org/package=shinyhelper

Meyer F and Perrier V, 2022. shinybusy: Busy indicators and notifications for 'shiny' spplications. R package version 0.3.1. https://CRAN.R-project.org/package=shinybusy

Pebesma E, 2018. Simple features for R: standardized support for spatial vector data. The R Journal 10 (1), 439–446. https://doi.org/10.32614/RJ-2018-009

Pebesma EJ and Bivand RS, 2005. Classes and methods for spatial data in R. R News 5 (2). https://cran.r-project.org/doc/Rnews/

Perrier V, Meyer F and Granjon D, 2022. shinyWidgets: Custom inputs widgets for shiny. R package version 0.7.3. https://CRAN.R-project.org/package=shinyWidgets

R Core Team, 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Tuomola J, Marinova-Todorova M and Hannunen S, 2023. NoBa Land Cover Retriever - A tool for retrieving land cover data needed in statistical assessment and planning of quarantine pest surveys. Finnish Food Authority, Helsinki, Finland. The app is available at http://www.noba-lcr.rahtiapp.fi/ and the source code at Zenodo http://doi.org/10.5281/zenodo.7560539

Statistics Finland, 2022. Centres for economic development, transport and the environment (ELY), Statistics Finland. The material was downloaded from Statistics Finland's interface service on 13 September 2022 with the license CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/deed.en)

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K and Yutani H, 2019. “Welcome to the tidyverse.” Journal of Open Source Software, 4(43), 1686. https://doi.org/10.21105/joss.01686

Annex C – Estimating the prevalence of *Bursaphelenchus mucronatus*

C.1. Introduction

The prevalence estimation of *Bursaphelenchus mucronatus* in wood and *Monochamus* samples collected in the pine wood nematode (PWN) surveys was done using a web application that was specifically designed for this purpose (Marinova-Todorova et al. 2023). The application can be used to estimate the apparent prevalence of *B. mucronatus* in individually tested wood samples and *Monochamus* adults, or pooled wood or *Monochamus* samples. In addition to *B. mucronatus*, the application can also be used to estimate the apparent prevalence of other organisms.

The application can be used to calculate:

1) The prevalence estimate, and its confidence interval for the whole data, and for each considered stratum (e.g., years, administrative regions, or tree or *Monochamus* species).
2) Pairwise comparisons of the prevalence for the considered strata.

In the NoBa-PWN project the apparent prevalence of *B. mucronatus* in wood and *Monochamus* samples was estimated separately for each country using data from all years without taking into account the potential differences in the prevalence in different tree and *Monochamus* species.

C.1.1. Technical details of the application

The application is written with R version 4.2.1 (R Core Team 2022) and its package ‘shiny’ (Chang et al. 2022). The apparent prevalence of *B. mucronatus* in individually tested samples is estimated with exact binomial test using the binom.test function from the R package ‘stats’ (R Core Team 2022). The apparent prevalence of *B. mucronatus* in pooled samples is estimated with maximum likelihood estimation test using the PoolPrev function from the R package ‘PoolTestR’ (McLure et al. 2021). Comparisons of prevalences of the pooled data are calculated with the pooledBinDiff function from the ‘binGroup’ package using Firth's Correction method to compute the point estimation, and Skew-Corrected Score method to compute the confidence interval estimation (Zhang et al. 2010).

Data must be uploaded to the application as comma separated csv files in which

- For individually tested samples, the number of samples is in one column and the number of positive samples in another column.
- For pooled samples, the number of items in each pooled sample is in one column and the presence/absence of *B. mucronatus* in the samples indicated in another column with 1 and 0 respectively.
- Data for estimating the prevalence in different stratum (e.g., years, administrative regions, tree or *Monochamus* species) is in its own column.
- The column titles are written without spaces and special characters.
- All rows where the total number of samples is zero are deleted.
- When the number of positive samples is zero, that is indicated with 0.

- Point is used as a decimal separator.

The application is available at https://b-mucronatus-prevalence-estimation.rahtiapp.fi/ and its source code has been published at EFSA's Knowledge Junction repository (http://doi.org/10.5281/zenodo.7800771) under the GNU General Public License version 3 (https://www.gnu.org/licenses/gpl-3.0.html).

C.2. Results

Table 1: Apparent prevalence of *B. mucronatus* in wood samples.

Country	Sample type	Estimate	Lower limit of 95% CI	Upper limit of 95% CI	Number of samples	Number of positive samples
Estonia	Symp[(a)]	0.078	0.066	0.091	1879	146
Finland	Symp[(a)]	0.123	0.111	0.145	2876	353
Lithuania	Symp[(a)]	0.053	0.029	0.087	266	14
Lithuania	Asymp[(b)]	0.015	0.005	0.035	334	5
Norway	Symp[(a)]	0.009	0.007	0.011	8244	75
Sweden	Symp[(a)]	0.082	0.073	0.092	3286	270

(a): Wood with signs of *Monochamus* activity

(b): wood without signs of *Monochamus* activity.

Table 2: Apparent prevalence of *B. mucronatus* in *Monochamus* samples.

Country	Estimate	Lower limit of 95% CI	Upper limit of 95% CI	Number of samples	Number of positive samples
Estonia	0.14	0.122	0.16	1285	180
Finland	0.157	0.117	0.204	287	45
Norway	0.017	0.008	0.031	575	10
Lithuania	0.036	0.008	0.143	26	2
Sweden	0.071	0.055	0.092	197	54

C.3. References

Chang W, Cheng J, Allaire J, Sievert C, Schloerke B, Xie Y, Allen J, McPherson J, Dipert A and Borges B, 2022. shiny: Web application framework for R. R package version 1.7.2. https://CRAN.R-project.org/package=shiny

Marinova-Todorova M, Tuomola J and Hannunen S, 2023. A web application for estimating the apparent prevalence of *Bursaphelenchus mucronatus* in wood and *Monochamus* samples. Finnish Food Authority, Helsinki, Finland. The app is available at https://b-mucronatus-prevalence-estimation.rahtiapp.fi/ and the source code at http://doi.org/10.5281/zenodo.7800771

McLure A, O'Neill B, Mayfield H, Lau C and McPherson B, 2021. PoolTestR: An R package for estimating prevalence and regression modelling for molecular xenomonitoring and other applications with pooled samples. Environmental Modelling and Software, 145, 105158. https://doi.org/10.1016/j.envsoft.2021.105158

R Core Team, 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Zhang B, Bilder C, Biggerstaff B, Schaarschmidt F and Hitt B, 2018. binGroup: Evaluation and experimental design for binomial group testing. R package version 2.2-1. https://CRAN.R-project.org/package=binGroup

Annex D – The confidence in pest freedom gained in past PWN surveys in five Nordic-Baltic countries

Table 1: The sensitivity of annual surveys considering that the aim of the survey was to provide evidence of pest freedom for trading partners (i.e., trade facilitation survey).

YEAR	Estonia	Finland	Lithuania	Norway	Sweden
2000		0.9301		0.0899	
2001		0.0604		0.0899	
2002	0.511	0.0502		0.0899	
2003	0.8652	0.5973		0.0899	
2004	0.6318	0.7235		0.0899	
2005	0.1869	0.6995		0.0899	
2006	0.1675	0.6594		0.0899	
2007	0.1772	0.7489		0	
2008	0.0291	0.7563		0.0908	
2009	0.0574	0.7712		0.0904	
2010	0.3162	0.7466		0.1061	
2011	0.4222	0.7309		0.0896	
2012	0.7123	0.7119		0.1155	
2013	0.697	0.7112		0.1054	0.5804
2014	0.601	0.7391		0.0061	0.408
2015	0.7728	0.6144	0.0136	0.1682	0.5426
2016	0.7614	0.6334	0.0159	0.09	0.4901
2017	0.8449	0.6492	0.0127	0.0208	0.6297
2018	0.7975	0.5854	0.011	0.1792	0.7099
2019	0.8341	0.6749	0.0152	0.128	0.6765
2020	0.8136	0.7505	0.0133		0.5911
2021	0.8083	0.7841	0.0186		0.5924

Table 2: The sensitivity of annual surveys considering that the aim of the survey was to detect possible PWN invasions early enough to enable successful eradication (i.e., early detection survey).

YEAR	Estonia	Finland	Lithuania	Norway	Sweden
2000		0.0152		0.0301	
2001		0.0005		0.0301	
2002	0.0193	0.0004		0.0301	
2003	0.0644	0.0056		0.0301	
2004	0.0338	0.0082		0.0301	
2005	0.0057	0.0079		0.0301	
2006	0.0062	0.0069		0.0301	
2007	0.0059	0.0079		0	
2008	0.0009	0.0082		0.0276	
2009	0.0025	0.0082		0.0228	
2010	0.0159	0.0076		0.0339	
2011	0.0217	0.0077		0.04	
2012	0.0435	0.0073		0.0193	
2013	0.0483	0.0072		0.0196	0.0601
2014	0.033	0.0077		0.002	0.0238
2015	0.0445	0.0053	0.0033	0.0499	0.034
2016	0.0495	0.0058	0.0044	0.0189	0.0282
2017	0.0632	0.0063	0.0031	0.0038	0.0424
2018	0.056	0.0047	0.0031	0.0478	0.0518
2019	0.06	0.006	0.0036	0.0443	0.0684
2020	0.0611	0.0084	0.0034		0.0433
2021	0.0578	0.0092	0.0044		0.0636

Table 3: The probability of pest freedom after the last annual survey considering that the aim of the survey was to provide evidence of pest freedom for trading partners (i.e., trade facilitation survey). For the time periods covered in the surveys, see Table 1 or 2.

The mean time between PWN invasions	Estonia	Finland	Lithuania	Norway	Sweden
2	0.8485	0.7827	0.0271	0.0002	0.5926
3	0.9075	0.8742	0.0732	0.005	0.7471
4	0.9335	0.9115	0.1201	0.0227	0.8169
5	0.9482	0.9317	0.1616	0.0545	0.8566
6	0.9575	0.9444	0.1969	0.0955	0.8821
7	0.964	0.9531	0.2267	0.1403	0.8999
8	0.9688	0.9595	0.252	0.1852	0.9131
9	0.9724	0.9643	0.2735	0.2281	0.9231

10	0.9753	0.9681	0.292	0.2679	0.9311
11	0.9777	0.9712	0.3081	0.3046	0.9376
12	0.9796	0.9737	0.3222	0.3379	0.9429
13	0.9812	0.9759	0.3345	0.3682	0.9475
14	0.9826	0.9777	0.3455	0.3957	0.9513
15	0.9838	0.9792	0.3553	0.4207	0.9546
16	0.9849	0.9806	0.3642	0.4434	0.9575
17	0.9858	0.9818	0.3721	0.4641	0.96
18	0.9866	0.9828	0.3793	0.4831	0.9623
19	0.9873	0.9838	0.3859	0.5005	0.9643
20	0.988	0.9846	0.3919	0.5164	0.9661
21	0.9886	0.9853	0.3974	0.5312	0.9677
22	0.9891	0.986	0.4025	0.5448	0.9692
23	0.9896	0.9867	0.4072	0.5574	0.9706
24	0.99	0.9872	0.4115	0.5691	0.9718
25	0.9904	0.9878	0.4156	0.58	0.9729
26	0.9908	0.9882	0.4194	0.5901	0.974
27	0.9911	0.9887	0.4229	0.5996	0.9749
28	0.9915	0.9891	0.4262	0.6085	0.9758
29	0.9918	0.9895	0.4293	0.6169	0.9767
30	0.992	0.9898	0.4322	0.6247	0.9775
31	0.9923	0.9902	0.4349	0.6321	0.9782
32	0.9926	0.9905	0.4375	0.6391	0.9789
33	0.9928	0.9908	0.4399	0.6457	0.9795
34	0.993	0.9911	0.4422	0.6519	0.9801
35	0.9932	0.9913	0.4444	0.6578	0.9807
36	0.9934	0.9916	0.4465	0.6634	0.9812
37	0.9936	0.9918	0.4484	0.6687	0.9817
38	0.9937	0.992	0.4503	0.6738	0.9822
39	0.9939	0.9922	0.4521	0.6786	0.9826
40	0.9941	0.9924	0.4538	0.6832	0.9831
41	0.9942	0.9926	0.4554	0.6876	0.9835
42	0.9943	0.9928	0.457	0.6918	0.9839
43	0.9945	0.993	0.4584	0.6958	0.9842
44	0.9946	0.9931	0.4599	0.6996	0.9846
45	0.9947	0.9933	0.4612	0.7033	0.9849
46	0.9948	0.9934	0.4625	0.7068	0.9853
47	0.995	0.9936	0.4638	0.7102	0.9856
48	0.9951	0.9937	0.465	0.7134	0.9859
49	0.9952	0.9938	0.4661	0.7166	0.9861
50	0.9953	0.994	0.4673	0.7196	0.9864

REPUBLIC OF ESTONIA AGRICULTURE AND FOOD BOARD
FINNISH FOOD AUTHORITY
SLU SWEDISH UNIVERSITY OF AGRICULTURAL SCIENCES
VKM Vitenskapskomiteen for mat og miljø
Norwegian Scientific Committee for Food and Environment

Table 4: The probability of pest freedom after the last annual survey considered that the aim of the survey was to detect possible PWN invasions early enough to enable successful eradication (i.e., early detection survey). For the time periods covered in the surveys, see Table 1 or 2.

The mean time between PWN invasions	Estonia	Finland	Lithuania	Norway	Sweden
2	0.0001	0	0.0254	0.0001	0.0132
3	0.0017	0.0005	0.069	0.0015	0.0491
4	0.0083	0.003	0.1137	0.007	0.094
5	0.021	0.0086	0.1533	0.0178	0.1381
6	0.039	0.0172	0.1871	0.0331	0.1781
7	0.0603	0.0283	0.2158	0.0514	0.2132
8	0.0833	0.041	0.2401	0.0713	0.2438
9	0.1069	0.0548	0.2609	0.0919	0.2704
10	0.1303	0.069	0.2788	0.1124	0.2937
11	0.1529	0.0833	0.2943	0.1324	0.3141
12	0.1745	0.0975	0.3079	0.1516	0.3322
13	0.1949	0.1113	0.32	0.17	0.3482
14	0.2142	0.1247	0.3306	0.1875	0.3625
15	0.2323	0.1376	0.3402	0.2039	0.3753
16	0.2493	0.15	0.3487	0.2195	0.3869
17	0.2652	0.1618	0.3565	0.2342	0.3974
18	0.2801	0.1731	0.3635	0.248	0.407
19	0.2941	0.1839	0.3699	0.2609	0.4157
20	0.3072	0.1941	0.3758	0.2732	0.4237
21	0.3195	0.2038	0.3811	0.2847	0.431
22	0.331	0.2131	0.3861	0.2956	0.4378
23	0.3419	0.2219	0.3907	0.3059	0.4441
24	0.3521	0.2303	0.3949	0.3156	0.45
25	0.3618	0.2384	0.3989	0.3248	0.4554
26	0.3709	0.246	0.4025	0.3335	0.4605
27	0.3795	0.2533	0.406	0.3417	0.4652
28	0.3877	0.2602	0.4092	0.3496	0.4696
29	0.3954	0.2669	0.4122	0.357	0.4738
30	0.4027	0.2732	0.415	0.3641	0.4777
31	0.4097	0.2793	0.4177	0.3708	0.4814
32	0.4163	0.2851	0.4202	0.3773	0.4849
33	0.4226	0.2907	0.4226	0.3834	0.4882
34	0.4287	0.2961	0.4249	0.3892	0.4913
35	0.4344	0.3012	0.427	0.3948	0.4943
36	0.4399	0.3061	0.429	0.4002	0.4971
37	0.4451	0.3109	0.431	0.4053	0.4998

38	0.4501	0.3154	0.4328	0.4102	0.5023
39	0.4549	0.3198	0.4345	0.4149	0.5047
40	0.4595	0.324	0.4362	0.4194	0.507
41	0.4639	0.328	0.4378	0.4237	0.5092
42	0.4681	0.332	0.4393	0.4279	0.5114
43	0.4722	0.3357	0.4408	0.4319	0.5134
44	0.4761	0.3394	0.4422	0.4358	0.5153
45	0.4798	0.3429	0.4435	0.4395	0.5172
46	0.4834	0.3463	0.4448	0.4431	0.5189
47	0.4869	0.3495	0.446	0.4465	0.5207
48	0.4903	0.3527	0.4472	0.4498	0.5223
49	0.4935	0.3558	0.4483	0.4531	0.5239
50	0.4966	0.3588	0.4494	0.4562	0.5254

Annex E – Description of capacity building and impact assessment of the project

The main aim of the project "Assessing the confidence in pest freedom gained in the past pine wood nematode surveys" (NoBa-PWN) was to build capacity for statistical analysis and planning of quarantine pest surveys, especially in the participating countries. The impact of the project was evaluated based on how well this aim was achieved. The description of capacity building and the achievements of the project are listed below.

E.1. Increasing the capacity for statistical analysis and planning of quarantine pest surveys

To increase the capacity for statistical analysis and planning of quarantine pest surveys in the participating countries, the project organized five workshops for the project members on various themes related to the subject. The workshops allowed the exchange of experience and expertise on statistical analysis and planning of surveys among the partner organizations. The project also organized a webinar on one of the project deliverables, namely the NoBa LCR application, for plant health officials of the participating countries.

In addition, openly accessible web applications NoBaSURV-PWN and NoBa Land Cover Retriever (NoBa LCR) were developed in the project. NoBaSURV-PWN enables assessments of the confidence in pest freedom gained in past official PWN surveys and the NoBa LCR facilitates the retrieval of CORINE Land Cover data needed in the statistical assessment and planning of quarantine pest surveys in the participating countries. Furthermore, a web application for estimating the apparent prevalence of *Bursaphelenchus mucronatus* in wood and *Monochamus* samples was developed for the needs of the project.

In order to enhance the utilization of Expert Knowledge Elicitation (EKE) in the project, project members underwent training through an online course "Steering an Expert Knowledge Elicitation (EKE)", provided by Lund University and hosted by EFSA. Although EKE was ultimately not used in the project, the course significantly strengthened the project members' capacity to apply EKE techniques in their future work.

The capacity built in the project was concretely applied by assessing the confidence in pest freedom gained in the past PWN surveys of the participating countries with the web applications developed in the project. The results of these assessments are presented in this report (section 3).

To increase the capacity for statistical analysis and planning of quarantine pest surveys among the global community on pest surveillance, comprehensive instructions on how to assess the statical confidence of past PWN surveys with the web-application built in the project (NoBaSURV-PWN) were written as part of this report (section 2). Furthermore, all materials from the workshops and webinars, as well as the source codes of the web applications, were published in EFSA's Knowledge Junction repository in Zenodo, and the project and its deliverables were presented at several national and international seminars and events.

E.2. Project deliverables

E.2.1. Workshops and webinars, and their materials

- Workshop on survey designs in the participating countries, 4 October 2021, online. Only project members participated in this workshop. In the workshop the PWN survey designs in the partnering countries were presented and discussed.
- Introduction to the assessment of the statistical confidence of pine wood nematode surveys, 29–30 November 2021, Helsinki, Finland. Only project members participated in this workshop. In the workshop several topics, such as terms and concepts, design prevalence, the sensitivity of annual surveys and the probability of freedom after the last annual survey were presented and discussed. In addition, the workshop included a practical exercise with the FinnSURV-Assess PWN web application (https://doi.org/10.5281/zenodo.3842358). Materials of the workshop are available at https://doi.org/10.5281/zenodo.5819713.
- Workshop on design prevalences, prevalence estimation, risk-based survey option etc., 30–31 May 2022, Saku, Estonia. Project members, and Olaf Mosbach-Schulz from EFSA, participated in this workshop. During the workshop, several topics were presented and discussed, including the 2-step approach, design prevalence, risk-based survey design options, and the relative risk of the risk factor levels. Additionally, the web application for estimating the apparent prevalence of *B. mucronatus* in wood and *Monochamus* samples was demonstrated. By utilizing the application, the prevalence of *B. mucronatus* in wood and *Monochamus* samples in the participating countries was estimated. Furthermore, the first version of the NoBa LCR was presented and tested during the workshop. Olaf Mosbach-Schulz highlighted the benefits of using EKE and suggested some steps in the analysis where the method could be applied. However, the parameters for which EKE was planned to be used in the project (i.e., the density of wood objects suitable for sampling and the density of *Monochamus* adults), were not needed for the assessments (since they were done assuming sampling from the Poisson distribution) and therefore EKE was not used in the project. Materials of the workshop are available at https://doi.org/10.5281/zenodo.6602388.
- Webinar on NoBa Land Cover Retriever application, 15 November 2022, online. The webinar was specifically targeted for plant health officials from the participating countries. In total, over 20 participants attended the webinar. The main focus of the webinar was to introduce and demonstrate the NoBa LCR application. Time was also allocated for comments and open discussion. The overall feedback received from the participants was highly positive, with plant health officials expressing their appreciation for the application's practicality in their work related to plant health surveys. Materials of the webinar are available at https://doi.org/10.5281/zenodo.7325787.
- Workshop on determining design prevalences, 9 January 2023, online. The workshop was intended for the project members only, with the main objective of preparing the protocol for obtaining the design prevalences from risk managers in the respective countries. Materials of the workshop are available at https://doi.org/10.5281/zenodo.7500787.
- Workshops/meetings with risk managers on setting the design prevalences, in all participating countries, several dates between 10 January and 14 February 2023. 1–2 workshops/meetings were organised per partnering country to i) explain the methodology and terminology and ii) set the design prevalences. The set design prevalences were subsequently utilized in the assessments conducted as part of the project.

- Workshop on NoBaSURV-PWN application, 15 February 2023, online. Project members, and Sybren Vos from EFSA, participated in this workshop. In the workshop, the NoBaSURV-PWN web application was introduced and demonstrated. After the workshop, the sensitivity of annual PWN surveys and the probability of pest freedom after the last annual survey were assessed with the application for each partnering country.

E.2.2. Web applications and their source codes

- NoBaSURV-PWN - A tool for assessing the confidence in pest freedom gained in official pine wood nematode surveys. Finnish Food Authority, Helsinki, Finland. The app is available at https://nobasurv-pwn.rahtiapp.fi/ and the source code at Zenodo https://doi.org/10.5281/zenodo.7766617.
- NoBa Land Cover Retriever - A tool for retrieving land cover data needed in statistical assessment and planning of quarantine pest surveys. The app is available at http://www.noba-lcr.rahtiapp.fi/ and the source code at Zenodo http://doi.org/10.5281/zenodo.7560539.
- A tool for estimating the apparent prevalence of *B. mucronatus* in wood and *Monochamus* samples. The app is available at https://b-mucronatus-prevalence-estimation.rahtiapp.fi/ and the source code at Zenodo http://doi.org/10.5281/zenodo.7800771.

E.2.3. Posters and presentations

- Presentation of the project in the annual meeting of the Nordic-Baltic plant health officials (7 December 2021, online)
- Presentation of the project in meeting of the EFSA Scientific Network for Risk Assessment in Plant Health (10 December 2021, online)
- Poster of the project in ONE 2022 conference (21–24 June 2022, Brussels & online)
- Poster of the project in the meeting of the International Pest Risk Research Group (10-13 October 2022, Athens)
- Presentation of NoBa Land Cover Retriever app in a webinar for plant health officials (15 November 2022, online)
- Poster of the project in Ruokavirasto’s Science Day (23 November 2022, Helsinki & online)
- Presentation of NoBa Land Cover Retriever app in meeting of the EFSA Scientific Network for Risk Assessment in Plant Health (9 December 2022, online)
- Presentation of NoBa Land Cover Retriever app in the annual training for plant health inspectors in Finland (24–25 January 2023, online)
- Presentation of NoBa Land Cover Retriever app in the EPPO workshop for inspectors on risk-based sampling and inspection (26–28 April 2023, Bern)

E.2.4. Other deliverables

- Data on the Nordic-Baltic pine wood nematode surveys published at https://doi.org/10.5281/zenodo.7793987

E.3. Project sustainability and long-term impacts

The field of statistical analysis and planning of quarantine pest surveys is a young discipline that still has significant room for growth and development. In that spirit, this project was initiated to focus on building capacity, especially in the participating countries but also elsewhere, by developing and applying easy-to use tools and identifying future challenges in the field, and its sustainability and long-term impacts within the discipline have the potential to become substantial. One contributing factor to its lasting influence is the open availability of all project deliverables, promoting transparency, enabling reproducibility, and facilitating the integration of the project findings into future research and practices.

The partnering organizations are well-positioned to share the knowledge and skills acquired in the project through various avenues, such as organizing workshops and training programs dedicated to the topic. For instance, webinars on NoBa LCR and NoBaSURV-PWN applications will be organized in 2023 and 2024 as part of the EFSA tailor-made activity "Training on tools and methods for risk assessment", which falls under the Focal Point framework. These dissemination efforts ensure that the project outputs reach a broader audience, benefiting professionals and stakeholders beyond the initial project participants.

Furthermore, the established collaboration between the partnering organizations can extend into the future. This ongoing collaboration can take the form of joint research projects and continued knowledge exchange. By maintaining this partnership, the project's impact can be sustained, allowing for further advancements in the field of statistical analysis and planning of quarantine pest surveys.