

PALEO-AKTUEEL

Het Groninger Instituut voor Archeologie presenteert zijn onderzoek

32



In dit nummer oa

Een uitzonderlijk kralensnoer uit de late bronstijd
Kledij van de bewoners van het terpengebied in de vroege middeleeuwen

Een raadselachtige Romeinse koepeloven
De identiteit en ondergang van de bekendste tjalk van Flevoland

Met de jaarlijkse uitgave van *Paleo-aktueel* geven medewerkers en studenten van het Groninger Instituut voor Archeologie en geassocieerde onderzoekers inzicht in recent of lopend onderzoek.

Aan dit nummer werkten mee: Stijn Arnoldussen, Peter Attema, Xandra Bardet (correctie Engelse samenvattingen), Siebe Boersma (vormgeving en omslagontwerp), Nathalie Brusgaard, René Cappers, Martijn Eickhoff, Merit Hondelink, Hans Huisman, Jos Kleijne, Flip Kramer (redactie-coördinatie), Martijn van Leusen, Arnoud Maurer, Johan Nicolay, Richard Paping, Daan Raemaekers en Mans Schepers.

Foto omslag: Remco Bronkhorst, Peter Attema en Steve Hayward tijdens de systematische survey van Mandra Vecchia (foto W. de Neef, UGent). Zie artikel Attema *et al.*

ISBN 9789493194533

ISSN 1572-6622

Website: www.paleo-aktueel.nl

Adres van de redactie

Rijksuniversiteit Groningen
Groninger Instituut voor Archeologie (GIA)
Poststraat 6 9712 ER Groningen
Tel.: 050 363 6712
gia@rug.nl

Adres van de uitgever

Barkhuis Publishing
Kooiweg 38 9761 GL Eelde
Tel. 050 3080936 fax 050 3080934
info@barkhuis.nl www.barkhuis.nl



**rijksuniversiteit
 groningen**

**groninger instituut
 voor archeologie**

© GIA.

www.rug.nl/let/onderzoek/onderzoekinstututen/gia/publications

Paleo-aktueel 32

Rijksuniversiteit Groningen / Groninger Instituut voor Archeologie (GIA)
University of Groningen / Groningen Institute of Archaeology
& Barkhuis Publishing
Groningen, 2022

Inhoud

VETWINNING IN DE SWIFTERBANTCULTUUR. EEN INTERPRETATIE VAN EEN VINDPLAATS TE NIEUWEGEIN-BEATRIXSLUIS Joshua Veldhuis	1
EEN BRONZEN HIELBIJL UIT EMMEN Stijn Arnoldussen, Eduard de Jonge, Bertil van Os, Gert van Oortmerssen & Hannie Steegstra	7
HOE BETROUWBAAR ZIJN ONZE ARCHEOLOGISCHE VELDVERKENNINGEN IN ITALIË EIGENLIJK? Suzanne Hoen, Jildou Bruinsma, Evelien Witmer & Martijn van Leusen	13
BOTANISCHE MACRORESTEN UIT SLOTEN ALS SPIEGEL VOOR HET LANDSCHAP Marjon Huiting, Annisa Spier & Mans Schepers	23
HET POLLINO ARCHAEOLOGICAL LANDSCAPE PROJECT (PALP) Peter Attema, Wieke de Neef, Antonio Larocca & Arnoud Maurer	33
KOPER, LOOD EN TIN - EEN UITZONDERLIJK KRALENSNOER UIT DE LATE BRONSTIJD UIT BORGER Wijnand van der Sanden & Bertil van Os	43
OP EEN VEELBEWOOND EILAND: EEN (DRONE) PILOTSTUDIE NAAR EFFECTIEVE SITEKARTERING IN ARCTISCH CANADA Jelke Take	51
EEN RAADSELACHTIGE ROMEINSE KOEPELOVEN UIT HENGELO (OV.) Adrie Ufkes	61
KLEDIJ VAN DE BEWONERS VAN HET TERPENGEBIED IN DE VROEGE MIDDELEEUWEN Anne Vrielink	69
DE ZEEHOND OF FIDUCIE? UITSLUITSEL OVER DE IDENTITEIT EN ONDERGANG VAN DE BEKENDSTE TJALK VAN FLEVOLAND Yftinus van Popta	79
CONTAMINATIE IN DE QUEEN ANNE? EEN EERSTE AANZET TOT HET VERKLAREN VAN DE AANWEZIGHEID VAN WILDE PLANTENRESTEN IN SCHEEPSWRAKKEN Morvenna van Rijn & Yftinus van Popta	89

DE VERDWENEN EEMS, EEN PARTICIPATIEPROJECT IN HET GRENSGEBIED Stijn Arnoldussen, Jana Esther Fries, Henny Groenendijk, Marion Heumüller, Hans Peeters & Wim Vuijk	101
STENEN VAN BETEKENIS: HUNEBEDDEN ALS ERFGOED EN ALS ACTOR Daan Raemaekers, Femke Bosscher & Harrie Wolters	109

Hoe betrouwbaar zijn onze archeologische veldverkenningen in Italië eigenlijk?

Suzanne Hoen¹, Jildou Bruinsma², Evelien Witmer³ & Martijn van Leusen⁴

Systematische en intensieve veldverkenningen spelen sinds de jaren 80 van de vorige eeuw een belangrijke rol in het karteren van de oppervlaktearcheologie van grote delen van de mediterrane wereld. Een recent uitgebreid overzicht wordt gegeven in Attema *et al.* 2020. Hoewel de benadering zelf al veel langer bestaat - al in de jaren 30 van de vorige eeuw voerden Duitse archeologen in Noordwest-Europa zgn. *Landesaufnahmen* uit - was de uitvoeringswijze 'extensief' of 'topografisch', wat wil zeggen dat er gericht gezocht werd naar substantiële (vaak monumentale) vindplaatsen door specifieke delen van het landschap zoals heuveltoppen te bezoeken. Er was daarom weinig reden om aandacht te besteden aan de zichtbaarheidsproblematiek: ook in moeilijk terrein kon zo'n substantiële vindplaats bijna altijd wel gedetecteerd worden.

Vanaf begin jaren '70 gaan, onder invloed van de zich dan versnellende rationalisering van de landbouw en uitbreiding van bebouwde kommen, systematische en vlakdekkende veldverkenningen een belangrijke rol spelen in het erfgoedbeleid in de Verenigde Staten en delen van Europa (met Stichting RAAP als Nederlandse voorloper). Omdat de te karteren pre- en protohistorische resten vaak weinig substantieel zijn, komt er dan veel meer aandacht voor de methodologische aspecten van het veldverkennen en ontwikkelt zich de systematische *non-site survey*, waarbij de dichtheid en samenstelling van artefacten over het hele landschap wordt gekarteerd. Clive Orton's handboek *Sampling in Archaeology* (2000) is de neerslag van deze *New Archaeology*-benadering, die vanaf begin jaren '80 ook in de mediterrane wereld steeds vaker wordt toegepast. De bekende

grote Groningse veldverkenningprojecten in Midden- en Zuid-Italië (Pontine Region Project, ca. 1990-2012; Raganello Archaeological Project, 2000-2009; zie Attema *et al.* 2010) vallen in methodologisch opzicht te herleiden tot Graeme Barkers (1995) *Biferno Valley Survey*.⁵ Op enkele technologische verbeteringen na (met name de nauwkeuriger plaatsbepaling met behulp van GPS) zijn doel en benadering van de archeologische veldverkenning in de mediterrane wereld sindsdien nauwelijks veranderd: namelijk, het registreren van een gedetailleerd, representatief en compleet beeld van de oppervlaktearcheologie van de gekozen regio door middel van het systematisch veldlopen door, in bijna alle gevallen, teams bestaande uit archeologiestudenten en materiaalspecialisten.

Naarmate de veldverkenneren zich meer gingen interesseren in het voorkomen van archeologische materialen in lage dichtheden, werden vragen over de betrouwbaarheid van de gebruikte methoden belangrijker. De fictie dat veldlopers steeds in staat zijn om betrouwbare waarnemingen te doen, is niet meer dan dat: een fictie. Alle praktiserende surveyarcheologen zijn het erover eens dat een reeks factoren samenspannen om de representativiteit van de genomen steekproef te ondermijnen (Attema *et al.* 2020). De meest algemeen erkende factor is de vegetatiegraad, maar ook factoren als bodemkleur, stenigheid, zonnestand en -intensiteit en wijze van bodembewerking kunnen het resultaat van een veldverkenning vertekenen. Het gaat dus niet alleen om kwantitatieve effecten (waarin een bepaald percentage van het oppervlaktearchief niet wordt waargenomen), maar ook en vooral om kwalitatieve (waarbij de gedane



Fig. 1. Een impressie van de veldexperimenten in de Raganellovallei uit 2014. In het eerste experiment (boven) voeren de deelnemers de verkenning volgens de standaardmethode uit en worden de vondsten per individueel transect geregistreerd. In het tweede experiment (onder) worden eerst alle vondsten gemarkeerd die de op standardsnelheid lopende deelnemers opmerken, waarna deze verzameld en geregistreerd worden op basis van de afstand tot de centrale lijn van het transect; vervolgens wordt hetzelfde transect nogmaals met de 'stofzuigmethode' gelopen om alle gemiste vondsten alsnog te registreren.

waarnemingen niet representatief zijn voor het werkelijke oppervlakte-archief). Kennis over de betrouwbaarheid van de eigen onderzoeksmethode kan op twee manieren verzameld worden: in gecontroleerde experimenten waarbij de imperfecte detectie aan de hand van eerder 'ingezaaide' materialen wordt gemeten, en in veldexperimenten waarbij de detectiegraad van de eigen methode bepaald wordt aan de hand van een 100% opname

van hetzelfde gebied. Dergelijke experimentele studies begonnen al in de vroege jaren 80, met Stephen Shennans pogingen om de betrouwbaarheid van zijn eigen East Hampshire Survey te bepalen (Shennan 1985); twee andere belangrijke experimentele studies zijn die van Robert Schön (2002) en Edward Banning (2006, 2011). Helaas kunnen de uitkomsten van die studies niet direct gebruikt worden om de uitkomsten van bestaande

verkenningprojecten, zoals die van het GIA in Midden- en Zuid-Italië, te corrigeren: daarvoor zijn de methodologische verschillen tussen de projecten te groot en ontbreekt de noodzakelijke standaardisatie van de documentatie. Om die reden heeft de laatste auteur in 2014 met hulp van tweedejaarsstudenten zelf twee veldexperimenten uitgevoerd, waarover wij hier rapporteren.

De veldexperimenten van juli 2014

De experimenten werden in de zomer van 2014 uitgevoerd in het kader van het tweedejaarsvak Veldwerk II van de vakgroep Klassieke en Mediterrane Archeologie van de Rijksuniversiteit Groningen in het stroomgebied van de Raganello, een rivier in het uiterste noorden van de regio Calabrië – de ‘teen’ van de Italiaanse laars (fig. 1). Het eerste experiment had tot doel te bepalen in hoeverre verschillen in de ervaring en vaardigheden van individuele veldlopers een rol spelen in de uitkomsten van een standaard veldverkenning. De zeven deelnemers hebben daarvoor in totaal 280 standaardtransecten van 50 meter lengte, verdeeld over verschillende landschaps- en landgebruikstypen, met de standaard wandelsnelheid gelopen en volgens standaard protocol alle vondsten groter dan 1 cm² verzameld. Begin- en eindtijd en de identiteit van de deelnemer werden daarbij geregistreerd.⁶

Al deze vondsten zijn vervolgens als volgt geëvalueerd om de verschillen tussen de deelnemers aan de standaardsurvey te kunnen bestuderen:

- Op hoofdtype - 1 Daktegels en opslagvaatwerk, 2 Gedraaid gebruiksaardewerk uit de klassieke periode, 4 Handgemaakt aardewerk uit de protohistorie, en 5 Fijne waar uit de klassieke periode. Overige vondsttypen (steen, obsidiaan, bot, metaal, glas en modern aardewerk) zijn wel geregistreerd maar niet meegenomen in de analyse.

- Op grootte – de vondsten werden ingepast in cirkeldiagrammen met doorsnedes van 1.25, 2.5, 5, 10 cm of groter dan 10 cm. Vondsten kleiner dan klasse 1.25 cm zijn uitgesloten van de analyse.
- Op hoofdkleur – Zwart, Grijs, Donkerbruin, Rood, Oranje, en Bleek. Veel twijfelgevallen, waarbij de kleur bijvoorbeeld als Rood-Bruin is genoteerd, moesten om die reden uitgesloten worden van de analyse.
- Op aanwezigheid van vorm of decoratie – bodem-, rand- en oorfragmenten, of met decoratie of glazuur, werden aangemerkt als diagnostisch.

Het tweede experiment had tot doel de zgn. detectiecurves van de deelnemers te bepalen, die aangeven hoe het aandeel waargenomen vondsten daalt in relatie tot de kijkafstand. Banning (2006) heeft dergelijke curves al eens in een kunstmatige setting bepaald, maar de uitkomsten daarvan leken weinig realistisch. De deelnemers aan ons experiment liepen hiervoor langs 35 met een meetlint uitgezette transecten van 50 meter en markeerden alle vanaf die lijn waargenomen vondsten met plastic lepels; daarna werden deze vondsten verzameld in stroken van 25 cm breed, gerekend vanaf de centrale lijn, en geëvalueerd als in experiment 1. Vervolgens zijn exact dezelfde transecten opnieuw gelopen, maar nu om met een ‘stofzuigmethode’ alle resterende vondsten te verzamelen en te classificeren.

Resultaten

Een hoofddoel van de experimenten was het achterhalen van de representativiteit van de standaardverkenning: idealiter zou elke vondstgroep een even grote kans moeten hebben om in een standaardverkenning opgeraapt te worden, want dan zouden de verzamelde vondsten representatief zijn voor het oppervlaktearchief als

Fig. 2. Van boven naar beneden:

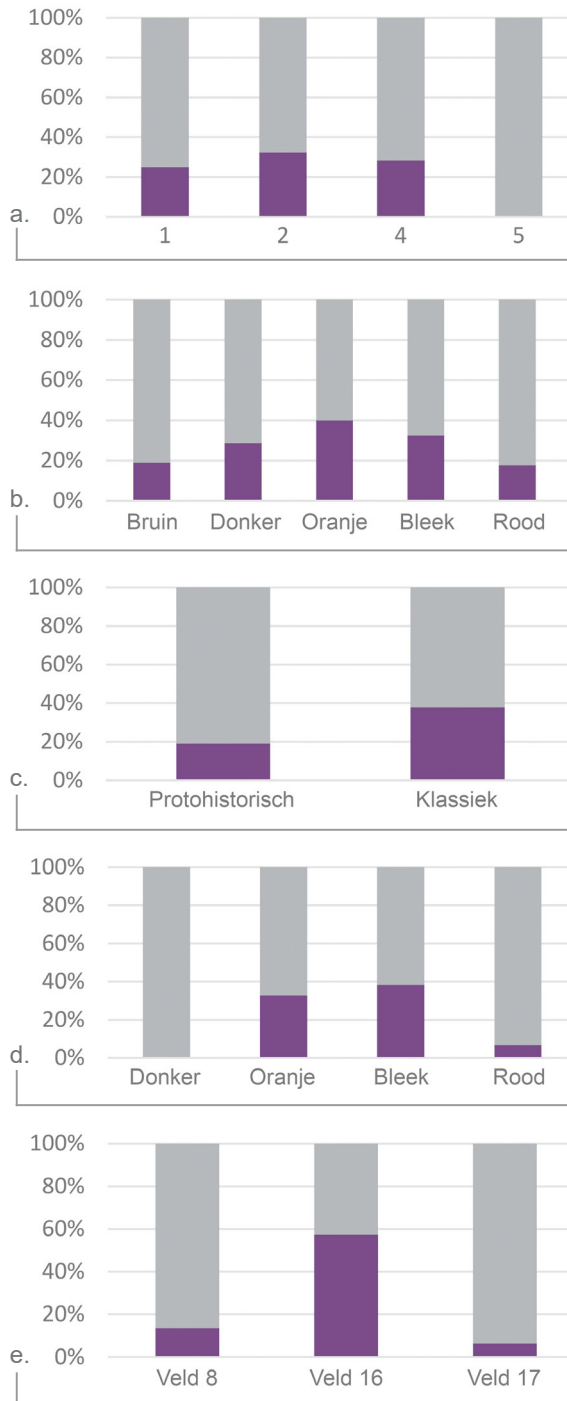
a. Verhoudingen tussen standaard- en totale survey van de 35 transecten van experiment 2, per vondstklasse (1 Bouw- en opslagmaterialen, 2 Gedraaid ruw en gedepureerd aardewerk, 4 Handgevormd aardewerk, en 5 Fijne waar).

b. Aandeel van het oppervlaktearchief dat per kleurgroep in de standaardmethode wordt opgepakt (paars).

c. Dezelfde gegevens gegroepeerd per tijdsperiode.

d. Percentage opgepakte scherven per kleur - veld 17.

e. Aandeel met standaardmethode gevonden diagnostische scherven ten opzichte van het totaal, voor drie velden uit experiment 2.



geheel. Uit experiment 2 (fig. 2a) blijkt dat van de belangrijkste groepen tussen 25 en 32% van het totaal wordt verzameld. Er is geen verontrustend verschil tussen de klassieke vondstgroepen 1 en 2 en het protohistorische aardewerk 4; het aantal vondsten bij 5 is te laag om conclusies uit te trekken. Hierdoor kan de samenstelling van de standaardverzameling vergeleken worden met die van het totale oppervlaktearchief.

Het voorgaande lijkt in tegenspraak met een ander resultaat uit experiment 2, namelijk dat vondstgroepen die qua kleur sterk afwijken van die van de bodem eerder worden opgeraapt dan groepen waarvoor dat niet geldt. Uit figuur 2b en c blijkt dat het klassieke oranje en bleek-beige aardewerk een ongeveer tweemaal hogere kans heeft om in de standaardverkenning opgeraapt te worden dan het donkerbruin-rode aardewerk uit de protohistorie (38 tegen 19 procent).

Dit 'bevoordelen' van klassieke vondsten lijkt zelfs nog sterker te gebeuren onder slechte zichtbaarheidsomstandigheden: onze gegevens voor veld 17 (fig. 2d) laten zien dat de detectie van oranje en bleek materiaal daar weinig onder lijdt, terwijl donkergekleurde scherven helemaal niet meer gevonden worden.

Naast kwantitatieve representativiteit hebben we ook behoefte aan het dateren van de aangetroffen vondstscatters door middel van zgn. diagnostische vondsten – en hoe meer we er daarvan vinden, hoe beter. Uit experiment 2 (fig. 2e) blijkt dat het verhogen van de surveyintensiteit, tot maximaal 100%, de hoeveelheid diagnostische vondsten kan verdubbelen tot vertienvoudigen. Dit resultaat rechtvaardigt de normale surveypraktijk, waarin eenmaal aangetroffen aardewerkconcentraties meteen of enkele dagen later opnieuw met hogere intensiteit onderzocht worden.

Looplijnen en detectiecurves

Met behulp van experiment 2 konden we ook bepalen hoe de vondstdetectie afhangt van de afstand tot de looplijn van de deelnemers – de zgn. detectiecurve. Meestal wordt bij benadering

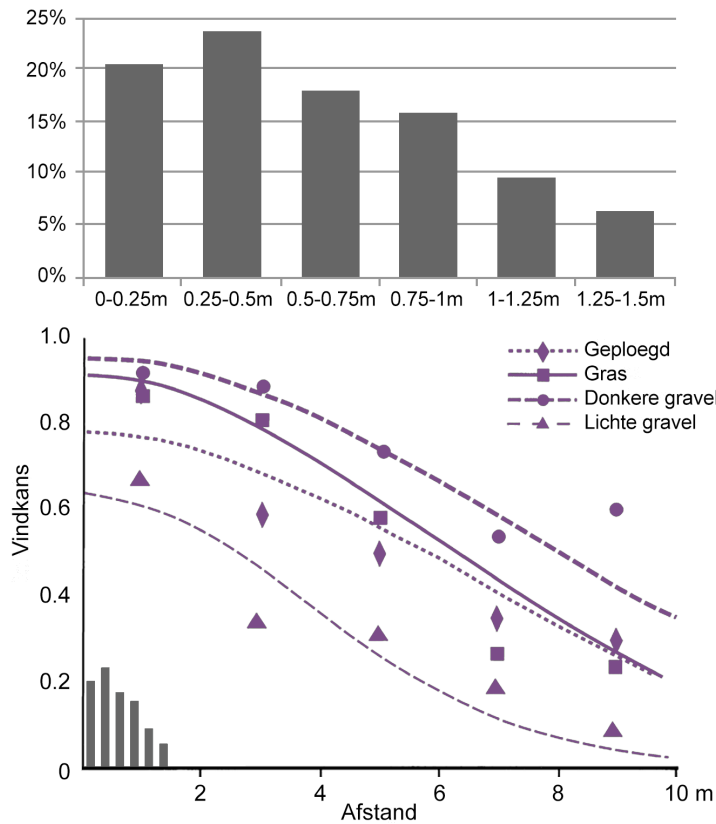


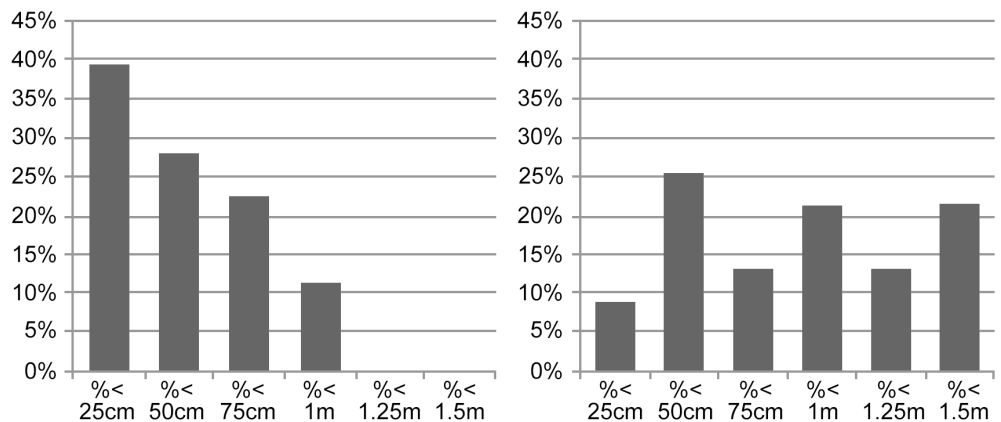
Fig. 3. Boven: detectiecurve van vondstaantallen in experiment 2. Onder: diezelfde curve ingepast in Bannings detectiecurves voor steengoed onder verschillende kunstmatige omstandigheden (Naar Banning et al. 2006, fig.7). Bannings detectiegraad is het hoogst op 'donkere gravel' en het laagst op 'lichte gravel'.

aangenomen dat 100% van de vondsten binnen 1m van de centrale lijn gedetecteerd wordt, maar dat is een fictie. In figuur 3 worden de door Banning in een kunstmatig experiment gevonden curves vergeleken met die uit ons eigen experiment. Er zijn grote verschillen: onze deelnemers detecteren al op 2 m vrijwel geen vondsten meer, de detectiegraad is overal minder hoog dan bij Banning, en het hoogste aandeel vondsten wordt niet op de looplijn zelf gedaan maar net ernaast, tussen 25 en 50 cm van de looplijn af. Dit laatste is mogelijk het gevolg van de neiging van de meeste deelnemers om al lopend de blik beurtelings naar links en rechts te wenden, waardoor de vondsten 'recht vooruit' er relatief bekaaid vanaf komen.

Ook in de experimenten van Robert Schön tijdens de Eastern Korinthia Archaeological Survey (2002, p. 230) bleken significante verschillen in detectie op te treden, waarbij afhankelijk van de zichtbaarheidsomstandigheden tussen de 10 en 17.5% meer vondsten gedaan werden binnen 2 m van de looplijn, dan tussen 2 en 4 m van de looplijn. Onze data zijn veel nauwkeuriger dan de zijne, en veel realistischer dan die van Banning.

In het voorgaande hebben we alle gegevens uit experiment 2 samengenomen, maar we kunnen ook onderzoeken of er interessante verschillen zijn tussen de detectiecurves van individuele deelnemers. Figuur 4 toont twee zeer verschillende

Fig. 4. Links de snel aflopende detectiecurve voor transect 17, rechts de meer evenredige voor transect 4.



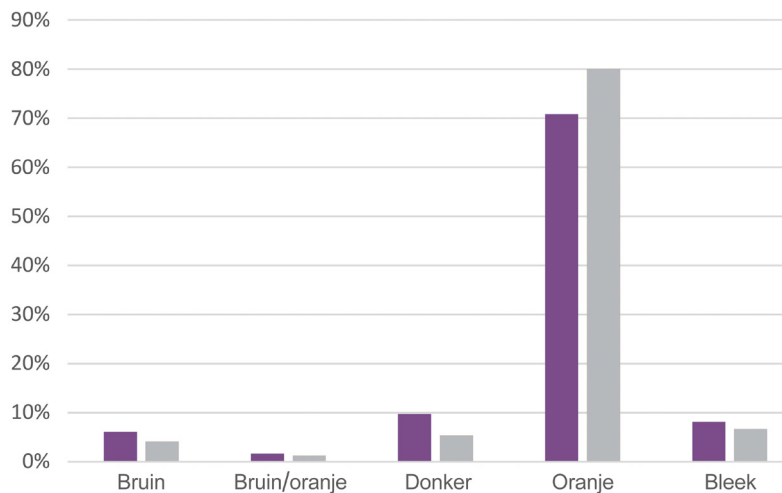


Fig. 5. Proporties van vondstkleuren, opgepakt door ervaren (paars) vs. onervaren (grijs) deelnemers.

detectiecurves waarbij de ene deelnemer zich concentreert op de dichtstbijzijnde delen van zijn transect, terwijl de ander zijn aandacht blijkbaar eerlijk verdeelt over de hele breedte van het transect.

Ervaring telt – soms

Mocht uit de experimenten blijken dat er belangrijke verschillen tussen ervaren en onervaren deelnemers zijn, dan zou dat aanleiding kunnen geven om de teams zorgvuldiger dan voorheen te trainen en samen te stellen. In eerste instantie leek dat echter niet het geval: met betrekking tot vondstgrootte, vondstgroep en diagnosticiteit zijn er in experiment 1 geen significante verschillen gevonden tussen beide groepen, en ook wanneer we individuele deelnemers met elkaar vergelijken zijn er geen consistente verschillen. Die verschillen bestaan er echter wel met betrekking tot de vondstkleur (fig. 5): ervaren deelnemers dragen vooral bij door de detectie van de relatief slecht zichtbare kleuren bruin en donker.

Dingen die we niet gepland hadden

Omdat van alle 17 experimentele velden ook de zichtbaarheid volgens de standaardmethode is geregistreerd, met een score van 1 (slecht) tot 5 (optimaal), kon een onverwachte bijdrage geleverd worden aan de veel gestelde vraag naar het effect

van de zichtbaarheid op de uitkomst van een veldverkenning. Zoals hierboven bleek, lijkt een lage zichtbaarheid vooral de al bestaande ‘voorkeur’ voor tegen de achtergrondkleur afstekende vondstgroepen te versterken. Andere verbanden hebben we niet kunnen vaststellen.

Omdat in experiment 1 steeds de begin- en eindtijd per transect geregistreerd zijn, hebben we ook kunnen vaststellen dat ervaren deelnemers de survey meer dan tweemaal zo snel kunnen uitvoeren als onervaren deelnemers. Dit gegeven kan van belang zijn wanneer het onderzoek vraagt om een hogere snelheid of een grotere dekkingsgraad.

Discussie

Omdat de hier besproken experimenten in een realistische veldsituatie hebben plaatsgevonden, konden niet alle variabelen onder controle gehouden worden. Velden verschillen onderling van elkaar, maar soms zijn er ook binnen een veld verschillen in de dichtheid en verspreiding van archeologica, net als in bodemkleur, stenigheid en zichtbaarheid. Er zijn dus meerdere verklaringen denkbaar voor sommige aangetroffen verschillen tussen de deelnemers en tussen de standaard- en totale verzamelingen. Idealiter zouden alle archeologica van tevoren ‘ingezaaid’ moeten zijn, zodat we precies zouden weten wat het effect van die veldvariabelen is – maar dan zou het experiment heel anders opgezet moeten worden.

Voor een correcte analyse van de experimentele gegevens moeten we de juiste statistische basis nog zien te bepalen. Bij veel analyses in dit artikel hebben we te kleine vondstgroepen die zeker zouden leiden tot insignificante uitkomsten al uitgesloten, maar voor een nog te verschijnen wetenschappelijke publicatie moet dit methodologische probleem nog worden opgelost.

Onze experimenten bevestigen wat iedere veldverkenner wel weet, namelijk dat met het verhogen van de surveyintensiteit ook de kans op het vinden van zgn. diagnostische vondsten toeneemt. Het heeft dus zin om, nadat uit de standaardverkenning is gebleken dat er ergens

een vondstconcentratie aanwezig is, deze in een vervolgstap intensiever te onderzoeken.

Om de gebleken lagere vindkans voor sommige vondstgroepen te corrigeren, zijn verschillende aanvullende strategieën denkbaar. In de Groningse veldverkenningen is vanaf 2000 de zgn. *adaptive sampling*-strategie (zie Orton 2000: 34-38) ingezet, waarmee bepaald kan worden of zeldzame protohistorische vondsten die tijdens de standaardverkenning aangetroffen zijn, wellicht in feite het ‘topje van de ijsberg’ zijn – dat wil zeggen dat ze eigenlijk onderdeel van een (in eerste instantie niet waargenomen) vondstconcentratie zijn. In ongeveer 30% van de gevallen waarin deze strategie is toegepast blijkt dit ook inderdaad het geval te zijn. Voor weer andere vondstgroepen is het misschien zelfs nodig om een aparte, specialistische survey uit te voeren: zo wordt vuursteen in een aardewerksurvey mogelijk stelselmatig over het hoofd gezien omdat de deelnemers zich instellen op het detecteren van bepaalde aan aardewerk voorbehouden vormen, kleuren en texturen.⁷

Conclusie

Het doel van de in dit artikel besproken experimenten was om te bepalen hoe betrouwbaar de uitkomsten van onze standaard systematische intensieve veldverkenningen in het mediterrane gebied eigenlijk zijn. Wij zien in de uitkomsten twee redenen tot zorg, en daarmee ook twee mogelijkheden voor verbetering. Ten eerste bevestigen de experimenten dat verschillen in ‘intrinsieke vindbaarheid’ ervoor zorgen dat sommige vondstgroepen (hier het handgemaakte aardewerk uit de protohistorie, maar het kan net zo goed gelden voor categorieën als vuursteen en vroegmiddeleeuws aardewerk) beduidend minder opgeraapt worden dan andere. Dit effect wordt versterkt onder lage zichtbaarheidsomstandigheden, maar kan beduidend verminderd worden door alleen gebruik te maken van ervaren deelnemers. Bovendien kan de verkenning daarmee ongeveer tweemaal sneller uitgevoerd worden.

Ten tweede zijn de verschillen tussen de individuele deelnemers beduidend groter dan we zouden wensen. De algemene surveytheorie (Orton 2000) gaat ervan uit dat alle deelnemers niet alleen een gelijke vaardigheid hebben in het detecteren van de voorkomende vondstgroepen, maar dat ze daarbij ook een even brede strook aan beide zijden van hun transect bestrijken. Uit experiment 2 is niet alleen gebleken dat die strook veel minder breed is dan Bannings experimenten deden vermoeden, maar ook dat deelnemers sterk kunnen verschillen in hoeveel aandacht ze besteden aan de verder weg gelegen delen van die strook.

Op basis hiervan kunnen we ons twee nieuwe experimenten voorstellen die meer of helderder licht kunnen werpen op de betrouwbaarheid van de standaardverkenning. Beide vragen echter wel een veel grotere investering in het opzetten en uitvoeren dan wij op dit moment kunnen verantwoorden. Ten eerste zouden in een nieuw experiment veel, zo niet alle verstorende variabelen uitgeschakeld kunnen worden door gebruik te maken van velden waarin alleen een van tevoren bekende, ingezaaide mix van vondstgroepen aanwezig is. Het effect van zichtbaarheidsfactoren in deze velden kan dan zuiver gemeten worden. Om statistisch significante resultaten te bereiken voor elke combinatie van zichtbaarheidsfactoren, die andere surveyleiders kunnen gebruiken om de vertekeningen van hun eigen teams beter in te schatten, zou hiervoor dan wel een groot aantal transecten verkend moeten worden.

In een tweede experiment zouden we de verschillen tussen individuele deelnemers beter kunnen gaan meten. Ook hiervoor zouden tevoren ingezaaide velden nodig zijn, maar ditmaal met een groter aantal deelnemers om te kunnen voldoen aan de statistische geldigheidseisen bij het onderzoeken van variabelen als hoeveelheid en aard van de ervaring, maar bijvoorbeeld ook lengte, conditie en zicht van de deelnemers. Ook het effect van training kan hiermee vastgesteld worden: wanneer zijn de deelnemers voldoende

ervaren om alle vondstgroepen met een gegeven, van tevoren bepaalde, mate van betrouwbaarheid te detecteren? In onze huidige dataset zien we bijvoorbeeld al aanwijzingen dat onervaren studenten in de loop van twee weken duidelijk beter worden in het vinden van het ‘lastige’ protohistorische aardewerk.

Dankwoord

De oorspronkelijke experimenten werden in 2014 door Martijn van Leusen en Evelien Witmer ontworpen en uitgevoerd met behulp van studenten Archeologie aan de universiteiten van Groningen en Gent, die we hierbij graag nog eens danken voor hun enthousiasme en inzet. De vondstbeschrijvingen, in eerste instantie uitgevoerd door meerdere teamleden, zijn naderhand gecontroleerd en gecorrigeerd door (destijds) promovenda Wieke de Neef, waarvoor wij haar zeer dankbaar zijn. Digitalisering en een eerste analyse van de gegevens is direct na de verkenning gedaan door Evelien Witmer, zodat de voorlopige resultaten van de experimenten al in september 2014 gepresenteerd konden worden op het Landscape Archaeology congres in Rome. Nadien heeft Jildou Bruinsma in het kader van haar bachelorscriptie zeer veel tijd besteed aan het opschonen en waar mogelijk corrigeren van de ingevoerde gegevens, het controleren van de door Witmer uitgevoerde analyses, en het doen van aanvullende analyses. In 2020 heeft Suzanne Hoen in het kader van haar masterscriptie ten slotte een deel van deze analyses nog eens nagelopen en aangevuld met enkele analyses van de verschillen tussen de deelnemers. In de loop van dit langdurige onderzoek zijn enkele malen deelresultaten gepresenteerd aan de International Mediterranean Survey werkgroep; wij zijn de leden van deze, tweemaal per jaar samenkomende internationale gemeenschap dankbaar voor hun vragen en kritiek.

Noten

1. Suzanne Hoen (Antea Groep, suzannehoen@outlook.com).
2. Jildou Bruinsma (x.jildou@gmail.com).
3. Evelien Witmer (RAAP Oost, evelienwitmer@gmail.com).
4. Martijn van Leusen (GIA, p.m.van.leusen@rug.nl).
5. In het oostelijke Middellandse Zeebekken werd de methodologische basis gelegd in de Boeotia Survey (Bintliff en Snodgrass 1985).
6. Tussen deze transecten in zijn nog 86 extra transecten verkend, waarbij de deelnemers de opdracht kregen zo langzaam te lopen als ze zelf nodig achtten om het transect goed te kunnen inspecteren. Deze gegevens, waaraan enige methodologische nadelen kleven, zijn nog niet geanalyseerd.
7. Om dit te onderzoeken is, eveneens in 2014, een speciaal vuursteenverkenningsexperiment uitgevoerd, waarover elders gepubliceerd zal worden.

How reliable are our Italian field surveys really?

This article presents the main results of three student-thesis based analyses of data obtained in field survey experiments carried out by the last author in Calabria (southern Italy) in 2014. Two separate experiments were conducted to study the differences between individual field walkers, and the ability of the GIA standard fieldwalking protocol to achieve a representative sample of the surface archaeological record. The analyses confirm and quantify the variability between walkers, as well as some of the survey biases that had already been suspected, and provide arguments for improving surveying practice in the Mediterranean. They also alleviate some concerns with respect to other parameters, such as sherd size.

Literatuur

- Attema, P., J. Bintliff, M. van Leusen, Ph. Bes, T. de Haas, D. Donev, W. Jongman, E. Kaptijn, V. Mayoral, S. Menchelli, M. Pasquinucci, S. Rosen, J. García Sánchez, L. Gutierrez Soler, D. Stone, G. Tol, F. Vermeulen, A. Vionis, 2020. A Guide to Good Practice in Mediterranean Surface Survey Projects, *Journal of Greek Archaeology* 5, 1-62.
- Attema, P., G.J. Burgers & P.M. van Leusen, 2010. *Regional Pathways to Complexity Settlement and Land-Use Dynamics in Early Italy from the Bronze Age to the Republican Period* (Amsterdam Archaeological Studies 15). Amsterdam, Amsterdam University Press.
- Banning, E.B., A. Hawkins & S. Stewart, 2006. Detection Functions for Archaeological Survey. *American Antiquity* 71(4), 723-742.
- Banning, E.B., A.L. Hawkins & S.T. Stewart, 2011. Sweep widths and the detection of artifacts in archaeological survey. *Journal of Archaeological Science* 38(12), 3447-3458.
- Barker, G., 1995. *A Mediterranean Valley: Landscape Archaeology and Annales History in the Biferno Valley*. London and New York, Leicester University Press.
- Bintliff, J.L. & A. Snodgrass, 1985. The Cambridge/Bradford Boeotian expedition: the first four years. *Journal of Field Archaeology* 12(2), 123-161.
- Orton, C., 2000. *Sampling in Archaeology* (Cambridge Manuals in Archaeology). Cambridge, Cambridge University Press.
- Schön, R., 2002. *Seeding the Landscape: Experimental Contributions to Regional Survey Methodology*. Bryn Mawr College, PhD-thesis.
- Shennan, S., 1985. *Experiments in the Collection and Analysis of Archaeological Survey Data: the East Hampshire survey*. Sheffield, University of Sheffield.

