目前，在预测水果农场的财务状况时，气候变化以外的因素更为重要。

Fabian Obster[1,2], Heidi Bohle[3] 和 Paul M. Pechan[3*]

[1] 慕尼黑联邦国防军大学工商管理系，德国慕尼黑[2] 慕尼黑路德维希-马克西米利安大学统计系，德国[3] 传播与媒体研究所，慕尼黑路德维希-马克西米利安大学，德国*通讯作者

关键词

气候变化，水果种植者，机器学习，预测分析，调查

摘要

本研究使用机器学习和统计建模方法，分析气候变化对突尼斯和智利水果种植者财务福祉的影响。分析基于对 801 名农民进行的面对面访谈。研究探讨了三个问题。首先，气候变化影响是否会影响农场的财务表现。其次，如果气候变化没有影响，那么哪些因素对于预测农场的财务福祉很重要。第三，确定观察到的农场财务福祉影响是否由预测变量之间的相互作用所致。这是首份直接比较气候变化与其他可能影响水果农场财务福祉因素的报告。 某些气候变化因素，具体而言是气温升高和降水减少，会在区域层面影响水果种植者自我感知的财务福祉。具体来说，气温升高和降水减少会对智利农场的财务福祉产生可测量的负面影响。这种影响在突尼斯较不明显。在智利观察到了气候影响的差异，但在突尼斯没有。然而，气候变化对于预测农场财务福祉仅有较小的重要性，尤其是对于财务状况已经良好的农场。更重要的因素主要包括突尼斯的对信息来源的信任和先前农场所有权。其他重要因素包括农场规模、所使用的水资源管理系统以及所种植水果作物的多样性。此外，在财务状况良好和不佳的农场之间，部分重要因素有所不同。因素间相互作用可能改善或恶化农场的财务福祉。结果表明，将监督式机器学习与经典统计工具结合起来进行农业领域的数据分析是有用的。

引言

气候变化对果树产量和农场经济福祉的影响

气候变化会影响作物，预计未来许多重要作物的产量将会下降 (1,2)。例如，气温升高会导致全球主要作物的产量下降 (3)。这类气候影响可能并将对粮食的供应和营养价值产生不利影响 (4)。本文重点关注气候变化对智利和突尼斯具有重要营养价值和经济价值的作物——樱桃和桃树 (5,6) 的影响。 这两种果树作物的生殖器官可能因气候变化而受损，导致可收获果实的数量和质量下降 (7,8,9)。因此，冬季气温升高会影响果树的需冷量，进而导致芽、花和坐果发生变化 (10,11,12,13,14)。同样，坐果和果实发育期间气温升高，会导致果实生长和成熟发生变化 (9,15,16)。高温与水资源可用性下降相结合，会影响果实产量和品质 (7,17,18)。这些影响会因果树栽培品种和物种而异。也有观察发现，极端事件（冰雹、风、霜冻）会影响物理环境并造成果树作物损害 (7,19,20,21)。 这些气候事件因地区而异，对各地粮食生产和作物的影响程度也不尽相同。

虽然可以估算气候变化对水果作物品质和产量的影响，但要预测／预估气候变化对农场收入（营收）的影响则困难得多，且结果具有很高的不确定性（22）。然而，经济估算至关重要，因为主要是农民需要动用自己的财务资源，采取必要的适应措施，以减轻气候变化（或其他任何潜在危害）对作物的影响。大多数气候变化经济影响模型，都是基于估算气候变化对作物产量的影响（依据气候和作物模拟模型），再将这些信息转化为对农场财务表现的可能预测。然而，这类分析基于多项假设，且很少将适应措施组合、社会经济因素以及区域差异等其他因素纳入考量（23）。衡量气候变化对农业影响时，最常用的经济模型之一是李嘉图方法，该方法关注土地价值和农业收入（24,25）。横截面回归分析和面板回归分析是首选的分析工具（26）。无论采用何种方法和分析类型，若遗漏了 可能直接或间接影响作物／农场收入／营收的变量，气候变化财务影响评估就会存在很高的不确定性。若能估算各变量对农场经济

表现的重要性，同时呈现分析的整体可预测性指标，将大大有助于降低这种不确定性。

机器学习在农业中的应用

预测分析基于现有数据对未来进行预测。在经典统计学中，回归分析也具有预测能力。然而，有些情况下，回归分析不足以处理生成的数据集或需要回答的特定问题，或者关于自变量与因变量之间存在线性函数的假设并不成立。这尤其适用于数据集中存在复杂变量交互作用的情况。在这种情况下，人工智能（AI），尤其是机器学习，便成为补充传统统计分析的实用工具（27,28,29）。除了发挥补充作用外，机器学习还能够同时分析大型数据集和多个变量，因此有助于降低重要变量在数据分析过程中被遗漏的可能性。

因此，机器学习能够为农业大数据分析提供新的可能性，也就不足为奇了（30,31,32,33,34,35,36,37）。事实上，机器学习在农业中的潜在应用范围十分广泛，尤其是与气候变化等其他研究领域结合时（38,39,40,41）。然而，解决农业问题十分复杂。例如，某种新的作物品种在特定气候条件下是否确实能提高产量和农场收入，不仅可能取决于其遗传性状，还可能取决于许多其他因素，例如与生物物理条件和农场管理相关的因素（42）。这意味着数据集中可能存在复杂而深层的交互作用。使用传统统计方法分析此类数据，很快便会变得困难重重。这正是机器学习的强大能力能够得到充分发挥之处（36,43）。分析中不仅可以纳入微气候效应、土壤结构和质量等生物物理变量，还可以纳入当地土地利用、城乡农场用水可及性、水资源相关政策、农场规模、人口统计数据和市场准入等社会经济变量，从而能够对农业价值链的每个环节进行分析（32, 44, 45）。仅针对一个农场，生成的数据集就可能包含数百万种数据点组合。重要的是，对此类数据进行分析可以回答适应措施是否能在特定气候条件下有效维持或提高作物产量和农场收入，提供干预措施对预期结果的相对重要性信息，并有助于未来的（产量）预测（38, 40, 46, 47, 48）。遗憾的是，单独采用机器学习方法可能会限制数据的可解释性。对于支持向量机、深度神经网络、随机森林或提升树等复杂机器学习算法而言，情况尤其如此。尽管

存在事后可解释性方法，可用于近似这类黑箱模型的运作方式，但要理解和解释导致结果的确切过程，并没有简单直接的方法。对于旨在深入理解与预期结果相关的过程或因素的研究问题而言，这可能是一个重大缺陷。为解决这一问题，我们在此提出一种（混合）方法，将基于广义线性模型的数据集分析与机器学习策略相结合，例如交叉验证、提升法和组变量选择。这种方法的输出保留了可解释性，遵循数据的分组结构，同时仍能与最先进的机器学习算法相竞争。本文的数据分析部分将详细介绍这一策略。

研究方法与目标

本文旨在评估气候变化可能对水果农场财务状况造成的影响。我们采用的方法是依靠农户自我报告，了解他们过去经历的气候变化，以及这些经历是否对农场的财务福祉产生了影响。相关信息通过面对面访谈收集。由于后续数据分析所用信息由农民提供，本文实际上报告的是农民感知到的财务福祉。

我们结合运用经典统计分析和机器学习，以考虑此类分析所需数据集的高维性，其中包括许多分组自变量。后者采用了一种新颖方法，可分析单个及成组的自变量，从而确定它们对结果变量的相对重要性和预测能力。

本文回答三个研究问题。第一，气候变化是否影响了农场的财务状况。第二，如果气候变化对农场财务福祉并不重要，哪些因素可能影响这一结果变量。第三，因素交互作用是否会影响农场财务福祉。

方法

研究区域

本研究选取了四个地理和气候条件各异的地区，其中两个位于突尼斯，两个位于智利。在突尼斯，选取的是 Mornag 和 Reueb 桃种植区。

在智利，选取的是 Rengo 和 Chillán 樱桃种植区。两国选定的地区在气候影响暴露的程度和类型上各不相同。

数据收集

本研究使用了两种数据收集工具：a）焦点会议；b）对智利樱桃种植户和突尼斯桃农进行面对面调查。要入选焦点会议和调查，农民必须拥有农场、管理并在农场工作，且农业活动收入占其总收入的70%以上。

调查

问卷准备

调查问卷以英语编制，并翻译成突尼斯阿拉伯语和智利西班牙语。为检查是否存在不一致之处，译文随后被回译成英语。我们与智利的 Qualitas AgroConsultores 和突尼斯的 Elka Consulting 协商，在12名农民中对调查进行了预测试。根据他们的反馈，以及突尼斯和智利研究同事的意见，我们删除了部分问题，并对其他问题进行了改写。

访谈方法

在突尼斯和智利选定地区随机抽取的樱桃种植户和桃农，经过预筛选，以符合数据收集资格标准（见上文）。随后，共对符合筛选标准的农民进行了 801 次面对面访谈–其中包括突尼斯的 401 名桃农（Mornag 地区 201 名、Regueb 地区 200 名）以及智利的 400 名樱桃种植户（Rengo 地区 200 名、Chillán 地区 200 名）。访谈时长约一小时，直接在农场对农民进行。收获结束后，Elka Consulting 于 2018 年秋季在突尼斯开展访谈，Qualitas AgroConsultores 则于 2019 年春季在智利开展访谈。 我们就调查实施及涉及人类受试者的数据使用事宜，向本研究所（ICMR、LMU）征求了指导意见。各国均按照适用的数据收集程序收集农场数据。调查参与者均为数据收集提供了知情同意。未收集任何可识别个人身份的数据，从而确保完全匿名。

测量

在面对面访谈中，研究人员要求农民回答与2009年至2018年期间气候变化及其对农场造成的气候影响相关的一系列选择题、开放式问题、李克特量表题和“是/否”题，并询问他们过去、目前及计划采取的适应

措施。汇总农户访谈数据后，对所得数据集进行了错误检查，并整合为 Excel 格式，以便进一步开展数据分析。

数据分析

待分析的变量

气候变化可能威胁智利和突尼斯农场的财务福祉。研究考察了四种不同的气候变化因素：气温、降水、极端天气和干旱。

- 气温。研究询问农民，过去10年间气温是上升、下降、保持不变，还是变得难以预测。
- 降水。研究询问农民，过去10年间降水是增加、减少、保持不变，还是变得难以预测。
- 极端天气。研究询问农民，过去10年间风、冰雹或霜冻等极端天气事件的发生频率是增加、减少、保持不变，还是变得难以预测。
- 干旱。我们询问农民，过去 10 年中干旱期或干旱的发生频率是增加、减少、保持不变，还是变得难以预测。

除气候变化外，某些群体或个别农场相关变量也可能单独或与气候威胁相互作用，从而影响农场财务福祉。分析中使用的所有变量、相应群组及调查问题概述见扩展数据图 1 和 2。

- 变量群组。我们将分析重点放在可能对农场财务福祉至关重要的农场变量群组（资产）上。这些群组包括：
    - o 自然因素（地理区域） o 人类因素（教育、年龄、性别、知识） o 社会因素（对信息的依赖／使用、对信息来源的信任、社区、科学或宗教） o 生物物理／人造因素（农场规模、农场使用的水资源管理系统、作物多样性、适应措施） o 经济因素（农场债务、农场绩效、对果园收入的依赖） o 气候经历 o 收入损失

上述变量的选择依据是五种资源／资本可持续性模型，该模型探讨可持续财富创造这一概念（49,50）。

- 个体变量。上述列出的分组变量也进行了单独评估。

  此外，我们还考察了其他变量，这些变量可能单独或与气候威胁相互作用，从而影响农场财务福祉。

- 因变量。**结果变量"农场的财务福祉"**分为三类：表现良好或非常好；表现不佳也不算差（**"中性"**）；以及表现不佳或很差。在整个分析过程中，财务福祉变量被编码为两个独立变量。第一个变量**称为"高幸福感"，用于比较财务状况良好**或非常好的农民与财务状况中性或不佳的农民（参照类别）；**第二个变量称为"低**幸福感**"**，用于区分财务状况不佳的农民与财务状况中性、良好或非常好的农民（参照类别）。这使我们能够区分导致农民财务状况不佳的过程和导致农民财务状况良好的过程，因为财务状况既不好也不差的农民始终属于参照类别。

分析策略

研究问题一：过去的气候变化经历是否会影响农场的财务福祉。

我们采用统计方法来确定自变量对农场财务福祉的影响。由于两个结果变量**"高幸福感"和"低幸福感"**均为二元变量，因此我们使用逻辑回归，分析了适应措施和过往经验对结果的优势比及其相关的p值和置信区间。

研究问题二：除气候变化外，哪些因素可能对农场的财务福祉很重要?

这一研究问题带来了重大挑战。数据集中存在许多可能产生影响的变量。其中一些可能与结果变量相关，另一些则可能无关。与结果变量无关的变量会造成不必要的背景**"噪声"**，因为广义线性模型在高维情况下往往会过度适应数据（即所谓的过拟合）。在极端情况下，如果自变量的数量多于观测值的数量，就无法拟合线性模型。解决这一问题的方法是进行变量选择，然后只将选出的变量纳入模型。目前的做法是依据文献和专家知识进行这项选择。事实上，数据的收集过程总是隐含着变量选择。不过，最终仍可能留下大量可能产生影响的变量。在这种情况下，可以结合统计学与机器学习来进行变量选择。我们采用了基于模型的提升（51），但也可以使用 Lasso（52）等其他策略

。基于模型的提升策略是通过仅添加最能提升整体模型效果的变量，来改进给定模型。如果进一步更新也无法得到“更好”的模型，就会停止添加变量。

重要的是，在某些情况下，变量组对模型而言可能比个体变量更重要。为此，我们采用了稀疏组提升法（53）。在稀疏组提升法中，模型可以在个体变量和变量组之间进行选择。由此可以针对所选变量或变量组与农场财务福祉之间的关联提出新的假设。区分变量组与个体变量的重要性，有助于问卷设计：如果个体变量比变量组更重要，那么问卷中只需纳入重要的个体变量。这样可以大幅缩短问卷，同时不损失信息。反过来，变量组可能会提供有关变量交互作用的信息。

研究问题三：农场财务福祉方面观察到的效应，是否源于调节效应和/或变量之间更复杂的关系？

我们分析了所有变量对农场财务福祉的（成对）交互效应。我们借助基于模型的提升法评估变量之间的交互作用，从而比较它们对结果变量的相对重要性。需要注意的是，如果数据集中有 p 个变量，那么数据集中可能存在 0.5* p*(p-1) 种交互作用，这会导致维度噪声问题进一步加剧。不过，这种穷举法有望识别出重要的调节效应或变量加性效应，从而增进我们对导致该结果的过程的理解。

根据所提出的研究问题，上文所述的数据分析复杂程度仍可能不够。在这种情况下，应使用不可解释的黑箱机器学习模型。将这些机器学习模型的预测性能与可解释的混合模型和统计模型进行比较，可以表明所需的分析复杂程度。如果混合模型在预测能力方面优于黑箱模型（即曲线下面积更高），则无需进一步提高复杂程度。反之，未来研究的目标应是探究如何解释这些复杂性，例如通过高度非线性关系或高阶交互作用来解释。

用于数据评估的模型

统计模型

我们使用广义线性模型（54）来回答干预措施是否对关注的结果产生了影响。由于结局变量为二元变量，因此采用逻辑回归计算优势比及其相应的p值和置信区间。

机器学习

我们比较了多种常用的机器学习模型，以确保用于分析的模型在预测能力方面具有竞争力。所有使用的模型列于扩展数据图3。与基于模型的提升模型和逻辑回归不同，这些机器学习模型无法帮助我们深入了解数据。

混合统计与机器学习预测模型

我们决定采用基于模型的提升来为预测模型筛选变量。提升迭代次数通过使用训练数据进行25折交叉验证来控制。这一超参数控制效应惩罚（平滑性）和正则化（变量选择）（51）。变量选择在4000次迭代以内完成。在我们的案例中，效应量即优势比通过岭正则化收缩至零。这样更便于解释结果，因为只需分析对结果最重要的变量，模型不会考虑无关变量。由于影响变量可以按组聚类（如情境定义中所述），我们采用稀疏组提升（53）作为基于模型的提升的扩展。所选方法使最终模型及变量的解释方式与广义线性模型（55）类似。另一种可行的替代方法是使用 Lasso 和稀疏组套索回归（56）。

模型评估

数据中 70% 的观测值被随机分配到训练数据集，其余 30% 被分配到测试数据集，用于最终评估。

模型评估依据测试数据上的受试者工作特征曲线下面积进行。对于二元结局变量，我们在每个概率阈值下评估了两项主要性能指标：第一，正确识别出财务状况良好的农场的比率；第二，正确识别出财务状况不佳的农场的比率，由此绘制受试者工作特征曲线（ROC）。ROC曲线下面积（AUC）通过考量预测模型计算出的所有可能概率阈值，同时纳入这两个比率。我们还计算了准确率作为补充指标，即分类模型在测试数据集中正确识别或预测的农民所占的百分比。尽管这一

指标不像AUC那样能在不平衡数据中平衡真阳性率和真阴性率，但由于其解读直观，因此仍被采用。

所有数据分析均使用统计编程环境 R 完成，图表则使用 R软件包 ggplot2（57）制作。

结果

用于数据评估的预测模型选择

在着手回答这三个研究问题之前，我们比较了不同的预测模型，以确定哪个模型的预测能力最佳，因此应当用于数据分析（表1）。除智利和突尼斯合并样本中的低财务健康外，随机森林（rf）在智利和突尼斯合并样本以及智利和突尼斯各自的样本中，往往优于所有其他模型。所选模型的ROC曲线概览见扩展数据图4。提升决策树（gbm）的表现与稀疏组提升（sqb）和基于模型的提升（mb）相近。在所有情况下，神经网络（nn）的表现都不如sgb和mb。仅包含气候变化经历及其财务影响的广义线性模型（glm），其预测性能低于sgb和mb。然而，使用提升法拟合glm（基于模型的提升-mb），并纳入更多与农场对气候变化的脆弱性及地理位置相关的变量后，与仅包含这些变量的glm相比，准确率和曲线下面积往往有所提高。纳入所有自变量之间的交互作用（mb-int）并未改善基于模型的提升的预测结果。结果表明，仅像glm那样考虑气候变化经历及其财务影响，不足以解释这两个财务福祉变量。因此，必须考虑其他变量。在与

准确率 福祉			**mb**	Mb int 广义线		**rf**	**nn**	**gbm**
智利	High	0.65	0.675		0.642	0.733	0.575	0.683
智利	Low	0.833	0.833		0.833	0.842	0.833	0.817
智利 & 突尼斯	High	0.71	0.693	0.685	0.618	0.734	0.556	0.705
智利 & 突尼斯	Low	0.809	0.809		0.822	0.817	0.822	0.793
突尼斯	High	0.595	0.579		0.545	0.645	0.529	0.57
突尼斯	Low	0.76	0.744		0.736	0.769	0.529	0.727
AUC	福祉 Sgb		**mb**	Mb int 广义线		**rf**	**nn**	**gbm**
智利	High	0.655	0.717		0.687	0.757	0.603	0.727
智利	Low	0.830	0.802		0.676	0.837	0.642	0.773
智利 & 突尼斯	High	0.733	0.723	0.731	0.627	0.796	0.619	0.758
智利 & 突尼斯	Low	0.763	0.733		0.637	0.746	0.721	0.735
突尼斯	High	0.663	0.661		0.537	0.710	0.616	0.658
突尼斯	Low	0.596	0.562		0.579	0.614	0.492	0.579

表1. 农场财务高福祉和低福祉的预测能力。所有拟合模型的准确率和曲线下面积（AUC）均使用来自智利和突尼斯的测试数据进行评估。相应的受试者工作特征曲线见扩展数据图4。缩写说明见正文。

可解释模型相比，像随机森林这类能够处理深层交互作用和复杂关系的模型，在某些情况下或许能略微提高对高福祉的预测准确率和曲线下面积；但对于低福祉的预测，较简单的模型似乎已经足够。由于本研究需要对数据进行解释，因此后续数据分析选用了sgb。

气候变化与农场财务福祉

首先，我们评估了经历气候变化是否会对农场财务福祉产生影响。降雨减少和气温升高与农场财务福祉下降有关（图 1）。总体而言，智利和突尼斯经历过降雨量减少的农民，其财务状况良好的可能性显著低于未经历降雨量减少的农民（0.635，p=0.020）。在较小程度上，这两个国家的气温升高也降低了农场财务状况良好的可能性（0.751，p=0.11）。在所研究的任何地区，干旱频率增加和极端天气经历均未对农场财务福祉产生显著影响。气温升高或降雨减少的影响在智利比在突尼斯更为明显。因此，在某些情况下，气候变化带来的负面经历会降低农场财务福祉，但这些负面经历的影响可能因国家而异。

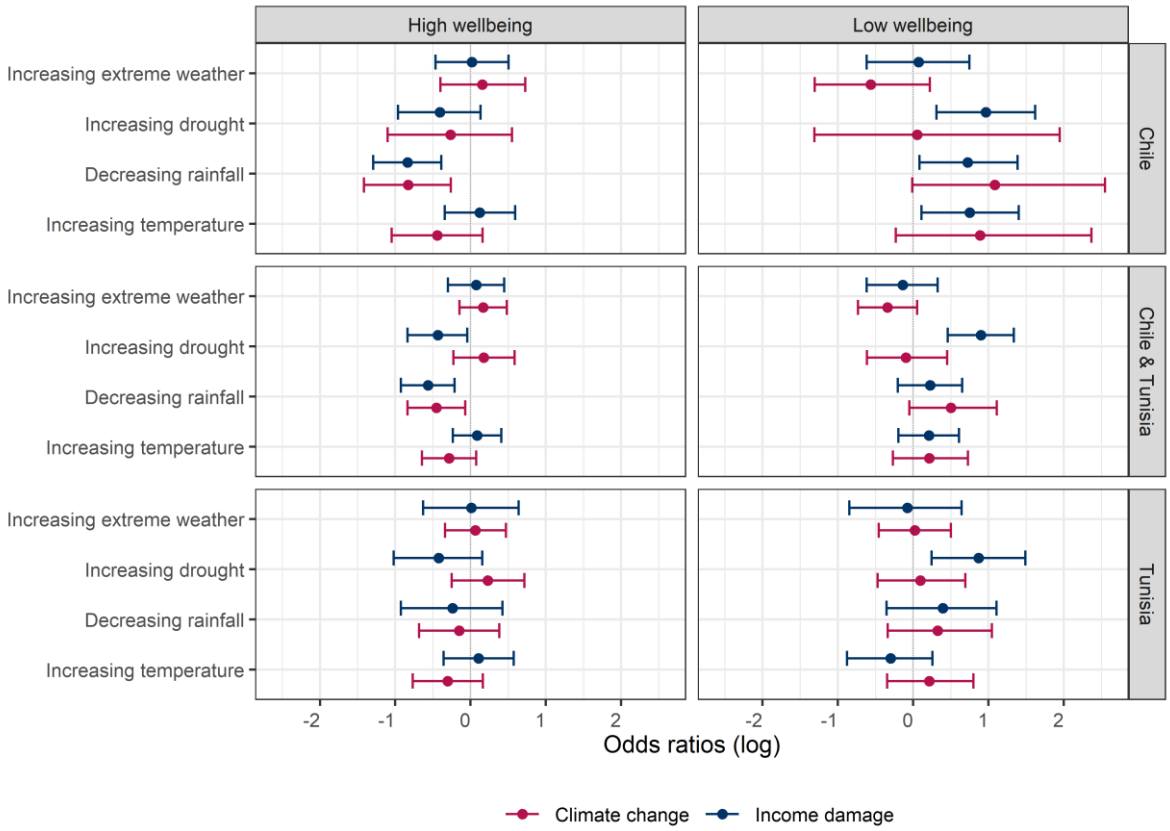


图1. 智利和突尼斯农场经历气候变化及作物财务损失对财务福祉的影响。基于逻辑回归的优势比（OR）置信区间（更多数据见扩展数据表5）。

其次，我们研究了特定气候变化影响造成的作物财务损失在多大程度上与整体农场财务福祉相关。结果表明，对于财务状况良好的农场，只有与降雨减少相关的收入影响与农场高幸福感显著相关（智利和突尼斯合并数据为 0.568，p=0.002；智利为 0.434，p<0.001）。对于财务状况不佳的农场，与气温升高相关的收入影响与农场低幸福感显著相关（智利为 2.119，p=0.021）；干旱更频繁也与之显著相关（智利和突尼斯合并数据为 2.457，p<0.001；智利为 2.623，p=0.003；突尼斯为 2.385，p=0.006）。降雨减少，尤其是在智利，似乎在一定程度上有助于解释农场的低幸福感。值得注意的是，经历干旱与低或高财务福祉并无显著关联，但干旱造成的财务影响往往与农场财务福祉显著相关。

农场财务福祉的变量重要性

稀疏组提升（sgb）算法使模型能够在预测建模中选择单个自变量或分组自变量（图2）。箭头方向表示一个组内所有变量对农场财务福祉产生的附加效应量（对数优势），

从而形成一个潜变量。对于高财务福祉，箭头朝上表示组变量值总体增加会提高高财务福祉的概率，而箭头朝下则表示高福祉的概率降低。同理，对于低财务福祉，箭头朝上表示组变量值增加会提高低福祉的概率。因此，社会资产增加或水平较高会提高农场高福祉的概率。请注意，对于非序数变量或变量组，没有添加箭头。

智利和突尼斯共有的农场高财务福祉最重要预测因素包括社会资产（对信息的依赖／使用、对信息来源的信任、社区、科学或宗教）和生物物理资产（农场规模、农场使用的水资源管理系统、所种作物的多样性），以及一个个体变量——农场拥有年限（图2）。自然资产（区域差异）几乎只对智利而言是重要的预测因素。先前农场所有权以及人力资产组（包括教育、年龄、性别、知识）则是仅对突尼斯而言的重要因素。智利和突尼斯共有的农场低财务福祉最重要预测因素包括区域差异、收入影响和财务资产组；例如，农场债务增加以及对果园收入的依赖，会提高农场低幸福感的可能性。若干因素仅会提高突尼斯农场低幸福感的可能性，包括社会资产组和生物物理资产组、农场所有权年限、干旱以及种植的品种。仅对智利而言，重要的个体因素是使用水井和农场管理年限。

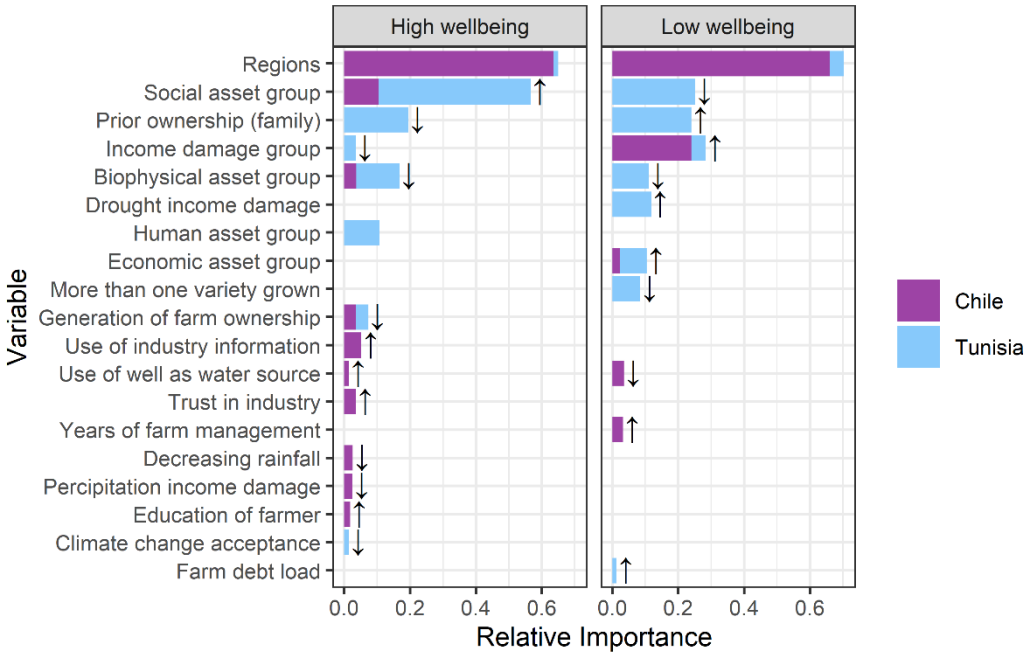


图2. 对农场财务福祉贡献最大的变量。分别针对智利和突尼斯，以及财务表现良好和较差的农场建立的稀疏组提升模型。

请注意，有些因素既是高财务福祉的预测因素，也是低财务福祉的预测因素，只是影响方向相反。例如，在智利，井水使用增加会提高高幸福感的可能性，同时降低低幸福感的可能性。在突尼斯，先前的家庭所有权会降低高幸福感的可能性，同时提高低幸福感的可能性。生物物理资产则是个例外：它们会降低高幸福感和低幸福感的可能性，这表明使用生物物理资产（如适应措施）只对帮助低财务福祉的农民有用。

变量交互作用的重要性

我们考察了自变量之间的交互作用是否会改变农场财务福祉模型的结果（图3）。尽管包含变量交互作用的模型，其预测能力不如仅包含加性效应的模型（表1），但每种交互作用的重要性仍揭示了数据集之间有趣且重要的相互依赖关系。其中一个结果是，考虑其他交互作用后，地区变量似乎不那么重要。社会资产和人力资产内部及彼此之间的交互作用似乎与农场的财务福祉有关，尤其是与信息使用和信任相关的交互作用。涉及适应措施、对气候变化的当前评估以及教育的交互作用也相对重要。这些交互作用揭示了变量之间的相互依赖关系，也表明某些变量可能存在混杂效应和中介效应。

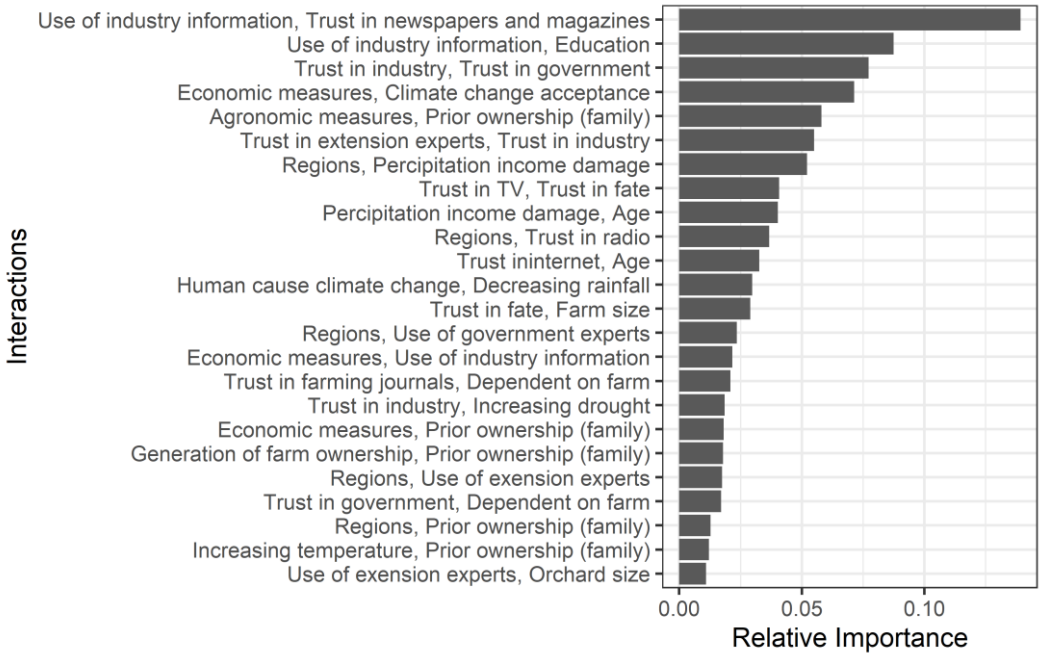


图 3. 对农场财务福祉影响最大的交互变量。智利和突尼斯合并数据的分量提升模型。

图4展示了一些可能影响农场财务福祉的重要交互作用。在智利和突尼斯，如果不使用报纸作为信息来源，那么气温升高或降水减少时，高幸福感的概率会下降（图4，左上）。然而，农民使用报纸时，智利和突尼斯的财务福祉不会因气温升高或降水减少而明显下降（图4，左下）。事实上，无论气温是否升高或降水是否减少，使用报纸都会提高农场财务福祉的概率：使用报纸消除了降水减少或气温升高对财务状况良好的任何负面影响。对行业的信任也观察到了类似的效应（图4，右上和右下）。农民，尤其是突尼斯的农民，如果信任行业这一信息来源，比不信任行业的农民更有可能获得良好的财务状况，无论他们是否经历了降水减少。然而，在突尼斯，对行业的信任似乎不会改变气温升高对高幸福感的影响；而在智利，与不信任行业相比，信任行业会加剧气温升高对农场财务福祉的负面影响。

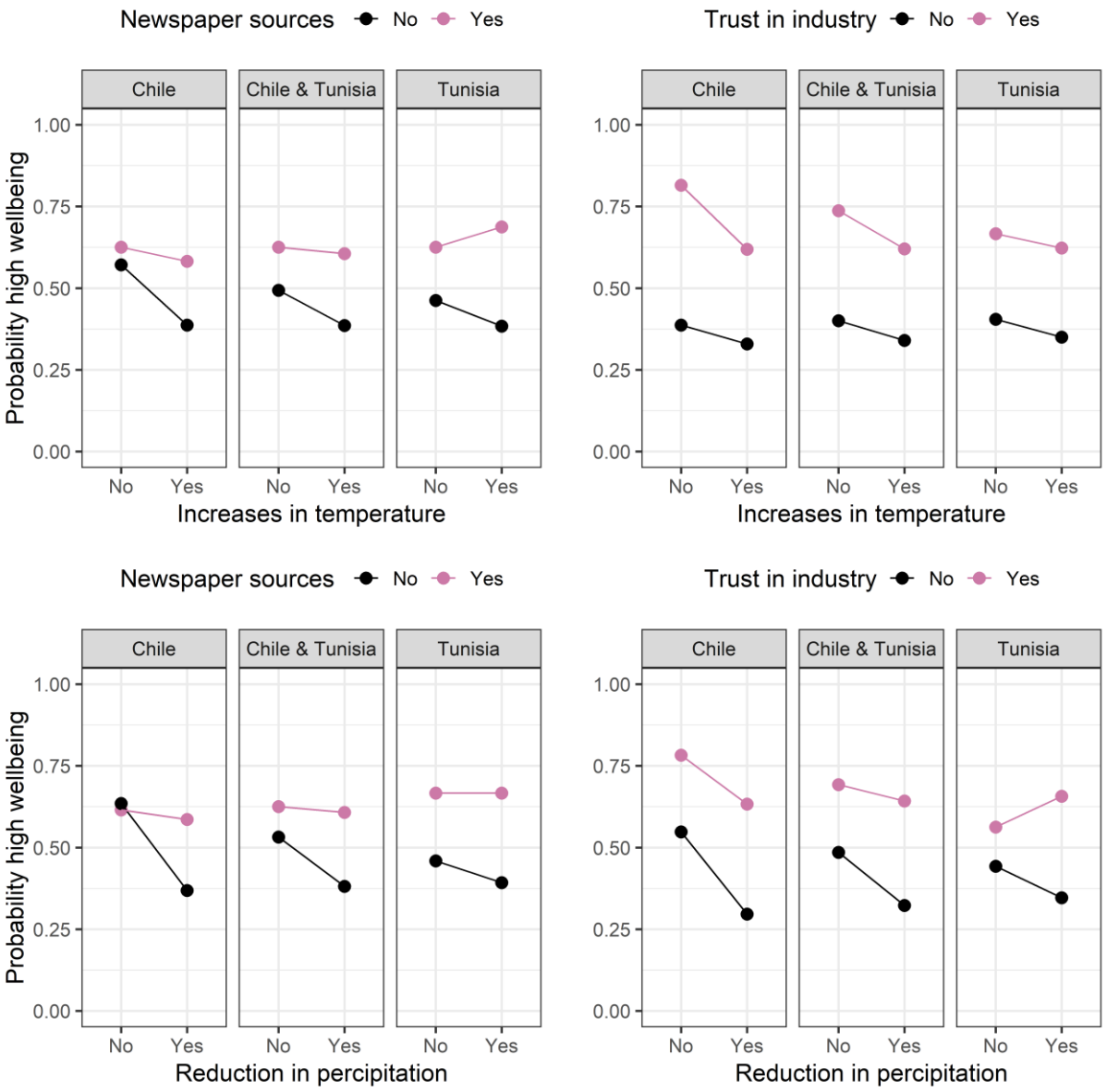


图4. 根据国家、气候变化因素、报纸使用情况和对行业的信任之间的交互作用，农场财务福祉较高的概率。

对媒体的信任、行业信息利用情况和农场财务福祉表明，无论来自哪个国家，不信任媒体且不使用行业信息的农民，财务状况良好的可能性最低（图5，左上）。信任媒体来源但仍不使用行业信息的农民，财务状况则明显更好。财务状况良好的可能性最高的农民，是那些信任媒体并使用行业信息的人；信任因素与信息利用之间发挥了协同作用。信任对财务福祉的重要性，可通过对行业、专家和政府的信任所产生的影响来说明。因此，对行业的信任与对专家的信任发挥了协同作用（图5，右上），对政府的信任与对行业的信任也是如此（图5，左下）。

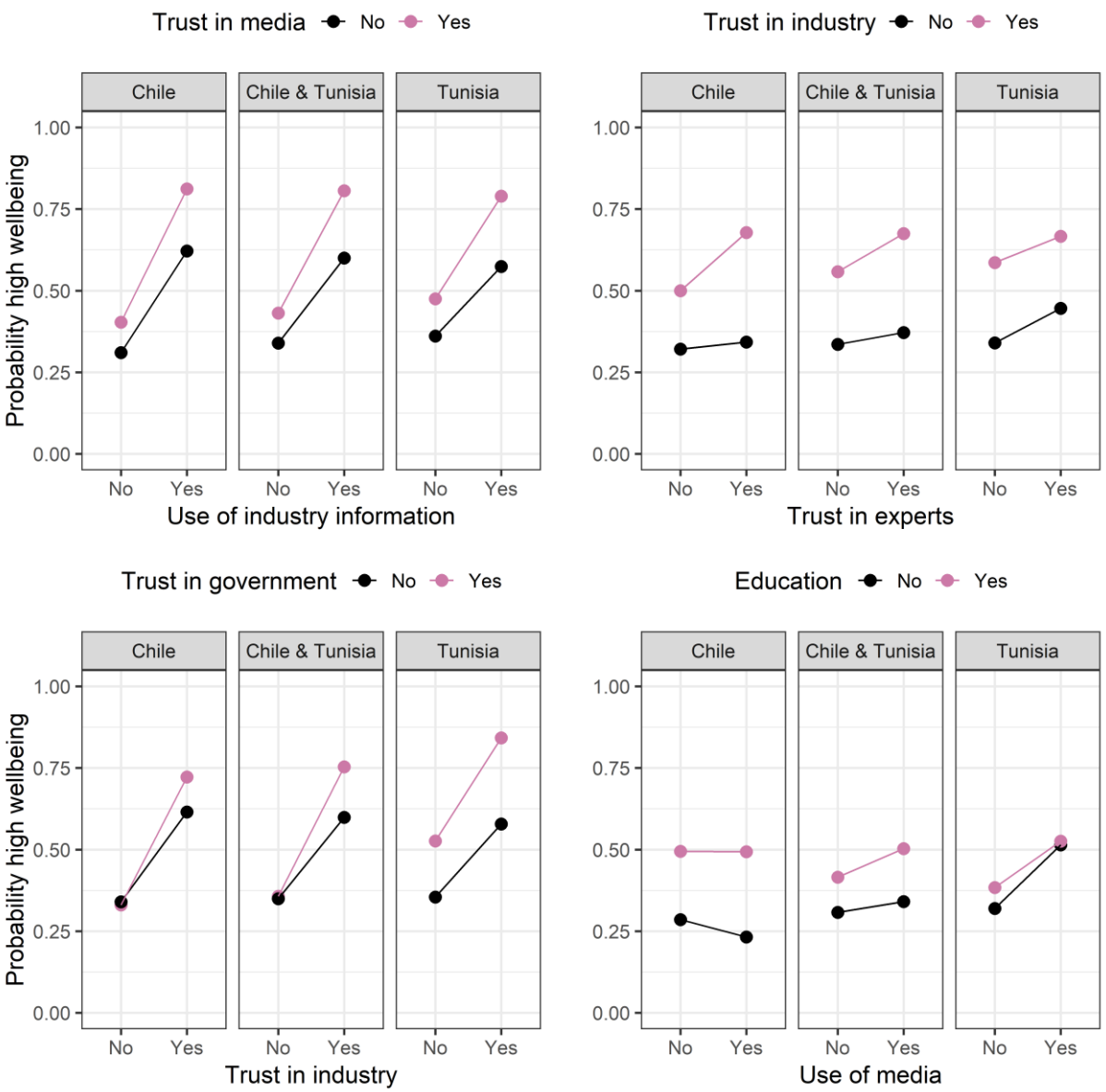


图5. 基于国家、信任、信息来源的使用情况和教育之间的交互作用，农场具有高财务福祉的概率。

在所有情况下，信任行业、专家或政府的农民，财务状况良好的可能性都高于不信任其信息来源的农民。其他交互作用，例如教育与媒体利用，在智利也对农场财务福祉产生积极的调节作用，但在突尼斯则没有：受教育程度较高且使用媒体的农民，财务状况良好的可能性往往高于受教育程度较低的农民（图5，右下）。

讨论

气候变化与农场财务福祉

此前发表的数据表明，气候变化会对作物产量和农场收入产生不利影响（3, 47, 58, 59, 60, 61, 62）。我们的研究结果显示，只有某些类型的气候变化会对水果农场财务状况产生负面影响。气温升高和降雨量减少很可能会损害水果农场财务状况，而极端气候事件似乎并未起到这样的作用。因此，尽管农民曾讨论过霜冻或冰雹事件可能造成的水果损害（63），但此类（目前仍相对罕见的）严重事件似乎不会影响农场中长期收入前景。事实上，水果种植者更担心干旱问题（以及随之而来的未来水资源可用性）（63），这与本文的研究结果一致：干旱频率增加会对农场收入和农场财务福祉产生负面影响。

然而，根据农民的自述，目前气候变化对水果农场财务状况的影响较小。事实上，与气候变化因素相比，其他变量更能预测水果农场财务状况。在智利，农场所在地是衡量农场财务状况是否良好的重要指标，智利中部的农场比智利南部的农场经营状况更好。在突尼斯，家族持有超过一代的农场财务状况较差。智利和突尼斯也有一些重要的共同预测因素。例如，在智利和突尼斯，信息获取以及对信息来源的信任，对于预测农场财务福祉而言都比气候变化更重要。这些共同因素有助于预测财务表现较高和较低的农场：信息获取越充分、对信息来源越信任，农场财务表现就越好，反之亦然。另一方面，我们还发现，在突尼斯和智利，某些因素对于财务状况较差或较好的农场更为重要。事实上，气候变化因素对财务状况不佳的农场影响更大。这表明，改善或维持农场财务表现时，需要考虑农场特定情况。因此，针对财务状况较差或较好的农场，可能需要采用不同的预测因素。最后，如上所述，不仅需要考虑单个因素，还需要考虑因素的协同效应，因为因素间相互作用会影响它们对农场财务福祉的单独或综合重要性。这与其他研究结果一致：财务状况较好的农场更有可能抵御气候影响（48,64），而因素间的相互依存关系可能促使人们采取应对气候变化的行动（65）。

尽管本文所呈现的结果表明，气候变化目前尚未对水果农场的财务福祉产生重大影响，但未来情况可能会发生变化。这

一点从本文分析的气候变化趋势中显而易见：随着气温升高、降水减少，果园财务绩效可能会下降。气温预测表明，未来突尼斯和智利的冬季夜间温度将持续升高。这将导致冬季冷量不足，并可能引发果树物候问题，因此很可能需要改变种植的果树类型（14,66）。同样，突尼斯降水量和水资源可用性的减少也将对未来产生重大影响。果树所需的大部分灌溉用水来自地下输水管道。如果这些管道中的水资源枯竭或变成盐化水，突尼斯农民可能会面临作物产量大幅下降（67）。在智利，安第斯山脉水流量减少、城市化加剧以及作物种植不当，可能造成水资源分配瓶颈（68,69）。这些后续预测与农民对未来气候的预测一致：大多数人担忧干旱和水资源可用性的影响（63）。要解决这些问题，就需要实施具体的适应措施，以降低水果农场未来面临的气候变化脆弱性；农民和政府都需要愿意为这些措施付费（64,70）。就此而言，影响果园财务绩效的某些因素具有区域特异性，因此应收集广泛的区域性数据集，而非依赖大数据分析的全国性数据集。

混合建模在气候变化对农业影响研究中的应用

预测建模为数据关系提供了新的见解，可用于提出和检验新的假设。我们发现，利用能够反映相关理论要求的有限数据集，统计模型可以对气候变化、中介变量与结果变量之间的关系作出恰当预测。不过，我们也证明，将统计模型与特定的机器学习方法（如提升法和交叉验证）相结合，能够大幅提高（广义线性）统计模型的预测能力。该混合模型仍可通过变量重要性和优势比进行解释，但对于通过数据驱动过程筛选出的变量，基于 F 统计量和 t 统计量的经典推断并不适用（71）。尽管如此，正如前文所述，混合模型的结果确实能带来新的见解，可用于构建假设，再通过经典统计方法加以检验。虽然随机森林（一种典型的黑箱模型）的表现优于稀疏组提升模型，但我们认为，这种改进通常不足以弥补可解释性的损失。采用类似的分析方法时，神经网络的表现略优于逻辑回归（72）。最终，我们选择预测性稀疏组提升模型和逐分量提升模型来分析当前数据。前一种模型表明，区域性或跨区域变量对于预测水果农场的财务表现是否良好十分重要。后一种模型则揭示，以农场财务福祉为结果变量时，各种变量与其之间的相互作用并非简单的一对一

关系。相反，某些变量似乎会调节那些可能直接影响结果变量的变量。

我们对混合模型的使用经验表明，尤其是在需要平衡预测性能提升（通常需要更大的数据集）与模型可解释性损失时，研究问题和可用的建模工具将决定应收集数据的范围和复杂程度、重点应放在区域性还是跨区域数据集上，以及可进行的分析类型和深度。与传统回归分析相比，机器学习使我们能够纳入范围更广的自变量，大幅提升对农场`s财务福祉的预测能力，并使数据呈现更加清晰。我们认为，通过对广义线性模型进行分组分量逐步提升，我们的混合方法能够在高维情境下生成有用的预测，同时保留变量重要性和优势比等基本可解释性。这样便能提出新的假设并构建模型，留待未来研究验证或否定。未来研究的关键挑战在于找到恰当的平衡：一方面采用基于理论的方法，在调查设计和由此形成的数据集中纳入数量有限、可能相关的变量；另一方面采用黑箱方法，依靠对尽可能多的数据点进行深度挖掘。

数据和代码可用性

用于开展分析的数据及相应代码可在补充材料中找到，也可在 GitHub 上获取（https://github.com/FabianObster/pasit_financial_wellbeing）

参考文献

1. 政府间气候变化专门委员会，2019。《决策者摘要》。载于：《气候变化与土地：政府间气候变化专门委员会关于气候变化、荒漠化、土地退化、可持续土地管理、粮食安全和陆地生态系统温室气体通量的特别报告 [P.R. Shukla、J. Skea、E. Calvo Buendia、V. Masson-Delmotte、H.- O. Pörtner、D. C. Roberts、P. Zhai、R. Slade、S. Connors、R. van Diemen、M. Ferrat、E. Haughey、S. Luz、S. Neogi、M. Pathak、J. Petzold、J. Portugal Pereira、P. Vyas、E. Huntley、K. Kissick、M. Belkacemi、J. Malley（编）]。剑桥大学出版社，剑桥，（2019）。

2. 政府间气候变化专门委员会，2022。《2022年气候变化：影响、适应与脆弱性》。政府间
D.C
S; L)22）。3. 赵等。气温升高使四项独立估算中的全球主要作物产量下降。《美国国家科学院院刊》114，9326-9331（2017）。4. Fanzo, J., McLaren, R., Davis, C. & Choufani, J. 如何在气候变化与变率背景下确保人人获得营养。GCAN政策简报1。国际食物政策研究所（IFP Pulawska,I., Lang, G.A.（编）。《樱桃：植物学、生产与用途》。英联邦农业局国际组织。（2017）。6. Manganaris, G. A., Minas, I., Cirilli, M., Torres, R., Bassi, D., & Costa, G. 面向未来的桃：重新审视一种特色作物。《园艺科学》305，111390（2022）。7. Predieri, S., Dris, R., Sekse, L., & Rapparini, F. 环境因素与果园管理对甜樱桃产量和品质的影响。《食品、农业与环境杂志》。
11-17页（2011）。9. Ghrab, M., BenMimoun, M., Masmoudi, M.M. & BenMechlia

《园艺科学》49，254-259（2014）。11. 张, J., Ferguson, L., & Whiting, M. D. 温度对

108762. （2019）。13. Penso G.A., Citadin, I., Scariotto, S., Santos, C., Junior,

14. Fernandez, E., Whitney, C., Cuneo,I.F. & Luedeling,E. 2020. 21世纪智利落叶果树

G. & DeJong, T.M. 春季气温对早期桃果实生长有重大影响。《园艺科学与生物技术杂志》

《园艺科学》20，1030-1042（1985）。18. Alae-Carew, C., Nicolaou, S., Bird, F.A.

对水果、坚果和种子的产量与营养品质的影响：系统综述。《环境研究快报》15. 023 020）。19. Botzen, W.J.W., Bouwer, L.M. & van den Bergh, J.C.J.M. 气候变化与冰雹灾害：实证证据及其对农业和保险的影响。《环境与资源经济学研究》32，341–362

境的

]。Ⅱ 会第一和第二工作组特别报告。剑桥大学出版社，剑桥，英国及纽约，美国，第109-230页。21. Lauren, E., Parker, A., McElrone, J., Ostoja,S.M. & Forrestel, E.J. 极端高温对多年生作物的影响及维持未来生产的策略。《植物科学》295，110397（2020）。22. Nelson, G. 等。全球情景中的农业与气候变化：为何模型结果不一致。《农业经济学》45，85-101（2014）。23. Lampe, M. 等。全球农业长期情景为何存在差异？AgMIP全球经济模型比较综述。《农业经济学》45，3-20（2014）。24. R. Mendelsohn, R., Nordhaus, W.D. & Shaw, D. 全球变暖对农业的影响：李嘉图分析。《美国经济评论》84，753-771（1994）。25. Mendelsohn, R.O. & Massetti, E. 运用横截面分析衡量气候影响对农业的作用：理论与证据。《环境经济学与政策评论》11，280-298（2017）。26. Carter, C., Cui, X., Ghanem, D., & Mérel, P. 识别气候变化对农业的经济影响。《资源经济学年度评论》10，361-380（2018）。27. Kononenko, I. 用于医学诊断的机器学习：历史、现状与展望。《医学人工智能》23，89-109（2001）。28. Bishop, C.M.《模式识别与机器学习》。施普林格（2006）。29. Bzdok, D., Altman, A. & Krzywinski, M. 统计学与机器学习。《自然方法》15，233-234（2018）。

30. McQueen, R.J., Garner, S.R., Nevill-Manning, C.G. & Witten, I.H. 将机器学习应用于农业数据。《农业计算机与电子学》12, 275–293 (1995). 31. González-Recio,O., Guilherme J.M., Rosa, G.J.M. & Gianola,D. 用于复杂性状全基因组预测的机器学习方法与预测能力指标。《畜牧科学》166, 217-231 (2014). 32. Coble, K. H., Mishra, A.K., Ferrell, S. & Griffin, T. 农业大数据：未来的挑战。《应用经济学观点与政策》40, 79-96 (2018). 33. Kamilaris, A. & Prenafeta-Boldú, F. X. 农业中的深度学习：一项调查。《农业计算机与电子学》147, 70-90 (2018). 34. Liakos, K. G., Busato, P., Moshou, D., Pearson. S. & Bochtis. D. 农业中的机器学习：综述。《传感器》18. 2674 (2018). 35. Zhu [illegible] 杂志》11, 32- 44 (2018). 36. Ansarifar, J., Wang, L. & Archontoulis, S.V. 用于作物产量预测 [illegible] Z. 作物以良的机器学习方法：利用表型和基因型大数据。《植物生理学杂志》257, 153[illegible] 易受 [illegible] 202[illegible] and, M. & Moradi, H. 运用机器学习模型、遥感和地理信息系统，研究气候变化和土地利用变化对洪水概率的影响。《水文学杂志》595, 125663 (2021). 40. Gu Y. e [illegible] 候以预测水稻产量。《生态指标》120, 106935 (2021). 41. Rolnick,D. et al. 运用机器学习应对气候变化。《ACM计算调查》55, 文章 42 (2022). 42. 政府间气候变化专门委员会。《2022年气候变化：影响、适应与脆弱性》。政府间气候变化专门委员会第六次评估报告第二工作组的贡献 [Pörtner, H.O. et al（编）]。剑桥大学出版社。即将出版 (2022). 43. Crane-Droesch. A. 用于作物产量预测和农 Var [illegible] . & Catal, C. 运用机器学习预测作物产量：系统性文献综述。《农业计算机与电子学》177, 105709 (2020).

45. Meshram, V. et al. 农业领域中的机器学习：现状调查。生命科学中的人工智能 1, 100010 (2021). 46. Mark, H.S. et al. 使农业适应气候变化。《美国国家科学院院刊》 104 10601-10606 (2007). 47. Khan, T., Sherazi, H.H.R., Ali, M., Letchmunan, S. & Butt 基于深度学习的生长预测系统：中国农业的一个用例。《农学》11, 1551 (2021). 48. Shahhosseini, M., Hu, G., Huber, I. & Archontoulis, S.V. 结合机器学习与作物建 Por 《背景中的世界：超越可持续发展的商业论证》。剑桥大学产业项目。(2003). www.cisl.cam.ac.uk/publications/the-world-in-context. 访问日期：12/09/14. 50. Ivory, S. & Brooks, S.B. 企业可持续发展的最新概念化：五种资源的可持续性。载于论文集。英国管理学会（BAM），英国管理学会2018年年会，布里斯托尔，英国 (2018). 51. Hothorn, T., Buehlmann, P., Kneib, T., Schmid, M. & Hofner, B. Mboost：基于模型的提升。R软件包版本 2.9-7 (2022). https://CRAN.R- project.org/package=mboost. 52. Tibshirani, R. 通过套索进行回归收缩与变量选择。《英国皇家统计学会杂志：B辑（方法论）》58, 267-288 (1996). 53. Obster, F. & Heumann, C. 稀疏组提升——无偏的组与变量选择。预印本见 https://arxiv.org/abs/2206.06344 (2022). 54. Nelder, J. A. R. & Wedderburn, W. M. 广义线性模型。《英国皇家统计学会杂志：A辑（综合）》135, 370–84 (1972). 55. Hofner, B. et al. R中的基于模型的提升：使用R软件包mboost的实用教程。《计算统计学》29, 3–35 (2014). 56. Simon, N., Friedman, J., Hastie, T. & Tib Wic 弱分析的优雅图形。施普林格纽约出版社。(2016). https://ggplot2.tidyverse.org. 58. Lobell,D.B. & Gourdji, S.M. 气候变化对全球作物生产力的 . & Aw-Hassan, A. 气候变化对中业农场收入保障的影响：综合建模方法。《农业、生态系统与环境》 188, 245-255 (2014). 60. Abraham, T.W. & Fonta, W.M. 气候变化与尼日利亚北部农民的适应融资。《金融创新》4, 11 (2018).

61. Dalhaus, T. 等. 极端天气对苹果品质的影响. Scientific Reports 10, 7919 (2020). 62. El Y[illegible]bi, A. 等. 气候变化引发的物候扰动对摩洛哥苹果园的潜在脆弱性：对产量和果实品[illegible]
Obs[illegible]
Obs[illegible]气候变化的因素. (2023). 65. van Valkengoed, A.M[illegible]萃分析. Nature Climate Change 9, 158–163 (2019). 66. Benmou[illegible] Luedeling, E., Ghrab, M., & Ben Mimoun, M. 冬季冷量大幅减少对突尼斯水果和坚果果园的影响 Climatic Change 162, 1249–1267 (2020). 67. Ver[illegible], J., McDonnell, R., Libert, C. & Konishi, Y. 突尼斯农业部门的气候变率、干旱与干旱管理. 世界银行集团，114页，
[illegible]
Bro[illegible]
(20[illegible]具. Scientific Reports 11, 17704 (2021).

致谢

作者谨向 PASIT 项目成员致谢，感谢他们在调查准备期间提供的反馈。

作者贡献

概念构思：P.M.P. 和 H.B. 调查：P.M.P. 和 H.B. 数据整理：P.M.P. 和 H.B. 正式分析：F.O. 和 P.M.P. 写作— F.O. 和 P.M.P.。所有作者均审阅并批准了稿件。

资金支持

本研究在项目“物候与社会对气温升高 –对突尼斯、智利和德国水果生产的气候影响”（PASIT；

利益冲突

作者声明不存在利益冲突。

扩展数据图

1. 分析中使用的所有自变量及其对应组别概览

分析（括号中的数字表示参照类别）

变量名称	类别*	n	组	原始问卷量表
农艺措施	是（否）	647	生物物理资产	二分的
经济措施	是（否）	464	生物物理资产	二分的
技术措施	是（否）	721	生物物理资产	二分的
使用水井作为水源	是（否）	231	生物物理资产	二分的
农场规模	>7公顷（≤7公顷）	283	生物物理资产	区间
果园规模	>2公顷（≤2公顷）	318	生物物理资产	区间
种植多个品种	是（否）	508	生物物理资产	二分的
其他产品	是（否）	571	生物物理资产	二分的
地区	智利中部	200	自然资产	名义的
地区	突尼斯中部	200	自然资产	名义的
地区	突尼斯北部	201	自然资产	名义的
地区	智利南部	200	自然资产	名义的
收入投资比例	≥80%（<80%）	137	经济资产	二分的
农场债务负担	负债沉重（其余）	96	经济资产	二分的
依赖农场的家庭成员	>2 (≤ 2)	528	经济资产	计数
家庭农场参与	>2 (≤2)	203	经济资产	计数
气候变化接受度	是（否）	676	人力资产	二分的
人为导致的气候变化	是（否）	685	人力资产	二分的
气候变化导致极端事件	是（否）	755	人力资产	二分的
农民年龄	>50 (≤50)	438	人力资产	计数
农民性别	男（女）	680	人力资产	二分的
农民受教育程度	≥ 小学（未上过小学）	632	人力资产	二分的
农场所有权世代	≥3 (0)	218	人力资产	计数
农场所有权世代	2 (0)	130	人力资产	计数
农场所有权世代	1 (0)	229	人力资产	计数
先前所有权	家庭（其他）	399	人力资产	二分的
农场管理年限	>10 (≤10)	437	人力资产	计数
报纸和杂志的使用	4,5 (1,2,3)	95	社会资产	李克特五点量表
使用农业期刊	4,5 (1,2,3)	161	社会资产	李克特五点量表
使用电视	4,5 (1,2,3)	415	社会资产	李克特五点量表
使用广播	4,5 (1,2,3)	219	社会资产	李克特五点量表
互联网使用	4,5 (1,2,3)	319	社会资产	李克特五点量表
农业推广专家的使用	4,5 (1,2,3)	346	社会资产	李克特五点量表
政府专家使用	4,5 (1,2,3)	166	社会资产	李克特五点量表
邻居利用情况	4,5 (1,2,3)	313	社会资产	李克特五点量表
行业信息利用情况	4,5 (1,2,3)	192	社会资产	李克特五点量表
农场协会利用情况	4,5（1,2,3）	97	社会资产	李克特五点量表
对报纸和杂志的信任	4,5 (1,2,3)	174	社会资产	李克特五点量表
对农业期刊的信任	4,5 (1,2,3)	291	社会资产	李克特五点量表
对电视的信任	4,5 (1,2,3)	329	社会资产	李克特五点量表
对广播的信任	4,5 (1,2,3)	241	社会资产	李克特五点量表
对互联网的信任	4,5 (1,2,3)	319	社会资产	李克特五点量表
对推广专家的信任	4,5 (1,2,3)	433	社会资产	李克特五点量表
对政府工作人员的信任	4,5 (1,2,3)	268	社会资产	李克特五点量表
对邻居的信任	4,5 (1,2,3)	319	社会资产	李克特五点量表
对行业的信任	4,5 (1,2,3)	215	社会资产	李克特五点量表
对农业协会的信任	4,5 (1,2,3)	184	社会资产	李克特五点量表
对政府机构的信任	4,5 (1,2,3)	213	社会资产	李克特五点量表
对其他农民的信任	4,5（1,2,3）	168	社会资产	李克特五点量表
对我的宗教的信任	4,5 (1,2,3)	236	社会资产	李克特五点量表
对命运的信任	4,5 (1,2,3)	268	社会资产	李克特五点量表
气温升高	是（否）	629	气候经历	二分的
降雨量减少	是（否）	659	气候经历	二分的
干旱加剧	是（否）	671	气候经历	二分的
极端天气增加	是（否）	542	气候经历	二分的
温度导致的收入损失	4,5（1,2,3）	294	收入损失	李克特五点量表
降水导致的收入损失	4,5（1,2,3）	219	收入损失	李克特五点量表
干旱导致的收入损失	4,5 (1,2,3)	161	收入损失	李克特五点量表
极端天气导致的收入损失	4,5 (1,2,3)	187	收入损失	李克特五点量表

*括号中的数字表示参照类别

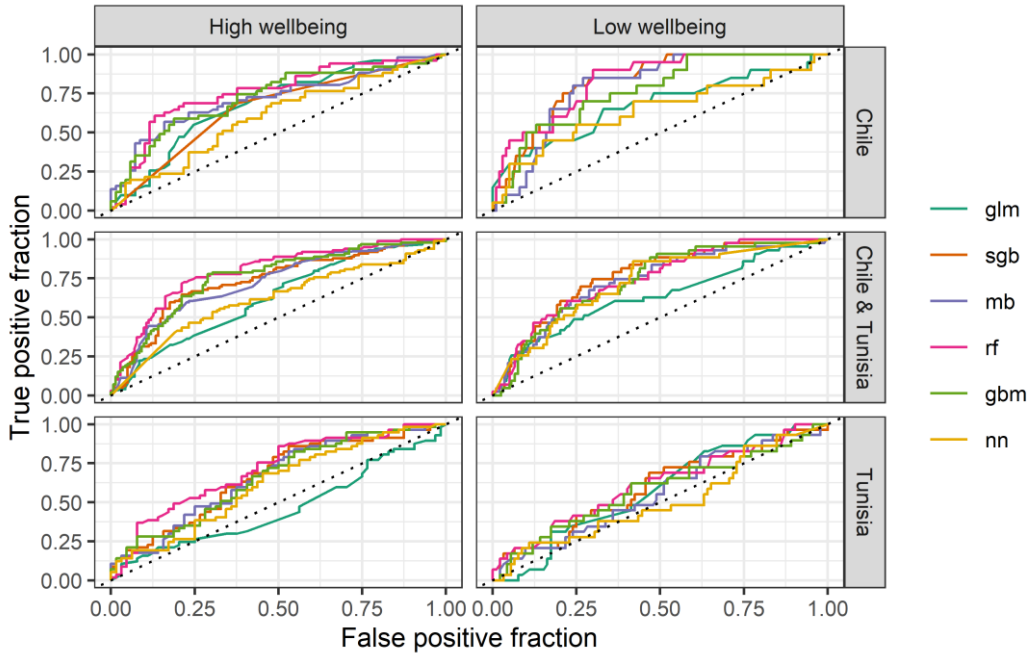
扩展数据图 **1 |** 分析中使用的所有自变量及其对应组别概览

变量名称	问题
农艺措施	请从以下列表中选择您在过去10年中已经采取的适应措施，以减少气候变化对 您的农业生产经营的影响。可选择多个答案。 a. 改变树木疏伐和修剪方式，以减少冰雹和雨害 b. 使用化学处理促进芽萌发并控制花期 c. 种植/嫁接需冷量低的新品种 d. 种植/嫁接耐旱品种 e. 种植/嫁接早熟或晚熟品种
技术措施	请从以下列表中选择您在过去10年中已采取的适应措施，以减少气候变化对 您农业生产经营的影响。可选择多个答案。 b. 使用化学处理促进芽萌发并控制花期 f. 安装防护网，以防冰雹、日灼和热害 g. 安装灌溉系统（例如树木滴灌） h. 提高灌溉系统的效率，以减少用水量和能源成本
经济措施	请从以下列表中选择您在过去 10 年中已采取的、旨在减轻气候变化对 您的农业生产经营造成影响的适应措施。可选择多个答案。 i. 停止耕种部分土地 j. 出售或出租部分或全部农场财产 k. 购买作物损失保险 l. 在农场外工作以补充农场收入（您和/或您的配偶）
使用水井作为水源	您的农场用于农业生产的水源有哪些（例如水井、河水）
农场规模	农场规模（公顷）
果园规模	您的桃树/樱桃树果园规模（公顷）
种植多个品种	种植的品种
其他产品	其他农产品
地区	智利中部
地区	智利南部
地区	突尼斯北部
地区	突尼斯中部
收入投资比例	近年来，您将年度农场总收入的多少百分比投资于适应措施，以减少气候变化对
已投资	您的农场运营造成的影响？从 0–20% 到 80–100% 的五点李克特量表
农场债务负担	您的农场业务负债程度如何？无负债、负债较轻、负债中等、负债较重
依赖农场的家庭成员	参与农场活动的家庭成员人数
依赖农场	
家庭农场参与	在您农场工作的家庭成员人数
气候变化接受度	全球气候没有变化。正确，不正确。
人为导致的气候	人类活动，例如燃烧化石燃料，是气候变化的一个重要原因
变化	正确，不正确
气候变化的成因	气候变化可能导致极端天气事件的强度增加、发生频率提高，例如冰雹、洪水、霜冻和强风。
极端天气事件	正确，错误
农民年龄	Age
农民性别	性别 男性，女性
农民受教育程度	教育程度：1. 未受教育，2. 小学未毕业，3. 小学毕业，4. 中学未毕业，5. 中学毕业，6. 技术教育（请注明专业领域），7. 大学（请注明专业领域），8. 大学研究生（请注明专业领域），9. 其他
农场经营代数	您的家庭拥有这座农场已有多少代?
所有权	
先前所有权	这座农场在您之前由谁拥有?
农场管理年限	您管理这座农场多少年了?
报纸和	您多久使用以下信息来源来了解如何应对气候变化对农业生产的影响?
杂志的使用	全国性报纸或杂志 五点李克特量表：从完全不使用到经常使用
使用农业期刊	农业期刊 五点李克特量表：从完全不使用到经常使用
使用电视	电视 五点李克特量表：从完全不使用到经常使用
使用广播	广播五点李克特量表，范围从完全没有到非常频繁
互联网使用	互联网五点李克特量表，范围从完全没有到非常频繁
农业推广专家的使用	农业推广人员五点李克特量表，范围从完全没有到非常频繁
政府专家使用	政府工作人员/专家，采用从完全不使用到非常频繁的五点李克特量表
邻居利用情况	邻居/社区，采用从完全不使用到非常频繁的五点李克特量表
行业信息利用情况	农业行业/出口行业，采用从完全不使用到非常频繁的五点李克特量表
农场协会利用情况	农民协会/合作社 五点李克特量表：从完全不使用到非常频繁
对报纸和	您在多大程度上信任以下信息来源能够提供可靠的信息，告诉您如何保护农场免受可能的气候变化
杂志的信任	影响？全国性报纸或杂志 五点李克特量表：从完全不信任到非常信任
对农业期刊的信任	农业期刊：五点李克特量表，从完全不频繁到非常频繁
对电视的信任	电视：五点李克特量表，从完全不频繁到非常频繁
对广播的信任	广播：五点李克特量表，从完全不频繁到非常频繁
对互联网的信任	互联网五点李克特量表：从完全没有到非常频繁
对推广专家的信任	农业推广人员五点李克特量表：从完全没有到非常频繁
对政府工作人员的信任	政府工作人员/专家五点李克特量表：从完全没有到非常频繁
对邻居的信任	邻居/社区：采用五点李克特量表，从完全没有到非常频繁
对行业的信任	农业行业/出口行业：采用五点李克特量表，从完全没有到非常频繁
对农业协会的信任	农民协会/合作社：采用五点李克特量表，从完全没有到非常频繁
对政府的信任	我信任政府机构能够帮助我保护农场，使其免受气候变化未来影响。
机构	从非常不同意到非常同意的五点李克特量表 我信任其他农民能为我提供建议，帮助我选择应采取哪些适应措施，以减少气候变化对我的农场造成的未来影响。
对其他农民的信任	从非常不同意到非常同意的五点李克特量表 相比科学，我更信任自己的宗教能指导我如何保护农场，使其免受气候变化未来影响的冲击
对宗教的信任	从非常不同意到非常同意的五点李克特量表
对命运的信任	我相信命运会指导我如何保护农场，使其免受气候变化未来影响的冲击。从非常不同意到非常同意的五点李克特量表
气温升高	近年来，我观察到农场的气温：1.升高了，2.没有变化，3.降低了，4.变得难以预测
降雨量减少	近年来，我观察到农场的降雨量：1.增加了，2.没有变化，3.减少了，4.变得难以预测 近年来，我观察到农场的干旱期和干旱
干旱加剧	1.增加了，2.没有变化，3.减少了，4.变得难以预测 近年来，我观察到我农场的极端天气事件
极端天气增加	1.增加了，2.没有变化，3.减少了，4.变得难以预测
气温收入	如果近年来您的农场因水果生长季气温变化而遭受桃/樱桃损害，您会认为
损失	这些作物损害对当年农场收入的影响有多严重？五点李克特量表：从完全不严重到非常严重
冬季气温收入	如果近年来您的农场因树木休眠期冬季气温变化而遭受桃/樱桃损害，您会认为
损害	这些作物损害对当年农场收入的影响有多严重？五点李克特量表：从完全不严重到非常严重 近年来，如果您的农场因水果生长季降雨量变化而遭受桃/樱桃损害，您会如何
降水导致的收入损失	评价作物损害对当年农场收入的影响有多严重？请使用五点李克特量表进行评分：从完全不严重到非常严重 近年来，如果您的农场因水果生长季干旱期和干旱的变化而遭受桃/樱桃损害，您认为这种损害有多
干旱导致的收入损失	严重？您会如何评价作物损害对当年农场收入的影响？请使用五点李克特量表进行评分：从完全不严重到非常严重
极端天气导致的收入	如果近年来，您在水果生长季因极端天气事件的变化而遭受农场桃/樱桃损害，您会如何评价
损失	该作物损害对您当年农场收入的影响有多严重？五点李克特量表：从完全不严重到非常严重

扩展数据图 **2 |** 本稿所采用的相关调查问题

模型	简称	软件包	主要超参数	可解释性
逻辑回归	glm	base	-	可解释
基于模型的提升不含交互作用	mb	mboost	Nu=0.3，mstop = 3000	可解释
稀疏组提升	sgb	mboost	Nu=0.3，alpha = 0.5，mstop = 3000	可解释
基于模型的提升含交互作用	mb int	mboost	Nu= 0.3，mstop = 3000	可解释
随机森林	Rf	randomForest	树=500，尝试次数=7	不可解释
梯度提升机	gbm	gbm	树=100，交互深度=3	不可解释
神经网络	nn	神经网络	隐藏层：1，逻辑激活函数	不可解释

扩展数据图 **3 |** 数据分析所用模型概览。



扩展数据图 **4 |** 概览　　智利所选模型的 ROC 曲线概览　　和突尼斯数据集

	智利 & 突尼斯	智利 & 突尼斯	智利	智利	突尼斯	突尼斯
	优势比（p值）	优势比（p值）	优势比（p值）	优势比（p值）	优势比（p值）	优势比（p值）
气候变化经历	高幸福感	低幸福感	高幸福感	低幸福感	高幸福感	低幸福感
气温升高	0.751 (0.11)	1.238 (0.400)	0.644 (0.150)	2.428 (0.167)	0.739 (0.201)	1.237 (0.466)
降雨减少	0.635 (0.020)	1.650 (0.089)	0.437 (0.004)	2.950 (0.086)	0.860 (0.577)	1.384 (0.353)
干旱频率增加	1.29 (0.394)	0.906 (0.715)	0.766 (0.523)	1.052 (0.949)	1.257 (0.353)	1.099 (0.749)
极端天气频率	1.18 (0.302)	0.709 (0.086)	1.170 (0.584)	0.569 (0.146)	1.068 (0.751)	1.022 (0.930)
收入影响						
气温升高	1.092 (0.592)	1.234 (0.305)	1.131 (0.606)	2.119 (0.021)	1.113 (0.652)	0.741 (0.299)
降雨减少	0.568 (0.002)	1.254 (0.297)	0.434 (<0.001)	2.064 (0.028)	0.789 (0.489)	1.481 (0.289)
干旱频率增加	0.647 (0.031)	2.457 (<0.001)	0.664 (0.143)	2.623 (0.003)	0.656 (0.158)	2.385 (0.006)
极端天气日益频繁频率	1.08 (0.699)	0.871 (0.565)	1.019 (0.940)	1.074 (0.837)	1.009 (0.976)	0.926 (0.838)

扩展数据图 **5 |** 智利和突尼斯农场的财务福祉。基于逻辑回归得出的优势比及相应的 p值