Cropland Capture – A Game for Improving Global Cropland Maps

Tobias Sturn
International Institute for Applied Systems Analysis
Earth Observation Systems
Laxenburg, Austria
sturn@iiasa.ac.at

Michael Wimmer
Institute of Computer Graphics and Algorithms
Vienna University of Technology, Austria
wimmer@cg.tuwien.ac.at

Carl Salk
International Institute for Applied Systems Analysis
Earth Observation Systems
Laxenburg, Austria
salk@iiasa.ac.at

Christoph Perger
International Institute for Applied Systems Analysis
Earth Observation Systems
Laxenburg, Austria
perger@iiasa.ac.at

Linda See
International Institute for Applied Systems Analysis
Earth Observation Systems
Laxenburg, Austria
see@iiasa.ac.at

Steffen Fritz
International Institute for Applied Systems Analysis
Earth Observation Systems
Laxenburg, Austria
fritz@iiasa.ac.at

ABSTRACT

Current satellite-derived global land-cover products, which are crucial for many modelling and monitoring applications, show large disagreements when compared with each another. To help improve global land cover (in particular the cropland class), we developed a game called Cropland Capture. This is a simple cross-platform game for collecting image classifications that will be used to develop and validate global cropland maps in the future. In this paper, we describe the game design of Cropland Capture in detail, including aspects such as simplicity, efficiency in data collection and what mechanisms were implemented to ensure data quality. We also discuss the impact of incentives on attracting and sustaining players in the game.

Categories and Subject Descriptors

D.2.2 [**Software Engineering**]: Design Tools and Techniques; K.8.0 [**Personal computing**]: General-*Games*

Keywords

Cropland Capture, Serious Games, Crowd-Sourcing, Geo-Wiki, Landspotting, Global Cropland Maps

1. INTRODUCTION

Over the last decade, a number of global satellite-derived land-cover products have been developed, e.g., MODIS [17], GlobCover [20] and GLC-2000 [21]. These products are important for monitoring, assessment and modeling purposes, yet when compared with each another, they show huge spatial disagreements [13]. For example, MODIS, GlobCover and GLC-2000 have a forest and cropland disagreement of 893 million hectares, and estimates of cropland differ by up to 20%, which makes these maps highly uncertain [26]. Knowledge of cropland extent is crucial for applications in the field of food security, e.g., to assess yield and production gaps or to estimate potential yield losses that could occur as a result of wide-spread drought or other anomalies that negatively affect crop production [14].

To improve these global land-cover products, the Geo-Wiki [25] tool was developed. Geo-Wiki is a visualization, crowdsourcing and validation platform in which volunteers classify high-resolution satellite imagery. Through a number of different Geo-Wiki crowdsourcing campaigns, more than 250,000 land-cover pixels (or areas) of varying sizes have been collected, which were used for both the development of hybrid land-cover products [15] and for validating existing maps [27]. However, it has been quite challenging to gather data through these campaigns, where the incentives used were small prizes and co-authorship on scientific publications.

Therefore, as a way of increasing crowd participation in Geo-Wiki, we experimented with a number of different game prototypes in the LandSpotting project [29]. These included a strategy game running on Facebook that was based on the game Civilization, a tower defence game [28], a tagging game, and a game in which users arranged tiles into pictures, while simultaneously classifying land cover. From these games we learned a great deal about what works and the need to simplify the game mechanics as much as possible.

In this paper, we present our latest and so far most successful game, called Cropland Capture (Figure 1), a simple cross-platform application for collecting information about the presence or absence of cropland on the Earth's surface. The data collected will be used to improve global cropland maps in the future.

First, the design goals and game mechanics of Cropland

Figure 1: Cropland Capture running on multiple platforms.

Capture are described, in particular how we attempted to make the game as efficient as possible, how correctness was ensured through the mechanism of output agreement and which incentives for playing the game were provided. This is followed by a presentation of the results of the game. Here we show that the game is very efficient in collecting image classifications. The correctness of the data that were collected are then discussed, in particular, how much the mechanism of output agreement was able to preserve correctness. We then discuss the effect of the incentives on persuading players to play the game and how the game was received by the players. Finally, conclusions are provided and some information about how the lessons learned from Cropland Capture are being implemented in the next game, which will be launched in the spring of 2015.

2. RELATED WORK

The idea of using games for a scientific purpose (GWAP, games with a purpose) is not new, and some games have already been able to very impressively help science. One of the most famous examples is *FoldIt* [23, 24], a puzzle-like game in which the players fold proteins. Within ten days, the FoldIt players have, for example, accurately determined the crystal structure of M-PMV, an AIDS-like virus infecting apes, which had been an unsolved problem for scientists despite 15 years of previous research effort. Just recently, with the help of the 230,000 FoldIt players, a new algorithm for protein folding was developed that outperforms previously published methods [4]. The game *EyeWire* [8], which is a game about mapping the brain, has already been played by more than 160,000 people.

The process of obtaining help from a large group of people is often called *crowdsourcing* [7] or *citizen science* [22], which is the broader involvement of citizens in a range of scientific activities from data collection to data analysis and research design. A successful example of a citizen-science project is *Galaxy Zoo* [2], in which the users help to discover and classify new galaxies. Galaxy Zoo 1 was able to collect over 50 million classifications done by more than 150,000 users. As can be seen from the games and citizen-science projects described above, there are already many projects yielding impressive scientific results.

Several types of game mechanics have already been developed to gather correct meta-information through games. One group of GWAPs are the so-called *output agreement* games, one example being the *ESP* game [11]. In this game, two random players have to label a given image (the input). The more they agree on the labels (the output), the higher the scores they achieve – hence output agreement. As of July 2008, 200,000 players had contributed more than 50 million labels in this game.

Another group of GWAPs are described as *input agreement* games, for example, *TagATune* [3]. In TagATune, a sound clip (the input) is presented to two random players, who in return give a series of labels to the other player as output. The two players win the game if they both correctly agree on whether they have heard the same input sound or not.

A third group of GWAPs are called *inversion problem* games, for example, the game *Verbosity* [12]. One player, the describer, is given a word as input and has to describe this word to the guesser. The two players only get points if the output of the guesser represents the input of the describer. For example, the describer describes the word "milk" with "white, something to drink, people usually eat cereal with it". Thus, facts are collected as a side effect of playing the game.

In the image-annotation game *Kisskissban* [6], a third player, the blocker, is added, who competes with two other collaborative players, the couple. While the couple tries to find consensual descriptions of an image, the blocker's mission is to prevent the couple from reaching consensus. The blocker will try to detect and prevent coalition between the couple. Therefore, these efforts naturally form a player-level cheat-proofing mechanism. To evade the restrictions set by the blocker, the couple would endeavour to produce a more diverse set of image annotations.

Although the methods described above are good ways for getting users to deliver input that they agree upon, they do not always guarantee correct data. As we will show, players of Cropland Capture were sometimes wrong even if the majority of people agreed on the answer. One reason for this was that the players were not trained sufficiently in understanding exactly what is meant by cropland. Another reason was that the game mechanics allowed a strong influence of some players who played the game frequently but who were sometimes wrong. This is possible because frequent players determine the initial classification for many tiles and are thereby able to influence other players. This means that the output-agreement mechanic alone is not sufficient and other mechanics must be implemented to guarantee correctness.

In the LandSpotting project [29, 28], where we already created a number of serious-game prototypes for collecting data on land cover, we used a disjoint game-design approach described by Markus Krause [18, 19], where the task to be solved should be part of the mechanics, but not the dominant element. For example, we created a strategy game where the task to classify land cover was only a part of the mechanics of a classic strategy game. Our motivation at the time was to take already successful games, especially social games like FarmVille, as a reference and to try to copy these already proven game mechanics to create games with a scientific purpose that would appeal to a large audience. In Cropland Capture, we did not use a disjoint game design. We describe the reasons for this and present the game design and mechanics of Cropland Capture in the next section.

3. CROPLAND CAPTURE

Cropland Capture is a cross-platform serious game played in a browser or on mobile devices like the iPhone, iPad and/or Android devices for gathering data on the presence or absence of cropland, with the eventual goal of improving global cropland maps. It can be downloaded from the AppStore[1] or GooglePlay Store[2] or can be played in a browser[3]. Cropland Capture ran for 25 weeks from November 15th 2013 to May 9th 2014, after which a series of prizes were awarded, but it continues to be available for playing.

Cropland Capture is described in detail in this section. First, the game-design goals are presented. This is followed by a description of the game mechanics, the efficiency of the game and how we try to guarantee correctness in the answers provided by the players. Finally, we examine the incentives for playing the game before presenting the results.

3.1 Design Goals

In our previous LandSpotting project [29], we found that while a disjoint game-design approach can be used to collect information on land cover, we were not very successful in attracting a large number of people with these games. One reason was that we were only a small team and there was only a short development time (below one year). The games developed in LandSpotting could therefore not compete with games with similar game mechanics such as the latest Civilization game or Plants vs. Zombies. People are very selective when it comes to games and entertainment, and they tend to choose games that offer more fun and entertainment. Also, some players might have felt that there was a "hidden agenda" in the LandSpotting games, i.e., that they were being used as "cheap workers".

For this reason, we did not use a disjoint game design in Cropland Capture. Instead, we defined the task to classify land cover as the central mechanic of the game. This approach felt much more straightforward and more honest to us.

The fundamental design goals we wanted to achieve in Cropland Capture are summarized below:

- Simplicity: The game should be very simple and easy to understand, so that it can be played by everyone, everywhere.
- Efficiency: The game should be very efficient in collecting data in the sense that as much data as possible should be collected in a certain time frame.
- Correctness: The game should ensure that players classify the images correctly.
- Incentives: The game should provide enough incentive to attract and sustain the players.

In the following, we describe how these fundamental design goals were implemented in Cropland Capture.

3.2 Simplicity

We used very basic game mechanics to achieve the goal of simplicity: Players are presented with an image (either satellite images or ground-based pictures) and asked whether they see any evidence of cropland in the image. This is illustrated in Figure 2, where a screenshot of the game is shown. The center shows the land-cover image, and the top shows the question "Is there any cropland in the red box?". The players can now answer with *yes*, *no*, or if they are unsure, with *maybe*. After the player has selected an answer, the next image appears seamlessly. Thus, the base mechanic of the game is very simple and easy to understand.

Figure 2: Classifying an image in Cropland Capture.

3.3 Efficiency

Efficiency is often defined as the ratio r=P/C, where P is the output produced per C, which is the amount of resources consumed. Here, C corresponds to the time a player spends with the game. In order to optimize this ratio, we tried to collect classifications as quickly as possible using a classification method that is highly intuitive. On mobile devices, the player can swipe the image with a finger to the correct category *yes*, *no* or *maybe*, where these categories are located

[1]https://itunes.apple.com/au/app/cropland-capture/id694689972?mt=8
[2]https://play.google.com/store/apps/details?id=iiasa.cropcapture
[3]http://www.geo-wiki.org/games/croplandcapture/

on different sides of the screen. The image is swiped to the left side if there is no cropland in the red box, to the right side if there is cropland, and to the bottom if the player is not sure. In the browser version, the player can click with the mouse on the categories or use the cursor keys left, right and down to play much faster. After a very short amount of time, the player will learn that left means *no*, right means *yes* and down means *maybe* without the need to look at the categories. Thus, the images can be classified without having to look away from them, which saves a lot of time.

This is also the reason that we ask the same question "Is there any cropland in the image?" throughout the whole game. Otherwise, the player would have to constantly read the question again, which would also cost more time.

Since the classification of the images can be performed very quickly, we had to optimize the loading times of the images in order to maintain a constant stream of images. We load many images simultaneously in many background threads. Depending on how much memory is available on the platform, we load up to 10 images. This fast loading is also one reason (in addition to the great advantage that we know exactly what the players have classified) why we use images rather than a platform like Google Earth to display the images in the game; the satellite images might change over time, and we would not be sure of exactly what images the players classified.

An additional feature introduced was the removal of images once enough classifications had been received for an image. This avoids collecting redundant classifications that do not provide new information about the presence or absence of cropland. In this way, we also use the time of the players more efficiently.

3.4 Correctness

To ensure that players make correct classifications, we used *output agreement*, which means that the players had to agree with other players on the classifications they made in order to score points. In the game, the player scores one point for each correct answer and loses one point for each wrong answer. Answers of *maybe* result in neither gain nor loss of points.

In order to determine if an answer is correct, we started with a small pool of images that were validated by experts. To automatically expand the pool of images, 10% of the images that the players received had not been classified at all. For those images, we assumed that the answer from the first player was correct. For the images already classified by other players, we calculated the average, and if more than 85% of players agreed on a classification, we assumed the answer to be that of the majority. If no agreement was found, we treated this image as unclassified and the player was given a point regardless of the answer provided. It has to be noted that these thresholds of 10% and 85% were intuitively chosen by us and more mathematical analysis has to be used in follow up games to optimally adapt these numbers. It would for example also be possible to use a dynamic system in which the thresholds dynamically adjust based on the performance of the players.

Figure 3: Background graphic of Cropland Capture which says "Help Science" to serve as an intrinsic motivator.

We implemented some additional quality-assurance features to ensure that players who answer randomly cannot influence the results. For example, only players who correctly classified the preclassified reference images received unclassified images. Moreover, only the answers from these players were then uploaded to the database and used in the output-agreement calculation. Another quality-assurance feature is that the ratio of cropland to non-cropland images that each player received was roughly 50%. This ensured that players who always chose the same answer consistently or answered randomly would not progress and their score would stay the same.

To further improve data quality, we added the feature that the players could challenge an answer if they thought they were correct but were penalized for an incorrect answer. These images were then sent to an expert for analysis. If the players were actually deemed to be correct, they gained 5 points, but if they were wrong, they lost 3 additional points.

3.5 Incentives

In order to motivate people to play Cropland Capture, we provided intrinsic and extrinsic incentives for the players. As an intrinsic motivator, we told players that by playing the game, they can help science, since the data we collect will be used to improve global cropland maps. These maps are crucial for food-security applications and the monitoring of famine during times of severe drought events. In Figure 3, we employed a background graphic in the game which says "Help science".

We also created a trailer video[4] of Cropland Capture, which explains in a funny way why people should play the game to help science.

Throughout the game, we highlighted how much of the Earth the players have already helped us to classify. This incentive was intended to make players feel good, to show that they have an impact on scientific research and that their contribution is very important.

As extrinsic motivators, we added a leaderboard and offered two forms of prizes. Each week during the last five weeks

[4]https://www.youtube.com/watch?v=T0Lmt7yXw2k

of the Cropland Capture competition, one answer was randomly chosen, and the player who submitted the answer was awarded a small prize, such as a fitness monitor or a compass. Thus, each additional classification increased a player's chance of winning the weekly prize. We used this sort of prize so that players with fewer contributions had a chance to win the prize and that these players would not stop playing if they noticed that they did not have a chance to win the competition anymore.

The second set of prizes was awarded at the end of the Cropland Capture competition in a draw. The top 3 players at the end of each competition week automatically qualified for the draw. Scores were reset at the beginning of each week so players would have a new chance to be entered into this final draw. After the Cropland Capture competition was finished, we randomly picked three winners from these top weekly winners, who then became our overall winners and were awarded bigger prizes such as smartphones and tablets. We used this lottery system to keep players engaged throughout the competition even if they were not in the top three in one week. We picked the top three winners each week instead of only one winner so that more people would compete to get into the top three each week.

We used Twitter[5] as the communication channel to talk to the players. Each week, we tweeted the top 3 winners who were entered into the draw and the lucky weekly winner who received a small prize during the last five weeks of the competition. When players had collectively reached a magic number of how much land they had already classified, we immediately tweeted these facts in order to acknowledge their contributions and encourage them further. We also posted when articles about Cropland Capture were published and when updates were made to the game.

In order to keep the players engaged, we wrote monthly newsletters to the players, which provided the latest information about the game and how much land area the players had already helped to classify. We compared the area that had already been classified with countries, e.g., "You have already classified an area bigger than Australia", so that the players had a clearer image of their contribution.

We constantly updated the game throughout the competition to show the players that they are very important to us and that we are listening to their proposals for the game. For example, we added a chat channel to the game so that the players could talk to each other and form a community. We also added geography quiz questions to the game, which popped up randomly to add a surprise element to the game. If the players were able to correctly answer the question, they received 10 bonus points. Examples of such questions were: "Which of these countries is the biggest?" or "What is the capital of Botswana?". Thus, the players also learned new information while playing the game.

4. RESULTS

In this section, we give general statistics and examine the efficiency of the game and the correctness of the obtained data. We also discuss how the game was received by the players.

4.1 General

In total, Cropland Capture was played by 3,314 players, who together provided 4,648,659 classifications of 187,673 unique images. Of these images, 98,411 were satellite images with areas ranging from 0.06 km^2 to 1 km^2, and 89,232 were landscape pictures taken on the ground. This huge number of classifications represented a big success when compared to previous Geo-Wiki campaigns, which collected 30K to 80K classifications per campaign.

We also collected information on the type of device used to play the game. The players made 588,177 ratings using Apple's iPhone 5, 619,457 using other iPhone types, 709,396 using iPads, 1,682,291 using the browser version, and 1,049,338 using Android devices. This was quite surprising, as we thought that the browser version would be by far the most successful platform. Instead, the iOS version collected 1,917,030 classifications in total, which is more than the browser version. This is surprising as we did not get featured on the Apple App Store, which might have explained this result. It is also surprising that the iOS version collected nearly twice as many ratings as the Android version, although the Android phone market is much bigger, with nearly 80% market penetration.

4.2 Efficiency of the Game

We have already defined efficiency as P/C where P is the number of images classified and C is the amount of time a player spends in the game. In order to calculate how efficient Cropland Capture is, we stored the amount of time needed by the players to do each classification differentiated by platform. Thereby the timer was always started as soon as an image was shown to the player and stopped as soon as the image has been classified. Because the next image is immediately shown after an image has been classified, there is no more additional time spent in the game. Therefore, in order to calculate the efficiency, the value C is just a sum of all the times measured.

For example, for 0.8% of the images, the players needed more than 10 seconds to do a classification, which is an efficiency bellow 0.1 images per second, for 0.2% of the images, the players needed more than 30 seconds, which represents an efficiency bellow 1/30, and for 0.12% of the images, the players had an efficiency bellow 1/60, i.e., they needed more than a minute to do a classification. The slowest classification was done in 3.2 days. We assume that for these very slow classifications, the players have taken a pause from the game because it is unreasonable to assume that they needed so much time to determine whether or not there was cropland in the image. Indeed, the majority of classifications was done much faster. 97% of the classifications were done in less than 4 seconds (0.25 images per second), and 92% were done in less than 2.5 seconds (0.4 images per second). We will, therefore, only use classifications below this threshold of 2.5 seconds in the analyses that follow.

Figure 4 shows the time that the players needed to classify an image on the different platforms with animations enabled. No images were classified under 500 ms, and the majority of classifications were done between 900 and 1100 ms (ef-

[5]http://www.twitter.com/cropcapture/

ficiency 0.9 to 1.1). The Android version is slightly faster than the other versions, but the players classified the images (with animations enabled) at nearly the same speed on all platforms.

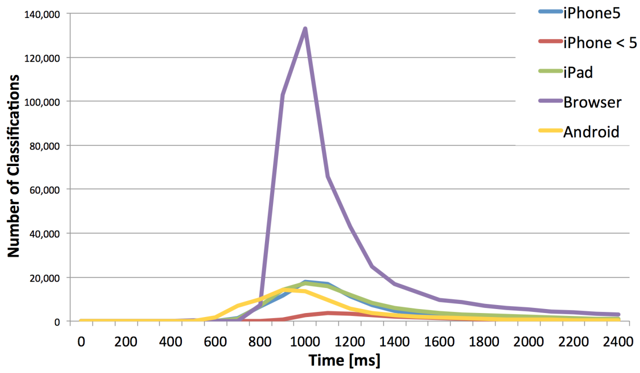

Figure 4: The time needed for making a classification on different platforms with animations enabled.

With animations enabled, the image moves smoothly to the side after the player has selected a category. This animation looks nice but takes 500 ms to complete. For this reason, we updated Cropland Capture in the middle of the competition and turned the animations off to see if the efficiency improved. In Figure 5, the new times needed for making a classification with animations disabled are shown. As can be seen from the figure, the efficiency has drastically improved to around 1.6 to 2.5 – for the browser version, in which the players now only need around 400 to 600 ms or half of the time needed with animations enabled. The other platforms have also improved. For the Android and iOS versions, the efficiency is now around 1.1 to 1.4, i.e., the players need around 700 to 900 ms to make a classification. Some players still needed 900 to 1100 ms, most likely because they did not update their application. The mobile versions, in which the players classified by dragging the images to the sides of the screen, improved by 300 ms after the update. The reason why the mobile versions only improved by 300 ms instead of the 500 ms of the browser version is because the dragging mechanic itself requires around 200 ms to complete and is slower than just clicking a cursor key. This shows that although animations can provide a nice visual appeal, they can also consume time that could be better spent. In our case, the animations made up nearly 50% of the time in the browser version and around 30% in the mobile versions.

4.3 Correctness

As has already been discussed in Section 3.4, one method for improving the correctness of the collected data was to give the players an equal number of cropland and non-cropland pictures so that the players would not just select the same answer all the time. In total, there were 2,246,683 classifications with the answer yes, and 2,314,725 with the answer no, which shows that we have been successful in providing a nearly equal set of cropland/non-cropland pictures. Clearly we did not provide enough incentives for the players to select the maybe answer, as only 87,251 (1,87%) maybe answers were given.

In order to create more certainty, on average 25 classifications from different users were collected for each image. Figure 6 shows the agreement of the players by number of classifications. As can be seen from the figure, the majority of images have a > 90% agreement, indicating that it was easy for the players to identify the presence or absence of cropland.

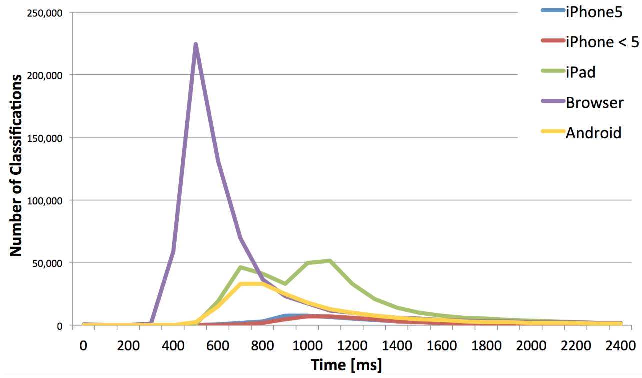

Figure 5: The time needed for making a classification on different platforms without animations.

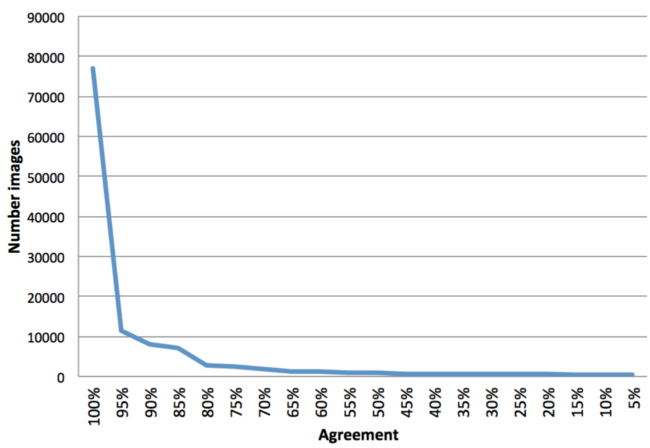

Figure 6: Agreement among players. Only images with more than 10 classifications were included.

To determine whether landscape or satellite images generated more crowd agreement, we calculated the agreement on positions where we have both satellite images and landscape pictures and compared them with each other. The results showed that the landscape pictures had an average agreement of 96.37%, while the satellite images had a 97.52% agreement. This indicates that the satellite images were slightly easier to classify, although not significantly so. Analyzing the agreement for different satellite image zoom levels (250 m = avg. agreement 97.57%, 500 m = 97.41%, 1 km = 97.44%), it can also be concluded that there are no meaningful differences between them.

Although the players agreed on the majority of images, this does not mean that the resulting classifications are always correct. As already described in Section 3.4, players could challenge an answer if they did not agree with the answer that the other players gave. The contested images were then sent to experts who validated the images. If the players were right, they received 5 additional points. In total, players contested 7,599 images. We now consider only contested images for which at least 30 classifications were collected and on which the players agreed with one another more than 95% of the time. This results in a total of 164 contested images, of which the experts agreed with the crowd on 151 of the images (82%), but disagreed on 13 of the images (18%).

We have taken a closer look at these 13 images. For the majority, the players classified these as cropland, yet they were actually pasture. However, this difference is often difficult to distinguish from satellite images, even for experts. For example, the image shown in Figure 7, which most likely shows pasture, was defined as cropland by 45 players, whereas only 1 person said that the image did not contain any cropland. This shows that although the players agree with each other most of the time, it does not necessarily mean that they are always correct. Output agreement alone is therefore not sufficient and guidance by experts is required.

Figure 7: Satellite image of most likely pasture that was classified as cropland by 45 players.

4.4 Reception by the Players

In this section, we discuss how Cropland Capture was received by the players. First, the player behavior is analyzed indirectly by looking at how many classifications have been done by the players in the 25 weeks in which the game ran. We then present the results of a questionnaire where the players were asked about Cropland Capture, including what their incentives were for playing the game.

Figure 8 shows the number of classifications per week. Interestingly, most peaks can be linked directly to press releases. In week 3, an IIASA press release and blog post about the game was published, which brought some initial players. The highest peak can be seen in week 5, when National Public Radio wrote a blog about the game. This resulted in a

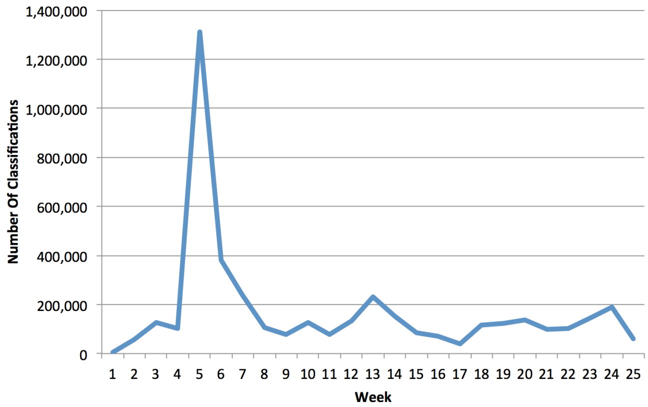

Figure 8: The number of classifications collected per week.

massive amount of data collected following this publication. In week 13, an article in the Guardian mentioned Cropland Capture as one of the top 10 serious games. This shows that press releases about the game have a big effect on the number of classifications collected. What is also interesting is what did not have an effect. In the last 5 weeks of Cropland Capture, weekly prizes were introduced in the hope that the Cropland Capture competition would end in a peak, but as can be seen, these prizes did not have a big impact on the number of classifications collected.

At the end of the competition, a questionnaire was administered in which players were asked what their motivations were for playing the game and what improvements could be made to the game; 22 players responded.

All 22 people said that helping science was a motivation for playing the game. Seven players said they played because the game was fun, while only five players said that the prizes were a motivating factor. Four people said that the competition was a motivation.

One person said that the random questions, which were added later to the game, were penalizing them too much and that they were slowing down the land-cover classifications. Three people said that images should not be repeated so much and that images of bad quality should be removed faster from the game. One person said that it would be good if more explanation could be given about what will be done with the data collected through the game. Two people mentioned that they sometimes gave the wrong answer because they knew that the game would treat the correct answer as wrong and they did not want to lose points.

5. CONCLUSIONS AND FUTURE WORK

We have presented Cropland Capture, a simple cross-platform game for collecting data on cropland, which will be used to improve global cropland maps in the future. As has been shown, the game is very efficient and has been able to collect a huge amount of data from a relatively small number of players. We clarified that although the agreement among the players is very high, this does not mean that the players are always correct, as they sometimes disagree with experts, especially in distinguishing cropland from pasture. Thus, the output agreement sometimes produced the opposite effect desired, i.e., resulted in the collection of incorrect data. We have shown that the press and other media are very important for increasing user participation, and we found that prizes were not the main motivation for why people played the game. It is also clear that the 'maybe' answers were not used that often; more incentives should be given to the players to select this answer when the images are very difficult to classify. Instead, players were driven to score points and therefore chose 'yes' or 'no' over 'maybe' even if the answer might turn out to be wrong.

We will use the data collected from the game to improve our current global hybrid cropland map, which integrates many existing cropland products. The data will also be used to improve cropland extent in those countries where only global maps are available, and in the further validation of this layer and other products in the future.

Given the success of this game, we are currently working on the next version of the game for gathering image classifications. The next game will use the same simple game mechanics where players swipe the images to the 'yes', 'no' and 'maybe' answers or use the keyboard arrow keys to make the classifications. However, there will be a series of changes. For example, the next game will not use the output agreement mechanic where players only score points if they agree with other players. Instead, we will use a mechanic involving expert agreement, which means that the players must agree with the expert classifications. The players will receive points for all images they classify. The points the players get are always incremented by 1, e.g., they will receive 1 point after classifying the first image, for the second image 2 points, for the third image 3 points etc. Once in a while, an image which has already been classified by an expert will be given to the player. If they do not agree with the experts, we will tell them that they are wrong and the score they will get for the next image they classify will start at 1 again. This should have several positive effects. We can teach the players how they should validate, the players will not be influenced by other players, and the reward for classifying an image correctly will go higher and higher the longer the player plays, meaning that the maybe ratings will hopefully be used more if the players are unsure.

6. ACKNOWLEDGMENTS

This work was supported by the CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS) under the Gaming contract (Collecting data via gaming to produce improved land cover products), the EU FP7 funded SIGMA project (No. 603719) and the ERC CrowdLand project (No. 612755).

7. REFERENCES

[1] C. Lintott, K. Schawinski, J. Vandenberg, et al. Galaxy zoo. *Monthly notices of the Royal Astronomical Society*, 389:1179–1189, 2008.

[2] D. Clery. Galaxy Zoo volunteers share pain and glory of research. *Science*, 333 (6039), 173-175, 2011

[3] E. Law, L. Ahn, R. Dannenberg, and M. Crawford. Tagatune: A game for music and sound annotation. *Proceedings of ISMIR*, pages 361–364, 2007.

[4] F. Khatib, S. Cooper, D. Baker, et al. Algorithm discovery by protein folding game players. *Proceedings of National Academy of Sciences USA*, 2011.

[5] G. Anthes. Html5 leads a web revolution. *Communications of the ACM*, 55:16–17, July 2012.

[6] Ho Chien-Ju, Chang Tao-Hsuan, Lee Jong-Chuan, et al. Kisskissban: A competitive human computation game for image annotation. *ACM SIGKDD Newsletter*, 12:21–24, June 2010.

[7] J. Howe. The rise of crowdsouring. *Wired Magazine*, 14 (6), 2006.

[8] J. Kim, M. Greene, K. Lee, M. Richardson, et al. Space-time wiring specificity supports direction selectivity in the retina. *Nature*, 509, 331-336, May 2014

[9] J. Paavilainen. Critical review on video game evaluation heuristics. *Proceedings of Future Play*, pages 56–65, 2010.

[10] K. Pittman, M.C. Hansen, I. Becker-Reshef, et al. Estimating global cropland extent with multi-year MODIS data. *Remote Sensing*, 2(7):1844-1863, 2010.

[11] L. Ahn and L. Dabbish. Labeling images with a computer game. *Proceedings of ACM SIGCHI*, pages 319–326, 2004.

[12] L. Ahn, M. Kedia, and M. Blum. Verbosity: A game for collecting common-sense knowledge. *Proceedings of ACM SIGCHI*, pages 75–78, 2007.

[13] L. See and S. Fritz. A method to compare and improve land cover datasets. *IEEE Transactions on Geoscience and Remote Sensing*, pages 1740–1746, July 2006.

[14] L. See, S. Fritz, L. You, M. Herrero, et al. Highlighting continued uncertainty in global land cover maps. *Global Food Security*, 2014

[15] L. See, D. Schepaschenko, M. Lesiv, I. McCallum, et al. Building a hybrid land cover map with crowdsourcing and geographically weighted regression. *ISPRS Journal of Photogrammetry and Remote Sensing*, 10.1016/j.isprsjprs.2014.06.016, 2014

[16] L. Yu, J. Wang, N. Clinton, et al. 30 m global cropland extent derived through multi-source data integration. *International Journal of Digital Earth*, 6(6):521-533, 2013.

[17] M.A. Friedl, D. Sulla-Menashe, B. Tan, et al. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, 114(1): 168-182, 2010.

[18] M. Krause, A. Takhtamysheva, M. Wittstock, and R. Malaka. Frontiers of a paradigm: exploring human computation with digital games. *Proceedings of the ACM SIGKDD Workshop on Human Computation*, pages 22–25, 2010.

[19] M. Krause and J. Smeddinck. Human computation games: A survey. *Proceedings of the 19th European Signal Processing Conference*, September 2011.

[20] P. Defourny, C. Vancustem, P. Bicheron, et al. GLOBCOVER: A 300m global land cover product for 2005 using ENVISAT MERIS time series. *Proceedings of the ISPRS Commission VII Mid-Term Symposium: Remote Sensing: from Pixels to Processes*, Enscede NL, 2006.

[21] P. Mayaux, H. Eva, J. Gallego, A. Strahler, M. Herold

et al. Validation of the global land cover 2000 map. *IEEE Transactions on Geoscience and Remote Sensing*, 44(7), 1728-1737.

[22] R. Bonney, C. Cooper, J. Dickinson, S. Kelling, T. Phillips, K. Rosenberg, and J. Shirk. Citizen science: A developing tool for expanding science knowledge and scientific literacy. *BioScience*, 59 (11), 977-984, 2009

[23] S. Cooper, A. Treuille, J. Barbero, Z. Popovic, et al. The challenge of designing scientific discovery games. *FDG*, pages 40–47, 2010.

[24] S. Cooper, F. Khatib, A. Treuille, Z. Popovic, et al. Predicting protein structures with a multiplayer online game. *Nature*, 466:756–760, 2010.

[25] S. Fritz, I. McCallum, C. Perger, et al. Geo-wiki.org: The use of crowdsourcing to improve global land cover. *Remote Sensing*, pages 345–353, August 2009.

[26] S. Fritz, L. See, I. McCallum, C. Schill, M. Obersteiner, M. Velde, et al. Highlighting continued uncertainty in global land cover maps. *Environmental Research Letters*, 6:4:044005, 2011.

[27] S. Fritz, L. See, I. McCallum, C. Schill, M. Velde, M. Obersteiner, et al. Downgrading recent estimates of land available for biofuel production. *Environmental Science and Technology*, 47(3):1688-1694, February 2013.

[28] T. Sturn, D. Pangerl, L. See, et al. Landspotting: A Serious iPad Game for Improving Global Land Cover. *GI-Forum*, Salzburg, Austria, July 2013.

[29] T. Sturn, M. Wimmer, P. Purgathofer, S. Fritz. Landspotting - Games for Improving Global Land Cover. *FDG*, Chanica, Greece, May 2013.