Mon. Not. R. Astron. Soc. **000**, 1–**??** (2012) Printed 15 August 2013 (MN LaTeX style file v2.2)

Milky Way Project : Clouds and Holes Paper

R. Simpson[1], C. J. Lintott[1,2], C. E. North[3]

[1]*Oxford* [2]*Adler* [3]*Cardiff*

Accepted 20xx. Received 20xx.

ABSTRACT

Abstract

1 INTRODUCTION

Infrared dark clouds (IRDCs) were first observed as dark regions silhouetted against the mid-infrared (MIR) background (**?**). Subsequent observations showed them to have low temperatures and high densities ($T \lesssim$K, $n_H > 10^5\,\mathrm{cm}^-3$, e.g. **???**). The IRDC absorbs the background light and causes a dip in the MIR sky brightness. They are thought to be the earliest observable formation stages high-mass stars and stellar clusters.

With MIR data alone, however, it is impossible to distinguish this absorption from a region of inherently lower background emission. Far-Infared (FIR) observations allow IRDCs, which appear bright at wavelengths above 24 μm, to be distinguished from regions of lower emission, which remain dark.

2 INPUT CATALOGUE

Describe catalogue (Fuller & Peretto) Spitzer & Herschel Hi-Gal observations

Number of objects etc.

3 ANALYSIS

Brief outline of analysis

3.1 Milky Way Project

The Milky Way Project[1] was established in 2010 as a citizen science interface to data from the *Spitzer* GLIMPSE survey primarily as a search for 'bubbles' associated with massive star formation. This effort was successful, and a catalogue of more than 5000 such bubbles which expanded on previous efforts by professional astronomers was published by **?** and used for a statistical analysis of bubble distribution by **?**. Inspired by this success, a second interface was added to the site in order to address the problem of identifying true IRDCs.

As with the previous interface, this new part of the site[2] makes use of the Zooniverse Application Programming Interface (API) originally built for Galaxy Zoo (**?**) and which supports a large number of similar citizen science projects. This API is primarily responsible for serving images and recording classifications provided by volunteers, who are required to be logged in for their work to be recorded. The interface itself is built in JavaScript and HTML5. Following a short tutorial, an image is selected from the database[3] and presented to the volunteer who may label it as a CLOUD, a HOLE or an INTERMEDIATE case by selecting one of three buttons, as shown in figure 1. Examples of both clouds and holes are provided during the tutorial phase and can be reviewed at any time; these examples are shown in 2 in both colour schemes available to classifiers. Once an image is classified, the volunteer is shown another image and presented with the opportunity to discuss the first image with other classifers[4].

DETAILS OF IMAGES AND IMAGE CONSTRUC-

Figure 1. The interface for the Clouds part of the Milky Way Project, as seen by classifiers who had to sort each image into one of three categories.

[1] http://www.milkywayproject.org
[2] http://www.milkywayproject.org/clouds

[3] Volunteers see an image they have not yet classified, selected randomly from those with the fewest classifications in the database. This algorithm for task assignment has the advantage of ensuring that images have approximately the same number of classifications at all times, facilitating preliminary data analysis.
[4] See http://talk.milkywayproject.org

TION.. Participants can chose to alter the colour palette presented (DO WE RECORD IF PEOPLE USE THIS OPTION?)

The clouds project was launched on XXXXXXX and ran until XXXXXXX. 3544 logged-in users provided classifications. However, approximately half of the classifications were received from those who were not logged in. The most active user has completed 59020 classifications and is one of three users to have seen every image provided. (HOW DO WE HANDLE THE REPEATS IN DATA REDUCTION). 3253 classifiers provided more than five classifications (91.9%, compared to 25% in the previous incarnation of the Milky Way Project), 1843 more than fifty (52.0% compared to 5.7%) and 168 five hundred classifications or more.

DETAILS OF NON-LOGGED IN USERS

MWP interface and initial results (or in results section?)
User and classification numbers and duration of dataset used

Raw classifications? Histogram of results?

Thumnail examples of clouds, holes and unknowns

3.2 Experts and Training data

Definition of training data definition

Thumbnails of training data and classifications

Expert versus general results (plot)

3.3 Analysis sequence

The procedure Growth charts Results of MC runs Any threshholds

4 RESULTS

Results from analyis (cloudiness chart)

Histogram of classifications before & after

Thumnail examples of clouds, holes and unknowns

Full table of results

5 CONCLUSIONS

6 ACKNOWLEDGEMENTS

<Zooniverse acknowledgements>

Herschel is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA.

SPIRE has been developed by a consortium of institutes led by Cardiff Univ. (UK) and including: Univ. Lethbridge (Canada); NAOC (China); CEA, LAM (France); IFSI, Univ. Padua (Italy); IAC (Spain); Stockholm Observatory (Sweden); Imperial College London, RAL, UCL-MSSL, UKATC, Univ. Sussex (UK); and Caltech, JPL, NHSC, Univ. Colorado (USA). This development has been supported by national funding agencies: CSA (Canada); NAOC (China); CEA, CNES, CNRS (France); ASI (Italy); MCINN (Spain); SNSB (Sweden); STFC, UKSA (UK); and NASA (USA).

This research made use of APLpy, an open-source plotting package for Python hosted at http://aplpy.github.com

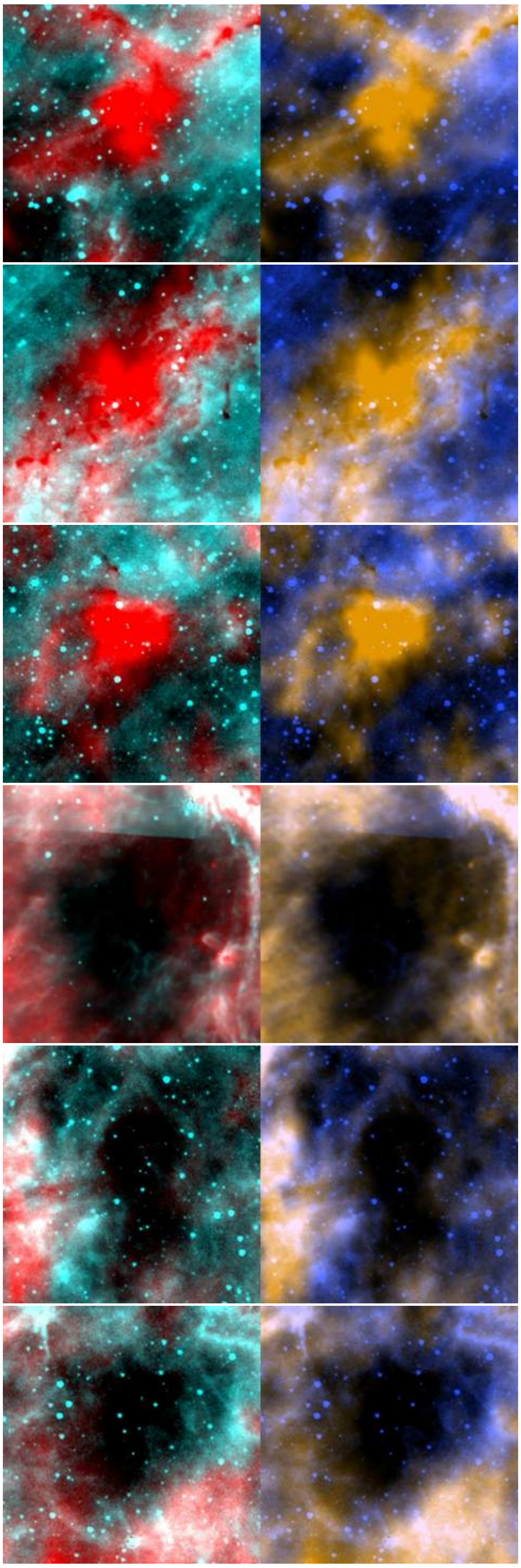

Figure 2. Top three rows : Three of the nine examples of true clouds used in the tutorial. Bottom three rows : Examples of holes given in the tutorial. The tutorial did not include examples of 'intermediate' images.