1 Title: The VISTA Galactic Plane Survey

PI: P.W. Lucas (1) & E.L. Martin (2); CoIs are listed in Section 9.
(1) University of Hertfordshire, England; (2) Instituto de Astrofisica de Canarias, Spain

1.1 Abstract:(10 lines max)

Near infrared data are required to probe the many parts of the galaxy which suffer high extinction. An ESO-public near infrared survey of the northern plane, the UKIDSS Galactic Plane Survey, has recently begun and optical public surveys covering the whole galactic plane are presently underway: the southern VPHAS+ survey and northern IPHAS survey. We propose a 5 band (ZYJHKs) survey of the southern galactic plane to complete the infrared and optical coverage of the galaxy, providing a legacy database with the same latitude range as these existing surveys ($|b| < 5$) and 4 magnitudes better sensitivity than 2MASS. Southern coverage is invaluable since most of the galactic disc is located at southern declinations. We describe how optical, near infrared and mid-infrared data for the whole galaxy provide the means to statistically investigate the star formation process in unprecedented detail, probe galactic structure for young and old populations, determine cluster distances and estimate spectral types and luminosity classes for many field stars and cluster members.

2 Description of the survey: (Text: 3 pages, Figures: 2 pages)

2.1 Scientific rationale:

A large area near infrared (IR) survey of the galactic plane (GPS) will be beneficial to many areas of stellar astrophysics. Using 1 epoch of data at ZYJHKs and 2 further epochs only in the Ks band it will:

- detect all O stars in the Galaxy to $d = 30$ kpc with extinction $A_V < 36$ mag and all B stars with $A_V < 12$ mag.
- provide deep stellar luminosity functions (LFs) for large numbers of star formation regions, detecting any environmental dependence in the Initial Mass Function (IMF) and measuring the cluster LF as well.
- map galactic structures such as the disc, the inner bar and the southern warp in unprecedented detail.
- determine distances and extinctions towards clusters through star counts and colours.
- statistically describe luminosity class, approximate spectral type and extinction for stars along every line of sight, and estimates for bright individual stars, using several reddening independent colour indices.
- provide a Legacy Database of $\sim 3 \times 10^9$ stars that will enable brief, rarely observed evolutionary stages of normal stars that are poorly understood to be studied. Eg. FU Ori-type protostars, LBVs, and AGB stars that are undergoing unstable shell burning will be detected via their large amplitude variability in the 3 epochs.
- provide essential prior data on X-ray transients in the quiescent state to understand these systems and guide follow up when flares occur.
- pick out nearby stars and brown dwarfs to $d \approx 100$ pc through proper motion measurements.
- discover hundreds of galaxies in the Zone of Avoidance, mapping the structure of the local Supercluster and determining whether the hypothetical "Great Attractor" is visible in ordinary baryonic matter.
- discover and map the structure of far more post-AGB stars, PPN and PNe than are presently known.
- be an invaluable resource combined with other existing and planned surveys including VPHAS+, IPHAS, GAIA, UKIDSS-GPS, SPITZER-GLIMPSE, WISE, MSX, ASTRO-F, HERSCHEL Hi-Gal, SCUBA-2 GPS, SCUBA-2 SASSy and the INTEGRAL-IBIS survey.

The proposed VISTA-GPS is conceived as a Legacy database and Atlas, with the goal of providing complete coverage of the Galactic Plane in combination with the UKIDSS Galactic Plane Survey of the northern plane (www.ukidss.org). VISTA-GPS will be complemented by several other major surveys of the galactic plane which are sensitive to stellar photospheres, eg. VST/VPHAS+ SPITZER-GLIMPSE and WISE (see section 3).

Survey Area, Depth and Concept

VISTA-GPS will cover the southern plane from galactic longitude $l = 225$ to $l = 20$, in latitude range $|b| < 5$. This provides a small overlap with the UKIDSS GPS survey. The latitude coverage is the same as that of several complementary surveys, see section 3. We propose a survey depth which is close to the confusion limit in the galactic plane, and will in fact reach that limit in the mid-plane, see section 4. Five photometric bands, ZYJHKs, are desirable in order to derive basic stellar characteristics (see below). These fluxes should be obtained at the same epoch in order to minimise the effects of variability on source colour and derived properties.

In addition to the main survey at ZYJHKs in the first epoch *we propose to observe at 2 further epochs at Ks band* in order to provide variability data for all sources and provide proper motions for all nearby sources. 3 epochs are a minimum to identify large amplitude variables with confidence and to distinguish variables from high proper motion stars in crowded galactic fields, where source confusion is significant. Proper motion data also assist detection of brown dwarfs in the plane, where single-epoch photometric identification is difficult [1].

The variability data will address a key science goal, which is the detection of very rare and brief phases of stellar evolution such as transient events in AGB stars and FU Orionis outbursts in protostars, which are presently poorly understood due to the small numbers of known objects in each class. The 3 epochs of observation must be spread over a period of at least 5 years to provide useful proper motion precision and increase the chances of detecting these types of high amplitude variable.

Below we describe a sample of the other main science drivers which influence the survey design.

Star Formation

A consensus is beginning to develop concerning the main evolutionary stages of low mass star formation but the picture becomes very uncertain for the youngest objects and for massive stars. High mass YSOs swiftly settle on to the Main Sequence and blow away their obscuring circumstellar material, becoming visible in the IR and then in the optical. Their presence in a cluster can profoundly influence the evolution of lower mass YSOs by triggering more star formation or by photoionising the envelopes of low mass systems which are still forming. Hence we might expect the IMF to be influenced by environmental variables such as the presence or absence of massive stars, as well as cluster density, metallicity and initial turbulence and angular momentum. VISTA-GPS will measure the luminosity function to very low masses in hundreds of star formation regions (SFRs). It will provide a statistically useful sample covering *all types of SFR* at each evolutionary stage, detecting any correlations between IMF and environment.

The distribution of SFRs in the galaxy is at present inferred primarily from low resolution CS mapping, mid-IR *MSX* data and far-IR *IRAS* data, eg.[2],[3],[4]. We wish to learn how the density of SFRs varies with distance from the galactic centre and search for spatial variations in the cluster LF. VISTA-GPS will also yield a global measurement of the cluster LF, which may well include new very massive clusters similar to Westerlund 1 [5] and NGC3603, which appear to be akin to the Super Star Clusters found in starburst galaxies.

The combination of optical VPHAS+ data, near IR VISTA-GPS data and mid-IR data from SPITZER-GLIMPSE and WISE will be very helpful in providing a comprehensive view of each SFR, showing the spatial differences between more evolved optically visible parts of the SFR (often HII regions) and the more obscured molecular parts which are visible only in the near and mid-IR. The Hα mapping of VPHAS+ and detection of mid-IR excesses due to discs will establish YSO status for individual sources more efficiently than VISTA-GPS data alone. However VISTA-GPS will detect the largest number of sources, owing to the effect of extinction in the optical and inferior spatial resolution and sensitivity in the mid-IR, eg. Ks=18.4 in VISTA-GPS compared to L$'$=15.4 in SPITZER-GLIMPSE. 40-50% of YSOs detected at J, H and Ks may be identified with confidence via Ks band excesses or via variability in the 3 Ks band epochs.

The Atlas and Galactic Structure

By imaging in 5 near IR bands it will be possible to construct an Atlas of the galaxy with 3-D information about the distances of stellar clusters and the location of obscuring dust clouds. In Figure 1 we illustrate how the extinction structure of the southern star formation region M17 was mapped with JHK data. The Atlas will map the extinction throughout the survey region, in SFRs and quiescent clouds. The distances of stellar clusters can be determined by calculating an average extinction value and then finding the best fit stellar isochrone.

A southern galactic survey is essential to properly explore galactic structure, since the inner regions of the galaxy and the far side of the galaxy (including the southern warp at $l \approx 250^\circ$) all lie at southerly declinations. Major galactic features such the central galactic bar and the warp in the outer galaxy can be detected and distances calculated either by using distinctive features such as the 'red clump' giant population in a two colour diagram or using star counts as a function of reddening, eg. [6],[7]. Previous large scale surveys (DENIS & 2MASS) have provided much useful data on galactic structure but their limited spatial resolution and sensitivity restrict them to observations of relatively luminous and old giant stars, with less information about faint distant clusters and younger stellar populations. Comparison of galactic structures as mapped out by old giant stars, young disc stars and previous 21 cm radio studies may detect changes in galactic structure caused by galactic mergers, probing the accretion history of the galaxy. We emphasise that the near IR extinction law is almost uniform throughout the interstellar medium [8] in contrast to the spatial variations seen in the visible. This is because dust grain radii are much smaller than the wavelength of observation for $\lambda \geq 1$ μm.

In addition, VISTA-GPS data in all 5 of the ZYJHKs bands makes it possible to estimate luminosity class (distinguishing dwarfs from giants), a crude spectral type (hot, sun-like, cool or very cool) and an extinction value for many individual normal field stars with 5 filter detections. This is done using reddening independent colour indices (see Figure 2), which break the serious degeneracy caused by reddening in conventional J-H, H-K near IR diagrams. These indices have limited value individually, but by using several indices, some including optical VPHAS+ data, it will be possible to check the results for consistency. Much more secure statistical results will be obtained for the whole population along each line of sight, as has previously been done at blue optical wavelengths[9]. VPHAS+ Hα photometry will distinguish spectral types near A0 from other early types by their strong Hα absorption. Note that the Z and Y bands are very important for this approach. Y band fluxes are needed in the most useful reddening independent index that can be constructed from the ZYJHKs filters, see Figure 2. Z band fluxes are also used in some of these indices. More importantly, the Z band has better sensitivity than the YJHKs bands if $A_V < 5$ and it is the most temperature sensitive of the 5 bands.

Investigating the Galactic X-ray Source Population

VISTA-GPS data will determine the nature of large numbers of currently unidentified X-ray point sources, many of which appear to be Cataclysmic Variables (CVs) and Active Stars[10]. These are often obscured from view by extinction in the optical. In addition *XMM-Newton* and CHANDRA continue to locate many new supernova remnants, SFRs and X-ray binaries, all of which will be detected by VISTA-GPS in statistically useful numbers. The X-ray emission from SFRs will aid in the identification of very young stars and clusters (5-20 Myr) detected by VISTA but lacking significant extinction or IR excesses after star formation is completed.

VISTA-GPS "prior data" on X-ray transients during quiescence is vital to understand them and plan follow up when they draw our attention through an X-ray flare. IR colours allow, by comparison of data taken when sources were in quiescence and active, the unraveling of the physical properties of systems such as low mass X-ray binaries and the details of the outburst events. VISTA-GPS will provide photometry and colours for several hundred known X-ray binaries. Many more will be identified via variability in a multipass survey, which is crucial in cases where there are several possible IR counterparts in an X-ray location box a few arcsec wide.

2.2 Immediate objective:

The aim is a photometric catalogue of the southern plane at 3 epochs. The first will be at ZYJHKs, to be followed by two further epochs at Ks only, to provide variability and proper motion data, with intervals of at least 2 years between each epoch for any given sky location. 2nd and 3rd epoch data may be obtained in non-photometric conditions, since the Vista Data Flow System team report that very good relative photometry within each field can be performed through thin cirrus to detect variability. 2MASS data can also be used for Ks band photometric calibrations of thin cirrus data. Exposure times of 80s at ZYJH and 60s at Ks yield 5 σ depths of Z=21.6, Y=21.0, J=20.5, H=19.6, K=18.4 (Vega system), though confusion will reduce survey depth by ~ 1 mag within 10° of the galactic centre.

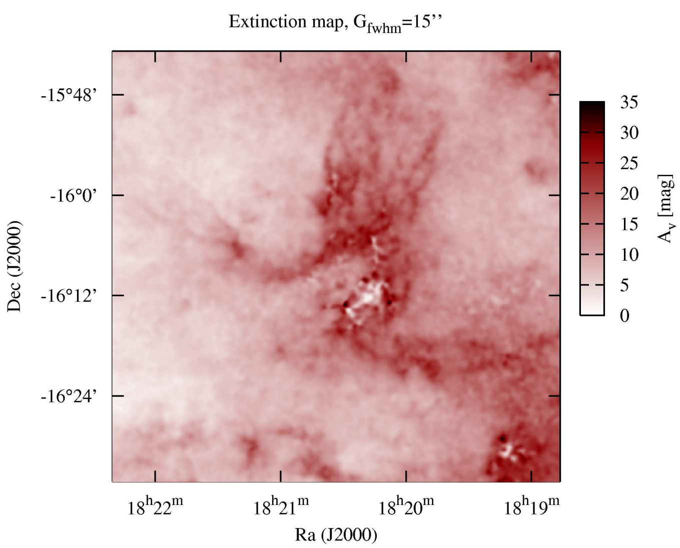

Figure 1: Extinction map of M17 (the Omega Nebula), a distant high mass star formation region. This map was produced from JHK UKIDSS commissioning data by averaging the reddening solutions for each stars into 15 arcsec bins, assuming M dwarf colours, (a good first-order assumption for most sources in SFRs). The molecular cloud around the HII region inside the Ω shape is clearly defined.

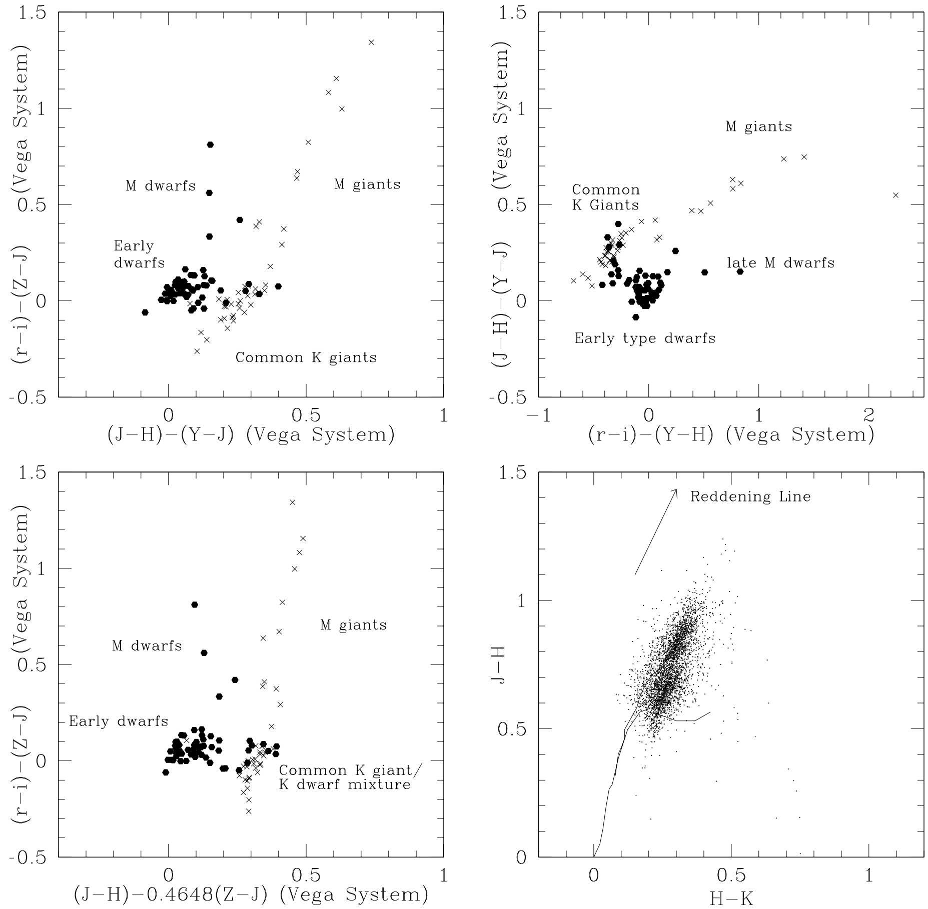

Figure 2: Reddening independent colour indices (*top panels and lower left)*. These were constructed from data in the UKIDSS photometry paper [11], using the Cardelli, Clayton & Mathis extinction law [8], which follows a power law that is independent of the R_V parameter at $\lambda > 0.9$ μm. The upper panels (which require Y band data) quite clearly distinguish both dwarfs (circles) from giants (stars) and early types from M types. *(lower left panel)* One of the many other reddening independent indices that do not use Y band and give less clear separation between dwarfs and giants. *(lower right)* A conventional 2-colour diagram, using UKIDSS data for a typical galactic field at l=+28, plotting 17000 stars with photometric errors <0.05 mag. The reddening vector creates highly problematic degeneracy between main sequence dwarfs (lower curve) and giants (upper curve).

References

(**1**) Burgasser A., Kirkpatrick J.D., et al., 2003, AJ 125, 850

(**2**) Bronfman L, Casassus S, May J, Nyman L A, 2000, A&A 358, 521. (**3**) Casassus S, Bronfman L, May J, Nyman L A, 2000, A&A 358, 514. (**4**) Lumsden, S.L., Hoare, M.G., et al., 2002, MNRAS 336, 621. (**5**) Clark J.S., et al., 2005, A&A 434, 949 (**6**) Lopez-Corredoira M., Hammersley P.L., et al., 2001, A&A 373, 139. (**7**) Lopez-Corredoira M., Cabrera-Lavers A., Garzon F., Hammersley P.L., 2002, A&A 394, 883. (**8**) Cardelli J A, Clayton G C, Mathis J S, 1989, ApJ 345,245. (**9**) Harris J., Zaritsky D., Thompson I., 1997, AJ 114 1933. (**10**) Motch et al., 2006, A&A (in prep). (**11**) Hewett P.C., Warren S.J., Leggett S.K., Hodgkin S.T., 2006, MNRAS (in press), astro-ph/0601592.

3 Are there ongoing or planned similar surveys? How will the proposed survey differ from those? (1 page max)

We are aware of the existence of a proposal led by Minniti to do a 20x18 deg variability study in the galactic bulge. Our two proposals, if both selected, should jointly consider how to make and schedule their observations in the common 20x10 sq deg subset of this area, and share the data, to maximise the science output for both surveys, and maximise efficient use of VISTA.

The VHS hemisphere proposal uses ZYJK bands and will aid VISTA-GPS by defining colour-magnitude sequences in different stellar populations with negligible reddening. If VHS is approved it would be useful to apply these relations to VISTA-GPS.

Co-PI Lucas is PI of the UKIDSS Galactic Plane Survey (www.ukidss.org/surveys) which began in May 2005 and is covering the northern plane down to a southerly declination limit of Dec=-15° (corresponding to $l < 230$ and $l > 15$), with the same latitude limits ($|b| < 5$) and JHK sensitivity as this survey. VISTA-GPS will therefore complement the UKIDSS GPS by completing the coverage of the galaxy. Both surveys are public to ESO members and the archive in Edinburgh will permit simultaneous and sophisticated and searches of both databases using SQL queries. Note that the present UKIDSS GPS plan also includes a narrow southern extension at $|b| < 2$ from l=15 down to l=-2. However it is unclear whether these observations will actually be made: no observations of this southern strip were made in 2005 due to the difficulty of obtaining good image quality near the galactic centre due to the high airmass when observing from Mauna Kea. If there is no success in 2006 the UKIDSS GPS southern strip will be abandoned, since it would be done better by VISTA.

The VST/VPHAS+ and INT/IPHAS optical surveys cover the southern and northern hemispheres respectively, also with a $|b| < 5$ limit. Co-PI Martin is a member of the VPHAS+ and IPHAS teams and CoI Drew is the PI of these surveys. In the mid-IR the SPITZER-GLIMPSE survey has covered 220 square degrees of the mid-plane already and the rest of the plane is expected to be observed in the WISE NASA Explorer mission in 2009, shortly after the completion of the first phase of VISTA-GPS. We highlight these optical and mid-IR surveys, since the main data products are photometry of stellar photospheres, which is also true of VISTA-GPS.

In addition, valuable complementary surveys of the Galactic Plane in the far IR and submm wavebands will be conducted during the period of VISTA-GPS. MIPSGAL is a SPITZER MIPS survey covering the SPITZER-GLIMPSE 220 square degree region at 24 and 70 μm. The far IR satellite *Akari* (ASTRO-F) has recently been launched successfully and will survey the whole sky in 3 bands at 50–150 μm. The 850 μm JCMT SCUBA-2 survey "SASSY" of the whole sky has recently been approved (starting in early 2007). It will cover all of the Galactic Plane at Dec$> -40°$ and $|b| < 5$ in the first 2 years, thereby providing substantial overlap with VISTA-GPS. Finally the proposed Hi-Gal survey with HERSCHEL aims to map the Galactic Plane at $|b| < 2$, $l = 0$–360 at several wavelengths in the 70–500 μm range. These datasets have relatively poor spatial resolution (typically 15 arcsec) but will aid VISTA-GPS by providing the locations of the most deeply embedded (Class 0) protostars and determining the physical conditions and chemical make up of the ambient medium in SFRs.

4 Observing strategy: (1 page max)

To reach the $5-\sigma$ sensitivity limits referred to in section 2.2 requires a time on object of 80 s at ZYJH and 60 s at Ks, according to the Exposure Time Calculator. This assumes seeing of 0.8 arcsec and an 1.2 arcsec software aperture, optimised for maximum sensitivity. Data from UKIDSS and coIs on this proposal show that these limits border on the confusion limit, depending how this is defined (see http://wiki.astrogrid.org/bin/view/VISTA/VGPS). Very similar limits are reached by the UKIDSS Galactic Plane Survey at JHK. At the shorter wavelengths the greater sensitivity of VISTA partly compensates for greater extinction, providing similar source counts in all filters. The slightly shorter time on object at Ks band reflects the fact that confusion is greatest in this band and the fact that the 3 Ks band epochs will eventually provide a 180 s total time.

These exposure times permit good observing efficiency. We desire 4 independent pixels at YJH (2 jitters at each of the 2 pawprints covering a particular pixel) in order to provide robust photometry, minimising effects of bad pixels and any flat fielding errors. At Z and Ks band we accept only 2 independent pixels (observed on different pawprints during the same OB), since the overheads associated with jittering otherwise lead to very poor observing efficiency with 60 s time on sky. Only 2 independent background limited pixels (40 s DIT) are possible at Z band for 80 s time on object, if we adopt the recommended minimum 40 sec DIT for Z. The 3 Ks band epochs will provide additional independent pixels by introducing small shifts in the OBs at each epochs. Detailed strategy, and time to cover a tile is shown in the table. Efficiencies are from the ETC v1.1.

Filter	Method	Time on object	Elapsed Time	Efficiency
Z band	40 s DIT, NDIT=1, Nexp=1, Njitter=1, Npaw=6	80 s	332 s	72.3%
Y band	20 s DIT, NDIT=1, Nexp=1, Njitter=2, Npaw=6	80 s	410 s	58.5%
J band	10 s DIT, NDIT=2, Nexp=1, Njitter=2, Npaw=6	80 s	422 s	56.9%
H band	10 s DIT, NDIT=2, Nexp=1, Njitter=2, Npaw=6	80 s	422 s	56.9%
K band	10 s DIT, NDIT=3, Nexp=1, Njitter=1, Npaw=6	60 s	284 s	63.4%

The first epoch ZYJHKs data will be taken together in a single OB that covers 1 tile, in order to minimse the effects of variability. The OB typical duration is 2020 s or 33.7 min. This includes 1876 s (ETC v1.1 value) and 5×30 s for filter changes (moving to adjacent filters in the wheel wherever possible to minimise the time).

The 2 Ks-only epochs use OBs which cover 4 adjacent tiles with no filter changes. The OB duration is as follows: 284 s for each tile (from the ETC v1.1) multiplied by 4=1136 s = 18.9 min.

Observing Constraints

Clear time is requested for the first epoch at ZYJHKs to provide reliable photometry. Photometric time would be preferable but the combination of 2% 'photometric' time as defined by ESO and 0.8 arcsec seeing (sse below) is too rare for a large survey.

Thin cirrus time is accepted for the 2 additional epochs at Ks band only. The Vista Data Flow System team reports based on UKIDSS data are that very good relative photometry within each field can be performed through thin cirrus to detect variability. They also report that 2MASS data can be used for absolute photometric calibrations of both the clear and thin cirrus data at JHKs, with a precision of $\approx 2\%$.

0.8 arcsec seeing or better in the focal plane (at Ks band) is required in order to minimise the effect of source confusion. (This is identical to the UKIDSS GPS limit). This will lead to fairly uniform resolution throughout the survey, since the 0.34 arcsec pixels and an instrumental contribution of up to 0.51" to the seeing mean that very good natural seeing conditions provide only a small improvement in the effective resolution.

Bright time is accepted for all observations. The moon has negligible effect on sensitivity at JHKs and occasional small sensitivity losses must be accepted at Z and Y if the survey is to be completed in a timely fashion.

Timing of Epochs

The 3 epochs of observation for each tile must be observed at least 2 years apart to provide useful proper motion and variability data (see science case and Estimated Observing Time).

5 Estimated observing time:

An example of the possible allocations per semester is shown in the table below, assuming for simplicity that the survey begins at full operating efficiency at the start of Period 79.

Period	Time (h)	Mean RA	Moon	Seeing	Transparency
P79	169	10h	bright	0.8	clear
P80	121	10h	bright	0.8	clear
P81	169	15.5h	bright	0.8	clear
P82	121	10h	bright	0.8	clear
P83	24	10h	bright	0.8	thin
P84	17	15.5h	bright	0.8	thin
P85	24	10h	bright	0.8	thin
P86	17	15.5h	bright	0.8	thin
P87	24	10h	bright	0.8	thin
P88	17	15.5h	bright	0.8	thin
P89	24	10h	bright	0.8	thin
P90	17	15.5h	bright	0.8	thin

5.1 Time justification: (1 page max)

According to the ETC a time on object of 80s at ZYJH and 60s at Ks is required to yield 5 σ depths of Z=21.6, Y=21.0, J=20.5, H=19.6, K=18.4 (Vega system), assuming 0.8 arcsec seeing and a seeing-optimised 1.2 arcsec software aperture. This is the desired depth as indicated in the Science Case, being close to the confusion limit in most fields (see http://wiki.astrogrid.org/bin/view/VISTA/VGPS), and equal to the UKIDSS depth at JHK. Note that VISTA has better sensitivity at Z band than the VST has at z', which is why Z is requested in VISTA-GPS rather than VPHAS+. Note that the definition of the confusion limit is somewhat arbitrary, and there is no sudden point at which longer time on object ceases to be useful. If it is defined as >1 source per 50 resolution elements it will formally be reached at b=0 at most longitudes but unconfused photometry is still provided for the majority of sources down to the sensitivity limit of Ks=18.4, and even overlapping sources can often have separate fluxes extracted using profile fitting photometry.

Efficiency (time on sky divided by OB duration) is 56.4% for the ZYJHKs OBs (1140 s on sky and 2020 s duration including filter change overheads, see previous section). For the Ks-only OBs the efficiency is 63.3%.

The area of the survey is 1550 deg^2, corresponding to a 155 degree long strip of the galactic plane which is 10 degrees wide, located south of Dec=-12°. The OB durations are 33.7 minutes for ZYJHKs OBs in the first epoch, which cover a 1.5 deg^2 tile, and 18.9 minutes for the Ks-only OBs in the other 2 epochs, which cover 4 tiles or 6 deg^2. Assuming that the 1 degree wide tiles are optimally rotated with the Cass rotator to span the latitude range in exactly 10 tiles then the area can be tiled with 100% efficiency. A simple calculation then gives a total observing time requirement of 743 hours, composed of 580.4 hours for the ZYJHKs first epoch and 81.4 hrs for each of the Ks-only epochs.

Note that we have made no allowances for photometric calibration observations, which we assume will be made in standard fields at regular intervals each night for all public surveys.

The Galactic Plane at Dec$< -12°$ is located in the RA range $7.5h < RA < 18h$. The optimal times for observation at low airmass are from mid-January to early July, leading to a fairly even split between the ESO winter and summer Observing Periods, but a slightly greater proportion of time in the Winter Periods, which run from 1st April to 30 September.

All observations can use Bright Time and a ≤ 0.8 arcsec seeing constraint in the focal plane in the Ks band, where source confusion will be worst. This will typically correspond to slightly worse seeing in the shorter wavelength bands.

It seems reasonable that whole survey region may be observed at ZYJHKs in the first 2 years of VISTA operation (308 hours per year) depending on the number of surveys which are awarded time. If this plan is endorsed then the Ks only data for the 2nd epoch should be taken in years 3 and 4 (just 41 hours per year) to provide a 2 year interval between epochs. The 3rd epoch (also Ks only) would than take place in years 5 and 6, also taking only 41 hours per year.

6 Data management plan: (3 pages max)

6.1 Team members:

Name	Function	Affiliation	Country
CASU (VDFS) team	Pipeline processing	University of Cambridge	UK
CASU (VDFS) team	Data Quality Control-I	University of Cambridge	UK
J. Emerson	VDFS Coordinator	Queen Mary University of London	UK
WFAU (VDFS) team	Science Archive	University of Edinburgh	UK
WFAU (VDFS) team	Data Quality Control-II	University of Edinburgh	UK
	This VISTA-GPS Survey specific tasks		
Lucas	Data Quality Control-III	University of Hertfordshire	UK
Lucas & Hoare	Algorithm Development	U Herts and U Leeds	UK

The CASU (VDFS) team consists of Irwin, Lewis, Hodgkin, Evans, Bunclark, Gonzales-Solares, Riello. The WFAU (VDFS) team consists of Hambly, Bryant, Collins, Cross, Read, Sutorius, Williams.

6.2 Detailed responsibilities of the team:

We will use the VISTA Data Flow System (VDFS; Emerson et al. 2004, Irwin et al. 2004, Hambly et al. 2004) for all aspects of data management, including: pipeline processing and management; delivery of agreed data products to the ESO Science Archive; production of a purpose-built IVOA compliant science archive with advanced datamining services; enhanced data products including federation of VISTA survey products with SDSS survey products. Standardised agreed data products produced by VDFS will be delivered to ESO, with copies remaining at the point of origin. The VDFS is a collaboration between the UK Wide Field Astronomy Units at Edinburgh (WFAU) and Cambridge (CASU) coordinated by the VISTA PI (QMUL) and funded for VISTA by PPARC. The VDFS is a working systems-engineered system that is already being successfully employed for the UKIRT WFCAM surveys as a test bed for the VISTA infrared surveys, and which is sufficiently flexible as to be applicable to any imaging survey project requiring an end-to-end (instrument to end-user) data management system. We emphasise the track record over the last decade of both the Cambridge and Edinburgh survey units in processing and delivering large-scale imaging datasets to the community as exemplified by the WFCAM Early Data release (EDR, (*http://surveys.roe.ac.uk/wsa/dboverview.html*) Lawrence et al 2006, Dye et al 2006).

6.3 Data reduction plan:

The data reduction will be using the VDFS, operated by the VDFS team, and augmented by individuals from VISTA-GPS, especially for product definition and product Quality Control. We divide the plan into two distinct but intimately related parts: pipeline processing and science archiving. Much greater detail can be found in the SPIE papers cited previously,

6.3.1 Pipeline processing

The Cambridge Astronomy Survey Unit (CASU) are responsible for the VDFS pipeline processing component which has been designed for VISTA and scientifically verified by processing wide field mosaic imaging data from UKIRT's NIR mosaic camera WFCAM and is now routinely being used to process data from the WFCAM at a rate of up to 250GB/night. It has also been used to process ESO ISAAC data e.g. the FIRES survey data and a wide range of CCD mosaic camera data.

The pipeline is a modular design allowing straightforward addition or removal of processing stages and will have been tested on a range of input VISTA datasets. The standard processing includes: instrumental signature removal – bias, non-linearity, dark, flat, fringe, cross-talk; sky background tracking and homogenisation during image stacking and mosaicing – possible extras may include removal of other 2D systematic effects from imperfect multi-sector operation of detectors; assessing and dealing with image persistence from preceding exposures if necessary; combining frames if part of an observed jitter sequence or tile pattern; producing a consistent internal photometric calibration to put observations on an approximately uniform system; standard catalogue generation including astrometric, photometric, shape and Data Quality Control (DQC) information; final astrometric calibration based on the catalogue with an appropriate World Coordinate System (WCS) placed in all FITS headers; photometric calibration for each generated catalogue augmented by monitoring of suitable pre-selected standard areas covering the entire field-of-view to measure and control systematics; frames and catalogue supplied with provisional calibration information and overall morphological classification embedded in FITS files; propagation of error arrays and use of confidence maps; realistic errors on selected derived parameters; nightly extinction measurements in relevant passbands; pipeline software version control – version used recorded in FITS header; processing history including calibration files recorded in FITS headers.

6.3.2 Science archiving

The concept of the science archive (SA, Hambly et al. 2004 and references therein) is key to the successful exploitation of wide field imaging survey datasets. The SA ingests the products of pipeline processing (instrumentally corrected images, derived source catalogues, and all associated metadata) into a database and then goes on to curate them to produce enhanced database-driven products. In the VDFS science archive, the curation process includes, but is not limited to, the following: individual passband frame association; source association to provide multi-colour, multi-epoch source lists; global photometric calibration; enhanced astrometry including derivation of stellar proper motions; consistent list-driven photometry across sets of frames in the same area; cross-association with external catalogues; and generation of new image products, e.g. stacks, mosaics and difference images etc., all according to prescriptions set up for a given survey programme. Archive curation includes quality control procedures, as required and led by the public survey consortium, and supported by archive team members. All these features are available in the context of a continually updating survey dataset from which periodic releases (as required by the community) can be made. Moreover, end-user interfaces were catered for from the beginning in the VDFS design process, and the philosophy has always been to provide both simple and sophisticated interfaces for the data. The former is achieved via simple point-and-click web forms, while the latter is achieved via exposing the full power of the DBMS back-end to the user. To that end, full access to Structure Query Language and the relational organisation of all data are given to the user.

We have developed a generalised relational model for survey catalogue data in the VDFS. The key features to note are the normalised design with merged multi-waveband catalogue data (the table of most use for scientific queries) being part of a related set of tables that allow the user to track right back to the individual source images if they require to do so; and also that the merged source tables (as derived either from individually analysed images, or consistently across the full passband set available in any one field) are seamless, and present the user with a generally applicable science-ready dataset. Similar relational models describe the organisation of all data in the science archive (image, catalogue, calibration metadata, etc.) - see Hambly et al. (2004) and references therein. The science archive has a high-speed query interface, links to analysis tools such as TopCat, and advanced new VO services such as MySpace. Data products are being successfully ingested into the WFCAM Science Archive (WSA) in Edinburgh, with the EDR in Feb 2006, and the WSA concept was also demonstrated on the SuperCOSMOS Science Archive (SSA). *http://surveys.roe.ac.uk/ssa/*

It will be possible to simultaneously query data tables for the VISTA-GPS and UKIDSS GPS archives in Edinburgh, using the freeform SQL query tool that is already in place to link surveys such as the UKIDSS Large Area Survey and the SLOAN DR2 release.

6.4 Expected data products:

Instrumentally corrected frames (pawprints, tiles etc) along with header descriptors propagated from the instrument and processing steps (science frames and calibration frames)
Statistical confidence maps for each frame
Stacked image data for jittered observations
Derived catalogues (source detections from science frames with standard isophotal parameters, model profile fitted parameters, image classification, etc.)
Data Quality Control database
Database-driven image products (stacks, mosaics, difference images, image cut-outs)
Frame associations yielding a survey field system; seamless, merged, multi-colour, multi-epoch source catalogues with global photometric calibration, proper motions (where appropriate)
Source remeasurement parameters from consistent list-driven photometry across all available bands in any one field
A Proper Motion catalogue for all sources detected at 3 epochs.

6.5 Algorithm development

The UKIDSS GPS team is presently engaged in the ongoing development of source characterisation algorithms to optimise the exploitation of Galactic Plane Surveys, eg. dereddening algorithms, galaxy detection in the plane etc. This work is led by Lucas & Hoare and is ongoing subject to the availability of resources. These algorithms will be made publically available for expoitation of VISTA-GPS data as well as the UKIDSS GPS.

6.6 General schedule of the project:

The VDFS system will make reduced data available to ESO as required by the end of the semester following the observations.

We anticipate delivering 4 public archive releases, which will be quality controlled by the VISTA-GPS team after passing through the VDFS. Quality control is expected to take 3 months, since the UKIDSS experience is that this is essential to ensure a reliable and fully functional public relational database. Each release will represent a significant milestone toward completion of the survey: a DR1 after 1 year of data has been processed, a DR2 after the first ZYJHKs epoch has been completed and processed, and DR3 and DR4 releases when each of the Ks-only epochs have been completed.

At the start of the survey the VDFs process is likely to take 6 months. By the end of the first year of operation we anticipate that reduced data can be released to the PIs for quality control within 2 months of raw data arriving in the UK.

Optional reprocessing of data based on improved knowledge of the instrument would also be considered

References:
Dye S. et al, 2006 in prep, The UKIDSS Early Data Release
Emerson J.P. et al., 2004, “VISTA data flow system: overview”, in Optimizing scientific return for astronomy through information technologies, eds. P.J. Quinn & A. Bridger, Proc. SPIE, vol. 5493, 401
Hambly N.C. et al., 2004, “VISTA data flow system: survey access and curation; the WFCAM science archive”,

in Optimizing scientific return for astronomy through information technologies, eds. P.J. Quinn & A. Bridger, Proc. SPIE, vol. 5493, 423

Irwin M.J. et al., 2004, "VISTA data flow system: pipeline processing from WFCAM and VISTA", in Optimizing scientific return for astronomy through information technologies, eds. P.J. Quinn & A. Bridger, Proc. SPIE, vol. 5493, 411

Lawrence A. et al, 2006 in prep, The UKIRT Infrared Deep Sky Survey

7 Envisaged follow-up: (1 page max)

We plan to carry out follow-up observations of selected sources uncovered by the VISTA GPS. The survey depth is well suited for spectroscopic follow-up with VLT instruments such as FORS and ISAAC. In particular, we foresee FORS spectroscopic follow-up of O and B-type stars that will yield distances to new OB associations and clusters, as well as physical parameters for the stars. FORS observations of galaxies in the zone of avoidance will allow us to determine their distances and amount of dark matter. We also foresee ISAAC spectroscopic follow-up of brown dwarf candidates to estimate their ages, distances and masses. Using NACO and the LGS, high spatial resolution imaging of embedded clusters will be possible, as well as the search for brown dwarf binaries, and the study of the morphology of post-AGB, PPN and PNes.

Infrared multi-object spectrographs, eg. the planned KMOS instrument for the VLT, will be invaluable to provide spectra of representative samples of objects in each young cluster in order to determine the level of background contamination and measure cluster ages. This is needed to determine reliable Luminosity Functions and Mass Functions by locating cluster members on stellar isochrones.

8 Other remarks, if any: (1 page max)

List of CoIs. We generally list one CoI per institution except where two or more people have made a major contribution to this case.

Name	Affiliation
Janet Drew	Imperial College, London
Martin Lopez-Corredoira	Instituto de Astrofisica de Canarias
Paulo Garcia	University of Portugal
Bertrand Goldman	MPIA, Heidelberg
Teresa Giannini	Rome Observatory
Jochen Eisloeffel	Thueringer Landessternwart
Paul Groot	Nijmegen University
Juan Fabregat	Universidad de Valencia
Xavier Delfosse	Observatoire de Grenoble
Thierry Forveille	Canada France Hawaii Telescope
Leonardo Bronfman	University of Chile
Jim Emerson	Queen Mary College, London
Nigel Hambly	Wide Field Astronomy Unit, Royal Observatory Edinburgh
Andy Longmore	Royal Observatory Edinburgh
Nic Walton	Cambridge Astronomical Surveys Unit, Cambridge University
Richard de Grijs	Institute of Astronomy, Cambridge
Melvin Hoare	Leeds University
Anja Schroeder	Leicester University
Tim Naylor	Exeter University
Mike Barlow	University College London
Albert Zijlstra	Manchester University
Glenn White	Open University
Andrew Gosling	Oxford University
Katherine McGowan	Southampton University
Andy Adamson	Joint Astronomy Centre, Hawaii
Reba Bandyopadhyay	University of Florida
Mark Thompson	University of Hertfordshire
Mark Cropper	Mullard Space Science Laboratory
John Lucey	Durham University