



Technical Note

The IWG7 4MOST Galactic Pipeline (4GP): Development Report 2

Document No.: MST-TNO-XXX-YYYYY-GGGG-HHHH

Issue No.: 1.00

DocuShare Document URL: <https://ds-web.aip.de/docushare/dsweb/Services/Document-ZZZZ>

Prepared by

2017-12-15

Dominic Ford



Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 3 of 32

Table of contents

1	Scope	5
2	Applicable Documents (AD)	5
3	Reference Documents (RD).....	5
4	Introduction.....	6
5	The 4MOST Galactic Pipeline (4GP)	7
6	The Cannon.....	7
7	Constructing a training set	8
8	Constructing a test set	9
9	Choice of abundances to fit and censoring spectra.....	13
10	Mapping performance across parameter space	16
11	Performance histograms.....	19
12	Next steps.....	20
	Appendix A : Mapping performance across the HR diagram (LRS)	22
A. 1	T_{eff}	22
A. 2	$\log(g)$	22
A. 3	[Fe/H]	23
A. 4	[Ca/H]	23
A. 5	[Mg/H]	24
A. 6	[Ti/H]	24
A. 7	[Si/H]	25
A. 8	[Na/H]	25
A. 9	[Ni/H]	26
A. 10	[Cr/H]	26
	Appendix B : Mapping performance across the HR diagram (HRS).....	27
B. 1	T_{eff}	27
B. 2	$\log(g)$	27
B. 3	[Fe/H]	28
B. 4	[Ca/H]	28
B. 5	[Mg/H]	29
B. 6	[Ti/H]	29
B. 7	[Si/H]	30
B. 8	[Na/H]	30



Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 4 of 32

B. 9 [Ni/H].....	31
B. 10 [Cr/H].....	31
Appendix C : List of Acronyms	31

List of Figures

Figure 8-1: The distribution of Teff values in the training and test samples.	11
Figure 8-2: The distribution of log(g) values in the training and test samples.	11
Figure 8-3: The distribution of [Fe/H] values in the training and test samples.	12
Figure 8-4: The distribution of stellar parameters in the training set. Stars are colour-coded by metallicity.	12
Figure 8-5: The distribution of stellar parameters in the test set. Stars are colour-coded by metallicity.	13
Figure 9-1: RMS offset in the Cannon's determination of Teff, log(g) and [Fe/H], when fitting 3 or 10 parameters to LRS spectra.	15
Figure 9-2: RMS offset in the Cannon's determination of Teff, log(g) and [Fe/H], when fitting 3 or 10 parameters to HRS spectra.	16
Figure 10-1: The RMS offsets in each parameter for giants and dwarfs in two metallicity bins, based on synthetic LRS observations.	18
Figure 10-2: The RMS offsets in each parameter for giants and dwarfs in two metallicity bins, based on synthetic HRS observations.	19
Figure 11-1: Histograms of the offsets in Teff, log(g) and [Fe/H] as estimated from LRS and HRS spectra with differing SNRs.	20

List of Tables

Table 9-1: The time taken to run the Cannon when fitting 3 and 10 parameters to 4MOST LRS and HRS spectra.	13
---	----



Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP): Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 5 of 32

1 Scope

We present an update on IWG7's development of a spectral analysis pipeline for the 4MOST Galactic Surveys. We describe proof-of-principle tests to investigate whether machine-learning techniques can deliver abundances to the accuracy required by the Surveys, as set out in the IWG7 Management Plan [AD1]. We demonstrate that the algorithm tested can already meet the accuracy requirements for all stellar parameters and many abundances, while others still need further refinement.

The tests presented here use one specific machine-learning technique, the Cannon (Ness et al. 2015 [RD5]; Casey et al. 2016 [RD1]). However, the framework we have developed for performing these tests is highly flexible, and could be used to test other algorithms also. We anticipate that it is quite likely that improved algorithms will have become available by the time 4MOST begins observing.

2 Applicable Documents (AD)

The following applicable documents (AD) of the exact issue shown form a part of this document to the extent described herein. In the event of conflict between the documents referenced herein and the contents of this document, the contents of this document are the superseding requirement.

AD ID	Document Title	Document Number	Issue	Date
[AD1]	IWG7 Management Plan	MST-PLA-PSC-20307-09237-0001	0.07	05.01.2017

3 Reference Documents (RD)

The following reference documents (RD) contain useful information relevant to the subject of the present document.

RD ID	Document Title	Document Number	Issue	Date
[RD1]	Casey, A. et al. 2016 – http://adsabs.harvard.edu/abs/2016arXiv160303040C			
[RD2]	Preliminary Results of IWG7 for radial velocities, atmospheric parameters and chemical abundances (Kordopatis et al. 2016)			
[RD3]	Ford, D. 2017, The IWG7 4MOST Galactic Pipeline (4GP): Development Report (July 2017)			
[RD4]	Ho, A. et al. 2017, ApJ, 841, 40 – http://adsabs.harvard.edu/abs/2017ApJ...841...40H			



Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP): Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 6 of 32

RD ID	Document Title	Document Number	Issue	Date
[RD5]	Ness, M. et al. 2015, ApJ, 808, 16 – http://adsabs.harvard.edu/abs/2015ApJ...808...16N			
[RD6]	Ruchti G. et al. 2016, MNRAS, 461, 2174 – http://adsabs.harvard.edu/abs/2016asc1.soft06001R			
[RD7]	Gustafsson B. et al. 2008, A&A, 486, 951 – http://adsabs.harvard.edu/abs/2008A%26A...486..951G			

4 Introduction

The analysis of observations from large spectroscopic surveys poses a substantial computational challenge. The highly multiplexed observing modes of multi-object spectrographs such as 4MOST means that they can observe thousands of objects in a single pointing. The 4MOST Galactic Pipeline (4GP) will need to be able to process tens of thousands of spectra per day in need to keep up with the flow of data from the telescope.

Whilst this means that large, statistically significant populations of stars can be observed in a very short time, it also means that a highly automated, fast and robust analysis pipeline will be needed. It is infeasible to propose doing manual abundance analysis on more than a tiny proportion of the stars 4MOST will observe.

Machine learning has recently emerged as a technique which scales affordably to such large data volumes. Such techniques rely on a set of training spectra of objects with known properties, which they use to learn how each pixel within the spectrum correlates with each parameter to be derived. The number of training spectra required rises with the number of parameters being fitted, but typically a few thousand are used. In the tests presented here, we use 3,310 training objects.

Once trained, machine-learning algorithms can rapidly estimate the properties of any other object based on its spectrum. We show here that ten parameters can be extracted from a continuum-normalised spectrum in less than half a second using a single desktop computer.

In their simplest form, these techniques are *data driven*, since everything they know about spectra is empirically derived from the training set. They are not pre-programmed with any prior knowledge of where important lines lie, or which lines are produced by which elements. In Section 9, however, we will demonstrate that their performance may be improved under certain conditions if we supply additional information about where the most important lines are.



Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 7 of 32

This report presents an update on IWG7’s development of a spectral analysis pipeline for the 4MOST Galactic Surveys. Previously, Kordopatis et al. (2016) [RD2] and Ford (2017) [RD3] reported tests of whether the Cannon could reproduce the stellar parameters and abundances of a sample of synthetic giant stars whose properties were taken from the APOGEE source catalogue. In this report we extend the analysis to cover simultaneous fitting of both dwarf and giant stars with a wide range of spectral types and metallicities.

5 The 4MOST Galactic Pipeline (4GP)

The framework we used to perform the tests described in this report is the 4MOST Galactic Pipeline (4GP). This is a collection of Python modules which use a common data format to store and manipulate spectra and their associated metadata. It makes it easy to pass spectra between a range of spectral synthesis and processing tools including Turbospectrum, the 4MOST Facility Simulator (4FS) and the Cannon, without the need for manual data format conversion. It includes the ability to store spectra in libraries and search for them by arbitrary metadata constraints, making it easy to create new tests on subgroups of stars filtered from larger samples.

The framework is available in two repositories on GitHub, and includes step-by-step installation instructions. The first repository contains the Python modules which provide programmatic interfaces for creating and manipulating libraries of spectra, including wrappers for passing them to various analysis tools:

<https://github.com/dcf21/4most-4gp>

The second repository contains python scripts which utilise these modules to synthesise the spectra and perform the tests described in this report:

<https://github.com/dcf21/4most-4gp-scripts>

6 The Cannon

The Cannon analyses spectra by forming an internal model of how the continuum-normalised flux in each pixel i within a spectrum can be expressed as a polynomial function of the N parameters being fitted. For example, the expression for the flux in pixel i might include the terms

$$f_i = a_0 + a_1 T_{\text{eff}} + a_2 \log(g) + a_3 [\text{Fe}/\text{H}] + \dots + a_n T_{\text{eff}} \log(g) + \dots + \text{noise} \quad [1]$$

The exact set of included terms is flexible, but previous works (e.g. Ness et al. 2015 [RD4]; Casey et al. 2016 [RD1]) have included a set of N terms which are linear in each parameter, plus $N(N+1)/2$ quadratic terms which contain every possible product of pairs of parameters. The coefficients a_j are unique to every pixel in the spectrum, which are described by different polynomials.



Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 8 of 32

This approach assumes that all the spectra being processed are sampled at a common set of wavelengths, such that each pixel is sensitive to the same features in each. In practice, this means that the spectra of stars with non-zero radial velocities must be resampled after being shifted into the rest frame of the star.

The noise term is introduced to describe any additional scatter in the value of each pixel which cannot be fitted by any of the terms in the polynomial. It is modelled as a Gaussian distribution, whose width is treated as a free parameter. It allows a probability to be ascribed for any given flux value being observed in the spectrum of a star with a given set of stellar parameters and abundances.

When the machine-learning algorithm is being trained, the coefficients a_j and the widths of the noise distributions are free parameters, while the stellar parameters and abundances associated with each spectrum are assumed to be perfectly known. The flux of each spectrum in each pixel is also known, subject to some uncertainty. The best-fitting values for the coefficients are derived by a least-squares maximum-likelihood approach.

The noise distributions act as a means to separately weight each pixel. Spectral features produced by elements which the Cannon is not trying to fit will vary in a way which is not described by any of the terms in its internal model, and so they will be assigned a high scatter.

Once the algorithm is trained, parameters can be inferred from other spectra by inverting the fitting process. The polynomials describing the fluxes at each pixel are treated as a set of simultaneous equations which can be solved to derive the best-fitting set of parameter values. The pixels are treated as independent random variables, and the probabilities that a given set of parameters might produce each of the pixel fluxes are combined to derive an overall likelihood for the observed spectrum. The most likely set of parameter values is derived by maximising this likelihood.

7 Constructing a training set

Until 4MOST is operational, we do not have any observed spectra which match 4MOST's wavelength span and resolution. Consequently, to run tests on the wavelength bands in which 4MOST will observe, we are restricted to using synthetic spectra. We adjust these to account for 4MOST's wavelength-dependent sensitivity using the 4MOST Facility Simulator (4FS).

Spectra are produced for both 4MOST LRS and HRS and tested separately. In both cases, we stitch the three arms of the 4MOST spectrographs into a single spectrum for combined analysis by the Cannon.

We use 4FS to inject a small amount of noise into each spectrum, such that its median SNR/pixel is 250 in the window 6180 – 6680 Å. This continuum window is selected because it is within the wavelength range of both LRS and HRS.

Our decision to train the Cannon on slightly noisy rather than clean spectra was suggested by Andy Casey. In clean spectra, very weak lines become visible which could never be observed



Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 9 of 32

in practice. The Cannon may learn to use these to infer abundances, even though they are never seen in real spectra.

To synthesise each spectrum, we use Turbospectrum with MARCS model atmospheres (Gustafsson et al. 2008 [RD7]) and line lists provided by Bengt Edvardsson (private communication). Further details of how these spectra are processed through 4FS can be found in our previous development report (Ford 2017 [RD3]).

A good training set needs to sample all the regions of parameter space in which the Cannon is going to be tested, to prevent extrapolation of its predictive model into regions where it is poorly constrained. To create such a sample, with a physically realistic distribution of properties and broad range of spectral types and metallicities, we take stellar parameters and abundances from stars observed with UVES in the Gaia-ESO Data Release 5 (DR5) catalogue, subject to the following criteria:

- $\text{SNR} > 20$
- UVES target
- Uncertainty in $T_{\text{eff}} < 100 \text{ K}$
- Uncertainty in $[\text{Fe}/\text{H}] < 0.15 \text{ dex}$
- Uncertainty in $\log(g) < 0.2 \text{ dex}$
- Radial-velocity uncertainty $< 10 \text{ km/s}$

These criteria leave a sample of 3,310 stars. In practice, errors in the Gaia-ESO parameter determination do not propagate into our analysis, since we do not use the observed spectra at all; we only use their parameters to synthesise our own physically realistic sample of stars.

Where abundances are null in the catalogue, we scale solar abundances with $[\text{Fe}/\text{H}]$ to provide an abundance for Turbospectrum to use. However, in our library of synthetic spectra, we flag these abundances and exclude them from any tests where that particular abundance is being fitted. This prevents the Cannon from learning spurious correlations between these elements and iron, which might lead it to use Fe lines as a proxy for them.

Figure 8-1, Figure 8-2, Figure 8-3 and Figure 8-4 show the distribution of stellar parameters in the training sample.

8 Constructing a test set

Several criteria define a good sample of stars with which to comprehensively test the Cannon’s performance. It should be confined to the same regions of parameter space which are covered by the training set, but should also include statistically significant numbers of stars in every region of parameter space of interest to the 4MOST surveys.

To create such a sample, we divide parameter space up into a number of regions:

- Turn-off dwarf stars ($T_{\text{eff}} > 6000$; $\log(g) > 3$)
- Dwarf stars ($T_{\text{eff}} < 6000$; $\log(g) > 3$)
- Giant stars ($\log(g) < 3$)



Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 10 of 32

We further divide parameter space into four metallicity regimes, as follows:

- $[\text{Fe}/\text{H}] > 0$
- $-0.5 < [\text{Fe}/\text{H}] < 0$
- $-1 < [\text{Fe}/\text{H}] < -0.5$
- $[\text{Fe}/\text{H}] < -1$

Within each of the 12 regions defined above, we synthesise 400 test stars to give a total of 4,800 test spectra. To generate physically plausible stellar parameters and abundances for each star, we select a random star from the training set as a seed for each test star, and apply perturbations to each parameter. These perturbations follow a Gaussian distribution with a standard deviation of 100K in T_{eff} and 0.1 dex in $\log(g)$ and all abundances. This means that the test stars differ from the training stars, but never deviate far from regions of parameter space in which the Cannon has seen at least one training spectrum.

One of the goals of our tests is to determine how the Cannon performs at high noise levels, and so we use 4FS to produce nine differently-noisy versions of each spectrum at the following values of SNR/pixel, defined as previously in the window 6180 – 6680 Å:

10, 20, 50, 80, 100, 130, 180, 250, 5000

Figure 8-1, Figure 8-2, Figure 8-3 and Figure 8-5 show the distribution of stellar parameters in the test sample. Note that the test set, by construction, contains a much higher number of low-metallicity stars than the training set. This is deliberate as we need a sizeable number of stars to test the Cannon’s performance in this regime.

Note also that in Figure 8-1 the temperature distribution of the dwarfs in the test set has an anomalous peak at 6000 K. This artefact is produced by the definition of our test sample, which artificially enhances the number of turn-off stars in the sample. We intend to correct this in future tests, but are confident it does not significantly influence any of the results presented in this report.

Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 11 of 32

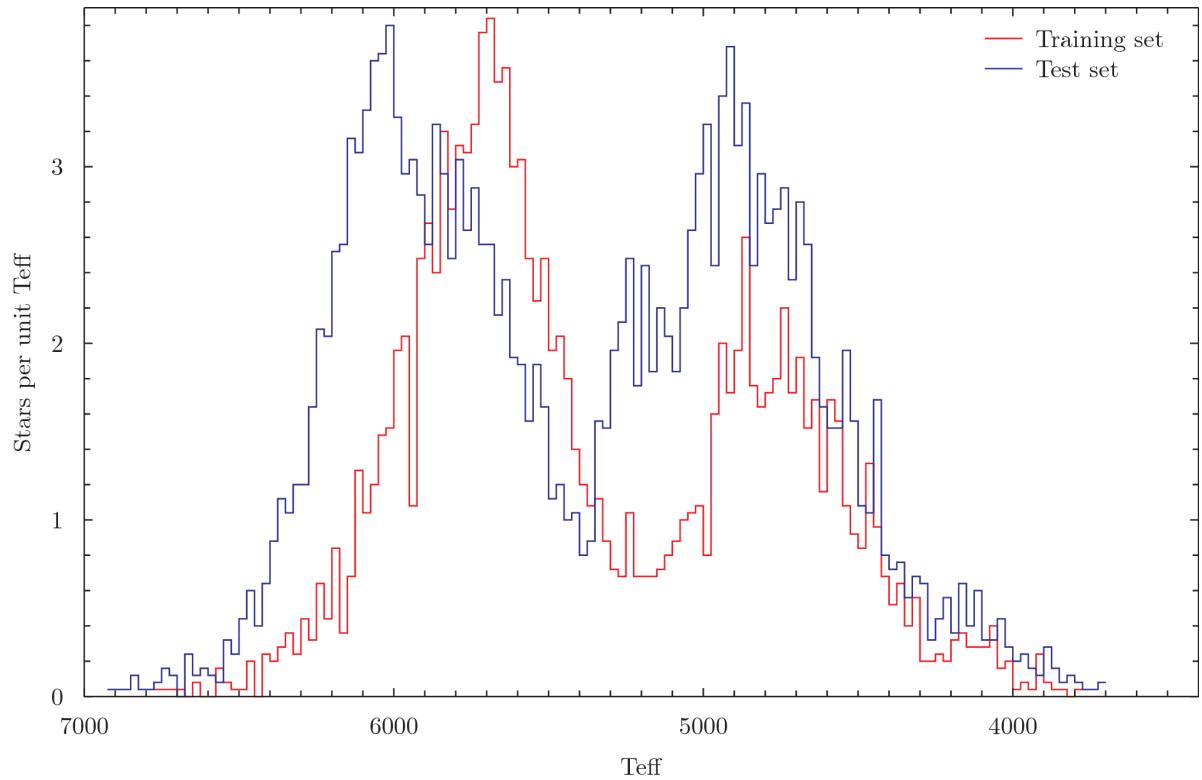


Figure 8-1: The distribution of T_{eff} values in the training and test samples.

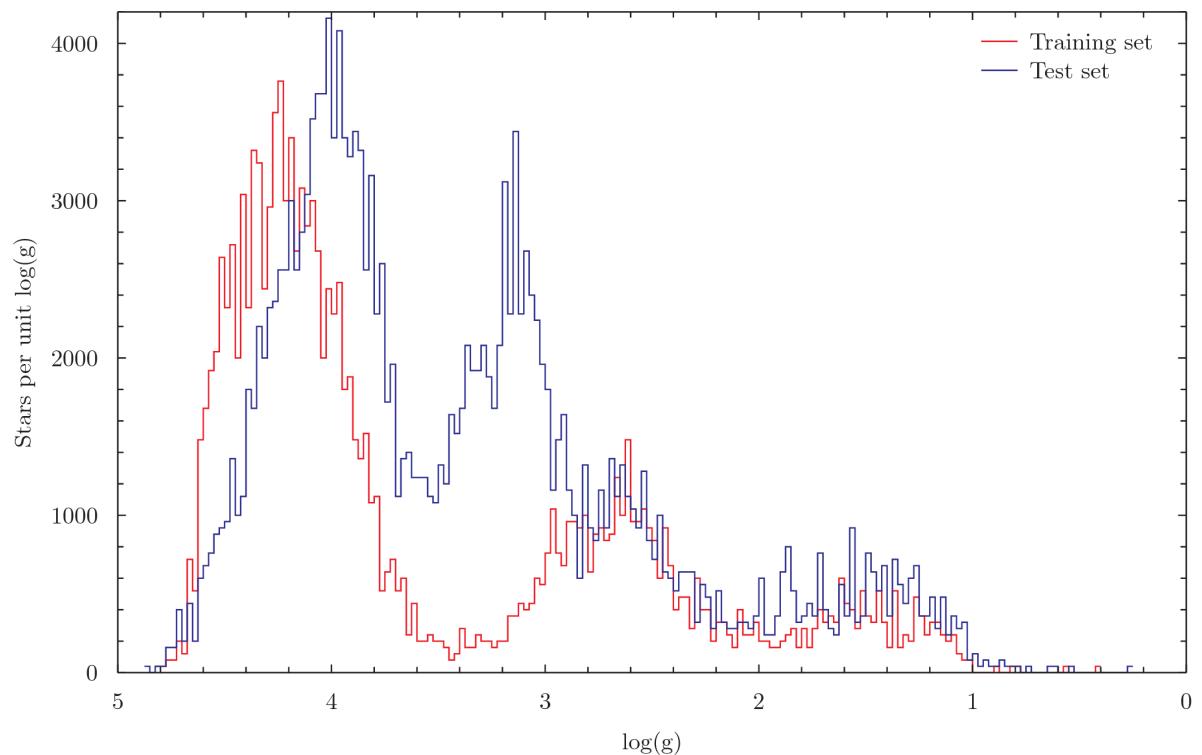


Figure 8-2: The distribution of $\log(g)$ values in the training and test samples.

Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 12 of 32

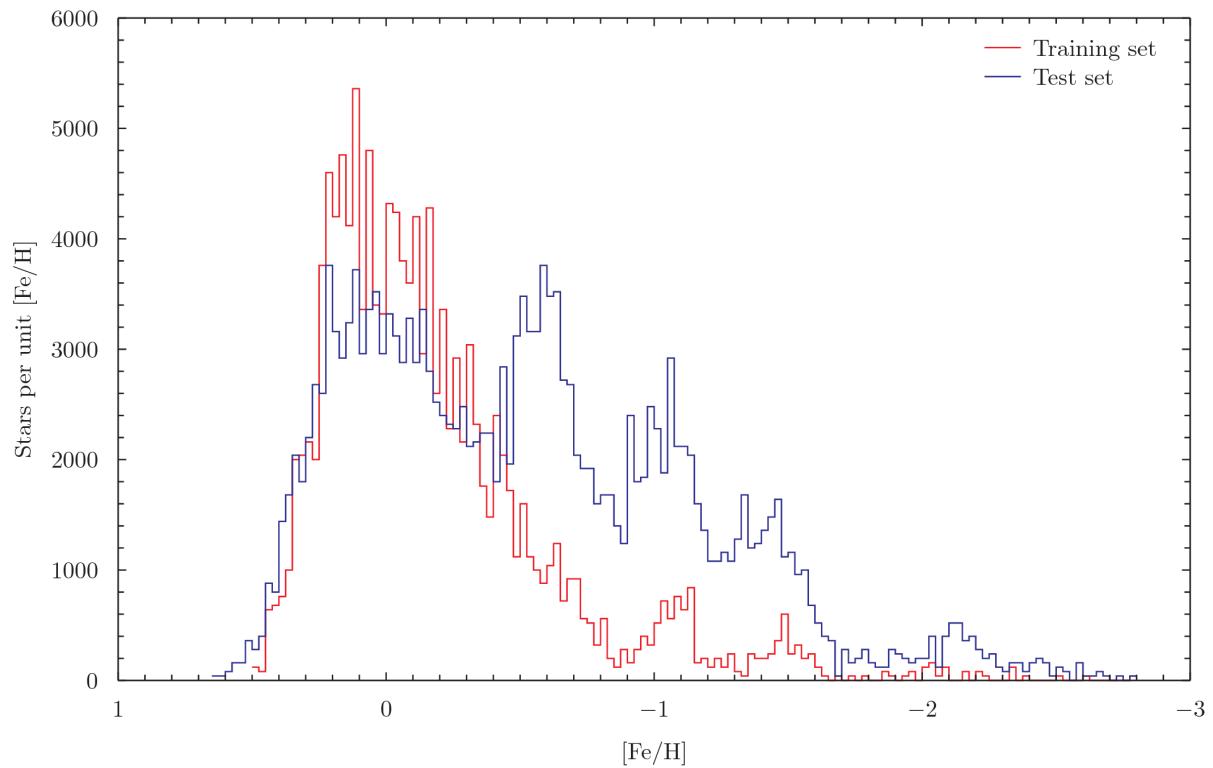


Figure 8-3: The distribution of [Fe/H] values in the training and test samples.

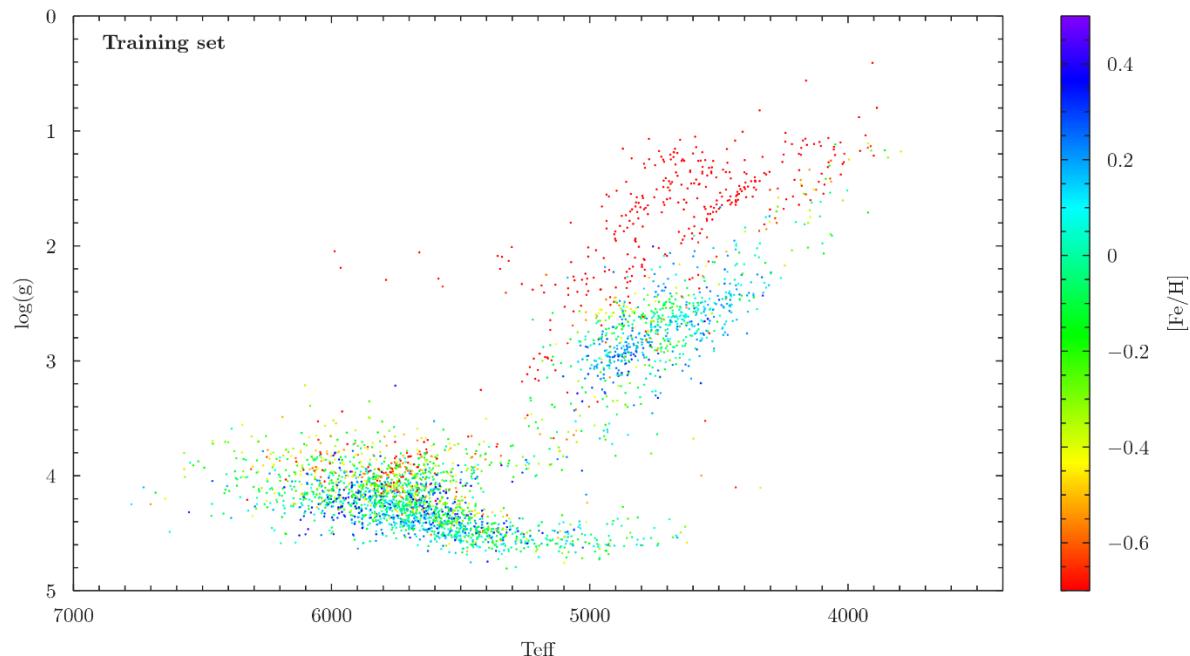


Figure 8-4: The distribution of stellar parameters in the training set. Stars are colour-coded by metallicity.

Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 13 of 32

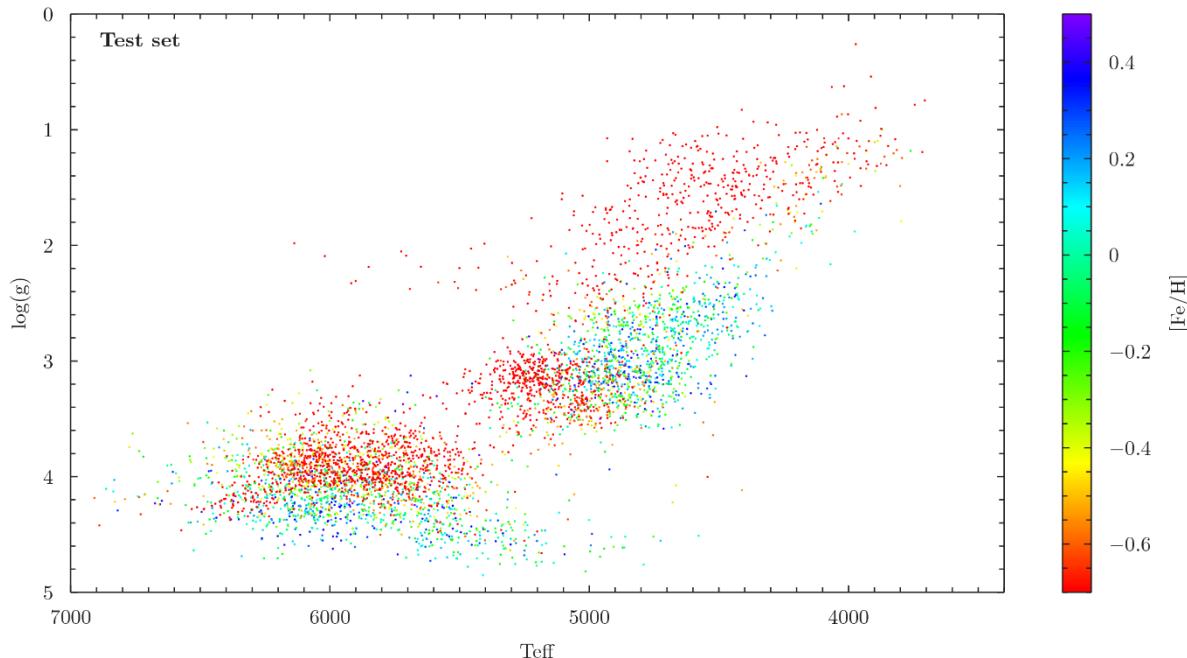


Figure 8-5: The distribution of stellar parameters in the test set. Stars are colour-coded by metallicity.

9 Choice of abundances to fit and censoring spectra

The number and selection of parameters which the Cannon tries to fit makes a significant impact, both on the amount of time it takes to run, and the quality of the results. Table 9-1 below shows the speed of the Cannon running on a quad-core Intel Core i7-6700 desktop. In all cases, the training set contains 3,310 stars, and the test set 4,800 stars. In the first case, the Cannon is only fitting three parameters: T_{eff} , $\log(g)$ and $[\text{Fe}/\text{H}]$. In the latter case, the Cannon fits ten parameters in total: T_{eff} , $\log(g)$, $[\text{Fe}/\text{H}]$, $[\text{Ca}/\text{H}]$, $[\text{Mg}/\text{H}]$, $[\text{Ti}/\text{H}]$, $[\text{Si}/\text{H}]$, $[\text{Na}/\text{H}]$, $[\text{Ni}/\text{H}]$, $[\text{Cr}/\text{H}]$.

Table 9-1: The time taken to run the Cannon when fitting 3 and 10 parameters to 4MOST LRS and HRS spectra.

	Training time [sec]	Test time [sec / star]
3 parameters (LRS)	99	0.0776
3 parameters (HRS)	216	0.0803
10 parameters (LRS)	1099	0.422
10 parameters (HRS)	1326	0.453

Figure 9-1 and Figure 9-2 show the root-mean-square offsets in the Cannon's determination of the parameters of each of the test stars, as a function of noise level for LRS and HRS respectively. Here, we express SNR per Å rather than per pixel, for consistency with previous reports. Note that this creates the impression that HRS needs a higher SNR than LRS to



Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 14 of 32

achieve comparable results, since HRS has approximately 8 pixels / Å in the measurement window 6180 – 6680 Å, while LRS has approximately 3 pixels / Å.

In the uncensored cases, the Cannon is allowed to see the entirety of the wavelength range observable by 4MOST. In the censored cases, we use the list of important lines compiled by Ruchti et al. 2016 [RD6] to create a mask, and only allow the Cannon to see the pixels in windows extending 1 Å to either side of these lines.

Larger windows are created for H-alpha, between 6470 and 6650 Å, for the Mg B line between 5100 and 5250 Å, and for the Ca I lines between 6152 and 6172 Å. The latter two regions are important indicators for $\log(g)$. This censoring scheme reduces the risk of the Cannon learning to use weak lines which may be undetected in noisy spectra.

When training for any given element, we allow the Cannon to see only the H-alpha and H-beta lines, plus the lines of elements which are being fitted.

Motivation for such a masking scheme can be found in Ho et al. 2017 [RD4] among other studies, who found that when the Cannon was used to fit LAMOST spectra, it was learning to use diffuse interstellar bands (DIBs) as a proxy for alpha-element abundance. This correlation arises because the alpha-rich thick disk of the Milky Way is on average more distant than the thin disk.

Even though DIBs are not present in our synthetic spectra, it is likely that similar constraints will necessitate some censoring in the analysis of real 4MOST observations. For this reason, we include censoring in our tests.

In Figure 9-1 and Figure 9-2, we see that 10-parameter fits invariably produce significantly better fits than 3-parameter fits. This implies that polynomials formed of only three variables do not have adequate flexibility to describe the way in which spectra vary across parameter space. This might be the case, for example, if many of the lines being used are blended with weaker lines from other elements.

We might expect the 3-parameter fits to $\log(g)$ to be particularly bad, since Ruchti et al. (2016 [RD6]) report that surface gravity determination requires the Mg B triplet at 5100 – 5250 Å or the Ca I line at 6152 – 6172 Å, together with independent lines with which to measure the Mg or Ca abundance. Our three-parameter fits have no flexibility to disentangle $\log(g)$ from either of these abundances.

Contrary to any fears that censoring might reduce the amount of data available to the Cannon, censoring significantly improves our results from HRS spectra, especially at low SNR. For LRS, censoring makes our $\log(g)$ estimates worse, but improves T_{eff} . The changes in all parameters are substantially smaller than the target accuracies with which we are trying to extract them.

In the remainder of this report, we restrict ourselves to presenting results from censored 10-parameter fits.

Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 15 of 32

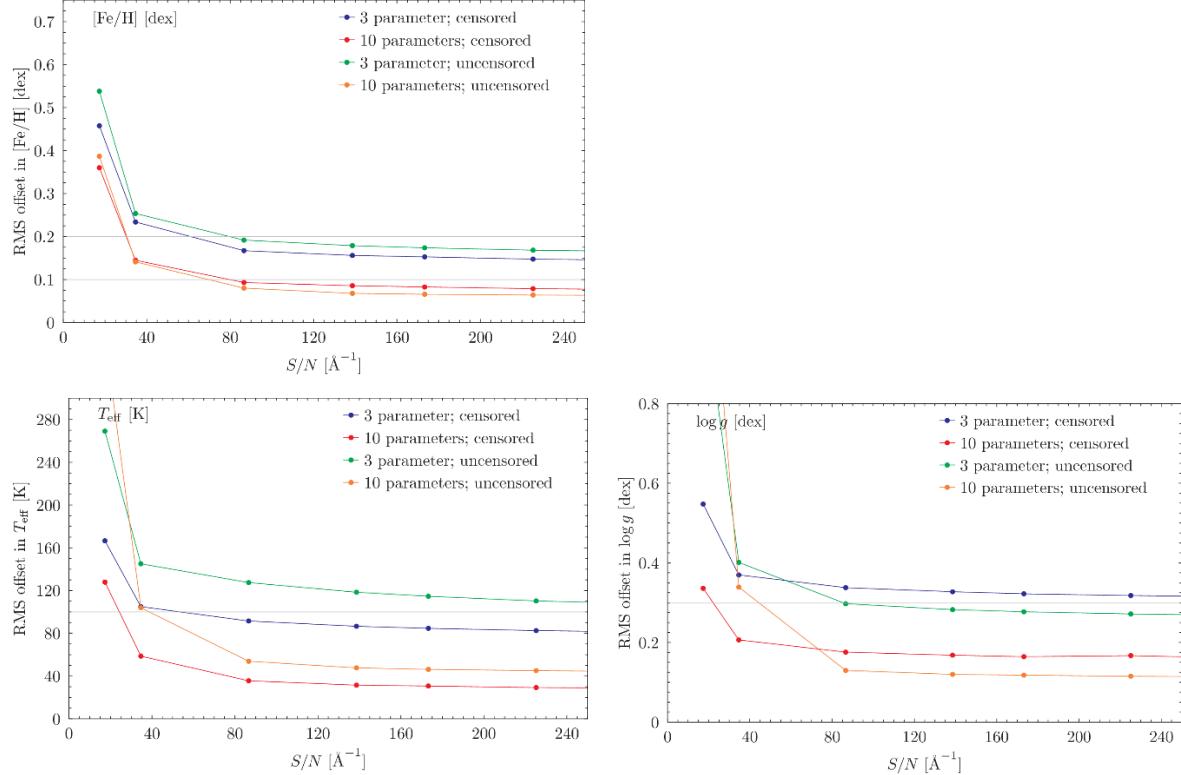


Figure 9-1: RMS offset in the Cannon's determination of T_{eff} , $\log(g)$ and $[\text{Fe}/\text{H}]$, when fitting 3 or 10 parameters to LRS spectra.

Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 16 of 32

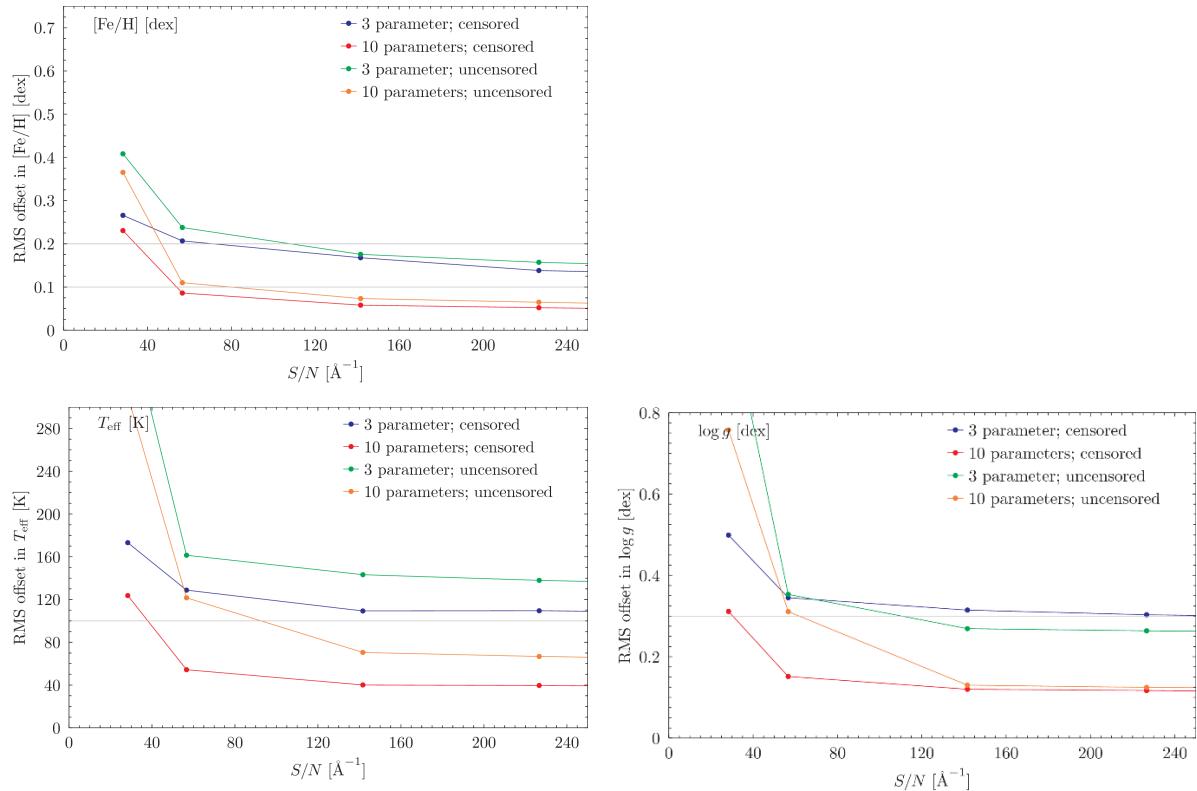


Figure 9-2: RMS offset in the Cannon's determination of T_{eff} , $\log(g)$ and $[\text{Fe}/\text{H}]$, when fitting 3 or 10 parameters to HRS spectra.

10 Mapping performance across parameter space

In this section we divide the test set into dwarfs and giants along the line $\log(g) = 3.25$, and into two metallicity bins along the line $[\text{Fe}/\text{H}] = -1$. For each of these four samples, we plot the RMS offset in each stellar parameter and abundance as a function of SNR, as in the previous section.

Figure 10-1 and Figure 10-2 show the results for 4MOST LRS and HRS respectively. For each stellar parameter and abundance, horizontal lines mark the target accuracies listed in the IWG7 Management Plan [AD1]. These are:

- 100 K accuracy in T_{eff}
- 0.3 dex accuracy in $\log(g)$
- 0.1 dex accuracy in all elemental abundances

For all elemental abundances, we also include a second horizontal line marking a weaker target accuracy of 0.2 dex.



Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 17 of 32

For stars with metallicities $[\text{Fe}/\text{H}] > -1$, we meet or come close to meeting the target accuracy of 0.1 dex in all the elements tested. For low-metallicity stars, we achieve much lower accuracy, however.

In Appendices A and B we further analyse the Cannon’s performance as a function of T_{eff} and $\log(g)$. We plot the test stars in a $T_{\text{eff}}\text{-}\log(g)$ plane, and colour code the lowest SNR at which the Cannon successfully extracts the parameters of each test star. The values plotted on the horizontal and vertical axes are the “true” stellar parameters used to synthesise the spectra, while the colours indicate how well the Cannon performs for each star.

Each stellar parameter and abundance is plotted separately, since some parameters – e.g. $\log(g)$ – are comfortably extracted to within the target precision at an $\text{SNR}/\text{\AA}$ of 40, while other more difficult abundances do not reach the target precision of 0.1 dex even at the highest SNR tested.

Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP): Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 18 of 32

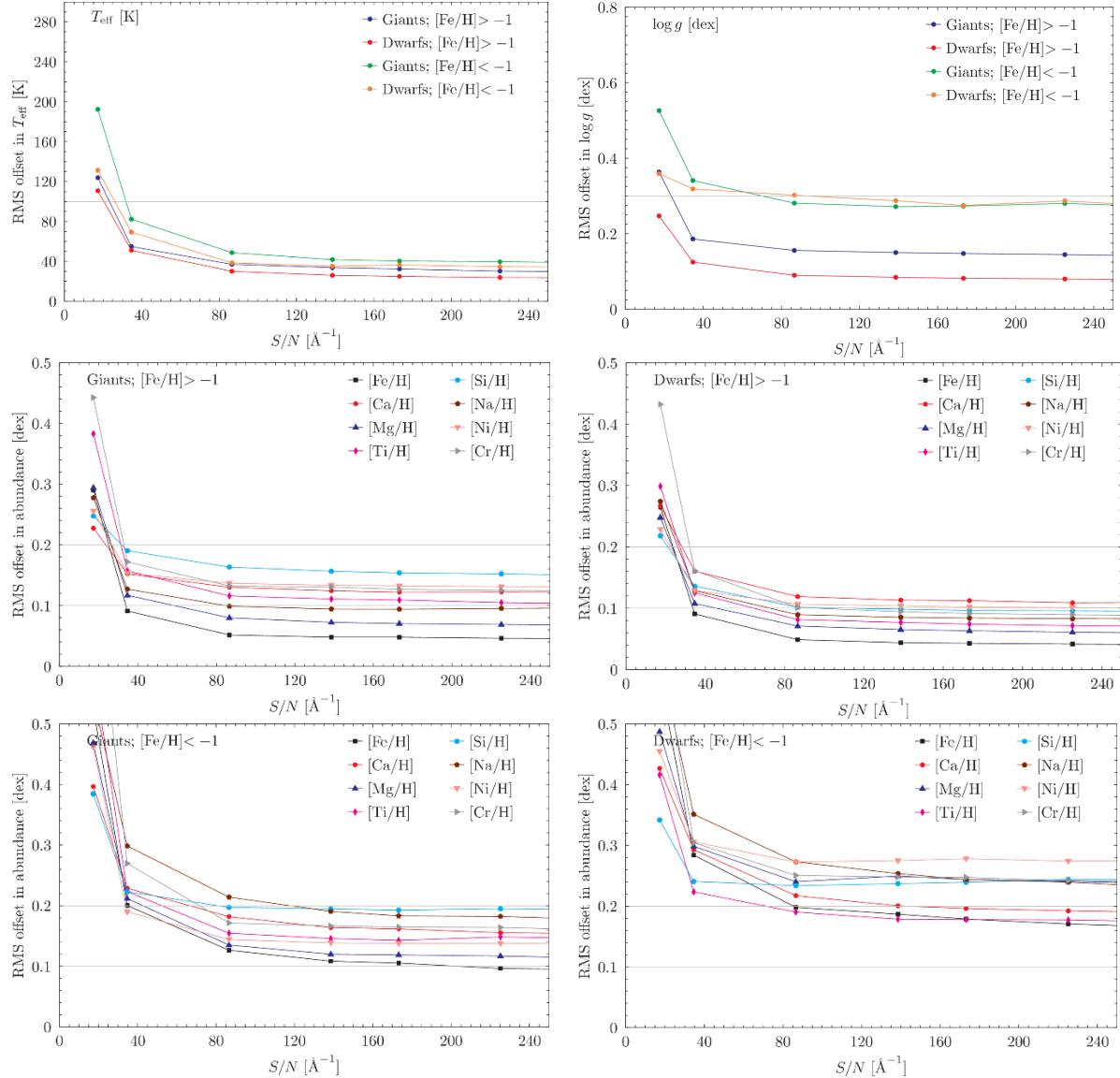


Figure 10-1: The RMS offsets in each parameter for giants and dwarfs in two metallicity bins, based on synthetic LRS observations.



Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP): Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 19 of 32

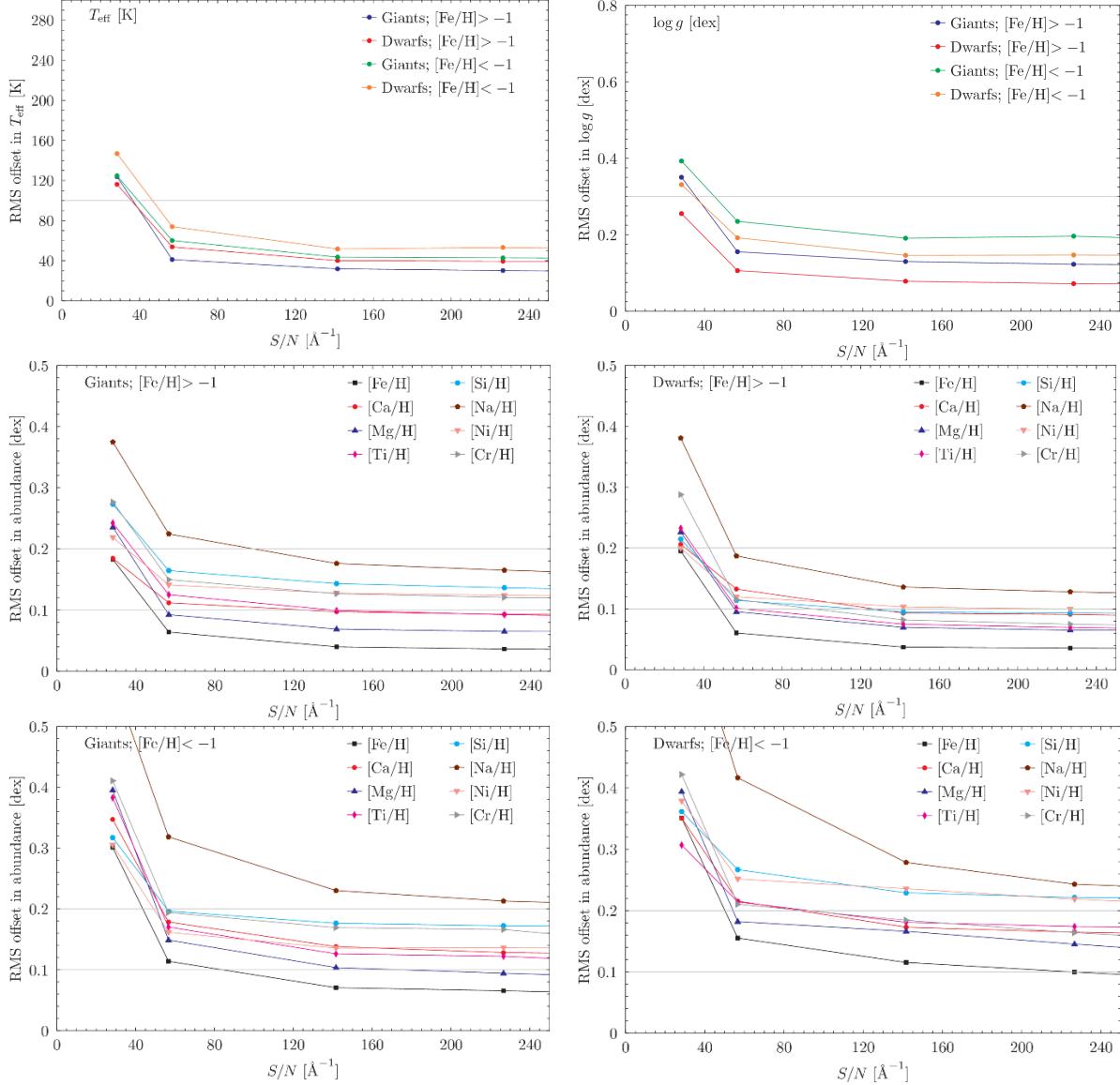


Figure 10-2: The RMS offsets in each parameter for giants and dwarfs in two metallicity bins, based on synthetic HRS observations.

11 Performance histograms

The plots in the previous sections have shown the RMS offsets in each of the stellar parameters and abundances that we have tried to extract using the Cannon. This is not sensitive, however, to any systematic biases which may be present in the Cannon's output when it is operated on noisy data.

In Figure 11-1 we show histograms of the absolute offsets in T_{eff} , $\log(g)$ and $[\text{Fe}/\text{H}]$ at each noise level. We define each offset to be the Cannon-derived value minus the “true” value used to synthesise the spectrum.

Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP): Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 20 of 32

For both LRS and HRS, systematic biases are present in [Fe/H] at SNR/pixel below 50, though these become negligible at higher SNRs. Although this behaviour is within the accuracy requirements set out in the IWG7 Management Plan, in practice 4MOST will undoubtedly produce many spectra at low SNRs, and we would like to understand the origin of these biases better.

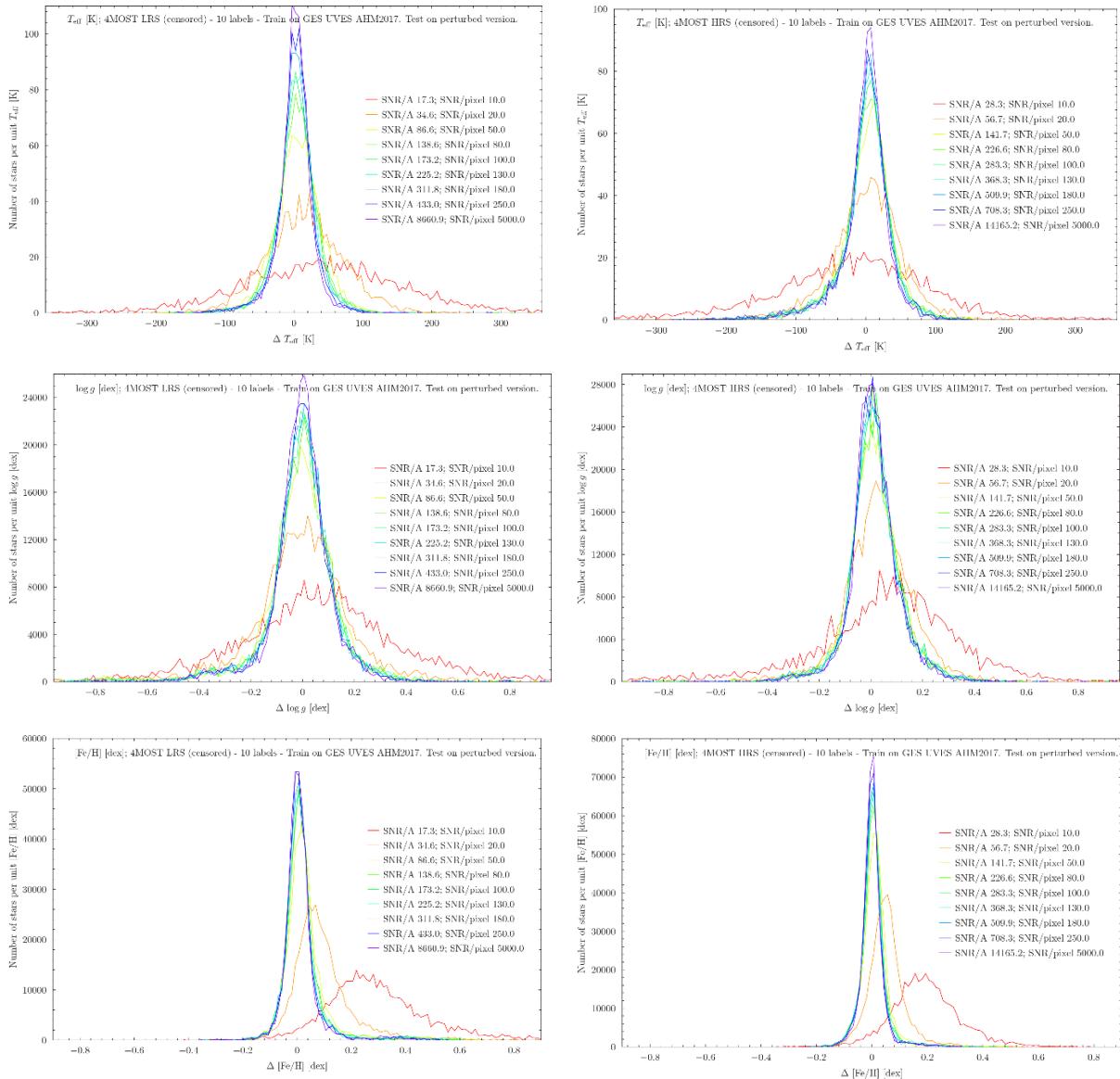


Figure 11-1: Histograms of the offsets in T_{eff} , $\log(g)$ and [Fe/H] as estimated from LRS and HRS spectra with differing SNRs.

12 Next steps

The tests presented in this document have shown the machine learning techniques are a plausible approach to analysing 4MOST observations. However, our tests have been working



Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 21 of 32

on a highly idealised synthetic data. In coming months, we plan to add several further effects into our testing framework:

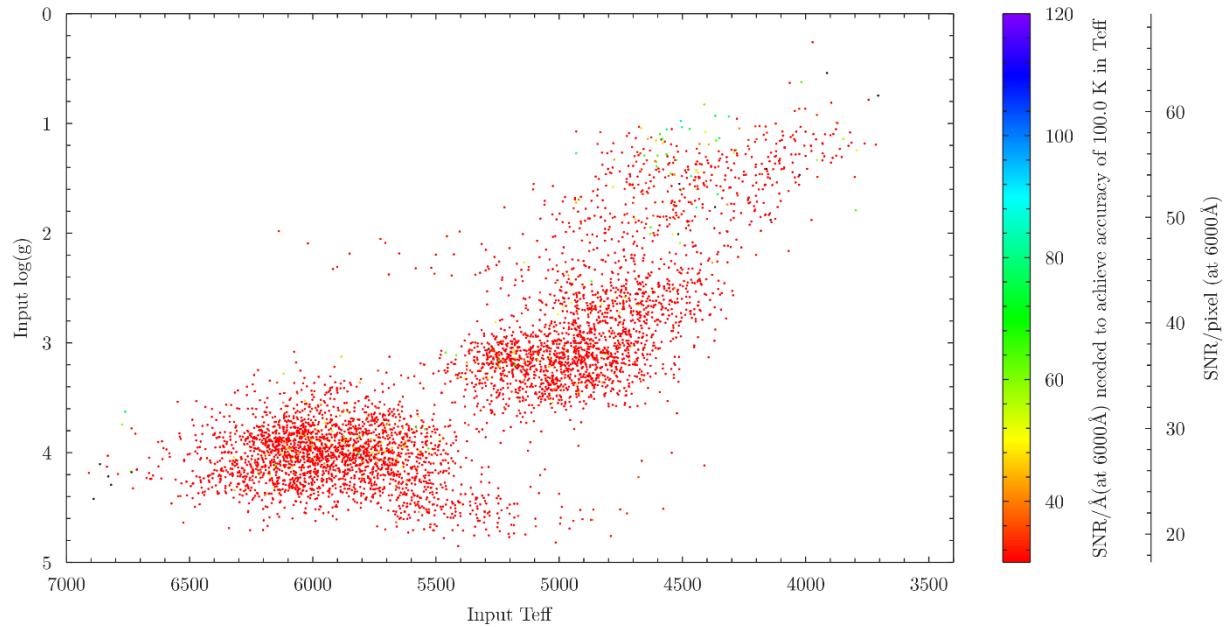
- **Continuum normalisation.** The Cannon requires spectra to be accurately continuum normalised, but the spectra we receive from DMS will not be normalised. Previous surveys, including APOGEE and RAVE have used the Cannon itself to assist in normalisation, by using its internal model to identify continuum pixels through which a smooth function may be fitted. Any inaccuracy in this process may severely impair the Cannon’s performance, and so we need to test whether a similar scheme can be used for 4MOST spectra.
- **Reddening** has the effect of increasing the noise level towards the blue end of continuum normalised spectra. We plan to test whether the Cannon can tolerate working on a mixture of spectra with varying degrees of reddening.
- **Additional elements** need to be added into our tests – including for example oxygen, lithium, barium and europium.
- **Radial velocities (RVs).** In addition to extracting stellar parameters and abundances, the IWG7 pipeline will also need to extract RVs to a target precision of 1-2 km/s for LRS. We need to ensure that the process of resampling spectra onto a common raster of wavelengths in the rest frame of the star (a requirement for using the Cannon; see Section 6) does not impair the Cannon’s performance.

Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 22 of 32

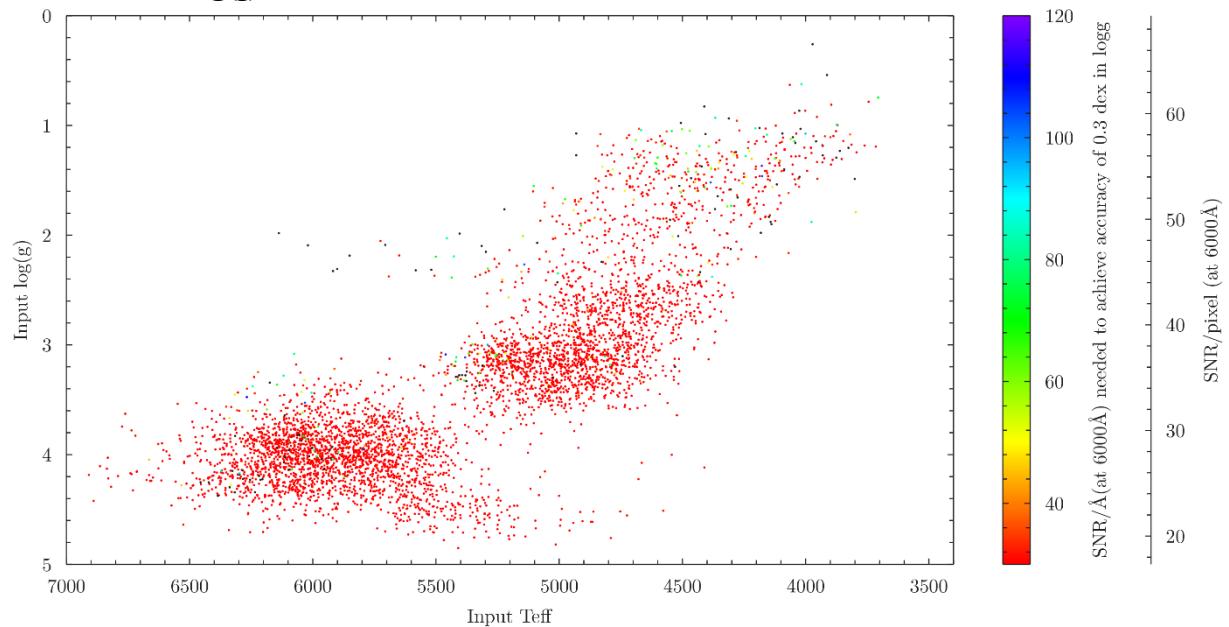
Appendix A: Mapping performance across the HR diagram (LRS)

See Section 10 for more information.

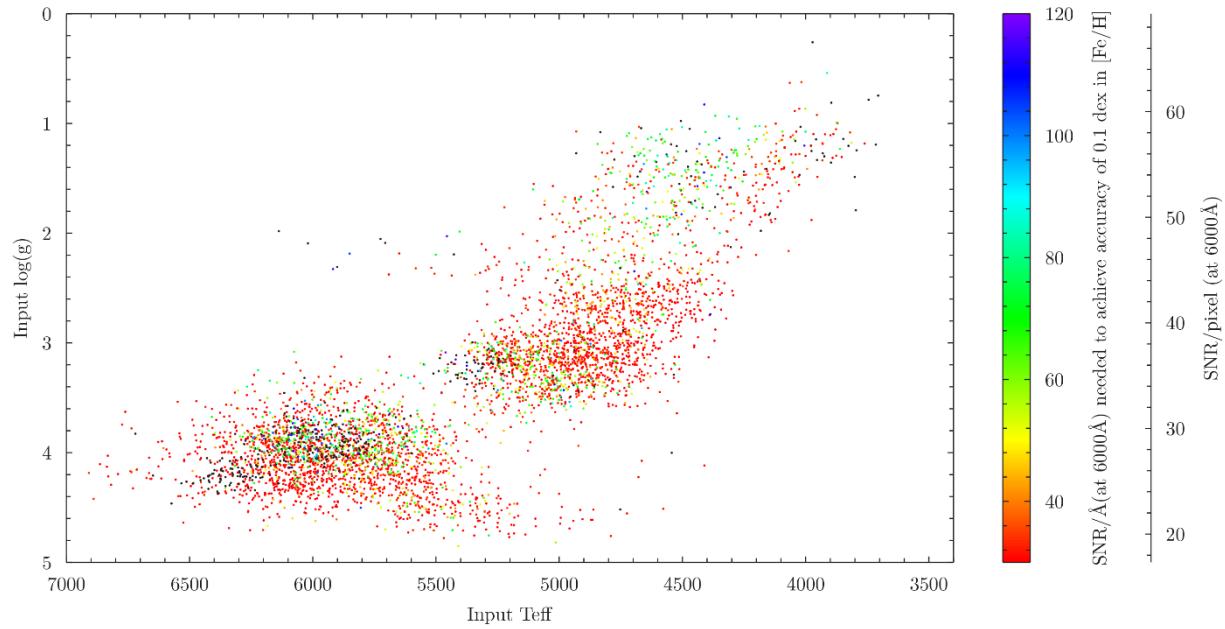
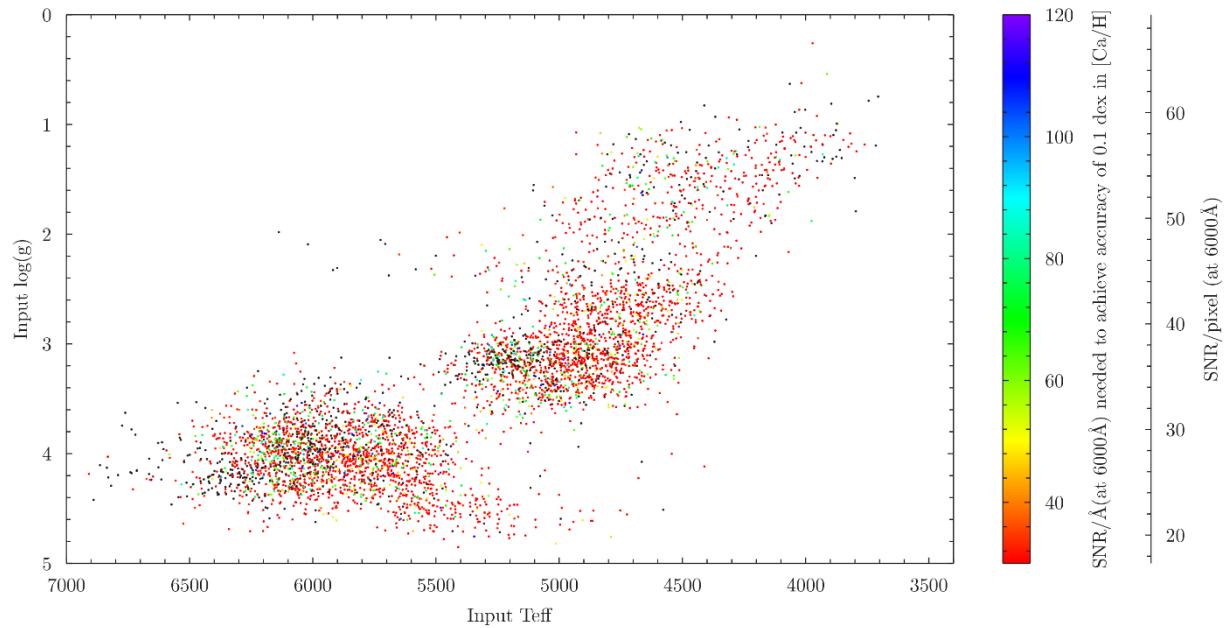
A. 1 T_{eff}



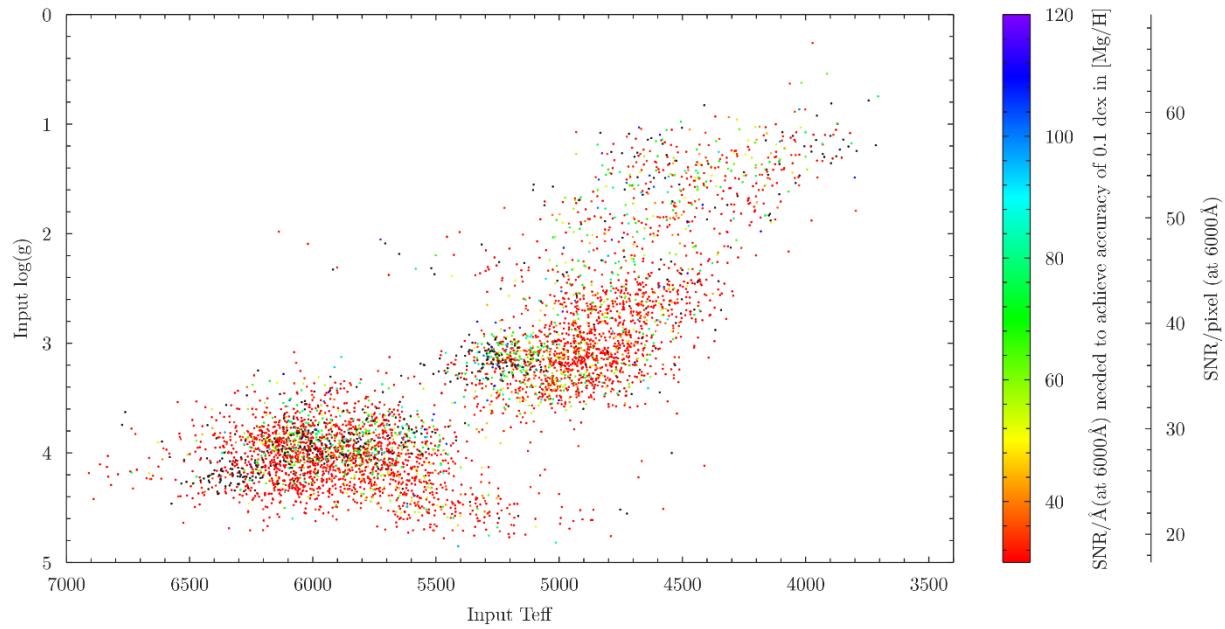
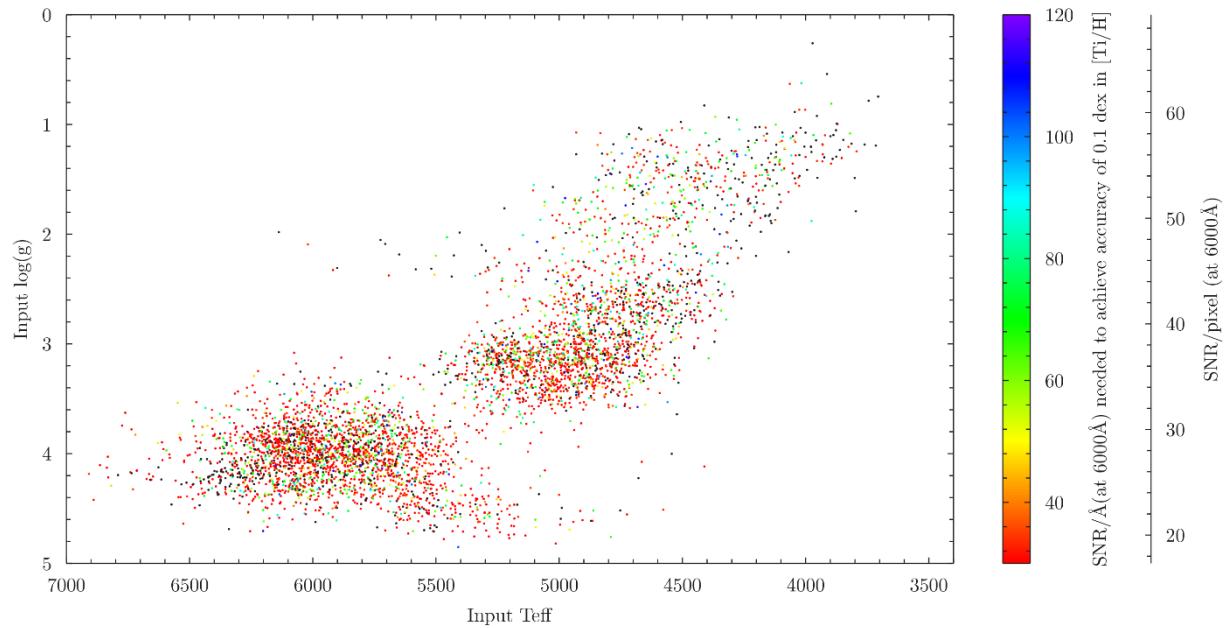
A. 2 $\log(g)$



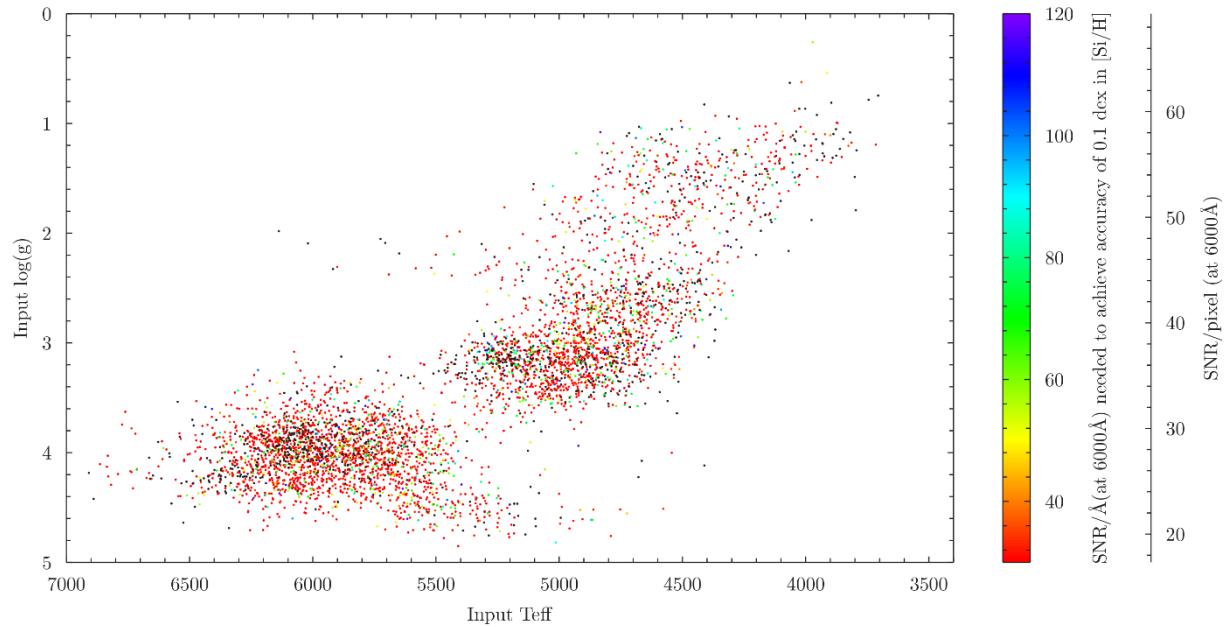
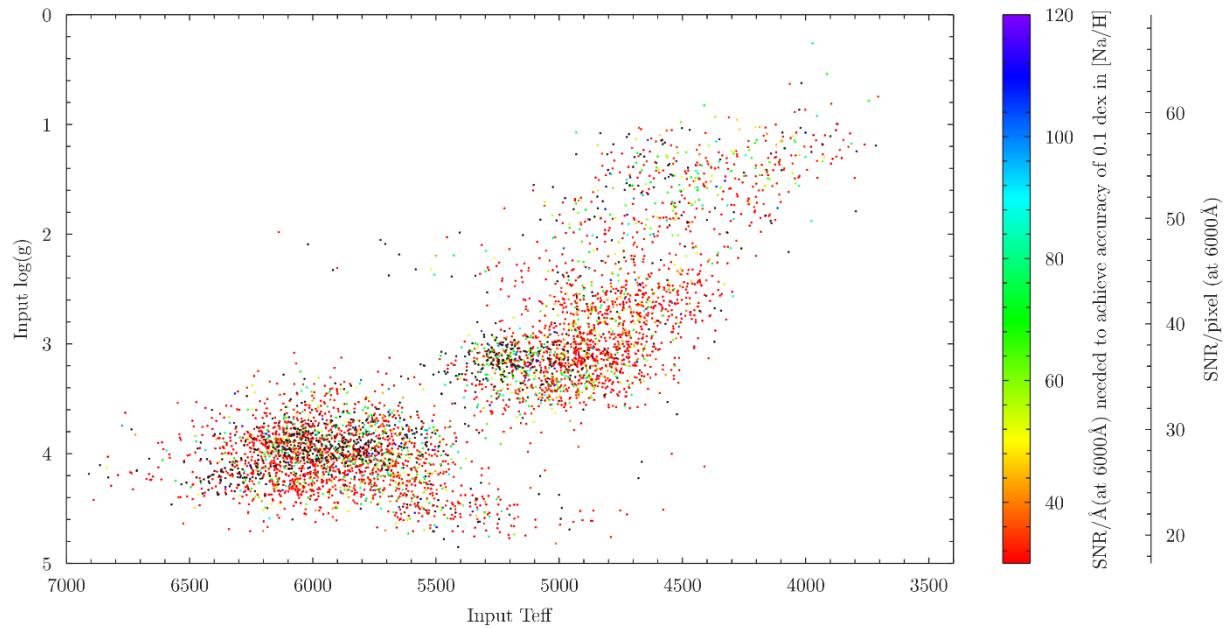
Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 23 of 32

A. 3
[Fe/H]

A. 4
[Ca/H]


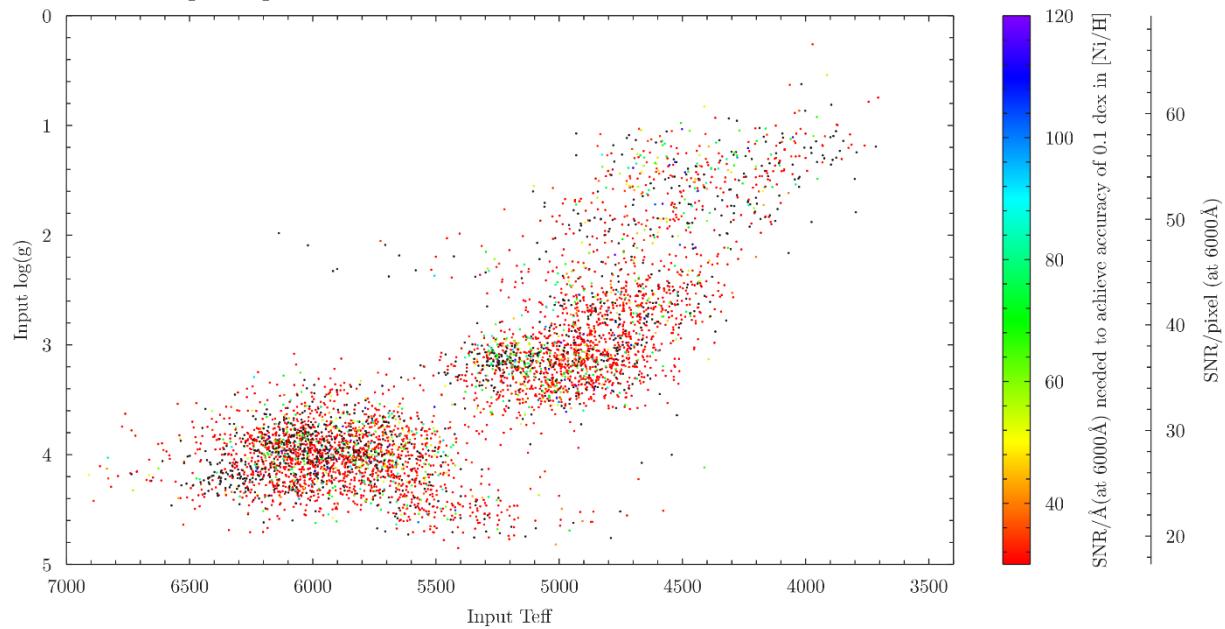
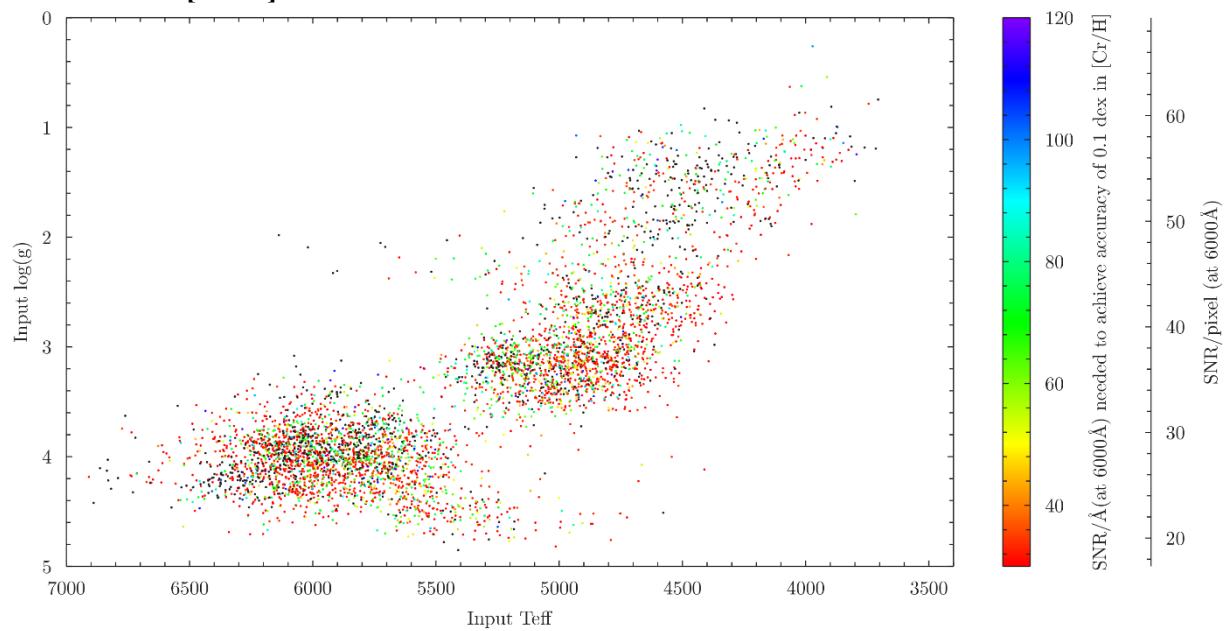
Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 24 of 32

A. 5
[Mg/H]

A. 6
[Ti/H]


Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 25 of 32

A. 7
[Si/H]

A. 8
[Na/H]


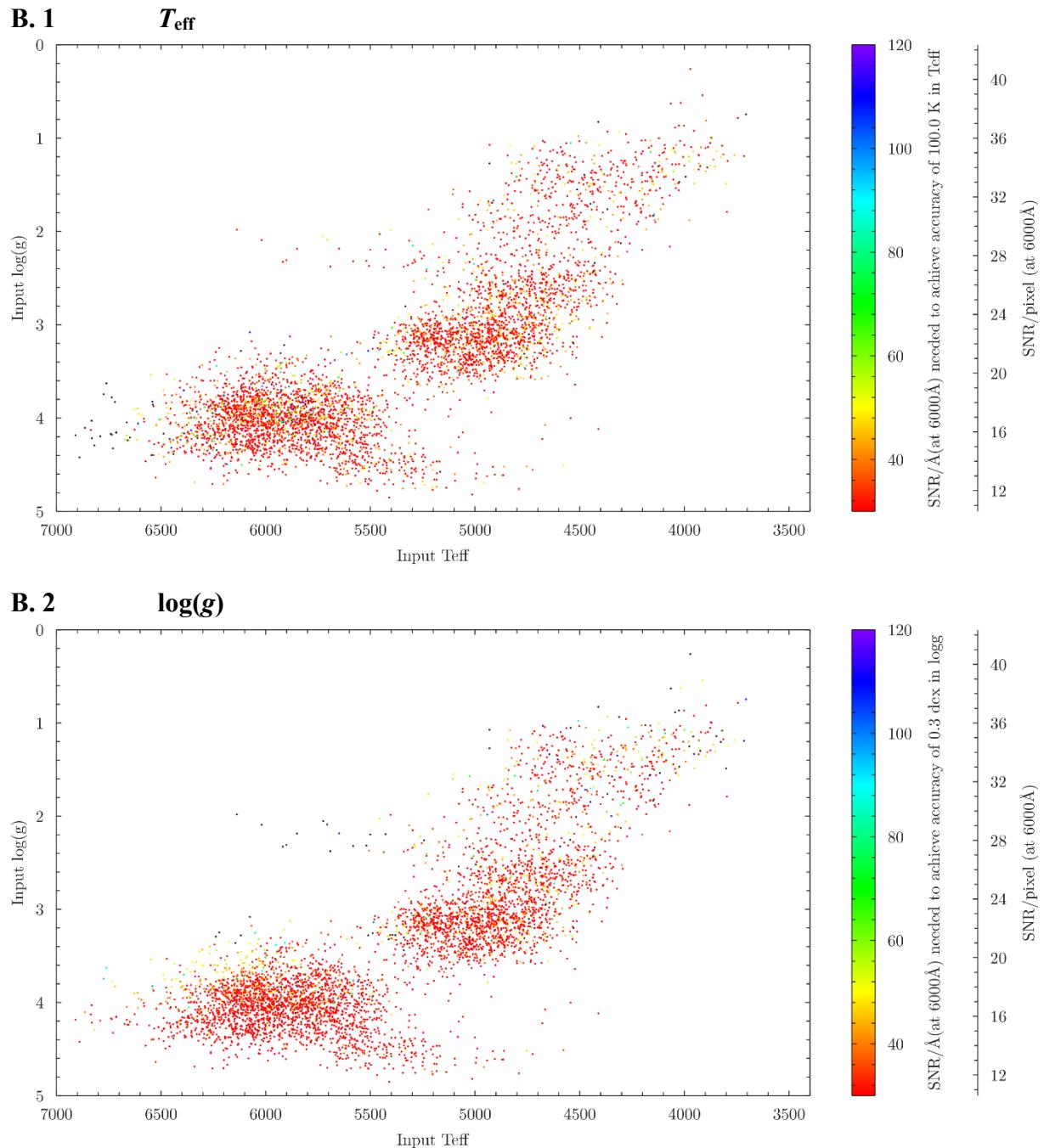
Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 26 of 32

A. 9
[Ni/H]

A. 10
[Cr/H]


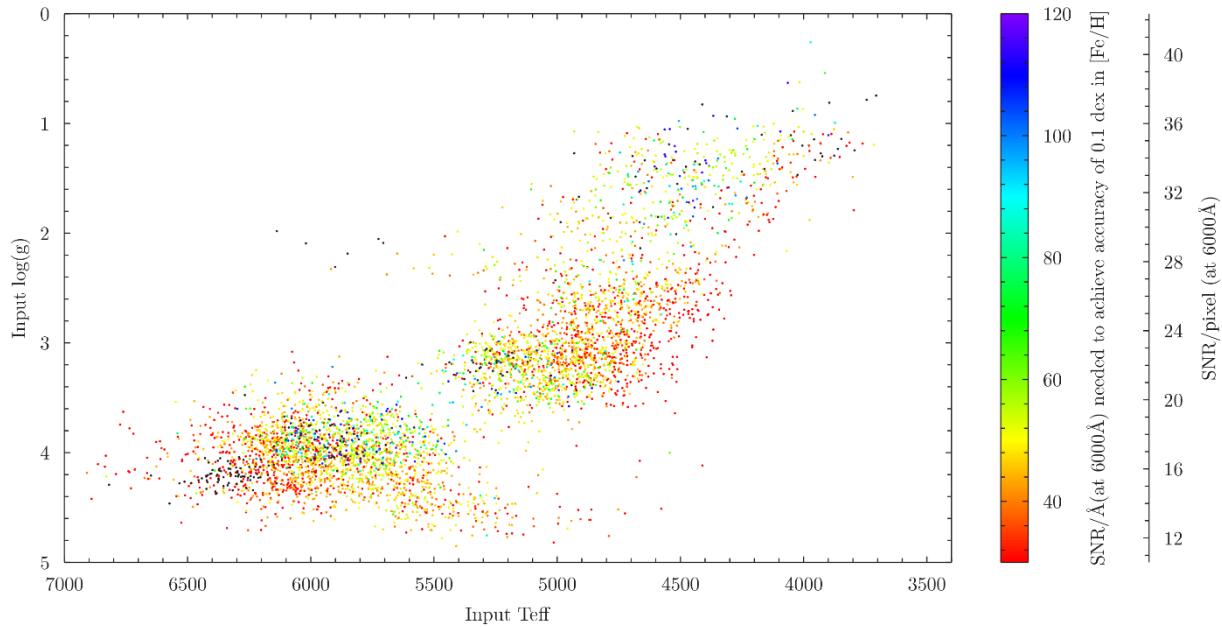
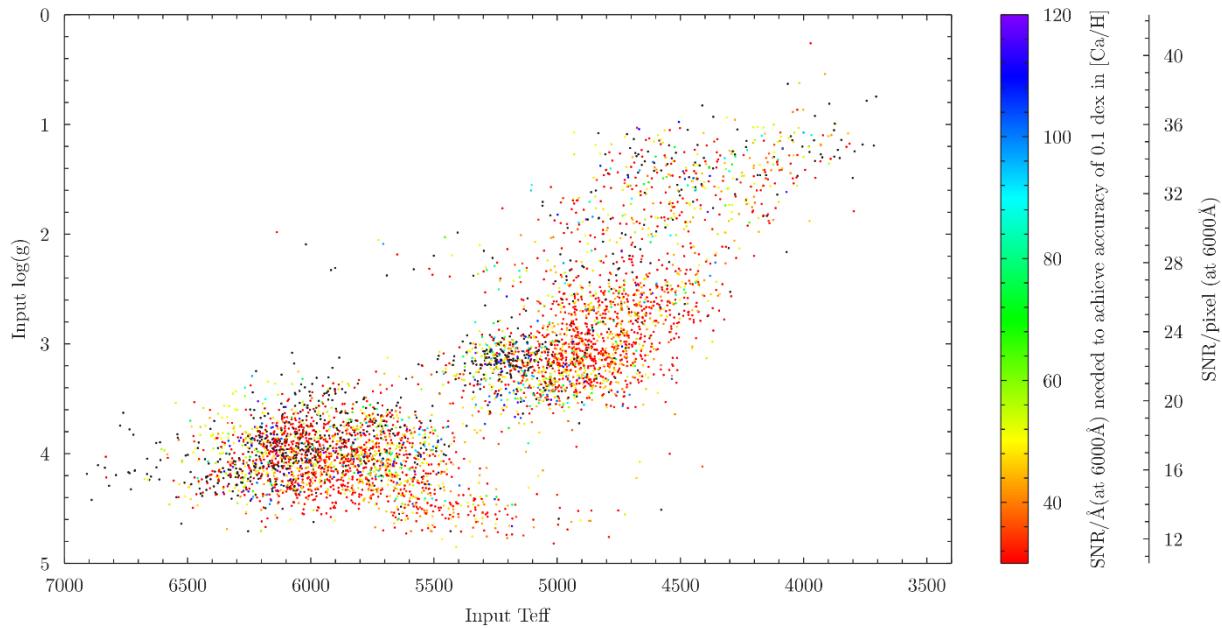
Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 27 of 32

Appendix B: Mapping performance across the HR diagram (HRS)

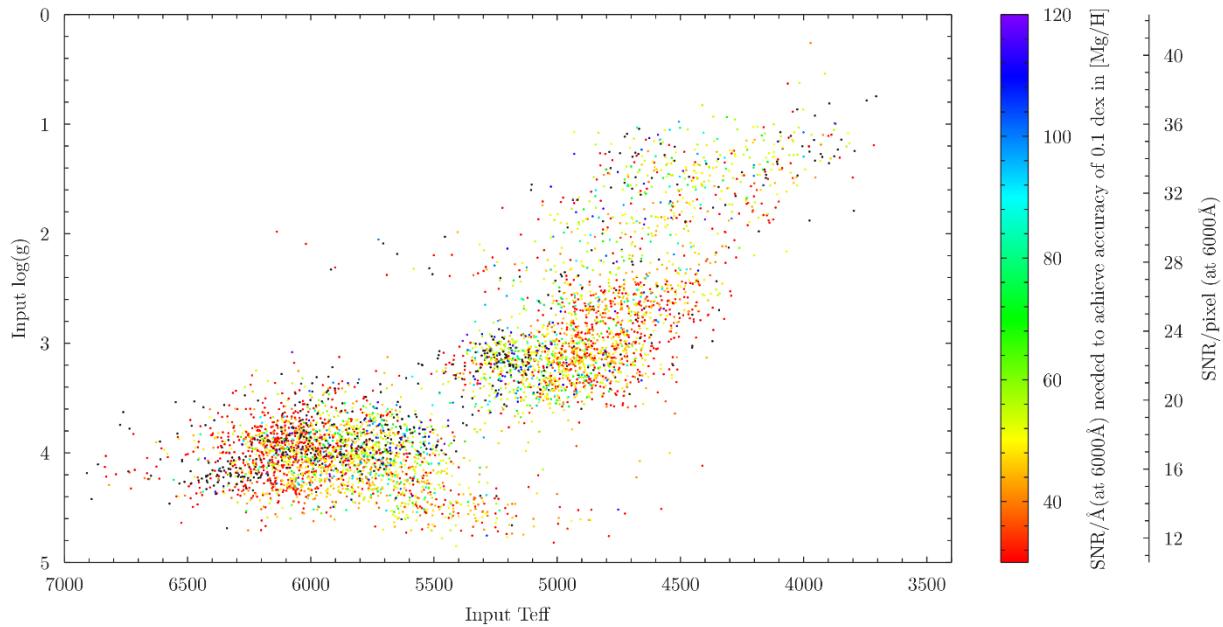
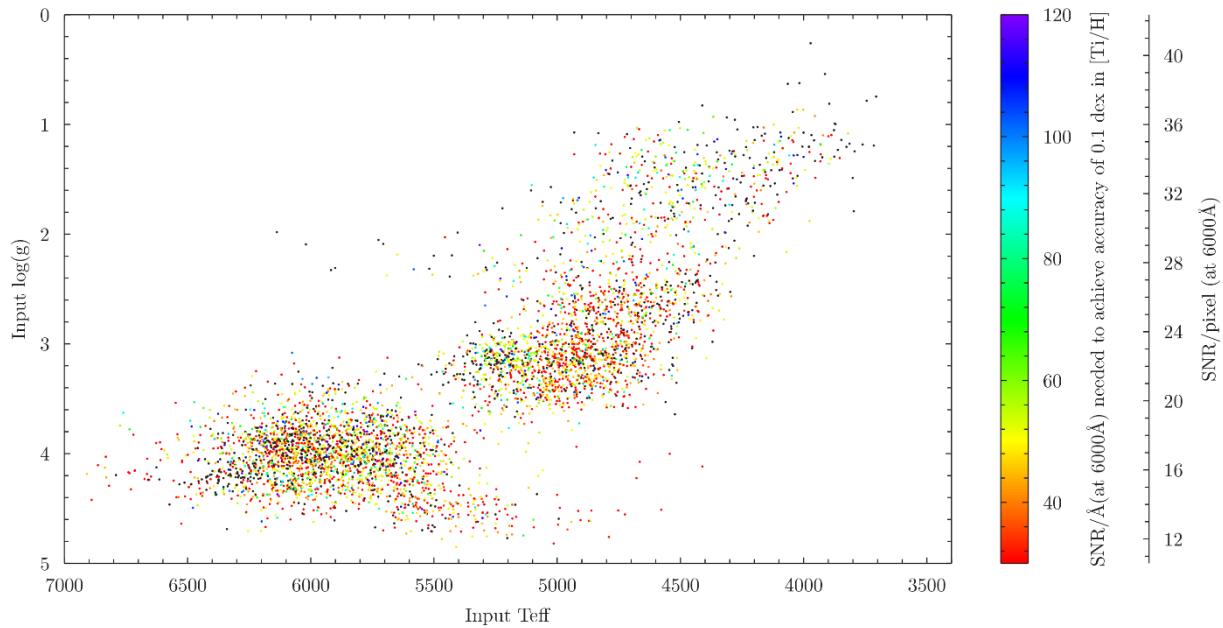
See Section 10 for more information.



Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 28 of 32

B. 3
[Fe/H]

B. 4
[Ca/H]


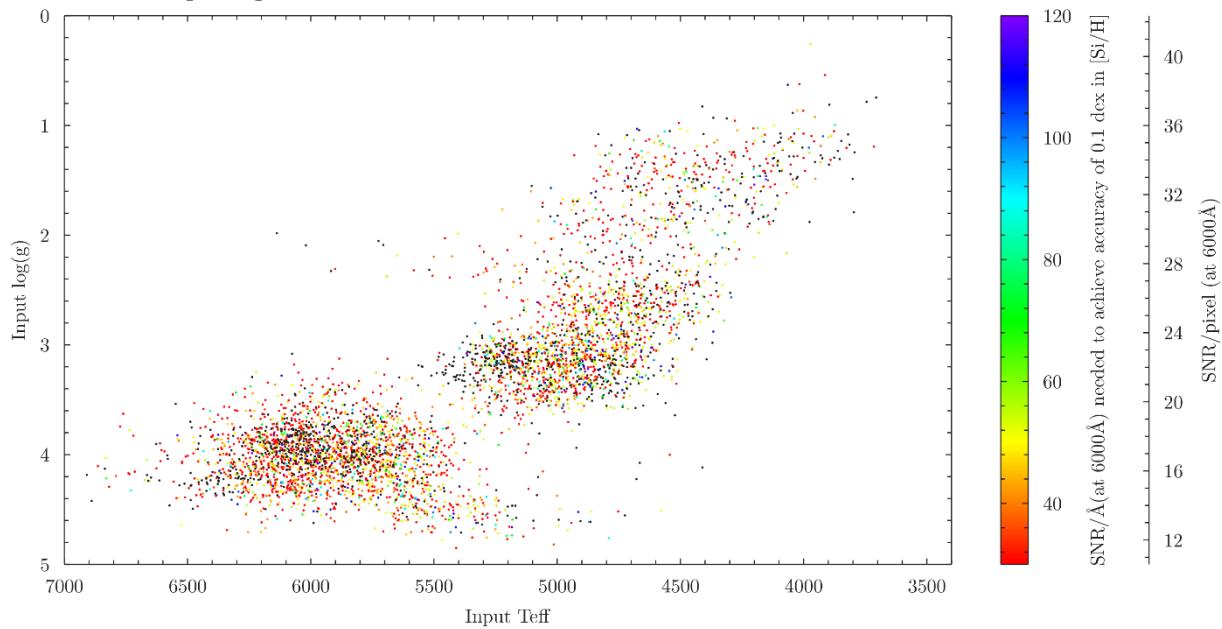
Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 29 of 32

B. 5
[Mg/H]

B. 6
[Ti/H]


Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 30 of 32

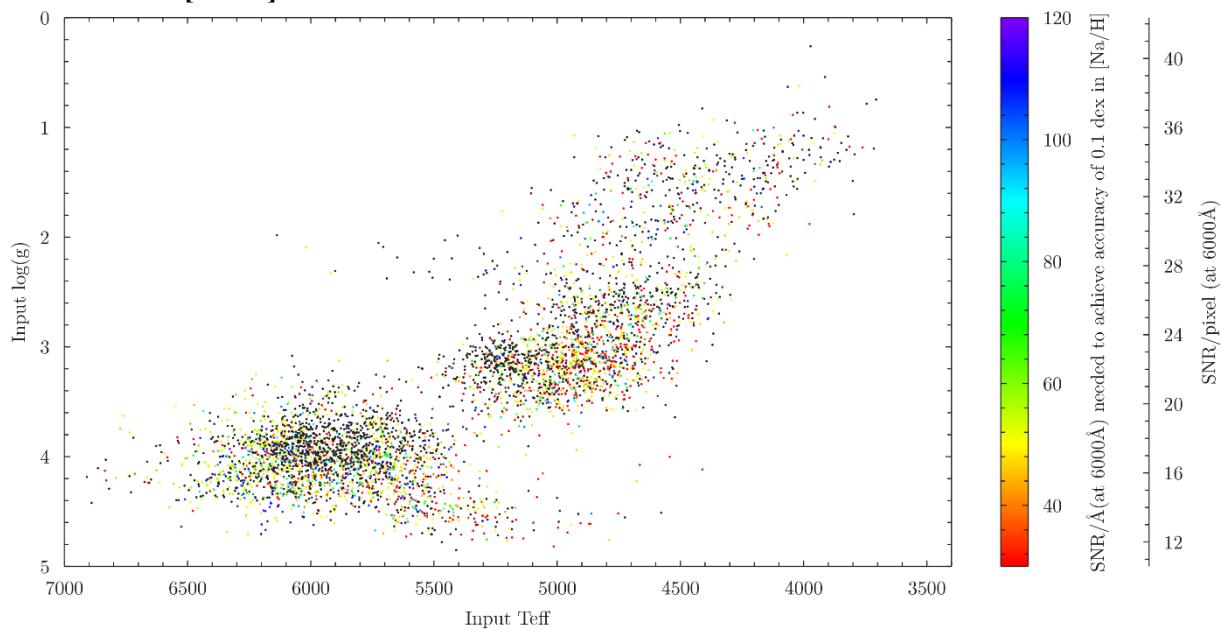
B. 7

[Si/H]

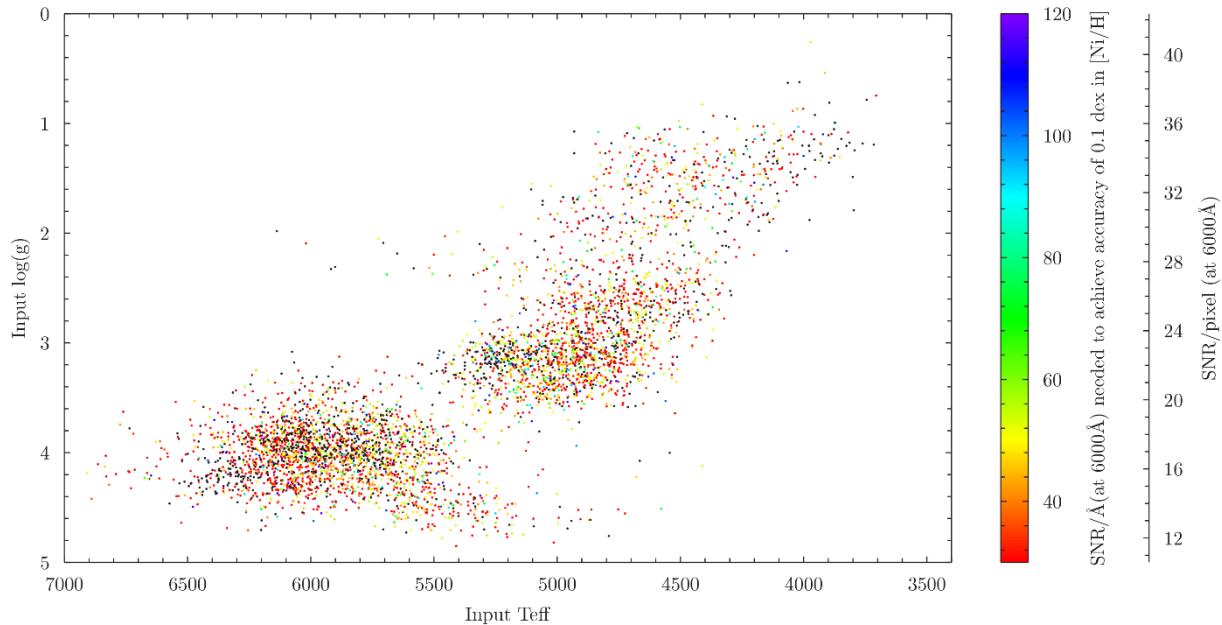
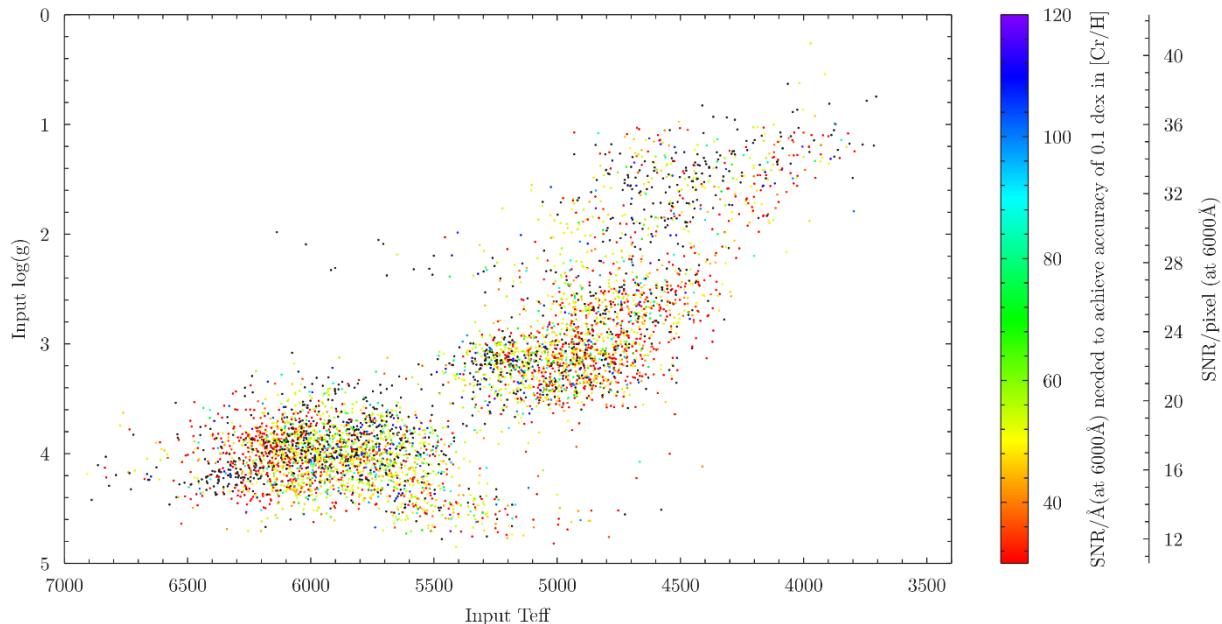


B. 8

[Na/H]



Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP): Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 31 of 32

B. 9
[Ni/H]

B. 10
[Cr/H]


Appendix C: List of Acronyms

List of Acronyms

List of Acronyms	
4FS	4MOST Facility Simulator
4GP	4MOST Galactic Pipeline



Doc. Title: The IWG7 4MOST Galactic Pipeline (4GP):	
Development Report 2	
Doc no.: MST-TNO-XXX-YYYYY-GGGG-HHHH	
Issue no.: 1.00	Date: choose date
Document status	Page 32 of 32

List of Acronyms

4MOST	4-metre Multi-Object Spectroscopic Telescope
APOGEE	The Apache Point Observatory Galactic Evolution Experiment
DIB	Diffuse Interstellar Bands
HRS	High-Resolution Spectrograph
IWG	Infrastructure Working Group
LAMOST	The Large Sky Area Multi-Object Fibre Spectroscopic Telescope
LRS	Low-Resolution Spectrograph
RMS	Root mean square
RV	Radial velocity
SNR	Signal-to-noise ratio