Measurement of the e^+p charged current cross sections with the ZEUS detector at HERA and QCD analysis of deep inelastic scattering data.

C. Collins-Tooth

Imperial College, London

A thesis submitted for the degree of

Doctor of Philosophy of the University of London

March 2004

Abstract

An account of lepton spin polarisation at HERA is presented, including a summary of two test beam experiments on the transverse polarimeter conducted during the 2001-2002 HERA shutdown. Unpolarised cross sections for e^+p charged current deep inelastic scattering at a centre-of-mass energy of 318 GeV are then measured using a total integrated luminosity of $60.9\,\text{pb}^{-1}$ collected during 1999 and 2000 with the ZEUS detector at HERA. The total cross section and single differential cross sections with respect to Q^2, x and y for $Q^2 > 200\,\text{GeV}^2$ were measured. The reduced double differential cross section was also measured in bins of x and Q^2 in the ranges $280 < Q^2 < 17000\,\text{GeV}^2$ and $0.008 < x < 0.42$. The Standard Model is shown to agree with the measured cross sections. The 1999-2000 ZEUS charged and neutral current cross sections are then used in a set of NLO QCD fits. The ZEUS fitting model is investigated and modified, then a fit with 1994-2000 ZEUS data is performed. The improvement to the fit offered by the extra ZEUS data is largely countered by the changes to the fitting model, however the end result still shows overall improvement in the fit uncertainties and has central values which are more consistent with published global fits.

Contents

List of Figures

List of Tables

Preface

After presenting an overview of the theory of deep inelastic scattering (DIS) and the ZEUS detector in the first two chapters, the thesis moves on to describe more closely the experimental and phenomenological investigations performed by myself, both as part of a group and as an individual.

Chapter three describes my contribution to the measurement of lepton polarisation at HERA. In this field, I worked as part of a group of approximately ten individuals drawn from the ZEUS, H1 and HERA-B experiments. My main contribution was to participate in the construction and testing of the silicon microstrip detector for the Transverse Polarimeter. This involved several test beam experiments (conducted at CERN and DESY) to gather initial data. My contribution also included overall analysis of the data to obtain information on the performance of the silicon.

Chapter four outlines the extraction of the unpolarised charged current deep inelastic scattering (CC DIS) cross sections using the ZEUS detector. In this area of my analysis, I worked as part of a group of approximately five people. This analysis was largely based on previous extractions of CC DIS cross sections, with cuts and bin definitions being extremely similar. One particular feature of my analysis was that I required 'good' tracks to have at least two hits in the stereo layers of the ZEUS central tracking detector; a feature not required by the overall ZEUS analysis.

Chapter five then describes how I extracted the proton parton distribution functions using both charged and neutral current data. In this section of my thesis I worked closely with Amanda Cooper-Sarkar, who provided much inspiration and oversight on the phenomenological aspects and implications of the fits. I used a completely separate analysis routine to provide some level of independence between the fits produced by myself and Amanda. The work presented begins by replicating two pre-existing ZEUS fits, to act as a check that my fitting procedure was working correctly. Fit parameters are then variously freed and fixed to examine the impact on the resulting distributions.

1 The Standard Model & Deep Inelastic Scattering

1.1 Introduction

The Standard Model (SM) [1] of particle physics has established itself as a successful and verifiable set of theories with which modern physics is largely able to describe fundamental particles and their interactions. Central to the SM is the assumption that matter is composed of point-like spin $1/2$[1] particles which fall into two families; quarks and leptons. In addition, there are considered to be four fundamental interactions (excluding gravity, these are usually considered to be mediated by a set of gauge bosons):

- The electromagnetic interaction, which is described using quantum electrodynamics (QED);
- The weak interaction;
- The strong interaction, which is described using a non-abelian gauge theory called quantum chromodynamics (QCD);
- Gravity, which is neglected at the energy scales relevant to modern particle physics. This interaction is different to the previous three, in that no field theory exists which is consistent with general relativity.

Electroweak theory [2] unifies the electromagnetic and weak interactions. The combination of electroweak interaction and QCD forms the SM.

It is important to realise that the SM is not the whole story. The limits of the model's validity are constantly being probed. Ultimately the breakdown of the SM at the Planck scale (for particles or systems so massive and pointlike that the gravitational energy of a system becomes comparable to the mass energy) puts a

[1] Note that "natural units", where $\hbar = c = 1$ have been used throughout this thesis.

limit on the viability of the SM.

Recently, observations [3–6] have revealed that neutrinos have a small but non-zero mass, which is not accounted for by the SM. It seems that at the very least, the SM is incomplete in its present state.

1.2 Deep Inelastic Scattering

Experimental evidence was of fundamental importance in the development of the SM. Since the SLAC-MIT results of the late 1960's [7, 8] where 7 and 17 GeV electrons were collided with hydrogen, the proton has been known to possess an internal structure with constituents resembling point-like particles. In the SLAC-MIT experiments, events where the proton was destroyed (termed "inelastic" scattering) gave rise to far more events with a large momentum transfer than would be expected with a continuous distribution of matter. This effect was theorised to be the product of a granular proton structure, akin to the conclusions drawn by Rutherford in his alpha particle scattering experiments on thin foil. In addition, the SLAC-MIT events with large momentum transfer between electron and proton were also termed "deep" because the gauge boson mediating the interaction (and transferring the momentum) had a small wavelength, able to resolve structures smaller than the proton itself. The events were collectively termed "deep inelastic scattering" (DIS).

At about the same time as the SLAC-MIT experiments, Feynman [9] proposed that hadrons (such as protons) were composed of an infinite number of fundamental particles, which he named *partons.* The experimental results and Feynman's theory were brought together in the quark parton model (QPM), where the proton is assumed to consist of a collection of almost free point-like spin 1/2 partons. It was natural to identify these partons with the quarks proposed earlier (and independently) by Gell-Mann and Zweig [10]. The proton was considered to be composed of three "valence" quarks (uud), which give the flavour and charge properties, and a "sea" of quark-antiquark pairs with no net flavour or charge.

In the QPM, deep inelastic scattering can be kinematically characterised by the following momenta (see figure 1):-

- k, k' the incoming and outgoing lepton four-momenta;
- P the incoming proton four-momentum;
- P' the final hadronic state four-momentum;
- $q = k - k' = P - P'$ the four-momentum transfer between lepton and proton.

DIS is then usually described using the following quantities formed from the momenta:-

- $Q^2 = -q^2$ the negative of the four-momentum transfer squared;
- $x = \frac{Q^2}{2P.q}$ the Bjorken scaling variable;
- $y = \frac{q.P}{k.P}$ the fractional energy transfer to the proton in its rest frame.

The Q^2 of the interaction indicates the probing power of the mediating gauge boson, with higher values of Q^2 corresponding to shorter wavelength gauge bosons (which probe the proton at smaller scales).

Neglecting the particle masses, these variables are related by $Q^2 = xys$, where $s = (k+P)^2$ is the square of the centre of mass energy. The quantities Q^2, x and y can then be reconstructed from any two of four measured quantities: the angles of the scattered electron and jet (θ, γ_h respectively, in figure 1), and the measured energies of the scattered electron and jet.

The value of y is related to the electron scattering angle in the electron-quark centre of mass frame ($\theta^\star$) by

$$1 - y = \frac{1 + \cos\theta^\star}{2}. \tag{1}$$

The variable x can be interpreted in the QPM as the fraction of the proton momentum carried by the struck quark, hence its value lies between 0 and 1. If

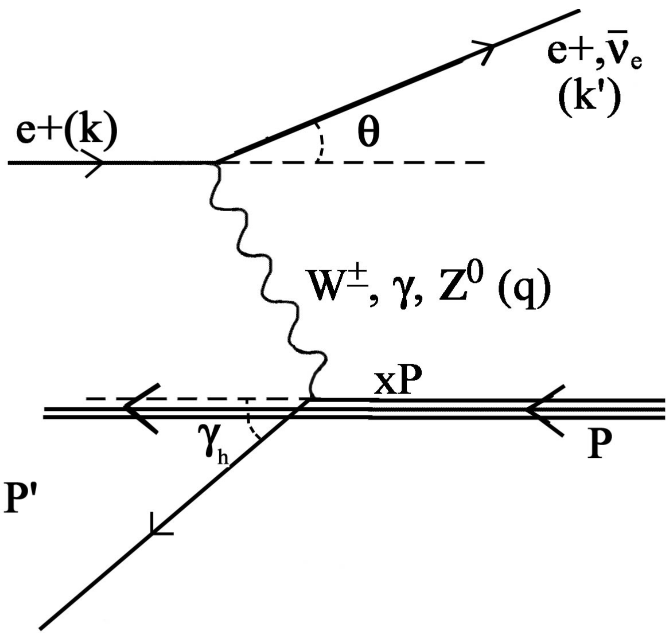

Figure 1: *Kinematics of e^+p DIS, showing the lepton and jet scattering angles relative to the incident beams, θ and γ_h.*

the quarks in the proton are assumed to be non-interacting in the naïve QPM, then the *parton distribution functions* (PDFs), which define the probability of the struck quark being of a particular flavour, should only be a function of x and should not depend on the distance scale to which the proton is probed (given by Q^2). This behaviour was predicted by Bjorken [11], and the independence of the structure functions (formed from PDFs) from the value of Q^2 was named *Bjorken scaling.*

Evidence (for example, see [12]) later emerged that the scaling was not perfect and that a logarithmic dependence on Q^2 is actually observed. In addition, the quarks appeared to carry only about half of the proton momentum [13]. The scaling violation and apparently missing momentum are explained by the theory of quantum chromodynamics (QCD) [14]. In QCD, the quarks interact via the strong force. As has already been mentioned, this interaction is mediated by a set of gauge bosons, which are called called gluons. Gluons interact with quarks, but also with themselves via their colour charge. Quarks may radiate gluons, and gluons may form quark-antiquark ($q\bar{q}$) pairs. These continuous dynamic processes of gluon radiation and $q\bar{q}$ production inside the proton generate the $q\bar{q}$ sea at small x. When this is considered, the scaling violations can be explained, and the missing proton momentum accounted for by the gluons.

At high x, the cross sections are sensitive mainly to the valence quarks. The density of the valence quarks will fall at increasing Q^2 due to gluon radiation, resulting in a decrease in the observed structure functions. At low x the cross section is dominated by the $q\bar{q}$ sea produced by the splitting of gluons, so as Q^2 increases, the number of $q\bar{q}$ pairs increases. This leads to structure functions which rise with increasing Q^2 at low x.

The lepton-proton interaction is considered to occur via the exchange of the electroweak gauge bosons (as shown in figure 1), which can be either electrically neutral (a photon or a Z^0), or charged (a W^+ or W^-). These interactions are known as neutral current (NC) and charged current (CC) deep inelastic scattering respectively. The electroweak gauge bosons interact only with the quarks in the proton, and not with the gluons. In CC DIS, the neutrino in the final state escapes without detection. This provides the characteristic missing transverse momentum signal since, when events are examined, the hadronic jet does not appear to be balanced by the (undetected) neutrino. This thesis focusses on the use of the ZEUS detector in HERA to select these CC DIS events, extract cross sections and perform analysis on the results.

1.3 Cross sections, structure functions and parton distribution functions

The positron and electron CC DIS double differential cross sections are given by [15]:

$$\frac{d^2\sigma^{CC}_{Born}(e^{\pm}p)}{dxdQ^2} = \frac{G_F^2 M_W^4}{4\pi x(Q^2+M_W^2)^2}\left[Y_+F_2^{CC}(x,Q^2) - y^2F_L^{CC}(x,Q^2) \mp Y_-xF_3^{CC}(x,Q^2)\right] \quad (2)$$

where $Y_{\pm} = 1 \pm (1-y)^2$. G_F and M_W are the Fermi coupling constant and mass of the W boson respectively. The structure functions F_2^{CC} and xF_3^{CC} are then defined in the quark parton model as follows:

$$F_2^{CC}(x,Q^2) = \sum_q xq(x,Q^2) + \sum_{\bar{q}} x\bar{q}(x,Q^2) \tag{3}$$

$$xF_3^{CC}(x,Q^2) = \sum_q xq(x,Q^2) - \sum_{\bar{q}} x\bar{q}(x,Q^2) \tag{4}$$

where, for e^+p scattering, the sums run over quarks $q = d, s, b$ and antiquarks $\bar{q} = \bar{u}, \bar{c}, \bar{t}$ (for e^-p scattering, the sums run over $q = u, c, t$ and $\bar{q} = \bar{d}, \bar{s}, \bar{b}$). The parton distribution functions (PDFs), $q(x,Q^2)$, give the probablility of the struck quark being of a particular flavour with a momentum fraction x at a particular Q^2. At HERA, the beam energies are not large enough to give significant contributions from t or b quarks. The structure function F_L^{CC} quantifies the contribution of longitudinally polarised W bosons to the cross section. In the quark parton model $F_L^{CC} = 0$, so that substituting eqns. 3 and 4 into equation 2, it is possible to obtain the values of the e^+p and e^-p cross sections in terms of the PDFs.

$$\frac{d^2\sigma_{Born}^{CC}(e^+p)}{dxdQ^2} = \frac{G_F^2 M_W^4}{2\pi(Q^2+M_W^2)^2}[(\bar{u}+\bar{c}) + (1-y)^2(d+s)] \tag{5}$$

$$\frac{d^2\sigma_{Born}^{CC}(e^-p)}{dxdQ^2} = \frac{G_F^2 M_W^4}{2\pi(Q^2+M_W^2)^2}[(u+c) + (1-y)^2(\bar{d}+\bar{s})] \tag{6}$$

It should be noted that $(1-y)^2$ in equations 5 and 6 is proportional to $(1+cos\theta^\star)^2$ (using equation 1). This means that for $e^{+(-)}p$ CC DIS, the double differential cross section for scattering from quarks (antiquarks) has an angular dependence, which is proportional to suppression by a factor of $(1-y)^2$.

Another quantity, the *reduced* double differential cross section $\tilde{\sigma}$, is also used throughout this thesis, in particular in the PDF determinations. It is defined as follows;

$$\tilde{\sigma}(e^{\pm}p) = \left[\frac{G_F^2 M_W^4}{2\pi x(Q^2+M_W^2)^2}\right]^{-1} \cdot \frac{d^2\sigma_{Born}^{CC}(e^{\pm}p)}{dxdQ^2} \tag{7}$$

1.4 Electroweak radiative corrections

The expression (2) for the CC DIS cross section at Born level must be modified to account for the presence of higher order electroweak interactions. In the analysis of the experimental data, this additional contribution to the cross section has to be separated from the lowest order electroweak Born cross section. By far the largest contribution to the higher order terms comes from the lepton initial state radiation (ISR). This and other CC DIS $\mathcal{O}(\alpha)$ electroweak radiative effects are depicted in figure 2. There are also additional contributions from the W self energy, fermionic and bosonic loops and two boson exchanges, which are depicted in figure 3.

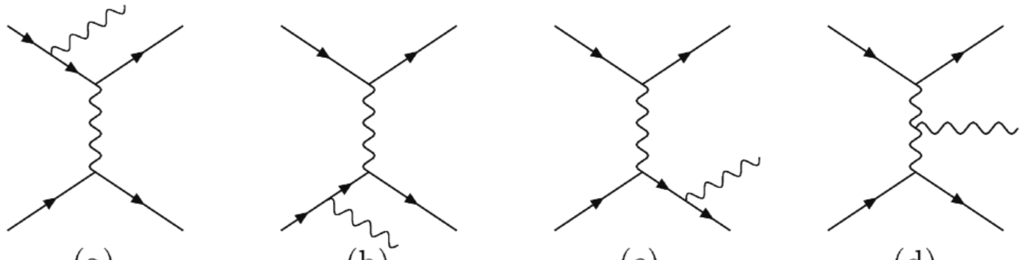

Figure 2: *Corrections applied to the CC DIS Born level cross section at $\mathcal{O}(\alpha)$. Initial state radiation from the incoming lepton (a) and parton (b), final state radiation from the struck parton (c) and radiation from the exchanged boson (d).*

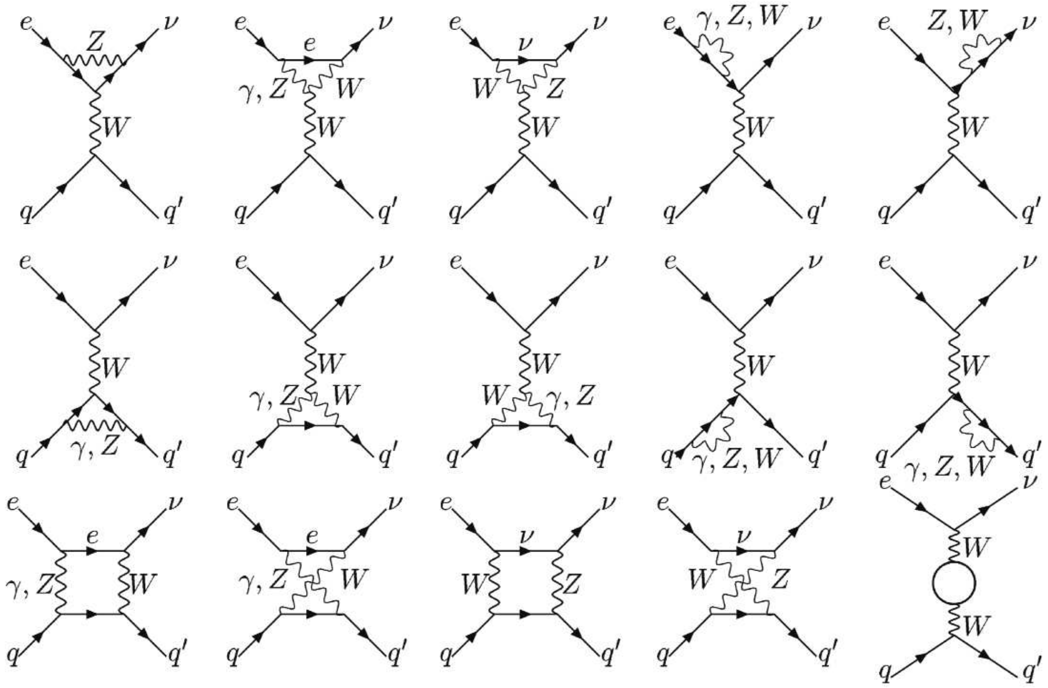

Figure 3: *Additional Electroweak corrections applied to the CC DIS Born level cross section.*

2 HERA and the ZEUS detector

2.1 The HERA accelerator

The Hadron Elektron Ring Anlage (HERA) is the largest accelerator at the Deutsches Elektronen Synchrotron (DESY) laboratory in Hamburg. Protons and positrons (or electrons) are collided at the H1 and ZEUS interaction points depicted in figure 4. Hera B and Hermes use only a single beam colliding with a fixed target. In 1999 and 2000, the lepton and proton beam energies were 27.5 GeV and 920 GeV respectively. The charged current analysis in this thesis focuses on 60.9 pb^{-1} of e^+p data collected by ZEUS during this period (see figure 5). HERA was configured to deliver bunch crossings every 96 ns, with 220 bunches in each beam. Unfilled bunches are left in order that the partner bunches in the other beam (termed 'pilot bunches') may be used in the estimation of non $e^{\pm}p$ backgrounds.

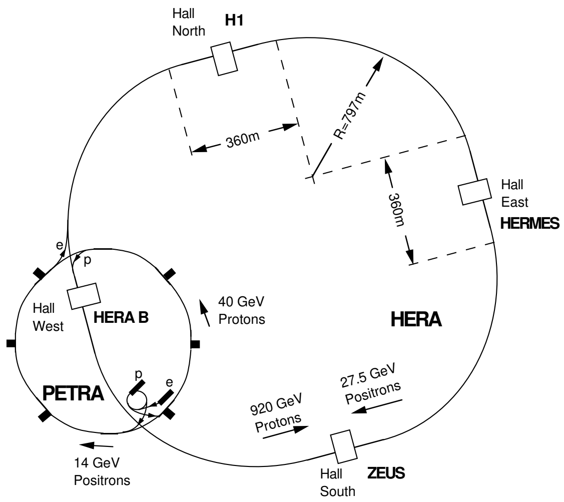

Figure 4: *The HERA accelerator and its experiments.*

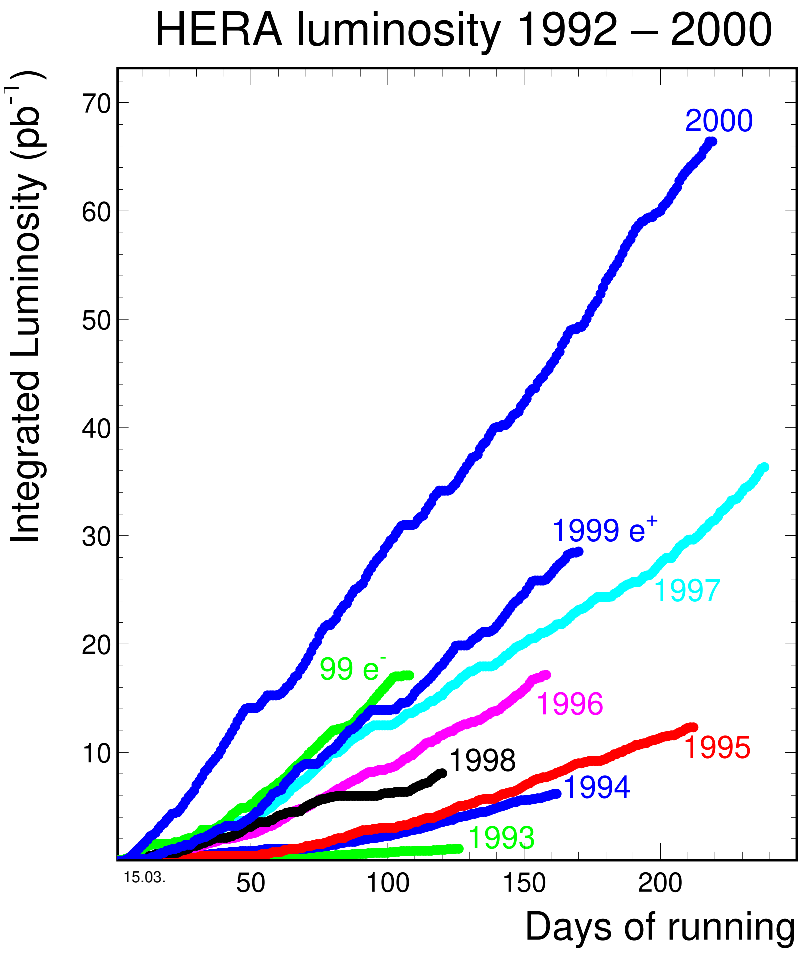

Figure 5: *Luminosity over the years collected by ZEUS.*

2.2 The ZEUS detector

The ZEUS detector [16] (also see figure 6) surrounds the $e^{\pm}p$ interaction point with solid angle coverage of about 99.7%. The detector is arranged with tracking detectors close to the interaction point, with surrounding calorimeters and a muon system at the outside of the detector. The most relevant components of ZEUS used in this analysis are the central tracking detector (§2.2.1), the uranium-scintillator calorimeter (§2.2.2) and the luminosity monitor (§2.2.3).

The ZEUS coordinate system is defined as follows. The proton beam direction is defined as the positive z direction, with the origin at the nominal interaction point. The direction from the interaction point to the centre of the HERA ring is defined as positive x. The positive y direction is then defined so as to complete a right-handed orthogonal coordinate system. In spherical polar coordinates, the polar angle θ is defined with respect to the positive z axis, whilst the azimuthial angle ϕ is defined in the $x - y$ plane with respect to the positive x axis.

2.2.1 The central tracking detector

The central tracking detector (CTD) [17] is a cylindrical drift chamber arranged around the nominal interaction point. It has the ZEUS z axis coincident with its longitudinal axis and tracks charged particles as they traverse the magnetic field of 1.43 T provided by a thin superconducting coil. The tracks are used to determine particle momenta and the $e^{\pm}p$ interaction vertex position. The CTD is composed of 72 layers of anode wires, grouped into 9 'superlayers'. Layers 2,4,6 and 8 are inclined at angles around 5^o relative to the to the z axis (see figure 7). These 'stereo layers' provide a method of determining the z trajectory of tracks. Depending on the number of superlayers traversed, the resolution provided by the stereo layers in the z-direction is approximately 2 mm [18]. For trigger purposes, the so-called "z-by-timing" information is used [19], in which superlayers 1,3 and

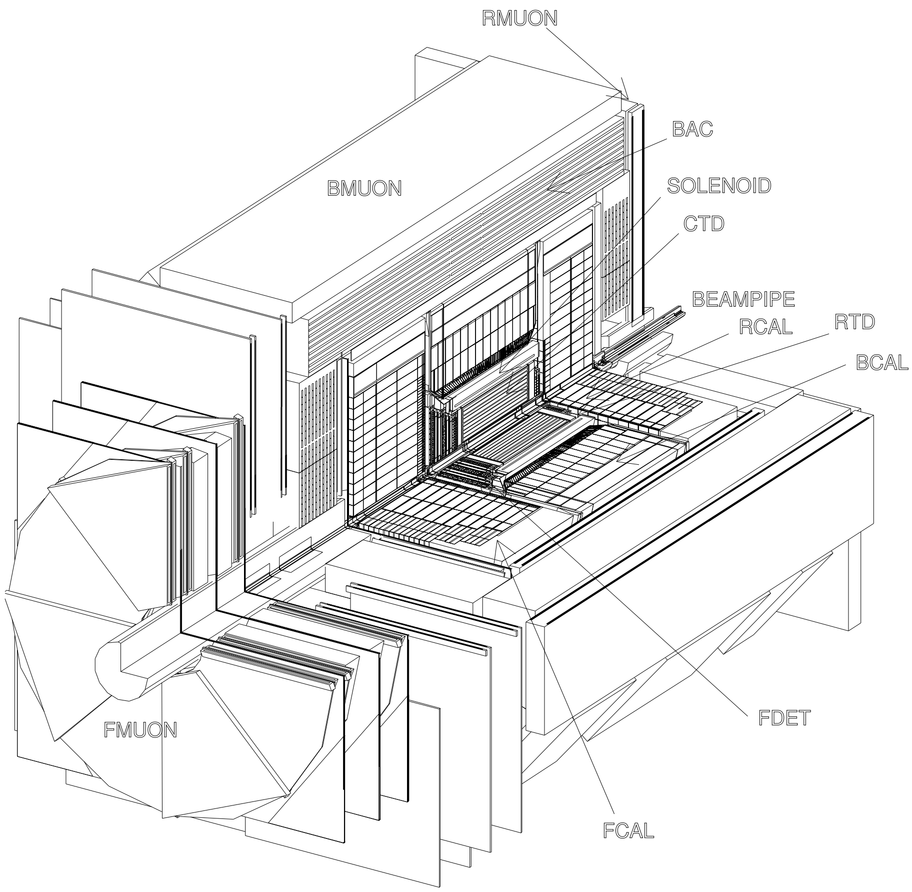

Figure 6: *A schematic of the ZEUS detector.*

5 determine the z-vertex from timing information with a resolution of $\sim 4-12$ cm, depending on the track multiplicity of the event [20]. The innermost superlayer of the CTD, SL1, covers the angular range $11.3^o < \theta < 168.2^o$, whilst SL9 covers $36.1^o < \theta < 142.6^o$, giving typical position resolution in the $r-\phi$ plane of $220-240$ μm (depending on the polar angle θ). This gives a transverse momentum resolution of $\sigma(P_T)/P_T = 0.0058P_T \oplus 0.0065 \oplus 0.0014/P_T$ (P_T in GeV) for tracks traversing all 9 superlayers [21].

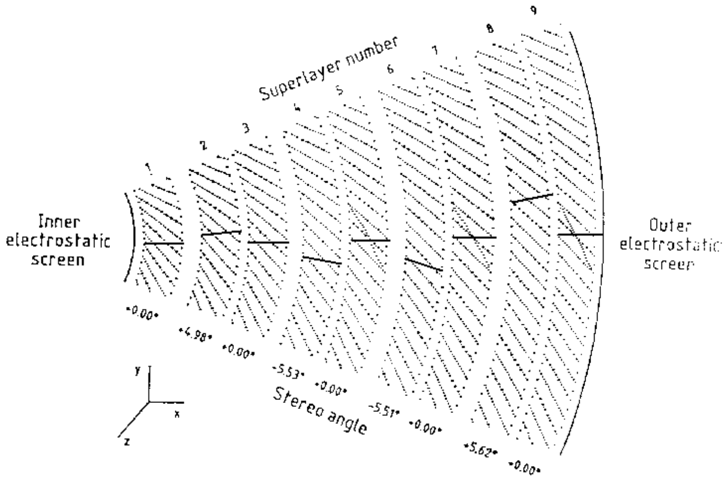

Figure 7: *An octant of the CTD showing the nine superlayers and stereo angles.*

2.2.2 The calorimeter

The ZEUS calorimeter (CAL) is a uranium-scintillator sampling calorimeter. It is composed of three sections, namely the forward (FCAL), barrel (BCAL) and rear (RCAL) calorimeters [22–24]. The individual sections are divided into layers. The FCAL and BCAL each have two hadronic calorimeter (HAC) layers HAC1 and HAC2, and an inner layer which forms the electromagnetic calorimeter (EMC). The RCAL has only one HAC layer, plus an inner EMC layer. An overview of the calorimeter layout is shown in figure 8.

Each layer is composed of cells. In the FCAL and BCAL, the inner face of each

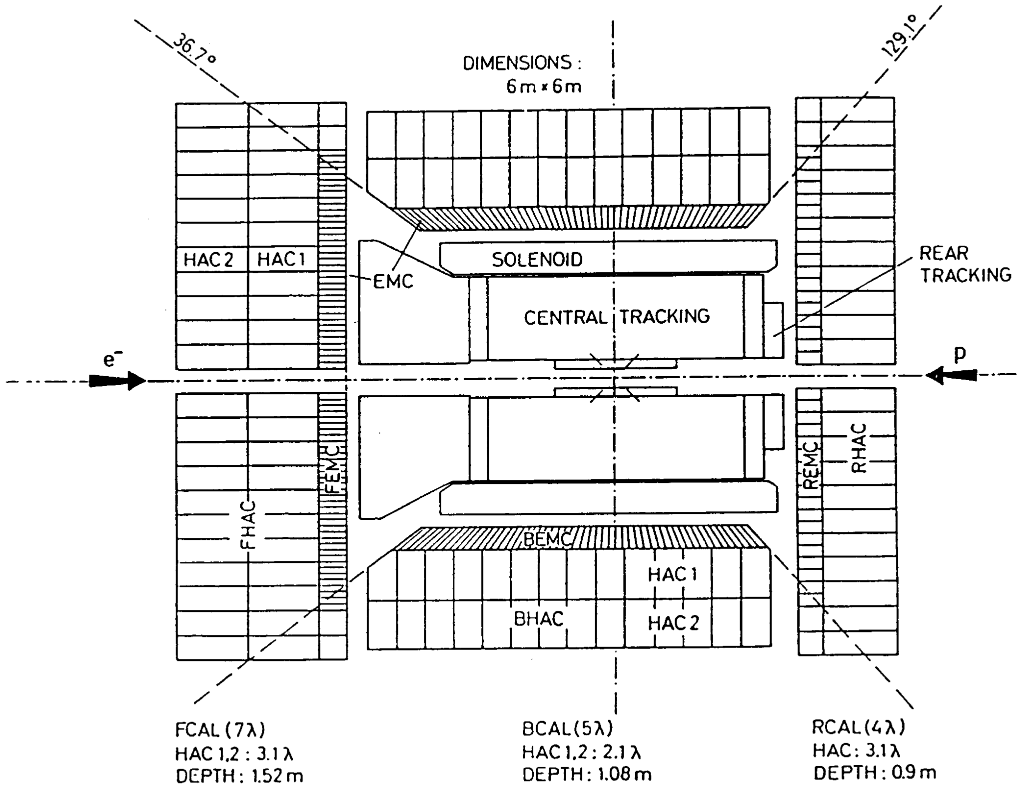

Figure 8: *A simplified side view of the ZEUS calorimeter, showing the three sections and the layers of EMC and HAC cells.*

cell measures $5 \times 20\,\text{cm}^2$ for the EMC and $20 \times 20\,\text{cm}^2$ for the HAC. In the RCAL, the EMC cells measure $10 \times 20\,\text{cm}^2$. In the FCAL and BCAL, a HAC1 cell is combined with a HAC2 cell and four EMC cells, to form a "tower". In RCAL, there is only one HAC layer, and each HAC cell is combined with two EMC cells to form an RCAL tower. Linear groups of towers are then combined into "modules".

Each cell is composed of layers of depleted uranium, having a thickness of 3.3 mm, interleaved with layers of plastic scintillator of thickness 2.6 mm. The scintillations are guided to photomultiplier tubes (PMTs) by wavelength shifters, and each cell is read out by two identical PMTs. This arrangement can be seen for the FCAL in figure 9.

The EMC thickness corresponds to approximately 25 (21 in the BCAL) radiation lengths (X_0) for electromagnetic showers, or approximately 1 interaction length (λ) for hadronic interactions. The total thickness of each calorimeter section varies (to reflect the lopsided event geometry) from 7.1λ in FCAL, to 5.3λ in BCAL and 4λ in RCAL.

The CAL provides timing information for energy deposits, at a resolution of approximately 1 ns for energy deposits over 4.5 GeV [22, 24]. The CAL single particle energy resolutions, as measured under test beam conditions, are $\sigma(E)/E = 0.18/\sqrt{E}$ for leptons and $\sigma(E)/E = 0.35/\sqrt{E}$ for hadronic deposits (with E in GeV). The measured uranium activity allows section-to-section intercalibration at the 1% accuracy level, and long-term absolute calibration at the 0.5% level [22].

2.2.3 The luminosity monitor

The cross section measurements made in this analysis are dependent upon knowledge of the total integrated luminosity captured by the ZEUS detector. The

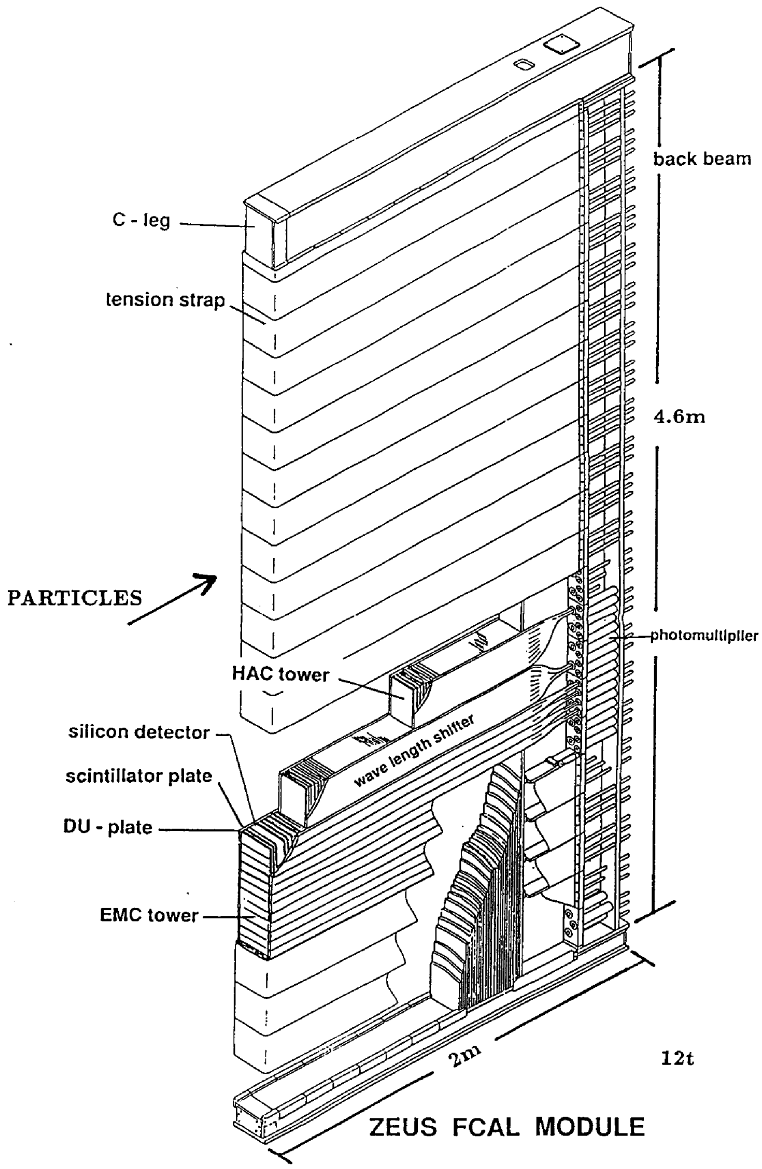

Figure 9: *FCAL towers combined into a module. The wavelength shifters resemble paddles and run parallel to the module sides along the length of each tower, feeding the scintillation light to the photomultiplier tubes.*

luminosity is measured using the Bethe-Heitler process: $ep \to e\gamma p$. The cross section for this process is accurately known (to approximately 0.3% [25]).

The final state particles are emitted at very small angles relative to the lepton and proton beam directions, and so escape the ZEUS detector by passing down the beam pipes. Photons continue down the proton beam pipe until it curves away at approximately $z = -92.5$ m, whereupon they leave the beam pipe through a copper-beryllium window and traverse a graphite shield, to enter a vacuum tube. A photon detector (Lumi-γ), positioned at $z = -107$m, finally detects them. The photon detector is a lead-scintillator sampling calorimeter with face dimensions $18 \times 18\,\text{cm}^2$, and depth equal to 22 radiation lengths (X_0). Scintillator strips laid in the x and y directions are set in the detector at a depth of $3X_0$, and provide information on the shower position with a resolution of approximately 3 mm.

The rate of detection of Bethe-Heitler photons is used, together with detector and analysis cut acceptances, to estimate the total integrated luminosity measured by the ZEUS detector. The collected luminosity was measured with an uncertainty of 2.25% for this analysis.

The Lumi-γ detector is protected against synchrotron radiation (produced by the bending of the positron beam) by the copper-beryllium window and the $3X_0$ graphite shield. During beam injection, the detector is additionally protected by a lead shield.

2.2.4 Trigger

The ZEUS detector uses a three stage trigger [26] which is schematically shown in figure 10. This reduces the potential input rate of 10.42 MHz to a triggered data rate of ~ 5 Hz, by removing backgrounds such as beam gas events or cosmic muons.

The first level of the trigger (FLT) is necessarily hardware based (due to the short bunch crossing time of 96 ns). It is required to have an output rate to the second level trigger (SLT) of the order of $\lesssim$ 200 Hz. Each component involved in an FLT decision (such as the CTD or CAL) coarsely processes its own data and sends signals to the global first level trigger (GFLT). The GFLT uses the component information to make the decision as to whether to keep or discard the event. Whilst the GFLT decision is made using the coarse data, the full event information is stored in digital and analogue pipelines for potential use. The GFLT makes the decision to keep or discard the event based on a logical OR of 64 trigger "slots", each of which has been set up with differing selection criteria.

In charged current DIS event selection, the GFLT makes an event decision primarily by using two key components,- the calorimeter FLT (CFLT) [27] and the CTD FLT [28]. In the CFLT, calorimeter cells are not read out individually as they are in the offline selection. Instead, calorimeter towers are used. A crude estimate of the energy in each EMC or HAC section of a tower is made using a fixed fraction (5%) of the PMT charge, which goes directly to the trigger readout. This method allows fast data processing, but reduces the resolution of the calorimeter. The FLT tower energy is zero-suppressed below 500 MeV, meaning that towers with energy below this threshold are ignored. The tower energies are then digitised on a coarse logarithmic scale, which leads to a reduction in accuracy of the calculated quantities (such as transverse energy) in the FLT. In the CTD FLT, there is insufficient time to perform full track fitting or to use the stereo layers. Instead, the z-by-timing system is used, together with groups of hits to form crude tracks which are analysed to see if they point towards the nominal vertex position. The total number of tracks and the fraction pointing towards the vertex are then used to decide if the event should be discarded.

If a positive decision is made by the GFLT, an accept signal is sent to all components so that they begin to read out the pipelined data. The event then passes to the SLT, where the input rate of $\lesssim$ 200 Hz is reduced to an output rate around $\sim$ 35 Hz. In a similar manner to the FLT, the data from each component are sent to the GSLT, which makes the decision to keep or discard the event. In the SLT, there is sufficient time to use the full spatial and energy resolution of the calorimeter, and CTD tracks are also reconstructed (though still without using the stereo layers for the z-vertex determination). If the GSLT issues an accept signal, the full event information is sent to the event builder which combines information from all detector components and sends the complete event to the third level trigger (TLT).

The TLT reduces the input rate of around ~ 35 Hz to a rate of $\sim\ 5 - 10$ Hz, so that it is possible to store events for offline analysis. The TLT input data are very similar to the offline data (differences arise, for example in offline calibrations). Simplified versions of the offline reconstruction code are run by a farm of computers and physics-based filters used to select events, which are then written to tape.

2.2.5 Veto detectors

The C5 [29] detector is positioned around the beam pipe at $z = -1.2$ m. It consists of four scintillator plates, arranged in two layers, and surrounded by tungsten. It assists in detecting background events, such as beam gas, and is used in the first level trigger.

The veto wall [16] is also used to reduce background. It consists of an instrumented 87 cm thick iron wall positioned at $z = -7.5$ m. Events in which particles enter the detector and pass through the veto wall are rejected.

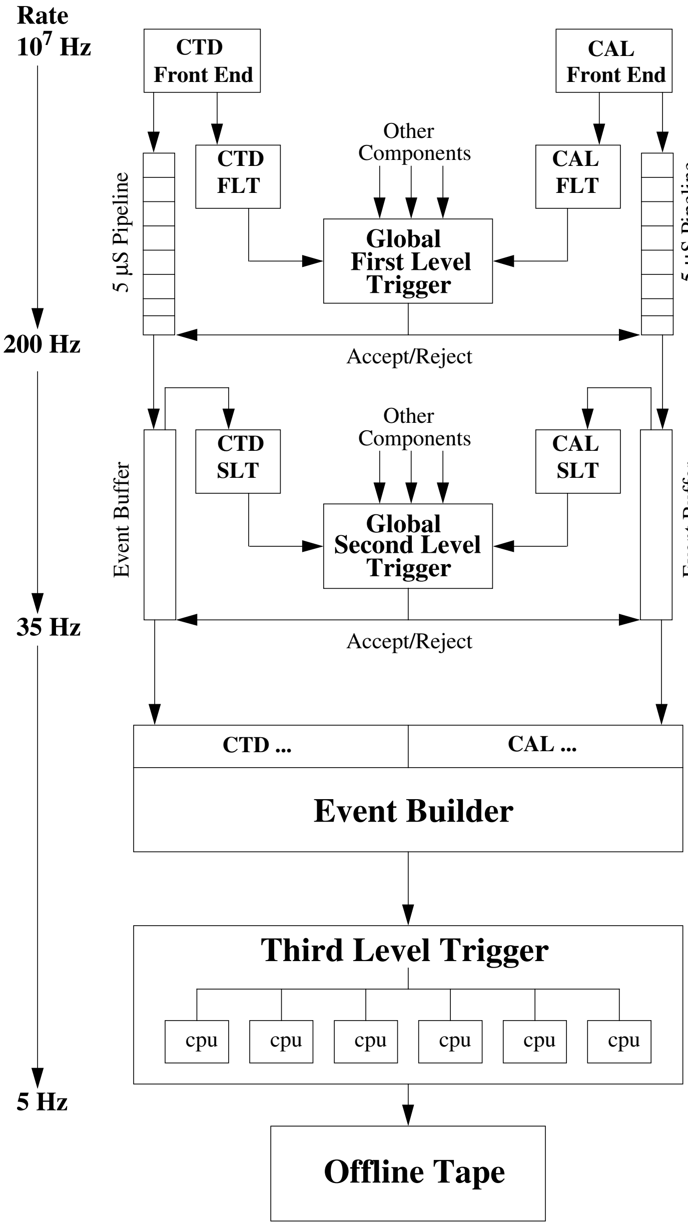

Figure 10: *A schematic overview of the three stage trigger used by the ZEUS detector.*

3 Lepton polarisation

During the HERA shutdown in 2001-2002, preparations were made to enable cross section measurements to be made using spin polarised leptons. Although charged current cross sections are extracted later in this thesis using ZEUS data, it is important to note that these cross sections are calculated using unpolarised leptons. Only runs taken since late 2003 will allow the polarised cross sections to be determined.

To begin with, a short account of spin polarisation and measurement is presented. The changes made during the 2001-2002 shutdown to allow polarised cross section measurements to be made are then outlined. Particular attention is focussed on the Transverse Polarimeter (TPOL), which was upgraded with the aim of reducing the systematic uncertainty on the polarisation measurements to the 1% level. The changes made to the transverse polarimeter were initially investigated in two test beams, one at DESY in January 2001 [30] and another at CERN in July of the same year [31].

3.1 Polarisation principles

Relativistic charged particles in curved portions of a storage ring emit synchrotron radiation in a direction tangential to the ring. This effect causes significant beam energy losses at HERA, in particular in the lepton ring. A small fraction of the radiated photons cause the lepton spin to flip. This effect was first documented at Novisibirsk by Loskutov, Korovina, Sokolov and Ternov [32,33]. For particle spins aligned along a uniform magnetic field, the $\uparrow\downarrow$ and $\downarrow\uparrow$ spin flip rates differ. For positrons, this leads to a build up of spin polarisation parallel to the guide field (electrons become spin-polarised antiparallel to the guide field). The polarisation builds up exponentially in time [34]:

$$P(t) = P_{ST}(1 - e^{-t/\tau_{ST}}) \tag{8}$$

where τ_{ST} is the Sokolov-Ternov polarisation build-up time and P_{ST} is the equilibrium polarisation, which has an asymptotic value of 92.4% at HERA.

The magnetic fields applied in real storage rings are much more complex than the Sokolov-Ternov case. In addition, the emission of synchrotron radiation produces stochastic kicks of the lepton energy, resulting in orbit oscillations. In practice, the orbit oscillations and complex magnetic fields cause diffusion of the lepton spins. For practical purposes the model can be adapted to include depolarisation effects characterised by a time constant τ_{dep} , which serves to limit the maximum achievable polarisation P_{max} and build-up time τ.

$$P_{max} = P_{ST}\frac{\tau_{dep}}{\tau_{ST} + \tau_{dep}} \tag{9}$$

$$\tau = \tau_{ST}\frac{\tau_{dep}}{\tau_{ST} + \tau_{dep}} \tag{10}$$

Historically at HERA, maximum achievable polarisations have been of the order of 50%, with overall build-up times of around 20 minutes [34].

3.2 Towards polarised cross sections

The polarisation process outlined in the previous section produces beams which have a net spin polarisation in a direction transverse to the beam. During the 2001-2002 shutdown at HERA, new spin rotators were provided before and after both of the H1 and ZEUS interaction regions, adding to the existing rotators surrounding the HERMES detector. The spin rotators rotate the spin axis from transverse to longitudinal. The leptons thus have their spins aligned longitudinally (i.e. in the direction of travel) as they transit the interaction regions. (Longitudinally polarised leptons are required by H1 & ZEUS to extract polarised cross sections). The spins are rotated back to the transverse direction by the spin rotators positioned after the interaction regions, so that polarisation can continue to be induced in the curved portions of the storage ring.

The measurement of the lepton beam polarisation is obtained using two polarimeters, the LPOL (longitudinal polarimeter) and the TPOL (transverse polarimeter). The positions of the two polarimeters (and spin rotators) are shown in figure 11. The LPOL is situated near the HERMES detector, in between two spin rotators, where the lepton polarisation vector is oriented along the beam axis. The TPOL is situated in the west section of HERA, and measures the (Sokolov-Ternov induced) transverse polarisation directly. The data gathered by ZEUS and H1, coupled with accurate knowledge of the state of the lepton beam polarisation, will allow polarised cross sections to be measured by these two collaborations in future running.

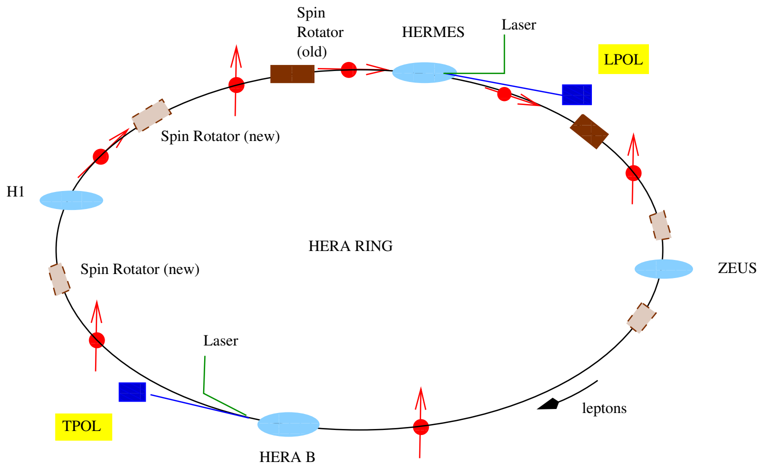

Figure 11: *A schematic overview of the HERA lepton ring, and the position of the major components involved in polarisation studies.*

During the 2001-2002 shutdown, members of the ZEUS collaboration were primarily involved (for polarisation purposes) in the upgrade of the TPOL to reduce the uncertainty of the polarisation measurements made. Before describing the upgrade, the operation of the TPOL prior to the 2001-2002 shutdown will be outlined. This will be followed by a summary of the test beams performed on the upgraded TPOL during the 2001-2002 HERA shutdown. A short summary of the physics potential of polarised measurements is then given.

3.2.1 TPOL operation prior to 2001-2002 shutdown

A detailed account of the TPOL operation prior to the 2001-2002 shutdown exists in [35], however a summary will be presented here. A schematic overview of the TPOL is shown in figure 12. The TPOL calorimeter is situated ~ 200 m north of the west interaction point (HERA-B), whilst the laser apparatus is housed above ground. Light from the Ar-Ion continuous wave laser with a wavelength of 514.5 nm and a power of 10 W is guided into the HERA tunnel and focussed onto the lepton beam at a crossing angle of 3.1 mrad. The laser power provides a probability of a Compton interaction with a lepton bunch of approximately 1/1000. The HERA bunch crossing time of 96 ns thus provides an interaction rate of approximately 10 kHz, with a negligible probability for an overlay of two (or more) simultaneous scattering processes. The laser light is circularly polarised up to values around 99.5% by a Pockels cell. At the TPOL interaction point, the transversely polarised lepton beam interacts with the circularly polarised laser photons. The photons are backscattered via the Compton effect into a calorimeter situated 65 m downstream, where the lepton beam arcs away under the influence of the storage ring dipole magnets. The calorimeter is described in detail in [35], but the main feature to note is that it is segmented into upper and lower halves to enable lepton polarisation measurements.

A typical TPOL measurement lasts approximately one minute. For about 20 seconds, the laser is stopped by a shutter. This allows measurement of the background which is dominated by bremstrahlung. For the remaining 40 seconds, the shutter is opened and the laser is alternated between left and right handed circular polarisation states at a frequency of 90 Hz. Data are then accumulated by the calorimeter for the three states “laser left”, “laser right” and “laser off”.

The key principle on which the TPOL operates is that the Compton scattering cross section for circularly polarised photons incident on transversely polarised leptons is asymmetric with respect to the orbital plane of the leptons [35]. The

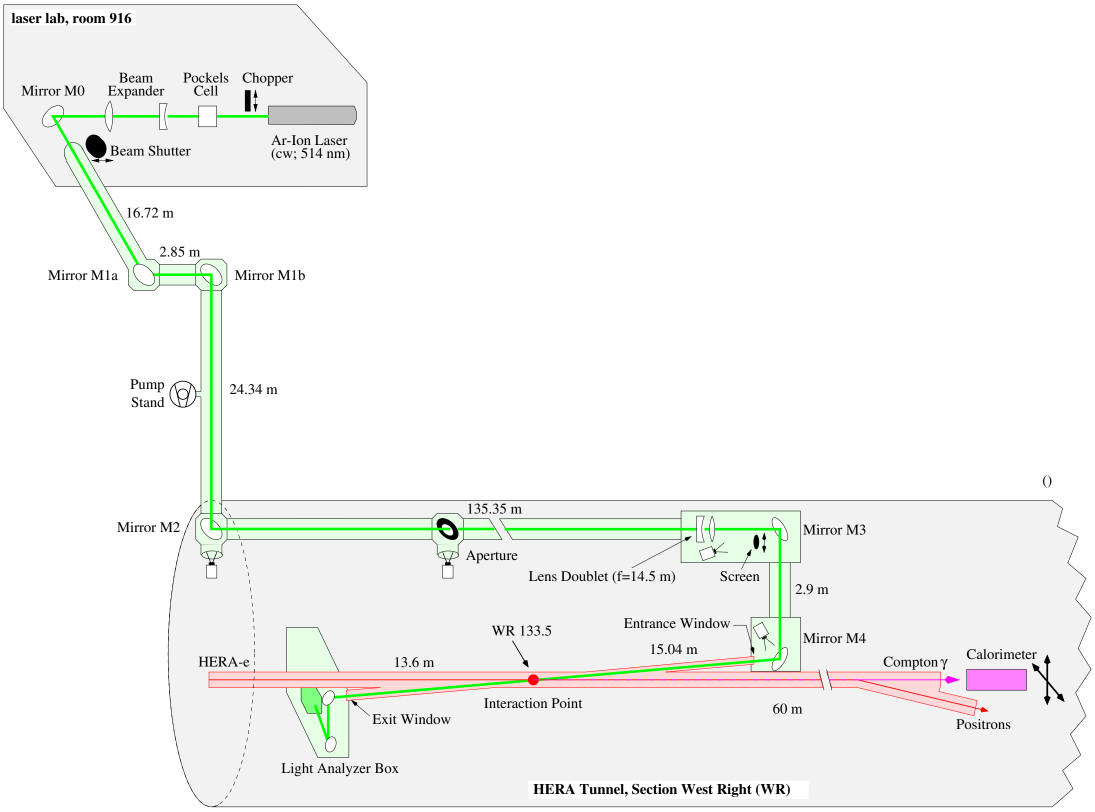

Figure 12: *The TPOL and the HERA tunnel. The calorimeter detects Compton scattered photons from the lepton beam.*

magnitude of the asymmetry is proportional to the vertical polarisation of the leptons. At the TPOL calorimeter this effect causes an up-down asymmetry in the photon distribution, meaning that the energy distribution measured in the two halves of the TPOL calorimeter is actually a function of the lepton polarisation.

Online measurements of the lepton polarisation utilise a quantity, η, which is defined as follows:

$$\eta = \frac{E_u - E_d}{E_u + E_d} \tag{11}$$

Where $E_{u,d}$ are the energies measured in the 'up' and 'down' halves of the calorimeter. The value of η is related to the y-position of the photon entry into the calorimeter and also to the energy of the photon. This relationship (prior to the 2001-2002 shutdown) was evaluated using simulations and test beam data, which provided the so-called "$\eta - y$-transformation" for a particular photon energy. Prior to the 2001-2002 shutdown, this transformation was the dominant source of systematic uncertainty on the polarisation measurements.

The calculation of the lepton polarisaton P_y was then made as follows:
As a function of photon energy (E_γ), the change in the mean value of η when "laser left" and "laser right" data are used was calculated:

$$\Delta\eta(E_\gamma) = \frac{<\eta_L> - <\eta_R>}{2}. \tag{12}$$

The value of $\Delta\eta(E_\gamma)$ is related to the lepton transverse polarisation by the following equation:

$$\Delta\eta(E_\gamma) = P_y.\frac{S_{3,L} - S_{3,R}}{2}.\Pi_\eta(E_\gamma). \tag{13}$$

$S_{3,L}$ and $S_{3,R}$ are the degrees of left and right circular polarisation of the laser light respectively. The value of the "analysing power" $\Pi_\eta(E_\gamma)$ is equal to the shift expected when $P_y.\frac{S_{3,L}-S_{3,R}}{2} = 1$. It is derived from the $\eta - y$-transformation, and as such, contains the dominant systematic uncertainty on the polarisation measurements. Finally P_y is obtained by fitting the measured $\Delta\eta(E_\gamma)$ to $P_y.\frac{S_{3,L}-S_{3,R}}{2}.\Pi_\eta(E_\gamma)$, with $\left[P_y.\frac{S_{3,L}-S_{3,R}}{2}\right]$ as a free parameter.

During the 2001-2002 shutdown, two testbeams were performed with the aim of demonstrating that the $\eta - y$-transformations could be derived at the same time as polarisation data are taken (i.e. without the need for pre-calibrated $\eta - y$-transformations).

3.2.2 Test beam at DESY in January 2001

The details of the experimental setup and the performance of the apparatus in the test beam are presented fully in [30]. A summary is presented here. The aim of the test beam was to investigate the viability of using a silicon strip detector to determinine the TPOL calorimeter $\eta - y$-transformation in-situ, i.e. whilst HERA data are taken, with a view to reducing the systematic uncertainty on the polarisation measurement to a value around 1%.

The DESY II ring was used for this test beam. It provided a source of electrons at energies between 1 and 6 GeV. These were directed through a lead 'preshower' and the small prototype silicon detector (used for the $\eta - y$ measurement) into the spare TPOL calorimeter [36], as depicted in figure 13. This test also included a movable fibre detector positioned in front of the silicon detector and aligned horizontally with the prototype silicon strips. For most of the tests using the prototype TPOL silicon detector, the fibre was kept out of the beam. The fibre's intended purpose was to make periodic measurements to ascertain the amount of radiation damage sustained by the silicon detector. A silicon 'telescope' [37] (which included two scintillation counters to act as a trigger) was also included in the test beam to enable the performance of the TPOL prototype silicon detector to be analysed.

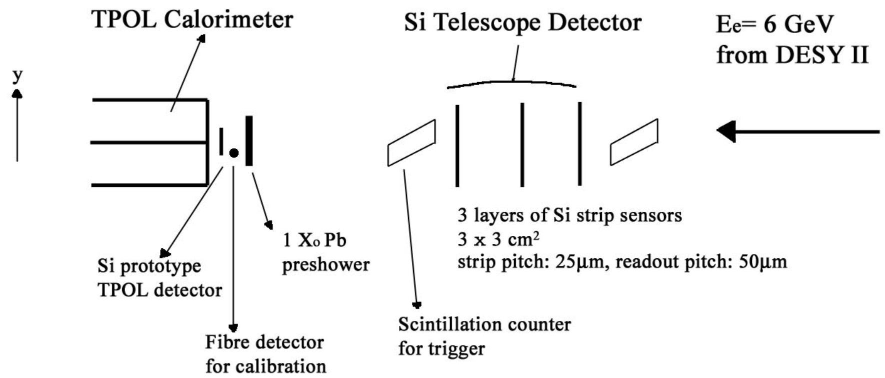

Figure 13: *A schematic view of the DESY test area.*

The prototype silicon detector measured $1 \times 1\,\text{cm}^2$, with an active area of $0.8 \times 0.8\,\text{cm}^2$. It had 100 horizontal strips at a pitch of $80\,\mu\text{m}$ and was read out by a single APV25 chip [38]. The lead preshower was of a thickness of 1 radiation length ($1\ X_0\text{Pb} = 5.6\,\text{mm}$). The preshower is necessary in the HERA tunnel to convert backscattered Compton photons into electron-positron pairs which can be detected by the silicon. The preshower was removed for some of the measurements made in the test beam, in order to understand the performance of the silicon.

The fibre, prototype silicon and calorimeter were housed in a light tight box mounted on a moveable table, so that the relative electron elevation could be varied for the tests. The trigger and readout electronics required to successfully obtain data were housed in a rack ~ 3 m from the TPOL.

Runs were taken both with and without the preshower in position. The $\eta - y$ transformations derived using the test beam data are shown in figure 14. In these plots, the $\eta - y$-transformation was fitted to a three parameter exponential function (for the preshower runs, the fit was only performed down to $-0.1\,\text{cm}$ because the table was $0.2\,\text{cm}$ lower than in the no-preshower runs):

$$\eta(y) = sign(y)\left[1 - e^{-|y|/P_3} + P_1\left(e^{-|y|/P_3} - e^{-|y|/P_2}\right)\right] \tag{14}$$

The DESY testbeam in January 2001 successfully showed that the prototype silicon detector was able to produce a rudimentary $\eta - y$-transformation whilst the TPOL calorimeter gathered data. Silicon parameters such as resolution and efficiency were also investigated during the DESY test beam. The detection efficiency was found to be uniform across the face of the detector, averaging around 98%. It was found that the observed differences between strips actually firing in the TPOL prototype silicon and those predicted to fire (using tracks through the telescope) were Gaussian distributed with a width of $132\,\mu\text{m}$. This was consistent with good performance of the TPOL silicon when material in front of the apparatus was taken into consideration.

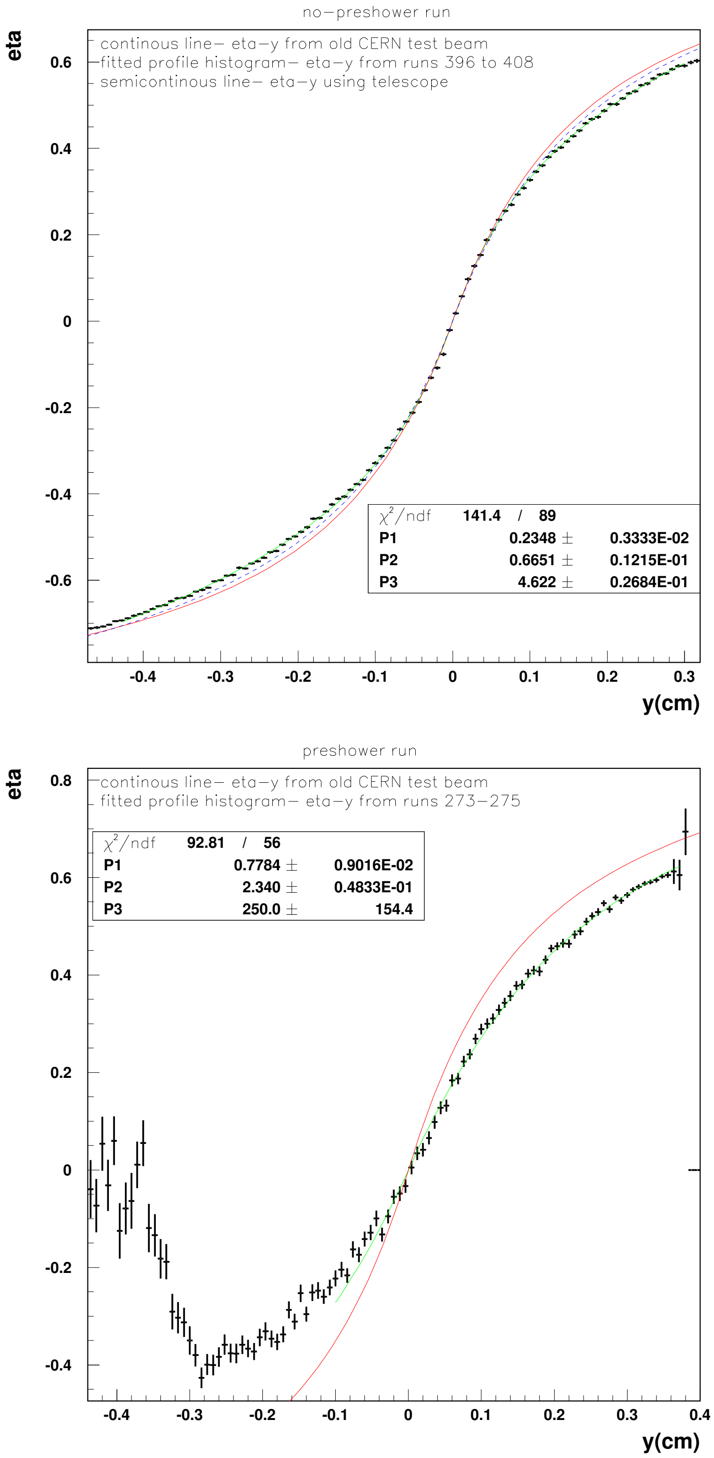

Figure 14: *The fitted $\eta - y$ profile histograms for the two data sets with preshower removed and inserted. The old CERN test beam $\eta - y$ transformation is shown in both plots. An $\eta - y$ transformation derived using the telescope is shown in the top plot. The inclusion of the preshower causes a flattening of the $\eta - y$ distribution (for the preshower runs, the fit was not expected to agree beyond -0.1 cm due to the table position being 0.2 cm lower than in the no-preshower runs).*

3.2.3 Test beam at CERN in July 2001

After the successful test beam at DESY, higher energy measurements and a test of the full-sized TPOL silicon were still required to be made. The TPOL spare calorimeter used in the DESY tests, the fibre and full-sized silicon were transported in their light-tight housing to the CERN SPS ring, which provided a source of electrons and positrons with energies ranging from $6 - 50\,\mathrm{GeV}$ (this being the region of interest for the TPOL). The results of this test beam are presented fully in [31]. A summary is presented here.

The silicon detector used for this testbeam consisted of two sub-detectors. One measuring y position and the other being designed to measure x position. The silicon detector y position measurement is vital for the TPOL in-situ $\eta - y$-transformation, whilst the intended x measurement would offer complete information on the cluster size and position on the calorimeter face. The y detector used for the test beam had a pitch of $80\,\mu\mathrm{m}$, with 768 strips, measuring approximately $6 \times 6\,\mathrm{cm}^2$. The x detector was not ready in time for the test, so a dummy detector of the same thickness ($300\,\mu\mathrm{m}$) was used, in order to keep the amount of material in front of the calorimeter as consistent as possible with what would be present in the HERA tunnel. The y detector was read out by six APV25 chips identical to the one used in the DESY tests. A schematic diagram of the test beam setup is shown in figure 15.

Whilst performing an online analysis of the data obtained from the CERN testbeam, it was noticed that some of the silicon strips appeared to give little or no signal. Worse still, a subset of these strips gave intermittent signals. Strips either side of the dead and intermittent strips also appeared to record extra hits. A full account of the tests made (some of them at DESY, after the CERN testbeam) and the strips found to be defective is given in [31]. The silicon responses received from a typical run are shown in figure 16. Dead strips are clearly evident, in both upper and lower histograms, along with the extra hits on either side which can clearly be seen in the magnified lower histogram.

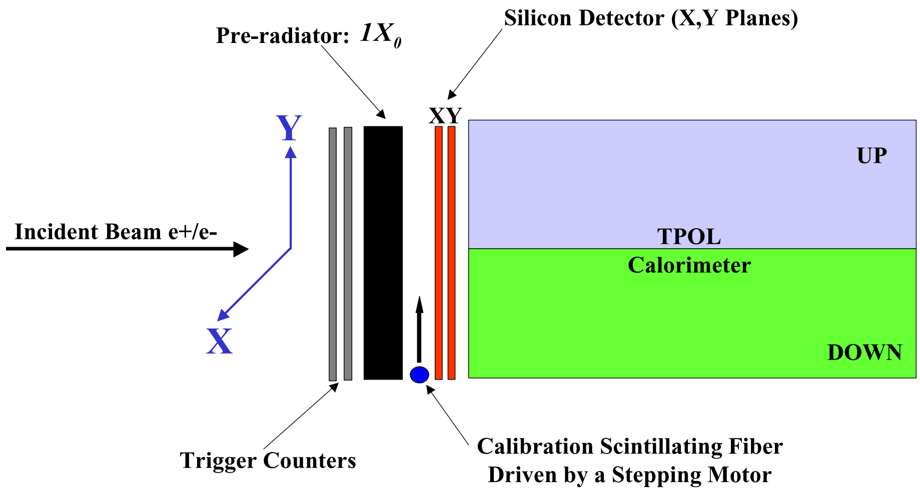

Figure 15: *A diagram showing the TPOL setup at the CERN test beam. The silicon x detector used was a dummy replacement.*

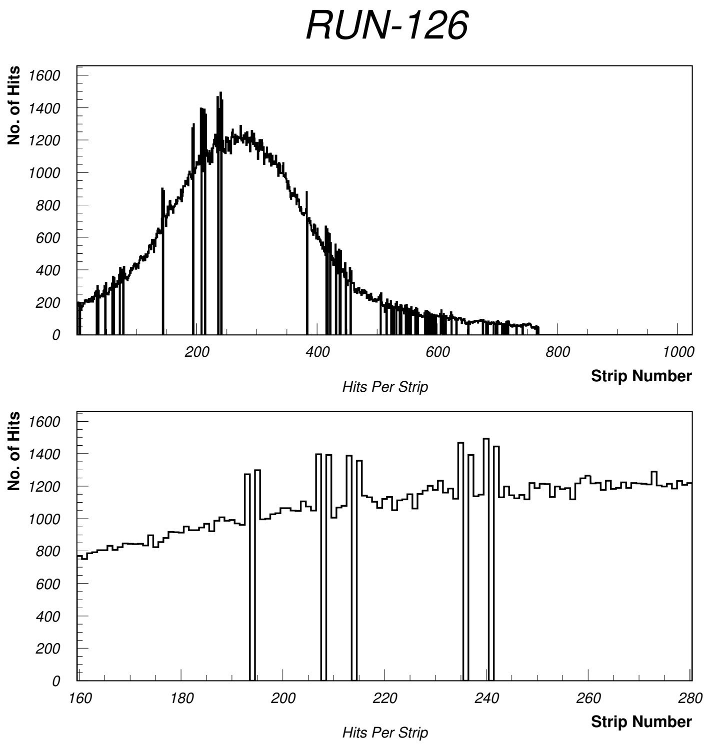

Figure 16: *(Top) Spikes of higher response and dead strips appear in the beam profile captured by the silicon detector. (Bottom) A magnification of the area between strips 160 and 280 shows five dead strips and the strips either side giving higher responses.*

The defective strips covered approximately 11% of the detector. All the defective strips bar two had even strip numbers (one of the two odd-numbered defective strips was never bonded and so was recorded as defective prior to the test). Later investigations showed that the problems emanated from the bonding process used to connect the "pitch adaptor" to the APV25 chips. The pitch adaptor was positioned in between the silicon detector and APV25 chips to provide tracks which converted the silicon strip pitch of 80 μm to the APV25 input channel pitch of 44 μm. It was manufactured from two layers glued together with an adhesive (each layer having an APV25 input pitch of 88 μm). The lower layer (used for odd strips) was made from printed circuit board (PCB) material having a thickness of 500 μm. The upper layer (used for even strips) was made from copper coated Kapton having a thickness of 100 μm. For the dead strips, the electrical contact of the bond on the Kapton surface was found to be the cause of the problem, effectively leaving them unbonded. Charge leakage from the dead strips to the strips on either side also appeared consistent with the extra hits which were observed during the test beam.

The bond failures on the Kapton layer appear to have occurred as a result of several factors. The narrow width of the metallised tracks is the first factor identified in [31] as a possible cause. Bonds failed more often for thinner tracks. The amount of ultrasonic energy applied to the bond appears to be much more crucial for tracks any narrower than 50 μm. The tracks on the Kapton layer were indeed narrower (38 − 40 μm) than those on the PCB layer (50 μm). Another possible cause was that the Kapton may have flexed, whilst force was applied to create a firm bond. This effect may have been due to the relative thinness of the Kapton (100 μm) and also possibly to an uneven application of the adhesive layer underneath.

The CERN test beam data were still analysed with a view to demonstrating that in-situ $\eta - y$-transformations could be obtained for the calorimeter at various energies. Each of the silicon dead strips (and the two adjacent "extra hit" strips) were removed from the analysis. A typical $\eta - y$ plot produced using 10 GeV positrons is shown in figure 17. The old CERN test beam $\eta - y$ transformation is also shown for reference. This was expected to differ from the new fit primarily due to the presence of the lead preshower which serves to flatten the transformation (i.e. it gives a lower value of η, due to the higher likelihood of the incident particle showering before the calorimeter is reached).

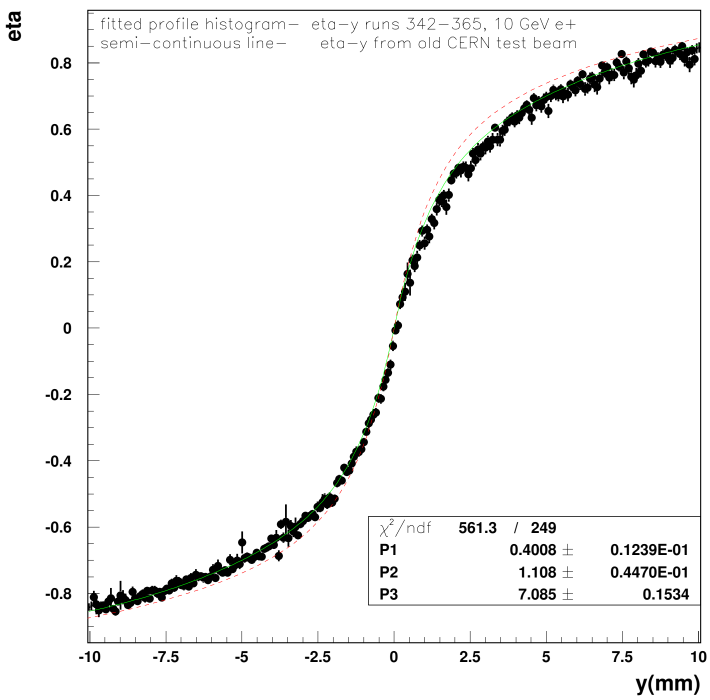

Figure 17: *A fitted $\eta - y$ profile histogram for 10 GeV positrons. The old CERN test beam $\eta - y$ transformation function is plotted for comparison.*

3.2.4 Installation in the HERA tunnel

Since the CERN test beam in July 2001, a new silicon detector has been produced and installed in the HERA TPOL (located in the tunnel). It is identical in most respects to the one tested at CERN in July 2001, however the pitch adapter (connecting the silicon to the APV25 readout chips) is now made of glass, with aluminium tracks. Additionally, the bonding process has been improved to eliminate the intermittent connections experienced during the CERN test beam. A full account of the bonding tests made can be found in [31]. Furthermore, the silicon detector having vertical strips (and thus measuring the x coordinate of the incident particles) has now been included, to replace the dummy x detector used in the CERN test beam. This means that both the x & y coordinates of clusters can now be recorded, but the primary interest still lies in the measurement of the y coordinate for polarisation purposes. The x detector has a pitch of 120 μm, but only has alternate strips bonded to the pitch adapter (and thus read out by the APV25 chips). The TPOL spare calorimeter used in the two test beams at DESY and CERN has also been installed in the tunnel, replacing the original calorimeter, which was identical in manufacture. Figure 18 shows a sample of the data taken in the HERA tunnel [39]. The plots show that the dead strip problem observed at CERN has disappeared. Additionally, it can be seen that the x detector is functional.

3.2.5 Tests of the SM using polarised lepton data

The use of the spin rotators coupled with accurate knowledge of the polarisation state of the lepton beam will allow new tests of Electroweak theory, such as the determination of a direct limit on the mass of the right handed W boson from polarised CC DIS cross section measurements (The SM requires that the W boson is left handed). The polarised CC DIS cross sections (in the QPM, with $F_L = 0$) are given by [15]:

$$\frac{d^2\sigma^{CC}_{Born}(e^+p)}{dxdQ^2} = (1+P)\frac{G_F^2 M_W^4}{2\pi(Q^2+M_W^2)^2}[(\bar{u}+\bar{c})+(1-y)^2(d+s)] \tag{15}$$

$$\frac{d^2\sigma^{CC}_{Born}(e^-p)}{dxdQ^2} = (1-P)\frac{G_F^2 M_W^4}{2\pi(Q^2+M_W^2)^2}[(u+c)+(1-y)^2(\bar{d}+\bar{s})] \tag{16}$$

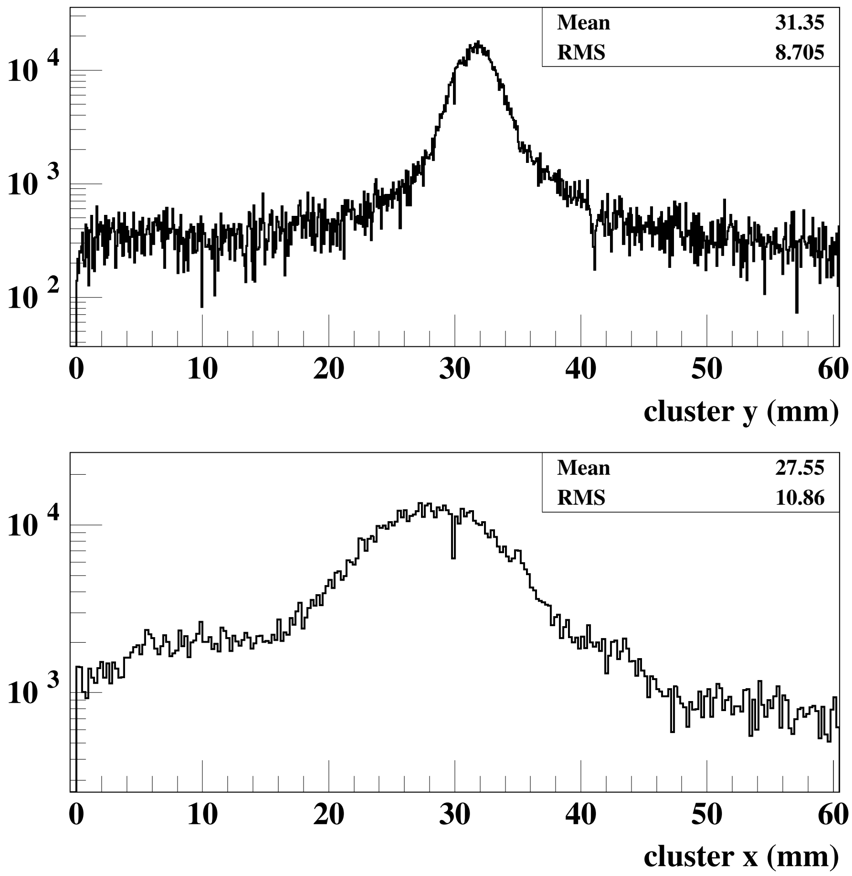

Figure 18: *The beam profile taken using the TPOL silicon detector in the HERA tunnel.*

where the variables have the meanings defined in chapter 1 and P $(=\frac{N_R-N_L}{N_R+N_L})$ is the relative polarisation of the lepton beam. At full polarisation (e.g. for positrons at $P=-1$), the cross section should pass through zero (if the physics is governed by the SM). This will enable a direct estimate to be made on the mass of the right handed W boson.

Neutral current cross section measurements can also be used to probe the SM, for example by measuring the light quark NC couplings to the Z^0 boson, as is proposed in [40]. A summary is presented here for reference. The cross section for NC events in polarised lepton-proton scattering can be written as:

$$\frac{d^2\sigma^{NC}_{Born}(e^{\pm}p)}{dxdQ^2} = \frac{2\pi\alpha^2}{xQ^4}\left(H_0^{\pm} + PH_P^{\pm}\right) \tag{17}$$

Where the effect of the polarised lepton beam is contained in the functions $H_P^{\pm}$ and α is the QED coupling constant. In the QPM with $F_L = 0$, the functions $H_{0,P}^{\pm}$ may be written as:

$$H_{0,P}^{\pm} = Y_+F_2^{0,P} \mp Y_-xF_3^{0,P} \tag{18}$$

where the factors $Y_{\pm}$ are defined as in chapter 1. The functions $F_i^{0,P}$ are the unpolarised and polarised structure functions respectively:

$$F_2^0 = \textstyle\sum_q x(q+\bar{q})A_q^0 \quad ; \quad xF_3^0 = \displaystyle\sum_q x(q-\bar{q})B_q^0 \tag{19}$$

$$F_2^P = \textstyle\sum_q x(q+\bar{q})A_q^P \quad ; \quad xF_3^P = \displaystyle\sum_q x(q-\bar{q})B_q^P \tag{20}$$

A_q^0, B_q^0 are given by:

$$A_q^0 = \quad e_q^2 - 2e_qv_qv_e\chi_Z + (v_q^2+a_q^2)(v_e^2+a_e^2)\chi_Z^2 \tag{21}$$

$$B_q^0 = \quad -2e_qa_qa_e\chi_Z + 4v_qa_qv_ea_e\chi_Z^2 \tag{22}$$

A_q^P, B_q^P are given by:

$$A_q^P = 2e_q v_q a_e \chi_Z - 2(v_q^2 + a_q^2) v_e a_e \chi_Z^2 \quad (23)$$

$$B_q^P = 2e_q a_q v_e \chi_Z - 2v_q a_q (v_e^2 + a_e^2) \chi_Z^2 \quad (24)$$

where e_q is the fractional quark charge relative to the positron. χ_Z (from [40]) is proportional to the ratio of the Z^0 and photon propagators:

$$\chi_Z = \left(\frac{Q^2}{Q^2 + M_{Z^0}^2}\right)\left[\frac{1}{sin^2 2\theta_W}\right] \quad (25)$$

v_f and a_f are the vector and axial couplings of the quarks and leptons to the Z^0 and θ_W is the weak mixing angle. For a fermion f having third component of weak isospin T_3^f (see table 1), the vector and axial couplings are given by:

$$a_f = T_3^f \quad \text{and} \quad v_f = T_3^f - 2e_f sin^2\theta_W \quad (26)$$

Deviations from the SM may appear in the quark couplings a_q and v_q when the polarised NC DIS cross sections are measured.

leptons	e_e	T_3^e	quarks	e_q	T_3^q
e_L^-	-1	$-\frac{1}{2}$	u_L, c_L, t_L	$\frac{2}{3}$	$\frac{1}{2}$
e_R^-	-1	0	u_R, c_R, t_R	$\frac{2}{3}$	0
e_L^+	1	$\frac{1}{2}$	d_L, s_L, b_L	$-\frac{1}{3}$	$-\frac{1}{2}$
e_R^+	1	0	d_R, s_R, b_R	$-\frac{1}{3}$	0

Table 1: *The electromagnetic charge and third component of weak isospin for the leptons and quarks.*

4 Charged Current analysis

In this chapter, the unpolarised charged current DIS cross sections are extracted using 1999-2000 e^+p ZEUS data. The process of determining the cross sections begins with a discussion of event simulation (§4.1), which is used to model both signal (CC DIS) and background events. The event reconstruction is then detailed (§4.2), where various corrections to the data are outlined and the calculation of event parameters is summarised. The selection of events is then discussed (§4.3) and the selection cuts are applied to obtain the charged current signal. Finally, the selected events are binned according to their values of Q^2, x and y and the cross sections calculated (§4.4).

4.1 Monte Carlo simulation

Monte Carlo (MC) generation is used to make large simulated data samples. These are used to estimate the detector acceptance and resolution (which are too complex for analytical calculation) and also to estimate the number of background events passing the selection cuts (see §4.3). From the measured numbers of real data events and estimates of the acceptance, resolution and background contributions from the MC, the cross sections are calculated.

The MC events are generated and passed through a simulation of the detector (MOZART[2]) and trigger (CZAR[2]). Events are then reconstructed using the same code as is used for the real data events (ZEPHYR[2]), and finally the same selection cuts are made for both MC and data.

This procedure is performed both for signal (charged current) MC and for background MC. For all MC events the "true" event information is known, which is free of biasses introduced by the measurement process itself. In addition to the "true" event information, the measured quantities are available for comparison.

[2]More information can be found on these packages in [16]

This makes it possible to estimate the acceptance, resolution and background contamination of the final sample and thus calculate the cross sections.

MC is also used to estimate many of the systematic uncertainties in the final cross section measurements.

4.1.1 Charged Current DIS Monte Carlo

The generation of CC DIS events is carried out in a three stage process;

- the hard lepton-nucleon scattering process
- QCD cascades (showering)
- hadronisation

The hard lepton-nucleon scattering process is simulated using DJANGOH 1.1 [41], with CTEQ5D [42] parton distribution functions (PDFs). DJANGOH is actually an interface between two programs; HERACLES 4.6.1 [43,44] and LEPTO 6.5 [45] (optionally with ARIADNE 4.10 [46] to perform QCD cascades). The electroweak corrections to allow for initial and final state radiation, two boson exchange, and propagator corrections (see chapter 1) are simulated using HERACLES, whilst LEPTO and ARIADNE simulate the hard scattering process and QCD cascades respectively.

ARIADNE was chosen for this analysis because it gives a better description of hadronic final states than the MEPS model contained in LEPTO [47,48]. MEPS describes the hard scatters using first order matrix elements, and the higher order corrections are described by the parton shower model [49]). ARIADNE employs the "Colour Dipole Model" (CDM), where quark, antiquark pairs ($q\bar{q}$) are treated as a colour dipole. This means the struck quark and the proton remnant are viewed as a dipole, which can radiate gluons. The radiated gluons themselves have colour, and the combination $q\bar{q}g$ are considered to be two independent dipoles (qg and $\bar{q}g$). Gluons may then be radiated from this combination of two dipoles, and so on, until the parton virtuality drops under a threshold, whereupon hadronisation begins. Boson-gluon fusion is included in ARIADNE as a separate process, since the CDM can not describe it. The absence of a struck quark means there is no initial $q\bar{q}$ dipole to radiate.

Although ARIADNE was used in this analysis, as a systematic check a MC sample was also generated using MEPS in order to quantify the uncertainty introduced by the choice of parton shower model (see §4.4.4.3).

The result of the QCD cascade (via ARIADNE or MEPS) is a set of coloured quarks and gluons. Hadronisation is then performed using JETSET 7.4 [50]. This uses the Lund string model to create colourless hadrons which are detected by the experiment.

At all stages, generation of events was performed using electroweak parameters from the Particle Data Group (PDG) [51].

To ensure adequate events were generated at high Q^2, several samples of charged current MC were generated. Each sample had kinematic cuts imposed at the generator level as described in table 2. The MC sample at the lowest Q^2 corresponds to 9.1 times the integrated data luminosity.

MC sample Q^2_{gen} cut	No. of events generated,N_{gen}	$\sigma(pb)$	$L(pb^{-1})$
$Q^2_{gen} > 10\,\mathrm{GeV}^2$	24990	45.202	553
$Q^2_{gen} > 100\,\mathrm{GeV}^2$	25000	39.774	629
$Q^2_{gen} > 5000\,\mathrm{GeV}^2$	15000	3.200	4688
$Q^2_{gen} > 10000\,\mathrm{GeV}^2$	4995	0.683	7313
$Q^2_{gen} > 20000\,\mathrm{GeV}^2$	5000	0.0619	80775

Table 2: *The cuts imposed on the various MC samples at generator level. σ is the cross section, and L is the equivalent integrated luminosity.*

4.1.2 Background Monte Carlo

The main physics processes contributing to the background in the selected events were modelled in MC, and put through the selection process in order to quantify the level of contamination in the final event sample.

4.1.2.1 Neutral Current DIS

NC events (see figure 19) are an expected background to the CC DIS signal. If the final state is poorly measured it may have large missing transverse momentum. In addition, cross sections for NC DIS are typically orders of magnitude larger than the CC cross sections. The NC samples used in this analysis were generated using DJANGOH with ARIADNE, and more detail is provided in table 3.

4.1.2.2 Photoproduction

Photoproduction (PHP) involves the exchange of a quasi-real ($Q^2 \simeq 0$) photon between positron and proton. Again, if the final state products are mismeasured, these events can appear to have missing transverse momentum. Consequently these events can pass the charged current selection criteria, and form a background to the signal.

There are two lowest order PHP Feynman diagrams, which are labelled “Direct” and “Resolved”, as illustrated in figure 20. Direct PHP occurs when the exchanged photon participates in the hard scatter, producing two jets plus the proton remnant. In Resolved PHP, the photon acts as a source of partons, and one of these partons interacts with the proton. In this case, only a fraction of the photon momentum participates in the hard scatter, and the end products include two jets and the photon and proton remnants.

The PHP events were required to satisfy some cuts in order to increase the efficiency of generation. The effectiveness of the cuts was determined by studies, such as those reported in [52]. Either the total transverse energy was required to exceed 30 GeV; or the total transverse momentum was required to exceed 6 GeV. In addition, the events were required to satisfy $Q^2 < 4\,\mathrm{GeV}^2$. These cuts reject a large fraction of the generated events, and so the cross sections provided by the generator must be multiplied by the selection efficiency to obtain the effective cross section for the PHP processes including cuts. The effective cross sections are listed in table 3. The imposed generator level cuts introduce a systematic uncertainty into the overall normalisation of the PHP sample, which is investigated in §4.4.4.6. Both direct and resolved PHP samples used in this analysis were generated with HERWIG 5.9 [53].

MC sample	No. of events generated, N_{gen}	$\sigma(pb)$	$L(pb^{-1})$
Direct PHP $E_T > 30\,\mathrm{GeV}$ or $P_T > 6\mathrm{GeV}$	1.2×10^6	2800	429
Resolved PHP $E_T > 30\,\mathrm{GeV}$ or $P_T > 6\mathrm{GeV}$	3.6×10^6	11900	303
NC DIS $Q^2_{gen} > 40\,\mathrm{GeV}^2$	99970	26260	3.81
NC DIS $Q^2_{gen} > 100\,\mathrm{GeV}^2$	939419	8122.0	115.7
NC DIS $Q^2_{gen} > 400\,\mathrm{GeV}^2$	111785	1167.7	95.7
NC DIS $Q^2_{gen} > 1250\,\mathrm{GeV}^2$	49911	197.54	252.7
NC DIS $Q^2_{gen} > 2500\,\mathrm{GeV}^2$	23995	58.915	407.3
NC DIS $Q^2_{gen} > 5000\,\mathrm{GeV}^2$	23990	14.846	1616
NC DIS $Q^2_{gen} > 10000\,\mathrm{GeV}^2$	23926	2.794	8563
NC DIS $Q^2_{gen} > 20000\,\mathrm{GeV}^2$	23946	0.310	77245
NC DIS $Q^2_{gen} > 30000\,\mathrm{GeV}^2$	11990	0.0544	220404
NC DIS $Q^2_{gen} > 40000\,\mathrm{GeV}^2$	12000	0.0109	1100917
NC DIS $Q^2_{gen} > 50000\,\mathrm{GeV}^2$	11990	0.00212	5655660

Table 3: *Summary of the background MC samples used, with the cuts imposed on the various MC samples at generator level. σ is the cross section, and L is the equivalent integrated luminosity.*

4.1.3 Event weights and vertex distribution

The MC events generated in each sample (see tables 2 and 3) are assigned a weight, so that when multiplied by the luminosity of the measured data, the expected number of observed events is obtained. The vertex distribution used in the MC was set equal to a measurement made using an unbiassed sample of low Q^2 NC events in the 1999-2000 running period, as described in [54]. The weight assigned to an event which is generated at Q^2_{gen} is determined as follows:

$$w = \left(\sum_{Q^2_{cut}=10,100,5k,10k,20k}^{Q^2_{cut}<Q^2_{gen}} \frac{N_{gen}(Q^2_{cut})}{\sigma(Q^2_{cut})} \right)^{-1} \tag{27}$$

where the sum runs over samples for which Q^2_{cut} is less than the event's Q^2_{gen}.

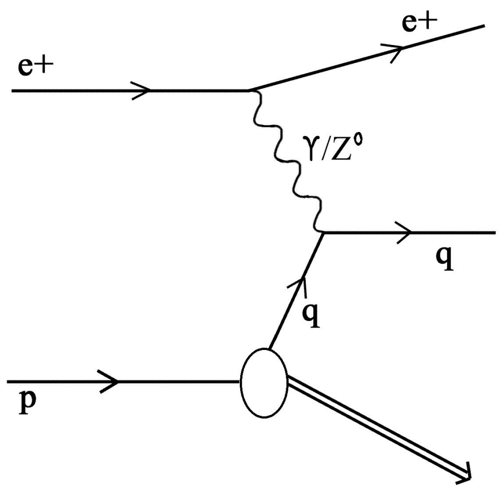

Figure 19: *Feynman diagram of* e^+ *P neutral current DIS at lowest order.*

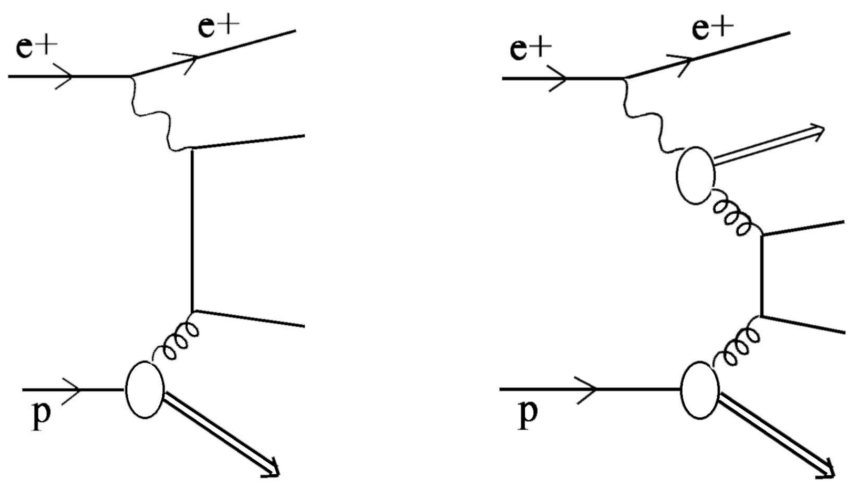

Figure 20: *Feynman diagrams of direct (left) and resolved (right) photoproduction.*

4.2 Event reconstruction

The kinematic variables Q^2, x, y (defined in chapter 1), and the primary interaction vertex must be derived from the event information in the detector. In addition, variables used for event selection must be corrected for a number of detector effects, such as noise, dead material and other energy corrections. The fully reconstructed event information is used in the offline selection process (see sections §4.3.3..4.3.5).

4.2.1 The Jacquet-Blondel method

The kinematic variables of charged current events must be determined from the hadronic information, since the neutrino is not detected. A key quantity used in the determination of these variables is the *missing transverse momentum.* This is defined as follows:

$$\vec{\not{P}}_T \equiv \left(\sum_i -p_x^i \ , \quad \sum_i -p_y^i \right) \tag{28}$$

where the sums run over all the detected final state particles, and the values p^i are the detected momenta of each particle i resolved along the specified direction. Usually, we require information on the magnitude of the missing transverse momentum, which is defined as:

$$\not{P}_T \equiv |\vec{\not{P}}_T| \equiv \sqrt{\left(\sum_i -p_x^i\right)^2 + \left(\sum_i -p_y^i\right)^2} \tag{29}$$

In addition the quantity

$$\delta \equiv \sum_i \left(E^i - p_z^i\right) \quad \left(= \sum_i \left[E^i - E^i cos\theta_i\right]\right) \tag{30}$$

is also needed. E^i and θ_i represent the measured energies and polar angles of the detected final state particles respectively. δ changes little when particles are lost through the forward beam pipe, since for these particles $E^i \simeq p_z^i$. Also, it is worthwhile noting that for a perfect detector which was able to detect the final state antineutrino, and neglecting the positron and proton masses, $\delta = \delta_{perf}$ would equal twice the incident positron beam energy $2 \times E_e$, since by conservation of 4-momentum we obtain

$$\delta_{perf} = E_e - (-E_e) + E_p - E_p = 2E_e. \tag{31}$$

In NC DIS (where the absence of an antineutrino means that the final state particles are detectable), the value of δ calculated using the final state products does indeed peak at twice the incident positron beam energy [55], though the peak is smeared by detector effects and processes such as initial state radiation (see chapter 1). For e^+p CC DIS, the observed δ is essentially zero with a tail, due to the escape of the undetected antineutrino.

The Jacquet-Blondel estimators of Q^2, x, y are then given by [56]:

$$Q^2_{JB} = \frac{\not{P}_T^2}{1 - y_{JB}} \tag{32}$$

$$y_{JB} = \frac{\delta}{2E_e} \tag{33}$$

$$x_{JB} = \frac{\not{P}_T^2}{s y_{JB}(1 - y_{JB})} \tag{34}$$

where s is the square of the centre of mass energy, as described in §1.2. Additionally, the hadronic angle γ_h is calculated as follows:

$$\cos\,\gamma_h = \frac{\not{P}_T^2 - \delta^2}{\not{P}_T^2 + \delta^2}. \tag{35}$$

4.2.2 Calorimeter noise, hot cells, holes and sparks

To successfully reconstruct events, noise and background caused by the ZEUS detector itself must be removed. For instance, the uranium in the calorimeter is radioactive. This and other noise, originating from the photomultiplier tubes (PMTs), or even the signal processing equipment must be removed where possible. Initially, a cell energy cut of 60 (110) MeV is applied to suppress noise in the EMC (HAC) sections. This threshold is set higher, to 100 (150) MeV for isolated cells in the EMC (HAC). These thresholds were determined to correspond to approximately four standard deviations of the noise distributions [57].

Occasionally, a photo-multiplier tube (PMT) in the CAL fires, resulting in a “spark” which mimics a large signal. Fortunately, each cell of the CAL is monitored by two PMTs, so sparks can generally be rejected by analysing the imbalance in the two readings. Noise studies [58, 59] have shown that removing cells where $|E_{PMT1} - E_{PMT2}| > 0.49 \times |E_{PMT1} + E_{PMT2}| + 0.03\,(\text{ GeV})$ is effective in identifying sparks.

A list of “hot” cells is mantained by the calorimeter group. These cells fire more frequently and measure more energy than their peers in randomly selected “passthrough” events. Hot cells are removed from the event information if the energy recorded is not sufficiently large (a threshold is set for each cell, [59]).

In the special case where one of the PMTs monitoring the cell is dead, the cell is labelled as “bad”. The reconstruction software uses the list of “hot” cells and the isolation of the cell either to set the dead PMT energy equal to the live one or to remove the cell.

Finally, when both PMTs are dead, no information can be gleaned from that cell. This is known as a “hole”, and any information for this cell is ignored.

4.2.3 Clustering of energy deposits

In event reconstruction, it is advantageous to combine individual calorimeter cells so as to produce objects which resemble as closely as possible the final state particles or jets. Several different methods of cell combination are used in the reconstruction and selection:

4.2.3.1 Condensates

Condensates are formed by gathering hit calorimeter cells, and ordering them in decreasing energy. The cell with the highest energy is used as a seed for a condensate, and cells which share a common side are added from the list of free cells if they have been hit. The process continues until the condensate stops growing, and the cells in the condensate are removed from the list of available cells. The new list of available hit cells is used and the cell with the highest energy in the new list is used as a seed for the next condensate. The position of a condensate is equal to the energy weighted average of the cell locations, its energy is simply the sum of the constituent cells' energies.

Condensates have a number of disadvantages, including the fact that cells diagonally adjacent to a hit cell are not automatically included.

4.2.3.2 Islands

An improvement on the condensates algorithm is achieved by using a two stage process to form "islands" of energy in the calorimeter [60]. Initially, "cell islands" or groups of cells in the same layer (HAC1,HAC2,EMC) are formed. This is done by connecting a cell to the highest energy neighbouring cell if the neighbour has a greater energy (adjacent and diagonally adjacent cells are included in this step). The cell islands in different layers are then combined into "cone islands". The formation of cone islands starts in the outermost layer of the calorimeter (HAC2) and works inwards, combining HAC2 islands with HAC1 and EMC islands and then HAC1 islands with EMC islands.

4.2.4 Calorimeter energy corrections

To minimise biasses, and keep the detector resolution as high as possible, the energies measured by the calorimeter cells must be accurately calibrated.

Two routines, RCALCORR v6.3 [61] and CORANDCUT [62] are used to make these corrections.

4.2.4.1 RCALCORR corrections

Initially, the RCALCORR v6.3 [61] routine is employed to perform electromagnetic energy scale corrections to calorimeter cells. Measured BCAL energies are scaled up by 5%. In the RCAL, the 44 inner cells (nearest the beampipe) were assigned individual correction factors ranging from +13% to −5%, whilst the outer cells were scaled up by a common factor of 2.2%. These factors were determined in [63], using 96/97 neutral current data.

4.2.4.2 CORANDCUT corrections

The CORANDCUT [62] routine is then used to make several calorimeter corrections to enable the hadronic final state to be reconstructed:

Dead material correction

Energy lost in the dead material interposed between the vertex and calorimeter is corrected for in the reconstructed cone islands. The correction is based on a detailed map of the detector and Monte Carlo studies [64]. Electromagnetic and hadronic islands (defined according to the fraction of energy measured in the EMC and HAC) are corrected for separately when the cone island has an energy less than 7 GeV. Below this value, the hadronic island energy was found to be overestimated. In [62] and in low energy beam tests [65], it is proposed that this effect occurs due to ionisation losses of low energy hadrons in the calorimeter, and to the different zero suppression cuts in the EMC and HAC sections.

Particles crossing the boundaries between FCAL, BCAL and RCAL sections are also corrected for the "super-crack" effect. In the boundary regions, particles lose energy in support structures and other uninstrumented material. The corrections applied for this effect were deduced using events with a single narrow hadronic jet to probe the cracks.

The resulting overall correction applied to each island was thus a function of the amount of dead material, and also of the type and energy of the island.

Backsplash

Following the dead material correction, another correction is made for the "backsplash" effect. This manifests itself as small energy deposits in the RCAL and rear half of the BCAL and is caused by a particle travelling in the forward direction scattering off inactive material. It can lead to an over estimation of the hadronic angle γ_h, for events with low y where most of the hadronic energy is deposited in the forward region. The method described in [62] is used to remove these deposits.

Low energy cone islands ($< 3\,\mathrm{GeV}$) with polar angles above a threshold, γ_{max} (which is itself dependent on the value of the hadronic angle γ_h), are removed from the event, and γ_h recalculated. Then, a new value of γ_{max} is determined from this value of γ_h, and more islands removed if their polar angles are greater than γ_{max}. Typically the process ends after 2 or 3 iterations, as no more low energy islands are detected. The value of $\gamma_{max}(\gamma_h)$ was determined from neutral current Monte Carlo so that less than 1% of non-backscatter islands are excluded. CORANDCUT uses the following values (for 920 GeV protons):

$$\gamma_{max} = 0.149 + 1.369 \times \gamma_h \quad (\gamma_h < 1.9) \tag{36}$$

$$\gamma_{max} = 2.750 + 0.328 \times (\gamma_h - 1.9) \quad (\gamma_h > 1.9) \tag{37}$$

Calorimeter hadronic energy scale

After the application of noise suppression, the RCALCORR energy rescaling, the dead material and backsplash corrections, there is still a discrepancy between data and Monte Carlo in the hadronic final state energy.

The hadronic scale and uncertainty were determined in [66,67], using a sample of neutral current DIS events. In neutral current events, the scattered lepton is still present, and this allows measurement of the kinematic variables Q^2_{DA}, x_{DA}, y_{DA} using the "double angle" method [68,69], which is independent of the measured energy. The double angle transverse momentum can be obtained using these kinematic variables ($P_{T,DA} = \sqrt{Q^2_{DA}(1 - y_{DA})}$), and compared to the measured hadronic transverse momentum $P_{T,had}$.

By forming the ratio $P_{T,had}/P_{T,DA}$ for the data and Monte Carlo, and plotting it against the fraction of the island's energy measured in the EMC, the analysis determines the correction factors and uncertainties.

To allow measurements which were (as far as possible) independent for the FCAL and BCAL, single jet neutral current events, where the jet was localised to only one of the two calorimeters, were used for the analysis. The results of the correction can be seen in figure 21.

Hadronic activity in the RCAL is typically very low, so to ascertain the hadronic scale and uncertainty for this portion of the calorimeter, "diffractive" events with hadronic activity exclusively in the RCAL were used in [66]. The same procedure as described above was applied to determine the correction factor and uncertainty for the RCAL.

The correction factors used are listed in table 4. After applying hadronic energy scale corrections, the remaining scale uncertainty in BCAL and FCAL is 1%, and for RCAL it is 2%. These uncertainties form part of the systematic checks performed in §4.4.4.2.

	RCAL	BCAL	FCAL
EMC	1.000 ± 0.02	1.003 ± 0.01	1.024 ± 0.01
HAC	1.000 ± 0.02	1.044 ± 0.01	0.941 ± 0.01

Table 4: *The factors used for the hadronic energy scale correction (determined in [66]).*

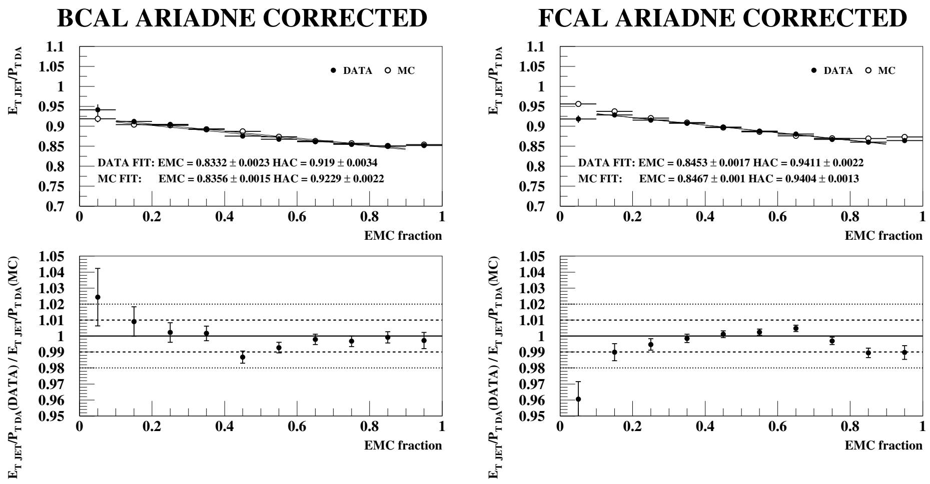

Figure 21: *Determination of the hadronic energy scale for the FCAL and BCAL. The corrected ratio of transverse jet momentum to the prediction of the double angle method (top). The ratio of data to MC (bottom) allows determination of the uncertainty. Reproduced from reference [67].*

4.2.5 Vertex reconstruction

Tracks and vertices are reconstructed off line by the VCTRAK package, which is described in [70, 71]. Track reconstruction was performed in "CTD-only" mode, meaning that the forward and rear tracking detectors (FTD,RTD) were not used. This avoids several problems, such as detector alignment inaccuracies, and possible charged particle production between detectors.

Reconstructed tracks were used to find the primary interaction vertex, and also secondary vertices (which may be caused by particle decay or secondary interaction with the beam pipe). This is the so-called "Multi-Vertex" algorithm. The process begins by testing pairs of tracks[3] to see if they are compatible with a common vertex. Pairs formed are then joined into "clusters" of tracks, which again consist of tracks loosely compatible with a common vertex. Tracks are assigned to as many different clusters as possible. A fit is then performed on each cluster to determine the common vertex, and tracks dropped if they contribute too much to the vertex χ^2.

[3]tracks must have superlayer 1 hits in the CTD

The vertices of surviving clusters are treated as possible primary or secondary vertex candidates.

The initial tracks are then used to find the best possible candidate for the primary vertex. Each track is used as a "seed" which is assumed to originate from the primary vertex. In this step, the proton beam line ($x = y = 0, \sigma_x = \sigma_y = 0.7\,\mathrm{cm}$) is used as a virtual point on the track, and all tracks found to be compatible with the primary vertex being tested are assigned to the candidate.

Tracks are then dropped from vertices until no track belongs to more than one vertex, and the tracks assigned to secondary vertices point away from the primary vertex.

The candidate is then studied, along with the secondary vertex candidates and a score assigned, which is based upon items which include the number and length of tracks originating at each vertex, and the proximity of the primary vertex to the $x = y = 0$ beam line.

Finally, a vertex "full fit" [72] is performed on the highest scoring configuration using the tracks, and both primary and secondary vertices, which provides accurate vertex positions and track momenta.

For charged current interactions, the primary vertex z-position reconstructed using VCTRAK was found to have a resolution of $\sim 1\,\mathrm{cm}$.

The azimuthial hadronic angle calculated with respect to the nominal vertex position is defined as γ_0. It is used to divide the charged current sample into two types. Those with $\gamma_0 \geq 0.4$ radians ($\sim 23^o$) lie within the geometrical CTD acceptance, and as such the VCTRAK routine can be successfully used to reconstruct the primary vertex position. For events with $\gamma_0 < 0.4$, typically little or no tracking information exists and so the vertex can not be accurately reconstructed using the CTD. For these events, the FCAL timing method [73] is used. This involves timing the energy deposits in the FCAL to estimate the z-coordinate of the interaction vertex. Neutral current DIS events were used to calibrate the timing on a run-by-run basis. The resolution of the z-vertex using this method was found to be approximately $\sim 10\,\mathrm{cm}$ for the low γ_0 charged current sample.

Figure 22 shows that events with a low hadronic angle populate the high-x low-y region of phase space, and as such extend the region over which the cross section measurements can be made.

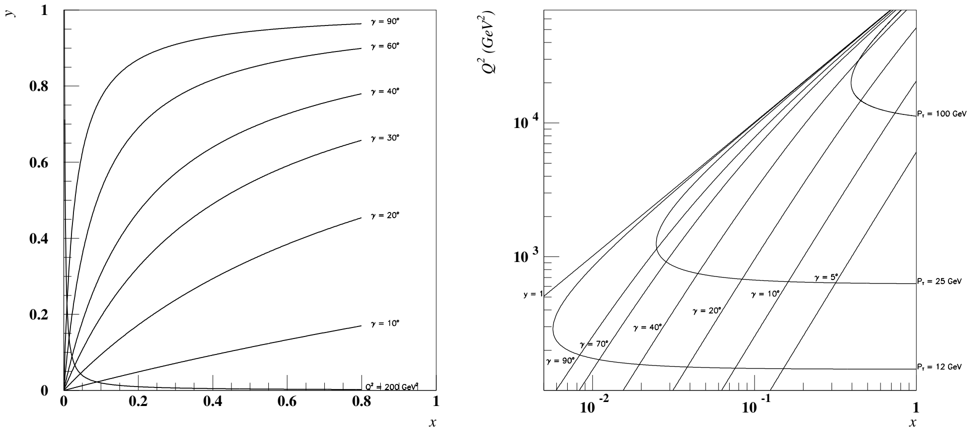

Figure 22: *Plot showing isolines of the hadronic angle in the $x - y$ plane (left). The transverse momentum and hadronic angle in the $x - Q^2$ plane (right).*

4.3 Event selection

The characteristic experimental property of charged current DIS events is missing transverse momentum $\not{P}_T$, in the calorimeter. This occurs due to the escape of the final state neutrino. The high solid-angle coverage of the ZEUS detector (99.7%) means that it is particularly well suited to measuring $\not{P}_T$, and the charged current trigger is primarily designed to select events based on this quantity. Unfortunately, a considerable amount of background is also included with this selection. In fact, after the trigger selection, the initial event sample is dominated by background events. It is necessary to remove as much of this background as possible (whilst keeping the signal in the sample), in order to accurately measure the charged current DIS cross sections. We begin the description of event selecton by outlining the backgrounds which we are attempting to remove (§4.3.1). Then, the three stage trigger, and offline selection cuts are detailed (§4.3.2..4.3.5).

4.3.1 Backgrounds

4.3.1.1 Neutral Current DIS

The neutral current (NC) cross section is typically orders of magnitude larger than the charged current cross section. Usually, NC events do not pass the cuts on $\not{P}_T$ because the detected final state products are balanced in terms of transverse momentum, however NC events where some of the final state products are mismeasured or not detected, can result in the apperance of an event with a large $\not{P}_T$. This may be due to the positron or a jet passing through inactive material in the calorimeter, or the mismeasurement of a jet because a sizeable fraction of it was lost in the beam pipe.

4.3.1.2 Cosmic and halo muons

Although the ZEUS detector is 30 m underground, muons produced in the upper atmosphere frequently reach it, giving rise to top-to-bottom tracks through the detector and often giving rise to events displaying large $\not{P}_T$. Similarly, large $\not{P}_T$ events are also generated by muons produced upstream of the detector, inside HERA. Upstream interactions between the particle beams and the accelerator components or residual beam gas, give rise to muons travelling horizontally through the detector in the direction of one of the beams.

4.3.1.3 Beam gas events

Inside the detector, the beams can interact with residual gas in the beam pipe, and sometimes with the beam pipe itself. The level of contamination introduced by these events can, in theory, be deduced by using events from 'pilot' bunches, where one of the beams has no corresponding particles in the other beam to collide with. Typically, however, these events have high track multiplicities. They also deposit large amounts of energy in the ring of FCAL towers closest to the beam pipe (the innermost ring).

4.3.1.4 Photoproduction events

As described in §4.1.2.2, there are two types of Photoproduction (PHP) events. Both direct and resolved PHP can mimic CC signal when the event products are undetected in a crack region of the calorimeter, or when a jet is partially lost down the beam pipe.

4.3.2 Trigger selection

The first stage of background removal occurs in the trigger selection. The ZEUS trigger is described in §2.2.4, and consists of a three stage selection process. For charged current candidate events, trigger slots in the first, second and third level triggers (F,S,TLT) are used. Candidate events must pass at least one slot in each level of the trigger.

Trigger selections are primarily based on measured energies and momenta, but it is important to note that these quantities are not identical to those used in the offline selection. In the FLT for example, they are calculated using only calorimeter towers, not islands (see §2.2.4). They also assume a nominal vertex position. None of the corrections made to the offline quantities (outlined in §4.2) are made at this stage. To remind the reader of the difference between trigger and offline variables, the trigger quantities are labelled with the superscripts F,S,TLT.

4.3.2.1 First Level Trigger

Two trigger slots aimed at the selection of CC events were used. They both primarily select events with large missing transverse momentum.

- slot 60

($\not{P}_T^{FLT} > 5\,\text{GeV}$ .and. $E_T^{FLT}(-2ir) > 5\,\text{GeV}$ .and. CTD tracks ≥ 1.)
.or.
($\not{P}_T^{FLT} > 8\,\text{GeV}$ .and. ($E_{FCAL}^{FLT} > 10\,\text{GeV}$ .or. CTD tracks ≥ 1)).

- slot 61

$\not{P}_T^{FLT} > 3\,\text{GeV}$ .and. $E_{FCAL}^{FLT}(ir) < 1.25\,\text{GeV}$ .and. CTD tracks ≥ 1

where E_T, the total transverse energy is defined as:

$$E_T = \sum_i \sqrt{(p_x^i)^2 + (p_y^i)^2} \tag{38}$$

and the sum runs over calorimeter cells. $E_T^{FLT}(-2ir)$ is the transverse energy measured in the FLT, excluding the calorimeter cells forming the inner two rings of the FCAL, and $E_{FCAL}^{FLT}(ir)$ is the energy measured in the inner ring of the FCAL. It should be noted that slot 60 allows both low and high-γ_0 candidate events to pass, since if no CTD tracks are discovered (low-γ_0 events), a correspondingly large energy deposit in FCAL should be observed, together with the missing transverse momentum.

To supplement the selection efficiency, events passing at least one of four FLT slots based on CAL energy are admitted. This also allows an independent cross check on the trigger efficiency, as described in §4.4.4.8.

- slot 41

$E_T^{FLT} > 21\,\mathrm{GeV}$

- slot 42

($E^{FLT} > 15\,\mathrm{GeV}$ .or. $E_{EMC}^{FLT} > 10\,\mathrm{GeV}$ .or. $E_{BEMC}^{FLT} > 3.4\,\mathrm{GeV}$ .or. $E_{REMC}^{FLT} > 2\,\mathrm{GeV}$) .and. CTD tracks ≥ 1.

- slot 43

$E_T^{FLT} > 11.5\,\mathrm{GeV}$ .and. CTD tracks ≥ 1.

- slot 44

$E_{BEMC}^{FLT} > 4.8\,\mathrm{GeV}$ .and. ($E_{REMC}^{FLT} > 3.4\,\mathrm{GeV}$ .or. CTD tracks ≥ 1).

where EMC, BEMC and REMC refer to the electro-magnetic calorimeter energies of the whole CAL, BCAL and RCAL respectively.

4.3.2.2 Second Level Trigger

The SLT has timing information available to it. This is used to remove events in which the calorimeter hits do not occur in a manner consistent with an e^+p

interaction. One slot, EXO-SLT-4, is used to distinguish CC events from background. Its logic is complex, but in a similar manner to the FLT, this SLT slot allows both low and high-γ_0 candidate events to pass, by using either CTD tracks or energy deposits in the FCAL together with $\not{P}_T$. Its logic can be summarised by:

- EXO-SLT-4

($|t_G| < 7$ ns)
.and.
(($\not{P}_T^{SLT} > 6\,\text{GeV}$ .and. $E_T^{SLT}(-2ir) > 6\,\text{GeV}$ .and. a vertex fitted CTD track)
.or.
($\delta > 6\,\text{GeV}$ .and. $(\not{P}_T^{SLT})^2 > 2.25 \times E_T^{SLT}$ .and. a vertex fitted CTD track).
.or.
($\not{P}_T^{SLT} > 9\,\text{GeV}$ .and. $\not{P}_T^{SLT}(-ir) > 8\,\text{GeV}$ .and. $E_{FCAL}^{SLT} > 20\,\text{GeV}$)
.or.
($\not{P}_T^{SLT} > 9\,\text{GeV}$ .and. $(\not{P}_T^{SLT})^2 > 2.31 \times E_T^{SLT}$ .and. $E_{FCAL}^{SLT} > 80\,\text{GeV}$))
$|t_G|$ is the averaged global CAL timing, and is cut off at 7 ns to be consistent with CC events as opposed to background ($t_G = 0$ is calibrated to the e^+p collision time, such that particles travelling at the speed of light from the nominal interaction point are sampled at each cell locally at t=0 [74]).

4.3.2.3 Third Level Trigger

Full event information is available to the TLT. Two slots designed to capture CC events are used and any CC candidate events are required to pass at least one of these slots. Both use tighter timing cuts, designed to remove cosmic muons passing from top-to-bottom in the detector. Again, both slots require missing transverse momentum together with CTD tracks or large energy deposits in FCAL (indicating a low-γ_0 event). The slots can be summarised by the following logic:

- EXO-TLT-2

$\not{P}_T^{TLT} > 6\,\mathrm{GeV}$ .and. $|t_u - t_d| < 8$ ns .and. a vertex fitted CTD track found.

- EXO-TLT-6

EXO-SLT-4 .and. $\not{P}_T^{TLT} > 8\,\mathrm{GeV}$ .and. $|t_u - t_d| < 8$ ns .and.
($E_{FCAL}^{TLT} > 10\,\mathrm{GeV}$ .or. a vertex fitted CTD track).

where t_u, t_d are the timings from the upper and lower halves of the BCAL respectively. The $|t_u - t_d|$ cut removes cosmic muon background. EXO-TLT-6 notably requires at least 10 GeV in FCAL if there are no vertex fitted CTD tracks.

4.3.3 Offline selection based on a CTD vertex

Variables, such as $\not{P}_T$ are available in their fully corrected forms (as per §4.2) for the offline selection. This is in contrast to the previous section, where uncorrected trigger variables were used for initial selection purposes. Events with a hadronic angle calculated with respect to the nominal interaction point (γ_0), greater than or equal to 0.4 (i.e. within the geometrical CTD acceptance) are required to have:

4.3.3.1 $\not{P}_T > 12\,\mathbf{GeV}$ and $\not{P}_T(-ir) > 10\,\mathbf{GeV}$

Missing transverse momentum is the characteristic signal of charged current DIS. Figure 23 shows the distribution of $\not{P}_T$ and $\not{P}_T(-ir)$. The cuts applied are shown as hatched regions. Events in these regions are removed because they fail only the cut applied in the plot, but pass all other cuts. The $\not{P}_T(-ir)$ cut is applied to remove background events where large $\not{P}_T$ comes from the beam pipe region of the FCAL.

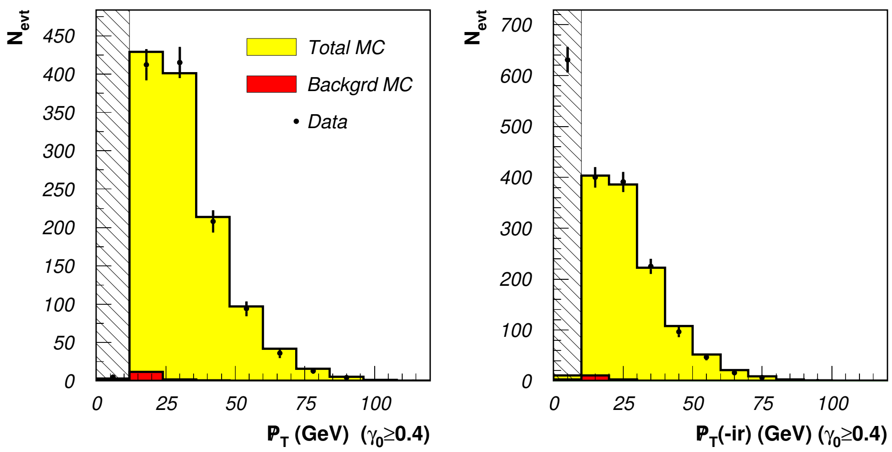

Figure 23: *The distribution of $\not{P}_T$ (left) and $\not{P}_T(-ir)$ (right) for data (closed circles) and total MC (yellow histogram) in the region $\gamma_0 \geq 0.4$. The red histogram shows the contribution of the background MC to the total MC. All other cuts have been applied to the events in these plots. Above the thresholds, good agreement between data and MC is observed.*

4.3.3.2 $\not{P}_T/E_T > 0.55$ for $\not{P}_T < 20$ GeV; and $\not{P}_T/E_T > 0.4$ for $20 \leq \not{P}_T < 30$ GeV

For multi-jet events and events where the end products are not well collimated (such as beam gas and photoproduction), the value of $\not{P}_T$ relative to the total transverse energy will be reduced. This cut is designed to reject these uncollimated events, especially in the lower $\not{P}_T$ region where photoproduction is concentrated.

No cut on $\not{P}_T/E_T$ is applied in the region of $\not{P}_T \geq 30$ GeV, where the background is negligible, and one risks the removal of multi-jet CC events in the high-$\not{P}_T$ region.

Figure 24 shows the distribution of $\not{P}_T/E_T$ for the three regions $\not{P}_T < 20$, $20 \leq \not{P}_T < 30$ and $\not{P}_T \geq 30$ GeV after all other cuts are applied.

4.3.3.3 $|Z_{vtx}| < 50$ cm

Most events originate within ~ 10 cm (in the z-direction) of the nominal interaction point, which is a consequence of the bunch lengths supplied by HERA. Events occurring more than 50 cm from the nominal interaction point are excluded, to ensure accurate reconstruction and to remove background events. The distribution of data and MC events along the z-axis is shown in figure 25.

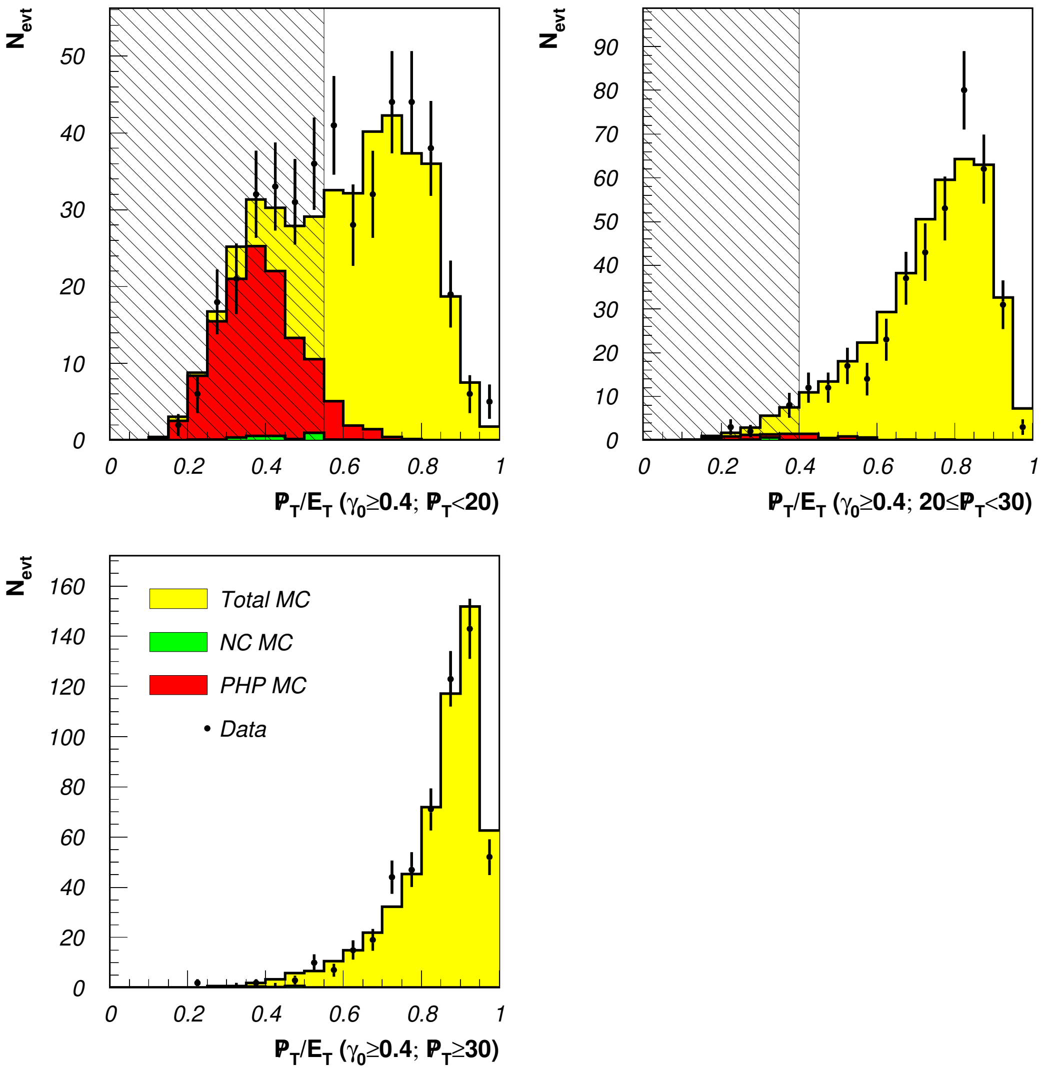

Figure 24: *The distribution of* $\not\!\!P_T/E_T$ *for the three regions* $\not\!\!P_T < 20$, $20 \leq \not\!\!P_T < 30$ *and* $\not\!\!P_T \geq 30$ *GeV after all other cuts are applied. Data events are shown as closed circles, whilst the sum of all MC contributions is shown as the yellow histograms. PHP and NC contributions to the total MC are shown in red and green respectively. The hatched area shows those events that pass all other cuts but are rejected by the* $\not\!\!P_T/E_T$ *cut.*

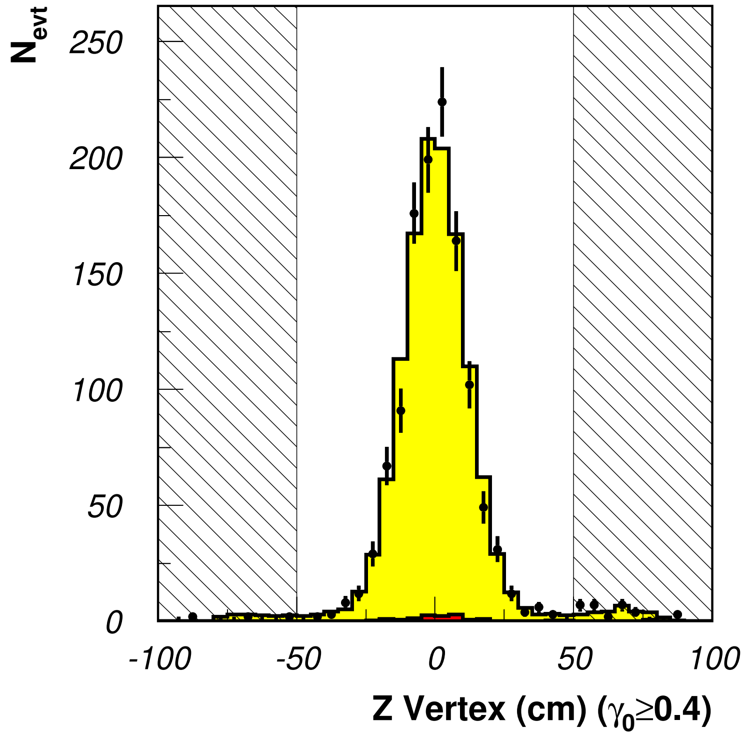

Figure 25: *The distribution of the interaction vertex as measured by the CTD, for events passing all other cuts in the* $\gamma_0 \geq 0.4$ *region. Data events are shown as closed circles, whilst total MC is shown by the yellow histogram. Background contributions to the total MC are shown in red. The hatched area shows those events that pass all other cuts, but are rejected by the* Z_{vtx} *cut.*

4.3.3.4 $N_{good} \geq 1$

N_{good} is the number of good tracks measured in the CTD. A good track is defined as one having the following attributes:

- the distance to the beamline is less than 1.5 cm;
- the track must have a transverse momentum larger than 0.2 GeV;
- the polar angle of the track must be in the range $15 - 164^o$;
- the track must have registered hits in at least two stereo layers of the CTD;
- the track must be associated with the primary interaction vertex by the VCTRAK reconstruction package.

Requiring an event to have at least one good track in the CTD reduces background, such as cosmic/halo muon events. In this case, a muon induced bremstrahlung shower may cause tracks to be reconstructed in the CTD, though they are unlikely to fulfil the criteria of a "good" track. Figure 26 shows the distribution N_{good} in data and MC.

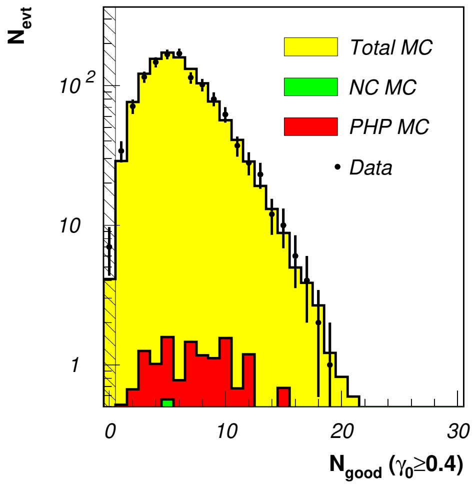

Figure 26: *The distribution of the number of "good" tracks measured by the CTD, for events passing all other cuts. Data events are shown as closed circles, whilst total MC is shown by the yellow histogram. PHP and NC background contributions to the total MC are shown in red and green respectively. The hatched area shows those events that pass all other cuts, but are rejected by the N_{good} cut.*

4.3.3.5 $N_{good} > 0.25 \times (N_{trk}^{tot} - 20)$

The number of good tracks must form a significant portion of the total number of tracks in the CTD, N_{trk}^{tot}. Beam gas and beam pipe interactions typically produce large numbers of poor quality tracks, and these events can be seen in the cut region of the data in figure 27. They are not reflected in the CC MC.

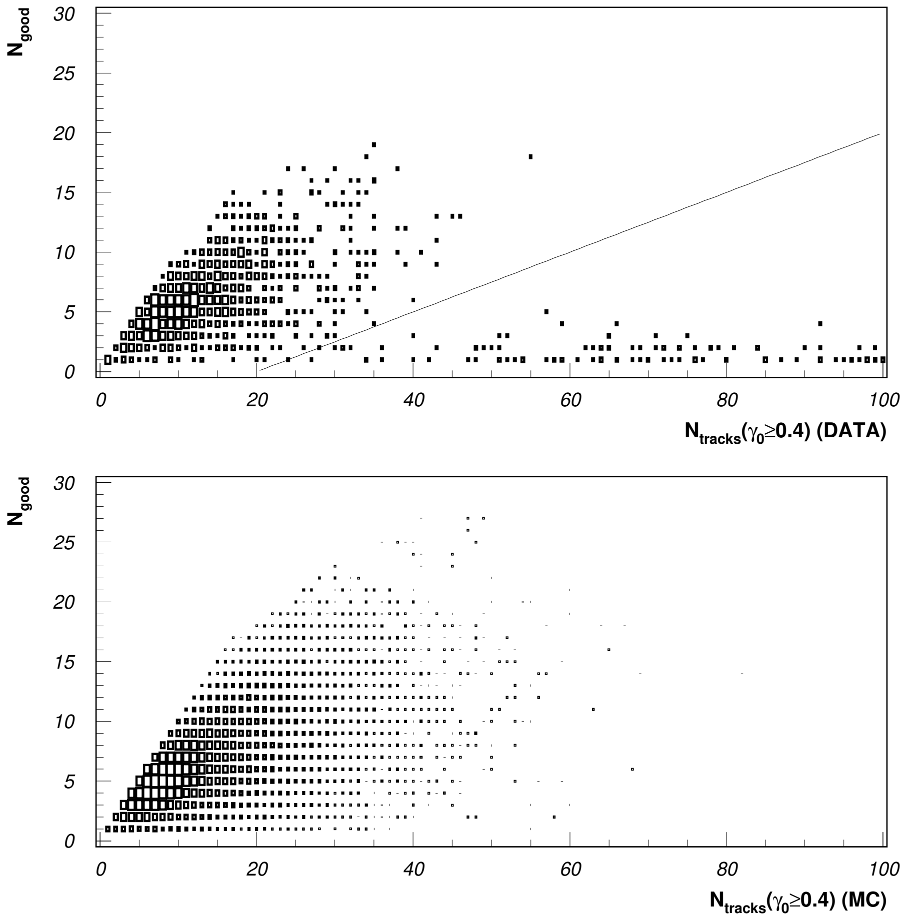

Figure 27: *The distribution of "good" vs. total CTD tracks for data (top) and signal MC (bottom) after all other cuts have been applied. High multiplicity with a relatively low number of good tracks indicates background.*

4.3.3.6 $|\Delta\phi| < 0.5$ ($\not{P}_T < 20\,\mathbf{GeV}$) and $|\Delta\phi| < 2.0$ ($\not{P}_T \geq 20\,\mathbf{GeV}$)

The azimuthial angle of the missing transverse momentum (ϕ), can be calculated using energy deposits in the CAL and separately using the "good"[4] CTD tracks. This cut imposes the requirement that the two angles agree reasonably well. The cut is tighter for events with $\not{P}_T < 20\,\mathrm{GeV}$, as the data in this region show a slight excess over MC. The data events and MC are shown for these two regions in figure 28.

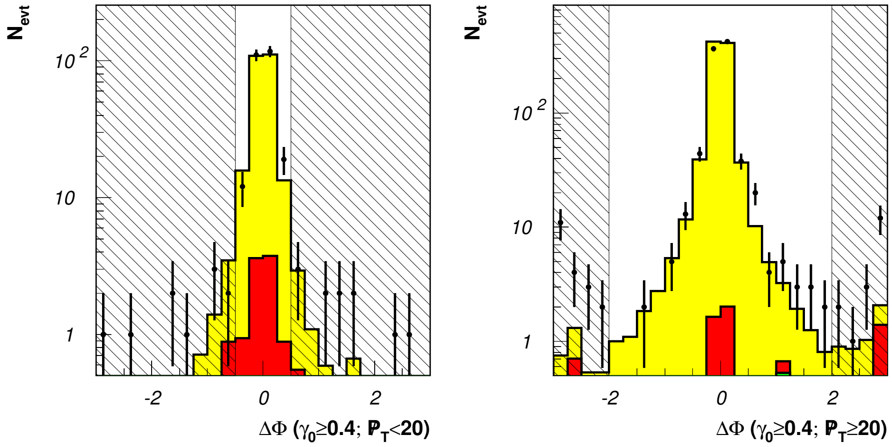

Figure 28: *The difference in azmuthial angle of the $\not{P}_T$ calculated using the CAL and good CTD tracks in the regions $\not{P}_T < 20$ GeV (left) and $\not{P}_T \geq 20$ GeV (right). Data events are shown as closed circles, whilst total MC is shown as the yellow histogram. The PHP background contribution to the total MC is visible as the red histogram.*

[4]Good CTD tracks defined as per §4.3.3.4.

4.3.3.7 Track Imbalance

In photoproduction events, $\not{P}_T$ can be caused by a jet escaping through a crack region of the CAL. In that case, the direction of $\not{P}_T$ is in the direction of the poorly measured jet. Some tracks in these events point in the opposite direction to the $\not{P}_T$, but crucially other tracks point in the direction of the $\not{P}_T$. Charged current signal events should not display tracks pointing in the direction of the $\not{P}_T$.

"Up" and "down" tracks can be defined with respect to the detected momentum. The "up" tracks are those in the *opposite* direction (± 0.5 rad) to the $\not{P}_T$. These tracks are expected to exist for CC signal and background. The "down" tracks are defined as those in the direction (± 0.5) of the $\not{P}_T$, and are not expected to be present for CC signal events. This situation is depicted in figure 29.

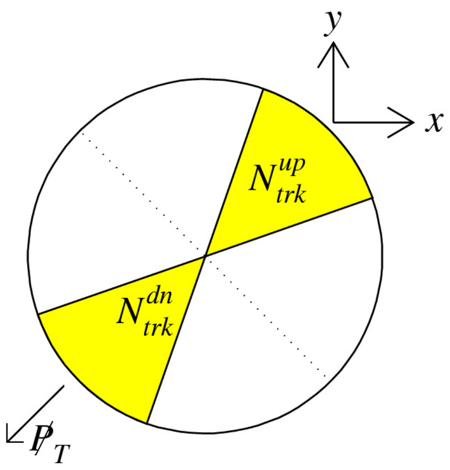

Figure 29: *The distribution of "up" and "down" tracks, together with the direction of the missing transverse momentum.*

The numbers of these tracks ($N_{trk}^{up}, N_{trk}^{dn}$) are then used to select CC signal events, by requiring (for $\not{P}_T < 20$) that:

- $N_{trk}^{dn} \leq 1$
- $Imb_{trk} = (N_{trk}^{up} - N_{trk}^{dn})/(N_{trk}^{up} + N_{trk}^{dn}) > 0.7$

where Imb_{trk} quantifies the track imbalance, which should be large for CC signal events. These two quantities (N_{trk}^{dn} and Imb_{trk}) are shown in figure 30.

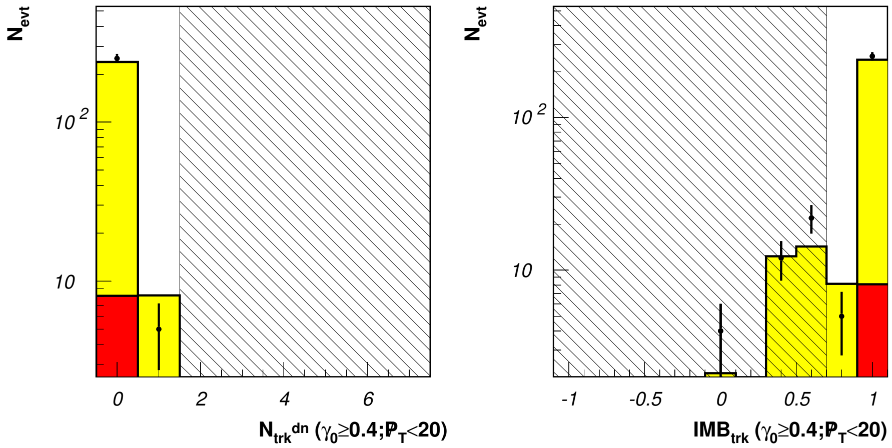

Figure 30: *The distribution of "up" and "down" tracks after all other cuts have been applied. Data events are shown by closed circles, and the total MC by the yellow histograms. The PHP background contribution to the total MC is shown in red.*

4.3.4 Offline selection based on a CAL vertex

Events with γ_0 less than 0.4 are required to pass the following cuts:

4.3.4.1 $\not{P}_T > 14\,\mathbf{GeV}$ and $\not{P}_T(-ir) > 12\,\mathbf{GeV}$

The cuts on missing transverse momentum, and missing transverse momentum minus the FCAL inner ring were set slightly higher for low-γ_0 events, due to the lack of information from the CTD. The distributions of these two quantities are shown in figure 31.

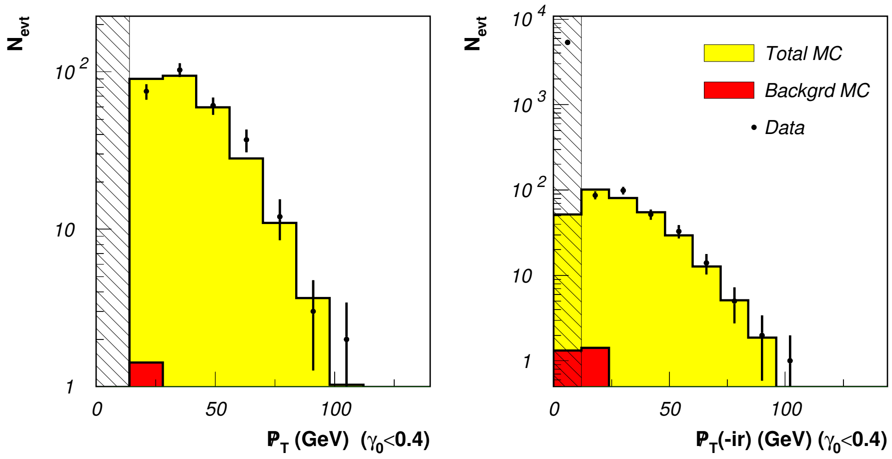

Figure 31: *The distribution of $\not{P}_T$ (left) and $\not{P}_T(-ir)$ (right) for data (closed circles) and total MC (yellow histogram) for $\gamma_0 < 0.4$. The red histogram shows the contribution of background to the total MC. All other cuts have been applied to the events in these plots.*

4.3.4.2 $|Z_{vtx}^{FCAL}| < 50\,\mathbf{cm}$

Events with $\gamma_0 < 0.4$ typically are not able to use CTD track information to reconstruct the interaction vertex accurately. Instead, the FCAL timing information is used to reconstruct the z position of the vertex, as described in [73]. Using this information, the z vertex is required to be within 50 cm of the nominal interaction point. Figure 32 shows the agreement between distributions of Z_{vtx}^{FCAL} for data and MC events after all other cuts have been applied.

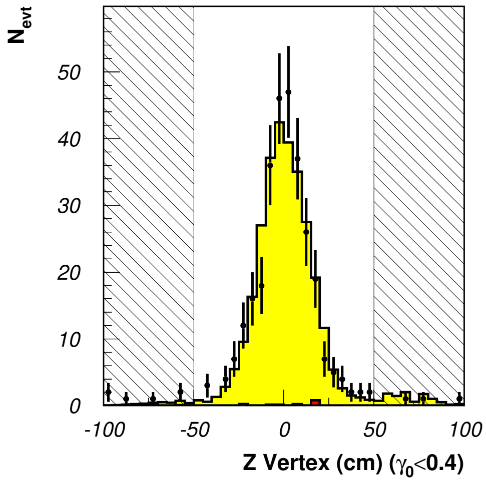

Figure 32: *The distribution of the z vertex obtained using FCAL timing in the region $\gamma_0 < 0.4$ for data (closed circles) and total MC (yellow histogram) after all other cuts have been applied. The PHP background contribution to the total MC is visible as the red histogram, and the hatched region represents events removed by this cut.*

4.3.4.3 FCAL condensate energy fractions

Three cuts were applied to the FCAL, in order to remove background where too much, or too little energy came from one of the three sections (EMC, HAC1 or HAC2) for it to resemble a CC event. Condensates (or clusters of adjacent cells passing an energy cut, see §4.2.3.1) are first isolated in the FCAL, and then the cuts applied to the maximum energy condensate, if its energy exceeded 2 GeV. The events were thus selected if they satisfied the following conditions:

- $E_{FEMC}^{maxc}/E_{FCAL}^{maxc} \geq 0.15$ .and.
- $E_{FHAC1}^{maxc}/E_{FCAL}^{maxc} \geq 0.1$ .and.
- $E_{FHAC2}^{maxc}/E_{FCAL}^{maxc} \leq 0.6$

For $E_{FCAL}^{maxc} > 2$. The distributions of these quantities for data and MC can be seen in figure 33.

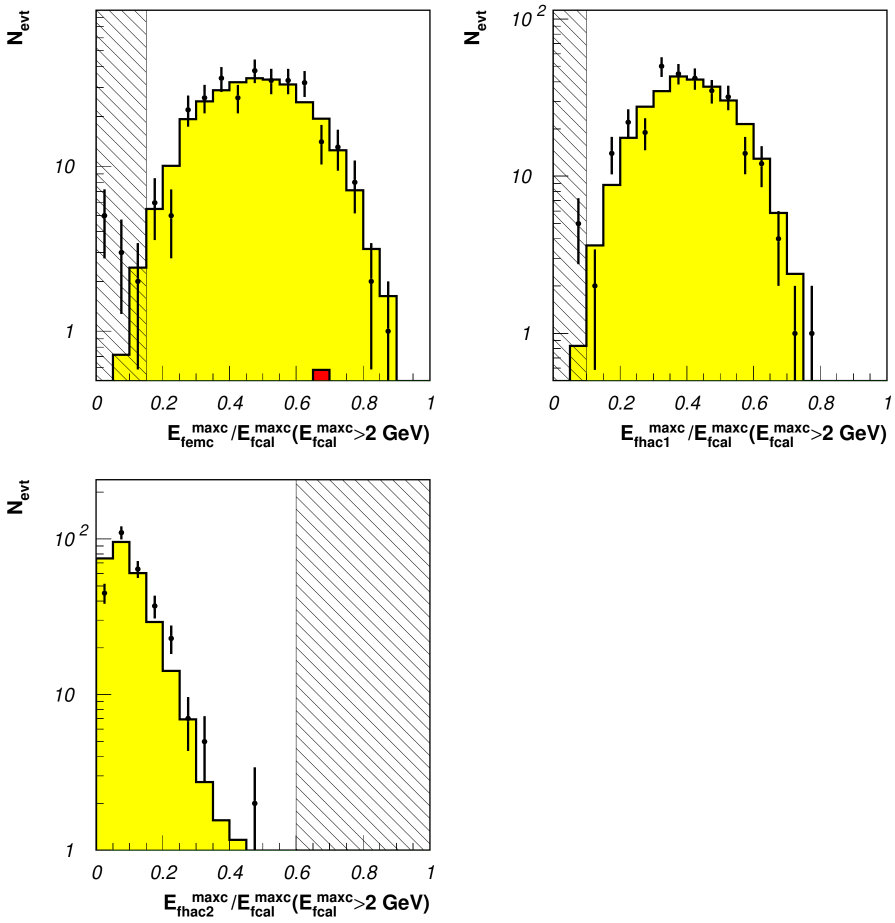

Figure 33: *The distributions of the energy fractions in the FCAL maximum energy condensate for data (closed circles) and total MC (yellow histograms) in the $\gamma_0 < 0.4$ region after all other cuts have been applied. The PHP background contribution to the total MC is visible as the red histogram.*

4.3.5 Kinematic and final cleaning cuts

The final event sample is obtained by applying cuts to limit the kinematic region and remove specific backgrounds. In addition, runs where the data were not suitable for physics analysis (for example where a component was not functional), were removed from the CC selection process and were not included in the overall data luminosity calculation.

4.3.5.1 $Q^2 > 200\,\mathrm{GeV}^2$ and $y < 0.9$

The kinematic region was restricted to events with $Q^2 > 200\,\mathrm{GeV}^2$ and $y < 0.9$. Figure 34 shows the distributions of these two quantities. The y cut is needed to remove background, and because the x and Q^2 resolutions become worse at large y. Using equation 32, we obtain equation 39, which shows that the Q^2 resolution is poor at high y.

$$\frac{\Delta Q^2}{Q^2} = \sqrt{4.\left(\frac{\Delta \not{P}_T}{\not{P}_T}\right)^2 + \left(\frac{\Delta y}{1-y}\right)^2} \tag{39}$$

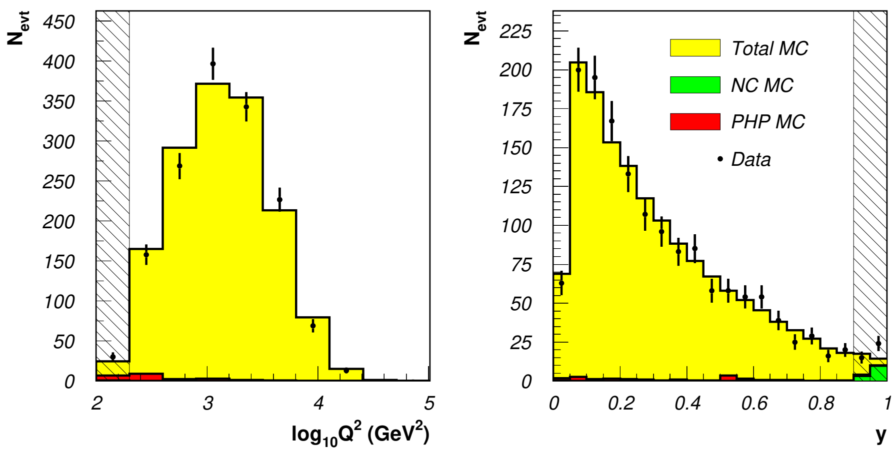

Figure 34: *The distributions of the reconstructed Q^2 and y values for data (closed circles) and total MC (yellow histograms) after all other cuts have been applied. The PHP and NC background contributions to the total MC are visible as the red and green histograms respectively*

4.3.5.2 Neutral Current DIS rejection

Specific cuts are applied in order to remove poorly measured NC events from the sample. To be classified as NC (and thus be rejected), an event must satisfy all of the following:

- $\delta > 20\,\mathrm{GeV}$

  δ, or $E - P_z$ is distributed around twice the positron energy for NC events (twice the positron energy $= 2E^{e+} = 55\,\mathrm{GeV}$).

- $\not{P}_T < 30\,\mathrm{GeV}$

  Though they may have some $\not{P}_T$ due to poor measurement, NC events are expected to have $\not{P}_T$ close to zero.

- $\wp_{sinistra} > 0.9$

  $\wp_{sinistra}$ is the probability that an positron has been found by SINISTRA. SINISTRA [75] is an electron/positron finder which uses a neural network to identify the final state positron using the calorimeter. For an event to be classified as NC, SINISTRA must identify a track to which it ascribes a $> 90\%$ probability of being created by the scattered positron.

- $E^{e+}_{scat} > 4\,\mathrm{GeV}$

  The scattered final state positron energy (E^{e+}_{scat}) is required to be greater than 4 GeV.

- $E_{nincon} < 5\,\mathrm{GeV}$

  E_{nincon} is the energy not associated with the positron candidate in an η, ϕ cone of radius 0.8. This cut requires that the positron is *isolated* . Energy deposits not associated with the positron itself, in a region around the positron (as described, in an $\eta - \phi$ cone of radius $R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.8$) are required to be less than 5 GeV. η is the pseudo-rapidity of the positron, defined by:

$$\eta = -ln\ \tan(\frac{\theta}{2}) \tag{40}$$

- if positron is in CTD acceptance ($15 < \theta < 164^o$)

  - $d_{trk} < 15\,\mathrm{cm}$

  The positron identified in the calorimeter is required to have an associated track passing within 15 cm, if the positron was within the CTD acceptance.

  - $P^{e+}_{trk}/E^{e+}_{cal} > 0.25$

  The positron track momentum (P^{e+}_{trk}) must be of a size similar to the energy deposit (E^{e+}_{cal}) associated with it in the calorimeter. A threshold of at least 25% is chosen.

- if positron is outside CTD acceptance

  - $P^{e+}_T > 20\,\mathrm{GeV}$

  If the positron is not in the CTD acceptance, little or no tracking information exists. For the event to qualify as an NC candidate, the positron is required to have a large transverse momentum (P^{e+}_T). This is the transverse momentum which is 'missing' in the CC signal.

The distributions of these quantities can be seen in figures 35 and 36. The NC cuts are all required to be satisfied, in order to reject the most obvious NC candidates, without rejecting CC signal. The hatched areas represent the regions where events fail to qualify as NC, and are *not* rejected.

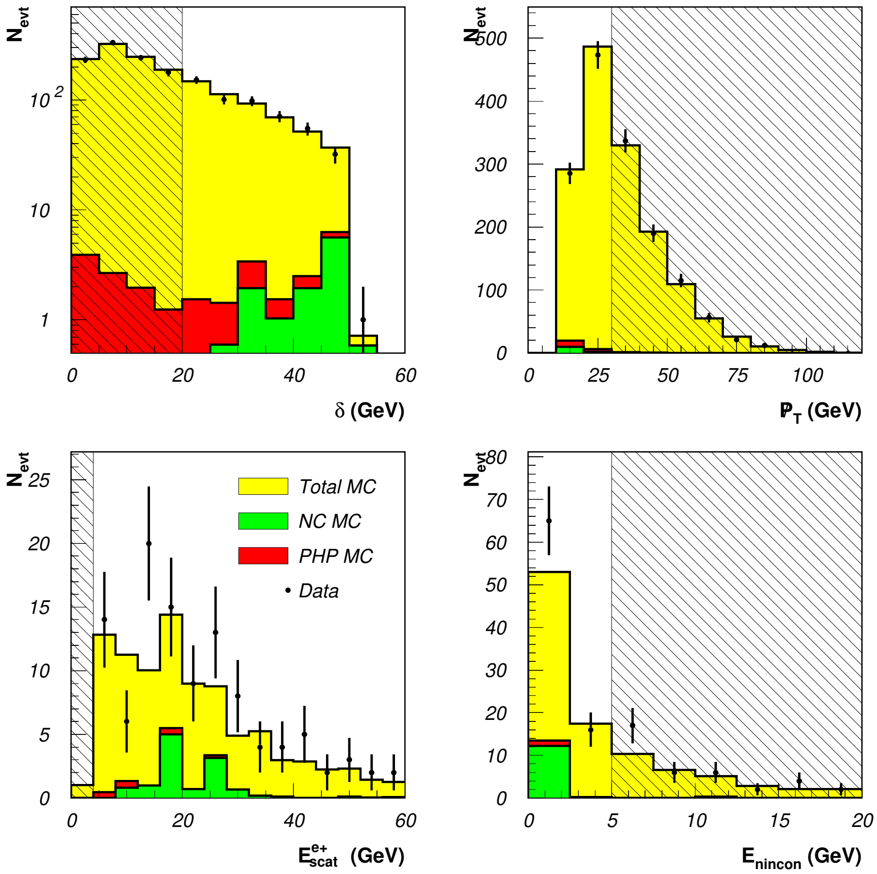

Figure 35: *The distributions of the quantities used to identify NC background for data (closed circles) and total MC (yellow histograms). The PHP and NC background contributions to the total MC are visible as the red and green histograms respectively. All cuts except those for NC DIS are applied. Candidate NC events are in the unhatched regions, but to be removed as NC, must be identified as such in all the relevant plots. This effectively shows where the NC events are located in the sample.*

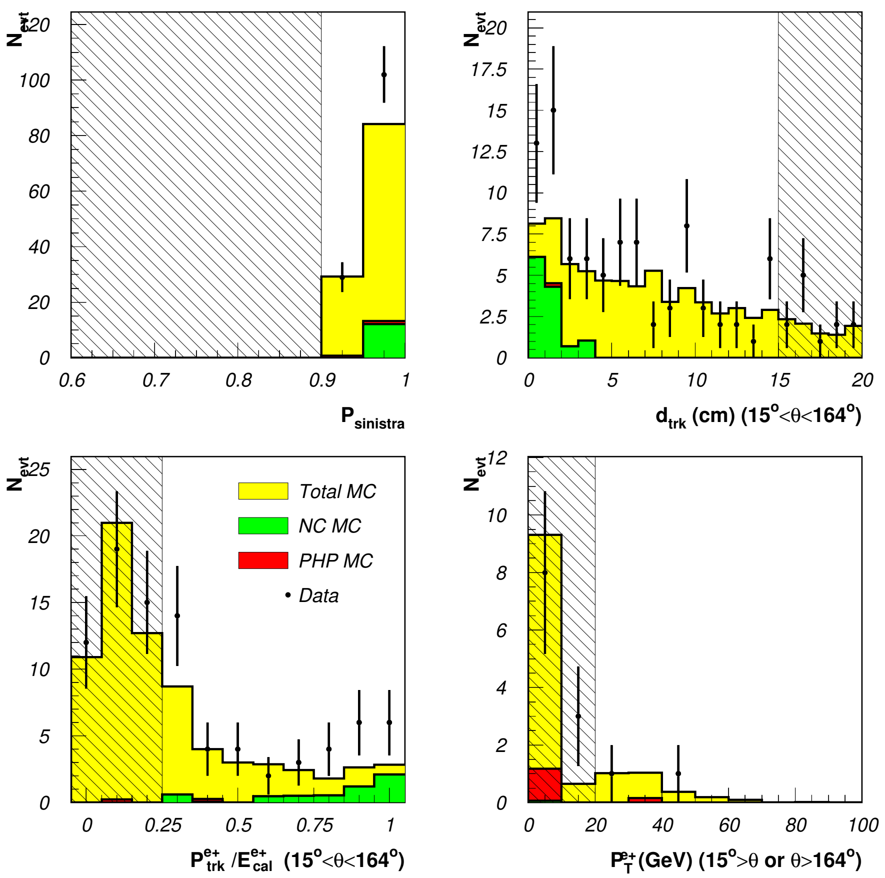

Figure 36: *The distributions of the quantities used to identify NC background for data (closed circles) and total MC (yellow histograms). The PHP and NC background contributions to the total MC are visible as the red and green histograms respectively. All cuts except those for NC DIS are applied. Candidate NC events are in the unhatched regions, but to be removed as NC, must be identified as such in all the relevant plots. This effectively shows where the NC events are located in the sample.*

4.3.5.3 Muon rejection

A program called MUFFIN [76, 77] was used to reject halo and cosmic muon events, primarily by identifying characteristic long narrow energy deposits in the calorimeter. MUFFIN also employs algorithms which use the ZEUS muon chambers, calorimeter timing, backing calorimeter and tracking information from the CTD to reject muons. In addition, for events with a BCAL energy exceeding 2 GeV, the fraction of energy in the hadronic section of the BCAL was required to be less than 90% of the total energy in the BCAL. This distribution can be seen in figure 37

Finally, events are removed if the transverse energy from one cell contributes more than 50% of the energy in the calorimeter. This distribution can be seen in figure 38.

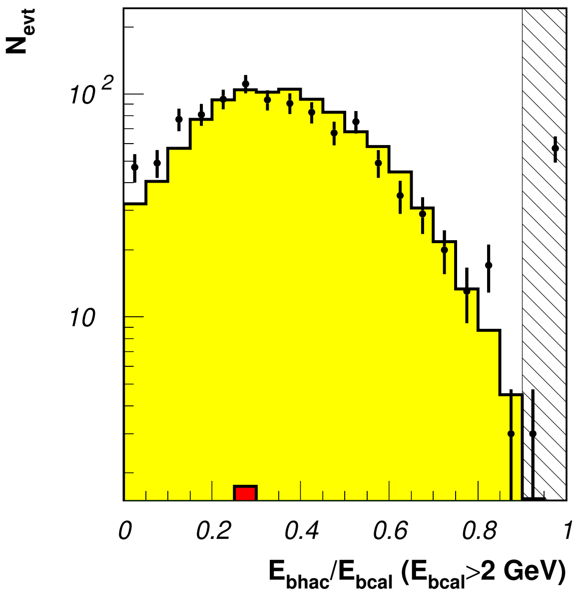

Figure 37: *The distribution of BCAL HAC energy relative to the total energy in BCAL after all other cuts have been applied for events with total BCAL energies greater than 2 GeV. Data events are shown as closed circles, whilst total MC is shown as a yellow histogram. The PHP contribution to the total MC is visible as a red histogram.*

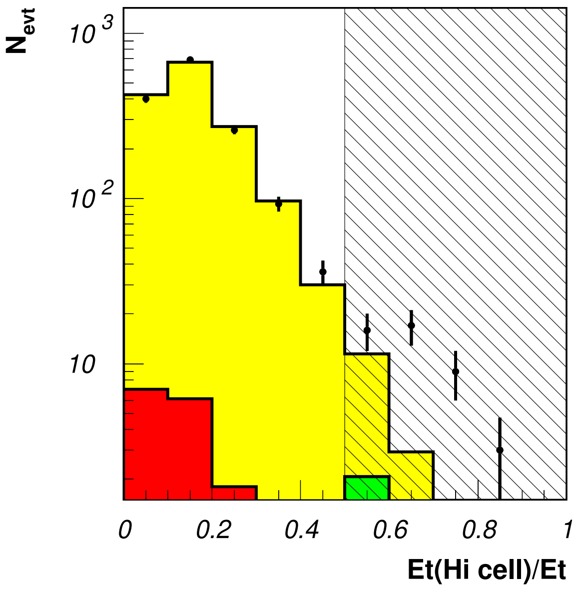

Figure 38: *The distribution of the ratio of transverse energy from the highest cell value to the total transverse energy in the calorimeter for data (closed circles) and total MC (yellow histogram). The PHP and NC background contributions to the total MC are visible as the red and green histograms respectively. All other cuts have been applied.*

4.3.6 Final event sample

The final event sample contained 1184 events in the high-γ_0 region and 293 events in the low-γ_0 region. As a first impression, this compares favourably to the 1206 and 288 charged current signal events predicted by the MC simulation in these regions (although at this stage the number of real data events quoted includes both signal and background events).

4.4 Cross section measurements

In the previous sections the event simulation, reconstruction and the selection of charged current signal have been discussed. This section will focus on the extraction of the charged current cross sections from the sample. In order to do this, the sample is divided into bins. It is desirable to use as many bins as possible to obtain information about the cross section shape, but the number of bins that can be defined is limited by the statistical precision of the data and resolution of the reconstruction. Measurements of the single differential cross sections with respect to Q^2, x and y are made ($\frac{d\sigma^{CC}_{Born}}{dQ^2}$, $\frac{d\sigma^{CC}_{Born}}{dx}$ and $\frac{d\sigma^{CC}_{Born}}{dy}$ respectively). In addition the reduced double differential cross section $\tilde{\sigma}$ (see equation 7) is measured as a function of x and Q^2.

4.4.1 Bin definitions

The number of bins into which the data are divided, their sizes and boundaries are important factors in the process of making cross section measurements. To assess the suitability of the bins using MC data, the acceptance $\mathcal{A}$, efficiency $\mathcal{E}$, and purity $\mathcal{P}$, of a bin must now be defined:

$$\mathcal{A} = \frac{\text{no. of events measured in bin}}{\text{no. of events generated in bin}} \tag{41}$$

$$\mathcal{E} = \frac{\text{no. of events generated and measured in bin}}{\text{no. of events generated in bin}} \tag{42}$$

$$\mathcal{P} = \frac{\text{no. of events generated and measured in bin}}{\text{no. of events measured in bin}} \tag{43}$$

where the number of events measured in a bin is the number of events in the bin satisfying the selection criteria of §4.3, and the number of events generated in a bin is the number of events in the kinematic region, (prior to any selection) calculated using the "true" values of the kinematic variables (available for MC), as opposed to the measured ones used in the selection process for data and MC.

The efficiency, acceptance and purity are related by the equation $\mathcal{A} = \mathcal{E}/\mathcal{P}$.

The acceptance $\mathcal{A}$ for each bin accounts for both the geometrical detector acceptance and the acceptance of the selection made in §4.3. An acceptance ~ 1 is desirable, since this reduces systematic uncertainties introduced by the MC simulation.

Events may not have been generated in the bin in which they are measured. This is known as *migration*. If the bin width is properly chosen, so that the resolution of a quantity is small compared to the bin size, then migration effects between bins (e.g. bin-to-bin correlations) and the statistical uncertainties will be minimised.

The efficiency $\mathcal{E}$ tells us the fraction of events which are generated in a bin and are then subsequently measured in that bin. The purity $\mathcal{P}$ measures the ratio of the number of events which really belong in the bin to the actual number of measured events, and so is reduced by migrations into a bin from other bins. High efficiency and purity will yield a good measurement, where migrations are small, and the selection efficiency for charged current signal is high.

For ease of comparison with previous analyses [52, 78], the bins used for the single differential cross sections in Q^2, x and y were left unchanged. For the measurement of $\frac{d\sigma^{CC}_{Born}}{dQ^2}$, 9 bins were used in total. The lowest and highest bins being wider in $log_{10}Q^2$ than the middle seven bins. These two bins spanned the ranges $Q^2 = 200 - 400\,\mathrm{GeV}^2$ and $Q^2 = 22500 - 60000\,\mathrm{GeV}^2$ respectively. The middle seven bins were of equal width in $log_{10}Q^2$, between $Q^2 = 400\,\mathrm{GeV}^2$ and $Q^2 = 22500\,\mathrm{GeV}^2$. The width of the bin for the lowest Q^2 events was originally chosen due to the background contamination being worse at low Q^2. The width of the bin for the highest Q^2 events was chosen because only a limited amount of data existed. The cross section $\frac{d\sigma^{CC}_{Born}}{dQ^2}$ is quoted at the logarithmic centre of each bin (i.e. at the bin centre on a log_{10} scale).

The resolution of the Jacquet-Blondel estimators improves rapidly for increasing

x and for this reason, three bins of equal width in $log_{10}x$ were chosen at low x, between $x = 0.01$ and $x = 0.1$ and four, narrower, bins of equal width at high x, between $x = 0.1$ and $x = 1$. The differential cross section $\frac{d\sigma^{CC}_{Born}}{dx}$ is quoted at the logarithmic centre of each bin, except for the highest x bin, where the cross section falls extremely rapidly. In this case, the cross section is quoted at $x = 0.65$.

The rapid change of the cross section $\frac{d\sigma^{CC}_{Born}}{dy}$ at low-y means that it is desirable to use narrower bins in this region. For this reason two narrow bins of equal width in y, between $y = 0$ and $y = 0.2$ were used. Then, at higher-y, five larger bins of equal width in y were used, between $y = 0.2$ and $y = 0.9$. The cross section is quoted at the centre of each bin.

The resolutions of the Jacquet-Blondel estimators, for the three single differential cross sections, in the bins described above, are shown in figures 39, 40, 41. With the exception of the highest Q^2, x and y bins, the resolutions are generally centred around zero with narrow widths, so no bias of the measured variables is expected, and migrations are expected to be low. The highest bins would be expected to exhibit larger uncertainties due to the shoulders which project outside the bin boundaries.

It is important to note at this stage that initial and final state radiation has been taken into account for the true MC values. When the positron (or, *much* less frequently, the proton) radiates a photon before the main interaction, then the initial momentum of the positron/proton is changed. This modified momentum is the one actually used in the model as the true MC variable. Similarly, the proton remnant may radiate a photon after the main interaction, modifying the hadronic remnant momentum. Again, the true MC momentum used for the proton just after the hard scatter is the sum of the photon and remnant momenta.

Figure 42 shows the acceptance, efficiency and purity for each of the single differential bins defined above. The purity and acceptance exceed 40% in all bins.

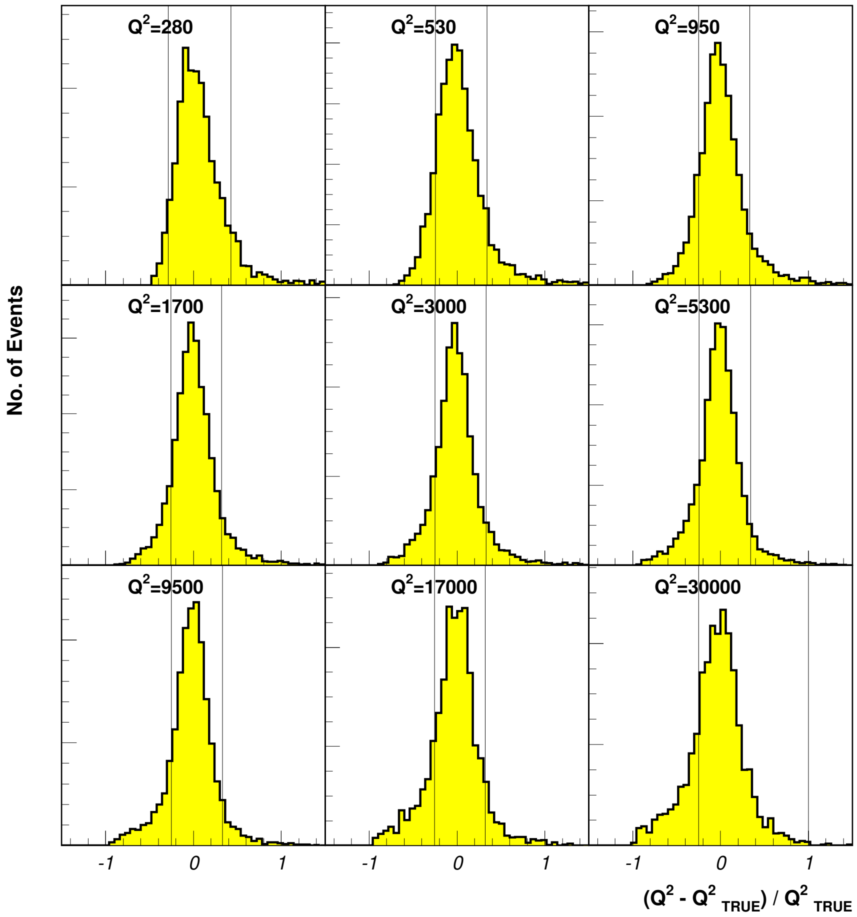

Figure 39: *The resolution of the measured Q^2 in the bins used for the extraction of $d\sigma_{Born}^{CC}/dQ^2$. The lines indicate the bin boundaries.*

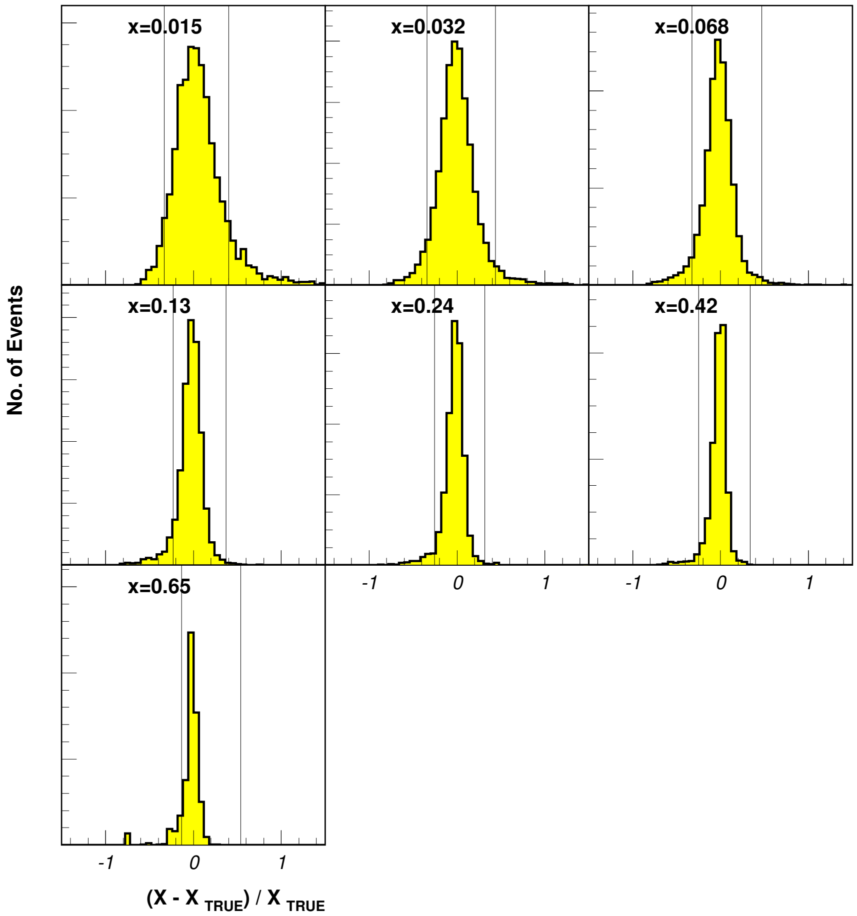

Figure 40: *The resolution of the measured x in the bins used for the extraction of $d\sigma^{CC}_{Born}/dx$. The lines indicate the bin boundaries.*

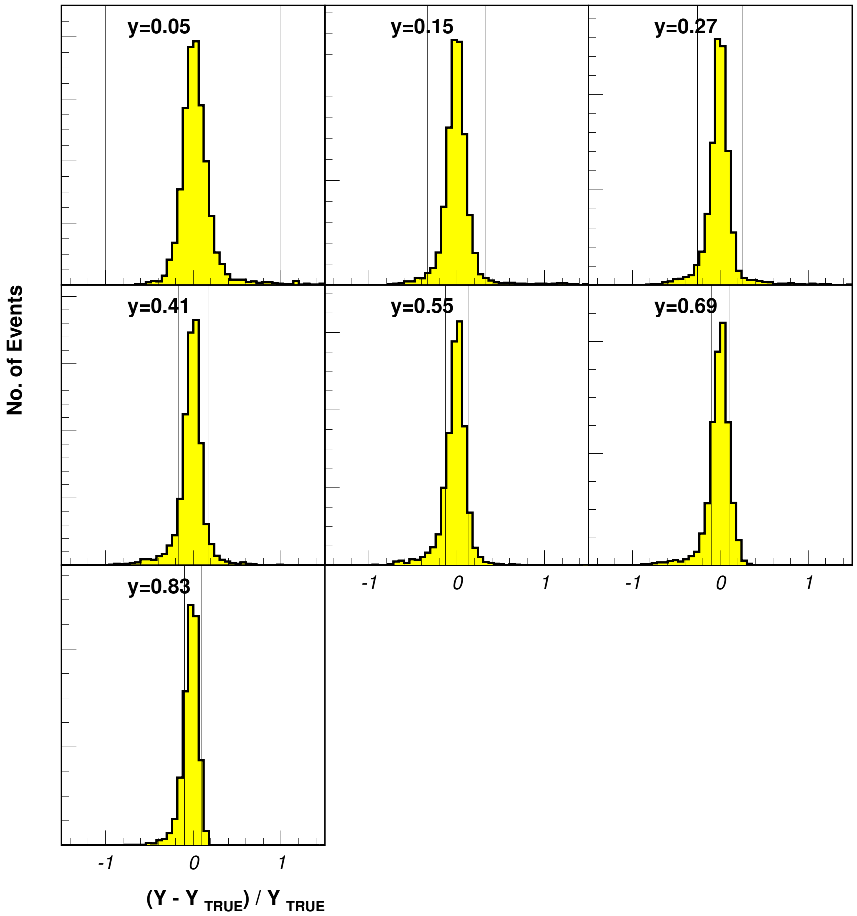

Figure 41: *The resolution of the measured y in the bins used for the extraction of $d\sigma^{CC}_{Born}/dy$. The lines indicate the bin boundaries.*

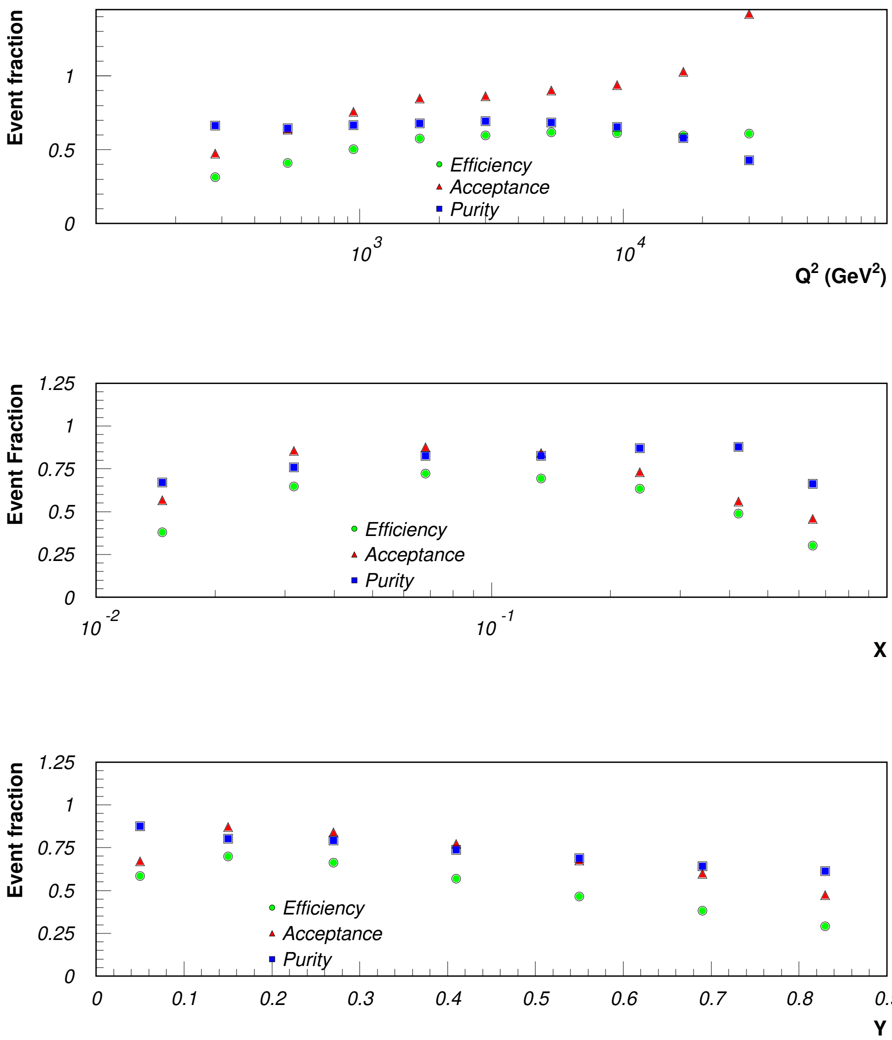

Figure 42: *The acceptance, efficiency and purity for the bins used in the determination of the single differential cross sections, shown at the bin centres.*

The bins used for the extraction of the double differential cross section were based on the bins used for the single differential cross section extraction. These bins, together with the selected data sample are collected in table 7, and shown in figure 43, which also shows the acceptances, efficiencies and purities of the bins. The resolutions of the quantities Q^2, x and y are also shown in figures 44, 45 and 46. Again, the resolutions are generally narrow, and centred around zero. The acceptances and purities are above 35%, which was deemed to be an acceptable lower boundary for these smaller double differential cross section bins. Many of the acceptances and purities are much higher than this threshold, and improve rapidly with increasing x or Q^2. Finally, the average migration expected from each bin used to extract the cross section is shown in figure 47.

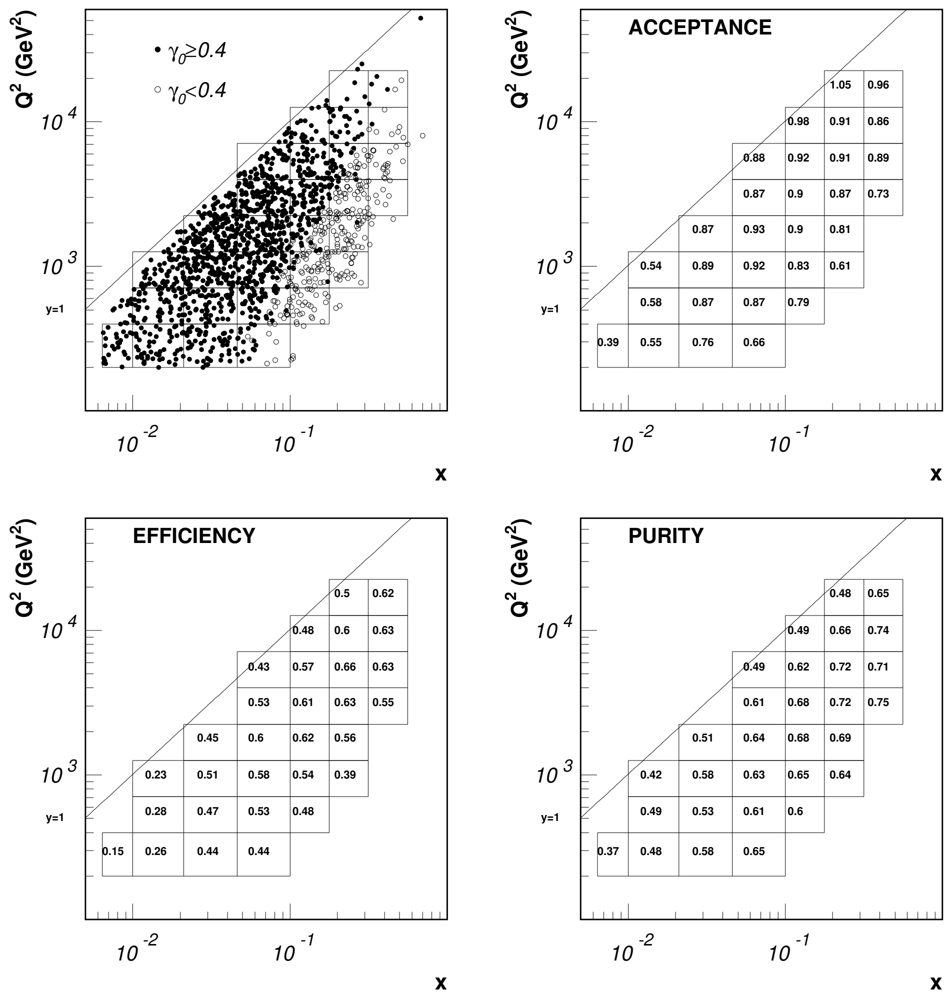

Figure 43: *The distribution of the event sample in the kinematic plane with bins used in the double differential cross section overlaid (top left). The closed circles correspond to events having* $\gamma_0 \geq 0.4$*, whilst open circles correspond to events having* $\gamma_0 < 0.4$*. Also shown are the acceptances, efficiencies and purities of the bins used in the double differential cross section determination.*

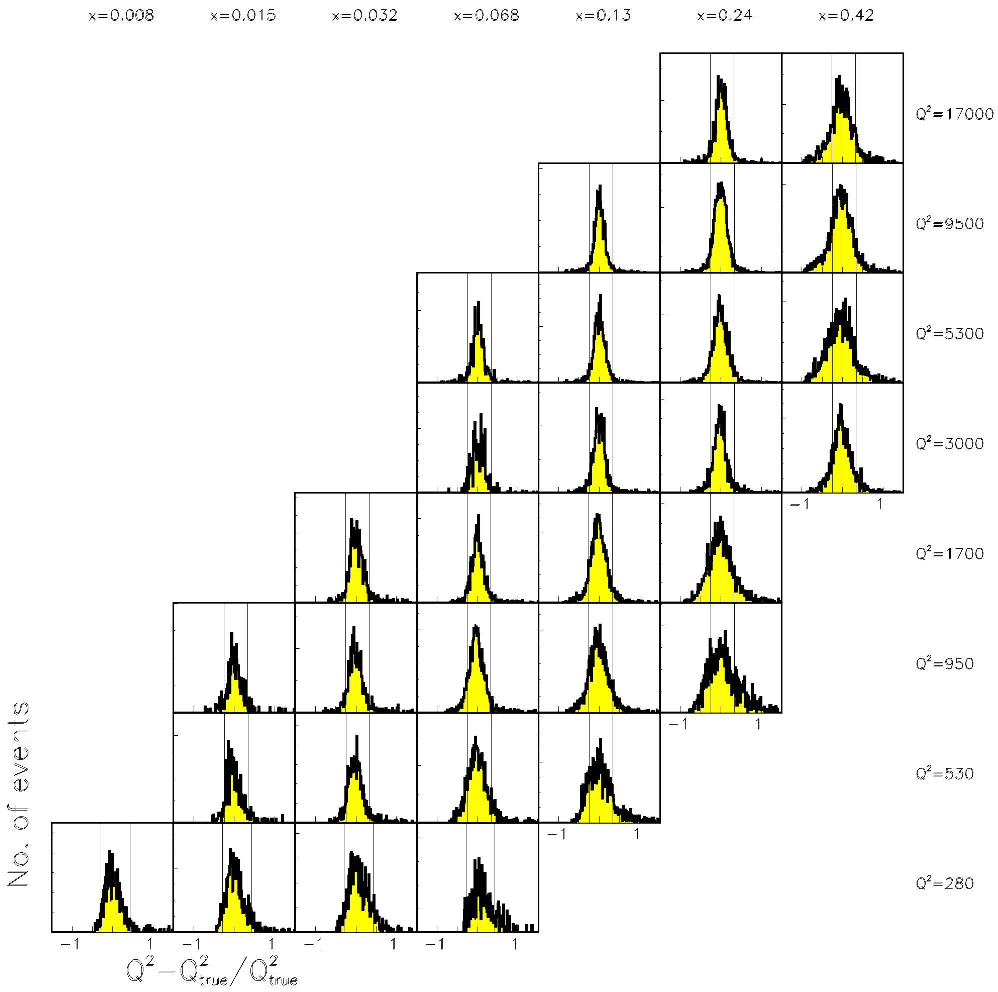

Figure 44: *The resolution of Q^2, $(Q^2 - Q^2_{true})/Q^2_{true}$ for each of the bins used in the determination of $\tilde{\sigma}$. The lines indicate the bin boundaries.*

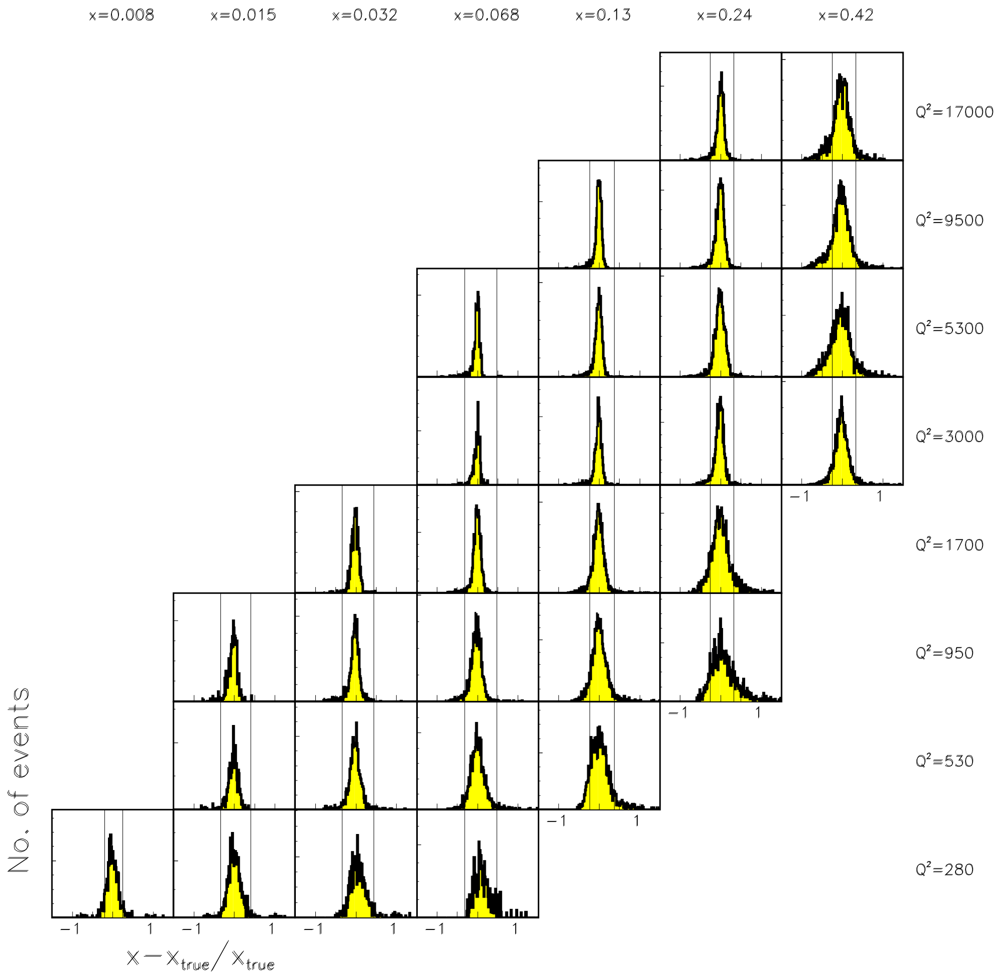

Figure 45: *The resolution of x,$(x - x_{true})/x_{true}$ for each of the bins used in the determination of $\tilde{\sigma}$. The lines indicate the bin boundaries.*

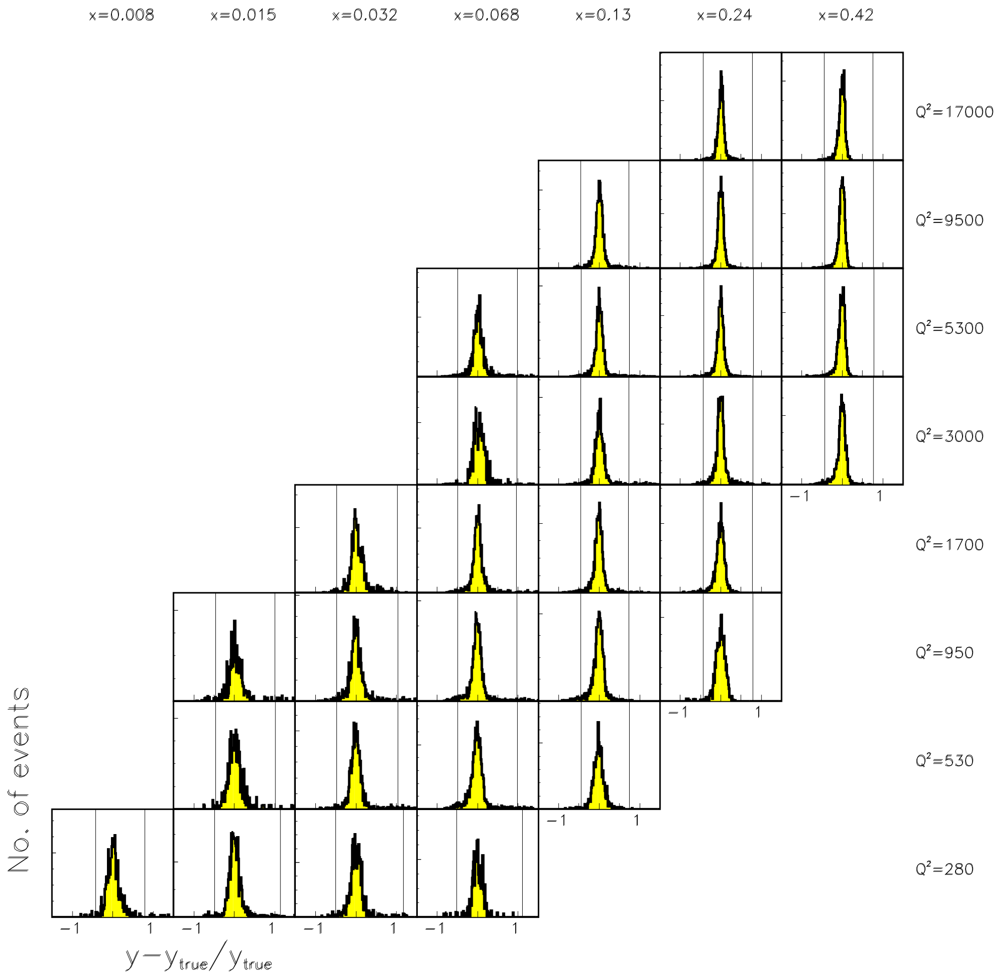

Figure 46: *The resolution of* y*,*$(y - y_{true})/y_{true}$ *for each of the bins used in the determination of* $\tilde{\sigma}$*. The lines indicate the bin boundaries.*

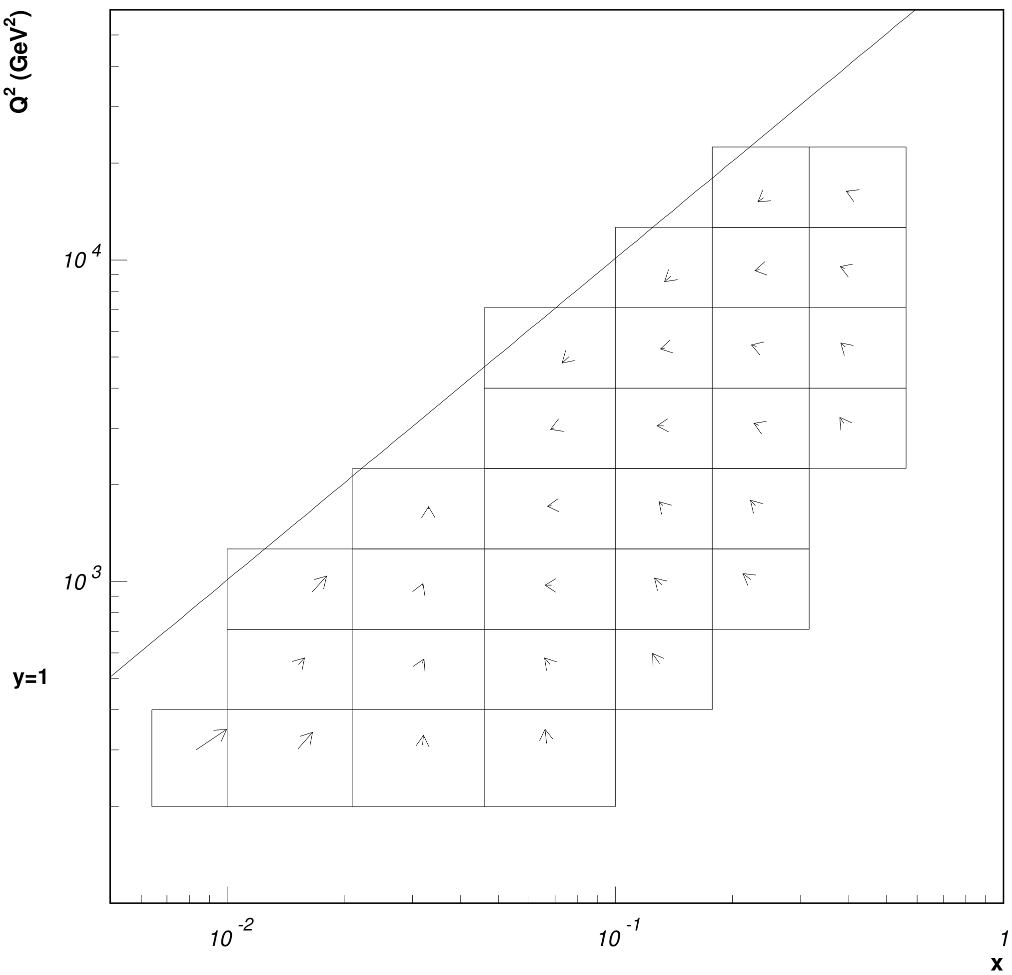

Figure 47: *The average migration expected in each of the bins used to extract the double differential cross section. The arrows go from mean generated x and Q^2 to the mean measured x and Q^2 in each bin.*

4.4.2 Method for cross section extraction

The single and double differential cross sections are extracted using 'bin-by-bin' unfolding, in which the measured number of events in each bin is corrected for detector acceptance, resolution and radiative effects.

The measured integrated cross section σ_{meas} in a bin is given by the expression

$$\sigma_{meas} = \frac{N_{obs} - N_{bg}}{\mathcal{L}\mathcal{A}} \tag{44}$$

where $\mathcal{L}$ is the integrated data luminosity ($60.9\,\mathrm{pb}^{-1}$), and $\mathcal{A}$ is the detector acceptance. N_{bg} represents the number of background events estimated using MC (each MC event having its weight multiplied by the data luminosity). N_{bg} must be subtracted from the observed number of signal data events in each bin, N_{obs}. The measured cross section includes electroweak radiative effects as discussed in §1.4. The correction factor $\mathcal{C}_{rad}$, used to obtain the Born level cross section from the radiative cross section is:

$$\mathcal{C}_{rad} = \frac{\sigma^{SM}_{Born}}{\sigma^{SM}_{rad}}. \tag{45}$$

The label SM refers to the Standard Model prediction. The correction factor is then applied to the measured cross section to obtain the measured Born level cross section.

$$\sigma^{meas}_{Born} = \sigma_{meas} \cdot \mathcal{C}_{rad} \tag{46}$$

$$= \frac{N_{obs} - N_{bg}}{N^{MC}_{obs}} \cdot \sigma^{SM}_{Born} \tag{47}$$

where N^{MC}_{obs} is the number of observed MC signal events in the bin.

The differential cross sections, quoted at a particular x, Q^2 point, are next calculated from the integrated cross section. To do this, (for example in Q^2), a 'bin centre correction' $\mathcal{C}_{bin}$, is defined.

$$\mathcal{C}_{bin} = \frac{\frac{d\sigma^{SM}_{Born}}{dQ^2}\big|_{Q_0^2}}{\sigma^{SM}_{Born}} \tag{48}$$

The differential cross section with respect to Q^2 at the point Q_0^2 is then simply

$$\left.\frac{d\sigma^{meas}_{Born}}{dQ^2}\right|_{Q_0^2} = \mathcal{C}_{bin} \;.\; \sigma^{meas}_{Born} \tag{49}$$

$$= \frac{N_{obs}-N_{bg}}{N^{MC}_{obs}} \;.\; \left.\frac{d\sigma^{SM}_{Born}}{dQ^2}\right|_{Q_0^2} \;. \tag{50}$$

4.4.3 Statistical uncertainty

The statistical uncertainty is calculated for each of the cross section points. For bins with 12 or more events, the statistical uncertainty assigned to the number of data events is $\sqrt{N_{obs}}$. In bins with fewer than 12 events observed, the statistical uncertainty on the number of data events was set equal to the Poisson 68.3% confidence interval limits, assuming no background [79]. The small contribution of the statistical uncertainty on the MC to the overall cross section statistical uncertainty was also included. For MC, the number of events is given by $\mathcal{L} \times \sum_i w_i$, where $\mathcal{L}$ is the integrated data luminosity, and w_i is the event weight (see eqn. 27). The sum runs over MC events in the bin. The MC statistical uncertainty is given by $\mathcal{L} \times \sqrt{\sum_i w_i^2}$. Finally, the statistical uncertainty on the differential cross sections is (again, the example of $d\sigma^{meas}_{Born}/dQ^2$ is used):

$$\Delta\frac{d\sigma^{meas}_{Born}}{dQ^2} = \sqrt{\left(\frac{\Delta N_{obs}}{N^{MC}_{obs}}\right)^2 + \left(\frac{\Delta N_{bg}}{N^{MC}_{obs}}\right)^2 + \left(\frac{\Delta N^{MC}_{obs}(N_{obs}-N_{bg})}{(N^{MC}_{obs})^2}\right)^2} \times \frac{d\sigma^{SM}_{Born}}{dQ^2} \tag{51}$$

4.4.4 Systematic uncertainties

Bin-by-bin unfolding of the cross sections relies on an accurate MC simulation of the data event distributions. If the event and detector simulations are poor, then the process of using the number of MC events in each bin to extract the cross sections may introduce bias to the cross section measurement. The quantities considered to introduce systematic uncertainty via this process fall roughly into two categories, "cut variations" and "parameter variations". In cut variations, the selection cut on a particular variable is tightened and relaxed (for both data and MC), and the cross sections are re-extracted. The difference in the extracted value for each bin is used as an estimate of the systematic uncertainty introduced by any inaccuracies in the MC simulation near the cut values. Parameter variations attempt to extract systematic uncertainties which are introduced by assumptions built into the simulation or reconstruction of events. In a similar manner to cut variations, the parameter in question is varied and the cross sections re-extracted to estimate the systematic uncertainty in each bin. A good example of this type of systematic uncertainty is the variation of the calorimeter energy scale. In this way, an estimate of each individual systematic uncertainty for every bin is made. The separate uncertainties which increase or decrease the cross section estimates are then combined in quadrature to give a total systematic uncertainty on the measurement for each kinematic bin.

Three of the sources of systematic uncertainty were identified as bin-to-bin correlated, and are labelled as δ_1, δ_2 and δ_3 below. These uncertainies generally exhibit trends when neighbouring bins are viewed alongside each other. It is advantageous to quantify the correlated systematic uncertainties separately (in addition to the total and uncorrelated systematic uncertainties), so that they may be treated properly in any further analysis, such as is performed in §5.

4.4.4.1 Variation of selection thresholds

The imposition of selection cuts on distributions which are not exactly the same can introduce biasses to the measured cross sections, as outlined previously. The cuts used to select both data and MC events were varied, and the cross sections re-evaluated.

A summary of the cut variations made can be found in table 5. For the majority of bins, it was found that tightening and relaxing the cuts by an amount reflecting the parameter resolution typically resulted in only small changes ($< 2\%$) to the measured cross sections. There were a few exceptions, which included varying the $\not{P}_T/E_T$ cut at $\not{P}_T < 20\,\text{GeV}$. This produced a change of 6% in the lowest Q^2 bin of the $d\sigma/dQ^2$ cross section, and up to 20% in the low double differential cross section bins[5].

One further exception was the variation of $\not{P}_T(-ir)$ at $\gamma_0 < 0.4$, where variations of up to 5% were observed in the highest two x bins of the $d\sigma/dx$ cross section, and bin thirteen[5] of the double differential cross section where a variation of 18% was observed.

Quantity	Threshold
$\not{P}_T$ ($\gamma_0 \geq 0.4$)	$> 12 \pm 1.2\,\text{GeV}$
$\not{P}_T$ ($\gamma_0 < 0.4$)	$> 14 \pm 1.4\,\text{GeV}$
$\not{P}_T(-ir)$ ($\gamma_0 \geq 0.4$)	$> 10 \pm 1.0\,\text{GeV}$
$\not{P}_T(-ir)$ ($\gamma_0 < 0.4$)	$> 12 \pm 1.2\,\text{GeV}$
$\lvert\Delta\phi\rvert$ ($\not{P}_T < 20\,\text{GeV}$)	$< 0.5 \pm 0.05$
$\lvert\Delta\phi\rvert$ ($\not{P}_T \geq 20\,\text{GeV}$)	$< 2.0 \pm 0.2$
$\not{P}_T/E_T$ ($\not{P}_T < 20\,\text{GeV}$)	$> 0.55 \pm 0.055$
$\not{P}_T/E_T$ ($20 \leq \not{P}_T < 30\,\text{GeV}$)	$> 0.4 \pm 0.04$
Proportion of tracks which are "good", N_{good}	$> 0.25 \times (N_{trk}^{tot} - [20 \pm 4])$

Table 5: *The variation of selection thresholds. The individual items are defined in the event selection, see §4.3.*

[5]Bins in the double differential cross section are numbered 1-30, corresponding to the rows in table 7. i.e. running from lowest Q^2, x through to highest Q^2, x by stepping up firstly in x and then Q^2.

4.4.4.2 Calorimeter hadronic energy scale

The hadronic energy scale of each section of the calorimeter (FCAL, BCAL or RCAL) and each layer (EMC or HAC) was varied by an amount corresponding to the relative uncertainty in each section (as determined in §4.2.4.2 and the study [67]). These variations produce some of the largest contributions to the total systematic uncertainty. The method for varying the individual energy scales as described in [66] was used to gain knowledge of the systematic uncertainties introduced on the cross section measurements.

For the first of the systematic uncertainties, all sections were scaled up and down together by their respective error margins (FCAL and BCAL $\pm1\%$, RCAL $\pm2\%$). The resulting variation in the measured cross sections is correlated from bin-to-bin, as seen in figure 48. It is labelled as δ_1 in the tables 6 and 7.

Next, the BCAL energy scale was decreased (increased) whilst the FCAL and RCAL energy scales were increased (decreased) by their respective uncertainties. Again, the resulting variation in the measured cross section is correlated from bin-to-bin, as can be seen in figure 49. This estimate of the systematic uncertainty is labelled as δ_2 in tables 6 and 7.

Finally, for each calorimeter section in turn (FCAL, BCAL and RCAL), the EMC and HAC energy fractions were altered by increasing the EMC fraction by 2% and simultaneously decreasing the HAC fraction by 2%, and vice-versa. This estimate of the systematic uncertainty introduced by HAC/EMC energy scale variation was much smaller than the other two energy scale uncertainties (generally $\leq5\%$), and was not considered to be bin-to-bin correlated. It was therefore included in the uncorrelated systematic uncertainty.

4.4.4.3 QCD cascade

The MC samples used to extract the cross sections were created using the Colour Dipole Model (CDM) contained within ARIADNE (see §4.1.1) to simulate the

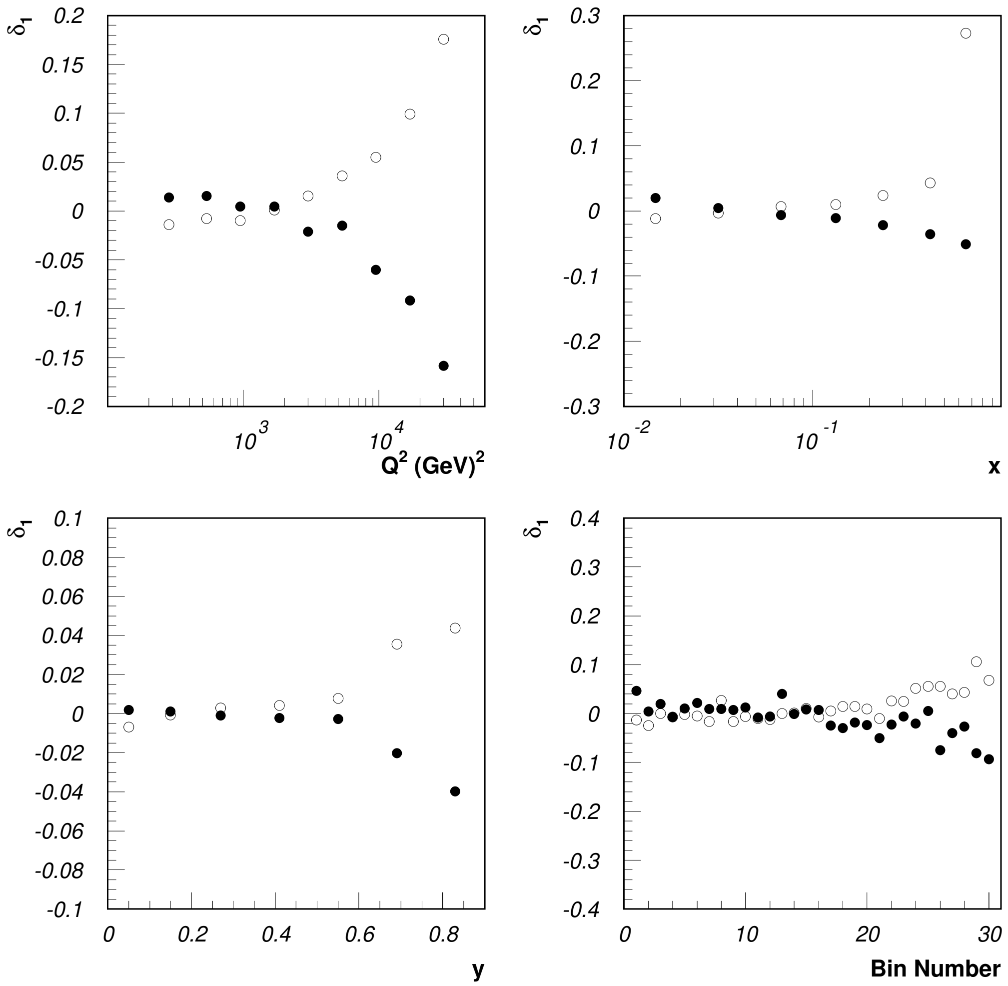

Figure 48: *The relative change in the measured cross sections for each bin due to the first estimate of the calorimeter energy scale systematic uncertainty. The FCAL, BCAL and RCAL were simultaneously scaled up (closed circles) and then down (open circles) by their respective uncertainties.*

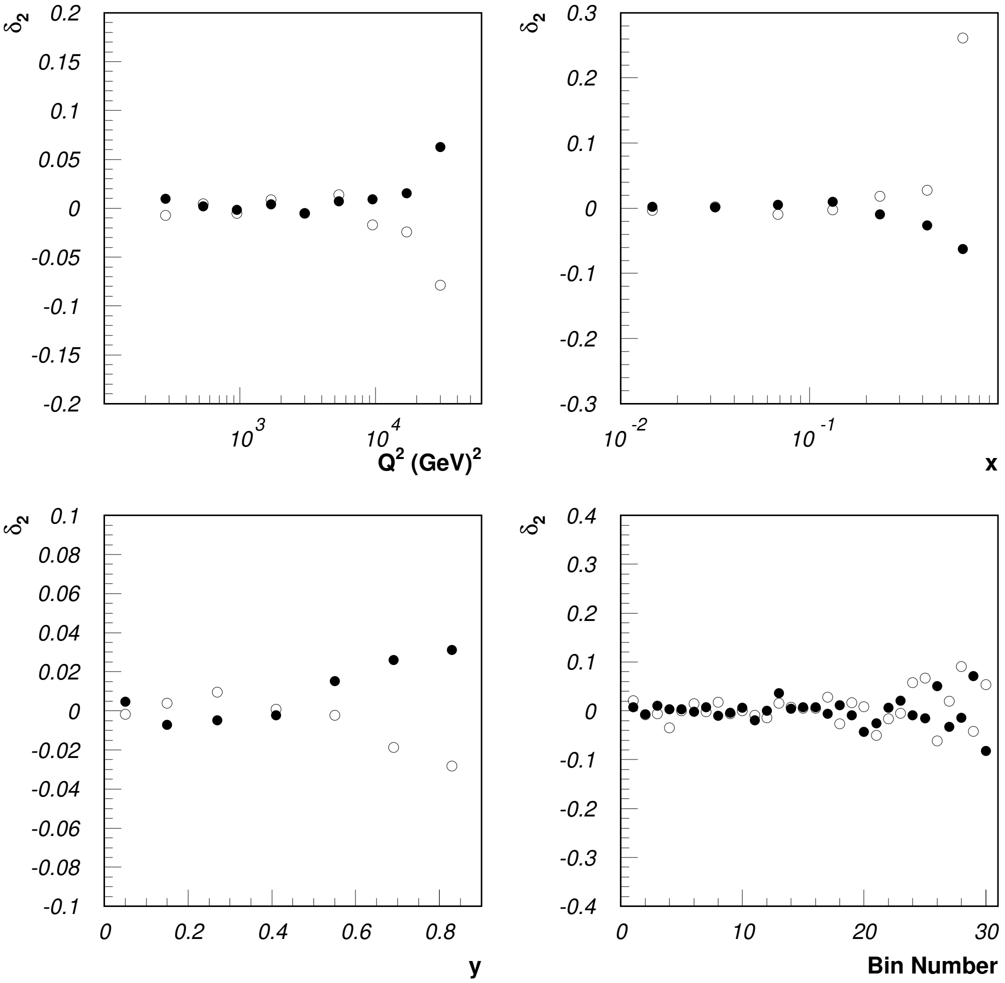

Figure 49: *The relative change in the measured cross sections for each bin due to the second estimate of the calorimeter energy scale systematic uncertainty. The BCAL scale was increased whilst FCAL and RCAL scales were simultaneously decreased by their respective uncertainties (open circles), and then vice versa (closed circles).*

QCD cascades (showering) in the detector. In addition, the MEPS model was used, and changes considered to be the systematic uncertainty introduced by the choice of showering model. As can be seen in figure 50, the contribution to the total systematic uncertainty is significant at high Q^2 and high x. In addition, the uncertainty was considered to be bin-to-bin correlated, and is labelled as δ_3 in tables 6 and 7.

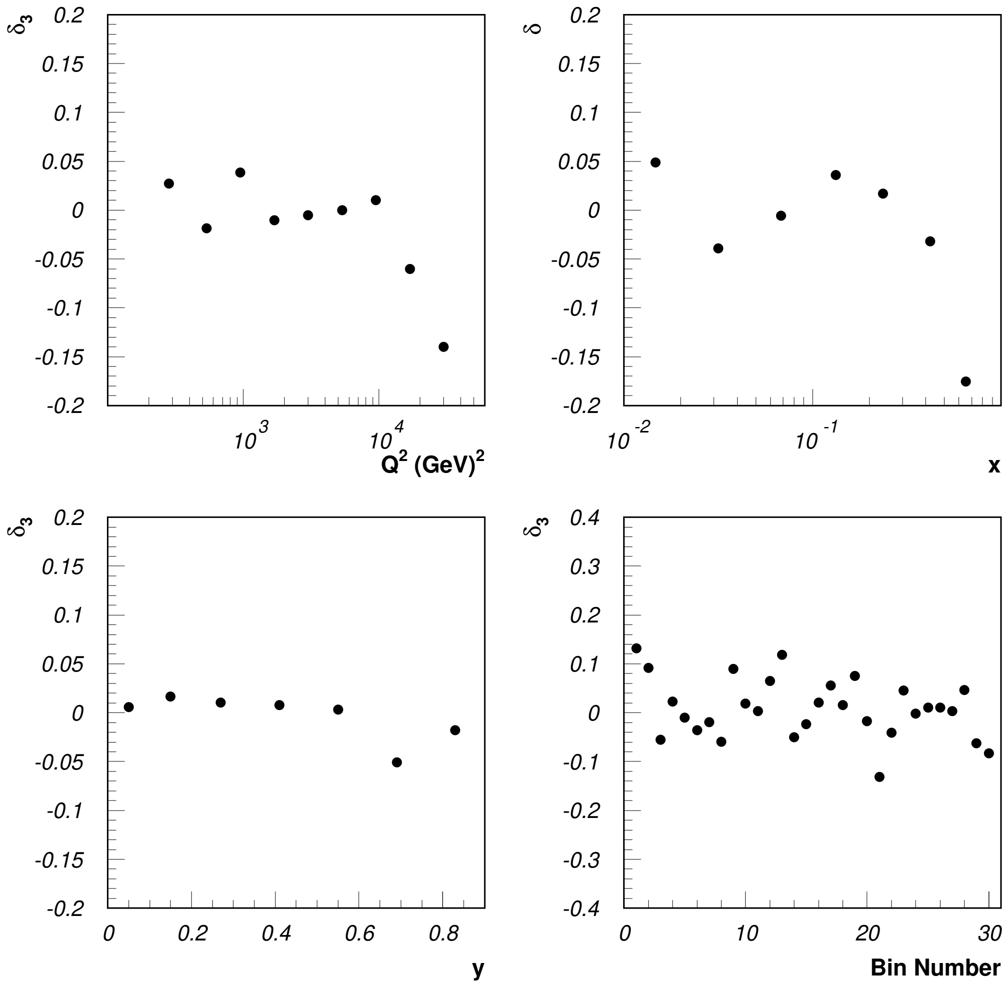

Figure 50: *The relative change in the measured cross sections for each bin when using MEPS instead of ARIADNE.*

4.4.4.4 Effect of F_L

The DJANGOH program neglects the longitudinal structure function (F_L) contribution to the double differential cross section when generating CC events. The effect of this assumption on the measured cross sections was tested by reweighting MC events by the ratio of the cross sections both with and without F_L. The reweighted events were then used to extract the cross sections, and the systematic uncertainty deduced from the changes. The systematic uncertainties introduced were under 2% in all bins, except in the highest y bins, where the uncertainty steadily increased to a maximum of 3%.

4.4.4.5 Parton distribution function dependence

The MC used in this analysis was generated using the CTEQ5D parton distribution functions (PDFs) [42]. The ZEUS-S fit [80] was used to examine the effect of the PDF choice on the cross sections by reweighting the MC events to the extremes of the cross section prediction allowed by the fit. The changes observed in the measured cross sections were all lower than 5%. These changes were included in the uncorrelated and total systematic uncertainties.

4.4.4.6 Photoproduction subtraction

The direct and resolved photoproduction (PHP) processes (modelled in MC), have cuts applied at the generator level (§4.1.2.2), which make it difficult to estimate the level of PHP contamination in the final data sample. A rudimentary estimate was made of the overall normalisation factor which should be applied to the PHP event weights by analysing the cut from §4.3.3.2 on $\not{P}_T/E_T$, for $\not{P}_T < 20\,\text{GeV}$. The cut at 0.55 is particularly efficient at removing PHP events, so it was temporarily removed whilst the normalisation of the PHP was investigated (the cut was replaced for the extraction of the cross sections). A fit was then performed to the data $\not{P}_T/E_T$ spectrum by varying the overall normalisation of the PHP. This can be seen in figure 51 plot (e).

In addition, the proportions of direct and resolved PHP events were also varied (figure 51, plots (a)-(i)), though the fits show little sensitivity to this factor, and equal proportions were actually used. The fit yielded a normalisation factor of 0.99 ± 0.56.

The systematic uncertainty on the overall cross sections was then estimated by varying the PHP normalisation within these limits. The effect was generally very small ($< 2\%$) for all cross section bins, except for the lowest Q^2 bin in the $d\sigma/dQ^2$ cross section, where the uncertainty was approximately 4%, and also in the first four double differential cross section bins, where the uncertainty was between 3 and 6%.

4.4.4.7 Vertex finding efficiency

Events at low γ_0 were not required to have a CTD vertex in order to be selected. These events can be analysed to see if the CTD actually detected the event vertex or not. The vertex finding efficiency ε_{CTD}, is defined as the ratio of events having a CTD vertex to the total number of events. It is shown as a function of γ_0 (the hadronic angle calculated using the nominal vertex) in figure 52. Data and MC efficiencies agree reasonably well, and both approach 100% near $\gamma_0 = 0.4$. The contribution to the systematic uncertainty from the CTD vertex finding efficiency is neglected.

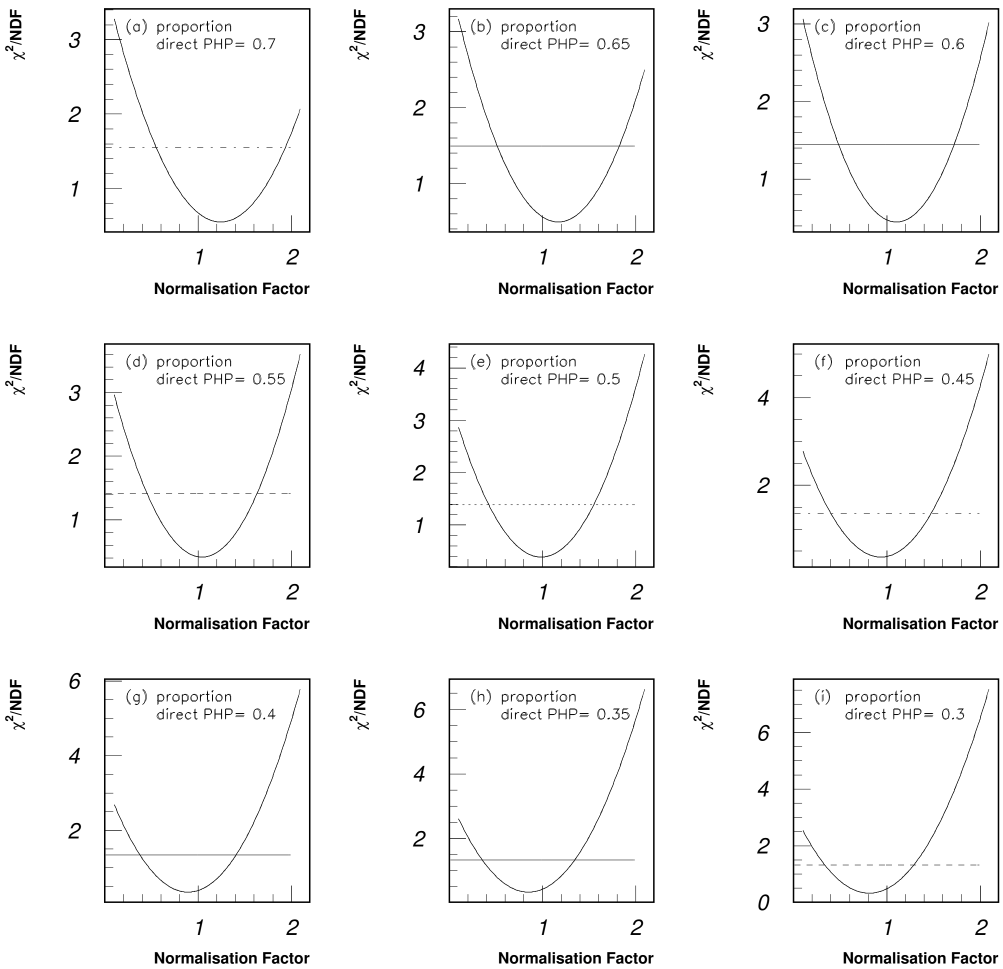

Figure 51: *The results of the fit to determine the overall normalisation factor for the PHP sample (plot (e)), and the sensitivity of the fit to the proportions of direct and resolved processes (plots (a)-(i)).*

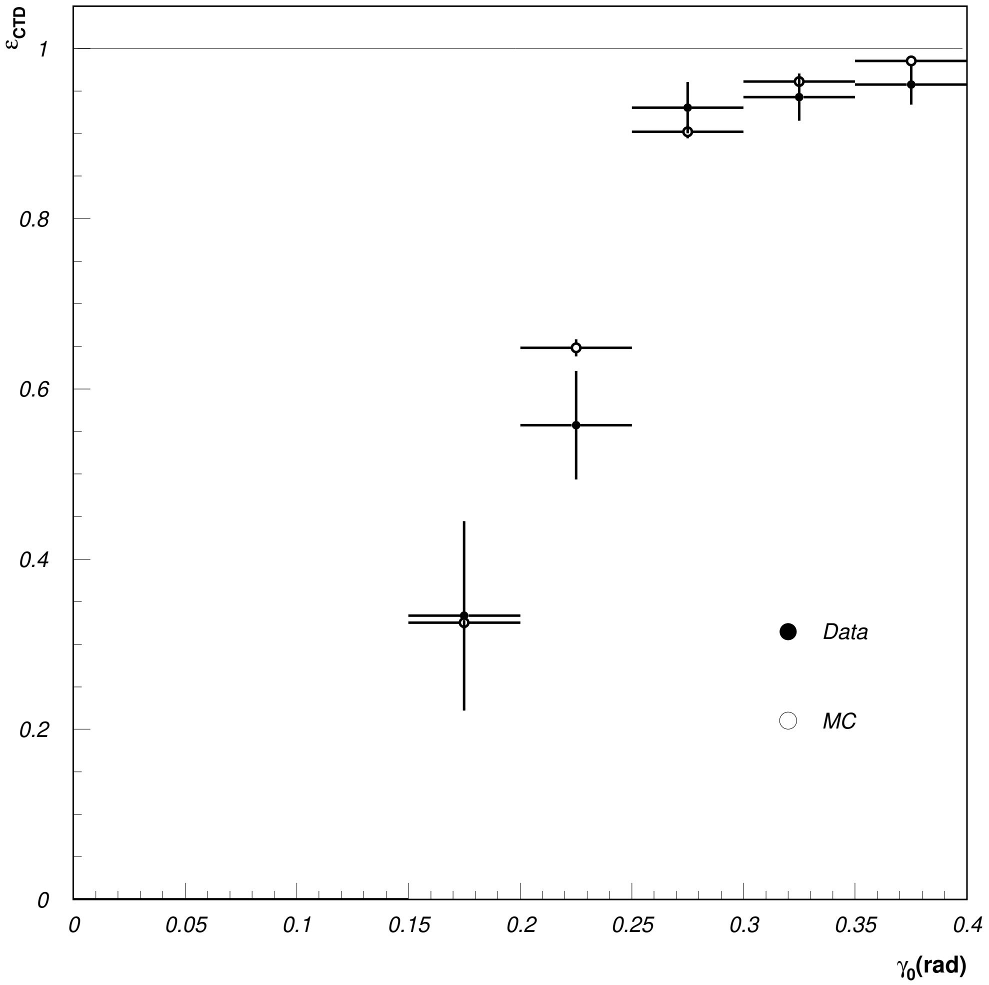

Figure 52: *The vertex finding efficiency of the CTD ε_{CTD}, at low values of γ_0. Both data (closed circles) and MC (open circles) have been included.*

4.4.4.8 Trigger efficiency

The ZEUS trigger system is described in §2.2.4, and the trigger selection used for charged current is outlined in §4.3.2. In the first level trigger (FLT) selection process, the quantity $\not{P}_T^{FLT}$ is coarsely reconstructed. At the second and third level trigger stages, the $\not{P}_T$ calculation is more accurate, since it uses the full information from the data pipelines. The efficiency of the charged current selection in the FLT based on $\not{P}_T^{FLT}$ was examined using an independent trigger slot which accepted events based on CAL energy (slot 42).

$$\varepsilon^{slot60\_61} = \frac{\text{No. of events passing slot 42 \& at least one of slots 60, 61}}{\text{No. of events passing slot 42}} \tag{52}$$

To investigate the efficiency at low values of offline $\not{P}_T$, the cuts on both low and high γ_0 events were relaxed to 8 GeV, and the offline cuts on $\not{P}_T(-ir)$ were removed completely. The efficiency was then calculated as a function of offline $\not{P}_T$, and the results can be seen for data and MC in figure 53. The fitted turn-on curves are of the form

$$\frac{1}{2}.\tanh\left(\frac{\not{P}_T - p_1}{p_2}\right) + \frac{1}{2} \tag{53}$$

where p_1 and p_2 are the free parameters of the fit. The MC events were reweighted to the data turn-on curve, in order to assess the systematic uncertainty introduced by the differing distributions. The resulting shift in the cross sections was under 1% in all bins, and was thus neglected.

4.4.4.9 Systematic uncertainty summary

The three correlated systematic uncertainties $\delta_1, \delta_2, \delta_3$, and the uncorrelated systematic uncertainties were combined to give the total systematic uncertainty in each bin. The individual uncertainties were added in quadrature separately for all of the positive and negative deviations, and the results are shown in figure 54 and listed in tables 6 and 7.

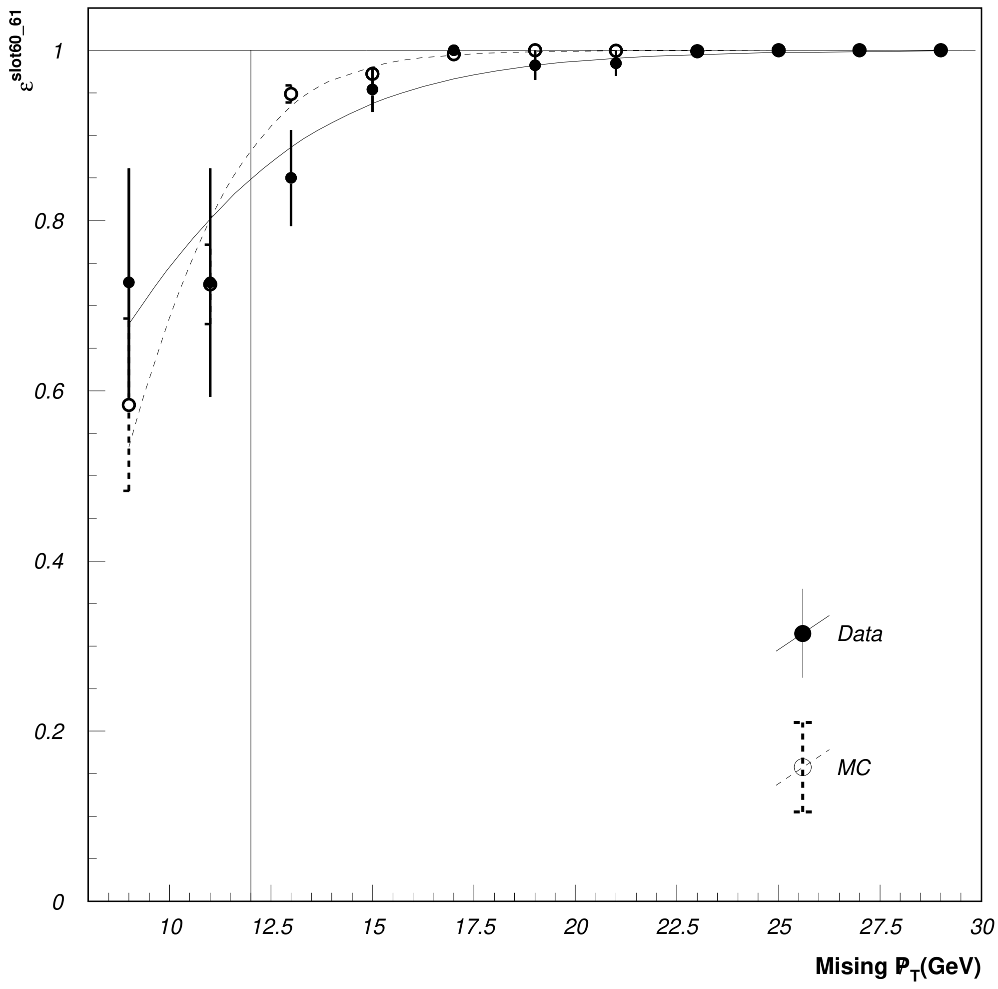

Figure 53: *The efficiency of slots 60,61 (ε^{slot60_61}), as a function of offline reconstructed missing transverse momentum. Both data (closed circles) and MC (open circles) have been included. The vertical line shows the lowest of the two offline selection thresholds for $\not{P}_T$ at 12 GeV.*

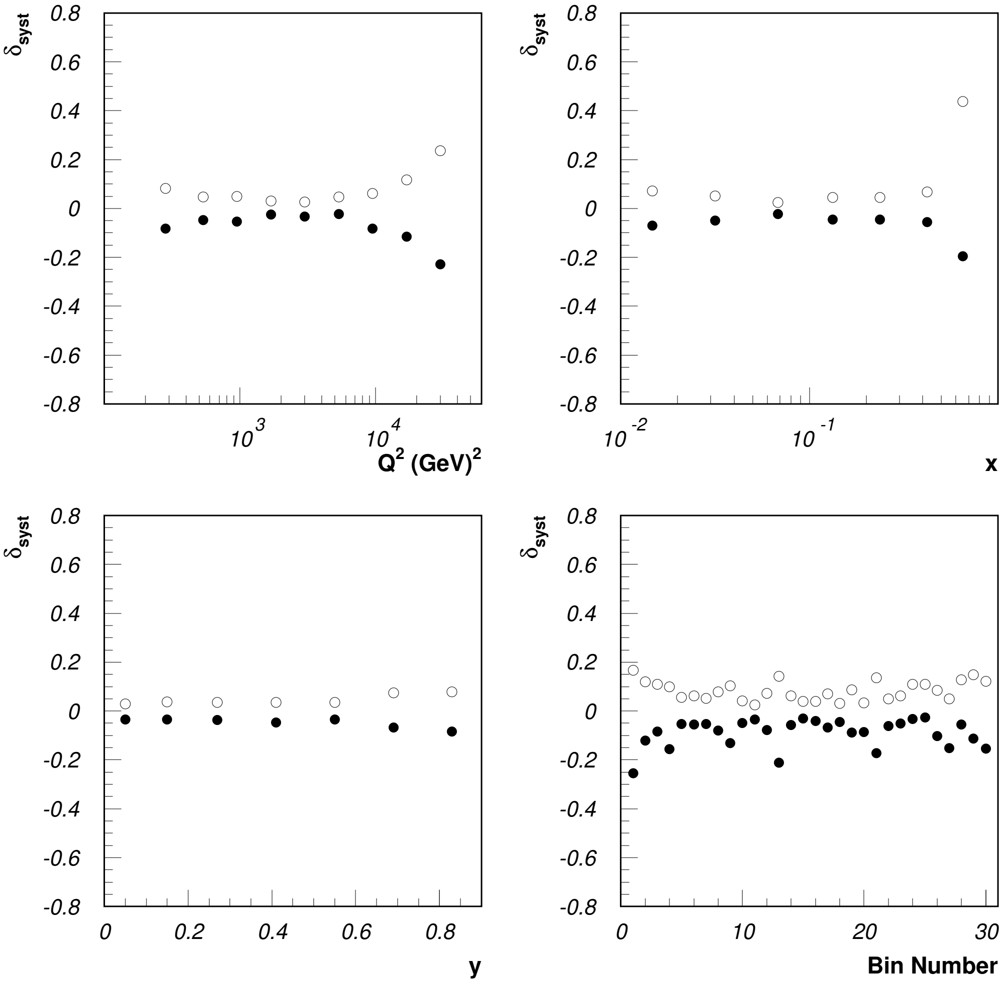

Figure 54: *The total positive (open circles) and negative (closed circles) systematic uncertainties for each bin.*

4.4.5 Cross section results

The total cross section for e^+p CC DIS in the region $Q^2 > 200\,\text{GeV}^2$ is

$$\sigma_{CC}^{e^+p}(Q^2 > 200\,\text{GeV}^2) = 36.1 \pm 1.0(\text{stat.})^{+1.0}_{-1.1}(\text{syst.})\,\text{pb}$$

where the systematic uncertainty includes the overall luminosity uncertainty of 2.25%. This compares well with the Standard Model prediction of $37.0^{+1.7}_{-0.8}\,\text{pb}$, which was evaluated using the ZEUS-S fit [80].

The single differential cross sections $\frac{d\sigma_{CC}^{e+p}}{dQ^2}$, $\frac{d\sigma_{CC}^{e+p}}{dx}$ and $\frac{d\sigma_{CC}^{e+p}}{dy}$ for $Q^2 > 200\,\text{GeV}^2$ are shown in figures 55, 56 and 57. The values, along with uncertainties are also listed in table 6. The bottom plot for each cross section shows the ratio of the measured cross section to the prediction made using the ZEUS-S fit. Also shown are the CTEQ6D [81] PDF parameterisations. The band represents the uncertainty on the cross section coming from the ZEUS-S PDF parameterisation.

The reduced double differential cross section, $\tilde{\sigma}$ (see equation 7) is shown as a function of Q^2 and x in figures 58 and 59. Also shown are the ZEUS-S, CTEQ6D and MRST(2001) [82] predictions. The values of $\tilde{\sigma}$ are collected in table 7. Values of $M_W = 80.419\,\text{GeV}$ and $G_F = 1.16639 \times 10^{-5}\,\text{GeV}^{-2}$ from the PDG [51] were used to calculate $\tilde{\sigma}$. Both the single and double differential cross sections are well described by the SM and are also in good agreement with the recent ZEUS analysis [83].

The results of the cross section determination were compared to the ZEUS analysis [83] in more detail. Some events rejected by this analysis were kept by the ZEUS analysis, and vice-versa. The differences were explained on an event-by-event basis, by the slightly differing selection criteria used in the two analyses.

From equations 5 and 7, to leading order in α_s, and neglecting bottom and top quarks, for e^+p scattering we obtain for the reduced cross section:

$$\tilde{\sigma} = x[(\bar{u} + \bar{c}) + (1 - y)^2(d + s)]. \tag{54}$$

For illustrative purposes, the contribution to the cross section made by the PDF combinations $x(\bar{u}+\bar{c})$ and $x(1-y)^2(d+s)$, (obtained from the NLO QCD ZEUS-S PDF parameterisations in the $\overline{MS}$ scheme) are shown in figure 59. Here, it can be seen that the charged current e^+p cross section is dominated by the d-quark density at high x. This implies that charged current measurements in this region should allow improved determination of the d-quark valence density of the proton, which is the topic of the next chapter of this analysis.

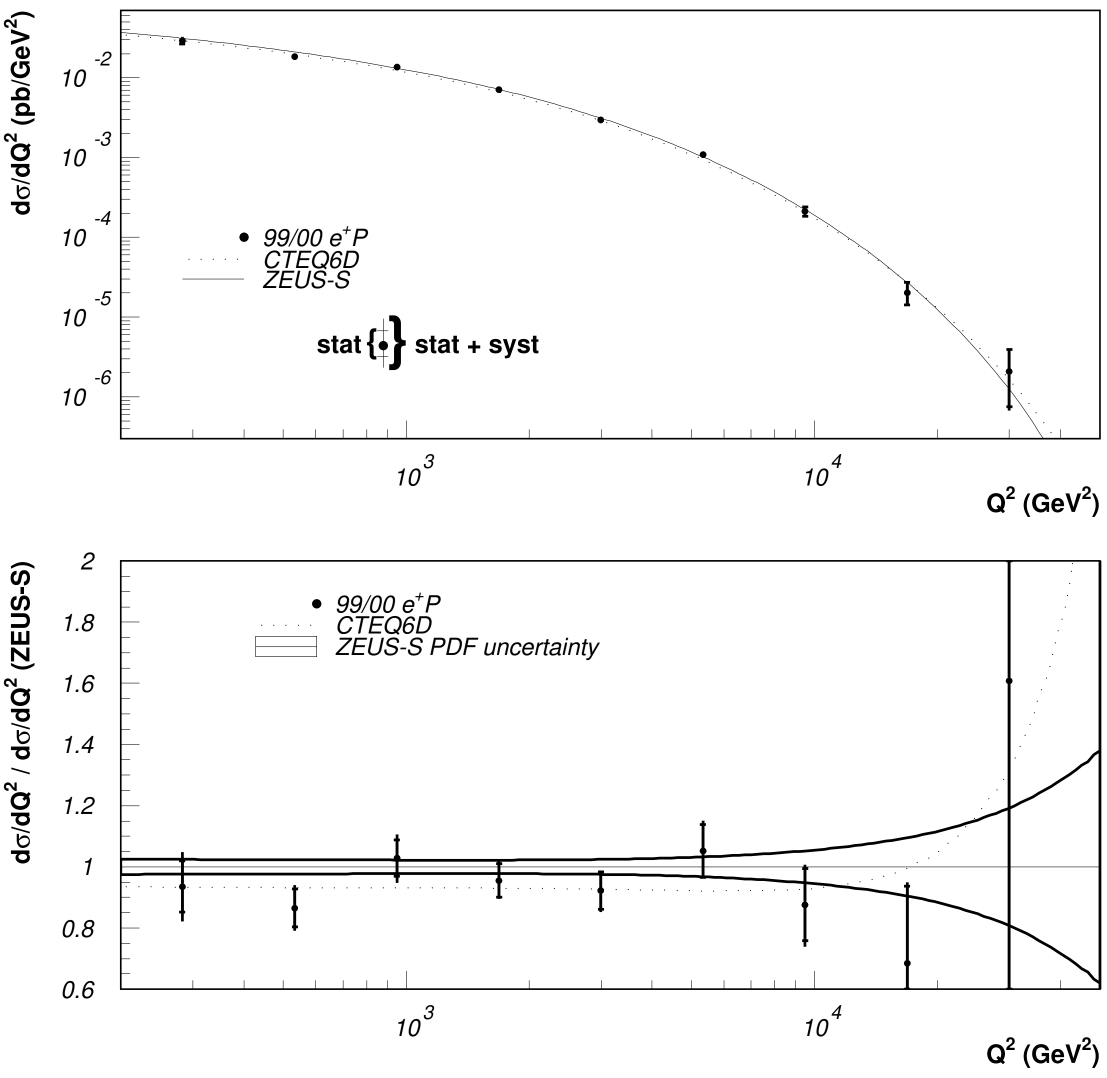

Figure 55: *The measured cross section $d\sigma/dQ^2$ (top). Also shown (bottom) is the ratio of the cross section to the ZEUS-S prediction.*

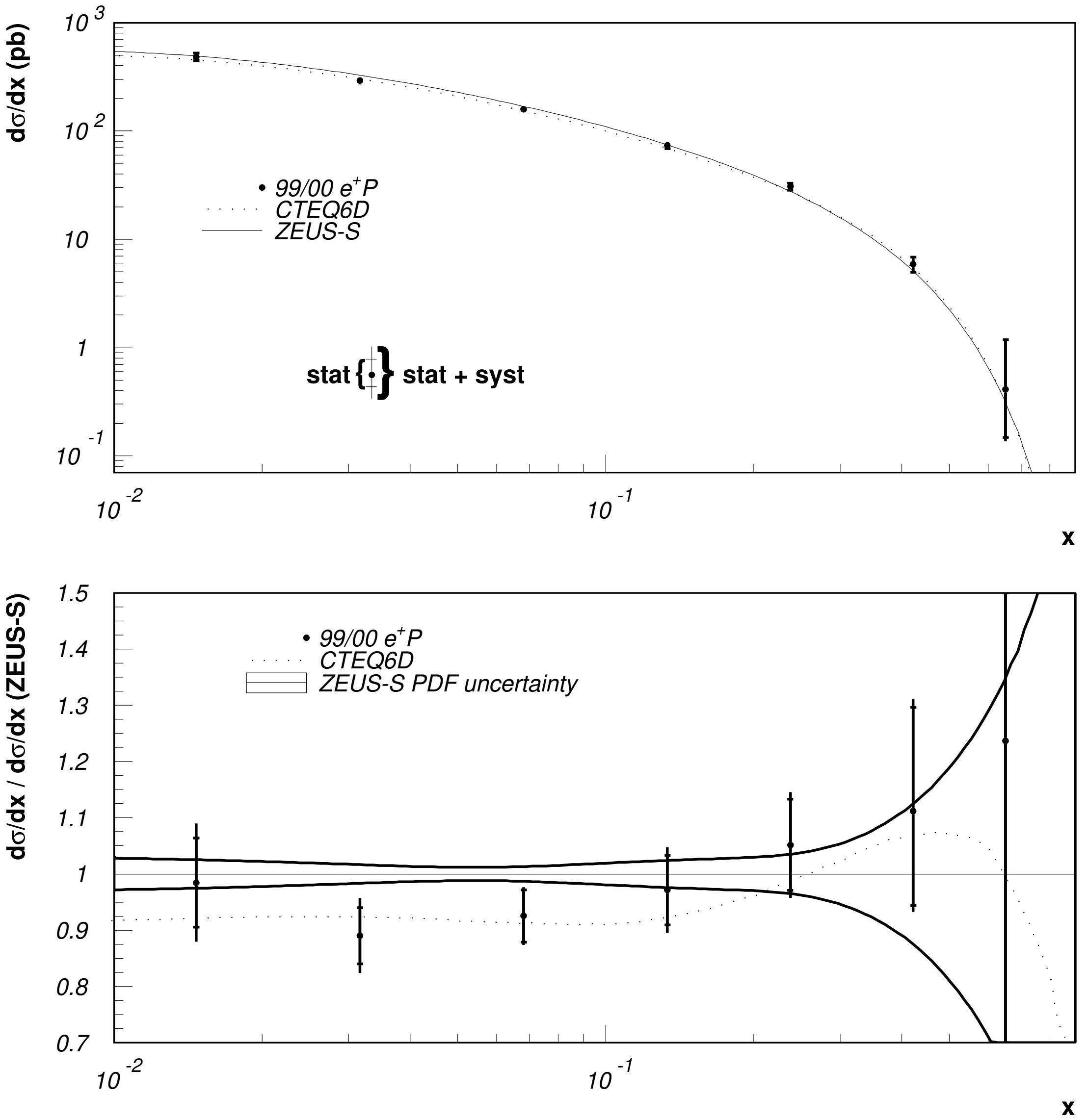

Figure 56: *The measured cross section $d\sigma/dx$ (top). Also shown (bottom) is the ratio of the cross section to the ZEUS-S prediction.*

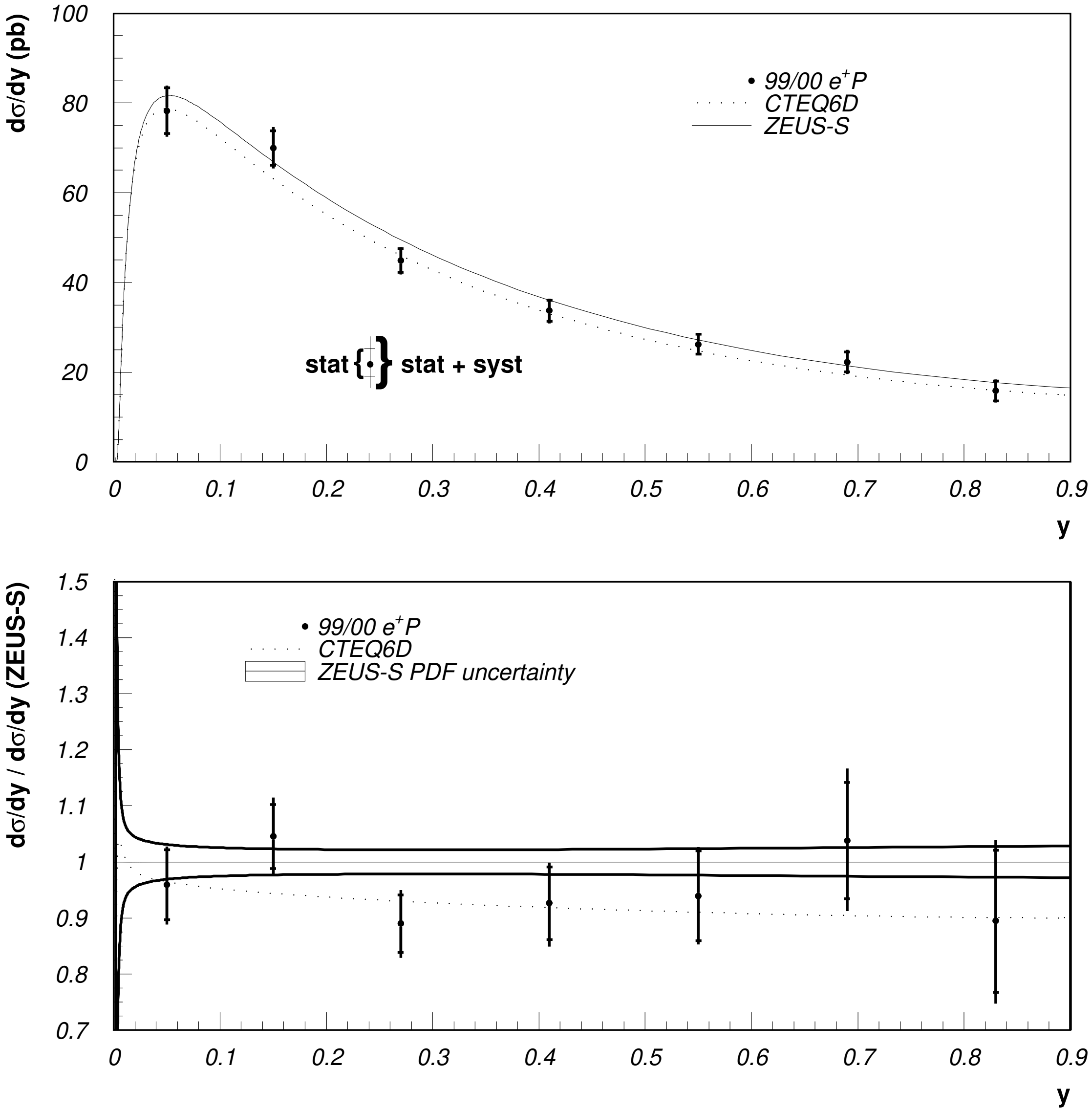

Figure 57: *The measured cross section $d\sigma/dy$ (top). Also shown (bottom) is the ratio of the cross section to the ZEUS-S prediction.*

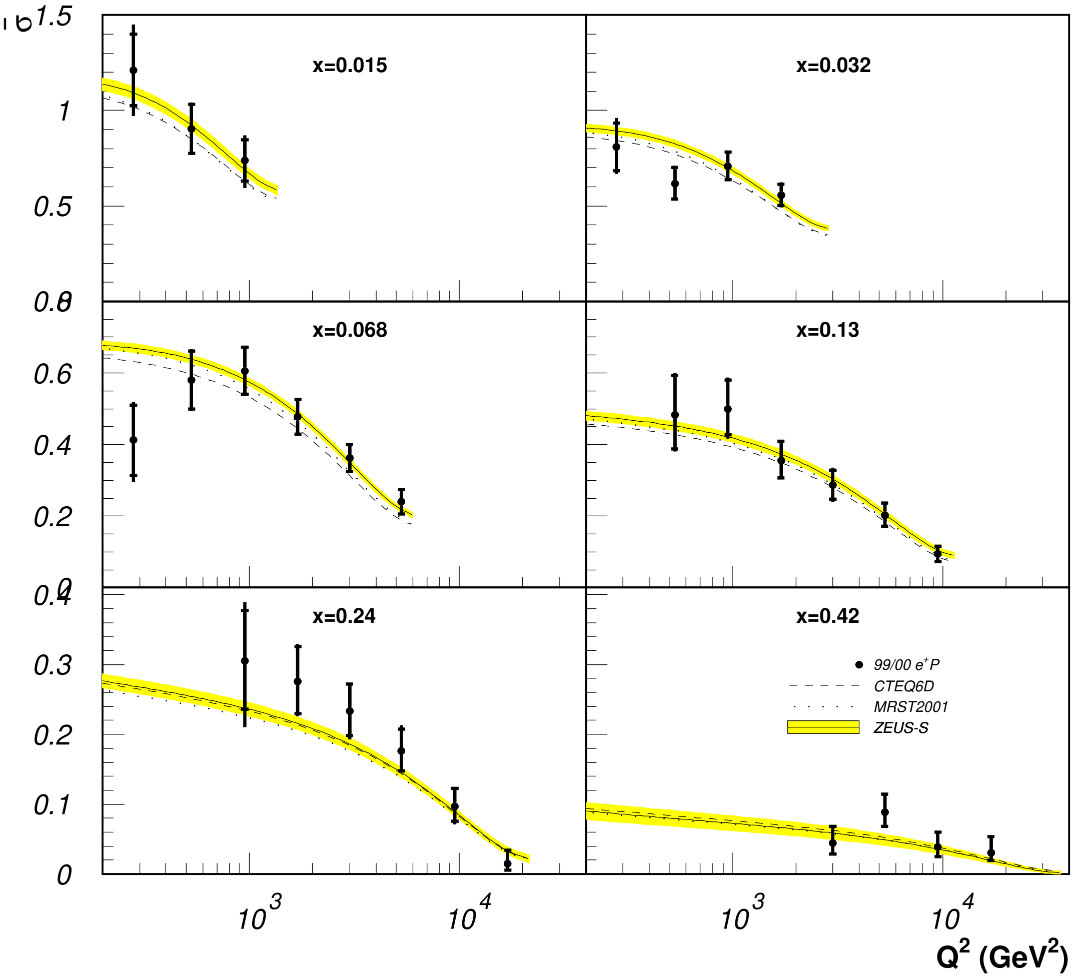

Figure 58: *The measured double differential reduced cross section $\tilde{\sigma}$, for $e^+p \to \overline{\nu_e}\, X$, in bins of fixed x.*

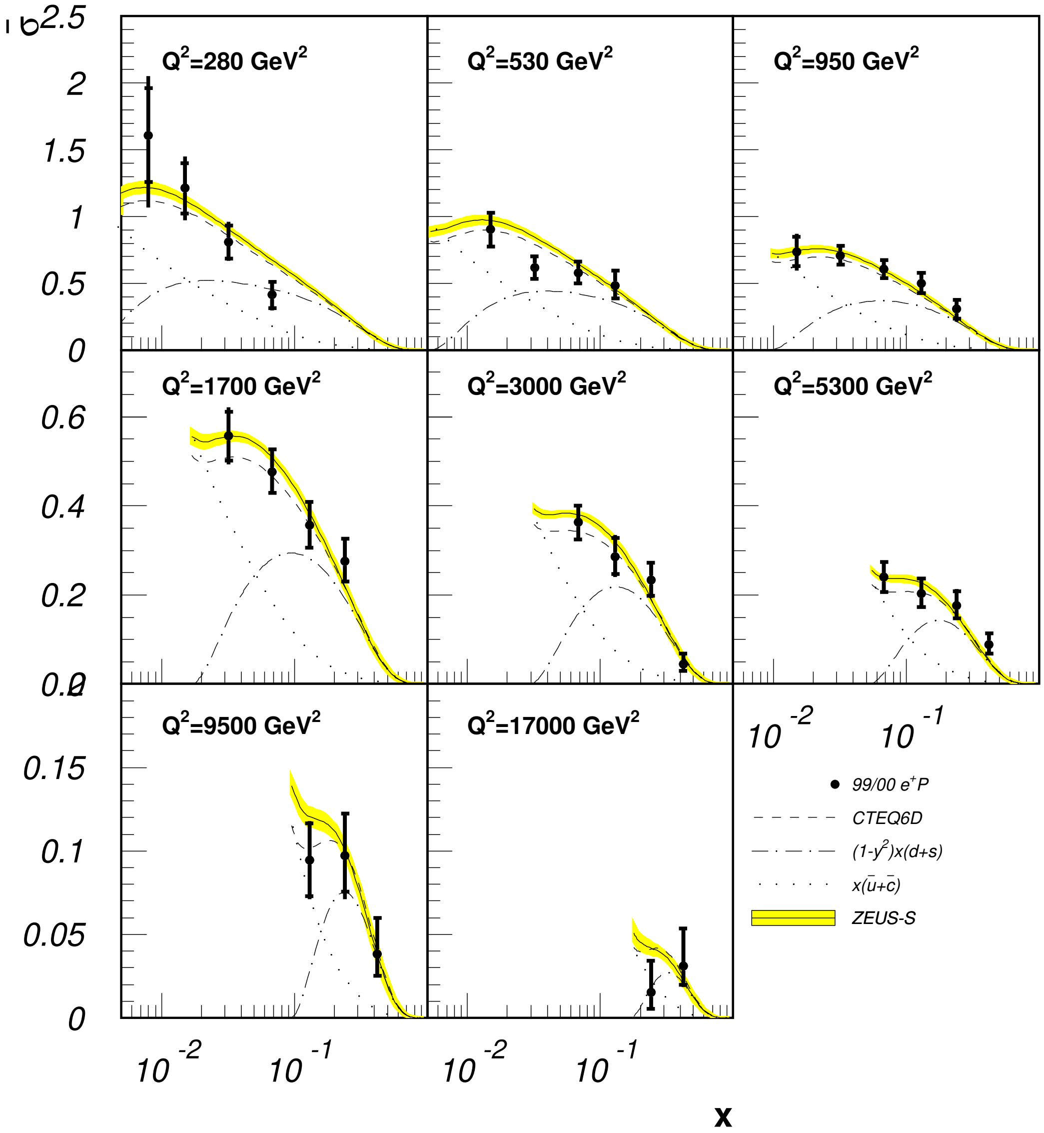

Figure 59: *The measured double differential reduced cross section $\tilde{\sigma}$, for $e^+p \rightarrow \overline{\nu_e}\, X$, in bins of fixed Q^2. The separate contributions of the PDF combinations $x(\bar{u}+\bar{c})$ and $(1-y)^2x(d+s)$ are also shown.*

$d\sigma/dQ^2$								
Q^2 range (GeV^2)	Q^2 (GeV^2)	N_{obs}	N_{bg}	$\delta_{unc}(\%)$	$\delta_1(\%)$	$\delta_2(\%)$	$\delta_3(\%)$	$d\sigma/dQ^2(\mathrm{pb}/\mathrm{GeV}^2)$
200-400	280	158	9.0	$^{+7.5}_{-7.7}$	$^{+1.4}_{-1.4}$	$^{+1.0}_{-0.7}$	± 2.7	$(2.92 \pm 0.26 \pm 0.24) \times 10^{-2}$
400-711	530	208	1.1	$^{+3.9}_{-4.3}$	$^{+1.5}_{-0.8}$	$^{+0.2}_{+0.5}$	∓ 1.8	$(1.84 \pm 0.13 \pm 0.09) \times 10^{-2}$
711-1265	950	321	3.2	$^{+3.1}_{-3.5}$	$^{+0.5}_{-1.0}$	$^{-0.2}_{-0.5}$	± 3.8	$(1.36 \pm 0.08 \pm 0.07) \times 10^{-2}$
1265-2249	1700	322	1.7	$^{+2.6}_{-2.3}$	$^{+0.5}_{-0.0}$	$^{+0.4}_{+0.9}$	∓ 1.0	$(7.04 \pm 0.41 ^{+0.21}_{-0.18}) \times 10^{-3}$
2249-4000	3000	241	0.4	$^{+2.0}_{-2.5}$	$^{-2.1}_{+1.5}$	$^{+0.5}_{-0.5}$	∓ 0.5	$(2.95 \pm 0.20 ^{+0.08}_{-0.10}) \times 10^{-3}$
4000-7113	5300	156	0.1	$^{+2.5}_{-1.7}$	$^{-1.5}_{+3.6}$	$^{+0.7}_{+1.4}$	± 0.0	$(1.08 \pm 0.09 ^{+0.05}_{-0.02}) \times 10^{-3}$
7113-12649	9500	57	0.2	$^{+2.4}_{-5.2}$	$^{-6.0}_{+5.5}$	$^{+0.9}_{-1.7}$	± 1.0	$(2.12 \pm 0.29 ^{+0.13}_{-0.17}) \times 10^{-4}$
12649-22494	17000	11	0.0	$^{+1.3}_{-2.7}$	$^{-9.2}_{+9.9}$	$^{+1.5}_{-2.4}$	∓ 6.0	$(2.00 ^{+0.73}_{-0.58} {}^{+0.24}_{-0.23}) \times 10^{-5}$
22494-60000	30000	3	0.0	$^{+3.2}_{-3.2}$	$^{-16}_{+18}$	$^{+6.3}_{-7.9}$	∓ 14	$(2.07 ^{+1.82}_{-1.31} {}^{+0.49}_{-0.47}) \times 10^{-6}$
$d\sigma/dx$								
x range	x	N_{obs}	N_{bg}	$\delta_{unc}(\%)$	$\delta_1(\%)$	$\delta_2(\%)$	$\delta_3(\%)$	$d\sigma/dx(\mathrm{pb})$
0.010-0.021	0.015	178	5.0	$^{+4.7}_{-4.9}$	$^{+2.0}_{-1.1}$	$^{+0.2}_{-0.3}$	± 4.9	$(4.86 \pm 0.39 \pm 0.34) \times 10^{2}$
0.021-0.046	0.032	354	5.3	$^{+3.2}_{-3.1}$	$^{+0.4}_{-0.4}$	$^{+0.1}_{+0.3}$	∓ 3.9	$(2.91 \pm 0.16 ^{+0.15}_{-0.14}) \times 10^{2}$
0.046-0.100	0.068	428	3.1	$^{+2.1}_{-2.0}$	$^{-0.6}_{+0.7}$	$^{+0.5}_{-1.0}$	∓ 0.6	$(1.60 \pm 0.08 \pm 0.04) \times 10^{2}$
0.100-0.178	0.130	258	0.1	$^{+2.4}_{-2.6}$	$^{-1.1}_{+1.0}$	$^{+1.0}_{-0.2}$	± 3.6	$(7.35 \pm 0.47 ^{+0.33}_{-0.34}) \times 10^{1}$
0.178-0.316	0.240	175	0.2	$^{+2.7}_{-3.5}$	$^{-2.2}_{+2.4}$	$^{-1.0}_{+1.9}$	± 1.7	$(3.06 \pm 0.24 \pm 0.14) \times 10^{1}$
0.316-0.562	0.420	45	0.0	$^{+3.2}_{-1.2}$	$^{-3.6}_{+4.3}$	$^{-2.6}_{+2.7}$	∓ 3.2	$(5.86 \pm 0.89 ^{+0.40}_{-0.33}) \times 10^{0}$
0.562-1.000	0.650	2	0.0	$^{+13}_{-3.0}$	$^{-5.1}_{+27}$	$^{-6.3}_{+26}$	∓ 18	$(4.11 ^{+7.71}_{-2.63} {}^{+1.80}_{-0.80}) \times 10^{-1}$
$d\sigma/dy$								
y range	y	N_{obs}	N_{bg}	$\delta_{unc}(\%)$	$\delta_1(\%)$	$\delta_2(\%)$	$\delta_3(\%)$	$d\sigma/dy(\mathrm{pb})$
0.00-0.10	0.05	263	4.2	$^{+2.9}_{-3.4}$	$^{+0.2}_{-0.7}$	$^{+0.5}_{-0.0}$	± 0.6	$(7.83 \pm 0.51 ^{+0.24}_{-0.28}) \times 10^{1}$
0.10-0.20	0.10	362	2.7	$^{+3.3}_{-2.9}$	$^{+0.1}_{-0.1}$	$^{-0.7}_{+0.4}$	± 1.7	$(7.00 \pm 0.38 ^{+0.26}_{-0.24}) \times 10^{1}$
0.20-0.34	0.27	318	2.7	$^{+3.2}_{-3.4}$	$^{-0.1}_{+0.3}$	$^{-0.5}_{+1.0}$	± 1.0	$(4.49 \pm 0.26 \pm 0.16) \times 10^{1}$
0.34-0.48	0.41	220	1.6	$^{+3.5}_{-4.6}$	$^{-0.2}_{+0.4}$	$^{-0.2}_{+0.1}$	± 0.8	$(3.37 \pm 0.23 ^{+0.12}_{-0.16}) \times 10^{1}$
0.48-0.62	0.55	151	2.4	$^{+3.0}_{-3.5}$	$^{-0.3}_{+0.8}$	$^{+1.5}_{-0.2}$	± 0.3	$(2.62 \pm 0.22 \pm 0.09) \times 10^{1}$
0.62-0.76	0.69	109	1.5	$^{+3.0}_{-3.6}$	$^{-2.0}_{+3.5}$	$^{+2.6}_{-1.9}$	∓ 5.1	$(2.23 \pm 0.22 ^{+0.16}_{-0.15}) \times 10^{1}$
0.76-0.90	0.83	54	0.9	$^{+5.3}_{-6.7}$	$^{-4.0}_{+4.4}$	$^{+3.1}_{-2.8}$	∓ 1.8	$(1.58 \pm 0.22 ^{+0.12}_{-0.13}) \times 10^{1}$

Table 6: *Values of the differential cross sections $d\sigma/dQ^2$, $d\sigma/dx$ and $d\sigma/dy$. The values listed are: the kinematic range of the measurement; the value of Q^2, x, y at which the cross section is quoted; the number of data events; the expected number of background events; the uncorrelated systematic uncertainty; the correlated systematic uncertainties $\delta_1, \delta_2, \delta_3$, defined in sections 4.4.4.2 & 4.4.4.3; and the measured cross section with statistical and total systematic uncertainties. The uncertainty on the measured luminosity of 2.25% is not included in the systematic uncertainty.*

$Q^2(\mathrm{GeV}^2)$	x	N_{obs}	N_{bg}	$\delta_{unc}(\%)$	$\delta_1(\%)$	$\delta_2(\%)$	$\delta_3(\%)$	$\tilde{\sigma}$
280	0.008	26	1.3	${}^{+9.0}_{-22}$	${}^{+4.7}_{-1.4}$	${}^{+0.7}_{-0.2}$	± 13	$(1.61 \pm 0.35\,{}^{+0.27}_{-0.41}) \times 10^{0}$
280	0.015	51	2.5	${}^{+7.6}_{-7.5}$	${}^{+0.4}_{-2.5}$	${}^{-0.7}_{-0.8}$	± 9.2	$(1.21 \pm 0.19\,{}^{+0.14}_{-0.15}) \times 10^{0}$
280	0.032	53	2.9	${}^{+9.1}_{-6.4}$	${}^{+1.9}_{-0.0}$	${}^{+1.0}_{-0.6}$	∓ 5.6	$(8.11 \pm 1.23\,{}^{+0.89}_{-0.69}) \times 10^{-1}$
280	0.068	23	1.8	${}^{+9.6}_{-15}$	${}^{-0.7}_{-0.7}$	${}^{+0.4}_{-3.4}$	± 2.3	$(4.13 \pm 0.97\,{}^{+0.41}_{-0.65}) \times 10^{-1}$
530	0.015	54	0.1	${}^{+5.5}_{-5.2}$	${}^{+1.0}_{-0.2}$	${}^{+0.3}_{-0.0}$	∓ 1.0	$(9.02 \pm 1.26\,{}^{+0.51}_{-0.48}) \times 10^{-1}$
530	0.032	59	0.7	${}^{+4.4}_{-4.0}$	${}^{+2.2}_{-0.5}$	${}^{-0.2}_{+1.4}$	∓ 3.6	$(6.18 \pm 0.83\,{}^{+0.39}_{-0.34}) \times 10^{-1}$
530	0.068	55	0.1	${}^{+4.7}_{-4.6}$	${}^{+0.9}_{-1.6}$	${}^{+0.7}_{-0.2}$	∓ 1.9	$(5.80 \pm 0.80\,{}^{+0.30}_{-0.31}) \times 10^{-1}$
530	0.130	27	0.0	${}^{+3.7}_{-5.2}$	${}^{+1.0}_{+2.7}$	${}^{-1.0}_{+1.7}$	∓ 5.9	$(4.84 \pm 0.96\,{}^{+0.37}_{-0.38}) \times 10^{-1}$
950	0.015	53	1.2	${}^{+5.3}_{-9.5}$	${}^{+0.7}_{-1.7}$	${}^{-0.4}_{-0.6}$	± 8.9	$(7.37 \pm 1.07\,{}^{+0.77}_{-0.97}) \times 10^{-1}$
950	0.032	107	1.1	${}^{+3.3}_{-4.5}$	${}^{+1.3}_{-0.6}$	${}^{+0.6}_{-0.0}$	± 1.8	$(7.10 \pm 0.71\,{}^{+0.29}_{-0.35}) \times 10^{-1}$
950	0.068	91	0.5	${}^{+2.6}_{-2.3}$	${}^{-0.8}_{-1.0}$	${}^{-1.9}_{-0.9}$	± 0.3	$(6.06 \pm 0.66\,{}^{+0.16}_{-0.21}) \times 10^{-1}$
950	0.130	49	0.0	${}^{+3.0}_{-3.9}$	${}^{-0.6}_{-1.2}$	${}^{+0.0}_{-1.5}$	± 6.5	$(5.00 \pm 0.74\,{}^{+0.36}_{-0.39}) \times 10^{-1}$
950	0.240	21	0.2	${}^{+5.4}_{-18}$	${}^{+4.0}_{-0.0}$	${}^{+3.6}_{+1.6}$	± 12	$(3.06 \pm 0.70\,{}^{+0.43}_{-0.65}) \times 10^{-1}$
1700	0.032	109	0.9	${}^{+3.5}_{-2.5}$	${}^{-0.1}_{+0.1}$	${}^{+0.5}_{+0.7}$	∓ 5.0	$(5.57 \pm 0.55\,{}^{+0.35}_{-0.31}) \times 10^{-1}$
1700	0.068	105	0.0	${}^{+2.9}_{-2.1}$	${}^{+0.8}_{-1.1}$	${}^{+0.7}_{+0.6}$	∓ 2.3	$(4.77 \pm 0.48\,{}^{+0.19}_{-0.15}) \times 10^{-1}$
1700	0.130	55	0.0	${}^{+3.2}_{-3.4}$	${}^{+0.7}_{-0.7}$	${}^{+0.7}_{+0.5}$	± 2.1	$(3.56 \pm 0.49\,{}^{+0.14}_{-0.14}) \times 10^{-1}$
1700	0.240	39	0.0	${}^{+3.2}_{-2.8}$	${}^{-2.4}_{+0.6}$	${}^{-0.6}_{+2.8}$	± 5.6	$(2.75 \pm 0.46\,{}^{+0.19}_{-0.19}) \times 10^{-1}$
3000	0.068	97	0.2	${}^{+1.7}_{-1.7}$	${}^{-3.0}_{+1.5}$	${}^{+1.2}_{-2.6}$	± 1.6	$(3.63 \pm 0.38\,{}^{+0.11}_{-0.17}) \times 10^{-1}$
3000	0.130	56	0.0	${}^{+3.6}_{-4.0}$	${}^{-1.9}_{+1.4}$	${}^{-0.9}_{+1.6}$	± 7.5	$(2.87 \pm 0.39\,{}^{+0.25}_{-0.25}) \times 10^{-1}$
3000	0.240	47	0.0	${}^{+2.5}_{-6.9}$	${}^{-2.3}_{+1.0}$	${}^{-4.3}_{+0.8}$	∓ 1.8	$(2.34 \pm 0.35\,{}^{+0.08}_{-0.20}) \times 10^{-1}$
3000	0.420	8	0.0	${}^{+3.7}_{-7.9}$	${}^{-5.0}_{-1.0}$	${}^{-2.6}_{-5.0}$	∓ 13	$(4.39\,{}^{+2.47}_{-1.51}\,{}^{+0.60}_{-0.75}) \times 10^{-2}$
5300	0.068	53	0.3	${}^{+1.3}_{-3.7}$	${}^{-2.2}_{+2.6}$	${}^{+0.6}_{-1.6}$	∓ 4.1	$(2.40 \pm 0.34\,{}^{+0.12}_{-0.15}) \times 10^{-1}$
5300	0.130	45	0.0	${}^{+2.8}_{-2.0}$	${}^{-0.6}_{+2.4}$	${}^{+2.0}_{-0.5}$	± 4.6	$(2.03 \pm 0.31\,{}^{+0.13}_{-0.10}) \times 10^{-1}$
5300	0.240	39	0.0	${}^{+7.6}_{-2.3}$	${}^{-2.1}_{+5.1}$	${}^{-0.9}_{+5.7}$	∓ 0.2	$(1.76 \pm 0.29\,{}^{+0.19}_{-0.06}) \times 10^{-1}$
5300	0.420	19	0.0	${}^{+6.4}_{-1.9}$	${}^{+0.5}_{+5.6}$	${}^{-1.6}_{+6.7}$	± 1.0	$(8.88 \pm 2.09\,{}^{+0.97}_{-0.24}) \times 10^{-2}$
9500	0.130	20	0.1	${}^{+3.9}_{-3.0}$	${}^{-7.5}_{+5.6}$	${}^{+5.0}_{-6.2}$	± 1.0	$(9.44 \pm 2.13\,{}^{+0.81}_{-0.96}) \times 10^{-2}$
9500	0.240	21	0.0	${}^{+2.1}_{-14}$	${}^{-4.0}_{+4.1}$	${}^{-3.2}_{+2.0}$	± 0.3	$(9.71 \pm 2.13\,{}^{+0.49}_{-1.47}) \times 10^{-2}$
9500	0.420	8	0.0	${}^{+6.2}_{-0.7}$	${}^{-2.7}_{+4.3}$	${}^{-1.4}_{+9.1}$	± 4.6	$(3.84\,{}^{+2.15}_{-1.30}\,{}^{+0.49}_{-0.21}) \times 10^{-2}$
17000	0.240	3	0.0	${}^{+4.2}_{-1.8}$	${}^{-8.1}_{+11}$	${}^{+7.1}_{-4.2}$	∓ 6.3	$(1.52\,{}^{+1.92}_{-0.97}\,{}^{+0.23}_{-0.17}) \times 10^{-2}$
17000	0.420	6	0.0	${}^{+1.8}_{-3.6}$	${}^{-9.3}_{+6.8}$	${}^{-8.2}_{+5.4}$	∓ 8.3	$(3.09\,{}^{+2.29}_{-1.12}\,{}^{+0.38}_{-0.47}) \times 10^{-2}$

Table 7: *Values of the reduced double differential cross section $\tilde{\sigma}$. The values listed are: the value of Q^2, x at which the cross section is quoted; the number of data events; the expected number of background events; the uncorrelated systematic uncertainty; the correlated systematic uncertainties $\delta_1, \delta_2, \delta_3$, defined in sections 4.4.4.2 & 4.4.4.3; and the measured cross section with statistical and total systematic uncertainties. The uncertainty on the measured luminosity of 2.25% is not included in the systematic uncertainty.*

5 Parton Distribution Function fitting

The paper [80] uses ZEUS and fixed-target data to investigate the quark and gluon densities of the proton. This is done by performing a next-to-leading-order QCD analysis, which fits assumed parton distributions to the observed cross-sections and values of F_2. In the paper, two fits are reported which initially will be replicated in this analysis, as a consistency check. The ZEUS-Standard (or ZEUS-S) fit is considered to be the most accurate since it includes ZEUS data and all fixed-target data where correlated systematic uncertainties have been published. The next fit to be replicated is the ZEUS-Only (or ZEUS-O) fit. This fit includes more ZEUS data than the ZEUS-S fit, but has no fixed-target data. Consequently it has smaller systematic uncertainties, but suffers from much larger statistical uncertainties. This analysis then moves on from replicating the fits presented in [80] to include the 1999-2000 ZEUS data and to investigate its consequences.

Before entering into the details of the replicated and new fits performed in this analysis, the method used to perform the fits must first be explained.

5.1 Fitting method

There are several methods of fitting experimentally obtained reduced cross sections and structure functions to obtain PDFs. These methods all involve the evaluation and minimisation of χ^2, calculated using the experimental data and the results of QCD analysis on a set of assumed PDF distributions. The PDFs are parameterised, and these parameters allowed to vary in order to minimise χ^2. The various methods of fitting differ in how they include correlated systematic errors (from the experimental data) in the minimisation. A good account of these methods can be found in [84].

In this analysis, MINUIT [85] has been used in the χ^2 minimisation, and the method introduced by Pascaud and Zomer [86], (also used in the ZEUS analysis [80]), has been used to fit the theoretical PDF parameterisation $f(x)$ in equation (55) to various datasets. The PDFs are parameterised in x using four parameters $\{p_i\}$ as follows:

$$xf(x) = p_1 x^{p_2}(1-x)^{p_3}(1+p_5 x) \tag{55}$$

This form allows freedom in low (p_2), high (p_3) and middling (p_5) x values. The *momentum* distributions $xf(x)$ parameterised are: $xu_v(x)$ u-valence; $xd_v(x)$ d-valence; $xS(x)$ total sea; $xg(x)$ gluon; and the difference between $\bar{d}$ and $\bar{u}$ contributions to the sea $x(\bar{d}-\bar{u}) \equiv x\Delta(x)$ (though this last distribution $x\Delta(x)$ only has any free parameters p_i in the replication of the global ZEUS-S fit, to be described in §5.2.1).

The momentum distributions (55) are evolved in Q^2 using QCDNUM [87] from an input value of $Q_0^2 = 7\,\mathrm{GeV}$, to give values over a grid of x, Q^2 points[6]. The choice of $Q_0^2 = 7$ is not significant; the input value of Q_0^2 has been varied in §5.9 to study its effect. In any case, QCDNUM performs the evolution to lower as well as higher Q^2 values. The evolution itself is performed using the DGLAP equations, and theoretical values for the structure functions F_2, F_L and xF_3 are calculated at NLO in the $\overline{\mathrm{MS}}$ scheme [88–90], with renormalisation and mass factorisation scales equal to Q^2.

The calculated structure function values depend on the treatment of heavy quarks in the sea. QCDNUM provides the option of (1) using a fixed flavour number scheme (FFN), where the number of flavours is fixed to 3, and charm is produced solely by boson-gluon fusion. In this scheme, there is no concept of a charm-quark parton distribution. (2) QCDNUM can generate the sea allowing the number of effective flavours to vary, but heavy quarks are only generated above their respective mass thresholds (i.e. they are treated as though they are massless).

[6] QCDNUM actually evolves the flavour singlet, gluon and non-singlet distributions. These were calculated at each stage of evolution and passed as arguments to QCDNUM.

This is called the zero-mass variable flavour number scheme (ZM-VFN). Both schemes have undesirable features - the FFN scheme fails to describe data for Q^2 far above the charm threshold. The ZM-VFN scheme (whilst being close to correct at large Q^2), fails in the threshold region for charm production. In this analysis the Thorne-Roberts non-zero mass variable flavour number scheme (TR-VFN) [91–94], was used (as was also done in [80]), to interpolate between Fixed Flavour Number behaviour of the sea (appropriate at $Q^2 \lesssim m_c^2$) and Zero Mass Variable Flavour Number behaviour (at $Q^2 \gg m_c^2$). The justification for use of the TR-VFN scheme for fits such as those performed in this analysis was largely laid out in [95]. The TR-VFN scheme prescribes that the evolution is performed by QCDNUM in the same way as is done in ZM-VFN, (charm quark spontaneously introduced at threshold), however, proprietary Thorne-Roberts routines are used to calculate the correct coefficient functions. These are then convoluted with the PDFs to obtain the final structure functions used in the fit.

Consistent with [80] and [96], the total sea $xS(x)$ at Q_0^2 is comprised of the up $xu_{sea}(x)$, down $xd_{sea}(x)$, strange $xs_{sea}(x)$ and charm $xc_{sea}(x)$ flavours in the following proportions:

$$xs_{sea}(x) = 0.2xS(x) \tag{56}$$

$$xu_{sea}(x) = 0.4xS(x) - 0.5xc_{sea}(x) - x\Delta(x) \tag{57}$$

$$xd_{sea}(x) = 0.4xS(x) - 0.5xc_{sea}(x) + x\Delta(x) \tag{58}$$

where $u_{sea}, d_{sea}, s_{sea}, c_{sea}$ include both quark and antiquark contributions to the sea. The charmed sea is generated as described by the TR-VFN scheme.

With the details of the structure function calculations having now been examined, it is necessary to explain how the data and calculated values are compared in the fits using χ^2, which is formed as follows:

$$\chi^2 = \sum_i \frac{[F_i(p,s) - F_i(meas)]^2}{\sigma_{i,stat}^2 + \sigma_{i,uncor}^2} + \sum_\lambda s_\lambda^2 \tag{59}$$

where

$$F_i(p,s) = F_i^{NLOQCD}(p) \;/\; \left(1 + \sum_\lambda \Delta_{i\lambda}^{sys} s_\lambda\right) \tag{60}$$

and $F_i(meas)$ represents a measured data point. The symbols $\sigma_{i,stat}$ and $\sigma_{i,uncor}$ represent the statistical and uncorrelated systematic errors on the data point i respectively. $F_i^{NLOQCD}(p)$ represents the prediction from NLO QCD using the parameters p (from the definition of the PDFs). $F_i^{NLOQCD}(p)$ is modified by the inclusion of correlated systematic errors as shown in equation (60). The one standard deviation uncertainty on a data point i, due to a source λ, is labelled $\Delta_{i\lambda}^{sys}$, and the parameters s_λ represent, for each source λ, independent Gaussian variables with a mean of zero and variance equal to one.

Initially, the parameters s_λ are set to zero (this includes only statistical and uncorrelated errors in χ^2). The minimisation is then performed to obtain the 'central values' ($p_{1..5}$) of the fit for each momentum distribution.

The Hessian matrix M_{ij} is then evaluated with respect to the parameters p;

$$M_{ij} = \frac{1}{2}\frac{\partial^2\chi^2}{\partial p_i \partial p_j}. \tag{61}$$

In the error analysis, the parameters s_λ are allowed to vary, and so a second Hessian matrix $C_{j\lambda}$ is also evaluated with respect to the central value parameters p and the correlated systematic error parameters s;

$$C_{j\lambda} = \frac{1}{2}\frac{\partial^2\chi^2}{\partial p_j \partial s_\lambda}. \tag{62}$$

The relevant covariance matrices V are then evaluated as described by Pascaud and Zomer [86]. The statistical (and uncorrelated systematic) covariance matrix

$V^{stat}(= M^{-1})$. The correlated systematic covariance matrix $V^{sys}(= M^{-1}CC^TM^{-1})$, and the total covariance matrix $V^{tot}(= V^{stat} + V^{sys})$.

The uncertainty on a quantity 'G' which is a function of the theoretical parameters p (for example a PDF) is then calculated by propagating the uncertainty:

$$< \Delta G^2 >= \sum_i \sum_j \frac{\partial G}{\partial p_i} V_{ij} \frac{\partial G}{\partial p_j} \tag{63}$$

where V_{ij} could be V^{stat}, V^{sys} or V^{tot} depending on the desired error band.

The fitting procedure described above thus provides a method of determining PDFs and their uncertainties using experimental data in the form of reduced cross sections and structure functions. The method will now be used to replicate two fits from [80] and also to perform new fits which include the 1999-2000 ZEUS data.

5.2 Replicating ZEUS-S and ZEUS-O fits

The ZEUS-S and ZEUS-O fits from the paper [80] will be used as a starting point for further work. To act as a consistency check, these fits will initially be replicated. A brief summary of these fits and the results obtained now follows.

5.2.1 ZEUS-S fit

The first of the two fits to be replicated is the ZEUS-S, or 'ZEUS-Standard' fit. This is considered to be the 'global' fit, as it incorporates (as well as ZEUS data) all fixed target data sets where full information on correlated systematic uncertainties is available:

- ZEUS Neutral Current F_2 e^+P data (1996-1997) [97].
- muon-proton scattering F_2 data from BCDMS [98] , E665 [99] and NMC [100].
- muon-deuterium data from NMC [100] and E665 [99].
- NMC data on F_2^{deut}/F_2^p obtained from simultaneous measurement of muons incident on symmetric targets of deuterium and hydrogen [101].
- $\nu, \bar{\nu}$-Fe xF_3 data from CCFR [102] in the range $0.1 \leq x \leq 0.75$.

The heavy target corrections to the CCFR data were made as suggested by MRST [103], according to the value of x. The deuterium data (E665, NMC and the NMC F_2^{deut}/F_2^p) were also corrected to the average of the free proton and neutron structure functions according to the value of x as prescribed by Gomez et al. [104].

The following cuts were made on the data:

- $Q^2 \geq 2.5\,\mathrm{GeV}^2$ to ensure perturbative QCD is applicable.
- $W^2 > 20\,\mathrm{GeV}^2$. W^2 is the square of the centre-of-mass energy of the boson-proton system,

$$W^2 = (p+q)^2 = m_p^2 - Q^2 + 2p.q \tag{64}$$

$$= m_p^2 + Q^2\frac{(1-x)}{x}. \quad (65)$$

This cut aids in reduction of nuclear mass effects [105], and also reduces the effect of higher twist (the fits were performed at leading twist), which becomes more prevelant at low W^2 [106].

- As noted above, CCFR data was restricted to the range $0.1 \leq x \leq 0.75$.

The data sets used were those published, apart from the BCDMS data, which was reduced by a factor of 2%. This has been common practice amongst other fitting groups for some time [103,106–109]. The normalisation uncertainties were included among the correlated systematic uncertainties.

In the fit, when calculating the distributions for $u_v(x,Q^2)$, $d_v(x,Q^2)$, $S(x,Q^2)$, $g(x,Q^2)$ the following parameters were fixed as was done with the previous fits in [80]:

- p_1u_v and p_1d_v were fixed by using the QCD number sum rules.
- p_1g was fixed using the momentum sum rule.
- p_2u_v and p_2d_v were fixed to a value of 0.5. These parameters govern the low-x behaviour of the u, d-valence distributions. Removal of the CCFR data for $x < 0.1$ leaves little data to determine the valence distributons at low-x. When the parameters are left free, values consistent with 0.5 are obtained.
- In the $x\Delta$ distribution, all parameters except the overall normalisation $p_1\Delta$ are fixed. $p_2\Delta = 0.5$, $p_3\Delta = p_3S + 2$ and $p_5\Delta = 0$.
- p_5g was set equal to zero. This choice maintains a positive gluon distribution at high-x, and historically stabilised the ZEUS-α_s fit in [80].

There are, therefore, a total of eleven free parameters in the ZEUS-S fit. The choice to fix $p_5g = 0$ was made in [80] essentially to stabilise fits where α_s was free, but is not absolutely necessary to fix $p_5g = 0$ when α_s is fixed. This choice has been controversial since the H1 fits (e.g. [110]) have a humpy shaped gluon, which requires a non-zero value of p_5g. In any case, it has negligible effect on the 'S-fit' detailed here. When it is freed, its value is $-0.5 \pm 1.9 \pm 1.1$ which is consistent with zero. Much more significant is the effect of p_5g on fits with ZEUS-only data, such as the fit performed in the next section. The effect of p_5g on this type of fit will be investigated in §5.7.

5.2.2 ZEUS-O fit

The second fit which this analysis seeks to replicate is the ZEUS-O or 'ZEUS-Only' fit. As might be expected, this fit contains only ZEUS data. The value of this fit lies in the removal of the large systematic uncertainties associated with the fixed target corrections, which are present in the ZEUS-S fit. The same data cuts were applied as in §5.2.1, but this fit now contains the following data sets:

- ZEUS Neutral Current F_2 e^+P (1996-1997 data) [97].
- ZEUS Neutral Current e^-P (1998-1999 data) [111]
- ZEUS Charged Current e^+P (1994-1997 data) [112] and [113].
- ZEUS Charged Current e^-P (1998-1999 data) [66].

The parameters entering into the fit were the same, apart from $p_1\Delta$, which was fixed to the value obtained from the 'S' fit. This is described in the ZEUS paper [80]; the exclusion of the fixed target data removes all constraints on the flavour content of the sea. This leaves ten free parameters in the ZEUS-O fit.

5.3 Results of the ZEUS-S and ZEUS-O fit replication

As can be seen by comparing figure 60 and also table 10 to the published fit in [80], the 'S' fit results are an excellent replication of the published data.

The 'O' fit results also replicate those produced in the paper [80]. Though the central fit values were not previously published for the ZEUS-O fit, they have been checked and found to be in agreement by private communication with the author of [80].

As can be seen from the two plots comparing the 'S' and 'O' fits (Figures 60 and 61), the u and d-valence distributions are more poorly determined in the 'O' fit, by a factor of about two. The d-valence distribution in particular has large bands of statistical uncertainty around the peak. This distribution is determined by high-Q^2 charged current e^+P data, so the addition of the 1999-2000 ZEUS data to this fit in §5.4 is of prime interest. It should be noted that the fixed target corrections made to the 'S' fit are eliminated in the 'O' fit, making the systematic uncertainty smaller in the 'O' fit, although the total uncertainty is larger.

In addition, it can be seen that the 'O' fit d-valence distribution is higher than the 'S' fit (and MRST(2001) [82]) distribution for high x, though the limited statistical precision of the 'O' fit means that it is still consistent with the 'S' fit. The u-valence distribution is also higher in the 'O' fit, but to a less pronounced extent. The differences arise from the central values of the fit ($p_{1..5}$), which are determined by the χ^2 minimisation. It is not altogether surprising that the fits have differing central values (and uncertainties), as they contain almost completely different data sets. As might be expected from the close replication of the two fits from the ZEUS analysis [80], these valence PDF differences are also present when the published ZEUS-S and ZEUS-O fits are compared, however they are investigated more fully in §5.5.

Figures 62 and 63 compare the 'S' and 'O' fit d-valence distributions at various Q^2 values using a linear abscissa and a logarithmic ordinate axis to emphasise the high-x region. As can be seen, the high-x uncertainty is greater for the 'O' fit. The 'O' fit uncertainty is also dominated by the uncorrelated and statistical contributions. This is in contrast to the high-x 'S' fit uncertainty, which is dominated by the correlated systematic uncertainties.

Looking at the high-x sea (as is done in figure 64, again using a linear abscissa and logarithmic ordinate axis to emphasise the high-x region), it is clear that the 'O' fit differs substantially from the MRST (and 'S' fit) predictions. This occurs because the fixed target data (BCDMS, NMC D/p ratio and CCFR xF_3) is excluded from the 'O' fit. One way around this problem is to fix the high-x sea parameter p_3S in the 'O' fit to the value obtained in the replication of the ZEUS-S fit (the 'S' fit), in which there was much more high-x sea information. This will be explored in §5.5 (for the current fit, the value of p_3S remains free). Another way of addressing the high-x sea difference is investigated in §5.6, where the BCDMS hydrogen data alone is added to the 'O' fit in order to stabilise the high-x sea.

<table>
<tr><th>Data
Set</th><th>No. of data
points</th><th>χ^2</th><th>χ^2 per
data point</th><th>No. of correlated
systematic uncertainties</th><th>No. of normalisation
uncertainties</th></tr>
<tr><td>ZEUS $F_2$96/97</td><td>242</td><td>203.10</td><td>0.84</td><td>10</td><td>2</td></tr>
<tr><td>BCDMS</td><td>305</td><td>288.22</td><td>0.94</td><td>5</td><td>5</td></tr>
<tr><td>NMC p</td><td>218</td><td>263.81</td><td>1.21</td><td rowspan="2">12</td><td rowspan="2">4</td></tr>
<tr><td>NMC D</td><td>218</td><td>201.18</td><td>0.92</td></tr>
<tr><td>NMC D/p</td><td>129</td><td>121.05</td><td>0.94</td><td>5</td><td>0</td></tr>
<tr><td>E665 p</td><td>47</td><td>54.459</td><td>1.16</td><td rowspan="2">7</td><td>2*</td></tr>
<tr><td>E665 D</td><td>47</td><td>44.135</td><td>0.94</td><td>1*</td></tr>
<tr><td>CCFR xF_3</td><td>57</td><td>22.557</td><td>0.40</td><td>18</td><td>0</td></tr>
<tr><td>TOTAL</td><td>1263</td><td>1198.51</td><td>0.95</td><td>57</td><td>14</td></tr>
</table>

Table 8: *χ^2 values for the data sets in the 'S' fit, obtained by adding all systematic and statistical uncertainties in quadrature at the minimum found by using the χ^2 of section 5.1. The systematic uncertainties for the p and D sets of NMC and E665 were taken together. Similarly, the normalisation uncertainties were the same for NMC p and D data. The star indicates that for E665, there is one common normalisation uncertainty, and one normalisation uncertainty each, for the targets p, D.*

<table>
<tr><th>Data
Set</th><th>No. of data
points</th><th>χ^2</th><th>χ^2 per
data point</th><th>No. of correlated
systematic uncertainties</th><th>No. of normalisation
uncertainties</th></tr>
<tr><td>ZEUS $F_2$96-97</td><td>242</td><td>168.26</td><td>0.70</td><td>10</td><td>2</td></tr>
<tr><td>ZEUS CC e^+P94-97</td><td>29</td><td>16.120</td><td>0.56</td><td>3</td><td>1</td></tr>
<tr><td>ZEUS NC e^-P98-99</td><td>92</td><td>60.477</td><td>0.66</td><td>6</td><td rowspan="2">1</td></tr>
<tr><td>ZEUS CC e^-P98-99</td><td>26</td><td>23.336</td><td>0.90</td><td>3</td></tr>
<tr><td>TOTAL</td><td>389</td><td>268.19</td><td>0.69</td><td>22</td><td>4</td></tr>
</table>

Table 9: *χ^2 values for the data sets in the 'O' fit, again obtained by adding all systematic and statistical uncertainties in quadrature at the minimum found by using the χ^2 of section 5.1.*

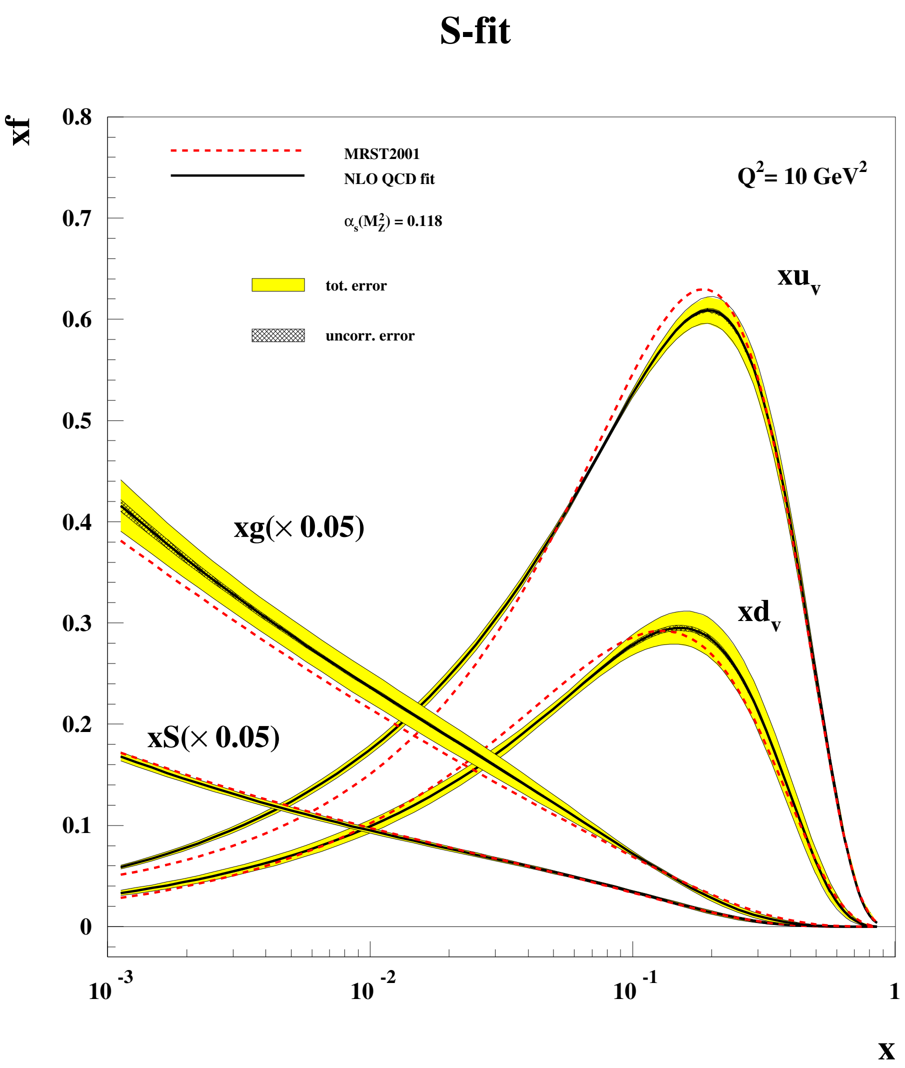

Figure 60: *The replication of the ZEUS-S fit, with MRST(2001) curves overlaid for reference.*

PDF	p_1	p_2	p_3	p_5
xu_v	*1.69	*0.5	$3.99 \pm 0.01 \pm 0.08$	$5.03 \pm 0.09 \pm 0.68$
xd_v	*0.96	*0.5	$5.35 \pm 0.10 \pm 0.48$	$6.2 \pm 0.4 \pm 2.3$
xS	$0.608 \pm 0.007 \pm 0.050$	$-0.233 \pm 0.002 \pm 0.012$	$8.9 \pm 0.2 \pm 1.2$	$6.8 \pm 0.4 \pm 2.1$
xg	*1.89	$-0.19 \pm 0.01 \pm 0.04$	$6.6 \pm 0.2 \pm 1.3$	*0.0
$x\Delta$	$0.27 \pm 0.01 \pm 0.07$	*0.5	*10.9	*0.0

Table 10: *PDF parameters at Q_0^2 from the replication of the ZEUS-S fit. The first uncertainty printed arises from the statistical and uncorrelated errors. The second uncertainty is the contribution from correlated systematic sources. Starred values were either fixed at the start of the fit or derived from the free parameters of the fit.*

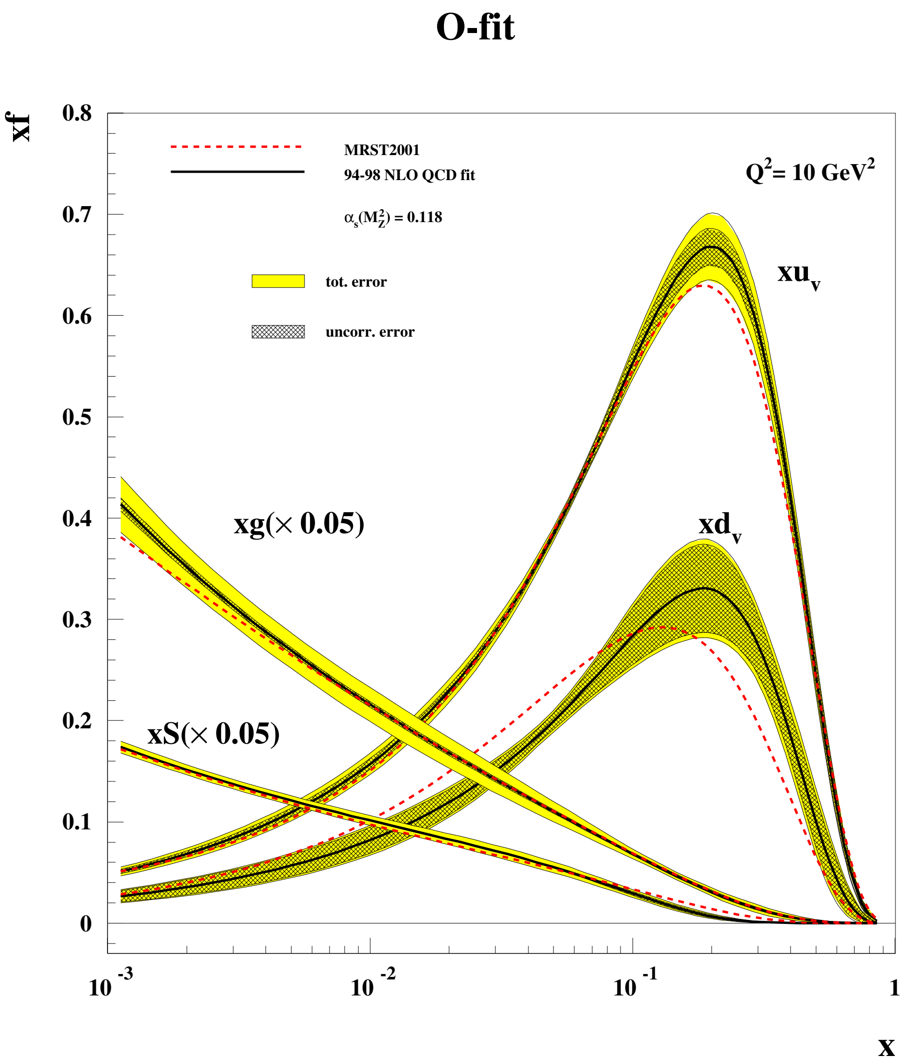

Figure 61: *The replication of the ZEUS-O fit, with MRST(2001) curves overlaid for reference. The d, u-valence distributions are higher at high-x than the ZEUS-S and MRST(2001) predictions.*

PDF	p_1	p_2	p_3	p_5
xu_v	*1.46	*0.5	$4.45 \pm 0.21 \pm 0.16$	$9.12 \pm 1.62 \pm 2.06$
xd_v	*0.76	*0.5	$4.87 \pm 0.97 \pm 0.57$	$9.7 \pm 8.1 \pm 4.0$
xS	$0.697 \pm 0.018 \pm 0.076$	$-0.220 \pm 0.004 \pm 0.017$	$15.0 \pm 1.7 \pm 2.5$	$14.5 \pm 3.2 \pm 7.7$
xg	*1.32	$-0.24 \pm 0.01 \pm 0.05$	$5.0 \pm 0.4 \pm 1.4$	*0.0
$x\Delta$	*0.27	*0.5	*17.0	*0.0

Table 11: *PDF parameters at Q_0^2 from the replication of the ZEUS-O fit. The first uncertainty printed arises from the statistical and uncorrelated errors. The second uncertainty is the contribution from correlated systematic sources. Starred values were either fixed at the start of the fit or derived from the free parameters of the fit.*

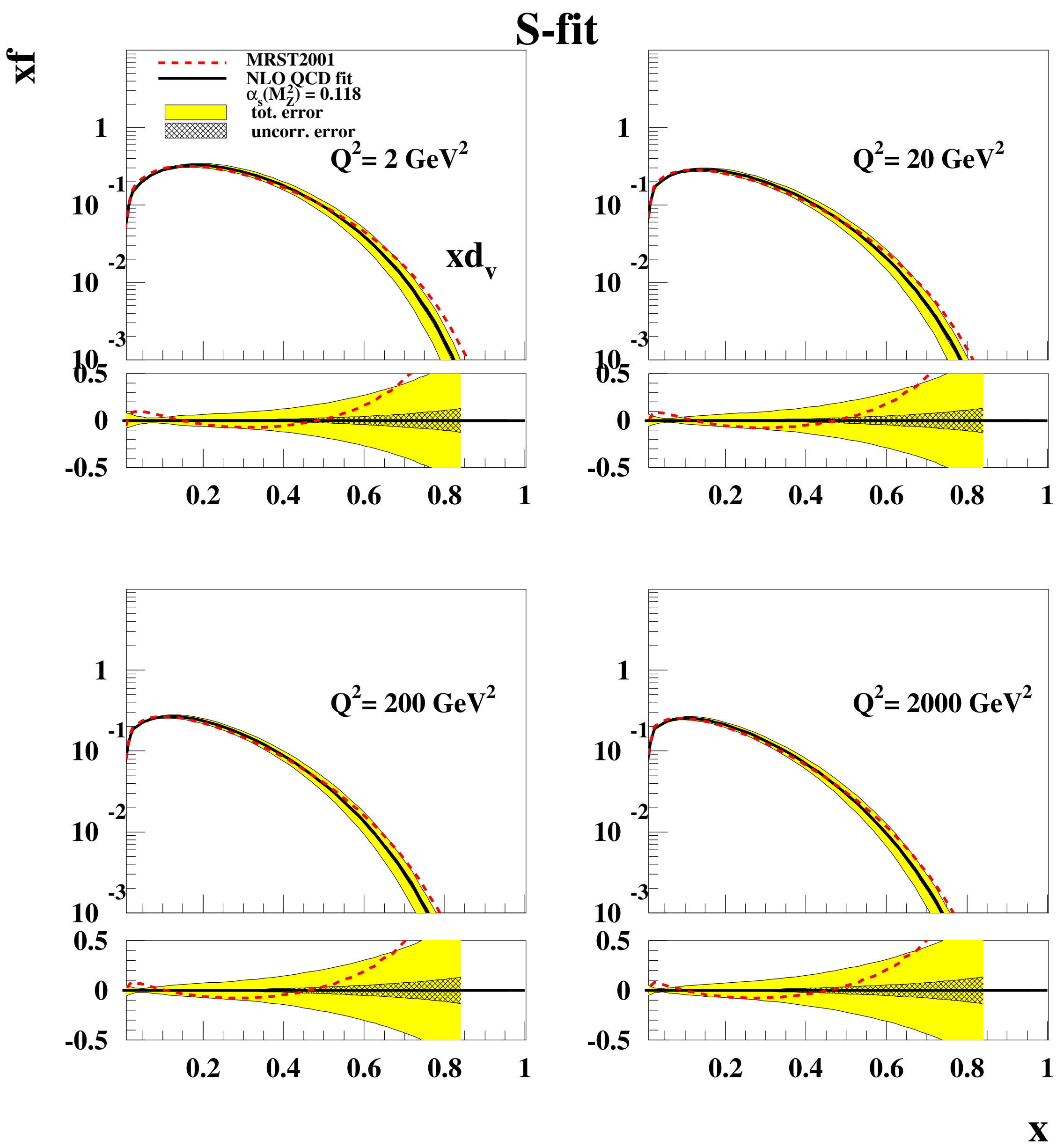

Figure 62: *The 'S' fit, showing d-valence plots on different scales, at various Q^2 values. The uncertainties are also shown for each Q^2 as fractional differences from the central value.*

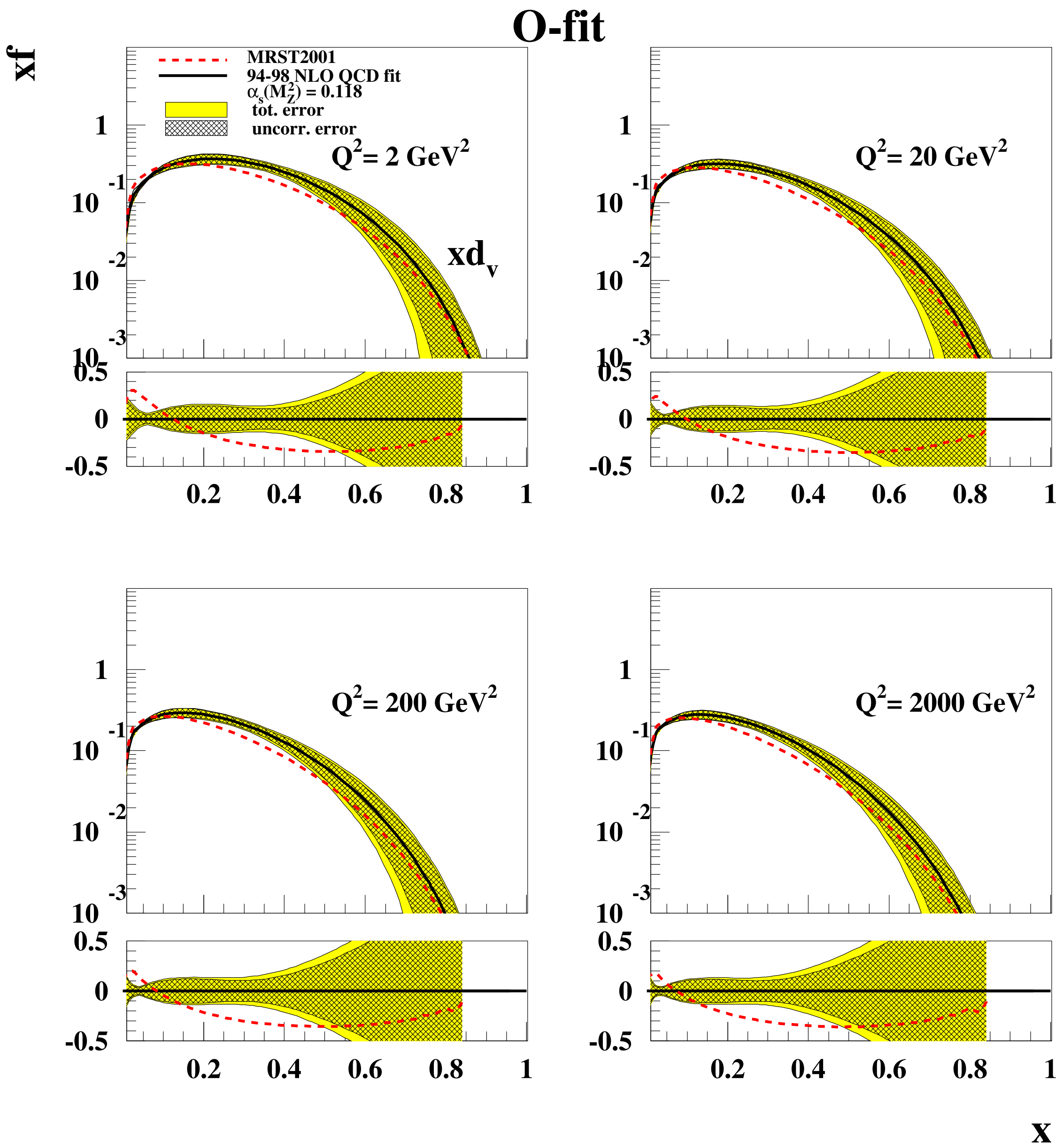

Figure 63: *The replication of the ZEUS-O fit, showing d-valence plots on different scales, at various Q^2 values. The uncertainties are shown for each Q^2 as fractional differences from the central value.*

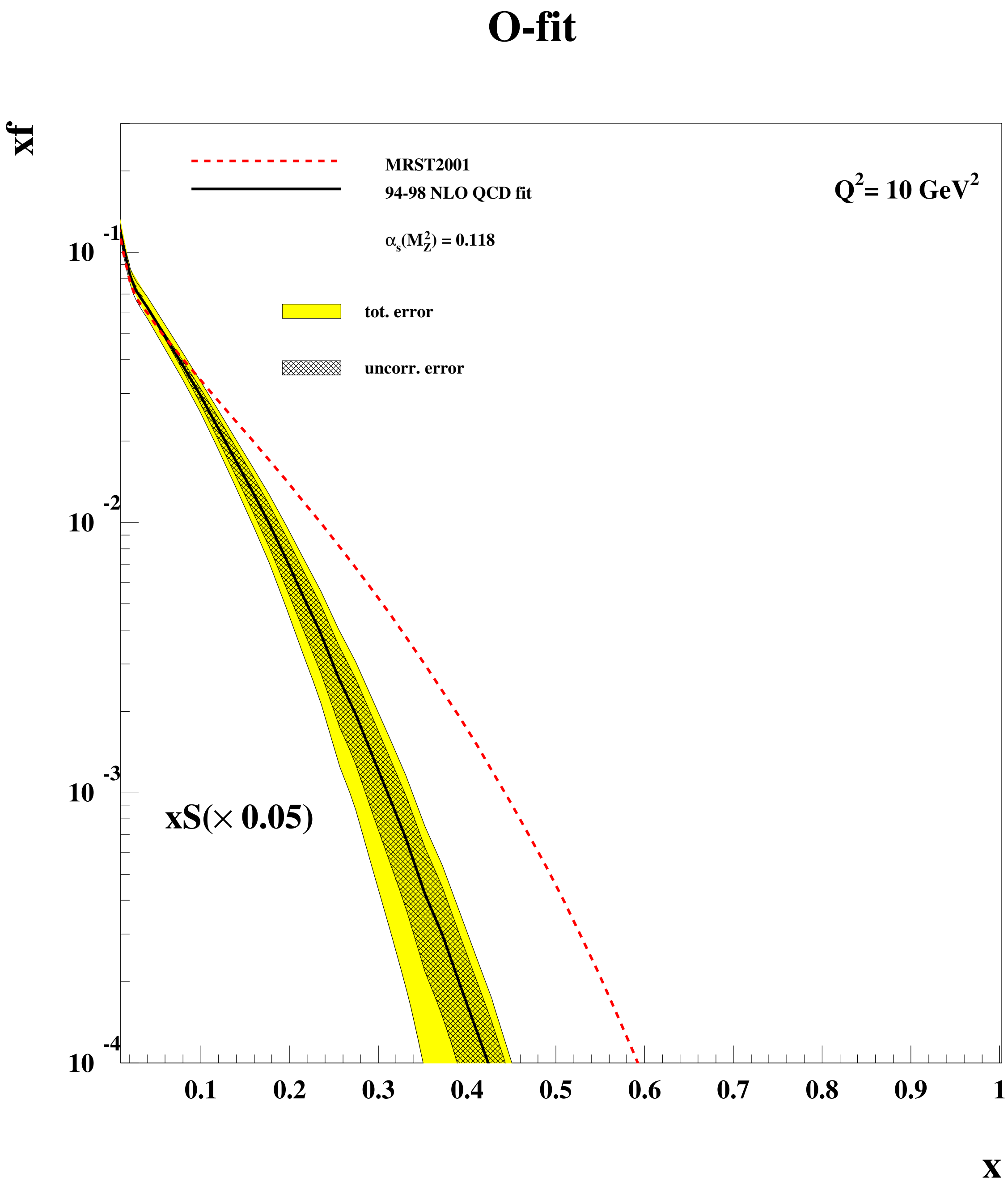

Figure 64: *The replication of the ZEUS-O fit, showing the behaviour of the high-x sea.*

5.4 The 1999-2000 data

Having successfully replicated the published ZEUS-S and ZEUS-O fits, the 1999-2000 ZEUS data are now included, in order to allow more accurate fits to be made. As mentioned in §5.2.2, the value of the ZEUS-O fit lies in the removal of the large systematic uncertainties associated with the fixed target corrections, which are present in the ZEUS-S fit. Since the ZEUS data are largely dominated by statistical uncertainty, it is hoped that the existing ZEUS-O fit can be improved by the addition of new ZEUS data.[7] In particular, it was noted in §5.3 that the d-valence distribution suffered from larger statistical uncertainties in the 'O' fit, and that the uncertainty could be reduced by the addition of e^+P charged current data. To this end, the ZEUS-O fit was repeated, but this time with the following extra data;

- ZEUS Charged Current e^+P (1999-2000 data) [83]
- ZEUS Neutral Current e^+P (1999-2000 data) [114]

making a total of six data sets in the fit. The data cuts and list of free parameters were kept the same as the original ZEUS-O fit (§5.2.2).

The data sets in the 1994-2000 'O' fit are listed in table 12 and also appear in figures 65 and 66. These figures also show the reduced neutral and charged current cross sections predicted by the fit. There is generally good agreement between data and fit prediction.

Figure 67 and table 13 show that the 1994-2000 'O' fit has central values which are comparable to the 1994-1998 'O' fit in figure 61. Looking at the uncertainties, the results of the new 1994-2000 'O' fit are encouraging; comparing the new 1994-2000 'O' fit (figure 68) to the 1994-1998 'O' fit (figure 63), we can see improvement in the d-valence fit uncertainty. So much improvement in fact, that

[7]It might be asked why these data are not added to the ZEUS-S fit. This has been done, but only provides marginal improvement to the fit.

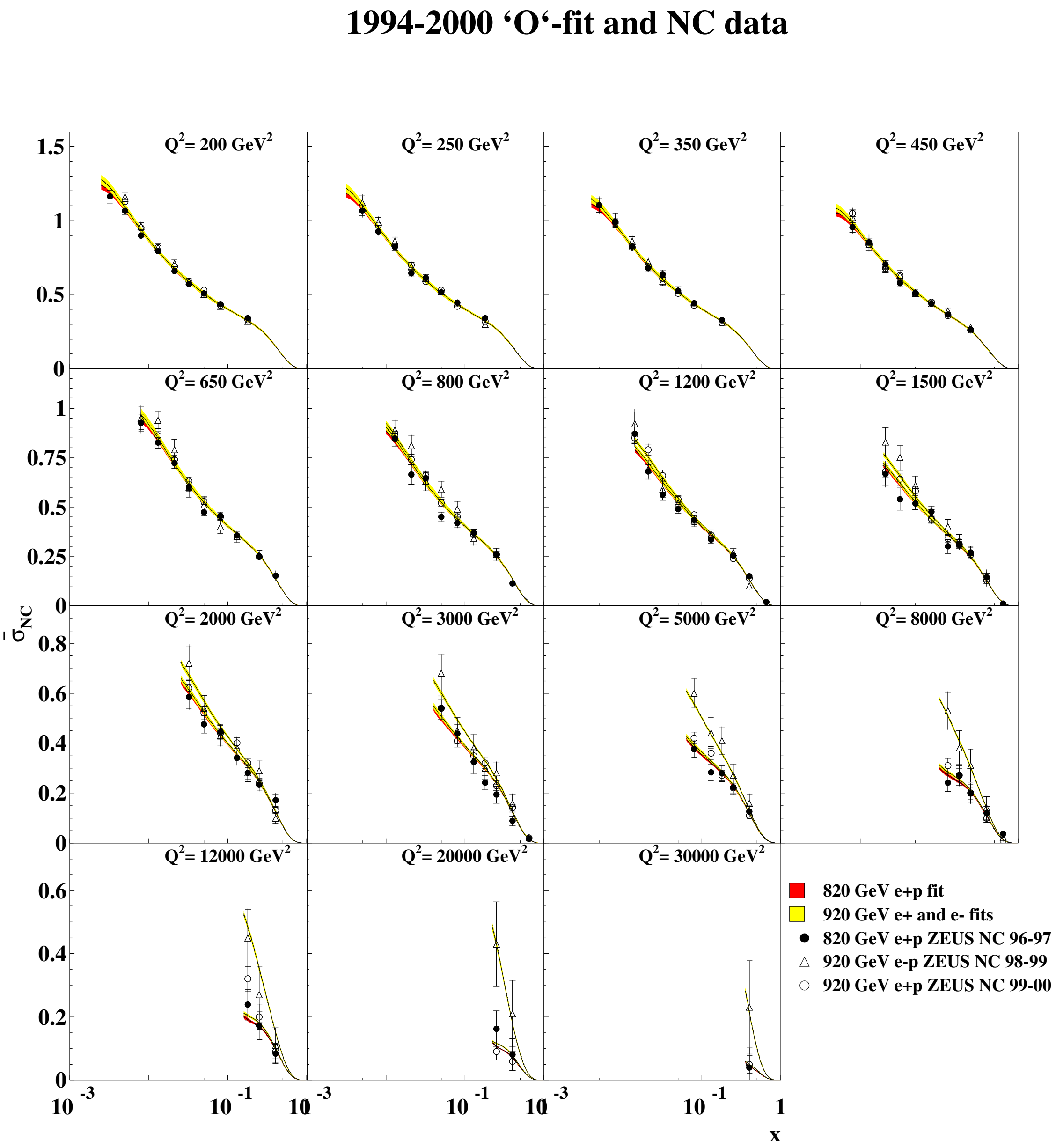

Figure 65: *The ZEUS reduced neutral current cross section data and 1994-2000 ‘O’ fit predictions in bins of fixed Q^2.*

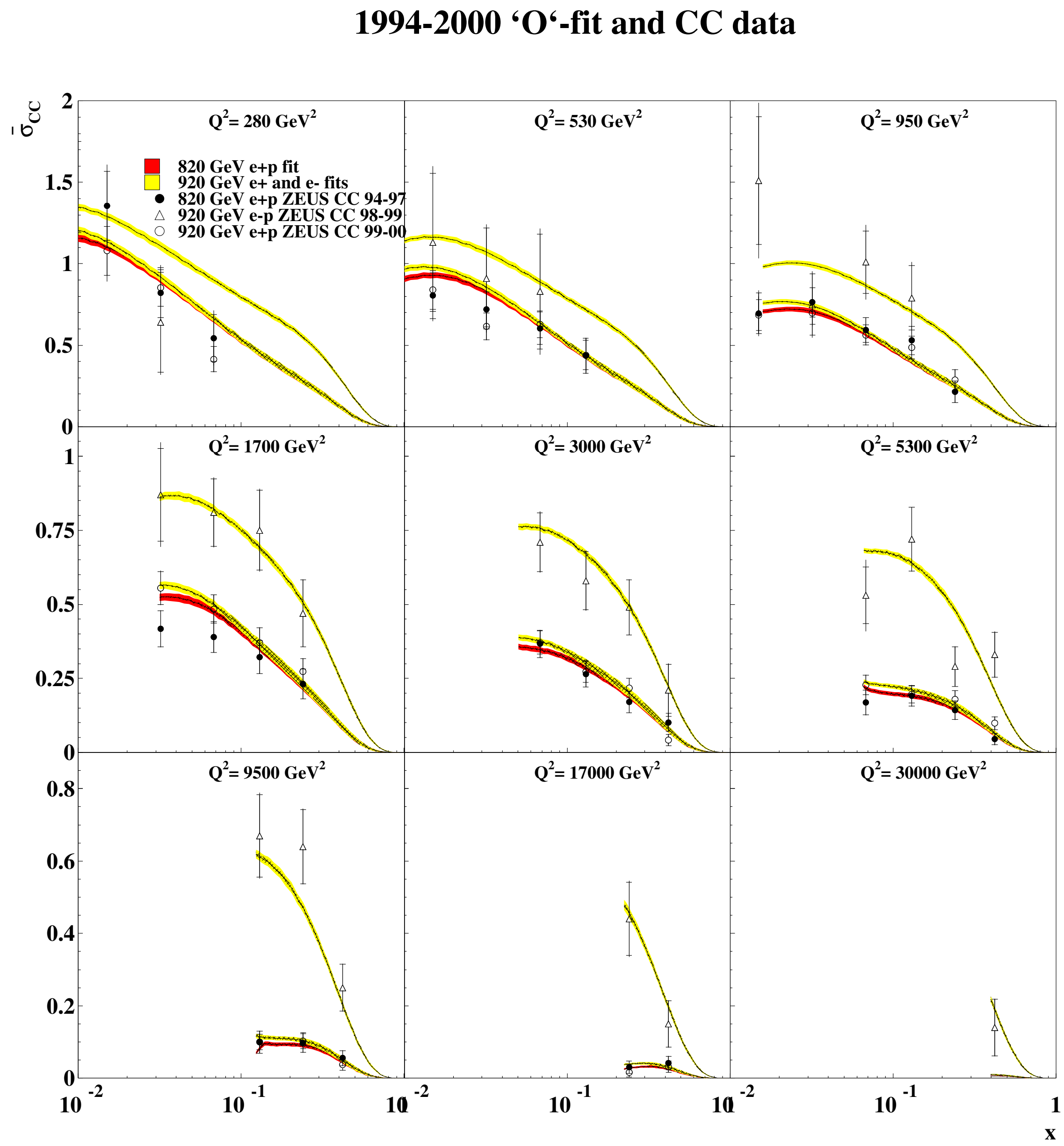

Figure 66: *The ZEUS reduced charged current cross section data and 1994-2000 ‘O’ fit predictions in bins of fixed Q^2.*

Data Set	No. of data points	χ^2	χ^2 per data point	No. of correlated systematic uncertainties	No. of normalisation uncertainties
ZEUS $F_2$96-97	242	174.51	0.72	10	2
ZEUS CC e^+P94-97	29	16.733	0.58	3	1
ZEUS NC e^-P98-99	92	55.751	0.61	6	1
ZEUS CC e^-P98-99	26	23.239	0.89	3	
ZEUS NC e^+P99-00	90	75.448	0.84	8	1
ZEUS CC e^+P99-00	30	32.238	1.07	3	
TOTAL	509	377.91	0.74	33	5

Table 12: χ^2 *values for the data sets in the 1994-2000 'O' fit.*

the high-x uncertainty is reduced to values similar to those seen in the global 'S' fit (figure 62). It should also be noted that the uncertainty is still predominantly statistical, making improvement possible with further data from HERA II. With the reduction in uncertainty afforded by adding the 1999-2000 data, the difference in the prediction of the 'O' fit and the 'S' fit central values has now become statistically significant. As stated previously, this is investigated in more detail in §5.5. Additionally, the prediction for the high-x sea is still very low (as in the 1994-1998 'O' fit). This will also be investigated in §5.5 and §5.6.

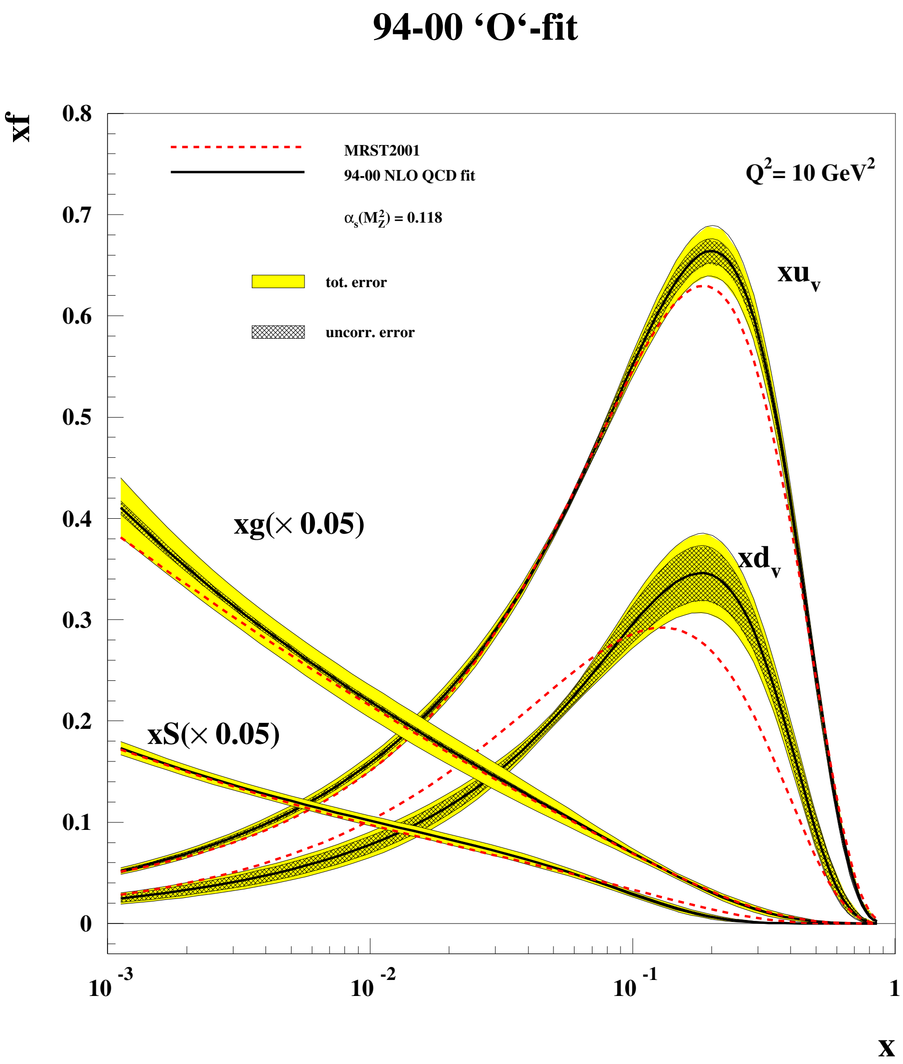

Figure 67: *The 1994-2000 'O' fit, including all six data sets, with MRST(2001) curves overlaid for reference.*

PDF	p_1	p_2	p_3	p_5
xu_v	*1.47	*0.5	$4.43 \pm 0.14 \pm 0.17$	$8.83 \pm 1.04 \pm 1.72$
xd_v	*0.70	*0.5	$5.27 \pm 0.54 \pm 0.37$	$13.3 \pm 6.3 \pm 5.9$
xS	$0.692 \pm 0.018 \pm 0.074$	$-0.220 \pm 0.003 \pm 0.016$	$15.5 \pm 1.2 \pm 3.0$	$15.8 \pm 2.5 \pm 7.9$
xg	*1.43	$-0.23 \pm 0.01 \pm 0.05$	$5.2 \pm 0.3 \pm 1.3$	*0.0
$x\Delta$	*0.27	*0.5	*17.5	*0.0

Table 13: *PDF parameters at Q_0^2 from the 94-00 'O' fit. The first uncertainty printed arises from the statistical and uncorrelated errors. The second uncertainty is the contribution from correlated systematic sources. Starred values were either fixed at the start of the fit or derived from the free parameters of the fit.*

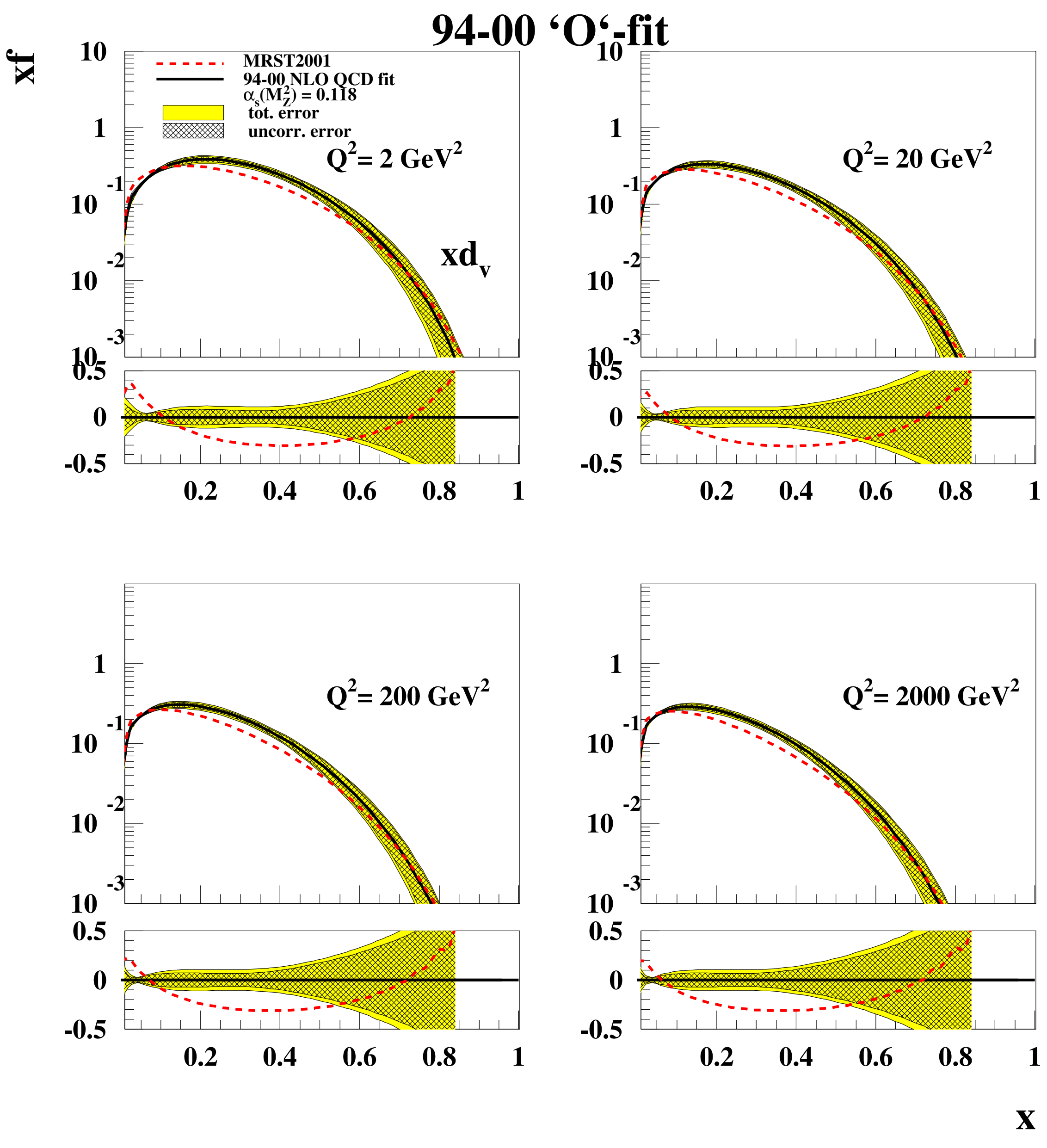

Figure 68: *d-valence plots at various Q^2 values for the 1994-2000 'O' fit with all six data sets. The uncertainties are shown for each Q^2 as fractional differences from the central value.*

5.5 Change of input parameterisation

The transition from 'S' fit to 'O' fit has caused changes in the u, d-valence and high-x sea distributions. These differences became statistically significant with the addition of the 1999-2000 ZEUS data to the 1994-1998 'O' fit. It should be noted that the u, d-valence and high-x sea distributions were first observed to differ in the ZEUS-S and ZEUS-O fits contained in [80].

When viewed in comparison with the 'S' fit and MRST prediction, it is clear that the assumptions made in the 1994-1998 and 1994-2000 'O' fits merit some investigation. The fixing and freeing of the fit parameters $p_{1..5}$ inevitably becomes more significant in the 'O' fits, since the ZEUS data essentially only provide information on the low-x sea and gluon, and the high-x valence distributions.

The decision to fix $p_2 u_v$ and $p_2 d_v$ to 0.5 was originally introduced in [80] because it was felt that there was little reliable information on the low-x valence distributions in either the ZEUS-S or ZEUS-O fits. This restriction will now be removed, allowing the parameters to go free, for both the 'S' and 'O' fits, though $p_2 u_v = p_2 d_v$ ($\equiv p_2 valence$) is still maintained, since there is no information to separate them.

The lack of information on the high-x sea in the 1994-1998 and 1994-2000 'O' fits is also addressed in sections 5.5.2 and 5.5.3 by fixing the value of $p_3 S$ in the 'O' fits to the value obtained in the 'S' fit, as was described in §5.3.

5.5.1 S-fit with $p_2 valence$ free

The 'S' fit of §5.2.1 was repeated, but $p_2 u_v = p_2 d_v$ ($\equiv p_2 valence$) was added to the list of free parameters. As can be seen in table 14, the χ^2 value for this fit has decreased relative to the original 'S' fit χ^2 value of 1198.5 from table 8. It should be noted that the fixed-target data measurements are made at modest Q^2 (well below the mass of the Z^0 squared), where γ-exchange dominates the measurements. At higher Q^2, Z^0-exchange becomes important, as does the $\gamma - Z^0$ interference term. In order to allow comparison of the ZEUS and fixed-target data at low-Q^2, only the portion of F_2 due to γ-exchange (F_2^{em}) is quoted for the ZEUS data. Figures 69 and 70 show the data entering into the fit, together with the fit predictions for F_2^{em} and $\bar{\sigma}_{NC}$ (the reduced neutral current cross section).

Figure 71 shows that the u and d-valence distributions are now much closer to the MRST prediction[8] near the peaks, when compared to the original 'S' fit distributions in figure 60.

Data Set	No. of data points	χ^2	χ^2 per data point	No. of correlated systematic uncertainties	No. of normalisation uncertainties
ZEUS $F_2$96/97	242	198.38	0.82	10	2
BCDMS	305	271.03	0.89	5	5
NMC p	218	268.71	1.23	12	4
NMC D	218	202.35	0.93		
NMC D/p	129	116.33	0.90	5	0
E665 p	47	54.374	1.16	7	2*
E665 D	47	43.632	0.93		1*
CCFR xF_3	57	15.856	0.28	18	0
TOTAL	1263	1170.66	0.93	57	14

Table 14: *χ^2 values for the data sets in the 'S' fit with $p_2$valence free, obtained by adding all systematic and statistical uncertainties in quadrature at the minimum found by using the χ^2 of section 5.1. The systematic uncertainties for the p and D sets of NMC and E665 were taken together. Similarly, the normalisation uncertainties were the same for NMC p and D data. The star indicates that for E665, there is one common normalisation uncertainty, and one normalisation uncertainty each, for the targets p, D.*

[8]The MRST prediction is used throughout this chapter as a useful benchmark.

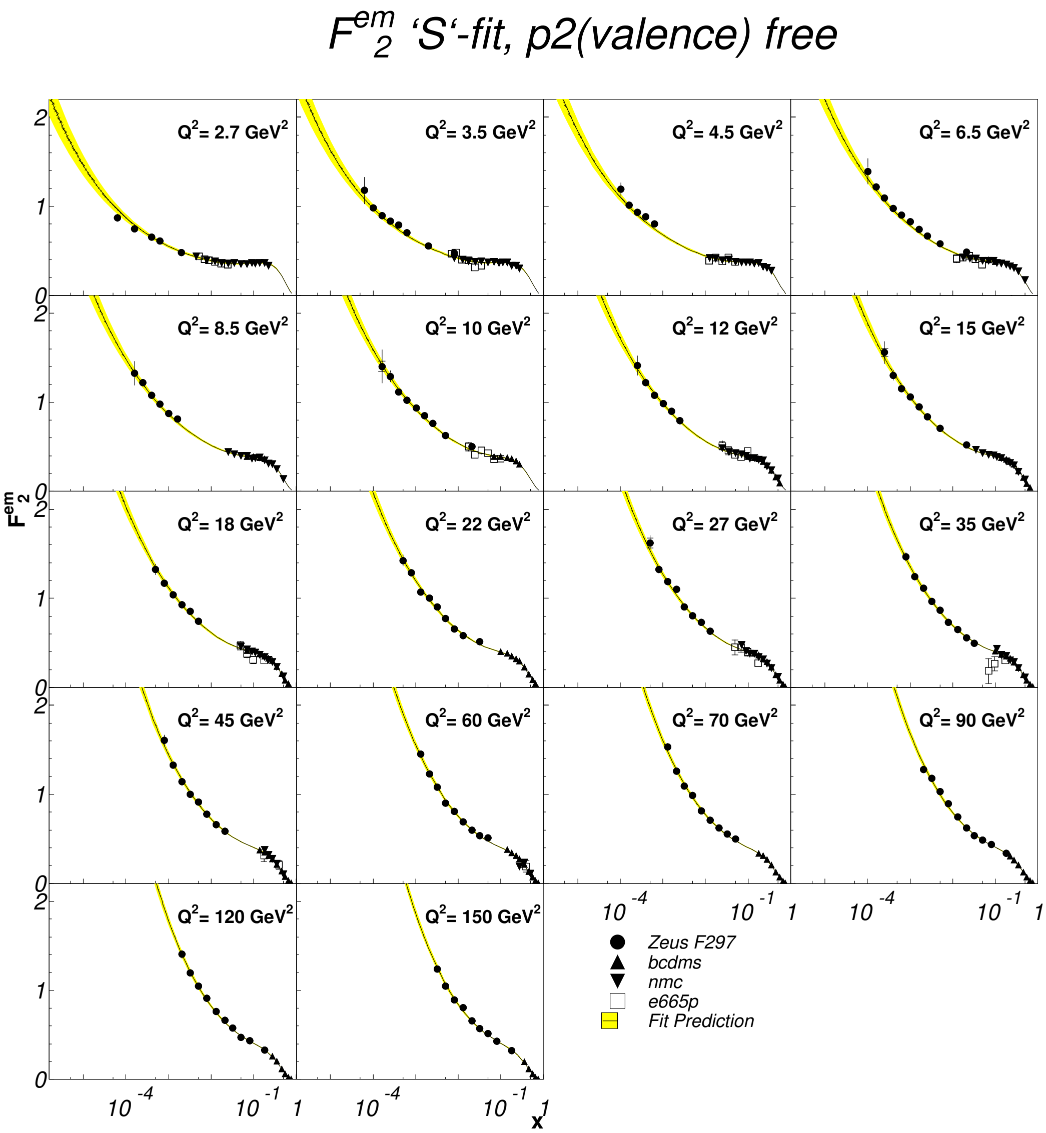

Figure 69: *The F_2^{em} fit prediction and data in the 'S' fit (with $p_2$valence free), in bins of fixed Q^2.*

‘S‘-fit, p2(valence)=free, and NC data

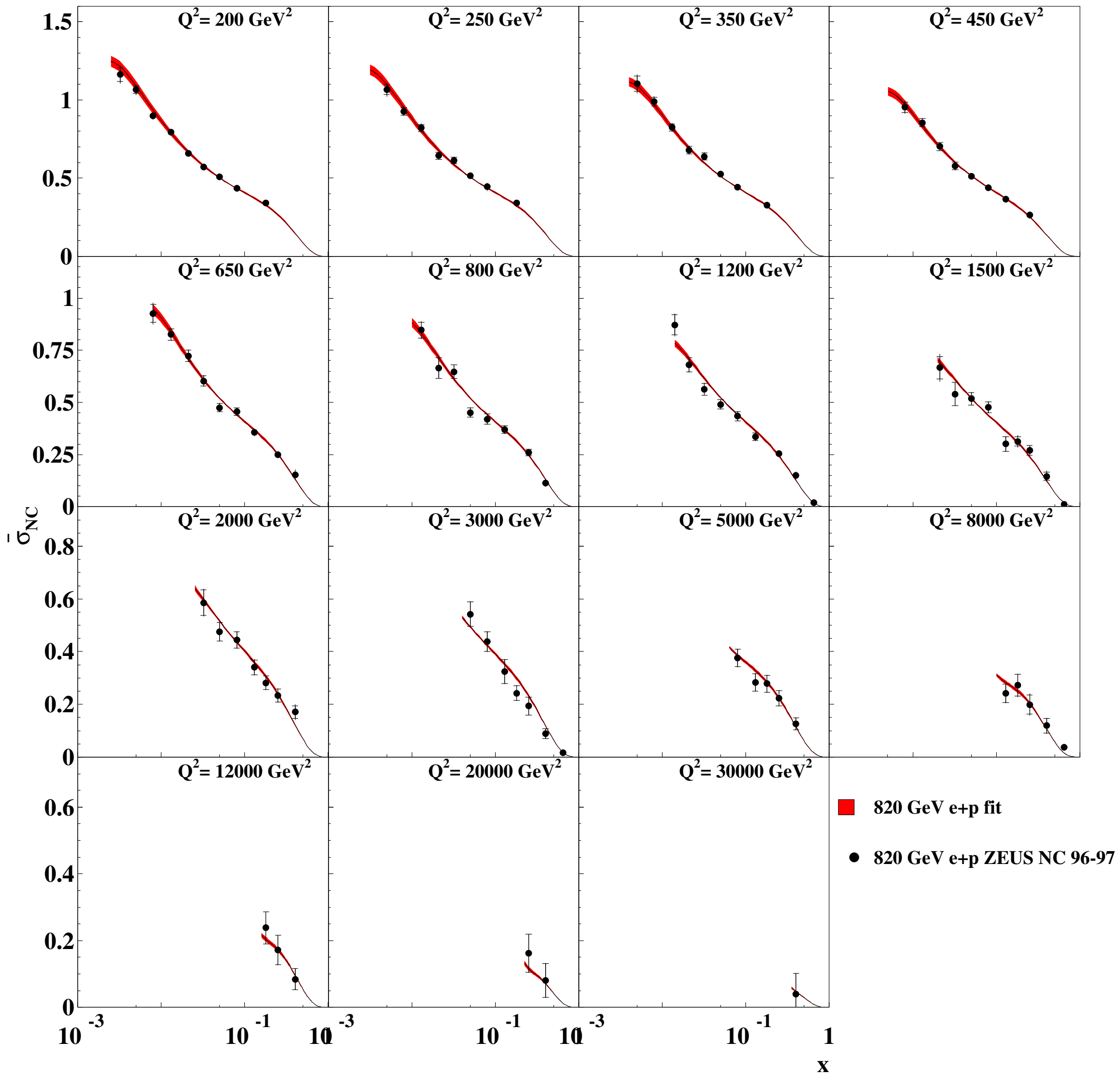

Figure 70: *The ZEUS reduced neutral current cross section data and ‘S’ fit (with $p_2$valence free) predictions in bins of fixed Q^2.*

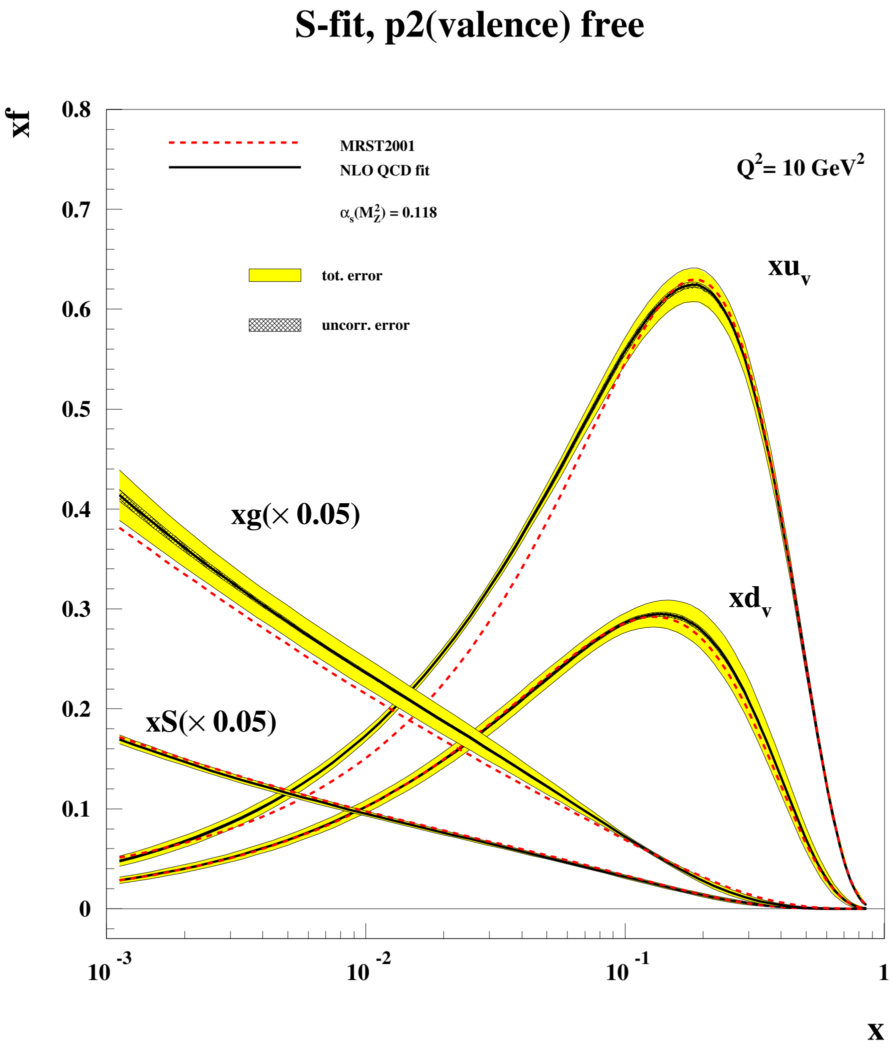

Figure 71: *The S-fit, with $p_2$valence free. MRST(2001) curves are overlaid for reference. This fit agrees more closely with MRST around the valence peaks than the original S-fit from §5.2.1.*

PDF	p_1	p_2	p_3	p_5
xu_v	*2.76	$0.60 \pm 0.01 \pm 0.04$	$3.90 \pm 0.02 \pm 0.11$	$2.27 \pm 0.22 \pm 0.69$
xd_v	*1.66		$4.64 \pm 0.19 \pm 0.63$	$1.4 \pm 0.5 \pm 1.7$
xS	$0.642 \pm 0.008 \pm 0.052$	$-0.227 \pm 0.002 \pm 0.012$	$7.4 \pm 0.4 \pm 1.7$	$2.7 \pm 0.7 \pm 2.3$
xg	*1.91	$-0.19 \pm 0.01 \pm 0.04$	$6.6 \pm 0.2 \pm 1.3$	*0.0
$x\Delta$	$0.29 \pm 0.01 \pm 0.07$	*0.5	*9.4	*0.0

Table 15: *PDF parameters at Q_0^2 from the S-fit with $p_2$valence free. The first uncertainty printed arises from the statistical and uncorrelated errors. The second uncertainty is the contribution from correlated systematic sources. Starred values were either fixed at the start of the fit or derived from the free parameters of the fit.*

5.5.2 The 1994-2000 'O-modified' fit - (1994-2000 O-fit with $p_2valence$ free and p_3S fixed)

The restriction on $p_2valence$ is now lifted for the 1994-2000 'O' fit, just as was done in the previous section for the 'S' fit. In addition, it was previously noted that the 'O' fits historically showed deviation from the 'S' fit (and comparable fits such as MRST(2001)) in the high-x sea region. For this reason, the 'O' fit is now also modified by fixing the central value of the high-x sea parameter p_3S, to the value obtained in the 'S' fit. For clarity, this fit will be known as the 1994-2000 'O-modified' fit.

Table 16 and figures 72, 73 and 74 show that the 1994-2000 'O-modified' fit gives an excellent description of the low-Q^2 F_2^{em} data, the high-Q^2 neutral current reduced cross section data and the high-Q^2 charged current reduced cross section data.

The improvement in the central values of the valence distributions is clearly seen when this fit (figure 75) is compared to the first 1994-2000 'O' fit in figure 67 (the first 1994-2000 'O' fit had the same parameterisation as the 1994-1998 'O' fit, but contained extra ZEUS data from 1999-2000). As a result of the new parameter fixing/freeing, the 1994-2000 'O-modified' fit now has central values which are more consistent with the 'S' fit and MRST(2001) curves. The improvement in the high-x sea is also visible when this fit (figure 77) is compared to the 1994-2000 and 1994-1998 'O'-fits (e.g. see figure 64).

Unfortunately, the size of the uncertainties on the valence distributions have grown slightly larger. The d-valence distribution in the 1994-2000 'O-modified' fit (figure 76) still shows improvement on the 1994-1998 'O' fit, however the 1994-2000 'O-modified' u-valence distribution has slightly larger uncertainties than the 1994-1998 'O' fit.

It is not entirely justifiable however, to compare the 1994-2000 'O-modified' fit and the 1994-1998 'O' distributions, as the parameters fixed and freed were not the same. To perform a valid comparison (and hence deduce the improvement offered by the 1999-2000 data), the 'O-modified' fit must be carried out on the 1994-1998 data sets (omitting the 1999-2000 data). This is the topic of the next section.

Data Set	No. of data points	χ^2	χ^2 per data point	No. of correlated systematic uncertainties	No. of normalisation uncertainties
ZEUS $F_2$96-97	242	174.51	0.72	10	2
ZEUS CC e^+P94-97	29	15.379	0.53	3	1
ZEUS NC e^-P98-99	92	57.455	0.62	6	1
ZEUS CC e^-P98-99	26	23.154	0.89	3	
ZEUS NC e^+P99-00	90	78.587	0.87	8	1
ZEUS CC e^+P99-00	30	33.415	1.11	3	
TOTAL	509	382.50	0.75	33	5

Table 16: *χ^2 values for the data sets in the 1994-2000 'O-modified' fit.*

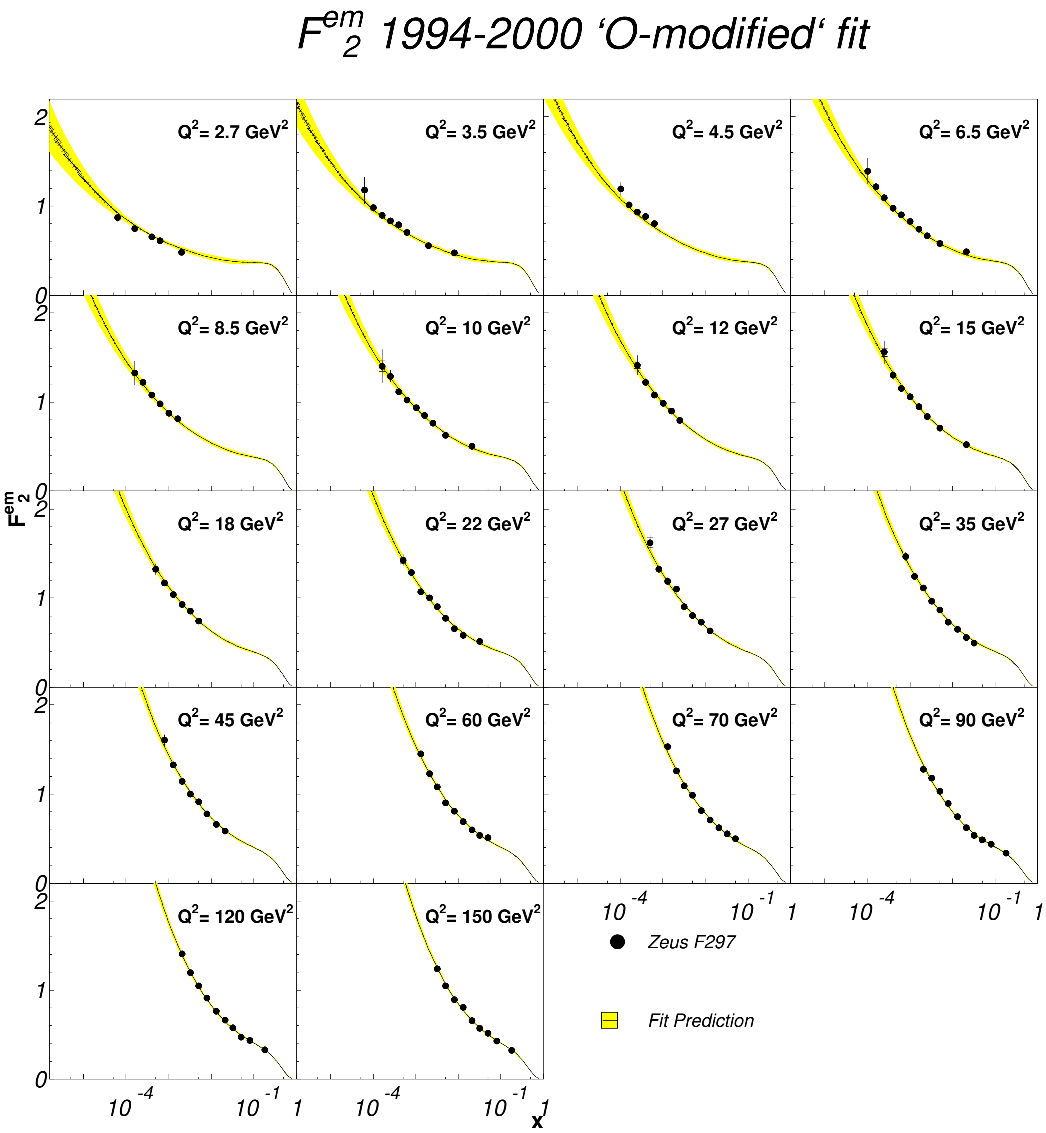

Figure 72: *The F_2^{em} fit prediction and ZEUS data in the 1994-2000 'O-modified' fit, in bins of fixed Q^2.*

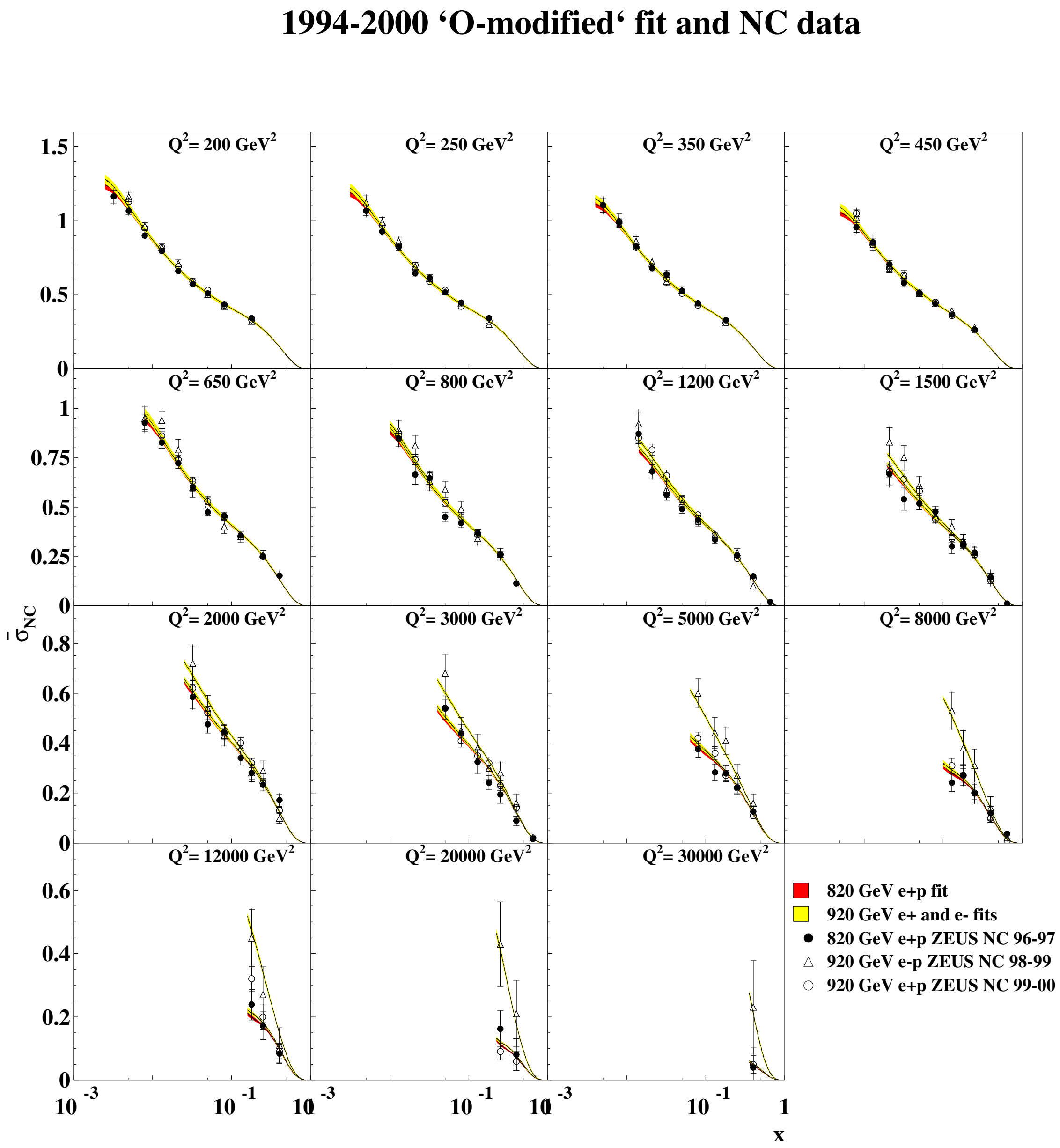

Figure 73: *The ZEUS reduced neutral current cross section data and 1994-2000 'O-modified' fit predictions, in bins of fixed Q^2.*

1994-2000 'O-modified' fit and CC data

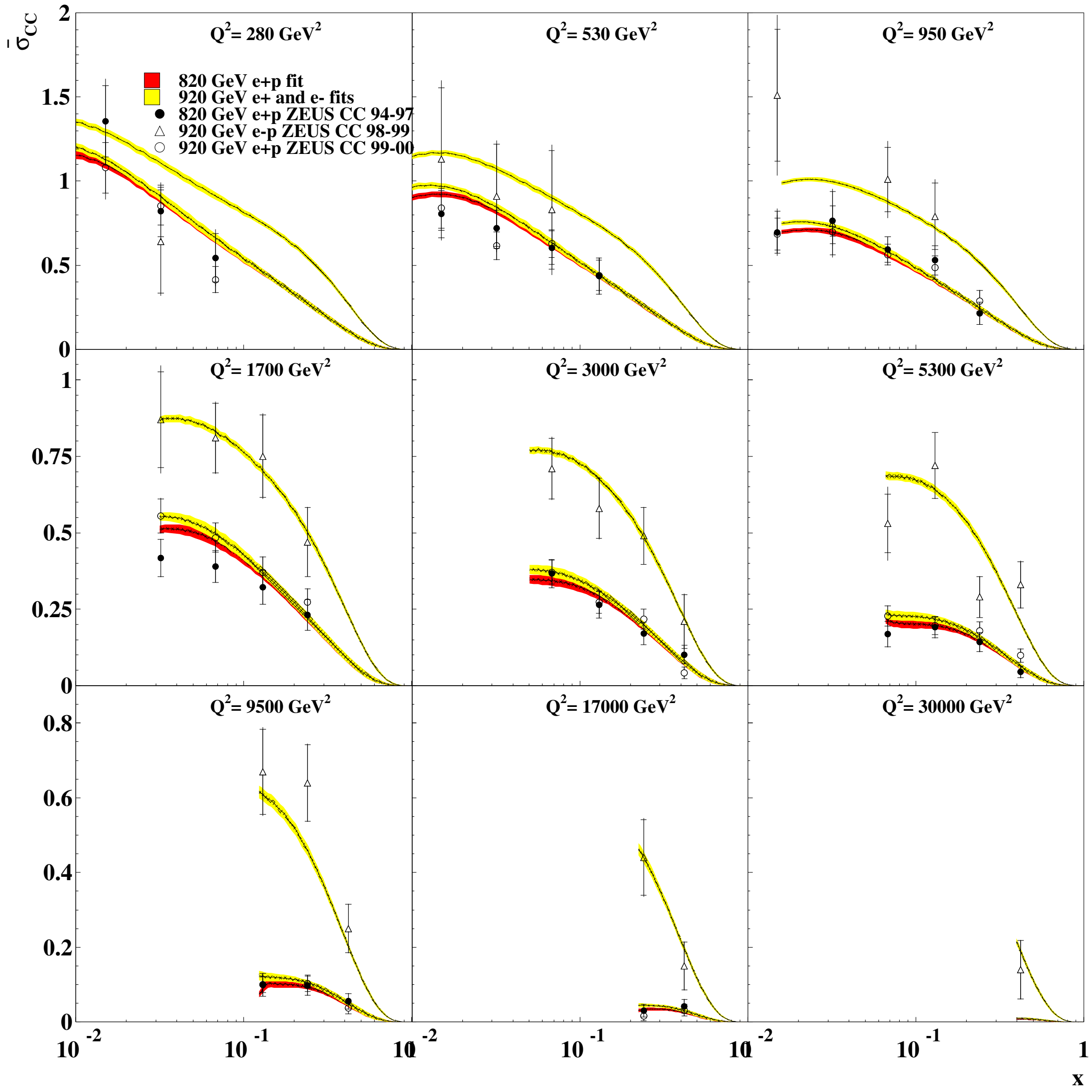

Figure 74: *The ZEUS reduced charged current cross section data and 1994-2000 'O-modified' fit predictions, in bins of fixed* Q^2.

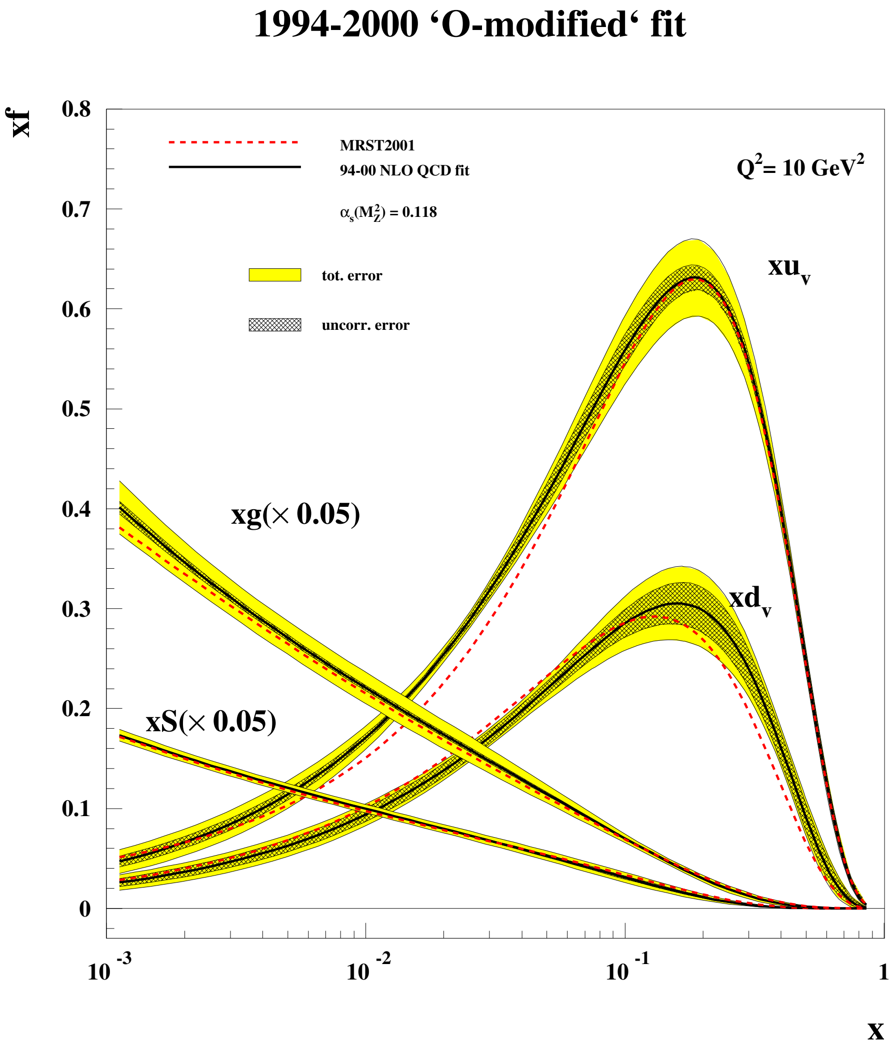

Figure 75: *The 1994-2000 'O-modified' fit with MRST(2001) curves overlaid for reference. This fit agrees more closely with MRST around the valence peaks than the first 1994-2000 'O' fit (§5.4), but it also has larger uncertainties.*

PDF	p_1	p_2	p_3	p_5
xu_v	*2.71	$0.60 \pm 0.05 \pm 0.08$	$3.96 \pm 0.16 \pm 0.20$	$2.57 \pm 1.14 \pm 1.07$
xd_v	*1.50		$4.42 \pm 0.73 \pm 0.79$	$2.2 \pm 2.1 \pm 1.6$
xS	$0.747 \pm 0.023 \pm 0.081$	$-0.210 \pm 0.004 \pm 0.015$	*8.9	$2.9 \pm 1.0 \pm 2.4$
xg	*1.57	$-0.21 \pm 0.01 \pm 0.05$	$5.6 \pm 0.3 \pm 1.2$	*0.0
$x\Delta$	*0.27	*0.5	*10.9	*0.0

Table 17: *PDF parameters at Q_0^2 from the 94-00 'O-modified' fit. The first uncertainty printed arises from the statistical and uncorrelated errors. The second uncertainty is the contribution from correlated systematic sources. Starred values were either fixed at the start of the fit or derived from the free parameters of the fit.*

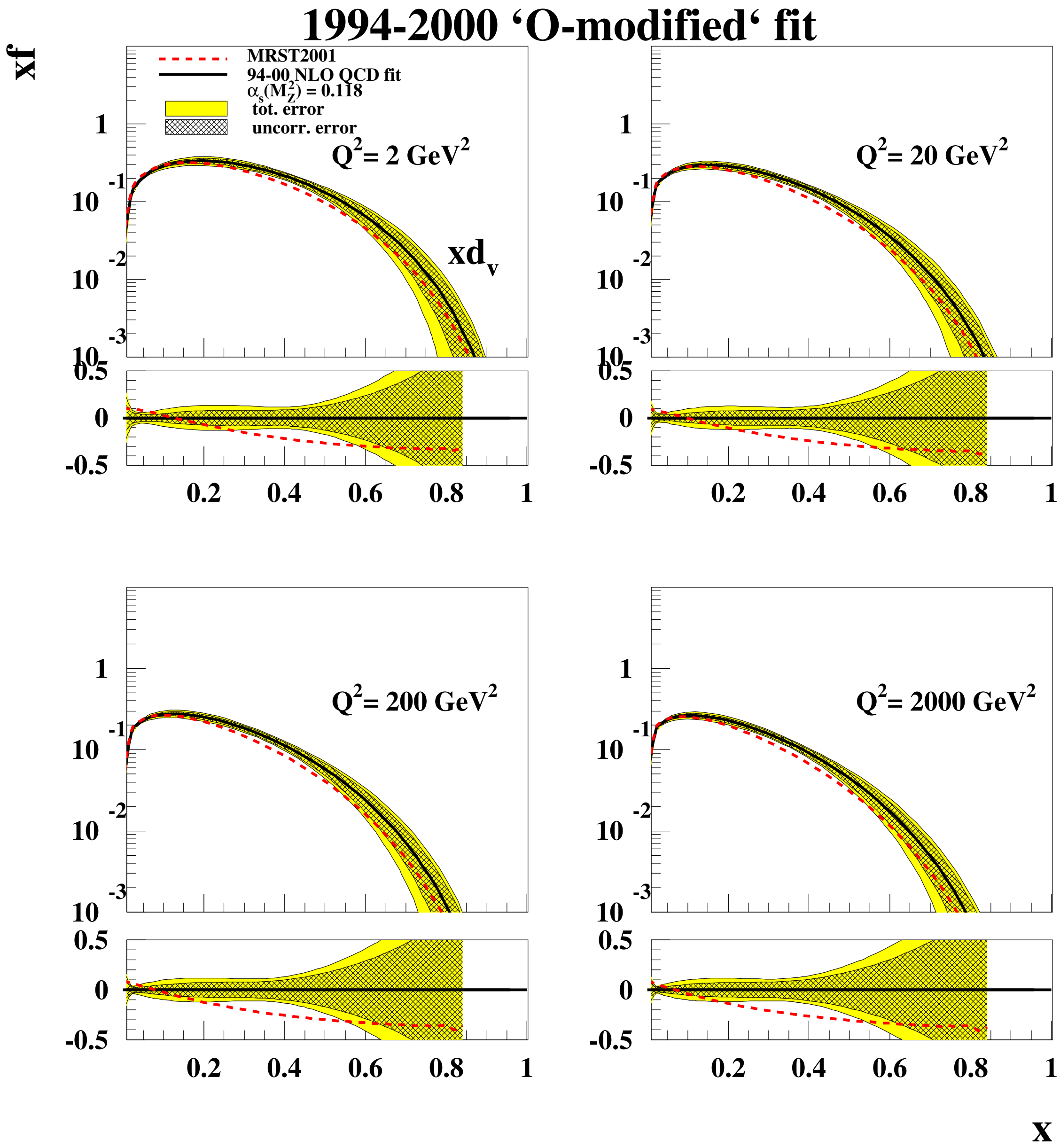

Figure 76: *The d-valence distributions at various values of Q^2 for the 1994-2000 'O-modified' fit, with MRST(2001) curves overlaid for reference. The uncertainties are also shown for each Q^2 as fractional differences from the central value. The uncertainty bands have increased in size relative to the 1994-2000 'O' fit (figure 68), however this fit still has a smaller high-x d-valence uncertainty than the 1994-1998 'O' fit (the ZEUS-O fit replication, figure 63).*

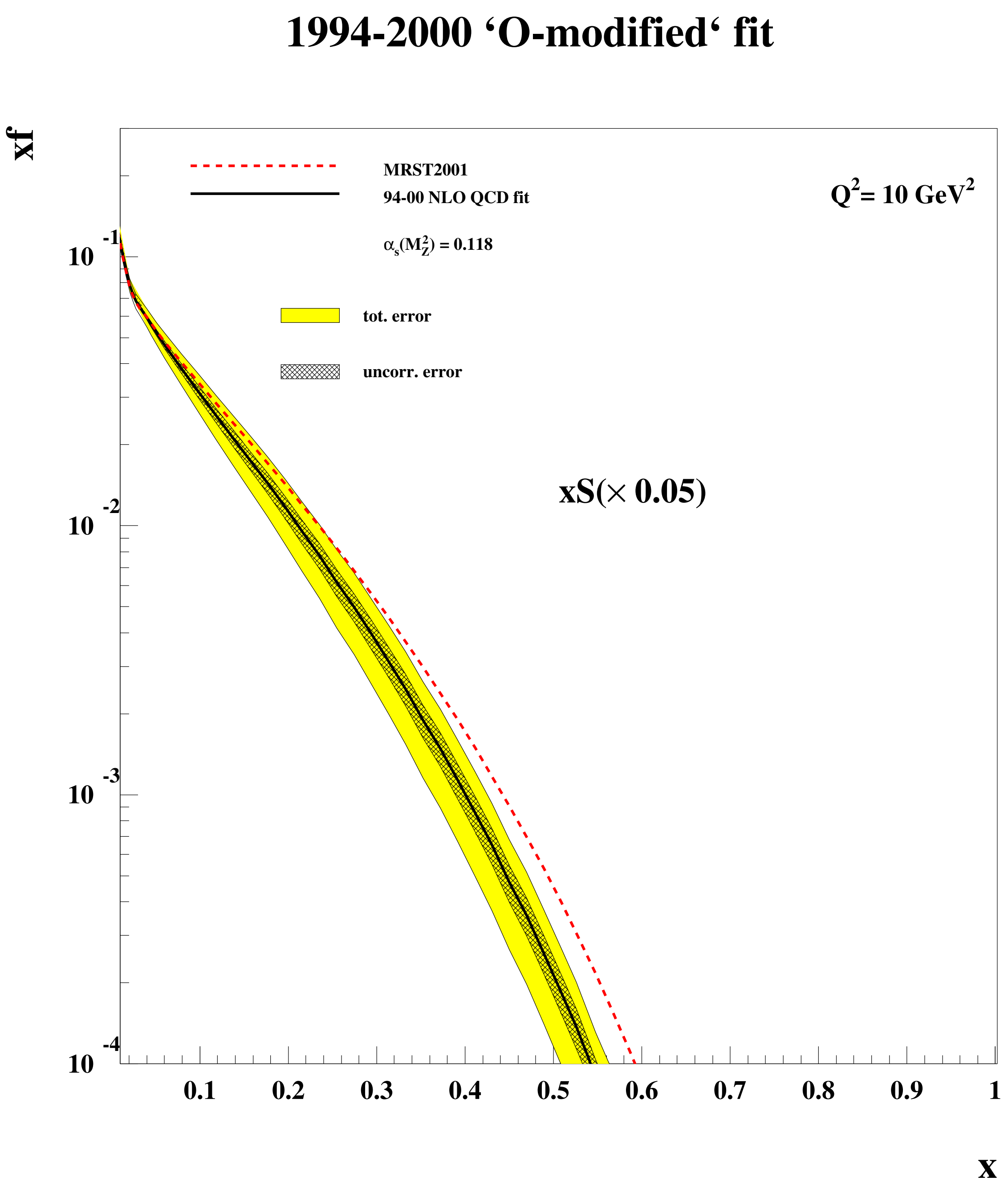

Figure 77: *The 'O-modified' fit high-x sea distribution with MRST(2001) values overlaid for reference.*

5.5.3 A retrospective 1994-1998 'O-modified' fit (1994-1998 'O' fit with $p_2valence$ free and p_3S fixed)

In order to make a valid assessment of the improvement offered by adding the 1999-2000 data, a comparison of the 1994-2000 'O-modified' fit with a 1994-1998 'O-modified' fit was attempted. Only the data sets from the 1994-1998 'O' fit were used, and the parameters fixed and freed were the same as in §5.5.2, namely the original 'O' fit parameters, plus $p_3S = 8.9$ and $p_2valence$ free.

It was discovered that with only the 1994-1998 data, the χ^2 minimisation finds two separate minima. These are displayed in table 20, figures 78, 79 and tables 18, 19. These two minima have largely consistent parameter values, apart from p_3d_v. It should be noted also that p_5d_v is quite different, but the fit uncertainties mean that the two values are not in disagreement. Both fits have PDFs much closer to ZEUS-S and MRST(2001) than the original 1994-1998 'O' fit. This double minimum collapses to a single minimum as soon as the 1999-2000 data are added.

Unfortunately, due to the double minimum structure witnessed in the 1994-1998 'O-modified' fit, it does not seem entirely possible to quantify the benefit provided by the 1994-2000 'O-modified' fit, however it seems that whatever configuration was used in the past, the 1994-2000 'O-modified' fit offers a d-valence determination with smaller uncertainties than any 1994-1998 'O' fit and is more consistent with the rest of the world data.

It is also possible to conjecture that the old 'O' fits (without $p_2valence$ free and p_3S fixed) might be considered naïve or over-optimistic (they needed more study of the model dependence), as they seem to predict smaller uncertainties in the valence distributions, whilst then tending to have central values far from the MRST(2001)/ZEUS-S fits.

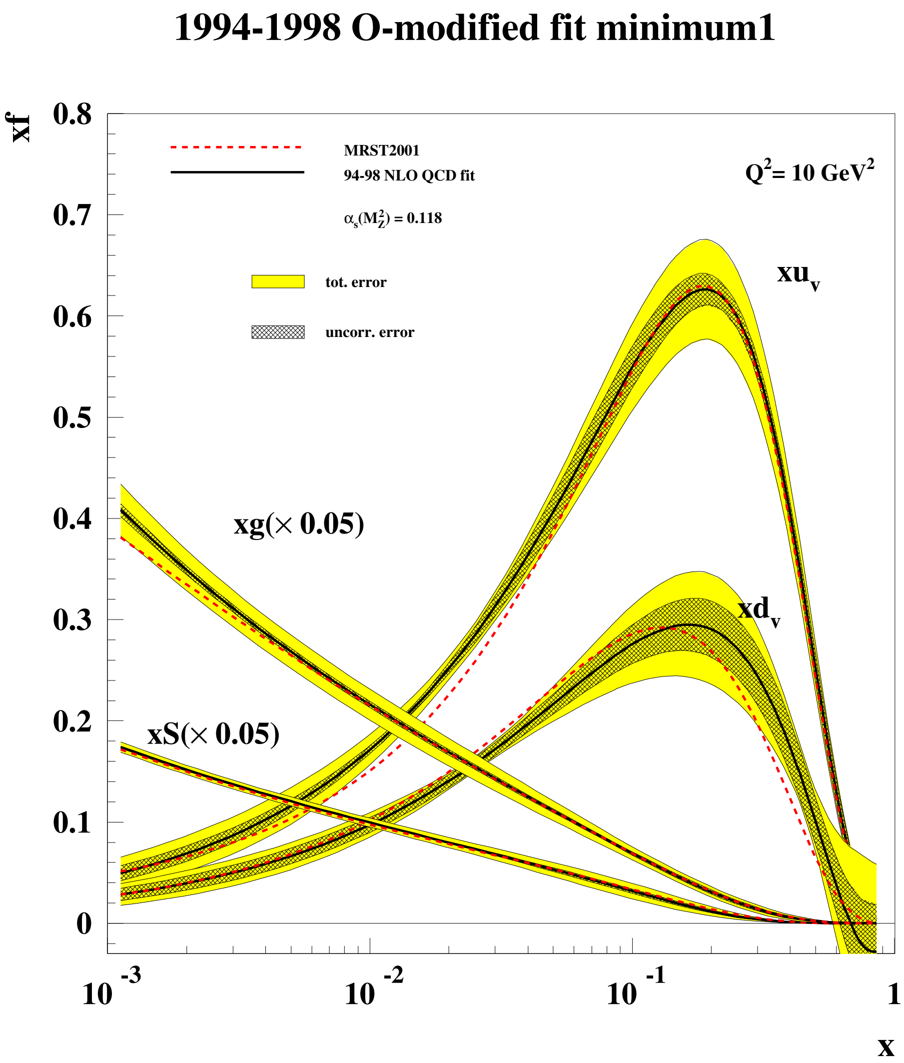

Figure 78: *The 1994-1998 'O-modified' fit (first χ^2 minimum), with MRST(2001) curves overlaid for reference.*

PDF	p_1	p_2	p_3	p_5
xu_v	*2.41	$0.58 \pm 0.07 \pm 0.10$	$3.92 \pm 0.26 \pm 0.16$	$3.08 \pm 1.83 \pm 1.37$
xd_v	*1.39		$1.21 \pm 0.62 \pm 0.72$	$-1.5 \pm 0.20 \pm 0.4$
xS	$0.730 \pm 0.027 \pm 0.087$	$-0.214 \pm 0.004 \pm 0.017$	*8.9	$3.42 \pm 1.29 \pm 3.64$
xg	*1.37	$-0.24 \pm 0.01 \pm 0.04$	$5.1 \pm 0.5 \pm 1.1$	*0.0
$x\Delta$	*0.27	*0.5	*10.9	*0.0

Table 18: *PDF parameters at Q_0^2 from the first minimum of the 94-98 'O-modified' fit. The first uncertainty printed arises from the statistical and uncorrelated errors. The second uncertainty is the contribution from correlated systematic sources. Starred values were either fixed at the start of the fit or derived from the free parameters of the fit.*

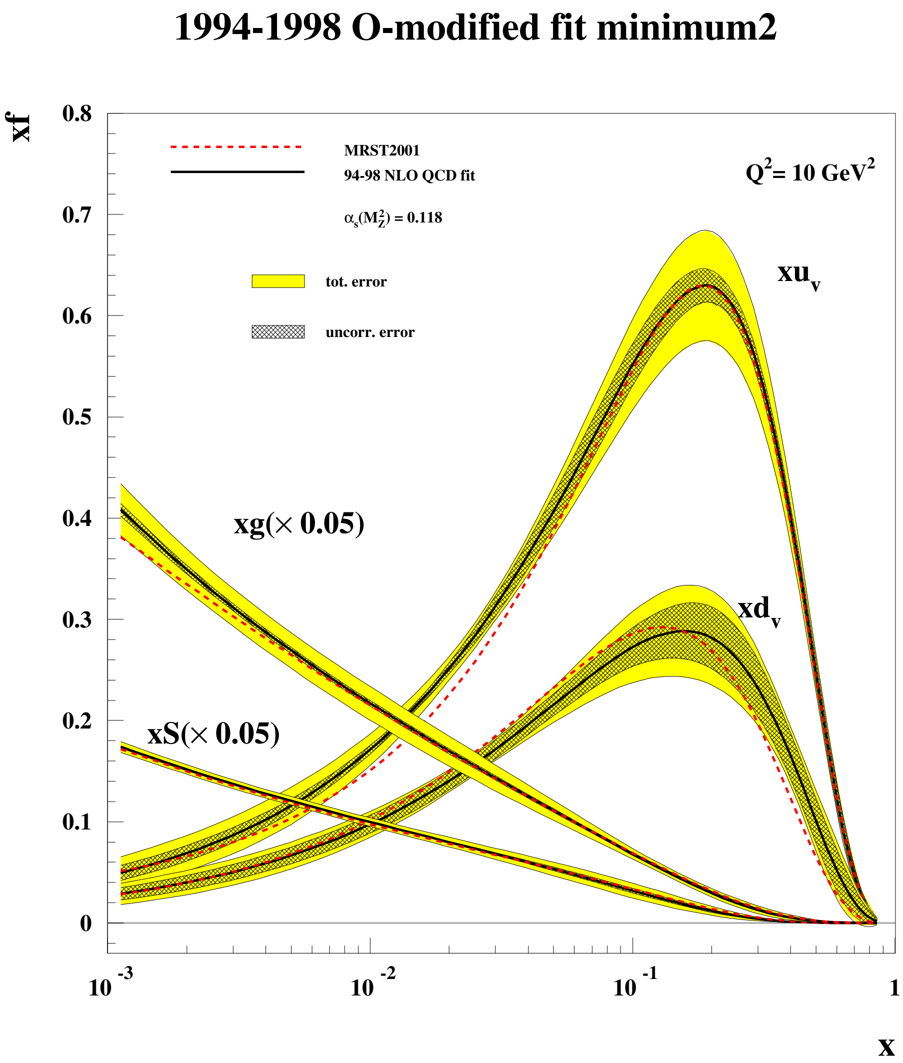

Figure 79: *The 1994-1998 'O-modified' fit (second χ^2 minimum), with MRST(2001) curves overlaid for reference.*

PDF	p_1	p_2	p_3	p_5
xu_v	*2.37	$0.58 \pm 0.07 \pm 0.10$	$4.03 \pm 0.21 \pm 0.31$	$3.41 \pm 1.78 \pm 1.42$
xd_v	*1.38		$3.69 \pm 1.41 \pm 1.18$	$1.0 \pm 2.8 \pm 1.5$
xS	$0.732 \pm 0.026 \pm 0.085$	$-0.213 \pm 0.004 \pm 0.017$	*8.9	$3.53 \pm 1.30 \pm 3.50$
xg	*1.41	$-0.23 \pm 0.01 \pm 0.04$	$5.3 \pm 0.4 \pm 1.3$	*0.0
$x\Delta$	*0.27	*0.5	*10.9	*0.0

Table 19: *PDF parameters at Q_0^2 from the second minimum of the 94-98 'O-modified' fit. The first uncertainty printed arises from the statistical and uncorrelated errors. The second uncertainty is the contribution from correlated systematic sources. Starred values were either fixed at the start of the fit or derived from the free parameters of the fit.*

Data Set	No. of data points	χ^2 min. 1	χ^2 min. 2	χ^2 per point min. 1	χ^2 per point min. 2
96-97 NC e^+P	242	168.70	169.02	0.70	0.70
98-99 NC e^-P	92	62.00	62.29	0.67	0.68
94-97 CC e^+P	29	15.71	15.28	0.54	0.53
98-99 CC e^-P	26	23.08	23.16	0.89	0.89
total	389	269.49	269.75	0.69	0.69

Table 20: *χ^2 values by data set for the two minima seen in the 1994-1998 'O-modified' fits.*

5.6 Adding BCDMS hydrogen data

It was initially thought that adding the fixed-target BCDMS hydrogen data [98] to the 'O' fits might stabilise the high-x sea distribution (instead of fixing p_3S), but this turned out not to be very useful. The BCDMS data stretches to much higher x than the other pure hydrogen data sets of NMC or E665, and whilst the addition of the data does stabilise the high-x sea, the effects on the valence distributions are severe. For the 1994-2000 'O' fit with BCDMS data, a reasonable fit to the data was seen, but the increase in systematic uncertainty introduced in the d-valence distribution was unacceptably large. The u-valence distribution also becomes quite incompatible with MRST. The results of adding the BCDMS data to the 1994-2000 'O' fit are shown in figure 80 and tables 21,22.

Data Set	No. of data points	χ^2	χ^2 per data point	No. of correlated systematic uncertainties	No. of normalisation uncertainties
ZEUS $F_2$96-97	242	177.33	0.73	10	2
ZEUS CC e^+P94-97	29	23.894	0.82	3	1
ZEUS NC e^-P98-99	92	57.515	0.63	6	1
ZEUS CC e^-P98-99	26	22.778	0.88	3	
ZEUS NC e^+P99-00	90	81.685	0.91	8	1
ZEUS CC e^+P99-00	30	44.609	1.49	3	
BCDMS	305	259.82	0.85	5	5
TOTAL	814	667.63	0.82	38	10

Table 21: *χ^2 values for the data sets in the 1994-2000 'O'+BCDMS fit.*

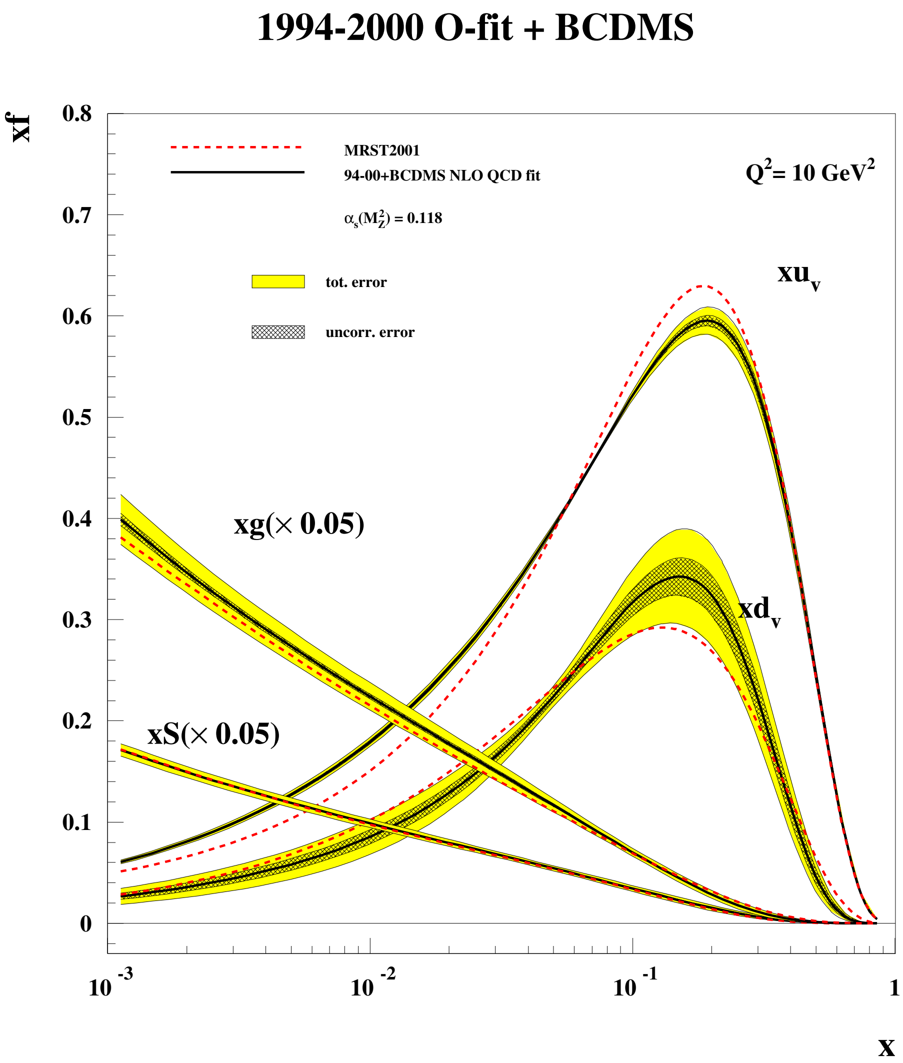

Figure 80: *The 1994-2000 'O' fit with BCDMS data. MRST(2001) curves are overlaid for reference.*

PDF	p_1	p_2	p_3	p_5
xu_v	*1.74	*0.5	$3.89 \pm 0.03 \pm 0.10$	$4.29 \pm 0.25 \pm 0.70$
xd_v	*0.75	*0.5	$6.68 \pm 0.37 \pm 0.69$	$17.0 \pm 4.9 \pm 11.4$
xS	$0.717 \pm 0.013 \pm 0.086$	$-0.214 \pm 0.003 \pm 0.015$	$6.50 \pm 0.71 \pm 0.52$	$1.54 \pm 0.76 \pm 1.27$
xg	*1.71	$-0.20 \pm 0.01 \pm 0.03$	$6.1 \pm 0.3 \pm 1.1$	*0.0
$x\Delta$	*0.27	*0.5	*8.50	*0.0

Table 22: *PDF parameters at Q_0^2 for the 94-00 'O' fit with BCDMS data. The first uncertainty printed arises from the statistical and uncorrelated errors. The second uncertainty is the contribution from correlated systematic sources. Starred values were either fixed at the start of the fit or derived from the free parameters of the fit.*

When the BCDMS data are added to the 1994-2000 'O-modified' fit, and p_3S is freed (to allow the BCDMS data to determine the high-x sea), the valence distributions become almost entirely incompatible with MRST(2001), but one interesting fact emerges: the value of p_3S determined by the fit is $9.1 \pm 3.0 \pm 4.0$. This is reassuringly close to the value obtained in the 'S' fit ($8.9 \pm 0.2 \pm 1.2$), and lends additional justification to the use of the 'S' fit value in the 'O-modified' fits.

5.7 The gluon parameterisation

As was previously mentioned in §5.2, fixing the parameter p_5g to zero in the 'S', 'O', and (lately) in the 'O-modified' fits has been controversial, since H1 perform their fits with p_5g free. This issue will now be investigated.

5.7.1 Freeing the middling-x gluon parameter, p_5g

When the 'S' fit with $p_2valence$ free (from §5.5.1) is performed with p_5g free as well (the 'S' fits include fixed-target data), values for p_5g close to zero are obtained. Other parameter values obtained all virtually unchanged. Table 23 shows the central values obtained for the 'S' fit with $p_2valence$ & p_5g free. Comparing these values with table 15, (showing the 'S' fit with just $p_2valence$ free) demonstrates that the effect of freeing p_5g on the 'S' fits is negligible.

Freeing p_5g for the fits involving ZEUS-only data (e.g. the 'O-modified' fits), results in larger p_5g central values (but with uncertainties still consistent with a p_5g of zero). The net result in the 'O-modified' fits is that a 'hump' appears in the gluon distribution for middling-x. This hump is also seen by H1 in their own fits [110], indeed it is much more pronounced. The issue therefore appears to encompass the ZEUS-only or H1-only fits, but has little effect on the fits which use all the fixed-target data.

For the 'O-modified' fits with p_5g free, it is not the presence of the hump which is worrying (the uncertainty bands around the hump encompass the MRST and ZEUS-S fit predictions, and are not inconsistent with a value for p_5g of zero). What is more concerning is the fact that the values of p_5g obtained also seem to cause the fits to have high values of p_3g. The form of the parameterisation at middling and high-$x \sim (1-x)^{p_3}(1+p_5x)$ thus means that a high value of p_3g rapidly forces the gluon to zero when x approaches one. This lack of high-x gluon is inconsistent with ZEUS-S and MRST fits, as is demonstrated in figure 82 (a linear abscissa and logarithmic ordinate have been used to emphasise the high-x region).

PDF	p_1	p_2	p_3	p_5
xu_v	*2.76	$0.60 \pm 0.01 \pm 0.04$	$3.90 \pm 0.02 \pm 0.11$	$2.27 \pm 0.22 \pm 0.70$
xd_v	*1.66		$4.64 \pm 0.19 \pm 0.63$	$1.4 \pm 0.5 \pm 1.7$
xS	$0.643 \pm 0.008 \pm 0.052$	$-0.227 \pm 0.002 \pm 0.012$	$7.4 \pm 0.3 \pm 1.7$	$2.7 \pm 0.6 \pm 2.4$
xg	*1.92	$-0.19 \pm 0.01 \pm 0.04$	$6.2 \pm 1.9 \pm 1.9$	$-0.5 \pm 1.9 \pm 1.1$
$x\Delta$	$0.29 \pm 0.01 \pm 0.07$	*0.5	*9.4	*0.0

Table 23: *The 'S' fit (with p_5g free, and $p_2$valence free) central values and uncertainties. The values obtained are almost exactly the same as table 15, where the only difference is that p_5g is fixed to zero. The fitted value of p_5g is close to zero, and thus there is no gluon 'hump' in this fit. This occurs because the fixed target data are included, which provides information on the high-x gluon.*

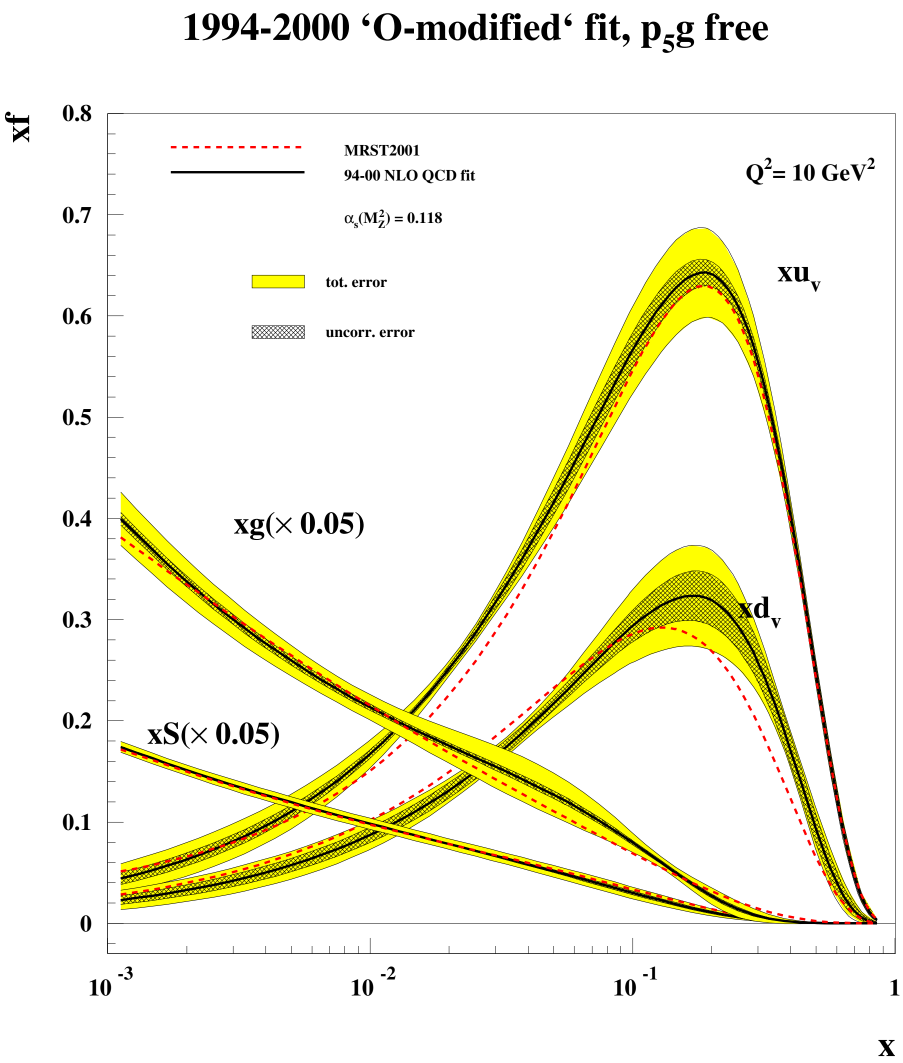

Figure 81: *The 1994-2000 'O-modified' fit with p_5g free. MRST(2001) curves are overlaid for reference.*

PDF	p_1	p_2	p_3	p_5
xu_v	*2.87	$0.62 \pm 0.06 \pm 0.11$	$4.03 \pm 0.18 \pm 0.20$	$2.63 \pm 1.33 \pm 1.63$
xd_v	*1.49		$4.78 \pm 0.69 \pm 1.03$	$3.7 \pm 2.5 \pm 2.5$
xS	$0.697 \pm 0.029 \pm 0.130$	$-0.221 \pm 0.006 \pm 0.026$	*8.9	$3.05 \pm 1.09 \pm 2.90$
xg	*0.96	$-0.28 \pm 0.02 \pm 0.07$	$11.5 \pm 0.9 \pm 4.4$	$19.8 \pm 6.3 \pm 26.0$
$x\Delta$	*0.27	*0.5	*10.9	*0.0

Table 24: *PDF parameters at Q_0^2 for the 94-00 'O-modified' fit with p_5g free. The first uncertainty printed arises from the statistical and uncorrelated errors. The second uncertainty is the contribution from correlated systematic sources. Starred values were either fixed at the start of the fit or derived from the free parameters of the fit.*

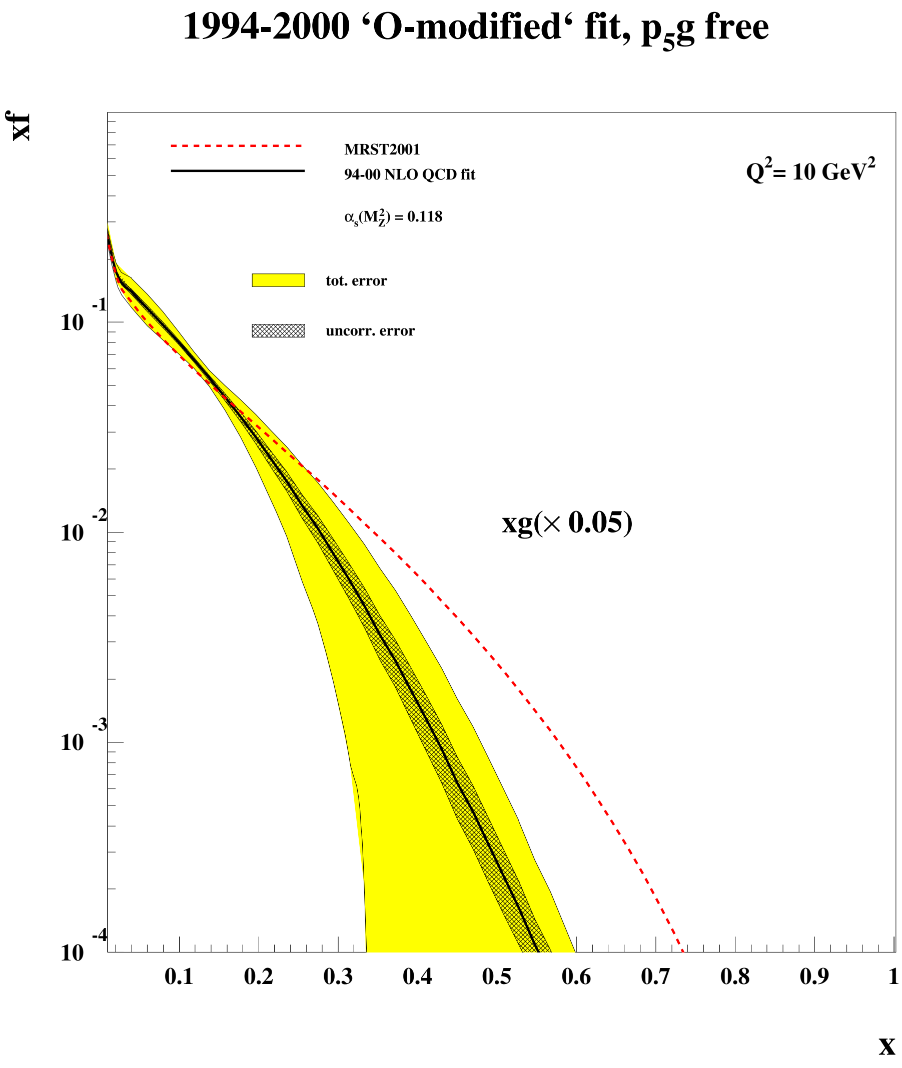

Figure 82: *The gluon distribution for the 1994-2000 'O-modified' fit, with p_5g free. The high-x attenuation of the fit is emphasised using a linear abscissa and a logarithmic ordinate. The MRST(2001) curve is overlaid for reference.*

5.7.2 The 1994-2000 'O-final' fit

From §5.7.1, it seems that freeing $p_5 g$ appears suitable in the 'S' fit scheme (with fixed target data), however, it appears unsuitable for the fits using ZEUS-only data. Historically, for both the ZEUS-S and ZEUS-O fit styles, $p_3 g$ (high-x parameter) has been free. When $p_5 g$ (middling-x parameter) is freed in addition, it is arguable that too much freedom is allowed for the gluon distribution in the fits with ZEUS-only data. The data are simply not able to drive the fit for the entire gluon distribution. For this reason, a new fit (with ZEUS-only data) is attempted which provides low and middling-x freedom for the gluon, but fixes the high-x gluon to the value obtained in the 'S' fit (§5.7.1), where data existed to drive the fit. For clarity, this new fit style will be called the 'O-final' fit.

Table 25 summarises the parameter changes made in the main fit styles containing ZEUS-only data. It should be noted that the parameterisation of the gluon in the 'O-final' fit now mimics the parameterisation of the sea introduced in the 'O-modified' fit. Both the sea and gluon have high-x values pegged to the 'S' fit values (where data existed to drive the fit), whilst middling & low-x sea and gluon parameters are free to be determined by the fit.

'O' fit $\longrightarrow$	'O-modified' fit	$\longrightarrow$ 'O-final' fit
-	+ $p_2u = p_2d (\equiv p_2 valence)$ + $p_2 valence$=free	+ $p_2u = p_2d (\equiv p_2 valence)$ + $p_2 valence$=free
-	+ $p_3 S$ fixed (with $p_5 S$ free)	+ $p_3 S$ fixed (with $p_5 S$ free)
-	-	+ $p_3 g$ fixed
-	-	+ $p_5 g$ free

Table 25: *Summary of the parameter changes made to the 'O' fit, for the 'O-modified' and 'O-final' fits.*

When the 1994-2000 'O-final' fit is performed (with the same six data sets and cuts used in the 1994-2000 'O' and 'O-modified' fits) the revised parameterisation of the gluon ($p_3 g$ fixed, $p_5 g$ free) provides a more satisfactory picture. Figure 83 and tables 26,27 illustrate that the 1994-2000 'O-final' fit has almost exactly the same χ^2 and parameter values as the 1994-2000 'O-modified' fit, except for the values of p_3g and p_5g which have been fixed and freed respectively in the 'O-final' fit. The result shows a value of p_5g close to zero, thus the humpiness of the gluon has disappeared. By fixing the value of p_3g, the gluon distribution is not attenuated to zero as quickly, improving the consistency with MRST(2001), as can be seen when figure 84 is compared to figure 82.

The uncertainty on the d-valence distribution is shown in more detail in figure 85. The uncertainty is smaller than it was in the 1994-1998 '0' fit (figure 63), though not as small as was originally hoped when the 1994-2000 'O' fit was first attempted in figure 68 (without any model changes). The d-valence distribution is thus more accurately determined than in the 1994-1998 'O' fit, without the need for fixed target data. Keeping the fit limited to ZEUS-only data has meant that the uncertainties on the d and u-valence distributions are largely statistical, implying that more data obtained from HERA II should limit the uncertainties further.

Figures 86, 87 and 88 show the fit predictions for F_2^{em}, plus the NC and CC reduced cross sections, together with the data entering into the fit. Excellent agreement between fit and data is observed.

<table>
<tr><th>Data
Set</th><th>No. of data
points</th><th>χ^2</th><th>χ^2 per
data point</th><th>No. of correlated
systematic uncertainties</th><th>No. of normalisation
uncertainties</th></tr>
<tr><td>ZEUS $F_2$96-97</td><td>242</td><td>174.08</td><td>0.72</td><td>10</td><td>2</td></tr>
<tr><td>ZEUS CC e^+P94-97</td><td>29</td><td>15.407</td><td>0.53</td><td>3</td><td>1</td></tr>
<tr><td>ZEUS NC e^-P98-99</td><td>92</td><td>57.540</td><td>0.63</td><td>6</td><td rowspan="2">1</td></tr>
<tr><td>ZEUS CC e^-P98-99</td><td>26</td><td>23.166</td><td>0.89</td><td>3</td></tr>
<tr><td>ZEUS NC e^+P99-00</td><td>90</td><td>78.573</td><td>0.87</td><td>8</td><td rowspan="2">1</td></tr>
<tr><td>ZEUS CC e^+P99-00</td><td>30</td><td>33.431</td><td>1.11</td><td>3</td></tr>
<tr><td>TOTAL</td><td>509</td><td>382.20</td><td>0.75</td><td>33</td><td>5</td></tr>
</table>

Table 26: *χ^2 values for the data sets in the 1994-2000 'O-final' fit.*

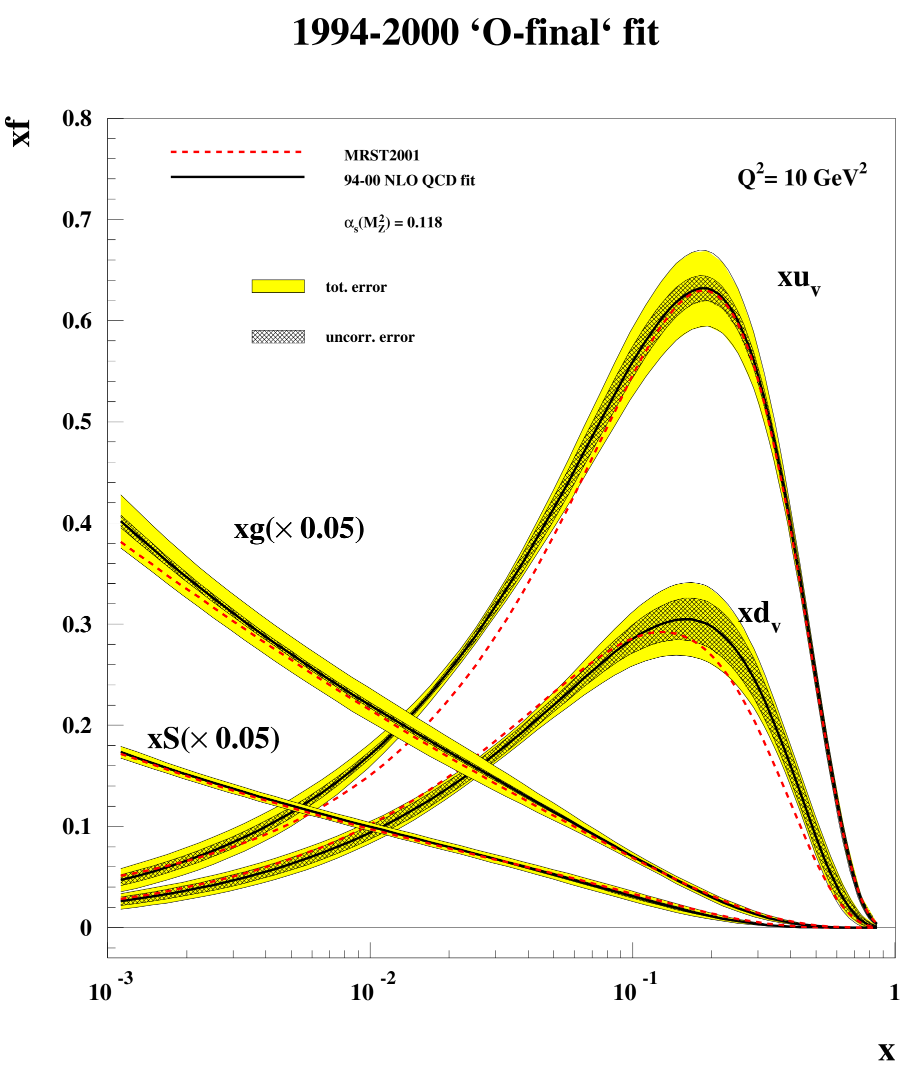

Figure 83: *The 1994-2000 'O-final' fit with MRST(2001) curves overlaid for reference.*

PDF	p_1	p_2	p_3	p_5
xu_v	*2.70	$0.60 \pm 0.05 \pm 0.08$	$3.97 \pm 0.16 \pm 0.20$	$2.62 \pm 1.15 \pm 1.08$
xd_v	*1.51		$4.44 \pm 0.73 \pm 0.77$	$2.2 \pm 2.0 \pm 1.5$
xS	$0.747 \pm 0.023 \pm 0.081$	$-0.210 \pm 0.004 \pm 0.015$	*8.9	$2.9 \pm 1.0 \pm 2.3$
xg	*1.54	$-0.22 \pm 0.01 \pm 0.05$	*6.2	$0.8 \pm 0.5 \pm 1.8$
$x\Delta$	*0.27	*0.5	*10.9	*0.0

Table 27: *PDF parameters at Q_0^2 for the 94-00 'O-final' fit. The first uncertainty printed arises from the statistical and uncorrelated errors. The second uncertainty is the contribution from correlated systematic sources. Starred values were either fixed at the start of the fit or derived from the free parameters of the fit.*

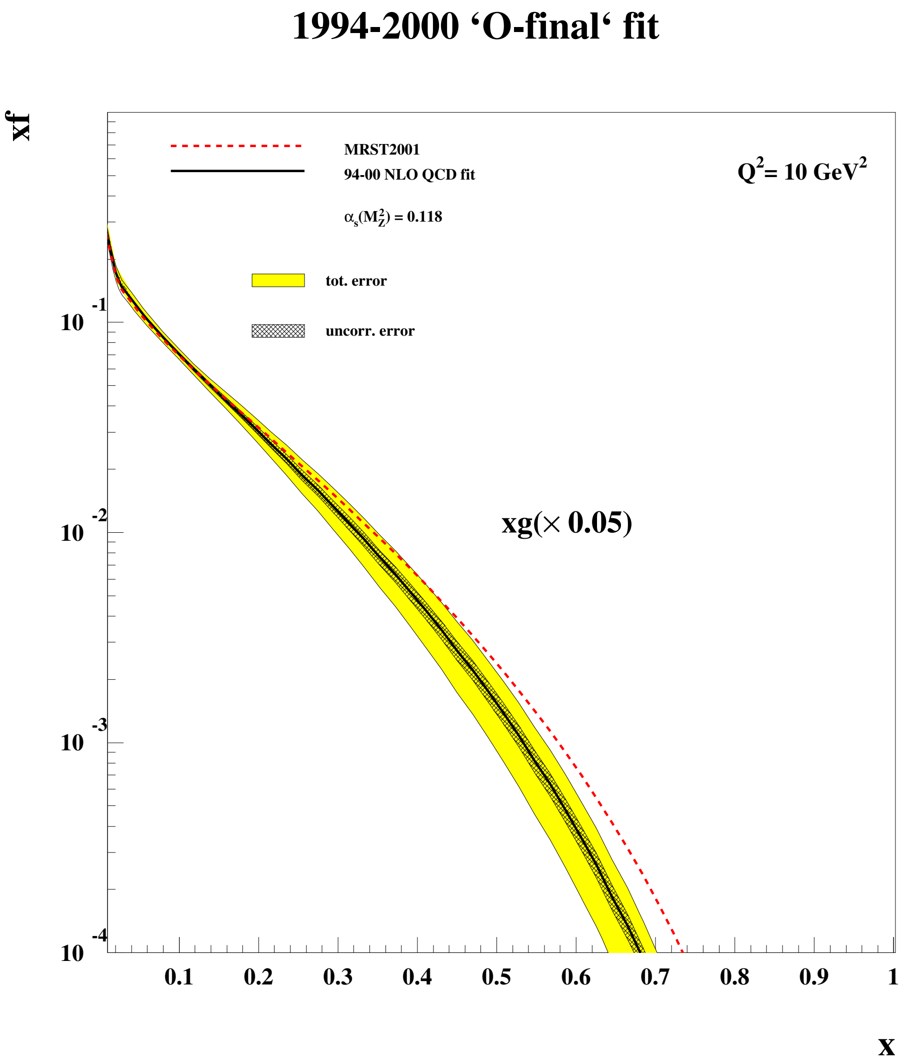

Figure 84: *The gluon distribution for the 1994-2000 ‘O-final’ fit. The high-x portion of the fit is emphasised using a linear abscissa and a logarithmic ordinate. The MRST(2001) curve is overlaid for reference.*

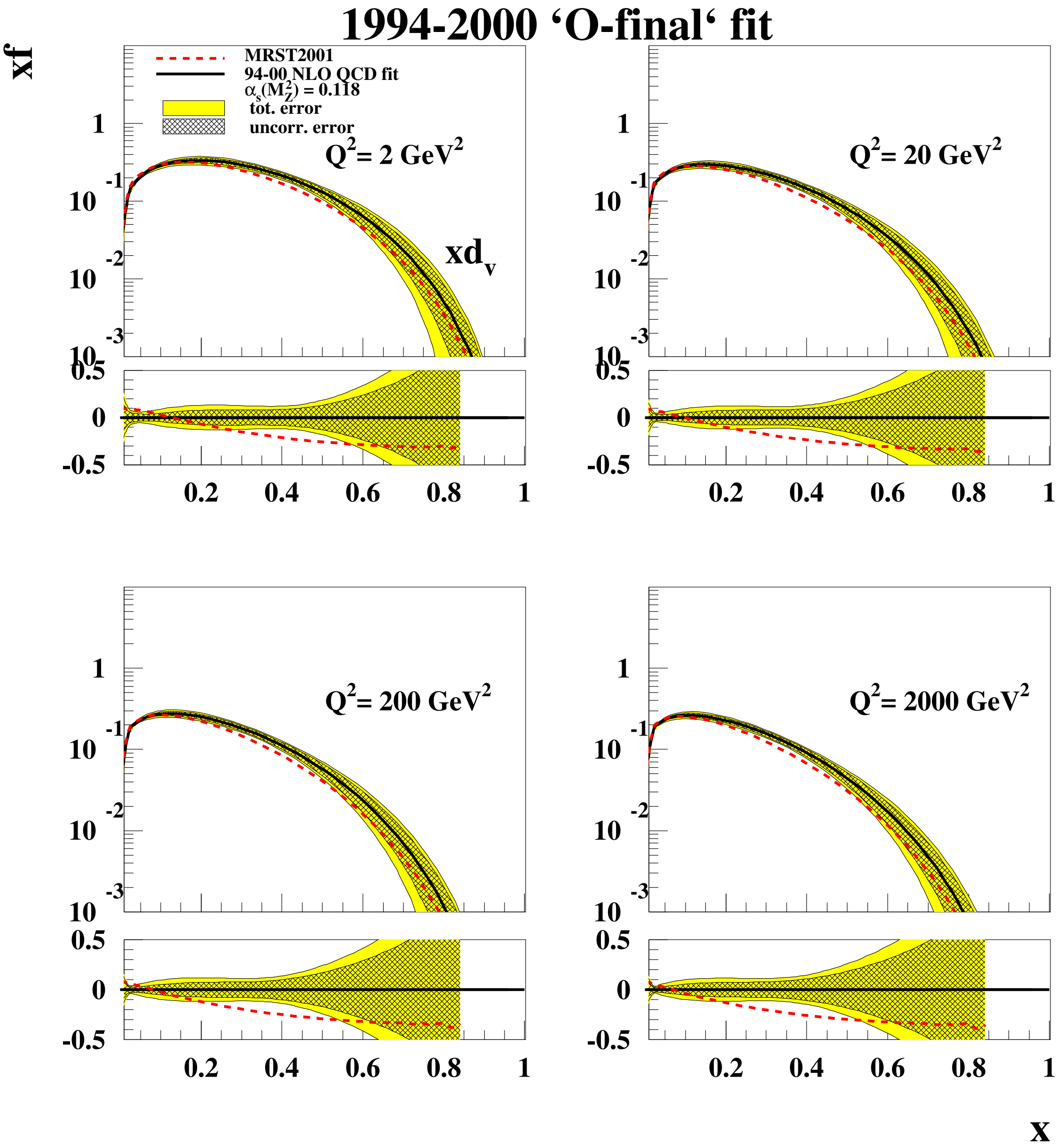

Figure 85: *d-valence plots at various Q^2 values for the 1994-2000 'O-final' fit. The uncertainties are shown for each Q^2 as fractional differences from the central value.*

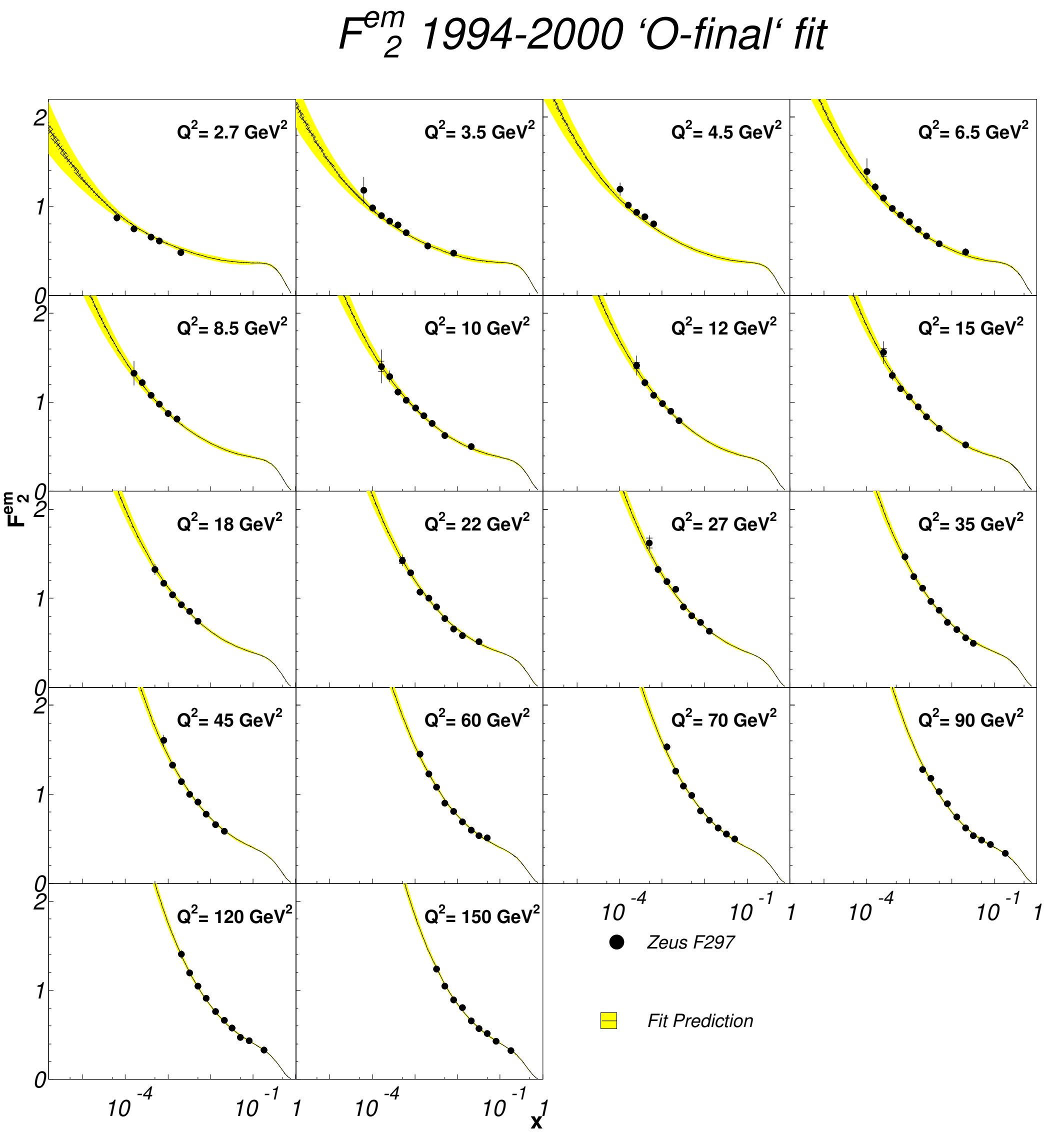

Figure 86: *The F_2^{em} fit prediction and ZEUS data in the 1994-2000 'O-final' fit, in bins of fixed Q^2.*

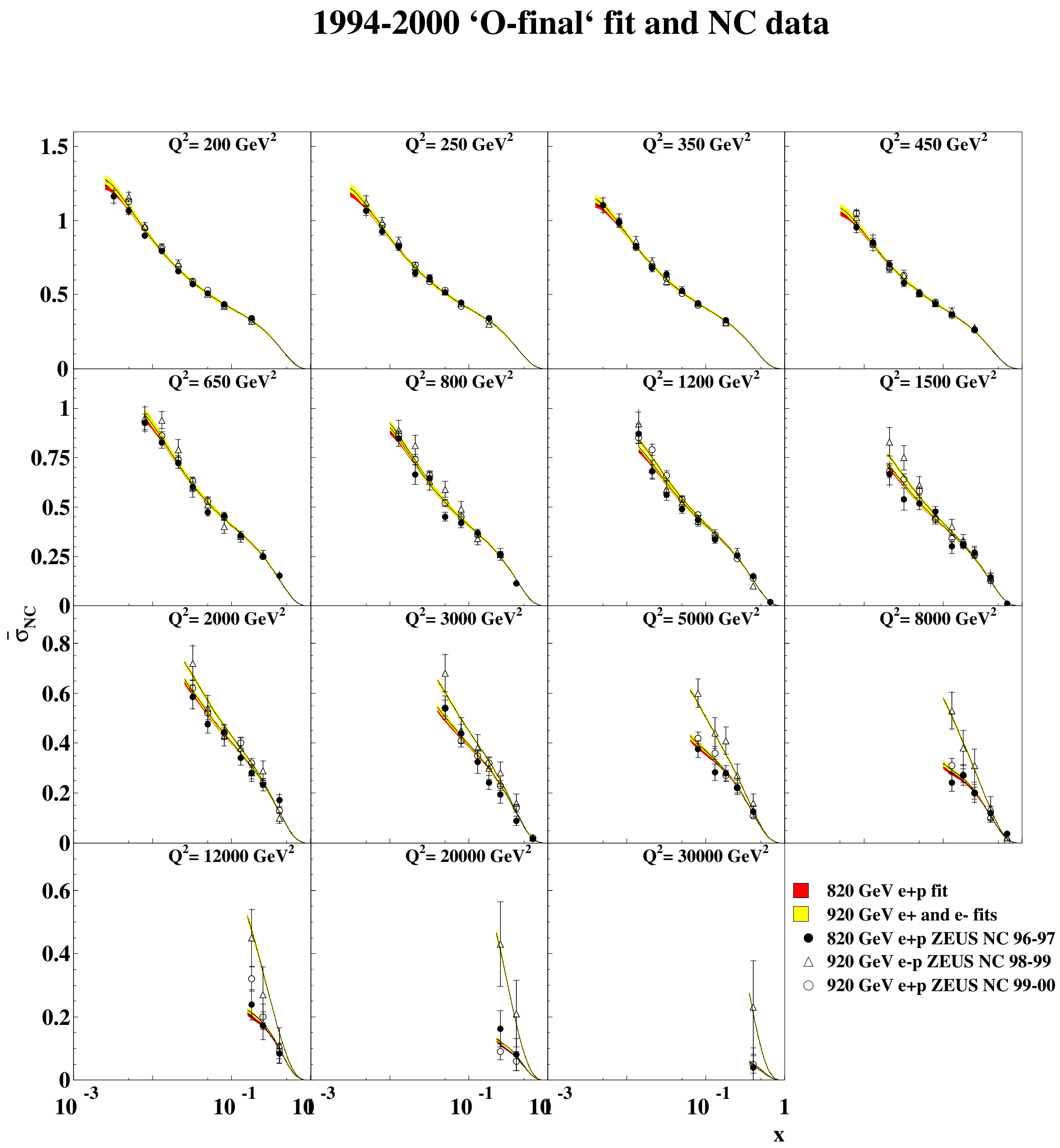

Figure 87: *The ZEUS reduced neutral current cross section data and 1994-2000 'O-final' fit predictions, in bins of fixed Q^2.*

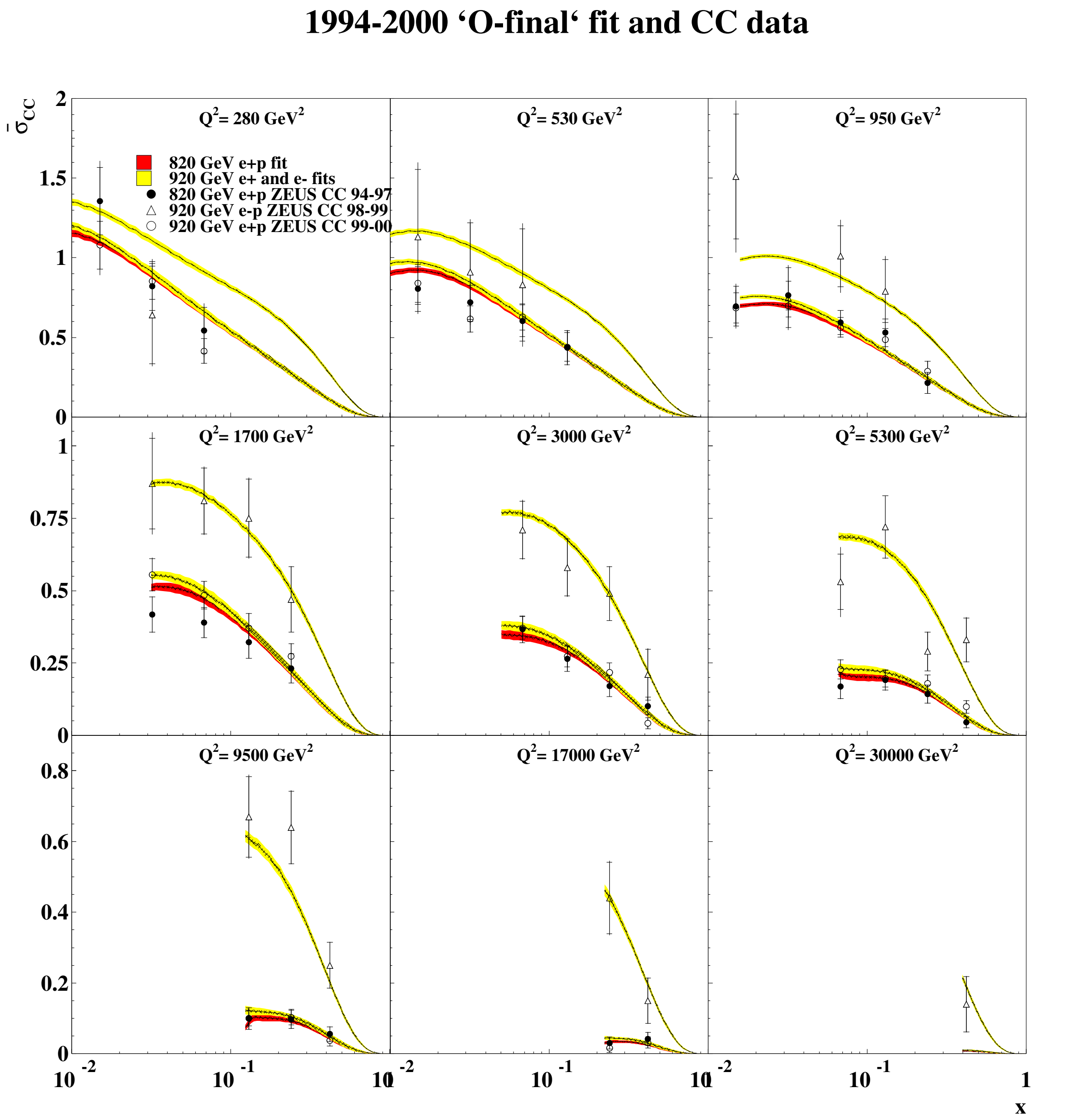

Figure 88: *The ZEUS reduced charged current cross section data and 1994-2000 'O-final' fit predictions, in bins of fixed* Q^2.

The prediction of the 1994-2000 'O-final' fit for F_2^{em} is extrapolated to low Q^2 in figure 89. Extra data are displayed (though not incorporated in the fit), namely the ZEUS shifted vertex (SVX) data [115], and the ZEUS beam-pipe-tracker (BPT) data [116]. The results are comparable with those from the previous analysis [80], where the lower limit to the applicability of the DGLAP equations was investigated. It seems that the 1994-2000 'O-final' fit is able to describe the data down to $Q^2 \sim 1.5\,\mathrm{GeV}^2$. Below this value, the fit fails to describe the data.

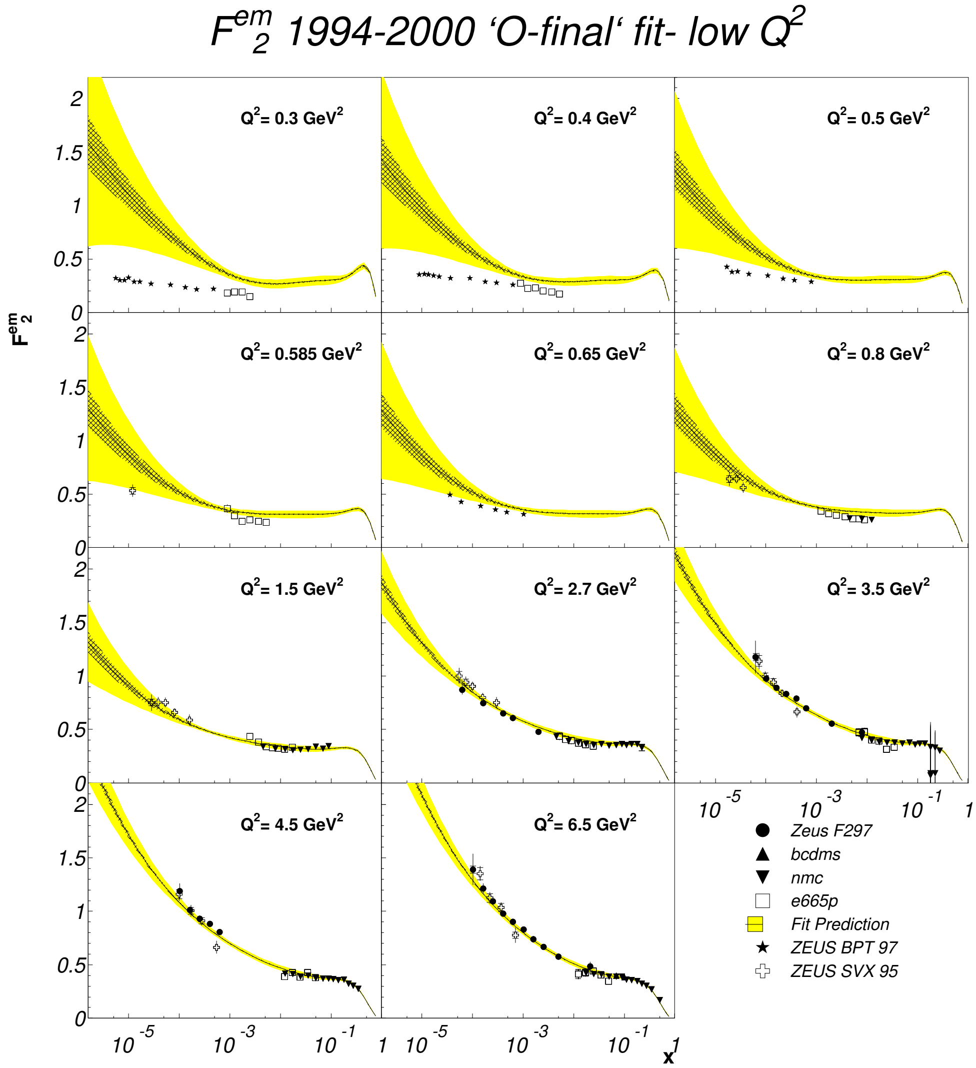

Figure 89: *The extrapolation of the 1994-2000 'O-final' fit prediction for F_2^{em} to the low-Q^2 region. The fit included the usual six ZEUS data sets, but NOT the BPT and SVX data (which is included on the plot, together with the fixed-target data). The extra fixed-target, BPT and SVX data on the plots serves only to illustrate where the fit ceases to be valid.*

5.8 Including H1 data

The gluon parameterisation in the 'O-final' fit allows freedom in the middling-x gluon parameter p_5g (which H1 has always allowed in the past). Additionally, work preceeding the 'O-final' fit has allowed improved fit predictions for the valence and sea distributions. To illustrate the effect of acquiring more statistics in the HERA kinematic region, a fit including both ZEUS and H1 data is now performed. Again, it is worth noting that the uncertainties in the valence distributions in the 1994-2000 'O-final' fit are largely statistical, meaning that extra HERA data should reduce this further. To this end, the 1994-2000 'O-final' fit was repeated with H1 data added. Specifically, the extra data included was:

- H1 minimum bias 97 e^+p NC [117]
- H1 low Q^2 96-97 e^+p NC [117]
- H1 high Q^2 94-97 e^+p NC [118]
- H1 high Q^2 94-97 e^+p CC [118]
- H1 high Q^2 98-99 e^-p NC [119], [110]
- H1 high Q^2 98-99 e^-p CC [119]
- H1 high Q^2 99-00 e^+p NC [110]
- H1 high Q^2 99-00 e^+p CC [110]

making a total of 14 data sets in the fit (including the six ZEUS data sets from the period 1994-2000).

Table 28 shows that the χ^2 per data point has risen relative to the same fit with ZEUS-only data, indicating the general incompatibility of the ZEUS and H1 data sets. In particular, the low-Q^2 ZEUS NC data seem to dislike being fitted along with the corresponding low-Q^2 H1 data. The χ^2 per data point of the ZEUS data rises from 174/242 in the 'O-final' fit to 218/242 when H1 data are included, similarly the highest values of χ^2 per data point in the H1 samples come from the low-Q^2 data.

The fit summary is shown in figures 90, 91 and table 29. Relative to the 'O-final' fit with ZEUS-only data, the uncertainties on the distributions have reduced, and the fit seems slightly less compatible with MRST(2001) as a result. The central values of the fit are generally comparable to the 'O-final' fit with ZEUS-only data, but a few important differences exist. The valence distributions are shifted higher when H1 data are included, and the low-x gluon parameter ($p_2 g$) is less negative. Interestingly, the middling-x gluon parameter $p_5 g$ is again low and consistent with zero, meaning there is little tendency to create the gluon 'hump' previously seen with H1 data. The d-valence distributions show marked decrease in uncertainty at high-x (comparing figures 91 and 85).

Data Set	No. of data points	χ^2	χ^2 per data point	No. of corr. syst. unc.	No. of norm. unc.
ZEUS $F_2$96-97	242	217.57	0.90	10	2
ZEUS CC e^+P94-97	29	16.372	0.56	3	1
ZEUS NC e^-P98-99	92	58.790	0.64	6	1
ZEUS CC e^-P98-99	26	23.499	0.90	3	
ZEUS NC e^+P99-00	90	89.822	1.00	8	1
ZEUS CC e^+P99-00	30	30.949	1.03	3	
H1 minimum bias 97 NC e^+p	55[9]	65.324	1.19	5	2[10]
H1 low Q^2 96-97 NC e^+p	80	112.50	1.41	0(+5)[11]	1
H1 high Q^2 94-97 NC e^+p	130	95.828	0.74	4(+1)[11]	1
H1 high Q^2 94-97 CC e^+p	25	18.580	0.74	1(+3)[11]	
H1 high Q^2 98-99 NC e^-p	139[12]	127.99	0.92	2(+4)[11]	1
H1 high Q^2 98-99 CC e^-p	28	17.621	0.63	1(+3)[11]	
H1 high Q^2 99-00 NC e^+p	147	137.28	0.93	0(+6)[11]	1
H1 high Q^2 99-00 CC e^+p	28	28.551	1.02	0(+4)[11]	
TOTAL	1141	1040.67	0.91	46	11

Table 28: *χ^2 values for the data sets in the 1994-2000 'O-final' fit, with H1 data included.*

[9] Of the published data points, twelve at low-Q^2 are removed by the data cut at $Q^2 = 2.5$.

[10] A common luminosity uncertainty of 0.5% is applicable to all the H1 data sets. This common uncertainty has been applied to all data sets in the minimisation, but is listed under the first H1 data set in the table. The reader is referred to [110] for the detail of the H1 luminosity uncertainties.

[11] Some correlated systematic uncertainties are applied by H1 which span several data sets. The first figure for each data set represents only the new (and therefore as-yet unlisted) correlated systematic uncertainties, whilst the figure in brackets represents the correlated systematics already listed separately in another data set. The reader is referred to [110] for the detail of the H1 correlated systematic uncertainties.

[12] The paper [119] supplies 126 data points, and an additional 13 data points in this sample came from the retrospective high-y analysis made in [110].

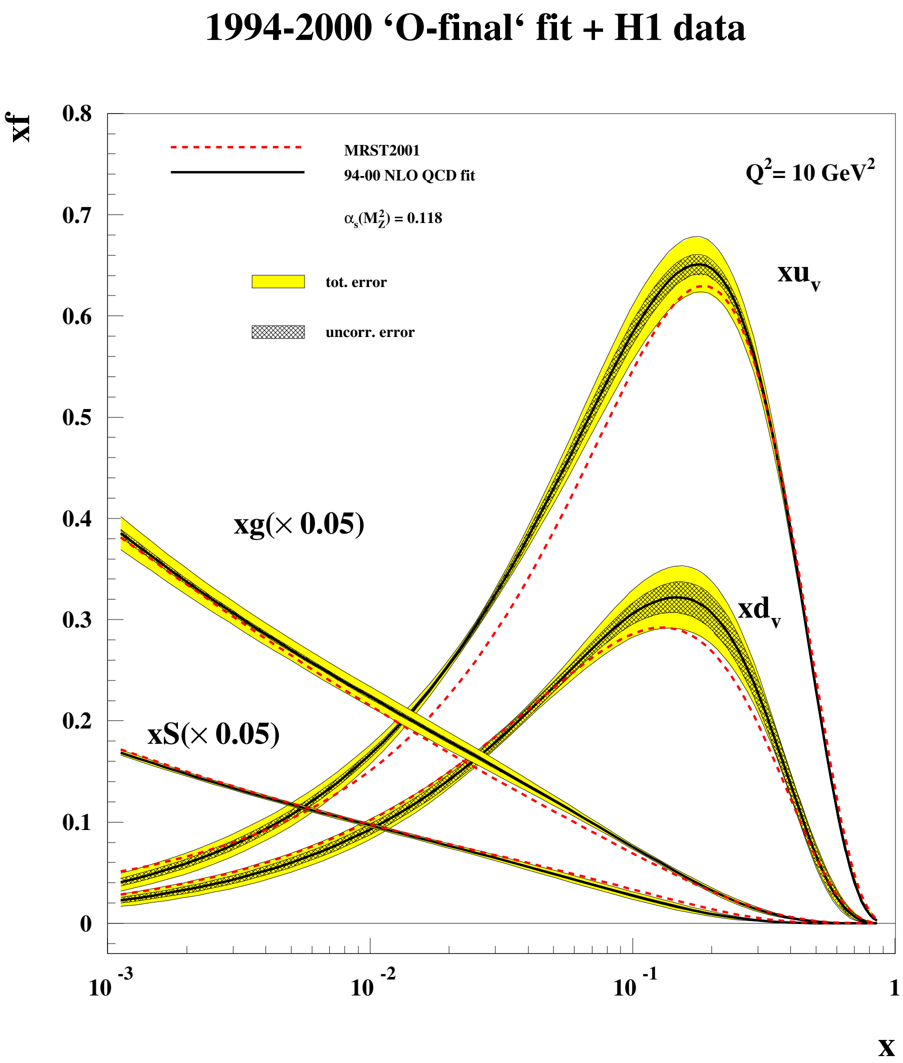

Figure 90: *The 1994-2000 'O-final' fit with H1 data added. MRST(2001) curves are overlaid for reference.*

PDF	p_1	p_2	p_3	p_5
xu_v	*3.47	$0.66 \pm 0.04 \pm 0.07$	$4.13 \pm 0.12 \pm 0.09$	$1.96 \pm 0.75 \pm 1.02$
xd_v	*1.95		$5.35 \pm 0.58 \pm 0.69$	$2.8 \pm 1.8 \pm 1.3$
xS	$0.745 \pm 0.016 \pm 0.038$	$-0.206 \pm 0.003 \pm 0.007$	*8.9	$1.5 \pm 0.7 \pm 1.8$
xg	*1.81	$-0.18 \pm 0.01 \pm 0.03$	*6.2	$0.6 \pm 0.3 \pm 1.1$
$x\Delta$	*0.27	*0.5	*10.9	*0.0

Table 29: *PDF parameters at Q_0^2 for the 94-00 'O-final' fit with H1 data added. The first uncertainty printed arises from the statistical and uncorrelated errors. The second uncertainty is the contribution from correlated systematic sources. Starred values were either fixed at the start of the fit or derived from the free parameters of the fit.*

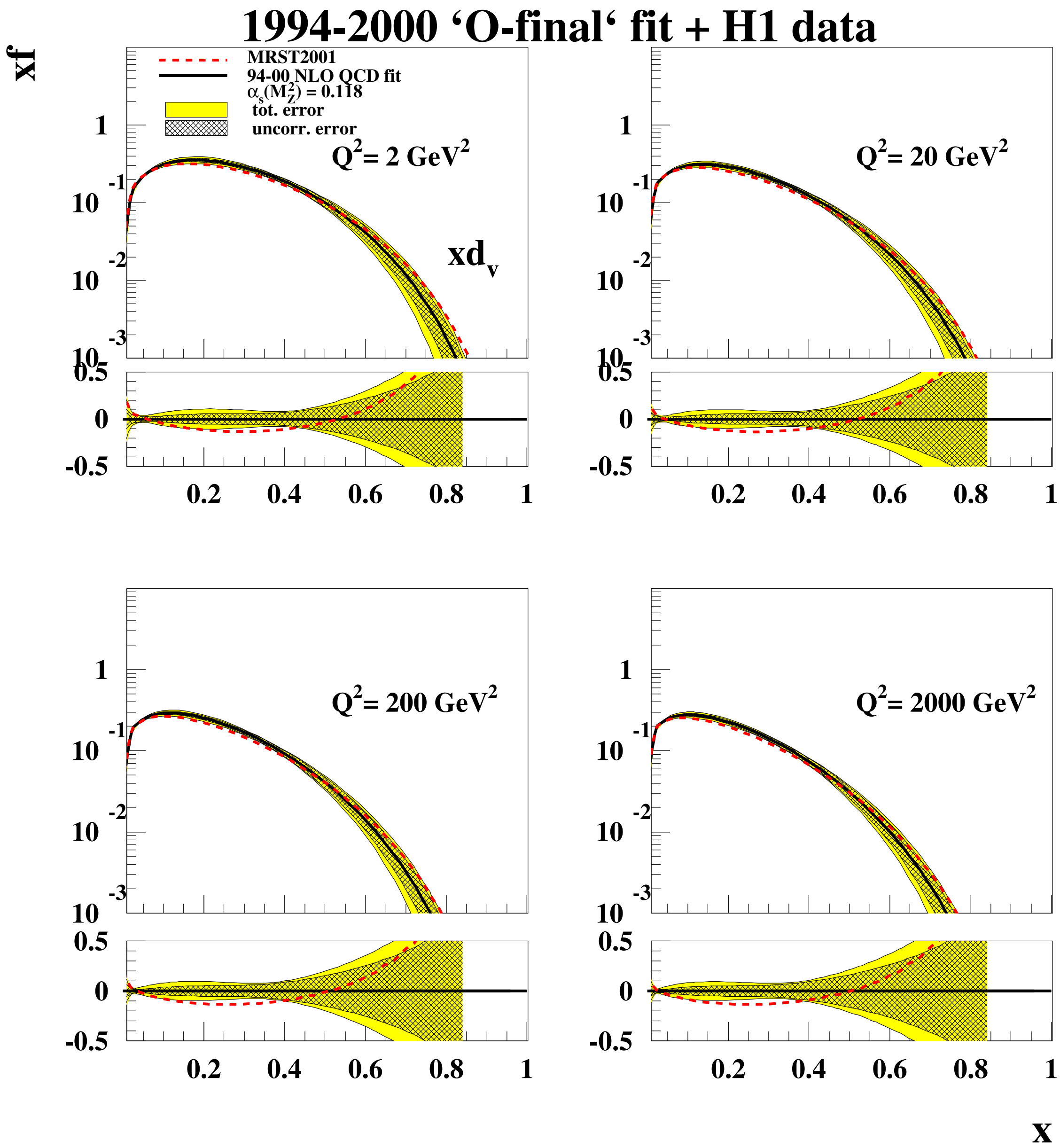

Figure 91: *d-valence plots at various Q^2 values for the 1994-2000 ‘O-final’ fit with H1 data included. The uncertainties are shown for each Q^2 as fractional differences from the central value.*

Figures 92, 93 and 94 show the ZEUS and H1 data, together with the fit prediction for the reduced NC cross section (low & high-Q^2) and the high-Q^2 reduced CC cross section respectively. In the low-$Q2$ NC region (figure 92), it appears as if there is a small systematic difference between the ZEUS and H1 data, which seems to tally with the already identified increase in χ^2 per data point for these samples when they are fitted together.

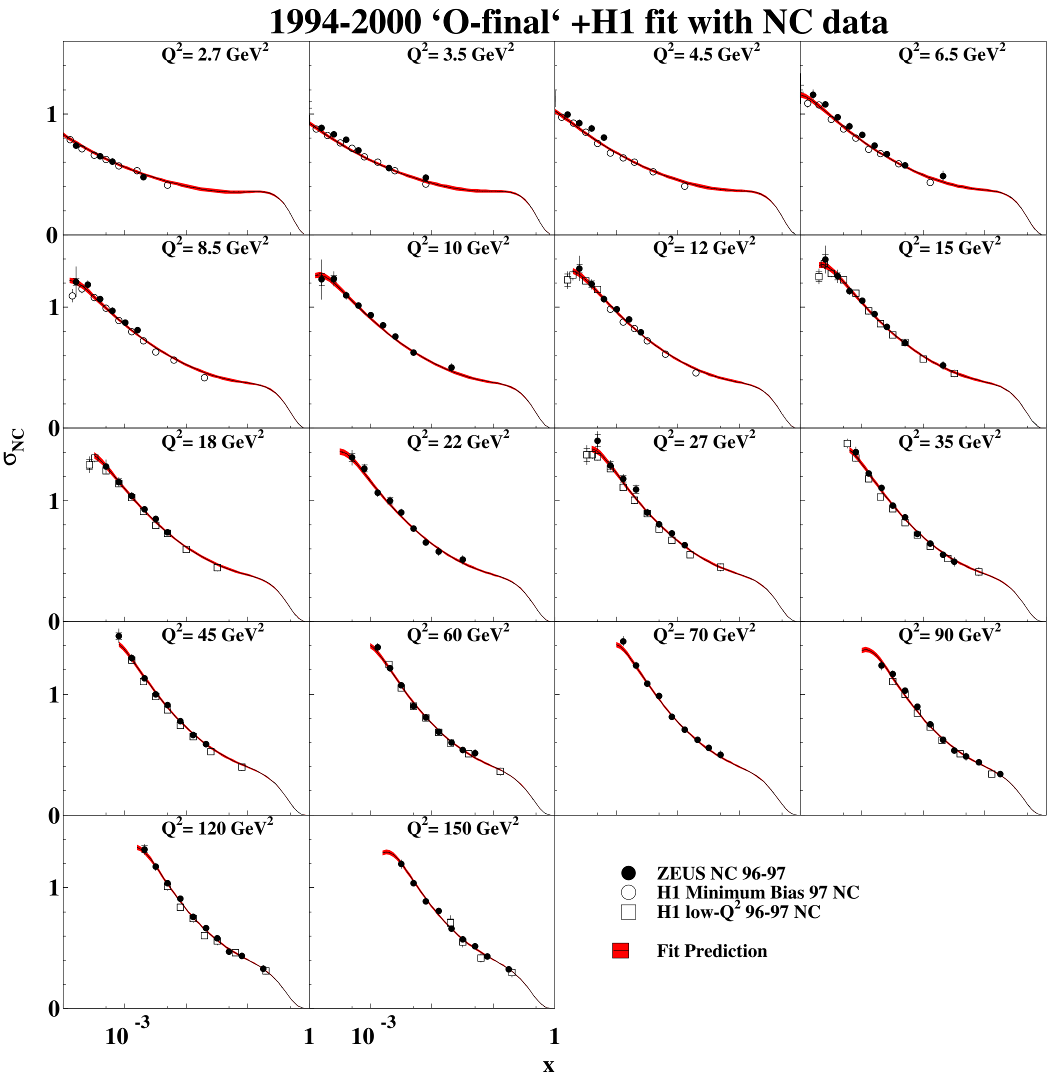

Figure 92: *The low-Q^2 reduced neutral current cross section fit prediction, together with the ZEUS and H1 data in the 1994-2000 'O-final'+H1 fit, in bins of fixed Q^2.*

1994-2000 'O-final' +H1 fit with NC data

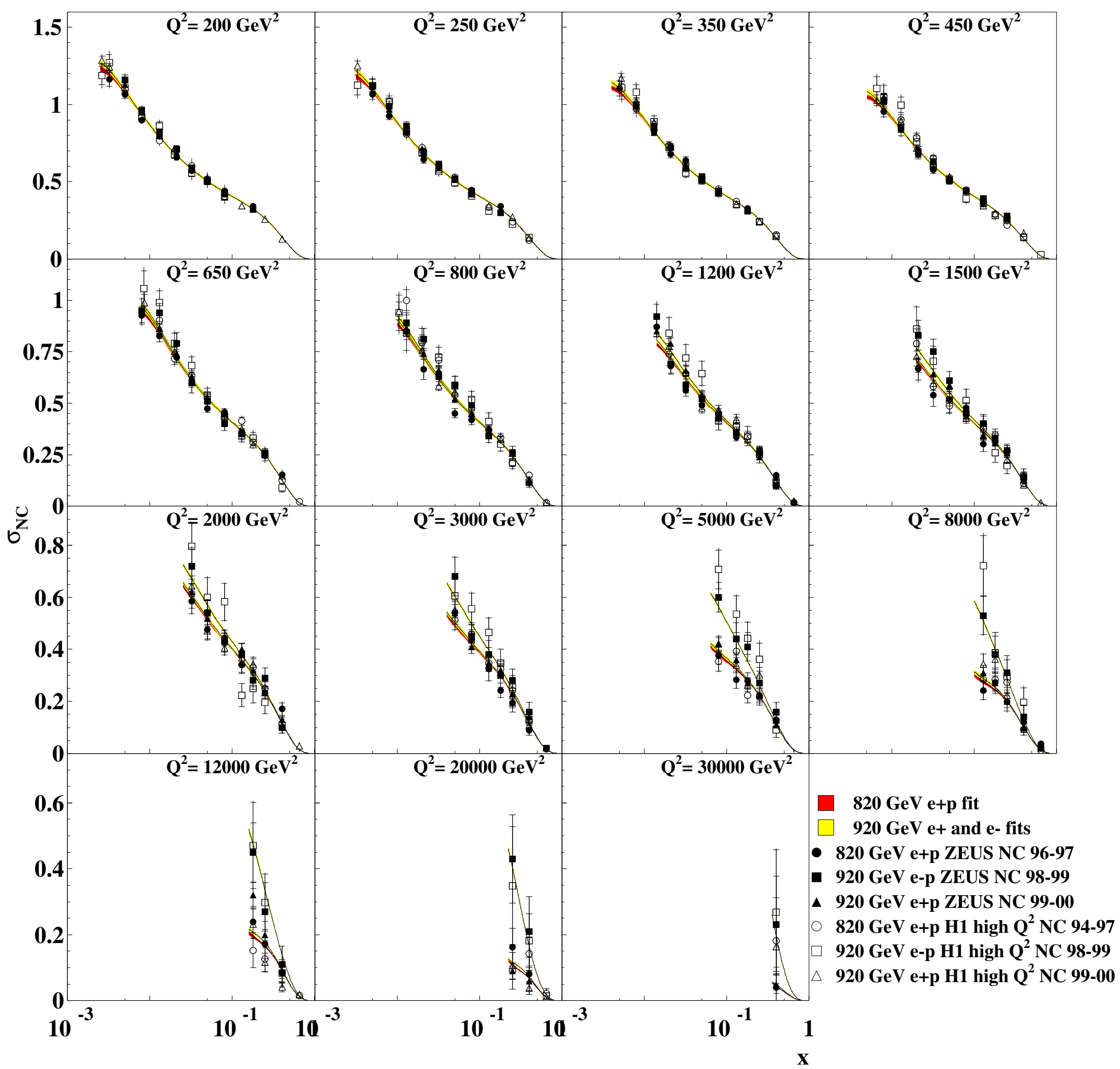

Figure 93: *The ZEUS and H1 reduced neutral current cross section data and 1994-2000 'O-final'+H1 fit predictions, in bins of fixed Q^2.*

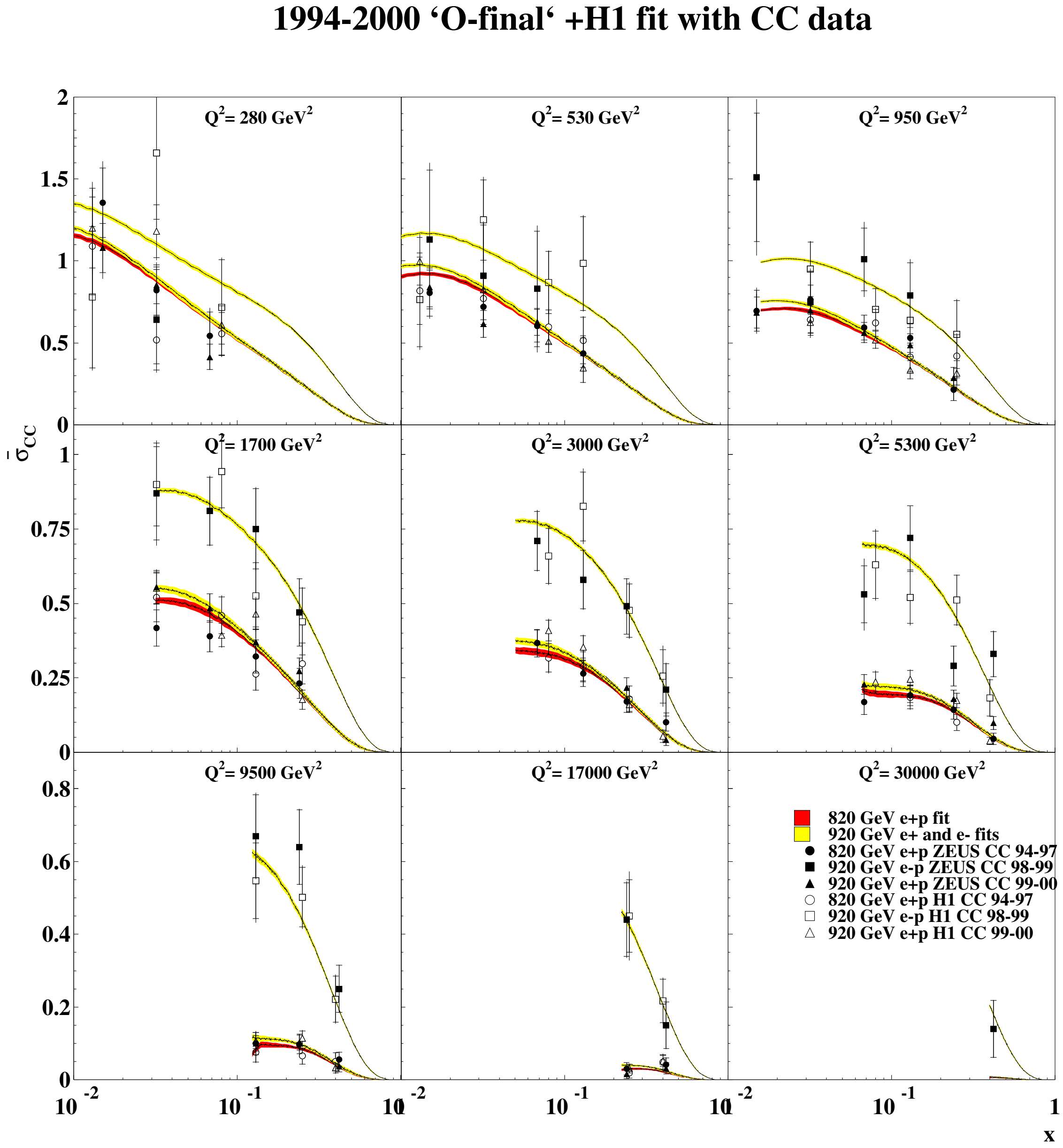

Figure 94: *The ZEUS and H1 reduced charged current cross section data and 1994-2000 'O-final'+H1 fit predictions, in bins of fixed Q^2.*

5.9 Model dependence of the 1994-2000 'O-final' fit

The 1994-2000 'O-final' fit (§5.7.2) has been established as being more consistent with the global fit and MRST(2001) than the other fits with ZEUS-only data which were tried. The dependence of this fit on the model assumptions is now tested.

5.9.1 Variation of x_{min} and Q^2_{min} cuts

The kinematic range of the data used in the 1994-2000 'O-final' fit was $6.3\times10^{-5} < x < 0.65$, $2.5 < Q^2 < 30000\,\mathrm{GeV}^2$. The minimum values of x and Q^2 for data entering into the fit were varied, and the effects on the central values of the PDFs noted for the following new cuts: $Q^2 > 4.5\,\mathrm{GeV}^2$; $x > 10^{-3}$.

<table>
<tr><th>PDF</th><th colspan="2">Δp_1</th><th colspan="2">Δp_2</th><th colspan="2">Δp_3</th><th colspan="2">Δp_5</th></tr>
<tr><th></th><th>x_{min}</th><th>Q^2_{min}</th><th>x_{min}</th><th>Q^2_{min}</th><th>x_{min}</th><th>Q^2_{min}</th><th>x_{min}</th><th>Q^2_{min}</th></tr>
<tr><td>xu_v</td><td>-</td><td>-</td><td rowspan="2">+0.02</td><td rowspan="2">+0.03</td><td>+0.01</td><td>-0.02</td><td>-0.23</td><td>-0.48</td></tr>
<tr><td>xd_v</td><td>-</td><td>-</td><td>+0.13</td><td>+0.15</td><td>+0.4</td><td>+0.4</td></tr>
<tr><td>xS</td><td>+0.038</td><td>-0.010</td><td>+0.009</td><td>-0.002</td><td>-</td><td>-</td><td>-0.66</td><td>-0.28</td></tr>
<tr><td>xg</td><td>-</td><td>-</td><td>+0.01</td><td>+0.07</td><td>-</td><td>-</td><td>-0.3</td><td>-1.3</td></tr>
</table>

Table 30: *Shift in central PDF values of the 94-00 'O-final' fit at $Q^2_0 = 7$ from the cut on $x_{min} = 0.001$, and $Q^2_{min} = 4.5$. Model uncertainties for fit parameters which were fixed or derived from free parameters of the fit are left blank (-).*

The most significant effect occurs in the gluon distribution (p_2g, p_5g) when the Q^2_{min} cut is changed, however the effect of the cuts is otherwise small compared to the statistical and systematic uncertainties.

Interestingly, investigations performed here (see figure 89) suggest that in this analysis, the DGLAP equations are valid down to $Q^2 \sim 1.5\,\mathrm{GeV}^2$, however a recent MRST analysis [120] suggests that a conservative limit of $Q^2 = 10\,\mathrm{GeV}^2$ might be more applicable for the DGLAP formalism. In any event, as data are removed from the fit, the model uncertainty appears to rise in this region.

5.9.2 Adding more middling-x freedom using p_4

The form of the parameterisation in equation (55) was extended in several fits from $(1 + p_5x)$ to $(1 + p_4\sqrt{x} + p_5x)$. This builds in extra flexibility in the intermediate-x range, where the PDFs change rapidly $(0.01 < x < 0.3)$. There never was any χ^2 advantage found when these parameters were freed and additionally, these fits had problems when the error analysis was made. The freeing of p_4 led to Hessian error matrices with negative eigenvalues. This is unphysical as the eigenvalues of these matrices correspond to the square root of a physical quantity (the error on the parameter combination defined by the eigenvector, the square root of which should always positive)[13]. It appears that there is simply insufficient data in the fits to justify an extra parameter controlling the middling-x region.

5.9.3 Changing the value of Q_0^2

The dependence of the χ^2 value on Q_0^2 was analysed, to investigate if changing the origin of evolution of the PDFs had any effect on the fit itself. Q_0^2 was shifted from its initial value of 7, to 4 and 10 GeV^2.

The changes observed were not significant, and this effect has been neglected in the model uncertainty.

5.9.4 Model uncertainties due to parameter fixing

The act of fixing parameters p_3S & p_3g in the 1994-2000 'O-final' fit introduces a model uncertainty, since the fixed parameters necessarily have uncertainties of zero. When the parameters are fixed, the inter-parameter correlations also change, meaning that the remainimg free parameters will have different central

[13]The reader is encouraged to see (for example) [121] to understand the consequences of the Hessian matrix having negative eigenvalues

values and uncertainties, in order to find the best miniumum in the new scheme. This means that the uncertainties are redistributed to a certain extent.

A conservative estimate of the uncertainty introduced by fixing parameters was made. This was done by running fits where the parameters previously fixed to the best central value estimate were changed such that they were fixed to the upper and lower bounds of the range allowed in the 'S' fit (where the parameters were free). The central values obtained from these new *shifted* central value fits were used as a crude estimate of the PDF uncertainty lost by fixing the parameter in the first place. Four fits were performed ($p_3 S$ up & down; $p_3 g$ up & down), and the estimates of the uncertainties combined for each parameter. The positive and negative deviations for each parameter were added separately in quadrature. The values are shown in table 31.

PDF	Δp_1	Δp_2	Δp_3	Δp_5
xu_v	-	± 0.01	$^{+0.10}_{-0.04}$	$^{+0.61}_{-0.39}$
xd_v	-		$^{+0.23}_{-0.08}$	$^{+1.05}_{-0.81}$
xS	$^{+0.007}_{-0.015}$	$^{+0.008}_{-0.004}$	-	$^{+1.78}_{-1.63}$
xg	-	$^{+0.00}_{-0.03}$	-	$^{+6.4}_{-1.8}$

Table 31: *Total shift in central PDF values of the 94-00 'O-final' fit at $Q_0^2 = 7$, resulting from the fits made with shifted central values. Model uncertainties for fit parameters which were fixed or derived from free parameters of the fit are left blank (-).*

5.9.5 Final model uncertainty

The final model uncertainty was estimated for each parameter by adding in quadrature the positive and negative individual model uncertainties. Figure 95 and table 32 illustrate the model uncertainties as they apply to the PDFs. The model uncertainty appears to be most significant in the gluon distribution.

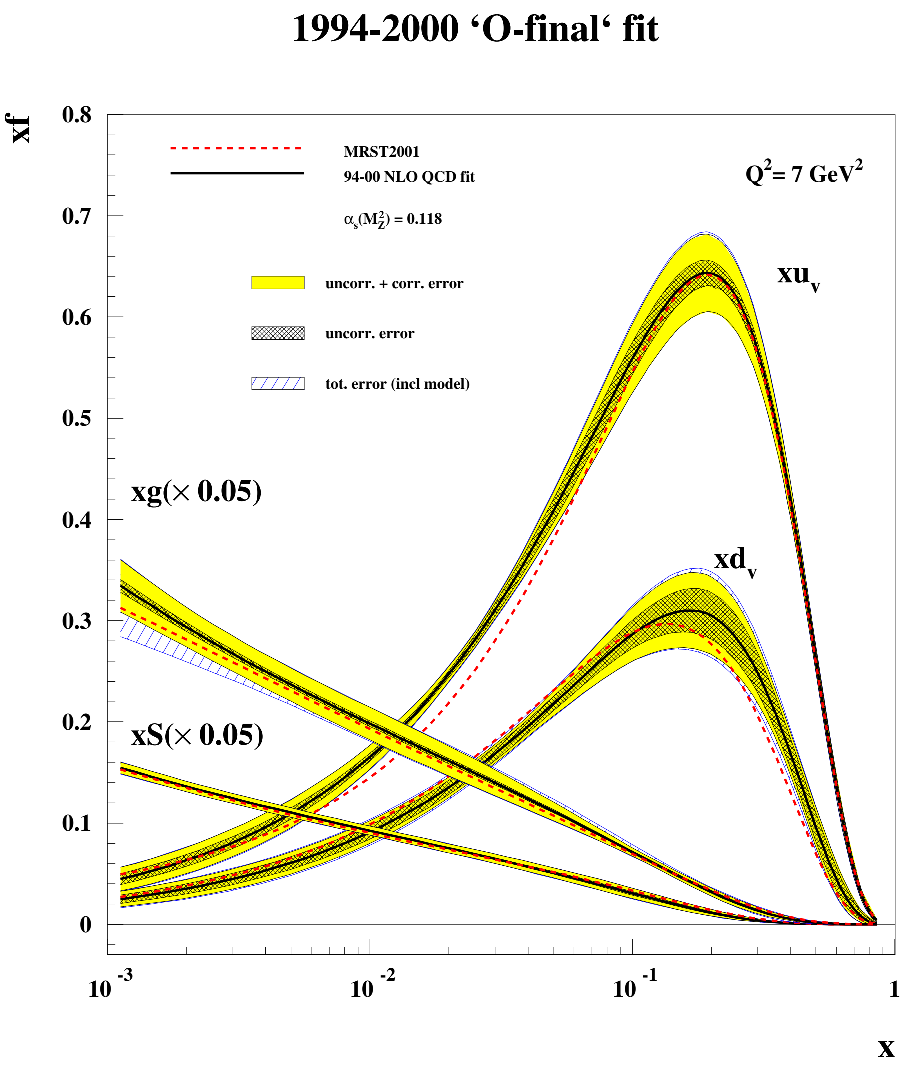

Figure 95: *The 1994-2000 'O-final' fit at $Q^2 = Q_0^2 = 7\,GeV^2$, with model uncertainties added. MRST(2001) curves are overlaid for reference.*

PDF	p_1	p_2	p_3	p_5
xu_v	*2.70	$0.60 \pm 0.05 \pm 0.08^{+0.04}_{-0.01}$	$3.97 \pm 0.16 \pm 0.20^{+0.10}_{-0.04}$	$2.62 \pm 1.15 \pm 1.08^{+0.61}_{-0.66}$
xd_v	*1.51		$4.44 \pm 0.73 \pm 0.77^{+0.30}_{-0.08}$	$2.2 \pm 2.0 \pm 1.5^{+1.17}_{-0.81}$
xS	$0.747 \pm 0.023 \pm 0.081^{+0.039}_{-0.018}$	$-0.210 \pm 0.004 \pm 0.015^{+0.012}_{-0.004}$	*8.9	$2.9 \pm 1.0 \pm 2.3 \pm 1.8$
xg	*1.54	$-0.22 \pm 0.01 \pm 0.05^{+0.07}_{-0.03}$	*6.2	$0.8 \pm 0.5 \pm 1.8^{+6.4}_{-2.2}$
$x\Delta$	*0.27	*0.5	*10.9	*0.0

Table 32: *PDF parameters at Q_0^2 for the 94-00 'O-final' fit. The first uncertainty printed arises from the statistical and uncorrelated errors. The second uncertainty is the contribution from correlated systematic sources. The third figure arises from the model uncertainties. Starred values were either fixed at the start of the fit or derived from the free parameters of the fit.*

5.10 Conclusions

The global fits, which include fixed-target data, have many correlated systematic uncertainties which dominate the PDF uncertainties. The published ZEUS-O fit [80] removes the fixed-target data and systematics, but suffers from lack of data and consequently its predictions have large bands of statistical uncertainty, particularly at high x.

By adding the 1999-2000 data sets to the 1994-1998 'O' fit, improvement in the uncertainty of the d-valence distribution is seen. However, the uncertainties appear to be generally over-optimistic. The fit central values lie far from the ZEUS-S published values [80], and the uncertainties (reduced by the 1999-2000 data) then imply disagreement between the fits. In light of the 1999-2000 ZEUS data, the 'O' fits thus required more investigation of the parameterisation dependence.

Performing the 1994-2000 'O-final' fit (i.e. fixing the parameters p_3S & p_3g, in addition to allowing $p_2valence$ and p_5g to go free), provides good improvement in fit central values and maintains some improvement in uncertainties over the 1994-1998 'O' fit. The 1994-2000 'O-final' fit, involving ZEUS-Only data, now appears much closer to MRST(2001) and the global 'S' fit than was previously observed in the published ZEUS-O fit [80]. The uncertainty of the d-valence distribution is also reduced relative to the published ZEUS-O fit. In general, the uncertainties are still largely statistical in nature, meaning that further ZEUS data should allow future improvement. Adding the H1 data to the 1994-2000 'O-final' fit does indeed improve the uncertainties, illustrating the impact further data in the HERA kinematic region may have.

The 1994-2000 'O-final' fit provides a successful description of the cross section and F_2^{em} data, and also appears to successfully predict the behaviour of F_2^{em} to Q^2 values as low as $\sim 1.5\,\mathrm{GeV}^2$.

Looking to the future, it may be possible to gain more information on the gluon and sea distributions in order that the high-x behaviour can be fully determined by ZEUS data alone. Further high-x data from ZEUS should constrain the sea distribution (adding H1 data did show improvement in the sea distribution uncertainties). In addition, work is currently ongoing to incorporate ZEUS jet and $F_2 charm$ data into the fits. This should provide information on the high-x gluon distribution, since boson-gluon fusion contributes strongly to both charm production and di-jet cross sections.

References

[1] F. Halzen and A.D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics.* John Wiley & Sons, Inc, 1984.

[2] A. Salam, *Elementary Particle Physics.* (1968). Reprinted in Gauge Theory of Weak and Electromagnetic Interactions, C.H. Lai, ed.(Singapore: World Scientific, 1981).

[3] Super-Kamiokande Coll., Y. Fukuda et al., Phys. Rev. Lett. **81**, 1562 (1998).

[4] Q.R. Ahmad et al., Phys. Rev. Lett. **89**, 011301 (2002).

[5] KamLAND Coll., K. Eguchi et al., Phys. Rev. Lett. **90**, 021802 (2003).

[6] SOUDAN 2 Coll., M. Sanchez et al., Preprint hep-ex/0307069, 2003. Subm. to Phys. Rev. D.

[7] E.D. Bloom et al., Phys. Rev. Lett. **23**, 930 (1969).

[8] M. Breidenbach et al., Phys. Rev. Lett. **23**, 935 (1969).

[9] R.P. Feynman, *Photon-Hadron Interactions.* Benjamin, New York, 1972.

[10] M. Gell-Mann, Phys. Lett. **8**, 214 (1964).

[11] J.D. Bjorken, Phys. Rev. **179**, 1547 (1969).

[12] M. Diemoz, M. Ferroni and E. Longo, Phys. Rep. **130**, 293 (1986).

[13] Donald H. Perkins, *Introduction to High Energy Physics.* Cambridge U. Press, 1972.

[14] T.P. Cheng and L.F. Li, *Gauge Theory of Elementary Particle Physics.* Clarendon Press (Oxford Science Publications), Oxford, UK, 1984.

[15] A.M. Cooper-Sarkar, R.C.E. Devenish and A. De Roeck, Int. J. Mod. Phys. **A 13**, 3385 (1998).

[16] ZEUS Coll., U. Holm (ed.), *The ZEUS Detector.* Status Report (unpublished), DESY (1993), available on `http://www-zeus.desy.de/bluebook/bluebook.html`.

[17] B. Foster et al., Nucl. Inst. Meth. **A 338**, 254 (1994).

[18] A. Quadt and O. Ruske, *A New Method to Measure Vertex Distributions* (unpublished). ZEUS-98-036, internal ZEUS note, 1998.

[19] N. Harnew et al., Nucl. Inst. Meth. **A 279**, 290 (1989).

[20] A. Quadt et al., Nucl. Inst. Meth. **A 438**, 472 (1999).

[21] R. Hall-Wilton et al., *The CTD Tracking Resolution* (unpublished). ZEUS-99-024, internal ZEUS-note, 1999.

[22] A. Andresen et al., Nucl. Inst. Meth. **A 309**, 101 (1991).

[23] M. Derrick et al., Nucl. Inst. Meth. **A 309**, 77 (1991).

[24] A. Bernstein et al., Nucl. Inst. Meth. **A 336**, 23 (1993).

[25] K. Olkiewicz and A. Eskreys, *Off-line Luminosity Calculation in the ZEUS Experiment in 1997, 1998 and 1999* (unpublished). ZEUS-99-44, 1999.

[26] ZEUS Coll., R. Carlin et al., Nucl. Inst. Meth. **A 379**, 542 (1996).

[27] W.H. Smith et al., Nucl. Inst. Meth. **A 355**, 278 (1995).

[28] G.P. Heath et al., Nucl. Inst. Meth. **A 315**, 431 (1992).

[29] J.A. Crittenden, *The C5 Upgrade* (unpublished). ZEUS-01-002, internal ZEUS note, 2001.

[30] C. Collins-Tooth et al., *The Transverse Polarimeter (TPOL) Test Beam at DESY in January-February 2001* (unpublished), June 2001, available on `http://www-zeus.desy.de/ZEUS_ONLY/zeus_notes/notes01.html`. ZEUS-01-019.

[31] T. Behnke et al., *The Transverse Polarimeter (TPOL) Test Beam at CERN in July-August 2001* (unpublished), October 2002, available on `http://www-zeus.desy.de/ZEUS_ONLY/zeus_notes/notes02.html`. ZEUS-02-019.

[32] I.M. Ternov, Yu.M. Loskutov and L.I. Korovina, Sov. Phys. JETP **14**, 921 (1962).

[33] A.A. Sokolov and I.M. Ternov, Sov. Phys. Dokl. **8**, 1203 (1964).

[34] D.P. Barber et al., Nucl. Inst. Meth. **A 338**, 166 (1994).

[35] D.P. Barber et al., Nucl. Inst. Meth. **A 329**, 79 (1993).

[36] M. Ruh et al., *Construction and Beam Test of a Spare Calorimeter for the HERA Transverse Polarimeter* (unpublished). Hermes Internal Note, 1997.

[37] M.Milite, *The ZEUS muon-VD and Testbeam Results.* Http://www.desy.de/m̃ilite/studentseminar.ps.gz. The transparencies shown at ZEUS student seminar.

[38] L.L. Jones et al., *Fifth Workshop on Electronics for LHC Experiments*, p. 162. (1999), available on `http://www.te.rl.ac.uk/med`.

[39] A. Tapper. Obtained via private communication.

[40] R.J. Cashmore et al., *Proc. the Workshop on Future Physics at HERA*, G. Ingelman, A. De Roeck and R. Klanner (eds.), Vol. 1, p. 163. DESY, Hamburg, Germany (1996).

[41] G.A. Schuler and H. Spiesberger, *Proc. Workshop on Physics at HERA*, W. Buchmüller and G. Ingelman (eds.), Vol. 3, p. 1419. Hamburg, Germany, DESY (1991).

[42] CTEQ Coll., H.L. Lai et al., Eur. Phys. J. **C 12**, 375 (2000).

[43] A. Kwiatkowski, H. Spiesberger and H.-J. Möhring, Comp. Phys. Comm. **69**, 155 (1992). Also in *Proc. Workshop Physics at HERA*, Ed. W. Buchmüller and G.Ingelman, (DESY, Hamburg, 1991).

[44] H. Spiesberger, *An Event Generator for ep Interactions at HERA Including Radiative Processes (Version 4.6)*, 1996, available on `http://www.desy.de/~hspiesb/heracles.html`.

[45] G. Ingelman, A. Edin and J. Rathsman, Comp. Phys. Comm. **101**, 108 (1997).

[46] L. Lönnblad, Comp. Phys. Comm. **71**, 15 (1992).

[47] ZEUS Coll., J. Breitweg et al., Eur. Phys. J. **C 6**, 239 (1999).

[48] J.R. Goncalo, *Simulation of the Hadronic Final State and Systematic Uncertainties in High-Q^2 NC DIS* (unpublished). ZEUS-02-004, 2002.

[49] M. Bengtsson, G. Ingelman and T. Sjöstrand, *Proc. HERA Workshop*, R.D. Peccei (ed.), Vol. 1, pp. 149–165. DESY, Hamburg, Germany (1987).

[50] T. Sjöstrand, Comp. Phys. Comm. **82**, 74 (1994).

[51] Particle Data Group, D.E. Groom et al., Eur. Phys. J. **C 15**, 1 (2000).

[52] A.D. Tapper, *Measurement of charged current deep inelastic scattering cross sections using the ZEUS detector at HERA*. Ph.D. Thesis, University of London, 2001. (unpublished).

[53] G. Marchesini et al., Comp. Phys. Comm. **67**, 465 (1992).

[54] M. Moritz, *Measurement of the High Q^2 Neutral Current DIS Cross Section at HERA*. Thesis, University Hamburg, 2001.

[55] J.R. Goncalo, *Measurement of the High-Q^2 Neutral Current Deep Inelastic Scattering Cross Sections with the ZEUS Detector at HERA*. Ph.D. Thesis, University of London, 2003.

[56] F. Jacquet and A. Blondel, *Proc. the Study of an eP Facility for Europe*, U. Amaldi (ed.), p. 391. (1979). Also in preprint DESY 79-48.

[57] P. de Jong, *Status of the Uranium Calorimeter Reconstruction Software* (unpublished). ZEUS-92-019, 1992.

[58] A. Savin, *Study of Calorimeter Noise in the 1996 Data* (unpublished). ZEUS-98-007.

[59] M. Martinez, *Calorimeter Noisy Cell Study for 1998-2000 ZEUS Data* (unpublished). ZEUS-00-016, internal ZEUS-note, 2000.

[60] G.M. Briskin, *Diffractive Dissociation in ep Deep Inelastic Scattering*. Ph.D. Thesis, Tel Aviv University, 1998. (Unpublished).

[61] N. Tuning and M. Wodarczyk, *rcalcorr.fpp, Calorimeter correction routine, PHANTOM Users Library*, available on `http://http://www-zeus.desy.de/~barbi/analysis/meson/routines/rcalcorr.fpp`.

[62] J. Grosse-Knetter, *Corrections for the Hadronic Final State* (unpublished). ZEUS-98-031, internal ZEUS-note, 1998.

[63] N. Tuning, *Proton Structure Functions at HERA*. Ph.D. Thesis, Amsterdam University, 2001.

[64] J. Grosse-Knetter, *Energy Correction for Islands* (unpublished). ZEUS-97-039.

[65] ZEUS Coll., Calorimeter group, A. Andresen et al., Nucl. Inst. Meth. **A 290**, 95 (1990).

[66] ZEUS Coll., S. Chekanov et al., Preprint DESY-02-064 (hep-ex/0205091), 2002.

[67] A.D. Tapper, *Calibration of the ZEUS Calorimeter for Hadrons and Jets* (unpublished). Talk Given at the Workshop on Energy Calibration of the ATLAS Calorimeter, 21-24 July 2002.

[68] S. Bentvelsen, J. Engelen and P. Kooijman, *Proc. Workshop on Physics at HERA*, W. Buchmüller and G. Ingelman (eds.), Vol. 1, p. 23. Hamburg, Germany, DESY (1992).

[69] K.C. Höger, *Proc. Workshop on Physics at HERA*, W. Buchmüller and G. Ingelman (eds.), Vol. 1, p. 43. Hamburg, Germany, DESY (1992).

[70] G.F. Hartner, *VCTRAK(3.07/04): Offline Output Information* (unpublished). ZEUS-97-064, internal ZEUS-note, 1997.

[71] G. F. Hartner, *VCTRAK Briefing: Program and Math* (unpublished). Zeus-98-058, internal ZEUS-note, 1998.

[72] P. Billoir and S. Qian, Nucl. Inst. Meth. **A 311** (1992).

[73] R. Pawlak, *Vertex reconstruction from FCAL timing* (unpublished). ZEUS-98-040.

[74] A. Caldwell et al., *The Time Calibration of the ZEUS Calorimeter* (unpublished). ZEUS-93-021, 1993.

[75] H. Abramowicz, A. Caldwell and R. Sinkus, Nucl. Inst. Meth. **A 365**, 508 (1995).

[76] A. Kruse and A. van Sighem, *MUFFINCUP- A Wrapper for MUFFIN User Manual and Documentation* (unpublished). ZEUS Note 98-025, 1998.

[77] A. Kruse, *Charged Current interactions at HERA*. Ph.D. Thesis, Amsterdam University, 1999.

[78] A. van Sighem, *Charged Current Deep Inelastic Scattering at HERA*. Ph. D., University of Amsterdam, 2000.

[79] G.J. Feldman and R.D. Cousins, Phys. Rev. **D57**, 3873 (1998).

[80] ZEUS Coll., S. Chekanov et al., Phys. Rev. **D 67**, 012007 (2003).

[81] J. Pumplin et al., Preprint hep-ph/0201195, 2002.

[82] A.D. Martin et al., Eur. Phys. J. **C**, 73 (2002).

[83] ZEUS Coll., S. Chekanov et al., Preprint DESY-03-093 (hep-ex/0307043), 2003. Subm. to Eur. Phys. J. C.

[84] A.M. Cooper-Sarkar, *Advanced Statistical Techniques in Particle Physics*, M.R. Whalley
and L. Lyons (eds.), pp. 197–205. University of Durham (2002), available on `http://www.ippp.dur.ac.uk/Workshops/02/statistics/proceedings.shtml`.

[85] F. James, *Minuit v94.1, CERN Program Library Long Writeup D506* (unpublished), available on `http://wwwinfo.cern.ch/asdoc/minuit/minmain.html`.

[86] C. Pascaud and F. Zomer, Preprint LAL-95-05, June 1995, available on `http://ccdb3fs.kek.jp/cgi-bin/img_index?9508298`.

[87] M. Botje, *QCDNUM version 16.12* (unpublished), available on `http://www.nikhef.nl/~h24/qcdcode/index.html`.

[88] G. Gurci, W. Furmanski and R. Petronzio, Nucl. Phys. **B 175**, 27 (1980).

[89] W. Furmanski and R. Petronzio, Phys. Lett. **B 97**, 437 (1980).

[90] W. Furmanski and R. Petronzio, Z. Phys. **C 11**, 293 (1982).

[91] R.S. Thorne, J. Phys. **G 25**, 1307 (1999).

[92] R.G. Roberts and R.S. Thorne, Phys. Lett. **B 421**, 303 (1998).

[93] R.G. Roberts and R.S. Thorne, Phys. Rev. **D 57**, 6871 (1998).

[94] R.G. Roberts and R.S. Thorne, Eur. Phys. J. **C 19**, 339 (2001).

[95] A.M. Cooper-Sarkar, *NLOQCD fits to ZEUS data* (unpublished). Internal ZEUS note, ZEUS-00-040, Nov. 2000.

[96] CCFR Coll., A.O. Bazarko et al., Z. Phys. **C 65**, 189 (1995).

[97] ZEUS Coll., S. Chekanov et al., Eur. Phys. J. **C 21**, 443 (2001).

[98] BCDMS Coll., A.C. Benvenuti et al., Phys. Lett. **B 223**, 485 (1989).

[99] E665 Coll., M.R. Adams et al., Phys. Rev. **D 54**, 3006 (1996).

[100] NMC Coll., M. Arneodo et al., Nucl. Phys. **B 483**, 3 (1997).

[101] NMC Coll., M. Arneodo et al., Nucl. Phys. **B 487**, 3 (1997).

[102] W.G. Seligman et al., Phys. Rev. Lett. **79** (1997).

[103] A.D. Martin et al., Eur. Phys. J. **C 4**, 463 (1998).

[104] J. Gomez et al., Phys. Rev. **D 49**, 4348 (1994).

[105] H. Georgi and H.D. Politzer, Phys. Rev. **D 14**, 1829 (1976).

[106] M. Virchaux and A. Milsztajn, Phys. Lett. **B 274**, 221 (1992).

[107] A.D. Martin, R.G. Roberts and W.J. Stirling, Phys. Lett. **B 387**, 419 (1996).

[108] A.D. Martin et al., Eur. Phys. J. **C 14**, 133 (2000).

[109] A.D. Martin et al., Nucl. Phys. Proc. Suppl. **79**, 105 (1999).

[110] H1 Coll.., Preprint hep-ex/0304003, 2003. Subm. to Eur. Phys. J. C.

[111] ZEUS Coll., Preprint DESY-02-113, DESY, 2002.

[112] ZEUS Coll., J. Breitweg et al., Eur. Phys. J. **C 12**, 411 (2000).

[113] ZEUS Coll., Eur. Phys. J. **C 27**, 305 (2003).

[114] ZEUS Coll., Preprint DESY-03-214, DESY, 2003. Subm. to Phys. Rev. D.

[115] ZEUS Coll., J. Breitweg et al., Eur. Phys. J. **C 7**, 609 (1999).

[116] ZEUS Coll., J. Breitweg et al., Phys. Lett. **B 487**, 53 (2000).

[117] H1 Coll., C. Adloff et al., Eur. Phys. J. **C 21**, 33 (2001).

[118] H1 Coll., C. Adloff et al., Eur. Phys. J. **C 13**, 609 (2000).

[119] H1 Coll., C. Adloff et al., Eur. Phys. J. **C 19**, 269 (2001).

[120] A.D. Martin et al., Preprint hep-ph/0308087, 2003.

[121] M. Botje, Preprint hep-ph/0110123, 2001.

Acknowledgements

I would like to thank my supervisor, Ken Long for his advice and help over the course of my research. Similarly, I would also like to thank Amanda Cooper-Sarkar for her input into the PDF fitting, Alex Tapper for his assistance in the Charged Current analysis and Fabio Metlica for help in the polarimetry section. I would like to thank my girlfriend, Susan Wallace, her parents and also my sister, Nikola for their support and generosity. I would especially like to thank my sister for being so understanding on the occasions I have had to lodge with her, ever good mannered and always willing to inflate the spare airbed. I would like to thank the DESY bistro for "culinary inspiration" and Fr. Susan Ketels for making my stay in Hamburg so pleasant. I am also grateful to my Father and Christine, without whose help this tome may never have been bound! I acknowledge the financial support of the Particle Physics and Astronomy Research Council.