HERAPDF 1.0 to 1.5

A M Cooper-Sarkar
Les Houches Feb 2011

Combination of ZEUS and H1 data and PDF fits to these data:

1. Inclusive cross-sections HERA-1 (1992-2000):arxiv:0911.0884 -improved constraints at low-x
2. F2(charm) data (preliminary)- constraints on the charm mass parameter, m_c(model)
3. Low energy runs – FL- 2007- (preliminary) –tension with low x, Q^2 data?
4. Inclusive cross-sections HERA-II (2003-2007)- (preliminary) -improved constraints at high-x
5. PDFs at NLO and NNLO,

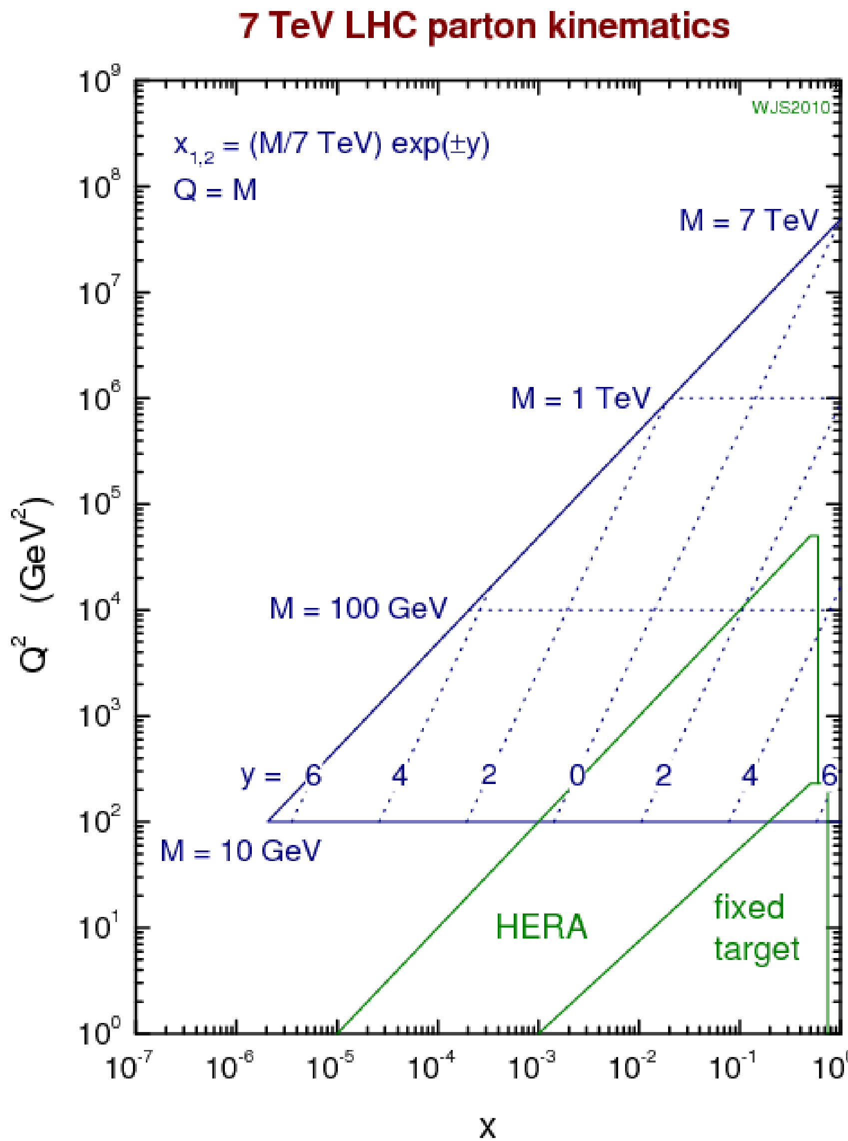
PDFs with different α_s =0.114-0.122, m_c =1.35-1.65, m_b = 4.3-5.0

Coming soon: fits with HERA jets data.

Why combine ZEUS and H1 data?

At the LHC we collide protons Protons are full of partons. Our knowledge of partons comes from Deep Inelastic Scattering data. HERA dominates these data and is most relevant for the kinematic region of early LHC data

We think we know how to extrapolate in Q^2 using (N)NLO QCD (using the DGLAP equations) but we don't a priori know the shapes of the parton distributions in x. The HERA data is our best guide



$$\frac{dq(x,Q^2)}{d\ln Q^2} = \frac{\alpha_s(Q^2)}{2\pi}\int_x^1 \frac{dy}{y}\left[Pqq(z)q(y,Q^2) + Pqg(z)g(y,Q^2)\right]$$

$$\frac{dg(x,Q^2)}{d\ln Q^2} = \frac{\alpha_s(Q^2)}{2\pi}\int_x^1 \frac{dy}{y}\left[\Sigma_q Pgq(z)q(y,Q^2) + Pgg(z)g(y,Q^2)\right]$$

DGLAP eqns

HERAPDF motivation

A substantial part of the uncertainty on parton distributions comes from the need to use many different input data sets with large systematic errors and questionable levels of consistency

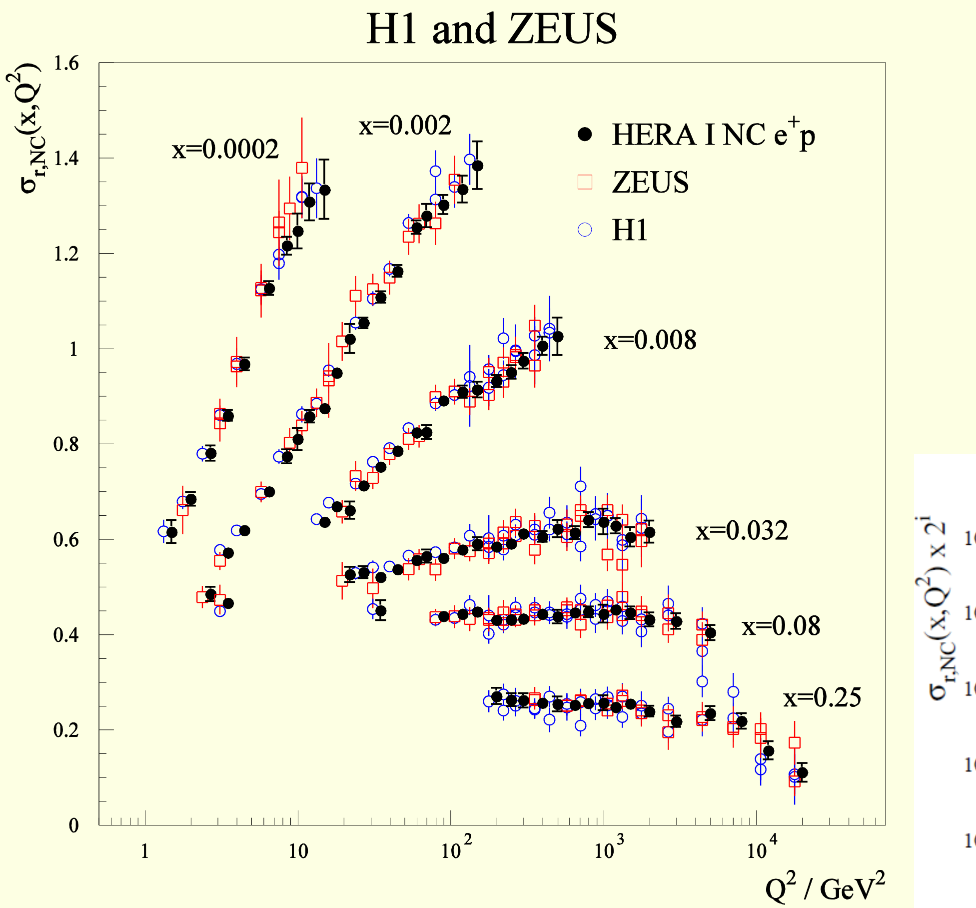
- Combining H1 and ZEUS data has provided a tool to study the consistency of the data and to reduce systematic uncertainties:

⇒ Experiments cross calibrate each other JHEP 1001.109 arxiv:0911.0884

- The combination method includes accounting for full systematic error correlations.

- The resulting combination is much more accurate than expected from the increased statistics of combining two experiments- it's like having a detector which combines the best features of each

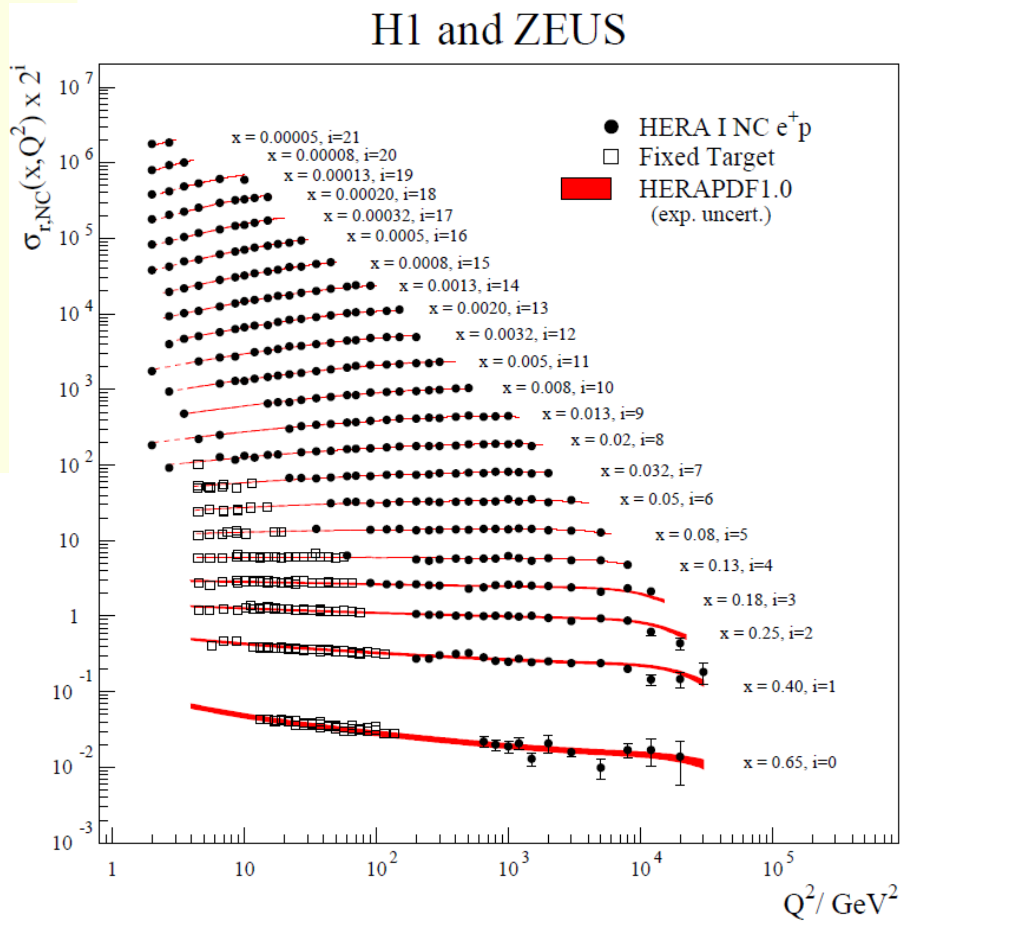
- The post-averaging systematic errors are smaller than the statistical across a large part of the kinematic plane and total errors as small as ~1.5%



This page shows NC e+ combined data

HERA data reach low- x values even for Q^2 ~ 100 GeV^2

Results of the combination compared to the separate data sets



These data are used for extracting parton distributions: HERAPDF1.0

Some of the debates about the best way of estimating PDF uncertainties concern the use of many different data sets with varying levels of consistency.

The combination of the HERA data yields a very accurate and consistent data set for 4 different processes: e+p and e-p Neutral and Charged Current reactions.

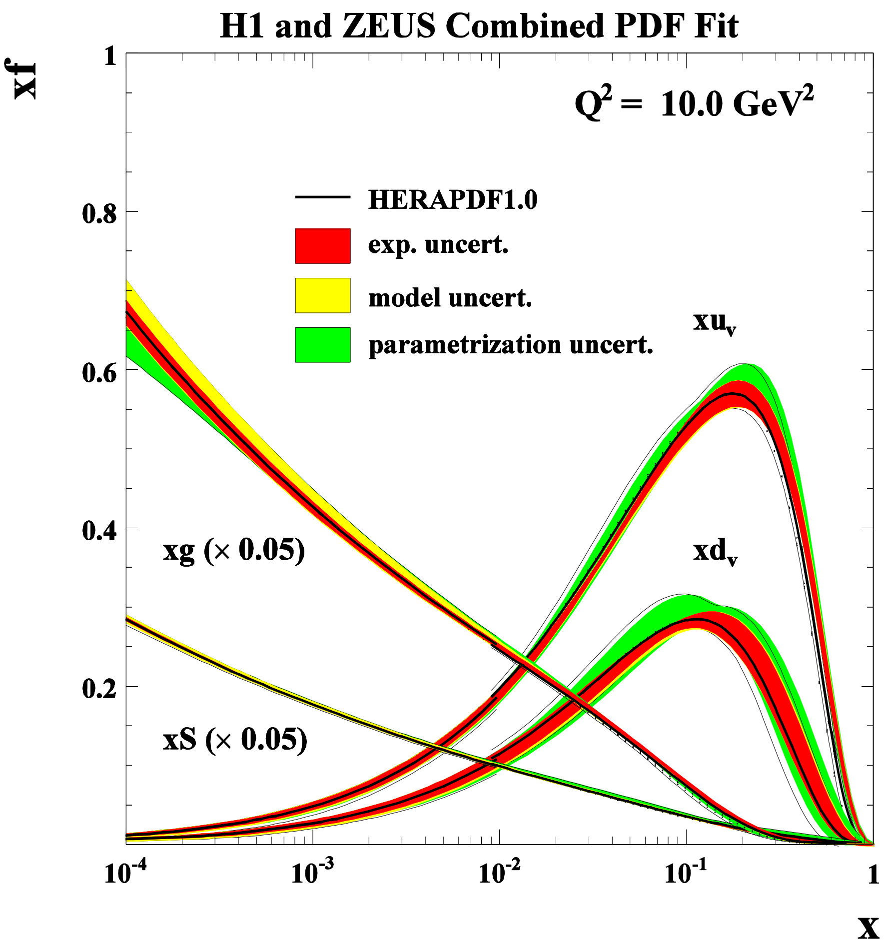
Whereas the data set does not give information on every possible PDF flavour it does:

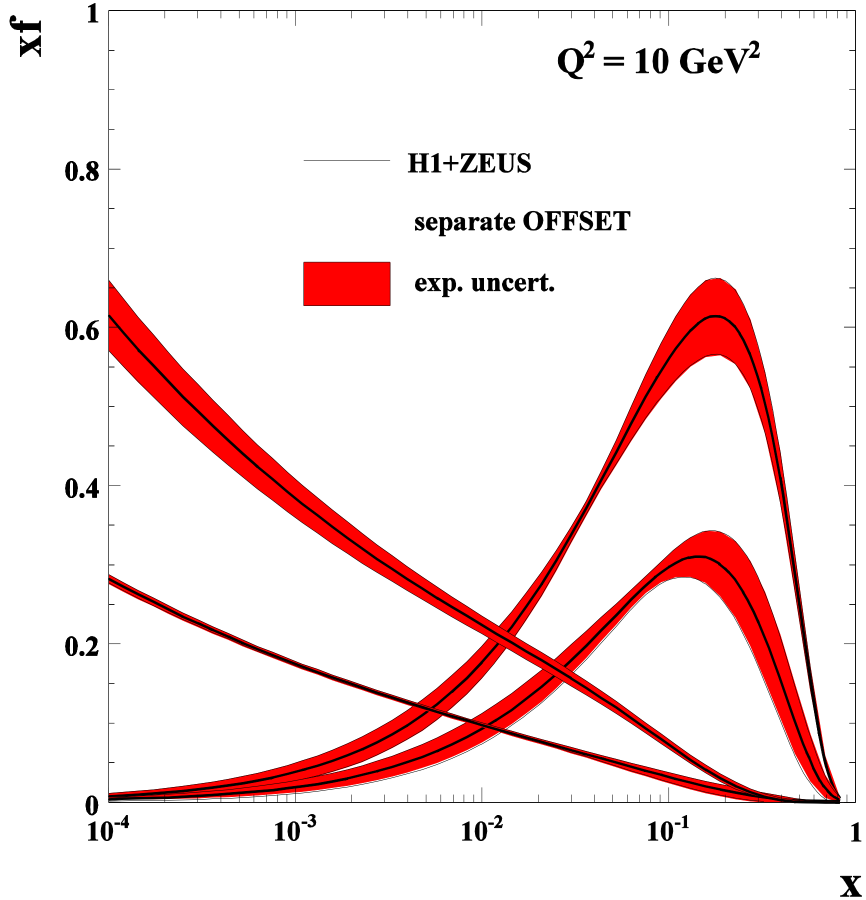
- Give information on the low-x Sea (NCe+ data)
- Give information on the low-x Gluon via scaling violations (NCe+ data)
- Give information on high-x u (NCe+/e- and CCe-) and d (CCe+ data) valence PDFs
- Give information on u and d-valence shapes down to $x \sim 3\ 10^{-2}$ (from the difference between NCe+ and NCe-)

NOTE the use of a pure proton target means d-valence is extracted without need for heavy target/deuterium corrections (1101.5148) or strong iso-spin assumptions these are the only PDFs for which this is true– also do not depend on assumptions on FL used to extract the fixed target F2 values that are usually fitted (1101.5261)

Furthermore, the kinematic coverage at low-x ensures that these are the most crucial data when extrapolating predictions from W, Z and Higgs cross-sections to the LHC

RESULTS for HERAPDF1.0 –arxiv:0911.0884



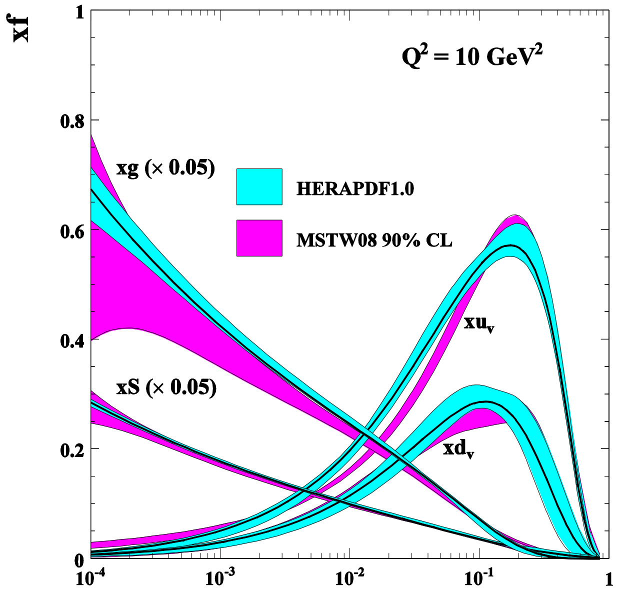


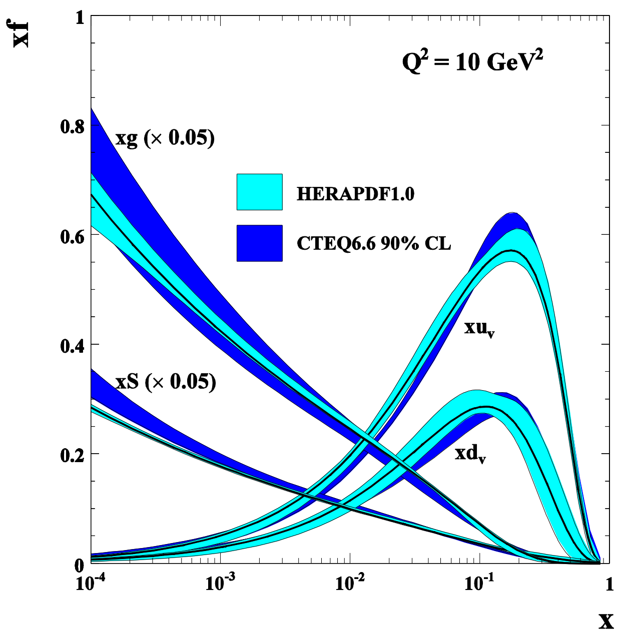
And here is a summary plot of the PDF results

Experimental uncertainties on PDFs are extracted with $\Delta\chi2=1$, and model and parametrization uncertainties are also evaluated.

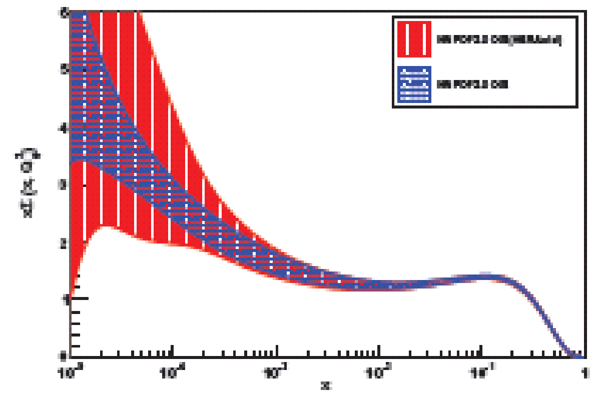
To appreciate how much better this is than uncombined HERA data compare the red experimental errors to this plot which shows the experimental errors for a smilar PDf fit to uncombined data

Compare to other PDF analyses

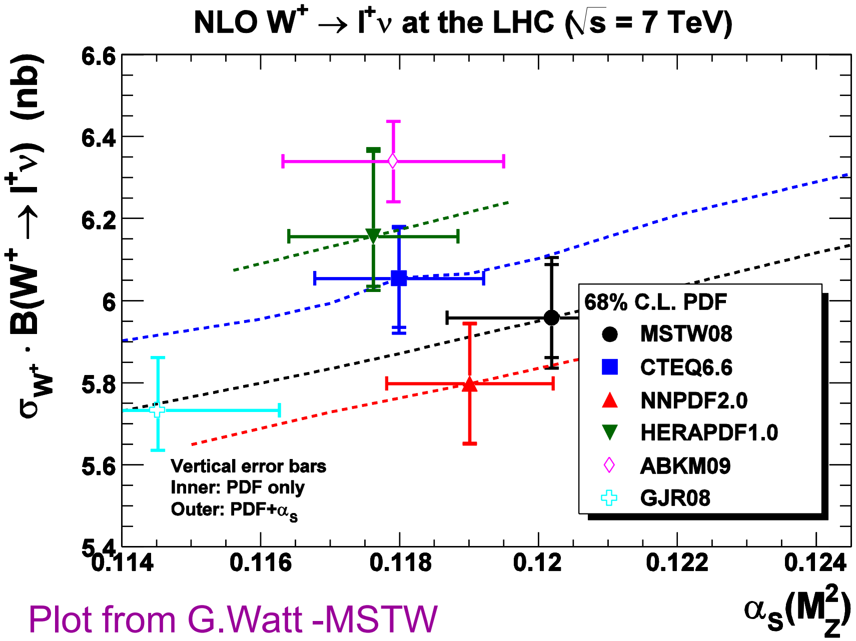




Effect of using HERA combined data on other PDf analyses



The NNPDF global PDF fitting group have incorporated the combined HERA data into their fit and here is the improvement to the Sea PDF- with uncombined HERA data you get the red- with combined you get the blue



Plot from G.Watt -MSTW

Comparisons of W+ cross-section as a function of $\alpha_S(M_Z)$

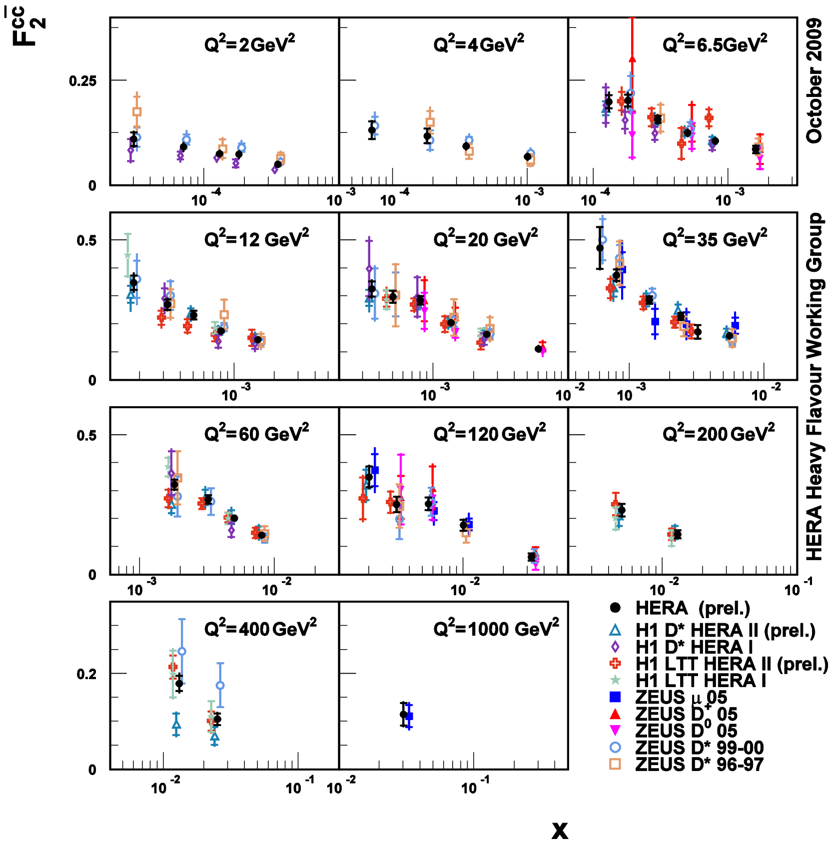
MSTW08
CTEQ66
HERAPDF1.0
NNPDF2.0
ABKM09
GJR08

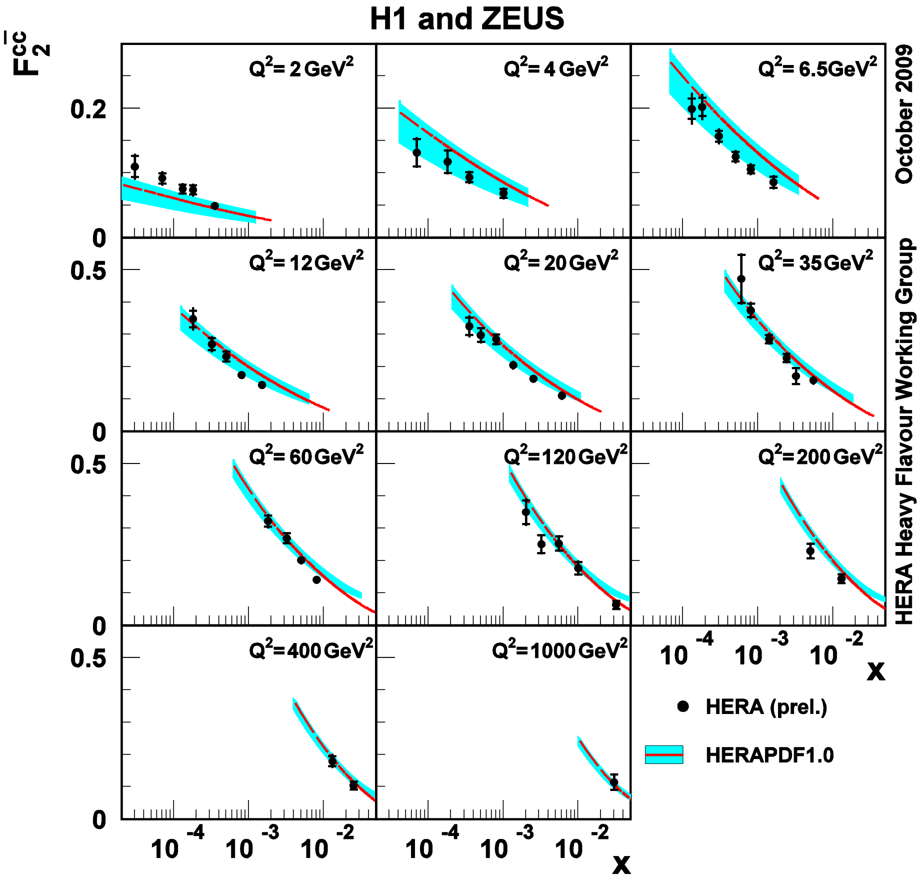
The PDF4LHC group has been considering all these PDFs at NLO

Recently the PDF4LHC group has been comparing predictions from modern PDFs. This plot shows the role that the uncertainty in the value of $\alpha_S(M_Z)$ plays in the overall uncertainty of predictions

This is not a large effect for W/Z production

But the value of m_c AND the scheme used to account for heavy quark production are..





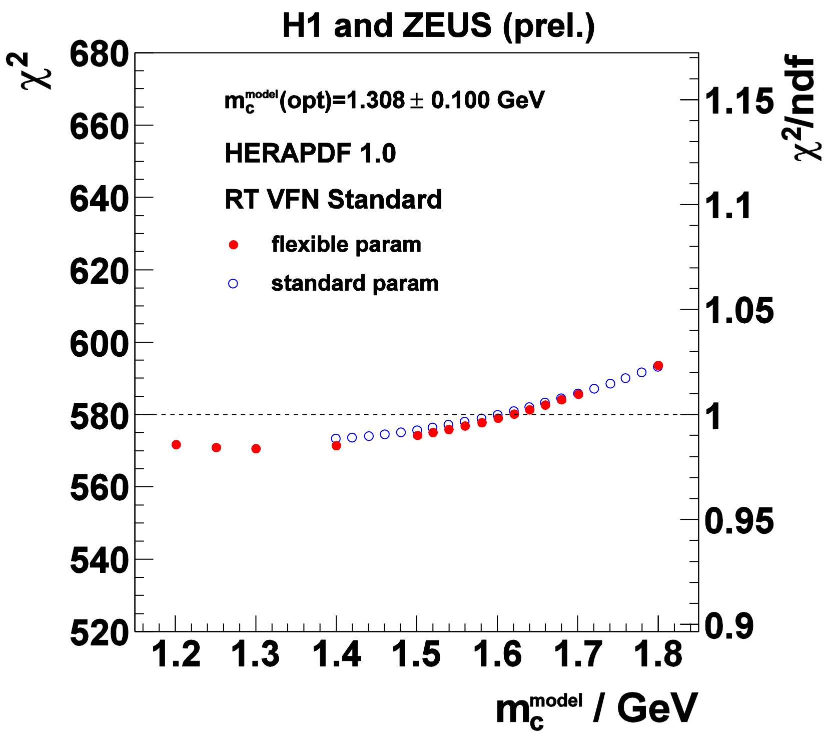
H1 and ZEUS have also combined charm data recently

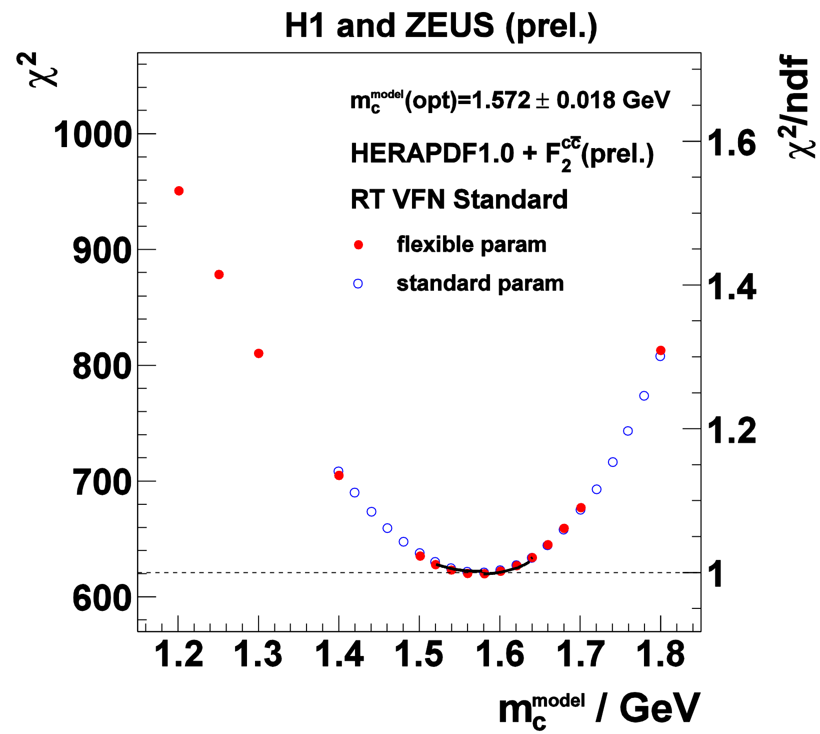
And the HERAPDF1.0 gives a good description of these data –within its error band-

The error band spans m_c=1.35 (high) to m_c=1.65 (low) GeV

The data show some preference for higher charm mass than the standard choice m_c=1.4 GeV

If we input the charm data to the PDF fit it does not change the PDFs significantly BUT

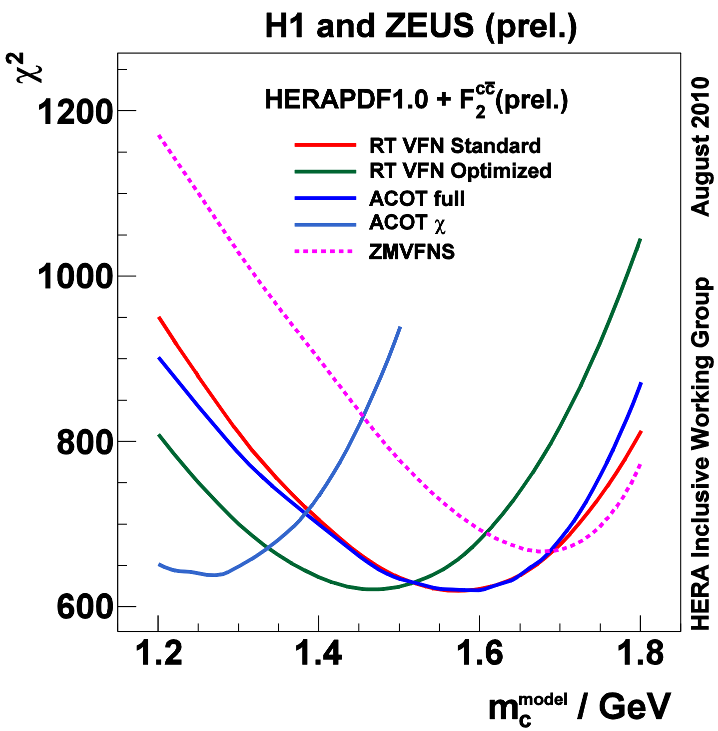




Before charm is input the χ2 profile vs the charm mass parameter is shallow..

After charm is input the χ2 profile vs the charm mass parameter gives

$m_c = 1.57 \pm 0.02$ GeV



But the HERAPDF uses the Thorne General Mass Variable Flavour Number Scheme for heavy quarks as used by MSTW08

This is not the only GMVFN

CTEQ use ACOT- χ

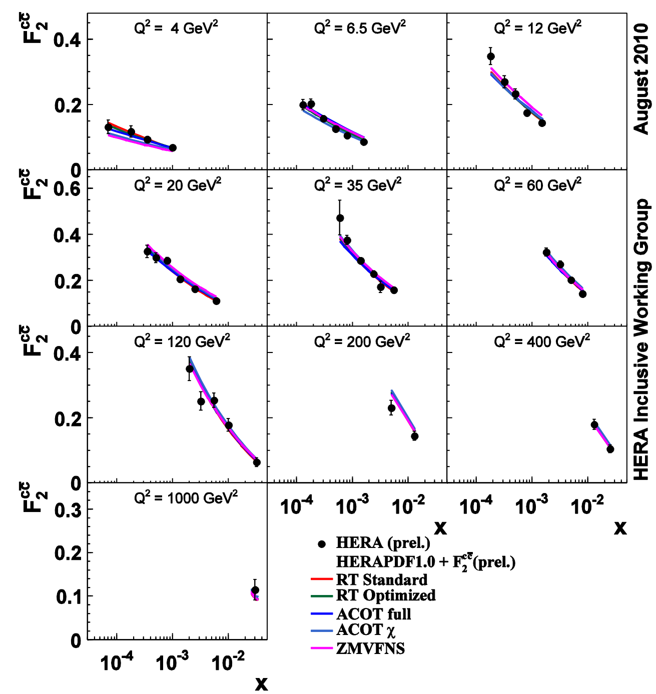
NNPDF2.0 use ZMVFN

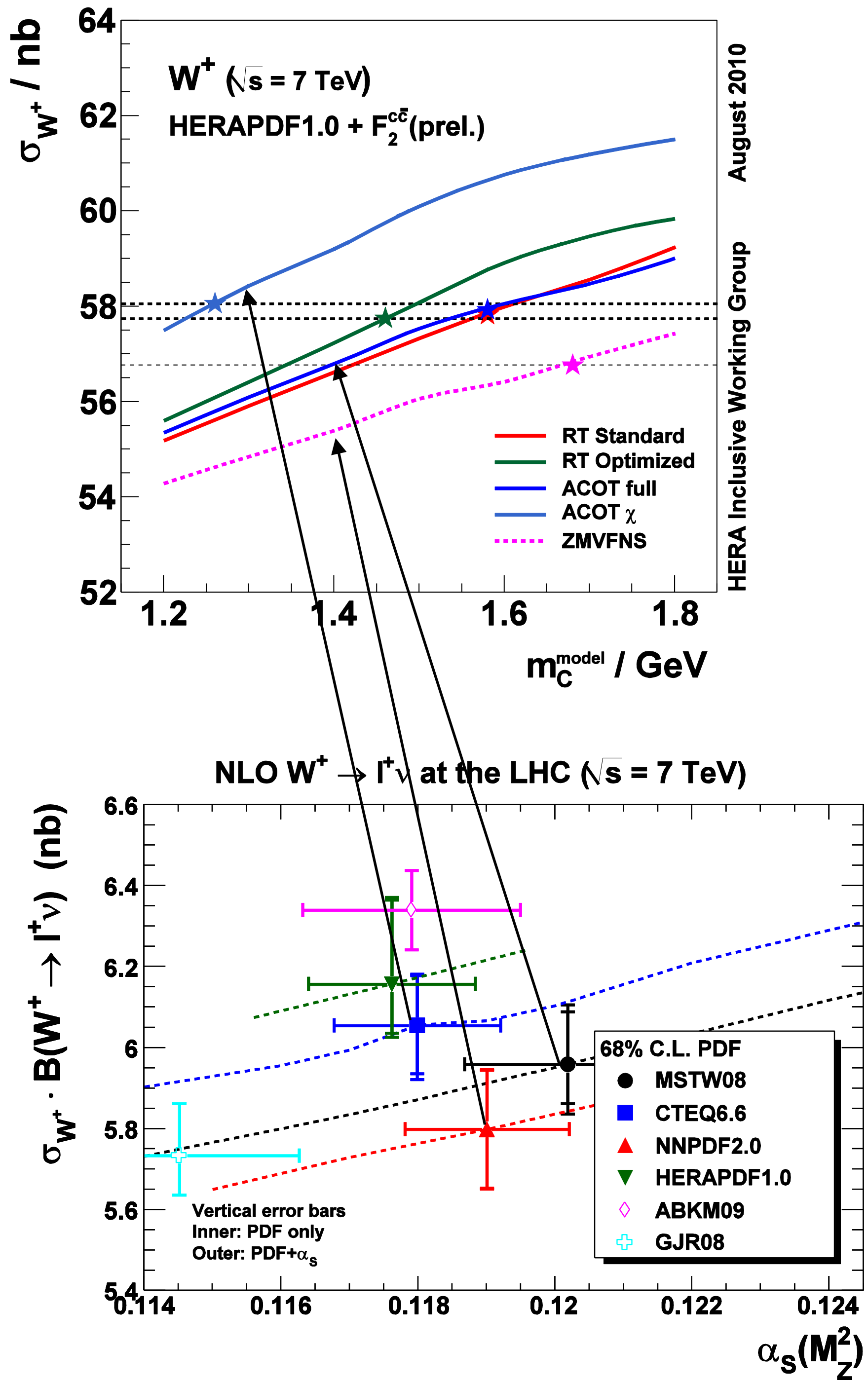
These all have different preferred charm mass parameters, and all fit the data well when used with their own best fit charm mass

We have re-analysed the HERAPDF+F2c data using several different heavy quark schemes

	m_c^{thr} (opt) / GeV	stat	syst	
RT stand	1.57	±0.02	+0.01	-0.03
RT optim	1.47	±0.02	+0.01	-0.03
ACOT full	1.58	±0.02	+0.02	-0.04
ACOT χ	1.25	±0.02	+0.02	-0.04
ZMVFNS	1.67	±0.02	+0.06	-0.06

Model and param. Errors included





We then use each of these schemes to predict W and Z cross-sections at the LHC (at 7 TeV) as a function of charm mass parameter

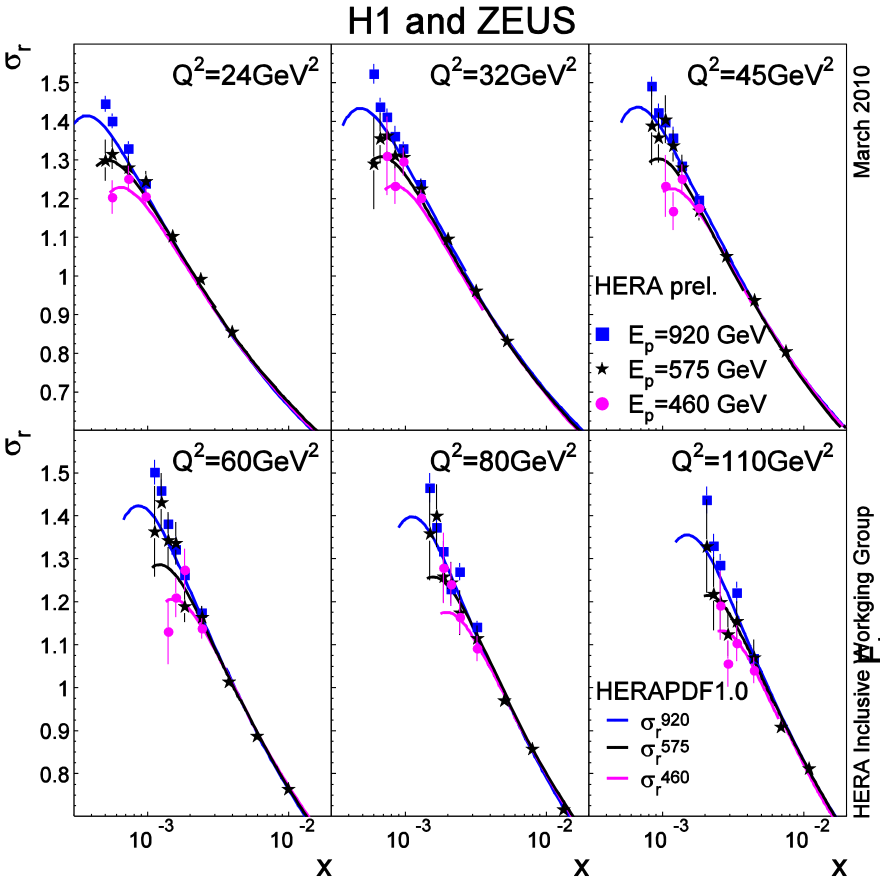
If a fixed value of mc is used then the spread is considerable (~7%)- but if each prediction is taken at its own optimal mass value the spread is dramatically reduced (~2%) even when a Zero-Mass (ZMVFN) approximation has been used

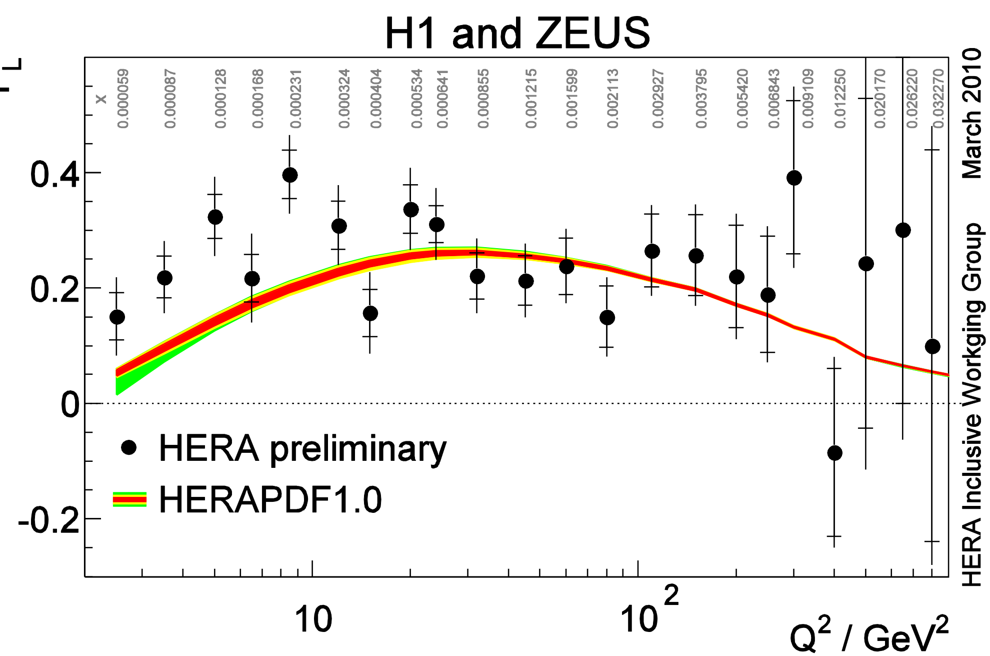
scheme	m_c^{model}(opt)	χ^2/dof	χ^2/ndp	σ_Z(nb)	σ_{W^+}(nb)	σ_{W^-}(nb)
RT Standard	$1.58^{+0.02}_{-0.03}$	620.3/621	42.0/41	$29.27^{+0.07}_{-0.11}$	$57.82^{+0.14}_{-0.22}$	$40.22^{+0.10}_{-0.15}$
RT Optimized	$1.46^{+0.02}_{-0.04}$	621.6/621	46.5/41	$29.17^{+0.07}_{-0.13}$	$57.75^{+0.14}_{-0.26}$	$40.15^{+0.10}_{-0.18}$
ACOT full	$1.58^{+0.03}_{-0.04}$	621.2/621	59.9/41	$29.28^{+0.10}_{-0.13}$	$57.93^{+0.18}_{-0.24}$	$40.16^{+0.12}_{-0.16}$
S-ACOT-χ	$1.26^{+0.02}_{-0.04}$	639.7/621	68.5/41	$29.37^{+0.08}_{-0.15}$	$58.06^{+0.16}_{-0.30}$	$40.23^{+0.11}_{-0.21}$
ZMVFNS	$1.68^{+0.06}_{-0.07}$	667.4/621	88.1/41	$28.71^{+0.19}_{-0.20}$	$56.77^{+0.33}_{-0.34}$	$39.46^{+0.24}_{-0.25}$
			differences	0.7% 2.3%	0.5% 2.3%	0.2% 2.0%

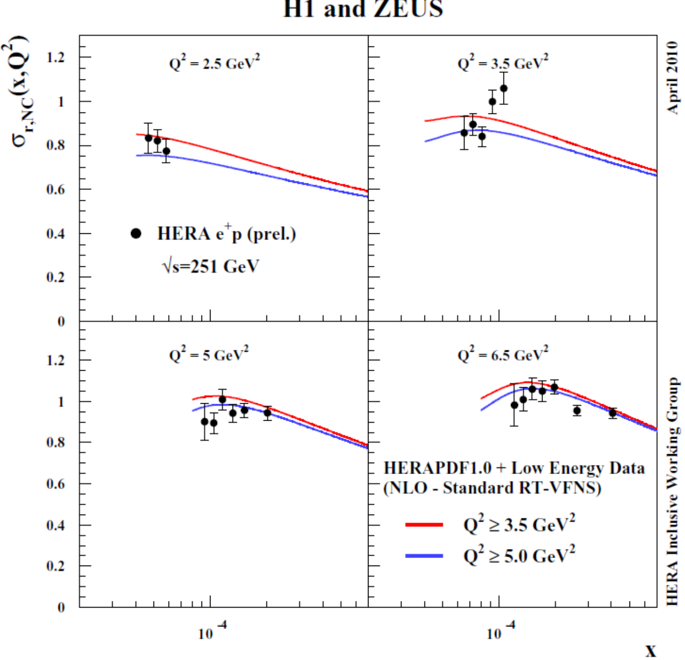
The PDFs MSTW08, CTEQ6.6, NNPDF2.0 do NOT use charm mass parameters at the optimal values- and this partly explains their differing predictions.

Note NNPDF2.1 HAS now moved upwards –heavy quarks scheme now used

H1 and ZEUS have also combined the e+p NC inclusive data from the lower proton beam energy runs (P_P = 460 and 575) and produced a common FL measurement

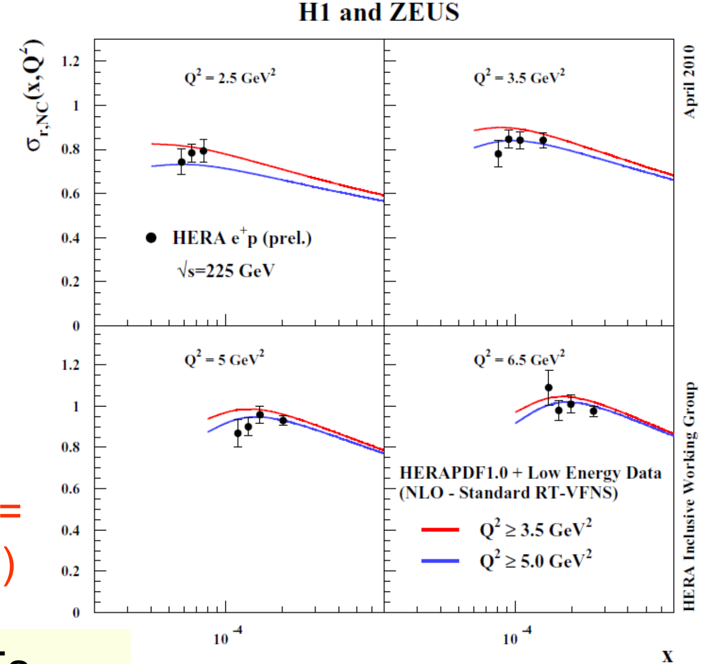






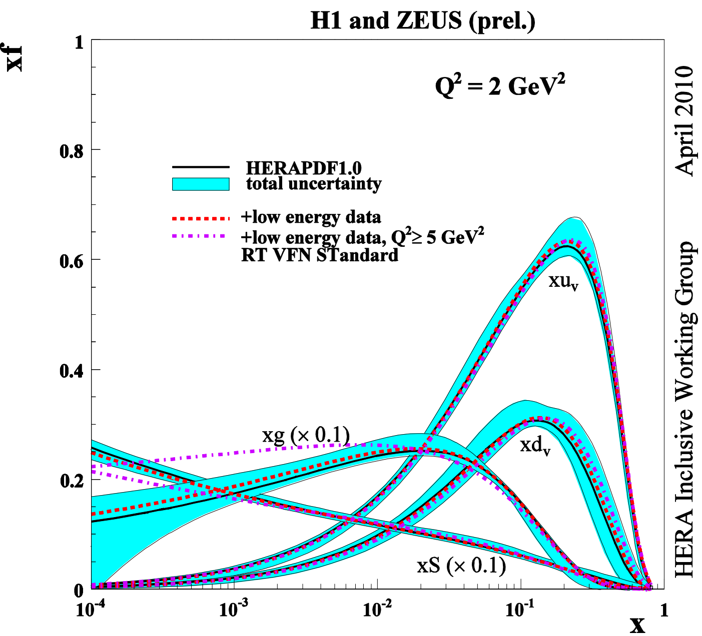
When the low energy data are input to the HERAPDF fit it becomes evident that the low Q^2/low-x data are not so well fit –

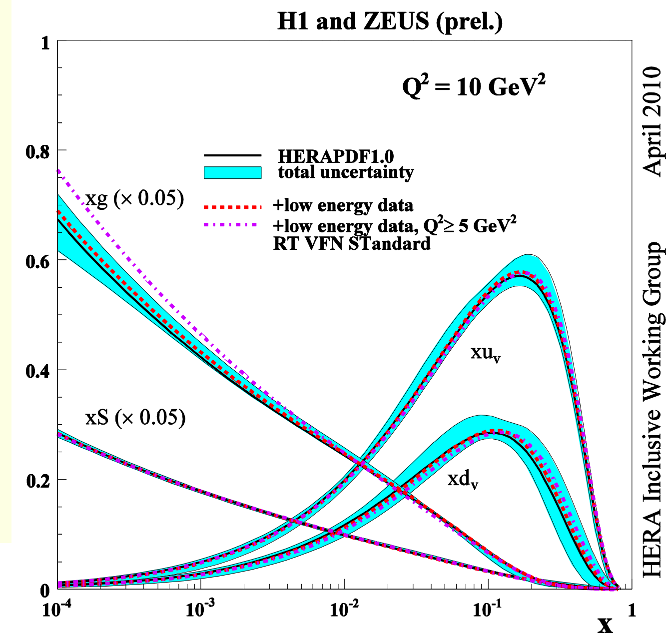
Imposing a harder Q^2 cut $Q^2 > 5$ improves the situation (X^2 /ndp = 1.1 decreases to 0.95)

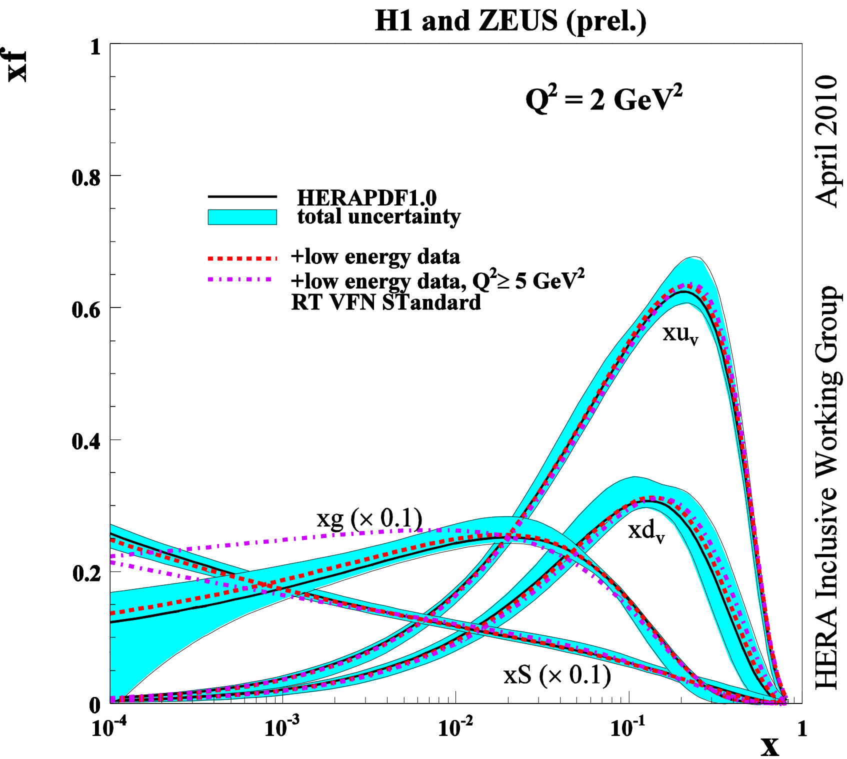


The resulting PDFs have a somewhat different shape- less valence-like gluon at low Q^2… steeper gluon at higher Q^2

This is also true if you make an x cut $x > 5\ 10^{-4}$ or a combined cut $Q^2 > 0.5\ x^{-0.3}$







Our Regge prejudices led us to think that the sea and gluon would have soft slopes at low x $\sim x^{-0.08}$ at the starting scale and THEN evolution would make them steeper. However at $Q^2 \sim 2$

the sea has a steeper slope $x^{-0.15}$

and the gluon is valence-like $x^{+0.2}$

If however we distrust the formalism for low x and Q^2 and we fit only data for $Q^2 > 5$

the sea has a softer slope $x^{-0.11}$

But the gluon is less valence-like $x^{+0.08}$

i.e. they are both closer to the Regge soft Pomeron value of -0.08

This implies that the 'true' gluon could be a little bit steeper than the HERAPDF1.0 gluon- or indeed CTEQ, NNPDF, MSTW gluons

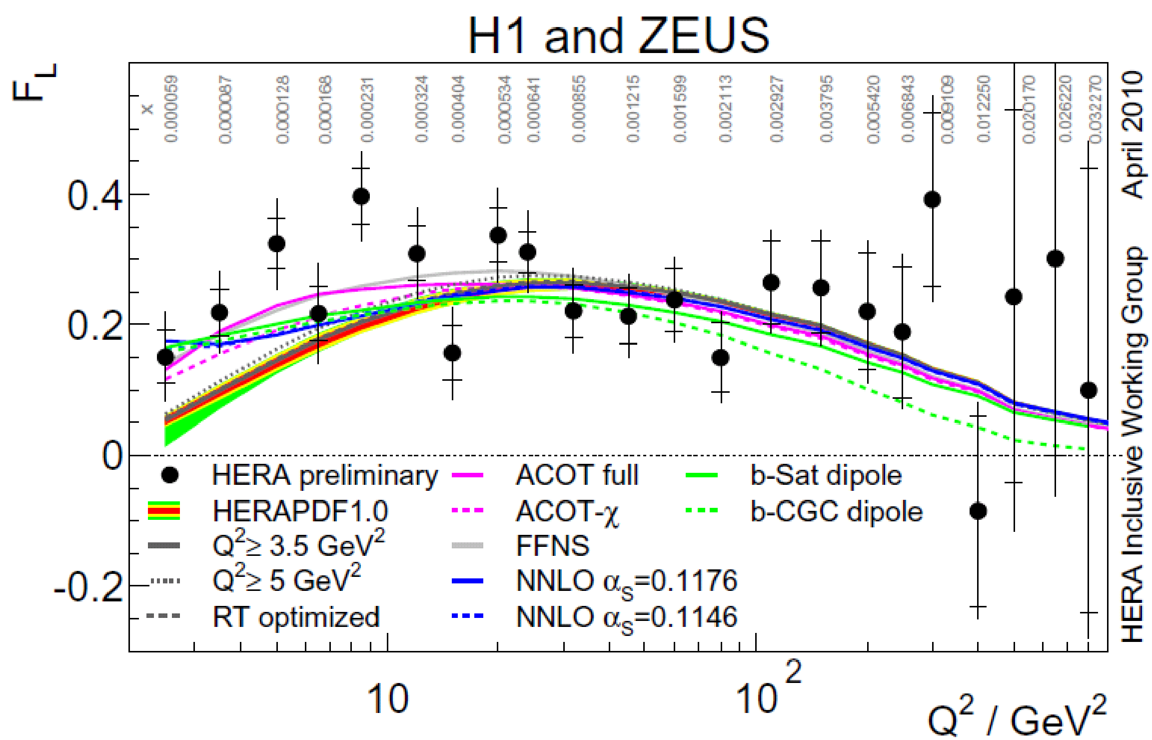
BUT NOTE there is no improvement from cutting high y. These x,Q^2 cuts do NOT have a big effect on the description of FL.

Changes of heavy quark scheme to ACOT, FFN

or a change from NLO to NNLO have a bigger effect on FL

Whereas NNLO does not improve the description of the low-x,Q2 cross-section data.. at least in TR scheme..

ACOT and FFN do improve the description- but it is improved further by x,Q2 cuts



HERAPDF1.0 is also available at NNLO for two values of $\alpha_S(M_Z)$:

0.1176 (standard ~PDG value) and 0.1145 (preferred by the data at NNLO)

NNLO is important for precision studies of cross-section uncertainties.

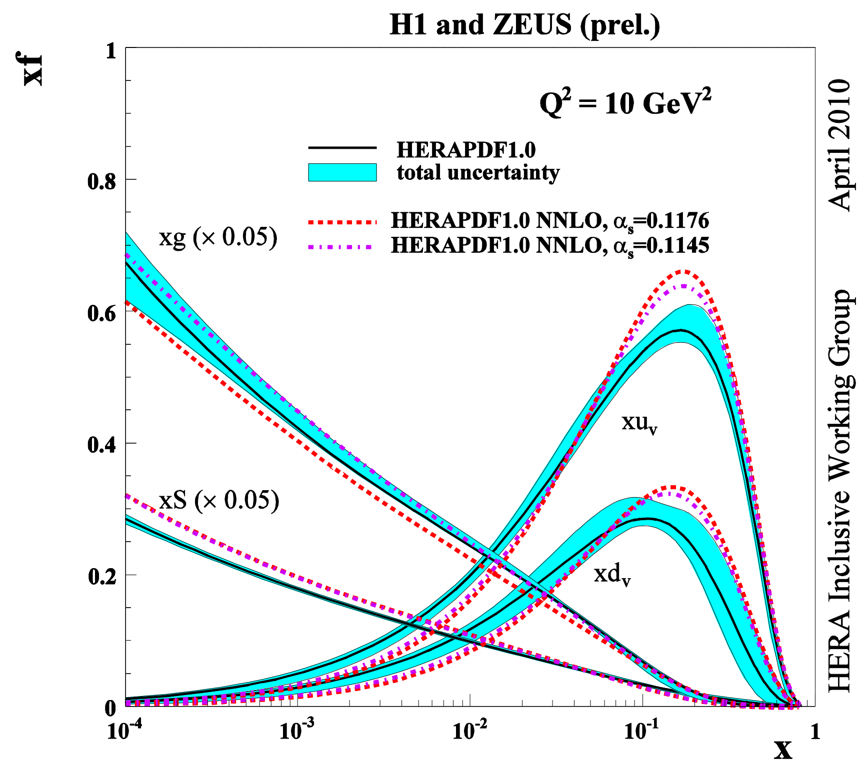
There are far fewer NNLO PDFS: MSTW08, ABKM

(Recent papers from 1101.1832 1101.5261

on Higgs production)

NOTE: NNLO has worse χ2 than NLO and does not fit low-x Q2 data better. The χ2 is also improved if low x, Q2 cuts are imposed.

In fact it is the 920 data which are worst fit at NNLO. Tension between the low and high-energy data shows up at low-x,Q2 and is not solved by moving to NNLO.



HERA-I	1992-2000	Ep=820,920 GeV
HERA-II	2003-2007	Ep=920, 460,575 GeV

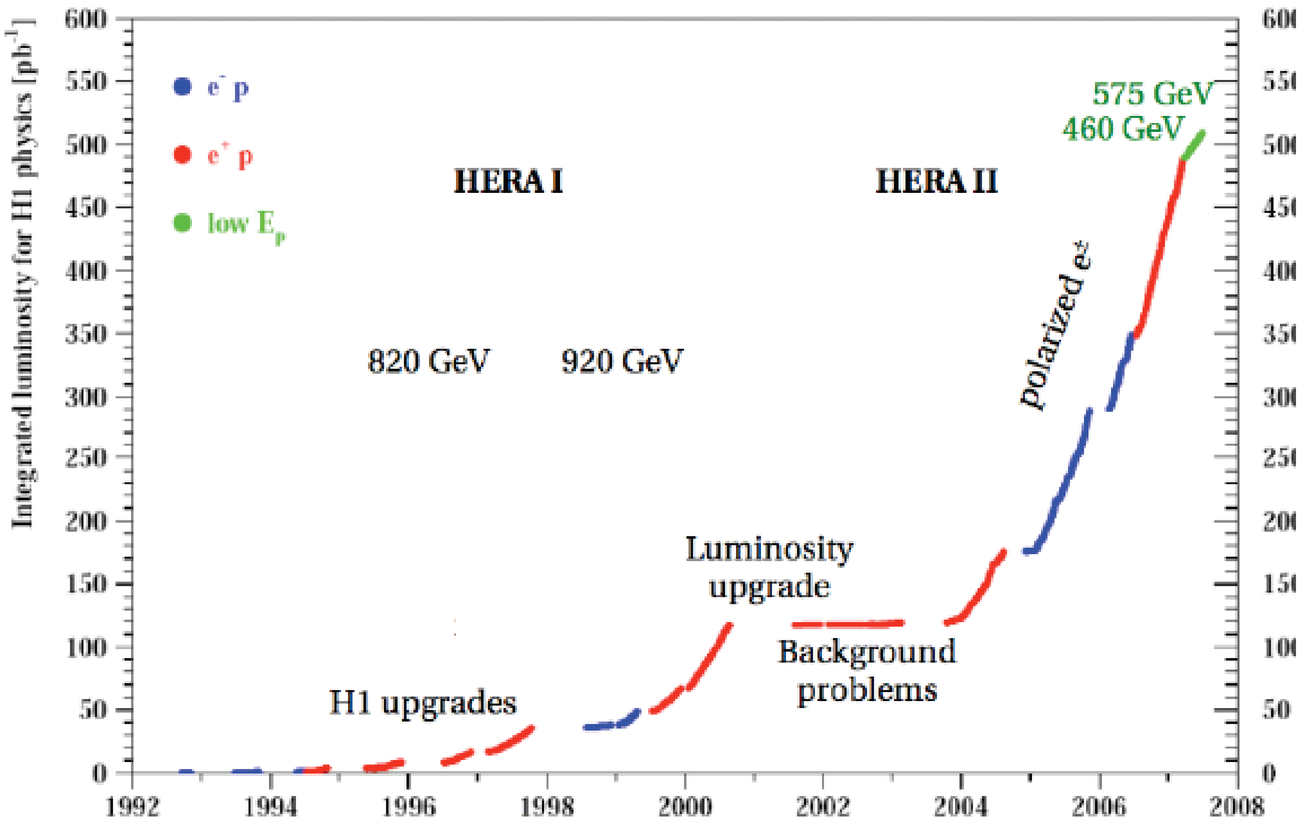
Registered ~1fb^{-1} of integrated luminosity of physics data.

HERA- I combination only ~250 pb^{-1} of data

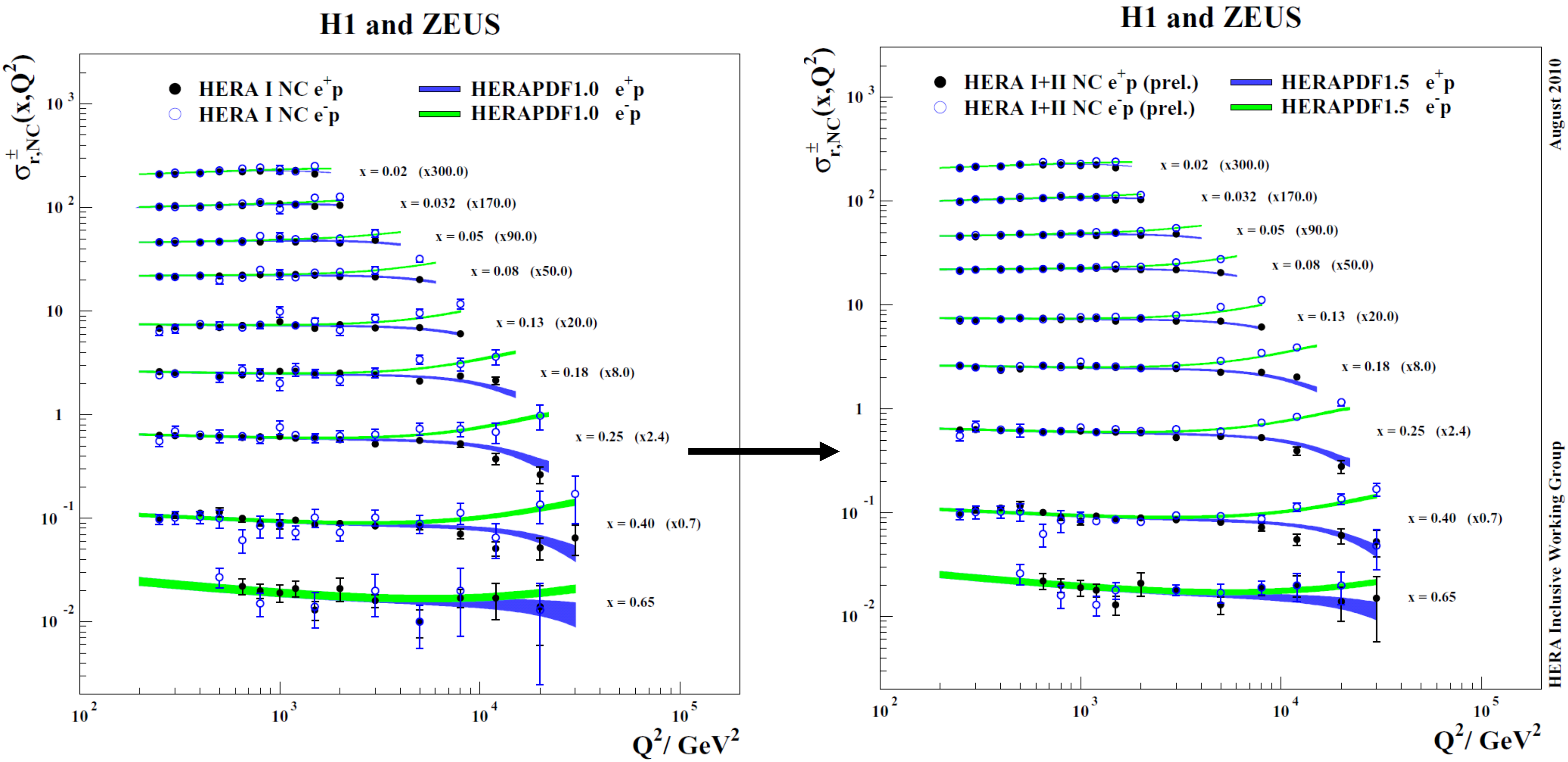
HERA-II gives 4 times as much data in total

The triggers were such that most of this is at higher x and Q^2

We have made a preliminary HERA-II combination- not all of the separate ZEUS and H1 inclusive data which go into this combination are yet published.

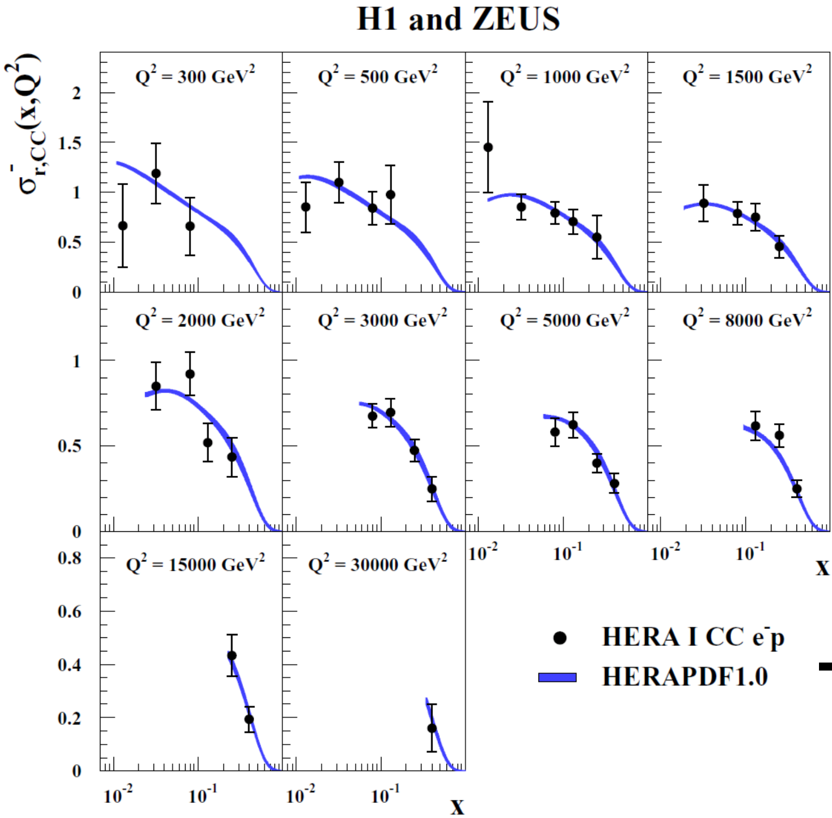


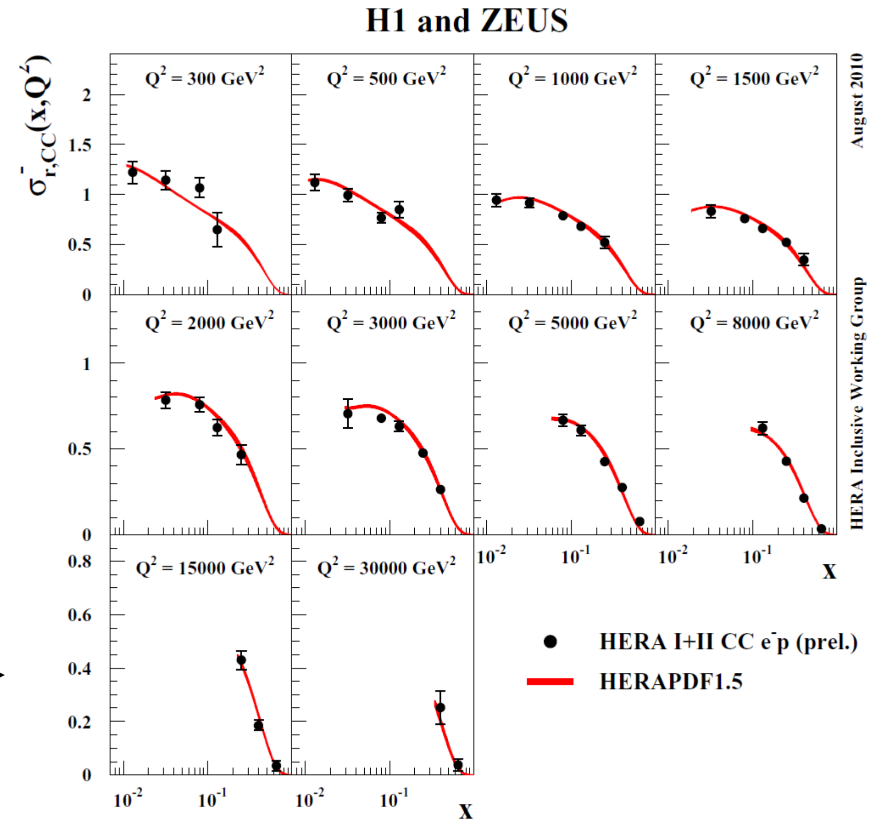
H1 and ZEUS have also combined preliminary high Q^2 HERA-II data along with the HERA-I data and HERAPDF1.0 has recently been updated to HERAPDF1.5 by including these data



The data on the left has been updated to the data on the right

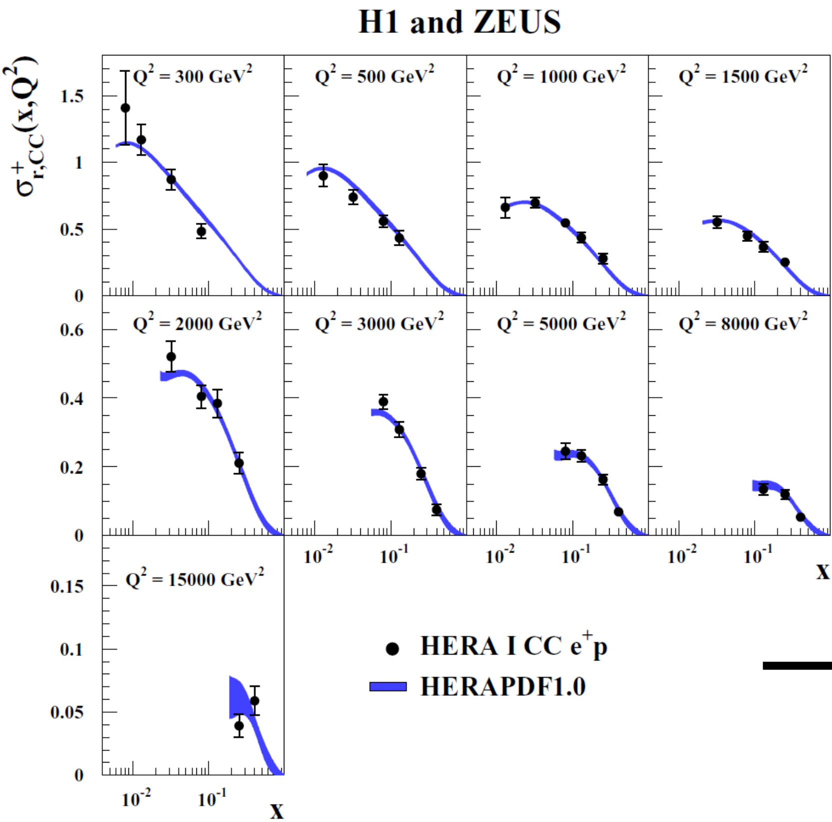
The HERAPDF1.0 fit on the left has been updated to the HERAPDF1.5 fit on the right

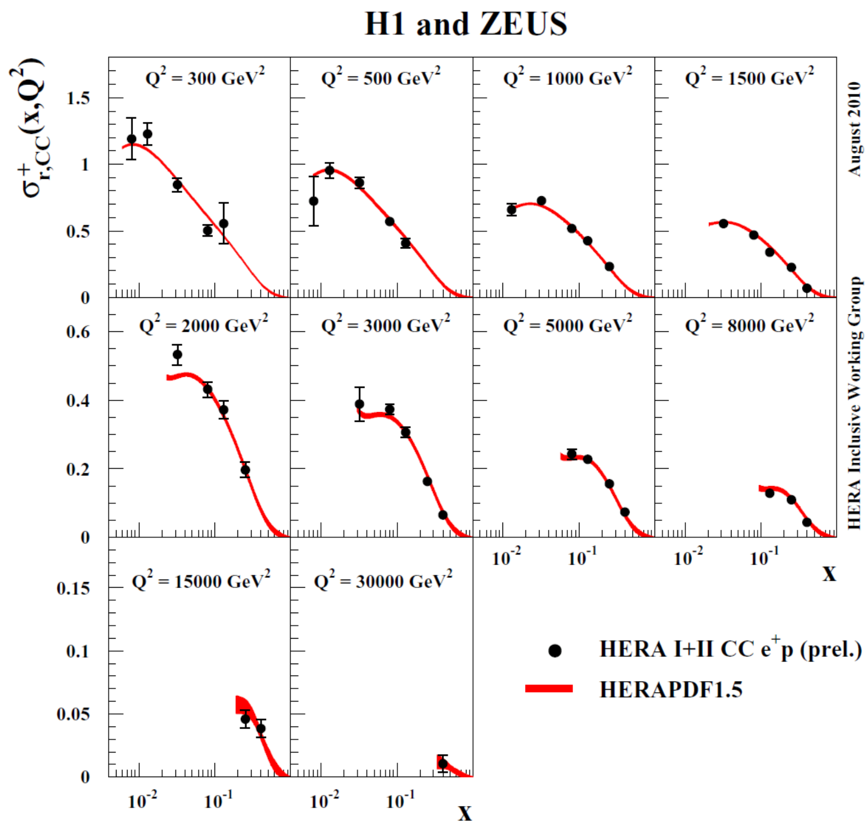




The data on the left has been updated to the data on the right

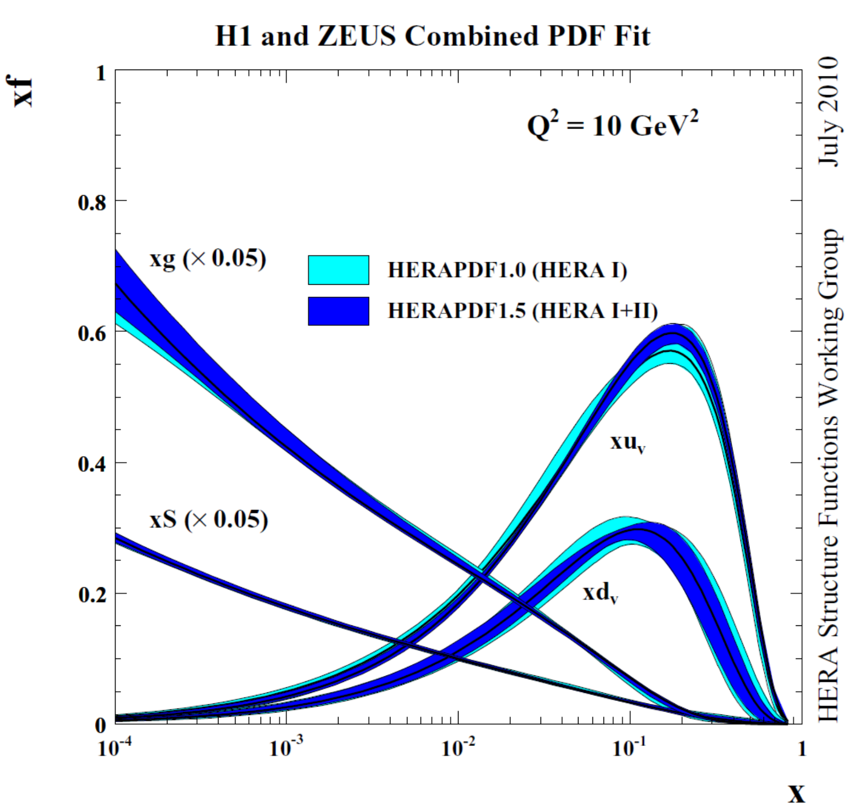
The HERAPDF1.0 fit on the left has been updated to the HERAPDF1.5 fit on the right

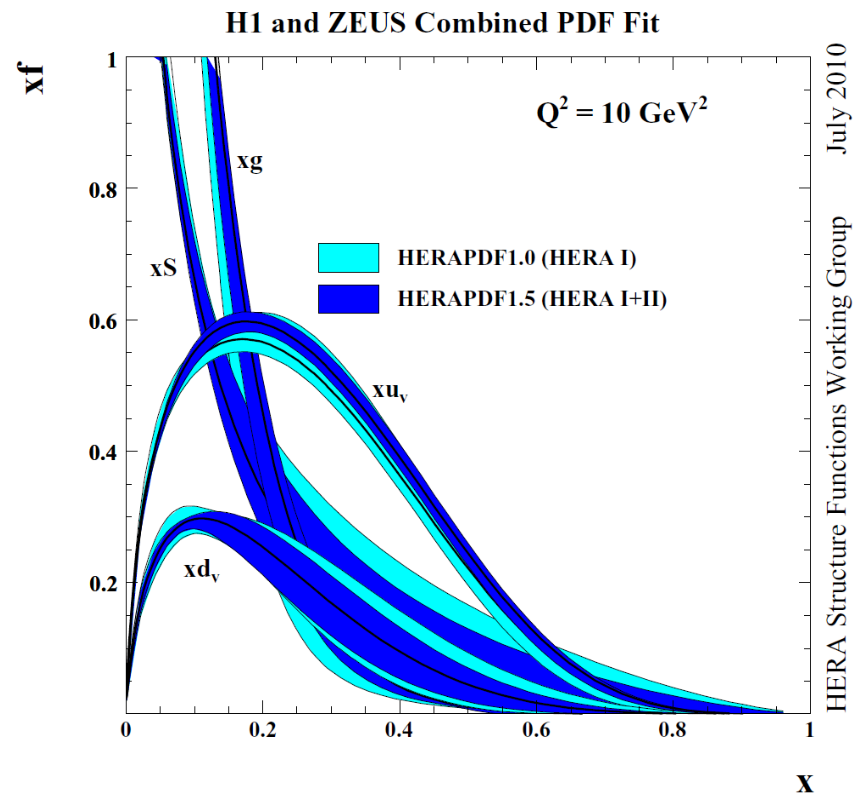




The data on the left has been updated to the data on the right

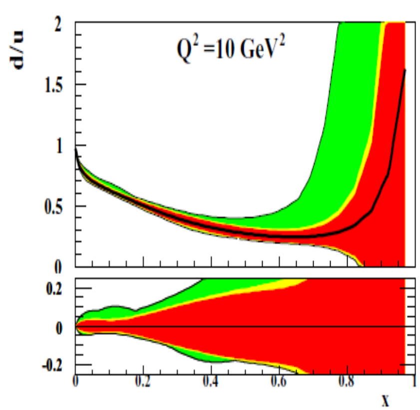
The HERAPDF1.0 fit on the left has been updated to the HERAPDF1.5 fit on the right





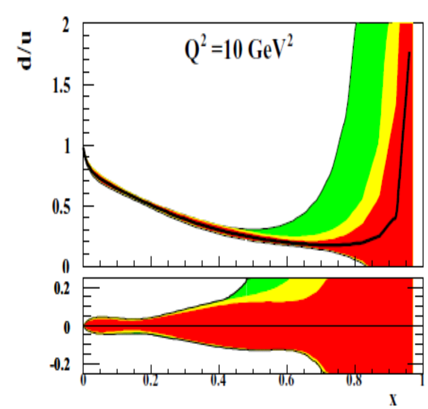
The PDF uncertainties have been reduced at high-x

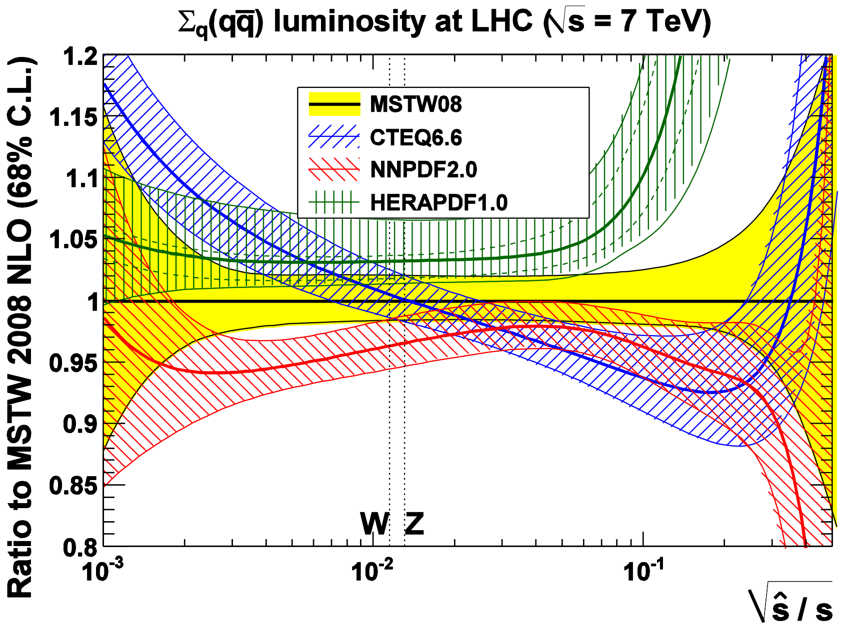
These plots show total uncertainties (model and parametrization included)

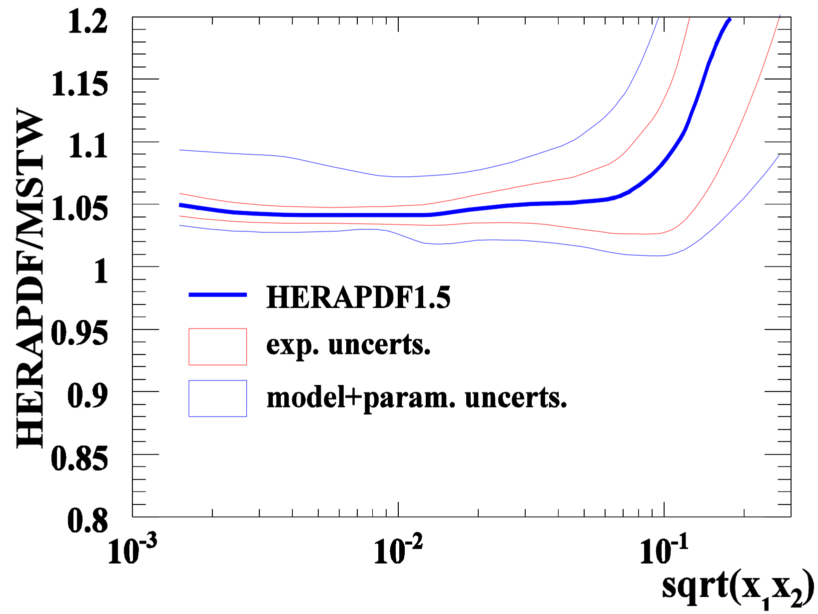


Improved determination of the d/u ratio at high-x.

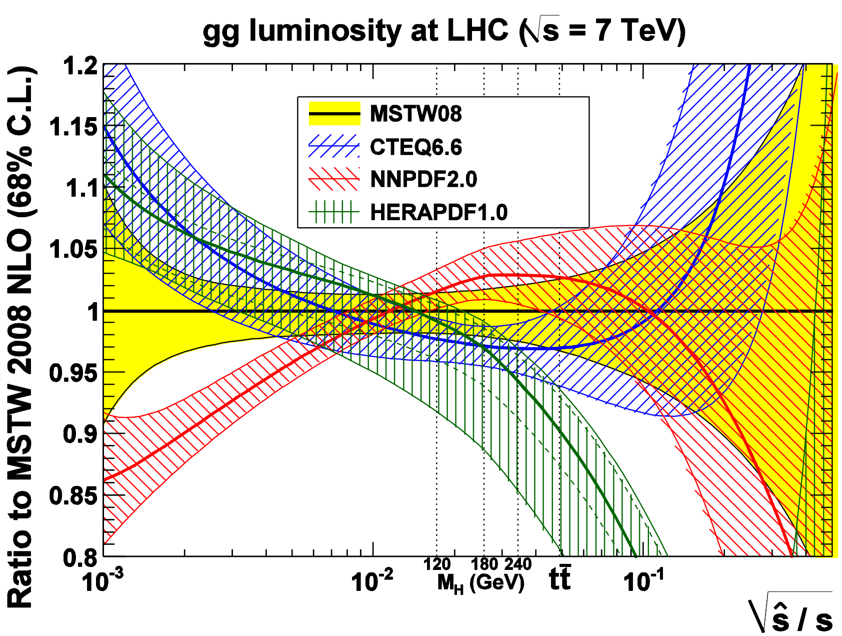
The only PDF which measures d in a proton rather than an isoscalar target

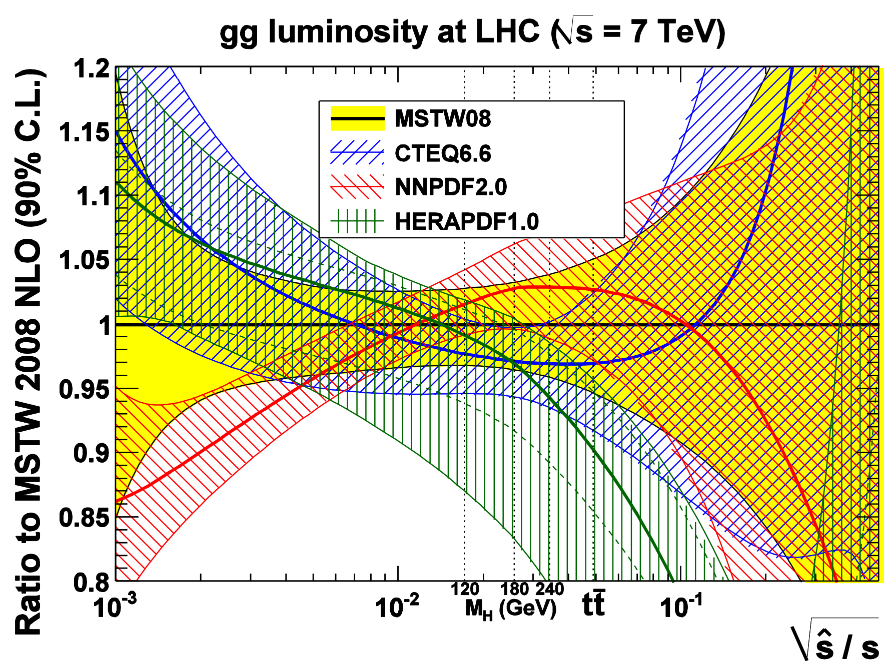




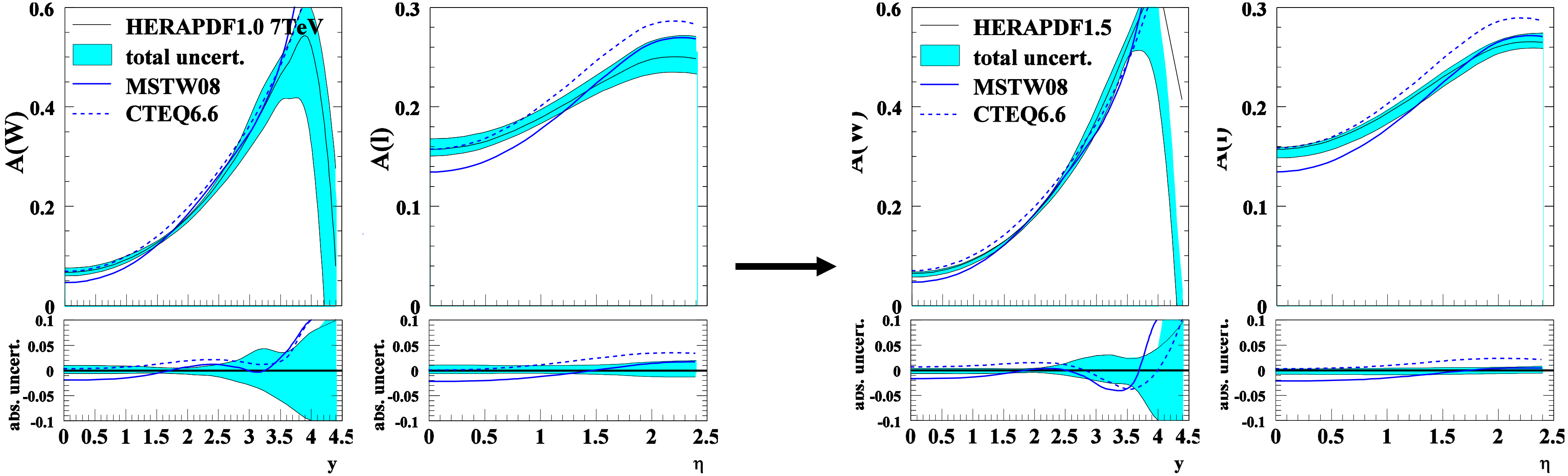


Plots from G.Watt

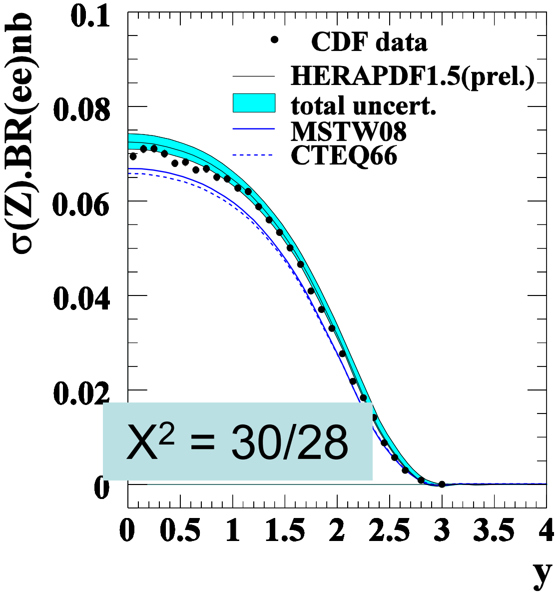


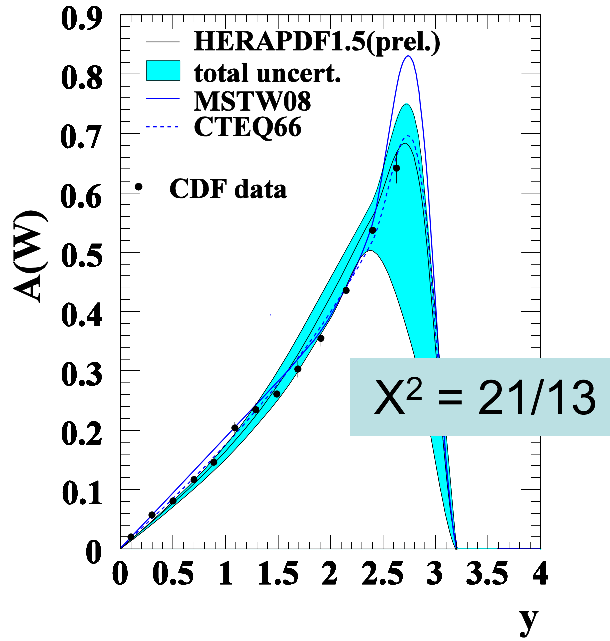


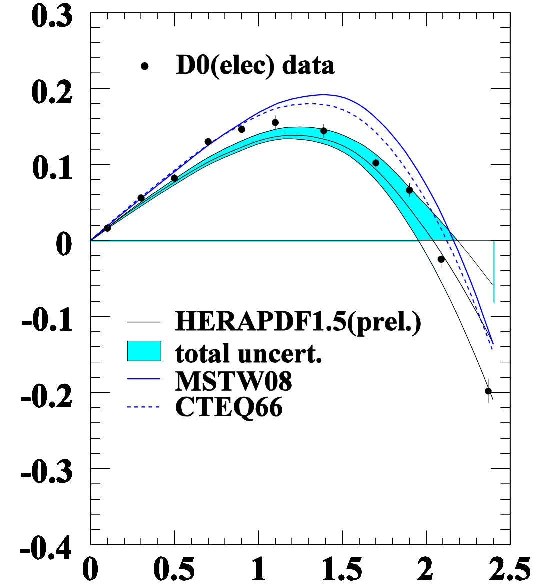
Thus HERAPDFs give predictions for W,Z production at LHC which are comparable to those of CTEQ6.6 ad MSTW08- although different in detail



This reduced high-x error of HERAPDF!.5 results in a reduced error at high rapidity for W/Z production at the LHC





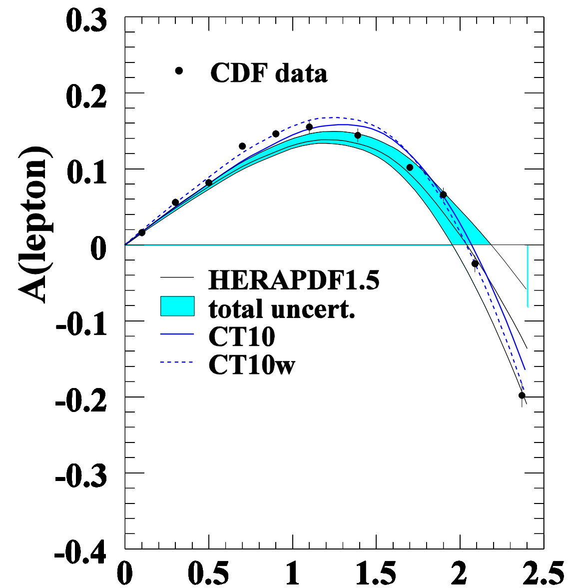


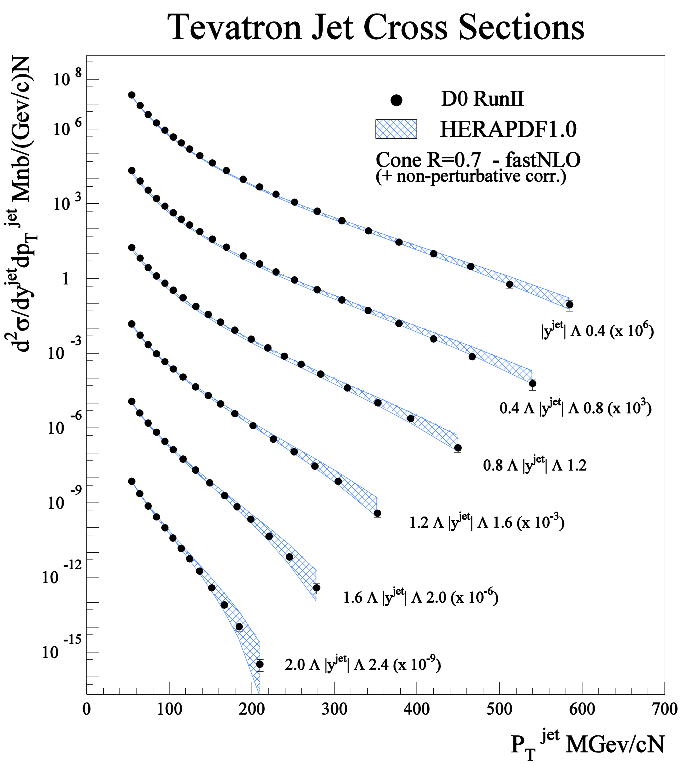
How about Tevatron data?

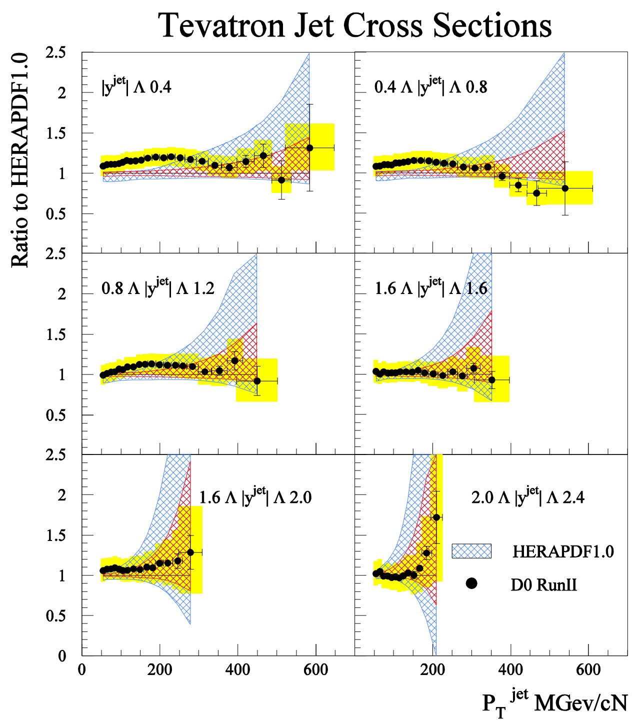
We don't include Tevatron data in the fits but we can describe it – WITHIN OUR ERROR BAND

Even the D0 electron asymmetry data is described- ~as well as other groups

Note some of the trouble comes from tension with NMC and BCDMS fixed target deuteron data- deuterium corrections are one possible explanation- the HERAPDF uses only proton data and is not subject to this uncertainty







How about Tevatron jets?

SUMMARY

Combination of ZEUS and H1 data and HERAPDF fits to these data:

1. Inclusive cross-sections HERA-1 (1992-2000):arxiv:0911.0884 -improved constraints at low-x HERAPDF1.0
2. F2(charm) data (preliminary)- constraints on the charm mass parameter, m_c(model)
3. Low energy runs – FL- 2007- (preliminary) –tension with low x, Q2 data?
4. Inclusive cross-sections HERA-II (2003-2007)- (preliminary) -improved constraints at high-x HERAPDF1.5
5. NLO and NNLO PDFs

PDFs with different α_s =0.114-0.122, m_c =1.35-1.65, m_b = 4.3-5.0

Coming soon: fits with HERA jets data.

extras

Where does the information on parton distributions come from?

CC e-p

$$\frac{d^2\sigma(e^-p)}{dxdy} = \frac{G_F^2\, M^4_W}{2\pi x(Q^2+M^2_W)^2}[x\,(u+c) + (1-y)^2 x\,(\bar{d}+\bar{s})]$$

CC e+p

$$\frac{d^2\sigma(e^+p)}{dxdy} = \frac{G_F^2\, M^4_W}{2\pi x(Q^2+M^2_W)^2}[x\,(\bar{u}+\bar{c}) + (1-y)^2 x\,(d+s)]$$

•We can use the reduced cross-sections to learn about high-x valence PDFs

For NC e+ and e-

$$\frac{d^2\sigma(e\pm N)}{dxdy} = \frac{2\pi\alpha^2 s}{Q^4}\; Y_+\; [\,F_2(x,Q^2) - \frac{y^2}{Y_+} F_L(x,Q^2) \pm \frac{Y_-}{Y_+} xF_3(x,Q^2)],\quad Y_\pm = 1 \pm (1-y)^2$$

$F_2 = F_2^\gamma - v_e P_Z F_2^{\gamma Z} + (v_e^2+a_e^2)P_Z^2 F_2^Z$

$xF_3 = - a_e P_Z xF_3^{\gamma Z} + 2v_e a_e P_Z^2 xF_3^Z$

Where $P_Z^2 = Q^2/(Q^2 + M^2_Z)\, 1/\sin^2\theta_W$, and at LO

$[F_2, F_2^{\gamma Z}, F_2^Z] = \Sigma_i\, [e_i^2, 2e_i v_i, v_i^2+a_i^2][xq_i(x,Q^2) + x\bar{q}_i(x,Q^2)]$

$[xF_3^{\gamma Z}, xF_3^Z] = \Sigma_i\, [e_i a_i, v_i a_i]\quad [xq_i(x,Q^2) - x\bar{q}_i(x,Q^2)]$

So that $xF_3^{\gamma Z} = 2x[e_u a_u u_v + e_d a_d d_v] = x/3\,(2u_v+d_v)$

Where $xF_3^{\gamma Z}$ is the dominant term in xF_3

The difference between NC e+ and e- cross-sections gives the valence structure function xF3 due to γ/Z interference and Z exchange

Note this is obtained on a pure proton target so

•No heavy target corrections

•No assumptions on strong isospin

(Unlike xF3 determined from neutrino scattering on heavy isocalar targets)

Theoretical framework

Fits are made at NLO in the DGLAP formalism -using QCDNUM 17.00

The Thorne-Roberts massive variable flavour number scheme is used (2008 version) and compared with ACOT

The staring scale Q^2_0 (= 1.9 GeV^2) is below the charm mass^2 (mc=1.4 GeV) and charm and beauty (mb=4.75) are generated dynamically

A minimum Q^2 cut $Q^2 > 3.5$ GeV^2 is applied to stay within the supposed region of validity of leading twist pQCD (no data are at low W^2)

Parametrisation and model assumptions

We chose to fit the PDFs for:

gluon, u-valence, d-valence and the Sea u and d-type flavours:

Ubar = ubar, Dbar = dbar+sbar (below the charm threshold)

To the functional form $xf(x,Q_0^2) = Ax^B(1-x)^C(1+Dx+Ex^2)$

The normalisations of the gluon and valence PDFs are fixed by the momentum and number sum-rules resp. Further constraints are:

B(d-valence) = B(u-valence), B(Dbar) = B(Ubar), low-x shape of Sea same for u-type+d-type

A(Ubar) = A(Dbar) (1-fs), where sbar = fs Dbar, so that ubar → dbar as x→ 0 (fs=0.31)

Uncertainties due to model assumptions are evaluated by varying input values

Variation of heavy quark masses:

Mc=1.35 to1.65 GeV (the pole-mass)

Mb= 4.3 to 5.0 GeV

Variaion of the sea fraction

Fs=s/(d+s) = 0.23 to 0.38

Variation of the minimum Q^2 cut on data entering the fit

$Q^2{}_{min}$= 2.5 to5.0 GeV^2

We also vary the value of the starting scale Q^2_0 from1.5 to 2.5 GeV^2: this is considered as a parametrisation uncertainty rather than a model uncertainty…

Parametrisation uncertainties- indicative, not exhaustive

The central fit is chosen as follows: start with a 9 parameter fit with all D and E parameters = 0 and then add D and E parameters one at a time noting the χ2 improvement. Chose the fit with the lowest χ2. This has E(u-valence) ≠ 0 and χ2/ndf = 574/582.

$$xf(x,Q_0^2) = Ax^B(1-x)^C(1+Dx+Ex^2)$$

This is the central fit

PDF	A	B	C	D	E
xg	sum rule	FIT	FIT	-	-
xu_{val}	sum rule	FIT	FIT	-	FIT
xd_{val}	sum rule	$=B_{u_{val}}$	FIT	-	-
$x\bar{U}$	$\lim_{x\to 0}\bar{U}/\bar{D}\to 1$	FIT	FIT	-	-
$x\bar{D}$	FIT	$=B_{\bar{U}}$	FIT	-	-

We then start with this 10 parameter fit and add all the other D and E parameters one at a time noting the χ2 improvement. It turns out that there is no significant further improvement in χ2 for 11 parameter fits.

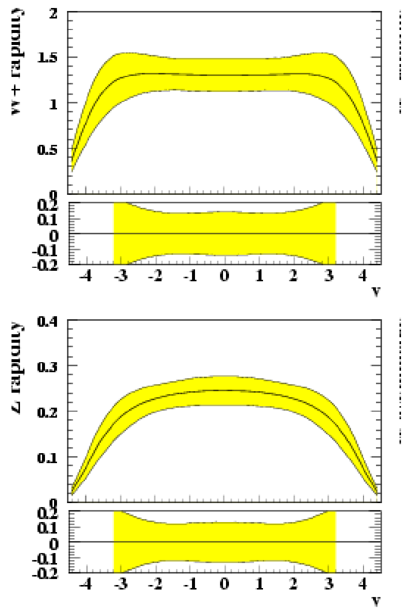
An envelope of the shapes of these 11 parameter fits is formed and used as a parametrization error. This gives the parametrization uncertainty at high-x.

Low-x parametrisation uncertainty is accounted for by the following additional variations:

1. Bdv free –this results in Bdv ≈ Buv
2. A negative gluon term: $-A\,x^B(1-x)^C$ is added to the usual gluon term, when the starting scale of the fit is lowered to Q^2_0=1.5 GeV^2 – this results in a small –ve gluon term but the gluon itself does not become negative in the kinematic range of the data

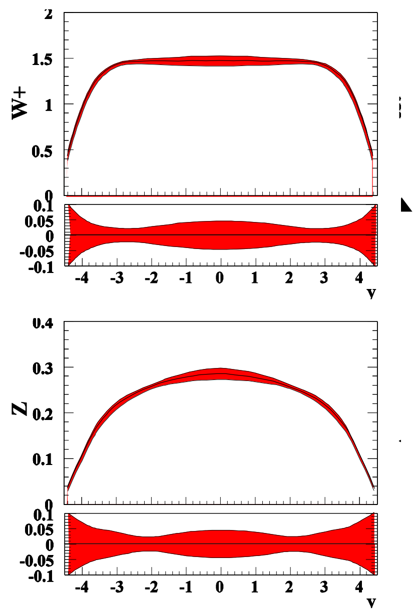
Consequences for W and Z production at the LHC

Look at predictions for W/Z rapidity distributions: Pre- and Post-HERA

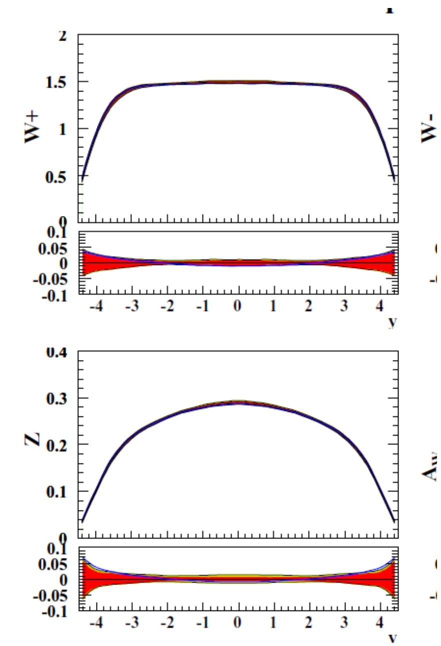


Note difference in scale for fractional errors

Just fixed target DIS data ~15% uncertainty



Separate HERA data sets~5% uncertainty

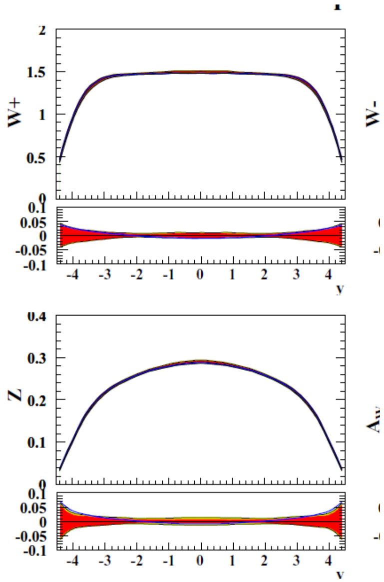


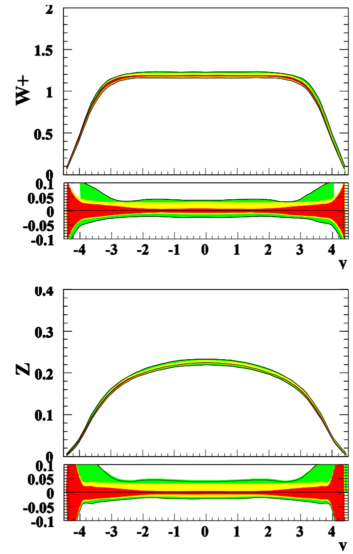
Combined HERA data set~1% uncertainty

Why such an improvement ?

It's due to the improvement in the low-x sea and gluon At the LHC the q-qbar which make the boson are mostly sea-sea partons And at $Q^2 \sim M_Z^2$ **the sea is driven by the gluon**

These illustrations at 14 TeV





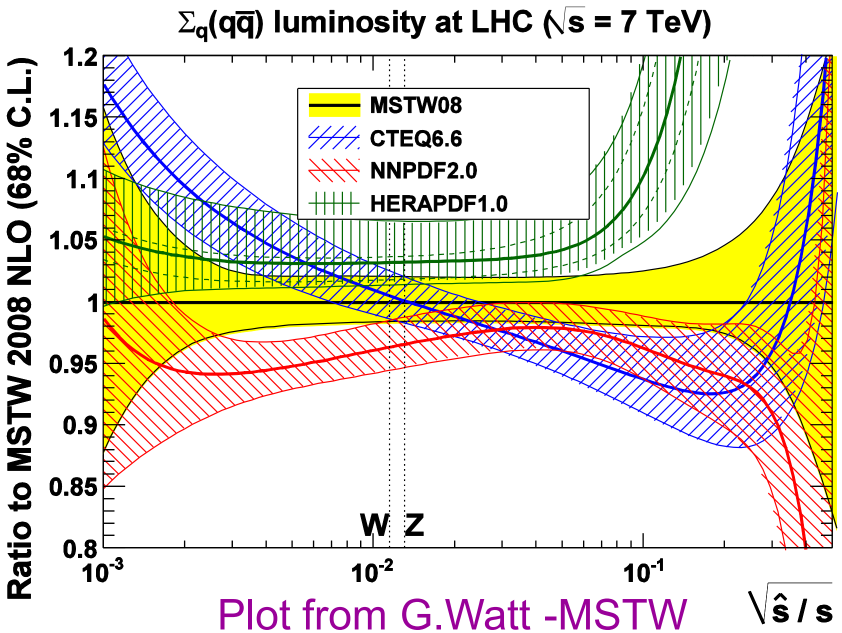
Model errors are the most signficant in the central region: m_c, m_b, fs, Q^2_{min}

m_c =1.35 – 1.65 GeV is the dominant contribution… but this can be improved if F2(charm) data are used…..

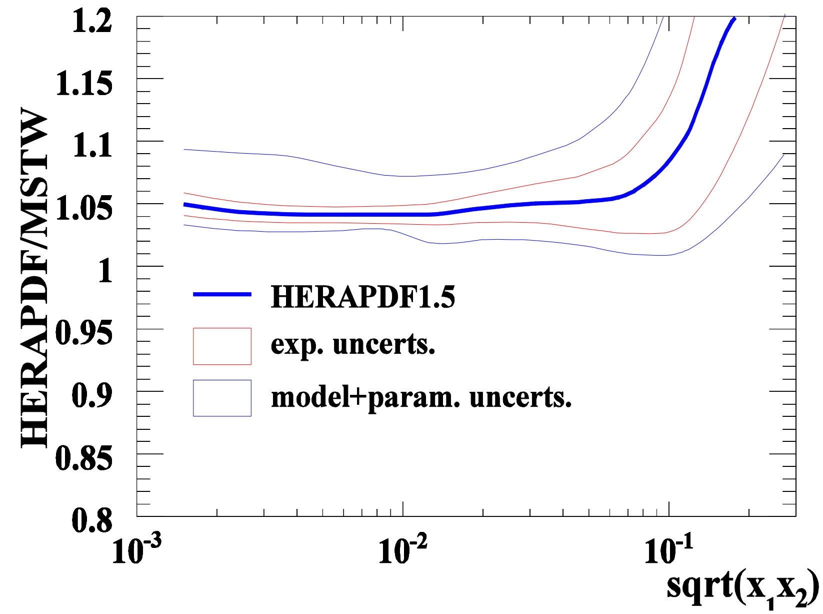
However PDF fitting should also include consideration of model errors and parametrisation errors

HERAPDF1.0 experimental plus model errors plus parametrisation

The PDF4LHC group has been comparing PDFs at the level of parton-parton lumiosities



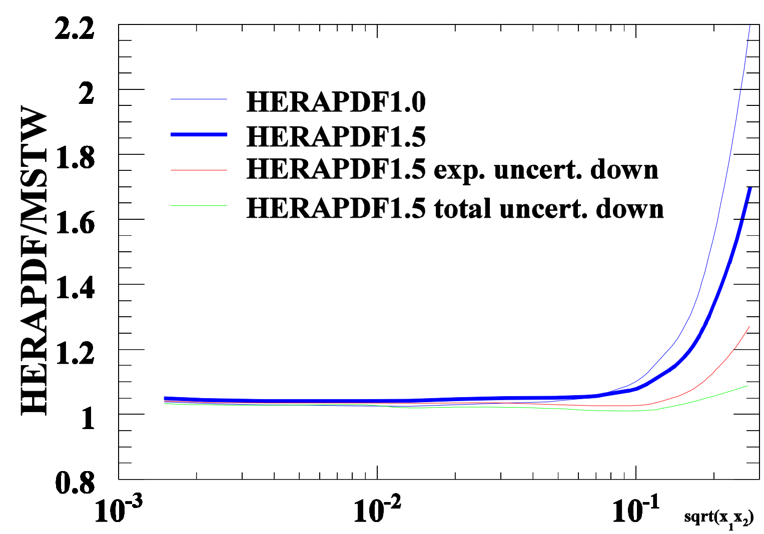
Plot from G.Watt -MSTW

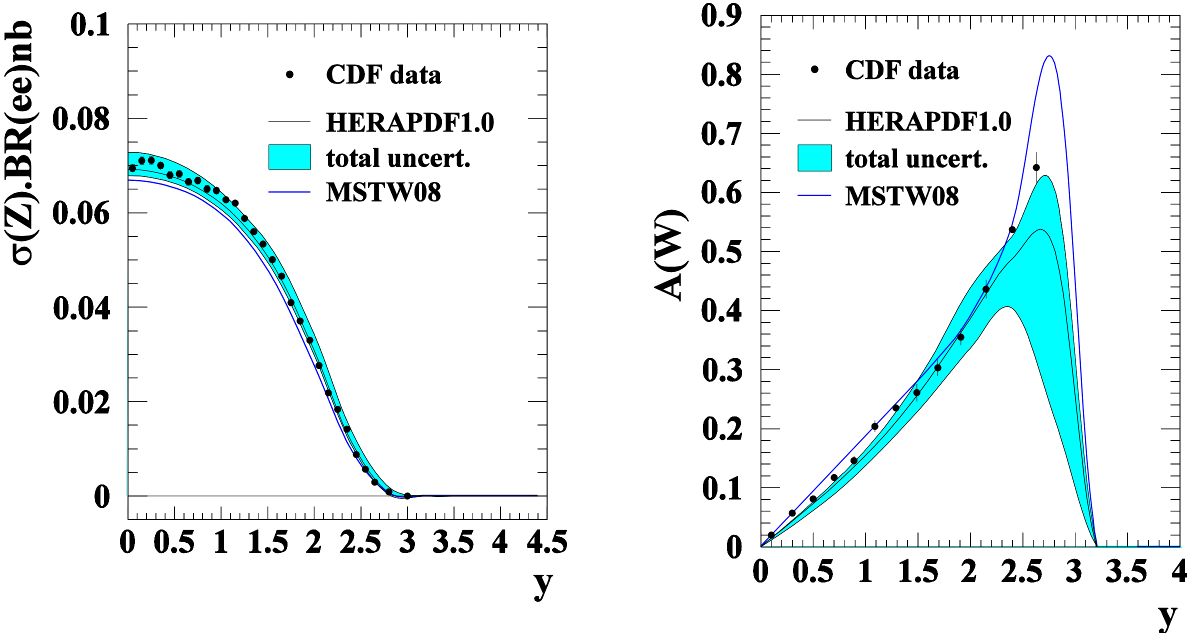


HERAPDF1.0 has a rather high q-qbar luminosity at high scale.

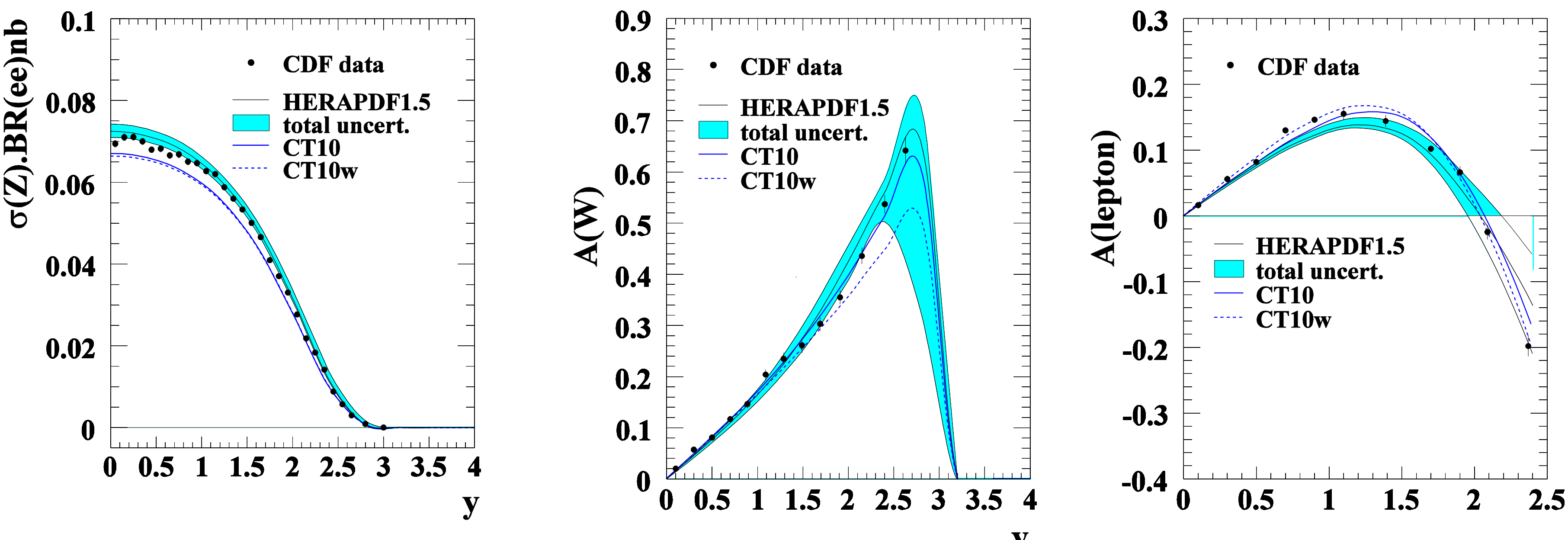
This is reduced in HERAPDF1.5

It is now closer to MSTW within uncertainties





Back up HERAPDF1.0

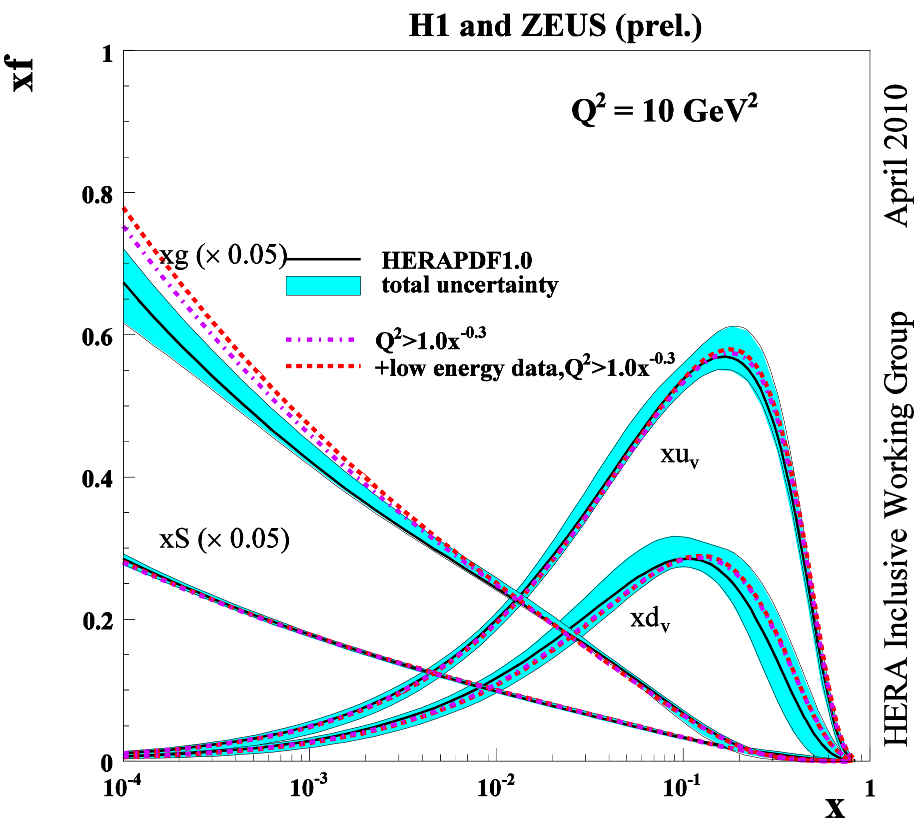


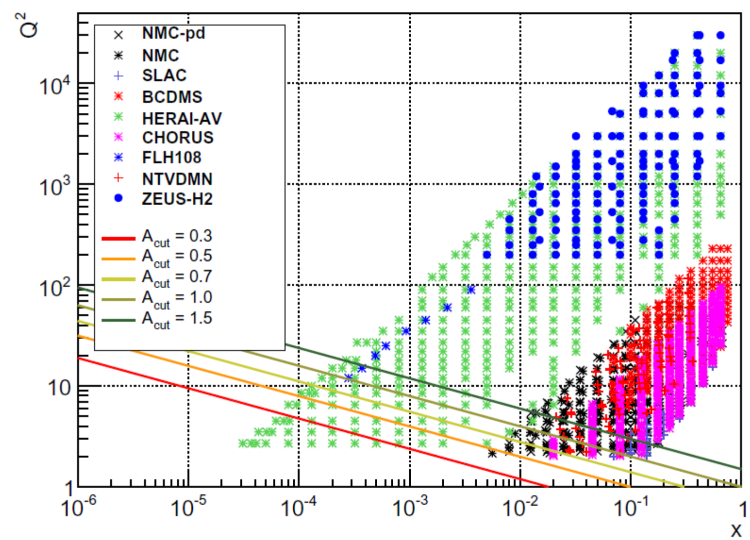
Back-up 15 compared to CT10

How hard do we need to cut such that analysis of just Ep=920 data

And analysis of lower energy data is once more in good agreement?

$Q^2 > 1.0\ x^{-0.3}$





This implies that the 'true' gluon could be a little bit steeper than the HERAPDF1.0 gluon- or indeed CTEQ6.6 or MSTW08 gluons

However this effect only starts to become important for $x < 10^{-3}$ so W/Z cross-sections at the LHC are only marginally affected- 1-1.5% up at 7 TeV