

Improving Long-Term MLS Roster Management with Survival Analysis

...

Kevin Minkus
@kevinminkus

Why do we care about prospect development?

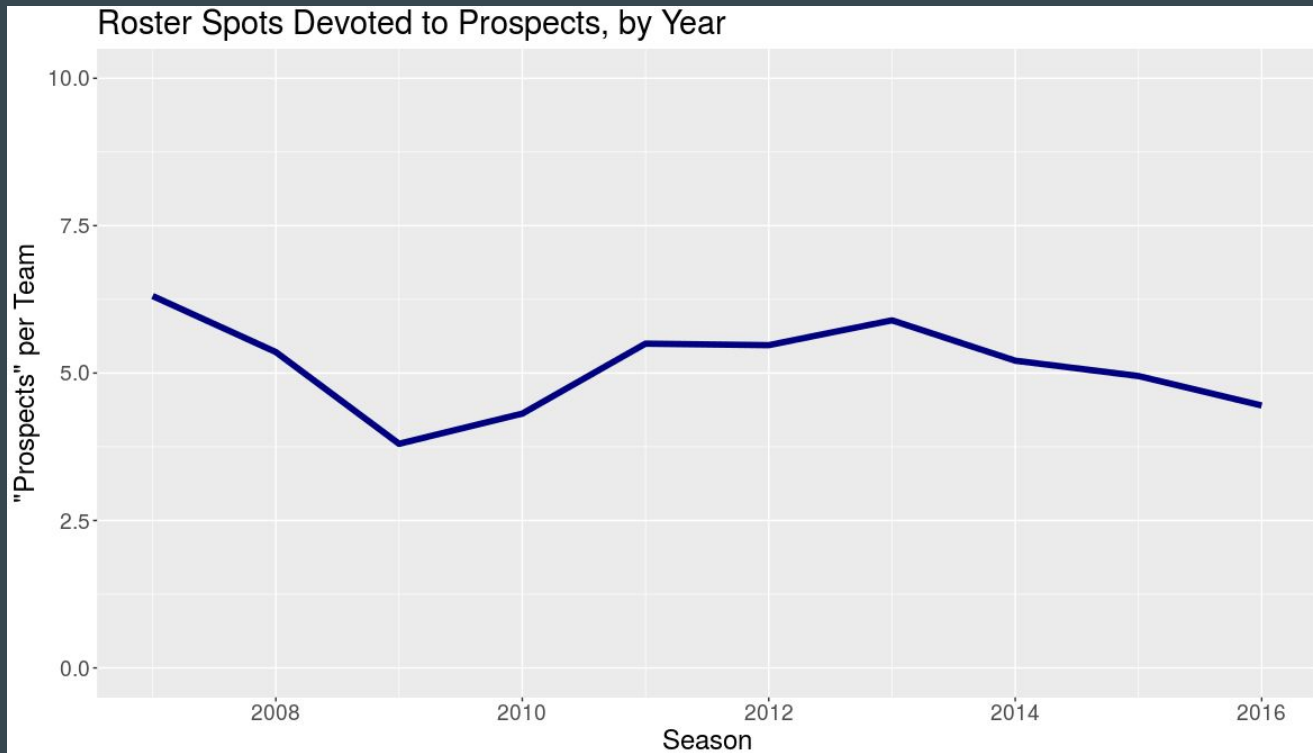
MLS Investment in Prospects

- 11 teams have a full USL second team
- 9 teams have a USL affiliate
- Every team has an academy
- More than \$50 million spent per year on MLS academies

Defining a 'Prospect'

- Player 21 or under
- Drawbacks
 - Diversity of backgrounds
 - Overly restrictive
- 20% of players

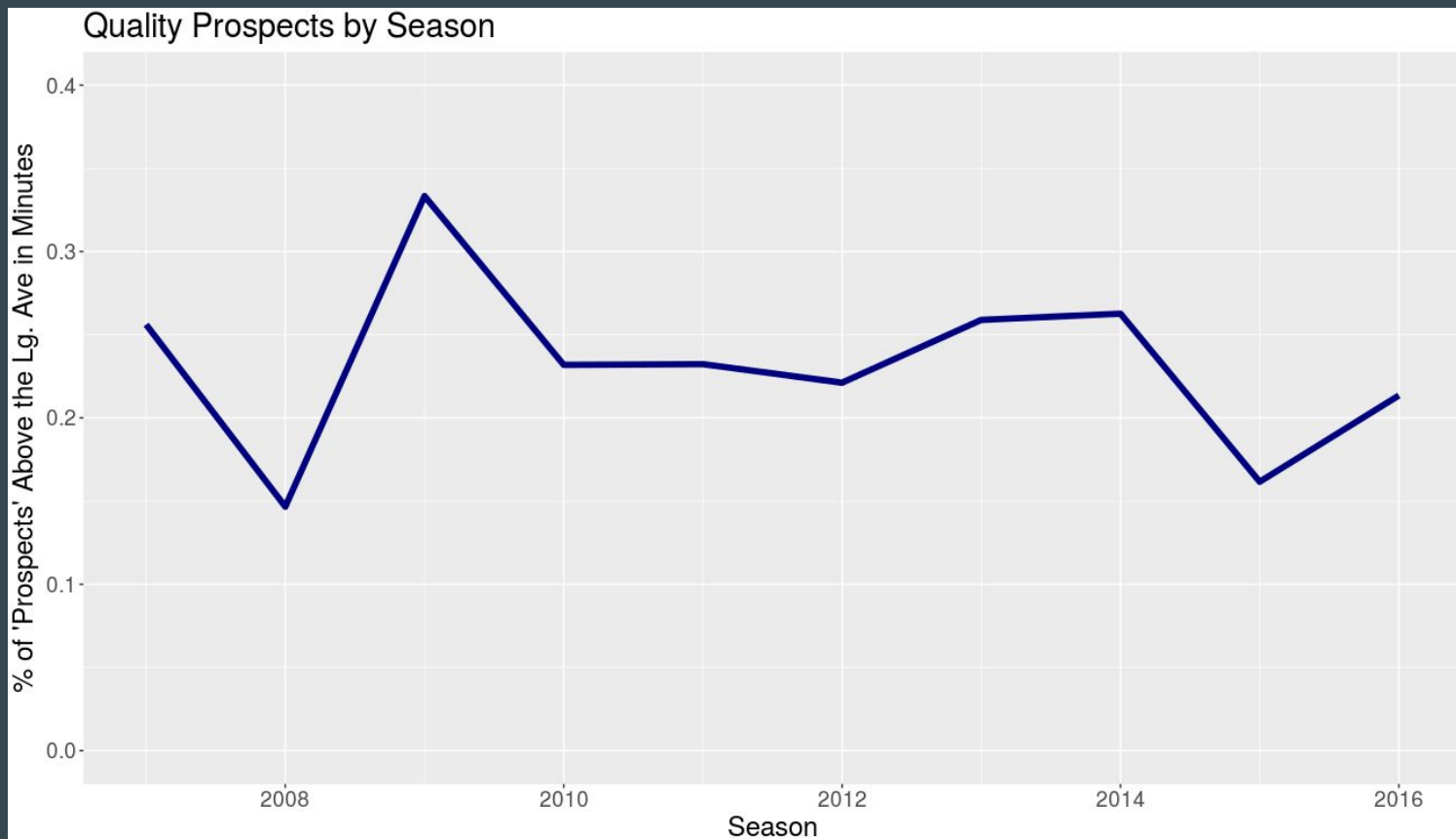
Are more prospects getting into the league?



Defining 'Quality'

- Minutes played in lieu of a one number metric

Are the prospects in the league getting better?

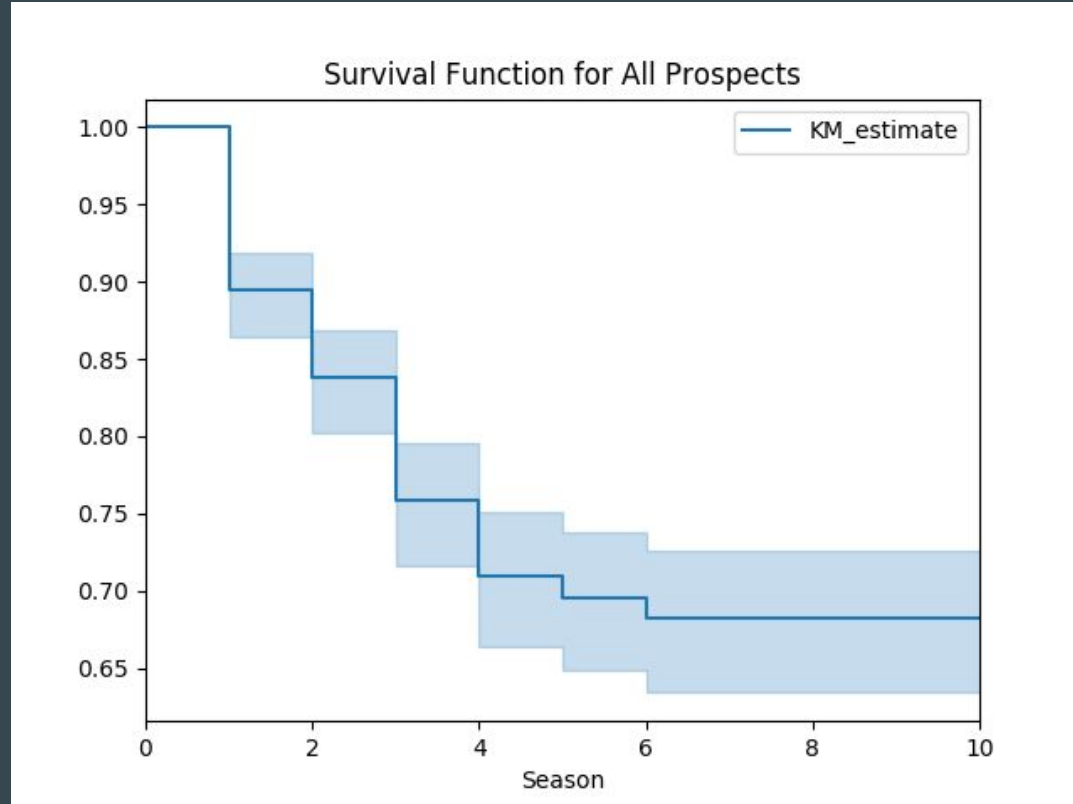


What does development look like for a prospect?

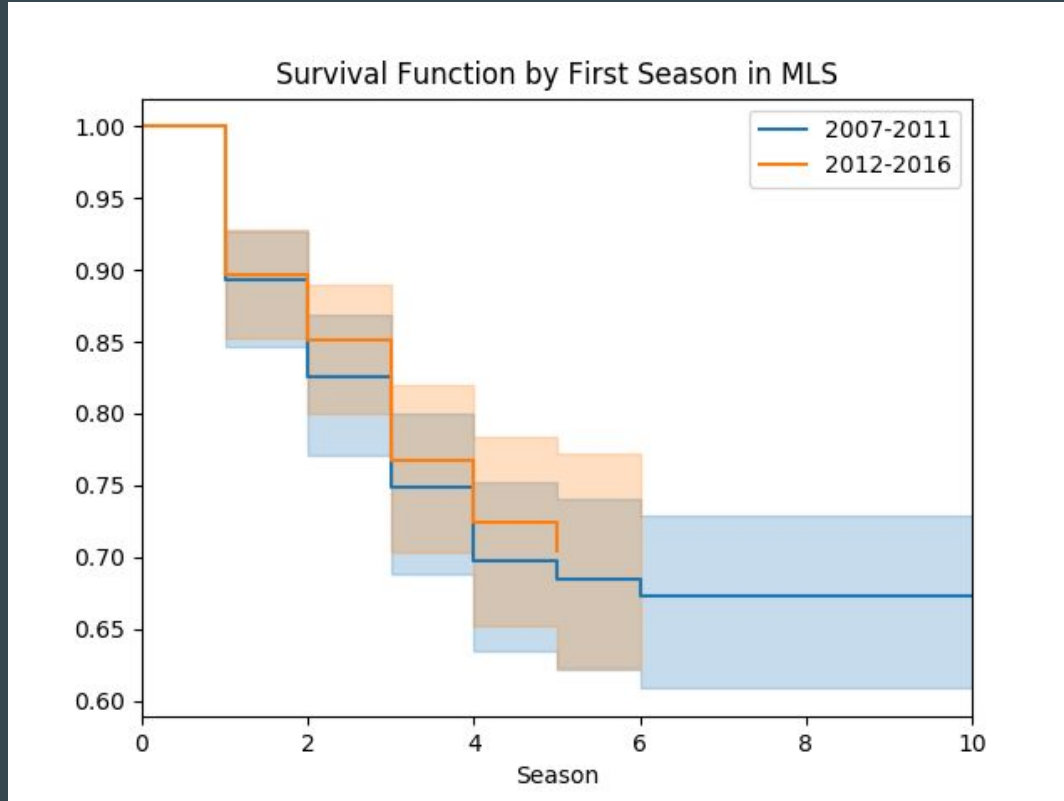
Defining “Starter-quality”

- A player in the 65th percentile or higher of minutes played in a season
- ~9 or 10 players per team

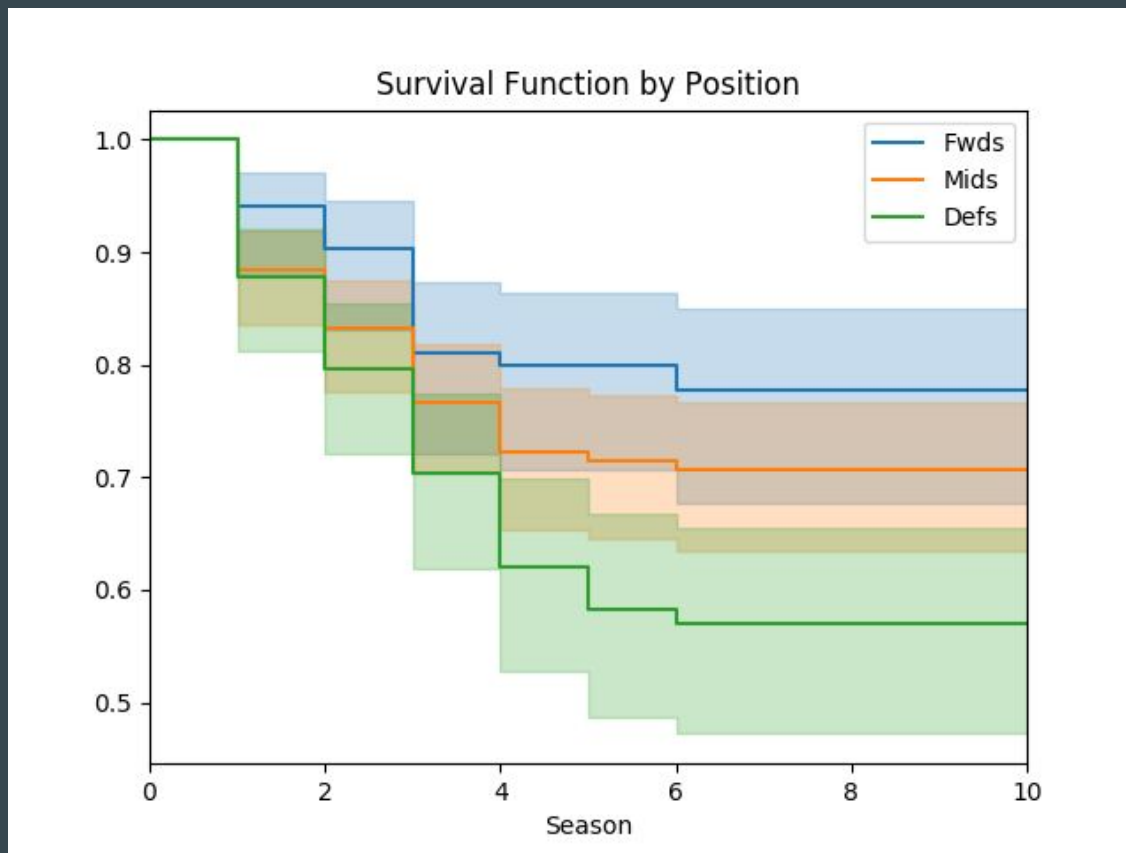
How do prospects currently develop?



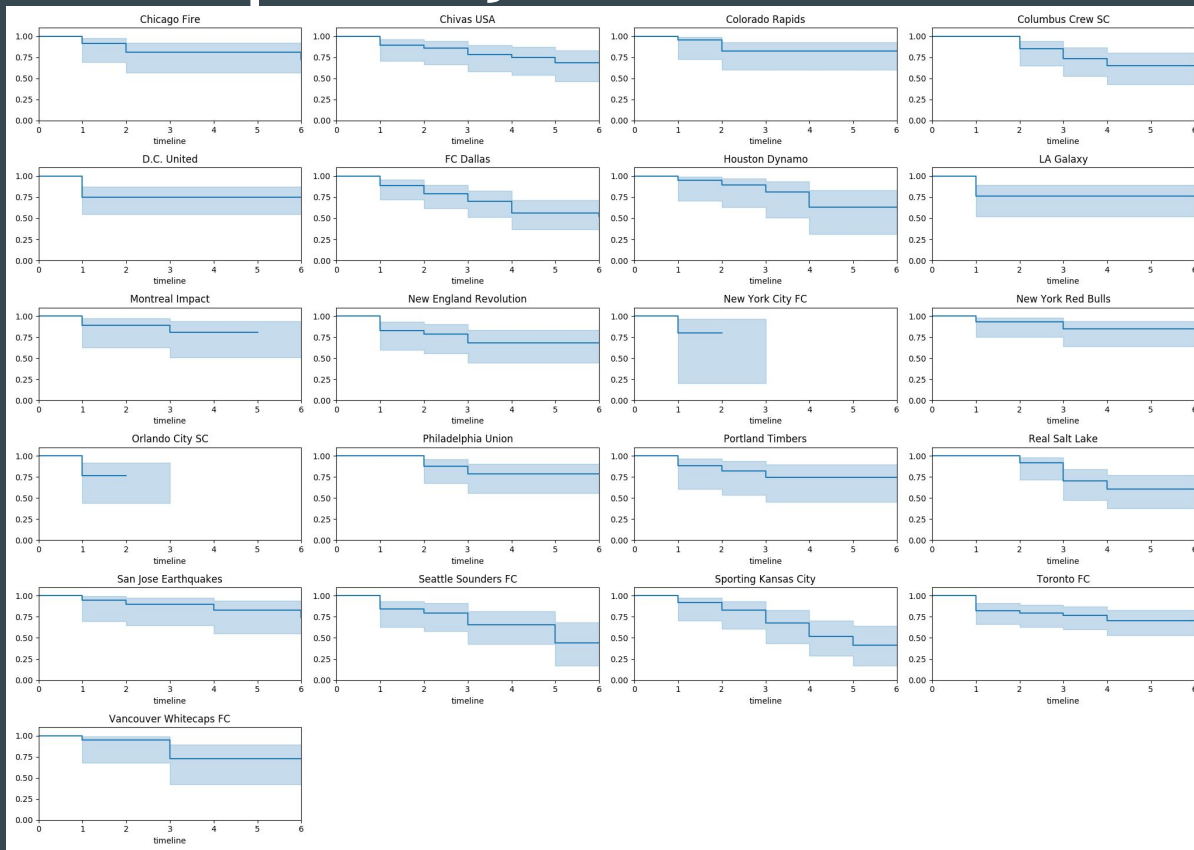
Prospect Development by Season



Prospect Development by Position



Prospect Development by Team



Building a Development Curve Model

Modeling

- Considerations

- Informative censoring
- Discrete time
- Team effects
- Age

- Methodology

- Variables - homegrown flag, draft flag, draft pick, USL2 team, position
- Model form
- Feature selection

Model

Concordance - 0.68

	exp(coef)	exp(-coef)	lower .95	upper .95
positionForward	0.4674	2.1393	0.2807	0.7785
positionMidfielder	0.6769	1.4773	0.4621	0.9914
draftflag	8.7716	0.1140	5.2636	14.6175
pick	0.9283	1.0772	0.9053	0.9519
entry.age	1.1314	0.8839	0.9813	1.3044
hgflag	1.8923	0.5284	1.0787	3.3197

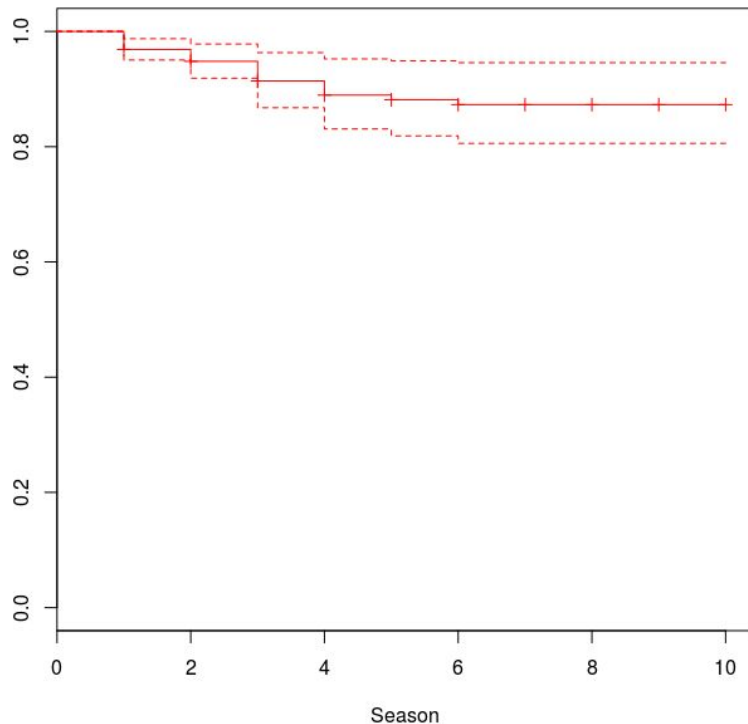
How could this get used in practice?

D.C. United Case Study

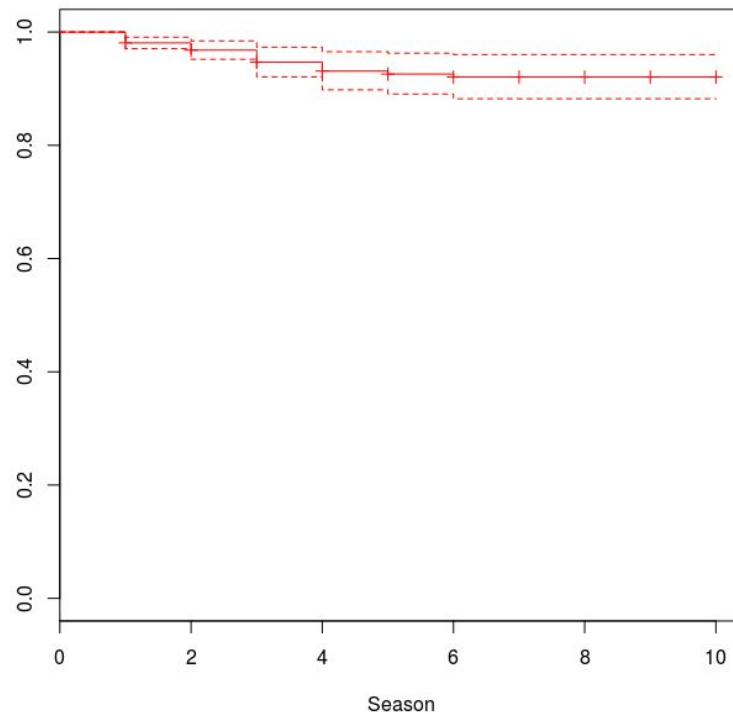
- Ian Harkes and Chris Durkin
- Add baseline hazard to Cox PH

Ian Harkes and Chris Durkin Survival Functions

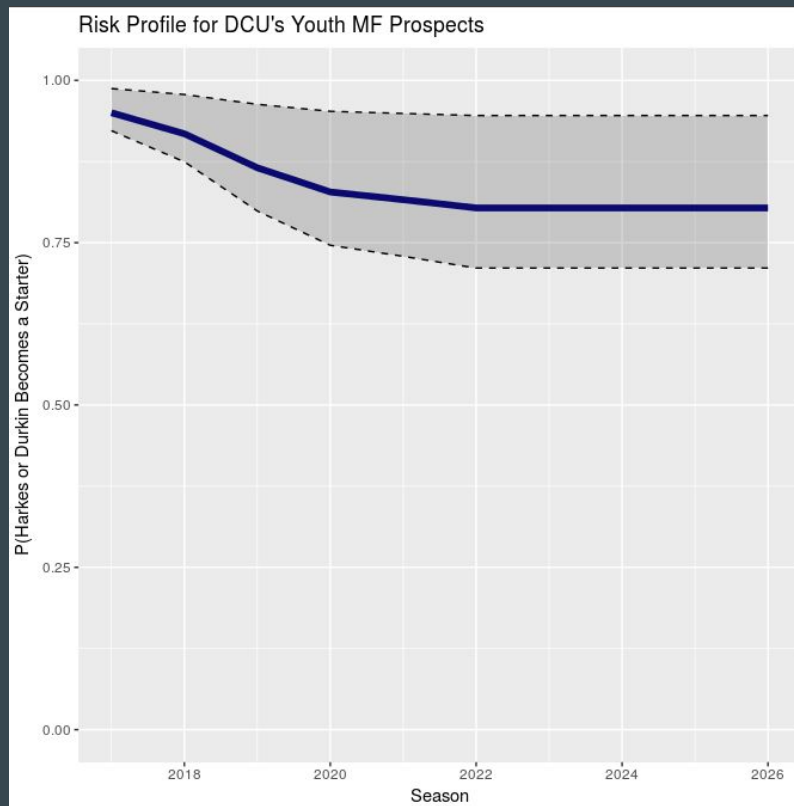
Ian Harkes Development Curve



Chris Durkin Development Curve



Combined Survival Curve



Future Applications

- Prospect acquisition
- Academies
- Other related problems

Links

<https://github.com/kcm30/prospect-development>

<http://data.princeton.edu/pop509/NonParametricSurvival.pdf>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3184466/>

<http://savvastjortjoglou.com/nfl-survival-analysis-kaplan-meier.html>

<http://www.kdd.org/kdd2017/papers/view/prospecting-the-career-development-of-talents-a-survival-analysis-perspecti>

Questions?