



UPPSALA
UNIVERSITET

UPTEC STS 23004

Examensarbete 30 hp

Februari 2023

Sporten bandy beskriven genom sportanalys hos IK Sirius Bandy

Viking Nilsson



UPPSALA
UNIVERSITET

The Sport of Bandy Described with the Use of Sports Analytics at IK Sirius Bandy

Viking Nilsson

Abstract

This essay seeks to describe the sport of bandy using sports analytics while working with the Swedish Elitserien team Sirius Bandy. The task entails data collection, finding suitable analytical models and presenting their results in a helpful manner. The essay will also seek to confirm or deny some hypothesis about the sport of bandy. An array of Python files and programs was created to help the user collect data from a televised live bandy game, create statistical models from the collected data and then automatically generate two types of presentations that presents the findings from these models. These presentations take the form of a *Game Report*, which can be created both in the half-time of a bandy game as well as quickly upon its conclusion, and as a *Season Report* which can be used to analyze more than one game. The code, along with two example reports, can be found in the following [GitHub](#). For more information on how to use the code see the README.md file.

When attempting to understand the sport of bandy with the help of these models some conclusions were made regarding the importance of shooting percentage, shots on goal and shot attempts. It has also been shown that duels between two or more players play a somewhat important role for the games' outcome but require more work to fully comprehend. No relation has been found between ball possession and the outcome of games. Some key aspects which differentiate play indoors and outdoors have also been identified, most notably an increase in the total number of duels and long passes. These differences are amplified when the ice quality outdoors worsens. More work is needed to fully comprehend the sport of bandy, and this is only a first step.

Teknisk-naturvetenskapliga fakulteten

Uppsala universitet, Utgivningsort Uppsala/Visby

Handledare: Jonas Holgersson Ämnesgranskare: David Sumpter

Examinator: Elisabéth Andresdottir

Populärvetenskaplig sammanfattning

Detta arbete söker att förklara sporten bandy med hjälp av sportanalys, både för att öka förståelsen kring sporten, men även för att hjälpa den svenska Elitserieklubben Sirius Bandys verksamhet. Arbetet har skett tillsammans med klubben. En rad påstående rörande sporten bandy kommer ställas upp, och rapporten syftar att med matematiska verktyg bekräfta eller dementera dessa påståenden.

I följande GitHub finns all kod som använts för rapporten, såväl som datan från alla matcher och två färdiga PowerPoint-presentationer där statistiken presenteras. Läsaren kan med fördel gå igenom denna, och i synnerhet README.md-filen i huvudmappen för en mer ingående teknisk förklaring av hur kodbasen används.

Sportanalys är ett område där olika matematiska modeller används för att beskriva sport, och för att skapa insikter om den. Detta är ett brett ämne som finns inom många olika sporter, och som de senaste 30 åren vuxit allt större. I bandy finns det lite eller ingen sportanalys gjord, vilket medför att detta arbete kommer hämta inspiration från sportanalys-modeller gjorda för andra sporter.

Då den offentligt tillgängliga datan inom svensk bandy är extremt begränsad bestod en stor del av arbetet i att skapa en metod för datainsamling. Detta mynnade ut i ett dataninmatnings-program där användaren definierar en rad händelser och sedan matar in dessa händelser. En fil genereras sedan där händelserna är tidsstämplade och där information om var på planen de skett samt av vilket lag.

Denna inmatningsprocess görs bäst från en TV-sänd bandymatch, och kan med lite träning göras i realtid under matchen helt utan att pausa. Efterarbetet som krävs för att den insamlade datan ska kunna användas uppgår till ett par minuter - hela processen tar således några minuter längre än matchtiden. Noterbart är dock den mentala ansträngningen som krävs för att göra denna inmatning; användarens arbetsminne belastas till den grad att det är svårt att samla data från fler än två matcher under en arbetsdag.

Datan från inmatningen användes sedan för att skapa diverse modeller som sökte beskriva sporten. Bland dessa finns diverse skottstatistik-modeller, närkamps-modeller och ett mått på lagens bollinnehav. För att skapa en klarare bild av sporten har flera åtgärder vidtagits. Lagens avslut delades in i sju kategorier som analyserades individuellt och hörstatistiken presenterades baserat på vilken sida hörnpassningen slås från. Vad gäller närkampsstatistiken presenterades den indelad efter var på isen närkampen ägt rum, samt vilket det bollförande laget var inför närkampen.

Det faktum att datainsamlingen är klar kort inpå att matchen är slut möjliggör för en automatiserad sammanställning av datan där en presentation levereras till laget i

såväl halvtid som kort inpå matchslut. Program skapades för att generera två typer av PowerPoint-presentationer; en matchrapport som sammanställer statistik från en enskild halvlek eller match, och en säsongsrappport där statistik från ett flertal matcher presenteras.

Båda rapporttyperna skapades i samråd med Sirius tränare och spelare, då ambitionen var att rapporterna skulle underlätta och förbättra förståelsen av matcherna, och att rapporterna skulle fungera som ett underlag vid samtal. Stor vikt lades vid läslighet och tydlighet, vilket medförde att diverse estetiska inslag prioriterades bort. Även andra anpassningar vad gäller presentationen gjordes för att underlätta för spelarna.

Vid matematisk analys av analysmodellerna framkom att skottförsök och skott på mål var en god indikator för resultat, och en låg utdelning på avsluten kan negativt påverka lags resultat under en längre period. Ett starkt närkampsspel på lagnivå hade visst korrelation med resultat, närkampsspelet föreföll vara en svåranalyserad del av bandy. Inget samband mellan bollinnehav och resultat upptäcktes.

Vad gäller skillnaderna mellan utomhus- och inomhusspel identifierades två förekomster; för det första att antalet långa passningar vid spel utomhus ökade markant, samt att antalet närkampssituationer ökade. Båda dessa skillnader blev än mer påtagliga då isen bedömdes vara dålig.

Sammanfattningsvis bör detta arbete ses som en introduktion till sportanalys inom bandy; det finns flera aspekter av spelet som än är outforskade, ytterligare efterforskning kan med fördel göras på de ämnen detta arbete avhandlat. Rapporten har kommit till vissa insikter om spelet, och utarbetat en anpassningsbar datainsamlings-metod som kan användas av andra personer och bandyföreningar i framtiden.

Förord och tack

Först och främst vill jag tacka alla på Sirius bandy. Det har varit otroligt roligt och lärorikt att se hur ni jobbar, både i med- och motvind. Jag vill speciellt tacka bandygym-gänget för trevligt häng på morgonen och för att jag fick förgylla er tillvaro med vacker musik. Ett extra tack riktar jag till min handledare Jonas Holgersson för otaliga timmars intressanta samtal om såväl bandy, som andra sporter, och samhället i stort.

Sedan vill jag tacka alla mina kursare under studietiden, speciellt Ted. ♥

Mest av allt vill jag dock tacka min sambo Lisa för allt stöd under hela utbildningen. Och min dotter Alice för all hjälp att koppla bort skolan på mornar, kvällar, helger och nätter, främst under utbildningens andra hälft.

Innehåll

1	Introduktion	1
1.1	Syfte och frågeställning	1
1.2	Kod	2
1.3	Källor	2
2	Bakgrund	4
2.1	Bandy som sport	4
2.1.1	Ute- och innespel	4
2.1.2	Hörnor i bandy	5
2.2	Sportanalys	5
2.2.1	Positiva insatser med negativa resultat	6
2.2.2	Bra spelare med höga siffror i bolltapp	6
2.2.3	Corsi	7
2.2.4	Förväntade mål, xG	7
2.3	Hypoteser	8
2.3.1	Att som lag vinna en stor del av närkamperna bidrar inte till att vinna matcher	8
2.3.2	Att som lag ha en hög corsiprocent bidrar inte till att vinna matcher	8
2.3.3	Att som lag ha ett högt bollinnehav bidrar inte till att vinna matcher	8
2.3.4	Skott utifrån har inte en lägre sannolikhet att bli mål än skott närmare från mål	8
2.3.5	Vid spel utomhus och på dålig is slår lagen inte fler långa passningar	9
2.3.6	Vid spel utomhus och på dålig is placeras närkamperna inte mer i mittzonerna och framför mål	9
2.4	Matematiska metoder	10
2.4.1	Hypotestestning av två urval med Z-test	10
2.4.2	Korrelationskoefficient	10
2.4.3	Nollhypotes-test med linjär regression	11
3	Avhandling	12
3.1	Datainsamling	12
3.2	Direktinsamlings-metod	12
3.3	Ställningstaganden och problematik vid datainsamling	14
3.3.1	Styrkor och problematik med insamlingsmetoden	14
3.3.2	Planens zonindelning	15
3.3.3	Avslut	16

3.3.4	Närkamper	17
3.3.5	Videokvalitet på TV-sändningarna	18
4	Resultat och analys	19
4.1	Dataunderlag och Sirius kräftgång under hösten 2022	19
4.2	Övergripande statistik	19
4.2.1	Resultat - övergripande	20
4.2.2	Analys - övergripande	20
4.3	Avslut och mål	20
4.3.1	Resultat - avslut	20
4.3.2	Analys - avslut	21
4.4	Fasta situationer	22
4.4.1	Resultat - fasta situationer	22
4.4.2	Analys - fasta situationer	22
4.5	Närkamper och dueller	23
4.5.1	Resultat - närkamper	23
4.5.2	Analys - närkamper	24
4.6	Hypoteser	26
4.6.1	Att som lag vinna en stor del av närkamperna bidrar inte till att vinna matcher	26
4.6.2	Att som lag ha en hög corsiprocent bidrar inte till att vinna matcher	28
4.6.3	Att som lag ha ett högt bollinnehav bidrar inte till att vinna matcher	30
4.6.4	Skott utifrån har inte en lägre sannolikhet att bli mål än skott närmare från mål	31
4.6.5	Vid spel utomhus och på dålig is slår lagen inte fler långa bollar	31
4.6.6	Vid spel utomhus och på dålig is placeras inte närkamperna mer i mittzonerna och framför mål	32
4.7	Presentationer till Sirius Bandy	35
4.7.1	Matchrapport	35
4.7.2	Säsongsrapport	36
5	Diskussion	38
5.1	Datainsamlingsmetoden	38
5.2	Sirius resultat	38
5.3	Rapporternas påverkan på Sirius Bandy	39
5.4	Vidare studier	39
5.4.1	Med detta dataunderlag	39
5.4.2	Med ett utökat dataunderlag	40

5.4.3	Med en annan insamlingsmetod	40
6	Slutsats	42
6.1	Svar på del-frågeställningarna	42
6.1.1	Hur kan relevant data samlas från bandymatcher?	42
6.1.2	Vilka sportanalys-modeller kan skapas från datan?	42
6.1.3	Hur ska resultaten från dessa metoder tolkas?	42
6.1.4	Hur ska dessa resultat presenteras på bästa sätt?	43
6.2	Svar på rapportens huvud-frågeställning	43

1 Introduktion

Linköpings universitet definierar sportanalys (engelska: “sports analytics”) som något som “använder data kopplad till sport för att vinna insikter om sporten”. Konceptet har funnits i mer än 50 år, till exempel har den nordamerikanska ishockeyligan NHL har hållit officiell plus/minus-statistik sedan 1959. De senaste åren har mer sofistikerade och avancerade modeller utvecklats för många olika sporter, bland annat baseball, amerikansk fotboll, ishockey, basket och fotboll [41] [22].

Klubbar i många olika sporter och länder har upptäckt att sportanalys kan vara till stor hjälp, och ofta kan det vara ett väldigt kostnadseffektivt sätt att stärka sin verksamhet. Föreningar som är tidiga med att använda det får ofta ett stort försprång framför klubbar som inte nyttjar sportanalys alls. Ett exempel på det är baseballaget Oakland Athletics som i början av 2000-talet hade en av MLB:s lägsta lönebudgetar, men som mycket tack vare en stark sportanalys-avdelning kunde konkurrera med lag vars lönebudgetar var nästan tre gånger så stora [23].

Inom den svenska högstaligan i bandy finns det nästintill ingen historik av att nyttja sportanalys; många lag analyserar inte matcherna mer än med penna och papper. Detta borde i teorin innebära att ett lag som investerade i sportanalys skulle kunna tillskansa sig en fördel jäntemot sina motståndare.

1.1 Syfte och frågeställning

Syftet med examensarbetet är att skapa sportanalysmodeller och tillvägagångssätt för att förstå och beskriva sporten bandy, och därmed hjälpa Sirius Bandys verksamhet. Målet är att skapa ett system som går att vidareutveckla och anpassa i framtiden. Det finns inget befintligt system för sportanalys alls, så arbetet kommer bestå i att skapa ett sådant från grunden.

Detta innebär att frågeställningen blir som följer:

- Hur kan sporten bandy, och mer bestämt IK Sirius spel, beskrivas med hjälp av sportanalys-modeller på ett vis som ökar förståelsen kring bandy och hjälper IK Sirius verksamhet?

För att kunna besvara detta kommer vi också behöva besvara följande frågor:

1. Hur kan relevant data samlas från bandymatcher?
2. Vilka sportanalysmodeller kan skapas från datan?
3. Hur ska resultaten från dessa metoder tolkas?
4. Hur ska dessa resultat presenteras på bästa sätt?

Det kommer även att ställas upp en rad hypoteser kring sporten bandy som rapporten sedan skall försöka bekräfta eller dementera.

1.2 Kod

All kod och data som krävs för att återskapa projektet finns i följande GitHub. Kodbasen är relativt stor, så den är indelad i filer. Alla funktioner och metoder har en förklarade docstring, så `help(funktion)` fungerar på samtliga funktioner. En mer genomgående förklaring av hur koden bör användas finns i README.md-filen. Läsaren kan med fördel läsa den filen parallellt med denna rapport.

I Githubens huvudmapp ligger även två PowerPoint-presentationer; en match- och en säsongsrapport. Dessa är exempel på de rapporter jag skapat till tränarna löpande under säsongen.

1.3 Källor

Källunderlaget för bandyanalys är i stort sett obefintligt, så kommer rapporten förlita sig på och ta inspiration från källor från andra sporter, främst fotboll och ishockey, då dessa sporter är de som oftast jämförs med bandy. Däremot kommer överförbara koncept från andra sporter även appliceras.

De akademiska artiklar som finns inom sportanalys är ofta tungt maskinlärnings-orienterade, eller fokuserar på en väldigt specifik aspekt, som i många fall kan vara svårapplicerade i ett såpass brett och generellt arbete som detta. Det har därför varit svårt att hitta relevanta akademiska källor till arbetet.

Däremot finns en uppsjö av välskrivna blogginlägg i diverse sporter; sportanalys är såpass nytt att många modeller och koncept utvecklats under en period när internetbloggande varit stort. Det har ofta utvecklats olika grupper på internet där likasinnade delat och diskuterat sina upptäckter. Inom dessa grupper har inte någon formell peer review-process existerat, men eftersom dessa bloggare varit aktiva på Twitter och andra sociala medier har det skapats ett informellt peer review-lik system, där författare har granskat och refererat till varandras arbeten. Flera av dessa bloggare har efter hand fått jobb som analytiker i klubbar i ligorna de skrivit om, eller som journalister på stora mediehus för att bevaka ligorna. Detta kan ses som en bekräftelse på att deras arbete varit tillförlitliga och hållit hög klass.

Det finns även diverse privatdrivna statistikhemsidor med högt anseende inom de olika sporterna. Dessa sidor erbjuder ofta olika typer av data och statistik såväl som artiklar om analys.

Arbetet har skett tätt med Sirius Bandy, jag som författare har i princip haft fri

tillgång till spelare och ledare inom klubben. Detta har innebär att inga formella intervjuer med personer inom klubben gjorts, däremot har daglig kontakt med framförallt Sirius tränare, men även med spelare förekommit. I rapporten kommer det refereras till påståenden från såväl tränare som spelare, dessa kommer emellertid aldrig hänvisa till en specifik intervju då några sådana aldrig ägt rum.

2 Bakgrund

Nedan följer en kort genomgång av sporten bandys historia, och några aspekter av spelet som bedömts vara extra viktiga och/eller intressanta. Sedan en rad sportanalys-modeller och koncept hämtade från andra sporter. Efter följer en rad hypoteser om bandy, framtagna tillsammans med Sirius tränare, som skall verifieras. Till sist en genomgång av de matematiska modeller som används i analysen.

2.1 Bandy som sport

Bandy är en klubb- och bollsport på is som har sitt ursprung i England på 1700-talet, men som idag främst spelas i Norden och Ryssland. Sporten spelas mellan två lag med elva spelare vardera. Samtliga spelare är skridskoförsedda, och alla utom målvakterna (en per lag) har klubba. Bollen är i serisrosa färg och drygt sex centimeter i diameter. Både plan och målburar är något mindre än sina motsvarigheter i fotboll, och matcherna spelas (likt fotboll) i 2×45 minuter. Precis som i fotboll låter man matchuret gå även när bollen inte är i spel och längre avbrott kompenseras genom tilläggstid [17].

Generellt är bandyn och fotbollen lika, både vad gäller spelförlopp och till regler, sporterna delar exempelvis offsideregeln. Det finns dock en rad skillnader, exempelvis tillåts flygande och obegränsade byten i bandy, något som är begränsat i fotboll [19].

Det faktum att spelarna är skridskoförsedda gör att bandyspelet är betydligt snabbare än det i fotboll, i detta hänseende har sporten likheter med ishockey. En annan likhet med ishockey är att spelarna har klubbor i båda sporterna. Dock skiljer sig bandy från hockey i det avseendet att man inte får spela lika fysiskt i bandy [31].

I Sverige har det funnits ett officiellt seriespel i bandy i över hundra år, den första SM-finalen spelades redan 1907 där IFK Uppsala vann. Sedan 2007 är Elitserien den hösta serien där de 13 lagen möter varandra två gånger. De tio bästa lagen går sedan till slutspel där lagen spelar bäst av fem-serier fram till finalen som spelas i en avgörande match på Studenternas IP i Uppsala [8] [2].

2.1.1 Ute- och innespel

Bandy har historiskt sett spelats utomhus, till en början på naturis och från 1900-talets andra hälft på konstis. 2003 byggdes Sveriges första inomhus-bandyhall i Edsbyn, och sedan dess har det tillkommit allt flera. Idag spelar alla utom fyra Elitserielag sina hemmamatcher inomhus. Utomhus blir spelet mer eller mindre påverkat av nederbörd, vind och temperatur, i vissa väderförhållanden kan domaren

bestämma att matchen ska spelas 3×30 istället för 2×45 minuter för att isen är i så dåligt skick [33] [5] [19].

Sirius, som är ett av de fyra utomhuslagen, spelar därmed drygt hälften av sina matcher utomhus. Tränarna har vid flera tillfällen sagt att man i sin förberedelse inte skiljer på hemma- och bortamatcher, utan istället inne- och utematcher för att matchbilden blir så påverkad. Det verkar inte finnas någon formell beskrivning på vad exakt som skiljer ute- och innespel, så rapporten kommer undersöka detta närmare.

2.1.2 Hörnor i bandy

Hörnor i bandy skiljer sig betydligt från motsvarigheten i fotboll. Vid en bandyhörna ställer det anfallande laget upp i en båge vid straffområdets ytterkant, medan det försvarande laget står bakom mållinjen vid målburen. När hörnpassningen slås från hörnflaggan får det försvarande laget rusa ut mot hörskyttarna för att försöka täcka undan skottet, och till skillnad mot i fotboll så blir det inte ytterligare en hörna ifall bollen täcks ut av en spelare i det försvarande laget, i sådana fall tillfaller bollen istället det försvarande lagets målvakt. Hur många försvarare som rusar respektive står kvar är upp till varje lag att själva välja, och det har genom historien varit olika hur stor del av laget som rusat. Det anfallande laget skjuter vanligen direkt, men kan vissa fall välja att först spela bollen för att komma runt försvararna [19].

En hörna anses vara en farlig målchans som uppstår mellan 15 och 22 gånger per match. Lag gör under en säsong vanligen mål på mellan 8 och 15 % av sina hörnor, men under kortare perioder kan denna effektivitet vara både betydligt lägre och högre än så. Sirius tränare pratar ofta om "hörnmatchen", och betraktar hörnor som en väldigt viktig, men något frikopplad del av spelet, likt hur man inom hockey ser på spel i numerärt över- och underläge [3].

2.2 Sportanalys

Sportanalys är ett relativt nytt område, speciellt på den nivån den är på idag. Många stora klubbar i olika sporter har numera stora sportanalys-avdelningar som de lägger mycket pengar på, och vissa anlitar utomstående konsultföretag för att hjälpa till med det. Kraven på sportanalytikerna har blivit mer matematiska och mindre sportfokuserade; i november 2021 sökte det amerikanska baseball-laget New York Mets en analytiker med krav på magisterexamen i datavetenskap eller statistik, men utan krav på baseball-kunskaper eller erfarenhet [29] [27].

Denna snabba utveckling inom sportanalys har till stor del påverkat spelet i de olika sporterna. Ett exempel på detta är att vissa hävdar att utvecklingen i den nordame-

rikanska basketlagen NBA, där lagen börjat skjuta betydligt fler trepoängs-skott och färre långa tvåpoängs-skott det senaste decenniet till stor del beror på sportanalys; lagen har sett att de förväntas göra fler poäng på trepoängarna än de långa tvåpoängarna och anpassat sitt spel därefter [36] [34].

Nedan följer ett antal koncept och modeller tagna från olika sporters sportanalys som har relevans i detta projekt.

2.2.1 Positiva insatser med negativa resultat

En nästintill paradoxal insikt som avancerad statistik och sportanalys medfört i hockey är hur vissa insatser som kan förefalla positiva i själva verket korrelerar med dåliga resultat. 2012 visade den kanadensiska hockeyjournalisten Cam Charron, som senare kom att arbeta som analytiker åt NHL-laget Toronto Maple Leafs i åtta säsonger och som nu skriver på sporthemsidorna *The Athletic* och *Daily Faceoff*, att ha många täckta skott korrelerar med att förlora matcher i NHL. Resonemanget är att det inte i sig är negativt att täcka skott, generellt sett är det till och med bättre att täcka ett skott än att släppa igenom det, däremot är det bästa att vara den som skjuter, vilket det täckande laget aldrig är [10] [1] [13][12].

2.2.2 Bra spelare med höga siffror i bolltapp

Detta resonemang från flera olika sporter liknar det ovanstående, men kan ses som det inverterade exemplet, är att de spelare med flest bolltapp ofta är bland de bästa i hela ligan. Totalt har de tio NBA-spelare med flest bolltapp (engelska: "turnovers") spelat 85 NBA all star-matcher (samtliga spelare har spelat minst två stycken), blivit utsedda till ligans MVP tio gånger och NBA-finalens MVP åtta gånger. Bland de tio spelare med flest bolltapp i NHL-historien (engelska: "giveaways") finns alla utom en av spelarna med bland de 35 högst betalda i ligans historia, även här har samtliga spelare blivit uttagna till all star-matchen [39] [40] [18].

Resonemanget är snarlikt det för positiva insatser med negativa resultat; bra spelare har ofta bollen, och kommer ibland tappa den. Detta blir extra påtagligt för de med en spelfördelande spelstil (exempelvis guards i basket och centrar och offensiva backar i hockey). Dessutom finns aspekten av att bra spelare ofta försöker göra svårare aktioner (exempelvis dribblingar) än dåliga spelare, och sätter sig i lägen där de kan komma att tappa bollen. Att gradera spelare utifrån vilka som har flest bolltapp kan därför bli intetsägende. Inom basketanalys har försök gjorts att kontextualisera bolltapp, men det finns ingen allmänt vedertagen modell [45].

2.2.3 Corsi

Corsi är ett av de mest centrala och äldsta begreppen inom ishockeyanalys och har sitt ursprung i den Edmontonbaserad hockeyanalysblogg-sfären under andra halvan av 2000-talet. Begreppet handlar om antal skottförsök, eller corsi-händelser, (CF), det vill säga skott på mål, utanför eller täckta, ett lag haft i relation till skottförsök emot (CA) vid spel fem mot fem. Corsi presenteras vanligen i procentform enligt ekvation 1 och används både på spelar- och lagnivå [26].

$$CF\% = \frac{CF}{CF + CA} \quad (1)$$

Resonemanget bakom statistiken är att eftersom skott bara skjuts in anfallszon kan corsi till viss del användas för att approximera puckinnehav. Det är också knutet till mål, men med en betydligt större urvalsstorlek. Detta gör att corsi bättre än gjorda mål förutspår framtida prestationer - på lång sikt bör spelare och lag med hög CF% få med sig bättre resultat än de med låg. Verktöget är inte perfekt, och har många nackdelar, exempelvis tas ingen hänsyn alls till typen av skott eller vilken spelare som skjuter, men det är en lättskapad modell, som tillsammans med andra datatyper kan användas för att beskriva sporten [20] [37].

Generellt gäller att spelare och lag med en hög CF% för spelet vilket oftast är bra. De absolut bästa spelarna i ligan har en CF% i de höga 50-talen; förra säsongen var det endast tre spelare i NHL (minst 50 spelade matcher) med en CF% över 60 % [35].

2.2.4 Förväntade mål, xG

Förväntade mål (engelska: expected goals, xG) finns idag i flera olika sporter, men har sitt ursprung inom fotbollen på 1990-talet. Begreppet försöker uppskatta kvaliteten på ett enskilt skott, och används ofta genom att summera alla skott under en match och presentera varje lags totala förväntade mål, antingen som antal eller i procentform. Resonemanget är att spelares och lags gjorda mål över tid bör gå mot deras xG medan den i en enskild match bättre kan återspegla matchbilden än faktiska mål [42] [32] [28] [9] [44].

Det finns ingen bestämd definition av exakt hur en xG-modell ska se ut, eller vad den ska ta hänsyn till, och olika modeller kommer ge snarlika, men inte exakt lika, resultat. Vanligen tar xG-modeller i fotboll hänsyn till avstånd till mål, vinkel, kroppsdel skottet skjuts med och vilken typ av framspelning som föranlett skottet. Det finns mer sofistikerade modeller som också väger in målvaktens och försvararnas position och vilken skotthöjd bollen skjuts från, men det finns nästintill

oändligt många parametrar en modell skulle kunna ta i beaktning. xG-modeller tar, likt corsi, ingen hänsyn till vem skytten är [38].

2.3 Hypoteser

Nedan följer en rad hypoteser runt sporten bandy som detta arbete söker besvara. Dessa är framtagna tillsammans med Sirius Bandys tränarstab under höstens gång.

2.3.1 Att som lag vinna en stor del av närkamperna bidrar inte till att vinna matcher

Vikten av ett starkt närkampsspel är något som Sirius tränare betonat. Motiveringen är att närkampsspel är viktigt såväl för att stoppa motståndarnas anfallsspel, som att stärka sitt anfallsspel genom boll-återerövring och att bibehålla innehav.

Under säsongen gång har detta reviderats; efter att ha sett hur närkamperna fördelats över banan är hypotesen att närkamper i försvars- och anfallszon är viktigare än närkamper i mittzon.

2.3.2 Att som lag ha en hög corsiprocent bidrar inte till att vinna matcher

Bandy har många liketer med ishockey, så rapporten undersöker corsibegreppet i en bandykontext. Rapporten söker även att jämföra corsi med skott på mål.

2.3.3 Att som lag ha ett högt bollinnehav bidrar inte till att vinna matcher

Studier inom fotboll har pekat på en stark korrelation mellan bollinnehav och att vinna matcher. En studie av 35 UEFA-ligor för säsongerna 2017 och 2018 visade att ligavinnaren i sitt hade ett bollinnehav på 57 %. En genomgång av data från säsongerna 2015-2019 för Champions League visade att laget med högst bollinnehav vann 295 matcher och förlorade 172 stycken. Då bandy till stor del liknar fotboll vore det intressant att undersöka detta förhållandet mellan vinst och bollinnehav även i bandy [14] [11].

2.3.4 Skott utifrån har inte en lägre sannolikhet att bli mål än skott närmare från mål

Skott utifrån har generellt ett lägre förväntat mål-värde i såväl fotboll som ishockey. Rapporten ämnar undersöka detta i bandy [44] [28].

2.3.5 Vid spel utomhus och på dålig is slår lagen inte fler långa passningar

Detta är allmänt vedertaget inom bandy att utomhusspel leder till att lagen slår fler långa passningar, då den dåliga isen gör det svårare att driva och spela bollen längs isen. Det vore intressant att se hur stor inverkan detta har på spelet. Enligt Sirius tränare fattar man vid något läge, om isen bedöms vara dålig, ett beslut om att ändra hela sin spelidé till att nästintill bara slå långa passningar, något som hela laget sedan gör. Rapporten söker att undersöka till vilken grad detta påverkar antalet långa passningar.

2.3.6 Vid spel utomhus och på dålig is placeras närkamperna inte mer i mittzonerna och framför mål

Detta hör samman med föregående punkt; spelarna får svårt att driva bollen och när de möter motstånd tvingas de slå långbollar mot motståndarnas straffområde istället för att anfälla via kanterna. Rapporten söker en skillnad i närkampernas placering på banan vid olika istyper.

2.4 Matematiska metoder

Nedan följer en kort genomgång av en rad matematiska metoder som kommer att användas.

2.4.1 Hypotestestning av två urval med Z-test

Givet två binära slumpvariabler som inhämtats oberoende av varandra från två populationer kan detta test avgöra ifall skillnaden mellan dessa är statistiskt signifikant [7].

Låt k_1, k_2 vara lyckade utfall dragna ur n_1, n_2 försök. Detta ger oss $p_1 = k_1/n_1, p_2 = k_2/n_2$. Nollhypotesen $H_0 : p_1 = p_2$ antas och testas den genom att beräkna z enligt ekvation 2:

$$z = \frac{p_1 - p_2}{\sqrt{p(1-p)(1/n_1 + 1/n_2)}} \quad (2)$$

Där p definieras enligt ekvation 3.

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} \quad (3)$$

Det erhållna z jämförs sedan med $z_{\alpha/2}$ för ett givet konfidensintervall, ifall $|z| > z_{\alpha/2}$ förkastas nollhypotesen H_0 . I tabell 1 finns en lista på olika $z_{\alpha/2}$ för olika konfidensintervall. Vanligen används 95 % eller 99 % [7].

Tabell 1: $z_{\alpha/2}$ för olika konfidensintervall.

Konfidensintervall	$z_{\alpha/2}$
80 %	1,282
85 %	1,440
90 %	1,645
95 %	1,960
99 %	2,576

2.4.2 Korrelationskoefficient

För att avgöra ifall det finns ett samband mellan två variabler kan dess korrelationskoefficient beräknas. Detta kan göras på flera olika sätt, men vanligen används Pearsons korrelationskoefficient. Den beräknas enligt ekvation 4

$$r_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \quad (4)$$

Där cov är covariansen och σ standardavvikelsen.

Korrelationskoefficienten kan anta värden $r \in [-1, 1]$, där värden nära 1 tyder på en stark positiv korrelation, värden kring 0 inte tyder på någon korrelation och värden nära -1 tyder på en stark negativ korrelation. Det finns en rad problem med korrelationskoefficienten, exempelvis går det att skapa scenarior där variablerna har ett tydligt samband men r nära 0. Utöver detta innebär korrelation inte kausalitet. Trots detta är det ett vanligt förekommande verktyg för att identifiera samband mellan variabler [25].

2.4.3 Nollhypotes-test med linjär regression

En linjär regressions-modell definieras som funktion där en eller flera oberoende prediktorvariabler x approximerar en responsvariabel y enligt ett linjärt och kontinuerligt samband, exempelvis enligt ekvation 5. För att hitta de värden på a som bäst approximerar y beräknas felet enligt någon metod, exempelvis minstakvadratmetoden, och detta fel minimeras sedan [24].

$$y = a_0 + a_1 x_1 \quad (5)$$

Linjär regression kan användas för att testa ifall det finns ett samband mellan en prediktor- och en responsvariabel. Detta görs genom att testa nollhypotesen H_0 att ett samband mellan prediktorvariabel x_n och responsvariabel y inte existerar, alltså $H_0 : a_n = 0$. För att göra detta beräknas t enligt ekvation 6. Det erhållna t -värdet kopplas sedan till ett P -värde. Vanligen förkastas nollhypotesen H_0 om $P < 0.05$ [24].

$$t = \frac{a_n - 0}{\sigma / \sqrt{N}} \quad (6)$$

Där a_n :s värde erhållits när modellen tränats, σ standardavvikelse och N är antal datapunkter.

3 Avhandling

I detta avsnitt följer en avhandling av projektets datainsamlingsdel. Det kommer beskrivas hur datainsamlingen gått till, vilka val och ställningstaganden som tagits samt varför. Den kod som hör till det här avsnittet finns i filen `get_data.py`.

3.1 Datainsamling

Projektets första del var att hitta ett system för datainsamling. Den statistik som finns tillgänglig offentligt bedömdes tidigt vara otillräcklig, då projektet eftersöker mer ingående data än endast skott, mål och frislägg. Varje Elitseriematch TV-sänds, och kan ses direktsänd såväl som i efterhand. Det beslutades att datainsamlingen skulle ske genom att titta på matcher och på något sätt anteckna vissa händelser. Detta är ett vanligt tillvägagångssätt för att samla data för sportanalys, och kan ta olika lång tid beroende på vilken typ av data som eftersöks. I ett medverkande i den nordamerikanska hockey-podcasten *The Hockey PDOcast* berättade Cam Charron, som numera skriver på sporthemsidorna *The Athletic* och *Daily Faceoff*, att det tar honom mellan fyra och sex timmar att samla data kring skott och puckförflyttning in i anfallszon från en ishockeymatch [6][4][15].

Eftersom rapporten skrivs tidigt i bandyanalys-processen gjordes valet att fokusera på en sådan typ av datainsamling som går att göra under match. Ambitionen var att denna insamling bara skulle ta några minuter längre än den faktiska matchens speltid, och att det därefter automatiskt skulle kunna genereras någon form av beskrivande statistik utifrån datan. Förhoppningen var inledningsvis att denna statistik skulle kunna presenteras för tränare och spelare i halvtid och/eller kort efter match.

3.2 Direktinsamlings-metod

Verktyget skall vara helt anpassningsbart både vad gäller vilken typ av data som insamlas samt i vilken form datan sparas. Eftersom datainsamlingen inledningsvis endast kommer göras av personer inom Sirius Bandy kan en liten inlärningskurva accepteras, och inga krav på ett visuellt gränssnitt ställs. Detta krav på anpassningsbarhet gjorde att det svårt att hitta ett färdigbyggt program för datainsamling, detta ihop med de obefintliga kraven på gränssnitt innebar att ett självskrivet program bedömdes bäst fylla kravbilen.

Det faktum att man i bandy låter matchuret gå även då bollen går ur spel är en stor fördel vid datainsamlingen. I programmet anger användaren när halvleken startar, och antecknar sedan i realtid när olika händelser sker. Programmet genererar en CSV-fil av storleken $(N + 1) \times 2$ där N är antalet antecknade händelser, första

kolumnen är texten som användaren fört in och andra kolumnen är tidsstämplar som börjar på 0 : 00 : 00. Datasamling från en 45 minuter lång halvlek resulterar i mellan 170 och 300 datapunkter

För att underlätta processen lades små förbättringsfunktioner till; exempelvis att programmet ska bortse från tomma inmatningar och möjlighet för användaren att ställa om klockan (vilket möjliggör för användaren att pausa TV-sändningen eller synkronisera tiden vid eventuella störningar på TV-sändningen), samt möjligheten att radera inmatningar.

Relativt snabbt blev jag såpass van vid programmet att jag från en TV-sändning kunde samla data från en match utan att behöva pausa. Däremot var det betydligt svårare att göra det på plats i en bandyhall. Det bestämdes därför att ifall samlingen ska ske i realtid måste matchen spelas upp på TV. En annan upptäckt som gjordes var hur mycket datainsamlingen belastade arbetsminnet och hur trött användaren blev av att göra det; inledningsvis var målet att datainsamlingen skulle uppgå till minst tre matcher à 90 minuter. Detta fick revideras till omkring tre halvlek istället.

En fördel med att datan sparas i CSV-filer är att man som användare kan läsa dem, och att det i vissa fall går att redigera filerna manuellt, något som gjorts ett fåtal gånger under projektets gång.

Programmet heter `collector_raw`. I figur 1 ses terminalen vid användning och CSV-filen som generas.

```
next event: avlsag sirius z5
next event: närkamp sirius z6
next event: passning strafområde sirius z8
next event: skott räddning sirius z8
next event: skottyp sirius z8 inlägg
next event: utkast vsk z7
next event: närkamp sirius z2
next event: passning lång sirius z8
next event: närkamp vsk z8
next event: brytning sirius z6
next event: mål straffområde sirius z8
next event: skottyp dribbling sririus z8
```

(a)

event	time
avlsag sirius z5	00:00:00
närkamp sirius z6	00:00:11
passning strafområde sirius z8	00:00:23
skott räddning sirius z8	00:00:28
skottyp sirius z8 inlägg	00:00:32
utkast vsk z7	00:00:36
närkamp sirius z2	00:00:50

(b)

Figur 1: `collector_raw`: (a) Inmatningar (b) CSV-fil.

För att organisera datan genererad av `collector_raw` på en form som passar skapades metoden `clean_csv`. Den går itererar igenom filen rad för rad och organiserar den i en CSV-fil av $(N + 1) \times 5$ där raderna i tur och ordning är tidsstämpel, lag, händelse, underhändelse, zon. Ifall användaren stavat fel eller

undgått att anteckna någon kategori ber programmet användaren att mata in den felaktiva eller saknade datan. Det tar omkring tre minuter för användaren att föra in den saknade informationen från en halvleks datainsamling via `clean_csv`. I figur 2 ses terminalen vid användning och CSV-filen som generas.

Figur 2: `clean_csv`: (a) Inmatningar (b) CSV-fil.

3.3 Ställningstaganden och problematik vid datainsamling

Nedan följer en genomgång av de problem och ställningstaganden som gjordes i samband med datainsamlingen.

3.3.1 Styrkor och problematik med insamlingsmetoden

De typer av händelser som eftersöktes bestämdes i samråd med tränarna, och utvecklades under projektets gång; en av fördelarna med det egenskapade inmatningsprogrammet var att händelser och underhändelser med lätthet kunde läggas till och tas bort. Det finns också möjlighet att utöka händelser och i efterhand lägga till dem matcher vars övrig data redan var insamlad.

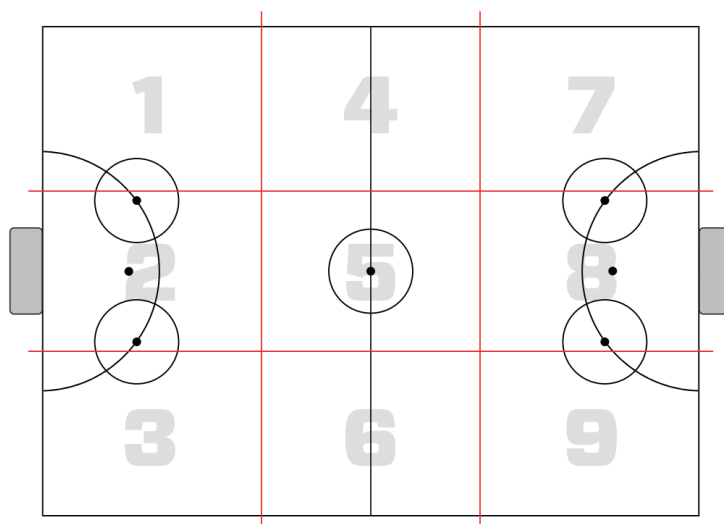
Ett fenomen som upptäcktes var att det var lättare att lägga till vissa typer av händelser än andra; händelser som sker då ett lag har besittning av bollen var generellt lättare att anteckna än händelser som sker i någon form av kampsituation. Detta exemplifieras av händelserna **skotttyp** (varje skottförsök kategoriseras efter hur det kommit till) och **40** (Sirius spelar hem bollen mot sin libero Sune Gustafsson); **skotttyp** sker alltid efter ett skott, och är en händelse som vanligen inmatas under tidspress då bollen ofta fortfarande är i spel, medan **40** nästintill alltid sker i ett läge när laget har bollen relativt obestritt. Det finns därför utrymme att addera händelser likt **40**, men inte fullt lika många som liknar **skotttyp**.

En begränsning med insamlingsmetoden är att det hela tiden eftersöks avslutande händelser; systemet bygger på att varje händelse är avslutad när den antecknas. Detta är dels för att inmatningsprogrammet är byggt så, men även för att underlätta för datainsamlaren. Processen kräver såpass stort fokus att användaren inte ska behöva ha några tankar i huvudet alls när den gör datainsamlingen. Ett alternativ hade varit att dela upp vissa händelser i start- och sluthändelser så att användaren antecknar när något börjar hända, och när detta slutat, men för att underlätta för inmataren valdes detta alternativ bort.

En sammanställd lista på händelser, underhändelser och deras definitioner finns i README.md-filen i följande GitHub.

3.3.2 Planens zonindelning

Spelplanen delades in i 3×3 zoner enligt figur 3 där Sirius anfaller mot zon 8.



Figur 3: Bandyplanen indelad i nio zoner.

Det kan finnas en tendens vid datainsamlingen att se zoner 2 och 8 som straffområdes-zoner, och därmed överskatta deras bredd vilket innebär att de inskränker på zoner 1 och 3, respektive 7 och 9. Det finns även en tendens att se zoner 4, 5 och 6 som mittzoner, och därför överskatta deras storlek gentemot zoner 1, 2, 3 och 7, 8, 9. Dessa tendenser bör tas i beaktning när zon-statistiken analyseras. Diskussioner fördes kring att omdefiniera ovanstående zoner för att återspegla detta, men valet föll på att inte göra det eftersom den mer raka indelningen med nio lika stora zoner ansågs vara mer lättförståelig för spelarna.

3.3.3 Avslut

Rörande avsluten eftersöks utfallet av varje skott; om det blir mål, skottet räddas, går utanför eller täcks av motståndarspelare. Detta gjordes genom att skapa ett underhändelse till skott som är dess utfall.

Vilken typ av skott som skjutits är även av intresse, samt hur skottet skapades. Från den övriga händelsedatan är det i vissa fall möjligt utläsa hur skottet skapades, exempelvis har data samlats rörande vissa typer av passningar, och alla fasta situationer - men dataunderlaget är långt ifrån så sofistikerat att det är möjligt att endast genom övrig data bedöma vilken typ av skott som skjutits för alla avslut. Det beslutades därför att skapa en egen händelse som kallad **skotttyp**, som för sig själv beskriver vilken typ av skott som skjutits. Ambitionen var att i efterhand kunna analysera skotten utifrån dessa typer. Följande sju kategorier identifierades:

1. **Friställande passning:** en längre passning når fram till en anfallsspelare i ett läge där denna antingen är fri, eller i ett sådant läge att den snabbt kan ta sig fram till ett ohotat skottläge i straffområdet. Ett exempel är Ted Bergströms 4-1-mål den 3 mars 2022 hemma mot Edsbyn.
2. **Inläggspassning:** bollen passas in i straffområdet från endera kant. Ett exempel är Nils Bergströms 1-4-mål den 25 november 2021 hemma mot AIK.
3. **Skott utifrån:** ett skott från håll, antingen ifrån straffområdets utkant eller helt utanför det. Ett exempel är David Thoréns 1-0-mål den 8 februari borta mot Edsbyn.
4. **Individuell dibbling in i straffområdet:** enskild spelare för bollen in i straffområdet och skapar skottläge. Ofta tydligt stark individuell prestation. Ett exempel är Jerker Ortmans 2-0-mål den 3 mars 2022 hemma mot Edsbyn.
5. **Centralt anfall med flera spelare:** ett anfall där flera spelare (ofta två eller tre) driver in bollen centralt i straffområde, ofta genom passningsspel såväl i som utanför straffområdet. Ett exempel är Matteus Liws 6-2-mål 28 oktober 2022 borta mot Vänersborg.
6. **Fast situation:** till exempel hörnor, frislage och straffar. Ett exempel är Kalle Mårtenssons 2-2-mål den 25 november 2022 hemma mot Gripens/Trollhättan.
7. **Retur:** skottet kommer som en direkt följd av ett annat skott. Ett exempel är Nils Bergströms 2-5-mål den 2 december hemma mot Sandviken.

Denna kategorisering kan förefalla godtycklig, vilket den i viss mån är; kategorierna är framtagna tillsammans med tränarna, och indelningen hade kunnat göras på andra sätt. Då dataunderlagets storlek är begränsat sötes antalet kategorier

begränsas för att varje skotttyps dataunderlag skulle bli nog stort.

Detta ställdes mot det faktum att kategorierna inte kunde vara för breda och därmed intetsägande. Den enda av ovanstående sju kategorier som kan upplevas vara för bred är **centralt anfall**. Vid en omkategorisering hade den skotttypen kunnat ses över. Indelningen är dock såpass robust att det under datainsamlingen bara var ett fåtal tillfällen då det rädde oklarheter om hur ett avslut skulle klassificeras.

3.3.4 Närkamper

Närkamp-noteringsarbetet var av stor vikt för Sirius. Tidigt identifierades tre typer av närkampssituationer:

1. **Närkamp**: en fysisk närkampssituation uppstår mellan den bollförande spelaren och minst en spelare i motståndarlaget. Syftet är att komma ur situationen med bollen under kontroll.
2. **Brytning**: en spelare lyckas med hjälp av sin positionering bryta en motståndarpassning och ta besittning av bollen.
3. **Bolltapp**: en spelare förlorar kontroll över bollen på grund av ett oprovcerat dåligt ingripande. Denna kategori har två underkategorier:
 - (a) **Passningsförsök**: spelaren försöker slå en uppenbart dålig passning som går till motståndare.
 - (b) **Tappad**: spelaren tappar bollen när han driver den, exempelvis genom en dålig dribbling.

Ett problem kan vara att det i vissa fall är oklart ifall ett lag fått bollen genom en **närkamp** eller en **brytning**; speciellt när flera defensiva spelare varit inblandade i bollerövringen. Detta problem hanterades genom att lägga ihop dessa två händelser i en samlingshändelse kallad **duell**.

Närkamper är möjligen den mest subjektiva datatypen som insamlas. Detta eftersom insamlaren vid varje potentiell närkampssituation måste göra en subjektiv bedömning av ifall det inträffade var en närkamp. Detta är ett väldokumenterat problem från andra sporter. I NHL är statistiken rörande tacklingar erkänt inkonsekvent; i de tre säsongerna 2018, 2019 och 2020 noterades det på Pittsburgh Penguins hemmamatcher 16,6 fler tacklingar än vid deras bortamatcher. Det finns även exempel på när hockeyjournalister uppmärksammat en stor diskrepans vad gäller brytningar (engelska: "takeaways") mellan den officiella statistiken från NHL och den de själva noterat [21] [43] [16].

I denna rapports fall är det bara en person som noterar närkamperna, så vad som

anses vara en närkamp bör vara relativt konsekvent. Den subjektiva bedömningen av vad en närkamp är bör uppmärksammas då närkampsstatistiken analyseras.

3.3.5 Videokvalitet på TV-sändningarna

Kvaliteten på de TV-sända matcherna är i vissa fall bristfällig. Allt som oftast är matcherna bara filmade med en kamera placerad vid mittlinjen. Detta, tillsammans med det faktum att bandybollen är notoriskt liten och svår att se, gör att det ibland varit svårt att exakt urskilja vad som hänt i straffområden.

Det har varit speciellt svårt att se ifall hörnskott täckts ut av rusande spelare, eller ifall skottet missat målet helt, detta kan heller inte härledas av övriga händelser eftersom båda fallen leder till målvaktens boll. Det har i vissa fall även varit svårt att se ifall ett skott räddats ut av målvakten, eller täckts ut av en försvarsspelare. Denna problematik påverkar skott på mål-statistikens pålitlighet, men rör inte corsi-händelsestatistiken.

Samma problem har uppstått vad gäller närkampssituationer vid bortre hörflaggor (zoner 3 och 9). Detta blir extra påtagligt på utomhusarenor när det är mörkt och/eller mycket nederbörd. I vissa inomhushallar kan kamerans placering göra det svårt att se exakt vad som händer längs sargen i mittzon (zon 4).

Sammantaget har det fungerat bra att följa matcherna via TV, men vissa, om än smärre, problem har det funnits.

4 Resultat och analys

Följande avsnitt avhandlar arbetet som gjordes för att sammanställa statistik utifrån den insamlade datan, samt resultatet av denna detta arbete. Lejonparten av den kod som används i dessa avsnitt finns i filerna `get_stats.py` och `compile_stats.py`. Efter detta följer en genomgång av hypoteserna presenterade i avsnitt 2.3. Slutligen avhandlas processen som varit för att skapa match- och säsongsrapporter till Sirius bandy. Den kod som avser presentations-avsnittet finns i `get_pp.py` och `get_plot.py`.

4.1 Dataunderlag och Sirius kräftgång under hösten 2022

Datainsamlingen är begränsad till uteslutande Siriusmatcher; alla matcher i grundserien 2022/23 (fram till 30 december, 15 matcher) och Svenska Cupen 2022 (tre matcher). Under hösten 2022 har resultaten inte varit med Sirius, då man av sina 18 matcher endast vunnit fem matcher, spelat oavgjort i en och förlorat tolv. Noterbart är dock att laget vann sina tre sista matcher under 2022. Statistiken nedan inkluderar alltid data från alla 18 matcher såvida inget annat anges.

Denna resultatmässiga snedfördellning skapar två problem. Dels, ifall Sirius under denna period underpresterat i det defensiva, offensiva eller i båda aspekter av spelet. Detta kan innebära att viss typ av statistik blir svårtolkad. Ett annat problem blir att den bollförande och återerövrings-inriktade spelstil som Sirius har kan framstå i dålig dager; förlorar Sirius på grund av sin spelstil, eller är det andra faktorer som gör att de förlorar?

Dessa frågor är svåra att besvara enbart med denna data, men hänsyn bör tas till dem när statistiken analyseras.

Däremot kan denna resultatmässiga snedvridning skapa en möjlighet till att förklara varför Sirius förlorat, och att med hjälp av statistiken peka ut specifika moment i spelet där Sirius varit sämre än sina motståndare.

4.2 Övergripande statistik

Olika typer av räknande statistik togs fram; antal mål, antal skott på mål och skottförsök samt antal närkamper och brytningar. En bollinnehavs-beräknande funktion skapades som utifrån att klassificera alla händelser i matchdatan som antingen innehavshändelser (såsom passning eller närkampsvinst), bollförlust-händelser (bolltapp eller offside), eller paushändelser (timeout eller mål). Programmet summerar sedan ihop tiden mellan de olika händelsetyperna och producerar båda lagens bollinnehav.

4.2.1 Resultat - övergripande

I tabell 2 finns övergripande statistik för Sirius respektive deras motståndare, både vad avseende faktiska antal och procentuell fördelning. Notera att skott i detta fall avser skott på mål.

Tabell 2: Övergripande statistik för Sirius och deras motståndare.

	Mål	Skott	Corsi	Dueller	Innehav
Antal					
Sirius	54	171	464	941	11:46:33
Motståndare	89	237	502	861	10:28:02
Procentfördelning					
Sirius	38%	42%	48%	52%	53%
Motståndare	62%	58%	52%	48%	47%

4.2.2 Analys - övergripande

Enligt den övergripande statistiken verkar Sirius vara starkare än motståndarna i såväl närkampsspelet som bollinnehavet. Däremot skjuter man lite färre skott, och har genomgående en lägre effektivitet. Motståndarna träffar målet vid 47,2 % av sina corsihändelser, medan Sirius bara gör det vid 36,9 % av sina. Sirius sedan mål på 31,6 % av sina skott på mål, medan motståndarna gör mål på 37,6 % av sina. Vad gäller procentfördelningen är Sirius sämre allt eftersom skotten kommer närmare mål; de har 48 % av corsihändelserna, 42 % skott på mål och 38 % mål.

4.3 Avslut och mål

Skottstatistiken delades upp per skotttyp. I tabell 3 finns skottsstatistik för de sju avslutstyperna definierade i avsnitt 3.3.3 gällande Sirius och i tabell 4 för motståndarna.

Då passningen som föranleder skottypen **inlägg** finns sparad togs statistik fram för att visa hur stor andel av dessa som lett till ett avslut, denna information finns i tabell 5 för båda lagen.

4.3.1 Resultat - avslut

Resultaten syns nedan. Notera att skottypen **fast** inkluderar skott på hörna, frislåg och straff, men inte returskott därifrån. En mer ingående sammanfattning av fasta situationer kommer senare i rapporten. Som skott används corsi-händelser, detta eftersom denna statistik är mer tillförlitlig enligt vad som skrivits i 3.3.5.

Skotttyp	Mål	Skott	Skottprocent	Skotttyp	Mål	Skott	Skottprocent
Fast	11	119	9,2 %	Fast	24	149	16,1 %
Centralt	10	56	17,9 %	Centralt	20	76	26,3 %
Utifrån	12	199	6,0 %	Utifrån	10	171	5,8 %
Inlägg	7	39	17,9 %	Inlägg	13	32	40,6 %
Friställande	6	16	37,5 %	Friställande	10	18	55,6 %
Retur	6	25	24,0 %	Retur	7	27	25,9 %
Dribbling	2	10	20,0 %	Dribbling	5	29	17,2 %
Totalt	54	464	11,6 %	Totalt	89	502	17,7 %

Tabell 3: Skottstatistik Sirius.**Tabell 4:** Skottstatistik motståndare.**Tabell 5:** Inläggspassningar för Sirius och deras motståndare.

	Mål/inläggsskott	%	Skott/inlägg	%
Sirius	7/39	17,9 %	39/201	19,4 %
Motståndare	13/32	40,6 %	32/158	20,3 %

4.3.2 Analys - avslut

Det är tydligt att Sirius i stora drag är betydligt mer ineffektiva än motståndarna vad gäller avsluten; med en total skottprocent motsvarande motståndarnas på 17,7 % hade Sirius gjort ytterligare 28 mål på sina 464 avslut, och därmed haft en målskillnad på -7 istället för -35. Vid närmare anblick kan syns att effektiviteten främst skiljer sig på skotttyperna **fast**, **centralt**, **inlägg** och **friställande**.

Resultaten från tabell 3 och tabell 4 kan användas som en primitiv förväntade mål-modell likt det som beskrivits i 2.2.4, där varje skotttyps förväntade mål-värde är den skottprocenten som anges i tabellen.

Vad gäller inläggsskotten visar tabell 5 att Sirius främsta problem inte varit att komma till inläggs-avslut; avslutsfrekvensen på inläggspassningar skiljer sig inte nämnvärt från motståndarna - däremot gör effektiviteten på avsluten det. Om sedan motståndarnas mer än dubbelt så höga effektivitet beror på ett poröst försvarsspel eller att Sirius haft låg effektivitet i avsluten är inte klart.

Sirius har även en betydligt lägre målfrekvens på frilägen. Noterbart här är denna statistik varit än mer snedfördelad under en stor del av säsongen; Sirius avslutade som bekant året med en tre matcher lång segersvit och gjorde då alla sina sex frilägesmål på sju frilägen, under denna period släppte man inte in ett enda frilägesmål på två frilägen. Detta innebär att Sirius under stor del av säsongen haft en frilägeseffektivitet på noll medan motståndarna gjort mål på mer än hälften av

sina lägen.

4.4 Fasta situationer

Den offentliga hörnstatistiken som finns att tillgå i bandy skiljer inte på höger- eller vänsterhörnor. Enligt tränarna är det dessutom olika ifall hörnstatistiken inkluderar mål gjorda på hörnretur, vilket gör den opålitlig. Vid insamlingen anger användaren från vilken zon hörnpassningen slås. Denna sammanställdes, och returmål inkluderas. Alla mål delades in i fyra typer beroende på spelform: spel-, hörn-, straff- och frislagsmål. Svenska Bandyförbundet för även statistik för självmål, men då dessa är så sällan förekommande (under hela Elitseriesäsongen 2021/22, inklusive kval och slutspel har förbundet bara noterat tre självmål) beslutades de undantas [4] [3].

4.4.1 Reslutat - fasta situationer

Hörnstatistik finns i tabell 6, mål per spelform finns i tabell 7.

Tabell 6: Hörnstatistik för Sirius och deras motståndare per sida.

	Vänster	Utdelning	Höger	Utdelning	Båda	Utdelning
Sirius	10/73	13,7 %	3/56	5,4 %	13/129	10,1 %
Motståndare	15/76	19,7 %	9/78	11,5 %	24/154	16,2 %
Totalt	25/149	16,8 %	12/134	9,0 %	38/283	13,4 %

Tabell 7: Mål per spelform för Sirius och deras motståndare.

	Spelmål	Hörnmål	Straffmål	Frislagsmål
Sirius	41	13	0	0
Motståndare	62	25	2	0
Totalt	103	38	2	0

4.4.2 Analys - fasta situationer

Hörnstatistiken visar att Sirius utdelning varit betydligt sämre än motståndarnas, man verkar även haft betydligt svårare att tillskansa sig hörnor från höger. Sirius släpper in mål på hörnor i det absolut översta skicket för vad som anses normalt, och gör mål med en relativt låg frekvens.

Ungefär vart fjärde bandymål verkar göras på hörna, och vad gäller mål i andra typer av fasta situationer är dessa nästintill obefintliga.

Noterbart är att det allmänt verkar vara svårare att göra mål på hörnor från höger, vilket kom som en överraskning för såväl spelare som tränare. Appliceras Z-testet enligt ekvation 2 med nollhypotesen $H_0 : p_{\text{höger}} = p_{\text{vänster}}$ för vardera lag blir resultatet $z_{\text{Sirius}} = 1,56$ och $z_{\text{Motståndare}} = 1,40$. Båda är klart mindre än $z_{\alpha/2} = 1,96$ som gäller för konfidensintervall på 95 %, $p_{\text{Sirius}} = 0,059$, $p_{\text{Motståndare}} = 0,081$.

Skillnaden i effektivitet mellan höger- och vänsterhörnor är därmed inte statistisk säkerställd. Det kan däremot vara av intresse att följa upp denna statistik framöver och se ifall skillnaden kvarstår.

4.5 Närkamper och dueller

Eftersom brytning ansågs vara en underkategori till närkamp (skillnaden mellan en brytning och en närkamp är avståndet till motståndaren), fattades beslutet att lägga ihop dem i kategorin **duell**. Framöver är alla resonemang rörande närkamper applicerbara på dueller. I tabeller och figurer anges ifall den presenterade datan rör dueller eller närkamper.

Närkamper är en kategori där man vid analyseringen bör vara uppmärksam de resonemang avhandlade i avsnitt 2.2.1 och 2.2.2. Att räkna antal vunna närkamper som något enbart positivt kan potentiellt leda till att en övervärdering av insatser som egentligen inte leder till framgångsrikt spel, då det faktum att ett lag har många vunna närkamper kan tyda på att de sällan har bollen. Med detta sagt är närkampsspelet något som Sirius lagt stor vikt vid, och därför en aspekt som mycket tid lagts på att analysera.

För att skapa en tydligare bild av närkampernas inverkan på matchbilden delades närkamperna upp utifrån vilket lag som haft bollinnehav inför närkampen; motivationen var att se hur många närkamper varje lag initierade, samt resultatet av dessa.

En ytterligare åtgärd som vidtagits för att kontextualisera närkampsdatan är att skilja på i vilken zon de skett.

4.5.1 Resultat - närkamper

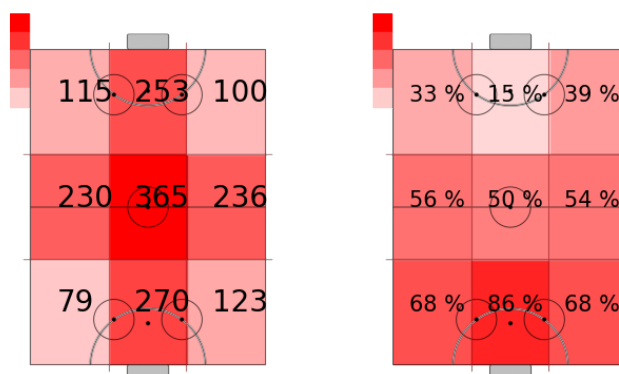
I tabell 8 finns antal vunna närkamper, lyckade brytningar och bolltapp för Sirius och motståndarna, både vad gäller antal och procentuell fördelning. Statistik beroende på innehav före finns i tabell 9. I figur 4 ses alla närkamper och brytningar per zon och i figur 5 andelen vunna närkamper och brytningar för Sirius per zon, Sirius anfaller uppåt i figurerna.

Tabell 8: Närkampsstatistik för Sirius och deras motståndare per typ.

	Närkamper	Brytningar	Bolltapp
Antal			
Sirius	802	139	94
Motståndare	705	156	47
Procentfördelning			
Sirius	53 %	47 %	67 %
Motståndare	47 %	53 %	33 %

Tabell 9: Tabell för duellstatistik baserat på lags innehav före och efter.

	Efter Sirius	Efter motståndare
Före Sirius	223	721
Före motståndare	718	140

**Figur 4:** Duellernas fördelning per zon. **Figur 5:** Andel vunna dueller Sirius.

4.5.2 Analys - närkamper

Sirius har genomgående varit starka i duellspelet samtidigt som man haft många bolltapp; däremot har upplevelsen under säsongen varit att de råa duellsiffrorna inte beskriver matchbilden särskilt bra. Tränarna har istället ansett att den zonspecifika datan från figurer 4 och 5 bäst beskriver spelet; lag som vunnit en stor del av närkamper och brytningar i anfallande straffområde har upplevts starka. Ett högt antal närkamper i offensiv zon, oavsett om vilket lag som vunnit dem, har också rimmat med att dominera matchbilden.

Närkampsstatistiken uppdelad efter innehav från tabell 9 har varit svårtolkad. Noterbart är att det bollförande laget vid de noterade duellerna förlorar bollen i omkring fyra av fem fall. Detta gör att argumentationen från avsnitt 2.2.1 går att tillämpa; att som lag vinna många närkamper tyder på att man ofta inte har bollen. Detta kan ha flera förklaringar; dels att det i närkampssituationer ofta är flera angripande spelare mot en bollförare, och att det som anfallsspelare kan vara svårt att ge understöd till den bollförande lagkamraten. En annan förklaring kan vara att närkamper i bandy först uppstår då den bollförande spelaren satt sig i en sårbar situation. Ett tredje alternativ handlar om den subjektiva bedömningen av vad en närkamp är som datainsamlaren gör; vissa gränsfallssituationer kan bedömas vara en närkamp endast ifall bollen byter ägare.

Vad som kan utläsas av denna statistik är att Sirius verkar varit betydligt bättre på att behålla bollen vid en närkampssituation, något som intuitivt känns positivt, men som inte nödvändigtvis är det. Statistiken kan också tolkas som ett mått på hur många närkamper varje lag initierat; genom att summera rad för rad i tabell 9 blir det klart att Sirius initierat 858 närkamper mot 944 för motståndarna. En argumentation kan föras för att det är bra att initiera närkamper, då det innebär en stress för motståndarna. Det är även möjligt att föra ett resonemang som bygger på att närkampssituationer ofta utlöses av det bollförande laget; det kommer uppstå fler närkampssituationer om ett lag driver bollen i offensivt straffområde jämfört med ifall de gör samma sak i defensivt straffområde. Med detta resonemang bör de högre siffrorna för motståndarnas initierade närkamper snarare ses som en fördel för Sirius, då det skulle tyda på att Sirius genom sitt spel skapat farliga målchanser. Detta resonemang liknar det i avsnitt 2.2.2.

Inläggspassnings-statistiken från tabell 5 kan användas som ett mått på lagens duellspel i eget straffområde; att hindra motståndarna från att komma till avslut vid en inläggspassning innebär vanligen att det försvarande lagets spelare i någon mån vinner en närkamp. Sirius och motståndarna tillåter skott på ungefär var femte inläggspassning; däremot gör motståndarna mål på sina inläggsskott mer än dubbelt så ofta som Sirius. Detta kan tolkas både som en kritik mot Sirius defensiva insatser i eget straffområde då man tillåter farliga inläggsskott, och som att Sirius målvakter är svagare på inläggsskott än motståndarna.

Generellt är närkampsspelet svårt att tyda; här får analysen ta hänsyn till olika motstående koncept, och utrymme finns för att komma till ytterligare insikter och skapa en djupare förståelse.

4.6 Hypoteser

Nedan följer test av hypoteserna från 2.3.

4.6.1 Att som lag vinna en stor del av närkamperna bidrar inte till att vinna matcher

Som diskuterats ovan visade sig närkamperna vara en svåranalyserad del av bandy. Det är dessutom inte självklart vad vinna matcher innebär i den här kontexten; jämförs bara matchens utfall (vinst eller förlust) innebär det att en 7-0-vinst likställs med en 5-4-vinst. Väljs istället målandel som parameter att jämföra mot likställs samma 7-0-match istället med en 1-0-seger. Det faktum att Sirius som lägger så stort fokus vid närkampsspelet, haft en så lång formsvacka under hösten medför även problem; Sirius presterar genomgående starka siffror vad gäller dueller, men har trots detta förlorat majoriteten av sina matcher.

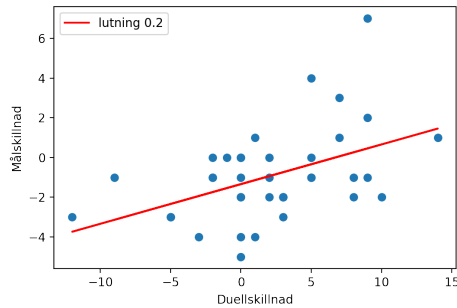
Det lag som gjort flest mål per halvlek har i snitt vunnit 51 % procent av duellerna, och 0,9 fler dueller än motståndarna.

Tabell 10: Lags resultat vid andel vunna dueller.

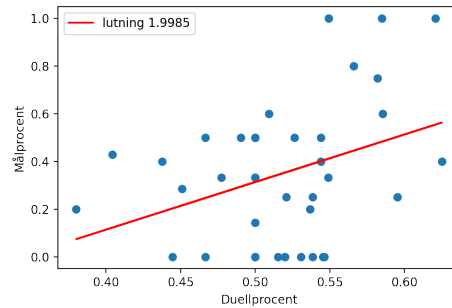
	Vinst	Förlust	Vinstprocent
>50 %			
Sirius	7	22	24 %
Motståndare	7	9	44 %
Totalt	14	31	32 %
>55 %			
Sirius	6	4	60 %
Motståndare	3	0	100 %
Totalt	9	4	69 %

I tabell 10 återfinns varje lags resultat halvleksvis vid olika andelar vunna dueller. I figur 6 syns Sirius duellskillnad mot målskillnad och i figur 7 finns Sirius duellprocent mot målprocent. Vid en överblick av resultaten i tabell 10 syns inget positivt samband mellan resultat och att vinna minst hälften av duellerna, snarare ett negativt samband. Dras istället gränsen vid 55 % av duellerna så blir dataunderlaget förvisso relativt litet, men laget som drar längsta sträcket vid närkamper gör också flest mål mer än dubbelt så ofta som inte.

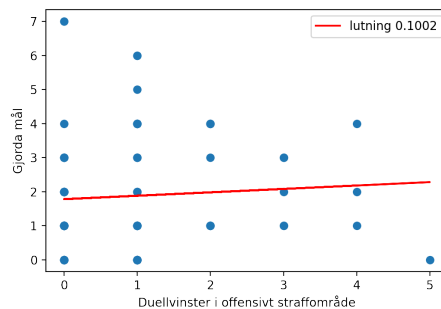
Korrelationskoefficienten mellan duellskillnaden och målskillnaden är $r = 0,458$. Detta tyder på en svag men inte obefintlig positiv korrelation. Figur 6 indikerar även det; trendlinjen är positiv. Jämförs istället duellprocenten med målprocenten



Figur 6: Duellskillnad mot målskillnad med trendlinje.



Figur 7: Duellprocent mot målprocent med trendlinje.



Figur 8: Duellvinster i offensivt straffområde mot gjorda mål med trendlinje.

är korrelationskoefficienten $r = 0.364$, och trendlinjen i figur 7 (notera att lutningen på trendlinjen är ovanligt hög eftersom spannet i y-led är betydligt större än på övriga).

För att vidare undersöka duellernas inverkan på resultatet appliceras nollhypotesen-test med linjär regression enligt avsnitt 2.4.3. När nollhypotesen mellan duellskillnad och målskillnad testas resulterar detta i ett t -värde på 3,093 vilket medför ett p -värde på 0,004. Nollhypotesen mellan duellskillnad och målskillnad kan således förkastas. När nollhypotesen testas för duellprocent och målprocent resulterar detta i ett t -värde om 2,346 vilket medför ett p -värde på 0,025. Nollhypotesen mellan duellprocent och målprocent kan därmed förkastas.

En teori är att duellvinster i offensivt straffområde (zon 8 för Sirius och zon 2 för motståndarna) bidrar till att göra många mål. Motiveringen är att duellvinster där direkt leder till farliga lägen. I figur 8 finns duellvinster i offensivt straffområde mot gjorda mål för både Sirius och motståndare mot gjorda mål med trendlinje. Korrelationskoefficienten mellan dessa två är $r = 0,071$, spridningen förefaller

vara slumpartad i figuren och trendlinjen är nästintill helt plan. Något samband mellan samband mellan antal duellvinster i offensivt straffområde och gjorda mål kan inte identifieras.

Sammantaget kan finns ett samband mellan att vinna många dueller och att vinna bandymatcher i dataunderlaget.

4.6.2 Att som lag ha en hög corsiprocent bidrar inte till att vinna matcher

För att kringgå problemet med definition av vunna matcher som presenterats ovan jämförs skottförsök med såväl antal gjorda mål som målfördelning. För att kontextualisera skottförsökens betydelse görs alla tester på skottförsök även för skott på mål.

Det lag som gjort flest mål per halvlek har i snitt haft 57 % av corsihändelserna och 3,1 corsihändelse mer än motståndarna. För skott på mål gäller 63 % och 2,8 fler skott på mål.

I tabell 11 återfinns lagens resultat vid olika corsiprocent och i tabell 12 finns lagens resultat vid olika skott på mål-procent. I figur 9 ses corsihändelser för mot gjorda mål och i figur 10 corsiprocent mot målprocent för båda lagen. I figur 11 ses skott på mål mot gjorda mål och i figur 12 skott på målprocent mot målprocent.

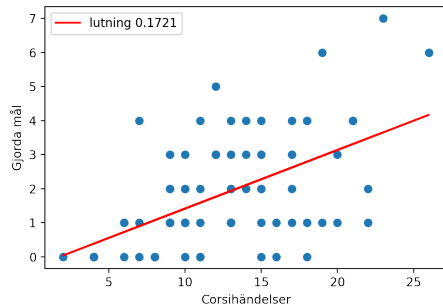
	Vinst	Förlust	V%
>50 %			
Sirius	5	5	50 %
Motståndare	14	1	93 %
Totalt	19	6	76 %
>55 %			
Sirius	4	3	57 %
Motståndare	7	0	100 %
Totalt	11	3	79 %

	Vinst	Förlust	V%
>50 %			
Sirius	6	2	75 %
Motståndare	22	1	96 %
Totalt	28	3	90 %
>55 %			
Sirius	4	1	80 %
Motståndare	17	1	94 %
Totalt	21	2	91 %

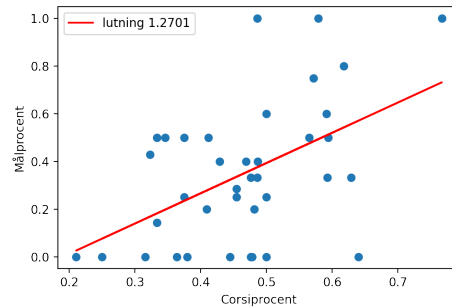
Tabell 11: Lags resultat vid corsiprocent.

Tabell 12: Lags resultat vid skott på mål-procent.

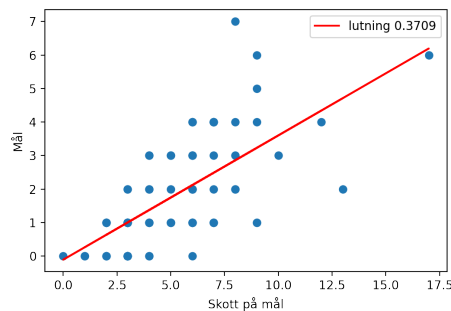
Från tabell 11 kan utläsas att ett övertag i corsihändelser förefaller ha ett samband med att göra flest mål. Då corsiovertaget ökar ökar även vinstprocenten, om än fast dataunderlaget är litet. Om denna data jämförs med resultaten vid övertag i skott på mål-statistiken från tabell 12 kan det konstateras att skott på mål förefaller vara en otroligt stark indikator; totalt sett vinner laget med flest skott på mål 90 %



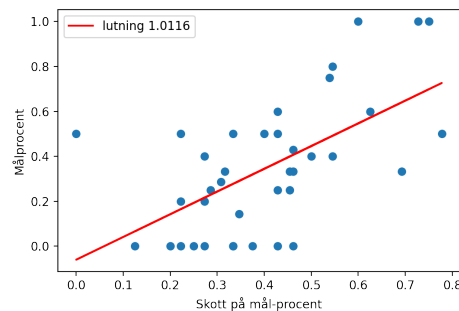
Figur 9: Corsihändelser mot mål med trendlinje.



Figur 10: Corsiprocent mot målprocent med trendlinje.



Figur 11: Skott på mål mot mål med trendlinje



Figur 12: Skott på målprocent mot målprocent med trendlinje.

av halvlekarna. Korrelationskoefficienten mellan corsihändelser och gjorda mål är $r = 0,485$ och skott på mål och gjorda mål är $r = 0,602$.

För att undersöka corsihändelsernas inverkan på resultatet appliceras nollhypotesen-test med linjär regression enligt avsnitt 2.4.3. När nollhypotesen mellan corsihändelser och gjorda mål testas resulterar detta i ett t -värde på 5,286 vilket medför ett p -värde på 0,000. Nollhypotesen mellan corsihändelser och gjorda mål kan således förkastas. När nollhypotesen testas för corsiprocent och målprocent resulterar detta i ett t -värde om 3,467 vilket medför ett p -värde på 0,001. Nollhypotesen mellan corsiprocent och målprocent kan därmed förkastas.

Dessa test genomförs även för skott på mål; nollhypotesen mellan skott på mål och gjorda mål resulterar i ett t -värde om 7,544, vilket medför ett p -värde på 0,000. Nollhypotesen mellan skott på målprocent och målprocent har ett t -värde på 4,446 och ett p -värde på 0,000.

Sammantaget är corsihändelser en resultatsindikator i vårt dataunderlag, men skott på mål är en starkare indikator. Detta är dock förväntat utfall, då mål är en delmängd av skott på mål, som i sin tur är en delmängd av corsihändelser. Detta kan uttryckas som $S_{\text{mål}} \subset S_{\text{skott}} \subset S_{\text{corsi}}$. Rapporten har i avsnitt 3.3.5 avhandlat fördelarna med corsihändelser gentemot skott på mål för detta dataunderlag.

4.6.3 Att som lag ha ett högt bollinnehav bidrar inte till att vinna matcher

Vid analysen av bollinnehav lyfts återigen frågeställningen kring det diffusa konceptet att vinna matcher. Av denna anledning provas bollinnehavet mot såväl målskillnad som målprocent.

Det lag som gjort flest mål per halvlek har i snitt haft 49 % av bollinnehavet.

I tabell 13 ses varje lags resultat halvleksvis vid olika andelar bollinnehav. I figurer 13 och 14 återfinns bollinnehavsprocent jämfört med målskillnad halvleksvis och bollinnehavsprocent jämfört med målprocent halvleksvis.

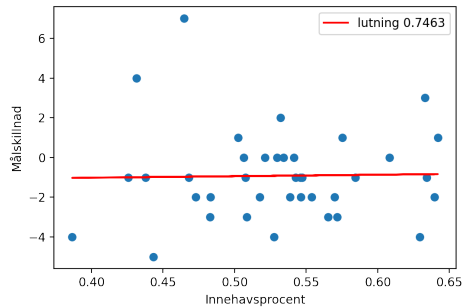
Tabell 13: Lags resultat vid andel bollinnehav.

	Vinst	Förlust	Vinstprocent
>50 %			
Sirius	4	12	25 %
Motståndare	7	2	78 %
Totalt	11	14	44 %
>55 %			
Sirius	2	5	29 %
Motståndare	3	1	75 %
Totalt	5	6	45 %

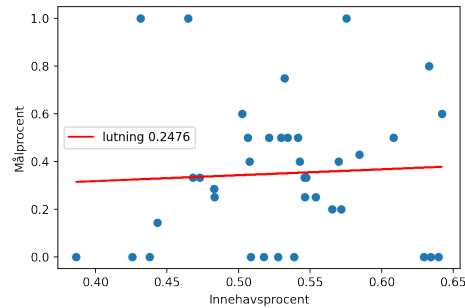
Vid en överblick av tabell 13 syns inget samband mellan bollinnehav och resultat, varken vid 50 % bollinnehav eller 55 %, lagen har nästintill identisk vinstprocent i båda fallen.

Korrelationskoefficienten för innehavsprocent mot målskillnad är $r = 0,020$ och för innehavsprocent mot målprocent $r = 0,090$. Spridningen i figurer 13 och 14 är förskjuten höger i x-led, då Sirius generellt har ett bollinnehavs-övertag, men det förefaller inte finnas något samband mellan x- och y-värden då båda trendlinjerna är nästintill helt flacka.

När nollhypotes-test med linjär regression enligt avsnitt 2.4.3 mellan bollinnehavsprocent och målskillnad appliceras resulterar detta i ett t -värde på 0,123 vilket



Figur 13: Innehavsprocent mot målskillnad med trendlinje.



Figur 14: Innehavsprocent mot målprocent med trendlinje.

medför ett p -värde på 0,903. Nollhypotesen mellan innehavsprocent och målskillnad kan således inte förkastas. När nollhypotesen testas för innehavsprocent och målprocent resulterar detta i ett t -värde om 0,318 vilket medför ett p -värde på 0,752. Nollhypotesen mellan innehavsprocent och målprocent kan därmed inte förkastas.

Resultatet blir att rapporten inte kan förkasta hypotesen att bollinnehav inte påverkar matchens resultat.

4.6.4 Skott utifrån har inte en lägre sannolikhet att bli mål än skott närmare från mål

Enligt tabeller 3 och 4 är skott utifrån den avslutstypen med lägst målprocent. Skotttypen fast, vars absoluta majoritet är hörnskott skjutna från straffområdets ytterkant, är näst lägst. Detta gör att rapporten inte kan fastslå att målfrekvensen för skott inte minskar med ökat avstånd från målet.

4.6.5 Vid spel utomhus och på dålig is slår lagen inte fler långa bollar

I tabell 14 finns antal långa passningar per match för ute- respektive inommatcher för varje lag för Sirius åtta utomhus- och tio inomhusmatcher.

Tabell 14: Långa passningar per match inomhus respektive utomhus.

	Utomhus	Inomhus
Sirius	14,4	5,3
Motståndare	16,1	8,9
Totalt	30,5	14,2

Vid två utomhusmatcher har spelarna upplevt isen som extra dålig; den 8 november mot Edsbyn och 19 november mot Västerås. Vid båda matcherna var det plusgrader och nederbörd, och vid matchen mot Västerås bedömdes isen så dålig att matchen spelades 3×30 minuter. I tabell 15 ses antal långa passningar per match för dessa två matcher, jämfört med resterande sex utomhusmatcher då isen bedömts vara bra.

Tabell 15: Långa passningar per match för bra respektive dålig is.

	Dålig is	Bra is
Sirius	21,5	12,0
Motståndare	16,5	16,0
Totalt	38,0	28,0

I detta dataunderlag slås det betydligt fler långa passningar utomhus än inomhus, även om bra iskvallitet utomhus jämförs med inomhusspel. Vid vad som kategoriserats som dålig iskvallitet slår lagen ännu fler långa passningar. Noterbart är att det till viss del verkar handla om en medveten strategi från lagen; Sirius slår färre långa passningar än sina motståndare såväl inomhus som på bra utomhusspel, medan de på dålig is slår fler. Detta kan tyda på att laget vid en viss iskvallitet ställer om hela sin offensiva strategi till att förlita sig mer på långa passningar, precis som tränarna sagt.

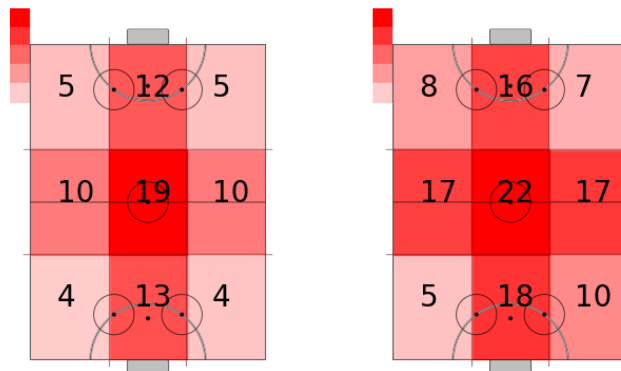
4.6.6 Vid spel utomhus och på dålig is placeras inte närkamperna mer i mittzonerna och framför mål

I figurer 15 och 16 finns snittet per match för antal dueller i planens nio zoner för inomhus- respektive utomhusspel. I figurer 17 och 18 ses snittet per match för antal dueller i planens nio zoner för utomhusspel vid bra respektive dålig is. I tabell 16 är en del av denna statistik sammaställd.

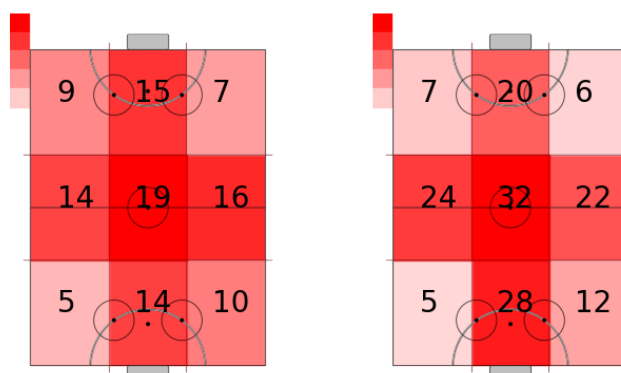
Tabell 16: Dueller per match och zon för olika istyper.

	Totalt	Hörnsoner	Mittsoner	Målsoner
Inomhus	82	22 %	48 %	30 %
Utomhus	120	25 %	47 %	28 %
Bra is	109	28 %	45 %	27 %
Dålig is	156	19%	50 %	31 %

Allmänt så är det totalt sett fler dueller vid spel utomhus än inomhus. Det är i sin tur fler närkamper vid dålig utomhusspel än vid bra utomhusspel.



Figur 15: Duellernas fördelning inne. **Figur 16:** Duellernas fördelning ute.



Figur 17: Duellernas fördelning utomhus vid bra is.

Figur 18: Duellernas fördelning utomhus vid dålig is.

I alla dessa fall är det markant färre närkamper i hörnzonerna (zoner 1, 3, 7 och 9) än i resterande zoner, de fyra hörnzonerna utgör 44 % av planens nio zoner, men i ovanstående data sker aldrig ens 30 % av duellerna i dem. Mittzonerna (zoner 4, 5 och 6) och målzonerna (zoner 2 och 8) är däremot kraftigt överrepresenterade vad gäller dueller för samtliga istyper. Denna snedfördelning kan till viss del förklaras av det som diskuterades i avsnitt 3.3.2, då en majoritet av duellerna sker i de zoner där en tendens att överskatta zonernas storlek påtalades.

Vad gäller hypotesen att duellernas placering skulle skilja sig markant vid spel inomhus och utomhus är detta inte självklart; duellernas placering vid spel inomhus och utomhus med dålig is är snarlika - med den skillnaden att det totalt sett är nästan dubbelt så många dueller i det senare fallet. Det som givet dataunderlaget kan konstateras är att utomhusspel, och framförallt utomhusspel på dålig is leder till fler närkampssituationer.

4.7 Presentationer till Sirius Bandy

En del av arbetets syfte är att hjälpa Sirius Bandy i sin verksamhet, och att presentera data för såväl tränare som spelare på ett pedagogiskt vis. Som nämndes i avsnitt 3.1 föll valet på en direktinsamlings-metod då detta skulle möjliggöra en presentation av statistik under matchers halvtidsvilor såväl som kort inpå dess slut.

Detta arbete mynnade ut i två typer av rapporter, en matchrapport med statistik från en halvlek eller match, och en säsongsrappport som presenterar statistik från en större mängd matcher. I GitHubens huvudmapp ligger två PowerPoint-presentationer; en av varje sort. Det tar ungefär fem sekunder att generera en matchrapport och tio att generera en säsongsrappport.

4.7.1 Matchrapport

Matchrapportens syfte är att på ett så lättläsligt vis som möjligt presentera statistik från matchen. Ambitionen var att tränarna skulle få underlag för samtal med spelarna i halvtid, och att den efter matchen skulle presentera en objektiv bild av spelet.

Detta krav på lättläslighet gjorde att all form av utsmyckning togs bort. Till en början användes lagens färger för presentera datan och presentationen hade en bakgrundsbild för att bidra till en bandyprofil; dessa togs bort och beslutet föll på vit bakgrund med röd text för Sirius (då det rimmade med de färgkodade skalorna i figur 4 och figur 5) och svart text för motståndarna. För att underlätta för läsaren presenteras alltid Sirius till vänster.

Rapporten innehåller närkampsstatistik likt de råa siffrorna från tabell 8, den zon-specifika statistiken från figur 4 och figur 5, vid denna presentation valde tränarna att de vunna närkamperna skulle presenteras i procent- istället för bråkform. Det var också viktigt för tränarna att Sirius alltid skulle anfalla uppåt i bild på dessa figurer, då de själva alltid ritar det på det viset. Närkampsstatistiken baserat på innehav från tabell 9 upplevdes som svårtolkad av såväl spelare som ledare.

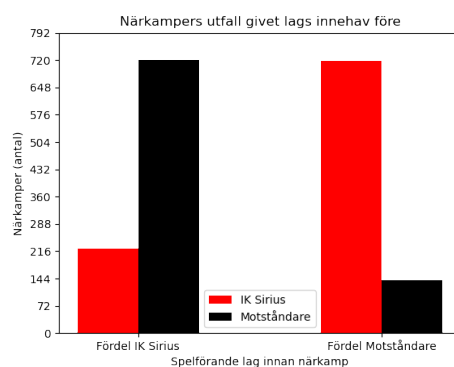
Hur denna data bäst bör visualiseras diskuterades ingående; till en början presenterades den enligt tabell 9. Inspirationen till detta var *confusion matrix*, som används för att presentera klassificeringsdata inom maskininlärning [30].

Flera spelare och ledare ansåg dock att denna presentation inte kändes intuitiv, och flera hade svårt att tolka den, även efter att de fått datan förklarad för sig. Andra typer av visuallisering diskuterades, bland annat punktform:

- Närkamper där Sirius haft bollen för och efter: X_1
- Närkamper där motståndaren haft bollen före och Sirius efter: X_2

- Närmkamper där motståndaren haft bollen för och efter: X_3
- Närmkamper där Sirius haft bollen före och motståndaren efter: X_4

Detta valdes bort då det upplevdes osammanhängande. Efter hand föreslog tränarna en presentation i stapeldiagrams-form enligt figur 19. Den överväldigande majoriteten av spelarna (delar av tränarstaben) föredrog denna presentation, men då vissa föredrog tabellen beslutade vi oss att behålla båda alternativen och alltid presentera dem sida vid sida.



Figur 19: Stapeldiagram för duellstatistik baserat på lags innehav före och efter.

Förutom ovan nämnd närmkampsstatistik finns även inläggsstatistik enligt tabell 5 inkluderad i matchrapporten.

Vid presentation av corsihändelser per skotttyp valde tränarna att skotttyperna skulle presenteras i bokstavsordning, och att skotttyper där ingen av lagen noterats för avslut skulle utelämnas.

För att skapa en mer nyanserad bild av lagens spelövertag över halvleken presenteras stapeldiagram med de båda lagens bollinnehav, skottförsök, vunna dueller och mål för varje $\frac{1}{N}$ del av halvleken (vanligen $N = 3$). Vidare skapades ett diagram över varje lags **långa anfall**, definierat som anfall där laget haft bollkontroll i minst 60 sekunder och inte förlorat bollen längre än tio sekunder åt gången, fördelade över halvleken. Under en period la Sirius stort fokus på sitt **40**-spel, så ett diagram med dessa händelsers spridning skapades även.

4.7.2 Säsongsrapport

Säsongsrapportens syfte är att vara ett analysverktyg för tränarstaben när ett flertal matcher ska summeras. Ambitionen är att den ska peka ut specifika delar av spelet där laget varit starka eller svaga samt att förklara resultaten.

Då både tränare och spelare vant sig vid matchrapporterna användes en liknande layout vid säsongsrapporten. Eftersom säsongsrapporten i regel summerar resultatet över flera matcher presenteras en majoritet av statistiken med tillhörande procentandels-fördelning.

I stort har säsongsrapporten många likheter med matchrapporten, det finns emellertid vissa skillnader: till skillnad mot i säsongsrapporten presenteras hörnstatistiken per sida, och statistiken per varje N -del av halvlekarna är utelämnade. Säsongsrapporten innehåller även skottstatistik-tabeller i likhet med tabell 3 och tabell 4.

5 Diskussion

Nedan följer en sammantagen diskussion om vad som tidigare presenterats i detta arbete.

5.1 Datainsamlingsmetoden

Det datainmatnings-verktyg som skapats möjliggör en stor datainsamling av en person, under en relativt liten tidsrymd. Detta verktyg har en viss inlärningskurva, men den är såpass liten att flera personer inom ett lag kan lära sig använda det och därefter samla data direkt under en match. På sikt skulle verktyget kunna utökas på ett sätt så data om fler parametrar samlas, och möjligheten finns för två insamlingar att ske parallellt, och sedan kombineras till ett större dataunderlag.

Verktyget kan även användas för att underlätta ytterligare datainsamling; möjligheten finns att nyttja det till att generera en tidsstämplad lista på när alla avslutar, för att sedan analysera dessa mer ingående.

Det finns utrymme att utveckla verktyget, lägga till, ta bort och omdefiniera händelser, men även i sin nuvarande form kan verktyget användas av olika bandylag.

5.2 Sirius resultat

En stor del av Sirius resultat kan förklaras med att de haft en låg skottprocent i jämförelse med motståndarna, framförallt vad gäller hörnor, frilägen, inlägg och centrala anfall. Man har förvisso haft lite färre avslut - men den (vad det verkar) avgörande skillnaden har legat i den sämre effektiviteten. Då det endast finns data från Sirius matcher är det omöjligt att med säkerhet avgöra ifall Sirius relativt sett låga skottprocent beror på att motståndarna lyckas göra mål i ovanligt hög utsträckning när de möter Sirius, eller ifall Sirius har låg utdelning. Corsihändelser förefaller vara en resultatindikator, men skott på mål har i detta dataunderlag varit en ännu bättre parameter. I båda dessa avseenden har Sirius varit svagare än sina motståndare.

De parametrarna där Sirius varit starka; duellspelet och bollinnehav har inte lika stor inverkan som skottförösk och skott på mål. Från vårt underlag verkar duellspelet först medföra resultat vid ett 55/45-övertag, och vad gäller bollinnehav har inget som helst samband mellan matchresultat identifierats.

5.3 Rapporternas påverkan på Sirius Bandy

Under säsongens gång har tränarstaben börjat förlita sig allt mer på matchrapporterna vid halvlekssamtal med spelarna. Vid årsskiftet var dataunderlaget stort nog för att skapa en bild av hur Sirius genomsnittliga motståndare presterar i de statistiska parametrar rapporten beskriver. För tränarna innebar det att man fick konkret data som satte prestationen under en halvlek eller match i relation till övriga säsongen; både vad gäller relativt Sirius prestationer men även motståndarnas.

Tränarna upplever att de med hjälp av statistiken kunnat presentera en konkret och objektiv beskrivning av matchen för spelarna. Detta underlag använts som en klargörande bild av lagets prestationer, och då datan upplevs som objektiv har den möjliggjort en mer rak, och värderingslös konversation. Samtal där spelare tidigare kan ha känt sig anklagade har när tränare kunnat hänvisa till ett statistiskt underlag varit öppnare.

Vad gäller säsongsrapporterna har dessa använts både som ett verktyg för att identifiera utvecklingsmöjligheter, där hörnor varit ett konkret exempel, men också för att peka ut specifika delar av spelet där laget presterar på en hög nog nivå, exempelvis vad gäller närkamper. Då träning i grunden är specifik har dessa rapporter underlättat för tränarna att rikta träningen mot specifika mål.

Ett konkret exempel på hur säsongsrapporten användes var för att analysera Sirius spel under segersviten vid årets slut; säsongsrapporten visade att Sirius under denna period hade skott- och räddningsprocent ungefär i linje med vad motståndarna presterat under de 15 föregående matcherna. Detta användes som motivation för spelarna att visa att deras prestationer inte varit ohållbart bra under segersviten.

Utifrån spelarnas perspektiv har rapporterna även medfört en upplevd ökad professionalitet som tidigare saknats inom bandyn. Det faktum att alla närkamper och skott noterats har för vissa spelare upplevts som en motiverande faktor.

5.4 Vidare studier

Nedan följer förslag till vidare studier som potentiellt skulle leda till en djupare förståelse av sporten bandy.

5.4.1 Med detta dataunderlag

Något det finns dataunderlag om, men som inte undersökt alls är utvisningars påverkan på spelet. Det har under säsongen noterat varje gång en spelare blivit utvisad, samt huruvida det var en fem eller tio minuter lång utvisning. Upplevelsen är att en utvisning inte påverkar spelet nästintill alls, men att spelbilden förändras

vid två mans underläge. Det vore intressant att se hur exempelvis skottfördelningen, bollinnehavet och närkampernas fördelning påverkades av numerärt överläge, och detta kan göras med det dataunderlag finns här samlat.

Det faktum att skott utifrån har så låg sannolikhet att direkt resultera i mål (omkring 6 %) men är så vanligt förekommande, i snitt mer än tio gånger per match och lag (detta är mer än hälften av alla avslut i icke-fast spelform) var uppseendeväckande. När denna låga effektivitet lyfts för spelare och tränare har de ofta svarat att skott utifrån ofta leder till farliga situationer.

Det vore intressant att se hur ofta det stämmer. Vad händer i situationen efter ett långskott? Hur ofta leder de till returskott eller hörnor? Hur ofta återerövrar det anfallande laget bollen? Ifall det visar sig att skotten allt som oftast leder till att motståndarna får bollen skulle detta kunna påverka i vilka lägen spelare tar avslut, och ha inflytelse över deras vilja att skjuta utifrån.

Ytterligare en aspekt är att analysera hur matchens ställning påverkar spelet. Upplevelsen är att spelet blir ytterst påverkat när matchen bedöms vara avgjord. Stämmer detta? Skjuts det färre skott och skapas det färre närkampssituationer ifall endera lag leder stort i matchens slutskede? Har lag i ledningsställning en benägenhet att backa hem och fokusera på det defensiva, och gör de i så fall färre mål? Om denna skillnad finns, hur stor måste ledningen vara för att det ska bli mätbart?

Enligt de upptäckter som gjorts gällande långa passningar vid utomhusspel och dålig is skulle dessa passningars inverkan kunna studeras; vad blir utfallet av dessa passningar, och är det en bra strategi att slå dem i hög utsträckning?

5.4.2 Med ett utökat dataunderlag

Att med denna insamlingsmetod utöka dataunderlaget till att inkludera andra lag än Sirius matcher skulle kontextualisera resultaten; exempelvis skulle detta potentiellt besvara frågorna rörande huruvida Sirius har låg skott- eller räddningsprocent. Dataunderlag från ytterligare lag skulle även kunna användas till att undersöka ifall olika lag genom sina spelstilar utmärker sig i den här typen av data; vissa lag upplevs spela väldigt mycket på kontring, syns det, exempelvis?

Skillnaden i hörnstatistik beroende på sida skulle även kunna undersökas vidare. Denna skillnad kunde inte statistiskt säkerställas i detta underlag, men med ett utökat dataunderlag vore detta möjligt.

5.4.3 Med en annan insamlingsmetod

Med en mer raffinerad insamlingsmetod finns det nästintill oändliga möjligheter att skapa analysmodeller. Exempelvis passningsnätverk inom bandy skulle potentiellt

beskriva sporten. Även spelarspecifik statistik skulle vara av intresse.

6 Slutsats

6.1 Svar på del-frågeställningarna

Nedan följer en redogörelse av de fyra underfrågeställningarna i avsnitt 1.

6.1.1 Hur kan relevant data samlas från bandymatcher?

Datainmatnings-verktyget beskrivet i avsnitt 3 kan efter en kortare inlärningskurva användas för att en person ska kunna inmata en rad olika tids-, position- och lagstämplade datapunkter från en TV-sänd bandymatch under tidsrymd bara är några minuter längre än matchtiden. Verktyget är anpassningsbart vilket medför att användaren själv kan utöka det och lägga till och ta bort händelser av intresse. Verktyget möjliggör även för användaren att kombinera den direktinsamlade datan med efterinsamlad data, ifall så önskas.

6.1.2 Vilka sportanalys-modeller kan skapas från datan?

Samtliga modeller beskrivna i avsnitt 4 är skapade med hjälp av data insamlad av verktyget beskrivet ovan. Utöver simplare räknade modeller såsom antal skott, corsihändelser och närkamper har en **CF%**-modell enligt ekvation 1 och ett verktyg för att beräkna varje lags bollinnehav skapats. Tabell 3 och tabell 4 kan användas som en primitiv förväntad mål-modell, enligt det beskrivet i avsnitt 2.2.4.

Vad gäller lagens närkampsspel har flera modeller skapat för att presentera detta; var på banan duellerna äger rum enligt figur 4, lagens procentuella duellvinster enligt figur 5 och närkampernas utfall givet lags innehav från tabell 9 och figur 19.

Förslag på fler modeller som potentiellt skulle kunna skapas finns i avsnitt 5.4.1.

6.1.3 Hur ska resultaten från dessa metoder tolkas?

Hur de skapade sportanalys-modellerna bör tolkas presenteras i avsnitt 5. Sammantaget är det tydligt att bandy är en komplex sport, och att stort arbete återstår för att fullt ut kunna förklara spelet. Det kan konstateras att lags skottprocent kan fälla ett avgörande matchernas utfall under en längre period; en stor del av Sirius dåliga resultat kan förklaras med den lägre effektiviteten i avsluten de haft under hösten. Att ha ett övertag gentemot motståndarna i skott på mål och skottförsök bidrar även till att vinna matcherna.

Närkampsspelet visade sig vara en svåranalyserad parameter; detta dataunderlag visar att det finns ett samband mellan ett övertag i närkampsspelet och att vinna

matcher. Ytterligare forskning på detta område kan behövas för att skapa en mer nyanserad bild av närkampernas påverkan på sporten.

Rapporten har inte kunnat fastslå att bollinnehav påverkar matchernas utfall.

Vilken inverkan isens kvalitet har på spelet har även analyserats. Enligt den insamlade datan resulterar utomhusspel i att lagen slår fler lång passningar. Vid vad som bedöms vara dålig iskvalitet ökar antalet dueller markant, och de placeras i större utsträckning framför målen.

6.1.4 Hur ska dessa resultat presenteras på bästa sätt?

I avsnitt 4.7 beskrivs processen som lett fram match- och säsongsrapporternas utformning. Dessa har varit det huvudsakliga medlet som använts för att förmedla resultaten till Sirius spelare och ledare. Genomgående har tydlighet prioriterats framför estetik; ambitionen är att spelare och ledare ska kunna ta till sig informationen på kortast möjliga tid.

Sammantaget har presentationsformen varit något som diskuterats ingående, och där hänsyn tagits till åsikter från spelare och ledare inom Sirius. Rapporternas utformning är ett exempel på hur analysmodellerna kan presenteras.

6.2 Svar på rapportens huvud-frågeställning

Rapportens frågeställning var:

- Hur kan sporten bandy, och mer bestämt IK Sirius spel, beskrivas med hjälp av sportanalys-modeller på ett vis som ökar förståelsen kring bandy och hjälper IK Sirius verksamhet?

Detta arbete syftade till att med sportanalys-modeller beskriva sporten bandy, och hjälpa Sirius Bandys verksamhet. För att skapa dessa modeller har ett datainsamlingsarbete gjorts; detta arbete har bestått i att avgöra vilka datapunkter som krävs för sportanalys-modellerna, att skapa ett tillvägagångssätt för datainsamling och sedan ett insamlingsarbete. Efter detta har en rad olika sportanalys-modeller skapats utifrån insamlad data. Resultatet från dessa modeller har sedermera analyserats. Slutligen har detta resultat presenterats i två olika PowerPoint-mallar.

Bandy är en mycket komplex sport, och det här arbetet har bara skrapat på ytan vad gäller att förklara den genom sportanalys. Med detta sagt, syftade detta arbete till att skapa en första grund av sportanalys-modeller som beskriver sporten bandy och som hjälper IK Sirius verksamhet, vilket har gjorts.

Referenser

- [1] The Athletic. *Author: Cam Charron*. URL: <https://theathletic.com/author/cam-charron/>. hämtad 2022-11-24.
- [2] Svensk Bandy. *Bandyfinalen historik*. URL: <https://old.svenskbandy.se/BANDYFINALEN/HISTORIK/>. hämtad 2022-11-24.
- [3] Svensk Bandy. *Elitrapport - målsnitt*. URL: <https://www.elitrapport.se/stats/>. hämtad 2023-01-03.
- [4] Svensk Bandy. *Startsida - Elitrapport*. URL: <https://www.elitrapport.se/games/>. hämtad 2022-12-09.
- [5] Svenska Bandyförbundet. *Bandyhallar*. URL: <https://old.svenskbandy.se/BANDY-INFO/anlaggningar/Bandyhallar>. hämtad 2022-11-24.
- [6] Bandygrytan. *2022 - Elitserien Herr (Alla)*. URL: <https://bandygrytan.se/#/competitions/2022/league/11829>. hämtad 2022-12-09.
- [7] George Casella & Roger L. Berger. *Statistical Inference 2nd Edition*. Duxbury Press, 2002. ISBN: 0-534-24312-6.
- [8] Bandyfeber - Mattias Bladh. *BANDYNS HISTORIA, DEL 2: GENOMBROTET*. URL: <https://www.bandyfeber.com/bandyns-historia-del-2-genombrottet/>. hämtad 2022-11-24.
- [9] Alex Borelli. *Analyzing NBA Shot Success Through Expected Points*. URL: <https://takeanumber.substack.com/p/analyzing-nba-shot-success-through>. hämtad 2022-12-09.
- [10] Cam Charron. *Are blocked shots really valuable?* URL: <https://canucksarmy.com/news/are-blocked-shots-really-valuable>. hämtad 2022-11-25.
- [11] Loïc Ravenel & Roger Besson Drs Raffaele Poli. *How important is ball possession in football?* URL: <https://www.football-observatory.com/IMG/sites/mr/mr41/en/>. hämtad 2023-01-03.
- [12] Eliteprospects. *Staff Profile: Cam Charron*. URL: <https://www.eliteprospects.com/staff/30355/cam-charron>. hämtad 2022-11-24.
- [13] Daily Faceoff. *Tags: Cam Charron*. URL: <https://www.dailyfaceoff.com/tag/cam-charron/>. hämtad 2022-11-24.
- [14] Vinícius Farias m. fl. "Relationship between ball possession and match outcome in UEFA Champions League". I: *Motricidade* 16 (jan. 2021), s. 1–7. doi: 10.6063/motricidade.18382.

- [15] Dimitri Filipovic. *The Hockey PDOcast*. 2022. URL: <https://open.spotify.com/episode/5bPWqdd3FphE4QKr9Qaeaq?si=dnQ5BTMxRgidY7jHj3nd=1>.
- [16] Dimitri Filipovic. *Tweet*. URL: <https://twitter.com/DimFilipovic/status/1352654297821925378?s=20&t=3l19DAW0yo6BDM3akLFw2w>. hämtad 2023-01-02.
- [17] Svenska Fotbollförbundet. *Spelregler för fotboll 2022*. URL: <https://www.svenskfotboll.se/48e359/globalassets/svff/bilderblock/aktiva/domare/spelregler-fotboll.pdf>. hämtad 2022-11-24.
- [18] Cap Friendly. *CAREER EARNINGS*. URL: <https://www.capfriendly.com/career-earnings>. hämtad 2023-01-02.
- [19] Federation of International Bandy. *Bandy Playing Rules*. URL: http://www.worldbandy.com/wp-content/uploads/2019/08/Int.-bandyregler_19-1.pdf. hämtad 2022-11-24.
- [20] Rob Mixer/ Columbus Blue Jackets. *FancyStats Breakdown: Corsi vs. Fenwick*. URL: <https://www.nhl.com/bluejackets/news/fancystats-breakdown-corsi-vs-fenwick/c-728734>. hämtad 2022-12-09.
- [21] Karim Kurji. *Breaking down the inconsistencies in counting hits between different NHL arenas*. URL: <https://thewincolumn.ca/2021/01/30/breaking-down-the-inconsistencies-in-counting-hits-between-different-nhl-arenas/>. hämtad 2023-01-02.
- [22] National hockey league. *Statistik från säsongen 1959/60*. URL: <https://www.nhl.com/stats/skaters?reportType=season&seasonFrom=19591960&seasonTo=19591960&gameType=2&filter=gamesPlayed,gte,1&sort=plusMinus&page=0&pageSize=50>. hämtad: 2022.08.24.
- [23] Michael Lewis. *Moneyball: The Art of Winning an Unfair Game*. W. W. Norton & Company, 2003.
- [24] Andreas Lindholm m. fl. *Machine Learning - A First Course for Engineers and Scientists*. Cambridge University Press, 2022. URL: <https://smlbook.org>.
- [25] Lund Research Ltd. *Pearson Product-Moment Correlation*. URL: <https://statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide.php>. hämtad 2023-01-03.

- [26] Bob McKenzie. *McKenzie: The real story of how Corsi got its name*. URL: <https://www.tsn.ca/mckenzie-the-real-story-of-how-corsi-got-its-name-1.10001>. hämtad 2022-12-09.
- [27] New York Mets. *Position: Analyst, Baseball Analytics*. URL: <https://blogs.fangraphs.com/instagrams/job-posting-new-york-mets-baseball-analytics-positions/>. hämtad 2022-11-25.
- [28] Money puck. *About Money puck*. URL: <https://moneypuck.com/about.htm>. hämtad 2022-12-09.
- [29] CJ Moore. *How men's college basketball teams are using outside analytics firms to find an edge*. URL: <https://theathletic.com/3631462/2022/09/29/college-basketball-analytics-business/>. hämtad 2022-11-25.
- [30] Sarang Narkhede. *Understanding Confusion Matrix*. URL: <https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62>. hämtad 2022-12-28.
- [31] Nationalencyklopedin. *Bandy*. URL: <http://www.ne.se/uppslagsverk/encyklopedi/l%C3%A5ng/bandy>. hämtad 2022-11-24.
- [32] NFELo. *What are Expected Points Added (EPA) in the NFL*. URL: <https://www.nfeloapp.com/analysis/expected-points-added-epa-nfl/>. hämtad 2022-12-09.
- [33] Lasse Nordström. *Vad har Sveriges första bandyhall betytt för Edsbyn?* URL: <https://old.svenskbandy.se/Nyhetsarkiv/vadharsverigesforstabandy>. hämtad 2022-11-24.
- [34] Harvard Petra. *How data analytics is revolutionizing the NBA*. URL: <https://d3.harvard.edu/platform-digit/submission/how-data-analytics-is-revolutionizing-the-nba/>. hämtad 2022-11-25.
- [35] Hockey Reference. *NHL Advanced Stats/Analytics*. URL: <https://www.hockey-reference.com/analytics/>. hämtad 2022-12-09.
- [36] Stephen Shea. *THE 3-POINT REVOLUTION*. URL: <https://shottracker.com/articles/the-3-point-revolution>. hämtad 2022-11-25.
- [37] Rink Stats. *Why we should be trying to do better than Corsi and Fenwick*. URL: <http://rinkstats.com/2014/05/corsi-and-fenwick-suck-or-why-we-should>. hämtad 2022-12-09.

- [38] StatsBomb. *Expected Goals xG FAQ*. URL: <https://statsbomb.com/soccer-metrics/expected-goals-xg-explained/>. hämtad 2022-12-09.
- [39] StatsMuse. *NBA player most career turnovers per game with a minimum of 400 games played*. URL: <https://www.statmuse.com/nba/ask/nba-player-most-career-turnovers-per-game>. hämtad 2023-01-02.
- [40] StatsMuse. *NHL most giveaways by a player all-time*. URL: <https://www.statmuse.com/nhl/ask?q=nhl+most+giveaways+all-time>. hämtad 2023-01-02.
- [41] Linköpings universitet. *Sports Analytics*. URL: <https://liu.se/en/research/sports-analytics>. hämtad: 2022.08.24. Min översättning.
- [42] S. Hilditch V. Barnett. "The Effect of an Artificial Pitch Surface on Home Team Performance in Football (Soccer)". I: *Journal of the Royal Statistical Society* 156.1 (1993), s. 39–50.
- [43] Michael Verardi. *How Arena Statistics Impact Fantasy Hockey*. URL: <https://lastwordonsports.com/hockey/2021/09/23/how-arena-statistics-impact-fantasy-hockey/>. hämtad 2023-01-02.
- [44] Jonny Whitemore. *What Are Expected Goals (xG)?* URL: <https://theanalyst.com/eu/2021/07/what-are-expected-goals-xg/>. hämtad 2022-12-09.
- [45] Evan Zamir. *How Should We Talk about Turnovers? Carefully*. URL: <https://www.goldenstateofmind.com/2012/2/18/2807894/how-should-we-talk-about-turnovers>. hämtad 2023-01-02.