

***Neue Aspekte bei der Sportsuche:
Der Sportpilot der ZBS***

Florian Seiffert
ZBSport

AGSB-Tagung Köln, 28.03.2007

Überblick – neue Aspekte

- Stand der Dinge
- Logfiles sind unsere Augen
- Analyse von Suchanfragen
- Verbesserungen durch Terminologie!?
- Was bringt die Zukunft?

SportPilot: Stand der Dinge

- Ziele:
 - ~~EINE Suche (Websuche, Katalogsuche, Datenbanksuche, ...)~~
 - eine einfache Suche
 - eine Suche für Laien (keine Experten)
 - „google like“

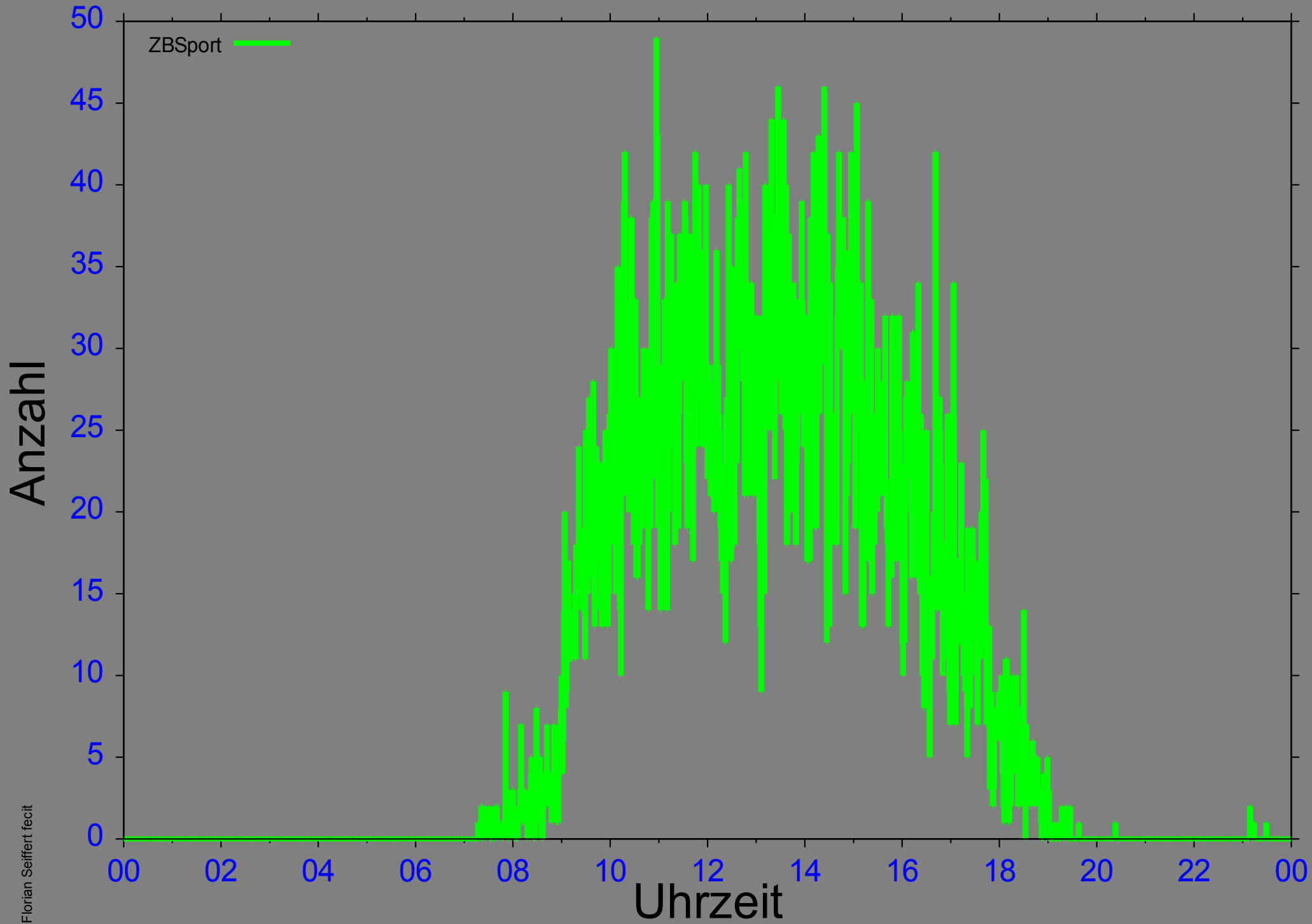
SportPilot: Stand der Dinge

- Inhalte
 - „alles zum Sport“
 - Sportpersonen des Munzinger Archivs (8.933)
 - Ein Link der ZBSport (10.161)
 - Eine Webseite der ZBSport (570)
 - Sportzitate (218)
 - Pressemitteilungen der DSHS (377)

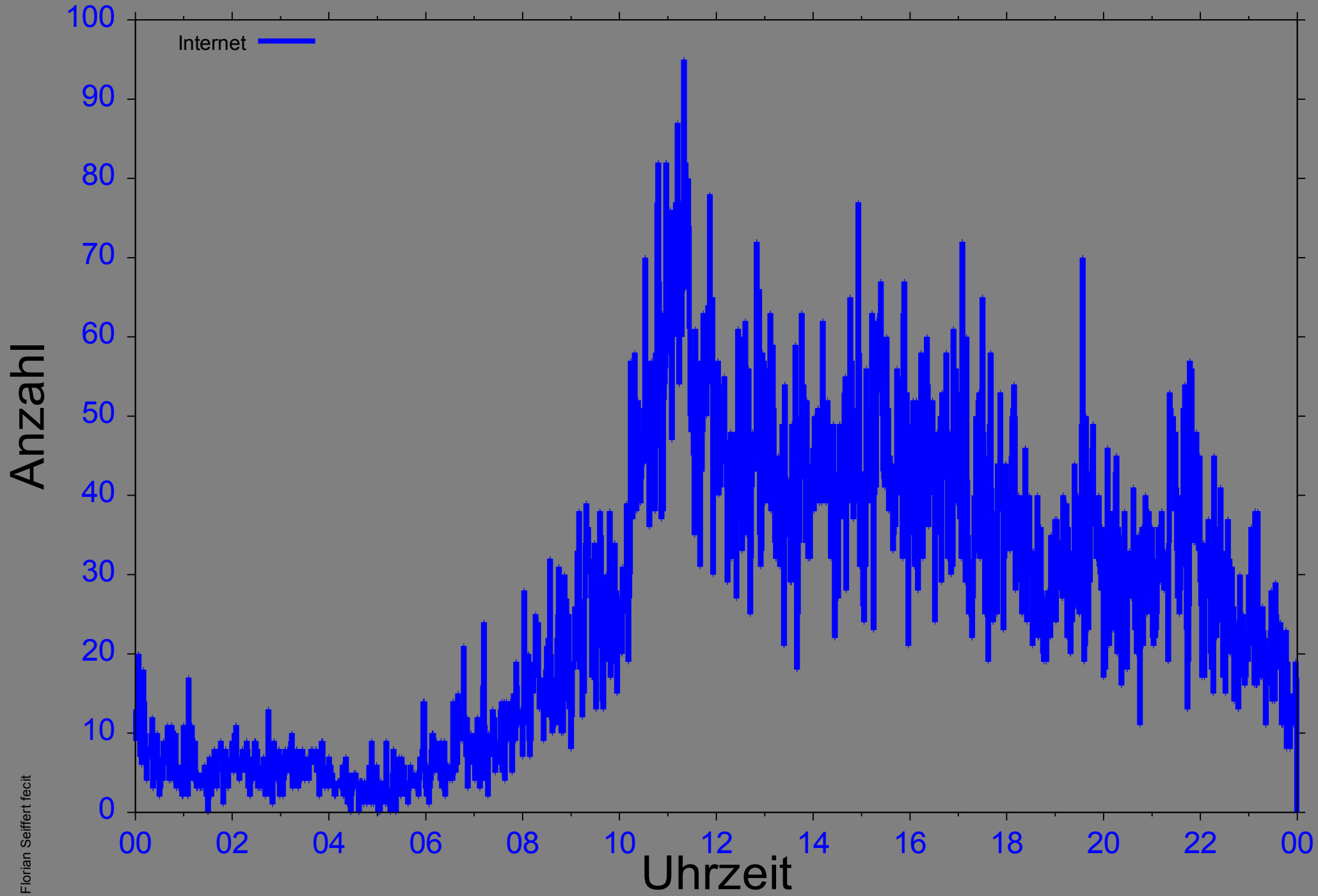
Logfiles sind unsere Augen

- Webserver und Datenbanken schreiben Informationen mit.
- Wer sie nicht auswertet, weiß wohl kaum, wie sein System benutzt wird.
- Wir werten sie aus, um zu erfahren:
 - wer arbeitet mit dem SportPilot?
 - was wird gesucht, welche Informationen erwartet Kundin?
 - müssen wir was verändern?
 - fehlen Inhalte?

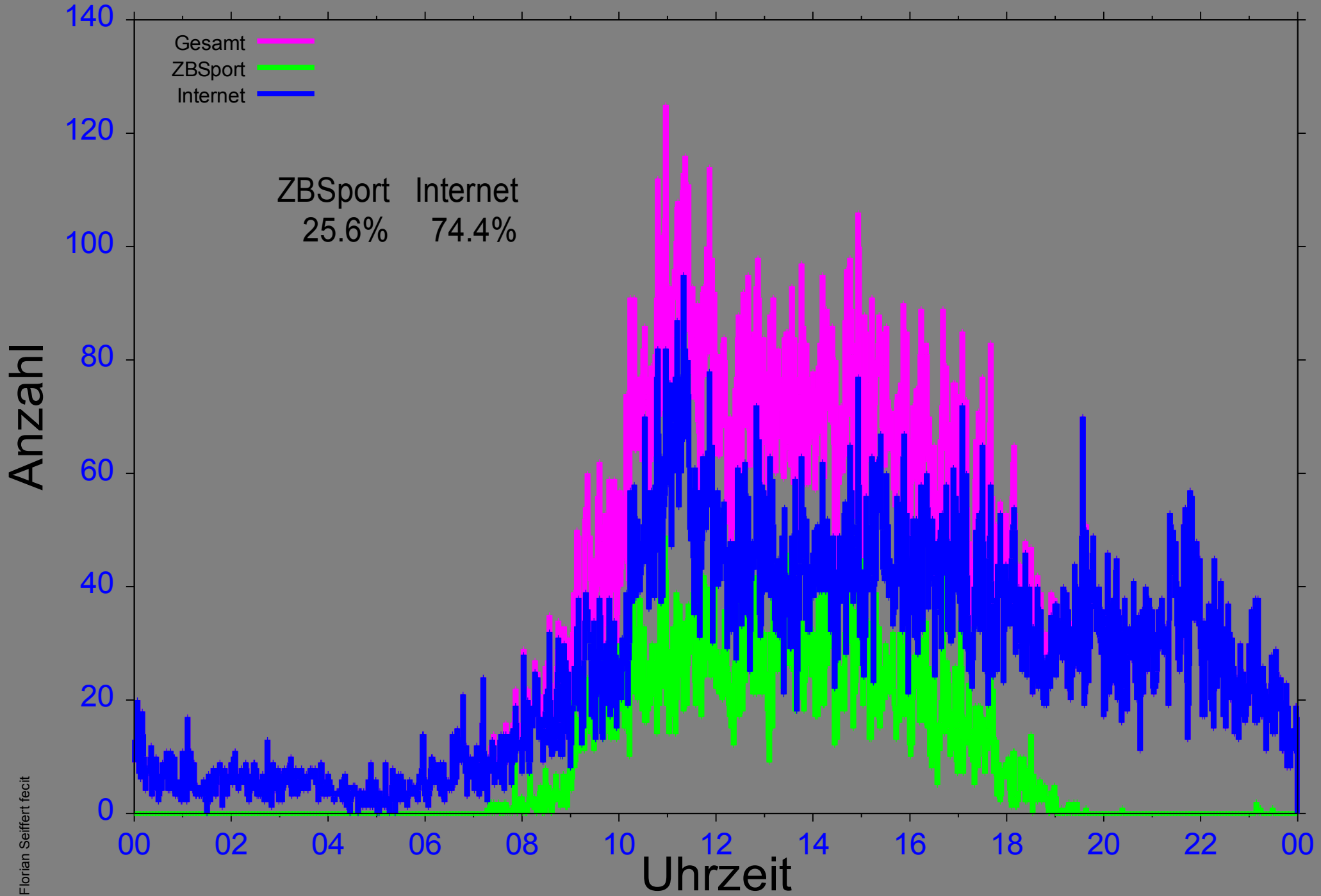
Anfragenverteilung Sportpilot



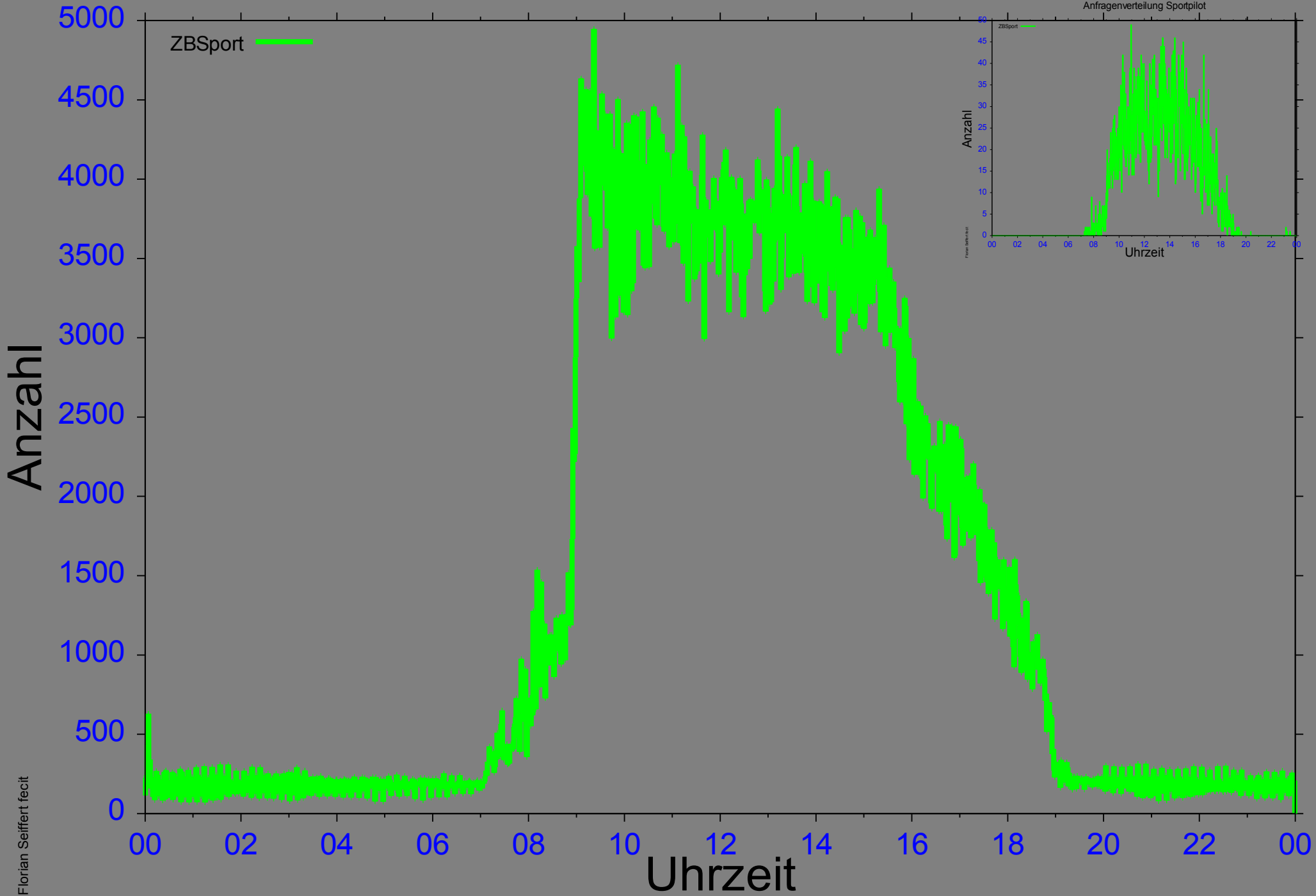
Anfragenverteilung Sportpilot



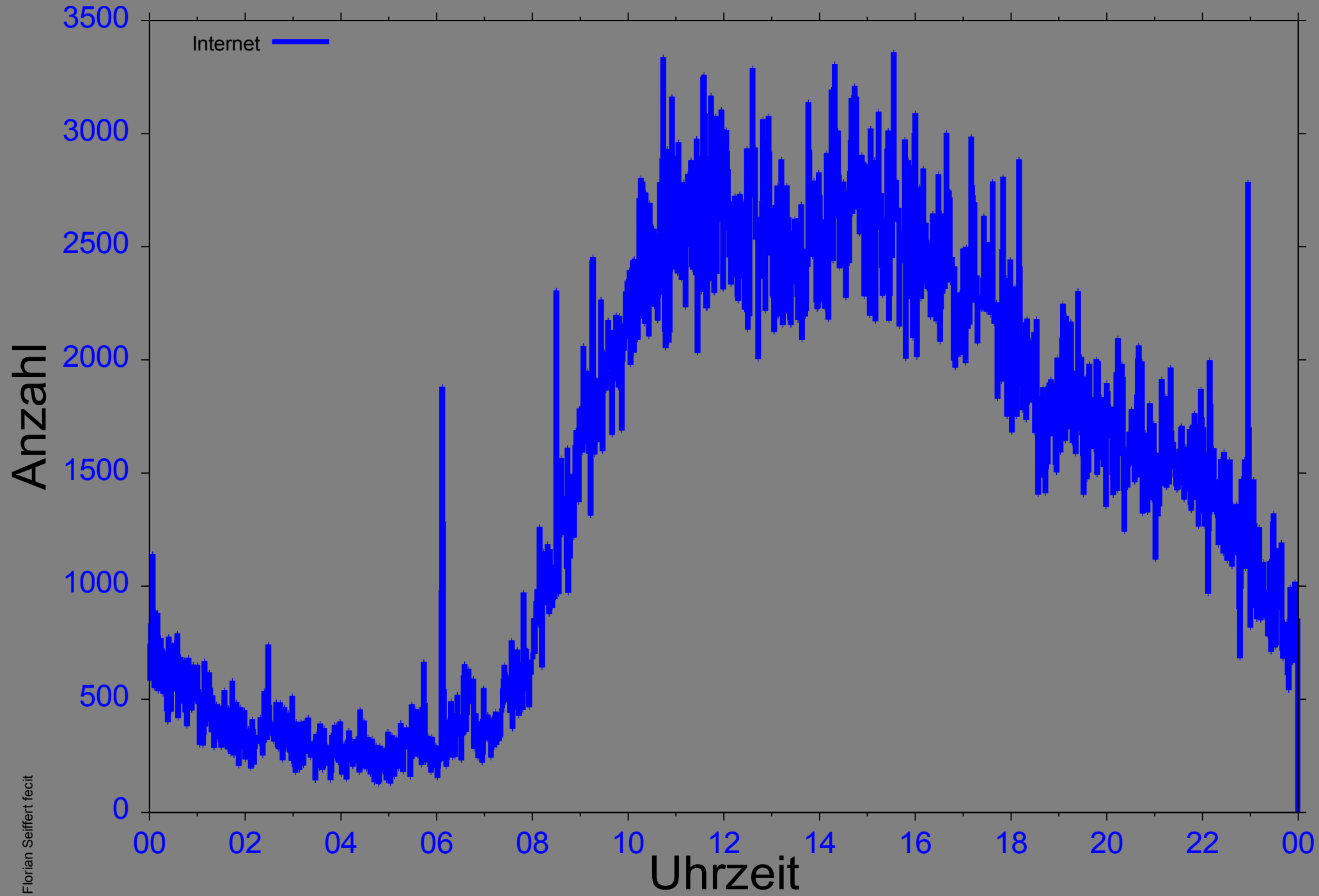
Anfragenverteilung Sportpilot



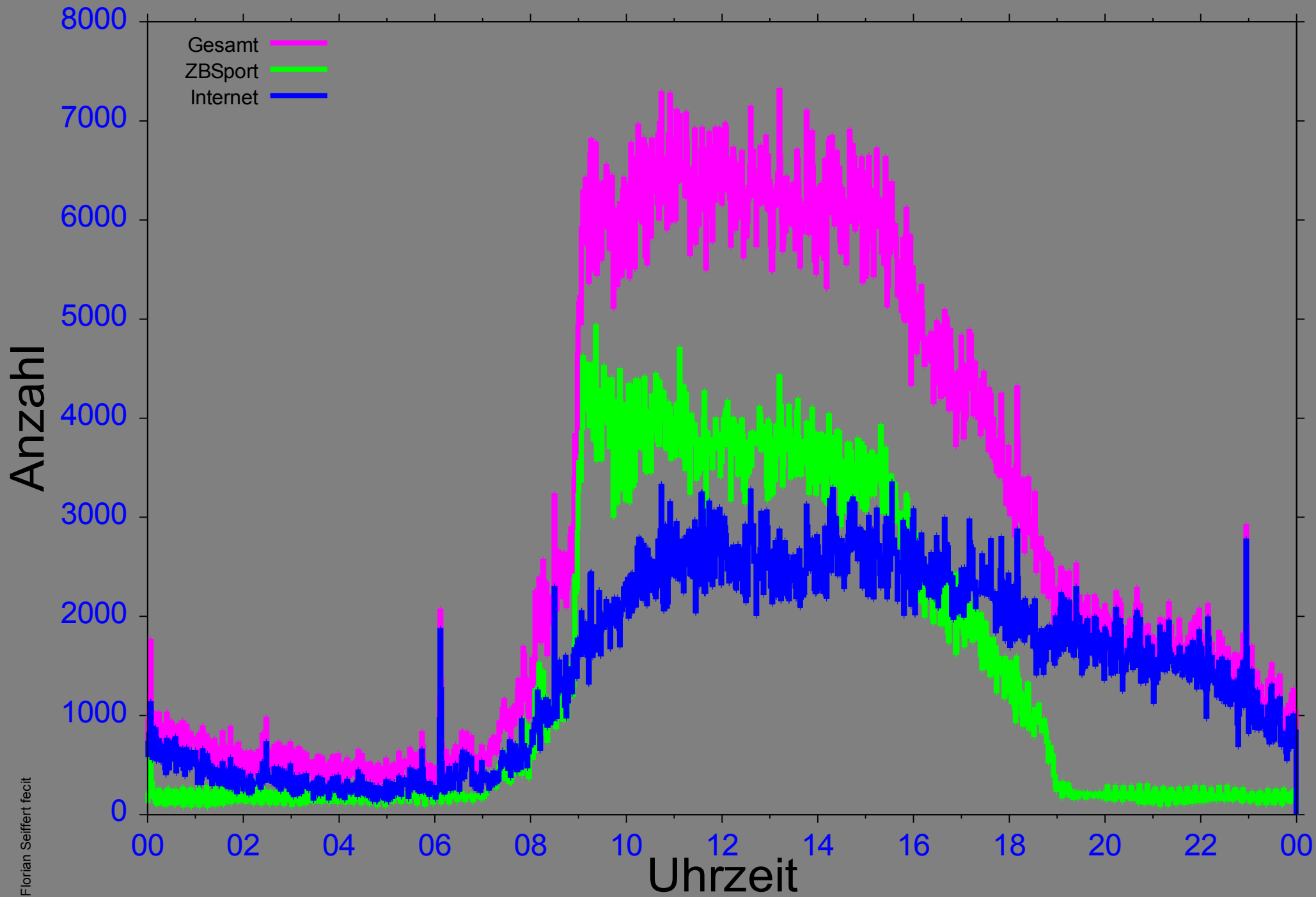
Webzugriffeverteilung Sportpilot



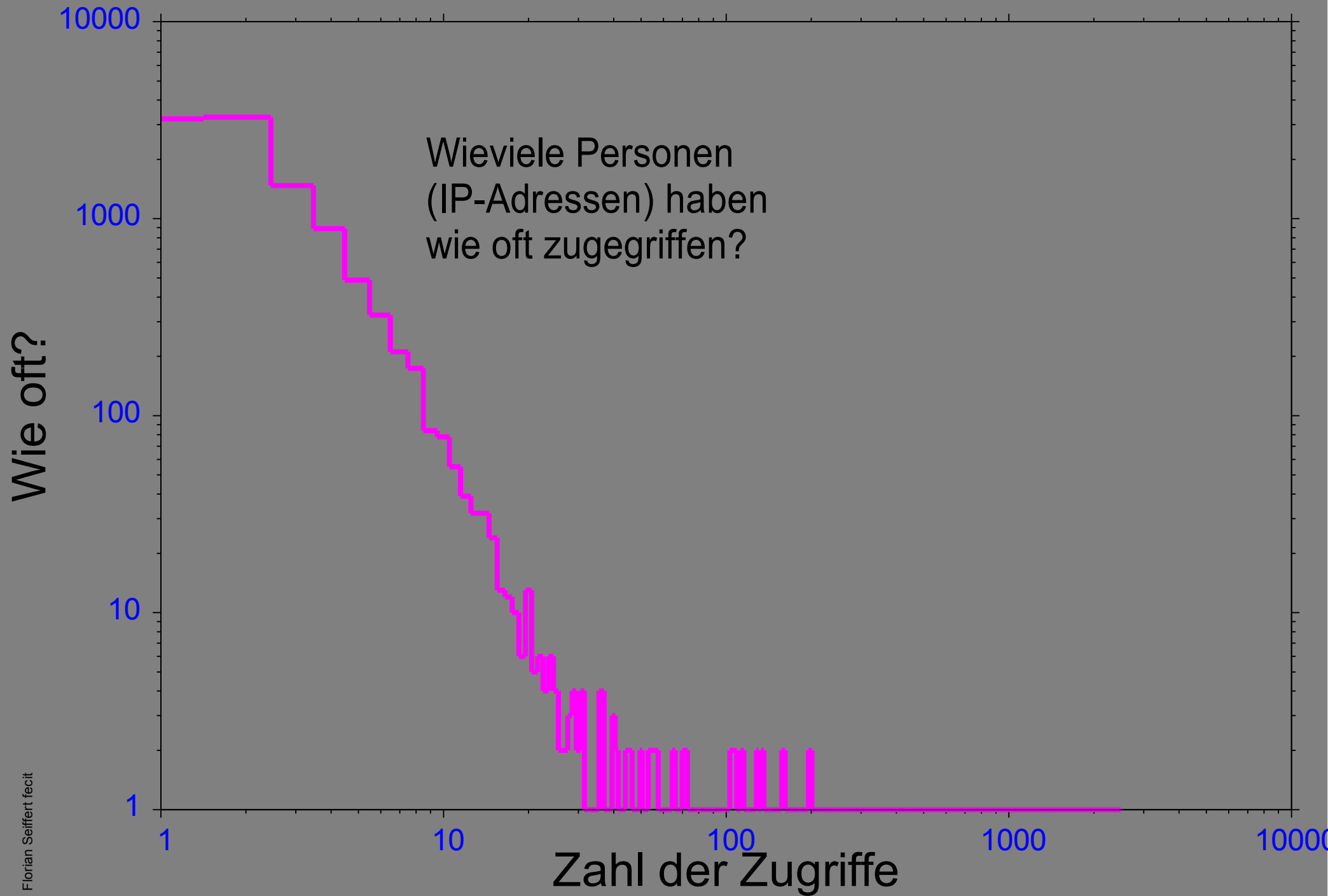
Webzugriffeverteilung Sportpilot



Webzugriffeverteilung Sportpilot



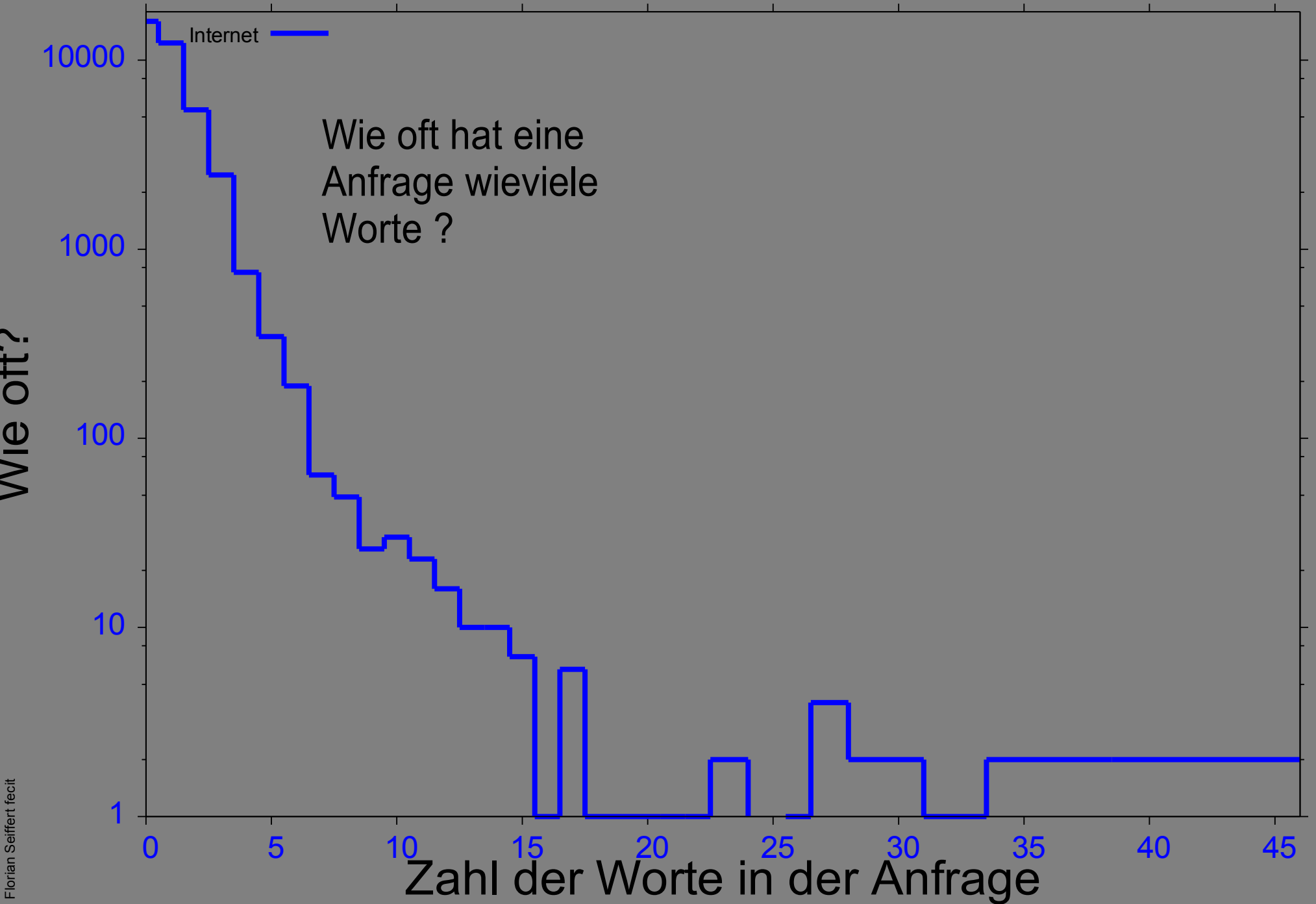
IP-Adresszugriffsverteilung Sportpilot



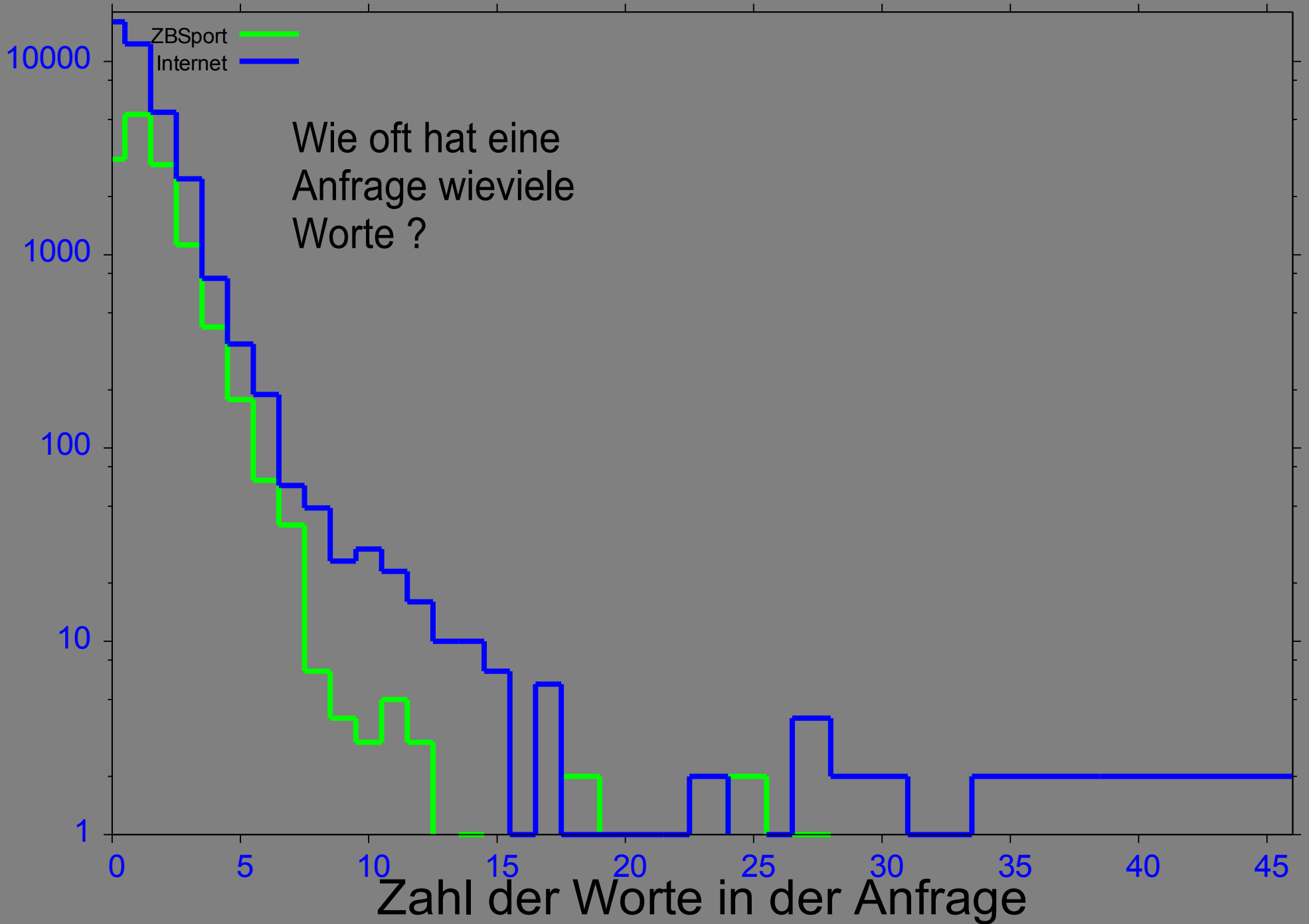
Wortverteilung Sportpilot



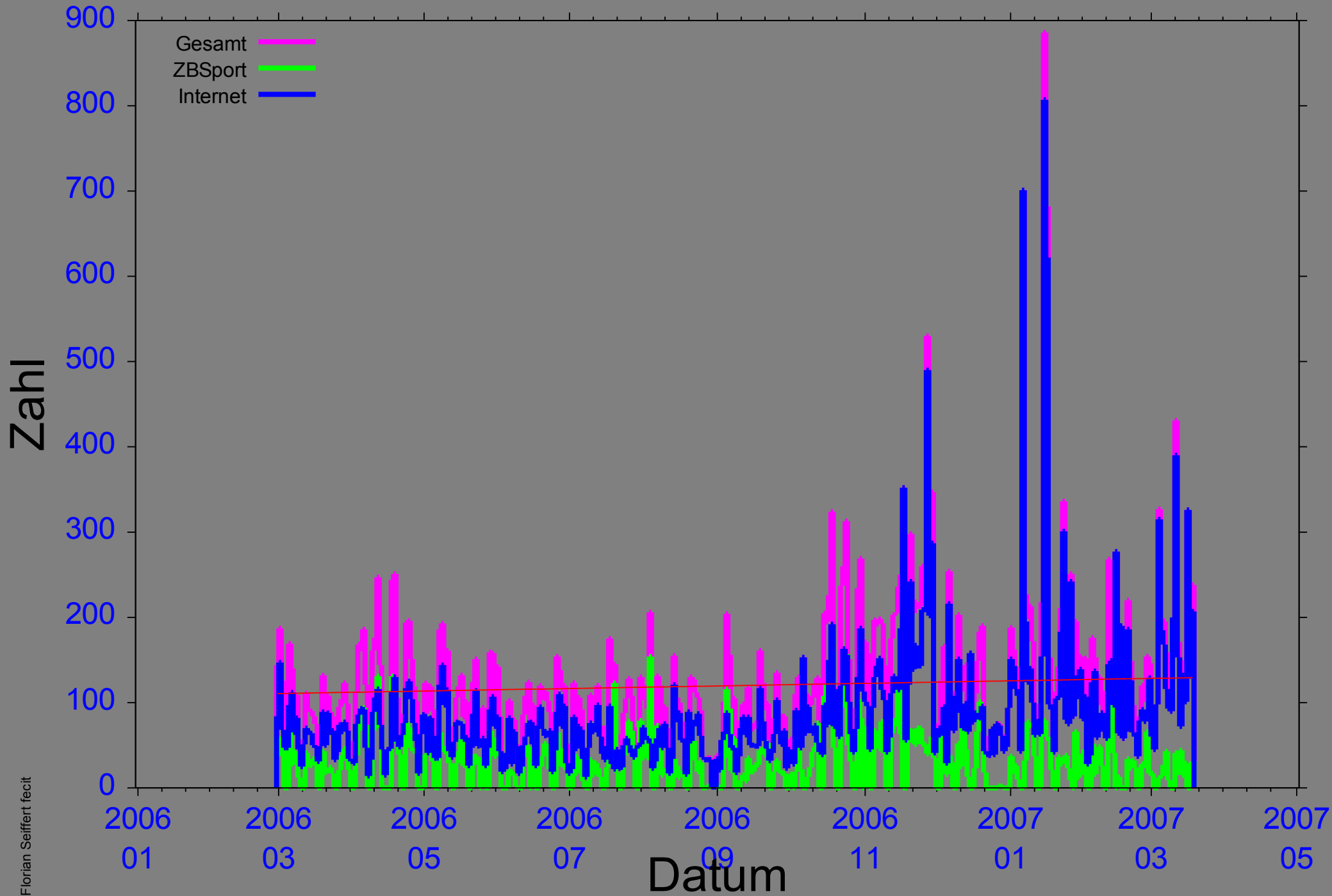
Wortverteilung Sportpilot



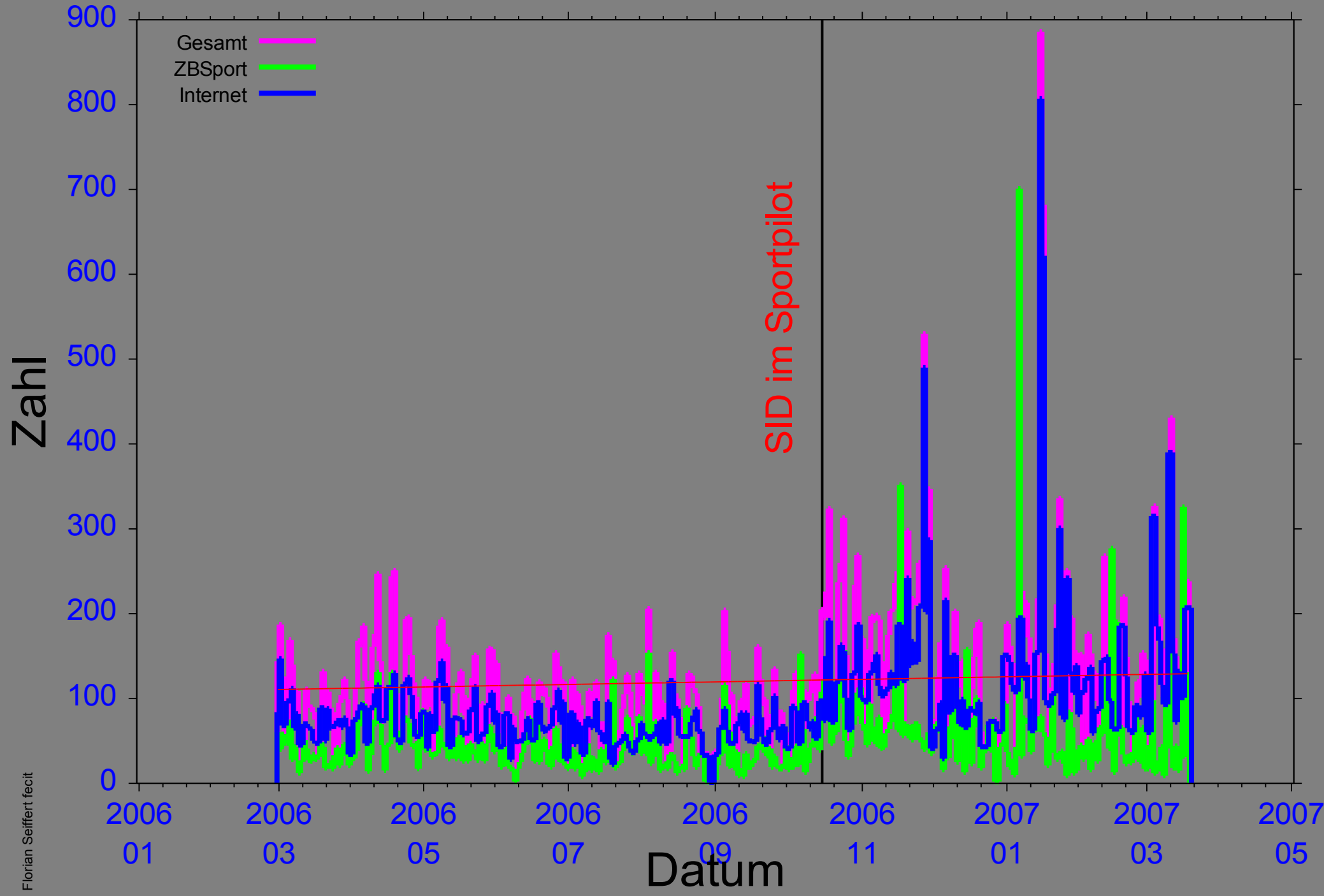
Wortverteilung Sportpilot



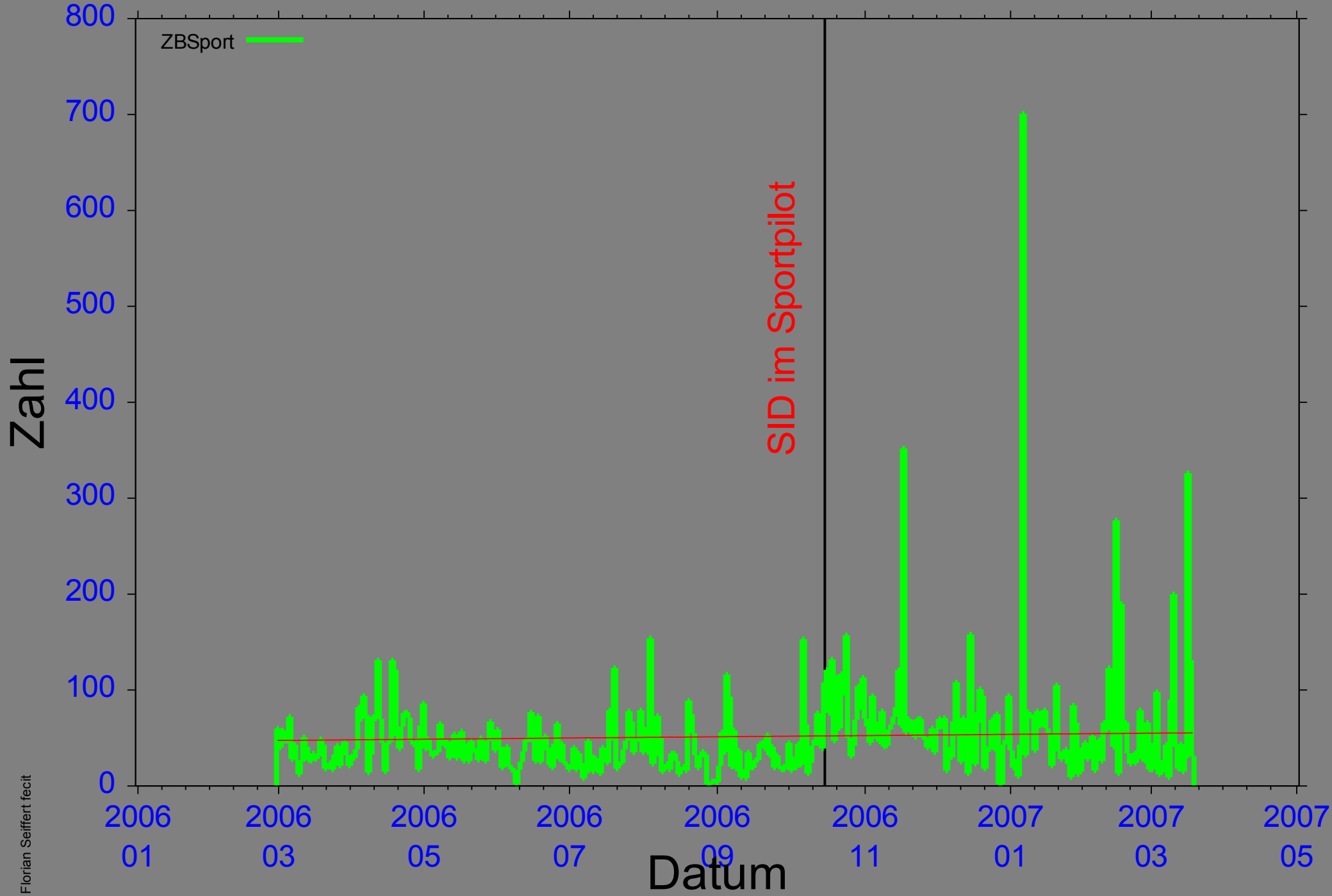
Sportpilot: Zugriffe pro Tag



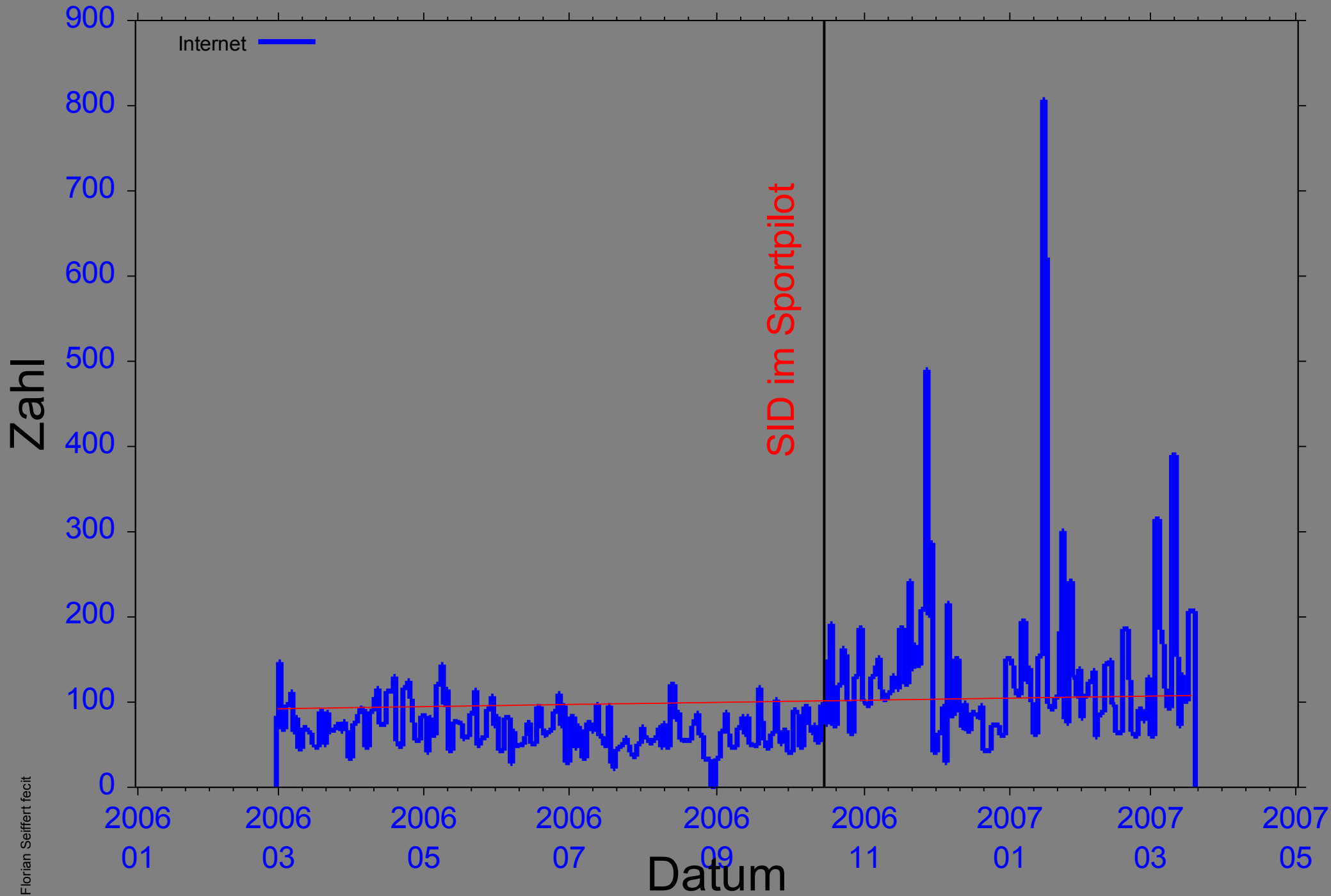
Sportpilot: Zugriffe pro Tag



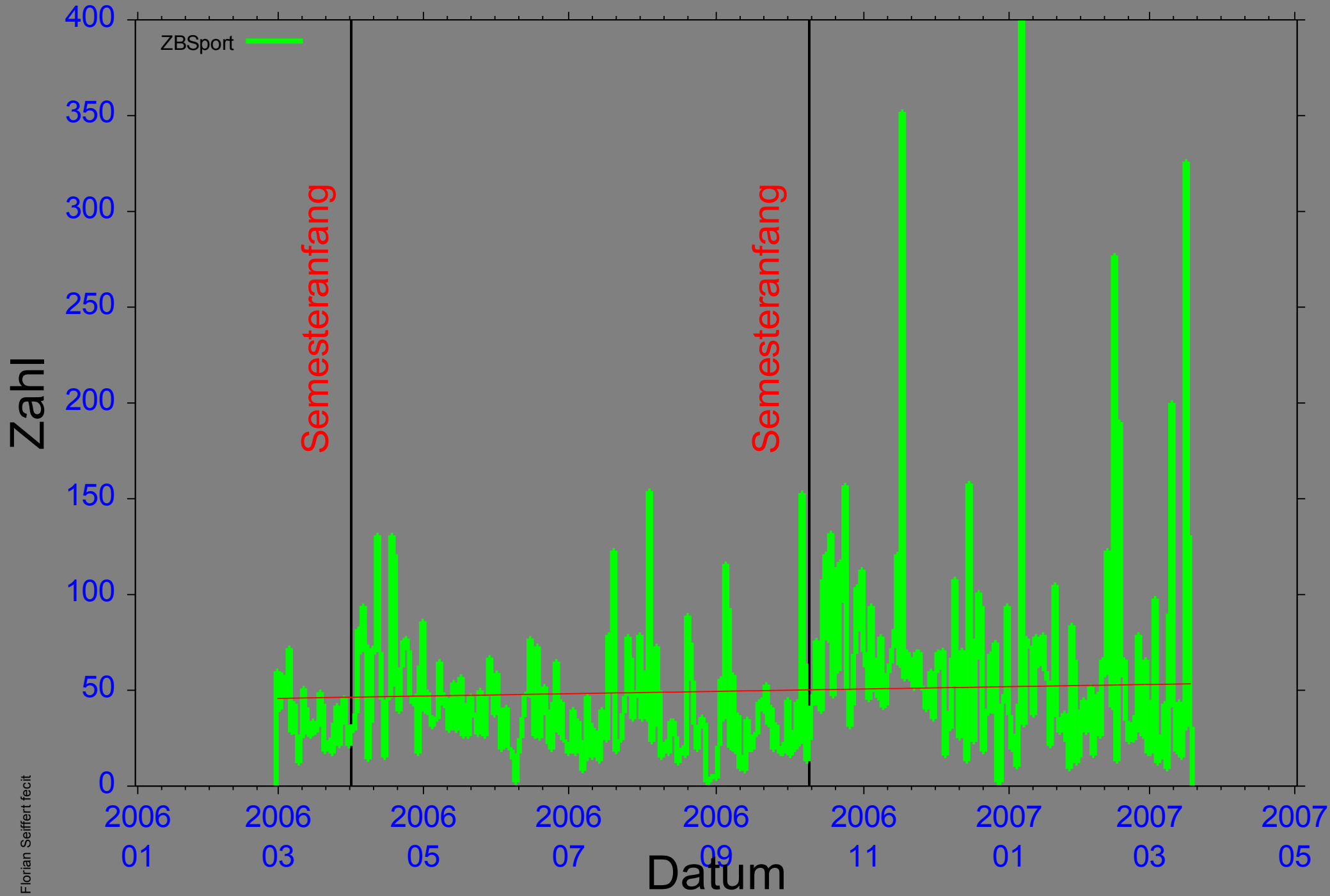
Sportpilot: Zugriffe pro Tag



Sportpilot: Zugriffe pro Tag



Sportpilot: Zugriffe pro Tag



Was lernen wir daraus?

- Drei Viertel der Zugriffe auf den Sportpilot kommen nicht aus der ZBSport !
- Zugriffe aus der ZBSport kommen ganz überwiegend von Studierenden.
- Viele fragen nur wenige mal, wenige fragen immer wieder.
- Viele fragen mit einem oder zwei Worten.
- Beste Stilllegungszeit (Backup) ist gegen 4 Uhr.
- Zahl der Zugriffe steigt (langsam).

Was lernen wir daraus?

- Drei Viertel der Zugriffe auf den Sportpilot kommen nicht aus der ZBSport !
- Zugriffe aus der ZBSport kommen ganz überwiegend von Studierenden.
- Viele fragen nur wenige mal, wenige fragen immer wieder.
- Viele fragen mit einem oder zwei Worten.
- Beste Stilllegungszeit (Backup) ist gegen 4 Uhr.
- Zahl der Zugriffe steigt (langsam).

SportPilot: Anfragen

- Bisher: 51.377 Anfragen
- Beispiele
 - „uefa und stadions“
 - wie fühlt der fan sich im stadion?
 - was will der fan im stadion sehen?
 - körperkultur + leipzig + zeitschrift
 - zeitschrift für sportmedizin
 - freiwald aufwärmen im sport
- Wir schauen uns ein paar der 7735 Einwortanfragen aus der ZBSport an.

Anfragen – 1 mal

- 1 KANAZAWA
- 1 GARDELIN
- 1 NIKEFOROS
- 1 ZENTNER
- 1 LETZTELTTER
- 1 FRIEDEL
- 1 WOLTERS,PETRA
- 1 LEXIK*
- 1 ECCENTRIC
- 1 BOSNIEN
- 1 SCHWANGERE
- 1 BÄNDERIß
- 1 WIAD
- 1 SOZIALPÄDAGOGIK
- 1 LAUFFIGUREN
- 1 EQUESTRIAN
- 1 ERNING
- 1 SCHWIMMENLERNEN
- 1 B41\$0000008982
- 1 CHIARELLO
-

Anfragen – 46 bis 542 mal

-
- 46 VOLLEYBALL
- 47 LEICHTATHLETIK
- 48 SPOLIT
- 48 TENNIS
- 51 BIOMECHANIK
- 53 FUßBALL
- 53 VERLÄNGERUNG
- 54 JUDO
- 57 ERNÄHRUNG
- 57 VORLESUNGEN
- 57 BASKETBALL
- 59 KLINSMANN
- 60 BENUTZERKONTO
- 60 PODOLSKI
- 61 BECKENBAUER
- 63 DIPLOMARBEIT
- 66 HANDBALL
- 70 AUFWÄRMEN
- 79 KRAFTTRAINING
- 95 SCHWIMMSPIELE
- 97 FUSSBALL
- 98 SCHWIMMEN
- 122 DOPING
- 142 DIPLOMARBEITEN
- 285 VORLESUNG
- 542 WEBSEITE

Anfragen – 46 bis 542 mal

Treffer in der Terminologiedatenbank

-
- 46 VOLLEYBALL
- 47 LEICHTATHLETIK
- 48 SPOLIT
- 48 TENNIS
- 51 BIOMECHANIK
- 53 FUßBALL
- 53 VERLÄNGERUNG
- 54 JUDO
- 57 ERNÄHRUNG
- 57 VORLESUNGEN
- 57 BASKETBALL
- 59 KLINSMANN
- 60 BENUTZERKONTO
- 60 PODOLSKI
- 61 BECKENBAUER
- 63 DIPLOMARBEIT
- 66 HANDBALL
- 70 AUFWÄRMEN
- 79 KRAFTTRAINING
- 95 SCHWIMMSPIELE
- 97 FUSSBALL
- 98 SCHWIMMEN
- 122 DOPING
- 142 DIPLOMARBEITEN
- 285 VORLESUNG
- 542 WEBSEITE

Einsatz von TerminologieDB sinnvoll?!

Terminologiedatenbank

The screenshot shows a Mozilla Firefox browser window with the following elements:

- Window Title:** Terminologie - Mozilla Firefox
- Menu Bar:** Datei, Bearbeiten, Ansicht, Chronik, Lesezeichen, Extras, Hilfe
- Address Bar:** http://localhost/cgi-bin/terminologie.pl
- Search Bar:** Google
- Bookmarks Bar:** DSHS, HBZ, misc, Myhome, Google, ZBS lokal, Poldi, ZBSportWiki, VifaWiki
- Page Title:** Terminologie
- Content:**
 - 14245 Sätze
 - Search input field
 - Buttons: Suche, Start, Help
 - Checkbox: tolerant
 - Erzwinge Treffer in: BISp Sponet Spowis FES zbs Syn_01
 - Verhindere Treffer in: BISp Sponet Spowis FES zbs Syn_01
 - Egal: BISp Sponet Spowis FES zbs Syn_01
- Footer:** Ein Skript von zbsport.de
- Status Bar:** Fertig

SportPilot: Verbesserungen durch Terminologie!?

- Der Einsatz einer Konkordanz als externes Expertinnenwerkzeug (Liste) ist beim SportPilot nicht sinnvoll !
- Wir brauchen einen Automatismus !
 - ENTWEDER: Verarbeitung der Suchanfrage
 - ODER: Verarbeitung der Datensätze

SportPilot: Plan

- Anreichern der SportPilot-Datensätze mit Synonymen
 - Recall erhöhen = Treffer ermöglichen , die sonst nicht erzielt worden wären
 - Kundin sieht, WARUM der Treffer erzielt wurde
 - Auto-Evaluation:
 - Mitloggen, was wurde gesucht und warum wurde ein Treffer erzielt!?
 - Tun wir das Richtige?
 - Ergebnisse: AGSB-Tagung 2008

SportPilot: Plan

In den nächsten Wochen:

- Katalog der ZBSport im SportPilot
- Auto Enrichment / Anreicherung
 - Umschlagbilder
 - Einzug von Verlagstexten automatisch vom Server der Deutschen National Bibliothek
- Kooperation mit Meyer & Meyer Verlag
 - Inhaltsverzeichnisse
 - Umschlagbilder
 - Volltexte