

Figure A1 Test Data Summary

<u>raw/source material</u> <u>name</u>	<u>collection</u> <u>name</u>	<u>size</u> reqs docs		<u>primary indexing</u>
Cranfield	C1400I	225	1400	manual from documents
UKCIS	U27000T	182*	27361	automatic from titles
	U27000P	182*	27361	automatic from titles via profiles
NPL	N11500A	93	11429	automatic from abstracts
	N11500T	93	11416	automatic from titles
Evans (INSPEC)	E2500T	39	2542	automatic from titles
	E2500P	39	2542	automatic from titles via profiles

* Most of the tests were done with a subset of requests having strict as opposed to fancy Boolean form for the original SDI profiles. These subsidiary collections with 75 requests are named U27000Tb and U27000Pb.

The C1400I, U27000P and E2500P requests are manual, the others automatic.

Figure A2 Test Raw Material

1. Cranfield : /^{see} Cleverdon, Mills and Keen 1966
Project on factors determining the performance of indexing systems.

Size

221 queries ; several subsets, especially 42 queries
1400 documents ; 200 documents

Subject

Aeronautics.

Indexing source

Full texts; &(title +) abstracts/titles.

Index languages

3 types, in 30 forms, all applied manually; abstracts and titles indexed automatically.

Requests

Authors of selected recent papers asked to state reason (in form of a question) for undertaking research leading to paper, and to provide other questions related to this research.

Relevance

By authors, for own cited papers; exhaustively by experts to obtain additional papers for author vetting. There were four relevance grades.* Relevance judgements were based on full text.

Document collection

Document set consisted of some recent papers, and their cited references, with some others.

2. UKCIS : see Barker, Veal and Wyatt 1974
Project on retrieval experiments based on Chemical Abstracts Condensates.

Size

193 queries; subset of 48
documents: CAC-1 11518 biochemistry, organic chemistry
CAC-2 15629 macromolecular, applied and physical chemistry
CBAC 1568 biochemistry
POST-J 1412 polymer science
POST-P 1442 polymer science.

Indexing source

Full texts/titles.

Index languages

Keywords applied manually; titles and digests effectively indexed automatically.

Requests

Formulations were from SDI service users and were current for the documents. Different versions were written for different data bases.

Relevance

Assessment of output (pooled if appropriate) by users, usually from titles plus digests, sometimes only titles. There were two relevance grades, and sometimes two non-relevance grades.

Document collection

Documents were taken from the current Chemical Abstracts Service files.

Comments

The assessed output averaged 230 items per query.

* the relevance grades do not include non-relevance as a grade.

3. NPL : see Vaswani and Cameron 1970
Project on statistical word associations and their use.

Size

93 queries
11571 documents.

Subject

Electronics and radio.

Indexing source

(Title +) abstracts.

Index languages

14, all automatic, representing various combinations of terms, weights and associations, for a 1000 stem vocabulary.

Requests

From subject experts, as suggested by the subject areas of abstracts randomly selected from the collection.

Relevance

By questioners, using a single grade, on abstracts for the pooled output of all the search strategies tested.

Document collection

Drawn from selected categories in an abstract journal.

4. Evans : see Evans 1975a,b
Project on methods of ranking and search strategy variations.

Size

55 queries ; several subsets, including 39 queries
20817 documents ; 2542 documents

Subject

Physics and electronics.

Indexing source

(Title +) abstracts.

Index languages

7, applied manually, representing various forms of profile term grouping and weighting; also Boolean.

Requests

Real, from recruits for a free SDI service.

Relevance

By questions on titles plus document index terms and classification codes, for the pooled output of top-ranking documents for each strategy. There were two grades.

Document collection

Current INSPEC SDI file.

A3 contd.

Requests

For all except the P collections request words are assimilated to stems in the document-derived vocabulary; words not so assimilated are rejected.

manually indexed requests:

C1400I 'significant' words were selected from the query texts following Cranfield instructions

U27000P, lists of word fragments (front and/or back truncated), words, E2500P and word strings (also end truncated) supplied

automatically indexed requests:

U27000T, words matched against the dictionary, hence processed like N11500A, documents with elimination of stop words, etc.
N11500T,
E2500T

Note The 1977 report by Sparck Jones and Bates covered experiments with old collections, which were somewhat heterogeneously processed. This report deals mostly with new collections processed in a systematic way. This apart from the old C1400I collection, and the profile (P) collections which require special treatment, the primary indexing for documents and requests is supplied as for the U27000T collection. The general approach to the provision of the formatted machine-readable data required for our programs is described in Section D of the 1977 report.

Misspellings are not corrected.

Figure A4 Test Collections

	Cranfield	UKCIS		NPL		Evans	
	C1400I	U27000T	U27000P	N11500A	N11500T	E2500T	E2500P
<u>Documents</u>							
max	102	24	see	105	56	19	see
min	5	1	Figure	1	1	1	Figure
nos	1400	27361	A5	11429	11416	2542	A5
tot	41841	180055		228087	65515	16120	
av	29.9	6.6		20.0	5.7	6.3	
<u>Requests</u>							
max	17	35	"	14	14	53	"
min	3	1		2	2	17	
nos	225	182		93	93	39	
tot	1783	1273		671	671	1263	
av	7.9	7.0		7.2	7.2	32.4	
<u>Relevance judgements</u>							
max	40	554	same	84	same	53	same
min	1	1		1		3	
nos	225	182		93		39	
tot	1614	10715		2083		899	
av	7.2	58.9		22.4		23.1	
<u>Inverted documents = terms</u>							
max	845	2345	see	2511	895	228	see
min	1	1	Figure	1	1	1	Figure
nos	2683	17537	A5	7491	3866	3288	A5
tot	41841	180055		228087	65515	16120	
av	15.6	10.3		30.4	16.9	4.9	
<u>Inverted requests</u>							
max	54	25	"	19	19	22	"
min	1	1		1	1	1	
nos	486	636		340	340	633	
tot	1783	1273		671	671	1263	
av	3.7	2.0		2.0	2.0	2.1	
<u>Inverted relevance judgements</u>							
max	8	7	same	6	same	5	same
min	1	1		1		1	
nos	831	7883		1735		633	
tot	1614	10715		2083		899	
av	1.9	1.4		1.2		1.4	

Figure A5 The Profile Collections

		UKCIS		Evans
		U27000P	U27000Pb	E2500P
<u>Documents</u>	max	-	-	-
	min	-	-	-
	nos	27361	27361	2542
	tot	-	-	-
	av	-	-	-
<u>Profiles</u>	max	158	86	112
	min	1	1	25
	nos	182	75	39
	tot	5353	1645	1871
	av	29.4	21.9	48.0
<u>Inverted documents</u>	-			
<u>Inverted profiles</u>	-			
<u>Profile terms per document</u>				
	max		8	7
	min		1	1
	av		1.2	1.2
<u>Documents per profile term</u>				
	max		2340	957
	min		1	1
	av		118.1	16.0

N.B. data includes NOT terms

Figure A6 Collection Comparison

<u>Collection</u>	No docu- ments	No terms	Av terms per doc	Av docs per terms	No reqs	No terms in reqs & % tot	Av terms per req	Av reqs per terms in reqs	Av docs per terms in reqs	No rel. docu- ments	No docs in rel- docs & % tot	Av rel- docs per req	Av reqs per rel- doc
C1400I	1400	2683	29.9	15.6	225	486 18.1	7.9	3.7	162.0	1614	831 59.4	7.2	1.9
U27000T	27361	17537	6.6	10.3	182	636 3.6	7.0	2.0	285.6	10715	7883 28.8	58.9	1.4
U27000P	27361	-	-	-	182	-	29.4	-	-	10715	7883 28.8	58.9	1.4
N11500A	11429	7491	20.0	30.4	93	340 4.5	7.2	2.0	567.4	2083	1735 15.2	22.4	1.2
N11500T	11416	3866	5.7	16.9	93	340 8.8	7.2	2.0	211.8	2083	1735 15.2	22.4	1.2
E2500T	2542	3288	6.3	4.9	39	591 18.0	32.4	2.1	29.2	899	633 24.9	23.1	1.4
E2500P	2542	-	-	-	39	-	48.0	-	16.0	899	633 24.9	23.1	1.4

Figure A7 Subsidiary Collections

		No docu- ments	No re- quests	No rel- docs	Av rel- docs per req
<u>Cranfield</u>	C1400e	700	225	834	3.7
	C1400o	700	225	780	3.5
<u>UKCIS</u>	U27000Te	13680	182		
	U27000To	13681	182		
	U27000Tf	11613	182	4058	22.3
	U27000Tl	15748	182	6657	36.6
	U27000T3/4	20521	182	8036 *	44.2 *
	U27000T1/2	13681	182	5358 *	29.4 *
	U27000T1/4	6841	182	2679 *	14.7 *
	U27000T1/8	3421	182	1339 *	7.4 *
	U27000T1/16	1711	182	670 *	3.7 *
	U27000TS1/4	6840	182	2702	14.8
	U27000Tb } U27000Pb }	27361	75	3739	49.9
	U27000Tbe } U27000Pbe }	13680	75	1837	24.5
	U27000Tbo } U27000Pbo }	13681	75	1902	25.4
	U27000Pbf	11613	75	1444	19.3
	U27000Pbl	15748	75	2295	30.6
	U27000Pb3/4	20521	75	2804 *	37.4 *
	U27000Pb1/2	13681	75	1870 *	24.9 *
	U27000Pb1/4	6841	75	935 *	12.5 *
	U27000Pb1/8	3421	75	467 *	6.2 *
	U27000Pb1/16	1711	75	234 *	3.1 *
U27000Pbs1/4	6840	75	902	12.0	
<u>NPL</u>	N11500Ae	5714	93	1022	11.0
	N11500Te	5706			
	N11500Ao	5715	93	1061	11.4
	N11500To	5710			
<u>Evans</u>	E2500Te } E2500Pe }	1271	39	443	11.4
	E2500To } E2500Po }	1271	39	456	11.7

* These figures are estimated as percentages of the whole set.

Note that ^{some requests} may not have any relevant documents in a given subset; if this is allowed for slightly different averages are obtained.

Figure A8 Term Retrieval Facts

<u>Collection</u>	Av terms per req	Av max match- ing terms per req	Av match- ing terms per doc	Av match- ing terms per rel- doc	No rel- docs retr. & % tot	Av docs retr. per req	Av rel- docs retr. per req
C1400I	7.9	5.7	1.7	3.5	1549 96.0	746.7	6.9
U27000T	7.0	2.8	1.1	1.4	4544 42.4	1841.4	25.0
U27000P	29.4	3.7	1.2	1.8	8844 82.5	2668.7	48.6
U27000Pb	21.9	2.9	1.1	1.4	2820 75.4	1481.5	37.6
N11500A	7.2	4.7	1.3	2.8	2034 97.6	3171.0	21.9
N11500T	7.2	3.7	1.1	1.8	1645 79.0	1329.0	17.7
E2500T	32.4	4.4	1.3	1.9	678 75.4	749.5	17.4
E2500P	48.0	4.2	1.2	1.9	725 80.6	477.6	18.6

The profile figures refer to simple term matching with any NOT terms removed from the profile term lists

Figure A9 Performance Representation Methods

A Document value (dv)

recall/precision figures obtained by averaging by numbers across matching values, with precision at standard recall levels subsequently obtained by linear interpolation.

B Recall cutoff (rc)

recall/precision figures obtained by averaging over precision computed for individual requests at standard recall levels using pessimistic interpolation (with fully and completely ranked output).

Document rank (dr)

recall/precision figures obtained as for dv, but averaging across matching ranks (fully and completely ranked output).

Precision rank (pr)

recall/precision figures obtained by averaging by numbers for specific ranks (fully and completely ranked output).

C Total retrieved (tr)

average relevant and total documents retrieved, by the lowest matching value, obtained by averaging by numbers.

Relevant rank (rr)

average numbers of relevant documents retrieved at specific ranks, obtained by average of numbers (fully and completely ranked output).

Cumulative requests (cr)

cumulative proportion of requests retrieving a relevant document by specific ranks (fully and completely ranked output).

Figure B1 Special Cases for Weighting Functions

Definition	Documents in which term occurs	Functions	Implications for document when term is:	
			present	absent
$r = 0$	non-relevant only	F1, F4	bad	indifferent
$n-r = 0$	relevant only	F4	good	indifferent
$R-r = 0$	all relevant & some others	F4	indifferent	bad
$N-n-R+r = 0$	some relevant & all others	F4	indifferent	good
$n-r = 0$ $R-r = 0$	} all relevant & no others	F4	good	bad
$r = 0$ $N-n-R+r = 0$				

bad : the document should never be retrieved, i.e. should be at bottom rank

good : the document should always be retrieved, i.e. should be at top rank

indifferent: the document should be unaffected, i.e. should be at the rank determined by its other terms

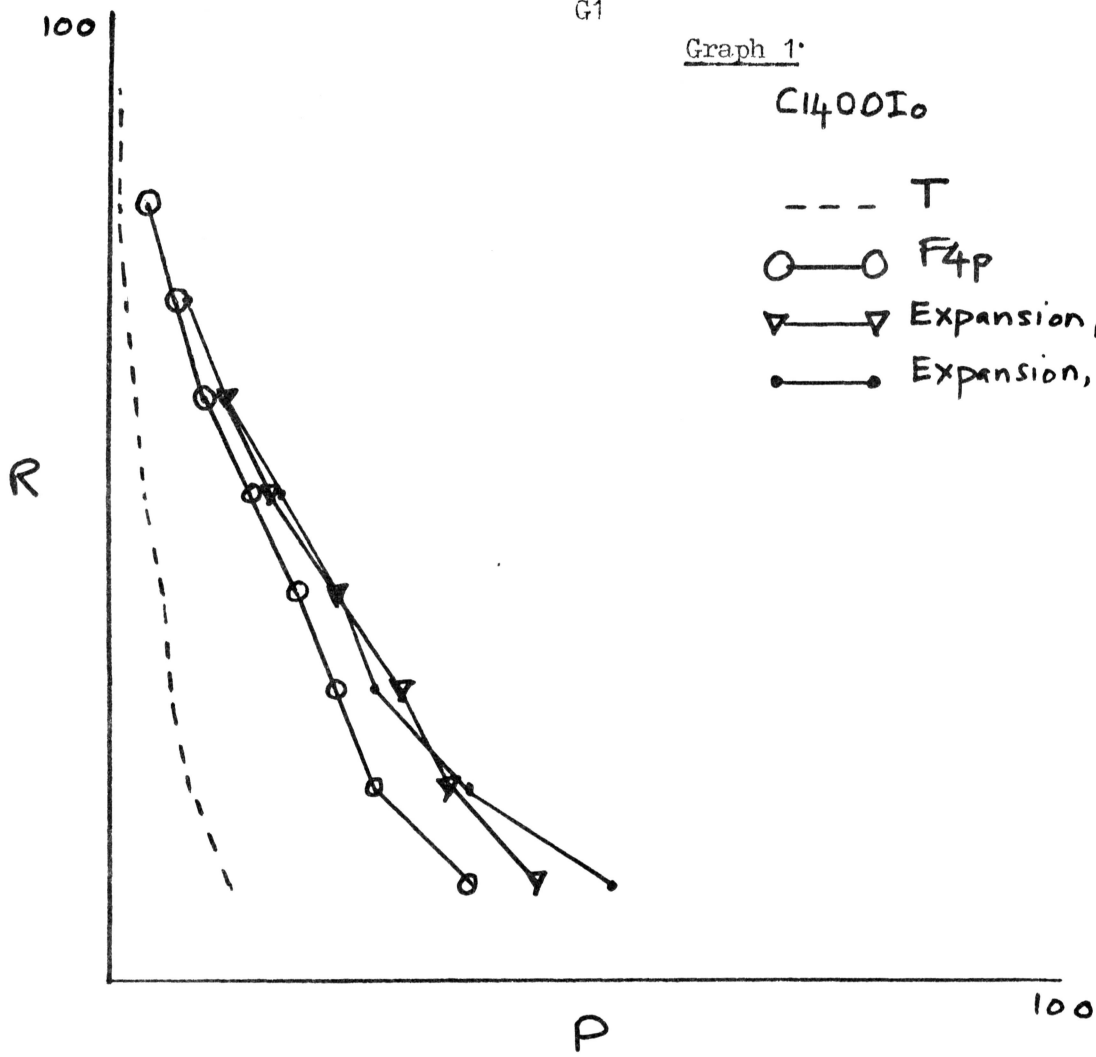
Figure B2 Numbers of Retrieved Documents

	score 1			recall 30%		
	av. rel	av. nrel	av. nrel per rel	av. rel	av. nrel	av. nrel per rel
<u>C1400Io</u>						
T	3.3	375.4	113.7	1.1	13.5	12.8
FIRST1 F4p	2.8	64.0	23.2	1.0	4.0	3.9
FIRST2 F4p	2.8	67.4	24.1	1.1	3.2	3.0
BEST1 F4p	2.5	48.8	19.8	1.1	4.3	4.0
BEST2 F4p	2.6	52.8	20.6	1.1	3.6	3.4
ALL F4p	2.8	64.1	22.7	1.1	3.3	3.1
ALL F4r	3.1	39.9	12.7	1.1	1.0	1.0
<u>U27000Fbo</u>						
T	18.9	724.2	38.3	7.6	173.3	22.7
FIRST1 F4p	10.5	114.5	11.0	7.6	20.9	2.7
FIRST2 F4p	12.0	146.5	12.2	7.7	23.9	3.1
FIRST3 F4p	14.4	183.7	12.8	7.6	25.9	3.4
BEST1 F4p	9.6	93.5	9.7	7.8	35.0	4.5
BEST2 F4p	11.7	116.9	10.0	7.7	22.6	2.9
BEST3 F4p	12.4	131.4	10.6	7.8	22.4	2.9
ALL F4p	18.7	335.5	18.0	7.8	9.6	1.2
ALL F4r	18.8	336.6	17.9	7.8	3.5	0.4

The slight variations in the numbers of relevant documents at standard recall are due to the use of interpolated document numbers.

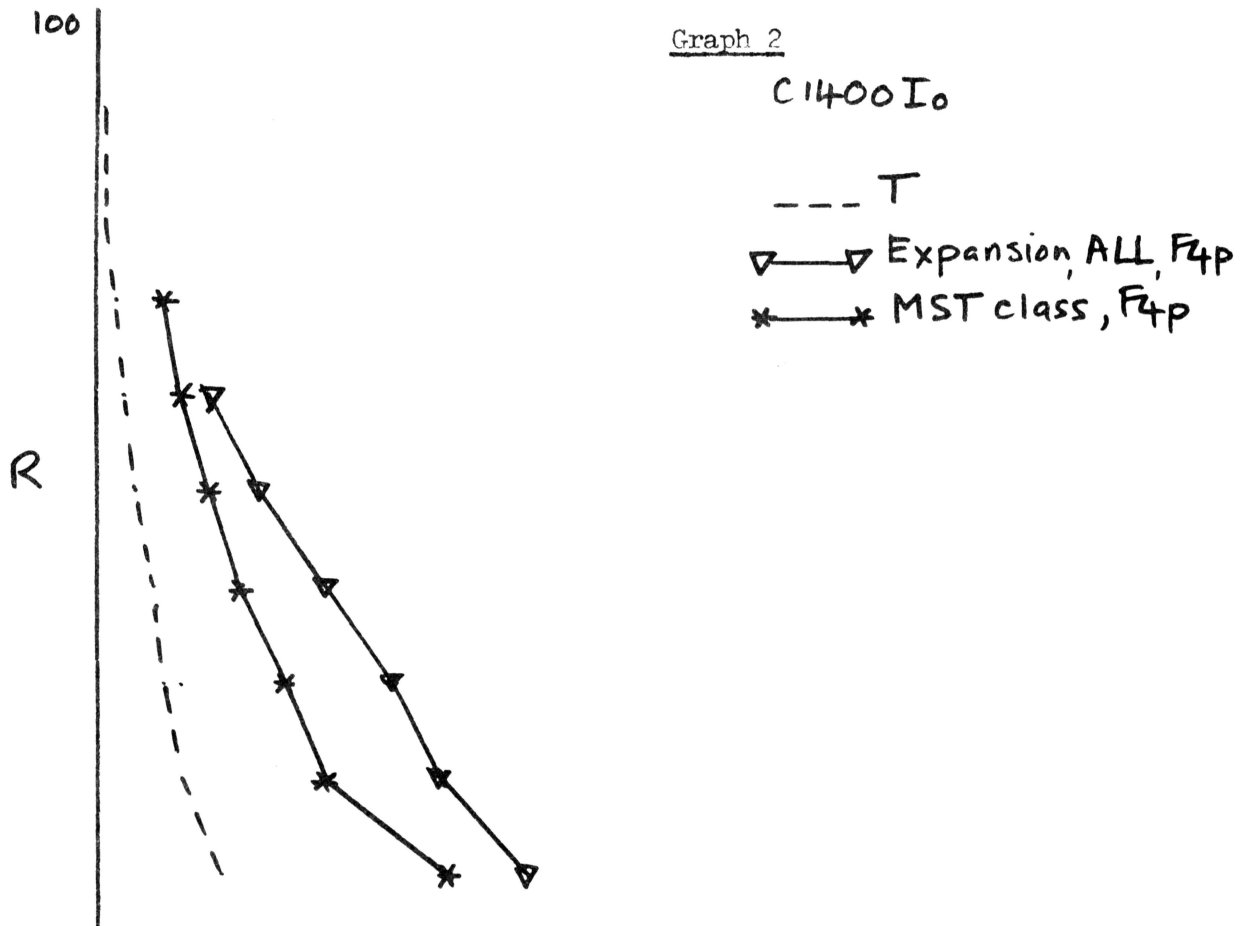
Graph 1

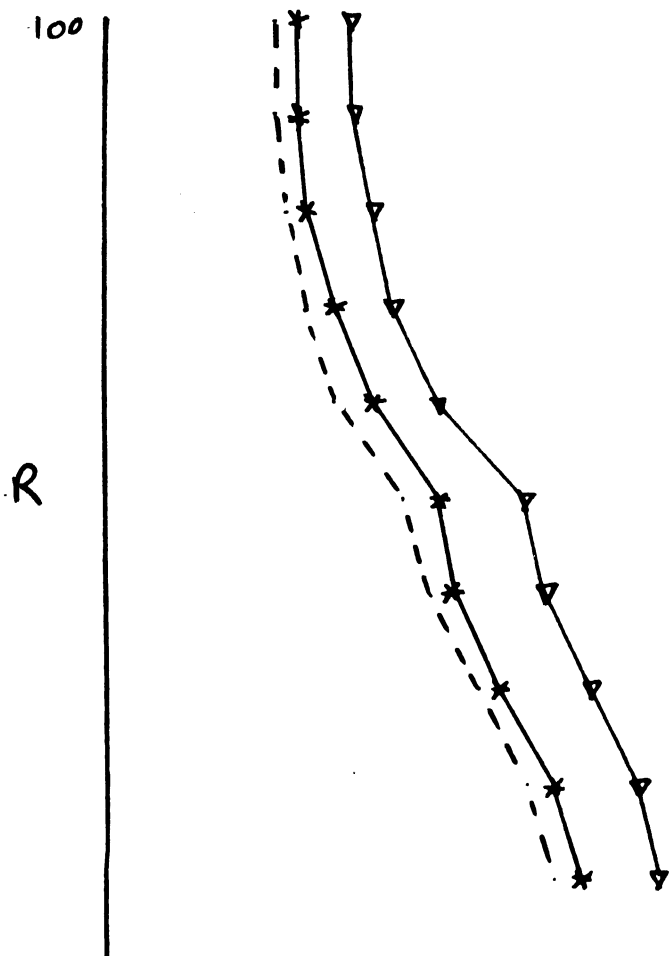
C1400I₀



Graph 2

C1400I₀





Graph 3

C1400 I₀

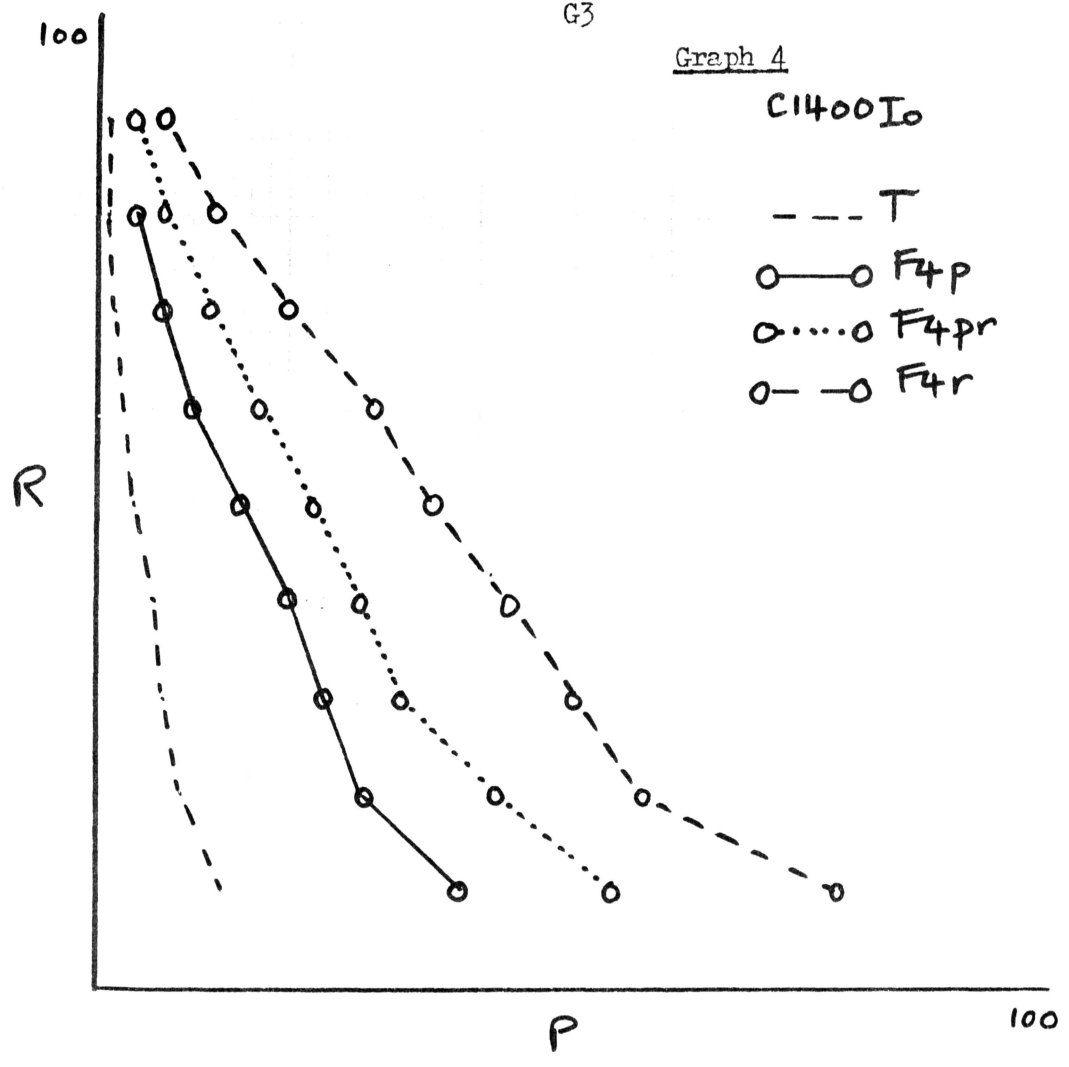
-- T

▽ Expansion, ALL, F4p

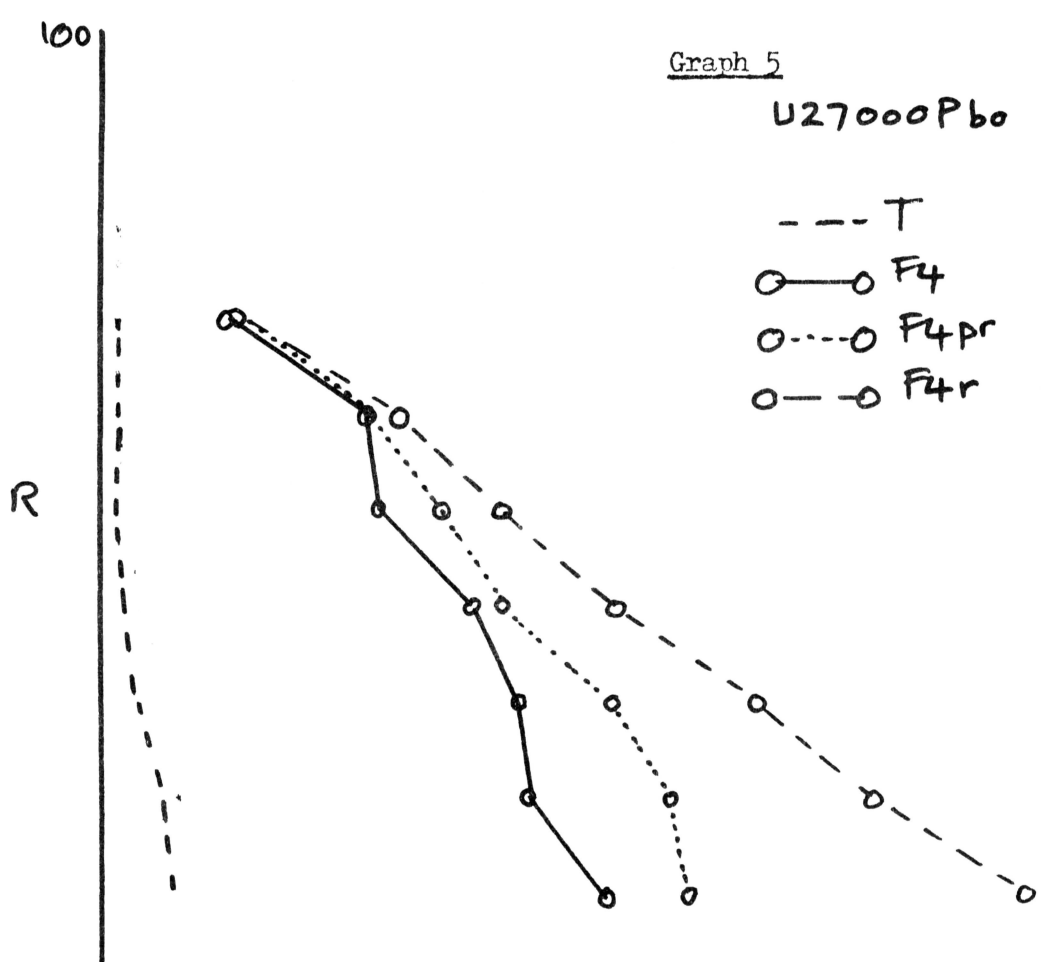
* MST class, F4p

recall cutoff

Graph 4

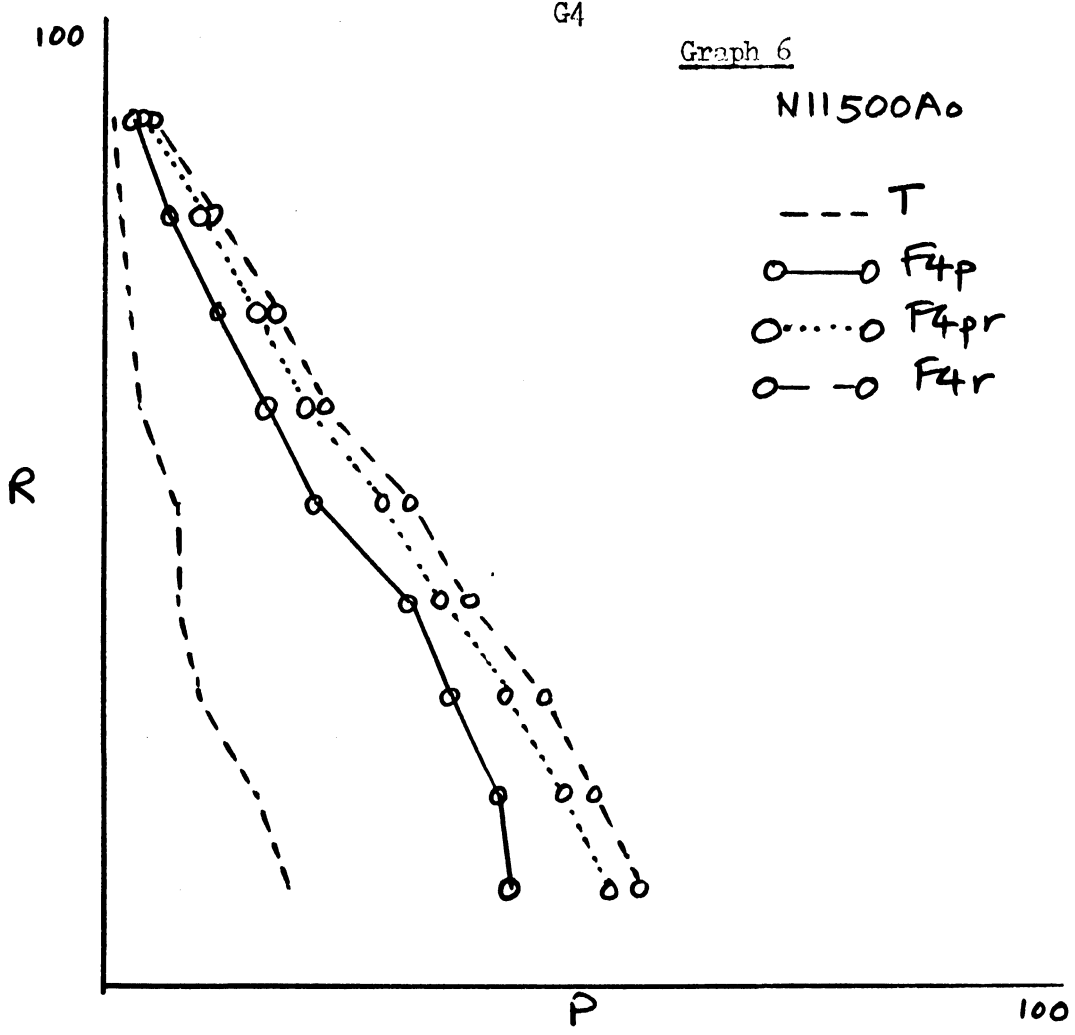


Graph 5



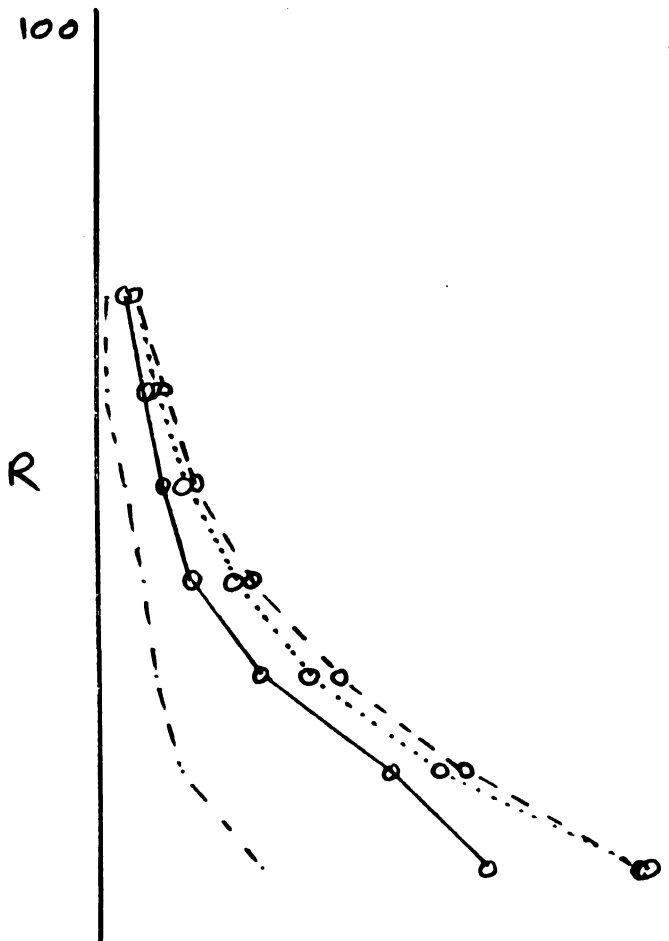
Graph 6

N11500A0



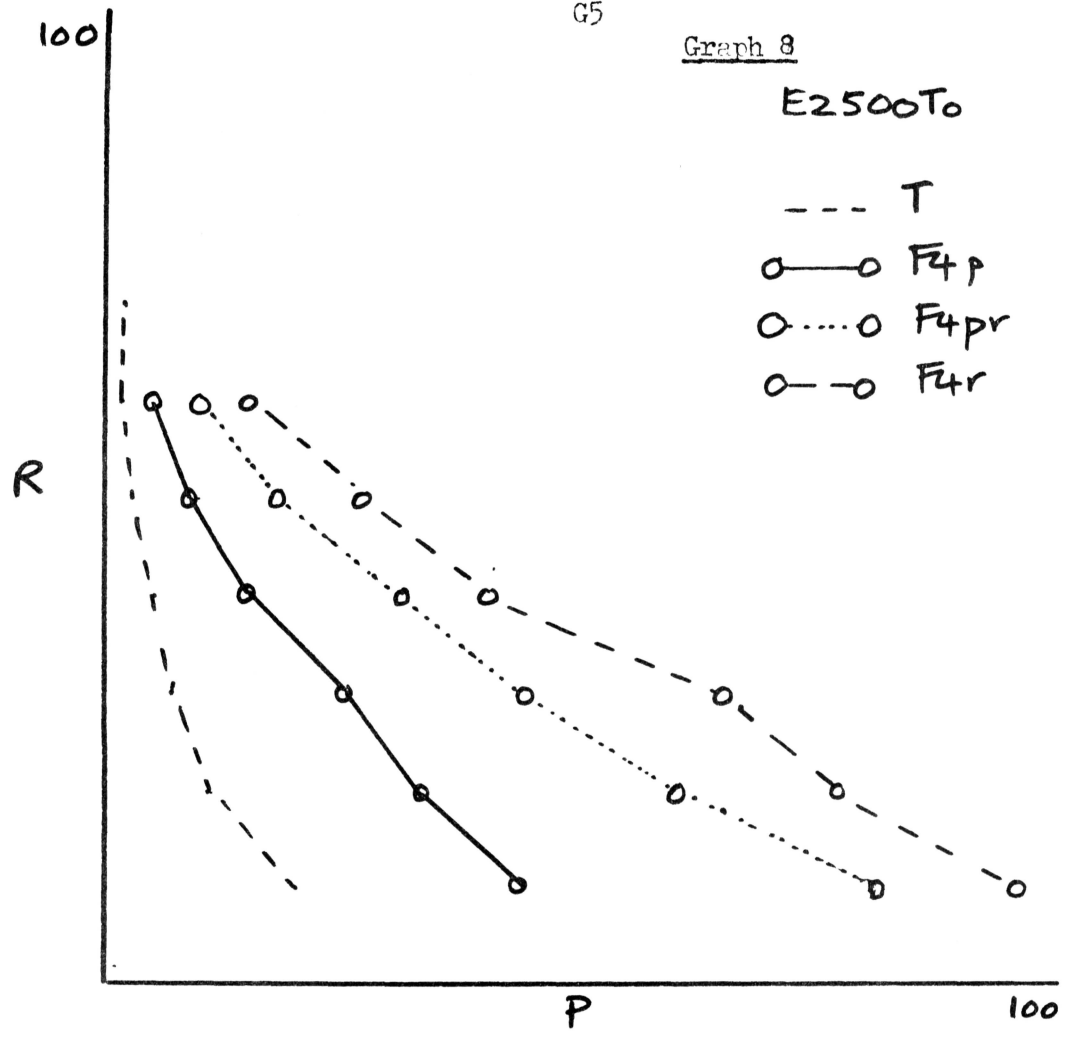
Graph 7

N11500T0



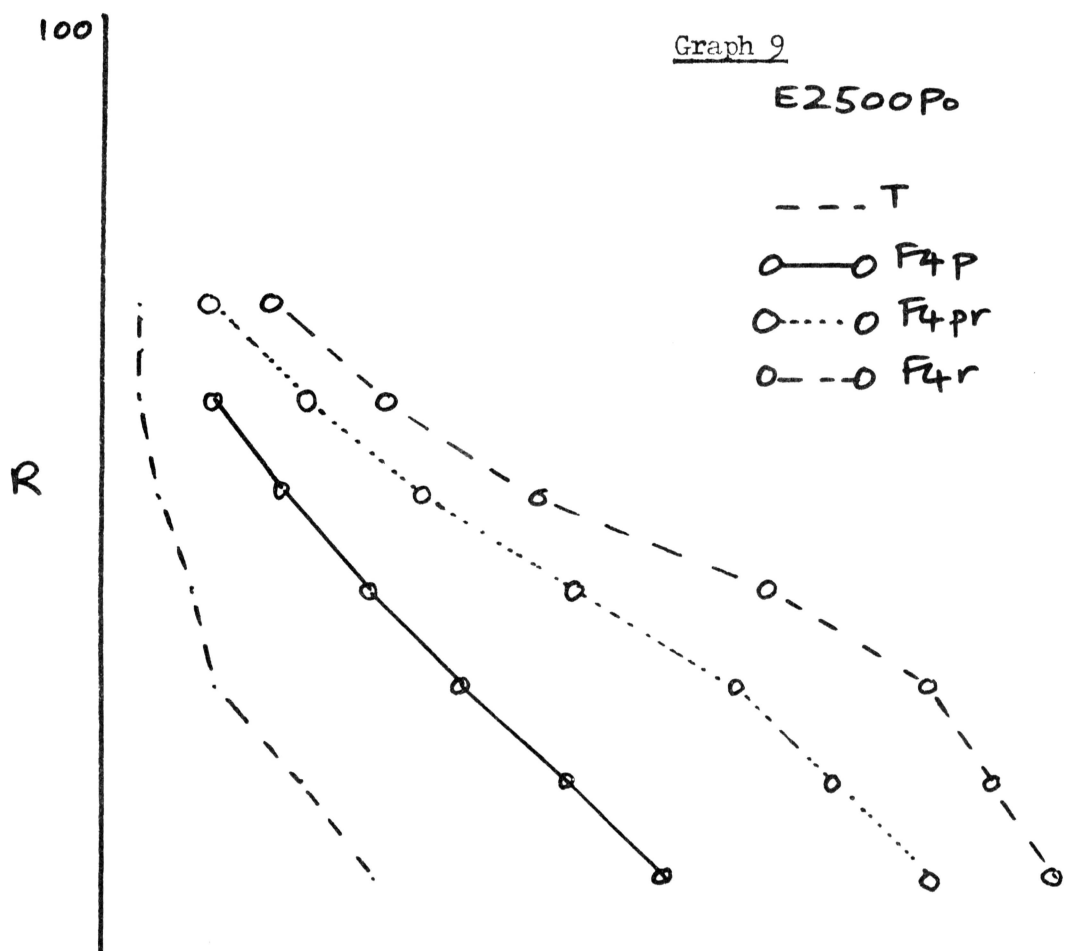
Graph 8

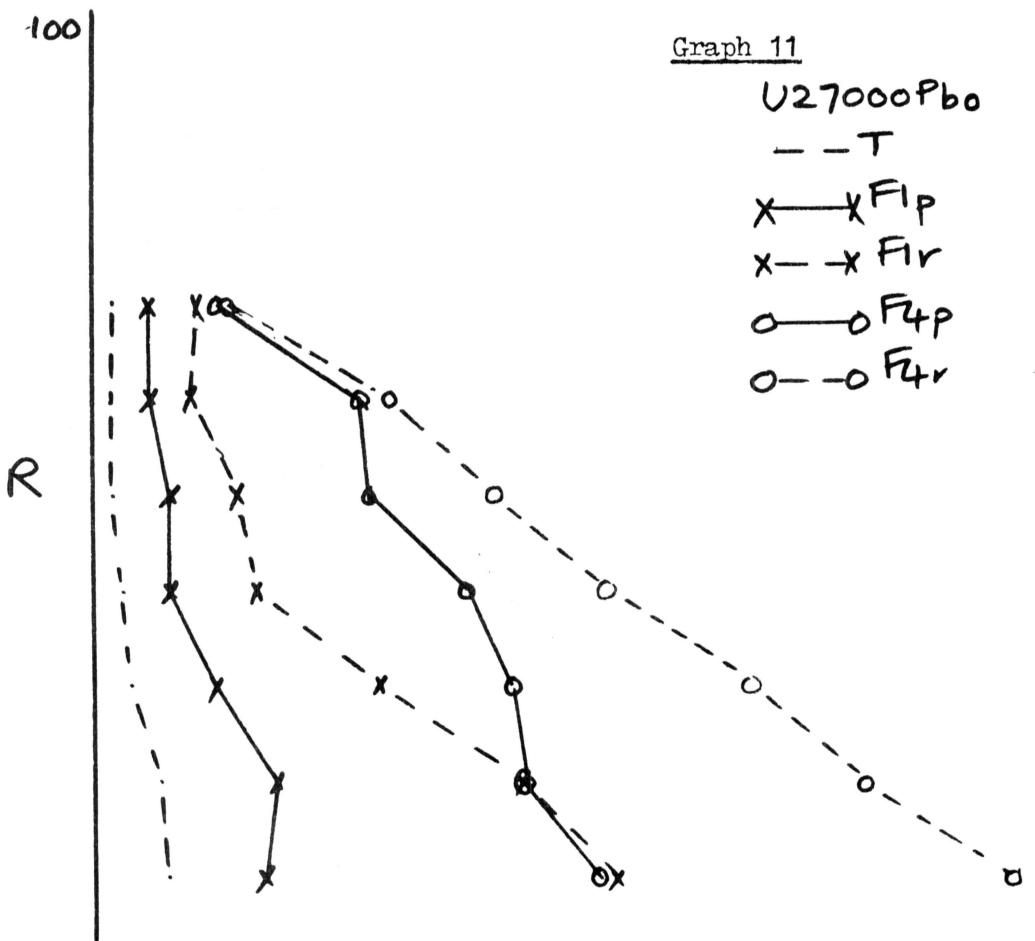
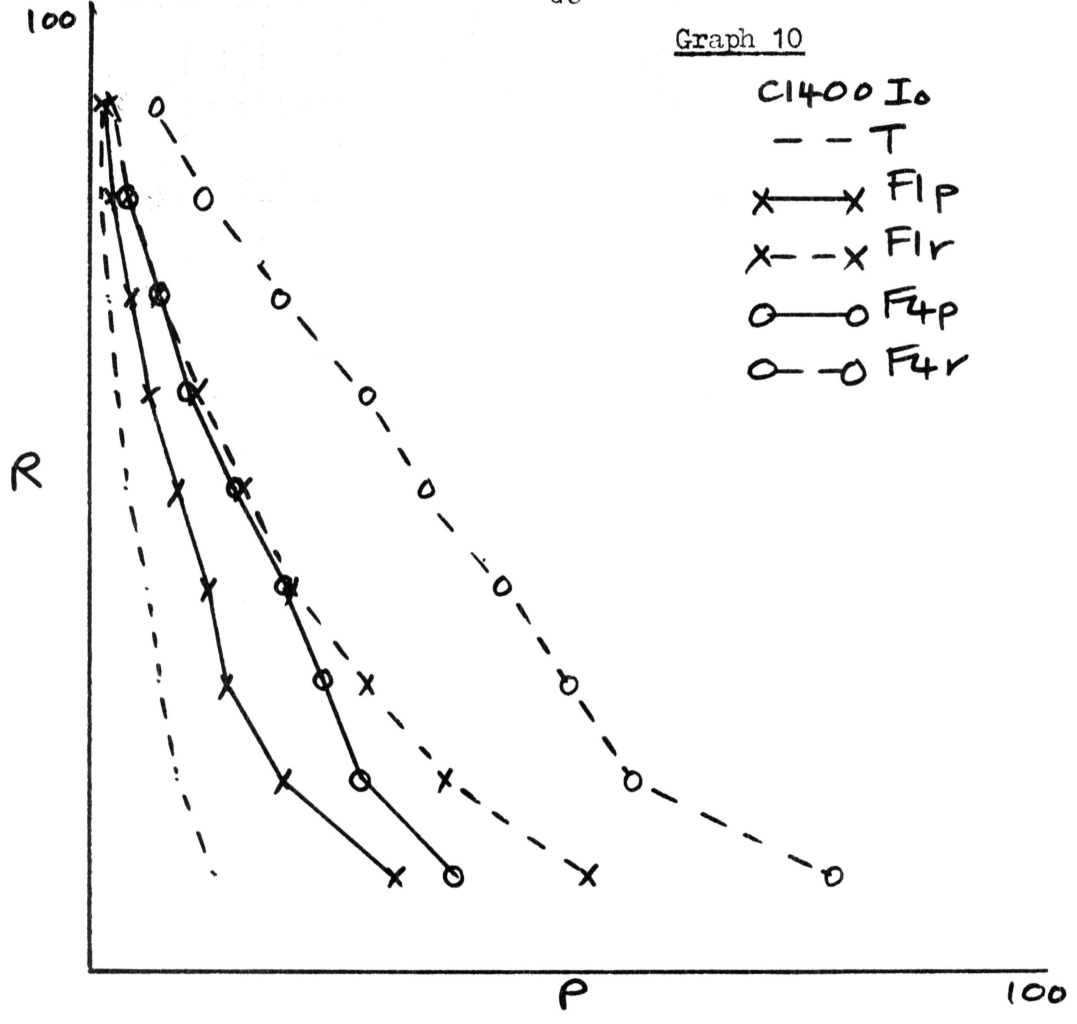
E2500T₀



Graph 9

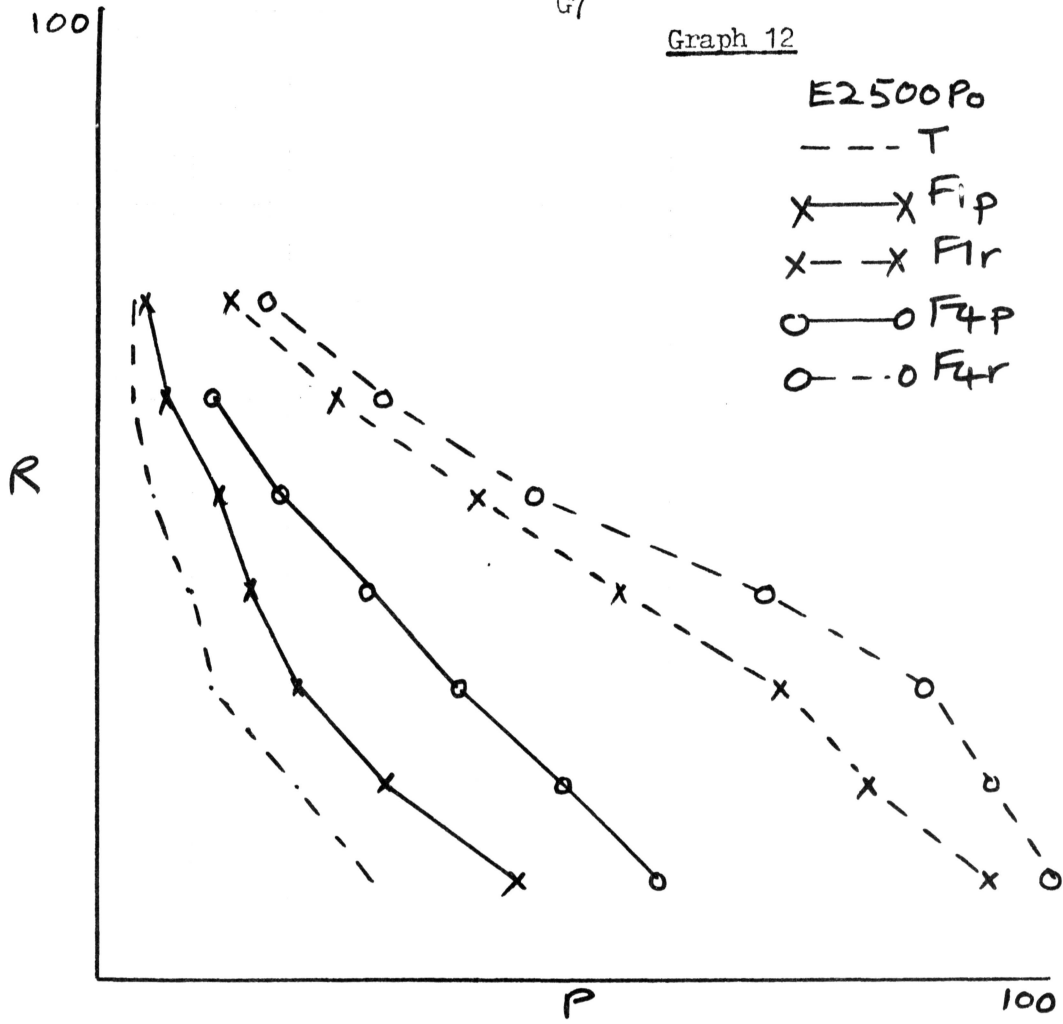
E2500P₀





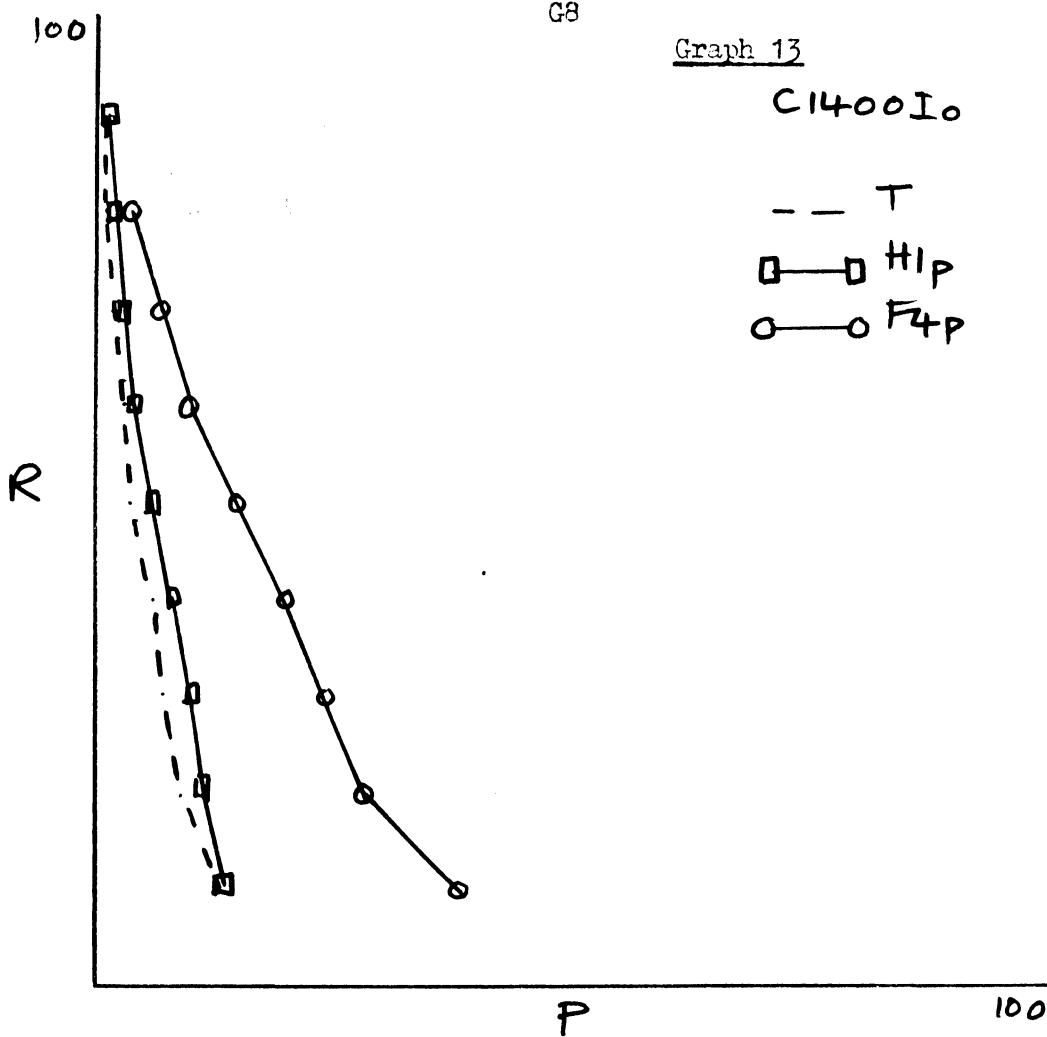
G7

Graph 12



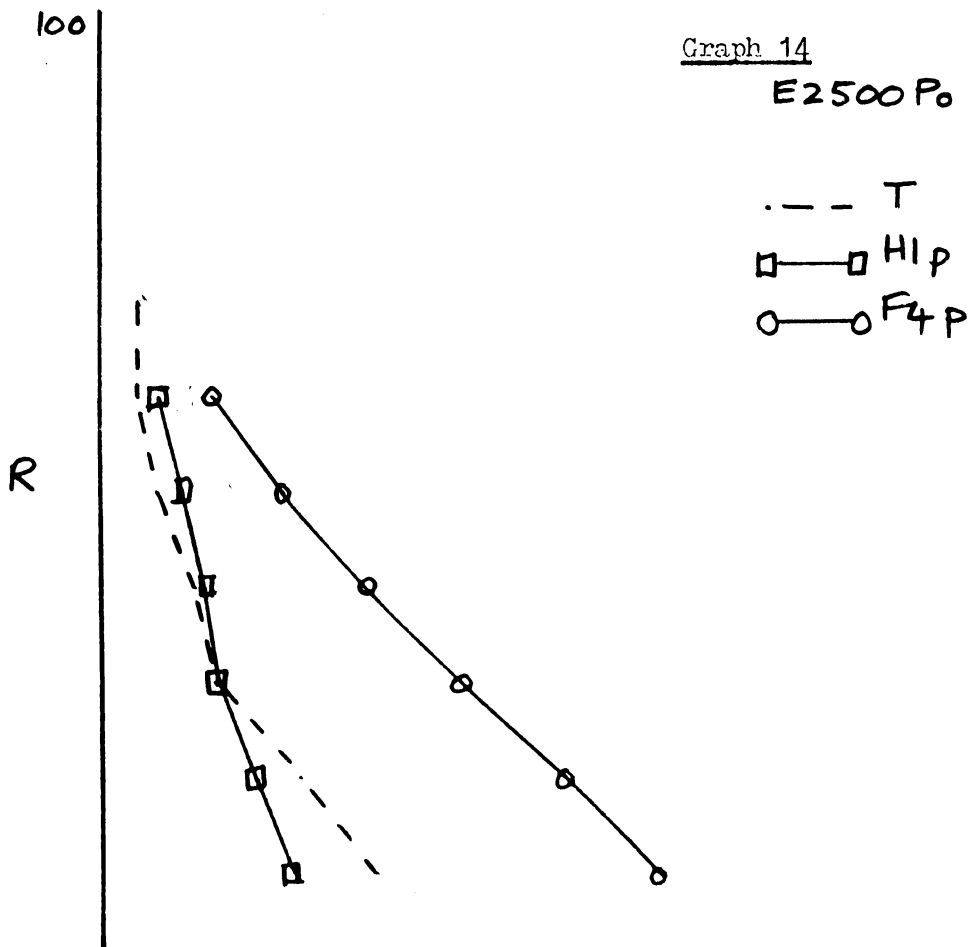
Graph 13

C1400I₀

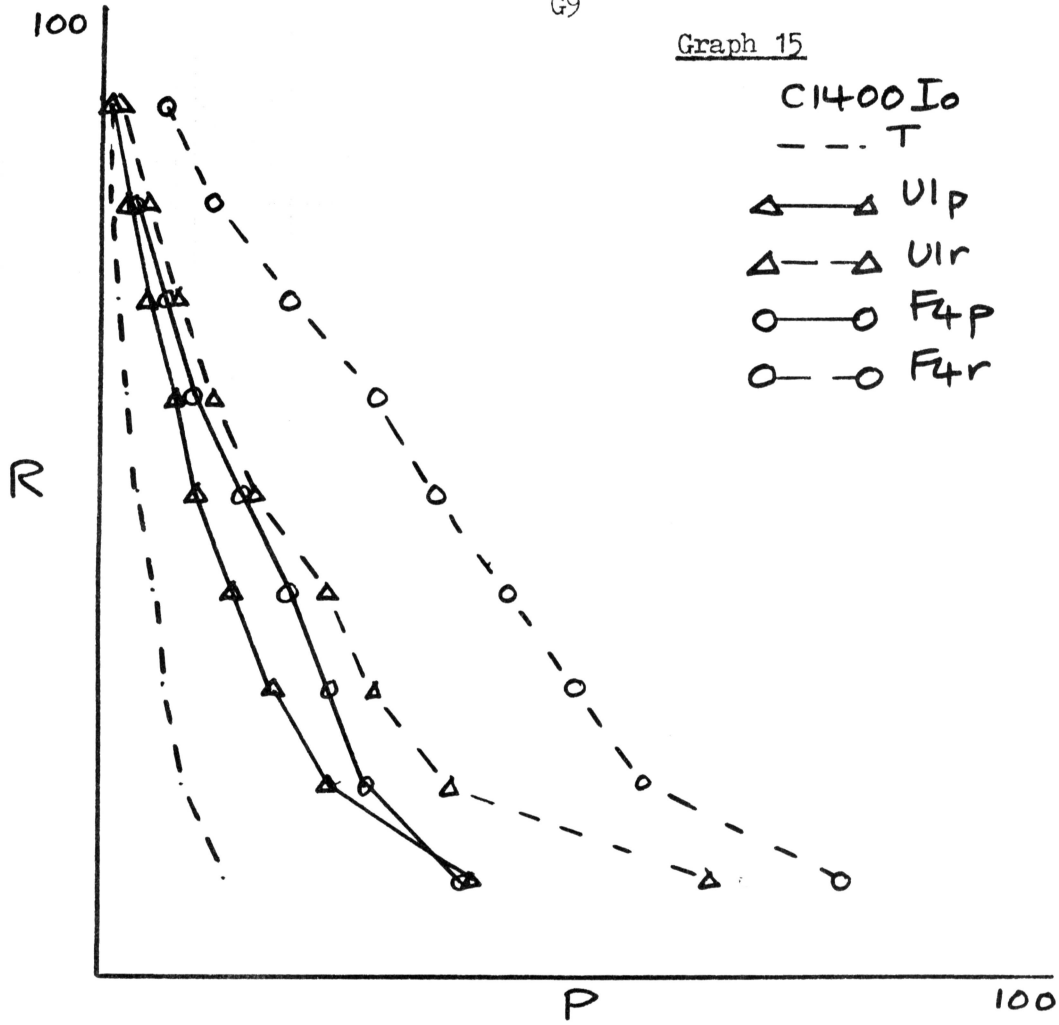


Graph 14

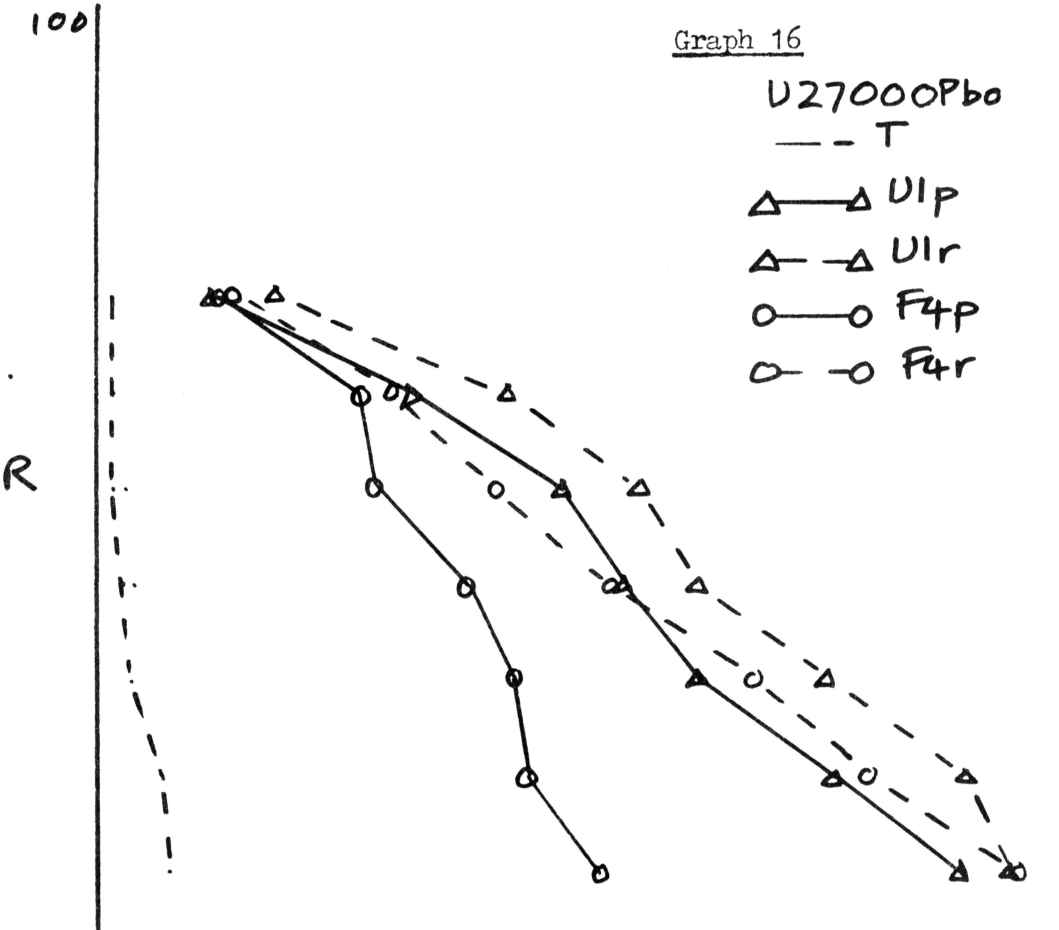
E2500P₀



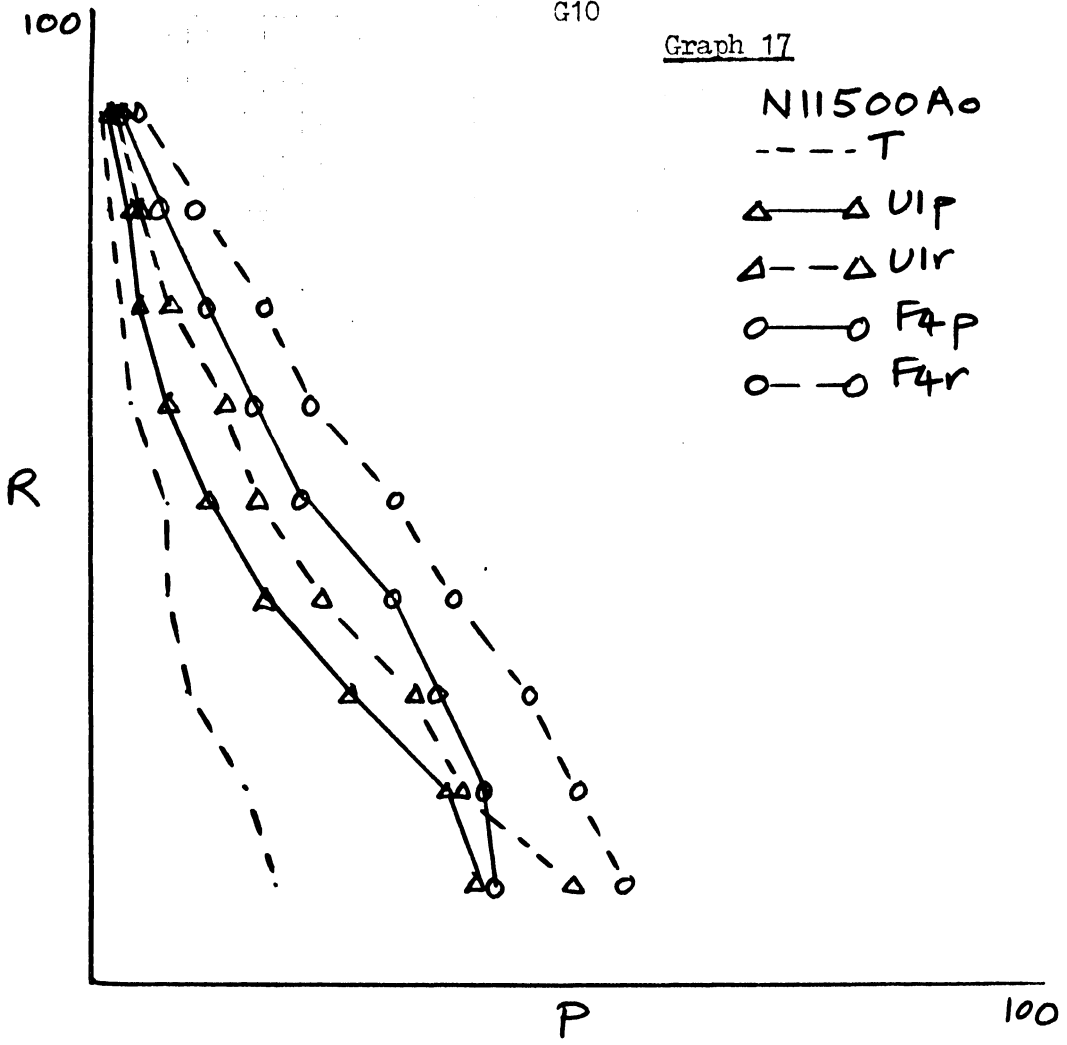
Graph 15



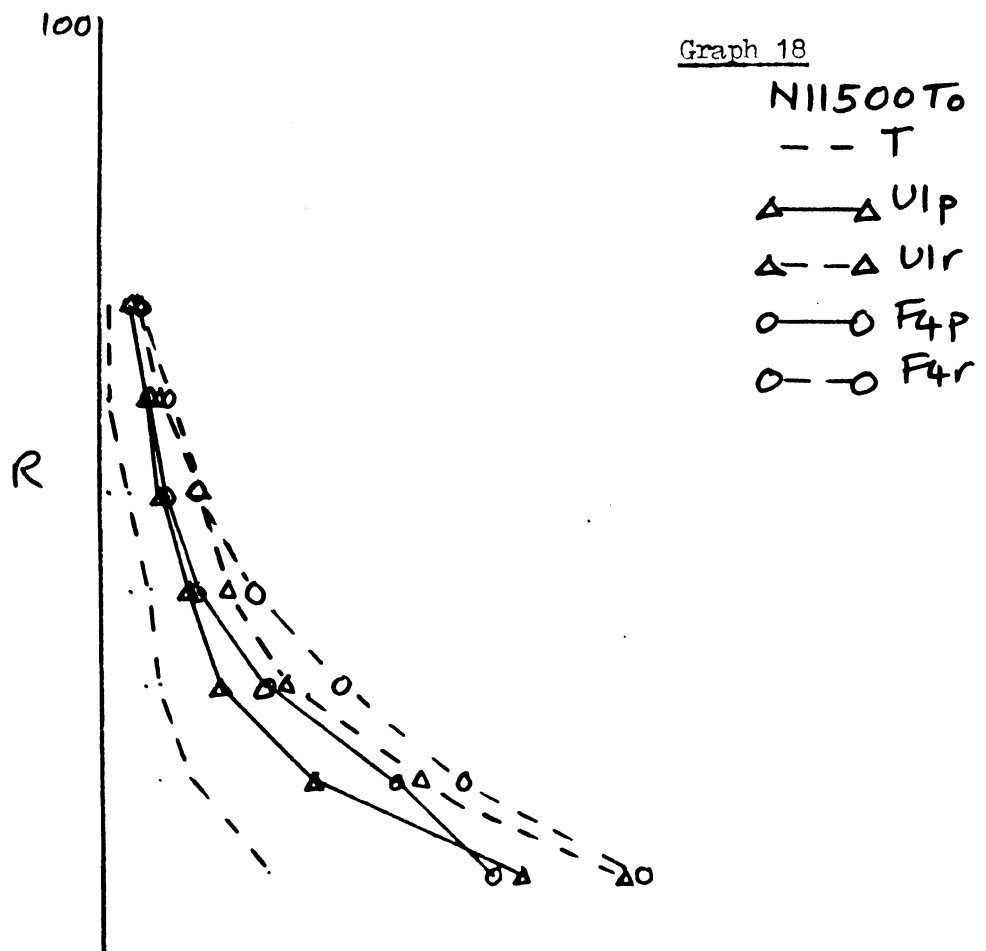
Graph 16



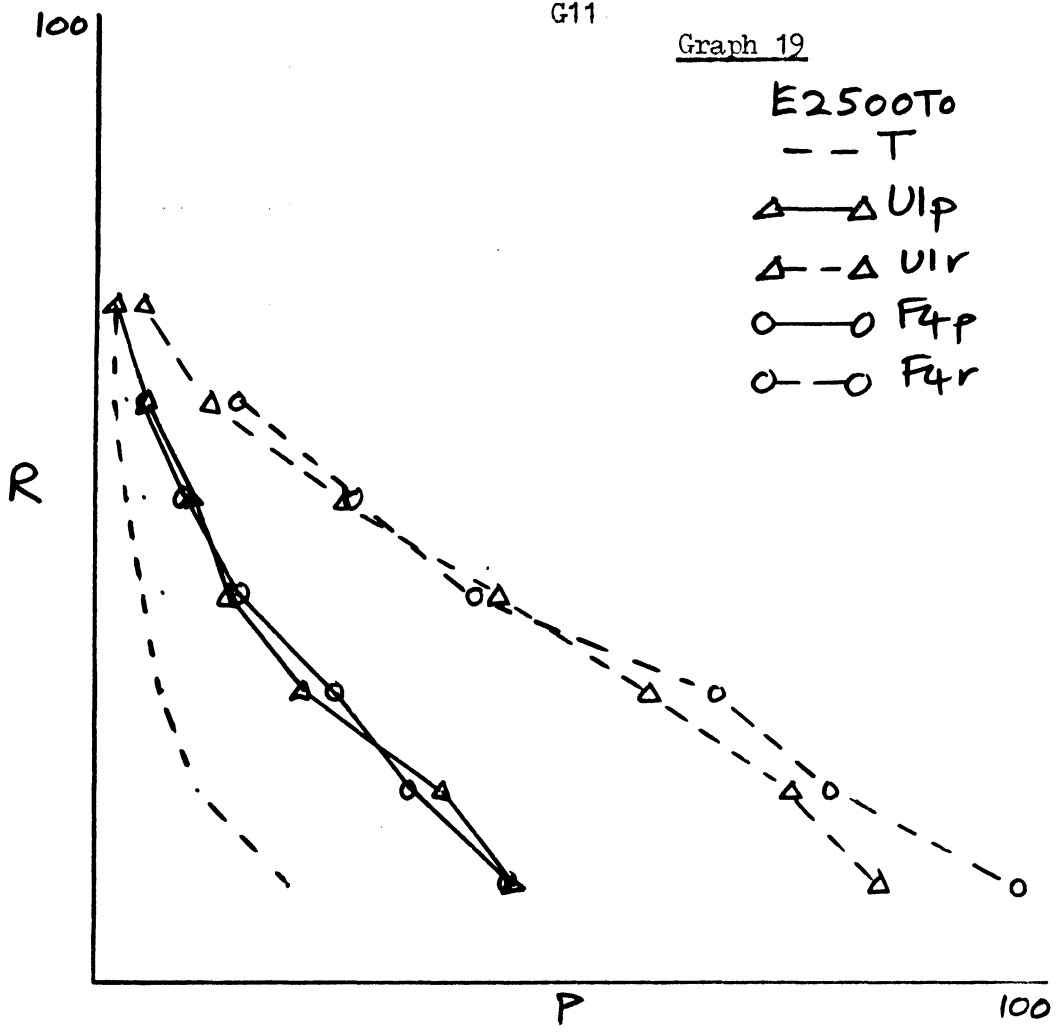
Graph 17



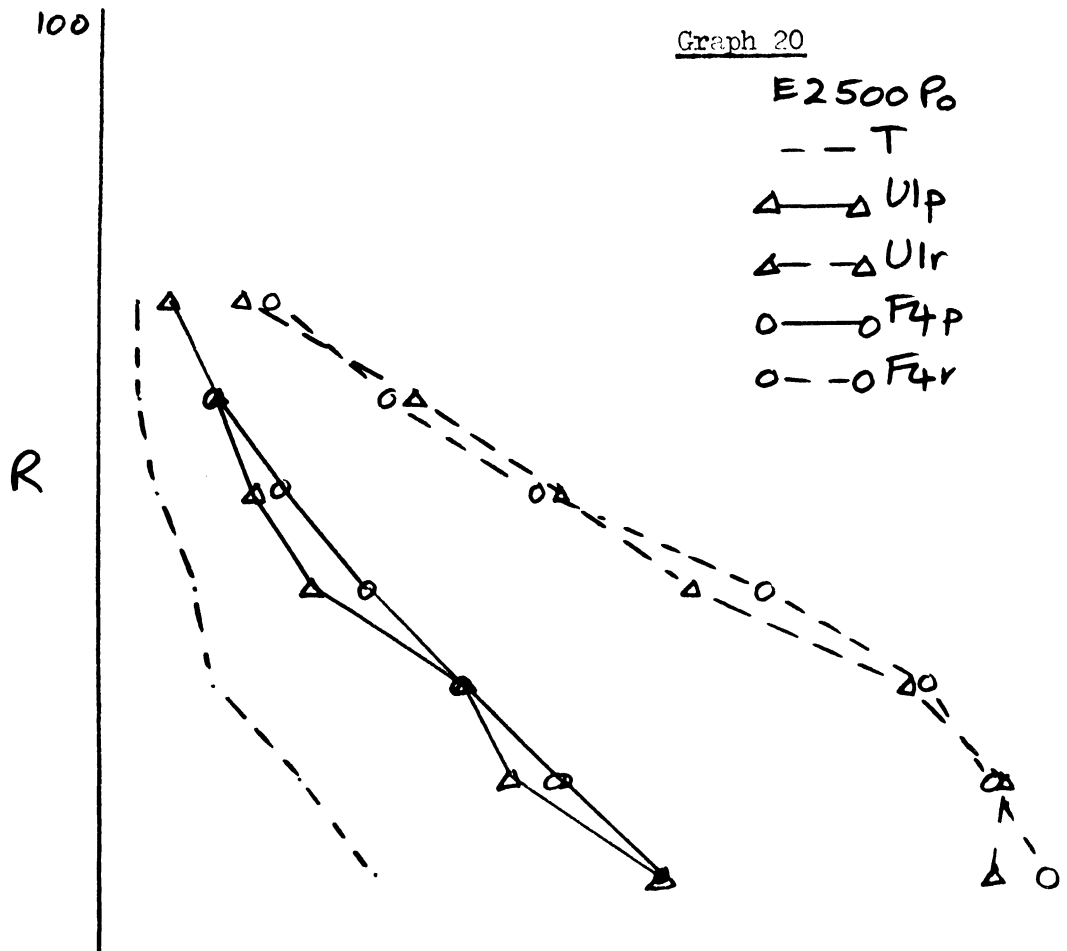
Graph 18

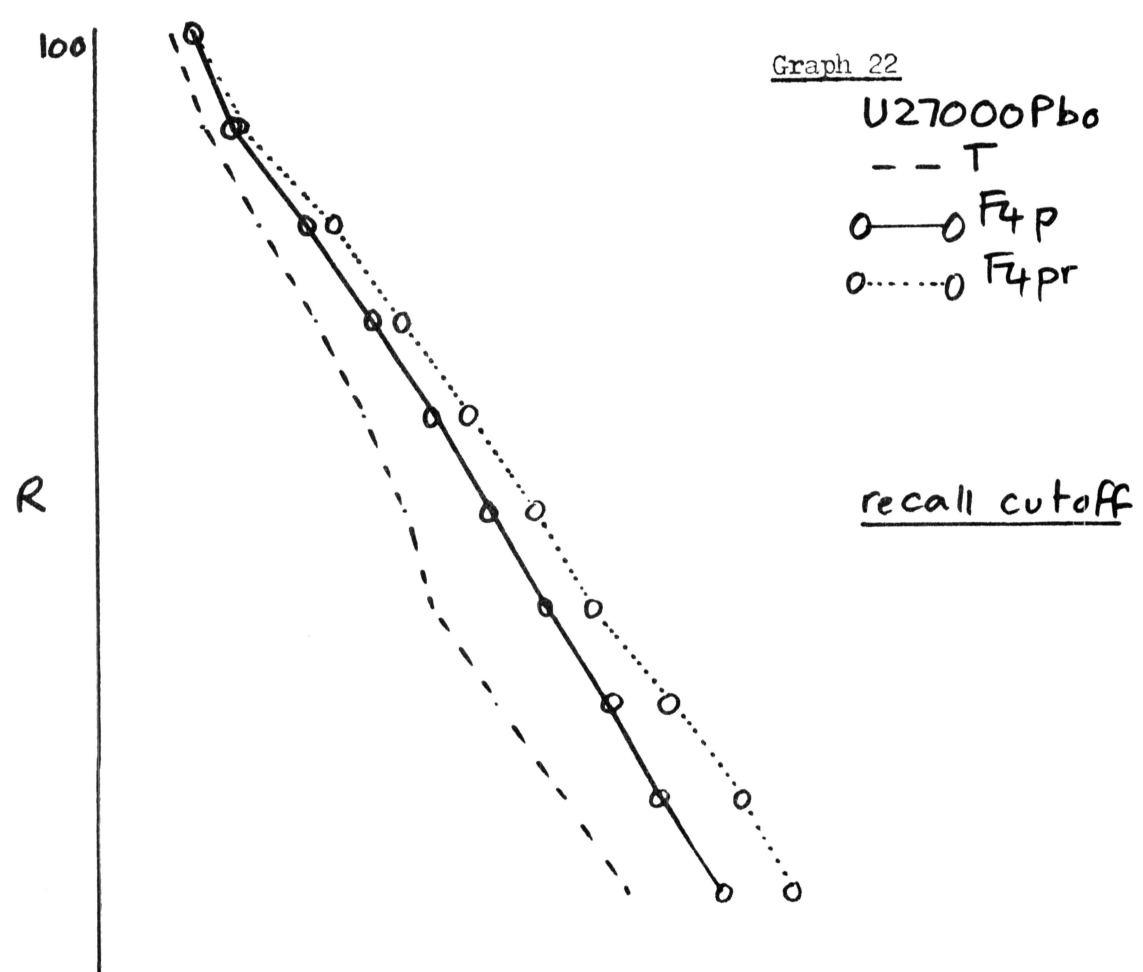
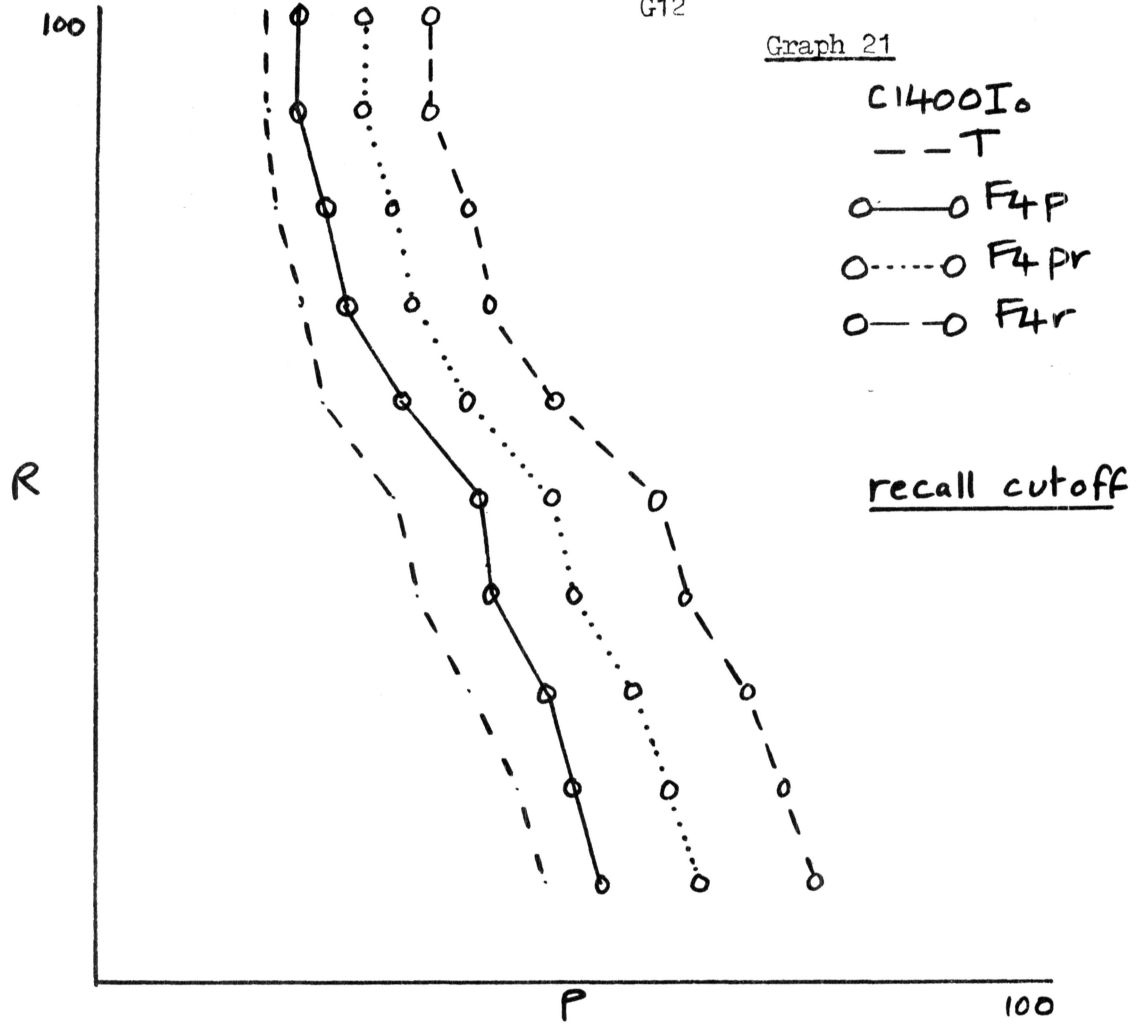


Graph 19

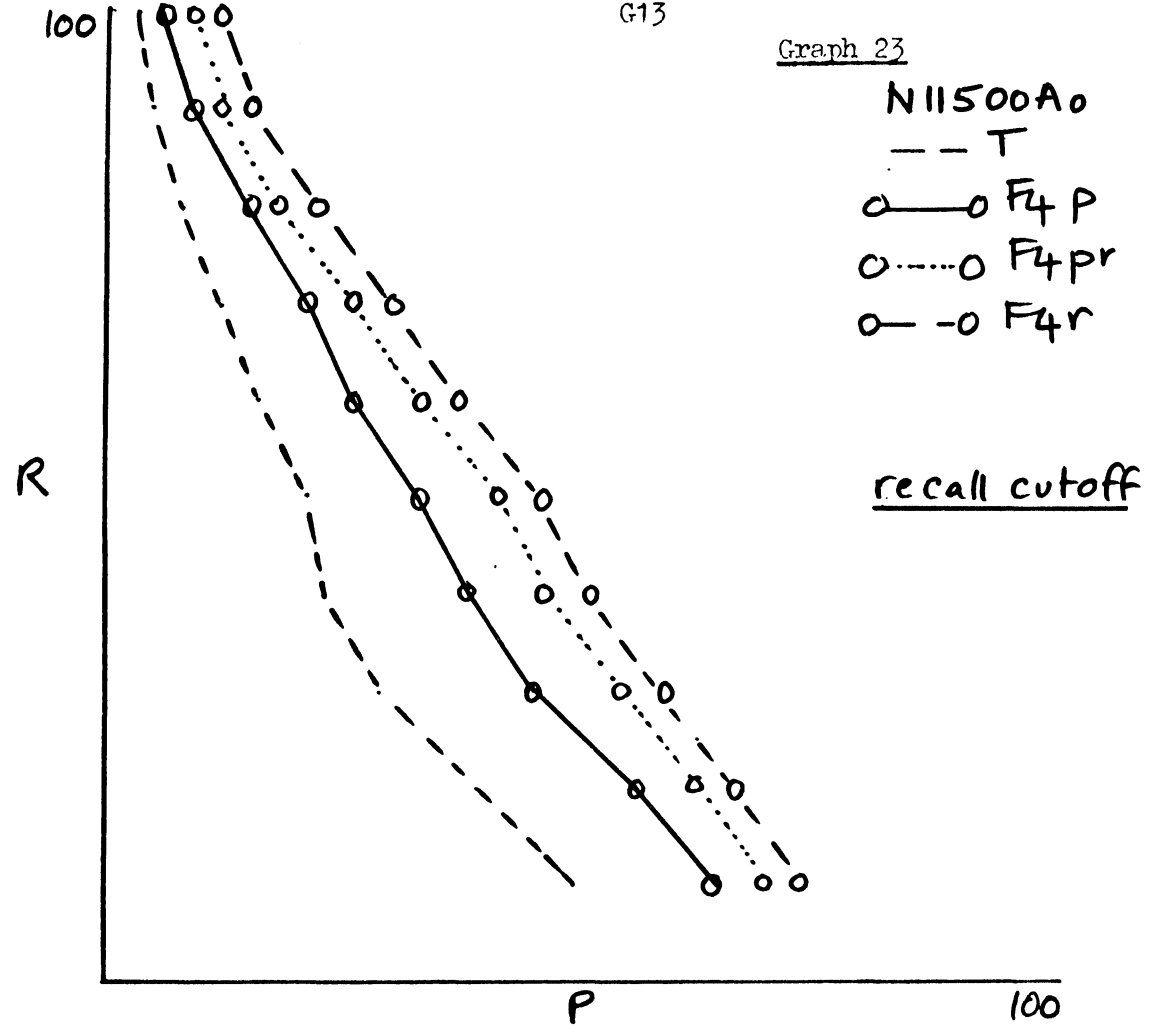


Graph 20

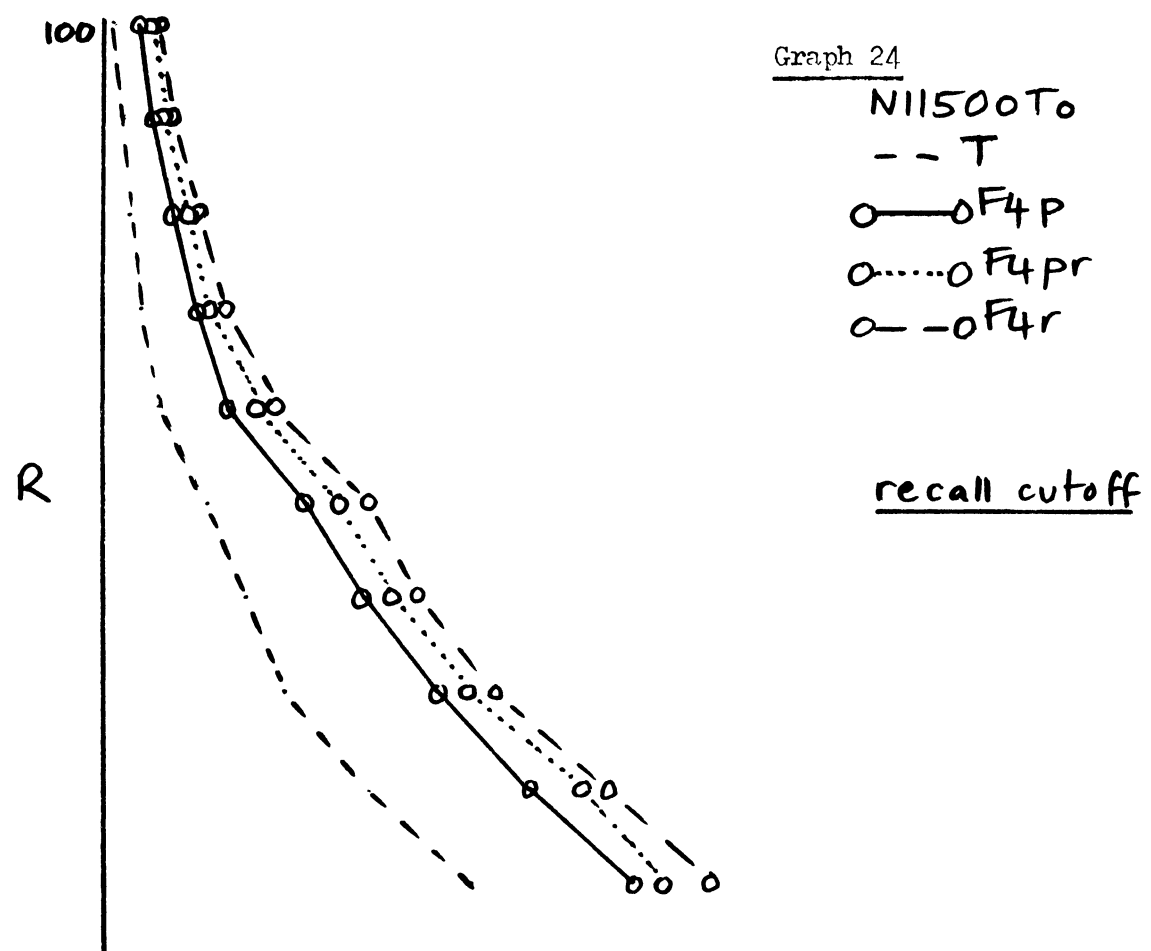




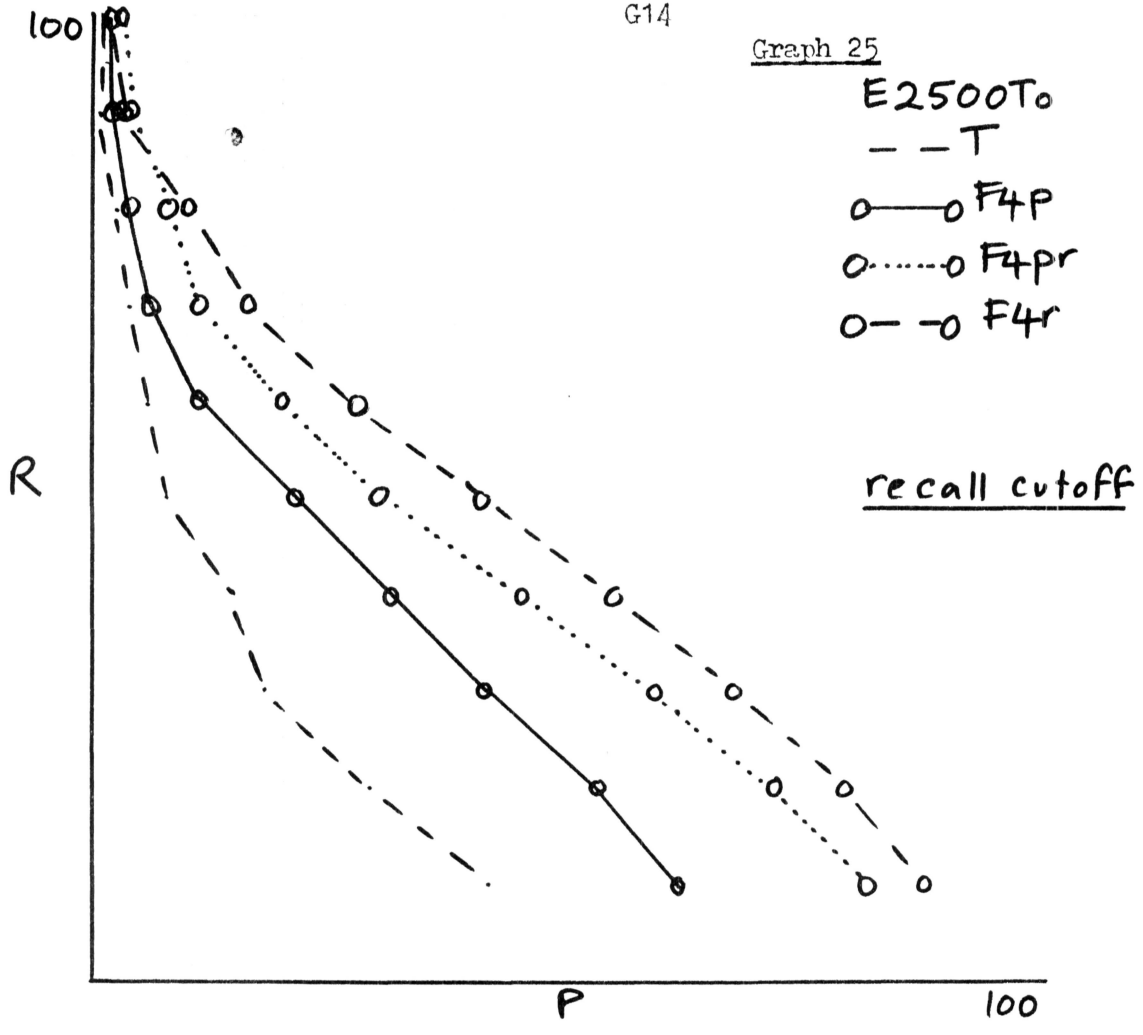
Graph 23



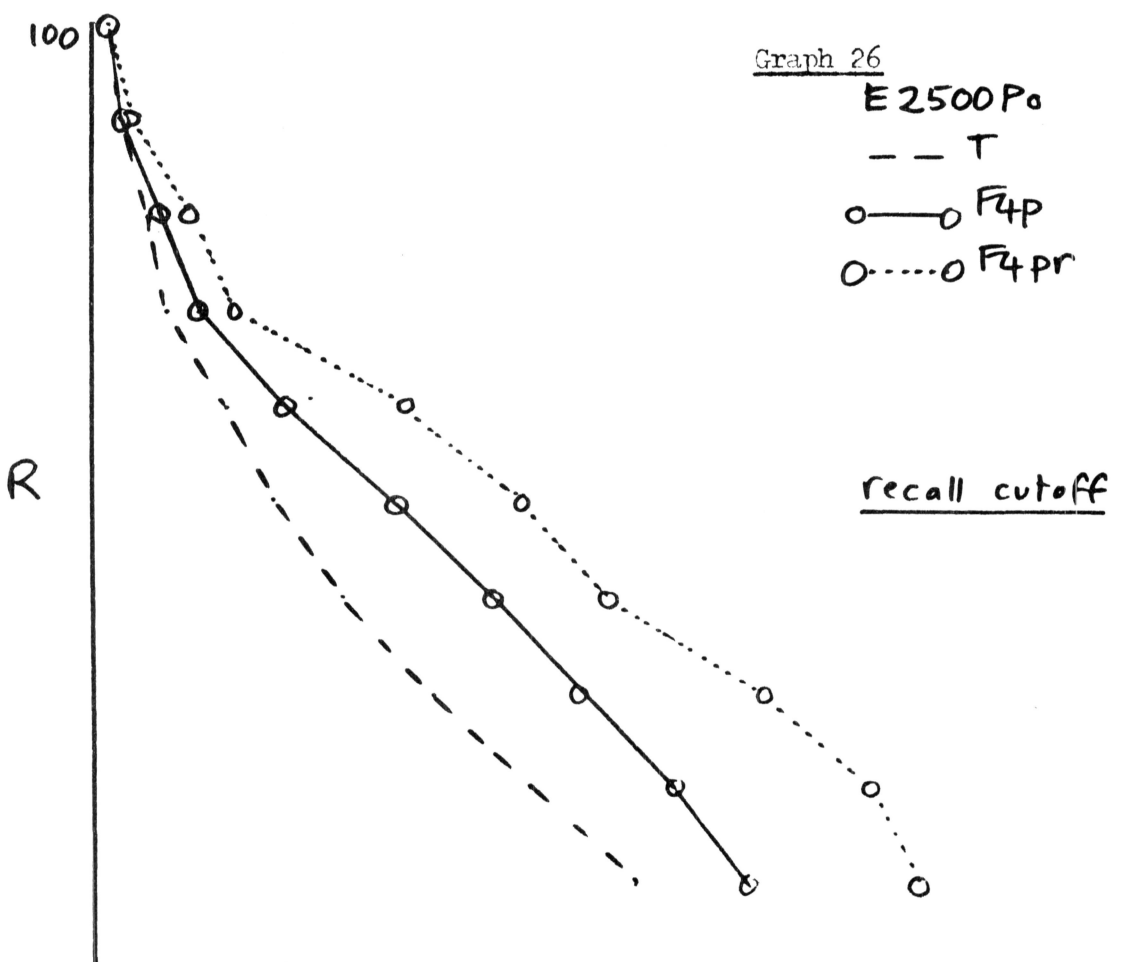
Graph 24

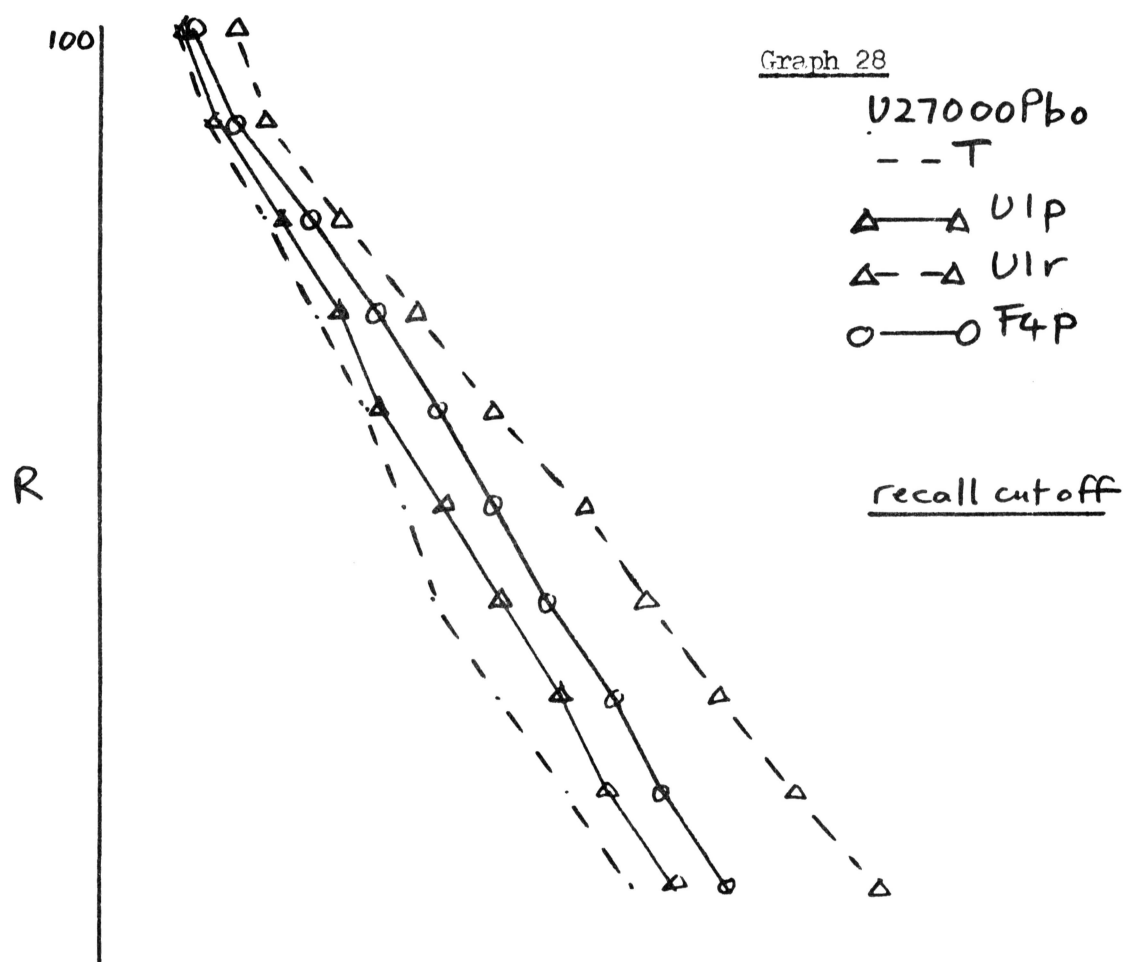
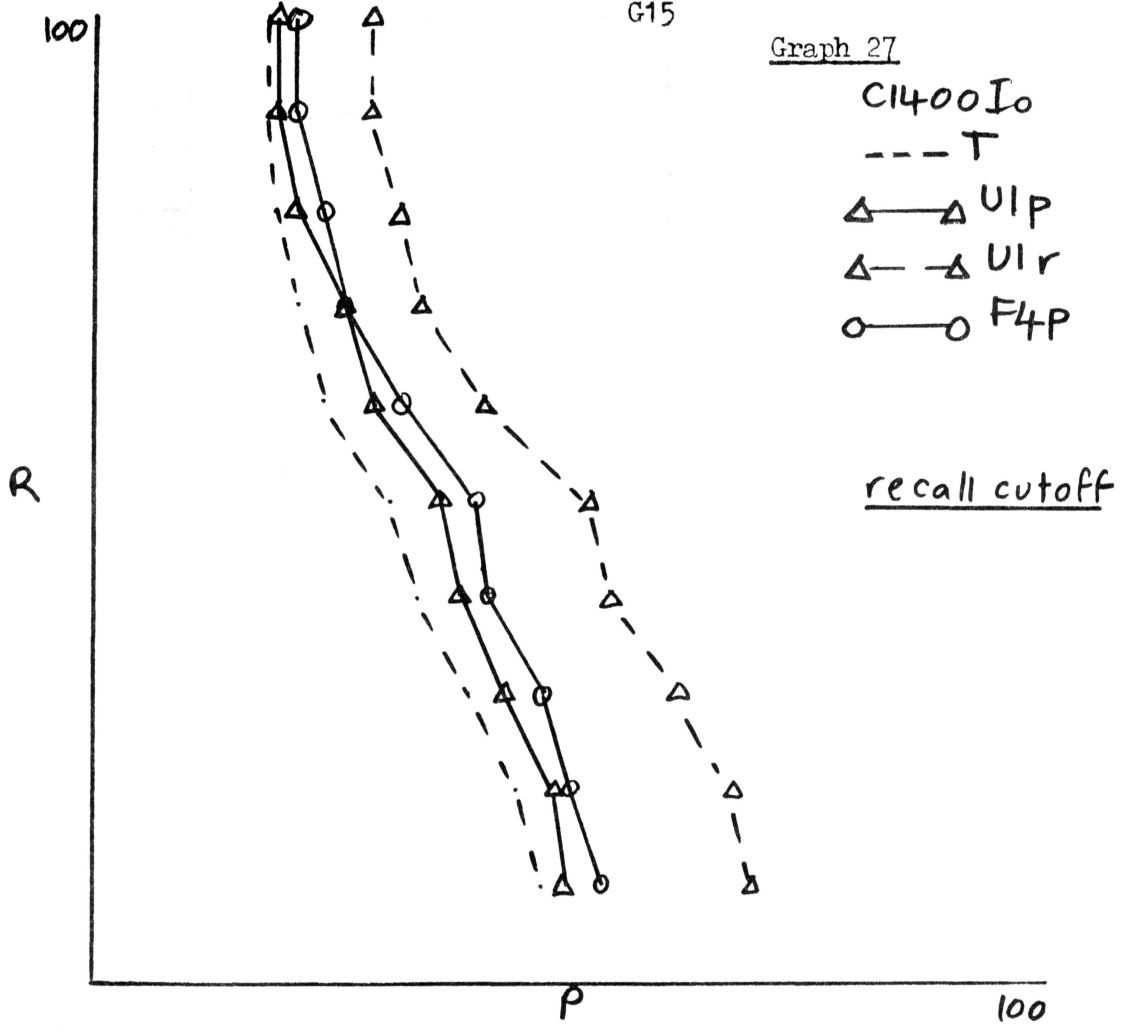


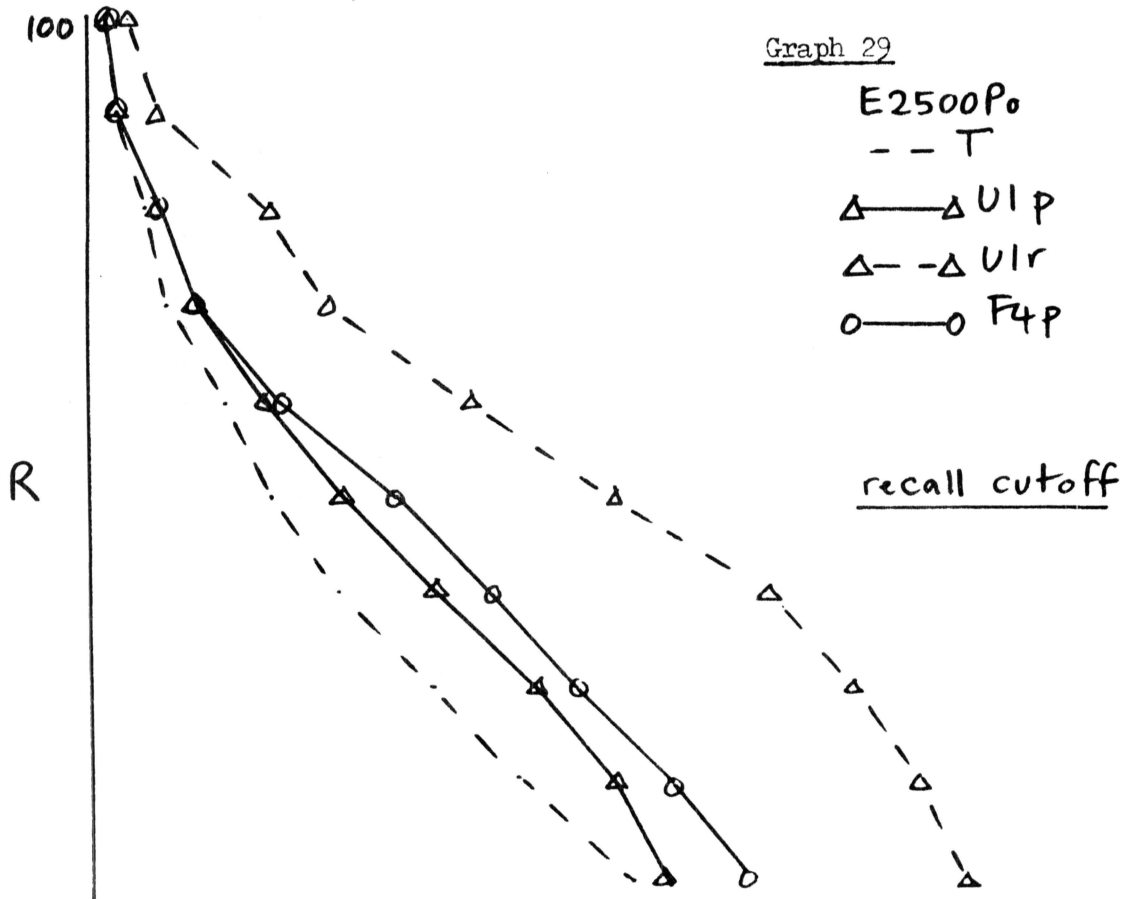
Graph 25



Graph 26







Graph 29

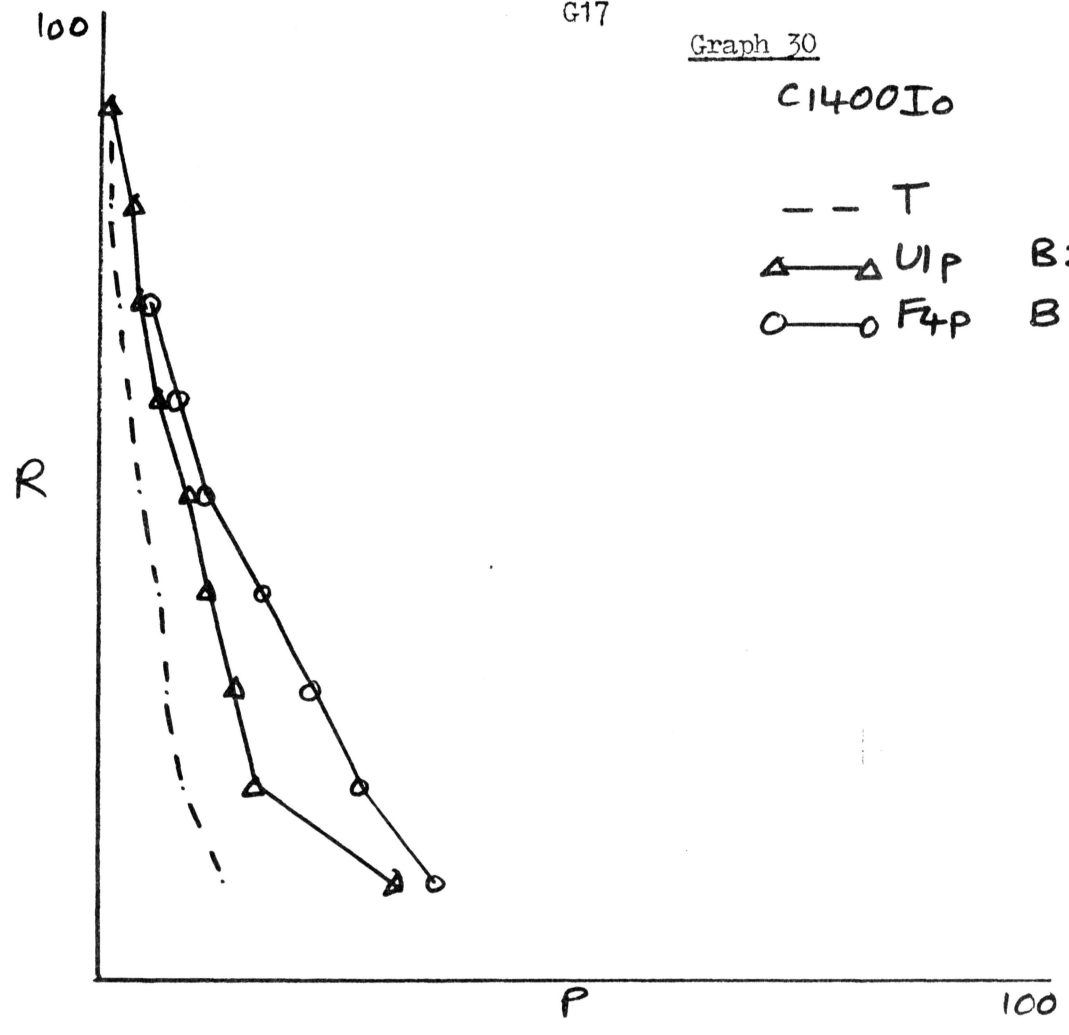
E2500Po
-- T
△—△ Ulp
△--△ Ulr
○—○ F4p

recall cutoff

Graph 30

C1400I₀

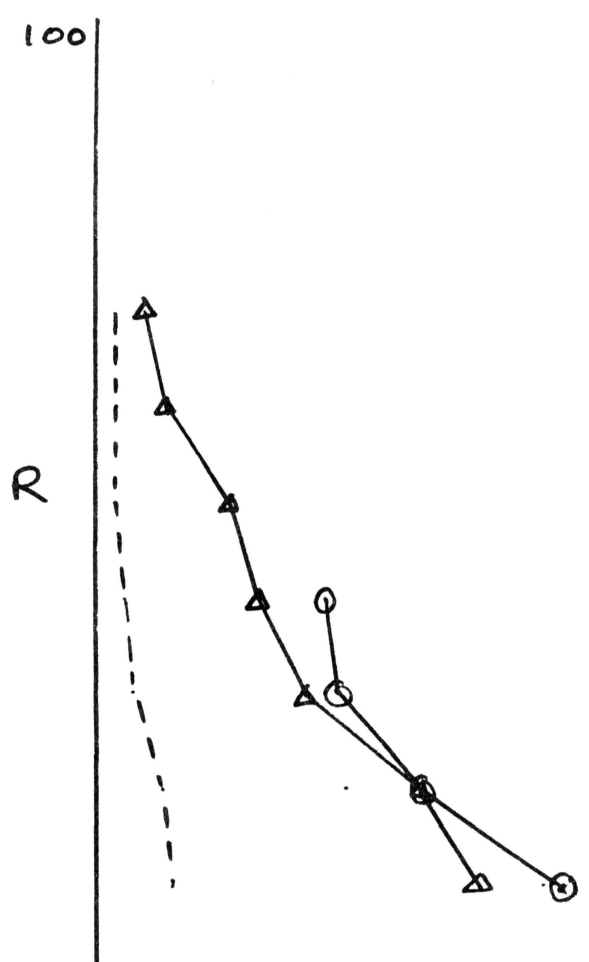
- T
- △---△ U1P B2
- F4P B2



Graph 31

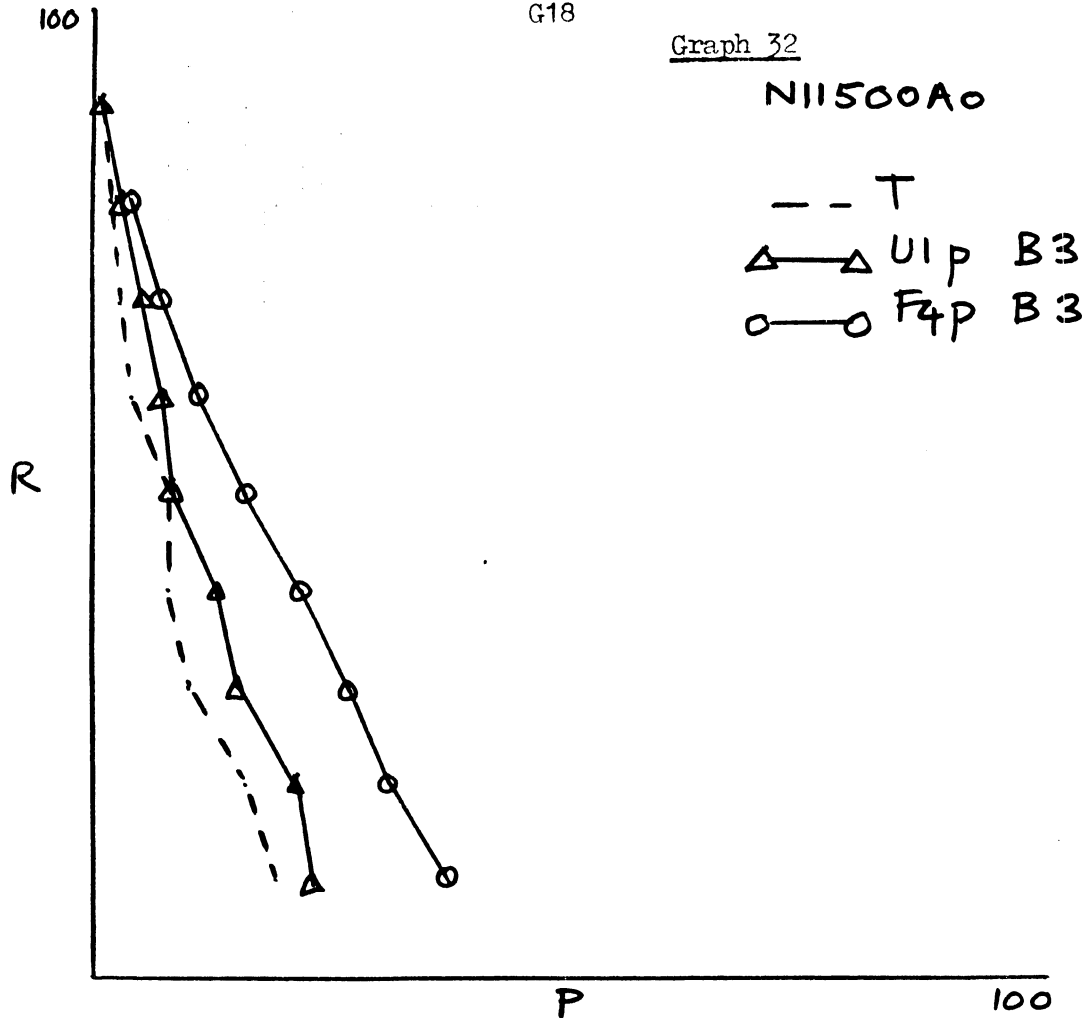
U27000Pb₀

- T
- △---△ U1P B3
- F4P B3



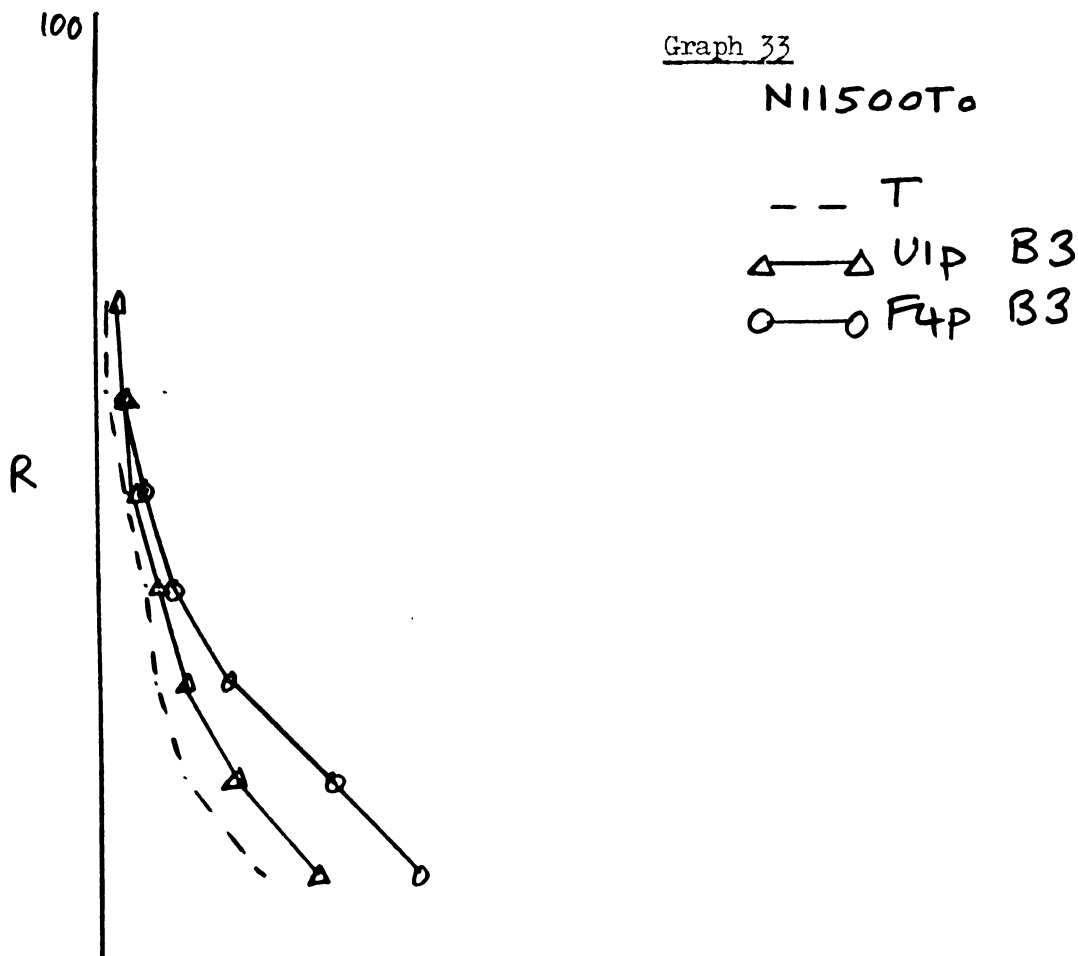
Graph 32

NI1500A0



Graph 33

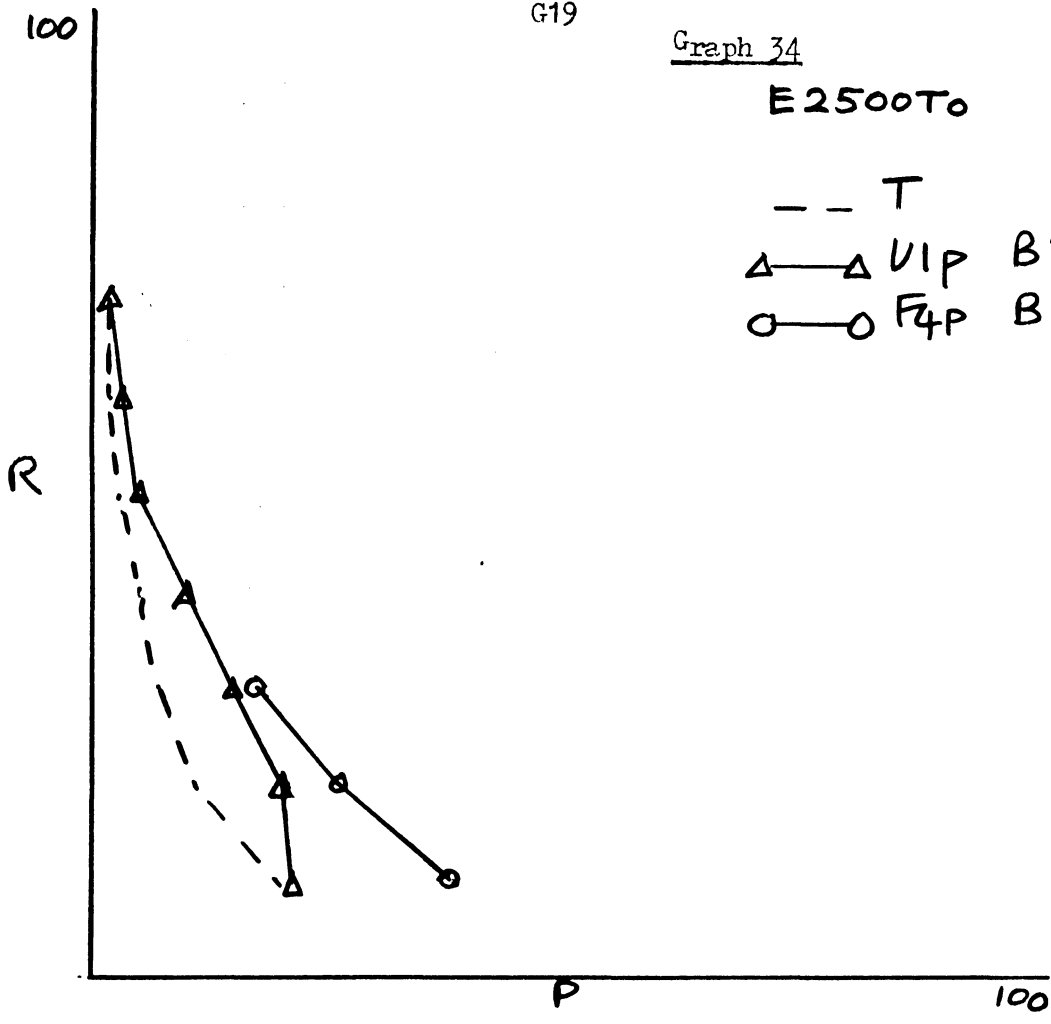
NI1500T0



Graph 34

E2500T₀

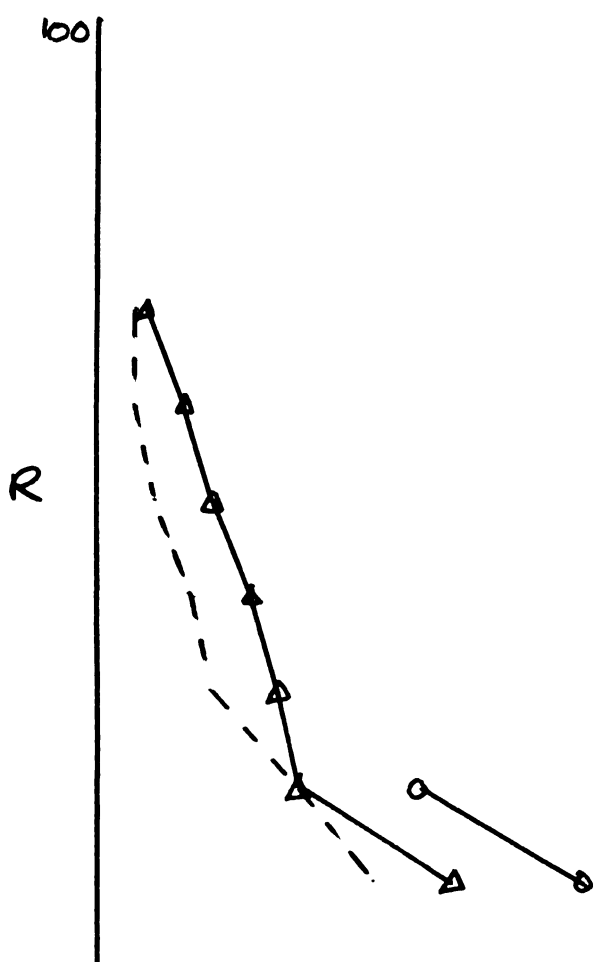
- - - T
 Δ-Δ VIP B3
 ○-○ F4P B3



Graph 35

E2500P₀

- - - T
 Δ-Δ VIP B3
 ○-○ F4P B3



Note on the organisation of the Run Tables

The performance evaluation methods used in the project are described in Chapter 4 of Section A and summarised in Figure A9. The organisation of the search output is discussed in Chapter 5 of Section A.

Each table groups runs for closely related indexing/searching strategies, with the parallel runs for different collections constituting a run set.

The tables are divided into two sets:

Main Tables: these contain recall/precision results obtained with the main performance representation method used, namely dv.

Secondary Tables: these contain different representations of results obtained with other performance representation methods, namely rc, dr, pr, tr, rr, and cr.

Any particular run (run set) can therefore have corresponding entries, for the different representation methods, in a Main Table and corresponding Secondary Tables.

There are 12 table groups of runs, as indicated in Section A, Chapter 5: thus R, for example, refers to a group of relevance weighting run sets. The corresponding Main and Secondary Tables are distinguished by the prefixes M and S respectively, with the latter further distinguished by the particular representation method used. For a given run set in the R group, say 9, we therefore refer to the particular search output representation figures of the Main Table as MR9, and to the figures of the rc representation as SrcR9.

In addition, again as described in Chapter 5 of Section A, particular indexing/searching strategies were tested with variant relevance sets, as opposed to the regular ones. The results essentially provide further alternative tables, involving corresponding Main and Secondary representations. The variant sets fall into two groups, labelled v1 and v2, and these labels are used to flag the appropriate tables within the Main and Secondary sets respectively: thus we may refer to run set v2/MR9 or v2/SrcR9. However as there are relatively few run sets involving the variants, their tables are summary ones following the complete sets of regular Main and Secondary Tables respectively.

An initial Key Table lists the table group and run set specifications, and the following Summary Table shows which runs have actually been done. The actual results for these all appear in the Main Tables, while those with corresponding Secondary Table entries (mainly rc other than tr) have a ' .

An Other Table, labelled O, gives some miscellaneous results.

The overall order of the tables is therefore:

Key Table	K
Summary Table	S
Main Tables, regular	M
" variant	v/M
Secondary Tables, regular	Sx
" variant	v/Sx
Other Table .	

SPECIFICATION OF TABLES

<u>Table</u>	<u>Description</u>	<u>Runs</u>
T	Terms, i.e. simple term matching of original requests	T1
W	Collection frequency weights, applied to original requests	W1, W1.1
R	Relevance weights, applied to original requests	R1-12, R2.1,2.2
S	Substitutions of relevant document terms for original requests	S1-5
SW	Substitutions with collection frequency weights	SW1-5
SR	Substitutions with relevance weights	SR1-5
C	Classifications enlarging original requests	C1-2
CW	Classifications with collection frequency weights	CW1-2
CR	Classifications with relevance weights	CR1-6
E	Expansions by relevant documents of original requests	E1-5
EW	Expansions with collection frequency weights	EW1-5
ER	Expansions with relevance weights	ER1-5

The individual runs in each table are listed below:

SPECIFICATION OF RUNS

<u>Table</u>	<u>Name</u>	<u>Description</u>	Name 1977
T	T1	Original request terms	(T1)
W	W1	Collection frequency weights, FO, using max n	
	W1.1	" " " " using N	
R	R1	Relevance weights, F1, predictive, i.e. F1p	(R4)
	R2	" " F4, " F4p	(R7)
	R3	" " F1, predictive retrospective, i.e. F1pr	(R8)
	R4	" " F4, " " F4pr	(R9)
	R5	" " F1, retrospective yardstick, i.e. F1r	(R10)
	R6	" " F4, " " F4r	(R13)
	R7	" " U1, predictive, i.e. U1p	
	R8	" " U3, " U3p	
	R9	" " U1, predictive retrospective, i.e. U1pr	
	R10	" " U1, retrospective yardstick, i.e. U1r	
	R11	" " U3, " " U3r	
	R12	" " H1, predictive, i.e. H1p	
	R2.1	" " F4, predictive, u=0 if n=0	
	R2.2	" " F4, " w=0,u=0 if n=0	
S	S1	F1 relevant document used as request	(T4)
	S2	B1 " " " " "	(T6)
	S3	F2 " " " " "	
	S4	B2 " " " " "	
	S5	ALL " " " " "	
SW	SW1	Combines S1 specification and W1 specification	
	SW2	" S2 " " " "	
	SW3	" S3 " " " "	
	SW4	" S4 " " " "	
	SW5	" S5 " " " "	
SR	SR1	Combines S1 specification and R2 specification	
	SR2	" S2 " " " "	
	SR3	" S3 " " " "	
	SR4	" S4 " " " "	
	SR5	" S5 " " " "	
C	C1	STARS classification enlarging request	(C7)
	C2	MST " " " "	
CW	CW1	Combines C1 specification and W1 specification	
	CW2	" C2 " " " "	
CR	CR1	Combines C1 specification and R2 specification	
	CR2	" " " " R7 "	
	CR3	" " " " R12 "	
	CR4	" C2 " " R2 "	
	CR5	" " " " R7 "	
	CR6	" " " " R12 "	
E	E1	F1 relevant document enlarging request	
	E2	B1 " " " "	
	E3	F2 " " " "	
	E4	B2 " " " "	
	E5	ALL " " " "	

contd.

EW	EW1	Combines E1 specification and W1 specification
	EW2	" E2 " " "
	EW3	" E3 " " "
	EW4	" E4 " " "
	EW5	" E5 " " "
ER	ER1	Combines E1 specification and R2 specification
	ER2	" E2 " " "
	ER3	" E3 " " "
	ER4	" E4 " " "
	ER5	" E5 " " "

NPL												
11500												
A						T						
gen. set	F1	F2	F3	B1	B2	B3	F1	F2	F3	B1	B2	B3
app. set	o	o	o	o	o	o	o	o	o	o	o	o
reqs	a	a	a	a	a	a	a	a	a	a	a	a
RUN												
R2					x	x'					x	x'
R7						x						x
R12												
R2.1					x	x					x	x
R2.2					x	x					x	x

Evans												
2500												
	T						P					
gen. set	F1	F2	F3	B1	B2	B3	F1	F2	F3	B1	B2	B3
app. set	o	o	o	o	o	o	o	o	o	o	o	o
reqs	a	a	a	a	a	a	m	m	m	m	m	m
RUN												
R2					x	x'		x'			x	x'
R7						x		x				x
R12												
R2.1					x	x		x				
R2.2					x	x						

Precision

Re-call	Cranfield 1400													UKCIS 27000				NPL 11500				Evans 2500					
	I		o		T		l	S $\frac{1}{4}$	b	P			b	o	l	S $\frac{1}{4}$	A		T		o		P		o		
	m	m	a	a	a	a	a	a	a	m	m	m	m	m	a	a	a	a	a	a	a	a	a	m	m	m	m
RUN																											
W1: 100	0	0	0		0				0		0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
90	1	1	0		0				0		0		0	2	2	0	0	0	0	0	0	0	0	0	0	0	0
80	2	2	0		0				0		0		0	5	4	0	0	0	0	0	0	0	0	0	0	0	0
70	3	3	0		0				0		2		2	6	6	1	1	1	1	1	1	1	1	1	1	1	1
60	5	5	0		0				0		3		3	9	8	2	2	2	2	2	2	2	2	2	2	2	2
50	6	7	0		0				0		8		8	12	12	3	3	3	3	3	3	3	3	3	3	3	3
40	8	8	1		1				0		11		10	15	15	5	5	5	5	5	5	5	5	5	5	5	5
30	12	11	3		3				2		12		11	18	18	8	8	8	8	8	8	8	8	8	8	8	8
20	16	15	5		5				6		12		12	23	24	17	16	16	15	15	15	15	15	15	15	15	15
10	24	23	11		9				14		15		13	30	29	24	24	29	29	29	29	29	29	29	29	29	29
W1.1 100			0								0		0										0				
90			0								0		0										0				
80			0								0		0										0				
70			0								0		0										0				
60			0								2		2										3				
50			0								3		3										5				
40			0								8		8										12				
30			1								9		9										14				
20			3								9		9										19				
10			5								10		10										28				
			8								13		12										40				

Precision

Re- call	Cran- field 1400		UKCIS 27000											NPL 11500				Evans 2500				
	I		T		l	S $\frac{1}{4}$	b	P			b	A		T		T		P				
	m	o	a	a	a	a	a	a	m	m	o	l	S $\frac{1}{4}$	a	a	a	a	a	a	m	o	
RUN																						
R1	100	0			6							0	0									0
	90	1			0							0	0									0
	80	2			0							0	0									0
	70	4			0							6	3									5
	60	6			0							6	3									7
	50	9			0							8	6									13
	40	12			2							8	6									16
	30	14			4							13	7									21
	20	20			5							19	8									30
	10	32			11							18	12									44
R2	100	0			0							0	0		0							0
	90	0			0							0	0		3							0
	80	4			0							0	0		7							0
	70	7			0							13	0		12							0
	60	10			0							28	16		17							12
	50	15			0							29	20		22							19
	40	20			2							39	23		32							28
	30	24			5							44	23		36							38
	20	28			14							45	36		41							49
	10	38			24							53	41		42							59
R3	100	0	0		0							0	0	0						0		0
	90	1	0		0							0	0	0						0		0
	80	3	0		0							0	0	0						0		0
	70	5	0		0							6	3	6						4		7
	60	8	0		0							6	7	6						9		10
	50	11	0		0							8	7	7						20		15
	40	14	1		1							9	6	10						27		23
	30	19	5		4							15	9	15						36		32
	20	27	9		6							22	18	20						53		45
	10	37	17		11							23	20	22						64		63
R4	100	0	0		0	0						0	0	0	0					0		0
	90	7	0		0	0						0	0	0	4					0		0
	80	11	0		0	0						0	0	10	16					0		0
	70	15	0		0	0						14	7	18	21					14		11
	60	21	0		0	0						28	41	34	29					22		22
	50	23	0		0	0						36	46	43	35					33		34
	40	28	2		2	0						42	51	50	35					43		50
	30	33	9		10	3						54	57	51	42					59		67
	20	40	25		27	9						60	62	59	48					70		77
	10	51	34		34	39						62	70	62	53					81		87
R5	100	0	0		0							0	0	0						0		0
	90	2	0		0							0	0	0						0		0
	80	4	0		0							0	0	0						0		0
	70	6	0		0							11	6	12						14		14
	60	9	0		0							10	12	13						25		25
	50	13	0		0							15	12	17						40		40
	40	18	2		3							17	11	24						55		55
	30	24	7		8							30	20	35						72		72
	20	32	14		14							45	42	51						81		81
	10	48	29		29							55	51	61						94		94

Precision

		UKCIS									
Recall		T					P				
generation set		$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$
application set		$S\frac{1}{4}$	$S\frac{1}{4}$	$S\frac{1}{4}$	$S\frac{1}{4}$	$S\frac{1}{4}$	$S\frac{1}{4}$	$S\frac{1}{4}$	$S\frac{1}{4}$	$S\frac{1}{4}$	$S\frac{1}{4}$
reqs		a	a	a	a	a	m	m	m	m	m
RUN											
R1	100							0			
	90							0			
	80							0			
	70							4			
	60							5			
	50							6			
	40							7			
	30							10			
	20							13			
	10							13			
R2	100	0	0	0	0	0	0	0	0	0	
	90	0	0	0	0	0	0	0	0	0	
	80	0	0	0	0	0	0	0	0	0	
	70	0	0	0	0	0	0	8	12	13	
	60	0	0	0	0	0	13	27	31	26	
	50	0	0	0	0	0	37	32	33	33	
	40	0	2	2	2	2	38	37	41	43	
	30	4	6	6	8	8	45	47	46	45	
	20	16	18	19	22	23	48	48	51	49	
	10	25	27	30	33	34	53	46	51	52	

Precision

Recall		NPL											
		A						T					
gen. set		F1	F2	F3	B1	B2	B3	F1	F2	F3	B1	B2	B3
app. set		o	o	o	o	o	o	o	o	o	o	o	o
reqs		a	a	a	a	a	a	a	a	a	a	a	a
RUN													
R2	100					0	0					0	0
	90					0	0					0	0
	80					5	4					0	0
	70					8	7					0	0
	60					13	11					0	3
	50					17	16					5	5
	40					22	22					7	8
	30					28	27					12	14
	20					32	31					26	25
	10					38	37					32	34
R7	100						0						0
	90						1						0
	80						3						0
	70						5						2
	60						7						3
	50						8						4
	40						13						6
	30						15						9
	20						21						14
	10						23						23
R12	100												
	90												
	80												
	70												
	60												
	50												
	40												
	30												
	20												
	10												
R2.1	100					0	0					0	0
	90					0	0					0	0
	80					4	4					0	0
	70					8	7					0	0
	60					13	11					0	3
	50					17	16					5	5
	40					22	22					8	8
	30					28	27					12	14
	20					32	31					25	25
	10					38	37					34	34
R2.1	100					0	0					0	0
	90					0	0					0	0
	80					5	4					0	0
	70					8	7					0	0
	60					13	11					0	3
	50					17	16					5	5
	40					22	22					8	8
	30					28	27					12	14
	20					32	31					25	25
	10					38	37					34	34

Precision

Recall		Evans											
		2500 T						P					
gen. set		F1	F2	F3	B1	B2	B3	F1	F2	F3	B1	B2	B3
app. set		o	o	o	o	o	o	o	o	o	o	o	o
reqs		a	a	a	a	a	a	m	m	m	m	m	m
RUN													
R2	100					0	0		0			0	0
	90					0	0		0			0	0
	80					0	0		0			0	0
	70					0	0		0			0	0
	60					0	0		0			0	0
	50					0	0		0			0	0
	40					0	0		0			0	0
	30					0	17		0			0	0
	20					26	26		32			0	34
	10					37	37		47			52	51
R7	100						0		0				0
	90						0		0				0
	80						0		0				0
	70						2		5				5
	60						3		7				9
	50						5		10				12
	40						10		15				16
	30						15		16				19
	20						20		19				21
	10						21		27				37
R12	100												
	90												
	80												
	70												
	60												
	50												
	40												
	30												
	20												
	10												
R2.1	100					0	0		0				
	90					0	0		0				
	80					0	0		0				
	70					0	0		0				
	60					0	0		0				
	50					0	0		0				
	40					0	0		0				
	30					0	18		22				
	20					26	26		36				
	10					37	37		51				
R2.1	100					0	0						
	90					0	0						
	80					0	0						
	70					0	0						
	60					0	0						
	50					0	0						
	40					0	0						
	30					0	0						
	20					0	15						
	10					23	24						

Precision

Re-call	Cran-field 1400		UKCIS 27000										NPL 11500				Evans 2500				
	I	o	T	o	l	S $\frac{1}{4}$	b	o	P	b	o	l	S $\frac{1}{4}$	A	o	T	o	T	o	P	o
reqs	m	m	a	a	a	a	a	a	m	m	m	m	m	a	a	a	a	a	a	m	m
RUN																					
R11	100																				
	90																				
	80																				
	70																				
	60																				
	50																				
	40																				
	30																				
	20																				
	10																				
R12	100																				
	90																				
	80																				
	70																				
	60																				
	50																				
	40																				
	30																				
	20																				
	10																				

M.A.

Precision

Recall		Cran-field 1400 I				UKCIS 27000 T											
		F1	F2	B1	B2	F1	F2	F3	B1	B2	B3	P b F1	F2	F3	B1	B2	B3
gen. set	app. set	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
reqs		m	m	m	m	a	a	a	a	a	a	m	m	m	m	m	m
RUN																	
R2	100																
	90																
	80																
	70																
	60																
	50				n.a.											n.a.	
	40																
	30																
	20																
	10																
R7	100																
	90																
	80																
	70																
	60																
	50																
	40																
	30																
	20																
	10																
R12	100																
	90																
	80																
	70																
	60																
	50				n.a.												
	40																
	30																
	20																
	10																

Recall, precision

Rank	Cranfield 1400		UKCIS 27000										NPL 11500				Evans 2500					
	I	o	T	o	l	S $\frac{1}{4}$	b	o	P	b	o	l	S $\frac{1}{4}$	A	o	T	o	T	o	P	o	
reqs	m	m	a	a	a	a	a	a	m	m	m	m	m	a	a	a	a	a	a	a	m	m
RUN																						
R11	5																					
	10																					
	20																					
	50																					
	100																					
R12	5																					
	10																					
	20																					
	50																					
	100																					

A.A.

Recall, precision

Rank	Cran-field 1400				UKCIS 27000											
	I				T							P b				
gen. set	F1	F2	B1	B2	F1	F2	F3	B1	B2	B3	F1	F2	F3	B1	B2	B3
app. set	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
reqs	m	m	m	m	a	a	a	a	a	a	m	m	m	m	m	m
RUN																
R2	5															
	10															
	20			n.a.												n.a.
	50															
	100															
R7	5															
	10															
	20															
	50															
	100															
R12	5															
	10															
	20			n.a.												
	50															
	100															

r = average relevant retrieved
 t = average total retrieved

		UKCIS									
		27000					P				
		T					b				
generation set		$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$
application set		$S\frac{1}{4}$	$S\frac{1}{4}$	$S\frac{1}{4}$	$S\frac{1}{4}$	$S\frac{1}{4}$	$S\frac{1}{4}$	$S\frac{1}{4}$	$S\frac{1}{4}$	$S\frac{1}{4}$	$S\frac{1}{4}$
reqs		a	a	a	a	a	m	m	m	m	m
RUN											
R1	r							9.1			
	t							368.4			
R2	r	n.a.	n.a.	n.a.	n.a.	n.a.	7.9	8.7	8.8	9.0	
	t						89.7	126.9	157.7	188.7	

		NPL											
		11500						T					
A													
gen. set		F1	F2	F3	B1	B2	B3	F1	F2	F3	B1	B2	B3
app. set		o	o	o	o	o	o	o	o	o	o	o	o
reqs		a	a	a	a	a	a	a	a	a	a	a	a
RUN													
R2	r					9.5	9.3					6.6	7.3
	t					213.5	221.1					182.1	224.2
R7	r						10.7						8.8
	t						6861						481.0
R12	r												
	t												
R2.1	r					9.5	9.3					6.7	7.3
	t					214.4	221.1					187.3	228.0
R2.2	r					9.5	9.3					6.7	7.3
	t					214.3	221.0					187.1	227.8

		Evans											
		2500						P					
		T											
gen. set		F1	F2	F3	B1	B2	B3	F1	F2	F3	B1	B2	B3
app. set		o	o	o	o	o	o	o	o	o	o	o	o
reqs		a	a	a	a	a	a	m	m	m	m	m	m
RUN													
R2	r					3.2	4.0		2.8			2.3	3.2
	t					15.7	26.4		10.1			7.0	12.8
R7	r						8.4		9.1				9.1
	t						34.3		22.4				22.0
R12	r												
	t												
R2.1	r					3.3	4.3		4.6				
	t					18.2	30.3		26.6				
R2.2	r					3.1	4.0						
	t					17.0	28.7						

REFERENCES

- BARKER, F. H., VEAL, D. C. and WYATT, B. K., 'Towards automatic profile construction', Journal of Documentation, 28, 44-55 (1972).
- BARKER, F. H., VEAL, D. C. and WYATT, B. K., Retrieval Experiments Based on Chemical Abstracts Condensates, Research Report No. 2, UKCIS, University of Nottingham (1974).
- BARRACLOUGH, E. D. et al., The Medusa Current Awareness Experiment, Computing Laboratory, University of Newcastle upon Tyne (1975).
- CLEVERDON, C. W., MILLS, J. and KEEN, E. M., Factors Determining the Performance of Indexing Systems, 2 vols., College of Aeronautics, Cranfield (1966).
- CROFT, W. B. and HARPER, D. J., 'Using probabilistic models of document retrieval without relevance information', Journal of Documentation, 35, 285-295 (1979).
- EVANS, L., Search Strategy Variations in SDI Profiles, Report R75/1, INSPEC, Institution of Electrical Engineers, London (1975a).
- EVANS, L., Methods of Ranking SDI and IR Outputs, Report R75/3, INSPEC, Institution of Electrical Engineers, London (1975b).
- HARPER, D. J. and VAN RIJSBERGEN, C. J., 'An evaluation of feedback in document retrieval using cooccurrence data', Journal of Documentation, 34, 189-216 (1978).
- HARPER, D. J., Relevance Feedback in Document Retrieval Systems: An Evaluation of Probabilistic Strategies, Ph.D. Thesis, Computer Laboratory, University of Cambridge (1980).
- MILLER, W. L., The Evaluation of Large Information Retrieval Systems with Application to MEDLARS, Ph.D. Thesis, University of Newcastle (1970).
- ROBERTSON, S. E., A Theoretical Model of the Retrieval Characteristics of Information Retrieval Systems, Ph.D. Thesis, University of London (1976).
- ROBERTSON, S. E. and SPARCK JONES, K., 'Relevance weighting of search terms', Journal of the American Society for Information Science, 27, 129-146 (1976).
- ROBERTSON, S. E., VAN RIJSBERGEN, C. J. and PORTER, M. F., 'Probabilistic models of indexing and searching', paper presented at the BCS/ACM Symposium on Research and Development in Information Retrieval, Cambridge, 1980; in press.
- ROBSON, A. and LONGMAN, J. S., 'Automatic aids to profile construction', Journal of the American Society for Information Science, 27, 213-223 (1976).

ROCCHIO, J. J., 'Relevance feedback in information retrieval', in The SMART Retrieval System, Ed. G. SALTON, Prentice-Hall, Englewood Cliffs, N.J. (1971).

SALTON, G. (Ed.), The SMART Retrieval System, Prentice-Hall, Englewood Cliffs, N.J. (1971).

SPARCK JONES, K., Automatic Keyword Classification for Information Retrieval, Butterworths, London (1971).

SPARCK JONES, K., 'A statistical interpretation of term specificity and its application in retrieval', Journal of Documentation, 28, 11-21 (1972).

SPARCK JONES, K., 'A performance yardstick for test collections', Journal of Documentation, 31, 266-272 (1975).

SPARCK JONES, K., 'Performance averaging for recall and precision', Journal of Informatics, 2, 95-105 (1977).

SPARCK JONES, K. and BATES, R. G., Research on Automatic Indexing 1974-1976, 2 vols., British Library Research and Development Report 5464, Computer Laboratory, University of Cambridge (1977).

SPARCK JONES, K., 'Search term relevance weighting given little relevance information', Journal of Documentation, 35, 30-48 (1979a).

SPARCK JONES, K., 'Experiments in relevance weighting of search terms', Information Processing and Management, 15, 133-144 (1979b).

SPARCK JONES, K., 'Search term relevance weighting - some recent results', Journal of Information Science, (1980).

VASWANI, P. K. T. and CAMERON, J. B., The National Physical Laboratory Experiments in Statistical Word Associations and Their Use in Document Indexing and Retrieval, National Physical Laboratory, Teddington (1970).

VAN RIJSBERGEN, C. J., 'A theoretical basis for the use of cooccurrence data in information retrieval', Journal of Documentation, 33, 106-119 (1977).

VAN RIJSBERGEN, C. J., HARPER, D. J. and PORTER, M. F., 'The selection of good search terms', paper presented at the Seventh Cranfield International Conference on Mechanised Information Storage and Retrieval Systems, 1979; in press.

VERNIMB, C., 'Automatic query adjustment in document retrieval', Information Processing and Management, 13, 339-353 (1977).

YU, C. T. and SALTON, G., 'Effective information retrieval using term accuracy', Communications of the ACM, 20, 135-142 (1977).

YU, C. T., LAM, K. and SALTON, G., 'Optimum term weighting in information retrieval using the term precision model', Department of Computer Science, Cornell University (1980).

GLOSSARY

- alternative request
different form of request primary indexing
- application (set)
document set used for searching with weighted requests
- baseline
simple term matching retrieval performance
- collection
see test collection
- collection frequency
number of documents (descriptions) in a collection containing a term
- collection frequency weight
term weight based on its collection frequency
- completely ordered
fully ordered search output including all documents in a collection
- component
of a retrieval system, in the most general sense
- cross check
comparison altering values of a variable other than the focal one
- description
characterisation of a document or request
- description exhaustivity
see indexing description exhaustivity
- document cutoff
basis for recall and precision computation
- exhaustivity
see indexing description exhaustivity
- factor
system component affecting performance
- fully ordered
ordered search output with one retrieved document per rank
- generation (set)
document set used to calculate term weights

- indexing description exhaustivity
 - number of index terms per description
- indexing factor
 - factor affecting the descriptions of documents and requests
- indexing mode
 - automatic or manual preparation of primary indexing
- indexing source
 - form of document or request used for indexing
- indexing vocabulary
 - term set used for collection primary indexing
- indexing vocabulary specificity
 - number of documents per term
- input factor
 - factor affecting input documents and requests
- matching condition
 - form of matching between requests and documents
- matching rank
 - point of comparison between request search outputs
- matching value
 - point of comparison between request search outputs
- material (difference)
 - difference between recall/precision graphs of at least 10%
- mode
 - see indexing mode
- noticeable (difference)
 - difference between recall/precision graphs of at least 5%
- ordered (output)
 - partially or fully ordered search output
- output factor
 - factor affecting search conduct and output
- output form
 - output as naturally generated by a search procedure
- output type
 - output as used for comparative evaluation
- parameter
 - environmental factor (not subject to control)
- partially ordered
 - ordered search output allowing more than one retrieved

document per rank

primary indexing
word or word-stem lists representing document and request sources

proposition
statement germane to a particular research topic and related to experiments

raw material
initial documents and request need statements

recall cutoff
basis for recall and precision computation

relevance feedback
retrieval techniques exploiting relevance assessments to control searching

relevance weight
term weight utilising relevance information

relevance variants
sets of relevant documents defined by different relevance grades

run
averaged search output for specific request, document and relevance sets, involving specific indexing and searching procedures

run set
specific run output for different collections

scanning strategy
form of search inspection of a collection

scoring criterion
type of scoring for a request-document match

setting
of a parameter

source
see indexing source

source material
forms of documents and request need statements exploited in indexing and relevance assessment

specificity
see indexing language specificity

subsidiary collection

variant, e.g. part of, a test collection

term
extracted word stem

term weighting
any weighting of index terms in a description

test collection
particular set of primary indexing descriptions for some
source material

topic
subject of investigation

unordered (output)
minimally ordered output i.e. retrieved/not retrieved

value
of a system variable

variable
system factor (subject to control)

variant collection
collection with relevance set alternative to main one

version
form of raw material represented by a specific collection

vocabulary
see indexing vocabulary

yardstick
performance defined by optimal weighting