

211617

PUBLIC VERSION
HIGHLY CONFIDENTIAL AND CONFIDENTIAL MATERIAL REDACTED

**BEFORE THE
SURFACE TRANSPORTATION BOARD**

OTTER TAIL POWER COMPANY

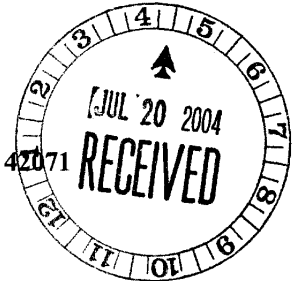
Complainant,

v.

**THE BURLINGTON NORTHERN AND
SANTA FE RAILWAY COMPANY,**

Defendant.

Docket No. 42071



ENTERED
Office of Proceedings

JUL 20 2004

Part of
Public Record

**COMPLAINANT'S REPLY TO
DEFENDANT'S MOTION TO STRIKE**

George Koeck, Esq.
General Counsel
Otter Tail Corporation
4334 18th Avenue, SW, Suite 200
Fargo, ND 58103
(701) 451-3567

Nicholas J. DiMichael, Esq.
Jeffrey O. Moreno, Esq.
Michael H. Higgins, Esq.
THOMPSON HINE LLP
1920 N Street, N.W., Suite 800
Washington, DC 20036
(202) 331-8800

Attorneys for Complainant

July 20, 2004

**PUBLIC VERSION – CONFIDENTIAL AND
HIGHLY CONFIDENTIAL MATERIAL REDACTED**

**BEFORE THE
SURFACE TRANSPORTATION BOARD**

OTTER TAIL POWER COMPANY,	)	
	)	
Complainant,	)	
	)	
v.	)	Docket No. 42071
	)	
THE BURLINGTON NORTHERN AND	)	
SANTA FE RAILWAY COMPANY,	)	
	)	
Defendant.	)	

**COMPLAINANT’S REPLY TO
DEFENDANT’S MOTION TO STRIKE**

Otter Tail Power Company (“Otter Tail”) hereby replies in opposition to “Defendant’s Motion to Strike,” filed on June 8, 2004 by The Burlington Northern and Santa Fe Railway Company (“BNSF”). In an order served on June 24, 2004, the Board extended the time for Otter Tail to submit this Reply until July 20, 2004.

BNSF’s Motion asks the Board to strike Otter Tail’s April 29, 2004 Rebuttal Evidence in six areas: (1) evidence relating to Otter Tail’s purchases of coal from certain southern PRB mines; (2) evidence relating to SARR capacity and operating statistics based on Otter Tail’s rebuttal string program;¹ (3) changes to Otter Tail’s configuration of the Fargo Yard; (4) evidence on alternatives to the RCAF-A and RCAF-U as an inflation index to be used in the DCF analysis; (5) updates to the calculation of the SARR’s cost of capital; and (6) Otter Tail’s statistical analysis of rerouted traffic. BNSF seeks to strike evidence in categories 1 through 5 on the ground that they exceed the proper scope of rebuttal, as set forth by the Board in *Duke*

¹ Contemporaneous with this Reply to BNSF’s Motion to Strike, Otter Tail has filed a Motion for Technical Conference that relates solely to BNSF’s arguments for striking Otter Tail’s rebuttal string program.

Energy Corp. v. Norfolk Southern Ry. Co., STB Docket No. 42069, slip op. at 13-14 (served Nov. 5, 2003) (“*Duke/NS*”). As to category 6, BNSF alleges that the evidence should be stricken for not meeting the admissibility standards for expert testimony in Rule 702 of the Federal Rules of Evidence.

BNSF’s Motion is a thinly disguised reply to Otter Tail’s Rebuttal Evidence. A Motion to Strike need only identify alleged violations of the Board’s standards for rebuttal evidence. BNSF has gone far beyond that by submitting argument and evidence (*e.g.* Verified Statements from three expert witnesses and one fact witness) in response to Otter Tail’s Rebuttal Evidence. The Board’s rules prohibit such a reply to a reply. 49 C.F.R. § 1104.13(c). The Board has proscribed the use of motions to strike as a vehicle for submitting an impermissible reply to a reply and it should do so again in this case.²

I. OTTER TAIL’S EVIDENCE OF COAL PURCHASES FROM MINES SOUTH OF CORDERO DIRECTLY RESPONDS TO BNSF’S REPLY EVIDENCE.

In its October 8, 2003 Reply Evidence and its March 22, 2004 Supplemental Reply Evidence, BNSF sought to exclude from the Otter Tail Railroad (“OTRR”) between 54 and 84 million tons of coal³ that originates at mines south of Cordero, and then moves south to Converse, because that traffic allegedly does not share facilities with the issue traffic. At pages

² See STB Docket No. AB-556 (Sub-No. 2X), STB Finance Docket No. 33385, *Railroad Ventures, Inc.—Abandonment Exemption—Between Youngstown, OH and Darlington, PA, in Mahoning and Columbiana Counties, OH and Beaver County, PA*; *Railroad Ventures, Inc.—Acquisition and Operation Exemption—Youngstown & Southern Railroad Co.*, 2000 STB Lexis 588 at *30, n.44 (Served Oct. 5, 2000); ICC Finance Docket No. 32229, *Burlington Northern Railroad Co.—Construction and Operation Exemption—Macon and Randolph Counties, MO*, 1993 ICC Lexis 225 at *2, n.2 (Served Nov. 17, 1993); ICC Finance Docket No. 31271; ICC Finance Docket 31230, *City of Colorado Springs and Metex Metropolitan District—Petition for Declaratory Order—Abandonment Determination*; *The Denver and Rio Grande Western Railroad Co.—Petition for Declaratory Order—Spur Track Determination*, 1989 ICC Lexis 78 at *27 (March 22, 1989); ICC Docket No. 40139, *Missouri Pacific Railroad Co. and Consolidated Rail Corporation—Joint Petition for Declaratory Order—Recovery of Transportation Charges*, 1987 ICC Lexis 66 at *1 (Nov. 9, 1987)

³ Specifically, BNSF seeks to exclude 54 million tons from Otter Tail’s base case, which was submitted in Otter Tail’s Opening Evidence, and 84 million tons from Otter Tail’s alternative case, which was submitted in Otter Tail’s Supplemental Opening Evidence.

I-19 through I-26 and III-A-23 through III-A-27 of its Rebuttal Evidence, Otter Tail demonstrated that BNSF was wrong factually, because this traffic *does share* facilities directly with the issue traffic; and moreover, even if it did not share facilities directly with the issue traffic (*i.e.*, what BNSF calls “second-layer” cross-over traffic), the point is moot because there is no requirement that second-layer cross-over traffic share facilities with the issue traffic. In partial support of the first argument, Otter Tail presented evidence of historical purchases of coal from mines south of Cordero, as recently as October 2003.⁴ This is the evidence that BNSF seeks to strike.

Although this evidence demonstrates with certainty that the issue traffic does share facilities with southbound coal traffic from mines south of Cordero, it is by no means essential to Otter Tail’s argument. As Otter Tail proved in its Rebuttal Evidence, the issue traffic is defined by the scope of the challenged tariff rates in the Complaint, which in this case includes rates from all PRB origins served by BNSF, including mines south of Cordero. There is no need for Otter Tail to present evidence of historical shipments from these mines to Big Stone to prove that argument.

Otter Tail is compelled to respond, however, to BNSF’s attempt to impugn Otter Tail’s motives for submitting evidence of its October 2003 movements of coal from the Black Thunder Mine to the Big Stone Station. BNSF charges Otter Tail with “gaming” the SAC results by scheduling shipments of Black Thunder coal in October 2003, only 10-14 days after BNSF filed its Reply Evidence, in which BNSF argued that all southbound coal from mines south of Cordero should be excluded from the OTRR because that coal does not share facilities with the issue traffic. Specifically, BNSF suggests that “the October 2003 shipments [from Black Thunder]

⁴ OTP Rebuttal Ev. at I-20 and III-A-24 through III-A-25.

were intended to game the SAC results by creating the appearance that Otter Tail intended to source future coal needs from the Black Thunder mine.”⁵ But, the attached Verified Statement of Ward Uggerud, a Senior Vice President for Otter Tail, proves beyond a doubt that the Black Thunder movements were planned *before* BNSF filed its Rebuttal Evidence on October 8, 2003, {{ [REDACTED] }}⁶ Thus, it is clear that Otter Tail’s decision to ship four trainloads of coal from Black Thunder in October 2003, for a test burn, could not have been influenced by BNSF’s Reply Evidence, but instead was part of Otter Tail’s search, in the ordinary course of business, for a coal supplier to Big Stone upon expiration of the current coal supply agreements {{ [REDACTED] }}⁷

Furthermore, the standards for rebuttal evidence in *Duke/NS* do not prohibit Otter Tail’s evidence of coal shipments from mines south of Cordero to Big Stone. BNSF correctly acknowledges that a “complainant may present further argument in support of the assumptions it made on opening in response to a defendant’s criticism of the complainant’s opening evidence.”⁸ This is precisely what Otter Tail did. But, BNSF then concocts an exception that appears nowhere in the *Duke/NS* decision. BNSF asserts that a “complainant cannot submit new evidence in support of its original assumptions unless it first shows that it could not have anticipated the need for such supporting evidence when it filed its opening evidence.”⁹ Thus, according to

⁵ BNSF Motion at 10.

⁶ See Uggerud V.S. at 2-4.

⁷ Although BNSF makes much of the fact that Otter Tail did not notify BNSF of these shipments on September 24, 2003, Otter Tail itself did not know it would schedule these movements until October 6, 2003, {{ [REDACTED] }} Uggerud V.S. at 4-5.

⁸ BNSF Motion at 5, citing *Duke/NS* at 14-15.

⁹ *Id.*

BNSF, Otter Tail was required to anticipate BNSF's line of attack and respond to it preemptively.

Such a "crystal ball" approach is absurd on its face. Regardless of this absurdity, BNSF cannot claim surprise or prejudice from Otter Tail's rebuttal evidence, because BNSF knew as well as Otter Tail what historical shipments Otter Tail had made from southern PRB mines to Big Stone. Therefore, BNSF was in a much better position to anticipate Otter Tail's rebuttal argument than Otter Tail was to anticipate BNSF's Reply.

Nevertheless, even if the Board were to adopt BNSF's concocted exception, Otter Tail has satisfied it. First, the evidence of Black Thunder movements in 2003 could not have been presented by Otter Tail in opening because those movements had not yet occurred. Second, as to all other evidence of historic shipments, Otter Tail had no reason to anticipate BNSF's "shared facilities" argument, since, as Otter Tail explained in its Rebuttal Evidence, BNSF would impose a legal standard that has never been adopted by the Board in prior cases and that directly contravenes the Board's determination in *TMPA*¹⁰ that the issue traffic is defined by all origins covered by the challenged tariff rate, regardless whether any coal has moved under that rate.¹¹ If there was any doubt about this fact, the Board erased it in the *Xcel* decision when it "calculated the SAC rate that the WCC would need to charge for the Xcel traffic at issue here from the PRB mines *included in Xcel's complaint* to its Pawnee steam electric generating plant near Brush, CO."¹² Just as the SAC rate applied to all PRB origins in the BNSF tariff for Pawnee, a SAC

¹⁰ *Texas Municipal Power Agency v. The Burlington Northern and Santa Fe Ry. Co.*, STB Docket No. 42056, slip op. at 10 (served March 21, 2003) ("*TMPA*").

¹¹ See OTP Rebuttal Ev. at I-20 through I-26.

¹² *Public Serv. Co. of Colorado d/b/a Xcel Energy v. The Burlington Northern and Santa Fe Ry. Co.*, STB Docket No. 42057, slip op. at 39 (served June 8, 2004) [emphasis added] ("*Xcel*").

rate will apply to all origins in the BNSF tariff for Big Stone. There can be no clearer statement that the issue traffic is defined by all traffic within the scope of the challenged tariff rates.

II. OTTER TAIL'S REBUTTAL STRING PROGRAM IS PROPERLY SUBMITTED UNDER THE *DUKE/NS* STANDARDS FOR REBUTTAL EVIDENCE.

Although BNSF purportedly has moved to strike Otter Tail's rebuttal string program as improper rebuttal under the *Duke/NS* standards, not once does BNSF explain how Otter Tail's rebuttal evidence on the string model contravenes those standards. Rather, BNSF has improperly used its Motion as a pretext to reply to Otter Tail's Rebuttal Evidence.

Otter Tail's rebuttal string model is entirely proper under the *Duke/NS* standards for rebuttal evidence. Throughout its evidentiary submissions, BNSF has made much of alleged "flaws" in Otter Tail's string model. Otter Tail's rebuttal evidence responds to these allegations either by refuting them or, in those instances when BNSF has identified a valid error in the string model, correcting the error. In *Duke/NS*, the Board defined the proper scope of rebuttal evidence as follows:

On rebuttal, as to those issues challenged by the railroad, the shipper may demonstrate that its opening evidence was feasible and supported, it may adopt the railroad's evidence, or in certain circumstances it may offer to refine its evidence to address issues raised by the railroad regarding its opening evidence. Where the railroad has identified flaws in the shipper's evidence but has not provided evidence that can be used in the Board's SAC analysis, or *where the shipper shows that the railroad's reply evidence is itself unsupported, infeasible or unrealistic, the shipper may supply corrective evidence.*¹³

¹³ *Duke/NS* at 14-15 [footnotes omitted] [emphasis added].

Otter Tail has presented extensive evidence that BNSF's use of the RTC Model is unsupported, infeasible and unrealistic.¹⁴ Under these circumstances, *Duke/NS* clearly permits Otter Tail to supply corrective evidence on rebuttal.

BNSF's objective is to sling as much mud at the string model as possible, regardless of merit or impact upon the final result, in order to foster a perception that the string model is hopelessly unreliable for its intended purpose.¹⁵ BNSF's Motion to Strike simply is another salvo in that mud brawl, and a salvo that is an impermissible reply to a reply.

Otter Tail can point to numerous examples of BNSF's meritless mud-slinging. In the most egregious example at page III.C-27 of its Reply Evidence, BNSF claims that the string model generates erratic and physically impossible speed fluctuations based on a graphical representation of the movements that the string model calculates for a loaded coal train in Exhibit III.C-5. However, BNSF has misrepresented the string model data by using an inconsistent time scale on the graph that compresses or elongates the time statistics in order to create the appearance that trains are stopping and starting instantaneously or experience widely fluctuating speeds.¹⁶ When BNSF's analysis is corrected to display the graph on a consistent

¹⁴ OTP Rebuttal Ev. at I-71 through I-73 and III-B-23 through III-B-29. For example, BNSF does not use the RTC Model to model the entire SARR, but only models segments, which results in trains that disappear between segments, an inability to link trains between segments, and trains that leave one segment and appear at the next segment either before or after their predicted departure times on the subsequent segment. Moreover, BNSF uses inconsistent inputs on its different segment simulations. BNSF's RTC Model simulation also uses incorrect locomotive horsepower ratings and contains impossible rates of acceleration and deceleration. In addition, BNSF has not provided the source code for the RTC Model, contrary to Board precedent. Nor has BNSF shown that the RTC Model is an accepted tool for developing and testing an operating plan for rail projects on the same scale as the OTRR.

¹⁵ BNSF has attempted to equate the string model with the string programs used in prior SAC cases. BNSF Motion at 11. Otter Tail's string model, however, was submitted after the string programs in those other cases and is an updated version that has fixed the problems that were identified in those prior cases. Therefore, the Board must consider Otter Tail's string model on its own merits.

¹⁶ OTP Rebuttal Ev. at III-C-26.

BNSF also suggests that Otter Tail's rebuttal string model has been drastically altered because there are over 100 lines of programming code that have been modified.²⁰ As Mr. Crowley explains in his Verified Statement, however, these code changes fall into just four categories.²¹ Over half of the changes, which fall into the first category, had no impact on the programming logic. Rather, they removed extraneous comments, codes and blank lines from the program in order to free up space for the programming code changes in the other three categories.²² The remaining forty-nine lines of code changes responded to BNSF's criticism of the string model's ability to track loading and unloading times; BNSF's criticism that the same track grade was applied to trains traveling in both directions; and an error that caused the program to lose track of the dispatch date for a limited subset of trains.²³ As any experienced computer programmer knows, the integrated nature of computer programming, imposed by the programming language, sometimes necessitates the modification of many lines of code in order to address just one issue. Thus, Otter Tail's modifications to the string model on rebuttal are not the massive rewrite of a "fundamentally flawed" program that BNSF attempts to portray.

Looking beyond those portions of BNSF's Motion that are an impermissible reply to a reply, the thrust of BNSF's argument is that the Board should reject any evidence that is based on a revised string program presented on rebuttal "because of the difficulty in evaluating the output and statistics generated by the string program."²⁴ According to BNSF, "the validity of a

²⁰ Leopold/Monks V.S. at 1-2.

²¹ Crowley V.S. at 9-10.

²² *Id.* at 10-11. As Mr. Crowley explains, the Visual FoxPro programming language of the string model has capacity limitations, which required Otter Tail to eliminate extraneous comments and codes from the program in order to create sufficient programming space to address BNSF's criticisms.

²³ *Id.* at 9-10.

²⁴ BNSF Motion at 11.

string program analysis and the statistics generated by the string program can be evaluated only with a detailed and extensive review of the program code and the output files.”²⁵ Because of the complexity of the string program, BNSF concludes that “the Board cannot permit the complainant to offer a new string program analysis on rebuttal and thereby avoid the scrutiny that would be required before any weight could be given to the string program analysis.”²⁶

BNSF’s position, however, would be highly prejudicial to Otter Tail, since it would deprive Otter Tail of any rebuttal opportunity with respect to the string model. This is procedurally improper and unfair, not to mention inconsistent with the *Duke/NS* standards for rebuttal evidence. In addition, the Board repeatedly has observed that rail rate cases are not ordinary commercial litigation because, in addition to being an adjudicator, the Board is charged with carrying out the national rail transportation policy, which requires the Board to make a fair and informed assessment of the reasonableness of a challenged rate.²⁷ Because Otter Tail has the burden of proof, it is entitled to an opportunity to respond to BNSF’s reply evidence. Therefore, the Board cannot strike Otter Tail’s evidence if it is “to achieve an appropriate balance so that the adversarial process works in a manner that provides an adequate record upon which the Board may make a fair and informed assessment of the reasonableness of a challenged rate.”²⁸

²⁵ *Id.* at 12. Ironically, although BNSF concedes the importance of being able to evaluate the program code, it has not produced the program code for its RTC Model, which has precluded Otter Tail from being able to scrutinize the RTC Model in any manner that is comparable to BNSF’s scrutiny of the string model. BNSF cannot have it both ways by insisting upon the need for an extensive review of the string model code while refusing to produce the RTC Model’s code.

²⁶ *Id.* at 14.

²⁷ STB Docket No. 42058, *Arizona Elec. Power Coop., Inc. v. The Burlington Northern and Santa Fe Ry. Co. and Union Pac. R.R. Co.*, slip op. at 2 (served Nov. 18, 2003). *See also*, *Xcel* at 22-23.

²⁸ *Xcel* at 23.

Nor is it fair or appropriate, as BNSF suggests, to strike Otter Tail's rebuttal string model because it allows for manual adjustments to trains.²⁹ This feature, in fact, is the strength of the string model over the RTC Model. The string model does not *require* manual adjustments, as BNSF asserts, but rather, it *allows* for them, which the RTC Model does not do. The string model could resolve every train conflict the same as the RTC Model, by adding infrastructure until no conflicts exist. But, this gold-plated approach is not consistent with the CMP objective to design a least cost, efficient SARR. Since many train conflicts can be resolved by holding trains at sidings and adjusting speeds, just as real-world dispatchers do, it is much more efficient to make these adjustments than to add costly infrastructure. The string model allows for manual adjustments in order to foster this more-efficient operation, and therefore, is superior to the RTC Model.³⁰ Thus, the Board cannot strike Otter Tail's rebuttal string model simply because additional effort may be necessary to review and verify Otter Tail's manual adjustments.³¹

Nevertheless, Otter Tail recognizes that BNSF has raised an important issue in its Motion that is entirely separate from whether the Board should strike any of Otter Tail's rebuttal evidence on the string model. It is in the interest of both the Board and the parties to confirm whether or not Otter Tail's proper rebuttal evidence modifications to the string model actually do correct the errors identified in BNSF's reply evidence. BNSF notes that "The Board relies on the parties to carry out [] complex evaluations of another party's evidence" and that it would be difficult for the Board to carry out such a review of the string model on rebuttal.³² Yet, the

²⁹ BNSF Motion at 12-13.

³⁰ Crowley V.S. at 18.

³¹ The review of Otter Tail's manual adjustments is not particularly onerous, since Otter Tail manually intervened to address train conflicts on only 1.42% of trains. OTP Rebuttal Ev. at III-B-36 through III-B-37. Furthermore, the string model output files clearly identify all manual interventions. Crowley V.S. at 17-18.

³² BNSF Motion at 13.

Board must conduct this review in order to provide procedural due process to Otter Tail. This creates a procedural conundrum that has adverse consequences for either party, depending upon the choice made by the Board.

Otter Tail has offered a solution to this conundrum in a “Motion for Technical Conference,” which it has filed contemporaneous with this Reply to BNSF’s Motion to Strike, that would allow the Board to evaluate Otter Tail’s rebuttal evidence on the string model with input from both parties. Although BNSF’s concerns do not establish any basis for the Board to strike Otter Tail’s rebuttal evidence on the string model, the Board can respond to those concerns through Otter Tail’s proposed Technical Conference.

III. OTTER TAIL PROPERLY SUBMITTED REBUTTAL EVIDENCE ON THE FARGO YARD.

BNSF also has moved to strike the portion of Otter Tail’s Rebuttal Evidence that reconfigures the Fargo Yard in order to perform intra- and inter-train (“I&I”) switching. On rebuttal, Otter Tail accepted BNSF’s argument that I&I switching was necessary for certain non-coal traffic, but in order to be consistent with SAC principles of least-cost efficiency, Otter Tail chose to perform the switching at the OTRR’s Fargo Yard, rather than its Glendive Yard as suggested by BNSF. There is nothing improper about Otter Tail’s evidence.

At pages III-B-15 through III-B-16 of its Rebuttal Evidence, Otter Tail reconfigured its Fargo Yard to perform I&I switching on an average of 2.7 non-coal trains per day. This evidence responded to BNSF’s Reply Evidence, at pages III.B-43 through III.B-46, wherein BNSF pointed out that, when Otter Tail rerouted non-coal traffic between Fargo and Snowden, it omitted facilities to perform I&I switching that BNSF currently performs at Minot, North Dakota on *some* of the rerouted traffic. Inexplicably, however, BNSF asserted that this I&I switching “needs” to be performed at the Glendive Yard. Otter Tail accepted the need to perform I&I

switching, but Otter Tail rejected BNSF's unexplained assertion that the I&I switching must be performed at the Glendive Yard, as both impracticable and infeasible.

Instead, Otter Tail chose to perform I&I switching at the Fargo Yard. The OTRR handles non-coal traffic between Fargo and Snowden, via Glendive. Since the OTRR handles only non-coal traffic in overhead train service, it is much more efficient and cost effective for the OTRR to perform I&I switching at either the beginning or end of the movement, than to interrupt the movement in the middle to perform I&I switching at the Glendive Yard.

In *Duke/NS*, the Board held that "where the shipper shows that the railroad's reply evidence is itself unsupported, infeasible or unrealistic, the shipper may supply corrective evidence."³³ Moreover, the Board stated that a "railroad's proposal would be unrealistic if, for example, it is not what the defendant railroad itself does in a comparable situation or what other railroads generally would do in that situation, or otherwise constitutes needless 'gold-plating.'"³⁴ Clearly, BNSF's selection of Glendive over Fargo for I&I switching is an unrealistic choice in light of the obvious inefficiencies it would create in the OTRR's operations, and BNSF's infeasible design of the Glendive Yard.³⁵ The *Duke/NS* standard prevents a railroad from positing an inefficient operation and then denying the shipper the right to expose those inefficiencies and propose a more efficient operation.

BNSF misleadingly attempts to distract the Board from this important point by alleging that Otter Tail is avoiding scrutiny of extensive cost and feasibility issues associated with its rebuttal modifications to the Fargo Yard in order to accommodate I&I switching operations. Specifically, BNSF alleges that, "since the Fargo Yard is located in downtown Fargo, it would be

³³ *Duke/NS* at 14-15 [footnote omitted].

³⁴ *Id.* at note 19.

³⁵ See *Crowley V.S.* at 27-29; OTP Rebuttal Ev. at III-B-16 and III-F-26.

necessary to determine...whether buildings would need to be demolished and whether overpasses and bridges would need to be expanded.”³⁶ But, these concerns all represent barriers to market entry.³⁷ BNSF has not demonstrated, or even argued, that it incurred any of these costs when it acquired the right-of-way for the Fargo Yard. This fact demonstrates that BNSF’s feasibility claim is a red-herring.

BNSF also inappropriately tries to equate Otter Tail’s rebuttal modifications to the Fargo Yard with rebuttal evidence that the Board struck in *Duke/CSX I*.³⁸ The complainant in that case attempted to change the entire location of an interchange yard and a tunnel. These changes significantly altered the configuration of the SARR, which measurably affected revenue allocations, and created one of the longest rail tunnels in the Eastern United States.³⁹ In contrast, Otter Tail’s addition of track to its Fargo Yard does not change the location of the yard or the length of the OTRR’s route, does not impact revenue allocations, and affects only the operations of the very few trains that receive I&I switching.⁴⁰

Furthermore, *Duke/CSX I* was decided prior to the Board’s refinement of its rebuttal standards in *Duke/NS*. In recognition of this fact, the Board’s final decision on the merits in *Duke/CSX II* did evaluate the complainant’s stricken evidence in light of the subsequently

³⁶ BNSF Motion at 16, note 21.

³⁷ The need to construct or relocate road crossings and overpasses is a well-established barrier to entry when the incumbent did not incur such costs, which typically is the case whenever the railroad preceded the road. *Arizona Pub. Serv. Co. and PacifiCorp v. The Atchison, Topeka and Santa Fe Ry. Co.*, 2 STB 367, 386, 410-11 (1997); STB Docket No. 42022, *FMC Wyoming Corp. v. Union Pac. R.R. Co.*, slip op. at 120, 127 (served May 12, 2000); *West Texas Utilities Co. v. The Burlington Northern R.R. Co.*, 1 STB 638, 672 (1996). Although the Board has not specifically addressed whether the need to demolish or relocate existing buildings within a right-of-way is a barrier to entry, in the absence of any evidence that such structures existed when BNSF originally acquired the Fargo Yard right-of-way, they too should be treated as barriers to entry under the same rationale as road crossings.

³⁸ STB Docket No. 42070, *Duke Energy Corp. v. CSX Transp. Inc.*, slip op. (served March 25, 2003) (*“Duke/CSX I”*).

³⁹ See Crowley V.S. at 23-24, citing *Duke/CSX I* at 2 and 4.

⁴⁰ Crowley V.S. at 24-25.

adopted *Duke/NS* standards and concluded that, even if the railroad's claims were unrealistic, thus justifying the complainant's rebuttal evidence, that evidence had failed to correct the deficiency pointed out by the railroad.⁴¹ *Duke/CSX I* has no relevance as to whether BNSF's reply evidence on I&I switching at the Glendive Yard in this case is feasible, supported, and realistic.

Finally, Otter Tail's modifications to its Fargo Yard are similar to rebuttal evidence modifications that the complainant made to its SARR yards in *WPL*.⁴² In that case, the complainant added 4.7 miles of track at two existing yards in response to criticism in the railroad's reply evidence. Because the complainant explained how its track additions would improve operations, the Board accepted the evidence. The Board clearly understood that the addition of track to an *existing* yard does not amount to an impermissible reconfiguration of the SARR.

IV. OTTER TAIL'S PRODUCTIVITY ADJUSTMENT TO THE DCF INFLATION INDEX IS A PROPER RESPONSE TO BOTH BNSF'S REPLY EVIDENCE AND RECENT BOARD DECISIONS.

BNSF seeks to strike two components of Otter Tail's rebuttal evidence on the inflation index for operating costs in the DCF model. First, BNSF challenges workpapers that Otter Tail has submitted containing articles on various ongoing efforts by the railroad industry to improve productivity. Second, BNSF asks the Board to strike Otter Tail's proposal for a modified inflation index. Although BNSF attacks each component separately, the first component both stands on its own and supports the second component. BNSF has not justified its request to strike any of Otter Tail's inflation index evidence from the record.

⁴¹ STB Docket No. 42070, *Duke Energy Corp. v. CSX Transp. Inc.*, slip op. at 14 (served Feb. 4, 2004) ("*Duke/CSX I*").

⁴² STB Docket No. 42051, *Wisconsin Power & Light Co. v. Union Pac. R.R. Co.*, slip op. at 65 (served Sept. 13, 2001) ("*WPL*").

In its Opening Evidence, Otter Tail advocated use of the RCAF-A to forecast the OTRR's future inflation. The RCAF-A is a well-established and accepted rail cost index published by the Board, which rendered it perfectly suitable for use in a stand-alone model. Thus, Otter Tail's evidence was feasible and supported. In fact, BNSF does not object to the RCAF, itself, but merely to Otter Tail's choice of the RCAF-A, which reflects productivity improvements, rather than the RCAF-U, which does not. On Reply, BNSF argued that the RCAF-U is more appropriate than the RCAF-A because the OTRR allegedly would not realize any productivity gains, since it is designed as a least-cost, most-efficient railroad from the outset.

Otter Tail's rebuttal evidence responds directly to the allegations in BNSF's Reply Evidence and to the most recent Board precedent.⁴³ First, Otter Tail's Rebuttal Evidence identifies the types of productivity improvements that the OTRR would experience along with the rail industry as a whole. BNSF's argument that Otter Tail was required to present this evidence on Opening is tantamount to requiring Otter Tail to anticipate and respond to BNSF's arguments before BNSF even presents them.

Second, Otter Tail presented evidence of rail productivity gains in response to the most recent Board precedent. In two decisions served after the filing of Opening and Reply Evidence in this case, the Board acknowledged that a SARR in fact is likely to experience productivity gains along with the rail industry as a whole, but to a lesser degree.⁴⁴ In those decisions, the Board concluded that, in the absence of an alternative to the RCAF-A or the RCAF-U that would

⁴³ See *Crowley V.S.* at 31-33, discussing how Otter Tail's Rebuttal Evidence responds to BNSF's Reply Evidence.

⁴⁴ *Carolina Power & Light Co. v. Norfolk Southern Ry. Co.*, STB Docket No. 42072, slip op. at 27-28 (served Dec. 23, 2003) ("*CPL*"); *Duke/CSX II* at 29-30.

better reflect the expected productivity gains of the SARR, it would adopt the RCAF-U because the productivity gains were more likely to be closer to the RCAF-U than the RCAF-A.

Otter Tail's Rebuttal Evidence properly responded to these decisions in two ways. First, Otter Tail presented evidence that the OTRR's productivity gains would be much closer to the RCAF-A than the RCAF-U. Therefore, in accordance with the Board's logic in those decisions, it should adopt the RCAF-A in the absence of a more appropriate alternative. Second, Otter Tail included an alternative indexing methodology that would be more accurate than either the RCAF-A or the RCAF-U by more precisely measuring just those productivity gains that the OTRR would experience. Contrary to BNSF's claims, Otter Tail *did not* use this alternative methodology in its DCF analysis.⁴⁵

On rebuttal, the complainant may demonstrate that its opening evidence was feasible and supported.⁴⁶ This is precisely what Otter Tail's rebuttal evidence on a DCF inflation index does. Furthermore, when a complainant shows that the railroad's reply evidence is itself unsupported, infeasible or unrealistic, the complainant may supply corrective evidence.⁴⁷ Otter Tail's Rebuttal Evidence also demonstrates that BNSF's evidence based on the RCAF-U is unsupported and unrealistic because the OTRR in fact would experience productivity gains. Moreover, the Board itself acknowledged this fact in its recent decisions, concluding that neither the RCAF-U or the RCAF-A accurately reflects a SARR's likely productivity gains. Thus, Otter Tail appropriately refined its evidence, on rebuttal, in response to these determinations.⁴⁸

⁴⁵ Crowley V.S. at 34-35.

⁴⁶ *Duke/NS* at 14-15.

⁴⁷ *Id.*

⁴⁸ At the oral argument held in *Pub. Service Co. of Colorado d/b/a Xcel Energy v. The Burlington Northern and Santa Fe Ry. Co.*, STB Docket No. 42057, on March 18, 2004, the Board expressly requested the parties to

V. CONSISTENT WITH BOARD PRECEDENT, OTTER TAIL UPDATED ITS COST OF CAPITAL TO REFLECT THE MOST RECENT DATA.

BNSF wrongly accuses Otter Tail of improperly modifying the OTRR's cost of debt in its Rebuttal Evidence. At pages III-G-1 to III-G-3 of its Rebuttal Evidence, Otter Tail updated its cost of debt to reflect a lower interest rate that would be available through refinancing in 1Q03. There is nothing improper about this modification to Otter Tail's Opening Evidence. Otter Tail merely updated its cost of debt consistent with Board precedent.

In *WTU*⁴⁹, the Board acknowledged that a SARR could refinance and lower its debt costs in the future. In response to BN's argument that the cost of debt should float throughout the DCF period, the Board observed:

Because this debt would be incurred during the construction period, and because debt financing establishes a set interest rate over a set time period, WTRR would not have to pay prevailing debt rates in periods following construction to prevent the suppliers of debt capital from withdrawing their investment. *Indeed, not only could WTRR continue to pay off its debt at the contracted rate, it could refinance and lower its debt costs if the opportunity provided itself in the future.*⁵⁰

Similarly, in *McCarty Farms*, the Board held that a SARR "would be free to choose a method of financing its debt to minimize that cost."⁵¹

Since a complainant cannot predict if or when interest rates will decline, allowing the SARR to refinance its debt, a complainant typically must file a petition to reopen a decision based on changed circumstances if and when the opportunity arises. However, during the time

propose alternative indexing methodologies in post-argument briefs. Otter Tail's Rebuttal Evidence also has been submitted in response to this expression of interest by the Board in alternatives to the RCAF-A and RCAF-U.

⁴⁹ *West Texas Utilities Co. v. The Burlington Northern R.R.*, 1 S.T.B. 638 (1996) ("*WTU*").

⁵⁰ *Id.* at 712 [emphasis added].

⁵¹ *McCarty Farms, Inc. v. Burlington Northern, Inc.*, 2 S.T.B. 460, 523, note 126 (1997) ("*McCarty Farms*").

that has elapsed since the filing of Opening Evidence in this proceeding, interest rates have declined significantly. This allowed Otter Tail, on rebuttal, to assume the OTRR would refinance its debt at lower rates in 1Q03, which is just what the Board assumed a least-cost, efficient railroad would do in *WTU* and *McCarty Farms*. It makes no sense to require Otter Tail to file a petition to reopen a case that has not even been decided yet.

There is ample precedent for updating the cost of capital in SAC proceedings even after the close of the administrative record. In *FMC*, the Board determined the 1998 railroad-industry cost of capital after the close of the record in order to reflect the most current data available.⁵² The Board did the same thing in *APS* in order to include 1995 in the average cost of equity, which had the effect of slightly reducing the average equity cost for the SARR.⁵³ Furthermore, the Board has updated other SAC inputs as new and better information became available prior to a final decision.⁵⁴ Thus, there is nothing inappropriate in Otter Tail's update of the OTRR's cost of debt on rebuttal.

VI. STATISTICAL ANALYSIS OF REROUTED TRAFFIC.

No where is BNSF's Motion to Strike more obviously an impermissible reply to a reply than its response to Otter Tail's statistical analysis of transit times on rerouted non-coal traffic moving between Fargo and Snowden. BNSF *does not* allege that Otter Tail's analysis exceeds the *Duke/NS* standards for the proper scope of rebuttal evidence. Rather, BNSF concocts a non-existent Board standard for expert evidence, based upon a distorted application of Rule 702 of the Federal Rules of Evidence. BNSF cites no Board precedent that adopts Rule 702; nor does

⁵² *FMC Wyoming Corp. v. Union Pac. R.R. Co.*, STB Docket No. 42022, slip op. at 178, note 248 (served May 12, 2000) ("*FMC*").

⁵³ *Arizona Pub. Serv. Co. v. The Atchison, Topeka and Santa Fe Ry. Co.*, 2 S.T.B. 367, 438 (1997) ("*APS*").

⁵⁴ See *Duke/NS* at 65; *Xcel* at 55; *CPL* at 16.

BNSF cite even a single judicial decision ever rejecting analyses similar to Otter Tail's rebuttal evidence. Although BNSF cites the Supreme Court's decision applying Rule 702 in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993), BNSF does not even mention, much less evaluate, any of the relevant factors that the Court identified. Thus, it is quite apparent that BNSF cursorily mentions Rule 702 simply as a pretext for responding to Otter Tail's rebuttal evidence, including the submission of new expert testimony and analyses. The Board should not permit this abuse of its procedures, by refusing to give any consideration to BNSF's improper evidence.

The testimony of BNSF's own expert witness, Matthew G. Mercurio, proves without a doubt that Otter Tail's statistical analyses are not "junk science," as BNSF so cavalierly contends.⁵⁵ Attachment B to Dr. Mercurio's verified statement is an article that acknowledges that the overlap methodology used in Otter Tail's analysis to compare the confidence intervals of the data is routinely used in many analyses. In fact, the authors of the article state that a "nonexhaustive search of abstracts and articles in the health sciences found more than 60 articles in which the method of examining overlap was used or recommended to demonstrate significance and/or lack of significance, either formally or informally" and they proceed to cite 22 specific examples of those articles.⁵⁶ Such widespread acceptance of Otter Tail's overlap methodology undoubtedly refutes any admissibility arguments under Rule 702, even if it were applicable. Because this fact comes from BNSF's own evidence, it confirms BNSF's ulterior, and improper, objective of using its Motion to Strike to reply to Otter Tail's rebuttal evidence.

⁵⁵ See BNSF Motion at 21.

⁵⁶ BNSF Motion, Verified Statement of Matthew G. Mercurio ("Mercurio V.S."), Attachment B, page 1.

If the Board gives any consideration to BNSF's impermissible reply evidence, it is important to note at the outset that BNSF's critique of Otter Tail's statistical analysis of the rerouted traffic is limited. Otter Tail presented *two* statistical analyses in support of its non-coal traffic reroutes between Fargo and Snowden at pages III-A-44 through III-A-52 of its Rebuttal Evidence. The first analysis, at pages III-A-45 through III-A-49, demonstrated that, despite *slightly* longer average transit times, the OTRR provides more consistent and reliable service than BNSF on the rerouted traffic and at lower rates. The second analysis, at pages III-A-49 through III-A-52, demonstrated that the OTRR's transit times, in fact, are not statistically different from BNSF's transit times on the rerouted traffic. BNSF's expert witness, Dr. Mercurio, clearly states that his testimony pertains *only* to Otter Tail's second statistical analysis:

In particular, I have been asked to review the statistical analysis contained in Otter Tail's Rebuttal Narrative *at III-A-49 through III-A-52*. Otter Tail offers this analysis to demonstrate that the transit times for rerouted non-coal traffic produced by Otter Tail's string program *are not "statistically different"* from real-world BNSF transit times for the same trains.⁵⁷

Thus, regardless whether or not the Board agrees with Dr. Mercurio's critique of Otter Tail's second analysis, Otter Tail's first analysis remains a statistically sound, unchallenged, and independent basis in support of the OTRR reroute of non-coal traffic between Fargo and Snowden.⁵⁸

In addition, as explained at pages 43-51 of Mr. Crowley's Verified Statement, BNSF's impermissible response to Otter Tail's second analysis is baseless. First, Dr. Mercurio suggests, without actually asserting, that BNSF's mean transit times are distorted by a "small number of

⁵⁷ Mercurio V.S. at 2 [emphasis added] [footnote omitted].

⁵⁸ See Crowley V.S. at 39-42.

outliers” that in fact favor the OTRR by comparison.⁵⁹ However, Mr. Crowley refutes this suggestion by eliminating the most extreme outliers from the BNSF transit time data between Fargo and Snowden.⁶⁰ Using both a conservative and a liberal definition of “outliers,” Mr. Crowley recalculated BNSF’s average transit times and determined that removing the outliers would reduce BNSF’s overall average transit time only marginally from 13.8 hours to either 13.3 hours (at three standard deviations) or 13.1 hours (at two standard deviations), compared to the OTRR’s overall average transit time of 14.4 hours. The standard deviations of these averages, 5.9 and 4.8 respectively, still substantially exceed the OTRR’s standard deviation of 1.9. Thus, even removing the BNSF’s outlier transit times, the OTRR still provides significantly more reliable and consistent service at only slightly longer transit times.

Second, Dr. Mercurio contends that Otter Tail should have performed a two-sample t-test of mean equality, instead of an overlap test, to determine whether the means of the two distributions are the same.⁶¹ But, as Mr. Crowley explains, Dr. Mercurio’s two-sample t-test is inappropriate because it is in the form of a Behrens-Fisher problem.⁶² A Behrens-Fisher problem is not appropriate when the ultimate objective is to measure the impact upon both the mean and the variance of individual transit times. In this case, Otter Tail is interested in both the impact on the mean and variance of the transit times, rather than just the average time for the rerouted traffic. Even though the difference in the overall mean transit time is 36 minutes (14.4 hours less 13.8 hours) in favor of BNSF, the variance of movements is 5.9 hours on the BNSF compared to only 1.8 hours on the OTRR. The OTRR’s reduction in transit time variation from one

⁵⁹ Mercurio V.S. at 3.

⁶⁰ Crowley V.S. at 44-45.

⁶¹ Mercurio V.S. at 11-13.

⁶² Crowley V.S. at 47-49.

movement to the next, not to mention its lower rates, far outweighs the value of a slight increase to the overall average transit time.

Third, in addition to his critique of Otter Tail's statistical methodology, Dr. Mercurio ventures into areas beyond his acknowledged expertise, to anecdotally critique Otter Tail's analysis. Specifically, he surmises that the distribution of the OTRR's transit times are artificially "tight", relative to the distribution of BNSF's transit times, because the string model "intentionally excludes real-world variation both in terms of events such as maintenance and other track outages and in terms of the time spent on routine tasks such as fueling, switching, and changing crews in the yards."⁶³ Dr. Mercurio is simply wrong for the reasons stated by Mr. Crowley at pages 41-42 of his Verified Statement. Furthermore, Dr. Mercurio does not claim to have any of the expertise required to make this assessment of the string model, and thus the Board should not give any credence to his unsubstantiated statements.

Finally, BNSF's vigorous assault upon Otter Tail's reroute of non-coal traffic between Fargo and Snowden is suspect by reason of its timing. The Board adopted its service efficiency standard for reroutes in *TMPA*, which was served on March 21, 2003. Nevertheless, when BNSF filed its Reply Evidence on October 8, 2003, it challenged only Otter Tail's reroute of the "Z-Train" non-coal traffic, despite having evidence at that time that most of the rerouted traffic experienced longer average transit times on the OTRR reroute. Thus, one can only conclude that BNSF itself recognized that shippers of non-Z-Train traffic considered other factors (*e.g.*, reliable and consistent service, reduced rates) to be a more important measure of transportation efficiency than average transit times.

⁶³ Mercurio V.S. at 4.

BNSF abruptly altered its position when it filed Supplemental Reply Evidence. BNSF decided to reject all of the rerouted non-coal traffic between Fargo and Snowden on the basis of the Board's *Duke/NS* decision, which had adopted a rebuttable presumption that reroutes longer than the current route are less efficient. Thus, it appears that BNSF attempted to take advantage of the subsequent adoption of a rebuttable presumption in *Duke/NS* to prove a fact that it did not even allege in its Reply Evidence, even though it possessed evidence of both greater distances and transit times on the OTRR reroute at that time (and in fact relied on that evidence to exclude Z-Trains). The Board should take note of this fact when it weighs the importance of more consistent and reliable service against slightly faster transit times in determining the efficiency of Otter Tail's reroute of non-coal traffic between Fargo and Snowden.

VII. CONCLUSION.

The Board should deny BNSF's Motion to Strike and disregard BNSF's impermissible reply to a reply. BNSF has failed to demonstrate that the five categories of alleged new evidence exceed the proper scope of rebuttal set forth in *Duke/NS*. Neither has BNSF demonstrated that a sixth category of allegedly unreliable expert analysis violates any applicable standard for expert testimony. Indeed, much of BNSF's Motion pays only lip service to the standards for proper rebuttal evidence, without even attempting to demonstrate how Otter Tail's evidence violates those standards. Clearly, BNSF's objective is not to strike Otter Tail's evidence, but to respond to it. The Board should not permit either result.

BNSF, however, has raised certain procedural matters of importance to both parties concerning the Board's review of Otter Tail's rebuttal string model, but which do not support a

Motion to Strike. Otter Tail has asked the Board to address those matters in a Motion for Technical Conference, filed contemporaneous with this Reply to BNSF's Motion.

Respectfully submitted,



George Koeck, Esq.
General Counsel
Otter Tail Corporation
4334 18th Avenue, SW, Suite 200
Fargo, ND 58103
(701) 451-3567

Nicholas J. DiMichael, Esq.
Jeffrey O. Moreno, Esq.
Thompson Hine LLP
1920 N Street, N.W., Suite 800
Washington, DC 20036
(202) 331-8800

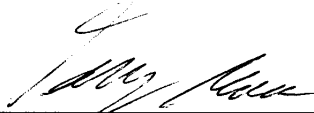
Dated: July 20, 2004

*Attorneys for Complainant
Otter Tail Power Company*

CERTIFICATE OF SERVICE

I hereby certify that on this 20th day of July, 2004, I served a copy of "Complainant's Reply to Defendant's Motion to Strike" by hand delivery to counsel for Defendant at the following address:

Samuel M. Sipe, Jr., Esq.
Anthony J. LaRocca, Esq.
Brooke L. Gaede, Esq.
Steptoe & Johnson
1330 Connecticut Avenue, NW
Washington, DC 20036



Jeffrey O. Moreno

WARD UGGERUD
VERIFIED STATEMENT

**BEFORE THE
SURFACE TRANSPORTATION BOARD**

OTTER TAIL POWER COMPANY,	)	PUBLIC VERSION
	)	
Complainant,	)	
	)	
v.	)	
	)	
THE BURLINGTON NORTHERN AND	)	Docket No. 42071
SANTA FE RAILWAY COMPANY,	)	
	)	
Defendant.	)	

VERIFIED STATEMENT OF WARD UGGERUD

1. My name is Ward Uggerud. I am over eighteen years of age and competent to provide this Verified Statement in support of Otter Tail Power Company's ("Otter Tail") Reply to the Motion to Strike filed by the Burlington Northern and Santa Fe Railway Company ("BNSF") in this proceeding. I previously sponsored a section of Otter Tail's Opening Evidence, filed June 13, 2003.

2. Since May of 2002, I have held the position of Senior Vice President for Otter Tail and my responsibilities include new business development, electric generation, environmental matters, and power marketing. I have over thirty years of experience working for Otter Tail, since starting as an engineer in July of 1971. I have knowledge of Otter Tail's actions relating to shipments of coal to the Big Stone electric generating station in September and October of 2003.

3. I have read Section II of the Highly Confidential Version of the Motion to Strike ("Motion") that BNSF filed in this proceeding on June 8, 2004. My testimony in this Verified Statement focuses on this section of the Motion, in which BNSF asks the Board to strike

evidence related to Otter Tail's past shipments of Black Thunder coal to Big Stone. Specifically, I wish to refute BNSF's spurious allegation that Otter Tail shipped coal from Black Thunder in October of 2003 in order to bolster its position on certain disputed issues in this proceeding—or as BNSF puts it, “to game the SAC results by creating the appearance that Otter Tail intended to source future coal needs from the Black Thunder mine.”¹

4. In fact, Otter Tail [REDACTED] [REDACTED] shipments of coal from Black Thunder in October of 2003 occurred in the ordinary course of business, as Otter Tail selected a fuel source for the station after its current contract with RAG Coal West, Inc. (“RAG”) expires [REDACTED].

5. Otter Tail contemplated burning Black Thunder coal at Big Stone during the years 2000 and 2001.² However, through a test burn conducted in June of 2000, Otter Tail discovered that [REDACTED]

[REDACTED] Thus, Otter Tail suspended shipments and burned only limited quantities of Black Thunder coal during those years.

6. In the fall of 2002, the ESP was successfully converted to an Advanced Hybrid Particulate Collector (“AHPC”) in conjunction with the U.S. Department of Energy for a full-scale demonstration project under the Power Plant Improvement Initiative. One of the goals of the AHPC project was to increase Big Stone's fuel flexibility, while maintaining improved particulate emission control.

¹ See Motion at 11.

² See Exhibit A, attached hereto, “Coal Supply Agreement” between Otter Tail and Arch Coal Sales Company, Inc., dated March 1, 1999.

³ See Exhibit B, attached hereto, “Memo from Tyler Albertson to Jeff Endrizzi,” dated July 12, 2000, produced at OTP 002345.

7. On September 8, 2003, Otter Tail issued a fuel supply Request for Proposal (“RFP”) for the Big Stone station for deliveries to begin on [REDACTED] Otter Tail submitted its RFP to a number of coal suppliers,⁵ [REDACTED] [REDACTED] Otter Tail issued this RFP exactly one month before BNSF filed its Reply Evidence in this proceeding.

8. On October 6, 2003, Otter Tail received a response from ARCH to its September 8th RFP, which proposed to supply coal from the Black Thunder mine.⁷ [REDACTED]
[REDACTED]

[REDACTED] Otter Tail received ARCH’s proposal two days before BNSF filed its Reply Evidence in this proceeding.

9. Given the previous operational issues that Big Stone encountered burning Black Thunder coal, and in order to confirm the coal’s suitability in the plant’s AHPC, Otter Tail decided that it would be necessary to conduct a test burn.

10. Accordingly, Otter Tail contacted ARCH for the purpose of arranging shipments of Black Thunder coal for use in a test burn at Big Stone. On October 13, 2003, confirmation was received from ARCH for the test burn of 4 to 5 train sets of Black Thunder

⁴ See Exhibit C, attached hereto, “Request for Proposal – Supply of Fuel,” dated September 8, 2003.

⁵ See Exhibit D, attached hereto, consisting of Request for Proposal cover letters from Otter Tail to [REDACTED], dated September 8, 2003.

⁶ See Exhibit E, attached hereto, Request for Proposal cover letter from Otter Tail to ARCH, dated September 8, 2003.

⁷ See Exhibit F, attached hereto, “Coal Sales Proposal” from ARCH to Otter Tail, dated October 6, 2003.

⁸ See Exhibit G, attached hereto, “Big Stone Plant Coal Bid Comparison Option 1” and “Big Stone Plant Coal Bid Comparison Option 2.”

coal beginning on October 17th.⁹ Four train sets were received and the test burn was conducted during October 20th to 27th. [REDACTED]

[REDACTED]

11. [REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

12. Otter Tail's shipments of Black Thunder coal in October of 2003 were not inconsistent with its previous notification to BNSF of coal shipments from RAG's Belle Ayr and Eagle Butte mines. When Mr. Keith Kelly of Otter Tail advised BNSF, on or about September 24, 2003, of coal shipments from Belle Ayr and Eagle Butte for October, Otter Tail's fuel supply RFP had been pending for just over two weeks. ARCH had not yet submitted its response proposing the Black Thunder mine as a fuel source. Accordingly, during September of 2003, neither Mr. Kelly, nor Otter Tail had any reason to advise BNSF that Otter Tail needed to ship coal from any mine other than Belle Ayr and Eagle Butte, Big Stone's primary fuel sources pursuant to the existing RAG contract. Indeed, no such need arose until October 6, 2003 when Otter Tail [REDACTED] decided to test burn Black Thunder coal. Accordingly, it is both wrong and misleading for Mr. Stephen C. Behn, Director of

⁹ See Exhibit H, attached hereto, "Email from Alan Godby to Dennis Bowman," dated October 13, 2003.

¹⁰ See Exhibit I, attached hereto, "Memo from Bill Swanson to Jeff Endrizzi," dated Nov. 10, 2003.

¹¹ See Exhibit J, attached hereto, [REDACTED]

Equipment and Service Planning for the BNSF Coal Marketing Group, to state that Otter Tail's decision to ship Black Thunder coal was "contrary to its plan, as of September 24, 2003."¹²

Plainly, the Black Thunder shipments resulted from new and unforeseen circumstances that arose subsequently to Mr. Kelly's communication with BNSF.

13. Similarly, there is no factual basis whatsoever for BNSF's accusation that Otter Tail ordered deliveries of coal from Black Thunder in October of 2003 in order to "game the SAC results" of this proceeding.¹³ On the contrary, Otter Tail made the decision to ship coal from Black Thunder in the ordinary course of business.

14. [REDACTED] e

[REDACTED]

[REDACTED]

¹² See Verified Statement of Stephen C. Behn at 2; Motion at 9-10.

¹³ See Motion at 10.

VERIFICATION

I, Ward Uggerud, verify under penalty of perjury, under the laws of the United States, that the foregoing is true and correct. Further, I certify that I am qualified and authorized to file this Verified Statement. Executed on this 23rd day of June, 2004.



Ward Uggerud
Senior Vice President
Otter Tail Power Company

ALL EXHIBITS HAVE BEEN REDACTED FROM THE PUBLIC VERSION OF THIS DOCUMENT.

THOMAS D. CROWLEY
VERIFIED STATEMENT

BEFORE THE
SURFACE TRANSPORTATION BOARD

OTTER TAIL POWER COMPANY	)	
	)	
Complainant,	)	
	)	
v.	)	Docket No. 42071
	)	
THE BURLINGTON NORTHERN	)	
AND SANTA FE RAILWAY COMPANY,	)	
	)	
Defendant.	)	

Verified Statement
of
Thomas D. Crowley
President
L. E. Peabody & Associates, Inc.

On Behalf
of
Otter Tail Power Company

Due Date: July 20, 2004

TABLE OF CONTENTS

	<u>PAGE</u>
I. INTRODUCTION	1
II. SUMMARY AND FINDING	4
III. RESPONSE TO BNSF’S CRITIQUE OF OTTER TAIL’s STRING MODEL	7
A. OTTER TAIL HAS NOT SUBSTANTIALLY CHANGED ITS STRING MODEL IN REBUTTAL	9
B. OTTER TAIL CAN REPLICATE THE RESULTS OF ITS REBUTTAL STRING MODEL PROGRAM	11
C. BNSF MISIDENTIFIED TRAINS WAITING ON PASSING SIDINGS AS COLLISIONS	13
D. THE STRING MODEL PROGRAM IS TRANSPARENT IN ITS OPERATION, AND IN THE DISPLAY OF ITS CODES	14
E. OTTER TAIL’S MANUAL INTERVENTIONS ARE DOCUMENTED, REASONABLE AND VERIFIABLE	16
F. THE SUPPOSED IMBALANCE OF LOADED AND EMPTY TRAINS IS A FUNCTION OF THE EFFICIENCY OF THE SARR AND NOT A PROBLEM WITH THE STRING DIAGRAM MODEL ...	19
IV. RESPONSE TO BNSF’S ASSERTIONS REGARDING THE FARGO YARD	21
A. THE CAPACITY ADDITIONS AT THE FARGO YARD WERE LIMITED IN SCOPE AND DO NOT CONSTITUTE RECONFIGURATION OF THE SARR	23
B. BNSF’S CLAIMS ABOUT DEMOLITION AND RELOCATION COSTS ARE UNFOUNDED	25

TABLE OF CONTENTS

	<u>PAGE</u>
C. BNSF'S ADDITION OF TRACK FOR I&I SWITCHING AT GLENDIVE IS UNFEASIBLE	27
V. RESPONSE TO BNSF'S ASSERTIONS REGARDING RAILROAD PRODUCTIVITY	30
A. OTTER TAIL'S REBUTTAL EVIDENCE WAS PRESENTED TO DIRECTLY REFUTE BNSF'S REPLY EVIDENCE	31
B. BNSF INCORRECTLY STATES THAT OTTER TAIL'S REBUTTAL INFLATION FORECAST IS BASED ON THE EIA'S FORECASTED CHANGES IN PRODUCTIVITY	34
VI. RESPONSE TO BNSF'S ASSERTIONS REGARDING THE SARR'S COST OF CAPITAL	36
VII. RESPONSE TO BNSF'S CRITIQUE OF OTTER TAIL'S TRANSIT TIME ANALYSIS	37
A. DR. MERCURIO HAS NOT CHALLENGED OTTER TAIL'S SERVICE CONSISTENCY ANALYSIS	39
B. DR. MERCURIO'S TRANSIT TIME ANALYSIS DOES NOT ACCURATELY REFLECT THE DATA OR ACTUAL RAILROAD OPERATIONS	43
C. DR. MERCURIO'S HYPOTHETICAL EXAMPLE IS MISLEADING AND FACTUALLY INCORRECT	46
D. DR. MERCURIO'S PERFORMANCE OF A TWO-SAMPLE T-TEST IS INAPPROPRIATE	47
E. OTTER TAIL'S TRANSIT TIMES FOR "M" TRAINS AND "P" TRAINS ARE COMPLETE AND COMPARABLE	49

I. INTRODUCTION

My name is Thomas D. Crowley. I am an economist and President of the economic consulting firm of L. E. Peabody & Associates, Inc. The firm's offices are located at 1501 Duke Street, Alexandria, Virginia 22314 and 5901 N. Cicero Avenue, Suite 504, Chicago, Illinois 60646. My qualifications and experience are attached as Part IV-4 to the Opening Evidence submitted in this proceeding by complainant Otter Tail Power Company ("Otter Tail") on June 13, 2003. I was responsible, along with Mr. Robert D. Mulholland, for the development of the traffic and revenues for the hypothetical stand-alone railroad submitted as part of Otter Tail's Opening and Rebuttal Evidence in this case. I was also responsible, along with Mr. Philip H. Burris, for the evidence related to the reasonableness test of the stand-alone railroad's divisions. In addition, I was responsible, along with Mr. Daniel L. Fapp, for the design and operation of the string program. Finally, I was also responsible for the evidence related to the route miles of the Otter Tail Railroad ("OTRR"), the non-road property investment, the discounted cash flow ("DCF") analysis and the results of the stand-alone cost ("SAC") analysis.

On June 8, 2004, the defendant in this case, The Burlington Northern and Santa Fe Railway Company ("BNSF") filed a "Motion to Strike" ("BNSF Motion") certain portions of Otter Tail's rebuttal evidence.^{1/} The BNSF Motion contends that Otter Tail submitted "new evidence" in its April 29, 2004 Rebuttal Evidence, that exceeds the scope of proper

^{1/} The BNSF Motion also included Verified Statements by Mr. Stephen C. Behn, Messrs. Karl M. Leopold and Terence J. Monks, and Dr. Matthew G. Mercurio.

rebuttal in a SAC case and other new evidence that does not meet the standards for the admissibility of expert testimony. According to the BNSF Motion, the new evidence that exceeds the proper scope of rebuttal includes the following: (1) evidence relating to Otter Tail's purchase of coal from certain southern PRB mines; (2) evidence relating to SARR capacity and operating statistics based on Otter Tail's rebuttal string program; (3) Otter Tail's alleged changes to the SARR's Fargo Yard configuration; (4) alleged new evidence on the inflation index to be used in the DCF analysis, and (5) Otter Tail's alleged modified calculation of the SARR's cost of capital.^{2/} The evidence that allegedly does not meet the admissibility standards for expert testimony relates to Otter Tail's statistical analysis of rerouted general freight traffic.^{3/}

I have been asked by Otter Tail to review the BNSF Motion, and the accompanying Verified Statements of Messrs. Karl M. Leopold and Terence J. Monks ("Leopold/Monks") and Dr. Matthew G. Mercurio, and evaluate BNSF's arguments as they pertain to evidence relating to SARR capacity and operating statistics produced by the string model program, evidence related to the Fargo Yard, evidence related to DCF inflation indexes and cost of capital, and evidence related to BNSF's critique of the statistical analysis of rerouted general freight traffic. Overall, the issues presented and conclusions reached in the BNSF Motion and accompanying Verified Statements are overreaching, and, in a great many instances, deal not with issues of supposed "new evidence," but rather constitute impermissible surrebuttal testimony.

^{2/} See BNSF Motion at 2 and 3.

^{3/} See BNSF Motion at 2 and 3.

A review of evidence submitted in Otter Tail's Rebuttal filings clearly shows that any so-called "new evidence" was in-fact evidence submitted either to directly refute BNSF's Reply Evidence, or to update statistical values with more current values as the STB has done in every recent SAC case. Moreover, much of the information BNSF provides in response to what it calls "new evidence" submitted by Otter Tail does nothing but rehash and extend upon BNSF's Reply Evidence, and appears to be an attempt by BNSF to buttress arguments clearly refuted by Otter Tail in its Rebuttal Evidence.

The remainder of this Verified Statement addresses, in detail, the issues summarized above under the following headings:

- II. Summary and Findings
- III. Response to BNSF's Critique of Otter Tail's String Model
- IV. Response to BNSF's Assertions Regarding the Fargo Yard
- V. Response to BNSF's Assertions Regarding Railroad Productivity
- VI. Response to BNSF's Assertions Regarding The SARR's Cost of Capital
- VII. Response to BNSF's Critiques of Otter Tail's Transit Time Analysis

II. SUMMARY AND FINDING

I have reviewed each of the issues raised in the BNSF Motion. None of the issues, either alone or in aggregate, represent new evidence submitted by Otter Tail in its Rebuttal Evidence. BNSF's claims are neither supported by its evidence or STB precedent, and in many cases, are nothing more than restatements of BNSF's Reply and Supplemental Reply Evidence arguments that Otter Tail disproved in its Rebuttal Evidence. My specific findings are summarized as follows:

1. BNSF's claim that Otter Tail substantially changed its string model program between Otter Tail's Supplemental Opening and Rebuttal Evidence is incorrect. The limited number of programming changes made to the program were in direct response to issues raised by BNSF in its Supplemental Reply Evidence.
2. BNSF's assertion that it cannot replicate the output of Otter Tail's Rebuttal string model run cannot be confirmed. I subsequently ran the program using the inputs provided as workpapers to the BNSF Motion and produced the same results as Otter Tail presented in its Rebuttal Evidence.
3. BNSF's identification of the trains that have head-on collisions is incorrect. The workpapers supporting BNSF's claim actually show one train passing on the OTRR's main line, while a second train waits on a siding.
4. BNSF's statements that the string program model and the model's outputs lack transparency are untrue and nothing but reiterations of BNSF's Reply and Supplemental Reply Evidence claims about the string model program. The fact that BNSF has been able to delve so deeply into the inner workings of the string model program, and modify the program on its own, proves that the program is easily understood and transparent.
5. Otter Tail's manual interventions and adjustments to the string model program's output do not require a substantial level of effort to review and confirm. Contrary to BNSF's assertions that the manual interventions are untested, any adjustments to the output are documented and supported, and accomplish the task of producing the least cost, most efficient stand-alone railroad.

6. BNSF's claim that the imbalance in the number of trains arriving and departing mines on the peak operating day leads to an understatement in operating costs is just another example of BNSF attempting to impute its own inefficiencies on the OTRR. The supposed imbalance is due to the OTRR operating in a more efficient manner than the BNSF, and Otter Tail's use of actual mine departure times to be assured of mine availability.
7. BNSF's claim that Otter Tail's addition of track to the OTRR's existing Fargo Yard constitutes improper reconfiguration of the SARR on Rebuttal is contrary to STB evidentiary guidelines and previous STB decisions. The additions made by Otter Tail were in direct response to issues raised by BNSF in its Reply Evidence, and were so limited in scope and scale to be nearly inconsequential. Moreover, Otter Tail could not adopt BNSF's proposed Reply solution to the problem because it was operationally infeasible.
8. BNSF's assertion that Otter Tail improperly included rebuttal evidence on railroad productivity is incorrect. The productivity information Otter Tail provided in its Rebuttal Evidence was direct support of its case-in-chief which was presented in its Opening Evidence, and did not lead Otter Tail to change its position between Opening and Rebuttal.
9. Contrary to BNSF's statements, Otter Tail did not use a modified RCAF-A forecast to index its operating costs on Rebuttal. Otter Tail utilized, in both its Opening and Rebuttal base case DCF models, unmodified forecasts of the RCAF-A produced by an independent, third-party forecaster.
10. Otter Tail's use of 2003 cost of railroad debt in its DCF model is not new evidence, as asserted by BNSF, but rather the updating of cost indices that is performed in every recent SAC case.
11. BNSF has not refuted Otter Tail's evidence that the OTTR will provide the same, or better, levels of service than is currently provided by the BNSF. Otter Tail demonstrated that the OTRR will provide more consistent, reliable service than that currently provided by the BNSF. Otter Tail also demonstrated that its transit times take into consideration operating and maintenance activities that BNSF claims artificially shrinks the variability of OTRR transit times.
12. The hypothetical example produced by Dr. Mercurio in an attempt to discredit Otter Tail's transit time analysis is flawed, and unrepresentative of typical railroad operations. Moreover, he uses the ill-conceived example to rebut a claim not made by Otter Tail.

13. Dr. Mercurio's use of a two-sample t-Test is inapplicable in this situation. Dr. Mercurio's application of the two-sample t-Test between BNSF and OTRR service levels is in the form of a Behrens-Fisher problem, which is inappropriate when testing for levels of service, as is the case here.
14. Dr. Mercurio's statements that OTRR manifest and premium manifest trains have artificially low transit times because Otter Tail failed to include time for switching is without merit. Otter Tail demonstrated in its Rebuttal Evidence that manifest trains moving on the OTRR system do not receive any switching between Fargo and Snowden as claimed by Dr. Mercurio. Otter Tail also demonstrated that the limited amount of I&I switching on premium manifest trains was accounted for in the OTRR transit times.

III. RESPONSE TO BNSF'S CRITIQUE OF OTTER TAIL'S STRING MODEL

The string model program used by Otter Tail to assess the capacity of the OTRR was described in Otter Tail's Opening Exhibit III-C-2 and in the underlying workpapers that were included in Otter Tail's Opening Errata and Supplemental Opening workpapers supporting Section III-C, including the specific computer program and all input and output files. The string model program is a sophisticated, yet straightforward, user-friendly model that incorporates the fixed physical facilities of the OTRR (single track/double track/passing siding locations, yards, locomotive fueling facilities, crew-change locations, etc.) and 13 different time variables that reflect the OTRR's train operations. The STB and BNSF have accepted and utilized simulation programs similar to Otter Tail's in previous stand-alone cost cases. In PPL, the BNSF accepted, and the STB utilized operating cost evidence derived in part from a string chart program similar to Otter Tail's string diagram program.^{4/}

In its Reply and Supplemental Reply Evidence, BNSF leveled seven (7) allegations against perceived infirmities of the string program model. These include:

1. The String Model Process is Laborious and Prone to Error;^{5/}
2. Loaded and empty trains are not "linked";^{6/}
3. The model does not detect "T-Bone" collisions;^{7/}

^{4/} See PPL at 12, "BNSF allocated the operating expenses PPL developed for the entire WMCRR between the north-south and western parts using PPL's computer programs."

^{5/} See BNSF Reply Supplemental Reply at III.B-4.

^{6/} See BNSF Reply at III.B-14.

^{7/} See BNSF Reply at III.B-13.

4. The number of trains exceeds the capacity of sidings and yards;^{8/}
5. Some trains spend too long on the system, while others do not appear at all;^{9/}
6. Otter Tail's model period is too short, and is not the OTRR's peak period;^{10/} and
7. The string diagram model does not produce intermediate files to reconcile transit times and service units.^{11/}

Otter Tail demonstrated in its Rebuttal Evidence that the vast majority of BNSF's claims were incorrect, and that those claims that were valid, were easily correctable.

BNSF is now coming back in its Motion making new claims about perceived infirmities in Otter Tail's string model program, or rehashing claims that it made in its Reply and Supplemental Reply testimony, and which Otter Tail refuted on Rebuttal. BNSF makes the following six (6) claims against Otter Tail's string model program: (1) that the string model program has been substantially changed between Otter Tail's Supplemental Opening Evidence and Rebuttal Evidence; (2) BNSF cannot replicate the results of Otter Tail's Rebuttal string model simulation; (3) Otter Tail's Rebuttal string model simulation has trains colliding head-on; (4) Otter Tail's string model program lacks transparency in its code and output; (5) a "massive audit" must be made of each string model run; and (6) there is an imbalance in loaded and empty trains returning to the mines on the peak day leading to an understatement in operating statistics.

^{8/} See BNSF Reply Narrative at III.B-13 to III.B-43.

^{9/} See BNSF Reply at III.B-16.

^{10/} See BNSF Reply at III.B-25 to III.B-26.

^{11/} See BNSF Supplemental Reply at III.C-7.

Each of BNSF's claims is incorrect as is discussed below.

**A. OTTER TAIL HAS NOT SUBSTANTIALLY
CHANGED ITS STRING MODEL IN REBUTTAL**

BNSF asserts that Otter Tail made over one hundred changes to the code of its string model program in Rebuttal, and modified all but one of the program's input files between Otter Tail's Supplemental Opening and Rebuttal filings.^{12/} These large number of changes, according to BNSF, indicate that Otter Tail has presented a new string model in Rebuttal, and that any evidence based on the model must be rejected.

In actuality, the limited number of programming changes Otter Tail made to the string model between Supplemental Opening and Rebuttal filings were in direct response to criticisms made by BNSF in its Reply and Supplemental Reply filings. Moreover, the so-called "modifications" of the input files were nothing more than re-saving the input files for Otter Tail's submission of its Rebuttal Evidence.

In its Reply and Supplemental Reply Evidence, BNSF asserted that there was a multitude of issues with Otter Tail's string model program. Otter Tail's Rebuttal Evidence addressed how the vast majority of BNSF's claims were meritless, and were nothing more than hyperbole and misinterpretations of the string model program, its code and its outputs.^{13/} Of the many claims BNSF made in its Reply and Supplemental Reply Evidence, Otter Tail identified three (3) claims raised by BNSF that were legitimate. First, BNSF pointed out the program was not accumulating mine loading times for all empty trains moving in the non-

^{12/} See BNSF Motion at 11 and Leopold/Monk VS at 1 and 2.

^{13/} See Otter Tail Rebuttal at III-B-20 to III-B-50 and III-C-17 to III-C-36

prevailing direction over the OTRR.^{14/} Second, BNSF indicated that while route grades were being applied to all trains, in some cases trains traveling in the opposite direction over the same section of track were experiencing the same grades.^{15/} Third, BNSF claimed that certain empty coal trains were missing from Otter Tail's Opening string model program.^{16/} Based on these criticisms made by BNSF in its Reply and Supplemental Reply testimony, Otter Tail changed 49 lines of code to address these minor issues.

Any remaining changes to the Otter Tail's rebuttal program were not changes to the programming logic of the program, but rather the removal of extraneous comments and codes, and blank lines from the program to allow room for the changes in response to BNSF's criticisms. The Visual FoxPro programming language that the string model program is written in, is a very stable tested program, but is hampered by capacity limitations. The 49 lines of code changes in Otter Tail's Rebuttal string diagram program that were made in response to BNSF's criticisms caused the program to exceed Visual FoxPro's maximum capacity. To gain the additional capacity required, Otter Tail eliminated extraneous comments and codes from the program to make additional room. The elimination of these lines of code had absolutely no impact on the operation of the program as anyone familiar with the Visual FoxPro language would immediately understand.^{17/}

^{14/} See Otter Tail Rebuttal at III-C-29.

^{15/} See Otter Tail Rebuttal at III-C-20.

^{16/} See Otter Tail Rebuttal at III-B-45.

^{17/} I have enclosed with this Verified Statement the electronic file "changes.xls" which categorizes and explains each change between Otter Tail's Supplemental Opening and Rebuttal string model programs.

In addition to their claim about the changes in the string program code, Leopold/Monk also state that they determined that all but one of the string model input files was “modified” between Otter Tail’s Supplemental Opening and Rebuttal.^{18/} Leopold/Monks base their claim on the electronic file “Folder diffs.htm.” which compares the “Date Modified” field from files in Otter Tail’s Supplemental Opening string model to the “Date Modified” field in files from Otter Tail’s Rebuttal string model. Because, according to Leopold/Monk, the “Date Modified” fields were different between Otter Tail’s Supplemental Opening and Rebuttal, the files were improperly modified.

Leopold/Monks’ claim is a red herring. Anyone who has worked with a computer that utilizes Microsoft Windows knows that the “Date Modified” date only represents the last time a file was saved. It does not indicate whether any information in a file has been modified. In preparing its Rebuttal Evidence, Otter Tail saved its Supplemental Opening string model input files, thus giving them a new “Modified” date. Leopold/Monk’s claim that Otter Tail modified virtually all of its input files is simply not true based on the evidence.

**B. OTTER TAIL CAN REPLICATE THE RESULTS
OF ITS REBUTTAL STRING MODEL PROGRAM**

Leopold/Monks claim that they ran Otter Tail’s Rebuttal string model program using the instructions and run-time parameters provided in Otter Tail’s Rebuttal Evidence, and were unable to replicate the string-model output files included in Otter Tail’s Rebuttal Evidence.^{19/} They conclude from this alleged discrepancy that Otter Tail either failed to provide correct

^{18/} See Leopold/Monk VS at 2.

^{19/} See Leopold/Monk VS at 3

instructions for running the string model program, or failed to provide the actual input codes or programs used to produce Otter Tail's Rebuttal string-model results.

In order to test Leopold/Monks' claim, I ran Otter Tail's Rebuttal string model program using the input files and programs presented to the STB in Otter Tail's April 29, 2004 Rebuttal Evidence. The outputs from this run were identical to outputs included in Otter Tail's Rebuttal Evidence.^{20/}

On June 14, 2004, Counsel for Otter Tail sent a Request for Workpapers to BNSF's Counsel seeking the electronic files BNSF used in its run of Otter Tail's string model program included in the BNSF Motion. BNSF forwarded the electronic workpapers under separate covers on June 30, 2004 and July 2, 2004. To test BNSF's claim, I ran Otter Tail's string model using the electronic workpapers produced by BNSF, and, once again, produced the same outputs as those included in Otter Tail's Rebuttal Evidence.

Given that I was able to produce results consistent with Otter Tail's Rebuttal Evidence using copies of the workpaper presented to the STB as part of Otter Tail's Rebuttal Evidence, and with the electronic files forwarded by BNSF as part of Otter Tail's Request for Workpapers, I am unable to definitively state the cause of BNSF's claim that it cannot reproduce the string model program results. There are several possible explanations though. BNSF's consultants may have erroneously "cross-contaminated" input files between Otter Tail's Supplemental Opening Evidence string model program and Otter Tail's Rebuttal program. Another

^{20/} Otter Tail's string model program, like BNSF's RTC program, is a deterministic and not stochastic model. Therefore, subsequent runs using the same input files and programs will produce the same results.

possibility is that BNSF attempted to modify the program in an effort to analyze it, and confused the outputs between its modified run and Otter Tail's Rebuttal outputs.^{21/} In either case, the simple fact is the results of Otter Tail's Rebuttal string model program are reproducible.

**C. BNSF MISIDENTIFIED TRAINS
WAITING ON PASSING SIDINGS AS COLLISIONS**

BNSF claims that it identified a head-on collision occurring on a single-track section of the OTRR between W. Echeta and E. Kendrick by reviewing the string diagram for this section of the system.^{22/} BNSF then concludes that Otter Tail has not addressed one of the alleged flaws in the program i.e., an inability to detect and resolve collisions. BNSF is wrong and no collision occurred between W. Echeta and E. Kendrick. In fact, the alleged support for BNSF's claim that two trains collided clearly refutes BNSF's claim.

BNSF included in its Motion workpapers the electronic file "String_47.xls," which it claims shows that train CDKMEOR91, an empty coal train moving in the prevailing direction, collided head-on with train GSEAKCM588S, a stack intermodal train moving in the non-prevailing direction. BNSF's workpaper clearly shows that train CDKMEOR91 arrived at W. Echeta at 4:54 pm on October 18. The W. Echeta station is at the end of a

^{21/} BNSF's consultants have a history of modifying and adding lines of code to string model programs submitted by shippers in STB rate cases. BNSF's consultants added several lines of code to Otter Tail's Opening string model program in this case to allegedly gather second-by-second speeds for a train operating over the SARR. See "BNSF's Responses and Objections to Otter Tail's Fifth Request for Workpapers" included as Exhibit No. 1 to this Verified Statement. BNSF's consultants did the same thing in the recent Xcel case when they modified Xcel's string model program. See "Surrebuttal Verified Statement of Karl M. Leopold and Terence J. Monks" filed with BNSF's "Defendant's Motion For Leave to File Limited Surrebuttal Evidence" with the STB on June 6, 2003.

^{22/} See BNSF Motion at 12 and Leopold/Monks VS at 2.

passing section of track between East and West Echeta. Train CDKMEOR91 held at W. Echeta until train GSEAKCM588S arrived at 5:00 pm and passed CDKMEOR91, while CDKMEOR91 waited on the passing siding. After GSEAKCM588S cleared, CDKMEOR91 departed W. Echeta and arrived at E. Kendrick at 5:32 pm.^{23/} Simply stated, no collision occurred between these two trains, as BNSF claims.

BNSF should have also known that train CDKMEOR91 did not collide with GSEAKCM588S by looking at the former's transit time between W. Echeta and E. Kendrick. BNSF's workpaper "String_47.xls" shows that trains moving in an easterly, or Loaded ("L") direction, travel between W. Echeta and E. Kendrick in about 16 to 19 minutes depending on train type and weight. Train CDKMEOR91, an empty coal train, took 37 minutes to move between the two stations, indicating that it waited at W. Echeta for GSEAKCM588S to pass. Simply stated, no collision occurred between these two trains as BNSF claims.

**D. THE STRING MODEL PROGRAM IS
TRANSPARENT IN ITS OPERATION,
AND IN THE DISPLAY OF ITS CODES**

BNSF asserts that Otter Tail's string program model lacks transparency because the operation of trains cannot be observed, and that the operating statistics for the modeled trains are produced through an obscure process that, according to BNSF, has produced erroneous results in the past.^{24/} Contrary to BNSF's claim, Otter Tail's string model program is completely transparent due to the fact that Otter Tail provided every bit of the program's

^{23/} I have included as Exhibit No. 2 a copy of BNSF's electronic file "String_47.xls" showing the arrival and departure times of the two trains, and a schematic of the OTRR track structure at these stations.

^{24/} See BNSF Motion at 12.

code for inspection. Unlike BNSF, which relies upon a compiled computer model to attempt to test the feasibility of only portions of its SARR configuration, Otter Tail provides every line of code, every input file and every output file from its string diagram model so that it can be inspected and critiqued. BNSF is able to dissect and scrutinize every minuscule portion of the string model program's operation, only because Otter Tail has provided the string model program's code to BNSF.

BNSF levies two specific criticisms about the supposed lack of transparency in the string program: 1) that the operations of the string program model cannot be observed, and 2) that the operating statistics are developed in an obscure manner. Both contentions are incorrect. Otter Tail's string diagram model produces detailed output files that record the dispatch and arrival time from SARR origin and destination, and arrival time at each station along each train's route of movement. In addition, graphical representations of each train's movement on the peak day are also produced. Through the use of these outputs, each train's movements across the SARR can be observed.

Moreover, in contrast to BNSF's contention, Otter Tail's string model program is so transparent that BNSF's consultants have been able to modify the program on their own in order to gather additional train operating information. As I indicated above, BNSF's consultants added lines of code to gather second-by-second speeds for trains operating over the OTRR.^{25/} Without access to Otter Tail's transparent string model program code, BNSF's consultants would not have been able to make these changes. In contrast, Otter Tail has

^{25/} See footnote 22, *supra*.

never been provided with information that would allow it to test the code of the RTC model, and to verify the accuracy of the data it generated.

BNSF also contends that Otter Tail's string model program aggregates operating statistics in a complex and unverifiable manner for individual trains. BNSF's claim is without merit. As Otter Tail explained on page III-C-29 of its Rebuttal Evidence, Otter Tail undertook a reconciliation of its transit times, loading hours, and delay times produced by the string model program, and bases its service unit calculations on this reconciliation. The reconciliation combines elapsed transit times, loading times and other applicable dwell times to confirm that all portions of the OTRR cycle times are accounted for on each train dispatched on the peak operating day.^{26/} This manual calculation was performed in direct response to criticisms contained in BNSF's Reply and Supplemental Reply Evidence that the aggregation of operating statistics was not clear.

BNSF's claim that the derivation of Otter Tail's rebuttal aggregated service units is "complex and unverifiable" is absurd considering the service units are calculated in a simple 16-column spread sheet.

**E. OTTER TAIL'S MANUAL INTERVENTIONS ARE
DOCUMENTED, REASONABLE AND VERIFIABLE**

BNSF contends that a "massive audit" must be made of each new string model program run, and that the STB cannot be expected to carry out such a review when a new string model program is submitted on rebuttal.^{27/} Moreover, BNSF asserts that the burden is compounded

^{26/} See Otter Tail Rebuttal electronic workpaper "Service Hours Reconciliation (Base Case).xls".

^{27/} See BNSF Motion at 13.

by the fact that the string program does not automatically identify all train conflicts, and that when conflicts are identified, they must be resolved by manual intervention.^{28/} BNSF's rhetoric regarding the burden of making manual adjustments is simply hyperbole. Moreover, they have nothing to do with whether Otter Tail improperly submitted new evidence in Rebuttal, but rather they simply reargue earlier arguments in an impermissible reply to Otter Tail's rebuttal evidence.

BNSF asserts that a "massive audit" must be made of each new string program run to determine whether problems such as overloaded sidings were overlooked, or whether manual adjustments made to resolve operating conflicts produced new problems. In fact, the output file that contains the string diagram and any manual adjustments is relatively small when compared to other files presented to the STB by both parties in this case. The file "STRING_47_BASE.xls," which contains Otter Tail's string diagram and manual adjustments, is only 2.9 megabytes in size, and consists of only five (5) worksheets of data. In comparison, BNSF's Reply electronic file "LUMs and Carmiles (Otter Tail-BNSF) v2.xls" in which the BNSF develops its operating statistics, is nearly twice as large, at 5.75 megabytes, and contains 12 worksheets of data that must be audited by the STB for completeness.

In addition, Otter Tail's output file "STRING_47_BASE.xls" clearly identifies the position of each train, and all manual interventions taken by Otter Tail. Each train's arrival time at each station is noted in the file, and any adjustments made by Otter Tail are clearly

^{28/} See BNSF Motion at 12.

marked in red. The review of this file is no more burdensome than the review given any other file produced by either party in a rate case.

As Otter Tail explained in its Rebuttal Evidence, manual interventions are performed hundreds of times per day in real world operations by railroad dispatchers to resolve operational conflicts.^{29/} Resolution of these conflicts can take two primary forms.

In one form, additional infrastructure can be added to the system in the guise of additional or lengthened passing sidings and yard tracks, or the addition of a second mainline track. In the second form, trains operating over the system can be slowed or held in yards or sidings to resolve any conflicts. BNSF's approach to designing a stand-alone railroad system implicitly only employs the first form of conflict resolution. In contrast to BNSF's approach, Otter Tail's string diagram process utilizes both forms of operating conflict resolution to develop an efficient, feasible system. The initial design of the stand-alone system was developed by Otter Tail's operating and engineering experts, and then this design was tested in the string-diagram model. Through an iterative process, Otter Tail identified conflicts along the OTRR system, and then, depending upon the number and position of these conflicts, either added or changed infrastructure to resolve the issue, or made manual interventions to fix the conflict. While BNSF's approach ends up building a gold-plated SARR by its ever continuing addition of infrastructure to resolve conflicts, Otter Tail's approach provides for the construction of a feasible, cost efficient stand-alone railroad.

^{29/} See Otter Tail Rebuttal at III-B-34.

F. THE SUPPOSED IMBALANCE OF LOADED AND EMPTY TRAINS IS A FUNCTION OF THE EFFICIENCY OF THE SARR AND NOT A PROBLEM WITH THE STRING DIAGRAM MODEL

BNSF takes issue with Otter Tail's manual reconciliation of operating units claiming that Otter Tail substantially understated its locomotive and railcar requirements. According to BNSF, Otter Tail's peak day is determined by the number of loaded, not empty, trains departing the mines on that day, and that by adding loading time to empty trains returning to the mines, Otter Tail is understating the OTRR's service units because there is an imbalance between loaded and empty trains.^{30/} In actuality, BNSF's assertion about how Otter Tail developed its peak day is wrong, and its argument about the supposed imbalance in trains is just another attempt to impute its real word inefficiencies on the OTRR.

BNSF's claim that Otter Tail determined its peak day based upon the number of loaded coal trains departing mines is false. As Otter Tail explained its Rebuttal Evidence at III-B-48, the OTRR's peak day was determined using the following process. First, Otter Tail identified all the loaded trains, both coal and general freight, originating on each day of the base year 2002 traffic. Next, these base year originations were adjusted to peak year 2021 levels through the application of Otter Tail's traffic forecast.^{31/} After applying the forecast, the day with the peak number of loaded coal and general freight trains entering the system was identified, and used as the basis of Otter Tail's simulation.^{32/} The OTRR's peak day was

^{30/} See BNSF Motion at 14 and Leopold/Monks VS at 4.

^{31/} Otter Tail held the non-coal traffic levels constant for the 20-year DCF period so that base year non-coal traffic is the same as peak year non-coal traffic.

^{32/} The application of this procedure is contained in Otter Tail's Opening electronic workpaper "Peak Coal Trains.xls."

not determined by the number of loaded coal trains departing the mines, but rather by the number of coal and general freight trains entering the stand-alone system. BNSF's claim to the contrary is false.

BNSF's claim about the supposed imbalance of trains departing and arriving at the mines is just an extension of the linked train issue addressed in Otter Tail's Rebuttal Evidence. As Otter Tail explained in Rebuttal, any computer model that analyzes round trip cycles in the SAC environment will result in some problem with the "linking" of the trains from one complete cycle to the next, i.e., balancing the number of loaded and empty trains returning to the origin mines.^{33/} There are only two ways to avoid such a problem. First, the stand-alone railroad could perform exactly like the railroad that it is replacing and ignore any economies and efficiencies resulting from the streamlined facilities and operations of the stand-alone railroad. This is completely contrary to the fundamental underpinnings of the Guidelines, SAC theory, and SAC applications. Alternatively, shippers could adjust to more efficient operations of the SARR by allowing empty trains to load upon their arrival at the mines. This approach avoids the appearance of returning trains waiting for seemingly long times at the mine for their "real world" loading slots to open, and allows the SARR to reflect its truly efficient operations by reducing equipment requirements due to shorter cycle times. The Board, however, has precluded this second (albeit more realistic) option by requiring Complainants to link the SARR operations to the real world.^{34/} The STB should not penalize Complainants for designing the SARR in accordance with this mandate.

^{33/} See Otter Tail Rebuttal at III-B-38.

^{34/} See TMPA at 75, "TMPA has not explained how the GCRR could avoid loading delays while on privately owned mine tracks that it would not control and that would also be used by other carriers."

IV. RESPONSE TO BNSF'S ASSERTIONS REGARDING THE FARGO YARD

Otter Tail describes the configuration of the OTRR at pages III-B-1 to III-B-4 of its Opening Evidence. In general, the OTRR starts at Converse Yard, Wyoming, and follows the BNSF's Orin Line in a northerly direction to Donkey Creek, Wyoming and then turns northwest, following the route of BNSF's mainline through Gillette and Sheridan, Wyoming, to Moran Junction, Montana. At Moran Junction, the line turns northeast and follows BNSF's mainline through Miles City, Montana to Glendive, Montana. At Glendive, the line turns east, and follows BNSF's mainline through Bismark and Jamestown, North Dakota to Fargo, North Dakota. At Fargo the OTRR turns southeast to Brushvale, Minnesota. From Brushvale to E. Breckenridge, Minnesota, the OTRR operates over the Red River Valley & Western Railroad ("RRVW") via trackage rights for 10.52 miles. The OTRR then continues southeast again following BNSF's mainline to Benson, Minnesota, where it turns west and follows the BNSF mainline to Big Stone, South Dakota to the spur that serves Otter Tail's Big Stone generating station where the OTRR terminates. In addition, the OTRR has seven branch lines, one of which is the Snowden Spur that runs from Glendive to Snowden, Montana, and transports predominantly general freight traffic moving between Snowden and Fargo, North Dakota. The OTRR also has 12 yards across its system, with two (2) principal yards at Donkey Creek, Wyoming, and Glendive, MT, and 10 interchange and staging yards

located at Campbell, Converse and Dutch, Wyoming, Moran Junction, Terry and Snowden, Montana, Fargo, Mandan and Casselton, North Dakota and Benson, Minnesota.^{35/}

BNSF's Reply Evidence accepted virtually all of Otter Tail's facility plan for the OTRR, including the location of the OTRR's yards, but claimed that Otter Tail understated the amount of mainline and yard capacity required by the OTRR.^{36/} One of the specific issues was the capacity of the Glendive Yard. BNSF asserts that it currently performs significant intra- and inter-train ("I&I") switching on trains moving between Fargo and Snowden. This traffic, which Otter Tail rerouted to the OTRR, would still require I&I switching, which BNSF proposes be performed at OTRR's Glendive Yard.

On Rebuttal, Otter Tail agreed that a limited number of trains traveling on the OTRR between Fargo and Snowden require I&I switching, but would perform the switching at OTRR's Fargo Yard instead of the Glendive Yard.^{37/} OTRR performs this I&I switching at its Fargo Yard because, as Otter Tail pointed out in its Rebuttal Evidence, it is operationally infeasible to perform the I&I switching at the Glendive Yard proposed by BNSF.

BNSF now claims in its Motion that Otter Tail's addition of capacity at the OTRR's Fargo Yard to handle the I&I switching is an impermissible reconfiguration of the stand-alone railroad, and must be stricken as new evidence submitted on Rebuttal.^{38/} As support

^{35/} See Otter Tail Opening Evidence at III-B-10 to III-B-12.

^{36/} See BNSF Reply Evidence at III-B-1 and Otter Tail Rebuttal Evidence at III-B-1. As Otter Tail discussed in its Rebuttal Evidence, BNSF incorrectly excluded rerouted general freight traffic moving between Fargo and Snowden, and the Snowden Spur from its Supplemental Reply Evidence.

^{37/} See Otter Tail Rebuttal Evidence at III-B-14. BNSF's claim that BNSF performs a "significant" amount of I&I switching on this traffic is an overstatement. As Otter Tail pointed out in its Rebuttal Evidence, on average, less than three (3) trains per day require I&I switching.

^{38/} See BNSF Motion at 16.

for its claim, BNSF cites to a decision in Duke/CSX regarding the STB's decision to strike the shipper's attempt to relocate a yard and tunnel in its rebuttal evidence. BNSF's claim is meritless. Otter Tail did not reconfigure its stand-alone railroad as the shipper in Duke/CSX attempted to do, but rather only added limited capacity at its existing Fargo Yard in direct response to evidence presented by BNSF in Reply. Moreover, BNSF's claims regarding its inability to reply to issues such as demolition costs and expansion of highways are moot because these issues are impermissible barriers to entry as Otter Tail already discussed in its Rebuttal Evidence, and must be disregarded. Finally, as Otter Tail pointed out in its Rebuttal Evidence, I&I switching could not be performed at Glendive, as proposed by the BNSF, because it is operationally infeasible to do so. I explain these issues further below.

**A. CAPACITY ADDITIONS AT THE FARGO YARD
WERE LIMITED IN SCOPE AND DO NOT
CONSTITUTE RECONFIGURATION OF THE SARR**

BNSF's primary argument for its claim that Otter Tail introduced new evidence by adjusting capacity at its Fargo Yard relies on its claim that adding additional track to the existing yard constitutes an impermissible reconfiguration of the SARR similar to the attempted movement of a yard and tunnel in the Duke/CSX case. An examination of the STB's Duke/CSX ^{39/} decision, as well as STB rulings in other cases, shows that BNSF's claim is clearly incorrect.

As the STB spells out in its Duke/CSX decision, the shipper in that case attempted to change the location of the interchange yard at the eastern terminus of the stand-alone railroad

^{39/} See STB Docket No. 42070, Duke Energy Corporation v. CSX Transportation, Inc., decided March 21, 2003.

from Fayette to Gauley, West Virginia, and to relocate a rail tunnel in the vicinity of Hagans, Virginia.^{40/} With regard to the interchange yard, the shipper in the Duke/CSX case stated that it decided to move the interchange yard to Gauley because of the huge grading costs that would be involved in constructing the yard at Fayette, a point at the bottom of the steepest and narrowest part of the New River Gorge. Moving the yard from Fayette to Gauley had a significant impact on the configuration of the SARR in that case, and on the SAC. First, moving the yard shortened the SARR's route by 8.3 miles, and moved to a location that would entail reduced grading costs. Second, relocating the yard altered revenue allocations between the SARR and the incumbent CSX.^{41/} Similarly, the relocation of the Smile/Hagans Tunnel also led to a significant redesign of Duke's SARR, creating one of the longest railroad tunnels in the Eastern United States.^{42/}

Otter Tail's addition of tracks at its existing Fargo Yard in no way resembles the reconfiguration of the SARR in the Duke/CSX case. First, no new yard locations are introduced and the length of the OTRR's route is unaltered.^{43/} Second, unlike Duke/CSX, the addition of capacity at the Fargo Yard has no direct impact on the revenue allocation between the OTRR and the incumbent BNSF. Third, the additional tracks at the Fargo Yard have no impact on the vast majority of OTRR's trains. The only impact is to the limited

^{40/} See Duke/CSX I at 2.

^{41/} See Duke/CSX I at 2.

^{42/} See Duke/CSX I at 4.

^{43/} One of CSX's complaints about the relocation of the Fayette yard was that the new location, Gauley, was an entirely new location, previously unmentioned in Duke's evidence. See Duke/CSX I at 3. In this situation, Otter Tail has from the beginning of the case placed a yard at Fargo, a location BNSF accepted in its Reply and Supplemental Reply Evidence. BNSF's claim that Otter Tail has presented new evidence on yard location is blatantly false.

number of trains that receive I&I switching.^{44/} Overall, there are virtually no similarities between the reconfiguration of the SARR in the Duke/CSX case and the addition of yard tracks to the existing Fargo Yard in this case.

Moreover, the addition of capacity to the OTRR's Fargo Yard is identical to the additional capacity added to SARR yards on Rebuttal in WPL. In its Reply Evidence in WPL, UP stated that the shipper did not provide enough capacity at two existing yards on the SARR system.^{45/} WPL responded to UP's criticisms by adding 4.7 miles of track to the two yards in its Rebuttal Evidence.^{46/} The STB accepted these additions stating that "WPL has explained how these additions would improve operations... and we accept WPL's amended evidence."^{47/} In the instant case, Otter Tail has added nearly the same amount of track to an existing yard as was added to existing yards in WPL. The STB clearly understood in WPL that the addition of tracks to an existing yard does not amount to reconfiguration of the basic SARR. BNSF's claims to the contrary must be summarily dismissed.

**B. BNSF'S CLAIMS ABOUT DEMOLITION
AND RELOCATION COSTS ARE UNFOUNDED**

BNSF claims that, by submitting evidence on Rebuttal about the additional track added at the Fargo Yard, Otter Tail avoids the scrutiny of many cost and feasibility issues,^{48/} such as whether buildings would need to be demolished, and overpasses and bridges would need

^{44/} In contrast, the relocation of the tunnel in Duke/CSX allowed for the avoidance of a need for a switchback track. See Duke/CSX I at 3.

^{45/} See WPL at 65.

^{46/} See WPL at 65.

^{47/} See WPL at 65 (cite omitted).

^{48/} See BNSF Motion at 16, Note 21.

to be expanded. BNSF is clearly overstating its case. BNSF had the opportunity to analyze Otter Tail's construction cost for the Fargo Yard in this case. Furthermore, Otter Tail clearly and distinctly rebutted BNSF's "feasibility issues" in Otter Tail's Rebuttal Evidence.

BNSF already has been afforded the opportunity to analyze and examine Otter Tail's cost for constructing the Fargo Yard. Otter Tail included in its Opening Evidence workpapers that detailed construction specifications assumed by Otter Tail, as well as Otter Tail's unit construction costs. BNSF had two opportunities to address any issues it had with Otter Tail's construction specifications and unit costs, once in its Reply Evidence, and once again in its Supplemental Reply Evidence. This is clearly distinct from the claims made by CSX in its motion to strike in the Duke/CSX case that it expended significant time and resources analyzing and critiquing the engineering and construction costs that would have been required to build and operate the original location, and a similar effort would be needed to respond to the location of the new yard.^{49/} Unlike in Duke/CSX, BNSF has had the opportunity to analyze and critique the basic construction of the Fargo Yard.

In addition, BNSF's claims about "feasibility issues" related to demolishing buildings and expanding overpasses and bridges have already been refuted by Otter Tail as impermissible barriers to entry in Otter Tail's Rebuttal. BNSF included costs in its Reply Evidence for yard demolition and relocating roadways. In its Rebuttal, Otter Tail clearly explained that these cost items are a barrier to entry and, as such, must be excluded from the OTRR's investment base.^{50/} Roads were not prevalent when the lines replicated by the

^{49/} See Duke/CSX I at 3.

^{50/} See Otter Tail Rebuttal Evidence at III-F-95.

OTRR were constructed. Therefore, BNSF, and its predecessors, did not incur costs to demolish, relocate and reconstruct roadways and bridges when rail yards were built. Furthermore, buildings were not demolished to make room for the rail yards built by BNSF and its predecessors. To the contrary, buildings, and entire towns in some cases, were constructed around the railroad and its yards. Consequently, BNSF, and its predecessors, did not incur costs to demolish buildings and foundations. BNSF's claim that it did not have the opportunity to address these issues is moot since they are clearly barriers to entry which are disallowed by the Guidelines.

**C. BNSF'S ADDITION OF TRACK FOR I&I
SWITCHING AT GLENDIVE IS UNFEASIBLE**

BNSF's Reply Evidence assumed that the limited I&I switching currently being performed on a small number of trains traveling between Fargo and Snowden on the BNSF's northern line via Minot would be performed at the OTRR's Glendive Yard.^{51/} In BNSF's Reply Evidence, BNSF added 10.1 miles of track to the Glendive Yard to perform this I&I switching. As Otter Tail discussed in its Rebuttal Evidence, BNSF's addition of the track to the Glendive Yard was operationally infeasible. Therefore, given the infeasibility of BNSF's Reply Evidence, Otter Tail had the opportunity under the STB's evidentiary guidelines to propose an alternative solution to that proposed by BNSF. Otter Tail chose as its alternative to expand the existing OTRR Fargo Yard because it was operationally more efficient.

^{51/} See BNSF Motion at 15.

The STB outlined and reiterated its guidelines for the submission of evidence in its Duke/NS decision.^{52/} The STB stated that on rebuttal a shipper may demonstrate that its opening evidence was feasible, it may adopt the railroad's evidence, or, in certain circumstances, it may offer to refine its evidence to address issues raised by the railroad regarding its evidence. Specifically, the STB opined "Where the railroad has identified flaws in the shipper's evidence, but has not provided evidence that can be used in the Board's SAC analysis, or where the shipper shows that the railroad's reply evidence is itself unsupported, infeasible or unrealistic, the shipper may supply corrective evidence."^{53/} In this case, BNSF's Reply Evidence is both infeasible and unsupported as discussed further below.

Otter Tail demonstrated in its Rebuttal Evidence that BNSF did not support the design of its Glendive Yard. BNSF's Reply Evidence adds approximately 10 miles of yard track and 3 miles of bypass track to Otter Tail's 27.04 miles of track at the Glendive Yard, a 48 percent increase in track miles. Even though BNSF proposed to increase track miles by 48 percent, it would increase the size of the yard by nearly 180 percent, from 114.4 acres to 319.8 acres.^{54/} In analyzing BNSF's Reply Evidence, Otter Tail found that BNSF's acreage proposals are simply text on a diagram, unsupported by any documentation. Otter Tail requested that BNSF provide workpapers underlying the calculation of acres for the Glendive Yard. In response, BNSF provided an electronic file titled "Yard_Acreage.dgn" which can

^{52/} See Duke/NS at 13 to 15.

^{53/} See Duke/NS at 15 (emphasis added).

^{54/} See Otter Tail Rebuttal Evidence at III-F-26.

only be read by an arcane software package which BNSF did not provide. As such, BNSF's calculations of yard acres for Glendive are not supported in this record.

In addition to not supporting its development of the Glendive Yard, BNSF's design of the Glendive Yard is infeasible. As Otter Tail discussed in its Rebuttal Evidence, BNSF's Reply track design for I&I switching at Glendive Yard reflects an additional 10.1 miles of track that would create an operational bottleneck on the lead tracks at each end of the yard.^{55/} Based on BNSF's proposed design and operation of the Glendive Yard, the I&I switching would conflict with every train entering and departing the yard while the switching was being performed.^{56/} The infeasibility of the operations proposed by the BNSF at the Glendive Yard made it incumbent upon Otter Tail to provide an alternative, feasible solution. This is exactly what Otter Tail did by performing the I&I switching at an expanded Fargo Yard, instead of at the Glendive Yard as BNSF proposed in its infeasible plan.

^{55/} See Otter Tail Rebuttal Evidence at III-B-16.

^{56/} See Otter Tail Rebuttal Evidence at III-B-16.

V. RESPONSE TO BNSF'S ASSERTIONS REGARDING RAILROAD PRODUCTIVITY

On Opening, Otter Tail adjusted operating expenses for inflation by using the RCAF-A that DRI-WEFA projected in its June 2001 forecast.^{57/} BNSF's Reply Evidence disagreed with Otter Tail's use of the RCAF-A stating that the RCAF-U was the appropriate index to adjust SARR operating expenses. BNSF provided two reasons as to why it believed that the RCAF-U was the appropriate index. First, because the OTRR is designed to be the most efficient, least-cost railroad, it would be inappropriate to impute to the OTRR improvements in productivity that the nation's Class I railroads are still in the process of achieving.^{58/} Second, because most railroad productivity improvements come from technological improvements that occur over time as assets are replaced, the OTRR could not be expected to gain these improvements since the OTRR is assumed to be built in its entirety over a three-year time period, and most of the investment is in assets with lives longer than the 20-year DCF period.^{59/} On Rebuttal, Otter Tail provided specific evidence as to why BNSF's arguments were incorrect, and that the OTRR would enjoy improvements in productivity over the 20-year DCF period.^{60/}

BNSF now asserts that Otter Tail's Rebuttal Evidence refuting BNSF's Reply claims is new evidence, and must be stricken.^{61/} BNSF also asserts that Otter Tail utilized a modified

^{57/} See Otter Tail Opening Evidence at III-G-3.

^{58/} See BNSF Reply Evidence at III.G-2.

^{59/} See BNSF Reply Evidence at III.G-2.

^{60/} See Otter Tail Rebuttal Evidence at III-G-3.

^{61/} See BNSF Motion at 16.

inflation forecast based on the coal transportation productivity factor contained in the Energy Information Administration's ("EIA") coal transportation rate forecast in its DCF analysis.^{62/}

Both of BNSF's assertions are incorrect as explained below.

**A. OTTER TAIL'S REBUTTAL EVIDENCE
WAS PRESENTED TO DIRECTLY
REFUTE BNSF'S REPLY EVIDENCE**

As mentioned earlier, Otter Tail adjusted future operating expenses for the impact of inflation by using the RCAF-A that DRI-WEFA projected in its June 2001 forecast.^{63/} The reason Otter Tail used the RCAF-A rather than the RCAF-U forecast in its DCF analysis is to reflect the productivity changes the OTRR could be expected to enjoy as changes in technology and equipment are introduced in the future. As the STB has indicated in previous SAC decisions, SARRs can expect to see increases in productivity over the 20-year DCF-period.^{64/}

BNSF's Reply Evidence attempts to refute the idea that a SARR will see any productivity over the 20-year DCF period. BNSF specifically states that, because the OTRR is designed to be the most efficient, least-cost railroad, it would be inappropriate to impute to the OTRR improvements in productivity that the nation's Class I railroads are still in the process of achieving. According to BNSF, these efficiencies are already built into the OTRR's base year operating expenses and thus, it is much more likely that future operating

^{62/} See BNSF Motion at 17.

^{63/} BNSF provided a June, 2003 DRI-WEFA, now renamed Global Insight, forecast of the RCAF-A and RCAF-U in its Reply Evidence. Because it is more current than the forecast used by Otter Tail in its Opening Evidence, Otter Tail adopted the June, 2003 forecast in its Rebuttal Evidence.

^{64/} For example, see Duke/CSX at 30, Duke/NS at 36-37 and CPL at 27-28.

expenses on the OTRR will inflate at the rate of the RCAF-U, not at the rate of the RCAF-A. BNSF also claimed most railroad productivity improvements come from technological advances that occur over time as assets are replaced. BNSF asserts that in the real world, railroad assets of all types are being replaced continuously so improvements in productivity occur across asset types, year by year; but, in the stand-alone analysis, the OTRR is assumed to be built in its entirety over a three-year time period, and most of the investment is in assets with lives longer than the 20-year DCF period. Therefore, under these circumstances, according to BNSF, it makes no sense to impute to the OTRR the productivity anticipated for the nation's railroad industry as a whole.

Otter Tail demonstrated through its Rebuttal Evidence that, contrary to BNSF's assertions, a SARR can achieve productivity through the 20-year DCF period. Otter Tail explained how the OTRR replaces its computer technology every five years, and with this change, new and more powerful, productive computers will be utilized.^{65/} Otter Tail also explained how, as the OTRR's traffic grows over the 20-year study period and new locomotives are added to the fleet, these locomotives will become more efficient and cost effective than current model locomotives.^{66/} In short, Otter Tail provided substantial evidence as to why BNSF's assertions regarding the lack of future productivity on stand-alone railroads are incorrect.

^{65/} See Otter Tail Rebuttal Evidence at III-G-4.

^{66/} See Otter Tail Rebuttal Evidence at III-G-5. BNSF's implicit belief that the OTRR would utilize 2001 vintage equipment to accommodate growth traffic in future years is nonsensical. History has shown that transportation equipment, including automobiles, trucks, locomotives and commercial aircraft, has seen dramatic improvements in performance and cost effectiveness over 20-year spans. To assume that similar gains will not be made in the 20-year DCF period is ridiculous.

The information Otter Tail provided in Rebuttal to support its Opening Evidence on the RCAF-A is completely consistent with the STB's evidentiary guidelines as outlined in the Duke/NS decision. The STB stated that, "On rebuttal, as to those issues challenged by the railroad, the shipper may demonstrate that its opening evidence was feasible and supported...".^{67/} Otter Tail's Rebuttal Evidence in this case clearly fits the STB's guidelines for rebuttal evidence. Otter Tail's Rebuttal Evidence regarding productivity supports Otter Tail's use of the RCAF-A to adjust operating expenses, and directly addresses the issues raised by the BNSF in challenging Otter Tail's Opening position.^{68/}

Furthermore, Otter Tail has not changed its position between its Opening and Rebuttal Evidence regarding the appropriateness of using the RCAF-A to index operating expenses. The evidence Otter Tail presented in Rebuttal regarding future railroad productivity directly supports Otter Tail's Opening Evidence regarding the appropriateness of using the RCAF-A. Otter Tail has not changed its case-in-chief in Rebuttal regarding this fact, and still relies on the use of the RCAF-A in its Rebuttal DCF model.^{69/}

^{67/} See Duke/NS at 14 (citation omitted).

^{68/} See also TMPA at 107, where the STB rejected the shipper's attempt to introduce a new maintenance of way plan on Rebuttal. In rejecting the shipper's evidence, the STB stated "TMPA's rebuttal evidence does not attempt to provide support for its case-in-chief presented in its opening evidence..." In this case, Otter Tail has specifically presented evidence to support its Opening case-in-chief.

^{69/} Only under the BNSF's draconian interpretation of the STB's evidentiary guidelines could Otter Tail's Rebuttal Evidence be considered new. The logical destination of BNSF's warped argument is that any data or argument presented in a shipper's rebuttal evidence in support of its primary case-in-chief, but not already presented in a shipper's opening evidence, must be disallowed.

**B. BNSF INCORRECTLY STATES THAT OTTER TAIL'S
REBUTTAL INFLATION FORECAST IS BASED ON
THE EIA'S FORECASTED CHANGES IN PRODUCTIVITY**

BNSF claims that Otter Tail uses an alternative inflation forecast based on the EIA's forecast of coal transportation productivity instead of relying upon the RCAF-A produced by DRI-WEFA/Global Insight to adjust operating expenses as it did in its Opening Evidence. Specifically, BNSF states, "While Otter Tail purports to present this alternative index as support for the 'appropriateness of using the RCAF-A in this case,' Otter Tail actually uses the supposed EIA productivity factor to develop a modified inflation index, which it then uses in its DCF analysis."^{70/} BNSF's claim is demonstratively incorrect.

As stated above, Otter Tail used the DRI-WEFA/Global Insight forecasts of the RCAF-A to adjust operating expenses in both its Opening and Rebuttal DCF models. Otter Tail made this perfectly clear in its Rebuttal Evidence.^{71/} This fact is confirmed by a review of Otter Tail's Rebuttal DCF Model Table B.^{72/} Contrary to BNSF's claim, Otter Tail used actual RCAF-A and DRI-WEFA/Global Insight forecast data to prepare its Rebuttal DCF analysis.

Otter Tail did include, in its Rebuttal electronic workpapers, an analysis of what the RCAF-A would look like assuming the use of the coal transportation productivity assumed by the EIA. As Otter Tail stated "The appropriateness of using the RCAF-A in this case can be tested by applying the productivity component of the EIA coal transportation rate forecasts."^{73/} This analysis was included in direct support for Otter Tail's case-in-chief use

^{70/} See BNSF Motion at 17 (citation omitted).

^{71/} See Otter Tail Rebuttal Evidence at III-G-10. "On Rebuttal, Otter Tail has used actual RCAF-A index values for time periods that actual data has been published, and it has also adopted the June 2003 DRI-WEFA forecast for the RCAF-A for adjusting future OTRR operating expenses."

^{72/} See Otter Tail Rebuttal electronic file "Exhibit-III-H-3-SP.123".

^{73/} See Otter Tail Rebuttal Evidence at III-G-11.

of the RCAF-A to index future operating expenses, and is perfectly legitimate evidence to present in Rebuttal. It did not replace the DCF model presented in Otter Tail's case-in-chief, which relied upon the DRI-WEFA/Global Insight RCAF-A forecast.

VI. RESPONSE TO BNSF'S ASSERTIONS REGARDING THE SARR'S COST OF CAPITAL

On Opening, Otter Tail calculated the SARR's cost of capital using separate factors for the cost of equity and the cost of debt.^{74/} On Rebuttal, Otter Tail updated its cost of equity and debt determinations to reflect more current cost of capital data that became available after the submission of Otter Tail's Opening Evidence. BNSF now claims, that while it is appropriate to update the cost of equity in the cost of capital determination, Otter Tail offers no justification for updating its cost of debt.^{75/}

BNSF clearly ignores the STB's preference for using actual cost of capital figures when they become available rather than estimates. For example, in McCarty, APS, and FMC, the STB updated cost of capital figures that became available after the submission of all evidence by parties in these proceedings.^{76/} Similarly, the STB has updated other SAC inputs as new and better information came available, including inflation indexes and tonnage forecasts.^{77/} The fact that Otter Tail used more current information that became available after the submission of Otter Tail's opening evidence should not preclude its use in Rebuttal.^{78/}

^{74/} See Otter Tail Opening Evidence at III-G-1.

^{75/} See BNSF Motion at 20.

^{76/} See McCarty at 522, APS at 438 and FMC at 178.

^{77/} See CPL at 16, Duke/NS at 65 and Xcel at 55.

^{78/} BNSF attempts to distort the issue by raising ancillary issues such as the ability to fund the debt or the cost financing fees in order to contend the updating is not possible. BNSF ignores the STB's decision in WTU allowing just such an approach used by Otter Tail in this case. See Otter Tail Rebuttal Evidence at III-G-2.

VII. RESPONSE TO BNSF'S CRITIQUE OF OTTER TAIL'S TRANSIT TIME ANALYSIS

The STB confirmed in its Xcel decision the right for a complaining shipper to reroute traffic that would be local to the stand-alone railroad, so long as the new route would meet shippers' needs.^{79/} In TMPA, the STB provided some insight in to what would constitute meeting shippers' needs by defining the notion as "...[providing] the same or superior level of service as the shipper receives under the contract..."^{80/}

In light of the STB's ruling in TMPA, and in previous STB decisions regarding rerouted traffic, Otter Tail included in its Opening Evidence an analysis demonstrating that there was no reasonably appreciable difference between the transit times for traffic moving between Fargo and Snowden on the OTRR, which is routed through Glendive, and on the BNSF, which is routed through Minot.^{81/} In Reply, BNSF accepted Otter Tail's rerouting of the Fargo-Snowden traffic for all classes of traffic, except for the highest priority intermodal traffic ("Z-Trains").^{82/} In its Supplemental Reply, BNSF reversed itself, and rejected the inclusion of any rerouted Fargo-Snowden traffic moving over the OTRR, claiming that the OTRR cannot provide the same level of service for any of the removed traffic as BNSF currently provides.^{83/}

^{79/} See Xcel at 20.

^{80/} See TMPA at 22.

^{81/} See Otter Tail Opening electronic file "Transit Time Comparison.xls."

^{82/} See BNSF Reply Evidence at III.A-41 to III.A-49.

^{83/} See BNSF Supplemental Reply Evidence at III.A-20.

In Rebuttal, Otter Tail responded to BNSF's Reply and Supplemental Reply arguments. Upon review of transit time levels and service requirements, Otter Tail agreed that Z-Train traffic should be excluded from the rerouted Snowden-Fargo traffic.^{84/} Otter Tail vehemently disagreed, however, with BNSF's Supplemental Reply claims regarding the remaining rerouted Fargo-Snowden traffic, and presented evidence to support inclusion of the remaining traffic in the OTRR traffic group. Specifically, Otter Tail proved that the OTRR provides more consistent service at lower rates and with only slightly higher transit times than the incumbent BNSF on the rerouted traffic. Otter Tail's evidence demonstrated that, because the OTRR's transit times have a smaller variance than BNSF's transit times, the OTRR provided more consistent levels of service.^{85/} Also, by performing a second analysis using the average and the standard deviations of the OTRR's and BNSF's transit times, Otter Tail showed that there is no appreciable difference between the OTRR's average transit times and the BNSF's average transit times.^{86/}

BNSF takes exception only to Otter Tail's second analysis, which shows that the OTRR's and the BNSF's average transit times for the Fargo-Snowden traffic are essentially the same. Through its witness, Dr. Matthew G. Mercurio, BNSF argues that the test Otter Tail used to equate Otter Tail's and BNSF's average transit times is flawed, and that a correct analysis shows that BNSF's average transit times are lower than the OTRR's for all but two traffic types, manifest trains ("M Trains") and priority manifest trains ("P Trains"). Even for

^{84/} See Otter Tail Rebuttal Evidence at III-A-37.

^{85/} See Otter Tail Rebuttal Evidence III-A-46.

^{86/} See Otter Tail Rebuttal Evidence at III-A-46.

those two traffic types, Dr. Mercurio argues that real world and statistical issues obscure the true transit time comparisons and that Otter Tail cannot prove that the OTRR is providing the same or better levels of service than BNSF.

Dr. Mercurio's revised statistical tests are flawed and cannot provide a statistically reliable answer to the question of whether the OTRR's and BNSF's average transit times are essentially the same. In addition, Dr. Mercurio extends his testimony beyond his acknowledged expertise into how real world railroads measure service levels, the workings of Otter Tail's string model program, and the traffic selected to move over the OTRR, which leads him to make clearly erroneous statements about all of these issues. Finally, Dr. Mercurio does not even attempt to refute Otter Tail's evidence that the OTRR provides more consistent and reliable service than that provided by the BNSF.

**A. DR. MERCURIO HAS NOT
CHALLENGED OTTER TAIL'S
SERVICE CONSISTENCY ANALYSIS**

Otter Tail presented two analyses in Rebuttal Evidence to demonstrate that the OTRR would provide at least the same levels of service as the BNSF on the rerouted traffic . The first test used average transit times and one (1) standard deviation of the transit times ("One Standard Deviation Test") to demonstrate that the OTRR provided more consistent levels of service with only slightly longer average transit times.^{87/} The second test, which utilized average transit times and three (3) standard deviations ("Three Standard Deviation Test") showed that the there was no statistical difference between the two railroads' average transit

^{87/} See Otter Tail Rebuttal Evidence at III-A-45.

times.^{88/} Dr. Mercurio only criticizes the Three Standard Deviation Test, and is silent on the statistical aspects of the One Standard Deviation Test.^{89/}

As Otter Tail explained in Rebuttal, BNSF's assumption that a lower average transit time means better service is incorrect.^{90/} Shippers have repeatedly stated that they desire reliability and consistency of service from railroads. Determining whether or not a railroad is providing reliable and consistent service requires more than an examination of average transit times. Average transit times only provide one characteristic of rail service, namely identifying the central mean tendency of the service. It is also necessary to determine the dispersion characteristics of the service. The amount of dispersion, as indicated by the standard deviation, identifies how far from the average a railroad's service can be expected to vary. The lower the standard deviation, the more closely grouped the data points are around the central mean characteristic. When measuring railroad services, the lower the standard deviation, the more reliable and consistent the railroad is in meeting its average transit time.^{91/} As Otter Tail showed in its Rebuttal Table III-A-2, OTRR is providing more consistent service than is the BNSF for the same types of traffic.

^{88/} See Otter Tail Rebuttal Evidence at III-A-48.

^{89/} See Mercurio V.S. at 2 ("In particular, I have been asked to review the statistical analysis contained in Otter Tail's Rebuttal Narrative at II-A-49 through III-A-52.").

^{90/} See Otter Tail Rebuttal Evidence at III-A-45.

^{91/} There is a key difference between the standard deviation of the distribution, which Otter Tail measured in Table III-A-2 of its Rebuttal Evidence, and the standard error of the mean that Dr. Mercurio calculated in his Table 1. The standard deviation of the distribution describes the width of the distribution or the dispersion of transit times, while the standard error of the mean provides insight to the precision of the mean estimate.

Instead of attempting to statistically disprove the One Standard Deviation Test, Dr. Mercurio ventures beyond his acknowledged expertise, in an attempt to demonstrate that the distribution of the OTRR's transit times is artificially "tight" relative to the distribution of BNSF transit times. Specifically, he makes erroneous statements about differences between BNSF and OTRR operations and Otter Tail's string model to claim that the string diagram model does not reflect real world items such as maintenance and other track outages, and intentionally excludes routine tasks such as fueling, switching and crew changes.^{92/} Dr. Mercurio's argument is incorrect.

Instead of attempting to impute artificial random failures as BNSF suggests, Otter Tail accounted for unscheduled track outages on the OTRR through its proper design and construction of the railroad structure and its operating plan, which is more than adequate to address unexpected occurrences. Therefore, as Otter Tail explained in its Opening Narrative, its experts developed the OTRR's capacity utilizing an 18-hour average operating day, which leaves 6 hours each day to allow for variations from expected normal operations.^{93/} Moreover, by utilizing actual mine dispatch and interchange arrival times in its string diagram model, Otter Tail already has included the impact of any unforeseen occurrences that impact the timing of train operations over the OTRR system, and reflected these occurrences in its operating statistics and operating costs. Finally, because the OTRR is a state of the art

^{92/} Mercurio VS at 4.

^{93/} See Otter Tail Opening at III-C-18. By designing the SARR around an 18-hour workday, Otter Tail has built into its system the capacity to handle planned maintenance and unexpected delays. In other words, the OTRR is designed to handle its traffic only using three-quarters of a day, leaving ample time available to accommodate delays and maintenance.

railroad, unlike the BNSF's rail lines, which have been in place for many years, the OTRR is constructed with new equipment and materials and should not face the number of unplanned outages that BNSF currently faces.^{94/} As Otter Tail explained in its Rebuttal Evidence, non-scheduled disruptions, by their very nature, cannot be identified and BNSF provided no evidence that these events would actually occur.^{95/}

Dr. Mercurio's assertion that Otter Tail intentionally excluded time spent on activities such as fueling, switching and crew changes also is false. Moreover, it is an assertion beyond his expertise to make. As Otter Tail's Rebuttal Evidence and workpapers show, Otter Tail included time in its string model simulation for changing crews, inspecting trains and fueling locomotives, interchanging trains to and from the BNSF, and coupling and uncoupling helper locomotives.^{96/} Dr. Mercurio has no basis to claim that Otter Tail intentionally excluded these operating delay times from its transit time calculations.^{97/}

^{94/} BNSF's position regarding maintenance outages and its position regarding railroad productivity are completely contradictory. On one hand, BNSF argues that the OTRR will not experience any productivity because it is constructing its railroad with new materials that will not be replaced during the 20-year DCF period. On the other, BNSF contends that OTRR has not allowed for maintenance windows to replace equipment as it fails.

^{95/} Otter Tail Rebuttal Evidence at III-C-35.

^{96/} See Otter Tail Rebuttal Evidence at III-C-29 for an explanation of the OTRR's delay times, and OTRR Rebuttal electronic file "ROUTE.DBF" showing the delay time inputs into the string diagram model.

^{97/} If these issues were true, one would expect they would have been addressed by BNSF's programming experts Leopold/Monks. Since they were made by someone who does not claim to have any programming experience or to have even reviewed the string diagram model, they can be summarily dismissed.

**B. DR. MERCURIO'S TRANSIT TIME
ANALYSIS DOES NOT ACCURATELY REFLECT
THE DATA OR ACTUAL RAILROAD OPERATIONS**

Dr. Mercurio begins his assault on Otter Tail's three standard deviation test by asserting that the analysis was likely biased in favor of Otter Tail because Otter Tail compared mean transit times, which, according to Dr. Mercurio, can be distorted by a relatively "small number of outliers."^{98/} Instead, Dr. Mercurio asserts that Otter Tail should have compared the medians of the OTRR's and BNSF's transit times.

Dr. Mercurio's rhetoric reveals that he has little experience analyzing railroad operations, and reflects little understanding of shipper needs and requirements. Dr. Mercurio has also offered no proof that statistical outliers influenced the results of Otter Tail's comparison of OTRR's and BNSF's mean transit times. In fact, a review of BNSF's average transit times when statistical outliers are removed shows that the OTRR still provides more consistent and reliable service than the BNSF, with only small increases in the difference of average transit times.

Dr. Mercurio's assertion that the correct way to compare OTRR and BNSF transit times is by comparing the median, and not the mean, of the times contradicts actual railroad operating practices and shipper requirements. In railroad operating statistics averages, not medians, are used to report the central tendency of operations. In addition, coal transportation contracts that include stipulated service requirements, in almost all cases

^{98/} See Mercurio VS at 3.

measure average transit or cycle times.^{99/} Similarly, when railroads report their own performance measures, they do not report median system train speed or median terminal dwell time, but rather they report average train speed and average terminal dwell time.^{100/}

To test Dr. Mercurio's claim that outliers distorted the presentation of BNSF's average transit times, I calculated the mean and standard deviation of BNSF's transit times between Fargo and Snowden with the outliers removed from the distribution.^{101/} Table 1 below compares the mean and standard deviation of Otter Tail's peak week transit times to the mean and standard deviation of BNSF's transit times with outliers removed.

^{99/} See, for example, Otter Tail Opening Workpapers at Vol. III, pp. 1823, 1828, 1832, 1841, 1852 and 1999.

^{100/} See for example the Association of American Railroads' ("AAR") Railroad Performance website at http://www.railroadpm.org/welcome_rpm.cfm which displays the AAR's determination of specific railroad's average train speeds and terminal dwell times.

^{101/} An outlier is defined as a subset of observations from a larger data set in which they are so far separated in value from the remainder of the observations that they give rise to the question of whether they came from the same process or not. See "Statistical Quality Control for Manufacturing Managers," by William S. Messina, John Wiley & Sons, 1987, page 4. There is no universal agreement as to what represents "so far separated in value from the remainder of the observations" to be considered an outlier, but practitioners will customarily classify outliers as those observations that are two (2) to three (3) standard deviations from the mean.

Table 1
**Comparison of BNSF and OTRR Average Transit Times (Hours)
 and Standard Deviations Between Fargo, ND and Snowden, MT**

Item (1)	<u>Average</u> (2)	<u>Standard Deviation</u> (3)
1. OTRR Transit Times ^{1/}	14.4	1.9
2. BNSF Transit Times With All Data ^{1/}	13.8	5.9
3. BNSF Transit Times With Outliers Beyond Three Standard Deviations Removed ^{2/}	13.3	4.5
4. BNSF Transit Times With Outliers Beyond Two Standard Deviations Removed ^{2/}	13.1	4.1

^{1/} **Source:** Otter Tail Rebuttal Evidence Table III-A-2.

^{2/} **Source:** Crowley VS electronic file "Confidence_Intervals_Revised.xls."

As Table 1 shows, eliminating the most extreme outliers from the BNSF transit time data set, i.e., those times that are beyond three standard deviations from the mean, lowers BNSF's average transit time by approximately one half of an hour; but the standard deviation of the average transit times are still about 2.5 times greater than that of Otter Tail. Even using a more liberal definition of outliers and eliminating transit times that are over two standard deviations from the mean shows that Otter Tail is still providing much more consistent and reliable service than is the BNSF. Table 1 clearly shows that BNSF's transit times have a much greater variance than those of the OTRR's.

Dr. Mercurio claims that Otter Tail should have known even through casual observation that there was no possible way the distributions as shown in his Figure 1 had the same

mean.^{102/} However, simple observation of histograms, which display data frequencies, provides no insight into the means of the data sets underlying the distributions. For example, consider two data sets which could represent railroad transit times. Data Set No. 1 consists of transit times equal to 1, 2, 3, 4 and 5 hours, and Data Set No. 2 consists of transit times equal to three observation of 3-hours each. Obviously the distributions of the two data sets are completely different, yet the mean value of the data sets are equal.

**C. DR. MERCURIO'S HYPOTHETICAL EXAMPLE
IS MISLEADING AND FACTUALLY INCORRECT**

In an attempt to disprove Otter Tail's Three Standard Deviation Test, Dr. Mercurio creates a self serving hypothetical example that he claims proves the logic behind Otter Tail's analysis is incorrect. First, he creates a straw man by postulating two hypothetical distributions that most likely would never represent the distributions of railroad transit times. Then he topples his straw man by stating that the distributions are obviously not identical. From this, he contends that Otter Tail's transit time analysis incorrectly concludes that the distributions of the OTRR and BNSF transit times are the same.^{103/} However, Otter Tail never concluded that the distribution of the two railroads' transit times were identical, only that there was a good chance the means of the transit times were comparable. Moreover, Dr. Mercurio's hypothetical example is so unrealistic and far-a-field from what one would see when observing the distribution of transit times of actual railroad operations as to be meaningless.

^{102/} See Mercurio VS at 4.

^{103/} See Mercurio VS at 5.

Further, Dr. Mercurio's hypothetical example of the flaws inherent in Otter Tail's conclusion, based on overlapping distributions, reveals his lack of insight into railroad operations. Dr. Mercurio uses two approximately normal distributions in his example. Clearly, as shown in Dr. Mercurio's Figure 1, long-run transit time data does not follow a normal distribution, but are skewed to the left. This is due to the existence of absolute minimum transit times. Because transit times cannot be negative, the distributions will tend to have "tails" on the right side of the distribution, but less so to the left. Also, because transit time and consistency of service are a customer service issue even for those commodities that are moving on the lowest priority trains, i.e. manifest trains, transit times will tend to skew towards the minimum of the distribution in the long run.^{104/}

D. DR. MERCURIO'S PERFORMANCE OF A TWO-SAMPLE T-TEST IS INAPPROPRIATE

Dr. Mercurio asserts that Otter Tail's Three Standard Deviation Test is invalid, and that Otter Tail should have instead employed a two-sample t-Test of mean equality to determine whether the OTRR's transit times are, on average, longer than those of the BNSF.^{105/} The two-sample t-Test, according to Dr. Mercurio, permits one to mathematically test the hypothesis that the means of the two distributions are the same.

In actuality, the two-sample t-Test carried out by Dr. Mercurio is inappropriate for this particular situation. Dr. Mercurio has set-up his two sample t-Test to address what has been

^{104/} In the short run, the skewness will not be apparent. This is true even more so with SARR's that have smaller traffic groups than real world railroads, are built new so there are less delays, and have more control over the operations of the railroad.

^{105/} See Mercurio VS at 11.

known as a Behrens-Fisher Problem. A Behrens-Fisher Problem is that of comparing the means of two populations when the ratio of the two populations is unknown. Dr. Mercurio's justification for using the two-sample t-Test is based on arguments taken from a paper that was written by Nathaniel Schenker and Jane F. Gentleman ("Schenker/Gentlemen").^{106/} Schenker/Gentleman's paper relies upon a mathematical statistics text by Peter J. Bickel and Kjell A. Doksum who reference the definitive work on Behrens-Fisher Problems by Dr. Henry Scheffe.^{107/} Dr. Scheffe explains that formulating a problem as a Behrens-Fisher Problem is appropriate when one is only concerned about testing the means of the populations without concern for the variances of the populations. As Dr Scheffe explains:

Formulation of a real problem of comparing two means as the Behrens-Fisher problem seems appropriate when the population variances are unknown, and, more importantly, we do not much care whether or not they are equal. An example where it seems appropriate is the comparison of the yields of two agricultural varieties; an example where it seems generally not appropriate is the comparison of the effect of a new drug with a control drug. In the first example one is interested in comparing the average yields over an area which is large relative to that of an experimental plot, but in the second example one is interested in comparing the effects of the drugs (including their variation) on individuals.^{108/}

The significance of this paragraph in the present case is the importance of the variance of the actual delivery time to the individual movements, and not simply the average time for all movements. The average difference in the estimated mean delivery time is 36 minutes,

^{106/} See Mercurio VA at 11 and Mercurio VS Attachment B.

^{107/} See Mercurio VS, Attachment B at 5. Excerpts from Drs. Bickel and Doksum test and Dr. Scheffe are included in Exhibit No. 3 to this VS.

^{108/} See Scheffé at 1500 (emphasis original).

or less than 5 percent, in favor of the current BNSF route; however, Otter Tail's proposed operations on the OTRR decreases the standard deviation of the individual movements from 5.9 hours (on the BNSF) to 1.8 hours (on the OTRR). When just in time delivery is important for interconnects and other scheduling reasons, the financial gain by the reduction in variation of individual delivery times (as measured by the estimated standard deviation) far outweighs the value of a minor increase in the overall average delivery time for all movements.

**E. OTTER TAIL'S TRANSIT TIMES
FOR "M" TRAINS AND "P" TRAINS
ARE COMPLETE AND COMPARABLE**

Notwithstanding the inappropriateness of using a two-sample T-test to compare the means of OTRR and BNSF transit times, Dr. Mercurio dismisses, out of hand, the two instances where his analysis shows that OTRR transit times are less than those of BNSF. As shown in Dr. Mercurio's Table 2, the OTRR's average transit times for "M" Trains and "P" Trains are significantly shorter than those for BNSF for the same train type. Dr. Mercurio parrots the claim made by BNSF, in its Supplemental Reply, that there is no basis for comparison between OTRR transit times and BNSF transit times for "M" Trains because the BNSF's transit times include a substantial amount of switching not included by Otter Tail.^{109/} Dr. Mercurio also dismisses shorter OTRR transit times for "P" Trains because Otter Tail allegedly did not include I&I switching in its calculation of transit times.^{110/} As Otter Tail

^{109/} See Mercurio VS at 12.

^{110/} Dr. Mercurio also claims that the sample size of the "P" trains is extremely small, and therefore would be inappropriate to draw strong conclusions regarding the apparent difference between "P" train transit times. See Mercurio VS at 12.

explained in its Rebuttal Evidence, and displayed in its Rebuttal workpapers, both of Dr. Mercurio's assertions are false.

Otter Tail explained in its Rebuttal Evidence that it strived to select only that traffic that moved in train-load service over the OTRR route, including "M" and "P" trains.^{111/} In its Reply Evidence, BNSF pointed out that some westbound "P" Trains receive a limited amount of I&I switching in Minot, but made no claims about switching on "M" Trains. BNSF's acceptance of these trains in its Reply traffic group indicates BNSF's implicit acceptance that these trains are moving over the OTRR system in trainload service. BNSF even produced, in its Reply Evidence, an analysis of transit times that includes those "M" and "P" trains that move between Fargo and Snowden intact.^{112/} If BNSF had proof that the "M" Trains Otter Tail included in the OTRR traffic base received any switching between Fargo and Snowden, it most assuredly would have raised the issue, and sought to have this traffic removed from the OTRR traffic group.^{113/} The fact that it did not, effectively demonstrates that this traffic moves intact between Fargo and Snowden.

Dr. Mercurio also claims that the transit time comparisons for "P" Trains are considered dubious because "P trains experience switching, which Otter Tail may not have accounted for correctly."^{114/} The switching that Dr. Mercurio alludes to is the limited I&I switching

^{111/} See Otter Tail Rebuttal Evidence at III-A-44.

^{112/} See BNSF Reply electronic file "Revised Transit Times Comparison.xls."

^{113/} For example, BNSF did point out in Reply that Otter Tail included a limited number of railcars (less than 1,000 in all of 2001) that originated or terminated on the BNSF's northern line between Fargo and Snowden. Otter Tail excluded these railcars from its Rebuttal Analyses. See Otter Tail Rebuttal Evidence at III-A-52.

^{114/} See Mercurio VS at 12.

BNSF performs on trains moving between Fargo and Snowden. As discussed above, Otter Tail explained in its Rebuttal Evidence that it agrees a limited amount of traffic received I&I switching, and accounts for this in its Rebuttal analyses.^{115/} Dr. Mercurio's claim is meritless.

^{115/} See Otter Tail Rebuttal at III-B-15.

VERIFICATION

COMMONWEALTH OF VIRGINIA)
)
CITY OF ALEXANDRIA)

I, THOMAS D. CROWLEY, verify under penalty of perjury that I have read the foregoing Verified Statement of Thomas D. Crowley, that I know the contents thereof, and that the same are true and correct. Further, I certify that I am qualified and authorized to file this statement.


Thomas D. Crowley

Sworn to and subscribed
before me this 19th day of July, 2004

Anthony V. Evanshaw III
Anthony Vincent Evanshaw III
Notary Public for the State of Virginia

My Commission expires: Sept. 30, 2007.
September 30, 2007

EXHIBIT 1

BEFORE THE
SURFACE TRANSPORTATION BOARD

OTTER TAIL POWER COMPANY,)

Complainant,)

v.)

THE BURLINGTON NORTHERN AND)
SANTA FE RAILWAY COMPANY,)

Defendant.)

Docket No. 42071

**BNSF'S RESPONSES AND OBJECTIONS TO
OTTER TAIL'S FIFTH REQUEST FOR WORKPAPERS**

Defendant The Burlington Northern and Santa Fe Railway Company hereby objects and responds to Otter Tail Power Company's Fifth Request For Workpapers.

BNSF objects to Otter Tail's requests to the extent they seek materials that were neither prepared by nor relied upon by BNSF witnesses in the development of BNSF's reply evidence in this proceeding. BNSF specifically objects to the following requests on grounds that, while generally styled as workpaper requests, these requests are in fact unauthorized discovery requests seeking additional information relating to BNSF's reply evidence or an elaboration on the substance of BNSF's reply evidence: Request Nos. 4; 6; 9; 10; 13; 14; 16; 17; 18; 19; 20; 22; 23; 24; 25; 26; 31; 34; 35; 38; 39; 40; 44; 45; 52; 63; 64; 65; 66; 67; 68; and 69. In addition, many of these requests are improper because the statements for which Otter Tail seeks workpaper support are obviously based on the witnesses' experience and knowledge of the relevant subject matter. BNSF further states that the workpapers used by BNSF to produce the evidence covered by these requests have already been submitted.

30. Please provide all workpapers supporting the derivation of, and the equations that use, the air drag coefficients, flange coefficients, journal constants, axle coefficients, cross-sectional areas, load/tare weight ratios, maximum air brake pressures and full service application brake pressures shown on pages 3 to 7 of BNSF Reply Exhibit III.B-8, and used by BNSF's witness Mr. Wheeler in his RTC simulations of the Gillette, WY to Converse, WY, Snowden, MT to Fargo, ND and Fargo, ND to Big Stone City, SD sections of the OTRR route.

BNSF Response: See BNSF Response No. 28 above.

Section III-C – Operating Plan

32. Please provide an uncompiled copy of the program referenced on page III.C-27 of BNSF's Reply Evidence, and used by BNSF's witness Mr. Monks to capture output from Otter Tail's string model to generate the graphs contained in BNSF Reply Exhibit III.C-5.

BNSF Response: Mr. Monks added several lines of code to Otter Tail's Errata electronic workpaper "otter_tail_string.prg" to capture the output of the model. Otter Tail's string program, with Mr. Monks' additions, is provided in the enclosed workpapers as "speedrun.prg."

33. At page III.C-28, BNSF states that Mr. Mueller observed activities at BNSF's Alliance Yard in January 2003. These observations are, in part, the basis for his estimates of time required for time assessments for trains in yards on the OTRR. Please provide copies of all of Mr. Mueller's notes and data or information collected related to his observations at Alliance Yard in January 2003.

BNSF Response: See BN.RP.WP.III.C-0009-0012.

35. Please provide all workpapers underlying BNSF's witness Mueller's assessment of the time that loaded coal trains would spend at the Glendive yard as outlined on pages III.C-37 to III.C-39 of BNSF's Reply Evidence.

BNSF Response: See Response No. 33 above.

36. At page III.C-40 BNSF states that Mr. Mueller observed inter/intra train switching activities at BNSF's Minot Yard in July 2003. These observations are, in part, the basis for his estimates of time required for these switching activities on the OTRR. Please provide copies of all of Mr. Mueller's notes and data or information collected related to his observations at Minot Yard in July 2003.

BNSF Response: A copy of Mr. Mueller's notes related to his observations at Minot Yard is attached hereto as BNSF.RP.SUPP.WP.III.C-01 to 04. Other information related to Mr. Mueller's conclusions with respect to intra-/inter-train switching at Minot Yard have been provided with BNSF's Reply Evidence.

PUBLIC VERSION

BEFORE THE
SURFACE TRANSPORTATION BOARD

PUBLIC SERVICE COMPANY OF
COLORADO D/B/A XCEL ENERGY

Complainant,

v.

THE BURLINGTON NORTHERN AND
SANTA FE RAILWAY COMPANY

Defendant.

Docket No. 42057

**SURREBUTTAL VERIFIED STATEMENT OF
KARL M. LEOPOLD AND TERENCE J. MONKS**

We are Karl M. Leopold and Terence J. Monks. We have been retained by The Burlington Northern and Santa Fe Railway Company (BNSF) in this proceeding to assist in the evaluation of Xcel's evidence. In this Surrebuttal Verified Statement, we address the continuing presence of fundamental flaws in the string model analysis Xcel presented in its May 19, 2003, Rebuttal Evidence.

In support of a Motion for Summary Disposition and Stay filed by BNSF on February 6, 2003, and in BNSF's April 4, 2003, Reply Evidence, we demonstrated that the first string program analysis submitted by Xcel in this proceeding contained flaws that rendered its output unreliable. On rebuttal, Xcel submitted a new string program analysis that supposedly fixed errors contained in the first presentation. As we describe below, Xcel's new analysis did not fix the errors in its original evidence and, indeed, introduced new errors that confirm the unreliability of the string program for assessing capacity requirements and for generating operating costs upon which costs are based.

PUBLIC VERSION

On the present record, neither the Board nor BNSF has a basis to determine whether the new grade and curve data are any more reliable than the old since neither set of data corresponds precisely to the electronic data or track profiles produced by BNSF in discovery. On both opening and rebuttal Xcel adjusted or manipulated BNSF's discovery data in some manner to produce its "grade.dbf" files. Whatever process Xcel employed is completely undocumented. In any case, as described below, Xcel's attempt to introduce new and different data is undermined by the failure of the string program properly to use that grade and curve data (or in fact *any* grade and curve data) for the overwhelming majority of train movements it models.

B. XCEL'S NEW STRING ANALYSIS DOES NOT PROPERLY USE THE NEW GRADE AND CURVE DATA IN ANY CASE

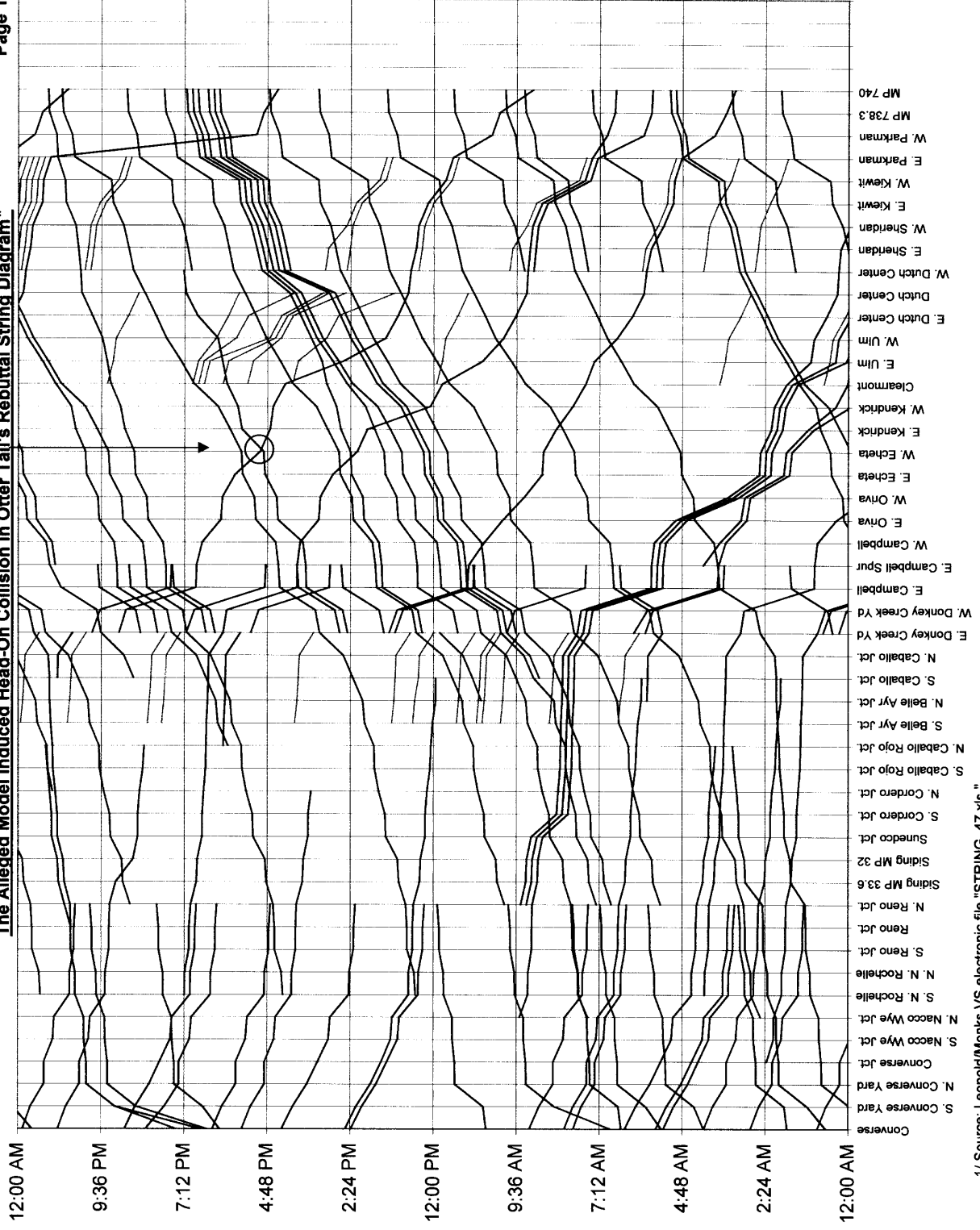
To determine the extent to which the new grade and curve data were responsible for Xcel's reduced transit times, we ran one of Xcel's trains through the string analysis to see what grade and curve values were being used by the program to calculate the train's movement.² The string program itself produces no intermediate grade or curve output showing what variables from the "grades.dbf" file were used during its process. For this reason, we developed additional code to show some of the internal calculations during the acceleration portion of the model, when the values for grade and curve are called upon by the program.³ Without adding this code it would have been impossible to track the activities associated with grade and curve assignments. Indeed, this code serves as a verification of a portion of the process relating to the development of acceleration that is being executed by the model.

² The specific program for Xcel's string analysis is contained in Xcel electronic work paper "pawnee_string.prg."

³ The modified string program is included with this Statement as BNSF electronic work paper "speeder.prg."

EXHIBIT 2

The Alleged Model Induced Head-On Collision in Otter Tail's Rebuttal String Diagram^{1/}



^{1/} Source: Leopold/Monks VS electronic file "STRING_47.xls."

REDACTED

Exhibit No. 2

Page 2 of 2

EXHIBIT 3

MATHEMATICAL STATISTICS

Basic Ideas and Selected Topics

PETER J. BICKEL

*University of California,
Berkeley*

KJELL A. DOKSUM

*University of California,
Berkeley*



PRENTICE HALL, Englewood Cliffs, New Jersey 07632

As an illustration, consider the following experiment designed to study the permeability (tendency to leak water) of sheets of building material produced by two different machines. From past experience, it is known that the log of permeability is approximately normally distributed and that the variability from machine to machine is the same. The results in terms of logarithms were (from Hald [11], p. 472)

x (machine 1)	1.845	1.790	2.042
y (machine 2)	1.583	1.627	1.282

We test the hypothesis H of no difference in expected log permeability. H is rejected if $|T| \geq t_{n-2}(1 - \frac{1}{2}\alpha)$. Here $\bar{y} - \bar{x} = -0.395$, $s^2 = 0.0264$ and $T = -2.977$. Since $t_4(0.975) = 2.776$, we conclude at the 5% level of significance that there is a significant difference between the expected log permeability for the two machines. The level 0.95 confidence interval for the difference in mean log permeability is

$$-0.395 \pm 0.368.$$

On the basis of the results of this experiment, we would select machine 2 as producing the smaller permeability and thus the more waterproof material. Again we can show that the selection procedure based on the level $(1 - \alpha)$ confidence interval has probability at most $\frac{1}{2}\alpha$ of making the wrong selection.

6.4.C. The Two-Sample Problem with Unequal Variances

In two-sample problems of the kind mentioned in the introduction to Section 6.4.B it may happen that the X 's and Y 's have different variances. For instance, a treatment that increases mean response may increase the variance of the responses. If normality holds, we are led to a model where $X_1, \dots, X_{n_1}$; $Y_1, \dots, Y_{n_2}$ are two independent $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ samples, respectively. As a first step we may still want to compare mean responses for the X and Y populations. This is the *Behrens-Fisher Problem*. Instead of deriving a L.R. test, we will give a natural solution keeping in mind the L.R. test we derived when $\sigma_1^2 = \sigma_2^2$.

Note that a complete, sufficient statistic for $\theta = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$ is $(\bar{X}, \bar{Y}, s_1^2, s_2^2)$, where

$$s_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_i - \bar{X})^2, \quad s_2^2 = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} (Y_j - \bar{Y})^2.$$

The M.L.E. (and U.M.V.U.E.) of $\Delta = \mu_2 - \mu_1$ is $D = \bar{Y} - \bar{X}$. The variance of D is

$$\sigma_D^2 = \text{Var}(\bar{X}) + \text{Var}(\bar{Y}) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

Thus D/σ_D has a $\mathcal{N}(\Delta, 1)$ distribution. Since σ_D^2 is unknown, it must be estimated. Its U.M.V.U. estimate is

$$s_D^2 = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}.$$

It is natural to try and use D/s_D as a test statistic for the one-sided hypothesis $H: \mu_2 \leq \mu_1$, $|D|/s_D$ as a test statistic for $H: \mu_1 = \mu_2$, and more generally $(D - \Delta)/s_D$ to generate confidence procedures. Unfortunately the distribution of $(D - \Delta)/s_D$ depends on σ_1^2/σ_2^2 for fixed n_1, n_2 . For large n_1, n_2 , by Slutsky's theorem and the central limit theorem, $(D - \Delta)/s_D$ has approximately a standard normal distribution. For small and moderate n_1, n_2 an approximation to the distribution of $(D - \Delta)/s_D$ due to Welch [20] works well.

Let $c = s_1^2/n_1 s_D^2$. Then Welch's approximation is $\mathcal{F}_k$ where

$$k = \left[\frac{c^2}{n_1 - 1} + \frac{(1 - c)^2}{n_2 - 1} \right]^{-1}.$$

When k is not an integer, the critical value is obtained by linear interpolation in the t -tables. The tests and confidence intervals resulting from this approximation are called *Welch's solutions* to the Behrens-Fisher problem. Wang [19] has shown the approximation to be very good for $\alpha = 0.05$ and $\alpha = 0.1$, the maximum error in size being bounded by 0.003.

Note that Welch's solution works whether the variances are equal or not. The L.R. procedure derived in 6.4.B which works well if the variances are equal can unfortunately be very misleading if they are not. A discussion of these and other issues connected with the Behrens-Fisher problem may be found in Scheffé [18]. See also Problem 6.4.9.

6.5. LIKELIHOOD RATIO PROCEDURES FOR BIVARIATE NORMAL DISTRIBUTIONS

If n subjects (persons, machines, fields, mice, etc.) are sampled from a population and two numerical characteristics are measured on each subject, then we end up with a bivariate random sample $(X_1, Y_1), \dots, (X_n, Y_n)$. Approximation arguments of the type given in Section 6.4.B and empirical data sometimes suggest that a reasonable model is one in which the two characteristics (X, Y) have a joint bivariate normal distribution, $\mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, with $\sigma_1^2 > 0, \sigma_2^2 > 0$.

6.5.A. Testing Independence, Confidence Intervals for ρ

The question "Are two random variables X and Y independent?" arises in many statistical studies. Some familiar examples are: X = weight, Y = blood pres-

Practical Solutions of the Behrens-Fisher Problem

HENRY SCHEFFÉ*

The Behrens-Fisher problem is that of comparing the means of two populations when the ratio of their variances is unknown. Solutions are discussed in terms of interval estimation and testing. Of the six solutions compared, three are preferred by the writer and judged practical, and for these, approximations are given for the power of the tests.

1. INTRODUCTION

The most frequently occurring problem in applied statistics is, in my opinion, the comparison of the means of two populations, say ξ_1 and ξ_2 . Let σ_1^2 and σ_2^2 denote the population variances. It is called the Behrens-Fisher problem if the ratio $\theta = \sigma_1^2/\sigma_2^2$ is unknown, and the assumption is added that the populations are normal. The normality assumption is of no practical importance for any of the solutions considered in this article, all of which, being based mainly on the difference of the sample means, are robust against its violation. I shall treat the problem first in terms of interval estimates for the difference $\delta = \xi_1 - \xi_2$ of the population means; the associated tests are discussed in Section 7. For each of the solutions considered, the probability that the interval covers δ is a function of θ only, and I shall denote it by $1 - \beta(\theta)$. The following further notation is used throughout: For the sample from the i th population ($i = 1, 2$), n_i denotes the sample size, $\bar{x}_i$ the sample mean, and s_i^2 the sample variance with denominator $f_i = n_i - 1$ d.f. (degrees of freedom). Let $d = \bar{x}_1 - \bar{x}_2$. Then d, s_1^2, s_2^2 are independently distributed, s_i^2 as $\sigma_i^2 \chi^2(f_i)/f_i$ where $\chi^2(f)$ denotes a chi-square variable with f d.f., d is $N(\delta, \sigma_d^2)$,

$$\sigma_d^2 = \lambda_1 \sigma_1^2 + \lambda_2 \sigma_2^2, \quad (1.1)$$

and $\lambda_i = 1/n_i$.

Formulation of a real problem of comparing two means as the Behrens-Fisher problem seems appropriate when the population variances are unknown, and, more importantly, we do not much care whether or not they are equal. An example where it seems appropriate is the comparison of the yields of two agricultural varieties; an example where it seems generally not appropriate is the comparison of the effect of a new drug with a control drug. In the first example one is interested in comparing the average yields over an area which is large relative to that of an experimental plot, but in the second example one is interested in comparing the effects of the drugs (including their variation) on individuals.

2. ELEMENTARY d -SOLUTION

Since $(d - \delta)/(\lambda_1 \sigma_1^2 + \lambda_2 \sigma_2^2)^{1/2}$ is $N(0, 1)$, a natural way to try to form a confidence interval for δ is by an interval of the form

* Henry Scheffé is professor, Department of Statistics, University of California, Berkeley. This research was partially supported by the Office of Naval Research under Contract N00014-67-A-0114-0004 (NR 043 030).

$$|\delta - d| \leq v(\lambda_1 s_1^2 + \lambda_2 s_2^2)^{1/2}, \quad (2.1)$$

where v is a constant, depending on n_1 , n_2 and the nominal confidence coefficient $1 - \alpha$. It seems to be not generally realized (at least not realized by me until recently) that the confidence coefficient for (2.1) is $1 - \alpha$ if we take $v = t_\alpha(f_0)$, where $t_\alpha(f)$ denotes the two-tailed α -point of the t -distribution with f d.f., and f_0 is the smaller of f_1 , f_2 . That the confidence coefficient is $1 - \alpha$ follows from results of Hsu [6] which will be described later. For any of the solutions considered in this article any lower bound of $1 - \beta(\theta)$ is called a *confidence level*; by the *confidence coefficient* I mean the g.l.b. (greatest lower bound). These definitions correspond to those of *significance level* and the *size* of tests in Section 7.

"Elementary d -solution" is any interval (2.1) where v is a constant, and "special d -solution" the case where $v = t_\alpha(f_0)$. That $\beta(\theta) < \alpha$ for the special d -solution was claimed by Elston [2] who suggested that for other reasons this solution was an improvement over the solution of Section 3 for which $\beta(\theta) \equiv \alpha$, his claim being based on a heuristic argument comparing the behavior of the two solutions. Hsu [6], in an analysis which is a model of mathematical rigor, derived a series expansion for the $\beta(\theta)$ of the elementary d -solution to prove that as θ increases from 0 to ∞ , $\beta(\theta)$ decreases to a minimum at $\theta = (n_1 f_1)/(n_2 f_2)$ and thereafter is increasing. (Hsu denoted $\beta(\theta)$ by $\beta_2(\theta)$ and the constant v by $u_0^{1/2}$; his series is finite when n_1 and n_2 are both odd.) The limiting values at $\theta = 0$ and ∞ are not given by his series but are $\Pr\{|t(f_i)| > v\}$ where $i = 2$ and 1, respectively, and $t(f)$ denotes a t -variable with f d.f. The larger of these two limiting values is $\Pr\{|t(f_0)| > v\}$. Hence the l.u.b. (least upper bound) of $\beta(\theta)$ is¹

$$\text{l.u.b. } \beta(\theta) = \Pr\{|t(f_0)| > v\}, \quad (2.2)$$

and therefore the confidence coefficient is $1 - \alpha$.

The g.l.b. of $\beta(\theta)$ is also of interest; it is the value of $\beta(\theta)$ at its minimum where $\theta = (n_1 f_1)/(n_2 f_2)$. But for this θ it is easily verified that the ratio $(d - \delta)/(\lambda_1 s_1^2 + \lambda_2 s_2^2)^{1/2}$ has exactly the t -distribution with $f_1 + f_2$ d.f., and hence

$$\text{g.l.b. } \beta(\theta) = \Pr\{|t(f_1 + f_2)| > v\}. \quad (2.3)$$

For any v the g.l.b. and the l.u.b. can thus be found from a table of the c.d.f. (cumulative distribution function or "probability integral") of the t -distribution.² For example, for the special d -solution with $\alpha = .05$, so that l.u.b. $\beta(\theta) = .05$, we find from the table that for $n_1 = 5$ and $n_2 = 15$, g.l.b. $\beta(\theta) = \Pr\{|t(18)| > t_{.05}(4)\} = .015$; while for $n_1 = n_2 = 10$, g.l.b. $\beta(\theta) = .036$. For $n_1 = 5$, $n_2 = 15$, Hsu calculated from his series that $u_0 = v^2 = 5.2$ makes $\beta(1) = .05$, and that for this v , l.u.b. $\beta(\theta) = .085$, g.l.b. $\beta(\theta) = .035$.

As the sample sizes approach equality the variation of $\beta(\theta)$ decreases, and as they become large it approaches zero; these statements may be made precise as follows. For the elementary d -solution let us hold the confidence coefficient $1 - \alpha$ fixed, so the variation of $\beta(\theta)$ is

$$V_n \beta(\theta) = \alpha - \text{g.l.b. } \beta(\theta). \quad (2.4)$$

¹ An easier derivation than Hsu's [6] of the bounds (2.2) and (2.3) has been given by Mickey and Brown [7].

² See, for example [8, Table 9].

I shall assume $f_1 \leq f_2$, so $f_0 = f_1$, the behavior of $\beta(\theta)$ being symmetrical for $f_1 \geq f_2$. Then

$$\text{g.l.b. } \beta(\theta) = \Pr\{|t(f_1 + f_2)| > t_\alpha(f_1)\}. \quad (2.5)$$

To justify the first statement I hold $f_1 + f_2 = f$ fixed and denote by f' the largest integer $\leq (f_1 + f_2)/2$. As f_1 increases from 1 to f' , $t_\alpha(f_1)$ decreases from $t_\alpha(1)$ to $t_\alpha(f')$, and hence we see from (2.5) that g.l.b. $\beta(\theta)$ increases from $\Pr\{|t(f)| > t_\alpha(1)\}$ to $\Pr\{|t(f)| > t_\alpha(f')\}$, and from (2.4) that $\forall \theta$ $\beta(\theta)$ decreases accordingly. To justify the second statement note from (2.5) that if both f_1 and $f_2 \rightarrow \infty$, then g.l.b. $\beta(\theta) \rightarrow \alpha$. However, if $f_2 \rightarrow \infty$ with f_1 fixed, then g.l.b. $\beta(\theta) \rightarrow \Pr\{|z| > t_\alpha(f_1)\} < \alpha$ where z is $N(0, 1)$.

Confidence intervals for the Behrens-Fisher problem lead to confidence sets for Fieller's problem of estimating the ratio of means ξ_1/ξ_2 in the case when σ_1^2/σ_2^2 is unknown. Elston [2] has shown that for this purpose the special d -solution has the advantage over the following solutions, to be considered later, of not requiring an iterative calculation: (a) The Welch-Aspin solution, (b) Welch's approximate t -solution, and (c) the Fisher-Behrens solution. However, since the lengths of the intervals for (a) and (b) are about the same, while that for (c) is longer and that of the special d -solution is longest, we might expect that in some sense the four corresponding confidence sets for ξ_1/ξ_2 might be of increasing size in the same order; but whether the differences of size are of any practical importance is another question.

3. AN IMPRACTICAL SOLUTION

All the interval estimates considered in this article are of the form

$$|\delta - d| \leq v s_d, \quad (3.1)$$

where s_d is an estimate of σ_d such that $E(s_d^2) = \sigma_d^2$, and v is a random variable (possibly a constant) that is a function of s_1^2/s_2^2 which does not vary much. For example, if $\alpha = .05$ and $f_0 \geq 3$, then $1.96 \leq v \leq 2.31$ for all the solutions considered.

In a study of a class of solutions of the Behrens-Fisher problem based on the t -distribution and for which $\beta(\theta) \equiv \alpha$, I [11, 12] produced a solution of the form (3.1) where $v = t_\alpha(f_0)$ and proved some optimum properties in the class of solutions. These articles were written before I had much consulting experience, and since then I have never recommended the solution in practice. The reason is that the estimate s_d requires putting in random order the elements of the larger sample, and the value of s_d and hence the length of the interval depend very much on the result of this randomization of the data. The effect of this in practice would be deplorable.¹ None of the other solutions considered in this article have this disadvantage. The interval of the special d -solution of Section 2 has the same root-mean-square length (if X is a random variable, $\text{r.m.s.}(X) = [E(X^2)]^{1/2}$), and is thus to be preferred. That its probability $1 - \beta(\theta)$ of covering δ exceeds $1 - \alpha$ (since θ cannot equal 0 or ∞) I regard as an additional

¹ If in choosing a randomized-blocks or Latin-square design at random one that looks too systematic is obtained, I believe a statistician would usually advise drawing again; he can justify this since the small proportion of designs thus excluded would have little effect on the probability calculations of interest to him, namely, those of the significance levels and power of his tests and the means and variances of his estimates. But what if the user of my confidence interval does not like the width he obtained and decides to randomize again?

advantage in interval estimation, although not necessarily in testing as indicated in Section 7. (One may question the practical importance of the virtue that, among the six solutions considered, these two are the only ones for which the confidence coefficient is known exactly.)

4. WELCH-ASPIN SOLUTION

This solution is of the form (3.1) with

$$s_d = (\lambda_1 s_1^2 + \lambda_2 s_2^2)^{1/2} \quad (4.1)$$

and v a random variable whose value is given by tables described below. The solution was derived by Welch [14] and further developed by Aspin [1] by purely formal calculations. Because of this, numerical verification of $\beta(\theta)$ is needed.⁴ I have found it reported only for the case $\alpha = .10$, $f_1 = f_2 = 6$, by Welch [16] who claimed that in this case $\beta(\theta)/2$ lies within .002 of .05; his method is explained in [16, p. 299].

The necessary tables for $\alpha = .10, .05, .02, .01$ were calculated by Aspin, Trickett, Welch, and James, and are reproduced in Pearson and Hartley [9, Table 11]. The tables are entered with f_1, f_2 , and

$$c = \lambda_1 s_1^2 / (\lambda_1 s_1^2 + \lambda_2 s_2^2).$$

The tables require f_0 , the smaller number of d.f., to be ≥ 6 when $\alpha = .10$, ≥ 8 when $\alpha = .05$, ≥ 10 when $\alpha = .02$ or $.01$. The lengths of the resulting confidence intervals are shorter than those of the special d -solution in the ratio $v/t_\alpha(f_0)$, and v is seen from the tables to fall between $t_\alpha(\infty)$ and $t_\alpha(f_0)$. For values of f_0 smaller than those allowed by the tables one may use the special d -solution, with confidence coefficient exactly $1 - \alpha$, or Welch's approximate t -solution of Section 5, or possibly as an approximation the Fisher-Behrens solution which is tabled for the f_i and α stated in Section 6.

5. WELCH'S APPROXIMATE t -SOLUTION

Welch [15] approximates the distribution of $(d - \delta)/s_d$, where s_d is given by (4.1), by a t -distribution with a suitable number f of d.f. as follows. The random variable s_d^2 is approximated as $k s_d^2 \chi^2(f)/f$, where k is a constant, by equating the first two moments, with the result that $k = 1$ and

$$f = (\lambda_1 \sigma_1^2 + \lambda_2 \sigma_2^2)^2 / [f_1^{-1}(\lambda_1 \sigma_1^2)^2 + f_2^{-1}(\lambda_2 \sigma_2^2)^2]. \quad (5.1)$$

The unknown f is then estimated as F , defined by replacing the σ_i^2 by s_i^2 in (5.1). F is a function of s_1^2/s_2^2 which may be written

$$F = [f_1^{-1}c^2 + f_2^{-1}(1 - c)^2]^{-1};$$

it varies between f_0 and $f_1 + f_2$ and of course is in general not an integer. Suppose we denote by $t_\alpha(F)$ the two-tailed α -point obtained by interpolating in the

⁴ It seems rather scandalous to me that for what are probably the two most used solutions of an important statistical problem, tables for which were published 20 and 30 years ago and are each included in one of the two most used compendia of statistical tables, one solution is based on a concept whose relevance most statisticians do not understand or do not accept, while the other can only be justified by numerical calculation of its operating characteristics which was done only once, 20 years ago, for one set of values of α, n_1, n_2 . Computer calculations of $\beta(\theta)$ for the Welch-Aspin solution are needed for other values of α, n_1, n_2 and especially for unequal n_1, n_2 .

t -tables for F d.f. Then Welch's approximate t -solution is of the form (3.1) with s_d given by (4.1) and v the random variable $t_\alpha(F)$. The confidence intervals will be shorter than those of the special d -solution in the ratio $t_\alpha(F)/t_\alpha(f_0)$ which varies between $t_\alpha(f_1+f_2)/t_\alpha(f_0)$ and 1. That they are sometimes shorter and sometimes longer than the Welch-Aspin intervals is shown by examination of the tables for the latter from which the difference is seen to be at most about two percent.

Welch [15] states that for the case $\alpha = .10$ and $f_1 = f_2 = 6$, $\beta(\theta)/2$ for his approximate t -solution varies between .048 and .051. Wang [13]⁶ calculated $\beta(\theta)$ for 15 values of θ for each of seven pairs (f_1, f_2) with $f_0 \geq 4$ and $\alpha = .10, .05, .01$. The greatest deviation of $\beta(\theta)$ from α was .004 (I have rounded to three decimals results from Wang's tables which she estimates to be correct within .0001), in the case $f_1 = 4, f_2 = 20, \alpha = .01$, where l.u.b. $\beta(\theta) = .014$. Otherwise $\beta(\theta)$ varied between $\alpha - .002$ and $\alpha + .003$. It appears that l.u.b. $\beta(\theta) - \alpha$ and g.l.b. $\beta(\theta) - \alpha$ do not vary much with α for $\alpha = .10, .05, .01$, so that in a relative sense the approximation is much better for larger α . I judge Wang's work will justify the conclusion that Welch's approximate t -solution, which requires only the ubiquitous t -tables, is a satisfactory practical solution of the Behrens-Fisher problem.

6. TWO OTHER SOLUTIONS

For completeness I mention briefly two other solutions. The Fisher-Behrens solution [3] is of the form (3.1) with s_d given by (4.1) and v given by tables calculated by Sukhatme, Fisher, and Healy and reproduced in Fisher and Yates [5] where Table VI is for $f_0 \geq 6$ and $\alpha = .05$ and .01, Table VI₁ is for all odd $f_1, f_2 \leq 7$ and $\alpha = .10, .05, .02$, and .01. The tables are entered with f_1, f_2 , and δ , where

$$\delta = \arctan[(\lambda_1 s_1^2)/(\lambda_2 s_2^2)]^{1/2} = \arcsin c^{1/2},$$

and the θ, n_i , and s_i of the tables and text denote our δ, f_i and $\lambda_i^{1/2} s_i$, respectively. I have symbolized by α the "significance level," the quotation marks indicating that the term is not used in the sense of Section 7, that is, as an upper bound for $\beta(\theta)$, but rather in a "fiducial probability" sense. Welch [14] calculated that the Fisher-Behrens interval should always include the Welch-Aspin interval, and the tables bear this out. The Fisher-Behrens interval is about eight percent (ten percent) longer when $f_1 = f_2 = 10, c = .5$, and $\alpha = .05$ (.01). According to Welch [14] the percent increase should be approximately proportional to $c(1-c)$ and to $f_1^{-1} + f_2^{-1}$; the v of the two solutions coincide in the limiting cases $c = 0$ or $c = 1$ or $f_1 = f_2 = \infty$. The above inclusion relation of the two intervals implies that $1 - \beta(\theta)$ for the Fisher-Behrens interval is $>$ than that of the Welch-Aspin interval, and the confidence coefficient is hence $\geq$.

The Student solution is based on the assumption that $\sigma_1^2 = \sigma_2^2$ and is thus not intended as a solution of the Behrens-Fisher problem, although it is sometimes used that way in practice. If $\theta = 1$, an unbiased estimate of σ_d^2 may be obtained from the pooled sample variances as

$$s_d^2 = (\lambda_1 + \lambda_2)(f_1 s_1^2 + f_2 s_2^2)/(f_1 + f_2);$$

⁶ I am indebted to Professor David L. Wallace for informing me of an early smaller version of these tables of $\beta(\theta)$ calculated by Dr. Wang and to Dr. Wang for letting me use them.

s_d is used for s_d in (3.1), while v is taken as $t_{\alpha}(f_1 + f_2)$. Note that $E(s_d^2) = \sigma_d^2$ if and only if $\theta = 1$ or $n_1 = n_2$. If $n_1 = n_2$, the Student solution coincides with the elementary d -solution (2.1) with $v = t_{\alpha}(f_1 + f_2)$, and so the results of Hsu about $\beta(\theta)$ described in Section 2 apply. From these it follows that as θ increases from 0 to ∞ , $\beta(\theta)$ decreases from the limiting value

$$\Pr\{|t(f_1)| > t_{\alpha}(2f_1)\} \quad (6.1)$$

to a minimum of α at $\theta = 1$ and thereafter increases to the limiting value (6.1) so that the confidence coefficient is 1 minus (6.1). For example if $\alpha = .05$ and $n_1 = n_2 = 5$, then the confidence coefficient is $1 - \Pr\{|t(4)| > 2.31\} = 1 - .082$; if $\alpha = .05$ and $n_1 = n_2 = 10$, the confidence coefficient is $1 - .065$. In general (when n_1 and n_2 are not necessarily equal) Hsu [6] calculated $\beta(\theta)$ for the Student solution as an infinite series (again finite if n_1 and n_2 are both odd; he denotes $\beta(\theta)$ by $\beta_1(\theta)$). From this he showed among other things that for $n_1 \neq n_2$, $\beta'(1) \neq 0$; hence it follows that g.l.b. $\beta(\theta) < \beta(1) = \alpha < \text{l.u.b. } \beta(\theta) = 1 - \text{confidence coefficient}$. For $\alpha = .05$, $n_1 = 5$, $n_2 = 15$, he calculated that $\beta(\theta)$ varies between .0024 and .317.

Unlike the other five solutions considered here, this solution of the Behrens-Fisher problem is asymptotically incorrect for large unequal sample sizes, elementary calculations showing that $\beta(\theta)$ may take on any value between 0 and 1 for any α ($0 < \alpha < 1$) and for suitable θ and suitably large n_1 , n_2 . The practical conclusion is that this solution should never be used unless n_1 and n_2 are equal, or nearly so. Perhaps even when n_1 and n_2 are equal but small, it should be used with an eye to the behavior of $\beta(\theta)$ discussed in connection with (6.1).

7. ASSOCIATED TESTS

An interval (3.1) may be associated with a test of the hypothesis $\delta = \delta_0$, the test accepting the hypothesis if and only if the interval covers δ_0 . Since $1 - \beta(\theta)$ is the probability that the interval covers δ_0 , the probability of Type I error is $\beta(\theta)$. If *significance level* of the test is defined as any upper bound of $\beta(\theta)$, and *size* as the l.u.b., then *significance level* = *1-confidence level*, and *size* = *1-confidence coefficient*. It might seem that from the behavior of $\beta(\theta)$ we cannot draw any conclusions about the probability of Type II error, but we can in the light of certain continuity considerations. Without loss of generality I will henceforth take $\delta_0 = 0$, so that the hypothesis tested is $H_0: \delta = 0$, and the power is $\Pr\{|d| > v s_d\}$. If v in (3.1) is assumed to be a continuous function of s_1^2/s_2^2 , then the power is a continuous function of the two parameters θ and $\delta' = \delta/\sigma_d$, where σ_d is given by (1.1). We may also think of H_0 as specifying $\delta' = 0$ without specifying θ . If for two solutions corresponding to subscripts 1 and 2, $\beta_1(\theta) < \beta_2(\theta)$ for some θ , then against alternatives consisting of this θ and δ' sufficiently close to 0, the second test will have greater power than the first. This may not be of any practical importance, since for δ' close to 0 neither test will have much power. Nevertheless, this seems to be the justification of the search for similar tests corresponding to $\beta(\theta) \equiv \text{constant}$.

Another easy conclusion that can be drawn about the power of tests associated with confidence intervals of the form (3.1) is that for any two solutions with the same s_d , and v denoted by v_1 and v_2 , if $v_1 < v_2$ (with probability one) then the power of the first test exceeds that of the second against all alterna-

tives. Thus, the Welch-Aspin test is more powerful than the Fisher-Behrens test which is more powerful than the special d -test.

I now consider the numerical evaluation of the power of some of the tests introduced. The power of the elementary d -test was calculated by Hsu [6] in the form of a doubly infinite series (finite for n_1 and n_2 both odd); from this series tables could presumably be calculated. However, for the power of this and the other tests we may be satisfied with approximations from existing tables, since the desired value is the probability of avoiding an error of the less serious kind, that is, of Type II, and we generally do not want this as accurately as $\beta(\theta)$, the probability of the more serious kind of error. The elementary d -test rejects H_0 if and only if $|d/s_d| > v$, where v is a constant, and s_d throughout the rest of this section denotes (4.1). The power can be approximated as

$$\Pr\{|t'(f, \delta/\sigma_d)| > v\}, \quad (7.1)$$

where $t'(f, \delta/\sigma_d)$ has the noncentral t -distribution with f d.f. given by (5.1) and noncentrality parameter δ/σ_d , by approximating s_d^2 by $\sigma_d^2 \chi^2(f)/f$ as in Section 5. For the special d -test $v = t_\alpha(f_0)$, and so the approximation (7.1) cannot be obtained by interpolating for f in tables of the power of the t -test where v would be $t_\alpha(f)$ instead of $t_\alpha(f_0)$; it requires interpolation for f in tables of the c.d.f. of the noncentral t -distribution like those of Resnikoff and Lieberman [10], f being the function of θ defined by (5.1).

For the Student test, where H_0 is rejected if $|d| > s_d t_\alpha(f_1 + f_2)$, the power was also calculated by Hsu as a doubly infinite series. It could be approximated by the same method used above, but since this is in general a bad test unless $n_1 = n_2$, I shall limit myself to this case. But this is a special case of the elementary d -test, so that the power can be approximated by (7.1) with $v = t_\alpha(2f_1)$, which can be found from the tables of Resnikoff and Lieberman.

I shall not discuss the power of my solution in Section 3 because I do not recommend it, nor that of the Fisher-Behrens test.* I shall treat together the power of the Welch-Aspin test and Welch's approximate t -test, both of which are of the form $\Pr\{|d|/s_d > v\}$, where v for the first test is obtained from the tables and v for the second test is the random variable $t_\alpha(F)$ defined in Section 5. The power for the first test should be approximated well by that of the second, since the entries in the tables are closely approximated by $t_\alpha(F)$, as remarked toward the end of Section 5. Using now the above approximation of $|d|/s_d$ as a noncentral t -variable, we would like to evaluate

$$\Pr\{|t'(f, \delta/\sigma_d)| > t_\alpha(F)\}. \quad (7.2)$$

However, this is not given by the c.d.f. of noncentral t because $t_\alpha(F)$ is not a constant. I propose to approximate (7.2) by replacing the random variable F in (7.2) by the (specified) constant f . This is the same kind of approximation that was made in the derivation of Welch's approximate t -solution in Section 5 where $\beta(\theta)$ was first approximated as

$$\Pr\{|t'(f, 0)| > t_\alpha(f)\}, \quad (7.3)$$

*I can imagine the old master fulminating [4] about anyone calculating the probability of Type II errors for his test.

and this was then approximated as α by replacing the random variable F in (7.3) by the (unknown) constant f . The resulting approximation for the power, namely $\Pr\{|\ell'(f, \delta/\sigma_d)| > t_{\alpha}(f)\}$, may be evaluated by interpolating for f in tables of the power of the t -test, such as Owen's [8], or in the chart in Pearson and Hartley [9, Table 10] or in the tables of the c.d.f. in Resnikoff and Lieberman [10]; here f is again the function of θ given by (5.1).

In conclusion I mention that instead of the intervals (3.1), one-sided intervals $\delta < d + vs_d$, or else $\delta > d - vs_d$, may be employed, or the associated one-tailed tests, rejecting $H_0: \delta = 0$ when $d < -vs_d$ or $d > vs_d$. The $\beta(\theta)$ are then halved for the same α because the distribution of the quotient $q = (d - \delta)/(vs_d)$ is symmetrical about $q = 0$. The above expressions for approximating the power of the tests then require obvious modifications.

REFERENCES

- [1] Aspin, Alice A., "An Examination and Further Development of a Formula Arising in the Problem of Comparing Two Mean Values," *Biometrika*, 35 (May 1948), 88-96.
- [2] Elston, R. C., "An Analogue to Fieller's Theorem Using Scheffé's Solution to the Fisher-Behrens Problem," *American Statistician*, 23, No. 1 (February 1969), 20-8.
- [3] Fisher, R. A., "The Fiducial Argument in Statistical Inference," *Annals of Eugenics*, 6 (December 1935), 391-8.
- [4] ———, "Statistical Methods and Scientific Induction," *Journal of the Royal Statistical Society, Series B*, 17 (November 1955), 69-78.
- [5] ——— and Yates, F., "Statistical Tables for Biological, Agricultural and Medical Research," Edinburgh: Oliver and Boyd, 1957.
- [6] Hsu, P. L., *Statistical Research Memoirs*, Department of Statistics, University College, London, 1938.
- [7] Mickey, M. R. and Brown, M. B., "Bounds on the Distribution Functions of the Behrens-Fisher Statistic," *The Annals of Mathematical Statistics*, 37 (June 1966), 639-42.
- [8] Owen, D. B., "The Power of Student's t -Test," *Journal of the American Statistical Association*, 60 (March 1965), 320-33.
- [9] Pearson, E. S. and Hartley, H. O., *Biometrika Tables for Statisticians*, Cambridge, Eng.: Cambridge University Press, 1966.
- [10] Resnikoff, G. J. and Lieberman, G. J., *Tables of the Non-Central t -Distribution*, Stanford, Calif.: Stanford University Press, 1957.
- [11] Scheffé, H., "On Solutions of the Behrens-Fisher Problem, Based on the t -Distribution," *The Annals of Mathematical Statistics*, 14 (March 1943), 35-44.
- [12] ———, "A note on the Behrens-Fisher Problem," *The Annals of Mathematical Statistics*, 15 (December 1944), 430-2.
- [13] Wang, Ying Y., "The Probabilities of the Type I Errors of the Welch Tests for the Behrens-Fisher Problem," to appear in *Journal of the American Statistical Association*.
- [14] Welch, B. L., "The Generalization of 'Students' Problem When Several Different Population Variances Are Involved," *Biometrika*, 34 (January 1947), 28-35.
- [15] ———, "Further Note on Mrs. Aspin's Tables and on Certain Approximations to the Tabled Function," *Biometrika*, 36 (December 1949), 293-6.
- [16] ———, "Note on Some Criticisms Made by Sir Ronald Fisher," *Journal of the Royal Statistical Society, Series B*, 18 (January 1956), 297-302.